Probing Molecular Structure with Two-Dimensional Infrared Spectroscopy

By

Daniel Ryan DeNeve

Thesis

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE

in

Chemistry

May 10th, 2019

Nashville, Tennessee

Approved:

Lauren E. Buchanan, Ph.D.

Michael P. Stone, Ph.D.

To my daughters, Olivia, Elizabeth, and Amelia

and

To my beloved wife, Shannon, endlessly supportive, loving, and understanding.

ACKNOWLEDGMENTS

This work would not have been possible without the financial support of The United States Army. I would like to thank the faculty at The United States Military Academy for their support, guidance, and for allowing me the opportunity to attend graduate school. I would also like to thank LTC Francisco Silverio, LTC William Moen, COL Jayson Gilberti (Ret.), CSM Guy Lohman, COL Matthew Neumeyer, COL Steven George (Ret.), CSM Brian Severino (Ret.), and COL Dean Bushnell (Ret.) for their mentorship and their inspiring example of leadership.

I am particularly grateful to all of my colleagues in Dr. Buchanan's group. I would especially like to thank Lauren, Will, Kelsey, Kathryn, Amanda, and Kayla for their patience, insight, and willingness to teach me many wide-ranging topics about chemistry. The members of my thesis committee have provided extensive personal and professional guidance and have taught me a great deal about teaching and scientific research. I would especially like to thank Dr. Lauren Buchanan, the chair of my committee. She is an outstanding teacher, scientist, and mentor who invests a great deal of time and energy in the personal and professional development of her students and every member of her group.

I would also like to thank my family for their support, love, and inspiration. To my loving and supportive wife, Shannon, and my three daughters, Olivia, Elizabeth, and Amelia, I want you to know that you are everything to me even if it may seem that I must demonstrate my love in seemingly strange ways by working late some nights and accruing many months and years spent away from home as I train, deploy, and fight for something that I hold so dear.

TABLES OF CONTENTS

Page

Chapter

LIST OF FIGURES

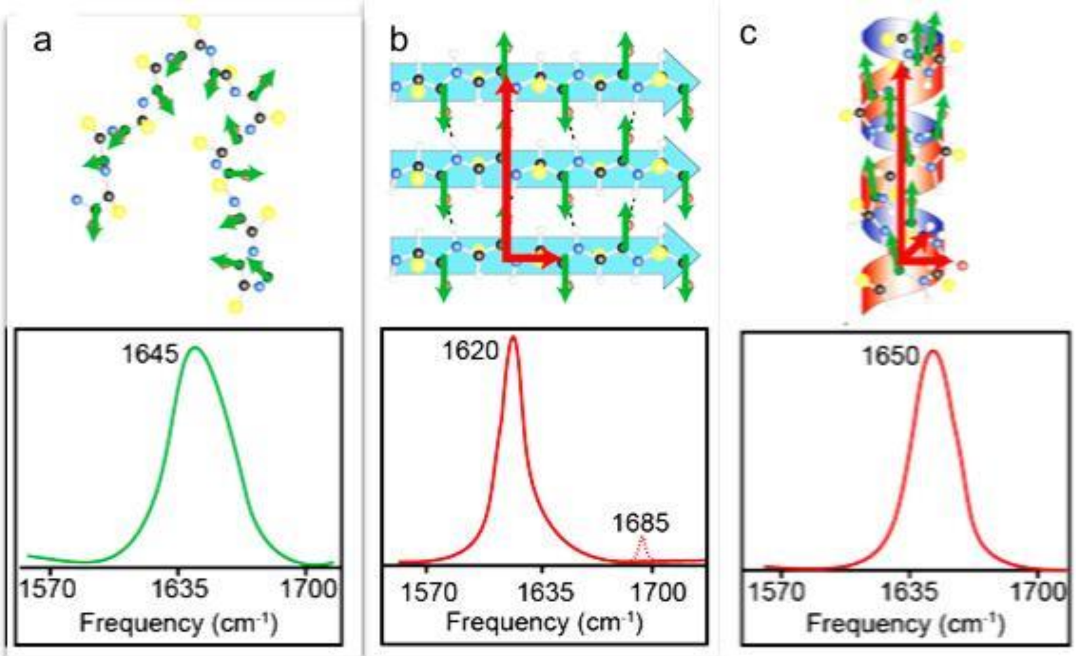
CHAPTER 1

Introduction

Drawing insight from biological processes where peptides, lipids, and various biomolecules will self-organize to form useful secondary structures, self-assembly represents a novel approach to create interesting nanoscale materials. Self-assembly is defined as a spontaneous organization of individual molecules into an ordered structure without the need for human intervention. Peptide self-assembly has seen biomedical applications such as tissue engineering and regeneration [1-2] and drug delivery [3-4]. In addition, peptide self-assembly has been adapted as a template to model the creation of metallic and semiconductor nanostructures [5-6]. By gaining insight into the mechanisms behind folding and self-assembly, new approaches to control the self-assembly process can be devised to design and create new functional nanomaterials.

While techniques such as x-ray diffraction [7], electron microscopy [8], NMR, and circular dichroism [9] have been used to examine the growth and structure of self-assembled peptide structures at the nano-scale, they generally lack either the temporal or structural resolution required to determine self-assembly mechanisms. Infrared (IR) spectroscopy is another well-established tool with wide-ranging applications in the physical and life sciences. Ultrafast two-dimensional infrared (2D IR) spectroscopy, a cutting-edge technique that has enjoyed great success in biophysical studies [10], demonstrates capabilities unmatched by conventional IR absorption spectroscopy due to inherently its higher sensitivity and enhanced temporal and spectral resolution. Through a sequence of excitation and probe pulses, 2D IR provides a powerful tool capable of examining how structure relates to function and how this relationship influences the fundamental chemistry serving as the basis of biological processes [11-12].

Infrared spectroscopy generally refers to the interaction of mid-IR light with matter to excite molecular vibrations. The spectral features serve as a molecular fingerprint that facilitates straightforward characterization. For peptide-based samples, the amide I mode is most commonly probed because it is sensitive to secondary structure [13]. The amide I mode is generated primarily by backbone carbonyl stretching, with lesser contributions from nitrogen bending and carbon-nitrogen stretching. The local frequency of the amide I vibration in aqueous solution is usually around 1645 cm^{-1}; in protein samples, a peak at this frequency corresponds to a disordered peptide. Changes in secondary structure can bring carbonyl groups into close proximity, leading to vibrational coupling. This causes the vibrational modes to delocalize over several amino acid residues and alters the frequency and linewidth of absorption peaks, creating distinct spectral features for each secondary structure. The formation of ß-sheets will cause a redshift to around 1620 - 1630 cm^{-1} due to alignment between amino acids in neighboring strands. The formation of α-helix structures will cause a blue-shift to around 1650 - 1660 cm^{-1}. Figure 1 demonstrates some of these relationships between the observed spectrum and the secondary structure.



Random/Disordered Beta Sheet Alpha Helix

Figure 1: A model peptide with individual amide I transition dipole moment indicated by green arrows and net dipole moment indicated by red arrows. Characteristic structures and linear infrared spectra are indicated for a) disordered peptide, b) parallel ß-sheets, and c) an α-helix. Adapted from reference [14].

The operation of the 2D IR experimental setup, however, requires constant vigilance and dedicated maintenance to realize its maximum potential. Daily maintenance and operation of the instrumentation, inherent limitations to sample lifetime, and the relatively high cost of ownership are all challenges which have limited the popularity and widespread implementation of 2D IR spectroscopy. Recent technical developments have improved the sensitivity, reliability, and speed of 2D IR experiments [15] and are opening up new opportunities to more easily probe molecular structures while mitigating some of the technical challenges imposed by earlier 2D IR experiments.

In chapter two, I will describe the theory and application of 2D IR. I will also discuss isotope labeling as it relates to 2D IR. In chapter three, I will describe the process of self-assembly, important applications, and the study of amphiphilic peptides to better understand the mechanisms of self-assembly. In chapter four, I will discuss the application of 2D IR to investigate hexagonal boron nitride. I will describe of the phonon-polariton interaction as it relates to optics and a unique class of materials. Initial results of the investigation will be discussed as well as future applications and directions. Chapter five will consider the application of 2D IR to investigate an ion channel for the influenza M2 protein and what can be discovered by continued investigation of this key drug target. Chapter six will detail all methods and lessons learned throughout my scientific experimentation and operation of instruments.

CHAPTER 2

2D IR Theory

2.1 Introduction

Scientific investigation regarding the structure and dynamics of any molecular system is difficult to address. Proteins tend to undergo structural fluctuations over vast time scales ranging from picoseconds to seconds [16]. While water molecules and ions can form or break a hydrogen bond on the order of one picosecond [17], the formation of secondary and tertiary structures can span a vast range of timescales, from microseconds to multiple days as shown in Figure 2. Sophisticated molecular dynamics simulations spanning the vast range in temporal orders of magnitude concerning chemical dynamics can quickly become computationally intractable. Experimental studies tend to favor particular instruments suited to the timescale of the dynamics under study. If the dynamics of the system under study are slow, nuclear magnetic resonance (NMR) spectroscopy provides the best structural information. Because the wavelength of x-rays are compatible with atomic distances, x-ray diffraction can also provide superior structural information as long as the sample can be crystallized. If the dynamics occur quickly, then optical spectroscopy can effectively probe the dynamics with a tradeoff in structural resolution. The ability to characterize the evolution of protein structure with bond-specific structural resolution and at all relevant time-scales is key to rigorously developing an understanding of their function. Structure and the evolution of structure is the most fundamental mechanism by which the function of a protein may be assessed. 2D IR spectroscopy has demonstrated its ability to track electron transfer in semiconductors [18], solvent dynamics [19-20], and the study of kinetics over varying time scales [21-22]. Understanding the kinetics of interconversion in a structural ensemble can help develop fundamental insights into protein functions [23].

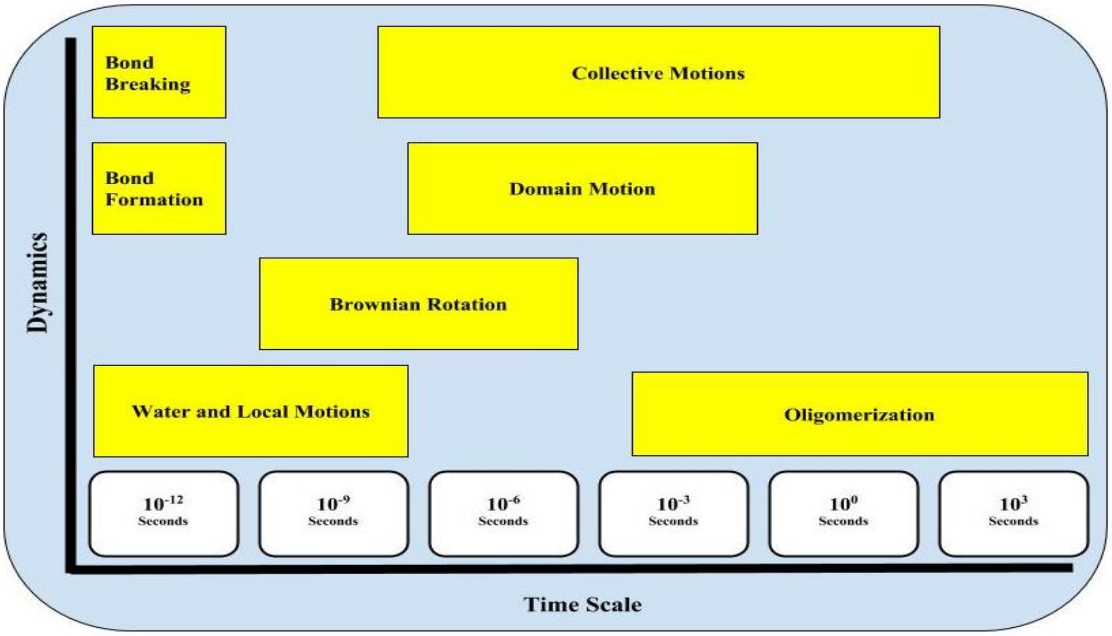


Figure 2: Relative timescales of protein and peptide dynamics.

The technology that makes 2D IR possible results from the theory of third-order nonlinear spectroscopy [24-25]. This technology, which requires the generation of ultrafast mid-IR pulses, has made significant progress in the past fifteen years [26-27]. The basic 2D IR experiment is implemented through the sequential interaction of a chromophore with three infrared laser pulses to generate a third-order polarization in a material which, in turn, generates a signal electric field. The signal generated in a vibrational-echo measurement is a function of the pulse delays between the excitation and signal fields. The first pulse from the laser generates a coherence between vibrational states which evolves over time before the second pulse terminates the evolution by the creation of a population. As a result, the delay between the first and second pulses is referred to as the coherence time. After a delay (referred to as the waiting time), a third laser pulse generates a new coherence which again evolves as a function of time (the detection time). This information can then be used to generate a spectral plot. The Fourier transform of this signal yields a 2D spectrum, which is a function of the frequencies of the coherence and detection time (ω_1 and ω_3 in Figure 3). Most commonly, 2D IR signals are detected by overlapping the signal field with a reference pulse, known as a local oscillator, to measure the amplitude and the phase of the generated signal. The combined fields are dispersed by a monochromator onto a mercury cadmium telluride (MCT) detector array to create an experimental Fourier transform over the detection time. This approach improves the strength of the detected signal while increasing the speed of data acquisition and reducing the influence of laser pulse instability [28].

2.2 Application of 2D IR

The top part of Figure 3 is a graphic representation of a 2D IR spectrum of a single vibrational mode in a molecule. A Fourier transform infrared spectroscopy (FTIR) spectrum of this mode would have an individual absorption peak centered about its vibrational frequency. The comparable 2D IR spectrum contains two out-of-phase transitions, known as diagonal peaks, along the diagonal line where pump and probe frequencies are equivalent ($\omega_1 = \omega_3$). The red peak appears along the diagonal line at the fundamental frequency of the vibrational mode, corresponding to a transition from the ground to the first excited state ($\nu = 0 \rightarrow 1$). The blue peak corresponds to an overtone transition from the first to the second excited state ($\nu = 1 \rightarrow 2$) and is offset from the fundamental peak along the probe axis by an anharmonic shift, Δ. 2D IR spectra allow the separation of homogeneous and inhomogeneous contributions to the vibrational linewidths. When identical vibrational modes exhibit a distribution of different frequencies, as is commonly observed in aqueous environments where molecules experience rapid hydrogen-bond exchange, the 2D IR peaks become elongated along the diagonal. In a more homogeneous environment, the 2D peaks will exhibit a nearly circular shape. The separation of homogeneous and inhomogeneous linewidths, particularly as the inhomogeneous frequency distribution becomes time-dependent, provides 2D IR with unique capabilities to investigate fast vibrational dynamics [29].

The bottom graphic of Figure 3 demonstrates a molecule with two vibrational modes. Multiple oscillators can be present on a peptide or protein and lead to off-diagonal features which are known as cross-peaks in 2D IR spectra. The ability to detect cross-peaks depends on the strength of the vibrational coupling between two modes. Cross-peaks represent coupled vibrational transitions. The off-diagonal anharmonicity separates cross-peaks by a distance of δ, as shown in Figure 3. The off-diagonal anharmonicity is often simplified in terms of the distance and angle between two transition dipoles because it depends on the coupling constant between two different modes. The ability to directly measure vibrational couplings and give insight into the molecular geometry between two coupled modes is an important and useful characteristic of 2D IR. Cross-peaks are observed when the laser pulses interact with modes in different locations, resulting in couplings and relative orientations that can be used to determine the structure. Solving Equation 1, shown below, allows the measurement of the relative angles between coupled transition dipoles (α and β) using a rigid rotor model. The application of Euler angles or spherical harmonics (by integrating the molecule over all possible orientations, $\int d\Omega$) allow the equation to be solved.

$$\langle(\hat{Z}\cdot\hat{\alpha})(\hat{Z}\cdot\hat{\alpha})(\hat{Z}\cdot\hat{\beta})(\hat{Z}\cdot\hat{\beta})\rangle = \frac{1}{4\pi}\int d\Omega \cos^2\theta_{Z\alpha}\cos^2\theta_{Z\beta}$$

[Equation 1]

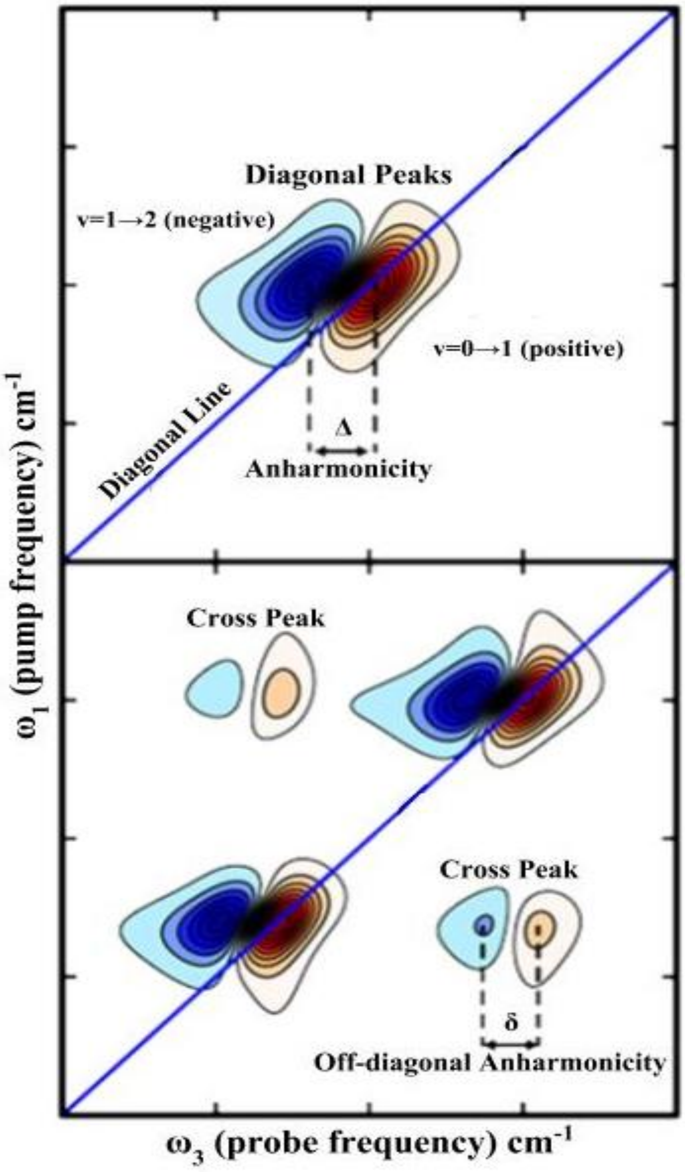


Figure 3: Characteristic features of a 2D IR spectrum illustrating cross peaks (bottom figure) and diagonal peaks (top figure). Cross peaks measure the coupling between molecular vibration and can provide structural information by modeling the structure dependence on the measured coupling. Diagonal peaks provide a fundamental frequency (lying on the diagonal) while the negative (blue) peak is shifted off the diagonal. Anharmonicity is a measure of the coupling between fundamental and overtone peaks. Spectra adapted from reference [10].

2.3 2D IR of Proteins

Because spectral features provide a rapid means of characterizing the secondary structure of proteins and peptides, the time evolution of spectra and the solvent environment can also provide useful information regarding structure. As discussed previously and shown in Figure 1, the amide I mode provides important information regarding the secondary structure of proteins. In the case of β-sheets, they may be oriented parallel or anti-parallel, depending on whether the strands run in the same or opposite directions from N- to C-terminus. Anti-parallel β-sheets are slightly more stable because their hydrogen bonding network is less strained. Despite this, FTIR cannot make a straightforward characterization as to whether a protein contains anti-parallel or parallel β-sheets due to the inherent configurational disorder present among proteins. Though challenging it is possible to make qualitative predictions as to whether β-sheets are parallel or anti-parallel by examining their 2D IR spectra [30]. Both parallel and anti-parallel β-sheets produce strong peaks centered near 1630 cm^{-1} (Figure 1b) in infrared spectra. The peak around 1630 cm^{-1} results from the in-phase oscillation of residues between adjacent strands. Anti-parallel β-sheets, however, typically exhibit a second, weaker peak around 1680 cm^{-1}, resulting from the in-phase oscillation of adjacent residues along the β-strands that is out-of-phase with respect to the hydrogen-bonded neighboring residue on the adjacent sheet. While this higher frequency peak tends to be more prevalent with anti-parallel β-sheets, some parallel structures can allow observation of this peak. While the spectral signatures of parallel versus anti-parallel β-sheet configurations have been established using model systems [31], factors such as local configurational disorder, the number of strands in the β-sheet, and the angle between the transition dipole moments all affect the delocalization of the vibrational excitation and thus the frequency, amplitude, splitting, and line shape of any observed diagonal and cross-peak features [32-33]. It is therefore important to confirm the assignment of parallel versus anti-parallel β-sheets by alternative means, such as isotope labeling.

The model of a perfect β-sheet, however, assumes a perfect order, an infinite length, and does not have any twists, turns, or fraying at the end of the peptide. To understand the significance of β-sheet or any other secondary structural peaks, it is important to have a baseline understanding of the vibrational Hamiltonian. The vibrational Hamiltonian is a measure of how vibrational couplings are affected by structural disorder and environmental fluctuations.

For proteins, the vibrational Hamiltonian is modeled in the basis set of the individual amide groups (also referred to as local modes). The fundamental vibrational frequencies of the amide groups appear on the diagonal of the matrix. Even in the example of a perfect, model β-sheet, the presence of water will still induce diagonal disorder due to fluctuating hydrogen bonds, resulting in inhomogeneous broadening of the peaks. While the frequency will be centered at 1645 cm^{-1}, diagonal disorder will be added to the Hamiltonian to make calculated spectra more realistic. Molecular dynamics simulations exist at the all-atom level which can accurately calculate these vibrational mode frequencies [34].

Calculation of the off-diagonal contribution must be performed for the coupling Hamiltonian by setting the coupling strengths between local modes. The coupling strength will depend on their relative distance, orientation, and mechanical anharmonicities between covalently bonded residues. The derivation of coupling strengths results from various models, *ab initio* calculations, and experiments [35]. Although more sophisticated models exist [24], the simplest model is that of transition dipole coupling. In this model, shown in Equation 2, the local mode transition dipoles interact with each other through dipole-dipole interactions. The coupling between two chromophores is given by β_{ij}, where μ is the dipole strength and r is the distance between dipole moments.

$$\beta_{ij}=\frac{1}{4\pi\varepsilon_0}\left[\frac{\vec{\mu}_i\cdot\vec{\mu}_j}{r_{ij}{}^3}-3\frac{(\vec{r}_{ij}\cdot\vec{\mu}_i)(\vec{r}_{ij}\cdot\vec{\mu}_j)}{r_{ij}{}^5}\right]$$

[Equation 2]

While the strongest coupling is between residues in adjacent β strands, other couplings within relative proximity are still influential and contribute to the coupling strength. Upon diagonalization of the Hamiltonian, an eigenstate can be generated for each oscillator, with each eigenstate potentially contributing to the vibrational spectrum. Because of the inherent symmetry of an idealized β-sheet, only two modes are observed (a lower frequency mode at approximately 1620 cm^{-1} and a higher frequency mode at approximately 1685 cm^{-1}). The remaining vibrational modes are unobserved but may contribute to the spectrum for more realistic β-sheet configurations as shown in Figure 4.

While the ideal β-sheet is a useful tool for developing an understanding of how to interpret structure based on the spectroscopic data, it is not exceedingly accurate nor is it a rigorous method used to interpret IR spectra [36]. This is primarily because the previously unobserved “dark” vibrational modes tend to increase in intensity for imperfect strands. As a result, peaks can broaden or shift in frequency and entirely new observable spectral features can appear.

As β-sheets do not start and stop abruptly, the edges of β-sheets tend to have a broader structural distribution due to variations in hydrogen bond strengths and phi/psi angles that may differ between local modes. Structural disorder, fraying, twist, turn, loops, and the size of the protein can also cause invisible modes to become more prevalent. This structural disorder can lead to a distribution of vibrational couplings. Solvent effects, such as fluctuating hydrogen bonds, will influence the entire protein and lead to a distribution of local mode frequencies. Both the diagonal and off-diagonal matrix elements in a vibrational Hamiltonian will differ as a result. Realistic systems which account for the effects of disorder would generate many Hamiltonians which must be diagonalized individually with the summation equating to the calculated spectra. The frequency difference, for example, of a hydrogen-bonded compared to a non-hydrogen-bonded amide-I mode can account for approximately 10 cm^{-1} of difference [36]. Differences in the local environment can also cause about 5-10 cm^{-1} of difference. These differences are similar to or greater than the coupling strengths and will change the frequencies of the β-sheet modes leading to more congested spectra as a result of the increased intensity of previously dark modes.

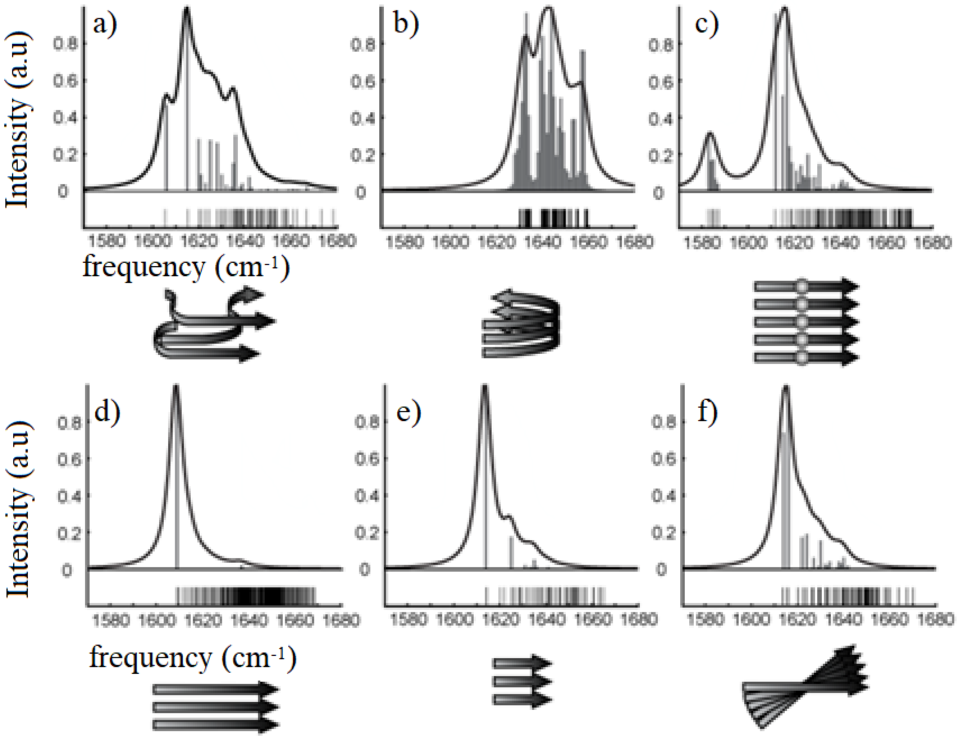


Figure 4: Simulated FTIR spectra based on characteristic secondary structures relevant to proteins and peptides. a) structurally disordered β-sheet, b) loop/turn structure, c) isotope labeled β-sheet (featuring isotope labeled modes centered around 1585 cm^{-1} and unlabeled features centered around 1620 cm^{-1}), d) perfect, infinite β-sheet, e) perfect, finite β-sheet, f) twisted β-sheet. Adapted from reference [36].

2.4 Isotope labeling

An important application of 2D IR spectroscopy is to use isotope labeling to study individual amino acids. Because vibrational frequencies are a function of the mass of the vibrating atoms, the introduction of isotope labels into amide carbonyls can serve as residue-specific structural probes. Isotope labeling is non-invasive and can spectrally isolate the residue by shifting the frequency of the vibrational mode of a peptide residue without altering the structure of the peptide. It is possible to methodically map the structure of a peptide or protein by incorporating isotope labels at different positions throughout a peptide. Because the amide I vibrational mode is primarily a result of the carbonyl stretching mode (C=O), it is advantageous to incorporate isotope labeling using a ^{13}C and ^{18}O label. The ^{13}C label shifts the amide I mode approximately 40 cm^{-1} lower in frequency. The 1% relative abundance of the ^{13}C isotope in nature, as well as potential side chain absorbances from other residues around 1600 cm^{-}1, can complicate interpretation of ^{13}C signals. Incorporation of a ^{13}C^{18}O carbonyl shifts the frequency approximately 55 cm^{-1} lower into a relatively empty region of the spectrum and diminishes the expectation of a naturally occurring isotope label down to approximately 0.002%, thus virtually guaranteeing that the only visible isotope-labeled residue observed is that which has been incorporated synthetically. Isotope labeling has proven to be a powerful tool in vibrational spectroscopy to determine molecular structure with bond-level resolution [37].

2D IR has some advantages over FTIR with respect to isotope labeling. The sensitivity of a 2D IR signal is inherently higher than linear IR because signal scales with the transition dipole strength as $|\mu|^4$, whereas linear IR signal scales as $|\mu|^2$ [24]. This will enhance the resolution of the isotope-labeled peak [38]. The local environment and secondary structure of labeled residues can be determined from line shape analysis of 2D IR spectra. Single-residue isotope labeling has been used to determine the orientation of membrane-bound peptides [39], while segmental isotope labeling of large proteins has been used to separately determine the secondary structure of the different domains [34].

When inserting an isotope label into a peptide, there is no guarantee that its frequency will appear precisely 55 cm^{-1} red-shifted from that of the corresponding unlabeled protein structures; instead the shift will depend on the position of the isotope label within the peptide because this will, as previously mentioned, affect how the residues coupe to one another. Thus, the frequencies observed through isotope labeling cannot be interpreted by simply subtracting the predicted isotopic shift from the characteristic frequencies of unlabeled protein structures [40]. Instead, to extract essential structural information from labeled 2D IR spectra, we must return to the vibrational Hamiltonian discussed in the previous section.

It is necessary to understand the vibrational modes of the peptides to be studied to help guide the design of the ideal isotope labeling scheme. As discussed previously, the model IR spectrum of a perfect β-sheet is quite simple. It will exhibit a strong peak around 1620-1630 cm^{-1} and a weaker peak around 1680-1690 cm^{-1} in the case of anti-parallel β-sheets.

The shift in frequency (shown in figure 4c) allowing an independent interpretation of an isotope-labeled peptide's structure can be more clearly understood by applying the previous discussion of the vibrational Hamiltonian. For an isotope-labeled Hamiltonian, a subtraction of approximately 55 cm^{-1} accounts for redshift due to isotope labeling from the diagonal elements corresponding to the isotope-labeled position(s). It should be noted that consideration of the off-diagonal elements will not change in value since isotope labeling does not change the fundamental structure of the strand. The coupling between labeled and unlabeled modes can be effectively ignored because the diagonal frequency shifts are sufficiently large compared to the coupling strength. The coupling between the isotope-labeled residues, however, cannot be ignored. As a result, it is possible to accurately block-diagonalize the Hamiltonian and treat isotope labeled residues independently from unlabeled residues. Just as with the unlabeled modes, more stable and rigid protein structures will tend to have the largest frequency shifts.

2.5 Conclusion

The physical model of 2D IR spectra shown in Figure 3 provides a means of interpreting observable characteristics such as frequencies, cross-peaks, and various 2D line shapes to develop a qualitative interpretation of the structural dynamics. A stronger and more quantitative understanding of 2D IR can be developed by learning the underlying concepts and models of vibrational coupling, the general effects of disorder and the local environment in realistic systems, and the application of isotope labeling to provide enhanced structural resolution. Because protein dynamics occur over vast time scales and experimental observation requires high-resolution structural dynamics to be captured in real time, 2D IR is uniquely suited to answer questions about the structural evolution of proteins to gain enhanced insight into the underlying biophysics as we enhance our understanding of protein structure as it relates to function. A 2D IR spectrum essentially provides a near-instantaneous snapshot of a protein's structural distribution. As we will see in the following chapter, molecular dynamics simulations will provide a complementary means of testing 2D IR experiments against simulations.

CHAPTER 3

Self-assembly and Amphiphilic Peptides

3.1 Introduction

Short-chain peptides are relatively simple to design and synthesize and represent an ideal model system for studying the process of self-assembly due to their relatively strong noncovalent interactions. The primary factors that drive self-assembly are complementarity in shape [41] and the effect of noncovalent interactions [42] such as hydrophobicity, hydrogen bonding, electrostatics, and van der Waals forces. By designing relatively simple systems which will undergo spontaneous organization into a stable and easily characterized macroscopic structure consisting entirely of non-covalent bonds, our goal has been to gain systematic insight into the fundamental mechanisms behind folding and self-assembly. Our study of short-chain peptides has focused on amphiphilic and surfactant-like peptides due to their relative simplicity and the rate at which they tend self-assemble in an aqueous solution.

Previous studies demonstrate that peptides which contain a hydrophobic head group and a hydrophilic tail group will self-assemble in an aqueous solution to form distinct nanostructures such as nanosheets, nanofibers, and nanorods. These distinctions are primarily due to the effects of hydrophobicity, which can drive the nonpolar region of each peptide towards one another to exclude water in a dry pocket. The shape and dimensionality of these supramolecular structures ultimately depend upon the electrostatic environment which is heavily influenced by the shape of the molecule and the polar head group [43].

Studies show that relatively small changes to the sequence of surfactant-like peptides can significantly alter the ultimate self-assembled nanostructure. Xu reported on such a series of surfactant peptides by synthesizing A_3K, A_6K, and A_9K (which differ only in the length of their hydrophilic alanine chain) [44]. These peptides self-assemble into three distinct nanostructures: small nanosheets for A_3K, long nanofibers for A_6K, and short nanorods for A_9K.

Figure 5: Transmission electron microscopy images of the nanostructures formed by the self-assembly of A_3K (left), A_6K (center), and A_9K (right) peptides. [44].

3.2 Molecular Dynamics Simulations

All-atom molecular dynamics (AA-MD) simulations have previously been used on various surfactant-like peptides as residue-specific investigations to determine whether AA-MD simulations agree with experimental models. This comparison aimed to reveal whether β-sheets form in parallel or anti-parallel configurations and how the sheets organize. Computational modeling and molecular dynamics simulations provide useful insight and supplementary information to experimental studies regarding the complex and dynamic electrostatic forces encountered on the surface of a peptide.

Computational simulations are an important complementary method that serves to validate structural models proposed from spectroscopic observations. Understanding the inter- and intra-molecular interactions that drive self-assembly of amphiphilic peptides is essential in determining how sequence modulates the self-assembly mechanism. By ruling out unstable configurations through simulation, other experimental characterization methods can continue to refine plausible intermediate and final configurations.

In 2014, Fileti conducted extensive AA-MD simulations to describe the structure and energetic potential of the self-assembled nanostructures of A_3K, A_6K, and A_9K amphiphilic peptides [45]. Figure 6 shows successful and unsuccessful starting structures probed in the study. While their results indicate that that A_3K can form stable nanosheets in both parallel and anti-parallel configurations, our 2D IR spectra indicate that the anti-parallel configuration is dominant.

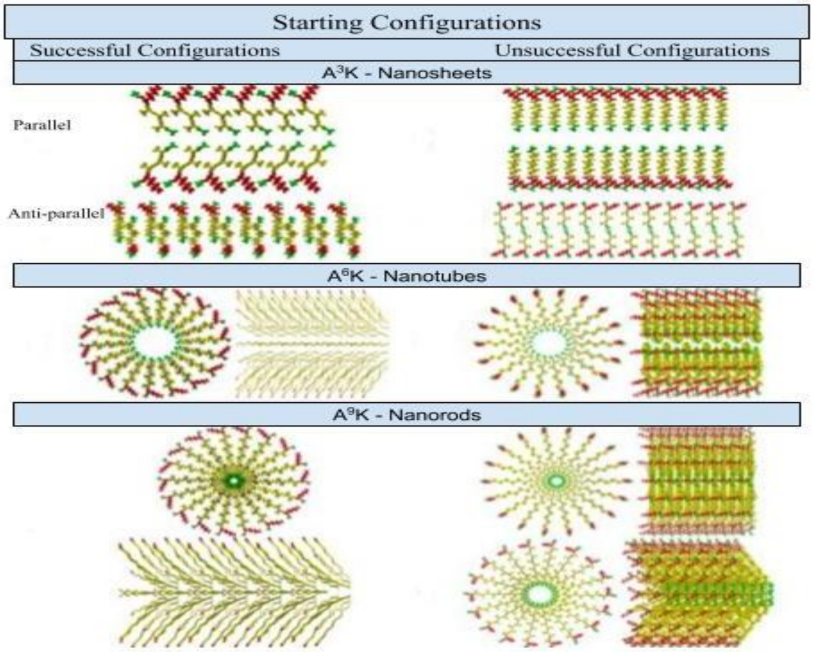


Figure 6: (left) A_nK successful starting structure and (right) A_nK unsuccessful starting structures, based on AA-MD simulation. Adapted from reference [45].

3.3 Results

My colleagues in the Buchanan lab have thus far synthesized a series of short-chain amphiphilic peptides: A_3K, A_6K, A_9K, V_6K, and V_6K_2. Chapter 6 discusses methods regarding solid-phase peptide synthesis, purification, and characterization. These peptides share common design characteristics: a polar headgroup consisting of 1-2 lysine residues and a non-polar tail consisting of 3-9 alanine or valine residues. All peptides were synthesized with an amidated C-terminus.

My initial attempts at investigating self-assembly were unsuccessful as the first peptide we synthesized, A_3K, failed to aggregate. Spectroscopic measurements indicated that our peptide maintained a signal indicating random coil structures for over 50 days until the peptide ultimately degraded. Figure 8 summarizes our results. While it typically takes approximately 7-days for the formation of β-sheets to occur, peptides synthesized using the A_3K amphiphilic peptide did not undergo self-assembly. We determined that N-terminal acetylation was required for self-assembly to occur. While there is no explicit mention of this fact in scientific publication, our follow-up spectroscopic measurements made it abundantly clear to us that a failure to perform the post-synthesis modification of N-terminal acetylation inhibits the formation of β-sheets for A_3K, A_6K, and A_9K amphiphilic peptides. Section 6.2 provides further discussion regarding incorporation methods of N-terminal acetylation.

Because the majority of eukaryotic proteins undergo N-terminal acetylation, the performance of this reaction is a critical step in ultimately assessing the relationship between structure and function for proteins [46]. The mediation of protein complex formation by N-terminal acetylation has received recent recognition [47]. While the biological role of N-terminal acetylation was initially presumed as a means to protect proteins from degradation, recent studies suggest that functional roles and implications are diverse among the thousands of proteins studied thus far [48].

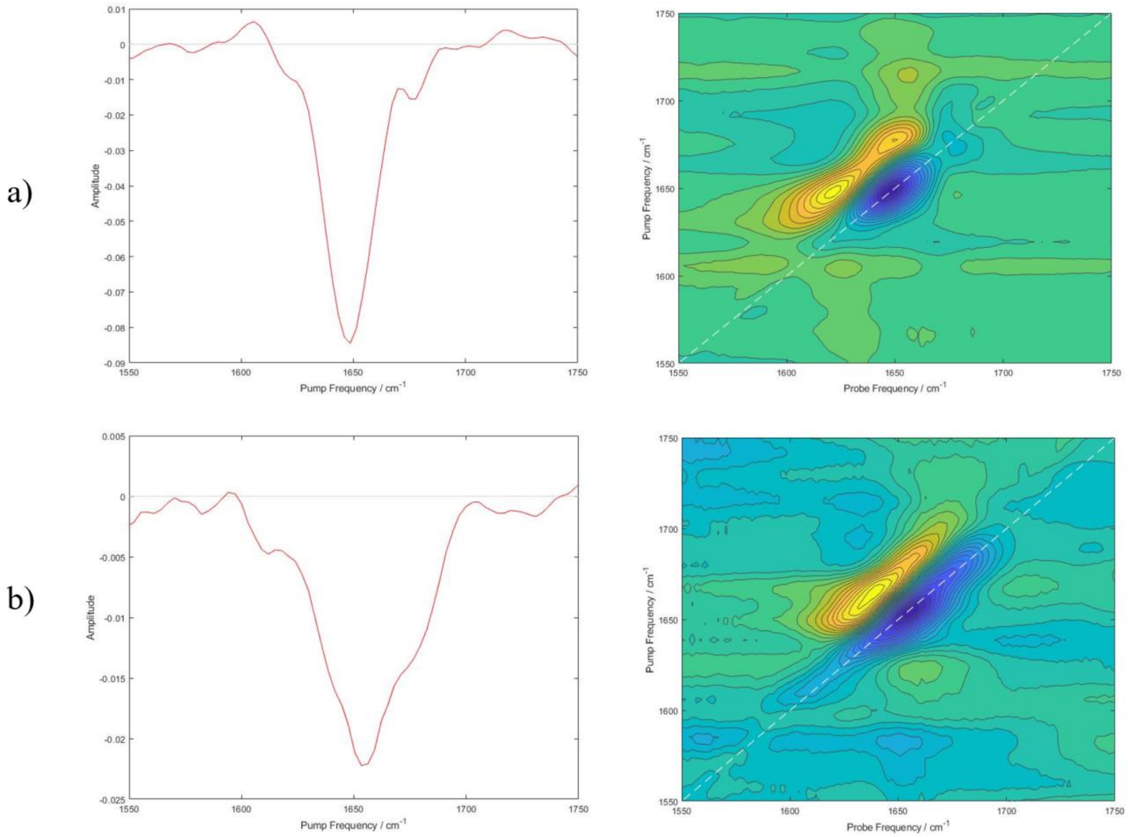


Figure 7: 2D IR spectra and diagonal slices of unacetylated A_3K after 10 and 56 days of dwell time. a) Diagonal slice (left) and 2D IR spectrum (right) of unacetylated A_3K, 10 mM concentration in D_2O. 13-day sample dwelling time. Spectrum acquired using parallel polarization and averaged over 50 scans. The peak is centered at approximately 1645 cm^{-1}, which indicates a disordered/random structure. b) Diagonal slice (left) and 2D IR spectrum (right) of unacetylated A_3K. 10 mM concentration in D_2O. 46-day sample dwelling time. Spectrum acquired using parallel polarization and averaged over 40 scans. The peak is centered at approximately 1655 cm^{-1}, which likely indicates a disordered/random structure with some broadening and fraying, perhaps due to degradation of the sample.

Subsequent syntheses of amphiphilic peptides performed with N-terminal acetylation demonstrated the formation of β sheets, as expected. Figure 9 illustrates an example spectrum of N-terminal acetylated A_3K. The lower frequency peak is centered at approximately 1629 cm^{-1}, which indicates a β-sheet structure. The higher frequency peak is centered around 1680 cm^{-1} and has a lower overall intensity since it manifests as the in-phase oscillation of adjacent residues along the β-sheets and is out-of-phase with respect to the hydrogen-bonded neighboring residue on the adjacent sheet. The diagonal is somewhat elongated due to additional resonances. The appearance of cross-peaks would more decisively confirm an anti-parallel classification of the β-sheets. A qualitative assessment of this peptide suggests anti-parallel β-sheets, but the experiment needs to be carried out with more rigor by allowing more overall scans to average out the spectrum and reduce noise and through the additional use of a perpendicular polarization scheme to amplify the signal of cross-peaks.

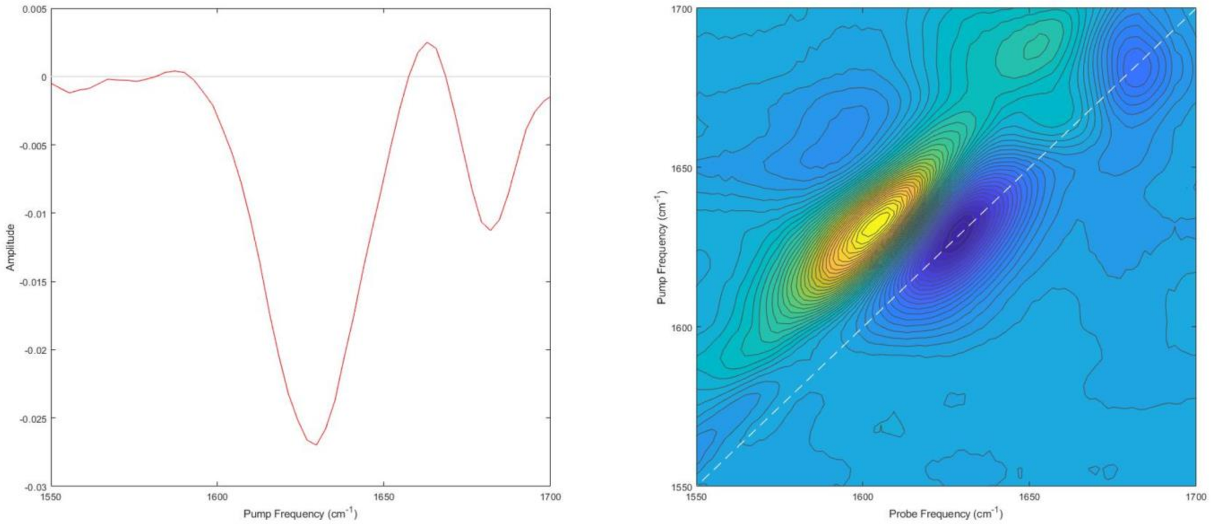


Figure 8: Diagonal slice (left) and 2D IR spectrum (right) of N-terminal acetylated A_3K showing aggregation after 13-day sample dwelling time. 25 mM concentration in D_2O. Spectrum acquired using parallel polarization and averaged over 20 scans. Most intense peak centered at 1630 cm^{-1} and less intense peak centered at 1682 cm-1

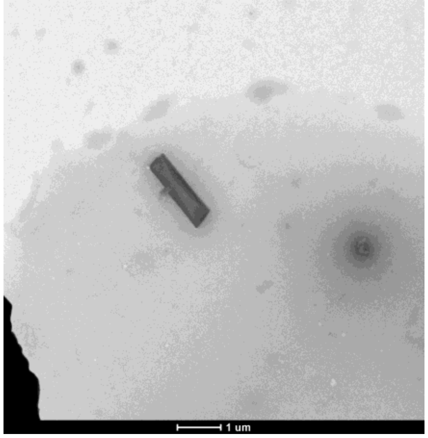


Figure 9: TEM image of A_3K. The image indicates stacked bilayers as described in simulation and other previously established experimental findings.

Peptide samples can be observed under transmission electron microscopy (TEM) by utilizing a uranyl acetate negative staining method originally described by Liu [65]. Observation under TEM can serve as a means of verifying the composition of the structure formed as a result of stacking beta sheets. Figure 9, shown above, demonstrates a recently acquired TEM image of A_3K matching imagery results seen from literature.

3.4 Conclusion and Future Directions

Recent increases in nanotechnology investment have brought about interest in the development of materials that can self-assemble into well-defined structures at the nanoscale. Although reproducible and ordered structures are common in biology, the production of well-ordered reproducible structures at the nanometer scale presents a significant engineering challenge for advancing materials science.

Now that our synthesis and characterization methods have been established and refined and acetylation has been incorporated into our peptide synthesis standard operating procedures, more isotope-labeled amphiphilic peptides will be produced, with subsequent interest placed on

V_6K, A_6K, V_6K_2, and A_9K. Each peptide will undergo variations in the labeling scheme to include a variation of the position of the isotope along the hydrophilic chain as well as a consideration for double labeling.

By studying incremental changes to variation in amino acid sequence and identity, we will continue to develop a better understanding of how peptide length, hydrophobicity, electrostatics, and hydrogen bonding all affect the self-assembly process. Upon completion of preliminary studies for the aforementioned peptides, we will focus on varying other parameters such as repeat length of polar and nonpolar regions, and the selection of amino acid residue. As our systematic approach to alter one variable at a time is sustained while continuing to build a more extensive database of surfactant-like peptides, we will continue to develop a better understanding of the variables influencing the mechanisms of self-assembly and come closer to the realization of our long-term goal of guiding the rational design of self-assembled peptide nanomaterials.

CHAPTER 4

Spectroscopic Investigation of Hexagonal Boron Nitride (hBN)

4.1 Introduction

Diffraction places limitations on conventional optics and the principal resolution limit must be in proportion to the measured wavelength of light in the material. For mid and far infrared measurements, the diffraction limit inherently becomes more problematic as the minimum spot size ranges from 10-100 µm as compared to approximately 200-400 nm for visible light. Because conventional optics employ materials with a positive dielectric constant, the only means of manipulating and confining plane waves is to employ dielectric materials with a high refractive index to confine light. For practical purposes, this can be a sufficient improvement to allow the use of some optical devices, but the fundamental diffraction limit still places a hard limit on performance as the wavelength of light increases.

In 2015, Dai performed sub-diffractional infrared nano-imaging through crystals of hBN by exploiting its hyperbolic properties to act as a ‘hyper-focusing lens’ and a multi-mode waveguide [49]. This surpassed then-current artificial hyperbolic materials by achieving lower loss and higher optical localization, allowing new opportunities for anisotropic layered insulators in infrared nanophotonics. Novel optics, light sources, and detectors can be realized by carefully designing polaritonic materials and nanostructures to confine light to the nanoscale and greatly enhance light-matter interactions. This can be achieved in the mid-IR to terahertz range by exploiting the vibrational polariton interaction and knowing the optical and electric properties of polaritonic materials.

4.2 Phonon-polaritons

In a material with a negative dielectric function, the refractive index will be imaginary and the electromagnetic fields will decay evanescently. While light would not propagate in such a material and its usefulness would be limited to only reflective optics, it is possible to expand the application of these materials by exploiting the formation of polaritons [50]. A polariton is a quasiparticle resulting from the coupling between light and an oscillating charge in a material. Phonon-polaritons arise when light interacts with a polar crystal lattice [52]. A broad range of materials exist that can support surface phonon-polaritons (SPhPs) from the mid-IR to the terahertz region of the electromagnetic spectrum (see Figure 10).

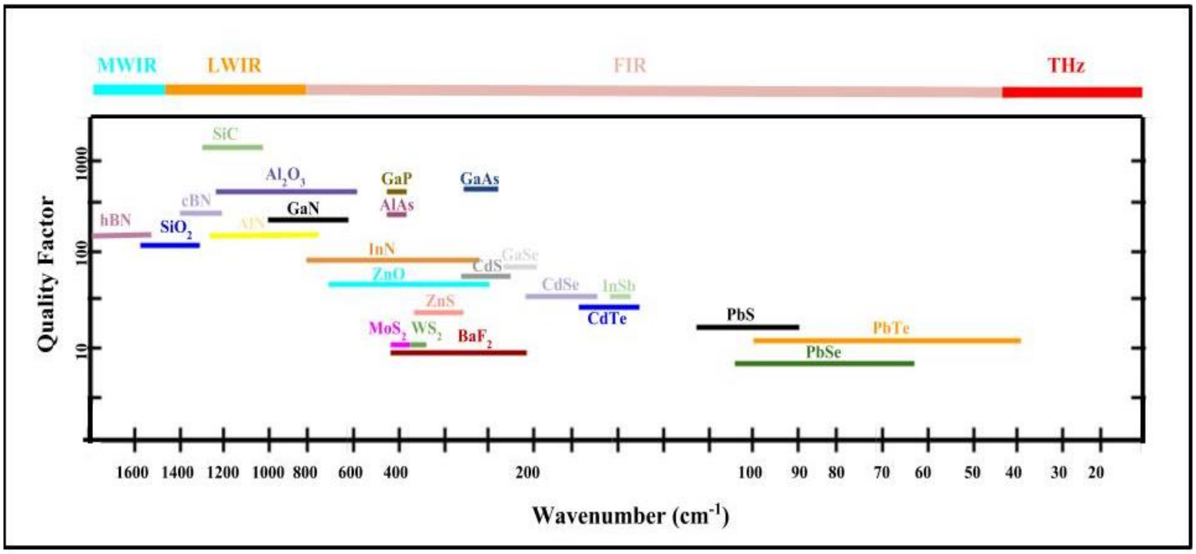


Figure 10: Spectral ranges of the reststrahlen bands for polar materials indicating the range where any material can support SPhPs. Adapted from reference [50].

A review article by Caldwell [51] suggested that exploitation of SPhPs could provide substantial advancements over current state-of-the-art infrared optics. While standard anisotropic materials have a dielectric function that behaves differently between the crystallographic axes of a material, an intriguing case arises when the dielectric function along orthogonal axes are not just different, but also opposite in sign. A material with the aforementioned opposite signs is called hyperbolic and supports polaritons with unusual properties that relate to the propagation of their electromagnetic waves with respect to the dielectric function. For a typical anisotropic material, frequency and spatial frequency components (ω and k), are related as in Equation 3 [52]. For a given frequency, the solutions to this equation can be represented in k-space by a surface detailing the possible values of k (i.e., polariton wavelengths) that are possible at that particular frequency for the material. The example of a sphere shown in Figure 11a is for an isotropic material. For an anisotropic material, all the axes exhibit the same sign, but because they have different values, they will form an ellipsoidal surface rather than a sphere to visually represent the values of k. When the signs for dielectric constants $\varepsilon_{x,y}$ and ε_z are different, then an open hyperbolic surface as seen in Figure 11b and Figure 11c emerges. Figures 11b and 11c differ in whether one axis (z) or two axes (x,y) are negative in sign, which is referred to as type 1 and type 2 response, respectively. Geometry-dependent localized vibrational polariton resonances can be established within the reststrahlen band of a material by carefully fabricating polar dielectric materials such as hBN [53]. Longitudinal vibrations occur in-phase with the direction of propagation and transverse vibrations occur perpendicular to the direction of propagation. The reststrahlen band is the region of the spectrum between the longitudinal optic (LO) and transverse optic (TO) vibrations where the real part of the dielectric function is negative.

$$\frac{k_x^2+k_y^2}{\varepsilon_z}+\frac{k_z^2}{\varepsilon_{x,y}}=\left(\frac{\omega}{c}\right)^2$$

[Equation 3]

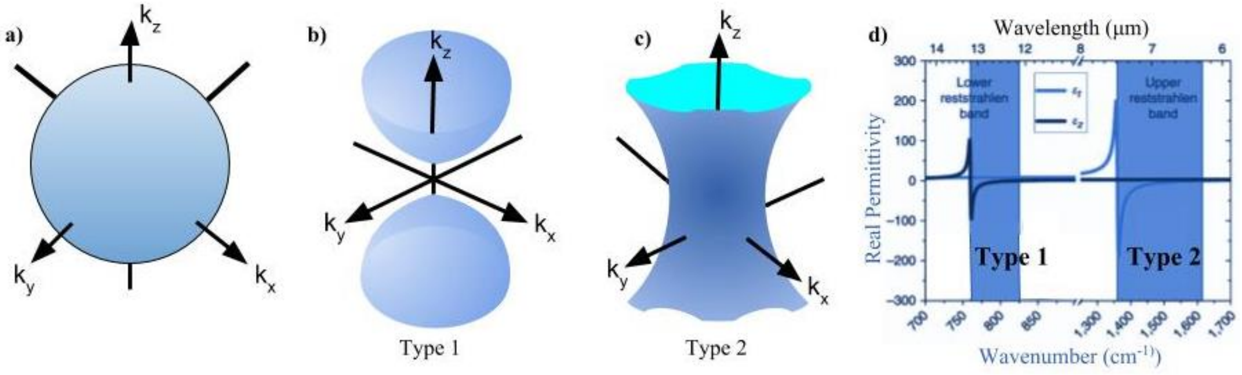
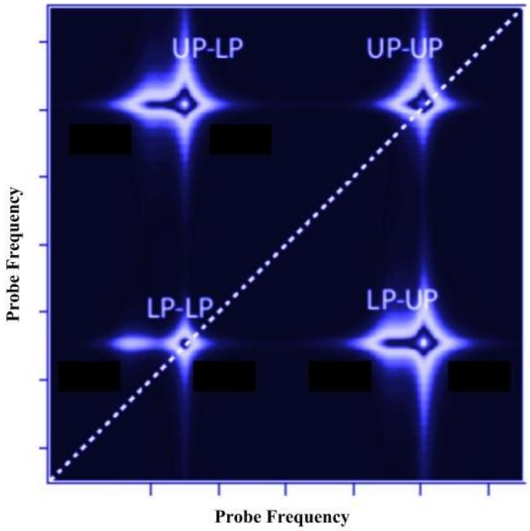


Figure 11: Frequency surfaces for a) isotropic material b) type 1 hyperbolic material c) type 2 hyperbolic material d) the dielectric function of hexagonal boron nitride, demonstrating the hyperbolic nature of the crystal structure. Adapted from reference [50].

4.3 2D IR of Polaritons

Recalling the previous discussion of 2D IR experiments, infrared spectra are spread into a second dimension and provide information on vibrational/phonon couplings and separate homogeneous and inhomogeneous dynamics as seen through line broadening. 2D IR spectra can also display cross peaks between coupled vibrational modes for polaritons. The investigation of these cross peaks may provide a powerful tool for studying the dielectric properties of 2D materials and potential interactions between LO and TO phonons.



[redacted] [redacted] [redacted] [redacted] [redacted] [redacted]

Figure 12: Example simulated 2D IR spectrum of an upper and lower polariton. Diagonal peaks represent a fundamental and overtone self-interaction while cross peaks represent an interaction that occurs off the diagonal. Adapted from reference [54].

In 2018, Dunkelberger applied 2D IR spectroscopy to observe light-matter interactions between vibrational polaritons for $W(CO)_6$ [55]. In resolving the interaction between upper and lower polaritons, 2D IR revealed unexpected dark states that traditional linear spectroscopy could not otherwise probe [56]. While upper and lower polaritons are defined as eigenstates and modeled in terms of a non-interacting (harmonic) exciton system, Figure 12 demonstrates self-interactions between the upper and lower polaritons respectively as well as cross-interactions between the upper and lower polaritons. Because it is difficult to distinguish between self-interaction and cross-interaction through one-dimensional spectroscopy, the self and cross-interactions between lower and upper polaritons have never been thoroughly investigated. Self-interaction results from pump mode polariton excitation with an external laser source. Self-interactions are thought to originate from the Coulomb and exchange interactions between the fermion components of the exciton-polariton.

In 2018, I attempted to observe the upper restrahlen band of an hBN sample under 2D IR with measurement centered around 1365 cm^{-1}. The sample, prepared by Tom Folland of the Caldwell group at Vanderbilt University, was fabricated by exfoliating a flake of hBN of approximately 4000 µm^2 in area onto a silicon substrate. The sample was attached and centered onto a modified sample cell holder with a calcium fluoride (CaF_2) window which is transparent in the mid-IR. The spectrum, shown below in Figure 13, appears to show a signal present in the area of interest in which the upper restrahlen band should be found. The spectrum also appears to exhibit significant scatter across the detection window. Further attempts should be made to verify this signal and suppress scatter by observing a new hBN sample and performing prior calibration.

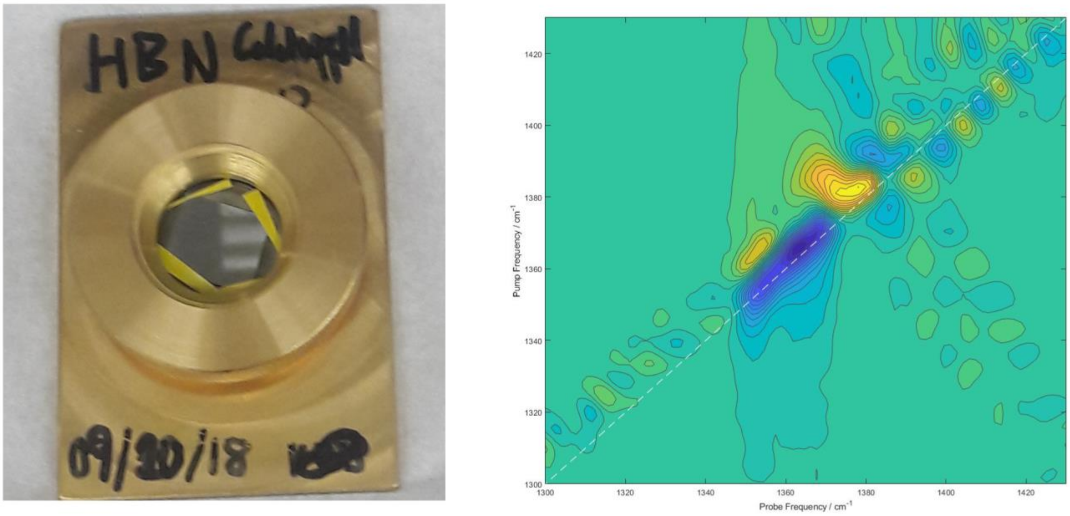


Figure 13: (left) A sample cell containing hBN fixed to a silicon substrate and centered on the viewing window. (right) 2D IR spectrum of hBN. There appears to be signal at approximately 1365 cm^{-1}, but there is also a significant amount of scattering across the diagonal.

Because the region centered at 1365 cm^{-1} had never been tested or calibrated before by the spectroscopic instruments of the Buchanan group, achieving the temporal and spatial overlap required to generate signal was a time-consuming task. Multiple instrument and software adjustments were needed to account for the change in path and dispersion caused by the redshift in output from the optical parametric amplifier. Calibration is an important step to reduce experimental variables and facilitate the acquisition of temporal and spatial overlap to achieve signal within the spectral region of interest. Calibration can also eliminate the possibility of self-interaction, which can occur if temporal resolution is inadequate and pump and probe pulses do not interact with the sample in the proper order. Obtaining 2D IR spectra in a different region of the mid-IR spectrum is not a simple feat and requires changing the grating angles, the pump and probe delay stage, the pump steering mirrors, the z-axis of the sample translation stage, and the group velocity dispersion correction within the pulse shaper software. All of this is to ensure that the pump and probe pulses are properly overlapped both spatially and temporally and focus directly on the sample. Calibration or lack thereof is additionally complicated by the fact that adjusting all these settings on an hBN flake can be quite tedious because the sample size area is only on the order of 4000 μm^2, which allows a minimal margin for error in placing the intersecting pump and probe beam on the correct position of the sample.

Due to the significant time and labor involved in fabricating an hBN sample and the numerous considerations that must be accounted for to ensure a sample is ready to observe under 2D IR, it is a more practical expenditure of time and resources to find an appropriate series of calibrant molecules and adjust all experimental setup conditions accordingly as the laser is redshifted from its known setup conditions in the Amide I spectral region at approximately 1626 cm^{-1} to the area of interest at the upper restrahlen band of hBN at approximately 1365-1370 cm^{-1}. This procedure is described in Chapter 6.

4.4 Manipulation of sample tilt angle to tune spectral frequency

In 2018, Wu and Fu proposed a method by which the control of electromagnetic wave propagation in hBN is possible by tilting the optical axis with respect to an incoming laser pulse [57]. This manipulation of the wavevector can be used to tune and strengthen the absorption in a sample of hBN. By stacking samples and slabs of hBN at different angles, it is possible to widen the wavenumber range and increase the absorptivity to tune absorption as desired within the reststrahlen bands. A combination of impedance matching and enlarged wavevectors at the sample interface contribute to the strong absorption. This reveals a new way to realize strong absorption in anisotropic materials and may lead to applications in areas related to thermal radiative energy harvesting and conversion [57].

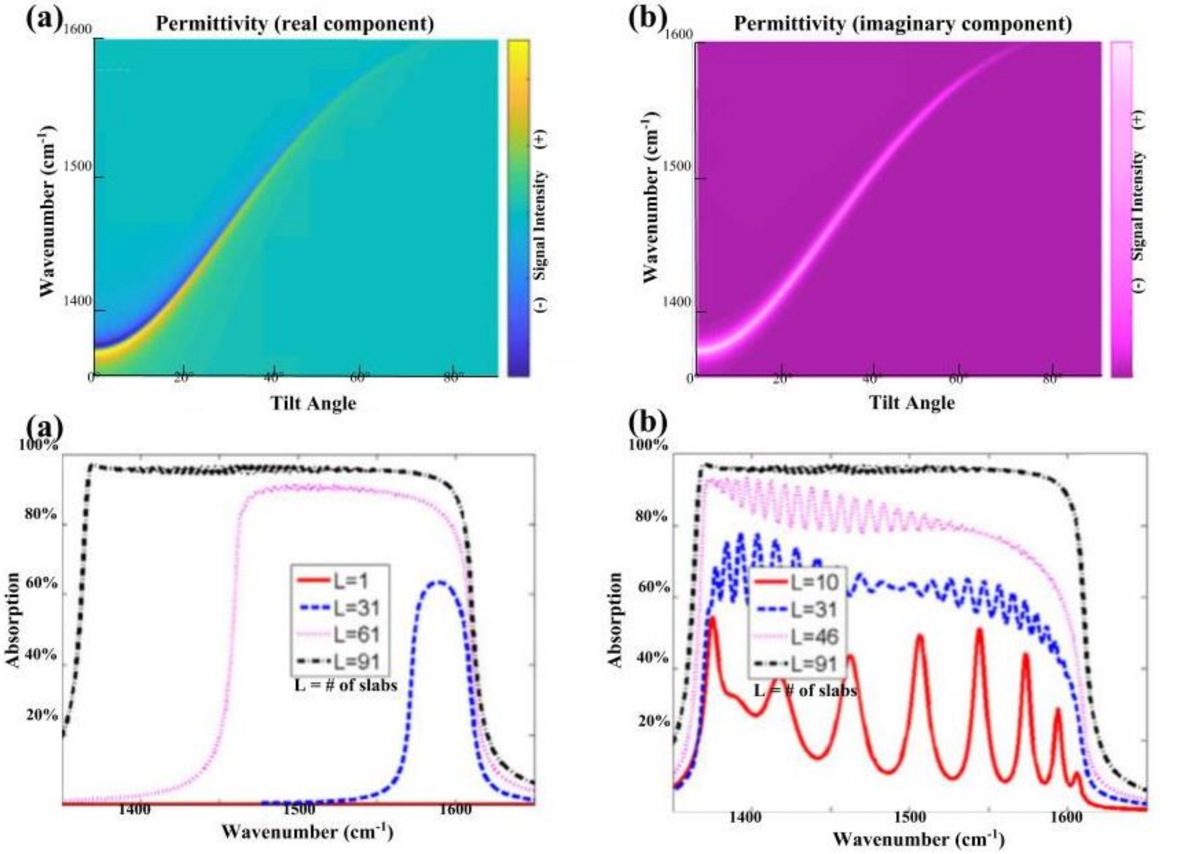


Figure 14: (top) Real and imaginary components of the permittivity with respect to tilt angle for hBN. (bot) Absorptivity of a layered structure vs. wavenumber for different numbers of slabs. (a) Difference in titled angles between adjacent slabs is 1°. (b) Difference in tilted angles between the slabs is 10°, 3°, 2°, and 1°, respectively. Adapted from reference [57].

Through the utilization of our current experimental setup, we can adjust the tilt angle of the sample by up to 7° and assess the change in the absorption of the sample. As can be seen from Figure 14, our ad-hoc method may be insufficient in demonstrating a significant shift in sample absorption tunability. A sample rotation stage would allow a more precise and variable means of experimentation (see Figure 15, right image). A sample rotation stage can incrementally tilt the sample up to 45°. The rotation adjustment, however, will require a slight translation in the z-axis (coincident with the incoming laser light), to account for the change in the intersection of light with the sample due to tilt.

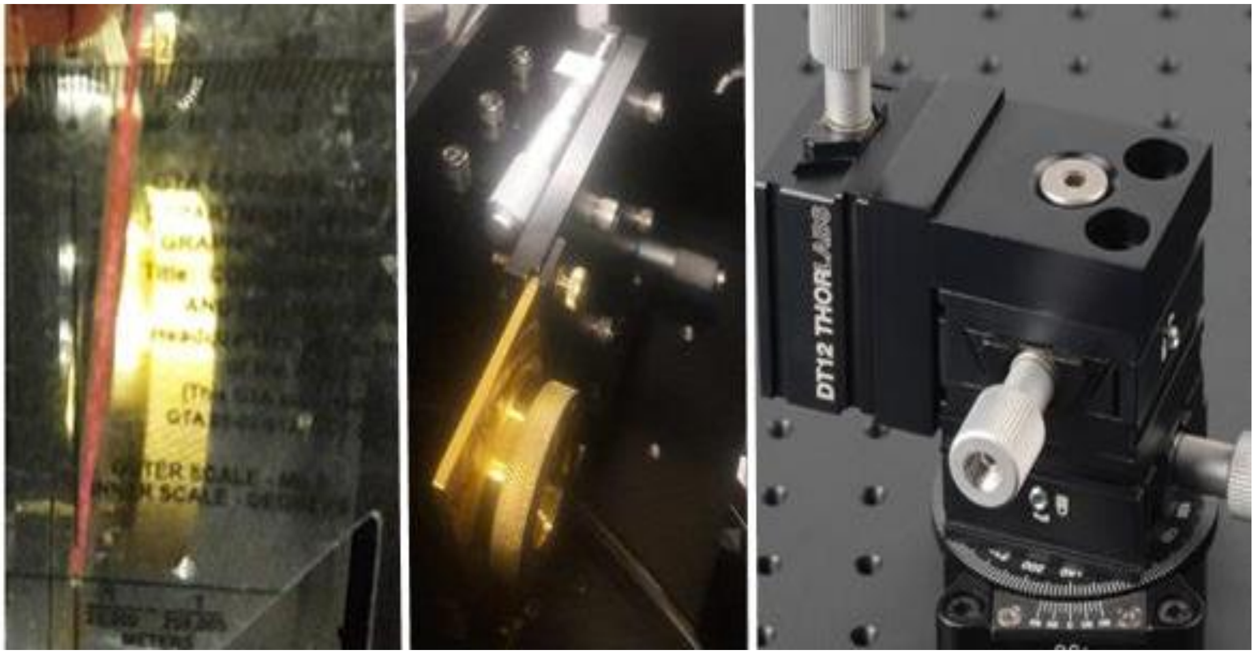


Figure 15: (left) *Ad hoc* adjustment to sample cell holder made by loosening the top sample holder screws and wedging a spacer in between sample and sample stage (center) to create an effective tilt angle of approximately 7°. Alternatively, a rotation stage (right) could be mounted to the sample holder and allow a greater range of tunability for the absorption spectra.

4.5 Conclusion and Future Directions

Previous studies have demonstrated that by adjusting the tilt angle of the sample, the cavity resonance could be tuned across the vibrational transition thereby allowing control of the polariton light-matter composition (see Figure 16). This work can be extended by performing polarization resolved measurements and performing 2D IR on different samples to track any emerging correlations between vibrational polaritons and uncoupled vibrations. Even if the interactions between the upper and lower polariton are out of the observable spectral window for the probe, observation of an upper and lower polariton cross peak can be made by stretching the pump frequency observation window by modifying the capabilities of the pulse shaper with the Phasetech software (see section 6.5) and generating a modified spectral plot with an extended pump frequency range. To fully observe both the upper and lower polariton bands in hBN, the current 2D IR experimental setup would require an additional OPA to generate light around 800 cm^{-1}. Fortunately, our group should be able to incorporate an additional OPA into the experimental setup in the future. Implementing this change is possible due to an excess of Solstice laser output and the modularity of the 2D IR pulse shaper cavity to allow conversion of the instrument to take place without unnecessarily hindering unrelated scientific experimentation.

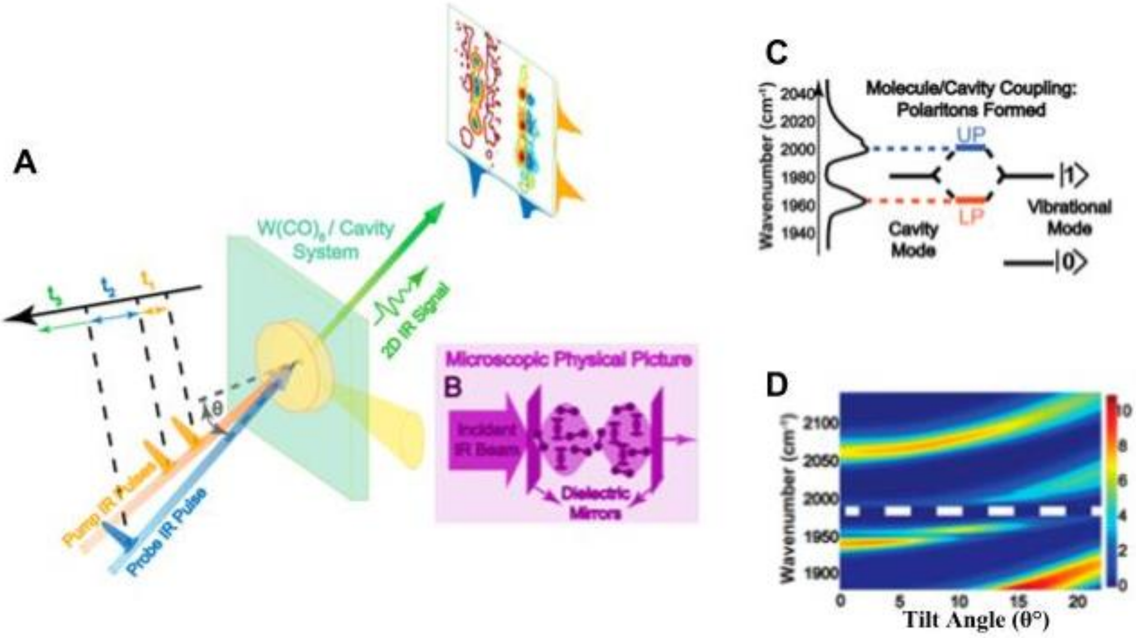


Figure 16: A) By adjusting the tilt angle, θ, the cavity resonance can be tuned across the vibrational transition which reveals a dispersive anti-crossing and allows control of the polariton light-matter composition (Figure D). Adapted from reference [55].

In addition, 2D materials can be conceivably fine-tuned by appropriately stacking the number of layers in a 2D material to create a van der Waals (vdW) material. By stacking layers appropriately, the properties of a vdW material can lead to the realization of expanded applications for 2D materials. While there are fabrication and economic challenges to overcome, sub-diffractional optical materials exploiting hyperbolic 2D materials such as hBN will likely result in improved optical components that are compact, tunable, and able to resolve the many interesting phenomena that occur in the infrared and terahertz regime. By continuing to develop an understanding of the role of dark states and how they interact with polariton states, it will be possible to develop better infrared photonic devices, improve laser cavity design and capabilities, and create better quantum mechanical and molecular modeling simulations [55]. Although there are very few published studies regarding hidden dark states and how they interact with bright polariton states, the application of 2D IR spectroscopy may be able to expand the few existing studies outside of the most basic systems and simulations. Although the task is daunting, the payoff of exploiting the vibrational-polariton interaction is too great not to warrant a continued investigation into this area of research due to the many exciting applications that can be realized with improved infrared and terahertz optical devices, advances in energy harvesting, and improved laser cavity design [57].

CHAPTER 5

Investigation of Influenza M2 Proton Channel

5.1 Introduction

The M2 proton channel of the influenza A virus is a popular drug target because it is necessary for the reproduction of the influenza virus. It is one of the smallest proton-selective channels found in nature. As a result, it is a model system for the study of selective, unidirectional proton transport across a membrane. While the structure has been probed using x-ray free electron laser diffraction techniques [59], there have thus far been no published studies on the Influenza A M2 proton channel examined using a method able to provide the structural and temporal resolution equal to that of 2D IR spectroscopy.

The time resolution of 2D IR spectroscopy, which employs approximately 50-60 femtosecond pulses, is more than capable of resolving protein dynamics. Because protein dynamics begin to occur on the order of 1 picosecond, 2D IR is uniquely suited to provide a method by which protein and peptide conformations can be observed in crystal structures to aid in refining molecular dynamics simulations and ultimately determine the functional mechanisms and configurations that operate in membrane proteins. Kratochvil, *et al.* investigated the KcsA potassium channel by applying $^{13}C^{18}O$ labels to the backbone carbonyls comprising the S1, S2, and S3 binding sites [58]. This labeling procedure enabled spectroscopic isolation of the ion binding sites to determine ion configurations in the K^+ ion channel selectivity filter. Kratochvil was able to rule out simulated configurations which were not observed in the spectra and piece together the fast evolution of structural distributions and multi-ion configurations through the combination of experimental data and MD simulations.

Our goal is the study of the 2D IR lineshapes of the M2 protein to look at the structure of water within the ion channel at different pH values. We want to perform this study while the M2 protein is in a crystal structure so that hydration can be decoupled from structural changes. It should be possible to gather information about the structural changes that occur as the H^+ ion channel is activated.

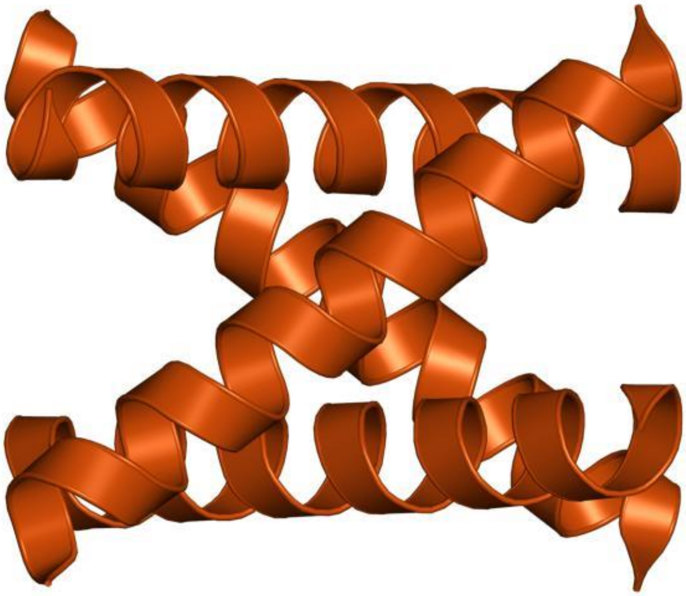


Figure 17: Closed-state structure of the M2 protein H^+ ion channel. Adapted from reference [59].

5.2 Results

Using the conditions outlined by DeGrado [59], two samples of matrix protein 2 from influenza A transmembrane (AM2TM) were prepared in a lipid cubic matrix at pH 8 and pH

5.5. These samples were observed under 2D IR to determine if a signal could be found in the Amide I mode for these proteins. Initial results were not promising, and a large amount of scattering was found throughout all the spectra for both samples. As a result, we attempted to suppress the amount of scattering by introducing quasi-crossed polarizer configuration using the method outlined by Helbing [60]. By using polarizers, the different fields in the pump-probe geometry can be independently controlled to allow enhanced anisotropic signals and suppression of scatter. This occurs because two independent measurements are carried out with identical scattering contributions but inverted 2D IR signals. By taking the difference between these readings, it is possible to suppress scattering measurements in data.

While the perpendicular polarization scheme did produce less congested spectra and did indeed suppress scatter, the spectra generated were still inconclusive in their ability to effectively generate a signal in the Amide I region of the spectra. We suspect that because the wavelengths of IR light used to interrogate this sample are of a similar size to the prepared crystals of the M2 protein, Mie scattering is occurring and distorting the 2D IR spectrum.

Several attempts were made to sonicate the samples, briefly, to determine if any signal could subsequently be found. Sonication must be performed carefully to avoid completely destroying the AM2TM crystals and limiting the usefulness of subsequent experimentation. There was no improvement to the 2D IR spectra after multiple sonication attempts. Our attempts at observing M2 crystals in a lipid matrix proved to be overambitious and we resolved to take a step back and attempt to produce samples more conducive to spectroscopic investigation.

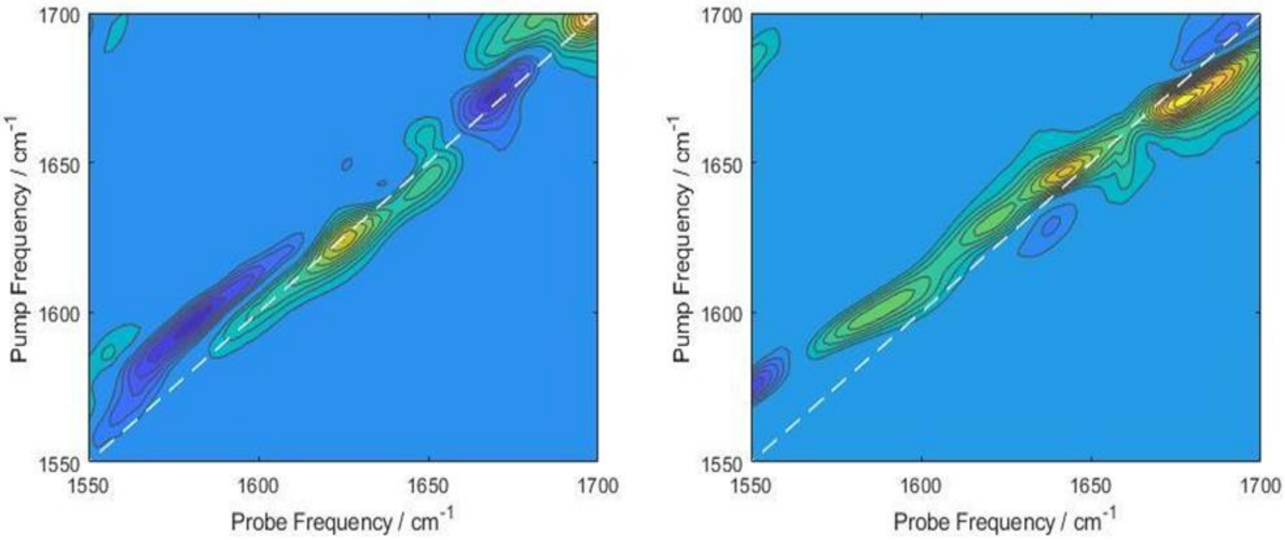


Figure 18: 2D IR spectra of M2 protein for influenza A at pH 5.5 (left) and pH 8.0 (right) using perpendicular polarization averaged over 10 scans. Characterization of both spectra was not possible due to a large amount of inherent Mie scattering. There is no clear signal from the amide I mode for any of the amide stretches present in the protein despite attempts to suppress scatter through perpendicular polarizer configuration.

A subsequent set of AM2TM samples were prepared in detergent as a stand-in for membrane-type environments with and without an isotope label on the G34 residue. The samples were prepared by DeGrado research group at UCSF and specific structural details are awaiting publication and Protein Data Bank (PDB) finalization (as of March 2019). Initial observations under 2D IR appear to be promising. The samples are well-resolved and do not show any appreciable scatter under parallel polarization. The unlabeled samples, however, appear to be practically identical to the G34 isotope-labeled samples under 2D IR. While the less-intense signal around 1610 cm^{-1} could potentially be characterized as an isotope label because it is approximately 50 cm^{-1} lower than the primary signal, this same signal was also detected on the unlabeled protein and is therefore more likely to be a sidechain interaction mode, arising from residues such as aspartic acid. One method to isolate any differences between two spectra is to subtract the normalized intensities to isolate the coupling due to the labeling of the G34 isotope. Difference spectra were generated for the entire sets of samples. Normalization of the intensity was applied by dividing the highest intensity throughout the entire array used to create the intensity for the pump and probe spectral plot. The contour count of the 2D IR spectrum was increased from 40 to 100 to provide stronger resolution for the difference spectrum. The resultant difference spectra shown in Figure 21 indicated slight differences between the two sets of samples. More testing will need to be done to determine the isotope labeling efficiency to account for the differences noted in the difference spectrum.

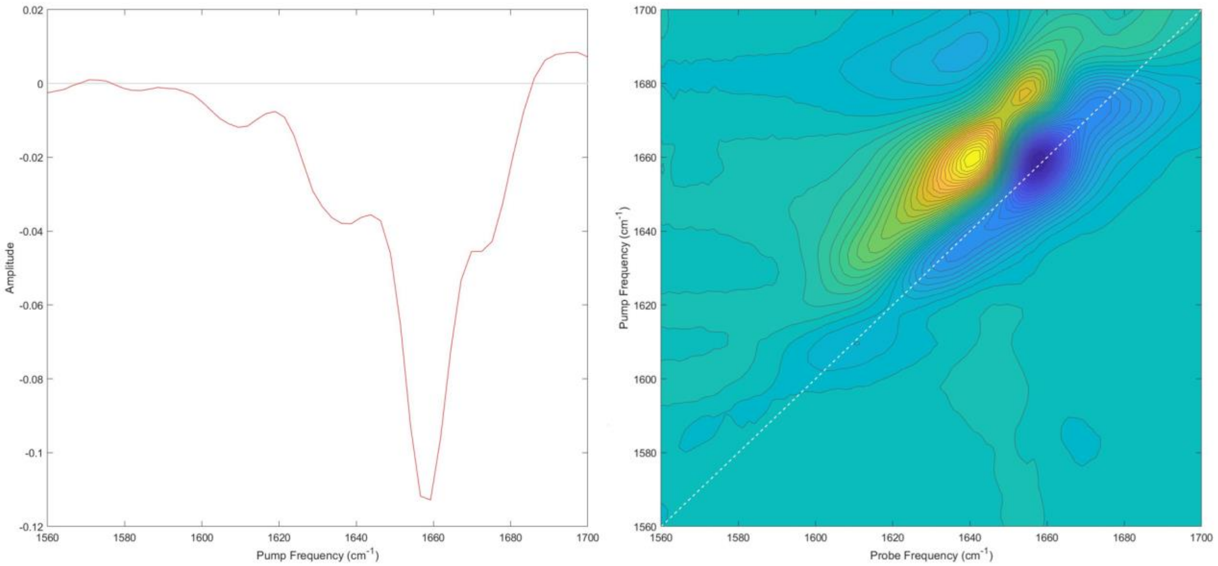


Figure 19: Diagonal slice and 2D IR spectrum of unlabeled, sample 1 of M2 protein for influenza A (lipid phase) 4.7mM concentration, pH 7.02, using parallel polarization averaged over 25 scans. Low intensity peak observed at 1610 cm^{-1}, most intense peak observed at 1658 cm^{-1}, and peak shoulder observed at 1673 cm^{-1}.

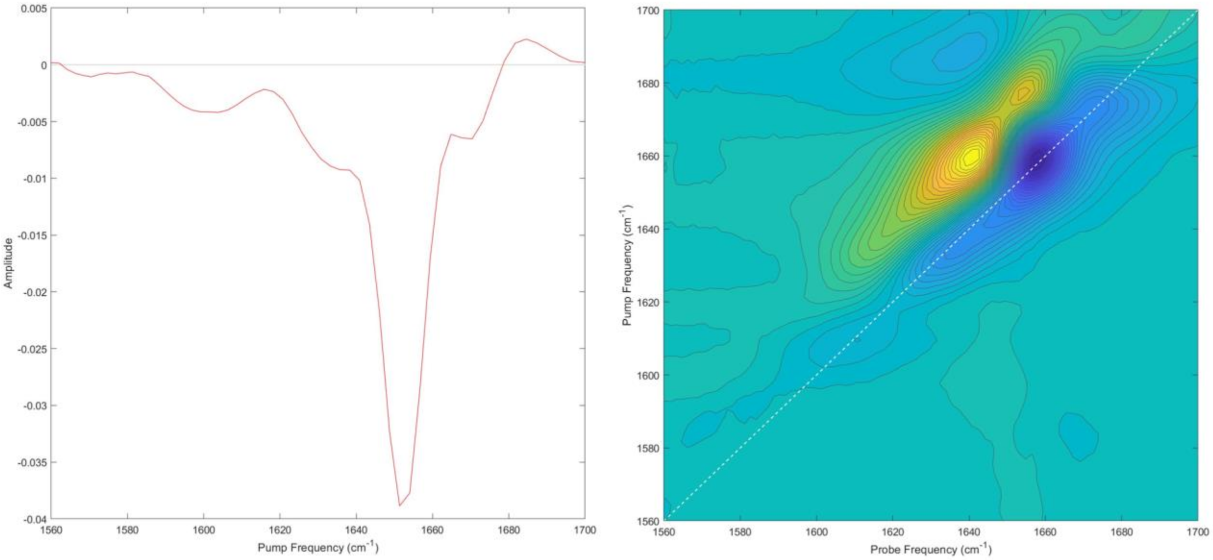


Figure 20: Diagonal slice and 2D IR spectrum of G34 labeled, sample 1 of M2 protein for influenza A (lipid phase) 4mM concentration, pH 7.02, using perpendicular polarization averaged over 179 scans. Low intensity peak observed at 1608 cm^{-1}, most intense peak observed at 1654 cm^{-1}, and peak shoulder observed at 1672 cm^{-1}.

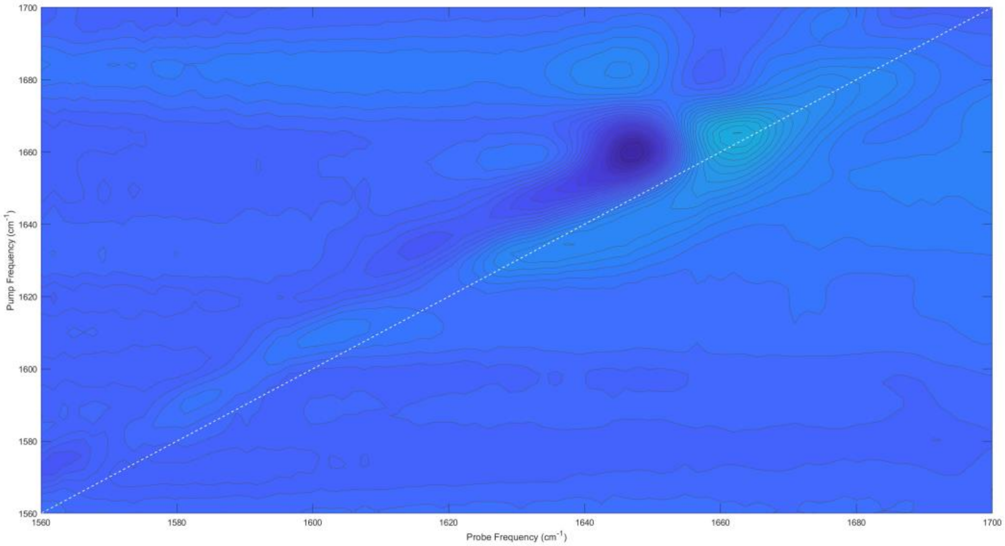


Figure 21: 2D IR difference spectrum of G34 labeled and unlabeled M2 protein for influenza A (lipid phase), sample 1, 4mM concentration, pH 7.02. 100 contours used to increase relative intensity. Negative peaks observed at 1605 cm^{-1}, 1636 cm^{-1}, and 1664 cm^{-1}.

5.3 Conclusion and Future Directions

In summary, over 93 spectra were taken of crystalline AM2TM protein of influenza A at pH 5.5 and pH 8.0 using both parallel and perpendicular configuration of polarization. Figure 18 shows two examples of the spectra. Although initial studies of the M2 protein were not promising, continued collaboration and investigation into this ion channel resumed with the subsequent examination of isotope-labeled and unlabeled configurations of the protein which were not in a crystalline state [61]. Initial results were promising in their ability to produce clear spectra and difference spectra between labeled and unlabeled spectra. Further investigation is necessary to obtain residue-specific 2D IR data to answer important questions regarding the structure of the M2 channel as it relates to point mutations which can alter the effectiveness of popular antiviral drugs such as amantadine. While the ultimate goal is to examine M2 in its crystalline state in order to decouple hydration effects from structural changes, our current course of action is to continue creating new sample conditions and ultimately progress back up to a crystalline conformation as we continue to gain insight into the structure of water within the ion channel. The variation of conditions will lead to an enhanced understanding of the overall structural evolution and function of the M2 protein.

Chapter 6

Experimental Methods

6.1 Introduction

During my time in the Buchanan group, I had the privilege and opportunity to learn a wide range of experimental techniques and enter the group at a point when methods and instrumentation procedures were beginning to be tested and established. I feel it is my professional obligation to capture all lessons learned during the group's preliminary years to help establish comprehensive systems and processes. Statistical process control and good stewardship over our resources not only builds upon our professional reputation as scientists, it can accelerate the learning curve and offer useful tactics, techniques, and procedures to future group members. I focused my efforts primarily on understanding and developing systematic methods to optimize spectroscopic calibration, maximize signal, and suppress scattering. Methods not covered in this chapter will be covered in the laser laboratory standard operating procedures document (SOP).

Preliminary research efforts involving 2D IR can be achieved in a much more efficient manner by developing a rigorous and intuitive understanding of how to better manipulate the 2D IR experimental setup through the software and optimization of all of the optics in the cavity to achieve maximum signal and ideal calibration. While there can be no substitute for hands-on training and experience working with optics, the use of SOPs and understanding standard operating conditions can accelerate experience gain. I will describe all the methods I have developed over my brief graduate career to perform peptide synthesis, characterization, and spectroscopic investigation.

6.2 Microwave-assisted Fmoc Solid-phase Peptide Synthesis

Microwave-assisted peptide synthesis is a cost-effective and reliable method of synthesizing peptides. Solid phase peptide synthesis was originally developed in 1963 by Bruce Merrifield [62] and has become the predominant technique for synthesizing peptides under approximately 50 residues. The use of microwave radiation accelerates reaction rates and allows for the automation of the peptide synthesis process through the use of a Liberty Blue automated peptide synthesis instrument. Solid-phase peptide synthesis involves the growth of a peptide chain from a solid support, usually a polystyrene (PS) bead. This support allows for the easy removal of reagents at each step of the synthesis process. While biological protein synthesis occurs from the N-terminus to C-terminus, solid phase peptide synthesis occurs from C-terminus to N-terminus. The amino acids are added to the chain sequentially by reacting the free amine of the chain with the free carboxylic acid of the subsequent amino acid in the series. To prevent undesired reactions, the amine of the subsequent amino acid is masked with a protecting group which consists of 9-fluoromethyloxycarbonyl (Fmoc). Because Fmoc is base-labile, the R-groups of the amino acids are masked with acid-labile (base-stable) protecting groups which prevent unwanted side reactions from occurring with the functional groups of the amino acid side chains during synthesis. Upon completion of Fmoc synthesis, deprotection of the amino acid sidechains and removal of the finished peptide from the resin can be completed using trifluoroacetic acid (TFA), as described in section 6.4.

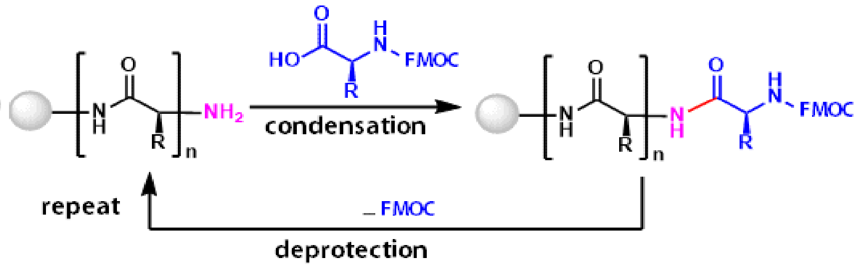


Figure 22: Fmoc solid-phase peptide synthesis reaction scheme. A Fmoc-protected PS resin is used. Deprotection occurs on the Fmoc group prior to the activation and coupling of the subsequent amino acid. This reaction repeats until the peptide is constructed from C-terminus to N-terminus.

N-terminal acetylation is an essential post-synthesis modification that can be added to the peptide synthesis procedure by applying the method originally established by Vogel in 2010 [63]. N-terminal acetylation is performed by reacting the 0.1 M scale finished peptide with 10% acetic anhydride in DMF.

6.3 Isotope Labeling of amino acids

Because the amide I vibrational mode is primarily a result of the carbonyl stretching mode (C=O), it is advantageous to incorporate isotope labeling using a ^{13}C and ^{18}O label. As previously discussed, the incorporation of a $^{13}C^{18}O$ carbonyl shifts the frequency approximately 55 cm^{-1} lower into a relatively empty region of the spectrum and diminishes the expectation of a natural occurrence of a $^{13}C^{18}O$ carbonyl down to approximately 0.002%.

While not impossible [64], it is generally only feasible from a financial standpoint to incorporate isotope labeling on amino acids with non-reactive hydrophobic side chains (alanine, glycine, isoleucine, leucine, phenylalanine, and valine). The reactive side chains of other amino acids must be protected during synthesis by purchasing amino acids with protecting groups already incorporated. Because side-chain protecting groups are acid-labile, they would be removed by the isotope labeling procedure which relies on acid-catalyzed chemistry.

^{13}C-C_{α} labeled amino acids can be purchased with or without Fmoc protection, but it is generally preferred to purchase the amino acids with Fmoc protection already incorporated since the experimental conditions for incorporating Fmoc labeling can decrease isotope labeling efficiency. Because the Fmoc group is base-labile, it will not be affected by the ^{18}O labeling procedure.

When necessary, the general reaction scheme for adding Fmoc protection is shown in Figure 23 below. A 1:1:1 ratio of amino acid: Fmoc-OSu:$NaHCO_3$ is added and sealed in a reaction vessel and allowed to mix and react overnight in a fume hood. On a 10 mmol scale which uses approximately 1 gram of amino acid, approximately 30 mL of H_2O:acetone is used. A 2 M $KHSO_4$ solution is then added the following day until the pH reaches approximately 2.0. At this point, the amino acid should precipitate. The peptide is then filtered, washed sparingly with H_2O, and then lyophilized.

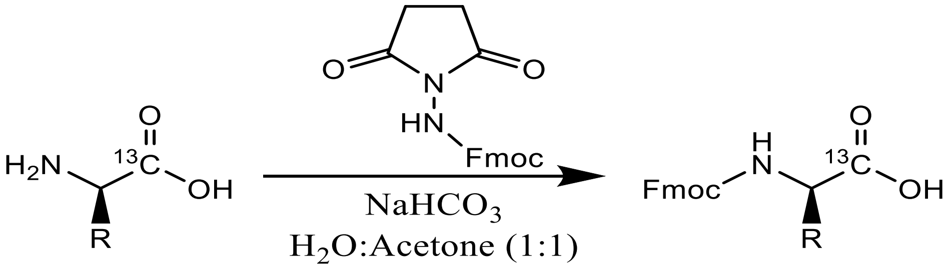


Figure 23: Reaction scheme for adding Fmoc protection to a ^{13}C-C_{α} labeled amino acid.

Figure 24 demonstrates the general reaction scheme for the isotope labeling procedure. The reaction is carried out using a Schlenk line (vacuum gas manifold) to maximize labeling efficiency. The Schlenk line is held at vacuum pressure and flame dried until the pressure reads below 50 mTorr. A nitrogen atmosphere is maintained throughout the reaction. For 1 gram of ^{13}C amino acid, 12 mL of 4 M HCl in dioxane is added to the reaction vessel. The reagents are refluxed at 150° C for 4 hours. The Schlenk line removes the dioxane until some precipitant is visible, but it is still kept very "wet". Drying is continued by finishing on a lyophilizer. The reaction can be repeated as necessary and final cleaning can be performed by washing the final product with a cold 3:1 mixture of ethyl acetate and hexane followed by lyophilization. The product is often precipitated with 8:1 hexane and ethyl acetate, and lyophilized. Desalting can also be performed as necessary.

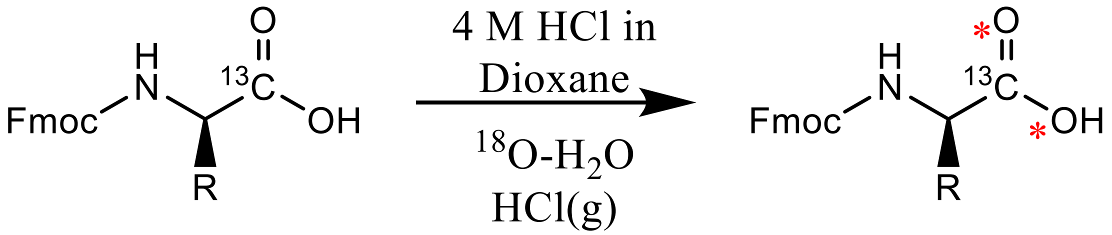


Figure 24: Reaction scheme for ^{18}O isotope labeling of Fmoc-^{13}C amino acid, where the red asterisks indicate isotope labeling sites.

The labeling efficiency can be checked with mass spectrometry. The labeling reaction can be repeated with fresh ^{18}O water until the desired efficiency is achieved. The mass to charge ratio measured in mass spectrometry should contain 3 peaks, representing fully unlabeled ($^{13}C^{16}O^{16}$)), half labeled ($^{13}C^{16}O^{18}O$), and fully labeled ($^{13}C^{18}O^{18}O$)

isotopes. The two oxygen atoms in the carboxyl group of an amino acid are chemically equivalent due to resonance. As a result, both oxygen atoms must undergo exchange to ^{18}O or the labeling efficiency in the synthesized peptide will be reduced. A minimum standard of 90% ^{18}O exchange is necessary for the selected residue or the isotope label will not be able to adequately couple in β-sheets [14]. The application of Equation 4 can easily calculate the exchange percentage. Our first attempt at performing an isotope exchange with the procedure above yielded an isotope labeling efficiency of 95.4 % on $^{13}C^{18}O$-phenylalanine. An electrospray ionization mass spectrometry (ESI-MS) instrument was used to perform Mass Spectrometry. After confirming the labeling efficiency, the isotope-labeled amino acid can be placed on the peptide synthesizer and inserted into the peptide sequence where desired.

$$Label\ \% = \frac{Height(^{13}C^{18}O^{18}O) + \frac{1}{2}Height\ (^{13}C^{18}O^{16}O)}{Height\ (^{13}C^{18}O^{18}O) + Height\left(^{13}C^{18}O^{16}O\right) + Height\ (^{13}C^{16}O^{16}O)}$$

[Equation 4]

6.4 Purification and Characterization of Synthesized Peptides

An illustration depicting the general process flow diagram for these purification and characterization procedures is shown below in Figure 25. Upon completion of peptide synthesis, the peptide resin is transferred from its reaction vessel into a centrifuge tube using dichloromethane to assist in the transfer between vessels. The resin is then dried by blowing off the dichloromethane with an appropriate flow of air or nitrogen. The resin can be safely stored in a freezer at -80° C until further use is necessary. Storing the peptide in this condition with the resin still attached effectively renders the peptide inert and improves the lifetime due to all of the reactive side chains retaining their protecting groups [14].

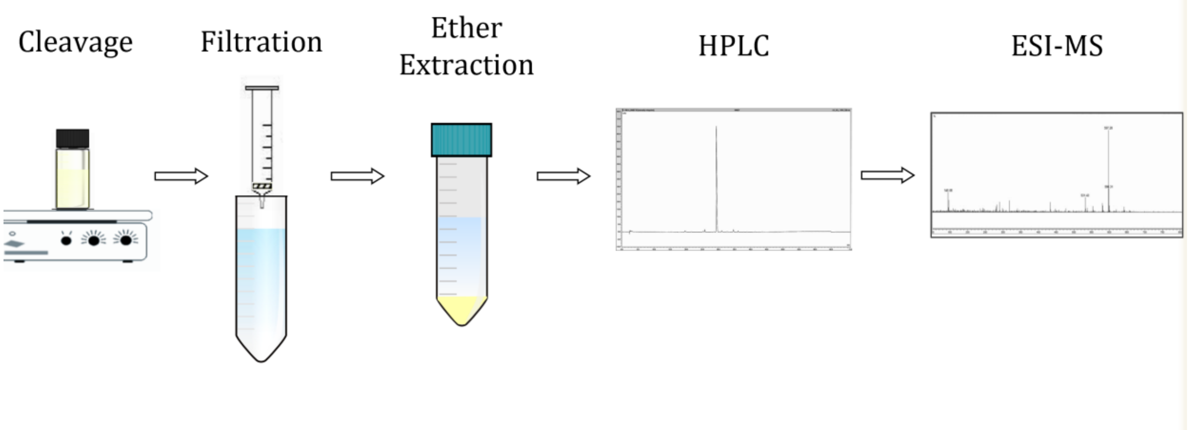


Figure 25: General process flow diagram for peptide purification and characterization following synthesis.

For a reaction at the 0.1 mmol scale, typically one-third of the resin is cleaved at a time. This amount should yield about 10 mg of peptide for A_3K and proportionately more for peptides of larger relative molecular weight than A_3K. For this amount of resin, 10 mL of cleavage cocktail is used. Although several different cleavage cocktails formulations have been tested, we have found that the cocktail giving the best yield for amphiphilic peptides consists of 90% trifluoroacetic acid (TFA), 5% ethanedithiol, 2.5% thioanisole, and 2.5% anisole. The primary constituent of the cocktail, TFA, removes all acid-labile side chain protecting groups and cleaves the peptide from the solid resin. Ethanedithiol, thioanisole, and anisole scavenge the mixture and bind the side chain protecting groups that have been removed from the peptide and prevent them from reattaching. Because of the noxious odor of the thiol groups contained in the cocktail, the cleavage is performed in a fume hood and anything that touches or interacts with the sample is submerged in bleach to help neutralize the odor.

For the cleavage, a 6-dram (22 mL) screw-cap glass vial is placed on a stirring plate and 300-500 mg of resin is added along with 5 mL of the prepared cleavage cocktail. The cleavage should proceed for three hours while gently swirling the vessel with a magnetic bar. It is important to not allow the cleavage reaction to go beyond 3 hours as this can lead to side chain reactions, particularly for samples containing cysteine or tryptophan side chains. As the reaction progresses, the solution in the vial will change color to a brownish-yellow or a bright yellow. This indicates the presence of scavenged protecting groups.

Upon completion of the cleavage reaction, the solution is then poured into a syringe which contains a 0.22-micron filter that has been punched out of a thick piece of filter paper and placed in the bottom of the syringe. The cleavage mixture is plunged through the syringe and into a centrifuge tube. The remaining 5 mL of cleavage cocktail is used to wash any remaining particulate out of the original screw-top glass vial and then discharge through the syringe again to rinse any remaining peptide through the resin-impregnated filter paper located at the bottom of the syringe. At this point, the solid resin should remain in the syringe and the peptide solution will now be in the centrifuge tube. The peptide solution is then blown dry with nitrogen and carefully heated using either a heat gun or water bath until the peptide is dry. It is important to exercise caution if utilizing a heat gun as this can quickly scorch the peptide or the centrifuge tube if the process is not monitored under steadfast observation.

As the peptide begins to dry, it is important to throttle the air down accordingly to preserve the yield of the peptide. It is not necessary to completely desiccate the peptide, but it should take on the appearance of a film once the process is complete. While the peptide dries, an aliquot of diethyl ether should be prepared by chilling it in the freezer for at least 30 minutes. If the peptide sample is in a 50 mL centrifuge tube, 15-20 mL of chilled ether should be sufficient in adding to the peptide. The peptide will precipitate out of the ether and any remaining scavengers and organic molecules will remain soluble. A spatula should be used to scrape along the walls of the centrifuge tube to ensure the peptide is sufficiently suspended within the ether. The centrifuge tube should then be capped and allowed to remain in an ice bath for approximately 1 hour. The cold ether-peptide mixture should then be centrifuged for 5 min at 5000 rpm. If possible, the centrifuge should be chilled and set to 0° C in the 1-hour preceding the centrifugation. The ether extraction process should be repeated as often as desired or deemed necessary. Three to four ether extractions are typical for most of our syntheses, and the peptide should appear colorless when complete. If desired, the molecular weight of the peptide can be checked using ESI-MS, but this is generally not good for the instrument as any remaining salts in the sample can potentially contaminate the mass spectrometry instrumentation and create a need for unscheduled maintenance on the electrospray tip and ion chamber.

The peptide is then purified using C18 reversed-phase high-performance liquid chromatography (RP-HPLC). We use a Waters XBridge C18 5µm, 250 x 4.6 mm analytical column. Having determined that scientific literature regarding specific methods for HPLC purification of amphiphilic peptides has a tendency to be either vague [64] or not-at-all mentioned, we have developed a method after extensive trial and error. Because peptide batches tend to be rather large and the use of autosampler capabilities can save on the order of 80 hours of otherwise tedious collection labor, it is important to develop a processing method that produces consistent and accurate results. It is important to wash the column before and after HPLC runs to prevent any buildup or previous aggregation from contaminating the purification process. Typically, both 220 nm and 280 nm wavelengths are monitored during elution to detect a peptide signal; the 220 nm wavelength monitors absorption due to the peptide backbone and the 280 nm wavelength monitors absorption resulting from residues with aromatic sidechains. As the amphiphilic peptides that we have tested thus far lack any aromatic rings, the 220 nm signal has been the primary means to detect and collect our peptide samples.

We start by dissolving approximately 2 mg of peptide per 1 mL of a 50:50 mixture of HPLC-grade methanol and water. The HPLC gradient is started from 0-3 minutes of 100% of H_2O with .045% TFA. From 3-5 minutes, the initial H_2O with .045% TFA is linearly decreased from 100% to 95% while a linear gradient is increased from 0-5% H_2O with 0.1% Acetonitrile (ACN). From 5-16 minutes, the 95% H_2O with .045% TFA is linearly decreased down to 40% and the 5% H_2O with 0.1% ACN is linearly increased to 60%. The gradient is held static from 16-18 minutes and the run finishes at 18 minutes. The flow rate is held constant throughout the HPLC run at 1.333 mL/min. The typical injection volume is 400 µL. We have found success at a range from 300 – 500 µL but generally decided that the best balance of throughput, signal, cleanest peaks, minimal use of solvent, and lowest amount of post-operational maintenance is achieved at an injection volume of 400 µL. It is essential to keep injection volumes consistent to ensure consistent results.

It is recommended that an initial run is collected for every batch of peptide and analyzed with ESI-MS in the event that the collected peak is contaminated or otherwise incorrectly characterized during the HPLC elution. Figure 26 demonstrates an example of an ESI-MS spectrum of purified acetylated A_3K. The highest peak appears at 402 Daltons, which indicates that the peptide is purified. It is possible for H^+, Na^+, or K^+ ions to bind with the peptide as well as any combination of the aforementioned ions in multiple charged states; the additional peaks generated as a result of this are not impurities. Because mass spectrometers are often public instruments, the instrumentation should be tested with a blank sample to determine whether impurities are present as well as assessed using a known calibration standard.

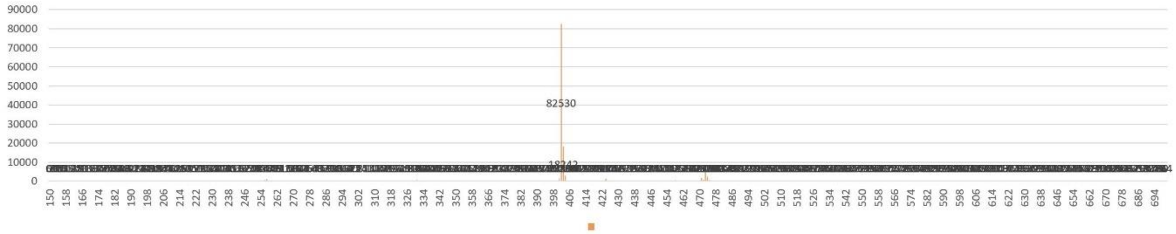


Figure 26: Excel histogram of an ESI-MS spectrum for N-Acetylated A_3K. Nominal mass values are recorded over a scan range from 150-700 Daltons.

It generally takes about 24 hours collect approximately 20 mg of peptide automatically. Higher throughputs can be achieved with a bigger column, but the lower throughput of an analytical column allows for an easier user experience to achieve narrow gradients for subsequent runs and is particularly suited to the task of achieving the highest possible yield for samples that are particularly expensive. One bad HPLC run on a larger preparatory column can result in a substantial loss of peptide. Figure 27, shown below, demonstrates our peptide signal with various amphiphilic peptides. As noted in the figure, the signal is consistent for various amphiphilic peptides and reliably generates the same signal at the same retention time for all subsequent runs.

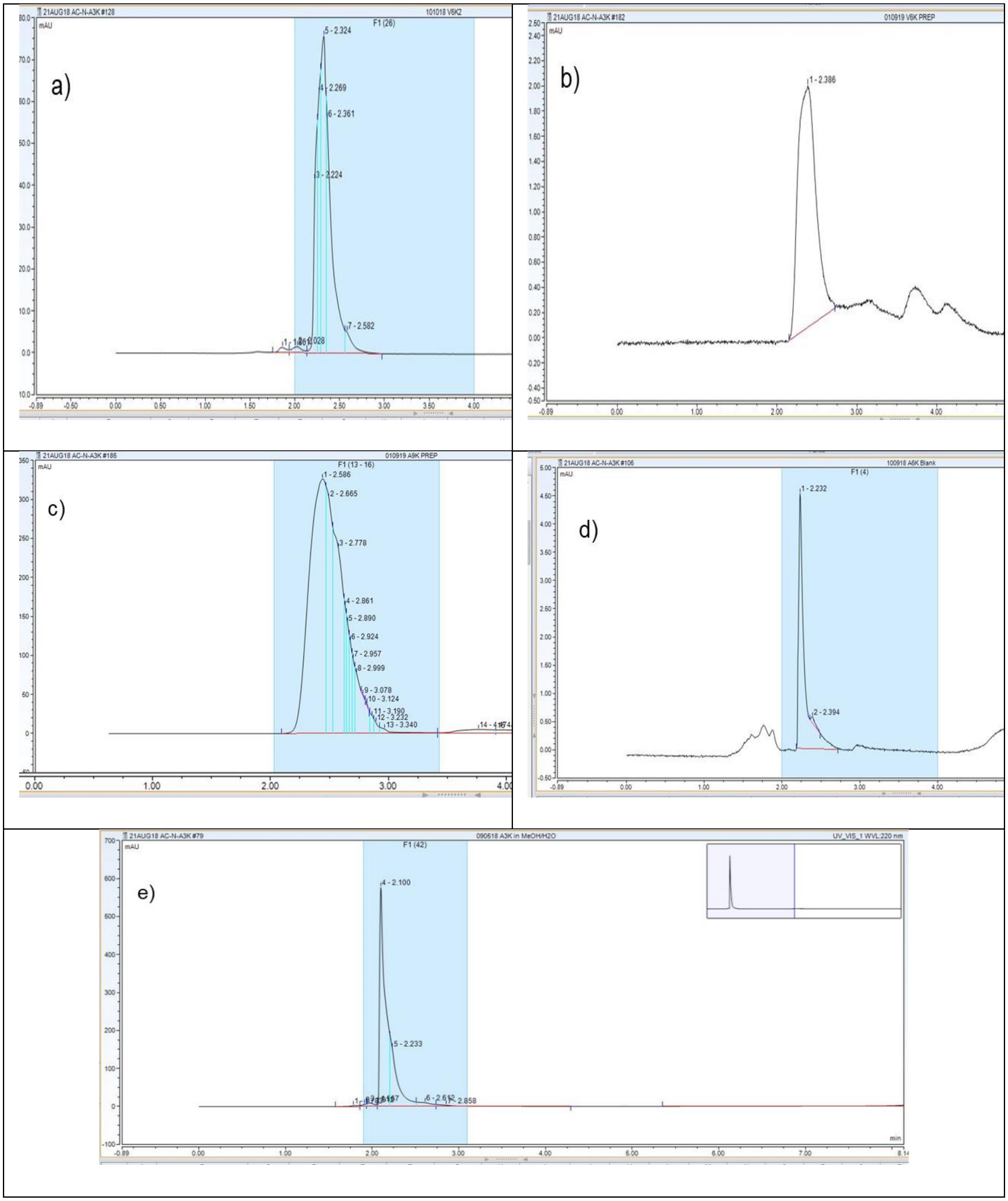


Figure 27: RP-HPLC chromatograms at 220nm of a) V_6K_2, b) V_6K, c) A_9K, d) A_6K, and e) A_3K. All amphiphilic peptides thus far demonstrated a signal of duration 1 minute or less starting around the 2-minute mark. This allows for a simple and efficient fraction collection.

Upon completion of HPLC and confirmation of purification through the use of mass spectrometry, the peptide samples from fraction collection are consolidated and freeze-dried using a lyophilizer. The peptide will likely contain trace amounts of TFA as a result of HPLC as well as the cleavage cocktail. TFA is visible in the mid-IR and gives a very strong absorption signal around 1673 cm^{-1}. This can interfere strongly in the characterization of α helices as well as β sheets. As a result, it is necessary to perform a counterion exchange, replacing the trifluoroacetate salt with the chloride anion. Because a peptide will tend to degrade around a pH of 1.0, it is important to use a concentration of HCl of no more than 0.1 M. The consolidated peptide is dissolved in 0.1 M HCl in water, allowed to exchange for 5 minutes, and then lyophilized. This process is repeated several times to ensure that all traces of TFA are removed, thus assuring a clean spectrum.

6.5 Preparation of Samples For 2D IR Spectroscopic Measurement

Upon purification, the peptide is typically divided into aliquots and stored as a dried powder in a -20° C freezer. Although it is difficult to subdivide and precisely measure, the most appropriate aliquot sample size tends to be on the order of 1 mg. When ready for an experiment, an aliquot is then dissolved in deuterated hexafluoroisopropanol (D-HFIP) to an approximate concentration of 4 µM. HFIP is an ideal solvent because it prevents aggregation and readily evaporates. Deuteration of this solvent exchanges the amide hydrogen atoms with deuterium and shifts the amide II mode (N-H bending) sufficiently that it will not compete with isotope-labeled amide I signals. The peptide should be allowed to dissolve completely in D-HFIP by sonicating and heating the sealed sample to around 40° C for approximately 4-6 hours. It is important to check the temperature while sonication is occurring to verify that the mechanical action of sonication is not heating the sample above 40° C and potentially degrading the peptide. Although it has never been our experience, if the peptide does not appear dissolved after 6 hours of sonication, it may be necessary to add more D-HFIP or wash with D_2O. The dissolved peptide should be clear. Any appearance of cloudiness or white chunks indicates that the peptide is not completely dissolved. Once the sample is dissolved, the D-HFIP should be removed by lyophilization overnight and then dissolved in D_2O.

A 2D IR experiment can generally be performed with approximately 2 µL of sample. Each sample is placed upon a 2 cm diameter cuvette window with a .025" thick Teflon ring spacer placed between another cuvette window. The cuvette windows are fastened together mechanically through screw threading in the sample cell holder. O-rings between the sample cell holder minimize exposure to the environment and keep the cuvette windows from becoming damaged. Every amphiphilic peptide, however, behaves slightly different within cuvette windows and the relatively low surface tension of an amphiphilic peptide in the presence of D_2O will likely create many instances in which the sample is lost due to the sample enigmatically moving outside of the path of the pump and probe beam during experimentation. The sample shift may also absorb onto the Teflon spacer as soon as the cuvette windows are tightened together from the mechanical action of the sample holder. It is important to plan on having a reserve of sample as a contingency to ensure scientific experimentation can carry on to completion. Once the amount of peptide to be used is determined, a concentration of peptide must also be chosen. Although we have been able to achieve a signal at concentrations as low as 10 mM, we have had the most success at a concentration of 25 mM as it generates a more intense spectrum. Concentrations at 50 mM have been attempted and work exceptionally well, but we have found that this is not the most economical use of a peptide, particularly in the case of isotope-labeled peptides.

6.6 Optimization of Laser and 2D IR Spectroscopic Signal

Our current experimental setup (see Figure 28) includes a laser system which consists of a commercial regenerative amplifier (Spitfire Ace by Spectra Physics) which is seeded by a Ti:Sapphire oscillator (Mai Tai by Spectra Physics). It also includes an optical parametric amplifier (TOPAS Prime by Light Conversion) with a difference frequency generation (DFG) attachment. Mid-IR light generated by the TOPAS then passes through the 2D IR spectrometer (2DQuick IR by PhaseTech) consisting of an arbitrary waveform generator (AWG), an acousto-optical modulator (AOM), a monochromator, and a HgCdTe (MCT) detector array which is kept chilled near liquid nitrogen temperature 76.8° C for enhanced sensitivity.

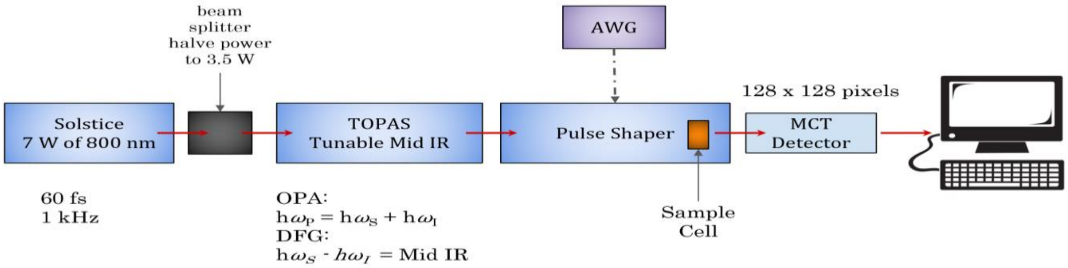


Figure 28: Simplified experimental setup of the laser system, electronics, and spectrometer.

The unmatched capabilities of 2D IR spectroscopy with respect to time resolution require precise temporal and spatial alignment. With our experimental setup allowing femtosecond pulses, structural fluctuations in protein structure and the local environment can be measured at a resolution of slightly less than 100 femtoseconds. Pulse shaping allows for easier alignment of the 2D IR experiment, much faster data collection, and the ability to rotate the individual phases of the pulses to reduce scatter.

Figure 29 shows a schematic diagram of our 2D IR spectrometer. A 10% beam splitter creates a distinct pump and probe path of beam diameter 0.4 mm (using the $1/e^2$ standard) which ultimately intersects as the pump and probe overlap inside the sample cell. The pump and probe beams are vertically aligned, parallel, and horizontally separated by approximately 2 cm as they approach the first parabolic mirror. A parabolic mirror focuses and spatially overlaps the beans onto the sample while the Phasetech software and delay stages can provide compensatory means to achieve temporal overlap. The position of the sample can be manipulated using a manual three-axis translation stage. Translation of the sample in the plane perpendicular to the direction of the laser pulses can allow the user to achieve the least amount of scattering and the strongest signal by choosing the best possible spot on the sample. This is typically done by translating the sample stage while observing the instantaneous 1D signal acquisition and attempting to minimize scattering while also maximizing signal in the spectral region of interest. Use of the translation stage also allows a small sample spot size, which is particularly important for preserving limited amounts of expensive isotope-labeled samples.

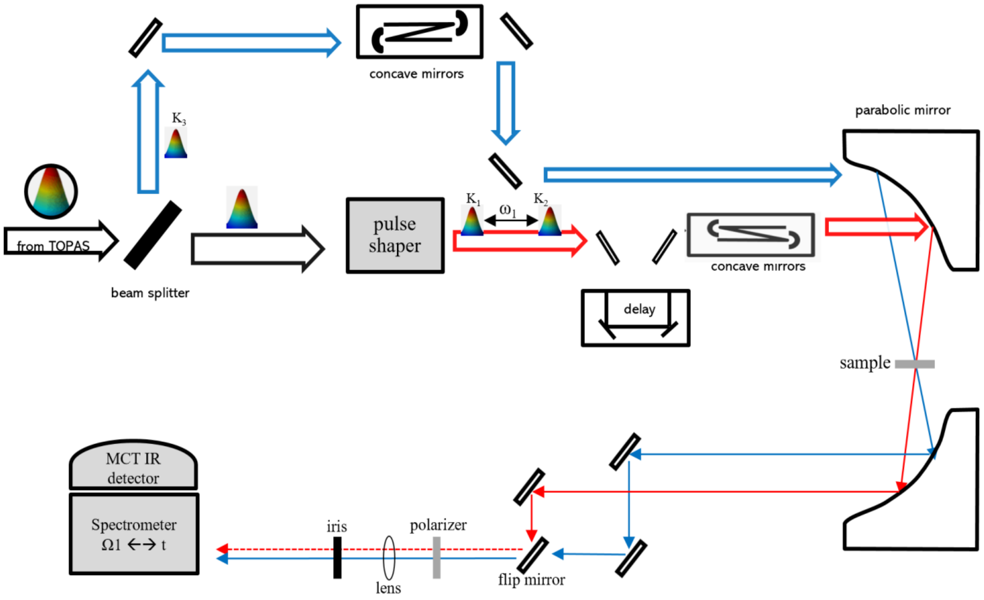


Figure 29: Schematic diagram of our 2D IR spectrometer design. A beam splitter separates the pump and probe pulses. The pulse shaper creates double pump pulses which allow 2D IR measurements. Time delay stages and the ability to translate the sample in the x, y, and z- planes enable tuning of spatial and temporal overlap of the pump and probe pulses at the position of the sample. The pump beam is generally dumped, but a flip mirror allows the detection of the pump beam for troubleshooting and calibration.

Delay stages and computer software are used to account for any delays in the laser trigger relative to the AWG trigger and then create an RF mask which is sent to the AOM to generate an acoustic wave used to shape appropriate mid-IR pulses. Both the laser trigger and AWG trigger are delayed so that the laser pulse arrives at the AOM coincident with the mask. Because the gratings in the shaper map the frequency of a pulse to the spatial dimension of the AOM, the propagation velocity of the acoustic wave in the AOM is also able to track the time-dependence of the RF mask produced by the AWG to the spatial dimension of the AOM. This allows a direct scaling between the different times in the RF mask and all of the pulse frequencies and can be used for frequency calibration of the shaper. The frequency calibration can be used to assign specific frequencies to the corresponding pixels in a 128x128 pixel MCT detector array. While frequency calibration is not required for daily and routine use of 2D IR spectroscopy, a fundamental understanding of how this calibration is performed within the software as well as the underlying mathematics is key to unlocking the full potential of the pulse shaper.

Because the position of a single-peak mask and the interval between multiple peaks are all known, the position of all peaks within the spectral window can be calculated and fit to a second-order polynomial which relates frequency and time in accordance with Equation 5. The frequency-time relationship can also be determined based on the geometry of the shaper in accordance with Equation 5. Equation 5 can also be used to theoretically determine the focal length and grating dispersion that is ideal for a specified application.

$$\nu = p_2(f_s t)^2 + p_1 f_s t + p_0$$

[Equation 5]

Equation 5: Second-order polynomial relationship between time and frequency where p_i are the fitting parameters, f_s is the sampling frequency of the AWG, t is time, and ν is frequency.

$$\nu = \left[\frac{1}{2\nu_0} + \frac{1}{Dc} \sin\left(\sin^{-1} \frac{Dc}{2\nu_0} - \tan^{-1} \frac{\nu_{ac}}{f}(t_0 - t) \right) \right]^{-1}$$

[Equation 6]

Equation 6: Relationship between frequency and the geometry of the shaper where ν_o and t_o are the frequency and time of an individual peak, D is the grating period, c is the speed of light, f is the focal length of the cylindrical mirror, and ν_{ac} is the acoustic wave velocity.

Outside of routine maintenance and signal confirmation, it is important to establish calibration parameters of the pulse shaper if the region of the spectrum under investigation is outside the spectral ranges of what is typically used in the investigation of the Amide I region. Once calibration parameters are established, it is generally sufficient to establish a known frequency using a single point. Any significant tuning or maintenance of the laser resulting in a realignment of optics can decrease the accuracy of the fitting parameters in the calibration and create potential distortion or attenuation of the signal in the spectrum. It is therefore necessary to verify and reestablish calibration parameters regularly to ensure maximum efficiency. Generally, we have found it sufficient to perform shaper calibration monthly.

It is possible to compensate for spectral dispersion in 2D IR measurements and therefore enhance the signal. The AOM contains a Germanium crystal which has a high index of refraction. The spectral phase of the pulse is altered by the transmission through the AOM due to the frequency dependence of the index of refraction. This frequency dependence can be described by a Taylor series expansion [66]. While the zero and first order terms do not affect the temporal profile of the pulse, the higher order terms do bear significance. Expansion terms higher than third-order, however, are negligible and can be disregarded. The second-order or group velocity dispersion (GVD) and the third-order dispersion (TOD) can be compensated by utilizing the AOM to apply a phase mask to the pulse. While it is possible to obtain a calculation of the GVD and TOD given the wavelength of the center of the pulse and the thickness of the Germanium crystal and the full width at half maximum (FWHM) of the input pulse, an exact calculation of the optimal TOD and GVD dispersion constants can become intractable because the final pulse is also dependent upon the initial dispersion and the alignment going into the pulse shaper. While a GVD and TOD calibration protocol is present in the software, we have found that these calculated settings are not as precise as that which can be achieved by trial and error and bounding local maxima by observing the 1D IR signal acquisition. Changing the GVD and TOD manually and then updating the AWG and masks allows the user to observe the signal present in the spectral acquisition and use it as a feedback mechanism to adjust by bounding and constraining the local maximum values for GVD and TOD constants based upon the symmetry and relative peak height of the fundamental and overtone signal.

The pump and the probe beams must achieve spatial and temporal overlap at the sample. Spatial overlap is best achieved by centering a 100 µm precision aperture of equal dimensions to the sample cuvettes inside the sample cell. The probe beam is typically spatially aligned first. Although not absolutely necessary, the pump beam is blocked to allow the probe's visible red guide beam to serve as a visual indicator of the probe beam going through the other side of the aperture. Once the visible red light is seen through the aperture by translating the sample stage on the plane perpendicular to the incident light, an independent pyroelectric detector can then used to make fine adjustments to obtain the highest possible intensity. We have also found that the MCT detector is more than capable and sensitive enough to perform this measurement as well. The maximum signal intensity typically seen on the MCT detector from the pump beam through the aperture is typically around 14-15 counts when the wavelength of mid-IR light is centered around 6150 nm.

After the probe beam has been spatially aligned and signal on the MCT detector has been maximized, the pump must be spatially aligned through the aperture without translating the position of the sample cell. This is achieved by blocking the probe beam, toggling the flip mirror prior to the MCT detector which will allow the pump beam to enter the detector, and the steering the pump mirror just prior to the first parabolic mirror to allow the beam to proceed through the aperture. The pump beam will be of higher intensity and beam width than the probe beam. The signal intensity will usually peak around 40-50 counts, and there may be a wider region where the peak appears at a maximum while adjusting the vertical and horizontal position of the pump reflection mirror. It is therefore important to interpolate the midpoint of the maximum in both the horizontal and vertical directions to the best of one's ability to achieve maximum overlap between the pump and the probe. Once the pump and probe beam are aligned, any subsequent translation of the sample stage in the plane perpendicular to the direction of the incoming light will not affect the spatial alignment. Subsequent translation in the z-direction (incident with the incoming light), however, will affect the spatial overlap of the pump and probe beam and must be otherwise compensated by translating the horizontal direction of the pump mirror to account for the translation of the sample.

Temporal overlap is first achieved through an understanding and analysis to equilibrate the lengths that must be traveled by the pump and probe after they are separated by the beam splitter. While temporal overlap can be initially set to an accuracy of several centimeters by measuring the pump and probe beam paths, delays must also be accounted for from transmission through optical materials. The additional time caused by transmission through the CaF_2 cuvette windows is relatively modest at 6.2 ps/cm, which is equivalent to an additional 1.9 mm in path length [67]. The germanium crystal in the AOM of the pulse shaper, however, has a significant effect due to its high refractive index in the mid-IR. The AOM must transmit through approximately 20 mm of Ge, which delays the pump pulse by approximately 470 ps and is equivalent to an additional path length of 141 mm [67]. Once an approximate match is made of the travel time between the pump and probe, a more precise temporal overlap can be found by translating the delay stage to maximize the signal of an appropriate sample for the corresponding wavelength of incident laser light. The measured value of the delay stage should be recorded for the given wavelength as it will not change to any appreciable degree assuming there are no other changes to the optics, beam path, pulse shaper, or 2D IR software.

6.7 Spectroscopic Calibration

Because changing the wavelength of the light in the TOPAS alters the path, dispersion, and timing of the pump and probe beams, the spatial and temporal overlap procedures mentioned in the previous section must be performed again if calibration or spectroscopic measurements are to be performed outside of the standard parameters set for 2D IR in the Amide I region of the spectrum. There are several steps necessary to achieve an ideal calibration. First, a suitable calibrant molecule must be chosen. The vast majority of our scientific research concerning 2D IR is done in the amide I region of the spectrum. The frequency axis for the probe on a 2D IR spectrum is set by the spectrometer and the MCT detector array. The calibration of the frequency at each pixel is usually performed by using ambient water vapor which can provide three strong absorption peaks at 1541, 1653, and 1699 cm^{-1}. The relative intensity of these peaks, spectral isolation, and spread across the observable spectral range is ideal for performing a calibration in the Amide I region of the spectra. N-Methyl-2-pyrrolidone (NMP) is an organic molecule with an absorption peak at 1650 cm^{-1} which is used to verify the calibration of the water peaks to their assigned pixels. A successful calibration will center the absorption peak of NMP between 1649-1651 cm^{-1}.

If water calibration fails, it is important to check the ambient humidity levels in the laboratory. Water calibration is typically performed with relative humidity levels between 14 - 40%. The HVAC and environmental controls will ensure the relative humidity will seldom rise above 40%. It is possible, however, for external environmental conditions to allow relative humidity to fall below the approximate 14% relative humidity threshold at which point the calibration becomes less accurate and ultimately impossible. Water calibration is typically performed at the beginning of the day, prior to any scientific experimentation. This is beneficial because the pulse shaper cavity will have had time overnight to build up ambient humidity and there should be sufficient ambient water molecules in the gas phase to allow for a viable signal. For calibration in the Amide I region, the typical sample will contain a 1:240 concentration of NMP in D_2O. The first few scans of a sample of the sample take place in ambient environmental conditions. The remaining scans are conducted using a compressed air dryer which delivers approximately 30-40 cubic feet/min of dry air to the shaper box capable which purges ambient air and decreases humidity to levels that are undetectable after 1 minute. Continued use of this compressed air dryer generates enough overpressure in the shaper box to prevent any measurable ambient humidity from entering the cavity and disrupting ongoing experimentation. We have found the ideal measurement of outlet pressure to be between 20 to 25 psi through the ¾” NPT pipe outlet which is then distributed throughout the shaper box as evenly as possible through Nalgene tubes which contain junctions and fluting to prevent any turbulent air flow in the pulse shaper cavity which could potentially kick up dust and damage sensitive optical components.

Successful performance of the daily calibration procedure is critical to ensuring accurate and reproducible spectra. For this reason, multiple calibrations may need to be performed in a single day until a satisfactory calibration spectrum is generated. All peaks must possess the requisite intensity and distinguishability to ensure that pixels are correctly assigned to spectral peaks and to prevent the 2D IR spectrum from becoming skewed as data points are interpolated between assigned pixels. Figure 30 illustrates a successful water calibration spectrum. Figure 31 illustrates an unsatisfactory water calibration spectrum. Although the daily performance of spectral calibration will generally see a relatively modest change or possibly no change to the assignment of pixels, it is nonetheless important to perform this calibration procedure daily to ensure that if unexpected results do arise, they can be appropriately attributed to appropriate biological or physical phenomena.

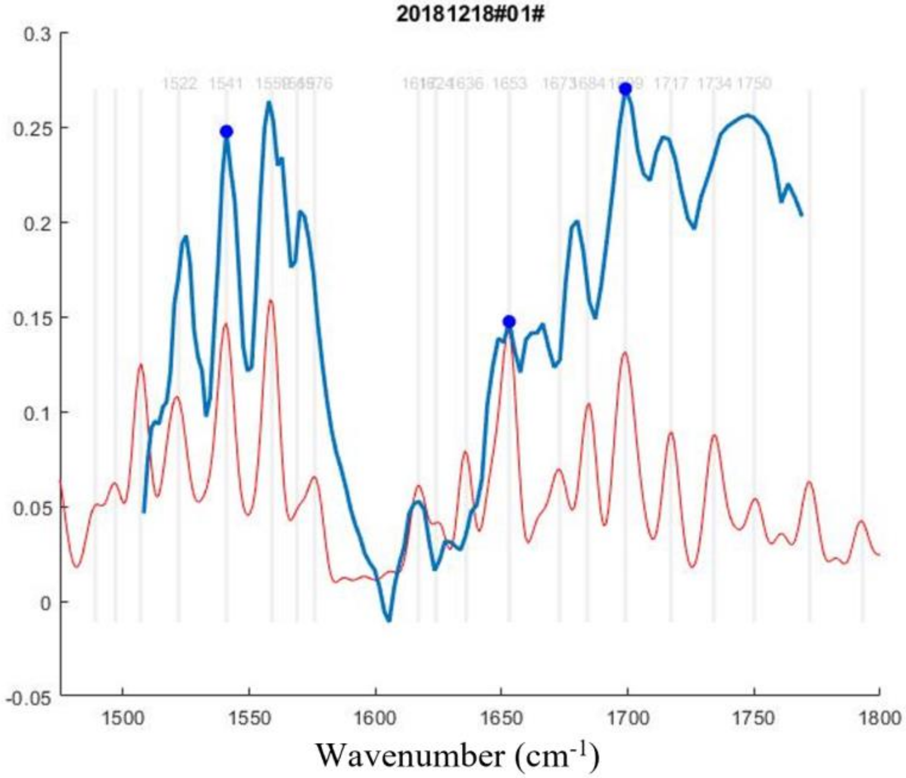


Figure 30: Water calibration spectrum acquired with relatively clean and distinct spectral features for the water spectrum. The red lines represent the calibration standard and serve as a point of reference for distinguishing and assigning peaks. The blue lines represent the calibration spectrum obtained through performing a scan with ambient humidity levels. The 3 dots assigned to the peaks at 1541, 1653, and 1699 cm^{-1} represent where the pixels will be set on the 128-pixel array on the MCT detector. Assigning these pixels correctly will ensure a successful calibration and enable subsequent experimentation throughout the remainder of the day.

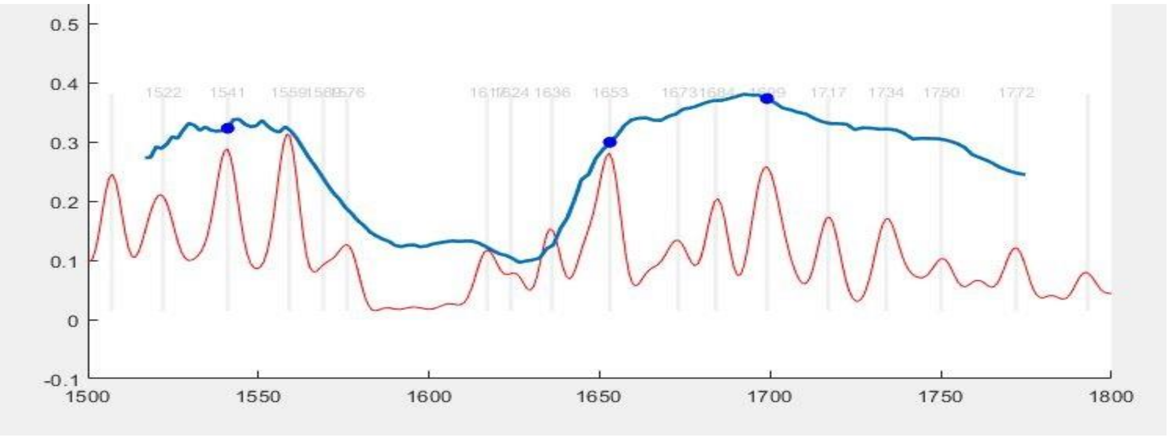


Figure 31: An unsatisfactory water calibration spectrum taken with a lack of distinct spectral features for the water spectrum. Although it is clear that water is present, it is impossible to accurately assign pixel values to the 1541, 1653, and 1699 cm^{-1} peaks. The red lines represent the calibration standard and serve as a point of reference for distinguishing and assigning peaks. The blue lines represent the calibration spectrum obtained through performing a scan with ambient humidity levels. Possible causes for this unsatisfactory spectrum could be: a lack of ambient humidity in the shaper cavity, purge air being turned on too soon, a lack of temporal or spatial alignment, and/or degradation/H_2O exchange of the NMP:D_2O sample.

Because the use of water in a calibration standard can yield inconsistent results and render the laser inoperable on days when the humidity is exceptionally low, we have attempted to find other molecules that are suitable for performing a calibration. 5-nitrosalicylaldehyde and 4-nitrobenzaldehyde (4NB) were both tested due to the relative intensities of their absorption peaks in the Amide I region of the mid-IR. The normalized FTIR spectra are shown below in Figure 32. Ultimately, 4NB appeared to be the better calibrant molecule because the peaks at 1543 cm^{-1} and 1710 cm^{-1} provided the best intensity and distinguishability for any of the combinations of calibrant molecule and solvent (shown below in Figure 33). For both 5-Nitrosalicylaldehyde and 4NB, it was difficult to find a suitable solvent. Toluene and benzene appeared to provide the best solubility among mid-IR transparent solvents. Concentrations between 300-400 mM seemed to provide the best signal to noise ratio. Concentrations higher than 400 mM had a tendency to not completely dissolve and present Mie scattering across the spectrum. Concentrations below 300mM became increasingly difficult to detect. While attempts to use 4NB as a calibration molecule were somewhat successful, there were several reasons why we ultimately decided that water is still a better calibration method. Toluene and benzene are hazardous and volatile. Sample cell preparation can be quite time-consuming as they must be prepared daily and under more stringent safety conditions. The calibration peaks under 2D IR are also not as clearly resolved as those of water. While our tests (shown below in Figure 33) indicated strong absorption peaks at 1557 cm^{-1} and 1706 cm^{-1}, the peak at 1617 cm^{-1} was more difficult to resolve relative to the other two and could only be seen with relatively high concentrations of 4NB.

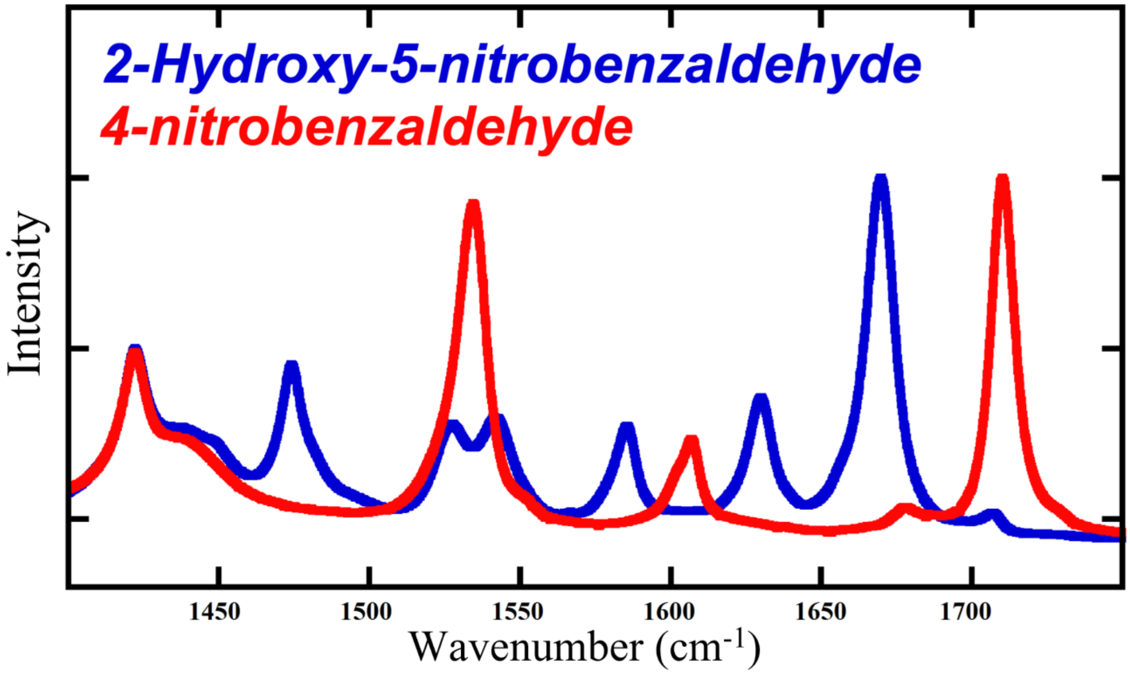


Figure 32: Normalized FTIR Spectra of 4NB and 5-Nitrosalicylaldehyde.

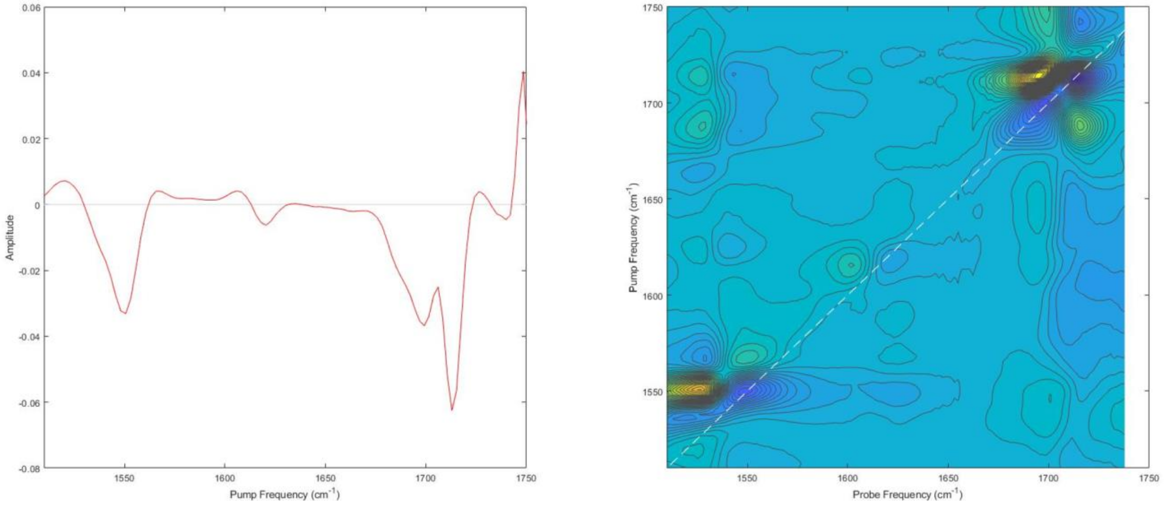


Figure 33: Diagonal slice (left) and 2D IR spectrum (right) of 300mM 4NB in Toluene. The 2D IR data were averaged over 40 scans. Peaks centered at 1557, 1617, and 1706 cm^{-1}

While investigating hexagonal boron nitride (hBN), it became necessary to generate mid-IR light centered at 1350 cm^{-1} on the TOPAS. This rather substantial shift from 1650 cm^{-1} to 1350 cm^{-1} alters the path length, timing, and dispersion of the laser and requires changes to the pulse shaper parameters as well. We have found that the best way to cover the 300 cm^{-1} redshift is to find an intermediate calibration molecule that is capable of allowing signal optimization at successive redshifted bounds for each distinguishable absorption peak. The optimum signal for NMP with the monochromator centered at 1650 cm^{-1}, for instance, has recorded values for grating angle 1, grating angle 2, group velocity dispersion, third order dispersion, the positions of the pump and probe delay stage, and the position of the z-axis of the translation stage. This is further complicated by the fact that the hBN flakes are rather small, with an average area of 4000 μm^2, and present a difficult target to hit for the intersecting pump and probe beams (see Figure 34).

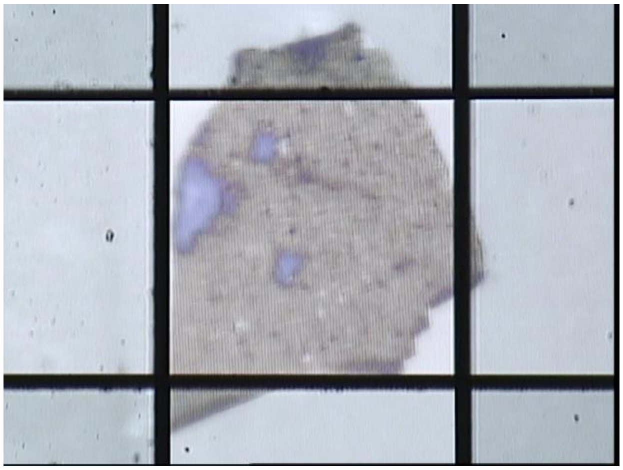
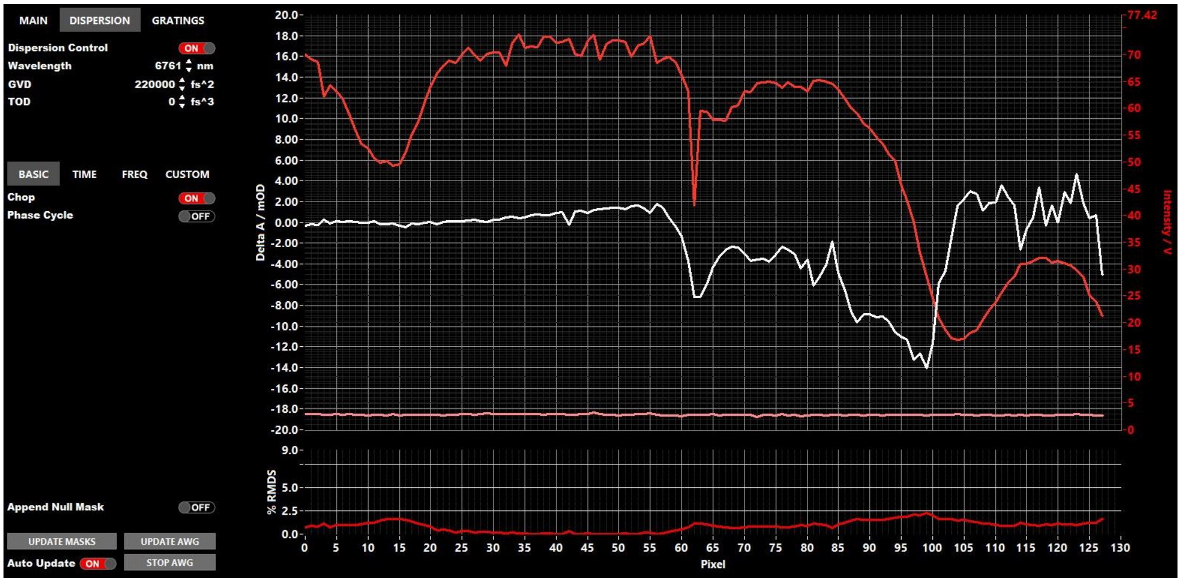


Figure 34: Microscopy image of a mechanically exfoliated hBN flake seen through a 65 µm grid.

An attempt was made at establishing new calibration parameters to account for redshift by taking the aforementioned intermediate steps between 1650 cm^{-1} and 1350 cm^{-1}. 4NB, with its strong absorption peak around 1543 cm^{-1}, and benzene, with a strong IR absorption around 1420 cm^{-1}, were both observed within the spectral window. These two molecules are just within the range of the detector array, and 4NB contains peaks relatively close to the 1626 cm^{-1} peak benchmark established by NMP calibration. A sample of 300 mM 4NB dissolved in benzene was created and the absorption peaks were observed in a 1D acquisition. With the initial experimental setup used for NMP calibration, it was possible to detect a signal for 4NB at 1595 cm^{-1} without any adjustments to the gratings, sample holder, delay stage, GVD or TOD parameters, or pump mirror. The optical parametric amplifier was set to center the laser pulses at an intermediate value of 1480 cm^{-1} and small adjustments were made to establish a signal for the peak around 1420 cm^{-1} for benzene while the signal for 4NB at 1595 cm^{-1} was still present. Relatively small adjustments were made that caused the signal to diminish for 4NB while adjusting the grating angles for the pulse shaper, moving the delay stage, adjusting the pump mirror, moving the sample translation stage, and adjusting the GVD and TOD parameters accordingly to obtain a signal and then maximize the signal for benzene. Because accounting for redshift to maximize the signal is a multivariate adjustment, this process of using intermediate calibration molecules greatly simplified the process of finding a signal in an area of the mid-IR that was previously unexplored. It is beneficial to extrapolate the measured adjustments to the probe translation stage, grating angles, z-translation of the sample cell holder, and GVD correction by using the associated values for these two known calibrant molecules and applying these adjustments to more efficiently obtain a signal for hBN and any other molecule under examination. Having reliable estimates can help constrain the time spent attempting to acquire a signal.



Dispersion Control ON
Wavelength 6761 nm
GVD 220000 fs^2
TOD 0 fs^3
Chop ON
Phase Cycle OFF
Append Null Mask OFF
Auto Update ON

Figure 35: IR scan of 5-Nitrosalicylaldehyde in benzene with the spectrum centered at 1480 cm^{-1}. Adjustments were made to instrumentation to obtain a signal for benzene (around pixel #100) as the signal for 5nitro diminished (around pixel #63). Observing both signals simultaneously allows for the measurement of the shift in temporal and spatial parameters as instrumentation and software adjustments are made to obtain maximum signal.

Taking the aforementioned series of steps will ensure the repeatability of future experimentation for varying 2D IR parameters and allow the user to more quickly adjust the experimental setup of the 2D IR system outside of its usual parameters in the Amide I region of the spectrum. This same method was used during the initial setup for the laser. $W(CO)_{6,}$ with a strong absorption around 2000 cm^{-1} and a long temporal lifetime was initially used to benchmark the performance of the laser. As it became necessary to perform scientific experimentation of the Amide I region of the infrared spectrum, acetyl aldehyde, with an absorption peak around 1743 cm^{-1}, was used as an intermediate to extrapolate adjustments to instrumentation necessary to ultimately simplify the acquisition and optimization of a signal for NMP at 1626 cm^{-1}.

REFERENCES

[1] Bury, M; et al.; Biomaterials, **2014**, 35, 9311-9321

[2] Hartgerink, J; Beniash, E; Stupp, S; Science, **2001**, 294, 1684-1688

[3] Herrera, L; Champion, J; Biomater. Sci., **2015**, 3, 787-799

[4] Boekhoven, J; et al.; RSC Adv., **2015**, 5, 8753-8756

[5] Tao, K; Gazit, E; et al.; Science, **2017**, 358 (6365)

[6] Kumar, R; Holmes, A; J. Am. Chem. Soc., **2011**, 132 (22), pp. 8564-8573

[7] Korevaar, P; Newcomb, C; Meijer, E; Stupp, S; J. Am. Chem. Soc., **2014**, 136, 8540-8543

[8] Unzueta, U; et al.; Biomaterials, **2012**, 33, 8714-8722

[9] Kelly, S; Jess, T; Price, N; Biochimica et Biophysica, **2005**, 1751, 119 – 139

[10] Zanni, M; Ghosh, A; Ostrander, J; Chem. Rev., **2017**, 117 (16), pp 10726–10759

[11] Kim, Y; Hochstrasser, R; et al.; Proc. Natl. Acad. Sci., **2009**, 106, 17751, 17756

[12] Woys, A; et al.; J. Am. Chem. Soc., **2011**, 133, 16062-16071

[13] Krimm, S; Bandekar, J; Adv. Protein Chem, **1986**, Vol. 38, pp 181-364

[14] Buchanan, L; PhD dissertation, *Mechanism of amyloid aggregation revealed by 2D infrared* spectroscopy, **2013**, University of Wisconsin, Madison

[15] Donaldson, M; Towrie, M; et al.; J. Phys. Chem. A**, 2018,** 122, 780−787

[16] Henzler-Wildman, K; Kern, D; Nature, **2007**, 450, 964-972

[17] Wernet, P; Nordlund, D; Cavalleri, M; Hirsch, T; Nilsson, A; Science, **2004**, 304, 995-999

[18] Bredenbeck, J; Helbing, J; Hamm, P, J. Chem. Phys., **2004**, 8, 1747-1756

[19] Finkelstein, I; Zheng, J; Ishikawa, H; Fayer, M; et al.; Phys. Chem. Chem. Phys., **2007**, 9, 1533-1549

[20] Ge, N; Hochstrasser, R; Phys. Chem. Comm., **2002**, 5, 17-26

[21] Middleton, C; Woys, A; Mukherjee, S; Zanni, M; Methods, **2010**, Sep;52(1): 12-22

[22] Zanni, M; Hochstrasser, R; Curr. Opin. Struct. Bio., **2001,** 11, 516-522

[23] Eaton, W; Jas, G; Lapidus, L; Hofrichter, J; et al.; Annu. Rev. Biophys. Biomol. Struct, **2000**, 29, 327-359

[24] Hamm, P; Zanni, M; *Concepts and Methods of 2D Infrared Spectroscopy, 1st ed,*; **2011**, Cambridge University Press, New York

[25] Mukamel, S.; *Principles of Nonlinear Optical Spectroscopy*, 1st. ed.; **1995**, Oxford University Press, New York

[26] Fayer, M; *Ultrafast Infrared Vibrational Spectroscopy*, **2013**, CRC Press, Florida

[27] Khalil, M; Demirdoven, N; Tokmakoff, A; J. Phys. Chem. A, **2003**, 107, 5258-5279

[28] Middleton, C; Woys, A; Mukherjee, S; Zanni, M; Methods, **2010**, 12-22

[29] Zanni, M; Hochstrasser, R; Curr. Opin. Struct. Bio., **2001,** 11, 516-522

[30] Pelton, J; McLean, L; Anal. Biochem, **2000**, 277, 167-176

[31] Cheatum, C; Tokmakoff, A; Knoester, J: J. Chem. Phys., **2004**, 120, 8201

[32] Hayashi T, Zhuang W, & Mukamel S; J. Phys. Chem. A, **2005**, 109(43):9747-9759

[33] Cheatum, CM ; Tokmakoff, A ; Demirdöven, Nurettin ; Chung, Hoi Sung; J. Am. Chem. Soc., **2004**, Vol.126(25), pp.7981-7990

[34] Reddy, A; Wang, L; Singh, S; Ling, Y; Buchanan, L; Zanni, M; Skinner, J; de Pablo, J; Biophys J, **2010**, 99, 2208

[35] Choi, J; Ham, S; Cho, M; J Phys. Chem. B, **2003**, 107, 9132

[36] Buchanan, L; Dunkelberger, E; Zanni, M; *Protein Folding and Misfolding*, **2012**, pp. 217-237, Springer, Berlin

[37] Arkin, I; Curr. Opin. Chem. Bio., **2006**, 10, 394-401

[38] Remorino, A; Hochstrasser, R; Acc. Chem. Res., **2012**, 45, 1896-1905

[39] Woys, A; Lin, Y; Zanni, M; et al.; J. Am. Chem. Soc., **2010**, 132, 2832-2838

[40] Moran, S; Buchanan, L; Zanni, M; et al.; Proc. Natl. Acad. Sci., **2012**, 109, 3329-3334

[41] Li, Y; Zhang, X; Cao, D; Scientific Reports, **2013,** 3, 3271

[42] Liu, Y; Zhang, L; Wei, W; Int. J. Nanomedicine, **2017**, 12, 659-670

[43] Lee, J; Macosko, C; Urry, D; Biomacromolecules, **2001**, 2, 170-179

[44] Xu, H; et al.; Langmuir, **2009**, 25, 4115-4123

[45] Fileti, E.; Colherinhas, G; J Phys. Chem. B, **2014**, 118 (42), 12215-12222.

[46] Starheim, K; Gebaert, K; Arnesen, T; Trends in Biochem. Sci., **2012**, V. 37, No. 4

[47] Scott, D; et al.; Science, **2011**, 334, 674-678

[48] Arnesen, T; PLoS Biology, **2011**, Vol 9, 5, e1001074

[49] Dai, S; et al.; Nature Communications, **2015,** 6, 6963

[50] Folland, T; Caldwell, J; Quantum Nano-Photonics, **2018**, pp. 235

[51] Caldwell, J; Lindsay, L; Giannini, V; et al.; Nanophotonics, **2015,** 4: p. 44-68

[52] Fisher, R; Gould, R; Physical Review Letters, **1969**, 22: p. 1093-1095

[53] Dunkelberger, A; Caldwell, J; et al; Nature Photonics, **2018,** 12, 50-56

[54] Takemura, N; Trebaol, S; Anderson, M; Kohnle, V; Léger, Y; Oberli, D; Portella-Oberli, M; Deveaud, B; Phys. Rev. B, **2015**, 92, 125415

[55] Xiang, B; Ribeiro, R; Dunkelberger, A; Xiong, Wei; et al.; PNAS, **2018**, 115 (19) 4845-4850;

[56] Takemura, N; Terbaol, S; Wouters, M; Physical Review B, **2014**, 90: 195307

[57] Wu, X; Fu, C; J. of Quantitative Spect. & Radiative Transfer, **2018**, 209 150-155

[58] Kratochvil, H; Zanni, M; et al; Science, **2016,** Vol. 353, Issue 6303, p.1040-1044

[59] Thomaston, J.; DeGrado, W; et al; PNAS, **2017,** 114 (51) 13357-13362

[60] Rehault, J; Helbing, J; Optics Express, **2012,** Vol.20, No. 19, 21676

[61] DeGrado, W; Thomaston, K; Protein Science, **2016**, Aug: 25(8):1551-4

[62] Merrifield, R; J. Am. Chem. Soc., **1963**, 85 (14), pp 2149–2154

[63] Nguyen, L.; Chau, J.; Perry, N.; de Boer, L.; Vogel, H. PLoS ONE. **2010**, 5, e12684

[64] Marek, P; Woys, A; Sutton, K; Zanni, M; Raleigh, D; Org. Lett., **2010**, 12(21): 4848-4851

[65] Meng, Q; Kou, Y; Ma, X; Liang, Y; Guo, L; Ni, C; Liu, K; Langmuir, **2012**, 28, 5017-5022

[66] Backus, S; Durfee, C; Murnane, M; Kapteyn, H; Rev. Sci. Inst., **1998**, 69, 1207-1223

[67] Middleton, C; Woys, A; Mukherjee, S; Zanni, M; Methods, **2010**, 52, 12-22