

**Elucidating molecular structure and interactions of  
disease-related noncovalent assemblies through ion  
mobility spectrometry – mass spectrometry**

A Dissertation Presented for the  
Doctor of Philosophy  
Degree  
The University of Tennessee, Knoxville

Amber LeAnn Hope Gray

August 2022

Copyright © by Amber LeAnn Hope Gray, 2022  
All Rights Reserved.

*This is for all the Steve Weibes in the world.*

## Acknowledgements

Throughout the past five years, there are many people who has helped me grow into the scientist and person I am today. The first person in that journey that I would like to extend my gratitude towards is my advisor, Professor Thanh D. Do, for his continued support, guidance, and leadership. I am grateful for all the encouragement and opportunities you have given me to travel to conferences to share my awesome research. I will forever be appreciative of your kindness and patience.

I would also like to thank my committee members, Dr. Bhavya Sharma, Dr. Ziling (Ben) Xue, and Dr. Rebecca Prosser for agreeing to serve on my doctoral committee and their helpful feedback.

I am grateful for all the friendships that I have forged during my time here at UTK. These include the Do group members, past and present. I will never forget our fun conversations, monthly events, and of course, all the margaritas we have drank together. I would also like to give a special thanks to my friends Eric Fussell, Grace Shelton-Sarabia, and Christian Scott, for always being down to grab lunch with me and our fun and insightful conversations that were (sometimes) about chemistry.

Lastly, I would like to thank my husband, Zachary Lawson, for being my biggest fan during this journey and all of the support he has given me.

## Abstract

Ion mobility spectrometry – mass spectrometry (IMS-MS) is a powerful gas-phase technique that is routinely employed in the investigations of amyloid oligomers and conformational studies due to its ability to separate species with the same mass-to-charge ratio ( $m/z$ ). The goal of this dissertation is to use IMS-MS as a primary platform to probe the conformational landscape of a macrocyclic protein, Cyclosporin A (CycA), and to characterize interactions of amyloidogenic assemblies. Through five studies presented within this document, the limits of IMS-MS are pushed by employing conditions which mimic biological environments.

Chapter 2 focuses on how biologically relevant metal cations influence CycA structure.  $\text{Ca}^{2+}$  promoted three conformations of CycA, which were mediated by cis-trans isomerization of the peptide backbone. The chameleonic effect exerted by  $\text{Ca}^{2+}$  helps explain the mechanistic pathways which CycA utilizes within the body to bind to receptors.

Chapter 3 uses phenol soluble modulin peptides to demonstrate the relevance of post-ion mobility dissociation, an event that can occur within IMS-MS instrumentation, in the interpretation of amyloid self-assembly data. A technique using post-IM dissociation was developed to discern differences in the toxicity of the phenol soluble modulin peptides by observing their interactions with detergent micelles.

Chapters 4 and 5 are concerned with the molecular crosstalk of amyloid assemblies with species they can encounter *in vivo*. Studies of amylin and CGRP, two peptides that are co secreted within the pancreas, showed that CGRP remodels amylin fibrillation. Furthermore, amylin fibrillation is inhibited via the formation of intermediary species upon complexation with  $\text{Zn}^{2+}$ . Additionally, investigations of the interactions between A $\beta$  [beta] and Tau showed that A $\beta$  [beta] disrupted the formation of toxic disulfide-bonded Tau dimers.

Lastly, the studies within Chapter 6 elucidate the feasibility of using small peptides, KV11 and its mutant G6W, as models for amyloid oligomers. Tissue viability assays showed that both peptides were able to replicate disease-like states that rival what is seen in neurodegenerative diseases within mouse brain tissue slices. Furthermore, IMS-MS and X-ray crystallography showed that G6W can assemble into a unique dodecameric assembly which can promote the release of disease-related neuropeptides from acute tissue slices.

# Table of Contents

|     |                                                                                                                                                                                                          |    |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 1   | Introduction .....                                                                                                                                                                                       | 1  |
| 1.1 | Motivation and overview.....                                                                                                                                                                             | 1  |
|     | Drift tube ion mobility: Theory .....                                                                                                                                                                    | 3  |
|     | Drift tube ion mobility: Application .....                                                                                                                                                               | 6  |
|     | Electrospray ionization .....                                                                                                                                                                            | 9  |
| 2   | Structural flexibility of <i>Cyclosporine A</i> is mediated by <i>cis-trans</i> isomerization and the chameleonic roles of calcium .....                                                                 | 11 |
| 2.1 | Disclosure.....                                                                                                                                                                                          | 11 |
| 2.2 | Abstract.....                                                                                                                                                                                            | 11 |
| 2.3 | Introduction .....                                                                                                                                                                                       | 12 |
| 2.4 | Materials and methods.....                                                                                                                                                                               | 16 |
|     | Ion-mobility spectrometry – mass spectrometry .....                                                                                                                                                      | 16 |
|     | Distance geometry modeling and theoretical CCS calculations .....                                                                                                                                        | 17 |
|     | NMR spectroscopy .....                                                                                                                                                                                   | 17 |
|     | Cyclophilin A expression .....                                                                                                                                                                           | 19 |
| 2.5 | Results and discussion .....                                                                                                                                                                             | 19 |
|     | IMS-MS reveals a unique compact conformation .....                                                                                                                                                       | 19 |
|     | Endogenous $\text{Ca}^{2+}$ binding to CycA induces <i>bent</i> , <i>closed</i> , and <i>open</i> conformations. NMR shows that CycA: $\text{Ca}^{2+}$ is open with all- <i>trans</i> peptide bonds..... | 27 |
|     | IMS–MS data show that $\text{Ca}^{2+}$ leads to more effective CycA:CypA complexation and structural changes of the protein .....                                                                        | 33 |
| 2.6 | Conclusion.....                                                                                                                                                                                          | 36 |
| 3   | Cytotoxicity of $\alpha$ -Helical, <i>Staphylococcus aureus</i> PSM $\alpha$ 3 investigated by post-ion-mobility dissociation mass spectrometry .....                                                    | 38 |
| 3.1 | Disclaimer.....                                                                                                                                                                                          | 38 |

|     |                                                                                                                                                                                                                                                              |    |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 3.2 | Abstract .....                                                                                                                                                                                                                                               | 38 |
| 3.3 | Introduction .....                                                                                                                                                                                                                                           | 39 |
| 3.4 | Materials and methods .....                                                                                                                                                                                                                                  | 42 |
|     | Ion mobility spectrometry – mass spectrometry .....                                                                                                                                                                                                          | 42 |
|     | Far-UV CD .....                                                                                                                                                                                                                                              | 42 |
|     | Peptide synthesis .....                                                                                                                                                                                                                                      | 43 |
| 3.5 | Results and discussion .....                                                                                                                                                                                                                                 | 44 |
|     | Losses of charged peptides from PSM $\alpha$ 3 oligomers due to post-IM dissociation:<br>problem and solution.....                                                                                                                                           | 44 |
|     | Losses of neutral <i>n</i> -Dodecyl- $\beta$ -D-maltoside (DDM) molecules from Peptide:DDM<br>heteroligomers due to post-IM dissociation: A strong correlation between toxicity<br>and the number of DDM molecules required to stabilize the complexes ..... | 53 |
| 3.6 | Conclusion .....                                                                                                                                                                                                                                             | 61 |
| 4   | $\alpha$ -CGRP disrupts amylin fibrillization and regulates insulin secretion: implications on<br>diabetes and migraine. ....                                                                                                                                | 63 |
| 4.1 | Disclaimer.....                                                                                                                                                                                                                                              | 63 |
| 4.2 | Abstract .....                                                                                                                                                                                                                                               | 63 |
| 4.3 | Introduction .....                                                                                                                                                                                                                                           | 64 |
| 4.4 | Materials and Methods.....                                                                                                                                                                                                                                   | 67 |
|     | Ion mobility spectrometry – mass spectrometry .....                                                                                                                                                                                                          | 67 |
|     | Replica exchange molecular dynamics .....                                                                                                                                                                                                                    | 68 |
|     | Transmission electron microscopy .....                                                                                                                                                                                                                       | 69 |
|     | Mouse islet culturing and harvesting .....                                                                                                                                                                                                                   | 69 |
|     | Liquid chromatography IMS-MS of single mouse islets .....                                                                                                                                                                                                    | 70 |
|     | Cell viability assays.....                                                                                                                                                                                                                                   | 71 |
| 4.5 | Results and Discussion .....                                                                                                                                                                                                                                 | 71 |
|     | IMS-MS reveals a CGRP $\alpha$ -helical intermediate that is stable upon interacting with<br>Zn <sup>2+</sup> .....                                                                                                                                          | 71 |

|                                                                                                                                         |     |
|-----------------------------------------------------------------------------------------------------------------------------------------|-----|
| Zinc complexation promotes $\beta$ -hairpin of amylin but favors the native structure of CGRP .....                                     | 74  |
| Amylin, CGRP, and their mixture do not show significant toxicity to cultured $\beta$ -cells .....                                       | 78  |
| <i>In silico</i> CGRP-amylin simulations support IMS-MS derived conformers .....                                                        | 83  |
| The amylin-CGRP mixture impairs insulin secretion but CGRP oligomers increase the Ins1 level .....                                      | 92  |
| CGRP in migraine and diabetes.....                                                                                                      | 97  |
| 4.6 Conclusion.....                                                                                                                     | 100 |
| 5 The overlooked protective function of non-amyloidogenic $A\beta_{42}$ on Tau pathology                                                | 101 |
| 5.1 Abstract.....                                                                                                                       | 101 |
| 5.2 Introduction .....                                                                                                                  | 102 |
| 5.3 Materials and Methods.....                                                                                                          | 105 |
| Ion mobility spectrometry – mass spectrometry .....                                                                                     | 105 |
| 5.4 Results and Discussion .....                                                                                                        | 105 |
| Tau <sub>PHF43</sub> dimers form via disulfide bonds.....                                                                               | 105 |
| Disulfide bonds are not essential for the formation of higher-ordered Tau <sub>PHF43</sub> oligomers .....                              | 110 |
| $A\beta_{42}$ disrupted Tau <sub>PHF43</sub> disulfide bonded dimers and formed $A\beta_{42}$ :Tau <sub>PHF43</sub> heteroligomers..... | 114 |
| $\beta$ -Sheet disruptor Cyclosporine A can interact with $A\beta_{42}$ and Tau <sub>PHF43</sub> .....                                  | 121 |
| 5.5 Conclusion.....                                                                                                                     | 123 |
| 6 Atomic view of an amyloid dodecamer exhibiting selective cellular toxic vulnerability in acute brain slices.....                      | 125 |
| 6.1 Disclaimer.....                                                                                                                     | 125 |
| 6.2 Abstract.....                                                                                                                       | 125 |
| 6.3 Introduction .....                                                                                                                  | 126 |

|     |                                                                                                                |     |
|-----|----------------------------------------------------------------------------------------------------------------|-----|
| 6.4 | Materials and methods.....                                                                                     | 128 |
|     | Ion mobility spectrometry – mass spectrometry .....                                                            | 128 |
|     | Liquid chromatography IMS-MS .....                                                                             | 128 |
|     | Tissue toxicity assay .....                                                                                    | 130 |
|     | X-ray crystallography .....                                                                                    | 130 |
|     | Liposome dye-release assay .....                                                                               | 130 |
| 6.5 | Results and discussion .....                                                                                   | 131 |
|     | Mouse brain slices exhibit selective regional vulnerability to KV11 and G6W .....                              | 131 |
|     | IMS-MS reveals that G6W can form large, structurally complex oligomers .....                                   | 134 |
|     | G6W assembles into a dodecameric asymmetrical autonomous unit.....                                             | 138 |
|     | G6W dodecamer can be identified with IMS-MS .....                                                              | 149 |
|     | Identifying molecules released from SCN/AH tissues may indicate neurochemical pathways that G6W acts upon..... | 162 |
| 6.6 | Conclusion .....                                                                                               | 165 |
| 7   | Conclusions .....                                                                                              | 167 |
| 7.1 | Summary .....                                                                                                  | 167 |
| 7.2 | Contributions to the field.....                                                                                | 169 |
| 7.3 | Future outlooks.....                                                                                           | 170 |
|     | Bibliography .....                                                                                             | 174 |
|     | Appendix .....                                                                                                 | 200 |
|     | Vita .....                                                                                                     | 213 |

## List of Tables

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |     |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table 2.1. Theoretical CCSs of modeled CycA and structures elucidated from past X-ray crystallography and NMR studies. <sup>a</sup> 22, 26, 46-57 .....                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 22  |
| Table 2.2. Experimental collisional cross sections of CycA and CypA species.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 24  |
| Table 5.1. Theoretical and observed monoisotopic $m/z$ of Tau <sub>PHF43</sub> oligomers. The bold numbers under the theoretical $m/z$ column correspond to the value that is closest to the observed $m/z$ .....                                                                                                                                                                                                                                                                                                                                                                                                          | 109 |
| Table 6.1. Experimental collisional cross section values for G6W oligomers.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 137 |
| Table 6.2. Comparison of CCS of shortest arrival time feature for G6W $n/z = 1/1$ . We cannot rely solely on isotopic spacing of the shortest arrival time feature in the 2D plot of G6W $n/z = 1/1$ due to ambiguity present. However, by calculating the experiment CCS values for various species, we can determine that the shortest arrival time feature must be an oligomer with $n = 6$ or greater. Oligomers smaller than $n = 6$ will have CCS that are smaller than oligomers of the same $n$ with a smaller $Z$ , which is not possible because CCS increases as $Z$ increases due to coulombic repulsions..... | 139 |
| Table 6.3. Neuropeptides upregulated in G6W treatments of the SCN+AH detected in releasates via LC-IMS-MS ( $p < 0.05$ ). .....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 164 |

## List of Figures

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |    |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 1.1. Schematic of Agilent 6560 IMS-QTOF. Ions are generated in the ionization source via electrospray ionization. They are then funneled into the drift cell, where they will separate based on their size, shape, and charge. Ions will then be guided into the collision cell via a hexapole and quadrupole ion guide. Ions are then separated based on their $m/z$ within the time-of-flight mass analyzer. ....                                                                                                  | 4  |
| Figure 1.2. (A) Visualization of how molecules of the same charge state (left panels) and noncovalently bonded assemblies (right panels) are separated within the drift cell. (B) Interpretation of the accompanying ATDs for each type of species.....                                                                                                                                                                                                                                                                     | 8  |
| Figure 1.3. Visualization of ESI mechanism. ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 10 |
| Figure 2.1. (A) Chemical structure of CycA. (B) X-ray crystal structure of the monomeric CycA:CypA complex (PDB 1CWA).....                                                                                                                                                                                                                                                                                                                                                                                                  | 13 |
| Figure 2.2. Representative mass spectrum of CycA in water. The blue highlighted region shows mainly $z = +1$ species, while the red highlight encompasses $z = +2$ species. Note that each mass spectral peak corresponding to a distinct species (e.g. $[\text{CycA} + \text{H}^+]^{+1}$ ) has several signals which arise from the multiple isotopes of the elements that the species is composed of. For proteins and peptides, the isotopic distribution occurs from the carbon-12 and carbon-13 isotopes. ....         | 20 |
| Figure 2.3. ATDs of (A) $[\text{CycA} + \text{H}^+]^{+1}$ and (B) $[\text{CycA} + 2\text{H}^+]^{+2}$ species seen within the mass spectrum of Figure 2.2. Note that the features seen within each ATD corresponds to all the isotopic signals of their respective mass spectral peak. The bottom and top x-axes display CCS and drift time, respectively. The blue highlight encompasses theoretical CCS values of CycA structures modeled here and by previous NMR and X-ray crystallography studies (see Table 2.1) ..... | 23 |
| Figure 2.4. Molecular surfaces and backbone traces of (A) DEKSAN: the X-ray structure of CycA in chloroform (the closed conformation) and (B) the most compact                                                                                                                                                                                                                                                                                                                                                              |    |

conformer of CycA obtained from DG modeling in vacuum. The latter contains three *cis*-amides: MeLeu9-MeLeu10, MeLeu10-MeVal11, and Sar3-MeLeu4..... 26

Figure 2.5. (A) ATDs corresponding to  $m/z = 1222$  in varying concentrations of  $\text{Ca}^{2+}$ . As  $\text{Ca}^{2+}$  concentration increases, so does the feature which corresponds to the dimer CycA in the complex with a  $\text{Ca}^{2+}$ . (B) 2D plot of drift time vs  $m/z$  showing the region  $m/z$  1220–1224 containing the overlapping isotopic distributions. (C) ATDs of  $m/z = 621$  are shown in varying environments. The top spectrum shows the ATD of the species in  $\text{H}_2\text{O}$ , the middle spectrum shows the ATD of the species in  $\text{H}_2\text{O}$  with added  $\text{Ca}^{2+}$ , and the bottom spectrum shows the ATD of the species in  $\text{H}_2\text{O}$  with added  $\text{K}^+$ . (D) 2D plot of drift time vs  $m/z$  from  $m/z$  620.8 to 623.0 in water shows three distinct conformations with the same isotopic spacing. .... 28

Figure 2.6. Proton spectra showing  $\text{H}^{\text{N}}$  (A) and *N*-methyl (B) regions of CycA samples in acetonitrile without (B-1 sample) and with  $\text{CaCl}_2$  (at a 1:1 ratio; sample B-3) at the indicated temperatures. The asterisks annotate common peaks among the samples and conditions..... 31

Figure 2.7. (A) Representative mass spectrum of CypA in  $\text{H}_2\text{O}$  (10  $\mu\text{M}$ ) with the charge state of each species labeled. The inset shows isotopic spacing of CypA with  $z = +15$ . (B) CCS of each ion species with its respective charge state for CypA shown in blue. The pink data points correlate to the theoretical CCS of three CypA structures (PDB IDs 5LUD,<sup>76</sup> 1AWR,<sup>77</sup> and 1AWV<sup>77</sup>). The inset shows the CCS of  $z = +8$  and  $+9$  of CycA:CypA, and the pink lines show the theoretical CCSs of the X-ray crystal structure (PDB 1CWA<sup>26</sup>) and NMR structure (PDB 3CYS<sup>81</sup>) of the complex. The black data points are for the CycA:CypA mixture with no metals added, the green points are the mixture with potassium, and the orange points are the mixture with calcium. (C) 1:1 mixture of CycA:CypA shown by a larger mass spectrum (black) with the major CycA peak labeled and the charge states for each CypA ion labeled. Teal mass spectrum within inset is the CycA:CypA mixture when CycA is in excess.

(D) Mass spectra of the CycA:CypA mixture (5:1 ratio) in the presence of  $\text{Ca}^{2+}$  (blue spectrum) and  $\text{K}^+$  (green spectrum). ..... 34

Figure 3.1. (Top) An image of a PSM $\alpha$ 3 WT dimer generated from the X-ray crystal structure (PDB ID: 5i55). The four mutated residues (F3, K9, F11, and G16) are shown in ball-and-stick form. The residues that contribute to stabilization of the dimer are shown in space-filling models (E2, K6, F10, K12, D14, K17, and N22). (Bottom) Primary sequences of WT, F3A, G16A, and K9P/F11P peptides. .... 40

Figure 3.2. Positive-mode ESI mass spectra of PSM $\alpha$ 3 peptides in 20 mM ammonium acetate. The peptide mass spectral peaks are labeled by their nominal oligomer-to-charge ratio ( $n/z$ ). The oligomers formed by these peptides are similar in order, thus the four mass spectrums are very similar. The peak labeled “tune ion” present in the WT and F3A mass spectra correspond to an ion present in a reference solution that is prepared from the Agilent API-TOF Reference Mass Solution Kit... 45

Figure 3.3. (A) Representative ATDs of F3A oligomers corresponding to the nominal  $n/z = 2/3, 3/4, 4/5$ , and  $5/6$ . (B) A 2D plot of  $m/z$  vs drift time for  $n/z = 2/3$ . (C) Possible assignments for the feature at  $\sim 24$  ms in the ATD of  $n/z = 2/3$ . (D) Deviation of experimental CCSs from isotropic predictions. .... 46

Figure 3.4. Ion-mobility data for WT, G16A, and K9P/F11P at  $n/z = 2/3, 3/4, 4/5$ , and  $5/6$ . A fewer number of features are resolved as oligomer order increases. Only the feature with the longest arrival time (the right most peak) is assigned in our analysis due to uncertainty in the identity of other features. Note that for K9P/F11P this trend deviates slightly in  $n/z = 4/5$ . The secondary structure of this peptide is shown to not have an  $\alpha$ -helix conformation, so oligomers of this species may be intrinsically disordered and therefore may behave differently from other helical peptides in the gas phase. The asterisk for WT  $n/z = 5/6$  denotes a feature that arises from overlapping isotopic spacing from a separate species. .... 47

Figure 3.5. CD time course spectra of 50  $\mu\text{M}$  (A) WT, (B) F3A, and (C) G16A in 20 mM ammonium acetate, pH 6.8, shows that each peptide retains its  $\alpha$ -helicity over a

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |    |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| two week time-period. (B) CD spectra of 50 $\mu\text{M}$ freshly dissolved PSM $\alpha$ 3 peptides in 20 mM ammonium acetate, pH 6.8. These spectra show that WT, F3A, and G16A adopt an $\alpha$ -helical secondary structure in solution, whilst K9P/F11P does not exhibit this characteristic.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 48 |
| Figure 3.6. 2D plots of $n/z = 1/4, 1/3, 1/2$ , and $1/1$ for WT. Note that at the same drift voltage, there is no overlap between the ATDs of $n/z = 1/4$ and $1/3$ with the ATDs of WT oligomers shown in Figure 3.4. On the other hand, the ATDs of nominal $n/z = 1/2$ and $n/z = 1/1$ overlap with the ATDs in Figure 3.4. The same observations can be made for other peptides.....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 51 |
| Figure 3.7. (A, B) Representative ATDs of peptide:DDM complexes corresponding to the nominal $[1 + m]/z$ ratios where “1” denotes peptide monomer, $m$ denotes the number of DDM molecules ( $m = 0-7$ ), and denotes $z$ the charge ( $z = 3$ ). Features are annotated as intact (gray), a loss of one DDM (purple), and a loss of two DDM (green). “M” denotes peptide monomer and “L” denotes DDM. Features present due to overlapping isotropic distributions of multiple species are labeled with an asterick and not discussed. (C) Peak ratios calculated for WT, F3A, G16A, and K9P/F11P. (D) Experimental CCSs and cross-section differences between WT complexes having $m$ and $m + 1$ DDM. ....                                                                                                                                                                                                    | 54 |
| Figure 3.8. ESI-MS in positive mode of PSM $\alpha$ 3 peptides and 200 $\mu\text{M}$ DDM in 20 $\mu\text{M}$ of ammonium acetate. The mass spectral peaks are labeled according to the number of DDM molecules that are complexed with the peptide monomer “intact” species with $z = +3$ . For example, the peaks labeled $[1+2]$ consists of the intact complex of the peptide monomer and 2 DDM molecules at a +3 charge state. Only those peptide-DDM complexes discussed in the main text and previous figures (Figure 3.7 A-B and Figure 3.9) are labeled. Also of note is the presence of the DDM monomer and dimer species ( $[1 \text{ DDM} + \text{Na}^+]$ and $[2 \text{ DDM} + \text{Na}^+]$ ) at $m/z = 533.3$ and $1043.6$ , respectively. Peptide concentrations are as follows: WT 14.6 $\mu\text{M}$ , F3A 15.4 $\mu\text{M}$ , G16A 3.5 $\mu\text{M}$ , and K9P/F11P 4.6 $\mu\text{M}$ . .... | 55 |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Figure 3.9. (A) Representative ATDs of $n/z = 1/3$ for all peptides in the absence of DDM.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |    |
| (B) Representative ATDs of F3A and G16A complexed with DDM. ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 56 |
| Figure 3.10. Cartoon representation of post-IM dissociation events. (A) A loss of a charged peptide from a large oligomer leading to a new feature at short arrival time in the ATD of the small oligomer. (B) A loss of a neutral DDM from a large complex leading to a new feature at long arrival time in the ATD of the small complex. ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 58 |
| Figure 3.11. Experimental CCSs and cross section differences between complexes having $m$ and $m+1$ DDM for (A) F3A, (B) G16A, and (C) K9P/F11P. The green, black, and red bars represent the $\Delta\text{CCS}$ between $m$ and $m+1$ (in $\text{\AA}^2$ ). ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 59 |
| Figure 3.12. 2D plot of $m/z$ vs drift time at $m/z$ 1598 of the WT. Features at $\sim 25.2$ ms and $\sim 21$ ms show a different isotopic spacing than the features of interest, indicating that they are not the same species. ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 60 |
| Figure 4.1. Primary sequences of CGRPs (red), amylin, (blue), and insulin (green). For mouse insulins, the residual difference between Ins1 and Ins2 is indicated in bold. In this work, we study human $\alpha$ -CGRP and human amylin. ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 66 |
| Figure 4.2. (A) Mass spectrum of amylin (10 $\mu\text{M}$ ) in 20 mM ammonium acetate. Each peak is labeled by its nominal $n/z$ ratio. The bottom spectrum exhibits oligomeric species with a relatively low intensity that is not easily identifiable in the top spectrum. The inset shows the isotopic spacing of $n/z = 1/3$ which exhibits the presence of a dimer with $z = +6$ . (B) The ATD of amylin at $m/z = 976$ . (C) Mass spectrum of CGRP (10 $\mu\text{M}$ ) in 20 mM ammonium acetate. The insrt shows the isotopic spacing of $n/z = 1/3$ which indicates that this mass spectral peak correlates to a triply charged dimer. (D) The ATD of CGRP at $m/z = 948$ and a REMD conformation of CGRP monomer with the segment 8–18 colored in red. This conformation is consistent with previous NMR and CD studies of CGRP in structure-inducing solvent. .... | 72 |

- Figure 4.3. (A and B) Representative ESI-mass spectra (in 50 : 50 MeOH : H<sub>2</sub>O) of (A) 10  $\mu$ M amylin and (B) 10  $\mu$ M CGRP, both in the presence of 0.5 mM Zn<sup>2+</sup>. (C) Representative ATDs of amylin, amylin:Zn, and amylin:Zn(H<sub>2</sub>O)<sub>5</sub>. (D) Representative ATDs of CGRP, CGRP : Zn, and CGRP : Zn(H<sub>2</sub>O)<sub>5</sub>. ..... 75
- Figure 4.4. Representative TEM images of (A and D) amylin, (B and E) CGRP, and (C and F) the mixture. The scale bar is 150 nm. .... 77
- Figure 4.5. Partial mass spectra of amylin:CGRP hetero- dimer and trimer and corresponding ATDs. For panels A – C, the top mass spectrum is of amylin alone and the bottom mass spectrum is of the amylin:CGRP mixture. Within the partial mass spectra, the mass spectral peaks boxed in blue correlate to amylin, red is CGRP, and green represents co-oligomeric species. (A - B) amylin:CGRP (1:1) heterodimers (C) amylin:CGRP (2:1) heterotrimer (D) ATDs of amylin dimer and trimer (left) and amylin:CGRP hetero- dimer and trimer (right). .... 79
- Figure 4.6. The mass spectrum of 1:1 amylin:CGRP (5:5  $\mu$ M) in 20 mM ammonium acetate at  $t = 0$ . Mass spectral peaks labeled “A+C” represents oligomers composed of 1:1 amylin:CGRP, “2A+C” represents oligomers made up of 2:1 amylin:CGRP, and “A+2C” represents oligomers of 1:2 amylin:CGRP. Note the asterisk next to the 1:1 amylin:CGRP species at  $z = +7$  and  $+8$  in the 2100 – 2700 mass range. These are species that are composed of 2:2 amylin:CGRP. .... 80
- Figure 4.7. Representative TEM images of the amylin:CGRP mixture with ZnCl<sub>2</sub> (A – C) and without ZnCl<sub>2</sub> (D – F). The samples with Zn had a mixture of material, including (A) long fibrils and (B-C) short fibrils. The short fibrils look similar in morphology to the globular aggregates (e.g. panel F). The samples without ZnCl<sub>2</sub> primarily exhibited globular aggregates, however there were some faint and indistinct fibrillar species present (D-E). The scale bar is a 150 nm. .... 81
- Figure 4.8. The mass spectrum of 1:1 amylin:CGRP (5:5  $\mu$ M) in 20 mM ammonium acetate after setting out at room temperature for one week. Only peaks corresponding to CGRP (blue) and amylin (orange) are labeled. Note that the

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |    |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| intensity of the amylin peaks are much lower than what is shown in Figure 4.6. This implies that amylin has fell out of solution to form globular aggregates. ....                                                                                                                                                                                                                                                                                                                                                                                                            | 82 |
| Figure 4.9. Cell viability assay using cultured $\beta$ -cells exposed to CGRP, amylin, and amylin : CGRP. 832/13 rat $\beta$ -cells were either left untreated (NT) or exposed to 1 ng mL <sup>-1</sup> IL-1 $\beta$ and 2 mM camptothecin (CT) compared with 5 $\mu$ M of amylin, CGRP, and amylin plus CGRP for 18 h. ****, $p < 0.0001$ versus NT using one-way ANOVA. $n = 4-6$ per group. When compared to CT, the effects of amylin, CGRP, or amylin : CGRP on cellular viability are minor, although amylin has a stronger effect than the other two conditions. .... | 84 |
| Figure 4.10. Cell viability assay using cultured $\beta$ -cells exposed to CGRP, amylin, and amylin:CGRP. 832/13 rat b-cells were either left untreated (NT) or exposed to 1 ng/mL IL-1b, 2 mM camptothecin (CT) compared with 20 and 200 nM of amylin, CGRP, or amylin plus CGRP for 18 h. ****, $p < 0.0001$ versus NT using one-way ANOVA. $n = 4-6$ per group. When compared to CT, there effects of amylin, CGRP, or amylin+CGRP on cellular viability on cell viability are minor, although amylin has a stronger effect than other two conditions. ....                | 85 |
| Figure 4.11. Representative structures of CGRP monomers obtained from REMD simulations and clustered together by residues 8-18. CCS values for these eigenstates are in excellent agreement with IM-MS data. ....                                                                                                                                                                                                                                                                                                                                                             | 87 |
| Figure 4.12. CGRP and amylin CCS from REMD simulations reveal small disordered CGRP-containing homodimers and heterodimer compared to larger $\beta$ -stranded amylin dimers. ....                                                                                                                                                                                                                                                                                                                                                                                            | 88 |
| Figure 4.13. Theoretical CCS values as a function of corresponding protein secondary structures from REMD, which highlight significantly disordered conformations. ...                                                                                                                                                                                                                                                                                                                                                                                                        | 89 |
| Figure 4.14. Representative structures of CGRP monomers ( <i>top</i> ), homodimers ( <i>middle</i> ), and heterodimers ( <i>bottom</i> ) obtained from REMD simulations. ....                                                                                                                                                                                                                                                                                                                                                                                                 | 90 |
| Figure 4.15. Molecular contact maps from REMD simulations reveal a range of disordered and hairpin conformations. ....                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 91 |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |     |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 4.16. Schematic of single islet analysis workflow. ....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 93  |
| Figure 4.17. (A) Extracted LC chromatogram showing Ins1 and Ins2. (B) Extracted partial LC-IMS-MS mass spectrum showing only Ins1. (C and D) 2D plots of arrival time vs. $m/z$ of Ins1 and Ins2 at $z = +6$ . ....                                                                                                                                                                                                                                                                                                                                                                                                     | 94  |
| Figure 4.18. Comparisons of insulin signals obtained from single islets exposed to (A and C) media contain mostly amylin/CGRP monomers and (B and D) media contain mostly amylin/CGRP oligomers. *** indicates $p < 0.05$ . ....                                                                                                                                                                                                                                                                                                                                                                                        | 96  |
| Figure 4.19. Comparisons of amylin and CGRP oligomers in “fresh” and “aged” incubating solutions. Amylin oligomers up to tetramers were detected, whereas CGRP oligomers up to dimers were detected in the two different solutions. The ion abundance of the oligomers increased in the aged peptides solutions. A paired t-test was performed to check for variation between the oligomers in the different solutions. *** indicates significant difference. ....                                                                                                                                                      | 98  |
| Figure 5.1. Schematic of the six different human Tau isoforms. Note that full length Tau contains two cysteines. The sequence of Tau <sub>PHF43</sub> is given, which encompasses part of the repeat 1 domain (R1), all of R3, and part of R4. ....                                                                                                                                                                                                                                                                                                                                                                     | 104 |
| Figure 5.2. (A) Representative positive mode-ESI mass spectrum of Tau <sub>PHF43</sub> freshly prepared ( $t = 0$ ). The inset reveals isobaric species that can be hidden underneath the same mass spectral peak. Peaks labeled blue denote isobaric species. (B) Mass spectrum of aged Tau <sub>PHF43</sub> ( $t = 24$ hours). (C) Mass spectral peaks of aged Tau <sub>PHF43</sub> + tris(2-carboxyethyl)phosphine (TCEP) (top panels) and aged Tau <sub>PHF43</sub> (bottom panels). The black arrows are pointing towards the monoisotopic peak which was used to determine the observed monoisotopic $m/z$ . .... | 106 |
| Figure 5.3. Schematic of possible disulfide bond formation present within Tau <sub>PHF43</sub> dimers, trimers, and tetramers. Noncovalent bonds (e.g. hydrogen bonding) are represented by a dashed line. Note that the position of the noncovalent bond is trivial. The bottom panel shows speculative disulfide bonding of large 4R Tau oligomers. ....                                                                                                                                                                                                                                                              | 111 |

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |     |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 5.4. (A) Mass spectral peaks of higher-ordered Tau <sub>PHF43</sub> oligomers. The monoisotopic peak is labeled. (B) 2D ion mobility plots of Tau <sub>PHF43</sub> $n/z = 1/4$ by itself (top panel) and in the presence of TCEP (bottom panel). .....                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 112 |
| Figure 5.5. 2D ion mobility plot at $m/z = 2245$ of Tau <sub>PHF43</sub> + TCEP (positive mode-ESI). Note that an unresolved dimer/trimer can be seen within the 2D ion mobility portion. Observation of isotopic spacing within the mass spectral peak confirms the unresolved 2D ion mobility plot does contain a dimer. ....                                                                                                                                                                                                                                                                                                                                                                                                     | 115 |
| Figure 5.6. (A) Representative positive mode-ESI mass spectrum of 10 $\mu$ M:10 $\mu$ M A $\beta$ <sub>42</sub> :Tau <sub>PHF43</sub> . Only Tau <sub>PHF43</sub> peaks are labeled. The peak in blue denotes higher-ordered species underneath the mass spectral peak. (B) Examples of isobaric mass spectral peaks of Tau <sub>PHF43</sub> in the presence of A $\beta$ . Note that there is no evidence of Tau <sub>PHF43</sub> dimers as in Figure 5.2C, only monomers. (C) 2D ion mobility plots (negative mode-ESI) of Tau <sub>PHF43</sub> by itself (top panel) and in the presence of A $\beta$ <sub>42</sub> (bottom panel). .....                                                                                        | 117 |
| Figure 5.7. Tau <sub>PHF43</sub> oligomers (top panels) and A $\beta$ <sub>42</sub> :Tau <sub>PHF43</sub> heteroligomers (bottom panels).....                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 118 |
| Figure 5.8. Mass spectrum of Tau <sub>PHF43</sub> + heparin that was left to incubate at room temperature over 48 hours (pink). The mass spectrum of Tau <sub>PHF43</sub> + A $\beta$ <sub>42</sub> + heparin, incubated for 48 hours, is shown in black as comparison. A $\beta$ <sub>42</sub> was added prior to mass spectral analysis. Note that the abundance of oligomers of Tau <sub>PHF43</sub> + heparin is much lower than that of the mixture containing the three components. This indicates that TAu <sub>PHF43</sub> has fibrillated, forming insoluble species not detected by the mass spectrometer. The soluble species of Tau are composed primarily of monomers, which are labeled (example shown by inset)..... | 120 |
| Figure 5.9. (A) ATDs of A $\beta$ <sub>42</sub> :CycA complex and of A $\beta$ <sub>42</sub> . (B) Mass spectral peaks with isotopic spacing of A $\beta$ <sub>42</sub> :CycA (2:1) complex, Tau <sub>PHF43</sub> :CycA complex (2:2), and Tau <sub>PHF43</sub> :CycA complex (1:1) in presence of A $\beta$ <sub>42</sub> .....                                                                                                                                                                                                                                                                                                                                                                                                    | 122 |

Figure 6.1. Tissue viability assessment under different experimental conditions.

Histogram bars indicate the percentage of DAPI-stained cells with different amounts of PI staining. High-intensity PI staining indicates dead cells, moderate PI staining indicates injured cells, and low to no PI staining indicates healthy cells.

(\* $p < .05$ ) The error bars represent the standard error mean..... 132

Figure 6.2. Representative mass spectra of (A) G6W and (B) KV11. Each mass spectral

peak is labeled by its nominal oligomer-to-charge ratio ( $n/z$ ). Note that the intensity in the colored regions of each mass spectra is adjusted to reveal less abundant oligomeric species. The 2D plots for  $n/z = 1/1$  are shown for both (C) G6W and (D) KV11, where the isotopic spacings can clearly be seen for the features within the ATDs. For G6W, notably, species mentioned in the text are marked with arrows. The data were collected using helium buffer gas..... 135

Figure 6.3. The 2D plots of (A) KV11 and (B) G6W are shown at  $n/z = 3/2$ . The ATDs show

two major features. The isotopic spacing for the faster feature clearly shows that  $Z = +4$ , which means that the oligomer represented by this feature must be a hexamer (i.e.  $n/z = 6/4 = 3/2$ ). ..... 136

Figure 6.4. The mass spectrum and 2D plot of G6W-Br5 at  $n/z = 1/1$ . The 2D plot of

G6W-Br5 showed that similar oligomers are still observed within this plot as the native G6W molecule..... 140

Figure 6.5. The G6W dodecamer crystal structure viewed (A, B) parallel and (C, D)

perpendicular to the pseudo  $6_4$  symmetry axis. Two highly twisted, antiparallel sheets (brown and gray) meet side-by-side at an interface composed of hydrophobic side chains including Trp6 residues, highlighted by van der Waals spheres in panels A and C. Backbone hydrogen bonds are highlighted by yellow lines. Shape complementary between sheets is poor, leaving voids filled by trifluoroacetic acid (TFA) and glycerol (GOL) highlighted by white circles. (B) A space-filling model reveals a charged outside surface (pink and blue spheres) and

(D) a cutaway view reveals a hydrophobic core (yellow spheres) with pockets for TFA and GOL binding..... 141

Figure 6.6. (A) A column of G6W dodecamers in the crystal. The box indicates the boundary of one dodecamer. Backbone hydrogen bonds are highlighted in yellow. Brown and gray colors distinguish the two  $\beta$ -sheets in the dodecamer. The cylinder represents the pseudo  $6_4$  symmetry axis. (B) We unrolled the  $\beta$ -sheets into a flat plane in order to highlight the location of the breaks in hydrogen bonding (red zeros). The count of hydrogen bonds between each pair of strands is indicated. (C) Orthogonal views of the dodecamer. (D) A projection down the symmetry axis highlights the deviation from ideal  $6_4$  symmetry in which the angles between antiparallel pairs of strands would be exactly  $60^\circ$ , and hydrogen bond counts (yellow numbers) would be evenly distributed around the axis. (E) Close-up view of the absence of hydrogen bonds in the break between dodecamers..... 143

Figure 6.7. Molecular modeling shows palmitic acid binding to hydrophobic cavity of G6W. .... 144

Figure 6.8. (A) A column of G6W-Br5 dodecamers in the crystal. The box indicates the boundary of one dodecamer. Backbone hydrogen bonds are highlighted in yellow. Brown and gray colors distinguish the two  $\beta$ -sheets in the dodecamer. The cylinder represents the perfect  $6_4$  symmetry axis. (B) We unrolled the  $\beta$ -sheets into a flat plane in order to highlight the regular, repeating hydrogen bonding pattern. The count of hydrogen bonds between each pair of strands is indicated. (C) Orthogonal views of the dodecamer. (D) A projection down the symmetry axis highlights ideal  $6_4$  symmetry in which the angles between antiparallel pairs of strands are exactly  $60^\circ$ , and hydrogen bond counts (yellow numbers) are evenly distributed around the axis. .... 146

Figure 6.9. (A) G6W molecules hydrogen-bond together as antiparallel  $\beta$ -sheet dimers. We show all 6 dimers of the dodecamer. Each dimer exhibits 10-12 backbone hydrogen bonds. (B) The dimer backbones superimpose well, but there is variability

in side chain conformation. (C) Dimers associate at varied angles and with varied numbers of hydrogen bonds. The unfilled sheet outline provides a reference. (D) the variation in angles between dimers is obvious in the superposition..... 147

Figure 6.10. Oligomeric assemblies of four homologous peptides (left to right: G6W, KV11, superoxide dismutase (SOD) residues 28-38, and SOD:28-38 with G37R hereditary mutation) are (A) viewed down the primary symmetry axis (black dot) and (B) perpendicular to the axis. Sequence homology among the four segments is illustrated on the top line. All feature (C) a dimeric building block: an antiparallel pair of  $\beta$ -strands joined by hydrogen bonds (yellow dashed line). (D) This dimeric unit exhibits structural variations in the backbone, and these contribute to the gross variations in the larger scale oligomeric assemblies. Compared to the G6W dimer, root mean square deviations (RMSDs) in 44 pairs of backbone atoms are 1.1, 1.6 Å with the closest related sequences G6W-Br5 and KV11. RMSDs are 1.5 and 1.6 Å for the more distantly related wild type and G37R SOD segments, and have fewer atoms aligned. (E) Deviations in the backbone are revealed in an orthogonal view of the aligned monomers. Residue 6 is highlighted by white colored bonds in most of the panels. .... 148

Figure 6.11. G6W liposome dye release. NR release results for liposomes composed of 100% PC, 10/90 PA/PC, and 10/90 DOTAP/PC after adding 4.5 mM G6W peptide. No significant decrease in NR fluorescence was observed over 24 hours, indicating G6W peptide did not lead to release of encapsulated dye..... 150

Figure 6.12. Dynamic light scattering of liposomes before and after G6W treatment. Raw DLS distribution curves for liposomes before and after peptide treatment. In all cases, stable liposomes with desired sizes were detected before peptide addition. After incubation with peptide for over 24 hrs, no change was observed in any of the samples tested. This further supported that G6W did not alter liposome self-assembly properties based on particle sizes. .... 151

Figure 6.13. Mass spectrum of 15  $\mu\text{M}$  G6W and 200  $\mu\text{M}$  DDM in 20 mM ammonium acetate. Each mass spectral peak is labeled by the number of DDM molecules attached to the respective G6W oligomer with the charge state of the complex. For example, the teal “10<sup>+3</sup>” represents a G6W monomer in complex with 10 DDM molecules with  $z = +3$ . Multiple oligomers of G6W (i.e., dimer, trimer, and tetramer) are observed to be complexed with DDM molecules. Note that all G6W-DDM species are not labeled within this mass spectrum..... 152

Figure 6.14. The experimental  $\text{CCS}_{\text{N}_2}$  of G6W monomer-, dimer-, trimer-, and tetramer – DDM complexes are shown in each respective plot by symbols. The line-of-best fit for experimental CCS is shown in dark blue for each complex. The isotropic CCS is shown by the solid line above the experimental CCS values. The isotropic values were calculated using  $n/z = 1/2, 2/1, 3/2$ , and  $4/2$  for the monomer, dimer, trimer, and tetramer species, respectively (the lowest charge-state species). The bars show the differences in CCS for G6W-DDM complexes with  $m$  and  $m+1$  number of DDM molecules. .... 153

Figure 6.15. Representative ATDs of the monomer-, dimer-, trimer-, and tetramer-DDM complexes illustrates post-IM dissociation (or lack of post-IM dissociation in case of the tetramer) of neutral DDM. Each ATD is labeled in the format of *[G6W oligomer order + # intact DDM molecules]/charge state*. For example, the ATD labeled *[1+3]/2* means that that ATD is extracted from the mass spectral peak representing a monomer with 3 DDM molecules with  $z = +2$ . (M = monomer, D = dimer, T = trimer, Tet = tetramer, L = DDM molecule). .... 155

Figure 6.16. Plots showing the experimental CCS of intact and dissociated (A) monomer- (B) dimer- and (C) trimer-DDM complexes are shown. .... 157

Figure 6.17. (A) Representative ATDs of G6W  $n/z = 1/1$  ( $m/z = 1,327$ ) in the presence of DDM (top) and absence of DDM (bottom) in nitrogen buffer gas. The shortest arrival time feature (blue arrow) is tentatively assigned as an octamer, decamer, or dodecamer. (M = monomer, D = dimer, T = trimer, Te = tetramer, Hex = hexamer).

(B) Possible oligomeric assignments of the shortest arrival time feature (blue arrow) in both Figure 6.2c and 6.17a and comparisons of  $\sigma_{\text{exp,He}}$  and  $\sigma_{\text{theory,He}}$  of the “double-helix” X-ray structures. Note that these CCS experimental values were obtained in helium in the absence of DDM. (C) 2D plot of G6W in the absence of DDM at  $m/z = 2,654$ . Each feature within the ATD is assigned according to isotopic spacing detailed by the 2D plot. Note that data shown is collected using helium as the drift gas. .... 158

Figure 6.18. Structural Models Derived from the X-ray Crystal Structure and the Corresponding Theoretical Cross Sections of G6W Oligomers. (A) The single-layer model, (B) The double-layer model, (C) Possible assignments to the largest oligomers observed at  $m/z$  1327 (nominal  $n/z = 1/1$ ) and the CCS values in helium. .... 161

Figure 6.19. Cross-comparison of putative neuropeptide mass spectra between the G6W-affected SCN media releasate, control SCN media releasate, and the G6W standard. .... 163

# 1 Introduction

## 1.1 Motivation and overview

The sampling of dynamic biomolecules is of utmost importance to help explain biological and disease-related mechanistic pathways that occur within living organisms. However, these systems are often difficult to investigate due to their heterogeneous nature. To obtain a more complete picture of the pathways utilized by these systems, one must use a technique that can isolate and capture the transient and often short-lived species involved. Ion mobility spectrometry - mass spectrometry (IMS-MS) is a technique that is routinely employed for the investigation of these transient systems.

IMS-MS combines two gas-phase techniques, ion mobility and mass spectrometry, which allows for the orthogonal separation of isobaric and isomeric species with the same mass-to-charge ( $m/z$ ) ratio. This allows investigations of inherently heterogeneous systems, such as small amyloid assemblies that can populate many different oligomeric states or macrocyclic peptides that occupy a wide range of many different conformations.

As an emerging technique, IMS-MS has been implemented as a primary tool for the structural investigation of proteins and biomolecular self-assembly studies. However, many of these investigations only utilize buffer conditions, failing to mimic the biological conditions that these species are found in. Moreover, IMS-MS data has primarily only been coupled with computational techniques, such as molecular dynamics. Thus, there are two motives that fuel the stories told herein this dissertation. (i) I want to push and improve the limits of IMS-MS technology. IMS-MS characterizations of Cyclosporin (specifically, Cyclosporin A or CycA) and small amyloid oligomers will be performed under biologically relevant conditions. By investigating the interactions of these species with species they are likely to come into contact with *in vivo*, such as salt cations or other biological factors. (ii) Data from condensed-phase techniques, such as X-ray

crystallography and NMR spectroscopy, will be used to leverage what is seen with IMS-MS. This will highlight the strengths and weaknesses of each technique.

The five studies discussed within this document utilized IMS-MS as a primary platform to investigate, revise, and answer questions across topics that are of broad biological and medical significance. We will begin in Chapter 2, where the structural landscape of macrocyclic peptide CycA, which can twist and bend via cis/trans isomerization of the peptide backbone, is investigated. CycA can theoretically encompass many different conformations, but despite being studied for the past several decades, only a few conformers have been found with condensed-phase techniques. Therefore, this makes CycA a model of interest to study with IMS-MS. How ionic species perturbs this system, such as biological relevant salts, will be investigated. This investigation is leveraged by the results obtained from NMR spectroscopy, illustrating the capability of IMS-MS as a technique that can preserve solution phase information in the gas-phase.

The remaining chapters will be concerned with using IMS-MS to investigate amyloid assemblies. Chapter 3 highlights a fundamental event that occurs within IMS-MS instruments called post-ion mobility (post-IM) dissociation. Despite the potential for post-IM dissociation, it is often not considered in the light of IMS-MS biomolecular self-assembly data. This investigation aims to propose a unique way to interpret data affected by this phenomenon by investigating how phenol soluble modulin peptides interact with a micelle-forming detergent.

Although the aggregation of proteins is notorious for its association with human disease, therapeutic efforts that target these aggregates are often unsuccessful. However, how protein aggregates are influenced by species they are likely to come across *in vivo* may be able to uncover mechanisms which can shed light on disease pathology and explain the correlation of seemingly unrelated diseases. Within Chapters 4 and 5, the molecular crosstalk of amyloid assemblies is investigated. Chapter 4 investigates amylin and the neuropeptide CGRP, peptides involved in Type II diabetes

and migraine, respectively. Through understanding the effects of the interactions between the two peptides, the inverse relationship of Type II Diabetes and migraine coincidence can be explained.

On the other hand, Chapter 5 is concerned with A $\beta$  and Tau, two proteins that are related through their occurrence in Alzheimer's Disease. The effects of A $\beta$  plaques on the formation of Tau tangles have been well-studied, but there are surprisingly few investigations on how small oligomers of both species interact with each other. It is shown how small A $\beta$  oligomers affect Tau dimers and fibrils and explains how their interactions may be significant to mechanisms underlying Alzheimer's Disease.

The fifth and final study discussed in Chapter 6 involves the investigation of a small peptide, termed G6W, which is derived from a well-known model of amyloid oligomers, called KV11. Through IMS-MS and X-ray crystallography, properties of G6W are uncovered, including the interactions it has with acute mouse brain tissue slices. Although both KV11 and G6W do not underlie any amyloid disease, the feasibility of using these models to induce effects within tissue slices that rival what is seen within actual neurodegenerative disease is investigated.

### **Drift tube ion mobility: Theory**

The premise of IMS-MS employs the combination of two gas-phase techniques, ion mobility spectrometry and mass spectrometry, from which two important quantities of charged gaseous species can be obtained: ion mobility and mass-to-charge ratio ( $m/z$ ). All IMS-MS data presented within this thesis was obtained from a 6560 Agilent ion mobility – quadrupole time-of-flight (IMS-QTOF). A schematic of the instrument can be seen in Figure 1.1. Drift cell ion mobility spectrometry is used to directly determine a species' ion mobility, a quantity which describes how fast an ion traverses through a high-pressure (approximately 4 torr) drift cell under the influence of a weak electric field (400 – 1700 V). Gaseous ions are first generated by an ionization source, such as electrospray ionization (ESI), collected into an entrance funnel, then pulsed into the drift

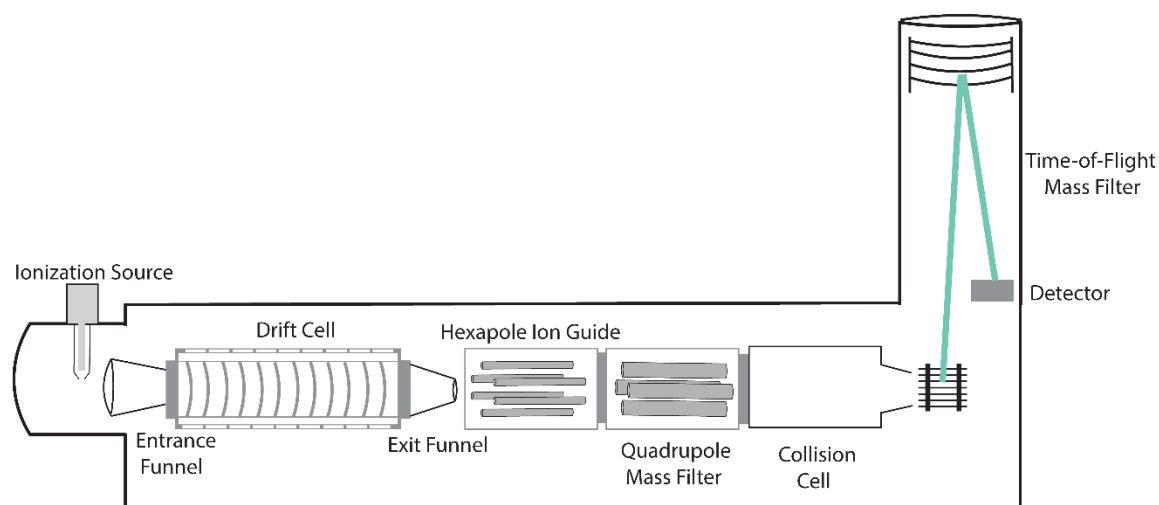


**Figure 1.1. Schematic of Agilent 6560 IMS-QTOF.** Ions are generated in the ionization source via electrospray ionization. They are then funneled into the drift cell, where they will separate based on their size, shape, and charge. Ions will then be guided into the collision cell via a hexapole and quadrupole ion guide. Ions are then separated based on their  $m/z$  within the time-of-flight mass analyzer.

cell. A voltage is applied to drift cell entrance and exit, and the ions are pulled forward through the drift cell by the electric field generated by the voltage difference. As the ions are traversing the drift cell, which is filled with an inert buffer gas, they will collide and interact with buffer gas molecules. The force on the ions originating from the electric field is negated by the frictional force from the collisions and interactions of the ions with the buffer gas. The offset of the two forces halts the acceleration of the ions and allows the ion to travel through the drift cell with a constant velocity proportional to the electric field, as shown by Equation 1.1.

$$v = KE \quad 1.1$$

Where  $v$  is the velocity,  $K$  is ion mobility, and  $E$  is the electric field. The ion mobility,  $K$ , is commonly reported as the reduced ion mobility,  $K_0$ , by correcting to standard temperature and pressure (Equation 1.2).

$$K_0 = K \frac{P}{P_0} \frac{T_0}{T} \quad 1.2$$

Where  $P$  is the pressure,  $T$  is the temperature, and  $P_0$  and  $T_0$  are the standard pressure and temperature (760 torr and 273.15 K), respectively.

The outcome of an ion mobility spectrometry experiment results in an arrival time distribution (ATD) spectrum, which details the arrival time of a species, or the time it takes the species to travel the drift cell and hit the detector. The arrival time is not to be confused with drift time. Drift time is the time it takes an ion to traverse the drift cell only, while the arrival time includes the time the ions spend outside the drift. The arrival time of a species ( $t_A$ ) at varying  $\frac{P}{V}$  quantities, where  $P$  is the pressure and  $V$  is the voltage of the drift cell, can be used to generate a linear plot detailing  $t_A$  vs  $\frac{P}{V}$ . The arrival time,  $t_A$ , is directly related to  $K_0$  as shown in Equation 1.3.

$$t_A = \frac{L^2 T_0}{P_0 T} \frac{1}{K_0} \frac{P}{V} + t_0 \quad 1.3$$

Where  $L$  is the length of the drift cell and  $t_0$  is the time the ions spend outside the drift cell. Thus, the drift time,  $t_d$ , can be obtained by Equation 1.4.

$$t_d = t_A - t_0 \quad 1.4$$

Under low electric-field conditions, the ion mobility is not a function of the electric field strength. As a result,  $K_0$  of an ion can be directly related to the average momentum transfer collision integral ( $\Omega$ ) seen in Equation 1.5.<sup>1</sup>

$$K_0 = \frac{3Ze}{16N_0} \sqrt{\frac{2\pi}{\mu kT}} \frac{1}{\Omega} \quad 1.5$$

Where  $Z$  is the charge of the ion,  $e$  is the elementary charge,  $N_0$  is the reduced gas number density,  $\mu$  is the reduced mass of the ion-buffer gas pair,  $k$  is the Boltzmann constant, and  $T$  is the temperature within the drift cell. The average momentum transfer collision integral,  $\Omega$ , details the momentum transfer between the ion-buffer gas pair.  $\Omega$  is approximate to the collisional cross section (CCS),  $\sigma$ , for a rigid sphere. The CCS is a value that can give information regarding an ion's size and shape. If the experimental CCS value is coupled to theoretical values obtained from computational modeling, then structural information of the species can be determined.

### **Drift tube ion mobility: Application**

Ion mobility spectrometry separates species based on three parameters: size, shape, and charge. When discussing the separation of ions, there are two general scenarios to consider. The first scenario is if our analytes were species of the same mass, but different sizes or shapes (e.g., isomers of a molecule with differing conformations). If a set of isomers have different sizes/shapes, then the more compact, smaller isomer would travel the drift cell faster than the larger, more extended isomer. This is because larger species will have more collisions with the buffer gas inside the drift cell, effectively slowing it down.

Alternatively, the second scenario considers non-covalently bounded identical species (e.g., amyloid oligomers). In this scenario, oligomers with a large  $n$  (where  $n$  is the number of units that the oligomer is composed of) will travel faster than oligomers of a smaller  $n$ . Large oligomers will have a higher charge state and thus will travel faster

than small oligomers with a lower charge state. According to the Mason Schamp Equation (Equation 1.5), the reduced mobility of an ion is directly proportion to its charge state,  $Z$ , but not to the mass of the molecule, which is part of the reduced mass,  $\mu$ . Therefore, if you double the charge of an ion, you double its mobility. On the other hand, if you were to double the mass of the ion, that does not double the ion's mobility. In other words, the effect of charge outweighs the effect of size. Figure 1.2 illustrates how to interpret ATD data based on these two scenarios.

It is important to note that the type of drift gas within the drift cell can affect the features seen within the ATD. For example, the CCS of species within an ATD will vary depending on the drift gas used. If a species is analyzed with helium (He) as the drift gas, then it would exhibit a smaller CCS than if that same species were to traverse a drift cell filled with nitrogen gas ( $N_2$ ). Along the same line, helium is typically the drift gas of choice when structural information of a species is to be obtained. This is because helium is the smallest drift gas one can use (hydrogen is not an option as a drift gas due to its flammability). Since helium is so small, it can “probe” the surface of an analyte more accurately than larger drift gases (e.g.,  $N_2$ ) and thus provide a more accurate CCS when compared to theoretical CCS from computational modeling. Although He may have this advantage, it is expensive to use in longer ion mobility experiments. On the other hand,  $N_2$  gas is cheap and abundant. Nitrogen gas also has the capability distinguish between protomers of benzocaine, or species with the same conformation but differ in the location of the proton (relevant to positive-mode ESI). This is explained due to the larger polarization of  $N_2$  compared to He, which will have a stronger interaction with the more polar tautomer.<sup>2</sup> Although the experiments in this thesis use either He or  $N_2$  as the drift gas, other types of gases such as argon or carbon dioxide can be used in ion mobility as well.

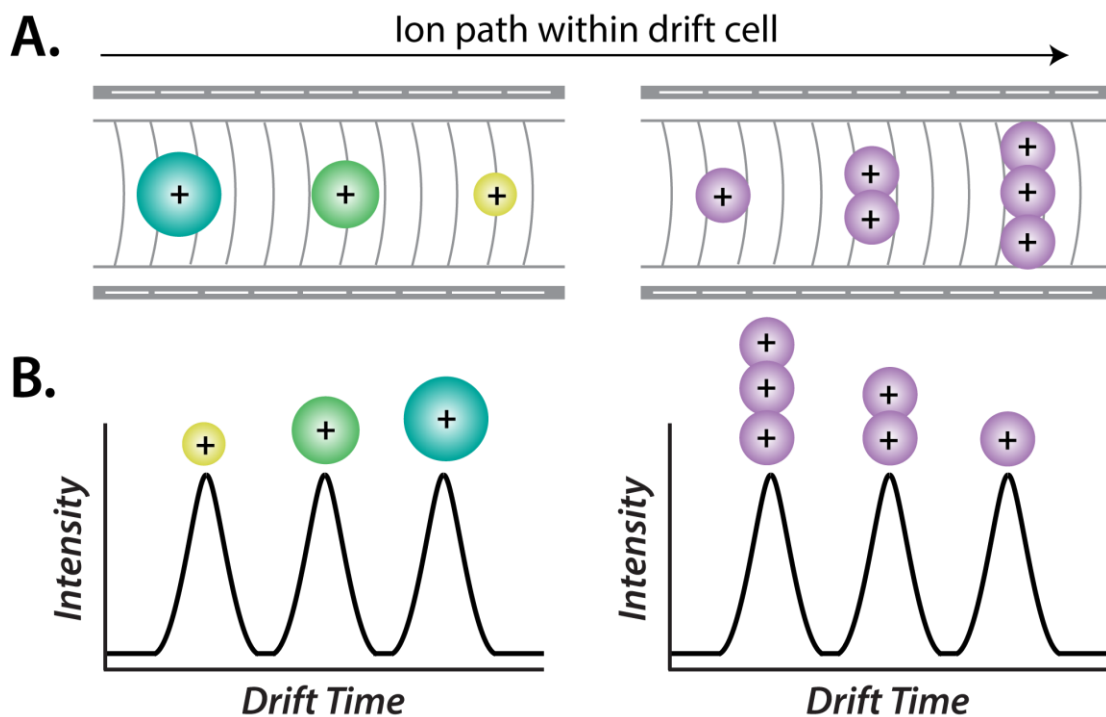


Figure 1.2. (A) Visualization of how molecules of the same charge state (left panels) and noncovalently bonded assemblies (right panels) are separated within the drift cell. (B) Interpretation of the accompanying ATDs for each type of species.

## Electrospray ionization

As mentioned prior, analytes that enter the drift cell (and subsequently, the mass analyzers) must first be ionized. Within our instrument, analytes in solution are gently transferred into the gas phase through an ESI-jet stream source. Since ESI is considered a “soft” ionization technique, little-to-no fragmentation of analyte molecules will occur, thus ensuring that weakly-bounded amyloid oligomers will stay intact and solution-phase properties are retained throughout the ionization process. Additionally, ESI allows for the population of multiply charged species which can allow for the observation of ions that are outside of the  $m/z$  range of the instrument.

The premise of ESI works by applying a high voltage (3500 - 3800 V) to analytes in solution which flow through a thin capillary needle. Within our ESI-source,  $N_2$  is used as a nebulizing or sheath gas to assist in higher flow rates of the sample. At the tip of the capillary, electrostatic forces and surface tension of the solution-phase sample compete to form a Taylor cone, which will emit highly charged droplets. As the size of the droplet decreases due to desolvation, the electrostatic repulsion of positive (or negative) charges outweighs the surface tension holding the droplet together. This results in Coulombic fission, in which the original droplet dissociates into many smaller droplets. This process repeats until only a stream of ions are left, which are then brought into the mass spectrometer inlet.<sup>3</sup> In the instrument used throughout this dissertation, the ESI spray is orthogonal to the mass spectrometer inlet. Heated nitrogen is used as a drying gas to further aid in evaporation of droplets and to increase instrument sensitivity. A schematic of the ESI-process can be found in Figure 1.3.

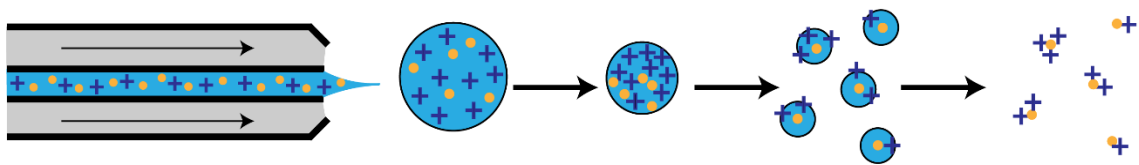


Figure 1.3. Visualization of ESI mechanism.

## 2 Structural flexibility of *Cyclosporine A* is mediated by *cis-trans* isomerization and the chameleonic roles of calcium

### 2.1 Disclosure

The following chapter has been adapted, with permission, from a journal article published in the *Journal of Physical Chemistry B* (Copyright 2021 American Chemical Society).<sup>4</sup> Modifications include the shortening the discussion of Cyclosporine A structures obtained from distance-geometry modeling and NMR spectroscopy, removal of the NMR tables and figures shown within the supporting information, and removal of the discussion comparing the bound CycA open conformation to the free CycA:Ca<sup>2+</sup> open conformation. This work was performed in conjunction with Carlos A. Steren, Guillermo A. Bermejo, Filippo Favretto, and Markus Zweckstetter. Carlos and Bermejo aided in the NMR efforts, Filippo and Markus synthesized the cyclophilin A protein, and Thanh Do performed the theoretical modeling of the CycA structures. The analysis of the IMS-MS data of CycA and its interactions with metals and cyclophilin A (CypA) were performed by myself under the guidance of Thanh Do.

### 2.2 Abstract

Falling outside of Lipinski's rule of five, macrocyclic drugs have accessed unique binding sites of their target receptors unreachable by traditional small molecules. Cyclosporin(e) A (CycA), an extensively studied macrocyclic natural product, is an immunosuppressant with undesirable side effects such as electrolytic imbalances. In this work, a comprehensive view on the conformational landscape of CycA, its interactions with Ca<sup>2+</sup>, and host-guest interactions with cyclophilin A (CypA) is reported through exhaustive analyses that combine ion-mobility spectrometry-mass spectrometry (IMS-MS), nuclear

magnetic resonance (NMR) spectroscopy, distance-geometry modeling, and NMR-driven molecular dynamics. Our IMS–MS data show that CycA can adopt extremely compact conformations with significantly smaller collisional cross sections than the closed conformation observed in CDCl<sub>3</sub>. To adopt these conformations, the macrocyclic ring has to twist and bend via *cis*–*trans* isomerization of backbone amides, and thus, we termed this family of structures the “bent” conformation. Furthermore, NMR measurements indicate that the closed conformation exists at 19% in CD<sub>3</sub>OD/H<sub>2</sub>O and 55% in CD<sub>3</sub>CN. However, upon interacting with Ca<sup>2+</sup>, in addition to the bent and previously reported closed conformations of free CycA, the CycA:Ca<sup>2+</sup> complex is open and has all-*trans* peptide bonds. Previous NMR studies using calcium perchlorate reported only the closed conformation of CycA (which contains one *cis* peptide bond). Here, calcium chloride, a more biologically relevant salt, was used, and interestingly, it helps converting the *cis*-MeLeu9–MeLeu10 peptide bond into a *trans* bond. Last, we were able to capture the native binding of CycA and CypA to give forth evidence that IMS–MS is able to probe the solution-phase structures of the complexes and that the Ca<sup>2+</sup>:CycA complex may play an essential role in the binding of CycA to CypA.

## 2.3 Introduction

Macrocycles are chemically diverse, strained molecules, some of which encompass a broad range of applications within the field of pharmaceuticals. Beyond Lipinski’s “rule of five”,<sup>5</sup> novel macrocycles have a unique ability to bind to open, flat, and groove-shaped binding sites such as protein–protein interfaces<sup>6</sup> in addition to exhibiting desirable pharmacokinetic properties such as passive membrane permeability and biostability.<sup>7–10</sup> Cyclosporin(e) A (CycA) is a well-characterized, fungal macrocyclic peptide<sup>11</sup> (Figure 2.1) that is profitable for its many medicinal uses. Most notably, CycA is used as an immunosuppressant to prevent the rejection of tissues in organ transplant patients as well as to treat symptoms of psoriasis and rheumatoid arthritis.<sup>12–15</sup> The mechanism in which CycA exhibits its immunosuppressive properties is through binding with its

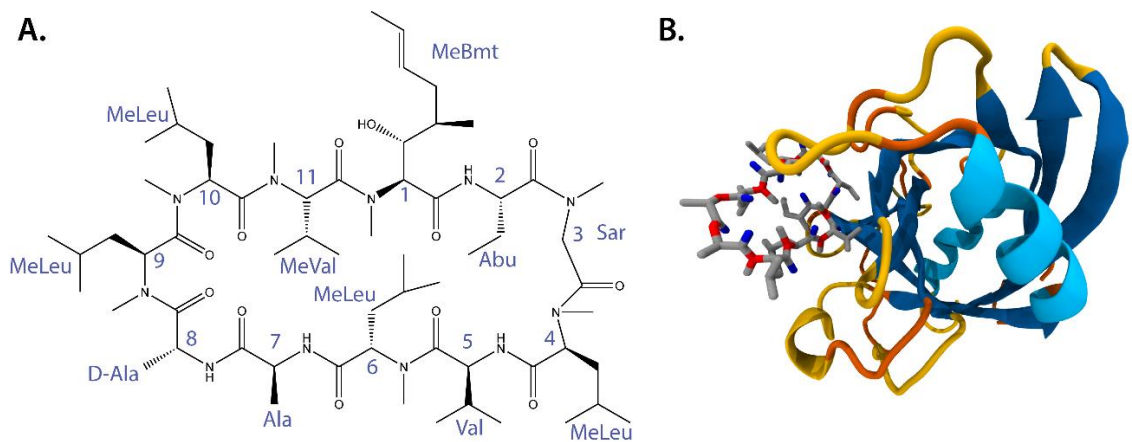


**Figure 2.1. (A) Chemical structure of CycA. (B) X-ray crystal structure of the monomeric CycA:CypA complex (PDB 1CWA).**

intracellular receptor cyclophilin A (CypA) (Figure 2.1).<sup>16</sup> Another name for CypA is peptidyl prolyl *cis/trans* isomerase A as it catalyzes the *cis/trans* isomerization of prolineimide peptide bonds and thereby regulates protein folding and trafficking.<sup>17</sup> Immunosuppression is ultimately achieved through T-cell inactivation once the CycA:CypA complex binds to the calcium and calmodulin-dependent serine/threonine protein phosphatase calcineurin.<sup>18</sup>

The structure of free CycA has been extensively studied at the molecular level with regard to solution-driven conformational isomerism using traditional solution [nuclear magnetic resonance (NMR) spectroscopy] and solid-state (X-ray crystallography) techniques.<sup>19-22</sup> It is well established that CycA exhibits so-called “chameleonic” behavior,<sup>23</sup> adopting conformations that depend largely on the environment. However, due to the poor solubility of CycA in aqueous solutions, most studies have been carried out in pure organic solvents (e.g., methanol, chloroform, and acetonitrile<sup>20, 21, 24, 25</sup>) that do not accurately represent biological conditions. Previously elucidated conformations can be classified as either an “open” or a “closed” structure. The closed conformation has four hydrogen bonds and a *cis*-amide between MeLeu9 and MeLeu10.<sup>22</sup> The “open” conformation, which is the form of CycA bound to its receptor CypA, has all *trans*-amide bonds and no hydrogen bonds.<sup>26</sup> The closed and open conformations were experimentally observed in separate experiments. In addition, according to molecular simulations, there are multiple closed and open conformations (i.e., “congruent” conformations<sup>27</sup>), indicating a complex conformational space with many transient structures.

Despite being regarded as one of the safest medicines and belonging to the World Health Organization’s list of essential drugs,<sup>28</sup> (24) CycA has been known to cause biochemical disturbances within the body such as the increased excretion of calcium and magnesium<sup>29</sup> and dyshomeostasis of electrolytes within the blood.<sup>30</sup> Due to these electrolytic imbalances, it has been of great interest to study the interaction of CycA with biologically relevant metal cations. Previous NMR and Fourier transform infrared

studies have elucidated that CycA has high affinities toward metal ions, which induce conformational changes in CycA.<sup>24, 25, 31</sup> In addition, the CycA:metal binding is stronger for divalent cations than for monovalent cations, with the largest known affinity being for calcium.<sup>25, 32</sup> CycA:Ca<sup>2+</sup> interactions may have biological implications on the electrolytic disturbances within CycA recipients. Previous NMR studies have used calcium perchlorate [Ca(ClO<sub>4</sub>)<sub>2</sub>]<sup>25</sup> instead of the more biologically relevant calcium chloride (CaCl<sub>2</sub>) as the source of Ca<sup>2+</sup>. This subtle difference, as will be discussed, may contribute to the conformation-induced activity of CycA upon binding to its target.

In the present study, our aims are twofold. First, we leverage what is known about free CycA and the CycA:CypA complex to demonstrate that ion-mobility spectrometry–mass spectrometry (IMS–MS) is an inexpensive and complementary approach to sample conformational space and host–guest interactions of macrocyclic drugs. Second, we connect the specific structures within CycA’s conformational ensemble to their abilities to selectively accommodate unique metal cations and bind CypA. We look further into the interactions between CycA and both endogenous and exogenous metals under aqueous conditions and how the macrocyclic drug binds to its protein target. IMS–MS is a powerful analytical tool that can preserve and isolate non-covalently bound complexes in the gas phase to allow the observation and distinction of isobaric species. Our recent successes in characterizing purely organic macrocycles indicate that IMS–MS retains solution-phase information.<sup>33, 34</sup> Drift-tube IMS allows the determination of collisional cross section (CCS) values of distinct conformers which can provide structural information about a species size and shape. Finally, we complement our IMS–MS data with NMR experiments under similar conditions to help resolve the discrepancies in the open versus closed form of CycA upon binding Ca<sup>2+</sup>. Our study illustrates the power of conformational analysis using both IMS–MS and NMR and at the same time sheds light on the strengths and weaknesses of each technique.

## 2.4 Materials and methods

### Ion-mobility spectrometry – mass spectrometry

The chemicals were of reagent grade and were used without further purification. CycA (98% purity, Tokyo Chemical Industry) was prepared for IMS–MS analysis by reconstituting the CycA solid in methanol at a 1 mM concentration and diluting it into liquid chromatography (LC)-grade water (Aldrich, OmniSolv) to 5–10  $\mu$ M. CypA was recombinantly produced (see below) and reconstituted to the desired concentration in LC-grade water. A mixture of CycA and CypA was prepared from their respective stocks in a 1:1 ratio into a solution of 20 mM ammonium acetate. Stock solutions of  $\text{CaCl}_2$  and KCl (VWR Chemicals) were prepared in LC-grade water and diluted to the desired concentrations.

IMS–MS experiments were performed on an Agilent 6560 IMS–MS QTOF (Agilent, Santa Clara). Analytes were ionized by electrospray ionization (ESI), and ions were subsequently pulsed into a helium-filled drift cell in a multifold experiment. An exit funnel and hexapole ion guide focused the ions into the QTOF mass spectrometer. Data were obtained in positive polarity over the course of 5.2 min with the drift cell voltages of  $\Delta V = 890, 790, 690, 590$ , and  $490$  V. For the CycA:CypA:metal mixture,  $\Delta V = 850, 750, 650, 550$ , and  $450$  V. The pressure of the drift cell was 3.94 Torr with the pressure differences between the drift tube and trap funnel being approximately 300 mTorr. Mass calibration and CCS validation were performed using the Agilent tune-mix.

Arrival time distributions (ATDs) and the resulting mass spectra were recorded with Agilent IM-MS Browser software and graphed with OriginPro. CCSs of ions were determined through relating the ion's arrival time ( $t_A$ ) to the change in drift cell voltage ( $\Delta V$ ) by Equation 3.1.<sup>1</sup>

$$t_A = \frac{L^2 T_0}{P_0 T} \frac{1}{K_0} \frac{P}{\Delta V} + t_0 \quad 2.1$$

Where  $L$  is the length of the drift cell (78.1 cm),  $K_0$  is the reduced mobility,  $P$  and  $T$  are the pressure and temperature of the buffer gas, respectively,  $T_0$  and  $P_0$  are the standard temperature and pressure, and  $t_0$  is the time that the ion spent outside of the drift cell.  $K_0$  is used to calculate the momentum-transfer collision integral through the Mason–Schamp (Equation 3.2), which is approximately the CCS ( $\sigma_{exp}$ ),<sup>1</sup>

$$\sigma_{exp} \approx \frac{3ze}{16N_0} \frac{1}{K_0} \sqrt{\frac{2\pi}{\mu k_B T}} \quad 2.2$$

where  $z$  is the ionic charge,  $e$  is the elementary charge,  $N_0$  is the number gas density,  $\mu$  is the reduced mass of the ion-buffer gas pair,  $k_B$  is the Boltzmann constant, and  $T$  is the temperature of the buffer gas.

### Distance geometry modeling and theoretical CCS calculations

Initial structures of the macrocycles were generated by the *Builder/Sketcher* module in VIDA v4.4.0.4 (OpenEye) starting from a 2D skeletal structure. The *macrocycle* module in OMEGA v3.0.0.1 (OpenEye)<sup>35, 36</sup> was used to generate ensembles of conformations, which were clustered based on root-mean-square deviation. Conformations in vacuum and water were generated by setting *dielectric\_constant* to 1.0 and 80, respectively. Theoretical CCSs were calculated using the trajectory method (CCS<sub>TJ</sub>) available in the MobCal package.<sup>37, 38</sup>

### NMR spectroscopy

Stock CycA was prepared in CD<sub>3</sub>OD (deuterated methanol; Aldrich) or CD<sub>3</sub>CN (deuterated acetonitrile; Aldrich), and stocks of CaCl<sub>2</sub> were prepared in D<sub>2</sub>O (Cambridge Isotope Laboratories, Inc.). One sample was prepared in CD<sub>3</sub>OD with 10% D<sub>2</sub>O/H<sub>2</sub>O at 5 mM of CycA (A-1). Three samples were prepared in CD<sub>3</sub>CN from their respective stocks: 5 mM of CycA (B-1), 5 mM CycA in 500  $\mu$ M Ca<sup>2+</sup> (B-2), and 5 mM CycA in 5 mM Ca<sup>2+</sup> (B-3). Due to the addition of D<sub>2</sub>O in the samples containing Ca<sup>2+</sup>, the same amount of D<sub>2</sub>O was added to the B-1 sample. Each sample contained less than 1% D<sub>2</sub>O.

NMR experiments were performed on a Varian VNMRS 600 MHz spectrometer equipped with a room-temperature  $^1\text{H}/^{13}\text{C}/^{15}\text{N}$  triple-resonance probe. 1D- $^1\text{H}$  and 2D gCOSY, eCOSY, zTOCSY, ROESYAD, NOESY,  $^{13}\text{C}$ -gHSQCAD, and  $^{13}\text{C}$ -gHMBCAD experiments were obtained for each sample. For the samples with  $\text{H}_2\text{O}$ , the experiments used the Presat water suppression sequence. The temperature was regulated at 25 °C unless indicated otherwise. 1D spectra were processed and analyzed with MestReNova software from Mestrelab Research. 2D experiments were processed with NMRPipe, distributed by the Institute for Bioscience and Biotechnology Research, and converted to the Sparky-UCSF format for their analysis with NMRFAM-Sparky, National Magnetic Resonance Facility at Madison.

NMR structure determinations were performed with the program Xplor-NIH<sup>39</sup> using protocols based on those described elsewhere.<sup>40</sup> Briefly, the structure calculation procedure (used for separate computations; see below) consists of three stages. Each stage generates 100 structures via a common molecular dynamics (MD)/simulated annealing scheme that implements a previously described force field,<sup>40</sup> with the following exceptions: (i) the addition of parameters extracted from high-resolution X-ray structures involving the “uncommon” amino acid residues Abu, Sar, MeLeu, MeVal, and MeBmt; (ii) the exclusion of statistical potentials,<sup>41, 42</sup> which are derived from known conformations of standard residues, due to the abundance of the uncommon ones in CycA (8 out of 11 residues, including 7 backbone N-methylated instances; see Figure 2.1). In the first stage, calculations are initiated from an extended configuration of CycA that excludes the N–C' bond between MeBmt1 and MeVal11. Second, the resulting lowest-energy structure from the previous stage is refined with the addition of a distance restraint between the above-mentioned atoms. The final stage performs further refinement, while properly accounting for the covalent geometry associated with the MeBmt1–MeVal11 peptide bond. All calculation stages implemented rotating frame Overhauser effect (ROE)-based distance restraints as the sole source of experimental information.

### Cyclophilin A expression

CypA was recombinantly produced and purified as reported previously.<sup>43</sup> Briefly, the gene of human CypA was cloned into a modified pET28a vector (Addgene), carrying a TEV cleavage site at the N-terminus of the protein to remove the His<sub>6</sub> tag. The plasmid was successively transformed in *Escherichia coli* BL21(DE3) cells (Novagen). Protein overexpression was induced adding 0.4 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside to the medium after an OD<sub>600</sub> of 0.6 was reached. The overexpression was carried on at 16 °C for 16 hours.

## 2.5 Results and discussion

### IMS-MS reveals a unique compact conformation

Figure 2.2 shows a representative ESI-mass spectrum of 10  $\mu$ M CycA in water. Due to the poor solubility of CycA in water, the IMS-MS sample was prepared via a stock solution in methanol and then diluted to the desired concentration in LC-MS water. The mass spectrum of CycA shows two distinct populations of mass spectral peaks with  $z = +1$  and  $z = +2$  charge states. Within each of these populations, there are various species including protonated species and metal adducts. The peaks are tentatively assigned based on the mass-to-charge ratio ( $m/z$ ); however, the identities of the adducts are later confirmed by directly adding the metal cations to the samples, as discussed in the next section.

Observations of sodiated and potassiated adducts of CycA are not surprising given that these ions are ubiquitous in MS solvents. However, the presence of any divalent cations (such as Ca<sup>2+</sup>, the subject of discussion in the next section) must be endogenously present within the sample since LC solvents are known to contain negligible amounts of these ions.

In the previous studies using IMS-MS, McLean and co-workers<sup>44</sup> reported that the protonated CycA ( $z = +1$ ) exhibits a CCS of 297 Å<sup>2</sup>, while Shapiro and co-workers<sup>45</sup> found

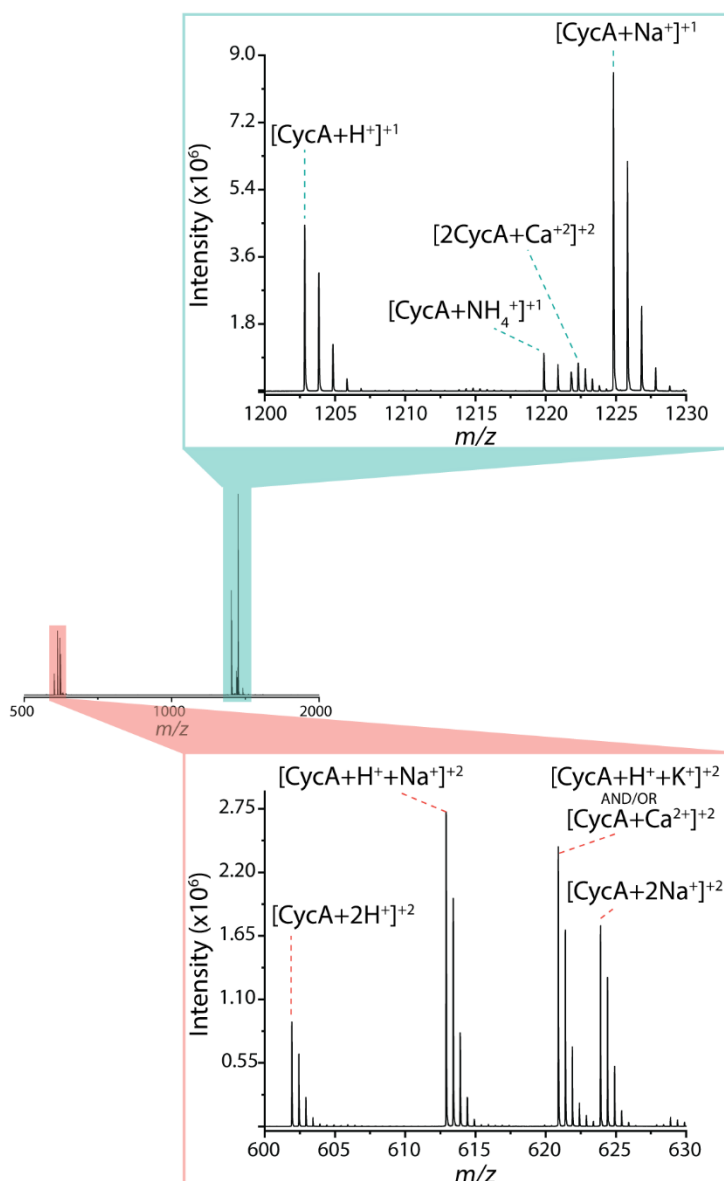


**Figure 2.2. Representative mass spectrum of CycA in water. The blue highlighted region shows mainly  $z = +1$  species, while the red highlight encompasses  $z = +2$  species. Note that each mass spectral peak corresponding to a distinct species (e.g.  $[\text{CycA}+\text{H}^+]^{+1}$ ) has several signals which arise from the multiple isotopes of the elements that the species is composed of. For proteins and peptides, the isotopic distribution occurs from the carbon-12 and carbon-13 isotopes.**

two conformers of the protonated CycA at  $z = +2$  with CCSs of 266 and 275 Å<sup>2</sup>. These data, when viewed together, contradict the perception that the CCS is directly proportional to charge state. As the charge state increases, so does the CCS; the size of the ion increases to accommodate like-charge repulsions. Of note, the CCS values from Shapiro and co-workers are significantly smaller than those calculated from any of the NMR and X-ray structures available (see Table 2.1). In such a study, the CCSs of CycA conformers were determined via a traveling-wave IMS instrument and then calibrating those values with references in helium.<sup>45</sup> In the current study, we obtained the absolute CCS values in helium gas using a multifield experiment with a drift-tube IMS–MS instrument as detailed below.

In the current study, we found two populations of singly charged CycA ( $z = +1$ ) centered at CCS values of 271 and 282 Å<sup>2</sup> (Figure 2.3), both with broad distributions ( $\pm 5$  Å<sup>2</sup>) indicating the presence of conformational ensembles. On the other hand, only one major population of doubly charged CycA ( $z = +2$ ) was observed, with a CCS of 297 Å<sup>2</sup> (Figure 2.3). Note that our data show the expected positive correlation between charge states and CCSs. Our largest observed CCS is in agreement with that reported by McLean and co-workers,<sup>44</sup> although associated with a higher charge. McLean and co-workers used matrix-assisted laser desorption/ionization as the ionization method which rarely produces multiply charged species. In their data, the sodiated CycA also has a smaller CCS than the free CycA at the same charge state, in agreement with our observation here (Table 2.2).

The highlighted regions within Figure 2.3 include CCS values calculated from all previous NMR and X-ray crystallographic structures (Table 2.1). Thus, Figure 2.3A suggests that IMS–MS captured a set of conformations that have similar CCSs to those from Shapiro and co-workers and are more compact than any solution or solid-state structures determined by NMR and crystallography. The structures are significantly more compact than the X-ray structure in chloroform (DEKSAN) whose CCS is 281 Å<sup>2</sup> (this population also has similar CCSs to the sodiated and some of the calcinated CycA

**Table 2.1. Theoretical CCSs of modeled CycA and structures elucidated from past X-ray crystallography and NMR studies.**<sup>a</sup> 22, 26, 46-57

| <b>X-ray crystallographic structures</b>                       |                    |                                         |                       |               |
|----------------------------------------------------------------|--------------------|-----------------------------------------|-----------------------|---------------|
| <i>PDB ID/CSD ID</i>                                           | <i>Chain</i>       | <i>CCS<sub>TJ</sub> (Å<sup>2</sup>)</i> | <i>Resolution (Å)</i> | <i>R9-R10</i> |
| 1cwa                                                           | C                  | 293.0                                   | 2.10                  | <i>trans</i>  |
| 2x2c                                                           | B                  | 288.3                                   | 2.41                  | <i>trans</i>  |
| 2x2c                                                           | F                  | 289.5                                   | 2.41                  | <i>trans</i>  |
| 2x2c                                                           | L                  | 287.9                                   | 2.41                  | <i>trans</i>  |
| 2x2c                                                           | P                  | 289.2                                   | 2.41                  | <i>trans</i>  |
| 2x2c                                                           | R                  | 290.5                                   | 2.41                  | <i>trans</i>  |
| 2x7k                                                           | B                  | 288.2                                   | 1.15                  | <i>trans</i>  |
| 2z6w                                                           | M                  | 291.1                                   | 0.96                  | <i>trans</i>  |
| 2z6w                                                           | N                  | 290.2                                   | 0.96                  | <i>trans</i>  |
| 2wfj                                                           | B                  | 287.3                                   | 0.75                  | <i>trans</i>  |
| 3bo7                                                           | E                  | 292.7                                   | 2.35                  | <i>trans</i>  |
| 3eov                                                           | C                  | 288.9                                   | 2.60                  | <i>trans</i>  |
| 3eov                                                           | D                  | 290.4                                   | 2.60                  | <i>trans</i>  |
| 3pmp                                                           | C                  | 290.3                                   | 1.47                  | <i>trans</i>  |
| 3pmp                                                           | D                  | 292.4                                   | 1.47                  | <i>trans</i>  |
| 4dgc                                                           | F                  | 291.3                                   | 2.65                  | <i>trans</i>  |
| 4dgc                                                           | G                  | 293.0                                   | 2.65                  | <i>trans</i>  |
| 4hy7                                                           | B                  | 290.7                                   | 1.20                  | <i>trans</i>  |
| 4jjm                                                           | E                  | 290.7                                   | 2.09                  | <i>trans</i>  |
| 4jjm                                                           | F                  | 292.2                                   | 2.09                  | <i>trans</i>  |
| 4yv9                                                           | E                  | 279.3                                   | 1.95                  | <i>trans</i>  |
| 4yv9                                                           | F                  | 285.0                                   | 1.95                  | <i>trans</i>  |
| 4yv9                                                           | G                  | 279.9                                   | 1.95                  | <i>trans</i>  |
| 4yv9                                                           | H                  | 279.6                                   | 1.95                  | <i>trans</i>  |
| <b>DEKSAN</b>                                                  | -                  | 280.8                                   | -                     | <i>cis</i>    |
| <b>NMR structures</b>                                          |                    |                                         |                       |               |
| <i>PDB ID</i>                                                  | <i>Solution</i>    | <i>Theoretical CCS (Å<sup>2</sup>)</i>  |                       |               |
| 1cya                                                           |                    | 298.5                                   |                       | <i>trans</i>  |
| -                                                              | CDCl <sub>3</sub>  | 284.6                                   |                       | <i>cis</i>    |
| -                                                              | CDCl <sub>3</sub>  | 286.9                                   |                       | <i>cis</i>    |
| <b>Modeled structures (distance geometry) within this work</b> |                    |                                         |                       |               |
|                                                                | <i>Environment</i> | <i>Theoretical CCS (Å<sup>2</sup>)</i>  |                       |               |
|                                                                | vacuum             | see discussion                          |                       | mixed         |
|                                                                | aqueous            | see discussion                          |                       | mixed         |

<sup>a</sup> R9-R10 annotates the isomerism of the peptide bond between MeLeu9 and MeLeu10.

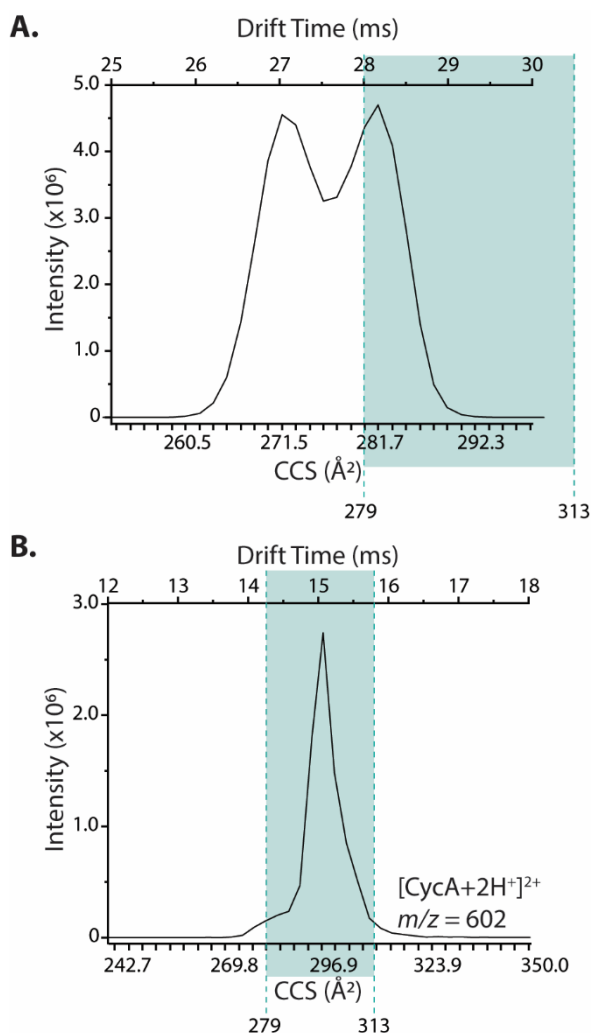


**Figure 2.3.** ATDs of (A)  $[\text{CycA} + \text{H}^+]^{1+}$  and (B)  $[\text{CycA} + 2\text{H}^+]^{2+}$  species seen within the mass spectrum of Figure 2.2. Note that the features seen within each ATD corresponds to all the isotopic signals of their respective mass spectral peak. The bottom and top x-axes display CCS and drift time, respectively. The blue highlight encompasses theoretical CCS values of CycA structures modeled here and by previous NMR and X-ray crystallography studies (see Table 2.1)

**Table 2.2. Experimental collisional cross sections of CycA and CypA species**

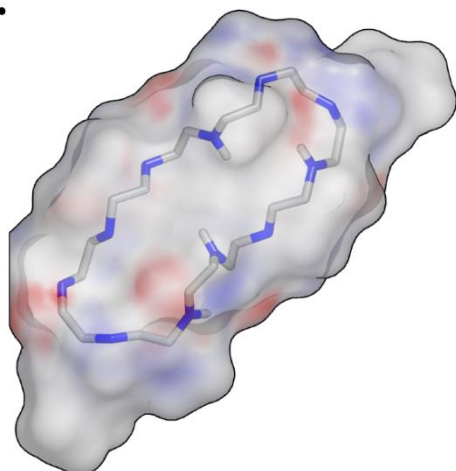
| Species                                                                          | Experimental CCS ( $\text{\AA}^2$ ) |
|----------------------------------------------------------------------------------|-------------------------------------|
| <b><i>Cyclosporin A</i></b>                                                      |                                     |
| $[\text{CycA} + \text{H}^+]^{+1}_{\text{closed}}$                                | 281.7                               |
| $[\text{CycA} + \text{H}^+]^{+1}_{\text{bent}}$                                  | 271.1                               |
| $[\text{CycA} + 2\text{H}^+]^{+2}$                                               | 297.2                               |
| $[\text{CycA} + \text{Ca}^{2+}]^{+2}_{\text{open}}$                              | 296.8                               |
| $[\text{CycA} + \text{Ca}^{2+}]^{+2}_{\text{closed}}$                            | 280.4                               |
| $[\text{CycA} + \text{Ca}^{2+}]^{+2}_{\text{bent}}$                              | 269.8                               |
| $[2\text{CycA} + \text{Ca}^{2+}]^{+2}$                                           | 444.4                               |
| $[\text{CycA} + \text{Na}^+]^{+1}$                                               | 267.2                               |
| <b><i>Cyclosporin A-Cyclophilin A Complex</i></b>                                |                                     |
| $[\text{CycA} + \text{CypA} + 8\text{H}^+]^{+8}$                                 | 1726.3                              |
| $[\text{CycA} + \text{CypA} + 9\text{H}^+]^{+9}$                                 | 1792.4                              |
| $[\text{CycA} + \text{CypA} + 8\text{H}^+]^{+8}$ (Presence of $\text{Ca}^{2+}$ ) | 1774.4                              |
| $[\text{CycA} + \text{CypA} + 9\text{H}^+]^{+9}$ (Presence of $\text{Ca}^{2+}$ ) | 1842.4                              |
| $[\text{CycA} + \text{CypA} + 8\text{H}^+]^{+8}$ (Presence of $\text{K}^+$ )     | 1767.7                              |
| $[\text{CycA} + \text{CypA} + 9\text{H}^+]^{+9}$ (Presence of $\text{K}^+$ )     | 1837.9                              |
| <b><i>Cyclophilin A</i></b>                                                      |                                     |
| $[\text{CypA}]^{+7}$                                                             | 1508*                               |
| $[\text{CypA}]^{+8}$                                                             | 1588.2*                             |
| $[\text{CypA}]^{+9}$                                                             | 2259                                |
| $[\text{CypA}]^{+10}$                                                            | 2427.1                              |
| $[\text{CypA}]^{+11}$                                                            | 2818.7                              |
| $[\text{CypA}]^{+12}$                                                            | 2983.9                              |
| $[\text{CypA}]^{+13}$                                                            | 3085.4                              |
| $[\text{CypA}]^{+14}$                                                            | 3196.9                              |
| $[\text{CypA}]^{+15}$                                                            | 3295.4                              |
| $[\text{CypA}]^{+16}$                                                            | 3390.8                              |
| $[\text{CypA}]^{+17}$                                                            | 3473.9                              |
| $[\text{CypA}]^{+18}$                                                            | 3586                                |
| $[\text{CypA}]^{+19}$                                                            | 3725.4                              |
| $[\text{CypA}]^{+20}$                                                            | 3749.5                              |
| $[\text{CypA}]^{+21}$                                                            | 3897.9                              |
| $[\text{CypA}]^{+8}$ (Presence of $\text{K}^+$ )                                 | 1679.2                              |
| $[\text{CypA}]^{+8}$ (Presence of $\text{Ca}^{2+}$ )                             | 1698.1                              |

\*Due to the low intensity of the mass spectral peaks, there may be slight error (~2%) in the reported CCS for low-charge state CypA ions ( $z = +7$  and  $z = +8$ ) and CycA:CypA complexes.

adducts which will be discussed later). In summary, the ATD of  $[\text{CycA} + \text{H}^+]^{+1}$  is a bimodal

distribution that contains at least two distinct families<sup>58</sup> (based on the widths of the ATD peaks). The first population, which has a slower mobility (larger CCS) is a combination of solution closed ( $\text{CCS} \leq 281 \text{ \AA}^2$ ) and open ( $\text{CCS} \geq 285 \text{ \AA}^2$ ) conformations. The lower limit of this CCS range (blue rectangle in Figure 2.3A) has a similar CCS to that of DEKSAN<sup>22</sup> (the *closed* conformation in  $\text{CDCl}_3$ ) and extends to partially cover the theoretical CCSs of CycA extracted from protein–ligand complexes. Of note, in all protein–ligand complexes, CycA adopts an *open* conformation.<sup>59</sup> The other population centered at  $271 \text{ \AA}^2$  cannot be explained by comparing the experimental CCS with those inferred from existing NMR and X-ray structures. To further investigate this compact population of structures, distance geometry (DG) modeling was used. Results of the modeling reveal a set of compact structures under vacuum conditions, including one structure shown in Figure 2.4, which has a theoretical CCS of  $270 \text{ \AA}^2$  and contains three *cis*-amide bonds and two internal hydrogen bonds between MeBmt1-DAla8 and Abu2-Val5. More importantly, this structure resembles a minor population of CycA in  $\text{CDCl}_3$ <sup>60</sup> and the NMR structures of CycA in the complex with  $\text{Mg}^{2+}$  and  $\text{Ce}^{3+}$  (perchlorate) in acetonitrile solved by Chowdhry<sup>25</sup> and Valensin<sup>24</sup> and co-workers, respectively, not only in the overall shape but also in the presence of several *cis*-amides. The *cis* Sar3–MeLeu4 and Val5–MeLeu6 amides were observed in the complex of CycA with  $\text{Mg}^{2+}$  and  $\text{Ce}^{3+}$ ,<sup>24</sup> respectively, suggesting that the *cis/trans* isomerization of backbone amides is critical for CycA to explore extremely compact structures. Our DG modeling shows that such compactness is achieved by twisting and bending of the macrocyclic backbone. In our modeling, compact structures with theoretical CCSs less than  $280 \text{ \AA}^2$  often contain two to three *cis*-amide bonds. The ring strains caused by twisting of the backbone are compensated by hydrogen bonding. These conformers may already exist in solution as minor populations, and the binding to divalent metals can shift the equilibrium to their favor. Herein, we refer to this class of conformations as “*bent*”-CycA in order to distinguish it

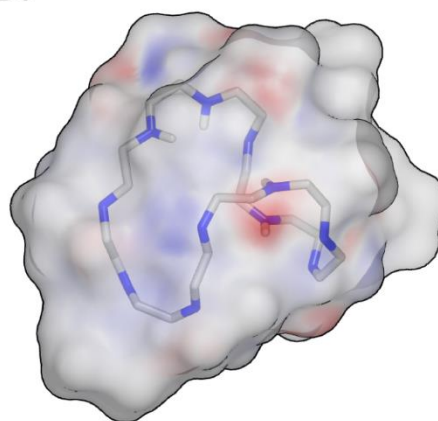
**A.**



**DEKSAN**

$$\text{CCS}_{\text{He}} = 281 \text{ \AA}^2$$

**B.**



***bent-CycA***

$$\text{CCS}_{\text{He}} = 270 \text{ \AA}^2$$

Figure 2.4. Molecular surfaces and backbone traces of (A) DEKSAN: the X-ray structure of CycA in chloroform (the closed conformation) and (B) the most compact conformer of CycA obtained from DG modeling in vacuum. The latter contains three *cis*-amides: MeLeu9-MeLeu10, MeLeu10-MeVal11, and Sar3-MeLeu4.

from the closed conformation with only one *cis*-amide (e.g., DEKSAN;<sup>22</sup> Figure 2.4). This class of conformers has more than one *cis*-amide, as exemplified by the structure shown in Figure 2.4B. In addition to the *cis* peptide bond between the residues MeLeu9–MeLeu10, Abu2–Sar3, and Sar3–MeLeu4, there are other two peptide bonds with high *cis* propensity. Seebach et al.<sup>60</sup> investigated [CS–NH<sup>MeBmt</sup>]CycA (a thiocyclosporin derivative of CycA with the C=O of MeBmt replaced with C=S), which showed that in addition to the closed conformation, there is a less abundant but still a major population, with a *cis* bond between Sar3 and MeLeu4.

**Endogenous Ca<sup>2+</sup> binding to CycA induces *bent*, *closed*, and *open* conformations. NMR shows that CycA:Ca<sup>2+</sup> is open with all-*trans* peptide bonds**

In Figure 2.2, there is a pair of well-resolved and reproducible peaks at  $m/z = 621$  and  $m/z = 1222$ . The isotopic spacings suggest that both are doubly charged adducts of the CycA monomer and dimer, respectively. The mass difference between these species and their corresponding bare, neutral CycA species is 40 amu. At first, we tentatively assigned them to be [CycA + H<sup>+</sup> + K<sup>+</sup>]<sup>2+</sup> ( $m/z$  621) and [2CycA + H<sup>+</sup> + K<sup>+</sup>]<sup>2+</sup> ( $m/z$  1222) since potassium (together with sodium) is often found in MS solvents. However, such assignments were ambiguous given the absence of the corresponding sodiated species for the dimer species. In fact, previous studies have shown that Na<sup>+</sup> and K<sup>+</sup> bind weakly to CycA.<sup>25, 31, 32</sup> Furthermore, direct addition of KCl to the sample (up to 0.5 mM) did not increase the intensities of the peaks. In fact, the peak at  $m/z$  1222 drastically decreased in intensity, whereas that at  $m/z$  1240 ([CycA + K<sup>+</sup>]<sup>+1</sup>) grew strongly. This led us to speculate that at least one peak, if not both peaks, could be a Ca<sup>2+</sup> adduct.

In the  $m/z$  1220–1224 region (Figure 2.5A,B), there are overlapping isotopic distributions ( $z = +1$  and  $+2$  with  $z = +1$  tentatively assigned to [CycA + NH<sub>4</sub><sup>+</sup>]<sup>+</sup>). To verify the existence of the Ca<sup>2+</sup>-bound CycA dimer, we prepared solutions of CycA with varying Ca<sup>2+</sup> concentrations to qualitatively examine peak ratios. The corresponding ATDs confirm that the feature at  $m/z$  1222 is a complex of the Ca<sup>2+</sup> ion with a CycA dimer as

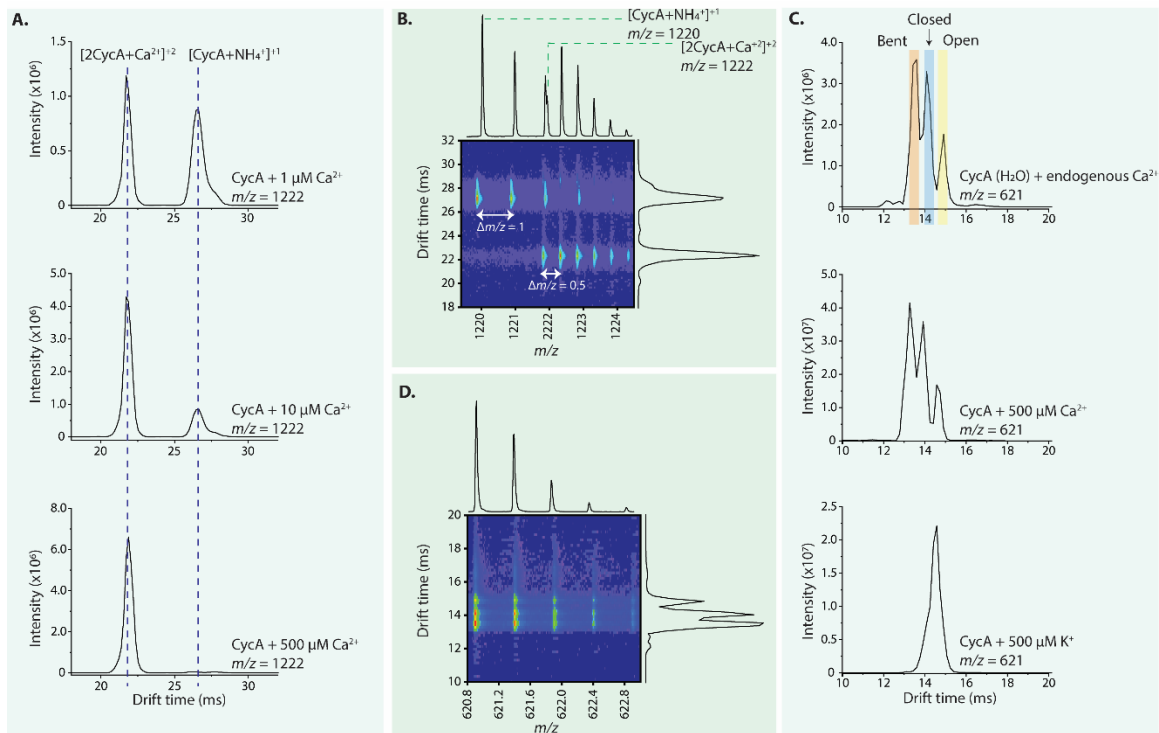


**Figure 2.5. (A)** ATDs corresponding to  $m/z = 1222$  in varying concentrations of  $\text{Ca}^{2+}$ . As  $\text{Ca}^{2+}$  concentration increases, so does the feature which corresponds to the dimer CycA in the complex with a  $\text{Ca}^{2+}$ . **(B)** 2D plot of drift time vs  $m/z$  showing the region  $m/z$  1220–1224 containing the overlapping isotopic distributions. **(C)** ATDs of  $m/z = 621$  are shown in varying environments. The top spectrum shows the ATD of the species in  $\text{H}_2\text{O}$ , the middle spectrum shows the ATD of the species in  $\text{H}_2\text{O}$  with added  $\text{Ca}^{2+}$ , and the bottom spectrum shows the ATD of the species in  $\text{H}_2\text{O}$  with added  $\text{K}^+$ . **(D)** 2D plot of drift time vs  $m/z$  from  $m/z$  620.8 to 623.0 in water shows three distinct conformations with the same isotopic spacing.

its intensity increases when more  $\text{Ca}^{2+}$  ( $\text{CaCl}_2$ ) was added. The  $\text{CycA-NH}_4^+$  feature can be seen as the control (Figure 2.5A). Further investigation of the 2D ion-mobility plot of  $m/z$  1222 also reinforces that the features correspond to two different species (Figure 2.5B). Indeed, Cirigliano and Cabrera<sup>61</sup> also observed the same dimeric complexes at a  $\text{Ca}^{2+}$  concentration as low as 0.1 mM. Here, we observed *endogenous*  $\text{Ca}^{2+}$  binding to dimeric CycA, which indicates (a) its strong affinity to CycA and (b) the existence of  $\text{Ca}^{2+}$  in its native source. No quantitative analysis was performed to determine the exact ratio of  $\text{CycA}:\text{Ca}^{2+}$  or the amount of endogenous  $\text{Ca}^{2+}$  in the CycA powder.

Similarly, at  $m/z$  621, the ATD shows three distinct conformations (Figure 2.5C). Since  $\text{K}^+$  is ubiquitous in MS analysis, it was reasonable to assign this peak as  $[\text{CycA} + \text{H}^+ + \text{K}^+]^{2+}$ , but the presence of the dimeric  $\text{CycA}:\text{Ca}^{2+}$  complex led us to believe otherwise. To determine if these conformations belong to the  $[\text{CycA} + \text{Ca}^{2+}]^{2+}$  species, two separate samples of CycA were prepared in 500  $\mu\text{M}$  aqueous solutions of either  $\text{Ca}^{2+}$  or  $\text{K}^+$ . The ATD of CycA in  $\text{Ca}^{2+}$  shows three distinct conformations, but only one conformation was observed in the CycA and  $\text{K}^+$  sample. The three conformations of  $[\text{CycA} + \text{Ca}^{2+}]^{2+}$  have experimental CCSs matching the bent, closed, and open conformations of free CycA. Further analysis of the 2D ion-mobility plot of drift time versus  $m/z$  (Figure 2.5D) shows that the isotopic spacings of the three conformations are the same, so each species must have the same charge state. This result ultimately suggests that the bent conformations can be introduced by  $\text{Ca}^{2+}$  binding (but not by  $\text{K}^+$  or  $\text{Na}^+$  binding). At the same time,  $\text{Ca}^{2+}$  also appears to preserve a population of open conformations. This is intriguing and demonstrates the chameleonic effect of  $\text{Ca}^{2+}$  upon binding of CycA (unlike other monovalent metals such as  $\text{Na}^+$  or  $\text{K}^+$ ; see Figure 2.5C). While the structural compactness driven by  $\text{Ca}^{2+}$  (similar to  $\text{Mg}^{2+}$ ) was reported in the literature,<sup>24, 25</sup> there has not been any report showing that  $\text{Ca}^{2+}$  can also bind to the open conformation of CycA. To further investigate this phenomenon, we turned to NMR.

Since previous NMR studies of metal-induced conformational changes in CycA used acetonitrile ( $\text{CD}_3\text{CN}$ ) as the solvent,<sup>24, 25, 32</sup> we carried out a set of experiments using the

same solvent and varying the concentrations of  $\text{CaCl}_2$ . At 5 mM CycA and no calcium (sample B-1), the closed conformations are favored (~55% population). They are evidenced by  $^1\text{H}$  and  $^{13}\text{C}$  chemical shift values that are similar to those previously reported for the closed form of CycA in chloroform<sup>22</sup> ( $\text{CDCl}_3$ ). The NMR structure elucidation of this conformer showed that it is very similar to the previously determined closed form (Figure 2.6). Thus, combined with our previous NMR findings (see above), these results suggest that the closed conformation of CycA is found in solvents of different polarities: methanol, chloroform, and acetonitrile.

Our data indicate that the long-established view on the competition between closed versus open conformations of CycA solely driven by solvent polarity is incomplete. If the closed conformation (with the *cis* MeLeu9–MeLeu10 peptide bond) remains significant in polar solvents, as shown by our NMR data here and the work by Ko and Dalvit,<sup>21</sup> then, how can CycA bind to CypA in an open, all-*trans* configuration? Given the abundance of the closed conformation in polar solvents (~19% in protic methanol and ~55% in aprotic acetonitrile), the conformational transition triggered by solvent polarity alone appears to be inefficient. Previous studies show that adding LiCl improves binding of CycA to CypA.<sup>62</sup> As discussed below, an updated mechanism may involve ions such as  $\text{Ca}^{2+}$  that upon interacting with CycA modulate the preference of closed versus open, all-*trans* conformations.

The HSQC spectrum obtained from NMR reveals that the closed conformation of CycA is dominant in acetonitrile (Sample B-1, ~46% population). However, when CycA- $\text{CaCl}_2$  are present in a 1:1 ratio (5 mM each, sample B-3), the  $^1\text{H}$  spectrum shows major changes compared to B-1 in both the backbone amide and *N*-methyl regions (Figure 2.6). Further analysis of 2D NMR spectra indicates the presence of a new conformer (~46% population at 25 °C and ~70% population at 5 °C) which was assigned to a CycA: $\text{Ca}^{2+}$  complex. Evidence of the interaction of  $\text{Ca}^{2+}$  with CycA in the CycA: $\text{Ca}^{2+}$  complex is the downfield shift of the amide proton of D-Ala 8 (to 9.88 ppm) as such shifts are observed in proteins that incorporate calcium. A similar downfield shift of

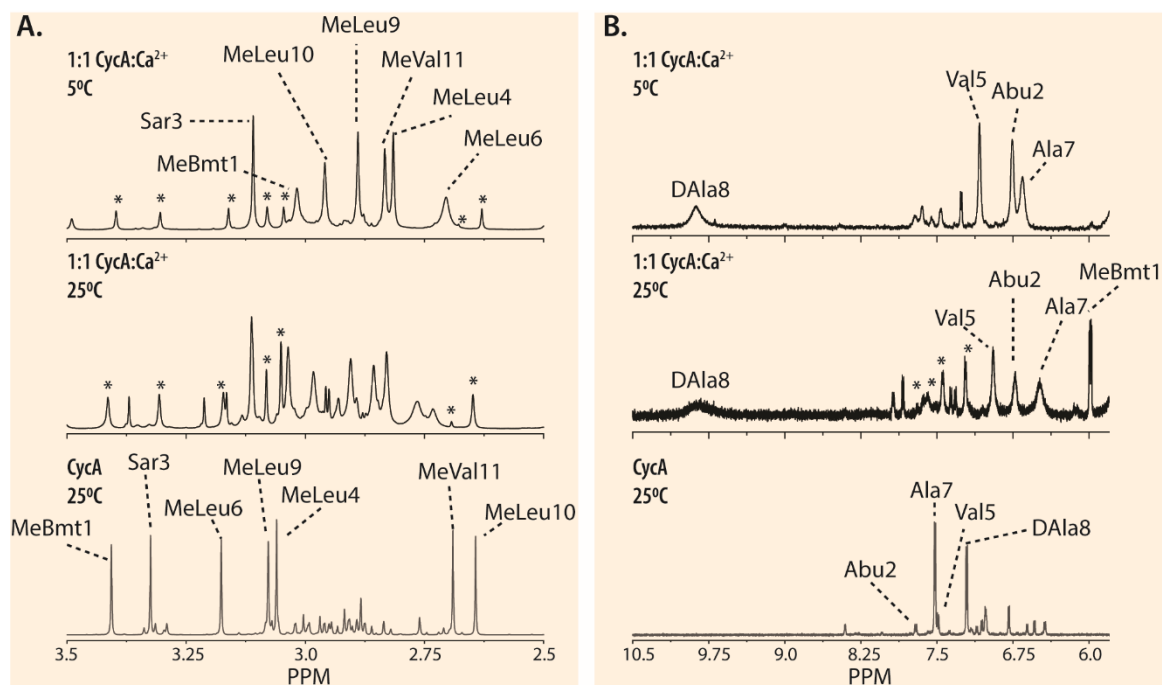


**Figure 2.6.** Proton spectra showing  $\text{H}^{\text{N}}$  (A) and  $\text{N}$ -methyl (B) regions of CycA samples in acetonitrile without (B-1 sample) and with  $\text{CaCl}_2$  (at a 1:1 ratio; sample B-3) at the indicated temperatures. The asterisks annotate common peaks among the samples and conditions.

one of the amide protons has been observed in a study of CycA:Ca<sup>2+</sup> in acetone [(CD<sub>3</sub>)<sub>2</sub>CO] by Dancer and Jones.<sup>63</sup> However, this study was performed with only 1D <sup>1</sup>H and <sup>13</sup>C NMR experiments with no assigned chemical shifts. More importantly, our ROESY spectra (whose chemical shifts are detailed within the supporting information of Gray et al.<sup>4</sup>) indicate that the peptide bonds of the CycA:Ca<sup>2+</sup> conformer are all-*trans*. The upfield shifts of the amide protons of Abu2, Val5, and Ala7 in CycA:Ca<sup>2+</sup> compared to free CycA indicate that the intra hydrogen bonds O...HN are broken. When compared to the results of published by Marjanović et al.<sup>64</sup> and Dancer et al.,<sup>63</sup> whom studied the CycA:Ca<sup>2+</sup> complex in THF and acetone respectively, our data collectively agree that CycA:Ca<sup>2+</sup> adopts the same peptide bond conformation (all-*trans*) in acetonitrile, acetone, and THF. That is, Ca<sup>2+</sup> can override solvent polarity to dictate the conformational preference of CycA.

A striking and major finding in this study lies in the differences between the structure of CycA:Ca<sup>2+</sup> (CaCl<sub>2</sub>) here and the one solved by Chowdhry and co-workers,<sup>25</sup> who used Ca(ClO<sub>4</sub>)<sub>2</sub>. Using perchlorate salt, they showed that the *cis* peptide bond between MeLeu9 and MeLeu10 was preserved, and no amide signals were downfield-shifted. As a result, the choices of metal salts are as important as, if not more than, those of solvents and metals. In addition, our data and the work by Bernardi et al.<sup>65</sup> suggest that chloride salts (CaCl<sub>2</sub>) shift the equilibrium from a closed to an open, all-*trans* conformation (compact to more extended structures), whereas perchlorate salts induce bent structures (compact to even more compact structures).

It is well documented that CycA increases cellular uptake of Ca<sup>2+</sup> <sup>66-69</sup> and, in some instances, is shown to increase intracellular Ca<sup>2+</sup> levels.<sup>70, 71</sup> We speculate two mechanisms relating to the interactions of CycA and Ca<sup>2+</sup> that may explain this phenomenon. First, CycA may act as a Ca<sup>2+</sup> transmembrane “carrier” (i.e., an ionophore<sup>32</sup>). Previous studies<sup>72-74</sup> suggested that a transition from the open to the closed form is critical for CycA to pass through the cell membrane. However, our structural characterization indicates a more complex picture. It would be of great

interest to determine whether the all-*trans*, open CycA:Ca<sup>2+</sup> would undergo a structural change similar to CycA/Dy<sup>3+</sup> in SDS micelles<sup>65</sup> in order to pass through the cell membrane. Of note, our IMS–MS data show that Ca<sup>2+</sup> also binds to the bent and closed conformations of CycA (Figure 2.5C). Complexing with metals can be another means for CycA to cross the membrane in which the chameleonic nature of the metal complex makes the process more efficient. CycA can then release the Ca<sup>2+</sup> passenger upon binding with CypA.

Second, our data and those of others<sup>24, 25</sup> show that Ca<sup>2+</sup> has high binding affinity toward CycA. Once free CycA enters the cell, it will complex with intracellular Ca<sup>2+</sup>, which would (a) decrease the concentration of free Ca<sup>2+</sup>, leading to an influx of extracellular Ca<sup>2+</sup>, and (b) affect the activity of PPlase via binding to CypA. To support the latter, we need to address whether Ca<sup>2+</sup> influences the abundance and structure of CycA:CypA complexes.

#### **IMS–MS data show that Ca<sup>2+</sup> leads to more effective CycA:CypA complexation and structural changes of the protein**

Before studying the interactions between CycA and its receptor, we first performed a brief characterization of CypA (Figure 2.7A). CypA has a molecular mass of approximately 18.2 kDa, and its mass spectrum in water shows several charge states with  $z = +10$  being the lowest charge state observed. In 20 mM ammonium acetate with pH = 6.8, charge states as low as +7 can be observed, consistent with the work by Barran and co-workers.<sup>75</sup> The plot of CCSs versus charge state and comparison to theoretical CCSs of CypA X-ray structures (PDB IDs 5LUD,<sup>76</sup> 1AWR,<sup>77</sup> and 1AWV<sup>77</sup>) show that the protein retains its native state at  $z = +7$  and  $+8$ . The large increase seen in charge states from  $z = +8$  to  $z = +9$  indicates the unfolding of the native CypA structure as extra charge states accrue, a phenomenon that is well documented for many proteins.<sup>78-80</sup>

To explore the binding of CypA and CycA, mixtures consisting of both species were prepared and analyzed subsequently with IMS–MS. At a 1:1 ratio, there were no signs of

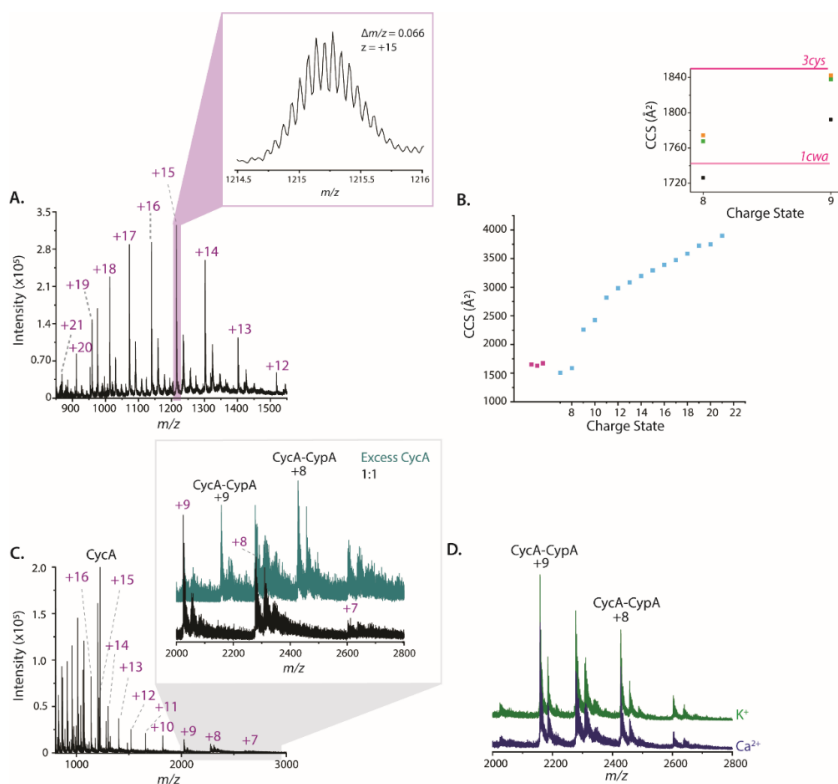


**Figure 2.7. (A)** Representative mass spectrum of CypA in H<sub>2</sub>O (10 μM) with the charge state of each species labeled. The inset shows isotopic spacing of CypA with  $z = +15$ . **(B)** CCS of each ion species with its respective charge state for CypA shown in blue. The pink data points correlate to the theoretical CCS of three CypA structures (PDB IDs 5LUD,<sup>76</sup> 1AWR,<sup>77</sup> and 1AWV<sup>77</sup>). The inset shows the CCS of  $z = +8$  and  $+9$  of CycA:CypA, and the pink lines show the theoretical CCSs of the X-ray crystal structure (PDB 1CWA<sup>26</sup>) and NMR structure (PDB 3CYS<sup>81</sup>) of the complex. The black data points are for the CycA:CypA mixture with no metals added, the green points are the mixture with potassium, and the orange points are the mixture with calcium. **(C)** 1:1 mixture of CycA:CypA shown by a larger mass spectrum (black) with the major CycA peak labeled and the charge states for each CypA ion labeled. Teal mass spectrum within inset is the CycA:CypA mixture when CycA is in excess. **(D)** Mass spectra of the CycA:CypA mixture (5:1 ratio) in the presence of Ca<sup>2+</sup> (blue spectrum) and K<sup>+</sup> (green spectrum).

complexation of the two species in the designated mass range, as shown by the black mass spectrum in Figure 2.7C.

Once CycA is present in excess within the CycA:CypA mixture, two CycA:CypA complexes can be observed (teal mass spectrum, Figure 2.7C). Interestingly, the CycA:CypA complexes were detected in low charge states, suggesting that CycA selectively binds to the folded CypA ion. Barran and co-workers<sup>75</sup> observed these same low charge states (from  $z = +7$  to  $+9$ ) of the CycA:CypA complexes with nanoESI-MS. This, along with the CCS data of free CypA versus charge state (Figure 2.7B), helps support the notion that the CycA:CypA complexes observed in IMS-MS closely resemble the complexes in solution. Previous studies have shown two crystal structures of CycA:CypA complexes: one contains a decameric CycA:CypA complex,<sup>82</sup> and the other shows a monomeric CycA:CypA complex.<sup>26</sup> In the monomeric structure, the binding of the two species is stabilized by hydrogen bonds mediated by solvent (water) contacts.<sup>26</sup> The binding between CycA and CypA is unique due to the conformational space that CycA can exploit, which allows binding to a receptor binding groove rather than within a binding pocket.<sup>6</sup> This subtle bonding was strong enough to be detected via IMS-MS, while still observing CypA at low charge states.

The experimental CCS and the theoretical CCS of the monomeric CycA:CypA complex (PDB 1CWA<sup>26</sup>) are in good agreement with each other as they have percent differences of  $-0.3$  and  $3.5\%$  for  $z = +8$  and  $z = +9$ , respectively. When CycA binds to CypA, the intramolecular hydrogen bonds are replaced by intermolecular interactions with the target, and CycA assumes the “open” conformation.<sup>59</sup> Thus, based on CCS comparison, CycA binds to CypA in its open form.

Next, to evaluate the possibility that  $\text{Ca}^{2+}$  would enhance the binding of CycA to CypA via its  $\text{Ca}^{2+}$ :CycA complex, we prepared samples of the CycA:CypA mixture (in 200 mM ammonium acetate, pH 6.8) in the presence of approximately 70  $\mu\text{M}$   $\text{Ca}^{2+}$  and  $\text{K}^+$ . The final ratio of the CycA:CypA/metal mixture is approximately 5:1:70. Additional instrument tuning was performed to yield qualitatively stronger CycA:CypA peaks.

However, there is no evidence of CycA:CypA:Ca<sup>2+</sup>:K<sup>+</sup> in the mass spectra.

Nonetheless, the IMS data (see Table 2.2) at the  $m/z$  values corresponding to the CycA:CypA complexes at  $z = +8$  and  $+9$  show that Ca<sup>2+</sup> and K<sup>+</sup> induce more extended structures of the complexes. The  $z = +8$  complex has an experimental CCS of 1726 Å<sup>2</sup>, which is similar to the theoretical CCS of the crystal structure (PDB ID 1CWA;<sup>26</sup> 1731 Å<sup>2</sup>). In the presence of Ca<sup>2+</sup> and K<sup>+</sup>, the CCSs increase to 1774 and 1768 Å<sup>2</sup>, respectively. At the  $z = +9$  charge state, the CCSs increase from 1792 Å<sup>2</sup> (no metal cations) to 1842 Å<sup>2</sup> (Ca<sup>2+</sup>) and 1838 Å<sup>2</sup> (K<sup>+</sup>). Our results suggest that the open conformation of the Ca<sup>2+</sup>:CycA complex may act as an intermediate to facilitate the strong and perhaps slightly different binding mode of CycA to CypA. This is an intriguing finding, indicating the chameleonic roles of Ca<sup>2+</sup> in CycA chemistry. First, it induces compact conformations that may promote membrane permeability of CycA. Then, when CycA is inside the cell, Ca<sup>2+</sup> can induce an all-*trans* CycA conformation that is optimal for CypA interactions.

## 2.6 Conclusion

In this study, we investigated the conformational space of bare and Ca<sup>2+</sup>-bound CycA. IMS–MS results unravel a unique set of conformations with at least two *cis*-amides, as shown by DG modeling. The *cis*-amides include, but are not limited to, MeLeu9–MeLeu10 and Sar3–MeLeu4. The set of conformations containing the *cis* Sar3–MeLeu4 amide was previously observed as the minor conformation in chloroform and when Mg<sup>2+</sup> binds to CycA.<sup>24</sup> A combination of IMS–MS and NMR experiments shows that Ca<sup>2+</sup>, which has a high affinity to CycA, induces two compact (bent and closed) conformations and one open, all-*trans* conformation. Interestingly, the open conformation of CycA:Ca<sup>2+</sup>, which has all-*trans* amides, is a dominant species in NMR experiments, possibly serving as an intermediate between the free (closed form) and bound (open form) forms of CycA. This particular conformation was not observed in previous NMR studies where Ca(ClO<sub>4</sub>)<sub>2</sub> was used.<sup>25</sup> Chloride salts, but not perchlorate salts, convert the closed conformation into an all-*trans*, open conformation. Our study illustrates the

important role of salt anions and metal cations in shaping the chemistry of macrocyclic molecules.

We rationalize the discrepancies from the previous studies regarding the types of CycA conformations complexing with metals. Until this work, what exist in the literature are independent results, and there is neither connection nor discussion on the reason(s) why different structures (open, closed, and bent) were obtained experimentally.

Last, we demonstrate that IMS–MS provides a robust and reliable method to quickly and simultaneously sample conformations of macrocyclic peptides and their interactions with guest ligands and protein targets. We demonstrate excellent agreement between gas-phase IMS–MS and solution-phase NMR and how they complement each other to provide a full description of the complex conformational space of macrocyclic molecules.

As an extension to this investigation of the conformational “dark space” of CycA, my lab group members, Aleksandra Antevska and Miranda Limbach, have worked on obtaining the structures of aqueous CycA via X-ray crystallography, NMR spectroscopy, and single-crystal neutron diffraction in aqueous environments. As mentioned previously, CycA exhibits very poor water solubility, so the predominant aqueous structure(s) of free CycA are currently a mystery. Their results show a unique conformer of CycA containing one *cis*-amide bond between residues 11 and 1, which is the main conformer in predominantly aqueous environment. Additional conformers were observed in 90/10 methanol/water. These recently discovered conformations of CycA may be relevant to help explain the binding process of CycA to protein targets, such as CypA.<sup>83</sup>

### **3 Cytotoxicity of $\alpha$ -Helical, Staphylococcus aureus PSM $\alpha$ 3 investigated by post-ion-mobility dissociation mass spectrometry**

#### **3.1 Disclaimer**

The following chapter has been adapted, with permission, from a journal article published in *Analytical Chemistry* (Copyright 2020 American Chemical Society).<sup>84</sup> Only minor modifications were made. I collected the IMS-MS data and was involved in the data analysis under the guidance of Thanh Do. This work was performed in conjunction with Professor Kristi L. Lazar Cantrell, where she aided in the synthesis of the PSM $\alpha$ 3 peptides and performed circular dichroism (CD) spectroscopy.

#### **3.2 Abstract**

Our knowledge of amyloid formation and cytotoxicity originating from self-assembly of  $\alpha$ -helical peptides is incomplete. PSM $\alpha$ 3 (where PSM stands for phenol-soluble modulins) is the only system where high-resolution X-ray crystallography and toxicity data are available. Oligomers of multiple  $\alpha$ -helical monomers are less stable than those of  $\beta$ -strands, partially due to the lack of a consistent hydrogen-bonding network. It is challenging to preserve such oligomers in the gas phase where mass-selected structural studies using ion-mobility spectrometry mass spectrometry (IMS-MS) could be performed. As the oligomers fall apart after exiting the drift cell of the mass spectrometer, novel features that have shorter (a loss of charged species) or longer (a loss of neutral species) arrival times than expected are present together with those from the intact species. By obtaining a complete data set of PSM $\alpha$ 3 peptides in solution and with *n*-dodecyl- $\beta$ -D-maltoside, a micelle-forming detergent, we are able to discern the dissociated from the intact oligomers and detergent-bound complexes and correlate the

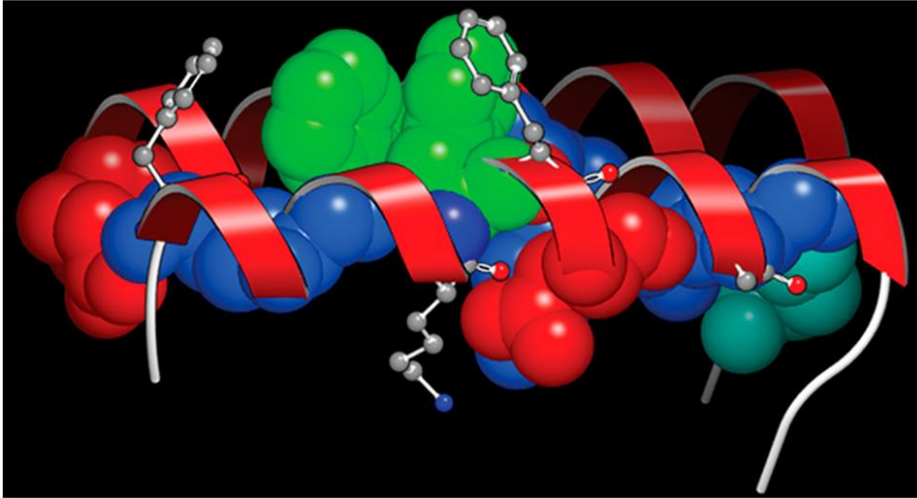
reported cytotoxicity to the peptide oligomeric structures and their interactions with membrane mimetics. The study sheds new insights into the interpretation of IMS-MS data from biomolecular self-assembly studies—an important and timely topic.

### 3.3 Introduction

Bacteria constitute a large domain of microorganisms found in a broad gamut of habitats including living organisms, metal surfaces, and radioactive waste. As a result, bacteria can influence human health, ecosystems, and industrial processes. Recently, functional amyloids have been observed in microbial biofilms produced from a wide variety of bacteria (e.g., *Pseudomonas* sp, *Bacillus subtilis*, and *Klebsiella pneumoniae*).<sup>85-90</sup> The bacterial amyloids serve a variety of functions in biofilm growth,<sup>91-94</sup> interaction with host tissues,<sup>95-97</sup> evasion of immune systems,<sup>98</sup> and aiding persistence in environment niches. Interestingly, bacterial amyloids, which extend laterally and entangle with each other, can provide a “railroad” for quorum-sensing metabolites to traffic through the extracellular matrix,<sup>99, 100</sup> and thus they can play an instrumental role in intercellular communication. The tensile strength of amyloid fibrils also aid the mechanical properties of the biofilm.<sup>94</sup>

In 2017, X-ray crystallography studies of the WT PSM $\alpha$ 3 peptide (Figure 3.1) in *Staphylococcus aureus* by Landau and co-workers<sup>101</sup> showed that the fibril core was made up of parallel  $\alpha$ -helix “rippled” sheets instead of  $\beta$ -strand flat sheets found in typical pathogenic amyloids. They showed that PSM $\alpha$ 3 assembly to fibrils helped the bacteria kill other cells. Mutations that disrupt the  $\alpha$ -helix also alleviate cytotoxicity (i.e., K9P/F11P), whereas other benign mutations retain similar toxic effects (i.e., F3A and G16A). These unexpected findings appear to contradict the general notion of human amyloid toxicity built upon  $\beta$ -sheet structures. Therefore, it has become apparent that our knowledge of amyloids and their toxicity is astonishingly incomplete.

The contribution of  $\alpha$ -helices to fibril formation (or well-ordered nanostructures in general) has been previously reported in our past work<sup>102, 103</sup> and by others.<sup>104-106</sup> For



|           |                        |
|-----------|------------------------|
| WT:       | MEFVAKLFKFFKDLLGKFLGNN |
| F3A:      | MEAVAKLFKFFKDLLGKFLGNN |
| G16A:     | MEFVAKLFKFFKDLLAKFLGNN |
| K9P/F11P: | MEFVAKLFPFPKDLLGKFLGNN |

Figure 3.1. (Top) An image of a PSM $\alpha$ 3 WT dimer generated from the X-ray crystal structure (PDB ID: 5i55). The four mutated residues (F3, K9, F11, and G16) are shown in ball-and-stick form. The residues that contribute to stabilization of the dimer are shown in space-filling models (E2, K6, F10, K12, D14, K17, and N22). (Bottom) Primary sequences of WT, F3A, G16A, and K9P/F11P peptides.

example, we have shown that an Apolipoprotein mimetic helix-turn-helix peptide can self-assemble into fibrils that shared the morphological characteristics of canonical cross- $\beta$  fibrils.<sup>102</sup> Circular dichroism (CD) measurements indicate no structural changes from  $\alpha$ -helix upon assembly and further analysis with IMS-MS showed that stabilization in the turn region promotes aggregation.<sup>102</sup> Another compelling piece of evidence indicating that  $\alpha$ -helical peptides can aggregate into fibrils is from a recent study<sup>103</sup> of customized peptides containing so-called chameleon sequences. Isolated chameleon sequences are intrinsically disordered, interconverting between  $\alpha$ -helix and  $\beta$ -strand conformations. The inherent conformational plasticity of the sequence can be constrained by the addition of flanking residues with a given secondary structure propensity. Shea and co-workers showed that the chameleon sequence with helical flanking residues aggregates rapidly into fibrils, whereas the chameleon sequence with flanking residues that favor  $\beta$ -conformations has weak aggregation propensity.<sup>103</sup> Unlike systems where conformational transitions occur from intrinsically disordered to  $\alpha$ -helical intermediates en route to  $\beta$ -sheet amyloid formation,<sup>107, 108</sup> the chameleon, PSM $\alpha$ 3, and a few other systems such as tau paired helical filaments<sup>109</sup> uniquely retained their  $\alpha$ -helical structures before and during their aggregation process evidenced by time-course CD data. The  $\alpha$ -helical monomers are also part of the fibrillar structures as shown by the CD and X-ray data of PSM $\alpha$ 3.<sup>101, 103</sup> Therefore, the aggregation of bacterial PSM $\alpha$ 3 into  $\alpha$ -helical fibrils is likely not an exception and other proteins can adopt similar amyloid structures. Exploration of this new class of amyloids will undoubtedly refine our prior knowledge on amyloids and their toxicity and introduce new structural classes of antibiotics.

### 3.4 Materials and methods

#### Ion mobility spectrometry – mass spectrometry

Ion mobility – mass spectrometry experiments were performed on an Agilent 6560 IMS-MS QTOF mass spectrometer. Lyophilized peptides were pretreated with hexafluoro-2-propanol (HFIP) and left in a dessicator to evaporate. Prior to analysis, the peptides were reconstituted in HPLC-H<sub>2</sub>O, diluted to approximately 20  $\mu$ M in a solution of 20 mM ammonium acetate, and directly injected into the mass spectrometer via syringe. Peptides are ionized by a jet-stream ESI source and ions are subsequently pulsed into a helium-filled drift cell in a multi-field experiment. All data was collected in positive polarity. The pressure of the drift tube is 3.940 Torr with the pressure differences between the drift tube and trap funnel being approximately 300 mTorr. Data was collected over the course of 5.2 minutes with  $\Delta V = 890, 790, 690, 590$ , and 490V. An exit funnel and hexapole ion guide focus the ions into the QTOF mass spectrometer. Arrival time distributions (ATD) and the resulting mass spectra were extracted with the Agilent IM-MS Browser software and graphed with Origin Pro. For each ATD, mobility features were assigned as discussed in the main text. Collisional cross section values were calculated in the same manner as discussed in Section 2.4.

#### Far-UV CD

CD was used to assess the secondary structure of WT PSM $\alpha$ 3 and the mutant peptides. CD data were collected on a Jasco J-810 CD spectrometer equipped with a temperature-controlled holder. The CD spectra were collected at 1 nm intervals from 260 to 190 nm with a 4 s response time and a 1 nm bandwidth. All spectra were collected in a 1 mm quartz cuvette at 295 K. The CD spectra of freshly dissolved ( $t = 0$ ) peptides were measured at a concentration of 50  $\mu$ M in 20 mM ammonium acetate, pH 6.8. The samples were incubated at room temperature and run again at  $t = 1$  week and  $t = 2$  week time points. 10 scans were taken for each peptide sample, and the data were

averaged. A buffer baseline was subtracted from the averaged data. Simultaneous measurements of CD signal and absorbance were taken to set the peptide concentration at 50  $\mu$ M. Since the peptides do not contain tryptophan or tyrosine the following procedure was followed. The concentration of peptide in solution,  $C$  (M), was calculated with the Beer-Lambert equation for each sample at 205 nm (Equation 5.1),

$$A_{205} = \epsilon_{205} \times l \times C \quad 5.1$$

The path length of the cuvette,  $l$ , was set to 0.1 cm.  $A_{205}$  was calculated by averaging the buffer corrected absorbance values of the peptide sample in solution across 10 scans at 205 nm. The molar extinction coefficient,  $\epsilon_{205}$  ( $M^{-1} \times cm^{-1}$ ), specific to each peptide sequence, was calculated by a protein calculation webserver (<https://spin.niddk.nih.gov/clore/>).<sup>110</sup>

### Peptide synthesis

WT PSM $\alpha$ 3 and mutant peptides (F3A, G16A, K9P/F11P) were synthesized using standard tert-butyloxycarbonyl (*t*-Boc) and hexafluorophosphate benzotriazole tetramethyl uronium (HBTU) manual solid-phase peptide synthesis on a 4-methylbenzhydryl amine resin. The peptides were cleaved from the resin at 0°C for 1 h in anhydrous HF containing 5% *p*-cresol followed by standard work-up.

Crude peptides were purified by reverse phase-high performance liquid chromatography (RP-HPLC) on a semipreparative C18 column (Phenomenex) using gradients of water [0.1% (v/v) TFA] and acetonitrile [0.1% (v/v) TFA]. Peptide purity was > 95% as determined by analytical RP-HPLC. The molecular masses of the peptides were verified by Electrospray Ionization (ESI) mass spectrometry.

### 3.5 Results and discussion

#### Losses of charged peptides from PSM $\alpha$ 3 oligomers due to post-IM dissociation: problem and solution

Figure 3.2 shows the ESI mass spectra of the four peptides at 20  $\mu$ M in 20 mM ammonium acetate (pH 6.8). Each mass spectral peak is annotated with a nominal oligomer size-to-charge ratio ( $n/z$ ). At first glance, there is no significant difference among the four spectra, leading us to further evaluate the ion-mobility data. Figure 3.3A,B shows the mobility data of F3A, and the same data for other peptides are shown in the Figure 3.4. In panel A of Figure 3.3, the arrival time distributions (ATDs) corresponding to the nominal  $n/z = 2/3, 3/4, 4/5$ , and  $5/6$  are shown. Each ATD contains multiple features, which could be assigned to either different conformers or oligomers. In panel B of Figure 3.3, the 2D plot of  $m/z$  vs drift time for  $n/z = 2/3$  ( $m/z$  1687) is shown. Given the isotopic spacing of 0.33  $m/z$ , the charge states for all these species are thought to be +3. However, the short arrival time features ( $t < \sim 24$  ms) cannot be dimers because their CCSs would be less than 700  $\text{\AA}^2$ , too small for a dimer given that the smallest monomer is 476  $\text{\AA}^2$  and the smallest possible dimer should be 757  $\text{\AA}^2$ .<sup>(26)</sup> Furthermore, the CD spectrum of F3A (as well as WT and G16A) consistently shows  $\alpha$ -helical signature over a period of 2 weeks (Figure 3.5), and the fibrils of WT and G16A were made of  $\alpha$ -helical subunits.<sup>101</sup> When taken together, the data indicate that unfolded and partially folded conformers are not abundant.

Another possibility, which we have reported,<sup>111</sup> is the post-IM dissociation of large oligomers,<sup>112-114</sup> but in the absence of activation voltages. The loss of a neutral species in post-IM dissociation was recently reported in Link et al., where we looked at the self-assembly of small macrocycles.<sup>34</sup> The feature corresponding to the post-IM dissociated species had a much longer arrival time than the intact species; at the same time, the arrival time was similar to that of its precursor ion. In QTOF instruments with a collision cell, there must be gas in this region to achieve reasonable ion transmission.

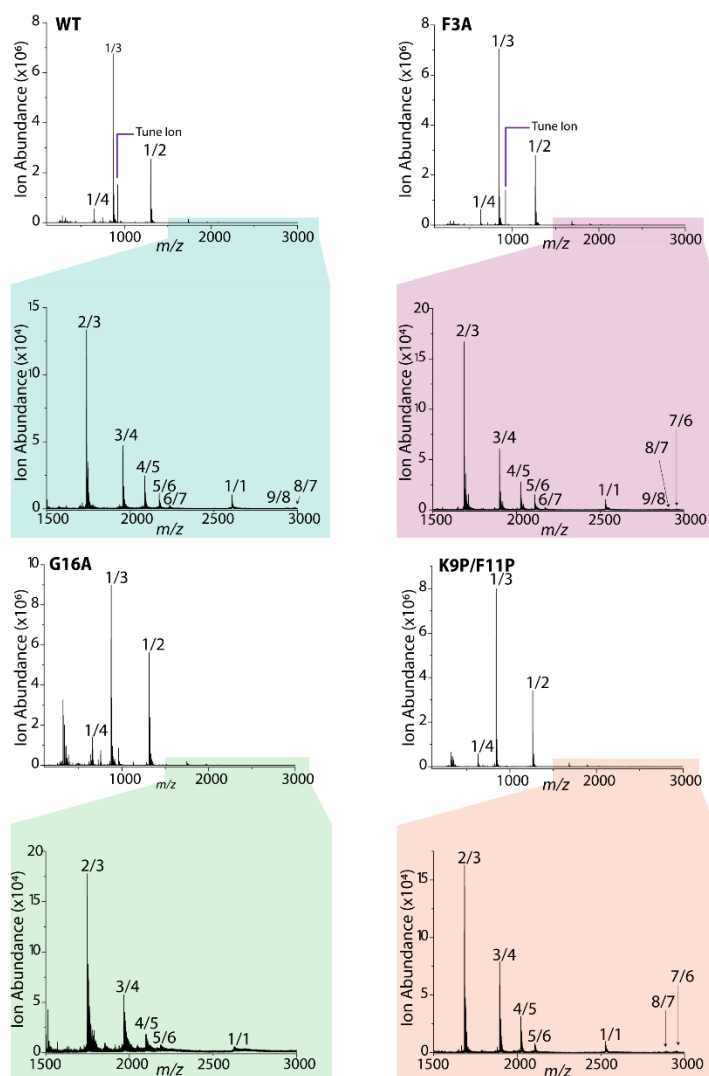


**Figure 3.2. Positive-mode ESI mass spectra of PSMα3 peptides in 20 mM ammonium acetate. The peptide mass spectral peaks are labeled by their nominal oligomer-to-charge ratio ( $n/z$ ). The oligomers formed by these peptides are similar in order, thus the four mass spectra are very similar. The peak labeled “tune ion” present in the WT and F3A mass spectra correspond to an ion present in a reference solution that is prepared from the Agilent API-TOF Reference Mass Solution Kit.**

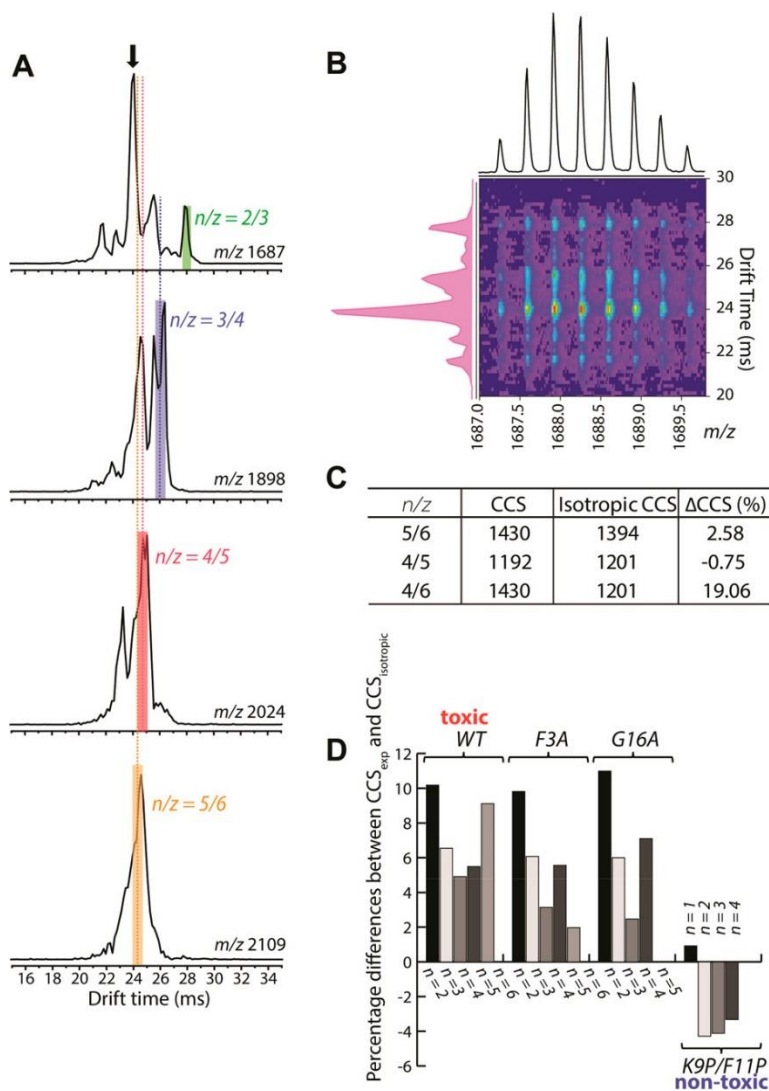


**Figure 3.3.** (A) Representative ATDs of F3A oligomers corresponding to the nominal  $n/z = 2/3, 3/4, 4/5$ , and  $5/6$ . (B) A 2D plot of  $m/z$  vs drift time for  $n/z = 2/3$ . (C) Possible assignments for the feature at  $\sim 24$  ms in the ATD of  $n/z = 2/3$ . (D) Deviation of experimental CCSs from isotropic predictions.

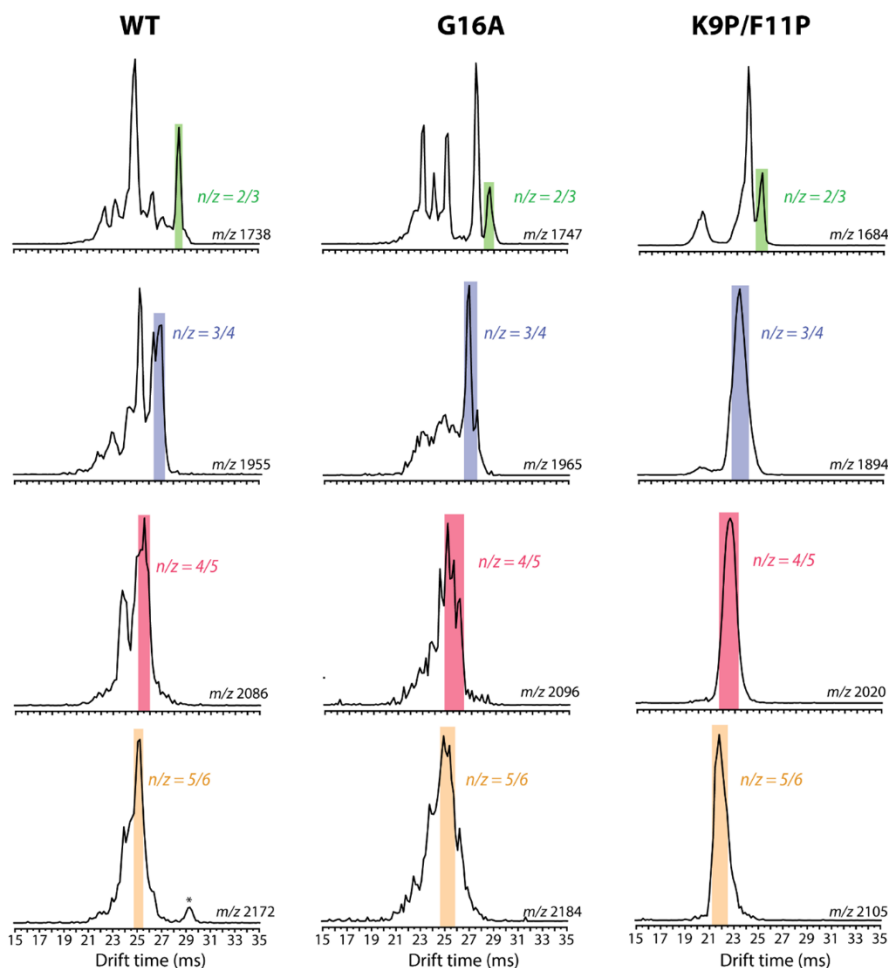


**Figure 3.4.** Ion-mobility data for WT, G16A, and K9P/F11P at  $n/z = 2/3$ ,  $3/4$ ,  $4/5$ , and  $5/6$ . A fewer number of features are resolved as oligomer order increases. Only the feature with the longest arrival time (the right most peak) is assigned in our analysis due to uncertainty in the identity of other features. Note that for K9P/F11P this trend deviates slightly in  $n/z = 4/5$ . The secondary structure of this peptide is shown to not have an  $\alpha$ -helix conformation, so oligomers of this species may be intrinsically disordered and therefore may behave differently from other helical peptides in the gas phase. The asterisk for WT  $n/z = 5/6$  denotes a feature that arises from overlapping isotopic spacing from a separate species.

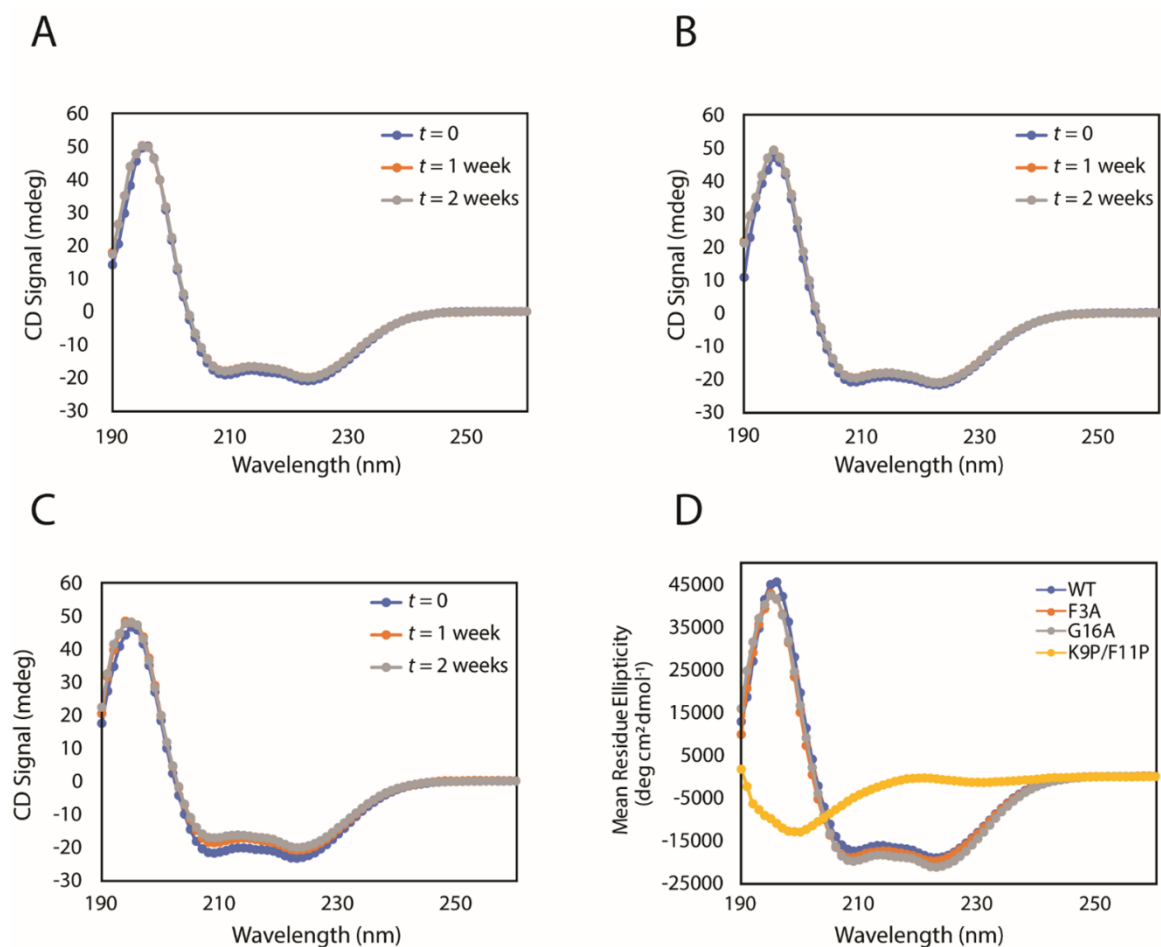


**Figure 3.5.** CD time course spectra of 50  $\mu$ M (A) WT, (B) F3A, and (C) G16A in 20 mM ammonium acetate, pH 6.8, shows that each peptide retains its  $\alpha$ -helicity over a two week time-period. (B) CD spectra of 50  $\mu$ M freshly dissolved PSM $\alpha$ 3 peptides in 20 mM ammonium acetate, pH 6.8. These spectra show that WT, F3A, and G16A adopt an  $\alpha$ -helical secondary structure in solution, whilst K9P/F11P does not exhibit this characteristic.

Dissociations of noncovalent complexes may occur at this stage. To overcome this issue, instrument modification is required,<sup>115</sup> which goes beyond the scope of this paper. We also note that the dissociations can also occur in an IM-quadrupole instrument (with no collision cell), as previously reported for hydrated diphenylalanine oligomers, evidently suggesting that weak interactions can be disrupted in the absence of a collisional cell.<sup>111</sup> Therefore, dissociation may happen anywhere between the exit funnel and the TOF detector.

Post-IM dissociation is an undesired event that occurs after the ions successfully traverse the drift tube but before they reach the detector. For example, an oligomer with  $n/z = 3/4$ , which travels as a +4 charge state species, could be dissociated into a triply charged dimer and a singly charged monomer after exiting the drift tube. The dissociated +4 trimer will be detected as a +3 dimer and a +1 monomer. The trimer will have a shorter arrival time than the intact +3 dimer because the effect of charges (+4 vs +3) always outweighs the effect of size, but both species would appear in the mass spectrum as triply charged dimers. At the same time, that “unexpected” dimer with +3 charges would have the same drift time (less than 0.1 ms difference) with the “intact” trimer with +4 charges.

Overall, the data shown in Figure 3.3A,B suggest that the loss(es) of charged peptides in post-IM dissociation provides a reasonable explanation for the detection of multiple features in the ATDs of the oligomers. Features at short arrival times in ATDs of small oligomers (e.g., those of  $n/z = 2/3$  with  $t < 28$  ms; Figure 3.3A) have similar drift times to those in the ATDs of larger oligomers, and thus are likely the results of post-IM dissociations. Only the longest drift time features, which are exclusive to each oligomer size and the least likely to overlap with other features, can be conservatively assigned to the smallest possible oligomers based on  $m/z$  and isotope spacing, as is consistent with common practices in IMS-MS measurements. For a large oligomer to have a similar drift time as the feature with the longest arrival time in a small oligomer such as  $n/z = 2/3$ , the large oligomer must be a low-charge-state species (in this case, they must have

fewer than one charge per peptide unit). It is well-established that low-charge species are stable because of fewer Coloumbic repulsions; thus, such stable species would be less prone to dissociating into small oligomers with more than one charge per peptide. There are three lines of supporting evidence. First, previous work, namely, those by Williams on supercharging agents<sup>116</sup> and by Bush on cation-to-anion proton-transfer reactions,<sup>117</sup> established the relationship between conformation and charge state. Higher charge state species tend to be more extended (unfolded) than lower charge state species (folded). Second, while dissociation of large biomolecular complexes (e.g., multimeric proteins) often occurs asymmetrically in terms of the charge with monomers often carrying half of the original charge,<sup>118-121</sup> we detected no monomer containing features whose arrival times overlap with the oligomers reported in Figure 3.3A that are consistent with this mechanism. Specifically, if the dissociations followed this channel, the most plausible monomers that could be post-IM products of the oligomers would have  $z = +3$  and  $+4$ . However, their arrival times are approximately 16 and 14 ms, respectively, which are much shorter than any features reported in Figure 3.3A. On the other hand, the ATDs of  $z = +1$  monomer showed multiple features with different isotope spacings and overlapping arrival times with the oligomers reported in Figure 3.3 (see Figure 3.6 showing the ATD of the WT as an example). Collectively, our data indicate that *asymmetric* charge partition upon dissociation did not occur at these oligomer levels and that the preferred pathway is via consecutive losses of singly charged monomers.

Finally, as the oligomers grow in size during aggregation, the ATDs of large oligomers exhibit narrower distributions and a fewer number of features than those of the smaller ones, which is consistent with larger oligomers being more stable than smaller ones. Since the ATD reflects the conformational distribution, a broad ATD would imply that many families of structures are present while a narrow ATD means that there are fewer structures, thus indicating the latter species are stable. Furthermore, in the context of

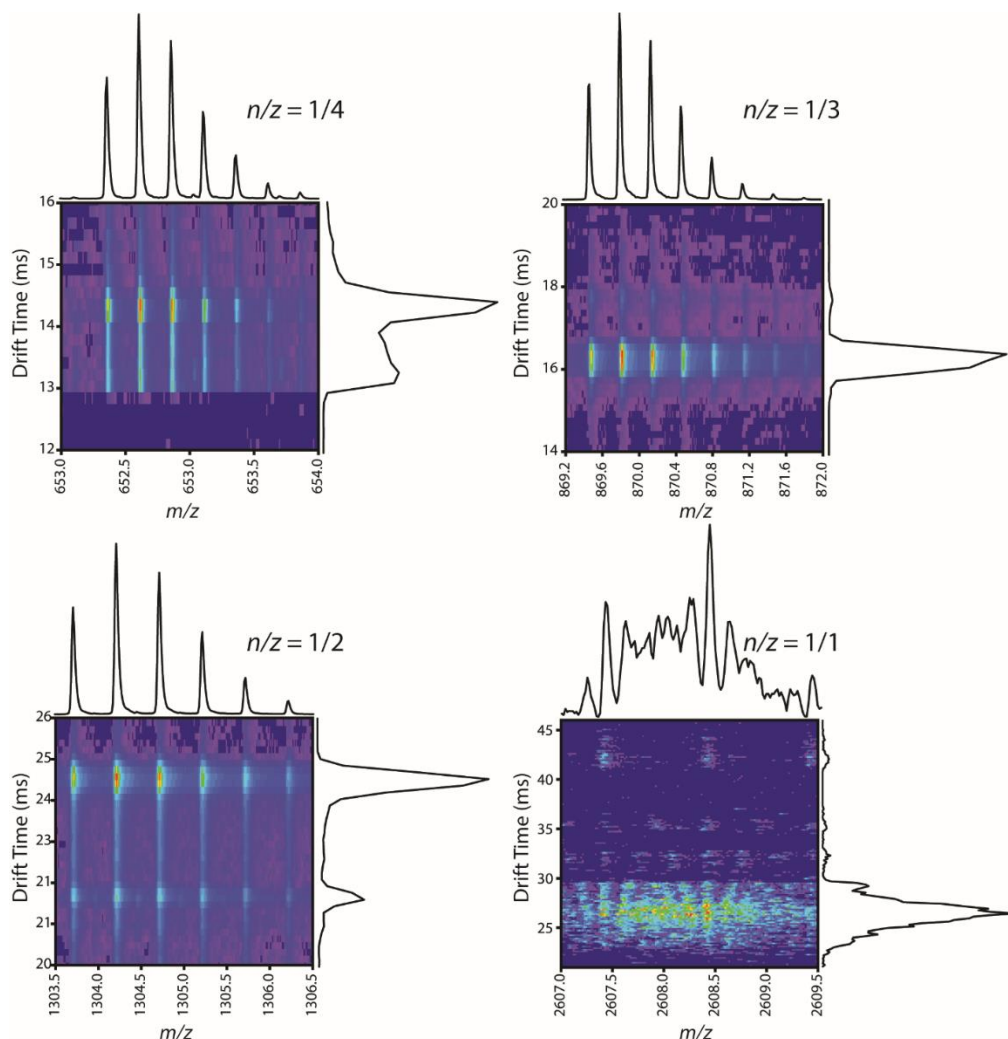


**Figure 3.6.** 2D plots of  $n/z = 1/4$ ,  $1/3$ ,  $1/2$ , and  $1/1$  for WT. Note that at the same drift voltage, there is no overlap between the ATDs of  $n/z = 1/4$  and  $1/3$  with the ATDs of WT oligomers shown in Figure 3.4. On the other hand, the ATDs of nominal  $n/z = 1/2$  and  $n/z = 1/1$  overlap with the ATDs in Figure 3.4. The same observations can be made for other peptides.

protein self-assembly, once the oligomers reach a certain size, it becomes stable in which dissociation is weaker than elongation or nucleation.<sup>122, 123</sup>

In short, careful consideration has to be made to accurately assign the features in complicated ATDs. In other words, if the ATDs of  $n/z = 3/4$ ,  $5/6$ , or  $6/7$  and isotope spacing information is not available, one may be tempted to assign the feature at  $\sim 24$  ms in the ATD of nominal  $n/z = 2/3$  (Figure 3.3A; labeled with a black arrow) as a tetramer with +6 charge (i.e.,  $n/z = 2/3 = 4/6$ ). This would yield an experimental CCS of  $1430 \text{ \AA}^2$ , which is  $\sim 19\%$  larger than the isotropic prediction ( $\sigma_n = \sigma_1 \times n^{2/3}$ ). (26,41) In this isotropic model, we used the experimental CCS of a triply charged monomer ( $n/z = 1/3$ ) whose CCS agrees well with that of an  $\alpha$ -helical conformer as the starting point. Typically, features are assigned the same nominal  $n/z$  for a specific ATD, but given the ATDs of higher mass species, the features can also be assigned to be a tetramer with a +5 charge state (CCS =  $1192 \text{ \AA}^2$ ;  $-0.8\%$  difference) or a pentamer with a +6 charge state (CCS =  $1430 \text{ \AA}^2$ ;  $+2.6\%$  difference). While both are reasonable assignments, there is not enough evidence to support one being better than the other. Thus, *such features are not assigned in our analysis because of this ambiguity.*

Figure 3.3D shows the CCS differences between the experimental CCSs and the predicted values using the isotropic models. Interestingly, the F3A and G16A mutants, which possess the same  $\alpha$ -helical structures as the WT, also show a similar pattern to the WT. The K9P/F11P mutant shows negative deviations, indicating this peptide behaves drastically different from the other three. In the original study by Landau and co-workers, they showed that K9P/F11P has the least cytotoxicity, followed by F3A, and then G16A and WT.<sup>101</sup>

**Losses of neutral *n*-Dodecyl- $\beta$ -D-maltoside (DDM) molecules from Peptide:DDM heteroligomers due to post-IM dissociation: A strong correlation between toxicity and the number of DDM molecules required to stabilize the complexes**

It is well-established that not all  $\alpha$ -helical peptides aggregate, and in fact, our study shows that the four peptides did not aggregate into fibrils under quiescent incubation at room temperature for 2 weeks. Landau and co-workers observed only fibril formation in 1 mM sample solutions after 1 week of constant shaking.<sup>101</sup> On the basis of our IMS-MS and CD data (Figure 3.2 and Figure 3.5), all four peptides formed oligomers in solution, in which WT, F3A, and G16A retained  $\alpha$ -helical structures. To further investigate the relationship between structures and cytotoxicity, we use DDM micelles to mimic the cellular membrane. DDM is a nonionic detergent that rapidly forms micelles in solution. It was one of the first detergents used in native MS, pioneered by the Robinson group, to study transmembrane proteins.<sup>124</sup> Given that we understand how to interpret post-IM dissociation data, the interactions between the peptides and DDM can be elucidated by determining the smallest number of DDM molecules required to stabilize a peptide:DDM complex.

Figure 3.7 shows the data obtained from WT or K9P/F11P peptides ( $\sim 5\text{--}15\ \mu\text{M}$ ) in  $20\ \mu\text{M}$  ammonium acetate, pH 6.8, where the DDM concentration is kept at  $200\ \mu\text{M}$ . In the mass spectra of the DDM-containing samples (Figure 3.8), we observed a series of peaks corresponding to the complexes of the peptide monomer (and sometimes dimer) with multiple DDM molecules. We compared the peaks in these mass spectra to those in the mass spectrum of DDM only to ensure they are not of DDM clusters. In Figure 3.9, we showed the ATDs of F3A and G16A collected from samples with and without DDM. In all cases, DDM complexation leads to the presence of multiple features at longer arrival times than those obtained in the absence of the detergent (panel A of Figure 3.9). As stated above, the experimental CCSs of  $n/z = 1/3$  for WT, F3A, and G16A without DDM ( $t_A = \sim 16\ \text{ms}$ ) are similar to the theoretical CCS of the X-ray  $\alpha$ -helical structure; therefore, those features at longer arrival times (see the first ATDs in Figure 3.7A,B) are

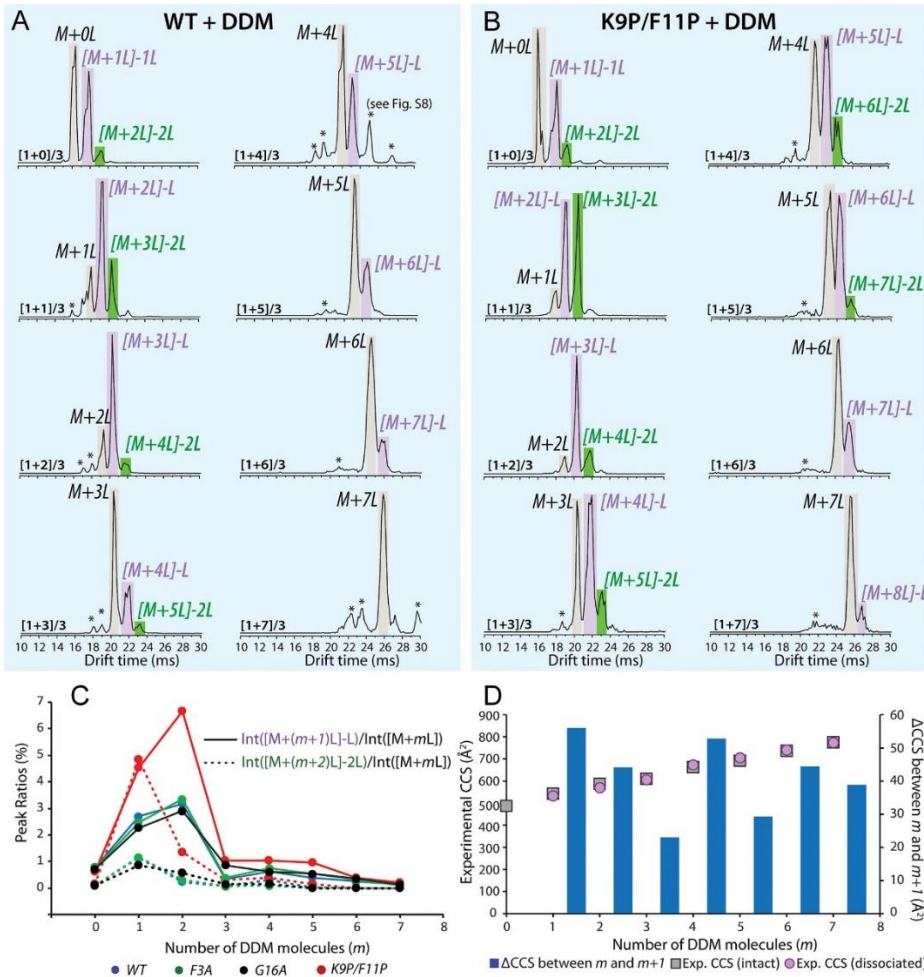


**Figure 3.7. (A, B) Representative ATDs of peptide:DDM complexes corresponding to the nominal  $[1+m]/z$  ratios where “1” denotes peptide monomer,  $m$  denotes the number of DDM molecules ( $m = 0-7$ ), and denotes  $z$  the charge ( $z = 3$ ). Features are annotated as intact (gray), a loss of one DDM (purple), and a loss of two DDM (green). “M” denotes peptide monomer and “L” denotes DDM. Features present due to overlapping isotropic distributions of multiple species are labeled with an asterisk and not discussed. (C) Peak ratios calculated for WT, F3A, G16A, and K9P/F11P. (D) Experimental CCSs and cross-section differences between WT complexes having  $m$  and  $m + 1$  DDM.**

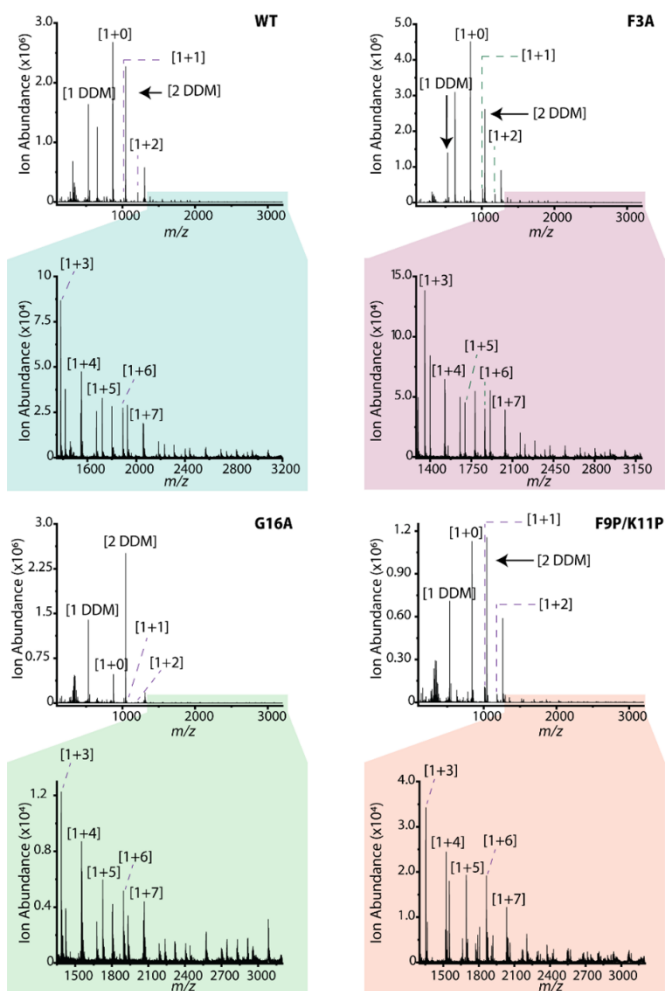


**Figure 3.8.** ESI-MS in positive mode of PSM $\alpha$ 3 peptides and 200  $\mu$ M DDM in 20  $\mu$ M of ammonium acetate. The mass spectral peaks are labeled according to the number of DDM molecules that are complexed with the peptide monomer “intact” species with  $z = +3$ . For example, the peaks labeled [1+2] consists of the intact complex of the peptide monomer and 2 DDM molecules at a +3 charge state. Only those peptide-DDM complexes discussed in the main text and previous figures (Figure 3.7 A-B and Figure 3.9) are labeled. Also of note is the presence of the DDM monomer and dimer species ([1 DDM + Na<sup>+</sup>] and [2 DDM + Na<sup>+</sup>]) at  $m/z = 533.3$  and 1043.6, respectively. Peptide concentrations are as follows: WT 14.6  $\mu$ M, F3A 15.4  $\mu$ M, G16A 3.5  $\mu$ M, and K9P/F11P 4.6  $\mu$ M.

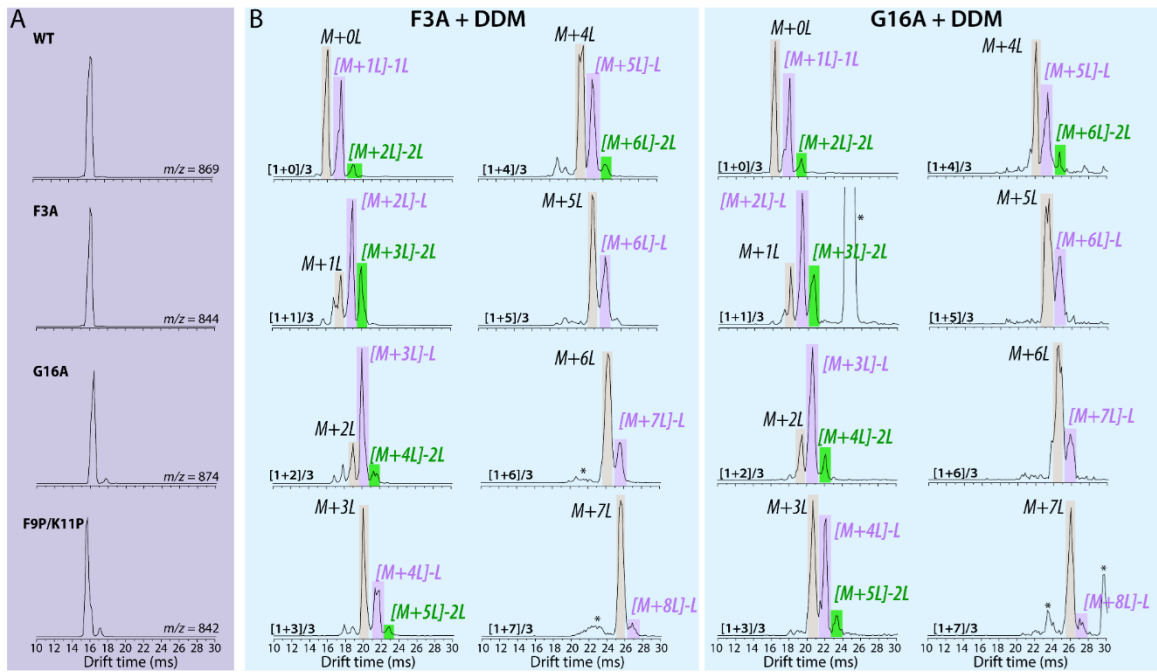


**Figure 3.9. (A) Representative ATDs of  $n/z = 1/3$  for all peptides in the absence of DDM. (B) Representative ATDs of F3A and G16A complexed with DDM.**

unlikely “new” conformations. We assign these features to neutral losses of one or two DDM molecules. For example, the  $[1 + 1]/3$  species (a triply charged complex of one peptide and one DDM) traverses the drift tube more slowly than  $[1 + 0]/3$  species because it has a larger mass. After the species exits the drift tube, a neutral loss of one DDM will have this species detected as  $[1 + 0]/3$  in terms  $m/z$  but at a longer arrival time than the monomer. The higher number of DDM loss, the longer (than expected) the drift time would be. This situation is opposite to the losses of one or several charged monomers, which yield multiple features at shorter arrival times than expected, as discussed earlier. This trend is consistent up to 7 DDM. Figure 3.10 shows a cartoon to summarize possible post-IM events and the logics behind our assignments.

It is expected that given a small number of DDM (e.g.,  $m = 1$  or  $2$ ), the peptide:DDM complex is not very stable and, thus, easily dissociates. However, as a large number of DDM are incorporated into the complex, the complexes become resilient, evidenced by the shortest arrival time peak (corresponding to the intact complexes) in each ATD becoming more intense. As shown in Figure 3.7A–C, the intact complex feature becomes the most dominant when  $m = 3$  for WT and  $m = 6$  for K9P/F11P. The behaviors of F3A and G16A are similar to those of the WT (see Figure 3.9 and Figure 3.11). We note that there are a few features (e.g., those in  $[1 + 4]/3$  of WT) that we labeled with asterisks. Those features are not part of the species of interest, but present because of the overlapping isotopic distributions among distinct species (for example, see Figure 3.12). When the peak ratios of the dissociated-to-intact complexes are taken (Figure 3.7C), the peak ratios of K9P/F11P are greater than those of the other peptides, indicating that the DDM:K9P/F11P complexes are much less stable than those of the other peptides.

Thus, we conclude that WT, F3A, and G16A require a fewer number of DDM to stabilize the complexes than K9P/F11P does. This may explain the difference in toxicity of the peptides; those that are easier to insert into the membrane should have higher toxicity.

Figure 3.7D shows the experimental CCSs of WT:DDM complexes obtained from the

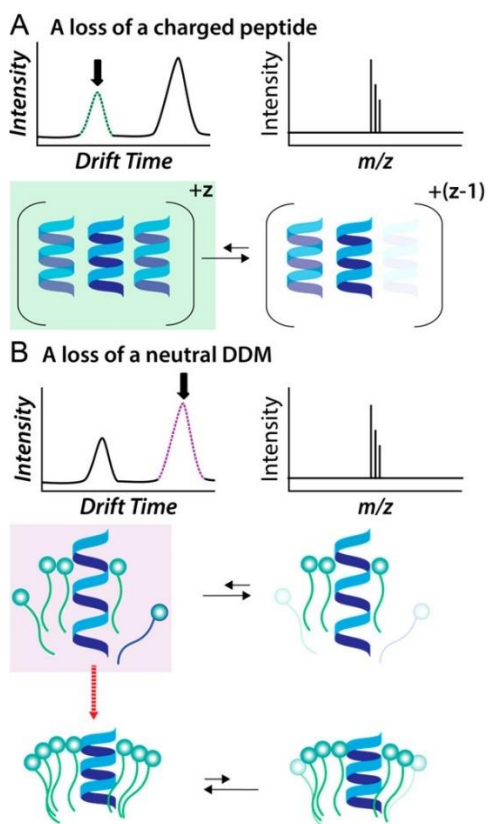


**Figure 3.10. Cartoon representation of post-IM dissociation events. (A) A loss of a charged peptide from a large oligomer leading to a new feature at short arrival time in the ATD of the small oligomer. (B) A loss of a neutral DDM from a large complex leading to a new feature at long arrival time in the ATD of the small complex.**

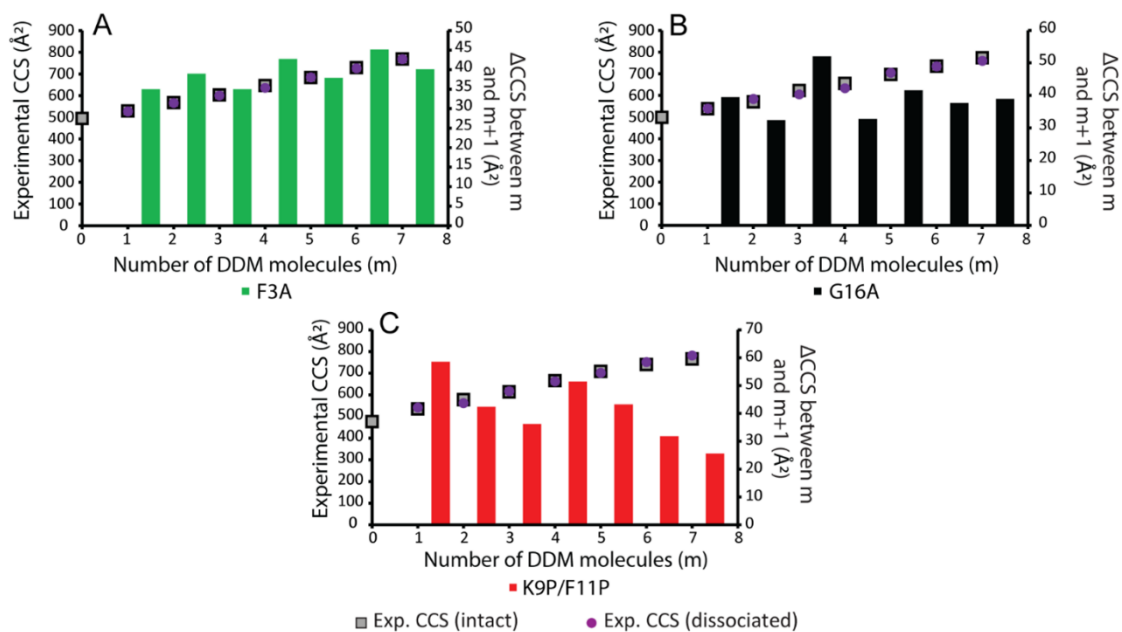


**Figure 3.11.** Experimental CCSs and cross section differences between complexes having  $m$  and  $m+1$  DDM for (A) F3A, (B) G16A, and (C) K9P/F11P. The green, black, and red bars represent the  $\Delta\text{CCS}$  between  $m$  and  $m+1$  (in  $\text{\AA}^2$ ).

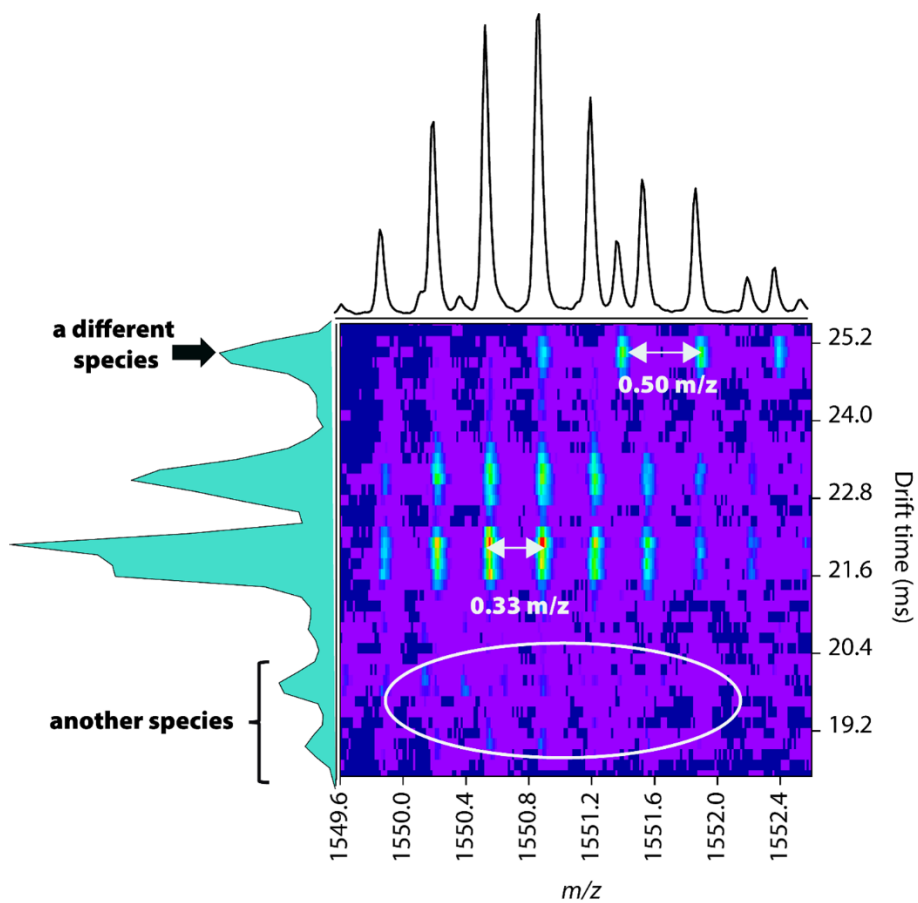


Figure 3.12. 2D plot of  $m/z$  vs drift time at  $m/z$  1598 of the WT. Features at  $\sim 25.2$  ms and  $\sim 21$  ms show a different isotopic spacing than the features of interest, indicating that they are not the same species.

features corresponding to the intact (gray) and dissociated (purple) species. The agreement between the two data sets support that our ATD assignments are reasonable. The same figure also illustrates an intriguing observation as we computed the differences in experimental CCSs of the peptide:DDM complexes. If the complexes grow directionally, it is expected that the contribution of a single DDM would get smaller. This is well-established in the studies of amyloid self-assembly.<sup>124-126</sup> Briefly, as the oligomers are formed, interactions between subunits will decrease the molecular surface of the oligomers. Bleiholder et al.<sup>125</sup> devised a linear model that considers contributions of each monomer in fibrillization as a function of oligomer size. Such a model is an example of directional growth. Here, starting from the 1:1 peptide:DDM complex, if the growth is directional, additions of DDM molecules will decrease the contribution of a DDM subunit to the total CCS of the complexes. However, if the contribution of a DDM monomer suddenly increases, that may imply a change in the type of growth. In our case, the contribution decreases from  $m = 1$  to  $m = 3$ , then increases, and remains relatively constant. Such trends may indicate the rearrangement of the complexes and account for the curvature of the micelles, which warrant a thorough investigation in the future.

### 3.6 Conclusion

In summary, we were able to take advantage of post-IM dissociation, an undesired phenomenon in IMS-MS, to connect structures and interactions of PSM $\alpha$ 3 peptides to their reported cytotoxicity. Our study provides a clear example that an accurate interpretation of IMS-MS data requires a complete set of mass spectral and mobility data, such that post-IM dissociation has to be seriously considered in studies of peptide self-assembly. It is possible for oligomers with completely different nominal  $n/z$  to appear in the same ATDs where they could be accidentally assigned as species with the same  $m/z$ . Using the technique, we demonstrate that *the toxicity of PSM $\alpha$ 3 is due to its  $\alpha$ -helicity, which renders the peptide prone to inserting into the membrane*. Lastly, post-

IM dissociations are inevitable and cannot be completely avoided or neglected. This is especially important as IMS-MS has been extensively and routinely utilized to study weakly bound noncovalent complexes.

## **4 $\alpha$ -CGRP disrupts amylin fibrillization and regulates insulin secretion: implications on diabetes and migraine.**

### **4.1 Disclaimer**

The following chapter has been adapted, with permission, from a journal article published in the *Chemical Science* (Copyright 2021 Royal Society of Chemistry).<sup>127</sup> Minor modifications were made, including the omission of several supporting figures regarding insulin. This work was performed in conjunction with Aleksandra Antevska, and the lab groups of Jason Collier, Zachary Levine, and Michael Karlstad. Aleksandra was involved in assessing the impact of amylin/CGRP on insulin secretion, Zachary provided the molecular dynamics simulations, Jason performed toxicity assays on  $\beta$ -cells, and Michael harvested the pancreatic islets. John Dunlap from the UTK microscopy center aided in obtaining transmission electron microscopy images of amylin, CGRP, and their mixtures. The IMS-MS data pertaining to amylin and CGRP was collected and analyzed by myself and Ben Link, an undergraduate student I helped mentor. Data collection and analysis were performed under the guidance of Thanh Do. Additionally, I helped train Aleksandra on our IMS-MS instrument and the software used to analyze the LC-MS data.

### **4.2 Abstract**

Despite being relatively benign and not an indicative signature of toxicity, fibril formation and fibrillar structures continue to be key factors in assessing the structure–function relationship in protein aggregation diseases. The inability to capture molecular cross-talk among key players at the tissue level before fibril formation greatly accounts for the missing link toward the development of an efficacious therapeutic intervention for Type II diabetes mellitus (T2DM). We show that human  $\alpha$ -CGRP remodeled amylin

fibrillization. Furthermore, while CGRP and/or amylin monomers reduce the secretion of both mouse Ins1 and Ins2 proteins, CGRP oligomers have a reverse effect on Ins1. Genetically reduced Ins2, the orthologous version of human insulin, has been shown to enhance insulin sensitivity and extend the life-span in old female mice. Beyond the mechanistic insights, our data suggest that CGRP regulates insulin secretion and lowers the risk of T2DM. Our result rationalizes how migraine might be protective against T2DM. We envision the new paradigm of CGRP:amylin interactions as a pivotal aspect for T2DM diagnostics and therapeutics. Maintaining a low level of amylin while increasing the level of CGRP could become a viable approach toward T2DM prevention and treatment.

### 4.3 Introduction

The onset and progression of many neurodegenerative diseases and chronic metabolic disorders are correlated with the accumulation of insoluble proteinaceous deposits.<sup>128-131</sup> The oligomeric antecedents have been considered as the cause of disease pathogenesis due to their cytotoxicity.<sup>132-136</sup> The possible toxic species are not limited to oligomers of individual proteins but cross-seeding of several distinct species.<sup>137-140</sup> Thus, it is important to account for the intervention of molecules present in the same environment and how they possibly affect the aggregation. When two peptides co-assemble, the new system may acquire unique properties that neither homologous counterparts could have, including cytotoxicity and/or aggressive fibrillization,<sup>140-142</sup> which can contribute to or exacerbate disease pathology.

For example, the pancreatic islets secrete several important prohormones to regulate blood glucose levels. Amylin, also known as islet amyloid polypeptide, is a 37-residue regulatory peptide belonging to the calcitonin peptide family and is co-secreted with insulin from pancreatic  $\beta$ -cells.<sup>143-145</sup> Amylin plays a role in glycemic regulation through delaying gastric emptying, promoting satiety, and suppressing insulin secretion. Amylin aggregation has long been considered a hallmark of type-2 diabetes mellitus (T2DM).<sup>146,</sup>

<sup>147</sup> Human amylin is an insulin suppressor of which aggregates damage islet-cells,<sup>148,</sup>  
<sup>149</sup> whereas rodent amylin, which is non-amyloidogenic, is not a modulator of insulin  
biosynthesis or secretion.<sup>150</sup> There are several hypotheses on what can trigger amylin  
aggregation (*e.g.*, pH,<sup>151, 152</sup> Zn<sup>2+</sup> concentration,<sup>153-156</sup> impaired insulin secretion,<sup>155, 157,</sup>  
<sup>158</sup> *etc.*), but there have been limited studies on the roles of calcitonin gene-related  
peptides (CGRPs) in the process.

Within the calcitonin gene family, CGRPs are widely expressed in the brain, spinal  
cord, and pancreas.<sup>159</sup> Two major isoforms of CGRPs are  $\alpha$ -CGRP and  $\beta$ -CGRP, which  
have been identified and are encoded from two distinct genes.<sup>160, 161</sup>  $\alpha$ -CGRP is found  
primarily in the central nervous system and peripheral nervous system tissues while  $\beta$ -  
CGRP is predominately located in the enteric nervous system.<sup>162</sup> Despite these  
differences,  $\alpha$ - and  $\beta$ -CGRPs are both released from the sensory nerves<sup>163</sup> and in a  
subpopulation of  $\delta$ -cells of islets.<sup>37</sup> CGRPs play an important role in both neural and  
vascular actions.<sup>164, 165</sup>

On the other hand, anti-CGRP treatments were recently approved to prevent and  
treat migraine,<sup>166</sup> a disease that may reduce the risk of T2DM.<sup>167</sup> In migraine, sensory  
nerve fibers convey the “pain” signal to the brain and evoke the release of vasoactive  
peptides including CGRPs.<sup>168, 169</sup> Therefore, the biological role of CGRPs certainly  
deserves thorough investigation in a broad context.

CGRPs and amylin share high sequence similarity (Figure 4.1). Similar to amylin,  
CGRPs also inhibit insulin secretion in both rodent and pigs,<sup>170-173</sup> possibly through  
interactions with  $\beta$ -cell receptors such as RDC1 and CRLR;<sup>174</sup> however, in a few studies,  
they were shown to stimulate insulin secretion.<sup>175-177</sup> Both peptides induce several  
proinflammatory actions. For example, both can dilate blood vessels<sup>178</sup> and compete for  
binding sites in the brain, skeletal muscle, and liver.<sup>179</sup> Surprisingly, their possible roles  
and synergistic effects in T2DM have not been thoroughly explored.

Studies of protein aggregation and characterization of oligomers existing in a  
dynamic equilibrium require sensitive analytical techniques capable of simultaneously

|                              | 1                      | 10         | 20                      | 30       |         |
|------------------------------|------------------------|------------|-------------------------|----------|---------|
| Human $\alpha$ -CGRP         | ACDTATCVT              | HRLAGLLSRS | GGVVKNNFVP              | TNVGSKAF |         |
| Human $\beta$ -CGRP          | ACNTATCVT              | HRLAGLLSRS | GGMVKSNFVP              | TNVGSKAF |         |
| Mouse $\alpha$ -CGRP         | SCNTATCVT              | HRLAGLLSRS | GGVVKDNFVP              | TNVGSEAF |         |
| Mouse $\beta$ -CGRP          | SCNTATCVT              | HRLADLLSRS | GGVLKDNFVP              | TDVGSEAF |         |
| Human Amylin                 | KCNTATCAT              | QRLANFLVHS | SNNFGAILSS              | TNVGSNTY |         |
| Mouse Amylin                 | KCNTATCAT              | QRLANFLVRS | SNNLGPVLPP              | TNVGSNTY |         |
| Human Insulin                | GIVEQCCTS              | ICSLYQLENY | CN                      |          | A-chain |
|                              | FVNQHLCGS              | HLVEALYLVC | GERGFFYTPK              | T        | B-chain |
| Mouse Insulin 1 <sup>2</sup> | GIVDQCCTS              | ICSLYQLENY | CN                      |          | A-chain |
|                              | FVKQHLCGP <sup>S</sup> | HLVEALYLVC | GERGFFYTPK <sup>M</sup> | <u>S</u> | B-chain |

**Figure 4.1.** Primary sequences of CGRPs (red), amylin, (blue), and insulin (green). For mouse insulins, the residual difference between Ins1 and Ins2 is indicated in bold. In this work, we study human  $\alpha$ -CGRP and human amylin.

isolating individual species. Ion mobility spectrometry-mass spectrometry (IMS-MS) has been demonstrated as a reliable technique for this task.<sup>180-182</sup> In IMS-MS, oligomers are identified *via* orthogonal mass to charge ( $m/z$ ) and drift time dimensions; the latter can yield valuable structural information. Amylin has been extensively studied with IMS-MS,<sup>158, 183-187</sup> and the main focus lies on structures of the monomers and dimers.<sup>158, 183-185, 187</sup> Here, in addition to traditional IMS-MS analysis of amylin and CGRP homo- and hetero-oligomers in solution, we perform liquid chromatography (LC)-IMS-MS of single mouse islets incubated with each peptide or their mixture to evaluate their synergistic effect on insulin secretion. This work is further complemented by atomistic replica exchange molecular dynamics (REMD) simulations, and transmission electron microscopy (TEM). In this contribution, we provide a comprehensive view on how CGRP disrupts amylin fibrillization, and at the same time, how CGRP regulates secretion of mouse Ins1 and Ins2 proteins, with implications on the protective effect of migraine against T2DM.

## 4.4 Materials and Methods

### **Ion mobility spectrometry – mass spectrometry**

The chemicals were purchased and used without further purification. Human amylin and human  $\alpha$ -CGRP peptides were purchased from AnaSpec (California, USA) at 98% purity. The amylin and  $\alpha$ -CGRP peptide powders were pretreated with hexafluoroisopropanol (TCI America) and the solvents were left to evaporate in a desiccator overnight. Prior to IMS-MS analysis, stocks of each peptide were prepared by reconstituting the lyophilized peptide in LC-grade water (Honeywell) and diluting to 10  $\mu$ M in a buffer of 20 mM ammonium acetate (VWR) pH = 6.8. The amylin:CGRP mixture was prepared by mixing the two peptide stocks in a 1 : 1 ratio and diluting the solution with 20 mM ammonium acetate to the final concentration of each peptide of 5  $\mu$ M. For the peptide samples containing  $ZnCl_2$ , we used 50:50 methanol:H<sub>2</sub>O instead of ammonium acetate. IMS-MS

experiments were performed on an Agilent 6560 IMS-QTOF (Agilent, Santa Clara). Briefly, the peptides were ionized by using an electrospray ionization source and their ions were then pulsed into a drift cell ( $P \approx 3.94$  Torr). After exiting the drift cell, the ions were guided through the QTOF mass spectrometer by a hexapole ion guide. Data were collected in helium, positive polarity, with  $\Delta V = 890, 790, 690, 590,$  and  $490$  V for amylin and  $\alpha$ -CGRP. The CCS values were calculated in a similar manner according to Section 2.4.

### **Replica exchange molecular dynamics**

Atomistic molecular dynamics simulations were run using a GROMACS 2020.1 integrator<sup>188</sup> and an AMBER99sb-ILDN protein force field<sup>189</sup> solvated with TIP3P water,<sup>190</sup> in order to match our prior work with amylin.<sup>191</sup>

The disulfide bond between cysteines was held constant using rigid constraints between CYX residues, which groups deprotonated CYS pairs into a single joint amino acid. Additionally, histidines were neutralized and only contained a hydrogen attached to the epsilon nitrogen atom. Newton's equations of motion were integrated over a 2 fs time step using a leapfrog algorithm. Short-range van der Waals and Coulomb potentials were truncated at 1.2 nm, while longer-ranged electrostatics were tabulated with particle mesh Ewald summation, which reduces computation with fast Fourier transforms. Cartesian periodic boundary conditions were also implemented in each direction to minimize the effects from unit-cell boundaries. Initial monomers of amylin and CGRP were generated using AVOGADRO 1.1.1<sup>192</sup> in a random-coil conformation, since monomers are intrinsically disordered. Seeding geometries for each replicate were selected at random from 20 ns equilibration simulations, which populated a wide range of structures, and thereby increased the diversity between replicates. Box dimensions ranged from 6–7 nm in length, while dimers were constructed by randomly pairing permutations of monomeric proteins and confirming that they remained dimerized for at least an additional 20 ns.

Temperature replica-exchange MD (T-REMD)<sup>193</sup> was carried out under an NVT ensemble using a Nose–Hoover thermostat<sup>194</sup> at temperatures between 290 and 370 K for 200 ns per replica (approx. 8.2  $\mu$ s of the total simulation time). This resulted in about 41 replicates per system, and the Monte Carlo exchange rate between replicates was empirically tuned to 25% (with exchanges attempted every 3 ps). While each replicate was run for 200 ns, analysis was only carried out from 100–200 ns, in order to minimize the impact of specific initial conditions. Convergence of REMD simulations was confirmed when the number of new protein conformations plateaued as a function of simulation time, indicating that only prior states were being sampled. Analysis of the room temperature replicate (300 K) from REMD was used to determine the structural and energy landscapes of protein monomers and dimers. Collision cross sections were calculated for each frame using the MOBCAL software package,<sup>195</sup> and protein clustering was carried out by grouping C $\alpha$  atoms together within a root-mean-square deviation of 2 nm. Analysis of the room temperature replicate (300 K) was used to determine the secondary structural propensities and energy landscapes of CGRP monomers, in addition to homo- and heterodimers. Additional computational details can be found in our prior work.<sup>191</sup>

### **Transmission electron microscopy**

The peptides for TEM analysis were prepared at the same concentrations as for IMS-MS analysis and stained with 1% phosphotungstic acid stain on a piece of parafilm. TEM images were acquired on a JEOL 1400-Flash transmission electron microscope (JEOL, Peabody) at 80 kV and recorded with a Gatan OneView camera.

### **Mouse islet culturing and harvesting**

Six 10 week-old male C57BL/6J mice (stock #000664) were ordered from the Jackson Laboratory (Bar Harbor, Maine) and allowed to acclimate to the animal facility conditions on a 12 h light cycle at  $22 \pm 1$  °C. The mice were provided standard chow and

drinking water *ad libitum* until they reached 12 weeks of age. At 12 weeks of age, the mice were euthanized *via* CO<sub>2</sub> asphyxiation followed by a cervical dislocation. All procedures were approved by the University of Tennessee Institutional Animal Care and Use Committee. Following isolation, the islets were incubated at 37 °C and 5% CO<sub>2</sub> for 24 hours in RPMI + D-glucose. After the 24 hour incubation, the islets were individually pipetted into Eppendorf tubes containing the blank solvent, solutions of individual peptides, or a mixture of the peptides.

Here we used a total of six mice for two sets of experiments. In the first set of experiments (3 mice), islets from the same mouse, except for those used as controls, were incubated with fresh amylin, fresh CGRP, or fresh mixture of amylin and CGRP. We used at least 3 islets from different mice per peptide condition. Because our LC-IMS-MS workflow was optimized for sub-single islet measurements, we can test multiple conditions on the same mouse, and obtain multiple (at least three) technical replicates from a single islet. In the second set of experiments (3 mice), we followed the same procedure as above. The only difference is that the islets were incubated with 4 day old amylin, CGRP, or the mixture.

### **Liquid chromatography IMS-MS of single mouse islets**

Stocks of amylin and CGRP were prepared as previously described and diluted with a 5% MeOH/0.1% formic acid solution to a final concentration of 10 µM. The concentration of each peptide in the amylin:CGRP mixture was 10 µM. Aliquots of 50 µL of each peptide solution (amylin, CGRP, and amylin:CGRP) were placed into Eppendorf tubes, wherein single islets were manually placed. Two sets of peptide solutions were used: a fresh solution, where the peptide solutions were prepared approximately one day before acquiring the islets and contains primarily peptide monomers, and then an “aged” solution, where peptide solutions were left at room temperature four days before islet acquisition and consists primarily of peptide oligomers. The “aged” solution was characterized with IMS-MS to determine the relative ratios of oligomers to monomers.

This was used to evaluate the effect of oligomers on the same islet system. For LC-IMS-MS analysis, an Agilent 1260 Infinity coupled to the Agilent 6560 IMS-MS was employed. In each run, 4.0  $\mu$ L of the sample was injected into a 2.1  $\times$  50 mm Zorbax Extended C18 reverse phase column (Agilent, Santa Clara). The mobile phase was composed of an aqueous solution of 0.1% formic acid in LC-grade water (solvent A) and acetonitrile (solvent B). The mobile phase gradient is as follows: 98% solvent A for 3 minutes, 80% solvent A (3–5 minutes), 65% solvent A (5–18 minutes), 50% solvent A (18–23 minutes), 25% solvent A (23–26 minutes), 10% solvent A (26–33 minutes), 10% solvent A was held for 2 minutes then followed by a rise to 98% (35–40 minutes) and kept for 5 minutes (40–45 minutes) for column equilibration. The column compartment was set to 40 °C. Three technical replicates were performed for each sample, with two consecutive blank injections between each sample to clean the column.

#### **Cell viability assays**

Stock solutions of amylin and CGRP in LC-grade water were added to cultures of 832/13 rat insulinoma cells to obtain the desired concentrations of 20, 200 nM, and 5  $\mu$ M and incubated overnight at 37 °C. Adenylate kinase release into the culture media was assessed using the Adenylate Kinase Assay Kit (Abcam #: ab228557) which measures the activity of the adenylate kinase released by damaged cells. Interleukin-1 $\beta$  (IL-1 $\beta$ ) and camptothecin (CT) were used as controls.

## **4.5 Results and Discussion**

### **IMS-MS reveals a CGRP $\alpha$ -helical intermediate that is stable upon interacting with Zn<sup>2+</sup>**

Figure 4.2A shows the representative ESI-mass spectrum of amylin. The mass spectral peaks are annotated with their nominal  $n/z$ . The top mass spectrum shows several peaks corresponding to small-order oligomers that are high in abundance. The bottom mass spectrum highlights the less intense mass spectral peaks that correspond to higher-order oligomers. In the designated mass range, amylin can form up to a

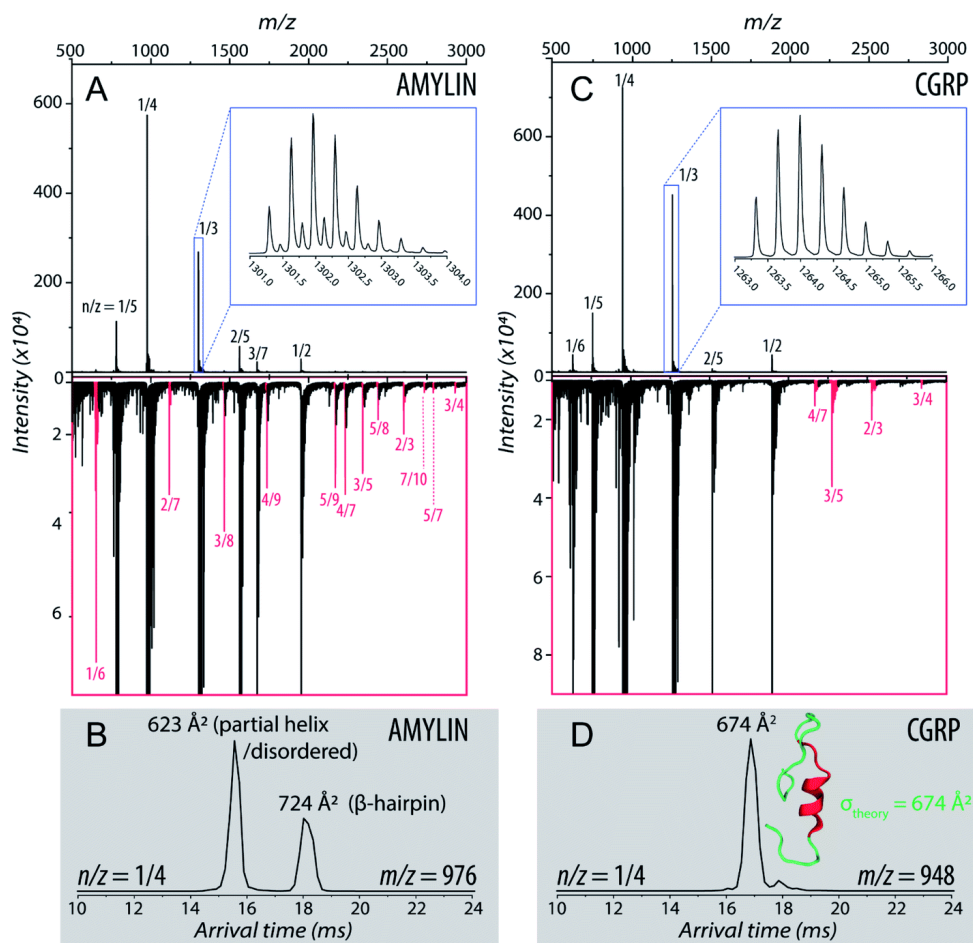


**Figure 4.2.** (A) Mass spectrum of amylin (10  $\mu$ M) in 20 mM ammonium acetate. Each peak is labeled by its nominal  $n/z$  ratio. The bottom spectrum exhibits oligomeric species with a relatively low intensity that is not easily identifiable in the top spectrum. The inset shows the isotopic spacing of  $n/z = 1/3$  which exhibits the presence of a dimer with  $z = +6$ . (B) The ATD of amylin at  $m/z = 976$ . (C) Mass spectrum of CGRP (10  $\mu$ M) in 20 mM ammonium acetate. The inset shows the isotopic spacing of  $n/z = 1/3$  which indicates that this mass spectral peak correlates to a triply charged dimer. (D) The ATD of CGRP at  $m/z = 948$  and a REMD conformation of CGRP monomer with the segment 8–18 colored in red. This conformation is consistent with previous NMR and CD studies of CGRP in structure-inducing solvent.

heptamer. However, upon analysis of the isotopic spacing and ion mobility data, oligomers larger than what is initially seen within each mass spectral peak can be observed. Isobaric oligomers may be hidden under the same nominal  $m/z$ . For example, the inset in Figure 4.2 shows the isotopic spacing of  $n/z = 1/3$ , indicating the presence of a monomer with a +3 charge state and a dimer with a +6 charge state (*i.e.*  $n/z = 2/6 = 1/3$ ).

Amylin, specifically the monomer ( $z = +4$ ), has been characterized and studied extensively with IMS-MS in the past.<sup>158, 183-185</sup> It has been shown by Bowers and Shea<sup>183-185</sup> that amylin populates two distinct families of stable conformations of amyloidogenic  $\beta$ -hairpin and a coil-helix. We note that the terms “helix” and “ $\beta$ -hairpin” were taken from the same studies. Figure 4.2B shows the arrival time distribution (ATD) of the amylin monomer ( $z = +4$ ) illustrating these two families of conformations. The collisional cross sections (CCSs) of these species were found to be 623 Å<sup>2</sup> for the coil-helix and 724 Å<sup>2</sup> for the  $\beta$ -hairpin. The difference in the CCS between these two structures is in good agreement with the data by Bowers and Shea.<sup>183-185</sup>

Unlike amylin,  $\alpha$ -CGRP (which we will refer to as “CGRP” hereafter) has not been characterized with IMS-MS. The mass spectrum of CGRP is shown in Figure 4.2C, indicating that this peptide formed several small oligomers with the largest one being a tetramer. Upon initial observation of the mass spectra, it is evident that CGRP formed smaller oligomers than amylin. Our data indicate that amylin readily assembles into oligomeric species and is much more prone to aggregation than CGRP.

Figure 4.2D shows the ATD of the CGRP monomer ( $z = +4$ ), which contains only one prominent feature with an arrival time at 16.89 ms. When compared to the amylin data, the CGRP structure falls in between the two major populations of amylin (coil-helix at 15.58 ms and  $\beta$ -hairpin at 18.05 ms). If CGRP were to adopt the same or similar conformations as amylin, then CGRP should have a slightly shorter arrival time (smaller CCS) than the respective amylin conformers due to having less mass. Therefore, this feature could correlate to a coil-helical CGRP conformer ( $\text{CCS}_{\text{exp}} = 674 \text{ Å}^2$ ). This is not

pure speculation; previous studies of CGRP in structure-inducing solvents (*e.g.*, trifluoroethanol (TFE)/water) demonstrated that residues 8–18 of the peptides exhibit a strong  $\alpha$ -helical propensity.<sup>196-199</sup> This segment encompasses the chameleon sequence “LAGLL” found in 8 different proteins.<sup>200</sup> However, it is entirely possible for several conformations with different secondary structures to have the same CCS, as discussed below together with the REMD simulations. Nonetheless, Figure 4.2D suggests that the structural preference of CGRP prevents the formation of the  $\beta$ -hairpin motif, thus inhibiting extensive CGRP aggregation.

### **Zinc complexation promotes $\beta$ -hairpin of amylin but favors the native structure of CGRP**

To further investigate the aggregation of CGRP and amylin under biologically relevant conditions, we added 0.5 mM  $\text{ZnCl}_2$  to the samples. We used the amylin data, which were previously reported,<sup>185</sup> as the control.  $\text{Zn}^{2+}$  is involved in the packaging, processing, and the secretion of insulin from  $\beta$ -cells.  $\text{Zn}^{2+}$  is found in mM concentration in pancreatic  $\beta$ -cells, reaching 10–20 mM in dense-core granules,<sup>201, 202</sup> and has been found to affect amylin fibrillization upon complexation. Previous studies showed that  $\text{Zn}^{2+}$  inhibited amylin fibril formation at low, physiologically relevant extracellular concentrations<sup>154, 157</sup> but promoted aggregation at high (intracellular) concentrations.<sup>185</sup>

Figure 4.3A shows the mass spectrum of amylin in the presence of  $\text{Zn}^{2+}$  (50:1  $\text{ZnCl}_2$ :amylin) which illustrates many Zn–amylin complexes. A majority of these species, such as the complexes of amylin containing Zn and  $\text{Zn}(\text{H}_2\text{O})_5$ , were observed by Bowers and coworkers.<sup>185</sup> The same types of complexes were observed for CGRP (Figure 4.3B).

Bowers and coworkers suggested that the amplification of amylin aggregation upon coordination with  $\text{Zn}^{2+}$  could be attributed to stabilization of the  $\beta$ -hairpin motif.<sup>185</sup> Similar findings are illustrated in Figure 4.3C where we show the ATDs of the bare amylin monomer ( $z = +4$ ) and two Zn:amylin complexes. The ATD of the bare amylin monomer has one major feature which corresponds to the coil-helix

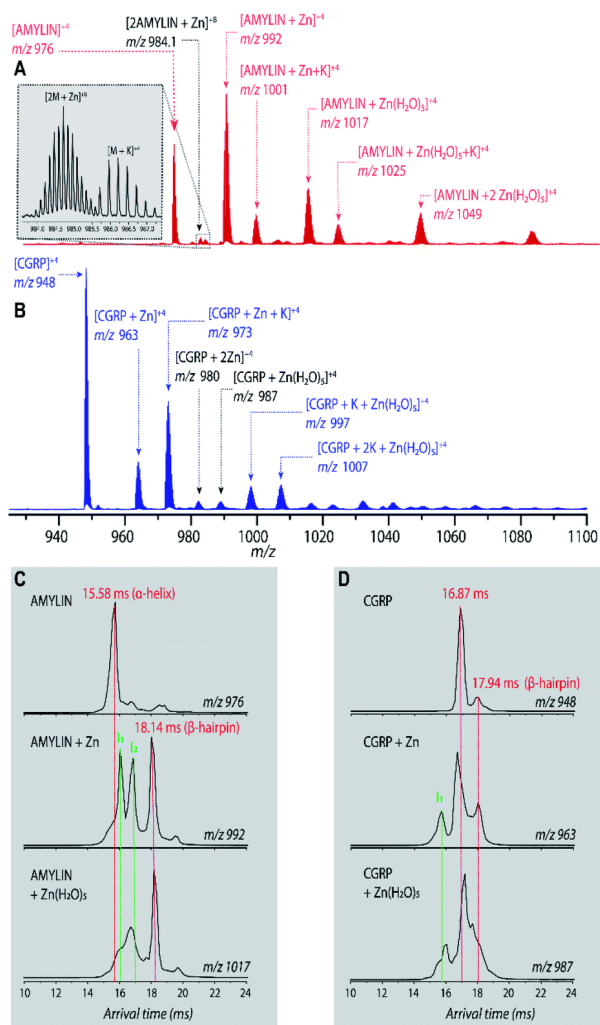


**Figure 4.3. (A and B) Representative ESI-mass spectra (in 50 : 50 MeOH : H<sub>2</sub>O) of (A) 10  $\mu$ M amylin and (B) 10  $\mu$ M CGRP, both in the presence of 0.5 mM Zn<sup>2+</sup>. (C) Representative ATDs of amylin, amylin : Zn, and amylin : Zn(H<sub>2</sub>O)<sub>5</sub>. (D) Representative ATDs of CGRP, CGRP : Zn, and CGRP : Zn(H<sub>2</sub>O)<sub>5</sub>.**

conformation. However, upon complexation of amylin with  $\text{Zn}^{2+}$ , the conformational distribution shifts to favor the  $\beta$ -hairpin, with this effect becoming more pronounced when complexed with the  $\text{Zn}(\text{H}_2\text{O})_5$  adduct.

We detected two intermediate species upon Zn–amylin complexation (annotated as  $I_1$  and  $I_2$  in Figure 4.3C). These intermediates were favored by  $\text{Zn}^{2+}$  complexation, competing with the amyloidogenic  $\beta$ -hairpin. Such competition could lead to an inhibition of fibril formation.<sup>153, 156</sup> These intermediates were not observed in the previous study<sup>185</sup> due to low mobility resolution of the instruments used. Here, their presence provides a rational explanation for the moderate inhibitory effect of  $\text{Zn}^{2+}$  on amylin aggregation. However, fibril formation was still observed in months-old samples of amylin with  $\text{ZnCl}_2$  as shown in Figure 4.3D.

Since CGRP shares similar biological properties with amylin, we evaluated its structures upon Zn complexation. Figure 4.3D shows the ATDs of the bare monomer ( $z = +4$ ), CGRP + Zn, and CGRP +  $\text{Zn}(\text{H}_2\text{O})_5$ . For bare CGRP (Figure 4.3D), there is only a single prominent feature, identical to that in Figure 4.2. However, once CGRP complexes with Zn, we note the addition of at least two new features. As mentioned in the previous section, the feature at  $\sim 18$  ms is assigned to the amyloidogenic  $\beta$ -hairpin motif. Observing the ATDs of the Zn–CGRP complexes, Zn slightly promotes the formation of this  $\beta$ -hairpin, albeit not as much as what was observed for amylin. In addition, a new compact species ( $t_a = \sim 16$  ms) in the same ATDs (Figure 4.3D; feature  $I_1$ ) exhibits a similar arrival time to that of the amylin's coil-helix. Given its short arrival time, this intermediate is likely a coil (disordered) structure. Notably, although Zn shifts the conformations of CGRP to a disordered and  $\beta$ -hairpin, the original structure of the CGRP is still the most dominant feature. Thus, the conformational space of CGRP is not dramatically changed upon  $\text{Zn}^{2+}$  complexation. The IMS-MS data help explain why CGRP does not aggregate as aggressively as amylin. Additionally, the TEM images (Figure 4.4B and E) show that CGRP formed neither small observable aggregates nor fibrils under any of the conditions.

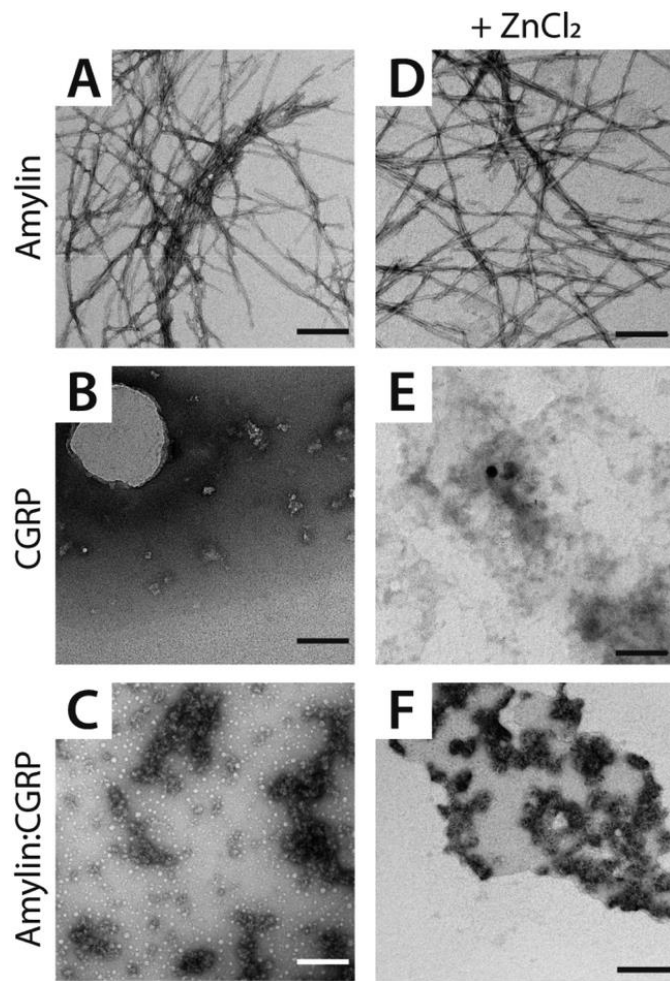


Figure 4.4. Representative TEM images of (A and D) amylin, (B and E) CGRP, and (C and F) the mixture. The scale bar is 150 nm.

### **Amylin, CGRP, and their mixture do not show significant toxicity to cultured $\beta$ -cells**

Since both amylin and CGRP are present within pancreatic islets, it would not be surprising if these peptides interacted in vivo to form amylin–CGRP hetero-oligomers. In the first approach, we identify and characterize the types of hetero-oligomers in solution using IMS-MS. Figure 4.5 shows the partial mass spectra of an amylin–CGRP heterodimer and heterotrimer. In each panel, the top mass spectrum is of amylin alone, while the bottom mass spectrum is of the amylin:CGRP mixture. Based on the nominal mass, the heterodimer consisted of amylin–CGRP in a 1:1 stoichiometric ratio and the heterotrimer was composed of 2:1 amylin–CGRP. Other species identified within the mass spectrum of the amylin–CGRP mixture is found in Figure 4.6. Based on  $m/z$  values, the heteroligomers contain more amylin than CGRP, suggesting that CGRP binds to amylin oligomers to remodel amylin aggregation. Additionally, the TEM images acquired from the amylin:CGRP mixture show the presence of small globular aggregates (Figure 4.4C and F). Only indistinct and faint fibrillar species can be observed among the predominantly globular aggregates in the TEM images of the mixture without  $\text{ZnCl}_2$  samples (see Figure 4.7).

Furthermore, the mass spectrum of the amylin:CGRP mixture at  $t = 1$  week does not contain the prominent amylin mass spectral peaks (Figure 4.8). This hints that amylin has fallen out of the solution to form insoluble species which connect to the observed globular aggregates. In other words, the globular aggregates are predominantly made of amylin. When  $\text{Zn}^{2+}$  is present, some distinct fibrils of various sizes can be observed along with the globular aggregates in the TEM images of the mixture (Figure 4.7). This agrees with our previous observation of the changed conformational spaces of the peptides upon  $\text{Zn}^{2+}$  coordination, which subsequently lessened the cross-talk between them. To determine the cytotoxicity of the amylin–CGRP aggregates, we studied the release of adenylate kinase, an intracellular enzyme marker of cell death, from cultured  $\beta$ -cells when exposed to amylin–CGRP aggregates and amylin fibrils. We assessed the toxicity of amylin and the amylin:CGRP mixture using two negative controls: the no-treatment

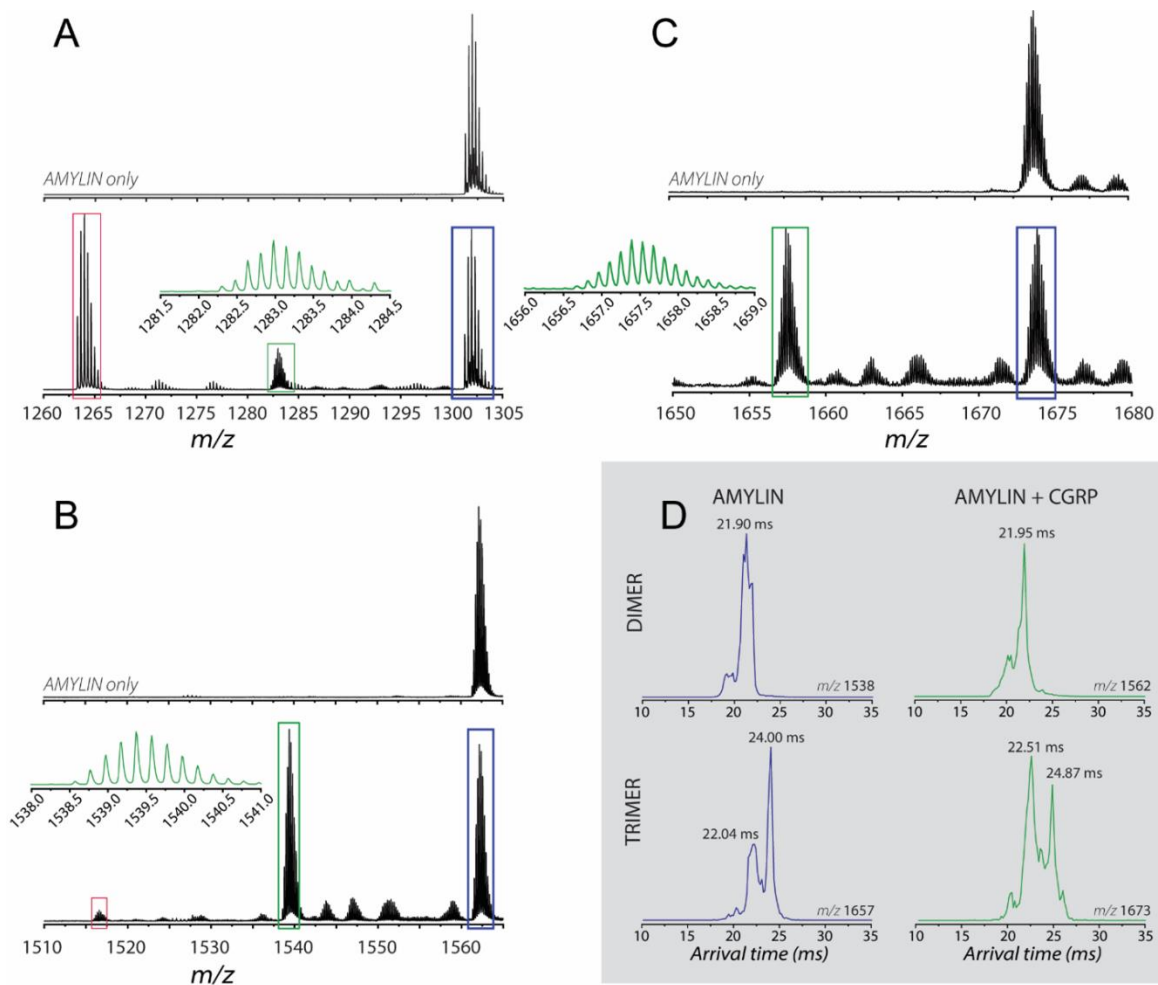


**Figure 4.5. Partial mass spectra of amylin:CGRP hetero- dimer and trimer and corresponding ATDs. For panels A – C, the top mass spectrum is of amylin alone and the bottom mass spectrum is of the amylin:CGRP mixture. Within the partial mass spectra, the mass spectral peaks boxed in blue correlate to amylin, red is CGRP, and green represents co-oligomeric species. (A - B) amylin:CGRP (1:1) heterodimers (C) amylin:CGRP (2:1) heterotrimer (D) ATDs of amylin dimer and trimer (left) and amylin:CGRP hetero- dimer and trimer (right).**

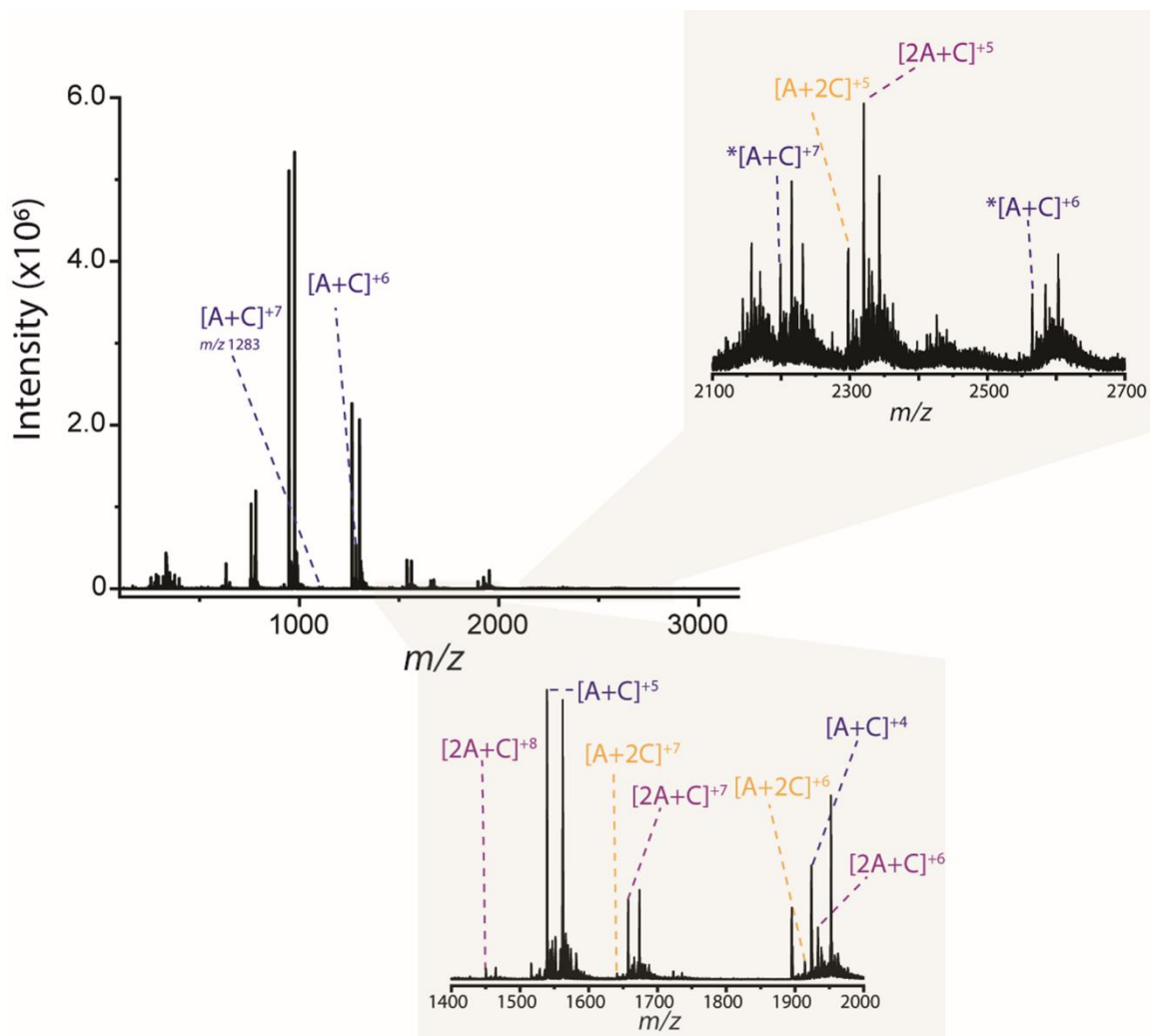


Figure 4.6. The mass spectrum of 1:1 amylin:CGRP (5:5  $\mu$ M) in 20 mM ammonium acetate at  $t = 0$ . Mass spectral peaks labeled “A+C” represents oligomers composed of 1:1 amylin:CGRP, “2A+C” represents oligomers made up of 2:1 amylin:CGRP, and “A+2C” represents oligomers of 1:2 amylin:CGRP. Note the asterisk next to the 1:1 amylin:CGRP species at  $z = +7$  and  $+8$  in the 2100 – 2700 mass range. These are species that are composed of 2:2 amylin:CGRP.

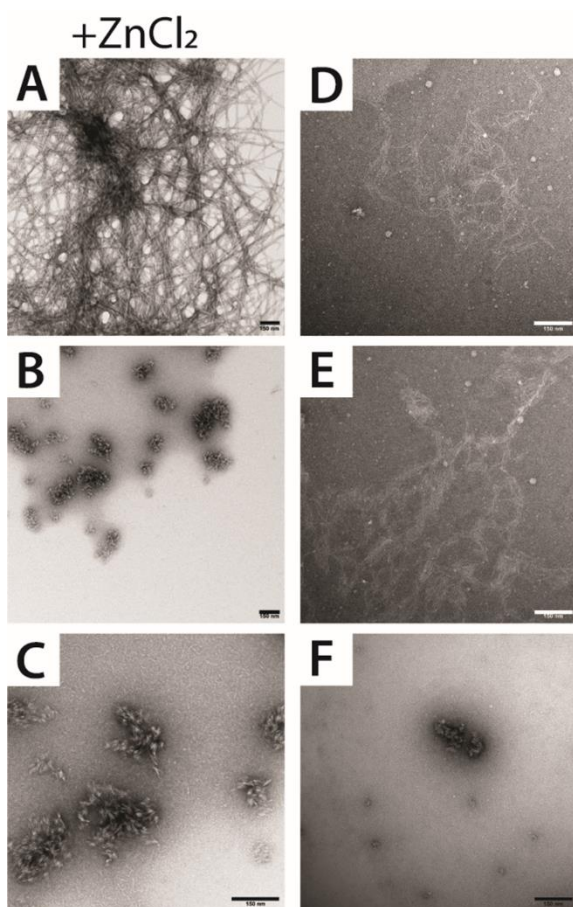


**Figure 4.7.** Representative TEM images of the amylin:CGRP mixture with ZnCl<sub>2</sub> (A – C) and without ZnCl<sub>2</sub> (D – F). The samples with Zn had a mixture of material, including (A) long fibrils and (B-C) short fibrils. The short fibrils look similar in morphology to the globular aggregates (e.g. panel F). The samples without ZnCl<sub>2</sub> primarily exhibited globular aggregates, however there were some faint and indistinct fibrillar species present (D-E). The scale bar is a 150 nm.

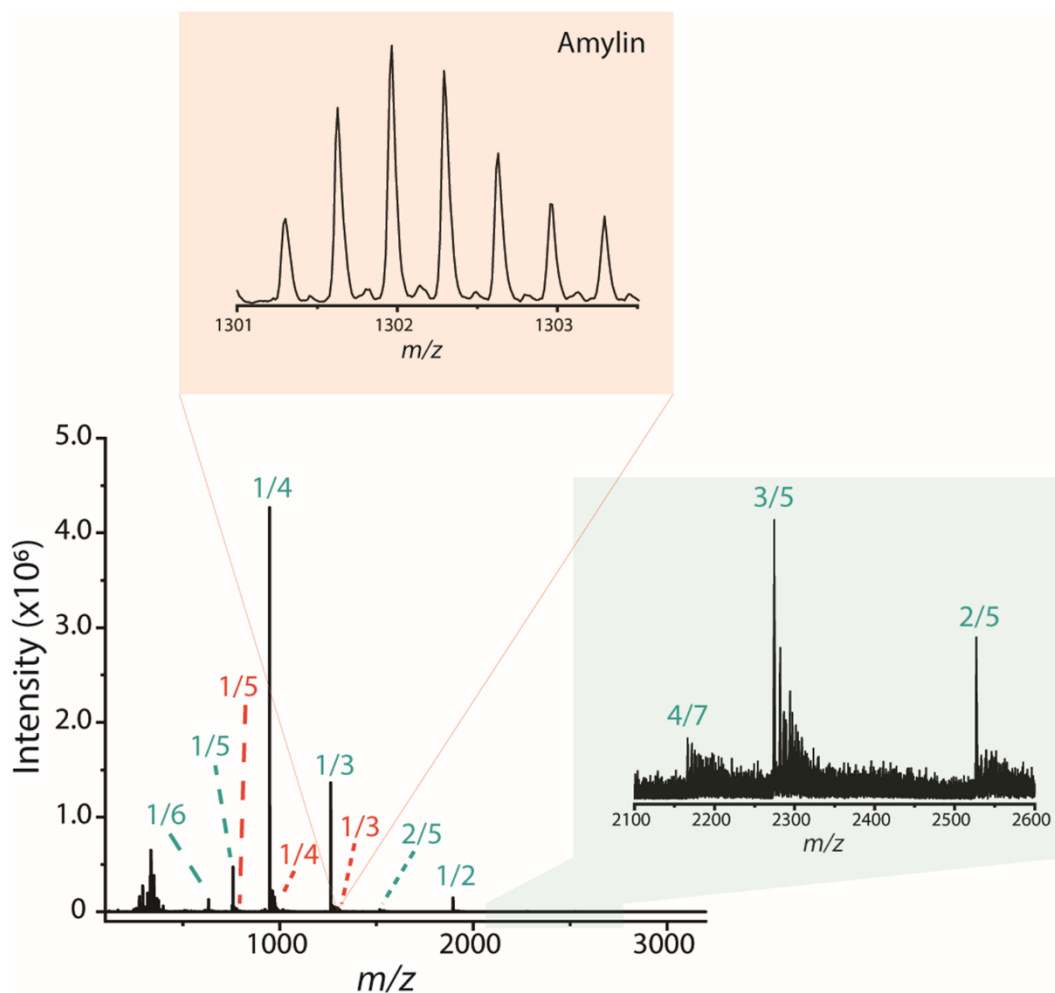


**Figure 4.8.** The mass spectrum of 1:1 amylin:CGRP (5:5  $\mu$ M) in 20 mM ammonium acetate after setting out at room temperature for one week. Only peaks corresponding to CGRP (blue) and amylin (orange) are labeled. Note that the intensity of the amylin peaks are much lower than what is shown in Figure 4.6. This implies that amylin has fallen out of solution to form globular aggregates.

condition and interleukin-1 $\beta$  (IL-1 $\beta$ ) treatment. IL-1 $\beta$  was used to indicate that insulin secretion can be reduced without significant toxicity.<sup>203</sup> For positive control, we used camptothecin (CT), a topoisomerase I toxin that damages eukaryotic cells, including  $\beta$ -cells.<sup>204</sup>

Previous investigations of amylin toxicity used micromolar concentrations (e.g., 0.5–8  $\mu$ M in Shigihara et al.<sup>205</sup> and ~2–8  $\mu$ M in Ritzel and Butler<sup>206</sup>) to assess toxicity. Here we performed three sets of experiments at 20 nM, 200 nM, and 5  $\mu$ M (see Figure 4.9 and Figure 4.10). In all cases, amylin, CGRP and their mixture did not result in significant toxicity. Nonetheless, amylin appeared to be more toxic than CGRP and amylin + CGRP. Regarding toxicity of amylin, our data are in line with those of Westwell-Roper et al.<sup>207</sup> who suggested that in a living organism, if amylin accumulates, it acts on the islet resident macrophage to create toxicity but not directly on the  $\beta$ -cells.

#### ***In silico* CGRP-amylin simulations support IMS-MS derived conformers**

To gain structural CGRP and amylin interactions, we carried out REMD simulations on CGRP monomers and heterodimers in solution. For amylin, since the early implicit-solvent REMD simulations by Bowers and Shea,<sup>183, 184</sup> newer force fields and explicit solvent models have been utilized to better intrinsically describe the disordered nature of amylin monomers and the  $\beta$ -hairpin dimer, including a recent study by Levine and Shea<sup>191</sup> on the interactions of amylin with mitochondrial derived peptides. Here the simulation protocol of CGRP simulations is identical to that of the reported amylin.

While theoretical and experimental CCS comparison has been a common approach in peptide structural studies using IMS-MS,<sup>180, 181, 208</sup> for intrinsically disordered peptides, matching experimental CCSs to model structures is challenging. In the gas phase, disordered structures can be denatured, while  $\alpha$ -helical structures can be over-stabilized.<sup>209</sup> Notably, the structure shown in the inset of Figure 4.2D was obtained from our REMD simulation, which shows a helical content consistent with previous NMR/circular dichroism studies in PFE/water (a condition that stabilizes the

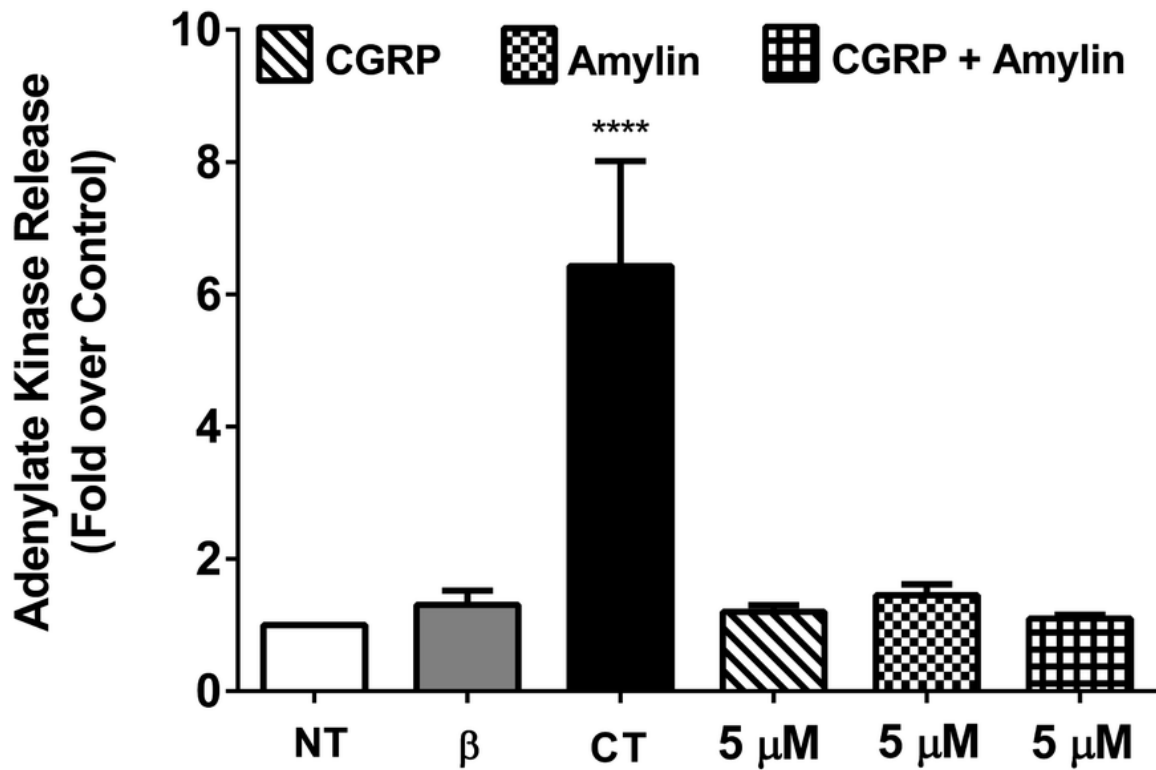


Figure 4.9. Cell viability assay using cultured  $\beta$ -cells exposed to CGRP, amylin, and amylin :CGRP. 832/13 rat  $\beta$ -cells were either left untreated (NT) or exposed to  $1 \text{ ng mL}^{-1}$  IL-1 $\beta$  and  $2 \text{ mM}$  camptothecin (CT) compared with  $5 \text{ }\mu\text{M}$  of amylin, CGRP, and amylin plus CGRP for 18 h. \*\*\*\*,  $p < 0.0001$  versus NT using one-way ANOVA.  $n = 4\text{--}6$  per group. When compared to CT, the effects of amylin, CGRP, or amylin :CGRP on cellular viability are minor, although amylin has a stronger effect than the other two conditions.

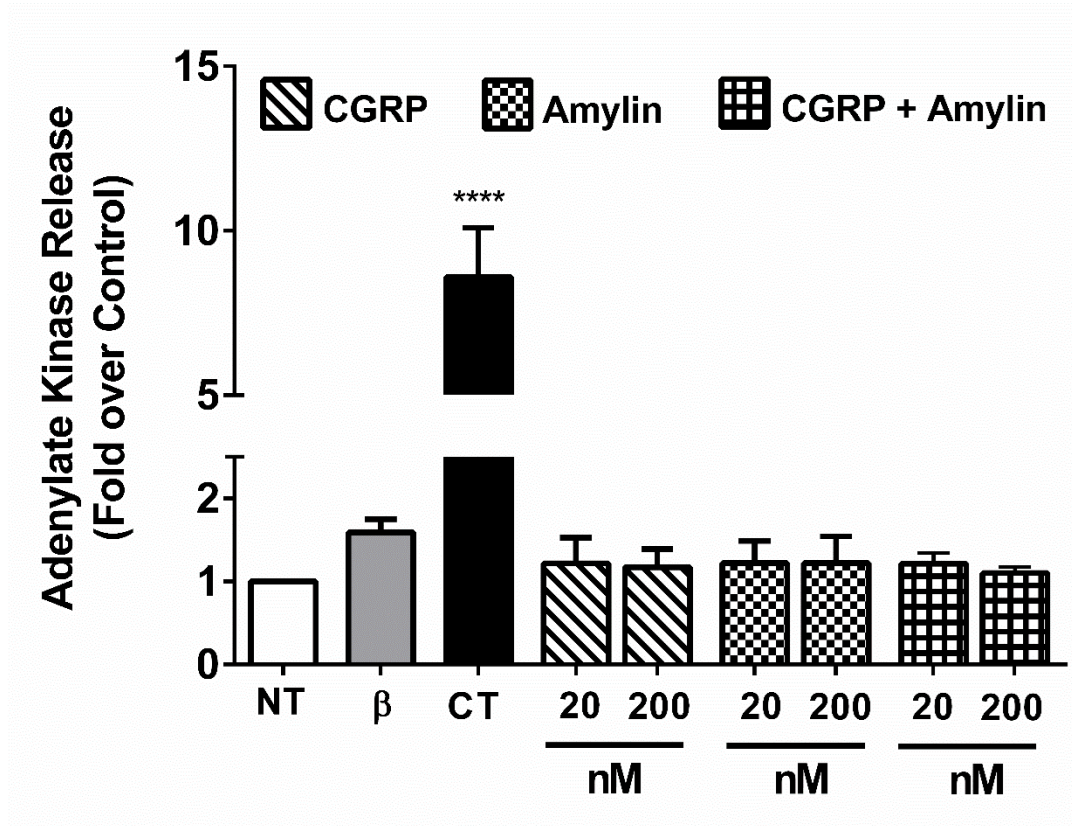


Figure 4.10. Cell viability assay using cultured  $\beta$ -cells exposed to CGRP, amylin, and amylin:CGRP. 832/13 rat b-cells were either left untreated (NT) or exposed to 1 ng/mL IL-1b, 2 mM camptothecin (CT) compared with 20 and 200 nM of amylin, CGRP, or amylin plus CGRP for 18 h. \*\*\*\*,  $p < 0.0001$  versus NT using one-way ANOVA.  $n = 4-6$  per group. When compared to CT, there effects of amylin, CGRP, or amylin+CGRP on cellular viability on cell viability are minor, although amylin has a stronger effect than other two conditions.

$\alpha$ -helix).<sup>196-198</sup> However, the population of this structure is very low given that the majority of CGRP monomers (from simulation) are disordered. Another way to understand the data is as follows. When we cluster protein states from REMD based just on residues 8–18 (corresponding to CGRP regions most likely to exhibit any secondary structures), we find that 60% of CGRP conformations adopt either a disordered +  $\beta$ -stranded structure, a completely disordered structure, or a disordered-helical structure (Figure 4.11). When we back-calculate CCS values for each of these eigenstates, we find an excellent agreement with the IMS-MS data. Finally, the range of theoretical CCSs for both CGRP/amylin monomers and dimers are in agreement with the experimental data (Figure 4.12, Figure 4.13, and Figure 4.14).

Here we used the REMD data to rationalize the inhibitory effect of CGRP on amylin aggregation. To better quantify these structural heterogeneities, we provide molecular contact maps (Figure 4.15) that indicate the average proximity of amino acids in each structure. CGRP and amylin monomers exhibit no obvious contact patterns since they tend to remain strongly disordered. However, a characteristic “X” pattern on these contact maps indicates the presence of a hairpin, which is notable in the amylin homodimer. While largely disordered, one of the CGRP proteins also appears to adopt a partial hairpin in the homodimer. Interestingly, in the CGRP–amylin heterodimer, the positions of the hairpin move to the exterior of each protein. In CGRP, the hairpin moves more towards the N-terminus while in amylin, the hairpin moves towards the C-terminus. As for intermolecular interactions, amylin tends to bind CGRP towards its C-terminus almost directly on the hairpin location. This bolsters the hypothesis that disordered CGRP tends to bind  $\beta$ -stranded amylin, which likely inhibits the ability to form regular fibrillar structures. Specifically, Asn31 in amylin regularly forms hydrogen bonds with residues 16–18 on CGRP. This interaction provides both molecules with stable  $\beta$ -sheets that are prevalent in almost all states, while the rest of the molecule is highly disordered. Of note, Asn31 mutations have been shown to reduce the  $\beta$ -sheet content in simulations of amylin protofilaments, suggesting that such interactions may

Clustering residues 8-18 reveals  $\beta$ -sheet, coil, and  $\alpha$ -helical clusters with predicted CCS values close to MS-derived data.

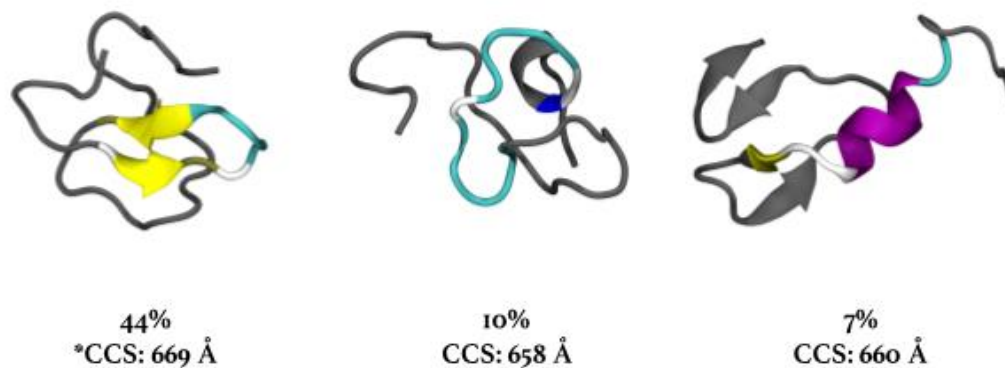


**Figure 4.11.** Representative structures of CGRP monomers obtained from REMD simulations and clustered together by residues 8-18. CCS values for these eigenstates are in excellent agreement with IM-MS data.

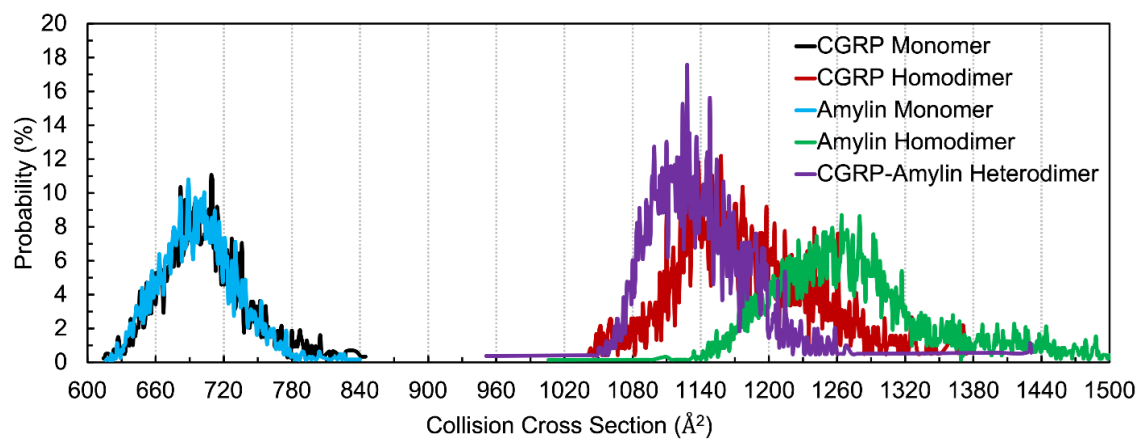


**Figure 4.12. CGRP and amylin CCS from REMD simulations reveal small disordered CGRP-containing homodimers and heterodimer compared to larger  $\beta$ -stranded amylin dimers.**

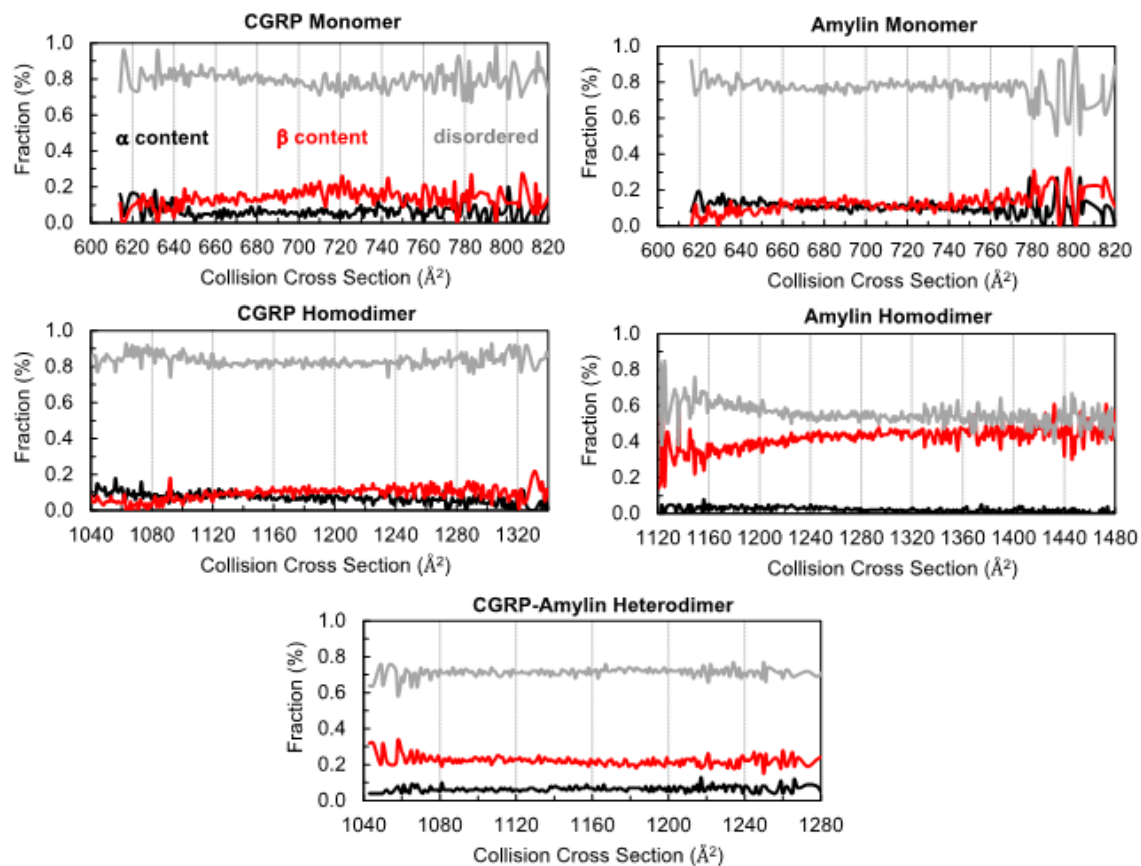


**Figure 4.13.** Theoretical CCS values as a function of corresponding protein secondary structures from REMD, which highlight significantly disordered conformations.

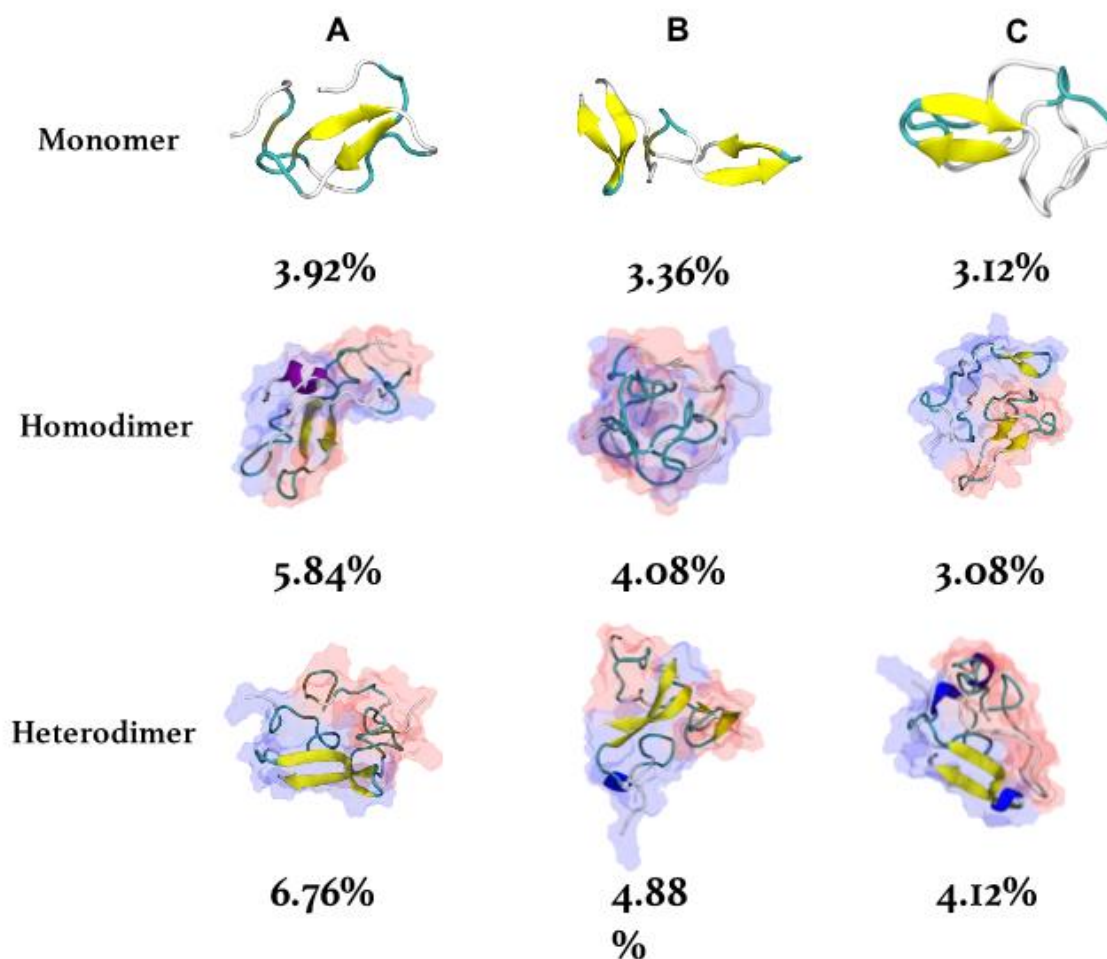


Figure 4.14. Representative structures of CGRP monomers (*top*), homodimers (*middle*), and heterodimers (*bottom*) obtained from REMD simulations.

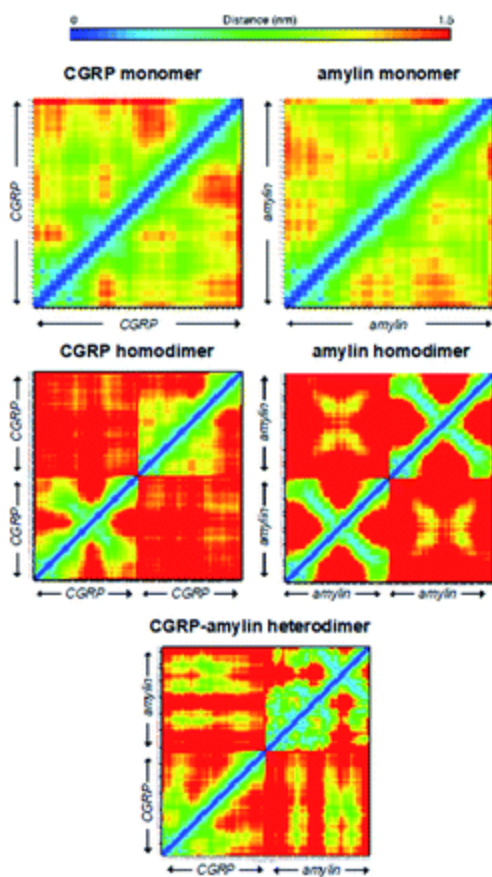


Figure 4.15. Molecular contact maps from REMD simulations reveal a range of disordered and hairpin conformations.

be important in forming oligomeric complexes with CGRP.<sup>210</sup> It is also possible that CGRP–amylin heterodimers compete against amylin homodimers, thereby inhibiting fibril formation at high concentrations. Taken together, these models help contextualize the changes in CCS values from experiments and provide mechanistic insights on how CGRP might bind to amylin to reduce fibrillization, thereby contributing to the formation of disordered oligomers.

### **The amylin-CGRP mixture impairs insulin secretion but CGRP oligomers increase the Ins1 level**

As previously mentioned, it has been well-established that both amylin and CGRP independently inhibit insulin secretion. As we determined that heterooligomers of amylin and CGRPs exist in solution, the next step is to investigate their effect on insulin secretion using healthy mouse islets as the model. We chose single mouse islets as the biological system because multiple islets from the same animals can be collected and assayed under different conditions (amylin, CGRP, and the mixture). By analyzing single islets, we reduced the number of required pancreases (animals). To address islet heterogeneity, three islets per animal per condition were assayed, and a total of six animals were used. Furthermore, with LC-IMS-MS, multiple LC injections from a single islet can be performed to yield at least three technical replicates. Each analysis requires approximately 150 cells and can separate mouse Ins1 and Ins2 proteins as discussed below. In comparison to traditional ELISA, our method is less expensive and requires a fewer number of pancreases. Figure 4.16 illustrates a schematic of the workflow.

Upon LC-IMS-MS analysis of the single islets, we detected both mouse Ins1 and Ins2 (Figure 4.17A). Ins1 was detected at an observed mass of 5799.668 Da (0.276 ppm difference from the theoretical mass of 5799.666 Da),<sup>211</sup> and at four different charge states (Figure 4.17B) which were identified unambiguously with isotopic spacings. The  $[M+6H]^{6+}$  species at  $m/z$  967.619 is the most intense, and the other charge states were at  $m/z$  829.963 ( $z = +7$ ), 1060.943 ( $z = +5$ ), and 1451.182 ( $z = +4$ ). Similarly, Ins2 was

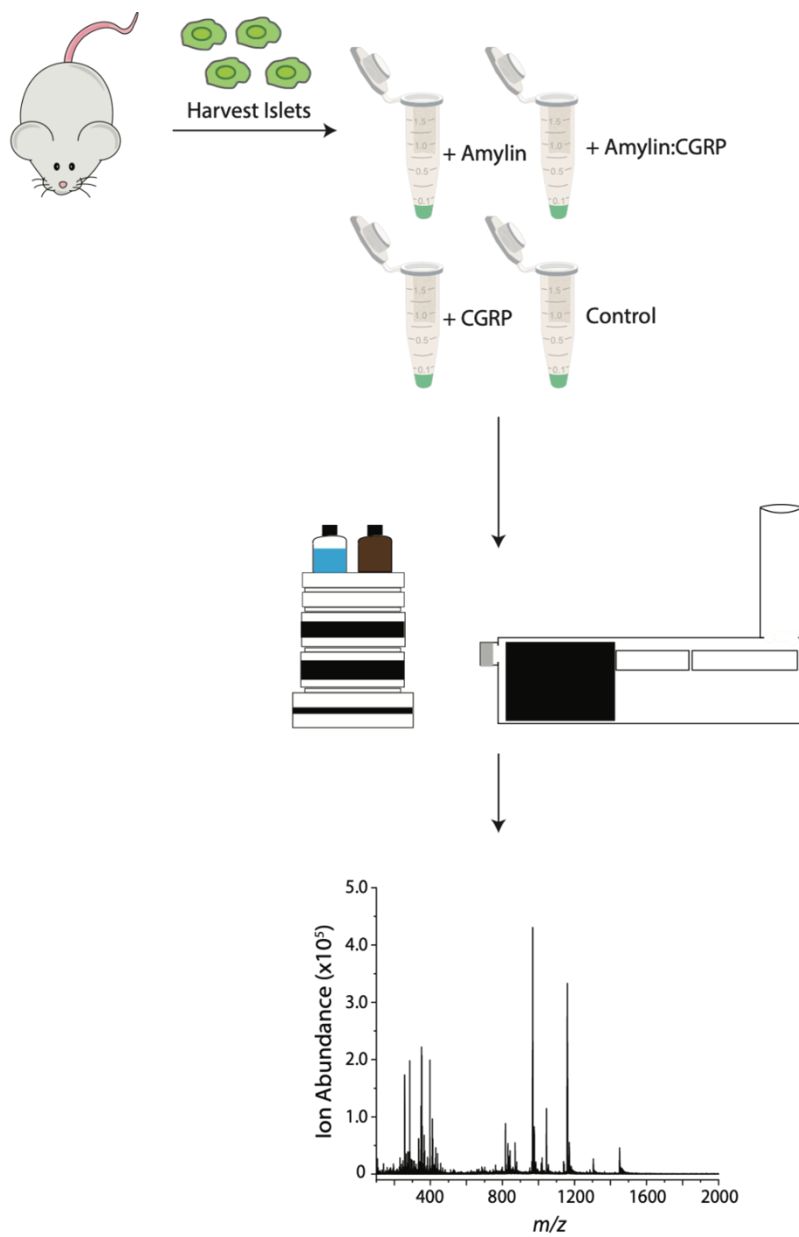


**Figure 4.16. Schematic of single islet analysis workflow.**

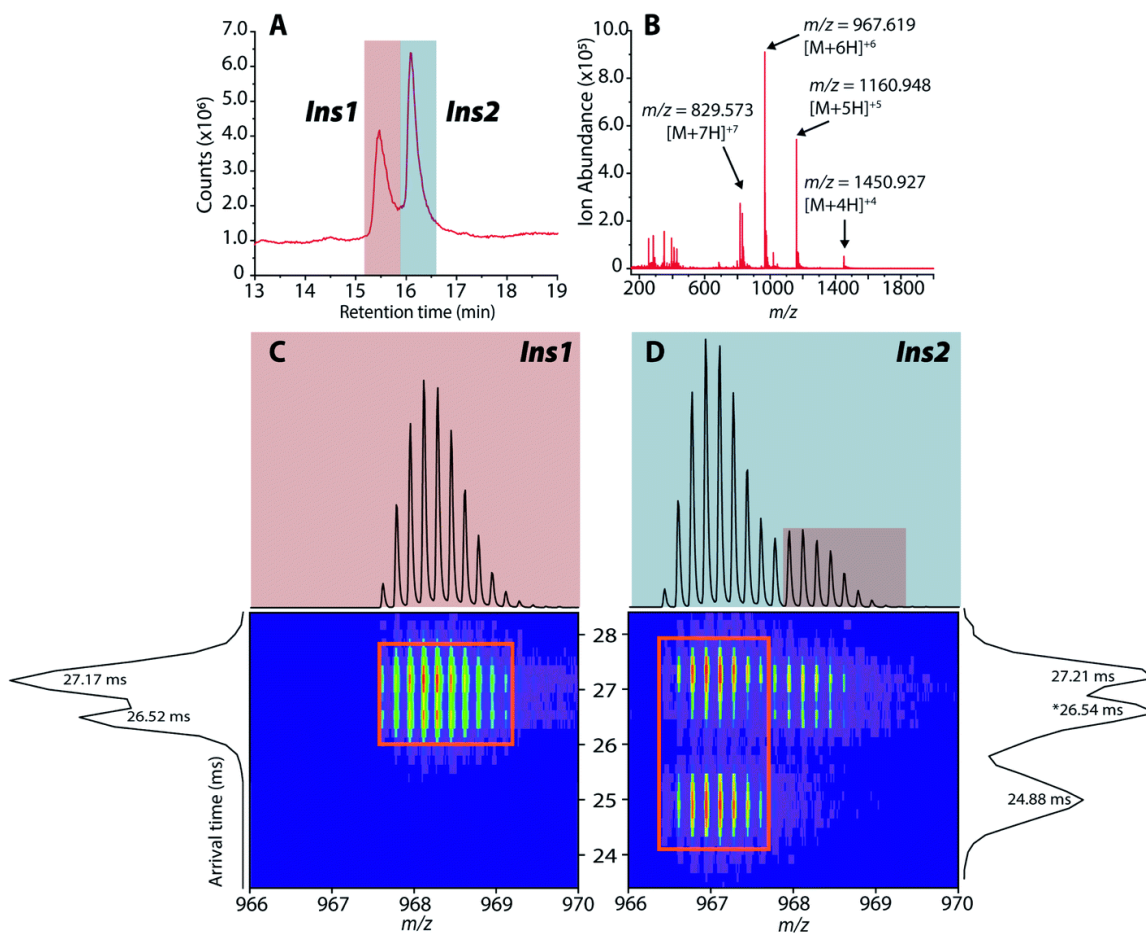


**Figure 4.17.** (A) Extracted LC chromatogram showing *Ins1* and *Ins2*. (B) Extracted partial LC-IMS-MS mass spectrum showing only *Ins1*. (C and D) 2D plots of arrival time vs.  $m/z$  of *Ins1* and *Ins2* at  $z = +6$ .

detected at an observed mass of 5793.5974 Da (2.261 ppm difference from the theoretical mass, 5793.6105 Da)<sup>212</sup>. In terms of CCSs, there is no reported experimental CCS of mouse insulins. However, there are reported values for human insulin by Bush and co-workers. The reported CCS of  $z = +4$  human insulin monomer is 772 Å<sup>2</sup>,<sup>208</sup> which agrees well with our experimental CCSs of Ins1 at the same charge state (CCS<sub>He,calibrated</sub> = 753 Å<sup>2</sup> and 783 Å<sup>2</sup>, CCSN2 = 902 Å<sup>2</sup> and 936 Å<sup>2</sup>).

The rodent (mouse and rat) insulin two-gene system is well reported. In mouse islets, expressions of Ins1 and Ins2 are similar (57:43% total insulin mRNA).<sup>213</sup> Ins1 differs from Ins2 by two amino acids in chain B (P9S and K29M) while other differences can be found at the proinsulin and gene levels.<sup>214</sup> Ins1 was shown to be more sensitive to glucose stimulation than Ins2,<sup>213</sup> and thus it had a more advantageous effect under scant food conditions. However, as the environment changes, Ins1 increased the risk of T1DM (insulin resistance), evidenced by nonobese diabetic (NOD) mice carrying only Ins1 gene developed T1DM as early as 10 weeks old.<sup>214</sup> On the other hand, mouse Ins2 is the homologue of human insulin, which exerts a protective effect against T1DM in NOD mice.<sup>215</sup> Therefore, in this work, the toxicity of amylin and the amylin:CGRP mixture, and their implications on T2DM should be evaluated based on the level of Ins2 rather than Ins1.

The 2D plots Figure 4.17C and D highlight the structural difference between Ins1 and Ins2. The arrival time at ~24–25 ms in panel D suggests that Ins2 either adopts a more compact conformation or is more prone to form oligomers. The biological assembly of Ins2 is a hexamer stabilized by Zn<sup>2+</sup>–histidine coordination.<sup>216</sup> Structural comparison of Ins1 and Ins2 is beyond the scope of this study, and will be a topic of future investigation.

Incubation of the islets with pure amylin and the amylin:CGRP mixture shows the lowest levels of insulin secretion, regardless of the peptide solution conditions (*i.e.*, “aged” vs. “fresh”; Figure 4.18). A one-way ANOVA analysis was performed to determine the statistical significance of our data followed by a Tukey test between each

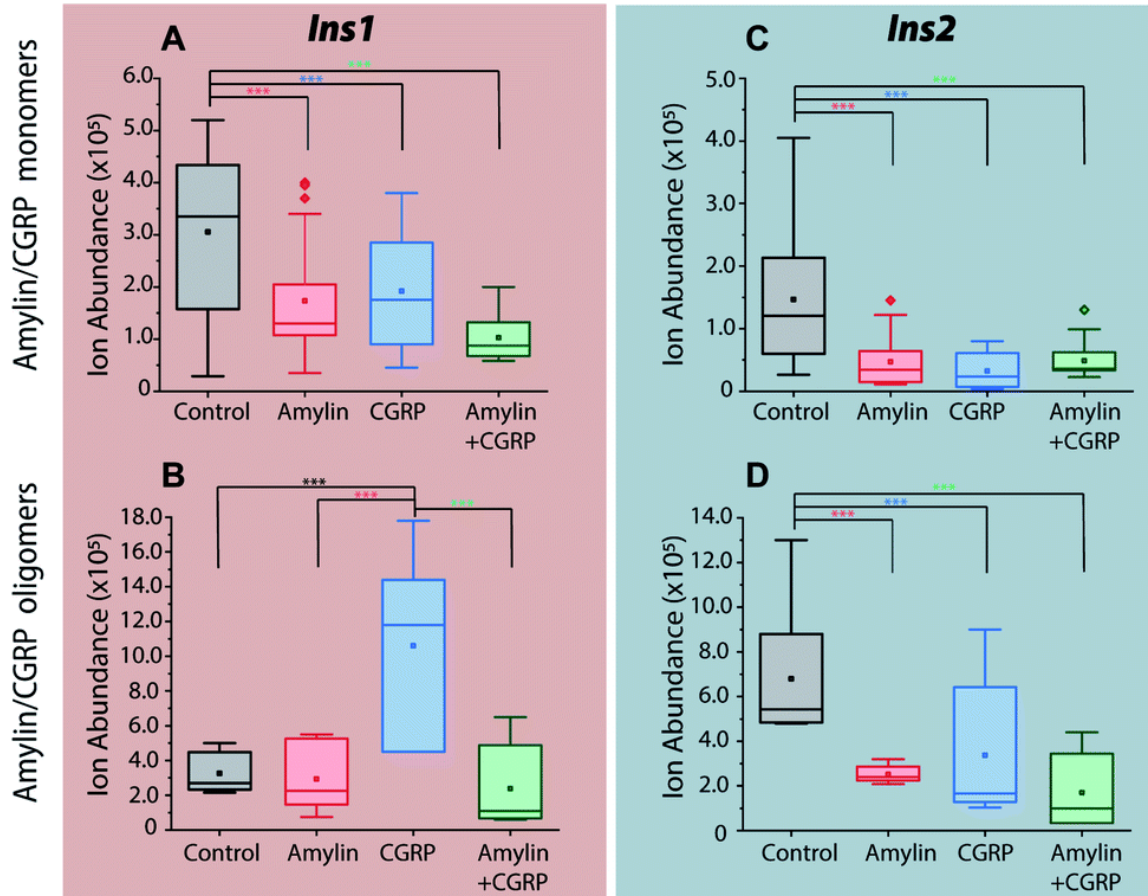


**Figure 4.18.** Comparisons of insulin signals obtained from single islets exposed to (A and C) media contain mostly amylin/CGRP monomers and (B and D) media contain mostly amylin/CGRP oligomers. \*\*\* indicates  $p < 0.05$ .

condition and the control. A qualitative comparison of amylin and CGRP oligomers in aged and fresh culture media is shown in Figure 4.19. The aged media contain significantly higher amounts of oligomers than the fresh media.

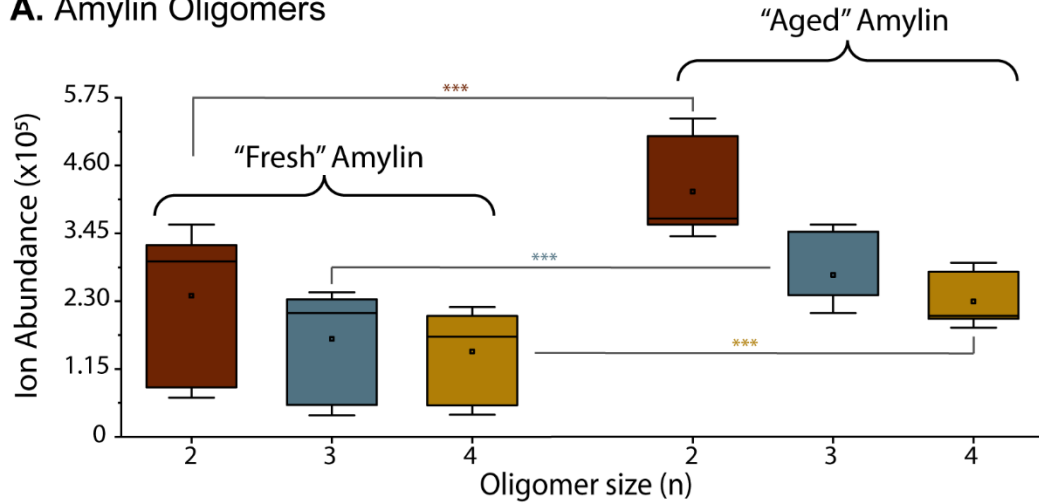
Interestingly, once the islets were placed under the conditions containing aged CGRP (more CGRP dimers), the Ins1 level increased, which contradicts the inhibitory effect of fresh CGRP. To our knowledge, this is the first time the dual effect of CGRP on insulin secretion was observed in isolated mouse islets, and this effect is specific to Ins1 and not Ins2. Similar effects were observed in animals with one insulin-gene system like dogs and pigs.<sup>175, 177</sup> Notably, Hermansen *et al.*<sup>175</sup> observed a similar trend to ours in which CGRP inhibits (dog) insulin secretion at low doses in a short time frame but increases it at high doses over a longer period. Unfortunately, earlier studies did not examine the possibility of how CGRP oligomers might behave differently from the CGRP monomer. The exact mechanism that allows for the dual effect of CGRP on Ins1 is not known.

In terms of Ins2, its inhibitory effect toward insulin secretion is consistent between the two media conditions. However, the effect is more profound in fresh CGRP. The data also confirm that CGRP is as effective as amylin in suppressing insulin secretion, but CGRP does not aggregate into fibrils.

### **CGRP in migraine and diabetes**

There have been several studies that have concluded that there is an inverse relationship between diabetes and migraine prevalence. Bonnet and coworkers<sup>167</sup> showed that women with active migraine had a lower risk of developing T2DM as opposed to women with no migraine history. Lund and co-workers<sup>217</sup> demonstrated that there is a markedly reduced prevalence of migraine among older patients with diabetes. These studies prompted us to investigate neuropeptides that are upregulated in migraine to determine their roles in islet chemistry and diabetes progression, and CGRP is one of those peptides. The elevated concentration of CGRP in migraine patients may exert a protective effect via suppressing insulin release. Templeman *et al.*<sup>218</sup> showed

### A. Amylin Oligomers



### B. CGRP Oligomers

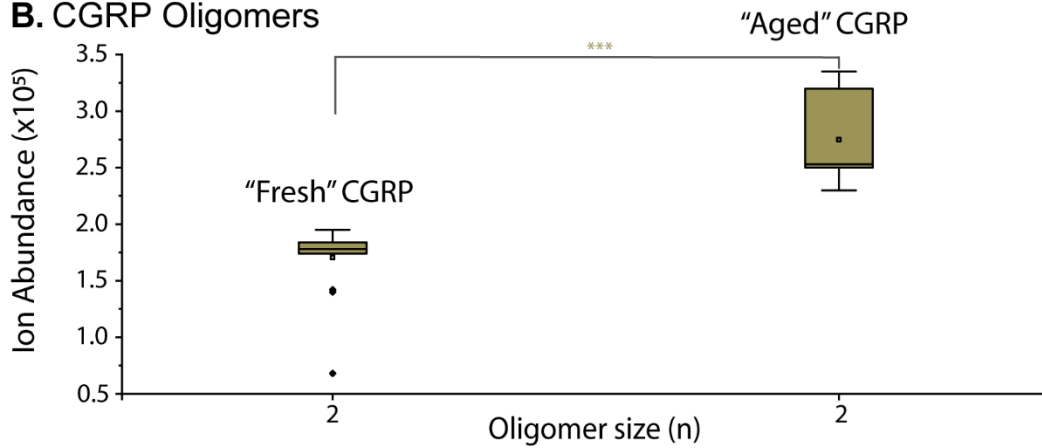


**Figure 4.19.** Comparisons of amylin and CGRP oligomers in “fresh” and “aged” incubating solutions. Amylin oligomers up to tetramers were detected, whereas CGRP oligomers up to dimers were detected in the two different solutions. The ion abundance of the oligomers increased in the aged peptides solutions. A paired t-test was performed to check for variation between the oligomers in the different solutions. \*\*\* indicates significant difference.

that decreased insulin led to lower fasting glucose and improved insulin sensitivity in aged mice. Thus, CGRP may play a critical role in regulating insulin release that can help lower the risk of diabetes.

Regarding toxicity of amylin, there are two possibilities for this peptide to generate negative effects on islet  $\beta$ -cells. The first is by direct toxicity, which our data do not support. The second is by activation of islet resident macrophages, which in turn produce inflammatory molecules in response to amylin uptake. Based on these two possibilities, our data are in line with those of Westwell-Roper et al.<sup>207</sup> who provided compelling evidence for amylin accumulation acting on the islet resident macrophage to promote inflammation-associated toxicity but not directly on the  $\beta$ -cells.

Our data show that CGRP could not regulate *Ins1* effectively, and interestingly NOD mice with only the *Ins1* gene develop T1DM as early as 10 weeks old.<sup>214</sup> Replacing *Ins1* with human insulin (the orthologue of mouse *Ins2*) protects the mice from diabetes.<sup>215</sup> To our knowledge, this is the first time that the dual effect of CGRP on insulin secretion has been investigated using isolated (mouse) islets. Moreover, because the biochemistry of mouse *Ins1* and *Ins2* is not the same, diabetes research using mouse models may need to distinguish these two isoforms.

We speculate that the use of some anti-CGRP medications for migraine treatment may lower the level of CGRP and increase the risk of developing T2DM. Specifically, some monoclonal CGRP antibodies which bind to CGRP<sup>219, 220</sup> could hinder CGRP function. This is based on the premise that mouse *Ins2* and human insulin behave similarly. Lowering the production of amylin while increasing the level of CGRP could be a viable approach toward T2DM prevention and treatment. Finally, we did not perform a clinical study to directly correlate migraine with diabetes, and this topic is beyond the scope of this paper. There have not been any clinical trials of this type. Nonetheless, we plan to further investigate the mechanism of CGRP on insulin secretion as well as other neuropeptides associated with migraine.

## 4.6 Conclusion

In summary, we studied the interactions that occur between amylin and CGRP in the context of T2DM. First, we show that the conformational preference of CGRP cannot be overridden by  $\text{Zn}^{2+}$  complexation. Such properties may prevent the extensive aggregation of CGRP. We also identified intermediate amylin:Zn states that could prevent aggregation at low  $\text{Zn}^{2+}$  concentrations. Second, we demonstrate that CGRP can hinder amylin fibrillization, leading to the formation of globular aggregates made predominantly of amylin. We show that amylin and the amylin:CGRP mixture do not show significant toxicity to pancreatic  $\beta$ -cells. However, amylin toxicity could be due to inflammation associated with islet resident macrophage as recently suggested by Westwell-Roper et al.<sup>207</sup> Since amylin is co-secreted with insulin, reducing the secretion of insulin is a strategy to lower amylin accumulation. Third, CGRP alone effectively reduces the secretion of Ins2, the orthologue of human insulin, but is less effective against Ins1. In fact, CGRP oligomers are able to stimulate the secretion of Ins1 which is linked to early diabetic development. Fourth, we demonstrate LC-IMS-MS of single islets as a viable and robust method to evaluate insulin secretion in the presence of exogenous peptides, which can become a platform to screen anti-aggregation drugs for T2DM. Finally, our study highlights the importance of surveying not only disease-implicated species, but also their probable co-assemblies that can arise in vivo, which provides a new look on disease etiology and novel insight into the diagnostics and treatments of diabetes.

## 5 The overlooked protective function of non-amyloidogenic A $\beta$ <sub>42</sub> on Tau pathology

This chapter will be adapted into a future publication with the proposed title:

*“The Overlooked Protective Function of Non-Amyloidogenic A $\beta$ <sub>42</sub> on Tau Pathology”*

Proposed authors:

Amber L. H. Gray, Thanh D. Do

All IMS-MS data were collected and analyzed by myself under the guidance of Thanh Do.

### 5.1 Abstract

The interplay between A $\beta$ <sub>42</sub> and Tau have long been a mystery for neurodegenerative research due to the temporal and spatial disconnect that the insoluble plaques of each serve within Alzheimer’s diseased brains. However, the consideration of small oligomeric species of each protein and how they interact with each other have surprisingly not been heavily highlighted in previous investigations, which may be due to the difficulty to isolate and study individual transient oligomers. Ion mobility spectrometry – mass spectrometry (IMS-MS) is one of the only experimental techniques that can capture this interplay and molecular crosstalk that can occur between A $\beta$ <sub>42</sub> and Tau before fibril formation. In this study, IMS-MS was used for the characterization of a Tau construct, termed Tau<sub>PHF43</sub>, which revealed that Tau has a propensity to form dimers cross-linked by disulfide bonds. The presence of disulfide bonding within Tau oligomers has been shown by previous investigations to be important to Tau pathology and toxicity. Interestingly, we observed that these disulfide-bonded Tau<sub>PHF43</sub> dimers, along with Tau fibrils, were disrupted in the presence of A $\beta$ <sub>42</sub>. Thus, we speculate that soluble A $\beta$ <sub>42</sub> may help serve as a neuroprotective factor that serves to break down toxic Tau species within the brain. Further experiments with cyclosporin A (CycA), a  $\beta$ -sheet disruptor, showed that this protective function of A $\beta$ <sub>42</sub> may result from its soluble

monomeric (or small oligomer) forms as opposed to A $\beta$ <sub>42</sub> fibrils. These results bring a new outlook to the mechanisms underlying Alzheimer's Disease.

## 5.2 Introduction

The self-assembly of proteins into amyloid fibrils are notorious in their association with human neurodegenerative disease.<sup>221-224</sup> Examples of amyloids include the proteins A $\beta$ <sub>42</sub> and Tau, both of which are thought to underlie the initiation and pathophysiology of Alzheimer's disease (AD).<sup>225-227</sup> However, despite the presence of insoluble A $\beta$ <sub>42</sub> plaques and Tau neurofibrillary tangles (NFTs) within the brains of those of Alzheimer's Disease, it is well accepted that it is instead soluble oligomers which precede fibril formation that are the toxic species of such disease.<sup>228-231</sup>

It has long been debated on how A $\beta$ <sub>42</sub> and Tau coincide *in vivo* to cause cellular death and damage exhibited in AD. Hardy and Higgins initially proposed the amyloid cascade hypothesis in the 1990s,<sup>232</sup> linking A $\beta$ <sub>42</sub> aggregation and amyloid plaques to a rise in intracellular Ca<sup>2+</sup> level as the major cause of cell death. The hypothesis was later modified to include the hyperphosphorylation of microtubule-associated Tau proteins (Tau) which induces the formation of NFTs in neurons of AD<sup>233, 234</sup> or other Tautopathies.<sup>235, 236</sup> Transgenic mice studies have supported the crosstalk between A $\beta$  and tau, showing that the level of NFTs dramatically increase in the presence of A $\beta$ <sub>42</sub>.<sup>237-239</sup>

However, one of the main issues underlying this "bullet" (A $\beta$ <sub>42</sub>) and "trigger" (Tau) hypothesis<sup>240</sup> results from the disconnect of spatial and temporal distribution of A $\beta$ <sub>42</sub> and Tau. Analysis have shown that the presence of NFTs precedes the A $\beta$ <sub>42</sub> amyloid plaques by approximately a decade.<sup>241</sup> Additionally, NFTs reside inside the neurons, while amyloid plaques accumulate as extracellular deposits. Despite this difference, it has been speculated that soluble A $\beta$ <sub>42</sub> may exist intracellularly,<sup>242</sup> specifically in the perikaryon of pyramidal cells in AD brains<sup>243</sup> Furthermore, physiological Tau can be

found within the extracellular space outside of neurons.<sup>244-246</sup> Therefore, soluble A $\beta$  and Tau have an ample opportunity to interact inside and outside of the cells.

To investigate this molecular cross-talk that occurs between soluble A $\beta_{42}$  and Tau, a technique that can separate and capture individual oligomeric species is required. Ion mobility spectrometry – mass spectrometry (IMS-MS) is an experimental technique that can isolate individual amyloid oligomers by obtaining arrival times of isobaric oligomeric species of the same mass-to-charge ratio ( $m/z$ ).

Surprisingly, there have not been many studies which have investigated the interactions between soluble A $\beta_{42}$  and Tau oligomers that may arise in the early disease state. To date, there is only one study, performed by the Bowers group, who investigated the interactions between short Tau and A $\beta$  model fragments with IMS-MS.<sup>247</sup> One facet of the current study aims to investigate how soluble A $\beta_{42}$  and Tau species may interact preceding the formation of tangles and fibrils. By understanding these initial interactions, we can gain insight into the mechanisms that occur to induce full-disease pathology. Furthermore, we investigated how cyclosporin A (CycA), a cyclic peptide known to interact with  $\beta$ -sheet and disrupt amyloid fibrils, interacts with both A $\beta_{42}$  and Tau.

Full-length Tau contains two N-terminal domains and a microtubule-binding region. The microtubule binding region consists of four repeats. Human Tau has six different isoforms, which differ by the number of two N-terminal domains and the presence of the second repeat (R2) within the microtubule binding region (Figure 5.1). The propensity of Tau to form paired helical filaments (PHF) depends on two five-residue sequences (PHF6; <sup>306</sup>VQIVYK<sup>311</sup> and PHF6\*; <sup>275</sup>VQIINK<sup>281</sup>) of which PHF6 is present within all Tau isoforms.<sup>248, 249</sup> These amyloid nucleating regions initiate PHF formation by promoting Tau's random-coil to a  $\beta$ -structure.<sup>248, 249</sup> The current study will specifically be utilizing a Tau construct, termed Tau<sub>PHF43</sub>, which exhibit similar aggregation behavior to the full-length Tau. Specifically, the formation of thin fibrils is promoted by dimerization and anionic factors such as heparin.<sup>248</sup>

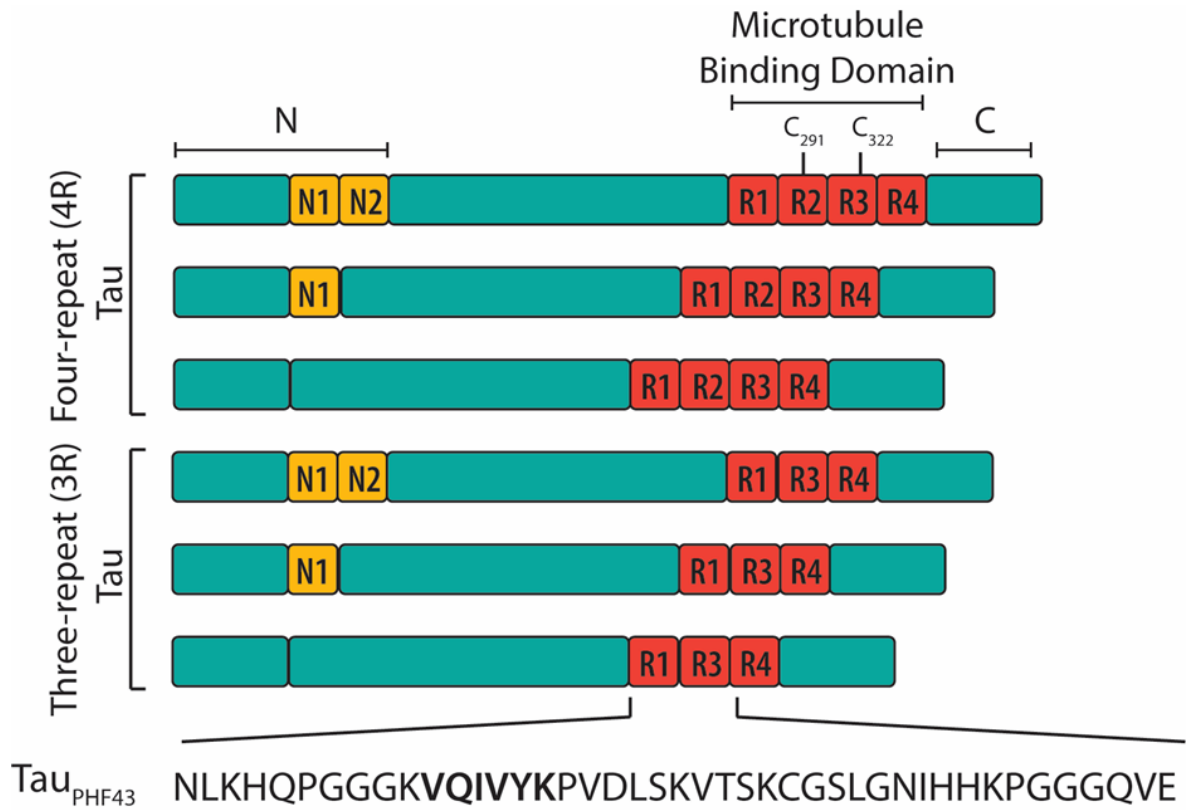


**Figure 5.1. Schematic of the six different human Tau isoforms. Note that full length Tau contains two cysteines. The sequence of Tau<sub>PHF43</sub> is given, which encompasses part of the repeat 1 domain (R1), all of R3, and part of R4.**

## 5.3 Materials and Methods

### Ion mobility spectrometry – mass spectrometry

A $\beta$ <sub>42</sub> (CAS 107761-42-2) was purchased from Genscript (Piscataway, NJ) at >95% purity. Tau<sub>PHF43</sub> was custom synthesized by Genscript at 95% purity, trifluoroacetic acid (TFA) was removed, and amino acid analysis was performed. The solid peptides were reconstituted with LC-MS water and diluted in 20 mM ammonium acetate solution to the desired concentrations. For the experiments involving heparin used heparin sodium salt derived from porcine intestinal mucosa (Sigma Aldrich) in which a 1.1 mM stock was made using LC-MS water and then diluted to the desired concentration. IMS-MS experiments were performed on an Agilent 6560 IMS-MS QTOF (Agilent, Santa Clara). Analytes were ionized by jet-stream ESI source and the ions were stored in an entrance funnel before subsequently pulsed into a nitrogen-filled drift cell in a multifold experiment. After traversing the drift cell, the ions entered an exit funnel and focused by a hexapole ion guide before reaching the QTOF mass analyzer. Data were collected over the course of 5.2 minutes with the drift cell voltages of  $\Delta V = 1449.5, 1350.6, 1250.2, 1150.5, \text{ and } 1050.3 \text{ V}$ . The pressure of the drift cell was maintained at 3.94 Torr with the pressure differences between the drift tube and trap funnel being approximately 300 mTorr.

## 5.4 Results and Discussion

### Tau<sub>PHF43</sub> dimers form via disulfide bonds

Figure 5.2 shows the representative mass spectra of the Tau<sub>PHF43</sub> construct (20  $\mu\text{M}$ ) which was freshly prepared ( $t = 0 \text{ hours}$ , panel A) and after 24 hours (panel B). The mass spectral peaks are annotated with their nominal oligomer-to-charge ratio ( $n/z$ ), or the smallest oligomer size-to-charge state that can be represented by that peak. The mass spectral peaks containing isobaric species (i.e., species with the same  $m/z$ ) are

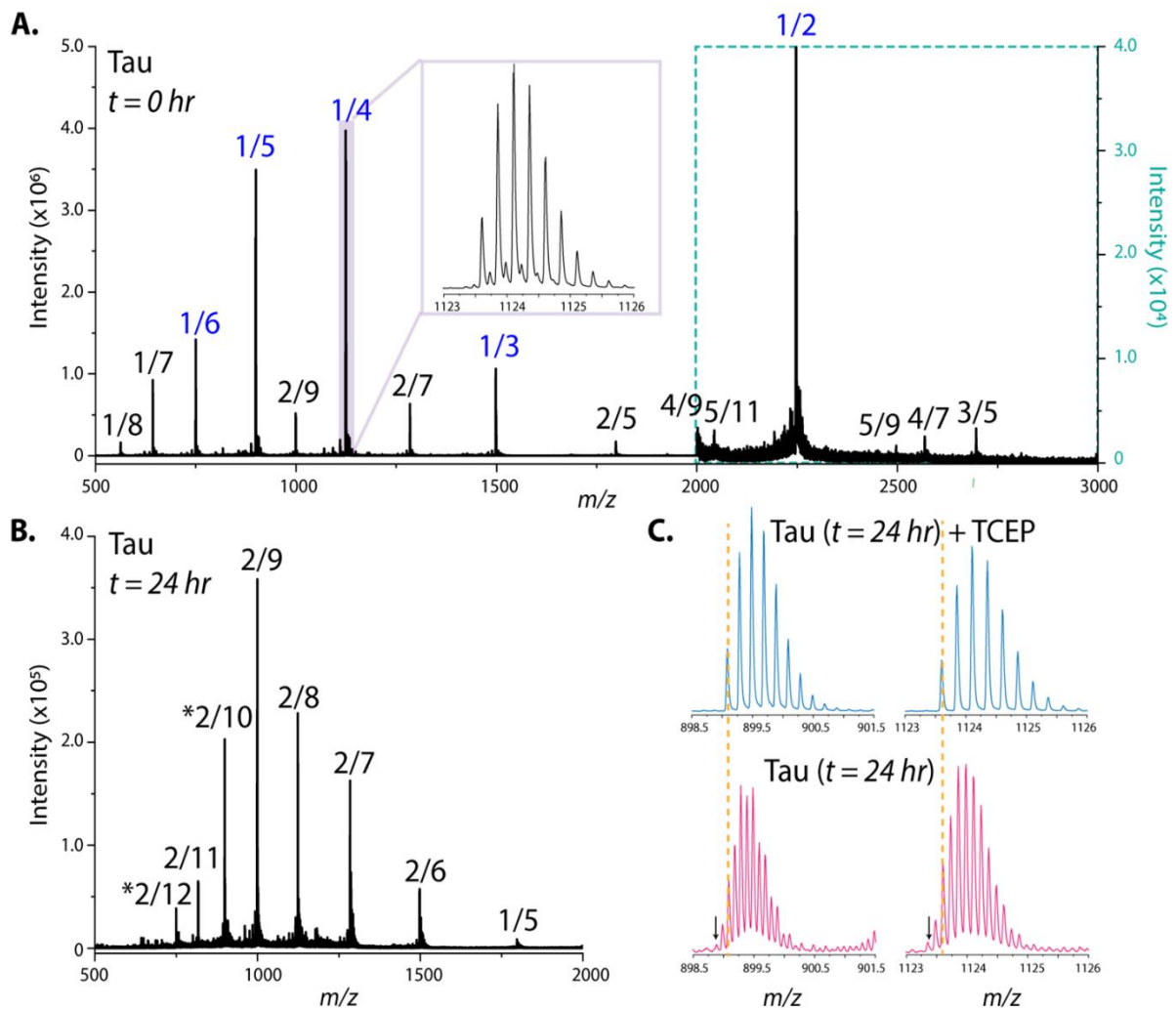


**Figure 5.2.** (A) Representative positive mode-ESI mass spectrum of  $\text{Tau}_{\text{PHF43}}$  freshly prepared ( $t = 0$ ). The inset reveals isobaric species that can be hidden underneath the same mass spectral peak. Peaks labeled blue denote isobaric species. (B) Mass spectrum of aged  $\text{Tau}_{\text{PHF43}}$  ( $t = 24$  hours). (C) Mass spectral peaks of aged  $\text{Tau}_{\text{PHF43}}$  + tris(2-carboxyethyl)phosphine (TCEP) (top panels) and aged  $\text{Tau}_{\text{PHF43}}$  (bottom panels). The black arrows are pointing towards the monoisotopic peak which was used to determine the observed monoisotopic  $m/z$ .

highlighted in blue. For example, the inset in A shows a mass spectral peak containing primarily a monomer with  $z = +4$ , however a dimer with  $z = +8$  ( $n/z = 2/8 = 1/4$ ) can be observed as well. The largest oligomer of Tau<sub>PHF43</sub> observed in the experiment is a pentamer.

When Tau<sub>PHF43</sub> (20  $\mu$ M) was allowed to set out at room temperature for 24 hours (i.e., aged Tau<sub>PHF43</sub>), the dimers become the predominant species within the mass spectrum (Figure 5.2B). For example, within the fresh Tau<sub>PHF43</sub> sample,  $n/z = 1/4$  was the dominant species. However, in the aged Tau<sub>PHF43</sub> sample it can be observed that the  $n/z = 2/9$  species now has the largest relative intensity compared to other species within the mass spectrum. Along a similar note, monomer species which contained isobaric dimers (see inset in Figure 5.2A) now only contained a dimer (unless marked with an asterisk in Figure 5.2B).

The propensity of Tau<sub>PHF43</sub> to form dimers appears to be controlled by the formation of intermolecular disulfide bonds. Tau<sub>PHF43</sub> contains one cysteine residue in the 27<sup>th</sup> position of the peptide. Each cysteine contains a sulfhydryl group, and under mild oxidizing conditions (such as air oxidation) a disulfide bridge can form between two Tau<sub>PHF43</sub> monomers, resulting in a dimer.

To show that Tau<sub>PHF43</sub> dimers are indeed formed via disulfide bridges, we exposed Tau<sub>PHF43</sub> dimers to reducing conditions. Specifically, we added 50  $\mu$ M tris(2-carboxyethyl)phosphine (TCEP), a reducing agent, to 20  $\mu$ M aged Tau<sub>PHF43</sub>. Figure 5.2C shows example mass spectral peaks of  $n/z = 1/4$  and  $n/z = 1/5$ . The top two mass spectral peaks (blue) are of aged Tau<sub>PHF43</sub> with TCEP, and the bottom two mass spectral peaks (pink) are of aged Tau<sub>PHF43</sub>. Based on isotopic spacing alone, only monomers can be observed within the sample containing TCEP. However, dimers are primarily observed within the sample containing only Tau<sub>PHF43</sub>.

To further support our argument, we noted that the monoisotopic peaks of the monomers within the Tau<sub>PHF43</sub> + TCEP samples appear at a slightly higher  $m/z$  than the dimers of the aged Tau<sub>PHF43</sub> sample. If the Tau<sub>PHF43</sub> dimers were formed without a

disulfide bond, then both species would have the same monoisotopic mass. To show that the presence of the disulfide bond was not an isolated instance of the dimer with  $z = +8$  and  $+10$ , we obtained the theoretical masses of Tau<sub>PHF43</sub> dimers with and without disulfide bond and calculated the theoretical  $m/z$  at various charge states. The theoretical  $m/z$  was then compared to the observed  $m/z$  to deduce if each dimer contains a disulfide bond or not. Table 5.1 shows that the listed dimers observed within the Tau<sub>PHF43</sub> sample contained disulfide bonds.

Full-length, 4R Tau isoforms contains two cysteines at positions 291 (located in the second repeat domain) and 322 (located in the third repeat domain) (Figure 5.1); therefore, the cysteines can participate in both intra- and intermolecular disulfide bridges.<sup>250-252</sup> It has been proposed that the two cysteines within 4R Tau are in different chemical environments and contribute differently to Tau's pathogenicity;<sup>253</sup> Cys-322 is embedded within the core of Tau fibrils and Cys-291 is located within the disordered region of the protein, forming the "fuzzy coat" of the fibrils.<sup>254</sup> Cys-322 (which is the cysteine in Tau<sub>PHF43</sub>) plays a dominant role in regulating tau stability.<sup>253</sup> Interestingly, previous studies showed that the presence of intermolecular disulfide bonds facilitates the aggregation of Tau.<sup>255-259</sup> Mandelkow and coworkers demonstrated that PHFs did not assemble when Tau contained two cysteines – possibly due to the formation of *intramolecular* disulfide bonds which stabilizes monomeric Tau.<sup>255, 260</sup> Disulfide bond formation has also been implicated in facilitating the aggregation of other amyloidogenic proteins, such as  $\beta_2$ -microglobulin.<sup>261</sup>

Furthermore, as the formation of disulfide bonds is facilitated by oxidizing environments, oxidative stress is a well-known event to occur within AD.<sup>262</sup> Saito et al. observed that both cysteines contribute to the accumulation of Tau under oxidative environments, suggesting that disulfide bond formation of Tau may contribute to disease pathology,<sup>263</sup> including the pathological spreading of Tau's NFTs to spread from one neuron to another as shown for K18-P301L (a tau isoform that encompass all four repeats but does not contain the N-terminal domain).<sup>264</sup>

**Table 5.1. Theoretical and observed monoisotopic  $m/z$  of Tau<sub>PHF43</sub> oligomers. The bold numbers under the theoretical  $m/z$  column correspond to the value that is closest to the observed  $m/z$ .**

| Condition           | Oligomer size | Charge State | Observed $m/z$ | Theoretical $m/z$ | Conclusion                |
|---------------------|---------------|--------------|----------------|-------------------|---------------------------|
| Disulfide bond      | 2             | 4            | 2246.18        | 2245.70           | No disulfide bond present |
| No disulfide bond   |               |              |                | <b>2246.21</b>    |                           |
| Disulfide bond      | 2             | 5            | 1796.76        | <b>1796.76</b>    | Disulfide bond present    |
| No disulfide bond   |               |              |                | 1797.16           |                           |
| Disulfide bond      | 2             | 6            | 1497.47        | <b>1497.47</b>    | Disulfide bond present    |
| No disulfide bond   |               |              |                | 1497.80           |                           |
| Disulfide bond      | 2             | 7            | 1283.68        | <b>1283.69</b>    | Disulfide bond present    |
| No disulfide bond   |               |              |                | 1283.98           |                           |
| Disulfide bond      | 2             | 8            | 1123.35        | <b>1123.35</b>    | Disulfide bond present    |
| No disulfide bond   |               |              |                | 1123.61           |                           |
| Disulfide bond      | 2             | 9            | 998.65         | <b>998.65</b>     | Disulfide bond present    |
| No disulfide bond   |               |              |                | 998.87            |                           |
| Disulfide bond      | 2             | 10           | 898.89         | <b>898.88</b>     | Disulfide bond present    |
| No disulfide bond   |               |              |                | 899.09            |                           |
| Disulfide bond      | 3             | 7            | 1925.46        | 1925.17           | No disulfide bond present |
| No disulfide bond   |               |              |                | <b>1925.46</b>    |                           |
| Disulfide bond      | 3             | 8            | 1684.89        | 1684.65           | No disulfide bond present |
| No disulfide bond   |               |              |                | <b>1684.91</b>    |                           |
| Disulfide bond (x1) | 4             | 9            | 1996.73        | 1996.51           | No disulfide bond present |
| Disulfide bond (x2) |               |              |                | 1996.29           |                           |
| No disulfide bond   |               |              |                | <b>1996.74</b>    |                           |

### **Disulfide bonds are not essential for the formation of higher-ordered Tau<sub>PHF43</sub> oligomers**

From Figure 5.2A, oligomers of Tau<sub>PHF43</sub> up to pentamer can be detected. As discussed previously, many studies have speculated the importance of disulfide bonds within Tau pathology and aggregation, and we showed that Tau<sub>PHF43</sub> dimers can form via intermolecular disulfide bonding. Since Tau<sub>PHF43</sub> contains only one cysteine residue, if disulfide bonds were present within a trimer, only one disulfide bond would be present. Likewise, a Tau<sub>PHF43</sub> tetramer could potentially contain up to two disulfide bonds. Both 3R (one cysteine per monomer) and 4R Tau (two cysteines) can aggregate and form NFTs, although the 3R isoform is slower.<sup>265</sup> It has also been suggested that oligomers grow through disulfide bond cross-linking.<sup>266</sup> Some tauopathies, such as Pick's Disease, exhibit inclusions that are composed up of solely 3R Tau fibrils.<sup>267</sup>

Of note, each 3R Tau monomer can neither form intramolecular disulfide bond nor grow larger than dimer via solely disulfide crosslinking. In other words, the growth of large oligomers of 3R Tau (and Tau<sub>PHF43</sub> here) should not rely on disulfide bond formation. On the other hand, the 4R Tau with two cysteines per monomer can propagate to large oligomers via disulfide crosslinking (Figure 5.3). It has been shown that the 3R isoform causes more profound axonal transport defects and locomotor impairments while the 4R isoform leads to greater neurodegeneration.<sup>268</sup> The difference in disulfide bond formation of 3R and 4R isoforms may contribute to isoform-specific mechanisms of tau-toxicity and the differential susceptibility of brain regions to different tauopathies.

To demonstrate that the formation of Tau<sub>PHF43</sub> oligomers is indeed independent from disulfide-bonded dimers, we investigated the composition of the Tau<sub>PHF43</sub> higher-ordered oligomers. Figure 5.4A shows the mass spectral peaks corresponding to trimers with  $z = +7$  and  $+8$  and tetramer with  $z = +9$ . The monoisotopic peak within each mass spectral peak is highlighted for each species. In a similar manner as previously, the observed monoisotopic  $m/z$  of the two trimers were compared to

|                      | Dimer | Trimer | Tetramer |
|----------------------|-------|--------|----------|
| Zero Disulfide Bonds |       |        |          |
| One Disulfide Bond   |       |        |          |
| Two Disulfide Bonds  |       |        |          |

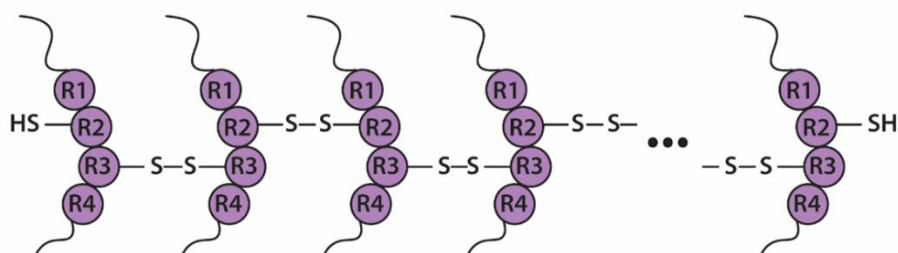


**Figure 5.3. Schematic of possible disulfide bond formation present within Tau<sub>PHF43</sub> dimers, trimers, and tetramers. Noncovalent bonds (e.g. hydrogen bonding) are represented by a dashed line. Note that the position of the noncovalent bond is trivial. The bottom panel shows speculative disulfide bonding of large 4R Tau oligomers.**

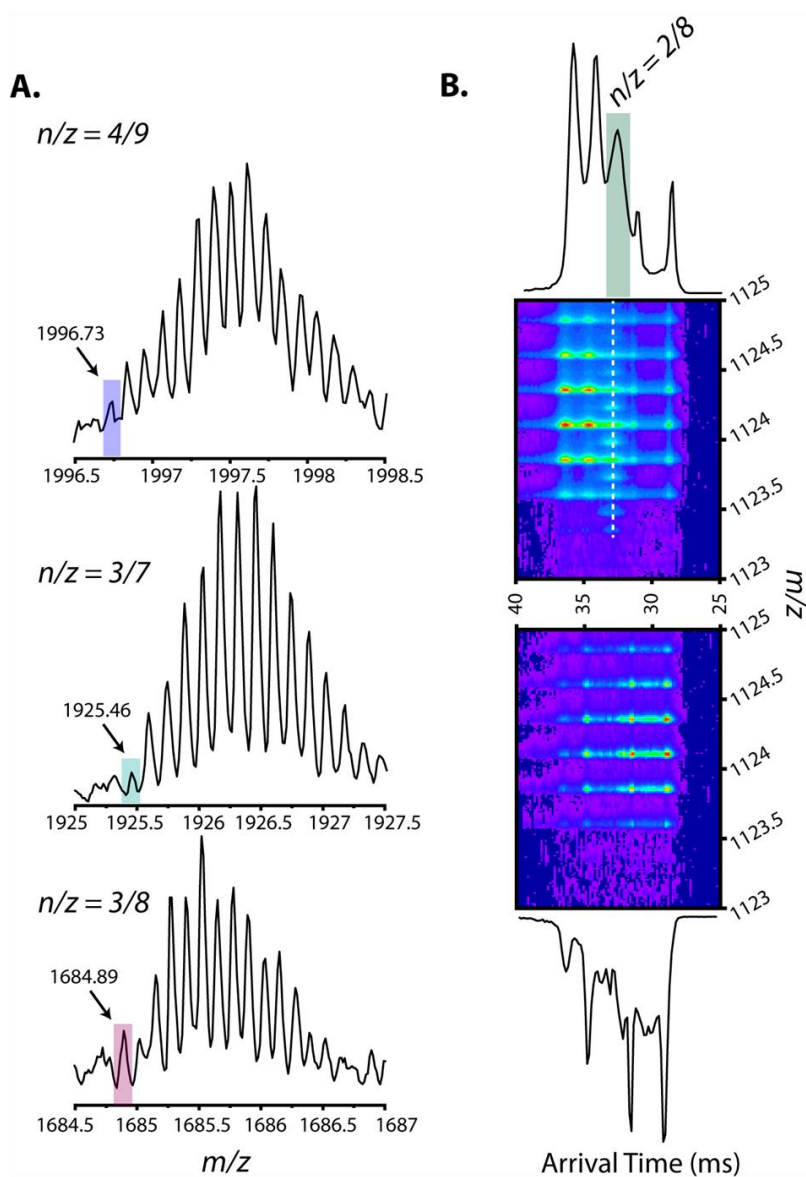


**Figure 5.4. (A) Mass spectral peaks of higher-ordered Tau<sub>PHF43</sub> oligomers. The monoisotopic peak is labeled. (B) 2D ion mobility plots of Tau<sub>PHF43</sub>  $n/z = 1/4$  by itself (top panel) and in the presence of TCEP (bottom panel).**

the theoretical  $m/z$  of a trimer containing either zero or one disulfide bond. Likewise, the observed monoisotopic  $m/z$  of the tetramer was compared to the theoretical  $m/z$  of a tetramer containing either zero, one, or two disulfide bonds. Table 5.1 shows these comparisons. It was observed that  $m/z$  of each trimer and tetramer matches the closest with species that contain no disulfide bonds. The pentamer was not analyzed due to the mass spectral peak being poorly resolved.

Figure 5.4B shows the 2D ion mobility plot of  $\text{Tau}_{\text{PHF43}}$   $n/z = 1/4$ . An isobaric dimer with  $z = +8$  can be observed along with the monomers. However, the isotopic spacing of the two shortest arrival time features (to the right of the dimer) are of a monomer with  $z = +4$ . This is inconsistent with ion mobility theory because higher charge state species should traverse the drift cell faster than lower charge state species, regardless of size. Therefore, the two shortest arrival time features seen within the 2D ion mobility plot of  $\text{Tau}_{\text{PHF43}}$  should have a charge state larger than  $z = +8$  to have a faster arrival time than the dimer. One possible explanation could be a phenomenon called post-ion mobility dissociation. Post-IM dissociation is where a weakly-bound species (such as amyloid oligomers) can successfully traverse the drift cell intact but dissociate before arriving to the mass analyzer. As a result of this phenomenon, the two shortest arrival time features must result from the dissociation of a large oligomer (e.g.,  $n/z = 3/12, 4/16, 5/20$ , etc.). A comprehensive overview of post-IM dissociation and its implication in interpreting amyloid self-assembly IMS-MS data has been published elsewhere.<sup>84</sup>

The  $\text{Tau}_{\text{PHF43}}$  dimer exhibited in Figure 5.4B contained a disulfide bond (Table 5.1). This is further supported when TCEP is added to the sample (Figure 5.4B, bottom) since it disrupts the dimer. However, although the dimer is disrupted the two prominent shortest arrival time features are still present. Since we have determined that these shortest arrival time features must be of an oligomer larger than a dimer, if they were stabilized by disulfide bonding, they should be disrupted upon addition of TCEP.

Mandelkow and coworkers determined that Cys-322 within peptide derivatives of  $\text{Tau}_{\text{PHF43}}$  does not directly promote the assembly of PHFs by itself but may enhance local

Tau concentration by facilitating  $\beta$ -sheet interactions.<sup>248</sup> The dimers may then act as a seed to nucleate further 3R Tau oligomerization. However, this raises the question of why the trimer nor tetramer doesn't exhibit any disulfide bonds. In other words, what mechanism occurs in which Tau<sub>PHF43</sub> dimer loses the disulfide bond upon the addition of other Tau<sub>PHF43</sub> units? To entertain this question, we searched within the mass spectrum of Tau<sub>PHF43</sub> under reduced conditions to find evidence of dimers without disulfide bonds. At  $m/z = 2245$ , an isobaric dimer species can be seen under reducing conditions (Figure 5.5). This alludes that there are two types of dimers that can be formed by Tau<sub>PHF43</sub>: oxidized (with disulfide bonds) and reduced (without disulfide bonds). This is not unheard of in biological complexes. For example,  $\beta_2$ -microglobulin forms two morphologically distinct forms of amyloid fibrils that presents themselves as thin and flexible filaments or needle-like amyloid fibrils.<sup>261, 269-272</sup> The type of  $\beta_2$ -microglobulin fibrils that are formed are determined early in the oligomerization pathway by reduced or oxidized  $\beta_2$ -microglobulin dimers.<sup>273</sup>

The presence of these two distinct dimers of Tau<sub>PHF43</sub> may help explain why 3R Tau aggregates much slower than its 4R counterpart (and thus why the Tau<sub>PHF43</sub> trimer and tetramer does not have a disulfide bond). Oligomerization of higher ordered oligomers may occur through the tau dimer with no disulfide bonds. Since Tau has a propensity to form disulfide bonded dimers, especially under oxidizing conditions typically present within neurodegeneration, this leaves a lower amount of Tau dimers that are bonded through noncovalent interactions as opposed to covalent disulfide bonds. This mechanism also explains the importance of disulfide bonds in the formation of higher-ordered oligomers for 4R tau species that has previously been speculated.<sup>266</sup>

#### **A $\beta_{42}$ disrupted Tau<sub>PHF43</sub> disulfide bonded dimers and formed A $\beta_{42}$ :Tau<sub>PHF43</sub> heterooligomers**

In AD, both Tau and A $\beta_{42}$  fibrils serve as a hallmark of disease. Therefore, it is logical to investigate how oligomers of Tau<sub>PHF43</sub> interacts with oligomers A $\beta_{42}$  to investigate

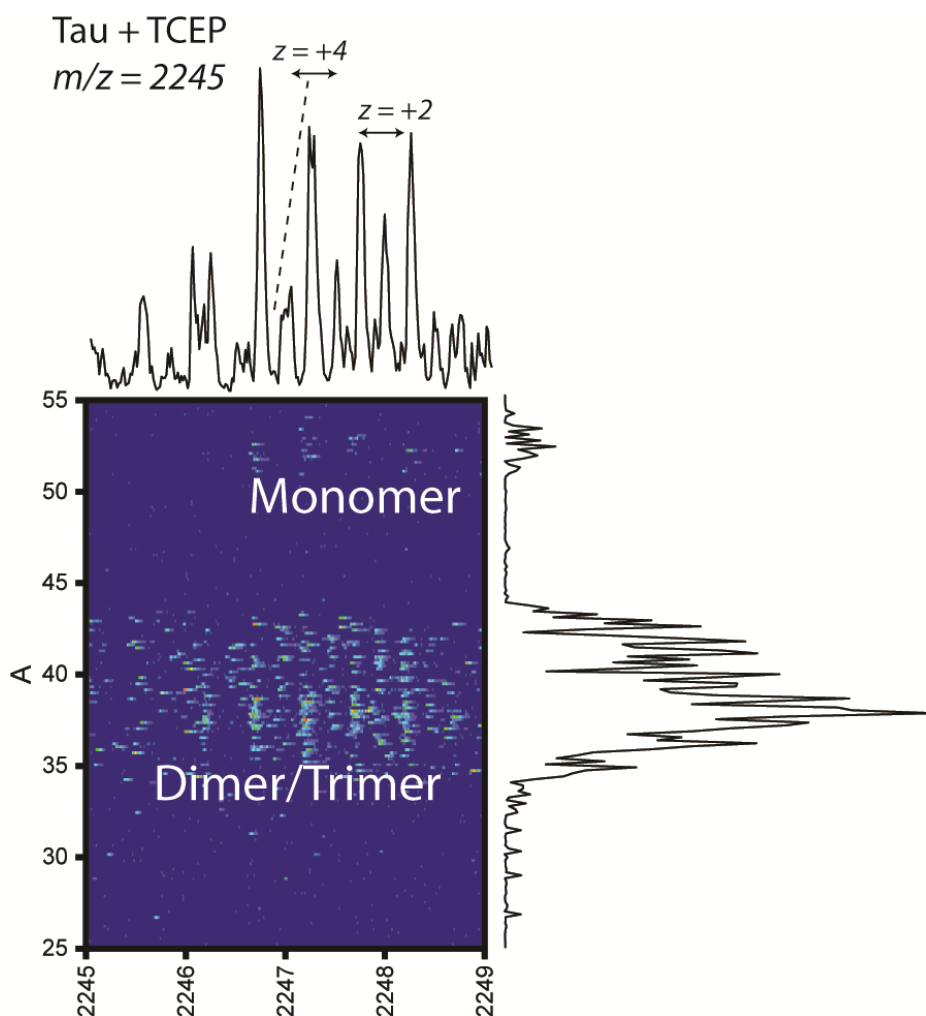


**Figure 5.5. 2D ion mobility plot at  $m/z = 2245$  of Tau<sub>PHF43</sub> + TCEP (positive mode-ESI). Note that an unresolved dimer/trimer can be seen within the 2D ion mobility portion. Observation of isotopic spacing within the mass spectral peak confirms the unresolved 2D ion mobility plot does contain a dimer.**

interactions that may occur in the early-disease state. Figure 5.6A shows the representative mass spectrum of aged Tau<sub>PHF43</sub> and A $\beta$ <sub>42</sub>. Only Tau<sub>PHF43</sub> mass spectral peaks are labeled by nominal  $n/z$ . Although Tau<sub>PHF43</sub> has been sitting out at room temperature for 24 h and is air oxidized, there are no Tau<sub>PHF43</sub> disulfide-bonded dimeric species present in the presence of A $\beta$ <sub>42</sub>. Examination of isotopic spacing of isobaric dimers observed in Figure 5.2B determined that these peaks represented only monomers (Figure 5.6B).

Figure 5.6C shows the 2D ion mobility plot of  $n/z = 1/2$  Tau<sub>PHF43</sub> (top panel) and Tau<sub>PHF43</sub> with and without A $\beta$ <sub>42</sub> (bottom panel). In pure Tau<sub>PHF43</sub>, the isotopic spacings of the  $n/z = 1/2$ ,  $2/4$ , and  $3/6$  can be observed. However, in the presence of A $\beta$ <sub>42</sub>, the intensities of the oligomer species are significantly lower. The feature at shorter arrival time than the dimer has an observed isotopic spacing of  $z = +2$  but should have the drift time of a larger species (i.e., trimer with  $z = +6$ ) that underwent post-IM dissociation to form a monomer with  $z = +2$  and other smaller species. This implied that larger oligomers of Tau<sub>PHF43</sub> may be unstable in the presence of A $\beta$ <sub>42</sub>.

Because large Tau<sub>PHF43</sub> oligomers, including the disulfide-bonded dimers, are susceptible to dissociation in the presence of A $\beta$ <sub>42</sub>, then we speculate that Tau<sub>PHF43</sub> may have a stronger affinity for A $\beta$ <sub>42</sub> than for itself. This is not of pure speculation - Figure 5.7 shows examples of Tau<sub>PHF43</sub>:A $\beta$ <sub>42</sub> complexes, and other studies have documented similar results. Do et al. found that Tau<sub>278-284</sub> and A $\beta$ <sub>25-35</sub> formed multiple Tau:A $\beta$  complexes in solution.<sup>247</sup> An earlier study by Bowers and co-workers using A $\beta$  and Tau fragments also showed an abundance of Tau<sub>278-284</sub> monomers in solution in the presence of A $\beta$ <sub>25-35</sub> despite the heterooligomers formation.<sup>247</sup> McGeer and coworkers showed that the binding affinity of A $\beta$  for Tau was almost 1000-fold higher than Tau for itself.<sup>242</sup>

Several studies suggested that the interaction of A $\beta$ <sub>42</sub> and Tau is a major factor in contributing to AD. For example, it has been shown that sub-toxic doses of soluble Tau and A $\beta$  affect memory and long-term potentiation within neurons once combined.<sup>274</sup> McGeer and coworkers speculated that soluble A $\beta$ :Tau binding may promote Tau

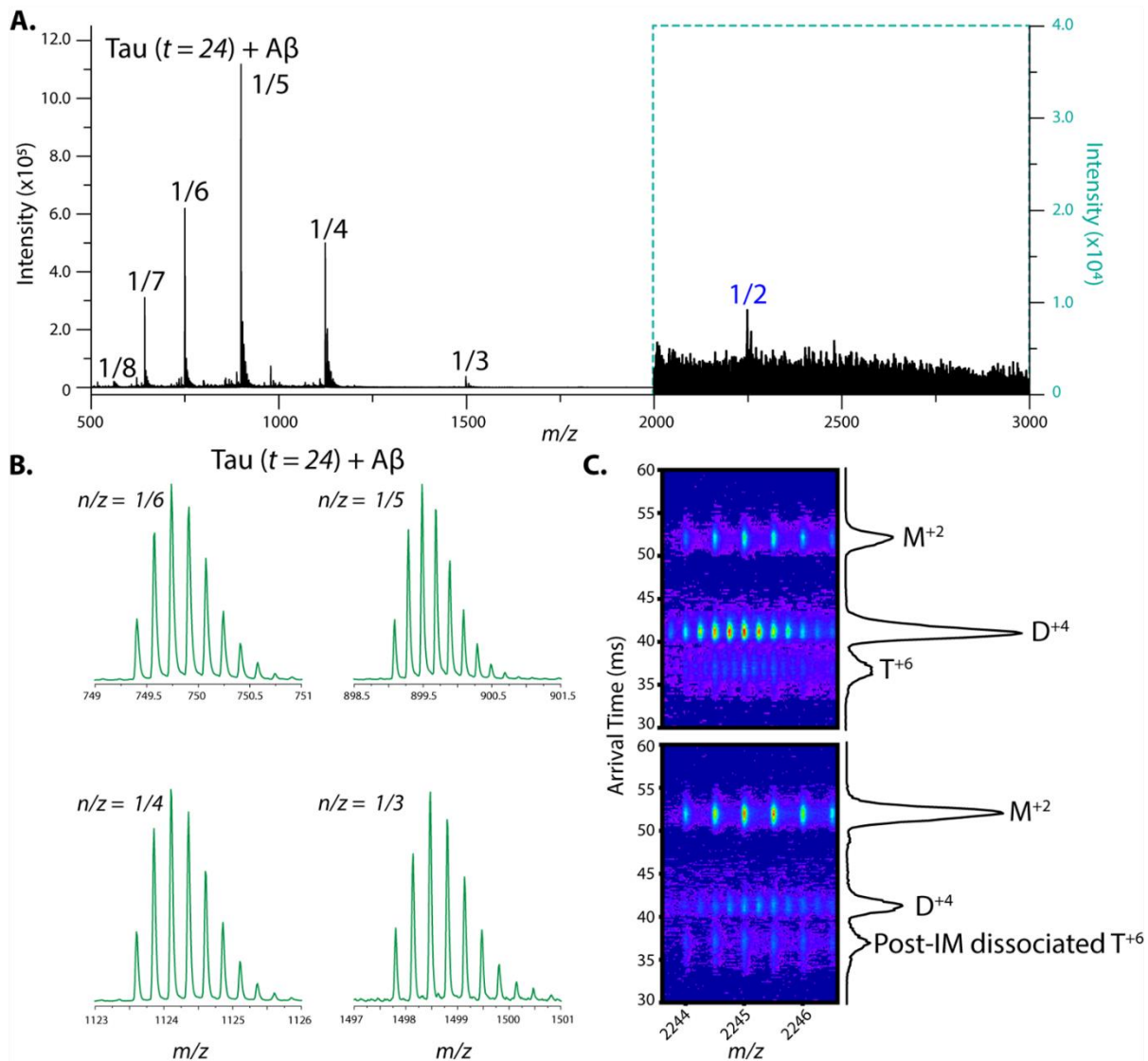


**Figure 5.6. (A)** Representative positive mode-ESI mass spectrum of 10  $\mu\text{M}$ :10  $\mu\text{M}$   $\text{A}\beta_{42}$ : $\text{Tau}_{\text{PHF43}}$ . Only  $\text{Tau}_{\text{PHF43}}$  peaks are labeled. The peak in blue denotes higher-ordered species underneath the mass spectral peak. **(B)** Examples of isobaric mass spectral peaks of  $\text{Tau}_{\text{PHF43}}$  in the presence of  $\text{A}\beta$ . Note that there is no evidence of  $\text{Tau}_{\text{PHF43}}$  dimers as in Figure 5.2C, only monomers. **(C)** 2D ion mobility plots (negative mode-ESI) of  $\text{Tau}_{\text{PHF43}}$  by itself (top panel) and in the presence of  $\text{A}\beta_{42}$  (bottom panel).

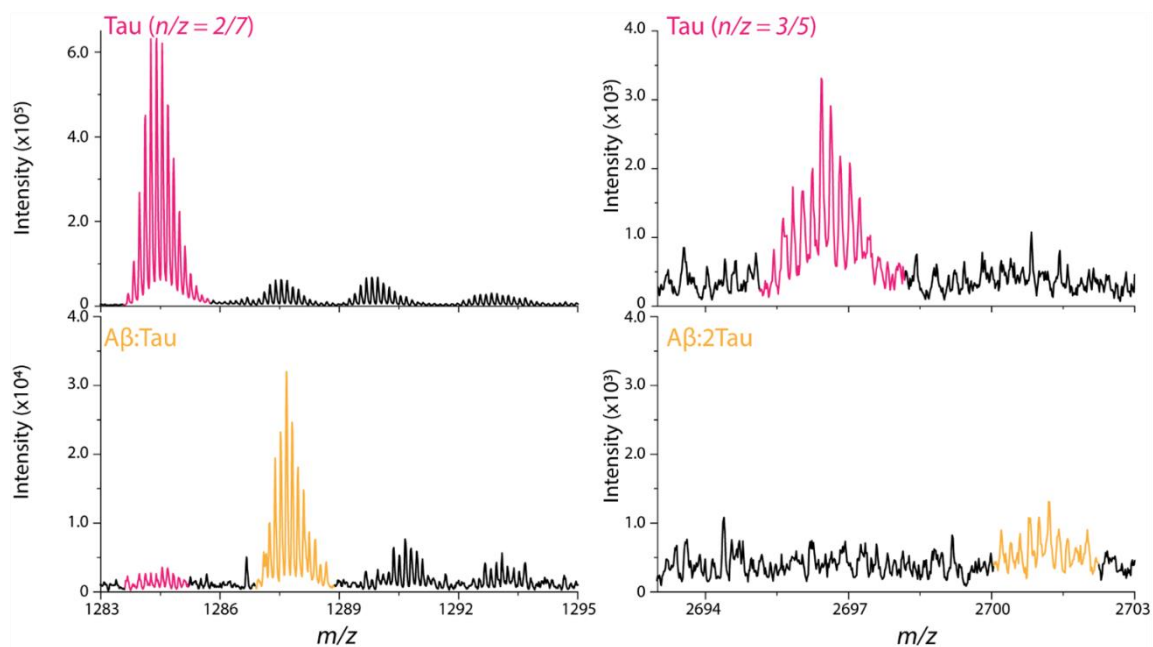


**Figure 5.7. Tau<sub>PHF43</sub> oligomers (top panels) and Aβ<sub>42</sub>:Tau<sub>PHF43</sub> heteroligomers (bottom panels).**

phosphorylation and A $\beta$  nucleation.<sup>242</sup> On the other hand, there is evidence that Tau and A $\beta$  heteroligomers can lead to the formation insoluble granular aggregates,<sup>247</sup> which could be inherently toxic themselves.

To further investigate the disruption of Tau<sub>PHF43</sub> species by A $\beta$ <sub>42</sub>, we studied how A $\beta$ <sub>42</sub> interacted with fibrils of Tau<sub>PHF43</sub> (as opposed to Tau<sub>PHF43</sub> soluble oligomers). Tau<sub>PHF43</sub> was allowed to incubate with heparin, an anionic factor which is known to induce Tau fibrillation. Figure 5.8 shows the mass spectrum of Tau<sub>PHF43</sub>, which has very low abundance of Tau<sub>PHF43</sub> soluble oligomers. This indicates that Tau<sub>PHF43</sub> has formed insoluble fibrils, which fell out of solution and thus cannot be detected with IMS-MS. An additional experiment was performed where fresh A $\beta$ <sub>42</sub> was added to the Tau<sub>PHF43</sub> fibrils. It is now seen within the mass spectrum of this mixture that soluble Tau<sub>PHF43</sub> species are very prominent. Upon inspection of isotopic spacing, Tau<sub>PHF43</sub> monomers are indeed the prevalent species.

This leads us to conclude that A $\beta$ <sub>42</sub> not only disassembles Tau<sub>PHF43</sub> oligomers (specifically Tau<sub>PHF43</sub> dimers), but also Tau<sub>PHF43</sub> fibrils. Interestingly, the intensity of mass spectral peaks corresponding to A $\beta$ <sub>42</sub> species was very low. There are two plausible explanations for this. First, since heparin is known to cause amyloidogenic proteins to fibrillate, heparin also induces A $\beta$ <sub>42</sub> aggregation. This would also explain the lack of A $\beta$ <sub>42</sub>:Tau soluble heteroligomers. However, due to the short time upon adding A $\beta$ <sub>42</sub> to the Tau<sub>PHF43</sub> fibrillated sample and mass spectral analysis, this is unlikely. This leads us to speculate that instead A $\beta$ <sub>42</sub> may be binding to insoluble Tau<sub>PHF43</sub> fibrils.

The “disassembly” of Tau<sub>PHF43</sub> fibrils and disulfide-bonded dimers by A $\beta$ <sub>42</sub> monomer may have relevance for AD pathology. It has also been speculated that there may be soluble A $\beta$ <sub>42</sub> present within neurons.<sup>242</sup> Many studies showed that the amount of Tau NFTs within neurons increase substantially in the presence of A $\beta$ <sub>42</sub> *fibrils*.<sup>237-239</sup> Moreover, previous studies have established that lysine acetylation, which impairs normal Tau function, is abolished by cysteine mutations.<sup>275</sup> In addition, cysteine-deficient Tau is more prone to ubiquitination, which could enhance its proteasomal

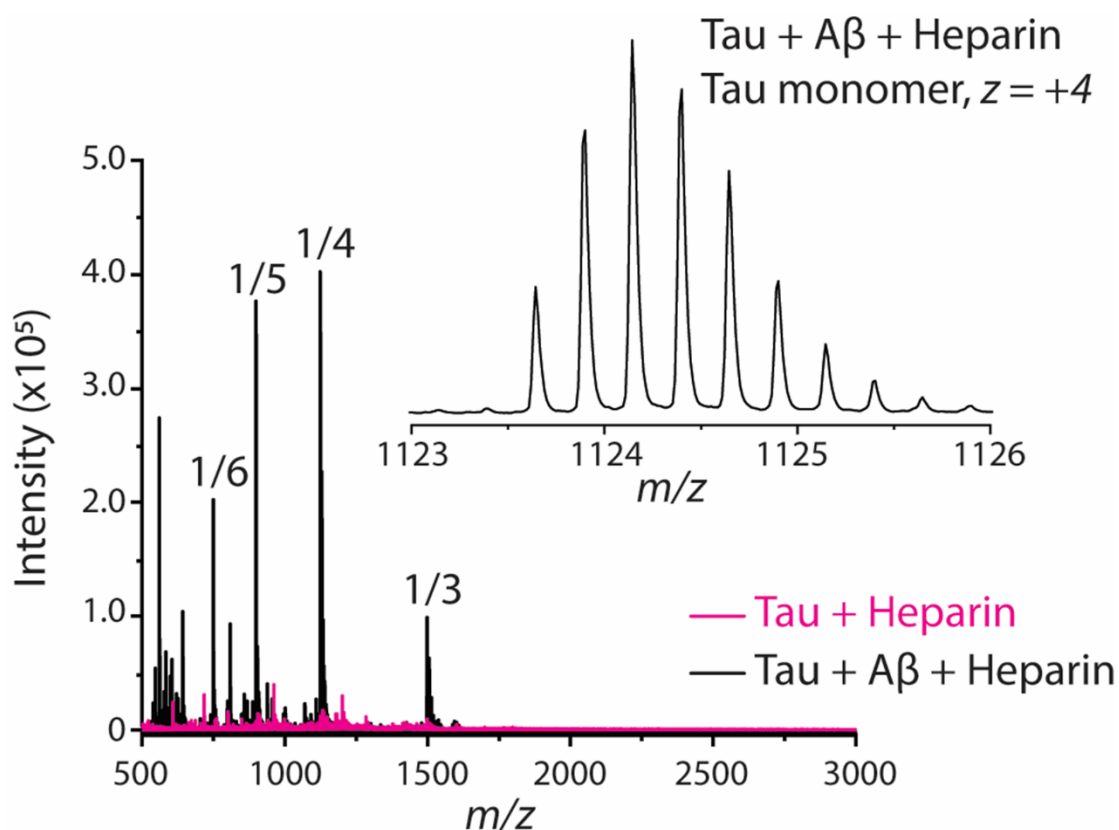


Figure 5.8. Mass spectrum of Tau<sub>PHF43</sub> + heparin that was left to incubate at room temperature over 48 hours (pink). The mass spectrum of Tau<sub>PHF43</sub> + Aβ<sub>42</sub> + heparin, incubated for 48 hours, is shown in black as comparison. Aβ<sub>42</sub> was added prior to mass spectral analysis. Note that the abundance of oligomers of Tau<sub>PHF43</sub> + heparin is much lower than that of the mixture containing the three components. This indicates that Tau<sub>PHF43</sub> has fibrillated, forming insoluble species not detected by the mass spectrometer. The soluble species of Tau are composed primarily of monomers, which are labeled (example shown by inset).

degradation<sup>253</sup> (i.e., less prone to accumulate and aggregate). Furthermore, Cys-322 (found in Tau<sub>PHF43</sub>) interacts with Cys-131 of the  $\beta$ -subunit of actin.<sup>276</sup> Thus, disulfide cross-linked tau dimers are not able to interact with actin. Finally, disulfide cross-linked Tau dimers induce tau pathology efficiently in primary neurons and also in tau transgenic mice (as shown for K18-P301L<sup>263</sup>). Collectively, our data suggest that A $\beta$ <sub>42</sub> monomeric species could exert a protective function on Tau because it prevents the formation of disulfide cross-linked dimers. However, the accumulation of A $\beta$ <sub>42</sub> leads to aggregation, and at the fibril state, its ability to disassemble the tau dimers is lost. In other words, while we arrive at a similar conclusion to previous work that A $\beta$  aggregation accelerates tau pathology, our rationale is different. A $\beta$ <sub>42</sub> itself does not promote tau aggregation, but the loss of the “protection” by A $\beta$  monomers make the Tau more vulnerable to loss of normal function and more prone to form toxic species.

#### **$\beta$ -Sheet disruptor Cyclosporine A can interact with A $\beta$ <sub>42</sub> and Tau<sub>PHF43</sub>**

*Could the protective effect of A $\beta$ <sub>42</sub> be enhanced by a  $\beta$ -sheet disruptor?* Here is our logics: if A $\beta$ <sub>42</sub> monomer is in fact Tau-protective, then its transition from intrinsically disordered to  $\beta$ -sheets should prolong its effect. Alternatively, if the disassembly of Tau was driven by the aggregation of A $\beta$ , then disrupting A $\beta$ 's  $\beta$ -sheet should have recovered the tau oligomers.

Here, we used CycA, an 11-residue macrocyclic peptide known for its immunosuppressive property, as the  $\beta$ -breaker. Recently, it has been reported that CycA can interact with amyloids to inhibit fibrillation and alleviate amyloid induced toxicity in neuroblastoma cells.<sup>277</sup> Kazmi et al. proposed that one of the potential mechanisms in which this inhibition of fibrillation may occur through binding of CycA with amyloidogenic protein monomers or oligomeric intermediates.<sup>277</sup>

We used IMS-MS to explore the two above hypotheses. Because CycA is lipophilic, we observed the complexes of CycA with both Tau and A $\beta$ . Notably, one of the more intense mass spectral peaks of this complex was observed at  $m/z = 1898$  (Figure 5.9).

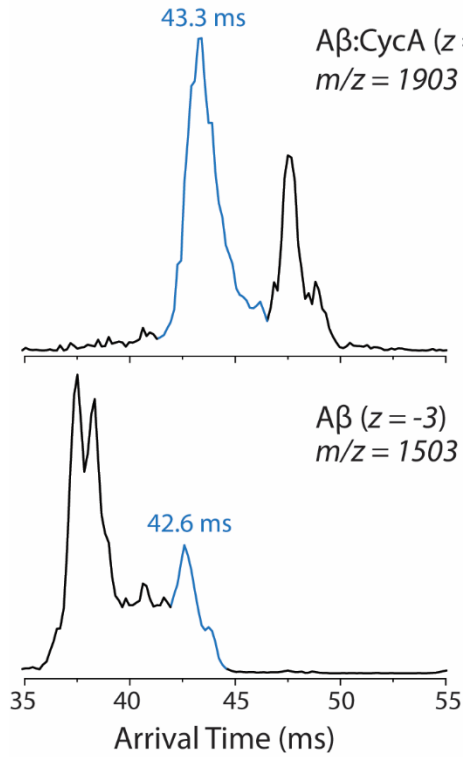
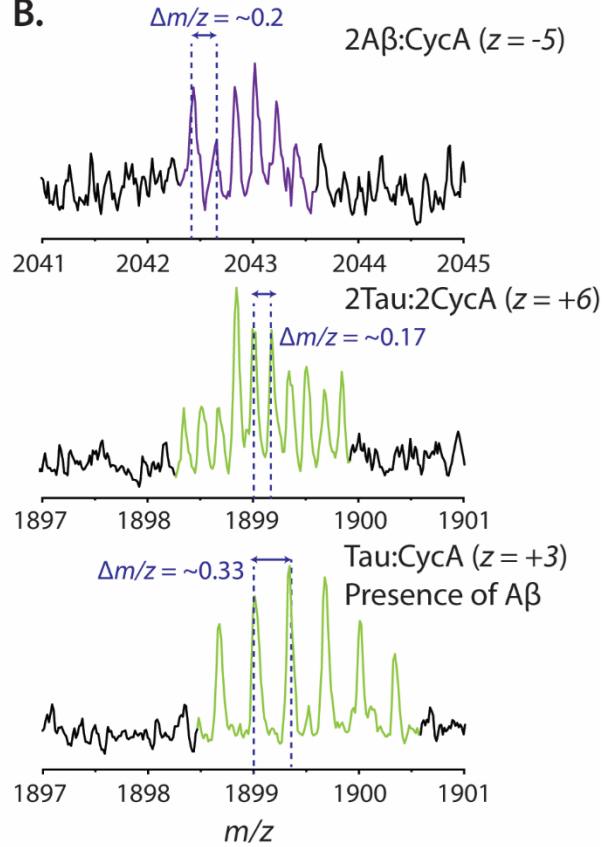
**A.****B.**

**Figure 5.9. (A) ATDs of Aβ<sub>42</sub>:CycA complex and of Aβ<sub>42</sub>. (B) Mass spectral peaks with isotopic spacing of Aβ<sub>42</sub>:CycA (2:1) complex, Tau<sub>PHF43</sub>:CycA complex (2:2), and Tau<sub>PHF43</sub>:CycA complex (1:1) in presence of Aβ<sub>42</sub>.**

Since the isotopic spacing of this species correlates to a species with  $z = +6$ , this must be of a dimeric unit of Tau<sub>PHF43</sub> and CycA (2 Tau<sub>PHF43</sub>:2 CycA). Interestingly, upon analysis of a mixture of Tau<sub>PHF43</sub>, A $\beta$ <sub>42</sub>, and CycA, the same peak at  $m/z = 1898$  now shows an isotopic spacing with  $z = +3$ , corresponding to a complex now containing only one Tau<sub>PHF43</sub> and one CycA unit. This showed that A $\beta$ <sub>42</sub> still has similar effects upon Tau<sub>PHF43</sub>, even in the presence of a  $\beta$ -breaker like CycA. Therefore, our data support that the protective effect of A $\beta$ <sub>42</sub> comes from its monomeric state (or small oligomers that are not yet large enough to nucleate aggregation).

Methodically, we also investigated the interactions between CycA and A $\beta$ <sub>42</sub>. The negative mode-ESI mass spectrum of a 1:1 mixture of A $\beta$ <sub>42</sub> and CycA showed that there is indeed binding between A $\beta$ <sub>42</sub> monomers with CycA at  $m/z = 1903$  (1 A $\beta$ <sub>42</sub> + 1 CycA with  $z = -3$ ). A $\beta$ <sub>42</sub> has a native charge state of -3, so the ATD of A $\beta$ <sub>42</sub> at  $n/z = -1/3$  can be compared to that of the A $\beta$ <sub>42</sub> + CycA complex. A $\beta$ <sub>42</sub> at  $n/z = -1/3$  has approximately two major features. When the ATDs of the A $\beta$ <sub>42</sub> monomer and the A $\beta$ <sub>42</sub> + CycA complex at  $z = -3$ , it is observed that there is overlap between one of the ATD features. Although A $\beta$  fibrils are highly ordered structures consisting of  $\beta$ -sheets, A $\beta$ <sub>42</sub> has been found to exhibit transient  $\alpha$ -helical character en route to fibrillation.<sup>278</sup> The overlap between the ATD features of the complex and monomer may support the notion that CycA stabilizes a conformer of A $\beta$ <sub>42</sub> that is not prone towards aggregation – thus, either slowing or eliminating fibrillation within this system. Collectively, the experiments with CycA supports that soluble, non-amyloidogenic A $\beta$ <sub>42</sub> is capable of disassembling disulfide cross-linked Tau dimers.

## 5.5 Conclusion

In summary, we characterized a Tau construct (Tau<sub>PHF43</sub>) previously reported by Mandelkow and coworkers<sup>248</sup> and investigated its interactions with A $\beta$ <sub>42</sub> via IMS-MS. First, we showed that Tau<sub>PHF43</sub> has a propensity to form disulfide cross-linked dimers. Previous studies have speculated the importance of these cross-linked dimers in Tau

pathology and toxicity. We were able to study the formation of higher ordered oligomers formed from Tau<sub>PHF43</sub> and examine them for the presence of disulfide bonds. The lack of disulfide bonds in these higher-ordered Tau<sub>PHF43</sub> oligomers may help explain why 3R Tau isoform (one cysteine per monomer hence no consecutive disulfide crosslinking beyond dimer) aggregate less than 4R Tau isoforms (two cysteines per monomer). Upon elucidation of interactions between Tau<sub>PHF43</sub> and A $\beta$ <sub>42</sub>, we saw that A $\beta$ <sub>42</sub> has a high tendency to bind to Tau<sub>PHF43</sub>, which has been reported previously. Surprisingly, we also found that A $\beta$ <sub>42</sub> disrupted disulfide bonds of the Tau<sub>PHF43</sub> dimers. Our data suggest that this may indicate a protective function of non-amyloidogenic A $\beta$  species (monomer or small oligomers) on maintaining normal Tau function and making Cys-322 available for actin binding. Experiments with CycA, a lipophilic macrocycle known to disrupt  $\beta$ -sheet structures, support our hypothesis on the overlooked protective function of A $\beta$ <sub>42</sub> monomer.

## **6 Atomic view of an amyloid dodecamer exhibiting selective cellular toxic vulnerability in acute brain slices**

### **6.1 Disclaimer**

The following chapter has been adapted, with permission, from a journal article published in *Protein Science*.<sup>279</sup> Only minor modifications were made, including the omission of supporting details regarding the X-ray crystal structure and tissue viability assay, and the addition of the detailed discussion regarding the G6W-DDM complexes and justification of the assignment of the features within the G6W  $n/z = 1/1$  in the presence of DDM (the original journal article has this information within the supporting document). This work was performed in conjunction with Rebecca Prosser and her lab group, Michael Sawaya, and the lab group of Michael Best. Rebecca Prosser performed the tissue toxicity assay, Michael Sawaya elucidated the atomic structure of G6W via X-ray crystallography, Thanh Do performed theoretical CCS calculations, and Michael Best performed the dye-release assays. All data pertaining to the IMS-MS analysis of G6W and KV11 were obtained and analyzed by myself under the guidance of Thanh Do.

### **6.2 Abstract**

Atomic structures of amyloid oligomers that capture the neurodegenerative disease pathology are essential to understand disease-state causes and finding cures. Here we investigate the G6W mutation of the cytotoxic, hexameric amyloid model KV11. The mutation results into an asymmetric dodecamer composed of a pair of 30° twisted antiparallel  $\beta$ -sheets. The complete break between adjacent  $\beta$ -strands is unprecedented among amyloid fibril crystal structures and supports that our structure is an oligomer. The poor shape complementarity between mated sheets reveals an interior channel for

binding lipids, suggesting that the toxicity may be due to a perturbation of lipid transport rather than a direct disruption of membrane integrity. Viability assays on mouse suprachiasmatic nucleus, anterior hypothalamus, and cerebral cortex demonstrated selective regional vulnerability consistent with Alzheimer's disease. Neuropeptides released from the brain slices may provide clues to how G6W initiates cellular injury.

### 6.3 Introduction

Transient oligomeric species formed during protein misfolding and aggregation have been implicated as the toxic agent for many neurodegenerative diseases, including Alzheimer's disease (AD),<sup>280, 281</sup> Huntington's disease,<sup>282</sup> Parkinson's disease,<sup>283</sup> and amyotrophic lateral sclerosis (ALS).<sup>284</sup> Conformational heterogeneity and co-aggregation of several proteins, such as the plaques of amyloid- $\beta$  (A $\beta$ ) and tangles of hyperphosphorylated tau observed in AD,<sup>285</sup> contribute to the complex etiology of the diseases. Further compounding the complexity of these diseases is the aspect of selective cellular/regional vulnerability, where deposits of protein aggregates are more prominent in specific regions of the brain and cause more damage to certain types of cells but not others.<sup>286</sup> For example, slow-firing motor neurons are more resilient than their fast-fatigable counterparts in ALS pathology.<sup>287-290</sup> Thus, cellular heterogeneity, even among similar cell types,<sup>288</sup> must also be taken into account when investigating such diseases. Understanding how certain cell types in some brain regions are able to resist amyloid toxicity is essential to developing novel therapeutic strategies.

Efforts to replicate how toxic oligomers initiate and propagate disease conditions have enlisted the use of cellular mimetics<sup>291, 292</sup> or relied on the outcomes of (colorimetric) cell viability assays.<sup>284, 293, 294</sup> However, the latter is often hypothesis-centric and performed on commercial cell lines, which do not account for cellular heterogeneity and inter-cellular communication. Acute tissue slices provide an alternative to commercial cell lines where different brain regions containing multiple

cell types can be evaluated while maintaining in vivo tissue structure. Identifying cell types that are resilient to amyloid toxicity may elucidate endogenous molecules that provide protection against the disease toxins. Moreover, to clarify how small but toxic amyloid oligomers affect brain tissues, amyloid models with known structures should be investigated. As we will demonstrate in this article, an attempt to disrupt one type of toxic oligomer may lead to the formation of different, larger, more structurally complex, and toxic oligomers. The multifaceted research path outlined here has not been fully exploited in AD research.

In this study, we started with the 11-residue segment of the  $\alpha$ B-crystallin protein (KVKVLGDVIEV), referred to here as KV11, previously shown by the Eisenberg group to form cylindrical hexamers (termed cylindrins).<sup>293</sup> Several 11-residue peptides from A $\beta$ <sup>290</sup> and TAR DNA-binding protein of 43 kDa (TDP-43) were proposed to form similar structural motifs.<sup>295, 296</sup> It was hypothesized that depending on the length and amino acid composition of the peptide, either cylindrin or a  $\beta$ -barrel could be formed. For example, Do et al. showed by IMS-MS and molecular dynamics simulation that A $\beta$  (25–35) forms an octameric  $\beta$ -barrel.<sup>295</sup>

Because the glycine in the sixth position is essential to the structural integrity of toxic amyloid-like cylindrin assemblies,<sup>284, 293</sup> it is logical to predict that a mutation of glycine to a bulky residue, such as tryptophan, would prevent cylindrin formation and alleviate cytotoxicity. Here, we will show that the opposite is true; as one type of toxic amyloid structures is disrupted, another type that is more toxic can be introduced. Experiments to identify and separate transient oligomers of KV11 and KV11<sup>G6W</sup> (hereafter referred to as G6W) reveal different species formed from these two distinct peptide species. Ion mobility spectrometry–mass spectrometry (IMS–MS) has been used previously for similar tasks.<sup>284, 295</sup> In IMS-MS, isobaric oligomers with the same *mass-to-charge* ratio (*m/z*) can be separated by measuring their drift times. We complement the IMS-MS work with X-ray crystallography to unambiguously determine some of the structure states of G6W. To gain insight into how these structures influence biochemistry, we

investigated how KV11 and G6W interact with healthy brain tissue. Specifically, we explored how G6W and KV11 differentially affect cells within distinct brain regions using viability assays with propidium iodide (PI). Using liquid chromatography (LC)-IMS-MS, we interrogated subtle changes in the peptides released from mouse brain slices incubated with G6W. Finally, ligand binding pockets discovered in our structure of G6W led us to investigate the effect of membrane mimetics on oligomerization and stability of G6W oligomers.

## **6.4 Materials and methods**

### **Ion mobility spectrometry – mass spectrometry**

KV11 and G6W peptides were acquired and pre-treated with hexafluoro-2-propanol (HFIP). After allowing HFIP to evaporate, the peptides were reconstituted in HPLC water and diluted to the desired concentration in 20 mM ammonium acetate. Data of KV11 and G6W oligomers with and without n-Dodecyl- $\beta$ -D-maltoside (DDM; Sigma) were directly injected via syringe into an Agilent 6560 IM-QTOF mass spectrometer (Santa Clara, CA) in a multi-field experiment. Helium or nitrogen was used as the drift gas and the pressure within the drift cell was approximately 3.94 torr. Positive-mode data were collected over approximately 5 minutes, with  $\Delta V = 890, 790, 690, 590,$  and  $490$  V (helium) or  $\Delta V = 1490, 1390, 1290, 1190,$  and  $1090$  V (nitrogen) within the drift cell. Mass spectra and ion mobility data were extracted with Agilent IMS-MS Browser and graphed with Origin Pro. Experimental CCS of pure peptides (and species complexed with DDM) were obtained in the same manner discussed in Section 2.4.

### **Liquid chromatography IMS-MS**

The media surrounding the SCN+AH brain slices (i.e., releasates) was concentrated via a speed vacuum and placed into a  $-20^{\circ}\text{C}$  freezer until LC-IMS-MS analysis. Prior to analysis, the releasate was reconstituted with approximately  $150\ \mu\text{L}$  of LC-MS water. For LC-IMS-MS analysis, an Agilent 1260 Infinity was coupled to the Agilent 6560 IMS-MS with

nitrogen as the drift gas. For each run, 1.50  $\mu$ L of the sample was injected into a 2.1x50 mm Zorbax Extended C18 reverse phase column (Agilent, Santa Clara). The mobile phases were composed of an aqueous solution of 0.1% formic acid in HPLC-grade water (solvent A) and acetonitrile (solvent B). The mobile phase gradient is as follows: 98% solvent A for 3 minutes, 98% solvent A (3 – 5 min), 80% solvent A (5 – 18 minutes), 65% solvent A (18 – 23 min), 50% solvent A (23 – 26 min), 5% solvent A (26 – 27 min), 5% solvent A (27 – 30 min), 98% solvent A (30 – 33 min). At the end of the run, 98% solvent A was kept for 5 minutes to allow for column equilibration. The temperature within the column compartment was set to 40°C. The multiplexed IMS-MS data was demultiplexed and multidimensionally smoothed with the Agilent de-multiplexing tool and Pacific Northwest National Laboratory Preprocessor tool (<https://omics.pnl.gov/software/pnnl-preprocessor>), respectively. At least six technical replicates were collected from each sample. A total of four biological replicates were used in this experiment.

To identify neuropeptides, a database containing neuropeptides were compiled in Agilent MassHunter Personal Compound Database and Library Manager Software. The neuropeptides were obtained from previous literature which identified neuropeptides expressed from rat SCN, however we used the mouse peptide sequence to compile the database. Agilent MassHunter Mass Profiler Software was then used to filter out low-intensity and noisy features within the LC-IMS-MS data. Data were then imported into Agilent MassHunter Mass Profiler Professional Software and features were annotated based on their mass through comparing LC-IMS-MS data to the neuropeptide database. Filtering was performed on frequencies to remove “one-hit wonders”. Statistical tests (Moderated T-test) and false discovery rate corrections (Benjamini & Hochberg’s) were performed to refine the list of potential markers. The presence of these features within each biological replicate were confirmed by manual inspection of the data and cross-compared with the control SCN/AH releasate as well as a G6W standard.

To calculate the experimental CCS of species detected via LC-IMS-MS, data were collected over a single voltage within the drift cell and CCS was calculated indirectly. To

do this, tune ions from Agilent Calibrant Tune Mix (Agilent, Santa Clara) were used as a reference for CCS. Agilent MassHunter Mass Profiler Professional Software was then used to determine the CCS of features within LC-IMS-MS samples.

### **Tissue toxicity assay**

Adult male C57BL/6 mice (Envigo; Indianapolis, IN) were housed under 12:12 light:dark cycles with food and water provided ad libitum. Brain slices containing either the SCN and the surrounding AH (two regions within the same brain slices) or the cortex were incubated in an interface brain slice chamber under static conditions either with Earl's Balanced Salt Solution (EBSS) or with EBSS containing approximately 300  $\mu$ M KV11 or 400 $\mu$ M G6W peptides for 4 hours. The media surrounding the slices was collected and stored for later analysis. The brain slices were maintained further for perfusion with PI, then fixed and sectioned for subsequent histological analysis. Fluorescent images of sectioned tissue were obtained with a Leica DM6000B microscope and analyzed with ImageJ software.

### **X-ray crystallography**

G6W and G6W-Br5 for X-ray crystallography trials were purchased from LifeTein (Somerset, NJ). Diffraction data of G6W and G6W-Br5 derivatives were obtained at beamline 24-ID-C of the Advanced Photon Source located at Argonne National Laboratory.

### **Liposome dye-release assay**

Liposomes consisting of 100% l- $\alpha$ -phosphatidylcholine (PC) and 90% PC and 10% of either l- $\alpha$ -phosphatidic acid sodium salt (PA) or 1,2-dioleoyl-3-trimethylammonium-propane, chloride salt was prepared, and fluorescence of liposome dye release was measured with a Cary Eclipse Fluorescence Spectrophotometer. A Malvern Zetasizer Nano ZS instrument was used for dynamic light scattering experiments.

## 6.5 Results and discussion

### Mouse brain slices exhibit selective regional vulnerability to KV11 and G6W

We investigated the relative toxicity of KV11 and G6W across three brain regions. KV11 has previously been shown to exhibit cytotoxicity using a wide range of commercial cell lines including HEK293, HeLa, PC12, and SH-SY5Y.<sup>293</sup> Here, using acute mouse brain slices, we focused on the cerebral cortex, the suprachiasmatic nucleus (SCN), and anterior hypothalamus (AH). The cerebral cortex is a brain region that exhibits a large degree of neurodegeneration in AD.<sup>297-300</sup> Conversely, the SCN, the locus of primary circadian clock in mammals,<sup>301-303</sup> is thought to be relatively resilient to AD, although disruptions in sleep and other circadian rhythms are increasingly acknowledged as early symptoms of AD.<sup>304-306</sup> The SCN is surrounded by the AH, of which little is known regarding its neurodegenerative vulnerability. Mouse brain slices containing cortex and SCN/AH region were isolated from healthy mice and incubated with either KV11 or G6W peptides (~300 or 400  $\mu$ M, respectively). After a 4-hour incubation, the media surrounding the tissue slices was removed and stored, and the slices were perfused with PI. Slices were then fixed, sectioned, and coverslipped with mounting medium containing DAPI. Cell viability within each brain slice (i.e., tissue viability) was assessed by quantifying DAPI and PI staining using fluorescence imaging.

Overall, as seen in Figure 6.1, it is clear that KV11 and G6W affected cell viability in all three brain regions. Focusing within the individual brain regions, 2-way ANOVA determined that G6W significantly decreased the percentage of healthy cells in all three brain regions compared to controls ( $p < .05$ ). G6W also significantly decreased the number of healthy cells in the SCN and AH compared to KV11 exposed tissue ( $p < .05$ ). KV11 decreased the percentage of healthy cells in the cortex compared to control conditions, but this did not reach statistical significance ( $p = .093$ ). G6W increased the number of injured cells in the SCN compared to control conditions and compared to KV11 ( $p < .05$ ). Finally, G6W significantly

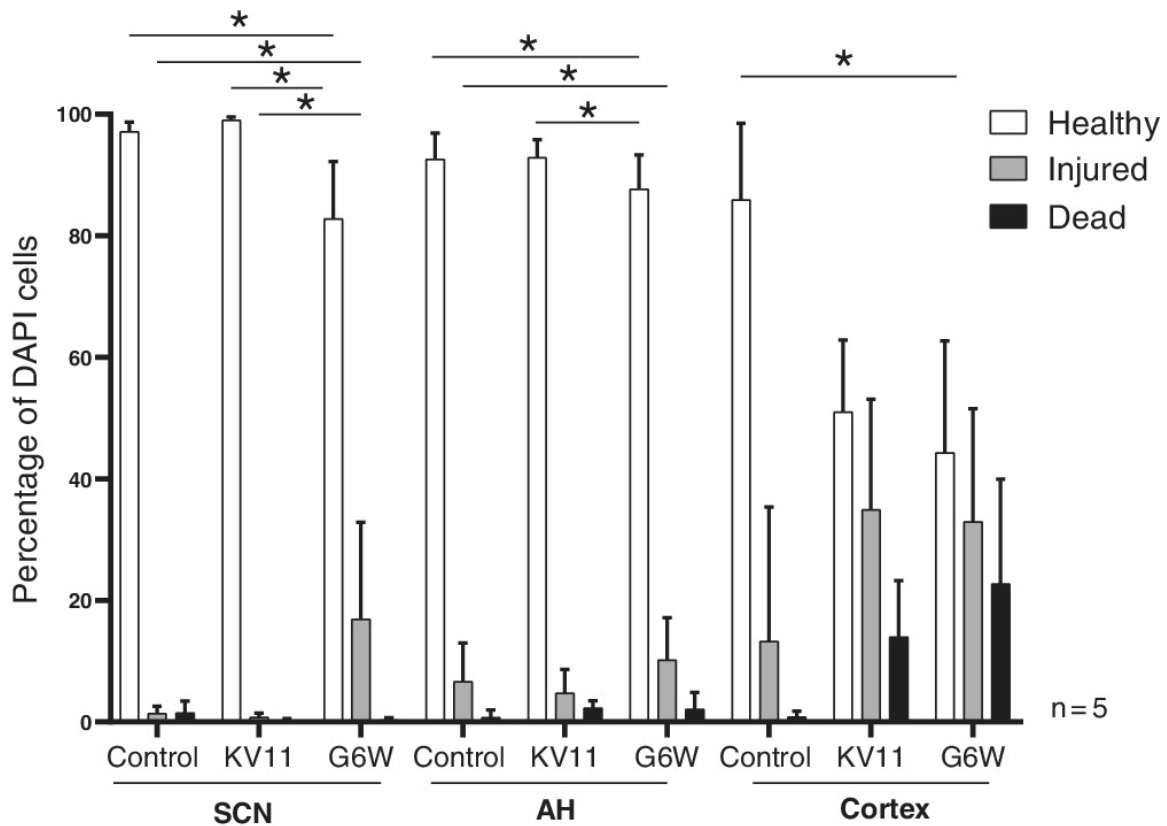


**Figure 6.1. Tissue viability assessment under different experimental conditions.** Histogram bars indicate the percentage of DAPI-stained cells with different amounts of PI staining. High-intensity PI staining indicates dead cells, moderate PI staining indicates injured cells, and low to no PI staining indicates healthy cells. (\* $p < .05$ ) The error bars represent the standard error mean.

increased the number of injured cells in the AH compared to the controls ( $p < .05$ ), while the increase in injured cells in the AH compared to KV11 did not quite reach statistical significance ( $p = .06$ ). Taken together, the results indicate that G6W is more toxic than KV11 at least in the SCN and AH, while the toxic effects of KV11 and G6W do not differ significantly between each other in the cortex. The results showing that the SCN, which is considered to be a resilient and robust structure within the brain,<sup>307</sup> is affected by amyloids in a manner similar to the cortex are somewhat surprising. Although there was no significant increase in SCN cell death at this time point, it is possible, for example, that toxic peptides are disrupting neuropeptide signaling pathways responsible for SCN neuronal health and synchronization.<sup>308-310</sup> The data also show that the AH exhibits similar vulnerability to amyloid peptides as the SCN. Few comparisons have been made between these adjacent areas, but cell culture studies have shown that SCN-derived cells exhibit greater excitotoxic resiliency compared to hypothalamus-derived cells.<sup>311</sup> These distinct results point to the importance of using intact tissues rather than cell cultures when investigating toxic susceptibility.

An important outcome of these experiments is their demonstration of cellular vulnerability to peptide toxicity during early stages of exposure (i.e., the first 4 hours) in all three brain regions. As discussed below, the mechanisms of toxicity are not known. Implicated proteinaceous agents of neurodegenerative disease, such as A $\beta$  and neurofibrillary tangles of paired tau filaments, are known to accumulate within and around cortex neurons.<sup>312</sup> Cortical neurons may exhibit properties that makes them more vulnerable to degradation and amyloid deposition. For example, differences in firing properties of different neuronal populations within layer II of the entorhinal cortex may contribute to susceptibility to neurodegeneration.<sup>313</sup> In future studies, it will be important to extend the duration of amyloid peptide exposure to determine if regional differences in toxic susceptibility become more apparent.

### IMS-MS reveals that G6W can form large, structurally complex oligomers

The representative mass spectra of KV11 and G6W under the same experimental conditions (20  $\mu$ M in 20 mM ammonium acetate pH 7) are shown in Figure 6.2A,B. Both peptides populate high-order oligomers annotated by nominal  $n/z$  where  $n$  is oligomer size and  $z$  is charge. Both mass spectra show intense peaks at  $n/z = 3/2$  of which IMS-MS data revealed the presence of trimers ( $z = +2$ ) and hexamers ( $z = +4$ ). The oligomers are unambiguously assigned based on the detected isotopic spacings of  $\Delta m/z = 0.5$  and 0.25, respectively (Figure 6.3).

Although at first glance, there are slight differences between the mass spectra of G6W and KV11 (e.g., the decapentamer species formed by G6W is not observed in KV11), the major difference between the two peptides can be observed in the arrival time distribution (ATD) of the nominal  $n/z = 1/1$  species. The 2D plot ( $m/z$  vs. drift time) of G6W (Figure 6.2C) shows a feature with an arrival time of  $\sim 20.30$  ms (red arrow), which has a distinct pattern of isotopic spacing not seen within the KV11 2D plot (Figure 6.2D). Within this spacing, one of the patterns indicates that there was a trimer with  $z = +3$ . The other isotopic spacing pattern suggests that there may also be a dimeric species with  $z = +2$  within this feature. However, if this feature were a dimer, then it would have a collisional cross-section (CCS,  $\sigma$ ) in helium of  $402 \text{ \AA}^2$ , which is inconsistent with the  $\sigma_{\text{exp,He}}$  values of other known dimeric species (Table 6.1). Post-ion-mobility dissociation, a phenomenon we have described in recent work,<sup>34, 84</sup> in which weakly-bound oligomeric species traverse the drift-cell successfully but dissociate before reaching the mass analyzer, may explain the ambiguity we are seeing within the isotopic spacing of this feature.

Additionally, there is a major feature within G6W at 16 ms (blue arrow) that is unique to this peptide. Due to the ambiguity present within the 2D plot, we could not rely solely on isotopic spacing to determine the oligomer size of this feature and therefore needed to incorporate additional data. As G6W oligomerizes to form higher-ordered oligomers, the spacing between the features decreases. For example, between  $n/z = 1/1$  and  $2/2$ ,

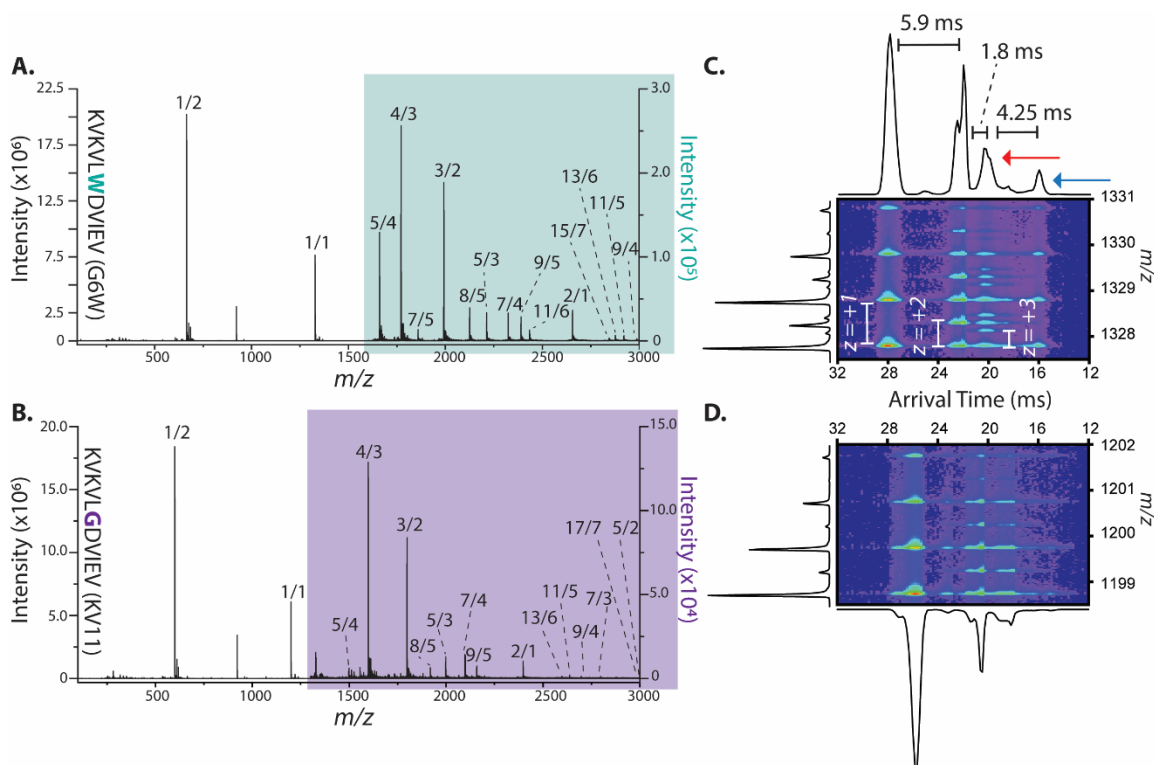


**Figure 6.2. Representative mass spectra of (A) G6W and (B) KV11.** Each mass spectral peak is labeled by its nominal oligomer-to-charge ratio ( $n/z$ ). Note that the intensity in the colored regions of each mass spectra is adjusted to reveal less abundant oligomeric species. The 2D plots for  $n/z = 1/1$  are shown for both (C) G6W and (D) KV11, where the isotopic spacings can clearly be seen for the features within the ATDs. For G6W, notably, species mentioned in the text are marked with arrows. The data were collected using helium buffer gas.

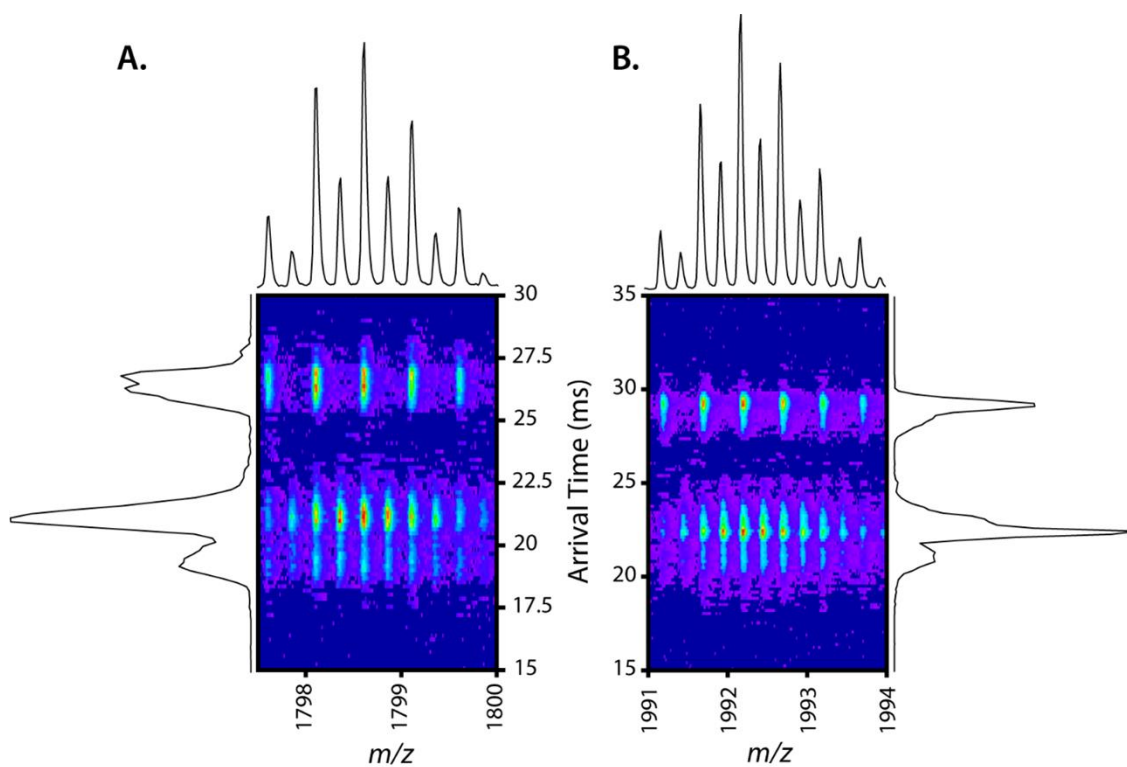


Figure 6.3. The 2D plots of (A) KV11 and (B) G6W are shown at  $n/z = 3/2$ . The ATDs show two major features. The isotopic spacing for the faster feature clearly shows that  $Z = +4$ , which means that the oligomer represented by this feature must be a hexamer (i.e.  $n/z = 6/4 = 3/2$ ).

**Table 6.1. Experimental collisional cross section values for G6W oligomers.**

| Oligomer Size ( <i>n</i> ) | <i>m/z</i> | Charge State ( <i>Z</i> ) | Experimental CCS <sub>He</sub> | Experimental CCS <sub>N<sub>2</sub></sub> |
|----------------------------|------------|---------------------------|--------------------------------|-------------------------------------------|
| 1                          | 1327.8     | 1                         | 278.2                          | 360.2                                     |
|                            | 664.4      | 2                         | 285.2                          | 368.2                                     |
| 2                          | 2655.1     | 1                         | 416.7                          | 518.3                                     |
|                            | 1327.8     | 2                         | 438.7                          | 543.4                                     |
| 3                          | 1991.2     | 2                         | 582.8                          | 707.8                                     |
|                            | 1327.8     | 3                         | 602.8                          | 730.6                                     |
| 4                          | 2655.1     | 2                         | 673.9                          | 811.8                                     |
|                            | 1770.1     | 3                         | 681.3                          | 820.2                                     |
|                            | 1327.8     | 4                         | 776.4                          | 928.7*                                    |
| 5                          | 2212.3     | 3                         | 780.5                          | 933.4                                     |
|                            | 1659.5     | 4                         | 814.8                          | 972.6                                     |
| 6                          | 2655.1     | 3                         | 883.3                          | 1050.7                                    |
|                            | 1991.2     | 4                         | 894.6                          | 1063.6                                    |
| 7                          | 2323.1     | 4                         | 1006.4                         | 1191.2                                    |
|                            | 1858.7     | 5                         | 1058.9                         | 1251.1                                    |
| 8                          | 2655.1     | 4                         | 1094.9                         | 1292.2                                    |
|                            | 2124.1     | 5                         | 1140.4                         | 1344.1                                    |
| 9                          | 2987.1     | 4                         | 1191.6                         | 1402.6                                    |
|                            | 2389.6     | 5                         | 1228                           | 1444.1                                    |
| 10                         | 2655.1     | 5                         | 1300.3                         | 1526.6                                    |
|                            | 2212.3     | 6                         | 1373.9                         | 1610.6                                    |
| 11                         | 2920.6     | 5                         | 1398.9                         | 1639.1                                    |
|                            | 2433.8     | 6                         | 1458.6                         | 1707.3                                    |
| 12                         | 1327.8     | 12                        | 1860.5                         | 2427.8                                    |
| 13                         | 2876.4     | 6                         | 1599.5                         | 1868.1                                    |
| 15                         | 2845       | 7                         | 1825.4                         | 2125.9                                    |

\*Data collected in nitrogen in presence of DDM.

there is a difference of 5.9 ms, while from 2/2 to 3/3 there is a difference of 1.8 ms (using helium as the buffer gas). The drift time difference between 3/3 and the shortest arrival time feature at 16 ms is 4.25 ms. As the oligomers grow in size, the contribution of a monomeric unit to the total oligomer cross-section gets smaller.<sup>295, 314, 315</sup> This led us to conclude that the species with the shortest arrival time feature must not be a tetramer. By comparing the  $\sigma_{\text{exp,He}}$  values of potential species that the shortest arrival time feature can represent (Table 6.2), the oligomer must be a very large oligomer ( $n > 6$ ). We will further discuss the assignment of this oligomer in the next sections in light of the X-ray crystallography and IMS-MS data on interactions of G6W with DDM detergent. Taking all these points into consideration, our IMS-MS data indicated that G6W oligomers are structurally more complex than KV11. G6W can populate two types of oligomers: those that have similar oligomer sizes (not necessarily the same structures) to KV11 (as shown in the analysis of  $n/z = 3/2$  species; Figure 6.3) and those that are distinct from KV11 (as seen in the ATDs of  $n/z = 1/1$ ; Figure 6.2).

### **G6W assembles into a dodecameric asymmetrical autonomous unit**

We used X-ray crystallography to investigate the relationship between the toxicity and the structure of G6W oligomers. Single-crystal diffraction data of both G6W and G6W-Br5 where the fifth residue (Leu) was replaced by l-(2-bromoallyl)-glycine were collected. We note that G6W-Br5 displayed similar oligomeric structures to G6W, as shown by IMS-MS data in Figure 6.4. The presence of heavy atoms (Br) allows a more accurate structural refinement.

The crystal structure of G6W reveals a dodecameric assembly distinct from amyloid fibrils. It consists of a pair of  $\beta$ -sheets twisted around a common axis, creating a double helix with approximate  $6_422$  symmetry (Figure 6.5A-C). The core of the oligomer prominently features 12 tryptophan side chains, which stack and stabilize the dodecamer through hydrophobic contacts with valine side chains of neighboring  $\beta$ -strands (Figure 6.5A,C,D).

**Table 6.2. Comparison of CCS of shortest arrival time feature for G6W  $n/z = 1/1$ . We cannot rely solely on isotopic spacing of the shortest arrival time feature in the 2D plot of G6W  $n/z = 1/1$  due to ambiguity present. However, by calculating the experiment CCS values for various species, we can determine that the shortest arrival time feature must be an oligomer with  $n = 6$  or greater. Oligomers smaller than  $n = 6$  will have CCS that are smaller than oligomers of the same  $n$  with a smaller  $Z$ , which is not possible because CCS increases as  $Z$  increases due to coulombic repulsions.**

| Oligomer size ( $n$ ) | Charge State ( $Z$ ) | Experimental CCS <sub>He</sub> |
|-----------------------|----------------------|--------------------------------|
| 4                     | 4                    | 620.3                          |
| 5                     | 5                    | 775.4                          |
| 6                     | 6                    | 930.4                          |

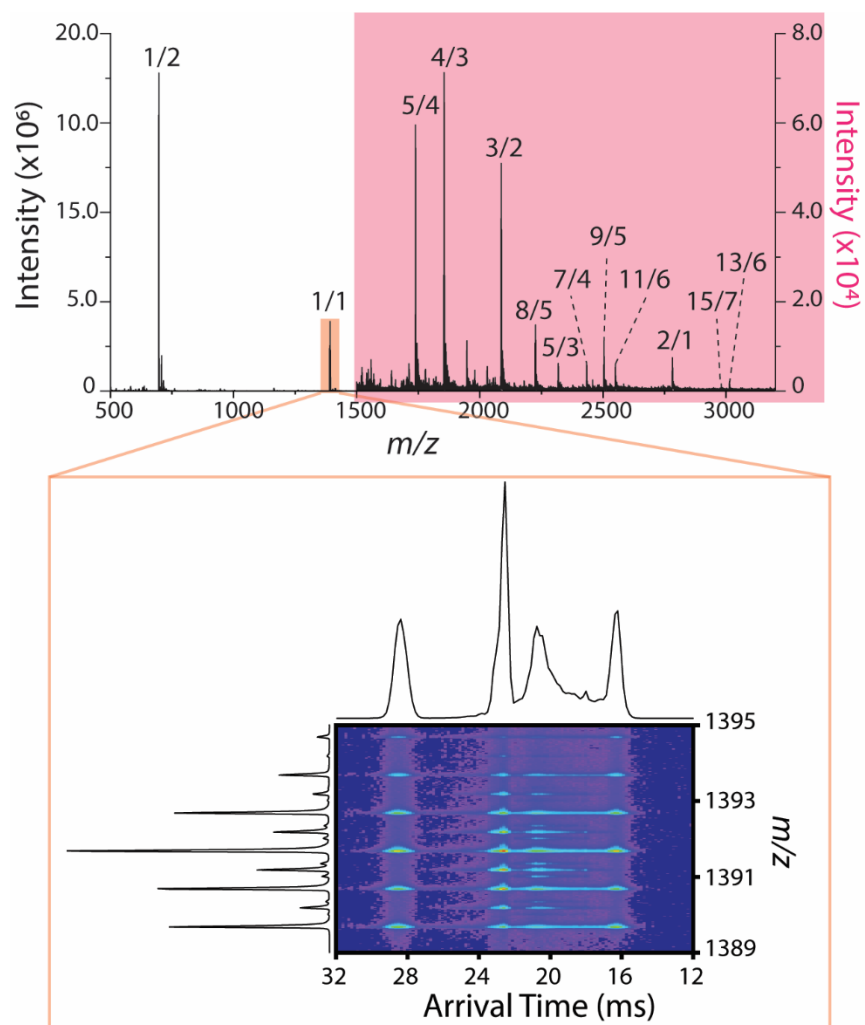


Figure 6.4. The mass spectrum and 2D plot of G6W-Br5 at  $n/z = 1/1$ . The 2D plot of G6W-Br5 showed that similar oligomers are still observed within this plot as the native G6W molecule.

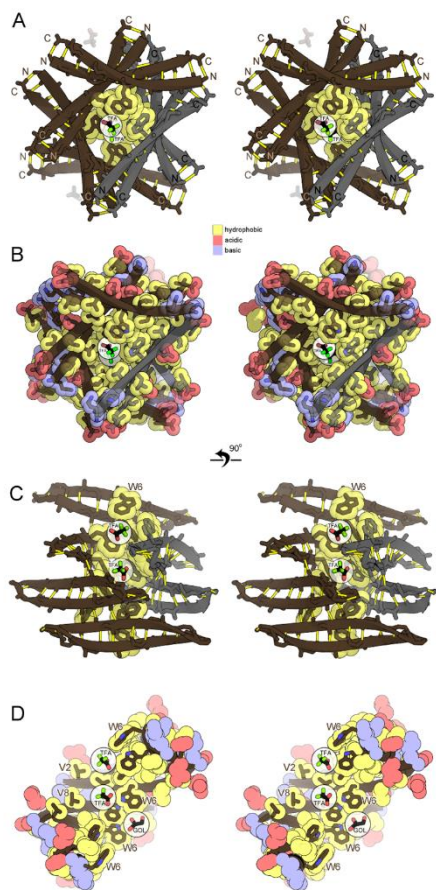


**Figure 6.5.** The G6W dodecamer crystal structure viewed (A, B) parallel and (C, D) perpendicular to the pseudo  $6_4$  symmetry axis. Two highly twisted, antiparallel sheets (brown and gray) meet side-by-side at an interface composed of hydrophobic side chains including Trp6 residues, highlighted by van der Waals spheres in panels A and C. Backbone hydrogen bonds are highlighted by yellow lines. Shape complementary between sheets is poor, leaving voids filled by trifluoroacetic acid (TFA) and glycerol (GOL) highlighted by white circles. (B) A space-filling model reveals a charged outside surface (pink and blue spheres) and (D) a cutaway view reveals a hydrophobic core (yellow spheres) with pockets for TFA and GOL binding.

The G6W structure exhibits four important distinctions from amyloid fibrils. The  $\beta$ -sheets are composed of antiparallel, rather than parallel strands as seen in most amyloid fibril structures. A survey of amyloid fibril literature reveals antiparallel sheet assemblies in only 2 out of ~80 fibril structures determined.<sup>316</sup> Moreover, the twist of the sheet is much greater in G6W than in amyloid fibrils. On average,  $\beta$ -sheets in G6W twist 30° per  $\beta$ -strand (Figure 6.5A). In contrast, most amyloid fibrils twist <2° per strand. Then, the  $\beta$ -sheets in G6W crystals are discontinuous; a complete break in hydrogen bonding occurs between  $\beta$ -strands every 50.6 Å along with the sheet (12 strands), separating adjacent oligomers (Figure 6.6). Such a complete break between adjacent  $\beta$ -strands is unprecedented among amyloid fibril crystal structures and supports the hypothesis that our structure is representative of an oligomer rather than a fibril. Lastly, and most significantly, the fit between mated sheets has poor shape complementarity, creating pockets for ligand binding. The  $\beta$ -sheets in G6W are mated side-by-side through contacts between side chains as is a characteristic of amyloid fibrils; however, the packing is not as tight-fitting as most amyloid steric zippers.<sup>317</sup> Indeed, gaps between sheets are filled with trifluoroacetate (TFA) molecules and glycerol (Figure 6.5).

Such voids in an otherwise stable oligomeric assembly create the potential for biological activity. If the TFA molecules are removed from the structure, a deep and narrow channel becomes evident. It is roughly the size and shape to bind lipids such as palmitic acid (Figure 6.7), or smaller metabolites. The G6W dodecamer shows a remote structural similarity to lipid transfer proteins (e.g., phosphatidylinositol transfer protein; PITP) and fatty acid binding proteins (FABP). The G6W is made of six pairs of anti-parallel  $\beta$ -strand dimers. PITP binding site (PDB ID 1T27)<sup>318</sup> is formed by the concave surface of an eight-stranded antiparallel  $\beta$ -sheet and two helices. The FABP (PDB ID 1HMR)<sup>319</sup> consists of 10 antiparallel  $\beta$ -strands and two short helices. The G6W dodecamer has a single nonpolar channel for binding of one acyl tail while PITP and FABP have a lipid headgroup binding pocket and hydrophobic channel(s) for fatty acid

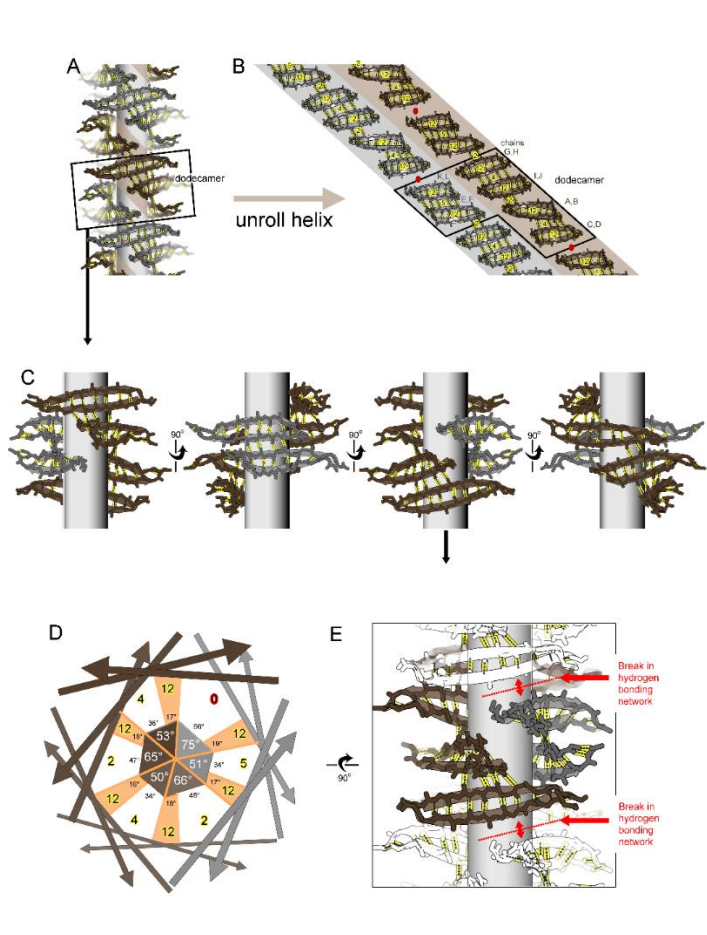


**Figure 6.6. (A)** A column of G6W dodecamers in the crystal. The box indicates the boundary of one dodecamer. Backbone hydrogen bonds are highlighted in yellow. Brown and gray colors distinguish the two  $\beta$ -sheets in the dodecamer. The cylinder represents the pseudo  $6_4$  symmetry axis. **(B)** We unrolled the  $\beta$ -sheets into a flat plane in order to highlight the location of the breaks in hydrogen bonding (red zeros). The count of hydrogen bonds between each pair of strands is indicated. **(C)** Orthogonal views of the dodecamer. **(D)** A projection down the symmetry axis highlights the deviation from ideal  $6_4$  symmetry in which the angles between antiparallel pairs of strands would be exactly  $60^\circ$ , and hydrogen bond counts (yellow numbers) would be evenly distributed around the axis. **(E)** Close-up view of the absence of hydrogen bonds in the break between dodecamers.

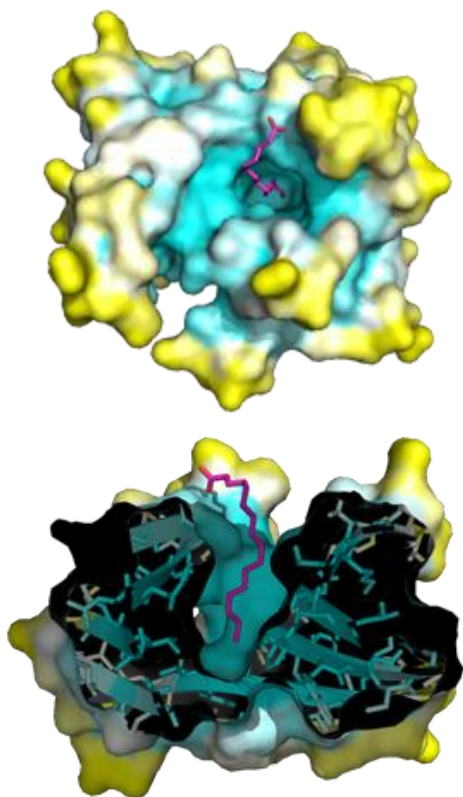


**Figure 6.7. Molecular modeling shows palmitic acid binding to hydrophobic cavity of G6W.**

tail(s). The toxic mechanism may allude to the role of Apolipoprotein E (APOE) in fat metabolism. Notably, the APOE gene variants are the most commonly associated with late-onset AD.<sup>320</sup>

Conformational flexibility appears to play a role in creating these binding pockets. Notably, symmetry is lacking between TFA binding sites (Figure 6.5D), suggesting asymmetry is important to form the pockets. Indeed, a perfectly symmetric conformation observed in the derivative G6W-Br5 (Figure 6.8) reveals no ligand binding. Conversion between symmetric and asymmetric dodecamers seems plausible given the modular construction of the dodecamer. G6W molecules are bonded tightly into antiparallel  $\beta$ -sheets by 12 backbone hydrogen bonds (Figure 6.9A). These dimer modules join together at different angles through different numbers of hydrogen bonds (Figure 6.9CD). We imagine that small adjustments to these joints, like those demonstrated here between G6W and G6W-Br5, might aid in binding a ligand through an induced-fit mechanism. Imperfect or mismatched symmetry has been noted as a characteristic of dynamic assemblies such as  $F_1F_0$  ATP synthase.<sup>321</sup>

The presence of an interior channel and flexibility in the dodecameric assembly of G6W molecules distinguish it from the structures of other amyloid oligomers. For example, multiple structures of the parent KV11<sup>293</sup> sequence have been determined, but no variation in the cylindrin architecture was observed and the interior was inaccessible to solvent. Structures of a more distant relative, a segment of superoxide dismutase segment, called “corkscrew”<sup>284, 322</sup> show variation in twist, but exhibited surface grooves rather than an interior channel. All these amyloid oligomers share a robust assembly of antiparallel pairs of  $\beta$ -strands (Figure 6.10). The dimer modules are held together by full backbone hydrogen bonding and hydrophobic forces between side chains. However, the fit between dimer units varies. The fit between dimers is tight in cylindrin, but poor in G6W, creating voids for ligand binding. This variation in architecture could account for the different patterns of toxic vulnerability incurred by KV11 and G6W.

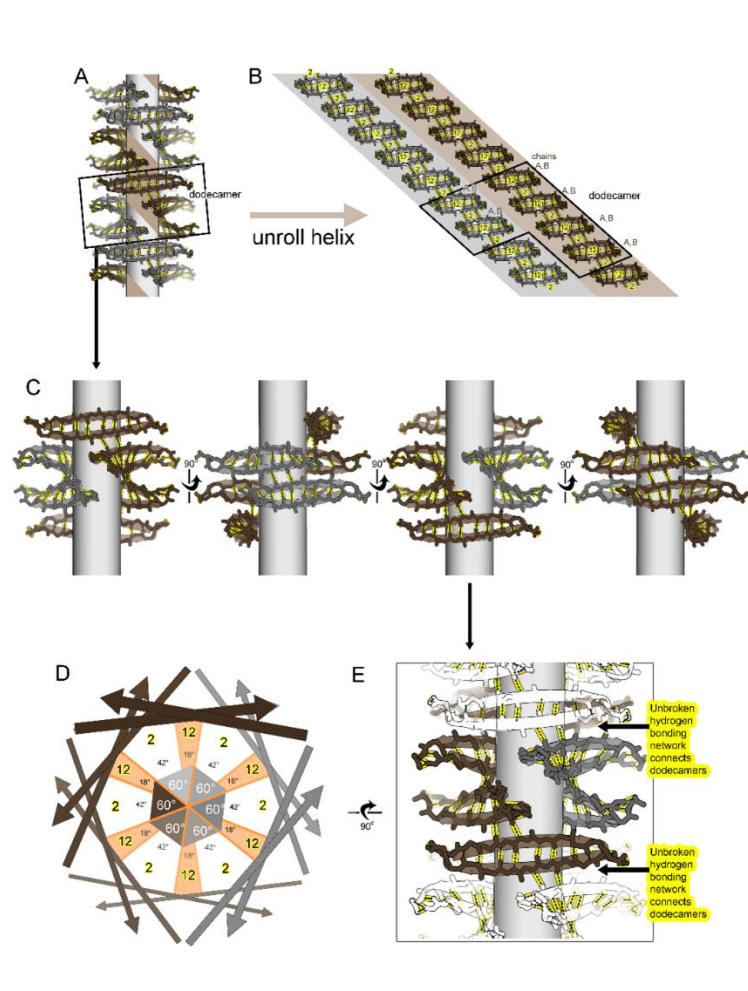


**Figure 6.8.** (A) A column of G6W-Br5 dodecamers in the crystal. The box indicates the boundary of one dodecamer. Backbone hydrogen bonds are highlighted in yellow. Brown and gray colors distinguish the two  $\beta$ -sheets in the dodecamer. The cylinder represents the perfect  $6_4$  symmetry axis. (B) We unrolled the  $\beta$ -sheets into a flat plane in order to highlight the regular, repeating hydrogen bonding pattern. The count of hydrogen bonds between each pair of strands is indicated. (C) Orthogonal views of the dodecamer. (D) A projection down the symmetry axis highlights ideal  $6_4$  symmetry in which the angles between antiparallel pairs of strands are exactly  $60^\circ$ , and hydrogen bond counts (yellow numbers) are evenly distributed around the axis.

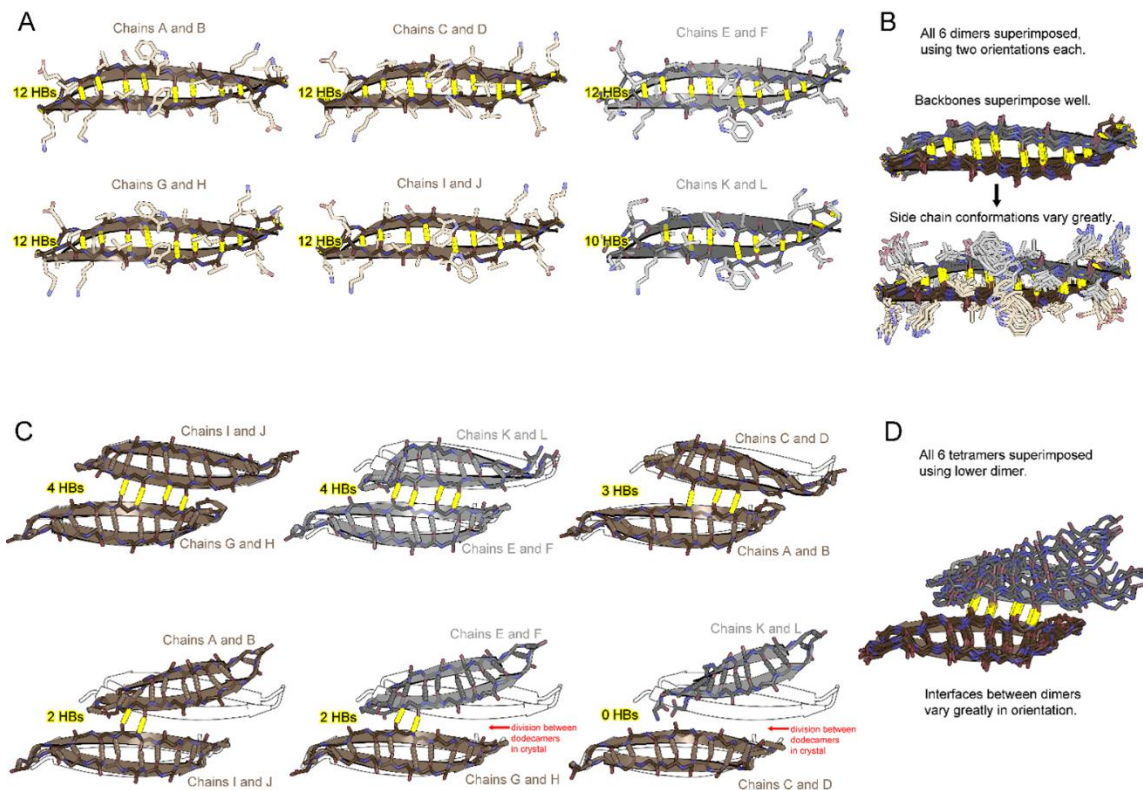


**Figure 6.9. (A) G6W molecules hydrogen-bond together as antiparallel  $\beta$ -sheet dimers. We show all 6 dimers of the dodecamer. Each dimer exhibits 10-12 backbone hydrogen bonds. (B) The dimer backbones superimpose well, but there is variability in side chain conformation. (C) Dimers associate at varied angles and with varied numbers of hydrogen bonds. The unfilled sheet outline provides a reference. (D) the variation in angles between dimers is obvious in the superposition.**

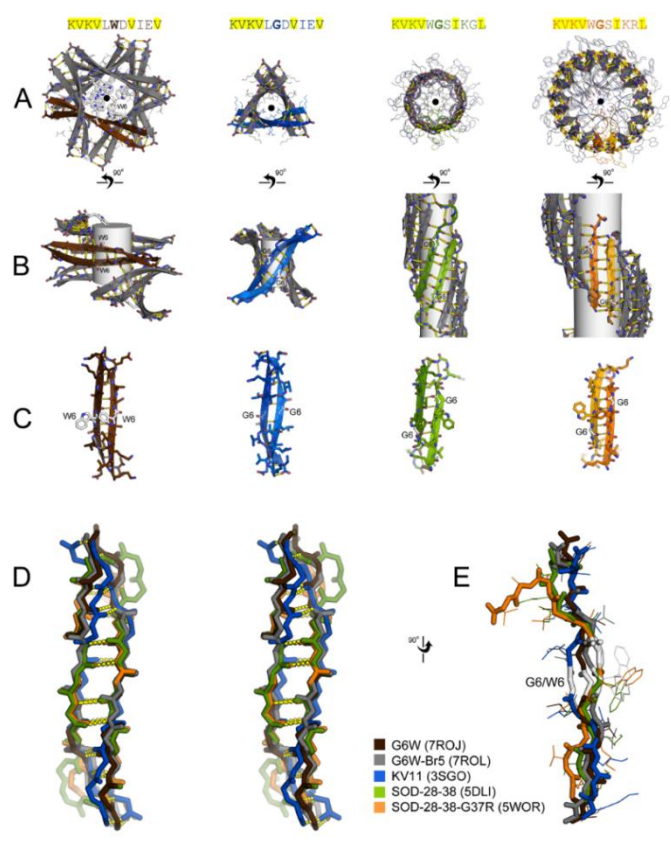


Figure 6.10. Oligomeric assemblies of four homologous peptides (left to right: G6W, KV11, superoxide dismutase (SOD) residues 28-38, and SOD:28-38 with G37R hereditary mutation) are (A) viewed down the primary symmetry axis (black dot) and (B) perpendicular to the axis. Sequence homology among the four segments is illustrated on the top line. All feature (C) a dimeric building block: an antiparallel pair of  $\beta$ -strands joined by hydrogen bonds (yellow dashed line). (D) This dimeric unit exhibits structural variations in the backbone, and these contribute to the gross variations in the larger scale oligomeric assemblies. Compared to the G6W dimer, root mean square deviations (RMSDs) in 44 pairs of backbone atoms are 1.1, 1.6 Å with the closest related sequences G6W-Br5 and KV11. RMSDs are 1.5 and 1.6 Å for the more distantly related wild type and G37R SOD segments, and have fewer atoms aligned. (E) Deviations in the backbone are revealed in an orthogonal view of the aligned monomers. Residue 6 is highlighted by white colored bonds in most of the panels.

Our discovery of a potential lipid-binding site and mechanism for ligand binding led us to speculate that G6W might exert its toxicity by binding and damaging the integrity of cell membranes. However, liposome dye-release experiments show that G6W does not induce membrane disruption (Figure 6.11 and Figure 6.12). KV11 also does not induce membrane leakage, but expectations for such activity seem lower for KV11 given its lack of obvious lipid-binding pockets.<sup>293</sup> We speculate that the oligomers can extract lipids from the membrane, but are unable to create a pore/channel sufficiently large to leak dye. With that in mind, we tested the hypothesis by studying the assembly of G6W with a membrane mimetic, MS-compatible detergent.

### **G6W dodecamer can be identified with IMS-MS**

While G6W does not induce toxicity via causing membrane leakage, here we show that interactions between G6W and membrane mimetics, such as n-Dodecyl- $\beta$ -D-maltoside (DDM), can affect the self-assembly and stability of G6W oligomers and give rise to very interesting observations. DDM is a nonionic detergent that forms micelles that mimic the cellular membrane and is used frequently in mass spectral analysis of intact membrane-protein complexes.<sup>323</sup> Additionally, we have previously shown that DDM can be used to explain the differences in toxicity of a set of phenol soluble modulin peptides (PSM $\alpha$ 3).<sup>84</sup>

The mass spectrum of 15  $\mu$ M G6W with 200  $\mu$ M DDM in 20 mM ammonium acetate shows that multiple G6W-DDM complexes were formed (Figure 6.13). Interestingly, some high-ordered G6W oligomers, such as the tetramer, were observed to be in complex with DDM molecules. This was not observed in our early study of PSM $\alpha$ 3 where the complexes consist of only one and at most two PSM $\alpha$ 3.<sup>84</sup>

Since a variety of G6W oligomers are in complex with DDM molecules, we would further like to understand how G6W oligomer size affects its interactions with DDM. we compared the experimental CCS of intact G6W-DDM complexes to the isotropic model (Figure 6.14).<sup>315</sup> The comparison coarsely elaborates on how DDM molecules

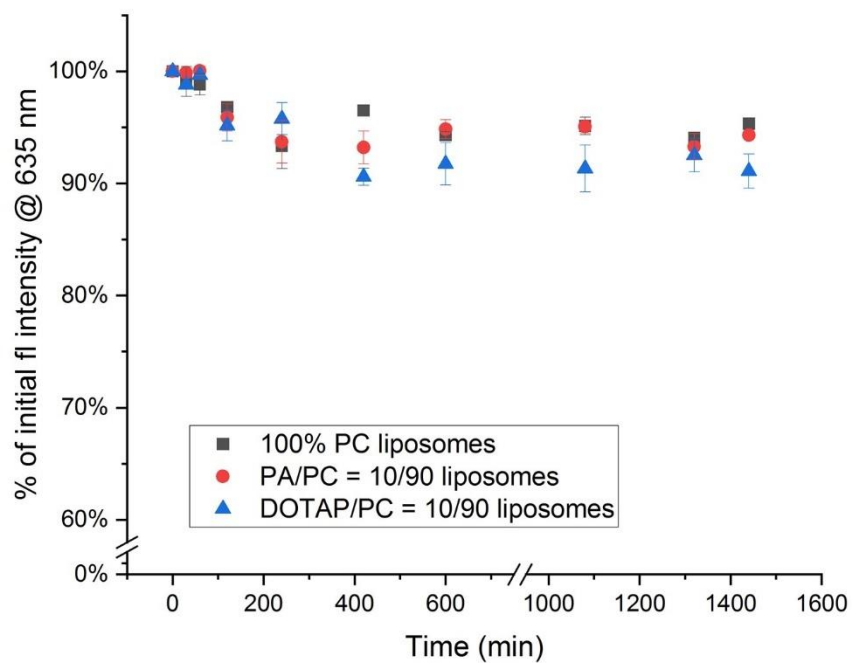


**Figure 6.11. G6W liposome dye release. NR release results for liposomes composed of 100% PC, 10/90 PA/PC, and 10/90 DOTAP/PC after adding 4.5 mM G6W peptide. No significant decrease in NR fluorescence was observed over 24 hours, indicating G6W peptide did not lead to release of encapsulated dye.**

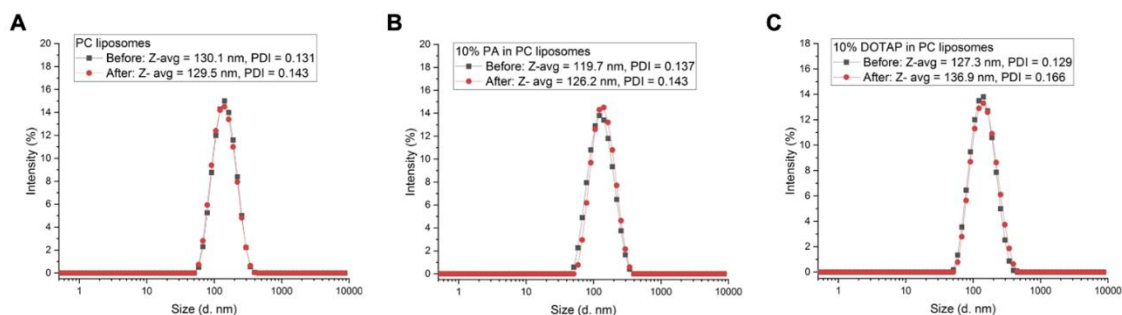


**Figure 6.12. Dynamic light scattering of liposomes before and after G6W treatment.** Raw DLS distribution curves for liposomes before and after peptide treatment. In all cases, stable liposomes with desired sizes were detected before peptide addition. After incubation with peptide for over 24 hrs, no change was observed in any of the samples tested. This further supported that G6W did not alter liposome self-assembly properties based on particle sizes.

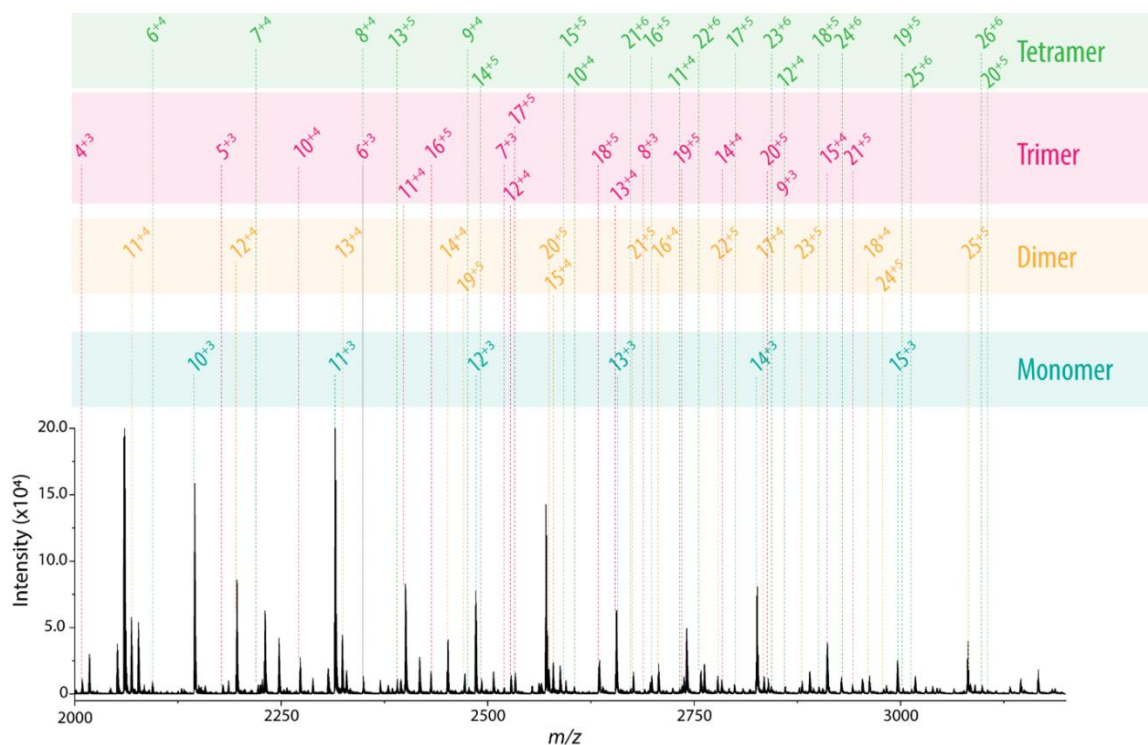


**Figure 6.13.** Mass spectrum of 15  $\mu\text{M}$  G6W and 200  $\mu\text{M}$  DDM in 20 mM ammonium acetate. Each mass spectral peak is labeled by the number of DDM molecules attached to the respective G6W oligomer with the charge state of the complex. For example, the teal “10<sup>+3</sup>” represents a G6W monomer in complex with 10 DDM molecules with  $z = +3$ . Multiple oligomers of G6W (i.e., dimer, trimer, and tetramer) are observed to be complexed with DDM molecules. Note that all G6W-DDM species are not labeled within this mass spectrum.

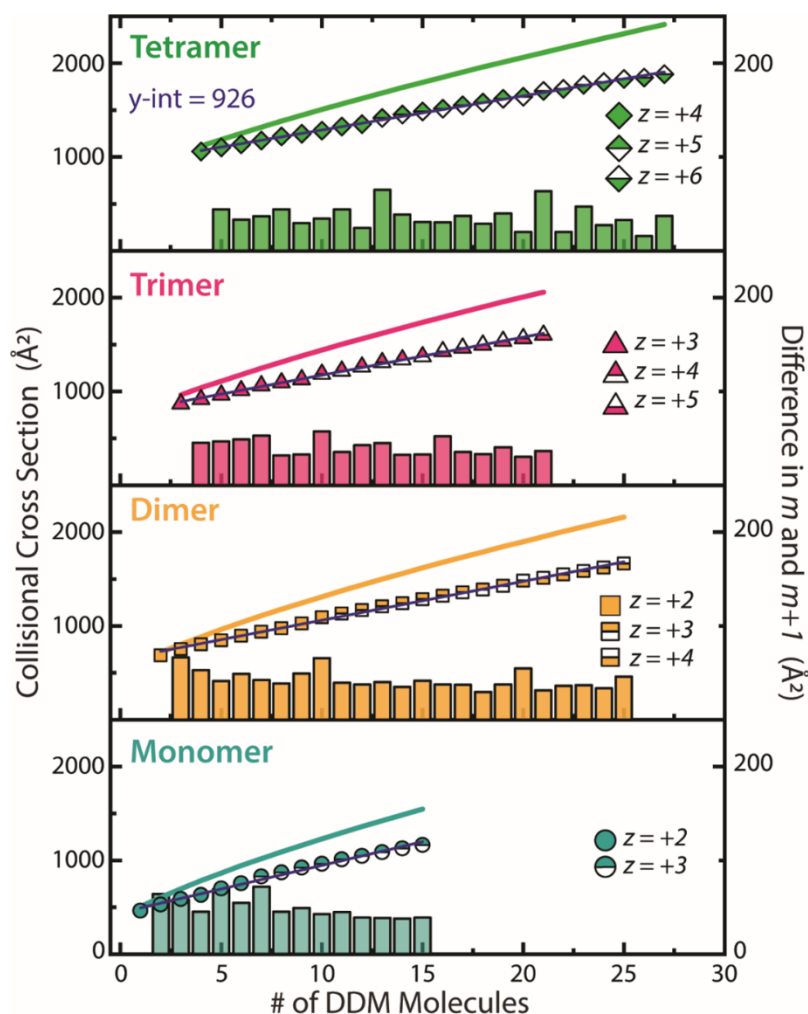
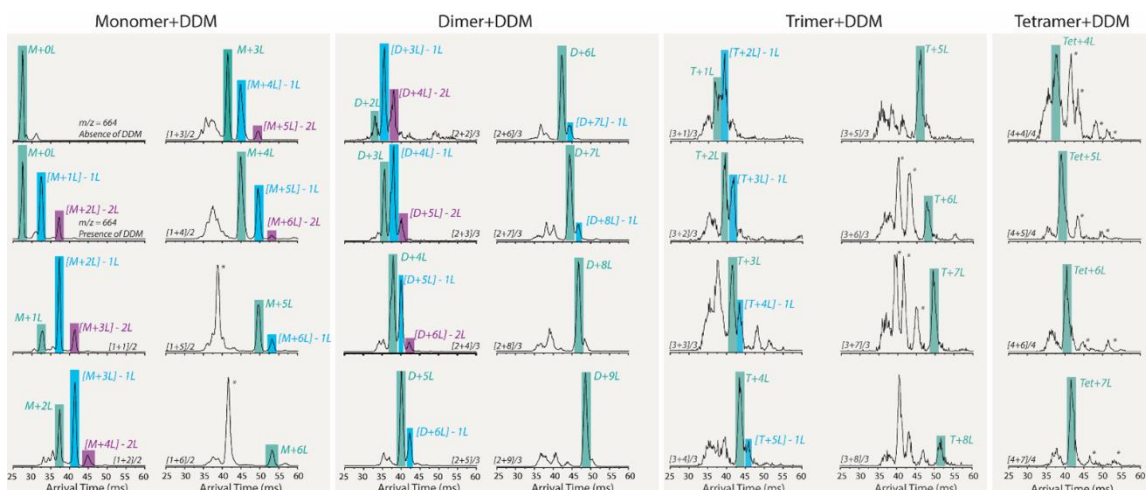


Figure 6.14. The experimental  $\text{CCS}_{\text{N}_2}$  of G6W monomer-, dimer-, trimer-, and tetramer – DDM complexes are shown in each respective plot by symbols. The line-of-best fit for experimental CCS is shown in dark blue for each complex. The isotropic CCS is shown by the solid line above the experimental CCS values. The isotropic values were calculated using  $n/z = 1/2$ ,  $2/1$ ,  $3/2$ , and  $4/2$  for the monomer, dimer, trimer, and tetramer species, respectively (the lowest charge-state species). The bars show the differences in CCS for G6W-DDM complexes with  $m$  and  $m+1$  number of DDM molecules.

assemble around G6W oligomers. Notably, for all surveyed G6W oligomers ( $n = 1$  to  $4$ ), the experimental CCS of the G6W-DDM complexes are well below those predicted by the isotropic model. The isotropic model assumes globular assembly,<sup>315</sup> and due to the poor agreement between the experimental CCS and isotropic CCS we can assume that the DDM molecules do not arrange themselves round the G6W oligomers in a spherical fashion.

Figure 6.14 shows evidence for rearrangement of DDM molecules around the G6W oligomers based on the number of DDM molecules in the complex. Examining the difference in CCS between a G6W-DDM complex with  $m$  number of DDM molecules and  $m+1$  DDM molecules, it can be observed that there are unexpected trends within the CCS differences. For example, the CCS difference for the G6W dimer with  $m = 2 - 4$  DDM molecules gradually decreases, and then there is a slight increase at  $m = 5$  DDMs. One would expect the difference between these CCS values to gradually decrease over time since the contribution of one DDM molecule would become less as the complex grows.<sup>295, 314, 315</sup> These erratic trends with each G6W-DDM complex may be due to how the complex rearranges itself to incorporate new DDM molecules or to adopt new structures. Similar trends were observed in other peptide systems such as PSM $\alpha$ 3.<sup>84</sup> Finally, our data also suggest that DDM can be used to increase the abundance of small oligomers (i.e.,  $n = 4, 6$ ) and may delay the formation of larger oligomers (i.e.,  $n = 8$  to  $12$ ).

As the order of the G6W oligomers within the DDM-bound complexes increases, we note that the complex is more stable. This is deduced from the lack of apparent post-IM dissociated species seen within the tetramer-DDM complexes but present extensively within the monomer-, dimer- and trimer- DDM complexes (Figure 6.15). As mentioned previously, post-IM dissociation is typically an undesired phenomenon whereby weakly-bounded oligomers successfully traverse the IMS drift cell intact but dissociate before they reach the mass analyzer. As a result, dissociated species will have the same drift time as the intact species but a different  $m/z$ .



An example of post-IM dissociation occurring within the G6W monomer-DDM complexes can be seen within the set of ATDs of the monomer-DDM complexes in Figure 6.15. The top ATD shows only one prominent feature of the G6W monomer with  $z = +2$  in the absence of DDM. However, the ATD exhibits three prominent features of this G6W monomer ( $z = +2$ ) in the presence of DDM. The shortest drift time feature is the bare monomer species seen in the top ATD, however the other two features to the right of the bare monomer are assigned as post-IM dissociated species of the monomer-DDM complex with the *neutral* loss of 1 or 2 DDM molecules. As an example, the feature labeled  $[M+1L]-1L$  (where M is the G6W monomer and L is DDM) traversed the drift cell as an intact complex ( $M+1L$ ), resulting in a slower drift time than the bare G6W monomer (larger species will have a longer drift time than smaller species of the same  $z$ ). However, before the  $M+1L$  species arrived to the mass analyzer, it dissociated and lost one *neutral* DDM molecule, leaving it with the same  $m/z$  as the bare G6W monomer but the same drift time as the intact  $M+1L$  species (seen within the third ATD). The agreement of CCSs within each dissociated monomer-DDM complex to the intact complexes are excellent (Figure 6.16), supporting our feature assignments within each ATD.

The lack of post-IM dissociation that occurs within the tetramer-DDM complexes (Figure 6.15) implies that these complexes must be more tightly held together to prevent extensive dissociation. From this, we can infer that larger G6W oligomers have stronger interactions with DDM membrane mimetics than smaller oligomers. However, we would like to note that the features assigned as intact tetramer-DDM complexes are broad, which may make it difficult to observe post-IM dissociated species if any are present.

We note that in the presence of DDM, the ATD of G6W's  $n/z = 1/1$  populates features that are otherwise absent or unresolved in the same ATD obtained in the absence of DDM (Figure 6.17A). There are no G6W-DDM complexes that would have  $m/z = 1,327$ , so these species must be oligomers of purely G6W (but initially) stabilized or promoted

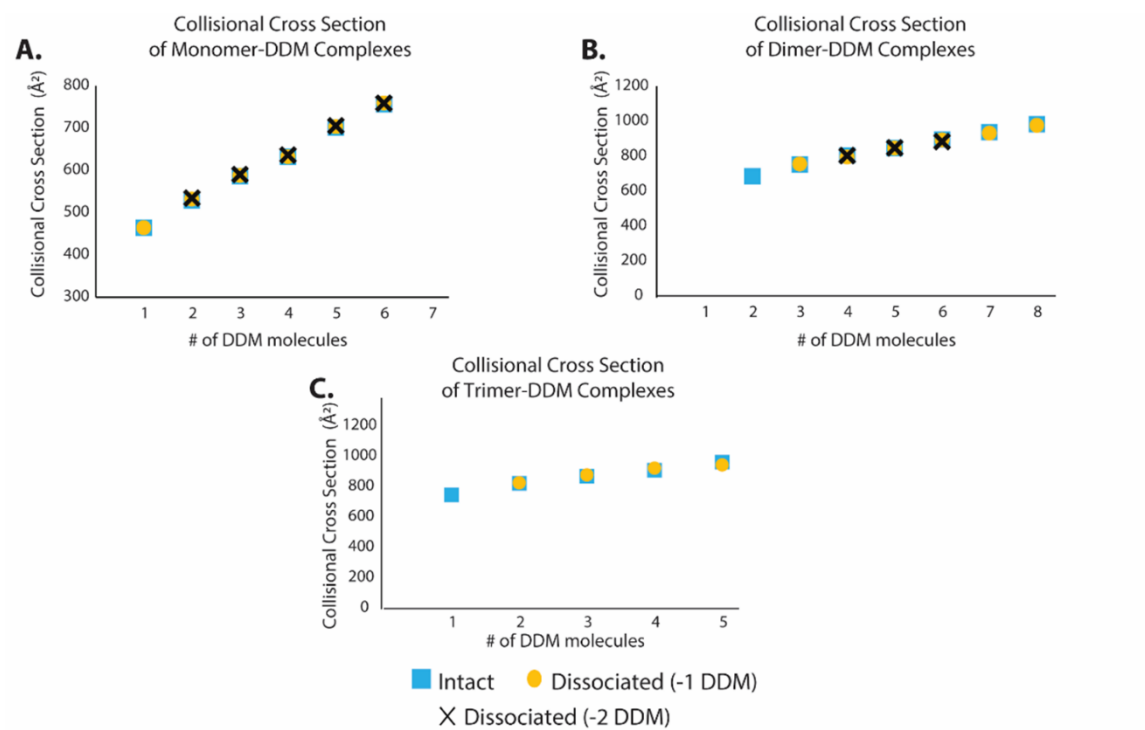


**Figure 6.16.** Plots showing the experimental CCS of intact and dissociated (A) monomer- (B) dimer- and (C) trimer-DDM complexes are shown.

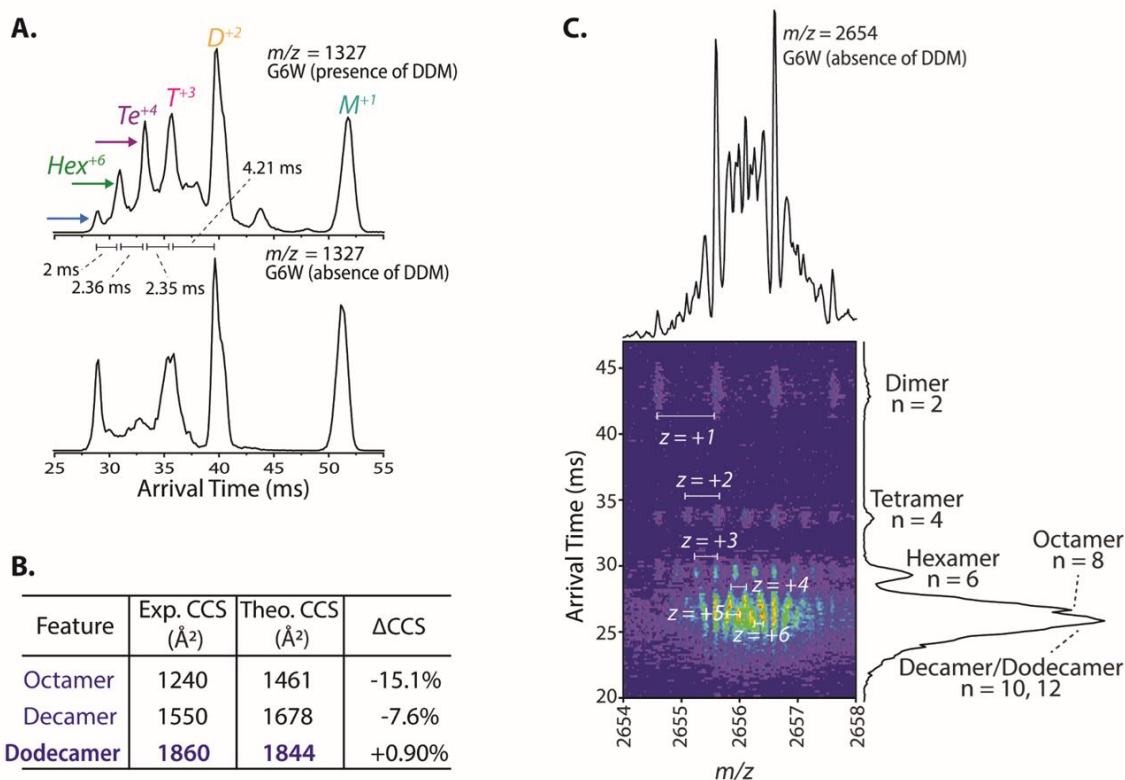


**Figure 6.17. (A)** Representative ATDs of G6W  $n/z = 1/1$  ( $m/z = 1,327$ ) in the presence of DDM (top) and absence of DDM (bottom) in nitrogen buffer gas. The shortest arrival time feature (blue arrow) is tentatively assigned as an octamer, decamer, or dodecamer. (M = monomer, D = dimer, T = trimer, Te = tetramer, Hex = hexamer). **(B)** Possible oligomeric assignments of the shortest arrival time feature (blue arrow) in both Figure 6.2c and 6.17a and comparisons of  $\sigma_{\text{exp,He}}$  and  $\sigma_{\text{theory,He}}$  of the “double-helix” X-ray structures. Note that these CCS experimental values were obtained in helium in the absence of DDM. **(C)** 2D plot of G6W in the absence of DDM at  $m/z = 2,654$ . Each feature within the ATD is assigned according to isotopic spacing detailed by the 2D plot. Note that data shown is collected using helium as the drift gas.

by DDM. The interactions between G6W oligomers and the DDM molecules would slow down the oligomerization of G6W, thus leaving a higher population of more intermediary species such as tetramers and hexamers. Because the interactions are generally weak, the DDM molecules are dissociated before and during IMS measurements. Furthermore, the experimental CCS values of DDM-bound complexes and the new features in the ATD with nominal  $n/z = 1/1$  allow for an extrapolation and determination of CCS values of the G6W tetramer and hexamer that are otherwise not observed in the ATD of  $m/z = 1,327$ .

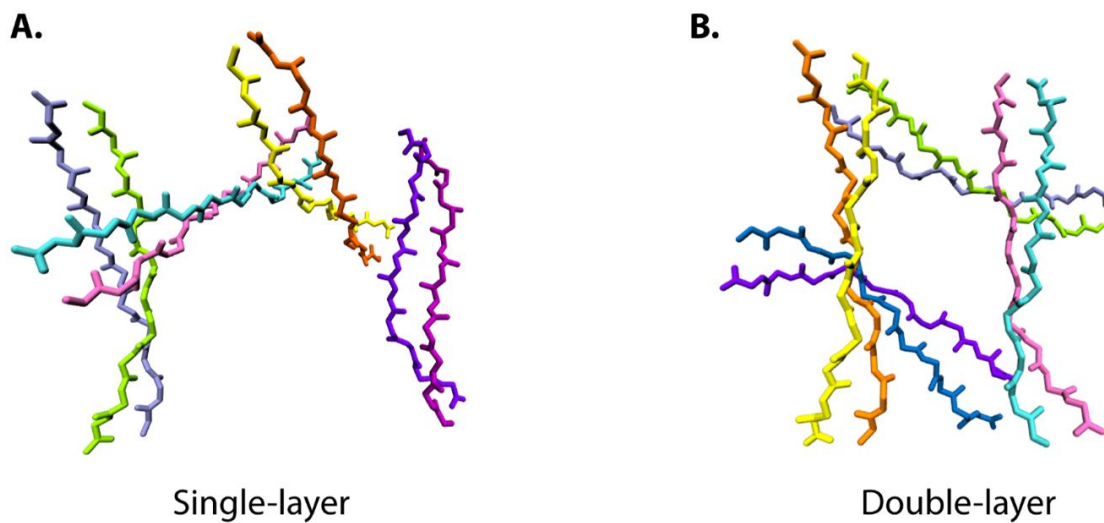
The presence of these novel features (i.e., oligomers larger than trimer in the ATD of nominal  $n/z = 1/1$ ) can be used to better assign the shortest arrival time feature (blue arrow in Figure 6.2C and Figure 6.17A). As noted previously, the arrival time of this feature suggested that this species may be an oligomer with  $n > 6$ . Initial assignments within the ATD are similar to those in helium shown in Figure 6.2C. However, now that more features are present within the ATD between the trimer ( $n/z = 3/3$ ) and shortest arrival time feature, there is more information to better assign these large oligomers. The features left of the trimer are assigned as a tetramer (purple arrow), a hexamer (green arrow), and a feature that can contain either an octamer or larger oligomer (blue arrow), respectively.

Within Figure 6.17, the features left of the trimer are assigned as a tetramer (purple arrow), a hexamer (green arrow), and a feature that can contain either an octamer or larger oligomer (blue arrow), respectively. The y-intercept of the linear regression line through the experimental CCS values of the G6W tetramer-DDM complexes would extrapolate the CCS of a bare tetramer with no bound DDM (Figure 6.14). Thus, the approximated CCS of bare tetramer from the data with DDM is  $926 \text{ \AA}^2$  which supports our assignment of the feature with a purple arrow in Figure 6.17A as a tetramer ( $\sigma_{\text{exp, N2}} = 929 \text{ \AA}^2$ ). Moreover, the green- and blue-arrow feature assignments are logical, based on an analysis of the differences in arrival time of each feature. For example, the difference in arrival time is significantly shorter between the dimer and trimer species

than the monomer and dimer species due to the smaller contribution of an additional G6W unit as the oligomer complex grows directionally.<sup>295, 314, 315</sup> On the other hand, the arrival time difference between the trimer and tetramer (2.35 ms) is approximately the same as the difference between the tetramer and the next feature at ~32 ms (2.36 ms; see Figure 6.17A). Thus, the feature at ~32 ms cannot be a pentamer because if this was the case, the difference in arrival time from the tetramer to the pentamer should have been smaller than that between the trimer and tetramer. Therefore, the feature at ~32 ms (green arrow) is assigned to be a hexamer. Note that going from  $n = 3$  to  $n = 4$ , the arrival time difference is 2.35 ms, and from  $n = 4$  to  $n = 6$  the difference is 2.36 ms. Thus, it appears that a contribution of a monomer subunit to the arrival time (and subsequently CCS) is reduced from ~2.35 ms/unit to ~1.18 ms/unit. Conservatively, we can expect the drift time difference per monomer unit going from  $n = 6$  to a larger oligomer to be smaller than 1 ms/monomer unit. Given the difference in drift time of  $n = 6$  (green arrow; Figure 6.17A) and the shortest drift time feature (blue arrow) is 2 ms, possible oligomers may include an octamer, a decamer, or even a dodecamer.

Comparisons of experimental  $\text{CCS}_{\text{He}}$  of potential oligomers for the shortest drift time feature (blue arrow, Figure 6.2C) and theoretical  $\text{CCS}_{\text{He}}$  are shown in Figure 6.17B. The theoretical CCS from the X-ray structure (Figure 6.18, example of the output of such calculations can be seen within the Appendix) were obtained in helium using the trajectory method.<sup>195, 324</sup> Note that we chose to do the comparison using the helium data because the Mobcal program has the most reliable parameters in helium but not nitrogen.<sup>195, 324</sup> We found that there was an excellent agreement between the experimental and theoretical CCS of the G6W dodecamer (<1% difference, Figure 6.17B). Collectively, this evidence points towards the shortest arrival time feature (blue arrow) in both Figure 6.2C and Figure 6.17B of being a dodecamer, which agrees with the X-ray model of the G6W oligomer.

To further support this assignment, we searched for similar oligomer distributions from mass spectral peaks at higher  $m/z$  (Figure 6.17C). Our reasoning is that an



**C.**

| Possible Assignment | Experimental $\text{CCS}_{\text{He}}$ | Theoretical $\text{CCS}_{\text{He}}$ |         |
|---------------------|---------------------------------------|--------------------------------------|---------|
|                     |                                       | 2-layer                              | 1-layer |
| 8-mer               | 1240                                  | 1461                                 | 1595    |
| 10-mer              | 1550                                  | 1678                                 | 1992    |
| 12-mer              | 1861                                  | 1844                                 | 2192    |

**Figure 6.18. Structural Models Derived from the X-ray Crystal Structure and the Corresponding Theoretical Cross Sections of G6W Oligomers. (A) The single-layer model, (B) The double-layer model, (C) Possible assignments to the largest oligomers observed at  $m/z$  1327 (nominal  $n/z = 1/1$ ) and the CCS values in helium.**

oligomeric species, if existing in solution and ionized by ESI, will populate multiple charge states. Low-charge state species will have more resolved isotope spacing patterns, allowing us to better identify the existence of such oligomers. Within Figure 6.17C, we see isotopic spacings that support (1) the existence of octamer, decamer, and dodecamer and (2) the shortest arrival time feature of  $m/z = 2,654$  (nominal  $n/z = 2/1$ ) being of a dodecamer. Since we can clearly see such large oligomers in the ATD of other mass spectral peaks, the assignment of the feature labeled with the blue arrow in both Figure 6.2C and Figure 6.17C as a dodecamer is not of pure speculation. This also strengthens our claim that oligomers larger than a decamer are difficult to resolve.

### **Identifying molecules released from SCN/AH tissues may indicate neurochemical pathways that G6W acts upon**

To further probe how the G6W oligomers may generate amyloid-disease, LC-IMS-MS was used to identify differences in the chemicals released from SCN/AH-containing brain slices incubated with G6W. In one study performed by Sweedler and co-workers,<sup>325</sup> rat SCN extracellular samples were collected by placing solid-phase extraction beads on the SCN slice. The beads were then prepared to be analyzed via mass spectrometry to identify neuropeptide release. Overall, the detection of peptides in releasates is challenging due to their very low concentration. In our study, we chose to sample the peptide contents released from the acute brain slices. A curated library containing the exact mass of mouse neuropeptides (obtained from the previous literature)<sup>326, 327</sup> were used to identify neuropeptide species in the SCN/AH-releasate. Putative neuropeptides significantly up-regulated in the presence of G6W are listed in Figure 6.19 and Table 6.3.

Peptides, such as tubulin polymerization-promoting proteins (TPPP) family member 3, were also found upregulated in the G6W SCN + AH releasate. The expression of TPPPs is increased in cortical amyloid plaques of AD patients.<sup>328</sup> Interestingly, eliminating TPPP expression in mutant mice has been found to affect the circadian timing system.<sup>329</sup> Disturbance in circadian rhythms is a common

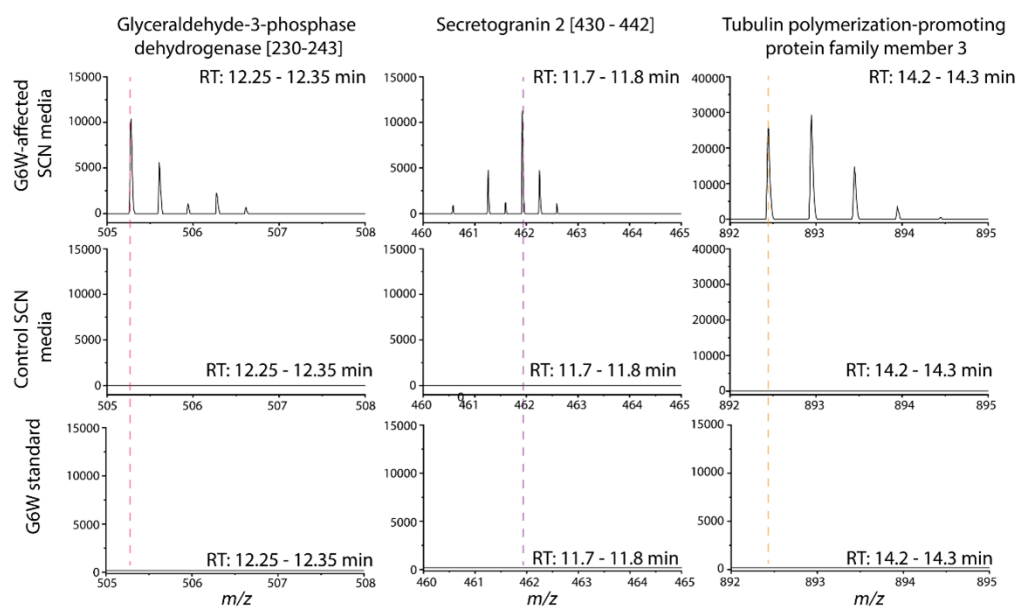


**Figure 6.19. Cross-comparison of putative neuropeptide mass spectra between the G6W-affected SCN media releasate, control SCN media releasate, and the G6W standard.**

**Table 6.3. Neuropeptides upregulated in G6W treatments of the SCN+AH detected in releasates via LC-IMS-MS ( $p < 0.05$ ).**

| Annotation                                                                   | Species       | $m/z$   | ppm   | $CCS_{exp, N_2} (\text{\AA}^2)$ |
|------------------------------------------------------------------------------|---------------|---------|-------|---------------------------------|
| Tubulin polymerization-promoting protein family member 3 [2-17] (Acetylated) | $[M+2H]^{2+}$ | 892.446 | -1.1  | 457.9                           |
| Glyceraldehyde-3-phosphase dehydrogenase [230-243]                           | $[M+3H]^{3+}$ | 505.277 | -14.7 | 497.8                           |
| Secretogranin 2 [430-442]                                                    | $[M+3H]^{3+}$ | 461.922 | 9.6   | 472.6                           |

symptom of AD, and could be linked to altered regulation of proteins, such as TPPP. Additional up-regulated peptides include glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and Secretogranin II, both of which have been investigated in the context of AD. Studies have shown that GAPDH can interact with neurodegenerative disease-associated agents,<sup>330</sup> such as the A $\beta$  protein precursor,<sup>331</sup> but its activity is significantly inhibited in AD brains.<sup>332</sup> Secretogranin II may serve as a potential marker for degradation of neuronal synapses in AD.<sup>333</sup> Future studies focused on these peptides may generate novel insights into AD processes and potential therapeutics.

## 6.6 Conclusion

In summary, our study presents a new model of an amyloid oligomer. X-ray crystallography reveals that the 11-residue peptide G6W (KVKVLWDVIEV) forms large oligomers (dodecamer) composed of 30° twisted antiparallel  $\beta$ -sheets with binding sites that can capture fatty acids or lipids. This is the most structurally complex and unique model of toxic amyloid oligomers reported to date. The differences in oligomer assembly and associated toxicity were investigated by IMS-MS and tissue viability assays. We demonstrated that KV11 and G6W generate distinct levels of cellular damage across regions of the mouse brain that are consistent with AD pathology. More importantly, our work demonstrates that toxic amyloid oligomers can adopt distinct structures, and the structural evolution from one to another can be a result of a single mutation. A mutation targeting one form of toxic amyloids may unexpectedly introduce a new form that is more toxic.

Additionally, we show that interactions between G6W and DDM, a cellular membrane mimetic, lead to an increased abundance of small G6W oligomers that allowed us to more unambiguously determine the sizes of G6W oligomers formed in the absence of DDM. Finally, we assessed functional changes induced by the amyloid model in the mouse SCN/AH by probing the release of various neuropeptides. While G6W is not an endogenous peptide, it shares a high sequence compatibility with the

amyloidogenic cores of A $\beta$  (i.e., A $\beta_{16-22}$ ; KLVFFAE)<sup>334-336</sup> and of tau (i.e., PHF6; VQIVYK),<sup>337, 338</sup> which contain aromatic residues known to be crucial for fibril formation and toxicity. Amyloidogenic fragments can exist in vivo as the result of peptidase activities. Collectively, the cytotoxicity of G6W resembles the amyloid effects in AD brains. Its dodecameric structure provides a new model of toxic amyloid oligomers with a novel mechanism to bind fatty acids or lipids.

## 7 Conclusions

### 7.1 Summary

Within this dissertation, the overarching motives was to (i) leverage and push IMS-MS to its limit as a tool to investigate noncovalent-bonded assemblies in biologically relevant environments and (ii) highlight the power of IMS-MS when combined with condensed-phase techniques such as NMR and X-ray crystallography.

We began in Chapter 2, where we were able to probe the conformational “dark-space” of CycA. Although CycA is a very well-characterized molecule, little is known about the short-lived conformations that this molecule can exhibit. We showed that IMS-MS was able to capture a unique and new conformation of CycA termed the “bent” conformation. It was determined that different metals can affect the conformational preference of CycA, which is of interest because CycA is known to disrupt electrolytic homeostasis within the body. More specifically,  $\text{Ca}^{2+}$  ions were shown to induce several CycA conformations (bent, closed, and open), which may be relevant to the binding of CycA to its protein receptor, CypA. NMR of CycA: $\text{Ca}^{2+}$  showed that the “open” conformer was predominant in solution. Ultimately, this study highlights the robustness of IMS-MS and how it can be used to capture subtle conformational changes of macrocyclic peptides and the interactions they can have with guest ligands and protein targets. The relevance of salt cations and anions transforming the conformational landscape of proteins are also demonstrated.

In Chapter 3, the differences in toxicity of four phenol soluble modulin peptides (called PSM $\alpha$ 3) were discerned with IMS-MS. We showed that a phenomenon called post-IM dissociation was occurring within our instrument, which makes ATD feature assignment of amyloid oligomers difficult. However, we were able to use this typically undesired phenomenon to develop a method to explain the differences of PSM $\alpha$ 3 toxicity by investigating their interactions with DDM micelles, which serve as a cellular

membrane mimetic. We found that the more toxic PSM $\alpha$ 3 peptides required fewer DDM molecules to stabilize them than the nontoxic peptide. This leads us to speculate that the toxic peptides can insert themselves more easily into the cellular membrane.

The remaining chapters are concerned with the investigation of noncovalently-bound assemblies of soluble protein aggregates which underlie many different human diseases. This discussion is first begun in Chapters 4 and 5, where the molecular cross-talk of related peptides and how they impact the behavior of each other are probed. Chapter 4 is concerned with a peptide called amylin, which is involved in the underlying etiology of Type II Diabetes, and  $\alpha$ -CGRP, a peptide that is cosecreted in the pancreas with amylin. Both amylin and  $\alpha$ -CGRP have very similar sequences and induce similar effects within the pancreas. Through an in-depth characterization of each peptide, we rationalized why amylin can fibrillate but  $\alpha$ -CGRP does not. It is observed that  $\alpha$ -CGRP remodels amylin fibrillation – disrupting amylin fibrils to form globular aggregates. Furthermore, we investigated the behavior of these peptides in a solution of  $\text{ZnCl}_2$  ( $\text{Zn}^{2+}$  is a relevant cation within pancreas islets) to show that intermediate conformers of amylin may provide an explanation as to why amylin does not fibrillate spontaneously within the pancreas. One important facet to this study may reveal the relationship between amylin and  $\alpha$ -CGRP in the context of migraine and diabetes. It was found that  $\alpha$ -CGRP helped regulate insulin secretion. The elevated levels of CGRP in migraine patients may help exert a protective effect by regulating insulin levels, thus reducing the risk of Type II Diabetes in those who suffer from migraine.

Similarly, Chapter 5 investigated the interactions between  $\text{A}\beta_{42}$  and  $\text{Tau}_{\text{PHF43}}$ , two proteins involved in Alzheimer's Disease. We showed that  $\text{Tau}_{\text{PHF43}}$  formed dimers which are crosslinked by disulfide bonds, and of which have been showed by previous studies to be important to the pathology and toxicity of Tau. However, it was observed that  $\text{A}\beta_{42}$  can disrupt these disulfide bonds of  $\text{Tau}_{\text{PHF43}}$  dimers and  $\text{Tau}_{\text{PHF43}}$  fibrils. Thus,  $\text{A}\beta_{42}$  oligomers may exhibit a protective effect within the brain to disrupt toxic Tau species.

Finally, Chapter 6 involves the investigation of G6W, a peptide which can self-assemble into a large dodecameric species that serves as a model for amyloid oligomers. IMS-MS showed subtle differences between G6W and KV11, another model for amyloid oligomers in which G6W was derived from. Interactions between G6W and DDM micelles helped support that G6W has the ability to form very large oligomers that is different from KV11. X-ray crystallography was used to discern that G6W assembles into asymmetric dodecamers. Upon comparison of the X-ray crystallography data and IMS-MS, we were able to elucidate that IMS-MS was able to capture these dodecameric species. Furthermore, we investigated the feasibility of G6W (and KV11) as a model to be used to study amyloid oligomers through tissue viability assays. These viability assays showed that both G6W and KV11 were able to induce this effect of selective cellular vulnerability, which is consistent with what is seen in Alzheimer's Disease. Additionally, neuropeptides (such as TAPP and secretogranin) were found to be upregulated in brain slices affected by G6W. Investigations of these neuropeptides may provide further insights into the mechanisms of Alzheimer's Disease and may provide clues to how G6W initiates cellular injury.

## **7.2 Contributions to the field**

Within this dissertation, IMS-MS was used as a primary platform in order to investigate biological assemblies to answer questions that are of a broad biological and medical significance. Although these projects utilized different biological species as the analyte of interest, the outcomes and contributions to the field from each study are cohesive and can be summarized as such:

- (I) The power and robustness of IMS-MS can be used as a tool to not only study the self-assembly of amyloidogenic species in a buffer solution, but also how these species can interact with micelle-forming detergents, living mice brain tissues, and salt-heavy environments. This in turn can provide insights to how

amyloid oligomers may behave within a more applicable biological environment.

- (II) Post-IM dissociation is an event that can happen routinely within IMS-MS instrumentation. Although Chapter 3 shed new insights on the interpretation of IMS-MS data affected by post-IM dissociation through studying PSM $\alpha$ 3, we were able to also witness this event occurring within other amyloidogenic systems (such as G6W and Tau<sub>PHF43</sub>). Most IMS-MS biomolecular self-assembly do not take this event into account when assigning ATD features, which can lead to incorrect feature assignments which can change the outcome of a study.
- (III) It is important to not only consider one toxic species implicated within a disease, but how these species interact with other assemblies they are likely to come across *in vivo*. By taking this into account, a larger picture can be obtained upon disease mechanisms and can perhaps provide new insights on disease relationships and future therapeutics.
- (IV) The results of IMS-MS data can be complemented and enhanced by coupling them to the results of other condensed-phase experimental techniques, such as NMR spectroscopy and X-ray crystallography. Doing such robust analysis spanning different experimental techniques provides strong evidence that gas-phase IMS-MS data retains solution and solid-phase information.

### 7.3 Future outlooks

The projects within this dissertation have laid a solid foundation for future investigations that stem from the outcomes of these studies. For example, there are many unanswered questions regarding the toxicity mechanism G6W dodecameric assembly (discussed within Chapter 6). We initially hypothesized that due to the presence of a hydrophobic cavity, G6W may be able to compromise the cellular membrane through the sequestration of fatty acids within the phospholipid bilayer. To

investigate lipid/fatty acid sequestration as a potential route of toxicity for G6W, we can compare the abilities of the asymmetric G6W dodecamer and symmetric G6W-Br5 dodecamer to bind to fatty acids. G6W-Br5 is a bromo-derivative of G6W that was initially used to further refine the X-ray structure of G6W. However, due to its symmetric shape it does not contain a binding cavity, allowing it to be used as a negative control for G6W. Exploring the chemistry that is essential of fatty acid binding to G6W will be a topic of my lab mate, Sarah Hirschbeck's, project. For example, how will G6W bind to different fatty acids and lipids demonstrating varying degrees of saturation? DPPC, POPC, and DOPC lipids have either zero, one, or two double bonds within their acyl chains, thus they can be used as candidate to investigate this. Additionally, how the length of the acyl chains within the lipids and how this may affect binding to G6W can be explored as well. To investigate the relevance of the G6W hydrophobic binding cavity to its toxicity, cellular viability assays can be performed.

Also related to the G6W project was the chemical profiling of the SCN+AH releasate, where the surrounding media of the SCN+AH tissues were obtained and analyzed through LC-IMS-MS to identify neuropeptides that was released in the presence of G6W. This experiment showed the feasibility of G6W to be used as a model for amyloid oligomers since it prompted the release of neuropeptides that are relevant within Alzheimer's Disease. However, we were only able to identify only several neuropeptide fragments, presumably due to the low concentration of neuropeptides within the releasate. Future studies can focus on optimizing this workflow in order to increase the number of biomolecules observed so tandem-mass spectrometry (MS/MS) can be used to confirm that the species observed is what it is being labeled. Ideas for this include homogenizing the SCN+AH tissue slice and then extracting biomolecules from this homogenate. Theoretically, this should increase the concentration of biomolecules and neuropeptides within the sample. Additionally, the neuropeptide library used to identify neuropeptides within the mouse SCN+AH releasate was based off neuropeptides known to exist within the rat SCN. Since only neuropeptide fragments were detected, it is

important to make sure that the neuropeptides within this mouse can be cleaved in the same positions as of the rat. Furthermore, these neuropeptide fragments had to manually be entered into the neuropeptide database, which made the inclusion of all possible post translational modifications extremely time-consuming. More robust data analysis software should be investigated for future iterations of similar experiments to scan a wide variety peptide (and other biomolecules) that are relevant to the mouse model that was used for this experiment in order to increase the chances of identifying analyte molecules.

The projects within this manuscript have consistently demonstrated that IMS-MS is sensitive to the oligomerization of peptides. Thus, it may be able to be used as a technique to screen short peptides in order to identify mutations that may trigger novel amyloid oligomer structures. For example, the IMS-MS data showed that G6W was able to populate large and complex features when compared to that of KV11 upon viewing of the 2D ion mobility plot of  $n/z = 1/1$ . Future studies can investigate short peptide fragments containing pathogenic mutations (for example, amyloidogenic segments of SOD1 containing familial mutations) to initially screen them to see if they may be able to populate toxic amyloidogenic oligomers. IMS-MS is an inexpensive and quick technique; thus it can save time and resources as opposed to initially analyzing these mutant peptides with X-ray crystallography.

Within Chapter 2, the molecular space of CycA, a very well characterized peptide, was investigated through IMS-MS and NMR spectroscopy. We showed the importance of metal binding in influencing the conformational landscape of CycA. Although we briefly investigated several metals within this study, such as  $\text{Na}^+$  and  $\text{K}^+$ , we primarily focused primarily on  $\text{Ca}^{2+}$ . To expand on this study, we could further investigate how biologically relevant cations interact with CycA to affect its structure. The CycA residues that these metals interact with can be elucidated by performing MS/MS, which would involve the fragmentation of the CycA-metal complex into smaller fragments through collision induced dissociation. These fragments can then be analyzed to discern the

position within the molecule that the metal interacts with. This can give information on how these metals induce *cis/trans* isomerization of the peptide backbone which CycA relies on to assume its many different conformations.

The interactions between A $\beta$ <sub>42</sub> and Tau<sub>PHF43</sub> was investigated within Chapter 5 and it was elucidated that A $\beta$ <sub>42</sub> “disassembles” toxic Tau<sub>PHF43</sub> species. Thus, it is speculated that A $\beta$ <sub>42</sub> may serve as a neuroprotective factor. It would be intriguing to further investigate this hypothesis. Preliminary investigations are already under way to understand how A $\beta$ <sub>42</sub>, Tau<sub>PHF43</sub>, and the mixture of both affect the SCN, AH, and cerebral cortex areas of the mouse brain through tissue viability assays. Furthermore, we can investigate the inherent toxicity of disulfide-bonding within the Tau<sub>PHF43</sub> species by performing viability assays among the native Tau<sub>PHF43</sub> and Tau<sub>PHF43</sub> mutants which do not contain the cysteine residue (and consequently will lose the ability to form disulfide bonds). Due to the hypothesis that disulfide bonding plays a role in Tau pathology from previous reports, it would also be interesting to know how disulfide bonds affect Tau oligomerization. For example, would 4R Tau constructs oligomerize faster than 3R Tau constructs? Would 4R Tau oligomers contain disulfide bonds, unlike the larger oligomers formed from Tau<sub>PHF43</sub>? These are important questions that high-resolution mass spectral data can provide which could give insight onto the differences of 4R vs. 3R Tauopathies.

## **Bibliography**

1. E. A. Mason and H. W. Schamp, *Annals of Physics*, 1958, **4**, 233.
2. S. Warnke, J. Seo, J. Boschmans, F. Sobott, J. H. Scrivens, C. Bleiholder, M. T. Bowers, S. Gewinner, W. Schöllkopf, K. Pagel and G. von Helden, *Journal of the American Chemical Society*, 2015, **137**, 4236-4242.
3. P. Kebarle and L. Tang, *Analytical Chemistry*, 1993, **65**, 972A-986A.
4. A. L. H. Gray, C. A. Steren, I. W. Haynes, G. A. Bermejo, F. Favretto, M. Zweckstetter and T. D. Do, *The Journal of Physical Chemistry B*, 2021, **125**, 1378-1391.
5. C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, *Advanced Drug Delivery Reviews*, 1997, **23**, 3.
6. B. C. Doak, J. Zheng, D. Dobritzsch and J. Kihlberg, *Journal of Medicinal Chemistry*, 2016, **59**, 2312.
7. A. T. Bockus, C. M. McEwen and R. S. Lokey, *Current Topics in Medicinal Chemistry*, 2013, **13**, 821.
8. A. R. Bogdan, N. L. Davies and K. James, *Organic and Biomolecular Chemistry*, 2011, **9**, 7727.
9. F. Giordanetto and J. Kihlberg, *Journal of Medicinal Chemistry*, 2014, **57**, 278.
10. D. S. Nielsen, H. N. Hoang, R. J. Lohman, T. A. Hill, A. J. Lucke, D. J. Craik, D. J. Edmonds, D. A. Griffith, C. J. Rotter and R. B. Ruggeri, *Angewandte Chemie International Edition*, 2014, **53**, 12059.
11. M. Dreyfuss, E. Härri, H. Hofmann, H. Kobel, W. Pache and H. Tschertter, *European Journal of Applied Microbiology and Biotechnology*, 1976, **3**, 125.
12. D. M. Canafax and N. L. Ascher, *Clinical pharmacology*, 1983, **2**, 515.
13. M. D. Colombo, N. Cassano, G. Bellia and G. A. Vena, *The Scientific World Journal*, 2013, **2013**, 805705.
14. J. Cush, P. Tugwell, M. Weinblatt and D. Yocum, *The Journal of Rheumatology*, 1999, **26**, 1176.
15. B. D. Kahan, *American Journal of Kidney Diseases*, 1984, **3**, 444.

16. N. Takahashi, T. Hayano and M. Suzuki, *Nature*, 1989, **337**, 473.
17. G. Fischer, B. Wittmann-liebold, K. Lang, T. Kiefhaber and F. X. Schmid, *Nature*, 1989, **337**, 476.
18. J. Liu, J. D. Farmer, W. S. Lane, J. Friedman, I. Weissman and S. L. Schreiber, *Cell*, 1991, **66**, 807.
19. H. Kessler, M. Köck, T. Wein and M. Gehrke, *Helvetica Chimica Acta*, 1990, **73**, 1818.
20. H. Kessler, H. R. Loosli and H. Oschkinat, *Helvetica Chimica Acta*, 1985, **68**, 661.
21. S. Y. Ko and C. Dalvit, *International Journal of Peptide and Protein Research*, 1992, **40**, 380.
22. H. R. Loosli, H. Kessler, H. Oschkinat, H. P. Weber, T. J. Petcher and A. Widmer, *Helvetica Chimica Acta*, 1985, **68**, 682.
23. A. Whitty, M. Zhong, L. Viarengo, D. Beglov, D. R. Hall and S. Vajda, *Drug Discovery Today*, 2016, **21**, 712.
24. F. Bernardi, E. Gaggelli, E. Molteni, E. Porciatti, D. Valensin and G. Valensin, *Biophysical Journal*, 2006, **90**, 1350.
25. J. A. Carver, N. H. Rees, D. L. Turner, S. J. Senior and B. Z. Chowdhry, *Journal of the Chemical Society, Chemical Communications*, 1992, **22**, 1682.
26. V. Mikol, J. Kallen, G. Pflügl and M. D. Walkinshaw, *Journal of Molecular Biology*, 1993, **234**, 1119.
27. J. Witek, B. G. Keller, M. Blatter, A. Meissner, T. Wagner and S. Riniker, *Journal of Chemical Information and Modeling*, 2016, **56**, 1547.
28. W. H. O. Geneva, *Model List of Essential Medicines, 21st List*, 2019, 2019.
29. C. T. Lee, H. Y. Ng, Y. H. Lien, L. W. Lai, M. S. Wu, C. R. Lin and H. C. Chen, *Journal of the American Society of Nephrology*, 2011, **34**, 87.
30. B. Einollahi, E. Nemati, Z. Rostami, M. Teimoori and A. R. Ghadian, *International Journal of Organ Transplantation Medicine*, 2012, **3**, 166.

31. R. A. Shaw, H. H. Mantsch and B. Z. Chowdhry, *International Journal of Biological Macromolecules*, 1994, **16**, 143.
32. R. M. Cusack, L. Grøndahl, D. P. Fairlie, G. R. Hanson and L. R. Gahan, *Journal of Inorganic Biochemistry*, 2003, **97**, 191.
33. I. W. Haynes, G. Wu, M. A. Haque, H. Li and T. D. Do, *Analytical Chemistry*, 2019, **91**, 13439.
34. B. A. Link, A. J. Sindt, L. S. Shimizu and T. D. Do, *Physical Chemistry Chemical Physics*, 2020, **22**, 9290.
35. P. C. D. Hawkins and A. Nicholls, *Journal of Chemical Information and Modeling*, 2012, **52**, 2919.
36. P. C. D. Hawkins, A. G. Skillman, G. L. Warren, B. A. Ellingson and M. T. Stahl, *Journal of Chemical Information and Modeling*, 2010, **50**, 572.
37. M. F. Mesleh, J. M. Hunter, A. A. Shvartsburg, G. C. Schatz and M. F. Jarrold, *The Journal of Physical Chemistry*, 1996, **100**, 16082.
38. A. A. Shvartsburg and M. F. Jarrold, *Chemical Physics Letters*, 1996, **261**, 86.
39. C. D. Schwieters, G. A. Bermejo and G. M. Clore, *Protein Science*, 2018, **27**, 26.
40. G. A. Bermejo and C. D. Schwieters, *Methods in Molecular Biology*, 2018, **1688**, 311.
41. G. A. Bermejo, G. M. Clore and C. D. Schwieters, *Protein Science*, 2012, **21**, 1824.
42. A. Grishaev and A. Bax, *Journal of the American Chemical Society*, 2004, **126**, 7281.
43. F. Favretto, J. D. Baker, T. Strohäker, L. B. Andreas, L. J. Blair, S. Becker and M. Zweckstetter, *Angewandte Chemie International Edition*, 2020, **59**, 5643.
44. C. R. Goodwin, L. S. Fenn, D. K. Derewacz, B. O. Bachmann and J. A. McLean, *Journal of Natural Products*, 2012, **75**, 48.
45. S. J. Hyung, X. Feng, Y. Che, J. G. Stroh and M. Shapiro, *Analytical and Bioanalytical Chemistry*, 2014, **406**, 5785.

46. M. E. C. Caines, K. Bichel, A. J. Price, W. A. McEwan, G. J. Towers, B. J. Willett, S. M. V. Freund and L. C. James, *Nature Structural & Molecular Biology*, 2012, **19**, 411.
47. B. M. Campos, M. L. Sforça, A. L. B. Ambrosio, M. N. Domingues, T. d. A. C. Brasil de Souza, J. A. R. G. Barbosa, A. F. P. Leme, C. A. Perez, S. B. M. Whittaker and M. T. Murakami, *Plant Physiology*, 2013, **162**, 1311.
48. S. W. Fesik, R. T. Gampe, H. L. Eaton, G. Gemmecker, E. T. Olejniczak, P. Neri, T. F. Holzman, D. A. Egan, R. Edalji and R. Simmer, *Biochemistry*, 1991, **30**, 6574.
49. K. Kajitani, M. Fujihashi, Y. Kobayashi, S. Shimizu, Y. Tsujimoto and K. Miki, *Proteins*, 2008, **70**, 1635.
50. M. Lammers, H. Neumann, J. W. Chin and L. C. James, *Nature Chemical Biology*, 2010, **6**, 331.
51. V. Parashar, C. Aggarwal, M. J. Federle and M. B. Neiditch, *Proceedings of the National Academy of Sciences*, 2015, **112**, 5177.
52. S. S. Sekhon, H. Kaur, T. Dutta, K. Singh, S. Kumari, S. Kang, S. G. Park, B. C. Park, D. G. Jeong and A. Pareek, *Acta Crystallographica. Section D, Biological crystallography*, 2013, **69**, 555.
53. C. M. Stegmann, R. Lührmann and M. C. Wahl, *PLoS One*, 2010, **5**, e10013.
54. C. M. Stegmann, D. Seeliger, G. M. Sheldrick, B. L. de Groot and M. C. Wahl, *Angewandte Chemie International Edition*, 2009, **48**, 5207.
55. V. Venugopal, A. K. Datta, D. Bhattacharyya, D. Dasgupta and R. Banerjee, *Acta Crystallographica. Section D, Biological crystallography*, 2009, **65**, 1187.
56. A. K. Wernimont, J. Lew, I. Kozieradzki, Y. H. Lin, X. Sun, C. Khuu, Y. Zhao, M. Schapira, C. H. Arrowsmith, A. M. Edwards, J. Weigelt, A. Bochkarev, R. Hui, J. D. Artz and T. Xiao.
57. P. Monzani, H. M. Pereira, K. P. Garamacho, G. V. Meirelles, G. Oliva and J. C. C. Cascardo, 2011.

58. T. Rezai, J. E. Bock, M. V. Zhou, C. Kalyanaraman, R. S. Lokey and M. P. Jacobson, *Journal of the American Chemical Society*, 2006, **128**, 14073.
59. C. Weber, G. Wider, B. von Freyberg, R. Traber, W. Braun, H. Widmer and K. Wuethrich, *Biochemistry*, 1991, **30**, 6563.
60. D. Seebach, S. Y. Ko, H. Kessler, M. Köck, M. Reggelin, P. Schmieder, M. D. Walkinshaw, J. J. Bölderli and D. Bevec, *Helvetica Chimica Acta*, 1991, **74**, 1953.
61. A. M. Cirigliano and G. M. Cabrera, *Rapid Communications in Mass Spectrometry*, 2014, **28**, 465.
62. M. Koeck, H. Kessler, D. Seebach and A. Thaler, *Journal of the American Chemical Society*, 1992, **114**, 2676.
63. R. Dancer, A. Jones and D. Fairlie, *Australian Journal of Chemistry*, 1995, **48**, 1835.
64. N. Marjanović, A. Čikoš and S. Koštrun, *Journal of Molecular Structure*, 2019, **1198**, 126929.
65. F. Bernardi, N. D'Amelio, E. Gaggelli, E. Molteni and G. Valensin, *The Journal of Physical Chemistry B*, 2008, **112**, 828.
66. R. A. Altschuld, C. M. Hohl, L. C. Castillo, A. A. Garleb, R. C. Starling and G. P. Brierley, *American Journal of Physiology*, 1992, **262**, H1699.
67. D. Bokemeyer, U. Friedrichs, A. Backer, H. J. Kramer and H. Meyer-Lehnert, *Journal of Clinical Investigation*, 1994, **72**, 992.
68. H. Meyer-Lehnert and R. W. Schrier, *Hypertension*, 1989, **13**, 352.
69. J. Pfeilschifter and U. T. Rüegg, *Biochemical Journal*, 1987, **248**, 883.
70. J. M. Frapier, C. Choby, M. E. Mangoni, J. Nargeot, B. Albat and S. Richard, *FEBS Letters*, 2001, **493**, 57.
71. H. G. Olbrich, H. Geerts, U. Waldmann, E. Mutschler, L. V. Donck, G. Kober and M. Kaltenbach, *Transplantation*, 1991, **51**, 972.
72. C. L. Ahlback, K. W. Lexa, A. T. Bockus, V. Chen, P. Crews, M. P. Jacobson and R. S. Lokey, *Future Medicinal Chemistry*, 2015, **7**, 2121.

73. C. K. Wang, J. E. Swedberg, P. J. Harvey, Q. Kaas and D. J. Craik, *The Journal of Physical Chemistry B*, 2018, **122**, 2261.
74. J. Witek, M. Mühlbauer, B. G. Keller, M. Blatter, A. Meissner, T. Wagner and S. Riniker, *ChemPhysChem*, 2017, **18**, 3309.
75. S. T. Esswein, H. V. Florance, L. Baillie, J. Lippens and P. E. Barran, *Journal of Chromatography A*, 2010, **1217**, 6709.
76. M. A. Wear, M. W. Nowicki, E. A. Blackburn, I. W. McNae and M. D. Walkinshaw, *FEBS Open Bio*, 2017, **7**, 533.
77. F. F. Vajdos, S. Yoo, M. Houseweart, W. I. Sundquist and C. P. Hill, *Protein Science*, 1997, **6**, 2297.
78. K. J. Laszlo, E. B. Munger and M. F. Bush, *Journal of the American Chemical Society*, 2016, **138**, 9581.
79. E. Segev, T. Wyttenbach, M. T. Bowers and R. B. Gerber, *Physical Chemistry Chemical Physics*, 2008, **10**, 3077.
80. K. B. Shelimov and M. F. Jarrold, *Journal of the American Chemical Society*, 1997, **119**, 2987.
81. C. Spitzfaden, W. Braun, G. Wider, H. Widmer and K. Wüthrich, *Journal of Biomolecular NMR*, 1994, **4**, 463.
82. G. Pflügl, J. Kallen, T. Schirmer, J. N. Jansonius, M. G. M. Zurini and M. D. Walkinshaw, *Nature*, 1993, **361**, 91.
83. M. N. Limbach, A. Antevska, D. S. Oluwatoba, A. L. H. Gray, X. B. Carroll, C. M. Hoffmann, X. Wang, M. W. Voehler, C. A. Steren and T. D. Do, *Journal of the American Chemical Society*, 2022.
84. A. L. H. Gray, A. Antevska, D. S. Oluwatoba, G. E. Schonfeld, K. L. Lazar Cantrell and T. D. Do, *Analytical Chemistry*, 2020, **92**, 11802-11808.
85. M. M. Barnhart and M. R. Chapman, *Annual Review of Microbiology*, 2006, **60**, 131.

86. L. Chai, D. Romero, C. Kayatekin, B. Akabayov, H. Vlamakis, R. Losick and R. Kolter, *Journal of Biological Chemistry*, 2013, **288**, 17559.
87. M. S. Dueholm, S. V. Petersen, M. Sonderkaer, P. Larsen, G. Christiansen, K. L. Hein, J. J. Enghild, J. L. Nielsen, K. L. Nielsen, P. H. Nielsen and D. E. Otzen, *Molecular Microbiology*, 2010, **77**, 1009.
88. D. M. Fowler, A. V. Koulov, W. E. Balch and J. W. Kelly, *Trends in Biochemical Sciences*, 2007, **32**, 217.
89. D. Romero and R. Kolter, *International Microbiology*, 2014, **17**, 65.
90. N. Van Gerven, R. D. Klein, S. J. Hultgren and H. Remaut, *Trends in Microbiology*, 2015, **23**, 693.
91. J. Y. Kim, S. Sahu, Y. H. Yau, X. Wang, S. G. Shochat, P. H. Nielsen, M. S. Dueholm, D. E. Otzen, J. Lee, M. M. Delos Santos, J. K. Yam, N. Y. Kang, S. J. Park, H. Kwon, T. Seviour, L. Yang, M. Givskov and Y. T. Chang, *Journal of the American Chemical Society*, 2016, **138**, 402.
92. D. Romero, C. Aguilar, R. Losick and R. Kolter, *Proceedings of the National Academy of Sciences*, 2010, **107**, 2230.
93. A. Taglialegna, I. Lasa and J. Valle, *Journal of Bacteriology*, 2016, **198**, 2579.
94. G. Zeng, B. S. Vad, M. S. Dueholm, G. Christiansen, M. Nilsson, T. Tolker-Nielsen, P. H. Nielsen, R. L. Meyer and D. E. Otzen, *Frontiers in Microbiology*, 2015, **6**, 1099.
95. E. A. Epstein and M. R. Chapman, *Cellular Microbiology*, 2008, **10**, 1413.
96. K. A. Kline, S. Falker, S. Dahlberg, S. Normark and B. Henriques-Normark, *Cell Host & Microbe*, 2009, **5**, 580.
97. K. Schwartz and B. R. Boles, *Current Opinion in Microbiology*, 2013, **16**, 93.
98. S. A. Tursi, E. Y. Lee, N. J. Medeiros, M. H. Lee, L. K. Nicastro, B. Buttarro, S. Gallucci, R. P. Wilson, G. C. L. Wong and C. Tukel, *PLOS Pathogens*, 2017, **13**, e1006315.
99. M. B. Miller and B. L. Bassler, *Annual Review of Microbiology*, 2001, **55**, 165.

100. T. Seviour, S. H. Hansen, L. Yang, Y. H. Yau, V. B. Wang, M. R. Stenvang, G. Christiansen, E. Marsili, M. Givskov, Y. Chen, D. E. Otzen, P. H. Nielsen, S. Geifman-Shochat, S. Kjelleberg and M. S. Dueholm, *Journal of Biological Chemistry*, 2015, **290**, 6457.
101. E. Tayeb-Fligelman, O. Tabachnikov, A. Moshe, O. Goldshmidt-Tran, M. R. Sawaya, N. Coquelle, J. P. Colletier and M. Landau, *Science*, 2017, **355**, 831.
102. T. D. Do, A. Chamas, X. Zheng, A. Barnes, D. Chang, T. Veldstra, H. Takhar, N. Dressler, B. Trapp, K. Miller, A. McMahon, S. C. Meredith, J. E. Shea, K. Lazar Cantrell and M. T. Bowers, *Biochemistry*, 2015, **54**, 4050.
103. B. Kim, T. D. Do, E. Y. Hayden, D. B. Teplow, M. T. Bowers and J. E. Shea, *The Journal of Physical Chemistry B*, 2016, **120**, 5874.
104. V. L. Anderson, T. F. Ramlall, C. C. Rospigliosi, W. W. Webb and D. Eliezer, *Proceedings of the National Academy of Sciences*, 2010, **107**, 18850.
105. S. Bera and E. Gazit, *Protein & Peptide Letters*, 2019, **26**, 88.
106. K. Broersen, D. van den Brink, G. Fraser, M. Goedert and B. Davletov, *Biochemistry*, 2006, **45**, 15610.
107. R. Kunjithapatham, F. Y. Oliva, U. Doshi, M. Perez, J. Avila and V. Munoz, *Biochemistry*, 2005, **44**, 149.
108. M. Pannuzzo, A. Raudino, D. Milardi, C. La Rosa and M. Karttunen, *Scientific Reports*, 2013, **3**, 2781.
109. M. Sadqi, F. Hernandez, U. Pan, M. Perez, M. D. Schaeberle, J. Avila and V. Munoz, *Biochemistry*, 2002, **41**, 7150.
110. N. J. Anthis and G. M. Clore, *Protein Science*, 2013, **22**, 851-858.
111. T. D. Do and M. T. Bowers, *Analytical Chemistry*, 2015, **87**, 4245.
112. J. R. Arndt, S. G. Kondalaji, M. M. Maurer, A. Parker, J. Legleiter and S. J. Valentine, *Biochemistry*, 2015, **54**, 4285.
113. S. R. Kirk, F. C. Liu, T. C. Cropley, H. R. Carlock and C. Bleiholder, *Journal of the American Society for Mass Spectrometry*, 2019, **30**, 1204.

114. S. L. Koeniger, S. I. Merenbloom, S. J. Valentine, M. F. Jarrold, H. R. Udseth, R. D. Smith and D. E. Clemmer, *Analytical Chemistry*, 2006, **78**, 4161.
115. T. Wyttenbach, P. R. Kemper, G. Baykut, M. A. Park and M. T. Bowers, *International Journal of Mass Spectrometry*, 2018, **434**, 108.
116. M. A. Zenaidee and W. A. Donald, *Analyst*, 2015, **140**, 1894.
117. M. M. Gadzuk-Shea and M. F. Bush, *The Journal of Physical Chemistry B*, 2018, **122**, 9947.
118. J. L. P. Benesch, F. Sobott and C. V. Robinson, *Analytical Chemistry*, 2003, **75**, 2208.
119. N. Felitsyn, E. N. Kitova and J. S. Klassen, *Analytical Chemistry*, 2001, **73**, 4647.
120. J. A. Madsen and J. S. Brodbelt, *Journal of the American Society for Mass Spectrometry*, 2010, **21**, 1144.
121. B. L. Schwartz, J. E. Bruce, G. A. Anderson, S. A. Hofstadler, A. L. Rockwood, R. D. Smith, A. Chilkoti and P. S. Stayton, *Journal of the American Society for Mass Spectrometry*, 1995, **6**, 459.
122. A. J. Dear, T. C. T. Michaels, G. Meisl, D. Klenerman, S. Wu, S. Perrett, S. Linse, C. M. Dobson and T. P. J. Knowles, *Proceedings of the National Academy of Sciences*, 2020, **117**, 12087.
123. T. C. T. Michaels, A. Saric, S. Curk, K. Bernfur, P. Arosio, G. Meisl, A. J. Dear, S. I. A. Cohen, C. M. Dobson, M. Vendruscolo, S. Linse and T. P. J. Knowles, *Nature Chemistry*, 2020, **12**, 445.
124. A. Laganowsky, E. Reading, J. T. Hopper and C. V. Robinson, *Nature Protocols*, 2013, **8**, 639.
125. C. Bleiholder, N. F. Dupuis, T. Wyttenbach and M. T. Bowers, *Nature Chemistry*, 2011, **3**, 172.
126. T. D. Do, N. E. LaPointe, R. Nelson, P. Krotee, E. Y. Hayden, B. Ulrich, S. Quan, S. C. Feinstein, D. B. Teplow, D. Eisenberg, J. E. Shea and M. T. Bowers, *Journal of the American Chemical Society*, 2016, **138**, 549.

127. A. L. H. Gray, A. Antevska, B. A. Link, B. Bogin, S. J. Burke, S. D. Dupuy, J. J. Collier, Z. A. Levine, M. D. Karlstad and T. D. Do, *Chemical Science*, 2021, **12**, 5853-5864.
128. A. Aguzzi and T. O'Connor, *Nature Reviews Drug Discovery*, 2010, **9**, 237-248.
129. D. Eisenberg and M. Jucker, *Cell*, 2012, **148**, 1188-1203.
130. C. A. Ross and M. A. Poirier, *Nature Medicine*, 2004, **10 Suppl**, S10-17.
131. C. Soto and S. Pritzkow, *Nature Neuroscience*, 2018, **21**, 1332-1340.
132. I. Benilova, E. Karran and B. De Strooper, *Nature Neuroscience*, 2012, **15**, 349-357.
133. J. Bieschke, M. Herbst, T. Wiglenda, R. P. Friedrich, A. Boeddrich, F. Schiele, D. Kleckers, J. M. Lopez del Amo, B. A. Grüning, Q. Wang, M. R. Schmidt, R. Lurz, R. Anwyl, S. Schnoegl, M. Fändrich, R. F. Frank, B. Reif, S. Günther, D. M. Walsh and E. E. Wanker, *Nature Chemical Biology*, 2011, **8**, 93-101.
134. C. G. Glabe, *Journal of Biological Chemistry*, 2008, **283**, 29639-29643.
135. R. Kaye, E. Head, J. L. Thompson, T. M. McIntire, S. C. Milton, C. W. Cotman and C. G. Glabe, *Science*, 2003, **300**, 486-489.
136. S. Lesné, M. T. Koh, L. Kotilinek, R. Kaye, C. G. Glabe, A. Yang, M. Gallagher and K. H. Ashe, *Nature*, 2006, **440**, 352-357.
137. M. Baldassarre, C. M. Baronio, L. A. Morozova-Roche and A. Barth, *Chemical Science*, 2017, **8**, 8247-8254.
138. T. D. Do, N. J. Economou, A. Chamas, S. K. Buratto, J. E. Shea and M. T. Bowers, *The Journal of Physical Chemistry B*, 2014, **118**, 11220-11230.
139. A. I. Ilitchev, M. J. Giammona, C. Olivas, S. L. Claud, K. L. Lazar Cantrell, C. Wu, S. K. Buratto and M. T. Bowers, *Journal of the American Chemical Society*, 2018, **140**, 9685-9695.
140. L. M. Yan, A. Velkova and A. Kapurniotu, *Current Pharmaceutical Design*, 2014, **20**, 1182-1191.
141. I. Kuperstein, K. Broersen, I. Benilova, J. Rozenski, W. Jonckheere, M. Debulpaep, A. Vandersteen, I. Segers-Nolten, K. Van Der Werf, V. Subramaniam, D. Braeken,

- G. Callewaert, C. Bartic, R. D'Hooge, I. C. Martins, F. Rousseau, J. Schymkowitz and B. De Strooper, *The EMBO Journal*, 2010, **29**, 3408-3420.
142. A. Vandersteen, M. F. Masman, G. De Baets, W. Jonckheere, K. van der Werf, S. J. Marrink, J. Rozenski, I. Benilova, B. De Strooper, V. Subramaniam, J. Schymkowitz, F. Rousseau and K. Broersen, *Journal of Biological Chemistry*, 2012, **287**, 36732-36743.
  143. D. L. Hay, S. Chen, T. A. Lutz, D. G. Parkes and J. D. Roth, *Pharmacological Reviews*, 2015, **67**, 564-600.
  144. T. Sanke, G. I. Bell, C. Sample, A. H. Rubenstein and D. F. Steiner, *Journal of Biological Chemistry*, 1988, **263**, 17243-17246.
  145. O. Schmitz, B. Brock and J. Rungby, *Diabetes*, 2004, **53 Suppl 3**, S233-238.
  146. Y. Kiriya and H. Nochi, *Cells*, 2018, **7**.
  147. J. F. Paulsson, A. Andersson, P. Westermark and G. T. Westermark, *Diabetologia*, 2006, **49**, 1237-1246.
  148. D. Raleigh, X. Zhang, B. Hastoy and A. Clark, *Journal of Molecular Endocrinology*, 2017, **59**, R121-r140.
  149. R. A. Ritzel, J. J. Meier, C. Y. Lin, J. D. Veldhuis and P. C. Butler, *Diabetes*, 2007, **56**, 65-71.
  150. S. Nagamatsu, M. Nishi and D. F. Steiner, *Diabetes Research and Clinical Practice*, 1992, **15**, 49-55.
  151. S. B. Chargé, E. J. de Koning and A. Clark, *Biochemistry*, 1995, **34**, 14588-14593.
  152. A. Pithadia, J. R. Brender, C. A. Fierke and A. Ramamoorthy, *Journal of Diabetes Research*, 2016, **2016**, 2046327.
  153. J. R. Brender, K. Hartman, R. P. Nanga, N. Popovich, R. de la Salud Bea, S. Vivekanandan, E. N. Marsh and A. Ramamoorthy, *Journal of the American Chemical Society*, 2010, **132**, 8973-8983.

154. X. Ge, A. Kakinen, E. N. Gurzov, W. Yang, L. Pang, E. H. Pilkington, P. Govindan-Nedumpully, P. Chen, F. Separovic, T. P. Davis, P. C. Ke and F. Ding, *Chemical Communications*, 2017, **53**, 9394-9397.
155. P. Nedumpully-Govindan and F. Ding, *Scientific Reports*, 2015, **5**, 8240.
156. S. Salamekh, J. R. Brender, S. J. Hyung, R. P. Nanga, S. Vivekanandan, B. T. Ruotolo and A. Ramamoorthy, *Journal of Molecular Biology*, 2011, **410**, 294-306.
157. J. R. Brender, E. L. Lee, K. Hartman, P. T. Wong, A. Ramamoorthy, D. G. Steel and A. Gafni, *Biophysical Journal*, 2011, **100**, 685-692.
158. A. C. Susa, C. Wu, S. L. Bernstein, N. F. Dupuis, H. Wang, D. P. Raleigh, J. E. Shea and M. T. Bowers, *Journal of the American Chemical Society*, 2014, **136**, 12912-12919.
159. F. A. Russell, R. King, S. J. Smillie, X. Kodji and S. D. Brain, *Physiological Reviews*, 2014, **94**, 1099-1142.
160. S. G. Amara, V. Jonas, M. G. Rosenfeld, E. S. Ong and R. M. Evans, *Nature*, 1982, **298**, 240-244.
161. M. G. Rosenfeld, J. J. Mermod, S. G. Amara, L. W. Swanson, P. E. Sawchenko, J. Rivier, W. W. Vale and R. M. Evans, *Nature*, 1983, **304**, 129-135.
162. P. K. Mulderry, M. A. Ghatei, R. A. Spokes, P. M. Jones, A. M. Pierson, Q. A. Hamid, S. Kanse, S. G. Amara, J. M. Burrin, S. Legon and et al., *Neuroscience*, 1988, **25**, 195-205.
163. B. Zhong, S. Ma and D. H. Wang, *In Vivo*, 2019, **33**, 1431-1437.
164. S. D. Brain, T. J. Williams, J. R. Tippins, H. R. Morris and I. MacIntyre, *Nature*, 1985, **313**, 54-56.
165. Z. Kee, X. Kodji and S. D. Brain, *Frontiers in Physiology*, 2018, **9**, 1249.
166. U. Arulmani, A. Maassenvandenbrink, C. M. Villalón and P. R. Saxena, *European Journal of Pharmacology*, 2004, **500**, 315-330.
167. G. Fagherazzi, D. El Fatouhi, A. Fournier, G. Gusto, F. R. Mancini, B. Balkau, M. C. Boutron-Ruault, T. Kurth and F. Bonnet, *JAMA Neurol*, 2019, **76**, 257-263.

168. P. L. Durham, *Headache*, 2006, **46 Suppl 1**, S3-8.
169. S. Iyengar, K. W. Johnson, M. H. Ossipov and S. K. Aurora, *Headache*, 2019, **59**, 659-681.
170. B. Ahrén and M. Pettersson, *International Journal of Pancreatology*, 1990, **6**, 1-15.
171. B. Leighton and G. J. Cooper, *Nature*, 1988, **335**, 632-635.
172. M. Pettersson and B. Ahrén, *Regulatory Peptides*, 1988, **23**, 37-50.
173. H. Tanaka, R. Kashiwagi and T. Koizumi, *Experimental and Clinical Endocrinology and Diabetes*, 2013, **121**, 280-285.
174. A. Martínez, S. Kapas, M. J. Miller, Y. Ward and F. Cuttitta, *Endocrinology*, 2000, **141**, 406-411.
175. K. Hermansen and B. Ahrén, *Regulatory Peptides*, 1990, **27**, 149-157.
176. T. Morishita, A. Yamaguchi, T. Yamatani, A. Nakamura, N. Arima, Y. Yamashita, H. Nakata, T. Fujita and T. Chiba, *Diabetes Research and Clinical Practice*, 1992, **15**, 63-69.
177. T. N. Rasmussen, M. Bersani, P. Schmidt, L. Thim, H. Kofod, P. N. Jørgensen, S. S. Poulsen and J. J. Holst, *Pancreas*, 1998, **16**, 195-204.
178. S. D. Brain, S. Wimalawansa, I. MacIntyre and T. J. Williams, *The American Journal of Pathology*, 1990, **136**, 487-490.
179. M. T. Galeazza, T. D. O'Brien, K. H. Johnson and V. S. Seybold, *Peptides*, 1991, **12**, 585-591.
180. G. Ben-Nissan and M. Sharon, *Current Opinion in Chemical Biology*, 2018, **42**, 25-33.
181. F. Lanucara, S. W. Holman, C. J. Gray and C. E. Eyers, *Nature Chemistry*, 2014, **6**, 281-294.
182. L. A. Woods, S. E. Radford and A. E. Ashcroft, *Biochimica et Biophysica Acta*, 2013, **1834**, 1257-1268.

183. N. F. Dupuis, C. Wu, J. E. Shea and M. T. Bowers, *Journal of the American Chemical Society*, 2009, **131**, 18283-18292.
184. N. F. Dupuis, C. Wu, J. E. Shea and M. T. Bowers, *Journal of the American Chemical Society*, 2011, **133**, 7240-7243.
185. A. I. Ilitchev, M. J. Giammona, J. N. Schwarze, S. K. Buratto and M. T. Bowers, *The Journal of Physical Chemistry B*, 2018, **122**, 9852-9859.
186. L. M. Young, P. Cao, D. P. Raleigh, A. E. Ashcroft and S. E. Radford, *Journal of the American Chemical Society*, 2014, **136**, 660-670.
187. L. M. Young, R. A. Mahood, J. C. Saunders, L. H. Tu, D. P. Raleigh, S. E. Radford and A. E. Ashcroft, *Analyst*, 2015, **140**, 6990-6999.
188. D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. Berendsen, *Journal of Computational Chemistry*, 2005, **26**, 1701-1718.
189. K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror and D. E. Shaw, *Proteins*, 2010, **78**, 1950-1958.
190. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *The Journal of Chemical Physics*, 1983, **79**, 926-935.
191. Z. A. Levine, K. Teranishi, A. K. Okada, R. Langen and J. E. Shea, *Journal of the American Chemical Society*, 2019, **141**, 14168-14179.
192. M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, *Journal of Cheminformatics*, 2012, **4**, 17.
193. Y. Sugita and Y. Okamoto, *Chemical Physics Letters*, 1999, **314**, 141-151.
194. W. G. Hoover, *Physical Review A*, 1985, **31**, 1695-1697.
195. M. F. Mesleh, J. M. Hunter, A. A. Shvartsburg, G. C. Schatz and M. F. Jarrold, *The Journal of Physical Chemistry*, 1996, **100**, 16082-16086.
196. A. L. Breeze, T. S. Harvey, R. Bazzo and I. D. Campbell, *Biochemistry*, 1991, **30**, 575-582.
197. B. Lynch and E. T. Kaiser, *Biochemistry*, 1988, **27**, 7600-7607.

198. M. C. Manning, *Biochemical and Biophysical Research Communications*, 1989, **160**, 388-392.
199. H. A. Watkins, D. L. Rathbone, J. Barwell, D. L. Hay and D. R. Poyner, *British Journal of Pharmacology*, 2013, **170**, 1308-1322.
200. W. Li, L. N. Kinch, P. A. Karplus and N. V. Grishin, *Protein Science*, 2015, **24**, 1075-1086.
201. B. Formby, F. Schmid-Formby and G. M. Grodsky, *Diabetes*, 1984, **33**, 229-234.
202. M. C. Foster, R. D. Leapman, M. X. Li and I. Atwater, *Biophysical Journal*, 1993, **64**, 525-532.
203. S. J. Burke, K. Stadler, D. Lu, E. Gleason, A. Han, D. R. Donohoe, R. C. Rogers, G. E. Hermann, M. D. Karlstad and J. J. Collier, *The American Journal of Physiology - Endocrinology and Metabolism*, 2015, **309**, E715-726.
204. J. J. Collier, S. J. Burke, M. E. Eisenhauer, D. Lu, R. C. Sapp, C. J. Frydman and S. R. Campagna, *PLoS One*, 2011, **6**, e22485.
205. N. Shigihara, A. Fukunaka, A. Hara, K. Komiya, A. Honda, T. Uchida, H. Abe, Y. Toyofuku, M. Tamaki, T. Ogihara, T. Miyatsuka, H. J. Hiddinga, S. Sakagashira, M. Koike, Y. Uchiyama, T. Yoshimori, N. L. Eberhardt, Y. Fujitani and H. Watada, *Journal of Clinical Investigation*, 2014, **124**, 3634-3644.
206. R. A. Ritzel and P. C. Butler, *Diabetes*, 2003, **52**, 1701-1708.
207. C. Y. Westwell-Roper, J. A. Ehses and C. B. Verchere, *Diabetes*, 2014, **63**, 1698-1711.
208. R. Salbo, M. F. Bush, H. Naver, I. Campuzano, C. V. Robinson, I. Pettersson, T. J. Jørgensen and K. F. Haselmann, *Rapid Communications in Mass Spectrometry*, 2012, **26**, 1181-1193.
209. A. Tkatchenko, M. Rossi, V. Blum, J. Ireta and M. Scheffler, *Physical Review Letters*, 2011, **106**, 118102.
210. F. Tofoleanu, Y. Yuan, F. C. t. Pickard, B. Tywoniuk, B. R. Brooks and N. V. Buchete, *The Journal of Physical Chemistry B*, 2018, **122**, 5657-5665.

211. E. T. Jansson, T. J. Comi, S. S. Rubakhin and J. V. Sweedler, *ACS Chemical Biology*, 2016, **11**, 2588-2595.
212. M. Ramström, C. Hagman, Y. O. Tsybin, K. E. Markides, P. Håkansson, A. Salehi, I. Lundquist, R. Håkanson and J. Bergquist, *European Journal of Biochemistry*, 2003, **270**, 3146-3152.
213. H. Roderigo-Milne, A. C. Hauge-Evans, S. J. Persaud and P. M. Jones, *Biochemical and Biophysical Research Communications*, 2002, **296**, 589-595.
214. N. Babaya, M. Nakayama, H. Moriyama, R. Gianani, T. Still, D. Miao, L. Yu, J. C. Hutton and G. S. Eisenbarth, *Diabetologia*, 2006, **49**, 1222-1228.
215. C. M. Elso, N. A. Scott, L. Mariana, E. I. Masterman, A. P. R. Sutherland, H. E. Thomas and S. I. Mannering, *PLoS One*, 2019, **14**, e0225021.
216. T. Blundell, G. Dodson, D. Hodgkin and D. Mercola, *Advances in Protein Chemistry*, 1972, **26**, 279-286.
217. L. I. Berge, T. Riise, O. B. Fasmer, O. Hundal, K. J. Oedegaard, K. Midthjell and A. Lund, *Epidemiology*, 2013, **24**, 129-134.
218. N. M. Templeman, S. Flibotte, J. H. L. Chik, S. Sinha, G. E. Lim, L. J. Foster, C. Nislow and J. D. Johnson, *Cell Reports*, 2017, **20**, 451-463.
219. H. Deng, G. G. Li, H. Nie, Y. Y. Feng, G. Y. Guo, W. L. Guo and Z. P. Tang, *BMC Neurology*, 2020, **20**, 57.
220. R. Manoukian, H. Sun, S. Miller, D. Shi, B. Chan and C. Xu, *The Journal of Headache and Pain*, 2019, **20**, 44.
221. M. Goedert, *Nature Reviews Neuroscience*, 2001, **2**, 492-501.
222. J. C. Rochet and P. T. Lansbury, Jr., *Current Opinion in Structural Biology*, 2000, **10**, 60-68.
223. D. J. Selkoe, *Physiological Reviews*, 2001, **81**, 741-766.
224. H. Y. Zoghbi and H. T. Orr, *Annual Review of Neuroscience*, 2000, **23**, 217-247.
225. I. W. Hamley, *Chemical Reviews*, 2012, **112**, 5147-5192.

226. L. I. Binder, A. L. Guillozet-Bongaarts, F. Garcia-Sierra and R. W. Berry, *Biochimica et Biophysica Acta*, 2005, **1739**, 216-223.
227. D. M. Walsh and D. B. Teplow, *Progress in Molecular Biology and Translational Science*, 2012, **107**, 101-124.
228. I. Benilova, E. Karran and B. De Strooper, *Nature Neuroscience*, 2012, **15**, 349-357.
229. R. Kayed and C. A. Lasagna-Reeves, *Journal of Alzheimer's Disease*, 2013, **33 Suppl 1**, S67-78.
230. C. Haass and D. J. Selkoe, *Nature Reviews Molecular Cell Biology*, 2007, **8**, 101-112.
231. S. Baglioni, F. Casamenti, M. Bucciantini, L. M. Luheshi, N. Taddei, F. Chiti, C. M. Dobson and M. Stefani, *The Journal of Neuroscience*, 2006, **26**, 8160.
232. J. A. Hardy and G. A. Higgins, *Science*, 1992, **256**, 184-185.
233. H. Braak and E. Braak, *Acta Neuropathologica*, 1991, **82**, 239-259.
234. J. Hardy, K. Duff, K. G. Hardy, J. Perez-Tur and M. Hutton, *Nature Neuroscience*, 1998, **1**, 355-358.
235. T. L. Spires-Jones, W. H. Stoothoff, A. de Calignon, P. B. Jones and B. T. Hyman, *Trends in Neurosciences*, 2009, **32**, 150-159.
236. M. Hutton, *Annals of the New York Academy of Sciences*, 2000, **920**, 63-73.
237. J. Götz, F. Chen, J. van Dorpe and R. M. Nitsch, *Science*, 2001, **293**, 1491-1495.
238. D. E. Hurtado, L. Molina-Porcel, M. Iba, A. K. Aboagye, S. M. Paul, J. Q. Trojanowski and V. M. Y. Lee, *The American journal of pathology*, 2010, **177**, 1977-1988.
239. J. Lewis, D. W. Dickson, W. L. Lin, L. Chisholm, A. Corral, G. Jones, S. H. Yen, N. Sahara, L. Skipper, D. Yager, C. Eckman, J. Hardy, M. Hutton and E. McGowan, *Science*, 2001, **293**, 1487-1491.
240. G. S. Bloom, *JAMA Neurology*, 2014, **71**, 505-508.

241. H. Braak and K. Del Tredici, *Advances in Anatomy, Embryology and Cell Biology*, 2015, **215**, 1-162.
242. J.-P. Guo, T. Arai, J. Miklossy and L. McGeer Patrick, *Proceedings of the National Academy of Sciences*, 2006, **103**, 1953-1958.
243. M. R. D'Andrea, R. G. Nagele, H. Y. Wang, P. A. Peterson and D. H. Lee, *Histopathology*, 2001, **38**, 120-134.
244. C. M. Karch, A. T. Jeng and A. M. Goate, *Journal of Biological Chemistry*, 2012, **287**, 42751-42762.
245. S. Magnoni, T. J. Esparza, V. Conte, M. Carbonara, G. Carrabba, D. M. Holtzman, G. J. Zipfel, N. Stocchetti and D. L. Brody, *Brain*, 2012, **135**, 1268-1280.
246. X. Chai, J. L. Dage and M. Citron, *Neurobiology of Disease*, 2012, **48**, 356-366.
247. T. D. Do, N. J. Economou, A. Chamas, S. K. Buratto, J.-E. Shea and M. T. Bowers, *The Journal of Physical Chemistry B*, 2014, **118**, 11220-11230.
248. M. von Bergen, P. Friedhoff, J. Biernat, J. Heberle, E. M. Mandelkow and E. Mandelkow, *Proceedings of the National Academy of Sciences*, 2000, **97**, 5129-5134.
249. P. Ganguly, T. D. Do, L. Larini, N. E. LaPointe, A. J. Sercel, M. F. Shade, S. C. Feinstein, M. T. Bowers and J. E. Shea, *The Journal of Physical Chemistry B*, 2015, **119**, 4582-4593.
250. N. Sahara, S. Maeda, M. Murayama, T. Suzuki, N. Dohmae, S. H. Yen and A. Takashima, *European Journal of Neuroscience*, 2007, **25**, 3020-3029.
251. K. Bhattacharya, K. B. Rank, D. B. Evans and S. K. Sharma, *Biochemical and Biophysical Research Communications*, 2001, **285**, 20-26.
252. Y. Furukawa, K. Kaneko and N. Nukina, *Journal of Biological Chemistry*, 2011, **286**, 27236-27246.
253. E. Prifti, E. N. Tsakiri, E. Vourkou, G. Stamatakis, M. Samiotaki and K. Papanikolopoulou, *The Journal of Neuroscience*, 2021, **41**, 797-810.

254. A. W. P. Fitzpatrick, B. Falcon, S. He, A. G. Murzin, G. Murshudov, H. J. Garringer, R. A. Crowther, B. Ghetti, M. Goedert and S. H. W. Scheres, *Nature*, 2017, **547**, 185-190.
255. S. Barghorn and E. Mandelkow, *Biochemistry*, 2002, **41**, 14885-14896.
256. J. Kuret, C. N. Chirita, E. E. Congdon, T. Kannanayakal, G. Li, M. Necula, H. Yin and Q. Zhong, *Biochimica et Biophysica Acta*, 2005, **1739**, 167-178.
257. M. A. Meraz-Ríos, K. I. Lira-De León, V. Campos-Peña, M. A. De Anda-Hernández and R. Mena-López, *Journal of Neurochemistry*, 2010, **112**, 1353-1367.
258. E. Sugino, C. Nishiura, K. Minoura, Y. In, M. Sumida, T. Taniguchi, K. Tomoo and T. Ishida, *Biochemical and Biophysical Research Communications*, 2009, **385**, 236-240.
259. H. Wille, G. Drewes, J. Biernat, E. M. Mandelkow and E. Mandelkow, *Journal of Cell Biology*, 1992, **118**, 573-584.
260. O. Schweers, E. M. Mandelkow, J. Biernat and E. Mandelkow, *Proceedings of the National Academy of Sciences*, 1995, **92**, 8463-8467.
261. Y. Ohhashi, Y. Hagihara, G. Kozhukh, M. Hoshino, K. Hasegawa, I. Yamaguchi, H. Naiki and Y. Goto, *The Journal of Biochemistry*, 2002, **131**, 45-52.
262. X. Zhu, A. K. Raina, H. G. Lee, G. Casadesus, M. A. Smith and G. Perry, *Brain Research*, 2004, **1000**, 32-39.
263. T. Saito, T. Chiku, M. Oka, S. Wada-Kakuda, M. Nobuhara, T. Oba, K. Shinno, S. Abe, A. Asada, A. Sumioka, A. Takashima, T. Miyasaka and K. Ando, *Human Molecular Genetics*, 2021, **30**, 1955-1967.
264. D. Kim, S. Lim, M. M. Haque, N. Ryoo, H. S. Hong, H. Rhim, D.-E. Lee, Y.-T. Chang, J.-S. Lee, E. Cheong, D. J. Kim and Y. K. Kim, *Scientific Reports*, 2015, **5**, 15231.
265. Q. Zhong, E. E. Congdon, H. N. Nagaraja and J. Kuret, *Journal of Biological Chemistry*, 2012, **287**, 20711-20719.
266. M. M. Haque, D. Kim, Y. H. Yu, S. Lim, D. J. Kim, Y.-T. Chang, H.-H. Ha and Y. K. Kim, *Amyloid*, 2014, **21**, 185-190.

267. B. Falcon, W. Zhang, A. G. Murzin, G. Murshudov, H. J. Garringer, R. Vidal, R. A. Crowther, B. Ghetti, S. H. W. Scheres and M. Goedert, *Nature*, 2018, **561**, 137-140.
268. M. A. Sealey, E. Vourkou, C. M. Cowan, T. Bossing, S. Quraishie, S. Grammenoudi, E. M. C. Skoulakis and A. Mudher, *Neurobiology of Disease*, 2017, **105**, 74-83.
269. D.-P. Hong, M. Gozu, K. Hasegawa, H. Naiki and Y. Goto, *Journal of Biological Chemistry*, 2002, **277**, 21554-21560.
270. N. M. Kad, N. H. Thomson, D. P. Smith, D. A. Smith and S. E. Radford, *Journal of Molecular Biology*, 2001, **313**, 559-571.
271. H. Katou, T. Kanno, M. Hoshino, Y. Hagihara, H. Tanaka, T. Kawai, K. Hasegawa, H. Naiki and Y. Goto, *Protein Science*, 2002, **11**, 2218-2229.
272. D. P. Smith and S. E. Radford, *Protein Science*, 2001, **10**, 1775-1784.
273. Y. Chen and N. V. Dokholyan, *Journal of Molecular Biology*, 2005, **354**, 473-482.
274. M. F, D. Puzzo, R. Piacentini, A. Staniszewski, H. Zhang, M. A. Baltrons, D. D. Li Puma, I. Chatterjee, J. Li, F. Saeed, H. L. Berman, C. Ripoli, W. Gulisano, J. Gonzalez, H. Tian, J. A. Costa, P. Lopez, E. Davidowitz, W. H. Yu, V. Haroutunian, L. M. Brown, A. Palmeri, E. M. Sigurdsson, K. E. Duff, A. F. Teich, L. S. Honig, M. Sierks, J. G. Moe, L. D'Adamio, C. Grassi, N. M. Kanaan, P. E. Fraser and O. Arancio, *Scientific Reports*, 2016, **6**, 19393.
275. T. J. Cohen, D. Friedmann, A. W. Hwang, R. Marmorstein and V. M. Y. Lee, *Nature Structural & Molecular Biology*, 2013, **20**, 756.
276. M. Martinho, D. Allegro, I. Huvent, C. Chabaud, E. Etienne, H. Kovacic, B. Guigliarelli, V. Peyrot, I. Landrieu, V. Belle and P. Barbier, *Scientific Reports*, 2018, **8**, 13846.
277. S. Kazmi, A. A. Mujeeb and M. Owais, *Scientific reports*, 2018, **8**, 17283-17283.
278. D. K. Klimov and D. Thirumalai, *Structure*, 2003, **11**, 295-307.
279. A. L. H. Gray, M. R. Sawaya, D. Acharyya, J. Lou, E. M. Edington, M. D. Best, R. A. Prosser, D. S. Eisenberg and T. D. Do, *Protein Science*, 2021, **31**, 716-727.

280. C. Haass and D. J. Selkoe, *Nature Reviews Molecular Cell Biology*, 2007, **8**, 101-112.
281. R. Kaye and C. A. Lasagna-Reeves, *Journal of Alzheimer's Disease*, 2013, **33**, S67-S78.
282. Yujin E. Kim, F. Hosp, F. Frottin, H. Ge, M. Mann, M. Hayer-Hartl and F. U. Hartl, *Molecular Cell*, 2016, **63**, 951-964.
283. N. Bengoa-Vergniory, R. F. Roberts, R. Wade-Martins and J. Alegre-Abarategui, *Acta Neuropathologica*, 2017, **134**, 819-838.
284. S. Sangwan, A. Zhao, K. L. Adams, C. K. Jayson, M. R. Sawaya, E. L. Guenther, A. C. Pan, J. Ngo, D. M. Moore, A. B. Soriaga, T. D. Do, L. Goldschmidt, R. Nelson, M. T. Bowers, C. M. Koehler, D. E. Shaw, B. G. Novitsch and D. S. Eisenberg, *Proceedings of the National Academy of Sciences*, 2017, **114**, 8770.
285. J. Ghiso and B. Frangione, *Advanced Drug Delivery Reviews*, 2002, **54**, 1539-1551.
286. H. Fu, J. Hardy and K. E. Duff, *Nature Neuroscience*, 2018, **21**, 1350-1358.
287. S. Boill  e, C. Vande Velde and Don W. Cleveland, *Neuron*, 2006, **52**, 39-59.
288. L. Brichta and P. Greengard, *Frontiers in Neuroanatomy*, 2014, **8**, 152.
289. F. Roselli and P. Caroni, *Neuron*, 2015, **85**, 901-910.
290. S. Saxena and P. Caroni, *Neuron*, 2011, **71**, 35-48.
291. N.   sterlund, R. Moons, L. L. Ilag, F. Sobott and A. Gr  slund, *Journal of the American Chemical Society*, 2019, **141**, 10440-10450.
292. M. Serra-Batiste, M. Ninot-Pedrosa, M. Bayoumi, M. Gair  , G. Maglia and N. Carulla, *Proceedings of the National Academy of Sciences*, 2016, **113**, 10866.
293. A. Laganowsky, C. Liu, M. R. Sawaya, J. P. Whitelegge, J. Park, M. Zhao, A. Pensalfini, A. B. Soriaga, M. Landau, P. K. Teng, D. Cascio, C. Glabe and D. Eisenberg, *Science*, 2012, **335**, 1228.
294. E. Tayeb-Fligelman, O. Tabachnikov, A. Moshe, O. Goldshmidt-Tran, M. R. Sawaya, N. Coquelle, J.-P. Colletier and M. Landau, *Science*, 2017, **355**, 831.

295. T. D. Do, N. E. LaPointe, R. Nelson, P. Krotee, E. Y. Hayden, B. Ulrich, S. Quan, S. C. Feinstein, D. B. Teplow, D. Eisenberg, J.-E. Shea and M. T. Bowers, *Journal of the American Chemical Society*, 2016, **138**, 549-557.
296. V. Laos, T. D. Do, D. Bishop, Y. Jin, N. M. Marsh, B. Quon, M. Feters, K. L. Cantrell, S. K. Buratto and M. T. Bowers, *ACS Chemical Neuroscience*, 2019, **10**, 4112-4123.
297. R. A. Armstrong, P. L. Lantos and N. J. Cairns, *Neuroscience Letters*, 2001, **298**, 53-56.
298. B. C. Dickerson, A. Bakkour, D. H. Salat, E. Feczko, J. Pacheco, D. N. Greve, F. Grodstein, C. I. Wright, D. Blacker, H. D. Rosas, R. A. Sperling, A. Atri, J. H. Growdon, B. T. Hyman, J. C. Morris, B. Fischl and R. L. Buckner, *Cerebral Cortex*, 2009, **19**, 497-510.
299. B. N. Dugger and D. W. Dickson, *Cold Spring Harbor Perspectives in Biology*, 2017, **9**, a028035.
300. P. Giannakopoulos, P. R. Hof, J.-P. Michel, J. Guimon and C. Bouras, *Brain Research Reviews*, 1997, **25**, 217-245.
301. R. M. Buijs, C. Escobar and D. F. Swaab, in *Handbook of Clinical Neurology*, eds. R. M. Buijs and D. F. Swaab, Elsevier, 2013, vol. 117, pp. 173-191.
302. J. A. Mohawk, C. B. Green and J. S. Takahashi, *Annual Review of Neuroscience*, 2012, **35**, 445-462.
303. L. P. Morin, *Experimental Neurology*, 2013, **243**, 4-20.
304. F.-X. Fernandez, S. Kaladchibachi and D. C. Negelsbach, *Experimental Gerontology*, 2021, **147**, 111258.
305. E. G. Stopa, L. Volicer, V. Kuo-Leblanc, D. Harper, D. Lathi, B. Tate and A. Satlin, *Journal of Neuropathology & Experimental Neurology*, 1999, **58**, 29-39.
306. J. Van Erum, D. Van Dam and P. P. De Deyn, *Sleep Medicine Reviews*, 2018, **40**, 17-27.
307. M. P. Butler and R. Silver, *Journal of Biological Rhythms*, 2009, **24**, 340-352.

308. H. Dardente, J. S. Menet, E. Challet, B. B. Tournier, P. Pévet and M. Masson-Pévet, *Molecular Brain Research*, 2004, **124**, 143-151.
309. R. P. Irwin and C. N. Allen, *European Journal of Neuroscience*, 2010, **32**, 1497-1506.
310. J. W. Mitchell, N. Atkins, J. V. Sweedler and M. U. Gillette, *Frontiers in Neuroendocrinology*, 2011, **32**, 377-386.
311. K. Bottum, E. Poon, B. Haley, S. Karmarkar and S. A. Tischkau, *Experimental Biology and Medicine*, 2010, **235**, 237-246.
312. V. Garcia-Marin, L. Blazquez-Llorca, J. R. Rodriguez, S. Boluda, G. Muntane, I. Ferrer and J. Defelipe, *Frontiers in Neuroanatomy*, 2009, **3**, 28.
313. A. M. Stranahan and M. P. Mattson, *Neural Plasticity*, 2010, **2010**, 108190.
314. S. L. Bernstein, N. F. Dupuis, N. D. Lazo, T. Wyttenbach, M. M. Condrón, G. Bitan, D. B. Teplow, J.-E. Shea, B. T. Ruotolo, C. V. Robinson and M. T. Bowers, *Nature Chemistry*, 2009, **1**, 326-331.
315. C. Bleiholder, N. F. Dupuis, T. Wyttenbach and M. T. Bowers, *Nature Chemistry*, 2011, **3**, 172-177.
316. M. R. Sawaya, M. P. Hughes, J. A. Rodriguez, R. Riek and D. S. Eisenberg, *Cell*, 2021, **184**, 4857-4873.
317. D. S. Eisenberg and M. R. Sawaya, *Annual Review of Biochemistry*, 2017, **86**, 69-95.
318. M. D. Yoder, L. M. Thomas, J. M. Tremblay, R. L. Oliver, L. R. Yarbrough and G. M. Helmkamp, *Journal of Biological Chemistry*, 2001, **276**, 9246 - 9252.
319. A. C. Young, G. Scapin, A. Kromminga, S. B. Patel, J. H. Veerkamp and J. C. Sacchettini, *Structure*, 1994, **2**, 523-534.
320. C.-C. Liu, T. Kanekiyo, H. Xu and G. Bu, *Nature Reviews Neurology*, 2013, **9**, 106-118.
321. W. Junge, O. Pänke, D. A. Cherepanov, K. Gumbiowski, M. Müller and S. Engelbrecht, *FEBS Letters*, 2001, **504**, 152-160.

322. S. Sangwan, M. R. Sawaya, K. A. Murray, M. P. Hughes and D. Eisenberg, *Protein Science*, 2018, **27**, 1231-1242.
323. A. Laganowsky, E. Reading, J. T. S. Hopper and C. V. Robinson, *Nature Protocols*, 2013, **8**, 639-651.
324. A. A. Shvartsburg and M. F. Jarrold, *Chemical Physics Letters*, 1996, **261**, 86-91.
325. N. G. Hatcher, N. Atkins, S. P. Annangudi, A. J. Forbes, N. L. Kelleher, M. U. Gillette and J. V. Sweedler, *Proceedings of the National Academy of Sciences*, 2008, **105**, 12527.
326. J. E. Lee, N. Atkins, N. G. Hatcher, L. Zamdborg, M. U. Gillette, J. V. Sweedler and N. L. Kelleher, *Molecular & Cellular Proteomics*, 2010, **9**, 285-297.
327. J. E. Lee, L. Zamdborg, B. R. Southey, N. Atkins, J. W. Mitchell, M. Li, M. U. Gillette, N. L. Kelleher and J. V. Sweedler, *Journal of Proteome Research*, 2013, **12**, 585-593.
328. E. Drummond, S. Nayak, A. Faustin, G. Pires, R. A. Hickman, M. Askenazi, M. Cohen, T. Haldiman, C. Kim, X. Han, Y. Shao, J. G. Safar, B. Ueberheide and T. Wisniewski, *Acta Neuropathologica*, 2017, **133**, 933-954.
329. E. Barbato, R. Darrah and T. J. Kelley, *Journal of Circadian Rhythms*, 2021, **19**, 5.
330. J. R. Burke, J. J. Enghild, M. E. Martin, Y. S. Jou, R. M. Myers, A. D. Roses, J. M. Vance and W. J. Strittmatter, *Nature Medicine*, 1996, **2**, 347-350.
331. H. Schulze, A. Schuler, D. Stüber, H. Döbeli, H. Langen and G. Huber, *Journal of Neurochemistry*, 1993, **60**, 1915-1922.
332. D. A. Butterfield, S. S. Hardas and M. L. Lange, *Journal of Alzheimer's Disease*, 2010, **20**, 369-393.
333. J. Marksteiner, W. A. Kaufmann, P. Gurka and C. Humpel, *Journal of Molecular Neuroscience*, 2002, **18**, 53-63.
334. A. H. Boyd-Kimball D, Reed T, Sultana R, Butterfield DA, *Chemical Research in Toxicology*, 2004, **17**, 1743-1749.

- 335. R. Cukalevski, B. Boland, B. Frohm, E. Thulin, D. Walsh and S. Linse, *ACS Chemical Neuroscience*, 2012, **3**, 1008-1016.
- 336. M. Genji, Y. Yano, M. Hoshino and K. Matsuzaki, *Chemical and Pharmaceutical Bulletin*, 2017, **65**, 668-673.
- 337. N. Ait-Bouziad, A. Chiki, G. Limorenko, S. F. Xiao, D. Eliezer and H. A. Lashuel, *Journal of Biological Chemistry*, 2020, **295**, 7905-7922.
- 338. P. Ganguly, T. D. Do, L. Larini, N. E. LaPointe, A. J. Sercel, M. F. Shade, S. C. Feinstein, M. T. Bowers and J. E. Shea, *JOURNAL OF PHYSICAL CHEMISTRY B*, 2015, **119**, 4582-4593.

## Appendix

Below is an example of an output of what is obtained from theoretical modeling of a species via the trajectory method using the Mobcal program. Note that this is specifically the output for the G6W dodecamer. The theoretical CCS is obtained from the summary near the end of the output labeled "Average TM cross section".

*input file name = Nat-12mer-H.mfj*

*input file label = TITLE*

*number of coordinate sets = 1*

*number of atoms = 1428*

*coordinates in angstroms*

*correction factor for coordinates = 1.0000E+00*

*using a uniform charge distribution*

*mass of ion = 1.4952D+04*

*center of mass coordinates = 8.3090E+00,-1.6300E+00, 1.6033E+01*

*Lennard-Jones scaling parameters: eo= 1.3400E-03 ro= 3.0430E+00*

*dipole constant = 2.3642E-59*

*mobility constant = 9.2532E-21*

*temperature = 2.9800E+02*

*using RANLUX*

*RANLUX LUXURY LEVEL SET BY RLUXGO : 3 P= 223*

*RANLUX INITIALIZED BY RLUXGO FROM SEEDS 5013486 0 0*

*structural asymmetry parameter = 1.0489*

*mobility calculation by MOBIL4 (HS scattering)*

*number of Monte Carlo trajectories = 5000000*

*maximum number of reflections allowed = 30*

*average PA mobility = 3.6127E-05*

*inverse average PA mobility = 2.7680E+04*

*average PA cross section = 1.4837E+03*

*maximum number of reflections encountered = 30*

*order cross section*

|    |             |
|----|-------------|
| 1  | 1.57608E-17 |
| 2  | 1.78225E-17 |
| 3  | 1.82996E-17 |
| 4  | 1.84538E-17 |
| 5  | 1.85300E-17 |
| 6  | 1.85708E-17 |
| 7  | 1.85955E-17 |
| 8  | 1.86126E-17 |
| 9  | 1.86238E-17 |
| 10 | 1.86321E-17 |
| 11 | 1.86390E-17 |
| 12 | 1.86430E-17 |
| 13 | 1.86477E-17 |
| 14 | 1.86506E-17 |
| 15 | 1.86527E-17 |

16 1.86555E-17  
17 1.86567E-17  
18 1.86587E-17  
19 1.86605E-17  
20 1.86621E-17  
21 1.86629E-17  
22 1.86630E-17  
23 1.86639E-17  
24 1.86654E-17  
25 1.86656E-17  
26 1.86664E-17  
27 1.86665E-17  
28 1.86667E-17  
29 1.86675E-17  
30 1.86680E-17

*average EHS mobility = 2.8714E-05*

*inverse average EHS mobility = 3.4827E+04*

*average EHS cross section = 1.8668E+03*

*mobility calculation by MOBIL2 (trajectory method)*

*global trajectory parameters*

*sw1 = 5.0000E-05      sw2 = 5.0000E-03*

*dtsf1 = 5.0000E-01      dtsf2 = 1.0000E-01*

*inwr = 1              ifail = 100*

*maximum extent orientated along x axis*

*coordinates rotated by ROTATE*

*theta= 0.0000E+00 phi= 1.9648E+02 gamma= 3.1079E+02*

*along x axis emax = -2.1970E-03eV rmax = 2.7888E+01A r00 = 2.7776E+01A*

*along y axis emax = -5.3519E-03eV rmax = 1.7252E+01A r00 = 1.6610E+01A*

*along z axis emax = -3.6225E-03eV rmax = 2.2556E+01A r00 = 2.2109E+01A*

*t\* = 1.9164E+01*

*set-up gst integration - integration over velocity*

| <i>pgst</i> | <i>wgst</i> | <i>v</i> | <i>ke/kt</i> | <i>gst^5*</i> | <i>frac of</i>            |
|-------------|-------------|----------|--------------|---------------|---------------------------|
|             |             |          |              |               | <i>exp(gst^2/tst) sum</i> |

|                   |                   |                   |                   |                   |                   |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| <i>2.3085E+00</i> | <i>5.8270E-03</i> | <i>5.8683E+02</i> | <i>2.7808E-01</i> | <i>4.9645E+01</i> | <i>2.9135E-03</i> |
| <i>2.8879E+00</i> | <i>8.2406E-03</i> | <i>7.3412E+02</i> | <i>4.3519E-01</i> | <i>1.2999E+02</i> | <i>9.9473E-03</i> |
| <i>3.2683E+00</i> | <i>1.0093E-02</i> | <i>8.3081E+02</i> | <i>5.5737E-01</i> | <i>2.1356E+02</i> | <i>1.9114E-02</i> |
| <i>3.5691E+00</i> | <i>1.1654E-02</i> | <i>9.0728E+02</i> | <i>6.6469E-01</i> | <i>2.9792E+02</i> | <i>2.9987E-02</i> |
| <i>3.8253E+00</i> | <i>1.3030E-02</i> | <i>9.7242E+02</i> | <i>7.6357E-01</i> | <i>3.8171E+02</i> | <i>4.2329E-02</i> |
| <i>4.0528E+00</i> | <i>1.4273E-02</i> | <i>1.0303E+03</i> | <i>8.5710E-01</i> | <i>4.6405E+02</i> | <i>5.5981E-02</i> |
| <i>4.2602E+00</i> | <i>1.5417E-02</i> | <i>1.0830E+03</i> | <i>9.4706E-01</i> | <i>5.4432E+02</i> | <i>7.0825E-02</i> |
| <i>4.4528E+00</i> | <i>1.6481E-02</i> | <i>1.1319E+03</i> | <i>1.0346E+00</i> | <i>6.2205E+02</i> | <i>8.6774E-02</i> |

4.6340E+00 1.7481E-02 1.1780E+03 1.1205E+00 6.9686E+02 1.0376E-01  
4.8065E+00 1.8426E-02 1.2218E+03 1.2055E+00 7.6842E+02 1.2171E-01  
4.9720E+00 1.9326E-02 1.2639E+03 1.2900E+00 8.3645E+02 1.4059E-01  
5.1321E+00 2.0185E-02 1.3046E+03 1.3744E+00 9.0073E+02 1.6034E-01  
5.2878E+00 2.1009E-02 1.3442E+03 1.4590E+00 9.6100E+02 1.8094E-01  
5.4401E+00 2.1803E-02 1.3829E+03 1.5443E+00 1.0171E+03 2.0234E-01  
5.5898E+00 2.2568E-02 1.4209E+03 1.6304E+00 1.0688E+03 2.2453E-01  
5.7375E+00 2.3308E-02 1.4585E+03 1.7177E+00 1.1159E+03 2.4747E-01  
5.8839E+00 2.4025E-02 1.4957E+03 1.8065E+00 1.1582E+03 2.7113E-01  
6.0297E+00 2.4722E-02 1.5328E+03 1.8971E+00 1.1955E+03 2.9551E-01  
6.1752E+00 2.5399E-02 1.5698E+03 1.9898E+00 1.2277E+03 3.2057E-01  
6.3210E+00 2.6059E-02 1.6068E+03 2.0849E+00 1.2545E+03 3.4630E-01  
6.4677E+00 2.6703E-02 1.6441E+03 2.1828E+00 1.2758E+03 3.7268E-01  
6.6158E+00 2.7331E-02 1.6818E+03 2.2839E+00 1.2913E+03 3.9969E-01  
6.7659E+00 2.7945E-02 1.7199E+03 2.3887E+00 1.3008E+03 4.2733E-01  
6.9184E+00 2.8546E-02 1.7587E+03 2.4976E+00 1.3042E+03 4.5558E-01  
7.0742E+00 2.9135E-02 1.7983E+03 2.6114E+00 1.3010E+03 4.8442E-01  
7.2339E+00 2.9712E-02 1.8389E+03 2.7306E+00 1.2912E+03 5.1384E-01  
7.3984E+00 3.0278E-02 1.8807E+03 2.8562E+00 1.2743E+03 5.4384E-01  
7.5687E+00 3.0833E-02 1.9240E+03 2.9892E+00 1.2500E+03 5.7439E-01  
7.7461E+00 3.1379E-02 1.9691E+03 3.1310E+00 1.2180E+03 6.0550E-01  
7.9319E+00 3.1916E-02 2.0163E+03 3.2830E+00 1.1779E+03 6.3715E-01  
8.1282E+00 3.2443E-02 2.0662E+03 3.4475E+00 1.1292E+03 6.6933E-01  
8.3373E+00 3.2962E-02 2.1194E+03 3.6271E+00 1.0712E+03 7.0203E-01  
8.5626E+00 3.3473E-02 2.1767E+03 3.8258E+00 1.0035E+03 7.3525E-01  
8.8087E+00 3.3977E-02 2.2392E+03 4.0489E+00 9.2502E+02 7.6897E-01  
9.0824E+00 3.4473E-02 2.3088E+03 4.3044E+00 8.3487E+02 8.0320E-01  
9.3944E+00 3.4962E-02 2.3881E+03 4.6052E+00 7.3169E+02 8.3791E-01

9.7630E+00 3.5444E-02 2.4818E+03 4.9737E+00 6.1357E+02 8.7312E-01  
 1.0224E+01 3.5920E-02 2.5991E+03 5.4549E+00 4.7768E+02 9.0880E-01  
 1.0868E+01 3.6389E-02 2.7628E+03 6.1638E+00 3.1911E+02 9.4495E-01  
 1.2087E+01 3.6853E-02 3.0725E+03 7.6231E+00 1.2615E+02 9.8157E-01

*set up b2 integration - integration over impact parameter*

*minimum value of  $(1-\cos X) = 5.0000E-04$*

| <i>gst</i> | <i>b2max/ro2</i> | <i>b/A</i> |
|------------|------------------|------------|
| 2.3085E+00 | 1.2070E+02       | 3.3431E+01 |
| 2.8879E+00 | 1.1490E+02       | 3.2618E+01 |
| 3.2683E+00 | 1.1210E+02       | 3.2218E+01 |
| 3.5691E+00 | 1.1010E+02       | 3.1930E+01 |
| 3.8253E+00 | 1.0870E+02       | 3.1726E+01 |
| 4.0528E+00 | 1.0750E+02       | 3.1550E+01 |
| 4.2602E+00 | 1.0650E+02       | 3.1403E+01 |
| 4.4528E+00 | 1.0570E+02       | 3.1285E+01 |
| 4.6340E+00 | 1.0500E+02       | 3.1181E+01 |
| 4.8065E+00 | 1.0430E+02       | 3.1077E+01 |
| 4.9720E+00 | 1.0370E+02       | 3.0988E+01 |
| 5.1321E+00 | 1.0310E+02       | 3.0898E+01 |
| 5.2878E+00 | 1.0260E+02       | 3.0823E+01 |
| 5.4401E+00 | 1.0210E+02       | 3.0748E+01 |
| 5.5898E+00 | 1.0170E+02       | 3.0688E+01 |

|            |            |            |
|------------|------------|------------|
| 5.7375E+00 | 1.0120E+02 | 3.0612E+01 |
| 5.8839E+00 | 1.0080E+02 | 3.0551E+01 |
| 6.0297E+00 | 1.0040E+02 | 3.0491E+01 |
| 6.1752E+00 | 1.0000E+02 | 3.0430E+01 |
| 6.3210E+00 | 9.9700E+01 | 3.0384E+01 |
| 6.4677E+00 | 9.9300E+01 | 3.0323E+01 |
| 6.6158E+00 | 9.9000E+01 | 3.0277E+01 |
| 6.7659E+00 | 9.8600E+01 | 3.0216E+01 |
| 6.9184E+00 | 9.8300E+01 | 3.0170E+01 |
| 7.0742E+00 | 9.7900E+01 | 3.0109E+01 |
| 7.2339E+00 | 9.7600E+01 | 3.0063E+01 |
| 7.3984E+00 | 9.7300E+01 | 3.0016E+01 |
| 7.5687E+00 | 9.6900E+01 | 2.9955E+01 |
| 7.7461E+00 | 9.6600E+01 | 2.9908E+01 |
| 7.9319E+00 | 9.6300E+01 | 2.9862E+01 |
| 8.1282E+00 | 9.5900E+01 | 2.9800E+01 |
| 8.3373E+00 | 9.5600E+01 | 2.9753E+01 |
| 8.5626E+00 | 9.5200E+01 | 2.9691E+01 |
| 8.8087E+00 | 9.4800E+01 | 2.9628E+01 |
| 9.0824E+00 | 9.4400E+01 | 2.9566E+01 |
| 9.3944E+00 | 9.3900E+01 | 2.9487E+01 |
| 9.7630E+00 | 9.3400E+01 | 2.9409E+01 |
| 1.0224E+01 | 9.2800E+01 | 2.9314E+01 |
| 1.0868E+01 | 9.2000E+01 | 2.9187E+01 |
| 1.2087E+01 | 9.0500E+01 | 2.8949E+01 |

*number of complete cycles (itn) = 10*

*number of velocity points (inp) = 40*

*number of random points (imp) = 2500*

*total number of points = 1000000*

*trajectory lost: b = 6.5942E-10 v = 3.0725E+03*

*summary of mobility calculations*

| <i>cycle</i> | <i>cs/A^2</i> | <i>avge cs/A^2</i> | <i>Ko^-1</i> | <i>avge Ko^-1</i> |
|--------------|---------------|--------------------|--------------|-------------------|
| 1            | 1.8431E+03    | 1.8431E+03         | 3.4384E+04   | 3.4384E+04        |
| 2            | 1.8480E+03    | 1.8456E+03         | 3.4477E+04   | 3.4431E+04        |
| 3            | 1.8388E+03    | 1.8433E+03         | 3.4304E+04   | 3.4388E+04        |
| 4            | 1.8436E+03    | 1.8434E+03         | 3.4395E+04   | 3.4390E+04        |
| 5            | 1.8495E+03    | 1.8446E+03         | 3.4504E+04   | 3.4413E+04        |
| 6            | 1.8461E+03    | 1.8449E+03         | 3.4441E+04   | 3.4417E+04        |
| 7            | 1.8376E+03    | 1.8438E+03         | 3.4281E+04   | 3.4398E+04        |
| 8            | 1.8388E+03    | 1.8432E+03         | 3.4305E+04   | 3.4386E+04        |
| 9            | 1.8492E+03    | 1.8439E+03         | 3.4499E+04   | 3.4399E+04        |
| 10           | 1.8420E+03    | 1.8437E+03         | 3.4364E+04   | 3.4395E+04        |

*average values for q1st*

| <i>gst2</i> | <i>wgst</i> | <i>q1st</i> |
|-------------|-------------|-------------|
| 5.3291E+00  | 5.8270E-03  | 1.8919E+01  |
| 8.3399E+00  | 8.2406E-03  | 1.8165E+01  |
| 1.0682E+01  | 1.0093E-02  | 1.7706E+01  |
| 1.2738E+01  | 1.1654E-02  | 1.7569E+01  |

1.4633E+01 1.3030E-02 1.7389E+01  
1.6426E+01 1.4273E-02 1.7039E+01  
1.8149E+01 1.5417E-02 1.7075E+01  
1.9827E+01 1.6481E-02 1.6838E+01  
2.1474E+01 1.7481E-02 1.6841E+01  
2.3102E+01 1.8426E-02 1.6522E+01  
2.4721E+01 1.9326E-02 1.6406E+01  
2.6338E+01 2.0185E-02 1.6235E+01  
2.7961E+01 2.1009E-02 1.6454E+01  
2.9594E+01 2.1803E-02 1.6241E+01  
3.1245E+01 2.2568E-02 1.6126E+01  
3.2919E+01 2.3308E-02 1.6034E+01  
3.4621E+01 2.4025E-02 1.6212E+01  
3.6357E+01 2.4722E-02 1.6039E+01  
3.8133E+01 2.5399E-02 1.5914E+01  
3.9956E+01 2.6059E-02 1.6012E+01  
4.1832E+01 2.6703E-02 1.5879E+01  
4.3769E+01 2.7331E-02 1.5893E+01  
4.5777E+01 2.7945E-02 1.5723E+01  
4.7865E+01 2.8546E-02 1.5771E+01  
5.0044E+01 2.9135E-02 1.5657E+01  
5.2329E+01 2.9712E-02 1.5624E+01  
5.4737E+01 3.0278E-02 1.5527E+01  
5.7286E+01 3.0833E-02 1.5443E+01  
6.0002E+01 3.1379E-02 1.5493E+01  
6.2915E+01 3.1916E-02 1.5444E+01  
6.6067E+01 3.2443E-02 1.5658E+01  
6.9511E+01 3.2962E-02 1.5481E+01

7.3318E+01 3.3473E-02 1.5403E+01  
7.7593E+01 3.3977E-02 1.5337E+01  
8.2490E+01 3.4473E-02 1.5181E+01  
8.8254E+01 3.4962E-02 1.5391E+01  
9.5316E+01 3.5444E-02 1.5473E+01  
1.0454E+02 3.5920E-02 1.5159E+01  
1.1812E+02 3.6389E-02 1.4855E+01  
1.4609E+02 3.6853E-02 1.5145E+01

*mean OMEGA\*(1,1) = 6.3377E+01*

*standard deviation = 1.4497E-01*

*standard error of mean = 4.5845E-02*

*f value for second order correction= 1.0000E+00*

*(integrations for second order correction are not  
accurate, check if correction becomes significant)*

*omega\*12 = 6.1839E+01 omega\*13 = 5.9927E+01 omega\*22 = 5.0053E+01*

*u2 = 1.8670E+03 w = 2.6770E-04 delta = 1.7437E-09*

*average (second order) TM mobility = 2.9074E-05*

*inverse average (second order) TM mobility = 3.4395E+04*

*average TM cross section = 1.8437E+03*

## SUMMARY

*program version = junkn.f*

*input file name = Nat-12mer-H.mfj*

*input file label = TITLE*

*temperature = 2.9800E+02*

*using RANLUX with seed integer = 5013486*

*structural asymmetry parameter = 1.0489E+00*

*mobility calculation by MOBIL4 (HS scattering)*

*number of Monte Carlo trajectories = 5000000*

*maximum number of reflections encountered = 30*

*inverse average PA mobility = 2.7680E+04*

*average PA cross section = 1.4837E+03*

*inverse average EHS mobility = 3.4827E+04*

*average EHS cross section = 1.8668E+03*

*mobility calculation by MOBIL2 (trajectory method)*

*trajectory parameters*

*sw1 = 5.0000E-05    sw2 = 5.0000E-03*

*dtsf1 = 5.0000E-01    dtsf2 = 1.0000E-01*

*number of complete cycles (itn) = 10*

*number of velocity points (inp) = 40*

*number of random points (imp) = 2500*

*total number of points = 1000000*

*inverse average (second order) TM mobility = 3.4395E+04*

*average TM cross section = 1.8437E+03*

*standard deviation (percent) = 2.2875E-01*

*number of failed trajectories = 0*

## **Vita**

Amber LeAnn Hope Gray was born in Manchester, KY on December 6, 1994. She grew up in Clay County Kentucky, where she graduated from Clay County High School in May 2013. She attended The University of the Cumberlands as an undergraduate student and obtained her Bachelor of Science in both biology and chemistry in May 2017. Afterwards, she joined the chemistry department as a graduate student at the University of Tennessee, Knoxville in Fall 2017. She started her research under Thanh D. Do in Spring 2019. Amber will graduate in August 2022 with her Ph.D. in Analytical Chemistry. Amber plans on pursuing a career within the pharmaceutical industry.