# Developing nonperturbative spectroscopic approaches for characterizing heterogenous ensembles during protein self-assembly

By

William Weeks

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

In

Chemistry

May 12, 2023

Nashville, Tennessee

Approved:

Lauren E. Buchanan, Ph.D.

Michael P. Stone, Ph.D.

Prasad L. Polavarapu, Ph.D.

Charles R. Sanders, Ph.D.

Copyright © 2022 William Weeks

All Rights Reserved

To my beloved wife, Keelan: Without your endless support and fathomless love, this work would not be possible.

And to my darling daughter, Brigid: Dream big princess!

# Acknowledgements

I would like to thank Vanderbilt and the Vanderbilt Chemistry Department for providing the funding for my research and means to present my research at domestic and international conferences.

I have many influential people to thank for my success in my career, none more so than Dr. Lauren Buchanan. I could not have asked for a better mentor, teacher, and friend. From trying to grasp quantum chemistry to raising an infant, her guidance has been invaluable through my graduate career. Similarly, I want to thank Dr. Sharon Hamilton, my undergraduate advisor, for stirring my interests in research and all the conferences adventures.

I'd also like to thank past and present members of committee: Dr. Jens Miller, Dr. Claire McCabe, Dr. Charles Sanders, Dr. Prasad Polavarapu, and Dr. Michael Stone for their instruction and mentorship. Additionally, I'd like to thank the members of VINSE for their support. I'd like to especially thank Dr. James McBride so all the help in electron microscopy and Dr. Dmitry Koktsh for all the analytical support.

I've made many enduring friends and colleagues during my time at Vanderbilt. I'd like to thank Nathan Spear and Christopher Sharp for their support and companionship over the past five years. I'd also like to thank my past and present lab members: Kelsey Webb, Daniel DeNeve, Kayla Hess, Dr. Lindsey Miller, Amanda Cao, Ali Fullilove, and Jessica Gaetgens.

Lastly, I want to thank my entire family for their continuous love and support.

**Table of Contents**

## List of figures

**Figure** **Page**

# List of Abbreviations

2D IR- Two-dimensional infrared spectroscopy

$|\mu|$- Transition dipole moment

μm- Micrometer

μM- Micromolar

Aβ42- Amyloid beta polypeptide 42

CD- Circular dichroism

$cm^{-1}$- Wavenumber

DIC- Diisopropylcarbodiimide

DLS- Dynamic light scattering

DMF- Dimethyl formaldehyde

DMSO- Dimethylsulfoxide

FMOC- Fluorenylmethyloxycarbonyl protecting group

HPLC- High performance liquid chromatography

IR- Infrared

kV- Kilovolts

MADLS- Multi-angle dynamic light scattering

mm- Micrometer

mM- Micromolar

nm- Nanometer

NMA- N-methylaniline

NMR- Nuclear magnetic resonance

OD- Optical density

RNA- Ribonucleic acid

SEC- Size exclusion chromatography

SEM- Scanning electron microscopy

Ser- L-Serine

TDS- Transition dipole strengths

TEM- Transmission electron microscopy

ThT- Thioflavin

# Chapter 1

# Introduction

## 1.1. Introduction

Biomolecular self-assembly is a powerful phenomenon that occurs naturally or can be employed strategically to create novel biomaterials that have revolutionized modern science. Self-assembly of many types of biomolecules is ubiquitous in nature. Common examples are the formation of monomers of lipids forming oil droplets in water or ribosomal proteins coalescing with RNA to form functional ribosomes(1). These processes are non-enzymatic and rely on the passive interactions of the biomolecules such as hydrogen bonding, electrostatic interactions, and van der Wal's forces for successful assembly(2, 3) (Figure 1.1). A deep understanding of these passive interactions is needed to rationally design novel biomaterials with specific functionality. Biomaterials, materials other than food or drugs that are in contact with tissue or biological fluids, have developed many aspects of modern medicine. This includes nearly all biomolecules such as lipids, protein, and RNA. The applications of biomaterials are extensive in science and currently range from tissue engineering as cell scaffolds in regenerative medicine to plastic surgery and dental augmentation(4–6). For example, peptide building blocks used as the base of 3D cell scaffolding have been shown to increase cell adherence and growth stimulating factors compared to scaffolds formed via traditional decellularized extracellular matrix components(7). The main advantage of self-assembling peptides lies in the ability to readily extend the scaffold to include more biologically active sequences further facilitating cell growth(8). The versatility of protein and peptide self-assembly is demonstrated throughout nature by the production of a dazzling diverse array of protein structures such as collagen, pearl, and even optical waveguides from just 20 basic amino acids(1). By observing these processes, we can begin to design synthetic biomaterials geared toward a specific biological function. Generally, there are two accepted strategies used in the fabrication of biomaterial design: a 'top-down' approach, in which biomaterials are fabricated

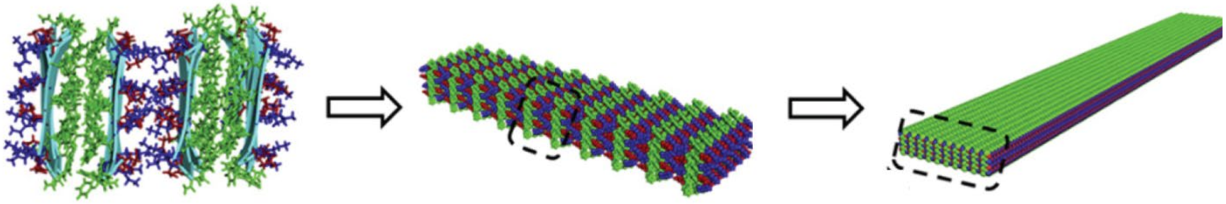

**Figure 1.1:** Example of peptide self-assembly driven via passive forces. Monomer species arrange to form bilayers as the hydrophobic residues (green) are held stable via hydrophobic interaction and π-bond stacking. The charged hydrophobic residues (blue and red) contribute strong hydrogen bonding with each other as well as the solvent. These bilayers propagate and elongate into flat protein fibrils driven entirely by self-assembly. Adapt from reference 3.

by stripping down a complex entity into component parts and a ‘bottom up’ approach in which the biomaterial is assembled molecule by molecule, or monomer by monomer. The latter approach holds more appeal due to the higher diversity of functionality granted to the fabricated material. However, this approach requires a deep understanding of the monomeric characteristics and therein an understanding of the passive interactions that drive the peptide aggregation mechanism. Here, limitations and challenges arise in identifying, isolating, and analyzing the monomeric structures and their self-assembling pathways. My research has focused on developing nonperturbative spectroscopic methods capable of analyzing these characteristics in peptide and protein systems in real time. The results and methods presented herein constitute progress toward developing a ‘bottom up’ approach in protein biomaterial design by revealing in-depth information previously only accessible via molecular dynamic simulations.

### 1.2. Infrared spectroscopy in proteins

Traditional methods of analyzing protein structure during self-assembly include X-ray crystallography, circular dichroism (CD), and nuclear magnetic resonance (NMR) spectroscopy. X-ray crystallography yields the highest structural resolution of these techniques. However, the protein have to be able to form stable crystalline state. Therefore, X-ray crystallography is limited to protein that will crystalize and is a static technique yielding no information regarding self-assembly kinetics or oligomer formation during assembly(9, 10). NMR is arguably the most widely utilized method for probing protein structure due to the diversity of experiments including multiple dimensional analysis and both solution-phase and solid-state experiments. Although sample requirements are less rigorous than X-ray crystallography, NMR still typically requires homogeneous sample populations to assign specific signal peaks to the corresponding protein component(11). This considerably compounds the analysis complexity of most protein samples as well as the time required to acquire data with an acceptable signal-to-noise ratio(12). Optical techniques such as CD, fluorescence, and infrared spectroscopy are largely advantageous yielding the ability to dynamically monitor protein structure and kinetics during self-assembly in solution. Florescent dyes sensitive to certain peptide structures have been used to track structural changes. For example, Thioflavin T (ThT) fluoresces upon binding to specifically amyloidogenic β-sheets(13). The growth of fibrils composed of these β-sheets can be monitored by analyzing the fluorescence intensity over time. However, recent studies show the binding of ThT can alter the structure of protein aggregates, if only slightly(14). These small changes in structure results in skewed kinetic data as well as inactive signals of oligomer or intermediate structure during self-assembly. Electronic and vibrational CD use circularly polarized light to identify the secondary structures within a protein ensemble(15, 16). While both techniques yield some amount of structural and kinetic data, they are bulk techniques lacking the ability to generate residue-level structural information. This residue-specific data is critical in understanding passive interactions such as hydrogen bonding, hydrophilic interactions, and π-bond stacking that drive protein self-assembly. Additionally, although both yield varying levels of temporal information, neither is fast enough to measure the timescale of these interactions as hydrogen bonding exchanges on the sub-picosecond level.

To overcome these inherent challenges in probing protein structure and dynamics, I implement infrared (IR) spectroscopy. When used with ultrafast pulsed laser systems and isotope labels, IR spectroscopy gains the structural and temporal resolution necessary to study self-assembly in real time. IR spectroscopy is able to probe secondary structure directly from peptide ensembles without implementing bulky spin tags or fluorescent dyes, both of which have been shown to alter the peptide structure and aggregation mechanism(14, 17). Paired with strategic isotope-labeling schemes, IR spectroscopy gains the desired residue-specific information needed to truly understand the protein structure and dynamics(18).

When probing protein structure via IR spectroscopy, the most common vibrational mode analyzed is the amide I' mode generated by the backbone carbonyl stretching motions(19). The native amide I' mode of a disordered peptide typically absorbs broadly around 1650 $cm^{-1}$. When peptides adopt organized secondary structures, such as β-sheets or α-helixes, the amide modes are held in close spatial proximity to each other and are able to couple. The vibrational coupling of several amide groups results in a larger vibrational mode delocalized across several residues (20). The amount of vibrational coupling varies depending on the secondary structure and alters the frequency at which the corresponding amide I' mode absorbs, yielding characteristic spectral features for each structure(19, 20). The strongest couplings result from the tightly formed β-sheets. Since the carbonyls are held aligned with each other, the perpendicular stretching modes result in a narrow linewidth absorption shifted to around 1620-1630 $cm^{-1}$. Antiparallel β-sheets yield an additional absorption peak at 1680 $cm^{-1}$ indicative of the perpendicular stretching modes(19–21). This vibrational mode is much weaker compared to its parallel counterpart and only appears in highly ordered antiparallel β-sheets. In contrast, weaker coupling between residues in the twist of α-helical structures result in a broader lineshapes absorbing around 1635-1655 $cm^{-1}$(22, 23).

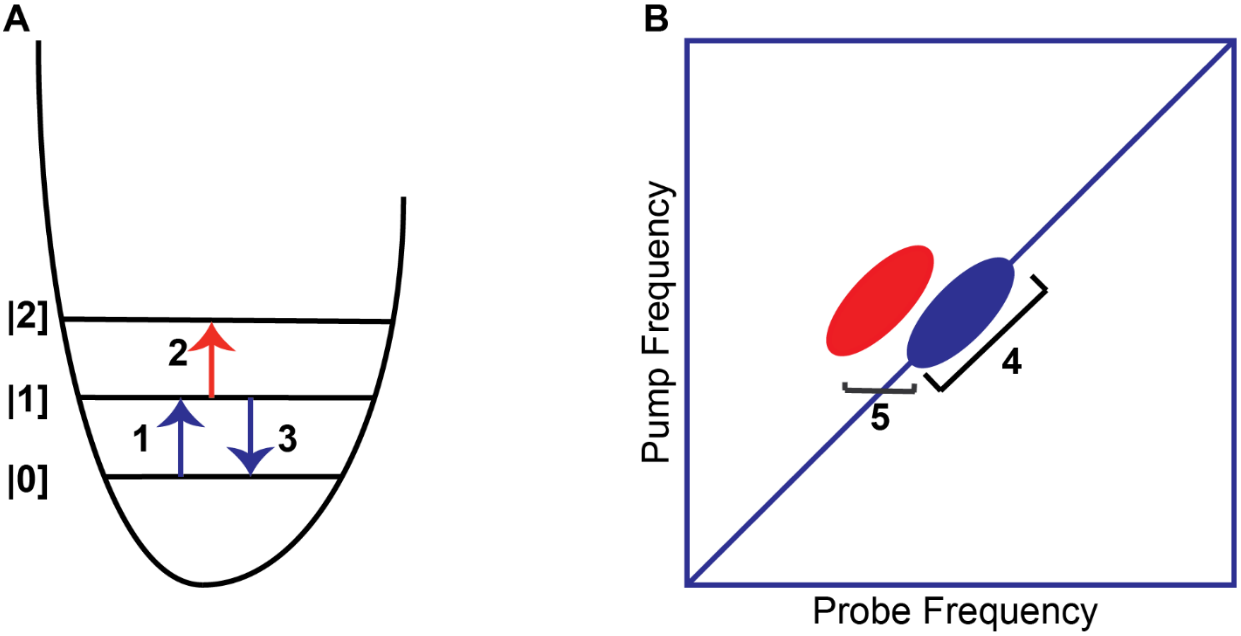

**Figure 1.2:** 2D IR spectroscopy overview. The fundamental peak (B, blue oval) appears along the diagonal center at the intersection of the same pump-probe frequency representing the fundamental ground state transition (A1) as well as the bleach or stimulated emission (A3). The overtone peak (B, red oval) arises from the excited state absorbance (A2). The distance between the fundamental and overtone peaks (B5) is a measure of the anharmonicity of the vibrational mode. The peak lineshapes (B4) reflect the solvent environment of the vibrational mode.

To further enhance both spatial and temporal resolution, I employ two-dimensional infrared (2D IR) spectroscopy via a pump-probe system utilizing pulse shaping technology(24). 2D IR spectroscopy retains the same basic theory of linear IR spectroscopy while spreading the vibrational spectrum over two frequency dimensions allowing for resolution of vibrational modes previously hidden in linear IR spectra(25) (Figure 1.2). The revealed spectral features often correspond to weaker coupled residues otherwise enveloped in linear IR spectroscopy. Additional spectral features created also include lineshapes, anharmonicity, and increase sensitivity to amide I' mode frequency and intensity which can be directly interpreted from the 2D IR spectrum(26, 27). This improved sensitivity is largely due to the additional interactions of the transition dipole moment ($|\mu|$) with the IR light. While the transition dipole strength (TDS) in linear IR spectroscopy scales as $|\vec{\mu}|^2$, the TDS, and therefore the sensitivity, increases in 2D IR spectroscopy as the 2D IR dipole moment scales as $|\vec{\mu}|^4$ due to increased interactions of light(19).

### 1.3. Using transition dipoles strengths to analyze protein structure

As previously mentioned, each secondary structure has a characteristic range of frequencies rather than a single corresponding frequency. This variation arises from the sensitivity of the vibrational frequency to small differences in coupling and electrostatic environment (28). For example, native short amphiphilic peptides forming β-sheets are more likely to absorb closer to 1620 $cm^{-1}$ while longer amyloidogenic peptides also forming β-sheets are likely to absorb around 1630 $cm^{-1}$(29). Additionally, even the frequency ranges are not truly unique to each secondary structure as frequency ranges overlap. The range of disordered and α-helix structures are overlapping between 1645-1650 $cm^{-1}$. In theory, it is possible to fit the amide I' absorption band to differentiate the vibrational modes(30). However, this presents several difficulties due to the structural disorder and complex lineshapes among other problematic characteristics.

In my research, I exploit transition dipole strengths in order to identify unique secondary structures in seemingly homogeneous protein ensembles. The transition dipole strength derives from Beer's law as seen in Equation (1.1) where A is absorbance, ε is the extinction coefficient, c is the protein concentration, and l is path length(31). Equation (1.2) shows the direct relationship between the transition dipole moment and the extinction coefficient. Transition dipole strengths can be more sensitive to secondary structure than the absorption frequency alone. The characteristic frequency shifts which denote the individual secondary structures result from the strong coupling of amide I' local modes. However, in many cases the redistribution of the oscillator

$$A = \epsilon \times c \times l \tag{1.1}$$

$$\epsilon = |\vec{\mu}|^2 \tag{1.2}$$

strengths (known as delocalization) occurs but is too weak to produce coupling capable of causing frequency shifts(28, 31). Transition dipole strengths scale linearly with the delocalization and are therefore more sensitive to small structural changes otherwise undetected by absorbance frequency alone(32). The extent of vibrational delocalization is affected by several factors, including the peptide/ aggregate size, the overall secondary structure, and the amount of structural disorder. Therefore, analyzing amide I' frequency shifts in conjunction with TDS values reveal previously hidden structural variation and strengthen spectral model assignment. For example, disordered peptides generally absorb around 1645 $cm^{-1}$, which corresponds to the native amide I′ vibrational frequency. Depending on their local environment, α-helices absorb between 1635 and 1655 $cm^{-1}$ and thus, due to significant spectral overlap, can be difficult to distinguish from disordered structures by frequency alone. Yet, the highly ordered hydrogen-bonding network in α-helices brings the individual backbone amide groups into close enough proximity to couple(34, 35). This coupling causes the amide I′ mode to delocalize over multiple amide units. Since TDS scales linearly with vibrational delocalization, this leads to an overall increase in the TDS of the amide I' mode compared to that of a disordered protein(30). The lower limit of residues over which the amide I' mode is delocalized can be determined from the TDS value. In a disordered structure, the amide I' mode is localized onto individual residues yielding a TDS value of 0.12 $D^2$. It has been previously determined the amide I' mode can delocalize across up to 3.5 residues in α-helix structures compared to the highly ordered β-sheet in which delocalization can stretch up to 12.5 residues(30, 33).

### 1.4. Applications

By combining the direct analysis of frequency absorbance, lineshapes, and anharmonicity with the ability to calculate TDS values from spectra taken within the same data set, I was able to obtain highly sensitive structural information from protein and peptide ensembles. Additionally, isotope-labeling schemes add residue-specific information previously only suggested through molecular dynamic simulations. To demonstrate these capabilities, I analyze the effects of N-terminal acetylation on small model amphiphilic peptide. Through small, approximately 40 amu acetyl capping unites at the N-terminus, the AcKFE8 peptide adopts a unique left-handed helical formation compared to the flat ribbon morphology of the unacetylated counterpart(36). Upon analysis, I found that the native 2D IR spectra of each peptide variant were nearly identical. With a strategic isotope-labeling scheme, I was able to uncover the presence of two unique β-sheet configurations only differentiated through the amount of terminal off-set. Both β-sheet configurations exist in the AcKFE8 peptide creating the left-handed helical morphology while the nonacetylated variant only contain β-sheets with N-terminal staggering shown in chapter 3. Furthermore, with the incorporation of TDS calculations, I reveal the presence of early-stage oligomers in AcKFE8 in chapter 4. This suggests TDS values not only improve structural assignment, but also reveal additional structure in otherwise seemingly homogeneous peptide ensembles.

In chapter 5, I attempt to apply these techniques to a more complex family of amyloidogenic proteins, amyloid beta 1-42 (Aβ42). While previous research has shown acetylation of the lysine side chains within Aβ42 changes the final morphologies and cytotoxicity, the mechanism or secondary structure composition is unknown(37, 38). I use a combination of 2D IR spectroscopy and TDS calculation to reveal the largely β-sheet composition of each variant while also revealing the weaking of the delocalization across the structure as a function of lysine acetylation. I have also attempted to track the aggregation kinetics of each Aβ42 variant through 2D IR spectroscopy and ThT fluorescence to distinguish.

In a final collaborative study, I analyze the effects of lipid micelles and small molecules inhibitors on the structure and aggregation of the C99 transmembrane protein. Through 2D IR spectroscopy experiments, I reveal considerable structural transition from β-sheet dominant populations to α-helix structures based on micelle concentration. By varying the lipid concentration, I found a critical micelle concentration relative to the C99 concentration which induces the structural change from β-sheet to α-helix motifs. Additionally, I was able to detect weakening of β-sheet structures by introducing a small molecule found via high throughput screening. These findings were compared to and supported by 2D NMR results previously collected under matching conditions.

### 1.5. Summary

The combination of 2D IR spectroscopy with transition dipole strength is a powerful tool for analyzing protein structure and aggregation. By observing the presence of multiple TDS values within a single peptide aggregate, I was able to reveal early-stage oligomer formation within a single peptide aggregate. This technique has the possibility for early screening of early-stage oligomers before a deeper more time consumer, and more expensive, experiment course is set. These techniques move science closer to understanding how small changes in primary structure effect the passive interactions that control the self-assembly of proteins and peptides. This information will prove pivotal in the future rational design of peptide and protein biomaterials.

### 1.6. References

1. Zhang, S. 2003. Fabrication of novel biomaterials through molecular self-assembly. *Nat. Biotechnol.* 21:1171–1178:10.1038/nbt874.

2. Webber, M.J., E.A. Appel, E.W. Meijer, and R. Langer. 2016. Supramolecular biomaterials. *Nat. Mater.* 15:13–26:10.1038/nmat4474.

3. Zhou, P., L. Deng, Y. Wang, J.R. Lu, and H. Xu. 2016. Different nanostructures caused by competition of intra- and inter-β-sheet interactions in hierarchical self-assembly of short peptides. *J. Colloid Interface Sci.* 464:219–228:10.1016/J.JCIS.2015.11.030.

4. Leonor, I.B., A.I. Rodrigues, and R.L. Reis. 2016. Designing biomaterials based on

biomineralization for bone repair and regeneration. In: Biomineralization and Biomaterials. Elsevier. pp. 377–404.

5. Webber, M.J., O.F. Khan, S.A. Sydlik, B.C. Tang, and R. Langer. 2015. A Perspective on the Clinical Translation of Scaffolds for Tissue Engineering. *Ann. Biomed. Eng.* 43:641–656:10.1007/S10439-014-1104-7/TABLES/1.

6. Freed, L.E., G. Vunjak-Novakovic, R.J. Biron, D.B. Eagles~, D.C. Lesnoy~, S.K. Barlow', and R. Langer. 1994. Biodegradable Polymer Scaffolds for Tissue Engineering. .

7. Miller, R.E.E., A.J.J. Grodzinsky, E.J.J. Vanderploeg, C. Lee, D.J.J. Ferris, M.F.F. Barrett, J.D.D. Kisiday, and D.D.D. Frisbie. 2010. Effect of self-assembling peptide, chondrogenic factors, and bone marrow-derived stromal cells on osteochondral repair. *Osteoarthr. Cartil.* 18:1608–1619:10.1016/j.joca.2010.09.004.

8. Genové, E., C. Shen, S. Zhang, and C.E. Semino. 2005. The effect of functionalized self-assembling peptide scaffolds on human aortic endothelial cell function. *Biomaterials*. 26:3341–3351:10.1016/j.biomaterials.2004.08.012.

9. Wiltzius, J.J.W., S.A. Sievers, M.R. Sawaya, D. Cascio, D. Popov, C. Riekel, and D. Eisenberg. 2008. Atomic structure of the cross-β spine of islet amyloid polypeptide (amylin). *Protein Sci.* 17:1467–1474:10.1110/PS.036509.108.

10. Barry, M.E., M.J. Rynkiewicz, E. Pavadai, W. Lehman, and J.R. Moore. 2022. Destabilizing tropomyosin-troponin I interactions increases calcium sensitivity. *Biophys. J.* 121:46a-47a:10.1016/j.bpj.2021.11.2482.

11. Petkova, A.T., R.D. Leapman, Z. Guo, W.M. Yau, M.P. Mattson, and R. Tycko. 2005. Self-propagating, molecular-level polymorphism in Alzheimer's β-amyloid fibrils. *Science (80-. ).* 307:262–265:https://www.science.org/doi/10.1126/science.1105850.

12. Paravastu, A.K., R.D. Leapman, W.-M. Yau, and R. Tycko. 2008. Molecular structural basis for polymorphism in Alzheimer's -amyloid fibrils. *Proc. Natl. Acad. Sci.* 105:18349–18354:10.1073/pnas.0806270105.

13. Ban, T., D. Hamada, K. Hasegawall, H. Naiki, Y. Goto, K. Hasegawa, H. Naiki, and Y. Goto. 2003. Direct Observation of Amyloid Fibril Growth Monitored by Thioflavin T Fluorescence. *J. Biol. Chem.* 278:16462–16465:10.1074/jbc.C300049200.

14. Di Carlo, M.G., V. Minicozzi, V. Foderà, V. Militello, V. Vetri, S. Morante, and M. Leone. 2015. Thioflavin T templates amyloid β(1–40) conformation and aggregation pathway. *Biophys. Chem.* 206:1–11:10.1016/j.bpc.2015.06.006.

15. Measey, T.J., and R. Schweitzer-Stenner. 2011. Vibrational circular dichroism as a probe of fibrillogenesis: The origin of the anomalous intensity enhancement of amyloid-like fibrils. *J. Am. Chem. Soc.* 133:1066–1076:10.1021/JA1089827/SUPPL_FILE/JA1089827_SI_001.PDF.

16. Kelly, S.M., T.J. Jess, and N.C. Price. 2005. How to study proteins by circular dichroism. *Biochim. Biophys. Acta - Proteins Proteomics*. 1751:119–139:10.1016/j.bbapap.2005.06.005.

17. Jassoy, J.J., C.A. Heubach, T. Hett, F. Bernhard, F.R. Haege, G. Hagelueken, and O. Schiemann. molecules Site Selective and Efficient Spin Labeling of Proteins with a Maleimide-Functionalized Trityl Radical for Pulsed Dipolar EPR Spectroscopy. 10.3390/molecules24152735.

18. Strasfeld, D.B., Y.L. Ling, R. Gupta, D.P. Raleigh, and M.T. Zanni. 2009. Strategies for extracting structural information from 2D IR spectroscopy of amyloid: application to islet amyloid polypeptide. *J. Phys. Chem. B.* 113:15679–91:10.1021/jp9072203.

19. Hamm, P., and M. Zanni. 2011. Concepts and Methods of 2D Infrared Spectroscopy. Cambridge: Cambridge University Press.

20. Kim, Y.S., and R.M. Hochstrasser. 2009. Applications of 2D IR Spectroscopy to Peptides, Proteins, and Hydrogen-Bond Dynamics. *J. Phys. Chem. B*. 113:8231–8251:10.1021/jp8113978.

21. Baiz, C.R., and A. Tokmakoff. 2015. Structural Disorder of Folded Proteins: Isotope-Edited 2D IR Spectroscopy and Markov State Modeling. *Biophys. J.* 108:1747–1757:10.1016/j.bpj.2014.12.061.

22. Buchanan, L.E., E.B. Dunkelberger, and M.T. Zanni. 2012. Examining Amyloid Structure and Kinetics with 1D and 2D Infrared Spectroscopy and Isotope Labeling. In: Fabian H, D Naumann, editors. Protein Folding and Misfolding: Shining Light by Infrared Spectroscopy. Berlin, Heidelberg: Springer Berlin Heidelberg. pp. 217–237.

23. Moran, S.D., A.M. Woys, L.E. Buchanan, E. Bixby, S.M. Decatur, and M.T. Zanni. 2012. Two-dimensional IR spectroscopy and segmental 13C labeling reveals the domain structure of human γD-crystallin amyloid fibrils. *Proc. Natl. Acad. Sci. U. S. A.* 109:3329–3334:10.1073/pnas.1117704109.

24. Shim, S.H., and M.T. Zanni. 2009. How to turn your pump-probe instrument into a multidimensional spectrometer: 2D IR and Vis spectroscopies via pulse shaping. *Phys. Chem. Chem. Phys.* 11:748–761:10.1039/b813817f.

25. Middleton, C.T., A.M. Woys, S.S. Mukherjee, and M.T. Zanni. 2010. Residue-specific structural kinetics of proteins through the union of isotope labeling, mid-IR pulse shaping, and coherent 2D IR spectroscopy. *Methods*. 52:12–22:10.1016/j.ymeth.2010.05.002.

26. Middleton, C.T., L.E. Buchanan, E.B. Dunkelberger, and M.T. Zanni. 2011. Utilizing

lifetimes to suppress random coil features in 2D IR spectra of peptides. *J. Phys. Chem. Lett.* 2:2357–2361:10.1021/jz201024m.

27. Ghosh, A., J.S. Ostrander, and M.T. Zanni. 2017. Watching Proteins Wiggle: Mapping Structures with Two-Dimensional Infrared Spectroscopy. *Chem. Rev.* 117:10726–10759:10.1021/acs.chemrev.6b00582.

28. Lomont, J.P., J.S. Ostrander, J.-J. Ho, M.K. Petti, and M.T. Zanni. 2017. Not All β-Sheets Are the Same: Amyloid Infrared Spectra, Transition Dipole Strengths, and Couplings Investigated by 2D IR Spectroscopy. *J. Phys. Chem. B*. 121:8935–8945:10.1021/acs.jpcb.7b06826.

29. Yung, S.K., L. Liu, P.H. Axelsen, and R.M. Hochstrasser. 2009. 2D IR provides evidence for mobile water molecules in β-amyloid fibrils. *Proc. Natl. Acad. Sci. U. S. A.* 106:17751–17756:10.1073/pnas.0909888106.

30. Grechko, M., and M.T. Zanni. 2012. Quantification of transition dipole strengths using 1D and 2D spectroscopy for the identification of molecular structures via exciton delocalization: Application to α-helices. *J. Chem. Phys.* 137:184202:10.1063/1.4764861.

31. Dunkelberger, E.B., M. Grechko, and M.T. Zanni. 2015. Transition Dipoles from 1D and 2D Infrared Spectroscopy Help Reveal the Secondary Structures of Proteins: Application to Amyloids. *J. Phys. Chem. B*. 119:14065–14075:10.1021/acs.jpcb.5b07706.

32. Petti, M.K., J.P. Lomont, M. Maj, and M.T. Zanni. 2018. Two-Dimensional Spectroscopy Is Being Used to Address Core Scientific Questions in Biology and Materials Science. *J. Phys. Chem. B*. 122:1771–1780:10.1021/acs.jpcb.7b11370.

33. Dunkelberger, E.B., L.E. Buchanan, P. Marek, P. Cao, D.P. Raleigh, and M.T. Zanni. 2012. Deamidation Accelerates Amyloid Formation and Alters Amylin Fiber Structure. *J.*

*Am. Chem. Soc.* 134:12658–12667:10.1021/ja3039486.

34. Baiz, C.R., M.E. Reppert, and A. Tokmakoff. 2013. An Introduction to Protein 2D IR spectroscopy. In: Fayer MD, editor. Ultrafast Infrared Vibrational Spectroscopy. CRC Press. pp. 361–397.

35. Reddy, A.S., L. Wang, Y.S. Lin, Y. Ling, M. Chopra, M.T. Zanni, J.L. Skinner, and J.J. De Pablo. 2010. Solution Structures of Rat Amylin Peptide: Simulation, Theory, and Experiment. *Biophys. J.* 98:443–451:10.1016/J.BPJ.2009.10.029.

36. Hwang, W., D.M. Marini, R.D. Kamm, and S. Zhang. 2003. Supramolecular structure of helical ribbons self-assembled from a β-sheet peptide. *J. Chem. Phys*. 118:389–397:10.1063/1.1524618.

37. Ngo, S.T., P.H. Nguyen, and P. Derreumaux. 2021. Cholesterol Molecules Alter the Energy Landscape of Small Aβ1–42 Oligomers. *J. Phys. Chem. B*. 16:acs.jpcb.1c00036:10.1021/acs.jpcb.1c00036.

38. Wang, H., L. Duo, F. Hsu, C. Xue, Y.K. Lee, and Z. Guo. 2020. Polymorphic Aβ42 fibrils adopt similar secondary structure but differ in cross-strand side chain stacking interactions within the same β-sheet. *Sci. Rep.* 10:5720:10.1038/s41598-020-62181-x.

# Chapter 2

# Experimental Methods

## 2.1. Introduction

In my time with the Buchanan research group, I have been privileged to experience a wide range of interdisciplinary experimental techniques. This includes alignment and optimization of multibeam laser apparatus, microwave assisted synthesis of peptides, application of dynamic light scattering (DLS), and development of protein sample preparation for electron microscopy. I have included specific details and procedures for the synthesis and purification of acetylated peptides, Additionally, I have detailed the sample conditions and preparation methods for analysis using 2D IR spectroscopy with TDS acquisition, DLS, and both transmission electron microscopy and cryogenic scanning electron microscopy. While the basic theory and data collection methods are presented in this chapter, I have included a stepwise example calculating a TDS spectrum from data collected via 2D IR spectroscopy in Appendix 2. Additionally, I have included several troubleshooting processes for the Liberty Blue peptide synthesizer I have found useful in Appendix 3.

## 2.2. Microwave assisted FMOC sold-phase peptide synthesis

Our lab is fortunate to utilize the automated Liberty Blue peptide synthesizer from CEM in the synthesis of our target protein and peptides. Traditional methods, such as bacterial expression, still hold value in many research situations. However, we find the convenience of speedy synthesis and high functionality to fit our research purposes extremely well.  Microwave assisted solid-phase peptide synthesis catalyzes the deprotection and coupling of single amino acids onto a solid resin base using microwaves to heat the reactants in the reaction chamber. The resin is usually composed of a high molecular weight nonreactive polymer such as polyethylene glycol or polystyrene. In most cases, the type of resin used is dependent on the desired C-terminus chemistry since the sequence is constructed from C-terminus to the N-terminus. In my research, I use a Rink Amide resin that yields a capped amidated N-terminus for the KFE8 family of peptides. In contrast, I use a preloaded polystyrene-based resin from CEM when synthesizing the recombinant Aβ42 sequence. Preloaded resins are purchased with the C-terminal amino acid already coupled to the polystyrene. In the case of Aβ42, the Wang resin was purchased coupled with an alanine residue. This yields the native carboxylic acid on the C-terminus. Regardless of resin type, the synthesis progresses under the same basic pattern. The resin and amino acids are all protected with an FMOC group. Since the FMOC is extremely base liable, 10% piperazine solutions are used to deprotect both resin and amino acids. The open carboxylic acids are then activated with an activator (diisoporopylcarbodimide, DIC) and an activator base (Oxyma). The resulting free amine from the resin attacks the carbonyl of the amino acid creating the peptide bond. Dimethyl formaldehyde (DMF) is used a general solvent and washes any excess reagent and biproducts through a filter to retain the resin with the newly added amino acids. The exact concentrations of piperazine, DIC, Oxyma, and amino acids are dictated by the desired synthesis scale and are calculated through the Liberty Blue software. These same cycles have been published and described in detail by associates of CEM as well as Dr. Lauren Buchanan and others from the Buchanan group(1–4).

In my work, two aspects of protein synthesis scheme are more unique in my research. The first is inducing N-terminal acetylation in the peptide sequence using the microwave assisted Liberty Blue systems. To accomplish successful acetylation, I employ 10% w/w solutions of acetic anhydride in DMF. Volumes equivalent to the additions of the amino acids are added after the final deprotection of the N-terminal amino acid. The acetic anhydride attacks the open amine and substitutes the acetyl group on the alpha carbon.

This method requires a nonreactive amino acid as reactive side chains with acid liable protecting groups may induce cleavage and side product production. The second aspect of synthesis specific to my work is the synthesis of amyloidogenic amyloid beta and the acetylated variants. Unlike the N-terminal acetylation, side chain acetylation should be performed on the desired amino acid prior to incorporation in the protein sequence. While it should be fairly simple to induce acetylation within the lysine amino acid, the acetylated lysine is readily available for purchase at reasonable prices. The purchased acetylated lysine can be directly incorporated without quantification or purification. Aβ42 presents several challenges in solid-phase synthesis due its amyloidogenic nature and composition.  While solid-phase synthesis is highly efficient, this efficiency begins to decline as the number of amino acids approach and passed 50 amino acids. Additionally, repeating amino acids that are hydrophobic significantly decrease coupling efficiency. While Aβ42 is generally amphiphilic, the amyloidogenic nature can cause the sequences to begin to fold and aggregate on the reaction column during synthesis. This makes the coupling cite less accessible to subsequent amino acids further lowing coupling efficiency. Several measures are taken to counteract this phenomenon and increase peptide product. First, a low-loading preloaded Wang resin is used. This smaller based resin takes up less room in the reaction vessel and allows for more room to build the peptide chains. Secondly, I implement two pseudo-proline dipeptide residues into the sequence at sites that have been shown to cause problematic folding of the peptide chain (Figure 2.1)(1). The pseudo-prolines remove hydrogen bond donors and induce kinks into the linear chain preventing the premature aggregation. Upon cleavage from the resin, the pseudo-prolines are broken, and the native amino acids are arranged in the sequence as designed. A more detailed description of the microwave settings for amino acid coupling, deprotection, and helpful troubleshooting are described in Appendix 3.

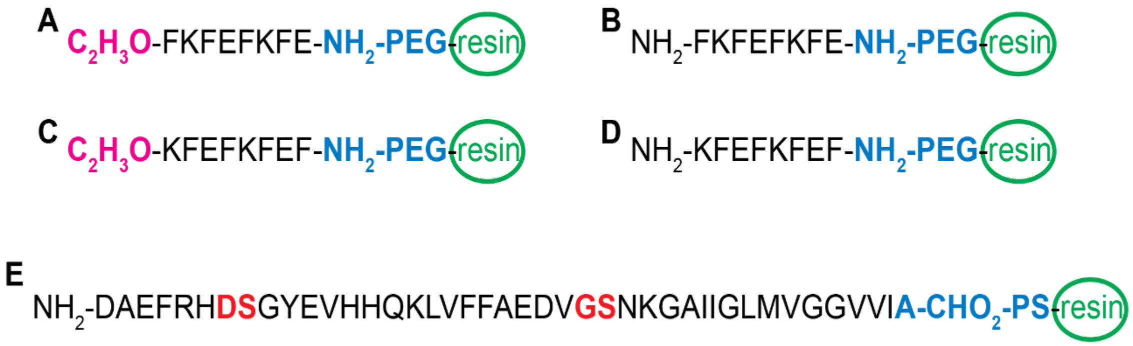

**Figure 2.1**: Primary sequence of the synthesized peptides. The KFE8 family of peptides (1.1A-D) only vary in N-terminal acetylation and the residue at the N-terminus. All KFE8 peptides were synthesized on Protide resin with a polyethalene glycol linked resulting an amidated C-terminus upon cleavage. Aβ42 (1.1E) was synthesized with a preloaded Wang resin with polystyrene linked yielding the native open carboxylic acid C-terminus when cleaved. To prevent aggregation during synthesis, pseudo-proline residues were substituted at the positions in red.

After synthesis, the resin should be dissolved in dichloromethane and transferred to a 50 mL centrifuge tube. The DCM should be slowly and carefully evaporated till the resin is fully dry. The peptide on resin is extremely stable as the amino acid side chain protecting groups are still in place. Any resin not used in the cleavage process should be stored at -20 °C. For both KFE8 peptides and Aβ42, I used a TFA based cleavage cocktail to cleave the resin and side chain protecting group. All the used amino acid side chain protection group are acid liable and therefore unaffected through the base deprotection during synthesis. An array of savaging molecules can be added to the cleavage cocktail depending on the amino acids and protecting groups being cleaved. These scavenger molecules bind to the cleaved side chain protecting groups to ensure they do not re-attach and no side products are formed. I used a solution of 90% TFA, 5% ethanedithiol, 2.5% anisole, and 2.5% thioanisole for cleavage. I have found that 10 mL of cleavage cocktail solution is sufficient to cleave a third of the resin produced in a 0.1 mmol scale synthesis. The cleavage should be carried out in a glass 6-dram vial stirred at room temperature for approximately 2.5 hours. The starting solution should be a bright yellow-green color. The color will darken overtime to a yellowish brown. This color changed signifies the cleavage of the protecting groups. Once cleavage is complete, the solution is filtered through crude Teflon filters to remove the cleaved resin. The peptide solution should be filtered directly into ice-cold diethyl ether at approximately 3 times the amount of cleavage cocktail used during cleavage. The addition of the cleavage cocktail tail to ether is extremely exothermic and should be done in an ice bath. Allow the filtered solution to chill in the ether for approx. 45 minutes. The protein solution is then centrifuged at 5000 rpm for 5 minutes. After centrifugation, the peptide should be crashed out in the bottom of the tube and the ether decanted carefully into waste. An equivalent amount of ice-cold ether should be added again, the peptide dissolved as fully as possible and placed back on the ice bath for an additional 45 minutes. These ether washes are performed for at least 3 cycles or until the thiol smell is largely dissipated. Immediately after the final ether is decanted, the peptide should be dissolved in a 50/50 v/v acetonitrile (ACN)/water solution and lyophilized. The ACN/water solution acts as a “fluffing” step and results in a pure white lyophilized powder ready for purification.

## 2.3. Purification of synthesized peptides

While the cleavage and fluffing steps are generally the same for the KFE8 peptides and Aβ42 variants, the purification of each peptide is drastically different. The purification of the short 8 amino acid KFE8 peptides is fairly straightforward via high performance liquid chromatography (HPLC). Briefly, a 2mg/mL solution of the fluffed peptides in 50/50 v/v DMSO/water solution is filtered through a 0.2 μm Teflon syringe filter into a HPLC vial, approximately 2 mL per vial. We use a prep scale reverse phase C18 column with a two-buffer gradient: solvent A being 100% water with 0.045% HCl and solvent B as 90% ACN

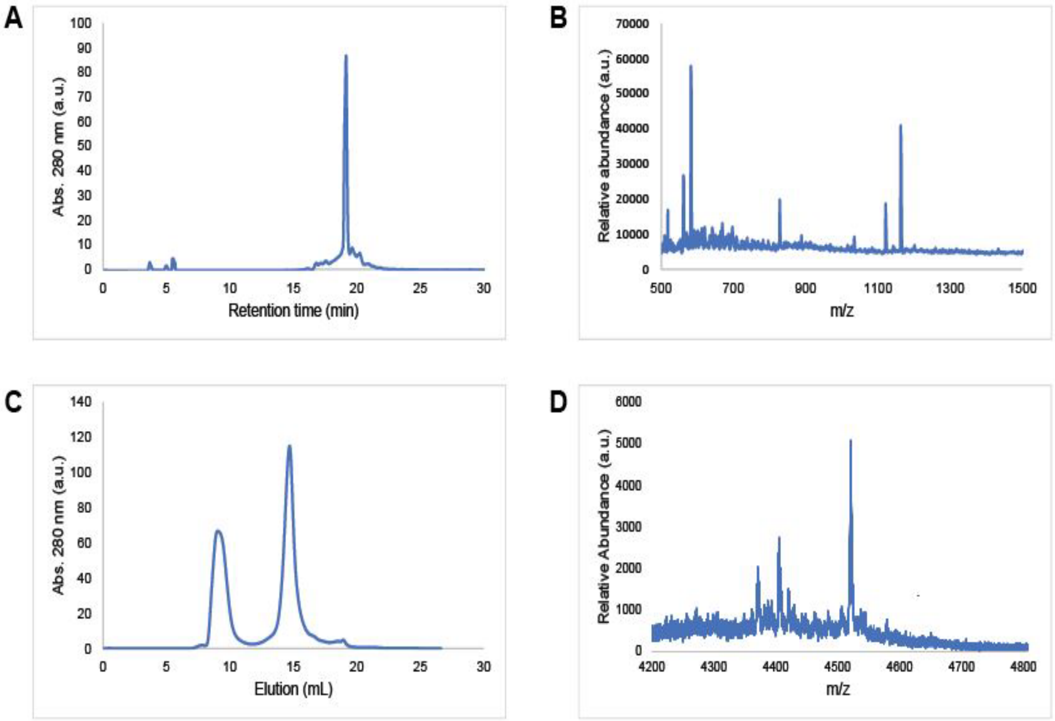

**Figure 2.2**: Example of purification and quantification of proteins. The KFE8 peptides, AcKFE8 demonstrated here, were purified via HPLC (1.2A) and the masses were check via direct inject ESI-MS (1.2B). Multiple mass-to-charge peaks are present in the spectrum as the two most prominent peaks at 1159 and 582 amu. Aβ42-WT was purified via SEC (1.2C). The aggregates peak at 10 mL elution was separated and discarded from the purified monomer peak at 15 minutes of elution. Since the Aβ42-WT protein was not soluble in appropriate solvents for ESI, we used MALDI-MS (1.2D) to check the mass shown at 4515 amu.

with 0.045% HCl. The retention time is mentored via UV-VIS wavelengths at 220 nm and 280 nm. The largest signal at 220 nm corresponds to the peptide backbone carbonyls with weak signal from the phenylalanine side chains is detected at 280 nm. With the full gradient running from 5% to100% solvent B over 30 minutes, the major protein peak elutes at approximately 16 minutes. Each peak is collected on the autosampler and saved for analysis via mass spectrometry. For the short KFE8 peptides, we use a direct injection method on an electrospray ionization mass spectrometer. The peaks corresponding to pure peptide are combined and lyophilized before re-solvating with known amounts of deionized water. The stock protein concentration is determined by UV-VIS absorbance at 280nm and calculated using Beer's law. After lyophilizing the solution, now with an accurately known peptide mass, $d_6$- hexafluroisopropanol (HFIP) is used to disaggregate the peptide at approximately 0.5 mM. This stock in HFIP is used to make aliquots for the desired experiments and lyophilized down to powder. The disaggregated peptide should be stored at -20 ºC until ready to use.

We originally aimed to purify the cleaved amyloid beta peptide through the same methods described above. However, despite trying many solvents and gradients, the peak resolution in HPLC chromatographs were very low and had corresponding low yields of purified peptides. We found amyloid beta is only largely soluble under basic conditions and would precipitate in the HCl counterion used in the gradients. We replaced HCl with 0.1% $NH_4OH$ to improve the peptide solubility. However, this did not improve peak resolution nor recovery yield. Upon recommendations from colleges and past publications, we turned to size exclusion chromatography (SEC) for purification. SEC chromatography separates molecules based on aggregates size rather than chemical affinity like HPLC. I found SEC chromatography to be far more advantageous for several reasons. First, we are able to load up to 4 mg of unpurified protein on the column with near 80% recovery over 40-minute runs. Secondly, we can separate the monomer and aggregate forms while simultaneously purifying the protein. To that end, we use a Superdex 75 30/100 SEC column with a linear gradient. The sole running buffer we found to yield the highest recovery is 20 mM Tris buffered to 8 pH. All solvents were filtered through 0.2 µM filters and degassed overnight at 4 ºC. Solvents were ran through the column at 0.5 mL per minute and elution was detected via UV-VIS at 280 nm. Just prior to loading, approximately 4 mg of Aβ42 was dissolved in 60 mM NaOH. The NaOH helps to dissolve and disaggregate the peptide to produce the most monomer possible. However, the peptide should not be left in NaOH for over 15 minutes to prevent possible side chain reactions. After collection, fractions containing the peptide are immediately desalted using Waters Oasis HLB 3cc/60 mg SPE cartridges and a vacuum manifold. The sorbent is first activated with 1 mL of methanol. The sorbent is then equilibrated with 2 mL of water and the peptide fractions loaded onto the cartridge, up to 3 mL of fractions from SEC per cartridge. The salt is washed off with 1-2 mL of water depending on the amount of peptide solution loaded unto the column. Finally, the peptide is eluted off with 2-3mL of 50/50 v/v solution of ACN/water with 0.1% $NH_4OH$. The concentration of the eluted peptide is calculated via UV-VIS and aliquots made accordingly. The individual aliquots are dissolved in HFIP at approximately 0.1 µg/µL. The peptides were left in HFIP and stored at -20 ºC until needed. Immediately before experiments, the HFIP would be lyophilized off yielding disaggregated Aβ42.

### 2.4. Dynamic light scattering methods and sample preparation

In addition to 2D IR spectroscopy and TDS calculations, I implement dynamic light scattering techniques as another form of optical analysis that tracks aggregate size over time. Since 2D IR spectroscopy only probes the secondary structure without the use of isotope-labels, it can be difficult to determine to supramolecular aggregates size during self-assembly. DLS is capable of determining the size of particles on a massive size range. This technique is applicable to particles ranging from 1 nm to approximately 6-10 µm which is extremely difficult to achieve using other techniques. This fills a gap in the detection of aggregates that are too small for some optical spectroscopy techniques and too large for electron microscopy. DLS detects the shift in scattered light by the Brownian motion of the target particle in solution yielding a hydrodynamic radius indicative of the aggregates size (5). The most precise measurements are generated from homogeneous spherical particles, a vary rare occurrence in protein systems. The hydrodynamic radius of non-spherical biomolecules is given as the radius of a sphere encapsulating the molecule(5, 6). This accounts for some degree of error when analyzing biomolecules like proteins. However, DLS is extremely sensitive to the dispersion of particles sizes, i.e., the homogeneity of the aggregates within a protein ensemble. Within this key principle is the most useful aspect of DLS as it allows for monitoring changes in aggregate size and the number of aggregates varying in size dynamically during self-assembly. Therefore, we employ DLS to analyze the KFE8 peptides during self-assembly as complementary data paired with spectra from 2D IR spectroscopy.

For DLS experiments, we use a Malvern Panalytical Zetasizer Nano Zs provided by the Vanderbilt Institute of Nanoscience and Engineering (VINSE). This instrument uses a continuous wave 800 nm laser as the light source. Rather than only detecting the scatter light from one angle, we imply multi-angle dynamic light scattering (MADLS). In this technique, the scattered light is detecting from three positions: front scatter (7°), side scatter (90°), and back scatter (170°). Each correlation function generated from the three positions are averaged yielding a much more accurate representation of the aggregates present in the protein ensemble. All protein samples were ran at the same concentration used in 2D IR spectroscopy experiments for a direct comparison of data. DLS experiments were conducted at room temperature in disposable 10 mm cuvettes. Each measurement was composed of 10 scans each integrated over 10 seconds per scan. Three measurements per peptide samples was averaged for each timepoint.

## 2.5. Electron microscopy sample preparation

### 2.5.1. Transmission electron microscopy

To visualize the supramolecular structures and motifs of the peptide aggregates, I employ several types of electron microscopy. Transmission electron microscopy (TEM) is widely used to image small peptide aggregates. However, protein and other biological samples present several inherent difficulties in electron imaging. Since the peptide aggregates themselves are not very conductive, heavy metal stains are used to create contrast in the peptide samples. EM is also conducted under extremely high vacuum so steps must be taken to ensure samples are extremely dry and adhered will to the sample grids. To acquire TEM images, we use a FEI Tecnai Osiris microscopy operating at 200 kV. Peptide samples are prepared on 400 copper mesh grids coated with formvar. This is a hydrophobic coating that helps biomolecules adhere more readily to the metal grids. I have found it helpful to have the grids glow discharged using a plasma cleaner. This creates a net positive charge on the surface of the grids and allows for a more uniform protein adhesion on the grid surface. The same peptide samples used in 2D IR experiments are immediately loaded onto the grids by adding a 10 µL droplet of peptide solution onto a flat grid suspended in the microscopy grade tweezers. The droplet is left on the gride for 120 seconds and excess solvent is wicked aways using a Kimwipe. To stain the peptides, I use a 2% depleted uranium acetate stain. A 10 µL drop of stain is placed on the grid and left on the grid for 120-240 seconds depending on the peptide concentration and morphologies to be imaged. Excess stain is then wicked off the grid and the grid placed in grid box inside a desiccator for at least 1-2 hours before inserting into microscope. The resulting micrographs are analyzed using ImageJ to determine the peptide aggregate dimensions.

### 2.5.2. Cryogenic scanning electron microscopy

As literature has suggested, imaging amyloid fibrils can be difficult and often require complex mixing of heavy metal staining techniques. Additionally, we found the Aβ42 concentration used in 2D IR spectroscopy experiment was too high for TEM measurements and therefore was not a feasible method for direct comparison of the structural motifs formed during self-assembly. To acquire high resolution images of Aβ42, we employ cryogenic SEM through VINSE. A new cryogenic sample chamber was purchased and equipped on the FEI Helios Nanolab G3CX focused ion beam scanning electron microscopy. Rather than using grids with carbon film supported by copper, we use copper rings of comparable size with an empty center. By dipping the rings in the sample protein solution used to conduct 2D IR experiments, a thin solvent film is created inside the ring like a bubble. This bubble is quickly submerged in liquid nitrogen slushy. This is liquid nitrogen that has been subject to high vacuum causing some solid nitrogen ice to form, lowering the temperature significantly. The decreased temperature allows for almost instantaneous vitrification and prevents the formation of structure ice crystals. It is imperative the ice formed remains amorphous as to not scatter the electron beam and create contrast with the peptide fibrils. Once vitrified, the copper rings are loaded on the plunger loaded holder and pushed into the cryogenic chamber. Here, the sample stage is gradually heated to allow the top layer of ice to evaporate, exposing the peptide structures. The sample is then sputter coated with pure platinum to make the peptide fibrils connective to the electron beam. Once the sample is coated, the sample is moved into the SEM stage and imaged. The rendered micrographs yield a much higher resolution of larger peptide morphologies compared to native TEM or SEM images (Figure A1.12).

### 2.6. Calculating transition dipole strength via 2D IR

The mathematical derivation of obtaining TDS values from infrared spectroscopy, through a mix of FTIR and 2D IR as well as solely through 2D IR, has been previously described in detail(7–9). Here, I want to highlight the methods I used to calculate TDS spectra completely from 2D IR spectroscopy with pulse shaping technology. It is important to note peptide systems that have very similar 1D and 2D spectra may be poor candidates for TDS calculations as the accuracy of the values relies on the spectra subtraction.

$$OD = -\log\left(\frac{voltage_{sample}}{voltage_{solvent}}\right) \qquad 2.1$$

First, 2D IR spectra of the protein sample and a calibrant molecule as well as a blank sample corresponding to each solvent (buffer or pure $D_2O$) are collected per the usual conditions. In this case, I use spectra with 20 scans, each scan is averaged over 60 individual scans collected within approximately 60 seconds each. Calibrant molecules are chosen with known, well-established TDS values and absorbance frequency similar to the target protein. I use 40 L-Serine (Ser) in $D_2O$ as the calibrant spectrum. The voltage (1D spectrum) of both the sample and buffer are then extracted from the collected data. This requires an array detector. In this case we use a Mercury Cadmium Tellurium detector cooled to cryogenic temperature. Using the voltages, the optical density (OD) is calculated via equation 2.1.

$$|\mu|^2 = \frac{\frac{\Delta OD_{sample}}{OD_{sample}}}{\frac{\Delta OD_{\max(calibrant)}}{OD_{\max(calibrant)}}} \times |\mu_{calibrant}|^2 \qquad 2.2$$

The generated OD spectrum will likely have a distorted baseline. The OD spectrum is fit to a second order polynomial trendline near the frequency of the sample or calibrant and subtracted to flatten the baseline. Using the same method, the baseline correction should also be performed for the calibrant OD spectrum. Next, the diagonal slice of the 2D IR spectrum of both the calibrant and target protein is acquired of the ΔOD. The ΔOD and OD generated spectra of the sample are divided to yield a TDS spectrum of the sample. Rather than fully divide the calibrant spectra, the maxima values from the ΔOD and OD spectra is found and divided as a single value used a scalar. Finally, the calculated TDS sample spectrum is multiplied by the known TDS value of the calibrant. The TDS value of L-serine used herein is 0.2 $D^2$. The final equation for calculated TDS spectrum is shown in equation 2.2.

## 2.7. References

1. Vanier, G.S. 2013. Microwave-assisted solid-phase peptide synthesis based on the fmoc protecting group strategy (CEM). *Methods Mol. Biol.* 1047:235–249:10.1007/978-1-62703-544-6_17/FIGURES/00172.

2. Coantic, S., A.E. Gilles, S. Ae, and J. Martinez. Microwave-assisted Solid Phase Peptide Synthesis on High Loaded Resins. 10.1007/s10989-008-9123-6.

3. Webb, K. 2022. Examining Peptide and Protein Structure and Dyanmics with Two-Dimensional Infrared Spectroscopy. .

4. DeNeve, D. 2019. Probing Molecular Structure with Two-Dimensional Infrared Spectroscopy. .

5. Arzenšek, D., R. Podgornik, and D. Kuzman. 2010. Dynamic light scattering and application to proteins in solutions. In: Seminar. . pp. 1–18.

6. Aggeli, A., G. Fytas, D. Vlassopoulos, T.C.B.B. McLeish, P.J. Mawer, and N. Boden. 2001. Structure and dynamics of self-assembling β-sheet peptide tapes by dynamic light scattering. *Biomacromolecules*. 2:378–388:10.1021/bm000080z.

7. Grechko, M., and M.T. Zanni. 2012. Quantification of transition dipole strengths using 1D and 2D spectroscopy for the identification of molecular structures via exciton delocalization: Application to α-helices. *J. Chem. Phys.* 137:184202:10.1063/1.4764861.

8. Lomont, J.P., J.S. Ostrander, J.-J. Ho, M.K. Petti, and M.T. Zanni. 2017. Not All β-Sheets Are the Same: Amyloid Infrared Spectra, Transition Dipole Strengths, and Couplings Investigated by 2D IR Spectroscopy. *J. Phys. Chem. B*. 121:8935–8945:10.1021/acs.jpcb.7b06826.

9. Dunkelberger, E.B., M. Grechko, and M.T. Zanni. 2015. Transition Dipoles from 1D and

2D Infrared Spectroscopy Help Reveal the Secondary Structures of Proteins: Application to Amyloids. *J. Phys. Chem. B*. 119:14065–14075:10.1021/acs.jpcb.5b07706.

# Chapter 3

# Investigating the effects of N-terminal acetylation on KFE8 self-assembly with 2D IR spectroscopy

## 3.1. Abstract

Peptide self-assembly is an exciting and robust approach to create novel nanoscale materials for biomedical applications. However, the complex interplay between intra- and intermolecular interactions in peptide aggregation means that minor changes in peptide sequence can yield dramatic changes in supramolecular structure. Here, we use two-dimensional infrared spectroscopy to study a model amphiphilic peptide, KFE8, and its N-terminal acetylated counterpart, AcKFE8. Two-dimensional infrared spectra of isotope-labeled peptides reveal that AcKFE8 aggregates comprise two distinct β-sheet structures although KFE8 aggregates comprise only one of these structures. Using an excitonic Hamiltonian to simulate the vibrational spectra of model β-sheets, we determine that the spectra are consistent with antiparallel β-sheets with different strand alignments, specifically a two-residue shift in the register of the β-strands. These findings bring forth new insights into how N-terminal acetylation may subtly impact secondary structure, leading to larger effects on overall aggregate morphology. In addition, these results highlight the importance of understanding the residue-level structural differences that result from changes in peptide sequence to facilitate the rational design of peptide materials.

## 3.2. Introduction

Self-assembling peptides are promising candidates for biomedical applications (1, 2) in areas such as drug or vaccine delivery (3–5), tissue regeneration (6–8), and bioimaging(9), due to their diverse functionalization and amphiphilicity (10–15).Complex drug-delivery systems combine carrier molecules and targeting ligands with the desired drug, making tunable self-assembling peptides ideal candidates for these applications. However, it can be difficult to predict a final three-dimensional structure, much less functionality, based solely on a linear sequence of amino acids. Studies have shown that relatively small changes to the sequence of amino acids can dramatically alter the final self-assembled nanostructures (16–21). For example, Xu et al. reported a series of cationic peptide surfactants $A_mK$ (m= 3, 6, 9) that differ only in the repeat length of the alanine chain (22). These peptides self-assemble into three distinct nanostructures, ranging from flat nanosheets ($A_3K$) to long nanofibers ($A_6K$) to short nanorods ($A_9K$). In another example, Cui et al. studied four constitutional isomers based on the tetrapeptide, $V_2E_2$, attached to an alkyl tail (23). They found that molecules with alternating hydrophobic and hydrophilic residues (VEVE or EVEV) self-assemble into flat nanostructures although non-alternating isomers (VVEE and EEVV) form cylindrical nanofibers. The KFE8 family of octapeptides, which comprises varying sequences of lysine, phenylalanine, and glutamic acid residues, is a classic example of the complexity in predicting sequence-structure relationships. Numerous studies have explored the effects of varying the patterning of the residues, changing the identity of the charged residues, introducing heterochiral blocks, or capping the termini (20, 21, 24–27). Although all variants form extended β-sheet structures, their supramolecular morphologies exhibit varying lengths, widths, and degrees of helicity or twisting. One particularly interesting variant, AcKFE8 ($COCH_3$-FKFEFKFE-$NH_2$), forms a left-handed helical ribbon. Researchers have explored the potential of this unique chiral morphology for biomedical applications (2, 8, 9, 28): in vivo, AcKFE8 exhibits an adjuvant effect when conjugated to other epitope-targeting domains (4, 29). Since neither AcKFE8 nor the various conjugated substrates have shown adjuvant effects alone, current research suggests self-assembling peptides are the key in activation of immunological responses of the epitope species. The supramolecular morphology of AcKFE8 has been well characterized (21, 24, 30) using imaging techniques such as atomic force microscopy and transmission electron microscopy (TEM); however, the detailed arrangement of the peptides within the aggregates remains elusive. Although Fourier-transform infrared (FTIR) and circular dichroism spectra indicate that the monomers are forming antiparallel β-sheets, the specific alignment of the strands has not been experimentally determined (24, 31). Thus, it is unclear why AcKFE8 forms a unique left-handed helical ribbon although its non-acetylated analogue, KFE8 ($NH_2$-FKFEFKFE-$NH_2$), forms ribbons with a flat morphology. Molecular dynamics (MD) simulations suggest that the AcKFE8 aggregates comprise two antiparallel β-sheet ribbons that stack to create a double helical morphology with a 20-nm pitch (24, 25). From these studies were proposed two possible antiparallel β-sheet conformers for these helical ribbons, varying only in the register of the strands, but there exists no direct experimental evidence for the proposed structures. The unique morphology of AcKFE8 demonstrates immense potential for biomedical applications, making it imperative to understand precisely how acetylation affects the underlying organization of peptide monomers to create this left-handed helical ribbon and thus elucidate how other peptide sequences could be driven to adopt such a structure.

In this study, we use two-dimensional infrared (2D IR) spectroscopy and isotope labeling to extract structural details of aggregates of both AcKFE8 and its non-acetylated analogue, KFE8, for comparison. 2D IR spectroscopy has been used extensively to characterize the structure and aggregation dynamics of other peptides that form β-sheet-rich aggregates, including amyloid fibrils (32–40). When peptides form organized secondary structures, vibrational coupling between backbone amide I’ modes results in frequency shifts that can be used to analyze both secondary and tertiary structures (41–44). These vibrational couplings are highly sensitive to both the distance and orientation between backbone residues. Thus, when β-sheet peptides are prepared with site-specific $^{13}C^{18}O$ labeling of individual backbone carbonyls, shifting the vibrational frequency of the labeled amide I0 mode by 55 $cm^{-1}$ (45–47), we can determine the precise alignment of β-strands with single-residue resolution. 2D IR spectra of labeled AcKFE8 exhibit two isotope-labeled peaks compared with a single peak present in KFE8. Using transition dipole coupling (TDC) calculations of model β-sheets to simulate vibrational spectra, we demonstrate that these spectral features are consistent with one of the structures previously proposed by MD simulations. These results demonstrate that 2D IR spectroscopy is a powerful tool for studying protein self-assembly that provides molecular insights that would be difficult or impossible to obtain with traditional experimental techniques.

## 3.3. Materials and methods

### 3.3.1. Materials

All reagents were used as purchased, excluding modifications made to 1-$^{13}$C-phenylalanine-OH, as described in the section on preparation of $^{13}C^{18}O$-phenylalanine. Piperazine was purchased from Alfa Aesar (Ward Hill, MA, USA). Sodium bicarbonate (NaHCO3), potassium bisulfate (KHSO4), dimethylformamide, diethyl ether, acetone, acetonitrile, hydrochloric acid (12.1 N), acetic anhydride, 4-nitrobenzaldehyde, and methanol were purchased from Fisher Scientific (Pittsburgh, PA, USA). N-9-fluoroenylmethoxycarbonyl (FMOC)-protected amino acids, Oxyma, and Rink Amide ProTide resin were purchased from CEM (Matthews, NC, USA). N,N0 -diisopropylcarbodiimide was purchased from Oakwood Chemical (Estill, SC, USA). 1-$^{13}$C-L-phenylalanine (99% enriched), $^{18}$O-H2O (98% enriched), and $D_2O$ (99% enriched) were purchased from Cambridge Isotope Laboratories (Tewksbury, MA, USA). 2% uranyl acetate solution was purchased from Electron Microscopy Sciences (Hatfield, PA, USA). All other reagents were purchased from Millipore Sigma (Burlington, MA, USA).

### 3.3.2. Solid-phase peptide synthesis and purification

AcKFE8 and KFE8 were synthesized on a Liberty Blue microwave peptide synthesizer (CEM, Matthews, NC, USA) using standard FMOC solid-phase peptide synthesis with piperazine deprotection and diisopropylcarbodiimide/Oxyma activation (48). Rink Amide ProTide resin was used as the solid support to produce an amidated C-terminus in both peptides. After final deprotection, 10% acetic anhydride was used to acetylate the N-terminus of AcKFE8 before cleaving the peptide off the resin. A solution of 95% trifluoroacetic acid, 2.5% triisopropylsilane, and 2.5% deionized water was used to cleave the peptides from the resin and remove side-chain protecting groups. The resin was removed via filtration, followed by precipitation of the crude peptide with ice-cold diethyl ether. Crude peptide was dissolved in a 50/50 water/acetonitrile solution and purified via reverse-phase high performance liquid chromatography (Ultimate 3000, Thermo Fisher Scientific, Waltham, MA, USA) with a binary gradient of water (solvent A) and 100% acetonitrile (solvent B) with 0.045% HCl (v/v) as the counterion. The gradient was varied from 0% to 90% Solvent B over 40 minutes whereas ultraviolet-visible absorbance was monitored at 214 nm and 280 nm. AcKFE8 eluted around 15 minutes whereas KFE8 eluted around 18 min. Molecular weight and purity of both peptides were confirmed via electrospray ionization mass spectrometry (Orbitrap XL Penn, Thermo Fisher, WA, USA). Purified peptide was lyophilized (Labconco, Kansas City, MO, USA) and stored at 20 ºC. Isotope-labeled peptides were synthesized and purified in the same manner, using $^{13}C^{18}O^{18}O$-phenylalanine-OH, preparation described below, at the Phe-5 position during solid-phase peptide synthesis.

### 3.3.3. Preparation of $^{13}C^{18}O$-phenylalanine

FMOC protection was added to the 1-$^{13}C$-phenylalanine using a 1:1:1 mol ratio of 1-$^{13}C$-phenylalanine-OH, $NaHCO_3$, and FMOC-succinimide, which was stirred overnight at room temperature in a 50:50 deionized water/acetone solution. The reaction was quenched with 2 M potassium sulfate to a pH of 2. The precipitated amino acid was vacuum filtered and washed sparingly with ice-cold deionized water. To introduce $^{18}$O labeling to the FMOC-protected amino acid, an acid-catalyzed $^{18}$O exchange was conducted under N2 atmosphere on a Schlenk line as previously reported (48). Briefly, 1 g of FMOC-1-$^{13}$C-phenylalanine-OH was reacted with 1 g $^{18}$OH$_2$ in 10 mL dioxane and 2 mL HCl in dioxane. The reaction was refluxed at 150 ºC for 6 h. Most of the solvent was removed under vacuum on the Schlenk line followed by lyophilization until completely dry. The process was repeated once to achieve R 90% $^{18}$O labeling efficiency, as verified by electrospray ionization mass spectrometry. The final labeled amino acid was washed with 10 mL of cold 3:1 (v/v) ethyl acetate:hexane solution. Most of the solution was removed via rotary evaporation. The amino acid was precipitated with cold 8:1 (v/v) hexane:ethyl acetate solution, vacuum filtered, and lyophilized to completely dry powder.

### 3.3.4. Sample Preparation

All peptide samples were prepared using deuterated solvents to ensure the amide I' mode was not obstructed by the strong water bending mode in 2D IR measurements. Stock solutions were prepared by dissolving the purified proteins in D$_2$O. After determining the concentration of the stock solution by using the absorbance at phenylalanine at 280 nm (49), the peptide was lyophilized and redissolved in deuterated hexafluoroisopropanol (d-HFIP) (98% enriched) at a concentration of 1 mM. The d-HFIP solution was allowed to sit overnight before sonication for 4 h to ensure that the peptides were disaggregated, and the backbone amides were fully deuterated. Finally, d-HFIP was removed via lyophilization, and the dry powder stored at 80 ºC. In accordance with previously published studies, samples were prepared by dissolving the lyophilized protein in unbuffered $D_2O$ at a final pD of 3 (24, 25, 30). As AcKFE8 has been shown to form the left-handed helical aggregates at concentrations ranging from 0.1 to 4 mM (30), a concentration of 1 mM was selected for these studies to achieve an acceptable signal-to-noise ratio in all experiments.

### 3.3.5. TEM

TEM micrographs were taken using a Tecnai Osiris transmission electron microscope (FEI, Hillsboro, OR, USA) operating at 200 kV. Ultrathin 400 copper mesh grids with formvar (Ted Pella, Redding, CA, USA) were dip coated in a 1 mM peptide solution to adhere peptide to grids. The grids were stained with 2 uranyl acetate solutions and dried overnight. ImageJ, a free Java-based image processing program developed by the National Institutes of Health and Laboratory for Optical and Computational Instrumentation (University of Wisconsin), was used to process all TEM micrographs. Width and pitch measurements of all peptide ribbons were determined for 100 individual peptide ribbons to calculate averages and standard deviations with a 95% confidence interval.

### 3.3.6. DLS

Dynamic light scattering measurements were performed on a Nano ZS zetasizer (Malvern Panalytical, Worcestershire, UK). Measurements were averaged over 300 scans and analyzed using a count size distribution. Malvern Analytical software calculates the average number of particles in solution within a size range based on the intensity of the light scattered within the size range using a size correlation function. This size distribution is plotted via a Gaussian function. The zetasizer utilizes an 800-nm continuous wave laser with a back-scatter detector at a 45º angle from samples. Ten centimeters pathlength polystyrene cuvettes (Fisher Scientific, Pittsburgh, PA) were used to measure samples.

### 3.3.7. 2D IR spectroscopy

A detailed description of 2D IR data collection and processing methods are described elsewhere (48). Briefly, 800-nm pulses (7 mJ, 1 kHz, 60 fs) were generated by a single box ultrafast amplifier (Solstice, SpectraPhysics, Milpitas, CA, USA). A 50/50 beamsplitter was used to direct half of the beam to pump an optical parametric amplifier with difference frequency generation (TOPAS-Prime, SpectraPhysics, Milpitas, CA, USA). The resulting mid-IR light (6,100 nm, 25 mJ, 1 kHz, 70 fs) was directed into the 2D IR spectrometer (2DQuick IR, PhaseTech Spectroscopy, Madison, WI, USA). 2D IR spectra were collected using a parallel (ZZZZ) beam polarization. The t1 time delay between pump pulses was scanned from 0 to 2.54 ps in 23.8-fs steps, although the t2 waiting time between pump and probe was held constant at 0 fs. The signal was directed into a monochromator (Princeton Instruments, Trinton, NJ) and dispersed onto a mercury cadmium telluride focal-plane array detector (PhaseTech Spectroscopy, Madison, WI, USA), which can achieve a spectral resolution of 2.1 cm$^{-1}$. The data was collected using the QuickControl software provided by PhaseTech and processed using custom MATLAB scripts. Peptide samples were measured by placing 5 mL of peptide solution between two $CaF_2$ windows (Crystran, Poole, Dorset, UK) separated by a 50-mm Teflon spacer.

### 3.3.8. Construction of model β-sheet aggregates

Flat antiparallel β-sheets were constructed from β-strands of eight residues each with an interstrand distance of 4.77 A˚. To simulate different strand registers, alternating strands were shifted by integer multiples of 6.5 A˚, which corresponds to the length of two residues in a β-sheet. Methods for simulating the helical β-sheet ribbons were adapted from (24). Helical aggregates were constructed by taking flat β-sheets comprising 200 strands and positioning the first strand at the desired radius (inner radius = 22.5 A˚; outer radius = 35.4 A˚) along the y-axis. The β-sheet extended along the x-axis, with individual strands oriented parallel to the z-axis and shifted so that the middle of the strand was at z = 0. The β-sheet was then rotated about the y-axis by an angle such that the pitch of each helix would be 19.4 nm. The helix radius (r) and the pitch (h) are related by Equation 3.1,

$$r = \frac{h}{2\pi tan\theta} \tag{3.1}$$

where θ is the pitch angle. To maintain a common pitch between the two helices, the y-axis rotation angle was 53.9º for the inner helix and 41.1º for the outer helix. Each strand in a β-sheet was then individually translated in the x-direction such that the middle of the strand had an x-coordinate of 0 followed by a rotation about the z-axis to produce the helical structure. The purpose of the translation before rotation was to obtain the desired radius after rotation (by construction, only the first strand in the β-sheet had the proper radius to the z-axis). The angles that were used to rotate about the z-axis were calculated to maintain 4.77 Aº between consecutive strands. Each strand in the inner and outer helix was rotated by an additional 7.2º and 5.8º, respectively, relative to the previous strand in the helix. The way the helix was constructed resulted in different hydrogen bond distances because successive strands were no longer parallel to one another. In a flat β-sheet with an interstrand distance of 4.77 A˚, the hydrogen bond distances between adjacent strands were 1.8 A˚. However, after the strands were rotated to create the helix, the hydrogen bond distances reached up to 2.3 A˚. We employed a simple steepest descent energy minimization to better align the strands and alleviate the elongated hydrogen bond distances. The energy minimization was completed using GROMACS v.2020.3 with standard GROMOS 53a6 bond, angle, and dihedral parameters within each strand. The interstrand hydrogen bond distances were fixed using a Lennard-Jones interaction ($\sigma$= 0.255 nm ε =2.0 kJ/mol) between the amide nitrogen on one strand and the carbonyl oxygen on a neighboring strand. These parameters were chosen because they had minimal effect on the strand structure and hydrogen bond distances in the β-sheet structure. The energy minimization on the double helix decreased the maximum hydrogen bond distance to roughly 2.0 Aº. Although this distance could be decreased further by increasing the ε Lennard-Jones parameter, we decided against this because larger ε parameters could lead to more significant changes to the overall structure of each strand and the purpose of the energy minimization was only used to slightly re-align the strands after the helix was constructed.

### 3.3.9. Simulation of vibrational spectra

FTIR and 2D IR spectra were calculated using COSMOSS, an open-source Matlab script available on GitHub that simulates vibrational spectra by creating an excitonic coupling Hamiltonian from structural input files (50). Flat aggregates were calculated using 20 strands, although helical aggregates were simulated using approximately 105 strands (approximately 50 strands per sheet, equivalent to one full helical turn). Unlabeled residues were assigned a local mode frequency of 1650 $cm^{-1}$ based on the experimental uncoupled amide I0 frequency, although $^{13}C^{18}O$-labeled residues were redshifted by 55 $cm^{-1}$ to 1595 $cm^{-1}$ (41). Coupling constants between residues were calculated according to a TDC model. However, this model overestimates coupling between adjacent, covalently bound residues, so nearest neighbor couplings were set to a value of 0.8 $cm^{-1}$, which was derived from quantum mechanical calculations (47, 51). To account for sources of environmental disorder, such as hydrogen bonding between the amide groups and water (52), the local mode frequencies were varied randomly around the assigned values using a Gaussian distribution with a full width at half maximum of 10 $cm^{-1}$ (36). Simulated spectra were generated by averaging over 100 samples. The calculated spectra exhibited a consistent 12–16 $cm^{-1}$ red shift of both the unlabeled and labeled amide I’ modes compared with experimental spectra, suggesting that our calculations consistently predict stronger vibrational couplings than we observe experimentally. This is not surprising, as the calculations are performed on ideal β-sheets models with no structural disorder. Real peptide aggregates likely exhibit some amount of structural disorder that alters the distance and angles between vibrational modes. Such structural variations would change the coupling strengths between residues, an effect that is not accounted for in our calculations (47). We have applied a constant 14 $cm^{-1}$ blue shift to the calculated spectra to allow for more straightforward comparison with experiments.

## 3.4. Results and discussion

Although the size and supramolecular morphology of AcKFE8 aggregates have been reported previously, KFE8 has yet to be investigated under similar conditions. The difference in supramolecular structure during and after aggregation is immediately apparent when imaged with TEM (Fig. 3.1). In bright-field TEM, regions of higher density appear darker than regions of lower density. Thus, the regions of alternating light and dark contrasts along the length of the aggregated AcKFE8 ribbons (Fig. 3.1 A) indicate periodic changes in thickness, a clear indication of rotation. The use of light contrasts in TEM micrographs has been used previously as an accurate means to measure protein periodicity (53–55). The AcKFE8 twisted ribbons were determined to have widths of 15.7 ± 5.4 nm (Fig. 3.1 B, solid line) and pitches of 14.5 ± 0.5 nm. In contrast, the non-acetylated variant self-assembled to form narrow, flat peptide ribbons with widths of 8 ± 1.3 nm (Fig. 3.1, C and D, solid line), with no change in contrast along the length of the aggregates. The flat ribbons also showed a tendency to stack and aggregate into larger ribbon bundles

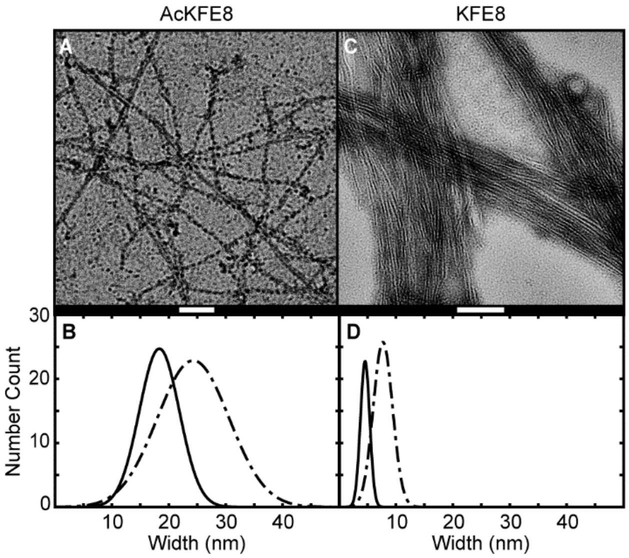

**Figure 3.1**: TEM micrographs of final morphology of AcKFE8 (A) and KFE8 (C) after 24 h aggregation. White scale bars are equal to 100 nm. Size distribution of peptide ribbon widths comparing measurements from DLS (dotted dashed line) and TEM micrographs above (solid lines) of AcKFE8 (B) and KFE8 (D) after 24 h aggregation.

that can reach over 100 nm in width, although helical ribbons remain separate in solution. These aggregate structures were found to persist for at least 1 week after initial aggregation (appendix 1, figure S1.1).

Previous studies have demonstrated that DLS is capable of detecting protein aggregates before they are clearly visible by electron microscopy (56). Therefore, to better understand the evolution of AcKFE8 and KFE8 aggregates, we used DLS to track the hydrodynamic radius of the protein particles at discrete timepoints throughout a 24-h aggregation window (57). DLS analysis generally assumes that all particles in the scattering volume can be approximated as spheres. Clearly, this is not the case for the ribbon aggregates observed via TEM. However, DLS studies of amyloid protein fibers commonly approximate the fibers as cylinders, in which case the effective hydrodynamic radius calculated from DLS can be interpreted in terms of the fibers' width (58–60). The fiber lengths are not reported in these measurements, as they exceed the limit of detection for DLS. As one would expect, the hydrodynamic radii of the peptide ribbons after 24 h of aggregation are larger than the mean widths measured by TEM (Fig. 3.1, B and D). The general trend still holds, however, as the AcKFE8 ribbons are wider and have a broader distribution than those formed by KFE8. DLS measurements from earlier time points showed no significant variation from the measurements at 24 hours; thus, we conclude that the β-sheet ribbons form rapidly (in less than 5 min after initiating aggregation) and their supramolecular morphology remains unchanged, even when the samples were rechecked after a week of aggregation. Although TEM and DLS reveal distinct differences in the size and supramolecular morphology of the peptide aggregates, they do not possess sufficient structural resolution to indicate why N-terminal acetylation causes such dramatic changes in the aggregate structures.

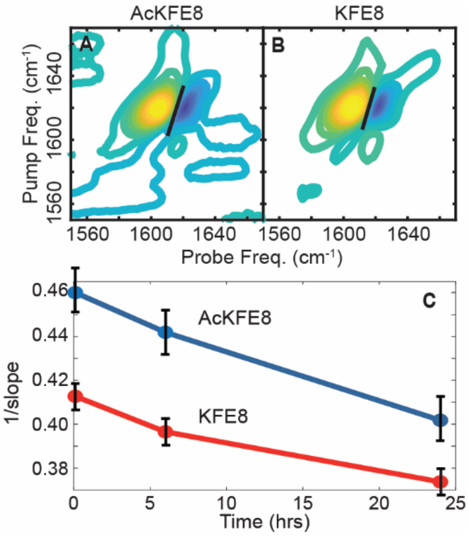

**Figure 3.2**: 2D IR contour maps of AcKFE8 (A) and KFE8 (B) with nodal slope line (black line). Inverse nodal slope regression (C) of AcKFE8 (blue line) and KFE8 (red line) over 24 h aggregation is shown. Error bars for each time point represent the standard deviation over three measurements.

To determine the residue-level structural differences that underlie the dramatically different supramolecular morphologies of AcKFE8 and KFE8, we collected 2D IR spectra of both peptides at three time points during aggregation: 5 min, 6 h, and 24 h. Figure 3.2

shows the 2D IR spectra for each peptide variant after 24 h. In these spectra, peaks appear in pairs along the vertical pump axis; within each pair, the blue peak represents the $|0 \rightarrow 1|$ fundamental vibrational transition and the yellow peaks represents the $|1 \rightarrow 2|$ overtone vibrational transition. Both AcKFE8 (Fig. 3.2 A) and KFE8 (Fig. 3.2 B) exhibit strong amide I' peak pairs at a probe frequency of 1622 $cm^{-1}$. This vibrational frequency is characteristic of highly ordered β-sheet structures (41, 44) and is present immediately after self-assembly is initiated (appendix 1, Fig. S1.1). In soluble monomeric proteins that adopt β-sheet structures, the frequency of the amide I' mode tends to correlate directly with the number of β-strands: as the number of strands increases, the amide I' mode is further red shifted due to increasing delocalization of the vibration across the β-strands (61). For large, β-sheet-rich protein aggregates, however, the β-sheets are sufficiently extensive that the frequency shift appears to reach an asymptotic limit. Instead, the precise amide I' frequency of these aggregates is determined by a variety of structural factors that affect vibrational coupling strengths, including interstrand spacings and the relative orientations of the amide I' groups (38). The nearly identical 2D IR spectra in Fig. 3.2, A and B suggest that residues in AcKFE8 and KFE8 experience, on average, nearly identical vibrational couplings and thus must adopt very similar β-sheet structures.

Although the vibrational frequencies were identical for both peptides and remained constant over 24 h of aggregation, lineshape analysis revealed slight differences between variants. Here, we use inverse slope, defined as the 1/slope of the nodal line between the fundamental and overtone peaks, to quantify the inhomogeneity of the 2D IR lineshapes. Inverse slope values can vary between 0 (for purely homogenous, round lineshapes) and 1 (for inhomogeneous lineshapes that are elongated along the diagonal). Samples rarely exhibit purely homogeneous or inhomogeneous lineshapes, and thus, the inverse slope serves as a measure of the ratio of inhomogeneous to homogenous contributions (62).

The spectral lineshapes, which reflect the distribution of vibrational frequencies within the sample, are directly related to the structural distribution of the proteins (36, 41). The inverse slope of both variants decreases over the course of 24 h (Fig. 3.2 C). As the frequency of the amide I' vibrational mode does not change during this period, we attribute the increased lineshape homogeneity to a subtle annealing of the ribbon structures, resulting in increased structural homogeneity, rather than a significant change in the structure of the β-sheets (38). In addition, AcKFE8 exhibits a consistently higher inverse slope than KFE8, indicating a higher degree of inhomogeneity throughout the course of aggregation. Thus, although amide I' frequencies suggest that AcKFE8 and KFE8 must adopt nearly identical β-strand configurations, the lineshapes reveal that AcKFE8 monomers experience more variation around this ‘‘average'' structure. This result agrees with the measured TEM and DLS distributions and indicates that AcKFE8 is more heterogeneous on both the molecular and supramolecular scale. The higher inverse slope could signal the presence of multiple, slightly different strand alignments, as suggested by MD simulations (24–26).

To test this hypothesis, we employ site-specific isotope labeling to probe the detailed alignment of the β-strands in each aggregate. Isotope labels are ideal structural probes, as they do not perturb peptide structure or dynamics. Introducing a $^{13}C^{18}O$ label into a backbone amide I' group redshifts the frequency of that oscillator by 55 cm$^{-1}$, isolating it from the other residues in the peptide (36, 39, 41, 63). This allows the vibrational couplings between specific residues, which are exquisitely sensitive to both the distance and relative orientation of the amide I' groups (44, 64), to be measured directly. In one of the strand alignments proposed from MD simulations, the N-terminus is slightly exposed (Fig. 3.3 A), although in the second strand alignment, the register of the β-strands is shifted by two residues such that the C-terminus is exposed (Fig. 3.3 B). Comparing these strand

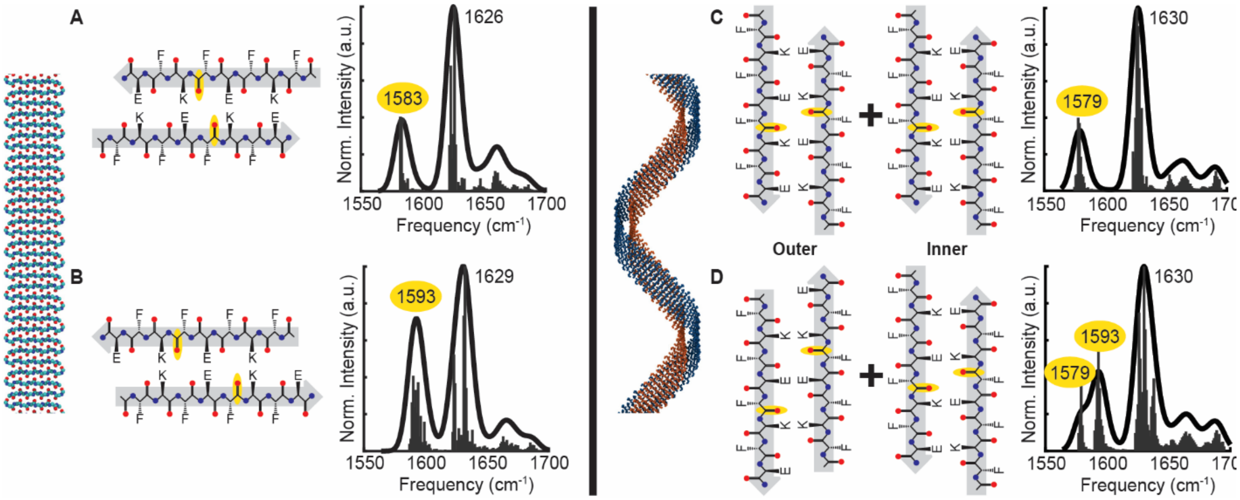

**Figure 3.3**: Simulated antiparallel β-sheet structures and corresponding calculated FTIR spectra. Flat β-sheets were simulated from either N-terminal exposed (A) and C-terminal exposed (B) strand alignments. Left-handed double helical β-sheets were simulated using the N-exposed strand alignment for the inner β-sheet and either the N-exposed (C) or C-exposed (D) strand alignment for the outer β-sheet. The location of the $^{13}C^{18}O$-labeled F5 residue is highlighted in yellow in each β-strand. The frequencies of the unlabeled and $^{13}C^{18}O$-labeled amide I' modes are labeled in the simulated spectra, with the isotope-labeled mode highlighted in yellow.

alignments, we can see that residue F5 is well aligned across the strands in the N-exposed alignment but staggered between strands in the C-exposed alignment. Thus, we would expect residue F5 to experience strong coupling (+9.6 cm$^{-1}$ according to TDC) in the N-exposed alignment, although coupling will be negligible (-0.2 cm$^{-1}$ according to TDC) in the C-exposed alignment.

To confirm the sensitivity of residue F5 to the two strand alignments, we simulated their vibrational spectra (50). The isotope-labeled mode appears at 1583 cm$^{-1}$ in the N-terminal strand alignment (Fig. 3.3 A), which is significantly red shifted compared with the 1593 cm$^{-1}$ peak calculated for the C-exposed strand alignment (Fig. 3.3 B). For comparison, we calculated an isotope-diluted spectrum in which only a single strand contained the $^{13}C^{18}O$ label at Phe-5 (Fig. 3.5 A); in this scenario, all coupling to the isotope-labeled amide I’ mode is eliminated, and the labeled mode appears as a broad, weak feature at its native or local mode, frequency of 1594 $cm^{-1}$ (33, 65). These results match our predictions that residue F5 is strongly coupled in the N-exposed alignment and virtually uncoupled in the C-exposed alignment. Thus, the coupling at residue F5 should serve as a sensitive reporter of the β-strand alignment in both variants.

Although these were the most favorable structures according to MD simulations (24, 25), we tested other strand alignments to ensure that 2D IR spectra could distinguish between any possible structures. Two new structures were simulated with an additional two-residue shift in register toward either the N- or C-terminus; single-residue register shifts were neglected as they would eliminate the favorable pi-stacking interactions between phenylalanine residues. The calculated spectra exhibit significant disordered features, as increased staggering of the strands causes the termini to extend far past the core β-sheet structure (Fig. S4, A and B). The experimental 2D IR spectra show minimal contributions from disordered structures (Fig. 3.4, A and B); as such, these additional structures are discounted.

To determine how helical twisting of the β-sheet aggregates would alter coupling, we simulated the vibrational spectra of the two helical β-sheet aggregates that have been proposed as likely structures for AcKFE8 (24, 25). In both models, the inner β-sheet is formed with the N-exposed strand alignment; the models differ, however, in whether the outer β-sheet contains N-exposed (Fig. 3.3 C) or C-exposed (Fig. 3.3 D) strands. The latter model was initially proposed (24), but subsequent MD studies suggested that the former was more energetically favorable (25). Our calculations show that the double N-exposed helix (Fig. 3.3 C) produces spectral features that look nearly identical to those of a single, flat N-exposed β-sheet (Fig. 3.3 A). The mixed N-exposed and C-exposed helix (Fig. 3.3 D) produces spectral features that look like a sum of the N-exposed (Fig. 3.3 A)

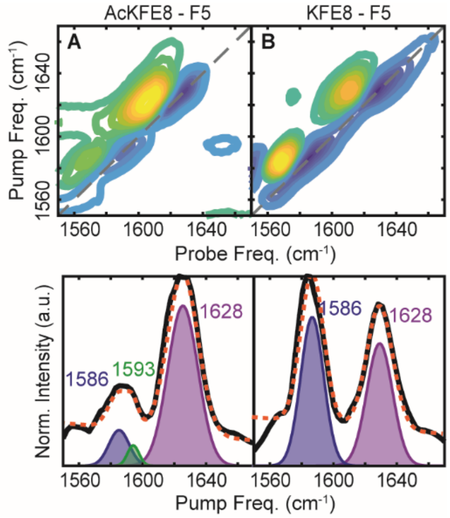

**Figure 3.4**: 2D IR spectra and pump slice amplitudes of single-isotope labeled AcKFE8 (A) and KFE8 (B). The slices were fit to a sum of Gaussian functions, with the combined fit of the traces (dotted orange line) and fits for the unlabeled (purple) and $^{13}C^{18}O$-labeled (green and blue) peaks plotted individually.

and C-exposed alignments (Fig. 3.3 B). Thus, we conclude that helical twisting of the β-sheets does not significantly alter coupling at residue F5.

Experimental 2D IR spectra of AcKFE8 and KFE8 labeled with $^{13}C^{18}O$ at residue F5 are shown in Fig. 3.4. Linear traces of the 2D IR spectra are calculated using the pump slice amplitude method (66) to reduce spectral artifacts and allow better comparison with calculated 1D IR spectra. Both variants exhibit an unlabeled amide I' mode centered at 1628 cm$^{-1}$, which is 6 cm$^{-1}$ higher than observed for the unlabeled species (Fig. 3.2, A and B). This shift can be attributed to isotope labeling of residue F5 disrupting the coupling of the β-sheets. Introduction of the heavier isotopes shift the frequency of the labeled mode by 55 cm$^{-1}$, which is sufficient to eliminate coupling between labeled and unlabeled oscillators. As residue F5 is located in the center of the β-strands, the isotope label effectively sections the β-sheet into two smaller sheets (one with four residues per strand and one with three), which yield broader, higher frequency amide I' peaks (47, 65). KFE8 exhibits an isotope-labeled feature at a probe frequency of 1586 cm$^{-1}$ (Fig. 4 B), indicative of strong coupling at residue F5, which is consistent with an N-terminal exposed strand alignment (Fig. 3.3 A).

The experimental spectrum of AcKFE8 exhibits a broad feature in the isotope-labeled region between 1580 and 1600 cm$^{-1}$ (Fig 3.4 A). This feature could not be fit to a single Gaussian peak but did fit well to a sum of two Gaussian peaks centered at 1586 cm$^{-1}$ and 1593 cm$^{-1}$. These features closely match our calculated spectra for the double helical structure containing both N-exposed and C-exposed strand alignments (Fig. 3.3 D), with the 1586 cm$^{-1}$ isotope mode corresponding to the N-exposed strand alignment of the inner β-sheet and the 1593 cm$^{-1}$ isotope mode corresponding to the C-exposed alignment of the outer β-sheet. In addition to creating frequency shifts, vibrational coupling delocalizes the vibrational mode of multiple oscillators and thus redistributes the transition dipole distribution (67). Recent studies have demonstrated that transition dipole strengths are an even more sensitive measure of coupling between peptide residues than frequency shifts (68). Furthermore, these studies have shown that 2D IR is exquisitely sensitive to changes in transition dipole strength (67, 68). Unfortunately, the weak signal strength of the isotope label in AcKFE8 makes it impossible to accurately quantify the transition dipole strength of this feature. Thus, we must restrict our analysis to qualitative comparison of the signal intensities. The significant difference in the labeled amide I’ peak intensity between AcKFE8 and KFE8 (Fig. 3.4) suggests that residue F5 experiences much stronger coupling in KFE8 than in the acetylated variant. This supports our structural assignments as residue F5 is strongly coupled across all strands in the N-exposed alignment. In contrast, at most half of the labeled F5 residues are strongly coupled in AcKFE8: the N-exposed inner helix would maintain coupling, although the C-exposed outer helix would exhibit minimal coupling. Furthermore, it has been shown that transition

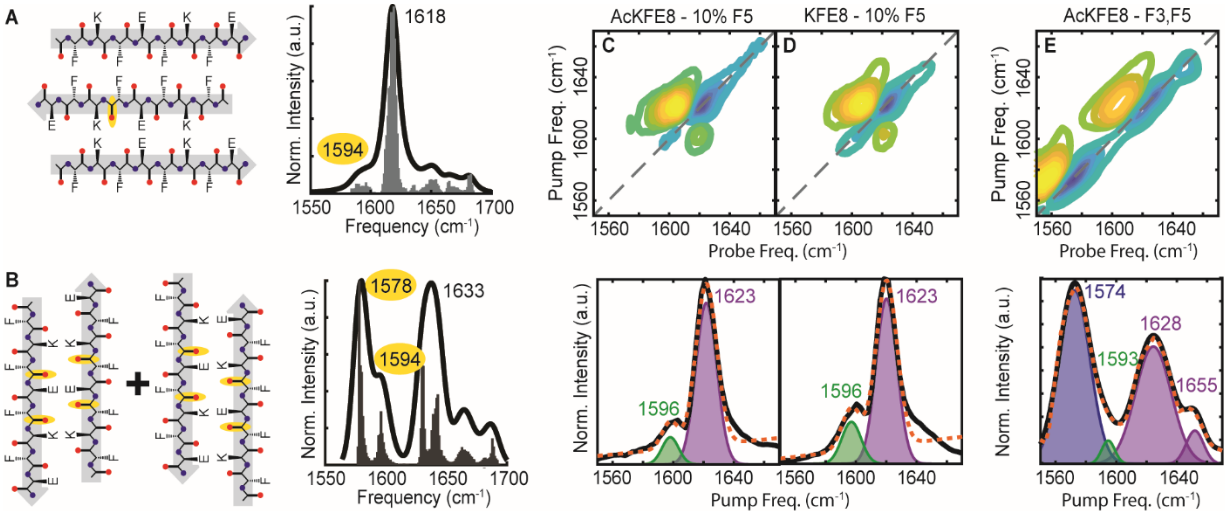

**FIGURE 3.5:** Simulated and experimental spectra of isotope-diluted and doubly isotope-labeled peptides. Simulations of an antiparallel β-sheet with 10% isotope labeling (A) exhibit a fully uncoupled $^{13}C^{18}O$ peak at 1594 cm$^{-1}$. comparable to the experimental spectra collected for isotope diluted samples of both AcKFE8 (C) and KFE8 (D). Simulations of the mixed N-exposed and C-exposed helix doubly labeled with $^{13}C^{18}O$ at residues F3 and F5 (B) predict an intense coupled peak at 1578 cm$^{-1}$ and a weak uncoupled peak at 1594 cm$^{-1}$, which closely matches experimental spectra of double-labeled AcKFE8 (E). The locations of the $^{13}C^{18}O$-labeled residues are highlighted in yellow in each β-strand. The frequencies of the unlabeled and $^{13}C^{18}O$-labeled amide I' modes are labeled in the simulated spectra, with the isotope-labeled mode highlighted in yellow. Experimental pump slice amplitudes were fit to a sum of Gaussian functions, with the combined fit of the traces (dotted orange line) and fits for the unlabeled (purple) and $^{13}C^{18}O$-labeled (green and blue) peaks plotted individually.

dipole strength is more sensitive than the overall mode frequency to subtle changes in coupling that arise from structural variations, such as interstrand spacings and dihedral angles (38). Thus, although both simulations and experiments exhibit the same coupled frequency for isotope-labeled F5, the dramatically reduced intensity of the coupled peak in AcKFE8 may reflect increased disruption of the coupling induced by the helical twist.

To confirm that the peaks observed in Fig. 3.4 truly arise from vibrational coupling between F5 residues, the AcKFE8 and KFE8 were prepared under isotope dilution conditions. If the isotope-labeled peaks shift back to the local mode frequency of 1595 $cm^{-1}$ (Fig. 3.5 A), we can attribute their redshifted frequencies in pure samples to vibration coupling; if one or both peaks remain unchanged in frequency upon isotope dilution, we must consider that other effects, such as solvent environment or hydrogen bonding, are responsible for their low frequencies (33, 39). For samples containing only 10% isotope-labeled peptides, the resulting 2D IR spectra are identical for both variants with a single, weak peak at 1596 $cm^{-1}$ (Fig. 3.5, C and D). Thus, we confirm that the two isotope-labeled modes in AcKFE8 arise from two different coupling constants between F5 residues: 10 $cm^{-1}$ and 3 $cm^{-1}$. These values align well with predicted coupling strengths within antiparallel β-sheet and with our calculations.

In addition, we synthesized another variant of AcKFE8 that is doubly labeled at residues F3 and F5 to determine whether coupling could be restored. Simulations of the mixed N-exposed and C-exposed double helix (Fig. 3.5 B) show with labeling at residues F3 and F5 show an extremely intense isotope peak at 1578 $cm^{-1}$, indicating strong coupling between the majority of the labeled residues. The correlates to coupling between residues F3 and F5 in the C-exposed strand alignment and between all F5 residues in the N-exposed strand alignment. There is also a much weaker peak at 1594 $cm^{-1}$, indicating the presence of uncoupled isotope labels. This is not surprising, as residue F3 does not align within the N-exposed strand alignment. In stark contrast, the intensities are reversed in the simulated spectrum of a double helix containing only N-exposed strands (Fig. S4 C): the uncoupled peak is more intense than the coupled peak, as residue F3 is never fully coupled. The experimental spectrum of double-labeled AcKFE8 (Fig. 5 E) also exhibits an extremely intense isotope peak at 1574 $cm^{-1}$, which confirms that coupling is re-established. When fitting the experimental spectrum, we found that including a small contribution from uncoupled isotope modes at 1593 $cm^{-1}$ did improve the overall goodness of fit (see comparison in Fig. S5) but could not be confidently assigned due to significant overlap with the intense coupled peak. Nevertheless, we can conclude that AcKFE8 adopts both N-exposed and C-exposed strand alignments whereas KFE8 adopts only the N-exposed alignment. Although this is in general agreement with the MD simulations (24, 25), it contradicts their final conclusion that the purely N-exposed double helical structure was more likely. Instead, AcKFE8 adopts the structure determined to be less energetically favorable in the second study (25), a double helix comprising both C-exposed and N-exposed strand alignments.

## 3.5. Conclusion

N-terminal acetylation is a relatively minor variation in primary sequence that drastically changes morphology and functionality. Although the acetylated and non-acetylated variants generated nearly identical spectra in unlabeled 2D IR spectroscopy studies, site-specific isotope labeling revealed multiple β-sheet configurations within the seemingly homogenous aggregates formed by AcKFE8. In combination with TDC calculations of model antiparallel β-sheets, we show the unique twisted ribbon morphology in AcKFE8 arises from the formation of two β-sheets with different strand registers. In contrast, the unacetylated variant KFE8 only adopts a single β-strand alignment and forms flat ribbons. We hypothesize that KFE8 must maintain an N-exposed strand alignment to allow for solvent stabilization of the charged N-terminus. Acetylation eliminates this N-terminal charge, loosening the constraints on strand alignment and allowing either the N- or C-terminus to be exposed to solvent. These results increase our understanding of how changes in linear sequence directly affect peptide secondary structure. Our approach can be applied readily to the systematic study of a broad range of sequence variations, including those that occur naturally as post-translational modifications or the mutations associated with different phenotypes in amyloid disease

(69, 70). A better link between primary sequence and structure will aid both the understanding of human disease and the rational design of protein biomaterials using self-assembling peptides.

### 3.6. References

1. Lee, Trinh, Yoo, Shin, Lee, Kim, Hwang, Lim, and Ryou. 2019. Self-Assembling Peptides and Their Application in the Treatment of Diseases. *Int. J. Mol. Sci.* 20:5850:10.3390/ijms20235850.

2. Rad-Malekshahi, M., L. Lempsink, M. Amidi, W.E. Hennink, and E. Mastrobattista. 2016. Biomedical Applications of Self-Assembling Peptides. *Bioconjug. Chem.* 27:3–18:10.1021/acs.bioconjchem.5b00487.

3. Eskandari, S., T. Guerin, I. Toth, and R.J. Stephenson. 2017. Recent advances in self-assembled peptides: Implications for targeted drug delivery and vaccine engineering. *Adv. Drug Deliv. Rev.* 110–111:169–187:10.1016/j.addr.2016.06.013.

4. Rudra, J.S., T. Sun, K.C. Bird, M.D. Daniels, J.Z. Gasiorowski, A.S. Chong, and J.H. Collier. 2012. Modulating Adaptive Immune Responses to Peptide Self-Assemblies. *ACS Nano*. 6:1557–1564:10.1021/nn204530r.

5. Wu, Y., P.K. Norberg, E.A. Reap, K.L. Congdon, C.N. Fries, S.H. Kelly, J.H. Sampson, V.P. Conticello, and J.H. Collier. 2017. A Supramolecular Vaccine Platform Based on α-Helical Peptide Nanofibers. *ACS Biomater. Sci. Eng.* 3:3128–3132:10.1021/acsbiomaterials.7b00561.

6. Miller, R.E.E., A.J.J. Grodzinsky, E.J.J. Vanderploeg, C. Lee, D.J.J. Ferris, M.F.F. Barrett, J.D.D. Kisiday, and D.D.D. Frisbie. 2010. Effect of self-assembling peptide, chondrogenic factors, and bone marrow-derived stromal cells on osteochondral repair. *Osteoarthr. Cartil.* 18:1608–1619:10.1016/j.joca.2010.09.004.

7. Bury, M.I., N.J. Fuller, J.W. Meisner, M.D. Hofer, M.J. Webber, L.W. Chow, S. Prasad, H. Thaker, X. Yue, V.S. Menon, E.C. Diaz, S.I. Stupp, E.Y. Cheng, and A.K. Sharma. 2014. The promotion of functional urinary bladder regeneration using anti-inflammatory nanofibers. *Biomaterials*. 35:9311–9321:10.1016/j.biomaterials.2014.07.057.

8. Moore, A.N., and J.D. Hartgerink. 2017. Self-Assembling Multidomain Peptide Nanofibers for Delivery of Bioactive Molecules and Tissue Regeneration. *Acc. Chem. Res.* 50:714–722:10.1021/acs.accounts.6b00553.

9. Ni, M., and S. Zhuo. 2019. Applications of self-assembling ultrashort peptides in bionanotechnology. *RSC Adv.* 9:844–852:10.1039/c8ra07533f.

10. Herrera Estrada, L.P., and J.A. Champion. 2015. Protein nanoparticles for therapeutic protein delivery. *Biomater. Sci.* 3:787–799:10.1039/c5bm00052a.

11. Molino, N.M., and S.-W. Wang. 2014. Caged protein nanoparticles for drug delivery. *Curr. Opin. Biotechnol.* 28:75–82:10.1016/j.copbio.2013.12.007.

12. Hartgerink, J.D., E. Beniash, and S.L. Stupp. 2001. Self-Assembly and Mineralization of Peptide-Amphiphile Nanofibers. *Science (80-. ).* 294:1684–1688:10.1126/science.1063187.

13. Genové, E., C. Shen, S. Zhang, and C.E. Semino. 2005. The effect of functionalized self-assembling peptide scaffolds on human aortic endothelial cell function. *Biomaterials*. 26:3341–3351:10.1016/j.biomaterials.2004.08.012.

14. Bowerman, C.J., and B.L. Nilsson. 2012. Review self-assembly of amphipathic β-sheet peptides: Insights and applications. *Biopolymers*. 98:169–184:10.1002/bip.22058.

15. Boekhoven, J., R.H. Zha, F. Tantakitti, E. Zhuang, R. Zandi, C.J. Newcomb, and S.I. Stupp. 2015. Alginate–peptide amphiphile core–shell microparticles as a targeted drug

delivery system. *RSC Adv.* 5:8753–8756:10.1039/C4RA16593D.

16. Sun, Y., Z. Qian, C. Guo, and G. Wei. 2015. Amphiphilic Peptides A $_6$ K and V $_6$ K Display Distinct Oligomeric Structures and Self-Assembly Dynamics: A Combined All-Atom and Coarse-Grained Simulation Study. *Biomacromolecules*. 16:2940–2949:10.1021/acs.biomac.5b00850.

17. Bourbo, V., M. Matmor, E. Shtelman, B. Rubinov, N. Ashkenasy, and G. Ashkenasy. 2011. Self-assembly and Self-replication of Short Amphiphilic β-sheet Peptides. *Orig. Life Evol. Biosph.* 41:563–567:10.1007/s11084-011-9257-y.

18. Vauthey, S., S. Santoso, H. Gong, N. Watson, and S. Zhang. 2002. Molecular self-assembly of surfactant-like peptides to form nanotubes and nanovesicles. National Academy of Sciences.

19. Altman, M., P. Lee, A. Rich, and S. Zhang. 2018. Conformational behavior of ionic self-complementary peptides. .

20. Lee, N.R., C.J. Bowerman, and B.L. Nilsson. 2013. Effects of Varied Sequence Pattern on the Self-Assembly of Amphipathic Peptides. *Biomacromolecules*. 14:3267–3277:10.1021/bm400876s.

21. Clover, T.M., C.L. O'Neill, R. Appavu, G. Lokhande, A.K. Gaharwar, A.E. Posey, M.A. White, J.S. Rudra, C.L. O'neill, R. Appavu, G. Lokhande, A.K. Gaharwar, A.E. Posey, M.A. White, and J.S. Rudra. 2020. Self-Assembly of Block Heterochiral Peptides into Helical Tapes. *J. Am. Chem. Soc.* 142:19809–19813:10.1021/jacs.9b09755.

22. Xu, H., J.J.J. Wang, S. Han, J.J.J. Wang, D. Yu, H. Zhang, D. Xia, X. Zhao, T.A. Waigh, and J.R. Lu. 2009. Hydrophobic-Region-Induced Transitions in Self-Assembled Peptide Nanostructures. *Langmuir*. 25:4115–4123:10.1021/la802499n.

23. Cui, H., A.G. Cheetham, E.T. Pashuck, and S.I. Stupp. 2014. Amino acid sequence in constitutionally isomeric tetrapeptide amphiphiles dictates architecture of one-dimensional nanostructures. *J. Am. Chem. Soc.* 136:12461–12468:10.1021/ja507051w.

24. Marini, D.M., W. Hwang, D.A. Lauffenburger, S. Zhang, and R.D. Kamm. 2002. Left-Handed Helical Ribbon Intermediates in the Self-Assembly of a β-Sheet Peptide. *Nano Lett.* 2:295–299:10.1021/nl015697g.

25. Hwang, W., D.M. Marini, R.D. Kamm, and S. Zhang. 2003. Supramolecular structure of helical ribbons self-assembled from a β-sheet peptide. *J. Chem. Phys*. 118:389–397:10.1063/1.1524618.

26. Zhou, P., L. Deng, Y. Wang, J.R. Lu, and H. Xu. 2016. Different nanostructures caused by competition of intra- and inter-β-sheet interactions in hierarchical self-assembly of short peptides. *J. Colloid Interface Sci.* 464:219–228:10.1016/J.JCIS.2015.11.030.

27. Lee, N.R., C.J. Bowerman, and B.L. Nilsson. 2013. Sequence length determinants for self-assembly of amphipathic β-sheet peptides. *Biopolymers*. 100:738–750:10.1002/bip.22248.

28. Fan, T., X. Yu, B. Shen, and L. Sun. 2017. Peptide Self-Assembled Nanostructures for Drug Delivery Applications. Hindawi.

29. Rudra, J.S., B.N. Banasik, and G.N. Milligan. 2018. A combined carrier-adjuvant system of peptide nanofibers and toll-like receptor agonists potentiates robust CD8+ T cell responses. *Vaccine*. 36:438–441:10.1016/j.vaccine.2017.12.017.

30. Bowerman, C.J., D.M. Ryan, D.A. Nissan, and B.L. Nilsson. 2009. The effect of increasing hydrophobicity on the self-assembly of amphipathic β-sheet peptides. *Mol. Biosyst.* 5:1058:10.1039/b904439f.

31. Bibian, M., J. Mangelschots, J. Gardiner, L. Waddington, M.M. Diaz Acevedo, B.G. De Geest, B. Van Mele, A. Madder, R. Hoogenboom, and S. Ballet. 2015. Rational design of a hexapeptide hydrogelator for controlled-release drug delivery. *J. Mater. Chem. B*. 3:759–765:10.1039/C4TB01294A.

32. Strasfeld, D.B., Y.L. Ling, S.-H. Shim, and M.T. Zanni. 2008. Tracking Fiber Formation in Human Islet Amyloid Polypeptide with Automated 2D-IR Spectroscopy. *J. Am. Chem. Soc.* 130:6698–6699:10.1021/ja801483n.

33. Shim, S.H.S.-H., R. Gupta, Y.L. Ling, D.B. Strasfeld, D.P. Raleigh, and M.T. Zanni. 2009. Two-dimensional IR spectroscopy and isotope labeling defines the pathway of amyloid formation with residue-specific resolution. *Proc. Natl. Acad. Sci. U. S. A.* 106:6614–6619:10.1073/pnas.0805957106.

34. Kim, Y.S., and R.M. Hochstrasser. 2009. Applications of 2D IR Spectroscopy to Peptides, Proteins, and Hydrogen-Bond Dynamics. *J. Phys. Chem. B*. 113:8231–8251:10.1021/jp8113978.

35. Dunkelberger, E.B., L.E. Buchanan, P. Marek, P. Cao, D.P. Raleigh, and M.T. Zanni. 2012. Deamidation Accelerates Amyloid Formation and Alters Amylin Fiber Structure. *J. Am. Chem. Soc.* 134:12658–12667:10.1021/ja3039486.

36. Moran, S.D., A.M. Woys, L.E. Buchanan, E. Bixby, S.M. Decatur, and M.T. Zanni. 2012. Two-dimensional IR spectroscopy and segmental 13C labeling reveals the domain structure of human γD-crystallin amyloid fibrils. *Proc. Natl. Acad. Sci. U. S. A.* 109:3329–3334:10.1073/pnas.1117704109.

37. Buchanan, L.E., J.K. Carr, A.M. Fluitt, A.J. Hoganson, S.D. Moran, J.J. de Pablo, J.L. Skinner, and M.T. Zanni. 2014. Structural motif of polyglutamine amyloid fibrils discerned with mixed-isotope infrared spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* 111:5796–

801:10.1073/pnas.1401587111.

38. Lomont, J.P., J.S. Ostrander, J.-J. Ho, M.K. Petti, and M.T. Zanni. 2017. Not All β-Sheets Are the Same: Amyloid Infrared Spectra, Transition Dipole Strengths, and Couplings Investigated by 2D IR Spectroscopy. *J. Phys. Chem. B*. 121:8935–8945:10.1021/acs.jpcb.7b06826.

39. Buchanan, L.E., M.M. Maj, E.B. Dunkelberger, P.-N.N. Cheng, J.S. Nowick, and M.T. Zanni. 2018. Structural Polymorphs Suggest Competing Pathways for the Formation of Amyloid Fibrils That Diverge from a Common Intermediate Species. *Biochemistry*. 57:6470–6478:10.1021/acs.biochem.8b00997.

40. Lomont, J.P., K.L. Rich, M. Maj, J.J. Ho, J.S. Ostrander, and M.T. Zanni. 2018. Spectroscopic Signature for Stable β-Amyloid Fibrils versus β-Sheet-Rich Oligomers. *J. Phys. Chem. B*. 122:144–153:10.1021/acs.jpcb.7b10765.

41. Hamm, P., and M. Zanni. 2011. Concepts and Methods of 2D Infrared Spectroscopy. Cambridge: Cambridge University Press.

42. Ghosh, A., J.S. Ostrander, and M.T. Zanni. 2017. Watching Proteins Wiggle: Mapping Structures with Two-Dimensional Infrared Spectroscopy. *Chem. Rev.* 117:10726–10759:10.1021/acs.chemrev.6b00582.

43. Ganim, Z., S.C. Hoi, A.W. Smith, L.P. Deflores, K.C. Jones, and A. Tokmakoff. 2008. Amide I two-dimensional infrared spectroscopy of proteins. *Acc. Chem. Res.* 41:432–441:10.1021/ar700188n.

44. Demirdöven, N., C.M. Cheatum, H.S. Chung, M. Khalil, J. Knoester, and A. Tokmakoff. 2004. Two-dimensional infrared spectroscopy of antiparallel β-sheet secondary structure. *J. Am. Chem. Soc.* 126:7981–7990:10.1021/ja049811j.

45. Buchanan, L.E., E.B. Dunkelberger, and M.T. Zanni. 2012. Examining Amyloid Structure and Kinetics with 1D and 2D Infrared Spectroscopy and Isotope Labeling. In: Fabian H, D Naumann, editors. Protein Folding and Misfolding: Shining Light by Infrared Spectroscopy. Berlin, Heidelberg: Springer Berlin Heidelberg. pp. 217–237.

46. Woys, A.M., A.M. Almeida, L. Wang, C.-C. Chiu, M. McGovern, J.J. de Pablo, J.L. Skinner, S.H. Gellman, and M.T. Zanni. 2012. Parallel β-Sheet Vibrational Couplings Revealed by 2D IR Spectroscopy of an Isotopically Labeled Macrocycle: Quantitative Benchmark for the Interpretation of Amyloid and Protein Infrared Spectra. *J. Am. Chem. Soc.* 134:19118–19128:10.1021/ja3074962.

47. Strasfeld, D.B., Y.L. Ling, R. Gupta, D.P. Raleigh, and M.T. Zanni. 2009. Strategies for extracting structural information from 2D IR spectroscopy of amyloid: application to islet amyloid polypeptide. *J. Phys. Chem. B*. 113:15679–91:10.1021/jp9072203.

48. Middleton, C.T., A.M. Woys, S.S. Mukherjee, and M.T. Zanni. 2010. Residue-specific structural kinetics of proteins through the union of isotope labeling, mid-IR pulse shaping, and coherent 2D IR spectroscopy. *Methods*. 52:12–22:10.1016/j.ymeth.2010.05.002.

49. Roeters, S.J., A. Iyer, G. Pletikapiä, V. Kogan, V. Subramaniam, and S. Woutersen. 2017. Evidence for Intramolecular Antiparallel Beta-Sheet Structure in Alpha-Synuclein Fibrils from a Combination of Two-Dimensional Infrared Spectroscopy and Atomic Force Microscopy. *Sci. Rep.* 7:1–11:10.1038/srep41051.

50. Ho, J.-J., A. Ghosh, T.O. Zhang, and M.T. Zanni. 2018. Heterogeneous Amyloid β-Sheet Polymorphs Identified on Hydrogen Bond Promoting Surfaces Using 2D SFG Spectroscopy. *J. Phys. Chem. A*. 122:1270–1282:10.1021/acs.jpca.7b11934.

51. Hahn, S., S.-S. Kim, C. Lee, and M. Cho. 2005. Characteristic two-dimensional IR spectroscopic features of antiparallel and parallel β-sheet polypeptides: Simulation

studies. *J. Chem. Phys.* 123:084905:10.1063/1.1997151.

52. Myshakina, N.S., Z. Ahmed, and S.A. Asher. 2008. Dependence of Amide Vibrations on Hydrogen Bonding. *J. Phys. Chem. B*. 112:11873–11877:10.1021/jp8057355.

53. Marini, M., T. Limongi, A. Falqui, A. Genovese, M. Allione, M. Moretti, S. Lopatin, L. Tirinato, G. Das, B. Torre, A. Giugni, F. Cesca, F. Benfenati, and E. Di Fabrizio. 2017. Imaging and structural studies of DNA–protein complexes and membrane ion channels. *Nanoscale*. 9:2768–2777:10.1039/C6NR07958J.

54. Beales, P.A., N. Geerts, K.K. Inampudi, H. Shigematsu, C.J. Wilson, and T.K. Vanderlick. 2013. Reversible assembly of stacked membrane nanodiscs with reduced dimensionality and variable periodicity. *J. Am. Chem. Soc.* 135:3335–3338:10.1021/ja311561d.

55. Chen, F., R. Strawn, and Y. Xu. 2019. The predominant roles of the sequence periodicity in the self-assembly of collagen-mimetic mini-fibrils. *Protein Sci.* 28:1640–1651:10.1002/pro.3679.

56. Georgalis, Y., E.B. Starikov, B. Hollenbach, R. Lurz, E. Scherzinger, W. Saenger, H. Lehrach, and E.E. Wanker. 1998. Huntingtin aggregation monitored by dynamic light scattering. *Proc. Natl. Acad. Sci.* 95:6118–6121:10.1073/pnas.95.11.6118.

57. Aggeli, A., G. Fytas, D. Vlassopoulos, T.C.B.B. McLeish, P.J. Mawer, and N. Boden. 2001. Structure and dynamics of self-assembling β-sheet peptide tapes by dynamic light scattering. *Biomacromolecules*. 2:378–388:10.1021/bm000080z.

58. Murphy, R.M., and M.M. Pallitto. 2000. Probing the kinetics of β-amyloid self-association. *J. Struct. Biol.* 130:109–122:10.1006/jsbi.2000.4253.

59. Streets, A.M., Y. Sourigues, R.R. Kopito, R. Melki, and S.R. Quake. 2013. Simultaneous Measurement of Amyloid Fibril Formation by Dynamic Light Scattering and Fluorescence

Reveals Complex Aggregation Kinetics. *PLoS One*. 8:e54541:10.1371/journal.pone.0054541.

60. Hill, S.E., J. Robinson, G. Matthews, and M. Muschol. 2009. Amyloid Protofibrils of Lysozyme Nucleate and Grow Via Oligomer Fusion. *Biophys. J.* 96:3781–3790:10.1016/j.bpj.2009.01.044.

61. Zandomeneghi, G., M.R.H. Krebs, M.G. McCammon, and M. Fändrich. 2009. FTIR reveals structural differences between native β-sheet proteins and amyloid fibrils. *Protein Sci.* 13:3314–3321:10.1110/ps.041024904.

62. Fenn, E.E., and M.D. Fayer. 2011. Extracting 2D IR frequency-frequency correlation functions from two component systems. *J. Chem. Phys.* 135:074502:10.1063/1.3625278.

63. Buchanan, L.E., E.B. Dunkelberger, H.Q. Tran, P.-N.P.N. Cheng, C.C.C.-C. Chiu, P. Cao, D.P. Raleigh, J.J. De Pablo, J.S. Nowick, and M.T. Zanni. 2013. Mechanism of IAPP amyloid fibril formation involves an intermediate with a transient -sheet. *Proc. Natl. Acad. Sci.* 110:19285–19290:10.1073/pnas.1314481110.

64. Paul, C., J. Wang, W.C. Wimley, R.M. Hochstrasser, and P.H. Axelsen. 2004. Vibrational Coupling, Isotopic Editing, and β-Sheet Structure in a Membrane-Bound Polypeptide. *J. Am. Chem. Soc.* 126:5843–5850:10.1021/ja038869f.

65. Kim, Y.S., L.L.L. Liu, P.H. Axelsen, and R.M. Hochstrasser. 2008. Two-dimensional infrared spectra of isotopically diluted amyloid fibrils from A 40. *Proc. Natl. Acad. Sci.* 105:7720–7725:10.1073/pnas.0802993105.

66. Valentine, M.L., Z.A. Al-Mualem, and C.R. Baiz. 2021. Pump Slice Amplitudes: A Simple and Robust Method for Connecting Two-Dimensional Infrared and Fourier Transform Infrared Spectra. *J. Phys. Chem. A*. 125:6498–6504:10.1021/acs.jpca.1c04558.

67. Grechko, M., and M.T. Zanni. 2012. Quantification of transition dipole strengths using 1D and 2D spectroscopy for the identification of molecular structures via exciton delocalization: Application to α-helices. *J. Chem. Phys.* 137:184202:10.1063/1.4764861.

68. Dunkelberger, E.B., M. Grechko, and M.T. Zanni. 2015. Transition Dipoles from 1D and 2D Infrared Spectroscopy Help Reveal the Secondary Structures of Proteins: Application to Amyloids. *J. Phys. Chem. B*. 119:14065–14075:10.1021/acs.jpcb.5b07706.

69. Iyer, A., S.J. Roeters, N. Schilderink, B. Hommersom, R.M.A.A. Heeren, S. Woutersen, M.M.A.E.A.E. Claessens, and V. Subramaniam. 2016. The Impact of N-terminal Acetylation of α-Synuclein on Phospholipid Membrane Binding and Fibril Structure. *J. Biol. Chem.* 291:21110–21122:10.1074/jbc.M116.726612.

70. Adhikari, R., M. Yang, N. Saikia, C. Dutta, W.F.A. Alharbi, Z. Shan, R. Pandey, and A. Tiwari. 2020. Acetylation of Aβ42 at Lysine 16 Disrupts Amyloid Formation. *ACS Chem. Neurosci.* 11:1178–1191:10.1021/acschemneuro.0c00069.

# Chapter 4

# Label-free detection of β-sheet polymorphism using 2D IR spectroscopy

## 4.1. Abstract

The ability to detect and characterize multiple secondary structures or polymorphs within peptide and protein aggregates is crucial to treatment and prevention of amyloidogenic diseases, production of novel, biomaterials, and many other applications. Here I report a label-free method to distinguish multiple β-sheet configurations within a single peptide aggregate using two-dimensional infrared spectroscopy. By calculating the transition dipole strength (TDS) spectrum from the ratio of linear and two-dimensional signals, we can extract maximum TDS values which provide higher sensitivity to vibrational coupling, and thus specifics of protein structure, than vibrational frequency alone. TDS spectra of AcKFE8 aggregates reveal two distinct β-sheet structures within fibers that appear homogeneous by other techniques. Furthermore, TDS spectra taken during early stages of aggregation show additional peaks that may indicate the presence of more weakly coupled β-sheet structures. These results demonstrate a unique and powerful spectroscopic method capable of distinguishing multiple oligomeric and polymorphic motifs throughout the aggregation using only native vibrational modes.

## 4.2. Introduction

Understanding how peptides and proteins assemble into aggregated structures is critical to research areas ranging from medicine to biomaterials. For example, the uncontrolled misfolding of proteins into amyloid fibrils, extended β-sheet structures that stack and twist along the fibrillar axis, is associated with wide range of human diseases, including Alzheimer's disease, Parkinson's disease, and type II diabetes(1–6). These fibrils exhibit a high degree of heterogeneity, even when formed from identical protein sequences, and the polymorphs can have widely varying levels of cytotoxicity (6–8). The development of effective therapeutics relies on finding specific inhibitors of each polymorph or species common to all polymorphic species. To further complicate drug design, increasing evidence indicates that prefibrillar oligomers, not the final amyloid fibrils, constitute the toxic species in many amyloid diseases(9, 10).Both on-pathway and off-pathway oligomers can form, creating a highly heterogeneous mixture of structures at early stages of aggregation(11–15). Thus, therapeutic routes rely on identifying early-stage polymorphic oligomers and understanding their mechanisms of aggregation. Furthermore, the ability to experimentally probe the formation of peptide and protein aggregates is required both to facilitate the rational design of new protein biomaterials and to validate structural predictions from computational methods(16–18). Standard techniques such as X-ray crystallography and solid-state nuclear magnetic resonance (NMR)(19, 20) require single aggregate morphologies to be isolated. These static techniques yield high structural detail but often miss shortlived oligomeric species and provide limited information about aggregation pathways(21–23). Common optical techniques for elucidating oligomer structures and aggregation kinetics involve labeling the monomers with site-specific tags such as spin labels or fluorescent dyes(3, 24). However, these bulky probes can potentially perturb the aggregation pathways, skewing the resulting structural and kinetic information. Two-dimensional infrared (2D IR) spectroscopy allows for characterization of protein structures based on the frequency of the backbone amide I′ modes- different secondary structures produce different vibrational couplings between the backbone amides, resulting in characteristic frequency shifts(25). Site-specific isotope labeling of the backbone carbonyls with $^{13}C$ and/or $^{18}O$ enables residue-level structural resolution, allowing oligomers and polymorphs to be characterized without perturbing the aggregation pathway(7, 26–29). While 2D IR spectroscopy of isotope labeled samples is exquisitely sensitive to molecular structure, some knowledge of potential aggregate structures, from either molecular modeling or other experimental techniques, is required to design isotope labeling schemes efficiently as distinct spectral signatures only arise when the labels come into close proximity(30). Sophisticated molecular dynamics techniques (replica exchange, umbrella sampling, metadynamics, etc.) with explicit solvent are generally required to sample the full configuration space of protein aggregation pathways, representing a significant computational expense(31–33). This emphasizes the need for experimental methods capable of detecting polymorphism in peptides, particularly during early aggregation. Here, we demonstrate a label-free method for detecting polymorphic secondary structures throughout the aggregation process by calculating transition dipole strength (TDS) spectra.

## 4.3. Methods and materials

### 4.3.1. Sample Preparation

Purified peptide was dissolved in deuterated hexafluoroisopropanol (dHFIP, 98% enriched) at 1 mM and allowed to sit overnight to ensure the peptide was completely disaggregated. The d-HFIP was removed via lyophilization and dry protein stored at -80 ºC. 2D IR samples were prepared by dissolving the lyophilized protein in unbuffered $D_2O$ (Cambridge Isotope Laboratories, Tewksbury, MA, USA). at a final concentration of 1 mM and pD of ~3, as previously reported.1 5 µL of peptide solution was placed between two $CaF_2$ windows (Crystran, Poole, Dorset, UK) separated by a 50 µ72m Teflon spacer.

### 4.3.2. 2D IR spectroscopy experimental setup

The 2D IR system has been described in detail previously(29). Briefly, the system consists of TOPAS-prime optical parametric amplifier pumped by a Solstice Ace (SpectraPhysics, Milpitas, CA, USA). Difference frequency generation is used to produce mid-IR light (6100 µm, 25 µJ, 1 kHz, 70 fs). 2D IR spectra were collected using a parallel (ZZZZ) beam polarization in a 2DQuick IR spectrometer (PhaseTech Spectroscopy, Madison, WI, USA). QuickControl software provided by Phasetech was used to collect data by scanning the delay between pump pulses from 0-2.54 ps in 23.8 fs steps, while the waiting time between pump and probe was 0 fs. Signal was collected with a monochromator (Princeton Instruments, Trinton, NJ) and mercury cadmium telluride focal-plane array detector (PhaseTech Spectroscopy, Madison, WI, USA), which can achieve a spectral resolution of 2.1 cm$^{-1}$. All data processing used custom MATLAB scripts.

### 4.3.3. Transition dipole strength spectra calculation

The mathematical formalism for calculating actual TDS spectra using 2D IR spectroscopy has been described in detail previously by Zanni and collaborators.2–4 The pump-probe geometry used for this study allows both two-dimensional and linear infrared spectra to be obtained simultaneously. The sample ΔOD is extracted directly from the 2D IR spectrum by taking the diagonal slice. The OD is calculated by collecting the probe transmission spectrum from laser shots when the pump is off for both the sample and the blank buffer, according to equation 4.1.

$$OD = -log\left(\frac{I_{sample}}{I_{buffer}}\right) \qquad 4.1$$

Generally, the OD spectrum must be baseline corrected using a second-order polynomial fit. OD and ΔOD spectra were also obtained for a L-serine (Alfa Aesar, Ward Hill, MA, USA), which served as the calibrant molecule for this study. Only the OD and ΔOD values at the peak maximum ($\omega_{max}$) of the calibrant are required to scale the TDS spectrum for the sample, which is calculated according to equation 4.2.

$$d(\omega) = \frac{\dfrac{\Delta OD_{sample(\omega,\omega)}}{OD_{sample(\omega)}}}{\dfrac{\Delta OD_{calibrant(\omega_{max},\omega_{max})}}{OD_{calibrant(\omega_{max})}}} \qquad 4.2$$

Note that the spectrum is also scaled by $I_{pump}$, the pump spectrum, to ensure that the measurement is frequency independent. In general, the TDS spectrum should be non-negative across the full spectral window and the baseline should approach 0 on either side of the sample peak. However, noise in the baseline of either the OD or ΔOD spectrum can create negative dips in the TDS baseline. As long as these remain small (< 5% of the peak maximum) and the overall baseline remains near 0, such dips should not affect analysis of the TDS magnitudes.

### 4.4. Results and discussion

Most infrared studies of proteins rely on analyzing the frequency shifts that arise from vibrational coupling between amide groups. However, such coupling also redistributes oscillator strengths, a process known as vibrational delocalization, and leads to changes in the TDS of the amide I′ mode. The TDS of a vibrational mode is directly related to its extinction coefficient and has been shown to be sensitive to variations in coupling that do not cause measurable frequency shifts(34–36). For example, disordered peptides generally absorb around 1645 $\text{cm}^{-1}$, which corresponds to the native amide I′ vibrational frequency. Depending on their local environment, α-helices absorb between 1635 and 1655 $\text{cm}^{-1}$ and thus, due to significant spectral overlap, can be difficult to distinguish from disordered structures by frequency alone(34, 35). Yet, the highly ordered hydrogen-bonding network in α-helices brings the individual backbone amide groups into close enough proximity to couple. This coupling causes the amide I′ mode to delocalize over multiple amide units. As TDS scales linearly with vibrational delocalization, this leads to an overall increase in the TDS of the amide I′ mode compared to that of a disordered protein. The extent of vibrational delocalization is affected by a number of factors, including the peptide/ aggregate size, the overall secondary structure, and the amount of structural disorder(37–39). Thus, if we can accurately measure TDS, we can gain unique insight into protein structure that is not possible from simple frequency analysis.

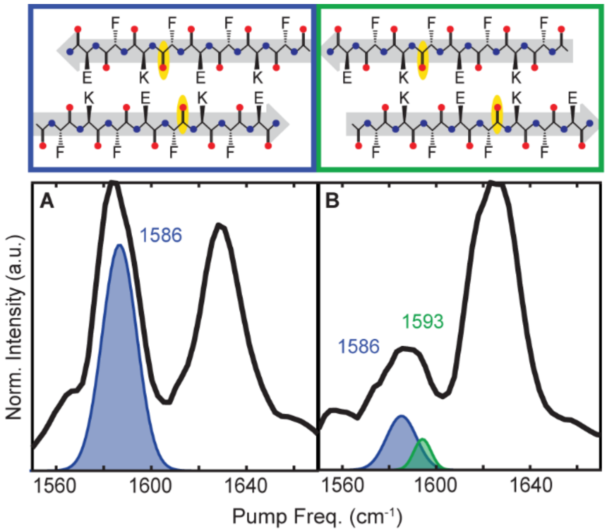

**Figure 4.1**. (top) N-terminal (blue) and C-terminal (green) exposed β-strand alignments proposed for KFE8 and AcKFE8 aggregates, with the isotope-labeled Phe-5 residue highlighted in yellow. (bottom) 2D IR traces of KFE8 (A) and AcKFE8 (B) labeled at Phe-5, with Gaussian fits highlighting peaks arising from Phe-5 within the N-terminal (blue) and C-terminal (green) exposed β-sheets. Reproduced with permission from ref 29. Copyright 2022 Elsevier.

While linear absorption techniques such as FTIR can be used to measure TDS under certain conditions, as in the case of uncoupled vibrational modes or coupled modes that are spectrally well-resolved, FTIR is largely insensitive to TDS in cases where the vibrational couplings are smaller than the line widths(34). This insensitivity arises because the integrated area of an absorption spectrum is conserved even when coupling causes vibrational delocalization(30). In contrast, nonlinear spectroscopies such as 2D IR are highly sensitive to TDS because the integrated area increases with increased vibrational delocalization, whether the coupling also results in a frequency shift or not(34, 35). Yet, extracting TDS values from 2D IR spectroscopy alone is challenging because, in accordance with Beer's law, both the beam overlap volume and sample concentration must be known. The latter can be particularly difficult to define in samples that are intrinsically heterogeneous, such as protein aggregates. Instead, 1D and 2D IR spectra can be collected simultaneously in the pump−probe geometry and their ratio used to determine TDS independent of sample concentration and other experimental variables. This approach has been described previously (34, 35) and an example calculation is given in Appendix 2.

Because of its increased sensitivity to specifics of vibrational coupling, and thus peptide structure, TDS spectra can detect the presence of distinct aggregate structures that may be hidden within 1D and 2D IR spectra due to overlap in frequency. To demonstrate this ability, we examine the KFE8 family of amphiphilic octapeptides as a model of polymorphism. Previously, we showed that AcKFE8 ($COCH_3$- FKFEFKFE-$NH_2$) aggregates contain two distinct β-sheet arrangements while KFE8 ($NH_2$-FKFEFKFE-$NH_2$) aggregates comprise only one β-sheet arrangement despite differing by only a single acetyl group at the N-terminus (29). Spectra of unlabeled AcKFE8 and KFE8 were identical; the differing strand alignments became apparent only when site-specific $^{13}C^{18}O$ labels were incorporated, yielding two isotope labeled modes consistent with two distinct antiparallel β-sheets differing only in alignment of their strands (Figure 4.1). While KFE8 and AcKFE8 both form β-sheets in which the N-terminus is solvent-exposed, the acetylated variant also forms β-sheets in which the β-strand register shifts by two residues to expose the C-terminus to solvent. Subtle variations in β-sheet alignment can alter the nearest-neighbor and interstrand couplings between residues, thus affecting the delocalization of amide I′ modes. Given that TDS is more sensitive than the vibrational frequency to such differences, TDS spectra of KFE8 and AcKFE8 aggregates should be able to detect the two β-sheet alignments even in unlabeled samples.

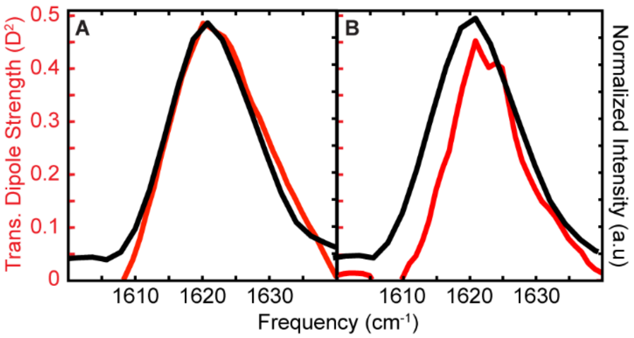

**Figure 4.2:** TDS spectra (red) and 2D IR linear traces (black) of KFE8 (A) and AcKFE8 (B) after 24 h of aggregation.

In Figure 4.2, we overlay 2D IR linear traces (black) calculated via the pump slice amplitude method with the corresponding calculated TDS spectra (red) for KFE8 and AcKFE8 after 24 hours of aggregation. While the linear slices are nearly identical for both variants, the TDS spectra differ considerably. The TDS spectrum of KFE8 closely matches the linear spectrum with a single peak at 1621 cm$^{-1}$ (Figure 4.2A). The TDS value for a single, uncoupled amide I′ mode is 0.12 D$^2$,(34) which we confirmed by measuring TDS spectra of fully disaggregated KFE8 and AcKFE8 (Figure S1.6). Thus, the maximum TDS value of 0.48 D$^2$ observed for KFE8 corresponds to delocalization across at least four β-strands. While this number may seem low, as peptide fibrils comprise thousands of β-strands and extend to micrometers in length, both structural disorder and solvent dynamics can decrease delocalization. For example, TDS measurements of amyloid fibrils have shown delocalization to vary from 4 to 12 strands even for the same peptide sequence.(36)

In contrast, TDS spectra reveal two distinct peaks underlying the vibrational transition for AcKFE8 (Figure 4.2B). The lower frequency peak at 1621 cm$^{-1}$ matches that observed for KFE8, while a unique peak appears at 1625 cm$^{-1}$. On the basis of both the higher frequency and the lower magnitude of this peak, we can conclude that it represents a β-sheet configuration with weaker coupling than the 1621 cm$^{-1}$ peak. By comparison with the data from the isotope-labeled experiments shown in Figure 4.1, we can assign the TDS peak at

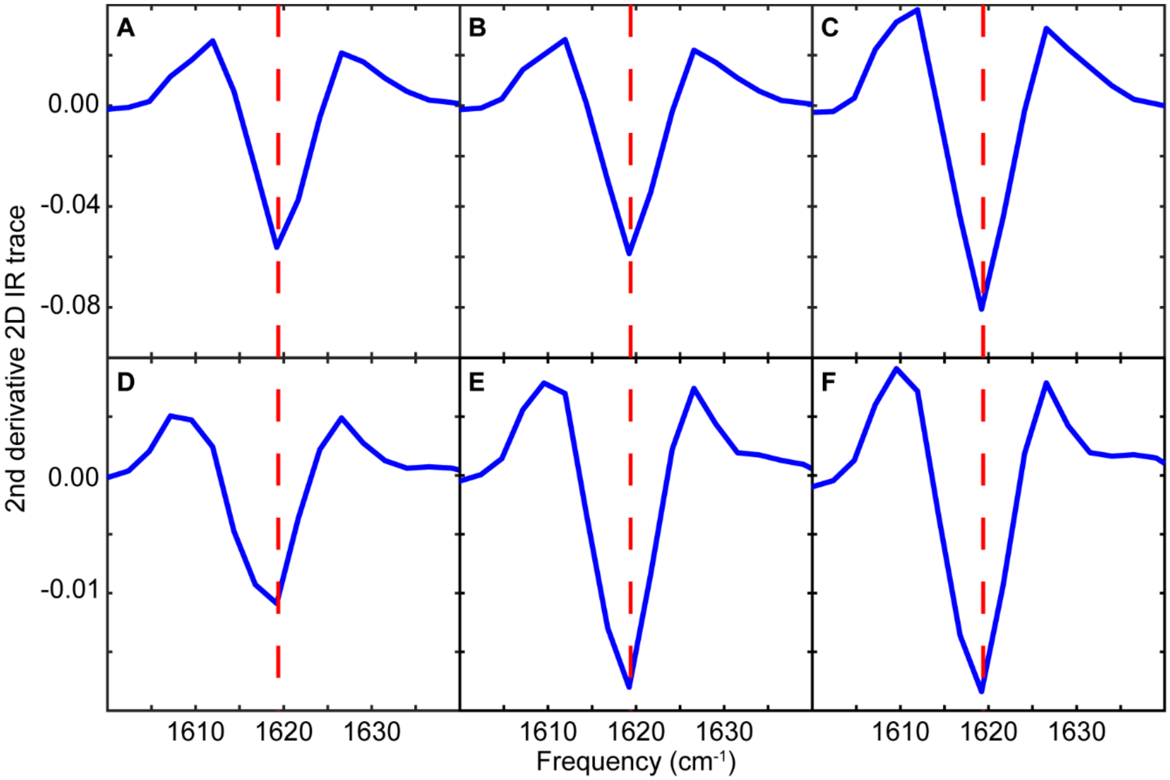

**Figure 4.3:** Second derivative spectra calculated from 2D IR linear traces. KFE8 after 5 minutes (A), 6 hours (B), and 24 hours (C) of aggregation and AcKFE8 after 5 minutes (D), 6 hours (E), and 24 hours (F). Only a single peak at 1620 cm$^{-1}$ is present in both KFE8 and AcKFE8 at all timepoints. Dashed red lines are used to highlight the consistency of the peak frequency between samples.

1625 cm$^{-1}$ present only in AcKFE8 to β-sheets in which the strands are aligned to expose the C-terminus to solvent (Figure 4.1, green), while the TDS peak present at 1621 cm$^{-1}$ in both variants arises from β-sheets in which the strands are aligned to expose the N-terminus (Figure 4.1, blue).Beyond detecting the distinct β-strand alignments present in the final aggregates, TDS spectra reveal additional β-sheet polymorphism in AcKFE8 during early stages of aggregation. While the 2D IR linear traces of AcKFE8 are identical at all time points measured (Figure S1.7), TDS spectra exhibit two additional shoulders at early aggregation times that are not readily apparent after 24 h (Figures 4.3 and S1.8). No such shoulders are observed in the TDS spectra of KFE8 (Figure S1.9), which remains relatively constant across all time points. As these local maxima are closely spaced, second-derivative spectra were

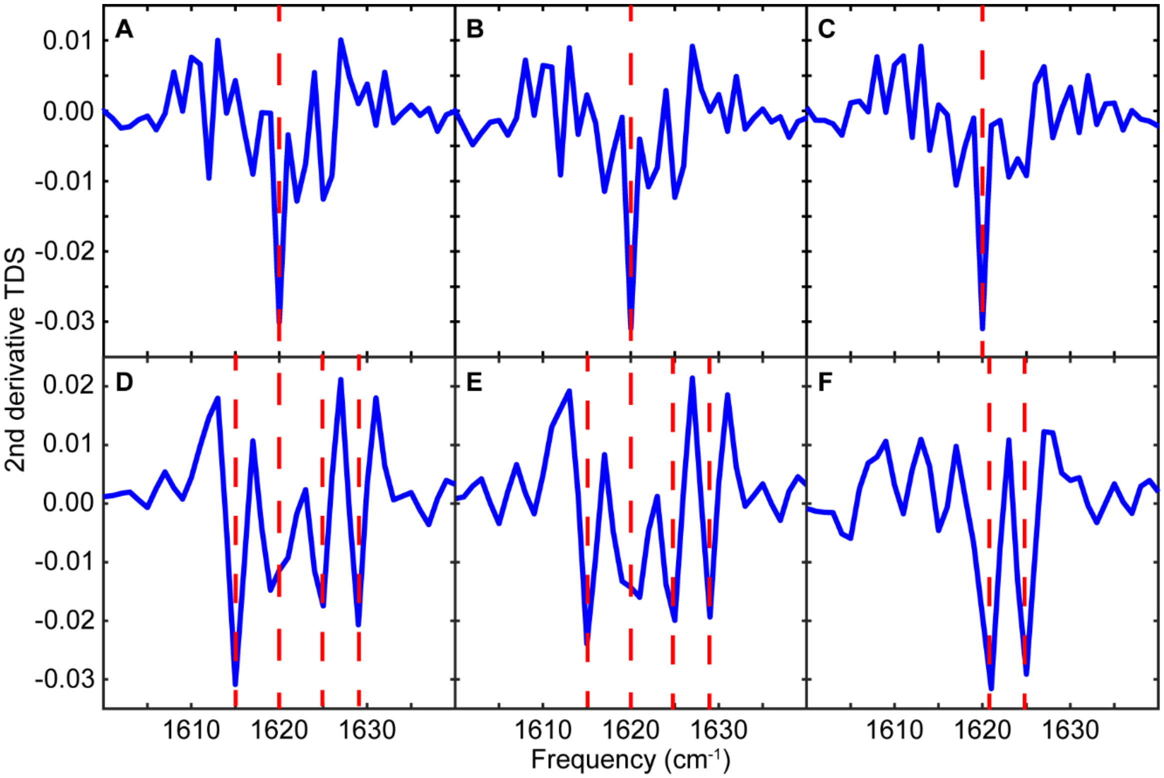

**Figure 4.4**: Second derivative spectra of the calculated TDS traces. KFE8 after 5 minutes (A), 6 hours (B), and 24 hours (C) of aggregation exhibits a consistent single peak at 1620 cm$^{-1}$. AcKFE8 after 5 minutes (D) and 6 hours (E) exhibits four peaks at 1615, 1620, 1625, and 1629 cm$^{-1}$, while only the 1620 and 1625 cm$^{-1}$ peaks persist after 24 hours (F). Dashed red lines are used to highlight peak positions in each sample.

calculated to confirm the presence of multiple peaks in the TDS spectra(40). For KFE8, a single distinct minimum was observed at 1620 cm$^{-1}$ in the second-derivative spectra of both the 2D IR (Figure 4.3A−C) and TDS traces (Figure 4.4A−C), confirming that KFE8 comprises only a single β-sheet configuration. While AcKFE8 similarly exhibits a single minimum in the 2D IR second-derivative spectra (Figure 4.3D−F), multiple minima appear within the second-derivative spectra calculated from the TDS traces (Figure 4.4D−F). These confirm the existence of two TDS peaks at 1620 and 1625 cm$^{-1}$ after 24 h as well as peaks at 1615 and 1629 cm$^{-1}$ at earlier times

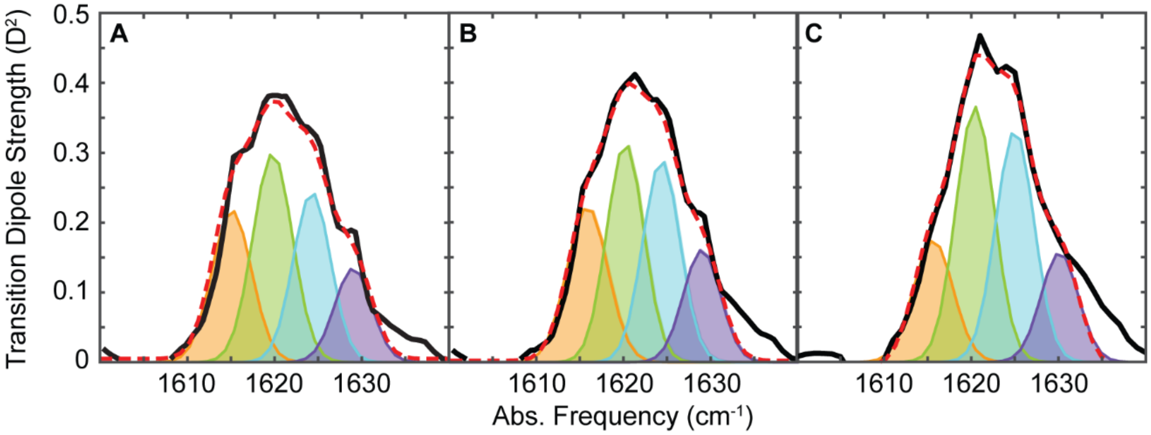

**Figure 4.5:** TDS of AcKFE8 after 5 min (A), 6 h (B), and 24 h (C) of aggregation. TDS (black) are fit to a sum of four Gaussian functions (red dotted). Individual Gaussian contributions centered at 1615 cm$^{-1}$ (purple), 1620 cm$^{-1}$ (blue), 1625 cm$^{-1}$ (green), and 1629 cm$^{-1}$ (orange) are shown.

To further aid analysis of the overlapping peaks, TDS spectra of AcKFE8 were fit to a sum of four Gaussian functions (Figure 4.5). Fit line widths were kept constant between time points, and the center frequencies were constrained to the spectral window of 1610−1635 cm$^{-1}$. Ultimately, all spectra were best fit with Gaussians centered at 1615, 1620, 1625, and 1529 cm$^{-1}$. These fits reveal that the TDS features at 1615 and 1629 cm$^{-1}$ may, in fact, be present in the final aggregate, although they were not apparent in the second-derivative spectra but are overshadowed by the stronger features at 1620 and 1625 cm$^{-1}$. The presence of four peaks suggests that AcKFE8 residues experience four distinct sets of average β-sheet couplings.

From isotope labeling (Figure 4.1), we know that two of these correspond to β-sheets with different strand registers that comprise the double-helical morphology of AcKFE8(29). As the other two cannot be differentiated with isotope labeling scheme used previously, they may represent similar strand alignments with subtle variations in structure, such as different degrees of rotation between strands. As aggregation progresses, the intensities of the TDS peaks at 1620 and 1625 cm$^{-1}$ increase to final values of 0.40 and 0.36 D$^2$, respectively, while the 1615 and 1629 cm$^{-1}$ peaks remain fairly consistent or decrease slightly to final values of 0.17 and 0.16 D$^2$, respectively. It is critical to remember that TDS peak intensities do not report population; by taking the ratio of 1D and 2D IR spectra, any concentration dependence is removed(34, 35). Instead, the TDS maxima correspond to the extent of vibrational coupling and delocalization. Thus, the β-sheet structures represented by the 1620 and 1625 cm$^{-1}$ peaks become more organized over time, and the corresponding amide I′ modes ultimately delocalize over at least three strands, while the 1615 and 1629 cm$^{-1}$ structures exhibit delocalization over less than two strands and become less organized or even disappear entirely.

## 4.5. Conclusion

As shown here, TDS spectra serve as a remarkably sensitive tool capable of detecting small variations in protein structure without the need for special probes or vibrational labels. This ability greatly facilitates the application of IR spectroscopy to large, complex aggregates which can be expensive to model computationally and difficult to study by traditional biophysical methods. Once the presence of multiple TDS peaks identifies the existence of distinct structural elements, isotopes or other vibrational labels can be used to resolve these structures at the single-residue level. Additionally, TDS spectra can be obtained with the same temporal resolution as standard 2D IR spectra, enabling the detection of oligomeric species during early stages of aggregation. Such an approach could be invaluable in distinguishing the multitude of on- and off-pathway oligomers believed to be responsible for cytotoxicity in amyloid disease(9–15). Beyond protein aggregates, native peptides and proteins contain an array of secondary structures, including β-sheets, α-helices, 3,10-helices, and disordered regions. Each secondary structure will have a distinct set of vibrational couplings, yet IR frequency alone is often insufficient to distinguish various helical and coiled structures. As TDS is more sensitive to variations in coupling than frequency alone,(34–36)

TDS spectra will allow such mixtures to be resolved. Thus, TDS spectra be used to efficiently characterize protein secondary structure, screen protein aggregates for the existence of polymorphic structures, and aid the rational design of both peptide biomaterials and aggregation inhibitors by providing insights into the aggregation mechanism.

### 4.6. References:

1. Ho, J.-J., A. Ghosh, T.O. Zhang, and M.T. Zanni. 2018. Heterogeneous Amyloid β-Sheet Polymorphs Identified on Hydrogen Bond Promoting Surfaces Using 2D SFG Spectroscopy. *J. Phys. Chem. A*. 122:1270–1282:10.1021/acs.jpca.7b11934.
2. Roeters, S.J., A. Iyer, G. Pletikapiä, V. Kogan, V. Subramaniam, and S. Woutersen. 2017. Evidence for Intramolecular Antiparallel Beta-Sheet Structure in Alpha-Synuclein Fibrils from a Combination of Two-Dimensional Infrared Spectroscopy and Atomic Force Microscopy. *Sci. Rep.* 7:1–11:10.1038/srep41051.
3. Wang, H., L. Duo, F. Hsu, C. Xue, Y.K. Lee, and Z. Guo. 2020. Polymorphic Aβ42 fibrils adopt similar secondary structure but differ in cross-strand side chain stacking interactions within the same β-sheet. *Sci. Rep.* 10:5720:10.1038/s41598-020-62181-x.
4. Rudra, J.S., S. Mishra, A.S. Chong, R.A. Mitchell, E.H. Nardin, V. Nussenzweig, and J.H. Collier. 2012. Self-assembled peptide nanofibers raising durable antibody responses against a malaria epitope. *Biomaterials*. 33:6476–6484:10.1016/j.biomaterials.2012.05.041.
5. Buchanan, L.E., J.K. Carr, A.M. Fluitt, A.J. Hoganson, S.D. Moran, J.J. de Pablo, J.L. Skinner, and M.T. Zanni. 2014. Structural motif of polyglutamine amyloid fibrils discerned with mixed-isotope infrared spectroscopy. *Proc. Natl. Acad. Sci.* 111:5796–5801:10.1073/pnas.1401587111.

6. Bousset, L., L. Pieri, G. Ruiz-Arlandis, J. Gath, P.H. Jensen, B. Habenstein, K. Madiona, V. Olieric, A. Böckmann, B.H. Meier, and R. Melki. 2013. Structural and functional characterization of two alpha-synuclein strains. *Nat. Commun.* 4:2575:10.1038/ncomms3575.

7. Buchanan, L.E., M.M. Maj, E.B. Dunkelberger, P.-N.N. Cheng, J.S. Nowick, and M.T. Zanni. 2018. Structural Polymorphs Suggest Competing Pathways for the Formation of Amyloid Fibrils That Diverge from a Common Intermediate Species. *Biochemistry*. 57:6470–6478:10.1021/acs.biochem.8b00997.

8. Adhikari, R., M. Yang, N. Saikia, C. Dutta, W.F.A. Alharbi, Z. Shan, R. Pandey, and A. Tiwari. 2020. Acetylation of Aβ42 at Lysine 16 Disrupts Amyloid Formation. *ACS Chem. Neurosci.* 11:1178–1191:10.1021/acschemneuro.0c00069.

9. Ferreira, S.T., M.N.N. Vieira, and F.G. De Felice. 2007. Soluble protein oligomers as emerging toxins in Alzheimer's and other amyloid diseases. In: IUBMB Life. John Wiley & Sons, Ltd. pp. 332–345.

10. Kreutzer, A.G., and J.S. Nowick. 2018. Elucidating the Structures of Amyloid Oligomers with Macrocyclic β-Hairpin Peptides: Insights into Alzheimer's Disease and Other Amyloid Diseases. *Acc. Chem. Res.* 51:706–718:10.1021/acs.accounts.7b00554.

11. Dear, A.J., G. Meisl, A. Šarić, T.C.T. Michaels, M. Kjaergaard, S. Linse, and T.P.J. Knowles. 2020. Identification of on- and off-pathway oligomers in amyloid fibril formation. *Chem. Sci.* 10.1039/C9SC06501F.

12. Karamanos, T.K., M.P. Jackson, A.N. Calabrese, S.C. Goodchild, E.E. Cawood, G.S. Thompson, A.P. Kalverda, E.W. Hewitt, and S.E. Radford. 2019. Structural mapping of oligomeric intermediates in an amyloid assembly pathway. *Elife*. 8:10.7554/eLife.46574.001.

13. Tornquist, M., R. Cukalevski, U. Weininger, G. Meisl, T.P.J. Knowles, T. Leiding, A. Malmendal, M. Akke, and S. Linse. 2020. Ultrastructural evidence for self-replication of alzheimer-associated Aβ42 amyloid along the sides of fibrils. *Proc. Natl. Acad. Sci. U. S. A.* 117:11265–11273:10.1073/pnas.1918481117.

14. Fusco, G., S.W. Chen, P.T.F. Williamson, R. Cascella, M. Perni, J.A. Jarvis, C. Cecchi, M. Vendruscolo, F. Chiti, N. Cremades, L. Ying, C.M. Dobson, and A. De Simone. 2017. Structural basis of membrane disruption and cellular toxicity by a-synuclein oligomers. *Science (80-. ).* 358:1440–1443:10.1126/science.aan6160.

15. Cremades, N., S.I.A. Cohen, E. Deas, A.Y. Abramov, A.Y. Chen, A. Orte, M. Sandal, R.W. Clarke, P. Dunne, F.A. Aprile, C.W. Bertoncini, N.W. Wood, T.P.J. Knowles, C.M. Dobson, and D. Klenerman. 2012. Direct observation of the interconversion of normal and toxic forms of α-synuclein. *Cell*. 149:1048–1059:10.1016/j.cell.2012.03.037.

16. Sinha, N.J., M.G. Langenstein, D.J. Pochan, C.J. Kloxin, and J.G. Saven. 2021. Peptide Design and Self-assembly into Targeted Nanostructure and Functional Materials. *Chem. Rev.* 121:13915–13935:10.1021/acs.chemrev.1c00712.

17. Zhao, Y., W. Yang, C. Chen, J. Wang, L. Zhang, and H. Xu. 2018. Rational design and self-assembly of short amphiphilic peptides and applications. *Curr. Opin. Colloid Interface Sci.* 35:112–123:10.1016/j.cocis.2018.02.009.

18. Matsuura, K. 2017. Construction of Functional Biomaterials by Biomolecular Self-Assembly. *Bull. Chem. Soc. Jpn.* 90:873–884:10.1246/bcsj.20170133.

19. Paravastu, A.K., R.D. Leapman, W.-M. Yau, and R. Tycko. 2008. Molecular structural basis for polymorphism in Alzheimer's -amyloid fibrils. *Proc. Natl. Acad. Sci.* 105:18349–18354:10.1073/pnas.0806270105.

20. Paravastu, A.K., I. Qahwash, R.D. Leapman, S.C. Meredith, and R. Tycko. 2009. Seeded growth of β-amyloid fibrils from Alzheimer's brain-derived fibrils produces a distinct fibril structure. *Proc. Natl. Acad. Sci. U. S. A.* 106:7443–7448:10.1073/pnas.0812033106.

21. Gao, Y., C. Guo, J.O. Watzlawik, P.S. Randolph, E.J. Lee, D. Huang, S.M. Stagg, H.X. Zhou, T.L. Rosenberry, and A.K. Paravastu. 2020. Out-of-Register Parallel β-Sheets and Antiparallel β-Sheets Coexist in 150-kDa Oligomers Formed by Amyloid-β(1–42). *J. Mol. Biol.* 432:4388–4407:10.1016/j.jmb.2020.05.018.

22. Tay, W.M., D. Huang, T.L. Rosenberry, and A.K. Paravastu. 2013. The Alzheimer's amyloid-β(1-42) peptide forms off-pathway oligomers and fibrils that are distinguished structurally by intermolecular organization. *J. Mol. Biol.* 425:2494–2508:10.1016/j.jmb.2013.04.003.

23. Huang, D., M.I. Zimmerman, P.K. Martin, A.J. Nix, T.L. Rosenberry, and A.K. Paravastu. 2015. Antiparallel β-sheet structure within the C-terminal region of 42-residue Alzheimer's amyloid-β peptides when they form 150-kDa oligomers. *J. Mol. Biol.* 427:2319–2328:10.1016/j.jmb.2015.04.004.

24. DiMaio, J.T.M.M., D.M. Raymond, and B.L. Nilsson. 2017. Display of functional proteins on supramolecular peptide nanofibrils using a split-protein strategy. *Org. Biomol. Chem.* 15:5279–5283:10.1039/C7OB01057E.

25. Barth, A. 2007. Infrared spectroscopy of proteins. *Biochim. Biophys. Acta - Bioenerg.* 1767:1073–1101:10.1016/j.bbabio.2007.06.004.

26. Strasfeld, D.B., Y.L. Ling, R. Gupta, D.P. Raleigh, and M.T. Zanni. 2009. Strategies for extracting structural information from 2D IR spectroscopy of amyloid: application to islet amyloid polypeptide. *J. Phys. Chem. B.* 113:15679–91:10.1021/jp9072203.

27. Buchanan, L.E., E.B. Dunkelberger, and M.T. Zanni. 2012. Examining Amyloid Structure and Kinetics with 1D and 2D Infrared Spectroscopy and Isotope Labeling. In: Fabian H, D Naumann, editors. Protein Folding and Misfolding: Shining Light by Infrared Spectroscopy. Berlin, Heidelberg: Springer Berlin Heidelberg. pp. 217–237.

28. Shim, S.H.S.-H., R. Gupta, Y.L. Ling, D.B. Strasfeld, D.P. Raleigh, and M.T. Zanni. 2009. Two-dimensional IR spectroscopy and isotope labeling defines the pathway of amyloid formation with residue-specific resolution. *Proc. Natl. Acad. Sci. U. S. A.* 106:6614–6619:10.1073/pnas.0805957106.

29. Weeks, W.B., C.J. Tainter, and L.E. Buchanan. 2022. Investigating the effects of N-terminal acetylation on KFE8 self-assembly with 2D IR spectroscopy. *Biophys. J.* 121:1549–1559:10.1016/j.bpj.2022.03.003.

30. Hamm, P., and M. Zanni. 2011. Concepts and Methods of 2D Infrared Spectroscopy. Cambridge: Cambridge University Press.

31. Buchanan, L.E., E.B. Dunkelberger, H.Q. Tran, P.-N.P.N. Cheng, C.C.C.-C. Chiu, P. Cao, D.P. Raleigh, J.J. De Pablo, J.S. Nowick, and M.T. Zanni. 2013. Mechanism of IAPP amyloid fibril formation involves an intermediate with a transient -sheet. *Proc. Natl. Acad. Sci.* 110:19285–19290:10.1073/pnas.1314481110.

32. Laio, A., and M. Parrinello. 2002. Escaping free-energy minima. *Proc. Natl. Acad. Sci.* 99:12562–12566:10.1073/pnas.202427399.

33. Piana, S., and A. Laio. 2007. A Bias-Exchange Approach to Protein Folding. *J. Phys. Chem. B*. 111:4553–4559:10.1021/jp067873l.

34. Grechko, M., and M.T. Zanni. 2012. Quantification of transition dipole strengths using 1D and 2D spectroscopy for the identification of molecular structures via exciton

delocalization: Application to α-helices. *J. Chem. Phys.* 137:184202:10.1063/1.4764861.

35. Dunkelberger, E.B., M. Grechko, and M.T. Zanni. 2015. Transition Dipoles from 1D and 2D Infrared Spectroscopy Help Reveal the Secondary Structures of Proteins: Application to Amyloids. *J. Phys. Chem. B*. 119:14065–14075:10.1021/acs.jpcb.5b07706.

36. Lomont, J.P., J.S. Ostrander, J.-J. Ho, M.K. Petti, and M.T. Zanni. 2017. Not All β-Sheets Are the Same: Amyloid Infrared Spectra, Transition Dipole Strengths, and Couplings Investigated by 2D IR Spectroscopy. *J. Phys. Chem. B*. 121:8935–8945:10.1021/acs.jpcb.7b06826.

37. Demirdöven, N., C.M. Cheatum, H.S. Chung, M. Khalil, J. Knoester, and A. Tokmakoff. 2004. Two-dimensional infrared spectroscopy of antiparallel β-sheet secondary structure. *J. Am. Chem. Soc.* 126:7981–7990:10.1021/ja049811j.

38. Paul, C., J. Wang, W.C. Wimley, R.M. Hochstrasser, and P.H. Axelsen. 2004. Vibrational Coupling, Isotopic Editing, and β-Sheet Structure in a Membrane-Bound Polypeptide. *J. Am. Chem. Soc.* 126:5843–5850:10.1021/ja038869f.

39. Hahn, S., S.-S. Kim, C. Lee, and M. Cho. 2005. Characteristic two-dimensional IR spectroscopic features of antiparallel and parallel β-sheet polypeptides: Simulation studies. *J. Chem. Phys.* 123:084905:10.1063/1.1997151.

40. Susi, H., and D. Michael Byler. 1983. Protein structure by Fourier transform infrared spectroscopy: Second derivative spectra. *Biochem. Biophys. Res. Commun.* 115:391–397:10.1016/0006-291X(83)91016-1.

## Chapter 5

## Investigating the aggregation of amyloid precursor protein secretions using 2D IR

### 5.1. Abstract

The amyloid precursor protein has been the target of many drug inhibitor studies as APP is the immediate precursor of the β-carboxyl terminal fragment (C99) transmembrane protein and therein, the precursor to the amyloid beta polypeptides. Amyloid beta (Aβ42) and other derivative amyloidogenic proteins are key factors in causing neurogenerative diseases such as Alzheimer's and Parkinson's Disease. As the amyloid beta polypeptides are the active substrate in these diseases, it is equally critical to understand the aggregation of the C99 transmembrane protein in lipid rich environments. Here, I use 2D IR spectroscopy to elucidate the variation of secondary structures induced by lipid micelles in C99. Additionally, I also analyze the effects of a prescreened small molecule inhibitor on C99 aggregation. Current research also suggests curtain off-pathway oligomer species caused by lysine acetylation may be more cytotoxic than the final aggregates. In this work, I use two-dimensional infrared (2D IR) spectroscopy to study the aggregation kinetics of the wild type Aβ42, Aβ42-K16, and Aβ 42-KKAc analyzing the variation in oligomers species contributing to the level of toxicity.

### 5.2. Introduction

The amyloid precursor protein (APP) gene residues in a small gene family including the human genes APLP1 and APLP2, Appl in flys, and apl-1 in worms(1). While each of these genes encodes a membrane protein composed of a large extracellular domain and a short cytoplasmic region, only the APP gene yields the amyloid beta (Aβ) domain. In order to produce the Aβ polypeptides, nearly the entire ectodomain is removed via a β or α-secretion. This results in the α- or β-carboxyl terminal fragment (CTF or C99) tether to the membrane(2). The β-CTF protein then undergoes a γ-secretase cleavage to yield the neurodegenerative Aβ polypeptide (40, 42, and 43) while the α-CTF produces the p3 protein through the same cleavage process(3). Although APP has been assigned several physiological roles, the actual function of APP remains convoluted and unclear. Here, I focus on the two main secretion productions, C99 and amyloid beta 42 (Aβ42).

### 5.2.1. β-carboxyl terminal fragment

The transmembrane β-carboxyl terminal fragment, also known as the C99 protein, has long been associated with the onset of Alzheimer's disease(4). The C99 fragment is composed of the 99 C-terminal residues (672-770) from APP. As both C99 and APP are integral parts of the membrane lipid bilayer, it is critical to understand the folding mechanism and structures in the presence of lipids(2, 4–7). Previous 2D NMR studies by Dr. James Hutchison in the Sanders research group at Vanderbilt medical school revealed monomeric C99 is composed of a disordered N-terminus with a membrane associated α-helix and a flexible loop structure on the N-terminal extracellular space(2). Following a helical transmembrane domain, C99 forms a disordered intracellular loop and a membrane associated α-helix. This monomeric structure was formed in the presence of anionic lyso-myristoyl-phosphatidylglycerol (LMPG) lipid micelles. While the monomeric structure was elucidated, the aggregated structures were not revealed nor were any structural effects elucidated from varying concentrations of lipid micelles. Here, I use 2D IR spectroscopy to analyze the effects of the lipid concentration in a collaborative effect.

As C99 is the direct precursor for Aβ42, the Sanders group have aimed to inhibit the γ-secretion resulting in the amyloid beta polypeptide production. To this end, a host of small molecule binding candidates were screened through molecular dynamics simulations. N-(pyridine-3-ylmethyl)aniline (VUx96) was found to induce significant structural change in 2D NOESY analysis conducted by Manuel Castro of the Sanders lab. To further probe the changes induced in secondary structure, I employ 2D IR spectroscopy to analyze C99, as well at the truncated derivatives C74 and C55, bond to varying concentrations of VUx96.

### 5.2.2. Amyloid beta polypeptide 42

Recent studies regarding Aβ polypeptide's role in Alzheimer's disease (AD) has taken two major directives. The first is analyzing the aggregate structures formed under the separate aggregation pathways. The other involves probing the functionality of the different aggregates within the brain. To design inhibitors and drug receptors, the first directive holds the most interest. However, aggregation pathways are inherently complex and only increase in complexity due to the amyloidogenic nature of Aβ. The aggregation pathway is capable of producing multiple oligomeric species yielding protofibrils and amyloid fibrils(8–11). This study focuses on the most dominant species of Aβ polypeptide peptides responsible for the plaques in the brain (Aβ42). It has been previously shown that the cytotoxicity of Aβ42 is directly linked to the aggregate structure(12). In addition to the multiple species produced from the γ-secretase (Aβ39-Aβ43), the Aβ family of proteins are susceptible to common post-translation modification such as phosphorylation and acetylation(12–14). Acetylation has gained interest as it has been shown to have vital function in DNA repair, cellular signaling, and protein aggregation and self-assembly(15). Acetylation has already been identified as a new therapeutic target in the Tau protein, another protein proposed to cause AD(16). Aβ42 can be acetylation at two positions, the lysine 16 and lysine 28 positions. The lysine 28 residue has been shown to form a salt bridge between residues A42 or D23 which stabilizes the amyloid fibrils. Adhikari et al. showed by acetylating the K16 and both K16 and K28 (KKAc) positions, the fibril structures did not form, Instead the aggregate structures became largely amorphous however, the cytotoxic of both lysine variants increased significantly(12). While the supramolecular morphologies were clearly different there is little known about how the acetylation changes the secondary structure composition during self-assembly. In this study, I use 2D IR spectroscopy to establish aggregation kinetics of Aβ42 and acetylated mutations demonstrating the vast effects of small changes in primary sequence to amyloidogenic protein oligomer formation and aggregation kinetics

## 5.3. Methods and Materials

### 5.3.1. C99 sample preparation

Stock solutions of C99 (200 μM), C74 (200 μM), C55 (180 μM), VUx96 (200 mM), buffer (25 mM $NaPO_4$ 75 mM NaCl) and deuterated 20% lyso-myristoyl-phosphatidylglycerol (LMPG) were provided by the Sanders lab. All experiments were conducted at final concentrations of 75 μM protein, 6.25% detergent, and 5% deuterated DMSO with varying amounts of VUx96: 10 mM, 4 mM, and 2 mM. Blank samples were prepared under the same conditions while withholding the respective protein to maintain the same level of VUx96/DMSO and detergent. Using our custom Matlab script, absorption spectra of the blank background samples were subtracted from the protein containing samples to enhance the protein absorption region (1590-1690 $cm^{-1}$). All spectra were internally normalized to ensure equal scaling.

### 5.3.2. Aβ42 sample preparation

Wild type (WT) Aβ42, Aβ42-K16Ac (K16), and Aβ42-K16AcK28Ac (KKAc) were synthesized via microwave assisted solid phase synthesis and purified via size exclusion chromatography as described in chapter 2. Aliquots were fully dissolved in hexafluroisopropanol (HFIP) and let sit at RT overnight to ensure the highest degree of de-aggregation. The aliquots were lyophilized for at least 24 hours to ensure all HFIP was removed. All Aβ42 and acetylated variant experiments were conducted at a final concentration of 150 μM in 20 mM deuterated sodium phosphate buffer at 7.4 pH.

### 5.3.3. 2D IR spectroscopy

A detail description of 2D IR data collection and processing methods are described elsewhere (17). Briefly, 800 nm pulses (1 kHz, 60fs, 7mJ) were generated by a single box ultrafast amplifier (Solstice, SpectraPhysics, Milpitas, CA, USA). A 50/50 beamsplitter was used to direct half of the beam to pump an optical parametric amplifier with difference frequency generation (TOPAS, SpectraPhysics, Milpitas, CA, USA). The resulting mid-IR light (6100 nm, 1 kHz, 70 fs, 25 μJ) was directed into the 2D IR spectrometer (2DQuick IR, PhaseTech Spectroscopy, Madison, WI, USA). In the spectrometer, the mid-IR light is split into pump (90%) and probe (10%) beams. The pump beam was passed through a germanium acousto-optic modulator (Ge AOM) pulse-shaper to generate pairs of pump pulses with varying time delays between them. The pump and probe pulses were overlapped at the sample to generate the signal. The signal was directed into a monochromator (Princeton Instruments, Trinton, NJ) and dispersed onto a mercury cadmium telluride focal-plane array (FPA) detector (PhaseTech Spectroscopy, Madison, WI, USA). The data was collected using the QuickControl software provided by PhaseTech and processed using custom MATLAB scripts. The FPA detector was calibrated daily was a 4-nitrobenzaldehyde (4NBA, 500 mM in toluene) standard. 4NBA absorbs at 1535 cm$^{-1}$, 1605 cm$^{-1}$, and 1711 cm$^{-1}$, which spans the isotope and amide I’ spectral regions. Samples were measured by placing 5 μL of sample solution between two $CaF_2$ windows (Crystran, Poole, Dorset, UK) separated by a 50 μm Teflon spacer. All experiments were conducted at room temperature.

## 5.4. Results and Discussion

### 5.4.1. Lipid micelle concentrations effects on C99 aggregation

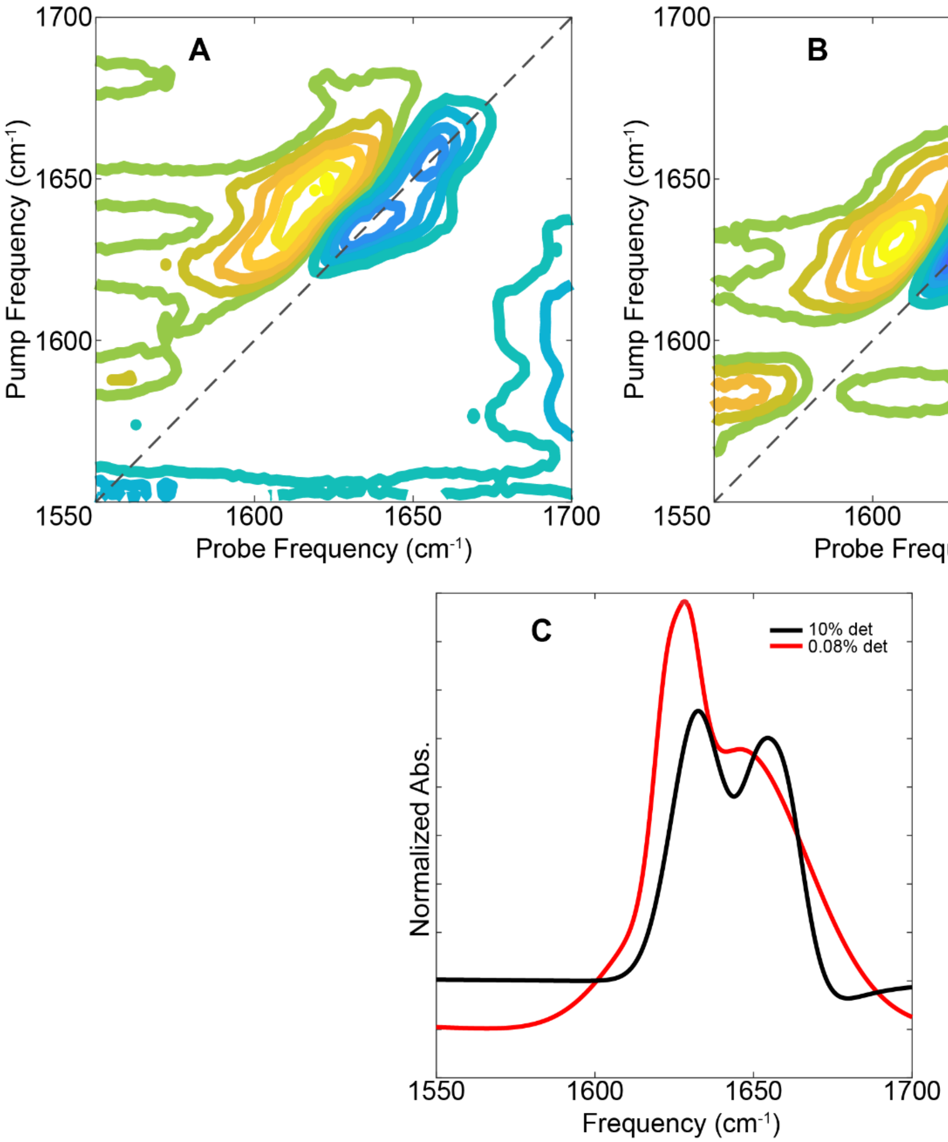

**Figure 5.1**: 2D IR contour maps of 150 μM C99 containing 10% LMPG detergent (A) and 0.08% LMPG detergent (B) show the fundamental and overtone absorption as well as the lineshapes variation due to the increase in detergent. The linear traces through the dashed line (C) yield a more direct comparison of the amide I' absorbance.

C99 samples containing 10% of the LMPG detergent were shown to absorb at two frequencies denoting the presence of two distinct amide I' modes. The amide I' absorptions at 1656 cm$^{-1}$ indicates α-helical secondary structure while the absorption at 1634 cm$^{-1}$ indicates β-sheet structure characteristics (Figure 5.1 A and C)(18, 19). While the intensities look to be equally, this does not mean the number of secondary structures is respectively equal within the same sample. However, there is a clear, strong conversion to β-sheet secondary structure as the detergent concentration decreases. At 0.08% detergent, the β-sheet absorption is slightly red shifted to 1630 cm$^{-1}$ and appears more intense as the α-helical peak transition to a shoulder of the main β-sheet absorbance (Figure 5.1 B and C). Both lineshapes of both 2D IR peaks in each sample are broad pointing toward a heterogenous contribution of secondary structures and solvent interactions (Figure 5.1A and B). Fraying of peptide aggregates toward the ends of structures is also likely to contribute the broadness of the peaks seen in the contour maps and linear trace(18, 20). To better understand how the lipid micelle concentration effects the secondary structures the LMPG concentration was varied from 0.08 to 10% wit a step at 0.325% LMPG while all other sample variables were held constant. The spectra were found to be nearly identical from 2 to 10% LMPG. However, significant differences in the ratio of the intensity of the β-sheet peak and α-helix peak were found around 1.5% LMPG. The β/α ratios were calculated by dividing the local maxima intensity of the β-sheet peak around 1630 cm$^{-1}$ by the local maxima intensity α-helix peak around 1650 cm$^{-1}$ generated the linear trace of the fundamental peak in each detergent dilution sample (Figure 5.2).

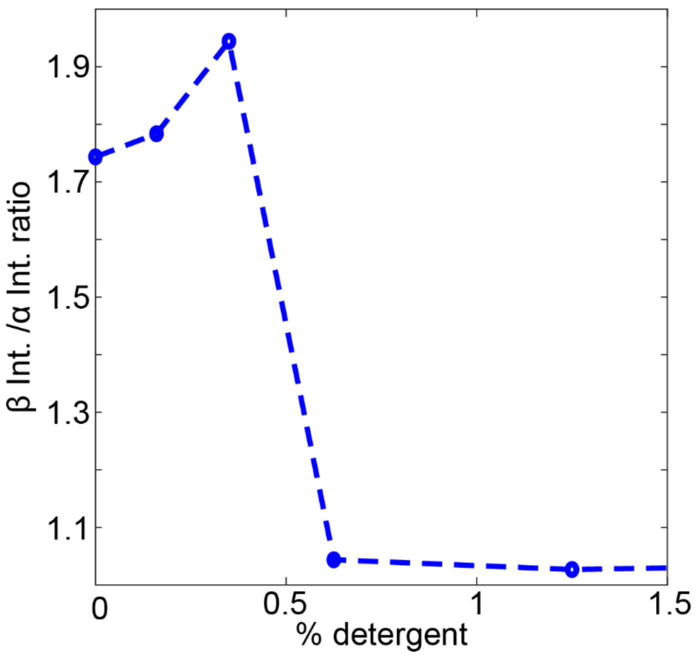

**Figure 5.2:** The ratio of β-sheet to α-helix intensity plotted as a function of the varied concentration of LMPG micelles.

This revealed a critical detergent concentration to at approximately 0.45% LMPG (between 0.325 and 0.65%) denoted by the major step in intensities as the secondary structure convert.

### 5.4.2. Structural effects of VUx96 on C99

The vehicle sample (no VUx96) shows two distinct modes at 1632 cm$^{-1}$ and 1650 cm$^{-1}$ indicative of both β-sheet and α-helical secondary structures, respectively, within the aggregate structure. As the VUx96 small molecule binds in large concentrations (10 mM), I observe each mode blue shifts in absorbance frequency by approximately 6 cm$^{-1}$. While both secondary structures are still present, the blue shift to higher wavenumbers represents the weakening of coupling among the transition dipoles(21). The absorbance frequency is sensitive to the proximity of oscillators (backbone carbonyls) therefore the blue shifts can represent an increase in organizational heterogeneity among structures (22). It is difficult to predict the actual structure change, however the weakened coupling is often related to increased

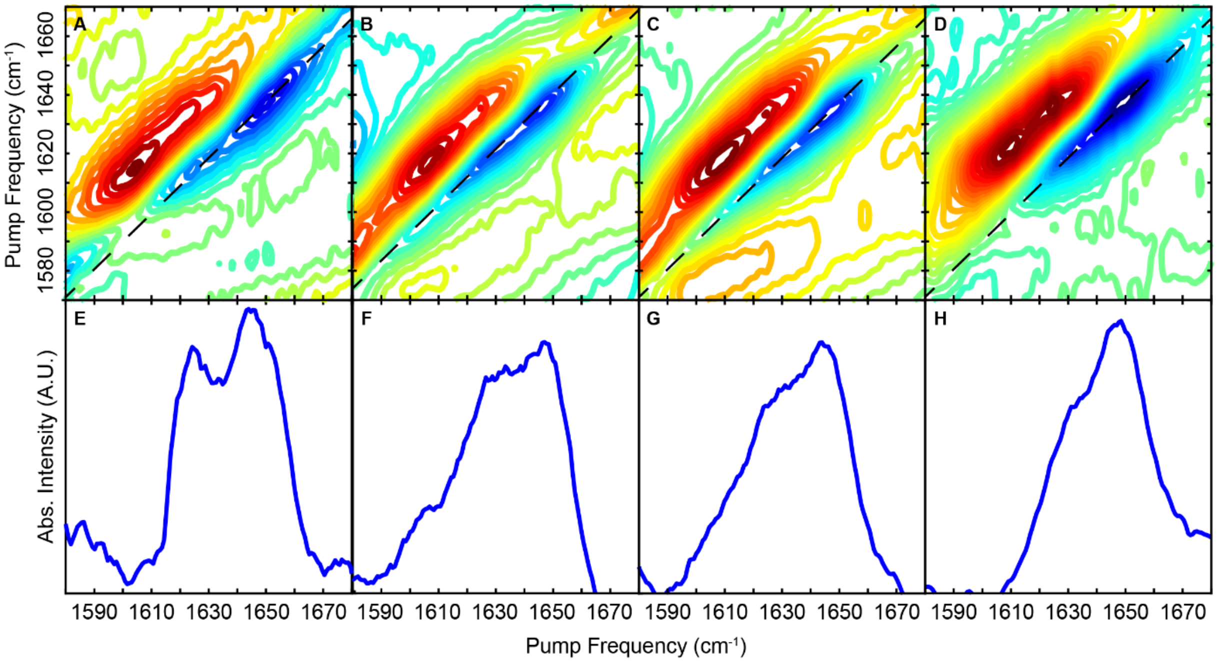

**Figure 5.3:** 2D IR contour maps and corresponding linear traces reveal that VUx96 disrupts and inhibits the formation of C99 β-sheet morphologies by varying relative amounts of VUx96 (A,E: no VUx96, B,F: 2 mM VUx96, C,G: 4 mM VUx96, D,H: 10 mM VUx96).

distance between β-strands, decreased periodicity in α-helices and fraying at the end of structures. The intensity ratio of the two modes also shifts when associated to VUx96. The mode at 1556 cm$^{-1}$ becomes more dominant in the spectra indicating VUx96 may be more disruptive to β-sheet structure than α-helices. This trend remains constant as the amount of VUx96 is decreased to 4 mM however the absorbance frequencies return closer to the values of the vehicle sample. This follows the theory of VUx96 being more perturbing to β-sheet structure as the ratio of the modes remains closer to the high VUx96 measurements. As the concentration of VUx96 is lowered further, the modes return to closely resemble the vehicle absorption. Therefore, the perturbation of C99 by VUx96 is largely concentration dependent with seemingly no meaningful perturbation below approximately 2 mM.

The truncated C99 (C74 and C55) were analyzed under the same condition as C99. Under most conditions, both C74 and C55 follow similar trends when exposed to VUx96 (appendix 1, Figure S1.10 and S1.11). In general, the 1630 $cm^{-1}$ mode is more rigorous depleted indicating VUx96 is more perturbative/targeting to β-sheets. The mode at 1652 $cm^{-1}$ representing α-helical structure is largely unperturbed. However, the mode does have a slight blue shift as well as spectral broadening. Spectral broadening mostly occurs from an increased range of oscillators experiencing different environment conditions(23). For example, embedded structures will have low rates of hydron bonding exchange while solvent exposed structure will have much higher rates. When structures have both embedded and exposed oscillators, the absorbance mode is broadened. While all spectra are normalized, VUx96 usually has the most intense absorption and therefore used for the standard value. However, VUx96 absorbance frequency change varies depending on the binding to protein, therefore some of the spectral subtraction look better than others. In almost as cases, the 10mM VUx96 samples are rough, particularly in the contour maps. This is largely due to the scattering of the light from the VUx96 molecules and the shifting absorption values. While possible work arounds include designing new polarization schemes or increased averaging, I wanted to keep condition equivalent over all spectra. Unfortunately, the native infrared spectra cannot be used to make quantified measurements of the amount of one structure vs the other. I am working on a method for measuring the transition dipole strengths of oscillators which is much more sensitive to the delocalization of the normal modes(24). These methods would allow for quantification (60% α-helix etc.). However, I wasn’t able to get accurate measurements for these sample largely due to other spectral contribution from the buffer and other components.

### 5.4.3. Investigating the effects of acetylation on amyloid beta 42

I first monitored the changes in the secondary structure of Aβ42 as the protein folds from the disordered starting state to the fully aggregated state using 2D IR spectroscopy. The 2D contour maps immediately show variety of structural differences between the amyloid beta variants (Figure 5.4). The Aβ42-WT, previously reported to form amyloid fibrils demonstrates a narrow peak centered at 1630 cm$^{-1}$ which is common among β-sheets within amyloidogenic proteins(22). The elongated lineshapes of the fundamental and overtone peaks show a varying degree of solvent effects suggesting the core of fibrils may be shielding from the strong hydrogen bonding effects(25, 26). In contrast, the Aβ42-K16 species shows a very broad absorption mode. While the mode is much broader, it is still centered at 1630 cm$^{-1}$ indicative of amyloid β-sheet structures. Rather than ellipsoidal, the peak pairs appear more spherical in natural. With the aggregate shown to be largely amorphous, this suggests similar solvent effects throughout the structure. Such attributes are often seen in porous structures with solvent readily exchanging evenly. Finally, the Aβ42-KKAc species closely resembles the wild-type spectrum. While the overtone and fundamental peaks are similarly elongated, there is a noticeable shoulder around 1645 cm$^{-1}$ in the double acetylated variant. Such shoulders are indicative of mixed secondary structure populations with amyloidogenic β-sheets and α-helixes and/or disordered structures within the peptide ensemble(19, 22, 24). While this native 2D IR study reveals several structural characteristics, a more resolved picture is needed to draw specific conclusion. This increased resolution is gained with the incorporation of isotope labels. However, isotope labeling schemes are difficult to predict for amyloidogenic structure that may contain may different oligomer species. To further probe the unlabeled protein, I employ TDS calculations. As expected, the main TDS peak is centered at the β-sheet frequency similar to the vibrational frequency shown above. However, the TDS intensity of this peak decreases as the degree of acetylation

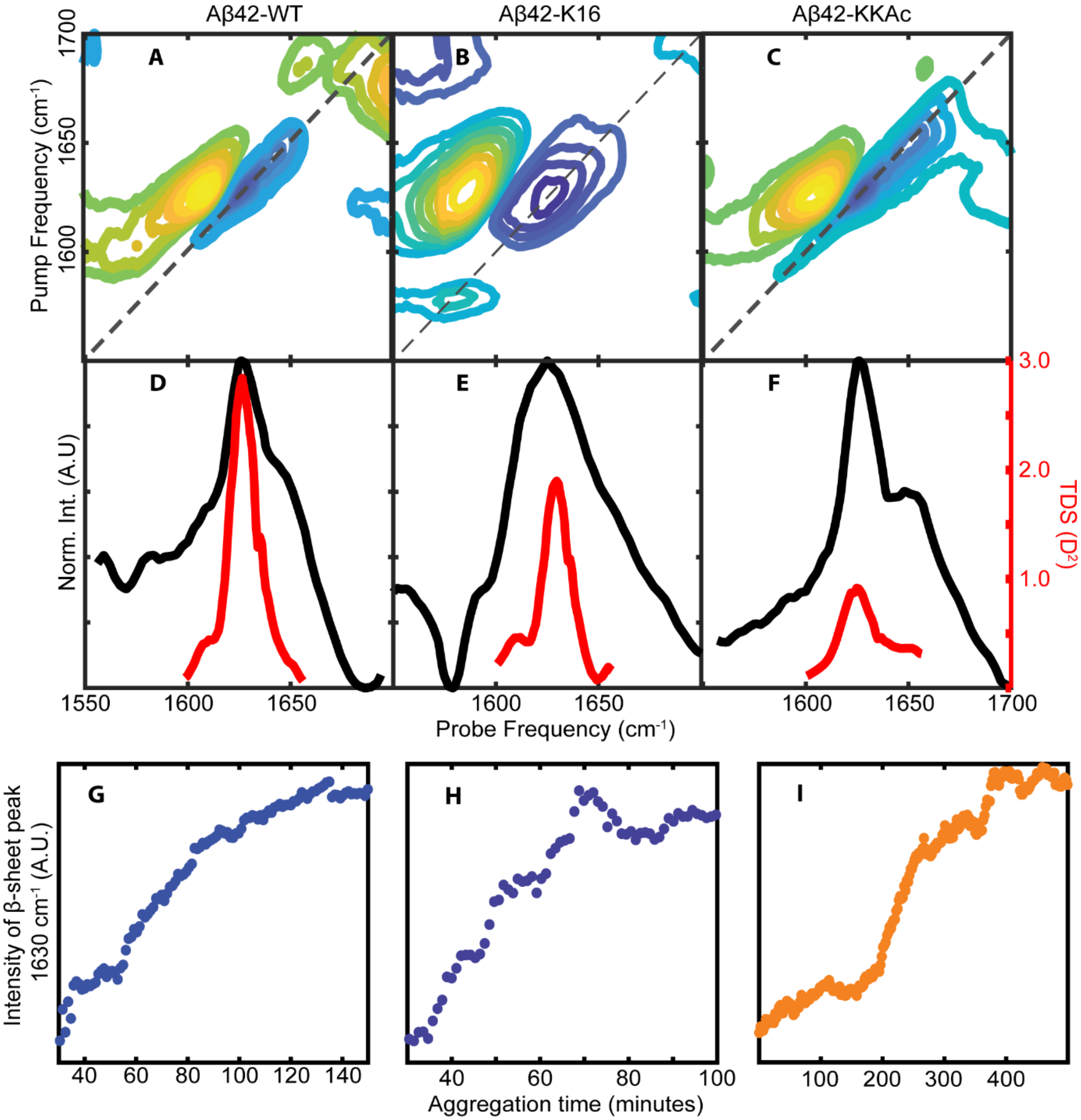

**Figure 5.4**: 2D IR contour maps and corresponding linear traces (black) and calculated TDS (red) of Aβ42-WT (A and D), Aβ42-K16 (B and E), and Aβ-KKAc (C and F) after 6 hours of aggregation. Aggregation kinetics are shown by plotted the intensity of the predominant β-sheet signal over time for Aβ42-WT (G), Aβ42-K16 (H) and Aβ42-KKAc (I).

increases. More specifically, the TDS is halved for every lysine that is acetylated. The wild-type TDS is shown to be approximately 2.8 $D^2$ while the K16 variant decreases to 1.6 $D^2$ and lastly the KKAc species decreases to 0.8 $D^2$. This systematic decrease in coupling strengths suggests that the acetylation is significantly hinder the vibrationally coupling by disrupting the surface over which the coupling can delocalize.

In addition to probing structure, 2D IR spectroscopy is also capable of analyzing the aggregation kinetics. The aggregation kinetics were tracked by plotting the intensity of the most prominent structural peak over the total aggregation before an equilibrium was reached noting the end of aggregation (Figure 5.4). I chose to monitor the 1630 $cm^{-1}$ corresponding to the amyloid β-sheet structures. All three variants were imaged via electron microscopy under the same conditions mentioned above to track the larger scale changes in supramolecular morphologies due to the acetylation of the lysine side chains. It is important these results are preliminary and are not a complete assessment of the full kinetic timeline. Ideally, the starting points of these graphs would have more lag phase (remain at the same intensity before increasing) to ensure the protein is in a completely disordered state(20, 27). As both the Aβ42-WT and Aβ42-K16 species begin to immediately increase in intensity, a fully disordered/disaggregated starting monomer state cannot be guaranteed.

## 5.5. Conclusion

This chapter demonstrates the wide applicability and novel insights provided by 2D IR spectroscopy in complex macro biomolecules ensembles beyond the detailed insights of model peptides. When studying amyloidogenic protein systems such as C99 and Aβ42, it is critical to understand a complete picture of aggregation from a true monomer state to the intermediate and fully aggregated structure. Finding the true monomeric starting states is the next major hurdle to solve in protein spectroscopy. Since the lifetimes of monomeric species are often incredibly fast, the sub-picosecond temporal resolution of 2D IR spectroscopy will serve as an excellent analysis method in future work. This work also demonstrates to the ability to analyze different types of biomolecules within the same aggregation ensembles.

### 5.6. References

1. Thinakaran, G., and E.H. Koo. 2008. Amyloid Precursor Protein Trafficking, Processing, and Function. *J. Biol. Chem.* 283:29615–29619:10.1074/jbc.R800019200.

2. Song, Y., K.F. Mittendorf, Z. Lu, and C.R. Sanders. 2014. Impact of Bilayer Lipid Composition on the Structure and Topology of the Transmembrane Amyloid Precursor C99 Protein. *J. Am. Chem. Soc.* 136:4093–4096:10.1021/ja4114374.

3. Lichtenthaler, S.F., D. Beher, H.S. Grimm, R. Wang, M.S. Shearman, C.L. Masters, and K. Beyreuther. 2002. The intramembrane cleavage site of the amyloid precursor protein depends on the length of its transmembrane domain. *Proc. Natl. Acad. Sci.* 99:1365–1370:10.1073/PNAS.032395699.

4. Ousson, S., A. Saric, A. Baguet, C. Losberger, S. Genoud, F. Vilbois, B. Permanne, I. Hussain, and D. Beher. 2013. Substrate determinants in the C99 juxtamembrane domains differentially affect γ–secretase cleavage specificity and modulator pharmacology. *J. Neurochem.* 125:610–619:10.1111/JNC.12129.

5. Winkler, E., F. Kamp, J. Scheuring, A. Ebke, A. Fukumori, and H. Steiner. 2012. Generation of Alzheimer disease-associated amyloid β 42/43 peptide by γ-secretase can be inhibited directly by modulation of membrane thickness. *J. Biol. Chem.* 287:21326–21334:10.1074/jbc.M112.356659.

6. Yamaguchi, T., T. Suzuki, T. Yasuda, T. Oishi, N. Matsumori, and M. Murata. 2012. NMR-based conformational analysis of sphingomyelin in bicelles. *Bioorg. Med. Chem.*

20:270–278:10.1016/J.BMC.2011.11.001.

7. Mourtas, S., B. Mavroidi, A. Marazioti, M. Kannavou, M. Sagnou, M. Pelecanou, and S.G. Antimisiaris. 2020. Liposomes Decorated with 2-(4′-Aminophenyl)benzothiazole Effectively Inhibit Aβ 1–42 Fibril Formation and Exhibit in Vitro Brain-Targeting Potential. *Biomacromolecules*. 21:4685–4698:10.1021/acs.biomac.0c00811.

8. Gunn, A.P., C.L. Masters, and R.A. Cherny. 2010. Pyroglutamate-Aβ: Role in the natural history of Alzheimer's disease. *Int. J. Biochem. Cell Biol.* 42:1915–1918:10.1016/j.biocel.2010.08.015.

9. Zhu, X., R.P. Bora, A. Barman, R. Singh, and R. Prabhakar. 2012. Dimerization of the Full-Length Alzheimer Amyloid β-Peptide (Aβ42) in Explicit Aqueous Solution: A Molecular Dynamics Study. *J. Phys. Chem. B*. 116:4405–4416:10.1021/jp210019h.

10. Ciudad, S., E. Puig, T. Botzanowski, M. Meigooni, A.S. Arango, J. Do, M. Mayzel, M. Bayoumi, S. Chaignepain, G. Maglia, S. Cianferani, V. Orekhov, E. Tajkhorshid, B. Bardiaux, and N. Carulla. 2020. Aβ(1-42) tetramer and octamer structures reveal edge conductivity pores as a mechanism for membrane damage. *Nat. Commun.* 11:3014:10.1038/s41467-020-16566-1.

11. Söldner, C.A., H. Sticht, and A.H.C. Horn. 2017. Role of the N-terminus for the stability of an amyloid-β fibril with three-fold symmetry. *PLoS One*. 12:e0186347:10.1371/journal.pone.0186347.

12. Adhikari, R., M. Yang, N. Saikia, C. Dutta, W.F.A. Alharbi, Z. Shan, R. Pandey, and A. Tiwari. 2020. Acetylation of Aβ42 at Lysine 16 Disrupts Amyloid Formation. *ACS Chem. Neurosci.* 11:1178–1191:10.1021/acschemneuro.0c00069.

13. Iyer, A., S.J. Roeters, N. Schilderink, B. Hommersom, R.M.A.A. Heeren, S. Woutersen,

M.M.A.E.A.E. Claessens, and V. Subramaniam. 2016. The Impact of N-terminal Acetylation of α-Synuclein on Phospholipid Membrane Binding and Fibril Structure. *J. Biol. Chem.* 291:21110–21122:10.1074/jbc.M116.726612.

14. Finder, V.H., I. Vodopivec, R.M. Nitsch, and R. Glockshuber. 2010. The Recombinant Amyloid-β Peptide Aβ1–42 Aggregates Faster and Is More Neurotoxic than Synthetic Aβ1–42. *J. Mol. Biol.* 396:9–18:10.1016/j.jmb.2009.12.016.

15. Mattson, M.P. 1997. Cellular actions of beta-amyloid precursor protein and its soluble and fibrillogenic derivatives. *Physiol. Rev.* 77:1081–1132.

16. Cook, C., J.N. Stankowski, Y. Carlomagno, C. Stetler, and L. Petrucelli. 2014. Acetylation: A new key to unlock tau's role in neurodegeneration. *Alzheimer's Res. Ther.* 6:1–8:10.1186/alzrt259.

17. Shim, S.H., D.B. Strasfeld, Y.L. Ling, and M.T. Zanni. 2007. Automated 2D IR spectroscopy using a mid-IR pulse shaper and application of this technology to the human islet amyloid polypeptide. *Proc. Natl. Acad. Sci. U. S. A.* 104:14197–14202:10.1073/pnas.0700804104.

18. Hamm, P., and M. Zanni. 2011. Concepts and Methods of 2D Infrared Spectroscopy. Cambridge: Cambridge University Press.

19. Lomont, J.P., K.L. Rich, M. Maj, J.J. Ho, J.S. Ostrander, and M.T. Zanni. 2018. Spectroscopic Signature for Stable β-Amyloid Fibrils versus β-Sheet-Rich Oligomers. *J. Phys. Chem. B*. 122:144–153:10.1021/acs.jpcb.7b10765.

20. Buchanan, L.E., E.B. Dunkelberger, and M.T. Zanni. 2012. Examining Amyloid Structure and Kinetics with 1D and 2D Infrared Spectroscopy and Isotope Labeling. In: Fabian H, D Naumann, editors. Protein Folding and Misfolding: Shining Light by Infrared

Spectroscopy. Berlin, Heidelberg: Springer Berlin Heidelberg. pp. 217–237.

21. Flanagan, J.C., and C.R. Baiz. 2019. Ultrafast pH-jump two-dimensional infrared spectroscopy. *Opt. Lett.* 44:4937:10.1364/OL.44.004937.

22. Lomont, J.P., J.S. Ostrander, J.-J. Ho, M.K. Petti, and M.T. Zanni. 2017. Not All β-Sheets Are the Same: Amyloid Infrared Spectra, Transition Dipole Strengths, and Couplings Investigated by 2D IR Spectroscopy. *J. Phys. Chem. B*. 121:8935–8945:10.1021/acs.jpcb.7b06826.

23. Buchanan, L.E., E.B. Dunkelberger, and M.T. Zanni. 2012. Protein Folding and Misfolding. 1st ed. Berlin, Heidelberg: Springer Berlin Heidelberg.

24. Dunkelberger, E.B., M. Grechko, and M.T. Zanni. 2015. Transition Dipoles from 1D and 2D Infrared Spectroscopy Help Reveal the Secondary Structures of Proteins: Application to Amyloids. *J. Phys. Chem. B*. 119:14065–14075:10.1021/acs.jpcb.5b07706.

25. Demirdöven, N., C.M. Cheatum, H.S. Chung, M. Khalil, J. Knoester, and A. Tokmakoff. 2004. Two-dimensional infrared spectroscopy of antiparallel β-sheet secondary structure. *J. Am. Chem. Soc.* 126:7981–7990:10.1021/ja049811j.

26. Ganim, Z., S.C. Hoi, A.W. Smith, L.P. Deflores, K.C. Jones, and A. Tokmakoff. 2008. Amide I two-dimensional infrared spectroscopy of proteins. *Acc. Chem. Res.* 41:432–441:10.1021/ar700188n.

27. Dunkelberger, E.B., L.E. Buchanan, P. Marek, P. Cao, D.P. Raleigh, and M.T. Zanni. 2012. Deamidation Accelerates Amyloid Formation and Alters Amylin Fiber Structure. *J. Am. Chem. Soc.* 134:12658–12667:10.1021/ja3039486.

## Chapter 6

## Summary and Future Work

### 6.1. Introduction

In this chapter, I aim to summarize the key contributions and research advances I have made in my graduate career as well as possible future research directions. I have had the privilege to apply in-depth physical chemistry techniques to interdisciplinary science. Integrating results from 2D NMR, light scattering, and electron microscopy with rapidly progressing techniques in 2D IR spectroscopy has made for a fulfilling graduate career in the research I've completed over the last five and a half year.

### 6.2. Investigating N-terminal acetylation on model amphiphilic peptides

My first project, with the help of Daniel DeNeve, investigating how hydrophobic amino acid repeats in primary sequence effect peptide structure yielded limited results (results found in the thesis of Cpt. Daniel DeNeve). At this point, the project direction was shifted toward the effects of N-terminal acetylation. In my research, both AcKFE8 and KFE8 were extensively studied through TEM, DLS, 2D IR spectroscopy, and transition dipole coupling simulations. By incorporating isotope-labels in 2D IR spectroscopy, I was able to precisely detect alignments of the β-sheet arrangements to single-residue resolution(1). The use of computational simulations greatly increased our confidence levels in our labeling scheme design and overall results of this study. The COSMOSS matlab script used for the vibrational spectra simulations is available open access through Github and should be used in vibrational studies in the future. In many cases, the coding can be further improved to encapsulate a larger range of structures including complex protein structures exported from the PDB database. This capability would have major implications in future work within vibrational spectroscopy.

In my research, I focused on the well-established AcKFE8 ($COCH_3$-FKFEFKFE-$NH_2$) and the non-acetylated species KFE8 ($NH_2$-FKFEFKFE-$NH_2$). However, simply shifting the amino acid sequence, such as in the sequence $COCH_3$-KFEFKFEF-$NH_2$, may have major implications on the self-assembly and structure (2–5). By having a charged residue (lysine) on the N-terminus, it would add the effects of electrostatics to changes already present caused by the N-terminal acetylation This would make for an interesting study to build off this work using largely established methodology.

The Vanderbilt Institute of Nanoscale Science and Engineering (VINSE) is an incredible useful resource for progressing scientific research. VINSE is always looking to push the frontier of material research, including protein biomaterials. With the help of VINSE, I've established useful methods for analyzing peptide structures through light scattering and cryogenic electron microscopy. There is still great potential to utilize both atomic force and fluorescence microscopy within the VINSE cores to generate data to strengthen observations made with the optical IR spectroscopy.

### 6.3. Probing amyloidogenic oligomers using transition dipole strengths

In Chapter 4, I demonstrated the increased sensitivity provided by transition dipole strengths (TDS) compared to vibrational frequency alone. In addition to detecting the separate β-sheet stacking arrangements within AcKFE8, TDS calculations also suggest the presence of polymorphic β-sheet species previously hidden in the vibrational spectrum of AcKFE8(6). The annealing of the two major β-sheet arrangements yield coupling strengths that overshadow the absorption contributions of the minor species. While the sensitivity of TDS allows for the detection of polymorphs, it can not specify the exact structure of the β-sheet aggregates without information obtained from isotope-labeled experiments. Theoretically, TDS values can be calculated for the isotope-labeled residue coupling as well(7, 8). Hoverer, as the isotope coupling strength is isolated to coupling across single residues, the intensity of the linear absorption is insufficient to distinguish signal from noise(9). While the pulse shaping implemented in our 2D IR experiments has many advantages, one disadvantages lies in only 5-10% of the beam is being used in the probe path. As the probe is solely responsible for the linear absorbance spectra, the signal is often low distorting the linear absorbance. It may be possible that data acquired through 2D IR BOXCAR set-ups would increase the linear absorbance resolution providing enough signal to allow for these calculations(10).

While TDS calculations showed the disruption of the coupling across the acetylated Aβ42 variant surface, I ideally would also like to use this same technique to calculate TDS values in C99 associated with detergent micelles. However, the LMPG detergent also absorbs in the amide I’ spectral region. To generate the 2D IR spectra of C99, LMPG spectra with the same concentration detergent and corresponding number of scans were subtracted from the C99/LMPG spectra. Therefore, TDS values that depend on the division of spectral values are assumed to be skewed and inaccurate. It may be possible to formalize a method to incorporate TDS calculations within subtracted spectra in future work.

The variation in secondary structure within Aβ42 is obvious from the 2D IR spectroscopy results presented in chapter 5. In the next step in this research would be to design isotope labeling scheme to isolate the structural differences within Aβ42 with single residue resolution. Modeling different isotope labeling schemes through COSMOSS would provide a screening method for successful designs and preliminary simulation data.

The ability to truly achieve a monomeric starting ensemble is equally important as probing the actual structure. I believe the concentration requirements (≥100 μM Aβ42) for sufficient signal detection in this research may have been too high for monomeric structures to survive for any observable timescales. To test this theory of concentration dependence, Thioflavin T fluorescence binding can be conducted to reveal an aggregation trend timeline and be further refined through 2D IR spectroscopy in the future(11–13). The insights provided by this work and future investigation will prove vital in understanding the forces that drive self-assembly and aid in the design of protein biomaterials.

## 6.4. References

1. Weeks, W.B., C.J. Tainter, and L.E. Buchanan. 2022. Investigating the effects of N-terminal acetylation on KFE8 self-assembly with 2D IR spectroscopy. *Biophys. J.* 121:1549–1559:10.1016/j.bpj.2022.03.003.

2. Altman, M., P. Lee, A. Rich, and S. Zhang. 2000. Conformational behavior of ionic self-complementary peptides. *Protein Sci.* 9:1095–1105:10.1110/ps.9.6.1095.

3. Zhou, P., L. Deng, Y. Wang, J.R. Lu, and H. Xu. 2016. Different nanostructures caused by competition of intra- and inter-β-sheet interactions in hierarchical self-assembly of short peptides. *J. Colloid Interface Sci.* 464:219–228:10.1016/J.JCIS.2015.11.030.

4. Morris, C., S. Cupples, T.W. Kent, E.A. Elbassal, E.P. Wojcikiewicz, P. Yi, and D. Du. 2018. N-Terminal Charged Residues of Amyloid-β Peptide Modulate Amyloidogenesis and Interaction with Lipid Membrane. *Chem. - A Eur. J.* 24:9494–9498:10.1002/chem.201801805.

5. Herrera Estrada, L.P., and J.A. Champion. 2015. Protein nanoparticles for therapeutic protein delivery. *Biomater. Sci.* 3:787–799:10.1039/c5bm00052a.

6. Weeks, W.B., and L.E. Buchanan. 2022. Label-Free Detection of β-Sheet Polymorphism. *J. Phys. Chem. Lett.* 13:9534–9538:10.1021/acs.jpclett.2c02292.

7. Dunkelberger, E.B., M. Grechko, and M.T. Zanni. 2015. Transition Dipoles from 1D and 2D Infrared Spectroscopy Help Reveal the Secondary Structures of Proteins: Application to Amyloids. *J. Phys. Chem. B.* 119:14065–14075:10.1021/acs.jpcb.5b07706.

8. Grechko, M., and M.T. Zanni. 2012. Quantification of transition dipole strengths using 1D

and 2D spectroscopy for the identification of molecular structures via exciton delocalization: Application to α-helices. *J. Chem. Phys.* 137:184202:10.1063/1.4764861.

9. Ding, B., A. Panahi, J.-J. Ho, J.E. Laaser, C.L. Brooks, M.T. Zanni, and Z. Chen. 2015. Probing Site-Specific Structural Information of Peptides at Model Membrane Interface In Situ. *J. Am. Chem. Soc.* 137:10190–10198:10.1021/jacs.5b04024.

10. Shim, S.H., D.B. Strasfeld, Y.L. Ling, and M.T. Zanni. 2007. Automated 2D IR spectroscopy using a mid-IR pulse shaper and application of this technology to the human islet amyloid polypeptide. *Proc. Natl. Acad. Sci. U. S. A.* 104:14197–14202:10.1073/pnas.0700804104.

11. Xue, C., T.Y. Lin, D. Chang, and Z. Guo. 2017. Thioflavin T as an amyloid dye: fibril quantification, optimal concentration and effect on aggregation. *R. Soc. Open Sci.* 4:160696:10.1098/rsos.160696.

12. Di Carlo, M.G., V. Minicozzi, V. Foderà, V. Militello, V. Vetri, S. Morante, and M. Leone. 2015. Thioflavin T templates amyloid β(1–40) conformation and aggregation pathway. *Biophys. Chem.* 206:1–11:10.1016/j.bpc.2015.06.006.

13. Watson, M.D., and J.C. Lee. 2019. N-Terminal Acetylation Affects α-Synuclein Fibril Polymorphism. *Biochemistry*. 58:3630–3633:10.1021/acs.biochem.9b00629.

# Appendix 1

## Supporting figures

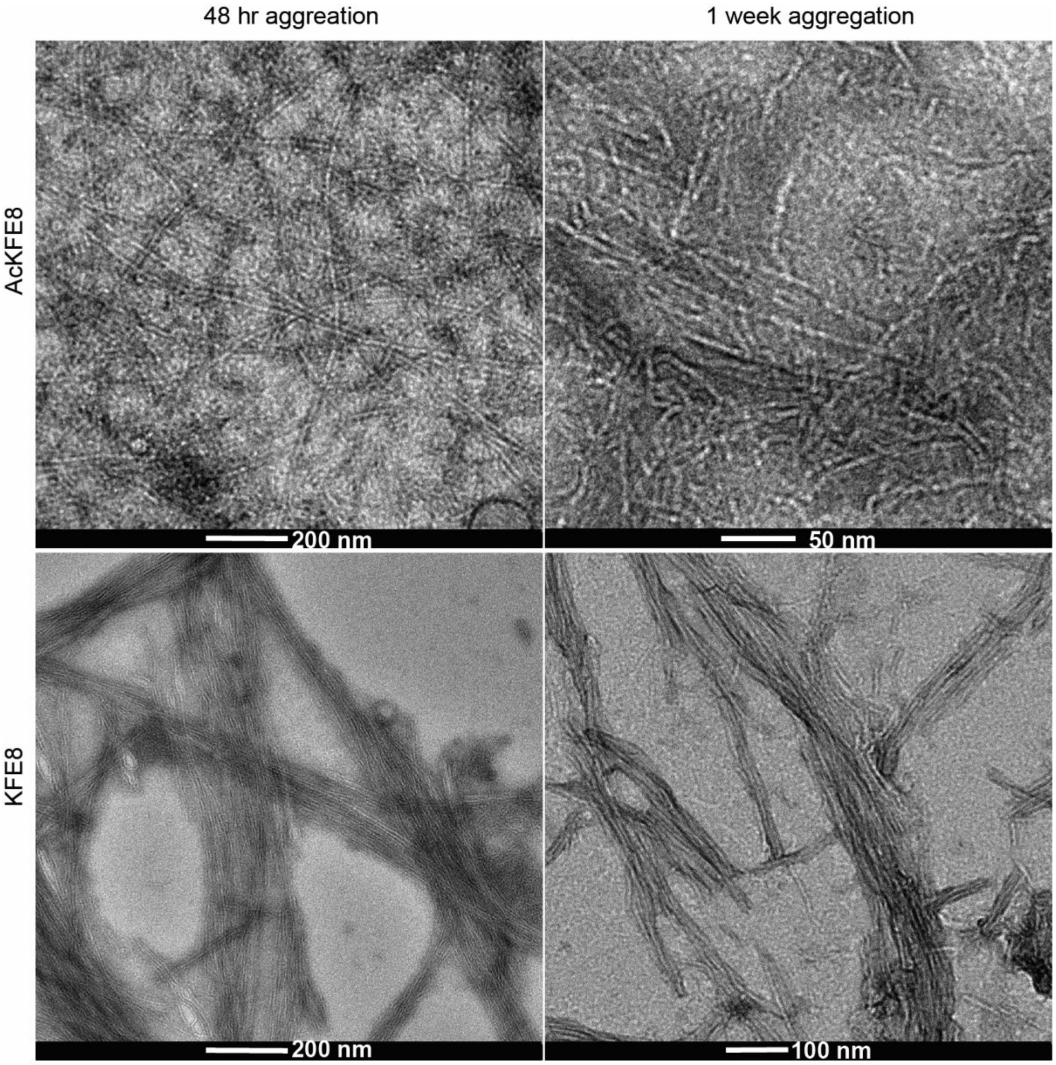

**Figure A1.1:** Transmission electron micrographs of AcKFE8 (top row) and KFE8 (bottom row) demonstrating that both helical (AcKFE8) and flat (KFE8) peptide ribbon morphologies are stable at 1 mM concentration in solution for up to a week.

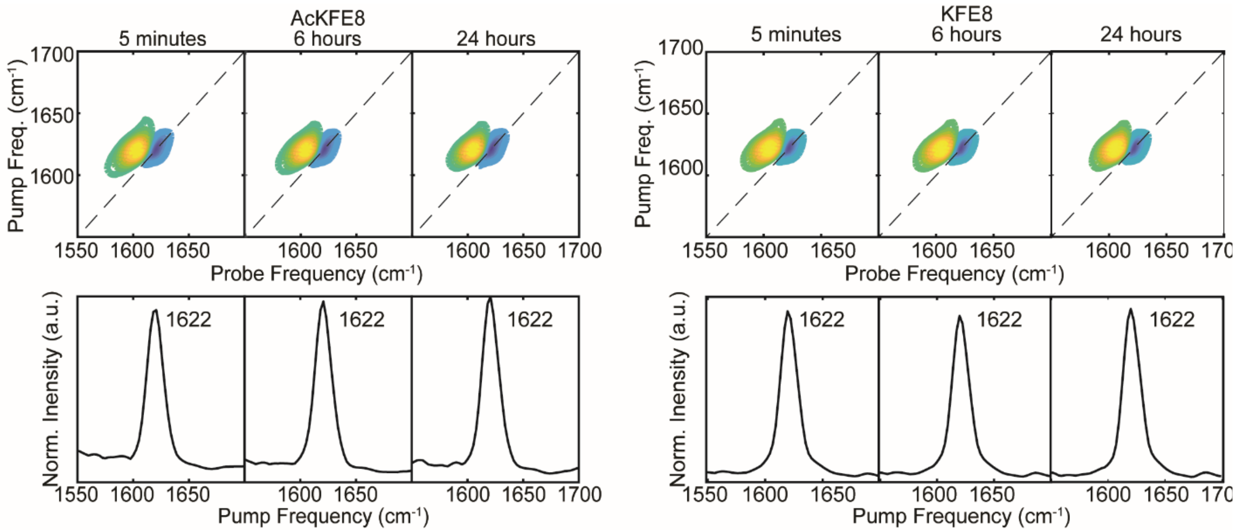

**Figure A1.2**: 2D IR spectra and corresponding one dimensional trace calculated via pump slice amplitude analysis of AcKFE8 (left) and KFE8 (right). The vibrational frequency remains unchanged in both peptides over 24 hours.

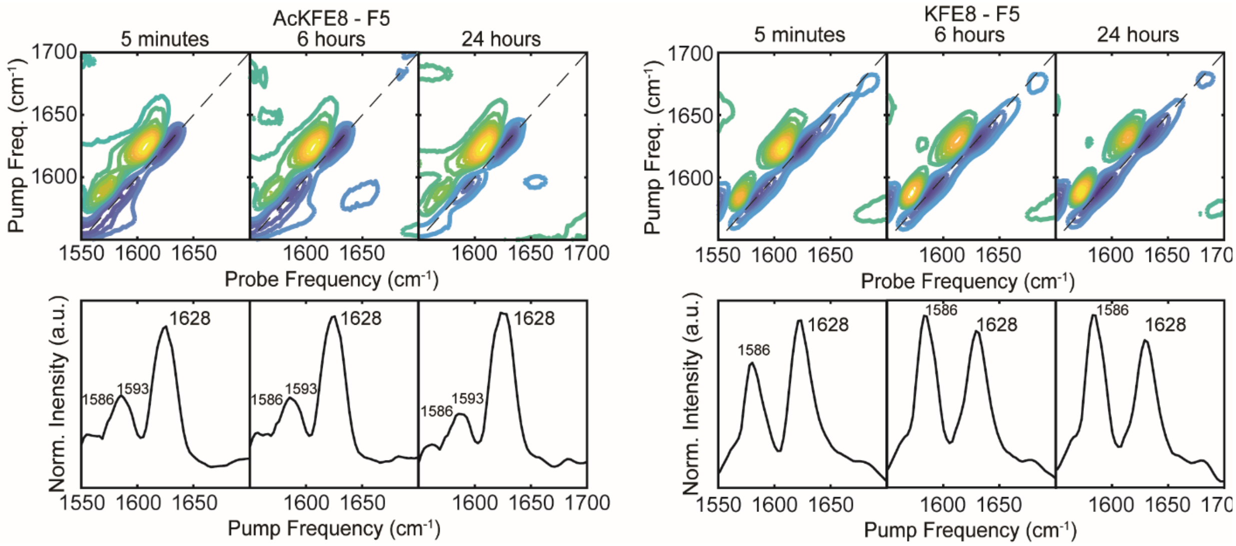

**Figure A1.3**: 2D IR spectra and pump slice amplitudes of AcKFE8 (left) and KFE8 (right) each containing a $^{13}C^{18}O$ label at residue F5. Although relative intensities change, frequencies remain constant over 24 hours.

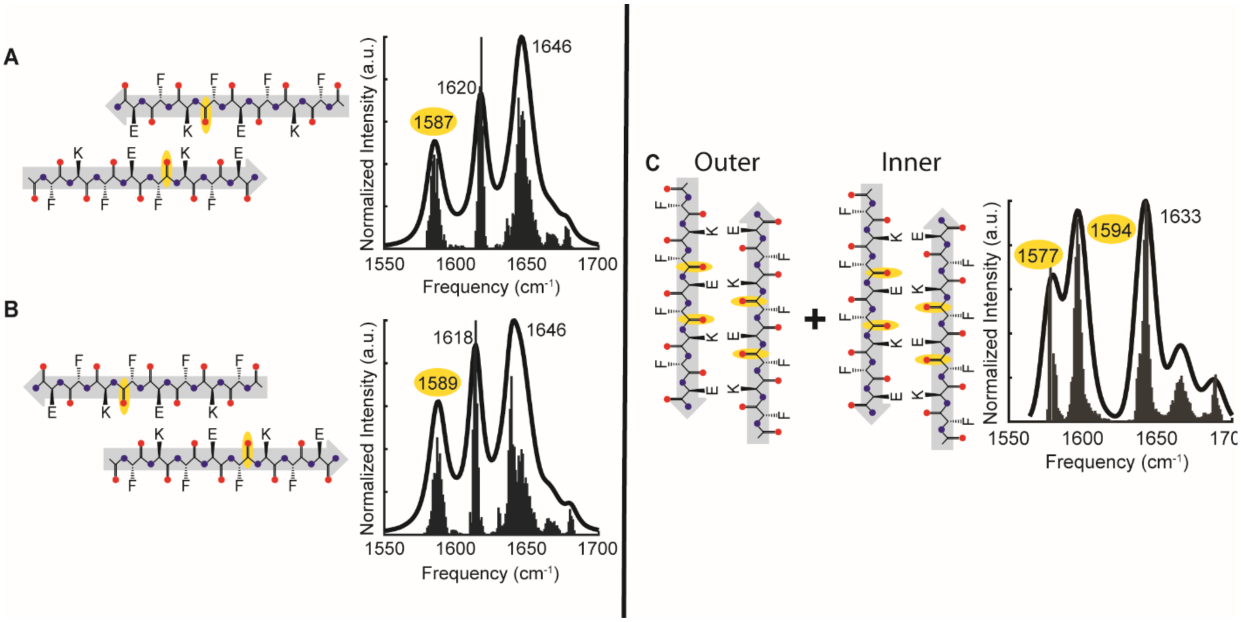

**Figure A1.4:** Additional COSMOSS simulations performed to rule out certain β-sheet stacking arrangements. Flat β-sheet stacking arrangements with increased overhang of the N-terminus (A) or C-terminus (B) lead to a dramatic increase of random coil features that would be highly visible via FTIR and 2D IR. Experimental spectra do not exhibit these features, so these strand alignments were discarded as possible structures. Simulated spectra of the double N-exposed helix in which residues F3 and F5 have been labeled (C) reveal both two isotope peaks in which the uncoupled peak at 1594 cm$^{-1}$ is more intense than the coupled peak at 1577 cm$^{-1}$, which does not match the experimental spectrum of doubly labeled AcKFE8.

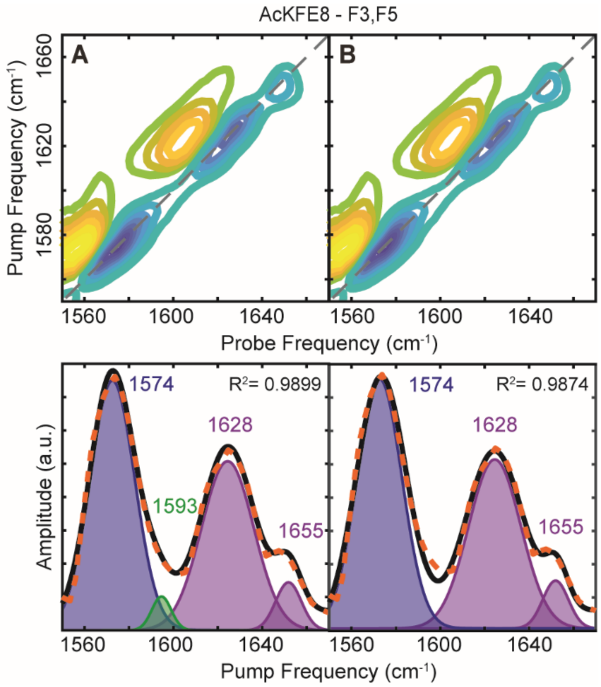

**Figure A1.5**: 2D IR contours and pump slice amplitudes of AcKFE8 doubly labeled at residues F3 and F5. The slices were fit to a sum of Gaussian functions, with the combined fit of the traces (dotted orange line) and fits for the unlabeled (purple) and $^{13}C^{18}O$-labeled (green and blue) peaks plotted individually. A better fit is achieved by including contributions from both coupled and uncoupled $^{13}C^{18}O$ modes (A) rather than just a single coupled isotope mode (B).

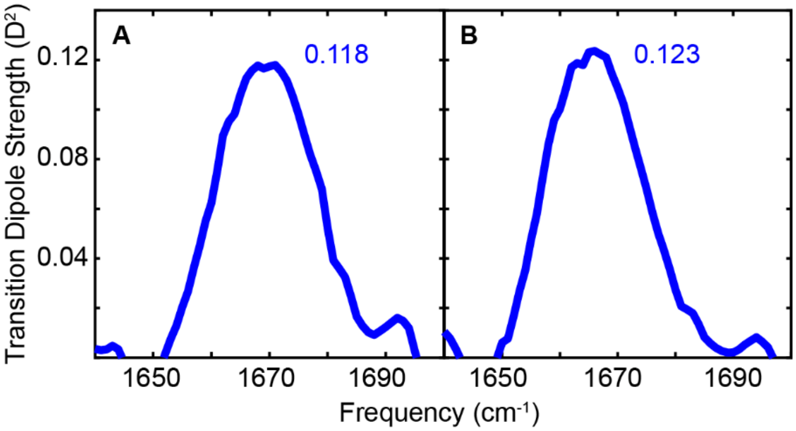

**Figure A1.6**: Calculated TDS spectra of KFE8 (A) and AcKFE8 (B) disaggregated in deuterated DMSO. The TDS values of both peptides match that of NMA 2, confirming that the amide I' modes are fully decoupled and thus the peptides are completely disaggregated. The peak frequency is blue-shifted from the standard disordered amide I' frequency of 1645 cm$^{-1}$ to the solvatochromic effect of DMSO (1, 2).

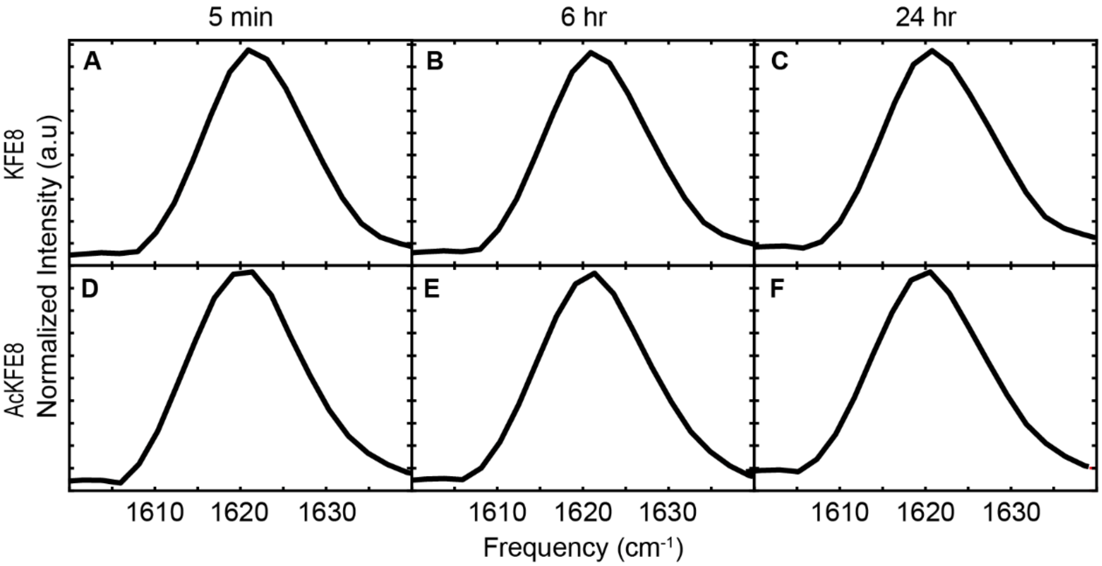

**Figure A1.7:** 2D IR linear traces of KFE8 (top) and AcKFE8 (bottom) at each timepoint.

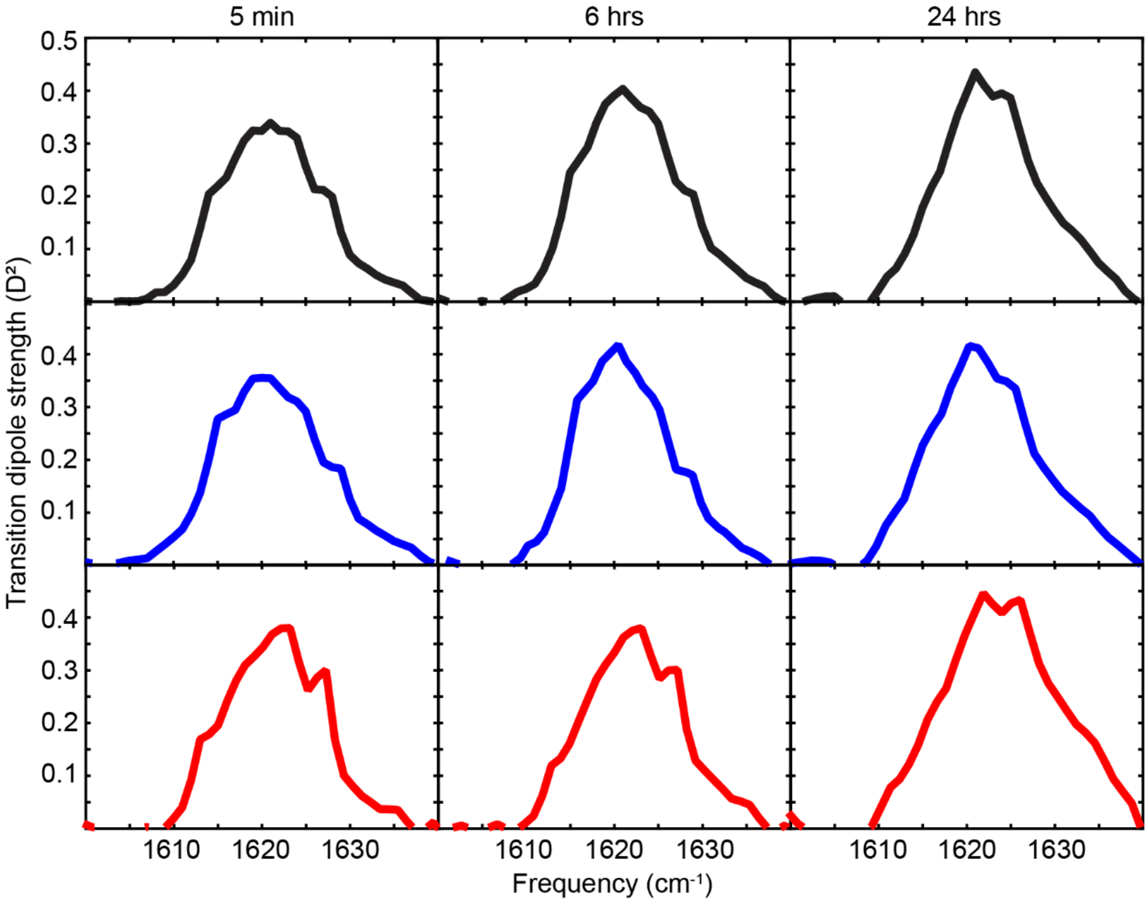

**Figure A1.8**: Calculated TDS spectra for three distinct samples of AcKFE8 after 5 minutes, 6 hours, and 24 hours of aggregation. Four maxima at 1615 cm$^{-1}$, 1621 cm$^{-1}$, 1625 cm$^{-1}$, and 1629 cm$^{-1}$ are reliably present at 5 minutes and 6 hours, while 1621 cm$^{-1}$ and 1625 cm$^{-1}$ features dominate after 24 hours. While the maximum TDS value for each feature differs slightly from sample to sample at early timepoints, indicating subtle differences in precise structure of the βsheets, 4 the 1621 cm$^{-1}$ and 1625 cm$^{-1}$ features consistently measure 0.43 D$^2$ and 0.4 D$^2$ after 24 hours.

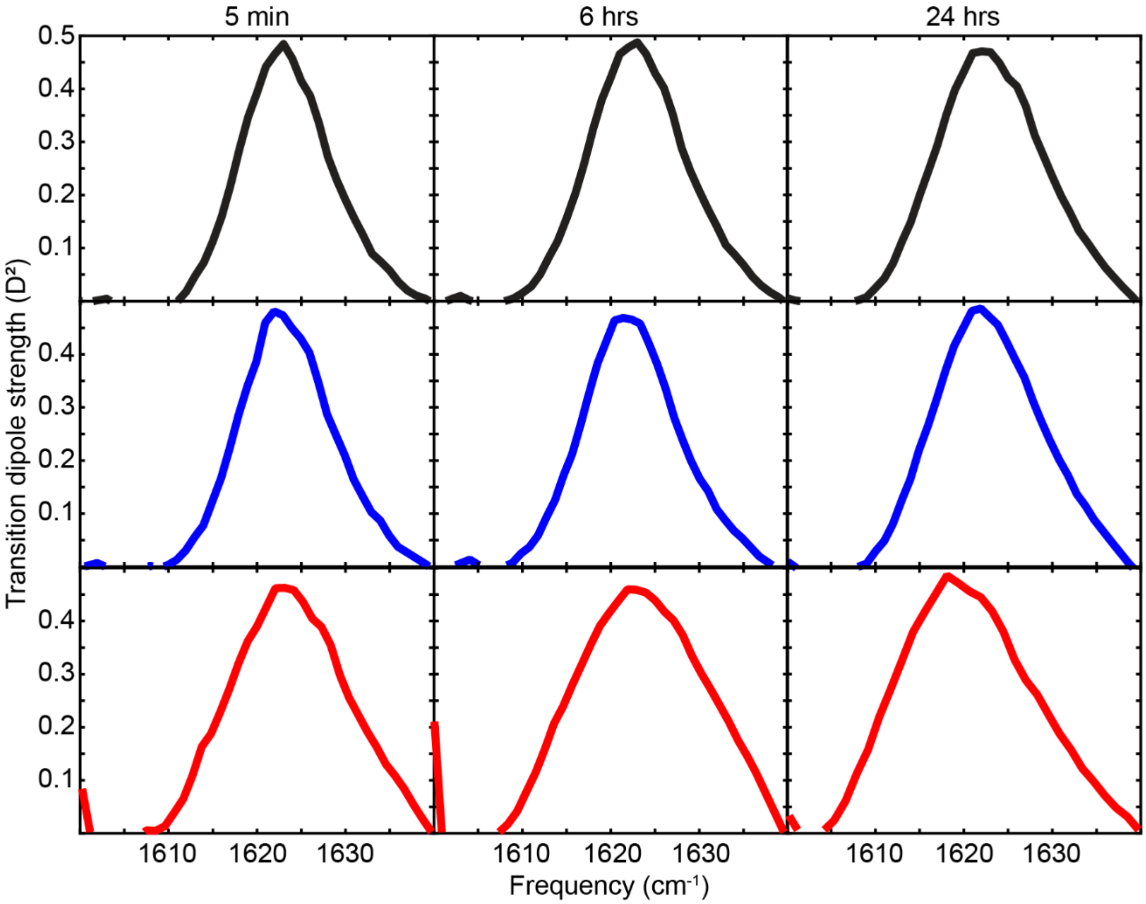

**Figure A1.9:** Calculated TDS spectra for three distinct samples of KFE8 after 5 minutes, 6 hours, and 24 hours of aggregation. The single peak is reproducible for all samples, with a maximum TDS of 0.49 ± 0.008 D$^2$ at 1621 cm$^{-1}$.

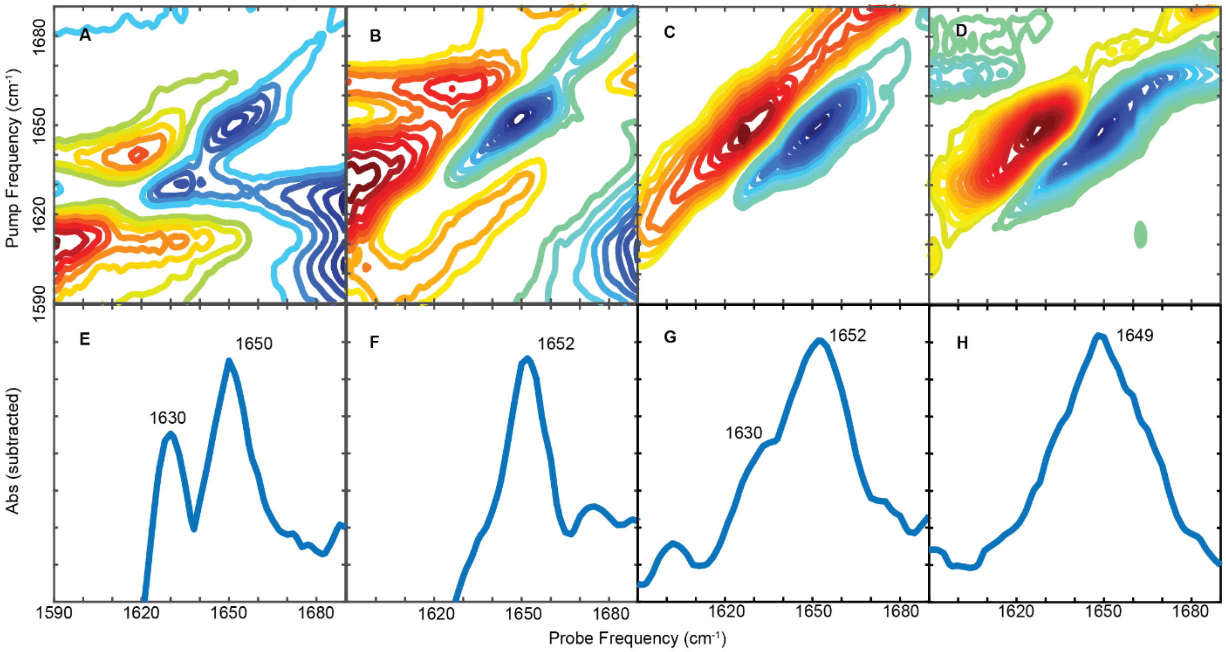

**Figure A1.10**: 2D IR contour maps and corresponding linear traces reveal that VUx96 disrupts and inhibits the formation of C74 β-sheet morphologies by varying relative amounts of VUx96 (A,E: no VUx96, B,F: 2 mM VUx96, C,G: 4 mM VUx96, D,H: 10 mM VUx96).

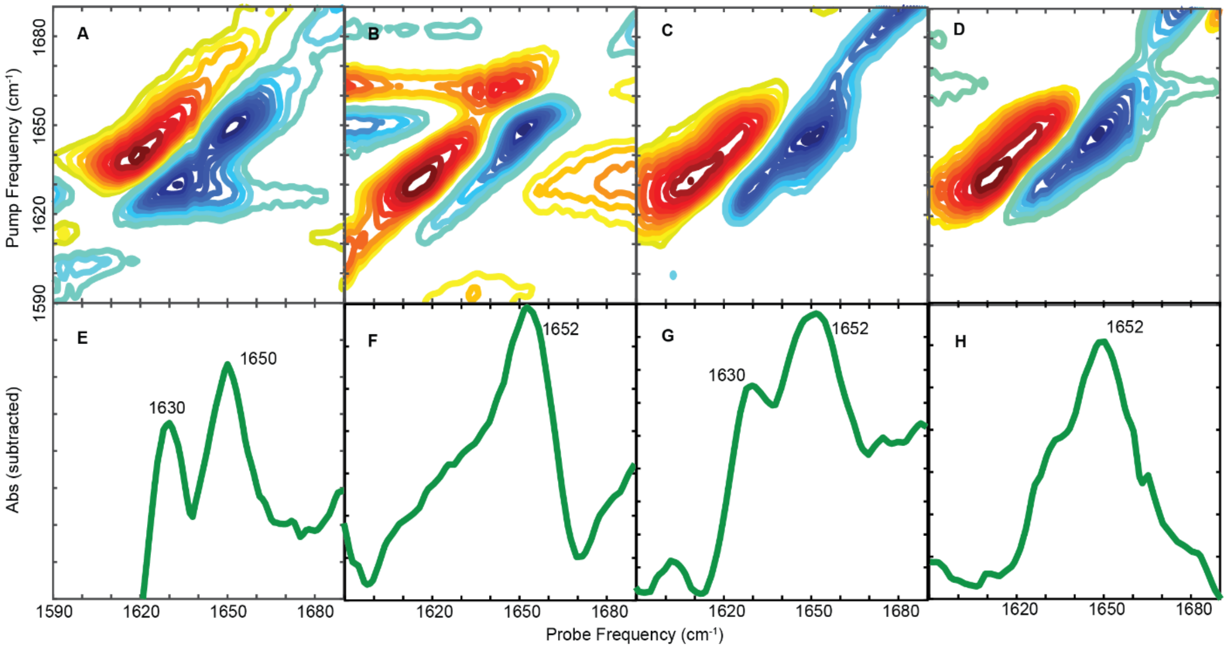

**Figure A1.11**: 2D IR contour maps and corresponding linear traces reveal that VUx96 disrupts and inhibits the formation of C55 β-sheet morphologies by varying relative amounts of VUx96 (A,E: no VUx96, B,F: 2 mM VUx96, C,G: 4 mM VUx96, D,H: 10 mM VUx96).

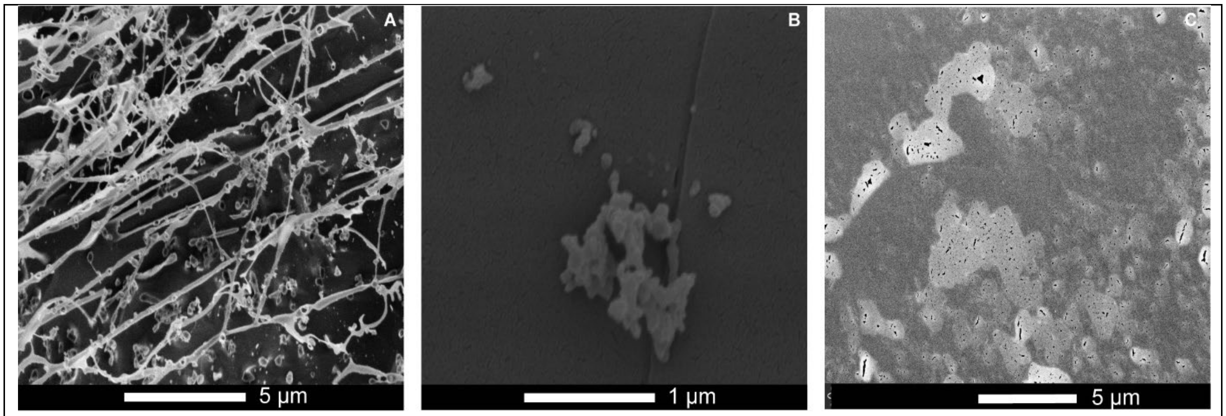

**Figure A1.12**: Cryo-SEM images of Aβ42-WT amyloid fibrils (A) and the amorphous undefined structures of Aβ42-K16 and Aβ42-KKAc

# Appendix 2

## Example TDS calculation of AcKFE8 with steps

1. 2D IR spectra composed of the same number of scans (I usually use 20 scans) should be obtained from the protein sample, the calibrant, and blank solutions of the buffer systems used for the protein and calibrant molecules. Here, both the protein and calibrant solvent is $D_2O$.

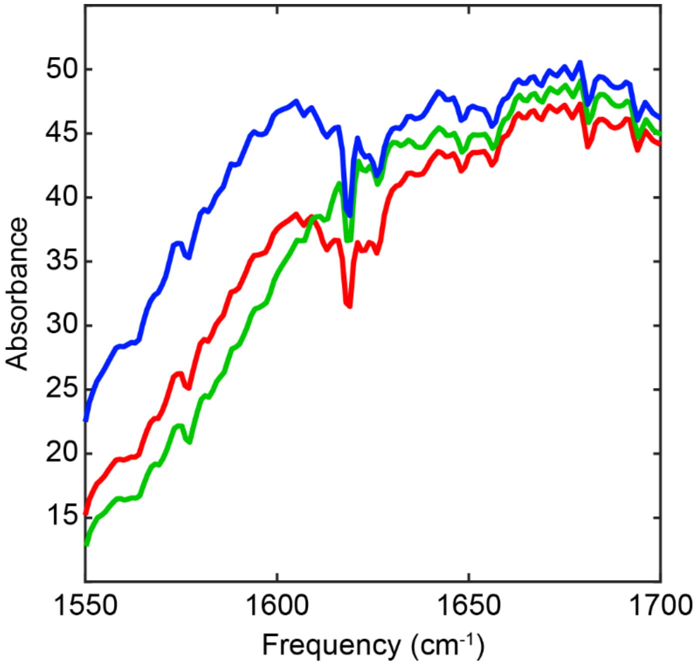

**Figure A2.1:** Averaged voltage measurements from the array detector of AcKFE8 (red), L-Serine (blue), and $D_2O$ (green)

2. The diagonal should be set through the center of the fundamental peak(s) as accurately as possible for both protein sample and calibrant. It is important to use the same slope and intercept values for the corresponding buffer scan.
3. Save the following variables as a .mat from the LEB_V2 matlab code: “pumpFreqs”, “probeFreqs”, “pumpRange”, “probeRange”, “linear”, and “sliceNeg”
4. Insert the .mat file name and path into the TDS code and load the data (TDS_V6_clean is currently the best code)

5. The "linear" variable is the voltage readings of 5 scans that are averaged into the avg_linear variable (Figure A2.1) and interpolated across the probeFreqs to match the correct frequency regime.
6. The sample and buffer avg_linear valves are divided and operated on by -log to yield the linear optical density (OD) (Figure A2.2)

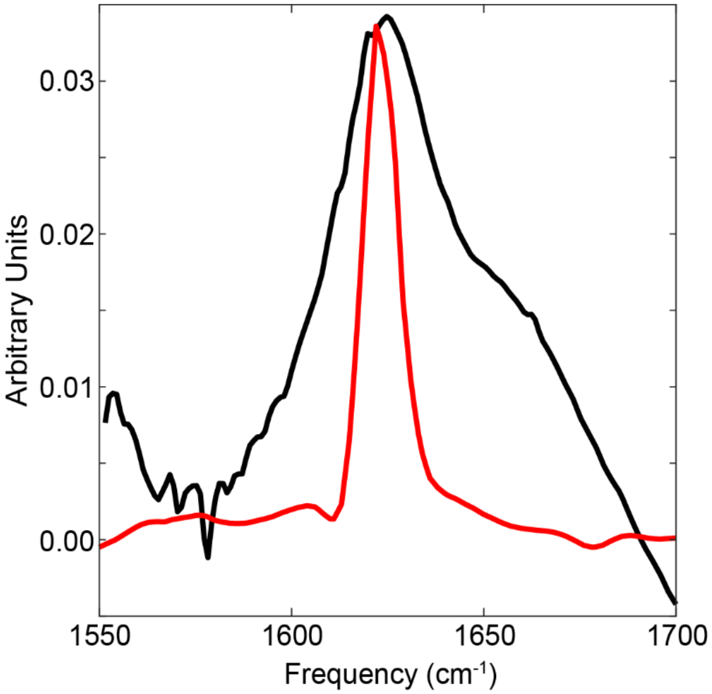

**Figure A2.2**: Calculated linear OD spectra of AcKFE8 (red) and L-Serine (black)

7. The linear OD (lin_OD) spectrum usually results in a distorted baseline. A polynomial function should be subtracted from the linear OD to correct the baseline and yield the fit_linOD variable. Two common ways to fit the lin_OD are to use a first degree polynomial fitting the edges of the spectrum, omitting the area of interest (1600-1660 $cm^{-1}$) or to use a second degree polynomial to fit a very small area just adjacent to the area of interest.
8. The "sliceNeg" variable undergoes a sign flip to make the values positive and is taken as the ΔOD in both calibrant and sample spectra. (Figure A2.3)

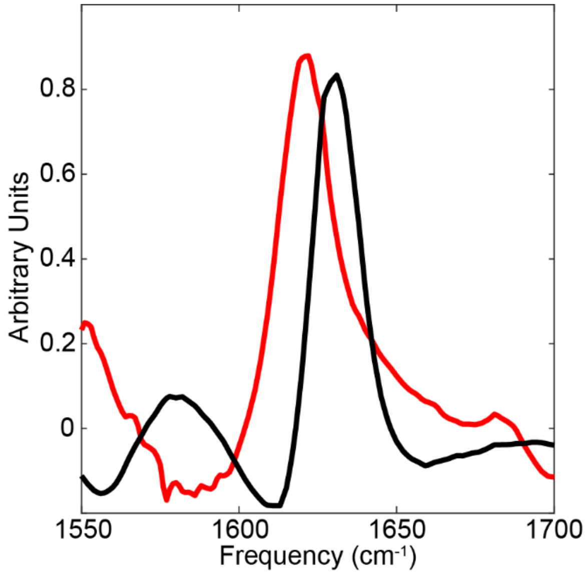

**Figure A2.3**: ΔOD spectra of AcKFE8 (red) and L-Serine (black)

9. The maximum ΔOD value from the calibrant spectrum is divided by the maximum value from the lin_OD spectrum and yield a single value.
10. The sample ΔOD spectrum is divided by the fitted lin_OD spectrum to yield a TDS spectrum. The values are scaled by dividing by the value in Step nine and multiplying the values by the known TDS value (0.2 for Ser) of the calibrant molecule. (Figure A2.4)

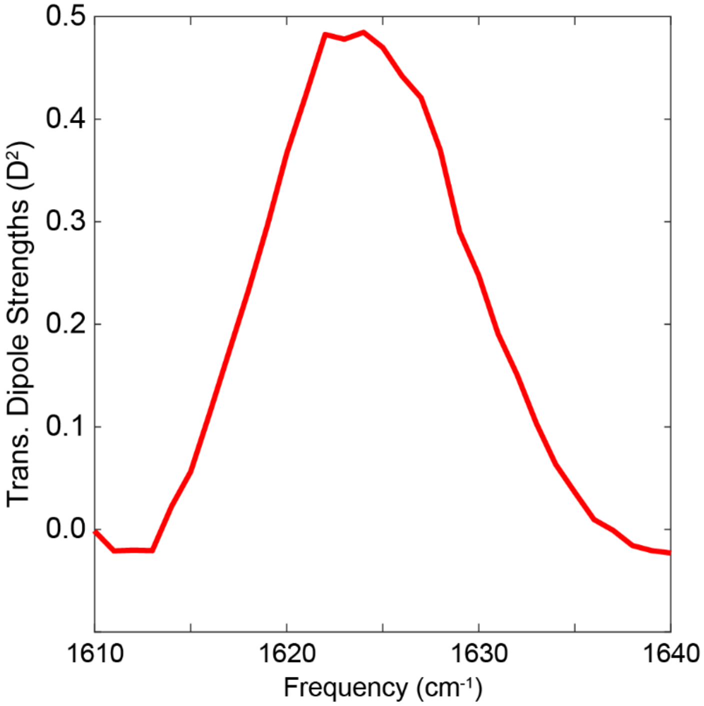

**Figure A2.4**: Final TDS spectrum of AcKFE8

# Appendix 3

## Troubleshooting for CEM Liberty Blue$^{TM}$ microwave-assisted solid-phase peptide synthesizer

I have collected useful checks and procedures for troubleshooting, maintaining, and repairing frequently occurring problems caused from regular use. These techniques have been provided from user manuals and correspondence from CEM personal and maintenance technicians. Shutdown procedures: The key to sustained synthesizer use is proper maintenance once a run is finished. First, each used amino acid position, activator and activator base should be backwashed and backpurged (options menu → maintenance → cleaning → select used positions). The tubes with pure DMF can then be replaced in their proper positions. Secondly, the deprotect solution must be backpurged as the piperazine solution is notorious for crashing out and causing clogs in the lines particularly during extended periods of no use (options menu → maintenance → backpurge → backpurge deprotect). Verify the reaction vessel is empty and run the operation. Watch for amply bubbles in the leftover deprotect solution to verify air is freely flowing with clogs. When prompted to replace the bottles, empty, and clean the deprotect bottle and select cancel so the system does not reprime the solution. Finally, preform the extended rotary wash (options menu → maintenance → wash → extended rotary wash) and depressurize the system before powering off and exiting out of the laptop software.

1. **PMAIN pressure check fails:** Most errors occur due to lose of pressure in the system. The “Pmain” check is the first error donating a pressure leak in the system. The synthesizer will not run without passing the Pmain pressure check. The Pmain check refers to the nitrogen manifold located in the lower instrument body through the back doors. The Pmain valve opens to fill the nitrogen manifold from the supply source. From the nitrogen manifold, nitrogen gas is distributed to the system where needed to pressurize the reservoir positions, purge reagents to the reaction vessel and to

pressurize the reaction vessel to drain it. If there is a leak, the system will have difficulty accomplishing these tasks. The most probable causes for this leak are:

a. Main solvent or deprotection bottle caps are loose. Check the O-rings inside the caps are secure and the threaded tube connections are tightly sealed on the outside of the cap

b. Fritted ferrules are leaking. These are the yellow connections/sealers on the end of the solvent lines inside the threaded tube connections.

    i. Checking Rotary 1: Remove the nut/tubing from the Pmain/R1 port. Screw the plug nug into the port. Run the Pmain leak test. If it is stable, the issue is with the fritted ferrules in this line. If it is not, the fritted ferrule is not the issue, and you need to go to next step (ii). In either case, remove plug nut and reattach the tubing into its port.

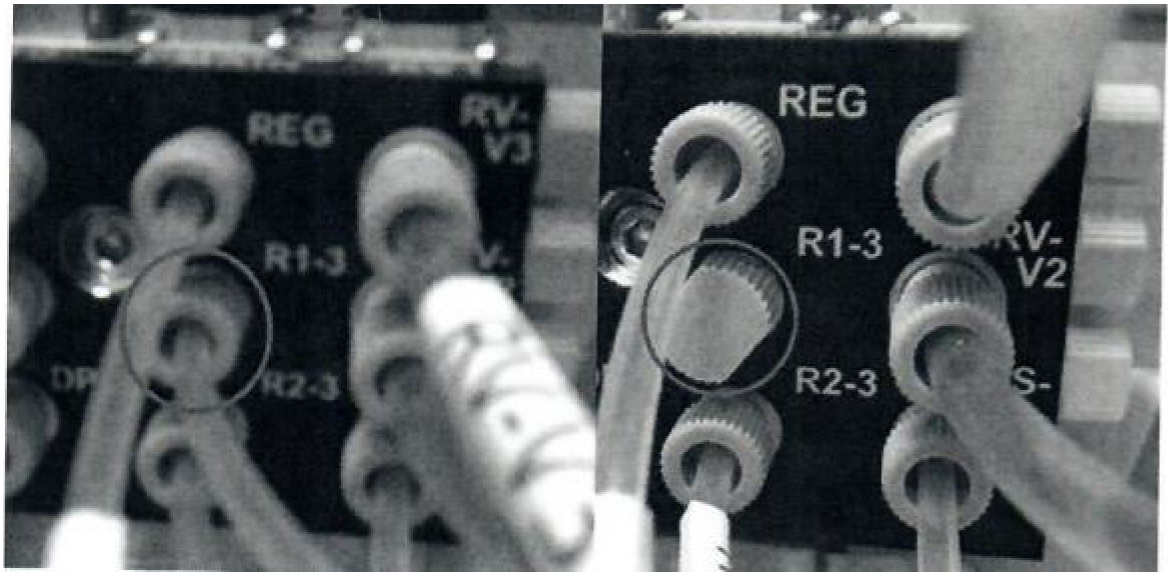

**Figure A3.1:** R1-3 line to replace with plug nut to check for leaks in rotary valve 1.

    ii. Checking Rotary 2: Remove the nut/tubing from the Pmain/R2 port. Screw the plug nut into this port. Run the Pmain leak test. If it is stable, the issue is with the fitted ferrule in this line, and you need to replace it. If not, the fritted ferrule was not the issue, and you should contact CEM product support for next steps. The leak is likely within the rotary valves.

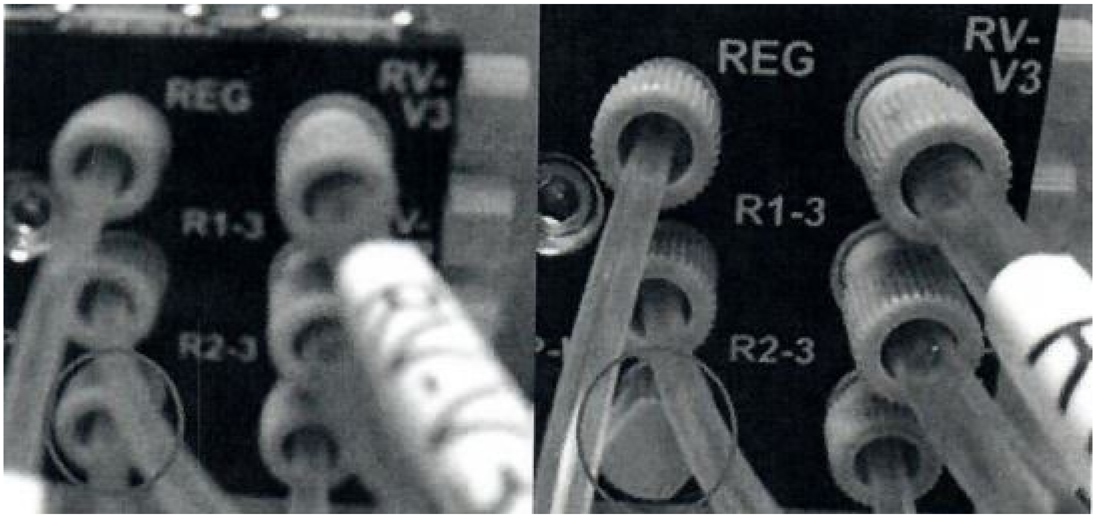

**Figure A3.2:** R2-3 line to replace with plug nut to check for leaks in rotary valve 2.

2. **Reaction vessel filter replacement**: While the Teflon filter in the bottom of the reaction vessel keeps the resin from washing away, it is not perfect by any measure. To ensure resin or other contaminants do not through the lines causing blockages, an additional elongated elliptical reaction vessel is located in the back of the synthesizer behind the nitrogen manifold. This filter should be changed every 1-2 years to prevent system errors and synthesis stoppages. Common symptoms of a clogged reaction vessel filter are when the reaction vessel fails to drain or takes too long to completely drain. To change the reaction vessel filter:
    a. Depressurize the Liberty Blue system (options menu → maintenance → miscellaneous → depressurize)
    b. Turn off Liberty Blue system

c. Rotate system to turn back swinging doors to you (may be easier to pull the system out of the hood onto a cart)
d. Locate the RV filter behind the pressure sensor manifold (PS2/PS3)
e. Unscrew the 4 Phillips head screws attaching the filter to the frame
f. Unscrew RV-V1 and RV-V2 tubing fittings
g. Remove the filter from behind pressure sensors
h. Install new filter in the same orientation as previous filter
i. Screw filter into place and re-attach RV-V1 and RV-V2 lines

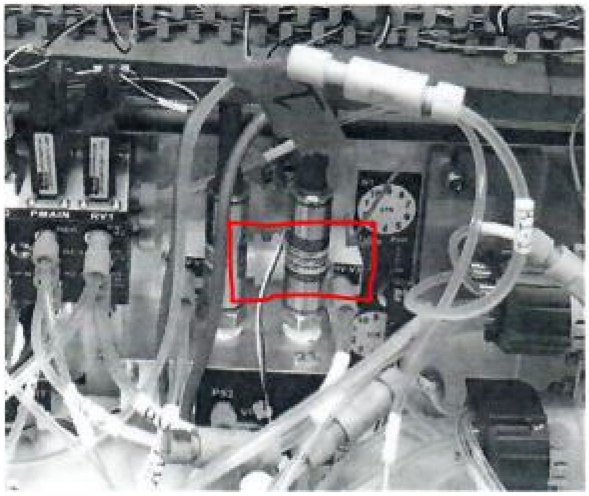

**Figure A3.3:** Location of the reaction vessel filter behind PS2/PS3 nitrogen manifold.