Chiral intermolecular interactions in solution, in cold rare gas matrices and at interfaces: chiroptical and interfacial force studies

by

Yanqing Yang

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

Doctor of Philosophy

Department of Chemistry

University of Alberta

Abstract

My PhD thesis centers on characterization of non-covalent interactions involving chiral molecules in solution, in cold rare gas matrices and at organic|water interfaces using a combined experimental and theoretical approach. In order to extract structural information and conformation distributions of different chiral molecules under different environments, several vibrational and vibrational optical activity spectroscopic tools, i.e., Infrared (IR), Raman, vibrational circular dichroism (VCD) and Raman optical activity (ROA), are utilized. To explain the experimental spectroscopic observations, first, systematic conformational searches of these molecular systems were performed using a recently developed conformational searching tool with the inclusion of polarizable continuum model of the appropriate environment, such as argon, and water and methanol solvents. Then the geometry optimizations and spectral simulations of the conformers identified were carried out at the B3LYP-D3BJ/def2-TZVP levels.

In Chapter 3, the conformational distribution and self-aggregation behavior of tetrahydro-2-furoic acid (THFA) were studied by using matrix isolation (MI)-VCD. The well-resolved experimental MI-IR and MI-VCD features in an argon matrix at 10 K, 24 K and 30 K allow one to identify the dominant monomeric and binary conformations. Interestingly, the main binary THFA structures observed consist of two *trans*-COOH THFA subunits where the hydroxyl and carbonyl groups of COOH are at opposite sides, in contrast to the usual double hydrogen-bonded ring binary structure which contains two

cis-COOH subunits. The work showcases the power of MI-VCD spectroscopy in revealing unusual structures formed in a cold rare gas matrix.

Another focus of my thesis research is on solvent effects in chiroptical measurements. In Chapter 4, I studied THFA in water under different pH conditions using VCD spectroscopy. The "clusters-in-a-liquid" solvation model which takes into account of both explicit and implicit solute-solvent interactions was applied to interpret the experimental IR and VCD spectra of THFA under acidic condition satisfactorily. The possible causes for the poorer performance of the above solvation model for THFA under basic condition were discussed. In Chapter 5, IR, Raman, VCD and ROA spectra of N-acetyl-L-cysteine in water and in methanol were measured. The performance of the "clusters-in-a-liquid" solvation approach was compared with that of the ab initio molecular dynamics (AIMD) approach. While both provide satisfactory agreements with the experimental data, the AIMD simulations perform better at the lower wavenumber region, likely because it includes some anharmonic effects whereas the static DFT approach is based on harmonic frequency calculations.

In Chapter 6, I explored a new research direction, i.e., chirality discrimination at a binary toluene (organic)/water(aqueous) interface. By using a combination of interfacial tension measurements and molecular dynamic simulations, Chirality discrimination at a binary toluene (organic)/water(aqueous) interface between R- or S-Tol-BINAP (2,2'-Bis(di-p-tolylphosphino)-1,1'-binaphthyl) molecules and the water-soluble serine chiral

species is examined for the first time. The molecular dynamics simulations show that differences in the interfacial tension values arise when a homochiral versus a heterochiral enantiomeric pairs are introduced at the interface, in agreement with the experimental results. Such observations prove that preferential interactions exist between different pairs of enantiomers at the binary interfaces and can be captured directly by using interfacial tension measurements.

Preface

This thesis is based on the research work done during my PhD study in the Department of Chemistry at the University of Alberta between September 2017 to July 2022 under the supervision of Prof. Yunjie Xu. The nature and extent of my contributions to the research work included in the thesis are briefly described below.

Chapter 3 is based on the published paper "Matrix Isolation-Vibrational Circular Dichroism Spectroscopic Study of Conformations and Non-Covalent Interactions of Tetrahydro-2-Furoic Acid", Y. Yang, J. Cheramy, Y. Xu, *ChemPhysChem* **2021**, *22*, 1336-1343. I did all the experimental measurements, theoretical calculations and data analyses, and wrote the first draft of the manuscript. J. Cheramy helped with reviewing and editing the manuscript. Prof. Y. Xu was the supervisory author and was involved in the conceptualization, funding acquisition, supervision, and writing the draft and its review and editing.

Chapter 4 is planned to be submitted shortly. I carried out all the experimental measurements, theoretical calculations, and data analyses, and wrote the initial manuscript draft. Prof. Y. Xu was the supervisory author and was involved in the conceptualization, funding acquisition, supervision, and writing the draft and its review and editing.

Chapter 5 is based on the published paper "Raman Optical Activity of N-Acetyl-L-Cysteine in Water and in Methanol: The "Clusters-in-a-Liquid" Model and *ab initio* Molecular Dynamics Simulations", Y. Yang, J. Cheramy, M. Brehm, Y. Xu, *ChemPhysChem* **2022**, e202200161. I carried out the static harmonic DFT calculations and data analyses, and wrote the main part of the initial manuscript draft. J. Cheramy did the experimental measurements and wrote a portion of the initial manuscript draft. Dr. M. Brehm was the corresponding author and was responsible for the AIMD simulations and the wrote the part about the AIMD results. Prof. Y. Xu was the corresponding author and was involved in the conceptualization, funding acquisition, supervision, and writing the draft and its review and editing.

Chapter 6 is based on the submitted manuscript titled “Chirality Discrimination at Binary Organic/Water Interfaces by Interfacial Tension Measurements and MD Simulations”, Y. Yang, X. Sun, M. R. Poopari, C. Jian, H. Zeng, T. Tang, Y. Xu, *ChemPhysChem.* **2022**. I completed all the experimental measurements and experimental data analyses, and wrote the first manuscript draft. Dr. X. Sun is an equal contributor and was responsible for the MD simulations and the related data analysis, and wrote a portion of the initial manuscript. Prof. T. Tang was the corresponding author, responsible for the MD simulations. Dr. M. R. Poopari helped with the initial experiment and also the reviewing of the manuscript. Prof. C. Jian helped with the reviewing of the manuscript. and its review and editing. Prof Y. Xu was the corresponding author and was involved in the conceptualization and supervision of the work, and in writing and editing the draft. Prof H. Zeng is the corresponding author and was involved in the conceptualization and supervision and reviewing of the manuscript.

Acknowledgements

I would like to first express my sincerely appreciation to my supervisor Prof. Yunjie Xu for her great support, precious suggestions, and guidance over the past five years, which I did benefit a lot from and help me grow up as an independent researcher. I would like to thank Prof. Wolfgang Jäger for all his valuable support, advice, and helpful discussions. I would like to also thank all my internal supervisory committee members, Prof. Alex. Brown and Prof. Michael J. Serpe, for your wonderful suggestions and comments during my five years PhD study, my candidacy exam, and my defense. I would like to thank my lab coordinator, Dr. Yoram Apelblat (CHEM 10X general chemistry lab) for his useful help and suggestion during my graduate studies. I would like to have a special thanks to Dr. Mohammad Reza Poopari (CHEM 37x physical chemistry lab coordinator and my lovely friend) for your wonderful suggestions and help during my most difficult time of my PhD studies. I would like to also give a special thanks to all my brilliant colleagues and friends in both Xu and Jäger group, Nate, Matthias, Leo, Jiao, Angelo, Amin, Joseph, Fan, Arsh, Moe, Bowei, Boyang, Guojie, Haolu, Qian, Jiarui, Mutasem, and Colton. I am also grateful for all the supports and valuable advice for the staff in electronic and mechanical shop on helping me to develop the new MI system.

I would like to thank all the incredible support and endless love from my family members, especially my grandparents, my father, and my mother. They are always encouraging and understanding me no matter what. Lastly, I would like to express my deep love and grateful to my girlfriend, Lingyu, for the understanding, supporting, caring, and staying with me as always. Without you, I will not have the passion and energy to finish my PhD journey.

Table of contents

List of Tables (Chapters)

List of Tables (Appendices)

Table Page

List of Figures (Chapters)

List of Figures (Appendices)

List of Symbols

+	Dextrorotatory
-	Levorotatory
λ	Wavelength
d	Optical path difference
$\Delta\rho$	Mass density difference
g	Gravity constant
R_0	Radius of curvature at the drop apex
β	Shape factor
D_E	Maximum diameter of the droplet
D_S	Diameter at distance D_E from the drop apex
k	Force constant
ν	Fundamental vibrational frequency
μ	Reduced mass
E_n	Vibrational energy of a given energy level n
$\hbar$	Reduced Planck's constant (also known as Dirac's constant), i.e., $h/2\pi$
ω	Angular frequency
D_{01}^{i}	Dipole strength of a given fundamental vibrational transition at the i^{th} normal mode
$\vec{\mu}$	Electric dipole moment operator
ψ_0	Wavefunction for the ground vibrational state
ψ_1^i	Wavefunction or the first excited vibrational state with respect to the i^{th} normal mode
R_{01}^{i}	Rotational strength of a given fundamental vibrational transition

	at the i^{th} normal mode
$\vec{m}$	Magnetic dipole moment operator
$cos\theta(i)$	The angle between the electric dipole and magnetic dipole transition moment vectors
ε	Molar absorptivity coefficient
$\Delta\varepsilon$	Differential molar absorptivity coefficient
E	Electric field
a	Electric dipole-electric dipole polarizability tensor
G'	Electric dipole-magnetic dipole optical activity tensor
A	Electric dipole-electric quadrupole optical activity tensor
ω_{ji}	The differential angular frequencies between the initial (i) and virtual state (j) of the molecule
Θ	Electric quadrupole
p_{xx}, p_{yy}, and p_{zz}	The diagonal components of the pressure tensor in the x, y and z directions

List of Abbreviations

BINAP	2,2′-Bis(diphenylphosphino)-1,1′-binaphthalene
CIP	Cahn-Ingold-Prelog
ORD	Optical rotatory dispersion
NMR	Nuclear magnetic resonance
ECD	Electronic circular dichroism
VCD	Vibrational circular dichroism
ROA	Raman optical activity
DFT	Density functional theory
ADF	Amsterdam DF
VOA	Vibrational optical activity
CP-FTMW	Chirped pulse Fourier transform microwave
IEF	Integral equation formalism
PCM	Polarizable continuum model
MI	Matrix isolation
THFA	Tetrahydro-2-furoic acid
LA	Lactic acid
NALC	N-acetyl-L-cysteine
AIMD	*ab initio* molecular dynamics
Tol-BINAP	2,2'-Bis(di-p-tolylphosphino)-1,1'-binaphthyl
CCD	Charged-coupled device
FT	Fourier transform
IR	Infrared radiation
FIR	Far-infrared

MIR	Mid-infrared
NIR	Near-infrared
MCT	HgCdTe
PEM	Photoelastic modulator
AC	Alternating current
DC	Direct current
LIA	Lock-in-amplifier
ST	Surface tension
IFT	Interfacial tension
BO	Born-Oppenheimer
MFP	Magnetic field perturbation
AAT	Atomic axial tensor
APT	Atomic polar tensor
CREST	Conformer-rotamer ensemble sampling tool
xTB	A semiempirical tight-binding quantum chemistry method
RMSD	Root-mean-square deviations
D3BJ	Grimme's dispersion correction with Becke Johnson damping
B3LYP	Becke, 3-parameter, Lee–Yang–Parr, Hybrid functional
Ar	Argon gas
sccm	Standard cubic centimeters per minute
MRR	Molecular rotational resonance
MTD	Meta-dynamics
AO	Atomic orbital
VA	Vibrational adsorption
NVPT	Nuclear velocity perturbation theory
QTAIM	A quantum theory of atoms in molecules
MD	Molecular dynamics

Chapter 1

Introduction

1.1 Importance of chirality and its detection

A molecule is chiral if its mirror image cannot be superimposed onto itself. The word *chirality* is derived from the Greek word, kheir, which means hand. A common example of chiral objects that one experiences in our daily life is our right and left hand. While they are mirrored images to each other, they are not superimposable if one puts the left hand on top of the right one. A chiral molecule and its mirror image are called enantiomers. They have the same chemical and physical properties in an achiral environment, but they differ in the spatial orientations of constituent atoms, a fact which might lead to completely different behavior in the chiral environment. For example, R-carvone has a mint-like smell, whereas S-carvone has a caraway-like smell. Such differences are typically achieved by different non-covalent interactions between them and the nasal receptors.[1] Apart from that, chirality also plays a significant role in the field of chemistry and biochemistry. Many important molecules in life and in pharmaceutical are chiral. For example, chemical formulas of two representative chiral drugs, i.e., thalidomide and ofloxacin, are given in Figure 1.1.1, where the opposite enantiomers tend to bind with target receptors differently, resulting in a different pharmacodynamic property. Thalidomide was widely used in late 1950s to early 1960s as an over-the-counter drug for treating nausea symptoms in pregnant women. It was later found while the R-enantiomers has the desire efficacy, the S-enantiomers led to some children born with phocomelia, a rare birth defect where an infant's arms and/or legs are severely shortened or sometimes completely absent.[2] Further research discovered that surprisingly, enantiopure thalidomide molecules racemize quickly *in vivo*.[3] The thalidomide tragedy led many countries around the world to issue stringent guidelines on the related testing of chiral drugs before they are introduced into the market.[4] Another example is ofloxacin. It is a fluoroquinolone antibiotic that is used to treat a number of bacterial infections, for instance, urinary tract infections. It was later found out that its S-enantiomer, levofloxacin, was twice as potent as ofloxacin *in vivo*. Moreover, the S-enantiomer also reduced the level of side effects that people might encounter within their

lower respiratory tract.[5] Therefore, it is crucial to investigate how the preferred conformation of a chiral molecule is changed in a chiral environment, for example, when interacting with a pair of enantiomers.

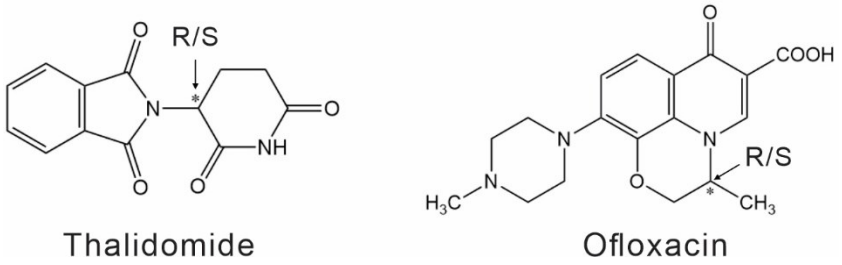


Figure 1.1.1. Chemical formula of the two representative chiral drugs: thalidomide and ofloxacin.

With the stringent regulations in terms of chiral drug development, it is not surprising that there has been much effort to develop spectroscopic tools for fast and reliable identification of chirality at every step of drug development. There are several spectroscopic methods that can be (and have been) used to achieve the goal, for example, X-ray crystallography,[6] nuclear magnetic resonance (NMR),[7] and chiroptical spectroscopies including optical rotatory dispersion (ORD), electronic circular dichroism (ECD),[8] vibrational circular dichroism (VCD),[9] and Raman optical activity (ROA),[10] and others, and more recently chiral tag rotational spectroscopy.[11]

All these techniques have their own advantages and disadvantages. For example, for NMR spectroscopy, one typically needs to synthesize a complex containing the chiral target and a chiral shift agent.[7] For X-ray measurements, a high-quality single crystal is needed. This may be difficult to achieve for some chiral molecules. In terms of ECD spectroscopy, the molecule of interest needs to contain a chromophore which can absorb the light in the UV-vis region.[12] VCD measures the difference in absorbance between the left and right circularly polarized light for a given fundamental vibrational transition of a chiral molecule, usually in the mid-infrared spectral region. The intensity ratios between VCD and IR tend to be much weaker, which is about 10^{-4} - 10^{-6}, compared to that of ECD (10^{-2} - 10^{-3}). Therefore, one spends more time on the measurement in order to achieve a good signal to

noise ratio. Similarly, ROA measures the difference in inelastic scattering of right and left circularly polarized light, although the intensity ratio of ROA versus Raman is also small, i.e., 10^{-3}. Since Raman is already a weak process,[13] one also needs to accumulate for some time to achieve good signal-to-noise ratio for an ROA spectrum.

The two techniques I used extensively in my PhD research are VCD and ROA. VCD spectroscopy has experienced some significant growth in both experimental and theoretical aspects in the past few decades.[14] Nowadays, there are many analytical instrument companies who manufacture VCD spectrometers, for instance, BioTools, Bruker, JASCO, and ThermoFisher. ROA, like its parent Raman spectroscopy, has several advantages in terms of applications in biochemistry, such as affording measurements of biomolecules directly in aqueous solution.[15] Moreover, an ROA spectrum can be measured down to around 200 cm^{-1} routinely. Such low cm^{-1} would be difficult to achieve with a FT-VCD instrument. Therefore, the ROA spectral analysis can often provide complementary information with VCD measurements. ROA can be measured in several different polarization forms. One common form, termed SCP-ROA, uses unpolarized incident radiation and then measures the intensity difference in the right minus left circularly polarized Raman scattered radiation by chiral molecules.[9a, 9b] This is the form utilized in the BioTools ROA spectrometers, the only commercialized ROA instrument currently available.

For these chiroptical measurements which are commonly done in solution, one needs to simulate the corresponding spectra and compare to the experimental ones in order to extract the information about the absolute configurations and other structural properties of a chiral molecule of interest. To do this, one turns to the help of theoretical calculations. Density functional theory (DFT) simulations of infrared and VCD, and Raman and ROA are implemented in some computational software packages, such as Gaussian[16] and Amsterdam DF (ADF).[17] More detailed information on the instrumentation of VCD and ROA spectroscopy and on DFT simulations will be discussed in Chapter 2. In addition, the

theoretical calculations for ECD and ORD are usually done nowadays using time-dependent density functional theory (TDDFT) and the calculation key words are available in many electronic structure software packages such as Gaussian.[18] In terms of comparison between experiment and theory, an ECD spectrum tends to have fairly broad spectral features, making detailed assignment challenging. In ORD, one basically compares the calculated and experimental values. If there are multiple conformations with different ORD values, it would be difficult to draw a reliable conclusion based on just one number. VCD and ROA have multiple vibrational bands which are relatively narrow, offering the opportunity to extract the rich structural information, for example structural information of individual conformers. This also makes it more challenging to properly interpret the experimental VCD and ROA spectra.

Recently, there has been an exciting development associated with measurements of ROA under (near) resonance, leading to the discovery of a new type of chiral Raman spectroscopy which we term ECD-circularly polarized (CP) Raman, i.e., eCP-Raman.[19] I was involved in some of the research activities as a minor author, although this result is not included in the current thesis. These publications point out the importance to recognize the contributions associated with other chiroptical events when the system of interest in under (near) resonance. For the ROA studies presented in this thesis, the chiral molecules are far off resonance, and one therefore does not need to be concerned about the potentially additional chiroptical events.

1.2 Non-covalent intermolecular interactions involving chiral solute and solvent (water and methanol) molecules

From the above discussion, it is clear that properly modeling chiral molecules in solution is a crucial step in extracting structural information from the VCD and ROA measurements. Below I give an overview of chiral solute···water interactions since this is one focus of my thesis.

Water is essential to our life. Water is a common medium for many biological processes in nature. It was recognized in recent years that water directly participates in a vast number of biological events, for example enzyme functions, molecular recognition, and protein assembly,[20] facilitated by non-covalent interactions such as hydrogen bonding interactions. At the molecular level, the chiral molecule···water interactions have been extensively studied in the gas phase using several different conformer specific spectroscopic tools. For example, Simons and co-workers applied IR-UV and UV-UV hole burning double resonance technique[21] to probe conformation preference of monosaccharides in the gas phase. They found that with the addition of just one single water molecule, such preference could be altered drastically. Using cavity and chirped pulse-Fourier transform microwave (CP-FTMW) spectrometers, Sanz, Alonso, Xu and others have investigated the conformational landscapes of complexes of some small chiral organic molecules with water, such as alanine[22], methyl glycidate[23], methyl lactate,[24] tetrahydro-2-fuoric acid,[25] and β-propiolactone, [26] as well as water complexes with a series of transiently chiral fluoroalcohol molecules.[27] One interesting outcome of these studies is the (barrierless) large amplitude motions in these water containing complexes. The significance of these motions in modeling chiral solute···water interactions in an aqueous solution will be further elaborated below.

The advantage for the vibrational optical activity (VOA) spectroscopies, which include VCD and ROA, is that they are not only very sensitive to the absolute configuration, but also to small structural changes. Such small structural changes may be brought on by non-covalent interactions, for example, between a chiral solute molecules and surrounding water solvent molecules. To properly interpret the measured experimental VOA features, it was already recognized in 2010s that explicit solvation effects needed to be included in the theoretical VOA simulation in addition to the implicit effects.[9c,28] Without a proper account of the solvation, a less than satisfactory agreement with the experiment would make the spectral interpretation more difficult.[29]

In 1930s, Onsager and others built the very early form of implicit solvation model in which they treated the electric dipole moment of solvent molecules in a spherical cavity as a function of dielectric constant.[30] Nowadays, more universal implicit solvation models are used to model the bulk solvent environment. For example, the solvents are treated as a conductor-like continuum in the COSMO model.[31] Other forms of the polarizable continuum model (IEF-PCM) treat the solvent as a dielectric environment,[32] where an integral equation formalism (IEF) version of the PCM model is used.[33] It is, however, more challenging to interpret VOA spectra for some chiral molecules of interest in water with only implicit solvation model, for example serine[34] and N-acetyl-L-cysteine[35] in water.

Since a water molecule can act as both hydrogen bond donor and acceptor, it will tend to form extensive hydrogen bonding networks among themselves and also with the solute. How to model such explicit hydrogen bonding interaction is a challenge. Our research group had carried out a series of VCD studies for some small organic molecules in water.[36] Surprisingly, some distinctive VCD features were captured in the water bending regions in all of these studies.[36] How could an achiral solvent have their own VCD band features? This is because water forms explicit H-bond(s) with the chiral solute molecule and some of such associations last a fairly long time so that one can treat the chiral solute···$water_n$ complexes as new species in VOA simulations. The observation of such induced solvent VCD features indicates there is a strong preference for certain binding topologies with water and also such associations exist for a substantial amount of time. Otherwise, it would not be possible to observe such chirality transferred VCD bands. From the significant findings described in the above studies, our group developed a new model to account for the solvent effect, which is called 'clusters-in-a-liquid' model.[37] The model proposed that the main experimental features in aqueous solution observed belong to the long-lived chiral solute···water complexes, rather than the chiral solute by itself. Since water is part of the long-lived chiral complexes, it is not surprising to observe VCD features at the water bending region.[36] To account for such solvent effects, one needs to build the chiral

solute···water$_n$ complexes and place such complexes in the implicit solvent model to mimic the bulk phase environment. How large should these chiral solute···water$_n$ complexes be? There are different opinions (see some general reviews below). Our research group advocates for small clusters with only one or two water molecules at each binding sites since small clusters tend to last longer than those with many water molecules simultaneously. The "clusters-in-a-liquid" model was also applied to other small protic solvent molecules such as methanol, where VCD features of methyl mandelate[38] and methyl lactate[39] were simulated and compared to the experiment.

Zehnacker and co-workers have recently reviewed some of the solvation approaches used to account for the solvent effects.[40] One critical question in the current field is how to generate the chiral hydration clusters that can be used to model the chiral solute···water interactions in aqueous solution. Currently, one common approach is to use molecular dynamics (MD) simulations to generate a large number of possible clusters. Then the snapshots of some smaller clusters, i.e., a solvent layer with a certain radius from the chiral solute center were considered. Often, tens to some thousands of such snapshots were extracted as initial candidates for the subsequent DFT geometry optimization and frequency calculations with the inclusion of PCM of the solvent.[41]

An exciting new development is the application of *ab initio* molecular dynamics (AIMD) simulations to the VOA solvation treatments.[42] The AIMD advantages are directly applicable to condensed phase systems, covering anharmonic effects to a certain extent, and reproducing line shapes, overtones, and combination bands in qualitatively correct manner.[42] One main focus of my thesis is to explore this research direction further using the 'clusters-in-a-liquid' model and also the newly developed AIMD approach for VOA in collaboration with one of the main developers in this area.

1.3 Probing chiral molecular aggregates using matrix-isolation vibrational circular dichroism (MI-VCD) measurements

The matrix isolation technique is a well-established one which has been extensively applied in IR spectroscopy.[37] For flexible chiral molecules with many possible conformations, the extensive non-covalent interactions with solvent molecules in solution can often lead to fairly broad IR and VCD spectral features with low signal intensity, making the final assignment of the conformational distributions difficult. Therefore, it is desirable to combine MI and VCD techniques together. With the MI-VCD technique, one could often obtain a much narrower bandwidth in the corresponding MI-IR and MI-VCD spectra compared to that of in solution, which will help us obtain more detailed information on the conformational distributions of the chiral molecule of interest. In addition, MI-VCD replaces the solvent with a cold rare gas matrix, avoiding the solute···solvent interactions and eliminating the need to consider a huge number of solute···solvent conformations, allowing one to focus only on those of the solute species alone.

Although MI-VCD experiments are challenging (see description in Chapter 2), some MI-VCD observations have been reported, for example, on conformational distributions and structural information of amino acids,[43] chiral alcohol, amine and acid,[44] and chiral complexes with NH_3.[45] One main goal of using the MI-VCD approach in my thesis research is to compare the dominant monomeric and binary conformations of a chiral acid in cold rare gas matrices with those obtained in a jet expansion and in solution. This comparison allows one to appreciate environmental effects on the conformational preference of chiral molecules and to follow the changes thus induced.

1.4 The importance of investigating the intermolecular interactions of the chiral molecules at the liquid-liquid interface

An exciting new direction I have pursued in my PhD study is to examine how the underlying non-covalent interactions between chiral molecules work at some liquid-liquid interfaces to facilitate chirality discrimination events at interfaces. Liquid-liquid interfacial behavior is of significant importance to the chemical and biological associated areas of

research, such as food emulsions[46] and drug delivery.[47] Non-covalent intermolecular interactions may generate much different changes at a liquid-liquid interface than in bulk solution, leading to some novel phenomena which are not observed in the bulk phase. For example, homochiral and heterochiral enantiomeric pairs might have different binding preference. As a result, rarely seen chemical phenomena could be observed at the liquid-liquid interface.[48] In particular, I aim to probe the underlying chirality discrimination/ recognition phenomena at the organic|water interfaces in my thesis research.

1.5 Outline of the thesis

The present thesis is divided into six chapters. Chapter 1, the current chapter, provides an overview of the chiroptical and interfacial tension force studies included in this thesis. The reminder of this thesis is outlined as below.

In Chapter 2, I describe briefly several instruments used in the thesis research and the general background of IR, VCD, Raman and ROA spectroscopy. The main instruments include a FT-IR and FT-VCD spectrometer and a Raman and ROA spectrometer, The working principles of the associated optical components are illustrated. Furthermore, the matrix isolation (MI) technique and the benefits when combining this technique with the FT-IR and FT-VCD instrument are also discussed. Finally, the pedant drop method is illustrated to show how the liquid-liquid interfacial force measurement are performed at the interface of two different bulk phases.

Chapter 3 reports the conformational distributions and self-aggregation behavior of tetrahydro-2-furoic acid (THFA) by using the MI-VCD spectroscopic method together with a theoretical model. We examined the conformational preference of the THFA molecule and also its self-aggregation behavior at three different matrix temperatures. The final observations are then compared to the related gas phase and solution studies to showcase the power of MI-VCD spectroscopy in revealing the unusual aggregation behavior of chiral molecules in cold rare gas matrices.

In Chapter 4, the THFA species in aqueous solutions under several different pH environments was studied using a combined VCD experimental and theoretical approach. Here we applied the 'clusters-in-a-liquid' model to account for solvent effects. Interesting conformational properties of THFA were compared to the previous reports in a jet expansion,[11b,25] in CCl_4[40] and in cold rare gas matrices.[49]

In Chapter 5, Raman and ROA spectra of a flexible chiral molecule, N-acetyl-L-cysteine (NALC), were measured in water and in methanol. As for NALC molecule, it contains multiple hydrogen bonding sites that can attach with the protic solvent molecules, which is a good chiral model system that can be used to evaluate the solvent effects. Two different solvation approaches are used. One is the 'clusters-in-a-liquid' model where both explicit and implicit solvation effects of water and methanol were evaluated. Furthermore, AIMD simulations were also applied to Raman and ROA measurements in these two solvents. This is the first time the AIMD approach was applied to a flexible molecule, after the previous report[50] on propylene oxide, a rigid chiral molecule.

In Chapter 6, chirality discrimination at a binary toluene (organic)|water interface between R- or S-2,2'-Bis(di-p-tolylphosphino)-1,1'-binaphthyl (Tol-BINAP) molecules and the water-soluble serine chiral species was examined using a combination of interfacial tension measurements (IFT) and MD simulations. This is the first combined experimental and MD study related to chirality discrimination/recognition at the liquid-liquid interface, to the best of our knowledge. The differences in IFT for the homochiral and heterochiral combinations across the organic|water interfaces were captured experimentally and theoretically.

In Chapter 7, the final chapter of my thesis, I summarize the general conclusions of my thesis work and also discuss potential future improvements and further work related to the new chirality discrimination method we developed in this thesis.

References

[1] J. C. Brookes, A. P. Horsfield, A. M. Stoneham, *J. R. Soc. Interface* **2009**, *6*, 75-86.

[2] J. H. Kim, A. R. Scialli, *Toxicological Sciences* **2011**, *122*, 1-6.

[3] L. A. Nguyen, H. He, C. Pham-Huy, *Int J Biomed Sci.* **2006**, *2*, 85-100.

[4] F. O. Kelsey, *Teratology* **1998**, *38*, 221-226.

[5] R. Davis, H. M. Bryson, *Drugs* **1994**, *4*, 677-700.

[6] (a) K. Kozlovskaya, E. Ovchinnikova, J. Kokubun, A. Rogalev, F. Wilhelm, F. Guillou, F. de Bergevin, A. F. Konstantinova, V. E. Dmitrienko, *Crystals* **2021**, *11*, 1389; (b) Q. He, H. Gomaa, S. Rohani, J. Zhu, M. Jennings, *Chirality* **2010**, *22*, 707-716.

[7] (a) G. R. Weisman, *In Asymmetric Synthesis*, J. D. Morrison Ed., Academic Press: New York, **1983**, *1*, 153; (b) J. M. Seco, E. Quiñoá, R. Riguera, *Chem. Rev*. **2004**, *104*, 17-117.

[8] (a) G. Pescitelli, S. Gabriel, Y. Wang, J. Fleischhauer, R. W. Woody, N. Berova, *J. Am. Chem. Soc.* **2003**, *125*, 7613-7628; (b) N. Berova, P. L. Polavarapu, K. Nakanishi, R. W. Woody, *Comprehensive chiroptical spectroscopy*, **2012**, Wiley: Chichester.

[9] (a) L. D. Barron, *Molecular light scattering and optical activity*, **2009**, Cambridge University Press; (b) L. A. Nafie, *Vibrational optical activity principles and applications*, **2011**, Wiley: Chichester; (c) G. Yang, Y. Xu, *Top Curr Chem* **2011**, *298*, 189-236.

[10] (a) F. Zhu, N. W. Isaacs, L. Hecht, L. D. Barron, *Structure* **2005,** *13*, 1409-1419; (b) V. Parchaňský, J. Kapitán, P. Bouř, *RSC Adv.* **2014**, *4*, 57125-57136.

[11] (a) B. H. Pate, L. Evangelisti, W. Caminati, Y. Xu, J. Thomas, D. Patterson, C. Perez, M. Schnell, *Quantitative chiral analysis by molecular rotational spectroscopy*, Chiral analysis, **2018**, Elsevier, Amsterdam, Netherlands, pp 679-729; (b) F. Xie, N. A. Seifert, W. Jäger, Y. Xu, *Angew. Chem. Int. Ed.* **2020**, *59*, 15703-15710; *Angew. Chem.* **2020**, *132*, 15833-15840.

[12] D. Padula, G. Pescitelli, *Molecules* **2018**, *23*, 128.

[13] D. A. Long. *The Raman effect: a unified treatment of theory of Raman scattering by molecules*, **2002**, Wiley: Chichester.

[14] (a) T. B. Freedman, X. Cao, R. K. Dukor, L. A. Nafie, *Chirality* **2003**, *15*, 743-758; (b) P. J. Stephens, F. J. Devlin, J. J. Pan, *Chirality* **2008**, *20*, 643-663; (c) P. L. Polavarapu, *Chirality* **2012**, *24*, 909-920.

[15] L. D. Barron, L. Hecht, I. H. McColl, E. W. Blanch, *Mol. Phys.* **2004**, *102*, 731-744.

[16] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, et al., Gaussian 16, Revision C.03; Gaussian, Inc.: Wallingford, CT, **2019**.

[17] Amsterdam Density Functional Program; Vrije Universiteit: Amsterdam, The Netherlands; http://www.scm.com. E. J. Baerends, T. Ziegler, J. Autschbach, D. Bashford, A. Bérces, F. M.

Bickelhaupt, C. Bo, P. M. Boerrigter, L. Cavallo, D. P. Chong, L. Deng, R. M. Dickson, D. E. Ellis, M. van Faassen, L. Fan, T. H. Fischer, C. Fonseca Guerra, M. Franchini, A. Ghysels, A. Giammona, S. J. A. van Gisbergen, A.W. Götz, J. A. Groeneveld, O. V. Gritsenko, M. Grüning, S. Gusarov, F. E. Harris, P. van den Hoek, C. R. Jacob, H. Jacobsen, L. Jensen, J. W. Kaminski, G. van Kessel, F. Kootstra, A. Kovalenko, M. V. Krykunov, E. van Lenthe, D. A. McCormack, A. Michalak, M. Mitoraj, S. M. Morton, J. Neugebauer, V. P. Nicu, L. Noodleman, V. P. Osinga, S. Patchkovskii, M. Pavanello, P. H. T. Philipsen, D. Post, C. C. Pye, W. Ravenek, J. I. Rodríguez, P. Ros, P. R. T. Schipper, G. Schreckenbach, J. S. Seldenthuis, M. Seth, J. G. Snijders, M. Solà, M. Swart, D. Swerhone, G. te Velde, P. Vernooijs, L. Versluis, L. Visscher, O. Visser, F. Wang, T. A. Wesolowski, E. M. van Wezenbeek, G. Wiesenekker, S. K. Wolff, T. K. Woo, A. L. Yakovlev.

[18] P. J. Stephens, F. J. Devlin, J. R. Cheeseman, M. J. Frisch, *J. Phys. Chem. A* **2001**, *105*, 5356-5371.

[19] (a) G. Li, J. Kessler, J. Cheramy, T. Wu, M. R. Poopari, P. Bouř, Y. Xu, *Angew. Chem. Int. Ed.* **2019**, *58*, 16495-16498; *Angew. Chem.* **2019**, *131*, 16647-16650; (b) T. Wu, G. Li, J. Kapitán, J. Kessler, Y. Xu, P. Bouř, *Angew. Chem. Int. Ed.* **2020**, *59*, 21895-21898; *Angew. Chem.* **2020**, *132*, 22079-22082; (c) G. Li, M. Alshalalfeh, Y. Yang, J. R. Cheeseman, P. Bouř, Y. Xu, *Angew. Chem. Int. Ed.* **2020**, *60*, 22004-22009; *Angew. Chem.* **2021**, *133*, 22175-22180; (d) G. Li, M. Alshalalfeh, J. Kapitán, P. Bouř, Y. Xu, *Chem. Eur. J.* **2022**, e202104302.

[20] (a) D. Laage, T. Elsaesser, J. T. Hynes, *Chem. Rev.* **2017**, *117*, 1069-10725; (b) M. Maurer, C. Oostenbrink, *J Mol Recognit.* **2019**, *32*, e2810; (c) M. Ahmad, W. Gu, T. Geyer, V. Helms, *Nat Commun* **2011**, *2*, 261.

[21] P. Çarçabal, R. A. Jockusch, I. Hunig, L. C. Snoek, R. T. Kroemer, B. G. Davis, D. P. Gamblin, I. Compagnon, J. Oomens, J. P. Simons, *J. Am. Chem. Soc.* **2005**, *127*, 11414-11425.

[22] V. Vaquero, M. E. Sanz, I. Peña, S. Mata, C. Cabezas, J. C. López, J. L. Alonso, *J. Phys. Chem. A* **2014**, *118*, 2584-2590.

[23] J. T. A. Gall, J. Thomas, F. Xie, Z. Wang, W. Jäger, Y. Xu, *Phys. Chem. Chem. Phys.* **2017**, *19*, 29508-29515.

[24] (a) P. Ottaviani, B. Velino, W. Caminati, *Chem. Phys. Lett.* **2006**, *428*, 236-240; (b) N. Borho, Y. Xu, *Phys. Chem. Chem. Phys.* **2007**, *9*, 1324-1328; (c) J. Thomas, O. Sukhorukov, W. Jäger, Y. Xu, *Angew. Chem. Int. Ed.* **2014**, *53*, 1156-1159 ; *Angew. Chem.* **2014**, *126*, 1175-1178.

[25] (a) F. Xie, S. Mahendiran, N. A. Seifert, Y. Xu, *Phys. Chem. Chem. Phys.* **2021**, *23*, 3820-3825; (b) F. Xie, X. Ng, N. A. Seifert, J. Thomas, W. Jäger, Y. Xu, *J. Chem. Phys.* **2018**, *149*, 224306.

[26] C. Pérez, J. L. Neill, M. T. Muckle, D. P. Zaleski, I. Peña, J. C. Lopez, J. L. Alonso, B. H. Pate, *Angew. Chem. Int. Ed*. **2015**, *54*, 979-982; *Angew. Chem.* **2015**, *127*, 993-996.

[27] (a) J. Thomas, Y. Xu, *J. Chem. Phys.* **2014**, *140*, 234307; (b) W. Huang, J. Thomas, W. Jäger, Y. Xu, *Phys. Chem. Chem. Phys.* **2017**, *19*, 12221-12228; (c) B. Wu, A. S. Hazrah, N. A. Seifert, S. Oswald,

W. Jäger, Y. Xu, *J.Phys. Chem. A* **2021**, *125*, 10401-10409; (d) A. Shahi, E. Arunan, *Phys. Chem. Chem. Phys.* **2015**, *17*, 24774-24782; (e) B. Wu, N. A. Seifert, S. Oswald, W. Jäger, Y. Xu, *ChemPhysChem* **2022**, doi.org/10.1002/cphc.202200348..

[28] (a) S. Gobi, E. Vass, G. Magyarfalvi, G. Tarczay, *Phys. Chem. Chem. Phys*. **2011**, *13*, 13972-13984; (b) J. Sadlej, J. Cz. Dobrowolski, J. E. Rode, *Chem. Soc. Rev*. **2010**, *39*, 1478-1488; (c) J. Haesler, I. Schindelholz, E. Riguet, C. G. Bochet, W. Hug, *Nature* **2007**, *446*, 526-529.

[29] K. J. Jalkanen, R. M. Nieminen, M. Knapp-Mohammady, S. Suhai, *Int. J. Quantum Chem.* **2003**, *92*, 239-259.

[30] L. J. Onsager, *Am. Chem. Soc.* **1936**, *58*, 1486.

[31] A. Klamt, G. J. Schuurmann, *Chem. Soc. Perkin Trans.* **1993**, *2*, 799.

[32] J. Tomasi, B. Mennucci, R. Cammi, *Chem. Rev.* **2005**, *105*, 2999-3094.

[33] E. Cances, B. Mennucci, J. Tomasi, *J. Chem. Phys.* **1997**, *107*, 3032.

[34] P. Zhu, G. Yang, M. R. Poopari, Z. Bie, Y. Xu, *ChemPhysChem* **2012**, *13*, 1272-1281.

[35] (a) M. R. Poopari, Z. Dezhahang, G. Yang, Y. Xu, *ChemPhysChem* **2012**, *13*, 2310-2321; (b) M. R. Poopari, Z. Dezhahang, Y. Xu, *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **2013**, *136*, 131-140.

[36] (a) M. Losada, P. Nguyen, Y. Xu, *J. Phys. Chem. A* **2008**, *112*, 5621-5627; (b) M. Losada, Y. Xu, *Phys. Chem. Chem. Phys.* **2007**, *9*, 3127-3135; (c) M. Losada, H. Tran, Y. Xu, *J. Chem. Phys.* **2008**, *128*, 014508; (d) G. Yang, Y. Xu, *J. Chem. Phys.* **2009**, *130*, 164506. (e) A. S. Perera, J. Cheramy, C. Merten, J. Thomas, Y. Xu, *ChemPhysChem* **2018**, *19*, 2234-2242.

[37] A. S. Perera, J. Thomas, M. R. Poopari, Y. Xu, *Front. Chem.* **2016**, *4*, 1-17.

[38] M. R. Poopari, Z. Dezhahang, Y. Xu, *Phys. Chem. Chem. Phys.* **2013**, *15*, 1655-1665.

[39] Y. Liu, G. Yang, M. Losada, Y. Xu, *J. Chem. Phys.* **2010**, *132*, 234513.

[40] K. Le Barbu-Debus, J. Bowles, S. Jähnigen, C. Clavaguéra, F. Calvo, R. Vuilleumier, A. Zehnacker, *Phys. Chem. Chem. Phys.* **2020**, *22*, 26047-26068.

[41] (a) S. T. Mutter, F. Zielinski, J. R. Cheeseman, C. Johannessen, P. L. A.Popelier, E. W. Blanch, *Phys. Chem. Chem. Phys.* **2015**, *17*, 6016-6027; (b) S. Ghidinelli, S. Abbate, J. Koshoubu, Y. Araki, T. Wada, G. Longhi, *J. Phys. Chem. B* **2020**, *124*, 4512-4526.

[42] M. Thomas, Theoretical modeling of vibrational spectra in the liquid phase. *Springer International Publishing*, **2017**.

[43] G. Pohl, A. Perczel, E. Vass, G. Magyarfalvi, G. Tarczay, *Phys. Chem. Chem. Phys.* **2007**, *9*, 4698-4708.

[44] (a) C. Merten, Y. Xu, *ChemPhysChem* **2013**, *14*, 213-219; (b) C. H. Pollok, C. Merten, *Phys. Chem. Chem. Phys.* **2016**, *18*, 13496-13502; (c) A. S. Perera, J. Cheramy, M. R. Poopari, Y. Xu, *Phys. Chem. Chem. Phys.* **2019**, *21*, 3574-3584.

[45] C. Merten, Y. Xu, *Angew. Chem. Int. Ed.* **2013**, *52*, 2073-2076; *Angew. Chem.* **2013**, *125*, 2127-2130.

[46] A. G. Gaonkar, *JAOCS* **1989**, *66*, 1090–1092.

[47] G. T. Vladisavljević, M. Shimizu, T. Nakashima, *J. Membr. Sci.* **2006**, *284*, 373-383.

[48] (a) H. Watarai, N. Sawadam T. Sawada (eds.), *Interfacial Nanochemistry*, Kluwer Academic /Plenum Publishers, New York, **2005**; (b) A. G. Volkov, D. W. Deamer, D. L. Tanelian, V. S. Markin (eds.), *Liquid Interfaces in Chemistry and Biology*, Wiley, New York, **1998**.

[49] Y. Yang, J. Cheramy, Y. Xu, *ChemPhysChem* **2021**, *22*, 1336-1343.

[50] M. Brehm, M. Thomas, *J. Phys. Chem. Lett.* **2017**, *8*, 3409-3414.

Chapter 2

Experiment and Theory

2.1 Introduction

In my current thesis, I have utilized the vibrational spectroscopic tools, for example, IR, VCD, Raman and ROA spectroscopy, to investigate the non-covalent interactions and also the absolute configurations of the chiral molecules in solution. These techniques are not only sensitive to the chirality, but also highly sensitive to the small structural changes induced by non-covalent intra- and intermolecular interactions. To extract structural information from the experimental results, it is necessary to utilize theoretical calculations to simulate the corresponding IR, Raman, VCD and ROA spectra and to further use them to interpret the experimental observations. Both the instrumentation and theoretical aspects will be concisely described below.

To extend the scope of our measurements, I also combined our current IR and VCD spectroscopic method with the matrix isolation (MI) technique. The MI-VCD technique affords one to avoid solute-solvent interactions and a better control of the sample composition, allowing identification of individual isolated monomeric forms or the molecular clusters. The MI-VCD set up will be discussed in Section 2.2.2.

Finally, we also developed a new method to probe chirality discrimination at the organic|water interfaces using interfacial force tension (IFT) measurements. The method applied to extract IFT values will be briefly described below together with some theoretical background information.

2.2 Instrumentation

2.2.1 History of VCD and ROA instruments

The first experimental VCD measurement was reported by Holzwarth and co-workers in 1974, where both enantiomers and the racemic mixture of 2,2,2-trifluoro-1-phenylethanol in neat solution was measured in the C-H stretching region. Moreover, R-neo-pentyl-l-d-chloride was also measured in the C-D stretching region.[1] The VCD of the tartaric acid in the O-H stretching region was reported by Sugeta and co-workers.[2] Two years later, Nafie and co-workers revisited the same chiral molecule in the C-H stretching

region by using the VCD spectrometer with the addition of two ZnSe modulators. Such changes enhanced the performance in the mid-infrared spectral region.[3] Later, they extended the measured spectral region to O-H and N-H stretching regions for many other chiral molecules.[4] These original VCD measurements were carried out using a dispersive scanning instrument, where only a limited spectral region can be recorded. In 1979, the first application of a FT-VCD measurement of camphor in CCl_4 solution in the C-H stretching region was reported.[5] This experiment demonstrated the power of a FT-IR/VCD spectrometer with better signal-to-noise ratio achieved in less measuring time when compared to the conventional dispersive scanning method.[5] Some further improvements, such as a HgCdTe liquid-nitrogen cooled detector and a CaF_2 polarizer, were incorporated in the FT-IR/VCD instrument, affording measurements in the mid-infrared fingerprint region (900-1600 cm^{-1}).[6] Although some initial commercial attempts to sell the VCD instruments by BioRad (Digilab) and Nicolet (Thermo Nicolet) failed in the 1980s due to some instrumentation problems,[7] Bomen/BioTools successfully marketed the first commercialized standalone Chiral*IR* FT-VCD spectrometer in the mid-1990s.[8] Today, VCD accessories and standalone spectrometers are available from many other companies, for example, Bruker, BioTools, Jasco and Thermo-Electron.

The first experimental ROA spectrum was reported by Barron and co-workers in 1973, where both enantiomers of α-phenylethylamine and α-phenylethanol in neat liquid were measured in the spectral region between 290 cm^{-1} and 400 cm^{-1} with the 90 degrees scattering arrangement.[9] These experimental observations were later validated by Hug and co-workers in 1975.[10] Initially, it was very difficult to obtain a good ROA spectrum because of the poor sensitivity and long acquisition time. Some significant improvements were implemented in the development of the ROA instrument. Barron and other co-workers found out that the ROA intensity could be maximized if one adopts the 180 degrees backscattering or zero degrees which is in forward direction.[11] It was later implemented by Barron, Hug and others and discovered that the signal to noise ratio of the measured

ROA spectrum could be improved by eight times when compares to the 90 degrees scattering arrangement.[12] Apart from that, the same research group later replaced some optical components and the detector with the newly innovated technologies, such as back thinned charged-coupled device (CCD) detector and high efficiency notch and edge filters.[13] In 2003, Hug[14] made some additional improvements on eliminating optical offsets by introducing the other pair of enantiomer using a half-wave plate, where flipping the handedness of the circularly polarized light. Such concept is also known as the generating of virtual enantiomer. The ROA instrument is now commercialized and manufactured by BioTools.

2.2.2 VCD Instrumentation

A FT-IR/VCD Spectrometer: Infrared radiation covers the region of the electromagnetic radiation spectrum from 10 cm^{-1} to 14000 cm^{-1}. The IR waves are shorter than the microwaves and radio waves, but longer than the visible light. The IR can be further split into three distinctive spectral regions based on the wavelength, which are far-infrared (FIR), 10-400 cm^{-1}, mid-infrared (MIR), 400-4000 cm^{-1}, and near-infrared (NIR), 4000-14000 cm^{-1}. The fingerprint region is located in the MIR region which is from 400-1500 cm^{-1} region. This region behaves like human's fingerprint, where it has unique and distinctive vibrational bands associated with different functional group in the molecules, for example amide I and amide II region from amino acids and proteins. These vibrational bands are mainly consisting of a combination of stretching and bending motions, which will be very helpful to use to extract structural information from the chiral molecules of interest through observed experimental results in the fingerprint region.

All the IR and VCD measurements in the thesis were carried out using the FT-IR/VCD instrument, where the schematic diagram is shown in Figure 2.1. The FT-IR/VCD instrument consists of a FT-IR spectrometer (Vertex 70, Bruker) and a VCD module (PMA 50, Bruker).

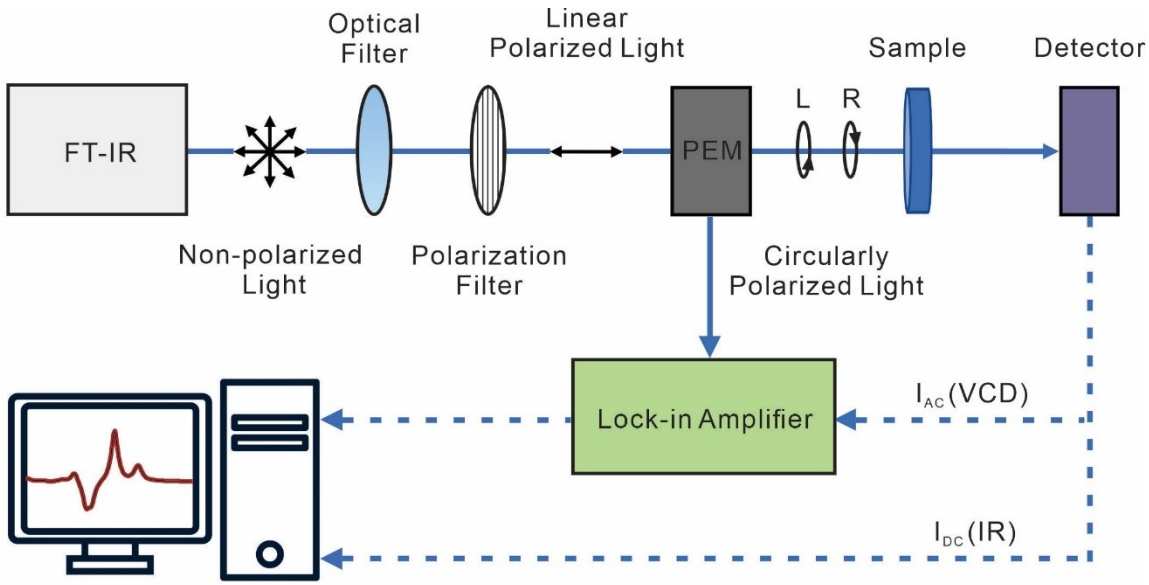


Figure 2.1. The schematic diagram of the FTIR-VCD spectrometer.

The current FTIR-VCD instrument is equipped with a common mid-IR light source, i.e., a globar (a U-shaped silicon carbide piece). It is air-cooled, with no other special cooling requirement. The IR radiation generated from the globar was then passed through a Michelson interferometer. The working principle of a Michelson interferometer is demonstrated in Figure 2.2. Generally, the parallel IR light rays are generated after passing through a collimating mirror, then the light rays hit the beam splitter where half of the light is transmitted to a fixed mirror, and the other half is reflected to a movable mirror. The light beams from both paths are then reflected back and recombined at the beam splitter with a path difference of 2d. Based on the phase differences between the two reflected beams, they will constructively or destructively interfere with each other.

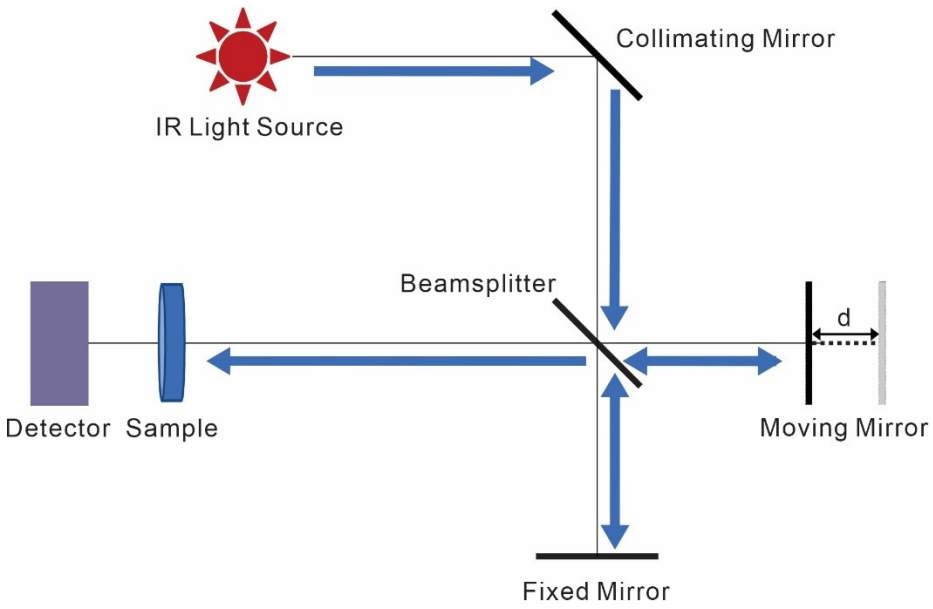


Figure 2.2. The schematic diagram of the Michelson interferometer in the FTIR-VCD instrument.

If the reflected beam from the fixed and the moving mirror are in phase, constructive interference results, and the following equation applies:

$$2d = n\lambda \tag{2.1}$$

where 2d is the optical path difference, n is the integer number and λ is the wavelength of the light. If the optical path difference is zero, i.e., n = 0, the recombined reflected light from both the movable mirror and the fixed mirror will be in phase with each other at all points, and the wave amplitude will add up in such a way to produce a much stronger intensity beam. Furthermore, the recombined light beam will also be in phase if the optical path difference is an integer multiple of the wavelength, i.e., n = 1, 2, 3, etc. On contrary, the recombined beams will have destructive interference if they are out of phase with each other, i.e., when the equation can be expressed as:

$$2d = (n + \frac{1}{2}\lambda) \tag{2.2}$$

when the optical path length is a half integer multiple of the wavelength, i.e., $\frac{1}{2}, \frac{3}{2}, \frac{5}{2},$ etc. Under such conditions, the recombined beams will be exactly out of phase and the

amplitude of the light beams from each mirror cancel each other out.

Next, the recombined IR beams pass through the sample and finally reach the detector. The detector we used is a HgCdTe (MCT) detector which operates under liquid N_2 cooling. The resulting recombined interferogram from IR light source, which is the combination of multiple cosine wave functions, will be received by the MCT detector. The further Fourier transformation (FT) will be applied to generate the frequency domain signal at each wavenumber.

For the FT-VCD part of the instrument, the unpolarized IR radiation from the FT-IR compartment is directed into the VCD chamber. It passes through the optical filter which has the different options for the cut-off value in the spectral region, for example above 1800 cm^{-1} or 3800 cm^{-1}. The randomly polarized IR light further goes through a wire grid linear polarizer where the polarizer grid lines run vertically, allowing the horizontally polarized component to pass through. The horizontally polarized light then reaches a photoelastic modulator (PEM). PEM is made of a transparent fused silica material and ZnSe, that is connected to a piezoelectric transducer. The key components in the PEM are shown in Figure 2.3. The main goal of the PEM is to convert the linearly polarized light into the left- and right-circularly polarized light with a 50 kHz modulation frequency. With this modulation technique, it can allow us to record very small VCD signals, typically 10^{-4} to 10^{-6} weaker than the parent IR intensities.

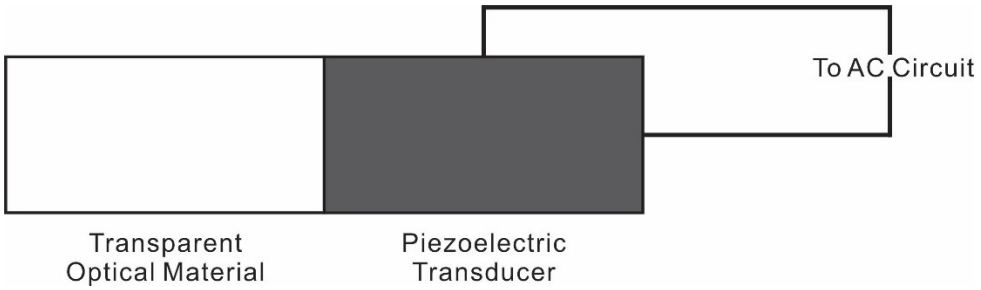


Figure 2.3. The key components of the PEM.

The modulator axis of the PEM is set at 45 degrees with respect to the plane of polarization of the incoming light which consists of two equally distributed horizontal and vertical components. The piezoelectric transducer is connected to a 50 kHz alternating current (AC) circuit to be stretched or compressed and become birefringent. As a result, the horizontal and vertical components of the incoming light travel with slightly different speeds when they pass through the PEM, leading to generation of circularly polarized light as shown in Figure 2.4. The important outcomes for the VCD measurement happen when the peak retardation reaches exactly one fourth of the wavelength of light. Generally speaking, the PEM works as an oscillating quarter waveplate to produce the right-circularly polarized light when it reaches to the maximum in the stretching cycle (i.e., the phase retardation is exactly 90 degrees), and left-circularly polarized light when it reaches the maximum in the compressing cycle (i.e., the phase retardation is exactly 270 degrees).

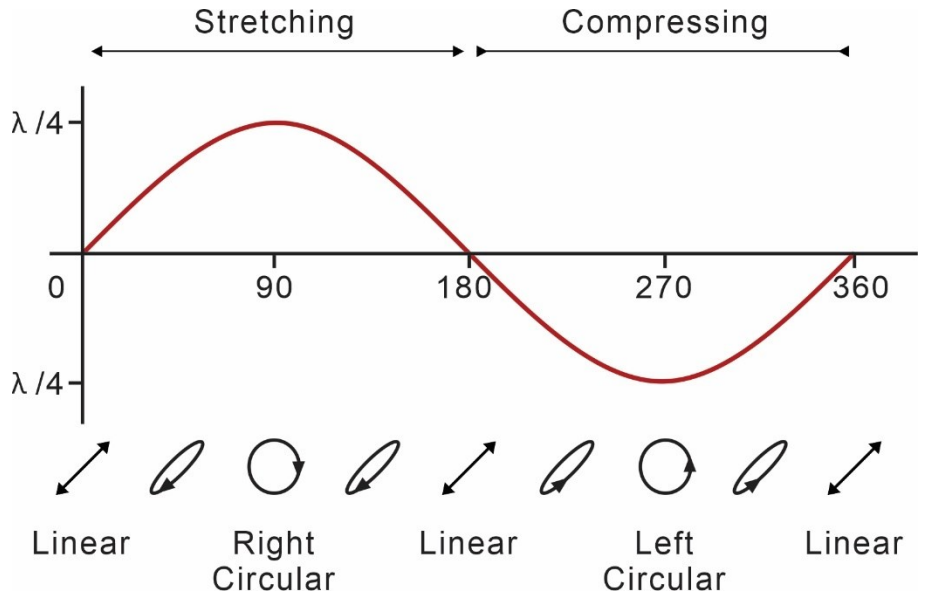


Figure 2.4. The principles of operation for the PEM. When there is no stress applied on the optical material, the speed of the horizontal and vertical components of the incoming light is not affected, resulting in the linearly polarized light where the polarization plane rotates 45 degrees in the same direction as the modulation axis of PEM. If the optical material is stretched, the horizontal component "lags" behind the vertical

component. On the other hand, if the optical material is compressed the vertical component "lags" behind the horizontal one. The phase difference between the horizontal and vertical components at any instant of time is called retardation.

The left- and right-circularly polarized light then propagate through the chiral sample and finally reach the MCT detector. In the detector, the IR beam intensity is converted to an electrical signal which consists of two intensity components. One is the low frequency direct current (DC) signal which is obtained by passing through a low pass filter. The DC signal is then directly sent to the electronics and generates the experimental IR spectrum. The other one is the high frequency alternating current (AC) signal which is obtained by passing through a high pass filter. Since the VCD signal is the difference in absorbance of the left and right circularly polarized light for a given fundamental vibrational transition, the VCD intensities are very small, ~10^{-4} to 10^{-6} of the corresponding IR intensity. To capture such small signals, a lock-in-amplifier (LIA) is used to demodulate the high frequency AC signal. Both the low frequency DC signal and the demodulated AC signal are then Fourier transformed to convert to the frequency domain spectrum. The final VCD spectrum is generated based on the ratio of the AC and DC Fourier transformed spectra.

Key requirements for VCD measurements: To obtain good IR and VCD spectra, several considerations need to be taken into account. First, one needs to choose a proper solvent based on the targeted chiral sample. In other words, the chiral solute must have a good solubility in the solvent you selected and also the solvent cannot have the strong absorption band in the region of interest, for example, 1000-1800 cm^{-1}. A solubility test is a good starting point if there is no data for the solubility of the chiral sample in a particular solvent. Once the appropriate solvent is chosen, the next important parameter one needs to decide is the concentration of the solution. Typically, a good VCD signal can be obtained if the IR absorbance is between 0.2-0.9. Otherwise, one might experience the saturation effects or

very weak VCD signals.

Based on the Beer-Lambert law, there are two approaches that can be used to achieve the desirable IR absorbance for the system of interest. One is to manipulate the solution concentration, and the other is to adjust the path length by using different sizes of spacers. Apart from that, it is also crucial to choose the right IR windows for the experimental measurements. The most common IR windows in the market are made of KBr, CsI, BaF_2, ZnSe, KRS-5, and CaF_2, where they all have their specific advantages. We typically use CaF_2 and BaF_2 where the former has a cut-off value of 1100 cm^{-1}, while the latter is at 850 cm^{-1}. Based on the pros and cons of each material for the IR window, BaF_2 is the best candidate that meets all the requirement of our investigations. It is unfortunately quite brittle and can break easily with thermal and mechanical shock.

For IR measurements, it usually takes less than a minute to obtain an IR spectrum with a good signal to noise ratio. Typically, a measurement with 32 scans is enough. However, it is not the case for the VCD measurements. One needs to spend a much longer measurement time, usually 3 to 6 hours measurement (13000-26000 scans) in order to capture the very weak VCD signals and achieve a good signal to noise ratio.

MI-VCD instrumentation: To obtain sharper VCD bands, one could apply the MI-VCD technique. The layout of the MI system in our laboratory is shown in Figure 2.5. It consists of three parts: a cryostat, a vacuum system, and a sample deposition system. The cryostat, i.e., a closed cycle cryogenic refrigerator, is composed of a cold head (DE 240SI) and a helium compressor (ARS-4HW) from Advanced Research Systems. This part of the MI system remains the same as it was originally put together in 2011.[15]

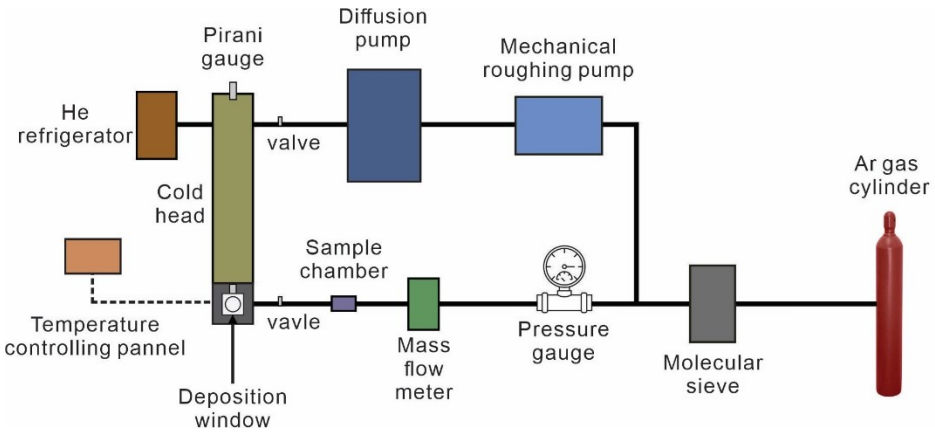


Figure 2.5. The schematic diagram of the MI system.

A new pumping system was put together during my thesis research where the turbo pump was replaced with a diffusion pump. The pumping system consists of a mechanical roughing pump (SD-450, Varian) and a diffusion pump (Edwards). The cold head needs to be first evacuated by the mechanical roughing pump to below 0.1 mbar before starting the diffusion pump. With the help of the diffusion pump, one can reach as low as 2.0×10^{-8} mbar at 10 K before deposition. Both liquid nitrogen and water-cooling systems are used to prevent oil from travelling up to the cold head while operating the diffusion pump. To control and monitor the temperature in the cold head, two silicon diode temperature sensors, which are wired to the temperature controller (Model 311, Lake Shore), are placed inside. One sensor is located on the deposition window and the other one is located on the cold head. The temperature difference between the two sensors is about 0.25-0.5 K at 333 K and around 3 K at 10 K.

The last component of the MI system is the sample deposition line. The residual water in the sample line system is removed by baking the sample lines and connecting them to some drying tubes containing molecular sieves (4 Å). A gas mixture was prepared by premixing a trace amount of the gas phase chiral target molecules with rare gas, for example, argon gas, at room temperature. The gas mixture is then introduced into the cold head through a needle and onto the cold window attached to the cold head. If the sample is

in solid form with a low vapor pressure, one could place the solid sample in a stainless-steel sample container and heat it to some desired temperature while controlling the mixing by using a mass flow meter (Model 1197A, MKS) to adjust the flow rate of the argon gas.

We monitor the sample mixture by taking IR spectra. If the target molecule vapor is too dilute, we can either reduce the flow rate of Ar gas or raise the sample temperature in order to achieve an IR absorbance of 0.2-0.9 in a reasonable time frame. The other important factor to be considered is the transparency of the window after sample deposition. Failing to obtain a transparent window would lead to greater light scattering. As a result, one may observe noticeable baseline elevation, making it very difficult to obtain a good quality VCD spectrum. To achieve a transparent window after sample deposition, one needs to manipulate the temperature of the cold surface and the deposition rate of the sample mixture. In addition, different degrees of freedom for the self-aggregation behavior of the target molecule can be investigated by increasing the temperature of the cold head in a stepwise manner.

2.2.3 ROA Instrumentation

All the Raman and ROA measurements in the thesis were carried out using the commercial ROA instrument, ChiralRAMAN-2X, from BioTools. The simplified layout of the instrument is shown in Figure 2.6. In general, it consists of two pathways. One is the incident light pathway, and the other is the scattered light pathway. First, I will work through each optical component in the incident light pathway and then in the scattered light pathway.

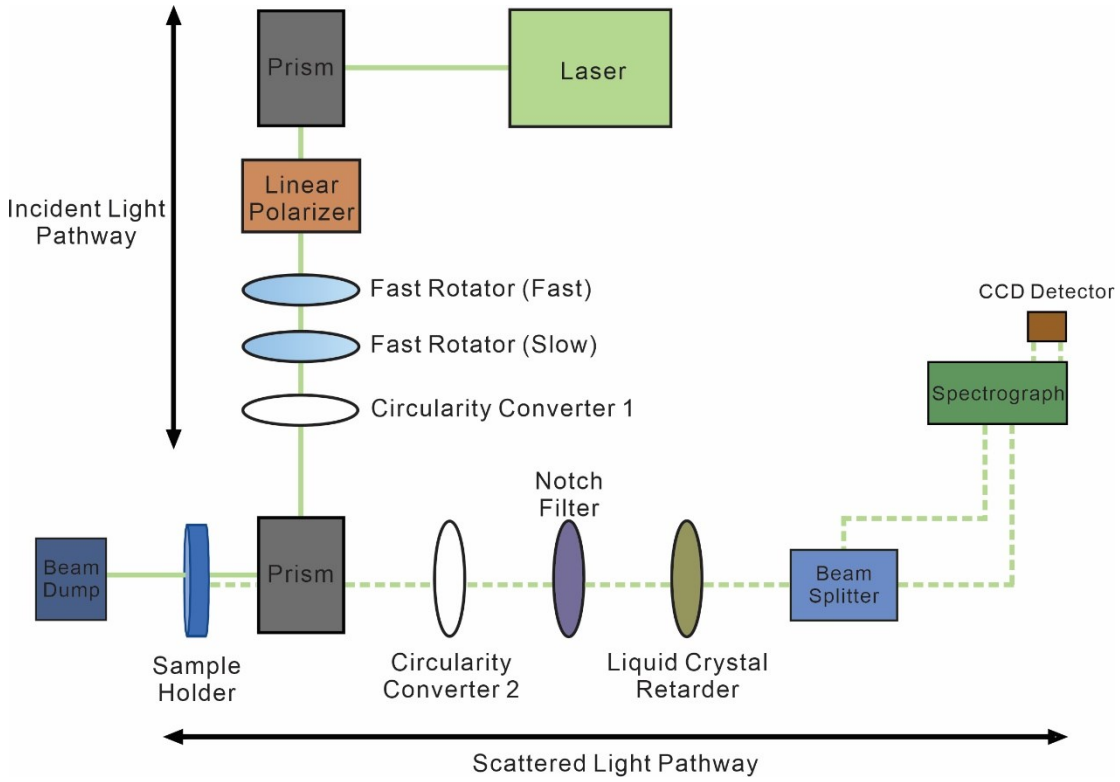


Figure 2.6. Simplified schematic diagram of ROA instrument.

Incident light pathway: The laser source we used in the current ROA instrument is a green diode laser (Laser Quantum) that has an excitation wavelength of 532 nm with adjustable power output from 0 to 2 W. As for the laser light, it is coherent and horizontally polarized. It is first passed through a prism which is used to scatter the light and change the direction of propagation. The laser light then goes through a linear polarizer, where any unwanted circularly polarized light is removed. The linear polarized light was further introduced to a set of fast rotators, where these two fast rotators act as two separate half waveplates and rotate at different speeds. The goal of this setup is to shift the polarization direction between horizontal and vertical, resulting in randomly polarized light. The randomly polarized light then reaches to the circularly converter one which also works as a half waveplate and is used to correct any unwanted circularly polarized light. The randomly polarized light finally reaches the second prism. In here, the propagation direction of the incident lights is

changed toward the sample cell, a quartz cuvette which can be placed in the mounted cell holder. In the meanwhile, they allow the parallel scattered light to pass through.

Scattered light pathway: The current ROA instrument is designed in the backscattering arrangement which maximizes the intensity of the ROA signals. In the scattered light pathway, only the back scattered light from the sample was collected and then reached the circularly converter two, another half waveplate. It flips the handedness of the scattered circularly polarized light and creates the "virtual enantiomers". This step is crucial for removing the optical offsets. The scattered light then passes through the notch filter (Kaiser Optical Systems Inc.) to get rid of the very strong Rayleigh scattering from the sample. It is then introduced to the liquid crystal retarder which acts as a quarter waveplate. At the liquid crystal retarder, the scattered left- and right-circularly polarized light is converted into s- and p-linearly polarized light with respect to the fast and slow axis of the liquid crystal retarder. Finally, both linearly polarized components reach the beam splitter. They are then split out into two different fiber optic branches, namely as the s- and p-branch. The spectrograph will separate the incoming scattered light based on its wavelength and it is finally collected by the CCD detector.

2.2.4 Pendant drop method

Surface tension measures the tendency of liquid surfaces at rest to shrink into the minimum surface area possible and is usually evaluated at air-liquid interfaces.[16] Interfacial tension (IFT) is generally associated with liquid-liquid interfaces. In the current thesis, I applied a technique called the pendant drop method for the IFT measurements. This is a convenient and versatile method to measure the interfacial behavior of two immiscible liquids. The schematic diagram of the experimental setup with the pendant drop method (Model 250, ramé-hart) is shown in Figure 2.7. It consists mainly of two sections. One is the image acquisition, and the other is image processing section.

First, the aqueous phase is placed inside a quartz cell which is placed on a 3-axis leveling specimen stage where one can freely move the position of the quartz cell into an optimized position. The organic sample is loaded into the glass syringe that is connected with a U-shaped needle. The tip of the needle is then immersed inside the aqueous medium to generate an organic phase droplet at the tip of the needle. The droplet is subsequently illuminated by means of a fiber optic illuminator. The shape of the organic phase droplet is finally captured by the CCD camera and then the image is processed using the DROPimage software.

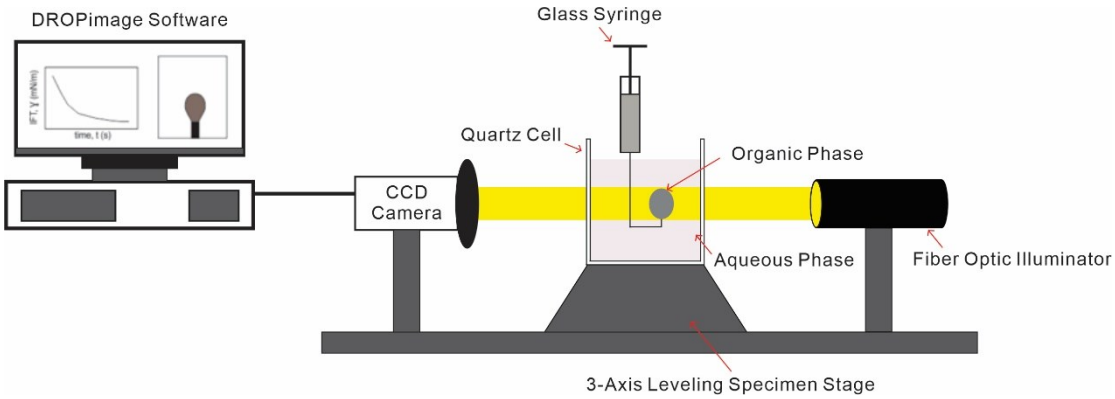


Figure 2.7. The schematic diagram of the experimental setup using the pendant drop method.

Second, the interfacial tension of a given system is determined through a two-step process. The size parameters of β and R_0 are first determined from the shape profile of the organic droplet that is captured by the CCD camera. In the next step, the interfacial tension is calculated by applying the following equation:

$$\gamma = \frac{[\Delta\rho \cdot g \cdot (R_0)^2]}{\beta} \qquad (2.3)$$

where Δρ is the mass density difference between the drop phase (e.g., toluene) and the surrounding medium (e.g., water), g is the gravity constant, R_0 is the radius of curvature at the drop apex and finally, β is the shape factor. Figure 2.8 shows the dimensions and

symbols used for the determination of interfacial tension.

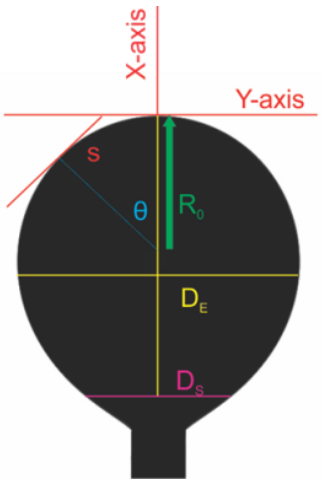


Figure 2.8. Droplet image with highlighted dimensions and symbols for the purpose of IFT calculation.

To calculate β, the parameter $\sigma = \frac{D_S}{D_E}$, where D_E is the maximum diameter of the droplet and D_S is the diameter at distance D_E from the drop apex, should be first calculated. Afterward, the calculated σ is then used in the following third-degree polynomial equation to obtain the corresponding shape factor parameter:

$$\beta = -0.12836 + 0.7577\,\sigma - 1.7713\,\sigma^2 + 0.5426\,\sigma^3 \qquad (2.4)$$

Then the R_0 parameter is calculated from the below equation with knowing the shape factor:

$$\frac{D_E}{2R_0} = 0.9987 - 0.1971\,\beta - 0.0734\,\beta^2 - 0.34708\,\beta^3 \qquad (2.5)$$

2.3 Theoretical

2.3.1 Basics of IR and VCD spectroscopy

Consider a simple harmonic oscillator model for a diatomic system, where the nuclei of the atoms in the diatomic molecule can be treated as point masses that are connected by a massless spring, which obeys Hooke's law. If an atom is pulled from the equilibrium with a displacement of *x*, a restoring force *F* which is proportional to the spring displacement is

then generated, $F = kx$. The vibrational potential energy for the diatomic molecule in which the nuclei are at the stationary position can be expressed as:

$$V = \frac{1}{2}kx^2 \tag{2.6}$$

where k is the force constant. Therefore, the fundamental vibrational frequency, ν, of the diatomic molecule can be calculated as the equation:

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}} \tag{2.7}$$

where μ is the reduced mass of the system, defined as below:

$$\mu = \frac{m_1 \cdot m_2}{m_1 + m_2} \tag{2.8}$$

The vibrational energy of a quantum mechanical harmonic oscillator is not a continuous function, but can be rather expressed as

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega, \quad (n = 0, 1, 2, 3, \dots) \tag{2.9}$$

where n is the vibrational quantum number, E_n is the vibrational energy of a given energy level n, $\hbar$ is the reduced Planck's constant, where $\hbar = \frac{h}{2\pi}$, and ω is the angular frequency, where $\omega = 2\pi\nu$. The lowest energy is $\frac{1}{2}\hbar\omega$ where the vibrational quantum number is equal to 0, which is also known as the zero-point energy level or ground vibrational state. The energy levels of the harmonic oscillator system are quantized and equally spaced and the transition is allowed as the vibrational quantum number changes by 1 ($\Delta n = \pm 1$) in the harmonic oscillator.

The IR intensity gives rise by the term called dipole strength, D_{01}^i, which is proportional to the electric dipole moment. In the harmonic approximation, the dipole strength for a given fundamental vibrational transition (from n = 0 to n = 1) at the i^{th} normal mode can be expressed as[17]

$$D_{01}^i = \left|\vec{\mu}_{01}^i\right|^2 = \left|\langle\psi_0|\vec{\mu}|\psi_1^i\rangle\right|^2 \tag{2.10}$$

Here, $\vec{\mu}$ is the electric dipole moment operator, ψ_0 is the wavefunction for the ground vibrational state, ψ_1^i is the wavefunction for the first excited vibrational state with respect

to the i^{th} normal mode, and the term $\langle\psi_0|\vec{\mu}|\psi_1^i\rangle$ is called the electric dipole transition moment.

For the VCD intensity of the given transition, it is proportional to the term called the rotational strength, R_{01}^i, which is the scalar product between the electric dipole and magnetic dipole transition moments, that can be written as

$$R_{01}^i = \langle\psi_0|\vec{\mu}|\psi_1^i\rangle \cdot \mathrm{Im}\langle\psi_1^i|\vec{m}|\psi_0\rangle = |\vec{\mu}_{01}^i| \cdot |\vec{m}_{01}^i| \cdot cos\theta(i) \qquad (2.11)$$

where $\vec{m}$ is the magnetic dipole moment operator. For the rotational strength expression, a new term, $\langle\psi_1^i|\vec{m}|\psi_0\rangle$, is considered which is the magnetic dipole transition moment. Therefore, when calculating the rotational strength, one needs to consider both electric dipole and magnetic dipole transition moments. It is relatively simple to simulate the dipole strength of the given vibrational transition of a normal mode by applying the Born-Oppenheimer (BO) approximation. It is, however, not the case for the magnetic dipole transition moment. One could lose the electronic contribution to a vibrational magnetic dipole transition under the BO approximation.[18] To better evaluate the magnetic dipole transition moments, some theoretical methods are used, for example, vibronic coupling[19], localized molecular orbital[20], charge flow model[21], and the coupled oscillator model[22]. Apart from these, Buckingham[23] and Stephens[24] made some significant advances by using the magnetic field perturbation (MFP) theory which works as a correction term that can be directly applied to the BO approximation. In this theory, the problems in the electronic contribution to the magnetic dipole transition moment can be corrected with consideration of the first derivatives of the ground state electronic wavefunction with respect to the nuclear displacements and the applied magnetic field.

As for the rotational strength expression, it can be further written as the scalar $|\vec{\mu}_{01}^i|$ $|\vec{m}_{01}^i|$ $cos\theta(i)$. The sign of the VCD signal is given by the $cos\theta(i)$ term which is the angle between the electric dipole and magnetic dipole transition moment vectors. Due to the different orientations between these two vectors, one could obtain different signs

(positive or negative) in the corresponding VCD spectrum. For the given vibrational mode i, $cos\theta$ is positive if $0^o < \theta < 90^o$ and $270^o < \theta < 360^o$. On the other hand, $cos\theta$ is negative if $90^o < \theta < 270^o$. However, when θ is exactly at 90^o or 270^o, one will observe zero intensity in the VCD spectrum since $cos\theta$ is equal to zero. It is also important to note that if θ is close 90^o or 270^o, even a very small perturbation in either experiment or theoretical simulation can result in altering the sign of the VCD signal.

To calculate the intensity of the simulated IR and VCD signals, the quantities called atomic axial tensor (AAT)[23] and atomic polar tensor (APT)[25] are used to evaluate the corresponding electric dipole and magnetic dipole transition moments at each normal mode. The MFP approach was also included into electronic structure packages such as, Gaussian and Dalton for calculated the AAT, APT, and the harmonic force field.[26] However, the calculated IR and VCD intensities for the transition associated to each vibrational mode is just a single delta function with a certain intensity value. To interpret the experimental IR and VCD spectra, a Lorentzian band shape with a desirable bandwidth is applied to the simulated results which make them comparable with the experiments.

The expression for converting the set of dipole and rotational strengths to the comparable IR and VCD spectra can be written as[8]

$$\varepsilon(\upsilon) = \frac{\upsilon}{9.184\times10^{-39}} \sum_i D_i f_i(\nu) \tag{2.12}$$

$$\Delta\varepsilon(\upsilon) = \frac{\upsilon}{2.296\times10^{-39}} \sum_i R_i f_i(\nu) \tag{2.13}$$

Here, ε and $\Delta\varepsilon$ (M^{-1} cm^{-1}) are the molar absorptivity coefficient and differential molar absorptivity coefficient, respectively, which are proportional to the dipole strength, D, and rotational strength, R, where both are in esu^2 cm^2. $f_i(\nu)$ is the normalized Lorentzian line shape function with a half-width of the vibrational band at half maximum, γ_i, at each vibrational mode i and transition frequency (υ). The line shape function is expressed as

$$f_i(\nu) = \frac{\gamma_i/\pi}{(\upsilon_i-\upsilon)^2+\gamma_i{}^2} \tag{2.14}$$

2.3.2 Basics of Raman and ROA spectroscopy

Raman spectroscopy is an optical technique where the incident monochromic light interacts with the sample of interest, which then scatters the photons instead of absorbing as in IR spectroscopy. The expression for the electromagnetic wave, E, oscillating with the frequency, v_0, can be written as

$$E = E_0 \cos(2\pi v_0 t) \tag{2.15}$$

where E_0 is the amplitude. The oscillating electric field will cause an induced dipole of the molecule, μ_{ind}, which is proportional to the polarizability of the molecule, α. The induced dipole can be expressed as

$$\mu_{ind} = \alpha E = \alpha E_0 \cos(2\pi v_0 t) \tag{2.16}$$

In general, the polarizability is a tensor and changes with time related to the nuclear coordinates and atomic displacements that oscillate for a given normal mode, Q_k, with a frequency, v_k, which can be then expressed as

$$\alpha = \alpha_0 + \left(\frac{\partial \alpha}{\partial Q_k}\right) \cos(2\pi v_k t) \tag{2.17}$$

By substituting Eq. 2.17 into 2.16, it yields

$$\mu_{ind} = E_0 \begin{Bmatrix} \alpha_0 \cos(2\pi v_0 t) + \frac{1}{2}\left(\frac{\partial \alpha}{\partial Q_k}\right) Q_k \cos[2\pi(v_0 + v_k]t + \\ \frac{1}{2}\left(\frac{\partial \alpha}{\partial Q_k}\right) Q_k \cos[2\pi(v_0 - v_k]t \end{Bmatrix} \tag{2.18}$$

From equation 2.18, one can see that the induced dipole consists of three cosines terms with different frequencies which are related to the polarizabilities. The first term is Rayleigh scattering, also known as elastic scattering, where the frequency of the scattered light remains unchanged when compared to the frequency of the incident light. The other two terms correspond to the inelastic scattering which is also known as Raman scattering. When the frequency of the scattered light is higher than the incident one, we call the process anti-Stokes scattering. The Stokes scattering, on the other hand, refers to the frequency of the scattered light which is lower than that of incident light. Such processes happen when the molecule is excited to a virtual state which can be regarded as a superposition state and

is never populated, by absorbing a photon and followed by relaxing to the ground state and emitting a photon. Depending on the initial and final vibrational states in the process, it will lead to the different scattering processes.

For the Raman intensity, one requires the electric dipole-electric dipole polarizability tensor, a, which can be expressed as[27]

$$\alpha_{\alpha\beta} = \frac{\hbar}{2}\sum_{j\neq i}\frac{\omega_{ji}}{\omega_{ji}^2-\omega^2} Re\left(\langle i|\hat{\mu}_a|j\rangle\langle j|\hat{\mu}_\beta|i\rangle\right) \quad (2.19)$$

where $\omega_{ji} = \omega_j - \omega_i$ is the differential angular frequencies between the initial (i) and virtual (j) states of the molecule.

For the ROA intensity, two additional tensors need to be considered. One is the electric dipole-magnetic dipole optical activity tensor, G', and the other is the electric dipole-electric quadrupole optical activity tensor, A, that are expressed as

$$G'_{\alpha\beta} = -\frac{\hbar}{2}\sum_{j\neq i}\frac{\omega_{ji}}{\omega_{ji}^2-\omega^2} Im\left(\langle i|\hat{\mu}_a|j\rangle\langle j|\hat{m}_\beta|i\rangle\right) \quad (2.20)$$

$$A_{\alpha,\beta\gamma} = \frac{\hbar}{2}\sum_{j\neq i}\frac{\omega_{ji}}{\omega_{ji}^2-\omega^2} Re\left(\langle i|\hat{\mu}_a|j\rangle\langle j|\widehat{\Theta}_{\beta\gamma}|i\rangle\right) \quad (2.21)$$

where m is the magnetic dipole, Θ is the electric quadrupole and α, β and γ are defined as Cartesian coordinates.

To calculate the intensity of the simulated Raman and ROA signals, one needs to carefully evaluate the corresponding tensors at each normal mode of the targeted molecule. To interpret the experiment Raman and ROA spectra, we utilized the same protocol as the one used for IR and VCD by applying a Lorentzian band shape with a desirable bandwidth on the simulated results which make them comparable with the experiments. For Raman and ROA spectra, we typically used 10 cm^{-1}.

2.3.3 Conformational searches and spectral simulations

To carry out a comprehensive conformational search for all the possible geometries of the target molecular system, we utilized the conformer-rotamer ensemble sampling tool (CREST) developed by Grimme and co-workers.[28] This code was built upon the previous

semiempirical tight-binding (TB) quantum chemistry method named as GFN-xTB.[29] As for the newly updated version of the CREST code[28], it is a very powerful tool for fast and reliable exploration and screening of the conformational space of mid- to large-sized molecules with up to about a thousand atoms. Moreover, a series of MD runs were automatically performed inside a CREST run to systematically explore the possible conformations for the molecule of interest. One of the built-in features is that it will delete all the redundant structures within a single CREST run. The other useful built-in feature is that the current code also offers a number of solvents available for users to consider the implicit solvation effect while doing the conformational search process. The current code has been successfully utilized by our group in a range of rotational and vibrational spectroscopic studies on its comprehensive conformational search of the potential candidates of the flexible, mid-size organic molecules and their associated clusters.[30] However, some unusual misses were also reported recently.[31]

To make the conformational search more complete, we typically built a few different starting geometries, and then carried out multiple CREST runs with each of the starting structures. Since the built-in function of the CREST code can only discard the redundant geometries within a single run, a further python program was used to remove the repeat structure between different CREST runs by calculating the root-mean-square deviations (RMSD) of different pairs of structures from all the searched outcomes. In general, if the RMSD value is zero for a given pair of calculated structure, it means they are identical. On the other hand, if the value becomes larger, it means more deviation is found between these two compared geometries. The resulting candidates were then optimized with a less expensive approach by using Molpro.[32] This loose optimization of the initial CREST candidates was done at the revPBE-D3/def2-SVP[33] level with the D3 empirical dispersion correction[34] and followed by a single point energy calculation at the Becke three-parameter Lee-Yang-Parr (B3LYP)[35] -D3/def2-TZVP level of theory. The filtering procedure (RMSD code) was revisited again to remove the remaining redundant structures.

In the next step, the conformers within a typical energy window of 10-15 kJ mol^{-1} relative to the global minimum were selected for the final geometry optimization and harmonic frequency calculations. Sometimes, we also increased the window to include higher energy conformers beyond 15 kJ mol^{-1} if some degree of kinetic trapping was suspected. All final geometric optimization and harmonic frequency calculations were performed at a higher level of level theory using Gaussian 16[36] with the built-in polarizable continuum model (PCM) for the implicit solvation effect.

The B3LYP hybrid functional was used to simulate all VOA spectra since it was reported to work well for hydrogen-bonded systems,[37] although some other functionals were also used. Generally, we used a large basis set such as cc-pVTZ, rather than a small basis set, such as 6-31G*.[38] Moreover, it is often beneficial to include the polarization and diffuse functions since additional functions could better describe the polarization of electron density in the system, providing a better prediction of the non-covalent interactions in the given system. In the current thesis, a number of basis sets were utilized, for instance, 6-311++G (2d, p), aug-cc-pVDZ, and def2-TZVP. Through a series of comparisons between experiment and theoretical simulations at different levels of theory, the B3LYP with D3 dispersion corrections and Becke-Johnson (BJ) damping[39], namely B3LYP-D3BJ, in combination with the def2-TZVP[33] basis set, generally gives the best agreement with experiment. Therefore, all the theoretical simulations in my thesis were performed at this level of theory.

To account for the specific chiral solute···solvent interactions, we applied the 'clusters-in-a-liquid' solvation approach.[40] To explore all possible binding sites at a chiral solute, we utilized the CREST code to search for possible chiral solute···$(solvent)_n$ complexes embedded in the implicit solvation continuum. Following the same procedure described above, all chiral solute-$(solvent)_n$ complexes were finally optimized with the PCM model to take into account of the implicit solvation effects.

References

[1] G. Holzwarth, E. C. Hsu, H. S. Mosher, T. R. Faulkner, A. Moscowitz, *J. Am. Chem. Soc.* **1974**, *96*, 251-252.

[2] H. Sugeta, C. Marcott, T. R. Faulkner, J. Overend, A. Moscowitz, *Chem. Phys. Lett.* **1976**, *40*, 397.

[3] L. A. Nafie, J. C. Cheng, P. J. Stephens, *J. Am. Chem. Soc.* **1975**, *97*, 3842-3843.

[4] L. A. Nafie, T. A. Keiderling, P. J. Stephens, *J. Am. Chem. Soc.* **1976**, *98*, 2715-2723.

[5] L. A. Nafie, M. Diem, D. W. Vidrine, *J. Am. Chem. Soc.* **1979**, *101*, 496-498.

[6] E. D. Lipp, C.G. Zimba, L. A. Nafie, *Chem. Phys. Lett.* **1982**, *90*, 1-5.

[7] P. Bouř, *ChemPhysChem* **2009**, *10*, 1983-1985.

[8] L. A. Nafie, *Nat Prod Commun* **2008**, *3*, 451-466.

[9] L. D. Barron, M. P. Bogaard, A. D. Buckingham, *J. Am. Chem. Soc.* **1973**, *95*, 603-605.

[10] W. Hug, S. Kint, G. F. Bailey, J. R. Scherer, *J. Am. Chem. Soc.* **1975**, *97*, 5589-5590.

[11] L. D. Barron, A. D. Buckingham, *Ann. Rev. Phys. Chem.* **1975**, *26*, 381-396.

[12] (a) L. Hecht, L. D. Barron, W. Hug, *Chem. Phys. Lett.* **1989**, *158*, 341-344; (b) L. D. Barron, L. Hecht, A. R. Gargaro, W. Hug, *J. Raman Spectrosc.* **1990**, *21*, 375-379.

[13] (a) L. D. Barron, A. R. Gargaro, Z. Q. Wen, *J. Chem. Soc., Chem. Commun.* **1990**, 1034-1036; (b) L. Hecht, L. D. Barron, A. R. Gargaro, Z. Q. Wen, W. Hug, *J. Raman Spectrosc.* **1992**, *23*, 401-411.

[14] W. Hug, *Appl. Spectrosc.* **2003**, *57*, 1-13.

[15] (a) C. Merten, Y. Xu, *Angew. Chem. Int. Ed.* **2013**, *52*, 2073-2076; *Angew. Chem.* **2013**, *125*, 2127-2130; (b) C. Merten, Y. Xu, *ChemPhysChem* **2013**, *14*, 213-219.

[16] K. R. J. Lovelock, *R. Soc. Open sci.* **2017**, *4*, 171223.

[17] G. Yang, Y. Xu, *Top Curr Chem* **2011**, *298*, 189-236.

[18] (a) N. V. Cohan, H. F. Hameka, *J. Am. Chem. Soc.* **1966**, *88*, 2136-2142; (b) T. R. Faulkner, C. Marcott, A. Moscowitz, J. Overend, *J. Am. Chem. Soc.* **1977**, *99*,8160-8168.

[19] (a) L. A. Nafie, T. B. Freedman, *J. Phys. Chem.* **1983**, *78*, 7108-7116; (b) R. Dutler, A. Rauk, *J. Am. Chem. Soc.* **1989**, *111*, 6957-6966; (c) D. Yang, A. Rauk *J. Chem. Phys.* **1992**, *97*, 6517-6534.

[20] (a) L. A. Nafie, T. H. Walnut *Chem. Phys. Lett.* **1977**, *49*, 441-446; (b) T. H. Walnut, L. A. Nafie, *J. Chem. Phys.* **1977**, *67*, 1501-1510.

[21] S. Abbate, L. Laux, J. Overend, A. Moscowitz, *J. Chem. Phys.* **1981**, *75*, 3161.

[22] G. Holzwarth, I. Chabay, *J. Chem. Phys.* **1972**, *57*, 1632.

[23] A. D. Buckingham, P. W. Fowler, P. A. Galwas, *Chem. Phys.* **1987**, *112*, 1-14.

[24] P. J. Stephens, *J. Phys. Chem.* **1985**, *89*, 748-752.

[25] W. B. Person, J. H. Newton, *J. Chem. Phys.* **1974**, *61*, 1040.

[26] (a) F. J. Delvin, P. J. Stephens, J. R. Cheesman, M. J. Frisch, *J. Am. Chem. Soc.* **1996**, *118*, 6327-6328; (b) P. Bouř, J. McCann, H. Wieser, *J. Chem. Phys.* **1998**, *108*, 8782-8789.

[27] L. D. Barron, *Molecular Light Scattering and Optical Activity*, **1982**, Cambridge University Press, Cambridge.

[28] P. Pracht, F. Bohle, S. Grimme, *Phys. Chem. Chem. Phys.* **2020**, *22*, 7169-7192.

[29] S. Grimme, C. Bannwarth, P. Shushkov, *J. Chem. Theory Comput.* **2017**, *13*, 1989-2009.

[30] (a) S. Oswald, N. A. Seifert, F. Bohle, M. Gawrilow, S. Grimme, W. Jäger, Y. Xu, M. A. Suhm, *Angew. Chem. Int. Ed.* **2019**, *58*, 5080-5084; *Angew. Chem.* **2019**, *131*, 5134-5138; (b) F. Xie, N. A. Seifert, W. Jäger, Y. Xu, *Angew. Chem. Int. Ed.* **2020**, *59*, 15703-15710; *Angew. Chem.* **2020**, *132*, 15833-15840; (c)Y. Yang, J. Cheramy, Y. Xu, *ChemPhysChem* **2021**, *22*, 1336-1343; (d) H. Wang, M. Heger, M. H. Al-Jabiri, Y. Xu, *Molecules* **2022**, *27*, 38.

[31] F. Xie, M. Fusè, A. S. Hazrah, W. Jäger, V. Barone, Y. Xu, *Angew. Chem. Int. Ed.* **2020**, *59*, 22427-22430; *Angew. Chem.* **2020**, *132*, 22613-22616.

[32] H. J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, *Wires Comput. Mol. Sci.* **2012**, *2*, 242-253.

[33] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297-3305.

[34] (a) S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104; (b) D. G. A. Smith, L. A. Burns, K. Patkowski, C. D. Sherrill, *J. Phys. Chem. Lett.* **2016**, *7*, 2197-2203.

[35] (a) A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098-3100; (b) C. Lee, W. Yang, R. Parr, *Phys. Rev. B* **1988**, *37*, 785-789.

[36] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, et al., *Gaussian 16*, Revision C.03; Gaussian, Inc.: Wallingford, CT, **2019**.

[37] (a) P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, *J. Phys. Chem.* **1994**, *98*, 11623-11627; (b) T. Kuppens, W. Langenaeker, J. P. Tollenaere, P. Bultinck, *J. Phys. Chem. A* **2003**, *107*, 542-553; (c) T. Kuppens, K. Vandyck, J. van der Eycken, W. Herrebout, B. J. van der Veken, P. Bultinck, *J. Org. Chem.* **2005**, *70*, 9103-9114; (d) T. Kuppens, W. Herrebout, B. van der Veken, P. Bultinck, *J. Phys. Chem. A* **2006**, *110*, 10191-10200; (e) L. Ducasse, F. Castet, A. Fritsch, I. Huc, T. Buffeteau, *J. Phys. Chem. A* **2007**, *111*, 5092-5098; (f) T. Brotin, D. Cavagnat, J. P. Dutasta, T. Buffeteau, *J. Am. Chem. Soc.* **2006**, *128*, 5533-5540.

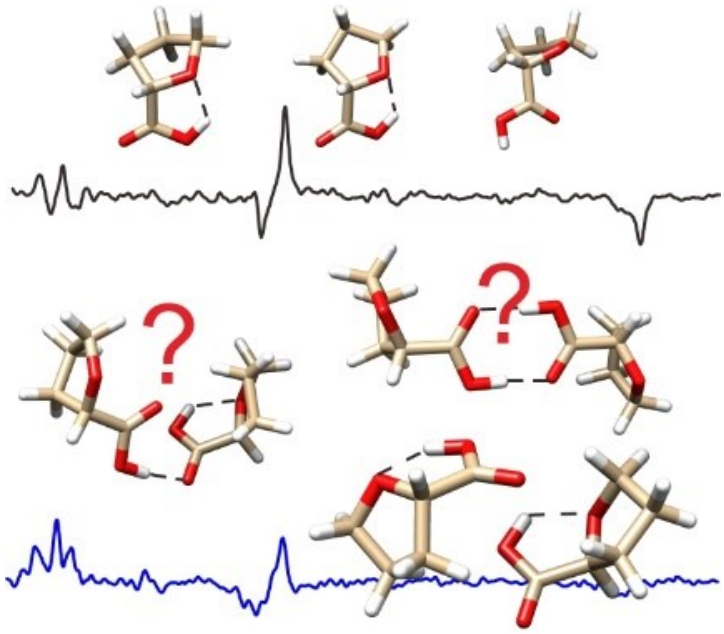
[38] T. Kuppens, K. Vandyck, J. van der Eycken, W. Herrebout, B. van der Veken, P. Bultinck, *Spectrochimica Acta Part A* **2007**, *67*, 402-411.

[39] A. D. Becke, E. R. Johnson, *J. Chem. Phys.* **2005**, *123*, 154101.

[40] A. S. Perera, J. Thomas, M. R. Poopari, Y. Xu, *Front. Chem.* **2016**, *4*, 1-17.

Chapter 3

Matrix Isolation-Vibrational Circular Dichroism Spectroscopic Study of Conformations and Non-Covalent Interactions of Tetrahydro-2-Furoic Acid



This chapter is directly copied and adapted the new format from the following publications:

Y. Yang, J. Cheramy, Y. Xu, *ChemPhysChem* **2021**, *22*, 1336-1343.

3.1 Introduction

Vibrational circular dichroism (VCD) spectroscopy is commonly used in the determination of absolute configurations of chiral molecules in solution.[1] Because of its unique sensitivity to conformations, in addition to chirality, it has become a powerful spectroscopic tool for investigating solute-solvent[2,3] interactions and solute-solute[4] interactions such as self-aggregation of chiral molecules. In some cases, the molecules of interest are quite flexible and have a large number of possible conformations. At the same time, they may also interact strongly with each other and/or with solvent molecules. All these factors tend to generate broad infrared (IR) bands and barely detectable VCD features because of the great degree of averaging.[2]

The combination of matrix isolation (MI) and VCD spectroscopy has been shown to be a powerful technique to overcome some of the above challenges.[5] With MI-VCD, one can study the molecular target by trapping it in a cold noble gas matrix (typically ~10 to 30 K). Since the solvent is now replaced by a cold inert gas matrix, one can avoid the solute-solvent interactions and have the molecular target largely in its isolated monomeric form, if desired. Often one can obtain MI-IR and MI-VCD spectra with a much narrower bandwidth than in solution and achieve detailed assignment and extract structural information and conformational distributions confidently. One may also explore VCD features of chirality transfer in non-covalently bonded species and of photoisomerization products.[6]

We report a combined experimental MI-IR, MI-VCD, and theoretical study of tetrahydro-2-furoic acid (THFA) and its aggregates. THFA is an important intermediate in the production of pharmaceuticals such as Alfuzosin, Faropenem and Tecadenoson.[7] The THFA monomer exhibits an interesting conformational landscape with a range of possible conformers. Recent rotational spectroscopic studies[8] indicate that THFA favors a *trans*-COOH configuration, which shows an intramolecular hydrogen (H)-bond between carboxylic OH and the ester oxygen, rather than the more common *cis*-COOH

configuration favored by most carboxylic acids.[9] In the previous VCD study of THFA in $CDCl_3$ and CS_2 solution, THFA molecules, like many other carboxylic acids, were reported to form a strong double H-bonded ring with two *cis*-COOH THFA subunits in solution.[10] Indeed, in many of the previous MI-IR studies of carboxylic acid dimers, for example formic acid dimer,[11] acetic acid dimer[12] and more recently glycolic acid dimer,[13] the binary structures consisting of two *cis*-COOH subunits forming the double H-bonded ring are by far the dominant binary species detected experimentally.

On the other hand, changes in experimental conditions or sometimes detection techniques may lead to discovery of new conformations and/or aggregation behaviors. For example, formic acid is known to mainly exist as *trans*-COOH at room temperature.[14] Using a long absorption path and also due to the much large dipole moment of the *cis*-COOH make it possible for observing the *cis*-isomer in the gas phase[15] or in cold rare gas matrices.[16] Very recently, new 2D-IR experiments demonstrate that formic acid exist as two distinct and long-living conformational isomers in aqueous solution at room temperature where the *trans*-COOH conformation is stabilized by hydration.[17] It would be of significant interest to examine the conformations of THFA and its aggregation behavior in rare-gas matrices under cryogenic conditions, and compare to the solution data. By using different matrix deposition temperatures and sample conditions, we aim to extract the possible conformations of the THFA molecule by itself and to follow the self-aggregation of THFA in a matrix environment. Interestingly, the dominant dimeric structures identified are very different from those in solution[10] and have not been observed experimentally before for THFA. The current results are also compared to a previous MI-VCD study of lactic acid[18] to highlight the diverse aggregation behavior of lactic acid and THFA in matrices under cryogenic conditions.

3.2 Results and Discussion

3.2.1 Overview of the Experimental MI-IR and MI-VCD Spectra

THFA, like many acids, is highly hydroscopic, making it quite challenging to obtain good quality MI-IR and especially MI-VCD spectra. A routine was developed after much trial and error and is described in the experimental section at the end. To maintain a nearly transparent deposited film on the cold window, a crucial requirement for a good quality VCD spectrum, we found it was best to keep THFA at room temperature and flow argon (Ar) slowly at 1 sccm (standard cubic centimeters per minute). Through these steps, we were able to remove the vast majority of water in the system without losing much THFA sample, and obtained good quality MI-IR and MI-VCD spectra. The experimental MI-IR and MI-VCD spectra at a deposition temperature of 10 K, 24 K and 30 K are presented in Figure 3.1. Please note that we use S-THFA throughout this paper for the presentations of the calculations and the experimental results.

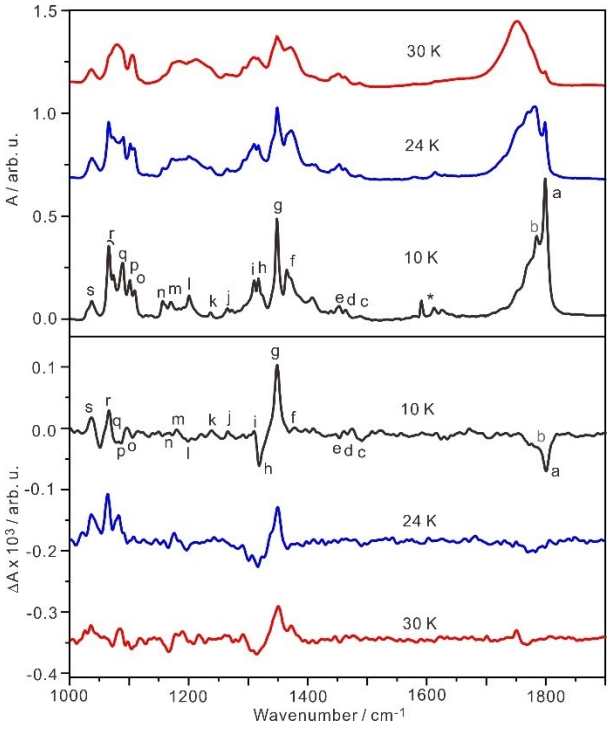


Figure 3.1. Experimental MI-IR and MI-VCD spectra of THFA recorded at the three different deposition temperatures. The main features belonging to the THFA monomer are labelled as "a-s" based on the assignment presented in the next section. "b" is likely due to larger aggregates (vide infra).

Interestingly, the IR spectral features in the three regions: 1000-1120, 1150-1250, and 1275-1400 cm^{-1}, remain very similar except some broadening/smoothing, while those in the 1700-1820 cm^{-1} region show some obvious intensity increases for bands to the lower frequency side as the deposition temperature increases. While it was previously reported that THFA exists mainly in the dimeric form in solution,[10] it is not clear whether there is already substantial amount of $(THFA)_2$ present in the gas phase before the deposition onto a cold cell window. This is also a topic of interest in spectroscopic studies of H-bonded species in a supersonic jet expansion.[19] We will examine this in the THFA case below.

3.2.2 Assignment of the MI-IR and MI-VCD Spectra at 10 K

Since some very sharp IR and VCD band features were detected at 10 K, similar to the case of lactic acid,[18] we first examined the contribution from the THFA monomer. We explored the possible conformational structures using the CREST code[20,21] under the PCM of Ar. The CREST details are described in the experimental/theoretical section. This conformational code has been shown to produce successfully a comprehensive set of structural candidates of organic molecules and their non-covalently bonded clusters,[22] although a deficiency was reported recently.[23]

For the THFA monomer, eight geometries were identified, similar to those reported previously[8,10] and were reoptimized at the B3LYP-D3BJ/def2-TZVP and B3LYP-D3BJ/6-311++G (2d,p) levels of theory without and with the inclusion of PCM of Ar. These eight geometries are named M1 to M8 based on their relative free energies, with M1 being the global minimum. While M1 and M2 adopt the *trans*-COOH configuration, M3 to M6 take on the *cis*-COOH conformation. M7 and M8 take on again the *trans*-COOH conformation but without the intramolecular H-bond, in contrast to M1 and M2. All eight optimized geometries are shown in Figure A1, Appendix A, while the three most stable monomeric conformers M1-M3 are also shown in Figure 3.2, as the representative *trans*- and *cis*-

COOH structures. The relative free energies and also the Boltzmann percentage factors at 298 K of M1-M8 at the B3LYP-D3BJ/def2-TZVP and B3LYP-D3BJ/6-311++G (2d,p) levels without and with the PCM of Ar are summarized in Table 3.1. In general, the differences in free energies and Boltzmann factors between the two different basis sets are small, with the stability ordering of M4 and M5 reversing, although they remain close in energy. On the other hand, it is interesting to note that the inclusion of the PCM of Ar increases the Boltzmann percentage of M1 by ~3 to 4% with both basis sets.

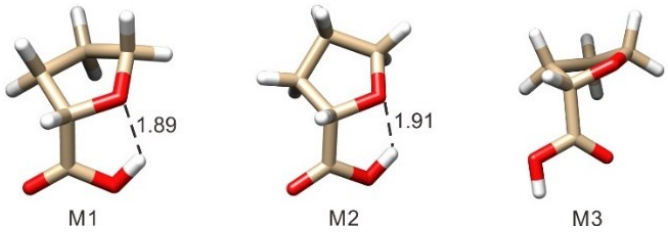


Figure 3.2. Optimized geometries of the two most stable *trans*- (M1 and M2) and one most stable *cis*-COOH (M3) conformers. M1 and M2 differ in their ring configurations and the related intramolecular H-bond lengths (in Å) are also indicated.

Table 3.1. The relative free energies, ΔG in kJ mol^{-1}, and Boltzmann percentage factors, B_f in %, of the eight THFA conformers at 298 K at the B3LYP-D3BJ/def2-TZVP, B3LYP-D3BJ/6-311++G(2d,p), B3LYP-D3BJ/def2-TZVP/ PCM(Ar), and B3LYP-D3BJ/6-311++G (2d,p)/PCM(Ar) levels of theory.

Conf.	def2-TZVP		6-311++G(2d,p)		def2-TZVP/Ar		6-311++G(2d,p)/Ar	
	ΔG	B_f	ΔG	B_f	ΔG	B_f	ΔG	B_f
M1	0.0	46.1	0.0	47.9	0.0	49.6	0.0	51.1
M2	0.8	33.0	0.9	33.1	0.9	34.5	0.9	34.8
M3	4.4	8.0	4.8	7.0	5.0	6.5	5.7	5.2
M4	5.6	4.9	6.0	4.2	6.7	3.3	7.0	3.0
M5	5.7	4.6	5.8	4.6	6.7	3.3	6.8	3.3
M6	6.5	3.4	6.7	3.2	7.3	2.6	7.4	2.6
M7	25.2	0.0	26.0	0.0	24.0	0.0	25.2	0.0

M8	26.6	0.0	27.2	0.0	25.8	0.0	26.3	0.0

The single monomer IR and VCD spectra of the six lowest energy THFA conformers with non-zero population contribution at room temperature were calculated. After that, the Boltzmann averaged monomer IR and VCD spectra at the four different levels of theory were calculated and are compared with the experimental MI-IR and MI-VCD spectra at 10 K in Figure 3.3. At a first glance, the averaged IR and VCD simulations of the monomeric conformers at all four levels of theory capture the main experimental features observed. This indicates that the THFA monomer is indeed the dominant species at 10 K. For example, the distinctive spectral features marked as "a", "f", "g", "h", "i" "I", "r" and "s" in the experimental IR and VCD spectra can be confidently associated with the calculated features. A noticeable difference among these four levels of theory is the relative IR intensity pattern of the two strong bands "f" and "g", with the B3LYP-D3BJ/def2-TZVP/PCM(Ar) providing the best agreement to the experiment. Many gas phase conformational studies, based on rotational spectroscopy, which is highly sensitive to subtle conformational changes, show that the B3LYP-D3BJ/def2-TZVP level of theory typically outperforms the Pople basis sets in reproducing those small structural differences.[8] For the above reasons and conciseness, we will use the simulated results at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level in our comparison with the experiment in the subsequent analyses of the THFA aggregates.

At the source temperature, M1 and M2 make up about 50% and 35% of the total monomer population, respectively. Not surprisingly, their IR and VCD features dominate the experimental spectra. The individual monomer IR and VCD spectra of M1-M6 at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level are depicted in Figure A2, Appendix A. As one can see, M1 and M2 exhibit very similar IR and VCD bands. The most noticeable evidence of M2 is the features labelled as "m" and "l", produced mainly by M2 which show a positive and a negative VCD feature, respectively, whereas M1 has two negative bands

around 1200 cm^{-1}. Another indication of the presence of M2 is the narrowly split IR feature, "r", generated by the slightly blue-shifted M2 band relative to the M1 band in this region, which is detected experimentally. These observed features confirm that M2 is slightly less abundant than M1, in good agreement with the Boltzmann factors predicted. The contribution of M3-M6, with abundances of 5.2% to 2.6%, is less obvious in the experimental data. Based on the theoretical, individual IR and VCD spectral features, the evidence of these higher energy conformers is mainly in the 1000-1120 cm^{-1} region (vide infra).

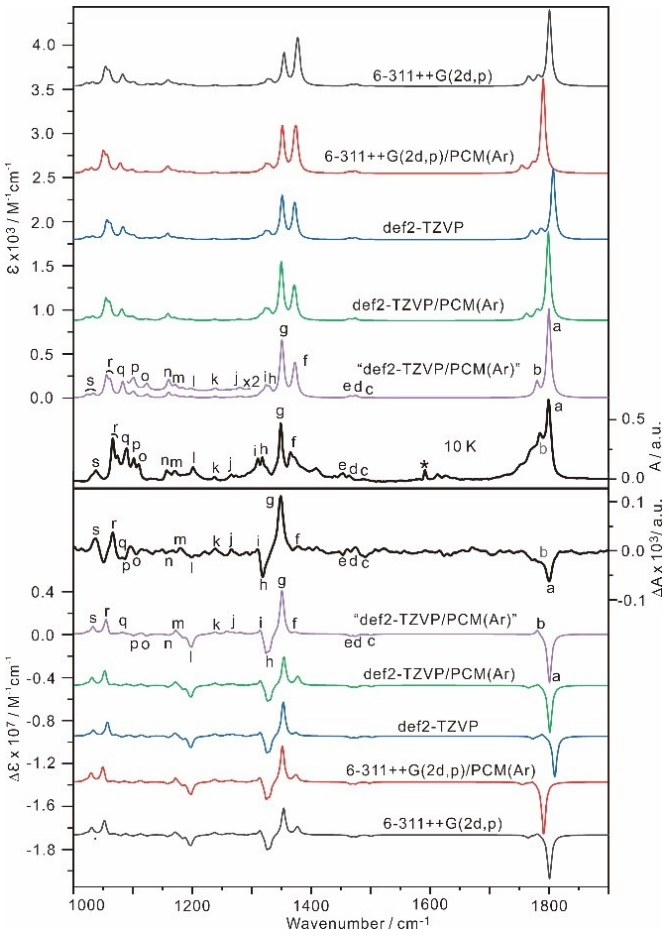


Figure 3.3. Comparisons of the experimental MI-IR and MI-VCD spectra of THFA at 10 K with the Boltzmann averaged IR and VCD spectra of the monomeric conformers at the four different levels of theory. The experimental units are on the right side. The extra weak bands in the 1600 cm^{-1} region are due to the

residual water in the system. The main band features are labelled “a-s” and gray colored “b” letter indicates that the experimental features are contaminated with substantial contribution from larger THFA aggregates. “def2-TZVP/PCM(Ar)” indicates a new Boltzmann averaged IR and VCD spectra with consideration of possible conformational relaxation in a cold Ar matrix. See the main text for the related discussions.

It is interesting to note that in the previous rotational spectroscopic study of THFA,[8] only M1, M2 and M3 were detected experimentally in a supersonic jet expansion, even though the instrument sensitivity was more than enough to observe the higher energy THFA conformers, M4-M6, based on their room temperature abundances. This observation could be understood by considering that the *cis*-COOH M4-M6 conformers differ only slightly from M3 in their ring configurations and can relax to M3 easily in a jet expansion. On the other hand, the relaxation barriers from the *cis*-COOH M3-M6 conformers to the *trans*-COOH M1 and M2 conformers are very high and are not expected to happen in a jet expansion or a cold matrix. We hypothesize that M4-M6 would relax to M3 in a cold Ar matrix, giving it a percentage abundance of ~15%. The new Boltzmann weighted IR and VCD spectra with M1:M2:M3=0.50:0.35:0.15 are also provided in Figure 3.3. The main differences between the new and the previous Boltzmann weighted IR and VCD spectra are in the 1000-1120 cm^{-1} region, while the spectral features in other frequency regions remain essentially the same. In this region, the new Boltzmann weighted IR and VCD spectra provide a better agreement with the experiment, allowing clearer match of the ‘o’ to ‘r’ bands, supporting the proposal of relaxation (or at least partial relaxation) of M4-M6 to M3 in the matrix. Such relaxation is expected to be even more complete at higher deposition temperatures. We therefore will use the new Boltzmann weighted monomer simulation in the discussions later on.

All main features observed at 10 K can be assigned to M1-M3, although contributions from larger aggregates are visible. For example, while the experimental IR carbonyl stretching band, “b”, is in a similar frequency region to the related band of M3, its

experimental intensity is way too high for M3. Further analyses (vide infra) demonstrate that the red shifted experimental features in this region are mainly due to larger aggregates, whereas the corresponding feature of M3 is buried underneath.

3.2.3 Unusual Dimeric Structures Revealed in the MI-IR and MI-VCD Spectra at Higher Deposition Temperatures

At higher deposition temperatures of 24 and 30 K, one can clearly see the emergence of the new IR and VCD features, such as the red-shifted C=O stretching bands at ~1800 cm^{-1} and broadening of many bands, suggesting formation of THFA dimers or even larger aggregates. By far, the most stable dimers are those with the classic double H-bonded ring of carboxylic acid dimers which consist of two *cis*-COOH THFA conformers, i.e., from M3-M6, leading to ten **Type 1** $(THFA)_2$ conformers. Three representative geometries of **Type 1** $(THFA)_2$, namely D33, D34, and D56, are shown in Figure 3.4, while all of their optimized geometries are provided in Figure A3, Appendix A. These dimers identified are similar to those reported previously although calculated at different levels of theory and under different conditions (in the gas phase or in solution) compared to the current case.[8b,10] Their relative free energies and Boltzmann population factors at 24 K, 30 K and 298 K of all ten **Type 1** binary conformers at the B3LYP-D3BJ/def2-TZVP without and with PCM of Ar are listed in Table 3.2. The inclusion of PCM(Ar) changed the ΔG values from the gas phase values, as observed in the case of the monomer. We note that the relative free energy ordering is modified noticeably among D34 to D56, although the most stable D33 and the least stable D66 remain in the same place.

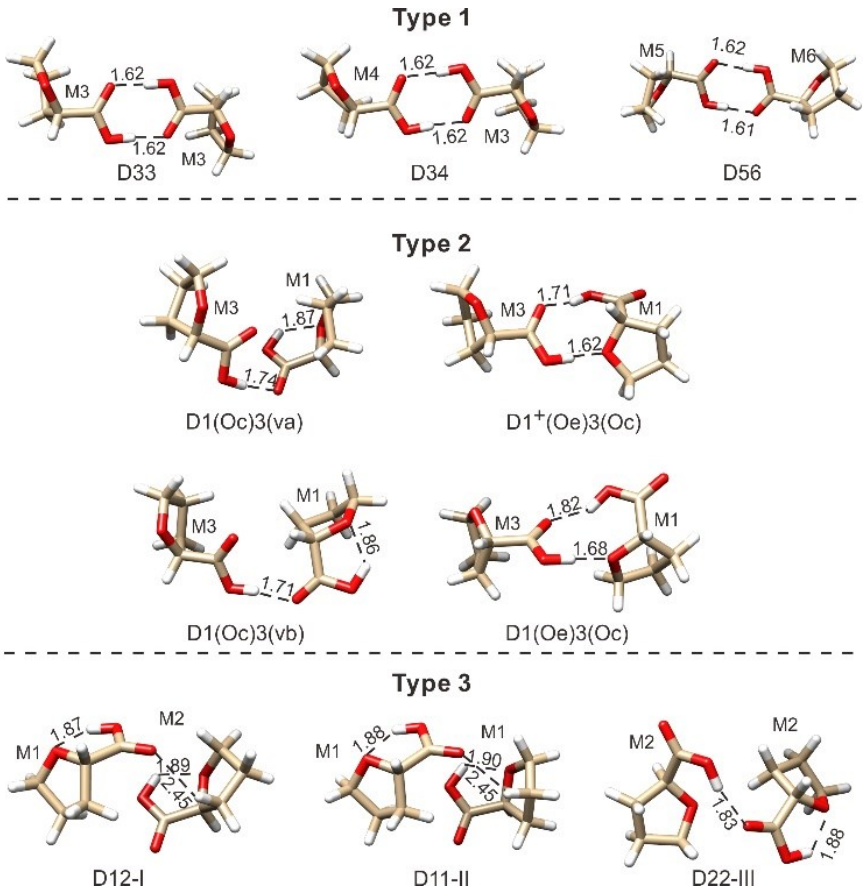


Figure 3.4. Representative conformational geometries of the three types of binary THFA aggregates. See the main text for the description of names.

Table 3.2. The relative free energies, ΔG in kJ mol^{-1}, and Boltzmann percentage factors, B_f in %, of the ten conformers of **Type 1** $(THFA)_2$ at 24 K and 298 K at the two levels of theory.[a]

$(THFA)_2$	B3LYP-D3BJ/def2-TZVP		B3LYP-D3BJ/def2-TZVP/Ar	
	ΔG (24K, 298K)	B_f (24K, 298K)	ΔG (24K, 298K)	B_f (24K, 298K)
D33	0.0, 0.0	98.4, 21.2	0.0, 0.0	96.9, 14.7
D34	1.1, 1.1	0.4, 13.6	0.8, 0.6	1.5 11.5
D35	0.9, 1.3	1.2, 12.6	0.8, 1.1	1.4, 9.4
D36	1.5, 1.8	0.0, 10.3	1.3, 0.7	0.2, 11.1
D44	2.3, 1.9	0.0, 9.8	1.7, 0.5	0.0, 12.0
D45	2.1, 2.2	0.0, 8.7	1.7, 1.1	0.0, 9.4
D55	2.9, 2.4	0.0, 8.0	2.7, 1.6	0.0, 7.7

D46	2.7, 2.5	0.0, 7.7	2.2, 1.0	0.0, 9.8
D56	3.6, 3.4	0.0, 5.4	3.1, 0.7	0.0, 11.1
D66	4.4, 5.1	0.0, 2.7	3.8, 3.8	0.0, 3.2

[a] The ΔG and B_f values at 30 K differ very slightly from those at 24 K and are not listed for conciseness.

The individual conformer IR and VCD spectra of the **Type 1** conformers and the related Boltzmann averaged spectra at 24K and 298 K are provided in Figure A4, Appendix A. Interestingly, the new IR and especially VCD features of the **Type 1** binary species predicted show very little agreement with the new experimental features appearing at 24 K. For example, the noticeably red-shifted IR C=O bands in the above 1650 cm^{-1} region and their related strong "W" shaped VCD feature (at 298K) appear to be missing in the corresponding experimental spectrum. On the other hand, these distinctive IR and VCD features at 298 K were observed experimentally in a $CDCl_3$ solution of THFA,[10] suggesting that the cold rare gas matrix is a much different environment compared to a solution. Further examination of the individual conformer IR and VCD spectra of **Type 1** conformers indicates that modifying the Boltzmann factors would not improve the agreement with the experiment either. For example, if we assume a Boltzmann temperature of 24 K for the **Type 1** species because the deposition temperature was at 24 K, the resulting IR and VCD spectra are essentially those of D33. Again, they show little resemblance to the new spectral features observed experimentally from 10 K to 24 K. Overall, it is clear that the **Type 1** conformers are not the dominant binary species under the present experimental condition, in contrast to the solution case[10] and also to the situation reported previously for lactic acid in Ar matrices.[18]

While the THFA monomer exists mainly in the *trans*-COOH conformations in a cold Ar matrix as discussed above, the double H-bonded topology utilized in the **Type 1** binary conformers, on the other hand, requires both subunits to be in the *cis*-COOH configuration. In a cold rare gas matrix, one may therefore expect a kinetically controlled formation process which would play an important role in the dimer formation processes. Two other

types of carboxylic acid binding motifs,[24] **Type 2** and **Type 3** structures which contain one *trans*- and one *cis*-CCOH subunits or two *trans*-COOH subunits, respectively, were generated in the systematic CREST conformational search with the inclusion of the PCM of Ar.

32 **Type 2** geometries with the OH groups pointing to either O atoms in the COOH group, or O of the ring were identified and further DFT optimization led to 30 **Type 2** conformers within an energy window of ~10 kJ mol^{-1}. The relative free energies and Boltzmann population factors at 24 K, 30 K, and 298 K of all 30 **Type 2** binary conformers at the two levels of theory with the PCM of Ar are listed in Table A1, Appendix A. As discussed in the previous section, M1-M3 are the main monomeric conformers whereas M4-M6 are essentially all converted to M3. We therefore focus only on the **Type 2** structures which are made of M1-M3. Seven most stable **Type 2** conformers made of M1 or M2 with M3 are summarized in Table 3.3. In additional, four example geometries of these **Type 2** geometries are depicted in Figure 3.4, while all seven of them are given in Figure A5, Appendix A, for completeness. The individual conformer IR and VCD spectra of the **Type 2** conformers are shown in Figure A6, Appendix A. We give each **Type 2** conformer a specific name to better differentiate them. For example, D1(Oe)3(Oc) indicates that this dimer is made of M1 and M3 where "Oe" or "Oc" represent the ether or carbonyl oxygen atom as the intermolecular H-bond acceptor, respectively. In D1(Oc)3(vb) and D1(Oc)3(va), the "va" and "vb" indicate two different kinds of non-covalent interactions between Oc and the tetrahydrofuran ring H atom (Figure 3.4). In some cases, an additional "+" is used, for example in 1+, indicating that the OH group is twisted further away from the OH direction in M1. It is interesting to note that all these structures exhibit one COOH···Oc intermolecular H-bond and some of them take on an additional intermolecular COOH···Oe H-bond. Overall, the **Type 2** binary conformers are at least 15 kJ mol^{-1} less stable than the **Type 1** structures.

Table 3.3. The relative free energies at 24K and 298 K, ΔG in kJ mol^{-1}, and Boltzmann percentage factors, B_f in %, of the **Type 2** and **Type 3** $(THFA)_2$ conformers at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.[a]

Type 2 conformers			**Type 3** conformers		
	ΔG (24, 298K)	B_f (24, 298K)		ΔG (24, 298K)	B_f (24, 298K)
$D1^+$(Oe)3(Oc)	0.0[b], 1.2	99.9, 21.9	D12-I	0.0[d], 0.0[e]	100.0, 23.0
D1(Oe)3(Oc)	1.4, 0.0[c]	0.1, 35.2	D11-II	1.7, 0.1	0.0, 21.7
$D2^+$(Oe)3(Oc)	2.0, 3.2	0.0, 9.7	D11-VI	2.5, 2.1	0.0, 9.8
D1(Oc)3(vb)	2.7, 2.6	0.0, 12.5	D22-III	4.6, 1.0	0.0, 15.1
D1(Oc)3(va)	3.1, 3.8	0.0, 7.6	D12-IV	5.7, 1.7	0.0, 11.5
D2(Oc)3(vb)	3.2, 3.1	0.0, 9.9	D11-V	6.9, 1.9	0.0, 10.6
D2(Oc)3(va)	5.2, 5.9	0.0, 3.2	D12-VIII	8.3, 4.7	0.0, 3.4
			D12-VII	8.6, 4.4	0.0, 3.9
			D12-IX	11.8, 7.7	0.0, 1.0

[a] The ΔG and B_f values at 30 K differ very slightly from the values at 24 K and are not listed for conciseness.

[b] $D1^+$(Oe)3(Oc) is the most stable **Type 2** conformer at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level under 24 K. Its relative free energy is 16.4 kJ mol^{-1} relative to the global minimum D33 at the same level.

[c] D1(Oe)3(Oc) is the most stable **Type 2** conformer at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level under 298 K. Its relative free energy is 15.1 kJ mol^{-1} relative to the global minimum D33 at the same level.

[d] D12-I is the most stable **Type 3** conformer at the B3LYP-D3BJ/def2-TZVP /PCM(Ar) level under 24 K. Its relative free energy is 22.6 kJ mol^{-1} relative to the global minimum D33 at the same level.

[e] D12-I is the most stable **Type 3** conformer at the B3LYP-D3BJ/def2-TZVP /PCM(Ar) level under 298 K. Its relative free energy is 23.7 kJ mol^{-1} relative to the global minimum D33 at the same level.

In general, the **Type 3** structures are formed by weaker non-covalent interactions between two *trans*-COOH THFA subunits. About 200 potential **Type 3** geometries were generated by CREST initially. Again, we focus only on those made of M1 and M2 in the

Type 3 structures. Subsequent DFT optimizations led to nine **Type 3** binary conformers within an energy window of ~8 kJ mol^{-1}. While many much higher energy conformers were also calculated, they were not included for further consideration because of their very limited contribution. Geometries of three example **Type 3** conformers are also shown in Figure 3.4, while all nine of them are depicted in Figure A7, Appendix A. We name the **Type 3** conformers based on their two subunits in Arabic numbers and their relative stability at 298 K in Roman numerals. For example, D12-I means that it is a **Type 3** conformer consisting of M1 and M2 and is the most stable **Type 3** conformer. The relative free energies and Boltzmann factors of these nine conformers are summarized in Table 3.3, while their individual theoretical IR and VCD spectra are depicted in Fig A8, Appendix A. The **Type 3** binary conformers are more than 22 kJ mol^{-1} less stable than the **Type 1** structures.

As discussed before, we focus on those binary conformers made of M1, M2, and M3 which are visibly present at 10 K spectra. The occurrence of M4-M6 is likely even rarer at 24 K or 30 K since the barriers for the relaxation of M4-M6 to M3 were reported to be low previously.[8] If the $(THFA)_2$ conformers were mainly generated in the gas phase prior to the deposition, the **Type 1** conformers would have dominated the experimental spectra since these conformers are strongly favored over the **Type 2** and **Type 3** conformers. This is, however, not the case, indicating that the binary species are largely produced in the cold rare gas matrices, rather than in the gas phase. On the other hand, if one assumes a purely kinetic controlled dimer formation process, one can estimate the kinetic abundances of the binary species. For example, the relative abundance of the **Type 2** conformers made on M1 and M3, which we denote as D(13) versus that of D(23) is 59%:41%, i.e. the ratio of M1:M2. Based on the kinetic model, one would expect the most contribution from **Type 3** which consists of D(11), D(12) and D(22), then **Type 2** which include D(13) and D(23) and very little **Type 1** which includes D(33).

We considered both kinetic and thermodynamic influences on the dimer formation in a

cold rare gas matrix to estimate the contribution of each binary conformers. For example, for **Type 2** D(13) and (23) conformers, we first applied Boltzmann factors at the deposition temperature to obtain the Boltzmann averaged D(13) IR and VCD spectra and do the same for D(23) conformers. Then we applied to the kinetic ratio of D(13):D(23) to the final simulated IR and VCD spectra of **Type 2** conformers. The same procedure was applied to the **Type 1** and **Type 3** conformers. In Figure 3.5, the simulated IR and VCD spectra of the THFA monomer, **Type 1, 2** and **3** dimers are shown and are compared to the experimental data at 24 K and at 10 K. While doing such comparison, we recognized that the contribution of D(22) should be more than that predicted kinetically, and empirically adjusted the weights among D(11):D(22):D(12) as 0.4:0.4:0.2 for **Type 3**. We also slightly adjusted the ratios of **Type 1**:**Type 2**:**Type 3** to 0.1:0.1:0.8, largely following the kinetic prediction of 0.03:0.26:0.71 while also taking into account the strong thermodynamical preference for **Type 1**. Finally, the monomer versus dimer ratio of 0.5:0.5 was used to simulate the final IR and VCD spectra for comparison with the experimental data at 24 K.

From Figure 3.5, one can see that the main experimental IR and VCD feature changes from 10 K to 24 K are largely captured by the simulations based on the monomeric and dimeric THFA conformers. In the 24 K IR spectrum, a few noticeable features, such as the noticeable decrease in the intensity ratio of the sharp monomer band in the C=O stretch region versus the broad dimer band, and the further red-shifted maximum, are re-produced. In the VCD spectrum, the red-shifted, negative C=O band is reproduced. The detailed VCD features in the carbonyl region are notoriously difficult to predict because these close features tend to be multi-signated and minor change in their relative frequencies can alter the appearance of the VCD features drastically. The three negative VCD features in the 1300-1350 cm^{-1} region are well simulated. It is also clear from both IR and VCD simulations that **Type 1** and **Type 2** cannot reproduce the important changes discussed above. It is interesting to point out that the **Type 3** conformers, the significantly less stable binary species among the three types, are responsible for new features associated with

$(THFA)_2$.

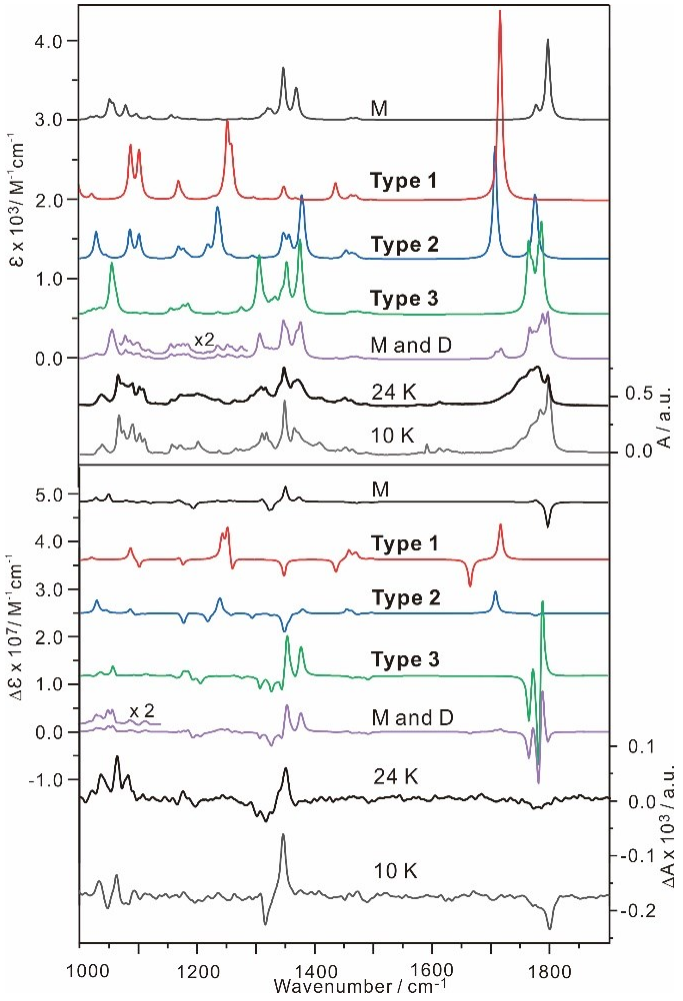


Figure 3.5. Comparisons of experimental MI-IR and MI-VCD spectra of THFA at 24 K and 10 K with the related simulated IR and VCD spectra of the averaged monomeric (M) and the **Type 1**, **2** and **3** $(THFA)_2$ species, as well as the final empirically weighted monomer (M) and dimer (D) spectra. See the detailed discussion in the main text. The experimental units are on the right side.

At 30 K deposition temperature, further line broadening and red-shift of the band maximum of the acid carbonyl band are observed. The carbonyl band features show a further reduction of the monomer contribution. These experimental changes can be largely accounted for by decreasing the monomer contribution and changing the relative

contribution of individual conformers to the final **Type 2** and **Type 3** contributions. In Figure A9, Appendix A, we compare the experimental data at 30 K with the simulated spectra discussed above with the weighting factors. Another interesting observation is that the IR spectrum of most bands appear smoother at 30 K than at 24 K but without any noticeable frequency shift. We speculate such appearance maybe be also due to some variation of the matrix environment at a higher temperature, experienced by each of these dimers and monomers.[25]

The observation of the dominant **Type 3** structures also indicates that there is little $(THFA)_2$ present in the gas phase before the deposition onto a cold cell window. If there were significant dimer formation in the gas phase already, one would expect mainly **Type 1** species, following the thermodynamic mechanism. The current study supports the conclusion reported for THFA in a previous rotational spectroscopic study.[8b]

It is interesting to comment on the very different aggregation behavior of lactic acid which was also studied under cryogenic matrix conditions using MI-IR and VCD spectroscopy.[18] In the lactic acid case, the *cis*-COOH configuration is strongly favored in its monomeric form. Consequently, the subsequent aggregation in cold matrices leads to the usual double H-bonded cyclic ring dimers which are favored by both the thermodynamically and kinetically controlled mechanisms. Therefore, one is able to use the MI technique to mimic/follow the aggregation behavior of lactic acid in liquid and in solution, one carboxylic acid molecule at a time. In the THFA case, the strong preference of *trans*-COOH in its monomeric form, coupled with the kinetically controlled dimer formation mechanism, leads to the very uncommon binary structures, drastically different from those reported in solution.[10] The THFA IR features initially led us to believe that the usual binary structures may be responsible and further simulations of larger aggregates based on **Type 1** binary structures were also carried out in an attempt to explain some discrepancy between the experiment and theory. On the other hand, the MI-VCD features

are pivotal in revealing these unusual structures and in guiding our search for these structures to properly account for the experimental observation.

3.3 Conclusion

In this study, we report the MI-IR and MI-VCD spectra of THFA at 10 K, 24 K and 30 K. The sharp IR and VCD features at 10 K allow clear identifications of the three main THFA monomeric conformations, M1-M3. These monomeric structures take on dominantly *trans*-COOH configuration and a small percentage of *cis*-COOH. The conformational relaxation within the structures of the *trans*- and *cis*-THFA groups in the cold rare gas matrices was discussed and a good agreement between the experimental and theoretical IR and VCD spectra is achieved. At higher deposition temperatures, noticeable THFA aggregation occur. Systematic conformational searches of possible binary aggregates were carried out using CREST, a new and powerful quantum chemistry code to identify all possible dimeric structures. The binary structures are separated into **Type 1**, **2**, **3** groups with energy gaps of ~15 to 22 kJ mol^{-1} between the most stable structures in **Type 1** and **2** and in **Type 1** and **3**, respectively. Interestingly, neither the **Type 1** structures, which feature the double H-bonded cyclic ring and are the dominant species in chloroform solution,[10] nor the **Type 2** geometries, which are the main species in a jet expansion,[8b] make much contribution under the current experimental condition. Rather, a new type of dimeric structures, **Type 3** conformers, which are the least thermodynamically stable among the three types, account for the majority of dimeric species in the observed IR and VCD spectra. The cold rare gas matrices provide a unique environment to forge the formation of these **Type 3** binary aggregates that are strongly favored by the proposed kinetically controlled formation mechanism. The results highlight the power of using MI-VCD spectroscopy in revealing uncharacteristic structures and in characterizing the aggregation behaviors of chiral molecules under rare gas cold matrix.

3.4 Experimental Section

3.4.1 Experimental

The matrix isolation systems consisted of a closed-cycle helium cryostat from Advanced Research Systems, Inc. (ARS 4HW compressor with a DE 204SI expander) and a stainless-steel vacuum line. The system is similar to the one described before,[5b] except the turbo pump was replaced with a diffusion pump equipped with a liquid-nitrogen cooling trap.

Both R-(+)-THFA (99%) and S-(-)-THFA (97%) were purchased from Sigma Aldrich and used without further purification. Both enantiomeric samples appear as oil at room temperature. The residue water in the sample system was removed by heating the drying tubes with molecular sieves (size: 4 Å) in the sample line, and by pumping the whole sample system for a day or two. The THFA liquid was then loaded into a sample reservoir situated right before a 1/16" flowing tube for deposition at room temperature and the whole system was pumped again overnight. Before deposition, Ar gas was passed through the sample to the cold head at room temperature while pumping the system for an hour or so. The flow rate of the Ar gas was held constant at 1 sccm by using a flow controller (MKS 1179 A). The Ar backing pressure was 8 psi. The depositions were done under three different temperatures (10 K, 24 K and 30 K) with a deposition time of 40 minutes to about an hour.

The experimental IR and VCD spectra were measured using a Bruker Vertex 70 supplemented with a PMA 50 module. All MI-IR and MI-VCD spectra were recorded with a spectral resolution of 2 cm^{-1}. The MI-VCD were collected with a total acquisition time of 12 h, or about 30000 scans. The final MI-VCD were generated by using the standard (S-R)/2 procedure. All the matrix IR spectra were background subtracted and the raw IR and VCD data for each enantiomer of THFA are provided in Figure A10, Appendix A. Although the CH-stretching region was shown before to offer some interesting structural information,[26] no MI-VCD was measured in the current study due to experimental

difficulty.

3.4.2 Computational

The Gaussian 16 program package[27] was used to perform all geometry optimizations, harmonic frequencies calculations and IR and VCD intensity calculations. All DFT calculations were undertaken using the B3LYP functional[28] with D3 dispersion corrections[29] and Becke-Johnson (BJ) damping[30], i.e. B3LYP-D3BJ. The basis set 6-311++G(2d,p) and def2-TZVP were used for the DFT calculations.[31] The Cartesian coordinates of the optimized geometries of the relevant THFA and $(THFA)_2$ conformers are provided in Table A2, Appendix A. The simulated IR and VCD spectra were generated by assigning a Lorentzian line shape with half-width at half maximum (HWHM) of 4 cm^{-1} to the predicted dipolar and rotational strength values. The frequencies are scaled with a factor of 0.975, only when compared to the experimental data directly.

To systematically explore the possible conformers of the THFA monomer and dimer, we utilized the CREST (conformer-rotamer ensemble sampling tool) code developed by Grimme and co-workers.[20,21] The CREST code, built upon the previous semiempirical tight-binding quantum chemistry method (GFN2-xTB),[21] can automatically perform multiple MD runs for conformational exploration. It has a built-in feature to compare and discard redundant geometries. The new code is a powerful tool for screening the conformational space of mid- to large-sized molecules with up to about a thousand atoms.

3.5 Acknowledgements

This work was funded by the Natural Sciences and Engineering Research Council of Canada and by the University of Alberta. We gratefully acknowledge access to the computing facilities by the Shared Hierarchical Academic Research Computing Network, the Western Canada Research Grid (Westgrid), and Compute/Calcul Canada. YX is a Tier I Canada Research Chair in Chirality and Chirality Recognition.

References

[1] (a) L. A. Nafie, Vibrational Optical Activity: Principles and Applications, John Wiley & Sons, Ltd., UK, **2011**; (b) L. D. Barron, *Molecular light scattering and optical activity, 2nd edn. Cambridge*, **2004**.

[2] (a) K. Le Barbu-Debus, J. Bowles, S. Jähnigen, C. Clavaguéra, F. Calvo, R. Vuilleumier, A. Zehnacker, *Phys. Chem. Chem. Phys.* **2020**, *22*, 26047-26068; (b) C. Merten, *Phys. Chem. Chem. Phys.* **2017**, *19*, 18803-18812; (c) A. S. Perera, J. Thomas, M. R. Poopari, Y. Xu, *Front. Chem.* **2016**, *4*, 1-17; (d) G. Yang, Y. Xu, *Top. Curr. Chem.* **2011**, *298*, 189-236.

[3] (a) V. P. Nicu, E. J. Baerends, P. L. Polavarapu, *J. Phys. Chem. A* **2012**, *116*, 8366-8373; (b) Y. Liu, G. Yang, M. Losada, Y. Xu, *J. Chem. Phys.* **2010**, *132*, 234513; (c) M. R. Poopari, Z. Dezhahang, Y. Xu, *Phys. Chem. Chem. Phys.* **2013**, *15*, 1655-1665.

[4] (a) G. Yang, Y. Xu, *Phys. Chem. Chem. Phys.* **2008**, *10*, 6787-6795; (b) M. Losada, H. Tran, Y. Xu, *J. Chem. Phys.* **2008**, *128*, 014508.

[5] (a) G. Tarczay, G. Magyarfalvi, E. Vass, *Angew. Chem. Int. Ed.* **2006**, *45*, 1775-1777; *Angew. Chem.* **2006**, *118*, 1807-1809; (b) C. Merten, Y, Xu, *ChemPhysChem* **2013**, *14*, 213-219; (c) C. H. Pollok, C. Merten, *Phys. Chem. Chem. Phys.* **2016**, *18*, 13496-13502.

[6] (a) C. H. Pollok, T. Riesebeck, C. Merten, *Angew. Chem. Int. Ed.* **2017**, *56*, 1925-1928; *Angew. Chem.* **2017**, *129*, 1952-1955; (b) C. Merten, Y. Xu, *Angew. Chem. Int. Ed.* **2013**, *52*, 2073-2076; *Angew. Chem.* **2013**, *125*, 2127-2130;

[7] (a) Y. Fujima, Y. Hirayama, M. Ikunaka, Y. Nishimoto, *Tetrahedron-Asymm.* **2003**, *14*, 1385-1391; (b) W. Chou, M. Chou, Y. Lu, S. Chen, *US Patent*, **2001**, 6316293(B1); (c) F. Codignola, M. Piacenza, *Italian Patent*, **1947**, ES179144(A1).

[8] (a) F. Xie, X. Ng, N. A. Seifert, J. Thomas, W. Jäger, Y. Xu, *J. Chem. Phys.* **2018**, *149*, 224306; (b) F. Xie, N. A. Seifert, W. Jäger, Y. Xu, *Angew. Chem. Int. Ed.* **2020**, *59*, 15703-15710; *Angew. Chem.* **2020**, *132*, 15833-15840.

[9] See for example, (a) W. Li, L. Evangelisti, Q. Gou, W. Caminati, R. Meyer, *Angew. Chem. Int. Ed.* **2019**, *58*, 859-865; *Angew. Chem.* **2019**, *131*, 869-875; (b) J. Thomas, A. Serrato III, W. Lin, W. Jäger, Y. Xu, *Eur. J. Chem.* **2014**, *20*, 6148-6153; c) M. Shome, N. Mishra, *Indian J. Adv. Chem. Sci.* **2016**, *4*, 56-67.

[10] T. Kuppens, W. Herrebout, B. Veken, P. Bultinck, *J. Phys. Chem. A* **2006**, *110*, 10191-10200.

[11] See for example, M. Halupka, W. Sander, *Spectrochim. Acta Part A* **1998**, *54*, 495-500.

[12] E. M. S. Maçôas, P. Myllyperkiö, H. Kunttu, M. Pettersson, *J. Phys. Chem. A* **2009,** *113,* 7227-7234.

[13] J. Ahokas, I. Kosendiak, J. Krupa, J. Lundell, M. Wierzejewska, *J. Mol. Struct.* **2018**, *1163,* 294-299.

[14] A. Aerts, A. Brown, F. Gatti, *J. Chem. Phys.* **2022**, *157*, 014306.

[15] M. Winnewisser, B. P. Winnewisser, M. Stein, M. Birk, G. Wagner, G. Winnewisser, K. M. T.

Yamada, S. P. Belov, O. L. Baskakov, *J. Mol. Spectrosc.* **2002**, *216*, 259-265.

[16] M. Pettersson, E. M. S. Maçôas, L. Khriachtchev, R. Fausto, M. Räsänen, *J. Am. Chem. Soc.* **2003**, *125*, 4058-4059.

[17] G. Giubertoni, O. O. Sofronov, H. J. Bakker, *J. Chem. Phys. Lett.* **2019**, *10*, 3217-3222.

[18] A. S. Perera, J. Cheramy, M. R. Poopari, Y. Xu, *Phys. Chem. Chem. Phys.* **2019**, *21*, 3574-3584.

[19] K. A. E. Meyer, M. A. Suhm, *J. Chem. Phys.* **2018**, *149*, 104307.

[20] S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 2847-2863.

[21] (a) S. Grimme, C. Bannwarth, P. Shushkov, *J. Chem. Theory Comput.* **2017**, *13*, 1989-2009; (b) S. Grimme, C. Bannwarth, S. Dohm, A. Hansen, J. Pisarek, P. Pracht, J. Seibert, F. Neese, *Angew Chem. Int. Ed.* **2017**, *56*, 14763-14769; *Angew. Chem.* **2017**, *129*, 14958-14964.

[22] See for example, (a) F. Xie, N. A. Seifert, M. Heger, J. Thomas, W. Jäger, Y. Xu, *Phys. Chem. Chem. Phys.* **2019**, 21, 15408-15416; (b) S. Oswald, N. A. Seifert, F. Bohle, M. Gawrilow, S. Grimme, W. Jäger, Y. Xu, M. A. *Angew. Chem. Int. Ed.* **2019**, *58*, 5080-5084; Suhm, *Angew. Chem.* **2019**, *131*, 5134-5138.

[23] F. Xie, M. Fusè, A. S. Hazrah, W. Jäger, V. Barone, Y. Xu, *Angew. Chem. Int. Ed.* **2020**, *59*, 22427-22430; *Angew. Chem.* **2020**, *132*, 22613-22616.

[24] (a) F. Madeja, M. Havenith, K. Nauta, R. E. Miller, J. Chocholoušová, P. Hobza, *J. Chem. Phys.* **2004**, *22*, 10544-10560; (b) M. Losada, H. Tran, Y. Xu, *J. Chem. Phys.* **2008**, *128*, 014508; (c) M. Albrecht, V. A. Soloshonok, L. Schrader, M. Yasumoto, M. A. Suhm, *J. Fluo. Chem*. **2010**, *131*, 495-504; (d) M. Shome, N. Mishra, *Indian J. Adv. Chem. Sci.* **2016**, *4*, 56-67; (e) G. Giubertoni, O. O. Sofronov, H. J. Bakker, *J. Phys. Chem. Lett.* **2019**, *10*, 3217-3222.

[25] A. Schriver, L. Schriver-Mazzuoli, A. A. Vigasin, *Vib. Spectrosc.* **2000**, *23*, 83-94.

[26] S. Abbate, L. F. Burgi, F. Gangemi, R. Gangemi, F. Lebon, G. Longhi, V. M. Pultz, D. A. Lightner, *J. Phys. Chem. A* **2009***, 113,* 11390-11405.

[27] Gaussian 16, Revision B.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, **2016**.

[28] A. D. Becke, *J. Chem. Phys*. **1993**, *98*, 5648.

[29] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.* **2011**, *32*, 1456-1465.

[30] A. D. Becke, E. R. Johnson, *J. Chem. Phys.* **2005**, *123*, 154101.

[31] D. Rappoport, F. Furche, *J. Chem. Phys.* **2010**, *133*, 134105.

Chapter 4

Vibrational Circular Dichroism Spectroscopy of Tetrahydro-2-Furoic Acid in Water under Different pH values

4.1 Introduction

Tetrahydro-2-furoic acid (THFA), a colorless oil, is reduced from 2-furoic acid.[1] In recent years, it has been widely used in pharmaceutical areas as a key reactant to produce drugs. For example, alfuzosin is one of the drugs which is synthesized from THFA and also it is used for the treatment of benign prostatic hyperplasia (BPH).[2] Moreover, THFA is also an important intermediate for the synthesis of faropenem, an antibiotic for the treatment of acute bacterial sinusitis.[3] It is of interest to gain a fundamental understanding on how the preferred conformations of THFA under different environments are existed and such knowledge may help us explore more possibilities of using THFA in pharmaceutical syntheses. It is also known that local environment can influence drug efficiency. For example, didanosine, an anti-retroviral drug, widely used to reduce the level of HIV in the blood, was found to degrade quickly under acidic condition, significantly reducing its pharmacological effect.[4] Since water is the common environment for many biological events,[5-7] it would be of significant interest to probe conformational distributions and the intermolecular hydrogen-bonding interactions of THFA in water under different pH environments.

Vibrational circular dichroism (VCD) spectroscopy has been widely used to investigate the absolute configuration and conformations of different kinds of chiral molecules in solution.[8,9] It measures the differences between the left and right circularly polarized IR radiation of a given chiral molecule accompanying a vibrational transition.[10] The intensity of the VCD signal is defined by the rotational strength, i.e., the imaginary part of the dot product of the electric and magnetic transition dipole moment vectors. VCD spectroscopy is highly sensitive to conformation changes and chirality of the chiral molecular systems. As a result, one can further apply VCD to investigate solute-solvent interactions and solute-solute interactions.[11-15] This is the spectroscopic method of choice used in the current study.

In terms of conformational properties of THFA, there have been a number of previous studies. Using VCD spectroscopy, Kuppens et al.[12] identified the dominant species of

THFA in chloroform where both the monomeric THFA conformers and its cyclic hydrogen-bonded THFA dimers (i.e., Type 1 dimers) were observed experimentally. Using jet-cooled rotational spectroscopy, Fan et al. studied the conformational distributions of THFA monomers [16] and its dimers [17] in great detail. While the *cis*-COOH configuration is generally favored by the majority of carboxylic acids,[18-20] THFA was reported to exist mainly in the *trans*-COOH configurations in the gas phase, thanks to the intramolecular H-bond formed between the carboxylic OH and the ester oxygen in the ring. Because of this somewhat unusual monomeric conformational distribution, both Type 1 THFA dimers, i.e., the cyclic hydrogen-bonded THFA dimers, which consist of two *cis*-THFA subunits, and Type 2 THFA dimers which consist of one *trans*- and one *cis*-THFA subunits, were detected with similar abundances in a supersonic jet expansion. Finally, in our recent matrix-isolation VCD study, we discovered the existence of the even less stable Type 3 THFA dimers which are made of two *trans*-THFA subunits.[21] This was discussed in Chapter 3. In addition, hydrogen-bonded binary complexes between THFA and water[22] and between THFA and propylene oxide [23] were also investigated using jet-cooled rotational spectroscopy. While in the former case, only the kinetically favored Type 2 monohydrate conformers consisting of a *trans*-THFA subunit were observed experimentally, in the latter case, both Type 1 and Type 2 binary chiral contact pairs were observed. Clearly, different chemical environments and/or experimental conditions can favor very different binary THFA conformations.

How to account for the water solvation effects is a subject of significant current research interest. One approach is the clusters-in-a-liquid model[24] which is largely based on the experimental detection of induced VCD features at the water bending regions,[25-27] suggesting that such chiral solute-water$_n$ complexes are relatively long-lived. This model utilizes small long-lived chiral solute-water$_n$ clusters to account for the explicit solute-solvent interactions and places such hydration clusters in the polarizable continuum model (PCM) of water to account for the bulk environment. It requires only modest computational

effort and has been applied to a number of molecular systems in solution successfully.[28,29] More recently, there have been a number of reports where the performances of the static, such as the clusters-in-a-liquid model, and the dynamic approaches were compared.[30,31]

In the current study, we will apply the clusters-in-a-liquid approach and utilize a recently developed conformational sampling tool, i.e., Conformer-Rotamer Ensemble Sampling Tool (CREST)[32] for the systematical conformational searches. The simulated results will be compared with the experimental data to extract the dominant THFA species and their preferred conformations in aqueous solutions. Furthermore, we will discuss the quality of the agreement and the causes of less satisfying agreement in some case. One focus point is the water-centered large amplitude motions which have been recently investigated using rotational spectroscopy.[33,34] These recent gas phase studies emphasized that the properties of a minimum and its related conformer can differ greatly where some barrierless large amplitude motions are involved. In the current study, we will examine the connection between the gas and solution phase studies and discuss the implication of the large amplitude motions in modelling water solvation effects by using small hydration clusters.

4.2 Experimental Section

4.2.1 Experimental

Both R-(+)-THFA (99%) and S-(-)-THFA (97%) were purchased from Sigma Aldrich and used without further purification. Both enantiomers were in oil-like liquid phase. All solutions were prepared in distilled water. The experimental IR and VCD spectra were measured by using a Bruker Vertex 70 supplemented with a PMA 50 module for polarization modulated measurements. In order to obtain a good VCD spectrum, the absorbance of the target IR bands should be between 0.2 and 0.8. All experimental IR and VCD spectra were measured by using a BaF_2 sample cell. No spacer was used for all the measurements. The concentration of the THFA solution is 5 M. As for the THFA solutions,

one varies the pH values ranging from 1.31 to 13.25 by adding a small amount of 3 M NaOH solution dropwise. All the raw VCD spectra were measured with a total acquisition time 3 hours (3×1 hr) and with a spectral resolution of 4 cm^{-1}. The final VCD spectra were generated by using the standard (S-R)/2 procedure.

4.2.2 Computational

For the systematic conformational searches, we utilized CREST with the inclusion of dubbed GBSA implicit solvation model of water.[32] CREST makes use of Grimme's semi-empirical extended tight-binding scheme, GFN-xTB (abbreviate for Geometries, Frequencies, and noncovalent interactions – extended Tight Binding).[35] It is a semiempirical method and is combined with a meta-dynamics (MTD) driven search algorithm to identify minima. The GFN2-xTB[36,37] basis set contains augmented atomic orbital (AO) with a second s-function for hydrogen and d-polarization function for heavier elements to enhance the modeling of hydrogen and hypervalent bonding. This is different from other most common semiempirical methods which only contain minimal AO basis set. In addition, the MTD simulations could be carried out in CREST program in the gas-phase or with the implicit solvation model.[38] This conformational tool has been utilized extensively in the rotational spectroscopic studies of non-covalently bonded species in our group[39-41] and has so far failed only in one case where CREST could not generate the heterochiral propylene oxide trimer candidate, even though the structure was determined experimentally using isotopic data.[42]

The CREST candidates were subjected to the subsequent geometry optimizations, harmonic frequencies calculations and IR and VCD intensity calculations using the Gaussian 16 program package.[43] All DFT calculations were undertaken by using the B3LYP functional[44] with D3 dispersion corrections[45] and Becke-Johnson (BJ) damping,[46] i.e., B3LYP-D3BJ. The basis sets used was def2-TZVP, which is developed by group of Ahlrichs,[47,48] to optimize the lowest-energy candidates of the neutral and deprotonated

THFA monomers, monohydrates and even larger water containing clusters.

To account for solvation effect, both the integral equation formalism (IEF) implicit PCM solvation model[49] and the explicit solvation model were included. The latter was based on the hydrogen bonded clusters of THFA with one water molecule and of deprotonated THFA with up to four water molecules. The dielectric constant of H_2O, ε = 78.3553, was used in PCM calculations. The simulated IR and VCD spectra were generated by assigning a Lorentzian line shape with half-width at half maximum (HWHM) of 4 cm^{-1} to the calculated dipole and rotational strength. The simulated IR and VCD spectra were compared with the experimental IR and VCD spectra with no scaling. Please note that we use the S enantiomer of neutral and deprotonated THFA throughout this paper for the presentations of the calculations.

4.3 Results and Discussion

4.3.1 Experimental IR spectra

One main goal is to probe the preferred conformations of the dominant THFA species in water under different pH environments. The IR spectra obtained at several pH values from 1.31 to 13.25 are summarized in Figure 4.1 in the frequency region from 1000 to 1500 cm^{-1} to avoid the strong water bending band in the 1600 cm^{-1} region. From Figure 4.1, one can see that the experimental IR spectra under different pH environments change significantly from low pH environments (pH=1.31, 2.62) to the high pH environments (pH > 5.70). Without the addition of strong additional acid or base, the pH of THFA in water is 1.31. With a dissociation constant (pKa) of 3.60, one can estimate that 0.5% of THFA is in the COO^- form and 99.5% is in the COOH (neutral) form at pH=1.31 according to the Henderson-Hasselbalch equation, $pH = pK_a + \log\{[A^-]/[HA]\}$. At pH=2.62, more than 90% of THFA molecules are in neutral form, generating very little change in the IR spectrum compared to that at pH=1.31.

As the pH increase to 3.91, the intensities of IR bands located at 1240 cm^{-1} and 1360

cm^{-1} became much weaker than at 1.31 or 2.62, and furthermore, two new IR bands emerged at ~1310 cm^{-1} and ~1420 cm^{-1}. This is because at pH=3.91, 33% of THFA molecules were estimated to exist as the COOH form and 67% as the COO^- form, a noticeable shift to the COO^- form from the low pH values. As pH further increases to 5.7 or even higher, the IR bands remain more or less the same, suggesting there are no new species generated. Comparing with the pH=3.91 spectrum, the band at 1240 cm^{-1} disappeared and the intensities of bands at 1310 cm^{-1} and 1420 cm^{-1} became larger.

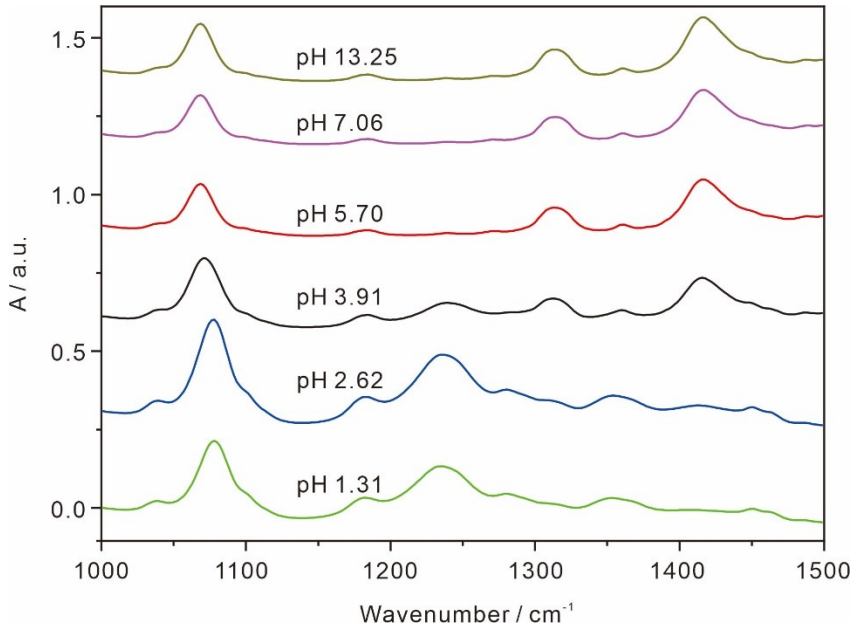


Figure 4.1. The experimental IR spectra of THFA in water with pH ranging from 1.31 to 13.25.

4.3.2 Simulated IR and VCD spectra of the most stable conformers of neutral and deprotonated THFA with the PCM of water

First, we examined the possible conformations of THFA in the neutral and deprotonated form. Following the recent gas phase study, we classified the neutral THFA monomer into the *cis*- and *trans*-COOH group where C=O and OH of the acid group are on the same or opposite side, respectively.[12] In addition, there are different THF ring puckering

configurations. The neutral THFA CREST candidates produced similar conformations in the PCM of water are nearly the same as those in the gas phase, and the geometries of the six most stable ones are depicted in Figure 4.2. The relative free energies and the Boltzmann factors are summarized in Table 4.1. The two most stable neutral THFA monomers, I and II, take on the *trans*-COOH configuration with their hydroxyl group pointing toward the oxygen atom on 5-membered ring forming an intramolecular hydrogen bond. It is difficult for them to self-aggregate or bind to water in these two conformations. The other four neutral THFA conformers have their OH and C=O groups pointing to the same direction, i.e., *cis*-THFA, facilitating the formation of the double cyclic hydrogen bonded ring in the THFA dimers or interaction with water.

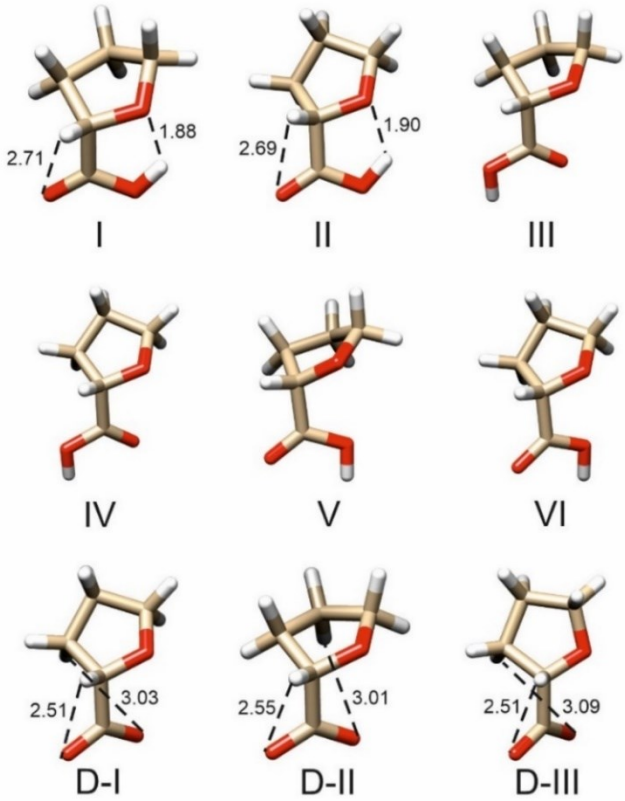


Figure 4.2. The most stable conformations of the neutral and the deprotonated THFA species predicted at the B3LYP-D3BJ/def2-TZVP level of theory with the PCM of water. The intramolecular hydrogen bond lengths (in Å) are indicated.

A similar conformational search was also carried out for the deprotonated THFA. In this case, the conformational landscape was simplified because the two Os in the COO^- moiety are equivalent and there is no possibility to form an intramolecular hydrogen bond either. Only three low energy conformers were identified, and they both have an intramolecular CH---O attractive contact and differ only slightly in their ring puckering conformation (Figure 4.2). We use D to abbreviate the deprotonated form of the THFA molecule and the Roman numbers to indicate the ordering of the relative free energies with I being the most stable. The relative Gibbs free energy and Boltzmann distribution factor of these two deprotonated conformations are given in Table 4.1.

Table 4.1. The relative free energies, ΔG, in kJ mol^{-1}, and Boltzmann percentage factors, Bf in %, of the conformers of the neutral and deprotonated THFA at 298 K at the B3LYP-D3BJ/def2-TZVP/PCM(water) levels of theory.

Conf.	**ΔG**	**B_f**
Neu-THFA		
I	0.0	54.7
II	0.9	38.7
III	7.9	2.3
IV	8.0	2.2
V	9.8	1.1
VI	10.0	1.0
VII	15.9	0.0
VIII	16.7	0.0
Dep-THFA		
I	0.0	39.5
II	0.5	32.6
III	0.9	27.9

In Figure 4.3, the simulated IR and VCD spectra of the six neutral THFA conformers are summarized together with their Boltzmann weighted averaged spectra. These are compared with the experimental spectra at pH=1.31. Moreover, we compare the observed

IR and VCD spectra at pH 1.31 with the six most stable neutral THFA conformer to evaluate the contribution of each conformer. Both experimental IR and VCD spectra are in poor agreement with the Boltzmann averaged ones. Since there has been much discussion about the accuracy of the DFT relative energies,[50] we also examined the contributions from each neutral THFA conformers and used empirical weights to attempt to improve the agreement. Here are the empirical population weights: 40% III, 54% IV, 2% V, 2% VI, 1% I and 1% II. As one can see in Figure 4.3, the empirical weighted IR and VCD spectra improve the agreement with the experimental result between 1000 and 1150 cm^{-1} region, but there is no good correlation for the rest of the region shown. Again, the agreement between experiment and theory is poor.

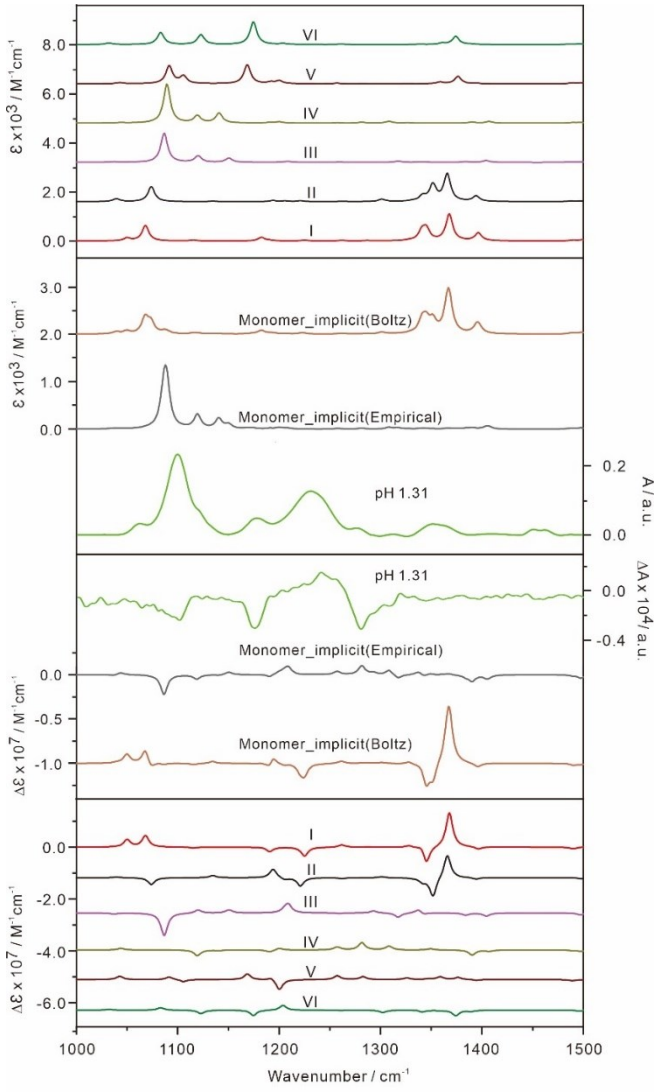


Figure 4.3. Comparisons of experimental IR and VCD spectra of THFA in water at pH 1.31 environment with the simulated Boltzmann averaged IR and VCD spectra of neutral THFA monomer the six most stable conformers at 298 K with the inclusion PCM of water at the B3LYP-D3BJ/def2-TZVP level of theory. Empirical population weighted IR and VCD spectra under were listed to aim the observation of the population of monomer.

Since the concentration used here is fairly high, we also considered the possible contribution of the THFA dimer in water. Here the dominant binary species is the Type 1 dimers which take on the common cyclic hydrogen bonded ring.[12] These Type 1 dimers differ slightly in their ring-puckering. Their relative free energies and Boltzmann population factors are provided in Table B1, Appendix B. In Figure 4.4, the Boltzmann averaged IR and VCD spectra of the THFA dimers are compared with the corresponding experimental data obtained at pH=1.31. As one can see from Figure 4.4, the dominant VCD features of the THFA dimers in the 1300-1400 cm^{-1} region are clearly not observed experimentally. We could therefore rule out the significant formation of the THFA dimers. This conclusion is also consistent with the previous study of carboxylic acids in several solvents such as chloroform, methanol, and acetonitrile.[51] While in chloroform, the less polar solvent, dimerization of the carboxylic acids was observed, in all the other more polar solvents, the explicit hydrogen bonding with the solvent molecules dominates and very little dimerization was observed. Since water is smaller and even more polar than those listed above, it is logical that we do not expect much dimerization and it is necessary to consider explicit interactions between THFA and water molecules.

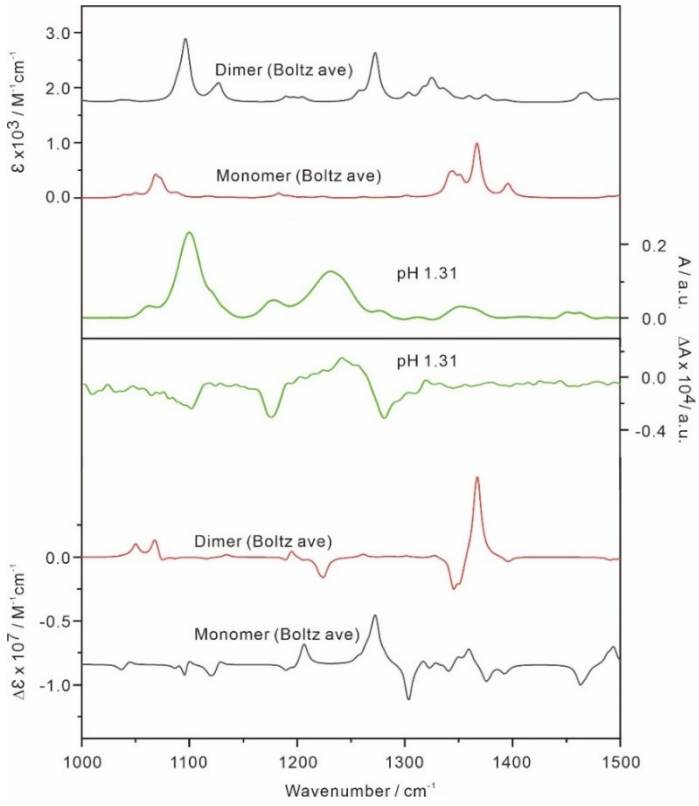


Figure 4.4. Comparisons of experimental IR and VCD spectra of THFA in water at pH 1.31 environment with the simulated Boltzmann averaged IR and VCD spectra of neutral THFA monomer and dimer with the inclusion PCM of water at the B3LYP-D3BJ/def2-TZVP level of theory.

4.3.3 The combined implicit and explicit solvation model of the IR and VCD spectra of THFA at pH=1.31

We applied CREST to generate conformational candidates of the binary complex of THFA and water. The optimized geometries of the six most stable THFA monohydrates are provided in Figure 4.5. Interestingly, by adding just one explicit water molecule, the lowest energy conformations of THFA are now all in the *cis*-COOH configuration. The *trans*-COOH monohydrates are about 8 kJ mol^{-1} less stable than the *cis*-COOH ones in the gas phase[22] and in the PCM of water, they are about 2 kJ mol^{-1} less stable, which are listed in Figure B1, Appendix B. We name the monohydrates based on the conformations of their THFA subunits, as well as the relative free energy ordering with 1 being the most stable.

To avoid any confusion, we also consider the naming separately between *cis*- and *trans*-THFA monohydrates. The relative free energies and Boltzmann factors distribution population of the six most stable neutral *cis*-THFA monohydrates are provided in Table 4.2.

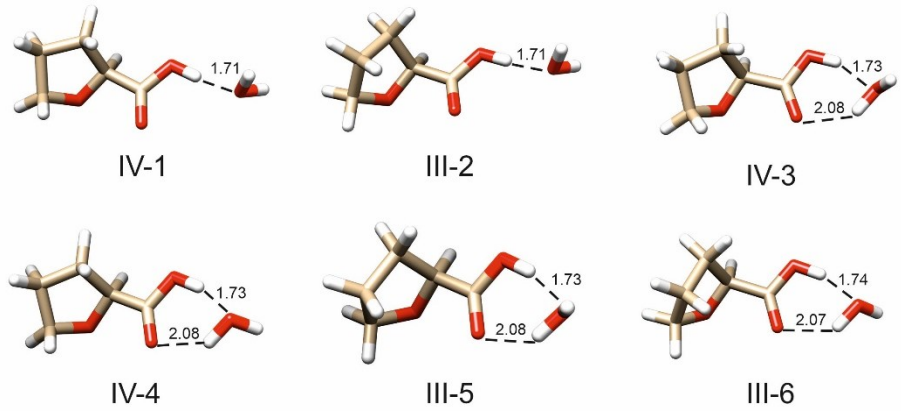


Figure 4.5. The most stable conformations of neutral THFA monohydrate at the B3LYP-D3BJ/def2-TZVP level of theory with the PCM of water. The intramolecular hydrogen bond lengths (in Å) are indicated.

Table 4.2. The relative free energies, ΔG, in kJ mol^{-1}, and Boltzmann percentage factors, Bf, in %, and empirical weight factor, Ef, in % for the conformers of the neutral THFA monohydrate at 298 K at the B3LYP-D3BJ/def2-TZVP/PCM (water) levels of theory.

Conf.	ΔG	B_f	E_f
IV-1	0.0	29.0	14.0
III-2	0.6	22.7	40.0
IV-3	1.6	15.5	7.0
IV-4	2.0	12.9	5.0
III-5	2.4	11.0	17.0
III-6	2.9	8.9	17.0

We should mention the simulated IR and VCD spectra of the *trans*-COOH monohydrate conformers (Figure B1, Appendix B) do not agree at all with the experimental observation and are therefore not included in the comparison below for conciseness. The simulated IR

and VCD spectra of the six individual *cis*-THFA monohydrate and the Boltzmann averaged spectra are provided in Figure 4.6, together with the experimental data. Obviously, the agreement between the experimental and simulated IR spectra improves drastically. The same can be said about the agreement for the VCD spectra. To facilitate discussions, we also number all main bands in IR and in VCD. For example, VCD features labelled with 1, 2, 4, 5, 6, and 10 are mainly from THFA-IV monohydrates, while bands labelled with 3, 7, 8, 9, 11 are contributed by THFA-III monohydrates.

There are some minor discrepancies between the simulated and the experimental VCD spectra. For instance, the relative intensity for band 3 and 4 in the experimental VCD spectrum was not perfectly captured. By using the empirically weighted IR and VCD spectra with a net population of 74% THFA-III monohydrate and 26% THFA-IV monohydrate, one can see some improvement in the agreement between experimental and theory in Figure 4.6. For example, the relative intensities for band 2 and 3 in simulated VCD spectrum improves. Moreover, the IR band patterns at 1200 – 1250 cm^{-1} correlate better with the experimental results. It is likely that the population of THFA-III monohydrates are higher than the original DFT prediction suggested.

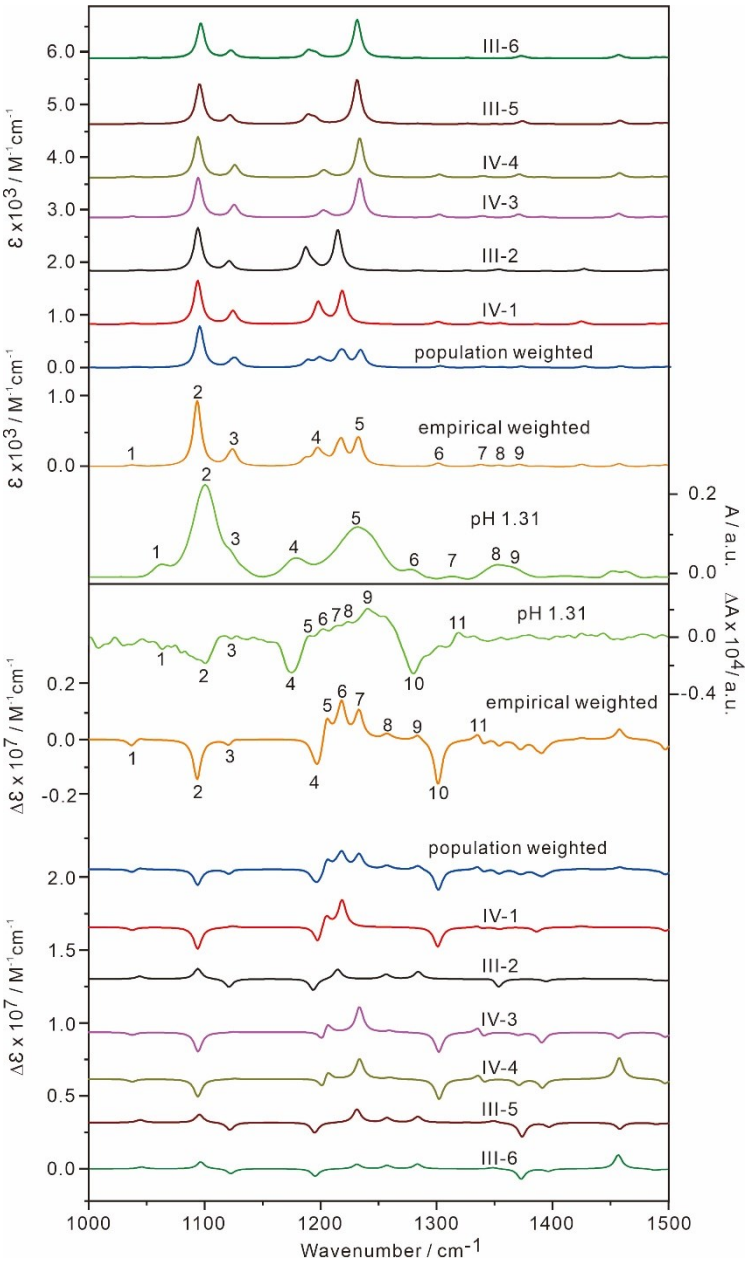


Figure 4.6. Comparisons of experimental IR and VCD spectra of THFA at acidic condition (pH 1.31) with the simulated IR and VCD Boltzmann averaged spectra of neutral THFA monohydrate under B3LYP-D3BJ/def2-TZVP/PCM (water) level of theory. Empirical population weighted IR and VCD spectra were also listed to aim the observation of conformation equilibrium of THFA monohydrate.

It is also interesting to compare with the previous rotational spectroscopic study of the

THFA monohydrates in a supersonic jet expansion.[22] In the supersonic jet expansion, only Type 2 THFA monohydrates, i.e., THFA I and II monohydrates, were observed experimentally. This is because the supersonic jet expansion provides a highly kinetic trapping environment and the most stable THFA I and II *trans*-COOH monomeric conformers could not convert freely to the *cis*-COOH configuration. The situation is much different in water because there is ample time to reach equilibrium which favors the more stable species.

4.3.4 The combined implicit and explicit solvation model of the IR and VCD spectra of THFA at pH=13.25

With the confidence we gained from the IR and VCD analyses of THFA in a highly acidic environment, we applied the same procedure to deal with THFA under highly basic condition. First, we searched for possible conformations of D-I, D-II and D-III monohydrates, The optimized geometries of these monohydrates with the PCM of water are shown in Figure 4.7 where the important intermolecular hydrogen and van der Waals bond lengths are indicated. The relative Gibbs free energies and corresponding Boltzmann population factors are summarized in Table 4.3. These monohydrates fall into two different categories: a) water is hydrogen-bonded to the COO^- group; b) water forms simultaneous contact with COO^- and the ether oxygen to form a seven membered hydrogen bonded ring.

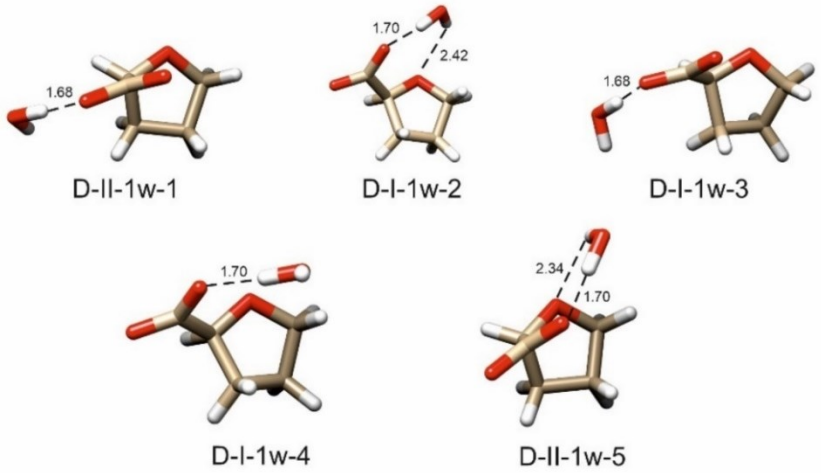


Figure 4.7. The most stable conformations of deprotonated THFA monohydrate at the B3LYP-D3BJ/def2-

TZVP level of theory with the PCM of water. The intramolecular hydrogen bond lengths (in Å) are indicated.

Table 4.3. The relative free energies, ΔG, in kJ mol^{-1}, Boltzmann percentage factors, B_f, in %, and empirical weight factor, E_f, in % for each conformer of deprotonated THFA-1w at 298K at the B3LYP-D3BJ/def2-TZVP/PCM (water) levels of theory.

Conf.	ΔG	B_f	E_f
D-II-1w-1	0.0	39.9	35.0
D-I-1w-2	2.0	17.7	30.0
D-I-1w-3	2.5	14.8	5.0
D-I-1w-4	2.5	14.8	15.0
D-II-1w-5	2.8	12.8	15.0

The individual IR and VCD spectra of the deprotonated THFA monohydrate conformers are depicted in Figure 4.8. The Boltzmann weighted IR and VCD spectra are compared with the experimental spectra at pH=13.25. First it appears that the simulated IR spectrum is in good agreement with the experimental one, capturing the changes observed from pH=1.31 to pH=13.25. All eight experimental IR bands labelled can be correlated to the simulated ones straightforwardly. This indicates that the dominant species have been correctly predicted. The agreement between the calculated and experimental VCD spectra, however, is not very good. For example, a distinctive negative with a broad shoulder has been observed for band 2 (C-O-C stretching) in the experimental VCD result, while a weaker and broader band were computed in the simulated spectrum. Moreover, the relative intensity for band 4, which consists of a mixture of CH_2 wagging, CH_2 twisting, CH_2 rocking and CH bending motions in the ring, is much weaker in the simulation when compared to the corresponding experimental VCD feature.

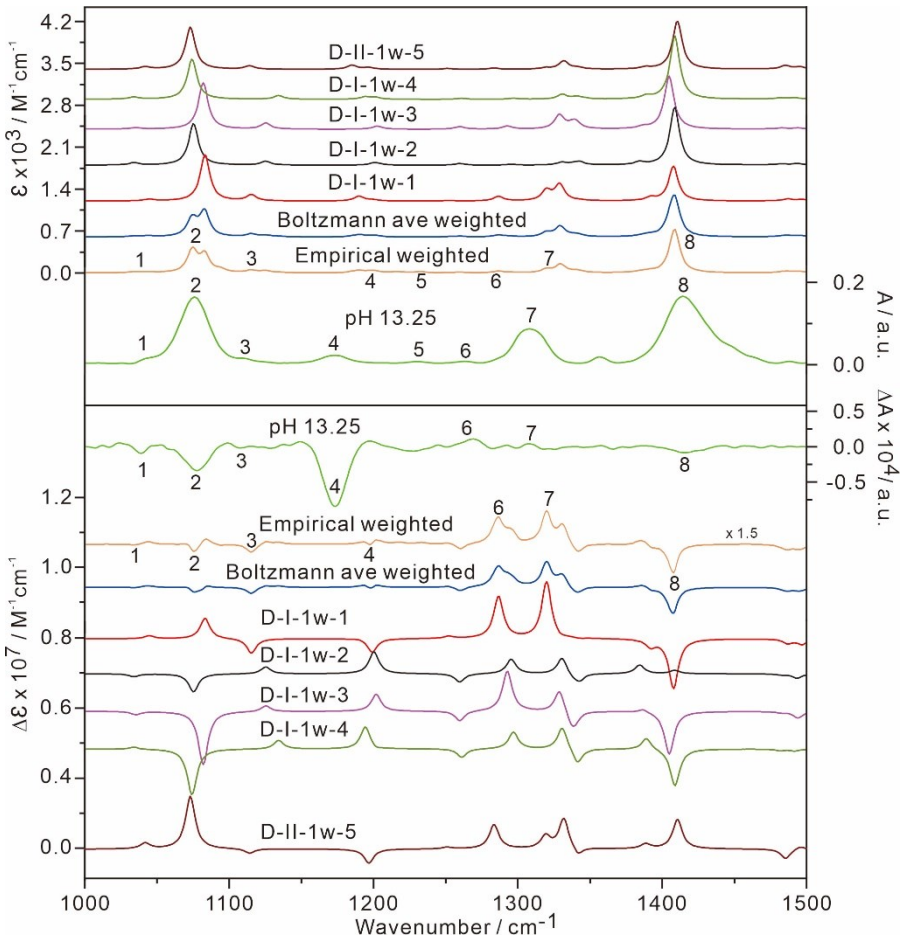


Figure 4.8. Comparisons of experimental IR and VCD spectra of THFA at basic condition (pH 13.25) with the simulated IR and VCD Boltzmann averaged spectra of deprotonated THFA monohydrate under B3LYP-D3BJ/def2-TZVP/PCM (water) level of theory and the simulated IR and VCD Boltzmann averaged spectra of deprotonated THFA monohydrate under B3LYP-D3BJ/def2-TZVP/PCM (water) level of theory as well as the final empirically weighted spectra.

We attempted to adjust the contributions of each deprotonated monohydrate conformers empirically: 35% D-I-1w-1, 30% D-I-1w-2, 5% D-I-1w-3, 15% D-I-1w-4 and 15% D-II-1w-5. Furthermore, we also included a small amount of neutral THFA monohydrates (0.95:0.05) to simulate the final empirical weighted IR and VCD spectra. While the agreement is very good for the IR spectra, and there is considerable improvement for the

VCD. Overall, the quality of the agreement is still much worse than for the neutral THFA case discussed above.

What could be the cause for the less satisfying agreement for VCD for the deprotonated species? Maybe it is necessary to generate larger hydrates because the species is now charged.[52] We therefore further increased the explicit water molecules one at a time from one to four and simulated all the IR and VCD spectra. The optimized geometries for the deprotonated THFA-2w, -3w, and -4w are provided in Figure B2, Appendix B and the relative free energies and Boltzmann factors for these systems are listed in Table B2, Appendix B. The Boltzmann averaged IR and VCD spectra are summarized in Figure B3, Appendix B, while the individual conformer spectra are given in Figure B4, B5 and B6, Appendix B for those with two, three and four water molecules, respectively. As one increases the number of explicit water molecules, the VCD bands in the 1050-1220 cm^{-1} region change greatly, while those in the 1250-1350 cm^{-1} region experience less dramatic changes whereas IR spectral patterns remain largely unchanged. Unfortunately, the additional consideration of these larger hydration clusters did not improve the agreement between the experimental and theoretical VCD spectra.

In search of a possible cause of the disagreement between theory and experiment, we examined some rotational spectroscopic studies of monohydrates with chiral molecules such as propylene oxide,[53] methyl lactate,[54] methyl glycidate[55] and 1-phenyl-2,2,2-trifluroethanol (PhTFE).[34] While in the first two cases, the spectroscopic properties predicted at the minima seemed to agree well with those observed, the situation was quite different in the later cases. In both monohydrates of methyl glycidate and PhTFE, some physical properties, such as electric dipole components predicted at the minima, differ drastically from the experimental ones which is measured at the ground vibrational state not at equilibrium. Such drastic deviations from the predicted equilibrium properties seem particularly common in monohydrates as it was reported in many rotational spectroscopic studies,[56,57,58] although the link to the solution modeling was rarely mentioned in these

studies. For example, in the case of PhTFE monohydrate, we found that one of the low energy minimum predicted is actually not a true conformer and while its simulated IR spectrum is very similar to the global minimum, its simulated VCD spectrum has opposite signs at several prominent bands.[34] Generally, if the potential energy surface is fairly flat near several minima, the usual static approach of using a harmonic frequency calculation based on minima to simulate the final VCD spectra breaks down. It is, however, difficult to predict when such breakdowns would happen. In addition, we also know from the previous gas phase studies that the THF ring puckering barriers are quite small in both the THFA monomers and dimers.[16,17] These evidences suggest that we may have experienced such a breakdown in applying the usual static DFT approach. Further proper simulation of the actual ground state properties is out of scope of the current work and will likely be pursued in collaboration with some theorists.

4.4 Conclusion

The detailed conformational distributions of the dominated THFA conformers under the highly acidic and basic environments were investigated by using IR and VCD spectroscopy, together with DFT calculations. Solely taking into account of the implicit water solvation failed to capture the experimental IR and VCD features. With the addition of just one water molecule explicitly to the model, a very good agreement with the experimental IR and VCD spectra was achieved, highlighting the importance of hydrogen bonding between solute and solvent molecules. The same approach provided good agreement with the experimental IR spectrum at pH=13.25, although the agreement for the VCD spectrum is less good. We hypothesized that some water-centered large amplitude motions maybe responsible to the less satisfying agreement. Overall, the results highlight the power of using VCD spectroscopy to determine the conformational distributions of chiral molecules in aqueous solutions.

4.5 Acknowledgements

This work was funded by the Natural Sciences and Engineering Research Council of Canada and by the University of Alberta. We gratefully acknowledge access to the computing facilities by the Shared Hierarchical Academic Research Computing Network, the Western Canada Research Grid (Westgrid), and Compute/Calcul Canada. YX is a Tier I Canada Research Chair in Chirality and Chirality Recognition.

References

[1] Z. Li, Y. Ding, W. Jiang, *Fenzi Cuihua*, **2005**, *19*, 131-135.

[2] U. R. Bapat, J. P. Potams, N. Subramanian, J. Valgeirsson, *PCT Int. Appl.* **2008**, 2008152514.

[3] H. Han, J. Jin, J. Liu, *Shenyang Yaoke Daxue Xuebao* **2001**, *18*, 20-22.

[4] W. C. Shyu, C. A. Knupp, K. A. Pittman, L. Dunkle, R. H. Barbhaiya, *Clin. Pharmacol. Ther.* **1991**, *50*, 503-507.

[5] G. A. Jeffrey, W. Saenger, Hydrogen Bonding in Biological Structure, Springer, Berlin, **1991**.

[6] P. Ball, H2O: A Biography of Water, Weidenfeld & Nicolson, London, **1999**.

[7] M. Yves, The Hydrogen Bond and the Water Molecule, Elsevier, New York, **2007**.

[8] P. L. Polavarapu, E. Santoro, *Nat. Prod. Rep.* **2020**, *37*, 1661-1699.

[9] D. Kurouski, *Anal. Chim. Acta*, **2017**, *990*, 54-66.

[10] L. D. Barron, *Molecular light scattering and optical activity, 2nd edn. Cambridge*, **2004**.

[11] M. Losada, P. Nguyen, Y. Xu, *J. Phys. Chem. A* **2008**, *112*, 5621-5627.

[12] T. Kuppens, W. Herrebout, B. Veken, P. Bultinck, *J. Phys. Chem. A* **2006,** *110,* 10191-10200.

[13] G. Yang, Y. Xu, *Phys. Chem. Chem. Phys.* **2008**, *10*, 6787-6795.

[14] C. Merten, C. J. Berger, R. McDonald, Y. Xu, *Angew. Chem. Int. Ed.* **2014**, *53*, 9940-9943; *Angew. Chem.* **2014**, *126*, 10098-10101.

[15] L. Weirich, K. Blanke, C. Merten, *Phys. Chem. Chem. Phys.* **2020**, *22*, 12515-12523.

[16] F. Xie, X. Ng, N. A. Selfert, J. Thomas, W. Jäger, Y. Xu, *J. Chem. Phys.* **2018**, *149*, 224306.

[17] F. Xie, N. A. Seifert, W. Jäger, Y. Xu, *Angew. Chem. Int. Ed.* **2020**, *59*, 15703-15710; *Angew. Chem* **2020**, *132*, 15833-15840.

[18] J. Thomas, A. Serrato III, W. Lin, W. Jäger, Y. Xu, *Eur. J. Chem.* **2014**, *20*, 6148.

[19] J. Thomas, M. J. Carrillo, A. Serato III, E. G. Schnitzler, W. Jäger, Y. Xu, W. Lin, *Chem. Phys. Lett.* **2016**, *661*, 210-214.

[20] M. Shome, N. Mishra, *Indian J. Adv. Chem. Sci.* **2016**, *4*, 56-67.

[21] Y. Yang, J. Cheramy, Y. Xu, *ChemPhysChem*, **2021**, *22*, 1336-1343.

[22] F. Xie, S. Mahendiran, N. A. Seifert, Y. Xu, *Phys. Chem. Chem. Phys.* **2021**, *23*, 3820-3825.

[23] F. Xie, N. A. Seifert, A. S. Hazrah, W. Jäger, Y. Xu, *ChemPhysChem* **2021**, *22*, 455-460.

[24] A. S. Perera, J. Thomas, M. R. Poopari, Y. Xu, *Front. Chem.* **2016**, *4*, 1-17.

[25] M. Losada, H. Tran, Y. Xu, *J. Chem. Phys.* **2008**, *128*, 014508.

[26] G. Yang, Y. Xu, *J. Chem. Phys.* **2009**, *130*, 164506.

[27] A. S. Perera, J. Cheramy, C. Merten, J. Thomas, Y. Xu, *ChemPhysChem* **2018**, *19*, 2234-2242

[28] M. R. Poopari, Z. Dezhahang, G. Yang, Y. Xu, *ChemPhysChem* **2012**, *13*, 2310-2321.

[29] M. R. Poopari, P. Zhu, Z. Dezhahang, Y. Xu, *J. Chem. Phys.* **2012**, *137*, 194308.

[30] S. Ghidinelli, S. Abbate, J. Koshoubu, Y. Araki, T. Wada, G. Longhi, *J. Phys. Chem B* **2020**, *124*,

4512-4526.

[31] K. Le Barbu-Debus, J. Bowles, S. Jähnigen, C. Clavaguéra, F. Calvo, R. Vuilleumier, A. Zehnacker, *Phys. Chem. Chem. Phys.* **2020**, *22*, 26047-26068.

[32] (a) P. Pracht, F. Bohle, S. Grimme, *Phys. Chem. Chem. Phys.*, **2020**, *22*, 7169-7192; (b) C. Bannwarth, E. Caldeweyher, S. Ehlert, A. Hansen, P. Pracht, J. Seibert, S. Spicher, S. Grimme, *WIREs Comput. Mol. Sci.* **2020**, *11*, e1493.

[33] B. Wu, N. A. Seifert, S. Oswald, W. Jäger, Y. Xu, *ChemPhysChem* **2022**, doi.org/10.1002/cphc.202200348.

[34] C. D. Carlson, A. S. Hazrah, D. Mason, Q. Yang, N. A. Seifert, Y. Xu, *J. Phys. Chem. A*, **2022**, jp-2022-058039 (submitted).

[35] C. Bannwarth, S. Ehlert, S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 1652-1671.

[36] S. Grimme, C. Bannwarth, P. Shushkov, *J. Chem. Theory Comput.* **2017**, *13*, 1989-2009.

[37] S. Grimme, C. Bannwarth, S. Dohm, A. Hansen, J. Pisarek, P. Pracht, J. Seibert, F. Neese, *Angew. Chem. Int. Ed.* **2017**, *56*, 14763-14769; *Angew. Chem.* **2017**, *129*, 14958-14964.

[38] S. Grimme, *J. Chem. Theory Comput*. **2019**, *15*, 2847-2862.

[39] F. Xie, N. A. Seifert, M. Heger, J. Thomas, W. Jäger, Y. Xu, *Phys. Chem. Chem. Phys.* **2019**, *21*, 15408-15416.

[40] (a) Q. Yang, C. D. Carlson, W. Jäger, Y. Xu, *J. Phys. Chem.* A **2022**, *126*, 2942-2949; (b) A. Insausti, J. Ma, Q. Yang, F. Xie, Y. Xu, *ChemPhysChem.* **2022**, doi.org/10.1002/cphc.202200176.

[41] N. A. Seifert, J. Thomas, W. Jäger, Y. Xu. *Phys. Chem. Chem. Phys.* **2018**, *20*, 27630-27637.

[42] F. Xie, M. Fusè, A. S. Hazrah, W. Jäger, V. Barone, Y. Xu, *Angew. Chem. Int. Ed.* **2020**, *59*, 22427-22430; *Angew. Chem.* **2020,** *132*, 22613-22616.

[43] Gaussian 16, Revision B.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, **2016**.

[44] A. D. Becke, *J. Chem. Phys*. **1993**, *98*, 5648.

[45] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.* **2011**, *32*, 1456.

[46] A. D. Becke, E. R. Johnson, *J. Chem. Phys.* **2005**, *123*, 154101.

[47] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.

[48] A. Schäfer, C. Huber, R. Ahlrichs, *J. Chem. Phys.* **1994**, *100*, 5829.

[49] J. Tomasi, B. Mennucci, R. Cammi, *Chem. Rev.* **2005**, *105*, 2999-3094.

[50] M. A. J. Koenis, Y. Xia, S. R. Domingos, L. Visscher, W. J. Buma, V. P. Nicu, *Chem. Sci.* **2019**, *10*, 7680-7689.

[51] K Bünnemann, C. H. Pollok, C. Merten, *J. Phys. Chem. B* **2018**, *122*, 8056-8064.

[52] S. T. Mutter, F. Zielinski, J. R. Cheeseman, C. Johannessen, P. L. A. Popelier, E. W. Blanch, *Phys. Chem. Chem. Phys.* **2015**, *17*, 6016-6027.

[53] Z. Su, Q. Wen, Y. Xu, *J. Am. Chem. Soc.* **2006**, *128*, 6755-6760.

[54] J. Thomas, O. Sukhorukov, W. Jäger, Y. Xu, *Angew. Chem. Int. Ed.* **2014**, *53*, 1156-1159; *Angew. Chem.* **2014**, *126*, 1175-1178.

[55] J. T. A. Gall, J. Thomas, F. Xie, Z. Wang, W. Jäger, Y. Xu, *Phys. Chem. Chem. Phys.* **2017**, *19*, 29508-29515.

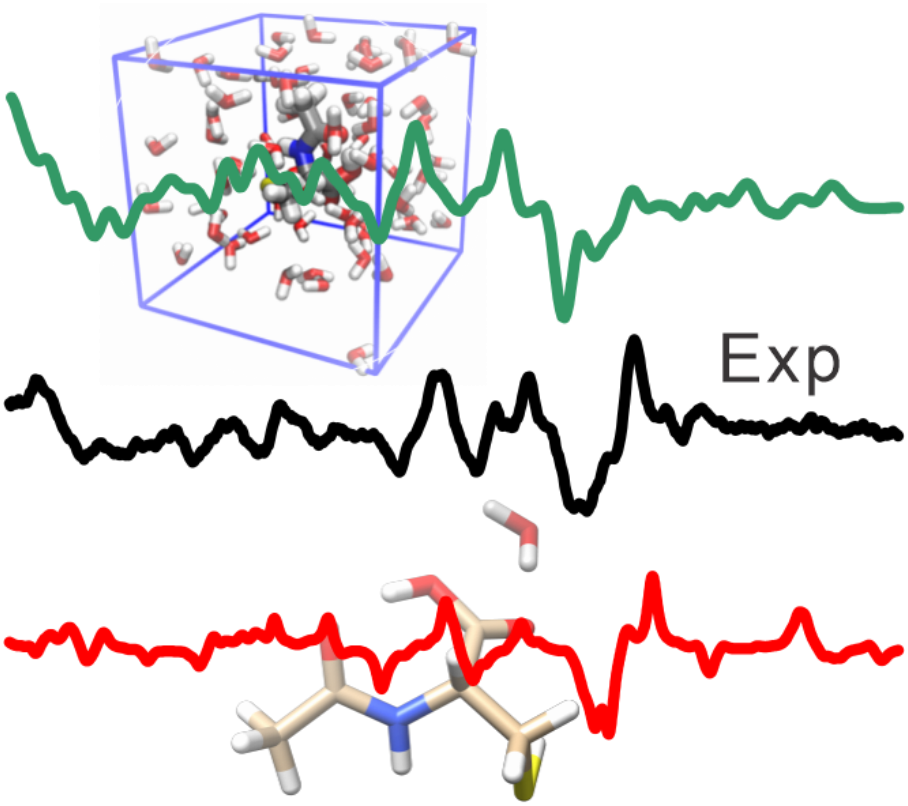
[56] B. Wu, A. S. Hazrah, N. A. Seifert, S. Oswald, W. Jäger, Y. Xu, *J. Phys. Chem. A*, **2021**, *125*, 10401-10409.

[57] W. Huang, J. Thomas, W. Jäger, Y. Xu, *Phys. Chem. Chem. Phys.* **2017**, *19*, 12221-12228.

[58] R. B. Mackenzie, C. T. Dewberry, R. D. Cornelius, C. J. Smith, K. R. Leopold, *J. Phys. Chem. A* **2017**, *121*, 855-860.

Chapter 5

Raman Optical Activity of N-Acetyl-L-Cysteine in Water and in Methanol: the "Clusters-in-a-Liquid" Model and ab initio Molecular Dynamics Simulations



This chapter is directly copied and adapted the new format from the following publication:

Y. Yang, J. Cheramy, M. Brehm, Y. Xu, *ChemPhysChem* **2022**, e202200161.

5.1 Introduction

For flexible chiral molecules, interactions with solvent molecules, such as water, can alter their conformational landscape significantly.[1] Such changes in conformational distribution have been extensively studied in recent years by using rotational spectroscopy where minor structural changes can be easily detected and recognized for example, in the case of methyl lactate versus methyl lactate-water clusters[2,3,4] and tetrahydro-2-furoic acid versus its monohydrates,[5,6] thanks to the high resolution capability of the rotational spectroscopic techniques used. To extract similar structural information from solution vibrational spectroscopic measurements is generally much more challenging because of the co-existence of many species in solution simultaneously and possible interconversion among them. Since biological environments are predominantly aqueous, it is of particular importance that one can model chiral vibrational measurements in water accurately. This is a crucial step in reliably extracting structural information including absolute configuration and conformation of chiral molecules, such as chiral drugs.[7] antibiotic,[8] cyclic peptides,[9] and amino acids and their derivatives[10,11] from chiral vibrational data.

In general, protic solvents, such as water and methanol, may interact significantly with chiral soluble molecules through medium to strong hydrogen bonding interactions which may lead to noticeable structural changes of the solute molecules[8] and even chirality transfer to surrounding water molecules.[12,13] To extract structural properties from chiral vibrational measurements, the accuracy of spectral simulations is, not surprisingly, a limiting factor. There have been considerable research efforts in this direction in the last few years and some were recently reviewed by Zehnacker and co-workers.[14] Several well-known challenges are related to the number of solvent molecules needed to be included and their locations/orientations.[15] To capture the sheer number of potentially important chiral molecule-solvent complexes, one common approach is to apply molecular dynamic (MD) simulations.[16-20] Typically, tens to several thousand snapshots of chiral molecule-

solvent complexes which contain the chiral solute and a few solvent molecules, up to two (or a few) layers, were taken, and were reoptimized at the DFT level with the inclusion of the polarizable continuum model (PCM) of the solvent. For example, in a recent study, the chiral solute was treated at the quantum mechanical level while the solvent layer with a radius of 6 Å from the solute was considered at the molecular mechanical level to evaluate vibrational optical activity of pantolactone.[19]

A somewhat simplified approach, i.e. the "clusters-in-a-liquid" solvation model,[21] emphasizes the long-lived solute-solvent clusters[22-24] and embeds these clusters in a solvation continuum. The induced solvent chirality features have provided the experimental evidence of the existence of long-lived solute-solvent clusters, and thus the foundation of the "clusters-in-a-liquid" model. Overall, all these approaches apply the static harmonic calculations in which harmonic frequencies are obtained from the Hessian matrix of a minimum energy structure, and in recent years anharmonic calculations of chiral vibrational spectra have also become possible.[25] While this approach works reasonably well in most cases, some shortcomings of the static calculations have been noted, in describing strong solute–solvent interactions associated with ionic liquids[26] and in describing the solvent dynamics.[14]

Ab initio molecular dynamics (AIMD) simulations[27] have been utilized for generation of infrared[28] and Raman[29] spectra for the last two decades. This protocol has several advantages, such as being directly applicable to condensed phase systems, covering anharmonic effects to a certain extent, and reproducing line shapes, overtones, and combination bands in qualitatively correct manner.[30] A noticeable shortcoming is the considerable computational overhead associated with the use of the Wannier localization.[31,32] This limitation was alleviated somewhat by using the Voronoi integration approach to obtain the electromagnetic moments via integration of the total electron density[33,34] so that a localization of molecular orbitals is no longer required. In terms of chiral vibrational spectroscopy, the first AIMD liquid vibrational circular dichroism (VCD)

spectrum was reported in 2016 by Thomas et al.,[35] based on a classical formalism for the magnetic moments. Later, Scherrer et al.[36] reported another AIMD liquid VCD spectrum, based on nuclear velocity perturbation theory (NVPT). Both approaches utilized the Fourier transform of time correlation functions along the trajectory to generate the VCD spectra, as suggested by Abbate in 1998.[37] The first AIMD Raman optical activity (ROA) spectrum was reported for propylene oxide, a rigid chiral molecule with a single conformation, in the gas phase in 2017.[38] In this work, some derivations were explicitly conducted under the assumption of a non-periodic system. Based on these derivations, Luber concluded that the application to periodic bulk phase systems would remain as an open project for the future.[38] Shortly after, the first AIMD liquid phase ROA spectrum of propylene oxide was predicted by Brehm et al.[39] The AIMD ROA method by Brehm and co-workers has since been implemented in the TRAVIS program package[40,41] and is now freely available to the scientific community.

In the current study, we focus on Raman and ROA of N-acetyl-L-cysteine (NALC), a conformationally flexible amino acid derivative,[42,43] in water and in methanol. With multiple hydrogen bonding sites to which protic solvent molecules can attach, NALC is a good model system for developing methods to account for induced conformational changes in solutions. ROA measures 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 conformations. The current commercial ROA instrument can measure vibrational transitions in a wide spectral range, routinely all the way down to ~ 200 cm^{-1}, allowing access to low frequency vibrational modes which may be coupled to low frequency water solvent bands.[44] Two different protocols will be utilized to simulate Raman and ROA spectra of NALC in water and in methanol. First, we will couple the "clusters-in-a-liquid" model with a recently developed conformational searching tool, CREST,[45] by Grimme and co-workers. Second, we will apply the AIMD approach described above.[39] To the best of our knowledge, in previous applications of the AIMD

ROA approach, the method was applied only to a rigid chiral molecule, propylene oxide. It would be very interesting to see how well this AIMD approach works for a flexible molecule like NALC, and this would only be the second application example of the AIMD liquid ROA approach. The simulated results from these protocols will be compared with the experimental data and with each other. The pros and cons of the protocols will be discussed.

5.2 Results and Discussion

5.2.1 Experimental Raman and ROA spectra of NALC in water and in methanol

The collected experimental Raman and ROA spectra for NALC in water and in methanol are shown in Figure 5.1. The two sets of data are for the most part nearly the same, save for the regions in which methanol produces very strong Raman scattering. Because strong Raman scattering of methanol can easily cause a saturation of the CCD detector implemented in the ROA instrument, the NALC in methanol solution was measured at a much lower incident laser power and for a much longer amount time than for the NALC in water solution. Despite this issue, the ROA spectra in the 200-1900 cm^{-1} region appear to be essentially the same between these two solutions after solvent subtraction. The raw spectra are shown in Figure C1, Appendix C. In the methanol solution, the strong Raman bands at ~1000 and 1460 cm^{-1} belong to the methanol solvent, and the solvent subtraction was not perfect because of the strong band intensity in these two regions. Overall, the similarity of the ROA spectra in these two solvents suggests that the solute-solvent interactions are alike for both solvents, even though water has a much stronger tendency to form extensive hydrogen bonded network in comparison to methanol. At the same time, water or methanol may cause noticeable conformational changes in the NALC molecule itself and certain NALC-solvent complexes may exist as long-lived species.

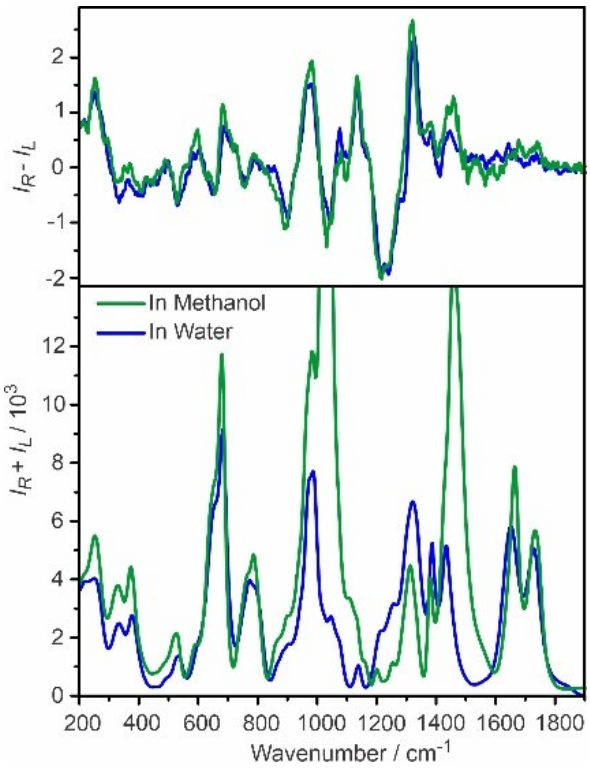


Figure 5.1. Experimental Raman (bottom) and ROA (top) spectra of the NALC molecule dissolved in water (blue) and in methanol (green). The very strong bands at 1000 and 1460 cm^{-1} are contaminated with the pure methanol solvent bands. See the main text for discussion.

To prevent interferences due to the addition of a buffer solution and/or acid and bases, none of them were added to the samples. The resultant pH of the NALC in water solution was measured to be about 2.5. Based on the pKa value of NALC and the simple relationship: pH= pKa+log{[A-]/[HA]}, one may estimate that the relative abundance of the neutral (COOH) relative to the deprotonated (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,[42] suggesting that the abundance of COO^- is considerably lower than the simple estimation implies. For this reason, in the current study, only the neutral form is modelled herein.

5.2.2 The NALC conformers and the simulations of Raman and ROA in water and in methanol

The CREST and subsequent DFT calculations generated ten low energy NALC conformers, five more than previously reported.[42] Among the ten geometries found, the six relevant ones (vide infra) are shown in Figure 5.2, while the others are provided in Figure C2, Appendix C. A closer look at the geometries in Figure 5.2 indicates that the most stable conformer takes on the *trans*-COOH configuration where the C=O and OH groups point in opposite directions, while the most stable *cis*-COOH conformer is ~3.4 kJ mol^{-1} less stable than the global minimum. The *trans*-COOH configuration allows the OH of the carboxylic group to form a strong intramolecular hydrogen bond with the O atom of the carboxamide group. A similar reason leads to the strong preference of the *trans*-COOH configuration in tetrahydro-2-furoic acid[5] and the detection of dimers which are made of *trans*-COOH in a supersonic jet expansion[56] and in a cold rare gas matrix.[46] A quantum theory of atoms in molecules (QTAIM) analysis shows that such an intramolecular hydrogen bond energy is about 40 kJ mol^{-1}, indicating a strong attractive force between the ether oxygen and OH on the carboxylic acid group. The same QTAIM analysis shows no specific bond critical point between the carboxylic acid oxygen and OH group for the *cis*-COOH configuration, further supporting the observation that the *trans*-COOH configuration is strongly preferred.

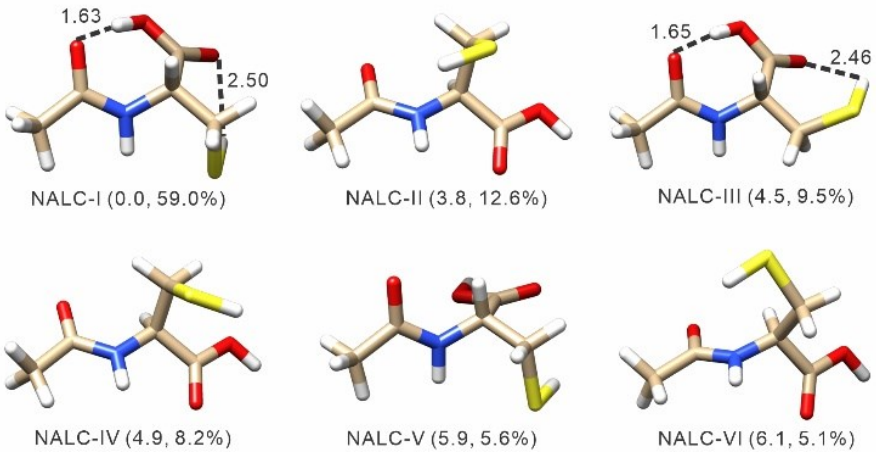


Figure 5.2. The six relevant, low energy NALC conformers optimized at the B3LYP-D3BJ/def2-TZVP level of theory with the PCM of water. Their relative free energies in kJ mol^{-1} at 298 K and also the calculated percentage Boltzmann abundances are listed in the brackets.

First, we applied the implicit solvation model to see how well this simple approach can reproduce the experimental Raman and ROA spectra. The individual Raman and ROA spectra of the six NALC conformers are shown in Figure 5.3, as well as the Boltzmann weighted spectra based on the relative free energies and the experimental spectra obtained in water. A further examination of the individual Raman and ROA spectra (Figure 5.3) indicates that the global minimum geometry appears to capture the main experimental features. This suggests that NALC-I is indeed the most dominant NALC monomeric conformation in water at room temperature. This fact is roughly in line with the prediction that NALC-I makes up about 60% of the total monomeric population. The inclusion of the other four high energy of conformers (Figure C2, Appendix C) worsens the agreement with the experimental data and are therefore not considered in the remainder of the paper. Overall, the Boltzmann averaged Raman and ROA simulations of the NALC monomers capture most of experimental features observed in the region between 700-1400 cm^{-1}. On the other hand, with only the implicit solvation model, some of the experimental features are still not accounted for. For example, in the 200-700 cm^{-1} region, the agreement between experiment and theory is still somewhat poor.

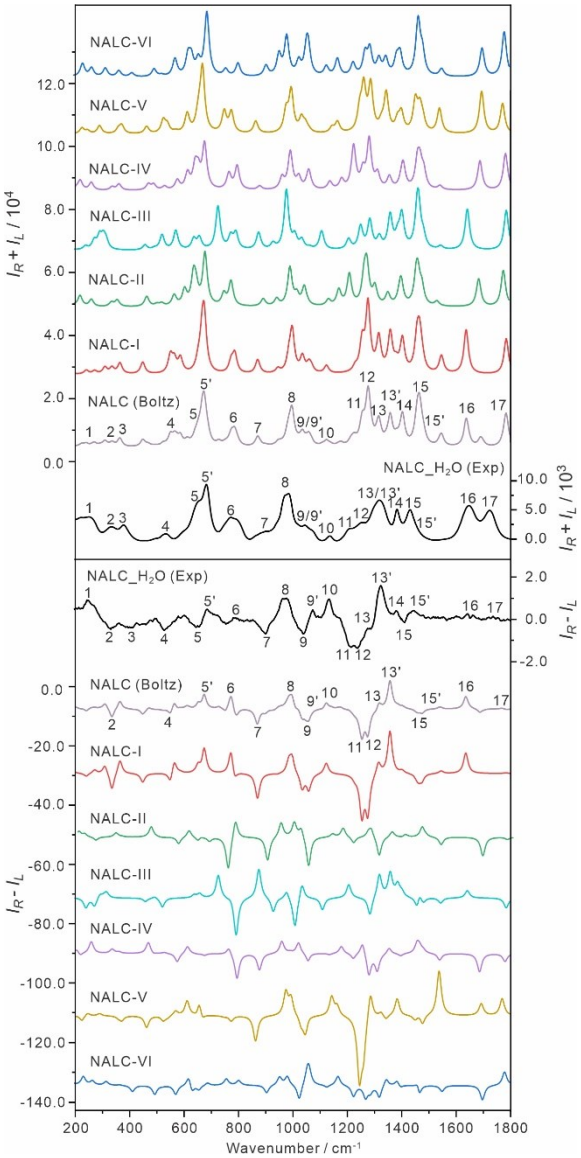


Figure 5.3. Individual Raman (top) and ROA (bottom) spectra of the six NALC conformers shown in Figure 5.2 at the B3LYP-D3BJ/def2-TZVP level with the PCM of water. The Boltzmann population weighted Raman and ROA spectra are also included, as well as the experimental data for comparison. The main features are labelled as "1-17" to guide comparisons. Note that not all experimental features have the corresponding simulated ones at this level.

The corresponding results for the simulated Raman and ROA spectra of the six conformers in methanol are provided in Figure C3, Appendix C. Not surprisingly, the simulated results obtained are very similar to those in water, and to the actual experimental

spectra. Similar deficiency was noted for the lower wavenumber region. Therefore, in the next step, we considered the explicit solute-solvent interactions, in addition to the implicit solvation modeling, in an effort to better capture the solvent effects and improve the agreement between the theory and the experimental Raman and ROA spectra.

5.2.3 Explicit solvation of NALC and the simulated Raman and ROA spectra in water

While the "clusters-in-a-liquid" solvation model[21] emphasizes the long-lived species, i.e. the dominant solute-solvent complexes in solution, there is no systematic guidance in this model on how to identify such long-lived species. For solvents, such as methanol and dimethyl sulfoxide, it was generally good enough to add one solvent molecule at one specific binding site.[47-49] The situation with water is considerably more complicated because water tends to form extensive hydrogen bonding network in bulk.[23,50] How to select the long-lived chiral solute-water$_n$ complexes becomes a challenging issue. In some of the previous VCD studies, some of us used the induced water VCD features at its bending region as the key features to guide the search of chiral solute-water$_n$ complexes.[21] In others, radial distribution functions from MD simulations and other approaches were used to estimate the number of closest water molecules at a specific binding site of the chiral solute to construct many chiral solute-water$_n$ complexes as the long-lived species.[18,51] It was recognized in some studies that the hydrated clusters, which contain water hydrogen bonded networks, tend to produce very large, multi-signated, induced water VCD features at the water bending region.[24] Such strong characteristic VCD signals were rarely detected experimentally, leading to doubt about the existence of the long-lived species with extensive water hydrogen bonded network. This is a point which we will address further later on. It appeared in these previous examples, the use of one water molecule at one specific binding site is enough. Since it is also less likely for all sites to have one water simultaneously, we therefore constructed 1:1 NALC-water complexes and used their weighted sum to account for the explicit water solvation effect.

For each NALC conformer, we applied the CREST code to search for possible low energy NALC monohydrate conformations under the PCM of water. The CREST candidates were then subjected to the fast optimization and single point calculation described in the Experimental section, and only the conformers which fall within an energy window of 15 kJ mol^{-1} relative to the global minimum were selected. The final optimization of the NALC monohydrates were done at B3LYP-D3BJ/def2-TZVP/ PCM(water) level. The relative free energies and Boltzmann population factors of all the NALC monohydrates are listed in Table 5.1. The NALC-water complexes obtained are named based on the conformational of the chiral solute, i.e. NALC-I to -VI, the number of water molecules, for example, -1w, and the relative free energy ordering, with 1 being the most stable one. Please note that the energy ordering is done within the family of each NALC conformer, i.e. within NALC-I-1w, NALC-II-1w, etc. For example, NALC-I-1w-2 means that the system contains NALC-I with 1 explicit water and the relative energy ordering is 2 compared to the most stable NALC I-1w-1.

Table 5.1. Relative free energy, ΔG (in kJ mol^{-1}) and Boltzmann population factor, B_f (in %) for each conformer in water, explicitly solvated by one water molecule, computed at B3LYP-D3BJ/def2-TZVP level with the PCM of water.

Conformer	ΔG	B_f	Conformer	ΔG	B_f
NALC-I-1w-1	0.0	32.8	NALC-III-1w-2	3.9	17.1
NALC-I-1w-2	0.6	25.6	NALC-III-1w-3	10.6	1.1
NALC-I-1w-3	1.1	21.4	NALC-III-1w-4	13.9	0.4
NALC-I-1w-4	4.8	4.7	NALC-IV-1w-1	0.0	95.5
NALC-I-1w-5	5.1	4.2	NALC-IV-1w-2	9.3	2.3
NALC-I-1w-6	5.2	3.9	NALC-IV-1w-3	10.0	1.7
NALC-I-1w-7	5.3	3.9	NALC-IV-1w-4	12.8	0.5
NALC-I-1w-8	6.6	3.0	NALC-V-1w-1	0.0	48.3

NALC-II-1w-1	0.0	52.5	NALC-V-1w-2	0.7	36.2
NALC-II-1w-2	0.8	38.4	NALC-V-1w-3	3.0	14.5
NALC-II-1w-3	4.9	7.2	NALC-V-1w-4	9.8	1.0
NALC-II-1w-4	8.8	1.5	NALC-VI-1w-1	0.0	57.8
NALC-II-1w-5	12.0	0.4	NALC-VI-1w-2	0.8	41.4
NALC-III-1w-1	0.0	81.4	NALC-VI-1w-3	10.4	0.8

The Boltzmann averaged Raman and ROA spectra of the NALC monohydrates of each NALC conformer are shown in Figure C4, Appendix C. In Figure 5.4, the Boltzmann averaged Raman and ROA spectra of NALC-I-1w are compared with the experimental data, and reproduce the major experimental features, indicating that this species is a dominant one in aqueous solution. To include the contributions of the monohydrates of NALC-II to VI, we empirically adjusted the weights of each set of the NALC monohydrates to best reproduce the observed spectra. The final empirically fitted Raman and ROA spectra of all monohydrates are also given in Figure 5.4. The empirically adjusted spectra consist of 60% NALC-I-1w, 10% NALC-III-1w, 10% NALC-IV-1w and 20% NALC-VI-1w.

The two strong bands in the 1600-1800 cm^{-1} region are assigned to the C=O stretching of the carbonyl group (lower cm^{-1}) and the C=O stretching of the carboxyl group (higher cm^{-1}). In the 800~1400 cm^{-1} region, the vibration bands predominantly correspond to those associated with various bending motions such as the S-H bend, C-H bend, CH_2 bend, CH_3 and N-H bend modes of NALC. While in the below ~800 cm^{-1} region, there are mostly molecular breathing vibrations and solvent O-H bending modes where these bands tend to overlap considerably. In some of the previous solvation studies, the vibrational modes of the explicit solvent molecules were frozen when simulating infrared and VCD spectra of the solute-solvent complexes, leading to solute only vibrational bands.[18] Here, the vibrations of the solvent molecules are considered at the same DFT level of theory as for the NALC itself. This is important because the experimental results appear to show evidence of contributions from these explicit solvent molecules.

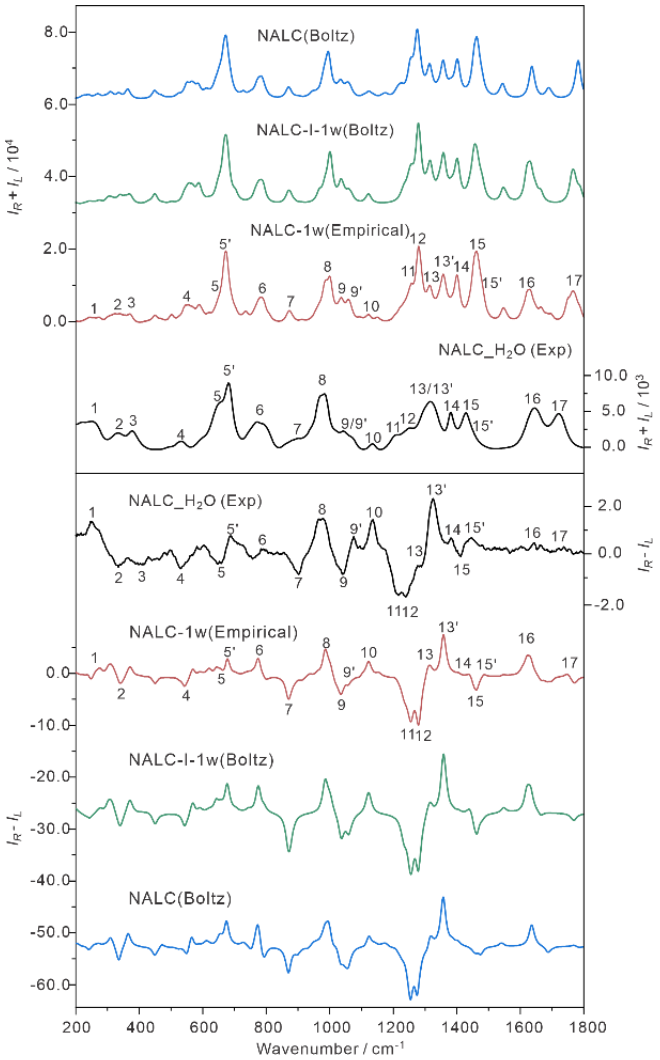


Figure 5.4. Comparisons of the experimental Raman and ROA spectra of NALC in water with three different simulations: the empirically weighted Raman and ROA spectra of all monohydrate conformers of NALC-I to -VI, the Boltzmann averaged Raman and ROA spectra of monohydrate conformers of NALC-I, and the Boltzmann averaged Raman and ROA spectra of the monomeric conformers at the B3LYP-D3BJ/def2-TZVP level of theory with the PCM of water. The main features are labelled as "1-17" to guide comparisons.

Overall, the simulated monohydrate spectra provide better agreement with the experimental spectra than those based on the monomeric NALC only. For example, the experimental ROA features in the region bracketed by the bands labelled as 6 and 7 are

well captured by the empirically weighted monohydrate ROA simulation, and less well by the implicit solvation only simulation. In this region, the bands are mainly assigned to the CH_2 wagging, COO bending, C-H bending, S-H bending of the main solute molecule, with small contribution from the O-H bending of the water molecule. In the ROA spectrum, the distinctive "+/-/+" features for the bands 14, 15, and 15' were well captured when we treat water explicitly, which are a mixture of the N-H bend, O-H bend, C-H bend and CH_3 bend motions. Moreover, in the carbonyl region, broader experimental Raman features were observed in the band 16-17 region, agreeing better with the weighted monohydrate ROA simulation than with the implicit model. These observations indicate that it is not enough to consider only how water induces small conformational changes in the solute molecule. Rather, it is necessary to capture the long-lived NALC-water complexes which contribute as a whole to the measured Raman and ROA spectra. While the simulated monohydrate spectra provide the best agreement with the experimental data so far, some deficiencies still exist, especially in the simulated ROA spectra in the region below 400 cm^{-1}.

As a next step, additional explicit water molecules are further added to make larger NALC-$water_n$ complexes with n =2, 3, 4, 5, and 6 in a stepwise manner. We utilized the same three tier approach in building these large water solvated clusters. Since NALC-I is the dominant NALC conformation based on the discussions above, we treated this conformer first. The Boltzmann weighted Raman and ROA spectra of NALC-I-1w, -2w, to -6w are summarized in Figure C5, Appendix C, together with the experimental spectra. Similarly, n=2, 3, 4, 5, and 6 water molecules were added to NALC-II, -IV, and -VI. The Boltzmann weighted Raman and ROA spectra of NALC-VI-2w, to -6w are given in Figure C6, Appendix C. As one can see, the ROA spectra exhibit some very strong bi- or multi-signated ROA features in the 1600-1800 cm^{-1} region, which are clearly not observed experimentally. Overall, these simulations indicate that these larger water-solvated clusters are unlikely to play a dominant role as the long-lived species, similar conclusions were made in some previous studies.[24]

5.2.4 Explicit solvation of NALC and the simulated Raman and ROA spectra in methanol

The same procedure described above for water was then applied to investigate the solvent effects for NALC in methanol. The relative free energies and Boltzmann population factors of NALC-I to NALC-VI with the inclusion of PCM of methanol are listed in Table C1, Appendix C. Not surprisingly, the relative free energy ordering is almost the same as those with the PCM of water. NALC-I is still the most stable one with the *trans*-COOH configuration, whereas NALC-VI is the least stable one. Interestingly, the stability order predicted switches for NALC-III and NALC-IV from the PCM of water to methanol. Please note that we keep the -I to -VI naming the same as in the water case to avoid confusion. The individual Raman and ROA spectra (not shown) of the NALC conformers show only very minor changes from those simulated with the PCM of water. In addition, the predicted geometries of NALC-I to -VI do not change much with the PCM of methanol instead of water. All these observations are consistent with the fact that the experimental Raman and especially ROA are very similar in both water and methanol.

We then explicitly added one methanol molecule to the NALC conformers. All the conformational searches and geometry optimization were carried out in a similar fashion as described for water. The relative free energy and Boltzmann population factors of all the 1:1 NALC-methanol binary conformers are listed in Table C2, Appendix C. Their simulated Boltzmann averaged Raman and ROA spectra of NALC-I-1m to NALC-VI-1m (m = methanol) conformers are shown in Figure C7, Appendix C. In general, visual inspection of the Boltzmann averaged Raman and ROA spectra with the consideration of either implicit or explicit solvent effects agree well with experiment (Figure 5.5). In line with the NALC in water case, NALC-I-1m reproduces the majority of the experimental Raman and ROA spectral features, proving that NALC-I is the dominant conformation in methanol solution. The same procedure is carried out to investigate the contributions of all

six NALC conformers with one explicit methanol and the empirically weighted results are also given in Figure 5.5.

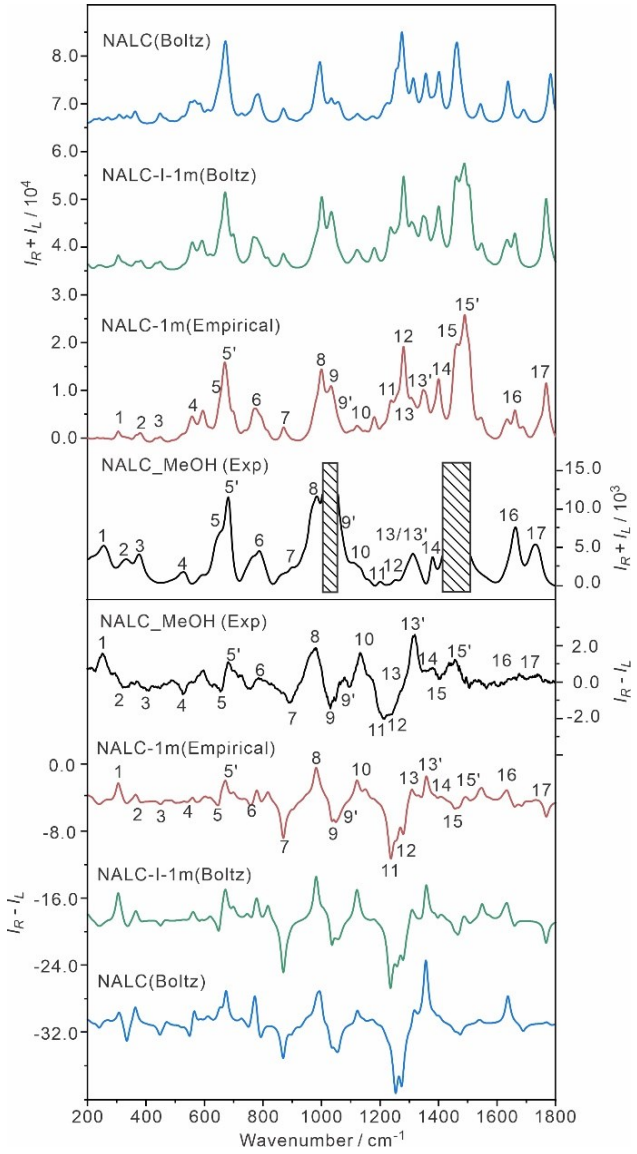


Figure 5.5. Comparisons of the experimental Raman and ROA spectra of NALC in methanol with the Boltzmann averaged Raman and ROA spectra of the monomeric conformers at the B3LYP-D3BJ/def2-TZVP level with the PCM of methanol. The empirically weighted Raman and ROA spectra of NALC-1m clusters and the Boltzmann averaged Raman and ROA spectra of NALC-I-1m clusters at B3LYP-D3BJ/def2-TZVP level of theory with the PCM of methanol are also provided. The main band features are numbered to guide

comparisons. The strong experimental Raman bands at 1020 cm^{-1} and 1460 cm^{-1} have severe solvent interference and are blocked out for clarity.

The empirically weighted Raman and ROA spectra were made of 70% NALC-I-1m, 5% NALC-II-1m, 5% NALC-III-1m, 5% NALC-IV-1m, 10% NALC-V-1m, and 5% NALC-VI-1m. Some promising improvements are achieved when compared to the simulation using only the implicit solvation model. For example, the experimental ROA bands 5 and 5' show a "-/+" spectral pattern, which are well captured when the explicit solvent effect is considered, while a "+/+" pattern was predicted in the implicit solvation model. Similarly, the experimental features in the 1410-1480 cm^{-1} (bands 14, 15 and 15') are better reproduced by the empirically weighted ROA spectrum of the NALC-methanol conformers than that of the NALC itself. These observations further prove the existence of long-lived solute-solvent clusters in experiment. However, the same deficiencies still exist in the simulated ROA spectra in the region below 400 cm^{-1}.

5.2.5 AIMD simulations of NALC in water and methanol

As mentioned above, while the static model provides good agreement with the experimental data, some deficiencies, especially in the simulated ROA spectra in the region below 400 cm^{-1}, still exist. These discrepancies may result from the solvent dynamics which are not taken into account in the current static approach based on the stable geometries of the solute-solvent clusters. In addition, it would be of considerable interest to apply the AIMD ROA approach developed to a flexible molecule, such as NALC. To compute the spectrum, a NALC molecule was placed in a cubic box of ~1300 pm edge length, together with 64 water molecules. A snapshot of the simulation cell is shown in Figure C8, Appendix C. All the simulation details are described in the Experimental Section.

One advantage of the AIMD-based approach to computing spectra when compared to the static–harmonic approach is the intrinsic conformer sampling due to the dynamics. We performed an AIMD simulation of NALC without any restraints (vide infra) a year ago and noticed that the molecule may change its conformation several times during the simulation. However, the NALC conformers in solution are relatively stable, so an AIMD simulation is significantly too short to sample all relevant conformers in one run. To overcome this, we decided to perform six independent AIMD simulations for the six conformers shown in Figure 5.2 in water (and similarly for the methanol case), and to predict vibrational spectra for each conformer separately. To avoid interconversion of the conformers during each AIMD simulation, harmonic restraints were used on relevant dihedral angles in order to keep the backbone geometry intact. First, the relevant dihedral angles which distinguish the conformers, referred to as ϕ_1 to ϕ_6, are identified and illustrated in Figure 5.6. The values of these dihedral angles for the six conformers NALC-I to NALC-VI are given in Table 5.2.

In the AIMD simulations for the six conformers, harmonic restraints with a spring constant of K = 0.005 Hartree rad^{-2} were placed on the six dihedral angles ϕ_1 to ϕ_6, using the values in Table 5.2 as equilibrium values. In contrast to constraints, restraints allow a certain motion of the restrained quantity around the equilibrium value but avoid large elongations. This keeps each of the six conformers in place without altering the backbone dynamics too much. The concept of restraining the backbone dihedral angles is also visualized in Figure 5.6, where an overlay of the conformer NALC-I over the whole AIMD production run is shown. It can be seen that despite the restraints, the backbone is not completely rigid and retains a certain degree of freedom, but the overall shape of the conformer is maintained.

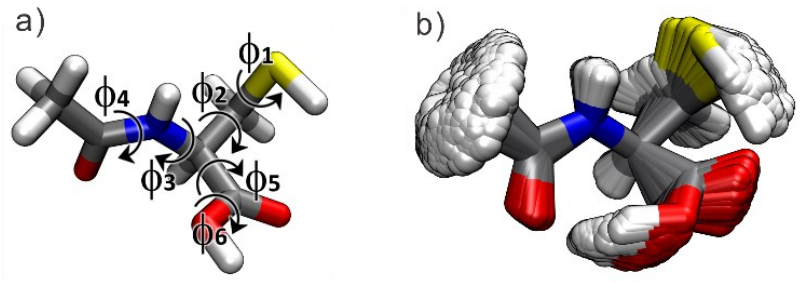


Figure 5.6. a) The six dihedral angles, ϕ_1 to ϕ_6, in NALC which were restrained to keep the conformers shown in Figure 5.2 intact. b) Overlay of the NALC structures along the trajectory for conformer NALC-I, which is kept intact by the dihedral angle restraints shown in a).

Table 5.2. The target values of the restraints on the dihedral angles ϕ_1 to ϕ_6 in NALC to keep the six conformers, NALC-I to NALC-VI, intact.

Dihedral	NALC -I	NALC -II	NALC -III	NALC -IV	NALC -V	NALC -VI
Φ_1	64.7	-64.9	-71.0	78.6	73.3	65.2
Φ_2	52.9	65.4	-166.6	62.6	59.1	-64.6
Φ_3	159.6	80.9	159.0	80.7	127.3	128.7
Φ_4	-172.9	178.3	-171.8	179.9	-179.4	179.2
Φ_5	-178.0	-58.3	-174.4	-58.6	150.0	-68.9
Φ_6	-1.7	179.4	-5.7	178.4	177.0	177.9

The AIMD Raman and ROA spectra of NALC-I to -VI in water are summarized in Figure C9, Appendix C. To compare the AIMD Raman and ROA spectra of NALC in water with the experimental ones, one needs to weigh individual conformers with their abundances. In the current study, we chose to use the empirically obtained conformational abundances and the reason is discussed above. The final predicted Raman and ROA spectra were created as a linear combination of the six conformer spectra with the empirically

determined coefficients (NALC-I : NALC-II : NALC-IV : NALC-V : NALC-VI = 0.21 : 0.17 : 0.10 : 0.05 : 0.47). The resulting predicted ROA spectrum of NALC in water is then compared with experiment in Figure 5.7. Apart from the strong positive peak at 1325 cm^{-1} which is not fully captured in the prediction, and apart from slightly wavenumber shifts in several bands, it can be seen that the simulated AIMD ROA spectrum predicts the experiment very well in terms of band pattern and also the relative intensity ratios. One particular improvement over the best static harmonic DFT simulated spectra is at the wavenumber region below 400 cm^{-1}. This is a remarkable achievement for such a complicated and flexible molecule.

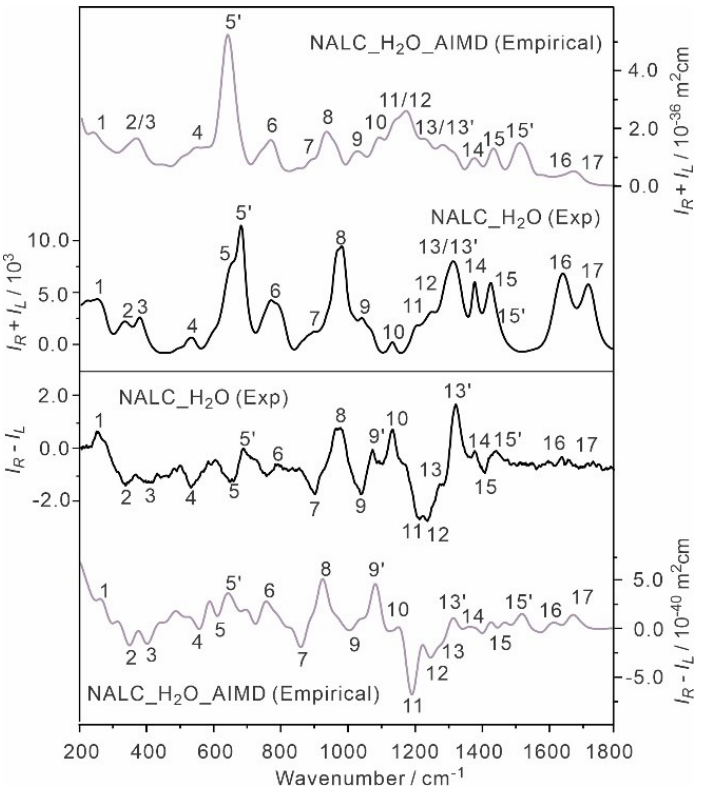


Figure 5.7. Comparison of the AIMD Raman and ROA spectra of NALC in water with the corresponding experimental spectra. The main band features are numbered to guide comparisons.

A similar simulation procedure was applied to generate the final AIMD Raman and ROA spectra in methanol. The AIMD Raman and ROA spectra of NALC-I to -VI in

methanol are summarized in Figure C10, Appendix C. The result is summarized in Figure 5.8. The empirical percentage abundances for NALC-I: NALC-II: NALC-VI = 0.42 : 0.16 : 0.42. One can see that the simulated Raman and ROA bands are in a good agreement with the experimentally observed band patterns and intensities. For example, the weak experimental ROA bands in the carbonyl region are well captured by the AIMD ROA spectrum, but not so well by the static DFT simulation. In addition, in the below 400 cm^{-1}, while some ROA band intensities were a bit too high in the AIMD spectrum, the shape patterns of band 1-3 are reasonably predicted. Similar to the case of NALC in water, the strong experimental intensity for the positive ROA peak at 1300 cm^{-1}, i.e., band 13', is somewhat low in the AIMD prediction.

Overall, it can be concluded that very similar NALC-solvent complexes exist in both water and methanol solutions because of their nearly the same experimental ROA spectra and also very similar simulated ROA features obtained by both the static harmonic and AIMD approaches. With regard to the conformational abundances, in a certain sense, one could argue that fixing a molecule's backbone geometry in an AIMD simulation by restraining certain dihedral angles is not a sensible way to go, as one loses one of the advantages of the AIMD approach—namely, the intrinsic conformer sampling. However, this conformer sampling is only effective if the energy barrier between all relevant conformers is sufficiently small. If this is not the case for NALC, the AIMD simulation time will be too short to obtain a meaningful average over the relevant conformers. As a result, the AIMD spectrum will heavily depend on the choice of the initial conformer, which should not be the case in a MD simulation with reasonable sampling. For these reasons, it is advantageous to consciously sacrifice the conformer sampling and run independent simulations for each conformer, and then apply the weighing procedure as we did for NALC in water and in methanol.

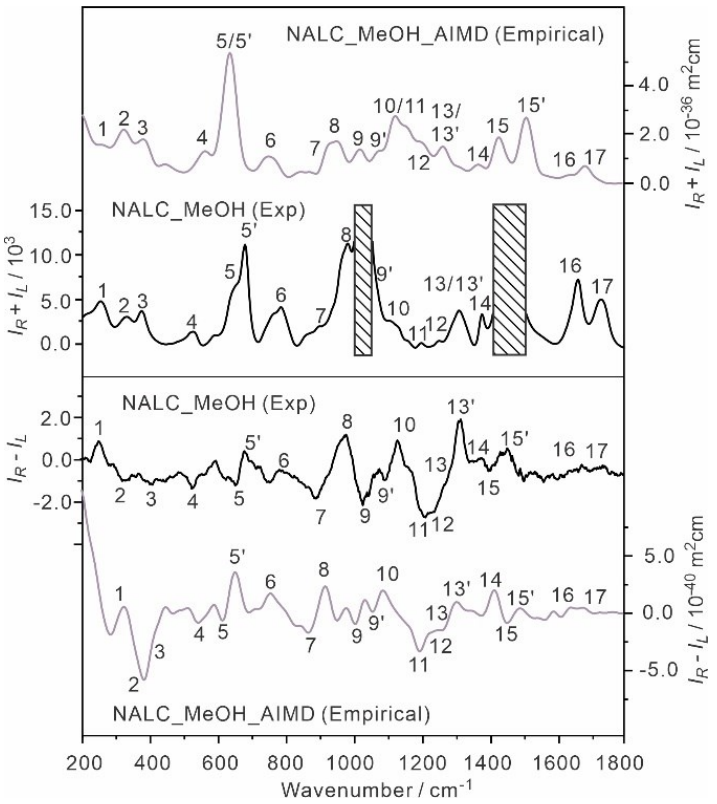


Figure 5.8. Comparison of the AIMD Raman and ROA spectra of NALC in methanol with the corresponding experimental spectra. The main band features are numbered to guide comparisons. The strong experimental Raman bands at 1020 cm^{-1} and 1460 cm^{-1} have severe solvent interference and are blocked out for clarity.

In this section, we make some comments about the pros and cons of the two different approaches used in the current study. Both the "clusters-in-a-liquid" model and AIMD approach provide simulated Raman and ROA spectra which are in reasonably good agreement with experimental ones in most of wavenumber regions. The "clusters-in-a-liquid" model captures the essence of the solute-solvent interactions. This static harmonic method, coupled with a fast and reliable conformational search tool, such as CREST, is easy to implement and can reproduce the main experimental ROA features well. In general, it is several orders of magnitude faster than the AIMD simulations because of the simplicity of the force calculations, which is more computationally economic and easier to investigate for a large and flexible system. Some noticeable improvements in the carbonyl region and

in the low wavenumber regions by the AIMD approach reveal the shortcoming of the static harmonic approach. Some of these particular regions mainly consist of vibrational modes of the explicit solvent molecules, highlighting the advantage of AIMD simulations in faithfully accounting for the explicit solute-solvent interactions. In addition, some main discrepancies in the AIMD Raman spectrum in water are in the 1100-1300 cm^{-1} region. This region mainly contains S-H bend, C-H bend, CH_2 bend, and N-H bend, which are associated with the backbone motions of NALC. While the harmonic restraints do allow a certain degree of freedom for the backbone movements, these restraints may be the source of some discrepancies for the final results.

The empirical conformational abundances estimated are somewhat different using these two different approaches. This is because each approach has its own shortcoming in capturing what is going on in solution. The “clusters-in-a-liquid” model is a considerably simplified model which does not take into account of the effects of many explicit solvent molecules except one in the current case. The AIMD simulations, on the other hand, take all these explicit solvent molecules into consideration. In principle, the "correct" abundances can be predicted from the AIMD simulations. However, this requires significantly longer simulations than we could afford here. It is our hope to demonstrate in future works how to compute the abundances directly from AIMD simulations.

In summary, it is now feasible to use AIMD simulation to compute Raman and especially ROA spectra for a larger and more complex molecule in the periodic bulk phase systems. This has not been demonstrated before to the best of our knowledge, and the current study extends the scope to the field of computational chiral vibrational spectroscopy.

5.3 Conclusions

In this study, the experimental Raman and ROA spectra were measured for NALC, a flexible chiral molecule, in water and in methanol. Very interestingly, the observed ROA

spectra are nearly identical, indicating that the chiral molecule-solvent interactions are essentially the same even though water is known to have a strong tendency to form extensive hydrogen bonded network. Both the "clusters-in-a-liquid" model and the AIMD approaches produced similar ROA simulations in these two solvents, consistent with the experimental observation.

To evaluate the main intermolecular interactions between NALC and solvent molecules, two very different solvation methods, were applied to simulate the corresponding Raman and ROA spectra. One of them is based on the "clusters-in-a-liquid" solvation model, which emphasizes the long-lived chiral molecule-solvent complexes, and combines the CREST conformational search strategy with static harmonic simulations. The other is the newly developed AIMD ROA approach which had only previously been applied to a rigid, single conformation, chiral molecule. NALC-I with the *trans*-COOH configuration was identified as a dominant conformation in both water and methanol. Some significant improvements were demonstrated by adding one solvent molecule, i.e., water or methanol, at the potential hydrogen binding sites of NALC, suggesting these 1:1 solute-solvent complexes are the long-lived solute-solvent clusters. The substantial efforts to include multiple water molecules simultaneously to NALC led to no improvement in the agreement between theory and experiment. This observation indicates that larger water solvated NALC clusters are not dominant species. The AIMD Raman and ROA spectra produce considerable improvements in the ROA spectra in the low wavenumber region and in the carbonyl stretch region over the static harmonic approach. In comparison to the static harmonic DFT simulations, the AIMD approach considers the influence of the entire bulk solvent and also some anharmonic effects, although it requires much more computational resources. The current study demonstrates the utility of the recently developed AIMD ROA approach for flexible chiral molecules, and the insights gained into the solute-solvent interactions using two very different modelling approaches, a subject of

important current interest in developing ROA spectroscopy as a versatile technique for extracting structural information of chiral molecules in solution.

5.4 Experimental Section

5.4.1 Experimental Details

NALC (purity≥99%) and methanol (purity≥99.8%) were obtained from Sigma-Aldrich and used without further purification. Sample solutions were prepared by dissolving NALC in either methanol (1.6 M) or in double distilled water (0.67 M). Raman and ROA spectra were measured using a BioTools ChiralRAMAN-2X spectrometer (532 nm excitation laser) in a back scattering configuration, designed with Hug's virtual enantiomer. [52] 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.

5.4.2 DFT Details

The conformer-rotamer ensemble sampling tool (CREST) code[45] was utilized to systematically explore the possible conformers of the NALC monomer and the 1:1 NALC-H_2O binary complex with the implicit solvation model[53] of solvent in place. Some larger NALC-$(H_2O)_{2\text{-}6}$ aggregates were also searched. The dubbed GBSA model of water or methanol were used in the searches. Built upon the previous semiempirical tight-binding (TB) quantum chemistry method by Grimme and co-workers, called GFN-xTB,[54] the new code is designed for fast and reliable exploration and screening of the conformational space of mid- to large-sized molecules with up to about a thousand atoms. It can automatically perform multiple MD runs for conformational exploration and has a built-in feature to compare and discard redundant geometries. It has been successfully applied to support

rotational spectroscopic studies of mid-sized organic molecules and their clusters[55,56] in identifying all low energy structural candidates, although occasional misses were also reported.[57]

A multi-tiered approach developed before[58] was utilized in the current study which includes several steps: 1) the generation of CREST candidates; 2) the DFT optimizations of the CREST candidates with a relaxed convergence criteria at the revPBE-D3/def2-SVP[59] level, with the empirical D3 dispersion correction;[60,61] 3) a single-point energy evaluation at the B3LYP[62,63]-D3/def2-TZVP level of the optimized structures in step 2). Steps 2 and 3 were done using Molpro;[64] 4) The final geometry optimization and harmonic frequency Raman and ROA calculations were performed using the Gaussian 16 package.[65] All calculations were done at the B3LYP-D3BJ/def2-TZVP level, with the Becke-Johnson (BJ) damping function,[66] coupled with the def2-TZVP[59] basis set. The implicit solvent was included using the integral equation formalism (IEF) version of the PCM to account for the bulk solvent environment (ε = 78.3553 for water and ε = 32.613 for methanol).[53] Some additional calculations were also carried out at the B3LYP/cc-pVTZ[67] level for geometry optimization and the related ROA calculations were performed using the same optimized geometries and with the aug-cc-pVDZ[68] basis set where 532 nm was set as the excitation wavelength, following the recommendation in Ref. [69]. Similar results were obtained with both functional/basis set combinations and all final results are presented using the B3LYP-D3BJ/def2-TZVP simulations. 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. The Raman and ROA intensity base unit is 10^{-40} m^2cm/sr where m is meter and sr is steradian for all spectral figures. No frequency scaling factors were used.

5.4.3 AIMD Calculations

The vibrational spectra of NALC were computed from the AIMD simulations of the relevant NALC conformers identified (vide infra) in explicit solvent. To prepare the

solvated structures, each relevant NALC conformer was put into a periodic cell of approximately 1.4 nm edge length (cubic), and 64 water molecules or 26 methanol molecules were randomly placed around the solute with the Packmol software.[70] To equilibrate the solvent shell, force field molecular dynamics simulations were performed with the LAMMPS program package.[71] During these runs, all atoms of the NALC molecule were kept fixed. The partial charges in NALC were obtained from RESP calculations,[72] while the atomic Lennard-Jones parameters were taken from OPLS–AA.[73] Methanol was also, modelled by OPLS-AA, while the TIP4P-EW force field was applied for water.[74]

Based on these pre-equilibrated structures, AIMD simulations were carried out with the program package CP2K,[75,76] using the Quickstep module[77] in conjunction with the orbital transformation (OT) algorithm.[78] The electron structure calculations were performed using the BLYP exchange-correlation functional,[62,63] together with the recent re-parametrization[61] of Grimme's D3 dispersion correction with Becke–Johnson damping, and the basis sets of the kind DZVP–MOLOPT–SR[79] together with Goedecker–Teter–Hutter (GTH) pseudopotentials.[80] The plane wave cutoff was set to 350 Ry with a REL_CUTOFF of 40. The temperature during the simulations was kept constant using a Nosé–Hoover chain thermostat[81-83] with a time constant of 100 fs. The integration time step was set to 0.5 fs in all cases. The simulations were running for 40 ps each, of which the first 15 ps were discarded as equilibration, so that the production run amounted for 25 ps. In the AIMD simulations, all relevant backbone dihedral angles of the conformers were kept approximately constant via harmonic restraints with a spring constant of 0.005 Hartree rad^{-2} in order to avoid the interconversion of conformers within each trajectory.

To obtain the vibrational spectra, the total electron density along each simulation was stored in the lossless, compressed bqb format[84] and subsequently processed using the Voronoi integration technique to obtain molecular electromagnetic moments.[33,34] The magnetic moments were obtained from the recently published classical approach.[35] In order

to obtain the ROA spectra, the recently published protocol for bulk phase systems was employed.[39] The required polarizabilities were obtained via two-sided finite differences of an external electric field with a field strength of $5.0*10^{-3}$ a.u. = $2.57*10^{9}$ V m^{-1}. All these steps were carried out with the TRAVIS program package.[40,41] The correlation functions for the spectra were computed with a temporal correlation depth of 1.15 ps.

Please note the DFT and the AIMD frequencies were not scaled since applying different empirical scaling factors to match the experimental ones in each case would make it difficult see the frequency differences among different approaches clearly. For example, one can see the effect on the band frequencies with and without the explicit water molecules clearly without scaling.

5.5 Acknowledgements

This research was funded by the Natural Sciences and Engineering Research Council of Canada, the Canada Foundation for Innovation, Alberta Enterprise of Advanced Education, and by the University of Alberta. We gratefully acknowledge access to the computing facilities by the Shared Hierarchical Academic Research Computing Network, the Western Canada Research Grid (Westgrid), and Compute/Calcul Canada. MB acknowledges financial support by the DFG through project Br 5494/1-1. YX is a Tier I Canada Research Chair in Chirality and Chirality Recognition.

References

[1] E. Brini, C. J. Fennell, M. Fernandez-Serra, B. H. Lee, M. Lukšič, K. A. Dill, *Chem. Rev.* **2017**, *117*, 12385-12414.

[2] N. Borho, Y. Xu, *Phys. Chem. Chem. Phys.* **2007**, *9*, 1324-1328.

[3] P. Ottaviani, B. Velino, W. Caminati, *Chem. Phys. Lett.* **2006**, *428*, 236-240.

[4] J. Thomas, O. Sukhorukov, W. Jäger, Y. Xu, *Angew. Chem. Int. Ed.* **2014**, *53*, 1156-1159; *Angew. Chem.* **2014**, *126*, 1175-1178.

[5] F. Xie, X. Ng, N. A. Seifert, J. Thomas, W. Jäger, Y. Xu, *J. Chem. Phys.* **2018**, *149*, 224306.

[6] F. Xie, S. Mahendiran, N. A. Seifert, Y. Xu, *Phys. Chem. Chem. Phys.* **2021**, *23*, 3820-3825.

[7] M. Gasior-Glogowska, K. Malek, G. Zajac, M. Baranska, *Analyst* **2016**, *141*, 291-296.

[8] D. P. Demarque, M. Kemper, C. Merten, *Chem. Commun.* **2021**, *57*, 4031-4034.

[9] C. Merten, F. Li, K. Bravo-Rodriguez, E. Sanchez-Garcia, Y. Xu, W. Sander, *Phys. Chem. Chem. Phys.* **2014**, *16*, 5627-5633.

[10] M. R. Poopari, P. Zhu, Z. Dezhahang, Y. Xu, *J. Chem. Phys.* **2012**, *137*, 194308.

[11] P. Zhu, G. Yang, M. R. Poopari, Z. Bie, Y. Xu, *ChemPhysChem* **2012**, *13*, 1272-1281.

[12] M. Losada, Y. Xu, *Phys. Chem. Chem. Phys.* **2007**, *9*, 3127-3135.

[13] M. Losada, P. Nguyen, Y. Xu, *J. Phys. Chem. A* **2008**, *112*, 5621-5627.

[14] K. Le Barbu-Debus, J. Bowles, S. Jähnigen, C. Clavaguéra, F. Calvo, R. Vuilleumier, A. Zehnacker, *Phys. Chem. Chem. Phys.* **2020**, *22*, 26047-26068.

[15] S. T. Mutter, F. Zielinski, J. R. Cheeseman, C. Johannessen, P. L. A. Popelier, E. W. Blanch, *Phys. Chem. Chem. Phys.* **2015**, *17*, 6016-6027.

[16] K. H. Hopmann, K. Ruud, M. Pecul, A. Kudelski, M. Dračínský, P. Bouř, *J. Phys. Chem. B* **2011**, *115*, 14, 4128-4137.

[17] S. Cardamone, B. A. Caine, E. Blanch, M. G. Lizio, P. L. A. Popelier, *Phys. Chem. Chem. Phys.* **2016**, *18*, 27377-27389.

[18] T. Giovannini, G. Del-Frate, P. Lafiosca, C. Cappelli, *Phys. Chem. Chem. Phys.* **2018**, *20*, 9181-9197.

[19] S. Ghidinelli, S. Abbate, J. Koshoubu, Y. Araki, T. Wada, G. Longhi, *J. Phys. Chem. B* **2020**, *124*, 4512-4526.

[20] V. Palivec, C. Johannessen, J. Kaminský, H. Martinez-Seara, *PLoS Comput. Biol.* **2022**, *18*, e1009678.

[21] A. S. Perera, J. Thomas, M. R. Poopari, Y. Xu, *Front. Chem.* **2016**, *4*, 1-17.

[22] G. Yang, Y. Xu, *J. Chem. Phys.* **2009**, *130*, 164506.

[23] M. Losada, H. Tran, Y. Xu, *J. Chem. Phys.* **2008**, *128*, 014508.

[24] A. S. Perera, J. Cheramy, C. Merten, J. Thomas, Y. Xu, *ChemPhysChem* **2018**, *19*, 2234-2242.

[25] C. Merten, J. Bloino, V. Barone, Y. Xu, *J. Phys. Chem. Lett.* **2013**, *4*, 3424-3428.

[26] A. Stark, M. Brehm, M. Brüssel, S. B. C. Lehmann, A. S. Pensado, M. Schöppke, B. Kirchner, *Top. Curr. Chem.* **2014**, *351*, 149-187.

[27] M. Thomas, M. Brehm, R. Fligg, P. Vöhringer, B. Kirchner, *Phys. Chem. Chem. Phys.* **2013**, *15*, 6608-6622.

[28] P. L. Silvestrelli, M. Bernasconi, M. Parrinello, *Chem. Phys. Lett.* **1997**, *277*, 478-482.

[29] A. Putrino, M. Parrinello, *Phys. Rev. Lett.* **2002**, *88*, 176401.

[30] M. Thomas, Theoretical modeling of vibrational spectra in the liquid phase. *Springer International Publishing*, **2017**.

[31] G. H. Wannier, *Phys. Rev.* **1937**, *52*, 191-197.

[32] N. Marzari, D. Vanderbilt, *Phys. Rev. B* **1997**, *56*, 12847-12865.

[33] M. Thomas, M. Brehm, B. Kirchner, *Phys. Chem. Chem. Phys.* **2015**, *17*, 3207-3213.

[34] M. Brehm, M. Thomas, *Molecules* **2021**, *26*, 1875.

[35] M. Thomas, B. Kirchner, *J. Phys. Chem. Lett.* **2016**, *7*, 509-513.

[36] A. Scherrer, R. Vuilleumier, D. Sebastiani, *J. Chem. Phys.* **2016**, *145*, 084101.

[37] S. Abbate, G. Longhi, K. Kwon, A. Moscowitz, *J. Chem. Phys.* **1998**, *108*, 50-62.

[38] S. Luber, *J. Chem. Theor. Comput.* **2017**, *13*, 1254-1262.

[39] M. Brehm, M. Thomas, *J. Phys. Chem. Lett.* **2017**, *8*, 3409-3414.

[40] M. Brehm, B. Kirchner, *J. Chem. Inf. Model.* **2011**, *51*, 2007-2023.

[41] M. Brehm, M. Thomas, S. Gehrke, B. Kirchner, *J. Chem. Phys.* **2020**, *152*, 164105.

[42] M. R. Poopari, Z. Dezhahang, G. Yang, Y. Xu, *ChemPhysChem* **2012**, *13*, 2310-2321.

[43] M. R. Poopari, Z. Dezhahang, Y. Xu, *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **2013**, *136*, 131-140.

[44] C. Merten, L. D. Barron, L. Hecht, C. Johannessen, *Angew. Chem. Int. Ed.* **2011**, *50*, 9973-9976; *Angew. Chem.* **2011**, *123*, 10149-10152.

[45] P. Pracht, F. Bohle, S. Grimme, *Phys. Chem. Chem. Phys.* **2020**, *22*, 7169-7192.

[46] Y. Yang, J. Cheramy, Y. Xu, *ChemPhysChem*, **2021**, *22*, 1336-1343.

[47] M. R. Poopari, Z. Dezhahang, Y. Xu, *Phys. Chem. Chem. Phys.* **2013**, *15*, 1655-1665.

[48] Y. Liu, G. Yang, M. Losada, Y. Xu, *J. Chem. Phys.* **2010**, *132*, 234513.

[49] L. Weirich, J. M. de Oliveira, C. Merten, *Phys. Chem. Chem. Phys.* **2020**, *22*, 1525-1533.

[50] S. A. Katsyuba, S. Spicher, T. P. Gerasimova, S. Grimme, *J. Chem. Phys.* **2021**, *155,* 024507.

[51] G. N. Simm, P. L. Türtscher, M. Reiher, *J. Comput. Chem.* **2020**, *41*, 1144-1155.

[52] W. Hug, *Appl. Spectrosc.* **2003**, *57*, 1-13.

[53] C. Bannwarth, E. Caldeweyher, S. Ehlert, A. Hansen, P. Pracht, J. Seibert, S. Spicher, S. Grimme, *WIREs Comput. Mol. Sci.* **2020**, *11*, e1493.

[54] S. Grimme, C. Bannwarth, P. Shushkov, *J. Chem. Theory Comput.* **2017**, *13*, 1989-2009.

[55] S. Oswald, N. A. Seifert, F. Bohle, M. Gawrilow, S. Grimme, W. Jäger, Y. Xu, M. A. Suhm, *Angew. Chem. Int. Ed.* **2019**, *58*, 5080-5084; *Angew. Chem.* **2019**, *131*, 5134-5138.

[56] F. Xie, N. A. Seifert, W. Jäger, Y. Xu, *Angew. Chem. Int. Ed.* **2020**, *59*, 15703-15710; *Angew. Chem.* **2020**, *132*, 15833-15840.

[57] F. Xie, M. Fusè, A. S. Hazrah, W. Jäger, V. Barone, Y. Xu, *Angew. Chem. Int. Ed.* **2020**, *59*, 22427-22430; *Angew. Chem.* **2020**, *132*, 22613-22616.

[58] H. Wang, M. Heger, M. H. Al-Jabiri, Y. Xu, *Molecules*, **2022**, *27*, 38.

[59] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297-3305.

[60] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.

[61] D. G. A. Smith, L. A. Burns, K. Patkowski, C. D. Sherrill, *J. Phys. Chem. Lett.* **2016**, *7*, 2197-2203.

[62] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098-3100.

[63] C. Lee, W. Yang, R. Parr, *Phys. Rev. B* **1988**, *37*, 785-789.

[64] H. J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, *Wires Comput. Mol. Sci.* **2012**, *2*, 242-253.

[65] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, et al., *Gaussian 16*, Revision C.03; Gaussian, Inc.: Wallingford, CT, **2019**.

[66] A. D. Becke, A. D. Johnson, *J. Chem. Phys.* **2005**, *123*, 154101.

[67] R. A. Kendall, T. H. Dunning Jr., R. J. Harrison, *J. Chem. Phys.* **1992**, *96*, 6796-6806.

[68] T. H. Dunning Jr, *J. Chem. Phys.* **1989**, *90*, 1007-1023.

[69] J. R. Cheeseman, M. J. Frisch, *J. Chem. Theory Comput.* **2011**, *7*, 3323-3334.

[70] L. Martínez, R. Andrade, E. G. Birgin, J. M. Martínez, *J. Comput. Chem.* **2009**, *30*, 2157-2164.

[71] S. Plimpton, *J. Comp. Phys.* **1995**, *117*, 1-19.

[72] C. I. Bayly, P. Cieplak, W. Cornell, P. A. Kollman, *J. Phys. Chem.* **1993**, *97*, 10269-10280.

[73] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc.* **1996**, *118*, 11225-11236.

[74] H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura, T. Head-Gordon, *J. Chem. Phys.* **2004**, *120*, 9665-9678.

[75] J. Hutter, M. Iannuzzi, F. Schiffmann, J. VandeVondele, *WIREs Comput. Mol. Sci.* **2014**, *4*, 15-25.

[76] T. D. Kühne, M. Iannuzzi, M. Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann, D. Golze, J. Wilhelm, S. Chulkov, M. H. Bani-Hashemian, V. Weber, U. Borštnik, M. Taillefumier, A. S. Jakobovits, A. Laz-zaro, H. Pabst, T. Müller, R. Schade, M. Guidon, S. Andermatt, N. Holmberg, G. K. Schenter, A. Hehn, A. Bussy, F. Belleflamme, G. Tabacchi, A. Glöß, M. Lass, I. Bethune, C. J. Mundy, C. Plessl, M. Watkins, J. VandeVondele,

M. Krack, J. Hutter, *J. Chem. Phys.* **2020**, *152*, 194103.

[77] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, J. Hutter, *Comput. Phys. Commun.* **2005**, *167*, 103-128.

[78] J. VandeVondele, J. Hutter, *J. Chem. Phys.* **2003**, *118*, 4365-4369.

[79] J. VandeVondele, J. Hutter, *J. Chem. Phys.* **2007**, *127*, 114105.

[80] S. Goedecker, M. Teter, J. Hutter, *Phys. Rev. B* **1996**, *54*, 1703-1710.

[81] S. Nose, *Mol. Phys.* **1984**, *52*, 255-268.

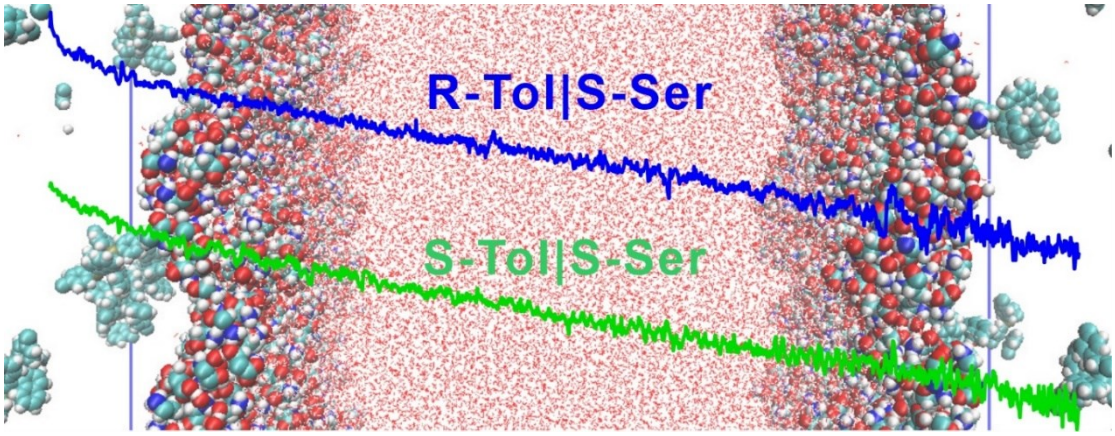
[82] W. G. Hoover, *Phys. Rev. A* **1985**, *31*, 1695-1697.

[83] G. Martyna, M. Klein, M. Tuckerman, *J. Chem. Phys.* **1992**, *97*, 2635-2643.

[84] M. Brehm, M. Thomas, *J. Chem. Inf. Model.* **2018**, *58*, 2092-2107.

Chapter 6

Chirality Discrimination at Binary Organic/Water Interfaces by Interfacial Tension Measurements and MD Simulations



This chapter is directly copied and adapted the new format from the following submitted paper:

Y. Yang, X. Sun, M. R. Poopari, C. Jian, H. Zeng, T. Tang, Y. Xu, *ChemPhysChem* **2022**.

6.1 Introduction

A molecule is chiral if its mirror image cannot be superimposed onto itself. Chirality discrimination/recognition, one of the most subtle aspects of molecular recognition, is a phenomenon in which a chiral molecule can distinguish a pair of enantiomers of another chiral molecule [1]. At the molecular level, non-covalent intermolecular interactions responsible for chirality recognition in prototype chiral molecular contact pairs have been extensively studied, for example using rotational[2,3] and infrared ion[4,5] spectroscopies. Such recognition events between chiral molecules have been utilized in NMR based, Mosher ester analysis for absolute configuration determination [6] and in the recent development of chiral tag rotational spectroscopy for enantiomeric excess determination[7]. In pharmaceutical and biotechnology fields, chiral recognition plays a crucial role in the separation and extraction technologies[8]. Chirality recognition also plays a key role in supramolecular architectures at the air-water interface[9] and such supramolecular structures often show greatly enhanced chiroptical responses [10]. For example, the enantiopure electroactive films based on a 2,2-biindole core were prepared and were reported to show powerful chirality manifestations[11].

More closely related to the current study are the reports of chirality recognition at surfaces and at liquid-liquid interfaces. Molecular force measurements have been utilized as a promising technique to discriminate a pair of enantiomers by means of quantifying their interaction forces at the contacted boundaries. For example, by chemical derivatization of a scanning probe tip with a chiral ligand, McKendry et al.[12] employed atomic force microscopy to distinguish between a pair of mandelic acid enantiomers arrayed on a surface by measuring the differences in the adhesion and frictional forces. Later, Kim and co-workers [13] designed a L-phenylalanine modified quartz cell microbalance to detect L- versus D-mandelic acid utilizing a vapor diffused molecular assembly reaction approach. In their experiment, L-mandelic acid was found to have a higher propensity toward L-phenylalanine than D-mandelic acid. By immersing α_1-acid-

glycoprotein (AGP) in water to create a micro liquid-liquid interface and monitoring the related cyclic voltammetry and differential pulse voltammetry, Lopes and Kataky[14] found that AGP has a greater binding affinity with S-propranolol than the R-enantiomer. An interesting concept of 'dynamical interface for chiral discrimination' was invoked by Shundo et al.[15], who demonstrated an enantioselective wetting of a chiral polymer film by monitoring the contact angle change with time. While the surface tensions (SF) were identical, it was found, on the other hand, that the contact angle of chiral liquids on the film was strongly associated to their chirality. Although not directly related to chirality recognition research, Adachi et al.[16] used the interfacial tension (IFT) changes at the toluene (with R or S- 2,2′-bis(diphenylphosphino)-1,1′-binaphthyl (BINAP)-water (with achiral $PdSO_4$ salt) interface at increasing salt concentration to monitor the formation of the Pd (II)$BINAP_2^+$ complex at the toluene-water interface. The IFT values observed were insensitive to the chirality of BINAP and relaxed to the same magnitudes as the concentration increased, as expected. More recently, Okada and co-workers[17] applied chromatographic data to reveal enhanced chiral recognition by β-cyclodextrin at liquid/liquid interfaces where the hexane-water IFT changes of β-cyclodextrin under different concentrations were used to establish its interfacial concentration in order to extract the interfacial equilibrium constants. Umemura et al. observed the enhanced chiral recognition signal by reorganized the elementary clay mineral platelets in the monolayers or multi-layered nanofilms through the Langmuir-Blodgett approach[18,19].

Liquid-liquid interfacial measurements have been widely used to understand the chemical and biological processes involving interfaces, such as food emulsions[20], drug delivery through liquid membrane[21], oil recovery[22-25], and induced aggregation processes[26], and how to control them. The interfacial molecular systems may have profound impacts on the chemical behavior and reactivity of solutes in ways which differ from those of a bulk liquid environment. As a result, chemical phenomena which were not observed in bulk, could be detected at the liquid-liquid interface[27,28]. In the present study, we demonstrated

and analyzed chirality recognition at the binary organic-water interface by monitoring its interfacial tension measurements using a state-of-the-art optical tensiometer with Tol-BINAP (2,2'-Bis(di-p-tolylphosphino)-1,1'-binaphthyl) and serine as the organic and aqueous solutes, respectively. Although molecular dynamics (MD) techniques have been utilized for studies of chiral molecules[29,30] and of systems containing complex mixtures, such as humic acid[31] in bulk solution and for investigations of IFT at the organic-water interface by Zeng and others[32,33], no MD study associated with chirality recognition events at an organic-water interface has been reported. It is therefore of considerable interest to carry out MD simulations of IFT at the organic-water interface including these chiral solutes to benchmark such calculations. Consistent results were obtained by IFT experimental measurements and by MD simulations, suggesting the utility of the latter for predicting chirality recognition outcome at organic-water interfaces. To the best of our knowledge, this is the first study on chirality discrimination by means of IFT measurements coupled with MD simulations.

6.2 Results and discussion

6.2.1 The IFT systems studied

To investigate chirality discrimination phenomena at the organic-water interface, the first step is to select appropriate aqueous- and organic-soluble chiral solutes. In this study, serine, one of the 20 naturally existing amino acids, which has been widely used in pharmaceutical applications[34], was selected as the aqueous solute. Tol-BINAP, which has been widely used as a precursor and a chiral catalyst in various organic syntheses[35,36], was chosen as the organic solute. Tol-BINAP exists in two stable axial chiral forms. This is because the interconversion barrier between the R- and S-forms of Tol-BINAP was found to be very high due to the large P substituents[37,38]. The substituents are spanned in all directions to minimize the steric effect, resulting in a relatively spherical shape.

The molecular systems as well as the compositions of droplets and surrounding

medium phases used for the current study are all tabulated in Table 6.1. In the remainder of the paper, we list the organic phase on the left and the aqueous phase on the right side of the symbol "|" which indicates the interface. Furthermore, if there are solutes involved in each phase, then only the solute is indicated, i.e., the solvents are implied in such cases. Overall, ten sets of experiments were designed to evaluate the chirality discrimination capability of the homochiral and heterochiral Tol-BINAP|Serine interfaces. The abbreviations of the systems introduced in Table 6.1 are used in the remainder of the paper when referring to a specific binary system.

To validate if MD simulations can correctly capture the above experimental observations, MD simulations were carried out on the systems of interest. The details of the simulated systems are shown in Table 6.1. Moreover, pH is another crucial factor to be considered since it can influence the charge state of a solute molecule and therefore the aggregation behavior of the native surface-active compounds at the interface substantially. It was reported previously that the molecules of interest become more ionized under low or high pH, they tend to accumulate more at the interface which enhance their surface activity and result in a lower IFT reading[39]. As elaborated later, both zwitterionic and neutral serine are expected under the current pH condition. The optimized geometries of a zwitterionic serine molecule, a neutral serine molecule and a Tol-BINAP molecule are shown in Figure D1, Appendix D.

Table 6.1. The systems investigated.

System	Droplet Phase[a]	Medium Phase[b]	# of molecules in simulated systems			
			# of Serine	# of Water	# of Tol-BINAP	# of Toluene
R-Tol-BINAP| Water	R-Tol-BINAP in Toluene	Pure water	-	35602	20	5973

S-Tol-BINAP| Water	S-Tol-BINAP in Toluene	Pure water	-	35475	20	5979
S/R-Tol-BINAP (50:50)|Water	S/R-Tol-BINAP in Toluene	Pure water	-	-	-	-
Toluene| R-Serine	Pure toluene	R-serine in water	1389	33046	-	5241
Toluene| S-Serine	Pure toluene	S-serine in water	1389	33046	-	5418
S-Tol-BINAP| S-Serine	S-Tol-BINAP in Toluene	S-serine in water	1389	33046	20	4507
R-Tol-BINAP| S-Serine	R-Tol-BINAP in Toluene	S-serine in water	1389	33046	20	4502
S/R-Tol-BINAP (50:50)|S-Serine	S/R-Tol-BINAP in Toluene	S-serine in water	-	-	-	-
R-Tol-BINAP| R-Serine	R-Tol-BINAP in Toluene	R-serine in water	1389	33046	20	4264
S-Tol-BINAP| R-Serine	S-Tol-BINAP in Toluene	R-serine in water	1389	33046	20	4262

a 9wt% of R-/ and S-Serine solutions were used. b 1000 ppm solutions of R-/ and S-Tol-BINAP were used.

6.2.2 Controlled experiments with a chiral solute in only one phase

Several controlled experiments were also carried out in which only one chiral solute was dissolved in either the aqueous or organic phase. First, the liquid-liquid interfacial

behavior between Tol-BINAP (droplet phase in toluene) and pure water (surrounding medium) was investigated. Experimentally, as shown in Figure 6.1, the pure Toluene|Water system has a stable IFT value of ~36.0 mN/m during the entire measurement time, which agrees with the previous literature[40]. With the addition of a chiral solute in the organic phase, a steep IFT decrease was observed during the first 400 s of the measurement time and then the IFT value quickly reached the plateau afterwards. It is important to note that both R-Tol-BINAP|Water and S-Tol-BINAP|Water systems converged to the same IFT values, indicating the absence of chirality discrimination between the Tol-BINAP and water molecules. This is as expected since the R-Tol-BINAP|Water and S-Tol-BINAP|Water systems are mirror images to each other and should have exactly the same IFT value. Interestingly, the racemic mixture of the R and S-Tol-BINAP also gave no detectable difference in its IFT value compared to those obtained for the enantiopure R- or S-Tol-BINAP systems. This indicates that there are no significant binding differences between the racemic and enantiopure Tol-BINAP with water molecules at the organic-water interface. It is important to emphasize that such an outcome is not known a priori since the racemic Tol-BINAP phase contains both homochiral and heterochiral interaction pairs or larger aggregates while the enantiopure phase contains only homochiral ones. In principle, such interaction differences in the same phase may also lead to different IFT values due to chirality discrimination originated in the same phase rather than at the interface, the subject of current study.

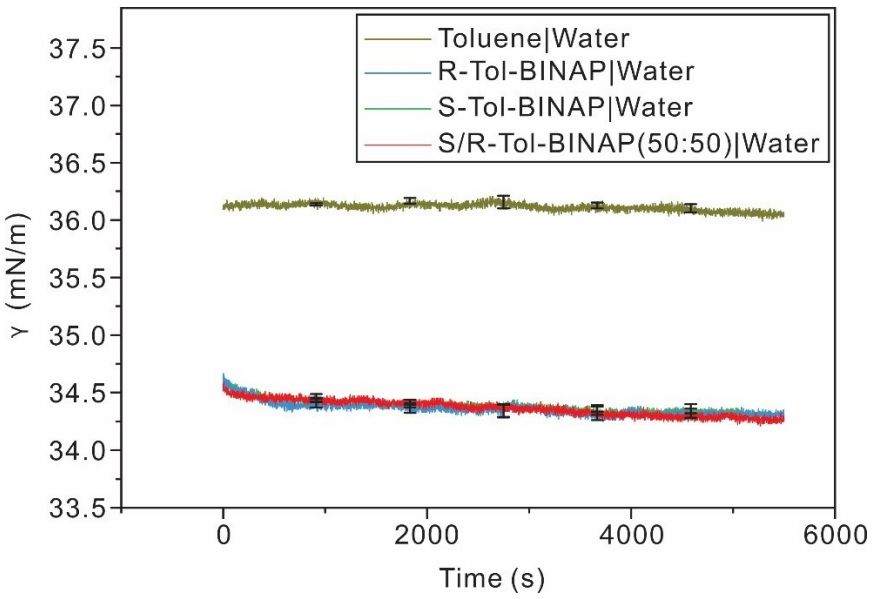


Figure 6.1. The experimental IFT plots of the Toluene|Water, S-Tol-BINAP|Water, R-Tol-BINAP|Water and S/R-Tol-BINAP (50:50 v/v)|Water systems. The last three IFT traces are essentially overlapped with each other.

The experimental and theoretical interfacial tension values for the above controlled experiments with S-Tol-BINAP|Water and R-Tol-BINAP|Water are summarized in Table 6.2. It is gratifying to note that the calculated IFT values for the last 30 ns of the MD simulations of the S-Tol-BINAP|Water and R-Tol-BINAP|Water systems turned out to be essentially the same, i.e., 34.5 ± 0.63 mN/m and 34.8 ± 0.68 mN/m, respectively, agreeing with the experimental observation.

Table 6.2. Experimental and theoretical interfacial tension (γ) values in mN/m for the controlled experiments.

Samples	γ-Theo	γ-Exp (5500 s)
Toluene|Water	-	36.03 ± 0.01
R-Tol-BINAP|Water	34.8 ± 0.68	34.32 ± 0.06
S-Tol-BINAP|Water	34.5 ± 0.63	34.28 ± 0.04

S/R-Tol-BINAP (50:50)|Water	-	34.26 ± 0.04
Toluene|R-Serine	42.3±1.85 (zwitterion) 18.8 ± 0.95 (neutral)	47.86 ± 0.11
Toluene|S-Serine	41.3 ± 2.05 (zwitterion) 19.1 ± 0.94 (neutral)	49.45 ± 0.04

For many chiral molecules including amino acids, one enantiomer is naturally occurring while the other is unnatural. The synthetic routes for the unnatural ones can be tedious and sometimes trace amounts of stabilizers may be added to these products. S-Serine is naturally occurring, whereas R-Serine is not. Since IFT can be very susceptible to the presence of impurities, we used only S-Serine in all the experiments. Out of curiosity, we did carry out the measurements of pure toluene droplet in an aqueous medium containing 9 wt% of either R- or S-Serine, and the resulting plots are depicted in Figure D2, Appendix D. As one can see, the IFT values of Toluene|S-Serine versus Toluene|R-Serine differ by about 1.6 mN/m. This was not too surprising because of the reason provided above.

We have also performed MD simulations for the above experimental systems. The initial models were generated based on the data published in a previous study[41,42]. The serine molecules have a tendency to self-associate in water.[42] This phenomenon was also observed in our simulation of the serine molecules confined in bulk water, as shown in Figure D3, Appendix D. It is gratifying to note that the MD calculations of the Toluene|R-Serine and Toluene|S-Serine systems provide the IFT values which are nearly the same within an acceptable error, consistent with the expectation. The MD interfacial tension values for the Toluene|R-Serine and Toluene|S-Serine solutions are also summarized in Table 6.2.

In the experimental measurements, the IFT value of ~49.5 mN/m for the Toluene|S-Serine system was found to be significantly higher than that of ~36.0 mN/m for the

Toluene|Water system. What is the mechanism for such a large increase in IFT when serine was added to the water phase? Since the serine molecules exist dominantly as zwitterions under neutral pH[43], their NH_3^+ and COO^- sites are solvated by more water molecules than the neutral serine molecules. It is therefore helpful to compare the zwitterions' IFT behavior with other well-known charged species. For example, it was reported that the IFT value of the Toluene|Water interface increases when increasing inorganic salt concentration to water. This is because water molecules tend to form the cage-like hydrogen-bonding interaction with ions, increasing the cohesive force between the water molecules and therefore the energy that is needed to generate a unit area of free water surface, resulting in a net increase in the related IFT[44]. Moreover, it was reported previously that more polarizable ions would lead to more repulsion at the interfacial surface[45,46]. Compared to the simple salt ions such as Na^+ and Cl^-, serine zwitterions are larger in size and are more polarizable. One can therefore expect adding serine would increase the corresponding IFT value greatly when compared to that of the pure Toluene|Water interface. Indeed, our MD predicted an IFT value of 42.3 mN/m for the interface with the serine zwitterions, a significant increase from the 18.8 mN/m value for the interface with the neutral serine. Actually, the calculated IFT value of the Toluene|neutral-Serine system falls below that of the pure Toluene|Water system, in contrary to the fact that the experimental IFT value of Toluene|Serine is significantly higher than that of Toluene|Water. Clearly, the experimental trends can only be captured by implementing the zwitterion form of serine in the simulations. This is in line with the previous vibrational spectroscopic studies where the zwitterionic form of serine was found to be the dominant one under the same pH condition as the current experiments[43]. Therefore, the zwitterion form of serine will be the only one used for the subsequent MD simulations.

In summary, the controlled experiments with the enantiopure (R- and S-Tol-BINAP) samples exhibit the same IFT values as expected, so do the associated enantiopure MD simulations. The MD outputs reveal that the calculated IFT values for the last 30 ns of

simulations of S-Tol-BINAP|Water and R-Tol-BINAP|Water systems turned out to be essentially the same, i.e., 34.5 ± 0.63 mN/m and 34.8 ± 0.68 mN/m, respectively. While the MD simulations of the Toluene|R-Serine and Toluene|S-Serine generated the same IFT behavior, the controlled experiments with the R- and S-Serine samples show somewhat different IFT behavior because of their different impurities as discussed above.

6.2.3 Effect of chiral discrimination with chiral solutes in both phases

To explore possible chirality discrimination with IFT measurements, a set of experiments were constructed where chiral molecules were introduced in both the aqueous and the organic phase. In these experiments, the IFT values of a droplet containing 1000 ppm of Tol-BINAP (S or R) in toluene were measured in the presence of 9 wt% solutions of serine (S or R) prepared in DI-water, or a total of four experiments. To avoid the complication of the serine impurity issue, we will compare the IFT trends obtained with the R-Serine and S-Serine separately, in the presence of organic-soluble chiral species, i.e., Tol-BINAP. Under such conditions, one can confidently evaluate the IFT differences introduced by the addition of R-Tol-BINAP versus S-Tol-BINAP and attribute such differences to chirality discrimination.

In the first set of experiments, the IFT values of R-Tol-BINAP|S-Serine and S-Tol-BINAP|S-Serine pairs were measured to evaluate the difference in the interfacial interactions between the homo- and heterochiral pairs at the interface. The experimental IFT values were collected for a total of 5500 s, as depicted in Figure 6.2, while their values at t = 5500 s are listed in Table 6.3. Initially (at t = 0 s), the difference between the IFT values of S-Tol-BINAP|S-Serine and R-Tol-BINAP|S-Serine was found to be around 0.95 mN/m. The difference stays fairly stable with an ending difference value of 1.00 mN/m at t = 5500 s. More importantly, at longer time, one could see the slopes of the three IFT-time traces decrease and appear to approach each of their equilibria slowly. The “drifts” observed, i.e., the long times at which the IFT-time traces reach equilibria, are hypothesized

to be associated with the time by which the solute molecules take to move to the interface and undergo rearrangements to eventually achieve the equilibrium. Similarly, long delays to reach a plateau were reported previously in a study of the dynamics of protein adsorption at an Oil|Water interface with different concentrations of a sodium salt. In fact, the system did not reach an equilibrium even after several hours, and this was attributed to a slow conformational rearrangement of the interfacial film.[47] Since the serine zwitterions are much larger than the simple salt ions such as Na^+ and Cl^- and are charged at both ends, the solvation shells would be even larger. As such, it is not surprising for the serine zwitterions to take a long time to reach the interface and for the whole system to take a long time to undergo the necessary rearrangement to finally reach an interface equilibrium.

From Figure 6.2, the lower IFT readings are associated with the S-Tol-BINAP|S-Serine system, i.e., the homochiral case. This result suggests the R-Tol-BINAP and S-Tol-BINAP molecules interact differently with the S-Serine molecules at the Organic|Water interface, which could be attributed to the difference in their intermolecular interactions in the vicinity of the interface.[48] With the controlled experiments with R-Tol-BINAP|water and S-Tol-BINAP|water which give the same IFT value, the differences observed between the IFT values of R-Tol-BINAP|S-Serine and S-Tol-BINAP|S-Serine are clearly due to chirality-specific discrimination at the interface.

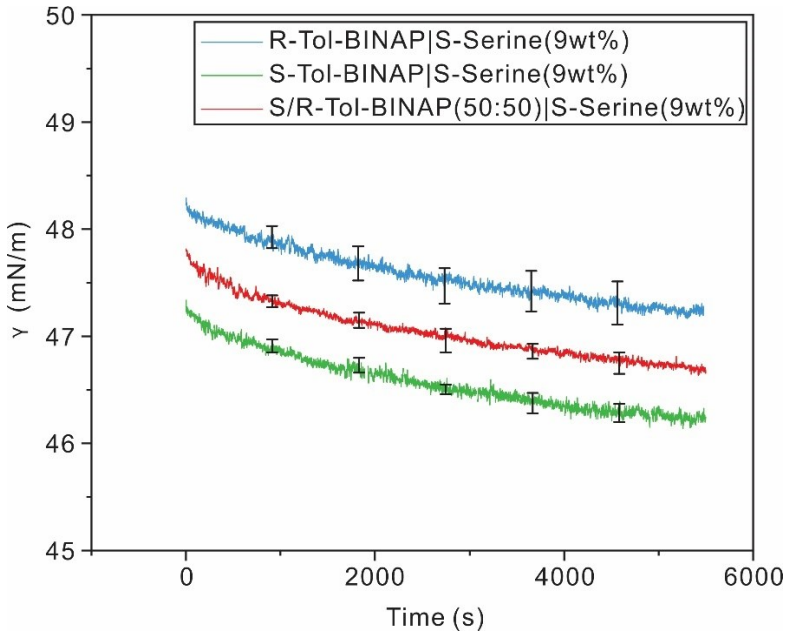


Figure 6.2. The IFT plots of S-Tol-BINAP|S-Serine (9wt%), R-Tol-BINAP|S-Serine (9wt%) and S/R-Tol-BINAP (50:50 v/v)|S-Serine (9wt%) systems.

Table 6.3. Experimental and theoretical interfacial tension (γ) values in mN/m for the R-Tol-BINAP, S-Tol-BINAP and S/R-Tol-BINAP (50:50 v/v) (in toluene) in 9 wt% S/(R)-Serine (in water) systems.

Sample	γ-Theo	γ-Exp (5500 s)
R-Tol-BINAP|S-Serine	41.5 ± 1.65	47.24 ± 0.21
S-Tol-BINAP|S-Serine	35.1 ± 2.01	46.24 ± 0.13
S/R-Tol-BINAP (50:50)|S-Serine	-	46.68 ± 0.12
R-Tol-BINAP|R-Serine	36.4 ± 2.00	44.89 ± 0.11[a]
S-Tol-BINAP|R-Serine	40.1 ± 2.02	44.34 ± 0.03[a]

[a] The R-serine measurements suffer from the impurity issue. See the extensive discussions about this issue in Sections 6.3.2 and 6.3.3.

Since the controlled experiment with the racemic Tol-BINAP|water provides the same IFT value as those of R- or S-Tol-BINAP|water, we moved on to investigate the IFT

response of the S/R-Tol-BINAP (50:50)|S-Serine interface. From Figure 6.2, one can see that the IFT value of S/R-Tol-BINAP (50:50)|S-Serine sits roughly in the middle of the values of R-Tol-BINAP|S-Serine and S-Tol-BINAP|S-Serine all the way from t = 0 s to t = 5500 s. This is expected since one would have roughly 50% homochiral and 50% heterochiral pairs at the interface, providing further proof that the difference in the IFT values observed for the R-Tol-BINAP|S-Serine and S-Tol-BINAP|S-Serine can be attributed to chirality discrimination at the Organic|Water interface. The results showcase the power of our current approach in recognizing the chirality of the chiral molecule in the organic phase once the proper controlled experiments are performed.

Does the above chirality recognition phenomenon persist if one replaces the chirality of serine from S to R in aqueous solution? In the second set of experiments with the R-Tol-BINAP|R-Serine and S-Tol-BINAP|R-Serine, noticeable differences in the IFT values between the two systems were recorded, as demonstrated in Figure D4, Appendix D. It is worth noting that, the gap between the beginning (t = 0 s) and the end (t = 5500 s) are fairly consistent with a difference of around 0.55 mN/m. As one can see from Figure D4, Appendix D, the lower IFT reading was associated with S-Tol-BINAP|R-Serine, i.e., the heterochiral pair, instead of the homochiral pair as in the case of R-Tol-BINAP|S-Serine vs. S-Tol-BINAP|S-Serine. We attribute the somewhat different result to the impurity in R-Serine as discussed in the above section.

6.2.4 Effect of chiral discrimination at the binary interface with MD simulation

To further prove our interpretation of the phenomena observed experimentally, the theoretical IFT values of the corresponding systems were investigated using MD simulations. Two snapshots of the S-Tol-BINAP|S-Serine and R-Tol-BINAP|S-Serine systems and the density profiles of all components in each corresponding systems are shown in Figure 6.3. To appreciate the meaning of the snapshots, we first look at top figure in Figure 6.3a where the density profiles of all components in the S-Tol-BINAP|S-Serine

system are presented. As one can see, the serine density profile has two pronounced peaks at the intersections of the water and toluene density profiles. This result is consistent with the fact that the serine molecules were located at the two interfaces of Toluene|Water, shown in below figure in Figure 6.3a. The same is true for that of R-Tol-BINAP|S-Serine system, which is presented in Figure 6.3b. The snapshots of other systems are also provided in Figure D5, Appendix D, as well as the associated density distribution profiles in Figure D6, Appendix D.

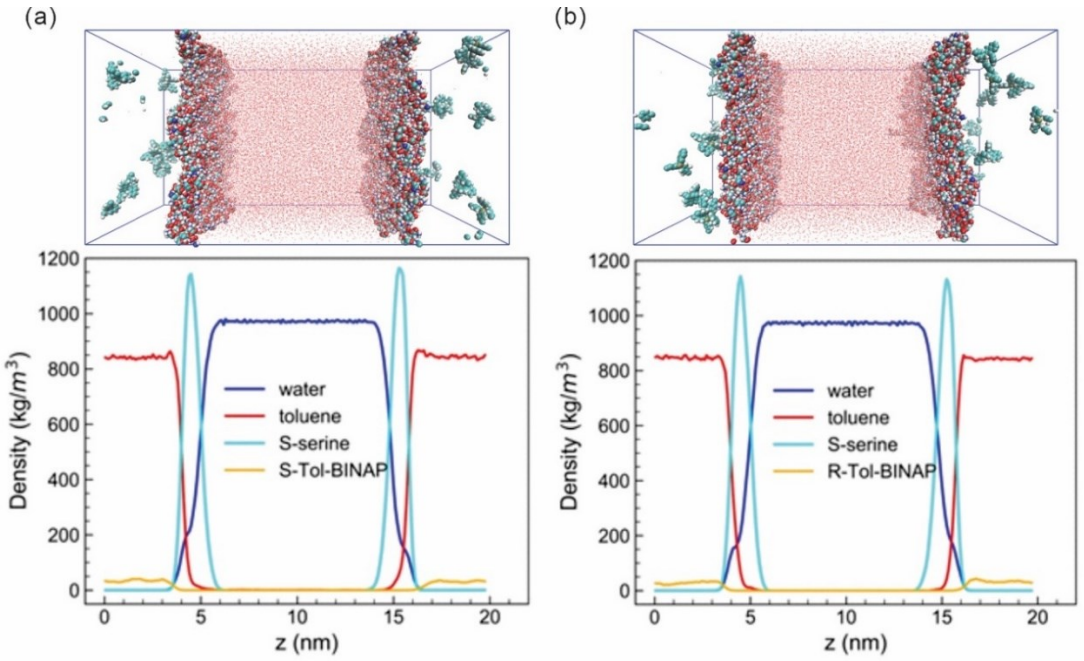


Figure 6.3. Final configurations (top) and the density profiles (bottom) averaged over the last 10 ns for all components in each corresponding system of (a) S-Tol-BINAP|S-Serine, (b) R-Tol-BINAP|S-Serine (Top: red line: water; cyan sphere: carbon; red sphere: oxygen; blue sphere: nitrogen; white sphere: hydrogen; toluene molecules are removed for clarity. Bottom: (blue: water, red: toluene, cyan: serine, orange: Tol-BINAP).

The IFT values calculated from the MD simulations are also shown in Table 6.3 to compare with the corresponding experimental values. For each system, the IFT value from simulations was lower than that from experiments, which might be attributed to the following facts. Firstly, the experimental IFT values have not yet reached the plateau and

were expected to be lowered with a much longer measurement time. In the simulations, the IFT values were obtained from the equilibrated systems where the interfacial films were formed and stable. Secondly, certain bulk concentrations of solutes were obtained in experiments, while the surface concentrations of solutes at the interface were difficult to be measured. The simulation systems were in nanoscale and focused on the interfacial regions so that the concentration of solutes at the interface might be higher than the experimental bulk concentration. Overall, the simulation results aimed to verify the trend between homo- versus heterochiral IFT values. For example, the experimentally observed IFT for homochiral pairs (S-Tol-BINAP|S-Serine) system was lower than that for heterochiral pairs (R-Tol-BINAP|S-Serine) system.

Looking closely at the MD simulation results, we counted how often the S-Tol-BINAP molecules are in close contact with the S-Serine or R-Serine molecules, and how often the R-Tol-BINAP molecules are in close contact with the S-Serine or R-Serine molecules at the interface as the simulation time goes on. A Tol-BINAP molecule and a serine molecule are said to be "in close contact" when the maximum distance between all atoms in the Tol-BINAP molecule and all atoms in the serine molecule was less than 0.35 nm, a cut-off distance frequently used for identifying H-bonding interactions[49] and other molecular contacts[50]. The number of close contacts during the last 30 ns of the simulations is presented in Figure 6.4a. Over the last 30 ns, the average number of contacts for S-Tol-BINAP|S-Serine was 0.98, while that of R-Tol-BINAP|S-Serine was 1.25. Clearly, the counts of the "close contacts" illustrates a preference for the formation of the heterochiral pairs than for the homochiral pairs at the interfaces.

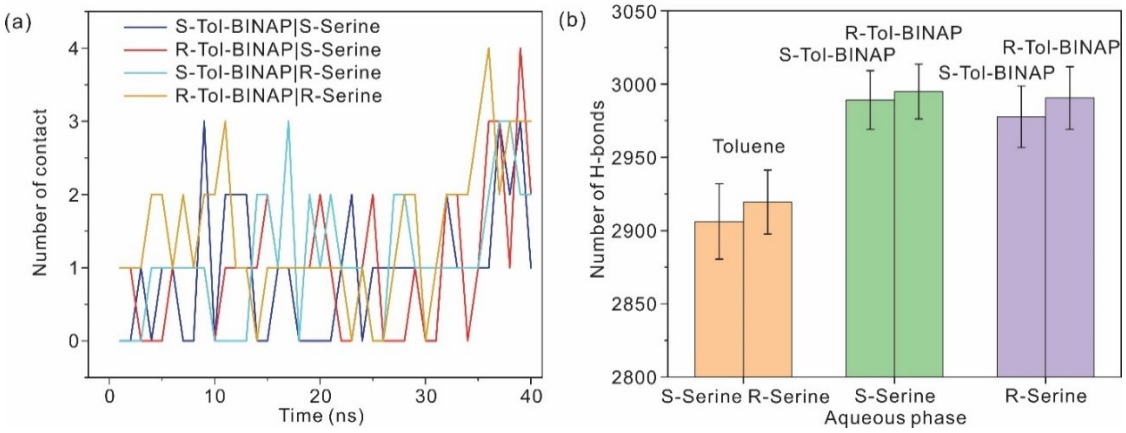


Figure 6.4. (a) Time evolution of the number of close contacts between the serine and Tol-BINAP molecules; (b) Number of H-bonds between the serine molecules; (Results in b are averaged over last 30 ns of simulation with the standard deviation shown as error bars.)

Within the narrow interfacial film (~1.2 nm), the number of hydrogen bonds between the serine molecules was also quantified and plotted in Figure 6.4b for six different simulation conditions: R-Serine or S-Serine in water combined with no Tol-BINAP, or R-Tol-BINAP or S-Tol-BINAP in the organic phase. Interestingly, the Serine molecules were found to form more hydrogen bonds when the Tol-BINAP molecules were present in the toluene phase, i.e., the cases of S-Tol-BINAP|S-Serine, R-Tol-BINAP|S-Serine, S-Tol-BINAP|R-Serine, R-Tol-BINAP|R-Serine, compared with the pure toluene cases, i.e., Toluene|S-Serine and Toluene|R-Serine. The decomposition of H-bonds within serine molecules are shown in Table 6.4. For each system, the number of H-bonds between $-COO^-$ and -OH has higher value than the other two types of H-bonds, while the value did not alter evidently when Tol-BINAP was added. In fact, the increase of number of H-bonds, as shown in Figure 6.4b, was attributed to the increased number of H-bonds between $-COO^-$ and $-NH_3^+$, and between $-NH_3^+$ and -OH. It indicated that the molecular assembly or arrangement of the Serine molecules at the interface was altered due to the presence of Tol-BINAP.

Table 6.4. H-bonds between different functionalities in serine molecules.

Sys. #	$-COO^-/-NH_3^+$	$-COO^-/-OH$	$-NH_3^+/-OH$
Toluene|S-Serine	276	1406	1212
S-Tol-BINAP|S-Serine	304	1408	1268
R-Tol-BINAP|S-Serine	299	1409	1284
Toluene|R-Serine	282	1404	1217
S-Tol-BINAP|R-Serine	301	1407	1263
R-Tol-BINAP|R-Serine	299	1406	1279

The IFTs for the six simulated systems are calculated and plotted as shown in Figure 6.5. The lowest IFT value of 35.1 ± 2.01 mN/m was found in S-Tol-BINAP|S-Serine. The IFT values for Toluene|S-Serine (41.3 ± 2.05 mN/m) and R-Tol-BINAP|S-Serine (41.5 ± 1.65 mN/m) were higher than the IFT for the S-Tol-BINAP|S-Serine system, agreeing with the experimentally observed trend. Overall, the MD models provide results in a very good agreement with the trends observed in the experiments with S-Serine.

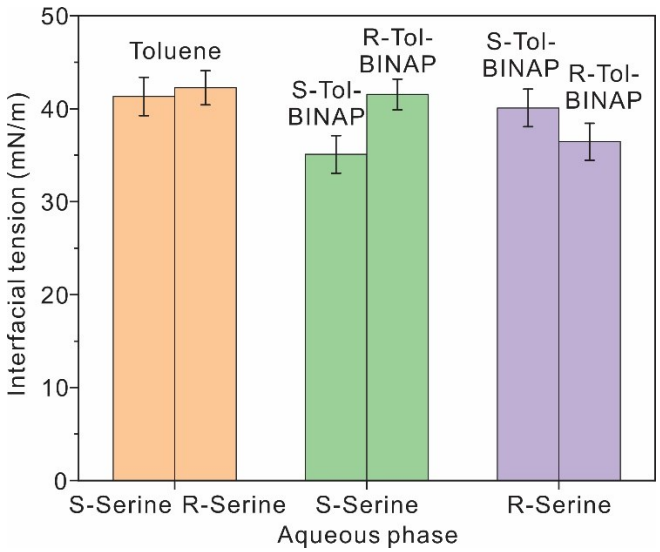


Figure 6.5. The IFT values of all six simulation systems. (Results are averaged over last 30 ns of simulation with the standard deviation shown as error bars.)

The MD simulations with R-Serine in water, also summarized in Table 6.3, produced results consistent with those predicted for S-Serine in water. For example, the homochiral R-Tol-BINAP|R-Serine has the lower IFT value of 36.4 ± 2.00 mN/m than the heterochiral S-Tol-BINAP|R-Serine which has a IFT value of 40.1 ± 2.02 mN/m. Furthermore, the IFT values calculated for the homochiral systems, i.e., R-Tol-BINAP|R-Serine and S-Tol-BINAP|S-Serine, are the same within the error bar, and the same conclusion can be drawn for the calculated IFT values of the heterochiral systems, i.e., R-Tol-BINAP|S-Serine and S-Tol|-BINAPR-Serine. From both the theoretical and experimental work discussed above, one can conclude confidently that different degrees of interactions are present in the homochiral and heterochiral systems, and that chirality discrimination exists at liquid/liquid interfaces. Since the experimental results with S-Serine, the naturally occurring enantiomer, are well predicted by the MD simulations, we take this as the indication that the S-Serine sample purchased has less impurity than the R-Serine sample. Therefore, further comparisons with simulations and the final conclusion will be based on the experimental data with S-Serine.

From the plots in Figure 6.4, one can see that there are many factors which may attribute to the final behaviors of the toluene/water interfaces studied. Generally, the final IFT value at an Organic|Water interface is controlled by the combined effects of the 'internal' interactions at the interfaces between serine molecules as discussed in Figure 6.4b, the 'external' interactions between serine and water and between serine and Tol-BINAP (Figure 6.4a). For example, more contacts at the interfaces between serine and Tol-BINAP in R-Tol-BINAP|S-Serine (the heterochiral system) than in S-Tol-BINAP|S-Serine (the homochiral system) decrease the adhesion between toluene and water, leading to higher IFT readings of the heterochiral system than the homochiral one, which were observed experimentally. On the other hand, the situation is more complex when comparing the IFT outcome for the Toluene|S-Serine, S-Tol-BINAP|S-Serine and R-Tol-BINAP|S-Serine systems. The higher number of the internal serine-serine interactions may cause a denser

serine layer and lead to the reduction of IFT in S-Tol-BINAP|S-Serine and R-Tol-BINAP|S-Serine when they are compared with Toluene|S-Serine. The combined effect of these two opposite factors of the serine-serine and serine-Tol-BINAP interactions results in a slightly higher IFT value of R-Tol-BINAP|S-Serine in comparison to that of Toluene|S-Serine, an experimental outcome which was also captured theoretically. For the S-Tol-BINAP|S-Serine homochiral system, the increase in the serine-Tol-BINAP interactions outweighs the increase in the serine-serine interactions, leading to a much lower IFT value of S-Tol-BINAP|S-Serine, also observed experimentally and theoretically.

6.2.5 Effect of concentration on IFT

In addition, we also investigated chirality discrimination phenomena at lower S-Serine concentrations, i.e., 6wt% and 4wt%. All the averaged IFT values for the observed systems are listed in Table D1, Appendix D, while the corresponding IFT plots are provided in Figure 6.6. As can be seen in Figure 6.6a and 6.6b, the heterochiral R-Tol-BINAP|S-Serine systems consistently show higher IFT values than the corresponding homochiral S-Tol-BINAP|S-Serine. Clearly, the same chirality discrimination trend at the Organic|Water interface observed with the 9wt% S-Serine were also observed under these more diluted conditions. Quantitatively, the lower serine concentrations brought about fewer internal/external interactions and consequently smaller IFT gaps between the homochiral S-Tol-BINAP|S-Serine and the heterochiral R-Tol-BINAP|S-Serine pairs. As shown in Figure 6.6, the same drifting behavior, i.e., a slow rate to reach a plateau, persists under the dilute conditions. The negative slopes become flatter, approaching to the equilibrium after 5000 s. Moreover, the IFT value of the racemic S/R-Tol-BINAP (50:50)|S-Serine falls roughly in the middle of the homochiral S-Tol-BINAP|S-Serine and the heterochiral R-Tol-BINAP|S-Serine traces. In summary, the IFT observations under the dilute conditions support the same conclusion of chirality discrimination at the Organic|Water interface achieved with the 9wt% serine solutions described above.

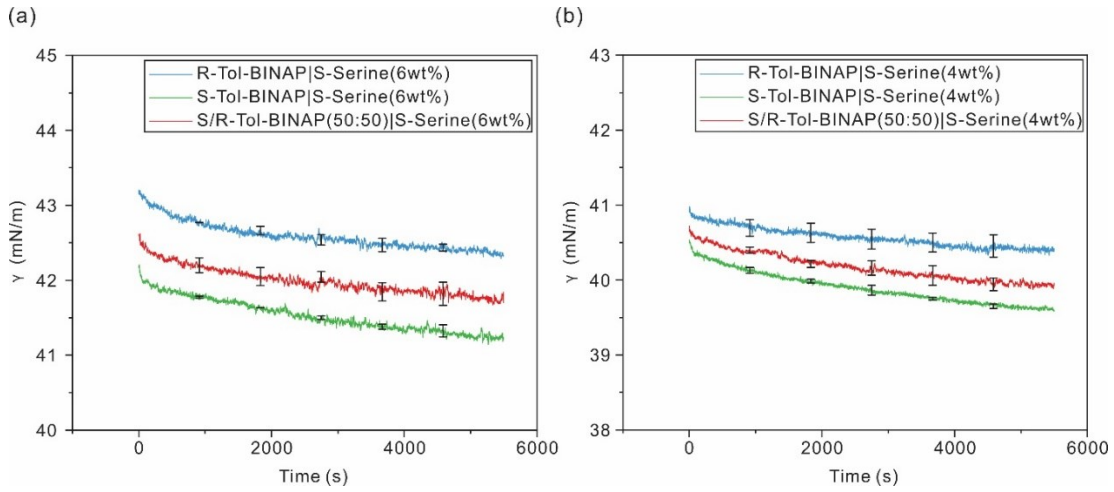


Figure 6.6. The IFT plot vs. time for S-Tol-BINAP|S-Serine,R-Tol-BINAP|S-Serine and S/R-Tol-BINAP (50:50 v/v)|S-Serine systems. (a) Under 6 wt% S-serine solution, (b) Under 4 wt% S-serine solution.

6.3 Conclusions

In this work, we demonstrated the general capability of interfacial tension measurements complemented with MD simulations in revealing stereochemical recognition of chiral-chiral molecular interactions at the Organic|Water binary interfaces. With S-serine in the aqueous solution, the S-Tol-BINAP molecules brought about a lower IFT value than the R-Tol-BINAP molecules, indicating a stronger intermolecular interaction for the homochiral S-serine and S-Tol-BINAP pairs at the interface. The chirality discrimination experiment with the racemic mixture of S/R-Tol-BINAP produced an IFT value which is roughly the average of the IFT values of the homochiral S-Tol-BINAP|S-Serine and heterochiral R-Tol-BINAP|S-Serine, firmly establishing that the IFT differences observed are due to chirality discrimination at the Organic|Water interface. Importantly, the associated MD simulations correctly predicted the experimental trend in the IFT difference between the R-Tol-BINAP and S-Tol-BINAP with S-serine. The MD simulations revealed that such outcome can be attributed to the combined effects of the internal interactions between serine molecules, the external interactions between serine and water, and the close contact interactions of the serine and Tol-BINAP molecules at the

interface.

Furthermore, our results demonstrated that the selectivity and discrimination capabilities are very much dependent on the structural configuration of the chiral species, for example whether serine exists as zwitterions or as neutral molecules. Based on the observation, the dominant serine species in aqueous phase is the zwitterion form, consistent with the previous observation reported in the vibrational optical activity spectroscopic study[43]. In the current report, the experiments were carried out under neutral pH environment. It would also be considerable interest to explore how the change of pH values, which will alter the dominant serine species and its charge state, influences the outcome of chirality discrimination at the Organic|Water interface, in addition to varying the shape of chiral molecules in the organic phase.

As demonstrated in the current study, chirality discrimination effects of the homo- and heterochiral Tol-BINAP|Serine pairs at the Organic|Water interfaces are captured macroscopically by using IFT measurements, with the further support of MD simulation. Such a combination offers a convenient way to monitor chirality discrimination process at interfaces, which has not been demonstrated before to the best of our knowledge. As highlighted in several recent reports and reviews [17,51-53], there are substantial interest in applying chiral discrimination at interfaces in the fields of electrochemical detection and biochemical sensing, and supramolecular chemistry for controlling properties of chiral aggregates. The current approach can be applied to examine some of these interfaces, for example those containing β-cyclodextrin or calixarenes and amino acids to gain insight into how different factors affect the chirality discrimination outcome. In summary, the current research approach developed in this study is fairly general and can be used to probe chirality discrimination at interfaces between many different interesting chiral molecular pairs, and to explore potentially very different degree of chirality discrimination at interfaces in comparison to in bulk.

6.4 Materials and methods

6.4.1 Materials

The liquid-liquid IFT measurements were carried out using a ramé-hart 250 series tensiometer instrument, equipped with the DROPimage software. All chemicals including S-Serine (≥99%), R-Serine (≥98%), S-(-)-2,2'-Bis(di-p-tolylphosphino)-1,1'-binaphthyl (Tol-BINAP) and R-(+)-Tol-BINAP were purchased from Sigma Aldrich and used without further purification. Their molecular structures are shown in Figure 6.7a and 6.7b. The racemic (S/R)-Tol-BINAP solution was prepared by mixing the pure S-Tol-BINAP and R-Tol-BINAP solution with a volume ratio of 50:50. Water for all solution preparation was first distilled, then deionized to a resistance of 18 MΩ-cm (Barnstead E-Pure system), and finally filtered through a 0.2 μm filter to remove any particulate matter. It is noted that the solution pH values were measured at the beginning and at the end of all the IFT measurements, and the pH values remain a consistent value of 5.8 for serine and 6.8 for deionized water. Therefore, the CO_2 intake did not play the dominant effect on the pH of the solution.

6.4.2 IFT measurements

The experimental setup and steps for IFT measurements have been reported in the previous literature[54-56], which is presented in Figure 6.7c. Briefly, the oil sample was first loaded into a syringe which was connected to a U-shape needle. The tip of the needle was then placed into a quartz cell (45 × 30 × 45 mm) filled with deionized water acting as a surrounding medium. An oil droplet was generated at the tip of the needle and was subsequently illuminated by means of a fiber optic illuminator. Finally, the shape of the droplet was captured by a CCD camera and analyzed by using the DROPimage software. The instrument was calibrated by using droplets of pure solvents, such as toluene or heptane, whose IFT values at their respective organic-water interfaces are well-documented in the literature. The potential interference from sources, such as external

vibration, thermal instability, and surface-activity impurities from the needle assembly was removed by the calibration procedure. In addition, an example schematic image of a chirality discrimination event facilitated by a homochiral and a heterochiral Tol-BINAP and serine pair at the molecular level is provided in Figure 6.7d. The measurement plots presented in the figures are the averaged IFT values out of two independent runs of a specific system and the variations between each run are presented as the error bar in each figure.

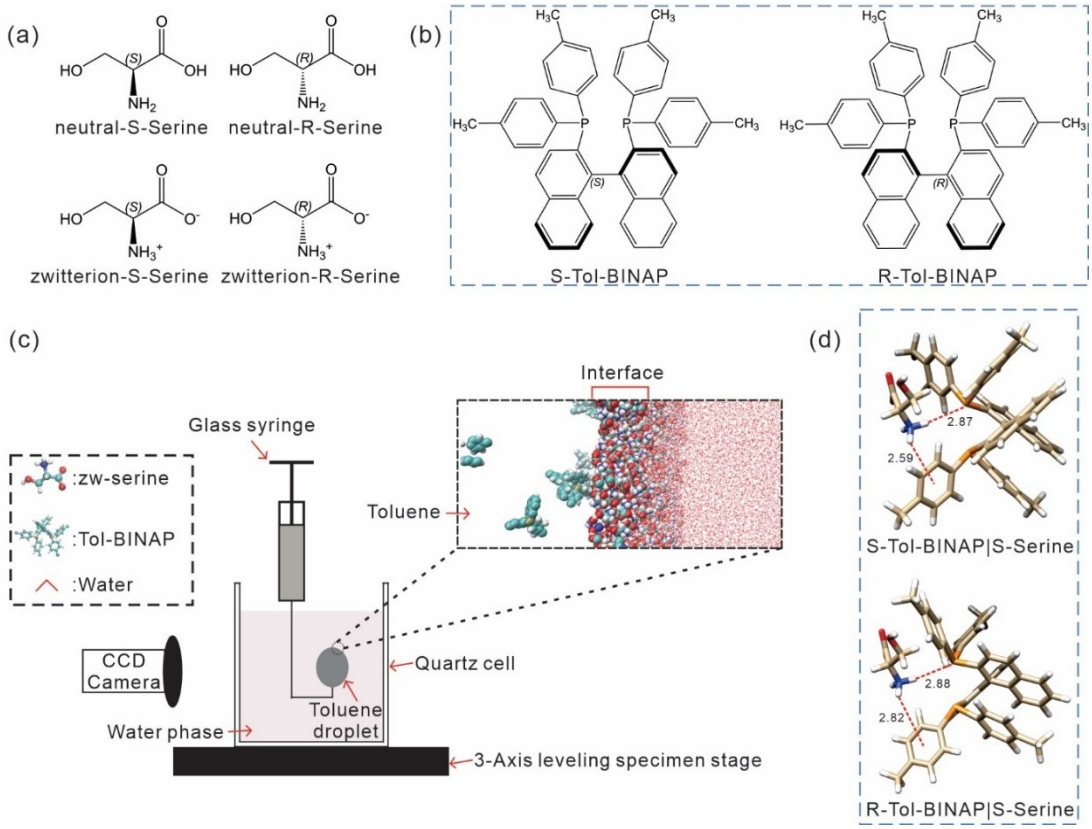


Figure 6.7. Molecular structures of (a) S-serine and R-serine in neutral and zwitterion form, and (b) S-Tol-BINAP and R-Tol-BINAP. (c) Schematic of typical experimental setup using pendant drop method and the atomic level illustration of the interface with the aid of MD simulation. (d) An example schematic image of the local arrangements of a homochiral and a heterochiral Tol-BINAP and serine pair for illustrating a chiral discrimination event which may happen at the interface.

6.4.3 MD simulations

The molecular structures for serine and Tol-BINAP molecules were built in Avogadro[57]. The geometries of these two molecules were first optimized by density functional theory (DFT) at the B3LYP[58]/6-31G(d,p) level with water as a solvent environment, using the Gaussian 16 electronic structure calculation package[59]. Then Automated Topology Builder (ATB)[60] was used to generate force field parameters compatible with the GROMOS force field parameter set 54A7[61]. Partial charges were calculated by using charges from electrostatic potentials using a grid-based method (CHELPG)[62] which is implemented in Gaussian 16. The force field parameters of toluene were validated in our previous works and directly adopted[63-65] and the single-point-charge (SPC) model was used for water. MD simulations were performed in the GROMACS package version 5.0.7[66-68]. A total of 33046 water molecules were placed in a 10×10×10 nm^3 box, then 1389 serine molecules were placed inside the water box and near the two surfaces in the x-y plane. The simulation box was then extended in the z-direction to the dimension of 10×10×20 nm^3 with the water box sheltered in the middle. The extended, surrounding space was directly solvated with pre-equilibrated toluene molecules if the organic phase was pure toluene. For systems with Tol-BINAP, the solute molecules were inserted randomly in the extended space and then solvated by toluene molecules. Each system underwent energy minimization using the steepest descent method, an NVT equilibration for 100 ps with position-restraints on the heavy atoms, and finally the production simulation. Temperature was controlled at 300 K by a velocity rescaling thermostat with a time constant (τ_T) of 0.1 ps. Pressure coupling was applied by using the Parrinello-Rahman[69] barostat with a time constant (τ_p) of 1.0 ps. LINCS[70] algorithm, Particle Mesh Ewald method for full electrostatics,[71] and the periodic boundary conditions in all directions were applied. The cut-offs for the van der Waals and electrostatic interactions were both 1.4 nm. During the production simulation, the $NP_{normal}AT$ ensemble was used, which applied a prescribed normal pressure in the z direction with a constant surface area in the x-y plane. All simulations had a time step of 2 fs and a total time of 40

ns. The IFT value was calculated using the following equations[72]:

$$\gamma = \frac{1}{2}(p_{zz} - \frac{p_{xx}+p_{yy}}{2})L_z \qquad (1)$$

where p_{xx}, p_{yy}, and p_{zz} are the diagonal components of the pressure tensor in the x, y and z directions, respectively, and L_z is the box length in the z direction.

6.5 Acknowledgements

This work was funded by the Natural Sciences and Engineering Research Council of Canada, the Future Energy Systems under the Canada First Research Excellence Fund, the Canada Foundation for Innovation, the Alberta Advanced Education & Technology Small Equipment Grants Program, the University of Alberta, and Compute Canada. HZ and YX are Tier I Canada Research Chair in Intermolecular Forces and Interfacial Science and in Chirality and Chirality Recognition, respectively.

References

[1] A. Zehnacker, M. A. Suhm, *Angew. Chem. Int. Ed.* **2008**, *47*, 6970-6992; *Angew. Chem.* **2008**, *120*, 7076-7100.

[2] F. Xie, N. A. Seifert, W. Jäger, Y. Xu, *Angew. Chem. Int. Ed.* **2020**, *59*, 15703-15710; *Angew. Chem* **2020**, *132*, 15833-15840.

[3] F. Xie, M. Fusè, A. S. Hazrah, W. Jäger, V. Barone, Y. Xu, *Angew. Chem. Int. Ed.* **2020**, *59*, 22427-22430; *Angew. Chem.* **2020**, *132*, 22613-22616.

[4] K. Hirata, Y. Mori, S. Ishiuchi, M. Fujii, A. Zehnacker, *Phys. Chem. Chem. Phys.* **2020**, *22*, 24887-24894.

[5] F. X. Sunahori, G. Yang, E. N. Kitova, J. S. Klassen, Y. Xu, *Phys. Chem. Chem. Phys.* **2013**, *15*, 1873-1886.

[6] T. R. Hoye, C. S. Jeffrey, F. Shao, *Nat. Protoc.* **2007**, *2*, 2451-2458.

[7] B. H. Pate, L. Evangelisti, W. Caminati, Y. Xu, J. Thomas, D. Patterson, C. Perez, M. Schnell, *Advances in Spectroscopy, Chromatography and Emerging Methods*, Second Edition; ed. P. L.Polavarapu, Elsevier, **2018**, pp 679-729.

[8] T. J. Ward, K. D. Ward, *Anal. Chem.* **2012**, *84*, 626-635.

[9] K. Ariga, *Phys. Chem. Chem. Phys.* **2020**, *22*, 24856-24869.

[10] H. Sato, *Phys. Chem. Chem. Phys.* **2020**, *22*, 7671-7679.

[11] S. Arnaboldi, T. Benincori, A. Penoni, L. Vaghi, R. Cirilli, S. Abbate, G. Longhi, G. Mazzeo, S. Grecchi, M. Panigati, P. R. Mussini, *Chem. Sci.* **2019**, *10*, 2708-2717.

[12] R. McKendry, M. E. Theoclitou, T. Rayment, C. Abell, *Nature* **2019**, *–391*, 566-568.

[13] H. Guo, J. Kim, S. Kim, S. Chang, W. Kim, *Langmuir* **2009**, *25*, 648-652.

[14] P. Lopes, R. Kataky, *Anal. Chem.* **2012**, *84*, 2299-2304.

[15] A. Shundo, K. Hori, T. Ikeda, N. Kimizuka, K. Tanaka, *J. Am. Chem. Soc.* **2013**, *135*, 10282-10285.

[16] K. Adachi, K. Chayama, H. Watarai, *Chirality* **2006**, *18*, 599-608.

[17] Y. Iimura, M. Fukuyama, A. Hibara, M. Harada, T. Okada, *J. Colloid Interface Sci.* **2017**, *508*, 469-475.

[18] Y. Umemura, A. Yamagishi, R. Schoonheydt, A. Persoons, F. D. Schryver, *J. Am. Chem. Soc.* **2002**, *124*, 992-997.

[19] R. H. A. Ras, Y. Umemura, C. T. Johnston, A. Yamagishi, R. A. Schoonheydt, *Phys. Chem. Chem. Phys.* **2007**, *9*, 918-932.

[20] A. G. Gaonkar, *JAOCS* **1989**, *66*, 1090-1092.

[21] G. T. Vladisavlević, M. Shimizu, T. Nakashima, *J. Membr. Sci.* **2006**, *284*, 373-383.

[22] P. Tavakoli, S. Z. Shadizadeh, F. Hayati, M. Fattahi, *Petroleum* **2020**, *6*, 293-303.

[23] C. Shi, B. Yan, L. Xie, L. Zhang, J Wang, A. Takahara, H. Zeng, *Angew. Chem. Int. Ed.* **2016**, *55*, 15017-15021; *Angew. Chem.* **2016**, *128*, 15241-15245.

[24] B. Yan, X. Wang, X. Zhang, S. Liu, M. Li, R. Ran, *J. Environ. Chem. Eng.* **2021**, *9*, 106607.

[25] D. Wang, D. Yang, C. Huang, Y. Huang, D. Yang, H. Zhang, Q. Liu, T. Tang, M. G. El-Din, T. Kemppi, B. Perdicakis, H. Zeng, *Fuel* **2021**, *286*, 119390.

[26] S. Reynaert, P. Moldnaers, J. Vermant, *Langmuir* **2006**, *22*, 4936-4945.

[27] H. Watarai, N. Sawada, T. Sawada (eds.), *Interfacial Nanochemistry*, Kluwer Academic/Plenum Publishers, New York, **2005**.

[28] A. G. Volkov, D. W. Deamer, D. L. Tanelian, V. S. Markin (eds.), *Liquid Interfaces in Chemistryand Biology*, Wiley, New York, **1998**.

[29] M. Kaminski, A. Kudelski, M. Pecul, *J. Phys. Chem. B* **2012**, *116*, 4976-4990.

[30] M. Losada, H. Tran, Y. Xu, *J. Chem. Phys.* **2008**, *128*, 014508.

[31] T. Lan, P. Wu, Z. Liu, M. Stroet, J. Liao, Z. Chai, A. E. Mark, N. Liu, D. Wang, *Environ. Sci. Technol.* **2022**, *56*, 917-927.

[32] X. Sun, D. Yang, H. Zhang, H. Zeng, T. Tang, *J. Phys. Chem. B* **2021**, *125*, 7556-7567.

[33] C. Jian, M. R. Poopari, Q. Liu, N. Zerpa, H. Zeng, T. Tang, *J. Phys. Chem. B* **2016**, *120*, 5646-5654.

[34] T. Arakawa, K. Tsumoto, Y. Kita, B. Chang, D. Ejima, *Amino Acids* **2007**, *33*, 587-605.

[35] D. A. Evans, R. J. Thomson, *J. Am. Chem. Soc.* **2005**, *127*, 10506-10507.

[36] D. Ferraris, B. Young, C. Cox, W. J. Drury III, T. Dudding, T. Lectka, *J. Org. Chem.* **1998**, *63*, 6090-6091.

[37] S. K. Burley, G. A. Petsko, *Science* **1985**, *229*, 23-28.

[38] Z. Dezhahang, C. Merten, M. R. Poopari, Y. Xu, *Dalton Trans.* **2012**, *41*, 10817-10824.

[39] A. Benderrag, M. Daaou, B. Bounaceur, B. Haddou, *Chemical Papers* **2016**, *70*, 1196-1203.

[40] J. Drelich, C. Fang, C. L. White, *Encycl. Surf. Colloid Sci.* **2002**, *3*, 3152-3166.

[41] S. P. Pinho, *J. Chem. Eng. Data* **2008**, *53*, 180-184.

[42] V. Scutelnic, M. A. S. Perez, M. Marianski, S. Warnke, A. Gregor, U. Rothlisberger, M. T. Bowers, C. Baldauf, G. von Helden, T. R. Rizzo, J. Seo, *J. Am. Chem. Soc.* **2018**, *140*, 7554-7560.

[43] P. Zhu, G. Yang, M. R. Poopari, Z. Bie, Y. Xu, *ChemPhysChem* **2012**, *13*, 1272-1281.

[44] C. Jian, M. R. Poopari, Q. Liu, N. Zerpa, H. Zeng, T. Tang, *Energy Fuels* **2016**, *30*, 10228-10235.

[45] E. R. A. Lima, B. M. de Melo, L. T. Baptista, M. L. L. Paredes, *Braz. J. Chem. Eng.* **2013**, *30*, 55-62.

[46] R. I. Slavchov, J. K. Novev, *J. Colloid Interface Sci.* **2012**, *387*, 234-243.

[47] C. J. Beverung, C. J. Radke, H. W. Blanch, *Biophys. Chem.* **1999**, *81*, 59-80.

[48] Z. M. Aman, E. P. Brown, E. D. Sloan, A. K. Sum, C. A. Koh, *Phys. Chem. Chem. Phys.* **2011**, *13*, 19796-19806.

[49] M. Abraham, D. van der Spoel, E. Lindahl, B. Hess, *GROMACS User Manual*, Version 2016, **2018**.

[50] T. Lan, H. Zeng, T. Tang, *J. Phys. Chem. C* **2019**, *123*, 22989-22999.

[51] X. Huang, K. Quan, D. Pei, J. Liu, D. Di, *Chirality* **2018**, *30*, 974-981.

[52] Y. Sun, D. Zhang, Y. Sheng, D. Xu, R. Zhang, M. Bradley, *Anal. Methods* **2021**, *13*, 2011-2020.

[53] W. Xu, M. Cheng, S. Zhang, Q. Wu, Z. Liu, M. K. Dhinakaran, F. Liang, E. G. Kovaleva, H. Li, *Chem. Commun.* **2021**, *57*, 7480-7492.

[54] E. Y. Arashiro, N. R. Demarquette, *Mater. Res.* **1999**, *2*, 23-32.

[55] C. E. Stauffer, *J. Phys. Chem.* **1965**, *69*, 1933-1938.

[56] J. D. Berry, M. J. Neeson, R. R. Dagastine, D. Y. C. Chan, R. F. Tabor, *J. Colloid Interface Sci.* **2015**, *454*, 226-237.

[57] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, G. R. Hutchison, *J. Cheminform* **2012**, *4*, 17.

[58] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648-5652.

[59] Gaussian 16 (Revision B.01), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, D. J. Fox, Gaussian, Inc., Wallingford CT, **2016**.

[60] A. K. Malde, L. Zuo, M. Breeze, M. Stroet, D. Poger, P. C. Nair, C. Oostenbrink, A. E. Mark, *J. Chem. Theory Comput.* **2011**, *7*, 4026-4037.

[61] C. Oostenbrink, A. Villa, A. E. Mark, W. F. van Gunsteren, *J. Comput. Chem.* **2004**, *25*, 1656-1676.

[62] C. M. Breneman, K. B. Wiberg, *J. Comput. Chem.* **1990**, *11*, 361-373.

[63] X. Sun, H. Zeng, T. Tang, *J. Mol. Liq.* **2020**, *310*, 113248.

[64] T. Lan, H. Zeng, T. Tang, *J. Phys. Chem. C* **2019**, *123*, 22989-22999.

[65] X. Sun, H. Zeng, T. Tang, *J. Colloid Interface Sci.* **2021**, *586*, 766-777.

[66] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. C. Berendsen, *J. Comput. Chem.* **2005**, *26*, 1701-1718.

[67] E. Lindahl, B. Hess, D. van der Spoel, *J. Mol. Model* **2001**, *7*, 306-317.

[68] H. J. C. Berendsen, D. van der Spoel, R. van Drunen, *Comput. Phys. Commun* **1995**, *91*, 43-56.

[69] H. Saito, H. Nagao, K. Nishikawa, K. Kinugawa, *J. Chem. Phys.* **2003**, *119*, 953-963.

[70] B. Hess, *J. Chem. Theory Comput.* **2008**, *4*, 116-122.

[71] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577-8593.

[72] A. R. van Buuren, S. J. Marrink, H. J. C. Berendsen, *J. Phys. Chem.* **1993**, *97*, 9206-9212.

Chapter 7

Conclusions and Future Work

7.1 Conclusion

In my PhD thesis research, I have focused on non-covalent interactions involving several chiral molecules using two main different approaches. The first one is by using the VCD and ROA spectroscopies complemented with theoretical calculations to probe the intermolecular interactions between chiral solute molecules and between chiral solute and achiral solvent molecules. Owing to the additional sensitivity of VOA spectroscopy to small conformational changes, I was able to characterize the conformational distributions of a chiral acid, THFA, in cold rare gas matrices, and in water and in methanol solutions for a more flexible chiral molecule, NALC. In the second approach, I utilized interfacial force measurements to probe the non-covalent interactions between homochiral and heterochiral molecular pairs at the organic|water binary interfaces and to extract information about the chirality discrimination behavior from the experimental data and the MD simulations.

In Chapter 3, I examined the MI-IR and MI-VCD spectral features of THFA in a cold rare gas matrix under three different deposition temperatures, i.e., 10 K, 24 K and 30 K. The purpose of this study was to investigate the conformations of THFA and also its self-aggregation behavior by increasing the temperature of the cold rare gas matrix in a gradual manner. The sharp IR and VCD features in cryogenic matrices allowed us to decode spectral information confidently. At 10 K, the THFA molecules exist mainly in the monomeric form in the cold matrices. The sharp experimental MI-IR and MI-VCD spectral features agree well with the theoretical simulation of the THFA monomeric conformers. At 10 K, M1 and M2 which adopt the *trans*-COOH configuration and differ only slightly in their ring puckering configurations are the dominant species in the experiment, while M3 (in the *cis*-COOH configuration) was found to exist in a small percentage. From the good agreement between the experiment and theory, one could further derive those the higher energy monomers experienced the conformational relaxation to the most stable *trans*-COOH and *cis*-COOH conformations, showing a somewhat different distribution

compared to the solution study.[1] As the matrix temperature increased, more THFA molecules aggregated, leading to some band broadening and new IR and VCD features. Extensive conformational searches were performed for the THFA dimer, by using a powerful conformational search code, CREST.[2] Several hundred CREST candidates were further optimized at the DFT level and sorted into three different groups, Types 1, 2 and 3. Type 1 structures consist of two *cis*-COOH THFA monomers formed in a double cyclic hydrogen bonded ring and are the most stable ones among the three types. Type 2 and 3 structures are made of one *cis*-COOH and one *trans*-COOH THFA, and two *trans*-COOH THFA monomers, respectively. They are about 15 to about 22 kJ mol^{-1} less stable than Type 1, which is the global minimum. A surprising outcome is that the experimental MI-IR and MI-VCD spectra at higher deposition temperature match those of Type 3 dominantly, with minor contributions from Type 1 and 2, which is very different from what was found in chloroform[1] or what was found in a supersonic jet expansion[3] where Type 1 and Type 2 dominate, respectively. With the unique cold rare gas matrix environment, one could capture the strongly kinetically favored formation mechanism. Therefore, this study showcases that the power of MI-VCD spectroscopy in revealing the rare structures formed in a cold rare gas matrix.

In Chapter 4, the neutral and deprotonated forms of THFA were studied in aqueous solution under different pH environments. To evaluate the main intermolecular interactions between the neutral form and the deprotonated form of THFA molecules in water, the "clusters-in-a-liquid" solvation model was used to emphasize the long-lived chiral solute···$water_n$ complexes. The conformational searches were done using the CREST conformational search code and were followed by the DFT geometry optimization and harmonic frequency calculations. In acidic conditions, the *cis*-COOH configurations of THFA were identified as the most dominant species for the neutral form of THFA in water. The significant improvement in the agreement between the experimental spectral features and theoretical simulation was achieved as one water molecule was added explicitly at the

potential binding sites around the neutral THFA molecules, supporting the hypothesis of the long-lived solute-solvent clusters. A similar approach was also applied for the deprotonated form of THFA in water. The main experimental IR features show very good agreement with the simulation results while the VCD features demonstrate an ok agreement. These comparisons suggest that the 1:1 solute-solvent complexes are the long-lived solute-solvent clusters in the experiment. Additional water molecules were added around THFA and the agreement with the experimental results did not improve, indicating that the hydrates of deprotonated THFA are not the dominant species. The current study highlights the power of using VCD spectroscopy to extract structural information for chiral molecules in aqueous solutions.

In Chapter 5, a similar strategy as in Chapter 4 was applied to account for solvation effects in water and in methanol for NALC, a highly flexible chiral ester. It is interesting to point out that very similar experimental ROA spectra were obtained in these two solvents, indicating that the long-lived solute···solvent interactions in two solvents are basically the same. Two different solvation methods, i.e., the "clusters-in-a-liquid" solvation model and the AIMD simulations, were utilized. The theoretical conformational searches reveal that NALC-I is the most dominant conformation in both water and methanol and adopts the *trans*-COOH configuration. The results based on the "clusters-in-a-liquid" solvation model provided a noticeable improvement when one solvent molecule, i.e., water or methanol were added explicitly at the possible hydrogen binding sites of NALC. Addition of more explicit solvent molecules to NALC provided no further improvement in the agreement between the experiment and theoretical simulation. On the contrary, some of the intense multi-signated signals generated were not observed in the experiment. The results further validated the "clusters-in-a-liquid model" where consideration of just one explicit solvent molecule tends to provide a satisfactory prediction. In the AIMD approach, the simulated Raman and ROA spectra show some improvements in the agreement with experiment at the lower wavenumber region and in the carbonyl region compared to the static harmonic

approaches. Such improvements are attributed to the inclusion of some anharmonicities and the bulk environment in the AIMD method. A drawback is that currently AIMD requires significant computational resources to complete.

In Chapter 6, the potential of IFT measurements for characterization of chirality discrimination/recognition at a binary organic|water interface was explored. By judicially selecting the appropriate chiral solutes and designing the essential control experiments, noticeable differences in the IFT values were observed for the homochiral versus heterochiral pairs at the binary organic|water interface, demonstrating that the IFT measurements can be used to characterize chirality-controlled properties at liquid-liquid interfaces. Specifically, a lower IFT value was observed experimentally with S-serine in the aqueous solution and S-Tol-BINAP in organic solution (i.e., the homochiral pair), which indicates a stronger intermolecular interaction for the homochiral S-serine and S-Tol-BINAP pairs at the interface. The experimental results were further supported by the associated molecular dynamics simulations where the same IFT trend was predicted. Such an outcome was further analyzed and could be attributed to the mutual effects of the internal interactions between serine molecules, the external interactions between serine and water, and the close contact interactions of the serine and Tol-BINAP molecules at the interface. Overall, chirality discrimination effects of the homo- and heterochiral Tol-BINAP|serine pairs at the organic|water interfaces were captured macroscopically by using IFT measurements for the first time. The current study provides new insight on discrimination/recognition of many different chiral molecular pairs at the organic|water interface.

7.2 Future Work

Further investigations of several aspects of the current thesis research could be beneficial. First, in the research activities conducted in Chapter 3 and 5, I tested the performance of several common basis sets used in dealing with solute···solute and

solute···solvent non-covalently bonded complexes. It appeared that there were some (noticeable) changes in the relative free energies, Boltzmann distribution factor, vibrational frequencies, and intensities among different basis sets. It would be of interest to do further systematic performance tests with a larger range of different basis sets and functionals to understand the effects better. Moreover, it would be of great interest to consider the anharmonicity effect in the theoretical simulations. A few recent studies demonstrated the improvement on the band pattern and intensity ratio in the calculated VOA spectra by including anharmonicity.[4]

The current and previous MI-VCD studies demonstrated its power in extracting structural information and in capturing some unusual structures formed under the cold rare gas matrix condition as well as the chirality transfer behavior.[5] To investigate a wider range of chemical systems, it would be worthwhile to introduce a second independent sample deposition line onto the cold head. In this case, one would be able to introduce the chiral solute and achiral solvent simultaneously in separate lines, allowing a greater control of the composition. Such a set up could provide us additional aspects on investigating the solute···solvent and mixed solute···solute interactions.

In Chapter 6, we demonstrated the capability of evaluating chirality discrimination between the different chiral-chiral molecular pairs at the organic|water binary interfaces using IFT measurements. Since this is a first study of the kind, detailed mechanisms for the observations are still not clear. It would be worthwhile to further study the detail mechanism for the significant increase of the IFT tension when zwitterionic forms of amino acid molecules are present in the aqueous phase. Moreover, it would also be of considerable interest to explore how the change of pH, which will alter the dominant serine species and its charge state, would influence the outcome of chirality discrimination at the organic|water interfaces, in addition to varying the shape of chiral molecules in the organic phase. Lastly, it would also be of interest to investigate how serine and Tol-BINAP molecules interact in the bulk by using VOA spectroscopy[6] and electronic circular

dichroism[7] and to compare to their binding preferences at the interfaces.

References

[1] T. Kuppens, W. Herrebout, B. Veken, P. Bultinck, *J. Phys. Chem. A* **2006**, *110*, 10191-10200.

[2] S. Grimme, C. Bannwarth, P. Shushkov, *J. Chem. Theory Comput.* **2017**, *13*, 1989-2009.

[3] F. Xie, N. A. Seifert, W. Jäger, Y. Xu. *Angew. Chem. Int. Ed*. **2020**, *59*, 15703-15710; *Angew. Chem.* **2020**, *132*, 15833-15840

[4] (a) P. T. Panek, C. R. Jacob, *Phys. Chem. Lett.* **2016**, *7*, 3084-3090. (b) K. Ziadi, *J. Raman Spectrosc.* **2021**, *53*, 222-236. (c) J. Bloino, M. Biczysko, V. Barone, *J. Phys. Chem. A* **2015**, *119*, 11862-11874.

[5] (a) A. S. Perera, J. Cheramy, M. R. Poopari, Y. Xu, *Phys. Chem. Chem. Phys.* **2019**, *21*, 3574-3584; (b) C. Merten, Y. Xu, *Angew. Chem. Int. Ed.* **2013**, *52*, 2073-2076; *Angew. Chem.* **2013**, *125*, 2127-2130.

[6] H. Watarai, K. Adachi, *Anal Bioanal Chem* **2009**, *395*, 1033-1046.

[7] S. Wada, K. Fujiwara, K. Monjushiro. H. Watarai, *Anal. Sci.* **2004**, *20*, 1489-1491.

Bibliography

[1] J. C. Brookes, A. P. Horsfield, A. M. Stoneham, *J. R. Soc. Interface* **2009**, *6*, 75-86.

[2] J. H. Kim, A. R. Scialli, *Toxicological Sciences* **2011**, *122*, 1-6.

[3] L. A. Nguyen, H. He, C. Pham-Huy, *Int J Biomed Sci.* **2006**, *2*, 85-100.

[4] F. O. Kelsey, *Teratology* **1998**, *38*, 221-226.

[5] R. Davis, H. M. Bryson, *Drugs* **1994**, *4*, 677-700.

[6] K. Kozlovskaya, E. Ovchinnikova, J. Kokubun, A. Rogalev, F. Wilhelm, F. Guillou, F. de Bergevin, A. F. Konstantinova, V. E. Dmitrienko, *Crystals* **2021**, *11*, 1389.

[7] Q. He, H. Gomaa, S. Rohani, J. Zhu, M. Jennings, *Chirality* **2010**, *22*, 707-716.

[8] G. R. Weisman, *In Asymmetric Synthesis*, J. D. Morrison Ed., Academic Press: New York **1983**, *1*, 153.

[9] J. M. Seco, E. Quiñoá, R. Riguera, *Chem. Rev*. **2004**, *104*, 17-117.

[10] G. Pescitelli, S. Gabriel, Y. Wang, J. Fleischhauer, R. W. Woody, N. Berova, *J. Am. Chem. Soc.* **2003**, *125*, 7613-7628.

[11] N. Berova, P. L. Polavarapu, K. Nakanishi, R. W. Woody, *Comprehensive chiroptical spectroscopy*, **2012**, Wiley: Chichester.

[12] L. D. Barron, *Molecular light scattering and optical activity*, **2009**, Cambridge University Press.

[13] L. A. Nafie, *Vibrational optical activity principles and applications*, **2011**, Wiley: Chichester.

[14] G. Yang, Y. Xu, *Top Curr Chem* **2011**, *298*, 189-236.

[15] F. Zhu, N. W. Isaacs, L. Hecht and L. D. Barron, *Structure* **2005,** *13*, 1409-1419.

[16] V. Parchaňský, J. Kapitán, P. Bouř, *RSC Adv.* **2014**, *4*, 57125-57136.

[17] B. H. Pate, L. Evangelisti, W. Caminati, Y. Xu, J. Thomas, D. Patterson, C. Perez, M. Schnell, *Quantitative chiral analysis by molecular rotational spectroscopy*, Chiral analysis, **2018**, Elsevier, Amsterdam, Netherlands, pp 679-729;

[18] F. Xie, N. A. Seifert, W. Jäger, Y. Xu, *Angew. Chem. Int. Ed.* **2020**, *59*, 15703-15710; *Angew. Chem.* **2020**, *132*, 15833-15840.

[19] D. Padula, G. Pescitelli, *Molecules* **2018**, *23*, 128.

[20] D. A. Long. *The Raman effect: a unified treatment of theory of Raman scattering by molecules*, **2002**, Wiley: Chichester.

[21] T. B. Freedman, X. Cao, R. K. Dukor, L. A. Nafie, *Chirality* **2003**, *15*, 743-758.

[22] P. J. Stephens, F. J. Devlin, J. J. Pan, *Chirality* **2008**, *20*, 643-663.

[23] P. L. Polavarapu, *Chirality* **2012**, *24*, 909-920.

[24] L. D. Barron, L. Hecht, I. H. McColl, E. W. Blanch, *Mol. Phys.* **2004**, *102*, 731-744.

[25] Gaussian 16, Revision B.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, **2016.**

[26] Amsterdam Density Functional Program; Vrije Universiteit: Amsterdam, The Netherlands; http://www.scm.com. E. J. Baerends, T. Ziegler, J. Autschbach, D. Bashford, A. Bérces, F. M. Bickelhaupt, C. Bo, P. M. Boerrigter, L. Cavallo, D. P. Chong, L. Deng, R. M. Dickson, D. E. Ellis, M. van Faassen, L. Fan, T. H. Fischer, C. Fonseca Guerra, M. Franchini, A. Ghysels, A. Giammona, S. J. A. van Gisbergen, A.W. Götz, J. A. Groeneveld, O. V. Gritsenko, M. Grüning, S. Gusarov, F. E. Harris, P. van den Hoek, C. R. Jacob, H. Jacobsen, L. Jensen, J. W. Kaminski, G. van Kessel, F. Kootstra, A. Kovalenko, M. V. Krykunov, E. van Lenthe, D. A. McCormack, A. Michalak, M. Mitoraj, S. M. Morton, J. Neugebauer, V. P. Nicu, L. Noodleman, V. P. Osinga, S. Patchkovskii, M. Pavanello, P. H. T. Philipsen, D. Post, C. C. Pye, W. Ravenek, J. I. Rodríguez, P. Ros, P. R. T. Schipper, G. Schreckenbach, J. S. Seldenthuis, M. Seth, J. G. Snijders, M. Solà, M. Swart, D. Swerhone, G. te Velde, P. Vernooijs, L. Versluis, L. Visscher, O. Visser, F. Wang, T. A. Wesolowski, E. M. van Wezenbeek, G. Wiesenekker, S. K. Wolff, T. K. Woo, A. L. Yakovlev.

[27] P. J. Stephens, F. J. Devlin, J. R. Cheeseman, M. J. Frisch, *J. Phys. Chem. A* **2001**, *105*, 5356-5371.

[28] G. Li, J. Kessler, J. Cheramy, T. Wu, M. R. Poopari, P. Bouř, Y. Xu, *Angew. Chem. Int. Ed.* **2019**, *58*, 16495-16498; *Angew. Chem.* **2019**, *131*, 16647-16650.

[29] T. Wu, G. Li, J. Kapitán, J. Kessler, Y. Xu, P. Bouř, *Angew. Chem. Int. Ed.* **2020**, *59*, 21895-21898; *Angew. Chem.* **2020**, *132*, 22079-22082.

[30] G. Li, M. Alshalalfeh, Y. Yang, J. R. Cheeseman, P. Bouř, Y. Xu, *Angew. Chem. Int. Ed.* **2020**, *60*, 22004-22009; *Angew. Chem.* **2021**, *133*, 22175-22180.

[31] G. Li, M. Alshalalfeh, J. Kapitán, P. Bouř, Y. Xu, *Chem. Eur. J.* **2022**, e202104302.

[32] D. Laage, T. Elsaesser, J. T. Hynes, *Chem. Rev.* **2017**, *117*, 1069-10725.

[33] M. Maurer, C. Oostenbrink, *J Mol Recognit.* **2019**, *32*, e2810.

[34] M. Ahmad, W. Gu, T. Geyer, V. Helms, *Nat Commun* **2011**, *2*, 261.

[35] P. Çarçabal, R. A. Jockusch, I. Hunig, L. C. Snoek, R. T. Kroemer, B. G. Davis, D. P. Gamblin,

I. Compagnon, J. Oomens, J. P. Simons, *J. Am. Chem. Soc.* **2005**, *127*, 11414-11425.

[36] V. Vaquero, M. E. Sanz, I. Peña, S. Mata, C. Cabezas, J. C. López, J. L. Alonso, *J. Phys. Chem. A* **2014**, *118*, 2584-2590.

[37] J. T. A. Gall, J. Thomas, F. Xie, Z. Wang, W. Jäger, Y. Xu, *Phys. Chem. Chem. Phys.* **2017**, *19*, 29508-29515.

[38] P. Ottaviani, B. Velino, W. Caminati, *Chem. Phys. Lett.* **2006**, *428*, 236-240.

[39] N. Borho, Y. Xu, *Phys. Chem. Chem. Phys.* **2007**, *9*, 1324-1328.

[40] J. Thomas, O. Sukhorukov, W. Jäger, Y. Xu, *Angew. Chem. Int. Ed.* **2014**, *53*, 1156-1159; *Angew. Chem.* **2014**, *126*, 1175-1178.

[41] F. Xie, S. Mahendiran, N. A. Seifert, Y. Xu, *Phys. Chem. Chem. Phys.* **2021**, *23*, 3820-3825.

[42] F. Xie, X. Ng, N. A. Seifert, J. Thomas, W. Jäger, Y. Xu, *J. Chem. Phys.* **2018**, *149*, 224306.

[43] C. Pérez, J. L. Neill, M. T. Muckle, D. P. Zaleski, I. Peña, J. C. Lopez, J. L. Alonso, B. H. Pate, *Angew. Chem. Int. Ed*. **2015**, *54*, 979-982; *Angew. Chem.* **2015**, *127*, 993-996.

[44] J. Thomas, Y. Xu, *J. Chem. Phys.* **2014**, *140*, 234307.

[45] W. Huang, J. Thomas, W. Jäger, Y. Xu, *Phys. Chem. Chem. Phys.* **2017**, *19*, 12221-12228.

[46] B. Wu, A. S. Hazrah, N. A. Seifert, S. Oswald, W. Jäger, Y. Xu, *J.Phys. Chem. A* **2021**, *125*, 10401-10409.

[47] A. Shahi, E. Arunan, *Phys. Chem. Chem. Phys.* **2015**, *17*, 24774-24782.

[48] B. Wu, N. A. Seifert, S. Oswald, W. Jäger, Y. Xu, *ChemPhysChem* **2022**, doi.org/10.1002/cphc.202200348.

[49] S. Gobi, E. Vass, G. Magyarfalvi, G. Tarczay, *Phys. Chem. Chem. Phys*. **2011**, *13*, 13972-13984.

[50] J. Sadlej, J. Cz. Dobrowolski, J. E. Rode, *Chem. Soc. Rev*. **2010**, *39*, 1478-1488.

[51] J. Haesler, I. Schindelholz, E. Riguet, C. G. Bochet, W. Hug, *Nature* **2007**, *446*, 526-529.

[52] K. J. Jalkanen, R. M. Nieminen, M. Knapp-Mohammady, S. Suhai, *Int. J. Quantum Chem.* **2003**, *92*, 239-259.

[53] L. J. Onsager, *Am. Chem. Soc.* **1936**, *58*, 1486.

[54] A. Klamt, G. J. Schuurmann, *Chem. Soc. Perkin Trans.* **1993**, *2*, 799.

[55] J. Tomasi, B. Mennucci, R. Cammi, *Chem. Rev.* **2005**, *105*, 2999-3094.

[56] E. Cances, B. Mennucci, J. Tomasi, *J. Chem. Phys.* **1997**, *107*, 3032.

[57] P. Zhu, G. Yang, M. R. Poopari, Z. Bie, Y. Xu, *ChemPhysChem* **2012**, *13*, 1272-1281.

[58] M. R. Poopari, Z. Dezhahang, G. Yang, Y. Xu, *ChemPhysChem* **2012**, *13*, 2310-2321.

[59] M. R. Poopari, Z. Dezhahang, Y. Xu, *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.* **2015**, *136*, 131-140.

[60] M. Losada, P. Nguyen, Y. Xu, *J. Phys. Chem. A* **2008**, *112*, 5621-5627.

[61] M. Losada, Y. Xu, *Phys. Chem. Chem. Phys.* **2007**, *9*, 3127-3135.

[62] M. Losada, H. Tran, Y. Xu, *J. Chem. Phys.* **2008**, *128*, 014508.

[63] G. Yang, Y. Xu, *J. Chem. Phys.* **2009**, *130*, 164506.

[64] A. S. Perera, J. Cheramy, C. Merten, J. Thomas, Y. Xu, *ChemPhysChem* **2018**, *19*, 2234-2242.

[65] A. S. Perera, J. Thomas, M. R. Poopari, Y. Xu,, *Front. Chem.* **2016**, *4*, 1-17.

[66] M. R. Poopari, Z. Dezhahang, Y. Xu, *Phys. Chem. Chem. Phys.* **2013**, *15*, 1655-1665.

[67] Y. Liu, G. Yang, M. Losada, Y. Xu, *J. Chem. Phys.* **2010**, *132*, 234513.

[68] K. Le Barbu-Debus, J. Bowles, S. Jähnigen, C. Clavaguéra, F. Calvo, R. Vuilleumier, A. Zehnacker, *Phys. Chem. Chem. Phys.* **2020**, *22*, 26047-26068.

[69] S. T. Mutter, F. Zielinski, J. R. Cheeseman, C. Johannessen, P. L. A.Popelier, E. W. Blanch, *Phys. Chem. Chem. Phys.* **2015**, *17*, 6016-6027.

[70] S. Ghidinelli, S. Abbate, J. Koshoubu, Y. Araki, T. Wada, G. Longhi, *J. Phys. Chem. B* **2020**, *124*, 4512-4526.

[71] M. Thomas, Theoretical modeling of vibrational spectra in the liquid phase. *Springer International Publishing*, **2017**.

[72] G. Pohl, A. Perczel, E. Vass, G. Magyarfalvi, G. Tarczay, *Phys. Chem. Chem. Phys.* **2007**, *9*, 4698-4708.

[73] C. Merten, Y. Xu, *ChemPhysChem* **2013**, *14*, 213-219.

[74] C. H. Pollok, C. Merten, *Phys. Chem. Chem. Phys.* **2016**, *18*, 13496-13502.

[75] A. S. Perera, J. Cheramy, M. R. Poopari, Y. Xu, *Phys. Chem. Chem. Phys.* **2019**, *21*, 3574-3584.

[76] C. Merten, Y. Xu, *Angew. Chem. Int. Ed.* **2013**, *52*, 2073-2076; *Angew. Chem.* **2013**, *125*, 2127-2130.

[77] A. G. Gaonkar, *JAOCS* **1989**, *66*, 1090-1092.

[78] G. T. Vladisavljević, M. Shimizu, T. Nakashima, *J. Membr. Sci.* **2006**, *284*, 373-383.

[79] H. Watarai, N. Sawadam T. Sawada (eds.), *Interfacial Nanochemistry*, Kluwer Academic /Plenum Publishers, New York, **2005**.

[80] A. G. Volkov, D. W. Deamer, D. L. Tanelian, V. S. Markin (eds.), *Liquid Interfaces in Chemistry and Biology*, Wiley, New York, **1998**.

[81] Y. Yang, J. Cheramy, Y. Xu, *ChemPhysChem* **2021**, *22*, 1336-1343.

[82] G. Holzwarth, E. C. Hsu, H. S. Mosher, T. R. Faulkner, A. Moscowitz, *J. Am. Chem. Soc.* **1974**, *96*, 251-252.

[83] L. A. Nafie, J. C. Cheng, P. J. Stephens, *J. Am. Chem. Soc.* **1975**, *97*, 3842-3843.

[84] L. A. Nafie, T. A. Keiderling, P. J. Stephens, *J. Am. Chem. Soc.* **1976**, *98*, 2715-2723.

[85] L. A. Nafie, M. Diem, D. W. Vidrine, *J. Am. Chem. Soc.* **1979**, *101*, 496-498.

[86] H. Sugeta, C. Marcott, T. R. Faulkner, J. Overend, A. Moscowitz, *Chem. Phys. Lett.* **1976**, *40*, 397.

[87] E. D. Lipp, C.G. Zimba, L. A. Nafie, *Chem. Phys. Lett.* **1982**, *90*, 1-5.

[88] P. Bouř, *ChemPhysChem* **2009**, *10*, 1983-1985.

[89] L. A. Nafie, *Nat Prod Commun* **2008**, *3*, 451-466.
[90] L. D. Barron, M. P. Bogaard, A. D. Buckingham, *J. Am. Chem. Soc.* **1973**, *95*, 603-605.
[91] W. Hug, S. Kint, G. F. Bailey, J. R. Scherer, *J. Am. Chem. Soc.* **1975**, *97*, 5589-5590.
[92] L. D. Barron, A. D. Buckingham, *Ann. Rev. Phys. Chem.* **1975**, *26*, 381-396.
[93] L. Hecht, L. D. Barron, W. Hug, *Chem. Phys. Lett.* **1989**, *158*, 341-344.
[94] L. D. Barron, L. Hecht, A. R. Gargaro, W. Hug, *J. Raman Spectrosc.* **1990**, *21*, 375-379.
[95] L. D. Barron, A. R. Gargaro, Z. Q. Wen, *J. Chem. Soc., Chem. Commun.* **1990**, 1034-1036.
[96] L. Hecht, L. D. Barron, A. R. Gargaro, Z. Q. Wen, W. Hug, *J. Raman Spectrosc.* **1992**, *23*, 401-411.
[97] W. Hug, *Appl. Spectrosc.* **2003**, *57*, 1-13.
[98] K. R. J. Lovelock, *R. Soc. Open sci.* **2017**, *4*, 171223.
[99] N. V. Cohan, H. F. Hameka, *J. Am. Chem. Soc.* **1966**, *88*, 2136-2142.
[100] T. R. Faulkner, C. Marcott, A. Moscowitz, J. Overend, *J. Am. Chem. Soc.* **1977**, *99*, 8160-8168.
[101] L. A. Nafie, T. B. Freedman, *J. Phys. Chem.* **1983**, *78*, 7108-7116.
[102] R. Dutler, A. Rauk, *J. Am. Chem. Soc.* **1989**, *111*, 6957-6966.
[103] D. Yang, A. Rauk *J. Chem. Phys.* **1992**, *97*, 6517-6534.
[104] L. A. Nafie, T. H. Walnut *Chem. Phys. Lett.* **1977**, *49*, 441-446.
[105] T. H. Walnut, L. A. Nafie, *J. Chem. Phys.* **1977**, *67*, 1501-1510.
[106] G. Holzwarth, I. Chabay, *J. Chem. Phys.* **1972**, *57*, 1632.
[107] A. D. Buckingham, P. W. Fowler, P. A. Galwas, *Chem. Phys.* **1987**, *112*, 1-14.
[108] P. J. Stephens, *J. Phys. Chem.* **1985**, *89*, 748-752.
[109] W. B. Person, J. H. Newton, *J. Chem. Phys.* **1974**, *61*, 1040.
[110] F. J. Delvin, P. J. Stephens, J. R. Cheesman, M. J. Frisch, *J. Am. Chem. Soc.* **1996**, *118*, 6327- 6328.
[111] P. Bouř, J. McCann, H. Wieser, *J. Chem. Phys.* **1998**, *108*, 8782-8789.
[112] L. D. Barron, *Molecular Light Scattering and Optical Activity*, **1982**, Cambridge University Press, Cambridge.
[113] P. Pracht, F. Bohle, S. Grimme, *Phys. Chem. Chem. Phys.* **2020**, *22*, 7169-7192.
[114] S. Grimme, C. Bannwarth, P. Shushkov, *J. Chem. Theory Comput.* **2017**, *13*, 1989-2009.
[115] S. Oswald, N. A. Seifert, F. Bohle, M. Gawrilow, S. Grimme, W. Jäger, Y. Xu, M. A. Suhm, *Angew. Chem. Int. Ed.* **2019**, *58*, 5080-5084; *Angew. Chem.* **2019**, *131*, 5134-5138.
[116] H. Wang, M. Heger, M. H. Al-Jabiri, Y. Xu, *Molecules* **2022**, *27*, 38.
[117] F. Xie, M. Fusè, A. S. Hazrah, W. Jäger, V. Barone, Y. Xu, *Angew. Chem. Int. Ed.* **2020**, *59*, 22427-22430; *Angew. Chem.* **2020**, *132*, 22613-22616.
[118] H. J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, *Wires Comput. Mol. Sci.* **2012**, *2*, 242-253.
[119] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297-3305.

[120] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J. Chem. Phys.* **2010**, *132*, 154104.

[121] D. G. A. Smith, L. A. Burns, K. Patkowski, C. D. Sherrill, *J. Phys. Chem. Lett.* **2016**, *7*, 2197-2203.

[122] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098-3100.

[123] C. Lee, W. Yang, R. Parr, *Phys. Rev. B* **1988**, *37*, 785-789.

[124] A. Aerts, A. Brown, F. Gatti, *J. Chem. Phys.* **2022**, *157*, 014306.

[125] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, *J. Phys. Chem.* **1994**, *98*, 11623-11627.

[126] T. Kuppens, W. Langenaeker, J. P. Tollenaere, P. Bultinck, *J. Phys. Chem. A* **2003**, *107*, 542-553.

[127] T. Kuppens, K. Vandyck, J. van der Eycken, W. Herrebout, B. J. van der Veken, P. Bultinck, *J. Org. Chem.* **2005**, *70*, 9103-9114

[128] T. Kuppens, W. Herrebout, B. J. van der Veken, P. Bultinck, *J. Phys. Chem. A* **2006**, *110*, 10191-10200.

[129] L. Ducasse, F. Castet, A. Fritsch, I. Huc, T. Buffeteau, *J. Phys. Chem. A* **2007**, *111*, 5092-5098.

[130] T. Brotin, D. Cavagnat, J. P. Dutasta, T. Buffeteau, *J. Am. Chem. Soc.* **2006**, *128*, 5533-5540.

[131] T. Kuppens, K. Vandyck, J. van der Eycken, W. Herrebout, B. van der Veken, P. Bultinck, *Spectrochimica Acta Part A* **2007**, *67*, 402-411.

[132] A. D. Becke, E. R. Johnson, *J. Chem. Phys.* **2005**, *123*, 154101.

[133] L. A. Nafie, Vibrational Optical Activity: Principles and Applications, John Wiley & Sons, Ltd., UK, **2011**.

[134] C. Merten, *Phys. Chem. Chem. Phys.* **2017**, *19*, 18803-18812;

[135] V. P. Nicu, E. J. Baerends, P. L. Polavarapu, *J. Phys. Chem. A* **2012**, *116*, 8366-8373.

[136] G. Yang, Y. Xu, *Phys. Chem. Chem. Phys.* **2008**, *10*, 6787-6795.

[137] G. Tarczay, G. Magyarfalvi, E. Vass, *Angew. Chem. Int. Ed.* **2006**, *45*, 1775-1777; *Angew. Chem.* **2006**, *118*, 1807-1809.

[138] Y. Yang, J. Cheramy, M. Brehm, Y. Xu, *ChemPhysChem* **2022**, *23*, e202200161.

[139] C. H. Pollok, T. Riesebeck, C. Merten, *Angew. Chem. Int. Ed.* **2017**, *56*, 1925-1928; *Angew. Chem.* **2017**, *129*, 1952-1955;

[140] Y. Fujima, Y. Hirayama, M. Ikunaka, Y. Nishimoto, *Tetrahedron-Asymm.* **2003**, *14*, 1385-1391.

[141] W. Chou, M. Chou, Y. Lu, S. Chen, *US Patent*, **2001**, 6316293(B1).

[142] F. Codignola, M. Piacenza, *Italian Patent*, **1947**, ES179144(A1).

[143] W. Li, L. Evangelisti, Q. Gou, W. Caminati, R. Meyer, *Angew. Chem. Int. Ed.* **2019**, *58*, 859-865; *Angew. Chem.* **2019**, *131*, 869-875.

[144] J. Thomas, A. Serrato III, W. Lin, W. Jäger, Y. Xu, *Eur. J. Chem.* **2014**, *20*, 6148-6153.

[145] M. Shome, N. Mishra, *Indian J. Adv. Chem. Sci.* **2016**, *4*, 56-67.

[146] T. Kuppens, W. Herrebout, B. Veken, P. Bultinck, *J. Phys. Chem. A* **2006**, *110*, 10191-10200.

[147] M. Halupka, W. Sander, *Spectrochim. Acta Part A* **1998**, *54*, 495-500.

[148] E. M. S. Maçôas, P. Myllyperkiö, H. Kunttu, M. Pettersson, *J. Phys. Chem. A* **2009,** *113,* 7227-7234.

[149] J. Ahokas, I. Kosendiak, J. Krupa, J. Lundell, M. Wierzejewska, *J. Mol. Struct.* **2018**, *1163,* 294-299.

[150] M. Winnewisser, B. P. Winnewisser, M. Stein, M. Birk, G. Wagner, G. Winnewisser, K. M. T. Yamada, S. P. Belov, O. L. Baskakov, *J. Mol. Spectrosc.* **2002**, *216*, 259-265.

[151] M. Pettersson, E. M. S. Maçôas, L. Khriachtchev, R. Fausto, M. Räsänen, *J. Am. Chem. Soc.* **2003**, *125*, 4058-4059.

[152] G. Giubertoni, O. O. Sofronov, H. J. Bakker, *J. Chem. Phys. Lett.* **2019**, *10*, 3217-3222.

[153] K. A. E. Meyer, M. A. Suhm, *J. Chem. Phys.* **2018**, *149*, 104307.

[154] S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 2847-2863.

[155] S. Grimme, C. Bannwarth, S. Dohm, A. Hansen, J. Pisarek, P. Pracht, J. Seibert, F. Neese, *Angew. Chem. Int. Ed.* **2017**, *56*, 14763-14769; *Angew. Chem.* **2017**, *129*, 14958-14964.

[156] F. Xie, N. A. Seifert, M. Heger, J. Thomas, W. Jäger, Y. Xu, *Phys. Chem. Chem. Phys.* **2019**, *21*, 15408-15416

[157] F. Madeja, M. Havenith, K. Nauta, R. E. Miller, J. Chocholoušová, P. Hobza, *J. Chem. Phys.* **2004**, *22*, 10544-10560.

[158] M. Albrecht, V. A. Soloshonok, L. Schrader, M. Yasumoto, M. A. Suhm, *J. Fluo. Chem*. **2010**, *131*, 495-504.

[159] A. Schriver, L. Schriver-Mazzuoli, A. A. Vigasin, *Vib. Spectrosc.* **2000**, *23*, 83-94.

[160] S. Abbate, L. F. Burgi, F. Gangemi, R. Gangemi, F. Lebon, G. Longhi, V. M. Pultz, D. A. Lightner, *J. Phys. Chem. A* **2009***, 113,* 11390-11405.

[161] A. D. Becke, *J. Chem. Phys*. **1993**, *98*, 5648.

[162] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.* **2011**, *32*, 1456-1465.

[163] D. Rappoport, F. Furche, *J. Chem. Phys.* **2010**, *133*, 134105.

[164] Z. Li, Y. Ding, W. Jiang, *Fenzi Cuihua*, **2005**, *19*, 131-135.

[165] U. R. Bapat, J. P. Potams, N. Subramanian, J. Valgeirsson, *PCT Int. Appl.* **2008**, 2008152514.

[166] H. Han, J. Jin, J. Liu, *Shenyang Yaoke Daxue Xuebao*. **2001**, *18*, 20-22.

[167] W. C. Shyu, C. A. Knupp, K. A. Pittman, L. Dunkle, R. H. Barbhaiya, *Clin. Pharmacol. Ther.* **1991**, *50*, 503-507.

[168] G. A. Jeffrey, W. Saenger, Hydrogen Bonding in Biological Structure, Springer, Berlin, **1991**.

[169] P. Ball, H2O: A Biography of Water, Weidenfeld & Nicolson, London, **1999**.

[170] M. Yves, The Hydrogen Bond and the Water Molecule, Elsevier, New York, **2007**.

[171] P. L. Polavarapu, E. Santoro, *Nat. Prod. Rep.* **2020**, *37*, 1661-1699.

[172] D. Kurouski, *Anal. Chim. Acta* **2017**, *990*, 54-66.

[173] C. Merten, C. J. Berger, R. McDonald, Y. Xu, *Angew. Chem. Int. Ed.* **2014**, *53*, 9940-9943; *Angew. Chem.* **2014**, *126*, 10098-10101.

[174] L. Weirich, K. Blanke, C. Merten, *Phys. Chem. Chem. Phys.* **2020**, *22*, 12515-12523.

[175] J. Thomas, M. J. Carrillo, A. Serato III, E. G. Schnitzler, W. Jäger, Y. Xu, W. Lin, *Chem. Phys. Lett.* **2016**, *661*, 210-214.

[176] F. Xie, N. A. Seifert, A. S. Hazrah, W. Jäger, Y. Xu, *ChemPhysChem* **2021**, *22*, 455-460.

[177] S. Abbate, L. Laux, J. Overend, A. Moscowitz, *J. Chem. Phys.* **1981**, *75*, 3161.

[178] M. R. Poopari, P. Zhu, Z. Dezhahang, Y. Xu, *J. Chem. Phys.* **2012**, *137*, 194308.

[179] C. D. Carlson, A. S. Hazrah, D. Mason, Q. Yang, N. A. Seifert, Y. Xu, *J. Phys. Chem. A* **2022**, jp-2022-058039 (submitted).

[180] C. Bannwarth, S. Ehlert, S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 1652-1671.

[181] Q. Yang, C. D. Carlson, W. Jäger, Y. Xu, *J. Phys. Chem.* A **2022**, *126*, 2942-2949.

[182] A. Insausti, J. Ma, Q. Yang, F. Xie, Y. Xu, *ChemPhysChem.* **2022**, doi.org/10.1002/ cphc.202200176.

[183] N. A. Seifert, J. Thomas, W. Jäger, Y. Xu. *Phys. Chem. Chem. Phys.* **2018**, *20*, 27630-27637.

[184] A. Schäfer, C. Huber, R. Ahlrichs, *J. Chem. Phys.* **1994**, *100*, 5829.

[185] M. A. J. Koenis, Y. Xia, S. R. Domingos, L. Visscher, W. J. Buma, V. P. Nicu, *Chem. Sci.* **2019**, *10*, 7680-7689.

[186] K Bünnemann, C. H. Pollok, C. Merten, *J. Phys. Chem. B* **2018**, *122*, 8056-8064.

[187] Z. Su, Q. Wen, Y. Xu, *J. Am. Chem. Soc.* **2006**, *128*, 6755-6760.

[188] R. B. Mackenzie, C. T. Dewberry, R. D. Cornelius, C. J. Smith, K. R. Leopold, *J. Phys. Chem. A* **2017**, *121*, 855-860.

[189] E. Brini, C. J. Fennell, M. Fernandez-Serra, B. H. Lee, M. Lukšič, K. A. Dill, *Chem. Rev.* **2017**, *117*, 12385-12414.

[190] M. Gasior-Glogowska, K. Malek, G. Zajac, M. Baranska, *Analyst* **2016**, *141*, 291-296.

[191] D. P. Demarque, M. Kemper, C. Merten, *Chem. Commun.* **2021**, *57*, 4031-4034.

[192] C. Merten, F. Li, K. Bravo-Rodriguez, E. Sanchez-Garcia, Y. Xu, W. Sander, *Phys. Chem. Chem. Phys.* **2014**, *16*, 5627-5633.

[193] K. H. Hopmann, K. Ruud, M. Pecul, A. Kudelski, M. Dračínský, P. Bouř, *J. Phys. Chem. B* **2011**, *115*, 4128-4137.

[194] S. Cardamone, B. A. Caine, E. Blanch, M. G. Lizio, P. L. A. Popelier, *Phys. Chem. Chem. Phys.* **2016**, *18*, 27377-27389.

[195] T. Giovannini, G. Del-Frate, P. Lafiosca, C. Cappelli, *Phys. Chem. Chem. Phys.* **2018**, *20*, 9181-9197.

[196] V. Palivec, C. Johannessen, J. Kaminský, H. Martinez-Seara, *PLoS Comput. Biol.* **2022**, *18*, e1009678.

[197] C. Merten, J. Bloino, V. Barone, Y. Xu, *J. Phys. Chem. Lett.* **2013**, *4*, 3424-3428.

[198] A. Stark, M. Brehm, M. Brüssel, S. B. C. Lehmann, A. S. Pensado, M. Schöppke, B. Kirchner, *Top.*

Curr. Chem. **2014**, *351*, 149-187.

[199] M. Thomas, M. Brehm, R. Fligg, P. Vöhringer, B. Kirchner, *Phys. Chem. Chem. Phys.* **2013**, *15*, 6608-6622.

[200] P. L. Silvestrelli, M. Bernasconi, M. Parrinello, *Chem. Phys. Lett.* **1997**, *277*, 478-482.

[201] A. Putrino, M. Parrinello, *Phys. Rev. Lett.* **2002**, *88*, 176401.

[202] C. Bannwarth, E. Caldeweyher, S. Ehlert, A. Hansen, P. Pracht, J. Seibert, S. Spicher, S. Grimme, *WIREs Comput. Mol. Sci.* **2020**, *11*, e1493.

[203] G. H. Wannier, *Phys. Rev.* **1937**, *52*, 191-197.

[204] N. Marzari, D. Vanderbilt, *Phys. Rev. B* **1997**, *56*, 12847-12865.

[205] M. Thomas, M. Brehm, B. Kirchner, *Phys. Chem. Chem. Phys.* **2015**, *17*, 3207-3213.

[206] M. Brehm, M. Thomas, *Molecules* **2021**, *26*, 1875.

[207] M. Thomas, B. Kirchner, *J. Phys. Chem. Lett.* **2016**, *7*, 509-513.

[208] A. Scherrer, R. Vuilleumier, D. Sebastiani, *J. Chem. Phys.* **2016**, *145*, 084101.

[209] S. Abbate, G. Longhi, K. Kwon, A. Moscowitz, *J. Chem. Phys.* **1998**, *108*, 50-62.

[210] S. Luber, *J. Chem. Theor. Comput.* **2017**, *13*, 1254-1262.

[211] M. Brehm, M. Thomas, *J. Phys. Chem. Lett.* **2017**, *8*, 3409-3414.

[212] M. Brehm, B. Kirchner, *J. Chem. Inf. Model.* **2011**, *51*, 2007-2023.

[213] M. Brehm, M. Thomas, S. Gehrke, B. Kirchner, *J. Chem. Phys.* **2020**, *152*, 164105.

[214] C. Merten, L. D. Barron, L. Hecht, C. Johannessen, *Angew. Chem. Int. Ed.* **2011**, *50*, 9973-9976; *Angew. Chem.* **2011**, *123*, 10149-10152.

[215] L. Weirich, J. M. de Oliveira, C. Merten, *Phys. Chem. Chem. Phys.* **2020**, *22*, 1525-1533.

[216] G. N. Simm, P. L. Türtscher, M. Reiher, *J. Comput. Chem.* **2020**, *41*, 1144-1155.

[217] B. Mennucci, J. Tomasi, R. Cammi, J. R. Cheeseman, M. J. Frisch, F. J. Devlin, S. Gabriel, P. J. Stephens, *J. Phys. Chem. A* **2002**, *106*, 6102-6113.

[218] R. A. Kendall, T. H. Dunning Jr., R. J. Harrison, *J. Chem. Phys.* **1992**, *96*, 6796-6806.

[219] S. Plimpton, *J. Comp. Phys.* **1995**, *117*, 1-19.

[220] C. I. Bayly, P. Cieplak, W. Cornell, P. A. Kollman, *J. Phys. Chem.* **1993**, *97*, 10269-10280.

[221] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc.* **1996**, *118*, 11225-11236.

[222] H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura, T. Head-Gordon, *J. Chem. Phys.* **2004**, *120*, 9665-9678.

[223] J. Hutter, M. Iannuzzi, F. Schiffmann, J. VandeVondele, *WIREs Comput. Mol. Sci.* **2014**, *4*, 15-25.

[224] T. D. Kühne, M. Iannuzzi, M. Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann, D. Golze, J. Wilhelm, S. Chulkov, M. H. Bani-Hashemian, V. Weber, U. Borštnik, M. Taillefumier, A. S. Jakobovits, A. Laz-zaro, H. Pabst, T. Müller, R. Schade, M.

Guidon, S. Andermatt, N. Holmberg, G. K. Schenter, A. Hehn, A. Bussy, F. Belleflamme, G. Tabacchi, A. Glöß, M. Lass, I. Bethune, C. J. Mundy, C. Plessl, M. Watkins, J. VandeVondele, M. Krack, J. Hutter, *J. Chem. Phys.* **2020**, *152*, 194103.

[225] J. Hutter, M. Iannuzzi, F. Schiffmann, M. Parrinello, T. Chassaing, J. Hutter, *Comput. Phys. Commun.* **2005**, *167*, 103-128.

[226] J. VandeVondele, J. Hutter, *J. Chem. Phys.* **2003**, *118*, 4365-4369.

[227] J. VandeVondele, J. Hutter, *J. Chem. Phys.* **2007**, *127*, 114105.

[228] S. Goedecker, M. Teter, J. Hutter, *Phys. Rev. B* **1996**, *54*, 1703-1710.

[229] S. Nose, *Mol. Phys.* **1984**, *52*, 255-268.

[230] W. G. Hoover, *Phys. Rev. A* **1985**, *31*, 1695-1697.

[231] G. Martyna, M. Klein, M. Tuckerman, *J. Chem. Phys.* **1992**, *97*, 2635-2643.

[232] M. Brehm, M. Thomas, *J. Chem. Inf. Model.* **2018**, *58*, 2092-2107.

[233] A. Zehnacker, M. A. Suhm, *Angew. Chem. Int. Ed.* **2008**, *47*, 6970-6992; *Angew. Chem.* **2008**, *120*, 7076-7100.

[234] K. Hirata, Y. Mori, S. Ishiuchi, M. Fujii, A. Zehnacker, *Phys. Chem. Chem. Phys.* **2020**, *22*, 24887-24894.

[235] F. X. Sunahori, G. Yang, E. N. Kitova, J. S. Klassen, Y. Xu, *Phys. Chem. Chem. Phys.* **2013**, *15*, 1873-1886.

[236] T. R. Hoye, C. S. Jeffrey, F. Shao, *Nat. Protoc.* **2007**, *2*, 2451-2458.

[237] T. J. Ward, K. D. Ward, *Anal. Chem.* **2012**, *84*, 626-635.

[238] K. Ariga, *Phys. Chem. Chem. Phys.* **2020**, *22*, 24856-24869.

[239] H. Sato, *Phys. Chem. Chem. Phys.* **2020**, *22*, 7671-7679.

[240] S. Arnaboldi, T. Benincori, A. Penoni, L. Vaghi, R. Cirilli, S. Abbate, G. Longhi, G. Mazzeo, S. Grecchi, M. Panigati, P. R. Mussini, *Chem. Sci.* **2019**, *10*, 2708-2717.

[241] R. McKendry, M. E. Theoclitou, T. Rayment, C. Abell, *Nature* **2019**, *391*, 566-568.

[242] H. Guo, J. Kim, S. Kim, S. Chang, W. Kim, *Langmuir* **2009**, *25*, 648-652.

[243] P. Lopes, R. Kataky, *Anal. Chem.* **2012**, *84*, 2299-2304.

[244] A. Shundo, K. Hori, T. Ikeda, N. Kimizuka, K. Tanaka, *J. Am. Chem. Soc.* **2013**, *135*, 10282-10285.

[245] K. Adachi, K. Chayama, H. Watarai, *Chirality* **2006**, *18*, 599-608.

[246] Y. Iimura, M. Fukuyama, A. Hibara, M. Harada, T. Okada, *J. Colloid Interface Sci.* **2017**, *508*, 469-475.

[247] Y. Umemura, A. Yamagishi, R. Schoonheydt, A. Persoons, F. D. Schryver, *J. Am. Chem. Soc.* **2002**, *124*, 992-997.

[248] R. H. A. Ras, Y. Umemura, C. T. Johnston, A. Yamagishi, R. A. Schoonheydt, *Phys. Chem. Chem. Phys.* **2007**, *9*, 918-932.

[249] G. T. Vladisavlević, M. Shimizu, T. Nakashima, *J. Membr. Sci.* **2006**, *284*, 373-383.

[250] P. Tavakoli, S. Z. Shadizadeh, F. Hayati, M. Fattahi, *Petroleum* **2020**, *6*, 293-303.

[251] C. Shi, B. Yan, L. Xie, L. Zhang, J Wang, A. Takahara, H. Zeng, *Angew. Chem. Int. Ed.* **2016**, *55*, 15017-15021; *Angew. Chem.* **2016**, *128*, 15241-15245.

[252] B. Yan, X. Wang, X. Zhang, S. Liu, M. Li, R. Ran, *J. Environ. Chem. Eng.* **2021**, *9*, 106607.

[253] D. Wang, D. Yang, C. Huang, Y. Huang, D. Yang, H. Zhang, Q. Liu, T. Tang, M. G. El-Din, T. Kemppi, B. Perdicakis, H. Zeng, *Fuel* **2021**, *286*, 119390.

[254] S. Reynaert, P. Moldnaers, J. Vermant, *Langmuir* **2006**, *22*, 4936-4945.

[255] H. Watarai, N. Sawada, T. Sawada (eds.), *Interfacial Nanochemistry*, Kluwer Academic/Plenum Publishers, New York, **2005**.

[256] A. G. Volkov, D. W. Deamer, D. L. Tanelian, V. S. Markin (eds.), *Liquid Interfaces in Chemistry and Biology*, Wiley, New York, **1998**.

[257] M. Kaminski, A. Kudelski, M. Pecul, *J. Phys. Chem. B* **2012**, *116*, 4976-4990.

[258] T. Lan, P. Wu, Z. Liu, M. Stroet, J. Liao, Z. Chai, A. E. Mark, N. Liu, D. Wang, *Environ. Sci. Technol.* **2022**, *56*, 917-927.

[259] X. Sun, D. Yang, H. Zhang, H. Zeng, T. Tang, *J. Phys. Chem. B* **2021**, *125*, 7556-7567.

[260] C. Jian, M. R. Poopari, Q. Liu, N. Zerpa, H. Zeng, T. Tang, *J. Phys. Chem. B* **2016**, *120*, 5646-5654.

[261] T. Arakawa, K. Tsumoto, Y. Kita, B. Chang, D. Ejima, *Amino Acids* **2007**, *33*, 587-605.

[262] D. A. Evans, R. J. Thomson, *J. Am. Chem. Soc.* **2005**, *127*, 10506-10507.

[263] D. Ferraris, B. Young, C. Cox, W. J. Drury III, T. Dudding, T. Lectka, *J. Org. Chem.* **1998**, *63*, 6090-6091.

[264] S. K. Burley, G. A. Petsko, *Science* **1985**, *229*, 23-28.

[265] Z. Dezhahang, C. Merten, M. R. Poopari, Y. Xu, *Dalton Trans.* **2012**, *41*, 10817-10824.

[266] A. Benderrag, M. Daaou, B. Bounaceur, B. Haddou, *Chemical Papers* **2016**, *70*, 1196-1203.

[267] J. Drelich, C. Fang, C. L. White, *Encycl. Surf. Colloid Sci.* **2002**, *3*, 3152-3166.

[268] S. P. Pinho, *J. Chem. Eng. Data* **2008**, *53*, 180-184.

[269] V. Scutelnic, M. A. S. Perez, M. Marianski, S. Warnke, A. Gregor, U. Rothlisberger, M. T. Bowers, C. Baldauf, G. von Helden, T. R. Rizzo, J. Seo, *J. Am. Chem. Soc.* **2018**, *140*, 7554-7560.

[270] C. Jian, M. R. Poopari, Q. Liu, N. Zerpa, H. Zeng, T. Tang, *Energy Fuels* **2016**, *30*, 10228-10235.

[271] E. R. A. Lima, B. M. de Melo, L. T. Baptista, M. L. L. Paredes, *Braz. J. Chem. Eng.* **2013**, *30*, 55-62.

[272] R. I. Slavchov, J. K. Novev, *J. Colloid Interface Sci.* **2012**, *387*, 234-243.

[273] C. J. Beverung, C. J. Radke, H. W. Blanch, *Biophys. Chem.* **1999**, *81*, 59-80.

[274] Z. M. Aman, E. P. Brown, E. D. Sloan, A. K. Sum, C. A. Koh, *Phys. Chem. Chem. Phys.* **2011**, *13*, 19796-19806.

[275] M. Abraham, D. van der Spoel, E. Lindahl, B. Hess, *GROMACS User Manual*, Version 2016, **2018**.

[276] T. Lan, H. Zeng, T. Tang, *J. Phys. Chem. C* **2019**, *123*, 22989-22999.

[277] X. Huang, K. Quan, D. Pei, J. Liu, D. Di, *Chirality* **2018**, *30*, 974-981.

[278] Y. Sun, D. Zhang, Y. Sheng, D. Xu, R. Zhang, M. Bradley, *Anal. Methods* **2021**, *13*, 2011-2020.

[279] W. Xu, M. Cheng, S. Zhang, Q. Wu, Z. Liu, M. K. Dhinakaran, F. Liang, E. G. Kovaleva, H. Li, *Chem. Commun.* **2021**, *57*, 7480-7492.

[280] E. Y. Arashiro, N. R. Demarquette, *Mater. Res.* **1999**, *2*, 23-32.

[281] C. E. Stauffer, *J. Phys. Chem.* **1965**, *69*, 1933-1938.

[282] J. D. Berry, M. J. Neeson, R. R. Dagastine, D. Y. C. Chan, R. F. Tabor, *J. Colloid Interface Sci.* **2015**, *454*, 226-237.

[283] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, G. R. Hutchison, *J. Cheminform* **2012**, *4*, 17.

[284] A. K. Malde, L. Zuo, M. Breeze, M. Stroet, D. Poger, P. C. Nair, C. Oostenbrink, A. E. Mark, *J. Chem. Theory Comput.* **2011**, *7*, 4026-4037.

[285] C. Oostenbrink, A. Villa, A. E. Mark, W. F. van Gunsteren, *J. Comput. Chem.* **2004**, *25*, 1656-1676.

[286] C. M. Breneman, K. B. Wiberg, *J. Comput. Chem.* **1990**, *11*, 361-373.

[287] X. Sun, H. Zeng, T. Tang, *J. Mol. Liq.* **2020**, *310*, 113248.

[288] T. Lan, H. Zeng, T. Tang, *J. Phys. Chem. C* **2019**, *123*, 22989-22999.

[289] X. Sun, H. Zeng, T. Tang, *J. Colloid Interface Sci.* **2021**, *586*, 766-777.

[290] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. C. Berendsen, *J. Comput. Chem.* **2005**, *26*, 1701-1718.

[291] E. Lindahl, B. Hess, D. van der Spoel, *J. Mol. Model* **2001**, *7*, 306-317.

[292] H. J. C. Berendsen, D. van der Spoel, R. van Drunen, *Comput. Phys. Commun* **1995**, *91*, 43-56.

[293] H. Saito, H. Nagao, K. Nishikawa, K. Kinugawa, *J. Chem. Phys.* **2003**, *119*, 953-963.

[294] B. Hess, *J. Chem. Theory Comput.* **2008**, *4*, 116-122.

[295] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L.G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577-8593.

[296] A. R. van Buuren, S. J. Marrink, H. J. C. Berendsen, *J. Phys. Chem.* **1993**, *97*, 9206-9212.

[297] P. T. Panek, C. R. Jacob, *Phys. Chem. Lett.* **2016**, *7*, 3084-3090.

[298] K. Ziadi, *J. Raman Spectrosc.* **2021**, *53*, 222-236.

[299] J. Bloino, M. Biczysko, V. Barone, *J. Phys. Chem. A* **2015**, *119*, 11862-11874.

[300] H. Watarai, K. Adachi, *Anal Bioanal Chem* **2009**, *395*, 1033-1046.

[301] S. Wada, K. Fujiwara, K. Monjushiro. H. Watarai, *Anal. Sci.* **2004**, *20*, 1489-1491.

Appendix A

Supporting Information for Chapter 3

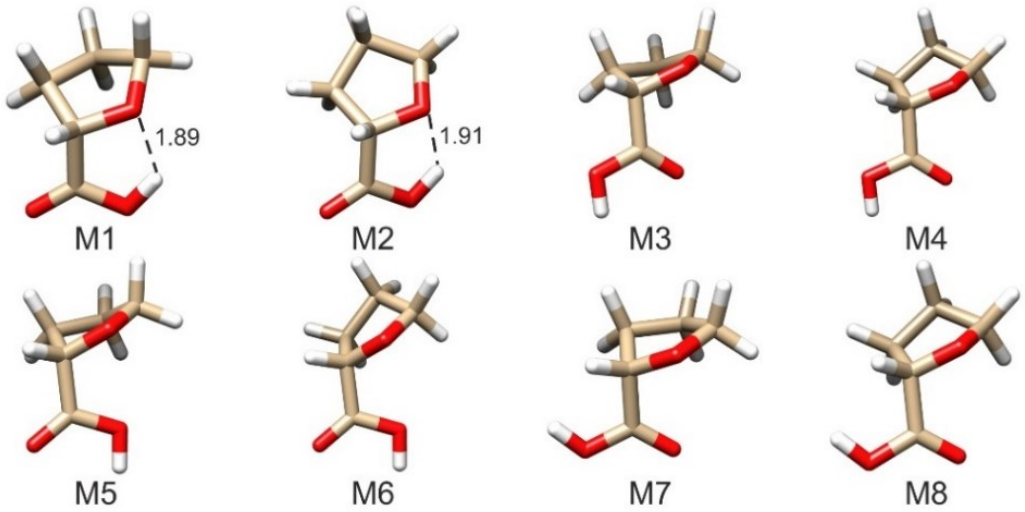


Figure A1. Optimized geometries of the eight THFA conformers at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

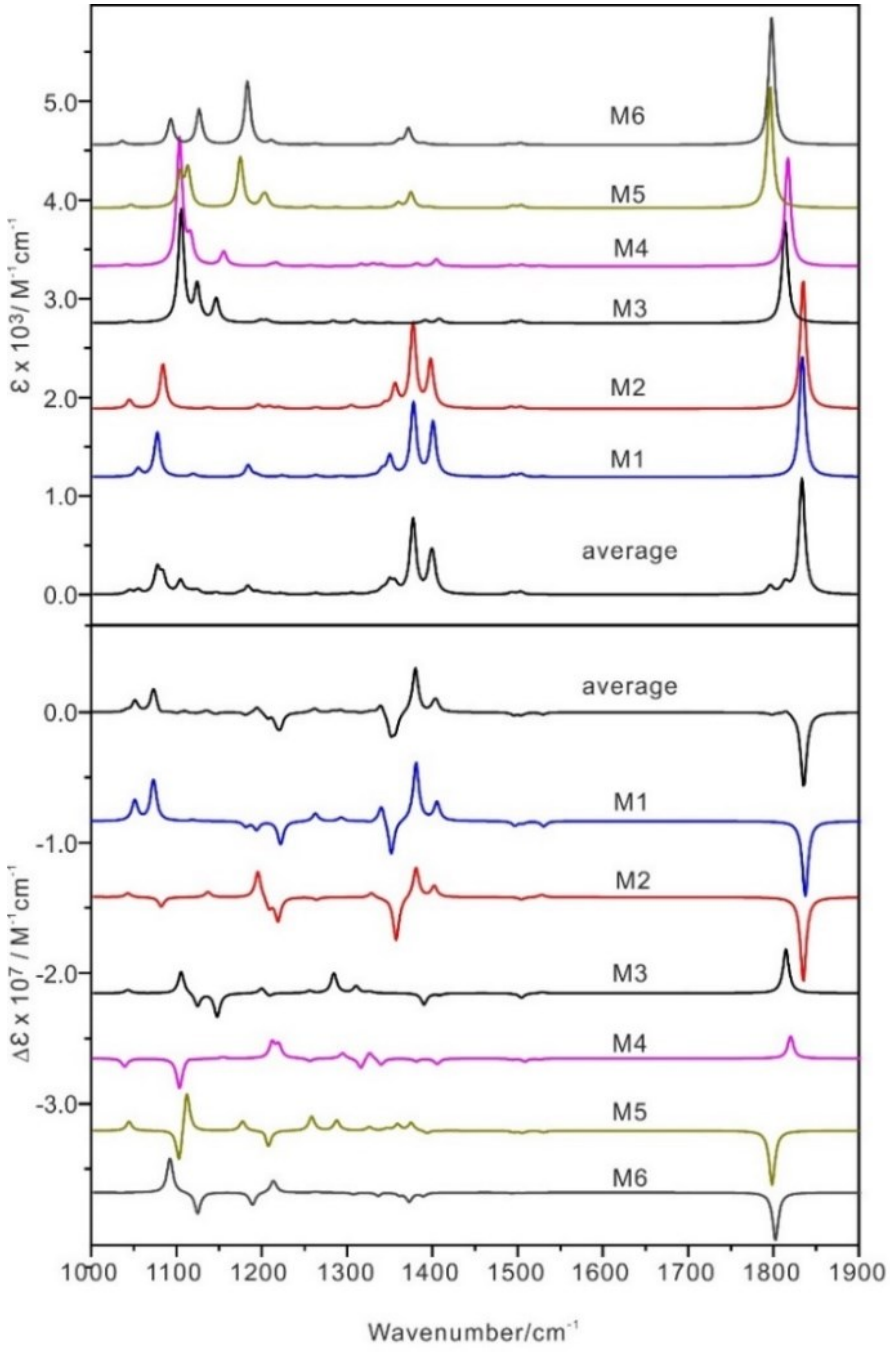


Figure A2. Simulated IR and VCD spectra of the six most stable THFA monomeric conformers, M1-M6 and the related Boltzmann averaged spectra at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory at 298 K.

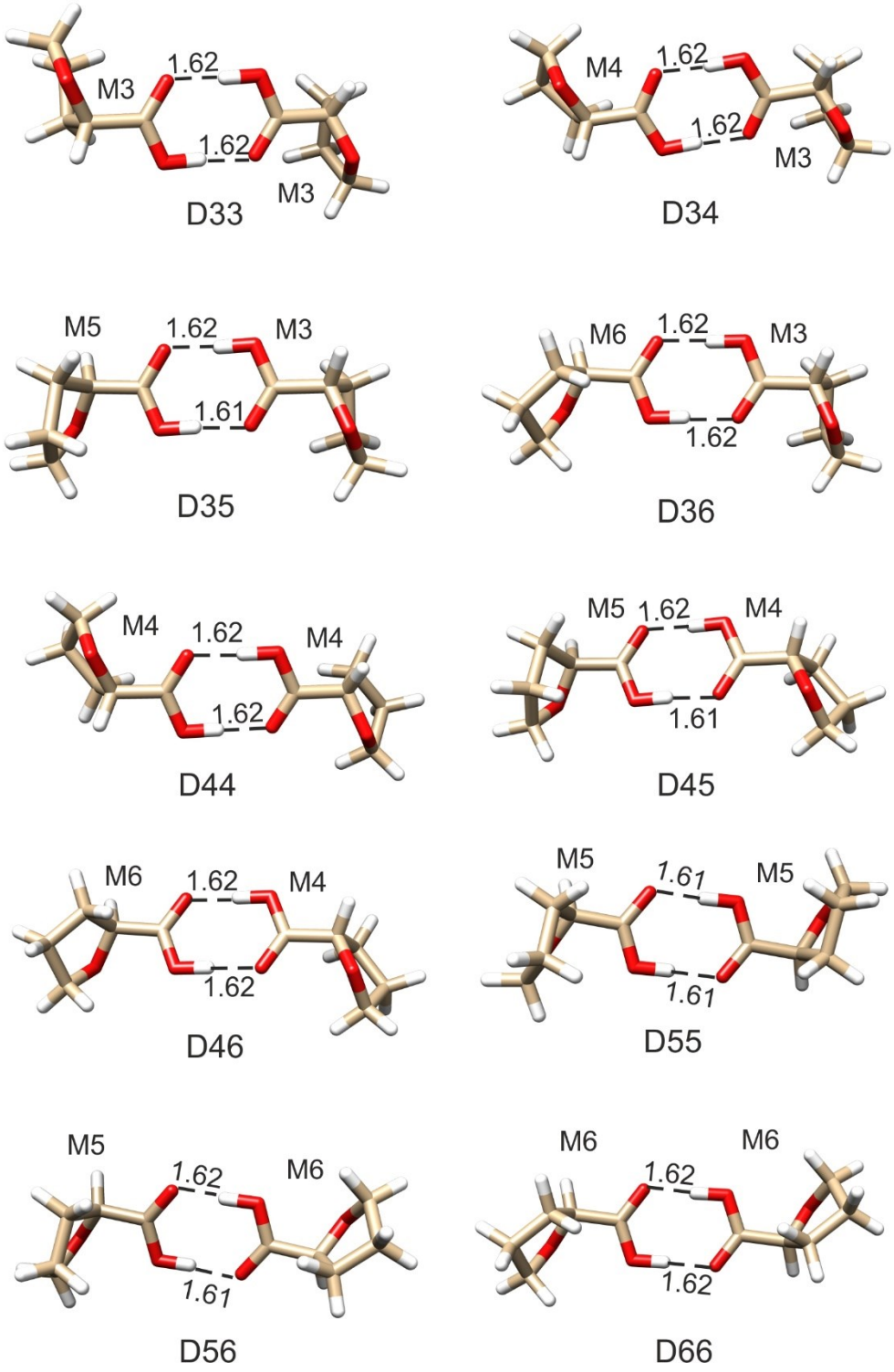


Figure A3. Optimized geometries of the ten **Type 1** binary THFA conformers at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory

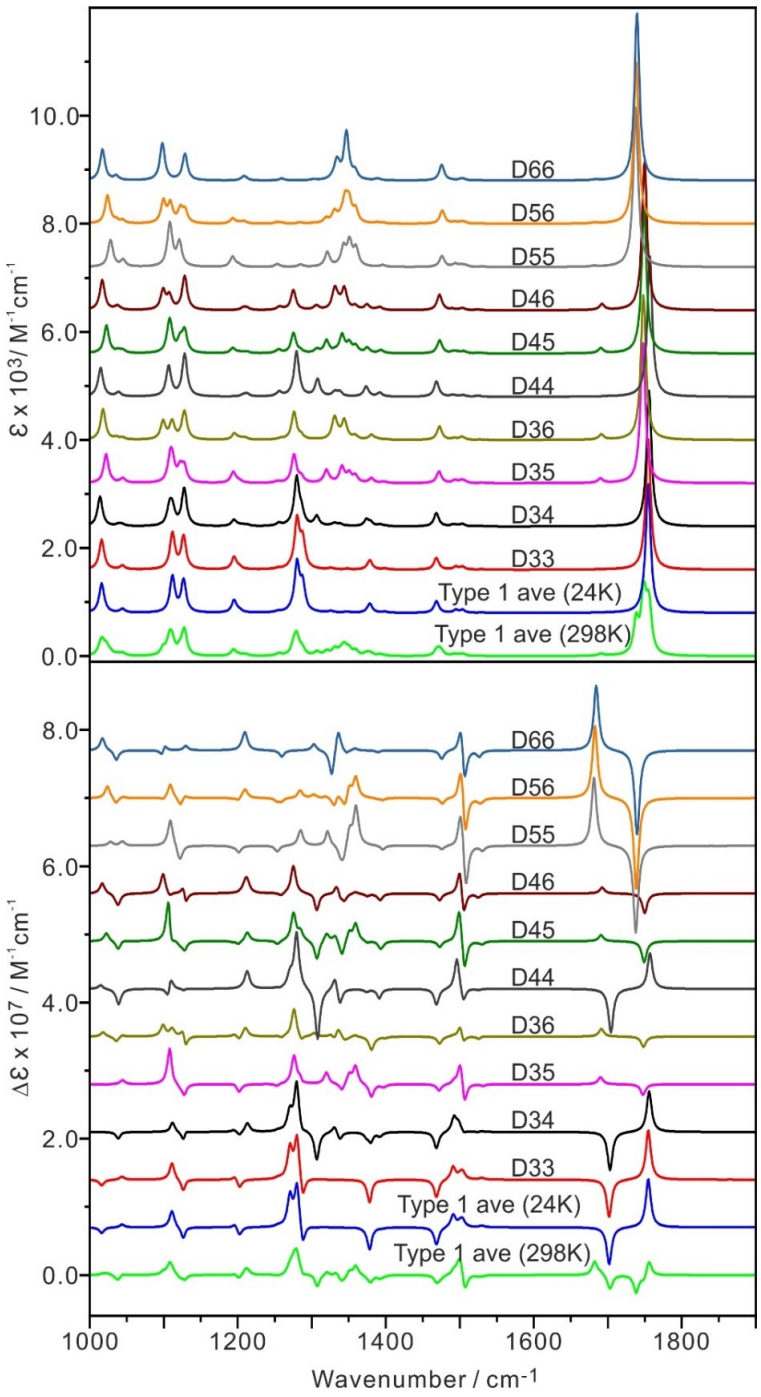


Figure A4. Theoretical MI-IR and MI-VCD spectra of the ten **Type 1** $(THFA)_2$ conformers at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory. The Boltzmann averaged spectra at 24 K and 298 K are also provided.

Table A1. The relative free energies, ΔG in kJ mol^{-1}, and Boltzmann percentage factors, B_f in %, of the conformers of the **Type 2** $(THFA)_2$ at 24, 30 and 298 K at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) and the B3LYP-D3BJ/6-311++G (2d,p)/ PCM(Ar) levels of theory.

Type 2 $(THFA)_2$ conformers	def2-TZVP/Ar		6-311++G(2d,p)/Ar	
	ΔG	B_f (24K, 30K, 298K)	ΔG	B_f (24K, 30K,298K)
D1+(Oe)3(Oc)	0.0[a]	99.6, 98.3, 9.0	0.0[b]	96.4, 92.3, 5.5
D1+(Oe)4(Oc)	1.2	0.2, 0.6, 4.2	1.4	0.1, 0.3, 2.7
D1(Oe)5(Oc)	1.3	0.1, 0.5, 5.7	0.7	3.1, 6.0, 5.2
D1(Oe)3(Oc)	1.4	0.1, 0.5, 14.4	1.3	0.1, 0.5, 7.1
D1(Oe)6(Oc)	1.8	0.0, 0.1, 3.8	1.2	0.3, 0.8, 4.5
$D2^+$(Oe)3(Oc)	2.0	0.0, 0.0, 4.0	2.0	0.0, 0.0, 2.7
D1+(Oe)5(Oc)	2.4	0.0, 0.0, 4.0	2.3	0.0, 0.0, 2.9
D1(Oc)3(vb)	2.7	0.0, 0.0, 5.1	2.3	0.0, 0.0, 6.6
D1(Oe)4(Oc)	2.8	0.0, 0.0, 2.9	2.5	0.0, 0.0, 2.6
D1+(Oe)6(Oc)	2.9	0.0, 0.0, 3.0	2.7	0.0, 0.0, 2.7
D2+(Oe)4(Oc)	3.1	0.0, 0.0, 2.4	3.2	0.0, 0.0, 1.7
D1(Oc)3(va)	3.1	0.0, 0.0, 3.1	2.8	0.0, 0.0, 3.3
D2(Oc)3(vb)	3.2	0.0, 0.0, 4.1	3.0	0.0, 0.0, 3.3
D1(Oc)5(vb)	3.5	0.0, 0.0, 3.5	2.4	0.0, 0.0, 20.9
D2(Oe)5(Oc)	3.8	0.0, 0.0, 1.7	3.2	0.0, 0.0, 1.8
D1(Oc)6(vb)	3.9	0.0, 0.0, 2.6	3.2	0.0, 0.0, 3.1
D1(Oc)4(vb)	4.0	0.0, 0.0, 3.3	3.7	0.0, 0.0, 3.5
D2(Oc)5(vb)	4.2	0.0, 0.0, 4.3	3.6	0.0, 0.0, 4.0
D2(Oc)4(vb)	4.2	0.0, 0.0, 4.4	4.1	0.0, 0.0, 3.0
D2(Oe)6(Oc)	4.3	0.0, 0.0, 1.2	3.6	0.0, 0.0, 1.9
D1(Oc)4(va)	4.3	0.0, 0.0, 2.6	4.2	0.0, 0.0, 1.6
D2(Oc)6(vb)	4.7	0.0, 0.0, 3.4	4.2	0.0, 0.0, 2.4
D2+(Oe)5(Oc)	4.8	0.0, 0.0, 1.6	4.6	0.0, 0.0, 1.4
D2+(Oe)6(Oc)	5.1	0.0, 0.0, 1.5	5.0	0.0, 0.0, 1.3
D2(Oc)3(va)	5.2	0.0, 0.0, 1.3	4.6	0.0, 0.0, 1.3
D2(Oc)4(va)	6.3	0.0, 0.0, 1.0	5.8	0.0, 0.0, 0.9
D1(Oc)5(va)	6.4	0.0, 0.0, 0.9	5.8	0.0, 0.0, 1.0
D1(Oc)6(va)	7.0	0.0, 0.0, 0.5	6.5	0.0, 0.0, 0.5

D2(Oc)5(va)	9.0	0.0, 0.0, 0.3	8.2	0.0, 0.0, 0.3
D2(Oc)6(va)	9.4	0.0, 0.0, 0.2	8.8	0.0, 0.0, 0.3

[a] $D1^{+}$(Oe)3(Oc) is the most stable **Type 2** conformer at the B3LYP-D3BJ/def2-TZVP/PCM(Ar). Its relative free energy, ΔG, is 14.9 kJ mol^{-1} relative to the global minimum D33 at 24 K.

[b] $D1^{+}$(Oe)3(Oc) is the most stable **Type 2** conformer at the B3LYP-D3BJ/6-311++G (2d,p)/PCM(Ar). Its relative free energy, ΔG, is 15.7 kJ mol^{-1} relative to the global minimum D33 at 24 K.

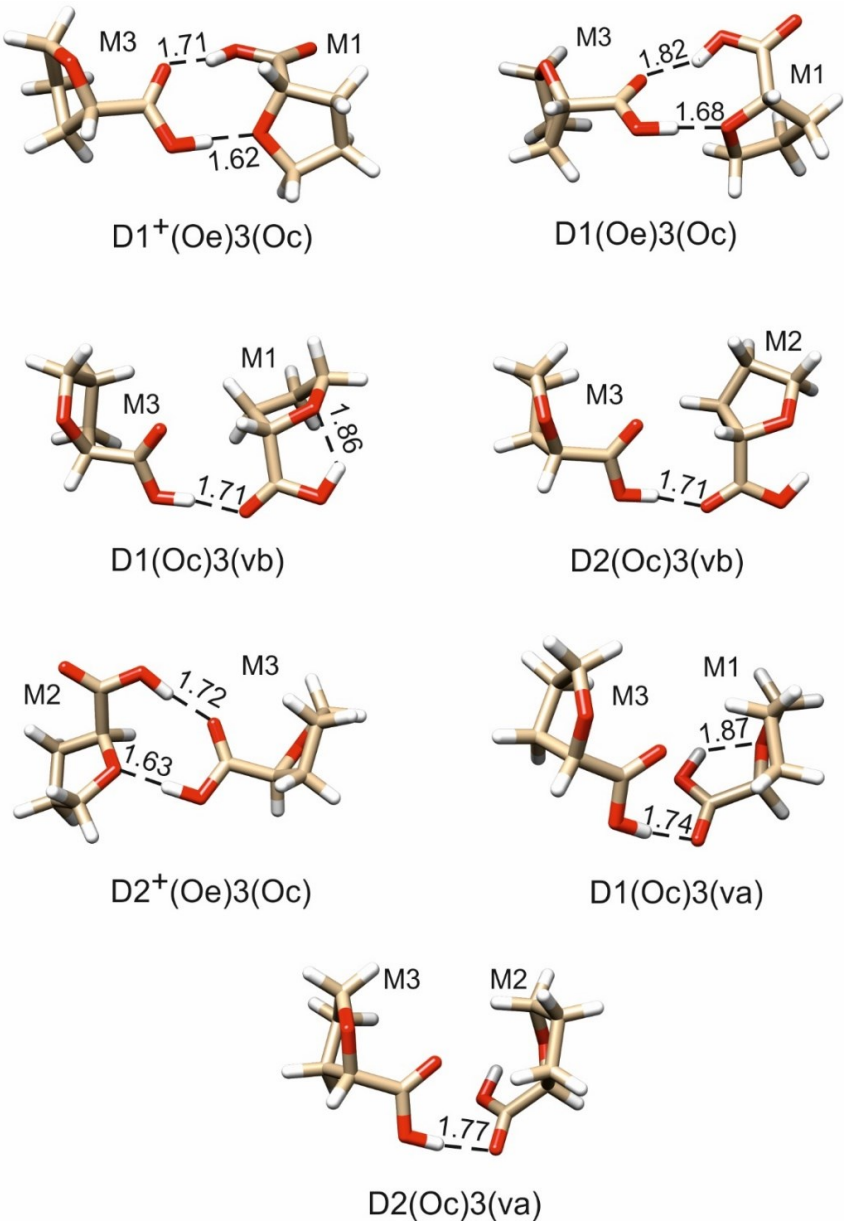


Figure A5. Optimized geometries of the seven **Type 2** binary THFA conformers at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

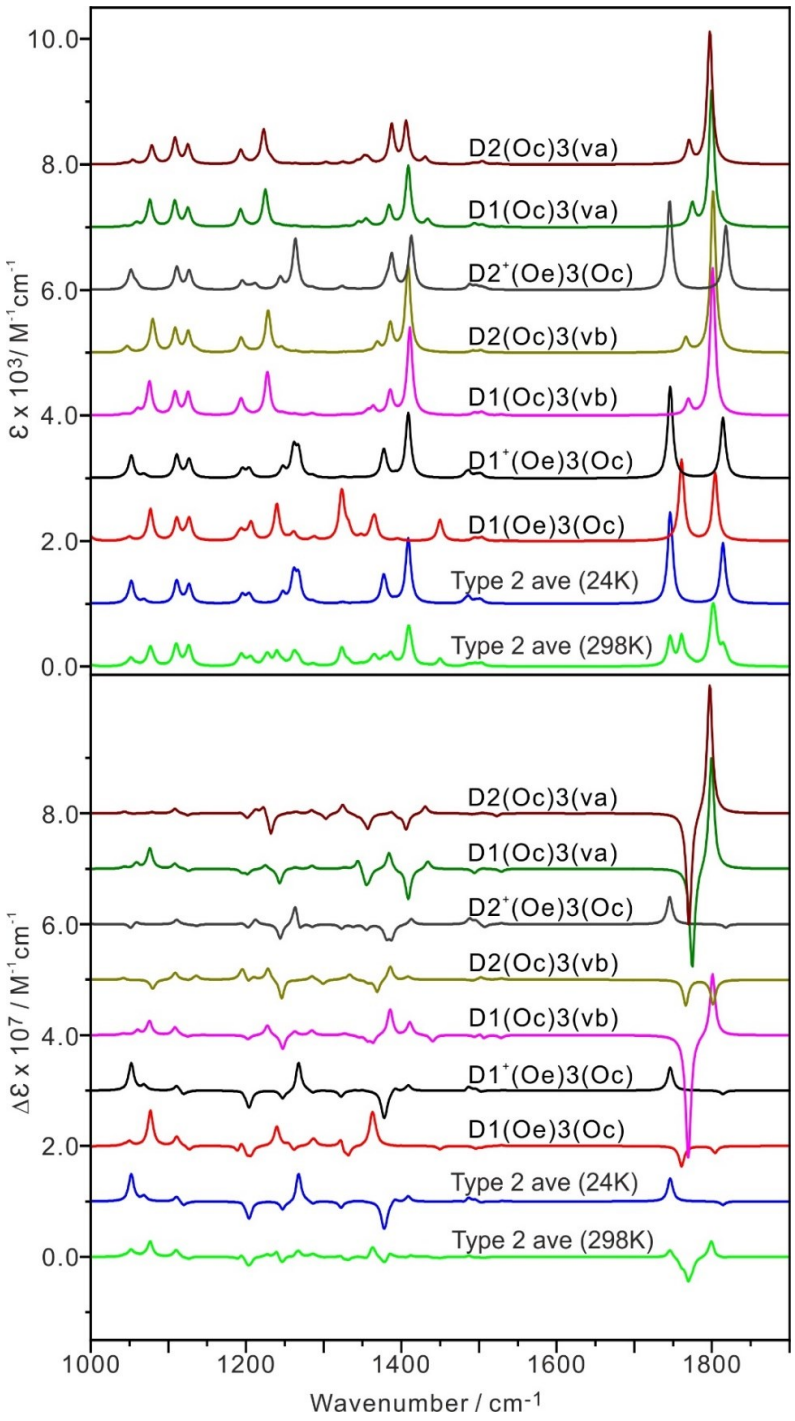


Figure A6. Theoretical MI-IR and MI-VCD spectra of the ten **Type 2** $(THFA)_2$ conformers at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory. The Boltzmann averaged spectra at 24 K and 298 K are also provided.

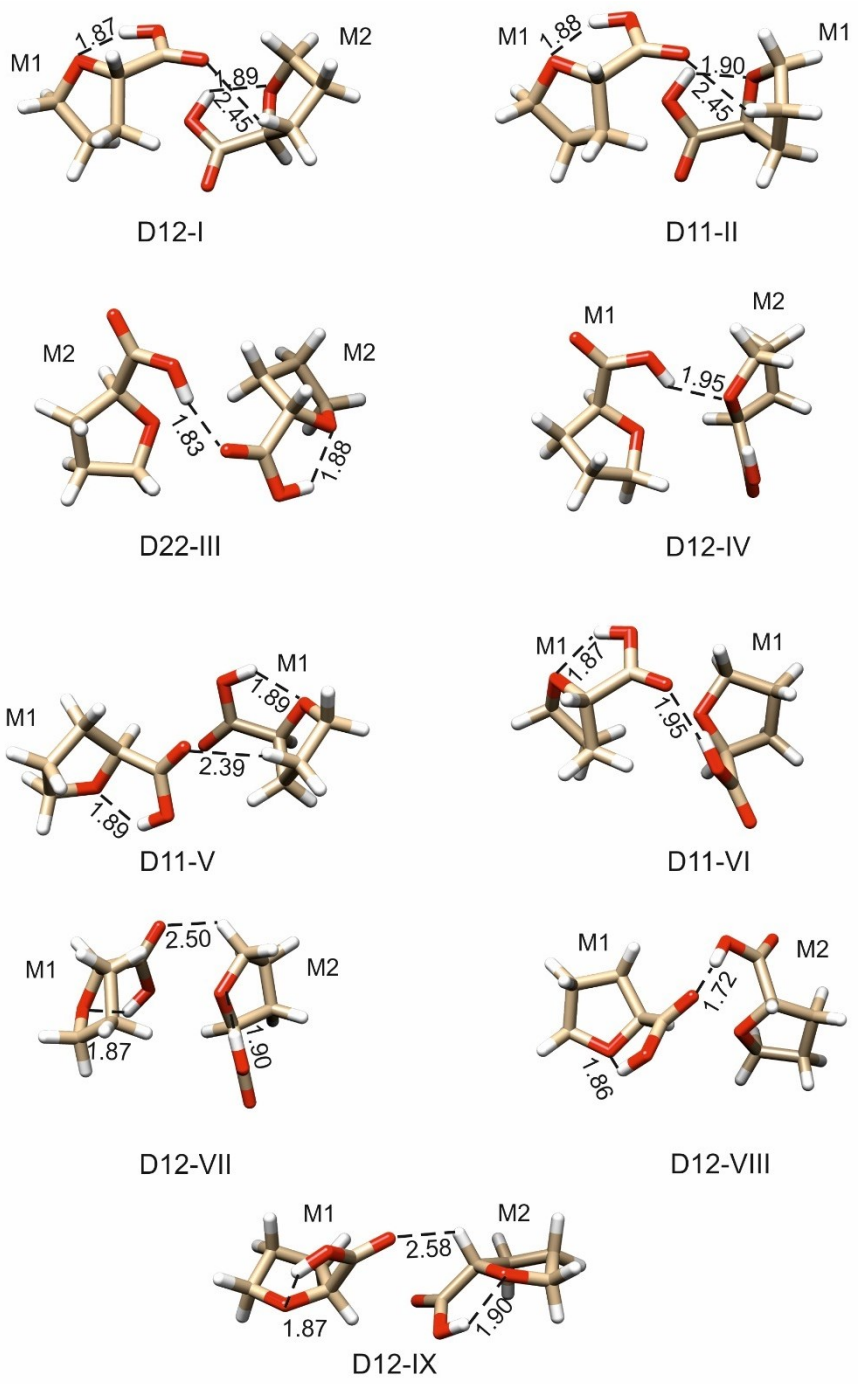


Figure A7. Optimized geometries of the nine **Type 3** binary THFA conformers at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

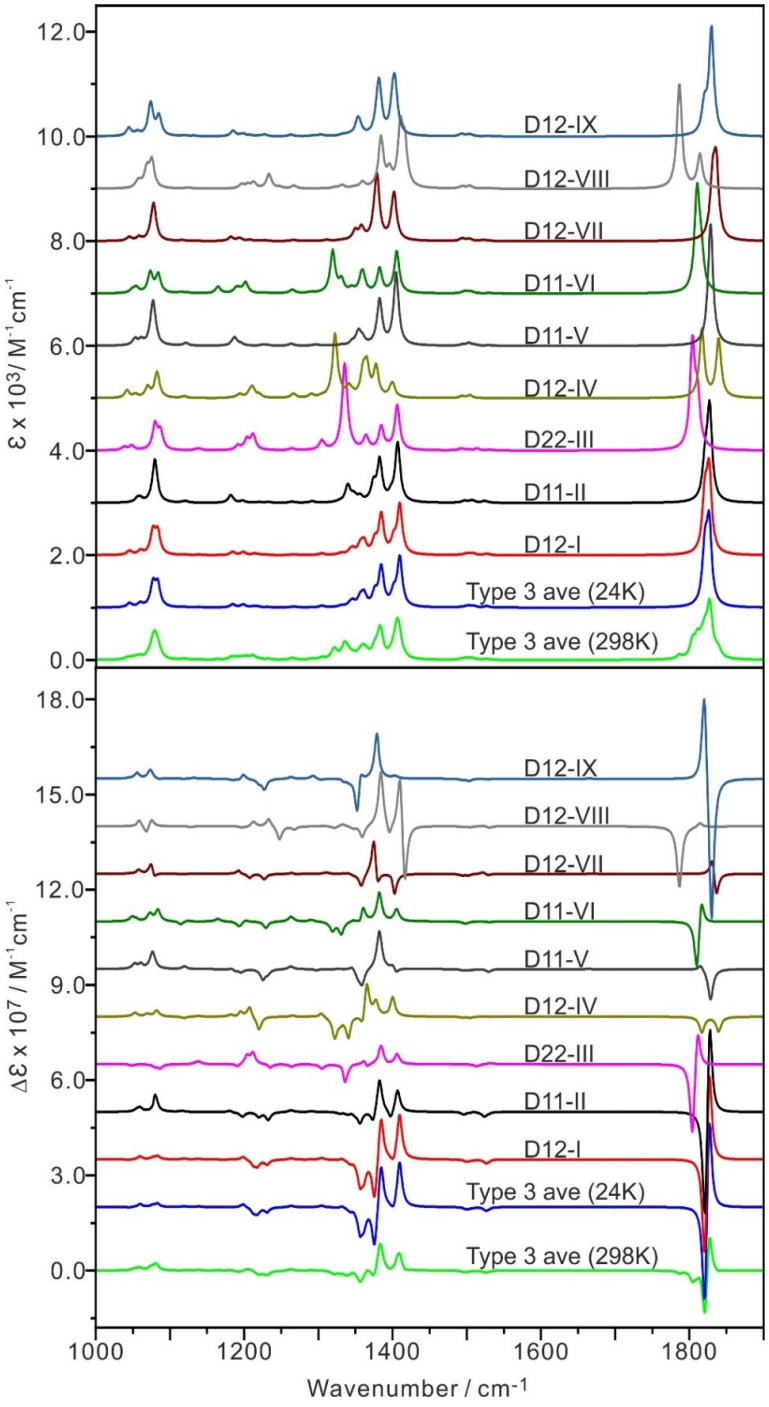


Figure A8. Theoretical MI-IR and MI-VCD spectra of the ten **Type 3** $(THFA)_2$ conformers at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory. The Boltzmann averaged spectra at 24 K and 298 K are also provided.

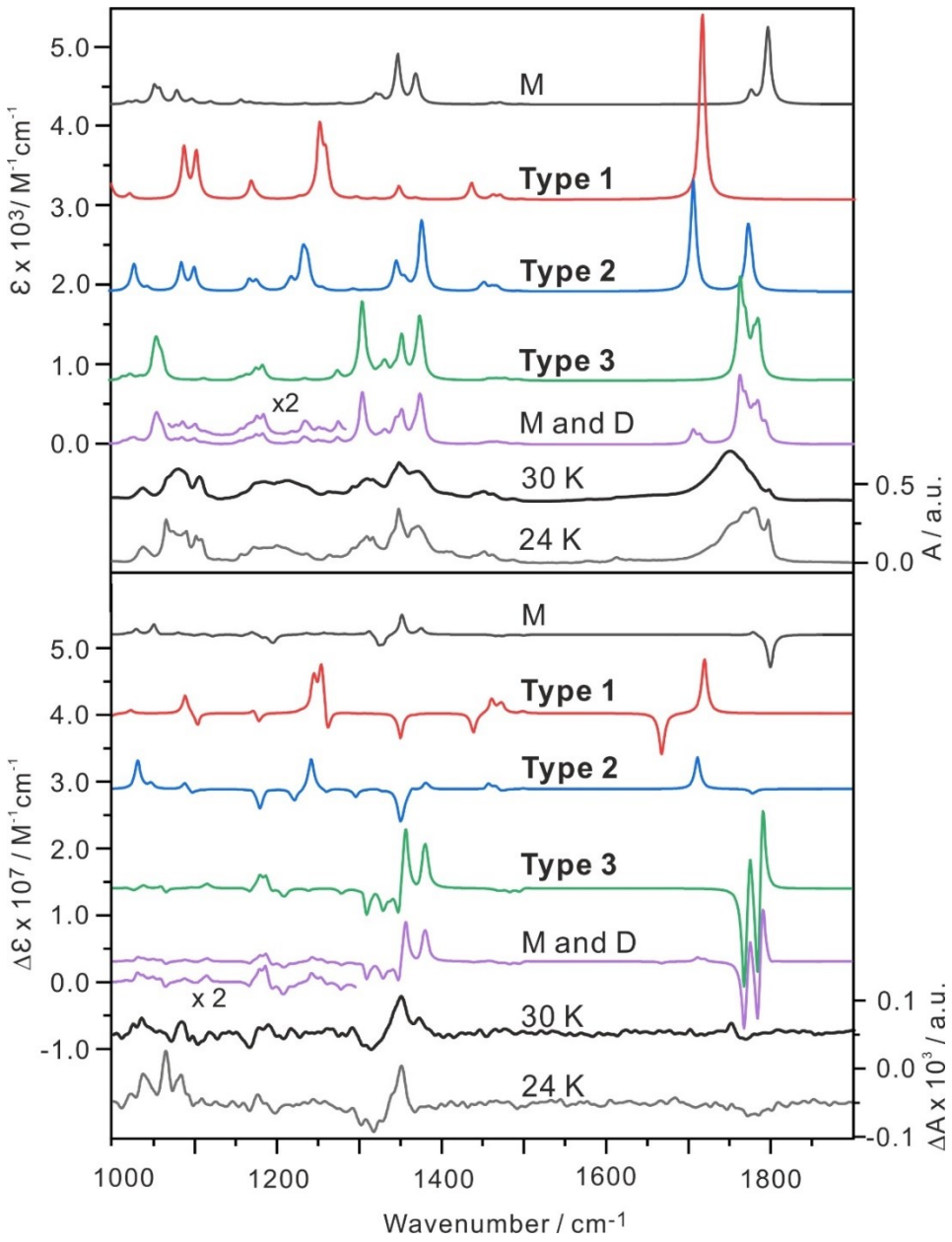


Figure A9. Comparisons of experimental MI-IR and MI-VCD spectra of THFA at 30 K and 24 K with the related simulated IR and VCD spectra of the monomeric and the **Type 1**, **2** and **3** $(THFA)_2$ species. In Type **2** and **3**, we have 0.8*D13+0.2*D23 and 0.6*D22+0.1*D11+0.3*D12, respectively. Type **1**:**2**:**3** = 0.015:0.15:0.80 and M:D = 0.2:0.8.

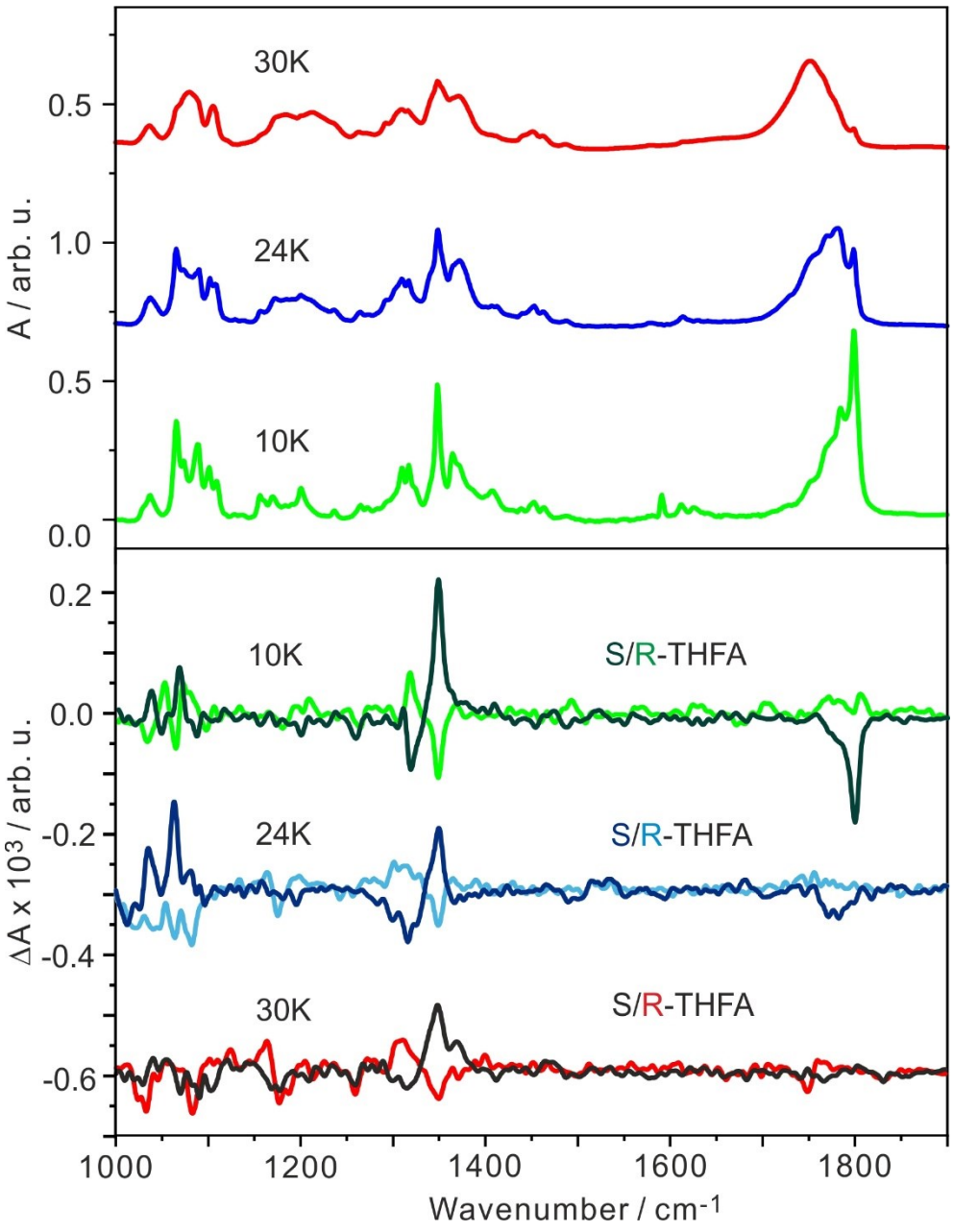


Figure A10. The experimental MI-IR and MI-VCD data of S- and R-THFA under three different deposition temperatures, 10 K, 24 K, and 30 K.

Table A2. Cartesian coordinates of the optimized conformers of THFA and $(THFA)_2$

Cartesian coordinates for M1 calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	0.6723	-1.01238	0.61338
O	-1.65091	-1.20286	-0.50927
O	-2.24748	0.90952	-0.09373
C	0.74779	1.33424	0.17449
C	-0.1093	0.19078	0.7377
C	1.69704	0.59991	-0.78212
C	1.92116	-0.72071	-0.0544
C	-1.44532	0.02126	0.00129
H	1.30924	1.80729	0.98241
H	0.13404	2.09478	-0.3044
H	-0.35281	0.34382	1.79209
H	1.21555	0.42833	-1.74788
H	2.62661	1.14164	-0.95846
H	2.15113	-1.55972	-0.71134
H	2.7098	-0.63766	0.70039
H	-0.85294	-1.72563	-0.29762

Cartesian coordinates for M2 calculated at B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	0.60108	-1.01943	0.50173
O	-1.81906	-1.15505	-0.42761
O	-2.32091	0.97285	0.0363

C	0.72951	1.23472	-0.26767
C	-0.10136	0.23652	0.55463
C	2.14183	0.64946	-0.15342
C	1.86958	-0.85104	-0.17066
C	-1.52802	0.07149	0.03564
H	0.64004	2.24894	0.11658
H	0.39622	1.24051	-1.30805
H	-0.17284	0.55136	1.59989
H	2.79976	0.96299	-0.96459
H	2.60403	0.9408	0.79282
H	1.78769	-1.2331	-1.19394
H	2.60951	-1.44328	0.36773
H	-1.0226	-1.70205	-0.28489

Cartesian coordinates for M3 calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	0.82526	-0.711	1.0218
C	1.90201	-0.70409	0.06136
C	1.62777	0.45692	-0.9032
C	0.66923	1.3511	-0.10348
H	2.83527	-0.57096	0.61547
H	1.92573	-1.67081	-0.4437
H	2.54038	0.97805	-1.19376
H	1.14164	0.09275	-1.80907
H	1.21784	2.00919	0.57322
H	0.02258	1.96786	-0.72766

C	-1.29734	-0.22842	-0.04483
O	-1.29133	-1.18251	-0.7755
O	-2.39783	0.53903	0.15816
H	-3.10653	0.16999	-0.39254
C	-0.10147	0.31334	0.73453
H	-0.46693	0.72667	1.67602

Cartesian coordinates for M4 calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	0.75949	-0.84159	0.87047
C	1.78977	-0.857	-0.12406
C	2.08322	0.61706	-0.38604
C	0.68502	1.25351	-0.29057
H	2.63026	-1.41331	0.28983
H	1.43259	-1.36572	-1.02642
H	2.73941	1.01072	0.3928
H	2.55619	0.79088	-1.35346
H	0.70047	2.25783	0.12976
H	0.21993	1.31368	-1.27593
C	-1.38664	-0.17985	-0.039
O	-1.53441	-1.15536	-0.7241
O	-2.38135	0.7112	0.20342
H	-3.16212	0.40381	-0.28401
C	-0.09224	0.26221	0.62925
H	-0.34131	0.71264	1.59169

Cartesian coordinates for M5 calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	0.75787	-0.68067	1.03184
C	1.88301	-0.72025	0.13128
C	1.63217	0.35477	-0.92898
C	0.71266	1.33964	-0.19473
H	2.78884	-0.51886	0.71153
H	1.95222	-1.72511	-0.28833
H	2.55599	0.81964	-1.27466
H	1.12249	-0.07218	-1.79418
H	1.29329	2.02182	0.42954
H	0.08601	1.93705	-0.85627
C	-1.39907	-0.03372	0.04635
O	-2.41966	0.60492	0.10284
O	-1.28965	-1.18213	-0.65497
H	-2.16004	-1.34728	-1.05161
C	-0.10001	0.40181	0.71948
H	-0.39977	0.89442	1.64593

Cartesian coordinates for M6 calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	0.6785	-0.80139	0.89112
C	1.78006	-0.89132	-0.02102
C	2.11422	0.55523	-0.36675
C	0.72012	1.19744	-0.42821
H	2.58134	-1.42546	0.48943

H	1.48452	-1.45949	-0.91105
H	2.7116	1.00323	0.43039
H	2.66304	0.6532	-1.30434
H	0.7156	2.25547	-0.17124
H	0.29787	1.09436	-1.42958
C	-1.47337	0.02888	0.05576
O	-2.45107	0.70168	0.26482
O	-1.48954	-1.0681	-0.73171
H	-2.40144	-1.17045	-1.04704
C	-0.09145	0.35747	0.5973
H	-0.25309	0.9055	1.52708

Cartesian coordinates for M7 calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	0.80407	-0.6413	1.07211
O	-2.48676	0.40807	0.04362
O	-1.22351	-1.2202	-0.78404
C	0.6451	1.34247	-0.18654
C	-0.12859	0.34697	0.70126
C	1.63295	0.41252	-0.90597
C	1.90584	-0.67551	0.13923
C	-1.30598	-0.26148	-0.07425
H	1.17007	2.05387	0.4544
H	0.00621	1.90389	-0.8695
H	-0.49986	0.81322	1.62049
H	1.16892	-0.0195	-1.79269

H	2.54242	0.92969	-1.21268
H	1.95793	-1.67272	-0.29904
H	2.82157	-0.48723	0.70623
H	-2.4135	1.11633	0.6969

Cartesian coordinates for M8 calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	0.74048	-0.79564	0.91344
O	-2.49159	0.58793	0.12212
O	-1.47169	-1.1732	-0.76737
C	0.66993	1.23446	-0.35627
C	-0.11436	0.28531	0.6023
C	2.07769	0.61259	-0.38692
C	1.79544	-0.84933	-0.056
C	-1.40012	-0.20476	-0.0707
H	0.66989	2.263	0.00361
H	0.2242	1.22784	-1.3521
H	-0.36262	0.78037	1.54787
H	2.56877	0.74266	-1.35196
H	2.71147	1.0552	0.38413
H	1.46358	-1.40863	-0.93676
H	2.63189	-1.37162	0.40712
H	-2.29619	1.28879	0.75803

Cartesian coordinates for D33 calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	4.1565	0.19648	-0.93671
C	4.63151	1.03608	0.1372
C	4.09496	0.4321	1.44341
C	3.75241	-1.0094	1.04002
H	5.72405	1.03186	0.09658
H	4.2734	2.05345	-0.02801
H	4.82357	0.47925	2.25291
H	3.19473	0.95873	1.76236
H	4.63839	-1.64653	1.06566
H	2.98127	-1.46871	1.65843
C	1.85215	-0.43603	-0.49312
O	1.47456	0.72723	-0.45612
O	1.05329	-1.48235	-0.54022
H	0.08966	-1.19644	-0.51831
C	3.32503	-0.82283	-0.42756
H	3.46657	-1.72431	-1.02541
O	-4.15656	-0.1963	-0.93671
C	-4.63152	-1.03608	0.13709
C	-4.09488	-0.43234	1.44336
C	-3.75234	1.00922	1.04021
H	-5.72407	-1.03182	0.09653
H	-4.27346	-2.05343	-0.02833
H	-4.82344	-0.47963	2.2529
H	-3.19464	-0.95905	1.76216
H	-4.63832	1.64636	1.06601
H	-2.98116	1.46841	1.65866

C	-1.85218	0.43611	-0.49314
O	-1.47459	-0.72716	-0.4563
O	-1.0533	1.48243	-0.54008
H	-0.08968	1.1965	-0.51817
C	-3.32504	0.82291	-0.42743
H	-3.46661	1.72449	-1.0251

Cartesian coordinates for D34 calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	-4.2089	0.02037	-0.94076
C	-4.66461	-1.08384	-0.13032
C	-4.13683	-0.83865	1.29097
C	-3.82532	0.66501	1.28683
H	-5.75715	-1.09045	-0.17131
H	-4.28704	-2.01338	-0.55927
H	-4.86193	-1.11482	2.05672
H	-3.22498	-1.41248	1.4606
H	-4.72439	1.25378	1.47811
H	-3.06245	0.95913	2.00762
C	-1.91696	0.55992	-0.33762
O	-1.51506	-0.56179	-0.6159
O	-1.14065	1.59624	-0.09641
H	-0.17078	1.33593	-0.14929
C	-3.39749	0.88434	-0.17599
H	-3.55916	1.90931	-0.51283
O	4.09256	0.58657	-0.79494

C	4.6061	0.97167	0.48526
C	4.87909	-0.35379	1.1907
C	3.69661	-1.22344	0.72626
H	5.49473	1.57755	0.31076
H	3.86104	1.5696	1.02257
H	5.82414	-0.77183	0.83884
H	4.92748	-0.25659	2.27601
H	3.96281	-2.2697	0.58705
H	2.87847	-1.18379	1.44742
C	1.80395	-0.19488	-0.57677
O	1.40213	0.90253	-0.21628
O	1.02757	-1.20454	-0.91448
H	0.0583	-0.95943	-0.80877
C	3.27888	-0.55747	-0.62272
H	3.43923	-1.22107	-1.47365

Cartesian coordinates for D35 calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	3.85198	-0.55303	1.21367
C	4.30384	-1.5314	0.25281
C	4.16831	-0.88665	-1.13422
C	4.11171	0.61348	-0.81125
H	5.34139	-1.77475	0.49716
H	3.6942	-2.43057	0.35473
H	4.9975	-1.1426	-1.79399
H	3.24079	-1.20525	-1.61126

H	5.11353	1.02505	-0.67496
H	3.59619	1.20737	-1.5659
C	1.88568	0.57859	0.34267
O	1.25398	-0.4648	0.24279
O	1.36499	1.7833	0.23384
H	0.38035	1.72994	0.03409
C	3.39507	0.60592	0.55166
H	3.63833	1.47827	1.15997
O	-3.85996	-0.53903	-1.20129
C	-4.38667	-1.46687	-0.22858
C	-4.22185	-0.80796	1.14772
C	-4.08496	0.68082	0.79916
H	-5.43606	-1.64519	-0.47882
H	-3.8383	-2.40688	-0.30977
H	-5.06633	-1.00855	1.8073
H	-3.31479	-1.16733	1.63487
H	-5.06401	1.14052	0.65113
H	-3.54346	1.26115	1.54599
C	-1.85346	0.59011	-0.35405
O	-1.21566	1.6363	-0.28707
O	-1.33791	-0.6082	-0.20974
H	-0.34824	-0.54995	-0.03827
C	-3.3636	0.61828	-0.56001
H	-3.57495	1.48653	-1.18528

Cartesian coordinates for D36 calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level

of theory.

O	-3.93866	-0.42358	-1.25276
C	-4.3659	-1.50618	-0.39828
C	-4.20152	-1.02053	1.04926
C	-4.1532	0.50656	0.89573
H	-5.40781	-1.72634	-0.64597
H	-3.75579	-2.3848	-0.61361
H	-5.01623	-1.35163	1.69346
H	-3.26374	-1.38752	1.46787
H	-5.15807	0.92764	0.82829
H	-3.62164	1.01357	1.70094
C	-1.95312	0.6065	-0.30268
O	-1.32035	-0.44042	-0.33321
O	-1.4293	1.79253	-0.07128
H	-0.44111	1.71849	0.10186
C	-3.46669	0.65443	-0.47468
H	-3.72292	1.58914	-0.97548
O	3.78025	-0.6338	1.14396
C	4.40888	-1.47139	0.16578
C	5.0542	-0.49202	-0.80865
C	4.00011	0.62456	-0.88539
H	5.11652	-2.11134	0.69243
H	3.65775	-2.09673	-0.33067
H	5.98809	-0.11206	-0.38929
H	5.26682	-0.93592	-1.78197
H	4.4245	1.61127	-1.06242

H	3.28489	0.41932	-1.68389
C	1.7975	0.54494	0.33005
O	1.16082	1.59224	0.38603
O	1.27913	-0.63108	0.06252
H	0.28799	-0.55583	-0.08899
C	3.30536	0.54286	0.50673
H	3.55176	1.39726	1.13806

Cartesian coordinates for D44 calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	4.14466	0.36422	-0.90351
C	4.63788	1.1023	0.22031
C	4.92531	0.03296	1.27052
C	3.76322	-0.95519	1.06128
H	5.51813	1.6518	-0.11198
H	3.8771	1.81237	0.56419
H	5.88132	-0.44832	1.05585
H	4.95965	0.43116	2.28539
H	4.04986	-1.99322	1.22029
H	2.93732	-0.733	1.73911
C	1.86642	-0.364	-0.48517
O	1.44537	0.78359	-0.44346
O	1.10837	-1.44032	-0.53537
H	0.13425	-1.19229	-0.50706
C	3.34706	-0.70005	-0.42174
H	3.525	-1.57194	-1.05278

O	-4.14465	-0.36539	-0.90302
C	-4.63812	-1.10187	0.22173
C	-4.92546	-0.03105	1.27046
C	-3.76312	0.95653	1.06007
H	-5.51844	-1.65166	-0.1099
H	-3.87751	-1.81162	0.56666
H	-5.88133	0.45015	1.05497
H	-4.96007	-0.42788	2.28585
H	-4.04946	1.99483	1.21778
H	-2.93728	0.73493	1.73817
C	-1.86635	0.36326	-0.48561
O	-1.44529	-0.78429	-0.4429
O	-1.10829	1.43954	-0.5366
H	-0.13417	1.19153	-0.50802
C	-3.34697	0.69942	-0.42265
H	-3.52483	1.57053	-1.05479

Cartesian coordinates for D45 calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	-3.92674	-0.44658	-1.24848
C	-4.46054	-1.4375	-0.34436
C	-4.29165	-0.87957	1.07554
C	-4.14025	0.62897	0.83406
H	-5.511	-1.58999	-0.60684
H	-3.91887	-2.3733	-0.49205
H	-5.13898	-1.1186	1.71845

H	-3.38878	-1.28133	1.53666
H	-5.11473	1.10755	0.71948
H	-3.59352	1.14915	1.62036
C	-1.90913	0.5989	-0.3232
O	-1.26119	1.63145	-0.18225
O	-1.40574	-0.61163	-0.26466
H	-0.41548	-0.57619	-0.08914
C	-3.41887	0.6567	-0.52616
H	-3.62128	1.56949	-1.08769
O	3.78962	-0.71473	1.12603
C	4.30726	-1.53153	0.06918
C	4.99256	-0.53964	-0.86628
C	4.04378	0.6715	-0.81369
H	4.98031	-2.25984	0.52064
H	3.48732	-2.05891	-0.43155
H	5.97462	-0.27527	-0.46951
H	5.12163	-0.92743	-1.8775
H	4.56281	1.62605	-0.87963
H	3.31893	0.63229	-1.62851
C	1.82854	0.49929	0.36773
O	1.18551	-0.52768	0.19972
O	1.32071	1.71452	0.34568
H	0.3356	1.68576	0.14321
C	3.33477	0.50262	0.56689
H	3.58531	1.3095	1.25708

Cartesian coordinates for D46 calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	3.84849	0.55201	-1.19793
C	4.48418	1.44799	-0.27783
C	5.11982	0.53067	0.76133
C	4.05574	-0.56881	0.91179
H	5.19806	2.04435	-0.8456
H	3.73858	2.11237	0.17425
H	6.05057	0.11503	0.36975
H	5.33577	1.03727	1.70278
H	4.47142	-1.54536	1.15381
H	3.34281	-0.30484	1.69494
C	1.85324	-0.54719	-0.30573
O	1.20588	-1.58932	-0.288
O	1.34764	0.64988	-0.12052
H	0.35597	0.59621	0.03744
C	3.36089	-0.57347	-0.48279
H	3.59728	-1.47118	-1.05506
O	-3.87255	0.60146	1.18578
C	-4.36553	1.52416	0.20732
C	-5.03318	0.63579	-0.83865
C	-4.09001	-0.58008	-0.889
H	-5.04626	2.20573	0.71653
H	-3.53369	2.09586	-0.21963
H	-6.02466	0.33812	-0.49215
H	-5.1386	1.12492	-1.80779

H	-4.61214	-1.52044	-1.05635
H	-3.35128	-0.46643	-1.68409
C	-1.8946	-0.52486	0.34011
O	-1.25144	0.51453	0.29544
O	-1.38367	-1.73035	0.19376
H	-0.39595	-1.67894	0.01029
C	-3.40393	-0.55144	0.51329
H	-3.66239	-1.42493	1.11357

Cartesian coordinates for D55 calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	4.16842	0.06468	-0.99467
C	4.75449	0.94001	-0.00781
C	4.23166	0.47966	1.36003
C	3.7855	-0.96372	1.08696
H	5.84087	0.8459	-0.0885
H	4.46974	1.96786	-0.23888
H	4.9928	0.54362	2.13788
H	3.3793	1.08779	1.66487
H	4.63195	-1.65142	1.13514
H	3.01316	-1.32329	1.76673
C	1.84189	-0.47235	-0.4307
O	0.96327	-1.32888	-0.41091
O	1.62204	0.82124	-0.44694
H	0.63393	1.01294	-0.44285
C	3.3118	-0.87308	-0.37546

H	3.3929	-1.8254	-0.90073
O	-4.16832	-0.06794	-0.99445
C	-4.75513	-0.93945	-0.00471
C	-4.23169	-0.47511	1.36152
C	-3.78539	0.96731	1.08381
H	-5.84145	-0.84455	-0.08549
H	-4.47144	-1.96832	-0.23249
H	-4.99249	-0.53666	2.13988
H	-3.37927	-1.08246	1.66782
H	-4.63183	1.65521	1.12964
H	-3.01304	1.32911	1.76237
C	-1.84176	0.4707	-0.43202
O	-0.96309	1.32725	-0.41373
O	-1.62197	-0.82291	-0.4453
H	-0.63388	-1.01463	-0.44063
C	-3.31163	0.87175	-0.3783
H	-3.39253	1.82237	-0.90668

Cartesian coordinates for D56 calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	-4.22117	0.23565	-0.96959
C	-4.80333	-0.91301	-0.31762
C	-4.29821	-0.90398	1.13192
C	-3.86804	0.55579	1.33464
H	-5.89018	-0.80941	-0.37769
H	-4.50262	-1.81193	-0.85846

H	-5.06515	-1.21775	1.84032
H	-3.44089	-1.56938	1.23925
H	-4.72354	1.185	1.58759
H	-3.10604	0.68906	2.1024
C	-1.90578	0.58703	-0.23712
O	-1.03789	1.40016	0.06487
O	-1.67028	-0.63249	-0.66099
H	-0.68017	-0.80696	-0.70594
C	-3.38094	0.93671	-0.07557
H	-3.46907	2.00554	-0.27364
O	4.09004	0.42119	-0.94322
C	4.77904	0.93476	0.20317
C	5.03366	-0.29391	1.06975
C	3.73464	-1.0945	0.88029
H	5.68362	1.42616	-0.15439
H	4.15066	1.66902	0.72061
H	5.88914	-0.85077	0.68195
H	5.22602	-0.047	2.11459
H	3.87671	-2.17228	0.93761
H	2.9994	-0.81288	1.63624
C	1.78921	-0.27693	-0.49319
O	0.92116	-1.1017	-0.76123
O	1.5549	0.95204	-0.09648
H	0.56553	1.12522	-0.04303
C	3.25897	-0.65568	-0.53719
H	3.35725	-1.45832	-1.26886

Cartesian coordinates for D66 calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	-4.14662	-0.13118	-1.01738
C	-4.82421	-0.97924	-0.08244
C	-5.09284	-0.07641	1.11685
C	-3.80579	0.76232	1.18133
H	-5.72314	-1.35146	-0.57333
H	-4.18382	-1.82711	0.18744
H	-5.95734	0.55972	0.91604
H	-5.27823	-0.63232	2.03679
H	-3.96273	1.76971	1.56289
H	-3.06441	0.27529	1.81737
C	-1.85318	0.42667	-0.37139
O	-0.99684	1.30525	-0.36006
O	-1.60199	-0.86154	-0.38124
H	-0.61036	-1.02972	-0.3805
C	-3.32794	0.78136	-0.30189
H	-3.43799	1.76731	-0.75471
O	4.14679	0.13044	-1.01751
C	4.82355	0.97956	-0.08298
C	5.0925	0.0777	1.11698
C	3.80607	-0.76204	1.18163
H	5.72232	1.3521	-0.57391
H	4.18254	1.82717	0.18622
H	5.95755	-0.5579	0.9169

H	5.27717	0.63438	2.03659
H	3.96393	-1.76925	1.56331
H	3.06434	-0.27563	1.81774
C	1.85334	-0.42699	-0.37109
O	0.99698	-1.30556	-0.35917
O	1.60217	0.8612	-0.3817
H	0.61055	1.02939	-0.38109
C	3.32811	-0.78171	-0.30154
H	3.43804	-1.76784	-0.75398

Cartesian coordinates for D1(Oc)3(va) calculated at the B3LYP-D3BJ/def2-TZVP /PCM(Ar) level of theory.

O	-3.27306	-0.83059	-0.99155
O	-1.20555	2.02838	-0.49108
O	-0.61856	-0.1448	-0.53217
C	-3.47363	0.46121	0.95421
C	-2.96364	0.45278	-0.49528
C	-3.17692	-0.97828	1.38988
C	-3.34205	-1.77546	0.08864
C	-1.46301	0.72157	-0.52447
H	-4.54676	0.65656	0.94907
H	-2.98892	1.21441	1.57291
H	-3.46053	1.19659	-1.11851
H	-2.15353	-1.05647	1.75641
H	-3.84655	-1.32447	2.17621
H	-2.55561	-2.52102	-0.03886

H	-4.31293	-2.27376	0.02957
H	-0.2289	2.17587	-0.43463
O	3.3435	-0.61741	0.88245
C	3.15212	-1.97176	0.41765
C	2.18959	-1.86226	-0.75374
C	2.60378	-0.53202	-1.3897
H	4.12216	-2.37686	0.1135
H	2.77387	-2.55615	1.25536
H	2.27661	-2.70168	-1.44251
H	1.15853	-1.79974	-0.40824
H	3.45332	-0.66544	-2.06011
H	1.79702	-0.05301	-1.93751
C	1.94818	1.25805	0.30455
O	1.51671	2.1502	-0.38985
O	1.51635	1.04384	1.53974
H	2.01934	0.26835	1.86614
C	3.03662	0.30397	-0.17366
H	3.91706	0.9131	-0.39095

Cartesian coordinates for D1(Oc)3(vb) calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	3.7757	0.08961	1.05579
O	1.11274	1.81013	-0.57063
O	1.00065	0.07517	0.85753
C	3.49201	-0.08179	-1.26188
C	3.15203	0.75516	-0.01883

C	3.45234	-1.50523	-0.69301
C	3.89789	-1.31177	0.7644
C	1.63974	0.83444	0.16545
H	4.49677	0.17854	-1.5969
H	2.80196	0.07983	-2.08846
H	3.54692	1.7696	-0.07554
H	2.43644	-1.89851	-0.72967
H	4.10039	-2.19103	-1.2373
H	3.27665	-1.87743	1.46031
H	4.94348	-1.58895	0.91807
H	0.12507	1.82389	-0.47778
O	-3.29693	-0.80537	0.94661
C	-3.56851	-1.99838	0.1844
C	-2.82245	-1.81375	-1.12862
C	-1.56367	-1.07029	-0.67348
H	-3.195	-2.86419	0.73936
H	-4.64888	-2.0875	0.07834
H	-2.60298	-2.75983	-1.6216
H	-3.40819	-1.19898	-1.81546
H	-0.81805	-1.75837	-0.27823
H	-1.0932	-0.48077	-1.45651
C	-2.40944	1.23327	0.02219
O	-1.58392	1.99048	-0.4357
O	-3.67771	1.5937	0.17393
H	-4.12769	0.81482	0.56476
C	-2.08774	-0.18907	0.46951

H	-1.35739	-0.11676	1.27454

Cartesian coordinates for D1(Oe)3(Oc) calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	3.55954	1.17747	-0.2052
O	0.81291	-0.57362	-1.61913
O	0.91076	0.50547	0.34639
C	3.61597	-1.16637	-0.26637
C	2.96134	0.08916	-0.86666
C	3.81517	-0.7413	1.19302
C	4.04776	0.77145	1.08614
C	1.45701	0.05292	-0.63981
H	4.57304	-1.33103	-0.76247
H	3.01134	-2.06468	-0.38324
H	3.14269	0.18512	-1.93745
H	2.91601	-0.94618	1.77359
H	4.64961	-1.25656	1.66678
H	3.51262	1.32619	1.85736
H	5.10772	1.03171	1.13302
H	-0.14403	-0.65149	-1.36672
O	-1.64115	-0.85038	-0.60363
C	-1.68082	-1.83415	0.46158
C	-2.99575	-1.59212	1.19601
C	-3.87339	-0.96038	0.11131
H	-1.64166	-2.8232	0.00051
H	-0.79968	-1.69293	1.08706

H	-3.41356	-2.51177	1.60226
H	-2.85118	-0.89163	2.01999
H	-4.27356	-1.7243	-0.5571
H	-4.69648	-0.36614	0.49858
C	-2.7404	1.29495	-0.00591
O	-3.70856	2.00835	0.03633
O	-1.56632	1.66789	0.4898
H	-0.83148	1.02465	0.36432
C	-2.86141	-0.09785	-0.64195
H	-3.1309	0.05277	-1.68727

Cartesian coordinates for $D1^{+}$(Oe)3(Oc) calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	-3.62099	0.02333	1.18618
O	-0.72577	-1.62326	-0.02869
O	-1.03434	0.57012	0.33993
C	-3.62659	-0.80784	-1.00956
C	-2.94682	-0.90182	0.36704
C	-3.97159	0.6839	-1.08104
C	-4.23711	1.0483	0.3842
C	-1.46823	-0.56299	0.24003
H	-4.52859	-1.42071	-0.99516
H	-2.99289	-1.15027	-1.82634
H	-3.03384	-1.89412	0.81024
H	-3.12445	1.25073	-1.466
H	-4.8312	0.8825	-1.71963

H	-3.80786	2.01344	0.65445
H	-5.30357	1.05556	0.6219
H	0.22977	-1.35075	-0.15165
O	1.78348	-0.86886	-0.2978
C	2.83911	-1.77852	-0.69265
C	4.13043	-1.14559	-0.19482
C	3.66589	-0.38796	1.05255
H	2.65055	-2.74644	-0.22031
H	2.79212	-1.89293	-1.77475
H	4.89638	-1.8899	0.0181
H	4.52242	-0.44227	-0.92902
H	3.57101	-1.06328	1.90454
H	4.32626	0.43299	1.31799
C	2.33187	1.47588	-0.08074
O	3.35703	2.06861	-0.2955
O	1.15057	1.94599	-0.46773
H	0.39827	1.3755	-0.17906
C	2.29252	0.11621	0.64015
H	1.58739	0.17299	1.47186

Cartesian coordinates for D2(Oc)3(va) calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	-3.09116	0.93811	0.97468
O	-1.08005	-1.9916	0.72093
O	-0.47975	0.16125	0.44045
C	-3.43177	-0.59839	-0.76385

C	-2.8247	-0.40414	0.63443
C	-3.16071	0.76681	-1.40596
C	-3.2382	1.73113	-0.21477
C	-1.32742	-0.68796	0.59998
H	-4.50265	-0.77809	-0.66076
H	-2.99287	-1.43296	-1.30775
H	-3.28178	-1.05174	1.38301
H	-2.16355	0.78401	-1.84559
H	-3.87957	1.013	-2.18637
H	-2.44762	2.4827	-0.24168
H	-4.20438	2.23845	-0.15616
H	-0.11371	-2.16678	0.6146
O	2.91217	0.81433	-1.04307
C	2.25971	1.98175	-0.49576
C	2.66671	1.98666	0.96848
C	2.65404	0.49276	1.32356
H	2.61165	2.83951	-1.06563
H	1.17605	1.88249	-0.59992
H	3.6682	2.4063	1.07764
H	1.97959	2.56289	1.58648
H	3.39776	0.23085	2.07238
H	1.67614	0.19002	1.68765
C	1.94622	-1.29403	-0.3641
O	1.64757	-2.17768	0.40707
O	1.43495	-1.22009	-1.58749
H	1.7884	-0.39314	-1.9761

C	2.94636	-0.19477	-0.02559
H	3.93885	-0.65185	-0.04747

Cartesian coordinates for D2(Oc)3(vb) calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	3.6566	-0.22896	1.19953
O	1.15874	1.88336	-0.20329
O	0.9538	-0.23207	0.53398
C	3.74394	0.35375	-1.06834
C	3.15998	0.73687	0.30044
C	3.72703	-1.17739	-0.99094
C	3.94411	-1.45266	0.50374
C	1.6355	0.72606	0.24537
H	4.765	0.73177	-1.13335
H	3.17467	0.757	-1.90428
H	3.48575	1.72234	0.63364
H	2.75808	-1.56094	-1.31015
H	4.49456	-1.63618	-1.61302
H	3.28485	-2.23822	0.8759
H	4.97861	-1.72391	0.72827
H	0.16963	1.85418	-0.27293
O	-3.45882	-0.69181	0.8083
C	-3.77156	-1.92645	0.12917
C	-2.48258	-2.34495	-0.56568
C	-1.85386	-0.99734	-0.92324
H	-4.1229	-2.63747	0.8758

H	-4.57504	-1.74448	-0.59175
H	-1.83696	-2.88919	0.12565
H	-2.66447	-2.97513	-1.43551
H	-0.77619	-1.03276	-1.0502
H	-2.30588	-0.58765	-1.8293
C	-2.43105	1.31485	-0.04271
O	-1.53436	2.04692	-0.39525
O	-3.6852	1.7427	0.04911
H	-4.20182	0.98302	0.3904
C	-2.22877	-0.14951	0.29852
H	-1.44802	-0.21709	1.05757

Cartesian coordinates for $D2^{+}$(Oe)3(Oc) calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	3.67886	0.80651	-0.78476
O	0.79228	-1.24152	-1.07429
O	1.04367	0.65148	0.10677
C	3.64434	-1.26346	0.31407
C	3.00408	-0.42778	-0.80683
C	3.93392	-0.19118	1.37131
C	4.19848	1.06681	0.5322
C	1.51408	-0.25814	-0.54613
H	4.56771	-1.7031	-0.06449
H	3.00329	-2.06837	0.67099
H	3.13009	-0.87885	-1.79132
H	3.06399	-0.04414	2.01062

H	4.78038	-0.45151	2.00528
H	3.70067	1.94575	0.94263
H	5.26455	1.27924	0.42618
H	-0.15542	-1.1167	-0.80646
O	-1.67917	-0.83621	-0.16701
C	-2.198	-1.60686	0.93045
C	-3.63435	-1.89434	0.52212
C	-4.05119	-0.60238	-0.20445
H	-1.56906	-2.48845	1.03712
H	-2.14508	-1.01294	1.8489
H	-3.66019	-2.74828	-0.15567
H	-4.27258	-2.11732	1.37606
H	-4.66884	-0.80595	-1.07604
H	-4.61022	0.07077	0.44093
C	-2.5679	1.46941	-0.01287
O	-3.49684	2.21115	-0.01325
O	-1.38594	1.82709	0.46294
H	-0.66035	1.17313	0.32915
C	-2.71325	0.05898	-0.60366
H	-2.60736	0.16023	-1.68497

Cartesian coordinates for D12-I calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	3.2842	0.85734	0.28221
C	3.83833	-0.45511	0.51724
C	2.80577	-1.43133	-0.02261

C	2.24369	-0.67449	-1.22946
H	4.79159	-0.53252	-0.01515
H	4.0243	-0.54682	1.58657
H	3.24083	-2.39114	-0.29716
H	2.01865	-1.60725	0.70858
H	2.87723	-0.81648	-2.10563
H	1.23207	-0.97628	-1.48611
C	0.98723	1.34967	-0.27668
O	0.00207	1.41174	-0.96525
O	1.00779	1.78919	0.98523
H	1.91346	1.6107	1.313
C	2.31854	0.79299	-0.77811
H	2.63643	1.44918	-1.59223
O	-2.6621	0.11776	1.15849
C	-3.02985	1.33774	0.47447
C	-3.6224	0.88085	-0.85044
C	-2.76632	-0.34412	-1.18459
H	-3.73238	1.86805	1.11592
H	-2.1366	1.94461	0.31788
H	-4.67006	0.59985	-0.72303
H	-3.56192	1.65265	-1.61674
H	-3.26513	-1.06166	-1.83182
H	-1.83536	-0.03703	-1.65677
C	-1.08273	-1.5225	0.32815
O	-0.67816	-2.42016	-0.36106
O	-0.3271	-0.95108	1.27701

H	-0.86806	-0.23047	1.66111
C	-2.48379	-0.93788	0.20501
H	-3.18639	-1.74133	0.4444

Cartesian coordinates for D11-II calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	3.40817	-0.51388	-0.23688
C	3.75989	0.88392	-0.29957
C	2.54082	1.62807	0.22
C	2.01748	0.67768	1.30114
H	4.64046	1.0533	0.32857
H	4.01623	1.11167	-1.3333
H	2.78711	2.61298	0.61385
H	1.79658	1.74969	-0.56463
H	2.53383	0.8434	2.24707
H	0.94976	0.77783	1.47508
C	1.21026	-1.45055	0.11657
O	0.22381	-1.75888	0.73251
O	1.3696	-1.74693	-1.17638
H	2.23966	-1.37645	-1.43012
C	2.38665	-0.71161	0.75146
H	2.76966	-1.36546	1.53901
O	-2.67524	-0.48405	-1.06271
C	-2.9753	-1.49383	-0.06735
C	-2.85938	-0.79562	1.28166
C	-3.25893	0.64414	0.94657

H	-3.98828	-1.86091	-0.25463
H	-2.26108	-2.30524	-0.18862
H	-3.50156	-1.24759	2.0366
H	-1.8289	-0.84186	1.63175
H	-4.34422	0.75166	0.9213
H	-2.85479	1.38758	1.62979
C	-1.28156	1.40733	-0.47877
O	-0.97977	2.42254	0.08967
O	-0.40378	0.70975	-1.21072
H	-0.86356	-0.10672	-1.49694
C	-2.69409	0.81505	-0.46396
H	-3.31112	1.47661	-1.07781

Cartesian coordinates for D22-III calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	1.35206	-0.4396	0.87461
C	1.72481	-1.83182	0.83486
C	3.2484	-1.84983	0.87038
C	3.59886	-0.55565	0.13465
H	1.2635	-2.32851	1.68897
H	1.33179	-2.27263	-0.08402
H	3.60977	-1.81133	1.90011
H	3.66528	-2.73671	0.39435
H	4.58714	-0.16463	0.36524
H	3.53118	-0.70093	-0.94558
C	2.18076	1.4671	-0.42099

O	2.8127	2.49049	-0.45163
O	1.22291	1.20486	-1.31291
H	0.76729	0.34843	-1.15616
C	2.48408	0.38832	0.61103
H	2.77931	0.90607	1.52719
O	-3.61137	-0.09506	-0.09736
C	-3.78536	0.50884	1.20137
C	-2.82511	1.68704	1.21815
C	-1.63434	1.14406	0.4249
H	-4.83484	0.78245	1.29655
H	-3.53526	-0.22059	1.97923
H	-3.27087	2.54356	0.70911
H	-2.5535	1.99147	2.22814
H	-1.03148	1.91751	-0.043
H	-0.97443	0.55006	1.05745
C	-1.54155	-1.01741	-0.94105
O	-0.43777	-1.01336	-1.42855
O	-2.17473	-2.14813	-0.63715
H	-3.05223	-1.88091	-0.29343
C	-2.32068	0.25276	-0.62456
H	-2.46477	0.78654	-1.56727

Cartesian coordinates for D12-IV calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	1.13096	-0.69151	0.62243
C	1.31934	-1.96549	-0.02879

C	2.69794	-1.89039	-0.66808
C	3.46675	-1.03766	0.34427
H	1.26659	-2.75728	0.72431
H	0.50608	-2.10676	-0.73865
H	3.14053	-2.87318	-0.82385
H	2.64387	-1.38296	-1.63358
H	3.7952	-1.64903	1.18583
H	4.33259	-0.52509	-0.06709
C	2.45632	1.27877	0.04195
O	3.45434	1.9473	0.02056
O	1.33541	1.65006	-0.58286
H	0.62319	0.98825	-0.44952
C	2.39442	-0.04526	0.81087
H	2.50905	0.20819	1.86745
O	-1.29413	0.74797	-0.6268
C	-1.9195	2.04581	-0.782
C	-2.68098	2.27343	0.51501
C	-3.13285	0.85675	0.87647
H	-1.12905	2.76907	-0.97215
H	-2.59197	2.01122	-1.64377
H	-2.01111	2.66957	1.28042
H	-3.51119	2.96788	0.39326
H	-3.35874	0.72027	1.93098
H	-4.0183	0.57655	0.30256
C	-2.30597	-1.38022	-0.02991
O	-2.75562	-2.21163	0.70646

O	-2.10194	-1.60165	-1.34015
H	-1.71154	-0.78319	-1.70116
C	-1.92721	0.01548	0.44404
H	-1.20426	-0.10326	1.25144

Cartesian coordinates for D11-V calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	2.98015	-0.3801	-0.97541
C	4.32	-0.07397	-0.53985
C	4.1859	0.44486	0.88675
C	2.98129	-0.34881	1.40068
H	4.91524	-0.99106	-0.57913
H	4.7472	0.65001	-1.23377
H	5.08877	0.28599	1.47488
H	3.96388	1.51422	0.88571
H	3.2813	-1.35342	1.70135
H	2.46178	0.12063	2.23246
C	1.09964	0.73177	0.07142
O	0.34538	1.02077	0.96325
O	1.11525	1.38925	-1.09394
H	1.80897	0.95982	-1.63387
C	2.09588	-0.42516	0.15823
H	1.5036	-1.33881	0.10225
O	-3.56806	-0.48485	0.55864
C	-3.87106	0.89912	0.83046
C	-2.79865	1.69829	0.10696

C	-2.54866	0.84879	-1.14342
H	-4.87305	1.11978	0.44812
H	-3.87042	1.03045	1.91184
H	-3.12589	2.70956	-0.13248
H	-1.88946	1.75755	0.70362
H	-3.27684	1.07848	-1.92225
H	-1.55152	0.97555	-1.55733
C	-1.48157	-1.33371	-0.32282
O	-0.62811	-1.53022	-1.14811
O	-1.38254	-1.77888	0.93204
H	-2.17983	-1.45446	1.39691
C	-2.77616	-0.58318	-0.63469
H	-3.31101	-1.19092	-1.36902

Cartesian coordinates for D11-VI calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	3.46542	0.091	0.34962
C	3.20512	-0.71069	1.52683
C	1.85802	-1.37728	1.28242
C	1.87196	-1.57106	-0.23588
H	4.01042	-1.44346	1.62603
H	3.2217	-0.05011	2.39318
H	1.75322	-2.3132	1.82929
H	1.04054	-0.71254	1.56476
H	2.47071	-2.44099	-0.50904
H	0.88739	-1.68502	-0.67982

C	1.62134	0.85456	-1.01649
O	0.70291	0.76732	-1.79146
O	1.90642	1.98659	-0.37327
H	2.67537	1.78219	0.19841
C	2.57886	-0.30083	-0.70979
H	3.16619	-0.46698	-1.61573
O	-0.95509	0.51739	0.67911
C	-1.3768	1.89601	0.58927
C	-2.86568	1.84357	0.27725
C	-3.30182	0.5789	1.02092
H	-1.18426	2.38314	1.55033
H	-0.77651	2.37707	-0.17938
H	-3.39324	2.73723	0.60807
H	-3.02531	1.73111	-0.79685
H	-3.43718	0.78719	2.08343
H	-4.21536	0.12773	0.64122
C	-2.26764	-1.25855	-0.38601
O	-3.08386	-2.14133	-0.37529
O	-1.48787	-1.02814	-1.4448
H	-0.86665	-0.28764	-1.29008
C	-2.08696	-0.3402	0.82662
H	-1.93097	-0.9953	1.68571

Cartesian coordinates for D12-VII calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	2.91413	-0.75347	0.85735

C	2.62706	-2.03093	0.25286
C	1.71812	-1.71949	-0.92535
C	2.26624	-0.37471	-1.41096
H	3.56733	-2.48788	-0.07063
H	2.17185	-2.6629	1.01391
H	1.74935	-2.49198	-1.69246
H	0.68603	-1.60871	-0.59521
H	3.11833	-0.52153	-2.07509
H	1.52618	0.23066	-1.92709
C	1.7347	1.31565	0.43445
O	1.4295	2.312	-0.16212
O	1.23585	1.01299	1.641
H	1.67069	0.17703	1.90736
C	2.73517	0.29413	-0.10803
H	3.68013	0.8247	-0.24317
O	-1.06622	0.30329	-0.86935
C	-1.43788	1.63321	-1.30606
C	-2.04651	2.30015	-0.08193
C	-2.7191	1.12415	0.62903
H	-0.53747	2.12747	-1.66254
H	-2.1606	1.54366	-2.12304
H	-1.25386	2.72261	0.53582
H	-2.74584	3.09281	-0.34568
H	-2.87889	1.28379	1.69243
H	-3.68289	0.89107	0.17137
C	-2.37547	-1.38089	0.28577

O	-2.8733	-1.92461	1.23161
O	-2.35724	-1.92616	-0.94234
H	-1.87035	-1.29542	-1.5082
C	-1.71995	-0.00994	0.37282
H	-0.97204	-0.04523	1.16658

Cartesian coordinates for D12-VIII calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	-2.7454	-1.11893	-0.69889
C	-3.86898	-0.40719	-1.26006
C	-3.74678	1.01678	-0.73892
C	-2.22838	1.20904	-0.69701
H	-3.79662	-0.44354	-2.35091
H	-4.77835	-0.92083	-0.95055
H	-4.24946	1.74094	-1.37826
H	-4.17231	1.09427	0.26393
H	-1.8409	1.45892	-1.68449
H	-1.89871	1.98074	-0.00663
C	-1.50599	-0.32927	1.21027
O	-0.72084	0.35188	1.83226
O	-2.22895	-1.27343	1.79654
H	-2.76701	-1.6744	1.08114
C	-1.73033	-0.1847	-0.29228
H	-0.78204	-0.44299	-0.76499
O	1.44329	-0.85624	-0.33981
C	2.31206	-1.54224	-1.24929

C	3.57495	-1.80133	-0.44304
C	3.70144	-0.50503	0.37121
H	1.79255	-2.44266	-1.57523
H	2.52011	-0.9117	-2.12108
H	3.42783	-2.66256	0.21103
H	4.44572	-1.99008	-1.07028
H	4.16716	-0.66135	1.34255
H	4.27773	0.24093	-0.17345
C	2.05046	1.43366	0.0344
O	2.82418	2.01226	-0.68259
O	0.9078	2.00391	0.42878
H	0.36521	1.40106	0.99146
C	2.25425	-0.01548	0.51025
H	1.86408	-0.12875	1.5226

Cartesian coordinates for D12-IX calculated at the B3LYP-D3BJ/def2-TZVP/PCM(Ar) level of theory.

O	3.45828	-0.24311	0.86472
C	4.41392	0.65535	0.26687
C	3.73592	1.19652	-0.98373
C	2.27341	1.2769	-0.53886
H	4.63729	1.45566	0.97902
H	5.32715	0.09333	0.07281
H	4.141	2.15858	-1.29503
H	3.84959	0.49402	-1.81245
H	2.08579	2.1788	0.04193

H	1.55461	1.25206	-1.35313
C	1.59427	-1.17509	-0.36889
O	0.52677	-1.18733	-0.9234
O	2.4	-2.24235	-0.33714
H	3.18852	-1.96166	0.17062
C	2.13781	0.0501	0.36839
H	1.47115	0.23905	1.20981
O	-2.93134	-1.03347	0.11516
C	-4.32945	-0.78415	0.32257
C	-4.68772	0.27647	-0.70697
C	-3.43689	1.16147	-0.6978
H	-4.85351	-1.7301	0.19546
H	-4.49917	-0.41813	1.34265
H	-4.8265	-0.18716	-1.68525
H	-5.59568	0.8226	-0.45262
H	-3.27014	1.68345	-1.63688
H	-3.49723	1.91267	0.0904
C	-1.29085	0.69204	0.60319
O	-0.70963	1.73424	0.45921
O	-1.08773	-0.10038	1.66566
H	-1.65622	-0.88442	1.53155
C	-2.30845	0.15868	-0.39643
H	-1.74464	-0.12006	-1.28674

Appendix B

Supporting Information for Chapter 4

Table B1. The relative free energies, ΔG, in kJ mol^{-1}, and Boltzmann percentage factors, B_f in %, of the conformers of the THFA dimers at 298 K at the B3LYP-D3BJ/def2-TZVP/PCM(water) levels of theory.

THFA dimers	ΔG	B_f
DIV-IV	0.0	20.3
DIV-VI	0.6	15.9
DIII-III	1.1	13.1
DIII-IV	1.3	12.0
DIII-V	1.8	9.8
DV-V	2.6	7.0
DIII-VI	2.7	6.8
DVI-VI	2.7	6.7
DIV-V	3.2	5.6
DV-VI	4.8	2.9

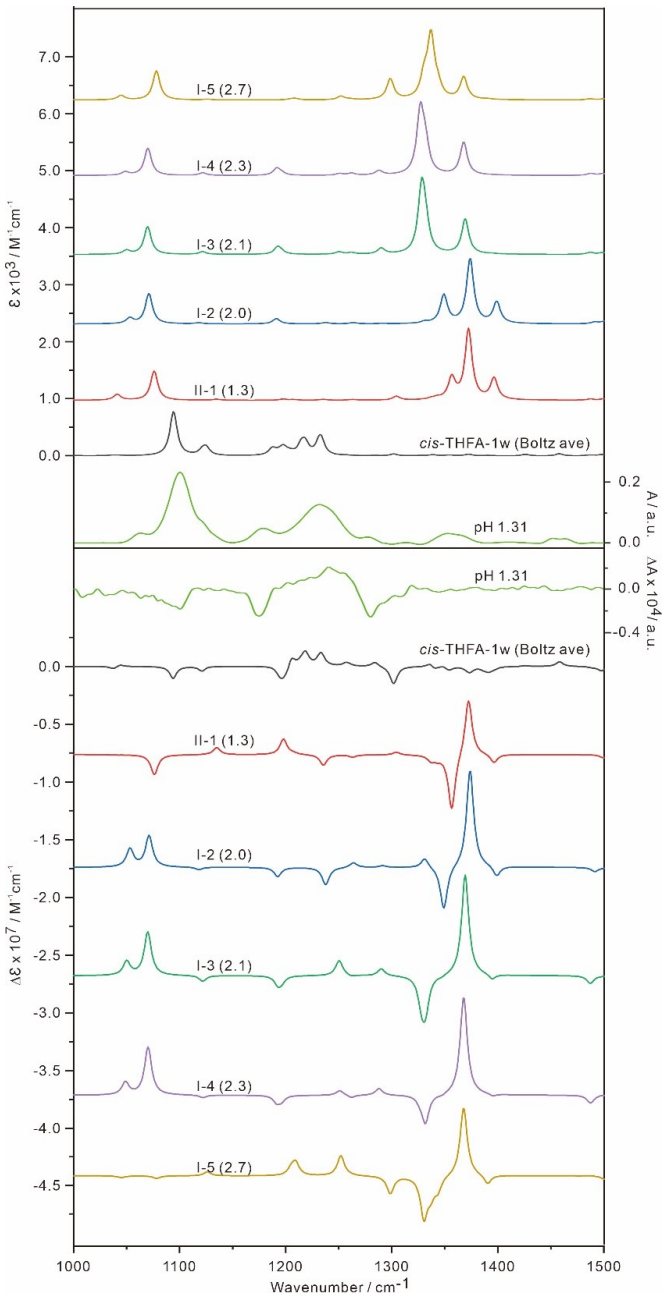


Figure B1. Comparisons of experimental IR and VCD spectra of THFA at acidic condition (pH 1.31) with the simulated IR and VCD Boltzmann averaged spectra of neutral *cis*-THFA monohydrates under B3LYP-D3BJ/def2-TZVP/PCM (water) level of theory. The most five stable *trans*-THFA monohydrates are also listed for comparison with the relative free energies with respect to the global minimum of *cis*-THFA monohydrates in the bracket.

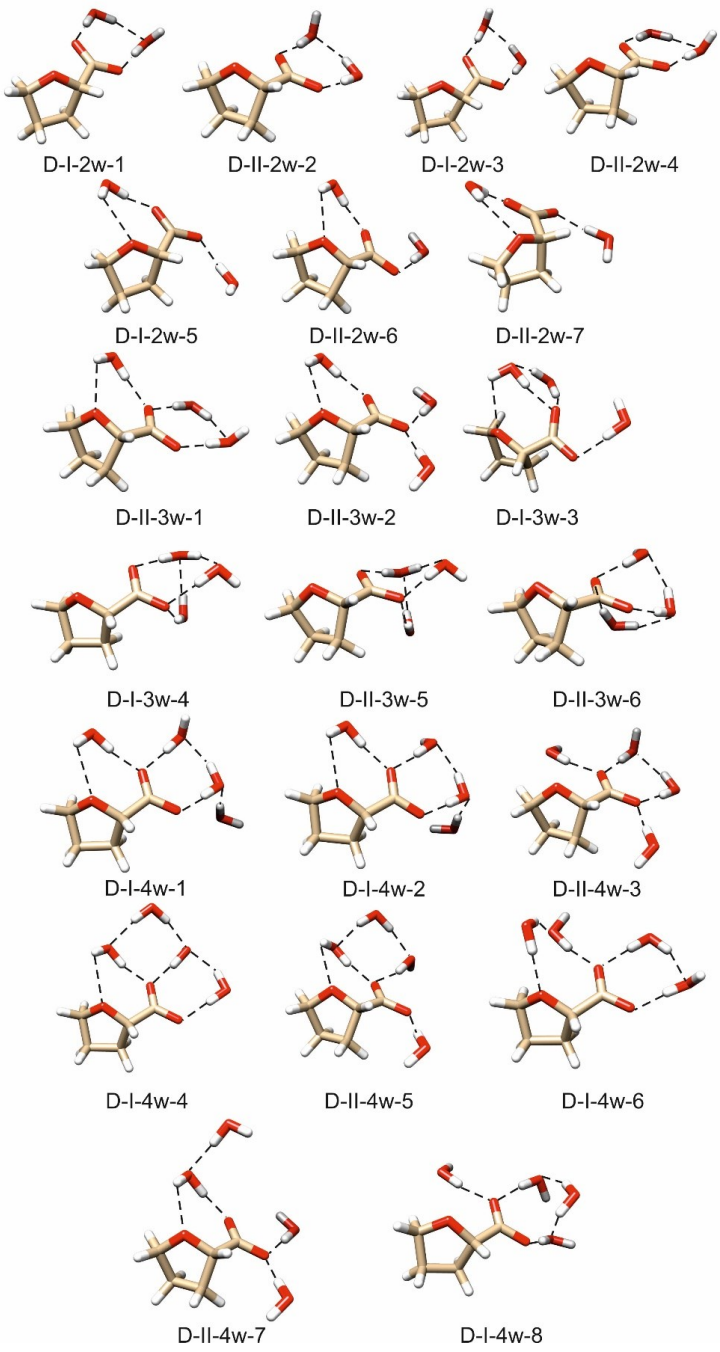


Figure B2. Optimized geimetries of seven deprotonated THFA-2w conformers, six deprotonated THFA-3w conformers and eight deprotonated THFA-4w conformers at the B3LYP-D3BJ/def2-TZVP/PCM (water) level of theory.

Table B2. The relative free energies, ΔG, in kJ mol^{-1}, and Boltzmann percentage factors, B_f in %, of the conformers of the deprotonated THFA-2w, deprotonated THFA-3w and deprotonated THFA-4w at 298 K at the B3LYP-D3BJ/def2-TZVP/PCM(water) levels of theory.

Conformer of deprotonated THFA-nW	**ΔG**	**B_f**
D-I-2w-1	0.0	39.0
D-II-2w-2	0.7	28.8
D-I-2w-3	3.4	10.1
D-II-2w-4	3.8	8.4
D-I-2w-5	4.8	5.6
D-II-2w-6	5.6	4.1
D-II-2w-7	5.6	4.0
D-II-3w-1	0.0	71.9
D-II-3w-2	4.7	11.0
D-I-3w-3	5.4	8.1
D-I-3w-4	7.4	3.7
D-II-3w-5	7.8	3.1
D-II-3w-6	8.6	2.2
D-I-4w-1	0.0	36.8
D-I-4w-2	1.5	20.3
D-II-4w-3	2.3	14.5
D-I-4w-4	2.8	11.7
D-II-4w-5	4.1	7.1
D-I-4w-6	5.0	4.8
D-II-4w-7	5.9	3.3

D-I-4w-8	7.9	1.5

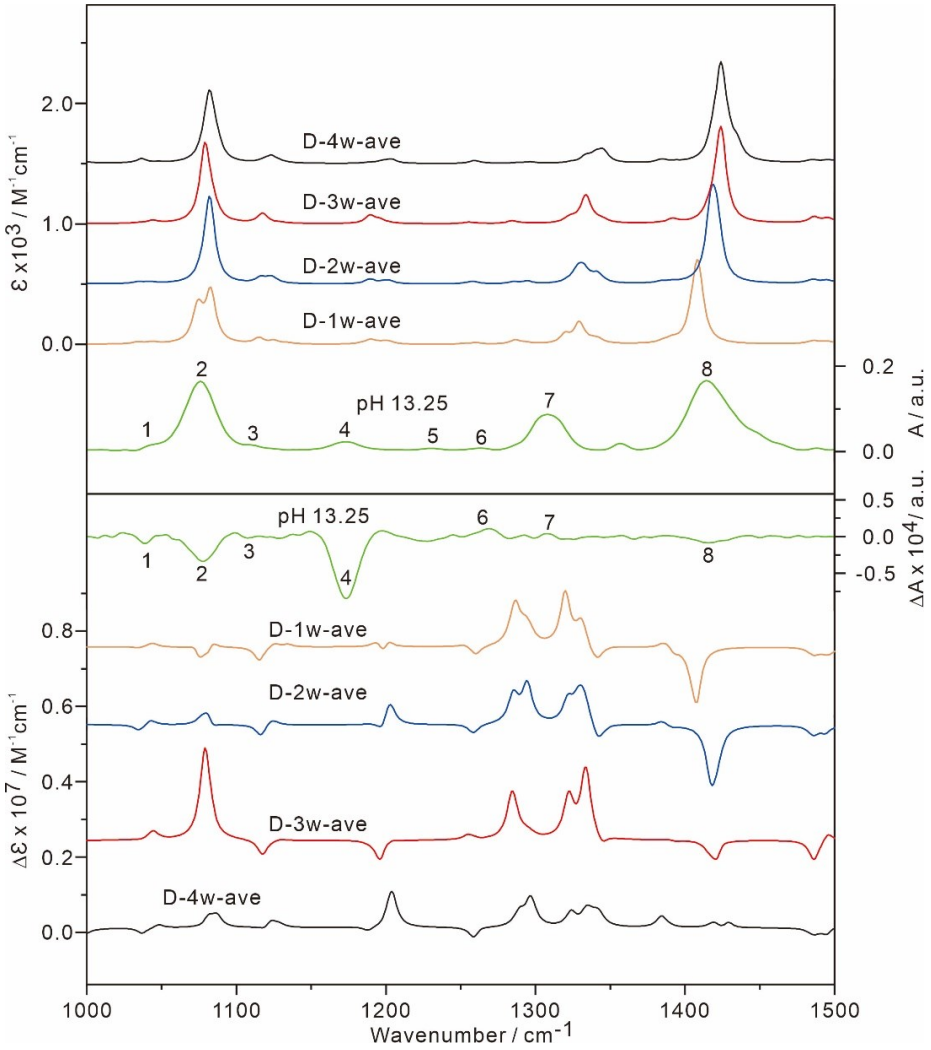


Figure B3. Comparisons of experimental IR and VCD spectra of THFA in water at pH 13.25 environment with the simulated Boltzmann averaged IR and VCD spectra of deprotonated THFA -1,2,3,4w clusters.

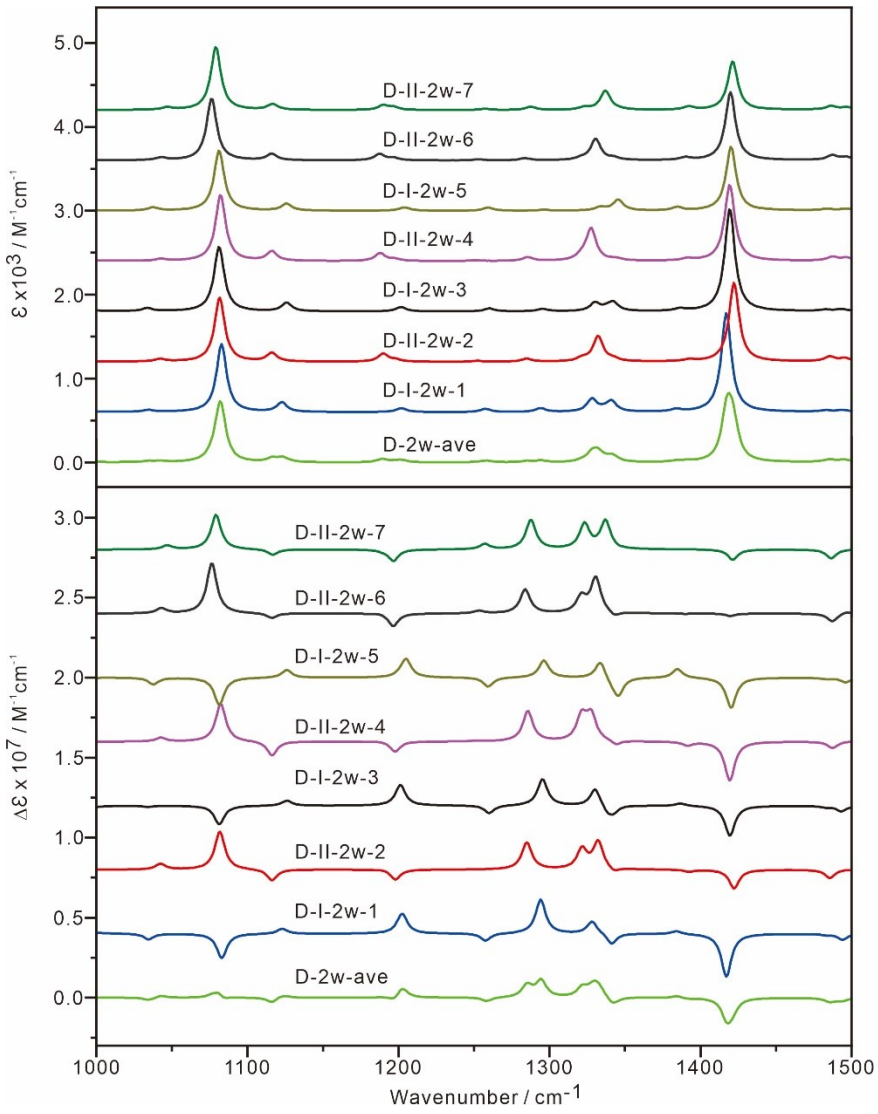


Figure B4. Theoretical IR and VCD spectra of the seven deprotonated THFA dihydrates conformers at the B3LYP-D3BJ/def2-TZVP/PCM(water) level of theory. The Boltzmann averaged spectra at 298 K are also provided.

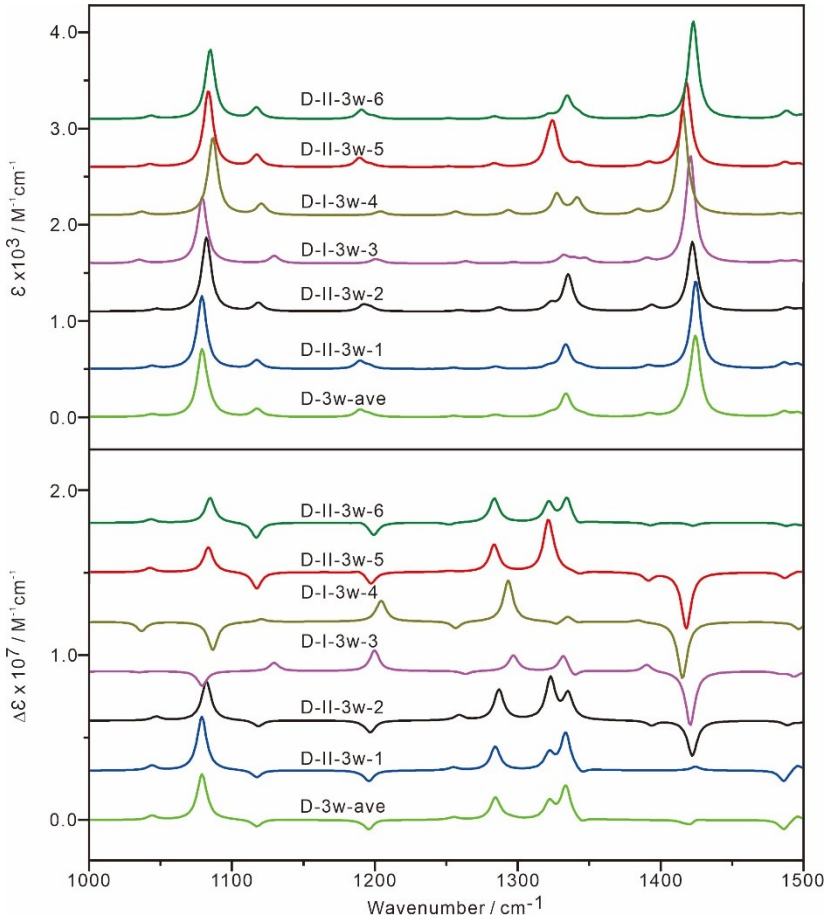


Figure B5. Theoretical IR and VCD spectra of the six deprotonated THFA trihydrates conformers at the B3LYP-D3BJ/def2-TZVP/PCM(water) level of theory. The Boltzmann averaged spectra at 298 K are also provided.

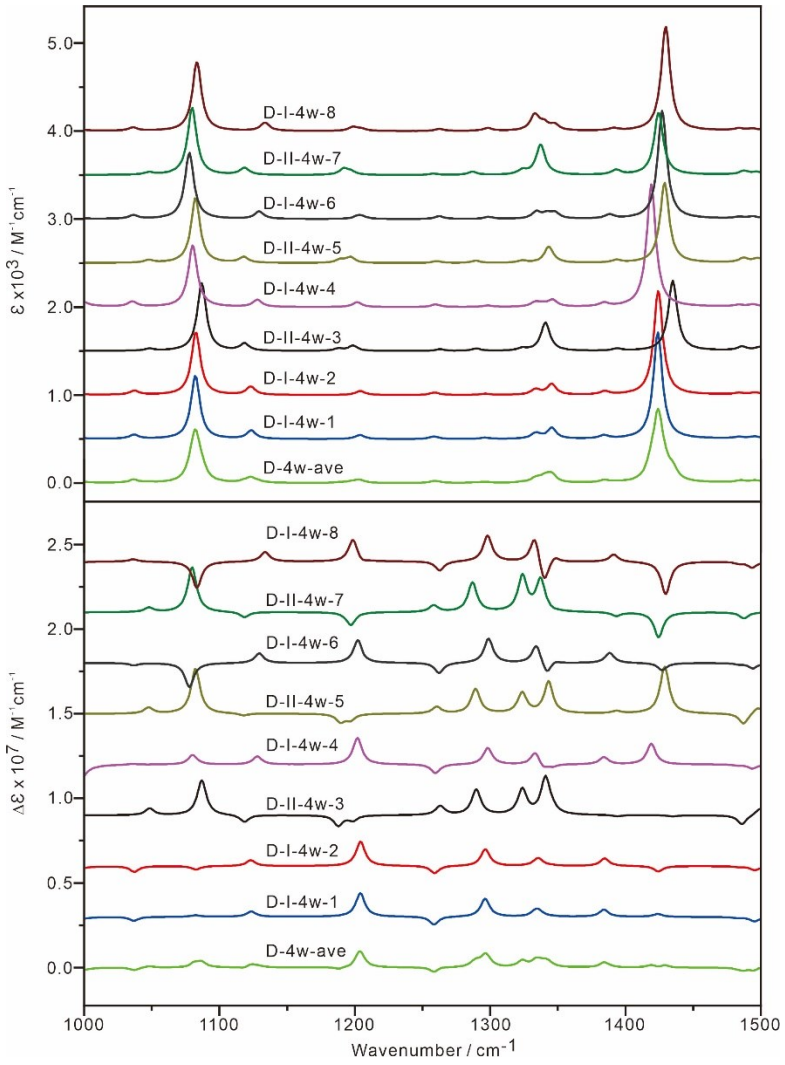


Figure B6. Theoretical IR and VCD spectra of the eight deprotonated THFA tetrahydrates conformers at the B3LYP-D3BJ/def2-TZVP/PCM(water) level of theory. The Boltzmann averaged spectra at 298 K are also provided.

Appendix C

Supporting Information for Chapter 5

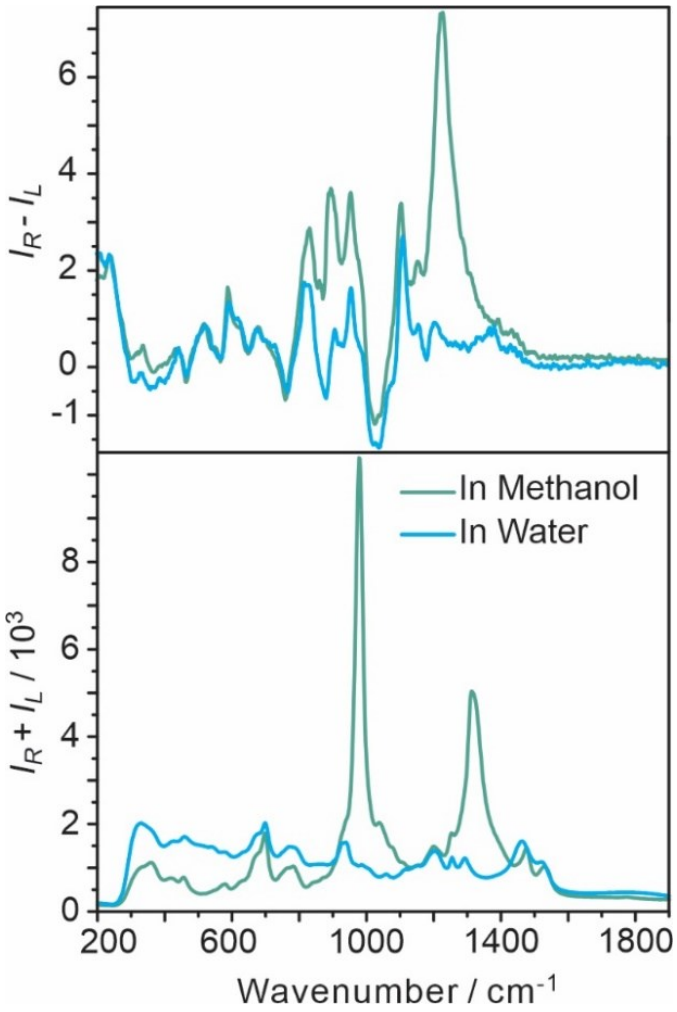


Figure C1. Raw experimental Raman (bottom) and ROA (top) spectra of the NALC molecule dissolved in water (blue) and methanol (green).

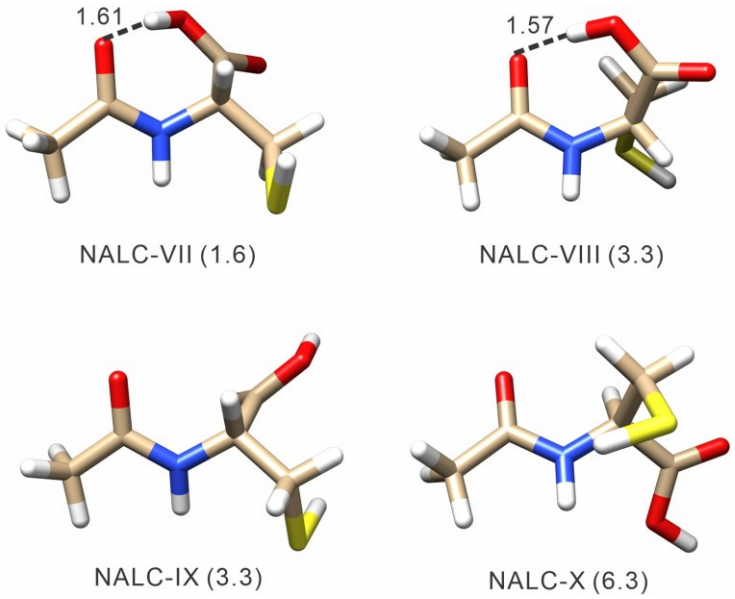


Figure C2. The four additional low energy NALC conformers optimized at the B3LYP-D3BJ/def2-TZVP level of theory with the PCM of water. The relative free energies (kJ mol^{-1}) to the global minimum are given in the brackets. As discussed in the main text, these conformers do not contribute much to the observed spectra.

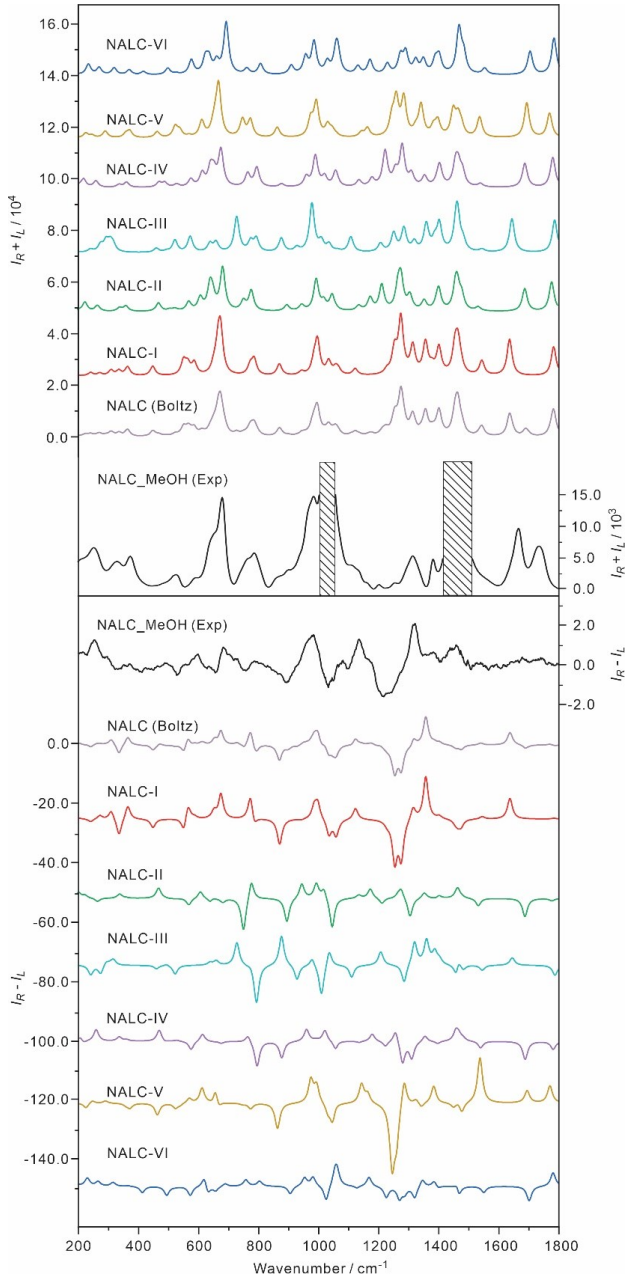


Figure C3. Individual Raman (top) and ROA (bottom) spectra of the six NALC conformers shown in Figure 5.2 at the B3LYP-D3BJ/def2-TZVP level with the PCM of methanol. The Boltzmann population weighted Raman and ROA spectra are also included, as well as the experimental data. The strong experimental Raman bands at 1020 cm^{-1} and 1460 cm^{-1} have severe solvent interference and are blocked out for clarity.

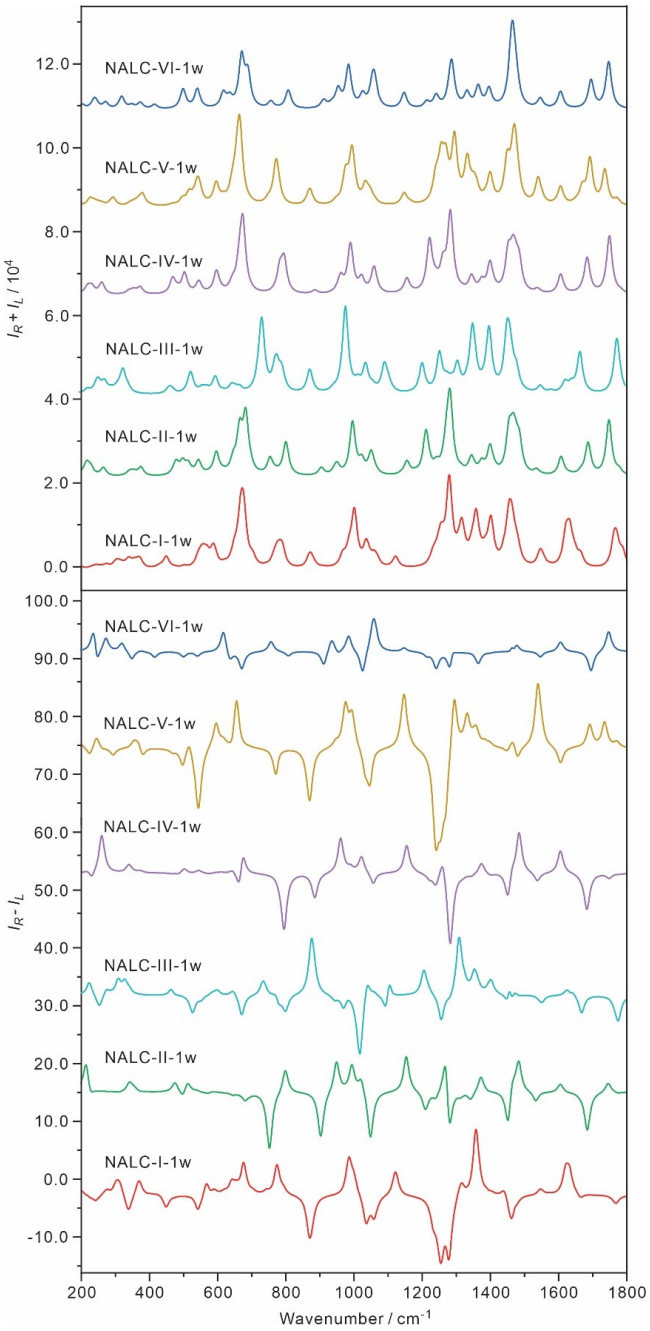


Figure C4. The Boltzmann averaged Raman and ROA spectra of each set of NALC monohydrates from NALC-I-1w to NALC-VI-1w at the B3LYP-D3BJ/def2-TZVP level with the PCM of water.

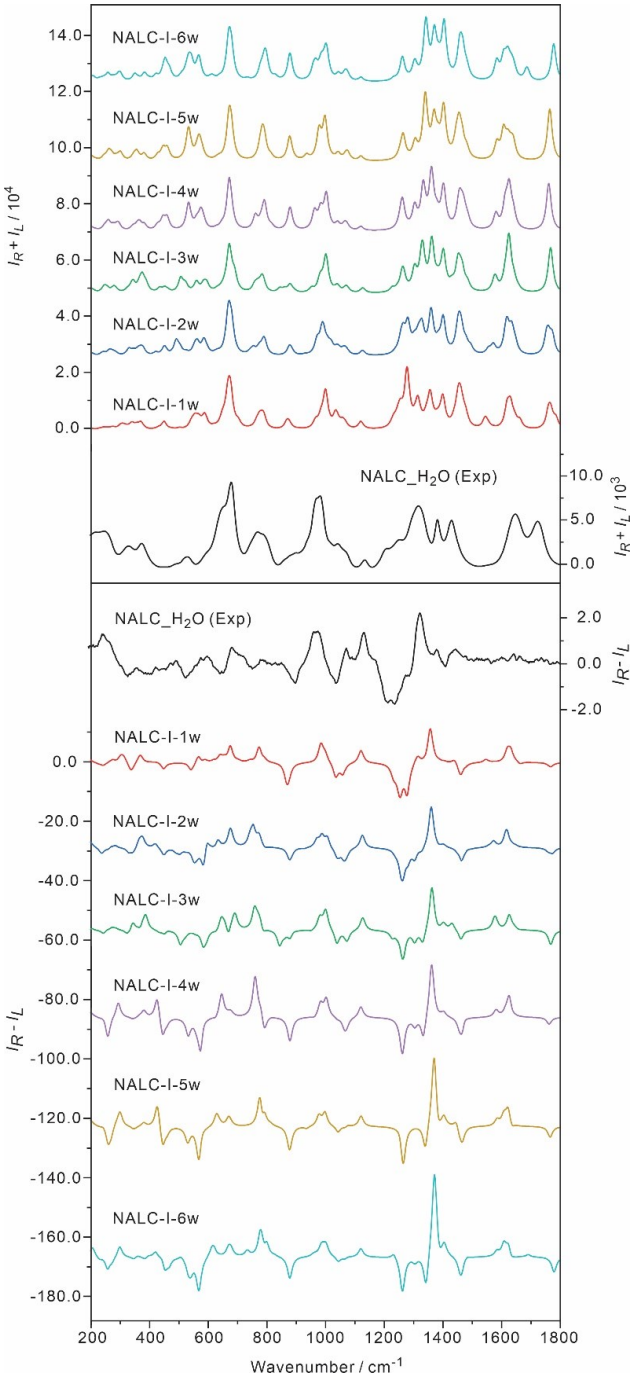


Figure C5. Comparison of the experimental Raman and ROA spectra of NALC in water with the Boltzmann weighted Raman and ROA spectra of conformers of NALC-I with 1 to 6 water molecules at the B3LYP-D3BJ/def2-TZVP level of theory with the PCM of water.

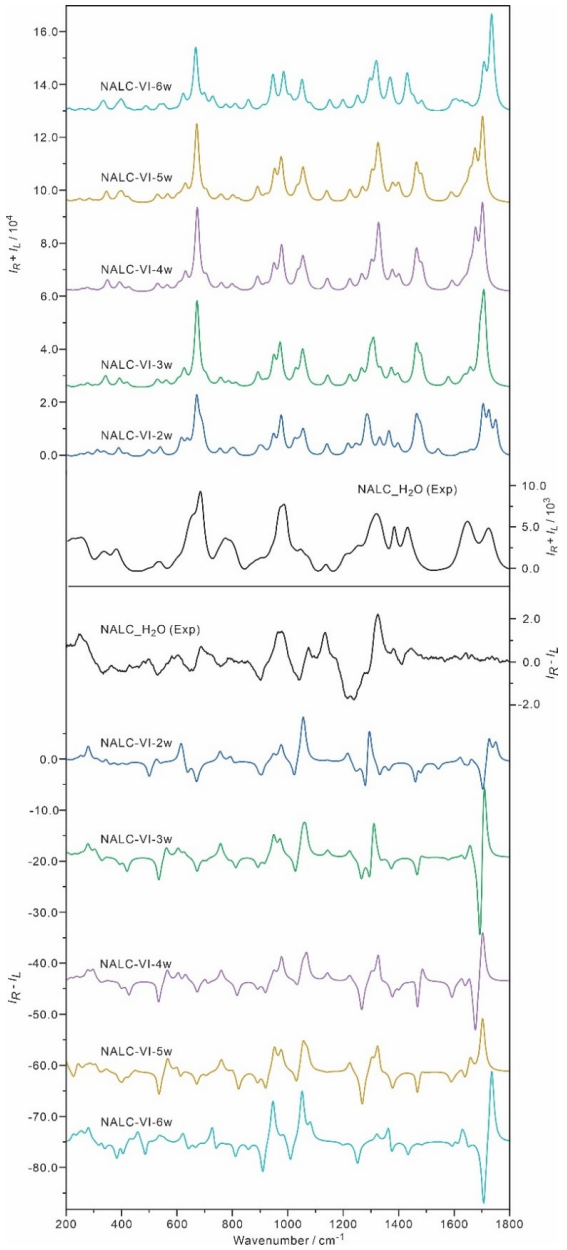


Figure C6. Comparison of the experimental Raman and ROA spectra of NALC in water with the Boltzmann weighted Raman and ROA spectra of conformers of NALC-VI with 1 to 6 water molecules. All the conformers are calculated at the B3LYP/cc-pVTZ/PCM (water) level for geometry optimization and the related frequencies calculations were performed using the same optimized geometries and with the B3LYP/aug-cc-pVDZ/PCM (water).

Table C1. The relative free energy, ΔG (kJ mol^{-1}) and the Boltzmann population factor, B_f (%) of the six NALC conformers at 298 K at the B3LYP-D3BJ/def2-TZVP level of theory with the PCM of methanol.

Conformer of NALC	**ΔG**	**B_f**
NALC-I	0.0	58.5
NALC-II	3.7	13.1
NALC-III*	4.7	8.8
NALC-IV*	4.6	9.0
NALC-V	5.6	6.0
NALC-VI	6.4	4.5

*The conformers -III and -IV are kept in the same order as those with the PCM of water for easy comparison since these two conformers are nearly iso-energy with the PCM of methanol.

Table C2. The relative free energy, ΔG (kJ mol^{-1}) and the Boltzmann population factor, B_f (%) of the 1:1 complexes of NALC-I to -VI with methanol at 298 K at the B3LYP-D3BJ/def2-TZVP level of theory with the PCM of methanol.

Conformer	ΔG	B_f	Conformer	ΔG	B_f
NALC-I-1m-1	0.0	30.6	NALC-II-1m-3	13.5	0.2
NALC-I-1m-2	0.5	25.4	NALC-III-1m-1	0.0	62.4
NALC-I-1m-3	1.5	16.8	NALC-III-1m-2	1.4	35.5
NALC-I-1m-4	3.5	7.4	NALC-III-1m-3	8.4	2.1
NALC-I-1m-5	3.6	7.3	NALC-IV-1m-1	0.0	35.9
NALC-I-1m-6	4.5	4.9	NALC-IV-1m-2	1.0	24.4
NALC-I-1m-7	6.0	2.7	NALC-IV-1m-3	1.4	20.3
NALC-I-1m-8	6.3	2.4	NALC-IV-1m-4	1.5	19.4
NALC-I-1m-9	7.3	1.6	NALC-V-1m-1	0.0	54.8
NALC-I-1m-10	8.6	0.9	NALC-V-1m-2	0.5	45.2
NALC-II-1m-1	0.0	62.6	NALC-VI-1m-1	0.0	61.2
NALC-II-1m-2	1.3	37.2	NALC-VI-1m-2	1.1	38.8

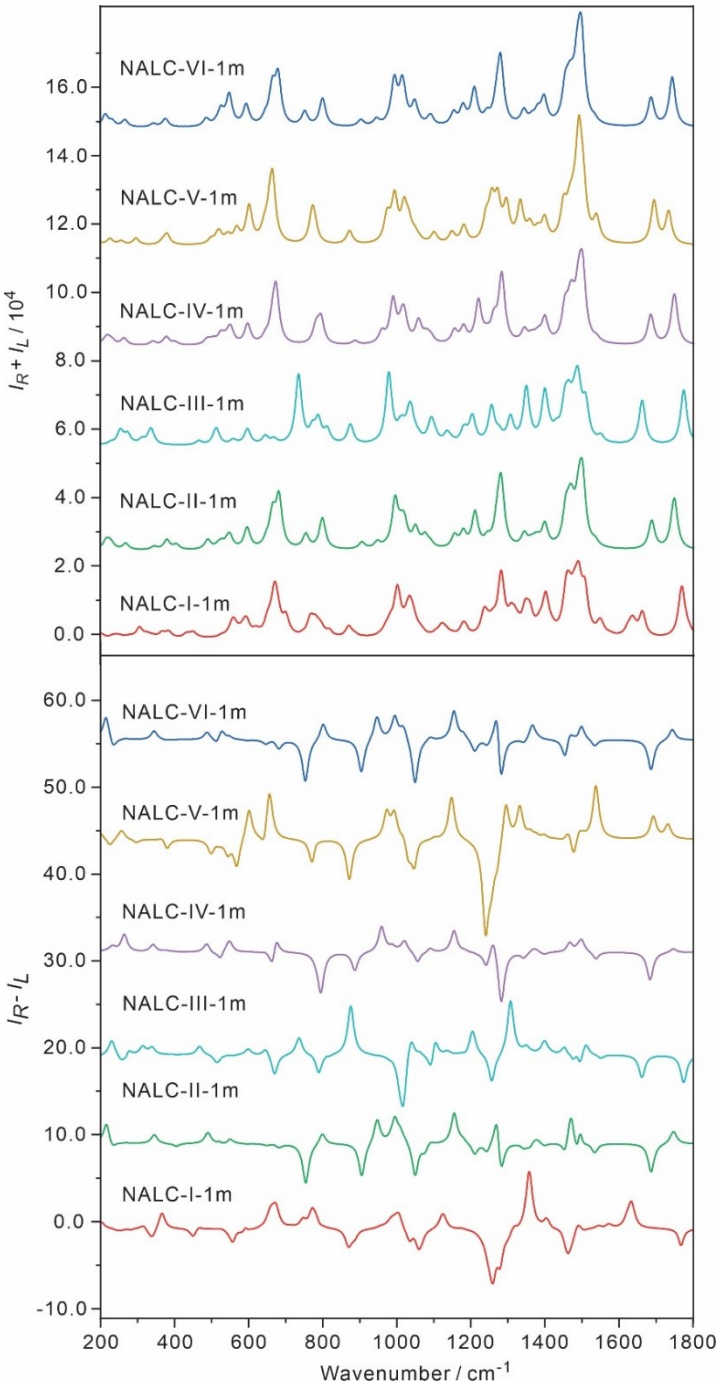


Figure C7. Boltzmann averaged Raman and ROA spectra of the conformers of NALC-I-1m to NALC-VI-1m at the B3LYP-D3BJ/def2-TZVP level of theory with the PCM of methanol.

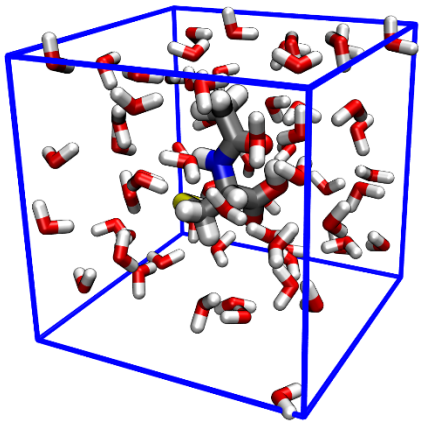


Figure C8. A snapshot of the AIMD simulation cell containing one NALC molecule in 64 molecules of water (1300 pm cubic cell size).

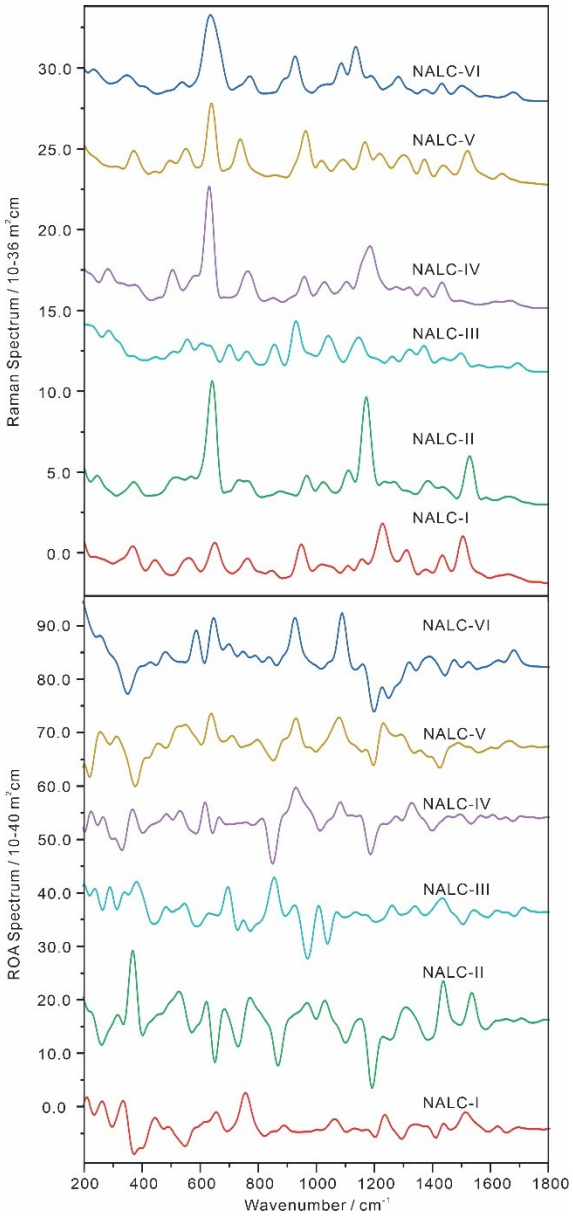


Figure C9. AIMD Raman and ROA spectra of the conformers of NALC in water.

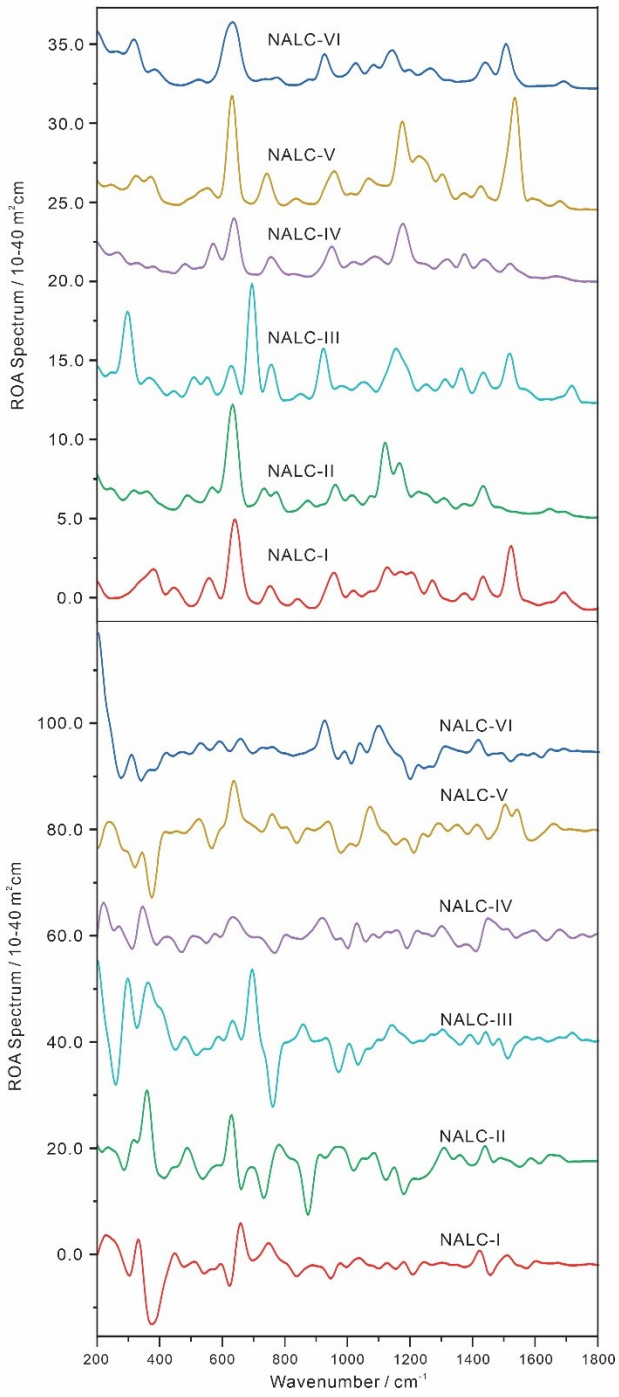


Figure C10. AIMD Raman and ROA spectra of the conformers of NALC in methanol.

Appendix D

Supporting Information for Chapter 6

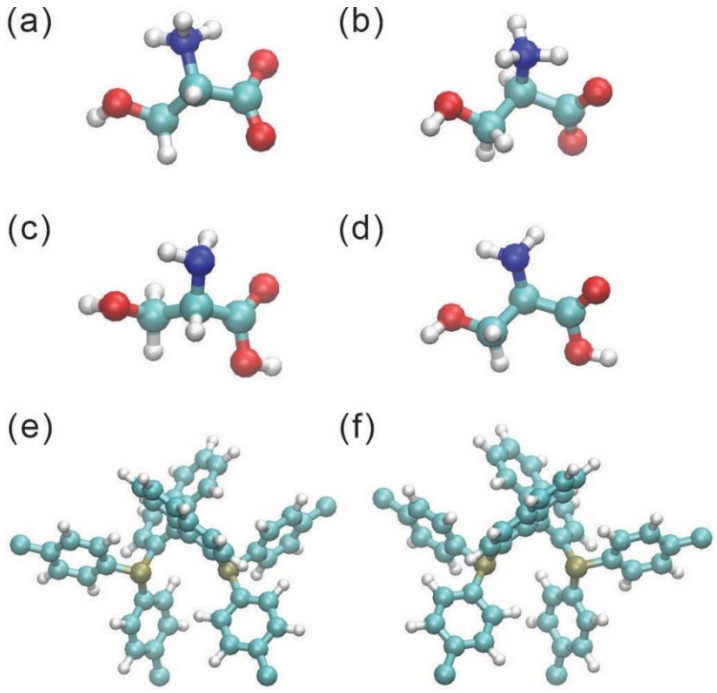


Figure D1. Optimized geometries of the zwitterionic S-Serine (a) and R-Serine (b) ions; neutral S-Serine (c) and R-Serine (d); neutral S-Tol-BINAP (e) and R-Tol-BINAP (f). (Cyan: carbon, red: oxygen, white: hydrogen, blue: nitrogen, and olive: phosphorus.)

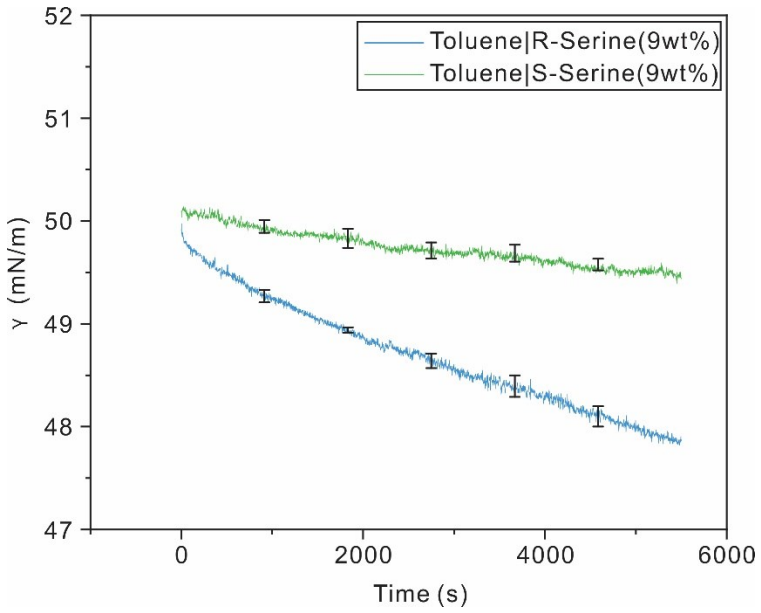


Figure D2. The IFT plots of the Toluene|S-Serine (9wt%) and Toluene|R-Serine (9wt%) systems.

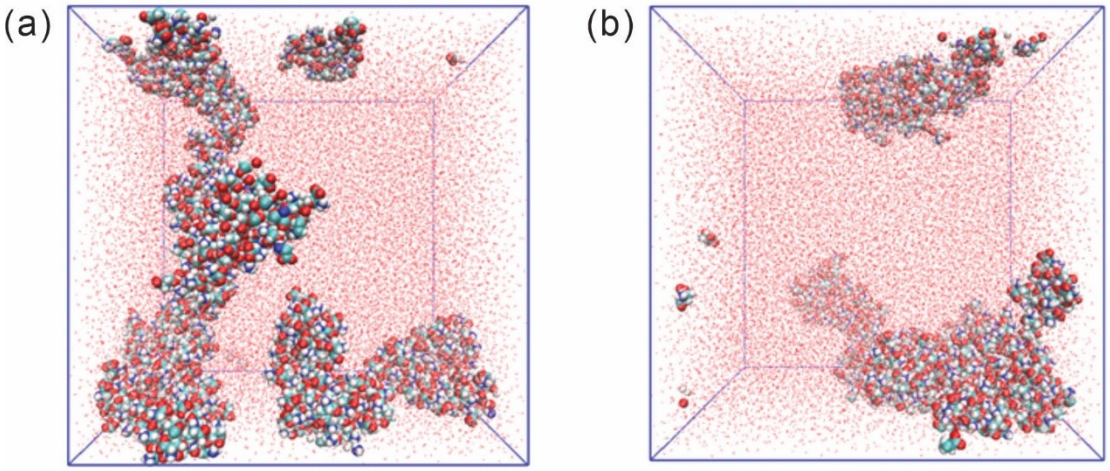


Figure D3. Final configurations for (a) S-Serine molecules in water, and (b) R-Serine molecules in water. (Red line: water; cyan sphere: carbon; red sphere: oxygen; blue sphere: nitrogen; white sphere: hydrogen.)

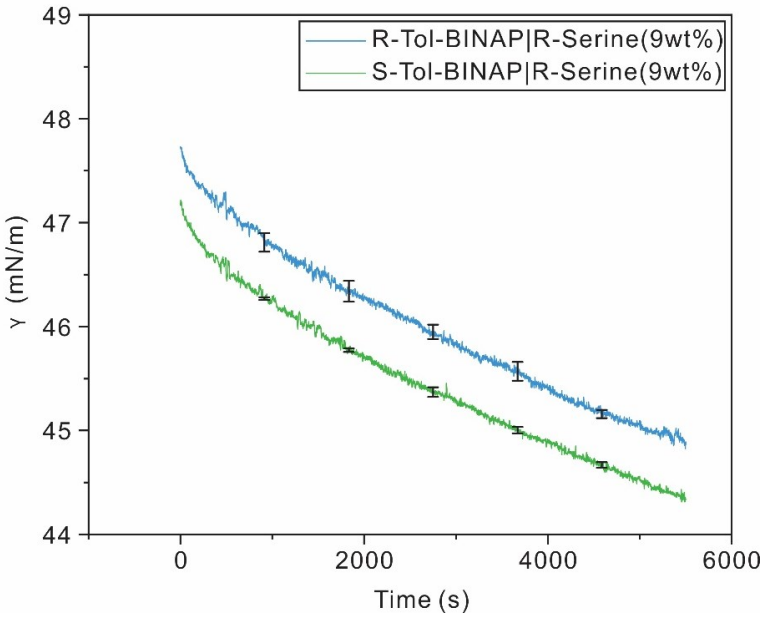


Figure D4. The IFT plot vs. time for S-Tol-BINAP|R-Serine (9wt%) and R-Tol-BINAP|R-Serine (9wt%).

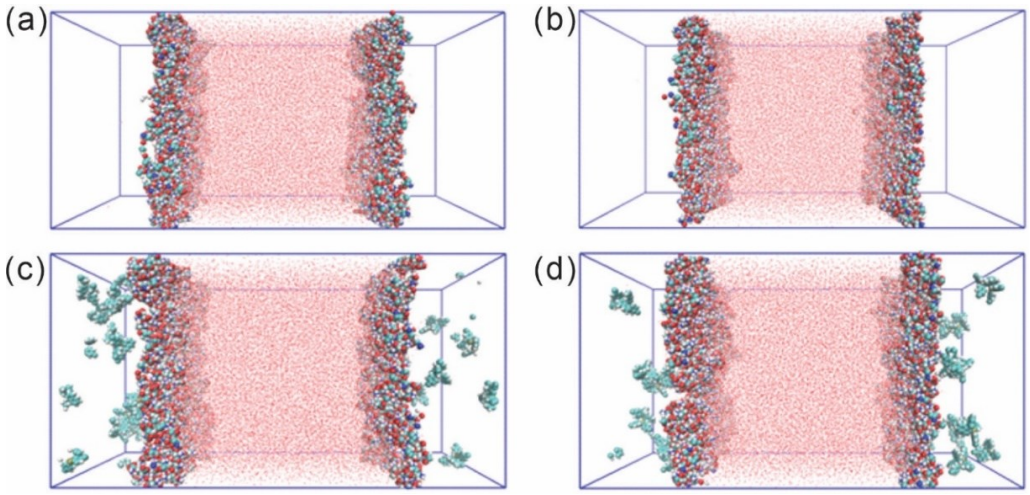


Figure D5. Final configurations of (a) Toluene|S-Serine, (b) Toluene|R-Serine, (c) S-Tol-BINAP|R-Serine, and (d) R-Tol-BINAP|R-Serine. (Red line: water; cyan sphere: carbon; red sphere: oxygen; blue sphere: nitrogen; white sphere: hydrogen; toluene molecules are removed for clarity.)

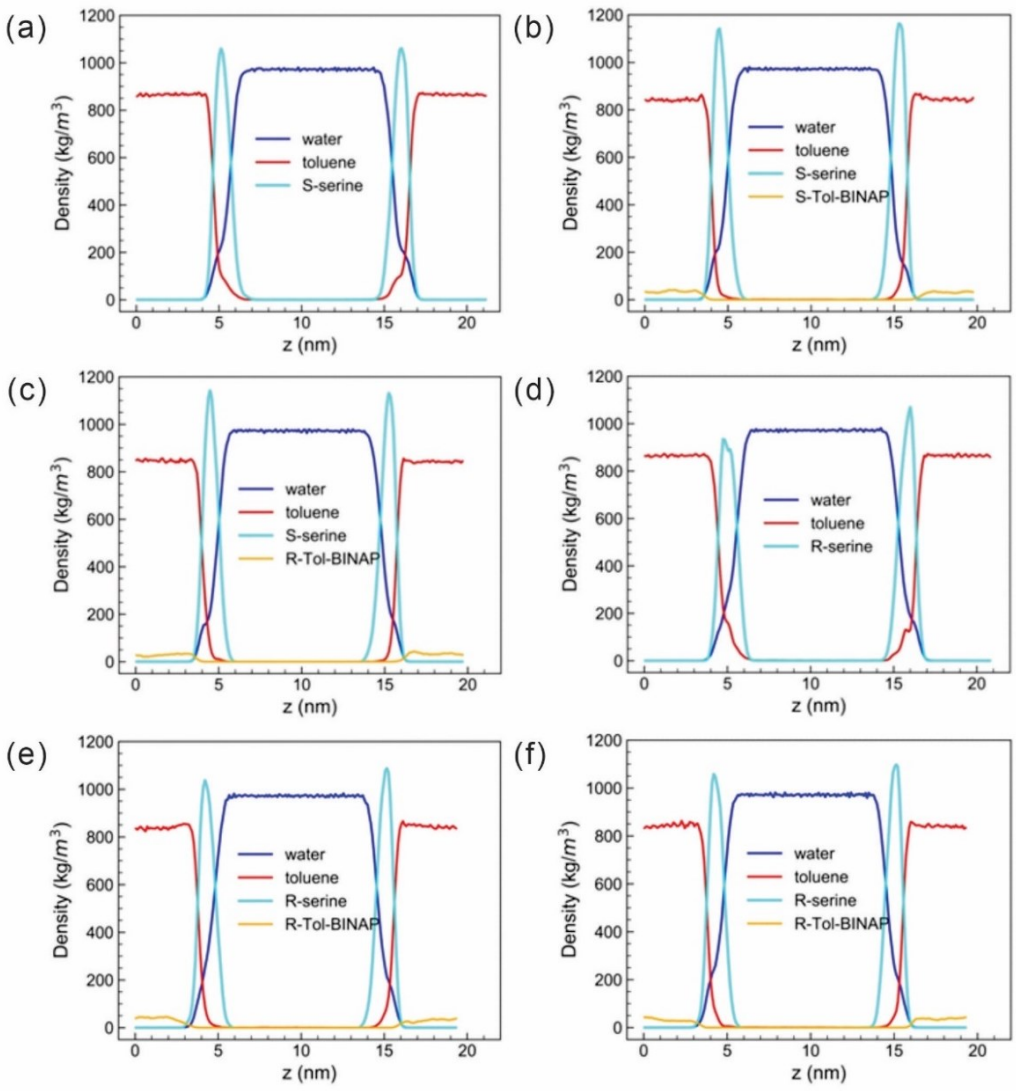


Figure D6. Density profiles averaged over last 10 ns for all components in (a) Toluene|S-Serine, (b) S-Tol-BINAP|S-Serine, (c) R-Tol-BINAP|S-Serine, (d) Toluene|R-Serine, (e) S-Tol-BINAP|R-Serine, and (f) R-Tol-BINAP|R-Serine. (Blue: water, red: toluene, cyan: serine, orange: Tol-BINAP.)

Table D1. Experimental interfacial tension (γ) in mN/m for the R-Tol-BINAP, S-Tol-BINAP and S/R-Tol-BINAP (50:50 v/v) (in toluene) in 6 and 4 wt% R/(S)-Serine (in water).

Samples	γ-Exp (5500 s)
R-Tol-BINAP|S-Serine (6wt%)	42.32 ± 0.10
S-Tol-BINAP|S-Serine (6wt%)	41.24 ± 0.09
S/R-Tol-BINAP (50:50)|S-Serine (6wt%)	41.70 ± 0.18
R-Tol-BINAP|S-Serine (4wt%)	40.39 ± 0.13
S-Tol-BINAP|S-Serine (4wt%)	39.60 ± 0.02
S/R-Tol-BINAP (50:50)|S-Serine (4wt%)	39.96 ± 0.09