

## ABSTRACT

LATSHAW II, DAVID CHRISTOPHER. Effects of Crowders, Inhibitors, and Interfaces on Peptide Aggregation Using Coarse-Grained Simulations. (Under the direction of Dr. Carol K. Hall).

Protein aggregation is a phenomenon in which proteins mis-fold and interact to form higher order structures that cause a variety of biological problems including neurodegenerative diseases like Alzheimer's and unwanted immune responses from aggregated therapeutic proteins. The specific aims of this work are to understand how forces like macromolecular crowding, aqueous-organic interfaces, and aggregation-inhibiting small molecules alter protein aggregation. To examine the effects of these forces on protein aggregation we used discontinuous molecular dynamics (DMD) simulations and an intermediate resolution protein model, PRIME20.

In our first study of macromolecular crowding, systems containing 192 A $\beta$ (16-22) peptides and hard sphere crowders of diameters 5Å, 20Å, and 40Å, at crowder volume fractions of 0.00, 0.10, and 0.20 were simulated. Results show that the addition of crowders to a system of peptides increases the rate of oligomer formation, shifting from a slow ordered formation of oligomers in the absence of crowders, similar to nucleated polymerization, to a fast collapse of peptides and subsequent rearrangement characteristic of nucleated conformational conversion with a high maximum in the number of peptides in oligomers as total crowder surface area increases. The rate of conversion from oligomers to fibrils also increases with increasing total crowder surface area, giving rise to an increased rate of fibril growth. In all cases, larger volume fractions and smaller crowders provide the greatest aggregation enhancement effects. We also show that the size of the crowders influences the formation of specific oligomer sizes.

To further examine the effect of crowding on protein aggregation we studied the effect of adding hydrophobic crowders to systems containing 48 A $\beta$ (16-22) peptides and crowders of diameters 5Å, 20Å, and 40Å, represented by hard spheres or square well/shoulder spheres, at a crowder volume fraction of 0.10. Results show that low levels of crowder hydrophobicity are capable of increasing fibrillation lag time and high levels of crowder hydrophobicity can prevent the formation of fibrils. Furthermore, at high levels of crowder hydrophobicity, the systems with 5Å crowders form only disordered oligomers while those with 40Å crowders form only  $\beta$ -sheets.

Next, we examined the effect of an aqueous/organic interface on the folding and aggregation of amphipathic peptides that represent therapeutic proteins. The systems contained either a single (KLLK)<sub>4</sub> peptide to study folding or 48 (KLLK)<sub>4</sub> peptides to study aggregation. Results show that the weak interface is capable of folding the randomly coiled peptide into an amphipathic  $\alpha$ -helix, while the strong interface disrupts intra-peptide hydrogen bonding and unfolds the peptides into random coil or  $\beta$ -hairpin conformations. Regardless of the number of proteins in the simulation, peptides bound irreversibly to the strong interface because of the side chain preferences for their respective phases. Peptides interacting with the weak interface were bound reversibly. In the multi-peptide case, the aggregation was highest when the peptides were initially in an  $\alpha$ -helical state in the presence of a weak interface and was lowest when the peptides were initially in a random coil state in the presence of a strong interface.

Finally, we present preliminary results on the aggregation of A $\beta$ (17-36) in the presence of naturally occurring phenolic aggregation inhibitors. Simulations were performed on systems containing 8 A $\beta$ (17-36) molecules and either 128 or 256 vanillin, resveratrol,

curcumin, or epigallocatechin-3-gallate (EGCG) molecules. During the initial stages of the simulations, we find that resveratrol binds most frequently with A $\beta$ , followed by curcumin, and finally vanillin. We also observe that curcumin and resveratrol are localized to the interior of aggregates, possibly preventing further aggregation into amyloid structures, while vanillin binds the exterior. Limited simulations on EGCG show that it is capable of binding monomers unlike the other molecules.

© Copyright 2014 by David Christopher Latshaw II

All Rights Reserved

Effects of Crowders, Inhibitors, and Interfaces on Peptide Aggregation Using Coarse-Grained Simulations

by  
David Christopher Latshaw II

A dissertation submitted to the Graduate Faculty of  
North Carolina State University  
in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

Chemical Engineering

Raleigh, North Carolina

2015

APPROVED BY:

---

Dr. Carol K Hall  
Committee Chair

---

Dr. Clay Clark

---

Dr. Gregory Reeves

---

Dr. Balaji Rao

---

Dr. Peter Fedkiw

## DEDICATION

*This dissertation is dedicated to my dad, David Latshaw, for teaching me perseverance and how to step back and put things in perspective and my mom, Kathy Latshaw, for her love, support, and encouragement without which I would not be where I am today. I would also like to dedicate this dissertation to my brother Jarrod Latshaw who is always there for me when I need him. Finally, I dedicate this work to my loving wife Alina Latshaw, who is truly my other half and has kept me going through graduate school. I cannot wait to spend the rest of our lives together.*

## **BIOGRAPHY**

David Latshaw was born in Staunton, Virginia to his parents David and Kathy Latshaw. After four years in Virginia, David's family moved to Audubon, Pennsylvania where he would remain until his sophomore year at Methacton High School when his family moved once more to Fountain Hills, Arizona. David finished his remaining years at Fountain Hills High School and, after graduating in 2005, attended Arizona State University to major in chemical engineering. After completing his Bachelor of Science in 2009 he began his graduate studies in the Department of Chemical and Biomolecular Engineering at North Carolina State University in the fall, advised by Dr. Carol Hall. After completion of his graduate studies in December 2014, David will join Janssen Pharmaceuticals in Malvern, Pennsylvania.

## ACKNOWLEDGMENTS

With much gratitude, I would like to thank Dr. Carol Hall for guiding me through my graduate studies. Most specifically, I would like to thank her for helping me learn how to develop my own research ideas without being caught up in unnecessary details, and also for helping me to refine my writing skills. I am also grateful to my committee members, Dr. Clay Clark, Dr. Peter Fedkiw, Dr. Balaji Rao and Dr. Gregory Reeves for their guidance and insight throughout my time at NC State. I would also like to thank Dr. Keith Gubbins and Dr. Erik Santiso for useful discussions on simulations techniques and Dr. Mookyung Cheon for his work on the PRIME20 force field and his discussions on our simulation work. Finally, I would like to thank Dr. Ted Randolph for inspiring our work on the behavior of amphipathic peptides at interfaces and for useful discussions on the topic.

I would like to thank Professor Robert M. Kelly and the National Institutes of Health Biotechnology Traineeship for the education and financial support that made part of this research possible. The knowledge and experience gained through this program has proved quite valuable. I would also like to thank the U.S. Department of Education Graduate Assistance in Areas of National Need (GAANN) Computational Science Fellowship for providing financial support. Lastly, I would like to thank the National Institutes of Health Research Grant for funding my work on protein aggregation.

I would also like to thank past and present members of the Hall research group. I want to thank Dr. Victoria Wagoner and Dr. Erin Phelps for introducing me to our simulation methods and getting me integrated into the group. I would also like to thank Dr. Emily Curtis for all the discussions on programming and talks about life, Dr. Abhishek Singh for the walks

and coffee, Steve Benner for discussing all things fitness, and David Rutkowski (The Swoletergeist®) for all the talks about video games. It helps to break the day up and put a smile on my face. To the newer members of our group David Rutkowski, Steve Benner, Binwu Zhao, Yiming Wang, and Kye Won Wang: Keep working hard and trust your ideas, you are here for a reason! To our post docs Dr. Emily Curtis, Dr. Xingqing Xiao, and Dr. Qing Shao: You already know what you are doing so keep up the good work!

Finally, I would like to thank my family and other friends. Although my family may not have understood exactly what I was researching in graduate school they always showed genuine interest and that was all I needed. Phil Schoch, Will Sweet, and José Eby thank you for allowing me to be your fourth roommate our first year. Dan Morales, Dennis McOwen, and Jeff Zurawski I had a blast living with you all too! Heath Johnson, Steph Lam, Christina McClure, Lettie Allen, and the Joshs Allen and McClure, thank you for giving a first year a chance even though I put my foot in my mouth a few times when we first met. You have all become great friends of mine and I hope it will always be that way. Last, but certainly not least, I would like to thank Alina, my wife. You have always supported me in every aspect of my life and I can't thank you enough. I could not be as happy as I am today without you.

## TABLE OF CONTENTS

<b>LIST OF TABLES .....</b>	<b>x</b>
<b>LIST OF FIGURES .....</b>	<b>xi</b>
<b>CHAPTER 1 .....</b>	<b>1</b>
<b>Motivation and Overview.....</b>	<b>1</b>
1.1 Motivation.....	2
1.2 Overview.....	7
1.2.1 Effects of Macromolecular Crowding on Amyloid Beta (16-22) Aggregation.....	7
1.2.2 Effects of Hydrophobic Macromolecular Crowders on Amyloid Beta (16-22) Aggregation.....	8
1.2.3 Folding and Aggregation of Amphipathic Peptides at an Aqueous- Organic Interface .....	9
1.2.4 Aggregation of Amyloid Beta in the Presence of Naturally Occurring Phenolic Inhibitors .....	10
1.2.5 Future Work.....	11
1.3 Publications.....	11
1.4 References.....	12
<b>CHAPTER 2 .....</b>	<b>17</b>
<b>Effects of Macromolecular Crowding on Amyloid Beta (16-22) Aggregation using     Coarse-Grained Simulations.....</b>	<b>17</b>
2.1 Introduction.....	20
2.2 Methods.....	28
2.2.1 Discontinuous Molecular Dynamics .....	28
2.2.2 PRIME20 force field .....	29
2.2.3 Crowder Model.....	30
2.2.4 Simulation Procedure .....	31
2.2.5 Oligomerization curve fit and parameters .....	32
2.2.6 Small oligomer free energy analysis .....	33

2.2.7 Fibrillization curve fit and parameters .....	35
2.3 Results.....	37
2.3.1 Aggregation in the presence of crowders .....	37
2.3.2 Effects of crowding on oligomerization .....	40
2.3.3 Effects of crowding on small oligomer formation .....	43
2.3.4 Effects of crowding on fibrillization .....	47
2.4 Discussion and Conclusions .....	51
2.5 Acknowledgements.....	55
2.6 References.....	56
<b>CHAPTER 3.....</b>	<b>74</b>
<b>Effects of Hydrophobic Macromolecular Crowders on Amyloid Beta (16-22)</b>	
<b>Aggregation using Coarse-Grained Simulations.....</b>	<b>74</b>
3.1 Introduction.....	77
3.2 Methods.....	80
3.3 Results and Discussion .....	82
3.3.1 Aggregation in the presence of hydrophobic crowders .....	82
3.3.2 Effect of hydrophobic crowders on disordered oligomer formation .....	84
3.3.3 Effect of hydrophobic crowders on $\beta$ -sheet formation.....	85
3.3.4 Effect of hydrophobic crowders on fibril formation .....	87
3.3.5 Global phase diagram peptides with hydrophobic crowders.....	89
3.4 Conclusions.....	90
3.5 Acknowledgements.....	93
3.6 References.....	94
<b>CHAPTER 4.....</b>	<b>109</b>
<b>Aggregation of Amphipathic Peptides at an Aqueous/Organic Interface using</b>	
<b>Coarse-Grained Simulations.....</b>	<b>109</b>
4.1 Introduction.....	112
4.2 Methods.....	117
4.2.1 Discontinuous Molecular Dynamics and PRIME20 Force Field .....	117

4.2.2 (KLLK) <sub>4</sub> Peptide and Interface Model .....	117
4.2.3 Cumulative Residency Time .....	121
4.2.4 Simulation Procedure – Single Peptide Simulations .....	121
4.2.5 Simulation Procedure – Single peptide interface simulations .....	122
4.2.6 Simulation Procedure – Multi-peptide interface aggregation simulations .....	123
4.3 Results and Discussion .....	124
4.3.1 Single peptide folding simulations .....	124
4.3.2 Single peptide interface simulations .....	125
4.3.3 Multi-peptide simulations .....	131
4.4 Discussion and Conclusions .....	136
4.5 Acknowledgements .....	139
4.6 References .....	140
<b>CHAPTER 5 .....</b>	<b>155</b>
<b>Aggregation of Amyloid Beta in the Presence of Naturally Occurring Phenolic</b>	
<b>Inhibitors using Coarse-Grained Simulations .....</b>	<b>155</b>
5.1 Introduction .....	157
5.2 Methods .....	162
5.2.1 Discontinuous Molecular Dynamics and PRIME20 Force Field .....	162
5.2.2 Inhibitor Model – Coarse-Graining .....	163
5.2.3 Inhibitor Model – Atomistic Simulations .....	164
5.2.4 A $\beta$ (17-36) Model .....	166
5.2.5 Simulation Procedure – Coarse-Grained Inhibitor Simulations .....	167
5.3 Results and Discussion .....	168
5.3.1 Peptides Only .....	168
5.3.2 Inhibitor-Peptide Binding During Lag Phase .....	169
5.3.3 Inhibitor-Peptide Interaction Comparison .....	170
5.4 Discussion and Conclusions .....	170
5.5 Acknowledgements .....	172
5.6 References .....	173

<b>CHAPTER 6</b> .....	<b>186</b>
<b>Conclusions and Future Work</b> .....	<b>186</b>
6.1 Conclusions.....	187
6.2 Future Recommendations .....	190
6.2.1 Interface Monolayer Simulations .....	191
6.2.2 Inhibitor Simulations .....	192
6.2.3 Expanding the PRIME20 Force Field .....	192
6.3 References.....	199

## LIST OF TABLES

<b>Table 2.1. Oligomer growth rate, maximum number of peptides in oligomers, and oligomer to fibril conversion time for crowder volume fractions <math>\phi=0.00, 0.10,</math> and <math>0.20</math> and crowder diameters <math>5\text{\AA}, 20\text{\AA},</math> and <math>40\text{\AA}</math> and total crowder surface area. ....</b>	<b>71</b>
<b>Table 2.2. The maximum fraction of oligomeric peptides in dimers and in hexamers for crowder volume fractions <math>\phi=0.00, 0.10,</math> and <math>0.20</math> and crowder diameters <math>5\text{\AA}, 20\text{\AA},</math> and <math>40\text{\AA}</math> and total crowder surface area.....</b>	<b>72</b>
<b>Table 2.3. Fibril primary nucleation rate, critical nucleation rate, and maximum growth rate, elongation rate, and secondary nucleation rate for crowder volume fractions <math>\phi=0.00, 0.10,</math> and <math>0.20</math> and crowder diameters <math>5\text{\AA}, 20\text{\AA},</math> and <math>40\text{\AA}</math> and total crowder surface area obtained by fitting simulation data to Equations 2.6-2.11. ....</b>	<b>73</b>
<b>Table 3.1. Crowder- side chain interaction energy parameters. ....</b>	<b>104</b>
<b>Table 4.1. Lysine-lysine and leucine-leucine attractive square well (negative) and repulsive square shoulder (positive) interaction strengths as a function their phase location.....</b>	<b>153</b>
<b>Table 5.1. Coarse grained groups for phenolic compounds.....</b>	<b>185</b>

## LIST OF FIGURES

<p><b>Figure 2.1. Snapshots of simulation progress for simulations. Top row: no crowders at <math>t^* =</math> (A) 0, (B) 400, (C) 1100, and (D) 2000. Bottom row 5Å crowders at crowder volume fraction <math>\phi=0.20</math> at <math>t^* =</math> (E) 0, (F) 90, (G) 220, and (H) 270.....</b></p>	<b>63</b>
<p><b>Figure 2.2. Number of peptides in oligomers vs. reduced time for (A) 5Å crowders (B) 20Å crowders and (C) 40Å crowders at <math>\phi=0.00</math> (green), 0.10 (orange), and 0.20 (purple) along with curve fits to an asymmetric double sigmoidal function. ....</b></p>	<b>64</b>
<p><b>Figure 2.3. Number of peptides in small oligomers of different sizes vs reduced time for no crowders.....</b></p>	<b>65</b>
<p><b>Figure 2.4. Number of peptides in small oligomers of different sizes vs reduced time at <math>\phi=0.10</math> for (A) 40Å crowders, (B) 20Å crowders, and (C) 5Å crowders.....</b></p>	<b>66</b>
<p><b>Figure 2.5. Number of peptides in small oligomers of different sizes vs reduced time at <math>\phi=0.20</math> for (A) 40Å crowders, (B) 20Å crowders, and (C) 5Å crowders.....</b></p>	<b>67</b>
<p><b>Figure 2.6. Difference in free energy between a disordered aggregate and free monomers vs number of peptides for crowder volume fractions <math>\phi=0.20, 0.15, 0.10,</math> and <math>0.05</math> with a crowder diameter of 40Å.....</b></p>	<b>68</b>
<p><b>Figure 2.7. Difference in free energy between a beta sheet and free monomers vs number of peptides at crowder volume fraction <math>\phi=0.20</math> for crowder diameters 5Å, 20Å, and 40Å.....</b></p>	<b>69</b>
<p><b>Figure 2.8. Number of peptides in fibrils vs. reduced time for A) 5Å crowders B) 20Å crowders and C) 40Å crowders at <math>\phi=0.00</math> (green), 0.10 (orange), and 0.20 (purple).....</b></p>	<b>70</b>
<p><b>Figure 3.1. Snapshots of the different types of aggregates observed as final configurations in our simulations with A) 5Å hard sphere crowder</b></p>	

case results in fibrils with many  $\beta$ -sheets B) 40Å hard sphere crowder case results in fibrils with few  $\beta$ -sheets C) 5Å crowders with  $5\epsilon$  interaction results in disordered oligomers and D) 40Å crowders with  $5\epsilon$  interaction results in  $\beta$ -sheet oligomers.....99

**Figure 3.2. Number of peptides in disordered oligomers vs. reduced time for crowder volume fraction  $\phi=0.10$  with crowder diameters (A) 40Å, (B) 20Å, and (C) 5Å.....100**

**Figure 3.3. Number of peptides in  $\beta$ -sheets vs. reduced time for crowder volume fraction  $\phi=0.10$  with crowder diameters (A) 40Å, (B) 20Å, and (C) 5Å. ...101**

**Figure 3.4. Number of peptides in fibrils vs. reduced time for crowder volume fraction  $\phi=0.10$  with crowder diameters (A) 40Å, (B) 20Å, and (C) 5Å. ...102**

**Figure 3.5. Global phase diagram summarizing the types of aggregates present in each simulation at crowder volume fraction  $\phi=0.10$  for different values of the crowder diameter, D, and interaction strength,  $X\epsilon$ . F=fibrils, B=ordered  $\beta$ -sheet oligomers, and D=disordered oligomers.....103**

**Figure S3.1. Number of peptides in disordered oligomers vs. reduced time for crowder volume fraction  $\phi=0.20$  with crowder diameters (A) 40Å, (B) 20Å, and (C) 5Å.....105**

**Figure S3.2. Number of peptides in  $\beta$ -sheets vs. reduced time for crowder volume fraction  $\phi=0.20$  with crowder diameters (A) 40Å, (B) 20Å, and (C) 5Å. ...106**

**Figure S3.3. Number of peptides in fibrils vs. reduced time for crowder volume fraction  $\phi=0.20$  with crowder diameters (A) 40Å, (B) 20Å, and (C) 5Å. ...107**

**Figure S3.4. Proposed global phase diagram summarizing the types of aggregates present in each simulation at crowder volume fraction  $\phi=0.20$  for different values of the crowder diameter, D, and interaction strength,  $X\epsilon$ . F=fibrils, B=ordered  $\beta$ -sheet oligomers, and D=disordered oligomers.....108**

**Figure 4.1. A) Folded native state  $\alpha$ -helix of (KLLK)<sub>4</sub> peptide. B) Extended conformation of (KLLK)<sub>4</sub> peptide in PRIME20. C) Interface**

simulation box. D) The discontinuous interface potentials experienced by hydrophobic leucine residues (orange) and hydrophilic lysine residue potential (cyan) at various distances ( $z$ -position) from the interface. The interface is located  $94.5\text{\AA}$  from the origin and the edge of the simulation box is  $126\text{\AA}$  from the origin. ....145

Figure 4.2. Number of  $\alpha$ -helical hydrogen bonds vs reduced temperature for the folding of a single (KLLK)<sub>4</sub> peptide. The temperatures used for each simulation type presented in this paper are circled for clarity: Single peptide random coil  $T^*=0.17$ , single peptide helix  $T^*=0.15$ , and multi-peptide aggregation  $T^*=0.165$ .....146

Figure 4.3. Number of native contacts ( $\alpha$ -helical hydrogen bonds) vs the  $z$ -coordinate of the center of mass of a single (KLLK)<sub>4</sub> peptide for an initial random coil secondary structure,  $T^*=0.17$ , at interface strengths 0.50 (green) and 1.00 (orange) and for an initial  $\alpha$ -helical secondary structure,  $T^*=0.15$ , at interface strengths 0.50 (purple) and 1.00 (pink). ....147

Figure 4.4. End to end distance vs the  $z$ -coordinate of the center of mass of a single (KLLK)<sub>4</sub> peptide for an initial random coil secondary structure,  $T^*=0.17$ , and interface strength 0.50 (green) and 1.00 (orange) and an initial  $\alpha$ -helical secondary structure,  $T^*=0.15$ , and interface strength 0.50 (purple) and 1.00 (pink).....148

Figure 4.5. Cumulative residency time of a single (KLLK)<sub>4</sub> peptide at an aqueous-organic interface vs reduced time for an initial random coil secondary structure,  $T^*=0.17$ , with interface strengths 0.50 (A) and 1.00 (B) and an initial  $\alpha$ -helical secondary structure,  $T^*=0.15$ , with interface strengths 0.50 (C) and 1.00 (D). ....149

Figure 4.6. Interfacial structures adopted by the (KLLK)<sub>4</sub> peptide: A)  $\alpha$ -helix, B)  $\beta$ -hairpin, and C) random coil. Interface location shown as dotted line. ....150

<b>Figure 4.7. (A) Fraction of peptides in oligomers of any type, (B) the fraction of peptides in helical bundles and (C) the fraction of peptides in <math>\beta</math>-sheets vs reduced time for a peptides that initially have random coil secondary structures at interface strengths 0.0 (gray), 0.50 (green) and 1.00 (orange),and peptides that initially have helical secondary structures at interface strengths 0.0 ( yellow) 0.50 (purple) and 1.00 (pink).</b> .....	<b>151</b>
<b>Figure S4.1. End to end distance vs the z-coordinate of the center of mass of a single (KLLK)<sub>4</sub> peptide for an initial random coil secondary structure, T*=0.17, and interface strength 0.50 (green) and 1.00 (orange) and an initial <math>\alpha</math>-helical secondary structure, T*=0.15, and interface strength 0.50 (purple) and 1.00 (pink)</b> .....	<b>154</b>
<b>Figure 5.1. Coarse-grained representations of A) resveratrol, B) EGCG, C) vanillin, and D) cucurmin to be used in PRIME20 simulations. The side bar defines each coarse-grained group</b> .....	<b>180</b>
<b>Figure 5.2. Pseudobonds used for A) Reseveratrol, B) EGCG, C)Vanillin, and D) Curcumin. Red bonds constrain benzene and benzopyran rings, and green, purple and blue bonds restrain spheres once, twice, and three times removed from the rings.</b> .....	<b>181</b>
<b>Figure 5.3. A<math>\beta</math>(17-36) U-loop for structure formed after 100 billion collisions for its last 20 frames: A) Top view, all peptides B) Side view, all peptides C) Top view, ordered peptides only and D) side view, ordered peptides only</b> .....	<b>182</b>
<b>Figure 5.4. Inhibitor-peptide complexes for A) vanillin, B) resveratrol, C) EGCG, and D) curcumin. Each peptide is a different color, and the inhibitors are colored by element to show their coarse-grained sites.</b> .....	<b>183</b>
<b>Figure 5.5. Radial distribution function g(r) vs. r for resveratrol (green), curcumin (orange) and vanillin (purple).</b> .....	<b>184</b>

<b>Figure 6.1. Simulation box with the proposed monolayer interaction boundary as a dotted line and the aqueous/organic interface as a solid line.....</b>	<b>195</b>
<b>Figure 6.2. Hydrogen bonding system from Fishman and Argos.<sup>1</sup> .....</b>	<b>196</b>
<b>Figure 6.3. Functions A) <math>E_r</math>, B) <math>E_p</math>, and C) <math>E_t</math> with example discontinuous boundaries to maintain compatibility with PRIME20. ....</b>	<b>197</b>
<b>Figure 6.4. Example of proposed coarse graining method using Phenylalanine. Similarly colored groups have the same interaction parameters. ....</b>	<b>198</b>

# **CHAPTER 1**

## **Motivation and Overview**

## 1.1 Motivation

Protein aggregation diseases are caused by the abnormal bundling of misfolded proteins into ordered aggregates, called amyloid fibrils. These diseases have become an important area of research due to their debilitating effects including, dementia, degradation of the central nervous system, and even death. Another issue related to protein aggregation is the aggregation of proteins designed as therapeutics. While the structure of therapeutic proteins is often well characterized in an aqueous environment, their secondary structures can be altered when they encounter an interface or membrane environment.<sup>1-4</sup> This conformational change can lead to reduced protein functionality and predispose proteins to misfolding and aggregation.<sup>5</sup> The research described here is aimed at providing a better understanding of molecular-level protein aggregation process, with the ultimate goal of contributing to the medical research community's efforts to develop a treatment or cure for the protein aggregation diseases and to aid in understanding therapeutic protein aggregation.

Amyloid fibrils are found in over 40 human disorders including Alzheimer's, Parkinson's, and the prion diseases.<sup>6</sup> Each disorder is associated with the aggregation of a distinct protein. In Alzheimer's Disease (AD), it is the amyloid  $\beta$  ( $A\beta$ ) peptide that aggregates to form oligomers (structures consisting of multiple monomeric  $A\beta$  subunits) and eventually fibrils and plaques. Although the fibrils and plaques that are found in the brains of AD patients were once thought to be the cause of AD, more recently attention has shifted to the oligomers as the toxic agent, one reason being that small concentrations of prefibrillar oligomers can cause neuronal cell death in the absence of fibrils.<sup>7</sup> This new view of AD etiology has catalyzed investigations of the peptide assembly pathway; the hope being that

this could reveal which parts of the aggregation pathway might serve as potential targets for drugs aimed at treating AD. Unfortunately, the causes of AD are still unknown, despite the fact that we have known about the existence of these aggregate structures for more than 100 years.<sup>8</sup>

The A $\beta$  peptide is formed in the body when the amyloid precursor protein (APP), which is located in the synapses of neurons, is cleaved in the transmembrane region by  $\gamma$ -secretase and in the extracellular region by  $\beta$ -secretase to form a 40 or 42 residue protein.<sup>9</sup> These are the most common lengths although other lengths can be formed. One important observation in AD patients is a high amount of A $\beta$ <sub>42</sub> relative to A $\beta$ <sub>40</sub>.<sup>10</sup> Despite the prominent role played by A $\beta$  in AD, its biological function remains unidentified. However, it has been proposed that A $\beta$  peptides might aid in maintaining optimal synaptic vesicle release probability to sustain proper brain function.<sup>11</sup> In other words, the presence of A $\beta$  might help regulate the propagation of nerve impulses between neurons. If specific amounts of A $\beta$  are used to regulate brain function, then an excess or deficiency of A $\beta$  could potentially lead to neurodegeneration. Although AD can develop at any time, a recent publication from the National Center for Health Statistics shows that death due to AD is 4000 times more likely in patients 85+ years old than in 45-54 year olds.<sup>12</sup> Since the elderly population is rising significantly, research on AD is more important than ever before.

Most protein aggregation studies are performed *in vitro*, since this is the most straightforward way to probe the biophysics underlying the assembly process. *In vitro* investigations do not, however, capture the influence of biomolecules on A $\beta$  assembly that would occur *in vivo*.<sup>13,14</sup> In fact the rate and extent of oligomer/fibril formation in a crowded

environment like the human brain can differ by orders of magnitude from that *in vitro*.<sup>15</sup> The intracellular and extracellular environments in the human brain are quite crowded with approximately 7% to >40% of the total volume occupied by a variety of macromolecules and structures.<sup>16-18</sup> We believe that learning how aggregation changes in a crowded environment could provide additional insight into how AD progresses.

Despite the vast amounts of research devoted to AD, there are yet no effective treatment options. We think that efforts to discover ways to prevent or treat AD would be aided by having a better understanding of the aggregation of A $\beta$  in the presence of small molecule inhibitors on a molecular level. A detailed account of the molecular events associated with formation of the various fibrillar structures and intermediates in the presence of aggregation inhibitors might provide insights on how to disrupt these processes and help us to identify more efficient aggregation inhibitors in the future.

The effectiveness and safety of therapeutic proteins like insulin<sup>19</sup>, antimicrobial peptides<sup>20,21</sup>, and monoclonal antibodies<sup>22</sup> can be significantly altered by aggregation. In fact, aggregation of therapeutic proteins can elicit an unwanted immune response from the body, leading to severe side effects, even death.<sup>22</sup> Antimicrobial peptides are an innate part of our body's immune system and have promise as therapeutics. These proteins are characterized by their amphipathic character, making them highly susceptible to unfolding and aggregation when they encounter interfaces.<sup>23-30</sup> Examining how model amphipathic peptides behave at an interface could aid in understanding how interfacial interactions alter amphipathic protein folding and aggregation.

Computer simulation offers us the opportunity to examine the structures and kinetics of protein aggregation on a molecular level. In molecular dynamics (MD) simulations, we get to observe the trajectories of a system of proteins and obtain insight into aggregation mechanisms that could not otherwise be obtained from laboratory work. Traditional protein simulations based on atomistic-scale force fields enable us to examine time scales on the order of tens of nanoseconds, but the aggregation of A $\beta$  occurs on timescales ranging from seconds for early oligomerization<sup>31</sup> to hours for fibrillar aggregates<sup>32</sup>, making the process inaccessible to traditional MD.

The timescale limitations of traditional MD can be circumvented by applying an alternate form of MD called discontinuous molecular dynamics (DMD).<sup>33,34</sup> In traditional MD the force associated with the potential, usually a continuous function of the interatomic separation, e.g. the Lennard Jones potential, must be recalculated every time the atoms are moved because the value of the potential varies. In DMD simulations, discontinuous potentials such as the hard sphere and square well potentials are used to simplify the calculation of interatomic interactions; these typically have only two values, either zero or the value of the interaction potential  $\epsilon$ , as opposed to the more complex Lennard Jones potential. The advantage of using DMD over traditional MD is that the atoms do not need to be moved at short regularly spaced time steps. Since the atoms in DMD move linearly and the potentials are discontinuous, the only time the simulation needs to recalculate velocities and positions is when a discontinuity is encountered. Therefore, the simulation can be advanced from collision to collision.

The DMD algorithm begins with the calculation of the times until the next collision for all atoms in the system based on knowledge of their current positions and velocities. Next, the shortest collision time is chosen and all atoms are advanced to their new positions. At the time of collision, the two atoms with the smallest collision time collide but the others do not. New velocities for the colliding atoms are calculated using conservation of momentum and conservation of kinetic energy, and the algorithm repeats until completion. There are several types of “collisions” that can occur. These include: excluded-volume events, covalent-bond events, pseudobond events, square-well hydrogen-bond events, and square-well hydrophobic-interaction events. Excluded volume events occur when two hard spheres collide and repel each other. A bonding event occurs when two spheres attempt to move outside of their assigned bond length, but are forced back together by an infinite repulsion. Square well events include well-capture, well-bounce, and well-dissociation “collisions” when a sphere enters, attempts to leave, or leaves the square well of another sphere.<sup>35</sup>

An additional way to alleviate the timescale limitations of atomistic MD is to coarse grain the peptide geometry by combining groups of atoms into “united atoms”, which are usually represented by spheres. One popular coarse-grained model used for studying peptides is a four-sphere per residue model in which each amino acid residue is represented by three backbone spheres and one side-chain sphere.<sup>36</sup> The four-sphere per residue model provides a balance of accuracy and simplicity that is ideal for DMD. Hall and coworkers were among the first to adopt a four-sphere per residue model. In 2004 they introduced a new intermediate resolution protein model, called PRIME (Protein Intermediate Resolution Model),

appropriate for homoproteins like polyalanine and polyglutamine. More recently, PRIME was extended to heteroproteins culminating in PRIME20 which describes the geometry and energetics for all 20 coarse grained amino acids. The parameters used in PRIME20 were derived utilizing a perceptron learning algorithm and a modified stochastic learning algorithm that compared the energy of native state proteins in the Protein Database (PDB) with those of a large number of decoy structures.<sup>37</sup> Using DMD in conjunction with the PRIME20 force field should provide additional molecular level insight to aid in the understanding of A $\beta$  and therapeutic protein aggregation.

## **1.2 Overview**

In this section a summary of the rest of the dissertation is provided. All chapters provide their own literature review and references.

### **1.2.1 Effects of Macromolecular Crowding on Amyloid Beta (16-22) Aggregation**

In chapter 2 we examine the effect of hard sphere crowding on protein aggregation. The systems contained 192 A $\beta$ (16-22) peptides and crowders of diameters 5Å, 20Å, and 40Å, represented here by simple hard spheres, at crowder volume fractions of 0.00, 0.10, and 0.20. Results show that both crowder volume fraction and crowder diameter have a large impact on fibril and oligomer formation. The addition of crowders to a system of peptides increases the rate of oligomer formation, shifting from a slow ordered formation of oligomers in the absence of crowders, similar to nucleated polymerization, to a fast collapse of peptides

and subsequent rearrangement characteristic of nucleated conformational conversion with a high maximum in the number of peptides in oligomers as total crowder surface area increases. The rate of conversion from oligomers to fibrils also increases with increasing total crowder surface area, giving rise to an increased rate of fibril growth. In all cases, larger volume fractions and smaller crowders provide the greatest aggregation enhancement effects. We also show that the size of the crowders influences the formation of specific oligomer sizes. In our simulations the 40Å crowders enhance the number of dimers relative to the numbers of trimers, hexamers, pentamers, and hexamers, while the 5Å crowders enhance the number of hexamers relative to the numbers of dimers, trimers, tetramers, and pentamers.

### **1.2.2 Effects of Hydrophobic Macromolecular Crowders on Amyloid Beta (16-22) Aggregation**

The effects of hydrophobic crowders on A $\beta$  (16-22) oligomer and fibril formation are presented in chapter 3. We simulated systems containing 48 A $\beta$ (16-22) peptides and crowders of diameters 5Å, 20Å, and 40Å, represented by hard spheres or spheres with square well/shoulder interactions, at a crowder volume fractions of 0.10. Results show that low levels of crowder hydrophobicity are capable of increasing fibrillation lag time and high levels of crowder hydrophobicity can fully prevent the formation of fibrils. Furthermore, at high levels of crowder hydrophobicity, the systems with 5Å crowders form only disordered oligomers while those with 40Å crowders form only  $\beta$ -sheets. Our simulations also show that hard sphere crowders increase the anti-parallel  $\beta$ -sheet content of fibrils while attractive

crowders reduce it. Finally we found that strong hydrophobic crowders have the ability to alter the secondary structure of the A $\beta$ (16-22) monomers.

### **1.2.3 Folding and Aggregation of Amphipathic Peptides at an Aqueous-Organic Interface**

In chapter 4 we examine the effect of an aqueous/organic interface on the folding and aggregation of a model amphipathic peptide (KLLK)<sub>4</sub>. The systems contain either a single (KLLK)<sub>4</sub> peptide to study folding or 48 (KLLK)<sub>4</sub> peptides to study aggregation. The peptides were initialized in random coil or  $\alpha$ -helical conformations and allowed to interact with either strong or weak interfaces. The interface strength governs how much the hydrophilic and hydrophobic amino acid side chains want to be in their respective phases. Results show that the weak interface is capable of folding the randomly coiled peptide into an amphipathic  $\alpha$ -helix, while the strong interface disrupts intra-peptide  $\alpha$ -helical hydrogen bonding and unfolds the peptides into random coil or  $\beta$ -hairpin conformations. Regardless of the number of proteins in the simulation, peptides bound irreversibly to the strong interface because of their side chain preference for their respective phases. When peptides interacted with the weak interface, they were bound reversibly. The (KLLK)<sub>4</sub> peptide was more likely to bind to the interface when it was initially in an  $\alpha$ -helical conformation than when it was a random coil. We believe this is because it has a larger hydrophobic moment, increasing the probability of interacting with the organic phase. In the multi-peptide case, the aggregation was highest when the peptides were initially in an  $\alpha$ -helical state in the presence of a weak

interface. This occurred because the increased local concentration of helices just outside the interface promoted the formation of helical bundles. Aggregation was lowest when the peptides were initially in a random coil state in the presence of a strong interface. The negligible hydrophobic moment in the random coil state meant that the peptides bound to the interface infrequently, making it difficult to form interfacial  $\beta$ -sheets. The most energetically favorable type of aggregate in the aqueous phase was the helical bundle, and the most energetically favorable type of aggregate at the interface was the  $\beta$ -sheet.

#### **1.2.4 Aggregation of Amyloid Beta in the Presence of Naturally Occurring Phenolic Inhibitors**

Chapter 5 presents preliminary work on the aggregation of A $\beta$ (17-36) in the presence of naturally-occurring small molecule aggregation inhibitors vanillin, resveratrol, curcumin, and epigallocatechin gallate (egcg) molecules. We describe our methodology for integrating small molecule inhibitors into the PRIME20 force field and examine early DMD simulations. These simulations include 8 A $\beta$ (17-36) molecules and either 128 or 256 vanillin, resveratrol, curcumin, or epigallocatechin-3-gallate (EGCG) molecules. When simulated in the absence of inhibitors, the A $\beta$ (17-36) peptide was capable of forming the U-loop structure characteristic of A $\beta$  protofilaments. During the lag phase and initial aggregation, we find that resveratrol binds most frequently with A $\beta$ , followed by curcumin, and finally vanillin. We also observed that curcumin and resveratrol are localized to the interior of aggregates, possibly preventing further aggregation into ordered amyloid structures, while vanillin binds

the exterior of aggregates. Limited simulations on EGCG show that it is capable of binding monomers unlike the other molecules.

### **1.2.5 Future Work**

Finally, in Chapter 6 we provide an outline for future work on interfacial monolayer formation and small molecule inhibitors. We also provide suggestions on how to expand the PRIME20 force field to include more detailed amino acid side chain representations and improve the flexibility of the hydrogen bonding.

## **1.3 Publications**

Chapters 2-5 are based on the following publications:

Chapter 2: D. C. Latshaw II, M. Cheon, and C. K Hall, “Effects of Macromolecular Crowding on Amyloid Beta (16-22) Aggregation using Coarse-Grained Simulations”, Accepted Journal of Physical Chemistry B

Chapter 3: D. C. Latshaw II and C. K Hall, “Effects of Hydrophobic Macromolecular Crowders on Amyloid Beta (16-22) Aggregation using Coarse-Grained Simulations”, in preparation

Chapter 4: D. C. Latshaw II and C. K Hall, “Folding and Aggregation of Amphipathic Peptides at an Aqueous-Organic Interface using Coarse-Grained Simulations”, in preparation

Chapter 5: D. C. Latshaw II and C. K Hall, “Aggregation of Amyloid Beta in the Presence of Naturally Occurring Phenolic Inhibitors using Coarse-Grained Simulations”, in preparation

## 1.4 References

1. Kim HJ, Decker EA, McClements DJ. Role of postadsorption conformation changes of beta-lactoglobulin on its ability to stabilize oil droplets against flocculation during heating at neutral pH. *Langmuir*. 2002;18(20):7577-7583. doi: 10.1021/la020385u.
2. Fang Y, Dalgleish DG. Conformation of beta-lactoglobulin studied by FTIR: Effect of pH, temperature, and adsorption to the oil-water interface. *J Colloid Interface Sci*. 1997;196(2):292-298. doi: 10.1006/jcis.1997.5191.
3. Husband FA, Garrod MJ, Mackie AR, Burnett GR, Wilde PJ. Adsorbed protein secondary and tertiary structures by circular dichroism and infrared spectroscopy with refractive index matched emulsions. *J Agric Food Chem*. 2001;49(2):859-866. doi: 10.1021/jf000688z.
4. Lefevre T, Subirade M. Formation of intermolecular beta-sheet structures: A phenomenon relevant to protein film structure at oil-water interfaces of emulsions. *J Colloid Interface Sci*. 2003;263(1):59-67. doi: 10.1016/S0021-9797(03)00252-2.
5. Wang W. Protein aggregation and its inhibition in biopharmaceutics. *International Journal of Pharmaceutics (Kidlington)*. 2005;289(1-2):1-30. doi: 10.1016/j.ijpharm.2004.11.014 ER.
6. Chiti F, Dobson CM. Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem*. 2006;75:333-366. doi: 10.1146/annurev.biochem.75.101304.123901.

7. Lambert M, Barlow A, Chromy B, *et al.* Diffusible, nonfibrillar ligands derived from A beta(1-42) are potent central nervous system neurotoxins. *Proc Natl Acad Sci U S A.* 1998;95(11):6448-6453. doi: 10.1073/pnas.95.11.6448.
8. Tanzi R, Bertram L. Twenty years of the alzheimer's disease amyloid hypothesis: A genetic perspective. *Cell.* 2005;120(4):545-555. doi: 10.1016/j.cell.2005.02.008.
9. Kang J, Lemaire H, Unterbeck A, *et al.* The precursor of alzheimers-disease amyloid-A4 protein resembles a cell-surface receptor. *Nature.* 1987;325(6106):733-736. doi: 10.1038/325733a0.
10. Chen Y, Glabe CG. Distinct early folding and aggregation properties of alzheimer amyloid-beta peptides A beta 40 and A beta 42 - stable trimer or tetramer formation by A beta 42. *J Biol Chem.* 2006;281(34):24414-24422. doi: 10.1074/jbc.M602363200.
11. Abramov E, Dolev I, Fogel H, Ciccotosto GD, Ruff E, Slutsky I. Amyloid-beta as a positive endogenous regulator of release probability at hippocampal synapses. *Nat Neurosci.* 2009;12(12):1567-U120. doi: 10.1038/nn.2433.
12. Xu J, Kochanek K, Sherry L, Murphy B, Tejada-Vera B. Deaths: Final data for 2007. *National Vital Statistics Reports.* 2010.
13. Gasser B, Saloheimo M, Rinas U, *et al.* Protein folding and conformational stress in microbial cells producing recombinant proteins: A host comparative overview. *Microbial Cell Factories.* 2008;7:11. doi: 10.1186/1475-2859-7-11.
14. Ignatova Z. Monitoring protein stability *in vivo*. *Microbial Cell Factories.* 2005;4:23. doi: 10.1186/1475-2859-4-23.

15. Ellis R. Macromolecular crowding: Obvious but underappreciated. *Trends Biochem Sci.* 2001;26(10):597-604. doi: 10.1016/S0968-0004(01)01938-7.
16. Fulton A. How crowded is the cytoplasm. *Cell.* 1982;30(2):345-347. doi: 10.1016/0092-8674(82)90231-8.
17. Zimmerman S, Trach S. Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of escherichia-coli. *J Mol Biol.* 1991;222(3):599-620. doi: 10.1016/0022-2836(91)90499-V.
18. Zimmerman S, Minton A. Macromolecular crowding - biochemical, biophysical, and physiological consequences. *Annu Rev Biophys Biomol Struct.* 1993;22:27-65. doi: 10.1146/annurev.bb.22.060193.000331.
19. Schernthaner G. Immunogenicity and allergenic potential of animal and human insulins. *Diabetes Care.* 1993;16:155-165.
20. Reddy K, Yedery R, Aranha C. Antimicrobial peptides: Premises and promises. *Int J Antimicrob Agents.* 2004;24(6):536-547. doi: 10.1016/j.ijantimicag.2004.09.005.
21. Bulet P, Stocklin R, Menin L. Anti-microbial peptides: From invertebrates to vertebrates. *Immunol Rev.* 2004;198:169-184. doi: 10.1111/j.0105-2896.2004.0124.x.
22. Hermeling S, Crommelin D, Schellekens H, Jiskoot W. Structure-immunogenicity relationships of therapeutic proteins. *Pharm Res.* 2004;21(6):897-903. doi: 10.1023/B:PHAM.0000029275.41323.a6.
23. Gerhardt A, Bonam K, Bee JS, Carpenter JF, Randolph TW. Ionic strength affects tertiary structure and aggregation propensity of a monoclonal antibody adsorbed to

- silicone oil-water interfaces. *J Pharm Sci.* 2013;102(2):429-440. doi: 10.1002/jps.23408.
24. Li J, Pinnamaneni S, Quan Y, Jaiswal A, Andersson FI, Zhang X. Mechanistic understanding of protein-silicone oil interactions. *Pharm Res.* 2012;29(6):1689-1697. doi: 10.1007/s11095-012-0696-6.
25. Felsovalyi F, Janvier S, Jouffray S, Soukiassian H, Mangiagalli P. Silicone-oil-based subvisible particles: Their detection, interactions, and regulation in prefilled container closure systems for biopharmaceuticals. *J Pharm Sci.* 2012;101(12):4569-4583. doi: 10.1002/jps.23328.
26. Bee JS, Schwartz DK, Trabelsi S, *et al.* Production of particles of therapeutic proteins at the air-water interface during compression/dilation cycles. *Soft Matter.* 2012;8(40):10329-10335. doi: 10.1039/c2sm26184g.
27. Rudiuk S, Cohen-Tannoudji L, Huille S, Tribet C. Importance of the dynamics of adsorption and of a transient interfacial stress on the formation of aggregates of IgG antibodies. *Soft Matter.* 2012;8(9):2651-2661. doi: 10.1039/c2sm07017k.
28. Badkar A, Wolf A, Bohack L, Kolhe P. Development of biotechnology products in pre-filled syringes: Technical considerations and approaches. *Aaps Pharmscitech.* 2011;12(2):564-572. doi: 10.1208/s12249-011-9617-y.
29. Majumdar S, Ford BM, Mar KD, Sullivan VJ, Ulrich RG, D'Souza AJM. Evaluation of the effect of syringe surfaces on protein formulations. *J Pharm Sci.* 2011;100(7):2563-2573. doi: 10.1002/jps.22515.

30. Gerhardt A, McGraw NR, Schwartz DK, Bee JS, Carpenter JF, Randolph TW. Protein aggregation and particle formation in prefilled glass syringes. *J Pharm Sci*. 2014;103(6):1601-1612. doi: 10.1002/jps.23973.
31. Bitan G, Kirkitadze M, Lomakin A, Vollers S, Benedek G, Teplow D. Amyloid beta-protein (A beta) assembly: A beta 40 and A beta 42 oligomerize through distinct pathways. *Proc Natl Acad Sci U S A*. 2003;100(1):330-335. doi: 10.1073/pnas.222681699.
32. Kaye R, Head E, Thompson J, *et al*. Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science*. 2003;300(5618):486-489. doi: 10.1126/science.1079469.
33. M.P. Nelson, D.J. Tildesley. *Computer simulation of liquids*. Oxford, NY: Oxford Science Publications; 1987:p. 101-108.
34. J.M. Haile. *Molecular dynamics simulation: Elementary methods*. Wiley-Interscience Publication; 1992:p. 104-141.
35. Smith A, Hall C. Assembly of a tetrameric alpha-helical bundle: Computer simulations on an intermediate-resolution protein model. *Proteins*. 2001;44(3):376-391. doi: 10.1002/prot.1103.
36. Wu C, Shea J. Coarse-grained models for protein aggregation. *Curr Opin Struct Biol*. 2011;21(2):209-220. doi: 10.1016/j.sbi.2011.02.002.
37. Cheon M, Chang I, Hall CK. Extending the PRIME model for protein aggregation to all 20 amino acids. *Proteins*. 2010;78(14):2950-2960. doi: 10.1002/prot.22817.

## **CHAPTER 2**

### **Effects of Macromolecular Crowding on Amyloid Beta (16-22) Aggregation using Coarse-Grained Simulations**

Chapter 2 is essentially a manuscript by David C Latshaw II, Mookyung Cheon, and Carol K Hall accepted by the Journal of Physical Chemistry B.

# Effects of Macromolecular Crowding on Amyloid Beta (16-22) Aggregation using Coarse-Grained Simulations

David C Latshaw II<sup>a</sup>, Mookyung Cheon<sup>b</sup>, and Carol K Hall<sup>a\*</sup>

<sup>a</sup>Department of Chemical and Biomolecular Engineering, North Carolina State University, 911 Partners Way, Raleigh, North Carolina 27695, United States

<sup>b</sup>Center for Proteome Biophysics, Department of New Biology, Daegu Gyeongbuk Institute of Science and Technology (DGIST), Daegu 711-873, Korea

## Abstract

To examine the effect of crowding on protein aggregation, discontinuous molecular dynamics (DMD) simulations combined with an intermediate resolution protein model, PRIME20, were applied to a peptide/crowder system. The systems contained 192 A $\beta$ (16-22) peptides and crowders of diameters 5Å, 20Å, and 40Å, represented here by simple hard spheres, at crowder volume fractions of 0.00, 0.10, and 0.20. Results show that both crowder volume fraction and crowder diameter have a large impact on fibril and oligomer formation. The addition of crowders to a system of peptides increases the rate of oligomer formation, shifting from a slow ordered formation of oligomers in the absence of crowders, similar to nucleated polymerization, to a fast collapse of peptides and subsequent rearrangement characteristic of nucleated conformational conversion with a high maximum in the number of peptides in oligomers as total crowder surface area increases. The rate of conversion from oligomers to fibrils also increases with increasing total crowder surface area, giving rise to an

increased rate of fibril growth. In all cases, larger volume fractions and smaller crowders provide the greatest aggregation enhancement effects. We also show that the size of the crowders influences the formation of specific oligomer sizes. In our simulations the 40Å crowders enhance the number of dimers relative to the numbers of trimers, hexamers, pentamers, and hexamers, while the 5Å crowders enhance the number of hexamers relative to the numbers of dimers, trimers, tetramers, and pentamers. These results are in qualitative agreement with previous experimental and theoretical work.

## 2.1 Introduction

Amyloid fibrils are found in over 40 human disorders including Alzheimer's, Parkinson's, and the prion diseases.<sup>1</sup> Each disorder is associated with the aggregation of a distinct protein. In Alzheimer's Disease (AD), it is the amyloid  $\beta$  ( $A\beta$ ) peptide that aggregates to form oligomers (structures consisting of multiple monomeric  $A\beta$  subunits) and eventually fibrils and plaques. Although the fibrils and plaques that are found in the brains of AD patients were once thought to be the cause of AD, more recently attention has shifted to the oligomers as the toxic agent, one reason being that small concentrations of prefibrillar oligomers can cause neuronal cell death in the absence of fibrils.<sup>2</sup> This new view of AD etiology has catalyzed investigations of the peptide assembly pathway leading from monomers to oligomers to fibrils, the hope being that this could reveal which parts of the aggregation pathway might serve as potential targets for drugs aimed at treating AD. Most of these studies are done *in vitro*, since this is the most straightforward way to probe the biophysics underlying the assembly process. *In vitro* investigations do not, however, capture the influence of biomolecules on  $A\beta$  assembly that would occur *in vivo*.<sup>3,4</sup> In fact the rate and extent of oligomer/fibril formation in a crowded environment like the human brain can differ by orders of magnitude from that *in vitro*.<sup>5</sup> The intracellular and extracellular environments in the human brain are quite crowded with approximately 7% to >40% of the total volume occupied by a variety of macromolecules and structures.<sup>6-8</sup> In this paper we use computer simulations to learn how crowding affects fibril and oligomer formation as well as aggregation mechanisms.

The influence of crowding agents on the aggregation of proteins has been the subject of a number of experimentally-based investigations. Hatters *et al.* examined the effects of adding dextran 10 at dextran volume fractions of up to  $\phi=0.11$  on the fibrillization of apolipoprotein C-II (apoC-II).<sup>9</sup> Despite the fact that a significant portion of the solution was occupied by dextran 10, the protein's fibril structure was unchanged compared to that in the absence of crowding. Another important finding was that dextran 10 did not interact directly with apoC-II, indicating that it was simply the volume excluded by dextran 10 that accelerated apoC-II fibrillization, not a crowder-peptide interaction. Uversky *et al.* did a comprehensive investigation of how different crowding agents influence the fibrillization of  $\alpha$ -synuclein.<sup>10</sup> Polyethylene glycol (PEG), dextran, ficoll, lysozyme, and bovine serum albumin (BSA) all decreased fibrillization lag time and increased the rate of fibrillization, but the effects were more pronounced with some crowders than with others. At constant concentration, crowders with longer chain lengths promoted aggregation faster than those with shorter chain lengths. Munishkina *et al.* explored the effect of macromolecular crowding on the aggregation pathways of four different proteins: a disordered form of S-carboxymethyl- $\alpha$ -lactalbumin, the natively-disordered  $\alpha$ -synuclein, bovine core histone, and the folded monomeric and hexameric forms of human insulin. They found that if the protein preferentially adopts a multimeric native state (the bovine histone core and hexameric human insulin both occur naturally in a multimeric state), the fibrillization of that protein is slowed because the multimeric species is stabilized by the crowders. They also found that the fibrillization of proteins that have a low degree of native structure is accelerated by crowding. They interpreted their results to mean that crowding may accelerate fibrillization

because it can promote the formation of a partially folded form of the protein that is highly amyloidogenic. Fung *et al.* examined the effect of adding the simple saccharides glucose, galactose, fructose, mannose, and sucrose on the aggregation of A $\beta$ 40 and A $\beta$ 42.<sup>11</sup> They found that the sugars that do not directly interact with A $\beta$  promote fibril formation, while the sugars that do interact with A $\beta$  (in their case through hydrogen bonding) promote nucleation and the formation of smaller protofibrils. In a similar vein, Sukenik *et al.* examined the effects of polyol osmolytes, glycerol and sorbitol, and PEG on the aggregation of a synthetic peptide, MET16, that folds into a  $\beta$ -hairpin and can aggregate into a fibrillar structure.<sup>12,13</sup> Variations in the molecular weight of PEG, glycerol, and sorbitol produced minimal variation in the fibrillization rate, but the polyol osmolytes increased the lag time for fibril formation and increased the fibrillar mass at equilibrium. The authors suggested that the strong polyol osmolyte effect was due to its distortion of the water hydrogen bond network which could change the preferred conformation of the peptides, altering their preferred aggregation state. Upon fitting circular dichroism and ThT fluorescence aggregation data to a simple kinetic model they discovered that addition of PEG leads to extensive fibril fragmentation, while addition of polyol osmolytes stabilized fibrillar structures by decreasing monomer dissociation.

The impact of crowding on protein aggregation has also been investigated using theory. The Minton group has shown that macromolecular crowding has different effects on protein aggregation, depending on what the rate limiting aggregation mechanism is.<sup>14,15</sup> Aggregation that is slow and reaction-limited is typically enhanced by the presence of crowders because the crowders increase the effective protein concentration and create

depletion forces between the proteins, while fast, diffusion-limited aggregation is hindered by crowders because diffusion decreases with increasing crowder concentration.<sup>16</sup> For example, since the dock-lock mechanism that is believed to govern A $\beta$  fibrillization is essentially a reaction-limited process, the presence of crowders should increase fibrillization.<sup>17</sup> In a series of papers by Kinjo and Takada, the effects of macromolecular crowding and chaperones on protein aggregation and folding were examined using density functional theory in conjunction with dynamic rate equations.<sup>18-20</sup> Aggregation, folded-unfolded protein reactions, and protein-chaperone binding reactions were modeled by diffusion and the dynamic rate equations. Proteins, crowders, and chaperones were modeled using hard spheres with square well potentials. They found that crowding enhanced the aggregation of the model unfolded proteins, but stabilized the model native state proteins so long as they were uniformly distributed in space. They also found that crowding accelerated the transition from unfolded to folded model protein if the folding rate was fast, and destabilized model proteins if the folding rate was slow, which is in agreement with Zhou *et al.*<sup>16</sup> More recently Minton examined theoretical models that incorporated a time-dependent macromolecular crowder concentration to mirror the observation that the concentration of soluble proteins in the human brain tends to increase linearly with time.<sup>21</sup> The major conclusion of this study was that rate constants for protein aggregation are undetectably small in the absence of crowders, and that the accumulation of crowders over time is what increases the rate constants to a level that actually induces aggregation. As the crowder concentration increased linearly with time, the aggregation rate constant also increased, at a minimum, exponentially.

Molecular-level simulations have also been used to examine the effect of crowding on protein aggregation. O'Brien *et al.* conducted atomistic simulations in implicit solvent to examine how crowder volume fraction, size, and shape affected the oligomerization of a 10-residue fragment of the transthyretin (TTR) protein.<sup>22</sup> The crowders were modeled as softly repulsive spheres of diameter 7Å, 12Å, and 22Å and as spherocylinders with diameter 7Å and length 23.1Å; the simulations were performed on a system containing 2 to 4 peptides at a concentration of 15 to 31 mM with crowder volume fractions of  $\phi=0.05$  to 0.20. They observed that adding crowders of any size and concentration to the simulation enhanced aggregation. One interesting result was that the addition of spherical crowders to the simulations destabilized TTR dimers in favor of trimeric and tetrameric oligomers. They also found that as the size of the crowding spheres increased, the level of aggregation enhancement decreased. Finally, spherocylinder crowders destabilized the oligomers to a larger extent than the spherical crowders, highlighting the importance of the shape of crowding molecules. Magno *et al.* simulated a system of 125 amphipathic 10-bead coarse-grained polypeptides designed specifically for studying the physics of fibril formation, i.e. the model peptide did not represent a specific sequence or have a defined length.<sup>23</sup> The peptide had a tunable energy parameter that could be shifted from an aggregation-prone state ( $\beta$ -state, which favors a cross- $\beta$  structure) to an aggregation-protected state ( $\pi$ -state, which favors a monomeric state). Their general conclusion was that crowding greatly accelerates the aggregation of peptides that have a reaction-limited aggregation mechanism but only modestly accelerates the aggregation of peptides that have a diffusion-limited aggregation mechanism. The former conclusion agrees with that of Ellis and Minton, but the latter

conclusion does not.<sup>14</sup> The Magno *et al.* polypeptide model is well suited for a fundamental study of protein aggregation in the presence of crowders, but it does not provide any information about the effects of crowding on more complex proteins. Co *et al.* used lattice Monte Carlo to simulate the effects of crowding and confinement on the fibrillization of 6, 10 and 24 peptides using a toy model with the sequence +HHPPHH-, where + is positive, - is negative, H is hydrophobic, and P is polar.<sup>24</sup> The sequence was designed to fold into a compact U-shape similar to a  $\beta$ -hairpin. Crowders were modeled as squares or rectangles on the lattice and confinement was modeled using a box with hard walls. They showed that crowding and confinement can decrease fibrillization lag time up to intermediate values of crowder surface area and confinement box length, after which the lag time begins to increase. Their results highlighted the complex nature of crowding, with longer fibrillization lag times at both high and low crowder concentrations, as predicted by Zimmerman and Minton<sup>8</sup>, and recently observed by Cabaleiro-Lago *et al.* using amine-modified polystyrene nanoparticles.<sup>25</sup>

The physical picture that emerges concerning the effects of crowders on protein aggregation is reminiscent of that concerning the effects of crowders on protein folding. Zhou *et al.* summarize the effects of crowding on protein folding in terms of the existence of an energy barrier to folding that is a function of the amount of space that a protein must occupy in order to fold.<sup>16</sup> If the folding pathway contains an intermediate that occupies more space than the initially unfolded protein then the rate of folding will be slowed because crowding prevents the protein from expanding; this raises the energy barrier to folding. If the folding pathway takes the protein through an intermediate that is more compact than the

unfolded protein, then the rate of folding tends to be enhanced by crowding because crowding forces the protein into a more collapsed conformation, effectively lowering the folding energy barrier. Additionally, intra-chain diffusion in large proteins can be decreased by the presence of crowders, slowing down the overall rate of folding if multiple regions must fold independently before the final tertiary structure is achieved.

These ideas can be generalized to the case of protein oligomerization and fibrillization. While proteins in dilute solution can form a wide variety of aggregate structures through many kinetic pathways, crowded systems have energetic penalties associated with forming aggregates that are not highly compact so the number of kinetic pathways tends to be smaller.<sup>26</sup> In the case of amyloidogenic proteins a given number of monomers typically occupies more space than an oligomer or fibril made up of the same number of proteins, making monomers less energetically favorable. Additionally, intrinsically disordered proteins which don't have a defined quaternary structure in dilute solution might be driven to interact with each other to form oligomers or fibrils in a crowded environment because they would then occupy less space. The same idea applies to a partially-folded protein. A partially-folded protein may preferentially interact with other partially-folded proteins in the presence of crowders because an aggregate of partially-folded proteins is more thermodynamically stable than a bunch of isolated partially-folded proteins. In this paper we apply a combination of discontinuous molecular dynamics (DMD) and the PRIME20 force field to examine how the aggregation of a multi-peptide system containing A $\beta$ (16-22) is impacted by macromolecular crowding. Although A $\beta$  is typically observed in its 40 or 42 residue form, our simulations focus on the A $\beta$ (16-22) peptide, which has been

shown to be a key sequence in the formation of A $\beta$  oligomers and fibrils, and has the ability to form fibrils on its own.<sup>27-29</sup> In our simulations we monitor the aggregation, oligomerization and fibril formation of a system containing 192 peptides using hard sphere crowders at crowding volume fractions of  $\phi=0.00, 0.10, \text{ and } 0.20$  and crowder diameters of  $D=5\text{\AA}, 20\text{\AA}, \text{ and } 40\text{\AA}$ . Since oligomeric structures have become important in the study of A $\beta$  toxicity, we also look closely at how crowding affects the stability of oligomers and their conversion to other larger species. We explore how crowding affects peptide association rates, and aggregation mechanisms. We compare our results to theoretical predictions of the effects of crowding on protein aggregation by O'Brien *et al.*, Munishkina *et al.*, Zhou *et al.*, and Zimmerman and Minton<sup>8,22,26,30</sup> and provide molecular level detail about how oligomer and fibril formation in a crowded medium differs from that in a dilute solution.

Our simulation results show that when crowders are added to a system of peptides they increase the rate of oligomer formation and the maximum number of oligomers that form. Oligomer formation shifts from being a slow process characterized by the templated addition of monomers to existing oligomers in the absence of crowders to a fast collapse and subsequent rearrangement that leads to a high maximum in the number of oligomers formed when crowders are present. Addition of crowders also increases the rate of conversion from oligomers to fibrils, giving rise to an increased rate of fibril growth. In all cases, larger crowder volume fractions and smaller crowder diameters provide the greatest enhancement of oligomerization and fibrillization. This enhancement is largely a consequence of the depletion forces between the peptides due to the crowders. These forces drive peptide-peptide association, making oligomers and fibrils more thermodynamically favorable than an

equivalent number of monomers. Additionally we have shown that, depending on the size of the crowders relative to the peptides, specific oligomer sizes can be stabilized. In our simulations with 40Å crowders, the dimers were stabilized and persisted longer relative to the trimers, tetramers, pentamers, and hexamers. We surmise that this is because the 40Å crowders have interstitial spaces that are large enough to easily accommodate dimers and therefore stabilize them. In contrast, in our simulations with 5Å crowders, the hexamers are preferentially formed compared to dimers, trimers, tetramers, and pentamers. We believe that with 5Å crowders, the formation of these larger oligomeric species is energetically favorable because oligomers allow peptides to take on a conformation that occupies less space than an equivalent number of isolated monomers. Since the oligomer occupies less space, a smaller number of crowders need to be displaced to make room for the structure, making it favorable.

## **2.2 Methods**

### **2.2.1 Discontinuous Molecular Dynamics**

The simulation method used in this work is discontinuous molecular dynamics (DMD), a fast alternative to traditional molecular dynamics.<sup>31</sup> In DMD, the potential is a discontinuous function of the interatomic separation, e.g. hard sphere and square well potentials, and since the atoms move linearly between collisions, the only time that the velocities and positions need to be recalculated is when a discontinuity in the potential is encountered. Therefore the simulation can be advanced from collision event to collision event. The types of discontinuous events in our simulations include hard sphere events, bond events, and square-well and square-shoulder capture and dissociation events.

### 2.2.2 PRIME20 force field

Coarse graining peptide geometry by combining groups of atoms into “united atoms” and then representing them as spheres is an additional way to alleviate the timescale limitations of atomistic MD. One popular coarse-grained model used for studying peptides is a four-sphere-per-residue model in which each amino acid residue is represented by three backbone spheres, one each for N-H, C-H, and C=O, and one side-chain sphere R.<sup>32</sup> The four-sphere per residue model provides a balance between accuracy and simplicity that is ideal for DMD. In 2004, our group introduced a new four-sphere per residue protein model called PRIME (Protein Intermediate Resolution Model), which was appropriate for homoproteins like polyalanine and polyglutamine.<sup>33-35</sup> In PRIME, all backbone bond lengths and bond angles are fixed at their ideal values, the distance between consecutive C $\alpha$  atoms is fixed so as to maintain the interpeptide bond in the trans configuration, and the side chains are held in positions relative to the backbone so that all residues are L-isomers.

More recently PRIME was extended to heteroproteins culminating in PRIME20 which describes the geometry and energetics for all 20 amino acids.<sup>36</sup> The interactions in PRIME20 simulations include excluded volume, hydrogen bonds, hydrophobic interactions, and charge interactions. PRIME20 also includes polar interactions, but A $\beta$ (16-22) does not have any polar amino acids. The parameters used in PRIME20 were derived utilizing a perceptron learning algorithm and a modified stochastic learning algorithm that compared the energy of native state proteins in the Protein Database (PDB) with those of a large number of decoy structures. The parameters used for the A $\beta$ (16-22) peptide in this work are taken from Cheon *et al.*<sup>37</sup>

### 2.2.3 Crowder Model

The crowders used in these simulations are modeled as hard spheres of diameter 5Å, 20Å and 40Å. Thus the crowder-peptide and crowder-crowder interactions are limited to excluded volume. The 20Å crowder diameter was selected because it is close to the N to C-terminal length of a fully extended A $\beta$ (16-22) peptide based on PRIME20 parameters. The 5Å and 40Å crowder diameters were selected to provide contrast for the effects of crowder size on aggregation. The 5Å crowders are small compared to A $\beta$ (16-22) so they have the ability to sit close to the peptide backbone and in between the side chains, while the 40Å crowders are significantly larger than A $\beta$ (16-22) and have larger interstitial spaces for the peptides to occupy. In these simulations the crowder volume fractions examined are  $\phi=0.00$ , 0.10, and 0.20; the crowder volume fraction is defined as  $\phi=N V_0/L^3$  where N is the number of crowders,  $V_0$  is the volume of a single crowder, and L is the simulation box length. These crowder volume fractions correspond to system densities of  $\rho=4\text{mg/mL}$ , 177mg/mL, and 350mg/mL respectively. The system density of 350mg/mL was chosen as a realistic reference for this work because it is the midpoint between the estimated densities of 300 mg/mL and 400mg/mL in cytoplasm and is commonly used as the density for crowding experiments. In order to calculate the internal density of a single crowder molecule necessary to achieve a total system density of 350 mg/mL we selected, a system with a crowder volume fraction of  $\phi=0.20$  and crowder diameter of 40Å as our reference because this seemed closest to physiologically relevant conditions. To make this system have a density of 350mg/mL, the internal density of a single 40 Å crowder was set at 1.04 Da/Å<sup>3</sup>. It follows then that the mass of the 40Å crowder is 34.9 kDa while the masses of the 20Å and 5Å crowders are 4.4 kDa

and 68 Da, respectively. These system parameters were selected based on the recommendations of Ellis and Minton, and Zimmerman and Trach for studying macromolecular crowding without exceeding our current computational limitations.<sup>7,14</sup>

#### **2.2.4 Simulation Procedure**

Our simulations proceed in the following way: Each simulation contains 192 A $\beta$ (16-22) peptides initially placed at random locations in a cubic simulation box with side lengths of  $L=400\text{\AA}$ , giving a peptide concentration of 5mM, and periodic boundary conditions. The crowders are then placed at random locations surrounding the peptides until the desired crowder volume fraction is achieved. The reduced temperature is defined as  $T^*=kT/\epsilon_{HB}$  where  $\epsilon_{HB}$  is the hydrogen bonding well depth. Velocities for each peptide bead and crowder are chosen at random from a Maxwell-Boltzmann distribution that is centered at the desired temperature. Initially, the temperature is set to  $T^*=0.50$ , a temperature high enough to denature the peptides so as to give them a random coil secondary structure. The system is then gradually cooled stepwise to  $T^*=0.193$  using a cooling scheme that lasts a total 14 billion collisions. The reduced temperature  $T^*=0.193$  was chosen because it is the transition temperature above which fibrillization does not occur for A $\beta$ (16-22). The simulations were performed in the canonical ensemble where the number of particles, temperature, and volume are fixed. The temperature is held constant using the Andersen thermostat method. Beads in the simulation experience random “ghost collisions” with “ghost particles” during which their velocity is reassigned to a random value from a Maxwell-Boltzmann distribution centered at the desired simulation temperature. Up to 15 independent simulations were run

for each set of conditions until the number of peptides in oligomers had decayed to 1/e, or 36.7%, of its maximum value. This amounts to ~45 billion collisions for the most crowded simulations because aggregation happens more rapidly, and ~165 billion collisions in the absence of crowders. In our simulations a peptide is defined as being part of an oligomer if it shares at least one hydrogen bond or one hydrophobic contact with another peptide in the oligomer. A peptide is defined as being part of a  $\beta$ -sheet if it shares at least 4 hydrogen bonds with another peptide in the  $\beta$ -sheet. A peptide is defined as being part of a fibril if it is in a  $\beta$ -sheet that shares at least 4 side chain interactions with another  $\beta$ -sheet in the fibril.

### 2.2.5 Oligomerization curve fit and parameters

Due to the transient nature of oligomers and the small number of peptides that are in an oligomer at any one time, oligomerization data sets can have large variations. To make the trends easier to see we have fit our data for the number of peptides in oligomers vs reduced time with an asymmetric double sigmoidal function.<sup>38</sup> Reduced time is defined as  $t^*=t/\sigma(k_bT/m)^{1/2}$ , where  $\sigma$  and  $m$  are the average bead diameter and average mass. The asymmetric double sigmoidal function is typically used to fit chromatography data because it can account for a sudden increase followed by a gradual decrease in a data set. We chose it because it nicely fits all our oligomerization data sets; to our knowledge there is no well-established equation for fitting oligomerization data sets vs time. The asymmetric double sigmoidal function is expressed in Equation 2.1

$$N_{olig}(t^*) = N_0 + A \left( \frac{1}{1 + \exp\left(\frac{t^* - t_{center} + w_1/2}{w_2}\right)} \right) \left( 1 - \frac{1}{1 + \exp\left(\frac{t^* - t_{center} + w_1/2}{w_3}\right)} \right) \quad (2.1)$$

where  $N_{\text{olig}}(t^*)$  is the number of peptides in an oligomer at time  $t^*$ ,  $N_0$  is the initial number of peptides in an oligomer,  $A$  is the amplitude,  $t_{\text{center}}$  is the center of the peak, and  $w_1$ ,  $w_2$ , and  $w_3$  are widths for the sigmoidal curves. Using this curve fit makes it straightforward to calculate the oligomer growth rate, the maximum number of peptides in oligomers, and the rate of conversion from oligomers to fibrils. To calculate the oligomer growth rate, we find the most linear portion of the curve between  $t^*=0$  and the time at which the number of peptides in oligomers reaches its maximum,  $t_{\text{peak}}$ , by calculating the  $R^2$  value. We then calculate the slope using the most linear portion of the curve. The maximum number of peptides in oligomers is taken as the maximum in the asymmetric double sigmoidal function, and the oligomer to fibril conversion time is calculated as the amount of time it takes for the number of oligomers to decay from its maximum value to  $1/e$ , or 36.7%, of its peak value. The average  $R^2$  value for our curve fitting of the number of peptides in oligomers vs reduced time with the asymmetric double sigmoidal function was 0.92. The average  $R^2$  value for our curve fitting of the number of peptides in dimers, trimers, tetramers, pentamers, and hexamers vs reduced time with the asymmetric double sigmoidal function was 0.83.

### **2.2.6 Small oligomer free energy analysis**

To understand if dimers, trimers, tetramers, pentamers, or hexamers are energetically favorable in the presence of crowders we analyze the change in free energy associated with forming small oligomers from monomers using a method introduced by O'Brien *et al.*<sup>22</sup> Using the Asakura and Oosawa theory for two bodies immersed in a solution of macromolecules along with scaled particle theory, O'Brien *et al.* found that  $\Delta G_{ij}(n)$ , the change in the Gibbs

free energy associated with transitioning from peptide structure j with n peptides to a different type of peptide structure i, also with n peptides is given by Equation 2.2.

$$\Delta G_{ij}(n) = \frac{k_b T \phi}{V_c} (V_{ex}^i(n) - V_{ex}^j(n)) \quad (2.2)$$

In Equation 2.2,  $V_{ex}^i(n)$  and  $V_{ex}^j(n)$  are the volumes excluded to the crowders by the peptide structures i and j,  $\phi$  is the crowder volume fraction, and  $V_c$  is the volume of a single crowder.<sup>39</sup> If  $\Delta G_{ij}(n)$  is negative, then species i is more energetically favorable than species j and it is more likely that species i occupies a small enough volume to fit within the interstitial spaces of the crowders. If  $\Delta G_{ij}(n)$  is positive, then species j is more energetically favorable than species i and it is more likely that species j occupies a small enough volume to fit within the interstitial spaces of the crowders. O'Brien *et al.* also developed expressions for the volume excluded to the crowders by monomers,  $V_{ex}^M(n)$ , disordered aggregates,  $V_{ex}^D(n)$ , and beta sheets,  $V_{ex}^\beta(n)$  of size n. These are given in Equations 2.3, 2.4, and 2.5 respectively.

$$V_{ex}^M(n) = \frac{4\pi n}{3} \left[ (R_g + R_c)^3 - R_g^3 \right] \quad (2.3)$$

$$V_{ex}^D(n) = \frac{4\pi}{3} \left[ (nR_g + R_c)^3 - (nR_g)^3 \right] \quad (2.4)$$

$$V_{ex}^\beta(n) = (l + 2R_c)(w + 2R_c)(h + 2R_c) - lwh - \left( 8R_c^3 - \frac{4}{3}\pi R_c^3 \right) \quad (2.5)$$

The volume excluded to crowders by monomers in Equation 2.3 is calculated by using the crowder radius  $R_c$  and approximating the peptide as a sphere with the peptide's radius of gyration  $R_g$ . In Equation 2.4 the disordered aggregate is approximated as a spherical globule

made of  $n$  peptides resulting in a larger sphere of radius  $nRg$ . In Equation 2.5 the  $\beta$ -sheet is approximated as a series of  $n$  connected rectangular parallelepipeds, each with length  $l$ , width  $w$ , and height  $h$ . For our calculations we take these to be the average dimensions of a single A $\beta$ (16-22) peptide in a  $\beta$ -sheet measured in our simulations to have length  $\sim 20\text{\AA}$ , width  $\sim 7.1\text{\AA}$ , and height  $\sim 4\text{\AA}$ .

### 2.2.7 Fibrillization curve fit and parameters

In order to characterize the formation of fibrils, the number of peptides in fibrils at time  $t$ ,  $N_{\text{fibril}}(t^*)$ , vs reduced time was fit to Equations 2.6-2.11. These equations were derived by Cohen *et al.* using fixed point analysis to model fibrillization (in the absence of crowders) in terms of the microscopic processes of primary fibril nucleation, fibril elongation, and secondary nucleation.<sup>40</sup> By fitting our data on the number of peptides in fibrils as a function of time given an initial peptide concentration,  $c_0$ , to these six equations we are able to extract the rate constants  $k_n$ , the primary fibril nucleation rate which describes the formation of a fibril from solution,  $k_+$  the fibril elongation rate which describes the addition of monomers to existing fibrils, and  $k_2$  the secondary fibril nucleation rate which describes the formation of secondary fibrillar structures, in our case the addition of  $\beta$ -sheets, to the fibril.

$$\frac{N_{\text{fibril}}(t^*)}{192} = 1 - \left( \frac{B_+ + A_+}{B_+ + A_+ e^{\kappa t^*}} \right) \left( \frac{B_- + A_+ e^{\kappa t^*}}{B_- + A_+} \right)^{\frac{g_{\infty}^2}{\kappa h_{\infty}}} e^{-k_{\infty} t^*} \quad (2.6)$$

$$\kappa = \left( 2k_+ k_2 c_0^3 \right)^{\frac{1}{2}} \quad (2.7)$$

$$A_{\pm} = \pm \frac{k_n}{2k_2c_0} \quad (2.8)$$

$$B_{\pm} = \left( \frac{1}{12} + \frac{k_n}{4k_2c_0} \right)^{\frac{1}{2}} \pm \left( \frac{1}{12} + \frac{k_n}{4k_2c_0} + \frac{k_n^2}{4k_2^2c_0^2} \right)^{\frac{1}{2}} \quad (2.9)$$

$$g_{\infty} = \left( 2k_+c_0 \left[ \frac{k_2c_0^2}{3} + k_nc_0 \right] \right)^{\frac{1}{2}} \quad (2.10)$$

$$h_{\infty} = \left( 2k_+c_0 \left[ \frac{k_2c_0^2}{3} + k_nc_0 + \frac{k_n^2}{k_2} \right] \right)^{\frac{1}{2}} \quad (2.11)$$

In these equations  $\kappa$  is an effective rate constant describing the secondary aggregation pathway comprised of secondary nucleation and fibril elongation,  $A_{\pm}$ ,  $B_{\pm}$ ,  $g_{\infty}$ , and  $h_{\infty}$  are constants that are determined by the rate constants  $k_n$ ,  $k_+$ , and  $k_2$  as well as the initial peptide concentration  $c_0$  of 5mM. These equations are rearranged versions of the equations that appear in the paper by Cohen *et al.* so as to contain only the rate constants of interest and the initial monomer concentration. The average R2 value for our curve fit of the number of peptides in fibrils vs reduced time to Equations 2.6-2.11 was 0.98 We have also calculated the lag time,  $t_{lag}$ , shown in Equation 2.12, critical primary fibril nucleation rate constant,  $k_{nc}$ , shown in Equation 2.13 and the maximum fibril growth rate,  $r_{max}$ , shown in Equation 2.14.

$$t_{lag} = \frac{\log(2k_2c_0/k_n) - e + 1}{\sqrt{2k_+k_2c_0^3}} \quad (2.12)$$

$$r_{\max} = 192 \sqrt{\frac{2k_+k_2c_0^3}{e^2}} \quad (2.13)$$

$$k_{nc} = \sqrt{\frac{k_+k_2c_0}{2e^2}} \quad (2.14)$$

The lag time is the amount of time before fibrillization begins, the maximal growth rate is the fastest fibril growth rate that occurs during the simulation, and the critical primary fibril nucleation rate constant is the value for the primary fibril nucleation rate constant above which there is no lag phase. Additionally we look at the nucleation time, which we define to be the first time point that has twenty consecutive non-zero values for the number of peptides in fibrils following it. Twenty consecutive time points equates to ~56 reduced time units. Here we use the term nucleation time in the place of lag time because our simulations operate above the supercritical peptide concentration and there is no lag time as calculated by Equation 2.12.

## 2.3 Results

### 2.3.1 Aggregation in the presence of crowders

Figure 2.1 shows snapshots from a simulation of 192 A $\beta$ (16-22) peptides in the absence of crowders (A-E) and in the presence of crowders (F-J) with crowder volume fraction  $\phi=0.20$  and crowder diameter 5Å (crowders have been removed for clarity). The peptide concentration is 5mM and the reduced temperature is  $T^*=0.193$ . In the simulation with no crowders the peptides have been colored so that all of the peptides in a given  $\beta$ -sheet

at the conclusion of the simulation have the same color and in the simulation with crowders the peptides have been colored so that all of the peptides in a given fibril at the conclusion of the simulation have the same color. At  $t^*=0$  (Fig. 1A and 1E) the initial peptide configuration for both simulations is random coils. In the simulation with no crowders some peptides begin to interact transiently, but do not form a stable fibril nucleus until  $t^*=400$  (Fig. 1B). The fibril nucleus is composed of two stable  $\beta$ -sheets colored green and light blue. At  $t^*=400$  the fibril has begun to elongate through monomer addition at the ends of the fibril; each  $\beta$ -sheet has lengthened, increasing the overall size of the fibril. At  $t^*=1100$  (Fig. 1C) three additional  $\beta$ -sheets have attached themselves to the fibril resulting in a much larger five sheet fibril. Finally, at  $t^*=2000$  (Fig.1D) all of the peptides in the simulation have integrated themselves into a single large fibril composed of six  $\beta$ -sheets. The simulation with crowders begins to form small oligomers very early in the simulation at  $t^*=90$  (Fig.1F). Soon after, the previously formed oligomers begin to rearrange themselves from disordered conformations into  $\beta$ -sheets and small fibrils while new oligomers are formed from the remaining free monomers. At  $t^*=220$  (Fig.1G) almost all of the free monomers have been integrated into an oligomer or fibril and they continue to reorganize from disordered conformations to  $\beta$ -sheets. Finally, at  $t^*=270$  (Fig. 1H) all of the disordered structures have reorganized into  $\beta$ -sheets resulting in six small fibrils.

By comparing the time scales for these two representative simulations, we can see that fibrillization is complete at  $t^*=2000$  for the simulation with no crowders and at  $t^*=270$  for the simulation with crowders. It is evident that the presence of crowders not only dramatically decreases the time scale of aggregation but, as can be seen from the snapshots,

the aggregation mechanisms are different. In the simulation without crowders a single long fibril forms through nucleated polymerization. In contrast, the simulation with crowders results in six small fibrils that form from smaller disordered oligomers characteristic of nucleated conformational conversion. These oligomers are initially disordered because the peptides are rapidly forced together by depletion forces. Since the peptides are forced together so quickly, they do not have the ability to orient themselves into a more favorable structure like a  $\beta$ -sheet. It is interesting to note that in the absence of fibrils, one large fibril is formed at  $t^*=2000$  but with crowding agents, six small fibrils form after only a fraction of the simulation time. It is unlikely that these six small fibrils would combine into a single large fibril if the simulation was run until  $t^*=2000$ . These small fibrils are very stable and energetically favorable, and while a few of them may combine to make a slightly larger fibril, we expect that much more computational time would be needed to observe a single large fibril forming from the smaller six. We speculate that this increase in the number of individual fibrils may be a general consequence of very crowded conditions. Support for this idea comes from simulations by Magno *et al.* who found that the number of supercritical oligomers (pre-fibrillar structures) increases with crowder concentration.<sup>23</sup>

As has been pointed out by other investigators, the dramatic effect of crowding on peptide oligomerization and fibrillization can be understood in part by appealing to the concept of depletion forces.<sup>22,41</sup> The depletion forces acting in our simulations can be approximated by adapting the expression introduced by Asakura and Oosawa to this case. The depletion potential,  $U(r)$ , between two peptides whose centers of mass are separated by distance  $r$  in the presence of crowders is<sup>42</sup>

$$U(r) = -\frac{\phi}{16R_c^3} k_B T \left[ 2(R_c + R_g)^3 - 3(R_c + R_g)^2 r - r^3 \right] \quad (2.15)$$

where  $\phi$  is the crowder volume fraction,  $R_c$  is the crowder radius,  $R_g$  is the radius of gyration of the peptide. This equation is only valid for  $R_c < r < R_g + R_c$ , and, although it is typically applied when  $R_c \ll R_g$ , provides a reasonable qualitative comparison of depletion forces for different crowder sizes. Since all of our simulations are run at the same temperature, simulation box volume, and peptide radius of gyration, the depletion potential will only change when the crowder volume fraction or the crowder radius change. As the crowder volume fraction increases, the depletion potential increases. As the crowder radius increases (at constant crowder volume fraction), the strength of the depletion potential decreases, although the range increases. The increased strength of the depletion potential for smaller crowders is what drives oligomerization, and by extension fibrillization, to occur at an accelerated rate compared to simulations in the absence of crowders.

### 2.3.2 Effects of crowding on oligomerization

Oligomerization is the first step in the aggregation process for A $\beta$ (16-22). Figure 2.2 shows the number of peptides in oligomers vs reduced time for crowder volume fractions  $\phi=0.00$  (no crowders), 0.10, and 0.20 at crowder diameters (A) 5Å, (B) 20Å, and (C) 40Å. The number of peptides in oligomers increases sharply as peptides begin to interact and then gradually decreases as the oligomers are converted to fibrils over time.

We begin our analysis of the oligomerization data by examining the rate of oligomer growth. Table 2.1 summarizes the results for the oligomer growth rate, maximum number of

peptides in oligomers, and oligomer to fibril conversion time at crowder volume fractions  $\phi=0.00$ , 0.10, and 0.20 and crowder diameters 5Å, 20Å, and 40Å. A useful measure of crowding conditions, in addition to crowder volume fraction and diameter, is the total crowder surface area, listed below the corresponding crowder volume fraction and crowder diameter in the table. Total crowder surface area is simply the surface area of a single crowder multiplied by the number of crowdors in the simulation. The first row of Table 2.1 shows the oligomer growth rate. At a constant volume fraction decreasing the diameter of the crowdors increases the oligomer growth rate. If crowder diameter is held constant, increasing the crowder volume fraction increases oligomer growth rate. The table shows that as the total crowder surface area increases, the growth rate increases monotonically. The highest growth rate is 3.21 peptides added per unit time for 5Å crowdors at a crowder volume fraction of  $\phi=0.20$  having a total crowder surface area of  $0.1536\text{Å}^2$ . This growth rate is more than five times greater than that in the absence of crowdors. The increase in oligomer growth rate with crowder volume fraction can be attributed to an increase in the effective concentration of the peptides. When crowdors are added to the simulation they exclude volume to the peptides making a large portion of the system inaccessible. The increase in oligomer growth rate when decreasing the crowder diameter can be attributed to the increase in depletion forces. As the size of the crowdors decreases at a fixed volume fraction, the magnitude of the depletion forces increases, further enhancing peptide-peptide interactions. Thus as either the crowder volume fraction increases or crowder diameter decreases, the peptides have a higher propensity to associate, leading to a higher oligomer growth rate.

At some point during oligomerization a fibril nucleates and the number of peptides in oligomers begins to decline as they are integrated into fibrils. This transition occurs when the number of peptides in oligomers reaches its maximum value. The second row of Table 2.1 shows the maximum number of peptides in oligomers. At a constant volume fraction, decreasing the diameter of the crowders increases the maximum number of peptides in oligomers. If crowder diameter is held constant, increasing the crowder volume fraction increases the maximum number of peptides in oligomers. Just like the oligomer growth rate, there is a monotonically increasing relationship between the maximum number of peptides in oligomers and the total crowder surface area. At the highest total crowder surface area, which corresponds to 5Å crowders and crowder volume fraction  $\phi=0.20$ , the maximum number of peptides in oligomers is 66.55, which is almost triple the amount in the absence of crowders. Similar to the oligomer growth rate, increasing the crowder volume fraction and decreasing crowder size increases the maximum number of peptides in oligomers. Once the peptides form an oligomer they are in a more energetically favorable state because they occupy less space than the equivalent number of free monomers. For this reason the peptides do not dissociate, leading to a higher maximum number of peptides in oligomers.

After the number of peptides in oligomers peaks, the oligomers begin to convert to fibrils. We characterize this process by calculating the oligomer to fibril conversion time. The third row of Table 2.1 shows the oligomer to fibril conversion time. At a constant crowder volume fraction, decreasing the diameter of the crowders decreases the oligomer to fibril conversion time. If crowder diameter is held constant, increasing the crowder volume fraction decreases the oligomer to fibril conversion time. As the total crowder surface area

increases, the oligomer to fibril conversion time decreases exponentially to its lowest value: 95.92 time units for 5Å crowders at a crowder volume fraction of  $\phi=0.20$ . Since fibrils occupy even less space than an equivalent number of peptides in an oligomer, they are more energetically favorable, providing a more thermodynamically stable structure than an oligomer. Larger depletion forces and a higher effective concentration make the oligomer to fibril conversion time shorter.

Summarizing thus far, the presence of crowders tends to increase oligomer formation. Increasing crowder volume fraction and decreasing crowder diameter increases the growth rate of oligomers and the maximum number of peptides in oligomers but decreases the oligomer to fibril conversion time. Although crowders promote rapid oligomerization early in the simulation, they also drive oligomers to form fibrils at a faster rate.

### **2.3.3 Effects of crowding on small oligomer formation**

Since small oligomers have been identified as toxic agents in Alzheimer's Disease, the dependence of the number of peptides in dimers, trimers, tetramers, pentamers, and hexamers on crowder diameter and crowder volume fraction is of interest. Here we focus on 5Å and 40Å crowders to see how the smallest and largest crowders impact small oligomer formation.

The oligomerization mechanisms observed in our simulations can be described most simply in terms of a step-growth mechanism. In a step growth mechanism monomers come together one by one to first form dimers, then trimers, then tetramers etc. which means that the maximum number of peptides in dimers is greater than the maximum number of peptides

in trimers and so on. In other words the smaller oligomers need to be formed before additional peptides can be added.

To begin our analysis we plot the curve fits to the number of peptides in dimers, trimers, tetramers, pentamers, and hexamers from our simulations vs reduced time using the asymmetric double sigmoidal function described in the Methods section. Figure 2.3 shows the number of peptides in dimers through hexamers vs reduced time for simulations with no crowders. The same data is shown in Figure 2.4 for  $\phi=0.10$  and in Figure 2.5 for  $\phi=0.20$ , with crowders of diameter 40Å (A), 20Å (B), and 5Å (C). Table 2.2 shows a summary of the fraction of oligomeric peptides that are dimers and hexamers at crowder volume fractions  $\phi=0.00$ , 0.10, and 0.20 and crowder diameters 5Å, 20Å, and 40Å. In the first row of Table 2.2, 40Å crowders have a larger fraction of oligomeric peptides in dimers than for 20Å and 5Å crowders. However, simulations with no crowders have an even higher fraction of oligomeric peptides in dimers than simulations with 40Å crowders. This indicates that the addition of crowders to a no-crowder simulation decreases the dimer content regardless of the properties of the crowders. It should be noted that for a given crowder volume fraction, larger crowders will have the smallest deviation from the no-crowder results due to their smaller surface to volume ratio. When comparing results from our simulations at the same crowder volume fractions but different crowder diameters, the 40Å crowders have the highest propensity to favor dimer formation. In the second row of Table 2.2, the 5Å crowders have a higher fraction of oligomeric peptides that are hexamers than the 20Å and 40Å crowder simulations.

Here we compare our conclusions from Table 2.2 to the free energy analysis introduced by O'Brien *et al.* presented in the Methods section.<sup>22</sup> The data presented in Table 2.2 shows that the fraction of oligomeric peptides in dimers at a fixed volume fraction is highest for simulations with 40Å crowders, and the fraction of oligomeric peptides in hexamers is highest for simulations with 5Å crowders. In order to interpret the change in free energy associated with transitioning from a monomer to an oligomer in the presence of 40Å crowders we will use Equation 2.3 for monomers with Equation 2.4 for disordered aggregates rather than Equation 2.5 for beta sheets because the peptides are typically in a more compact state rather than an extended conformation when confined in the space between 40Å crowders. Figure 2.6 shows a plot of the difference in free energy between a disordered aggregate of size  $n$  and  $n$  free monomers  $\Delta G_{DM}$  vs the number of peptides  $n$  for crowder diameter 40Å predicted by Equations 2.2-2.4 at crowder volume fractions  $\phi=0.00, 0.05, 0.10, 0.15,$  and  $0.20$ . We have included additional values for the crowder volume fraction beyond what we simulated to provide clarity. The negative values of  $\Delta G_{DM}$  for  $n = 2, 3,$  and  $4$  peptides and positive values for  $n = 5$  and  $6$  peptides for the 40Å crowders case suggests that the formation of disordered dimers, trimers, and sometimes tetramers is energetically favorable compared to the formation of larger pentamers and hexamers when compared to an equivalent number of free monomers.

Our result that 40Å crowders favor the formation of dimers, trimers and tetramers and that the 5Å crowders favor the formation of pentamers and hexamers can be understood based on the following arguments. We believe that the size of the interstitial spaces created by the 40Å crowders, (as compared to the smaller size interstitial spaces created by the 5Å

crowders) is commensurate with the sizes of dimers, trimers and tetramers and hence favors their formation. We have come to this conclusion based on the fact that negative values of  $\Delta G_{ij}(n)$  in Equation 2.2 mean that the oligomer structure is energetically favorable when surrounded by crowders and occupies a volume small enough that crowders do not need to be displaced when the oligomer is present. In addition, since the depletion forces are weakest for the 40Å crowders and the interstitial spaces are of limited size, there is little in the way of driving force to create larger oligomers. This is consistent with the following ideas which were mentioned earlier. If the oligomer occupies approximately the same volume or less than the interstitial spaces between the crowders, it is energetically favorable because the crowders do not need to move to accommodate the oligomer. If the volume of the oligomer is greater than the interstitial space it becomes energetically unfavorable because the oligomer no longer fits neatly into the space and the crowders must be moved in order to accommodate the oligomer. In the case of the smaller 5Å crowders, the interstitial spaces created are so small that no particular size oligomer is favored. However the depletion forces are quite sizeable, so that once any oligomer forms it tends to grow larger due to the large depletion forces. The latter effect can be seen in the following analysis of our simulations with 5Å crowders.

To continue our free energy analysis we now examine the case of 5Å crowders. In order to interpret the change in free energy associated with the transition from a monomer to an oligomer in the presence of 5Å crowders we will use Equation 2.3 for monomers with Equation 2.5 for  $\beta$ -sheet aggregates rather than Equation 2.4 for disordered aggregates. This is because the peptides typically adopt a more extended conformation when surrounded by

5Å crowders than by the larger crowders since the small crowders can sit closer to the peptide backbone. Applying Equations 2.2, 2.3, and 2.5 we arrive at Figure 2.7 which shows the difference in free energy between a  $\beta$ -sheet of size  $n$  and  $n$  free monomers  $\Delta G_{\beta M}$  vs the number of peptides  $n$  at crowder volume fraction  $\phi=0.20$ . For each crowder diameter, increasing the number of peptides  $n$  makes  $\Delta G_{\beta M}$  more negative, favoring the formation of a  $\beta$ -sheets over free monomers. Although the formation of larger aggregates is favorable for all sizes of crowders, the favorability of  $\beta$ -sheet formation increases more rapidly as the crowder volume fraction increases for the 5Å crowders than for 20Å and 40Å crowders because the change in free energy  $\Delta G_{\beta M}$  is significantly more negative [Data not shown]. Since small crowders create a larger depletion force between peptides, the attractive force promotes the formation of aggregates in favor of free monomers.

### 2.3.4 Effects of crowding on fibrillization

We now turn our attention to the formation of fibrillar structures which occurs after the oligomers have formed. Some of the mechanisms suggested to govern fibril formation are nucleated polymerization in which fibril growth does not occur until a nucleus is formed and growth occurs via monomer addition to the fibril, or nucleated conformational conversion in which monomers rapidly aggregate into oligomers, and then convert to fibrils over time.<sup>43,44</sup> Figure 2.8 shows the number of peptides in fibrils vs reduced time for crowder volume fractions  $\phi=0.00$  (no crowders), 0.10, and 0.20 at crowder diameters (A) 5Å, (B) 20Å, and (C) 40Å. The fibrils formed in the simulations at  $\phi=0.00$ , *i.e.* no crowders, have a relatively

linear growth rate over the 1200 reduced time units in Figure 2.8. However, once crowders are added, as in the  $\phi=0.10$  and  $0.20$  simulations, rapid fibrillization occurs at earlier times.

Table 2.3 summarizes the results for the kinetic constant parameters for the fibrillization fit to the model of Cohen *et al.* as described in the methods section: the primary fibril nucleation rate constant, critical primary fibril nucleation rate constant, maximum fibril growth rate, fibril elongation rate constant, and the secondary fibril nucleation rate constant at crowder volume fractions  $\phi=0.00$ ,  $0.10$ , and  $0.20$  and crowder diameters  $5\text{\AA}$ ,  $20\text{\AA}$ , and  $40\text{\AA}$ . The first row of Table 2.3 shows the primary fibril nucleation rate constant,  $k_n$ , which describes the formation of a fibril nucleus from a solution of monomers. All of the simulations have primary fibril nucleation rate constants that fall between  $0.20\text{ M}^{-1}\text{s}^{-1}$  and  $0.25\text{ M}^{-1}\text{s}^{-1}$ , except for the simulations with a crowder volume fraction of  $\phi=0.20$  with  $5\text{\AA}$  crowders, which has a primary fibril nucleation rate constant of  $0.46\text{ M}^{-1}\text{s}^{-1}$ . The highest rate of primary fibril nucleation occurs for the highest crowder volume fraction with the smallest crowders. This indicates that the large depletion forces at these conditions force nucleation to occur much more rapidly, causing multiple small fibrils to form rather than a single large fibril as shown in Figure 2.1.

The second row of Table 2.3 shows the critical primary fibril nucleation rate constants,  $k_{nc}$ , which can be compared to the primary fibril nucleation rate constant. The critical primary fibril nucleation rate constant is the value for the primary fibril nucleation rate constant above which there is no lag phase. In every case the primary fibril nucleation rate constant is greater than the critical value confirming that there is no lag phase in our simulations. Our hypothesis as to why there is no lag phase in our simulations is that there is

minimal, if any, monomer dissociation from oligomers and fibrils, indicating that both structures are more energetically favorable than a free monomer and that we are operating above the supercritical peptide concentration as described by Powers and Powers.<sup>45</sup> The maximum fibril growth rate,  $r_{\max}$ , for each type of simulation is shown in row three of Table 2.3. The maximum fibril growth rate has a monotonically increasing trend with total crowder surface area. Although the primary nucleation rate is approximately the same for all simulations, except at  $D=5\text{\AA}$  and  $\phi=0.20$ , we can see that the maximum growth rate does in fact increase with increasing crowder surface area. This trend indicates that although primary nucleation occurs at approximately the same rate for all conditions, the growth rate directly following nucleation increases with increasing crowder surface area.

Next we look at the fibril elongation rate,  $k_+$ , which describes the rate of fibril growth through monomer addition to the ends of the fibril and the rate of secondary fibril nucleation,  $k_2$ , which in our simulations is the rate of addition of a new  $\beta$ -sheet to a fibril. The fourth and fifth rows of Table 2.3 show the fibril elongation rate constant and secondary fibril nucleation. Smaller crowder diameters and larger crowder volume fractions increase both the fibril elongation rate and secondary nucleation rate. As the total surface area of the crowders increases, the fibril elongation rate constant and secondary nucleation rate also increase monotonically. The exception to the monotonically increasing trend is the  $D = 5\text{\AA}$   $\phi=0.20$  case. If a linear trend were to apply to these simulations we would expect an elongation rate constant of  $\sim 25\text{ M}^{-1}\text{s}^{-1}$  and a secondary nucleation rate constant of  $\sim 10\text{ M}^{-2}\text{s}^{-1}$ , but instead they are  $11.38\text{ M}^{-1}\text{s}^{-1}$  and  $4.87\text{ M}^{-2}\text{s}^{-1}$  respectively. We are unsure why this deviation occurred, but it is possible that it is because we need a more comprehensive equation to

describe fibrillization that includes mechanisms beyond primary nucleation, secondary nucleation, and elongation.

The fastest fibril growth rate in our simulations occurs at crowder volume fraction of  $\phi=0.20$  and diameter  $5\text{\AA}$  where primary and secondary nucleation are very high. We attribute the high rates of primary and secondary nucleation to the very high depletion forces that occur at these conditions. A high rate of primary and secondary nucleation should lead to formation of a large number of fibrils made of many  $\beta$ -sheets. This behavior is indicative of nucleated conformational conversion in which the peptides rapidly form disordered oligomers, and then reorganize over time to form fibrils. This is consistent with the snapshots of our simulations in Figure 2.1H for a crowder volume fraction of  $\phi=0.20$  and  $5\text{\AA}$  crowders. As crowder volume fraction decreases and the size of the crowders increase, the depletion forces become less prominent and the aggregation mechanism begins to shift toward slow ordered fibril growth characteristic of nucleated polymerization. Nucleated polymerization would have lower values of primary and secondary nucleation relative to fibril elongation because nucleation occurs much less often than in simulations with higher depletion forces. If nucleation occurs at a lower rate and fibril elongation dominates we would expect a smaller number of longer fibrils and that is exactly what we saw in our simulations in Figure 2.1E.

Nucleation time is the amount of time it takes until the first fibril begins to form. We define the nucleation time as the reduced time at which there are twenty consecutive non-zero values for the number of peptides in fibrils. Without crowders, fibril nucleation occurs after 38.0 reduced time steps. The fibril nucleation time shows no particular trend with increasing crowder volume fraction or crowder diameter for  $20\text{\AA}$  and  $40\text{\AA}$  crowders, varying

between 31.5 and 41.2 reduced time steps for 20Å crowders and 34.7 and 44.4 reduced time steps for 40Å crowders [Data not shown]. This indicates that the presence of 20Å or 40Å crowders does not provide enough excluded volume to force peptides down an aggregation pathway consistently and that the nucleation time is random and most likely dependent on the initial spatial distribution of peptides. We did not observe the decrease in nucleation time at high crowder volume fractions mentioned in the introduction. Our explanation for why we did not observe this behavior is that our peptide does not need to fold in order to be a part of a fibril. High crowder volume fractions and small crowders could prevent the peptides from adopting the proper conformation to be integrated into a fibril, but in our simulations there is no folding because A $\beta$ (16-22) is only 7 residues long, so this effect is not present. Additionally, the crowder volume fractions we studied may not be large enough to see the delay in nucleation time observed by others.

## 2.4 Discussion and Conclusions

Using the combination of DMD and our PRIME20 force field we have been able to simulate systems of coarse grained proteins that have realistic geometry and energetic parameters along with crowding spheres up to realistic volume fractions. Although previous studies have been performed on similar systems, we are not aware of any that match the scale and realism of the species involved in the simulations. The systems contained 192 A $\beta$ (16-22) peptides and crowders of diameters 5Å, 20Å, and 40Å, represented here by simple hard spheres, at crowder volume fractions of  $\phi=0.00$ , 0.10, and 0.20. Our results show that both crowder volume fraction and size have a large impact on fibril and oligomer formation. The

addition of crowders to a simulation without crowders increases the rate of oligomer formation and the peak number of oligomers that form. As the crowder volume fraction increases or the crowder diameter decreases, the increase in oligomer formation is accompanied by a shift from a slow ordered formation of oligomers, similar to nucleated polymerization, to a fast collapse and subsequent rearrangement that leads to the high maximum number of peptides in oligomers as is characteristic of nucleated conformational conversion. The rate of conversion from oligomers to fibrils also increases giving rise to an increased rate of fibril growth. Based on our analysis it appears there is not an abrupt transition from nucleated polymerization to nucleated conformational conversion while increasing crowder volume fraction or decreasing crowder size, rather the mechanism governing fibrillization changes gradually with the simulation conditions. In all cases, larger volume fractions and smaller crowders provide the largest enhancement of oligomerization and fibrillization. These results agree with those of O'Brien *et al.* in that adding crowders of any size or concentration to the simulation will enhance aggregation and as the size of the crowders increases the level of aggregation enhancement is diminished.<sup>22</sup> Although crowding is also expected to impact oligomerization and fibrillization through changes in peptide diffusion and viscosity, we have not analyzed those effects here.

We have also presented a free energy analysis of the formation of dimers, trimers, tetramers, pentamers, and hexamers in the presence of crowders. In our simulations the 40Å crowders have interstitial spaces that are large enough to easily accommodate the dimers and therefore stabilize these oligomers allowing them to persist longer relative to trimers, tetramers, pentamers, and hexamers when compared to systems of equivalent crowder

volume fractions but different crowder diameters. The depletion forces from the 5Å crowders are so great that the largest oligomers, in our case hexamers, are the most energetically favorable. Our analysis showed that in the presence of crowders it is possible for specific oligomers to be more energetically favorable than free monomers because they allow the peptides to adopt more compact conformations. This idea agrees with Munishkina *et al.* and their idea that in the presence of crowders, specific oligomer and fibril aggregation pathways are preferred because of the favorability of specific peptide structures and the fact that they may be more energetically stable than others.<sup>26</sup>

Since A $\beta$  fibrillization is thought to be a reaction-limited process, crowding should increase aggregation and that trend was observed.<sup>17</sup> One trend we did not observe in our simulations is the increase in fibrillization lag time associated with very high crowder volume fractions. We surmise that we did not observe this behavior because our peptide does not need to fold in order to be a part of a fibril. High crowder volume fractions and small crowders could prevent the peptides from adopting the proper conformation to be integrated into a fibril, but in our simulations there is no folding because A $\beta$ (16-22) is only 7 residues long, so this effect is not present. Additionally, the crowder volume fractions we studied may not be large enough to see the delay in nucleation time predicted by Zimmerman and Minton and observed in experiment by Cabaleiro-Lago *et al.* and observed in simulation by Co *et al.*<sup>8,24,25</sup>

Although the combination of DMD and our intermediate resolution protein model, PRIME20 has allowed us to simulate the aggregation of a large number of peptides up to physiologically relevant conditions, there are some inherent limitations to our approach. Since

the peptide studied is very short, only 7 residues, we are not able to get a picture of how the competition between folding and aggregation changes in the presence of crowders. In the future we hope to examine a longer protein sequence to focus on the effects of protein folding in addition to aggregation. Although we are unable to include hydrodynamic interactions in our DMD simulations, we believe that their inclusion would likely enhance the rate of oligomer and fibril formation beyond what we reported since long range hydrodynamic interactions typically reduce protein diffusion. In addition, since the peptide we are considering does not fold, intrapeptide hydrodynamic interactions would not come into play. Additionally, the model for the crowders that we have used here only takes crowder volume exclusion into account and does not capture the effects of nonspecific attractive interactions that may exist between proteins and crowders. A more detailed model might include these interactions to address how they might change the influence of crowding on aggregation. Finally, more complex crowder geometry might be necessary to increase the accuracy of our simulations. We have limited our study to spherical crowders, but crowders represented as sphereocylinders, polymer chains, or coarse-grained representations of real crowding molecules might increase the relevance of our simulations. In a forthcoming study we will examine how the addition of attractive crowders to a system of peptides affects aggregation and how different types of crowder-peptide interactions change the behavior of the system. We predict that the complex interplay between enthalpic and entropic effects imparted by attractive crowders should have a much different effect on aggregation than hard-core crowders, as shown by Kim and Mittal and Sapier and Harries.<sup>46,47</sup> Strongly

attractive crowdors would likely diminish the formation of oligomers and fibrils, counteracting the aggregation enhancement due to hard-core crowdors shown in this paper.

The major conclusions in our paper are not sensitive to our definition of fibrils and oligomers. For example our definition of an oligomer requires that at least two peptides share a side chain contact or a hydrogen bond. We considered breaking this into two classes of oligomer, disordered (primarily side chain contacts between chains) or ordered ( $\beta$ -sheet structure), but in these simulations disordered oligomers are very short lived and needlessly complicate the discussion. A change in the definition of oligomer, e.g., requiring more hydrogen bonds, would simply shift the curves to a later point in time as the ultimate structures formed are the same. The values of our calculated parameters would change slightly, but the overall trends would be preserved.

## **2.5 Acknowledgements**

This work was supported by the National Institutes of Health, USA, under grant GM056766 and EB006006 (C.K.H.) and the National Creative Research Initiatives (Center for Proteome Biophysics) of National Research Foundation/Ministry of Education, Science and Technology, Korea grant 2011-0000041 (M.C.).

## 2.6 References

1. Chiti F, Dobson CM. Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem.* 2006;75:333-366. doi: 10.1146/annurev.biochem.75.101304.123901.
2. Lambert M, Barlow A, Chromy B, *et al.* Diffusible, nonfibrillar ligands derived from A beta(1-42) are potent central nervous system neurotoxins. *Proc Natl Acad Sci U S A.* 1998;95(11):6448-6453. doi: 10.1073/pnas.95.11.6448.
3. Gasser B, Saloheimo M, Rinas U, *et al.* Protein folding and conformational stress in microbial cells producing recombinant proteins: A host comparative overview. *Microbial Cell Factories.* 2008;7:11. doi: 10.1186/1475-2859-7-11.
4. Ignatova Z. Monitoring protein stability *in vivo*. *Microbial Cell Factories.* 2005;4:23. doi: 10.1186/1475-2859-4-23.
5. Ellis R. Macromolecular crowding: Obvious but underappreciated. *Trends Biochem Sci.* 2001;26(10):597-604. doi: 10.1016/S0968-0004(01)01938-7.
6. Fulton A. How crowded is the cytoplasm. *Cell.* 1982;30(2):345-347. doi: 10.1016/0092-8674(82)90231-8.
7. Zimmerman S, Trach S. Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of escherichia-coli. *J Mol Biol.* 1991;222(3):599-620. doi: 10.1016/0022-2836(91)90499-V.

8. Zimmerman S, Minton A. Macromolecular crowding - biochemical, biophysical, and physiological consequences. *Annu Rev Biophys Biomol Struct.* 1993;22:27-65. doi: 10.1146/annurev.bb.22.060193.000331.
9. Hatters D, Minton A, Howlett G. Macromolecular crowding accelerates amyloid formation by human apolipoprotein C-II. *J Biol Chem.* 2002;277(10):7824-7830. doi: 10.1074/jbc.M110429200.
10. Uversky V, Cooper E, Bower K, Li J, Fink A. Accelerated alpha-synuclein fibrillation in crowded milieu. *FEBS Lett.* 2002;515(1-3):99-103. doi: 10.1016/S0014-5793(02)02446-8.
11. Fung J, Darabie A, McLaurin J. Contribution of simple saccharides to the stabilization of amyloid structure. *Biochem Biophys Res Commun.* 2005;328(4):1067-1072. doi: 10.1016/j.bbrc.2005.01.068.
12. Sukenik S, Politi R, Ziserman L, Danino D, Friedler A, Harries D. Crowding alone cannot account for cosolute effect on amyloid aggregation. *Plos One.* 2011;6(1):e15608. doi: 10.1371/journal.pone.0015608.
13. Sukenik S, Harries D. Insights into the disparate action of osmolytes and macromolecular crowders on amyloid formation. *Prion.* 2012;6(1):26-31. doi: 10.4161/pri.6.1.18132.
14. Ellis R, Minton A. Protein aggregation in crowded environments. *Biol Chem.* 2006;387(5):485-497. doi: 10.1515/BC.2006.064.

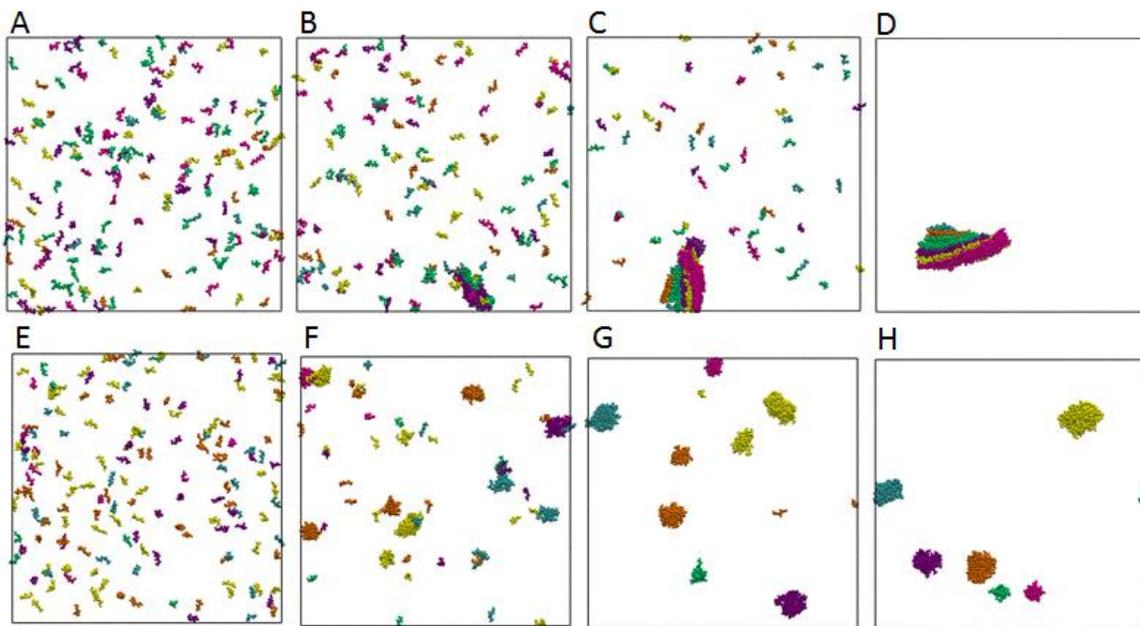
15. Minton A. Models for excluded volume interaction between an unfolded protein and rigid macromolecular cosolutes: Macromolecular crowding and protein stability revisited. *Biophys J.* 2005;88(2):971-985. doi: 10.1529/biophysj.104.050351.
16. Zhou H, Rivas G, Minton AP. Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences. *Annual Review of Biophysics.* 2008;37:375-397. doi: 10.1146/annurev.biophys.37.032807.125817.
17. Nguyen PH, Li MS, Stock G, Straub JE, Thirumalai D. Monomer adds to preformed structured oligomers of A beta-peptides by a two-stage dock-lock mechanism. *Proc Natl Acad Sci U S A.* 2007;104(1):111-116. doi: 10.1073/pnas.0607440104.
18. Kinjo A, Takada S. Competition between protein folding and aggregation with molecular chaperones in crowded solutions: Insight from mesoscopic simulations. *Biophys J.* 2003;85(6):3521-3531. doi: 10.1016/S0006-3495(03)74772-9.
19. Kinjo AR, Takada S. Effects of macromolecular crowding on protein folding and aggregation studied by density functional theory: Statics. *Physical Review E.* 2002;66(3):031911. doi: 10.1103/PhysRevE.66.031911.
20. Kinjo AR, Takada S. Effects of macromolecular crowding on protein folding and aggregation studied by density functional theory: Dynamics. *Physical Review E.* 2002;66(5):051902. doi: 10.1103/PhysRevE.66.051902.
21. Minton AP. The effect of time-dependent macromolecular crowding on the kinetics of protein aggregation: A simple model for the onset of age-related neurodegenerative disease. *Frontiers in Physics.* 2014;2(48).

- <http://www.frontiersin.org/biophysics/10.3389/fphy.2014.00048/abstract>. doi: 10.3389/fphy.2014.00048.
22. O'Brien EP, Straub JE, Brooks BR, Thirumalai D. Influence of nanoparticle size and shape on oligomer formation of an amyloidogenic peptide. *Journal of Physical Chemistry Letters*. 2011;2(10):1171-1177. doi: 10.1021/jz200330k.
  23. Magno A, Caflisch A, Pellarin R. Crowding effects on amyloid aggregation kinetics. *J Phys Chem Lett*. 2010;1(20):3027-3032. doi: 10.1021/jz100967z.
  24. Nguyen Truong Co, Hu C, Li MS. Dual effect of crowders on fibrillation kinetics of polypeptide chains revealed by lattice models. *J Chem Phys*. 2013;138(18):185101. doi: 10.1063/1.4804299.
  25. Cabaleiro-Lago C, Quinlan-Pluck F, Lynch I, Dawson KA, Linse S. Dual effect of amino modified polystyrene nanoparticles on amyloid beta protein fibrillation. *Acs Chemical Neuroscience*. 2010;1(4):279-287. doi: 10.1021/cn900027u.
  26. Munishkina LA, Ahmad A, Fink AL, Uversky VN. Guiding protein aggregation with macromolecular crowding. *Biochemistry (N Y)*. 2008;47(34):8993-9006. doi: 10.1021/bi8008399.
  27. Petkova A, Yau W, Tycko R. Experimental constraints on quaternary structure in alzheimer's beta-amyloid fibrils. *Biochemistry (N Y)*. 2006;45(2):498-512. doi: 10.1021/bi051952q.
  28. Tjernberg L, Naslund J, Lindqvist F, *et al*. Arrest of beta-amyloid fibril formation by a pentapeptide ligand. *J Biol Chem*. 1996;271(15):8545-8548.

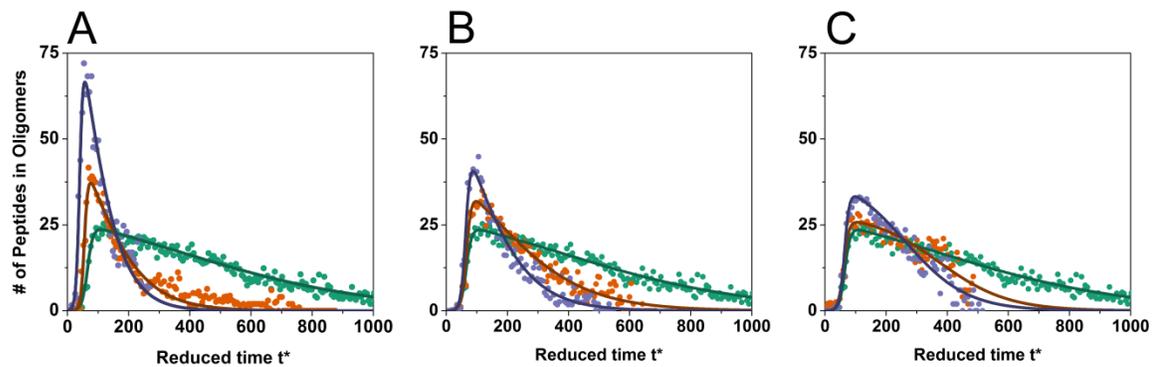
29. Balbach J, Ishii Y, Antzutkin O, *et al.* Amyloid fibril formation by A beta(16-22), a seven-residue fragment of the alzheimer's beta-amyloid peptide, and structural characterization by solid state NMR. *Biochemistry (N Y)*. 2000;39(45):13748-13759. doi: 10.1021/bi0011330.
30. Zhou H, Rivas G, Minton AP. Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences. *Annual Review of Biophysics*. 2008;37:375-397. doi: 10.1146/annurev.biophys.37.032807.125817.
31. Alder BJ, Wainwright TE. Studies in molecular dynamics .1. general method. *J Chem Phys*. 1959;31(2). doi: 10.1063/1.1730376.
32. Wu C, Shea J. Coarse-grained models for protein aggregation. *Curr Opin Struct Biol*. 2011;21(2):209-220. doi: 10.1016/j.sbi.2011.02.002.
33. Smith A, Hall C. Alpha-helix formation: Discontinuous molecular dynamics on an intermediate-resolution protein model. *Proteins-Structure Function and Genetics*. 2001;44(3):344-360.
34. Nguyen H, Hall C. Kinetics of fibril formation by polyalanine peptides. *J Biol Chem*. 2005;280(10):9074-9082. doi: 10.1074/jbc.M407338200.
35. Marchut A, Hall C. Side-chain interactions determine amyloid formation by model polyglutamine peptides in molecular dynamics simulations. *Biophys J*. 2006;90(12):4574-4584. doi: 10.1529/biophysj.105.079269.
36. Cheon M, Chang I, Hall CK. Extending the PRIME model for protein aggregation to all 20 amino acids. *Proteins*. 2010;78(14):2950-2960. doi: 10.1002/prot.22817.

37. Cheon M, Chang I, Hall CK. Spontaneous formation of twisted A beta(16-22) fibrils in large-scale molecular-dynamics simulations. *Biophys J*. 2011;101(10):2493-2501. doi: 10.1016/j.bpj.2011.08.042.
38. Origin (OriginLab N, MA). Asymmetric double sigmoidal function. <http://www.originlab.com/www/help/online/origin/en/UserGuide/Asym2Sig.html>.
39. Asakura S, Oosawa F. On interaction between 2 bodies immersed in a solution of macromolecules. *J Chem Phys*. 1954;22(7):1255-1256.
40. Cohen SIA, Vendruscolo M, Dobson CM, Knowles TPJ. Nucleated polymerization with secondary pathways. II. determination of self-consistent solutions to growth processes described by non-linear master equations. *J Chem Phys*. 2011;135(6):065106. doi: 10.1063/1.3608917.
41. Cheung M, Klimov D, Thirumalai D. Molecular crowding enhances native state stability and refolding rates of globular proteins. *Proc Natl Acad Sci U S A*. 2005;102(13):4753-4758. doi: 10.1073/pnas.0409630102.
42. Asakura S, Oosawa F. Interaction between particles suspended in solutions of macromolecules. *Journal of Polymer Science*. 1958;33(126):183-192. doi: 10.1002/pol.1958.1203312618.
43. Xue W, Homans SW, Radford SE. Systematic analysis of nucleation-dependent polymerization reveals new insights into the mechanism of amyloid self-assembly. *Proc Natl Acad Sci U S A*. 2008;105(26):8926-8931. doi: 10.1073/pnas.0711664105.

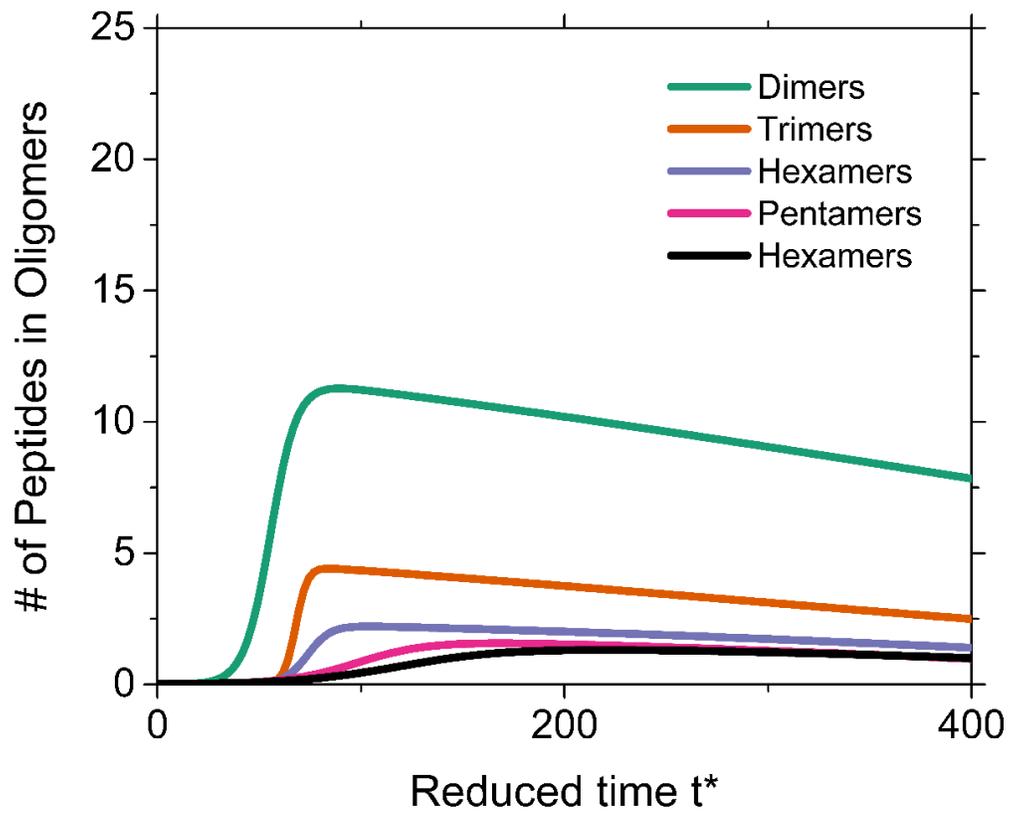
44. Lee J, Culyba EK, Powers ET, Kelly JW. Amyloid-beta forms fibrils by nucleated conformational conversion of oligomers. *Nature Chemical Biology*. 2011;7(9):602-609. doi: 10.1038/NCHEMBIO.624.
45. Powers ET, Powers DL. The kinetics of nucleated polymerizations at high concentrations: Amyloid fibril formation near and above the "supercritical concentration". *Biophys J*. 2006;91(1):122-132. doi: 10.1529/biophysj.105.073767.
46. Kim YC, Mittal J. Crowding induced entropy-enthalpy compensation in protein association equilibria. *Phys Rev Lett*. 2013;110(20):208102. doi: 10.1103/PhysRevLett.110.208102.
47. Sapir L, Harries D. Origin of enthalpic depletion forces. *Journal of Physical Chemistry Letters*. 2014;5(7):1061-1065. doi: 10.1021/jz5002715.



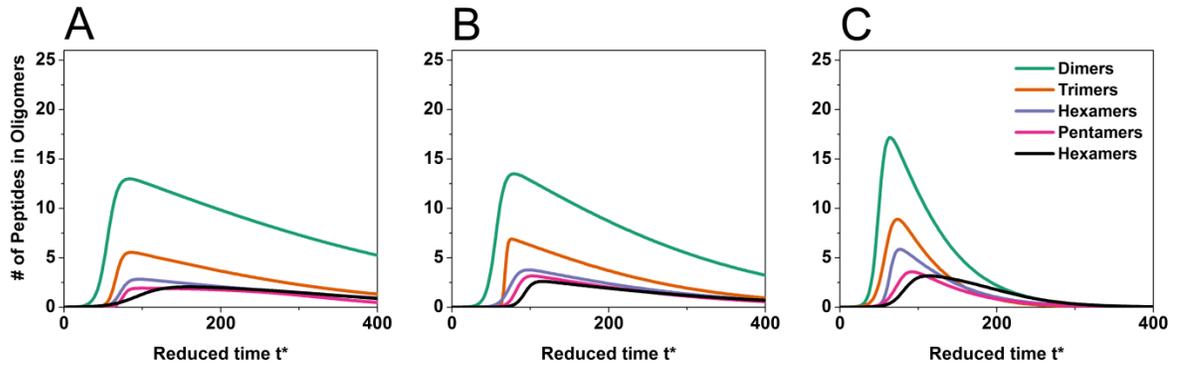
**Figure 2.1.** Snapshots of simulation progress for simulations. Top row: no crowders at  $t^* =$  (A) 0, (B) 400, (C) 1100, and (D) 2000. Bottom row 5Å crowders at crowder volume fraction  $\phi=0.20$  at  $t^* =$  (E) 0, (F) 90, (G) 220, and (H) 270.



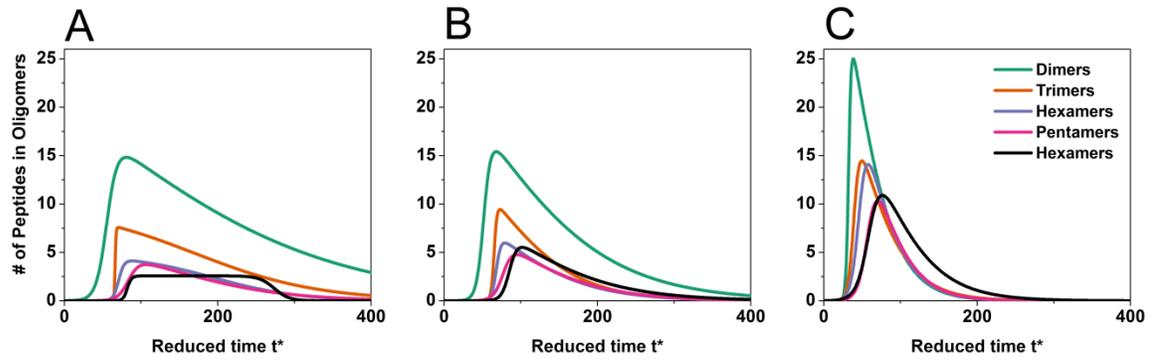
**Figure 2.2.** Number of peptides in oligomers vs. reduced time for (A) 5Å crowders (B) 20Å crowders and (C) 40Å crowders at  $\phi=0.00$  (green), 0.10 (orange), and 0.20 (purple) along with curve fits to an asymmetric double sigmoidal function.



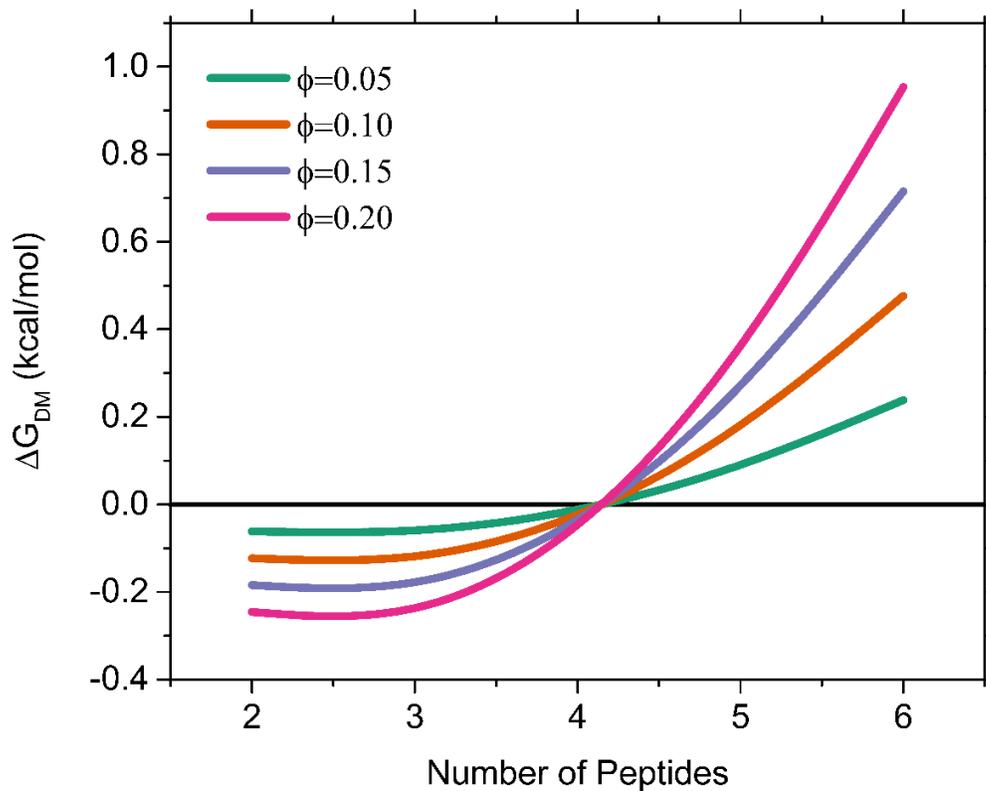
**Figure 2.3.** Number of peptides in small oligomers of different sizes vs reduced time for no crowders.



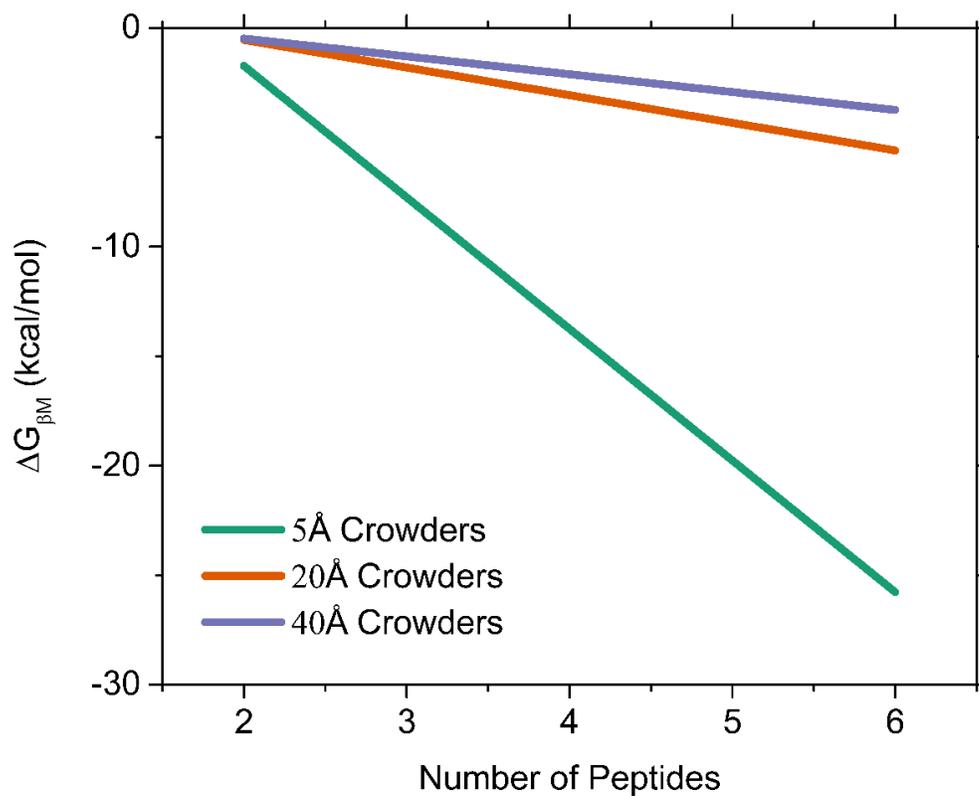
**Figure 2.4.** Number of peptides in small oligomers of different sizes vs reduced time at  $\phi=0.10$  for (A) 40Å crowders, (B) 20Å crowders, and (C) 5Å crowders.



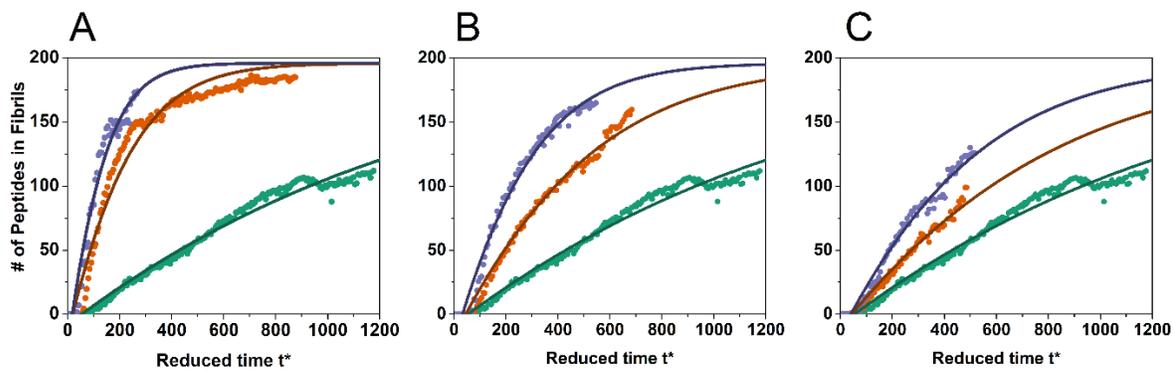
**Figure 2.5.** Number of peptides in small oligomers of different sizes vs reduced time at  $\phi=0.20$  for (A) 40Å crowders, (B) 20Å crowders, and (C) 5Å crowders.



**Figure 2.6.** Difference in free energy between a disordered aggregate and free monomers vs number of peptides for crowder volume fractions  $\phi=0.20$ , 0.15, 0.10, and 0.05 with a crowder diameter of 40Å.



**Figure 2.7.** Difference in free energy between a beta sheet and free monomers vs number of peptides at crowder volume fraction  $\phi=0.20$  for crowder diameters 5Å, 20Å, and 40Å.



**Figure 2.8.** Number of peptides in fibrils vs. reduced time for A) 5 Å crowders B) 20 Å crowders and C) 40 Å crowders at  $\phi=0.00$  (green), 0.10 (orange), and 0.20 (purple).

**Table 2.1.** Oligomer growth rate, maximum number of peptides in oligomers, and oligomer to fibril conversion time for crowder volume fractions  $\phi=0.00$ , 0.10, and 0.20 and crowder diameters 5Å, 20Å, and 40Å and total crowder surface area.

		Crowder Size, Crowder Volume Fraction, and Total Crowder Surface Area					
Kinetic parameters	No Crowders $\phi=0.00$	D=40Å $\phi=0.10$ $96 \times 10^{-4} \text{ Å}^2$	D=40Å $\phi=0.20$ $192 \times 10^{-4} \text{ Å}^2$	D=20Å $\phi=0.10$ $192 \times 10^{-4} \text{ Å}^2$	D=20Å $\phi=0.20$ $384 \times 10^{-4} \text{ Å}^2$	D=5Å $\phi=0.10$ $786 \times 10^{-4} \text{ Å}^2$	D=5Å $\phi=0.20$ $1536 \times 10^{-4} \text{ Å}^2$
Growth rate (peptides time <sup>-1</sup> )	0.61±0.30	0.74±0.42	1.06±0.68	1.00±0.54	1.43±0.37	1.48±0.47	3.21±1.06
Maximum number of peptides (peptides)	23.61±4.30	25.74±3.44	33.23±4.38	31.84±3.37	40.63±4.49	37.20±5.97	66.55±5.00
Conversion time (time)	600.67±187.58	346.58±67.78	249.89±11.81	250.08±42.67	145.44±31.46	137.84±40.74	95.92±12.80

**Table 2.2.** The maximum fraction of oligomeric peptides in dimers and in hexamers for crowder volume fractions  $\phi=0.00$ , 0.10, and 0.20 and crowder diameters 5Å, 20Å, and 40Å and total crowder surface area.

Crowder Size, Crowder Volume Fraction, and Total Crowder Surface Area							
Small Oligomer Fractions	No Crowders	D=40Å $\phi=0.10$	D=40Å $\phi=0.20$	D=20Å $\phi=0.10$	D=20Å $\phi=0.20$	D=5Å $\phi=0.10$	D=5Å $\phi=0.20$
	$\phi=0.00$	$96 \times 10^{-4} \text{ \AA}^2$	$192 \times 10^{-4} \text{ \AA}^2$	$192 \times 10^{-4} \text{ \AA}^2$	$384 \times 10^{-4} \text{ \AA}^2$	$786 \times 10^{-4} \text{ \AA}^2$	$1536 \times 10^{-4} \text{ \AA}^2$
Dimer:Total Oligomer Fraction	0.54	0.52	0.47	0.44	0.39	0.44	0.33
Hexamer:Total Oligomer Fraction	0.06	0.10	0.08	0.09	0.13	0.08	0.15

**Table 2.3.** Fibril primary nucleation rate, critical nucleation rate, and maximum growth rate, elongation rate, and secondary nucleation rate for crowder volume fractions  $\phi=0.00$ , 0.10, and 0.20 and crowder diameters 5Å, 20Å, and 40Å and total crowder surface area obtained by fitting simulation data to Equations 2.6-2.11.

Kinetic Parameters	Crowder Size, Crowder Volume Fraction, and Total Crowder Surface Area						
	No Crowders	D=40Å	D=40Å	D=20Å	D=20Å	D=5Å	D=5Å
	$\phi=0.00$	$\phi=0.10$	$\phi=0.20$	$\phi=0.10$	$\phi=0.20$	$\phi=0.10$	$\phi=0.20$
		$96 \times 10^{-4} \text{ \AA}^2$	$192 \times 10^{-4} \text{ \AA}^2$	$192 \times 10^{-4} \text{ \AA}^2$	$384 \times 10^{-4} \text{ \AA}^2$	$786 \times 10^{-4} \text{ \AA}^2$	$1536 \times 10^{-4} \text{ \AA}^2$
$k_n (\text{M}^{-1} \text{time}^{-1})$	0.21±0.06	0.25±0.10	0.24±0.06	0.20±0.09	0.25±0.05	0.25±0.09	0.46±0.18
$k_{nc} (\text{M}^{-1} \text{time}^{-1})$	0.03±0.01	0.04±0.01	0.07±0.01	0.07±0.01	0.12±0.02	0.15±0.03	0.26±0.02
$r_{max} (\text{peptides time}^{-1})$	0.05±0.02	0.08±0.02	0.13±0.02	0.13±0.02	0.24±0.04	0.29±0.05	0.51±0.04
$k_1 (\text{M}^{-1} \text{time}^{-1})$	1.40±0.44	2.17±0.35	3.31±0.40	3.42±0.43	5.85±0.78	13.48±2.94	11.38±2.03
$k_2 (\text{M}^{-2} \text{time}^{-1})$	1.45±0.46	2.53±0.66	4.03±0.64	3.97±0.64	7.41±1.58	4.87±4.55	17.53±10.33

## **CHAPTER 3**

### **Effects of Hydrophobic Macromolecular Crowders on Amyloid Beta (16-22) Aggregation using Coarse-Grained Simulations**

Chapter 3 is essentially a manuscript by David C Latshaw II, and Carol K Hall in preparation.

# Effects of Hydrophobic Macromolecular Crowders on Amyloid Beta (16-22) Aggregation using Coarse-Grained Simulations

David C Latshaw II and Carol K Hall\*

Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, North Carolina

## Abstract

In Alzheimer's disease (AD), the amyloid  $\beta$  ( $A\beta$ ) peptide aggregates in the brain to form progressively larger oligomers, fibrils and plaques. The aggregation process is strongly influenced by the presence of other macromolecular species, "crowders", which can exert forces on the proteins. One very common attribute of macromolecular crowders is their hydrophobicity. The effect of hydrophobic crowders on protein aggregation was examined by using discontinuous molecular dynamics (DMD) simulations in combination with an intermediate resolution protein model, PRIME20. The systems considered contained 48  $A\beta(16-22)$  peptides and crowders of diameters 5Å, 20Å, and 40Å, represented by hard spheres or spheres with square well/shoulder interactions, at crowder volume fractions of 0.10 and 0.20. Results show that low levels of crowder hydrophobicity are capable of increasing fibrillation lag time and high levels of crowder hydrophobicity can fully prevent the formation of fibrils. The types of structures that remain at the final stages of the simulations are summarized in a global phase diagram that shows fibril, disordered oligomer, or beta sheet phases in the space spanned by crowder size and crowder hydrophobicity. In

particular, at high levels of hydrophobicity, simulations with 5Å crowders result in only disordered oligomers and simulations with 40Å crowders result in only  $\beta$ -sheets. The presence of hard sphere crowders tends to increase the anti-parallel  $\beta$ -sheet content of fibrils while hydrophobic crowders reduce it. Finally, strong hydrophobic crowders alter the secondary structure of the A $\beta$ (16-22) monomers, bending them into a shape incapable of forming ordered  $\beta$ -sheets or fibrils. These results qualitatively agree with previous theoretical and experimental work.

### 3.1 Introduction

Amyloid fibrils are the pathological hallmark of a class of serious medical disorders known collectively as the amyloidoses. The best known of these is Alzheimer's, which afflicts 5.2 million Americans and 36 million people worldwide.<sup>1, 2</sup> In Alzheimer's Disease (AD), the amyloid  $\beta$  ( $A\beta$ ) peptide aggregates in the brain to form a variety of structures, from oligomers to fibrils, that ultimately cause cognitive impairment and memory loss.<sup>3</sup> The prevalence of AD, and the toll that it takes on its victims and their families, has spurred research into its molecular underpinnings, driven by the hope that this could help us learn how to cure or prevent AD. Most of these studies are conducted *in vitro* in order to hone in on the basic biophysics of protein aggregation without the complications arising from interactions with the surrounding media. This leaves out an important part of the story, however, since the media surrounding proteins *in vivo* is quite crowded and has an impact on the aggregation process.<sup>4, 5</sup> We have been taking a computational approach to examine how the presence of other biomolecules *in vivo* affects the basic biophysics of protein aggregation.

In a previous paper on macromolecular crowding, we used molecular-level computer simulations to examine how adding non-interacting crowders to a system of model proteins affected the kinetics and thermodynamics of the oligomerization and fibrillization process.<sup>6</sup> More specifically discontinuous molecular dynamics (DMD) combined with the PRIME20 force field was applied to a system comprised of 192  $A\beta(16-22)$  peptides and hard sphere crowders of diameters  $D=5\text{\AA}$ ,  $20\text{\AA}$ , and  $40\text{\AA}$  at crowder volume fractions  $\phi=0.00$ ,  $0.10$ , and  $0.20$  starting from a random configuration. The major findings of that paper were: (1) the addition of crowders to a system of peptides increases the rates of oligomerization and

fibrillization, with higher volume fractions and smaller crowders providing the largest enhancement effects, and (2) the aggregation mechanism changes from a relatively slow nucleated polymerization mechanism to a relatively fast nucleated conformational conversion mechanism as the crowder volume fraction increases or the crowder diameter decreases. The enhancement of oligomer and fibril formation was largely due to depletion forces between peptides induced by the surrounding crowders, making oligomer and fibrils more thermodynamically favorable than having an equivalent number of free monomers. We also showed that the size of the crowders could influence the formation of specific oligomers, with the 40Å crowders enhancing the formation of smaller oligomers and the 5Å crowders enhancing the formation of larger oligomers.

In this paper we apply discontinuous molecular dynamics (DMD) and the PRIME20 forcefield to examine how the aggregation of a system containing A $\beta$ (16-22) peptides is impacted by crowders with hydrophobic characteristics. The oligomerization and fibrillization of a system containing 48 peptides in the presence of crowders with square well/shoulder interactions with diameters  $D=5\text{\AA}$ ,  $20\text{\AA}$ , and  $40\text{\AA}$  over a range of crowder volume fractions is monitored. We compare our results on the formation of disordered aggregates, ordered  $\beta$ -sheet oligomers, and fibrils in the presence of hydrophobic crowders to theoretical and experimental results describing how hydrophobic crowders can alter protein aggregation.

The protein-protein association/folding enhancement associated with adding hard sphere crowders to a system of proteins can be diminished or even negated when peptide-crowder attractive interactions are also present, as would occur with hydrophobic crowders.<sup>7-</sup>

<sup>16</sup> The impact of these attractions on protein aggregation can be measured in terms of the difference between the protein-protein binding free energy when attractive crowders are present and when they are absent.<sup>7, 17, 18</sup> This difference can be further broken down into enthalpic and entropic contributions. Since the enthalpic contribution to the change in protein-protein binding free energy when attractive crowders are added is intrinsically temperature-dependent (because it arises due to the attractions between protein and crowder) while the entropic contribution is not (because it arises from excluded volume), there is a competition between the entropic and enthalpic effects arising from the addition of attractive crowders. The entropic effect dominates at high temperatures, in which case the crowders enhance protein aggregation as we saw in our previous paper. The enthalpic effect dominates at low temperatures, in which case the crowders effectively pull the aggregates apart and reduce the peptides' aggregation propensity. At a constant temperature and crowder size/volume fraction, the entropy-enthalpy competition depends on how strong the crowder-peptide interactions are. The entropic effect dominates at low crowder-peptide interactions and the enthalpic effect dominates at high crowder-peptide interactions. In this work we examine crowder-peptide systems at a constant temperature and evaluate how different crowder-peptide interaction strengths change the oligomerization and fibrillization of A $\beta$ (16-22). The results are summarized in a global phase diagram that shows the types of structures that appear (fibril, disordered oligomer, and  $\beta$ -sheet), at different crowder sizes and crowder hydrophobicities.

The magnitudes of the entropic and enthalpic contributions to the difference between the protein-protein binding free energy when attractive crowders are present and when they

are absent are highly dependent upon the system being studied. Entropic effects will be large if there is a high crowder volume fraction or the crowders have a small diameter leading to larger depletion forces.<sup>6, 19</sup> Enthalpic effects increase as the crowder-protein interactions increase.<sup>7-16</sup> The magnitude of the crowder-protein interactions is dictated by the solution properties, the protein's sequence, number of hydrophobic residues, solvent accessible surface area in the folded and unfolded states, and the crowder geometry. For example, a protein with a large number of hydrophobic residues or a large solvent accessible surface area will have stronger crowder-protein interactions than proteins lacking hydrophobic residues or having a small solvent accessible surface area. It should be noted that at very high hard sphere crowder volume fractions crowders can actually diminish aggregation rather than enhance it, but the effect depends on the protein of interest.<sup>20, 21</sup> In this work we will confine our attention to A $\beta$ (16-22) peptide; examination of sequence dependent changes in crowder-peptide interactions is reserved for future work.

## 3.2 Methods

In this paper we combine discontinuous molecular dynamics<sup>22</sup> (DMD), a fast alternative to traditional molecular dynamics, with a four-sphere per residue protein model developed in our group called PRIME20<sup>23</sup>. The crowders in these simulations are modeled as spheres of diameters 5Å, 20Å and 40Å. We simulate a crowder volume fraction of  $\phi = 0.10$  as in our previous paper where crowder volume fraction is defined as  $\phi = NV_0/L^3$  where N is the number of crowders,  $V_0$  is the volume of a single crowder, and L is the simulation box length. For additional details on the selection of these parameters please refer to our previous

paper.<sup>6</sup> The crowder-crowder interactions are restricted to excluded volume interactions because we do not want to deal with aggregation of the crowders. Crowders interact with the peptide side chains via square-well and square-shoulder interactions. Several sets of values for the square-well interaction strengths between the peptide side chains, K, L V F and A, and crowder spheres are examined to better characterize the effects of hydrophobic crowders on A $\beta$ (16-22) aggregation. The reference value for the crowder-side chain interaction energy,  $\epsilon$ , hereafter referred to as  $1\epsilon$ , was chosen to be the set of interaction energies between an alanine side chain and each amino acid side chain in the PRIME20 force field. For example, the values of the  $1\epsilon$  crowder-side chain interactions are the same as the values for the {A-K, A-L, A-V, A-F, A-A, and A-E} side chain interactions: {0.074, -0.148, -0.148, -0.148, -0.084, and 0.074}. The alanine-alanine side-chain interaction was selected to represent the crowder-side-chain interaction because it is a weak hydrophobic interaction similar to the interaction that a side chain would experience with other macromolecules. Table 3.1 lists all of the crowder-side chain interaction energies, with negative values signifying attractive square wells and positive values signifying repulsive square shoulders, as well as the ranges of the crowder-side chain interaction considered in this paper. It should be noted that the interaction range is the same regardless of the interaction strength.

The simulations were conducted using the following procedure. Forty eight A $\beta$ (16-22) peptides with random coil structures and crowders at the desired crowder volume fraction were placed initially at random locations in a cubic simulation box with sides of length of  $L=252\text{\AA}$ , giving a peptide concentration of 5mM. Periodic boundary conditions were maintained. The reduced temperature in our simulations, defined as  $T^*=kT/\epsilon_{HB}$  where  $\epsilon_{HB}$  is

the hydrogen bonding well depth was set to  $T^*=0.193$ . The simulations were performed in the canonical ensemble where the number of particles, temperature, and volume are fixed. At least three independent simulations were run at each set of conditions for a minimum of 65 billion collisions for the most crowded conditions (5Å crowders,  $\phi=0.10$ ) and up to 225 billion collisions in the absence of crowders. Since the crowded conditions greatly accelerate aggregation, less time is needed to reach the same level of aggregation as in simulations without crowders. In our simulations a peptide is defined as: (a) being part of a disordered oligomer if it shares at least one hydrogen bond or one hydrophobic contact with another peptide in the oligomer, (b) being part of an ordered oligomer ( $\beta$ -sheet) if it shares at least 4 hydrogen bonds with another peptide in a  $\beta$ -sheet, and (c) being part of a fibril if it is in a  $\beta$ -sheet that shares at least 4 side chain interactions with another beta sheet in the fibril. For additional details on the simulation methods please refer to our previous paper.<sup>6</sup>

### **3.3 Results and Discussion**

#### **3.3.1 Aggregation in the presence of hydrophobic crowders**

Figure 3.1 shows snapshots that summarize the four different types of aggregates that we observe as final configurations in our simulations of A $\beta$ (16-22) in the presence of hydrophobic crowders. Figure 3.1A is a snapshot of a fibril that we observe in simulations with 5Å hard sphere crowders; these typically contain many  $\beta$ -sheets with a small number of peptides per sheet. These fibrils form due to the large depletion forces associated with small diameter crowders. The large depletion forces cause aggregation to occur via nucleated

conformational conversion creating a large number of  $\beta$ -sheets within a given fibril. Aggregation via nucleated conformational conversion is characterized by the fast collapse of peptides into disordered oligomers that quickly rearrange into  $\beta$ -sheets and later fibrils. Figure 3.1B is a snapshot of a fibril that we see in simulations with 40Å hard sphere crowders; these typically contain a small number of  $\beta$ -sheets and a large number of peptides per sheet. The small magnitude of the depletion forces associated with 40Å crowders creates an environment favoring the nucleated polymerization aggregation mechanism and formation of fibrils with a small number of  $\beta$ -sheets. Fibrils created through the nucleated polymerization mechanism typically grow by adding peptides one-by-one to the ends of the fibrillar  $\beta$ -sheets rather than by adding new  $\beta$ -sheets, which leads to a small number of  $\beta$ -sheets per fibril. Figure 3.1C is a snapshot of the large disordered oligomer we observe in simulations with 5Å crowders and a crowder-peptide interaction energy of  $5\epsilon$ . This type of oligomer forms due to the combination of large depletion forces and very strong crowder-peptide interactions. The depletion forces promote the aggregation of the peptides and the crowder-peptide interactions pulls them into a bent shape, preventing the peptides from forming ordered structures and forcing them to remain in a disordered oligomer. Figure 3.1D is a snapshot of a single  $\beta$ -sheet oligomer that we observe in simulations with 40Å crowders and a crowder-peptide interaction energy of  $5\epsilon$ . Although  $\beta$ -sheets are typically an intermediate in the formation of fibrils, in our simulations at this set of parameters,  $\beta$ -sheets are the predominant aggregate species at the end of the simulations. The large curved surface of a 40Å crowder combined with the strong crowder-peptide interactions promotes the

nucleation of  $\beta$ -sheets on these surfaces, and then holds them tightly against the surface, preventing further aggregation into fibrils.

### 3.3.2 Effect of hydrophobic crowders on disordered oligomer formation

Here we examine how the presence of hydrophobic crowders impacts the formation of disordered oligomers. To put this discussion in perspective, it helps to recall that in our previous simulations of peptide aggregation with hard sphere crowders, disordered oligomers formed only transiently and then rapidly reorganized into more ordered  $\beta$ -sheets structures. Figure 3.2 shows the number of peptides in disordered oligomers vs. reduced time for crowder volume fraction  $\phi=0.10$  with crowder diameters (A) 40Å, (B) 20Å, and (C) 5Å. At a crowder volume fraction of  $\phi=0.10$  with 40Å crowders (Fig. 2A), almost no disordered oligomers form except for the most hydrophobic crowders at interaction strength  $5\epsilon$ , where some disordered aggregates form but quickly disappear. When the crowder diameter is decreased to 20Å (Fig. 2B) disordered oligomers that initially form are stabilized and persist through the simulation without reorganizing into ordered structures. Finally with 5Å crowders (Fig. 2C), disordered oligomers form at the  $4\epsilon$  and  $5\epsilon$  interaction strengths. At  $4\epsilon$  the disordered oligomers form but are gradually converted to other species, but at  $5\epsilon$  the number of peptides in disordered oligomers grows continually until almost all of the peptides are a part of a single disordered oligomer. The simulations with 5Å crowders require a larger number of crowders to achieve a crowder volume fraction of  $\phi=0.10$  than simulations with 20Å or 40Å crowders. With a large potential number of crowder-peptide interactions and a high crowder-peptide interaction strength, the most energetically favorable state for an

individual peptide would be one where it experiences the maximum number of attractive crowder interactions. Since the five central residues of A $\beta$ (16-22), LVFFA, are hydrophobic they are attracted to the crowders while the N-terminal and C-terminal residues, K and E, are repulsed by the crowders. Maximizing the number of L, V, F, and A contacts for each crowder leads to the peptide adopting a bent conformation. This stable bent conformation is what prevents disordered oligomers from rearranging into ordered  $\beta$ -sheets and fibrils. For each of these cases, increasing the crowder volume fraction to  $\phi=0.20$  causes a decreased amount of disordered oligomers to form during a given simulation [Figure S3.1]. A higher crowder volume fraction will promote the formation of more compact ordered structures and since disordered oligomers are typically a metastable intermediate on the way to forming ordered structures, the number of disordered oligomers that form during the simulation decreases.

### **3.3.3 Effect of hydrophobic crowders on $\beta$ -sheet formation**

Next, we will examine the formation of  $\beta$ -sheets, which can be formed by the association of monomers, or through rearrangement of disordered oligomers. Again to put this discussion in perspective, it helps to recall that in our previous simulations of peptide aggregation with hard sphere crowders,  $\beta$ -sheets were a stable intermediate structure that eventually led to the formation of fibrillar structures. Figure 3.3 shows the number of peptides in  $\beta$ -sheets vs. reduced time for crowder volume fraction  $\phi=0.10$  with crowder diameters (A) 40Å, (B) 20Å, and (C) 5Å. At a crowder volume fraction of  $\phi=0.10$  with 40Å crowders (Fig. 3A) the number of peptides in  $\beta$ -sheets over time for the 1 $\epsilon$  and 2 $\epsilon$  interaction cases are almost the

same as those for hard spheres. Increasing the interaction strength to  $3\epsilon$  increases the number of  $\beta$ -sheets that form and increasing it to  $4\epsilon$  increases that number to its peak value. When the interaction strength reaches  $5\epsilon$ , the number of peptides forming  $\beta$ -sheets diminishes. This is likely because the very strong crowder-peptide interactions restrict peptide diffusion. Comparing the number of peptides in  $\beta$ -sheets for  $40\text{\AA}$  crowder to  $20\text{\AA}$  and  $5\text{\AA}$  crowders reveals an important distinction. When the interactions are above  $2\epsilon$  for the  $40\text{\AA}$  crowders, the number of peptides in  $\beta$ -sheets is much larger than those for  $20\text{\AA}$  and  $5\text{\AA}$  crowders at the same interaction strength and crowder volume fraction. We believe that this is due to the fact that  $40\text{\AA}$  crowders have a larger surface area available to peptides and act as nucleation sites for  $\beta$ -sheets. The curved surface of the crowder allows  $\beta$ -sheets to nucleate on the surface and wrap around the crowder. Decreasing the crowder diameter to  $20\text{\AA}$  (Fig. 3B) results in a reduction of the number of peptides in  $\beta$ -sheets for each interaction strength. Further decreasing the crowder diameter to  $5\text{\AA}$  provides an additional reduction in the number of peptides in  $\beta$ -sheets, except for  $3\epsilon$  and  $4\epsilon$  where there is very little change compared to the simulations with  $20\text{\AA}$  crowders. Comparison of Figures 3.2C and 3.3C shows that the simulations with  $5\text{\AA}$  crowders and  $4\epsilon$  interactions have both disordered oligomers and  $\beta$ -sheets. Increasing the crowder-peptide interaction from  $4\epsilon$  to  $5\epsilon$  for  $5\text{\AA}$  crowders dramatically decreases the number of peptides in  $\beta$ -sheets to practically zero. This signals a shift from a mixture of disordered and  $\beta$ -sheet oligomers to disordered oligomers only. The increase in interaction energy from  $4\epsilon$  to  $5\epsilon$  provides the crowders with enough interaction energy to pull ordered  $\beta$ -sheet oligomers apart and prevent their formation. Increasing the crowder volume fraction to  $\phi=0.20$  [Figure S3.2] has minimal effect on  $\beta$ -sheet formation for

any crowder with an interaction energy less than or equal to  $2\varepsilon$ . For crowders with interactions greater than  $2\varepsilon$ , increasing the crowder volume fraction to  $\phi = 0.20$  decreases the number of peptides in  $\beta$ -sheets.

### 3.3.4 Effect of hydrophobic crowders on fibril formation

Finally, we examine the formation of fibrillar structures. Generally this step requires multiple  $\beta$ -sheets to have already formed and then to merge into a single larger structure. Figure 3.4 shows the number of peptides in fibrils vs. reduced time for crowder volume fraction  $\phi = 0.10$  with crowder diameters (A)  $40\text{\AA}$ , (B)  $20\text{\AA}$ , and (C)  $5\text{\AA}$ . Fibrils are the final aggregate structure formed during simulations in the absence of crowders, and with hard sphere crowders. In our simulations of systems with  $40\text{\AA}$  crowders and  $20\text{\AA}$  crowders, the peptides form fibrils at crowder volume fractions  $\phi = 0.10$  and  $0.20$  when the crowder-peptide interactions are  $1\varepsilon$  and  $2\varepsilon$  [For  $\phi = 0.20$  see Figure S3.3]. At an interaction strength of  $3\varepsilon$  or higher for these crowder diameters, no fibrils form. This indicates that for  $40\text{\AA}$  and  $20\text{\AA}$  crowders, the transition point where crowder-peptide interactions become more favorable than peptide-peptide interactions is between  $2\varepsilon$  and  $3\varepsilon$ .

The simulation results for fibril formation with  $5\text{\AA}$  crowders are much more complex than those for the  $40\text{\AA}$  and  $20\text{\AA}$  crowders. For  $5\text{\AA}$  crowders we observe fibril formation with hard sphere and  $1\varepsilon$  interactions, no fibril formation at  $2\varepsilon$  interactions, fibril formation with  $3\varepsilon$  and  $4\varepsilon$  interactions, and no fibril formation at  $5\varepsilon$  interactions. Since the  $5\text{\AA}$  crowders create much larger depletion forces, and by extension promote the formation of fibrils, we would expect fibril formation to occur, at the very least, for hard spheres and interactions of  $1\varepsilon$  and

2 $\epsilon$ . However, as the interaction strength increases from hard sphere to 1 $\epsilon$  and 2 $\epsilon$  the crowder-peptide interactions become more and more favorable, delaying peptide-peptide interactions and by extension the formation of  $\beta$ -sheets and fibrils. Our hypothesis as to why there are no peptides in fibrils for the 2 $\epsilon$  interaction strength is that their formation is delayed beyond the scope of our simulations. If the fibrillization lag time continued to increase with crowder interaction strength we would have expected no fibrils to form during the time scale we simulated for interaction strengths 3 $\epsilon$ , 4 $\epsilon$ , and 5 $\epsilon$ , but we did observe fibril formation for 3 $\epsilon$  and 4 $\epsilon$  interactions. To understand why this happened, refer to Figures 3.2 and 3.3. Simulations at 3 $\epsilon$  have a very low number of peptide in disordered oligomers and simulations at 4 $\epsilon$  simulations have a quick increase in the number of peptides in disordered oligomers followed by a gradual decrease. Both 3 $\epsilon$  and 4 $\epsilon$  simulations have a gradually increasing number of peptides in ordered  $\beta$ -sheets. The increased crowder-peptide interaction strength delays disordered oligomer formation and allows them to persist longer than they do with hard sphere crowders because the crowder-peptide interactions are competing with peptide-peptide interactions. The 4 $\epsilon$  interaction simulations delay the structural rearrangement longer than the 3 $\epsilon$  interaction simulations because crowder-peptide interactions are stronger but both still ultimately result in fibrils. We believe the fibrillization mechanism that occurs for 5 $\text{\AA}$  crowders with 4 $\epsilon$  interactions is similar to the nucleated conformational conversion mechanism in which peptides rapidly collapse into disordered oligomers then rearrange into  $\beta$ -sheets and eventually fibrils, but that in the presence of strong hydrophobic crowders, the structural rearrangement step is prolonged due to favorable crowder-peptide interactions.

### 3.3.5 Global phase diagram peptides with hydrophobic crowders

The results on the types of aggregates observed at a crowder volume fraction of  $\phi=0.10$  are summarized in a “global phase diagram”, which shows the types of structures ultimately formed in the space spanned by the crowder size and energy parameters. This is depicted in Figure 3.5.

Although fibrils are the most energetically favorable aggregate structures in our simulations without crowders, or with hard sphere crowders, adding hydrophobic characteristics to the crowders creates an attraction between peptides and crowders that alters their preferred structural conformation. As indicated in Fig 5, simulations with hard spheres and interaction strengths of  $1\epsilon$  or  $2\epsilon$  ultimately produced fibrils for all crowder sizes. Although there is a weak hydrophobic interaction between the peptides and crowders, it is not enough to deter the peptides from ultimately forming fibrils. However, increasing the crowder-peptide interaction strength to  $3\epsilon$  or greater makes a large difference. At interaction  $3\epsilon$  and greater,  $40\text{\AA}$  crowders form ordered  $\beta$ -sheets, not fibrils, as the end-point aggregate the simulation. We believe that when the large curved surface of the  $40\text{\AA}$  crowder is paired with strong hydrophobic interactions, the crowders serve as nucleation sites for the  $\beta$ -sheets. The surface of the crowder curves gently enough to allow  $\beta$ -sheets to sit flush against the surface, and the hydrophobic interaction holds the  $\beta$ -sheet in place, effectively restricting its movement and preventing fibril formation. As interactions increase to  $4\epsilon$  and  $5\epsilon$ ,  $5\text{\AA}$  crowders begin to promote the formation of disordered oligomers as the primary aggregate type in our simulations. At  $4\epsilon$  a large number of disordered oligomers form, but are then slowly converted to fibrils. We use the notation F/D in Figure 3.5 to denote the presence of

both disordered oligomers and fibrils, and omit B for  $\beta$ -sheet because there are no stable  $\beta$ -sheet oligomers at the conclusion of the simulations. At  $5\epsilon$ , disordered aggregates form immediately and continue to grow in size, ultimately winding up as the only aggregate in the system. For the  $20\text{\AA}$  crowders, at interactions  $3\epsilon$  and  $4\epsilon$ , we see primarily  $\beta$ -sheet oligomers and a further increase to  $5\epsilon$  also promotes the formation of disordered oligomers. We use the notation B/D in Figure 3.5 to denote the presence of both  $\beta$ -sheets and disordered oligomers. This behavior is intermediate between that of the  $40\text{\AA}$  and  $5\text{\AA}$  crowders. If the crowder volume fraction is increased to  $\phi=0.20$  the phase boundaries will shift to higher interaction strengths (to the right in Figure 3.5), but the boundaries do not shift up or down along the crowder size axis. Increasing the crowder volume fraction increases the entropic contribution to the peptide-peptide binding free energy making it such that even higher enthalpic contributions are necessary to overcome enhanced peptide-peptide interactions, hence shifting phase boundaries to the right. The change in crowder volume fraction does not shift the boundaries up or down because the entropic contribution to the peptide-peptide binding free energy only changes the relationship with the enthalpic crowder-peptide interaction energy. We believe that more simulation time may be necessary to refine our global phase diagram at crowder volume fraction  $\phi=0.20$ , but we propose one in Figure S3.4.

### 3.4 Conclusions

DMD/PRIME20 simulations have been used to analyze the oligomerization and fibrillization of a large system of  $A\beta(16-22)$  peptides immersed in a sea of hydrophobic crowders. The focus here has been on how the introduction of crowder-peptide interactions,

beyond the excluded volume interactions that we considered previously, alters the types of aggregates formed and how this changes as the crowder size and hydrophobicity change. We believe that the unique contribution of our work lies in our ability to distinguish between the different types of aggregate structures formed and to learn what properties of hydrophobic crowders lead to their formation. The combination of a large system size, relatively realistic peptide model, and hydrophobic crowders rather than universally attractive crowders gives us a molecular-level perspective of how a more realistic crowder environment may alter protein aggregation. The systems considered here contain 48 A $\beta$ (16-22) peptides and crowders of diameters 5Å, 20Å, and 40Å, represented here by simple hard spheres or hydrophobic crowders, at a crowder volume fraction of 0.10. Our results show that specific combinations of crowder size and interaction strengths can create an environment in which the predominant species at the end of the simulation are either disordered oligomers or ordered  $\beta$ -sheet oligomers rather than the fibrillar structures that we typically observe in our aggregation simulations.

Our results are consistent with the findings of other research groups. Friedman and Caflisch used a 10-bead coarse-grained polypeptide model (a generic model specifically designed for study of protein aggregation, i.e. not representative of any particular peptide) with a 3-bead surfactant consisting of a single hydrophilic head bead and two hydrophobic tails.<sup>24</sup> While their surfactant model is not the same as our crowder model, the attractive portion of their surfactant does play a similar role to the attractive portion of our crowder model. In their work they found that the addition of the surfactants increases the lag time until fibril formation, and this is the same as what we found in the case of our 5Å crowders at

a crowder volume fraction of  $\phi=0.10$ . They also found that the polymorphism of the fibrils was ultimately unaffected by the presence of the surfactants. The polymorphism of fibrils in our simulations is determined by the parallel/antiparallel  $\beta$ -sheet content of the fibril because this dictates the inter-sheet side chain contacts. In our simulations a fibril is in its “ideal” conformation if 100% of the peptides are in anti-parallel  $\beta$ -sheets. A brief analysis of the parallel/antiparallel  $\beta$ -sheet content of our simulations shows that fibrils formed in the absence of crowders have  $\sim 90\%$  antiparallel  $\beta$ -sheets, fibrils formed in the presence of  $5\text{\AA}$  hard sphere crowders and crowder volume fraction of  $\phi=0.10$  have  $\sim 95\%$  antiparallel  $\beta$ -sheets, and fibrils formed in the presence of  $5\text{\AA}$  crowders with a  $4\epsilon$  interaction and a crowder volume fraction of  $\phi=0.10$  have  $<80\%$  antiparallel  $\beta$ -sheets. This shows that hard sphere crowders have the ability to increase the order of fibrils, while strong hydrophobic interactions decrease fibril order because of the dominant crowder-peptide interactions.

Although most other simulation and theoretical work done on this subject focuses on how attractive crowders affect the binding of two proteins<sup>9, 10, 25, 26</sup>, not on the aggregation of multi-protein systems, we can still compare our results to these studies because the same principals apply; our simulations are simply on a larger scale. In general, other researchers have found that attractive crowders destabilize protein secondary structure. Our short 7-residue peptide does not fold into a defined secondary structure when we simulate it in the absence of crowders so it is difficult to confirm the change in secondary structure. However, we observe that in the presence of  $5\text{\AA}$  crowders with  $5\epsilon$  interactions the radius of gyration of the A $\beta$ (16-22) monomers is  $\sim 2/3$ , and the end-to-end is  $\sim 1/2$ , of those measurements in the absence of crowders. This indicates that the monomers are folded into a bent shape even

before they are integrated into a disordered oligomer. Generally speaking, attractive crowders have been found to decrease protein-protein binding in favor of crowder-protein interactions. We also observe this in our simulations, most specifically in the case of 40Å crowders with 5ε interactions. In those simulations, the peptides preferentially interact with crowders, and only after they are on the large curved surface of the crowder do they form β-sheets.

Experimental work examining the effects of attractive molecules on protein aggregation covers a range of different types of species including osmolytes<sup>7, 13-15, 27-29</sup>, saccharides<sup>30</sup>, small molecule inhibitors<sup>31-33</sup>, and quantum dots<sup>34</sup>. In general the presence of an attractive molecule is shown to reduce the measurable level of aggregation, which typically means the amount of fibrillar structures observed. Our measurement of the number of peptides in fibrils vs time (Figure 3.4) agrees with these findings, showing that hydrophobic crowders, in which the crowder-peptide interactions are primarily attractive, have the ability to reduce fibril formation as their interaction strength increases. We found that increasing the interaction strength (hydrophobicity) of the crowders increases the lag time of fibril formation until a critical interaction strength is reached. Past the critical interaction strength, crowder-peptide interactions dominate peptide-peptide interactions and no fibrils form. Instead, the high levels of crowder hydrophobicity promote only the formation of oligomers.

### **3.5 Acknowledgements**

This work was supported by the National Institutes of Health, USA, under grant GM056766 and EB006006 (C.K.H.)

### 3.6 References

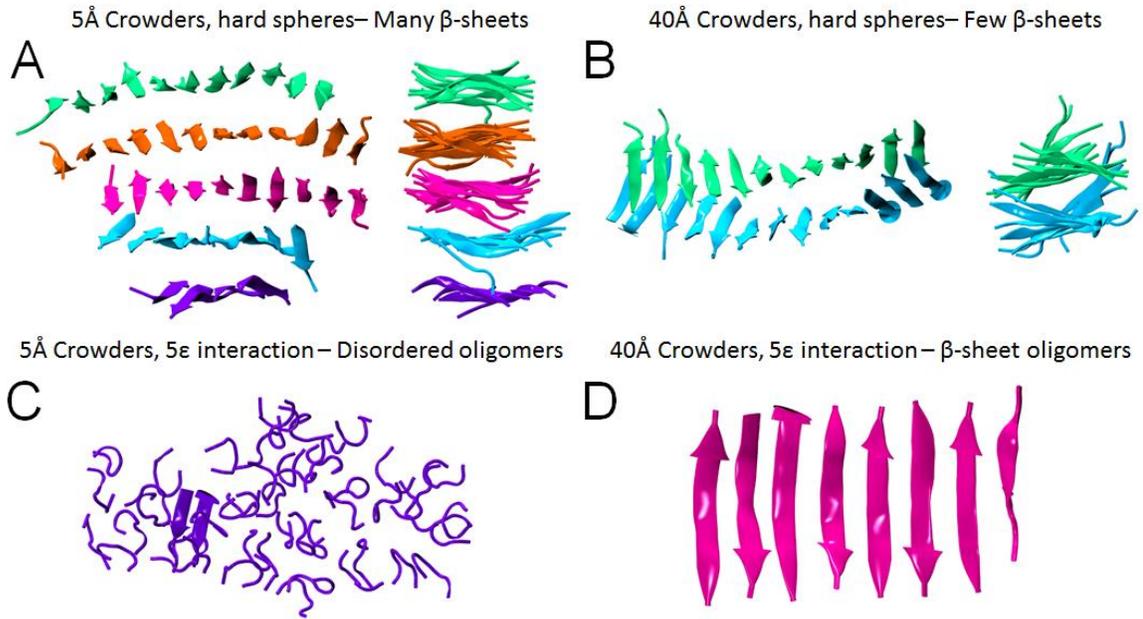
1. Alzheimer's facts and figures.  
[http://www.alz.org/alzheimers\\_disease\\_facts\\_and\\_figures.asp](http://www.alz.org/alzheimers_disease_facts_and_figures.asp). Accessed 11/17, 2014.
2. 2014 alzheimer's statistics. <http://www.alzheimers.net/resources/alzheimers-statistics/>. Accessed 11/17, 2014.
3. Ross C, Poirier M. Protein aggregation and neurodegenerative disease. *Nat Med*. 2004;10(7):S10-S17. doi: 10.1038/nm1066.
4. Gasser B, Saloheimo M, Rinas U, *et al*. Protein folding and conformational stress in microbial cells producing recombinant proteins: A host comparative overview. *Microbial Cell Factories*. 2008;7:11. doi: 10.1186/1475-2859-7-11.
5. Ignatova Z. Monitoring protein stability *in vivo*. *Microbial Cell Factories*. 2005;4:23. doi: 10.1186/1475-2859-4-23.
6. Latshaw DC, Cheon M, Hall CK. Effects of macromolecular crowding on amyloid beta (16-22) aggregation using coarse-grained simulations. *Journal of Physical Chemistry B*. <http://dx.doi.org/10.1021/jp508970q>. doi: 10.1021/jp508970q.
7. Jiao M, Li H, Chen J, Minton AP, Liang Y. Attractive protein-polymer interactions markedly alter the effect of macromolecular crowding on protein association equilibria. *Biophys J*. 2010;99(3):914-923. doi: 10.1016/j.bpj.2010.05.013.
8. Sarkar M, Lu J, Pielak GJ. Protein crowder charge and protein stability. *Biochemistry (N Y)*. 2014;53(10):1601-1606. doi: 10.1021/bi4016346.

9. Rosen J, Kim YC, Mittal J. Modest protein-crowder attractive interactions can counteract enhancement of protein association by intermolecular excluded volume interactions. *J Phys Chem B*. 2011;115(11):2683-2689. doi: 10.1021/jp200625k.
10. Kim YC, Mittal J. Crowding induced entropy-enthalpy compensation in protein association equilibria. *Phys Rev Lett*. 2013;110(20):208102. doi: 10.1103/PhysRevLett.110.208102.
11. Shen VK, Cheung JK, Errington JR, Truskett TM. Insights into crowding effects on protein stability from a coarse-grained model. *Journal of Biomechanical Engineering-Transactions of the Asme*. 2009;131(7):071002. doi: 10.1115/1.3127259.
12. Wojciehowski M, Cieplak M. Effects of confinement and crowding on folding of model proteins. *BioSystems*. 2008;94(3):248-252. doi: 10.1016/j.biosystems.2008.06.016.
13. Sukenik S, Harries D. Insights into the disparate action of osmolytes and macromolecular crowders on amyloid formation. *Prion*. 2012;6(1):26-31. doi: 10.4161/pri.6.1.18132.
14. Sukenik S, Sapir L, Harries D. Balance of enthalpy and entropy in depletion forces. *Current Opinion in Colloid & Interface Science*. 2013;18(6):495-501. doi: 10.1016/j.cocis.2013.10.002.
15. Sukenik S, Politi R, Ziserman L, Danino D, Friedler A, Harries D. Crowding alone cannot account for cosolute effect on amyloid aggregation. *Plos One*. 2011;6(1):e15608. doi: 10.1371/journal.pone.0015608.

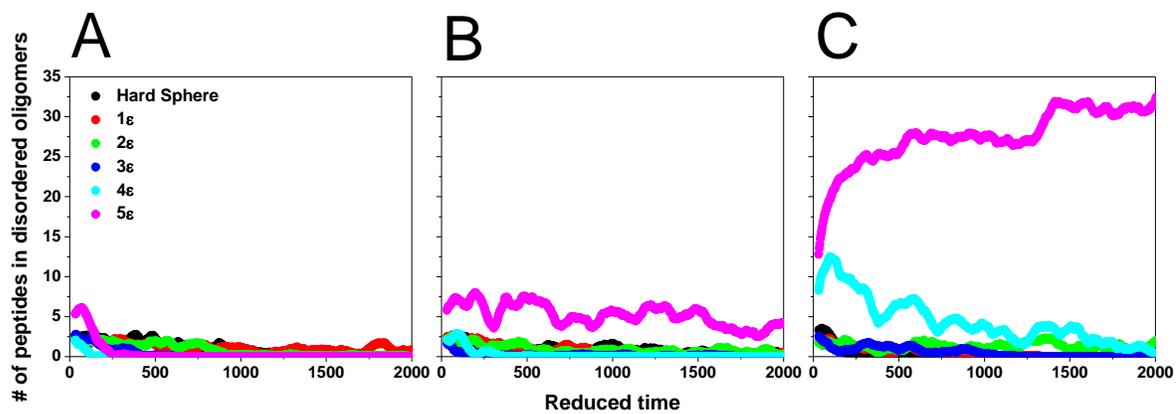
16. Sapir L, Harries D. Origin of enthalpic depletion forces. *Journal of Physical Chemistry Letters*. 2014;5(7):1061-1065. doi: 10.1021/jz5002715.
17. Minton A. Excluded volume as a determinant of macromolecular structure and reactivity. *Biopolymers*. 1981;20(10):2093-2120. doi: 10.1002/bip.1981.360201006.
18. Minton A. The effect of volume occupancy upon the thermodynamic activity of proteins - some biochemical consequences. *Mol Cell Biochem*. 1983;55(2):119-140. doi: 10.1007/BF00673707.
19. Hall D, Minton A. Macromolecular crowding: Qualitative and semiquantitative successes, quantitative challenges. *BBA-Proteins Proteomics*. 2003;1649(2):127-139. doi: 10.1016/S1570-9639(03)00167-5.
20. Cabaleiro-Lago C, Quinlan-Pluck F, Lynch I, Dawson KA, Linse S. Dual effect of amino modified polystyrene nanoparticles on amyloid beta protein fibrillation. *ACS Chemical Neuroscience*. 2010;1(4):279-287. doi: 10.1021/cn900027u.
21. Nguyen Truong Co, Hu C, Li MS. Dual effect of crowders on fibrillation kinetics of polypeptide chains revealed by lattice models. *J Chem Phys*. 2013;138(18):185101. doi: 10.1063/1.4804299.
22. Alder BJ, Wainwright TE. Studies in molecular dynamics .1. general method. *J Chem Phys*. 1959;31(2). doi: 10.1063/1.1730376.
23. Cheon M, Chang I, Hall CK. Extending the PRIME model for protein aggregation to all 20 amino acids. *Proteins-Structure Function and Bioinformatics*. 2010;78(14):2950-2960. doi: 10.1002/prot.22817.

24. Friedman R, Caflisch A. Surfactant effects on amyloid aggregation kinetics. *J Mol Biol.* 2011;414(2):303-312. doi: 10.1016/j.jmb.2011.10.011.
25. Minton AP. Quantitative assessment of the relative contributions of steric repulsion and chemical interactions to macromolecular crowding. *Biopolymers.* 2013;99(4):239-244. doi: 10.1002/bip.22163.
26. Kim YC, Bhattacharya A, Mittal J. Macromolecular crowding effects on coupled folding and binding. *The journal of physical chemistry.B.* 2014;118(44):12621-12629. doi: 10.1021/jp508046y.
27. Uversky VN, Li J, Fink AL. Trimethylamine-N-oxide-induced folding of alpha-synuclein. *FEBS Lett.* 2001;509(1):31-35. doi: 10.1016/S0014-5793(01)03121-0.
28. Ignatova Z, Gierasch LM. Inhibition of protein aggregation *in vitro* and *in vivo* by a natural osmoprotectant. *Proc Natl Acad Sci U S A.* 2006;103(36):13357-13361. doi: 10.1073/pnas.0603772103.
29. Natalello A, Liu J, Ami D, Doglia SM, de Marco A. The osmolyte betaine promotes protein misfolding and disruption of protein aggregates. *Proteins-Structure Function and Bioinformatics.* 2009;75(2):509-517. doi: 10.1002/prot.22266.
30. Fung J, Darabie A, McLaurin J. Contribution of simple saccharides to the stabilization of amyloid structure. *Biochem Biophys Res Commun.* 2005;328(4):1067-1072. doi: 10.1016/j.bbrc.2005.01.068.
31. Chebaro Y, Jiang P, Zang T, *et al.* Structures of A beta 17-42 trimers in isolation and with five small-molecule drugs using a hierarchical computational procedure. *J Phys Chem B.* 2012;116(29):8412-8422. doi: 10.1021/jp2118778.

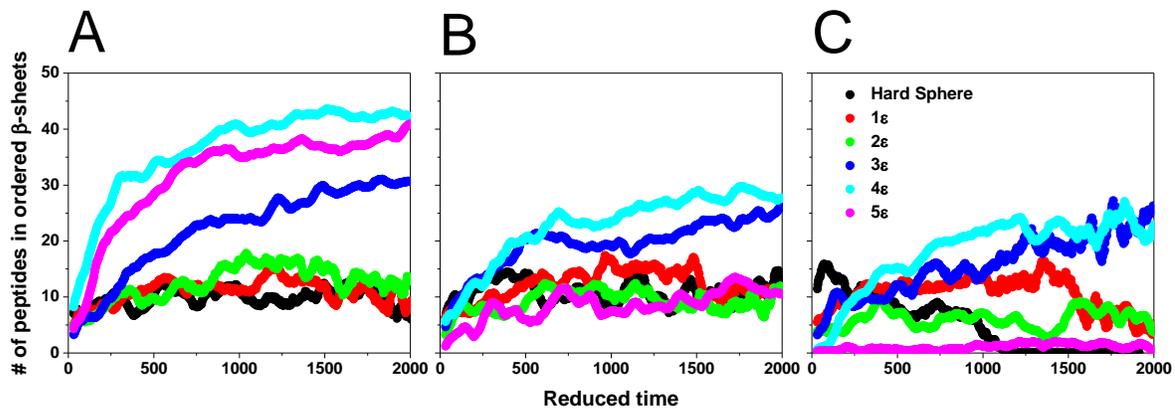
32. Reinke AA, Gestwicki JE. Structure-activity relationships of amyloid beta-aggregation inhibitors based on curcumin: Influence of linker length and flexibility. *Chemical Biology & Drug Design*. 2007;70(3):206-215. doi: 10.1111/j.1747.0285.2007.00557.x.
33. Feng Y, Wang X, Yang S, *et al*. Resveratrol inhibits beta-amyloid oligomeric cytotoxicity but does not prevent oligomer formation. *Neurotoxicology*. 2009;30(6):986-995. doi: 10.1016/j.neuro.2009.08.013.
34. Xiao L, Zhao D, Chan W, Choi MMF, Li H. Inhibition of beta 1-40 amyloid fibrillation with N-acetyl-L-cysteine capped quantum dots. *Biomaterials*. 2010;31(1):91-98. doi: 10.1016/j.biomaterials.2009.09.014.



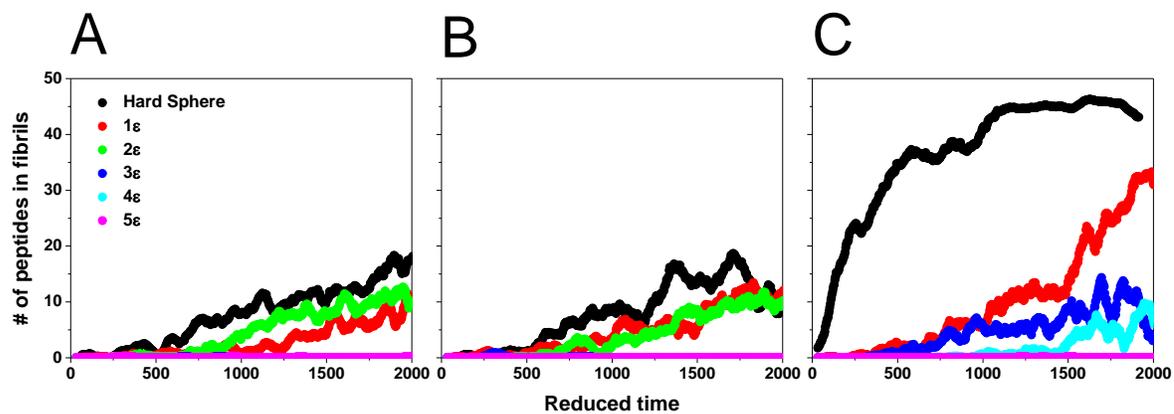
**Figure 3.1.** Snapshots of the different types of aggregates observed as final configurations in our simulations with A) 5Å hard sphere crowder case results in fibrils with many  $\beta$ -sheets B) 40Å hard sphere crowder case results in fibrils with few  $\beta$ -sheets C) 5Å crowders with 5 $\epsilon$  interaction results in disordered oligomers and D) 40Å crowders with 5 $\epsilon$  interaction results in  $\beta$ -sheet oligomers.



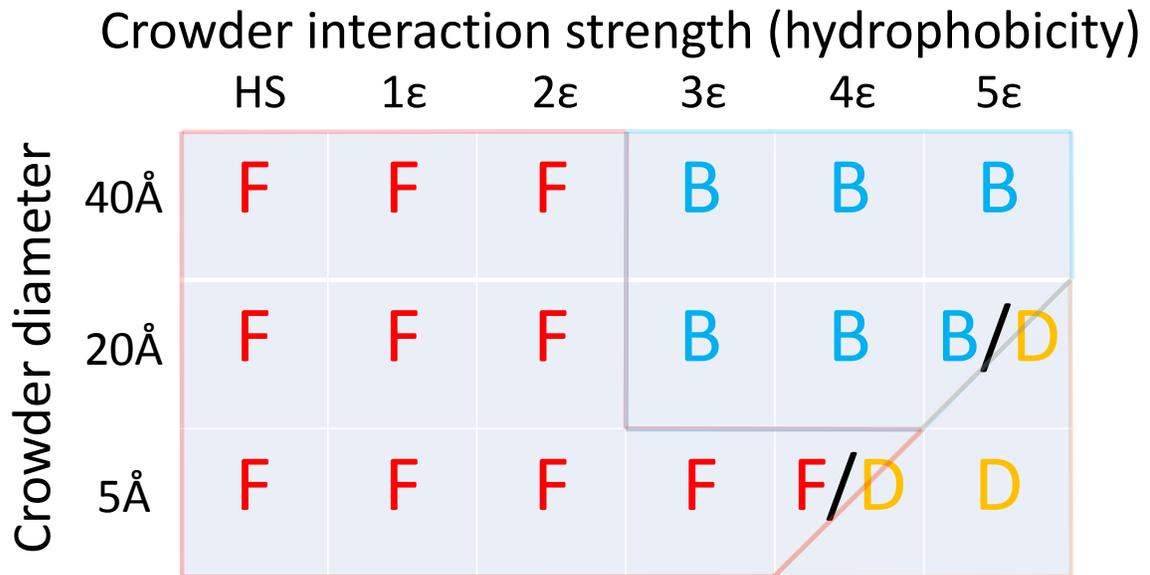
**Figure 3.2.** Number of peptides in disordered oligomers vs. reduced time for crowder volume fraction  $\phi=0.10$  with crowder diameters (A)  $40\text{\AA}$ , (B)  $20\text{\AA}$ , and (C)  $5\text{\AA}$ .



**Figure 3.3.** Number of peptides in  $\beta$ -sheets vs. reduced time for crowder volume fraction  $\phi=0.10$  with crowder diameters (A) 40Å, (B) 20Å, and (C) 5Å.



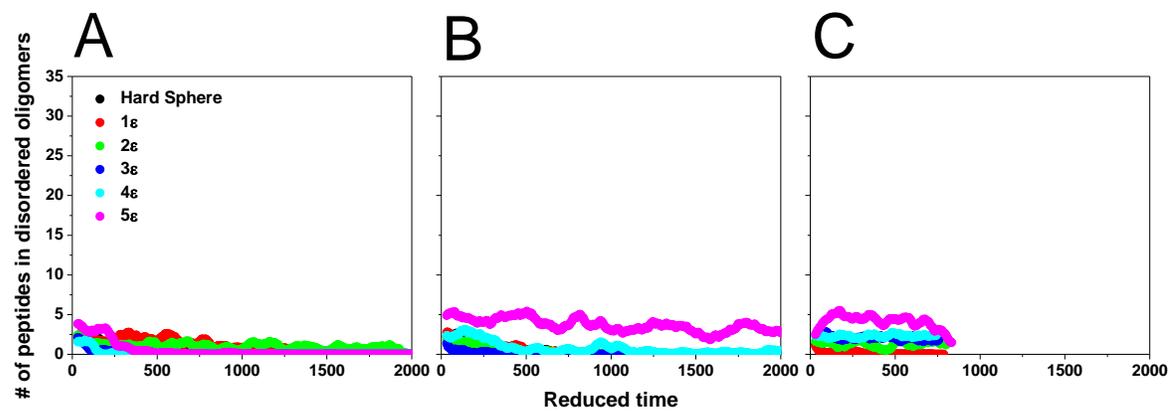
**Figure 3.4.** Number of peptides in fibrils vs. reduced time for crowder volume fraction  $\phi=0.10$  with crowder diameters (A)  $40\text{\AA}$ , (B)  $20\text{\AA}$ , and (C)  $5\text{\AA}$ .



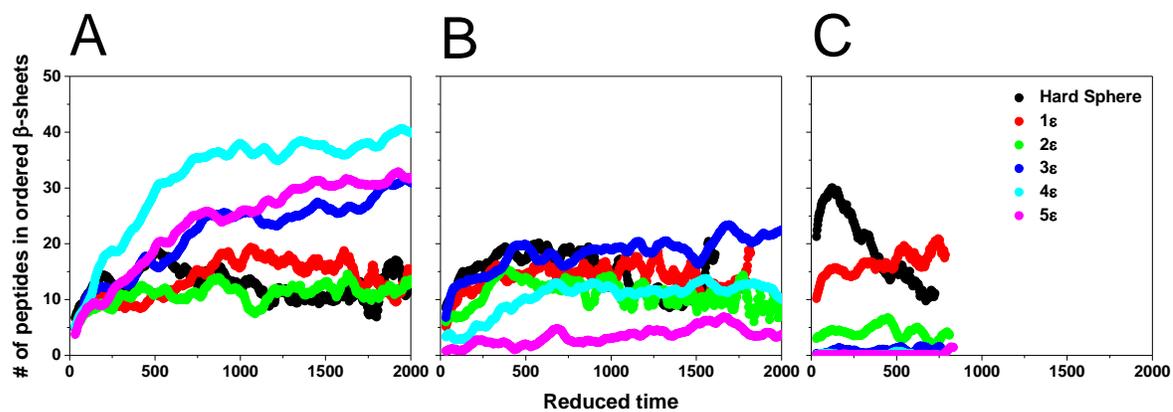
**Figure 3.5.** Global phase diagram summarizing the types of aggregates present in each simulation at crowder volume fraction  $\phi=0.10$  for different values of the crowder diameter,  $D$ , and interaction strength,  $X\epsilon$ . F=fibrils, B=ordered  $\beta$ -sheet oligomers, and D=disordered oligomers.

**Table 3.1.** Crowder- side chain interaction energy parameters.

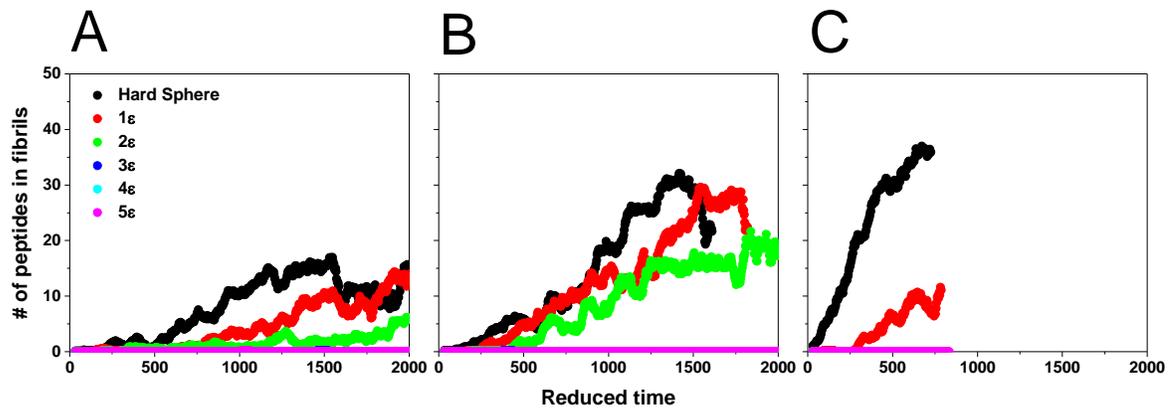
Amino Acid	Interaction strength $\epsilon$ and range $\lambda$ between crowder and side chains K, L, V, F, A and E					
	$1\epsilon$	$2\epsilon$	$3\epsilon$	$4\epsilon$	$5\epsilon$	$\lambda$
K	0.074	0.148	0.222	0.296	0.370	1.357
L	-0.148	-0.296	-0.444	-0.592	-0.740	1.436
V	-0.148	-0.296	-0.444	-0.592	-0.740	1.699
F	-0.148	-0.296	-0.444	-0.592	-0.740	1.752
A	-0.084	-0.168	-0.252	-0.336	-0.420	1.350
E	0.074	0.148	0.222	0.296	0.370	1.482



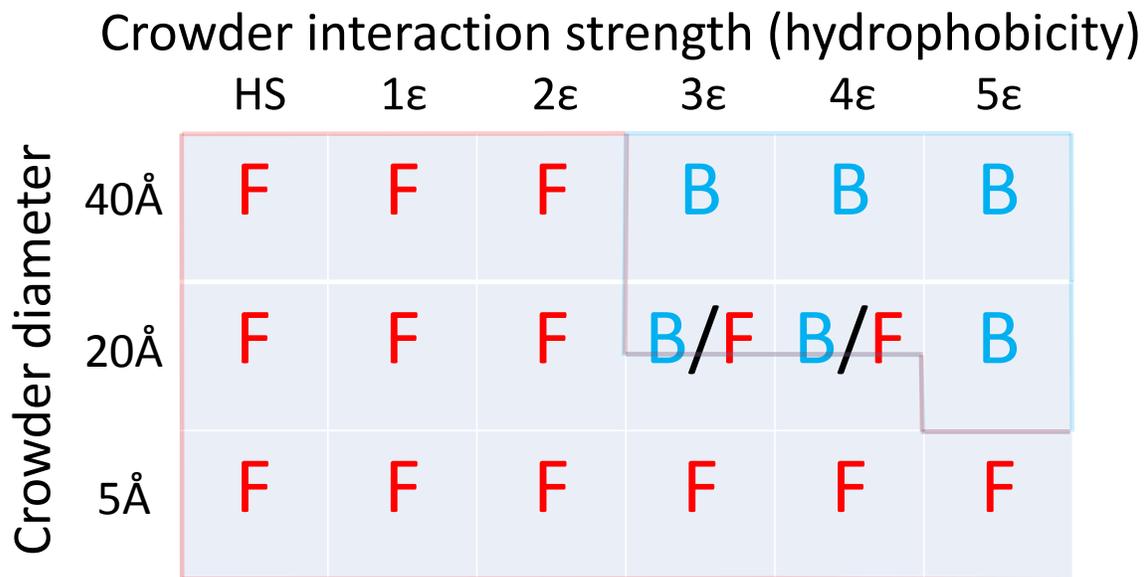
**Figure S3.1.** Number of peptides in disordered oligomers vs. reduced time for crowder volume fraction  $\phi=0.20$  with crowder diameters (A)  $40\text{\AA}$ , (B)  $20\text{\AA}$ , and (C)  $5\text{\AA}$ .



**Figure S3.2.** Number of peptides in  $\beta$ -sheets vs. reduced time for crowder volume fraction  $\phi=0.20$  with crowder diameters (A) 40Å, (B) 20Å, and (C) 5Å.



**Figure S3.3.** Number of peptides in fibrils vs. reduced time for crowder volume fraction  $\phi=0.20$  with crowder diameters (A)  $40\text{\AA}$ , (B)  $20\text{\AA}$ , and (C)  $5\text{\AA}$ .



**Figure S3.4.** Proposed global phase diagram summarizing the types of aggregates present in each simulation at crowder volume fraction  $\phi=0.20$  for different values of the crowder diameter,  $D$ , and interaction strength,  $X\epsilon$ . F=fibrils, B=ordered  $\beta$ -sheet oligomers, and D=disordered oligomers.

## **CHAPTER 4**

### **Aggregation of Amphipathic Peptides at an Aqueous/Organic Interface using Coarse-Grained Simulations**

Chapter 4 is essentially a manuscript by David C Latshaw II and Carol K Hall in preparation.

# Aggregation of Amphipathic Peptides at an Aqueous-Organic Interface using Coarse Grained Simulations

David C Latshaw II and Carol K Hall\*

Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, North Carolina

## Abstract

To examine the effect of an aqueous/organic interface on the folding and aggregation of amphipathic peptides, discontinuous molecular dynamics (DMD) simulations combined with an intermediate resolution protein model, PRIME20, were applied to a peptide/interface system. The systems contained either 1 or 48 (KLLK)<sub>4</sub> peptides in random coil or  $\alpha$ -helical conformations interacting with strong and weak interfaces. Results show that the weak interface is capable of folding a randomly coiled (KLLK)<sub>4</sub> peptide into an amphipathic  $\alpha$ -helix, while the strong interface disrupts intrapeptide hydrogen bonding and forces the peptides to adopt a random coil or  $\beta$ -hairpin conformation. We also found that when peptides interacted with the strong interface they adsorbed irreversibly due to the strong side chain preference for their respective phases. When peptides interacted with the weak interface, it was possible for them to adsorb and dissociate. The (KLLK)<sub>4</sub> peptide was more likely to adsorb to the interface when it was in an  $\alpha$ -helical conformation as opposed to a random coil because of its larger hydrophobic moment. We also studied the aggregation of 48 (KLLK)<sub>4</sub> peptides in the presence of an aqueous/organic interface. We found that the

conditions producing the highest amount of aggregation were  $\alpha$ -helical peptides interacting with a weak interface. Since the peptides adsorbed reversibly to the interface, the increased local concentration of helices promoted the formation of helical bundles. The most energetically favorable type of aggregate in the aqueous phase was the helical bundle, and the most energetically favorable type of aggregate at the interface is the  $\beta$ -sheet. The results are in excellent qualitative agreement with previous experimental and theoretical work.

## 4.1 Introduction

Commercial-level production of therapeutic proteins is a challenging task that involves several steps including production, purification, storage and delivery.<sup>1</sup> Many therapeutic proteins are stored and later delivered via glass-prefilled syringes in which the plunger is lubricated with a silicone oil.<sup>2</sup> Within these syringes are aqueous/organic interfaces due to plunger lubrication and aqueous-air interfaces from air bubbles arising during the filling process. While the structure of therapeutic proteins in aqueous environments is often well characterized, their secondary structures can be altered when an organic interface or membrane environment is encountered.<sup>3-6</sup> Such conformational changes can lead to reduced protein functionality and predispose proteins to misfolding and aggregation.<sup>7</sup> It is not surprising therefore that interfaces have been found to cause the unwanted aggregation of a variety of therapeutic proteins.<sup>8-15</sup> Not only does the aggregation of these proteins render them ineffective, it can elicit an unwanted immune response in patients, creating further complications.<sup>16</sup>

Antimicrobial peptides are part of our immune defense against pathogens and have promising value as therapeutic proteins.<sup>17,18</sup> This class of proteins typically has cationic and hydrophobic properties and adopts amphipathic  $\alpha$ -helices,  $\beta$ -sheets,  $\beta$ -hairpins, and larger mixed structures.<sup>18</sup> While these proteins come with a variety of sequences, their amphipathic character is a consistent theme. In this paper we use computer simulations to examine the effects of an aqueous/organic interface on the folding and aggregation of a (KLLK)<sub>4</sub> peptide which folds into an amphipathic  $\alpha$ -helix in the aqueous phase.

A number of simulation-based studies have been conducted on isolated short peptides interacting with an aqueous/organic interface. Chipot and Pohorille used atomistic molecular dynamics to simulate the translocation of an eleven-residue poly-leucine peptide across a water/hexane interface.<sup>19</sup> As the peptide passed through the interface into the hexane phase it folded into an  $\alpha$ -helix after 36ns and remained adsorbed to the interface rather than enter the bulk hexane phase. The peptide preferentially aligned parallel to the interface, although the perpendicular conformation was only slightly less energetically favorable. This result led them to propose that the mechanism of peptide insertion into membranes begins with the peptide initially aligning parallel to the interface and gradually shifting to a perpendicular orientation as it inserts into the membrane. In a later study, Chipot *et al.* performed simulations using an amphipathic peptide, LQQLLQQLQL, at the same water/hexane interface.<sup>20</sup> The peptide began interacting with the interface almost immediately, but did not fold into its ideal amphipathic  $\alpha$ -helix, despite a 161.5ns simulation trajectory. This was explained in terms of the large energetic penalties associated with transitioning from amphipathic to non-amphipathic conformations in the presence of the interface which increased the folding time scale of the peptide far beyond that of the poly-leucine peptide that they had simulated previously. In a paper by Anderson *et al.*, Monte Carlo simulations were used to determine how the presence of a water/oil interface affected the structure of a 27-residue heteropeptide.<sup>21</sup> Of the four location/conformation states that the protein could adopt (bulk/folded, bulk/unfolded, interface/folded and interface/unfolded), the interface/unfolded state was the most probable. This was because the peptide was able to partition its side chains into their respective phases without undergoing energetically unfavorable conformation

changes to reach the folded state. With the parameters they chose to represent oil, the peptide adsorbed to the interface irreversibly, mainly because the oil parameters favored this outcome. In a follow-on paper, Leonhard *et al.* investigated how a wider range of interface parameters changed the adsorption of the same 27-residue heteropeptide and also examined a 64-residue heteropeptide.<sup>22</sup> Variations in the strength and width of the interface produced conditions that led to no adsorption, reversible adsorption, and irreversible adsorption. They also showed that interfaces are more likely to adsorb peptides and alter peptide secondary structure than surfaces with similar properties.

The interaction of amphipathic peptides with interfaces has been studied experimentally. Maget-Dana *et al.* examined the properties of the peptides (LKKL)<sub>n</sub> and (LK)<sub>n</sub> at surfaces for several values of n; these two peptides form amphipathic  $\alpha$ -helices and  $\beta$ -sheets respectively, at an air/water interface.<sup>23</sup> The two peptides were found to have about the same equilibrium spreading pressure, but the (LKKL)<sub>n</sub> peptides were found to diffuse to and adsorb at the interface more quickly than the (LK)<sub>n</sub> peptides. Finally,  $\beta$ -sheet monolayers were found to be more stable and less compressible than  $\alpha$ -helix monolayers at the interface. In two papers by Wang *et al.* the interfacial properties of peptides that form both amphipathic  $\alpha$ -helices<sup>24</sup> and amphipathic  $\beta$ -sheets<sup>25</sup> were measured at a water/oil interface. They found that the peptides that adopted amphipathic  $\alpha$ -helices could be divided into two classes that differed in their ability to lower surface tension and dissociate from the interface. Furthermore they found that peptides that irreversibly bind interfaces have the ability to lower surface tension more than those that bind reversibly because of their strong interactions with each phase. In general, when peptides sit at an interface and interact with

both phases they lower the ability of each phase to bind with itself, reducing the tension of the interfacial surface. Wang *et al* suggested that the more hydrophobic content the peptide has, the stronger its interfacial binding and the higher its ability to lower surface tension. They also found that  $\beta$ -sheets were adsorbed irreversibly to the interface and lowered surface tension more than  $\alpha$ -helices. The conclusion that  $\beta$ -sheets are more stable at the interface is in good agreement with Maget-Dana *et al*.

In this paper we apply a combination of discontinuous molecular dynamics (DMD) and the PRIME20 force field to examine how the folding and aggregation of single and multi-peptide systems containing the (KLLK)<sub>4</sub> peptide are impacted by the presence of an aqueous/organic interface. The (KLLK)<sub>4</sub> peptide was selected because it is amphipathic in its native folded  $\alpha$ -helix and non-amphipathic when it is extended. Not only does this sequence offer a contrast in amphipathic behavior depending on its conformation, but it also has the ability to form multiple quaternary structures. It can form helical bundles as well as  $\beta$ -sheets, providing a rich variety of interface effects on amphipathic helices. In our simulations, we monitor the change in secondary structure and residence time of a single peptide as well as the aggregation of a system containing 48 peptides at aqueous/organic interfaces that have either a strong or weak preference for hydrophilic and hydrophobic amino acid side chains. In order to understand how different interface characteristics affect the folding/unfolding of amphipathic peptides, we monitor the number of native contacts, radius of gyration, end-to-end distance, and residence time of the peptides at the interface and in the bulk. We also explore the aggregation of peptides at the interface by monitoring the formation of helical bundles and  $\beta$ -sheets formed by the amphipathic peptide. We compare our results to

theoretical and experimental predictions of the effects of aqueous/organic interfaces on amphipathic peptides by Chipot *et al.*, Anderson *et al.*, Leonhard *et al.*, Maget-Dana *et al.*, and Wang *et al.*<sup>19-25</sup> and provide molecular-level detail on how an aqueous/organic interface affects the folding and aggregation of amphipathic peptides.

Highlights of our results are the following. The weak interface is capable of folding a randomly coiled (KLLK)<sub>4</sub> peptide into an amphipathic  $\alpha$ -helix, while the strong interface disrupts intra-peptide  $\alpha$ -helical hydrogen bonds and unfolds the peptides into a random coil or  $\beta$ -hairpin conformation. Regardless of the number of peptides in the simulation, peptides adsorbed irreversibly to the strong interface because of their side chain preference for their respective phases. Peptides that interact with the weak interface are reversibly adsorbed. The (KLLK)<sub>4</sub> peptide was more likely to bind to the interface when it was in an  $\alpha$ -helical conformation than when it was a random coil. This is likely because the  $\alpha$ -helix has a larger hydrophobic moment than the random coil, which increases the probability that it will interact with the organic phase. Aggregation was highest when peptides in an  $\alpha$ -helical initial conformation interacted with a weak interface. We hypothesize that the increased aggregation is due to the fact that the peptides could bind reversibly to the interface, increasing the local concentration of helices just outside the interface and promoting the formation of helical bundles. The lowest levels of aggregation were observed when the peptides were initially in the random coil state in the presence of a strong interface. This is likely because random coil peptides have a negligible hydrophobic moment, which decreases the probability that they will interact with the organic phase. A decreased frequency of interaction with the interface would make it difficult to form interfacial  $\beta$ -sheets. The most

energetically favorable type of aggregate in the aqueous phase was the helical bundle, and the most energetically favorable type of aggregate at the interface was the  $\beta$ -sheet.

## **4.2 Methods**

### **4.2.1 Discontinuous Molecular Dynamics and PRIME20 Force Field**

The simulation method used in this work is discontinuous molecular dynamics (DMD), a fast alternative to traditional molecular dynamics, coupled with the PRIME20 force field.<sup>26</sup> In DMD, the potential is a discontinuous function of the interatomic separation, e.g. hard-sphere and square-well potentials. Since the atoms move linearly between collisions, the only time that the velocities and positions need to be recalculated is when a discontinuity in the potential is encountered. This makes it faster than simulations with continuous potentials, where the simulation is advanced at regularly-spaced time steps on the order of picoseconds. The PRIME20 force field is based on a four-sphere-per-residue model; each amino acid residue is represented by three backbone spheres, one each for the N-H, C-H, and C=O, and one side-chain sphere R.<sup>27</sup> Detailed descriptions of these methods have been presented in other papers so only substantial modifications will be discussed here.<sup>28,29</sup>

### **4.2.2 (KLLK)<sub>4</sub> Peptide and Interface Model**

We chose the (KLLK)<sub>4</sub> peptide as the subject of our simulations because its native state is an amphipathic  $\alpha$ -helix and its extended conformation is non-amphipathic, i.e. the leucine and lysine residues alternate on each side of the peptide. We felt that this sequence would be ideal for studying amphipathic peptides because it offers the contrast between

amphipathic and nonamphipathic structures within the same peptide. Figure 4.1A and 4.1B show  $\alpha$ -helical and extended conformations of the (KLLK)<sub>4</sub> peptide using the PRIME20 protein representation. Figure 4.1C shows a sketch of the interface simulation box used in this paper with the blue area representing the aqueous phase, the orange area representing the organic phase and the solid black line representing the interface between the phases. Figure 4.1D shows the potential energy changes experienced by hydrophilic lysine and hydrophobic leucine side chain beads vs the z-position of the bead for the aqueous/organic interface. In Figure 4.1D the blue area represents the aqueous phase, the orange represents the organic phase and the grey area represents a mixed organic/aqueous phase. The interface potential is modeled using triple square-well and square-shoulder discontinuous potentials in order to be compatible with the other discontinuous potentials in the PRIME20 force field. The interface potential for a hydrophobic residue is represented by the orange line and the interface potential for a hydrophilic residue is represented by the blue line. The horizontal dotted line at  $y=0$  is zero potential energy.

The interface width was selected to be 4.5Å to mimic the simulation results of Ashbaugh *et al.*, who found that the width of a water/n-C<sub>18</sub> interface was 2-3 times the width of a water molecule or 3-4.5Å.<sup>30</sup> Using a triple well instead of a single well gives us a wider interface and allows us to have a mixed aqueous/organic phase with a more gradual transition to each pure phase, hence giving us a more realistic representation of the interface. The depth/height of each square well/shoulder was selected to mimic the density profile of water/n-C<sub>18</sub> in the paper from Ashbaugh *et al.* We tested interface profiles with up to seven wells to determine if the increased detail would provide a more realistic representation of the

interface, but ultimately determined that it was unnecessary since the results were the same but the simulation was much slower.

The total interface square well depth/shoulder height for amino acid  $i$ ,  $\varepsilon_i$ , is given by Equation 4.1

$$\varepsilon_i = 2\varepsilon_s \varepsilon_h \quad (4.1)$$

where  $\varepsilon_s$  is the interface strength, a user-defined parameter, and  $\varepsilon_h$  is the hydrophobicity score of the amino acid  $i$  side chain. In our simulations, we used two values of  $\varepsilon_s$ : 0.50 and 1.00, which will be referred to as the weak and strong interface, respectively. For these values of  $\varepsilon_s$  the peptides do not fully penetrate the interface and enter the organic phase. Each amino acid side chain bead was assigned a hydrophobicity score according to the scale developed by Kyte and Doolittle.<sup>31</sup> In their model the lysine side chain bead has a hydrophobicity score of -3.9, indicating its affinity for the aqueous phase, and leucine has a score of 3.8, giving it an affinity for the organic phase. We selected these two amino acids because their hydrophobicity scores were the closest in magnitude and we felt that it would lead to simulations in which the side chain attractions to each phase were balanced. For example: if we had selected alanine rather than leucine, the peptide would have a stronger preference for the aqueous phase because the magnitude of lysine's preference for the aqueous phase, 3.8, is much greater than alanine's preference for the organic phase, which is 1.8.

The energetic changes that an amino acid side chain undergoes as it approaches the interface can be understood by referring to Figure 4.1D. When the hydrophobic leucine side chain (orange line) is in the aqueous phase (blue) it is in an energetically unfavorable state,

experiencing a square-shoulder repulsion. When its z position increases, it enters mixed phase 1, followed by mixed phase 2 and then the organic phase lowering its potential energy as it enters each well. Thus as the leucine side chain follows this path it enters progressively more favorable energetic states. When the hydrophilic lysine side chain (blue line) is in the aqueous phase (blue) it is in an energetically favorable state. When its z position increases, it enters mixed phase 1, followed by mixed phase 2 and then the organic phase increasing its potential energy as it jumps up each shoulder. Thus as the lysine side chain follows this path, it enters progressively less favorable energetic states. We have also made modifications to the PRIME20 values for the lysine-lysine and leucine-leucine side-chain interactions to represent what happens when these two side chains interact in the organic phase. In the organic phase, electrostatic screening of the lysine-lysine interaction is minimal so we increase the repulsive interaction energy between lysine groups by 25% over its value in PRIME20. In the aqueous phase, the leucine-leucine interaction is attractive to mimic hydrophobic residues burying themselves away from water, but in the organic phase this behavior is expected to be absent and can be represented by a simple hard sphere interaction. Leucine-lysine interactions were the same in all phases. Thus the lysine-lysine and leucine-leucine interactions are gradually changed as the residues pass through the mixed phase of the interface according to the depth/ height of the well/shoulder. The interactions are summarized in Table 4.1 for clarity.

### 4.2.3 Cumulative Residency Time

Another important measure of how an aqueous/organic interface affects a peptide is the residency time of the peptide at the interface. Since the peptides take a random path in the simulation, it is possible that a peptide might not interact with the interface during a given time window. To measure how often the peptide interacts with the interface, and how long it stays there, we calculate the cumulative residency time. A peptide is counted as interacting with the interface if at least one of the peptide side chains has an interaction with the interface potentials. The cumulative residency time is the total amount of time that a peptide has spent adsorbed to the interface at a given point in the simulations. In a plot of cumulative residency time vs reduced time, a positive slope indicates that the peptide is adsorbed to the interface and a zero slope indicates that the peptide is in the aqueous phase. At a given reduced time on the x-axis, the total time the peptide has spent adsorbed to the interface can be obtained from the corresponding y-value of the cumulative residency time curve.

### 4.2.4 Simulation Procedure – Single Peptide Simulations

Prior to examining the effects of an aqueous/organic interface on the (KLLK)<sub>4</sub> peptide we needed to determine its folding temperature using the PRIME20 force field. The procedure is the following . A single isolated peptide is placed at a random location in a cubic box with sides of length  $L=252\text{\AA}$ . The reduced temperature is defined as  $T^*=kT/\epsilon_{\text{HB}}$  where  $\epsilon_{\text{HB}}$  is the hydrogen bonding well depth. Velocities for each peptide bead and are chosen at random from a Maxwell-Boltzmann distribution centered at the desired temperature. Initially, the temperature is set to  $T^*=0.30$ , a temperature high enough to

denature the peptide giving it a random coil secondary structure. The system is then gradually cooled to  $T^*=0.18$  in a stepwise fashion in increments of 0.02. From  $T^*=0.18$  to 0.14 the cooling increment is reduced to 0.005 to increase the resolution of the folding curve in what we anticipated to be the folding transition region. From  $T^*=0.14$  to  $T^*=0.10$  the cooling increment was returned to 0.02. Each temperature was run for 2 billion collisions. This cooling scheme resulted in 34 billion collisions, which is quite extensive for a single peptide. The temperature was held constant using the Andersen thermostat method. Beads in the simulation experience random “ghost collisions” with “ghost particles” during which their velocity is reassigned to a random value from a Maxwell-Boltzmann distribution centered at the desired simulation temperature. We ran three independent simulations to generate the folding data.

#### **4.2.5 Simulation Procedure – Single peptide interface simulations**

Single peptide simulations were run to determine how the presence of an aqueous/organic interface affected the folding/unfolding transition of the  $(KLLK)_4$  peptide. A single peptide was placed at a random location in the central aqueous phase of a cubic simulation box with side lengths of  $L=252\text{\AA}$  ( $126\text{\AA}$  from the origin), an interface location of  $189\text{\AA}$  ( $94.5\text{\AA}$  from the origin), and periodic boundary conditions, resulting in an overall peptide concentration of  $100\mu\text{M}$ , and  $245\mu\text{M}$  in the aqueous phase. As seen in Figure 4.1C, the aqueous phase has a smaller dimension in the z-direction than in the x and y-directions, resulting in a higher peptide concentration in the aqueous phase than in the simulation box as a whole. These simulations were run at two different reduced temperatures, depending on the

desired secondary structure of the peptide. If the peptide was started in a random coil the simulation was run at  $T^*=0.17$  to bias it toward that configuration. If the peptide was started in an  $\alpha$ -helix the simulation was run at  $T^*=0.15$  to preserve the helical structure. The decision to run each simulation at a temperature that biases it toward its starting configuration was made to accentuate the influence of the interface on the peptides. For example, if the peptide started in an  $\alpha$ -helix with a reduced temperature of  $T^*=0.15$  and it spontaneously unfolds at the interface, we can conclude that it was indeed the effects of the interface, and not a random event, that unfolded the peptide because the low temperature biases the secondary structure of the peptide toward an  $\alpha$ -helix. If we had run the simulation at a higher reduced temperature, the peptide would be more likely to unfold spontaneously in the aqueous phase, making our results unclear. We ran 5 independent simulations for each set of conditions until the peptide had interacted with the interface. Each simulation lasted ~15-20 billion collisions.

#### **4.2.6 Simulation Procedure – Multi-peptide interface aggregation simulations**

In multi-peptide interface aggregation simulations, 48 peptides are randomly positioned in the central aqueous phase of a simulation box with side lengths of  $L=252\text{\AA}$ , an interface location of  $189\text{\AA}$ , and periodic boundary conditions, resulting in an overall peptide concentration of 5mM, and 6.6mM in the aqueous phase. For the multi-peptide simulations, we also performed simulations in the absence of an interface to establish “background” readings that we could compare to our interface simulations. In order to make the peptide concentration in these simulations equivalent to the peptide concentration in the interface simulations, the multi-peptide simulations in the absence of an interface (aqueous phase

only) had 64 peptides in a simulation box with side lengths of  $L=252\text{\AA}$  and periodic boundary conditions, resulting in an overall peptide concentration of 6.6mM. We chose to run these simulations at  $T^*=0.165$ , slightly below the  $T^*=0.17$  used to bias the structure toward a random coil in single peptide simulations. This temperature was selected because it would not bias the peptides toward a folded structure and it would provide a balance between helical bundle and  $\beta$ -sheet formation. In these simulations, we define a peptide to be a part of a helical bundle oligomer if the peptide has at least six  $\alpha$ -helical hydrogen bonds and the peptides share at least four leucine-leucine side chain interactions with another peptide in the helical bundle. A peptide is defined to be a part of a  $\beta$ -sheet oligomer if it shares at least six interchain hydrogen bonds with another peptide in the  $\beta$ -sheet oligomer.

## 4.3 Results and Discussion

### 4.3.1 Single peptide folding simulations

In order to understand the temperature dependent folding behavior of the  $(KLLK)_4$  peptide we ran single peptide folding simulations as described in the Methods section. A plot showing the number of  $\alpha$ -helical hydrogen bonds (out of a possible 12) vs reduced temperature for our simulation data is shown in Figure 4.2. A single  $(KLLK)_4$  peptide was cooled from  $T^*=0.30$  to 0.10. From  $T^*=0.30$  to 0.20 the peptide remains in a random coil secondary structure with no  $\alpha$ -helical hydrogen bonds. When the temperature reaches  $T^*=0.17$ , on average only a single  $\alpha$ -helical hydrogen bond forms in the peptide. Further decreasing the temperature promotes  $\alpha$ -helix formation until the peptide is cooled to  $T^*=0.16$  where the peptide is 50% folded; this is the folding transition for the  $(KLLK)_4$  peptide. The

rapid increase in the number of  $\alpha$ -helical hydrogen bonds with decreasing temperature continues until the temperature is reduced to  $T^*=0.15$  where on average the peptide has ten  $\alpha$ -helical hydrogen bonds. Additional cooling from  $T^*=0.15$  to 0.10 only provides a modest increase of  $\sim 2$   $\alpha$ -helical hydrogen bonds.

Quantifying the number of  $\alpha$ -helical hydrogen bonds, which are also the number of native contacts of the folded peptide, lets us determine optimal temperatures at which to perform our other simulations. Selecting a temperature that allows a minimal amount of  $\alpha$ -helical hydrogen bonds to form is ideal for simulations where the objective is to maintain a random coil secondary structure. Similarly selecting a simulation temperature that allows the peptide to have almost purely  $\alpha$ -helical structure is ideal for keeping the peptide folded. Using this logic  $T^*=0.15$  is ideal for simulations that promote  $\alpha$ -helical secondary structure and  $T^*=0.17$  will maintain a primarily random coil secondary structure.

### 4.3.2 Single peptide interface simulations

To begin our analysis of the effects of an aqueous/organic interface on the  $(KLLK)_4$  peptide we wanted to know how the presence of the interface alters the peptide's secondary structure. Figure 4.3 shows the number of native contacts (number of  $\alpha$ -helical hydrogen bonds) vs the z-coordinate of the center of mass of the peptide initially in a random coil secondary, structure ( $T^*=0.17$ ) at interface strengths 0.50 (green) and 1.00 (orange) and for a peptide initially in an  $\alpha$ -helical secondary structure, ( $T^*=0.15$ ) at interface strengths 0.50 (purple) and 1.00 (pink). The values from the positive and negative sides of the simulation box were averaged for clarity with  $z=0\text{\AA}$  being the center of the box. Recall that the

simulations with a random coil initial configuration were run at a reduced temperature  $T^*=0.17$  to ensure that the peptide would remain in a random coil in the aqueous phase. From  $0\text{\AA}$  to  $\sim 80\text{\AA}$  the peptide has one to two native contacts indicating that the secondary structure was indeed a random coil. As the peptide moves beyond  $80\text{\AA}$  and approaches the strong interface at  $94.5\text{\AA}$ , the number of native contacts remains at about two (orange curve Figure 4.3). The lack of change in the number of native contacts tells us that the aqueous and interfacial behavior of this peptide are very similar. As the peptide moves beyond  $80\text{\AA}$  and approaches the weak interface at  $94.5\text{\AA}$  the number of native contacts gradually increases to a maximum of eleven (green curve Figure 4.3). The increase in the number of native contacts tells us that the peptide is folding into its native state at the interface. Simulations with an  $\alpha$ -helical initial configuration were run at a reduced temperature  $T^*=0.15$  to ensure that the peptide would remain helical in the aqueous phase. From  $0\text{\AA}$  to  $\sim 80\text{\AA}$  the peptide has nine to eleven native contacts indicating that the secondary structure does remain helical. As the peptide moves beyond  $80\text{\AA}$  and approaches the strong interface at  $94.5\text{\AA}$  the number of native contacts drops to about two (pink curve Figure 4.3). The rapid decrease in the number of native contacts indicates that the peptide is unfolding from its native helical structure to either a random coil or another conformation that is not the native folded state of the peptide. As the peptide moves beyond  $80\text{\AA}$  and approaches the weak interface at  $94.5\text{\AA}$  the number of native contacts remains at about eleven indicating that it remains folded (purple curve Figure 4.3). These single peptide simulations indicate that the strong interface creates an environment that promotes the unfolding of the peptide at the interface into a conformation

other than its native  $\alpha$ -helix while the weak interface allows the peptide to adopt its native state  $\alpha$ -helix at the interface.

Next we examine the end-to-end distance of the (KLLK)<sub>4</sub> peptide. The end-to-end distance, measured in angstroms, is the distance from the N-terminal to the C-terminal of the peptide. Figure 4.4 shows the end-to-end distance of the peptide vs the z-coordinate of the center of mass of the peptide for an initial random coil secondary structure,  $T^*=0.17$ , with interface strength 0.50 (green) and 1.00 (orange) and an initial  $\alpha$ -helical secondary structure,  $T^*=0.15$ , with interface strength 0.50 (purple) and 1.00 (pink). The values from the positive and negative sides of the simulation box were averaged for clarity with  $z=0\text{\AA}$  being the center of the box. Peptides starting in a random coil with both interface strengths (orange and green curves in Figure 4.4) have an end-to-end distance of  $\sim 20\text{\AA}$  in the aqueous phase from  $0\text{\AA}$  to  $80\text{\AA}$ . For both interface strengths as the peptide approaches the interface at  $94.5\text{\AA}$  the end to end distance increases to  $\sim 30\text{\AA}$  at  $z=85\text{\AA}$  and ultimately drops back down to  $24\text{\AA}$  at the interface for the peptide interacting with a weak interface (green curve) and  $16\text{\AA}$  with a slight jump to  $24\text{\AA}$  for the strong interface (orange curve). The end-to-end distance of a random coil peptide in the aqueous phase was  $20\text{\AA}$ . From Figure 4.1 we know that the random coil peptides interacting with a weak interface (green curve) fold into an  $\alpha$ -helix, so an end-to-end distance of  $24\text{\AA}$  corresponds to an  $\alpha$ -helical structure. The peptides in a random coil interacting with the strong interface (orange curve) had minimal native contacts according to Figure 4.1 so we believe the  $16\text{\AA}$ - $24\text{\AA}$  end-to-end distance distribution conveys a range of range of random coiled structures the (KLLK)<sub>4</sub> peptide can adopt. Peptides starting in an  $\alpha$ -helix with both interface strengths (purple and pink curves in Figure 4.4)

have an end to end distance of  $\sim 24\text{\AA}$  in the aqueous phase from  $0\text{\AA}$  to  $80\text{\AA}$ . The end-to-end distance remains the same as the peptide approaches the weak interface indicating that it remains helical, but as the peptide approaches the strong interface the end-to-end distance decreases to  $5\text{\AA}$ . These values are much different from both the helical and random coil values that we have already observed, indicating that a different secondary structure is forming. The very short end-to-end distance means that the N-terminal and C-terminal of the peptide are right next to each other, indicating that the peptide is in a very compact conformation. Upon further inspection, it appears that peptides in these simulations ultimately adopt a  $\beta$ -hairpin secondary structure. Similar trends were observed in the peptides radius of gyration shown in Figure S4.1.

Another important measure of how an aqueous/organic interface affects a peptide is the residency time of the peptide at the interface. Since the peptides take a random path in the simulation, it is possible that a peptide might not interact with the interface during a given time window. The residency times gives a good indication of how often the peptide interacts with the interface, and how long it stays there. In a plot of cumulative residency time vs reduced time, a positive slope indicates that the peptide interacts with the interface over a period of time and a zero slope indicates that the peptide is in the aqueous phase. Figure 4.5 shows a plot of the cumulative residency time vs reduced time for an initial random coil secondary structure,  $T^*=0.17$ , with interface strengths 0.50 (A) and 1.00 (B) and an initial  $\alpha$ -helical secondary structure,  $T^*=0.15$ , with interface strengths 0.50 (C) and 1.00 (D). Simulations with a weak interface behave very differently from simulations with a strong interface. In 9 out of 10 runs with a strong interface (Figures 4.5B and 4.5D), as soon as the

peptide encounters the interface it is irreversibly adsorbed. In the other simulation, the peptide simply never reached the interface. In simulations with a weak interface, the peptides have the ability to adsorb reversibly to the interface, as indicated by the changing slope of the cumulative residency time in Figure 4.5A and 4.5C. In simulations with a random coil starting conformation and a weak interface, the peptides have only one or two interaction events with the interface while the simulations that start in a helical conformation typically have many interaction events. The physical explanation for the difference in residency time events lies in the peptides' hydrophobic moment. The hydrophobic moment is a vector that describes the direction and magnitude of the peptide's hydrophobic character.<sup>32,33</sup> Since the helical peptide is amphipathic, the hydrophobic leucine and hydrophilic lysine residues are already partitioned onto separate sides of the peptide, creating a large hydrophobic moment pointing in the direction of all the leucine residues. When the orientation of the helical peptide is such that the hydrophobic moment is facing the organic phase when the peptide encounters the interface, the interaction is very favorable and the peptide can easily interact with the interface, increasing its residency time. A random coil peptide will have its hydrophobic leucine and hydrophilic lysine residues alternating along the backbone making the hydrophobic moment negligible. With no distinct hydrophobic moment it is unlikely that a randomly coiled (KLLK)<sub>4</sub> peptide will have an energetically favorable interaction with the interface, decreasing its residency time.

To compare the stability of the conformations adopted by the (KLLK)<sub>4</sub> peptide at the interface, we analyze their interpeptide hydrophobic interactions and intra-peptide hydrogen bonds. In Figure 4.6 we show the conformations that the (KLLK)<sub>4</sub> peptide adopts at the

interface A)  $\alpha$ -helix, B)  $\beta$ -hairpin, and C) random coil conformations with the interface location shown as a dotted line. The  $\alpha$ -helical conformation was adopted when the peptide was adsorbed to the weak interface, while the  $\beta$ -hairpin and random coil structures were adopted when the peptide was adsorbed to the strong interface. The  $\alpha$ -helix forms when the peptide is adsorbed to the weaker interface because the interface is too weak to disrupt the  $\alpha$ -helical hydrogen bonds, but strong enough to partition the side chains into each phase. The random coil and  $\beta$ -hairpin structures form when the peptide is adsorbed to the strong interface because the strong interface is strong enough to disrupt the  $\alpha$ -helical hydrogen bonds. Since there are two structures that form at the strong interface and each has all the side chains partitioned into their respective phases, we can determine which is the most energetically favorable by looking at their intrapeptide hydrogen bonds. The random coil peptide has no hydrogen bonds and the  $\beta$ -hairpin has up to 7 hydrogen bonds centered around the 9<sup>th</sup> lysine residue meaning that the  $\beta$ -hairpin is a more energetically favorable interface structure than the randomly coil structure. From this, we can conclude that the most stable interfacial structure at the weak interface is the  $\alpha$ -helix, since it is the only interfacial structure at a weak interface, and the most stable structure for the strong interface is the  $\beta$ -hairpin.

Ultimately, we want to know if the peptide interacts with the interface because it is folded, or if the peptide folds because it is interacting with the interface. The answer to this question depends on two factors: the secondary structure of the peptide and the strength of the interface. In our simulations the  $\alpha$ -helical (KLLK)<sub>4</sub> peptide is more likely to interact with the interface because it has a large hydrophobic moment while the randomly coil

conformation is in an “interface protected” state because its hydrophobic moment is negligible. We also found that a peptide could fold when interacting with the weak interface, but not the strong interface. The weak interface provides just enough of a driving force to promote the folding of the (KLLK)<sub>4</sub> peptide into an  $\alpha$ -helix, while the strong interface disrupts the  $\alpha$ -helical hydrogen bonds and pulls the structure into a random coil or  $\beta$ -hairpin. If the peptide has a random coil structure it folds because it interacts with the weak interface but if it a helical structure it interacts with the interface because it is folded.

### 4.3.3 Multi-peptide simulations

Now that we have quantified the structural changes and adsorption behavior of a single peptide at an aqueous/organic interface, we examine how the interface affects the aggregation of multiple peptides. Figure 4.7 shows (A) the fraction of peptides in oligomers of any type, (B) the fraction of peptides in helical bundles and (C) the fraction of peptides in  $\beta$ -sheets vs reduced time for peptides that initially have random coil secondary structures at interface strengths 0.0 (gray), 0.50 (green) and 1.00 (orange), and peptides that initially have helical secondary structures at interface strengths 0.0 (yellow) 0.50 (purple) and 1.00 (pink). By interface strength 0.0, we mean there is no interface present. In order to establish “baseline” aggregation data we ran simulations in the absence of an interface at peptide concentrations equivalent to those in the simulations with an interface as described in the Methods section. In these simulations, peptides starting in an  $\alpha$ -helix begin forming helical bundles immediately (yellow curve Figure 4.7B), while simulations with peptides in a random coil immediately began forming  $\beta$ -sheets (grey curve Figure 4.7C). This occurrence

is logical because the peptides are already in a conformation that predisposes them to the particular oligomer type. Peptides in an  $\alpha$ -helix only need to have leucine-leucine interactions to form a bundle, while random coil peptides only need to form backbone hydrogen bonds to turn into  $\beta$ -sheets. By  $t^*=25$ , simulations with a helical starting configuration (yellow curve Figure 4.7B) have the same fraction of peptides in  $\beta$ -sheets as the simulations starting in random coils (grey curve Figure 4.7C) and by  $t^*=75$  simulations with a random coil starting configuration have the same fraction of peptides in helical bundles as the simulations starting in helices (Figure 4.7B). After  $t^*=75$  these two simulations have approximately the same oligomer content (Figure 4.7A). Despite the different types of oligomers formed in these two types of simulations, their total oligomer content ends up being the same. Since the total oligomer content is the same we can conclude that the temperature selected for our multipolypeptide simulations,  $T^*=0.165$ , provides a balanced formation of helical bundles and  $\beta$ -sheets.

In the interfacial simulations with a helical starting conformation, there are two different sets of behavior. With the strong interface and peptides starting in an  $\alpha$ -helix (pink curve Figure 4.7A) the total number of peptides in oligomers is the same as in simulations without an interface (grey and yellow curves Figure 4.7A) until  $t^*=100$ . After  $t^*=100$  the simulations with an interface begin to have more peptides in oligomers than the simulation without an interface. The number of peptides in helical bundles for peptides starting in an  $\alpha$ -helix with a strong interface (pink curve Figure 4.7B) is identical to the number of peptides starting in an  $\alpha$ -helix in simulations without an interface (yellow curve Figure 4.7B) but there is a large deviation in the number of peptides in  $\beta$ -sheets beginning at  $t^*=100$ . After  $t^*=100$

the number of peptides in  $\beta$ -sheets for peptides starting in an  $\alpha$ -helix and a strong interface (pink curve Figure 4.7C) is higher than the corresponding number of peptides in simulations with an interface (Figure 4.7C grey and yellow curves). The increased  $\beta$ -sheet formation in the presence of the strong interface occurs because the helical peptides unfold into a  $\beta$ -hairpin or random coil structure at the interface. Once there are multiple peptides in these two conformations at the interface, they become highly susceptible to aggregating into  $\beta$ -sheets. The weak interface simulation with peptides starting in  $\alpha$ -helices (purple curve Figure 4.7A) has the highest number of peptides in oligomers, and all are helical bundles. As we observed in the single peptide simulations, helical peptides can interact with the weak interface and dissociate with the helix intact. This type of interaction leads to an increased local concentration of helices around the interface, promoting the formation of helical bundles.

In the simulations with a random coil starting conformation and an interface, the total number of peptides in oligomers (green and orange curves Figure 4.7A) is lower than at all other conditions. The weak interface with random coil starting configuration (green) has more total peptides in oligomers than the strong interface (orange) until  $t^*=175$  where they are equivalent. Prior to  $t^*=200$  the weak interface simulations have more peptides in helical bundles than the strong interface (Figure 4.7B) and less  $\beta$ -sheets than the strong interface simulations until  $t^*=50$  (Figure 4.7C). The number of helical bundles in both cases is much lower than in the simulations that have a helical starting configuration because the peptides need to fold before they can form bundles. The weak interface simulations have more peptides in helical bundles because the peptide can fold at the interface and dissociate. In this way, the interface accelerates single helix formation so that bundles are easier to form.

Ultimately, the simulations with a strong interface have more  $\beta$ -sheets because the peptides are pulled into the interface and remain in conformations susceptible to  $\beta$ -sheet formation ( $\beta$ -hairpins and random coils).

To compare the stability of helical bundles and  $\beta$ -sheets oligomers it is beneficial to estimate the energies of each structure based on the numbers of hydrophobic interactions and intra-peptide hydrogen bonds. First we consider the types of oligomers formed in the aqueous phase. The helical bundle is an oligomer formed when multiple helices have leucine-leucine side chain interactions and is only formed in the aqueous phase. Here we will consider a tetrameric helical bundle composed of four helices which is the ground state for the bundle structure. Each helix has 12  $\alpha$ -helical hydrogen bonds for a total of 48  $\alpha$ -helical hydrogen bonds and 8 leucine side chain beads for a total of 32 leucine side chain beads. A leucine-leucine side chain interaction is  $\sim 3\text{\AA}$  in the PRIME20 force field so we will assume that the leucine side chains of a given helix interact only with the adjacent helices in the bundle resulting in 32 leucine-leucine side chain interactions. Next we consider a  $\beta$ -sheet formed in the aqueous phase. An ideal 4-peptide  $\beta$ -sheet will have a total of 32 inter-peptide hydrogen bonds with 8 inter-peptide hydrogen bonds between each peptide, and the remaining 16 hydrogen bonding sites extending out from the edges of the  $\beta$ -sheet. Next, we consider a random coil peptide in the aqueous phase. In the aqueous phase, randomly coiled peptides tend to form  $\beta$ -sheets that have alternating leucine and lysine residues on a given side of the  $\beta$ -sheet. Although there are likely leucine-leucine, leucine-lysine, and lysine-lysine interactions in a  $\beta$ -sheet, only the leucine-leucine interactions decrease the energy of the configuration. The magnitude of the leucine-leucine interaction is much smaller than the

other energetically unfavorable pair interactions so the backbone hydrogen bonds are the largest contribution to the energy of the  $\beta$ -sheet. Based on this discussion we conclude that in the aqueous phase the helical bundles have more hydrogen bonds and a larger number of favorable leucine-leucine side chain interactions so we can conclude that the helical bundle is the most energetically favorable aggregate in the aqueous phase. At the interface only  $\beta$ -sheets form. The inter-peptide hydrogen bonds remain the same as in the aqueous phase, but the backbone is twisted in such a way that all the leucines are toward the organic phase and lysines are toward the aqueous phase. This rearrangement occurs to minimize the energy of the structure at the interface making  $\beta$ -sheets the only energetically favorable oligomer at the interface.

Finally, we would like to address the question: does peptide aggregation in the aqueous phase cause interaction with the interface, or do the peptides have to interact with the interface before they aggregate. Aggregation of the (KLLK)<sub>4</sub> peptide in the aqueous phase does not cause interaction with the interface, in fact it reduces aggregation at the interface. All of the aggregates that form in the aqueous phase have a negligible hydrophobic moment, putting them in an “interface-protected” state. However, when the (KLLK)<sub>4</sub> peptide interacts with the interface it forms stable  $\beta$ -sheet oligomers. Our simulation shows that the (KLLK)<sub>4</sub> peptide does not need to interact with the interface to aggregate because aggregation occurs at in both the aqueous phase and at the interface.

## 4.4 Discussion and Conclusions

Using the combination of DMD and our PRIME20 force field we were able to simulate the interaction of a single (KLLK)<sub>4</sub> amphipathic peptide and of 48 (KLLK)<sub>4</sub> peptides with an aqueous/organic interface. Many simulations of peptides at interfaces have been performed by other investigators but, to our knowledge, none have observed the spontaneous folding of amphipathic  $\alpha$ -helical peptides at interfaces or the aggregation of proteins at interfaces for such large systems. The peptides considered in the work are initially in random coil or  $\alpha$ -helical conformations and interact with strong and weak interfaces. Our results show that the weak interface is capable of folding a randomly coiled (KLLK)<sub>4</sub> peptide into an amphipathic  $\alpha$ -helix, while the strong interface disrupts intrapeptide hydrogen bonding and forces the peptides to adopt a random coil or  $\beta$ -hairpin conformation. We also found that when peptides interacted with the strong interface they adsorbed irreversibly due to the strong side chain preference for their respective phases. When peptides interacted with the weak interface they adsorbed reversibly. In the latter case the (KLLK)<sub>4</sub> peptide was more likely to bind to the interface when it was initially in an  $\alpha$ -helical conformation as opposed to a random coil because it had a larger hydrophobic moment. We also studied the aggregation of the (KLLK)<sub>4</sub> peptide in the presence of an aqueous/organic interface. We found that the conditions producing the highest levels of aggregation occurred when the peptides started out in helical conformations and interacted with a weak interface. Since the peptides could bind to the interface and reversibly dissociate, the increased local concentration of helices in the vicinity of the interface promoted the formation of helical bundles. The most energetically favorable type of aggregate in the

aqueous phase was the helical bundle, and the most energetically favorable type of aggregate at the interface was the  $\beta$ -sheet.

Our simulation results agree well with both the simulation and experimental literature we presented in the introduction. Chipot *et al.* found that both the polyleucine and the amphipathic LQQLLQQLQL peptide preferred to adopt a parallel conformation at the interface rather than perpendicular, with either the N or C terminal contacting the interface.<sup>19,20</sup> This is also what we observed and, based on our results, the partitioning of the leucine side chains into the organic phase is the driving force behind this preference. Even peptides in simulations with the weak interface preferred the parallel conformation to perpendicular because the additional 8 side chain-organic phase interactions reduced the energy of the peptide. Although Chipot *et al.* were unable to simulate the folding of the LQQLLQQLQL peptide at a water/hexane interface, we did observe the folding of the (KLLK)<sub>4</sub> peptide into its lowest energy  $\alpha$ -helical conformation at the interface. Our simulation results are also consistent with their conclusion that once the peptide has adsorbed to the interface it usually remains in some amphipathic conformation because the energetic penalties associated with transitioning to non-amphipathic conformations are too great. In an independent simulation not presented here we examined the trajectory of a (KLLK)<sub>4</sub> peptide folding at the interface and found that it does not actually fold while the majority of the peptide is adsorbed. Instead, it folds into the amphipathic helix as it is gradually pulled into the interface, bypassing the need to go through non-amphipathic intermediates while the peptide is adsorbed to the interface. To our knowledge, this is the first observation of this behavior.

Anderson *et al.* determined that the most probably conformation occupied by their 27-residue heteropeptide was when it was adsorbed to the interface, but unfolded.<sup>21</sup> This is also the most frequently occurring conformation in our simulations. We were also able to produce conditions that lead to reversible adsorption and irreversible adsorption. In preliminary work, we found conditions that lead to no adsorption, but they were not included in this paper because we deemed them physically unrealistic.

Maget-Dana *et al.* examined a peptide with the sequence (LKKL)<sub>n</sub>, which is the almost the same as the peptide used in our simulations, (KLLK)<sub>4</sub>, but they examined an air/water interface rather than our aqueous/organic interface.<sup>23</sup> They found that the  $\alpha$ -helical (LKKL)<sub>n</sub> peptide diffused through the aqueous phase more quickly and had a higher probability of adsorbing to the interface than the (LK)<sub>n</sub> peptide. In our discussion on the residency time of the (KLLK)<sub>4</sub> peptide at the interface we also came to the same conclusion. We believe that the amphipathic helix has a higher probability of interacting with the interface than randomly coiled peptides because it has a large hydrophobic moment. While the (LK)<sub>n</sub> peptide is amphipathic, it is much more flexible than a folded (LKKL)<sub>n</sub> peptide, likely leading to a less prominent hydrophobic moment and making it harder for the peptide to interact with the interface. We are also considering simulating a (KL)<sub>8</sub> peptide but reserve it for future work. Maget-Dana *et al.* also found that interfacial  $\beta$ -sheets were much more stable than interfacial  $\alpha$ -helices. We also found that once a  $\beta$ -sheet structure was formed at an interface it was adsorbed irreversibly while  $\alpha$ -helices could dissociate. Wang *et al.* identified two classes of  $\alpha$ -helical peptides: those that bind strongly to interfaces and those that bind weakly.<sup>24</sup> We believe that the (KLLK)<sub>4</sub> peptide falls into the weak binding category because

of its ability to reversibly bind to interfaces and the fact that it is only a single amphipathic helix. The strong binding example in their paper contained multiple amphipathic  $\alpha$ -helices within a single protein so if one helix dissociated, the other may remain at the interface making it difficult for the peptide to fully dissociate. They also found that interfacial  $\beta$ -sheets were adsorbed irreversibly, confirming our results.

## **4.5 Acknowledgements**

This work was supported by the National Institutes of Health, USA, under grant GM056766 and EB006006 (C.K.H.). We gratefully acknowledge helpful discussions with Professor Ted Randolph at the University of Colorado.

## 4.6 References

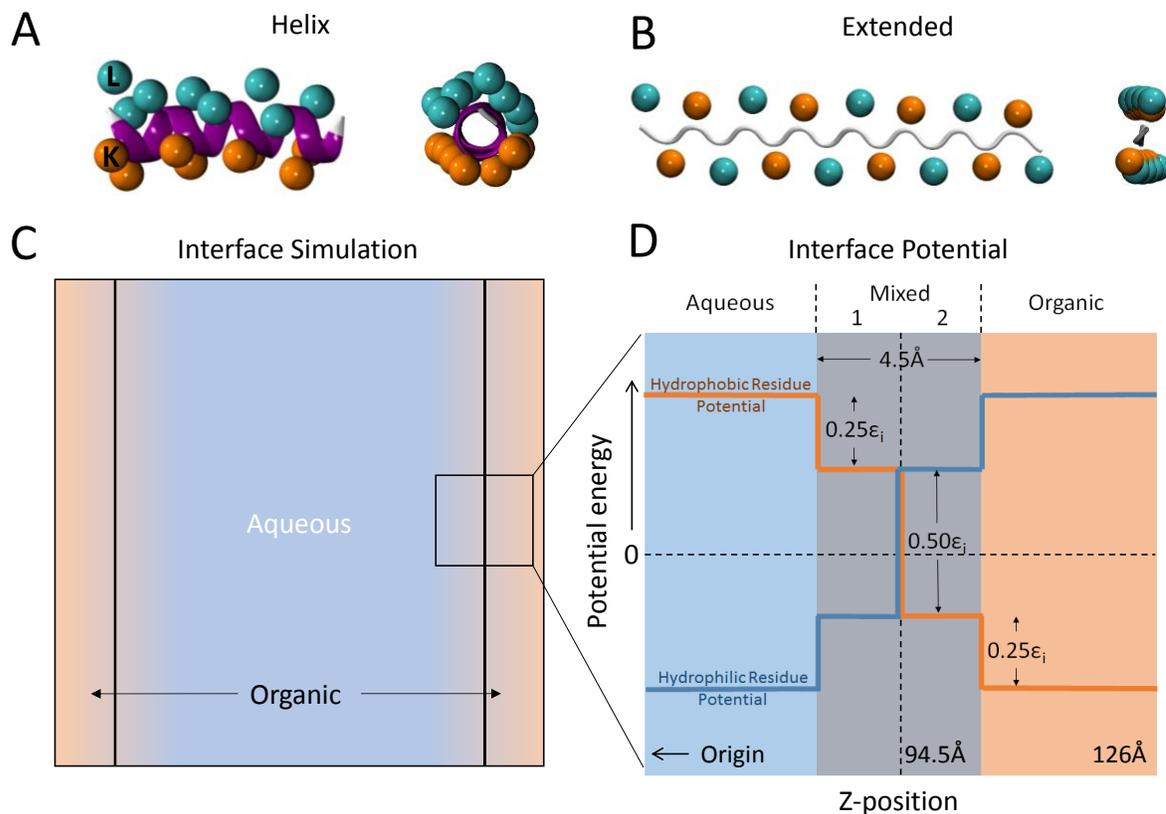
1. Wang W. Instability, stabilization, and formulation of liquid protein pharmaceuticals. *International Journal of Pharmaceutics (Amsterdam)*. 1999;185(2):129-188. doi: 10.1016/S0378-5173(99)00152-0 ER.
2. Jezek J, Darton NJ, Derham BK, Royle N, Simpson I. Biopharmaceutical formulations for pre-filled delivery devices. *Expert Opinion on Drug Delivery*. 2013;10(6):811-828. doi: 10.1517/17425247.2013.780023.
3. Kim HJ, Decker EA, McClements DJ. Role of postadsorption conformation changes of beta-lactoglobulin on its ability to stabilize oil droplets against flocculation during heating at neutral pH. *Langmuir*. 2002;18(20):7577-7583. doi: 10.1021/la020385u.
4. Fang Y, Dalgleish DG. Conformation of beta-lactoglobulin studied by FTIR: Effect of pH, temperature, and adsorption to the oil-water interface. *J Colloid Interface Sci*. 1997;196(2):292-298. doi: 10.1006/jcis.1997.5191.
5. Husband FA, Garrood MJ, Mackie AR, Burnett GR, Wilde PJ. Adsorbed protein secondary and tertiary structures by circular dichroism and infrared spectroscopy with refractive index matched emulsions. *J Agric Food Chem*. 2001;49(2):859-866. doi: 10.1021/jf000688z.
6. Lefevre T, Subirade M. Formation of intermolecular beta-sheet structures: A phenomenon relevant to protein film structure at oil-water interfaces of emulsions. *J Colloid Interface Sci*. 2003;263(1):59-67. doi: 10.1016/S0021-9797(03)00252-2.

7. Wang W. Protein aggregation and its inhibition in biopharmaceutics. *International Journal of Pharmaceutics (Kidlington)*. 2005;289(1-2):1-30. doi: 10.1016/j.ijpharm.2004.11.014 ER.
8. Gerhardt A, Bonam K, Bee JS, Carpenter JF, Randolph TW. Ionic strength affects tertiary structure and aggregation propensity of a monoclonal antibody adsorbed to silicone oil-water interfaces. *J Pharm Sci*. 2013;102(2):429-440. doi: 10.1002/jps.23408.
9. Li J, Pinnamaneni S, Quan Y, Jaiswal A, Andersson FI, Zhang X. Mechanistic understanding of protein-silicone oil interactions. *Pharm Res*. 2012;29(6):1689-1697. doi: 10.1007/s11095-012-0696-6.
10. Felsovalyi F, Janvier S, Jouffray S, Soukiassian H, Mangiagalli P. Silicone-oil-based subvisible particles: Their detection, interactions, and regulation in prefilled container closure systems for biopharmaceuticals. *J Pharm Sci*. 2012;101(12):4569-4583. doi: 10.1002/jps.23328.
11. Bee JS, Schwartz DK, Trabelsi S, *et al*. Production of particles of therapeutic proteins at the air-water interface during compression/dilation cycles. *Soft Matter*. 2012;8(40):10329-10335. doi: 10.1039/c2sm26184g.
12. Rudiuk S, Cohen-Tannoudji L, Huille S, Tribet C. Importance of the dynamics of adsorption and of a transient interfacial stress on the formation of aggregates of IgG antibodies. *Soft Matter*. 2012;8(9):2651-2661. doi: 10.1039/c2sm07017k.

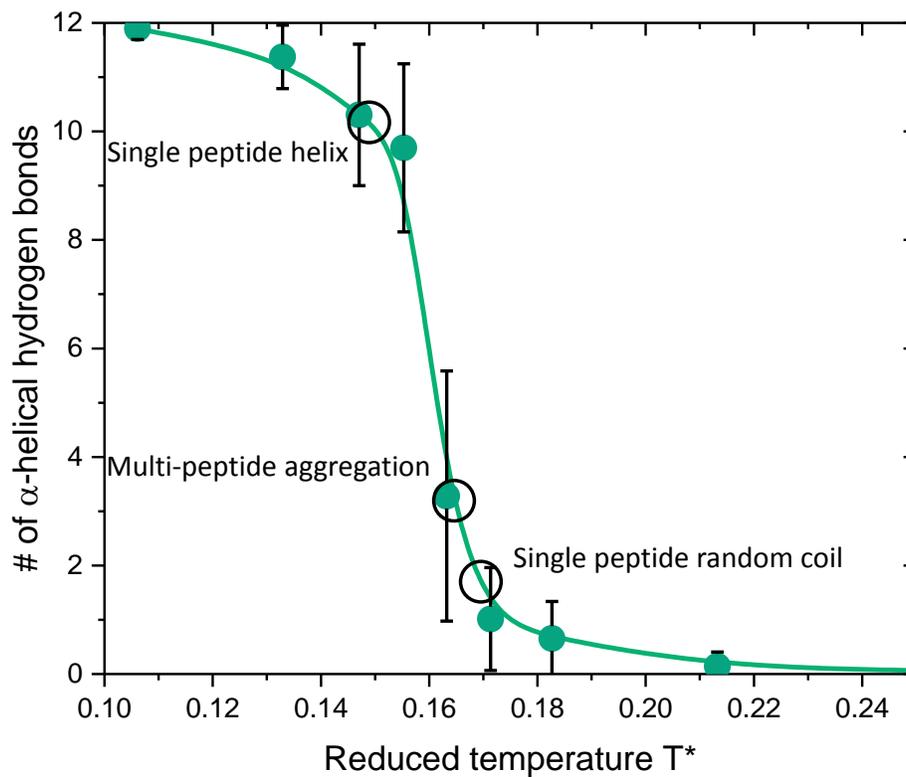
13. Badkar A, Wolf A, Bohack L, Kolhe P. Development of biotechnology products in pre-filled syringes: Technical considerations and approaches. *Aaps Pharmscitech*. 2011;12(2):564-572. doi: 10.1208/s12249-011-9617-y.
14. Majumdar S, Ford BM, Mar KD, Sullivan VJ, Ulrich RG, D'Souza AJM. Evaluation of the effect of syringe surfaces on protein formulations. *J Pharm Sci*. 2011;100(7):2563-2573. doi: 10.1002/jps.22515.
15. Gerhardt A, McGraw NR, Schwartz DK, Bee JS, Carpenter JF, Randolph TW. Protein aggregation and particle formation in prefilled glass syringes. *J Pharm Sci*. 2014;103(6):1601-1612. doi: 10.1002/jps.23973.
16. Hermeling S, Crommelin D, Schellekens H, Jiskoot W. Structure-immunogenicity relationships of therapeutic proteins. *Pharm Res*. 2004;21(6):897-903. doi: 10.1023/B:PHAM.0000029275.41323.a6.
17. Reddy K, Yedery R, Aranha C. Antimicrobial peptides: Premises and promises. *Int J Antimicrob Agents*. 2004;24(6):536-547. doi: 10.1016/j.ijantimicag.2004.09.005.
18. Bulet P, Stocklin R, Menin L. Anti-microbial peptides: From invertebrates to vertebrates. *Immunol Rev*. 2004;198:169-184. doi: 10.1111/j.0105-2896.2004.0124.x.
19. Chipot C, Pohorille A. Folding and translocation of the undecamer of poly-L-leucine across the water-hexane interface. A molecular dynamics study. *J Am Chem Soc*. 1998;120(46):11912-11924. doi: 10.1021/ja980010o.
20. Chipot C, Maigret B, Pohorille A. Early events in the folding of an amphipathic peptide: A multnanosecond molecular dynamics study. *Proteins-Structure Function and Genetics*. 1999;36(4):383-399.

21. Anderson R, Pande V, Radke C. Dynamic lattice monte carlo simulation of a model protein at an oil/water interface. *J Chem Phys.* 2000;112(20):9167-9185. doi: 10.1063/1.481537.
22. Leonhard K, Prausnitz J, Radke C. Three-dimensional lattice monte carlo simulations of model proteins. IV. proteins at an oil-water interface. *Langmuir.* 2006;22(7):3265-3272. doi: 10.1021/la052535h.
23. Maget-Dana R, Lelievre D, Brack A. Surface active properties of amphiphilic sequential isopeptides: Comparison between alpha-helical and beta-sheet conformations. *Biopolymers.* 1999;49(5):415-423. doi: 10.1002/(SICI)1097-0282(19990415)49:5<415::AID-BIP7>3.0.CO;2-J.
24. Wang LB, Atkinson D, Small DM. Interfacial properties of an amphipathic alpha-helix consensus peptide of exchangeable apolipoproteins at air/water and oil/water interfaces. *J Biol Chem.* 2003;278(39):37480-37491. doi: 10.1074/jbc.M303133200.
25. Wang L, Small D. Interfacial properties of amphipathic beta strand consensus peptides of apolipoprotein B at oil/water interfaces. *J Lipid Res.* 2004;45(9):1704-1715. doi: 10.1194/jlr.M400106-JLR200.
26. Alder BJ, Wainwright TE. Studies in molecular dynamics .1. general method. *J Chem Phys.* 1959;31(2). doi: 10.1063/1.1730376.
27. Wu C, Shea J. Coarse-grained models for protein aggregation. *Curr Opin Struct Biol.* 2011;21(2):209-220. doi: 10.1016/j.sbi.2011.02.002.

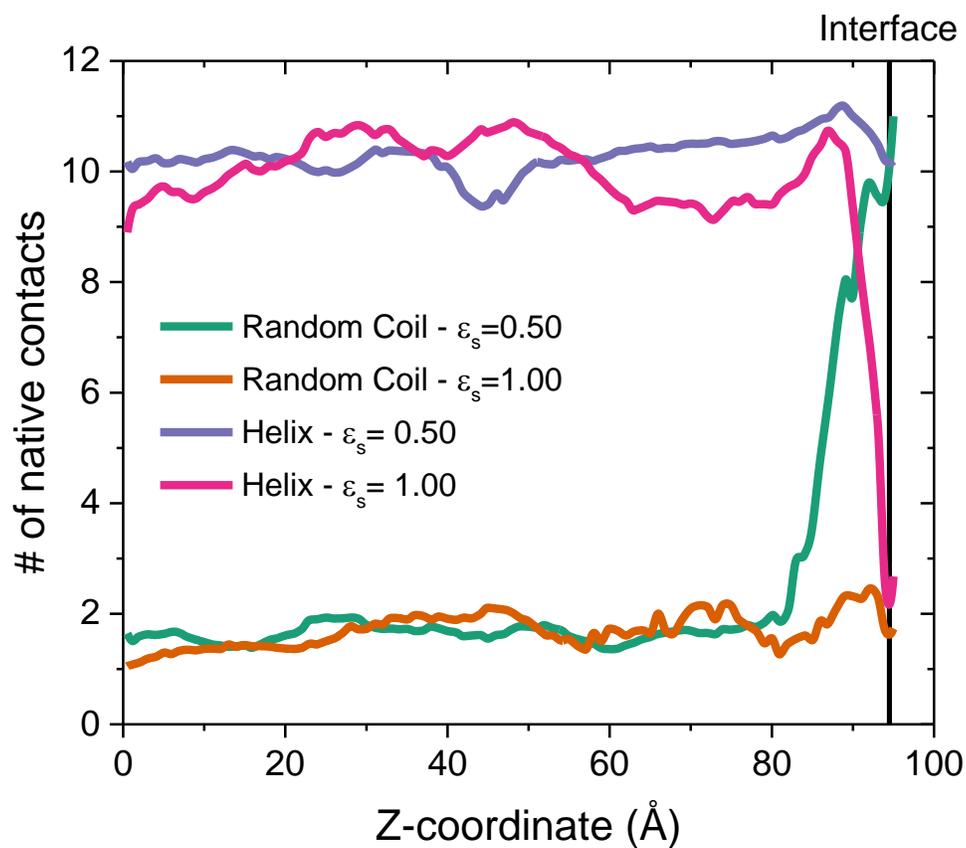
28. Smith A, Hall C. Alpha-helix formation: Discontinuous molecular dynamics on an intermediate-resolution protein model. *Proteins-Structure Function and Genetics*. 2001;44(3):344-360.
29. Smith A, Hall C. Assembly of a tetrameric alpha-helical bundle: Computer simulations on an intermediate-resolution protein model. *Proteins*. 2001;44(3):376-391. doi: 10.1002/prot.1103.
30. Ashbaugh HS, Pratt LR, Paulaitis ME, Clohery J, Beck TL. Deblurred observation of the molecular structure of an oil-water interface. *J Am Chem Soc*. 2005;127(9):2808-2809. doi: 10.1021/ja042600u.
31. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol*. 1982;157(1):105-132. doi: 10.1016/0022-2836(82)90515-0.
32. EISENBERG D, WEISS R, TERWILLIGER T, WILCOX W. Hydrophobic moments and protein-structure. *Faraday Symposia of the Chemical Society*. 1982(17):109-120.
33. Eisenberg D, Weiss RM, Terwilliger TC. The helical hydrophobic moment - a measure of the amphiphilicity of a helix. *Nature*. 1982;299(5881):371-374. doi: 10.1038/299371a0.



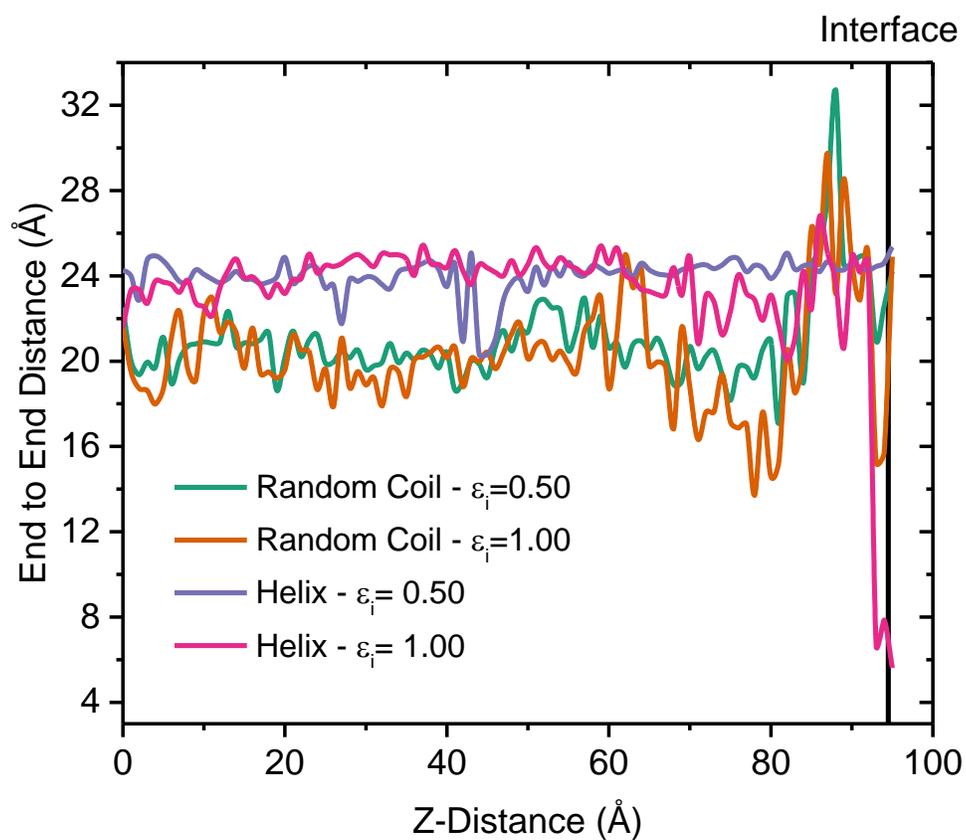
**Figure 4.1.** A) Folded native state  $\alpha$ -helix of  $(KLLK)_4$  peptide. B) Extended conformation of  $(KLLK)_4$  peptide in PRIME20. C) Interface simulation box. D) The discontinuous interface potentials experienced by hydrophobic leucine residues (orange) and hydrophilic lysine residue potential (cyan) at various distances ( $z$ -position) from the interface. The interface is located  $94.5\text{\AA}$  from the origin and the edge of the simulation box is  $126\text{\AA}$  from the origin.



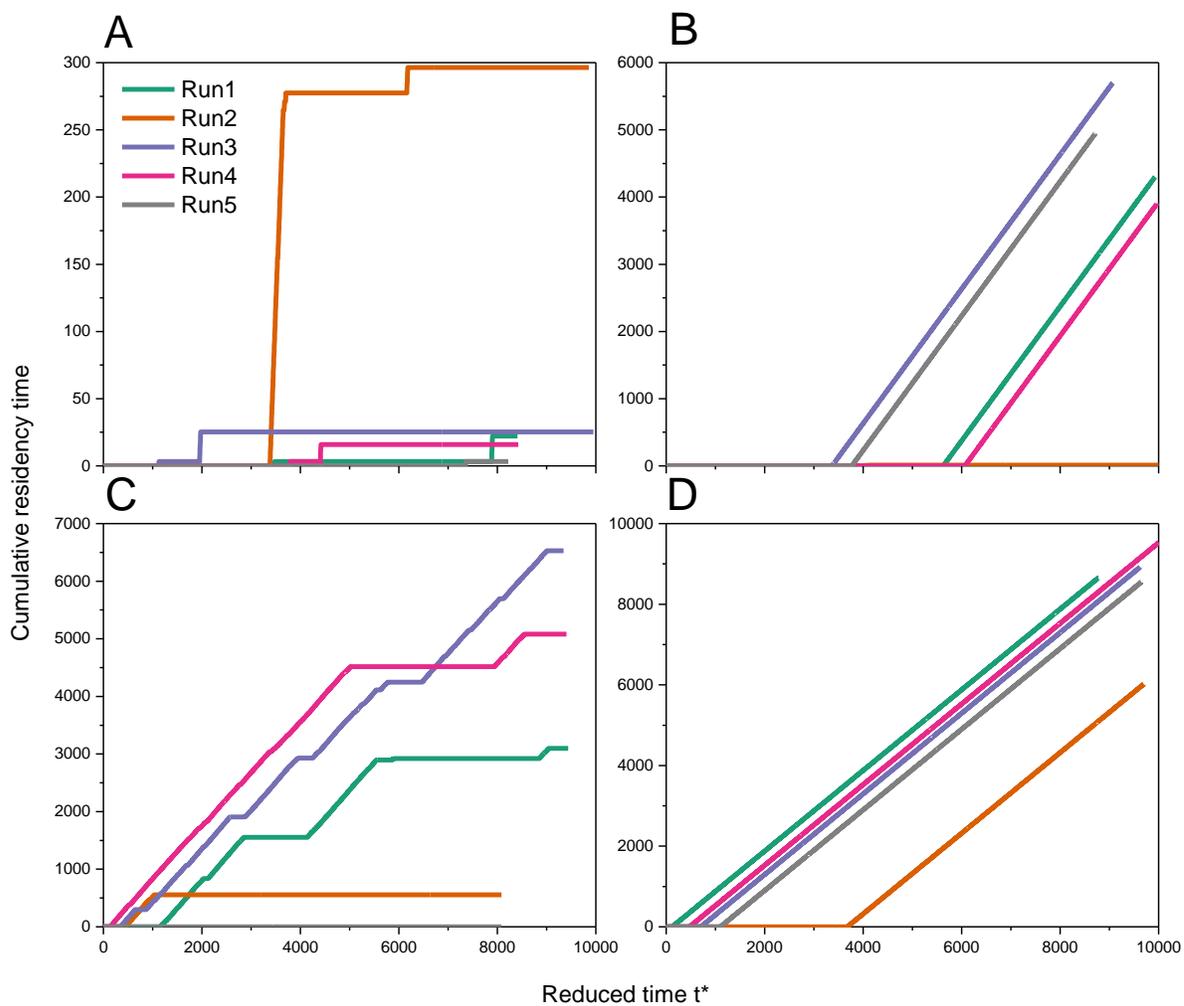
**Figure 4.2.** Number of  $\alpha$ -helical hydrogen bonds vs reduced temperature for the folding of a single (KLLK)<sub>4</sub> peptide. The temperatures used for each simulation type presented in this paper are circled for clarity: Single peptide random coil  $T^*=0.17$ , single peptide helix  $T^*=0.15$ , and multi-peptide aggregation  $T^*=0.165$ .



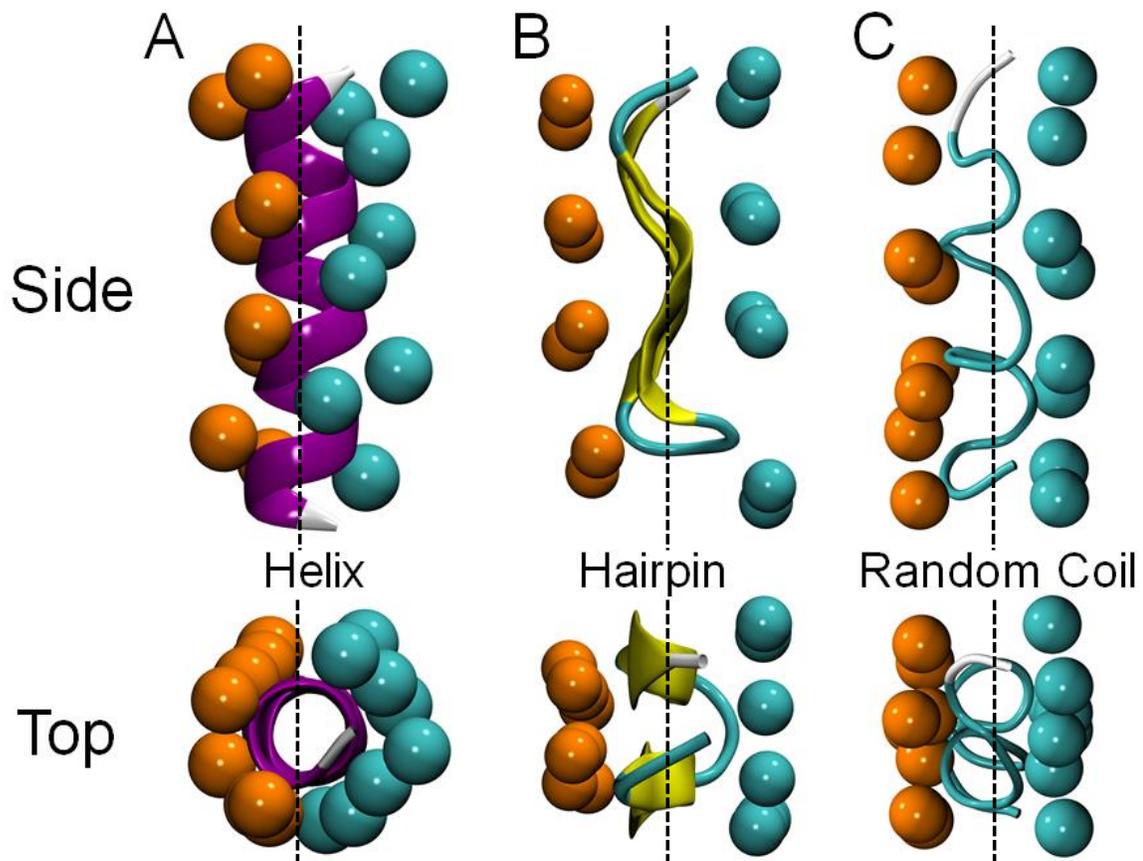
**Figure 4.3.** Number of native contacts ( $\alpha$ -helical hydrogen bonds) vs the z-coordinate of the center of mass of a single (KLLK)<sub>4</sub> peptide for an initial random coil secondary structure,  $T^*=0.17$ , at interface strengths 0.50 (green) and 1.00 (orange) and for an initial  $\alpha$ -helical secondary structure,  $T^*=0.15$ , at interface strengths 0.50 (purple) and 1.00 (pink).



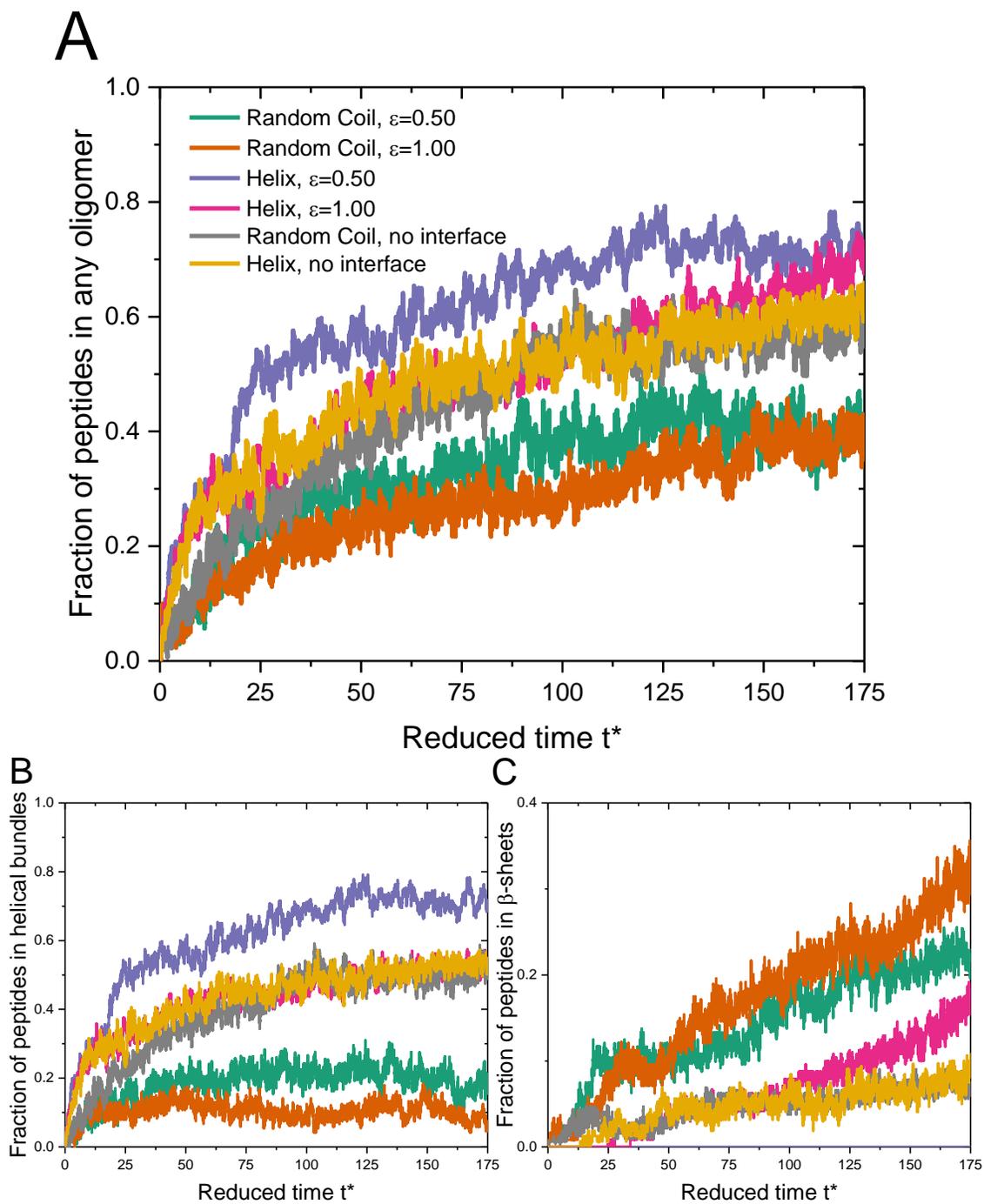
**Figure 4.4.** End to end distance vs the z-coordinate of the center of mass of a single (KLLK)<sub>4</sub> peptide for an initial random coil secondary structure,  $T^*=0.17$ , and interface strength 0.50 (green) and 1.00 (orange) and an initial  $\alpha$ -helical secondary structure,  $T^*=0.15$ , and interface strength 0.50 (purple) and 1.00 (pink).



**Figure 4.5.** Cumulative residency time of a single (KLLK)<sub>4</sub> peptide at an aqueous-organic interface vs reduced time for an initial random coil secondary structure,  $T^*=0.17$ , with interface strengths 0.50 (A) and 1.00 (B) and an initial  $\alpha$ -helical secondary structure,  $T^*=0.15$ , with interface strengths 0.50 (C) and 1.00 (D).



**Figure 4.6.** Interfacial structures adopted by the  $(KLLK)_4$  peptide: A)  $\alpha$ -helix, B)  $\beta$ -hairpin, and C) random coil. Interface location shown as dotted line.

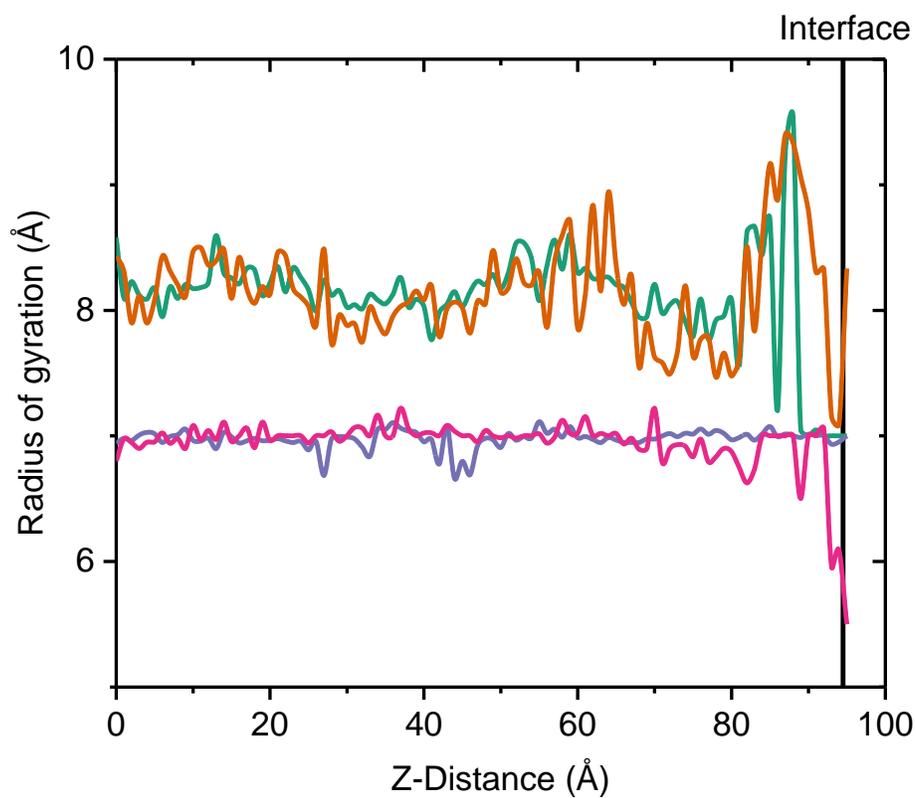


**Figure 4.7.** (A) Fraction of peptides in oligomers of any type, (B) the fraction of peptides in helical bundles and (C) the fraction of peptides in  $\beta$ -sheets vs reduced time for a peptides that initially have random coil secondary structures at interface strengths 0.0 (gray), 0.50 (green)

and 1.00 (orange), and peptides that initially have helical secondary structures at interface strengths 0.0 (yellow) 0.50 (purple) and 1.00 (pink).

**Table 4.1.** Lysine-lysine and leucine-leucine attractive square well (negative) and repulsive square shoulder (positive) interaction strengths as a function their phase location.

	<b>Aqueous Phase</b>	<b>Mixed Phase 1</b>	<b>Mixed Phase 2</b>	<b>Organic Phase</b>
<b>K-K</b> (reduced energy)	0.073	0.078	0.87	0.091
<b>L-L</b> (reduced energy)	-0.200	-0.150	-0.050	Hard Sphere



**Figure S4.1.** End to end distance vs the z-coordinate of the center of mass of a single (KLLK)<sub>4</sub> peptide for an initial random coil secondary structure,  $T^*=0.17$ , and interface strength 0.50 (green) and 1.00 (orange) and an initial  $\alpha$ -helical secondary structure,  $T^*=0.15$ , and interface strength 0.50 (purple) and 1.00 (pink).

## **CHAPTER 5**

### **Aggregation of Amyloid Beta in the Presence of Naturally Occurring Phenolic Inhibitors using Coarse-Grained Simulations**

Chapter 5 is essentially a manuscript by David C Latshaw II and Carol K Hall in preparation.

# Aggregation of Amyloid Beta in the Presence of Naturally Occurring Phenolic Inhibitors using Coarse-Grained Simulations

David C Latshaw II and Carol K Hall\*

Department of Chemical and Biomolecular Engineering, North Carolina State University,  
Raleigh, North Carolina

## Abstract

To examine the effect of naturally-occurring phenolic compounds on the aggregation of A $\beta$ (17-36), discontinuous molecular dynamics (DMD) simulations combined with an intermediate resolution protein model, PRIME20, were applied to a peptide/inhibitor system. Simulations were performed on systems containing 8 A $\beta$ (17-36) molecules and either 128 or 256 vanillin, resveratrol, curcumin, or epigallocatechin-3-gallate (EGCG) molecules. When simulated in the absence of inhibitors, the A $\beta$ (17-36) peptide was capable of forming the U-loop structure characteristic of A $\beta$  protofilaments. During the lag phase and initial aggregation, we find that resveratrol binds most frequently with A $\beta$ , followed by curcumin, and finally vanillin. We also observe that curcumin and resveratrol bind in the middle of aggregates, while vanillin binds the exterior, possibly preventing further aggregation into ordered amyloid structures. Limited simulations on EGCG show that it is capable of binding monomers unlike the other molecules.

## 5.1 Introduction

Alzheimer's Disease (AD) is a neurodegenerative disease that causes dementia, nervous system degradation, and death. Currently there are no therapeutic agents available for the treatment of AD despite great effort from the research community.<sup>1-3</sup> The pathological hallmark of AD is the aggregation of the amyloid  $\beta$  (A $\beta$ ) peptide. A $\beta$  monomers, which result from the enzyme-based cleavage of the amyloid precursor protein, APP, assemble extracellularly into progressively larger aggregate structures, monomers to oligomers to fibrils, and are eventually deposited in plaques in the brain. One promising avenue for AD treatment is based on identifying small molecule aggregation inhibitors that could prevent or reverse the aggregation of the A $\beta$  peptide. Many small molecules have been identified<sup>4</sup> or designed<sup>5</sup> as candidates for the treatment of AD. These molecules have been shown to act through a variety of mechanisms such as scavenging free radicals created by reactive oxygen species or reducing inflammatory responses associated with AD.<sup>6-12</sup> Another treatment mechanism provided by inhibitors is their ability to interact with amyloid structures so as to prevent and even reverse the aggregation of A $\beta$ . A subset of these molecules, phenolic compounds, has been shown to be very effective at preventing the formation of A $\beta$  amyloidogenic structures.<sup>12-14</sup> In this paper we use computer simulations to learn how the naturally-occurring phenolic compounds resveratrol, vanillin, curcumin, and epigallocatechin-3-gallate (EGCG) affect A $\beta$ (17-36) oligomer and fibril formation.

Many researchers have investigated the effectiveness of resveratrol<sup>9,10,15-20</sup>, vanillin<sup>21-23</sup>, curcumin<sup>7,8,19-22,24-28</sup>, and EGCG<sup>6,19,29-36</sup> as AD therapeutics. Resveratrol is a naturally-occurring compound found in the skin of red grapes and is often consumed in the form of red

wine. Although its concentration in red wine is about ten times lower than the amount shown to inhibit A $\beta$  aggregation *in vitro*<sup>18</sup>, Wang *et al.* showed that moderate consumption of Cabernet Sauvignon could also promote the clearance of A $\beta$  in transgenic mice.<sup>17</sup> Ladiwala *et al.* tested the ability of resveratrol to remodel five types of A $\beta$ (1-42) structures *in vitro*: monomers, soluble oligomers, non-toxic oligomers, fibrillar intermediates, and amyloid fibrils. They found that resveratrol was able to modify soluble oligomers, fibrillar intermediates, and amyloid fibrils into an off pathway structure that was disordered, non-toxic, and high in molecular weight. They suggested that resveratrol disassembles fibrillar structures by interacting with their  $\beta$ -sheet structures, residue stacking, and aromatic side chains. Circular dichroism measurements indicated that the soluble oligomer structure was closer to that of a monomer than the ordered, in-register,  $\beta$ -sheets of fibrils, yet resveratrol also remodeled these structures. They proposed that resveratrol might actually change these soluble oligomers to  $\beta$ -sheet like structures before disassembling them. Feng *et al.* tested the effects of resveratrol on A $\beta$ (1-42) aggregation *in vitro* and found that it could inhibit A $\beta$  fibril formation, disaggregate preformed fibrils, and inhibit oligomer cytotoxicity, but it could not inhibit the formation of the oligomers themselves which were instead stabilized.<sup>20</sup>

Vanillin is the principal component found in the extract from vanilla bean. It is much smaller than the other phenolic compounds considered in this paper and it shares aromatic substitution patterns with curcumin. The effectiveness of vanillin as an A $\beta$  aggregation inhibitor is less clear than other compounds considered in this paper, but it is included to clarify the necessity of having multiple aromatic groups in an A $\beta$  aggregation inhibitor. Papers from De Felice *et al.*<sup>37</sup> and Necula *et al.*<sup>4</sup> identify vanillin as capable of disrupting

both A $\beta$ (1-42) oligomer and fibril formation, while Ladiwala *et al.*<sup>23</sup> and Reinke and Gestwicki.<sup>21</sup> both indicate that vanillin has little to no effect on A $\beta$ (1-42) aggregation.

Curcumin is one of the principal components in the spice turmeric, known for its potent antioxidant and anti-inflammatory properties. It has been shown to disrupt A $\beta$  aggregation by direct interaction. Ono *et al.* found that curcumin was capable of inhibiting A $\beta$ (1-40) and A $\beta$ (1-42) fibril formation and destabilizing preformed fibrils.<sup>25</sup> They were unsure of the inhibition mechanism but proposed that curcumin might bind to the ends of fibrils destabilizing the structure and preventing further aggregation. Yang *et al.* also studied curcumin's effects on A $\beta$  and determined that it could inhibit formation of A $\beta$ (1-40) oligomers and fibrils, and prevent A $\beta$ (1-42) oligomer formation.<sup>26</sup> Although no direct evidence was presented, Yang *et al.* drew parallels to RS-0406, a small molecule selected by high throughput screening for inhibiting A $\beta$  aggregation<sup>5</sup>, suggesting that since curcumin and RS-0406 share similar aromatic groups and a hydrophobic linker perhaps their inhibitory mechanisms were similar. RS-0406 is a “ $\beta$ -sheet breaker” that inhibits amyloid fibril formation by disrupting  $\beta$ -sheet hydrogen bonds. Expanding on the idea of similarly structured aggregation inhibitors, Reinke and Gestwicki examined the effect of aromatic groups, aromatic group substitution patterns, and the length and flexibility of the chain linking each aromatic group on A $\beta$ (1-42) aggregation.<sup>21</sup> They found that at least two aromatic groups were required to inhibit A $\beta$  aggregation, those aromatic groups should have at least functional group capable of hydrogen bonding, a linker length of 6Å-19Å, and only one or two sp<sup>3</sup> hybridized carbons. These results put curcumin squarely into their “ideal” A $\beta$  aggregation inhibitor range. Simulations have also been performed to examine curcumin

binding to A $\beta$  monomers and aggregate structures. Ngo and Li used docking simulations and found that curcumin preferentially binds to the D1 terminal amino acid of A $\beta$ (1-40).<sup>27</sup> However, when they performed the same simulations with an A $\beta$ (9-40) fibril composed of two loop structures, curcumin bound S26 and K28 in a more central location inside the loop structure characteristic of A $\beta$  fibrils. Curcumin bound F20 in the Ngo and Li simulations with an A $\beta$ (9-40) fibril composed of three loop structures. Zhao *et al.* used atomistic molecular dynamics simulations to examine the effects of curcumin on the stability of an A $\beta$ (1-42) dimer.<sup>28</sup> Their major finding was that curcumin decreased  $\beta$ -sheet content in each monomer without altering the inter-peptide contacts. This decrease was what preceded the actual  $\beta$ -sheet breakage.

Epigallocatechin-3-gallate (EGCG) is an abundant molecule in green and white tea, with smaller amounts occurring in black tea. EGCG is similar to curcumin and resveratrol in that it is known for its potent antioxidant properties. Although EGCG contains three aromatic groups, it is structurally different from the other phenolic compounds mentioned above because it does not have a flexible carbon chain linker, making it more rigid. Bieschke *et al.* studied EGCG interactions with A $\beta$ (1-42) and found that EGCG can bind A $\beta$ (1-42) monomers and force them off pathway into large, disordered, non-toxic aggregates.<sup>29</sup> Their experiments showed that EGCG has the capability to completely disaggregate fibrils or remodel them without completely disassembling them to prevent further aggregation. Similar results were obtained by Ehrnhoefer *et al.*<sup>32</sup> and Bastianetto *et al.*<sup>33</sup> Wang *et al.* used isothermal titration calorimetry to understand how EGCG binds to A $\beta$ (1-42). Their major results were that while both hydrophobic interactions and hydrogen bonding contribute to

EGCG-A $\beta$ (1-42) interactions at low EGCG/A $\beta$ (1-42) ratios, the interactions are primarily due to hydrogen bonding and as the amount of EGCG increases the interactions shift toward hydrophobic interactions. Atomistic simulations have also been used to probe the molecular level details of EGCG interaction with A $\beta$ . Liu *et al.* found that non-polar interactions accounted for almost three quarters of the EGCG-A $\beta$ (1-42) binding energy and that the remaining energy was due to hydrogen bonding.<sup>34</sup> The residues with the strongest EGCG interaction were F4, R5, F19, F20, E22, K28, G29, L34-G37, and I41. Simulations by Zhang *et al.* confirmed the experimental results that EGCG remodels A $\beta$ (1-42) monomers and expanded upon that idea by showing that EGCG decreases  $\beta$ -sheet content and increases coil and  $\alpha$ -helix secondary structure.

The research described above on how these four phenolic compounds affect A $\beta$  aggregation can be summarized in the following way. Resveratrol is capable of preventing and reversing A $\beta$  fibril formation, and can transform toxic oligomers into a non-toxic disordered oligomer. Vanillin's impacts on A $\beta$  aggregation are somewhat unclear, results thus far are inconclusive with some saying that it can prevent oligomer and fibril formation, and others saying that it has no effect at all. Curcumin inhibits A $\beta$  oligomer and fibril inhibition by disrupting  $\beta$ -sheet formation through binding to the  $\beta$ -sheet structure characteristic of A $\beta$  fibrils. EGCG has been shown to bind to monomers, which is different from the other phenolic compounds considered here. It also binds to oligomers and fibrils, thereby decreasing aggregation.

Studies have been conducted that directly compare the effectiveness of these small molecules. Chebaro *et al.* used docking simulations to show that resveratrol and curcumin

and EGCG, among other inhibitors, have good binding affinity for A $\beta$ (17-42) trimers, most specifically residues 17-21.<sup>19</sup> Resveratrol and curcumin both preferred to bind to the lowest energy A $\beta$  oligomer tested, while EGCG preferred binding with high-energy structures. This may mean that the aggregation inhibition from EGCG stems from binding early intermediate structures rather than later stable ones. Experiments by Reinke and Gestwicki showed that vanillin was significantly less effective than resveratrol and curcumin at inhibiting A $\beta$  aggregation.<sup>3</sup> Finally, Feng *et al.* found curcumin to be effective, but not as effective as resveratrol at inhibiting fibril formation.<sup>20</sup> From these results we expect that the effectiveness of these inhibitors will be Resveratrol>Curcumin>EGCG>Vanillin.

## 5.2 Methods

### 5.2.1 Discontinuous Molecular Dynamics and PRIME20 Force Field

In this work, we use discontinuous molecular dynamics (DMD), a fast alternative to traditional molecular dynamics and the PRIME20 force field.<sup>38</sup> In DMD simulations, the only time that the velocities and positions need to be recalculated is when a discontinuity in the potential is encountered. In contrast, with continuous potentials, the simulations are advanced at regularly-spaced time steps. Since DMD advances the simulation on a collision to collision basis rather than at small predetermined time steps, the technique allows us to simulate much longer time scales. The PRIME20 force field uses a four-sphere-per-residue model in which each amino acid residue is represented by three backbone spheres, one each for N-H, C-H, and C=O, and one side-chain sphere R.<sup>39</sup> While reducing the resolution of the model decreases some detail, it also simplifies the system and allows us to simulate larger

system sizes. A detailed description of DMD and PRIME20 can be found in other papers from our group.<sup>40-42</sup> We will discuss the modifications we have made to these techniques needed to simulate inhibitors alongside proteins.

### 5.2.2 Inhibitor Model – Coarse-Graining

While PRIME20 and DMD are well suited for studying the aggregation of proteins, PRIME20 does not contain any geometric or energetic parameters for other types of molecules. In order to simulate the proposed phenolic compounds in DMD with proteins represented by PRIME20 we had to come up with a way to make an inhibitor model compatible with PRIME20. We began by coarse-graining each molecule in a way that would be compatible with the hydrogen bonding scheme in PRIME20 as well as preserving enough detail so that the inhibitor-peptide interactions are not lost. The hydrogen bonding scheme used for our simulations is discussed extensively in a paper from Smith and Hall.<sup>43</sup> Figure 5.1 shows the coarse-grained representations of A) resveratrol, B) EGCG, C) vanillin, and D) curcumin with a side bar defining each coarse-grained group.

Once our choice of coarse-grained groups was determined, we had to define their geometrical and interaction parameters to be consistent with the PRIME20 force field. PRIME20 already has parameters for C-H<sub>3</sub>, C-H, and C=O spheres so we based the parameters for each new group on those parameters with slight modifications. Since C-H<sub>3</sub> and C-H differ by two hydrogen atoms, we calculated the contribution of a single hydrogen atom to the diameter  $\sigma$  of coarse-grained groups in PRIME20 to be 0.150Å and the well width  $\lambda$  to be 0.300Å. Using these values, we established the parameters for C-H<sub>2</sub> and C by

scaling the diameters and well widths accordingly, giving us diameters of 2.550Å for C-H<sub>2</sub> and 2.250Å for C, and well widths of 5.100Å for C-H<sub>2</sub> and 4.500Å for C. The well depth  $\epsilon$  for these spheres is the same as the value for the alanine side chain but may be updated in future calculations. To determine the diameter for the O-H sphere in PRIME20 we scaled the N-H PRIME20 sphere diameter down by 2% to obtain 3.234Å. This scale-down was used to mimic the ratio of atomistic oxygen and nitrogen van der Waals radii, 1.520Å and 1.550Å respectively. The diameters for the O and H spheres were simply taken to be their atomistic values, 3.040Å and 2.18Å. We are aware that this will need future modification to preserve the scaling of atomic radii. The hydrogen bonding groups O-H, O and C=O all have the same well width and well depth  $\epsilon$  to maintain compatibility with our current hydrogen bonding routine. The H sphere was modeled as a hard sphere interaction type because it is only included in the model to maintaining hydrogen bonding compatibility. These values are our first attempt at integrating new coarse-grained spheres into our simulations and will be adjusted in the future. Please see future work for proposed modifications. The values for the mass, reduced mass, bead diameter  $\sigma$ , well diameter  $\lambda$ , square well depth  $\epsilon$ , and interaction type are summarized in Table 5.1 for each coarse-grained sphere.

### 5.2.3 Inhibitor Model – Atomistic Simulations

After defining the coarse-grained inhibitor representation and parameters, we had to determine the bond locations and lengths necessary to constrain the inhibitor molecules to have realistic structural dynamics. In order to do this, the atomistic representation for each inhibitor was constructed using Accelrys Discovery Studio<sup>44</sup> and the topology files were

generated using SwissParam<sup>45</sup>, an automated force field generator for small molecules. Using the newly generated geometry and force field parameters, we simulated a single molecule of each inhibitor using the GROMACS<sup>46</sup> simulation package, the GROMOS53a6<sup>47</sup> force field and SPC water. Each simulation was run with explicit water at 297K in a 5Åx5Åx5Å simulation box with periodic boundary conditions for a total of 10ns. Covalent bond lengths in our coarse-grained inhibitor molecule were assigned the values determined by SwissParam because we maintain an almost atomistic geometry, minus hydrogens, besides the O-H and C=O groups. The C to O-H (navy sphere to light blue sphere in Figure 5.1), H to C=O (grey sphere to red sphere in Figure 5.1) and C=O to C (red sphere to light blue sphere in Figure 5.1) bond lengths are simply the atomistic C to O, H to C, and C to C bond lengths. The trajectories allowed us to monitor the structural fluctuations of the molecules and determine where pseudobonds may need to be applied for our coarse-grained model. In PRIME20, pseudobonds are used to constrain coarse-grained sphere movement to realistic behavior. In PRIME20 there are only single bonds so we use these pseudo bonds to maintain the aromatic ring structure of the benzene and benzopyran rings as well as other C=C double bonds. Figure 5.2 shows the location of the pseudobonds used for A) resveratrol, B) EGCG, C) vanillin, and D) curcumin. Red bonds constrain benzene and benzopyran rings, and green, purple and blue bonds restrain spheres once, twice, and three times removed from the rings. The pseudo bonds chosen here are an initial guess and will be refined in future work. Pseudo bonds for benzene and benzopyran rings were allowed to fluctuate between 2.60Å and 2.92Å, pseudobonds for spheres once removed from a ring fluctuate between 2.21Å and 2.50Å, pseudobonds for spheres twice removed from a ring fluctuate between 2.24Å and

2.57Å, and pseudobonds for spheres three times removed from a ring fluctuate between 2.73Å and 3.80Å.

#### **5.2.4 A $\beta$ (17-36) Model**

In these simulations, we wanted to use a model peptide that was the smallest possible A $\beta$  peptide that was capable of forming the U-loop structure characteristic of amyloid fibrils. To determine what sequence would be best we turned to unpublished work from Dr. Mookyung Cheon using DMD and PRIME20 to simulate fibrillization of systems containing 8 A $\beta$ (17-42) peptides. In those simulations a U-loop structures characteristic of the fibrils observed experimentally for A $\beta$  were formed. In those simulations, Cheon used a double well potential to allow amino acid side chains to interact gradually and two slightly biased simulation parameters from PRIME20: parallel preference angular constraints for forming backbone hydrogen-bonds and an enhanced salt-bridge interaction between K28 and D23. During the course of his analysis he found that the F19-L34 and F19-I32 hydrophobic interactions on the interior of the loop were key for holding the structure together. Since adding a significant amount of inhibitors to our simulations would slow them down considerably, we decided not to use a double well potential for side chains, and instead slightly enhance the aforementioned hydrophobic interactions while keeping all side chain interactions as a single well. Since F19 and L34 were evidently important interactions, we extended the sequence considered to include the two residues before F19 and two residues after L34. The punchline was that the best peptide to simulate for this project was the A $\beta$ (17-36) peptide. In future work we will likely use the full length peptide.

### 5.2.5 Simulation Procedure – Coarse-Grained Inhibitor Simulations

The coarse-grained inhibitor-A $\beta$ (17-36) simulations proceed in the following way: 8 peptides are placed at a random locations in the simulation box and, if inhibitors are also being simulated, either 128 or 256 inhibitor molecules. The simulation box is cubic with side lengths of  $L=138.5\text{\AA}$  and periodic boundary conditions. This results in a peptide concentration of 5mM and inhibitor concentrations of either 80mM or 160mM. In these simulations the reduced temperature is defined as  $T^*=kT/\epsilon_{HB}$  where  $\epsilon_{HB}$  is the hydrogen bonding well depth. Velocities for each peptide and inhibitor sphere are chosen at random from a Maxwell-Boltzmann distribution that is centered at the desired temperature. Initially, the temperature is set to  $T^*=0.20$  and the temperature is held constant using the Andersen thermostat method. Spheres in the simulation experience random “ghost collisions” with “ghost particles” during which their velocity is reassigned to a random value from a Maxwell-Boltzmann distribution centered at the desired simulation temperature. For this initial work, we ran two independent sets of simulations for each inhibitor and inhibitor concentration as well as one extended peptide-only simulation to verify the efficacy of our A $\beta$ (17-36) model. Inhibitor simulations were run for ~10 billion collisions and the peptide-only simulation was run for ~100 billion. Inhibitor simulations were only run for 10 billion collisions because there are still issues to be worked out in the code. Simulations with EGCG were only run for 2 billion collisions since this model was created much later than the other inhibitors.

## 5.3 Results and Discussion

### 5.3.1 Peptides Only

To verify the effectiveness of our enhanced A $\beta$ (17-36) we ran one initial test to see whether or not the U-loop structure characteristic of amyloid fibrils would form. Figure 5.3 shows the A $\beta$ (17-36) U-loop structure formed after 100 billion collisions for its last 20 frames: A) Top view, all peptides B) Side view, all peptides C) Top view, ordered peptides only and D) side view, ordered peptides only. Figure 5.4A and 5.4B show the full configuration for all 8 peptides. The green, pink, blue, and yellow peptides are completely in  $\beta$ -sheets and while the purple peptide is attached to the structure through side chain contacts, it has yet to form the backbone hydrogen bonds necessary to lock it into a  $\beta$ -sheet with the other peptides. The fact that the purple peptide is attached to the structure, but it is not in a  $\beta$ -sheet indicates that it is attached through intersheet sidechain interactions with the yellow peptide. Figures 5.4C and 5.4D include only the ordered peptides for clarity. The lower portion of the grey peptide is in a  $\beta$ -sheet with the green peptide, but the red peptide is stuck between them, preventing the grey peptide from joining the U-loop structure. The other end of the red peptide hangs in a random coil in front of the other peptides. Finally, the orange peptide is bound to the grey peptide but is not yet locked in place with backbone hydrogen bonds. We anticipate that the red peptide will eventually detach itself, the grey peptide will form a  $\beta$ -sheet with the green peptide, and ultimately the red peptide will reattach to the end of the U-loop, completing the fully ordered structure.

### 5.3.2 Inhibitor-Peptide Binding During Lag Phase

The simulations including peptides and inhibitors ran only for ~10 billion collision and 2 billion collisions for EGCG. In our simulations this time period is usually associated with the aggregation lag phase during which there are primarily monomers (first 2 billion collisions) and the start of aggregation into oligomers (~10 billion collisions). Figure 5.4 shows inhibitor-peptide complexes for A) vanillin, B) resveratrol, C) EGCG, and D) curcumin. Each peptide is a different color, and the inhibitors are colored by element to show their coarse-grained sites. Figures 5.4A, 5.4B, and 5.4D were captured during the start of aggregation after 10 billion collisions and Figure 5.4C was captured while the peptides were isolated monomers after 2 billion collisions. Figure 5.5A shows two vanillin molecules interacting with an aggregate of four A $\beta$ (17-36) peptides, Figure 5.5B shows five resveratrol molecules interacting with an aggregate of three A $\beta$ (17-36) peptides, Figure 5.5C shows 3 EGCG molecules interacting with an A $\beta$ (17-36) monomer, and Figure 5.5D shows three curcumin molecules interacting with an aggregate of six A $\beta$ (17-36) peptides. While these snapshots are very early in the simulation there is one very important observation we can make about how the inhibitors are interacting with the peptides. In Figure 5.5A the vanillin molecules are interacting with the exterior of the A $\beta$ (17-36) aggregate, while in Figures 5.5B and 5D the curcumin and resveratrol molecules sit at the interior of the aggregate. Further work is needed to determine why this is occurring, but this could point to different inhibition mechanisms. It is also interesting to note that the EGCG molecules interact strongly with a monomer, whereas the other molecules interact with aggregates.

### 5.3.3 Inhibitor-Peptide Interaction Comparison

To quantify the inhibitor peptide interactions we calculated the peptide-inhibitor radial distribution functions for vanillin, resveratrol and curcumin taking data from the initial state of the simulation until the start of aggregation. (Those for EGCG were not calculated because of the limited simulation time). Figure 5.5 shows a plot of the radial distribution function  $g(r)$  between the center of mass of the peptides and the center of mass of the inhibitors vs.  $r$  for vanillin (purple), resveratrol (green), and curcumin (orange). Each radial distribution function has the same general shape because while the inhibitor geometry differs, each molecule currently has the same types of interactions: hydrophobic and hydrogen bonds. The first peak in each curve is due to hydrogen bonding between the inhibitor O-H, C=O, and O groups and the peptide backbone C=O and N-H groups and the larger broad peak comes from the hydrophobic interactions between the peptides and inhibitors. Our results indicate that more resveratrol is found near the A $\beta$ (17-36) peptide than curcumin, which is higher than vanillin. The higher number of resveratrol molecules near the A $\beta$ (17-36) peptide indicates that it likely binds better and may be the best aggregation inhibitor, followed by curcumin, and vanillin.

## 5.4 Discussion and Conclusions

Using the combination of DMD and our PRIME20 force field we were able to simulate systems of phenolic aggregation inhibitors with the A $\beta$ (17-36) peptide. Much simulation and experimental work has been put into aggregation inhibitor research to help combat AD, but none have looked at systems of aggregation inhibitors and peptide on this

large size scale. Atomistic simulations are limited to docking studies or molecular dynamics/Monte Carlo on systems consisting of only a few peptides, typically arranged in a preformed fibrillar structure, and inhibitors. For the first time we are able to observe how these inhibitors interact with systems of peptides starting in a random coil all the way to the formation of the U-loop structure characteristic of amyloid structures.

The limited results we have presented here show promise for our future work on this project. The U-loop fibril structure that we formed in our simulation with 8 A $\beta$ (17-36) peptides has very similar structural characteristics to the of A $\beta$ (1-42) protofilament identified by Luhrs *et al.* including the loop,  $\beta$ -sheet structure, and inter-sheet side chain interactions.<sup>48</sup> A longer sequence would be ideal for this study so in the future we may look at the full-length peptides. In our literature review, we determined that the effectiveness of vanillin as an aggregation inhibitor was inconclusive, while curcumin and resveratrol were found to be quite effective. In our snapshots of the initial stages of aggregation in our simulations, we showed that vanillin tends to bind the exterior of aggregates while curcumin and vanillin binds the interior. When we perform longer simulations in the future this difference may offer insight into why vanillin is not as effective as the other inhibitors. Finally, our plot of the peptide-inhibitor radial distribution function showed that resveratrol bound to A $\beta$  most frequently, followed by curcumin, and lastly vanillin. The experiments from Reinke and Gestwicki showed that vanillin was significantly less effective than resveratrol and curcumin at inhibiting A $\beta$  aggregation<sup>3</sup> and Feng *et al.* found curcumin to be effective, but not as effective as resveratrol at inhibiting fibril formation<sup>20</sup>. Although these results are from

different studies, our initial analysis is in agreement with their results for inhibitor effectiveness.

## **5.5 Acknowledgements**

This work was supported by the National Institutes of Health, USA under grant GM56766 and EB006006 (CKH).

## 5.6 References

1. Wolf LK. Getting to alzheimer's early. *Chemical & Engineering News [Online]*. 2014;91(11):41.
2. Kolata G. Doubt on tactic in alzheimer's battle. *The New York Times [Online]*. 2010. Available from: <http://www.nytimes.com/2010/08/19/health/19alzheimers.html>.
3. Kolata G. Drug trials test bold plan to slow alzheimer's. *The New York Times [Online]*. 2010. Available from: <http://www.nytimes.com/2010/07/17/health/research/17drug.html?pagewanted=all>.
4. Necula M, Kaye R, Milton S, Glabe CG. Small molecule inhibitors of aggregation indicate that amyloid beta oligomerization and fibrillization pathways are independent and distinct. *J Biol Chem*. 2007;282(14):10311-10324. doi: 10.1074/jbc.M608207200.
5. Nakagami Y, Nishimura S, Murasugi T, *et al*. A novel beta-sheet breaker, RS-0406, reverses amyloid beta-induced cytotoxicity and impairment of long-term potentiation *in vitro*. *Br J Pharmacol*. 2002;137(5):676-682. doi: 10.1038/sj.bjp.0704911.
6. Choi Y, Jung C, Lee S, *et al*. The green tea polyphenol (-)-epigallocatechin gallate attenuates beta-amyloid-induced neurotoxicity in cultured hippocampal neurons. *Life Sci*. 2001;70(5):603-614. doi: 10.1016/S0024-3205(01)01438-2.
7. Frautschy S, Hu W, Kim P, *et al*. Phenolic anti-inflammatory antioxidant reversal of A beta-induced cognitive deficits and neuropathology. *Neurobiol Aging*. 2001;22(6):993-1005. doi: 10.1016/S0197-4580(01)00300-1.

8. Lim G, Chu T, Yang F, Beech W, Frautschy S, Cole G. The curry spice curcumin reduces oxidative damage and amyloid pathology in an alzheimer transgenic mouse. *Journal of Neuroscience*. 2001;21(21):8370-8377.
9. Jang J, Surh Y. Protective effect of resveratrol on beta-amyloid-induced oxidative PC12 cell death. *Free Radical Biology and Medicine*. 2003;34(8):1100-1110. doi: 10.1016/S0891-5849(03)00062-5.
10. Savaskan E, Olivieri G, Meier F, Seifritz E, Wirz-Justice A, Muller-Spahn F. Red wine ingredient resveratrol protects from beta-amyloid neurotoxicity. *Gerontology*. 2003;49(6):380-383. doi: 10.1159/000073766.
11. Kim J, Lee HJ, Lee KW. Naturally occurring phytochemicals for the prevention of alzheimer's disease. *J Neurochem*. 2010;112(6):1415-1430. doi: 10.1111/j.1471-4159.2009.06562.x.
12. Sun AY, Wang Q, Simonyi A, Sun GY. Botanical phenolics and brain health. *Neuromolecular Medicine*. 2008;10(4):259-274. doi: 10.1007/s12017-008-8052-z.
13. Porat Y, Abramowitz A, Gazit E. Inhibition of amyloid fibril formation by polyphenols: Structural similarity and aromatic interactions as a common inhibition mechanism. *Chemical Biology & Drug Design*. 2006;67(1):27-37. doi: 10.1111/j.1747-0285.2005.00318.x.
14. Kim J, Lee HJ, Lee KW. Naturally occurring phytochemicals for the prevention of alzheimer's disease. *J Neurochem*. 2010;112(6):1415-1430. doi: 10.1111/j.1471-4159.2009.06562.x.

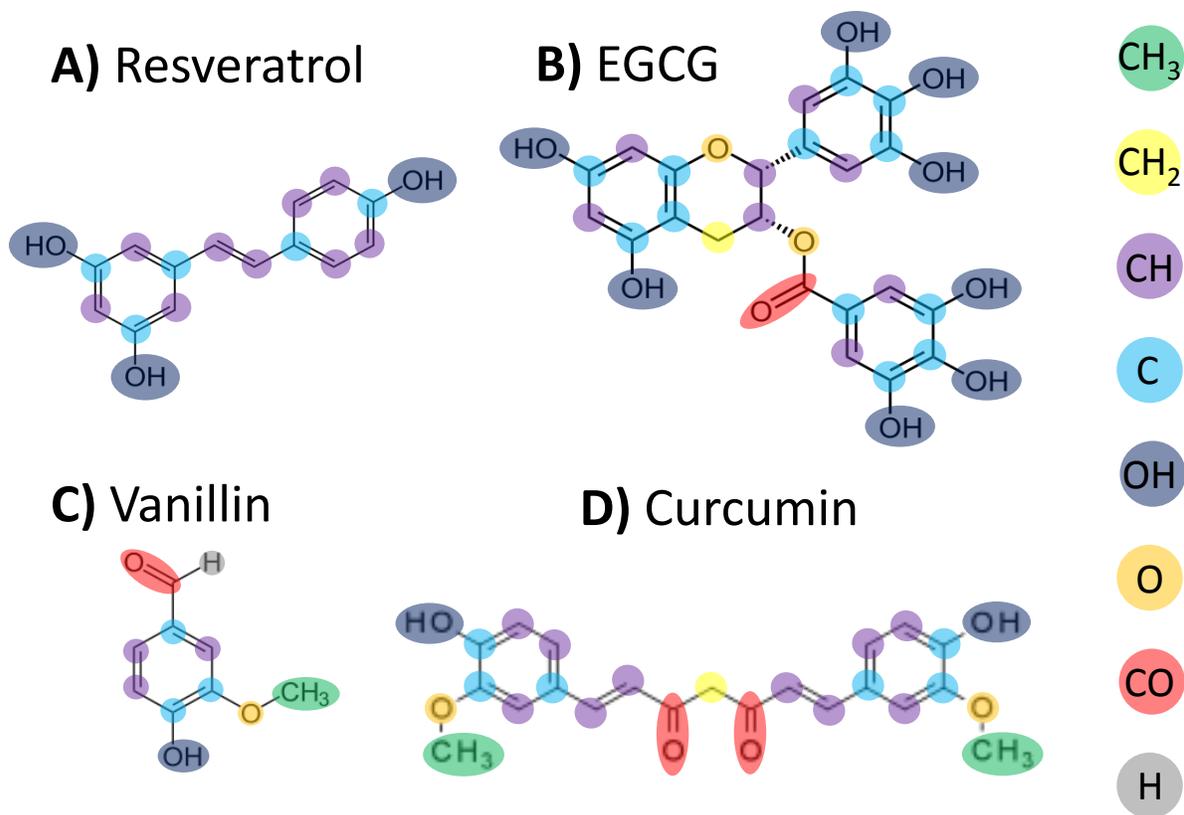
15. Karuppagounder SS, Pinto JT, Xu H, Chen H, Beal MF, Gibson GE. Dietary supplementation with resveratrol reduces plaque pathology in a transgenic model of alzheimer's disease. *Neurochem Int.* 2009;54(2):111-118. doi: 10.1016/j.neuint.2008.10.008.
16. Han Y, Zheng W, Bastianetto S, Chabot J, Quirion R. Neuroprotective effects of resveratrol against beta-amyloid-induced neurotoxicity in rat hippocampal neurons: Involvement of protein kinase C. *Br J Pharmacol.* 2004;141(6):997-1005. doi: 10.1038/sj.bjp.0705688.
17. Wang J, Ho L, Zhao Z, *et al.* Moderate consumption of cabernet sauvignon attenuates A beta neuropathology in a mouse model of alzheimer's disease. *Faseb Journal.* 2006;20(13):2313-2320. doi: 10.1096/fj.06-6281com.
18. Marambaud P, Zhao H, Davies P. Resveratrol promotes clearance of alzheimer's disease amyloid-beta peptides. *J Biol Chem.* 2005;280(45):37377-37382. doi: 10.1074/jbc.M508246200.
19. Chebaro Y, Jiang P, Zang T, *et al.* Structures of A beta 17-42 trimers in isolation and with five small-molecule drugs using a hierarchical computational procedure. *J Phys Chem B.* 2012;116(29):8412-8422. doi: 10.1021/jp2118778.
20. Feng Y, Wang X, Yang S, *et al.* Resveratrol inhibits beta-amyloid oligomeric cytotoxicity but does not prevent oligomer formation. *Neurotoxicology.* 2009;30(6):986-995. doi: 10.1016/j.neuro.2009.08.013.
21. Reinke AA, Gestwicki JE. Structure-activity relationships of amyloid beta-aggregation inhibitors based on curcumin: Influence of linker length and flexibility.

- Chemical Biology & Drug Design*. 2007;70(3):206-215. doi:  
10.1111/j.1747.0285.2007.00557.x.
22. Necula M, Kaye R, Milton S, Glabe CG. Small molecule inhibitors of aggregation indicate that amyloid beta oligomerization and fibrillization pathways are independent and distinct. *J Biol Chem*. 2007;282(14):10311-10324. doi:  
10.1074/jbc.M608207200.
23. Ladiwala ARA, Lin JC, Bale SS, *et al*. Resveratrol selectively remodels soluble oligomers and fibrils of amyloid A beta into off-pathway conformers. *J Biol Chem*. 2010;285(31):24228-24237. doi: 10.1074/jbc.M110.133108.
24. Garcia-Alloza M, Borrelli LA, Rozkalne A, Hyman BT, Bacskai BJ. Curcumin labels amyloid pathology *in vivo*, disrupts existing plaques, and partially restores distorted neurites in an Alzheimer mouse model. *J Neurochem*. 2007;102(4):1095-1104. doi:  
10.1111/j.1471-4159.2007.04613.x.
25. Ono K, Hasegawa K, Naiki H, Yamada M. Curcumin has potent anti-amyloidogenic effects for Alzheimer's beta-amyloid fibrils *in vitro*. *J Neurosci Res*. 2004;75(6):742-750. doi: 10.1002/jnr.20025.
26. Yang F, Lim G, Begum A, *et al*. Curcumin inhibits formation of amyloid beta oligomers and fibrils, binds plaques, and reduces amyloid *in vivo*. *J Biol Chem*. 2005;280(7):5892-5901. doi: 10.1074/jbc.M404751200.
27. Ngo ST, Li MS. Curcumin binds to A beta(1-40) peptides and fibrils stronger than ibuprofen and naproxen. *J Phys Chem B*. 2012;116(34):10165-10175. doi:  
10.1021/jp302506a.

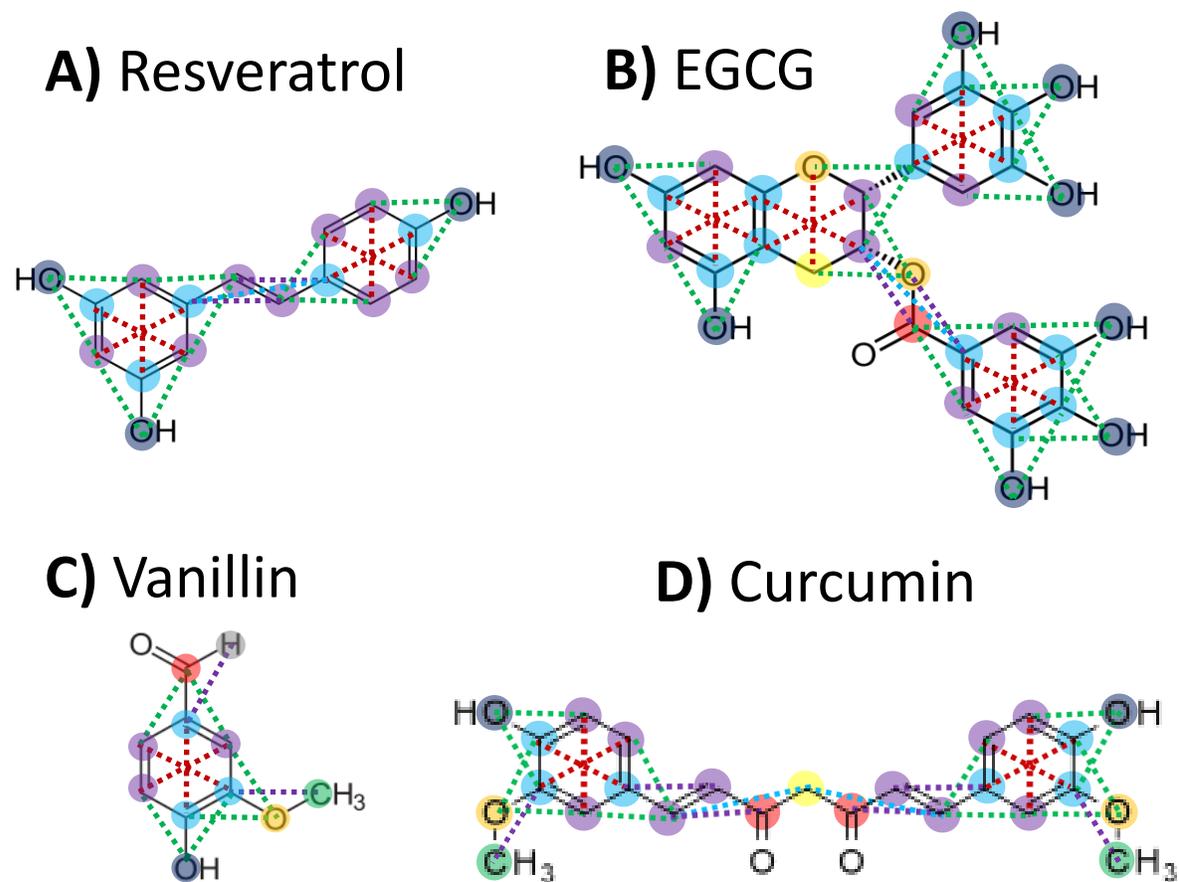
28. Zhao LN, Chiu S, Benoit J, Chew LY, Mu Y. The effect of curcumin on the stability of A beta dimers. *J Phys Chem B*. 2012;116(25):7428-7435. doi: 10.1021/jp3034209.
29. Bieschke J, Russ J, Friedrich RP, *et al*. EGCG remodels mature alpha-synuclein and amyloid-beta fibrils and reduces cellular toxicity. *Proc Natl Acad Sci U S A*. 2010;107(17):7710-7715. doi: 10.1073/pnas.0910723107.
30. Rezai-Zadeh K, Arendash GW, Hou H, *et al*. Green tea epigallocatechin-3-gallate (EGCG) reduces beta-amyloid mediated cognitive impairment and modulates tau pathology in alzheimer transgenic mice. *Brain Res*. 2008;1214:177-187. doi: 10.1016/j.brainres.2008.02.107.
31. Levites Y, Amit T, Mandel S, Youdim M. Neuroprotection and neurorescue against A beta toxicity and PKC-dependent release of non-amyloidogenic soluble precursor protein by green tea polyphenol (-)-epigallocatechin-3-gallate. *Faseb Journal*. 2003;17(3):952-+. doi: 10.1096/fj.02-0881-fje.
32. Ehrnhoefer DE, Bieschke J, Boeddrich A, *et al*. EGCG redirects amyloidogenic polypeptides into unstructured, off-pathway oligomers. *Nature Structural & Molecular Biology*. 2008;15(6):558-566. doi: 10.1038/nsmb.1437.
33. Bastianetto S, Yao Z, Papadopoulos V, Quirion R. Neuroprotective effects of green and black teas and their catechin gallate esters against beta-amyloid-induced toxicity. *Eur J Neurosci*. 2006;23(1):55-64. doi: 10.1111/j.1460-9568.2005.04532.x.
34. Liu F, Dong X, He L, Middelberg APJ, Sun Y. Molecular insight into conformational transition of amyloid beta-peptide 42 inhibited by (-)-epigallocatechin-3-gallate

- probed by molecular simulations. *J Phys Chem B*. 2011;115(41):11879-11887. doi: 10.1021/jp202640b.
35. Wang S, Liu F, Dong X, Sun Y. Thermodynamic analysis of the molecular interactions between amyloid beta-peptide 42 and (-)-epigallocatechin-3-gallate. *J Phys Chem B*. 2010;114(35):11576-11583. doi: 10.1021/jp1001435.
  36. Zhang T, Zhang J, Derreumaux P, Mu Y. Molecular mechanism of the inhibition of EGCG on the alzheimer A beta(1-42) dimer. *J Phys Chem B*. 2013;117(15):3993-4002. doi: 10.1021/jp312573y.
  37. De Felice FG, Vieira MNN, Saraiva LM, *et al*. Targeting the neurotoxic species in alzheimer's disease: Inhibitors of A beta oligomerization. *Faseb Journal*. 2004;18(12):1366-1372. doi: 10.1096/fj.04-1764com.
  38. Alder BJ, Wainwright TE. Studies in molecular dynamics .1. general method. *J Chem Phys*. 1959;31(2). doi: 10.1063/1.1730376.
  39. Wu C, Shea J. Coarse-grained models for protein aggregation. *Curr Opin Struct Biol*. 2011;21(2):209-220. doi: 10.1016/j.sbi.2011.02.002.
  40. Cheon M, Chang I, Hall CK. Extending the PRIME model for protein aggregation to all 20 amino acids. *Proteins*. 2010;78(14):2950-2960. doi: 10.1002/prot.22817.
  41. Nguyen H, Hall C. Kinetics of fibril formation by polyalanine peptides RID D-5346-2011. *J Biol Chem*. 2005;280(10):9074-9082. doi: 10.1074/jbc.M407338200.
  42. Marchut AJ, Hall CK. Spontaneous formation of annular structures observed in molecular dynamics simulations of polyglutamine peptides. *Computational Biology and Chemistry*. 2006;30(3):215-218.

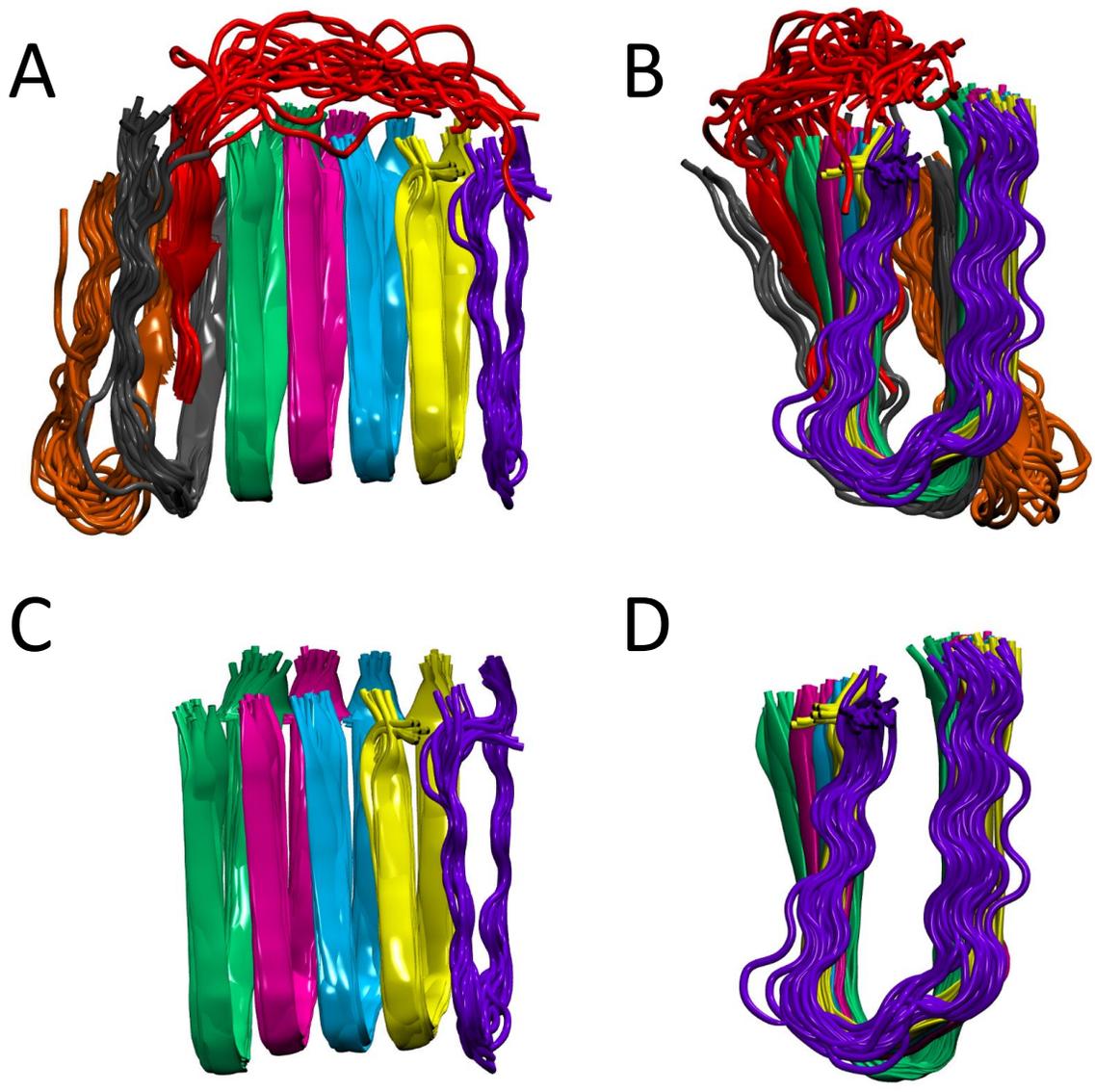
43. Smith A, Hall C. Alpha-helix formation: Discontinuous molecular dynamics on an intermediate-resolution protein model. *Proteins-Structure Function and Genetics*. 2001;44(3):344-360.
44. Accelrys Software Inc. *Discovery studio modeling environment, release 4.0*. San Diego: Accelrys Software Inc. 2013.
45. Zoete V, Cuendet MA, Grosdidier A, Michielin O. SwissParam: A fast force field generation tool for small organic molecules. *Journal of Computational Chemistry*. 2011;32(11):2359-2368. doi: 10.1002/jcc.21816.
46. Pronk S, Pall S, Schulz R, *et al.* GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*. 2013;29(7):845-854. doi: 10.1093/bioinformatics/btt055.
47. Oostenbrink C, Villa A, Mark A, Van Gunsteren W. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry*. 2004;25(13):1656-1676. doi: 10.1002/jcc.20090.
48. Luhrs T, Ritter C, Adrian M, *et al.* 3D structure of alzheimer's amyloid-beta(1-42) fibrils. *Proc Natl Acad Sci U S A*. 2005;102(48):17342-17347. doi: 10.1073/pnas.0506723102.



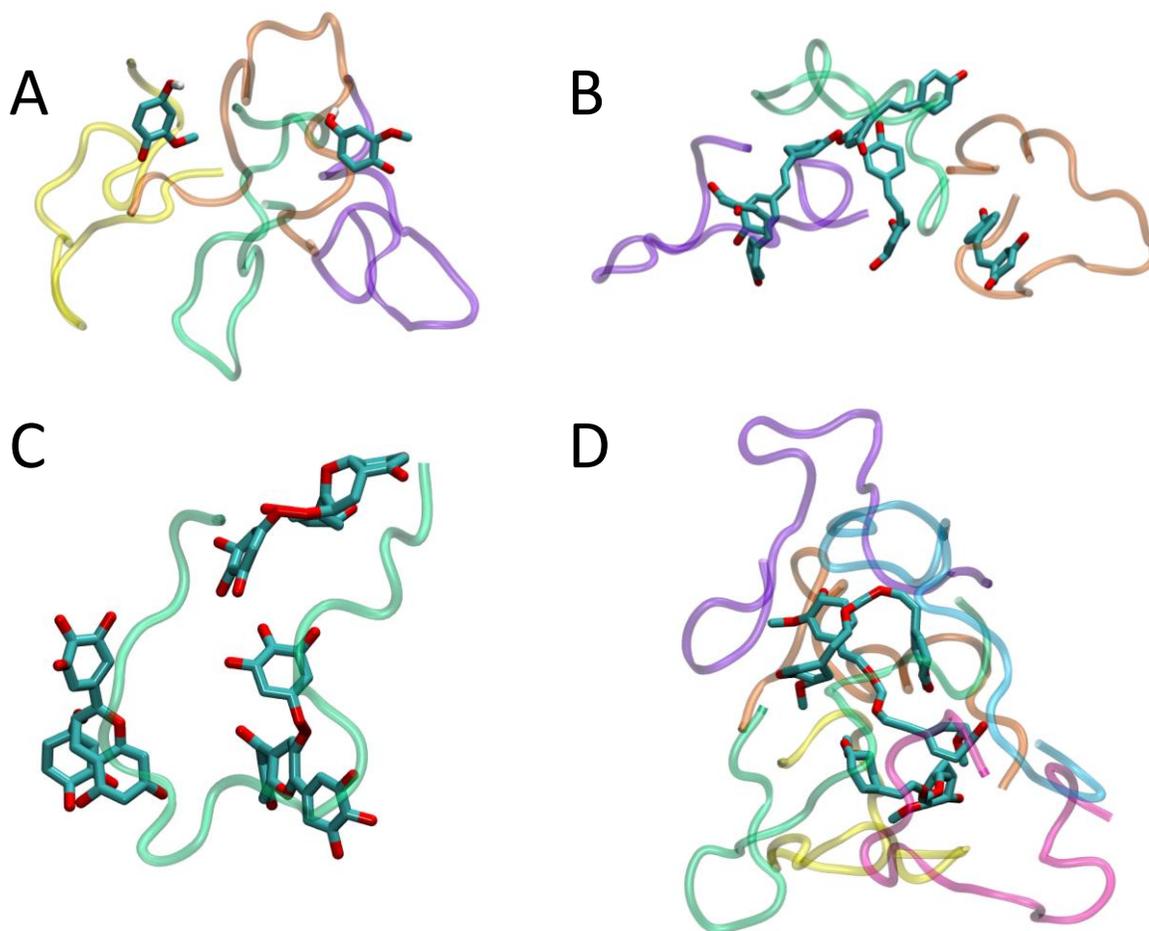
**Figure 5.1.** Coarse-grained representations of A) resveratrol, B) EGCG, C) vanillin, and D) curcumin to be used in PRIME20 simulations. The side bar defines each coarse-grained group.



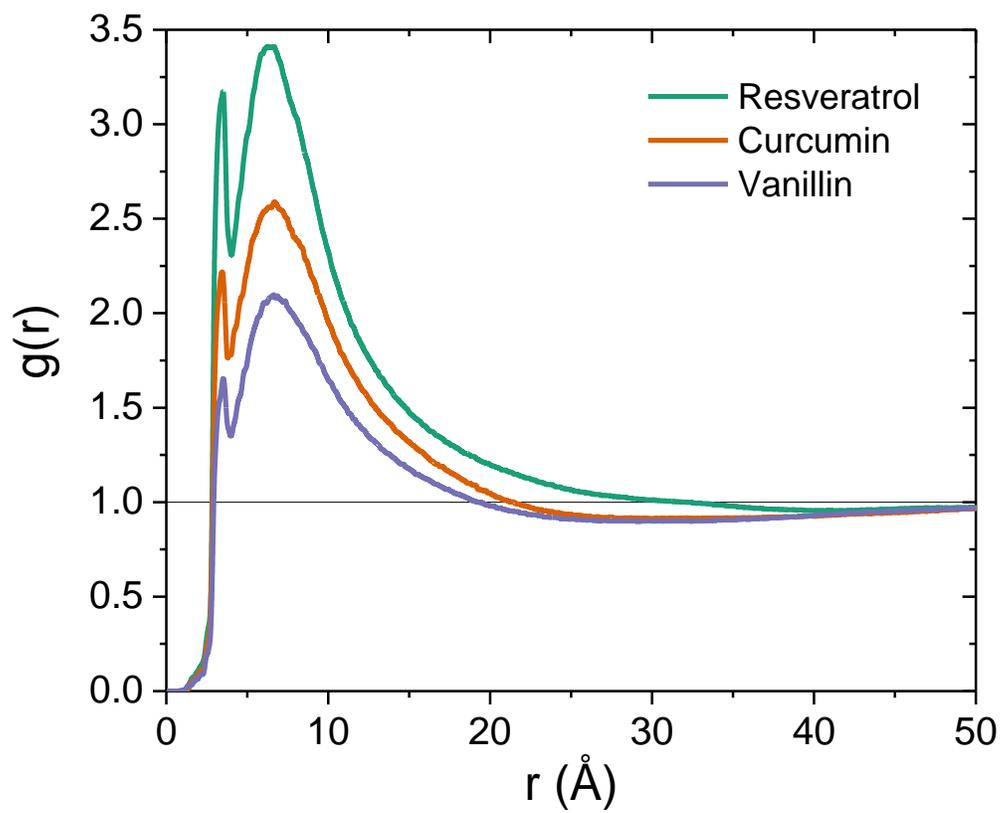
**Figure 5.2.** Pseudobonds used for A) Reseveratrol, B) EGCG, C)Vanillin, and D) Curcumin. Red bonds constrain benzene and benzopyran rings, and green, purple and blue bonds restrain spheres once, twice, and three times removed from the rings.



**Figure 5.3.** A $\beta$ (17-36) U-loop for structure formed after 100 billion collisions for its last 20 frames: A) Top view, all peptides B) Side view, all peptides C) Top view, ordered peptides only and D) side view, ordered peptides only



**Figure 5.4.** Inhibitor-peptide complexes for A) vanillin, B) resveratrol, C) EGCG, and D) curcumin. Each peptide is a different color, and the inhibitors are colored by element to show their coarse-grained sites.



**Figure 5.5.** Radial distribution function  $g(r)$  vs.  $r$  for resveratrol (green), curcumin (orange) and vanillin (purple).

**Table 5.1.** Coarse grained groups for phenolic compounds

CG Site	Atoms	Mass	Reduced Mass	$\sigma(\text{\AA})$	$\lambda(\text{\AA})$	$\epsilon(\text{reduced energy})$	Interaction type
1	CH <sub>3</sub>	15.035	1.000	2.700	5.400	0.084	Hydrophobic
2	CH <sub>2</sub>	14.027	0.933	2.550	5.100	0.084	Hydrophobic
3	CH	13.019	0.866	2.400	4.800	0.084	Hydrophobic
4	C	12.011	0.799	2.250	4.500	0.084	Hydrophobic
5	OH	17.007	1.131	3.234	4.500	1.000	Hydrogen Bond
6	O	15.999	1.064	3.040	4.500	1.000	Hydrogen Bond
7	CO	28.010	1.863	2.400	4.500	1.000	Hydrogen Bond
8	H	1.008	0.067	2.180	0.000	0.000	Hard Sphere

## **CHAPTER 6**

### **Conclusions and Future Work**

## 6.1 Conclusions

In the preceding chapters, we explored the the effects of outside forces on protein aggregation. We investigated the effects of hard sphere and hydrophobic crowders on A $\beta$ (16-22) aggregation. Furthermore, we studied the effects of an aqueous/organic interface on the folding an aggregation of an amphipathic  $\alpha$ -helix. Finally, we began in investigation into the effects of naturally occurring phenolic inhibitors on A $\beta$ (17-36) aggregation. Some of the major findings are summarized below.

**Chapter 2.** Using the combination of DMD and our PRIME20 force field we have been able to simulate systems of coarse grained proteins that have realistic geometry and energetic parameters along with crowding spheres up to realistic volume fractions. Although previous studies have been performed on similar systems, we are not aware of any that match the scale and realism of the species involved in the simulations. The systems contained 192 A $\beta$ (16-22) peptides and crowders of diameters 5Å, 20Å, and 40Å, represented here by simple hard spheres, at crowder volume fractions of  $\phi=0.00$ , 0.10, and 0.20. Our results show that both crowder volume fraction and size have a large impact on fibril and oligomer formation. The addition of crowders to a simulation without crowders increases the rate of oligomer formation and the peak number of oligomers that form. As the crowder volume fraction increases or the crowder diameter decreases, the increase in oligomer formation is accompanied by a shift from a slow ordered formation of oligomers, similar to nucleated polymerization, to a fast collapse and subsequent rearrangement that leads to the high maximum number of peptides in oligomers as is characteristic of nucleated conformational

conversion. The rate of conversion from oligomers to fibrils also increases giving rise to an increased rate of fibril growth. Based on our analysis it appears there is not an abrupt transition from nucleated polymerization to nucleated conformational conversion while increasing crowder volume fraction or decreasing crowder size, rather the mechanism governing fibrillization changes gradually with the simulation conditions. In all cases, larger volume fractions and smaller crowdors provide the largest enhancement of oligomerization and fibrillization.

**Chapter 3.** DMD/PRIME20 simulations have been used to analyze the oligomerization and fibrillization of a large system of A $\beta$ (16-22) peptides immersed in a sea of hydrophobic crowdors. The focus here has been on how the introduction of crowder-peptide interactions, beyond the excluded volume interactions that we considered previously, alters the types of aggregates formed and how this changes as the crowder size and hydrophobicity change. We believe that the unique contribution of our work lies in our ability to distinguish between the different types of aggregate structures formed and to learn what properties of hydrophobic crowdors lead to their formation. The combination of a large system size, relatively realistic peptide model, and hydrophobic crowdors rather than universally attractive crowdors gives us a molecular-level perspective of how a more realistic crowder environment may alter protein aggregation. The systems considered here contain 48 A $\beta$ (16-22) peptides and crowdors of diameters 5Å, 20Å, and 40Å, represented here by simple hard spheres or hydrophobic crowdors, at a crowder volume fraction of 0,10. Our results show that specific combinations of crowder size and interaction strengths can create an

environment in which the predominant species at the end of the simulation are either disordered oligomers or ordered  $\beta$ -sheet oligomers rather than the fibrillar structures that we typically observe in our aggregation simulations.

**Chapter 4.** Using the combination of DMD and our PRIME20 force field we were able to simulate the interaction of a single (KLLK)<sub>4</sub> amphipathic peptide and of 48 (KLLK)<sub>4</sub> peptides with an aqueous/organic interface. Many simulations of peptides at interfaces have been performed by other investigators but, to our knowledge, none have observed the spontaneous folding of amphipathic  $\alpha$ -helical peptides at interfaces or the aggregation of proteins at interfaces for such large systems . The peptides considered in the work are initially in random coil or  $\alpha$ -helical conformations and interact with strong and weak interfaces. Our results show that the weak interface is capable of folding a randomly coiled (KLLK)<sub>4</sub> peptide into an amphipathic  $\alpha$ -helix, while the strong interface disrupts intrapeptide hydrogen bonding and forces the peptides to adopt a random coil or  $\beta$ -hairpin conformation. We also found that when peptides interacted with the strong interface they adsorbed irreversibly due to the strong side chain preference for their respective phases. When peptides interacted with the weak interface they adsorbed reversibly. In the latter case the (KLLK)<sub>4</sub> peptide was more likely to bind to the interface when it was initially in an  $\alpha$ -helical conformation as opposed to a random coil because it had a larger hydrophobic moment. We also studied the aggregation of the (KLLK)<sub>4</sub> peptide in the presence of an aqueous/organic interface. We found that the conditions producing the highest levels of aggregation occurred when the peptides started out in helical conformations and interacted

with a weak interface. Since the peptides could bind to the interface and reversibly dissociate, the increased local concentration of helices in the vicinity of the interface promoted the formation of helical bundles. The most energetically favorable type of aggregate in the aqueous phase was the helical bundle, and the most energetically favorable type of aggregate at the interface was the  $\beta$ -sheet.

**Chapter 5.** Using the combination of DMD and our PRIME20 force field we were able to simulate systems of phenolic aggregation inhibitors with the A $\beta$ (17-36) peptide. Much simulation and experimental work has been put into aggregation inhibitor research to help combat AD, but none have looked at systems of aggregation inhibitors and peptide on this large size scale. Atomistic simulations are limited to docking studies or molecular dynamics/Monte Carlo on systems consisting of only a few peptides, typically arranged in a preformed fibrillar structure, and inhibitors. For the first time we are able to observe how these inhibitors interact with systems of peptides starting in a random coil all the way to the formation of the U-loop structure characteristic of amyloid structures.

## **6.2 Future Recommendations**

In this dissertation, we have discussed the effects of macromolecular crowding, interfaces, and inhibitors on protein aggregation using the PRIME20 force field. The scope of this work can be expanded and some future recommendations are summarized below.

### 6.2.1 Interface Monolayer Simulations

One important phenomenon that we did not observe in our studies of the folding and aggregation of amphipathic peptides at an interface was the formation of peptide monolayers at the aqueous/organic interface. We believe that the reason for this is twofold: First, there were not enough peptides in the simulation to have enough coverage of the interface to form a monolayer, and second the large amounts of aggregation in the aqueous phase prevented monolayer formation at the interface. In order to simulate the formation of monolayers at an aqueous/organic interface we propose the following ideas. We will modify the interface aggregation code to include a monolayer interaction boundary. From the center of the simulation box to the monolayer interaction boundary, peptide-peptide interactions will be turned off so that there is no aggregation in the aqueous phase, but peptides can still fold. Once the peptide passes the monolayer interaction boundary the peptide-peptide interaction will be turned back on so that peptide aggregation can occur, but only within a predefined distance set by the boundary. This will allow us to localize aggregation to the interface and observe monolayer formation. Tuning this difference may also allow us to observe different monolayer thicknesses. Secondly, we will use more proteins to obtain the interface coverage necessary to form a monolayer. While this will slow down the simulation, the effects will not be as drastic because the monolayer interaction boundary will turn off the slow aggregation steps in the aqueous phase. Different protein concentrations may also lead to different monolayer thicknesses. Figure 6.1 shows an approximate location of the proposed monolayer interaction boundary as a dotted line and the aqueous/organic interface is a solid line.

### 5.2.2 Inhibitor Simulations

While the work presented in Chapter 5 on inhibitor interactions with A $\beta$ (17-36) is a good start, there are many changes we would like to make before the project is complete to ensure more accurate simulations. First we would like to use either A $\beta$ (17-42) and A $\beta$ (1-42) for the simulations rather than the smaller peptide we performed our first simulations with. The simulation of the full length protein is a long term goal and will take some refinement, but using the same approach as we did for the A $\beta$ (17-36) peptide and enhancing key interactions may allow us to simulate the formation of the U-loop structure in a reasonable amount of time. Second, we need to refine the geometric and energetic parameters of the inhibitors themselves. There are currently some discrepancies in the model that do not make physical sense, such as the C bead being smaller than the O bead. We need to come up with better parameters for each coarse grained group while maintaining compatibility with PRIME20. The covalent and pseudo bonds also need to be refined as there are currently structural fluctuations that may not be realistic. In order to do this we will obtain radial distribution functions between all relevant atoms for each inhibitor and define pseudobonds unique to those pairs rather than generalized values as presented Chapter 5.

### 6.2.3 Expanding the PRIME20 Force Field

The PRIME20 forcefield has proved to be very good at predicting protein aggregation, but there are a few things that could be done to improve its accuracy and flexibility. First, the hydrogen bonding routine is written specifically to perform protein backbone-backbone hydrogen bonds between coarse-grained N-H and C=O groups. We

propose to develop a more generalized hydrogen bonding scheme to increase the flexibility of the model for future studies based on the STRIDE secondary structure assignment algorithm from Frishman and Argos.<sup>1</sup> This method will be beneficial because it is not based solely on protein geometry and allows for the incorporation of other hydrogen bonding groups. Figure 6.2 shows the hydrogen bonding system as described by Frishman and Argos. The hydrogen bonding energy  $E_{HB}$  is a function of the distance between hydrogen bonding groups,  $E_r$ , the orientation of the hydrogen bonding groups,  $E_p$ , and the orientation of the hydrogen atom relative to the lone pairs of the donor,  $E_t$ . Equations 6.1-6.4 describe these terms.

$$E_{HB} = E_r E_p E_t \quad (6.1)$$

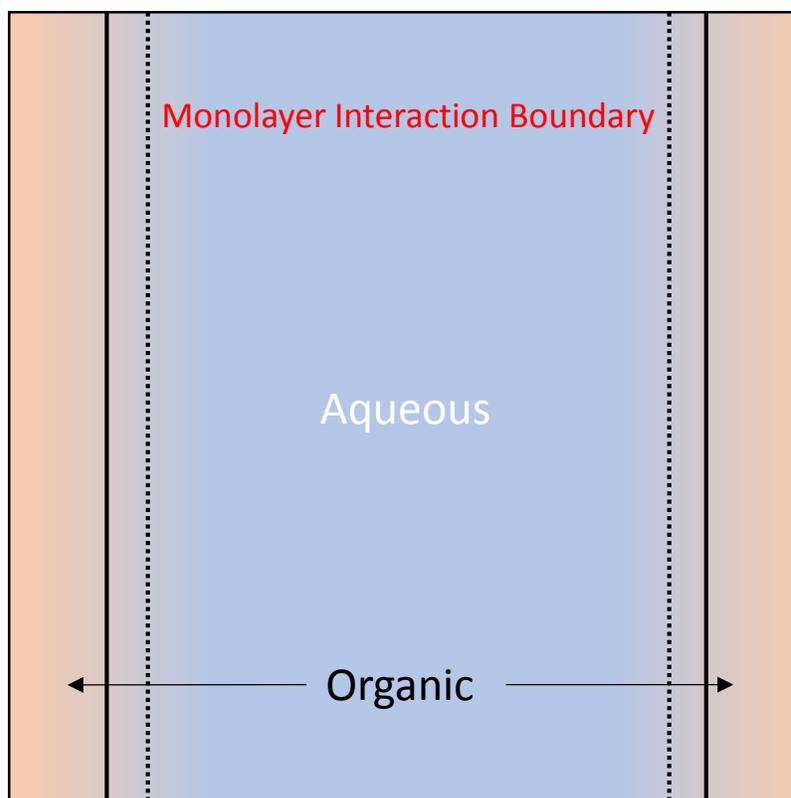
$$E_r = \frac{C}{r^8} - \frac{D}{r^6} \quad (6.2)$$

$$E_p = \cos^2 \theta \quad (6.3)$$

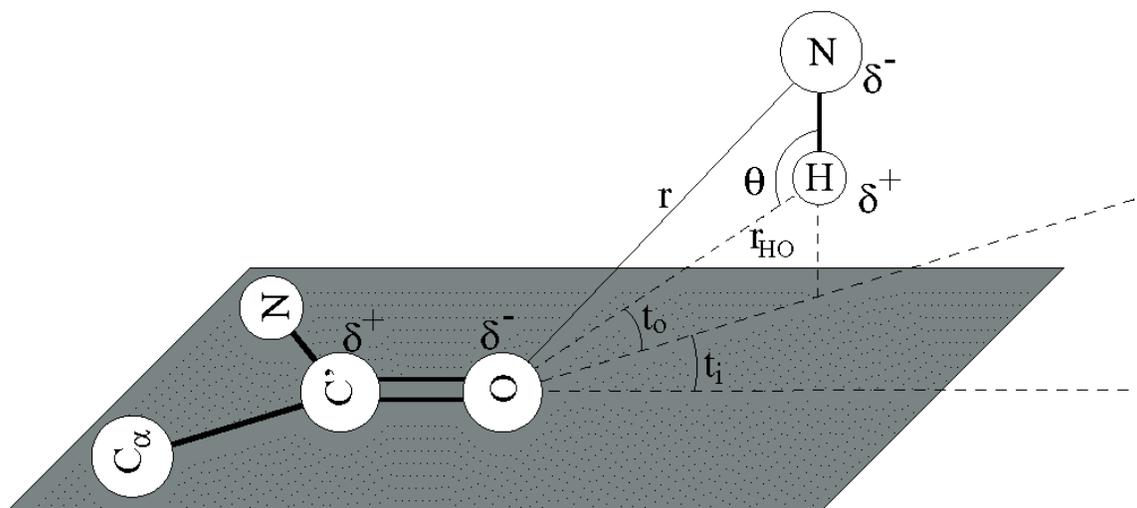
$$E_t = \left\{ \begin{array}{l} (0.9 + 0.1 \sin 2t_i) \cos t_o \text{ for } 0 < t_i < 90^\circ \\ \frac{K_1 (K_2 - \cos^2 t_i)^3 \cos t_o}{0 \text{ for } t_i > 110^\circ} \text{ for } 90^\circ < t_i < 110^\circ \end{array} \right\} \quad (6.4)$$

In these equations  $C = -3E_m r_m^8$ , and  $D = -4E_m r_m^6$ , where  $r$  is the distance between the donor and acceptor atoms participating in the hydrogen bond,  $E_m$  is the “optimal” hydrogen bond energy, and  $r_m$  is the optimum hydrogen bond length. The values for  $E_m$  and  $r_m$  could be easily made to match the current values used in PRIME20. They proposed an additional constraint that  $E_r = E_m$  for  $r < r_m$ . In Equation 6.4,  $K_1 = 0.9 / \cos^6 110^\circ$  and  $K_2 = \cos^2 110^\circ$ . Figure

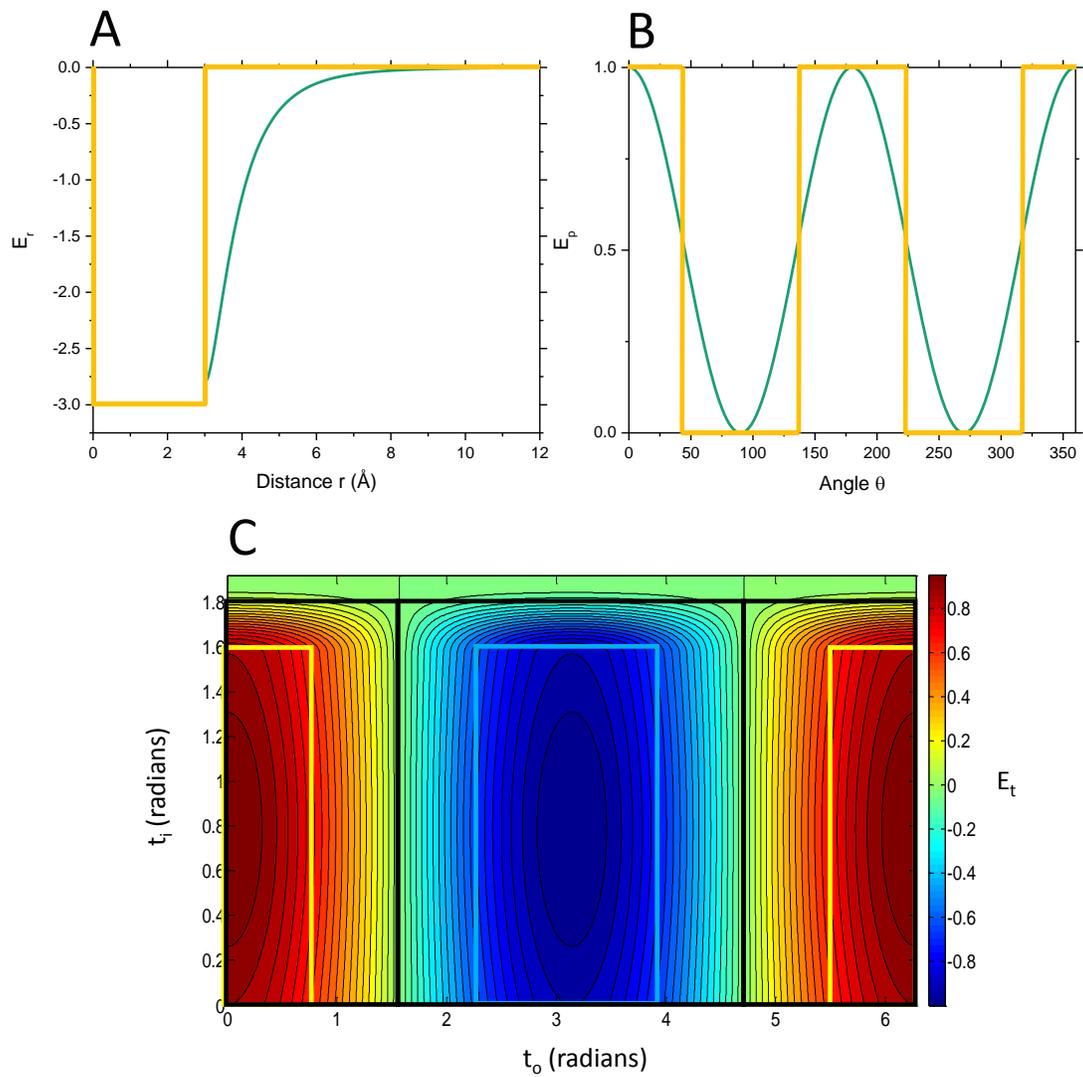
6.3 shows the functions A)  $E_r$ , B)  $E_p$ , and C)  $E_t$  with example discontinuous boundaries to maintain compatibility with PRIME20. New hydrogen bonding energies and interaction could be integrated into PRIME20 much more easily with this algorithm because it is not specific to peptide backbone-backbone hydrogen bonds. One final modification would be the geometry of the amino acid side chains. In unpublished work investigating D-form peptide aggregation inhibitors we found that the geometry for aromatic side chains is not well described by a single sphere and lacks pi-pi stacking interactions. Representing each ring in an aromatic side chain as three smaller beads, the least amount required to maintain a geometric plane, with the same interaction parameters would at the very least provide information about the stacking of aromatic side chains. Future work could expand this to include an orientation interaction as well. In general, representing each side chain with more detail but maintaining current interaction parameters may be an easy way to begin approaching the modification of PRIME20 to include more detail without limiting advantages of coarse-graining. Figure 6.4 shows an example of this method using Phenylalanine with similarly colored groups having the same interaction parameters.



**Figure 6.1.** Simulation box with the proposed monolayer interaction boundary as a dotted line and the aqueous/organic interface as a solid line.

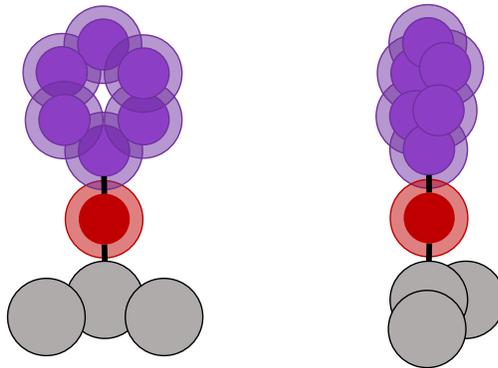


**Figure 6.2.** Hydrogen bonding system from Fishman and Argos.<sup>1</sup>

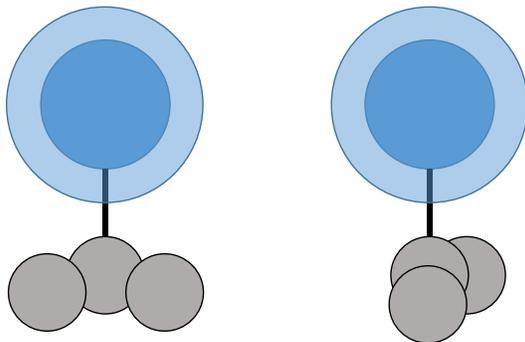


**Figure 6.3.** Functions A)  $E_r$ , B)  $E_p$ , and C)  $E_t$  with example discontinuous boundaries to maintain compatibility with PRIME20.

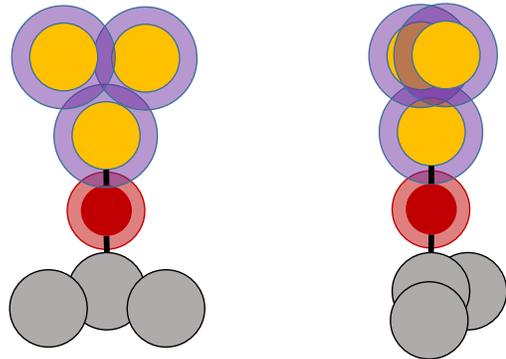
## Atomistic



## PRIME20



## Proposed Model



**Figure 6.4.** Example of proposed coarse graining method using Phenylalanine. Similarly colored groups have the same interaction parameters.

## 6.3 References

1. Frishman D, Argos P. Knowledge-based protein secondary structure assignment. *Proteins-Structure Function and Genetics*. 1995;23(4):566-579. doi: 10.1002/prot.340230412.