

Molecular Modeling of the Amyloid β -Peptide: Understanding the Mechanism of Alzheimer's Disease and the Potential for Therapeutic Intervention

Justin Alan Lemkul

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Biochemistry

David R. Bevan, Chair
Richard F. Helm
Sean F. O'Keefe
Webster Santos
Pablo Sobrado

March 7, 2012
Blacksburg, Virginia

Keywords: Simulations, Neurodegeneration, Thermodynamics, Protein-lipid interactions, Free energy calculations

Copyright 2012, Justin Alan Lemkul

Molecular Modeling of the Amyloid β -Peptide: Understanding the Mechanism of Alzheimer's Disease and the Potential for Therapeutic Intervention

Justin Alan Lemkul

ABSTRACT

Alzheimer's disease is the leading cause of senile dementia in the elderly, and as life expectancy increases across the globe, incidence of the disease is continually increasing. Current estimates place the number of cases at 25-30 million worldwide, with more than 5.4 million of these occurring in the United States. While the exact cause of the disease remains a mystery, it has become clear that the amyloid β -peptide ($A\beta$) is central to disease pathogenesis. The aggregation and deposition of this peptide in the brain is known to give rise to the hallmark lesions associated with Alzheimer's disease, but its exact mechanism of toxicity remains largely uncharacterized.

Molecular dynamics (MD) simulations have achieved great success in exploring molecular events with atomic resolution, predicting and explaining phenomena that are otherwise obscured from even the most sensitive experimental techniques. Due to the difficulty of obtaining high-quality structural data of $A\beta$ and its toxic assemblies, MD simulations can be an especially useful tool in understanding the progression of Alzheimer's disease on a molecular level.

The work contained herein describes the interactions of $A\beta$ monomers and oligomers with lipid bilayers to understand the mechanism by which $A\beta$ exerts its toxicity. Also explored is the mechanism by which flavonoid antioxidants may prevent $A\beta$ self-association and destabilize toxic aggregates, providing insight into the chemical features that give rise to this therapeutic effect.

This work has been supported by the Macromolecular Interfaces with Life Sciences (MILES) Integrative Graduate Education and Research Traineeship (IGERT) of the National Science Foundation under Agreement No. DGE-0333378, and by the Institute for Critical Technology and Applied Science (ICTAS) at Virginia Tech. Computing time on the SystemX and Athena supercomputers was provided by Advanced Research Computing at Virginia Tech.

Acknowledgments

I have found it difficult to find a way to properly thank everyone who has been supportive of this work and my development as a person and a scientist. I will do my best here in the limited space I am allowed.

First, I must thank my adviser, Dr. David Bevan. I will forever be grateful for the research and professional opportunities he has afforded me, as well as his willingness to teach me, advise me, motivate me, and push me to accomplish more than I ever thought possible. I joined the lab as a sophomore, having completed only basic biology and chemistry courses, but he took a chance on me early in my undergraduate career, and I have been ever grateful for it.

I must express my gratitude to my committee members, Dr. Richard Helm, Dr. Pablo Sobrado, Dr. Sean O’Keefe, and Dr. Webster Santos, for their continued commentary and constructive criticism of my work. Because of their support and continued interest in my projects, I have been able to achieve the goals I set for myself at the outset of my graduate career. The quality of this dissertation derives in large part from their diverse scientific viewpoints and high standards, and I appreciate the constant challenge of answering their questions and discussing the procedures involved in my studies.

Other members of the faculty have encouraged my development over the years and I would be remiss to not mention them here as well. I thank Dr. E. M. Gregory for always pushing me to be a true “card-carrying biochemist,” despite my computational tendencies. I thank Dr. Erin Dolan for teaching me to think critically and thoroughly about scientific problems in ways that no one else could. Dr. Diego Troya provided me with an insight into the molecular world that forever changed how I will think about life and its underlying physics and chemistry. I am also grateful to the MILES-IGERT co-directors, Dr. Susan Duncan and Dr. Tim Long, for constantly challenging me to step outside my biochemical comfort zone and apply my skills to critical problems. I thank the ICTAS program and its director, Dr. Roop Mahajan, for inspiring me to set high standards, and to not only meet them, but exceed them.

My path as a student and scientist began long before I entered college, and I must thank Mr. Edward Burt, my high school calculus teacher, for giving me one of the highest compliments I ever received, a phrase that has motivated much of my success: “I’m proud of you for sticking with it.” I had gotten a deceptive integral problem wrong (by nothing more than a negative sign), but nevertheless I hadn’t given up. His message came to mind every time something wasn’t working

out, be it in my undergraduate studies or in pursuit of my Ph.D. He wouldn't let me give up on myself, so I never did, either. To quit would be an insult to his faith in me, and so I kept going.

I had help along the way from colleagues and friends in the lab, especially Joe Allen, Nikki Lewis-Huff, and Somaye Badiyan, whose continued input on my work has been invaluable. Thanks for being my constant sounding board, source of sanity, proofreaders, and resource providers.

I thank all of my "e-colleagues" on the GROMACS users' list, especially Mark Abraham, Tsjerk Wassenaar, Chris Neale, Michael Shirts, David Mobley, and Xavier Periole, for teaching me so much, from my first forays into GROMACS in constructing a membrane protein system to the more advanced techniques that I used later on.

Without the support of friends and family, I would not have accomplished any of the work described here. My friends have been an integral part of my learning and living experience, and I owe a debt of gratitude (whether they know it or not) to Emily and Chris Schoonover, Kaitlin and Jay Read, Michelle Oppenheimer, Wyatt Chocklett, Sahil Khanna, Shivan Desai, Will Harkcom, Caroline Osborne, and many others who have contributed to my support network, both personal and professional.

To Jen and Will, thanks for setting the example I always tried to live up to and for teaching me what it means to be a Hokie. Without a path to follow, I surely would have been lost.

I thank my parents for always believing in me and providing me with every opportunity to succeed, but always giving me the freedom to use those opportunities as I saw fit.

Most especially, I thank my wife Kelly, for her unfailing faith in me that sustained me through even the most stressful times.

As I am about to graduate again, I am reminded of the circumstances of my first college graduation. Having endured an unimaginable tragedy on April 16, 2007, the Virginia Tech community came together in an astonishing display of what it means to be a member of this University. I am, and forever will be, a Hokie, a fact I state with pride wherever I go. On that fateful day, we were reminded, all too harshly, that life is precious and short. We must make the most of each day that we have. Live for 32.

Contents

1	Introduction	1
1.1	Alzheimer's Disease	1
1.1.1	Discovery and Definition of Alzheimer's Disease	1
1.1.2	The Impact of Alzheimer's Disease on Society	2
1.2	The Amyloid Hypothesis	2
1.2.1	Histological Observations	2
1.2.2	Discovery of $A\beta$ as the Principal Component of Senile Plaques (1985)	3
1.2.3	The Formation of the Amyloid Hypothesis (1992)	4
1.2.4	The Biology of $A\beta$	4
1.2.5	Structures of $A\beta$ Aggregates	10
1.2.6	Interactions of $A\beta$ with Membranes	11
1.2.7	The Role of $A\beta$ in Oxidative Stress	13
1.2.8	Current Therapeutics for Alzheimer's Disease	17
1.3	Organization of the Dissertation	19
1.4	Attribution	20
2	Molecular Dynamics Simulations	21
2.1	Newtonian Mechanics	21
2.2	Force Fields	23
2.2.1	Types of Force Fields	24
2.3	Time Scales of MD Simulations	25

3	Practical Considerations for Running Molecular Dynamics Simulations	26
3.1	Introduction	26
3.2	Comparison of Different Force Fields	26
3.2.1	Simulation Systems	27
3.2.2	Results of Force Field Comparison	28
3.3	Conclusions	37
4	Interactions Between Monomeric $A\beta$ and a Model DPPC Membrane	39
4.1	Introduction	39
4.2	Dynamics of the $A\beta_{40}$ Peptide	41
4.2.1	Methodology	41
4.2.2	Results	43
4.2.3	Discussion	48
4.3	$A\beta_{40}$ Disrupts Equilibrium Lipid Properties	52
4.3.1	Methodology	52
4.3.2	Results	53
4.3.3	Discussion	65
5	Interactions Between Monomeric $A\beta$ and Model Membranes and Lipid Rafts	75
5.1	Introduction	75
5.2	Methods	76
5.2.1	System Construction	76
5.2.2	Molecular Dynamics Simulations	78
5.2.3	Analysis	79
5.3	Results	79
5.3.1	Secondary Structure of $A\beta_{40}$	80
5.3.2	Hydrogen Bonding Between $A\beta_{40}$ and Lipids	89
5.3.3	Position and Orientation of $A\beta_{40}$ in Membranes	96
5.4	Discussion	104

5.5	Conclusions	107
5.6	Acknowledgments	108
6	Interactions Between Multimeric Aβ and Lipid Bilayers	109
6.1	Introduction	109
6.2	Methods	109
6.3	Results and Discussion	112
6.3.1	Dimerization of A β_{40} Peptides	112
6.3.2	Position and Orientation of A β_{40} in Membranes	114
6.3.3	Secondary Structure of A β_{40} Peptides	115
6.3.4	Effect of A β_{40} on Membrane Structure	116
6.4	Conclusions	117
7	Thermodynamics of Aβ Protofibril Assembly	121
7.1	Introduction	121
7.2	Steered Molecular Dynamics	121
7.3	Umbrella Sampling	122
7.4	Methodology	123
7.4.1	SMD Simulations	123
7.4.2	<i>In silico</i> Mutations	124
7.4.3	Umbrella Sampling Simulations	126
7.5	Results	127
7.5.1	Structural Details of A β Protofibrils	127
7.5.2	SMD Simulations	131
7.5.3	Umbrella Sampling Simulations	137
7.6	Discussion	138
7.7	Conclusions	141
8	Destabilizing Aβ Protofibrils with Morin	142
8.1	Introduction	142

8.2	Derivation of Morin Parameters	143
8.3	Simulations of $A\beta$ Protofibrils in the Presence of Morin	146
8.3.1	Methods	146
8.3.2	Results	151
8.4	Conclusions	188
9	Inhibiting Early $A\beta$ Aggregation Events with Morin	192
9.1	Introduction	192
9.2	Methods	193
9.2.1	System Construction	193
9.2.2	Simulation Protocol	193
9.3	Results	194
9.3.1	Simulations of $A\beta_{40}$ Monomers	194
9.3.2	Simulations of $A\beta_{42}$ Monomers	200
9.3.3	Effect of Morin on $A\beta_{40}$ Dimerization	203
9.3.4	Effect of Morin on $A\beta_{42}$ Dimerization	209
9.3.5	Binding of Morin to Fibril-derived $A\beta_{40}$ and $A\beta_{42}$ Dimer	214
9.3.6	Binding of Morin to an Oligomer-derived $A\beta_{42}$ Dimer	218
9.4	Discussion	220
9.5	Conclusions	224
9.6	Acknowledgments	225
10	Development of GridMAT-MD for Membrane Analysis	226
10.1	Introduction	226
10.2	Description	227
10.2.1	Processing the coordinate file	228
10.2.2	Calculating thickness	229
10.2.3	Calculating area per lipid headgroup	232
10.3	Validation Simulations	232

10.3.1	Methods	232
10.3.2	Results	233
10.4	Summary	234
11	Proper Parameterization of Small Molecules for the GROMOS96 Force Fields	237
11.1	Introduction	237
11.2	Methods	238
11.2.1	Condensed-phase and gas-phase systems	239
11.2.2	Interfacial amino acid systems	239
11.2.3	UDP-galactopyranose mutase	241
11.3	Results and Discussion	245
11.3.1	Condensed-phase and gas-phase systems	245
11.3.2	Interfacial amino acid systems	247
11.3.3	UDP-galactopyranose mutase	255
11.3.4	Strategies for charge assignment in GROMOS topologies	257
11.4	Conclusions	274
11.5	Acknowledgments	275
12	Conclusions	276
	Bibliography	279
	Appendix A Cholesterol Topology	319
	Appendix B Palmitoylsphingomyelin Topology	326
	Appendix C Ganglioside GM1 Topology	334
	Appendix D Example .mdp Files	353
D.1	Generic twin-range cutoff NPT simulation (GROMOS96)	353
D.2	Simulated annealing simulation for a membrane protein (GROMOS96)	355

D.3	NPT simulation involving morin (GROMOS96)	357
D.4	Steered MD simulation	359
D.5	Umbrella sampling simulation	361

Appendix E Perl Scripts **363**

E.1	Average multiple data files	363
E.2	Calculate x - y distances	366
E.3	Parse a data set	368
E.4	Plot the contents of hbmap.xpm	370
E.5	Plot secondary structure frequencies from ss.xpm	375
E.6	Create g_sham input	379
E.7	Create index groups of per-leaflet reference atoms	382
E.8	Measure effective chain length and chain tilt angle	386
E.9	Calculate the orientation of a lipid headgroup	396

List of Figures

1.1	Proteolysis of APP	6
1.2	The A β sulfuranyl radical	16
3.1	NMR structure of the A β ₄₀ peptide	27
3.2	Secondary structure, GROMOS96 43A1 + SPC water	29
3.3	Secondary structure, GROMOS96 53A6 + SPC water	29
3.4	Secondary structure, OPLS-AA + TIP4P water	29
3.5	Secondary structure, OPLS-AA + SPC water	30
3.6	Secondary structure, AMBER03 + TIP3P water	30
3.7	Secondary structure, AMBER03 + SPC water	31
3.8	Secondary structure evolution and probabilities, GROMOS96 43A1 + SPC water	32
3.9	Secondary structure evolution and probabilities, GROMOS96 53A6 + SPC water	33
3.10	Secondary structure evolution and probabilities, OPLS-AA + SPC water	34
3.11	Secondary structure evolution and probabilities, OPLS-AA + TIP4P water	35
3.12	Secondary structure evolution and probabilities, AMBER03 + TIP3P water	36
3.13	Secondary structure evolution and probabilities, AMBER03 + SPC water	37
3.14	Backbone RMSF	38
4.1	The structure of DPPC	41
4.2	Final snapshots of A β ₄₀ -DPPC systems	45
4.3	Secondary structure progression of A β ₄₀ -DPPC systems	46
4.4	Water leakage into the DPPC bilayer	48

4.5	Distortions in the DPPC membrane caused by $A\beta_{40}$	56
4.6	Partial densities along the z -axis	58
4.7	Membrane thickness composites	60
4.8	Area per lipid headgroup as a function of time, for simulation set A	62
4.9	Area per lipid headgroup as a function of time, for simulation set B	63
4.10	Area per lipid headgroup as a function of time, for simulation set C	63
4.11	Area per lipid headgroup as a function of time, for simulation sets E and P	64
4.12	Area per lipid headgroup as a function of time, for simulation sets OS and NS	65
4.13	Area per lipid headgroup as a function of distance, A1 and A2	66
4.14	Contacts between the $A\beta_{40}$ peptide in simulation A2 and the lipids of the top leaflet	67
4.15	Area per lipid headgroup as a function of distance, A3 and A4	68
4.16	Area per lipid headgroup as a function of distance, B1 and B2	69
4.17	Area per lipid headgroup as a function of distance, B3 and B4	70
4.18	Area per lipid headgroup as a function of distance, C1 and C2	71
4.19	Area per lipid headgroup as a function of distance, E1 and E2	72
4.20	Area per lipid headgroup as a function of distance, P1 and P2	73
4.21	Effective chain length and acyl chain tilt angle	74
5.1	Lipid structures	77
5.2	Secondary structure evolution for $A\beta_{40}$ model CH in POPC	81
5.3	Secondary structure evolution for $A\beta_{40}$ model CI in POPC	82
5.4	Secondary structure evolution for $A\beta_{40}$ model CH in POPS	83
5.5	Secondary structure evolution for $A\beta_{40}$ model CI in POPS	84
5.6	Secondary structure evolution for $A\beta_{40}$ model CH in POPC-POPE	85
5.7	Secondary structure evolution for $A\beta_{40}$ model CI in POPC-POPE	86
5.8	Secondary structure evolution for $A\beta_{40}$ model CH in rafts	88
5.9	Secondary structure evolution for $A\beta_{40}$ model CI in rafts	89
5.10	Free energy surfaces for $A\beta_{40}$ model CH in rafts	90
5.11	Free energy surfaces for $A\beta_{40}$ model CI in rafts	91

5.12	Helicity as a function of position	92
5.13	Hairpin formation induced by GM1	93
5.14	Starting structures of A β ₄₀ -raft and A β ₄₀ -raft + GM1 systems	97
5.15	Free energy surfaces	99
5.16	Structures of A β ₄₀ model CH peptides in POPC	100
5.17	Structures of A β ₄₀ model CI peptides in POPC	101
5.18	Structures of A β ₄₀ model CH peptides in POPS	102
5.19	Structures of A β ₄₀ model CI peptides in POPS	103
5.20	Structures of A β ₄₀ model CH peptides in POPC-POPE	104
5.21	Structures of A β ₄₀ model CI peptides in POPC-POPE	105
6.1	The starting configuration of the A β ₄₀ -POPC dimer system	111
6.2	Examples of “strong” and “weak” dimers	113
6.3	Distance matrices and images of A β ₄₀ for all trajectories in which dimers formed	119
6.4	Membrane thickness projections in the $x - y$ plane	120
7.1	A β protofibril with pulling axis	123
7.2	A β protofibril with key residues labeled	124
7.3	SMD pull rate comparison	127
7.4	A β protofibril structures	128
7.5	Internal hydration of the A β protofibril	129
7.6	Sidechain contacts in the A β protofibril	130
7.7	Twisting of the A β protofibril	130
7.8	Dissociation of peptide A in the WT A β protofibril	132
7.9	Dissociation pathways of WT and mutant protofibrils	134
7.10	Non-native packing of the K28A protofibril	135
7.11	PMF curves	139
8.1	Thermodynamic cycle for morin	144
8.2	The structure of morin, with atom names labeled.	146

8.3	Lennard-Jones contribution to ΔG	149
8.4	Coulombic contribution to ΔG	150
8.5	Snapshots from the end of Set I protofibril simulations	153
8.6	RMSD from Set I protofibril simulations	154
8.7	RMSF from Set I protofibril simulations	155
8.8	Backbone hydrogen bonds from Set I protofibril simulations	156
8.9	Asp23-Lys28 salt bridge distances from Set I protofibril simulations	157
8.10	Binding of morin to the $A\beta$ protofibril	158
8.11	Hydrophobic contacts between morin and the $A\beta_{42}$ protofibril in Set II simulations	159
8.12	Hydrogen bonds between morin to the $A\beta$ protofibril backbone	160
8.13	Hydrogen bonds between morin to the $A\beta$ protofibril sidechain groups	162
8.14	Snapshots from the end of Set II protofibril simulations.	163
8.15	RMSD from Set II protofibril simulations	165
8.16	RMSF from Set II protofibril simulations	166
8.17	Asp23-Lys28 salt bridge distances from Set II protofibril simulations	167
8.18	Backbone hydrogen bonds from Set II protofibril simulations	168
8.19	Initial placement of inserted morin and evolution of interactions over time	171
8.20	Snapshots from the end of Set III protofibril simulations.	172
8.21	Comparison of morin positions from Set II and Set III simulations.	172
8.22	RMSD from Set III protofibril simulations	173
8.23	RMSF from Set III protofibril simulations	174
8.24	Backbone hydrogen bonds from Set III protofibril simulations	175
8.25	Hydrogen bonds between morin to the $A\beta$ protofibril backbone	176
8.26	Hydrogen bonds between morin to the $A\beta$ protofibril sidechain groups	177
8.27	Asp23-Lys28 salt bridge distances from Set III protofibril simulations	178
8.28	Destabilization of the $A\beta$ protofibril by morin in the protofibril interior	181
8.29	Disruption of native salt bridges by morin	182
8.30	Morin inhibits the assembly of $A\beta$ protofibrils	183

8.31	Structures from the end of assembly simulations	184
8.32	Hydrogen bonds between free peptide and the untreated $A\beta$ protofibril	184
8.33	Hydrogen bonds between free peptide and the morin-bound $A\beta$ protofibril	185
8.34	Center-of-mass distance between free peptide and the control $A\beta$ protofibril	186
8.35	Center-of-mass distance between free peptide and the morin-bound $A\beta$ protofibril	186
8.36	Atomic contacts between free peptide and the control $A\beta$ protofibril	187
8.37	Atomic contacts between free peptide and the morin-treated $A\beta$ protofibril	187
9.1	$A\beta_{40}$ structures (M40 and M40 _M)	196
9.2	Gibbs free energy surfaces of R_g vs. $A\beta_{40}$ atomic contacts	197
9.3	$A\beta_{40}$ structures (M40 _{XM})	199
9.4	$A\beta_{42}$ structures (M42 and M42 _M)	200
9.5	Gibbs free energy surfaces of R_g vs. $A\beta_{42}$ atomic contacts	202
9.6	$A\beta_{42}$ structures (M42 _{XM})	204
9.7	Structures of D40 dimers	206
9.8	Structures of D40 _M dimers	207
9.9	Distance matrices and secondary structure frequency for D40 _M simulations	208
9.10	Structures of D42 dimers	210
9.11	Structures of D42 _M dimers	212
9.12	Distance matrices and secondary structure frequency for D42 _M simulations	214
9.13	Petkova $A\beta_{40}$ dimers	215
9.14	Lührs $A\beta_{42}$ dimers	217
9.15	Yu $A\beta_{42}$ dimers	219
10.1	The GridMAT-MD algorithm	228
10.2	The GridMAT-MD decision structure	230
10.3	Schematic diagram of the GridMAT-MD thickness calculation	231
10.4	Snapshots from the KALP ₁₅ trajectory	234
10.5	The KALP ₁₅ -DPPC system, at 50 ns, in a traditional rendering	235

10.6	Polygon representation of lipids in a membrane protein system	236
11.1	Small molecule charges and charge groups	240
11.2	Initial position of lysine in the cyclohexane-water interfacial system	241
11.3	Charges assigned to Ala, Asn, and Asp	242
11.4	Charges assigned to Ile, Lys, and Ser	243
11.5	Charges assigned to FAD	244
11.6	RDF for ethanol and <i>p</i> -cresol	247
11.7	Amino acid positions generated by GROMOS96 43A1	249
11.8	Amino acid positions generated by PRODRG	250
11.9	Electrostatic and Lennard-Jones interactions energies with water	251
11.10	Lysine sidechain RDF plots	252
11.11	Hydrogen bonds to water	254
11.12	Alignment of FAD structures	256
11.13	RMSF of the UGM backbone	258
A.1	Structure of cholesterol	319
B.1	Structure of palmitoylsphingomyelin	326
C.1	Structure of ganglioside GM1	334

List of Tables

1.1	Diseases Associated with Amyloidogenic Proteins	3
1.2	Mutations to APP and A β	8
3.1	Simulation Systems	28
3.2	Secondary Structure Content	31
3.3	Radius of Gyration	32
3.4	Intramolecular Contacts and Hydrogen Bonds	34
3.5	Backbone RMSD	36
4.1	Simulation Systems	42
4.2	pK _a Values, as Calculated by H++	43
4.3	Properties of Control DPPC Membranes	54
4.4	Properties of Melittin-Membrane Systems	54
4.5	Average Values of $-S_{CD}$ for A β_{40} -DPPC systems	57
4.6	Area per Lipid Headgroup	61
5.1	Average secondary structure content of A β_{40}	80
5.2	Hydrogen bonding between A β_{40} and lipids in all non-raft systems	94
5.3	Hydrogen bonding between A β_{40} and GM1 sugar moieties	95
5.4	Hydrogen bonding between A β_{40} and raft lipid groups	95
5.5	C-terminal tilt and position	98
6.1	Contents of the simulated systems	110
6.2	A β_{40} -lipid hydrogen bonds	114

6.3	Tilt angle and z-axis position of A β ₄₀ peptides in model membranes	115
6.4	Secondary structure content of A β ₄₀ peptides	116
7.1	Backbone RMSD values after 100 ns of MD simulation.	128
7.2	Results from Umbrella Sampling Simulations.	138
8.1	Nonbonded parameters for morin	147
8.2	Thermodynamic and log <i>P</i> data	148
8.3	Backbone RMSD from Set I protofibril simulations	151
8.4	Interpeptide backbone hydrogen bonds from Set I protofibril simulations	152
8.5	Interpeptide Asp23 - Lys28 distance from Set I protofibril simulations	152
8.6	Morin-backbone hydrogen bonds from Set II protofibril simulations	161
8.7	Morin-sidechain hydrogen bonds from Set II protofibril simulations	161
8.8	Backbone RMSD from Set II protofibril simulations	169
8.9	Asp23-Lys28 distances from Set II protofibril simulations	169
8.10	Interpeptide backbone hydrogen bonds from Set II protofibril simulations	170
8.11	Backbone RMSD from Set III protofibril simulations	170
8.12	Interpeptide backbone hydrogen bonds from Set III protofibril simulations	179
8.13	Morin-backbone hydrogen bonds from Set III protofibril simulations	179
8.14	Morin-sidechain hydrogen bonds from Set III protofibril simulations	179
8.15	Interpeptide hydrogen bonds in control Set IV simulations	185
8.16	Interpeptide hydrogen bonds in morin-treated Set IV simulations	185
8.17	Center-of-mass distances in control Set IV simulations	185
8.18	Center-of-mass distances in morin-treated Set IV simulations	186
8.19	Atomic contacts in control Set IV simulations	187
8.20	Atomic contacts in morin-treated Set IV simulations	188
9.1	Monomeric simulation system designators, contents, and trajectory lengths	194
9.2	Dimeric simulation system designators, contents, and trajectory lengths	195
9.3	Structural parameters for monomeric A β ₄₀ peptides	198

9.4	Structural parameters for monomeric A β ₄₂ peptides	201
11.1	Density and ΔH_{vap} for small molecules	246
11.2	Small molecule charges calculated by Antechamber	259
11.3	Alanine charges calculated by Antechamber	260
11.4	Isoleucine charges calculated by Antechamber	261
11.5	Asparagine charges calculated by Antechamber	262
11.6	Serine charges calculated by Antechamber	263
11.7	Aspartate charges calculated by Antechamber	264
11.8	Lysine charges calculated by Antechamber	265
11.9	Small-molecule charges calculated by Spartan '04	267
11.10	Alanine charges calculated by Spartan '04	268
11.11	Isoleucine charges calculated by Spartan '04	269
11.12	Asparagine charges calculated by Spartan '04	270
11.13	Serine charges calculated by Spartan '04	271
11.14	Aspartate charges calculated by Spartan '04	272
11.15	Lysine charges calculated by Spartan '04	273

Chapter 1

Introduction

1.1 Alzheimer's Disease

1.1.1 Discovery and Definition of Alzheimer's Disease

The neurodegenerative disorder now known as Alzheimer's disease was first described by German psychiatrist Alois Alzheimer, for whom the disease was later named. His patient, a woman named Auguste Deter, was exhibiting signs of memory loss, confusion, paranoia, drastic mood changes, and dementia. Following her death in 1906 at the age of 51, Dr. Alzheimer analyzed her records, performed a biopsy on her brain, and presented his findings in a seminar and subsequent publication in 1907, which has since been translated to English and republished [1]. Upon autopsy, Dr. Alzheimer noticed cerebral atrophy, proteinaceous plaques, and other tangled proteins in his patient's brain. These protein aggregates have become the main histological lesion in Alzheimer's disease. The elucidation of the identity of the molecular components of these aggregates was not to come for many years, during which time the neuronal atrophy Dr. Alzheimer described was considered a normal part of aging, though it was strange that a woman as young as Auguste Deter was experiencing symptoms more frequently seen in much older patients.

Until recent advances in brain imaging and the analysis of cerebrospinal fluid, the only tool available for diagnosing Alzheimer's disease was autopsy. Upon examination of the brain, neuronal dystrophy and death should be apparent as well as proteinaceous plaques, which, when stained, are visible under light microscopy. The molecular nature of these plaques and their precursors is the main focus of this work, and will be described in detail in the remainder of this Chapter.

1.1.2 The Impact of Alzheimer's Disease on Society

It is now known that Alzheimer's disease is not a normal part of aging, but rather a disease of elusive origin. It is the sixth leading cause of death in the United States and is the leading cause of senile dementia in the elderly [2]. Current estimates suggest that over 5.4 million individuals in the United States are afflicted by Alzheimer's disease, and perhaps upwards of 25 million individuals are affected worldwide [3, 4, 5]. The incidence of the disease is likely to increase over the coming years. With individuals living longer due to greater access to health care services and medication, the mean lifespan of people around the world is on the rise. Thus, the incidence of neurodegenerative disorders such as Alzheimer's disease will increase. Current projections predict that, in the absence of an effective therapeutic agent, the incidence of Alzheimer's disease among individuals 65 years or older may triple within the next 50 years [4].

The costs associated with Alzheimer's disease are significant; expenditures are in excess of \$183 billion annually in the United States alone, a figure that is expected to increase to \$1.1 trillion by the year 2050 [4]. Of these figures, 70% of the cost is paid by Medicare and Medicaid. The true economic impact of the disease is much higher, as this figure does not include payments made by private insurance companies or lost revenue due to missed work time for caregivers and family members of those afflicted with the disease. With the incidence of Alzheimer's disease on a continual rise, the economic burden this disease presents cannot be understated, and it is of critical importance to gain insight into the mechanism of this debilitating condition to more effectively treat it.

1.2 The Amyloid Hypothesis

1.2.1 Histological Observations

The term "amyloid" was coined by Rudolf Virchow in 1854. Virchow stained abnormal extracellular masses in brain samples with iodine and sulfuric acid, and upon the emergence of a violet color, he concluded that the masses were composed of cellulose or starch, though a distinction between these two species was not understood at the time. The name "amyloid" derives from the Latin "amylum" and Greek "amylon," meaning "starch" [6, 7]. In 1859, amyloid samples analyzed by Friederich and Kekule were found to contain protein [6, 7]. Thus emerged the more modern (and ultimately, correct) view that amyloid aggregates were principally composed of protein. The term "amyloid" now refers to a large class of diseases that result from the deposition of aggregated protein. Several of the most common amyloid diseases and their associated proteins are listed in Table 1.1.

When Alois Alzheimer examined the brain of his patient, Auguste Deter, he discovered tangled protein aggregates, some within the cell and others in the extracellular space in the tissue. Over time, it was discovered that the extracellular plaques were composed of a protein that became

Table 1.1: Diseases Associated with Amyloidogenic Proteins

Disease	Protein name
Alzheimer's	amyloid β -peptide
Parkinson's	α -synuclein
Diabetes mellitus type 2	islet amyloid polypeptide (IAPP)
Transmissible spongiform encephalopathy	PrP ^{Sc}
Huntington's	huntingtin
Atherosclerosis	apolipoprotein A1
Rheumatoid arthritis	serum amyloid A
Dialysis-related amyloidosis	β_2 microglobulin

known as the amyloid β -peptide ($A\beta$), while intracellular neurofibrillary tangles were composed of a hyperphosphorylated microtubule-associated protein, tau. Tau aggregation is not addressed here, as little is known about its implications on Alzheimer's disease, if there are any at all. Instead, the focus will be on the association of $A\beta$ with the progression of the disease.

1.2.2 Discovery of $A\beta$ as the Principal Component of Senile Plaques (1985)

Nearly 80 years after the first report of Alzheimer's disease, and over 100 years after the term "amyloid" was coined, two groups independently determined the identity of the core component of the extracellular plaques present in patients with Alzheimer's disease. The first group was led by Colin Masters, whose group reported the identity of a protein called " A_4 " in June, 1985 [8]. The name of the protein derives from "amyloid" and its molecular mass in the monomeric state (4 kDa). In that study, Masters *et al.* described a solubilization procedure for A_4 (low pH and strong denaturants) and the pH-dependence of its aggregation. The authors further concluded that the sequence of A_4 was identical to the aggregated protein observed in patients with Down syndrome, thus linking this disorder with Alzheimer's disease on a molecular level. It was also concluded that the A_4 protein was distinct from the scrapie protein, the causative agent in transmissible spongiform encephalopathy in goats and sheep.

In a subsequent work published in December, 1985, Wong *et al.* presented an immunohistochemical analysis of a protein they called " β protein" [9]. From the names assigned by Masters *et al.* and Wong *et al.* arose the modern name, the amyloid β -peptide, or $A\beta$. During the early analysis of this protein, various names were used, with some papers also using a hybrid of the first two names, " βA_4 ," among others that appear in the literature. The study by Wong *et al.* determined that antibodies could be raised against β protein, which were used to establish the presence of this protein in neural tissue of patients with Alzheimer's disease and Down syndrome, but not in healthy individuals.

1.2.3 The Formation of the Amyloid Hypothesis (1992)

The first formal proposal of what is now known as the “amyloid hypothesis” was presented by Hardy and Higgins in 1992 [10]. Citing the work of Masters *et al.* and Wong *et al.*, along with genetic evidence regarding $A\beta$ and its precursor, they proposed that the deposition of $A\beta$ in neural tissue was the pathological event in Alzheimer’s disease. The accumulation of $A\beta$ gave rise to the neuritic plaques observed in patients. Their hypothesis remains the prevailing view of the etiology of Alzheimer’s disease, though the underlying causes of the pathological behavior were unclear at the time, and to a large extent, remain so at present. The biological basis of the amyloid hypothesis is described in the following sections.

1.2.4 The Biology of $A\beta$

1.2.4.1 The Amyloid Precursor Protein

The human amyloid precursor protein (APP) is a single-pass, type I transmembrane protein with a large extracellular N-terminal domain and a smaller intracellular C-terminal domain. There are numerous isoforms of the protein (a result of alternative splicing), but 3 isoforms predominate, consisting of 695, 751, and 770 amino acids. The 695-residue form is the most prominent isoform of APP in the central nervous system; the other two forms are expressed in all tissues in the body [11]. APP is a member of a small gene family that includes APP-like proteins, APLP1 and APLP2 [11]. The $A\beta$ sequence within APP is unique to this gene; it is not present in APLP1 or APLP2. These three proteins evolved from a common ancestor and serve overlapping functions involved in neural development and the maintenance of synaptic plasticity [11]. The remaining discussion will focus directly on APP.

The APP gene is located on chromosome 21. In Down syndrome, this chromosome is present in triplicate, thus explaining the molecular similarities between the Alzheimer’s phenotype and that of patients with Down syndrome, that is, excess $A\beta$ deposition in the brain [8, 9]. APP is processed through the endoplasmic reticulum and Golgi apparatus on its way to the plasma membrane, and is extensively post-translationally modified [12]. It is during this cellular trafficking process that $A\beta$ is produced (discussed in Section 1.2.4.2).

Proteolysis of APP (Section 1.2.4.2) results in two principal products, in addition to $A\beta$ - a secreted extracellular domain called sAPP and the APP intracellular domain (AICD). The sAPP protein may take one of two forms, sAPP α or sAPP β , depending upon which protease cleaves it from APP (discussed in Section 1.2.4.2). sAPP functions in synaptic development and plasticity [13], cell differentiation [14], and neurite outgrowth [15]. sAPP may also bind to a variety of cellular receptors to modulate a number of pathways, as reviewed in [16]. The AICD has several putative functions, as well. There is evidence suggesting that it may act as part of a transcription factor complex with Fe65 (a nuclear adapter protein) and Tip60 (a histone acetyltransferase) [17], which then can interact with several different DNA-binding domains to modulate transcription of

a number of genes, including tumor suppressor protein Kall1 [18], the lipoprotein receptor LRP1 that affects apolipoprotein E and cholesterol metabolism [19], epidermal growth factor receptor (EGFR) [20], and a gene that produces a protease that can degrade A β [21]. Another study has suggested that C99, the precursor to AICD (produced after β -site cleavage but before γ -site cleavage (Section 1.2.4.2), functions as a cholesterol sensor in the plasma membrane [22]. Thus it is clear that the two major domains of APP serve necessary functions in the cell. Putative functions for A β , the intervening sequence between these two domains, will be discussed in Section 1.2.4.4.

1.2.4.2 Sequential Proteolysis Events Give Rise to A β

The A β peptide is produced from APP through sequential proteolysis by two enzymes, β -secretase and γ -secretase. The β -secretase activity is carried out by a protease known as β -site APP cleaving enzyme 1 (BACE-1) [23, 24], which is localized predominantly in endosome and the late Golgi [12, 25]. The γ -secretase activity has been attributed to a protein complex containing presenilin, nicastrin, APH-1, and PEN-2 [26, 27], and is localized at the plasma membrane and other organelles in specialized microdomains known as lipid rafts [28, 29]. The actions of β - and γ -secretase constitute what is known as the “amyloidogenic pathway,” that is, they give rise to the toxic A β peptide. Production of A β can be precluded by the action of a third protease, α -secretase, which is a member of the ADAM metalloproteinase family [30]. α -Secretase is also localized at the plasma membrane [31], cleaving APP in the middle of the A β sequence, precluding β -site cleavage and generating a non-amyloidogenic peptide known as p3. A schematic of the amyloidogenic and non-amyloidogenic pathways is illustrated in Figure 1.1.

The A β peptide exists in several forms, ranging in size from 38 to 43 residues. The most common forms of A β are 40 and 42 residues in length. The sequence of A β_{42} is:



In healthy individuals, A β_{40} comprises 90% of the total A β pool [32]. In the disease state, A β_{42} is the principal species and is the main component of senile plaques. The different alloforms arise because of stepwise cleavage by the γ -secretase complex [28, 33]. The γ -secretase complex is believed to process its substrate in a processive manner, first cleaving its substrate (C83 or C99, see Figure 1.1) and then trimming off several residues (three to four) at a time. Different lineages of A β peptides arise based on the initial site of proteolytic cleavage and subsequent processing. This initial cleavage by γ -secretase at the ϵ cleavage site (downstream from the C-terminus of A β) is imprecise, generating A β precursors, either A β_{48} or A β_{49} [34, 35, 36, 37]. Stepwise removal of tripeptides from these proteins give A β_{40} (from A β_{49}) or A β_{42} (from A β_{48}) [38], and other peptides ranging in length from 38 to 43 residues, most commonly, though species as long as 46 residues exist.

Mutations in the proteolytically active subunit of the γ -secretase complex (presenilin 1 or 2, depending on the isoform) elevate the A β_{42} /A β_{40} ratio [39, 40, 41, 42, 43, 44]. It has been proposed

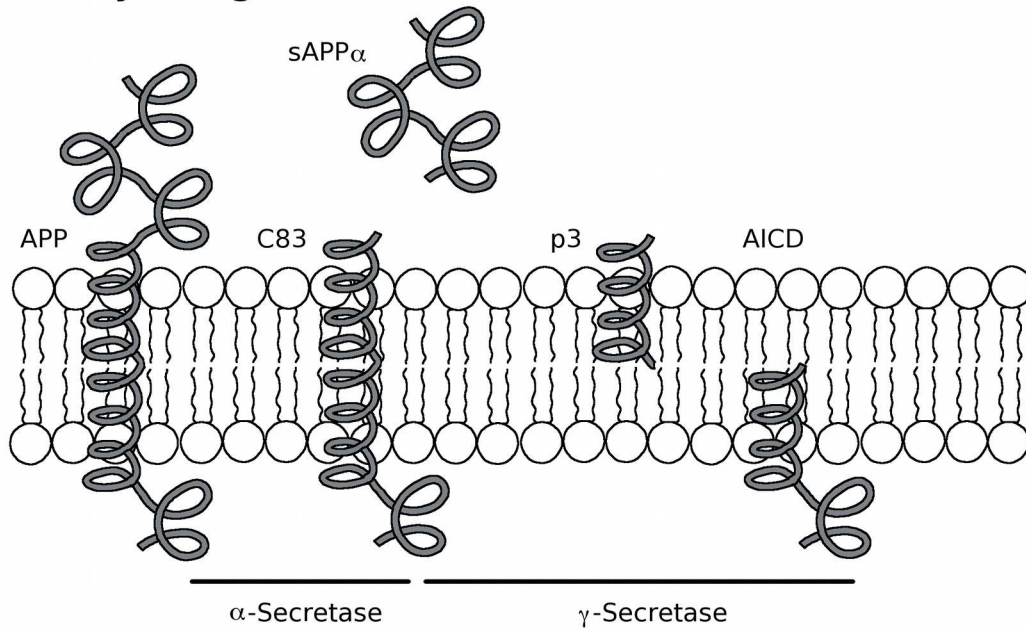
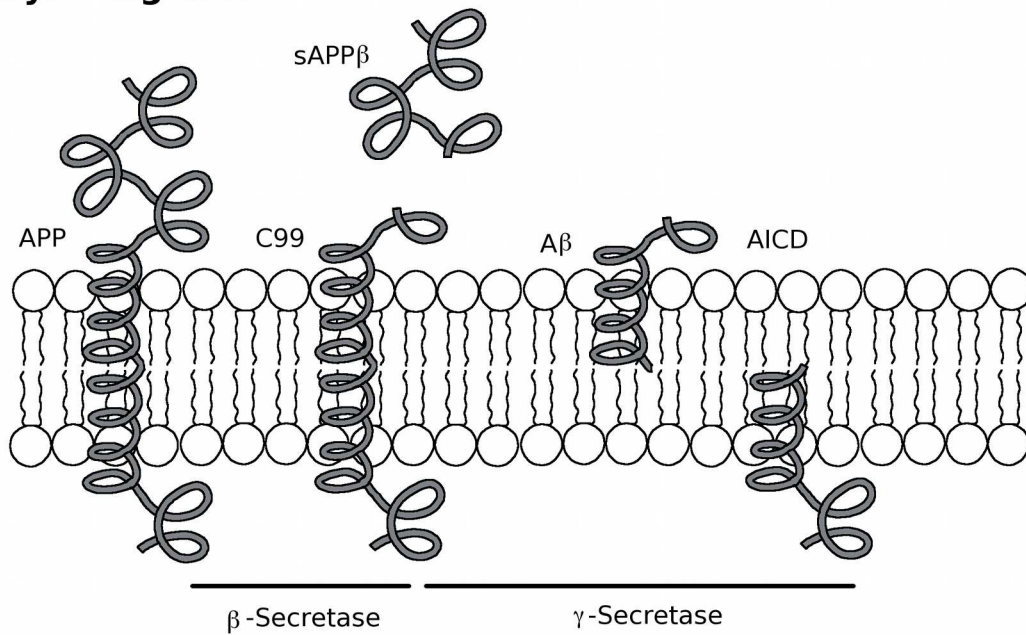
Non-Amyloidogenic**Amyloidogenic**

Figure 1.1: Proteolysis of APP. Both non-amyloidogenic and amyloidogenic pathways are shown. For illustrative purposes, the N-terminal domains (sAPP α and sAPP β) are not shown to scale. It is not implied that these proteolysis events occur in temporal or spatial proximity.

that the mutations lead to an enzyme that is catalytically slower, thus less processive, allowing longer forms of $A\beta$ to exit the enzyme active site [45, 46].

1.2.4.3 Mutations within APP and $A\beta$

The vast majority of Alzheimer's cases are sporadic, or otherwise not linked to inherited traits, with perhaps less than 1% of all cases being linked to a specific genetic predisposition [4]. There exist, however, a number of mutations both within the APP sequence (and some within the $A\beta$ sequence itself) that account for increased $A\beta$ toxicity or a more aggressive disease progression. Several of these aspects will be described here. Mutations also exist within the components of the γ -secretase complex that alter the manner in which APP is processed, but a discussion of these mutations (over 100 identified to date) is beyond the scope of this work.

Several of the most common mutations affecting $A\beta$ production and aggregation are listed in Table 1.2. Such mutations cause so-called "familial Alzheimer's disease" (FAD) and are named after the geographic regions in which they were first identified. These mutations may have an effect on the aggregation behavior of $A\beta$ itself or the manner in which APP is processed, giving rise to more aggregation-prone forms of $A\beta$.

1.2.4.4 Putative functions of $A\beta$

It has been difficult to determine the exact physiological function of $A\beta$. From an evolutionary standpoint, the acquisition of this peptide sequence in APP, but not in the genes of similar proteins (APLP1 and APLP2) would seem to indicate that it serves some necessary function, and that the disease associated with its deposition in neural tissue is a consequence of normal cellular processes gone awry.

It has been demonstrated that $A\beta$ decreases synaptic activity, the ability of neurons to communicate via electrical and chemical signals [59, 60]. This inhibition may be crucial to the prevention of excitotoxicity in normal cells, which could otherwise lead to seizures [61]. Another hypothetical role for $A\beta$ (one that also extends to APP) is in metal ion homeostasis. The $A\beta$ peptide has been shown to bind copper, iron, and zinc ions *in vitro* [62]. These results have suggested that $A\beta$ may have a unique role in controlling the levels of these pro-oxidant transition metals. As will be discussed in Section 1.2.7, $A\beta$ may promote oxidative damage through its interactions with these same metals when the metal:peptide ratio increases. Other investigators propose that $A\beta$ is simply an unfortunate byproduct of normal APP metabolism that is not efficiently cleared in the Alzheimer's brain [3].

A more recent hypothesis regarding the endogenous function of $A\beta$ is that it can act as part of the immune system by serving as an antimicrobial peptide [63]. Soscia *et al.* demonstrated that, *in vitro*, $A\beta$ possesses antimicrobial activity against numerous bacteria, including *Streptococcus*, *Staphylococcus*, and *Pseudomonas* species, among others. The authors further showed consid-

Table 1.2: Mutations to APP and A β

Mutation name	Mutation*	Effect(s)	Reference
Arctic	E22G	Increased A β protofibril formation	[47]
Austrian	T43I (T714I)	Increased A β_{42} /A β_{40} ratio	[48]
Dutch	E22Q	Increased fibrillogenesis Increased pathogenicity	[49]
Dutch/Iowa	E22Q/D23N	More potent pathogenicity than single mutant	[49]
Flemish	A21G	Increase in secreted A β Little fibril formation	[49]
Florida	I45V (I716V)	Increase in proportion of A β_{42} and A β_{43}	[50]
Indiana	V46F (V717F)	Risk factor for Alzheimer's disease?	[51]
Iowa	D23N	Increased fibrillogenesis Increased pathogenicity	[49]
Italian	E22K	More rapid aggregation Increased neurotoxicity	[52]
Italian	L34V	Severe blood vessel damage and hemorrhaging without plaque formation	[53]
Japanese	Δ E22	Increased fibrillogenesis Seeds A β_{39} fibrils	[54]
London	V46I (V717I)	Increased A β_{42} levels and plaque formation (mice)	[55, 56]
Spanish	I45F (I716F)	Early disease onset Increased A β_{42} /A β_{40} ratio	[57]
Swedish	KM(-1 - 2)NL	Increased BACE1 processing Increase in A β levels	[58]

*Residue numbering is based on position within the A β sequence. Residue numbers in parenthesis indicate residues outside of the normal A β sequence and are numbered as in the full-length APP. Negative numbers in the Swedish mutant are commonly used to indicate the position of the mutation relative to the β -secretase cleavage site (Asp1 in A β).

erable antimicrobial activity was present in brain homogenates of patients who had died from Alzheimer's disease. The antimicrobial activity was attributed to $A\beta$ after anti- $A\beta$ antibody treatment removed this activity. This antimicrobial hypothesis is particularly compelling, as it aligns with the fact that $A\beta$ kills neurons and has affinity for cell membranes (discussed in Section 1.2.6).

1.2.4.5 Degradation of $A\beta$ by Endogenous Proteases

Endogenous proteases exist that are capable of degrading free $A\beta$. One such candidate that has been identified is neprilysin, which, when inhibited, allows for the accumulation of $A\beta$ [64]. As stated previously, the expression of the neprilysin gene may be regulated by the AICD [21], and thus the interplay between $A\beta$ production and degradation may be self-regulating in normal, healthy individuals. Other putative $A\beta$ -degrading enzymes include endothelin-converting enzyme (ECE) [65] and insulin-degrading enzyme (IDE) [66].

1.2.4.6 The Role of Apolipoprotein E in $A\beta$ Production and Clearance

Though the most striking histological feature of Alois Alzheimer's initial characterization of Auguste Deter's brain was the tangled masses of proteins, he also reported abnormal lipid inclusions in the tissue as a final note in his initial report of the disease [1]. This feature of the underlying pathology of Alzheimer's disease received little attention until recently, but now the role of lipids in the disease progression is becoming more clear. Though there are numerous mutations associated with APP and $A\beta$ directly (discussed in Section 1.2.4.3), the single greatest genetic risk factor for the development of Alzheimer's disease turns out to be a particular form of apolipoprotein E (ApoE) [67]. It is worth noting, however, that certain ApoE genotypes do not guarantee the development of Alzheimer's disease, they only increase an individual's risk. Familial mutations in APP or $A\beta$ guarantee the eventual onset of the disease.

ApoE functions in the transport of lipoproteins, cholesterol, and lipid-soluble vitamins. It is the main cholesterol carrier in the brain. There are three major isoforms of the ApoE protein, ApoE- ϵ 2, ApoE- ϵ 3, and ApoE- ϵ 4 [68, 69]. The three isoforms differ only at positions 112 and 158 of the polypeptide chain. ApoE- ϵ 3 is considered the "normal" variant, with Cys112 and Arg158. Both ApoE- ϵ 2 and ApoE- ϵ 4 are considered dysfunctional forms of ApoE, with ApoE- ϵ 2 having Cys112 and Cys158 and ApoE- ϵ 4 having Arg112 and Arg158. ApoE- ϵ 2 has implications for atherosclerosis and is not known to be involved in the pathology of Alzheimer's disease. As such, ApoE- ϵ 4 will be the focus of this discussion.

ApoE- ϵ 4 affects the clearance of $A\beta$ from the brain. A protein called low-density lipoprotein receptor-related protein 1 (LRP1) is responsible for binding to $A\beta$ and transporting it across the blood-brain barrier (reviewed in [70]). Plasma forms of LRP1 are responsible for transporting $A\beta$ to the liver for degradation and clearance from the body. ApoE- ϵ 4 in the brain, but not ApoE- ϵ 2 or ApoE- ϵ 3, binds to LRP1 and inhibits this process, thus causing $A\beta$ to accumulate in the brain.

It is clear that cholesterol levels in the brain also play an important role in $A\beta$ generation. ApoE- $\epsilon 4$ most commonly binds to very low-density lipoproteins (VLDL), which contain high amounts of cholesterol. In contrast, the “normal” form of ApoE, ApoE- $\epsilon 3$, binds high-density lipoproteins (HDL), which have a lower cholesterol content. Thus, in addition to the potential role of ApoE- $\epsilon 4$ in retaining $A\beta$ in the brain, it may also function in altering cellular cholesterol levels that, in turn, affect APP processing and $A\beta$ generation. In one potential model (as reviewed in [71]), ApoE- $\epsilon 4$ binds VLDL and delivers them to the neuronal cell, thus increasing the amount of intracellular cholesterol. This cholesterol can remain in its free form in the plasma membrane, or can be converted to its esterified form and condensed in cytoplasmic lipid droplets. These droplets offer one possible explanation for the lipid inclusions first described by Alzheimer [1]. It has previously been shown that an increase in cellular cholesterol results in increased $A\beta$ production [72, 73, 74]. Subsequent studies demonstrated that the balance between free cholesterol and cholesteryl-esters is critical to understanding $A\beta$ generation; the formation of cholesteryl-esters is sufficient to up-regulate $A\beta$ production [75]. Free cholesterol may play a role in $A\beta$ generation, as well, since $A\beta$ production takes place in cholesterol-enriched lipid raft microdomains [76].

1.2.5 Structures of $A\beta$ Aggregates

To understand $A\beta$ aggregation, it is important to summarize the predominant forms in which the peptide may exist along the aggregation pathway from a soluble, monomeric peptide to the large fibrillar forms and plaques that are visible under a light microscope. Formation of $A\beta$ fibrils does not proceed linearly among intermediates; rather, the peptide may adopt conformations that lead to structures that can be classified into two groups, either “on-pathway” (those that lead to fibrils) or “off-pathway” (those that do not form fibrils). It is important to note that all known $A\beta$ aggregates are neurotoxic (to differing extents) and thus the structure of these species may provide insight into their role in causing the toxicity and neuronal loss associated with Alzheimer’s disease.

Based on structural studies in membrane-mimicking environments, it is currently believed that $A\beta$ is largely α -helical, with some disordered regions towards the termini, in the cell membrane [77, 78]. In solution, the $A\beta$ monomer is believed to be structurally dynamic. The tendency for the peptide to aggregate in aqueous solution has complicated efforts to obtain reliable structural information. The general view from studies of peptide fragments and model peptides is that $A\beta$ does not possess well-defined secondary structure, leading instead to a complex conformational ensemble of collapsed coil structures in aqueous media [79, 80]. The prevailing view of “on-pathway” aggregation dictates that partially structured, soluble $A\beta$ monomers aggregate to form low-molecular weight paranuclei, consisting of approximately six peptides. These paranuclei give rise to protofibrils (larger structures rich in β -strand content and approximately 5 nm in diameter). Protofibrils then accumulate linearly to form protofilaments, and ultimately fibrils (reviewed in [81] and references therein). Fibrils are approximately 10 nm in diameter and may extend up to 1 μm in length.

“Off-pathway” structures are those that do not ultimately give rise to fibrils, and compete for the

formation of “on-pathway” aggregates. These “off-pathway” structures are generally spheroidal [81] and lack large amounts of β -sheet content [82] that are characteristic of “on-pathway” aggregates.

The most toxic $A\beta$ species are soluble oligomers, as they cause rapid neuronal death *in vitro* [83] and lead to increased membrane conductance [84] and rapid calcium influx that triggers apoptotic pathways [85]. Monomeric $A\beta$ has relatively low toxicity [86] and fibrillated $A\beta$ is several orders of magnitude less toxic than oligomers [83]. Although fibrils have low neurotoxicity, they have the ability to impair synaptic function and may cause neuritic dystrophy [87]. Thus, the prevailing view of the amyloid aggregation cascade and its relationship to the neurotoxicity observed in Alzheimer’s disease is that the intermediates along the aggregation pathway are the most toxic form of $A\beta$ to neurons. Detailed structural characterization of these intermediates remains challenging, and a more complete description of their conformational features is required to explain the mechanism(s) of toxicity. The application of molecular simulation methods to these species can greatly augment the information obtained through *in vitro* and *in vivo* experimentation, as simulations can provide atomistic insight to explain structural details that are unavailable through most experimental techniques.

1.2.6 Interactions of $A\beta$ with Membranes

The plasma membrane has been identified as the principal target of $A\beta$ toxicity in Alzheimer’s disease (as well as the target of other amyloidogenic proteins in different disorders) [88]. As such, numerous experiments have been conducted in cellular and cell-free systems (*e.g.*, model membranes) to determine the mechanism of $A\beta$ toxicity. It has been established that the principal toxic species in the progression of Alzheimer’s disease are soluble, oligomeric assemblies of $A\beta$ [83, 84, 88, 89]. These oligomers can associate with cell membranes and lead to structural deformations [90, 91, 92] and ion leakage [85]. Soluble monomeric $A\beta$ species can also incorporate into lipid bilayers, disrupting the cohesive interactions between lipids [90]. This disruption leads to local thinning of the bilayer (reduced P-P distance) [92] and decreased attraction between the lipids. Thus far, no mechanism has been determined experimentally to explain how $A\beta$ causes these phenomena, illustrating the importance of using atomistic simulations to examine the fundamental interactions between $A\beta$ and the surrounding lipid matrix.

Accumulation of $A\beta$ also diminishes synaptic plasticity, the level of interconnectedness between neurons. When plasticity is severely impaired, so too are the signals sent between neurons. $A\beta$ fibrils and plaques are believed to be responsible for structural damage to neuronal axons, and thus diminished cell-cell communication [87]. These fibrils, while neurotoxic, are not the principal toxic species in Alzheimer’s disease [88, 93]. It has been shown that fibril assembly is facilitated by the presence of a lipid bilayer [94]. Since fibrils may serve as a dynamic reservoir of oligomers [95, 96] and may promote oxidative stress through binding of transition metals [62], their clearance from the brain may ameliorate some of the toxicity associated with Alzheimer’s disease. Because these fibrils are largely resistant to endogenous proteases [97], targeting them using small molecules becomes an attractive approach.

1.2.6.1 A β Ion Channels

Since the discovery that A β oligomers could enhance the leakage of ions across the plasma membrane [85], an explanation has been sought for this behavior. Two possibilities have emerged. The first possibility is that A β aggregates form discrete ion channels that promote the passage of cations, specifically Ca²⁺, into the cell. The second theory is that oligomeric A β species perturb the lipid bilayer in such a way that the ability of the membrane to serve as a conductance barrier is compromised. At this point in time, it is not clear whether or not A β ion channels exist, though it remains the prevailing theory. As such, it is important to review both arguments here in the context of understanding the larger role that A β plays on perturbing the cell membrane and promoting neurotoxicity.

The first account of cation-specific A β channels was reported by Arispe *et al.* [98], who found that incorporation of A β ₄₀ into liposomes induced discrete changes in membrane potential, which they concluded were a result of conductivity of cations. They concluded that a number of monovalent cations (Li⁺, Na⁺, K⁺, and Cs⁺) could permeate the A β ion channel, as well as Ca²⁺. Investigations into the nature of these ion channels continued over many years, and recently Quist *et al.* demonstrated the ability of several amyloidogenic proteins to form ion channels, suggesting a common mechanism of action among many diseases [99]. Many structural models exist to explain the nature of A β ion channels, but high-resolution structural data are elusive. The most complete picture to date of putative A β ion channels comes from atomic force microscopy (AFM) imaging and resulting computer-generated models [100, 101, 102].

During the investigation of A β ion channels, several studies emerged suggesting that discrete pore formation was not required for membrane conductance to occur. These studies argued instead that the presence of oligomeric forms of A β sufficiently destabilized the membrane such that ions were more likely to leak across. The membrane was thus still intact, but far more susceptible to ion permeation. This mechanism of toxicity was first proposed by Kaye *et al.* [84], who found linear changes in membrane conductance, rather than discrete, stepwise conductance changes that would be characteristic of an ion channel. The mechanism proposed by Kaye *et al.* was that soluble A β oligomers incorporated into their model membranes, causing slow leakage of ions. The effect was specific to oligomerized A β , not fibrillated A β , as an oligomer-specific antibody ameliorated this effect. Subsequent work [103, 104] suggested that conformational changes in A β upon partitioning from the aqueous phase into the membrane were responsible for weakening the membrane structure and, as a consequence, lowering the dielectric barrier and thus leading to ion conductivity.

A strong argument against A β oligomer-induced ion leakage has been made [105]. Capone *et al.* found that they could reproduce the findings of both Kaye *et al.* and Sokolov *et al.* simply by using hexafluoroisopropanol (HFIP), the organic solvent used in previous studies to disaggregate A β . The findings by Capone *et al.* suggest that in the absence of a discrete A β ion channel, the ion leakage was simply an artifact of residual HFIP used in preparing the A β for experimentation rather than an actual effect of the peptide itself. It is unclear, however, how the mere presence of

HFIP would explain the ability of an anti-A β antibody to ameliorate the ion permeation effects, as observed by Kayed *et al.* [84]. Thus, the controversy over the exact nature of membrane-destabilizing A β species that promote ion leakage remains, though it is clear through all of these studies that A β causes ion leakage that is ultimately detrimental to the health of neurons.

1.2.7 The Role of A β in Oxidative Stress

The persistence of oxidative stress and inflammation has been linked to a number of disease states of the body, such as cancer, diabetes, and obesity. Neurodegeneration is no exception, as in Alzheimer's disease, inflammation and oxidative damage persist in the brain. The A β peptide is intimately linked to this phenomenon. It contains several residues that can bind metal ions and participate in redox chemistry, including Asp1, His6, Tyr10, His13, His14, and Met35, and it can disrupt other cellular processes through its association with proteins and membranes. One of the most interesting observations that has been made in individuals afflicted with Alzheimer's disease is that A β plaques isolated from the brains of Alzheimer's patients contain high levels of copper ($\sim 25 \mu\text{g/g}$ plaque), zinc ($\sim 69 \mu\text{g/g}$), and iron ($\sim 53 \mu\text{g/g}$) [106]. Thus it is clear that A β is capable of binding these pro-oxidant transition metals, though it is unclear at this time whether these A β -metal complexes are causative of Alzheimer's disease or a consequence of other events.

1.2.7.1 Oxidative Stress and the Role of Transition Metals

1.2.7.1.1 Copper Numerous structural studies have concluded that A β is capable of binding Cu²⁺ via coordination with His6, His13, His14, and another residue, which may be Tyr10 [107], a carboxylate from Asp1 or Glu11 [108], or the carbonyl oxygen of Ala2 [109]. The precise role of Cu in the toxicity observed in Alzheimer's is not clear. Both pro-oxidant [62, 110, 111] and antioxidant [112, 113] models have been proposed for the A β -Cu complex, and a dual role has also been suggested [114].

In the pro-oxidant model of the A β -Cu complex, it is proposed that complex formation will ultimately result in the reduction of Cu²⁺ to Cu⁺ [110], generating a peptide radical (A β \cdot). The A β \cdot species can react with other cellular components (lipids, other proteins, *etc.*) to regenerate A β , at the expense of damaging the other molecule. The A β -Cu⁺ complex generates hydroxyl radical ($\cdot\text{OH}$) in the presence of H₂O₂, re-oxidizing Cu⁺ to Cu²⁺. This reaction process parallels Fenton chemistry.

The antioxidant model of the A β -Cu complex suggests that A β acts as a copper chelator [112, 113]. Since copper is very active in catalyzing redox reactions, this mechanism may serve to sequester free metal from solution. The persistence of oxidative stress in the disease state may be attributed to the progressive increase in A β concentration as it aggregates, since higher-order aggregates are more effective at producing toxic reactive oxygen species [114]. It has further been demonstrated that, at low concentrations of A β (nM - μM) in cell-free systems, A β -Cu complexes

are potent antioxidants [115, 116, 117]. The relevance of these findings to *in vivo* conditions remains unknown.

Copper, though toxic in high quantities, is essential to normal neuronal function, especially in regulating neuronal excitation. Upon excitation, presynaptic neurons (those that send signals) are capable of releasing high levels of copper (up to 250 μM) into the synaptic cleft, the region separating the presynaptic and postsynaptic neurons (those that receive signals) [110, 118]. It may be through these normal cellular occurrences that copper accumulates and associates with $\text{A}\beta$ to initiate reactive oxygen species (ROS) production. Copper that is released by the cell to modulate these signals may be sequestered outside the cell by $\text{A}\beta$, thus depleting the intracellular pool and inhibiting the activity of such antioxidant enzymes as Cu/Zn superoxide dismutase (as reviewed in [119]).

1.2.7.1.2 Zinc The role of Zn^{2+} with respect to $\text{A}\beta$ is complex. Low levels of zinc ($< 50 \mu\text{M}$) are generally regarded as neuroprotective, while higher levels (in the high μM - low mM range) are neurotoxic [120, 121]. Huang *et al.* demonstrated that 25 μM zinc preserves the α -helical conformation of $\text{A}\beta$ (thus inhibiting β -strand formation associated with structures that proceed towards insoluble fibrils), though these helical species are prone to reversible aggregation [122]. Higher levels of zinc (50 μM) produce disordered, non-fibrillar, $\text{A}\beta$ aggregates [123].

Zn^{2+} is known to bind to $\text{A}\beta$ at the same binding site as Cu^{2+} , thus competing with copper for these coordinating interactions. Since zinc is in higher abundance relative to copper in the brain, it is believed to be neuroprotective by inhibiting copper binding and thus the generation of free radicals [124]. Oxidative stress can lead to the release of large amounts of vesicle-bound zinc from within the cell [125, 126]. Upon this release, zinc ceases to be neuroprotective and instead contributes to the toxicity observed in Alzheimer's disease.

1.2.7.1.3 Iron The presence of increased iron in the brain of Alzheimer's patients also has important implications for the progression of amyloidosis. Total iron levels are elevated in aging brains [127, 128].

Some studies have suggested that $\text{A}\beta$ can bind iron, leading to the production of H_2O_2 , similar to the phenomena observed in the presence of copper [62]. Other studies have concluded that the affinity of $\text{A}\beta$ for iron is very low, and that the toxic effect of iron is that it slows the formation of $\text{A}\beta$ fibrils, allowing soluble, disordered species (which are far more toxic) to persist [129]. A high-affinity binding site within $\text{A}\beta$ for iron has yet to be definitively identified, and thus the conclusions reached by Liu *et al.* remain the most detailed analysis of $\text{A}\beta$ -iron interactions. Their work has concluded that, like zinc, iron plays multiple roles in the progression of Alzheimer's disease. At low levels, iron (specifically Fe^{2+}) participates in Fenton reactions to generate $\cdot\text{OH}$ radicals from H_2O_2 . At higher levels, iron interacts with $\text{A}\beta$, albeit weakly and through a mechanism that is not completely understood, to alter the structure of the $\text{A}\beta$ aggregates, inhibiting ordered fibril formation to generate more toxic, soluble oligomeric species.

1.2.7.2 Toxicity Associated with the A β -Binding Alcohol Dehydrogenase

A hypometabolic state is associated with Alzheimer's disease, and the degree of severity of hypometabolism is correlated with the severity of cognitive decline [130]. It has been shown that three enzymes involved in the Krebs cycle have significantly decreased activity - pyruvate dehydrogenase (PDH), isocitrate dehydrogenase (ICDH), and α -ketoglutarate dehydrogenase (KGDH) [131]. The commonality between these three enzymes is that they all catalyze decarboxylation reactions (the loss of CO₂), though a connection between these deficiencies and the progression of Alzheimer's has not been established.

The site of the majority of the energy production in the cell is the mitochondrion, an organelle with two membranes. The inner membrane is the site of oxidative phosphorylation and the electron transport chain. Experiments have demonstrated that APP can associate with the inner mitochondrial membrane and become stuck there by virtue of its amino acid composition, which contains an N-terminal targeting sequence to the mitochondrion [132]. Accumulation of APP in the inner mitochondrial membrane disrupts the function of the inner mitochondrial membrane proteins responsible for electron transport and ATP synthesis. The afflicted cells thus become energy-deficient.

Two alloforms of A β (A β ₄₀ and A β ₄₂) can also be deposited in the mitochondria. The accumulation of these peptides decreases proper function of electron-transport complexes III and IV, resulting in increased production of H₂O₂ [131]. These data indicate that A β can interact with components of the mitochondria and prevent the proper usage of molecular oxygen, diverting it from a pathway that normally produces water to one that generates H₂O₂.

More recently, an enzyme known as A β -binding alcohol dehydrogenase (ABAD) has been identified as having a central role in the depression of normal metabolic activity in individuals with Alzheimer's disease. Alcohol dehydrogenases are enzymes that oxidize alcohol functional groups to carbonyls, concomitant with the reduction of nicotinamide adenine dinucleotide (NAD⁺) to its reduced form, NADH. ABAD was discovered to bind A β , inhibiting its normal enzymatic activity. As a consequence, a metabolic intermediate known as 2-methyl-3-hydroxybutyryl-CoA builds up within the mitochondria. This compound is hydrolyzed to 2-methyl-3-hydroxybutyrate, which inhibits the Krebs cycle and the electron transport chain [130].

The crystal structure of ABAD in the presence of A β has been determined [133]. Although the A β -binding loop is missing from the structure, other important structural features of the enzyme are identifiable. There is a specific binding loop in ABAD, called L_D, that mediates A β binding [133]. It is a unique binding region, as the amino acid sequence represents an insertion relative to other homologous enzymes. The N-terminal region of A β is believed to be responsible for interacting at this site [133], leaving the C-terminal hydrophobic residues free to oligomerize and perhaps form heterogeneous aggregates within the mitochondria.

1.2.7.3 Oxidative Damage of A β and Other Cellular Components

The oxidative stress resulting from A β aggregation and neurodegeneration damages cellular macromolecules through mechanisms that are common to many inflammatory states. A β aggregation has been correlated with increased lipid peroxidation, protein oxidation and nitration, and oxidative DNA damage [134].

The mechanism of A β -induced lipid peroxidation is quite straightforward and is particularly relevant in light of the association of A β with membrane-mediated toxicity (Section 1.2.6). Met35 in the A β sequence is susceptible to oxidation in the presence of ROS. The abstraction of a single electron from the thioether moiety of methionine gives a sulfuranyl radical, which can then abstract a hydrogen atom from a nearby lipid. The lipid radical can then react rapidly with triplet oxygen to form a lipid peroxide (Figure 1.2) [134].

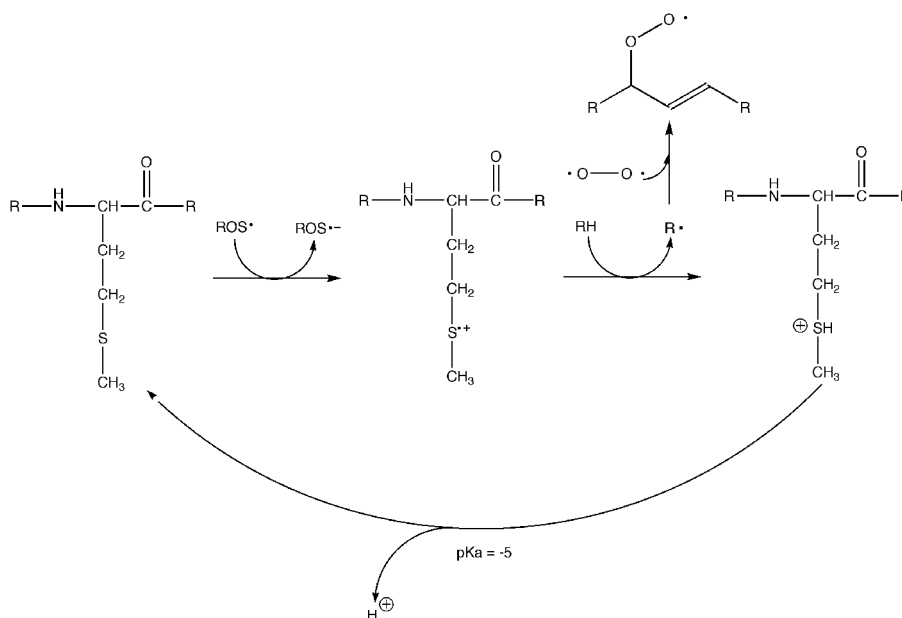


Figure 1.2: The A β sulfuranyl radical is formed by oxidation of the thioether moiety of Met35. Abstraction of a hydrogen atom from a nearby lipid molecule (RH) forms a strongly acidic species that loses a proton to restore the original methionine moiety. The lipid radical (R•) can undergo a number of fates, such as abstracting a hydrogen atom from a nearby lipid, removing an electron from Met35 of A β , or reacting with triplet oxygen to form a peroxide.

In addition to participating in sulfuranyl radical-mediated reactions, the Met35 residue can also become oxidized to methionine sulfoxide, commonly denoted Met35(O). A β found in senile plaques is frequently found in this oxidized state [135, 136, 137, 138], though it is not yet clear whether oxidation occurs before aggregation or as a consequence of the oxidizing environment that exists in the vicinity of A β plaques (discussed in Section 1.2.7.1). Watson *et al.* demonstrated that formation of Met35(O) destabilized the formation of β strand structures in the C-terminal region of

A β ₄₀ [139]. This region instead favored random coil configurations, thus implicating Met35(O) formation in inhibiting fibril-forming intermediates, which are rich in β -strand structures in the C-terminal region. The authors attributed this effect to entropic factors or the ability of Met35(O) to form hydrogen bonds with water, both of which would act to destabilize α -helical and β -strand structures. In contrast, the presence of Met35(O) in A β ₄₂ does not impede fibril formation [140], suggesting that the added hydrophobicity of Ile41 and Ala42 in the A β sequence can overcome the effects of oxidation of Met35. In contrast with the results of Watson *et al.*, a study conducted by Snyder *et al.* concluded that Met35(O) in A β ₄₀ accelerated aggregation [141], though differences in solution conditions (pH, ionic strength, *etc.*) could explain the different conclusions reached in these studies. Further experimentation and simulations are necessary to resolve such discrepancies.

Another possible outcome of increased ROS in the brain is damage to the tyrosine residue that resides at position 10 within the A β sequence. This residue is susceptible to covalent cross-linking [134], which strengthens intermolecular A β interactions by keeping individual peptide molecules in close proximity in aggregates. Physical damage inflicted by these aggregates increases ROS levels in the brain through the mechanisms discussed above, causing the oxidative state to persist.

1.2.8 Current Therapeutics for Alzheimer's Disease

1.2.8.1 Acetylcholinesterase Inhibitors

Existing therapies for Alzheimer's disease primarily utilize acetylcholinesterase (AChE) inhibitors. Since A β leads to the death of cholinergic neurons, one of the neurotransmitters that becomes depleted is acetylcholine (ACh). ACh is necessary for proper neural function, especially in areas of the brain related to memory and learning [142].

Evidence suggests that donepezil (Aricept[®], Pfizer Incorporated) performs well in improving cognitive ability in individuals with Alzheimer's disease [143], but support remains modest, and evidence is strongest only for mild to moderate forms of Alzheimer's disease [144]. Newer evidence suggests that donepezil may be effective in treating severe cases of Alzheimer's [145], but these data are limited. Aricept[®] is the most commonly prescribed pharmaceutical for treating Alzheimer's disease, according to Pfizer's statistics [146].

Other compounds, such as rivastigmine (Exelon[®], Novartis International AG) and galantamine (Reminyl[®], Janssen Pharmaceutica), also work by competitively inhibiting AChE. Combination therapies using these AChE inhibitors may include memantine, an antagonist of N-methyl D-aspartate (NMDA) glutamatergic receptors. Memantine functions in preventing Ca²⁺ influx into the cell [147], an event that has been linked to the deleterious actions of A β [85]. All three of these drugs (donepezil, rivastigmine, and galantamine) are associated with gastrointestinal side effects. Since these drugs are only moderately effective in treating downstream symptoms of Alzheimer's caused by A β toxicity, more effective therapeutics are necessary to treat the disease.

1.2.8.2 Inhibitors of β - and γ -Secretase

Since the symptoms of Alzheimer's disease manifest themselves as a result of the interactions between $A\beta$ and neurons, more recent efforts have focused on preventing the production of $A\beta$ from its precursor, APP. Inhibitors have been designed for β - and γ -secretases, but progress thus far has been slow and results are mixed.

β -Secretase (BACE-1) belongs to the general class of aspartyl proteases, which catalyze peptide bond cleavage through two active site aspartate residues. BACE-1 expression is highest in the brain, specifically neurons, as determined by mRNA [23, 24, 148] and enzyme activity [24] levels. APP is not the only substrate for BACE-1; others include sialyltransferase STGal6 I [149], P-selectin glycoprotein ligand-1 (PSGL-1) [150], and APP-like proteins [151]. Since β -secretase cleavage is required for the production of $A\beta$, it was quickly identified as a potential drug target. The question of whether or not such inhibition is safe remains to be answered. Cai *et al.* determined that *bace1*^{-/-} mice had diminished $A\beta$ levels and could survive normally [152], but Dominguez *et al.* reported partial lethality and impaired development in *bace1*^{-/-} mice [153].

γ -Secretase activity is attributed to a complex of proteins, which combined has the same activity as an aspartyl protease. The active site resides in the hydrophobic core of the plasma membrane, making structural and mechanistic studies difficult. Targeting γ -secretase appears to be a less viable option than targeting BACE-1, since the side effects of inhibiting γ -secretase are comparable to the symptoms of Alzheimer's disease. Deficiencies in presenilins 1 and 2 (PS1 and PS2), believed to be the catalytic components of the γ -secretase complex, result in deficits in learning, memory, and synaptic function, as well as neuronal death, in mice [154].

Questions remain about the viability of targeting BACE-1 and γ -secretase. If these proteases are inhibited, will side effects stemming from incomplete or inappropriate proteolysis of its other targets occur? An additional issue relates to the nature of both BACE-1 and γ -secretase. How does one design an inhibitor for BACE-1 and γ -secretase within the brain without affecting other aspartyl proteases in the target tissue and elsewhere? Furthermore, how does one deliver these synthetic molecules across the blood-brain barrier? Some progress has been made on these fronts, as reviewed in [155], but these challenges remain largely unsolved.

1.2.8.3 Targeting $A\beta$ Aggregation

The last main target for therapeutic molecules is the $A\beta$ peptide itself. Since it is $A\beta$ that leads most directly to the symptoms of neurodegeneration seen in Alzheimer's disease, it may be possible to block or reverse its aggregation and association with neurons. Since endogenous proteases can clear soluble $A\beta$ (discussed in Section 1.2.4.5), dissolution of aggregates may yield soluble entities that can be acted upon by these enzymes. Existing evidence suggests that such phenomena are possible, and work is being actively conducted to target $A\beta$ for degradation after it is produced. Ongoing research has examined synthetic molecules, antibodies, and dietary compounds for their potential roles in diminishing the $A\beta$ burden in the Alzheimer's brain.

In one of the first studies to use synthetic molecules to target $A\beta$, Gestewicki *et al.* [156] reported the use of a “Trojan horse” mechanism in designing bifunctional molecules that target $A\beta$. Two domains were present on each molecule, one that bound $A\beta$, and another that bound a chaperone that would target $A\beta$ for degradation. The results of that study indicated that it was indeed possible to use pathogenic aggregation as a target to derive novel therapeutics in treating Alzheimer’s disease. The development of tramiprosate (AlzhemedTM, Neurochem, Inc.) is based on similar principles. Alzhemed purportedly binds to soluble $A\beta$ and prevents its aggregation. The compound was demonstrated to be well-tolerated and effective in decreasing amyloid plaque levels in the brain [157]. However, more recent evidence suggests that the use of Alzhemed may exacerbate the intracellular aggregation of tau protein, which comprises the other major histopathological marker of Alzheimer’s disease [158]. The inhibition of one pathological marker, at the expense of an increase in another, suggests there may still be unresolved problems with Alzhemed, which has yet to receive FDA approval.

Experiments have shown that dietary antioxidants (flavonoids) can destabilize $A\beta$ aggregates, and even inhibit their formation [159]. These findings suggest a physical association between the flavonoids and $A\beta$, and that dietary intervention may be a useful component of the treatment for Alzheimer’s disease. Efficacy of these compounds in attenuating $A\beta$ toxicity has been demonstrated *in vitro* in cell culture [160] and *in vivo* using a transgenic mouse models of Alzheimer’s disease [161, 162] at the micromolar level.

The use of natural products such as flavonoid antioxidants is attractive for several reasons. First, targeting $A\beta$ aggregation has already been shown to be feasible and effective in decreasing cytotoxicity [156, 157]. Second, flavonoids are well-tolerated and non-toxic at the levels required to manifest an effect [163]. Third, these compounds readily cross the blood-brain barrier at therapeutically-relevant concentrations after oral administration [162], a significant challenge in designing brain-targeted therapeutics. Last, these compounds may also be able to ameliorate the oxidative stress that persists in the brain as $A\beta$ accumulates [163], as was discussed in Section 1.2.7.

1.3 Organization of the Dissertation

This dissertation will take the following form. Chapter 2 will discuss the basics of molecular dynamics (MD) simulations, and Chapter 3 will outline the justification for the force field chosen for the work conducted. The remaining chapters will describe all the experiments that were performed, starting with $A\beta$ -membrane systems (Chapters 4, 5, and 6), and moving to studies of $A\beta$ aggregates in solution, the first of which is a thermodynamic analysis of $A\beta$ protofibril stability (Chapter 7), followed by the interactions of a flavonoid (morin) with $A\beta$ monomers, dimers, and higher-order aggregates (Chapters 8 and 9). The development of the GridMAT-MD program (used for lipid analysis throughout this dissertation) is described in Chapter 10. Chapter 11 is dedicated to understanding the proper derivation of small molecule topologies for use in MD simulations,

information that pertains to several of the earlier Chapters. Finally, the overarching conclusions of this body of work are presented in Chapter 12.

1.4 Attribution

The work contained herein is my own, except for the specific elements described here, which were produced in collaborative efforts on projects that were co-authored. Copyrights for published articles are given in each Chapter to which they pertain. The GridMAT-MD program (described in Chapter 10) was written by William J. Allen (Ph.D., Virginia Tech Department of Biochemistry 2011). In Chapter 11, simulations of the UGM-FAD complex, simulations of liquid/vapor forms of *p*-cresol, hexane, and Spartan '04 calculations were performed by William J. Allen.

Chapter 2

Molecular Dynamics Simulations

2.1 Newtonian Mechanics

The underlying theory of molecular dynamics (MD) simulations is that classical Newtonian mechanics can be applied to molecular systems to obtain an accurate representation of the dynamics. That is, quantum effects, most notably bond breakage and formation, are not considered. Electrons are represented only implicitly through the invocation of the Born-Oppenheimer Approximation [164], which states that electrons move so rapidly during the period of nuclear motion that, over sufficiently long timeframes, their effects are considered uniform and there is no deformation of the electron cloud.

In the classical MD model, all atoms are represented as hard spheres connected by harmonic springs that represent bonds. These hard spheres represent the positions of the atomic nuclei. All electronic charges are centered on the position of this nucleus. That is, “point charges” are used to represent all atoms that bear charge (partial or full) due to ionization state or dipole.

The integration of Newton’s Laws gives rise to trajectories, linking atomic motions through time. Application of force to each atom causes an acceleration, dependent upon the atomic mass. This relationship was expressed in Newton’s Second Law:

$$\mathbf{F} = m\mathbf{a} \tag{2.1}$$

The motion of any particle under an applied force can be described by several fundamental relationships:

$$\frac{d^2\mathbf{r}_i}{dt^2} = \frac{\mathbf{F}_i}{m_i} \quad (2.2)$$

$$\frac{d\mathbf{r}_i}{dt} = \mathbf{v}_i \quad (2.3)$$

$$\frac{d\mathbf{v}_i}{dt} = \frac{\mathbf{F}_i}{m_i} \quad (2.4)$$

Thus it can be seen that current positions (\mathbf{r}), velocities (\mathbf{v}), and accelerations (\mathbf{a}) can be used to calculate the change in position applied to each particle in the system through integration. By calculating forces and integrating the resulting accelerations, velocities can be calculated (Equation 2.4), and likewise the integration of these velocities gives the change in position (Equation 2.3). These equations must be solved numerically, and thus a method for conducting these calculations is needed. The classical method for integrating Newtonian equations is the Verlet algorithm [165], given by the following expressions:

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \mathbf{v}(t)\delta t + \frac{1}{2}\mathbf{a}(t)\delta t^2 \quad (2.5)$$

$$\mathbf{r}(t - \delta t) = \mathbf{r}(t) - \mathbf{v}(t)\delta t + \frac{1}{2}\mathbf{a}(t)\delta t^2$$

The summation of the above two equations, implying time-reversibility of the algorithm, yields:

$$\mathbf{r}(t + \delta t) = 2\mathbf{r}(t) - \mathbf{r}(t - \delta t) + \mathbf{a}(t)\delta t^2 \quad (2.6)$$

The Verlet algorithm is of modest precision, and it is preferable to include velocities in the calculation of positions. The Velocity Verlet algorithm [166] was produced to solve this issue, and is widely used. It is often preferred over the original Verlet algorithm as it is mathematically equivalent and makes better use of a computer's finite precision. The equations used for the Velocity Verlet algorithm are as follows:

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \mathbf{v}(t)\delta t + \frac{1}{2}\mathbf{a}(t)\delta t^2 \quad (2.7)$$

$$\mathbf{v}(t + \delta t) = \mathbf{v}(t) + \frac{1}{2}[\mathbf{a}(t) + \mathbf{a}(t + \delta t)]\delta t \quad (2.8)$$

The other common integration method is the so-called “leap-frog” integrator. The name derives from the fact that positions and velocities are never calculated at the same time, rather velocities are calculated every $\frac{1}{2}\delta t$, while positions are calculated at integer multiples of δt :

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \mathbf{v}\left(t + \frac{1}{2}\delta t\right) \delta t \quad (2.9)$$

$$\mathbf{v}\left(t + \frac{1}{2}\delta t\right) = \mathbf{v}\left(t - \frac{1}{2}\delta t\right) + \mathbf{a}(t)\delta t \quad (2.10)$$

Velocities at time t can be approximated by the following expression:

$$\mathbf{v}(t) = \frac{1}{2} \left[\mathbf{v}\left(t - \frac{1}{2}\delta t\right) + \mathbf{v}\left(t + \frac{1}{2}\delta t\right) \right] \quad (2.11)$$

The leap-frog algorithm is very robust and has been shown to be sufficiently accurate. Thus it is commonly used in place of more computationally-expensive Verlet-type algorithms.

In all of the integration equations above, the term δt is present. This quantity is called the “time step” in an MD simulation. It reflects the amount of time allowed to pass before the potential energy equation is re-evaluated. Its magnitude is limited by the fastest vibrations in the simulated system, typically bonds involving hydrogen atoms. The potential energy equation is explored in Section 2.2.

2.2 Force Fields

MD simulations are dependent upon so-called “force fields,” which have two components: a potential energy function (also called a “functional form”) and a parameter set, describing atom types and bonded parameters. A typical functional form resembles the following:

$$V(r) = \frac{1}{2} \sum_{\text{bonds}} k_b(b - b_0)^2 + \frac{1}{2} \sum_{\text{angles}} k_\theta(\theta - \theta_0)^2 + \frac{1}{2} \sum_{\substack{\text{dihedral} \\ \text{angles}}} k_\phi[1 + \cos(n\phi - \delta)] + \sum_{\substack{\text{nonbonded} \\ \text{pairs}}} \left[\frac{A}{r^{12}} - \frac{C}{r^6} + \frac{q_1 q_2}{Dr} \right] \quad (2.12)$$

The force on an atom can be calculated by differentiating Equation 2.12 with respect to its position within the system, a vector quantity:

$$\mathbf{F}_i = -\frac{\partial V}{\partial \mathbf{r}_i} \quad (2.13)$$

These forces are imparted on all of the atoms in the system as part of the integration of the Newtonian equations and are what give rise to the motion of the atoms in the system, described in Section 2.1. Thus the functional form provides not only a means to directly describe the potential energy of the system, but a way to calculate the forces on the atoms to impart motion.

2.2.1 Types of Force Fields

There are two major types of MD force fields. The first type is called “atomistic,” as it describes the component molecules in terms of each of their individual atoms. The second type is called “coarse grain,” with molecules built not from discrete atoms, but from “coarse” particles that represent chemically-distinct functional groups.

Within the atomistic class of force fields, there are two sub-classes, all-atom and united-atom. All-atom force fields explicitly represent every atom of the system, including hydrogen. All-atom force fields are by far the most common, including popular parameter sets such as OPLS-AA [167, 168], AMBER [169, 170, 171], and CHARMM [172, 173]. The main drawback to all-atom force fields is that explicit representation of every single atom in the system requires a shorter time step (usually 1 fs) and can result in a prohibitive amount of atoms, leading to very slow calculations. With the rapid progress in both hardware and software, this limitation has become less of a problem.

United-atom force fields merge nonpolar hydrogen atoms into their connected carbon atoms, assigning a heavier mass to this atom, and allowing for a longer time step (usually 2 fs). For instance, a $-\text{CH}_3$ group would be represented by a single C atom with a mass of 15.035 amu and modified Lennard-Jones parameters to reflect the chemical nature of this group. Polar hydrogens are explicitly represented, as are aromatic hydrogens (in some parameter sets). The most commonly-used united-atom force fields are the GROMOS parameter sets [174].

Coarse-grain force fields further simplify the representation of molecules in a system. Instead of explicit atoms, coarse particles are used to represent the various functional groups present in the molecule(s). That is, several atoms are grouped together into a single coarse particle bearing the chemical characteristics of that functional group. Use of coarse grain particles reduces the level of detail available in a simulation, but allows substantially more data to be collected very quickly, as the time step available in these simulations is on the order of 20-40 fs. There are a number of coarse grain force fields, and the MARTINI parameter set [175, 176] is frequently used within GROMACS.

Each force field has a different functional form, as far as the description of the bonded and non-bonded parameters, but the general format of Equation 2.12 describes the salient features common to all MD functional forms.

2.3 Time Scales of MD Simulations

MD simulations seek to describe microscopic systems over short periods of time. While the folding of a complete protein may take upwards of microseconds to seconds, it is currently not feasible to generate reliable atomistic MD trajectories longer than several hundred nanoseconds.

The first MD simulation of a protein (bovine pancreatic trypsin inhibitor) was reported in 1977, containing approximately 500 atoms and lasting for just 9.2 ps [177]. With rapid advances in computer processors, disk storage, and software algorithms, it has become possible to simulate systems that are orders of magnitude larger (hundreds of thousands to millions of atoms) for far more extended timeframes. The systems described in the present work seek to address protein unfolding and aggregation over hundreds of nanoseconds for systems containing between approximately 20,000 and 100,000 atoms. It has recently been demonstrated that very long simulations (on the order of microseconds) may, in fact, result in oversampling of high-free energy states in proteins [178]. Thus, although protein unfolding and aggregation may occur over a longer period of time *in vitro* and *in vivo*, these timeframes remain inaccessible to atomistic MD simulations.

Chapter 3

Practical Considerations for Running Molecular Dynamics Simulations

3.1 Introduction

The force field used for any MD simulation is perhaps the most important consideration when preparing a system. Each force field has advantages and disadvantages, and is based upon different assumptions and derivation schemes. The force field chosen for the system of interest will also largely govern the remaining choices such as solvent model, electrostatics and Lennard-Jones calculation methods, time step, *etc.*

3.2 Comparison of Different Force Fields

To determine the most appropriate force field for the intended simulations, it was necessary to establish benchmarks for the reliability of the different, commonly-used force fields available as part of the GROMACS package [179, 180]. By default, GROMACS is capable of automatically generating topologies under for use with the OPLS-AA force field [167, 168] and several variants of the GROMOS force fields, including a modified version of GROMOS87 [181], as well as the 43A1 [182], 43A2 [183], 45A3 [184], 53A5, and 53A6 [174] GROMOS96 parameter sets. The CHARMM27 force field [173] has recently been implemented in GROMACS [185], as well, although only several years after the work presented here was initiated. The AMBER force fields have also been ported to GROMACS format; the available editions of the force field available include AMBER94 [169], AMBER99 [170], AMBER99 ϕ [186], and AMBER03 [171], among others.

3.2.1 Simulation Systems

To compare a variety of MD force fields, the $A\beta_{40}$ peptide in water was used as a model system, as it directly pertains to all of the simulations conducted in this body of work. The theory that underlies the GROMOS96 force fields is that the parameters are derived in order to accurately describe the interactions of amino acid side chains with polar and nonpolar media, thus describing protein folding and interactions of proteins with a variety of environments. Since the goals of this body of work include studying protein misfolding and aggregation in both aqueous and membrane environments, the GROMOS96 force fields were a logical choice. However, it was necessary to further justify this decision based on the observed dynamics of $A\beta$.

Different protein force fields are typically derived using a particular solvent model. As such, for different force fields, different water models were used, and the simple point charge (SPC) model was applied to all systems as a uniform standard and since it is commonly utilized in conjunction with GROMOS96 parameter sets, the desired force field. Treatment of electrostatics and van der Waals cutoff distances also differ between force fields, with some parameter sets being more sensitive to change than others. For this reason, some simulations were conducted with non-standard neighbor searching distances and nonbonded cutoffs.

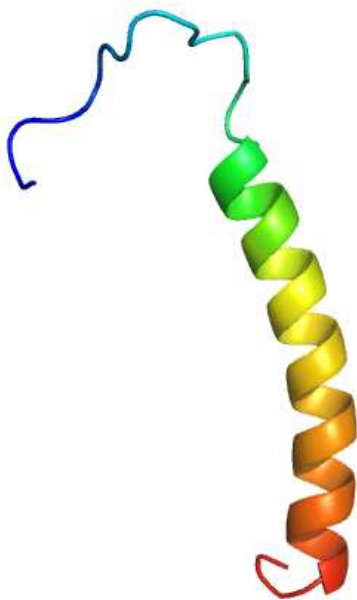


Figure 3.1: The structure of the $A\beta_{40}$ peptide, as determined by NMR in the presence of SDS micelles. The peptide is colored as a rainbow from blue (N-terminus) to red (C-terminus).

The model of the $A\beta_{40}$ peptide (Figure 3.1) was taken from PDB entry 1BA4 [77], determined by NMR spectroscopy in the presence of sodium dodecylsulfate (SDS) micelles. All simulated systems centered the protein in a rhombic dodecahedral box with a minimum solute-box distance of 1.0 nm to satisfy the minimum image convention. All titratable groups were assigned their typical proto-

nation states at pH 7.4 (for details see Chapter 4, Section 4.2.1), resulting in a net -3 charge on $A\beta$. Following the addition of solvent (water), Na^+ counterions were added such that each system had a net neutral charge. All systems were energy-minimized using the steepest descent method, after which they were equilibrated for 100 ps under an NVT ensemble, using weak coupling [187] to maintain temperature at 298 K, followed by 100 ps under an NPT ensemble, again applying weak coupling to the pressure of the system (1.0 bar). Production simulations were conducted under an NPT ensemble using the Nosé-Hoover thermostat [188, 189, 190] and Parrinello-Rahman barostat [191, 192] to maintain temperature and pressure, respectively, and to generate a rigorous NPT ensemble. All simulations utilized the smooth particle mesh Ewald (PME) algorithm [193, 194] for long-range electrostatics. Simulation parameters that were specific to each system are enumerated in Table 3.1.

Table 3.1: Simulation Systems

Force Field	Water Model	Neighbor searching cutoff (nm)	Electrostatics cutoff (nm)	van der Waals cutoff (nm)
GROMOS96 43A1	SPC	0.9	0.9	1.4
GROMOS96 53A6	SPC	0.9	0.9	1.4
OPLS-AA	TIP4P	0.9	0.9	1.4
OPLS-AA	SPC	0.9	0.9	1.4
AMBER03	TIP3P	0.8	0.8	0.8
AMBER03	SPC	0.8	0.8	0.8

3.2.2 Results of Force Field Comparison

Several criteria were used to analyze the structures of $A\beta$ during the simulations. These included secondary structure, radius of gyration, heavy atom contacts, intramolecular hydrogen bonds, root-mean-square deviation (RMSD), and root-mean-square fluctuation (RMSF).

3.2.2.1 Secondary Structure

Secondary structure content varied widely based on the force fields applied to $A\beta$ (Table 3.2). Circular dichroism (CD) spectroscopy studies of $A\beta$ in solution indicate that in water, $A\beta$ contains 1.1% α -helix, 60.4% β -strand, 31.4% random coil, and 7.1% β -turn [79]. None of the force fields were able to reproduce these experimental findings exactly. In general, application of the AMBER03 force field yielded high levels ($>36\%$) of α -helical content, although all the force fields tested here result in α -helical content that exceeds what is observed experimentally. The AMBER94 force field was shown to be “helix friendly,” over-stabilizing α -helical structures [195],

but subsequent revisions of the force field (AMBER99 and AMBER03) recalibrated the protein backbone terms to maintain the appropriate balance of secondary structure elements. It appears that the AMBER force fields still tend to predict greater levels of α -helical content than the other force fields. Final structures from independent 50-ns simulations under each force field are shown in Figures 3.2 - 3.7. In each figure, the A β peptide is shown in cartoon rendering, colored as a rainbow from blue (N-terminus) to red (C-terminus).

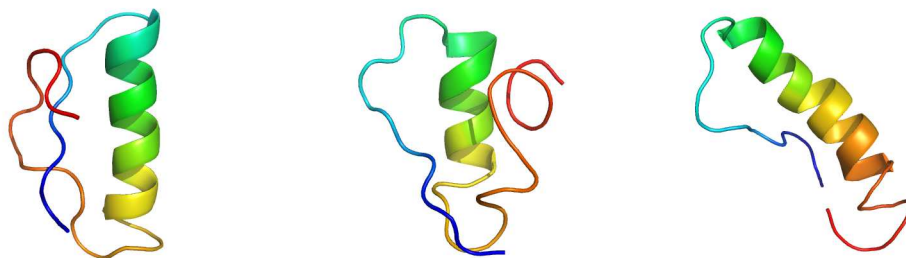


Figure 3.2: Structures of A β produced by the GROMOS96 43A1 force field in SPC water.

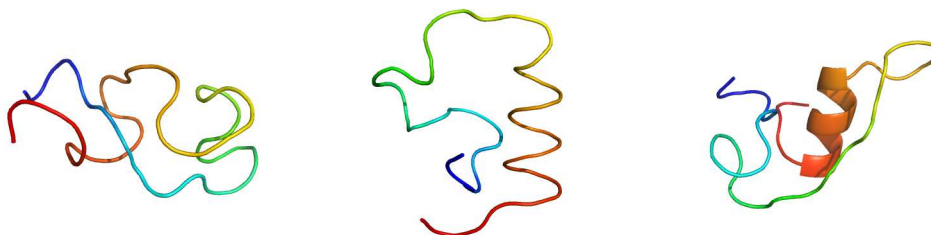


Figure 3.3: Structures of A β produced by the GROMOS96 53A6 force field in SPC water.

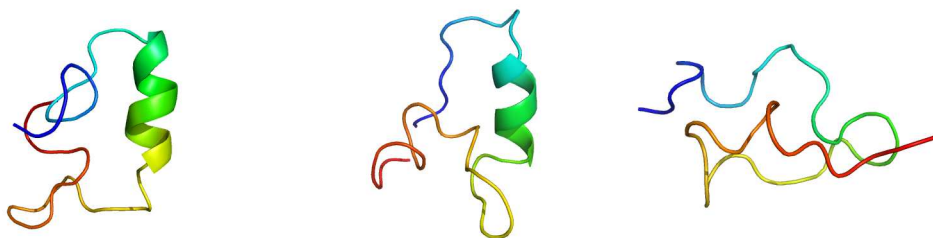


Figure 3.4: Structures of A β produced by the OPLS-AA force field in TIP4P water.

The β -strand content of A β in the simulations was also well below the experimentally-observed value. The GROMOS96 53A6 parameter set predicted the greatest level of β -strand content, although the wide deviations in the obtained values make it difficult to conclude that this particular parameter set is truly better than GROMOS96 43A1 or OPLS-AA (when combined with the SPC

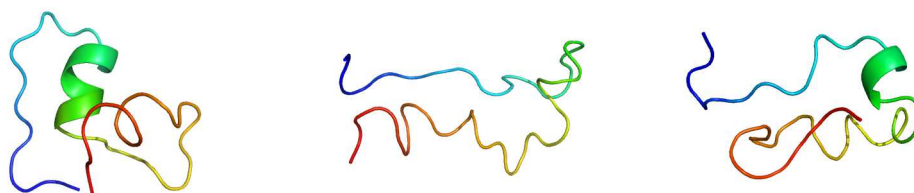


Figure 3.5: Structures of A β produced by the OPLS-AA force field in SPC water.

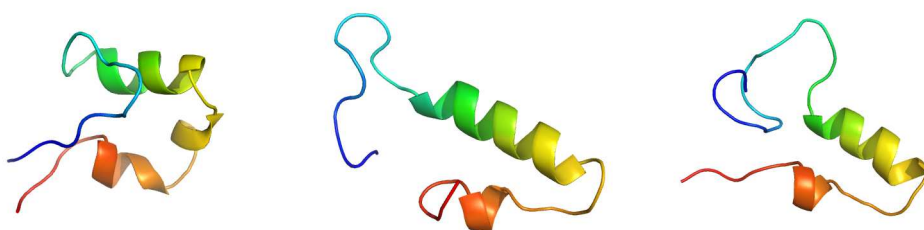


Figure 3.6: Structures of A β produced by the AMBER03 force field in TIP3P water.

water model, Table 3.2). There is some debate in the literature as to whether or not the GROMOS96 53A6 parameter set over-samples extended conformations such as β -strands. Matthes and de Groot [196] concluded that the application of GROMOS96 53A6 or OPLS-AA with PME generated an increased propensity for β -hairpin and extended structures. Conversely, Best *et al.* [197] determined that most force fields overestimate α -helical content with no bias towards β -structure content.

All force fields were able to reasonably predict the correct random coil content, given the standard deviations listed in Table 3.2. β -Turn content was consistently overestimated, although GROMOS96 43A1, GROMOS96 53A6, and AMBER03 + TIP3P produced reasonable results while the remainder of the force fields tested more drastically overestimated this secondary structure type.

Based on secondary structure, none of the force fields performed significantly better than any of the others. Some force fields outperformed the others in individual secondary structure categories, but none were consistently accurate. Since A β is known to be largely composed of β -strand and random coil elements, it appears that the GROMOS96 53A6 and OPLS-AA parameter sets (regardless of water model) perform comparably, based on the structural distributions observed (Figures 3.3 - 3.5). The structures produced by the AMBER03 force field sample a narrow distribution of structures and contain too much α -helical character, behavior that was also independent of the water model. Secondary structure evolution (both as a function of time and as probability per residue) for all systems described in this section are shown in Figures 3.8 - 3.13.

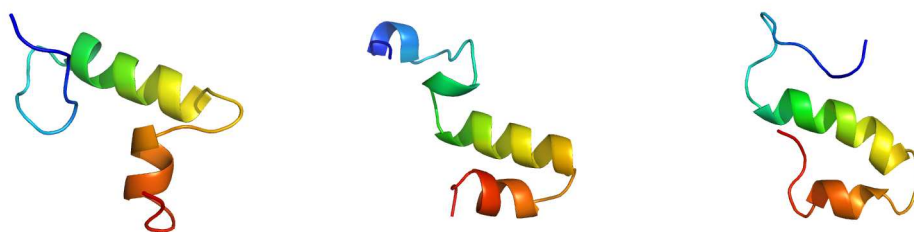
Figure 3.7: Structures of A β produced by the AMBER03 force field in SPC water.

Table 3.2: Secondary Structure Content

Structure	GROMOS96 43A1/SPC	GROMOS96 53A6/SPC	OPLS- AA/SPC	OPLS- AA/TIP4P	AMBER03/ TIP3P	AMBER03/ SPC
% α -helix	25.7 \pm 2.3	15.4 \pm 5.1	11.4 \pm 3.6	10.0 \pm 8.3	36.8 \pm 7.0	38.0 \pm 2.5
% π -helix	5.6 \pm 9.4	6.0 \pm 5.9	0.1 \pm 0.1	0.1 \pm 0.2	0.0 \pm 0.0	0.0 \pm 0.1
% 3_{10} -helix	2.0 \pm 2.7	0.1 \pm 0.1	6.6 \pm 3.3	4.0 \pm 2.9	4.7 \pm 2.7	3.9 \pm 1.5
% β -strand	3.0 \pm 3.1	5.0 \pm 6.9	1.0 \pm 0.6	0.0 \pm 0.0	0.0 \pm 0.0	0.0 \pm 0.0
% β -bend	17.1 \pm 4.3	15.7 \pm 5.1	20.4 \pm 5.0	29.2 \pm 4.1	16.3 \pm 4.2	12.7 \pm 3.1
% β -turn	11.4 \pm 2.5	15.2 \pm 7.7	22.3 \pm 4.5	19.8 \pm 6.5	13.3 \pm 1.3	19.9 \pm 6.8
% β -bridge	3.9 \pm 1.5	2.6 \pm 1.2	3.9 \pm 2.1	1.7 \pm 1.2	0.2 \pm 0.2	0.2 \pm 0.3
% Coil	31.3 \pm 4.7	40.0 \pm 9.2	34.3 \pm 2.8	35.2 \pm 3.7	28.6 \pm 1.9	25.3 \pm 4.7

3.2.2.2 Radius of Gyration

The radius of gyration, R_g , of a protein is defined as the root-mean-square distance of the N- and C-termini to its center of mass (Equation 3.1). This term can further be decomposed into terms that are aligned with the x -, y -, and z -axes. R_g represents the level of compaction of a protein structure. That is, smaller values of R_g indicate that the protein is compact, while larger values of R_g indicate extended conformations. Values of R_g obtained from the simulations under each of the force fields considered are listed below in Table 3.3. Standard deviations shown are from $n = 3$ simulations for each force field, and are calculated over the stable portion of the 50-ns trajectories (generally the last 25 ns).

$$R_g = \sqrt{\frac{\sum_i ||\mathbf{r}_i||^2 m_i}{\sum_i m_i}} \quad (3.1)$$

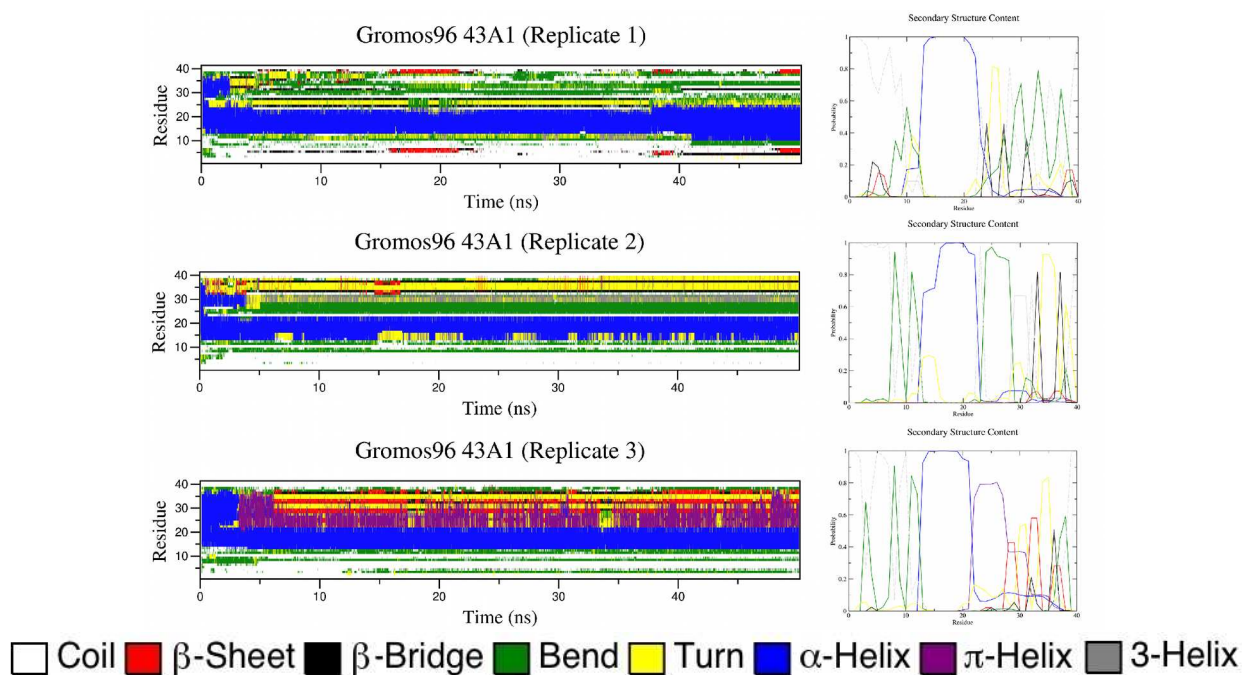


Figure 3.8: Left, secondary structure evolution of the three replicate systems simulated with the GROMOS96 43A1 parameter set. Right, probabilities of given secondary structural elements as a function of residue number.

Table 3.3: Radius of Gyration

Force Field	R_g (nm)
GROMOS96 43A1/SPC	0.97 ± 0.08
GROMOS96 53A6/SPC	1.05 ± 0.08
OPLS-AA/TIP4P	1.04 ± 0.04
OPLS-AA/SPC	1.03 ± 0.01
AMBER03/TIP3P	1.14 ± 0.10
AMBER03/SPC	1.16 ± 0.01

NMR studies conducted by Zhang *et al.* [80] have indicated that the $A\beta$ peptide adopts a collapsed coil structure in water, a behavior observed in all of the simulations conducted here. The peptide presents a large amount of hydrophobic surface area upon adopting this compact structure, perhaps the reason for its propensity to aggregate in solution. Based on the results of the R_g calculations conducted here, all of the force fields assessed predict similar configurations, with respect to the level of compaction in the structure. The AMBER force fields tended to predict the most elongated structures, but given the relatively wide standard deviation of the AMBER03 + TIP3P simulations, it is difficult to assess whether this difference is significant.

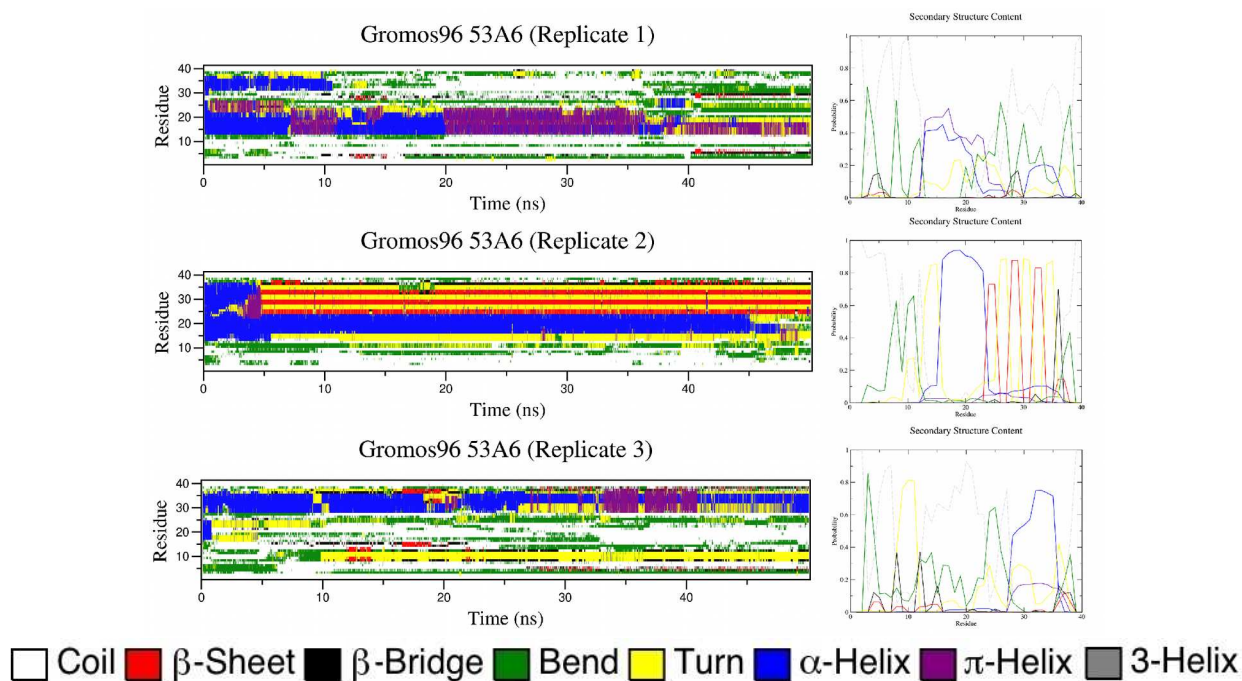


Figure 3.9: Left, secondary structure evolution of the three replicate systems simulated with the GROMOS96 53A6 parameter set. Right, probabilities of given secondary structural elements as a function of residue number.

3.2.2.3 Intramolecular Contacts and Hydrogen Bonds

The results presented in Section 3.2.2.2 indicate that under the influence of all force fields, the $A\beta$ peptide becomes more compact over time. Thus, it is expected that the number of intramolecular contacts should increase. The loss of α -helicity in favor of extended structures should also manifest itself in terms of changes in native hydrogen bonding. Listed in Table 3.4 are heavy atom contacts and hydrogen bonds present over the stable portion of the trajectory (again, generally the last 25 ns). Heavy atom contacts are used as the criterion for assessment rather than total contacts, as all-atom force fields (AMBER and OPLS-AA) will necessarily have more atomic contacts present when nonpolar hydrogen atoms are considered. As before, standard deviations shown are based on $n = 3$ simulations.

3.2.2.4 Structural Deviations

Two quantities that are frequently used to describe protein structures are root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF). RMSD of a system of N particles is calculated according to Equation 3.2:

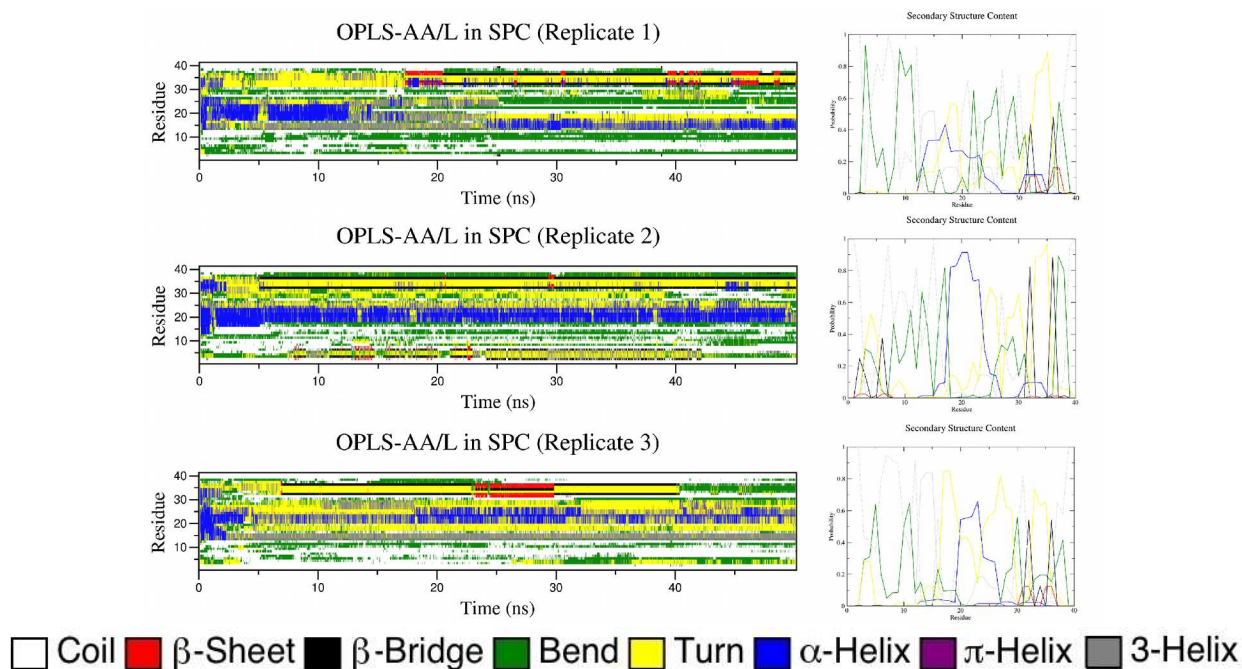


Figure 3.10: Left, secondary structure evolution of the three replicate systems simulated with the OPLS-AA parameter set in SPC water. Right, probabilities of given secondary structural elements as a function of residue number.

Table 3.4: Intramolecular Contacts and Hydrogen Bonds

Force Field	Number of contacts	Hydrogen bonds
GROMOS96 43A1/SPC	11044 ± 257	15.8 ± 3.0
GROMOS96 53A6/SPC	9588 ± 39	12.1 ± 5.2
OPLS-AA/TIP4P	9304 ± 106	10.8 ± 0.8
OPLS-AA/SPC	9399 ± 84	8.5 ± 1.3
AMBER03/TIP3P	9033 ± 180	13.4 ± 1.6
AMBER03/SPC	9414 ± 227	16.3 ± 2.0

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^{i=N} \delta_i^2} \quad (3.2)$$

In Equation 3.2, δ_i is the distance of particle i from its reference structure. Most algorithms for calculating RMSD include a least-squares fitting procedure to minimize RMSD. GROMACS allows for any subset of the atoms in the system to be the fitting group, and any other group to have the

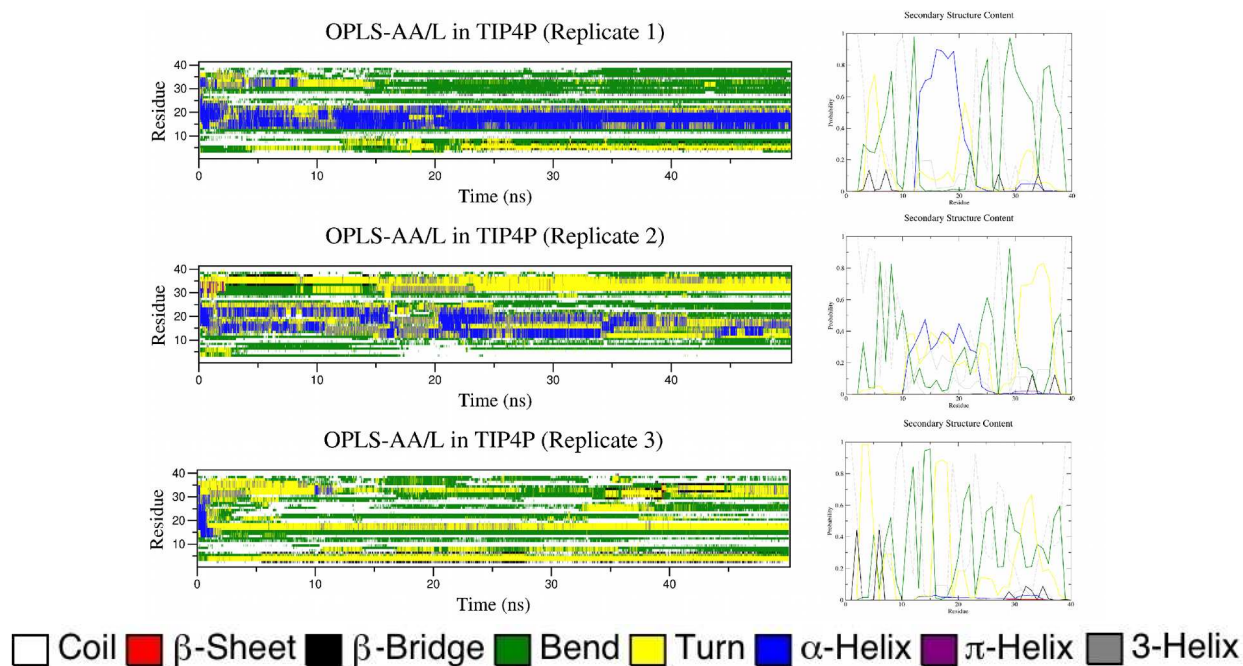


Figure 3.11: Left, secondary structure evolution of the three replicate systems simulated with the OPLS-AA parameter set in TIP4P water. Right, probabilities of given secondary structural elements as a function of residue number.

RMSD calculation performed on it. RMSD is averaged over N particles, thus giving values over time.

RMSF is calculated according to Equation 3.3:

$$RMSF = \sqrt{\frac{1}{T} \sum_{t_j=1}^T (x_i(t_j) - \tilde{x}_i)^2} \quad (3.3)$$

In Equation 3.3, T represents the time over which the RMSF is averaged, $x_i(t_j)$ is the position of particle i at time j , and \tilde{x}_i is the average position of particle i . RMSF values are also generally obtained after least-squares fitting and are averaged over time, yielding a single RMSF value for each particle. RMSF values are related to the flexibility of a protein structure, similar to the temperature factor (B-factor) calculated in crystallographic studies of proteins. Greater flexibility is indicated by greater values of the B-factor and thus the RMSF. The relationship between RMSF of particle i and its B-factor is shown in Equation 3.4:

$$\beta_i = RMSF_i \cdot \frac{8}{3} \pi^2 \quad (3.4)$$

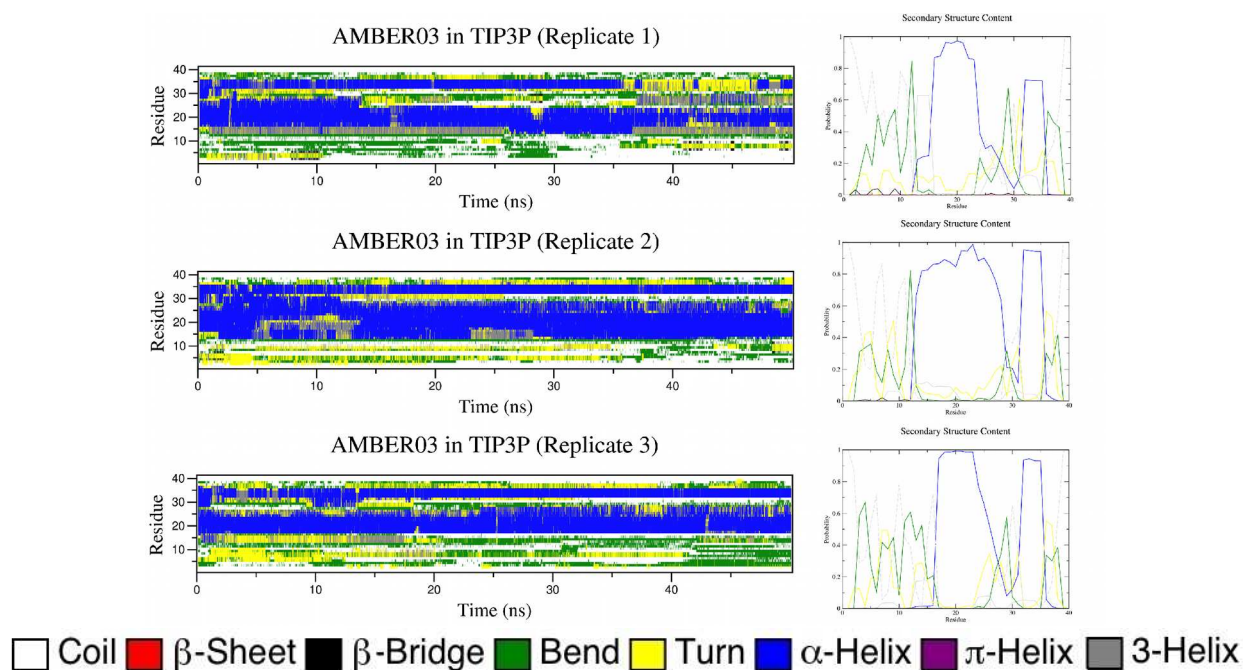


Figure 3.12: Left, secondary structure evolution of the three replicate systems simulated with the AMBER03 parameter set in TIP3P water. Right, probabilities of given secondary structural elements as a function of residue number.

Each of the force fields tested here gave comparable results for the stabilized value of RMSD of the backbone atoms of $A\beta_{40}$ (again, taken generally from the last 25 ns of each simulation). These results are listed in Table 3.5, with standard deviations taken from $n = 3$ simulations.

Table 3.5: Backbone RMSD

Force Field	RMSD (nm)
GROMOS96 43A1/SPC	0.93 ± 0.27
GROMOS96 53A6/SPC	0.98 ± 0.25
OPLS-AA/TIP4P	1.01 ± 0.03
OPLS-AA/SPC	0.92 ± 0.10
AMBER03/TIP3P	0.93 ± 0.10
AMBER03/SPC	0.84 ± 0.12

Plots of backbone RMSF per residue are shown in Figure 3.14. Values of RMSF were largely consistent within the replicates of each force field, with the exception of GROMOS96 53A6, which showed wider variation. In all simulations, the regions of greatest flexibility were terminal residues as well as especially polar regions of the $A\beta$ sequence, such as $S_8GYEVHHQK_{16}$ and $G_{25}SNKGAI_{32}$.

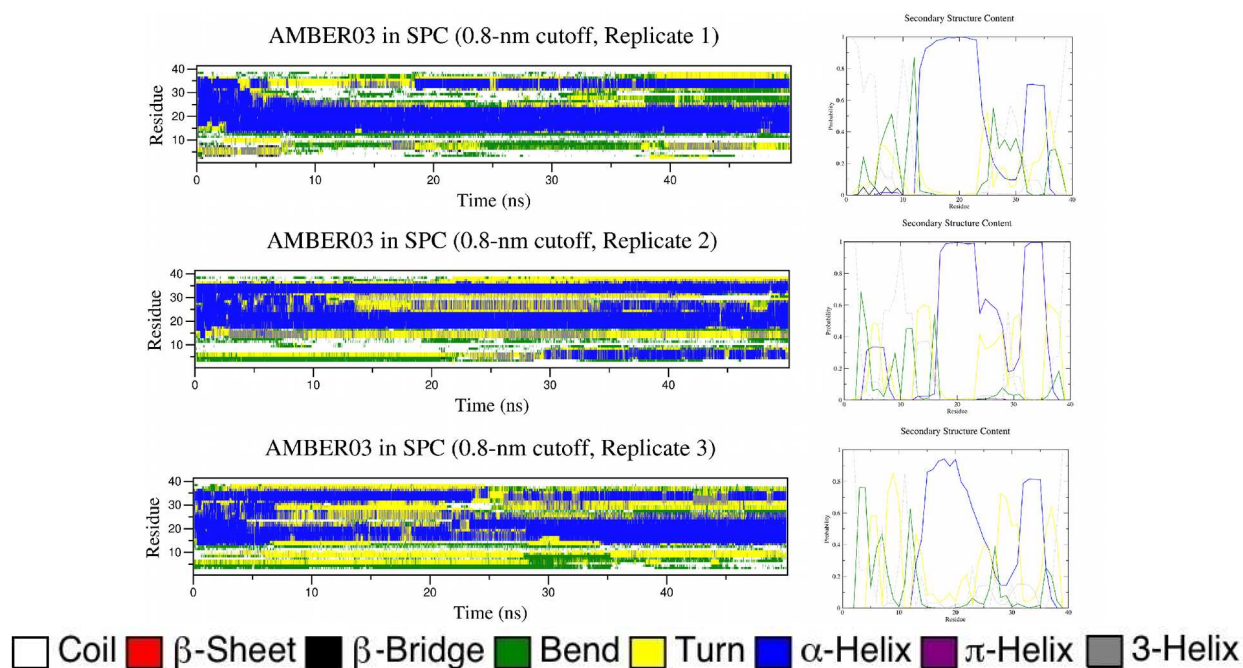


Figure 3.13: Left, secondary structure evolution of the three replicate systems simulated with the AMBER03 parameter set in SPC water. Right, probabilities of given secondary structural elements as a function of residue number.

3.3 Conclusions

The GROMOS96 53A6 parameter set is the most amenable to the simulations proposed in this body of work. Based on the data in this chapter, this parameter set most closely models the secondary structure content of $A\beta_{40}$, as determined experimentally [79]. Furthermore, this parameter set is compatible with a popular model for lipids, which will be explored later in Chapters 4 - 6.

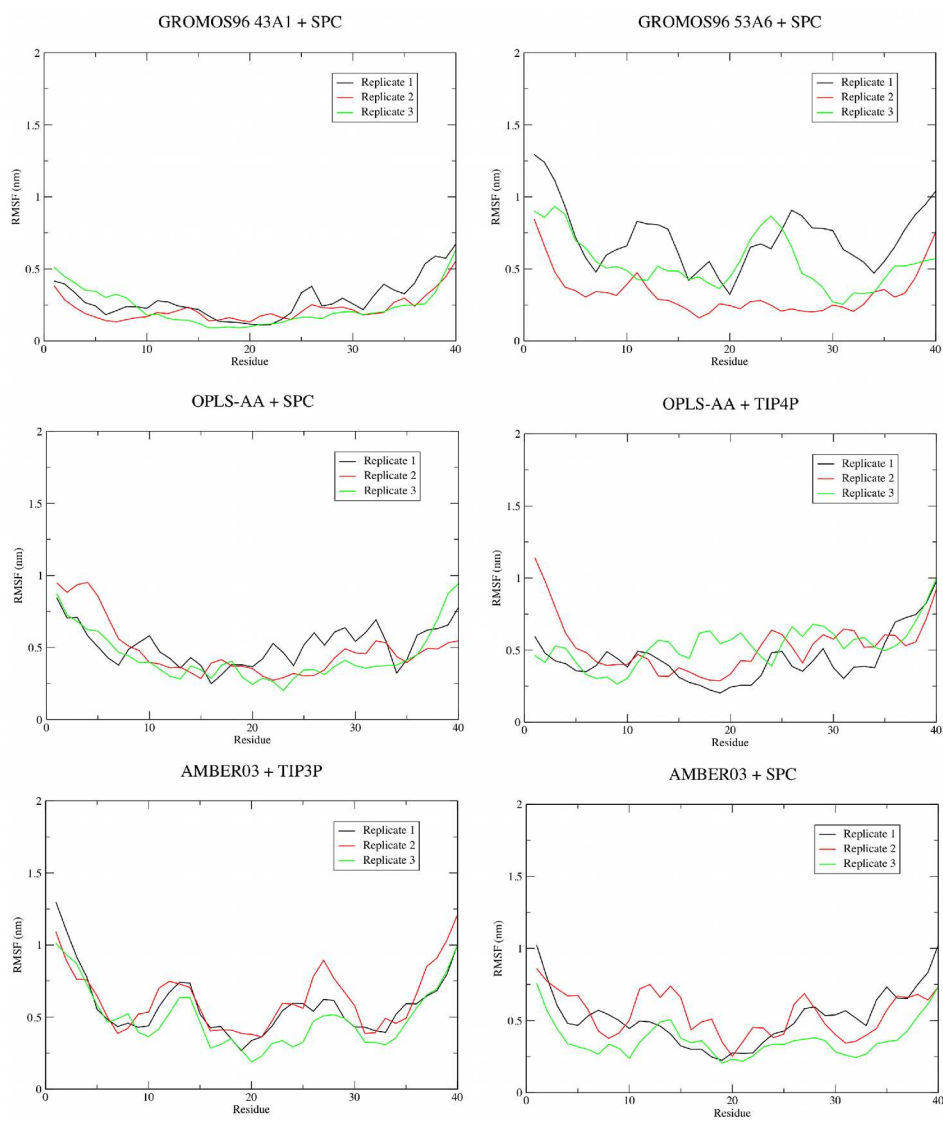


Figure 3.14: Backbone RMSF values, averaged per residue, for all force fields tested.

Chapter 4

Interactions Between Monomeric $A\beta$ and a Model DPPC Membrane

Copyright statement: The contents of this chapter are reprinted from *Archives of Biochemistry and Biophysics*, vol. 470, J. A. Lemkul and D. R. Bevan “A comparative molecular dynamics analysis of the amyloid β -peptide in a lipid bilayer” p. 54-63, copyright (2008) with permission from Elsevier, and from *FEBS Journal*, vol. 276, J. A. Lemkul and D. R. Bevan, “Perturbation of membranes by the amyloid β -peptide – a molecular dynamics study” p. 3060-3075, with permission from John Wiley and Sons.

Attribution: J. A. L. and D. R. B. designed the research. J. A. L. conducted the simulations and performed analysis. J. A. L. and D. R. B. wrote the papers.

4.1 Introduction

Since the toxicity of $A\beta$ is believed to be exerted through its interactions with the cell membrane [84, 88], it is important to understand the intrinsic physical and chemical features of the peptide that give rise to this toxicity. The experiments in this Chapter were conducted to assess the interactions of monomeric $A\beta$ in a variety of physiologically-relevant membrane models. The goal of such an analysis was to understand the mechanism by which $A\beta$ escapes its native membrane environment and/or exerts its toxicity.

Experiments have shed light on the structure of $A\beta_{40}$ and its position within the membrane [198]. The peptide has been found to be mostly α -helical in a membrane-mimicking environment, and its C-terminal helix is thought to be embedded in the lipid bilayer [77]. In a previous MD study on $A\beta_{40}$, Lys28 was assigned to the interface region between the bilayer and the aqueous environment [199]. That study found that $A\beta_{40}$ exited a dipalmitoylphosphatidylcholine (DPPC) bilayer within a 100-ns MD timeframe and much of the α -helical character of the peptide was

retained throughout the simulation. While a number of theoretical studies have been conducted on A β [199, 200, 201, 202, 203, 204, 205], the Xu *et al.* study [199] is the only one to our knowledge that had previously examined A β_{40} in an explicit bilayer environment prior to the conduction of the work presented here.

There is evidence to suggest that A β_{40} is positioned in the membrane such that Val24 is located at the interface instead of Lys28. An early study suggested that in the 99-residue C-terminal fragment (CTF), the peptide derived by β -secretase cleavage and prior to γ -secretase cleavage, Val46 (numbered such that Asp1 in A β is also Asp1 in CTF) is located within the membrane on the intracellular side of the bilayer. This positioning places the lysine triplet at residues 53 - 55 slightly away from the interface [206]. In 1998, Coles and colleagues proposed a number of possibilities for the location of A β within the plasma membrane based on the structure they solved [77]. Given this structure, if Lys53 - Lys55 were located where Tischer and Cordell [206] predicted them to be, Val24 would be positioned at the interface. The insertion depth of A β within a lipid bilayer could have important implications for its mechanism of exit and its propensity to do so.

An *in vitro* study of rat synaptic plasma membranes showed that A β peptides intercalated in the membrane [92]. That study found that soluble A β monomers were capable of penetrating deep into the bilayer core, whereas aggregates were found closer to the interface region. The peptides were found to reduce membrane thickness, of possible implication in the pathology of Alzheimer's disease. This insertion phenomenon has been shown to be strongly influenced by pH and the presence of metal ions, specifically Cu²⁺ and Zn²⁺ [207].

A later study by Yip and McLaurin used atomic-force microscopy to find that A β_{40} monomers were partially inserted into lipid bilayers rather than fully embedded in the bilayer core [208]. Furthermore, that study found that A β adopted a mostly random conformation in the presence of dimyristoylphosphatidylcholine (DMPC) micelles, which have zwitterionic headgroups. Fibrils were observed to form within the lipid headgroups, leading the authors to hypothesize that this fibril formation caused membrane disruption, which, in a cellular system, would lead to the death of the cell. These observations have implications regarding ionic balance across lipid bilayers in Alzheimer's patients. It is well-known that A β causes disruption of the ionic balance across neural cell membranes [103, 209], but the exact mechanism of this disruption has not been determined. Increased ion conductance across cell membranes has been observed in the presence of A β oligomers, which associate at the bilayer periphery [84]. This phenomenon was not observed when only soluble A β monomers were present. Another study observed that A β_{40} oligomers led to increased intracellular Ca²⁺ levels, which could signal apoptotic pathways [210].

A previous MD study showed that the behavior of A β fragments in aqueous solution is significantly affected by the presence of ions. It was demonstrated that ions had a pronounced effect on the intra- and inter-peptide interactions thought to be important in A β aggregation [211]. The presence of salt, specifically sodium chloride, has also been shown to affect the interactions between charged lipid headgroups [212]. Based on the nature of the interactions between A β and the bilayer, it is hypothesized that the presence of a physiologically-relevant salt concentration will affect the

behavior of A β at the membrane interface.

As a final note, it was shown by Ma *et al.* that pH plays a major role in the progression of amyloidosis [213]. The protonation states of ionizable sidechains on A β have implications on the rate of aggregate formation. The authors used NMR to determine the pK_a values of titratable sidechains in the 28-residue A β fragment and summarized the structural implications in different environments and at various pH values.

4.2 Dynamics of the A β ₄₀ Peptide

DPPC is a fully-saturated membrane lipid with a zwitterionic headgroup (Figure 4.1). It is commonly used in simulation studies and biophysical experiments to represent a membrane environment. For this reason, DPPC was chosen as the first lipid for analysis of A β -membrane interactions.

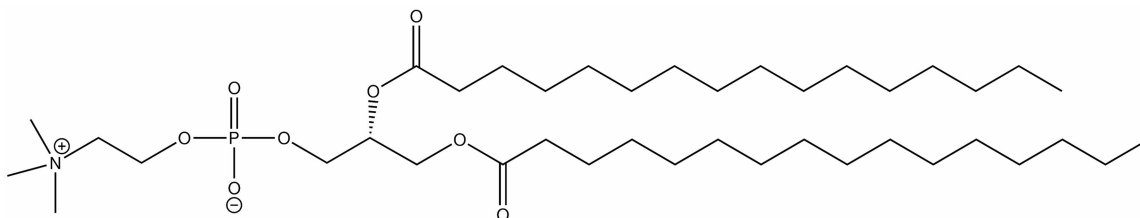


Figure 4.1: The structure of DPPC.

The simulations described in this section examine the effects of positioning (either Lys28 or Val24 at the membrane-water interface), ionic strength, pH (protonation state), and temperature on the interactions of A β ₄₀ with DPPC.

4.2.1 Methodology

Of importance in this study was the protonation state of certain amino acid residues of A β in both aqueous and membrane environments. To determine the pK_a values of the titratable residues in A β ₄₀, the structure file [PDB entry 1BA4 [77], the fourth of ten structures in the NMR ensemble determined at pH 5.1 in the presence of sodium dodecyl sulfate (SDS) micelles] was submitted to Virginia Tech's H++ server [214, 215]. This structure was chosen because its conformation was such that its N-terminal loop would not be initially located within or near the bilayer after insertion. Our simulations involved three models based on insertion depth - Models A and B located Lys28 at the interface, while Model C located Val24 at the interface. For pK_a calculations, the A β structure was separated into two parts, the segment of the peptide that would be exposed to an aqueous environment (residues 1-28 in Models A and B; residues 1-24 in Model C), and that which would be embedded in the bilayer (residues 29-40 in Models A and B; residues 25-40 in

Model C). For the aqueous segment, calculations were performed using an internal dielectric of 6 [216], external dielectric of 80 [216], ionic strength of 0.10 M, and a pH of 7 for Models A and C, and pH 5.1 for Model B. For the membrane-bound segment, calculations were performed with an internal dielectric of 6 [216], external dielectric of 2 [217, 218], ionic strength of 0 M, and pH of 7 to simulate a bilayer environment.

Ten simulation systems were prepared. Eight systems located Lys28 at the water-bilayer interface (Models A and B). Of these eight, four had deprotonated C-termini (A1, A2, B1, and B2 - see Table 4.1), and four were protonated (A3, A4, B3, and B4). Model A systems were prepared either with only enough counterions to give the system a net charge of zero (neutralizing counterions) or with 100 mM NaCl. Model B systems were prepared with only neutralizing counterions. The remaining two models located Val24 at the interface (Model C) with neutralizing counterions only and 100 mM NaCl. Both of these models had deprotonated C-termini. All preparation steps, minimizations, equilibrations, and production MD runs were conducted using facilities provided within the GROMACS package, version 3.3 [179], or scripts written in-house.

Table 4.1: Simulation Systems

Model	Ionic strength (mM)	DPPC lipids	Peptide net charge	Temp (K)	Na ⁺ ions	Cl ⁻ ions	H ₂ O molecules	Simulation ID
A	0*	122	-3	323	3	0	6912	A1
	100		-3		26	23	6873	A2
	0		-2		2	0	6922	A3
	100		-2		26	24	6878	A4
B	0	122	+1	300	0	1	6949	B1
	0		+1		0	1	6948	B2
	0		+7		0	7	6951	B3
	0		+7		0	7	6948	B4
C	0	120	-3	323	3	0	6928	C1
	100		-3		26	23	6888	C2

*An ionic strength of 0 implies that only counterions were added to these systems.

The A β_{40} peptide was solvated within a 6.5-nm cubic box of simple point charge (SPC) water and submitted for energy minimization. Water molecules were removed from the output structure to give the peptide that was used to construct the peptide/bilayer systems. Minimization *in vacuo* would likely have also yielded reasonable results.

Structure and topology files for a pre-equilibrated, 128-lipid DPPC bilayer solvated with 3655 water molecules were downloaded from the Tieleman group's website [219]. DPPC was chosen because it was used in a previous MD study [199], and it closely resembles systems used experimentally, such as DMPC and dodecylphosphocholine (DPC) micelles [213, 208], in terms of

surface charge and lipid composition. To construct the peptide/bilayer systems, a Perl script was used to overlap the coordinate files of the peptide and the DPPC bilayer and delete DPPC residues that had atoms within 1.3 Å of the peptide, such that any bad contacts at 1.4 Å could be resolved by energy minimization. The output structure was energy-minimized using the steepest descents method, first with position restraints on the heavy (non-H) atoms of A β ₄₀ ($k_{pr} = 1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$), and again without these restraints. The peptide/bilayer system was solvated by adding a slab of water to the extracellular face of the bilayer. This step was done to solvate the peptide and provide additional room for any potential movements of the N-terminal loop segment. The system was energy-minimized, after which counterions were added to compensate for the net charge on the system. To the systems with higher ionic strength, additional sodium and chloride ions were added, after which a final minimization was performed. Box dimensions following construction of the system were 6.5 nm in the lateral directions, with a height of approximately 10 nm.

4.2.2 Results

4.2.2.1 pK_a Calculations

Results of H⁺⁺ calculations are presented in Table 4.2. All models are predicted to have deprotonated C-termini, based on the very low pK_a values of these residues. The only other pK_a value that is shifted relative to the canonical value is that of Lys28 in Model C, which is elevated, and indicates that the residue is likely to be protonated under the given conditions, just as it is in the other models. All other residues have nearly canonical pK_a values. The results of these calculations compare well with the findings of Ma *et al.* [213]. While not perfectly matched, these calculated pK_a values lead to the same protonation state predicted by Ma *et al.* at physiological pH.

In Model C, both Lys28 and the C-terminal Val40 are embedded in the DPPC bilayer, but due to the dielectric continuum previously described, these residues do not experience the same dielectric environment. Thus, calculations were repeated as discussed above (see Section 4.2.1). The pK_a values for Lys28 remain above 11, and those for the C-terminus remain below 3, and thus the calculations predict that Lys28 will be protonated, and the C-terminus, deprotonated at all points within the DPPC bilayer.

4.2.2.2 Molecular Dynamics Simulations

Starting structures for each of the models are shown in Figure 4.2(a, b). Structures at the end of the simulations are illustrated in Figure 4.2(c - l). These images from the end of the trajectory provide representative structures, as few major conformational changes were observed after 50 ns of simulation. A description of the observations from the simulations is given below. The following terminology will be used to refer to various segments of the A β ₄₀ peptide: N-terminal loop (residues 1-14), helix 1 (residues 15-24), hinge region (residues 25-27), helix 2 (residues 28-36), and C-terminal loop (residues 37-40).

Table 4.2: pK_a Values, as Calculated by H++

Residue	Model A	Model B	Model C	Charge state at pH 7	Charge state at pH 5.1
N-terminus	8.4	9.0	8.4	+	+
Asp1	1.7	0.3	1.7	−	−
Glu3	4.0	3.6	4.0	−	−
Arg5	12.5	13.4	12.6	+	+
His6	6.8	7.0	6.8	0*	+
Asp7	3.6	2.1	3.6	−	−
Tyr10	11.5	12.3	11.5	0	0
Glu11	4.6	4.4	4.7	−	−
His13	6.4	6.5	6.5	0	+
His14	5.8	5.9	5.8	0	+
Lys16	10.6	11.4	10.7	+	+
Glu22	5.6	5.5	5.7	−	0 [†]
Asp23	3.6	2.7	3.9	−	−
Lys28	10.9	9.9	23.5	+	+
C-terminus	−2.8	−3.2	−3.6	−	−

*A charge state of “0” indicates that these residues were in neutral protonation states, although the histidine residues (in all models) are likely to be partially protonated (and thus contribute some positive charge) in an actual physiological system.

[†]Glu22 (in Model B) is likely to be partially deprotonated at pH 5.1 and contribute some negative charge.

4.2.2.2.1 Simulation Set A This set included models of the A β ₄₀ peptide at physiological pH and at two different ionic strengths. These simulations began with a widely-accepted construct for the positioning of the peptide in the membrane, which locates Lys28 at the water/bilayer interface.

Characteristic in this set of simulations was that A β ₄₀ remained partially embedded in the DPPC bilayer, such that Lys28 remained at the interface, with the exception of A3, wherein the peptide located itself such that Val24 was at the interface. Also typical of these simulations was loss of α -helicity in favor of random and β -structures in some portions of the peptides (Figure 4.3). For example, in A1, helix 1 lost all α -helicity within approximately 35 ns of simulation, with unwinding initiating at Glu22 and Asp23 (both deprotonated), a phenomenon predicted by Coles *et al.* [77]. Simulation A2 featured a complete loss of α -helicity within the first 10 ns of simulation, but was similar to A1 in that the peptide was located similarly within the bilayer. Simulation A3 differed from the other simulations in this group in that a central segment of this peptide retained its original α -helicity. Finally, the peptide in A4 was similar in position to A1 and A2, but retained α -helicity in helix 1. The positioning of A3 likely results from the increased hydrophobicity of its C-terminus by virtue of its protonation. The peptide in A4 also had a protonated C-terminus,

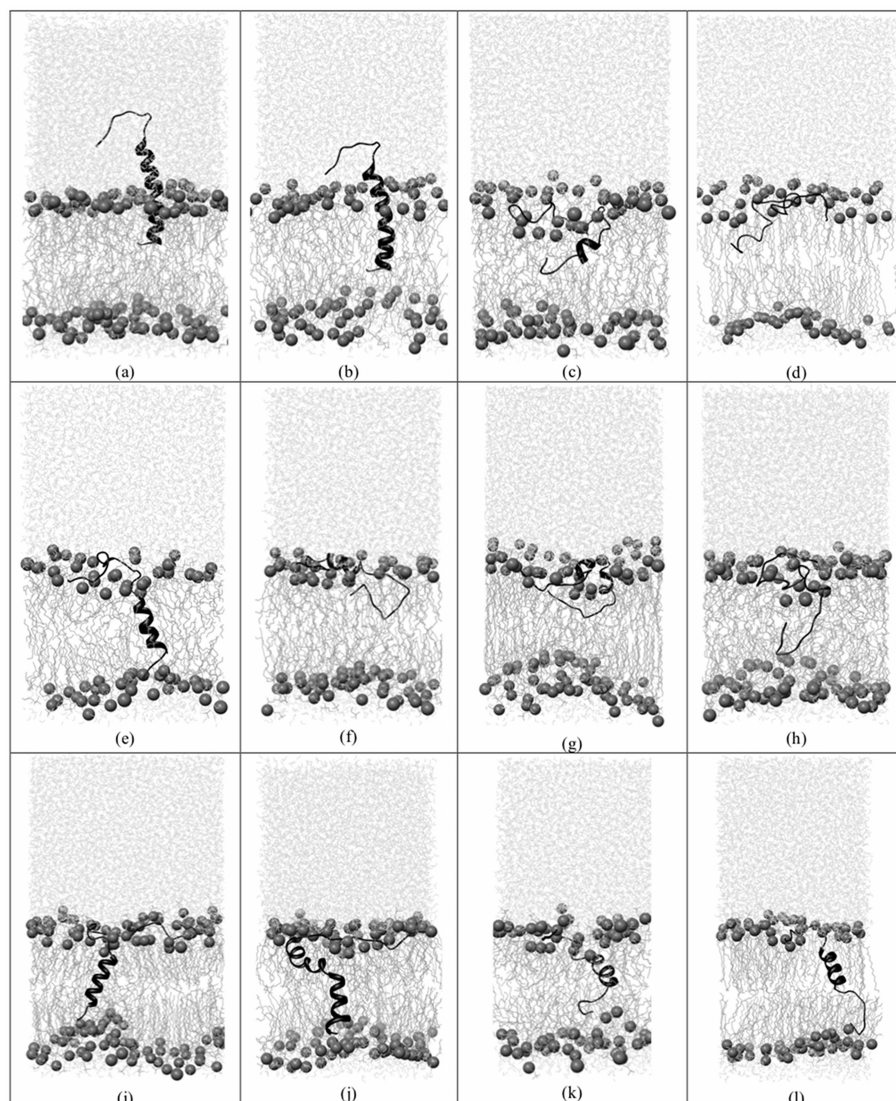


Figure 4.2: Structures from all trajectories: (a) Starting structure of Models A and B; (b) Starting structure of Model C; (c - l) Structures after 100 ns, which are representative of the trajectory: (c) A1, (d) A2, (e) A3, (f) A4, (g) B1, (h) B2, (i) B3, (j) B4, (k) C1, and (l) C2. The peptide is rendered as a ribbon, the DPPC bilayer as sticks with phosphorus atoms shown as spheres, and water molecules are shown as lines. Some DPPC molecules were deleted from in front of the peptide for clarity.

but may have been attracted more strongly to the aqueous environment due to the increased ionic strength.

Presence of $A\beta$ at the interface and the unwinding events we observed are in good agreement with the findings of Yip and McLaurin [208]. Simulation A3 agrees with experimental evidence that suggests $A\beta$ monomers are capable of penetrating the hydrophobic core of a lipid bilayer [92].

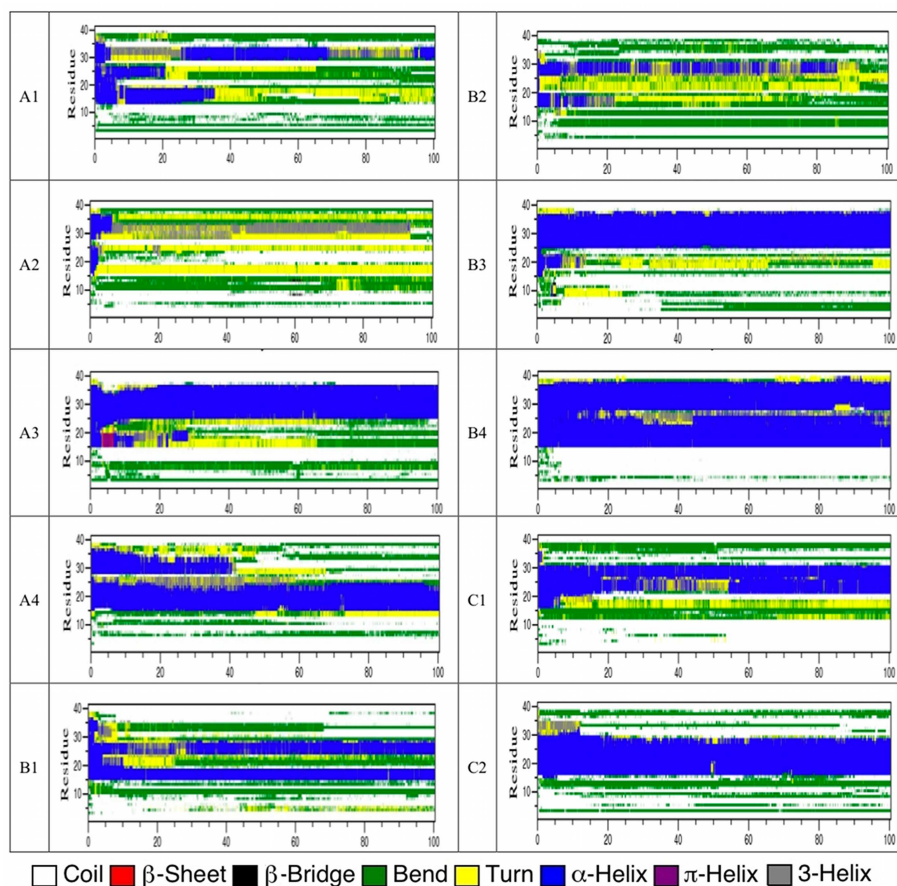


Figure 4.3: Secondary structure analysis for all trajectories. The x-axis in each diagram represents time in nanoseconds.

These findings differ from those of a similar work, wherein $A\beta_{40}$ exited a DPPC bilayer within 100 ns and maintained much of its original α -helicity [199].

4.2.2.2.2 Simulation Set B This simulation set was designed to examine the effects of protonation state (pH) and temperature, above and below the phase-transition temperature for DPPC, on the behavior of $A\beta_{40}$ and the surrounding lipid environment. The low pH environment is unrepresentative of the plasma membrane, but resembles that of the endosomal membrane environment and enables us to study the sensitivity of $A\beta$ -membrane interactions to changing protonation state.

None of the peptides in this simulation set exited the bilayer, and no β -strands were observed, but other β -structures developed over time (Figure 4.3). Simulation B1 featured two short segments that retained α -helicity throughout 100 ns. Simulation B2, carried out at a higher temperature, showed similar helical features, but helicity in residues 15-20 eventually disappeared over time. The $A\beta_{40}$ peptide in both B1 and B2 adopted positioning that placed Val24 at the interface, slightly deeper than the prevailing positioning in simulation set A. In simulations B3 and B4, in which

all ionizable residues in A β_{40} were protonated, the peptide became embedded even more deeply within the DPPC bilayer, such that Lys16 was at the interface throughout most of the trajectory. These peptides retained an appreciable amount of the original α -helical content and adopted nearly transmembrane orientations by the end of 100 ns.

The simulations conducted at 300 K, below the phase-transition temperature of DPPC (B1 and B3), clearly show the effects of the peptide on the surrounding lipids. While DPPC residues far from the peptide tended to extend their acyl chains and pack together, those near the peptide are noticeably more disordered, clearly showing a disruptive effect of the peptide on the surrounding membrane environment. This disruption will be further explored in Section 4.3.

Increased α -helicity as pH decreases has been observed experimentally [77, 213], and we have succeeded in simulating these observations. The NMR data from Coles *et al.* predicted α -helicity in residues 15-36, broken only by a hinge at residues 25-27 [77]. In simulations B1 and B2, conducted at a simulated pH that mimicked those of the experiment, we observed retention of helicity in residues 15-30, broken by the same hinge. Some α -helicity was lost in B2, likely due to the higher temperature of the simulation (323 K). The structural determination by Coles *et al.* was conducted at 298 K, slightly lower than the temperature of B1, and much lower than that of B2.

4.2.2.2.3 Simulation Set C The purpose of the final two simulations was to monitor the effects of deeper initial insertion depth of A β_{40} in the bilayer, a less widely-accepted construct that has some experimental backing [206]. In this simulation set, we examine the effects of initial insertion depth and ionic strength on a system under physiologically-relevant pH and ionic strength.

The peptides in both of these trajectories positioned themselves within the bilayer such that Lys16 was at the interface, an orientation that was maintained throughout the trajectory. Both simulations featured peptides that retained α -helicity in a central segment (Figure 4.3), with the increased ionic strength of C2 having no obvious effect on the behavior of the peptide. Positioning of the A β_{40} peptide in the core of the bilayer is in good agreement with the findings of Mason *et al.* [92]. We observed no β -strand structures in either trajectory, and the peptide did not exit the bilayer in the 100-ns timeframe.

An interesting phenomenon that occurred in both C1 and C2 was the persistence of water in the bilayer (Figure 4.4). Water molecules entered the bilayer and associated with the four charged groups in the hydrophobic interior (Glu22, Asp23, Lys28, and the C-terminus). This observation is similar to one made in an MD study of melittin [220], where water was thought to enter the bilayer and compromise the integrity of the membrane. This event compares well with experimental observations that membranes are disrupted in the presence of A β [91].

The behavior of the charged residues within the hydrophobic interior of the lipid bilayer is worth noting. The presence of charged residues within a lipid bilayer during simulations is not unprecedented. As can be seen from Figure 4.4, these charged residues appear to be capable of “snorkeling” to the interface region, a phenomenon that has been reported in other simulations of transmembrane helices [221, 222, 223]. Snorkeling was observed in all of our simulations in which

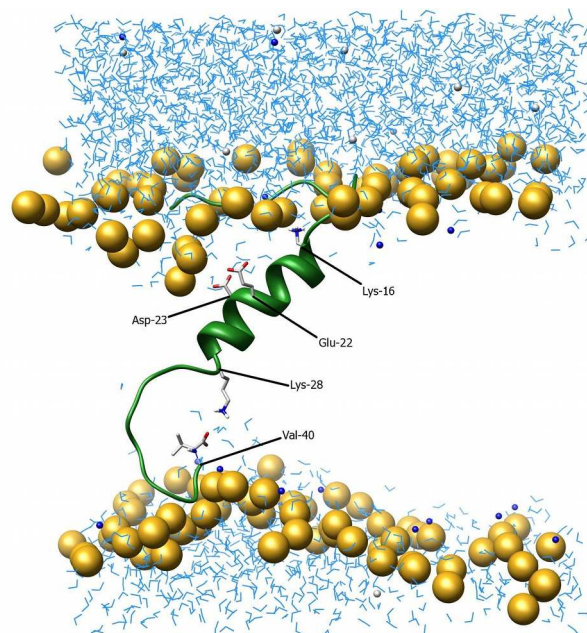


Figure 4.4: Water leakage into the lipid bilayer, illustrated using a snapshot at 80 ns from simulation C2. The bilayer has been removed for clarity, but phosphorus atoms are shown as large spheres for perspective. The peptide is shown as a ribbon, with charged side chains labeled, shown as sticks, and colored by element. Water molecules are shown as lines and ions are shown as small spheres (dark blue = Na⁺, light blue = Cl⁻).

charged residues were located within the core of the bilayer (A3, B3, B4, C1, and C2).

4.2.3 Discussion

In a recent review, Haass and Selkoe present a model for the generation of A β from APP that describes the association of A β with membrane compartments [224]. Although A β is depicted as exiting the membrane after γ -secretase cleavage, no model was proposed to describe the release of A β to the extracellular space. To develop a complete understanding of the role of A β in Alzheimer's disease, it is necessary to devise studies that consider A β in association with membranes. Atomic-level descriptions of A β association with lipid bilayers are not currently feasible using experimental methods, though several investigators have examined its structure in the presence of membrane-mimicking environments such as organic solvents, detergents, and phospholipid micelles [77, 208, 213, 225].

Computational methods are applicable to studies of peptide and protein interactions with phospholipid bilayers as models of cell membranes, and a review summarizing some of these studies has recently appeared [226]. Of particular relevance to our work are studies of single helix spanning peptides. For example, early simulation studies utilized melittin in a DPPC bilayer [220, 227].

Melittin was inserted in an α -helical conformation with the helix oriented perpendicular to the bilayer surface, but the helix did not completely span the membrane. During the MD simulation, melittin maintained its α -helical conformation, but was observed to adopt a 25° tilt relative to the membrane normal, and a bend of approximately 30° was seen in the helix near the N-terminus, which was the end extending farthest into the bilayer. An MD simulation of alamethicin in a POPC bilayer showed similar results in that a bending motion was observed in the helix, though no appreciable tilt of the peptide was reported [228]. These studies were necessarily short (< 2 ns) due to computational resources available at the time. Thus, it is not unexpected that more extensive conformational changes were not observed. More recently, simulations of model (WALP) $_n$ and (KALP) $_n$ transmembrane helices have been conducted. Kandaswamy and Larson [229] applied MD to studies of (KALP) $_n$ peptides in DLPC, DMPC, and DPPC bilayers, with a particular focus on understanding the effect of hydrophobic mismatch, in which the length of the hydrophobic transmembrane segment varied along with the width of the lipid bilayer. In simulations extending out to 200 ns, they observed that the peptides remained α -helical and that they tilted relative to the bilayer normal when necessary to accommodate the hydrophobic mismatch. Also of relevance are MD studies of the insertion of WALP peptides into membrane environments [230, 231]. Both studies started with an extended peptide and applied replica exchange MD to increase conformational sampling, though Im and Brooks [230] utilized a generalized Born implicit membrane model and Nymeyer *et al.* [231] used an explicit DPPC bilayer. Very distinct differences were observed in that in one case the peptide formed a helix prior to insertion [230] and in the other, the extended peptide inserted into the DPPC bilayer, after which it folded into a transmembrane helix [231]. These studies present an interesting comparison with our work in that they reveal the range of conformations that may be observed with membrane-associated peptides.

To our knowledge, only two computational studies examined A β in a membrane-like environment prior to our work. One of these, by Mobley *et al.* [205], examined A β_{40} and A β_{42} in an implicit membrane environment. Through the use of multiple Monte Carlo (MC) simulations, the authors report a number of insertion modes of A β , each with comparable energy. The other study, by Xu *et al.*, applied MD to studies of A β_{40} in an explicit lipid bilayer composed of DPPC [199]. The results presented by these investigators suggest that the exit pathway of A β_{40} from DPPC can be observed within 100 ns of MD simulation, and that much of the initial α -helicity of the peptide is maintained. Our work builds upon this previous research.

One consideration in initiating our studies was the starting conformation of A β_{40} . As we have noted, we selected a structure from an NMR study of A β_{40} in SDS micelles. The particular structure we selected was representative of the 10 reported, and it provided us with a starting conformation in which the N-terminal portion of the peptide extended away from the DPPC bilayer. A similar conformation was used by Xu *et al.* in their study, and this conformation also has been used in aqueous simulations [201]. Notably, a starting structure incorporating an α -helical segment in the bilayer is reasonable considering the preponderance of this structure in single membrane-spanning protein segments. Another consideration in our studies was the composition of the lipid bilayer. Although a number of different compositions might have been chosen, we elected to use DPPC because it was used in the Xu *et al.* study, thereby allowing some comparison of our work

with the only other prior computational study of A β in an explicit lipid environment.

Of note is the difference between the results we present and the work presented by Xu *et al.* [199], which may be attributed to differences in the force fields used or the way the system was prepared. For this study, we used an updated version of the GROMOS force field [174]. There are subtle differences between the force field used by Xu *et al.* and that of the present study. The GROMOS96 53A6 parameter set contains charges that generate slightly more polar entities (Tyr, His, *etc.*) than in the GROMOS87 force field. The Lennard-Jones parameters were also refined in GROMOS96 53A6. As a result, the nature and strength of the nonbonded interactions between the peptide and the bilayer may be slightly different. The peptide in the present study may have experienced stronger interactions with lipid headgroups, preventing diffusion into the bulk solvent. In preparing the system, our methodology differed in that we created a cavity for the peptide with void space that was closed during equilibration, allowing the lipids to pack around the A β_{40} peptide. The authors of the previous study report removing one DPPC molecule from the pre-equilibrated bilayer prior to A β_{40} insertion. This method may have resulted in different forces between the lipids and the inserted peptide than our method.

The fact that large conformational changes in the A β peptide were observed in a 100-ns timeframe is not surprising, as previous aqueous MD studies have witnessed unfolding of A β within 10 ns using the GROMOS87 force field [201] and in 50 ns using GROMOS96 [203]. Fibril assembly has been modeled as occurring in under 10 ns with the AMBER94 and CHARMM27 force fields [202], and A β oligomers have been assembled from peptide fragments under the CHARMM and CFF91 force fields [232].

We recognize that it is not possible to comprehensively sample all possible conformations over the course of a single 100-ns MD simulation. In fact, some of the structural differences that we observed among the simulations may arise from limitations in conformational sampling rather than from the conditions imposed during the simulations. Nevertheless, reasonably consistent results were observed under some conditions, from which we were able to make some general observations.

4.2.3.1 pH Effects

In simulations at physiologically-relevant pH (A1-A4, C1-C2), our results compare well with the findings of Coles *et al.*, who in addition to reporting the structure of A β_{40} at pH 5.1, found that at pH above 6.0, residues 15-24 of the peptide were capable of assuming an unstructured conformation [77]. We found that this disorder persisted in simulations A1, A2, and A4, and there was a transient loss of helicity in this region of the peptide in simulation C1.

In each of the low-pH simulations (B1-B4), A β_{40} retained an appreciable amount of α -helical character, except for simulation B2, which lost much of its helicity by the end of the trajectory. In the two simulations of fully-protonated A β_{40} , α -helicity was preserved in residues 25-36 in B3, and 15-36 in B4, broken only by a bend at residues 25-27. This finding suggests that protonation

of acidic residues and histidines helps stabilize helical structure in the presence of a membrane. The observation regarding B4 compares exactly with the structural predictions made by Coles *et al.* regarding A β in the presence of SDS micelles, although B4 represented a system pH lower than that of all the systems tested by Coles *et al.* [77]. Since the two systems (SDS micelles and DPPC bilayers) differ in both surface charge and lipid composition, there must be some common feature inherent to A β that allows this structure to be observed.

Unwinding of A β_{40} has been observed experimentally to be a pH-dependent process, and our simulations agree well with these observed data, in that in most of our low-pH simulations (B1-B4, where A β_{40} was highly protonated), much of the original α -helicity was retained along the trajectory. This contrasts with the more physiologically-relevant simulations (A1-A4, C1-C2), wherein we observed a greater degree of loss of α -helicity in favor of β -structures and random coils.

4.2.3.2 Insertion Depth

Models A and C in this study modeled the A β peptide under comparable conditions but at two different starting locations in the DPPC bilayer (Lys28 and Val24 at the interface, respectively). As observed in the Model C simulations, locating Val24 at the interface consistently caused the peptide to adopt a more transmembrane orientation. This phenomenon was observed once in the Model A series (simulation A3), but for the rest of the Model A simulations, A β_{40} remained partially embedded, with Lys28 at the interface. This occurrence may have been a result of two factors, the first of which is the fact that in A3, the C-terminus of the peptide was protonated, making it more hydrophobic and thus more likely to implant in the bilayer. Simulation A3 was conducted in the absence of additional ionic strength, which may be the reason for the difference between A3 and A4, both of which had protonated C-termini. A4 was simulated in the presence of ~ 100 mM NaCl, which may have kept the aqueous segment of the peptide from penetrating the bilayer.

4.2.3.3 Electrostatic Interactions and Ionic Strength

In the present study, we found a number of electrostatic interactions common among all of the simulations conducted. The most notable of these were cation- π interactions, which most frequently formed between the choline headgroups and residues Phe4, Tyr10, Phe19, and Phe20. In cases where A β_{40} embedded in the bilayer more deeply, cation- π interactions involving Phe19 and Phe20 were not observed, as these residues resided in the bilayer interior. Another important electrostatic interaction involves Arg5. This residue associated strongly with phosphate moieties of the lipid headgroups, and in some instances was found to embed into the ester region of the bilayer (A1, B3, and C2), where it remained for a significant portion of the trajectory. Aromatic interactions with lipid headgroups have been observed computationally to persist for several nanoseconds at a time, and positively-charged residues have also been found to embed themselves in the phosphate region of the bilayer [233].

The presence of a 100 mM ionic strength appeared to have little consistent effect on the properties of the peptide. In one case (A2), helix 1 of the peptide unfolded rapidly, but in another (A4), much of the helicity in this region was maintained, even after 100 ns. Helix 2 lost much of its α -helical character in both of these cases, indicating that the increased ionic strength may have had an effect on the dielectric environment near the membrane-water interface and thus the properties of the embedded helix. Simulation C2 was also conducted in the presence of ~ 100 mM NaCl, but exhibited no obvious differences relative to C1, conducted in the presence of only Na⁺ counterions.

In all simulations, we witnessed the elongation of a disordered segment across the surface of the lipid bilayer, within the phosphatidylcholine headgroups. This segment stretched from residues 1-16, 1-24, or 1-28, depending on how far the peptide was embedded in the bilayer over time. This finding agrees with experimental observations of Yip and McLaurin [208], who found amyloid fibrils within the lipid headgroups. Also related to these findings are the results obtained by Bokvist *et al.* [234], who used a combination of NMR and circular dichroism spectroscopy to find that a short segment of A β is embedded in membranes with neutral headgroups. These investigators also found that A β consists of mostly β -structures at pH 7.8, with very little α -helical character. These findings correlate with our simulations A1 and A2, which utilize the same protonation state studied by Bokvist *et al.* [234].

For amyloid plaques to form, A β peptides must exit the bilayer and self-associate. We did not observe exit of the peptide from the bilayer in our simulations, and at least two factors may account for this observation. One is that the length of time of the simulations simply was not sufficient to observe this behavior of the peptide. The second is that experimental evidence indicates that a nucleation site is needed to initiate the aggregation process. Specifically, the association of A β with gangliosides in lipid rafts and glycosaminoglycans may be important nucleation sites for A β aggregation [235, 236]. Perhaps A β in these specialized membrane regions would be more susceptible to exiting the bilayer. These types of systems will be discussed in Chapter 5.

4.3 A β_{40} Disrupts Equilibrium Lipid Properties

In this section, we aim to expand upon our previous work by examining the properties of lipid molecules surrounding the membrane-perturbing A β_{40} peptide. While it is known that A β can interact with the plasma membrane and assemble in this environment [208], a fundamental understanding of the molecular basis for this phenomenon is missing. Central questions still remain, especially regarding the intrinsic characteristics, both structural and chemical, of A β that allow it to disrupt the surrounding lipids. Crucial to understanding this phenomenon are detailed studies with atomic resolution, such as the simulations we present here. A greater knowledge of the most basic interactions between A β and a model membrane can lead to a more complete understanding of the membrane-aided assembly of A β and the resulting damage to cell membranes.

4.3.1 Methodology

Simulation systems were constructed and simulated as described in Section 4.2.1. Systems belonging to simulation set A were designed primarily to understand the effects of increased salt concentration. Simulation set B examined the effects of both protonation state of the peptide and temperature on the behavior of the system. Finally, the systems in simulation set C also examined the effects of increased salt, but also contrasted with simulation set A in that the A β peptide was placed more deeply in the membrane.

Two sets of negative control systems of pure DPPC bilayers were prepared by a similar method. These systems were designed to examine whether the additional solvation or increased ionic strength had any background effect on lipid dynamics. Two systems were prepared from this structure the original bilayer with the original water-to-lipid solvation ratio (“Original Solvation,” OS1), and this bilayer in the presence of 100 mM NaCl (OS2). In addition, three systems were prepared by placing an additional slab of water to one side of the bilayer to approximate the increased water-to-lipid solvation ratio and the system size present in the peptide-bilayer systems. Similarly to the OS simulation set, these systems contained either no salt (“New Solvation,” NS1 and NS3) or 100 mM NaCl (NS2). The solvation ratios of NS1, NS2, and NS3 closely, but do not exactly, match those of the simulated A β systems. They were instead designed to strike a balance between the dimensions of the system and the number of water molecules to examine whether or not the asymmetry of the system and increased solvation would affect the dynamics of the lipids. System details are summarized in Tables 4.1 and 4.3.

Positive control systems were prepared with melittin the presence of DPPC. The structure of melittin was taken from the crystal structure, PDB entry 2MLT [237]. Two orientations were prepared, one with melittin embedded in the DPPC bilayer, as in previous studies [220, 227] (“Embedded,” E1, and with 100 mM NaCl, E2) and the other with melittin parallel to the bilayer interface as in [238] (“Parallel,” P1, and with 100 mM NaCl, P2). These systems were prepared in the same manner as the A β -DPPC systems, giving starting configurations comparable to those presented in the original works. Details of these systems are presented in Table 4.4.

In each simulation, coordinates were saved every 2 ps, generating 50,000 data points per simulation. Analyses were conducted using tools within the GROMACS package, version 3.3 [179] (for deuterium order parameters) and code developed in-house [239] (for lipid tilt, effective chain length, area per lipid headgroups, and bilayer thickness). Averaging over time was conducted, when appropriate, to generate a time-dependent progression of these measurements.

The initial asymmetric orientation of A β relative to the DPPC bilayer creates an interesting situation when analyzing the properties of the surrounding lipid bilayer. Over time, the peptide interacts differently with each leaflet. Such a situation resembles that of melittin, whose interactions with lipids have been described experimentally [240, 241, 242, 243] and computationally [220, 227, 238].

Table 4.3: Properties of Control DPPC Membranes

Model	Ionic strength (mM)	DPPC lipids	Temp (K)	H ₂ O molecules	Solvation Ratio
OS1	0*	128	323	3655	28.6 : 1
OS2	100			3641	28.4 : 1
NS1	0			6907	54.0 : 1
NS2	100			6881	53.8 : 1
NS3	0	128	300	6907	54.0 : 1

*An ionic strength of 0 implies that only counterions were added to these systems.

Table 4.4: Properties of Melittin-Membrane Systems

Model	Ionic strength (mM)	DPPC lipids	Temp (K)	H ₂ O molecules	Solvation Ratio
E1	0*	119	323	7071	59.4 : 1
E2	100			6993	58.8 : 1
P1	0	118	323	7047	59.7 : 1
P2	100			7005	59.4 : 1

*An ionic strength of 0 implies that only counterions were added to these systems.

4.3.2 Results

4.3.2.1 Validity of Melittin-DPPC Controls Relative to Previous Work

The work of Bachar and Becker [227] and Bernèche *et al.* [238] provide meaningful reference points to our current simulations. We constructed systems that had initial configurations similar to those produced by the original investigators, but we simulated the systems in a somewhat different manner. We applied new equilibration schemes to pack the lipids around the peptides, using different force field parameters (GROMOS/Berger instead of CHARMM). We also conducted simulations that were far longer than in the original reports, 100 ns instead of 300-500 ps. The goal of this series of simulations was to produce data not only to validate our simulation setup, but also to serve as a basis for comparing the effects of A β ₄₀ on a DPPC bilayer in light of the observations made with respect to melittin.

Simulation E1 was inspired by the work of Bachar and Becker [227], and simulation E2 arose from our desire to examine the effects of increased ionic strength on the peptide-membrane systems. Even in our longer 100-ns simulations, the positioning and orientation of melittin at the end of the simulations was similar to that reported by these authors, although in our simulations melittin

became embedded more deeply in the bilayer and more disordered at its termini. The disordering at the termini was predicted by the authors of the original work, but was not observed in the timeframe they simulated [227].

With respect to the lipid properties, the most meaningful comparison between this work and the previous simulations arises with respect to lipid order. Deuterium order parameters describe the orientation of the lipid acyl chains, on average, relative to the bilayer normal. These parameters are calculated by the equation:

$$-S_{CD} = \left\langle \frac{3\cos^2\theta - 1}{2} \right\rangle \quad (4.1)$$

In Equation 4.1, θ represents the angle between the C-D bond and the bilayer normal, and the angle brackets denote that the values are averaged over all equivalent atoms, and over time.

We observed that the lipids nearest melittin experience a greater degree of disorder, whereas more distant lipids become more ordered relative to control simulations in the absence of melittin. This disordering effect is comparable to the results obtained in the original reports. In addition, the top leaflet of the bilayer, which interacts with melittin most strongly, was observed to be more disordered relative to the bottom leaflet, which experienced a greater degree of chain elongation and lipid packing. The authors of the original work divided the lipids in their bilayer into “tiers” based on the distance between the protein and lipid molecule center of mass. The authors presented the average value of $-S_{CD}$ for the “plateau region” of the acyl chain, which extends from carbons 4 to 8 (denoted $\langle -S_{CD} \rangle_{[4,8]}$). The values they reported are 0.157 ± 0.009 , 0.215 ± 0.006 , and 0.215 ± 0.006 for the first, second, and third tiers, respectively. We find very similar values of 0.144 ± 0.010 , 0.194 ± 0.007 , and 0.219 ± 0.005 for these same subsets of lipids. We attribute the small differences in these values to the use of different force fields, application of different equilibration schemes, and the length of our simulations, which is several orders of magnitude longer than that of the original work.

Similar conclusions can be made between our simulation P1 (i.e., starting with melittin parallel to the interface of the bilayer at the beginning of the simulation) and the work of Bernèche *et al.* [238]. We make the same observations with respect to the behavior of the lipids, that those of the top leaflet (which also interact most strongly with melittin) become very disordered relative to the lipids of the lower leaflet, overall, although our values for the deuterium order parameters are higher. In the original work, the average order parameter in the top leaflet was 0.149, and 0.188 (a difference of 21%) in the bottom leaflet. The corresponding values for these parameters from our simulations are 0.157 and 0.220 (29% difference), respectively. We attribute these differences to many of the same factors as described above with respect to the work of Bachar and Becker, and also the fact that the simulations conducted by Bernèche *et al.* utilized DMPC as the membrane lipid instead of DPPC, so some differences should be expected.

4.3.2.2 Deuterium Order Parameters of the A β_{40} -DPPC systems

In our work with A β_{40} , we examine a peptide that is primarily asymmetric with respect to its interactions with the lipid membrane. As such, we analyzed the deuterium order parameters of each leaflet separately, including the whole acyl chain and the “plateau region” described above. Taking the approach applied by Bachar and Becker, we analyzed the lipids in “tiers” at increasing distance from A β_{40} . The first and second tiers contained 20 lipids each, and the third tier contained the remaining lipids in the leaflet, between 18 and 24. The results of these calculations are presented in Table 4.5. Note that the tiered analysis does not apply to the DPPC-only controls; the values presented represent an average of the plateau region for each leaflet of the bilayer.

From these data, it can be seen that, overall, the $-S_{CD}$ values in the top leaflet of the A β_{40} -DPPC systems are lower than those in the bottom leaflet. One explanation for this phenomenon has been proposed by Tieleman *et al.* [244], wherein the lipids that interact strongly with the protein become increasingly tilted relative to the bilayer normal, causing the angle between the C-D bond and the normal to decrease. This occurrence was reported in simulations of melittin [227], and occurs here as well in the case of both melittin and A β_{40} . The lipids of the top leaflet tend to adopt an angle such that they become tilted with their headgroups pointing towards A β_{40} , and the lipids of the lower leaflet elongate to become more ordered, filling the void in the center of the bilayer (Figure 4.5). The results of simulations C1 and C2 reflect the fact that the peptides in these simulations interacted more or less symmetrically with both leaflets over time. The peptide became deeply inserted in the bilayer in a transmembrane orientation, with disordered N- and C-termini protruding through the lipid headgroups of both leaflets.

We also note that the lipids in the first tier tend to be more disordered than those of the second and third tiers. In fact, in most cases, the values of $\langle -S_{CD} \rangle_{[4,8]}$ increase as the distance between the peptide and the lipids increases. The presence of the A β_{40} peptide causes substantial disorder on the lipids with which it most closely interacts, while simultaneously resulting in an increase in order of the lipids that are further away. This behavior is dependent upon the conformation of the peptide. In cases where A β_{40} lost much of its initial α -helicity, the nearby lipids become more disordered and the more distant lipids increase in order. In cases where the peptide unfolds to a lesser extent (see for example, simulation B4), the distant lipids approach a value of $\langle -S_{CD} \rangle_{[4,8]}$ that is comparable to that of the relevant control (NS1), based on the average order parameter of Tier 3 in the two leaflets. We thus conclude from these data that A β interacts with the membrane in a manner similar to melittin with respect to its effects on the disordering of the surrounding lipids.

In addition, the $\langle -S_{CD} \rangle$ values for the top leaflet lipids of the peptide-membrane systems are primarily lower than the respective controls (NS1 or NS2), while the bottom leaflet lipids are more ordered than the controls. This behavior is a result of the fact that the top leaflet interacts strongly with the unfolded and charged portions of the peptides in each simulation, and is especially true in the case of the lipids closest to the peptide (Tier 1). The values of $\langle -S_{CD} \rangle$ for the controls are in good agreement with previous experimental and simulation studies [245, 246].

The contraction of the lipid headgroups, and concomitant disordering of the acyl chains of the

Table 4.5: Average Values of $-S_{CD}$ for A β_{40} -DPPC systems

System	Top Leaflet $\langle -S_{CD} \rangle_{[4,8]}$			Bottom Leaflet $\langle -S_{CD} \rangle_{[4,8]}$			Whole Acyl Chain	
	Tier 1	Tier 2	Tier 3	Tier 1	Tier 2	Tier 3	Top	Bottom
A1	0.163 (0.007)*	0.223 (0.008)	0.233 (0.007)	0.180 (0.005)	0.211 (0.009)	0.216 (0.006)	0.177	0.217
A2	0.209 (0.008)	0.274 (0.006)	0.285 (0.011)	0.254 (0.010)	0.235 (0.006)	0.263 (0.010)	0.223	0.260
A3	0.147 (0.008)	0.219 (0.005)	0.231 (0.004)	0.192 (0.006)	0.191 (0.005)	0.203 (0.009)	0.173	0.231
A4	0.188 (0.009)	0.250 (0.004)	0.273 (0.006)	0.197 (0.015)	0.247 (0.006)	0.250 (0.003)	0.207	0.238
B1	0.205 (0.019)	0.243 (0.016)	0.318 (0.017)	0.189 (0.019)	0.271 (0.020)	0.268 (0.018)	0.231	0.263
B2	0.134 (0.008)	0.213 (0.006)	0.227 (0.004)	0.165 (0.004)	0.193 (0.006)	0.203 (0.006)	0.167	0.213
B3	0.233 (0.007)	0.249 (0.013)	0.307 (0.014)	0.212 (0.008)	0.246 (0.014)	0.286 (0.010)	0.240	0.280
B4	0.160 (0.010)	0.220 (0.006)	0.239 (0.008)	0.220 (0.006)	0.191 (0.004)	0.197 (0.007)	0.176	0.213
C1	0.161 (0.010)	0.240 (0.005)	0.288 (0.007)	0.200 (0.008)	0.221 (0.005)	0.247 (0.006)	0.204	0.196
C2	0.183 (0.012)	0.284 (0.007)	0.321 (0.012)	0.232 (0.010)	0.269 (0.005)	0.275 (0.009)	0.240	0.245
OS1	0.216 (0.004)			0.215 (0.004)			0.185	0.185
OS2	0.232 (0.004)			0.231 (0.005)			0.201	0.201
NS1	0.217 (0.003)			0.217 (0.004)			0.188	0.188
NS2	0.242 (0.004)			0.243 (0.003)			0.213	0.214
NS3	0.252 (0.007)			0.253 (0.009)			0.229	0.229
E1	0.166 (0.007)	0.218 (0.004)	0.235 (0.005)	0.206 (0.010)	0.193 (0.004)	0.211 (0.005)	0.178	0.206
E2	0.217 (0.005)	0.252 (0.005)	0.245 (0.005)	0.225 (0.004)	0.253 (0.008)	0.232 (0.007)	0.210	0.245
P1	0.177 (0.013)	0.200 (0.005)	0.203 (0.004)	0.176 (0.012)	0.205 (0.008)	0.198 (0.005)	0.157	0.220
P2	0.202 (0.009)	0.228 (0.004)	0.247 (0.006)	0.205 (0.007)	0.229 (0.006)	0.228 (0.004)	0.187	0.252

*Mean shown, with (standard deviation), if applicable.

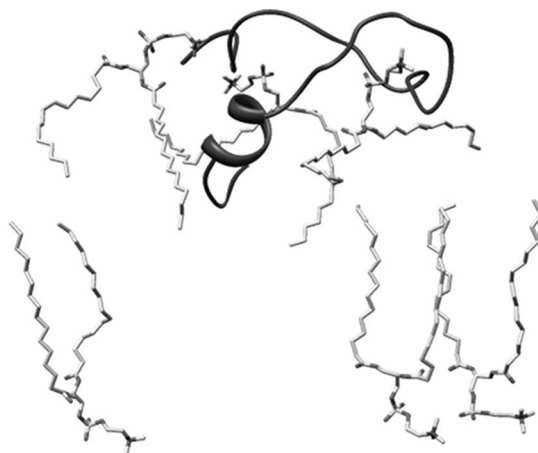


Figure 4.5: As shown in a snapshot from the end of trajectory A1, the $A\beta_{40}$ peptide causes DPPC lipids of the top leaflet of the bilayer to become more disordered, with their acyl chains becoming more parallel to the bilayer surface. Lipids in the bottom leaflet become more ordered, extending their acyl chains to fill in the growing void in the center of the bilayer. Representative lipids (sticks) near the peptide (ribbon) are shown.

lipids in closest contact with $A\beta$, results in no substantial changes in the overall density of the lipid bilayer. There is a slight increase in density among the lipids nearest $A\beta$ (likely a result of the strong interaction between $A\beta$ and the lipid headgroups, discussed below), but regions of slightly lower density exist to compensate for this more tightly-packed region. The bottom leaflet, which becomes more ordered over time, increases in density slightly. The top leaflet appears to be slightly less dense than the bottom leaflet as well as the control. Factoring in the presence of the protein and averaging between the two leaflets gives an overall result that the bulk density of the lipids in the peptide-membrane systems is not substantially different from that of the control (DPPC-only) systems (Figure 4.6).

4.3.2.3 Bilayer Thickness

It has been reported previously that monomeric $A\beta_{40}$ can intercalate into the hydrophobic core of reconstituted synaptic plasma membranes, resulting in a decrease in the thickness of the membrane [92]. To quantitatively assess this descriptor of membrane disruption, we measured the thickness of our simulated bilayers in terms of the P-P distance between the top and bottom leaflets of the bilayer, using GridMAT-MD (Chapter 10) [239]. Results are shown in Figure 4.7. Shown are time averages over the last 25 ns of each trajectory. The final conformation of the peptide is shown, placed at its average location over this time period. The final conformation of the peptide is representative of the last 50 ns of the trajectories, since most of the prominent secondary structure changes occurred during the first half of each simulation, described above in Section 4.2 [247].

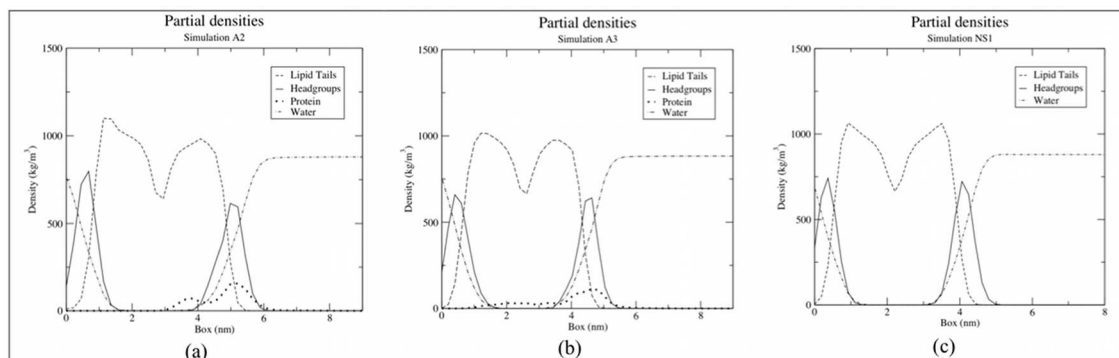


Figure 4.6: Density plots. (a) Simulation A2, wherein the peptide lost all organized secondary structure, (b) Simulation A3, in which the peptide retained much of its original α -helicity and became transmembrane, (c) Control simulation NS1.

The most striking observation overall is the amount by which the $A\beta_{40}$ peptide depresses the bilayer in its immediate vicinity, on the order of 1.0 nm. There are hydrogen bonds and favorable electrostatic interactions between the zwitterionic headgroups of the DPPC lipids and the backbone and charged residues of the peptide. The result of these interactions is that the lipid headgroups tilt substantially around the peptide, causing the acyl chains of the lipids to spread outward, more parallel to the surface of the bilayer (Figure 4.5, and discussed further below). Control simulations above the phase transition temperature (those without embedded peptides, at 323 K) show good agreement with the experimentally-determined thickness of 3.7 nm [248].

It is also observed that melittin can lead to a similar magnitude of bilayer thinning, on the order of 0.5 - 1.0 nm. This thinning only occurs in regions in which the peptide became more disordered over time. For simulations E1 and E2, these disordered segments were the N- and C-termini of the peptide, while in P1 it was the N-terminus, and in P2 the middle of the peptide became slightly disordered.

4.3.2.4 Area Per Lipid Headgroup

Experimental work has concluded that the average area per lipid headgroup for fully hydrated DPPC at 50°C is between 62 - 64 Å² [249, 250]. Previous simulations of DPPC that examined the effects of increased ionic strength have demonstrated that area per lipid headgroup decreases with increasing salt concentration, from 62.7 Å² in the absence of NaCl to 60.5 Å² in the presence of 100 mM NaCl [251]. The results from our control systems, averaged over the last 50 ns of simulation (Table 4.6), compare well with these findings. There is very little difference between DPPC systems at the original solvation ratio, and those in an asymmetric box with an increased amount of water (NS1, NS2, and NS3).

Determining the area per lipid headgroup in the presence of an irregularly-shaped protein presents a unique challenge, and we utilize the GridMAT-MD methodology, wherein each headgroup is

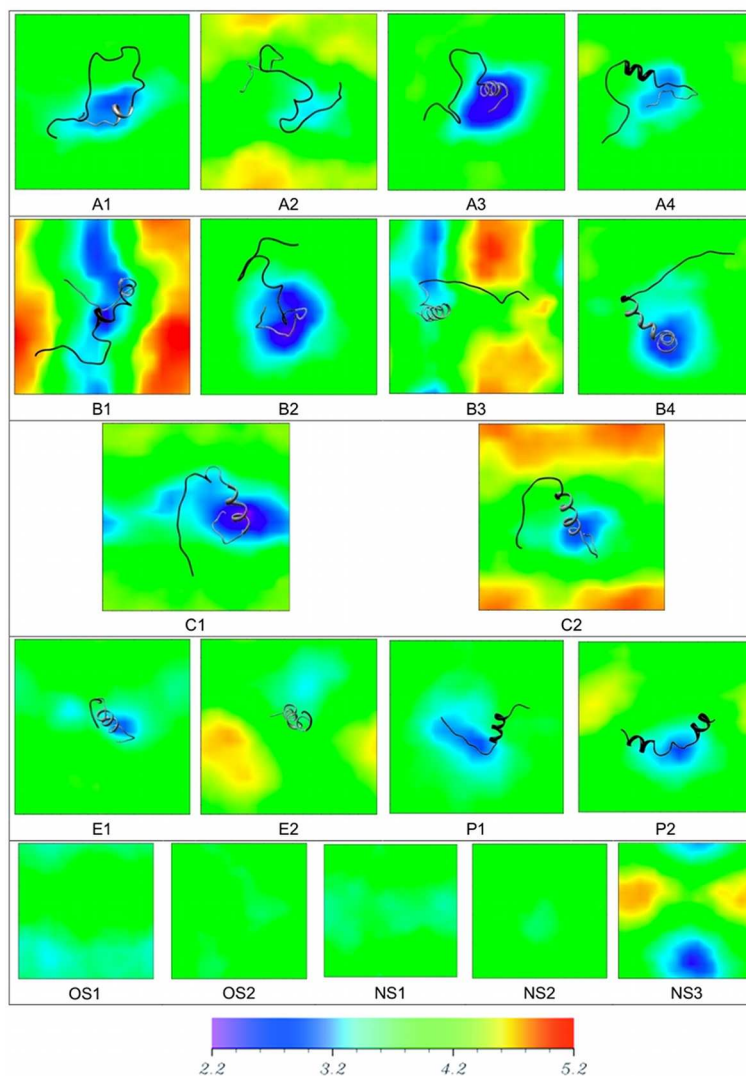


Figure 4.7: Bilayer thickness around the embedded peptides, taken from the average thickness over the last 25 ns of simulation. Peptide conformations are from the final frame of each simulation, which is representative of the final 50 ns of simulation time. For perspective, the embedded region of the peptide is colored gray, while the region exposed to the water-bilayer interface is colored black. The legend shows bilayer thickness in nm, mapped to the corresponding colors.

assigned to a polygon within the grid of the lateral bilayer surface (Chapter 10) [239]. From Table 4.6, a trend becomes clear. The area per lipid headgroup for lipids in the top leaflet is decreased substantially from the control simulations, while the area per lipid headgroup for lipids in the bottom leaflet is only slightly less than that of the controls. The reason for this result is connected with the observations regarding deuterium order parameters. The lipids of the top leaflet interact strongly with the backbone and charged amino acid sidechains of the disordered N-terminal segment of $A\beta$ via hydrogen bonds and electrostatic interactions. The result is that the lipids tilt

Table 4.6: Area per Lipid Headgroup

Simulation	Initial interfacial residue	Temperature (K)	Top Leaflet	Bottom Leaflet
A1	K28	323	52.7 \pm 1.4 (-17%)*	60.0 \pm 0.8 (-5.5%)
A2			49.0 \pm 1.7 (-17%)	54.5 \pm 1.1 (-8.1%)
A3			54.4 \pm 1.0 (-14%)	58.5 \pm 1.0 (-7.9%)
A4			49.0 \pm 0.8 (-17%)	57.4 \pm 1.0 (-3.2%)
B1	K28	300	46.8 \pm 0.6 (-20%)	55.2 \pm 0.5 (-6.0%)
B2		323	55.9 \pm 1.3 (-12%)	61.6 \pm 0.8 (-3.0%)
B3		300	48.0 \pm 0.9 (-18%)	54.1 \pm 0.8 (-7.8%)
B4		323	52.9 \pm 1.4 (-17%)	60.4 \pm 0.8 (-4.9%)
C1	V24	323	53.6 \pm 1.4 (-16%)	61.7 \pm 1.6 (-2.8%)
C2			49.7 \pm 1.5 (-16%)	55.9 \pm 1.4 (-5.7%)
OS1	N/A [†]	323		63.3 \pm 0.7
OS2		323		60.5 \pm 1.0
NS1		323		63.5 \pm 1.1
NS2		323		59.3 \pm 0.7
NS3		300		58.7 \pm 0.7
E1		W19	323	58.7 \pm 1.1 (-7.6%)
E2			56.0 \pm 0.8 (-5.6%)	55.4 \pm 1.3 (-6.6%)
P1	N/A		62.5 \pm 1.8 (-1.6%)	60.7 \pm 0.9 (-4.4%)
P2			57.1 \pm 1.1 (-3.7%)	55.7 \pm 0.8 (-6.0%)

*Mean \pm standard deviation, in \AA^2 , with % difference from controls shown in parenthesis. Averages taken over the last 50 ns of each trajectory.

[†]Not applicable. In the case of the OS and NS series, no peptide was present. For P1/P2, the entire peptide was initially located at the membrane-water interface.

substantially, reducing the vertical thickness of the top leaflet. This behavior requires the lipids of the lower leaflet to pack more tightly and extend their acyl chains to maintain the integrity of the membrane. The substantial tilt (and resulting disorder) of the top leaflet lipids and the slight increase in packing (and thus order) in the bottom leaflet lipids is reflected in the area per lipid headgroup.

The interaction between the N-terminal segment of A β and the DPPC headgroups develops over time. Following equilibration, the measured area per lipid headgroup in each system is close to the accepted experimental value (62 - 64 \AA^2). Upon contact between the N-terminal region of A β and the membrane-water interface (within 10 ns of simulated time), the area per lipid headgroup begins to rapidly decrease as the lipids associate with this disordered segment of the peptide (Figure 4.8 - 4.11). Unfolding of A β occurs over the first 50 ns of each simulation, after which the peptide conformation is largely unchanged [247]. Area per lipid headgroup for the control systems

(simulation sets OS and NS) remains steady over time at values appropriate for a fully hydrated DPPC bilayer under the given conditions (Figure 4.12).

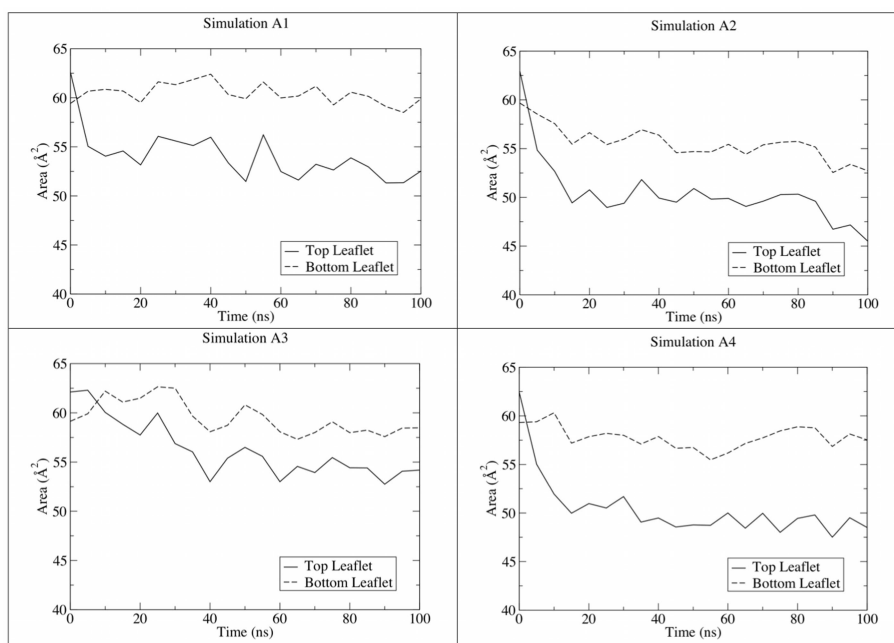


Figure 4.8: Area per lipid headgroup as a function of time, for simulation set A. After making contact with the DPPC headgroups (within 10 ns in all cases), the N-terminal segment of $A\beta$ attracts the lipids of the top leaflet, depressing their lateral area. The area per lipid headgroup in the bottom leaflet is decreased due to the increased order and packing in this leaflet, which is a consequence of the disordering of the top leaflet.

The lipids closest to $A\beta_{40}$ experience the greatest decrease in area per lipid headgroup. From Figure 4.13, it can be seen that lipids closest to the peptide have the smallest lateral area, while lipids further away tend to occupy areas close to the bulk value of DPPC. In Figure 4.13, lipids of the top leaflet were ordered according to their proximity to the center of mass of the $A\beta$ peptide. Thus, the closer lipids have the smaller residue designation. In the case of simulation A1, the area per lipid headgroup is largely constant at the outset of the simulation, fluctuating around a value of 62 \AA^2 . Over time, the lipids close to the peptide are drawn to it by the interactions described above, while more distant lipids maintain a more canonical value for their lateral area. This trend is apparent in all other simulations of $A\beta$ (Figures 4.15 - 4.18), except for A2. In simulation A2, the peptide unfolded to the greatest extent of any of the simulations, thus contacting the greatest number of lipids. The lipids closest to the peptide center of mass have a depressed value for their lateral area, as do some lipids that are more distant from this point. The explanation for this behavior is that, as the disordered N-terminal segment elongates through the lipid headgroups, it interacts with a greater number of lipids than in any other simulation. In Figure 4.13, lipids numbered from 41 - 49 are close to the peptide center of mass and are attracted by contact with the polar backbone, lipid residues 50 - 60 make van der Waals contacts with the amino acid side

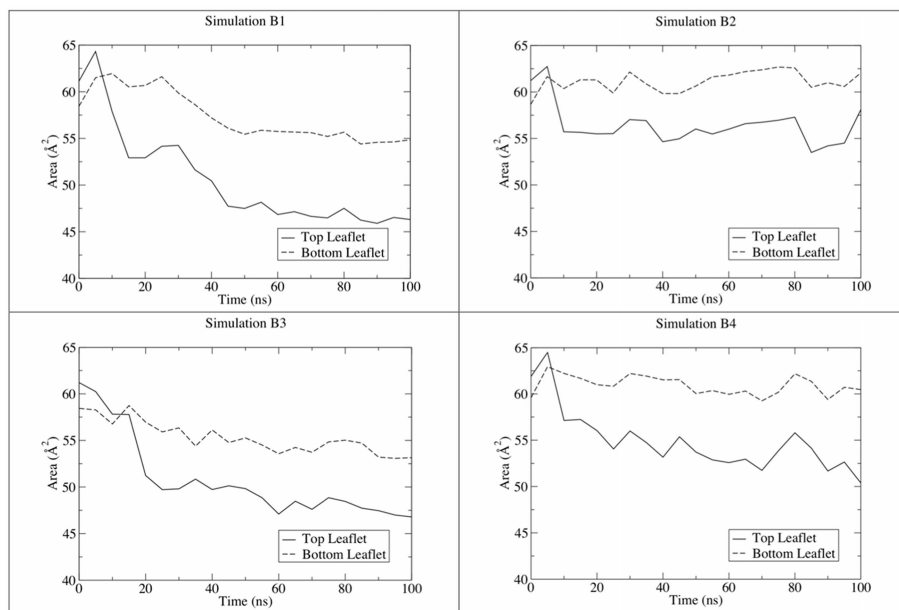


Figure 4.9: Area per lipid headgroup as a function of time, for simulation set B.

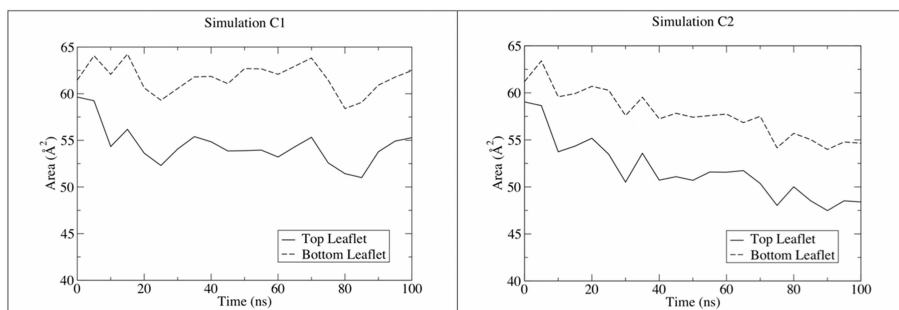


Figure 4.10: Area per lipid headgroup as a function of time, for simulation set C.

chains of $A\beta$, and lipid residues 60 - 70 interact strongly with the highly charged and disordered N-terminal segment of $A\beta$. This behavior is illustrated in Figure 4.14.

In all cases, area per lipid headgroup in the bottom leaflet was largely insensitive to proximity to the peptide, even in the simulations wherein the C-terminus of $A\beta$ interacted with the lipids of the lower leaflet (A3, C1, and C2). This observation indicates that the ability of $A\beta$ to condense nearby lipids lies primarily in its highly-charged, unstructured N-terminal segment.

Simulations of melittin showed similar behavior. Simulations E1 and E2 showed a slight decrease in area per lipid in the vicinity of the peptide (Figure 4.19). Since melittin largely maintains its secondary structure over time, the effects of the peptide on this parameter are less pronounced than in the case of $A\beta$. In simulations P1 and P2, wherein the entire peptide was in contact with the DPPC headgroups, the nearest lipids experienced a reduction in their lateral area, which we

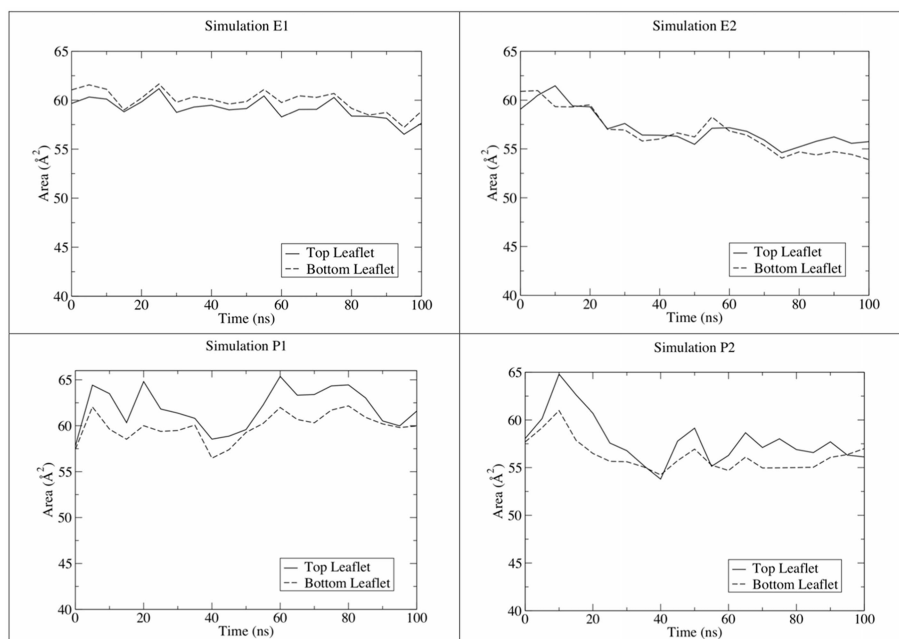


Figure 4.11: Area per lipid headgroup as a function of time, for simulation sets E and P.

attribute to hydrogen bonding between charged headgroup phosphates and the backbone of the small section of the peptide that became disordered over time (Figures 4.7 and 4.20).

4.3.2.5 Lipid Tilt and Effective Chain Length

The attraction between the lipids and unfolded regions of the $A\beta$ peptide described in previous sections gives rise to striking behavior of the lipid acyl chains. As mentioned above, the acyl chains of lipids near the peptide tilt substantially, increasing their disorder as the peptide draws them close to itself. To quantify this observation, two related parameters were measured - acyl chain tilt angle and effective chain length. We defined the acyl chain tilt angle as the angle formed between the bilayer normal and the vector defined by the first methylene carbon and the terminal methyl carbon on the acyl chain. A description of effective chain length has been proposed by Petrache *et al.* [252] (therein called “average chain length,” LC^*). This descriptor is simply defined as the distance along the bilayer normal between the first methylene carbon and the terminal methyl carbon. These two parameters, the tilt angle and the effective chain length, should be related under most circumstances, such that as the tilt angle increases (and the acyl chain becomes more parallel to the bilayer surface), the effective chain length should decrease.

Tilt angle and effective chain length have been analyzed for the systems simulated here as a function of distance from the peptide. There was no substantial difference in the results for the *sn*-1 and *sn*-2 chains, hence data presented and discussed here will be in direct reference to the *sn*-1 chain for the purpose of clarity. We find that the lipids in the top leaflet in closest contact with the

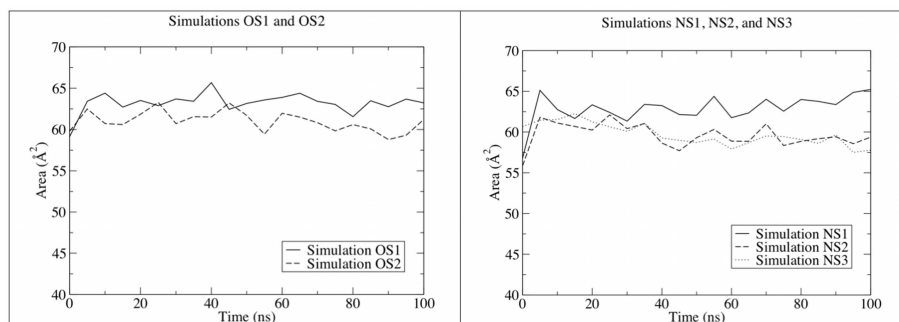


Figure 4.12: Area per lipid headgroup as a function of time, for control DPPC simulations.

peptide (typically those interacting with the disordered N-terminal segment) increase their tilt angle over time, while simultaneously decreasing their effective chain length (Figure 4.21). In other words, the strong attraction between the peptide backbone and charged residues draws the headgroups of nearby lipids away from other surrounding lipids, pulling the entire lipid more parallel to the surface of the bilayer. Regions of the most substantially tilted lipids correspond to those with the smallest area per lipid headgroup and greatest amount of disorder. In the bottom leaflet, lipid chains elongate, demonstrated by a small increase in effective chain length over time, as well as an overall reduction in tilt angle (Figure 4.21).

Lipid tilting was minimal in simulations involving melittin, which interacts more weakly with the lipid headgroups due to its smaller size and greater retention of secondary structure. These results demonstrate that more extensive lipid tilting is induced by the dynamic behavior of $A\beta$. Since $A\beta$ unfolds to a much greater extent than melittin, interacting with more lipids, it is able to cause greater disruption of canonical lipid dynamics and orientation. Control simulations of pure DPPC showed an effective chain length of approximately 1.25 nm in both leaflets, in agreement with the value proposed by Petrache *et al.* [252].

4.3.3 Discussion

According to the “amyloid hypothesis,” $A\beta$ is central to the development and progression of Alzheimer’s disease [224], but relatively little is yet known about how this small peptide interacts with a lipid membrane in the context of neurodegeneration. While a number of experiments have examined the properties of lipids in the presence of $A\beta$ [84, 90, 92], little detailed structural data exist to indicate how $A\beta$ is inducing these observed phenomena, providing the motivation for the present work. In order to gain a clear picture of how $A\beta$ disrupts a membrane environment, the peptide must be examined in detail with respect to its structural and chemical features and how they impact the surrounding lipid matrix. Previous work by Ambroggio *et al.* has suggested that $A\beta_{42}$ interacts strongly with a lipid environment, becoming part of it and disrupting interactions between the lipids, leading to structural deformation [90]. Although the focus of that work was $A\beta_{42}$, and we chose to investigate $A\beta_{40}$ in the present study, we believe there to be common features of both

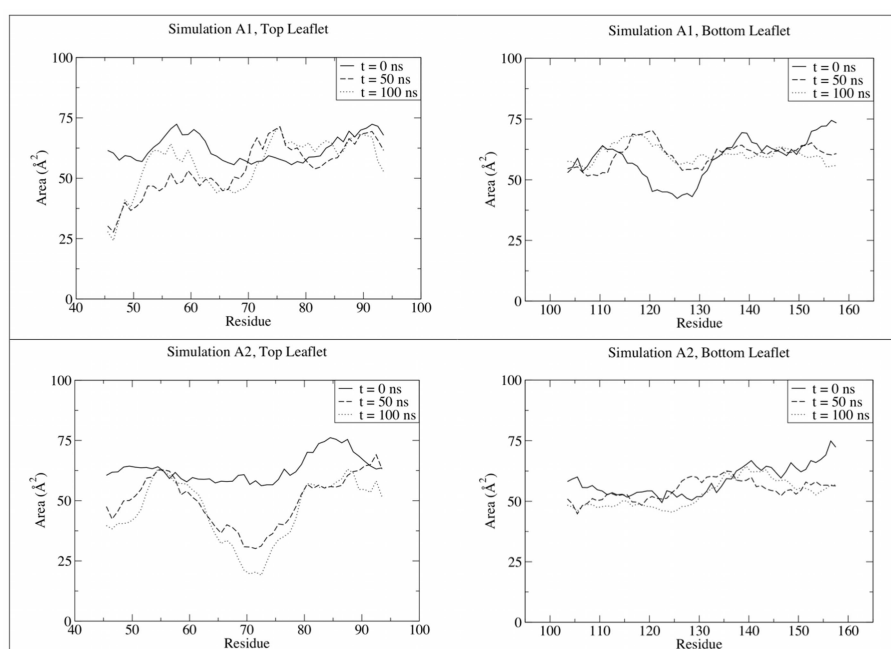


Figure 4.13: Area per lipid headgroup as a function of distance from the protein; shown are simulations A1 and A2 at each of three time points (0, 50, and 100 ns). Lipid residues are numbered such that those closest to the peptide have the lowest numbering, increasing as the lipids are further away from the peptide. For clarity, running averages of the data are shown, using a window of 10 data points.

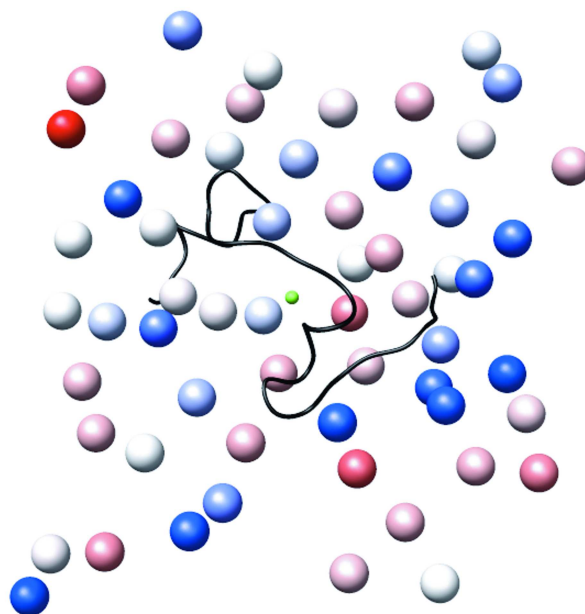


Figure 4.14: Illustration of the contacts between the $A\beta_{40}$ peptide in simulation A2 and the lipids of the top leaflet. The peptide is shown in a black ribbon, and each lipid is represented by the phosphorus of its headgroup, shown as spheres. The phosphorus atoms are colored according to the lateral area of the corresponding lipid, increasing as the colors change from blue to red. The small green sphere represents the peptide center of mass, illustrating that not all of the lipids closest to this point experience the greatest degree of association with the peptide in this simulation.

alloforms of the peptide that contribute to membrane perturbation.

The most substantial observations we present in this work are that the $A\beta_{40}$ peptide causes nearby DPPC lipids to become more tilted and disordered relative to controls, that the peptide is capable of reducing the area per lipid headgroup of the lipids with which it most directly interacts, and that it is capable of reducing the thickness of the membrane in its immediate vicinity. Taken as a whole, these data suggest interesting roles for the region of the peptide that is present in the extracellular environment, and that which remains embedded in the bilayer.

4.3.3.1 Lipid Tilt and Effective Chain Length

The factor that gives rise to much of the behavior discussed in this work is the tilting of lipids that are closest to the peptide present in each simulation. The $A\beta$ peptide is capable of drawing lipid headgroups to itself through electrostatic and hydrogen bonding interactions, weakening the interactions between these lipids and others that are more distant from $A\beta$. Nearby acyl chains tilt substantially over time (Figure 4.21), leading to a slight thinning of the hydrophobic core of the bilayer, manifested in a reduction in effective chain length of these same lipids. Lipids

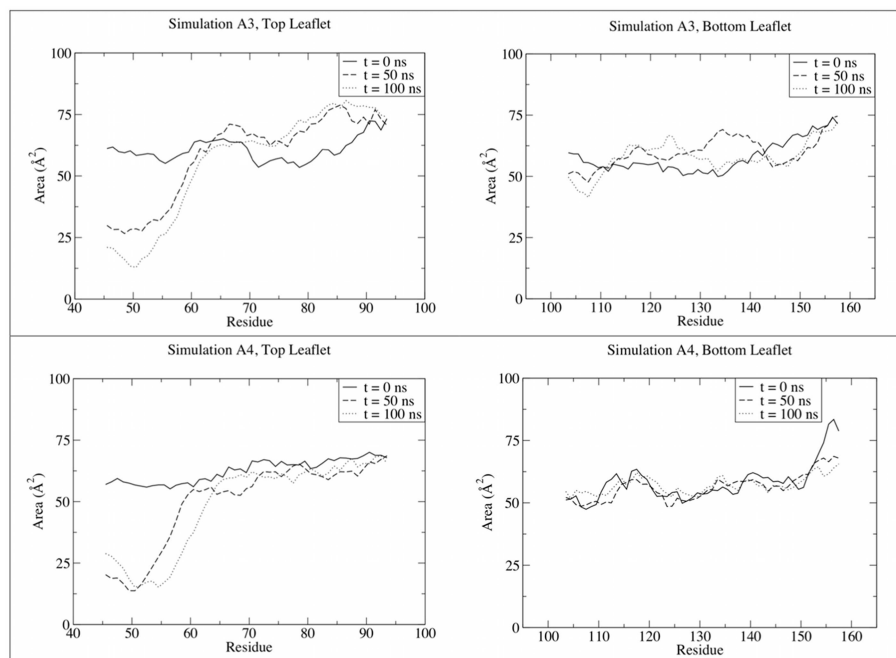


Figure 4.15: Area per lipid headgroup as a function of distance from the protein; shown are simulations A3 and A4 at each of three time points (0, 50, and 100 ns). Lipid residues are numbered such that those closest to the peptide have the lowest numbering, increasing as the lipids are further away from the peptide. For clarity, running averages of the data are shown, using a window of 10 data points.

that displayed the greatest degree of tilting also correspond to those with the smallest area per lipid headgroup and greatest amount of disorder. We attribute these observations to the ability of the $A\beta_{40}$ peptide, especially through its N-terminal disordered region, to bind lipid headgroups very closely to itself and draw them closely to each other. A more detailed description of this phenomenon follows in subsequent sections, detailing the effects of lipid tilting on each of the other parameters measured in this work.

4.3.3.2 Area Per Lipid Headgroup

The work described in Section 4.2 indicated that $A\beta_{40}$ interacts strongly with the water-bilayer interface region and became disordered in almost all simulations. We attributed this behavior to strong hydrogen bonding and electrostatic interactions between the peptide and the zwitterionic phosphatidylcholine headgroups. This behavior creates an interesting contrast with melittin, another membrane-disrupting peptide. When embedded, melittin still has a small portion of its structure exposed to the extracellular environment, but it is substantially shorter than the extracellular segment of $A\beta_{40}$, and it retains more organized secondary structure than $A\beta_{40}$ in the simulations we conducted. It can be seen from Table 4.6 that the effect of melittin on the area per lipid headgroup

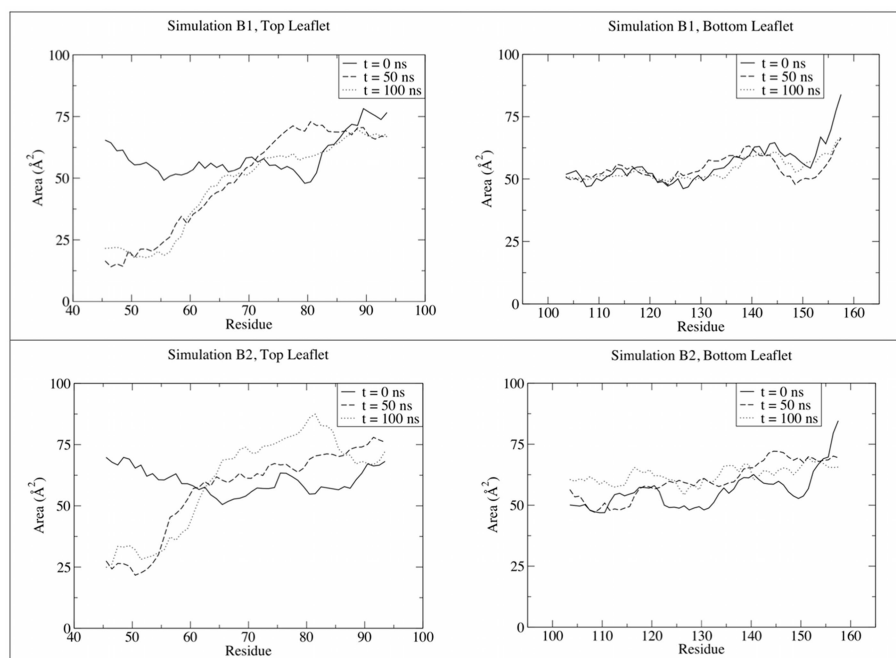


Figure 4.16: Area per lipid headgroup as a function of distance from the protein; shown are simulations B1 and B2 at each of three time points (0, 50, and 100 ns). Lipid residues are numbered such that those closest to the peptide have the lowest numbering, increasing as the lipids are further away from the peptide. For clarity, running averages of the data are shown, using a window of 10 data points.

of the individual bilayer leaflets is minimal, even when the whole peptide is positioned roughly parallel to the water-bilayer interface. There is some reduction in the area per lipid headgroup by melittin, but the difference in area per lipid headgroup between the leaflets is nearly indistinguishable. However, in the case of $A\beta_{40}$, there is often a prominent difference in the area per lipid headgroup between the leaflets, with the leaflet interacting with the extracellular (N-terminal) region of $A\beta_{40}$ experiencing a substantially reduced area per lipid headgroup relative to the intracellular leaflet. This fact is true even in the case of simulations A3, C1, and C2, in which the peptide adopted a transmembrane orientation, interacting with the lipid headgroups in the bottom leaflet, as well.

We propose that this behavior arises because of different interactions between the peptides and the lipid headgroups. In the case of $A\beta_{40}$, the peptide is capable of attracting lipid headgroups very close to its long, mostly disordered, N-terminal segment, tilting the lipids and arranging them very closely to each other. This finding is independent of protonation state and ionic strength of the surrounding aqueous medium, suggesting that electrostatic interactions and hydrogen bonding are likely involved in drawing lipid headgroups in, but that these interactions are non-specific. In other words, they are not sensitive to the protonation state of any particular residue or group of residues. They occur simply because there are many charged and polar amino acids in the extracellular

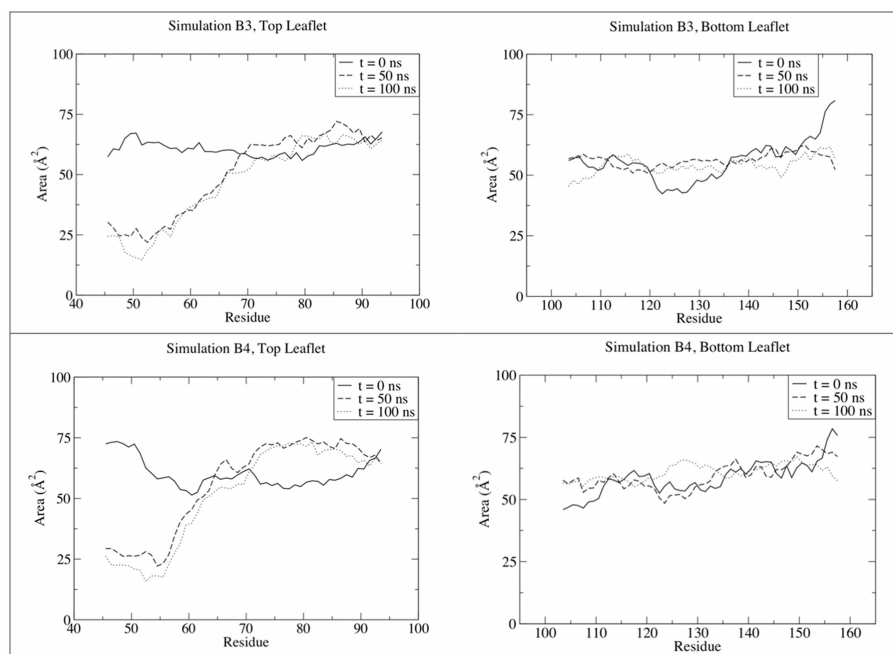


Figure 4.17: Area per lipid headgroup as a function of distance from the protein; shown are simulations B3 and B4 at each of three time points (0, 50, and 100 ns). Lipid residues are numbered such that those closest to the peptide have the lowest numbering, increasing as the lipids are further away from the peptide. For clarity, running averages of the data are shown, using a window of 10 data points.

region of $A\beta_{40}$, which interacts with the elements of the water-bilayer interface. While some lipids become oriented such that their headgroups interact with the C-terminal region of melittin, they tend to remain more dispersed compared to the lipids in the $A\beta_{40}$ simulations. That is, there are fewer lipids tightly associated with melittin than there are in the case of $A\beta_{40}$.

4.3.3.3 Disordering of Nearby Lipids

In our simulations, both $A\beta_{40}$ and melittin demonstrated the ability to cause disorder in nearby lipids. In the case of $A\beta_{40}$, we believe this disorder to be due to primarily two factors: (i) reduction in area per lipid headgroup, and (ii) the unfolding of the C-terminal, embedded, region of the peptide. As discussed above, $A\beta_{40}$ is capable of increasing the tilt of the surrounding lipids, thus disordering their acyl chains. The fact that melittin also causes a similar amount of disorder on the surrounding lipids leads to an interesting question. If, as we have proposed, melittin does not interact as strongly with the lipid headgroups as does $A\beta$, how does one account for the fact that both of these peptides can disorder the surrounding lipids to the same extent?

One possible answer to this question has been proposed by the authors of the original melittin

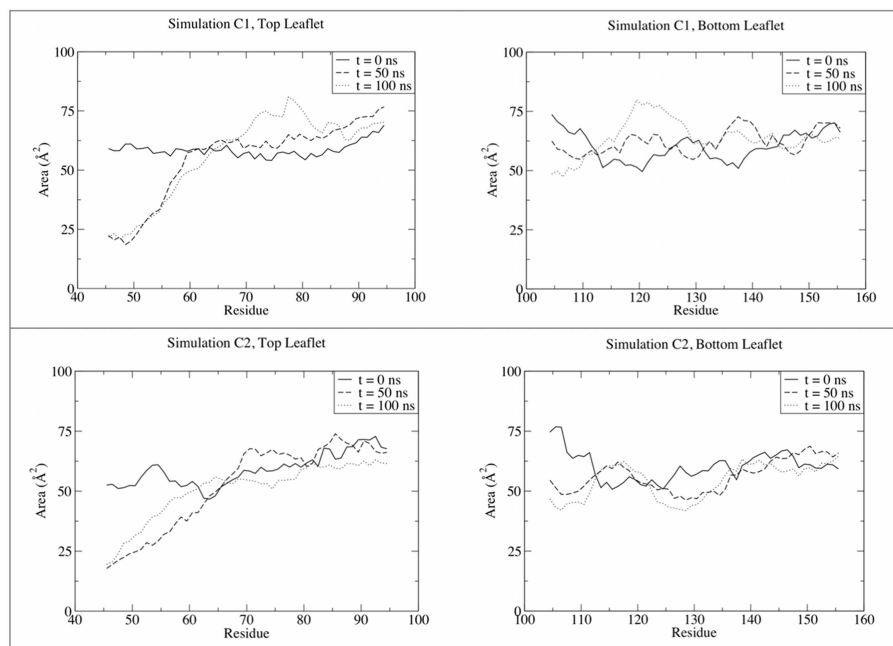


Figure 4.18: Area per lipid headgroup as a function of distance from the protein; shown are simulations C1 and C2 at each of three time points (0, 50, and 100 ns). Lipid residues are numbered such that those closest to the peptide have the lowest numbering, increasing as the lipids are further away from the peptide. For clarity, running averages of the data are shown, using a window of 10 data points.

simulations in which the peptide was embedded in the membrane [227]. They observed that the surrounding lipids tend to pack around the peptide and tilt their acyl chains, as the protein itself tilts. Another contribution to the disordering effect is the restriction of motion along the acyl chains once the lipids have packed around the protein. Since this behavior arises because of the tilt of the peptide and interactions between the embedded region of the peptide and the acyl chains, it presents an interesting insight into the interactions between asymmetrically-embedded peptides and nearby lipids. Much like melittin, $A\beta_{40}$ is embedded asymmetrically in the lipid bilayer, and over time, tilts with respect to the bilayer normal. As such, we attribute some of the disorder experienced by the nearby lipids to the motion and tilting of this embedded segment.

4.3.3.4 Thinning of the Bilayer in the Presence of $A\beta_{40}$

Since $A\beta_{40}$ is known to disrupt the integrity of the lipid membrane [84, 90], another parameter of interest is the local thickness of the bilayer. In Section 4.2.2.2, the ability of water to penetrate into the bilayer when $A\beta_{40}$ is present was discussed, and a more thorough examination of the bilayer thickness is appropriate here in the context of lipid parameters.

The experimentally-determined thickness of a fluid-phase DPPC bilayer (in terms of the P-P dis-

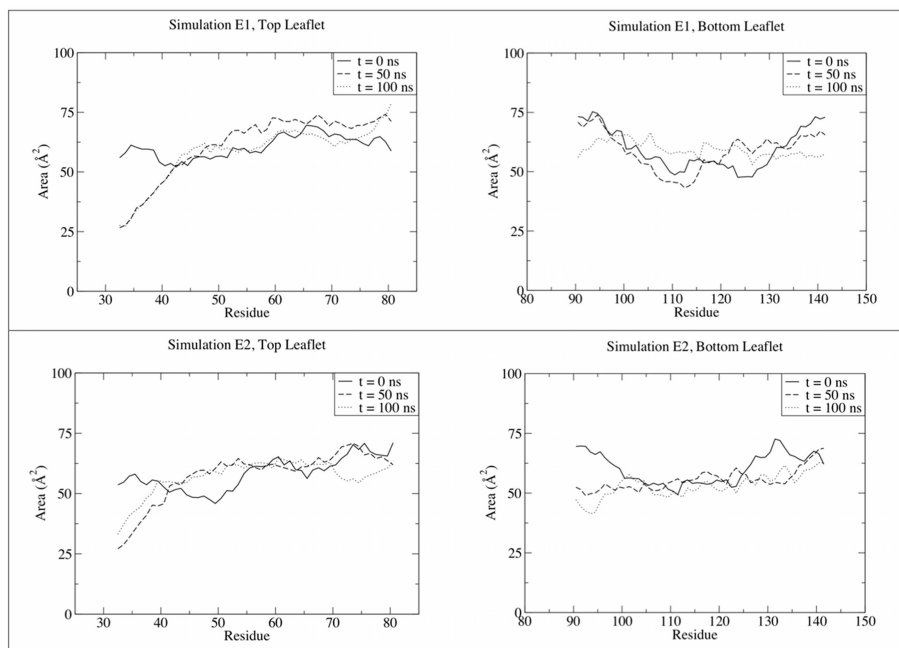


Figure 4.19: Area per lipid headgroup as a function of distance from the protein; shown are simulations E1 and E2 at each of three time points (0, 50, and 100 ns). Lipid residues are numbered such that those closest to the peptide have the lowest numbering, increasing as the lipids are further away from the peptide. For clarity, running averages of the data are shown, using a window of 10 data points.

tance), is 3.7 nm [248]. We achieve good agreement with this value in all of our control simulations conducted at 323 K (Figure 4.7). In the presence of $A\beta_{40}$ (at 323 K), however, the bilayer may become depressed between 1.0 - 1.5 nm, as determined by averaging the bilayer dimensions over the last 25 ns of each simulation. As with reduced area per lipid headgroup, decreased bilayer thickness is independent of peptide protonation state. Local thinning of the bilayer, of comparable magnitude, is observed in all of the $A\beta_{40}$ -DPPC simulations.

Similar thinning of the bilayer also occurs in the presence of melittin. While the bilayer may deform to decrease its thickness up to 1.0 nm (also determined by the same averaging discussed above), the deformed region is much smaller than in the case of $A\beta_{40}$. We attribute this observation to the fact that, unlike $A\beta_{40}$, melittin retains much of its secondary structure throughout the trajectory, thus remaining more compact. The unfolding of $A\beta_{40}$ makes it more accessible to a wider area, and thus more lipids, causing a more pronounced depression in the bilayer.

Hydrophobic mismatch likely plays a part in the local deformation of the bilayer. Systematic analysis of hydrophobic mismatch using KALP model peptides by Kandasamy and Larson illustrates that short, helical peptides, with charges placed within the lipid headgroups, can cause depression of bilayer thickness or tilting of the peptide to accommodate the size of the peptide [229]. The systems we present here also include elements of hydrophobic mismatch, since neither $A\beta_{40}$ nor

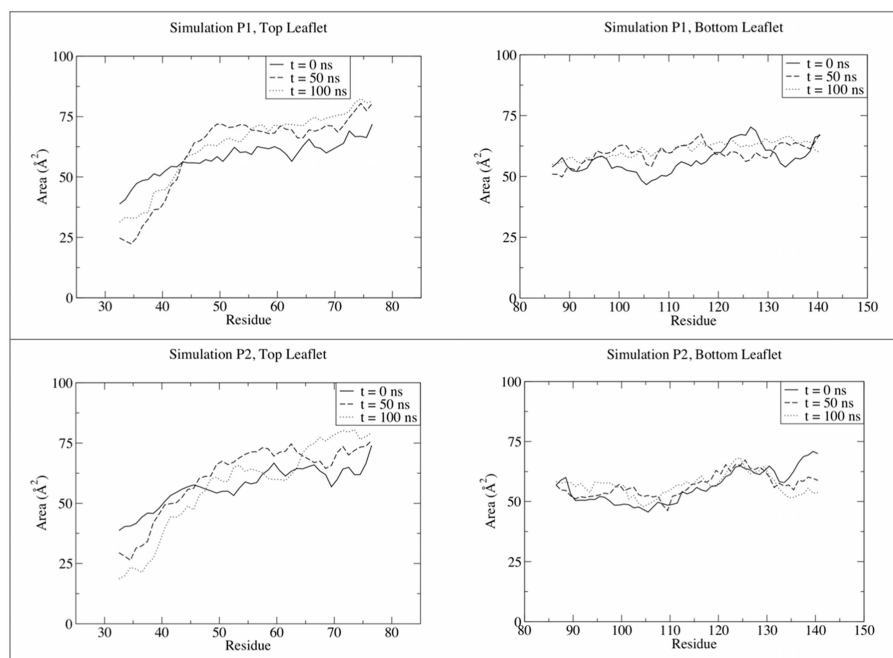


Figure 4.20: Area per lipid headgroup as a function of distance from the protein; shown are simulations P1 and P2 at each of three time points (0, 50, and 100 ns). Lipid residues are numbered such that those closest to the peptide have the lowest numbering, increasing as the lipids are further away from the peptide. For clarity, running averages of the data are shown, using a window of 10 data points.

melittin completely spans the bilayer in their initial configurations. This orientation positions the hydrophobic, embedded, regions of the peptide approximately halfway through the bilayer. Thus, it is not unreasonable to conclude that the reason the bilayer deforms is to accommodate this orientation. Even as the simulations progress, and the peptides alter their orientation, becoming more deeply embedded in some cases, the length of the short hydrophobic stretch remains, and is less than the dimensions of the hydrophobic core of the bilayer. This mismatch, anchored at the water-bilayer interface in many cases by charged amino acids such as Lys16 and Lys28 in $A\beta_{40}$, likely contributes to the deformation of the dimensions of the surrounding bilayer.

The observations we make from our simulations compare well with experimental observations. Early work by Mason *et al.* indicated that $A\beta_{40}$ was capable of penetrating into rat synaptic plasma membranes and decreasing bilayer thickness [92]. In addition, Kaye *et al.* report that, while oligomeric $A\beta$ species are primarily responsible for membrane permeability, monomeric and low molecular weight species can penetrate into the bilayer interior and cause thinning of 0.5 nm [84]. While the pathogenic agent of Alzheimer's disease is widely believed to be oligomeric $A\beta$ species, it is also important to understand the interactions of monomeric $A\beta$ with lipid bilayers. The results we have presented here potentially shed light on the molecular interactions that give rise to the experimentally-observed behaviors described above, including a new hypoth-

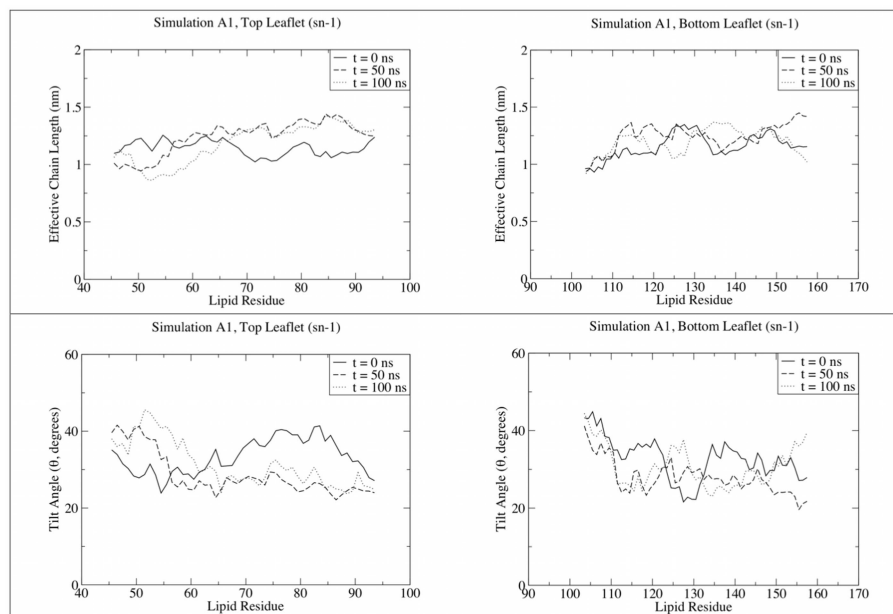


Figure 4.21: Effective chain length (top panel) and acyl chain tilt angle (bottom panel), as a function of distance from the protein, for simulation A1. The results are indicative of all the other simulations involving $A\beta_{40}$, and thus is a representation of the general conclusions discussed in the text. There is no substantial difference in the characteristics of the *sn-1* and *sn-2* chains, thus results for only the *sn-1* chain are presented here for clarity.

esis for a functional role of the N-terminal 16 residues of $A\beta$. While it has long been postulated that $A\beta$ is partially embedded in the hydrophobic core of the bilayer via its C-terminus, a functional role for its N-terminus has not yet been established. Our simulations suggest an important role for the N-terminus in associating with the surrounding lipid matrix. The preponderance of charged amino acid side chains and exposure of the disordered backbone interact favorably with the strongly polar environment of the membrane-water interface. It is not surprising that this region of the peptide may contribute to neurotoxicity. The N-terminal sequence is cleaved by α -secretase to generate the non-amyloidogenic sAPP α and p3 fragments from APP, precluding the production of $A\beta$ [252, 253]. The p3 peptide and $A\beta$ differ by only the absence or presence of these 16 N-terminal residues, and p3 is non-toxic whereas $A\beta$ is neurotoxic. The toxicity of $A\beta$ has been proposed to be exerted through its interactions with membranes [88], suggesting that this sequence of amino acids likely plays an important role in $A\beta$ -membrane interactions. We present here the molecular basis for the perturbation of lipids, due to both hydrophobic mismatch of the C-terminal region of $A\beta$ within the lipid bilayer and the favorable interaction of the N-terminal segment of $A\beta$ with the polar environment of the lipid headgroups. Understanding this simple system, monomeric $A\beta$ in a model membrane, provides an excellent starting point for more complex lipid systems and multiple $A\beta$ peptides in membrane systems. Such systems are explored in Chapters 5 - 6.

Chapter 5

Interactions Between Monomeric A β and Model Membranes and Lipid Rafts

Copyright statement: The contents of this chapter are reprinted with permission from J. A. Lemkul and D. R. Bevan (2011) “Lipid composition influences the release of Alzheimer’s amyloid β -peptide from membranes.” *Protein Sci.* **20** (9): 1530-1545. <http://dx.doi.org/10.1002/pro.678>

Attribution: J. A. L. and D. R. B. designed the research. J. A. L. conducted the simulations and performed analysis. J. A. L. and D. R. B. wrote the paper.

5.1 Introduction

Since the γ -secretase complex is localized in lipid rafts [29], which are liquid-ordered membrane microdomains enriched in cholesterol and sphingolipids, it stands to reason that it is this environment to which A β is exposed upon its production. The composition of these lipid rafts likely influences the structure and position of A β in the membrane, factors that may affect the release of A β from the membrane and the formation of toxic, low-molecular weight oligomers. Although the facile release of A β from membranes is often depicted schematically in the literature, little attention has been paid to the actual mechanism by which A β enters the extracellular space. A thorough characterization of A β -membrane interactions is necessary, and understanding the behavior of A β in the membrane at the atomic level may provide insight into the earliest events of Alzheimer’s disease.

To characterize A β -membrane interactions in a variety of systems, we explored a number of model membranes that contain many common membrane phospholipids, including systems containing single phospholipid components (POPC or POPS), a mixture of two lipid components (POPC/POPE), a model raft containing palmitoylsphingomyelin (PSM), cholesterol, and POPC, and a raft that included a physiologically-relevant concentration of ganglioside GM1. Sphin-

gomyelin (SM) and lipids containing phosphatidylcholine (PC), phosphatidylethanolamine (PE), and phosphatidylserine (PS) headgroups are among the most common components of eukaryotic cell membranes [254, 255]. Thus, we examined the interactions of A β_{40} with these lipids and combinations thereof to establish a thorough biophysical characterization of A β dynamics in the plasma membrane and to understand if any of these lipids had unique characteristics that were relevant to the release of the peptide from the membrane.

Ganglioside GM1 (Figure 5.1) is a component of lipid rafts and is enriched in the outer leaflet of the plasma membrane of cells in the central nervous system [256]. It has been shown that A β -GM1 complexes may act as seeds for further A β aggregation [235, 257, 258], indicating that interactions between A β and GM1 are important to the progression of Alzheimer's disease. A β -GM1 complexes form most readily in lipid raft environments that are enriched with cholesterol [258, 259]. Further, GM1 has been found to enhance the formation of β -strand structures in A β , thus promoting aggregation [235, 259]. While numerous studies have concluded that A β binds to the oligosaccharide headgroup of GM1 [260, 261, 262], the influence of GM1 on the ability of A β to exit the membrane is thus far unexplored, and while numerous experiments have been conducted to characterize the interactions between A β and GM1, the existing results lack the detail that can be obtained from molecular dynamics (MD) simulations.

5.2 Methods

5.2.1 System Construction

We conducted MD simulations of A β_{40} in model membranes consisting of pure POPC and POPS, and an equimolar mixture of POPC and POPE. Coordinates and topologies for pre-equilibrated POPC [228] and POPS [263] membranes (128 lipids each) were obtained from D. P. Tieleman's site [219] in the case of POPC, and by personal communication in the case of POPS. Coordinates for the binary POPC-POPE system (288 total lipids, 144 each of POPC and POPE) were taken from work by Leekumjorn *et al.* [264], with lipid topologies also from Tieleman's work [228, 244]. Lipid raft membranes contained POPC, PSM, and cholesterol, with and without 6.75 mol% GM1, which represents a physiologically-relevant concentration of GM1 [256]. The coordinates and topologies for the model lipid rafts were taken from work by Niemelä *et al.* [265]. Topologies for cholesterol and PSM were updated to replace GROMOS87 [181] atom types and charges with those of equivalent functional groups in the GROMOS96 53A6 force field (Appendices A and B) [174]. The topology for GM1 was created using nonbonded parameters derived by Berger *et al.* [266] for the acyl chain portion and those of the GROMOS96 53A6 force field for all other functional groups. The GROMOS96 53A6 force field was applied to the remaining components of the system. The topology for GM1 is given in Appendix C. From the initial raft configuration [265], a subset of lipids containing 222 lipids (74 each of POPC, PSM, and cholesterol) was extracted. Such a system is of sufficient size to accommodate the A β_{40} peptide and avoid spurious interactions across periodic boundaries during the MD simulations. To create a raft system con-

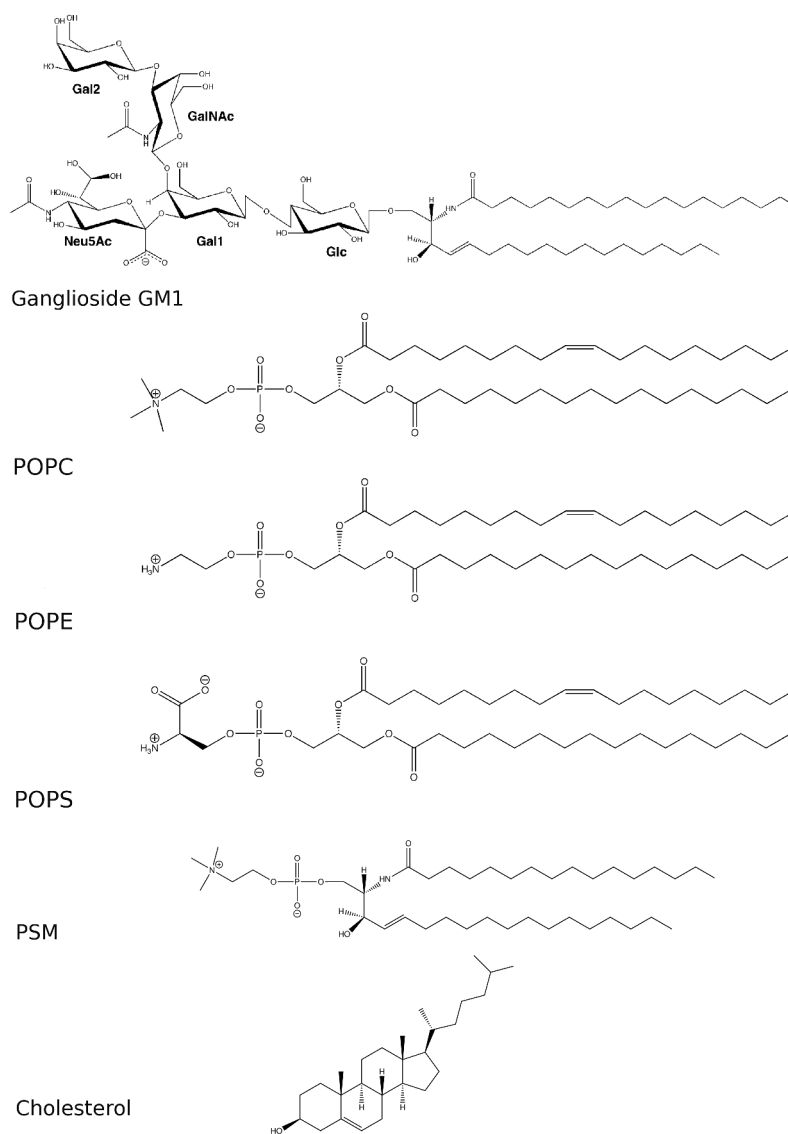


Figure 5.1: The structures of the lipids used in this work. For ganglioside GM1, the constituent sugar moieties are labeled as referred to in the main text: glucose (Glc), internal galactose (Gal1), N-acetylneuraminic acid (Neu5Ac), N-acetyl galactosamine (GalNac), and terminal galactose (Gal2).

taining ganglioside GM1, 15 PSM molecules in the extracellular leaflet of the raft were replaced with GM1 to give a total GM1 content of 6.75 mol% to reproduce the experimentally-determined 6.59% ganglioside content in lipid rafts *in vivo* [256] as accurately as possible. The rafts (both with and without GM1) were energy-minimized and equilibrated for 20 ns prior to inserting the $A\beta_{40}$ peptide. Residues 29-40 of $A\beta_{40}$ were aligned with the membrane normal, with Lys28 coincident with the average phosphate plane, in agreement with experimental evidence of this location [77].

Lipids were packed around the peptide using the InflateGRO method [267], modified in-house to accommodate multiple lipid types (in the case of raft and POPC/POPE systems). The remainder of the unit cell was filled with SPC water [268] and approximately 150 mM NaCl (including Na⁺ counterions).

The model of A β_{40} used here was taken from PDB entry 1BA4 [77], using the same configuration from the NMR ensemble as in our earlier work [247, 269]. This structure was determined in the presence of sodium dodecyl (SDS) micelles, and thus is assumed to be a representative configuration for A β_{40} in the presence of a membrane. The transmembrane domain of APP and the polypeptide sequence that encompasses the A β sequence is believed to be principally helical with disordered solvent-exposed regions [22]. For these reasons we believe this starting structure to be applicable to modeling A β_{40} inserted in a membrane. For reasons described previously [247], both neutral (model “CH”) and ionized (model “CI”) C-termini were considered. Given that the γ -secretase complex has a solvent-accessible active site [28], both protonation states are plausible and potentially relevant to the dynamics of A β . Other titratable groups were assigned their typical protonation states at pH 7.4. The net charge on the CH peptide was -2 , while that on the CI peptide was -3 .

5.2.2 Molecular Dynamics Simulations

All simulations were performed with GROMACS, version 4.0.7 [180]. Periodic boundary conditions were applied in all directions. The real-space contribution to the Coulombic potential was truncated at 1.2 nm, and long-range electrostatics were calculated using the smooth particle mesh Ewald method [193, 194], using cubic-spline interpolation and a Fourier grid spacing of 0.12 nm. Van der Waals interactions were truncated at 1.2 nm, and dispersion correction was applied to the energy and pressure terms. All bond lengths were constrained using the P-LINCS algorithm [270], allowing an integration timestep of 2 fs. Coordinates and energies were saved every 10 ps for analysis. Five independent simulations were produced for each raft system (Raft-CH, GM1-CH, Raft-CI, and GM1-CI) by assigning different random velocities from a Maxwell distribution at the outset of equilibration. Three independent simulations were generated for all other systems (PC-CH, PC-CI, PS-CH, PS-CI, PC/PE-CH, and PC/PE-CI).

Simulation systems were equilibrated in three phases, all of which employed position restraints ($k_{pr} = 1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$) on all heavy atoms of the A β_{40} peptide. Since the InflateGRO procedure removed all water molecules in the input configuration that needed to be added back later, the hydration of the equilibrated lipids was disrupted, requiring careful re-equilibration. The first phase employed an isochoric-isothermal (NVT) ensemble, with temperature controlled by the Berendsen weak coupling algorithm [187]. The NVT ensemble was applied for 100 ps, during which the temperature of the system was maintained at 100 K with a coupling constant of 0.1 ps. The protein, lipids, and solvent (including ions) were coupled separately. During NVT equilibration, position restraints were placed on lipid phosphorus atoms to restrict their motion to the $x - y$ plane. Following NVT equilibration, the restraints on the lipids were removed and simulated annealing

was performed to heat the system linearly from 100 K to 310 K over 500 ps under an isobaric-isothermal (*NPT*) ensemble. During annealing, the Berendsen weak coupling method [187] was used to control both temperature and pressure (1 bar). Coupling constants for temperature and pressure were 0.1 ps and 2.0 ps, respectively, and pressure was coupled semi-isotropically to allow independent deformations of the system in the $x - y$ and z -dimensions. Following annealing, 1 ns of *NPT* equilibration was performed, using the Nosé-Hoover thermostat [188, 189, 190] and Parrinello-Rahman barostat [191, 192]. Coupling constants were the same as during the *NPT* equilibration phase. Production simulations were carried out for 100 ns under the same *NPT* ensemble in the absence of any restraints.

5.2.3 Analysis

All data analysis was conducted using tools present within the GROMACS distribution. Secondary structure assignments were determined using the DSSP algorithm [271]. Images of A β ₄₀-membrane systems were generated with PyMOL [272]. Statistical outliers were identified using a Q test with 95% confidence. All values following \pm are standard deviations, unless otherwise noted.

5.3 Results

The systems analyzed here describe the interactions of the A β ₄₀ peptide with a variety of model membranes with different biophysical and biochemical properties. The POPC membrane allows us to examine the interactions of A β with a neutral, zwitterionic lipid with mixed acyl chains. The POPC/POPE membrane is much like the POPC membrane, but includes ethanolamine headgroups, which provide hydrogen bond donors by virtue of their primary amino groups. The POPS membrane contains headgroups with both hydrogen bond donor (amino group) and acceptor (carboxylate and phosphate) groups. The raft systems allow us to examine the interactions of A β ₄₀ with a model of the environment in which it is produced, with a specific focus on the impact of ganglioside GM1 enrichment in this membrane microdomain.

Our results lead us to introduce the following mechanism for the release of A β from the membrane environment following γ -secretase cleavage. The N-terminal residues of A β interact with clusters of GM1 through hydrogen bonding and other electrostatic interactions, inducing the peptide to adopt β -hairpin configurations and allowing the C-terminal embedded region to cross the membrane-water interface, aided by hydrogen bonding to nearby sugar moieties. These A β -GM1 interactions preclude other interfacial A β -lipid interactions involving the N-terminal residues that otherwise anchor the peptide to the membrane, which we observe in the simulations of A β ₄₀ in POPC, POPS, and POPC/POPE membranes. The stability of A β ₄₀ in POPC, POPS, and POPC/POPE indicates that A β may bind and insert into these membrane domains to exert its toxicity.

Throughout this section, simulations are categorized according to two factors, lipid type and A β ₄₀

model. There were two models, one wherein the C-terminal Val40 was protonated (model “CH”) and one wherein it was ionized (model “CI”). Lipids are identified according to headgroup (PC, PS, and PE). “Raft” refers to a lipid raft without GM1, while system identifiers with a “GM1” prefix indicate rafts that contain GM1. Individual simulations within each set are indicated by a numeral. For additional details, see Section 5.2.

5.3.1 Secondary Structure of A β_{40}

Average secondary structure content for all simulations performed here is listed in Table 5.1. The β -strand content reflects the sum of extended β -strands and isolated β -bridges, while “total helix” refers to the sum of α -, 3_{10} -, and π -helices.

Table 5.1: Average secondary structure content of A β_{40} in all model membranes over the last 50 ns of all trajectories.*

System	Coil	β -Strand	Bend	Turn	Total Helix
PC-CH	32 (4)	2 (2)	15 (8)	11 (5)	40 (16)
PC-CI	38 (7)	4 (5)	22 (8)	10 (4)	27 (15)
PS-CH	38 (3)	2 (3)	11 (2)	14 (8)	36 (6)
PS-CI	35 (10)	8 (8)	14 (6)	12 (3)	30 (8)
PC/PE-CH	40 (14)	5 (6)	21 (15)	11 (8)	22 (27)
PC/PE-CI	31 (6)	4 (3)	12.3 (0.7)	15 (6)	38 (7)
Raft-CH	35 (9)	0.1 (0.1)	12 (10)	13 (6)	34 (11)
Raft-CI	48 (9)	5 (8)	21 (10)	8 (2)	7 (5)
GM1-CH	33 (7)	10 (8)	25 (6)	13 (10)	20 (9)
GM1-CI	45 (15)	7 (7)	21 (3)	14 (7)	15 (10)

*All values are expressed as percentages, with standard deviations given in parenthesis.

5.3.1.1 A β_{40} in POPC

PC-containing lipids comprise the most common types of phospholipids found on the extracellular surface of eukaryotic plasma membranes [254, 255]. Our previous work [247, 269] modeled A β_{40} in a fully-saturated dipalmitoylphosphatidylcholine (DPPC) membrane, at elevated temperature (323 K) to maintain a fluid-phase membrane. POPC is a better model of a physiologically-relevant membrane lipid, as it can be simulated at 310 K and remain fluid.

In POPC, the A β_{40} peptide tended to lose some of its initial α -helicity (45% at the beginning of the simulation) over time in favor of random coil structures. In sets PC-CH and PC-CI, A β_{40} retained $40 \pm 16\%$ and $27 \pm 15\%$ total helicity, respectively. The lower helical content in the

Secondary structure

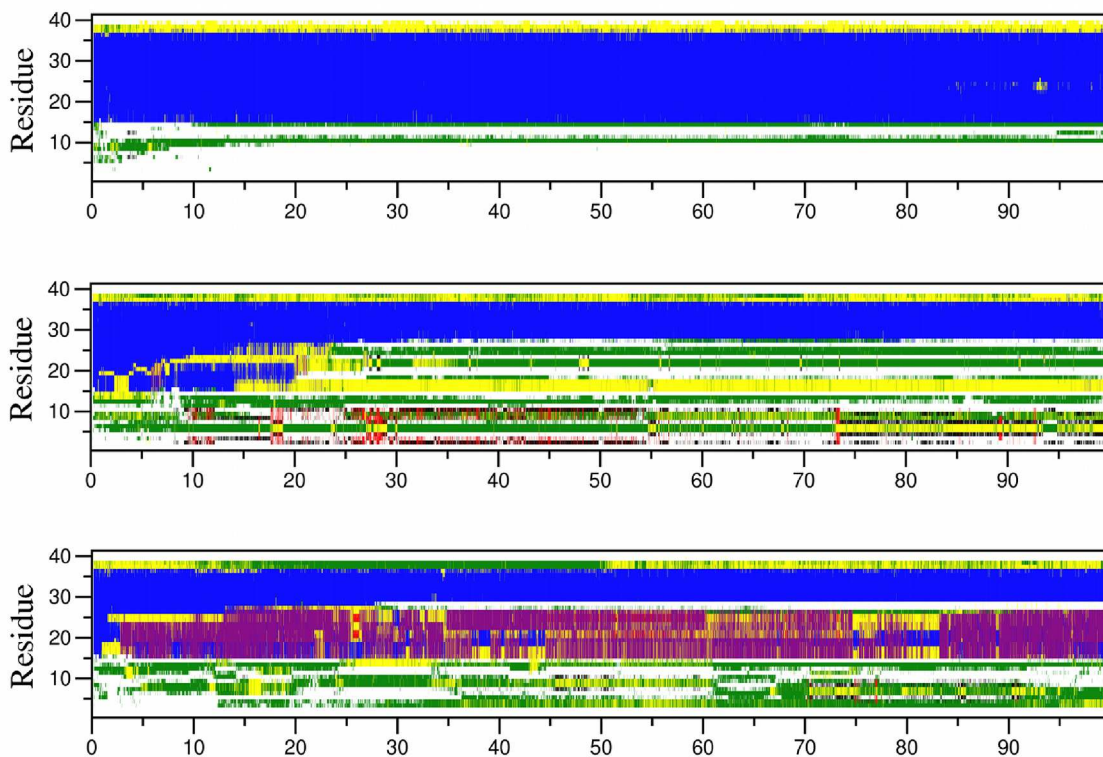


Figure 5.2: Secondary structure evolution for $A\beta_{40}$ model CH in POPC.

PC-CI simulations relative to the PC-CH simulations was principally due to the deprotonation of the C-terminus, which caused Val40 and the other hydrophobic C-terminal residues to disorder and approach, but not cross, the membrane-water interface. β -Strand content in the $A\beta_{40}$ peptides was low, $2 \pm 2\%$ for PC-CH simulations and $4 \pm 5\%$ for PC-CI simulations. These β -strands, which formed in four of the six simulations in this set, typically involved short stretches of amino acids (2-4 residues) at various positions within the N-terminal polar region. Snapshots from the end of the PC-CH and PC-CI simulations and secondary structure evolution are shown in Figures 5.2 and 5.3.

5.3.1.2 $A\beta_{40}$ in POPS

Though PS-containing phospholipids typically occur in the cytofacial leaflet of eukaryotic cell membranes [254, 255], anionic lipids are frequently used in biophysical characterization of $A\beta$ -membrane interactions [207, 234, 273, 274, 275]. It has been suggested that cellular injury may expose anionic lipids, like POPS, to the extracellular space [273]. Thus, $A\beta$ -POPS interactions are important to characterize, since the dynamics and unfolding of the peptide may be related to such

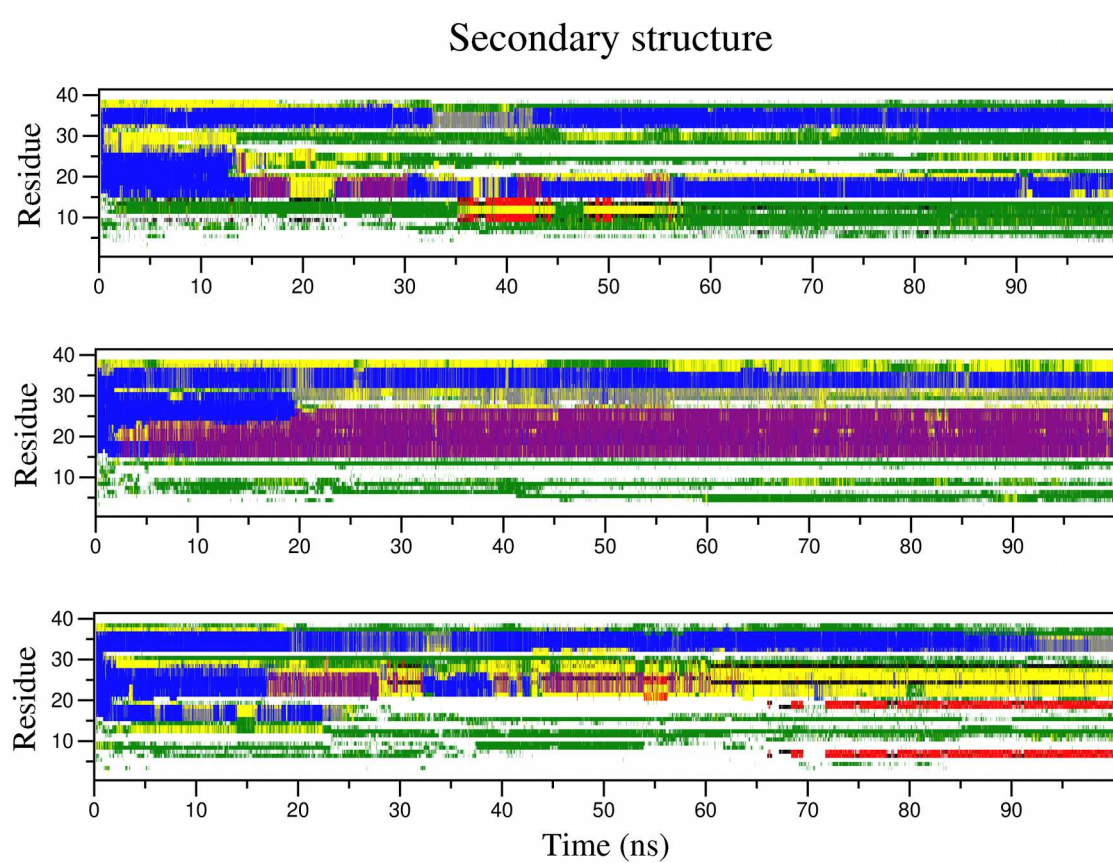


Figure 5.3: Secondary structure evolution for $A\beta_{40}$ model CI in POPC.

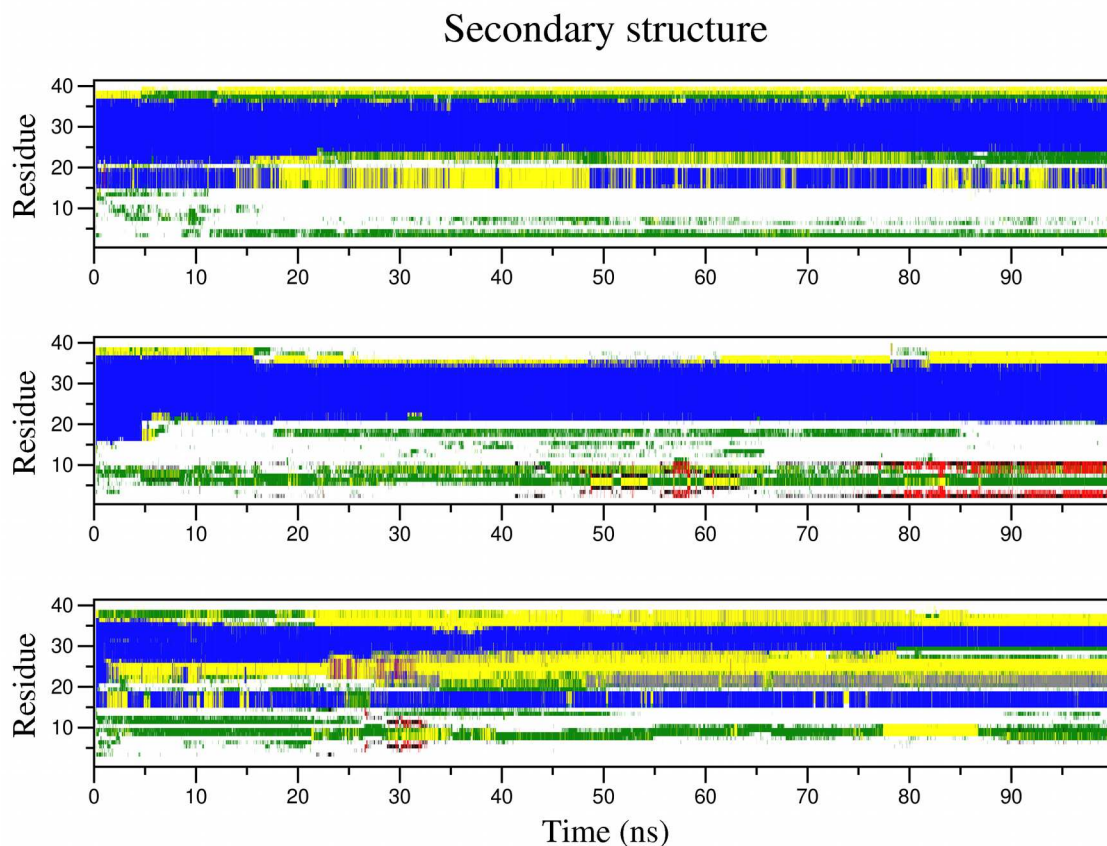


Figure 5.4: Secondary structure evolution for $A\beta_{40}$ model CH in POPS.

events.

The secondary structure content of $A\beta_{40}$ in POPS was similar to the results of $A\beta_{40}$ in POPC. The peptide retained $36 \pm 6\%$ and $30 \pm 8\%$ of its initial helicity, on average, in simulation sets PS-CH and PS-CI, respectively. The helical content is in good agreement with CD spectroscopy results of $A\beta_{40}$ inserted in anionic lipids [234]. Short β -strands developed in four of these six trajectories, with PS-CH simulations containing, on average, $2 \pm 3\%$ β -strand content, and PS-CI, $8 \pm 8\%$. As with simulations in POPC, N-terminal polar residues of $A\beta_{40}$ bound to the membrane-water interface, associating with the charged lipid headgroups of POPS. Snapshots from the end of the PS-CH and PS-CI simulations and secondary structure evolution are shown in Figures 5.4 and 5.5.

5.3.1.3 $A\beta_{40}$ in POPC/POPE

Addition of POPE to a POPC membrane (which has only hydrogen bond acceptor groups, phosphates) introduces many new hydrogen bonding partners for $A\beta$, a fact that may affect the secondary structure of the peptide and its ability to bind the membrane. POPE, by virtue of its pri-

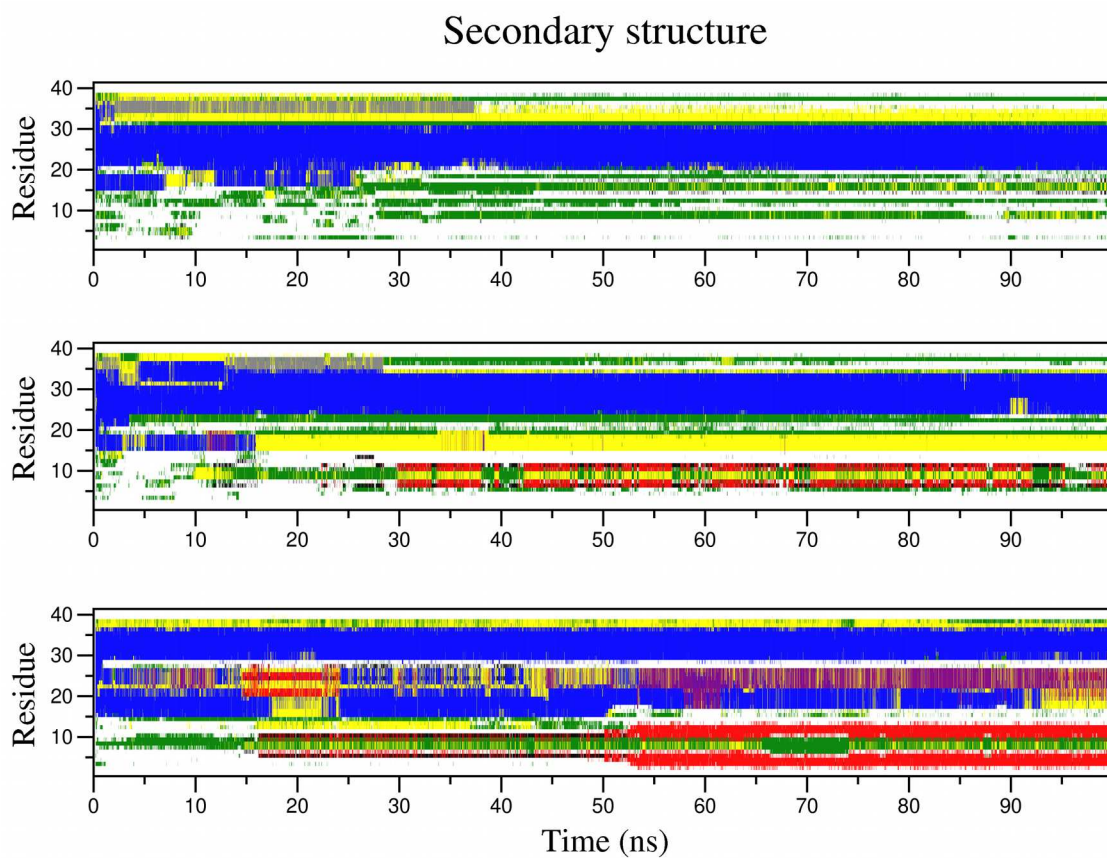


Figure 5.5: Secondary structure evolution for $A\beta_{40}$ model CI in POPS.

mary amine and phosphate groups, can engage in intermolecular (POPE-POPE, POPE-POPC, and POPE-A β) as well as intramolecular hydrogen bonding.

In the simulations of A β_{40} in POPC and POPS, model CH, with a protonated C-terminus, retained more of its initial helicity than model CI. In POPC-POPE membranes, however, this trend was reversed, with PC/PE-CH simulations averaging $21 \pm 27\%$ helicity, and PC/PE-CI simulations, $38 \pm 7\%$, although given the relatively wide standard deviations of these simulations, we cannot conclude that there is any significant difference between the different membranes with respect to helical content. In PC/PE-CH-1, A β_{40} lost all of its initial α -helicity, while in PC/PE-CH-2 and PC/PE-CH-3 and the entire PC/PE-CI set, at least some of the initial helicity was retained, as in the POPC and POPS simulations.

A β_{40} in POPC-POPE membranes developed similar β -strand content as the peptides in POPC and POPS systems, $5 \pm 6\%$ in the case of PC/PE-CH systems, and $4 \pm 3\%$ in PC/PE-CI. Of note in this simulation set was the development of a short β -hairpin (V $_{24}$ GSNKG $_{29}$) in A β_{40} in simulation PC/PE-CI-3. Ser26 and Asn27 formed the turn that connected the 4-residue hairpin. This structure evolved due to interactions of Asn27 and Lys28 sidechains with POPE headgroups. Hydrogen bonding stabilized the turn, allowing β -strands to evolve over the last 10 ns of simulation PC/PE-CH-3. Snapshots from the end of the PC/PE-CH and PC/PE-CI simulations and secondary structure evolution are shown in Figures 5.6 and 5.7.

5.3.1.4 A β_{40} in Raft Systems

The simulations of A β_{40} in POPC, POPS, and POPC/POPE presented here, as well as those of A β_{40} in DPPC published previously [247], indicate that the peptide possesses an intrinsic ability to form short β -strands and hairpins in a variety of membrane environments. Experiments have shown that ganglioside GM1 increases the level of β -strand content in A β [235, 259]. Toxic oligomeric species [276] and mature A β fibrils [277, 278] are rich in β -strand content, indicating that the formation of β -strand elements is central to the aggregation cascade.

We found that the conversion of α -helix to β -strand was amplified by the presence of GM1, principally through the formation of β -hairpins, which we observed in nine out of ten simulations of A β_{40} in GM1-containing rafts, all five GM1-CH trajectories and four GM1-CI trajectories (Figures 5.8 and 5.9). The β -hairpins occurred principally within the N-terminal 16 residues of A β_{40} and involved very specific residues, as discussed below in the context of hydrogen bonding. In one instance (simulation GM1-CI-5), a hairpin formed involving residues Val24-Gly25 and Lys28-Gly29, connected by a turn involving Ser26-Asn27. Formation of β -strand structures near glycines within residues 24-37 has previously been proposed as an important factor in the overall conversion of A β from α -helix to β -strand [199], but emergence of such structures has never before been observed in simulations of membrane-bound A β , indicating that this behavior is related to the presence of a suitable hydrogen bond donor/acceptor environment, such as GM1, or POPE, as we observed above in the case of simulation PC/PE-CH-3. This phenomenon is discussed in greater detail below. Conversely, in simulations of raft systems lacking GM1, we found that β -strands

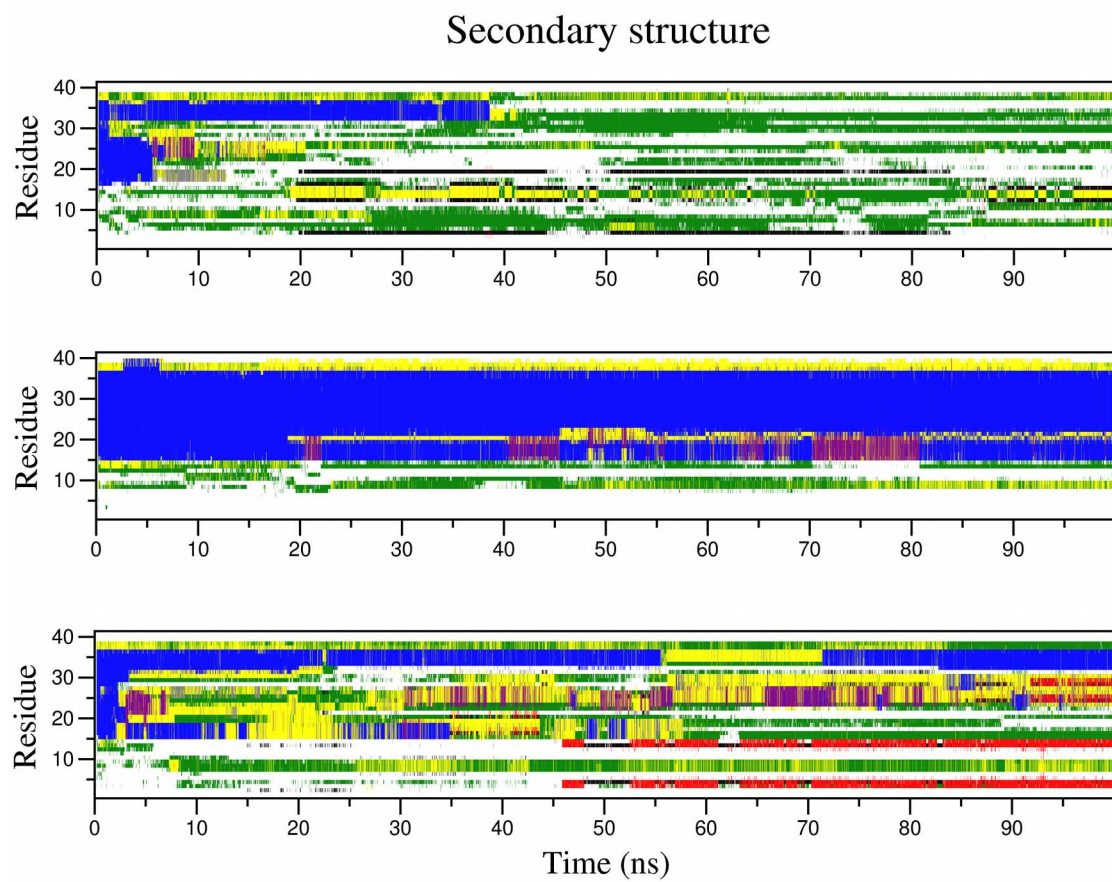


Figure 5.6: Secondary structure evolution for $A\beta_{40}$ model CH in POPC-POPE.

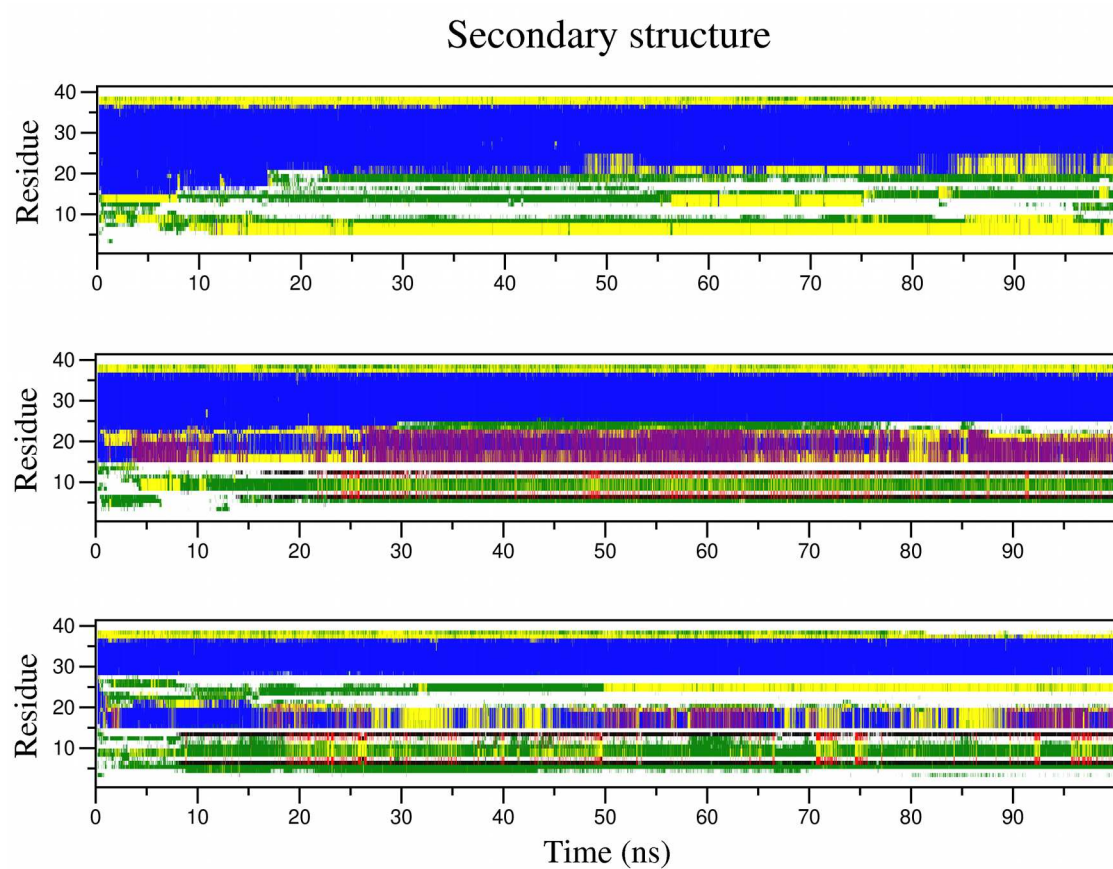


Figure 5.7: Secondary structure evolution for $A\beta_{40}$ model CI in POPC-POPE.

formed in only six out of the ten total trajectories (two Raft-CH and four Raft-CI, Figures 5.8 and 5.9).

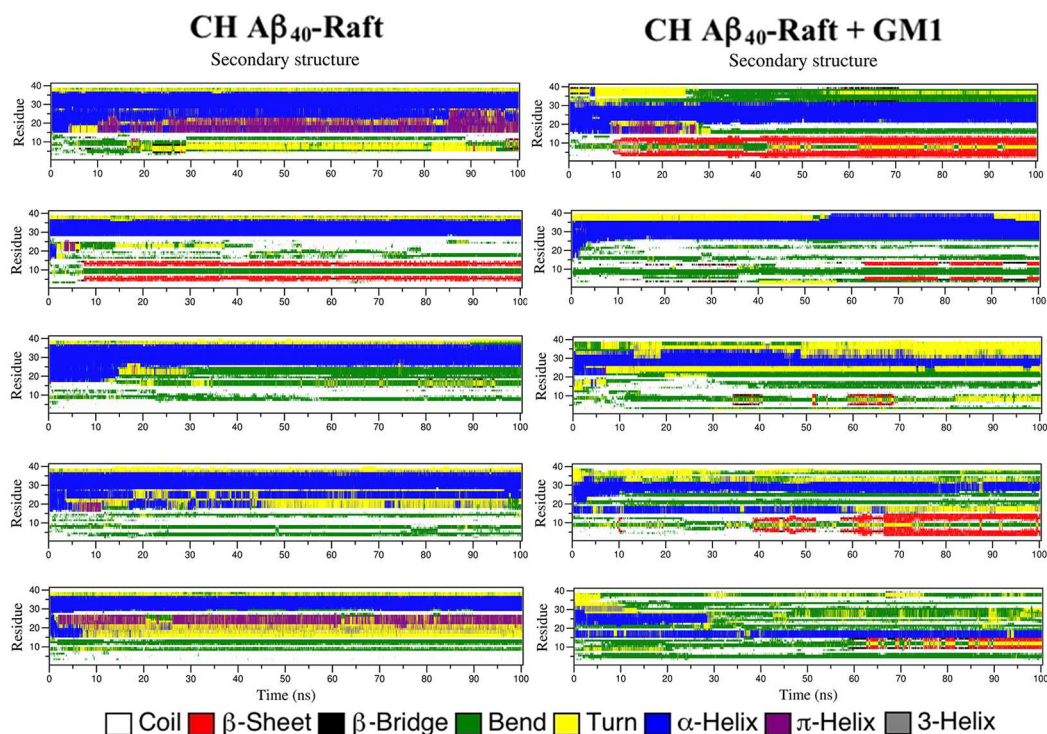


Figure 5.8: Secondary structure plots for $A\beta_{40}$ Raft-CH (left) and GM1-CH (right) systems.

Destabilization of helical regions of $A\beta_{40}$ occurred due to two principal factors in our simulations of $A\beta_{40}$ in a raft environment: (i) hydrogen-bonding interactions between N-terminal residues and GM1 to form β -hairpins and (ii) exposure of C-terminal residues 29-40 to solvent to form random coil structures. The C-termini of $A\beta_{40}$ in simulation sets GM1-CI and Raft-CI approached the membrane-water interface independently of the presence of GM1 (Figures 5.10 and 5.11) due to the ionized state of the C-terminal carboxylate of Val40, which snorkeled towards this polar environment. The result was a total helical (sum of α -, 3_{10} -, and π -helix) content of $7 \pm 5\%$ and $15 \pm 10\%$ for Raft-CI and GM1-CI simulation sets, respectively.

In the GM1-CH simulation set, interactions between $A\beta_{40}$ and GM1 reduced the helical content of $A\beta$ from $34 \pm 11\%$ in set Raft-CH to $20 \pm 9\%$ in set GM1-CH. GM1 caused $A\beta_{40}$ to rise further out of the membrane (discussed below), resulting in closer proximity of the C-terminal residues to the polar interface than in the CH set, causing them to become disordered. In one of the simulations, GM1-CH-5, the total helical content was reduced to 9.7% as residues 29-40 were extracted from the membrane and exposed to solvent (discussed below), destabilizing the initial α -helix in favor of random coil structures. The total helical content of residues 29-40 as a function of their center-of-mass position relative to the interface is shown in Figure 5.12. In general, helical content dropped sharply when residues 29-40 were within 1 nm of the interface, an environment

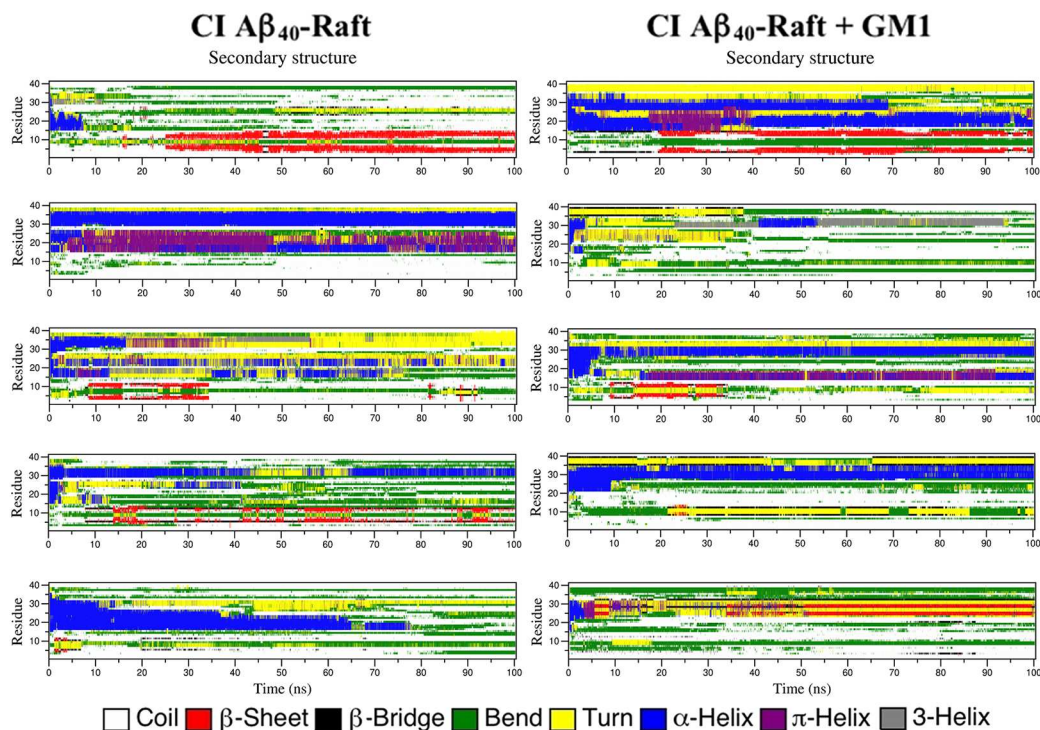


Figure 5.9: Secondary structure plots for $A\beta_{40}$ Raft-CI (left) and GM1-CI (right) systems.

that is much more polar than the membrane core.

Several representative examples of β -hairpins that formed during our simulations are shown in Figure 5.13. The longest β -hairpins (8-10 residues) formed at the interface of two or more GM1 molecules (Figures 5.13D and 5.13E), an event that was independent of the protonation state of the C-terminus of $A\beta_{40}$. These results provide insight into the mechanism underlying the experimental observations that $A\beta$ binds to GM1 clusters, thereby seeding further aggregation [235, 257, 258]. Gangliosides have been shown to dehydrate the membrane-water interface of phosphatidylcholine-containing lipid bilayers [279], and this phenomenon appears to play a functional role in terms of the unfolding of $A\beta$, facilitating the self-association of backbone amide groups that leads to β -strand hydrogen bonding. As a representative example, in simulation GM1-CH-4 (Figures 5.13A and 5.13B), a β -hairpin formed 1.3 nm above the membrane-water interface, at which point the density of water is 14.6% lower than bulk solution, while in raft systems lacking GM1, the density of water at this same vertical position above the interface is only 3.9% below bulk (Figure 5.13A). In simulation Raft-CH-2 (Figures 5.13A and 5.13C, wherein the position of the β -hairpin was representative of those formed in the Raft-CH and Raft-CI sets), a β -hairpin formed only 0.7 nm above the interface, a location where the water density is 9.4% lower than bulk.

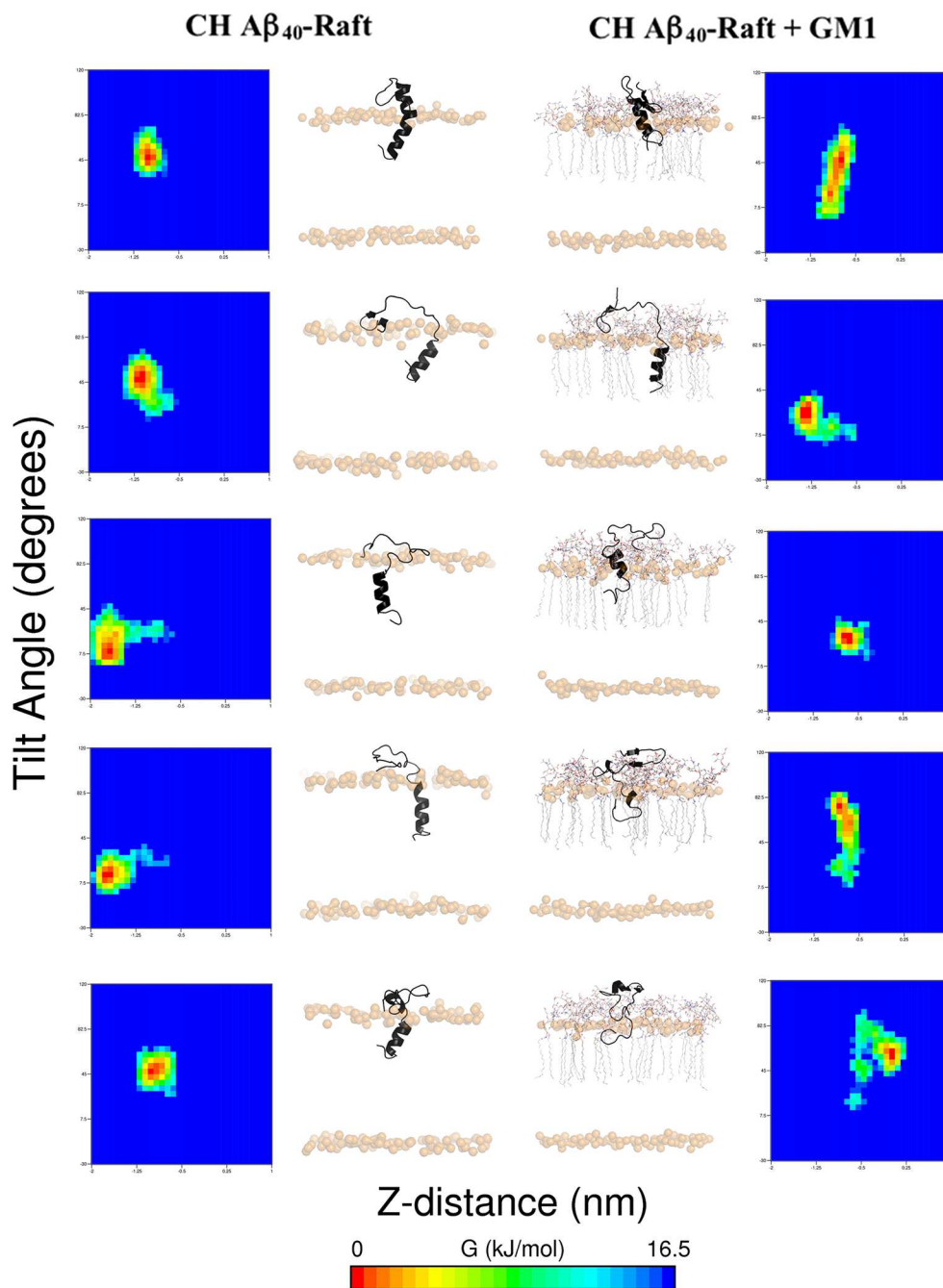


Figure 5.10: Free energy surfaces and snapshots at 100 ns for all Raft-CH and GM1-CH simulations. The $A\beta_{40}$ peptide is shown as a black cartoon and P atoms (POPC and PSM) are shown as translucent gold spheres. In the GM1-containing system, the GM1 molecules are shown as lines and colored by element. For clarity, other lipid atoms are not shown.

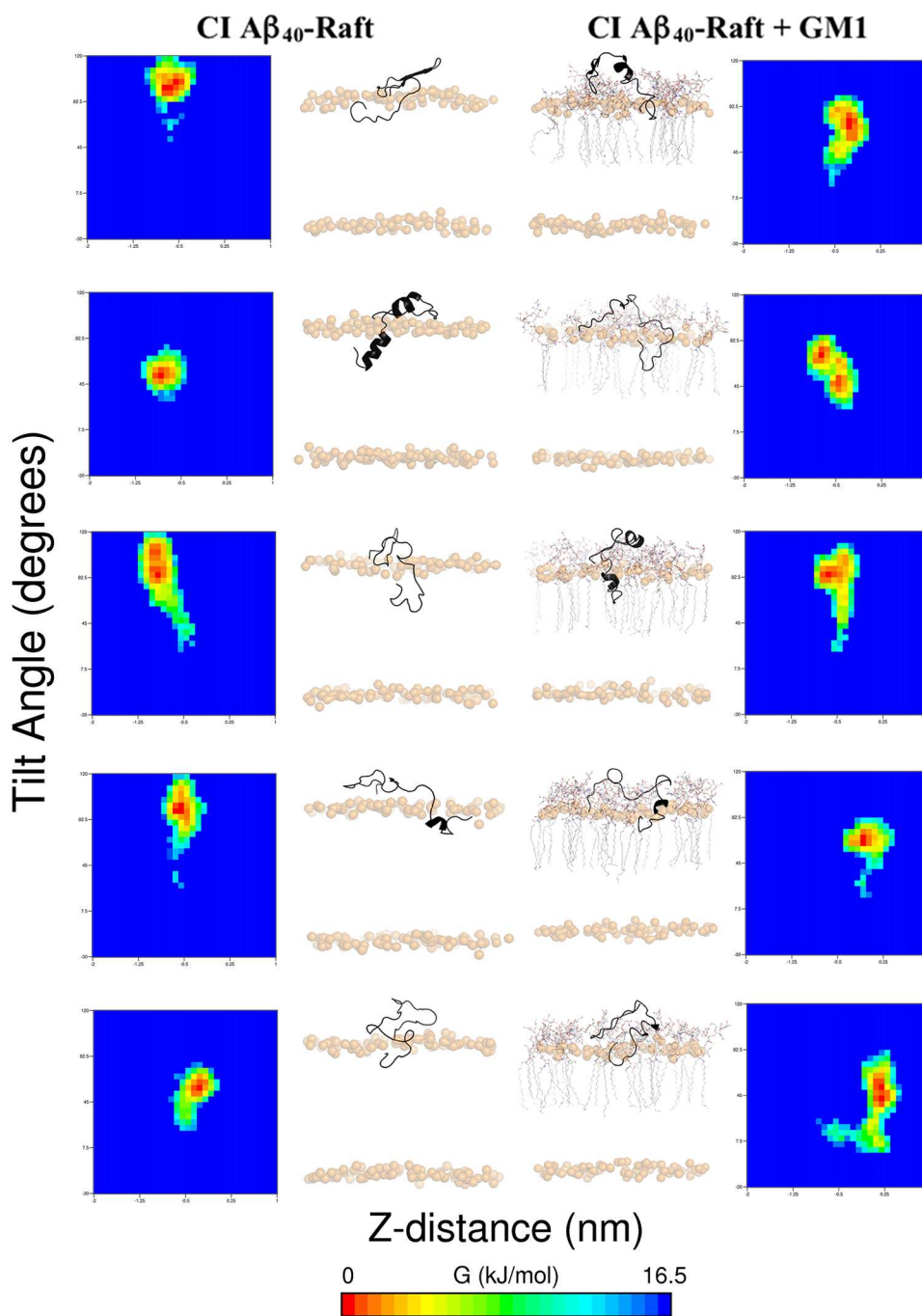


Figure 5.11: Free energy surfaces and snapshots at 100 ns for all Raft-CI and GM1-CI simulations. The $A\beta_{40}$ peptide is shown as a black cartoon and P atoms (POPC and PSM) are shown as translucent gold spheres. In the GM1-containing system, the GM1 molecules are shown as lines and colored by element. For clarity, other lipid atoms are not shown.

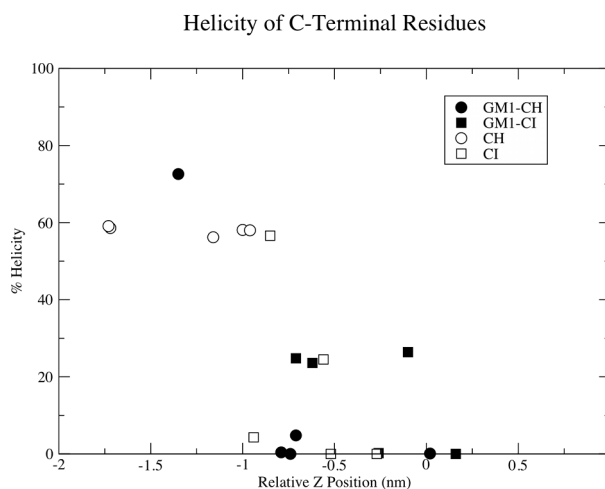


Figure 5.12: Total helicity of C-terminal residues 29-40 as a function of distance relative to the interface, defined as the average phosphorus position over time. The total helicity is defined as the sum of α -, 3_{10} -, and π -helical structures, averaged over the last 50 ns of simulation time.

5.3.2 Hydrogen Bonding Between $A\beta_{40}$ and Lipids

Table 5.2 summarizes backbone and sidechain hydrogen bonding between $A\beta_{40}$ and all phospholipids that were not part of raft membranes. In general, the sidechains of $A\beta_{40}$ formed more hydrogen bonds with surrounding lipids than did backbone groups, in agreement with experimental findings that $A\beta$ -membrane interactions are largely mediated by non-specific electrostatic interactions, such as those found between lipid headgroups and the charged and polar residues in the N-terminal region of $A\beta$ [280].

5.3.2.1 $A\beta_{40}$ in POPC

POPC lipids are only capable of serving as hydrogen bond acceptors, having phosphate and ester groups that can interact with hydrogen bond donors in $A\beta_{40}$. Hydrogen bonds involving backbone amide groups were increased in the PC-CI set relative to PC-CH, principally due to the lower α -helical content of the $A\beta_{40}$ peptide in these simulations relative to PC-CH. As described above, the ionized state of Val40 in the PC-CI series caused this residue to snorkel towards the membrane-water interface, destabilizing helicity in the C-terminal residues, as well as several other residues that reside at or below the interface. This destabilization led to exposed amide groups that interacted with POPC hydrogen bond acceptors, particularly the ester functional groups connecting the glycerol backbone of POPC to the acyl chains. In PC-CI simulations, the number of hydrogen bonds to glycerol esters averaged 4 ± 2 , while in PC-CH simulations, only approximately one hydrogen bond formed with these groups. Hydrogen bonding to phosphate moieties was comparable between the two simulation sets, 3 ± 1 for PC-CH systems and 5 ± 2 for PC-CI.

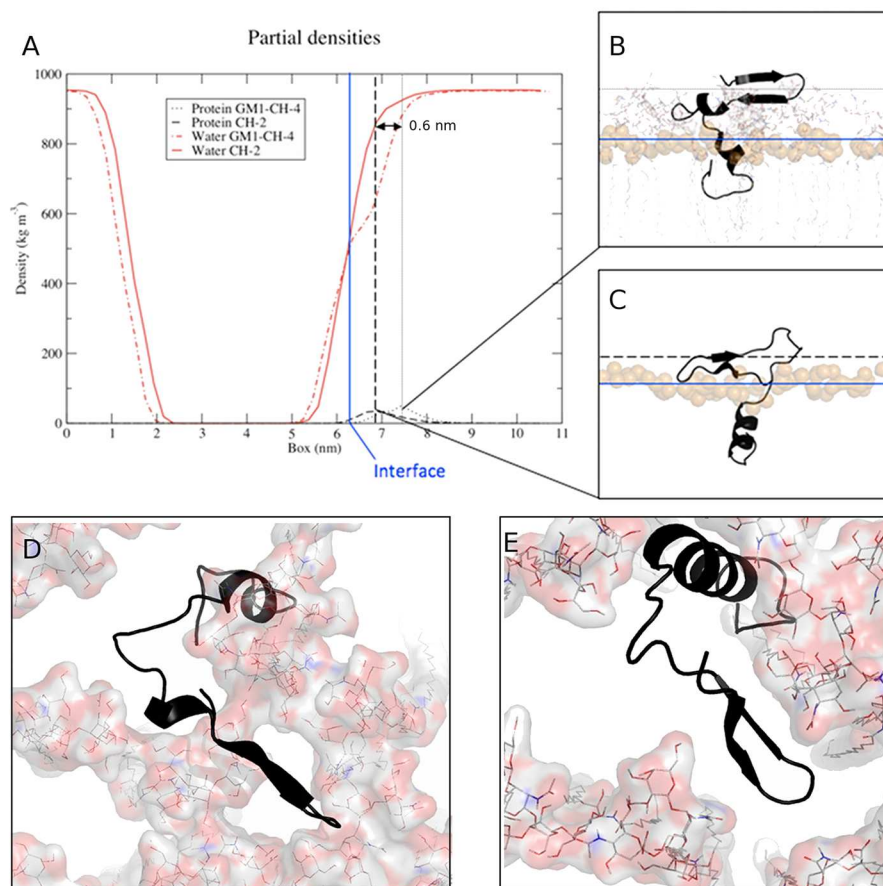


Figure 5.13: (A) Position of β -hairpin formation relative to the level of hydration along the z -dimension of the simulation box. The membrane-water interface is indicated as a solid blue vertical line, and the density maxima for the β -hairpins formed in simulations GM1-CH-4 and CH-2 indicated by dotted and dashed lines, respectively. (B) Formation of an 8-residue β -hairpin in simulation GM1-CH-4 and (C) formation of a 4-residue β -hairpin in simulation Raft-CH-2. Top-down views (along the z -axis) of two GM1-containing systems showing the locations of β -hairpins at the interface of multiple GM1 molecules in simulations (D) GM1-CH-4 and (E) GM1-CH-1.

Sidechain-POPC hydrogen bonds were decreased in PC-CI simulations relative to PC-CH, likely due to the backbone destabilization phenomena discussed above. In the PC-CI simulations, the backbone is more exposed to the surrounding lipids, causing their hydrogen bonding acceptor groups to be occupied more frequently by exposed amide functional groups, leaving fewer open positions for other sidechain moieties that may be present.

Table 5.2: Hydrogen bonding between A β ₄₀ and lipids in all non-raft systems.*

System		POPC	POPS	POPE
PC-CH	Backbone	4 (2)	–	–
	Sidechain	9 (1)	–	–
PC-CI	Backbone	8 (3)	–	–
	Sidechain	4.4 (0.8)	–	–
PS-CH	Backbone	–	6 (4)	–
	Sidechain	–	19 (2)	–
PS-CI	Backbone	–	6 (3)	–
	Sidechain	–	13 (1)	–
PC/PE-CH	Backbone	4 (2)	–	2 (2)
	Sidechain	4 (2)	–	11 (1)
PC/PE-CI	Backbone	3 (1)	–	4 (1)
	Sidechain	3.5 (0.3)	–	8 (4)

*All values are expressed as the average number of hydrogen bonds, with standard deviations given in parenthesis.

5.3.2.2 A β ₄₀ in POPS

Hydrogen bonding patterns in the PS-CH and PS-CI simulation sets were very similar, due largely to the fact that the A β ₄₀ peptide in these simulations adopted very similar secondary structures. Thus, the number of exposed backbone amides in each simulation was very similar. The slight increase in total helicity of A β ₄₀ in the PS-CI simulation set relative to those of PC-CI (Table 5.1) caused the peptides in the PS-CI set to form fewer backbone-POPS hydrogen bonds than backbone-POPC hydrogen bonds in the PC-CI set. The embedded C-terminal residues were somewhat more helical in PS-CI than in PC-CI due to repulsion between the negatively charged C-terminus and anionic interface. This electrostatic repulsion prevented the C-terminal residues from approaching the interface as closely as in the case of PC-CI simulations, leading to less destabilization of the helix and less overall hydrogen bonding involving backbone groups.

Numerous hydrogen bonds formed between A β ₄₀ sidechain moieties and POPS lipids, principally involving headgroup functional groups. Unlike POPC, which contains only hydrogen bond acceptors, POPS contains an amino group that serves as a hydrogen bond donor and a carboxylate group that is an additional acceptor. These two amino acid-derived functional groups are the principal sites of hydrogen bonding between A β ₄₀ sidechains and POPS lipids. Thus, while levels of backbone-lipid hydrogen bonding were comparable between the POPC and POPS systems, the presence of POPS lipids allowed for greater overall hydrogen bonding capacity, particularly with the sidechains of A β ₄₀. Thus, the N-terminal residues of A β ₄₀ are capable of extending along the interface, forming hydrogen bonds and other electrostatic interactions with several different moieties on the POPS lipids.

5.3.2.3 A β_{40} in POPC/POPE

A slightly smaller number of hydrogen bonds formed between A β_{40} and POPC lipids in the equimolar POPC/POPE membrane than in the pure POPC membrane (Table 5.2). This fact can be attributed to the presence of POPE lipids in the membrane. Though the backbone of A β_{40} participated in more hydrogen bonds with POPC than POPE, the sidechain moieties of A β_{40} formed significantly more hydrogen bonds to POPE than POPC, due largely to the presence of the primary amine in the POPE headgroup. A β_{40} -amine hydrogen bonds accounted for 54.4% of the total A β_{40} -POPE hydrogen bonds in the PC/PE-CH systems, and 63.1% in PC/PE-CI.

5.3.2.4 A β_{40} in Raft Systems

Ganglioside GM1 contains an oligosaccharide headgroup that provides numerous sites for potential hydrogen bonding to A β (Figure 5.1), substantially more per lipid than any of the other phospholipids analyzed above. In simulations of A β_{40} in GM1-containing rafts, A β_{40} formed hydrogen bonds with the GM1 molecules, most frequently with the Neu5Ac and Gal2 moieties (Table 5.3). This behavior occurred independently of the protonation state of the C-terminus, since hydrogen bonds primarily involved the polar N-terminal residues. In simulations of A β_{40} in lipid rafts lacking GM1, the backbone and sidechains of A β_{40} formed numerous hydrogen bonds to the phosphate moieties of the POPC and PSM lipids (Table 5.4). In GM1-CH systems, hydrogen bonds to POPC and PSM lipids were decreased by 68.0% and 43.4%, respectively, relative to Raft-CH systems. For GM1-CI models, hydrogen bonds to POPC and PSM were reduced by 58.8% and 30.7%, respectively. Hydrogen bonding to cholesterol and the sphingosine backbone groups of PSM was negligible (Table 5.4).

Table 5.3: Hydrogen bonding between A β_{40} and GM1 sugar moieties.*

System		Glc	Gal1	Neu5Ac	GalNAc	Gal2
GM1-CH	Backbone	0.3 (0.6)	0.3 (0.5)	1.1 (0.7)	1 (1)	1.5 (0.9)
	Sidechain	0.02 (0.03)	0.1 (0.2)	3 (2)	1 (1)	2 (1)
GM1-CI	Backbone	0.06 (0.05)	0.0 (0.0)	1 (1)	0.8 (0.8)	1.3 (0.9)
	Sidechain	0.02 (0.04)	0.3 (0.2)	3 (2)	0.6 (0.4)	1.8 (0.7)

*All values are expressed as the average number of hydrogen bonds, with standard deviations given in parenthesis.

The presence of GM1 thus reduced the incidence of interfacial interactions between A β and the other lipids. Our previous work [247, 269] indicated that the N-terminus of A β_{40} associates strongly with the membrane-water interface in a DPPC model system, principally through electrostatic interactions with the zwitterionic lipid headgroups. In the present study, analysis of three other non-raft phospholipid membranes indicates that A β associates strongly with these lipids, as

Table 5.4: Hydrogen bonding between $A\beta_{40}$ and POPC, PSM, and cholesterol for raft-only and GM1-raft systems.*

System		POPC PO ₄	POPC glycerol	PSM PO ₄	PSM NH	PSM OH	Chol
Raft-CH	Backbone	3 (1)	0.9 (0.7)	3 (1)	0.0 (0.0)	0.5 (0.8)	0.0 (0.1)
	Sidechain	5 (1)	2 (1)	5 (2)	0.0 (0.0)	0.4 (0.4)	0.3 (0.5)
GM1-CH	Backbone	0.9 (0.8)	0.0 (0.0)	0.9 (0.7)	0.9 (0.6)	0.0 (0.0)	0.3 (0.7)
	Sidechain	2.3 (0.9)	0.0 (0.0)	2.4 (0.9)	0.3 (0.4)	0.3 (0.4)	0.0 (0.0)
Raft-CI	Backbone	4 (1)	1.3 (0.9)	3.4 (0.7)	0.2 (0.4)	0.1 (0.1)	1 (2)
	Sidechain	5 (1)	1.3 (0.4)	5 (2)	0.0 (0.0)	0.2 (0.2)	0.0 (0.0)
GM1-CI	Backbone	2 (2)	0.2 (0.5)	2 (2)	0.9 (0.8)	0.1 (0.2)	0.1 (0.2)
	Sidechain	3 (1)	0.2 (0.2)	2.9 (0.9)	0.4 (0.3)	0.0 (0.0)	0.1 (0.1)

*All values are expressed as the average number of hydrogen bonds, with standard deviations given in parenthesis.

well. Hydrogen bonding to the membrane-water interface is amplified in POPC/POPE and POPS systems relative to the pure POPC membrane due to the additional hydrogen bonding capacity of the POPE and POPS lipids.

Previous work [280] has concluded that $A\beta$ associates with phospholipid headgroups through non-specific electrostatic interactions, but interactions between $A\beta$ and GM1 have been postulated to be based on specific interactions [281]. Our findings indicate that specific, favorable interactions between $A\beta$ and GM1 can effectively compete with non-specific electrostatic interactions between $A\beta$ and other nearby phospholipids that more often cause the N-terminal residues of $A\beta$ to otherwise adopt coil or bend configurations (Figures 5.8 and 5.9).

Examination of the most persistent hydrogen bonds formed between $A\beta$ and GM1 elucidates the roles of several important amino acids in these trajectories, most notably His6, Asp7, His13, and His14. Very recently, His13 and His14 have been identified as key residues for the binding of $A\beta$ to GM1 [282]. The Neu5Ac moiety of GM1 has previously been proposed as an important binding site for $A\beta$, an effect that is enhanced by the presence of a terminal galactose residue (Gal2) in the oligosaccharide headgroup [281]. In the five instances in which Asp7 interacted with either Neu5Ac or Gal2, a β -turn formed at this location, stabilizing the β -hairpin structure in nearby residues, a structure that has previously been proposed to contribute to neurotoxicity [283].

Regarding the histidine residues, in eight out of ten simulations of $A\beta_{40}$ in GM1-containing rafts (four in each set, GM1-CH and GM1-CI), at least one of these histidine residues formed persistent hydrogen bonds (lasting continuously for 10 ns or more) with at least one sugar moiety on GM1, and in six of these instances, the sugar moiety involved was Neu5Ac. In eight out of nine simulations that formed β -strands, at least one histidine residue was present in the β -strand, but in the absence of GM1, these residues were principally found in coil or bend configurations (Figures 5.8 and 5.9). The presence of carbohydrate groups and their roles as both hydrogen bond donor and

acceptor have been proposed to stabilize β -structures in the A β peptide, specifically in the vicinity of histidine residues at positions 6, 13, and 14 in the A β sequence [284, 285, 286]. In addition to these hydrogen-bonding interactions, in simulation GM1-CI-5, the formation of β -strand and turn structures described above was promoted by hydrogen bonding interactions formed by Gly29 with Neu5Ac, as well as Ser26 and Asn27 with Neu5Ac and GalNAc.

5.3.3 Position and Orientation of A β_{40} in Membranes

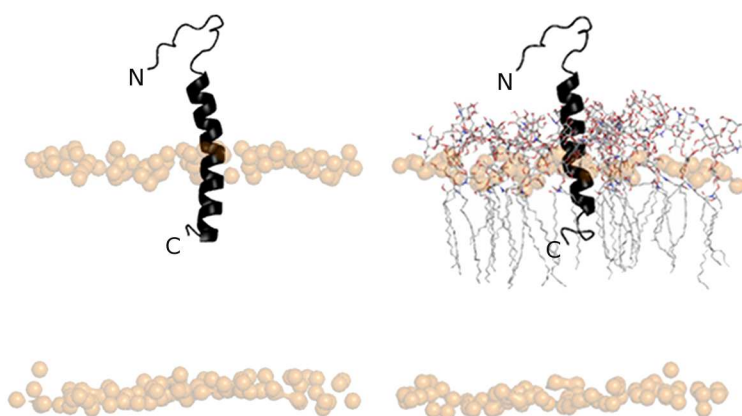


Figure 5.14: Starting structures of A β_{40} -raft (left) and A β_{40} -raft + GM1 (right) systems. The systems are rendered as in Figures 5.10 and 5.11.

The initial orientation of A β_{40} in all of the systems modeled here was prepared such that hydrophobic residues 29-40 were embedded in the hydrophobic core of the lipid membrane, with the helix axis aligned with the membrane normal (Figure 5.14). The position and orientation of this peptide segment over time can be described by two principal factors that influence solvent exposure and thus aggregation propensity: (i) the location of residues 29-40 with respect to the membrane interface (defined as the average phosphorus plane of the extracellular leaflet of the bilayer) and (ii) the tilt angle of the principal axis of these same residues with respect to the membrane normal. The interactions of A β with the surrounding lipid matrix influence these two variables, which can be plotted as free energy surfaces (Figure 5.15). The free energy of a configuration with respect to these two variables, tilt (θ) and z -position (z), can be described by the following expression:

$$\Delta G(z, \theta) = -k_B T [\ln P(z, \theta) - \ln P_{max}] \quad (5.1)$$

The probability distribution is determined by binning the tilt and z -coordinate data from the MD trajectories. The maximum of the distribution (P_{max}) is subtracted to give $\Delta G(z, \theta) = 0$ for the lowest point on the free energy surface.

Table 5.5 summarizes the average values of the tilt angle and relative z -position for all membranes studied in the present work.

Table 5.5: Average C-terminal tilt angles and vertical positions of A β_{40} residues 29-40 in all model membranes over the last 50 ns of all trajectories.*

System	Tilt Angle (degrees)	Relative Position (nm)
PC-CH	44 (27)	-1.3 (0.3)
PC-CI	101 (10)	-0.8 (0.3)
PS-CH	83.0 (0.4)	-1.4 (0.2)
PS-CI	87 (30)	-1.13 (0.05)
PC/PE-CH	37 (26)	-0.867 (0.009)
PC/PE-CI	71 (5)	-1.1 (0.3)
Raft-CH	36 (18)	-1.3 (0.4)
Raft-CI	80 (22)	-0.6 (0.3)
GM1-CH	48 (18)	-0.7 (0.5)
GM1-CI	69 (11)	-0.3 (0.3)

*All values are expressed with standard deviations given in parenthesis.

5.3.3.1 A β_{40} in POPC

Work by Bokvist *et al.* [234] indicated that A β_{40} can partially insert into zwitterionic membranes such as POPC, with its hydrophobic C-terminal residues oriented towards the hydrophobic interior of the membrane. Our results are in agreement with that finding. Though the embedded C-terminal residues of A β_{40} (residues 29-40) adopted a considerable tilt angle with respect to the bilayer normal, the center of mass of these residues remained well below the membrane-water interface (Table 5.5). This vertical distance was -1.3 ± 0.3 nm in the case of the PC-CH simulations, and -0.8 ± 0.3 nm in the case of PC-CI. Thus, despite the fact that the C-terminal Val40 snorkeled towards the membrane-water interface (in both PC-CH and PC-CI, but to a greater extent in PC-CI wherein the C-terminus was ionized), the hydrophobic C-terminal residues did not exit the membrane (Figures 5.16 and 5.17). A β_{40} was stably inserted in this PC-containing membrane, as Bokvist *et al.* [234] and others [208, 287] have predicted.

5.3.3.2 A β_{40} in POPS

In addition to proposing the partial insertion of A β_{40} into zwitterionic membranes, Bokvist *et al.* also demonstrated that A β_{40} could insert more deeply within anionic lipid membranes than those composed of neutral lipids. With respect to the C-terminal position relative to the interface, our simulations of A β_{40} in POPS showed that the A β_{40} peptide remained more deeply embedded in the

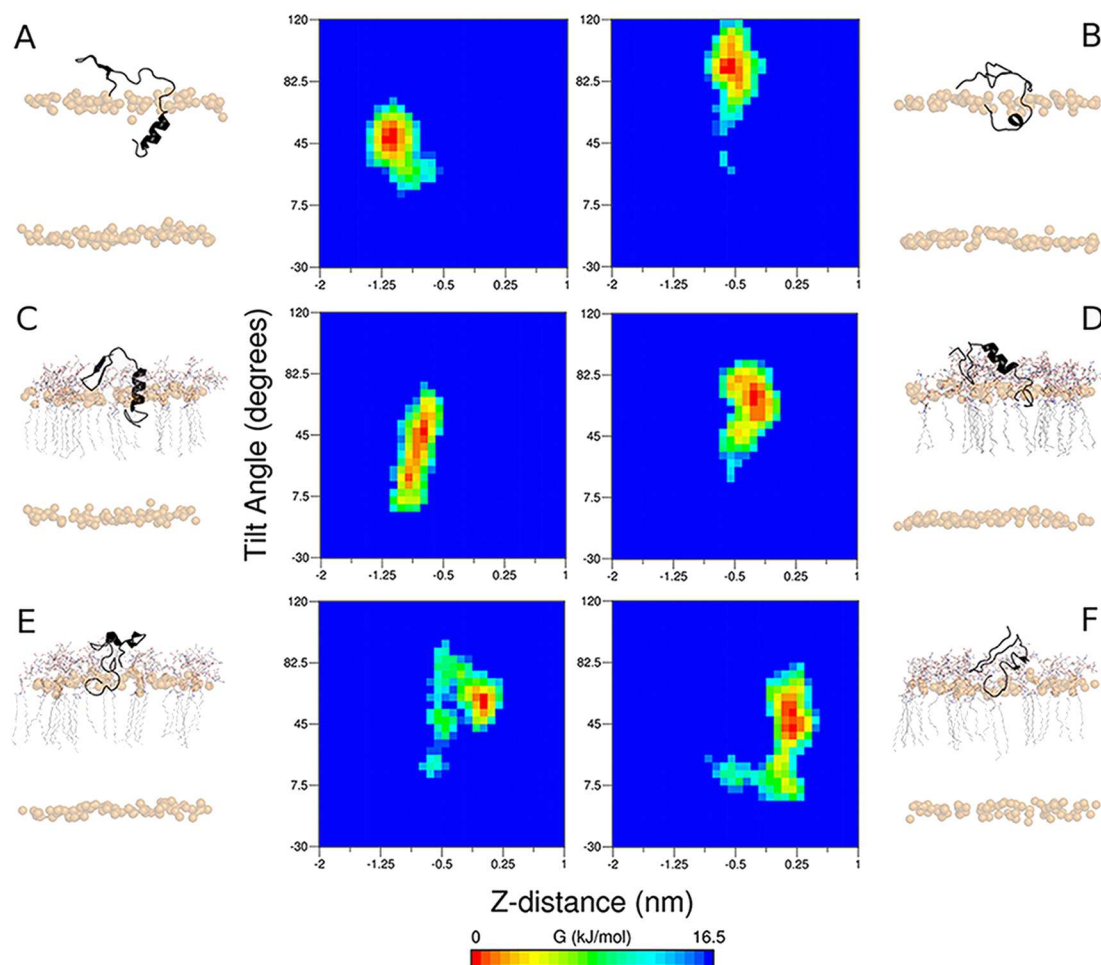


Figure 5.15: Free energy surfaces as a function of both the C-terminal tilt angle (y) and vertical position of $A\beta$ residues 29-40 relative to the average phosphate plane of the extracellular membrane leaflet (x). Representative images were chosen from each set for $A\beta_{40}$ in 1:1:1 POPC:PSM:cholesterol raft systems: (A) Raft-CH-2 and (B) Raft-CI-4; GM1-containing raft systems: (C) GM1-CH-1 and (D) GM1-CI-1; and GM1-containing raft systems wherein the free energy minimum resided above the average phosphate plane: (E) GM1-CH-5 and (F) GM1-CI-5. Rendered images from each system correspond to a configuration taken from the energy minimum of each plot, with the $A\beta_{40}$ peptide shown as a black ribbon, phosphorus atoms shown as transparent gold spheres, and GM1 molecules drawn as lines and colored by element.

membrane than in the case of the POPC systems, -1.4 ± 0.2 nm for PS-CH systems and -1.13 ± 0.05 nm for PS-CI systems. The significantly deeper insertion of the PS-CI systems relative to the PC-CI systems derived from the negative charge on the C-terminal Val40 residue. Though it had a tendency to snorkel towards the polar membrane-water interface (Figures 5.18 and 5.19), it was ultimately repelled by the net negative charge of this environment and remained within the

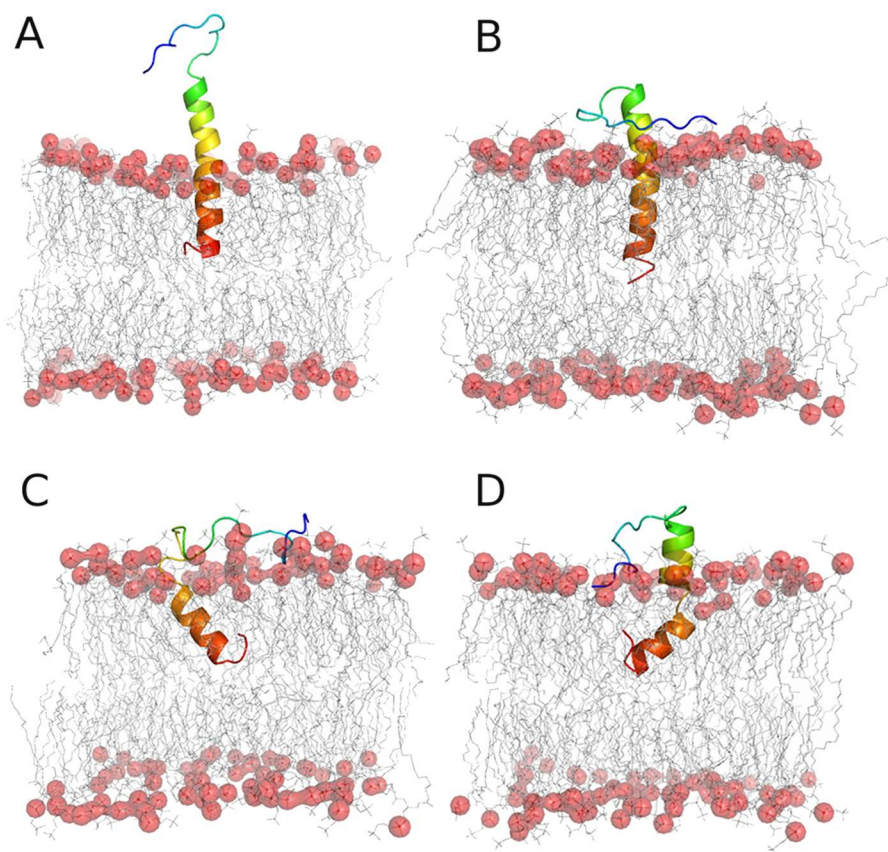


Figure 5.16: Structures of $A\beta_{40}$ model CH peptides in POPC: (A) Starting structure for the PC-CH trajectories and structures from the end of the three 100-ns PC-CH: (B) replicate 1, (C) replicate 2, and (D) replicate 3. The $A\beta_{40}$ peptide is rendered as a ribbon and colored as a gradient from N-terminus (blue) to C-terminus (red). POPC lipids are shown as gray lines, with P atoms shown as translucent red spheres. For clarity, water molecules are not shown.

hydrophobic core of the membrane.

5.3.3.3 $A\beta_{40}$ in POPC/POPE

In an equimolar mixture of POPC and POPE, $A\beta_{40}$ also remained embedded in the membrane, regardless of the protonation state of Val40. The average relative z -position of residues 29-40 was -0.867 ± 0.009 in the case of PC/PE-CH and -1.1 ± 0.3 nm for PC/PE-CI. As with POPC and POPS systems, these residues tilted with respect to the bilayer normal, $37 \pm 26^\circ$ in the case of PC/PE-CH systems, and $71 \pm 5^\circ$ in PC/PE-CI systems. These tilt angles are in good agreement with the findings of Ravault *et al.* [288], who studied the dynamics of a truncated form of $A\beta$, the

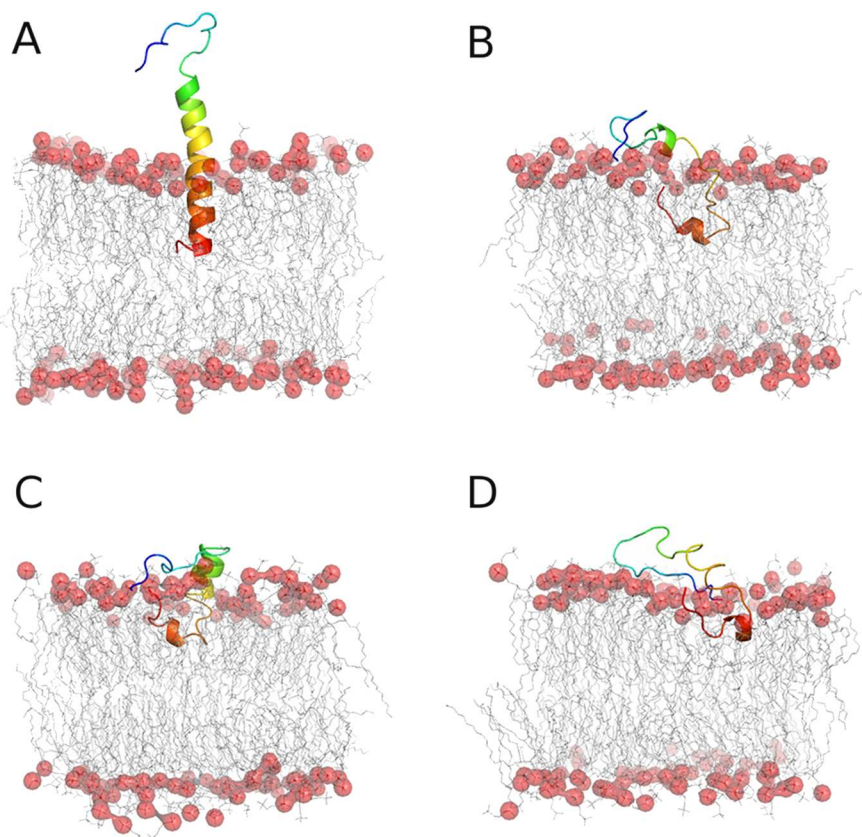


Figure 5.17: Structures of A β_{40} model CI peptides in POPC: (A) Starting structure for the PC-CI trajectories and structures from the end of the three 100-ns PC-CI: (B) replicate 1, (C) replicate 2, and (D) replicate 3. The A β_{40} peptide is rendered as a ribbon and colored as a gradient from N-terminus (blue) to C-terminus (red). POPC lipids are shown as gray lines, with P atoms shown as translucent red spheres. For clarity, water molecules are not shown.

peptide fragment from residues 29-42, in a 9:1 POPC:POPE membrane. To our knowledge, this is the only experimentally-characterized system that is comparable to the one we have studied here. Though Ravault *et al.* explored a slightly different system over a longer timeframe (ms and higher) than is accessible to atomistic simulations (ns - μ s), our results are in good qualitative agreement. In that study, and as shown by our results, the C-terminal residues of A β tended to localize just below the membrane-water interface, with the C-terminus tilted up to interact with this polar region (Figures 5.20 and 5.21).

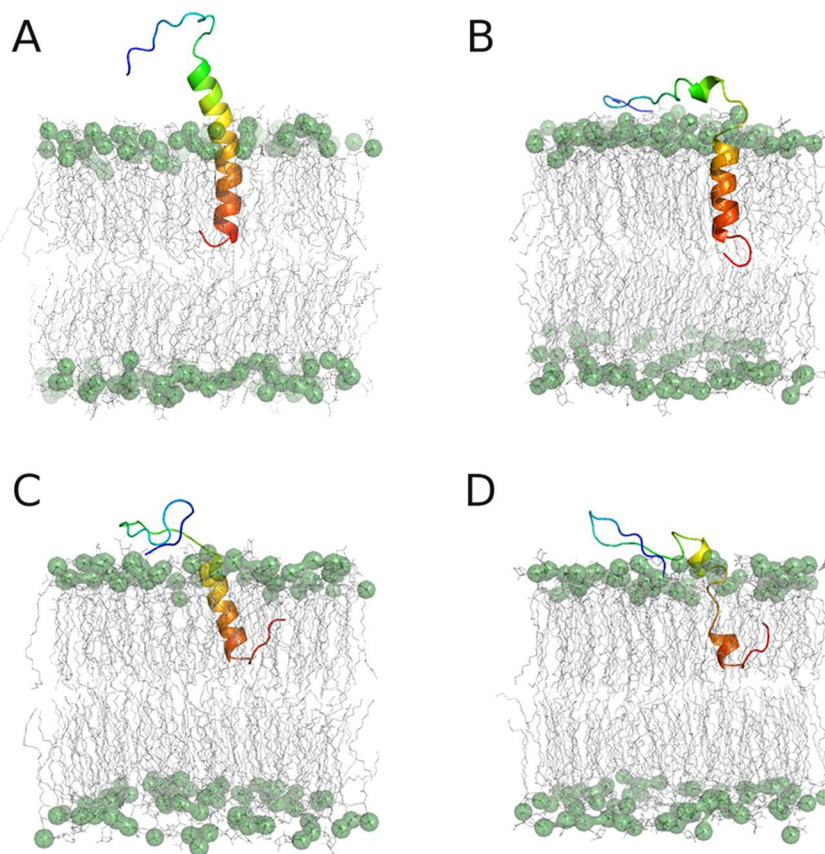


Figure 5.18: Structures of $A\beta_{40}$ model CH peptides in POPS: (A) Starting structure for the PS-CH trajectories and structures from the end of the three 100-ns PS-CH: (B) replicate 1, (C) replicate 2, and (D) replicate 3. The $A\beta_{40}$ peptide is rendered as a ribbon and colored as a gradient from N-terminus (blue) to C-terminus (red). POPS lipids are shown as gray lines, with P atoms shown as translucent green spheres. For clarity, water molecules are not shown.

5.3.3.4 $A\beta_{40}$ in Raft Systems

In lipid raft systems without GM1, the position of the center of mass of residues 29-40 remained well below the membrane-water interface, located at -1.3 ± 0.4 nm and -0.6 ± 0.3 nm for Raft-CH and Raft-CI systems, respectively, relative to the membrane-water interface (Figures 5.15A, 5.15B, 5.10, and 5.11). Some tilting and disordering of residues 29-40 was intrinsic to $A\beta$ in lipid rafts, as well, with Raft-CH systems tilting, on average, $36 \pm 18^\circ$ and Raft-CI systems, $80 \pm 22^\circ$. In the case of Raft-CI systems, the ionized state of Val40 caused the carboxylate moiety to snorkel towards the membrane interface in all five simulations, leading to greater tilt and proximity to the interface overall (Figure 5.11). However, based on the position of residues 29-40, $A\beta_{40}$ clearly remained embedded in all raft systems in the absence of GM1. The rigidity of the membrane itself

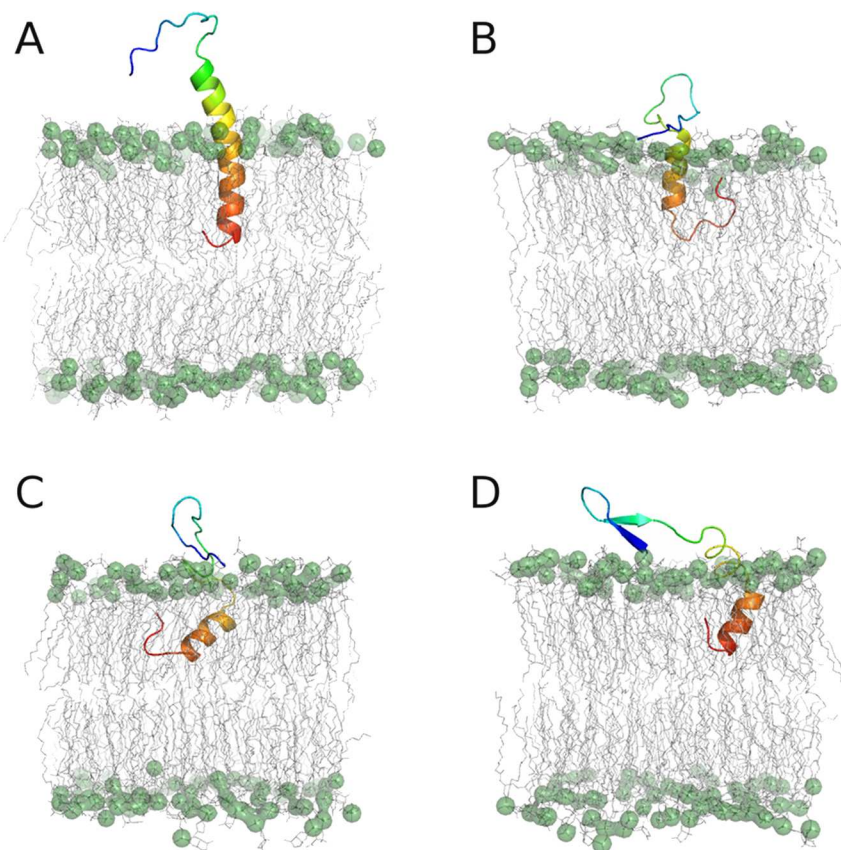


Figure 5.19: Structures of $A\beta_{40}$ model CI peptides in POPS: (A) Starting structure for the PS-CI trajectories and structures from the end of the three 100-ns PS-CI: (B) replicate 1, (C) replicate 2, and (D) replicate 3. The $A\beta_{40}$ peptide is rendered as a ribbon and colored as a gradient from N-terminus (blue) to C-terminus (red). POPS lipids are shown as gray lines, with P atoms shown as translucent green spheres. For clarity, water molecules are not shown.

may also be a contributing factor in the approach of the Raft-CI peptides to the interface. Unlike the more fluid PC, PS, and PC/PE membranes, the cholesterol-rich lipid raft was more densely packed and thus likely limiting the dynamics of $A\beta_{40}$. Our previous work [269] demonstrated a reciprocal effect of $A\beta_{40}$ on the surrounding membrane, such that nearby lipids tilted along with the peptide. In the case of an ordered raft domain, this tilting is disfavored.

In GM1-CH systems, the average position of residues 29-40 relative to the interface was at -0.7 ± 0.5 nm, an upward shift of $+0.6$ nm with respect to the Raft-CH systems. In GM1-CI systems, the average relative z -position was located at -0.3 ± 0.3 nm, a shift of $+0.3$ nm relative to the Raft-CI systems. Tilting of the C-terminal region occurred to a similar extent in GM1-containing systems as with the rafts lacking GM1, $48 \pm 18^\circ$ in GM1-CH systems, and $69 \pm 11^\circ$ in GM1-CI systems (Figures 5.15C and 5.15D). What is most notable in the simulations of the GM1-containing rafts

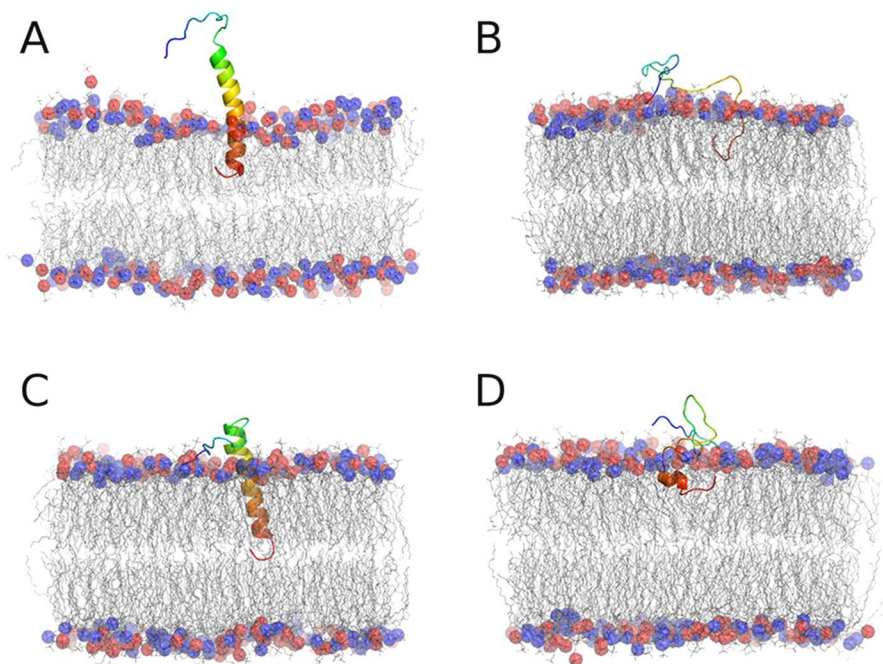


Figure 5.20: Structures of $A\beta_{40}$ model CH peptides in POPC-POPE: (A) Starting structure for the PC/PE-CH trajectories and structures from the end of the three 100-ns PC/PE-CH: (B) replicate 1, (C) replicate 2, and (D) replicate 3. The $A\beta_{40}$ peptide is rendered as a ribbon and colored as a gradient from N-terminus (blue) to C-terminus (red). Lipids are shown as gray lines, with P atoms shown as translucent red (POPC) and blue (POPE) spheres. For clarity, water molecules are not shown.

is that in one of the GM1-CH trajectories, residues 29-40 of the $A\beta$ peptide exited the membrane, establishing a free energy minimum at +0.06 nm above the membrane interface (Figure 5.15E). In two of the GM1-CI systems, $A\beta$ also exited the membrane, finding free energy minima at +0.25 nm (GM1-CI-5, Figure 5.15F) and -0.03 nm (GM1-CI-4). The free energy minimum of simulation GM1-CI-4 indicates that residues 29-40 of the peptide were effectively coincident with the membrane-water interface and spent substantial time above this interface during the trajectory, sampling configurations as high as +0.41 nm. As a result, the free energy basin from simulation GM1-CI-4 is broader, such that configurations at +0.06 nm above the interface are separated from the free energy minimum by only 1.38 kJ mol^{-1} , accessible by thermal fluctuation at 310 K ($RT = 2.58 \text{ kJ mol}^{-1}$). Snapshots from the end of all simulations and corresponding free energy surfaces are shown in Figures 5.10 and 5.11.

The contours of the free energy surfaces provide detail into the path along which $A\beta_{40}$ proceeded to the free energy minima in these trajectories. In the case of GM1-CH-5 (Figure 5.15E), the presence of GM1 caused residues 29-40 of $A\beta$ to tilt, cross the membrane interface, and reach

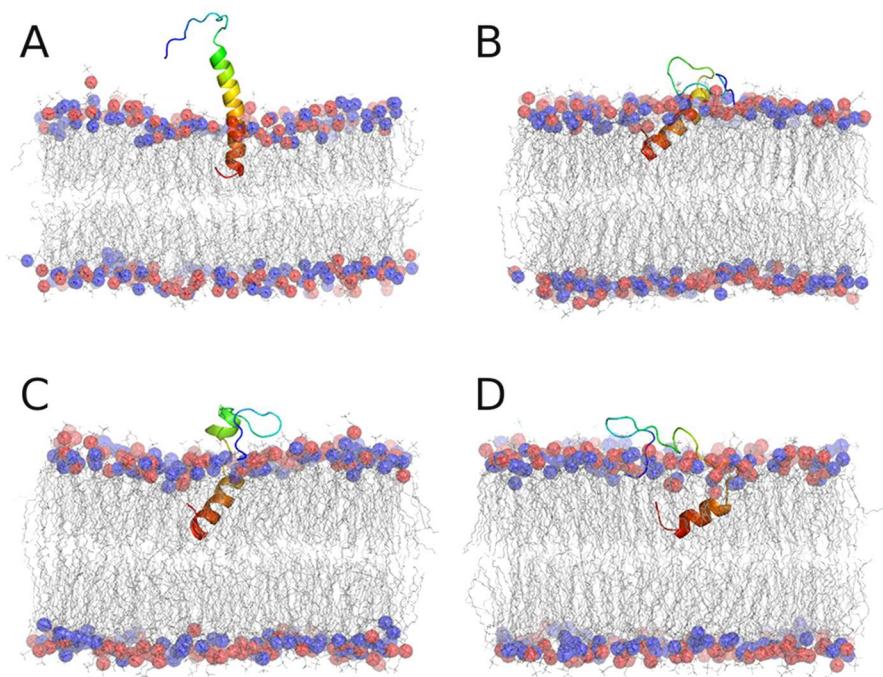


Figure 5.21: Structures of $A\beta_{40}$ model CI peptides in POPC-POPE: (A) Starting structure for the PC/PE-CI trajectories and structures from the end of the three 100-ns PC/PE-CI: (B) replicate 1, (C) replicate 2, and (D) replicate 3. The $A\beta_{40}$ peptide is rendered as a ribbon and colored as a gradient from N-terminus (blue) to C-terminus (red). Lipids are shown as gray lines, with P atoms shown as translucent red (POPC) and blue (POPE) spheres. For clarity, water molecules are not shown.

equilibrium approximately +0.06 nm above the membrane-water interface. In the case of GM1-CI-5 (Figure 5.15F), the $A\beta_{40}$ peptide rose out of the membrane before tilting, establishing its free energy minimum at +0.25 nm above the interface. In both GM1-CH-5 and GM1-CI-5, the backbone of hydrophobic amino acids in $A\beta$ formed hydrogen bonds with nearby GM1 molecules. In GM1-CH-5, the backbone of Gly38 and Val39 interacted with Glc and the ceramide backbone of GM1. Upon forming these hydrogen bonds, the C-terminal residues tilted up towards the interface, drawing this region of the peptide out of the membrane. In GM1-CI-5, the backbone of $G_{33}LM_{35}$ hydrogen bonded with Glc and Neu5Ac, pulling the peptide up before residues $V_{36}GGVV_{40}$ tilted at the membrane-water interface. Despite subtle differences in the progression of these two trajectories, a common feature emerges: hydrogen bonding groups, both donors and acceptors, on GM1 in the vicinity of the membrane-embedded residues of $A\beta$ compete for native backbone hydrogen bonding and allow hydrophobic residues 29-40 to approach, and ultimately cross, the membrane-water interface. The oligosaccharide headgroup of GM1 sequesters the N-terminal residues of $A\beta$ away from the membrane-water interface, blocking interactions (like those observed in the POPC, POPS, and POPC/POPE systems studied here) that otherwise anchor $A\beta$ in the membrane. We

note that in simulation PC/PE-CH-3, the presence of a hydrogen bond donor group (a primary amine) induced the formation of a β -hairpin, but this group did not enable A β_{40} to escape the membrane. Thus, it appears that GM1 possesses a unique ability to facilitate this process.

5.4 Discussion

All previous MD studies of A β in membranes [199, 205, 247, 269, 287, 289, 290] have consisted of only a single lipid type or an implicit model representing a membrane. In the present work, however, we have explored numerous explicit model membranes, including the most complex lipid environments in which A β has been simulated to date, rafts that correspond very closely to the lipid matrix that A β encounters upon its production following γ -secretase cleavage of APP. These simulations provide insight into the behavior of A β during a time period for which experimental evidence is limited, but critical for understanding the early events in the development of Alzheimer's disease.

A previous MD study conducted on A β_{40} in a model DPPC bilayer showed that A β could spontaneously exit the membrane environment [199], but those findings are in direct contrast to experimental [208, 234] and theoretical [205, 247, 287] demonstrations that A β_{40} remains partially inserted in such a lipid environment, with residues 29-40 embedded in the hydrophobic core of the bilayer. The discrepancy has been attributed to the use of better-quality force field models in the more recent studies [287]. Using this improved force field model, we were able to reproduce experimental behavior of A β_{40} in DPPC [247, 269]. We have applied that same force field model in the present study. It is important to note, however, that all force fields have some inherent limitations and biases that must be taken into account [196, 197], and that studies on structurally dynamic proteins such as A β_{40} require great care in interpreting results. We note that, despite our efforts to comprehensively sample conformations that A β_{40} may adopt through relatively long simulations (100 ns) and several replicates, our data reflect the considerable heterogeneity intrinsic to the peptide. Thus, we focus on broader aspects of our observations while still considering the fact that the current limitations of atomistic simulations may prevent us from obtaining a complete understanding of structural transitions that occur over much longer time scales *in vitro* and *in vivo*.

In all of the simple membrane models explored here (POPC, POPS, and POPC/POPE), A β_{40} remained embedded within the membrane, with residues 29-40 well below the membrane-water interface. Ionization of C-terminal Val40 tended to cause some destabilization of the initial α -helical configuration of this segment of the peptide as Val40 snorkeled towards the interface. This behavior was observed regardless of the lipid composition of the surrounding membrane. In POPS, the ability of Val40 to approach the interface was diminished due to charge repulsion between the anionic lipids and the ionized C-terminus. Snorkeling contributed to the tilt of the embedded C-terminal region, a behavior that has been suggested to have implications for the toxicity of A β exerted through its interactions with membranes [269, 275, 288].

The main secondary structural elements of A β_{40} in each of the simple model membranes were

helices and random coils, though short, transient β -hairpins were observed, principally within its N-terminal polar region. The POPC, POPS, and POPC/POPE membranes did not significantly stabilize long β -strands, though it appears that the membrane environment may be conducive to the formation of such structures. Furthermore, these β -hairpin structures involved many different residues within the N-terminal sequence and were variable in terms of their location and spacing by intervening bends or turns. Thus it appears that PC, PS, and PE lipids induce nonspecific β -strand formation.

In POPC, POPS, and POPC/POPE, A β_{40} anchored itself to the membrane through hydrogen bonding and other electrostatic interactions, in agreement with our previous findings regarding A β_{40} in DPPC [247, 269]. The increased hydrogen bonding between A β_{40} and POPS and POPC/POPE relative to POPC may explain why A β binds more tightly to these lipids than PC-containing lipids [234]. In POPS and POPC/POPE, there are simply more functional groups available to form hydrogen bonds with A β_{40} .

Clusters of ganglioside GM1 create a polar, dehydrated environment that extends more than 2 nm above the membrane-water interface. Experiments have shown that dehydrated environments facilitate amyloid aggregation by reducing the free energy barrier of desolvating the membrane interface [291]. The oligosaccharide headgroup of GM1 possesses hydrogen bond donors and acceptors that compete for protein-protein hydrogen bonds and destabilize native α -helical structures, a behavior that accounts for the increased rate of β -strand formation observed experimentally [291]. In our simulations, the formation of β -hairpins was stabilized by specific interactions between A β residues (His6, Asp7, His13, and His14) and GM1, most commonly Neu5Ac and Gal2, sugar moieties that have previously been predicted to be important in A β -GM1 binding [281]. The sequestration of the polar N-terminus of A β away from the membrane-water interface, within the oligosaccharide headgroup of GM1 clusters, precluded non-specific electrostatic interactions (like those of the POPC, POPS, and POPC/POPE systems) that otherwise anchor A β to the membrane [269]. Our results indicate that GM1 clusters are required for this behavior. In simulations of A β_{40} in rafts containing only one GM1 molecule at random locations, A β_{40} bound to GM1, but the N-terminal region was not sequestered above the membrane-water interface. As a result, A β_{40} adopted positions, orientations, and secondary structure content that was indistinguishable from the results of raft-only simulations (data not shown).

The ability of GM1 to promote the release of A β_{40} from a lipid raft was a result of the presence of hydrogen-bonding groups near the membrane-water interface. Interactions between the peptide backbone of residues 29-40 and the Glc and Neu5Ac moieties of GM1 can lead to disordering of the peptide and transport out of the hydrophobic lipid matrix. The numerous hydrogen bond donors and acceptors in the oligosaccharide headgroup of GM1 can effectively compete for the native intra-protein hydrogen bonding in residues 29-40, leading to destabilization of the helix and exit from the membrane. The phospholipid headgroups in rafts lacking GM1 have only hydrogen bond acceptors (phosphate groups), which, as demonstrated here, interacted with A β_{40} in a manner that did not lead to efficient release of A β_{40} from the membrane. From the present study and our previous work [247], it appears that the C-terminal residues of A β_{40} have some intrinsic capacity to approach, but not cross, the polar interface region, a process that was influenced to some degree

by the protonation state of the C-terminal Val40. This effect may be enhanced in rigid, densely packed raft domains.

In general, model CI peptides approached the membrane-water interface more closely than did model CH peptides. Deprotonated peptides likely have a greater affinity for the polar interfacial environment, contributing to their stability in this region and eventual release from the membrane. Model CH peptides were capable of considerable C-terminal tilting in all membranes examined here, placing Val40 close to the interface, indicating that proton exchange between the peptide and any water molecules at the interface may be possible. Thus, a dynamic equilibrium between protonated and deprotonated C-termini may exist *in vivo*, but this is not a phenomenon that can directly be addressed using MD simulations, as bonds cannot break or form under classical mechanics assumptions.

The embedded C-terminal residues lost some of their initial α -helicity as they moved away from the hydrophobic core of the membrane and became more proximal to the polar interfacial environment. Our data indicate that GM1 facilitates this process for both CH and CI models of the A β_{40} peptide, enhancing the release of A β_{40} from its native lipid raft environment and promoting the exposure of hydrophobic, aggregation-prone residues to the extracellular environment.

5.5 Conclusions

There is an apparent paradox in the amyloid hypothesis. If A β exerts its toxicity by stably binding to, and inserting in lipid membranes, how does A β leave the membrane in the first place and gain access to the aqueous extracellular space? Our results provide insight into this mechanism and a very specific lipid that may resolve this issue. A β -GM1 interactions promote the structural conversion of A β from α -helix to β -strand, while simultaneously providing a scaffold that allows the C-terminal hydrophobic region of A β (residues 29-40) to unwind, tilt upward, and exit the membrane, a behavior that we observed in several simulations described here. Membrane exit was not observed in any of the simulations of A β_{40} in POPC, POPS, or POPC/POPE. Rather, the stability of A β in these environments may shed light on the ability of A β to re-insert into membranes and exert its toxic, membrane-destabilizing effects.

The behavior described here may provide insight into the earliest events in the A β aggregation cascade, the release of the A β peptide following its production in lipid rafts. It should be noted that our results do not preclude the possibility that traumatic brain injury or other neuronal damage might also contribute to the increased secretion of A β , as has been suggested previously [292, 293]. Prior damage to the plasma membrane may indeed release large amounts of A β and initiate the aggregation cascade in a different manner than that which is observed in otherwise healthy individuals. The mechanism we propose here may be useful for designing small molecule inhibitors of A β -GM1 interactions, preventing the formation of aggregation-prone A β conformations and the release of A β from the plasma membrane during the earliest stages of Alzheimer's disease.

5.6 Acknowledgments

We thank Prof. Mikko Karttunen for providing lipid raft coordinates and topologies, Prof. D. Peter Tieleman for providing coordinates and topologies for the POPS membrane, and Sukit Leekumjorn for providing coordinates and topologies for the POPC/POPE membrane.

Chapter 6

Interactions Between Multimeric $A\beta$ and Lipid Bilayers

Copyright statement: The contents of this chapter are currently in preparation for publication.

Attribution: J. A. L. and D. R. B. designed the research. J. A. L. conducted the simulations and performed analysis. J. A. L. and D. R. B. wrote the paper.

6.1 Introduction

In the present work, we sought to expand upon the current understanding of $A\beta$ -membrane interactions in the context of peptide dimerization. We posed several questions. First, do $A\beta_{40}$ monomers form dimers within the membrane, and if so, is there any dependence on lipid composition in this process? Second, what is the nature of the interpeptide interactions, in terms of both residues involved in dimerization and any secondary structure changes? Finally, does $A\beta_{40}$ dimerization adversely affect the structure of the lipid membrane? The application of atomistic MD simulations allows us to probe these events with great detail.

6.2 Methods

The starting configuration for the $A\beta_{40}$ peptide was taken from work by Coles *et al.* [77], which we have used in our previous studies on $A\beta$ -membrane interactions [247, 269, 294]. The coordinates and topologies for the membranes used in this work (POPC, POPS, POPC/POPE, and rafts) were taken from previous studies [228, 244, 264, 265, 266], with coordinate files expanded as needed in the $x - y$ plane to produce membranes of suitable dimensions for the insertion of two peptides. Contents and dimensions of the simulation systems are given in Table 6.1. The force field applied

to the A β_{40} peptide was GROMOS96 53A6 [174]. These peptides were placed diagonally from one another in the $x - y$ plane of the membrane with a center-of-mass (COM) distance of 3.3 nm and a minimum distance between any pair of atoms of approximately 1.0 nm (Figure 6.1). These proteins were inserted into the membranes using the InflateGRO methodology [267].

Table 6.1: Contents of the simulated systems.

System	Lipid Content	Dimensions (nm)*
POPC-CH/CI	250 POPC	$8.75 \times 8.75 \times 12$
POPS-CH/CI	246 POPS	$8.5 \times 8.5 \times 12$
POPC/POPE-CH/CI	140 POPC, 140 POPE	$9.94 \times 9.94 \times 12$
Raft-CH/CI	127 POPC, 113 PSM, 121 cholesterol	$8.73 \times 8.73 \times 12$
GM1-Raft-CH/CI	125 POPC, 92 PSM, 119 cholesterol, 24 GM1	$8.75 \times 8.75 \times 12$

*Dimensions given are those of the energy-minimized systems, i.e. prior to equilibration or further simulation.

All further preparation steps and simulations were carried out using the GROMACS package [180], version 4.0.7. Following insertion into the lipid membrane, the unit cell was filled with SPC water [268] and ~ 150 mM NaCl, including 6 Na⁺ counterions to balance the net charge on each of the A β_{40} peptides. Each system was energy-minimized using the steepest descent algorithm and was subsequently equilibrated in three phases. During all phases of equilibration, position restraints (using $k_{pr} = 1000$ kJ mol⁻¹ nm⁻²) were applied to all protein heavy atoms. For the first phase of equilibration, an NVT ensemble was applied and the system equilibrated at 100 K. The Berendsen weak coupling algorithm [187] was used to control the temperature of the system. The peptides, lipids, and solvent (including ions) were coupled separately. NVT equilibration was carried out for 100 ps. Following NVT , each system was warmed from 100 K to 310 K linearly over 1 ns using an NPT ensemble at 1 bar of pressure. During this process, the Berendsen algorithm [187] was used to control both temperature and pressure. For the final phase of equilibration, an NPT ensemble was maintained at 310 K and 1 bar using the Nosé-Hoover thermostat [188, 189, 190] and Parrinello-Rahman barostat [191, 192]. Production MD simulations were carried out for 200 ns using this same ensemble, in the absence of any restraints. Three replicate simulations were conducted for each system, initiated using different random starting velocities at the outset of NVT equilibration.

Simulations were carried out with constraints on all bonds using the P-LINCS algorithm [270], allowing an integration time step of 2 fs. Long-range electrostatics were calculated using the smooth particle mesh Ewald (PME) method [193, 194]. All short-range nonbonded interactions (including van der Waals terms and the real-space contribution to PME) were truncated at 1.2 nm. The neighbor list was updated every 5 simulation steps. Dispersion correction was applied to energy and pressure terms to account for finite truncation of van der Waals interactions.

Analysis was conducted using programs within GROMACS or with our own software, GridMAT-MD [239]. Secondary structure was assigned using the DSSP algorithm [271], as implemented in

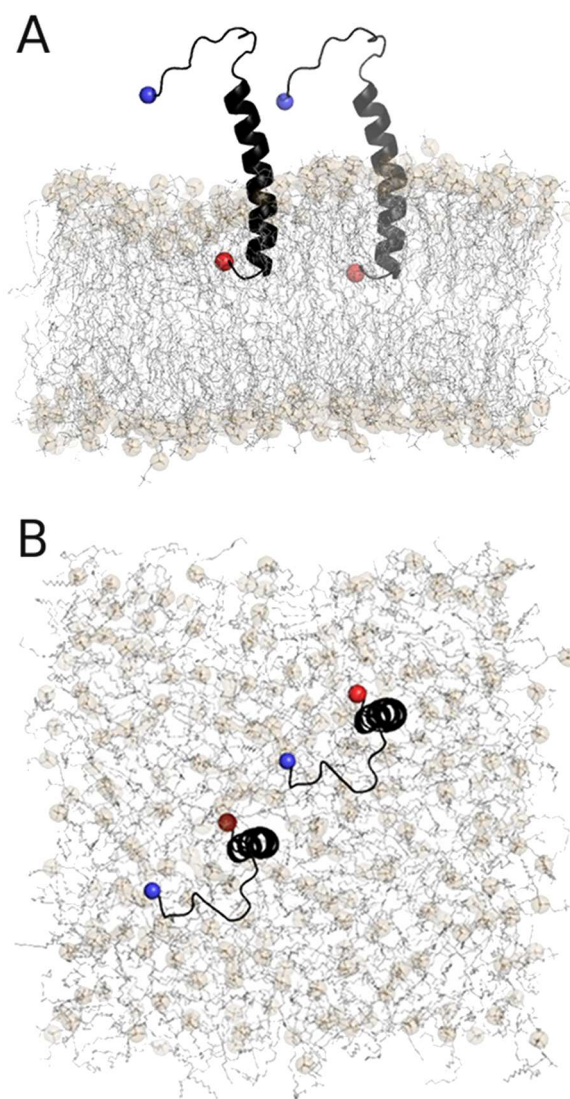


Figure 6.1: The starting configuration of the A β_{40} -POPC dimer system, as viewed (A) through the $x - y$ plane and (B) along the z -axis. The N- (blue) and C-termini (red) are indicated by spheres for perspective. The peptides are shown as black ribbons. Lipids are shown as gray lines with phosphorus atoms indicated by translucent gold spheres.

GROMACS.

Throughout this paper, simulations are referred to by the lipid type (POPC, POPS, POPC/POPE, Raft, and GM1-Raft) and the C-terminal protonation state of the A β_{40} peptide. Thus, peptides with protonated C-termini are called “CH” and those with ionized C-termini are called “CI.” Both protonation states were considered for reasons described previously [247, 294]. Individual simulations within a set are assigned a numeral.

6.3 Results and Discussion

Of interest in this work were several points related to both A β aggregation and the effects this process may have on the structure of the surrounding membrane and the dynamics of the peptides themselves. Thus, we analyze several points in turn: (i) dimer formation, (ii) position and orientation of A β_{40} within the membrane, (iii) secondary structure of A β_{40} , and (iv) membrane structure.

We begin the discussion of our findings with several important caveats. The starting structure for A β_{40} was determined in the presence of SDS micelles [77], and as such may not correspond directly to the structure of the peptide *in vivo*. We do believe, however, that this structure is a reasonable approximation of the structure of A β_{40} upon its release from the γ -secretase complex, as the domain of APP containing the A β sequence is a single-pass transmembrane helix, and thus it is likely that much of the A β_{40} structure is helical upon its generation. The initial placement of the two A β_{40} peptides in the systems constructed here is also arbitrary; an infinite set of configurations could be generated for the starting structure based on different rotational and translational transformations to the peptide structures. The starting configuration used here was a simple construct that enabled us to build a system of minimal size, to allow for longer simulations in a reasonable amount of time. As a final note, it is important to state that the timeframe explored in each of these simulations (200 ns) is short by physiological standards, but very long by computational metrics. Recent work by Zhao *et al.* [295] simulated an A β trimer for 1 μ s, but they conducted only one simulation. Our approach of using multiple 200-ns simulations provides greater statistical confidence in the results we have obtained.

6.3.1 Dimerization of A β_{40} Peptides

We defined a peptide dimer as occurring when a minimum of 100 interpeptide heavy atom contacts persisted over the final 100 ns of each simulation, and the $x - y$ COM distance between the peptides was within 2.75 nm. Simulations in which the A β_{40} peptides satisfied both of these criteria were called “strong” dimers; in cases where the contact criterion was met but the COM distance criterion was not, the resulting dimers were called “weak.” The latter case represents the situations in which the A β_{40} peptides adopted elongated, disordered structures that interacted over great distance. Strong dimers formed at close distance and established a greater number of contacts. There were no instances in which the distance criterion was satisfied but the contact criterion was not. Examples are shown in Figure 6.2.

The simulations of A β_{40} in POPC produced four dimers out of the six total simulations, three strong (all in the POPC-CH set) and one weak dimer in the POPC-CI set. Simulations of A β_{40} in POPS produced only two dimers, one strong (in the POPS-CI set) and one weak (in POPS-CH). In the POPC/POPE-CI set, two strong dimers formed, but no dimers were observed in the POPC/POPE-CH set. Thus, for these three simple model membranes, it appears that the ability of A β_{40} to dimerize can be expressed as a function of hydrogen bonding potential with the sur-

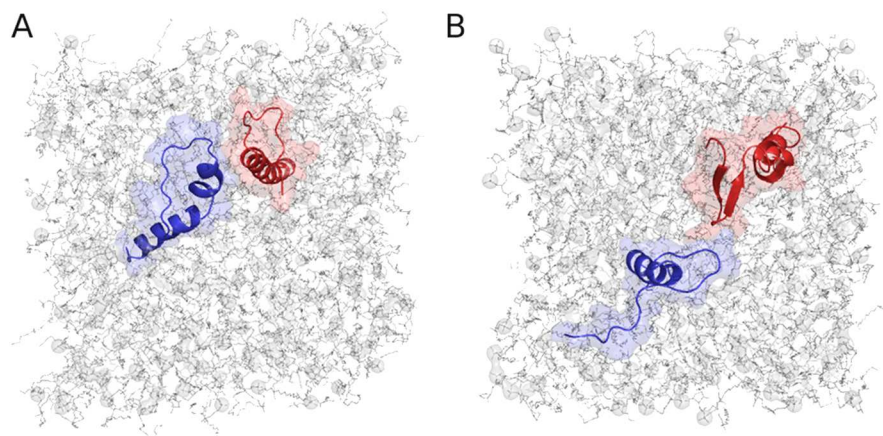


Figure 6.2: Examples of (A) a “strong” dimer, from the POPC-CH-3 trajectory, and (B) a “weak” dimer, from the POPS-CH-3 trajectory. Each peptide is shown with an overlaid translucent surface to give an impression of the overall shape of the molecules. Lipids are shown as gray lines, with phosphorus atoms indicated by translucent gray spheres.

rounding lipids. Lipids with greater numbers of hydrogen bonding groups such as POPS impeded $A\beta_{40}$ dimerization to the greatest extent. The presence of POPE in the POPC/POPE membrane also impeded dimerization, but not as strongly as in the case of POPS. $A\beta$ -lipid hydrogen bonds for the different systems simulated here are summarized in Table 6.2.

The $A\beta_{40}$ peptides formed numerous dimers in the raft environment without GM1. In the Raft-CH simulations, two strong dimers formed in the three replicate simulations. In the Raft-CI set, all three simulations produced strong dimers. In the GM1-Raft systems, the GM1-Raft-CH set produced one strong dimer, and the GM1-Raft-CI set produced two strong dimers. These results are not unexpected, given the fact that most of the lipids in the raft (both POPC and PSM) contained PC headgroups. The $A\beta_{40}$ peptides formed comparable numbers of hydrogen bonds with phospholipids in the POPC, Raft, and GM1-Raft simulations. Thus, the mobility of these peptides and their potential for interaction was comparable in these environments, allowing dimers to form with a greater frequency and strength than in the POPS and POPC/POPE systems. Fewer dimers were observed in the GM1-Raft systems relative to the Raft systems, likely due to the extensive interactions that emerged between $A\beta_{40}$ and GM1.

The interactions formed between $A\beta_{40}$ peptides upon dimerization were principally between residues in the N-terminal regions of the peptide. The C-terminal hydrophobic residues generally remained far apart, forming few, if any, direct contacts within the membrane core. To quantify these interactions, distance matrices were generated over the last 100 ns of each simulation (Figure 6.3). These findings are in good agreement with recent experiments conducted by Zhang *et al.* [296], who found that $A\beta$ aggregation in membranes was driven principally by residues 11-16 of the $A\beta$ sequence. They further determined that the hydrophobic C-terminal residues, while important for aggregation of $A\beta$ in the extracellular medium, were not involved in aggregation within the

Table 6.2: A β_{40} -lipid hydrogen bonds. Data are averaged over the final 100 ns of each trajectory, considering both peptides in all simulation sets. Shown are the resulting averages with standard deviations.

System	POPC	POPS	POPE	PSM	Chol	GM1
POPC-CH	12.7 \pm 3.0	-	-	-	-	-
POPC-CI	16.6 \pm 4.9	-	-	-	-	-
POPS-CH	-	22.8 \pm 6.6	-	-	-	-
POPS-CI	-	25.6 \pm 3.5	-	-	-	-
POPC/POPE-CH	8.9 \pm 2.7	-	13.4 \pm 3.6	-	-	-
POPC/POPE-CI	8.6 \pm 3.8	-	12.7 \pm 3.4	-	-	-
Raft-CH	5.7 \pm 1.8	-	-	6.2 \pm 2.2	0.7 \pm 0.6	-
Raft-CI	6.6 \pm 4.2	-	-	6.0 \pm 2.8	0.8 \pm 0.9	-
GM1-Raft-CH	6.6 \pm 2.4	-	-	2.3 \pm 1.5	0.3 \pm 0.4	12.0 \pm 5.8
GM1-Raft-CI	7.0 \pm 2.0	-	-	1.5 \pm 1.2	0.8 \pm 1.4	12.5 \pm 3.8

membrane, rather, these residues served to anchor the peptide in the membrane.

6.3.2 Position and Orientation of A β_{40} in Membranes

In our previous work [294], we found that monomeric A β_{40} interacted with GM1 in a manner that promoted the peptide to exit the membrane, a behavior that was not observed in the simulations of A β_{40} in the other model membranes. Thus, it was of interest in the present work to determine if the ability of A β_{40} to dimerize affected the stability of A β_{40} in the membrane. We previously defined an “exit” from the membrane as occurring when the COM of residues 29-40 of A β_{40} (initially embedded in the membrane) crossed above the average phosphate plane of the extracellular leaflet of the membrane. We utilize that same definition here. We have quantified the position and orientation of the A β_{40} peptide within each model membrane using two criteria: (i) the tilt angle of the principal axis through residues 29-40 and the z -axis, and (ii) the distance between the COM of residues 29-40 and the average phosphate plane. Results of these calculations are summarized in Table 6.3.

The values observed for tilt angle and relative position are very similar to our previous results for monomeric A β_{40} in these same membranes [294], indicating that the presence of multiple peptides does not affect the intrinsic ability of A β_{40} to tilt and rise towards the membrane-water interface. In these simulations, none of the A β_{40} peptides exited the membrane. Our previous work [247, 294] has suggested that A β_{40} should remain stably embedded in a variety of membranes, but that the presence of GM1 enhances the ability of A β_{40} to exit from raft membranes. We do not observe such behavior here in the context of the GM1-Raft simulations. We attribute this finding to the fact that there is only a small probability that exit will occur (30% in the previous work), and that

Table 6.3: Tilt angle and z -axis position of A β_{40} peptides in model membranes. Averages (with standard deviations) were calculated for both peptides in each system, averaged over the final 100 ns of each trajectory.

System	Tilt Angle (degrees)	Relative Position (nm)
POPC-CH	64 ± 41	-1.5 ± 0.5
POPC-CI	96 ± 14	-0.9 ± 0.2
POPS-CH	17 ± 3	-1.5 ± 0.2
POPS-CI	87 ± 16	-1.1 ± 0.2
POPC/POPE-CH	62 ± 25	-0.8 ± 0.1
POPC/POPE-CI	78 ± 44	-0.5 ± 0.3
Raft-CH	59 ± 26	-1.0 ± 0.3
Raft-CI	71 ± 20	-0.4 ± 0.2
GM1-Raft-CH	66 ± 16	-0.7 ± 0.4
GM1-Raft-CI	81 ± 26	-0.5 ± 0.3

dimerization should impede the ability of A β to exit the membrane, as A β monomers preferentially aggregate in the membrane rather than exit into the extracellular space [296]. We observed strong dimer formation in 50% of the simulations conducted here (one in the GM1-Raft-CH set and two in the GM1-Raft-CI set). Thus, our previous observations and those presented here are not in conflict. Whereas monomeric A β_{40} may leave the membrane via interactions with GM1 and can lead to extracellular aggregation, the presence of multiple A β_{40} peptides disfavors membrane release and instead leads to aggregation within the membrane.

6.3.3 Secondary Structure of A β_{40} Peptides

Previous structural and simulation studies of A β in membranes and membrane-mimicking environments have concluded that the peptide is principally helical, though some regions are disordered [77, 78, 297]. Our own results of monomeric A β_{40} have shown that transient β -strand formation is observed in all membrane environments, but β -strand content is enhanced by the presence of GM1 clusters [294]. These GM1 clusters provide a polar, dehydrated environment that facilitates the formation of β -hairpin structures, principally within in the N-terminal region of the peptide. Secondary structure content for the A β_{40} peptides in the simulations conducted here is summarized in Table 6.4.

In agreement with previous experimental and simulation data [77, 78, 199, 294, 297], the principal secondary structure elements of the A β_{40} peptides were random coil and helix. We consider total helical content here, as did in our previous work [294], rather than individual α -helix, π -helix, and 3_{10} -helix elements, as the definition and balance of these structures is not well defined for most force fields [196, 197, 298]. β -strand content was highest for the simulations of A β_{40} in rafts

Table 6.4: Secondary structure content of A β_{40} peptides, averaged over the last 100 ns of simulation time.

System	Coil	β -strand	Bend	Turn	Total Helix
POPC-CH	35 \pm 5	1 \pm 2	17 \pm 7	11 \pm 10	35 \pm 14
POPC-CI	48 \pm 6	4 \pm 5	18.2 \pm 0.9	11 \pm 2	19 \pm 1
POPS-CH	32 \pm 7	3 \pm 5	15 \pm 2	15 \pm 2	36 \pm 7
POPS-CI	40 \pm 6	3 \pm 2	23 \pm 5	13 \pm 3	22 \pm 5
POPC/POPE-CH	41 \pm 2	3 \pm 1	21 \pm 2	16 \pm 9	15 \pm 4
POPC/POPE-CI	41 \pm 14	1.5 \pm 0.8	24 \pm 2	13 \pm 8	20 \pm 10
Raft-CH	35 \pm 3	2 \pm 1	19 \pm 4	13 \pm 3	32 \pm 2
Raft-CI	49 \pm 5	6 \pm 2	20 \pm 7	15 \pm 4	10 \pm 5
GM1-Raft-CH	39 \pm 1	9 \pm 1	21 \pm 5	14 \pm 5	17 \pm 1
GM1-Raft-CI	44 \pm 4	4 \pm 4	25 \pm 3	17 \pm 2	10 \pm 6

containing GM1, in agreement with our previous results [294]. In all six of the simulations of A β_{40} in GM1-containing rafts, short β -hairpins emerged at various locations within the A β_{40} structures.

The fact that the secondary structure content of multiple A β_{40} peptides, including those that have dimerized, is similar to those in the monomeric state explored in our previous work suggests that A β_{40} dimers do not form extensive β -strand structures and remain principally disordered at this stage in the aggregation cascade. Structural studies have concluded that low-molecular weight A β aggregates are largely unstructured in solution, forming extensive β -strand structures in higher-order species [81]. Our results indicate that the same holds true for dimers of A β_{40} formed within the membrane. Emergence of more elongated β -strands and intermolecular β -sheets may require the presence of a larger number of A β peptides, a topic that bears further examination in the future.

6.3.4 Effect of A β_{40} on Membrane Structure

A central tenet of the “amyloid hypothesis” of Alzheimer’s disease is the ability of A β to cause structural destabilization of cell membranes that ultimately leads to cell death. We sought to determine if the presence of multiple A β_{40} peptides, dimerized or not, could impart structural changes on the model membranes considered here. We quantified the membrane thickness along the z -axis for each system using GridMAT-MD [239]. Results are shown in Figure 6.4.

The most notable deformations occurred in the case of the POPC membranes. In the vicinity of the A β_{40} peptides, the membrane thickness was reduced in excess of 1.0 nm from the remainder of the membrane at more distant locations. Such behavior is similar to our observations in the case of A β_{40} in DPPC [247]. Several of the simulations in the POPC/POPE-CH and POPC/POPE-CI systems produced similar results, though the thinned area is considerably smaller and concentrated

to the immediate vicinity around the A β_{40} peptides than in the case of pure POPC. The other membranes remained relatively unperturbed. We attribute these findings to two factors: (i) the ability of different lipids to form hydrogen bonds, and (ii) the presence of cholesterol in the Raft and GM1-Raft simulation sets.

Relative to the simulations of A β_{40} in POPC, the simulations of A β_{40} in POPC/POPE produced less pronounced thinning, and the simulations in POPS produced no noticeable thinning. The phosphatidylethanolamine headgroup of the POPE lipids is characterized by a primary amine capable of hydrogen bonding to A β_{40} (Table 6.2), POPC, or other POPE lipids. Whereas A β_{40} tilted and caused POPC lipids to disorder, leading to bilayer thinning, the intermolecular interactions between POPC and POPE lipids made this membrane less susceptible to such structural change, as was also the case for the POPS membrane. The phosphatidylserine headgroup contains two hydrogen bond acceptors (phosphate and carboxylate) and a donor (primary amine), thus giving it the greatest per-lipid hydrogen bonding capacity of any of the lipids examined here. The structure of the POPS membranes remained very stable over the course of the simulations carried out here, indicating that greater hydrogen bonding capacity within the lipids can reduce the structural destabilization imparted by A β_{40} .

Previous work by Qiu *et al.* has suggested that the presence of cholesterol in the membrane acts to protect the bilayer from A β toxicity [299]. In our simulations in the Raft and GM1-Raft simulation sets, cholesterol was present in a nearly equimolar ratio with the phospholipids. In cell membranes, raft domains are more rigid and tightly packed than other regions, and thus may be less susceptible to structural changes induced by the presence of A β . Thus, it appears that raft domains serve dual functions in the context of A β . The first was proposed in our earlier work [294], in which we found that monomeric A β_{40} could exit GM1-containing rafts. The second function, illustrated here, may be to concentrate A β , as we observed extensive formation of strong dimers in our simulations. These dimers may then diffuse to other regions of the cell membrane rich in phosphatidylcholine-containing lipids and lacking cholesterol in order to destabilize the cell membrane.

6.4 Conclusions

We have conducted an extensive analysis of A β_{40} dimerization in a variety of physiologically relevant model membranes. We find that A β_{40} can form dimers within the membrane, and that the interactions between the peptides that lead to the formation of these structures occur exclusively within the N-terminal polar region of the peptide sequence. Secondary structure content in the dimerized species was comparable to that of monomeric A β_{40} [294], indicating that more extensive β -strand formation likely only occurs in higher-order aggregates. Unlike the A β_{40} monomer, dimeric A β_{40} does not exit any of the model membranes examined here, in agreement with experimental results that suggest A β aggregation within membranes is favored over the monomeric state in solution [296]. We further find that peptide-induced membrane deformation is strongly dependent upon the lipid content of the model membrane. Whereas the POPC membrane was very

susceptible to A β_{40} -induced defects, lipids that are capable of more extensive inter-lipid hydrogen bonding (such as POPS and POPE) ameliorated this effect. Further, cholesterol-rich rafts were less susceptible to such deformations as well, in agreement with the results of Qiu *et al.* [299]. The present work has provided a greater understanding of the molecular events that lead to the aggregation of A β_{40} in physiological membranes, as well as greater detail into the membrane domains that are most susceptible to A β -induced toxicity.

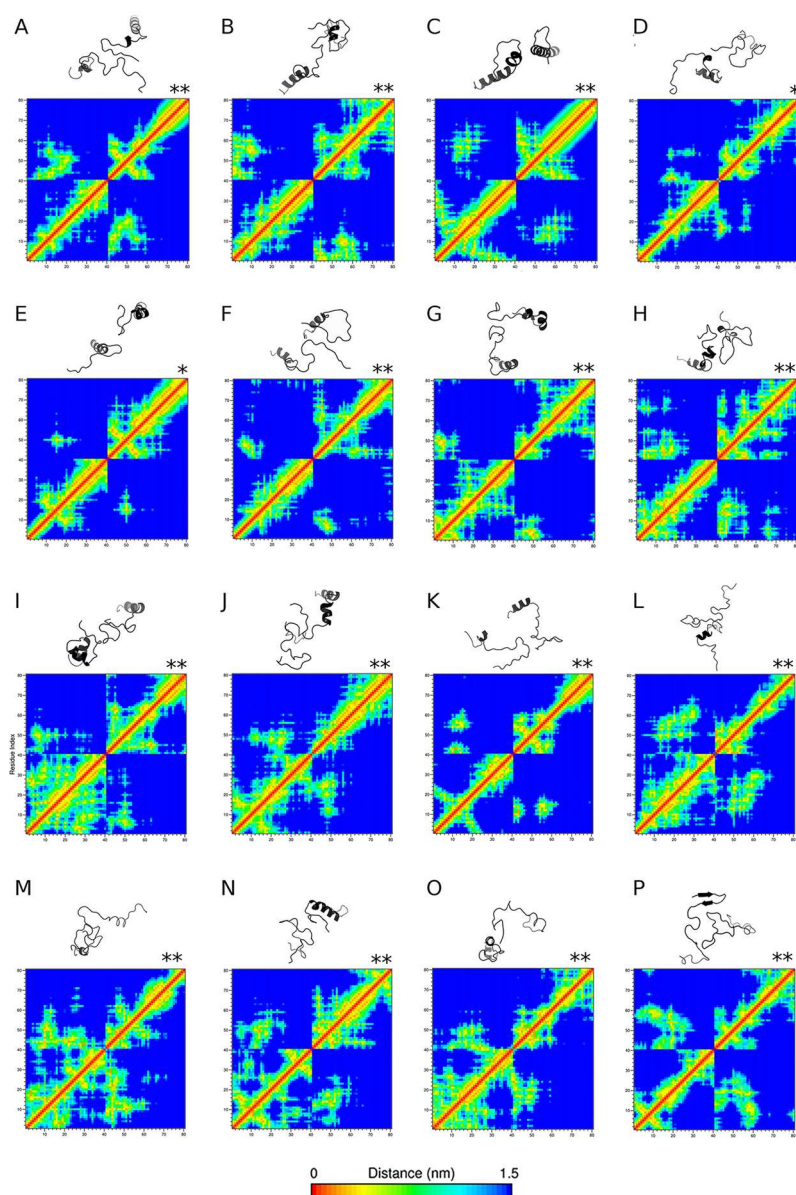


Figure 6.3: Distance matrices and images of $A\beta_{40}$ for all trajectories in which dimers formed: (A) POPC-CH-1, (B) POPC-CH-2, (C) POPC-CH-3, (D) POPC-CI-1, (E) POPS-CH-3, (F) POPS-CI-2, (G) POPC/POPE-CI-1, (H) POPC/POPE-CI-3, (I) Raft-CH-1, (J) Raft-CH-3, (K) Raft-CI-1, (L) Raft-CI-2, (M) Raft-CI-3, (N) GM1-Raft-CH-3, (O) GM1-Raft-CI-2, and (P) GM1-Raft-CI-3. Distance matrices represent the mean smallest distance between all residue pairs over the final 100 ns of each trajectory. Residue numbers are given on the x - and y -axes (peptide 1, residues 1-40 and peptide 2, residues 41-80). Peptide structures are taken from a representative snapshot at the end of each trajectory. Strong dimers are indicated by a double asterisk (**) and weak dimers are indicated by a single asterisk (*).

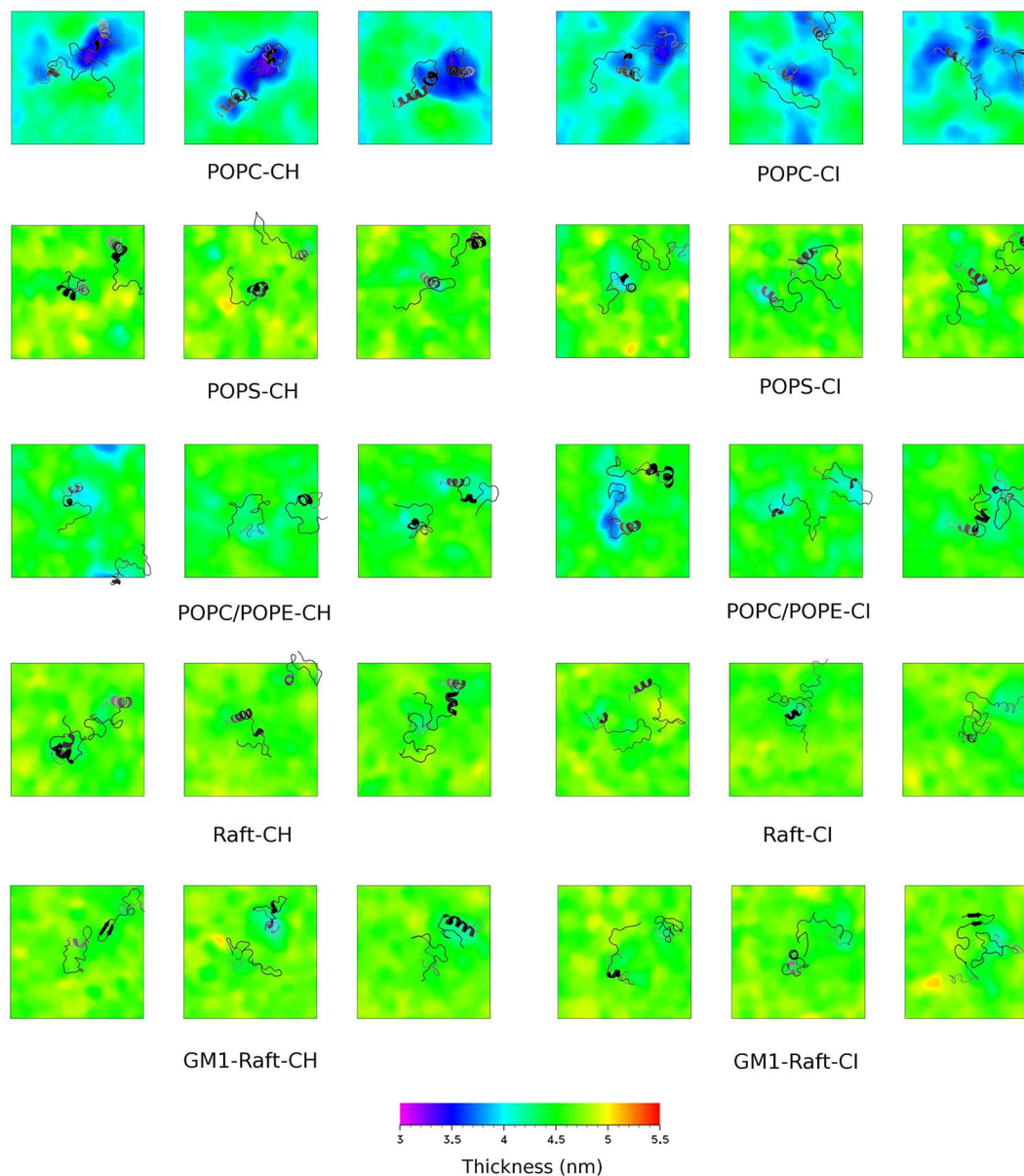


Figure 6.4: Membrane thickness projections in the $x-y$ plane. Thicknesses were calculated as the average over the last 50 ns of each trajectory, using snapshots every 1 ns. To define the membrane boundaries, phosphate atoms were used as reference atoms for the GridMAT-MD calculations. $A\beta_{40}$ structures are taken from a representative configuration from the end of each trajectory and are colored as a black to gray gradient to imply depth within the membrane. Black regions indicate regions of the peptide near the membrane-water interface, while gray residues are within the bilayer interior.

Chapter 7

Thermodynamics of A β Protofibril Assembly

Copyright statement: The contents of this chapter are reprinted with permission from J. A. Lemkul and D. R. Bevan (2010) “Assessing the Stability of Alzheimer’s Amyloid Protofibrils Using Molecular Dynamics.” *J. Phys. Chem. B* **114** (4): 1652-1660.

<http://pubs.acs.org/articlesonrequest/AOR-8EXDc22FjhYcZS7F3dAM> Copyright (2010) American Chemical Society.

Attribution: J. A. L. and D. R. B. designed the research. J. A. L. conducted the simulations and performed analysis. J. A. L. and D. R. B. wrote the paper.

7.1 Introduction

Amyloid fibrils represent a thermodynamically-stable form of many misfolded proteins. These assemblies are generally protease- and detergent-resistant, and thus they persist in tissues for long periods of time and are difficult to study experimentally. Understanding the intrinsic factors that contribute to the stability of A β assemblies may provide insight into the features of the peptide that can be targeted in order to destabilize these assemblies.

7.2 Steered Molecular Dynamics

Steered molecular dynamics (SMD) is a technique in which a biasing potential is applied to one or more species in a simulated system. The application of this force produces non-equilibrium trajectories, allowing for forces within the system to evolve in response to the biasing potential. The biasing potential can be applied to any atom(s) in the system, but is frequently applied to

the center of mass (COM) of a chosen molecule. The resulting forces are distributed over the remaining atoms in the molecule. In GROMACS, there are several methods for applying this force - harmonic (umbrella), constraint (rigid), and constant force.

When applying a harmonic potential to pull on a molecule, the force is applied via a “spring,” the position of which is defined by the COM of the pulled species, and a “dummy” particle that begins at this COM position. As the dummy particle is pulled, forces are calculated from Hooke’s Law:

$$F = k(x - x_0) \quad (7.1)$$

In Equation 7.1, k is the force constant of the spring, a parameter that controls how stiff the spring (harmonic potential) is. The values of x and x_0 are the length of the spring and its equilibrium length, respectively. Thus, as the spring stretches, the force increases, resulting in changes in the structure to which the force is applied. The native interactions within the structure (hydrogen bonding, hydrophobic packing, ionic interactions, *etc.*) resist this applied force, and net changes only occur once this restoring force is overcome.

Constraint pulling resembles the behavior of bonds to which holonomic constraints [270, 300] are applied. Instead of applying a harmonic potential, the spring is replaced by a rigid rod that pulls on the species of interest. Both constraint and harmonic pulling techniques are known as “constant-velocity” pulling, since the dummy particle moves at a constant velocity away from the pulled species, allowing the forces to change in response to alterations in the structure. Application of a constant force results in non-constant velocity of the dummy particle.

7.3 Umbrella Sampling

Since SMD is a path-dependent, non-equilibrium technique, it is desirable to devise methods for extracting equilibrium thermodynamic data such as ΔG , a state function, from these trajectories. The first of these methods was derived by Jarzynski [301], wherein multiple SMD simulations (at different pulling velocities) can be evaluated such that a resulting ΔG for the association or dissociation of the complex can be determined. The limitation of Jarzynski’s Equality is that its accuracy is dependent upon many factors, including the pulling rate, force constant (k) of the spring, and the number of SMD trajectories produced. Reasonable accuracy may require tens or hundreds of SMD simulations at a very slow pulling rate, making data collection slow.

The potential of mean force (PMF) is a quantity from which ΔG can be derived. To calculate PMF, a pulled species must be restrained at different distances from a reference position along a reaction coordinate, ξ . The force at each discrete distance away from the reference is averaged to generate a mean force, from which the derivative is taken to derive the potential energy associated with restraining the molecules at this given distance. A common technique for calculating PMF is umbrella sampling. Independent simulations are conducted along the reaction coordinate, ξ ,

which is typically a vector between a reference position (or molecule) and the pulled species. The distance between the reference and the pulled species is maintained by a harmonic potential. Forces are calculated using Hooke's Law. The harmonic (umbrella) potential allows for the pulled species to oscillate about the desired restraint distance in what is called a "sampling window." Each restraint distance along the reaction coordinate defines a sampling window, which is generically described as a region of configurational space that can be explored by the pulled species. This pulled species is then simulated for sufficient time to generate an ensemble of structures within each window. The resulting values of the PMF are unbiased using a number of algorithms, the most popular of which is known as the Weighted Histogram Analysis Method (WHAM) [302].

WHAM produces a continuous energy function, provided that the ensembles of structures in neighboring windows overlap. That is, they sample configurations of similar energy. The overlap in energy results in a continuous PMF curve. The difference between the maximum and minimum of this curve gives the value of ΔG for the process under consideration.

7.4 Methodology

7.4.1 SMD Simulations

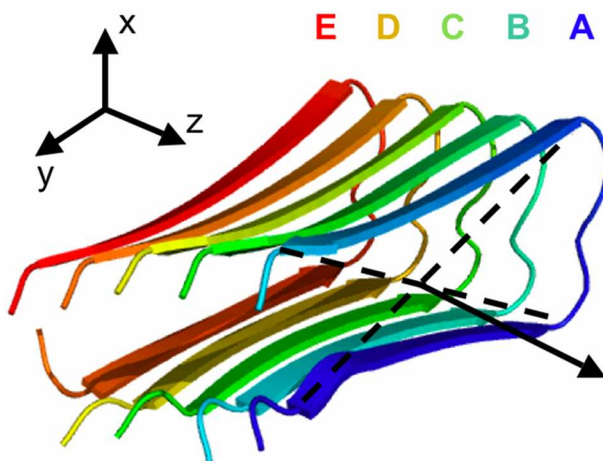


Figure 7.1: The A β protofibril, labeled according to the nomenclature used here, with the pulling direction (coincident with the z -axis) indicated

To analyze the stabilizing forces within the A β protofibril, SMD was applied to the wild-type structure such that a force was applied to peptide A directly along the z -axis, coincident with the protofibril axis (Figure 7.1). Different trajectories were generated by applying the pulling force directly to the C-terminal Ala42, Lys28 in the bend region, and the COM of peptide A. The results of these simulations were equivalent (that is, the evolution of structures and forces were the same),

and thus the discussion here will focus on the trajectory generated by applying the biasing force to the COM of peptide A. The application of the biasing force indicated that several residues may be involved in critical stabilizing interactions, which are described in the following sections. The residues identified for *in silico* mutation were Phe19, Asp23, Lys28, Ile32, and Leu34.

7.4.2 *In silico* Mutations

All mutants were generated by simply deleting the relevant sidechain atoms from the target residues. The mutations created were F19G, D23A, K23A, and double mutant I32G/L34G. Positions of these residues in the A β protofibril model are shown in Figure 7.2. Mutations to alanine were based on the standard biochemical alanine scanning mutagenesis technique. Mutations to glycine were designed to eliminate any hydrophobic packing effects that would still be present with an alanine mutant (exerted through the β -carbon).

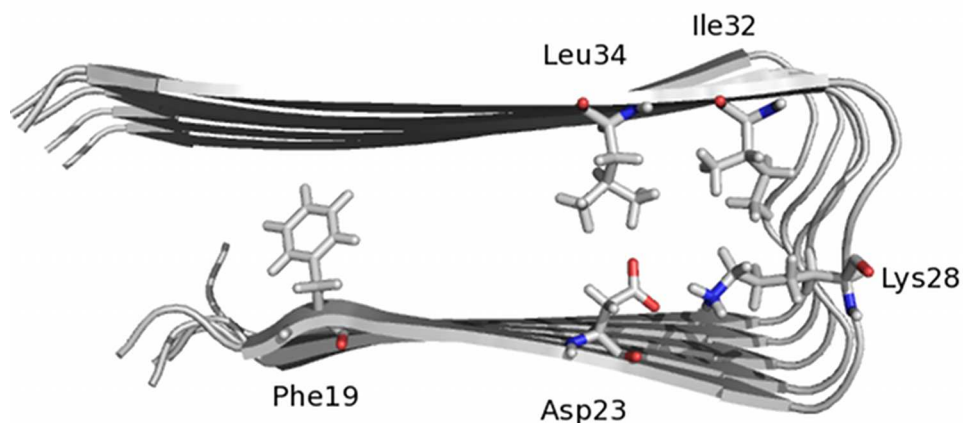


Figure 7.2: Positions of key amino acid residues in the A β protofibril model. The N-terminal 16 residues are not present in this structure.

7.4.2.1 System Construction and General Methodology

The model for the A β fibril chosen for this work was the core of the A β_{42} fibril, a protofilament structure determined using solid-state NMR [278]. In this model (PDB code 2BEG), the N-terminal 16 amino acids are missing. Masman *et al.* recently proposed that this core region of the fibril, comprising residues 17-42, is principally responsible for its stability [303]. As such, although only a portion of the A β protofilament, and thus the larger fibril as a whole, was modeled here, the observations made and conclusions reached likely extend to the mature, full-length A β fibril. Due to the small size of the assembly considered here (five peptides), the structure will be referred to as a protofibril, as it is the structure that comprises the repeat unit of a protofilament, and

ultimately the mature fibril. A representative structure, containing intact salt bridges, was chosen from the ensemble of structures in the PDB file as the starting structure for the simulations in this work. The mutation of Phe19 to glycine was motivated by the structural studies of Lührens *et al.*, who found that, despite the mutation, the A β F19G mutant formed fibrils, albeit of distinct morphology from the wild type [278]. The mutations of Asp23 and Lys28 to alanine were an attempt to assess the contribution of the inter-peptide salt bridges present in the hydrophobic core to the stability of the fibril by removing this favorable electrostatic interaction. Several studies have suggested the formation of this salt bridge contributes strongly to the stability of amyloid aggregates as higher-order structures develop [204, 304, 305]. By mutating Ile32 and Leu34 to glycine, we sought to probe hydrophobic packing interactions that may contribute to stabilizing the salt bridge and the hydrophobic interior of the fibril. It is not implied that any of these mutants should necessarily form fibrils, but these mutations are relevant here in the assessment of side chain contributions to stability.

The N-terminus of each peptide was capped with an acetyl group to give an uncharged terminus to mimic the full-length peptide, wherein additional amino acids would be present. This capping was done using the xLeap module of AmberTools [306]. The C-terminus of each peptide was deprotonated (-1 charge). All other titratable amino acids were assigned their canonical state at physiological pH. Parameters from the GROMOS96 53A6 parameter set [174] were applied to all species in the simulated system. Short-range nonbonded interactions were cut off at 1.4 nm, with long-range electrostatics calculated using the Particle Mesh Ewald (PME) algorithm [193, 194]. Dispersion correction was applied to energy and pressure terms to account for truncation of van der Waals terms. Periodic boundary conditions were applied in all directions.

7.4.2.2 Generation of SMD Starting Structures

In order to generate equilibrated starting structures for the pulling simulations, each protofibril (wild type and mutants) was placed in a cubic box of simple point charge (SPC) water [268], to which 100 mM NaCl was added, including neutralizing counterions. Following steepest descents minimization, each of the protofibril systems was equilibrated in two steps, with position restraints applied to peptide heavy atoms throughout. The first phase involved simulating for 50 ps under an isochoric-isothermal (NVT) ensemble. Protein and non-protein atoms were coupled to separate temperature coupling baths, and temperature was maintained at 310 K using the Berendsen weak coupling method [187]. Following NVT equilibration, 50 ps of isobaric-isothermal (NPT) equilibration were performed, also using weak coupling [187] to maintain pressure isotropically at 1.0 bar. Production MD simulations were conducted for 100 ns in the absence of any restraints. For this data collection period, the Nosé-Hoover thermostat [188, 189, 190] was used to maintain temperature, and the Parrinello-Rahman barostat [191, 192] was used to isotropically regulate pressure. This combination of thermostat and barostat ensures that a true NPT ensemble is sampled [188, 189, 190, 191, 192]. Simulations were conducted using the GROMACS package, version 4.0.2 [180].

Structures from the end of each of these trajectories (WT, F19G, I32G/L34G, K28A, and D23A) were used as starting configurations for pulling simulations. The protofibril structures were placed in a rectangular box with dimensions sufficient to satisfy the minimum image convention and provide space for pulling simulations to take place along the z -axis. As before, SPC water was used to represent solvent, and 100 mM NaCl was present in the simulation cell. All pulling simulations were conducted with the GROMACS package, version 4.0.4, to take advantage of pull code improvements and new features implemented during the course of the project. Equilibration was performed for 100 ps under an NPT ensemble, using the same methodology described above.

7.4.2.3 SMD Data Collection Simulations

Following equilibration, restraints were removed from all peptides except peptide B (Figure 7.1); it was used as an immobile reference for the pulling simulations. Position restraints have frequently been used in simulations of model A β fibrils to mimic the stability of a much larger structure [307, 308, 309]. For each of the protofibril structures, peptide A was pulled away from the core structure along the z -axis (Figure 7.1) over 500 ps, using a spring constant of 1000 kJ mol⁻¹ nm⁻² and a pull rate of 0.01 nm ps⁻¹ (0.1 Å ps⁻¹). Application of slower pulling rates (0.005 nm ps⁻¹ and 0.001 nm ps⁻¹) resulted in the production of nearly identical trajectories and similar force vs. time curves (Figure 7.3), and thus the faster pulling rate was applied to all systems here in order to expedite data collection while still preserving reliability of the results.

7.4.3 Umbrella Sampling Simulations

Following the SMD simulations, a final center-of-mass (COM) distance between peptides A and B of approximately 5.5 nm was achieved. From these trajectories, snapshots were taken to generate the starting configurations for the umbrella sampling [310, 311, 312] windows. An asymmetric distribution of sampling windows was used, such that the window spacing was 0.1 nm up to 2 nm COM separation, and 0.2 nm beyond 2 nm of COM separation. Such spacing allowed for increasing detail at smaller COM distance, and resulted in 31 windows. In each window, 10 ns of MD were performed for a total simulation time of 310 ns utilized for umbrella sampling. Analysis of results was performed with the Weighted Histogram Analysis Method (WHAM) [302]. The force constant for restraining the COM distance between peptides A and B was 1000 kJ mol⁻¹ nm⁻². Since the pulling direction between peptides A and B resulted in a constant increase in the COM distance between these two species, the “distance” pulling method in GROMACS was used. Had the COM distance changed signs (that is, the pulling species, peptide A, moved to the opposite side of the reference species, peptide B), then the “direction” or “position” methods would have been more appropriate.

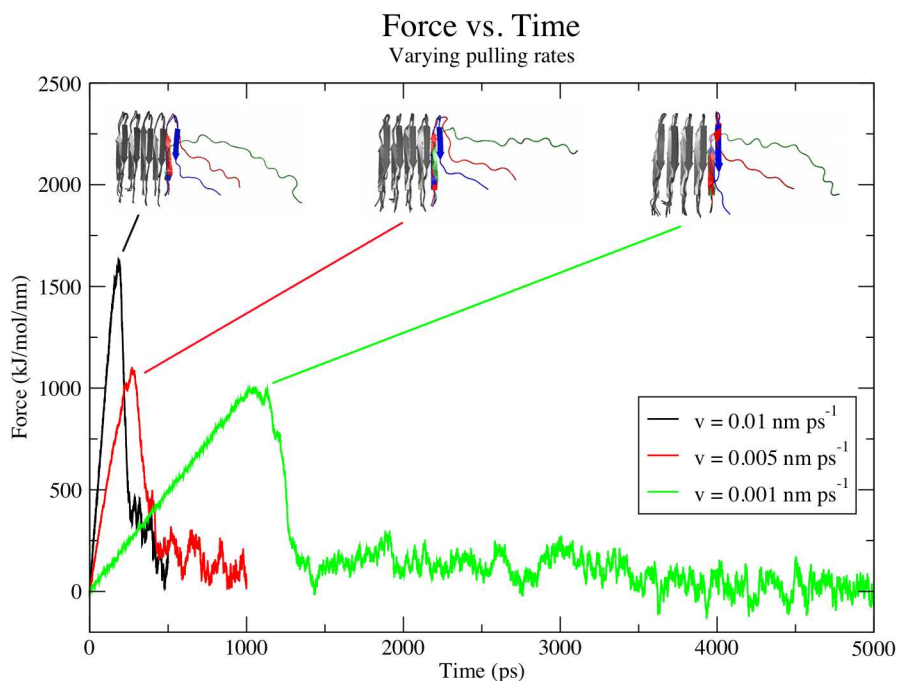


Figure 7.3: The dependence of SMD trajectories on the pull rate, v . The events observed in each trajectory are equivalent, as are the shapes of the force vs. time curves. Blue structures correspond to the points of maximum force in each trajectory. Red structures are time points after which the most significant stabilizing interactions had broken. Green structures are taken after the Asp23-Lys28 salt bridge between peptides A and B had broken.

7.5 Results

All results refer to amino acids by their position in the full-length A β peptide, for ease of reference to previous experiments, even though the N-terminal 16 residues were missing in the NMR structure used in these simulations. The β -strands in the protofibril are referred to as β_1 (extending from residues 18-26) and β_2 (residues 31-42), from the nomenclature devised by Lühns *et al.* [278]. A bend connects the two β -strands.

7.5.1 Structural Details of A β Protofibrils

Mutations in the A β sequence led to noticeable distortions in the structure of the protofibril over the course of 100-ns MD trajectories, although none of the peptides spontaneously dissociated from the core of the structure. Shown in Figure 7.4 are images of the starting structures used in the pulling simulations; these structures are the result of 100 ns of unrestrained MD following the equilibration protocol described above. Backbone root mean square deviation (RMSD) values are shown in Table 7.1. These values represent the average calculated over the period of the simulation

during which the RMSD was stable, generally the last 50 ns. Maintenance of the integrity of the hydrophobic core of the protofibril explains the observed structures. From Figure 7.4B and Table 7.1, it can be seen that the wild-type protofibril did not deviate substantially from the original NMR structure. Also in the wild-type structure, the Asp23-Lys28 salt bridges were maintained over time in six independent simulations (data not shown).

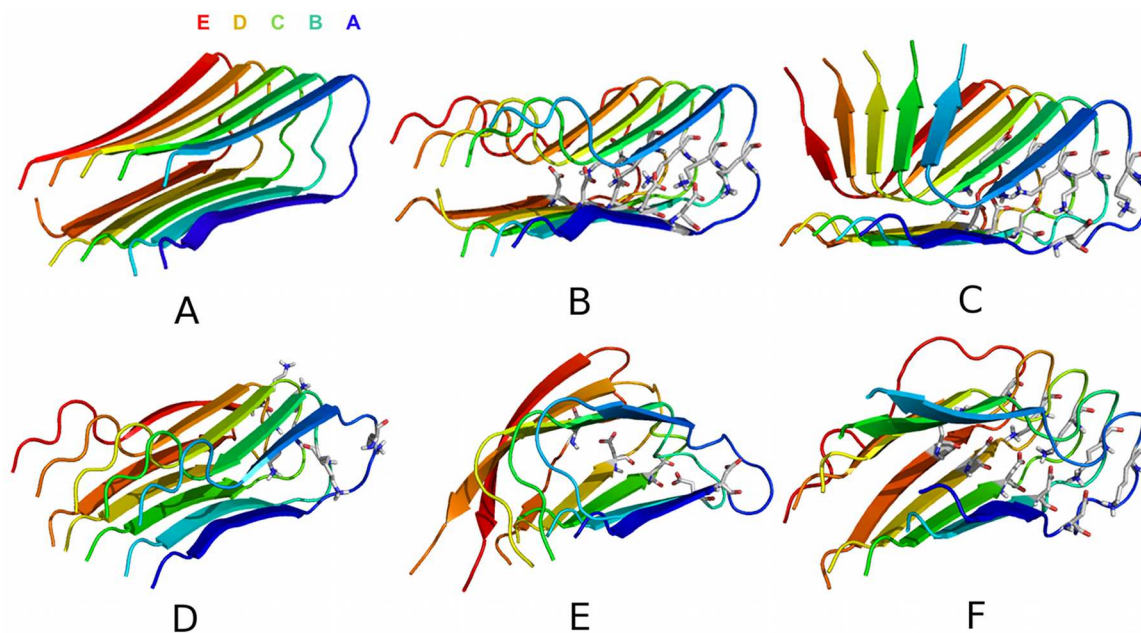


Figure 7.4: Protofibril structures (A) from NMR, with peptides labeled according to the nomenclature used here, and after 100 ns of simulation at 310 K, which served as the initial structures for pulling simulations: (B) WT, (C) F19G, (D) D23A, (E) K28A, and (F) I32G/L34G. All images include explicit stick representation of Asp23 and Lys28, unless mutated.

Table 7.1: Backbone RMSD values after 100 ns of MD simulation.

System	Backbone RMSD (Å), $\pm \sigma$
WT	2.65 ± 0.09
F19G	5.34 ± 0.17
D23A	2.82 ± 0.07
I32G/L34G	3.93 ± 0.25
K28A	5.93 ± 0.19

Mutation of Phe19 to glycine resulted in a pinched structure, with strand $\beta 2$ making contact with strand $\beta 1$ where the side chain of Phe19 normally packs against Gly38 (Figure 7.4C). Mutating Asp23 to alanine prompted the Lys28 residues of each peptide to orient their side chains into the aqueous solvent, thus removing the positively-charged residues from the hydrophobic core

(Figure 7.4D). This behavior led to minimal deformation of the protofibril structure. The K28A mutation caused the greatest changes to the protofibril structure. Since aspartate lacks a long, flexible side chain like that of lysine, this charged residue is less able to move outwards to the aqueous environment to solvate the negatively-charged carboxylate (Figure 7.4E). As such, the K28A protofibril, while still maintaining the overall integrity of its hydrophobic core, deforms substantially to maximize the exposure of Asp23 to solvent, which protrudes through the bend region into the solvent. In the case of the I32G/L34G mutant, the most significant observation is the increased solvation of the Asp23-Lys28 salt bridge, which leads to destabilization of the bend region (Figure 7.4F).

In the WT protofibril, an average of 6.8 water molecules was present in the solvent-accessible channel near the Asp23-Lys28 salt bridges (Figure 7.5), with a maximum solvation at times reaching 15 water molecules, as determined by the GROMACS accessory program “g_count” (created and distributed by Oliver Beckstein [313]). In the case of the I32G/L34G mutant, an average of 7.8 water molecules was present over time, peaking at a maximum of 30 water molecules. For the F19G protofibril, an average of 5.1 water molecules was present per timeframe, with solvation peaking at 13 water molecules. In the case of D23A, this solvent-accessible channel was absent due to the hydrophobic packing that evolved over the course of the 100-ns simulation. In the K28A protofibril, no solvent-accessible channel can be rigorously defined based on the structural changes that took place.

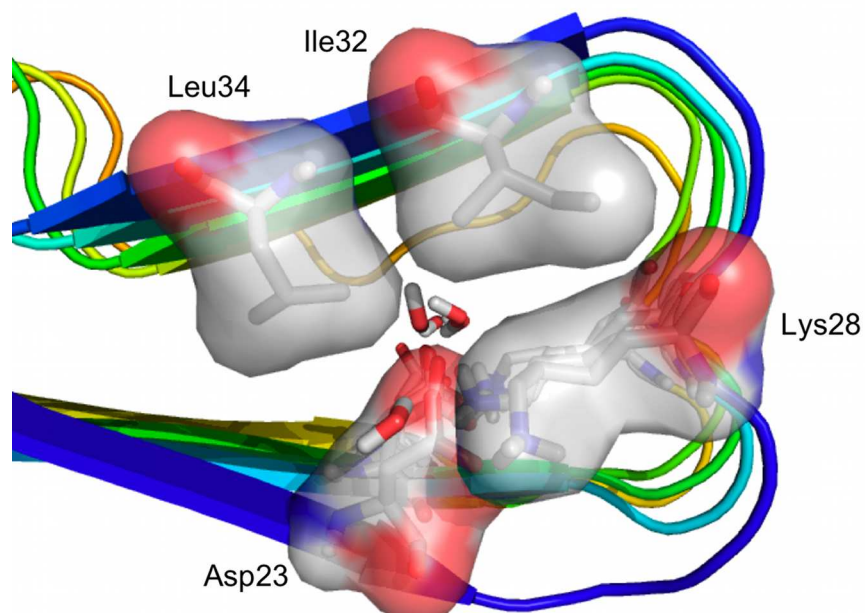


Figure 7.5: Native packing of Ile32 and Leu34 against the Asp23-Lys28 salt bridge, taken from the WT fibril configuration at the end of 100 ns of simulation. A narrow channel of water can be seen in the vicinity of the salt bridge, with nearby residues labeled.

Twisting of β -strands in A β fibrils relative to one another to increase side chain contacts (par-

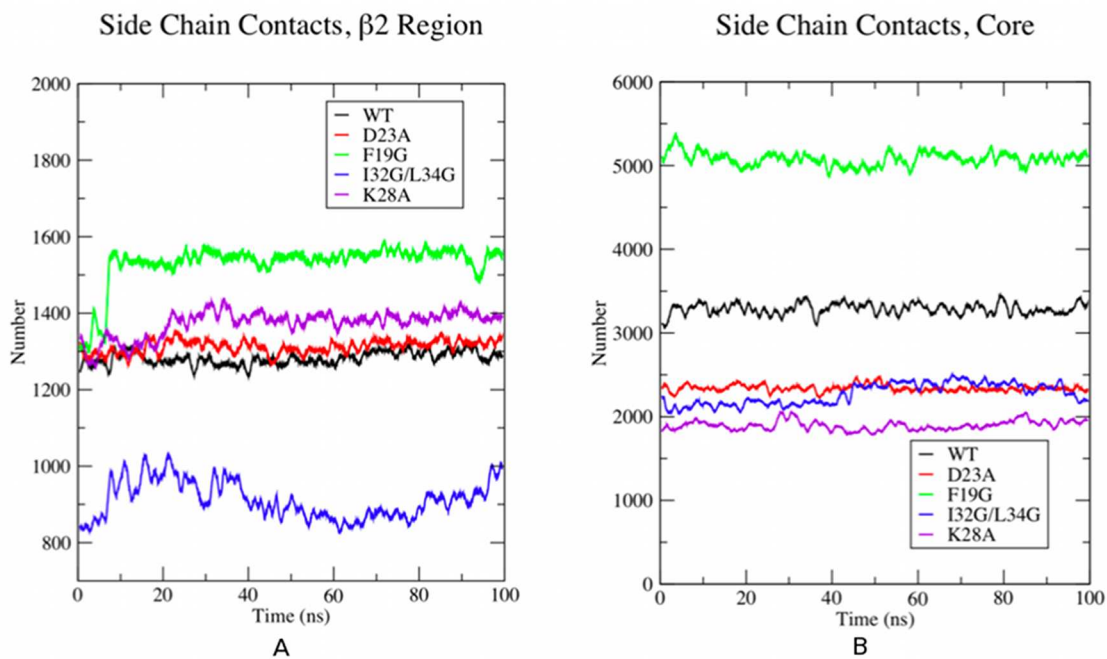


Figure 7.6: Formation of (A) β 2 and (B) protofibril core (residues 21-34) sidechain-sidechain contacts over time as the protofibril β -strands twist. For clarity, data are plotted as running averages within a 100-point sliding window.

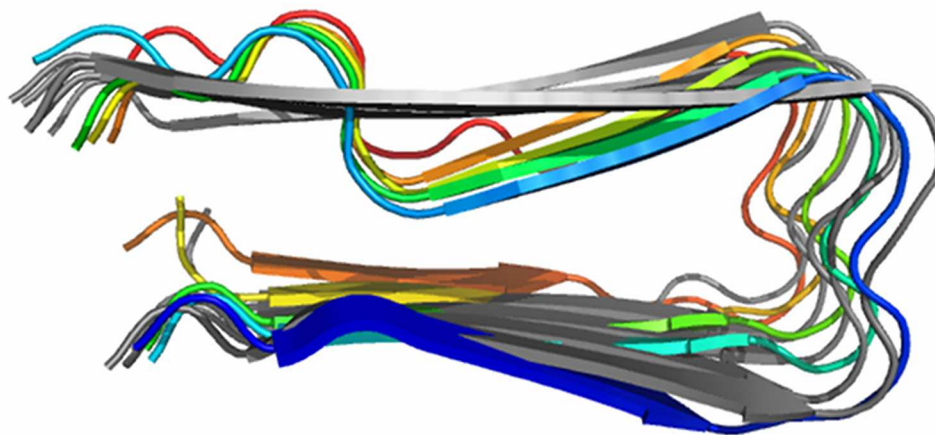


Figure 7.7: Twisting of β -strands in the A β ₄₂ protofibril structure. The NMR structure is shown in gray, with the final structure after 100 ns of MD simulation overlaid and colored by peptide, as in Figure 7.4.

ticularly in the β 2 region) has been proposed as a driving force for amyloid fibril stability and has been observed in previous MD simulations [202, 303, 314]. We find that in all cases, some twisting is apparent (Figures 7.4 and 7.7) and side chain contacts in the β 2 region quickly increase

(within 10 ns of simulation) and remain largely stable over the remainder of each of the trajectories (Figure 7.6). We note that a previous study by Masman *et al.* [303] observed a more pronounced twisting effect in the case of the WT protofibril than is observed here (Figure 7.7), a difference that may be attributed to the presence of additional residues in the previous study (full-length peptides instead of the truncated form utilized here) and the choice of force field used (OPLS-AA [168] in the previous work, and GROMOS96 53A6 [174] in this study). Contacts formed between side chain atoms in the β 2 region are comparable among the WT, F19G, D23A, and K28A protofibrils, with F19G forming somewhat more contacts than the WT due to the pinching at the C-terminus described above (Figure 7.6). The β 2 region of the I32G/L34G mutant forms substantially fewer total side chain contacts due to the instability of peptide E over the course of 100 ns of MD simulation. The most dramatic difference in terms of contacts are seen when considering the core residues that pack around the Asp23-Lys28 salt bridges, specifically residues 21-34, packing interactions whose importance will be described in the subsequent sections. While the F19G protofibril forms more contacts than any other protofibril in this core region, the other mutated structures (D23A, I32G/L34G, and K28A) form substantially fewer contacts than either the WT or F19G protofibrils (Figure 7.6), which likely contributes to the instability seen in these structures.

7.5.2 SMD Simulations

COM pulling, or steered molecular dynamics (SMD), simulations can be used to bias the behavior of a system towards a particular phenomenon that might otherwise be inaccessible on the time scale of conventional MD. SMD simulations have been successfully applied to the study of many processes, including protein-ligand binding, protein-protein interactions, and extraction of lipids from membranes [315]. Application of an external force to cause displacement in the simulated system allows for the calculation of work, a path-dependent quantity. The use of this pulling force perturbs the equilibrium of the system, thus precluding the calculation of thermodynamic quantities directly from the SMD trajectory without large errors. Methods proposed by Jarzynski [301] and the Weighted Histogram Analysis Method (WHAM) of Kumar *et al.* [302] are commonly used to extract equilibrium data (free energies) from the non-equilibrium SMD trajectories for a more direct, quantitative means of comparing two systems. While Jarzynski's method requires running multiple SMD simulations, the WHAM procedure calculates ΔG from a number of simulations conducted on configurations generated from a single SMD simulation. We use the latter approach here.

By pulling on the COM of peptide A of the A β ₄₂ protofibril, force builds up until a breaking point is reached, at which time critical interactions are disrupted, allowing peptide A to dissociate from the core protofibril structure. In each of the protofibrils in this study, the point of maximum force corresponded to the instant just before the interpeptide backbone hydrogen bonds between Gly33 and Met35 of peptide A and the corresponding residues of peptide B were broken (Figure 7.8), although the events leading up to the dissociation of each peptide differed, sometimes dramatically. It is for this reason that the force-time curves, generated by conducting COM pulling, by themselves are insufficient to conclusively determine the stability of each of the mutant protofibrils

in this study (in terms of ΔG), since the dissociation is a path-dependent process. Thus the maximum force in each pulling simulation is not necessarily comparable to that of other simulations unless the dissociation pathway is very similar. Umbrella sampling [310, 311, 312] is employed to calculate the ΔG of binding of peptide A to each of these protofibrils considered here.

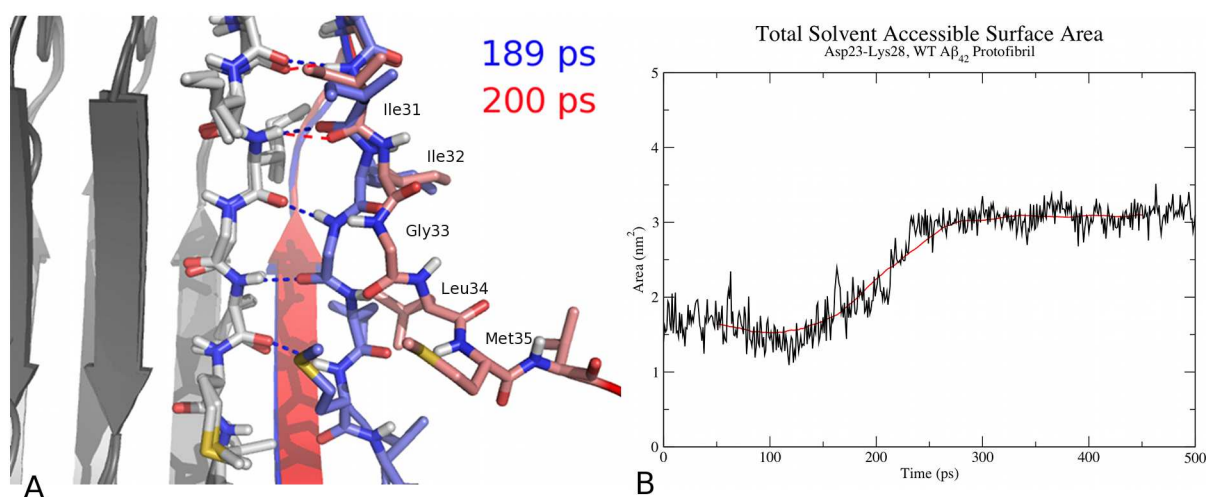


Figure 7.8: (A) Detail of the dissociation of peptide A from peptide B in the WT A β_{42} fibril. Carbon atoms in peptide A are tinted blue (189 ps, the point of maximum force in Figure 7.9(A)) or red (200 ps, as the force is decreasing), while carbon atoms in peptide B are gray. Other elements are colored according to their typical representation (N: dark blue, O: dark red, H: white, S: yellow). Hydrogen bonds are indicated as dashed lines, colored according to the snapshot in which they occurred, demonstrating the interactions that are sequentially broken over the dissociation pathway. (B) Total solvent accessible surface area of the Asp23-Lys28 salt bridge of peptide A of the WT fibril as the peptide is pulled away from the fibril structure. The running average of the data (100-point windows) is shown in red to illustrate the overall trend in the data.

The mature A β_{42} fibril is stabilized by both intrapeptide (hydrophobic packing of $\beta 1$ against $\beta 2$) and interpeptide (backbone hydrogen bonding, salt bridges, and side chain packing) interactions [278]. The use of COM pulling makes it possible to monitor the evolution of these types of interactions over time as peptide A is pulled away from the protofibril structure. Such a method serves a dual purpose here. It allows for qualitative assessment of the specific interactions that contribute to the stability of the protofibrils and it generates a series of configurations that are used later for umbrella sampling [310, 311, 312].

While the Asp23-Lys28 salt bridges in the A β_{42} fibril are primarily interpeptide interactions [278], their stability may be influenced by a number of intrapeptide interactions, including the overall rigidity of the peptide structure induced by hydrophobic packing around the salt bridge and the interaction between the side chains of residues in $\beta 1$ and $\beta 2$. The terminal peptide (peptide A) may be especially influenced by this packing phenomenon, as an intrapeptide salt bridge is also capable of forming. That is, Asp23 of peptide A can be involved in two salt bridges simultaneously in the

WT simulations described above - it can interact with both Lys28 of peptide B and Lys28 of peptide A. Both likely contribute to maintaining rigidity of the bend region, which, in turn, stabilizes the interaction of β 1 and β 2. These interactions are systematically explored in the following sections.

The discussion begins with the COM pulling simulations and analysis of the WT protofibril, which provides the rationale for *in silico* mutations that were made. The dissociation pathway of the mutant protofibrils is then described in two sections - descriptions of mutations with intact Asp23-Lys28 salt bridges, and those without.

7.5.2.1 Wild-type A β ₄₂ Protofibril

In the WT protofibril, after reaching the maximum force at 189 ps (blue structure in Figure 7.8, blue line in Figure 7.9A), several intrapeptide packing interactions, Leu34 against Ala21 and Ile32 against the aliphatic portion of the Lys28 side chain, were both disrupted within only 11 ps (red structure in Figure 7.8, red line in Figure 7.9A). These observations in the case of the WT protofibril highlight the importance of maintaining the packing of strand β 2 (containing Ile32 and Leu34) against strand β 1 (containing Asp23), especially in the vicinity of the Asp23-Lys28 salt bridge. The position of residues Ile32 and Leu34 define a narrow channel of water near the Asp23-Lys28 salt bridge (Figure 7.5), allowing a finite level of hydration to persist near these charged residues. Mutation of Lys28 to alanine substantially perturbs this region (Figure 7.10, and discussed below). The presence of internal hydration in the A β ₄₂ fibril has been proposed previously [202, 316]. Disruption of native packing in this region allows additional water molecules to interact with the Asp23-Lys28 salt bridge, causing it to become increasingly exposed to bulk solvent (Figure 7.8B). Within approximately 60 ps after reaching the maximum force, the salt bridge was entirely exposed to bulk solvent (green line, Figure 7.9A). It appears that a delicate balance must be maintained in order to preserve a narrow channel of water that acts to stabilize the Asp23-Lys28 interaction, with water molecules serving as hydrogen bond bridges between these residues. If these interactions are disrupted, peptide A quickly moves away from peptide B. Thus we generated a series of mutants (D23A, K28A, and I32G/L34G) to assess the contributions of the salt bridge and surrounding hydrophobic residues to the stability of the fibril as a whole. In addition, a mutant whose structure has been characterized experimentally (F19G) was analyzed.

The evolution of the protofibril in the SMD simulations was also examined using slower pulling rates (0.005 nm ps⁻¹ and 0.001 nm ps⁻¹) to verify that the faster pulling rate utilized (0.01 nm ps⁻¹) did not induce any artifacts. From Figure 7.3, it is clear that peptide A dissociated in the same manner under each of the applied pulling rates, giving effectively the same sequence of configurations, regardless of the pulling rate. The time required to achieve the maximum force, as well as the value of this maximum force, is dependent upon the pulling rate, as the structural changes are occurring at different rates. However, it is clear that a single main structural transition is produced in each case, corresponding to the pathway detailed above.

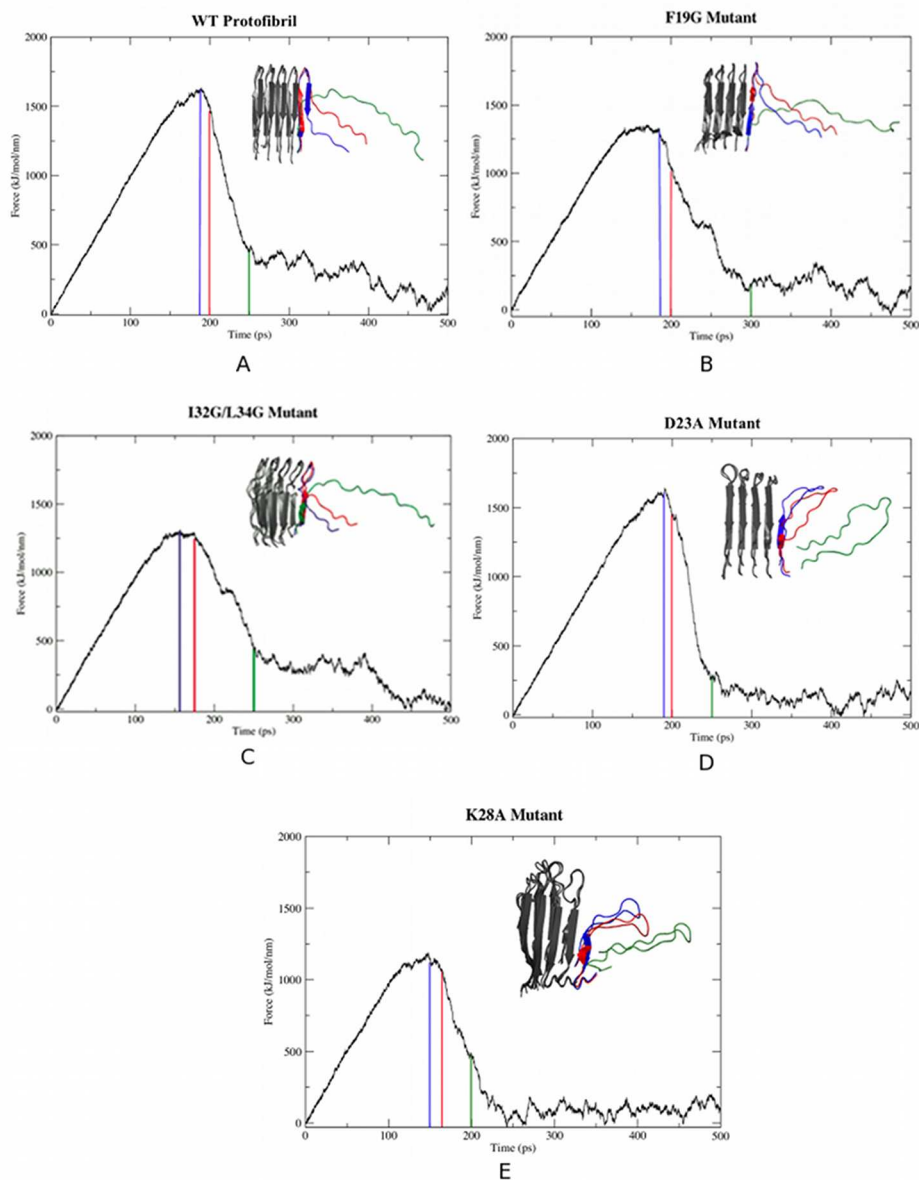


Figure 7.9: Dissociation pathways of $A\beta_{42}$ protofibrils, with corresponding plots of force vs. time: (A) WT, (B) F19G, (C) I32G/L34G, (D) D23A, and (E) K28A. In each panel, the blue line (and structure) corresponds to the point of maximum force, and the red and green lines and structures correspond to major structural transitions (WT, F19G, and I32G/L34G) or the dissociation of peptide A (D23A and K28A).

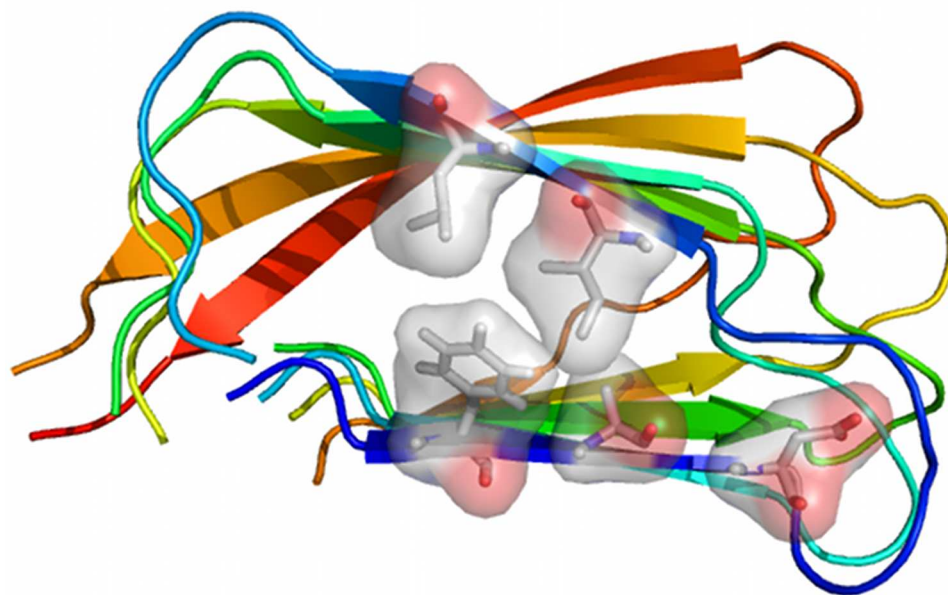


Figure 7.10: Non-native packing in the K28A mutant protofibril, taken from the final configuration of the system (100 ns). Shown in detail are Phe19, Ala21, Asp23, Ile32, and Leu34 of peptide A. In the WT protofibril, Leu34 normally packs against the side chain of Asp23, and Ile32 with Lys28.

7.5.2.2 Mutant Protofibrils with Intact Salt Bridges (F19G and I32G/L34G)

In the case of the mutant protofibrils that possessed intact salt bridges after 100 ns of MD simulation (F19G and I32G/L34G), applying force to peptide A resulted in dissociation pathways that largely followed that of the WT protofibril. Under the applied force, the C-terminal residues in the β 2 region dissociated first, followed by the bend region and β 1 for both F19G and I32G/L34G (Figure 7.9B and 7.9C, respectively). For this reason, some general observations of the forces required to destabilize these protofibrils relative to the WT can be made, although a rigorous quantitative comparison of forces is not possible since the pathways are not identical.

The dissociation of peptide A in the F19G mutant required less force than in the case of the wild type. The F19G mutation does not alter the native packing in and around the salt bridge, although hydration is slightly decreased (5.1 water molecules per timeframe) relative to the WT (6.8 water molecules per timeframe) over the course of the 100-ns simulation conducted prior to pulling. The reduction in force thus derives from the non-native packing interactions that arise between strands β 1 and β 2 in F19G. The deletion of the Phe19 side chain (which normally packs against Gly38 in the WT structure) caused a bend in the backbone of β 2 to generate a Gly19-Gly38 contact. This phenomenon strains the hydrogen bonds involving Val40 and Gly38, evidenced by the fact that less force is required to break them than in the case of the WT. The dissociation pathway of the F19G mutant resembled that of the WT protofibril, due largely to native packing of Ile32 and

Leu34 against the Asp23-Lys28 salt bridge, with the narrow water channel described previously.

When Ile32 and Leu34 are simultaneously mutated to glycine, the bulky hydrophobic side chains that normally pack against β 1 and the Asp23-Lys28 salt bridge are absent, such that when the hydrogen bonds involving Met35 break, there are fewer stabilizing interactions in this region to preserve the interaction of peptides A and B. The Asp23-Lys28 salt bridge is also more solvent-accessible as a consequence. The greater extent of hydration around the salt bridge contributes to this destabilization, and peptide A is able to dissociate from peptide B with substantially less force than in the case of the WT protofibril.

7.5.2.3 Mutant Protofibrils without Intact Salt Bridges (D23A and K28A)

To evaluate the contribution of the Asp23-Lys28 salt bridge to the binding of peptide A to the protofibril, both Asp23 and Lys28 were independently mutated to alanine. In the case of these mutants, peptide A dissociated from the protofibril along a dissociation pathway that was vastly different from that of the WT, F19G, and I32G/L34G structures, and thus the D23A and K28A mutants must be considered separately with respect to the nature of the forces required to cause this dissociation.

In the case of the D23A mutant, the predominantly polar bend region (G₂₅SNKGA₃₀) dissociates from the protofibril first (Figure 7.9D), likely due to the favorable exposure of Ser26, Asn27, and Lys28 to the aqueous solvent. Tight hydrophobic interactions between β -strands β 1 and β 2 (specifically Ile32-Ala23 and Leu34-Ala21) caused peptide A to remain compact as it dissociated from the rest of the protofibril structure. As interpeptide interactions in the β 1 and β 2 regions were simultaneously lost, the force, which was acting on both β 1 and β 2, decreased over a very short period of time (Figure 7.9D). Although the maximum force reached in both the WT and D23A pulling simulations is essentially identical (Figures 7.9A and 7.9D, respectively), peptide A followed a very different dissociation pathway in each case. Specifically, the force applied to the COM of the WT peptide A produces a sequential dissociation of the β -strands, first β 2, then β 1, while the same force applied to the D23A mutant acts to destabilize β 1 and β 2 simultaneously, following the dissociation of the bend region. Thus conclusions cannot be made regarding the stability of these structures purely from an examination of this path-dependent quantity, but will be explored later using umbrella sampling.

In the case of the K28A mutant, Asp23 protrudes through the backbone of the bend region, contributing to the destabilization of the bend (Figure 7.4E), and leading to non-native packing (Figure 7.10) prior to the application of any pulling force. The peak of the force curve (blue line, Figure 7.9E) corresponds to the disruption of inter-chain backbone hydrogen bonds involving Gly33 and Met35, as well as interpeptide Leu34 interactions, the side chains of which are also tightly packed against Phe19 and Ala21 in the core of the protofibril. Ile32 is also involved in this interaction in the K28A mutant, as the mutation destabilized the bend region and caused a shift in the alignment of the side chain contacts (Figure 7.10). Once this hydrophobic region is destabilized, the rest of peptide A quickly dissociated from the core of the protofibril. As was the case with the

D23A mutant, peptide A dissociated in a compact conformation resembling a β -hairpin.

The structural rearrangements induced by the K28A and D23A mutations caused a shift in the contacts made between β 2 residues and the central hydrophobic cluster (CHC, formed by packing between β 1 and β 2 in the region of residues L₁₇-VFFA₂₁). When Lys28 is mutated, Ile32 and Leu34 shift their positions to interact with this sequence. Only a minor rearrangement is observed in the case of D23A, wherein Ile32 packs against the mutant Ala23 residue instead of the aliphatic portion of Lys28, which was exposed to solvent in the mutant structure. It appears from these observations that hydrophobic forces alone are insufficient to maintain the stability of the protofibril.

7.5.2.4 Summary of Critical Interactions

The various mutations studied here illustrate the combination of hydrophobic and electrostatic forces that act to stabilize the A β ₄₂ protofibril. Mutations alter the native packing of the protofibril, highlighting the importance of regular side chain packing both between peptides, and within peptide A, the β 1- β 2 interactions described above. From these COM pulling simulations, several important observations become clear: (i) the shape of the force curve and the magnitude of the maximum force (Figure 7.9) are dependent upon the dissociation pathway, but are also loosely correlated with the extent of structural perturbation induced by the mutation, as measured by backbone RMSD (Table 7.1), (ii) protection of the Asp23-Lys28 salt bridge from excess hydration (through packing interactions with Ile32 and Leu34, an interaction stabilized by Gly33 and Met35 interpeptide backbone hydrogen bonding, as well as Leu34-Ala21 intrapeptide packing) is important to the stability of the protofibril, and (iii) the absence of the Asp23-Lys28 salt bridge leads to the simultaneous dissociation of β 1 and β 2 of peptide A from the A β protofibril, in contrast to the WT pathway, wherein β 2 and β 1 dissociated sequentially under the influence of the applied force. It becomes clear that preserving the integrity of the Asp23-Lys28 salt bridge, and thus the rigidity of the bend region, largely governs the stability of the interaction between peptides in the A β ₄₂ protofibril.

7.5.3 Umbrella Sampling Simulations

Umbrella sampling simulations are used to determine the ΔG of a particular event along a reaction coordinate, ξ . In this case, the reaction coordinate corresponds to the z-axis, coincident with the A β protofibril axis. By using 31 sampling windows along this axis, one-dimensional potential of mean force (PMF) curves were obtained for each protofibril (Figure 7.11), leading to the ΔG of binding for peptide A in each of the protofibrils in this study (Table 7.2). Error analysis was conducted using a bootstrap method described previously [317]. For the WT, F19G, D23A, and K28A PMF curves, the energy minimum occurred at a COM distance of 0.49-0.50 nm, in good agreement with WT peptide spacing determined by NMR [278]. For the I32G/L34G protofibril, the energy minimum is located at approximately 0.45 nm due to the fact that the double mutation allowed for closer interpeptide spacing because two bulky side chains were deleted.

Table 7.2: Results from Umbrella Sampling Simulations.

System	ΔG_{Bind} (kcal mol ⁻¹)*	$\Delta\Delta G_{Bind}$ (kcal mol ⁻¹)
WT	-50.5	-
F19G	-50.9	-0.4
D23A	-45.2	+5.3
I32G/L34G	-43.1	+7.4
K28A	-37.1	+13.4

*The error associated with energy minima is ± 0.2 kcal mol⁻¹ for each system.

Taking into account the magnitude of the standard error, mutation of Phe19 to glycine has no distinguishable effect on the free energy of binding between peptides A and B. In light of the fact that F19G mutant A β peptides form fibrils [278], this result is not surprising. Though structurally distinct from the WT fibril, the F19G mutant can form stable interactions between peptides, leading to fibril formation. Mutating Asp23 or Lys28 to alanine, however, substantially destabilizes the protofibril, with the K28A mutation having a more profound effect (Table 7.2). It is clear from these results that disrupting the salt bridge, especially by increasing its exposure to solvent in the core of the A β_{42} protofibril, would likely destabilize the structure as a whole. This finding is corroborated by the PMF data for the I32G/L34G mutant. In this protofibril, the Asp23-Lys28 salt bridge is intact but exposed to a greater amount of solvent than in the WT protofibril. The $\Delta\Delta G$ of binding for the I32G/L34G mutant is +7.4 kcal mol⁻¹, indicating that the increased solvation of the salt bridge, through the loss of hydrophobic interactions that rigidify this region and define the solvent-accessible channel, leads to destabilization.

7.6 Discussion

Numerous simulations have been conducted on A β fragments [204, 305, 318, 319, 320], monomeric full-length peptides [201, 203, 321, 322], protofilaments [202, 303], and fibrils [304, 307, 308, 309, 316], but as of yet a detailed thermodynamic analysis of the factors influencing the stability of the A β_{42} fibril under physiologically relevant conditions is lacking. We sought to elucidate the effects of various mutations on the stability of the A β_{42} protofibril as a means to understand the contributions of various structural features (hydrophobic packing and the Asp23-Lys28 salt bridge) on the stability of the mature fibril. In order to properly account for electrostatic interactions, we conducted all simulations using explicit solvent, since many implicit solvent models fail to accurately represent these interactions [204].

Previous simulations on the A β_{9-40} protofilament concluded that the stability of the mature A β fibril was due to inter- and intramolecular hydrophobic packing between the β -strands, as well as intermolecular Asp23-Lys28 salt bridges [202]. Contributing to the stability of the salt bridge is a

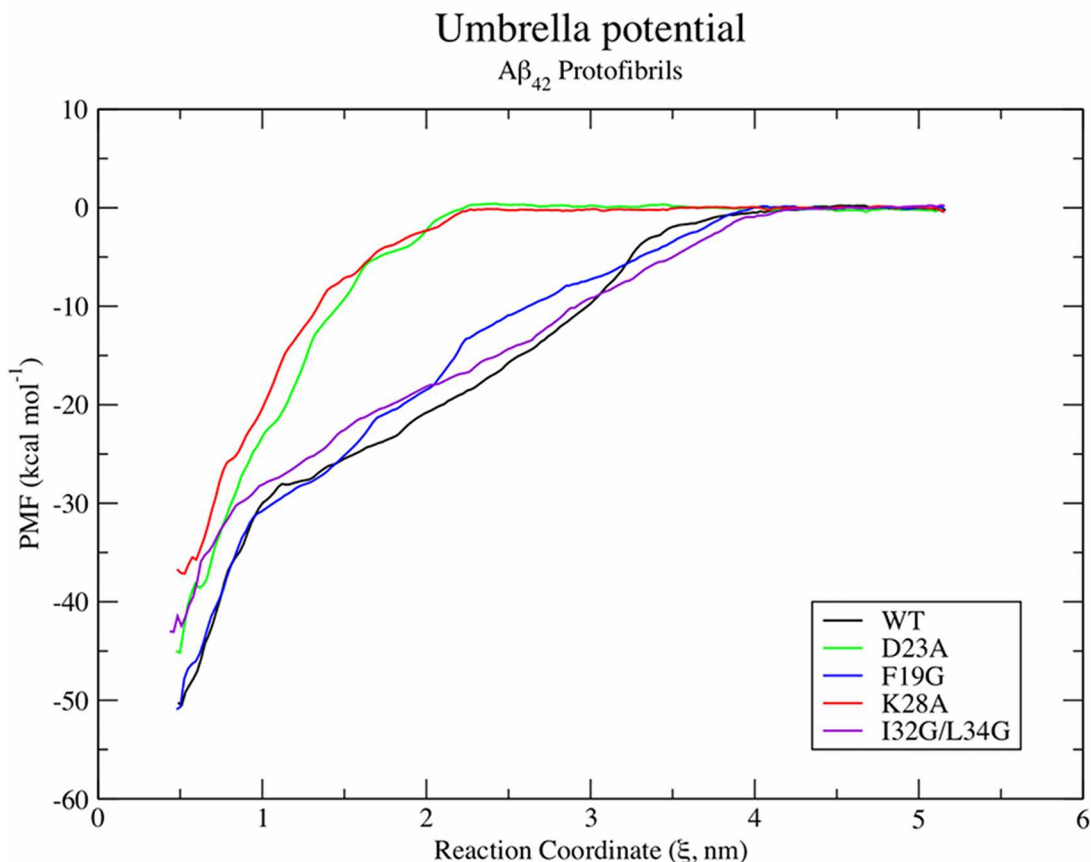


Figure 7.11: PMF curves for all the protofibrils considered here.

region of internal hydration within the core of the fibril, a finding that extends to simulations conducted by Tarus *et al.* [318], who found that, in the context of the A β_{21-30} fragment in solution, a single hydration shell favors the transient formation of the Asp23-Lys28 salt bridge over a competing Glu22-Lys28 salt bridge that may form in a more dehydrated environment. Increased solvation precludes the formation of stable salt bridges in the A β_{10-35} fragment [204]. Taken together with the results of replica-exchange MD simulations conducted by Melquiond *et al.* [305], who reported that the strength of this salt bridge increases in dimeric structures relative to monomers, MD simulations indicate that the contribution of the Asp23-Lys28 salt bridge (both intra- and intermolecular) to the stability of A β aggregates increases with the size of the aggregate. Thus, it is reasonable to postulate that the Asp23-Lys28 salt bridge contributes significantly to the binding energy of an incoming A β peptide onto a growing A β fibril.

In our SMD simulations of the WT structure, the event precipitating the release of peptide A from the larger protofilament structure was the increased exposure of the Asp23-Lys28 salt bridge to solvent by destabilizing the $\beta 2$ region extending from residues 32-35 (I₃₂GLM₃₅). It has been noted in several studies [202, 316] that a narrow channel of water persists inside the A β fibril in

the vicinity of the Asp23-Lys28 salt bridges, which we observe here, as well. Internal hydration near the Asp23-Lys28 salt bridge in the WT, F19G, and I32G/L34G structures persists over the course of the 100-ns trajectories conducted prior to pulling, with water molecules often bridging the amino and carboxylate functional groups of the respective side chains. When both Ile32 and Leu34 are mutated to glycine, additional water enters this channel, resulting in a weaker binding energy between peptides A and B. Thus, we propose that a delicate balance in the level of hydration of the salt bridge is key to its integrity. A small amount of water in the interior of the fibril stabilizes the Asp23-Lys28 interaction, but excess hydration of the salt bridge destabilizes it, allowing the Asp23 and Lys28 residues to interact with the bulk solvent. This event perturbs the bend region of peptide A, destabilizing the association between peptides A and B. This observation is also corroborated by the results of the SMD simulations conducted on the D23A and K28A mutants. In these two simulations, salt bridges could not be formed, and the bend region of peptide A dissociated from that of peptide B very quickly, pulling the β 1 and β 2 regions away simultaneously. By using umbrella sampling, we have determined that independent mutations of Lys28 and Asp23 to alanine decrease the binding energy of the terminal peptide by 26.5% and 10.5%, respectively, relative to the WT protofibril. Increased hydration in this internal channel, as demonstrated in our I32G/L34G mutant, decreases the magnitude of the binding energy by 14.6%. All of these reduced binding affinities are a product of decreased native contacts and the resulting instability of the bend region.

Recent simulations of the A β protofilament structures (both A β ₄₀ and A β ₄₂) concluded that the shape of the bend region influenced the packing behavior between the N-terminal and C-terminal β -strands [304]. Since the formation of this bend structure will dictate the nature of the hydrophobic contacts between the β -strands, its shape and the integrity of native packing may govern the stability of the fibril overall. Simulations of the A β _{9-40/42} found that the formation of a stable bend between residues 23 and 28 (D₂₃VGSNK₂₈) could promote β -strand formation of the central hydrophobic cluster (CHC) located from residues 17-21 (L₁₇VFFA₂₁), which otherwise exists as a random coil in solution [305]. To analyze the importance of packing in the CHC and the bend region in protofibrils that are still salt bridge-competent, we generated two mutants, F19G and I32G/L34G. Despite deformations in the packing of β 2 against the CHC (the Gly19-Gly38 contact and the resulting bend in the β 2 region), native-like packing of Ile32 and Leu34 against the salt bridges is observed in the case of the F19G mutant, resulting in native-like internal hydration and a ΔG of binding between peptides A and B that is essentially the same as that of the WT protofibril. This result makes sense in light of the structural studies of Lührens *et al.*, who determined that A β F19G mutant peptides could form fibrils, albeit morphologically-distinct from those produced by WT A β [278]. The hydrophobic contacts involving Ile32 and Leu34 in the WT and F19G protofibrils are obviously absent in the I32G/L34G mutant, resulting in increased hydration of the Asp23-Lys28 salt bridges, a phenomenon that diminishes the magnitude of the binding energy of peptide A to peptide B.

If the Asp23-Lys28 salt bridge is perturbed, native side chain contacts and packing in the CHC may not be preserved, and as a result, the protofibril structure is substantially less stable than in the case of the wild type. The packing interaction between the aliphatic portion of the Lys28 side

chain with that of Ile32 aids in maintaining the necessary balance of hydration. Together with Leu34, the Ile32-Lys28 interaction defines a core region of stability in the A β_{42} fibril. Previous experimental work has determined that the weak hydrophobic interaction between the aliphatic portion of a lysine side chain and other hydrophobic amino acids such as tryptophan and isoleucine can contribute to the stability of β -hairpin peptides [323, 324]. Our work is the first demonstration that this phenomenon is an important contribution to the stability of A β fibrils.

7.7 Conclusions

Using explicit solvent MD simulations, SMD, and umbrella sampling, we have demonstrated several key factors that are important to the stability of the A β_{42} fibril. In agreement with the results of Masman *et al.* [303], we find that packing of the $\beta 2$ region of the protofibril against $\beta 1$, especially in the vicinity of the Asp23-Lys28 salt bridge, is important to the stability of the protofibril. Expanding upon those findings, we propose that residues Ile32 and Leu34 pack in the core of the fibril to help maintain the stability of the Asp23-Lys28 salt bridge by allowing a narrow channel of water to be present while simultaneously preventing overexposure of these charged residues to solvent. If Ile32 and Leu34 become destabilized, the salt bridge is disrupted and quickly solvated, as in the case of our I32G/L34G mutant. The absence of a stable salt bridge, modeled by our D23A and K28A mutants, leads to substantially reduced binding energy between the terminal peptide and its binding partner. Thus, we propose that compounds designed to either interfere with the hydrophobic packing interactions of residues Ile32 and Leu34 or directly disrupt the Asp23-Lys28 salt bridge may be viable therapeutic candidates to destabilize A β fibrils in the treatment of Alzheimer's disease.

Chapter 8

Destabilizing $A\beta$ Protofibrils with Morin

Copyright statement: The contents of this chapter are reprinted with permission from J. A. Lemkul and D. R. Bevan (2010) “Destabilizing Alzheimer’s $A\beta_{42}$ Protofibrils with Morin: Mechanistic Insights from Molecular Dynamics Simulations.” *Biochemistry* **49** (18): 3935-3946. <http://pubs.acs.org/articlesonrequest/AOR-6ZvRyDCY4HS9VsG5s7c8> Copyright (2010) American Chemical Society.

Attribution: J. A. L. and D. R. B. designed the research. J. A. L. conducted the simulations and performed analysis. J. A. L. and D. R. B. wrote the paper.

8.1 Introduction

Though the principal toxic species in Alzheimer’s disease are believed to be soluble, oligomeric assemblies of $A\beta$ [84, 88, 325, 326], numerous studies have indicated that fibrils also contribute to neurotoxicity [83, 87, 278, 327]. It has also been noted that pro-oxidant metal ions, such as copper, zinc, and iron, are found at high levels in amyloid deposits in the brain [62], thus indicating that fibrils and plaques may contribute to the persistent oxidative stress that exists in the Alzheimer’s disease state. Flavonoid antioxidants may function to ameliorate this oxidative stress by chelating metals and scavenging radicals *in vivo*, but both *in vitro* and *in vivo* evidence suggests that flavonoids physically disrupt $A\beta$ aggregates, as well. It is this information that motivates the present study.

Given the difficulties in obtaining detailed structural data from oligomeric species and the relative abundance of structural information regarding mature fibrils [278, 328, 329, 330], analysis of fibrils lends itself well to detailed mechanistic studies, and thus these structures will be modeled here, although it is worth mentioning recent experimental success in deriving structural data from $A\beta$ oligomers [276], as well as several simulation studies that have been conducted on such systems to examine the $A\beta$ aggregation pathway [331, 332]. Knowledge of the reasons for fibril

susceptibility may also lead to a greater understanding of the interactions between flavonoids and A β in general, opening up the possibility of targeting the more toxic soluble oligomers, as well. Molecular dynamics (MD) simulations are useful in examining such atomic-level behavior that is otherwise obscured from most experimental techniques. Here, we employ explicit-solvent, atomistic simulations on the sub-microsecond time scale to study the interaction of morin with a model of the A β_{42} fibril.

8.2 Derivation of Morin Parameters

The GROMOS96 53A6 parameter set [174] utilizes free energy calculations to obtain the free energy of solvating amino acid side chain analogs in both polar and nonpolar media (water and cyclohexane, respectively). By modifying Lennard-Jones parameters and partial charges on the atoms, the resulting values were tuned to match experimentally-determined energies [333]. The goal of the parameter set is to accurately describe partitioning behavior of amino acids between hydrophobic and hydrophilic media, such that protein folding and dynamics can be accurately predicted.

Free energy calculations are typically conducted in one of two ways, either free energy perturbation (FEP), whereby one species is mutated to another and the resulting difference in energy ($\Delta\Delta G$) quantified, or thermodynamic integration (TI), whereby the change in free energy of the system (ΔG) is quantified to describe the introduction or removal of a specific chemical entity. The latter scheme is applied here. To conduct TI calculations, the change in the system is described by a coupling parameter λ , which describes the extent of change between two states. State A ($\lambda = 0$) typically describes the “fully-coupled” state, such that the entity to be removed is fully interacting with the surrounding system. State B ($\lambda = 1$) then describes the “fully-uncoupled” state, such that the decoupled entity does not interact with the system at all. By conducting simulations at discrete values of λ between 0 and 1, the area under the $\frac{\partial H}{\partial \lambda}$ vs. λ curve can be calculated. This quantity represents the ΔG for removing the species of interest from the system, and is expressed by the following:

$$\Delta G = \int_{\lambda=0}^{\lambda=1} \left\langle \frac{\partial H(\lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda \quad (8.1)$$

To properly describe the flavonoid morin in the 53A6 force field, parameters must be derived in a manner consistent with the original derivation of the parameter set. Since ΔG values for morin hydration are not available experimentally, an alternate derivation scheme had to be devised. Experimentally-determined values of $\log P$ (Equation 8.2) for various flavonoids are available [334, 335, 336]. Thus, it should be possible to calculate the free energy change of introducing morin into both octanol and water, with the resulting ΔG of transferring the molecule between

those media being equivalent to the partitioning energy, which keeps in the spirit of the derivation of the GROMOS96 53A6 parameter set.

$$\log P = \log \left(\frac{[solute]_{octanol}}{[solute]_{water}^{un-ionized}} \right) \quad (8.2)$$

The thermodynamic cycle used for these calculations is shown in Figure 8.1.

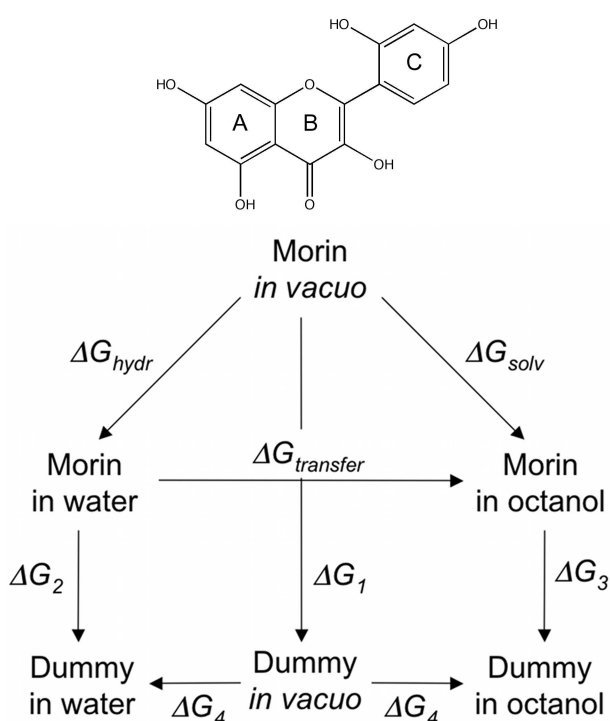


Figure 8.1: Thermodynamic cycle for the hydration of morin.

From the thermodynamic cycle in Figure 8.1, the following relationships can be established:

$$\Delta G_{hydr} = \Delta G_1 + \Delta G_4 - \Delta G_2 \quad (8.3)$$

$$\Delta G_{solv} = \Delta G_1 + \Delta G_4 - \Delta G_3 \quad (8.4)$$

Since ΔG_4 corresponds to the interaction energy of a non-interacting (dummy) species in solvent, this term in each equation is zero. Since ΔG_1 is common to both ΔG_{hydr} and ΔG_{solv} , the final expression becomes simply:

$$\begin{aligned}\Delta G_{transfer} &= \Delta G_{solv} - \Delta G_{hydr} \\ &= \Delta G_2 - \Delta G_3\end{aligned}\tag{8.5}$$

The remaining element required is a means to connect this value of ΔG with $\log P$. The classical thermodynamic relationship between ΔG and the equilibrium constant, K , is:

$$\Delta G = -RT \ln K\tag{8.6}$$

Thus the relationship between ΔG and $\log P$ is established:

$$\Delta G_{transfer} = -2.303RT \log P\tag{8.7}$$

Simulations were performed using the free energy code that is part of GROMACS, version 3.3.3 [179].

For each system, morin was centered within a cubic box. Gas-phase systems were set up using a non-periodic 10-nm box. For systems solvated in water, a 3-nm box was filled with SPC water [268]. For octanol-solvated systems, a 4-nm box was first filled with pre-equilibrated octanol, and enough water molecules were added randomly within the box to achieve the experimentally-determined 0.255 mole fraction of water:octanol at 298 K [337, 338]. This setup models the environment that morin would experience in a biphasic octanol-water system, from which $\log P$ can be calculated. To better sample the heterogeneous octanol-water systems, 3 simulations were conducted from different starting solvent configurations and random initial velocities. For simulations of morin in the gas phase and in water, only one simulation was conducted.

For each λ point, charges were turned off independently of the Lennard-Jones parameters. Scaling of the Lennard-Jones parameters utilized a soft-core potential [339], but no soft-core potential was applied to the scaling of the charges. A total of 21 λ -points was used for the integration (0, 0.005, 0.01, 0.02, 0.025, 0.05, 0.075, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 1). Hysteresis checks were conducted for each simulation, defining the fully solvent-coupled state as $\lambda = 1$, and the fully decoupled state as $\lambda = 0$. The λ values utilized in these simulations were complementary to those listed above (0, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.925, 0.95, 0.975, 0.980, 0.990, 0.995, 1). The asymmetry in the λ -point spacing was based on work by Hess and van der Vegt [340], who suggested that increased data collection at λ values near the fully coupled state was necessary to properly define the $\frac{\partial H}{\partial \lambda}$ curve for compounds that interact with the solvent through hydrogen bonding.

At each value of λ , the morin systems were energy-minimized using the steepest descents method, equilibrated for 10 ps under an isochoric-isothermal (NVT) ensemble, and then equilibrated for

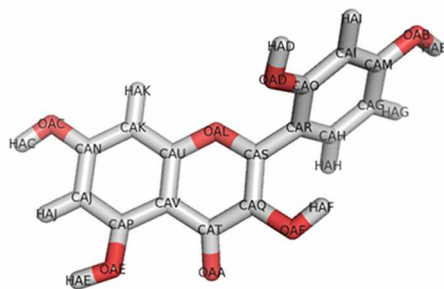


Figure 8.2: The structure of morin, with atom names labeled.

another 100 ps under an isothermal-isobaric (NPT) ensemble. Production simulations were conducted for 5 ns under an NPT ensemble. All procedures utilized a Langevin integrator [341] (GROMACS “sd” integrator) to maintain the temperature of the system at 298 K, and the NPT ensembles added the Berendsen weak coupling method [187] to maintain pressure at 1 bar. All bond lengths were constrained using LINCS [300], allowing an integration step of 2 fs. For gas-phase systems, no cutoffs were applied and the neighbor list was fixed. For condensed-phase systems, long-range electrostatics were calculated using the smooth Particle Mesh Ewald (PME) method [193, 194], with the real-space contribution to the Coulombic interactions truncated at 0.9 nm. Van der Waals interactions were truncated at 1.4 nm, with dispersion correction applied to the energy and pressure terms. The neighbor list was updated every 5 simulation steps (10 fs).

The parameters utilized for morin are shown in Table 8.1 (with atom names corresponding to those shown in Figure 8.2), which yielded the results shown in Table 8.2.

The plots of $\frac{\partial H}{\partial \lambda}$ vs. λ for decoupling morin in water are shown in Figures 8.3 and 8.4.

The area between the curves in Figures 8.3 and 8.4 are approximately zero, and the values of $\frac{\partial H}{\partial \lambda}$ at $\lambda = 0.5$ are equal in value and opposite in sign for both curves. These results indicate that there are no hysteresis artifacts in the data; the systems are stable.

Values for ΔG_{hydr} for several flavonoids have been calculated using *ab initio* quantum mechanics (QM) calculations [342, 343, 344]. Table 8.2 summarizes the results of TI calculations for morin, quercetin, and kaempferol, comparing the obtained values of $\log P$ and ΔG_{hydr} with previously-reported data. In general, there is good agreement between the QM-derived values of ΔG_{hydr} and those calculated by TI. The values of $\log P$ tend to be overestimated, but this is a known limitation of the GROMOS96 53A6 parameter set [345]. Additionally, the heterogeneity of the biphasic octanol-water system may introduce some error. Though several simulations were completed, the initial placement of water around the flavonoid solute will influence the result.

Table 8.1: Nonbonded parameters for morin

Atom name	Atom type	Charge	Charge group
OAF	OA	-0.548	1
HAF	H	+0.398	1
CAQ	C	+0.150	1
CAT	C	+0.424	2
OAA	O	-0.424	2
CAV	C	0.000	2
CAP	C	+0.150	3
OAE	OA	-0.548	3
HAE	H	+0.398	3
CAJ	CR1	-0.100	4
HAJ	HC	+0.100	4
CAN	C	+0.150	5
OAC	OA	-0.548	5
HAC	H	+0.398	5
CAK	CR1	-0.100	6
HAK	HC	+0.100	6
CAU	C	+0.250	7
OAL	OA	-0.500	7
CAS	C	+0.250	7
CAR	C	0.000	8
CAH	CR1	-0.100	8
HAH	HC	+0.100	8
CAG	CR1	-0.100	9
HAG	HC	+0.100	9
CAM	C	+0.150	10
OAB	OA	-0.548	10
HAB	H	+0.398	10
CAI	CR1	-0.100	11
HAI	HC	+0.100	11
CAO	C	+0.150	12
OAD	OA	-0.548	12
HAD	H	-0.398	12

Table 8.2: Thermodynamic and log P data

Species	log P (expt)	log P (TI)	ΔG_{hydr} (QM, kcal mol ⁻¹)	ΔG_{hydr} (TI, kcal mol ⁻¹)
Morin	1.27 \pm 0.07	2.63 \pm 0.42*	-32.10	-32.32
Quercetin	1.82 \pm 0.32	2.99 \pm 0.49 [†]	-32.83	-30.86
	2.74 \pm 0.14		-23.94	
			-33.03	
Kaempferol	3.11 \pm 0.54	3.99	N/A	-28.9
	3.35 \pm 0.15			

*Standard deviation from n = 4 simulations.

[†]Standard deviation from n = 3 simulations.

8.3 Simulations of A β Protofibrils in the Presence of Morin

8.3.1 Methods

The structure chosen to serve as a model of the mature A β_{42} fibril is a pentapeptide segment of the A β protofilament, as determined by solid-state NMR [278]. The structure features an N-terminal β -strand (termed β_1 , residues 18-26) and a C-terminal β -strand (β_2 , from residues 31-42), with a bend region connecting the two. While this model of the A β fibril has been a popular one for simulation studies and interpretation of experimental data, it should be noted that it has recently been reported that the conditions under which A β fibrils are grown may influence their structural characteristics [328, 329]. This pentameric structure is the repeat unit of the mature A β fibril, and in isolation, is best described as a “protofibril,” distinct from soluble oligomeric species of A β . Thus, in this work, we refer to the structure as such. The N-terminal 16 residues of each peptide are missing in this structure, so the N-terminus of each peptide was capped with an acetyl group, giving uncharged N-termini. The residue numbering scheme used here refers to amino acid positions in the full-length A β peptide, for ease of comparison to other experimental and theoretical studies on A β . Previous work [303, 346] has concluded that this region of A β_{42} (residues 17-42) is principally responsible for the stability of the mature fibril, and thus this core region, lacking the N-terminal 16 residues, serves as a suitable model of the full-length fibril and is similar to model systems previously used for simulation studies [202, 304, 307, 308, 316, 347]. Several different systems were prepared, and are described in Sections 8.3.1.1 - 8.3.1.4.

All simulations were conducted using the GROMACS package, version 4.0.2 [179, 270], with parameters from the GROMOS96 53A6 parameter set [174] applied to the protein, water, ions, and morin (discussed in Section 8.2). The water model used in all cases was SPC [268]. The NaCl concentration in all simulations was approximately 100 mM, after adding Na⁺ counterions sufficient to neutralize the net charge on the protofibril. In all cases, short-range non-bonded in-

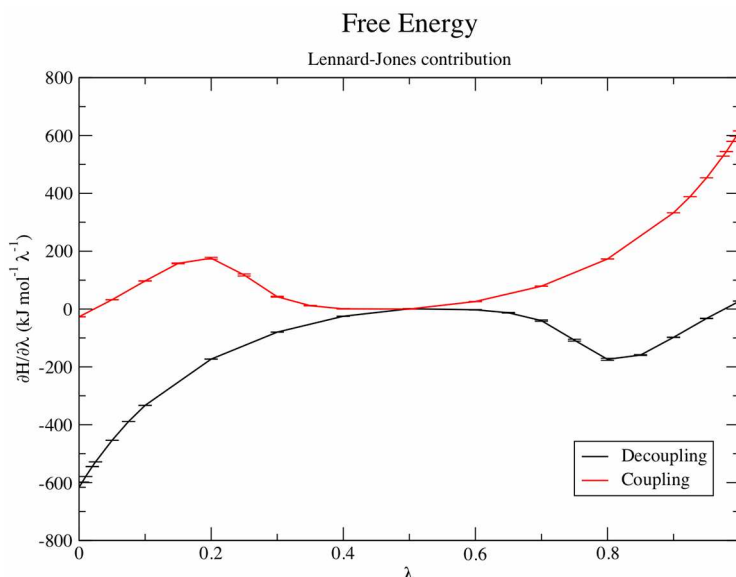


Figure 8.3: Lennard-Jones contribution to ΔG for alchemical transformations of morin in water. Error bars represent the standard deviation of each $\frac{\partial H}{\partial \lambda}$ point.

teractions were truncated at 1.4 nm, applying long-range dispersion correction to the energy and pressure terms to account for truncation of the van der Waals interactions. Long-range electrostatics were calculated using the smooth Particle Mesh Ewald (PME) method [193, 194]. All bond lengths were constrained with P-LINCS [180], allowing an integration time step of 2 fs. Periodic boundary conditions were applied in all directions.

In each simulation set, systems were simulated for at least 100 ns; exact times are noted in the detailed description that follows. General preparation steps were common to all systems, again, with minor differences noted below. Once built and energy-minimized using the steepest descents method, each system was equilibrated under NVT conditions for 50 ps at 310 K, the incubation temperature of the experimental assays [160], followed by 50 ps under NPT conditions, applying position restraints to the peptides. Weak coupling [187] was applied during equilibration, but for data collection, the Nosé-Hoover thermostat [188, 189, 190] and Parrinello-Rahman barostat [192, 191] were employed to generate a rigorous NPT ensemble.

8.3.1.1 Set I - Control $A\beta$ Protofibrils

The six simulations in Set I were the negative controls, consisting only of the $A\beta$ protofibril structure in an aqueous solution of 100 mM NaCl. Set I simulations serve to expose the inherent fluctuations in the $A\beta$ protofibril structure under physiologically-relevant conditions. To construct these systems, the pentameric structure was centered in a 6.9-nm cubic box, to which water and ions were added, as discussed above. Set I simulations were conducted for 100 ns.

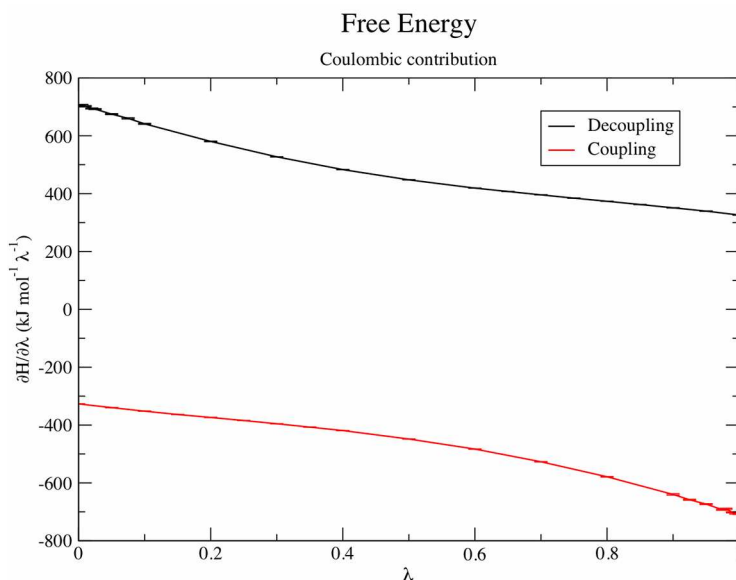


Figure 8.4: Coulombic contribution to ΔG for alchemical transformations of morin in water. Error bars represent the standard deviation of each $\frac{\partial H}{\partial\lambda}$ point.

8.3.1.2 Set II - $A\beta$ Protofibrils with Diffuse Morin

The six systems in Set II included 10 morin molecules in the solvent around the $A\beta$ protofibril structure, giving a 2:1 mole ratio of compound:peptide, as in the fibril destabilization assays conducted by Ono *et al.* [160]. These systems were prepared as in Set I, with the morin added at random positions within the unit cell prior to solvation and the addition of ions. These systems were simulated for up to 400 ns for reasons described in Section 8.3.2.

8.3.1.3 Set III - $A\beta$ Protofibrils with Inserted Morin

For the six simulations in Set III, a single molecule of morin was steered into the core of the $A\beta$ protofibril using the center-of-mass (COM) pulling capability of GROMACS. The target was a small opening in the protofibril present in the initial NMR structure, described in greater detail in Section 8.3.2. To these systems, nine additional morin molecules were inserted at random positions around the protofibril structure to give the same mole ratio as in Set II. All simulations were conducted for 400 ns.

8.3.1.4 Set IV - Assembly Mechanism of the $A\beta$ Protofibrils

The final set of simulations (Set IV) applied COM pulling to draw a free peptide towards the $A\beta$ protofibril structure, in the presence and absence of morin. The starting structure for the $A\beta$

protofibril in simulations IV-1, IV-2, and IV-3 (without morin) were taken from the final configuration of simulation I-1, a representative result of that set. For simulations IV-4, IV-5, and IV-6 (treated with morin), the protofibril and morin coordinates were taken from a representative snapshot of simulation II-5. A free peptide was generated by extracting the coordinates of one $A\beta$ peptide from the original NMR structure and simulating it for 100 ns in SPC water with 100 mM NaCl. This simulation was judged stable by examining the convergence of intrapeptide hydrogen bonds, solvent-accessible surface area (SASA), radius of gyration, root mean square deviation (RMSD), and the number of intrapeptide atomic contacts. This peptide, having adopted a collapsed-coil structure with little β -strand content resembling the structure described by Zhang *et al.* [80], was placed within a rectangular box 5.0 nm away from the closest peptide in the protofibril, directly along the z -axis. Prior to pulling, the system was equilibrated under an NPT ensemble for 100 ps. After equilibration, the free peptide was pulled towards the protofibril structure. During the pulling, position restraints were placed on the $C\alpha$ atoms of the center peptide (peptide C, Figure 7.4) of the protofibril to establish a fixed target location. After pulling the free peptide along the z -axis (directly coincident with the protofibril axis) to achieve a roughly 1.5-nm COM distance between the free peptide and nearest peptide in the protofibril, the biasing force was released and unrestrained simulations were conducted for 100 ns, with each independent simulation initiated with different starting velocities.

8.3.2 Results

8.3.2.1 Dynamics of Untreated $A\beta_{42}$ Protofibrils

The Set I simulations serve as a benchmark to understanding the inherent stability of the $A\beta_{42}$ protofibrils. Although similar simulations have already been performed [303], that work used a different force field, OPLS-AA [168], and thus it was important for us to generate our own set of controls using the GROMOS96 53A6 parameter set, which was chosen for the present study. Over the course of the six independent 100-ns simulations in this set, only small deviations from the initial structure were observed (Figure 8.5). Some twisting was apparent in the β_1 and β_2 regions, such that the β -strands were not entirely co-planar, allowing for increased side chain packing. This phenomenon has been proposed previously as a contributing factor to amyloid fibril stability [202, 303, 304, 314].

The peptides are labeled A through E, with peptide A in blue and peptide E in red in Figure 8.5, the same labeling scheme used in Figure 7.4. Peptide A has been proposed to be the end of the fibril to which incoming peptides attach in the unidirectional model of amyloid fibril growth [278].

On average, each peptide chain (across all six simulations) stabilized at an RMSD value of approximately 0.25 ± 0.01 nm, indicating a very low level of overall structural change (Figure 8.6, Table 8.3). From the RMSF data, it is clear that the most flexible regions of each peptide are the C-terminus (residues 40-42) and those residues in the vicinity of the bend region (residues 25-30) that connects the two β -strands in each peptide (Figure 8.7). Interchain backbone hydrogen bonds were

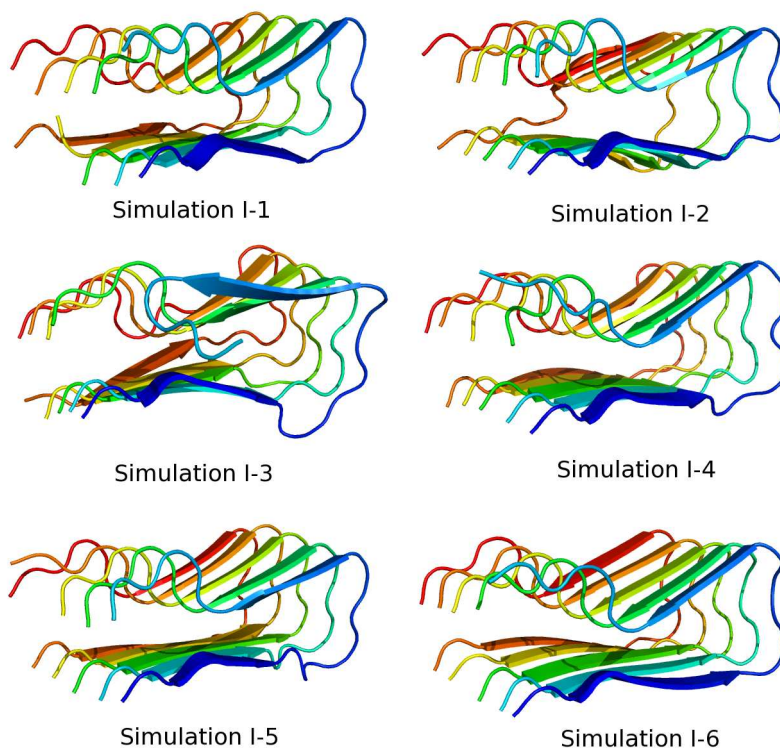


Figure 8.5: Snapshots from the end of Set I simulations (100 ns each).

Table 8.3: Backbone RMSD (nm, average over 50 ns $\pm \sigma$) from Set I protofibril simulations

System	Chain A	Chain B	Chain C	Chain D	Chain E
1	0.25 ± 0.02	0.24 ± 0.01	0.24 ± 0.01	0.27 ± 0.01	0.25 ± 0.02
2	0.25 ± 0.02	0.24 ± 0.01	0.24 ± 0.01	0.23 ± 0.01	$0.51 \pm 0.03^{*\dagger}$
3	$0.43 \pm 0.06^\dagger$	0.22 ± 0.02	0.21 ± 0.01	0.22 ± 0.01	0.28 ± 0.02
4	0.26 ± 0.03	0.24 ± 0.01	0.24 ± 0.01	0.24 ± 0.01	0.26 ± 0.02
5	0.27 ± 0.02	0.25 ± 0.01	0.25 ± 0.01	0.25 ± 0.01	0.28 ± 0.02
6	0.26 ± 0.04	0.25 ± 0.02	0.24 ± 0.02	0.24 ± 0.02	0.28 ± 0.03
Average	0.26 ± 0.01	0.24 ± 0.01	0.23 ± 0.01	0.24 ± 0.02	0.27 ± 0.01

*Average taken over the last 40 ns.

[†]Outlier, as determined by Q-test with 99% confidence. Not included in average statistics.

also very stable across all simulations, with an average of 22.0 ± 0.8 hydrogen bonds (averaged over all neighboring peptide pairs, A-B, B-C, C-D, and D-E) persisting throughout all of the 100-ns trajectories (Figure 8.8, Table 8.4). Also of interest in these simulations is the stability of the interpeptide Asp-Lys salt bridges. Across all six simulations, these interactions were very stable,

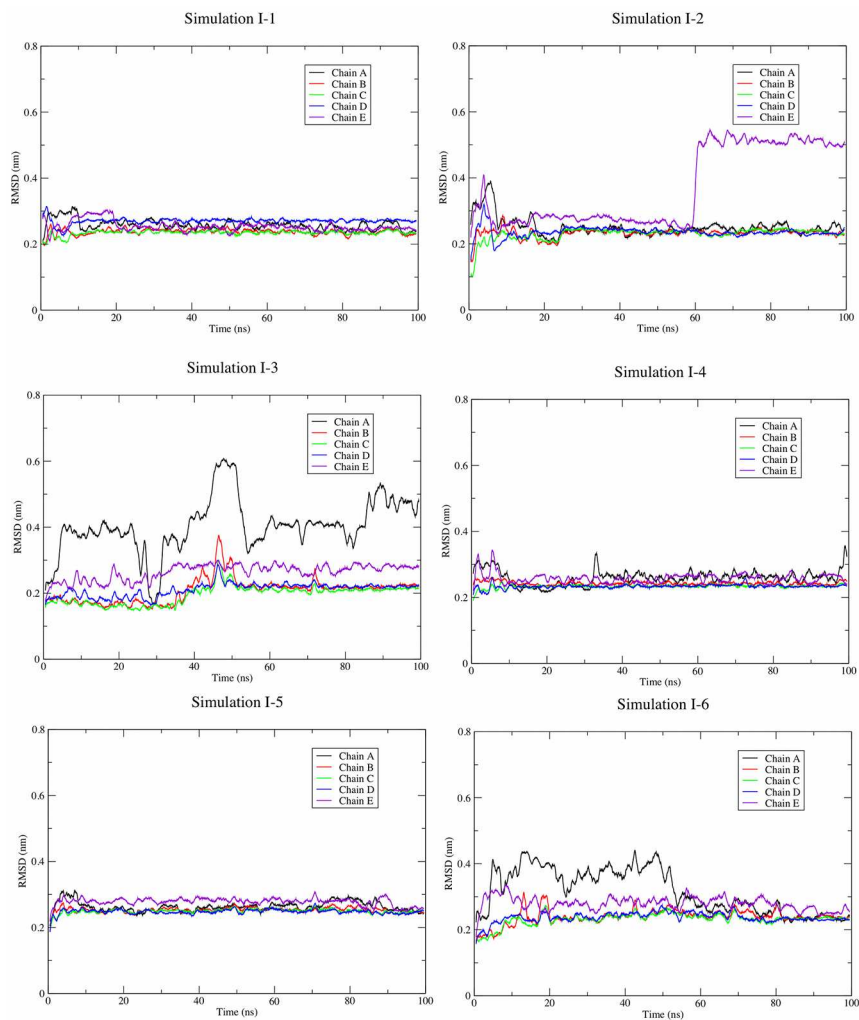


Figure 8.6: Backbone RMSD from Set I simulations. RMSD values for chain E (I-2) and chain A (I-3) were identified as outliers (see Table 8.3).

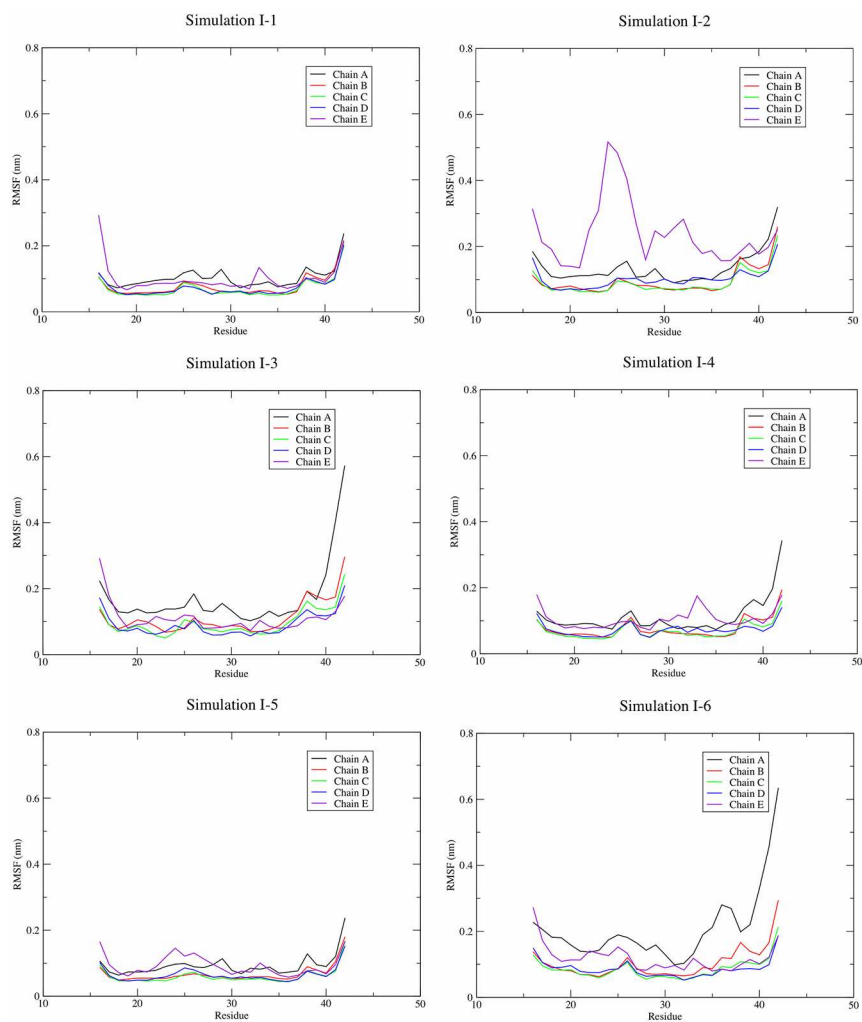


Figure 8.7: Backbone RMSF, averaged per residue for Set I simulations. As with RMSD, the RMSF values of chain E (I-2) and chain A (I-3) are considered outliers.

Table 8.4: Backbone hydrogen bond count (average over 50 ns $\pm \sigma$) from Set I protofibril simulations

System	Chain A - B	Chain B - C	Chain C - D	Chain D - E
1	21.7 \pm 1.5	22.0 \pm 1.5	22.5 \pm 1.3	22.0 \pm 1.6
2	21.5 \pm 1.7	23.0 \pm 1.5	22.8 \pm 1.5	16.0 \pm 1.4* [†]
3	17.7 \pm 1.6 [†]	22.9 \pm 1.4	23.3 \pm 1.4	21.3 \pm 1.7
4	20.7 \pm 1.6	22.8 \pm 1.4	22.8 \pm 1.4	21.4 \pm 1.6
5	21.3 \pm 1.5	22.2 \pm 1.5	22.8 \pm 1.5	21.1 \pm 1.8
6	20.2 \pm 2.5	22.1 \pm 1.6	22.1 \pm 1.6	21.8 \pm 1.7
Average	21.1 \pm 0.6	22.5 \pm 0.4	22.7 \pm 0.4	21.5 \pm 0.4

*Average taken over the last 40 ns.

[†]Outlier, as determined by Q-test with 99% confidence. Not included in average statistics.Table 8.5: Interpeptide Asp23 - Lys28 COO⁻ \rightarrow NH₃⁺ distances (nm, average over 50 ns $\pm \sigma$) from Set I protofibril simulations

System	Chain A - B	Chain B - C	Chain C - D	Chain D - E
1	0.34 \pm 0.06	0.31 \pm 0.02	0.31 \pm 0.02	0.34 \pm 0.06
2	0.45 \pm 0.02	0.32 \pm 0.03	0.32 \pm 0.03	0.70 \pm 0.29* [†]
3	0.38 \pm 0.06	0.30 \pm 0.02	0.31 \pm 0.02	0.38 \pm 0.04
4	0.35 \pm 0.08	0.30 \pm 0.02	0.31 \pm 0.02	0.38 \pm 0.08
5	0.43 \pm 0.01	0.30 \pm 0.01	0.31 \pm 0.01	0.35 \pm 0.05
6	0.38 \pm 0.10	0.30 \pm 0.02	0.31 \pm 0.02	0.42 \pm 0.12
Average	0.39 \pm 0.05	0.31 \pm 0.01	0.31 \pm 0.01	0.38 \pm 0.03

*Average taken over the last 40 ns.

[†]Outlier, as determined by Q-test with 99% confidence. Not included in average statistics.

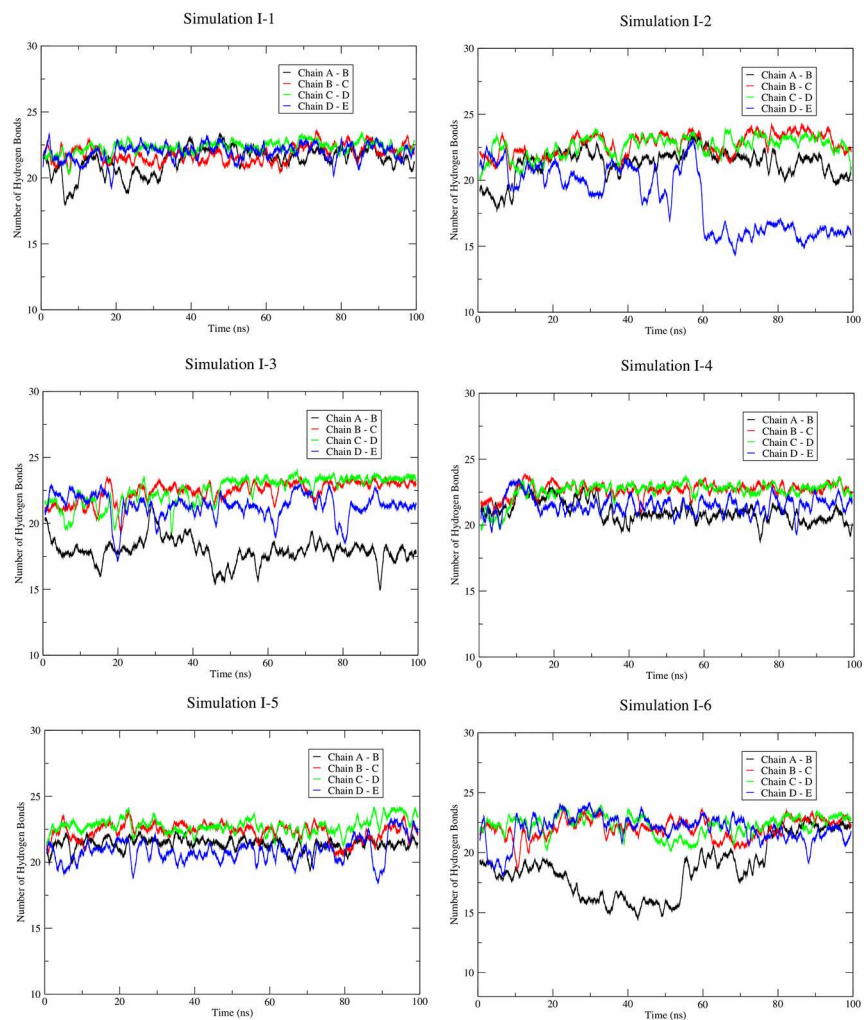


Figure 8.8: Backbone hydrogen bonds for Set I simulations. Decreased chain D-E hydrogen bonds (I-2) and chain A-B hydrogen bonds (I-3) are considered outliers (see Table 8.4).

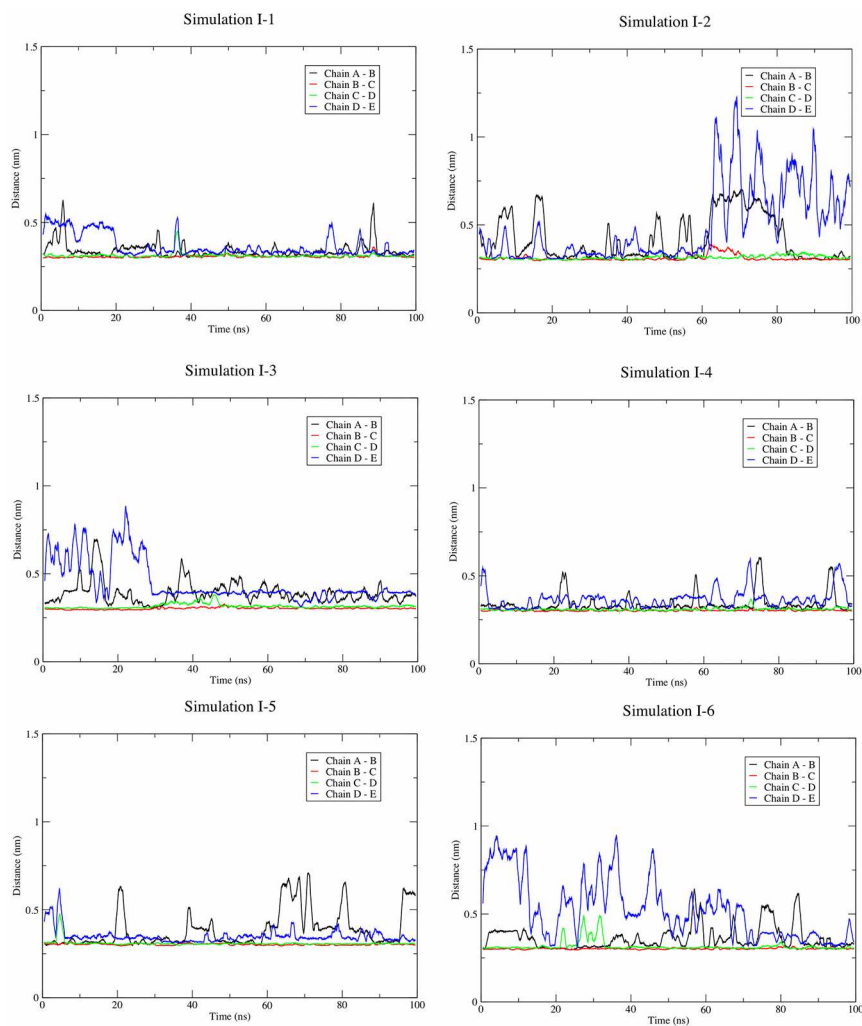


Figure 8.9: Asp23 - Lys28 ($\text{COO}^- \rightarrow \text{NH}_3^+$) distance for Set I simulations. The chain D-E salt bridge distance (I-2) was identified as an outlier (see Table 8.5).

with the average distance between the $N\zeta$ -amino groups and corresponding $C\gamma$ -carboxylates being approximately 0.35 ± 0.04 nm, averaged over all peptides and all six simulations. (Figure 8.9, Table 8.5).

8.3.2.2 The Influence of Morin in Solution with $A\beta_{42}$ Protofibrils

Simulations as part of Set II were designed to model the assay conditions (temperature, ionic strength of the solvent, and morin concentration) used by Ono *et al.* in their fibril destabilization assays [160]. The highest concentration of morin (and thus the highest mole ratio of morin: $A\beta$, 2:1) used in the *in vitro* work was chosen for the present study in order to maximize the likelihood of the desired phenomena being observed within a timeframe that is feasible using MD simulations. Simulations were conducted for sufficient time to allow the positions of morin to reach equilibrium around the $A\beta$ protofibril. That is, simulations were stopped only after the position of morin was largely unchanged over time (net change in position of < 0.001 nm over the last half of the trajectory, and fluctuations in the COM position of each morin molecule no greater than approximately 0.01 nm ps^{-1}) and other quantities related to protein stability (RMSD, RMSF, hydrogen bonds, salt bridge distances) had converged. These criteria resulted in three simulations of 100 ns in length (II-1, II-3, and II-6), two simulations of 200 ns (II-4 and II-5), and one simulation of 400 ns in length (II-2).

Over the course of these simulations, morin rapidly deposited onto the surface of the $A\beta$ protofibril (within 10-20 ns), establishing a large number of hydrophobic contacts with amino acid side chains in the protofibril (Figures 8.10 and 8.11), a phenomenon that is not unexpected, given the large hydrophobic surface area of the protofibril structure and the low solubility of morin in water. Over time, the contacts between morin and the $A\beta$ protofibril evolved to generate two principal binding modes. The first is what we will call “capping,” such that a network of morin molecules assembles at one end of the $A\beta$ protofibril, occupying the backbone hydrogen bonding groups of the peptides (A and E) that are exposed to solvent (Figures 8.10A and 8.12, Table 8.6). This capping was present, to varying extents, in simulations II-1, II-2, II-4, II-5, and II-6. In simulation II-3, morin aggregated on the hydrophobic surface presented by the C-terminal residues. Morin-sidechain hydrogen bonding also participated in stabilizing the association of morin with the $A\beta$ protofibril (Figure 8.13 and Table 8.7). Images from the final snapshot of these simulations are shown in Figure 8.14.

The second binding mode involves the partial penetration of morin into the hydrophobic interior of the protofibril, and was observed only in simulation II-2 (Figure 8.10B). During the 400-ns trajectory, it was observed that the interaction between morin and the Asp23 and Lys28 residues in peptides A and B persisted for over 82.5% of this simulation (from the period of 10-340 ns), although by the end of the simulation, the interaction of morin with these charged residues was not present. Full entrance of morin into the hydrophobic core of the $A\beta$ protofibril was not observed in this trajectory, but the result of simulation II-2 indicated the potential for this entrance event to occur, and is explored later in Set III.

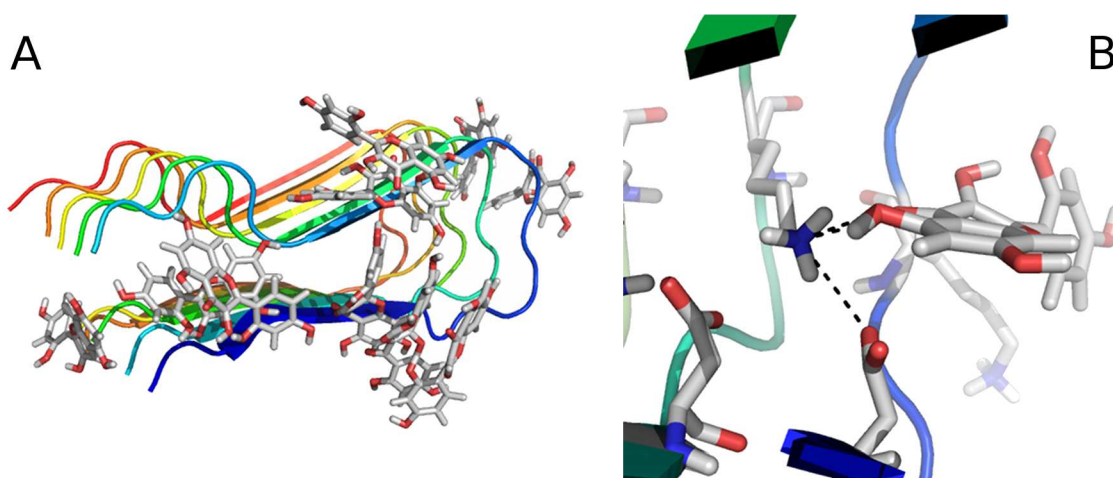


Figure 8.10: The two binding modes of morin to the $A\beta$ protofibril: (A) Capping, taken from a snapshot of simulation II-5 at 170 ns, and (B) insertion in the vicinity of the Asp23-Lys28 salt bridge, from a snapshot of simulation II-2 at 260 ns. Putative hydrogen bonds are shown as black dashed lines.

Table 8.6: Morin-backbone hydrogen bond count (average over last half of trajectory $\pm \sigma$) from Set II protofibril simulations

System	Chain A	Chain B	Chain C	Chain D	Chain E
1	2.2 ± 1.1	0.8 ± 0.3	0.1 ± 0.4	0.3 ± 0.5	2.7 ± 1.1
2	3.0 ± 1.5	0.3 ± 0.5	0.3 ± 0.6	0.4 ± 0.6	1.9 ± 1.4
3	3.3 ± 1.1	0.1 ± 0.2	0.6 ± 0.8	0.7 ± 0.6	0.1 ± 0.2
4	2.4 ± 1.0	0.1 ± 0.2	0.2 ± 0.4	0.1 ± 0.4	4.2 ± 1.4
5	$5.0 \pm 1.6^*$	0.5 ± 0.7	0.6 ± 0.7	0.2 ± 0.5	0.3 ± 0.6
6	0.1 ± 0.3	0.9 ± 0.3	0.4 ± 0.2	0.1 ± 0.3	4.9 ± 1.4

*Average taken over the last 50 ns.

Table 8.7: Morin-sidechain hydrogen bond count (average over last half of trajectory $\pm \sigma$) from Set II protofibril simulations

System	Chain A	Chain B	Chain C	Chain D	Chain E
1	1.8 ± 1.8	5.3 ± 2.0	4.0 ± 2.0	2.7 ± 1.7	1.8 ± 1.6
2	3.2 ± 1.4	4.1 ± 1.8	4.6 ± 1.9	3.3 ± 2.0	1.5 ± 1.4
3	3.2 ± 2.0	5.2 ± 2.2	5.0 ± 2.0	3.7 ± 2.0	2.1 ± 1.7
4	0.1 ± 0.4	3.0 ± 1.5	5.0 ± 1.9	3.3 ± 1.8	4.0 ± 1.7
5	3.7 ± 2.2	4.8 ± 1.8	4.6 ± 1.9	3.6 ± 1.9	1.0 ± 1.2
6	0.2 ± 0.6	3.8 ± 1.8	5.7 ± 2.0	4.7 ± 2.2	1.5 ± 1.4

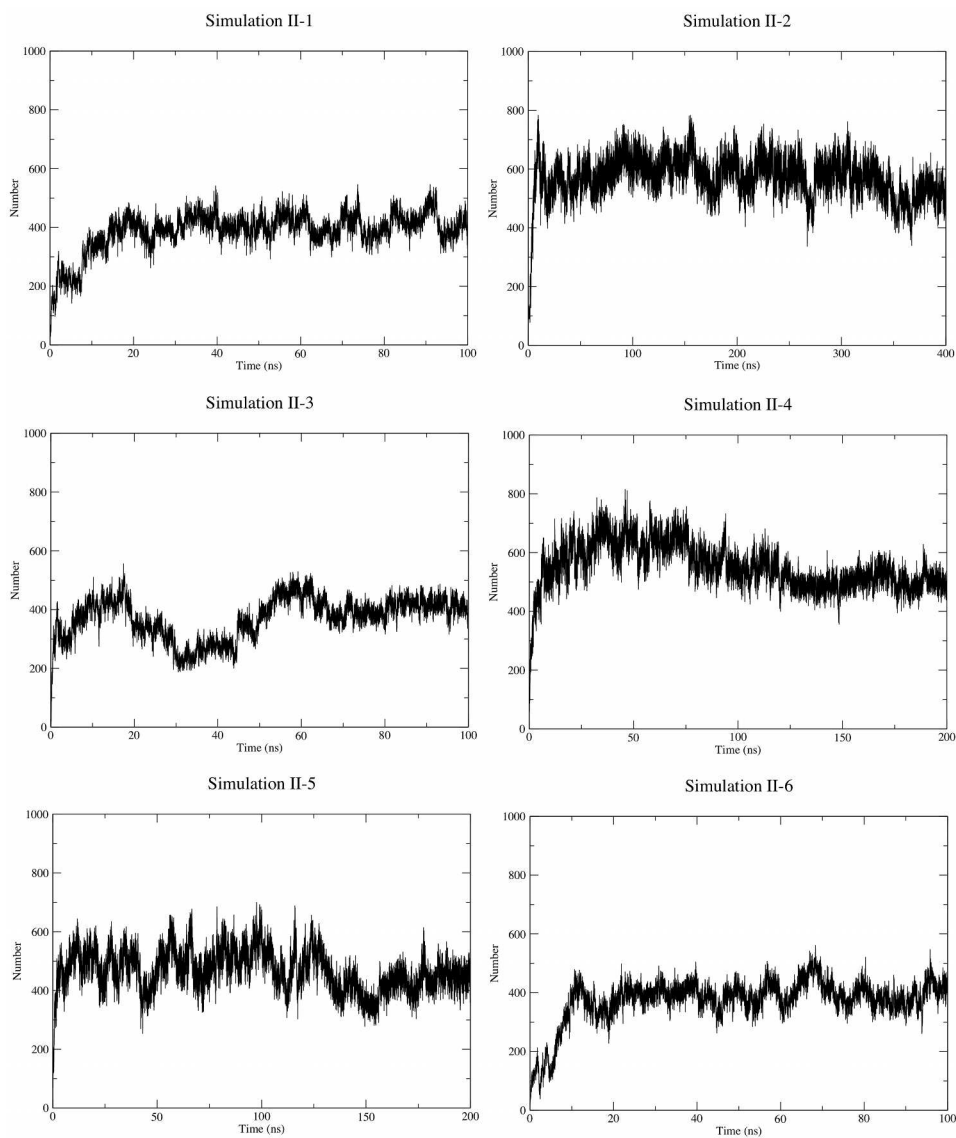


Figure 8.11: Hydrophobic contacts between morin and protofibril side chain atoms for Set II simulations. A hydrophobic contact was defined as any contact between a side chain carbon atom and any carbon atom in morin within a distance of 0.6 nm.

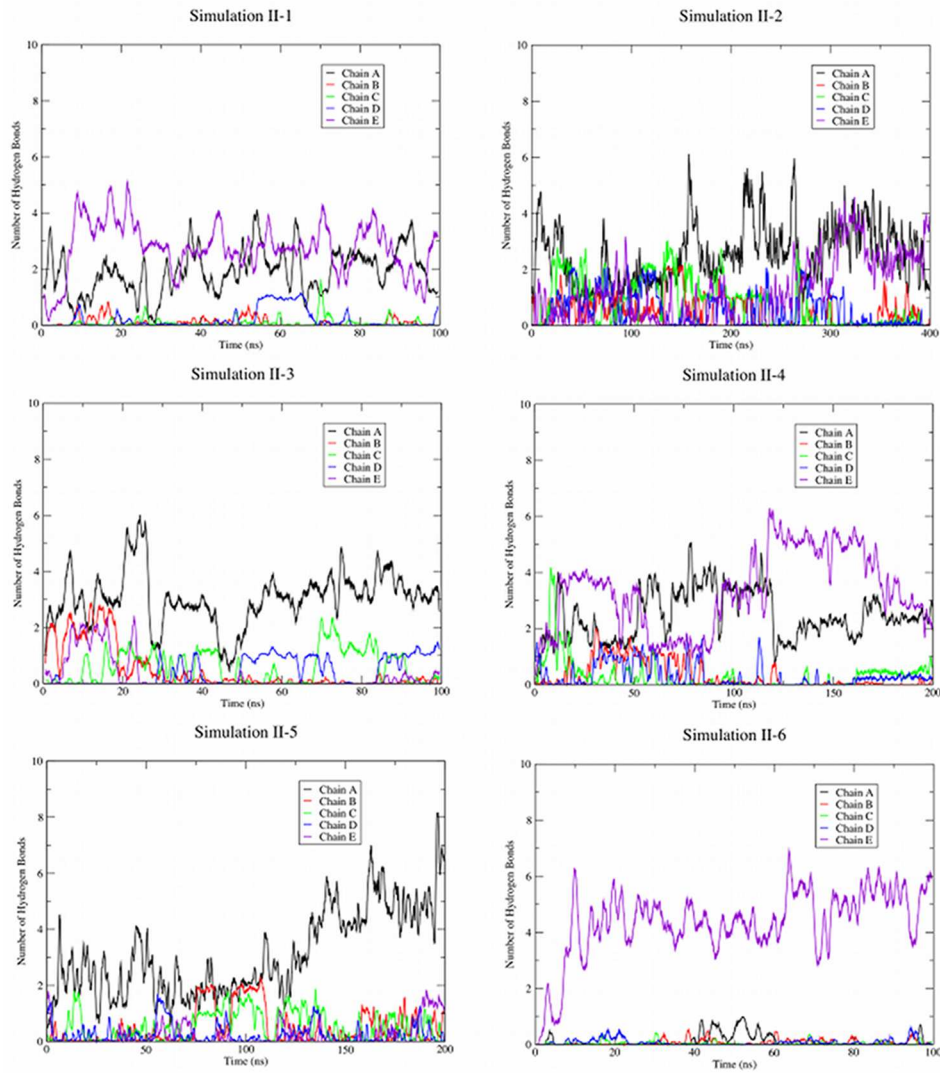


Figure 8.12: Hydrogen bonding between morin and the $A\beta$ protofibril backbone for all replicates in Set II.

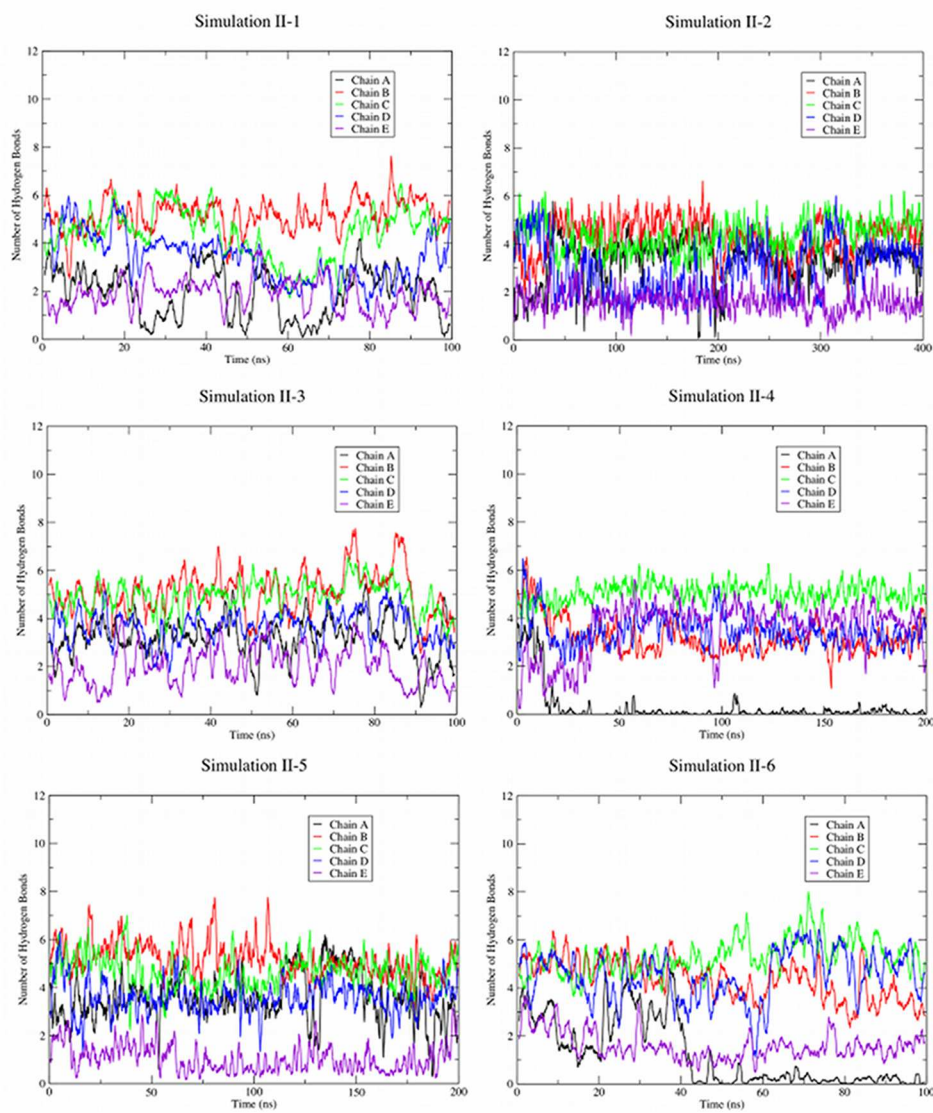


Figure 8.13: Hydrogen bonding between morin and the $A\beta$ protofibril sidechain groups for all replicates in Set II.

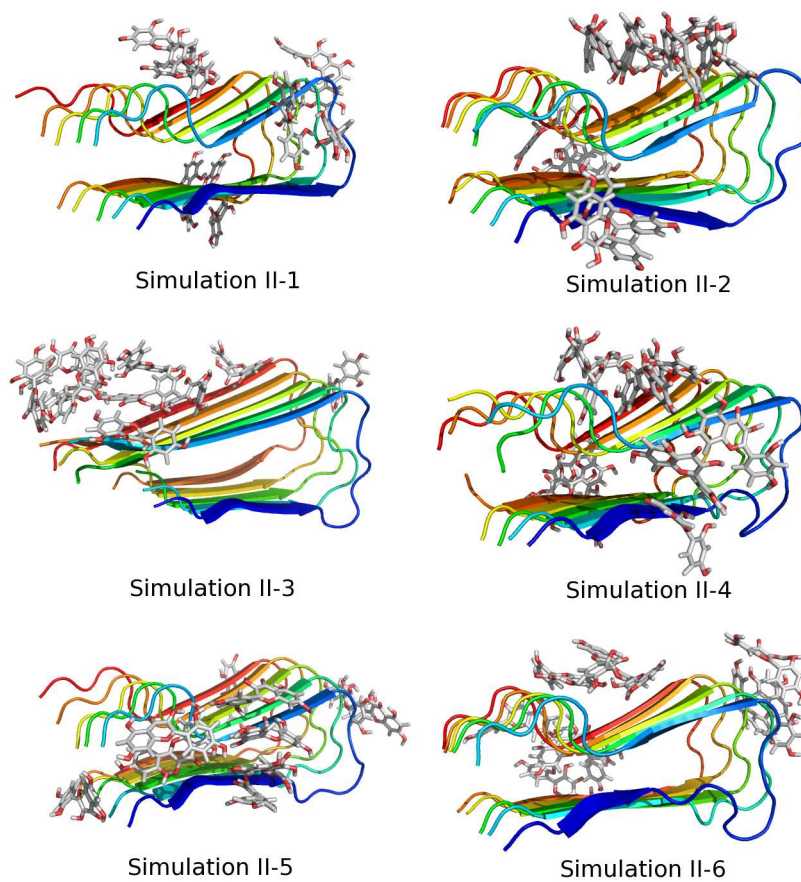


Figure 8.14: Snapshots from the end of Set II simulations.

Table 8.8: Backbone RMSD (nm, average over last half of trajectory $\pm \sigma$) from Set II protofibril simulations

System	Chain A	Chain B	Chain C	Chain D	Chain E
1	0.27 \pm 0.01*	0.27 \pm 0.03	0.26 \pm 0.03	0.25 \pm 0.02	0.29 \pm 0.06
2	0.26 \pm 0.02	0.24 \pm 0.02	0.24 \pm 0.01	0.23 \pm 0.01	0.25 \pm 0.02
3	0.37 \pm 0.03	0.32 \pm 0.02	0.30 \pm 0.02	0.29 \pm 0.02	0.25 \pm 0.03
4	0.28 \pm 0.02	0.23 \pm 0.01	0.25 \pm 0.01	0.26 \pm 0.01	0.28 \pm 0.02
5	0.27 \pm 0.03	0.24 \pm 0.01	0.24 \pm 0.01	0.24 \pm 0.01	0.25 \pm 0.02
6	0.27 \pm 0.01	0.24 \pm 0.01	0.24 \pm 0.01	0.25 \pm 0.01	0.26 \pm 0.02

*Average taken over the last 35 ns.

The binding of morin to the surface of the A β protofibril had little effect on the stability of the protofibril structure. When morin bound to the ends of the protofibrils, the RMSD and RMSF of the terminal peptide (chains A and E, with which morin is interacting most directly) did not increase substantially (Figures 8.15 and 8.16, Table 8.8). In two cases (simulations II-4 and II-6), the association of morin with peptide A attracted Asp23 of peptide A away from its native salt bridge with Lys28 in peptide B (Figure 8.17 and Table 8.9), destabilizing native backbone hydrogen bonding and indicating that peripheral association may cause some small level of destabilization of the bend region. In simulation II-4, this interaction decreased the number of backbone hydrogen bonds between peptides A and B to 17.3 ± 1.3 per timeframe, with similar results in simulation II-6 (17.9 ± 1.2) relative to the control value of this parameter, 21.1 ± 0.6 hydrogen bonds per timeframe (Figure 8.18 and Table 8.10). The partial insertion of morin in simulation II-2 caused a slight increase in the RMSF of the residues in the bend region between residues 25-30, likely due to the reorientation of Lys28 in peptide A towards the solvent, resulting in increased flexibility in this region of the peptide. Also in simulation II-2, the interaction between Lys28 of peptide B and Asp23 of peptide A was periodically destabilized (from 200 - 350 ns) by this partial morin insertion (Figure 8.17).

Capping networks of morin also destabilized this same salt bridge in the other simulations in Set II, though by a slightly different mechanism. Whereas the partial penetration of morin in simulation II-2 interfered with the Asp23-Lys28 salt bridge by competing for hydrogen bonding, capping networks destabilized the salt bridge by simply attracting Asp23 outward, towards the solvent. In some cases (simulations II-1 and II-2), this salt bridge destabilization was more periodic, and in other cases (simulation II-4 and II-6), it was more persistent (Figure 8.17). The stability of the Asp23-Lys28 salt bridge between peptides A and B in simulations II-3 and II-5 was comparable to the controls of Set I (Figure 8.17 and Table 8.9), indicating that deposition of morin in the β 2 region of the protofibril (in the case of simulation II-3) does not influence the stability of this interaction. In simulation II-5, the morin capping group formed at a greater distance from the bend region than in the other simulations, and thus few interactions between morin and Asp23 of peptide A were observed.

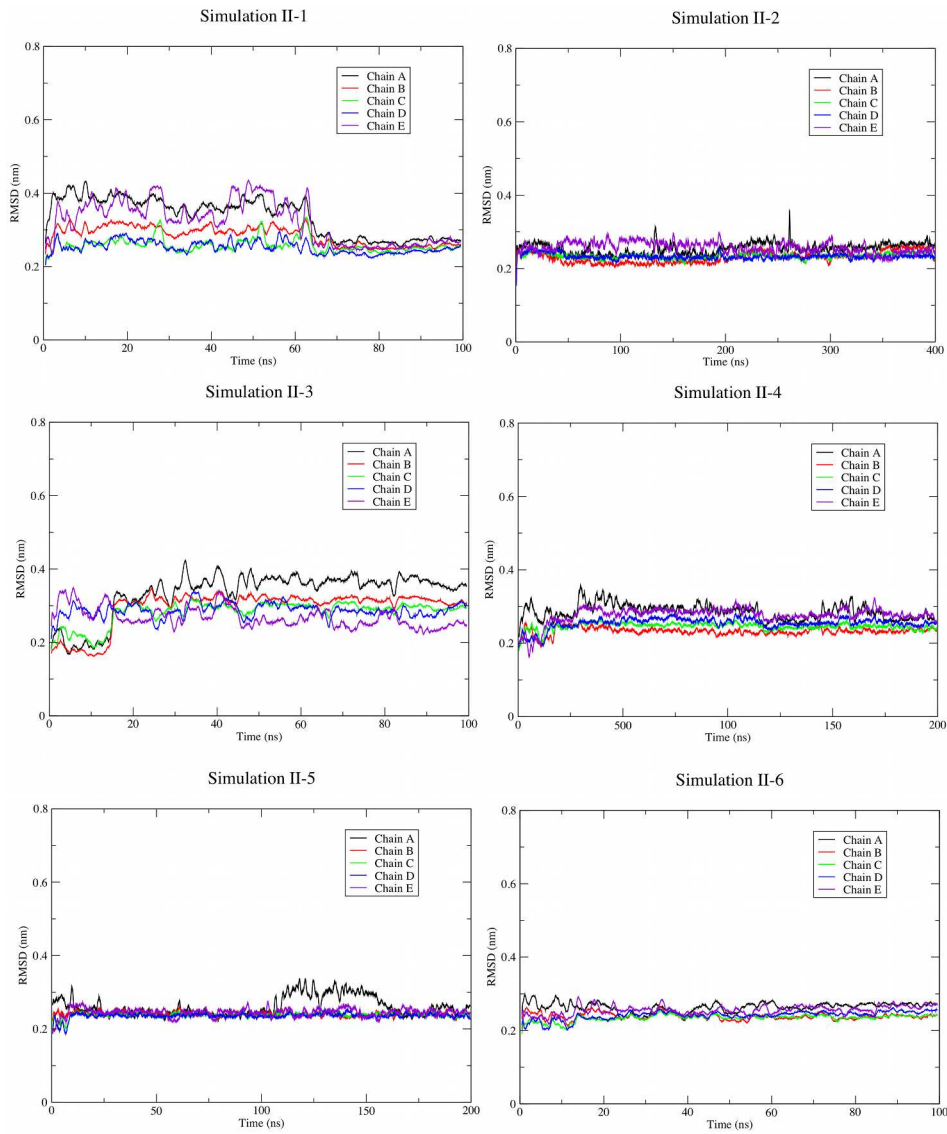


Figure 8.15: Backbone RMSD for Set II simulations.

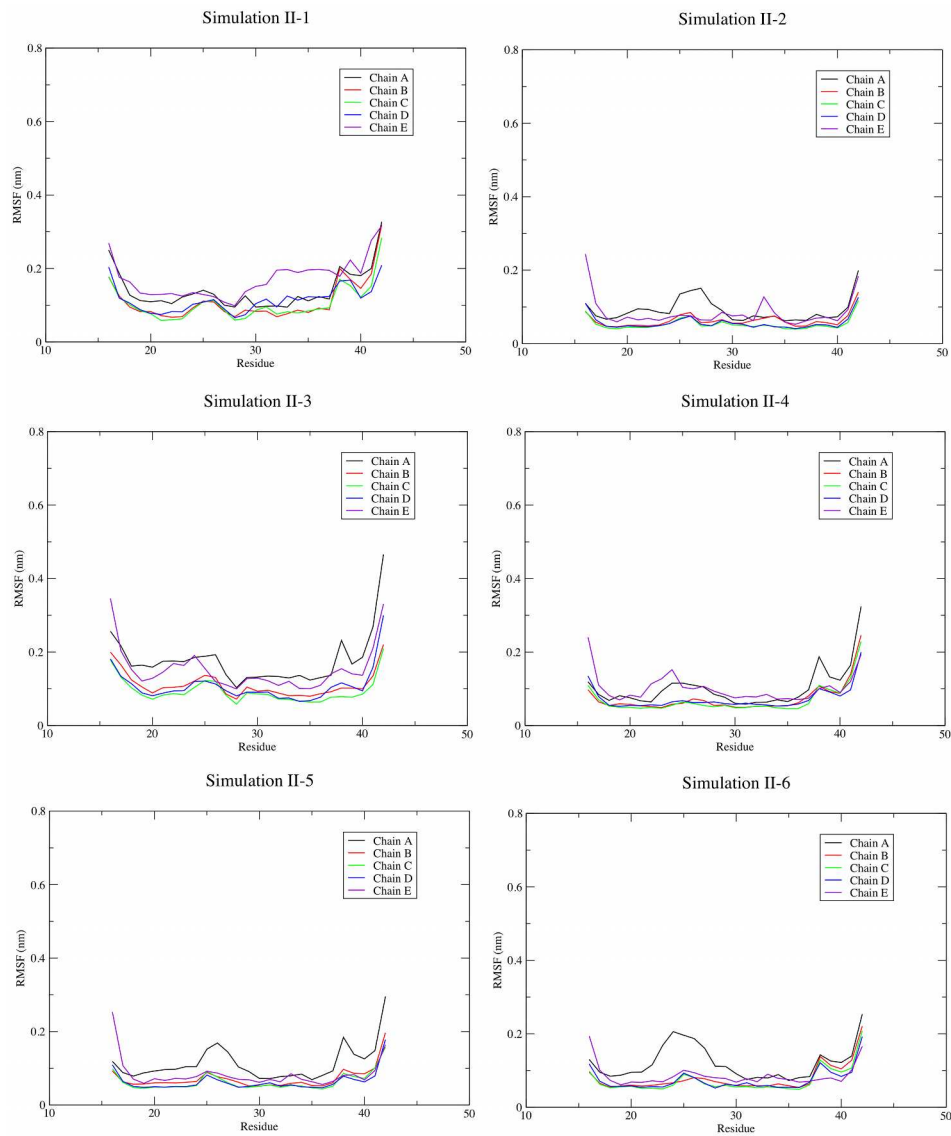


Figure 8.16: Backbone RMSF, averaged per residue for Set II simulations.

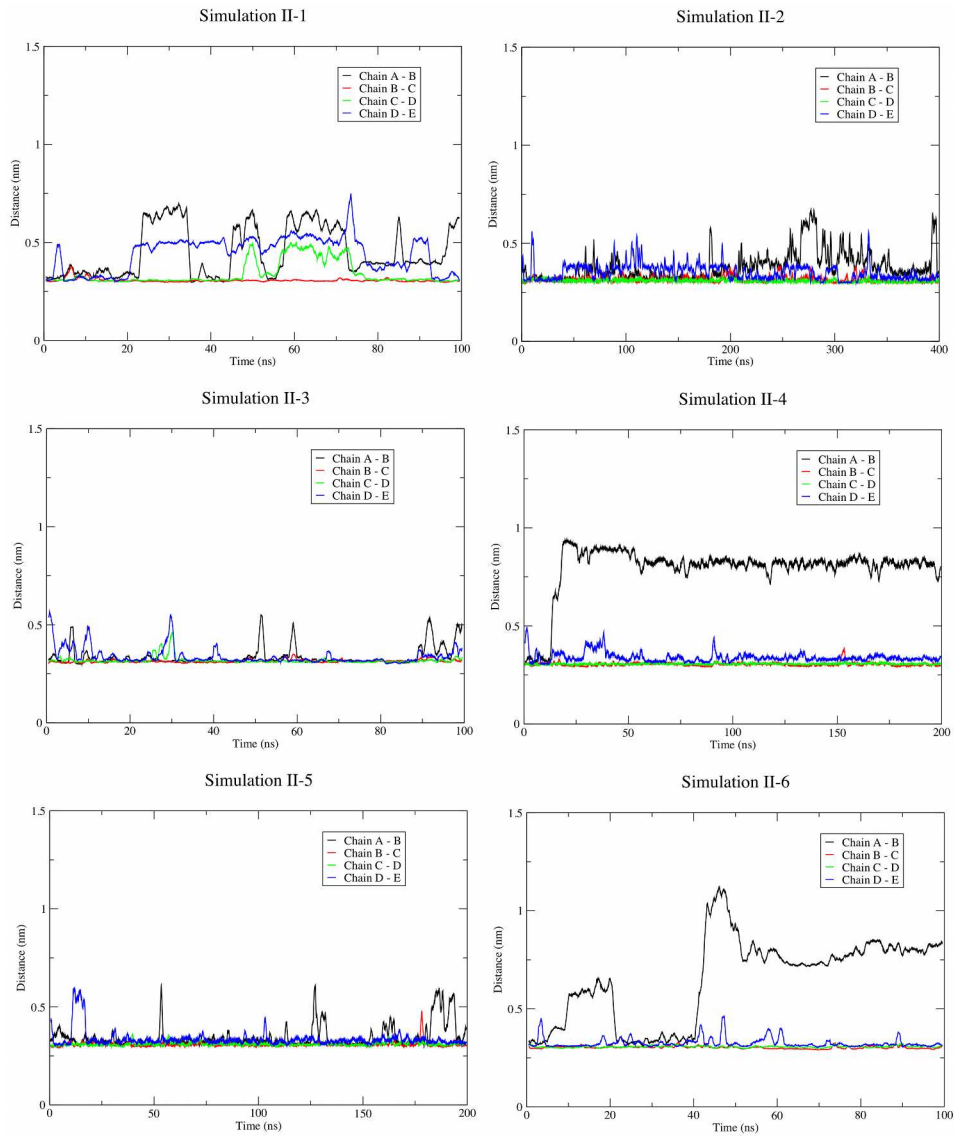


Figure 8.17: Asp23 - Lys28 (COO⁻ → NH₃⁺) distance for Set II simulations.

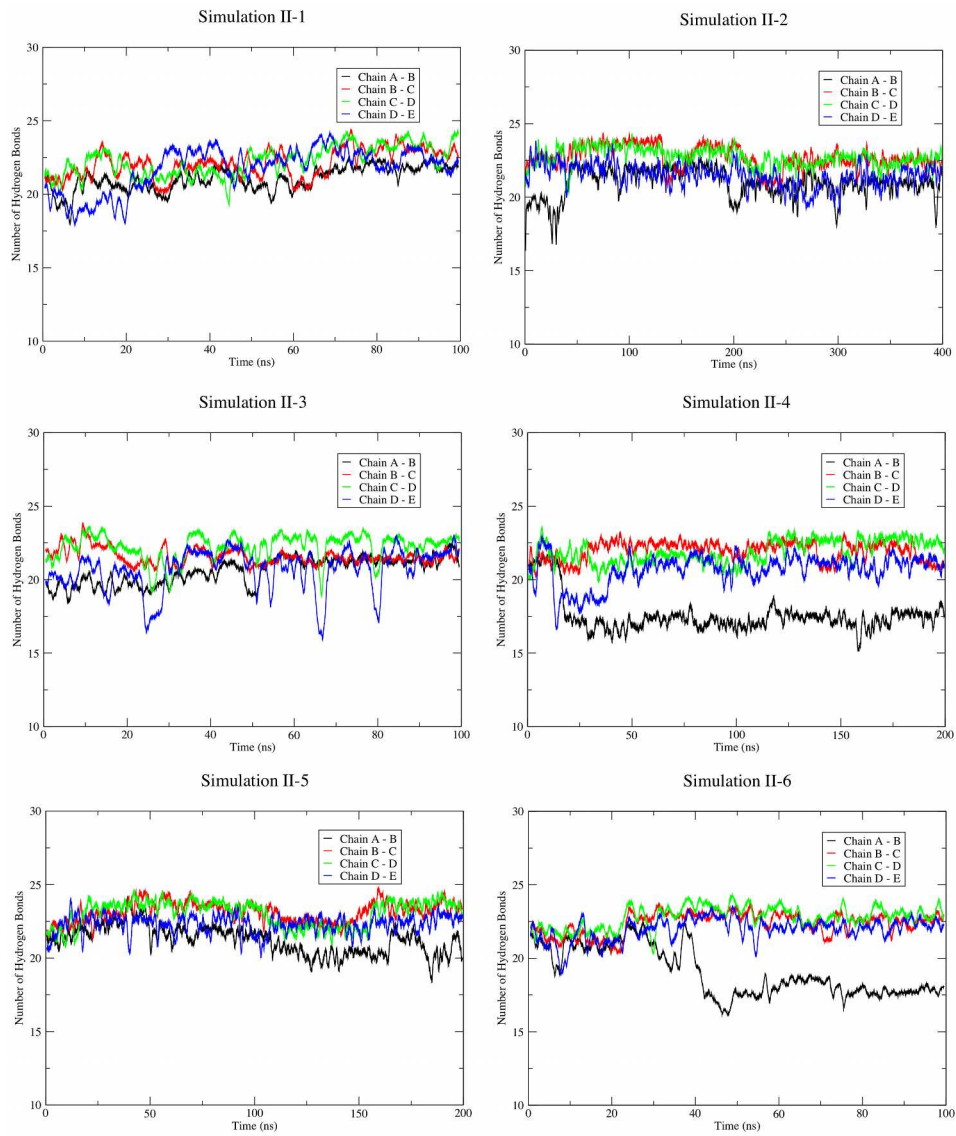


Figure 8.18: Backbone hydrogen bonds for Set II simulations.

Table 8.9: Interpeptide Asp23 - Lys28 $\text{COO}^- \rightarrow \text{NH}_3^+$ distances (nm, average over last half of trajectory $\pm \sigma$) from Set II protofibril simulations

System	Chain A - B	Chain B - C	Chain C - D	Chain D - E
1	0.48 ± 0.12	0.31 ± 0.02	0.37 ± 0.08	0.46 ± 0.10
2	0.42 ± 0.10	0.31 ± 0.03	0.31 ± 0.02	0.34 ± 0.04
3	0.35 ± 0.08	0.32 ± 0.02	0.31 ± 0.02	0.33 ± 0.04
4	0.81 ± 0.05	0.30 ± 0.02	0.31 ± 0.01	0.34 ± 0.03
5	0.36 ± 0.09	0.31 ± 0.03	0.31 ± 0.02	0.33 ± 0.04
6	0.78 ± 0.07	0.30 ± 0.02	0.31 ± 0.02	0.32 ± 0.04

Table 8.10: Backbone hydrogen bond count (average over last half of trajectory $\pm \sigma$) from Set II protofibril simulations

System	Chain A - B	Chain B - C	Chain C - D	Chain D - E
1	21.4 ± 1.5	22.5 ± 1.7	23.0 ± 1.5	22.5 ± 1.6
2	20.7 ± 1.5	22.3 ± 1.5	22.4 ± 1.4	21.1 ± 1.6
3	21.2 ± 1.5	21.4 ± 1.3	22.5 ± 1.6	20.8 ± 2.2
4	17.3 ± 1.3	21.9 ± 1.6	22.4 ± 1.5	21.0 ± 1.5
5	20.6 ± 1.6	23.0 ± 1.4	22.8 ± 1.5	22.3 ± 1.6
6	17.9 ± 1.2	22.5 ± 1.4	23.0 ± 1.3	22.1 ± 1.5

8.3.2.3 The Effects of Morin Penetration into $A\beta_{42}$ Protofibrils

The results of the simulations in Set II suggest that binding of morin to the surface of the $A\beta$ protofibril does little to destabilize its structure. We hypothesized that morin may exert its destabilizing effect by gaining entry into the hydrophobic core of the protofibril, a process that may evolve over long periods of time or occur with an energy penalty that cannot be overcome using conventional MD. Based on the results of simulation II-2, this hypothesis seems plausible, since one morin molecule partially inserted into the core of the protofibril, but perhaps the timeframe necessary to see full, spontaneous penetration exceeds that of even the sub-microsecond simulations conducted here. To accelerate the entrance of morin into the hydrophobic core of the protofibril, we applied center-of-mass (COM) pulling to a molecule of morin to place it into a small, solvent-accessible cavity present in the initial NMR structure (Figure 8.19A). This initial placement allowed for the assessment of the stability of the interaction between morin and the protofibril interior. In five of the six simulations in Set III, the initially-inserted morin molecule remained located within the core of the protofibril, and in some cases additional morin molecules penetrated in, as well (Figure 8.20) - one molecule in simulation III-4 and two molecules in simulations III-3 and III-5. This observation indicates that, while the initial penetration of the first morin molecule may occur over

Table 8.11: Backbone RMSD (nm, average over last half of trajectory $\pm \sigma$) from Set III protofibril simulations

System	Chain A	Chain B	Chain C	Chain D	Chain E
1	0.50 ± 0.05	0.36 ± 0.02	0.34 ± 0.02	0.35 ± 0.03	0.35 ± 0.03
2	N/A	N/A	N/A	N/A	N/A
3	0.50 ± 0.03	0.36 ± 0.02	0.30 ± 0.02	0.26 ± 0.02	0.29 ± 0.03
4	0.50 ± 0.03	0.33 ± 0.02	0.30 ± 0.02	0.26 ± 0.01	0.26 ± 0.01
5	0.52 ± 0.02	0.28 ± 0.02	0.24 ± 0.02	0.24 ± 0.02	0.29 ± 0.02
6	0.51 ± 0.03	0.30 ± 0.03	0.26 ± 0.02	0.25 ± 0.02	0.29 ± 0.02

Table 8.12: Backbone hydrogen bond count (average over last half of trajectory $\pm \sigma$) from Set III protofibril simulations

System	Chain A - B	Chain B - C	Chain C - D	Chain D - E
1	16.4 ± 2.5	22.5 ± 1.6	22.1 ± 1.6	18.6 ± 1.5
2	N/A	N/A	N/A	N/A
3	13.2 ± 1.5	21.2 ± 1.5	20.9 ± 1.5	18.9 ± 2.5
4	16.7 ± 1.6	22.0 ± 1.6	21.9 ± 1.6	22.0 ± 1.5
5	16.5 ± 1.6	21.3 ± 1.7	22.1 ± 1.5	21.1 ± 1.7
6	19.1 ± 1.7	21.0 ± 1.5	21.9 ± 1.7	22.4 ± 1.5

very long periods of time (as implied by the results of Set II), once morin has accessed the interior of the protofibril, other molecules may penetrate in, as well. In simulation III-2, the inserted morin molecule exited from the core of the protofibril within 10 ns. For this reason, we do not consider the data from this simulation, but we do note that while morin may be capable of entering into the $A\beta$ fibril, it may also exit with a finite probability.

Complete data sets for the simulations in Set III are available in Figures 8.21 - 8.27 and Tables 8.11 - 8.14. Over time, the position of the inserted morin molecule moved from its initial location in the center of the protofibril structure (Figure 8.19A,B) to interact principally with the Asp23-Lys28 interchain salt bridges of peptides A and B, within the solvent-accessible channel defined by Ile32 and Leu34 (Figure 8.19C,D). This location was preferred by morin in simulation II-2 in the case of the partially-inserted molecule (Figure 8.21).

A more detailed analysis reveals that hydrogen bonding interactions of morin, principally involving hydrogen bonding groups on the A and B rings of morin (Figure 8.1), with Asp23 and Lys28 lead to a disruption of the native Asp23-Lys28 salt bridge and backbone hydrogen bonding between peptides A and B (Figure 8.28A,B), promoting some disordering of the bend region and increased RMSF of residues 21-30 (Figure 8.23). This interaction is further stabilized by hydrophobic pack-

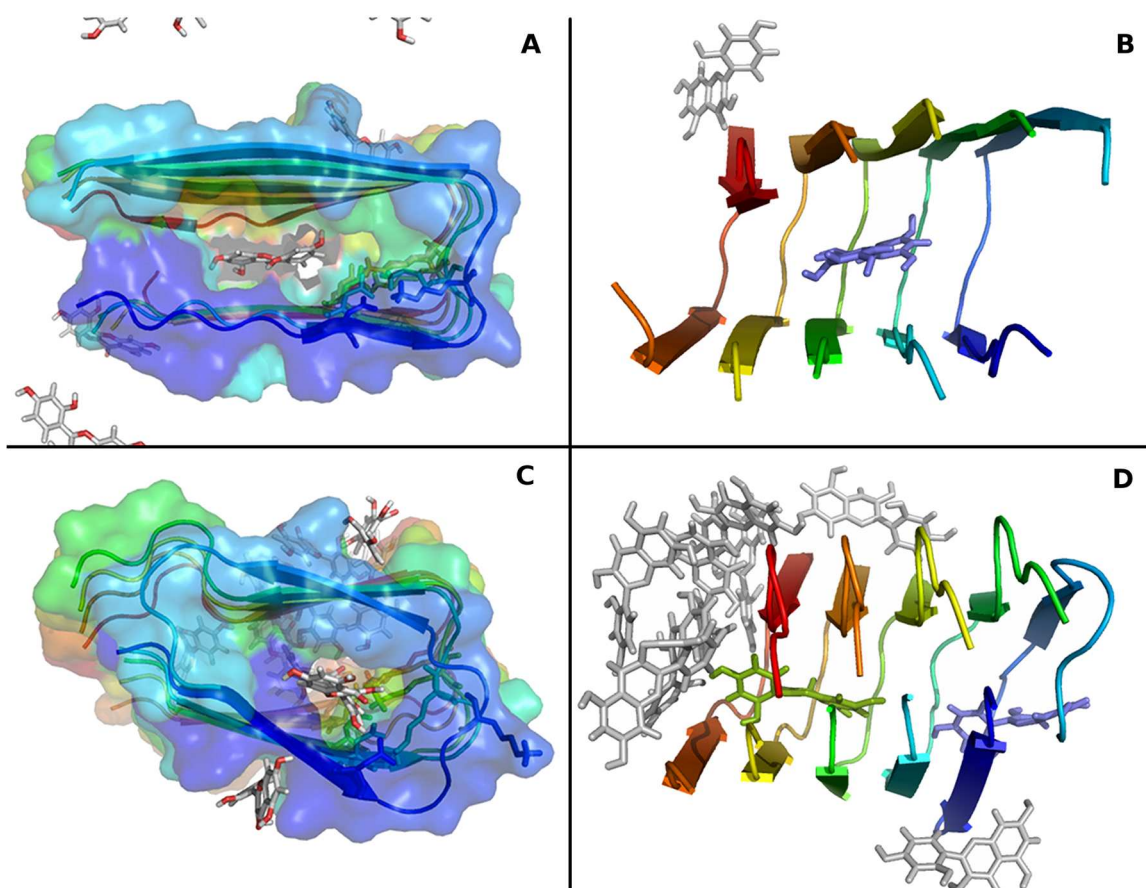


Figure 8.19: Initial placement of morin when inserted into the $A\beta$ fibril. (A) Surface rendering showing the internal cavity to which morin was targeted using COM pulling, and (B) an alternate view of the same starting structure with the initially-inserted morin molecule shown in light blue and another nearby morin molecule in light gray, (C) final position (400 ns) of the inserted morin molecule in simulation III-4, and (D) an alternate view of panel (C), with the initially-inserted morin molecule (light blue) and another morin molecule that entered the protofibril core (light green) shown. Other morin molecules are shown in light gray. In all panels, Asp23 and Lys28 are shown as sticks for reference. Note that, for clarity, not all ten morin molecules present in the unit cell are visible in all panels.

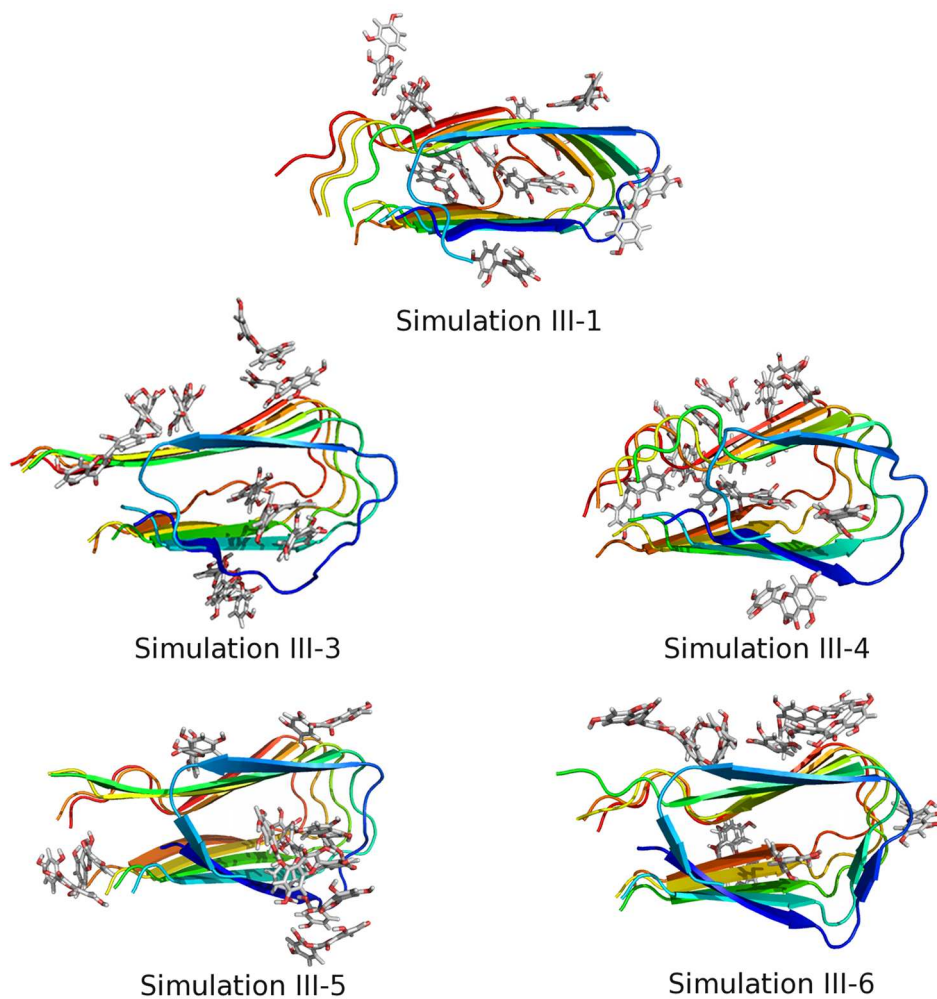


Figure 8.20: Snapshots from the end of Set III simulations.

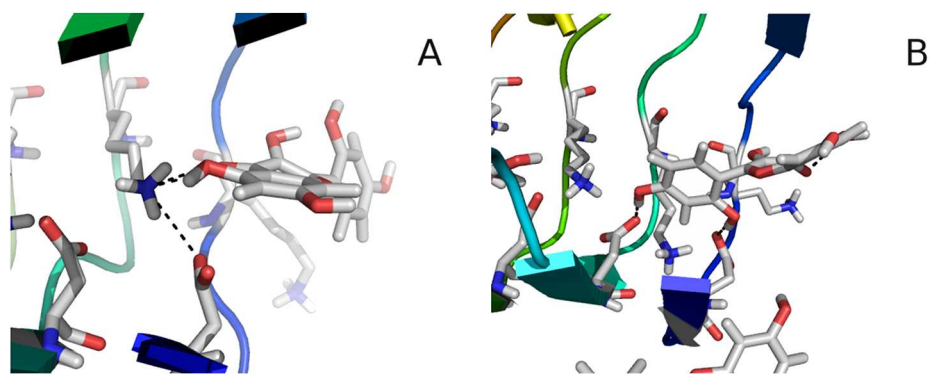


Figure 8.21: (A) Comparison between equilibrium positions of initially-diffuse morin from simulation II-2, and (B) initially-inserted morin from simulation III-4, representative of other simulations in this set.

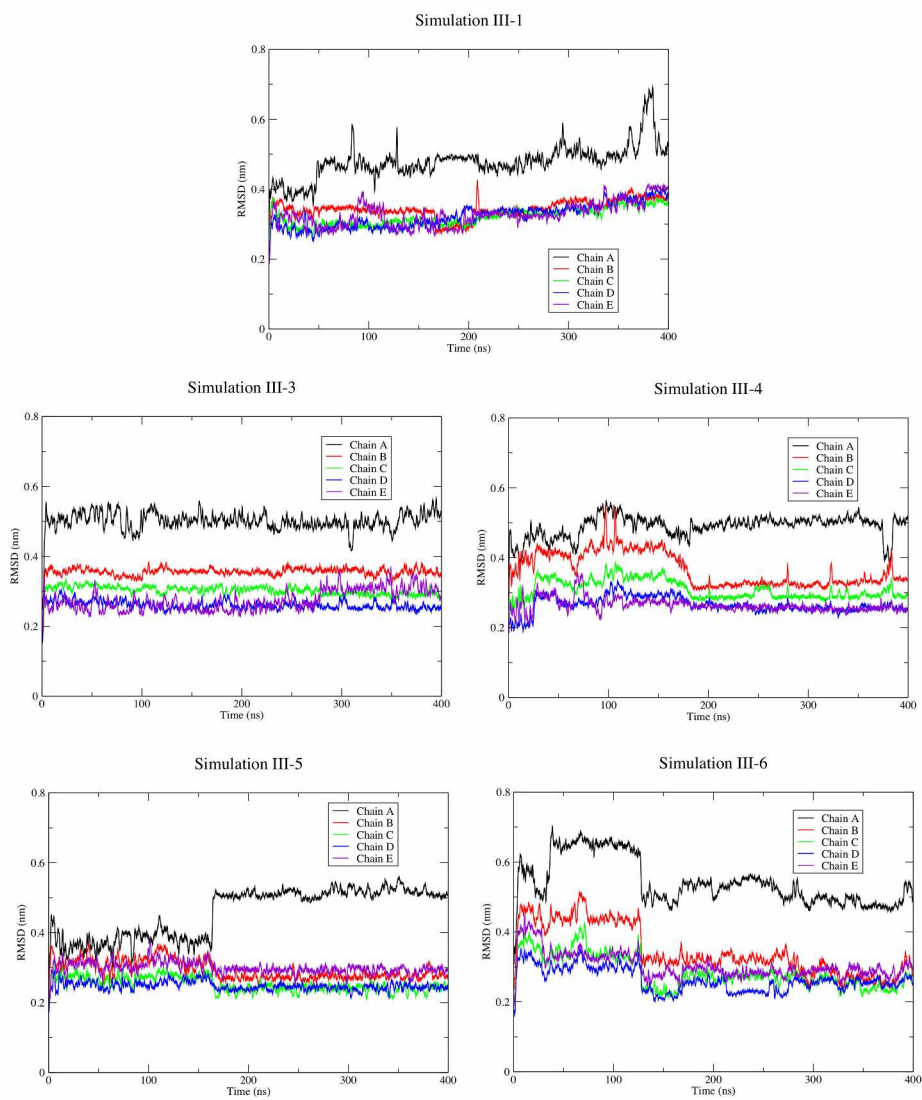


Figure 8.22: Backbone RMSD for Set III simulations.

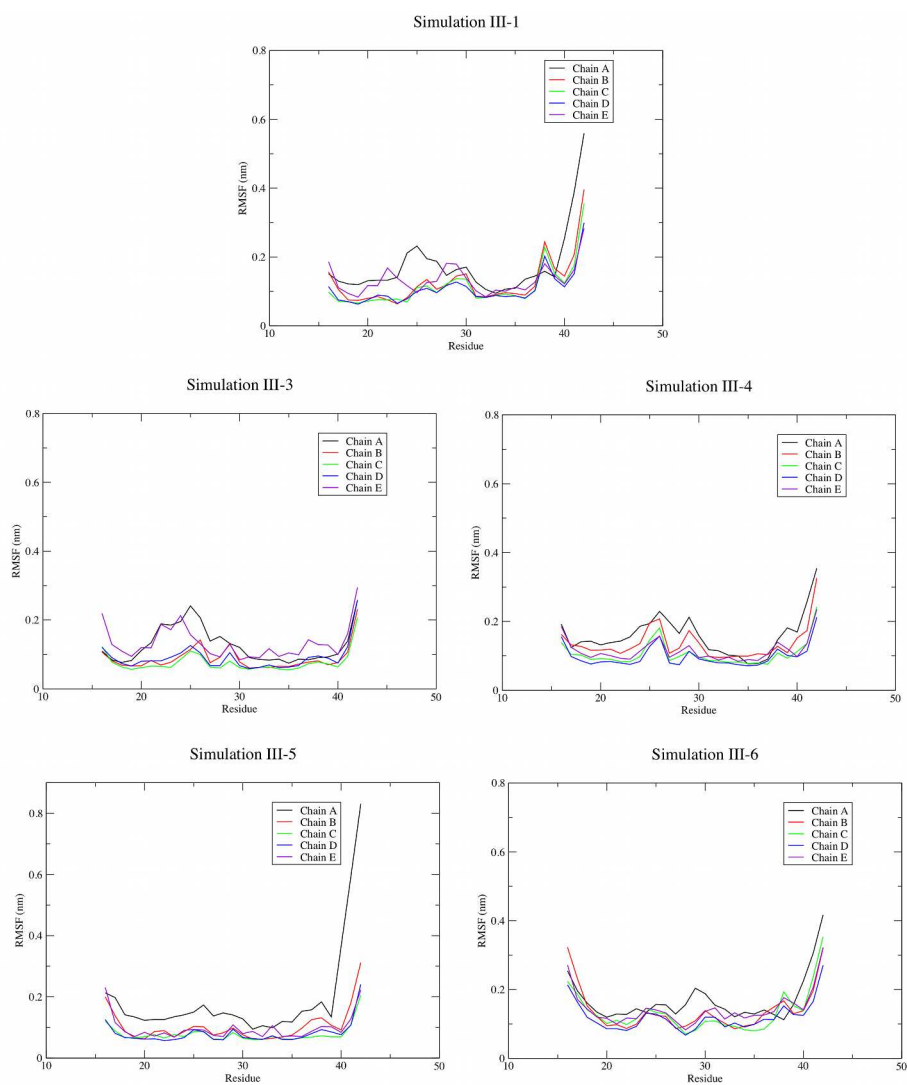


Figure 8.23: Backbone RMSF, averaged per residue for Set III simulations.

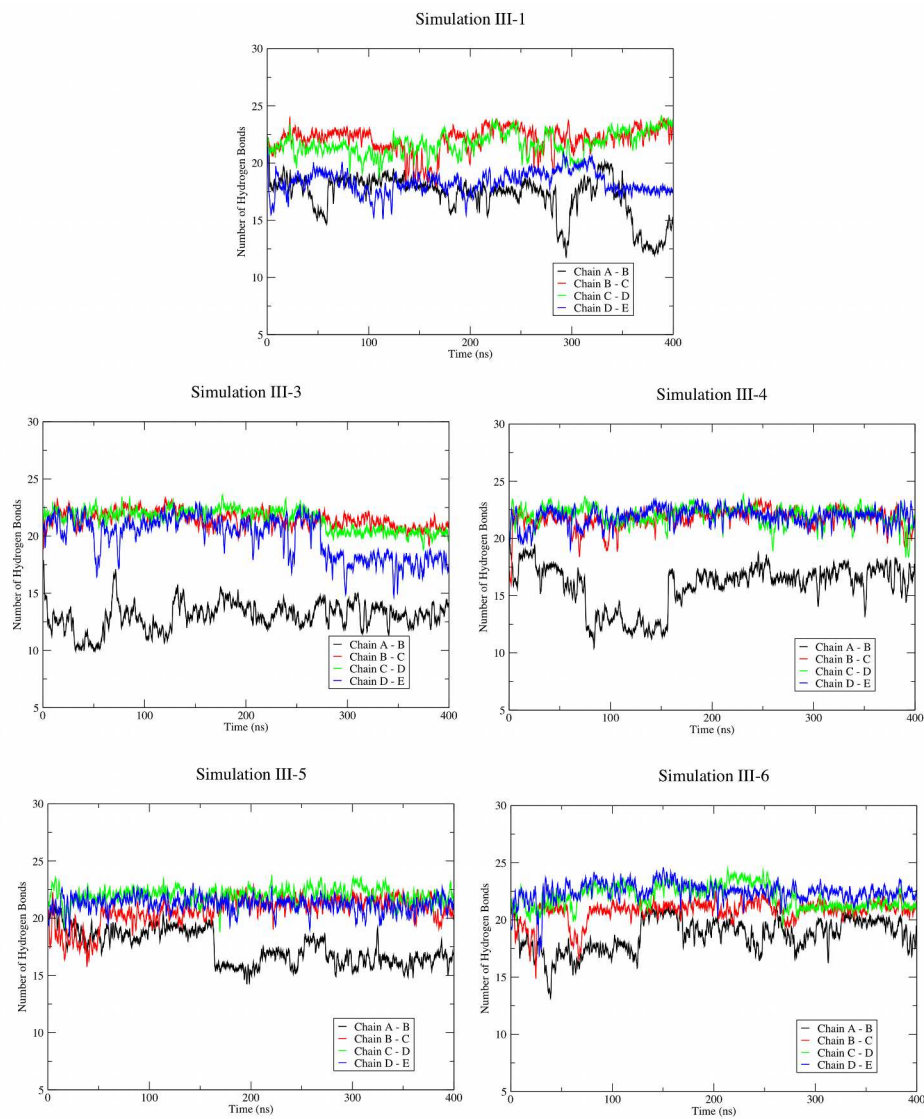


Figure 8.24: Backbone hydrogen bonds for Set III simulations.

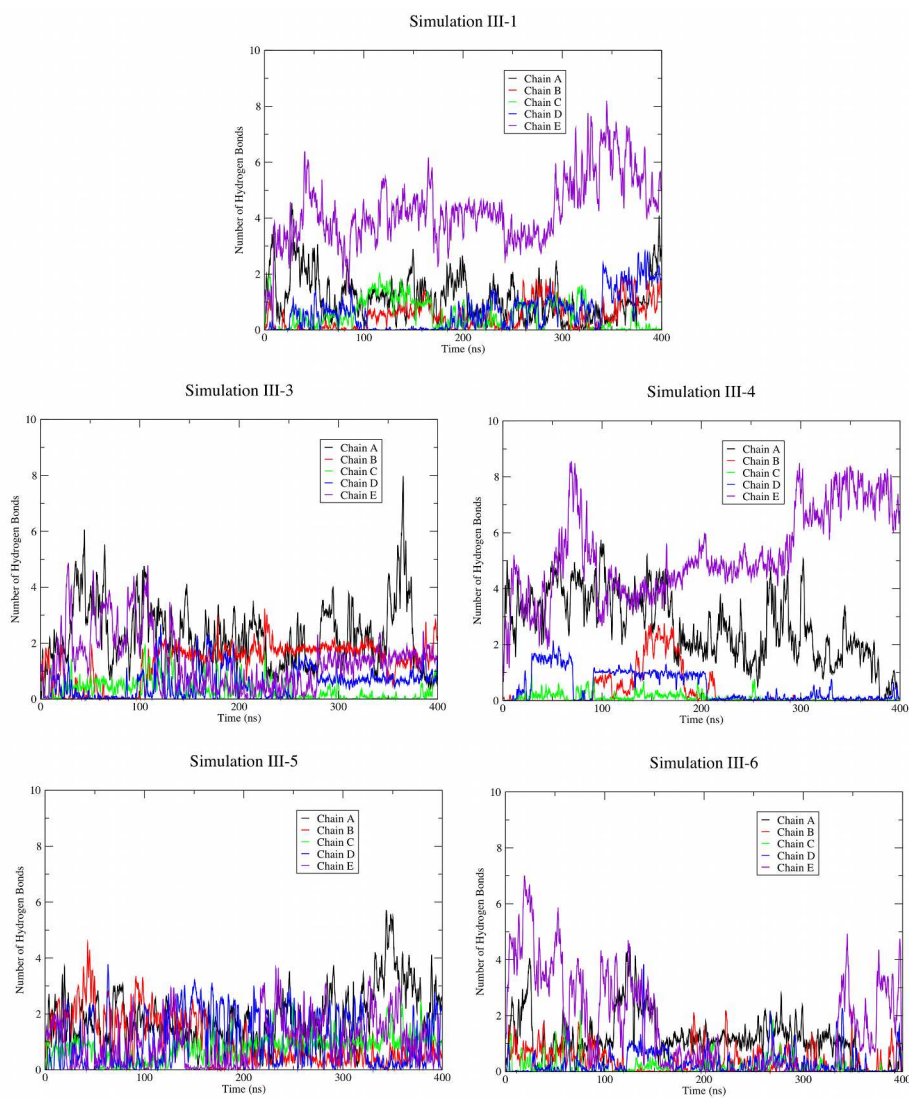


Figure 8.25: Hydrogen bonding between morin and the $A\beta$ protofibril backbone for all replicates in Set III.

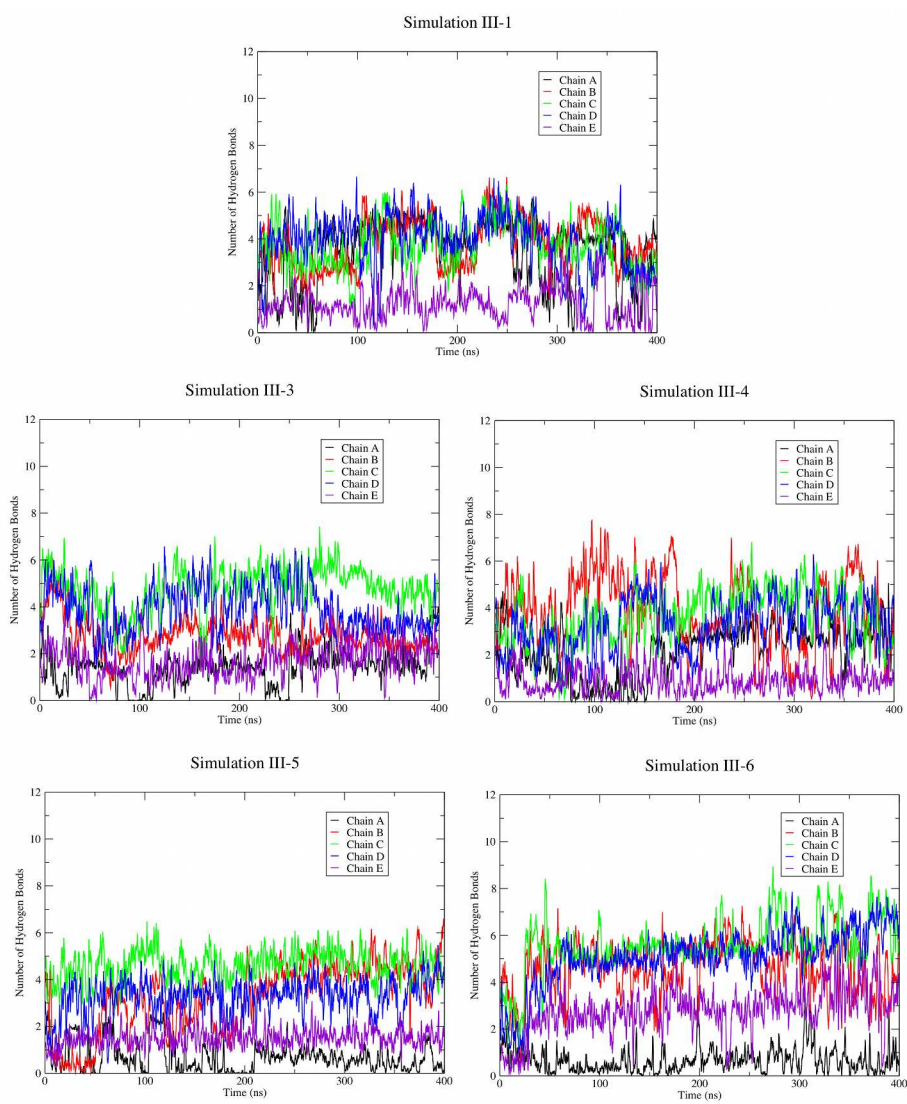


Figure 8.26: Hydrogen bonding between morin and the $A\beta$ protofibril sidechain groups for all replicates in Set III.

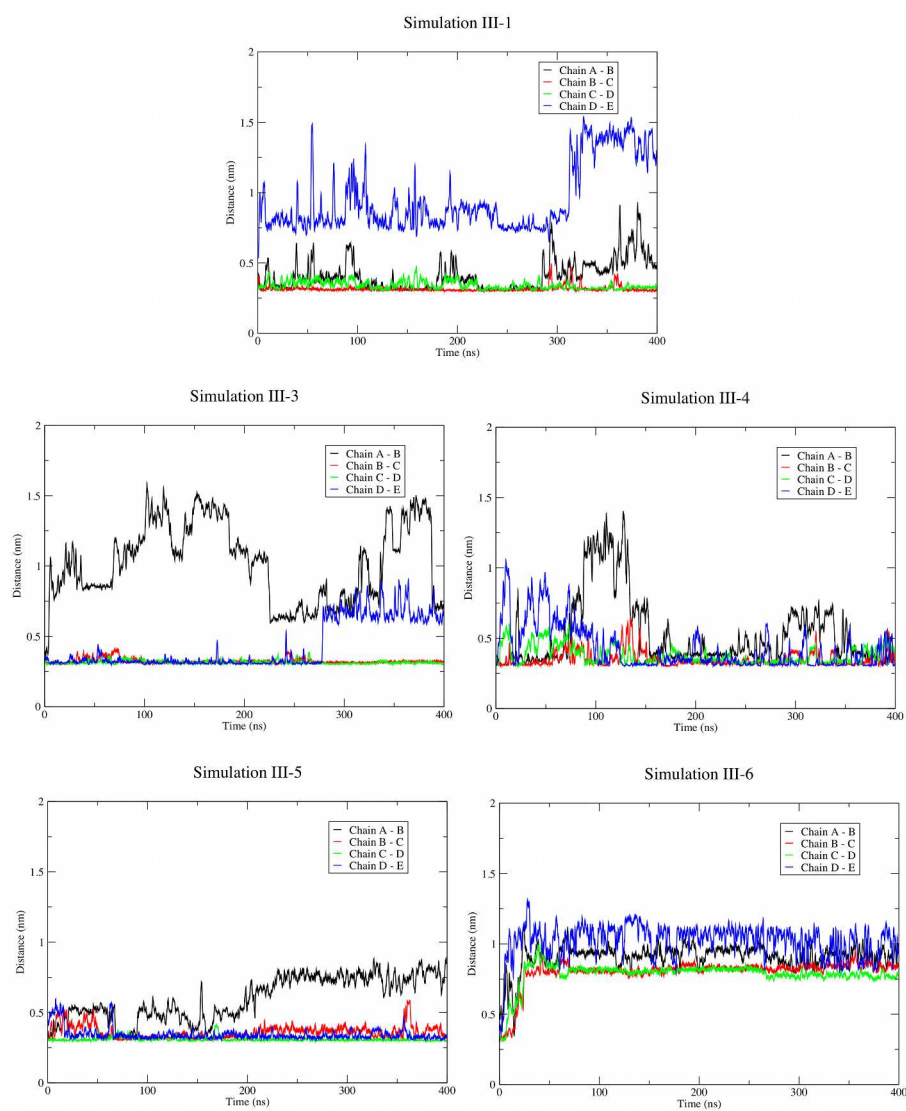


Figure 8.27: Asp23 - Lys28 ($\text{COO}^- \rightarrow \text{NH}_3^+$) distance for Set III simulations.

Table 8.13: Morin-backbone hydrogen bond count (average over last half of trajectory $\pm \sigma$) from Set III protofibril simulations

System	Chain A	Chain B	Chain C	Chain D	Chain E
1	0.9 ± 0.9	0.7 ± 0.7	0.4 ± 0.6	1.0 ± 0.8	4.7 ± 1.6
2	N/A	N/A	N/A	N/A	N/A
3	2.1 ± 1.5	1.7 ± 0.7	0.2 ± 0.4	0.7 ± 0.6	1.0 ± 0.9
4	1.9 ± 1.3	0.4 ± 0.2	0.0 ± 0.2	0.1 ± 0.3	6.2 ± 1.7
5	2.2 ± 1.4	0.5 ± 0.7	1.0 ± 0.8	1.0 ± 1.1	1.3 ± 1.1
6	0.9 ± 0.8	0.3 ± 0.6	0.2 ± 0.4	0.3 ± 0.6	0.8 ± 1.3

Table 8.14: Morin-sidechain hydrogen bond count (average over last half of trajectory $\pm \sigma$) from Set III protofibril simulations

System	Chain A	Chain B	Chain C	Chain D	Chain E
1	3.4 ± 1.9	4.1 ± 2.0	3.9 ± 2.1	3.7 ± 1.9	1.3 ± 1.5
2	N/A	N/A	N/A	N/A	N/A
3	1.6 ± 1.4	2.6 ± 1.5	5.0 ± 1.9	3.8 ± 1.8	1.9 ± 1.6
4	2.8 ± 1.4	3.3 ± 2.3	4.2 ± 2.1	3.6 ± 2.1	0.9 ± 1.2
5	0.6 ± 1.0	4.3 ± 1.9	4.8 ± 1.9	3.4 ± 1.9	1.5 ± 1.4
6	0.8 ± 1.1	4.8 ± 2.0	6.1 ± 2.0	5.8 ± 2.0	3.1 ± 1.9

ing with Ile32 and Leu34 (Figure 8.28C,D). A representative example of this destabilization is observed in simulation III-3. Within 25 ns, morin had moved from its initial placement and was interacting with the Asp23-Lys28 salt bridge. After approximately 80 ns, the distance between the charged moieties of these side chains was greater than 1.0 nm due to the disruption by morin (Figure 8.27); this distance persisted for much of the remaining simulation time, far larger than the Set I results that indicated the salt bridge distance should persist at approximately 0.35 ± 0.04 nm.

While causing the destabilization of the salt bridge, several other morin molecules associated with this disordered region of the $A\beta$ protofibril, and over time, these molecules interrupted the native backbone hydrogen bonds between peptides A and B (Figure 8.28), allowing an average of only 13.2 ± 1.5 hydrogen bonds to persist between these two peptides over the last 200 ns of simulation time, well below the control value of 21.1 ± 0.6 hydrogen bonds. Once these backbone-backbone hydrogen bonds were broken, they were solvated and did not tend to re-form. Few hydrogen bonds between morin and the peptide backbone persisted throughout the simulations in Set III (Table 8.13), indicating that morin was capable of competing for these native hydrogen bonds, but once exposed to solvent, the peptide backbone preferred to form hydrogen bonds with water. The result of these broken interactions was a dramatically increased backbone RMSD for peptide A of 0.50 ± 0.03 nm, double that of the controls in Set I.

It should be noted here that more hydrogen bonds formed between morin and the peptide backbone in the simulations in Set II than in Set III (Figures 8.12 and 8.25, Tables 8.6 and 8.13), principally due to the formation of the capping networks previously described, wherein morin occupied backbone amide groups at the ends of the protofibrils that are otherwise exposed to solvent. In the simulations in Set III, hydrogen bonds were principally formed between morin and backbone amides that form interpeptide hydrogen bonds, thereby disrupting these native interactions. These hydrogen bonds did not persist for long periods of time, however, and as the backbone was exposed to solvent, these hydrogen bonding groups were increasingly occupied by water molecules.

The observations enumerated above are largely shared among the other simulations that are part of Set III. In all of these cases, the backbone RMSD of peptide A (with which morin principally interacts) reached a stable value on the order of 0.50 nm (Figure 8.22), and the RMSF of the bend region increased substantially relative to the control results of Set I (Figure 8.23). In simulation III-6, a unique behavior was observed regarding the interpeptide salt bridges. The presence of morin in the core of the protofibril caused the Lys28 side chains to shift their orientation, such that they were oriented away from the “growing end” of the protofibril structure (Figure 8.27 and 8.29). Thus the native interdigitation of Lys28 with Asp23 of neighboring peptides was perturbed. The integrity of the interpeptide salt bridges and side chain packing around them has been proposed to be an important contribution in the stability of the $A\beta_{42}$ fibril [346], and thus it can be seen that the presence of morin in the protofibril interior, by interfering with native packing and by forming hydrogen bonds within the salt bridges, can interrupt these stabilizing interactions.

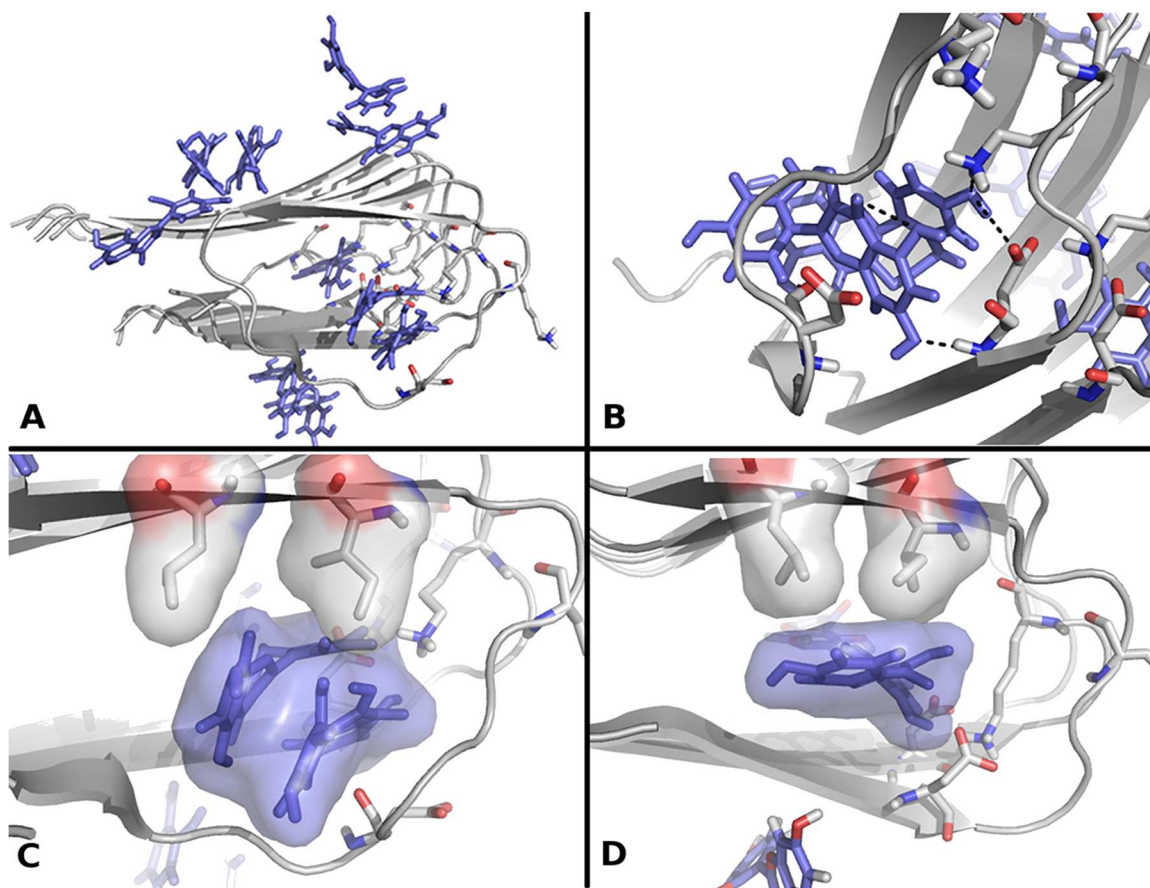


Figure 8.28: Destabilization of the $A\beta$ protofibril by morin in the protofibril interior. (A) Image from a snapshot at 400 ns of simulation III-3, showing the persistence of three morin molecules (light blue) within the protofibril core, causing a destabilization of peptide A. (B) Disruption of native backbone hydrogen bonds by morin from the same snapshot as in (A), with putative hydrogen bonds indicated as black dashes. Asp23 and Lys28 are shown as sticks and colored by element. Interactions between morin, Ile32 and Leu34 are shown in panels (C) and (D) for simulations III-3 and III-4, respectively, with surface representations showing hydrophobic packing. Protein residues are shown as sticks and colored by element; morin molecules are shown in sticks and colored light blue.

8.3.2.4 Morin Inhibits the Assembly of $A\beta_{42}$ Protofibrils

The simulations conducted as part of Set II indicate that the principal mode of interaction between morin and the $A\beta$ protofibril was capping, although partial penetration into the hydrophobic core was observed (simulation II-2). We hypothesized that the capping of the protofibril by morin could preclude the association of an incoming $A\beta$ peptide by blocking potential backbone hydrogen bonds and other side chain packing interactions that would lead to the growth of the protofibril. To test this theory, we employed COM pulling to promote the association of a free peptide with the

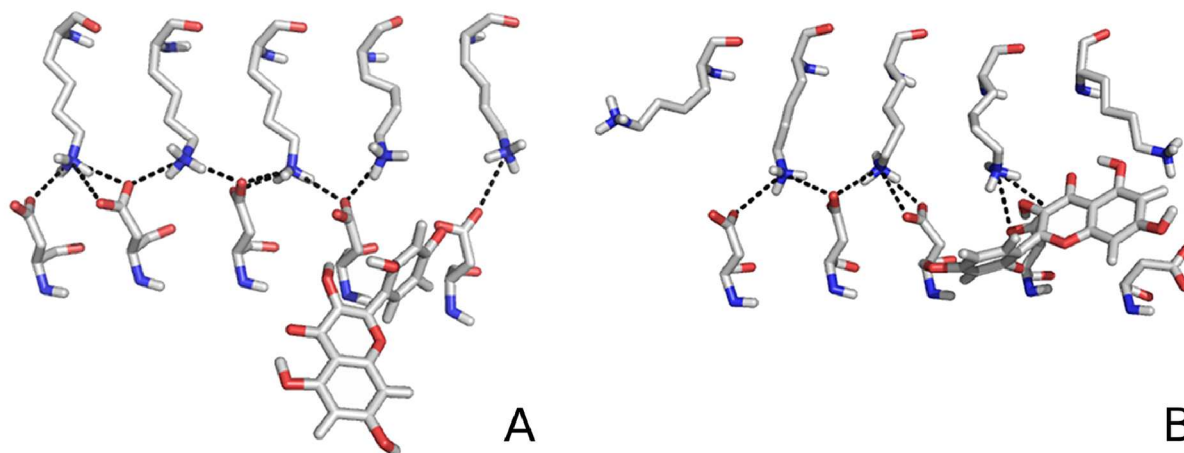


Figure 8.29: Disruption of the native interpeptide Asp23-Lys28 salt bridges in simulation III-6. (A) Initial structure after equilibration and (B) structure at 400 ns, showing the shift in position of the Lys28 residues. Black dashed lines show hydrogen bonding and electrostatic interactions, and the arrow points to the “growing end” of the fibril.

protofibril, either in the absence (simulations IV-1, IV-2, and IV-3) or presence (IV-4, IV-5, and IV-6) of morin.

The free peptides in both simulations were pulled toward the protofibril along the z -axis (coincident with the protofibril axis) until a COM separation between peptide A and the free peptide of approximately 1.4 nm was achieved, after which 100 ns of unrestrained MD was performed. In these simulations, the free peptide in simulations IV-1, IV-2, and IV-3 established close contact with the protofibril (Figures 8.30A and 8.31), while morin effectively blocked this association in simulations IV-4, IV-5, and IV-6 (Figures 8.30B and 8.31). In the case of simulation IV-6, the free peptide was directed away from the end of the protofibril and towards the N-terminal face of the structure as a result of its interactions with morin (Figure 8.31).

Over the course of 100 ns, the untreated systems established an average of 6.7 ± 1.7 hydrogen bonds per timeframe between the free peptide and chain A of the protofibril (Figure 8.32 and Table 8.15), while in simulations IV-4, IV-5, and IV-6, these peptides only formed 2.2 ± 1.9 hydrogen bonds during the same period (Figure 8.33 and Table 8.16). Morin was able to sequester the free peptide away from the protofibril, with a COM distance between the free peptide and chain A stabilizing at an average distance of 1.08 ± 0.52 nm in the untreated systems, while in the morin-treated systems, this distance averaged 1.24 ± 0.10 nm (Figures 8.34 and 8.35, Tables 8.17 and 8.18). The reduced contact between the free peptide and the protofibril was also manifested in the substantially reduced number of atomic contacts between the incoming peptide and chain A. In the untreated systems, 1274 ± 327 contacts were formed per timeframe, averaged over all systems (Figure 8.36 and Table 8.19), but in the morin-treated systems, only 579 ± 304 contacts were formed (Figure 8.37 and Table 8.20). A contact was defined as occurring between any atomic

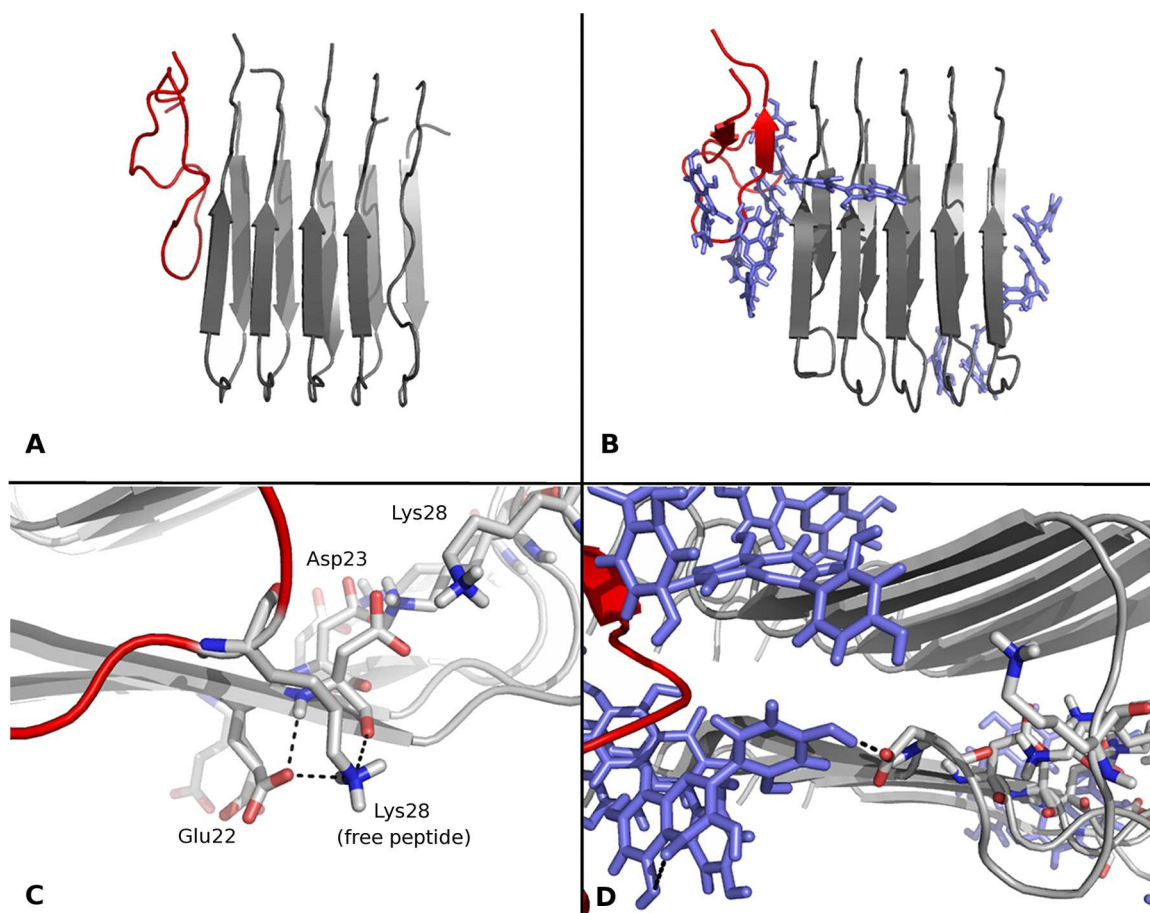


Figure 8.30: Inhibition of $A\beta$ protofibril assembly by morin. (A) Final structure of simulation IV-1 (representative of the untreated systems in this set), (B) final structure of simulation IV-4 (representative of the morin-treated systems), (C) details of the interactions that form in simulation IV-1, and (D) association of morin with Asp23 to block its interaction with the incoming peptide in simulation IV-4. In all panels, the protofibril is colored gray and the free peptide is colored red. In panels (C) and (D), putative hydrogen bonds are indicated by black dashed lines. Glu22, Asp23, and Lys28 are shown as sticks and colored by element in panels (C) and (D). Similarly, morin is depicted in panels (B) and (D) as sticks and colored light blue.

pair within a distance of 0.6 nm, a criterion we previously used when assessing the stability of $A\beta$ peptide-peptide interactions [346].

In simulations IV-1 and IV-2, Lys28 of the incoming peptide formed ionic interactions and hydrogen bonds with the side chain of Glu22 and backbone of Asp23 of chain A in the protofibril (Figure 8.30C). Additionally, in simulation IV-3, a salt bridge between Asp23 of the free peptide and Lys28 of chain A briefly formed. These interactions were blocked by a network of hydrogen-bonded morin molecules in simulations IV-4, IV-5, and IV-6, precluding a strong association between the incoming peptide and the fibril (Figure 8.30D).

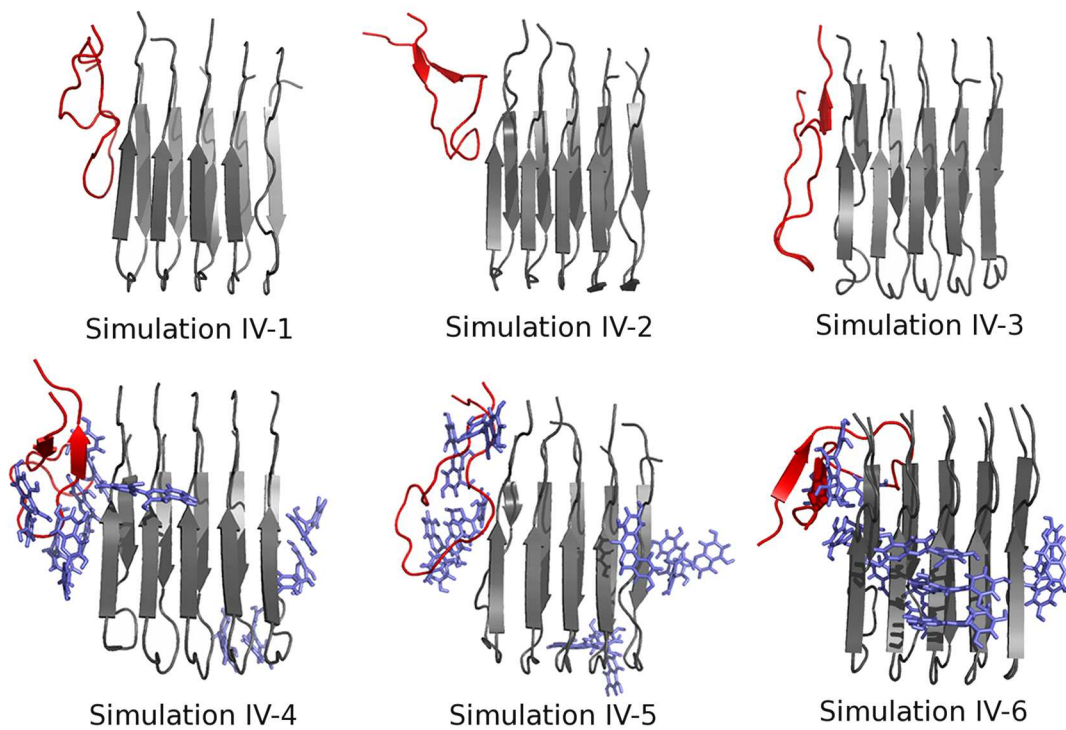


Figure 8.31: Snapshots from the end of Set IV simulations. Images are rendered as in Figure 8.30.

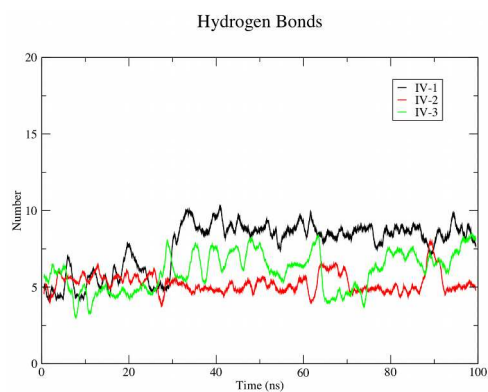


Figure 8.32: Formation of hydrogen bonds between the free $A\beta$ peptide and the protofibril in simulations IV-1, IV-2, and IV-3.

Table 8.15: Hydrogen bond count between the incoming peptide and the $A\beta$ protofibril in control Set IV simulations (average over last half of trajectory $\pm \sigma$)

System	Count
IV-1	8.6 ± 1.6
IV-2	5.3 ± 1.3
IV-3	6.3 ± 1.6

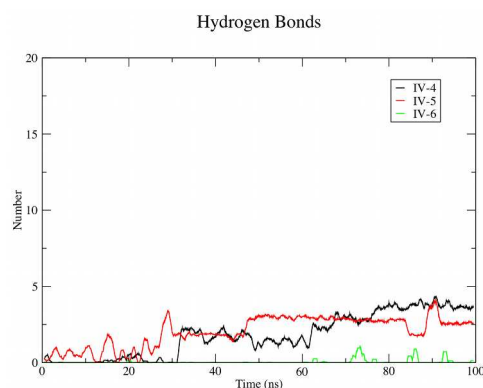


Figure 8.33: Formation of hydrogen bonds between the free $A\beta$ peptide and the morin-bound protofibril in simulations IV-4, IV-5, and IV-6.

Table 8.16: Hydrogen bond count between the incoming peptide and the $A\beta$ protofibril in morin-treated Set IV simulations (average over last half of trajectory $\pm \sigma$)

System	Count
IV-4	$3.7 \pm 1.7^*$
IV-5	2.8 ± 0.7
IV-6	0.1 ± 0.4

*Average over last 20 ns.

Table 8.17: Center-of-mass distance between the incoming peptide and the $A\beta$ protofibril in control Set IV simulations (average over last half of trajectory $\pm \sigma$)

System	Count
IV-1	0.95 ± 0.03
IV-2	1.65 ± 0.09
IV-3	0.64 ± 0.09

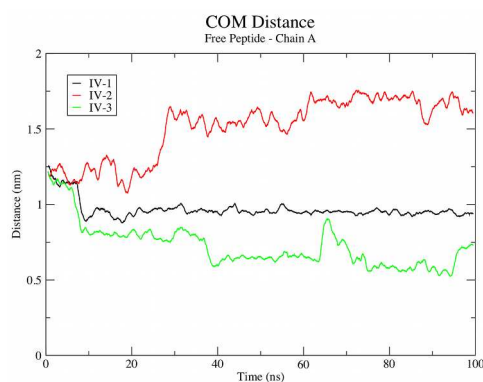


Figure 8.34: Center-of-mass distance between the free $A\beta$ peptide and the morin-bound protofibril in simulations IV-1, IV-2, and IV-3.

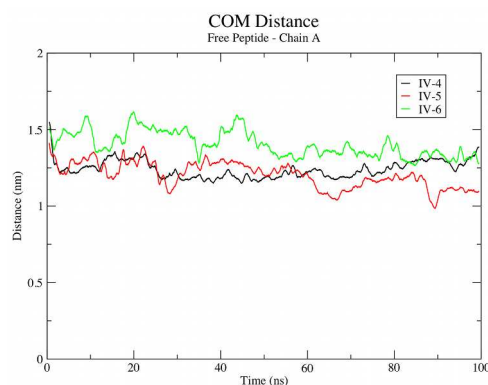


Figure 8.35: Center-of-mass distance between the free $A\beta$ peptide and the morin-bound protofibril in simulations IV-4, IV-5, and IV-6.

Table 8.18: Center-of-mass distance between the incoming peptide and the $A\beta$ protofibril in morin-treated Set IV simulations (average over last half of trajectory $\pm \sigma$)

System	Count
IV-4	1.24 ± 0.06
IV-5	1.15 ± 0.07
IV-6	1.34 ± 0.06

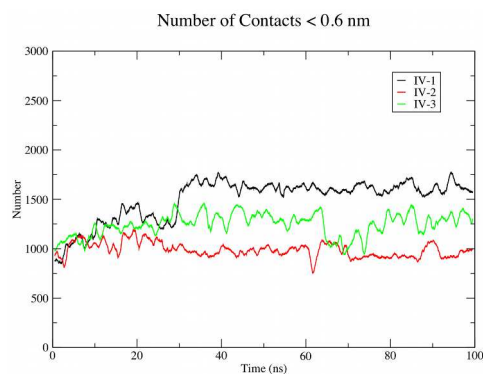


Figure 8.36: Atomic contacts between the free $A\beta$ peptide and the control protofibril in simulations IV-1, IV-2, and IV-3.

Table 8.19: Atomic contacts between the incoming peptide and the $A\beta$ protofibril in control Set IV simulations (average over last half of trajectory $\pm \sigma$)

System	Count
IV-1	1613 \pm 94
IV-2	960 \pm 94
IV-3	1248 \pm 149

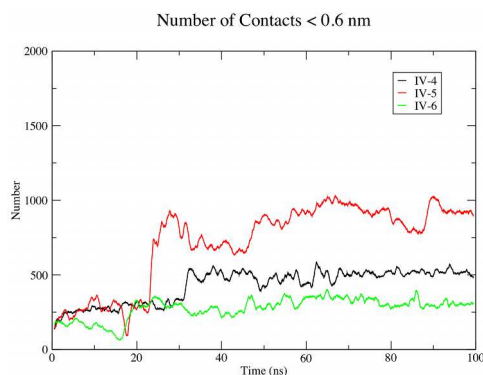


Figure 8.37: Atomic contacts between the free $A\beta$ peptide and the morin-bound protofibril in simulations IV-4, IV-5, and IV-6.

Table 8.20: Atomic contacts between the incoming peptide and the $A\beta$ protofibril in morin-treated Set IV simulations (average over last half of trajectory $\pm \sigma$)

System	Count
IV-4	501 \pm 59
IV-5	914 \pm 83
IV-6	321 \pm 58

Although complete attachment and in-register alignment of the incoming $A\beta$ peptide onto the protofibril was not observed in simulations IV-1, IV-2, or IV-3, it is not expected that this phenomenon would be seen within a timeframe accessible to MD simulations, although many native-like contacts formed over the course of simulation IV-3 (Figure 8.31). Fibril elongation is observed *in vitro* over the course of several hours, well beyond what can be simulated. Nevertheless, morin induces a striking reduction in the number of contacts and hydrogen bonds between the incoming $A\beta$ peptide and the protofibril, suggesting that indeed these “capping” networks are likely capable of inhibiting fibril elongation by blocking the attachment of an incoming peptide onto the growing end of an $A\beta_{42}$ fibril.

8.4 Conclusions

A number of *in vitro* [159, 160, 348] and *in vivo* [162, 349] studies have suggested that polyphenols (flavonoids) from red wine and other food products are effective in disrupting aggregates of $A\beta$, including fully-formed fibrils, rendering them less cytotoxic. The *in vivo* results further suggest that such intervention may improve cognitive ability in patients suffering from Alzheimer’s disease. Missing in the literature thus far is the detailed mechanism of how these flavonoids exert their beneficial effect, knowledge that is essential to the development of therapeutic compounds. Over the course of twenty-five explicit-solvent, atomistic MD simulations totaling 4.3 μ s of simulation time, we have probed the interactions of morin with a model of the $A\beta_{42}$ fibril.

MD simulations have been conducted on a number of models of the $A\beta$ fibril under various conditions [202, 303, 304, 307, 308, 316]; also among these are investigations on the interaction between $A\beta$ fibrils and aromatic compounds [350] and ibuprofen [347]. Clearly, there is great interest in understanding how small molecules might interact with, and ultimately destabilize, $A\beta$ assemblies. Polyphenolic compounds such as morin are attractive therapeutic candidates, as they are found in natural food products, are capable of crossing the blood-brain barrier [162, 349], and are non-toxic in clinically-relevant doses [163]. As such, interactions between flavonoids and α -synuclein [351] and the human islet amyloid polypeptide (hIAPP) [352] have already been studied for the purposes of treating a variety of conditions, but as of yet, flavonoid- $A\beta$ interactions remain largely uncharacterized.

In the context of inhibiting fibrillation, our results agree with a previous study conducted by Porat *et al.*, who used phenol red to inhibit hIAPP fibril growth [352]. In that work, aromatic interactions were identified as the basis for the inhibition of fibril growth, a conclusion that extends to the present study. We find that, when initially placed in solution around a preformed $A\beta_{42}$ protofibril, morin had a tendency to quickly deposit on the surface of the protofibril and form networks that were stabilized by hydrophobic interactions and hydrogen bonding. In several of the simulations in Set II, these networks developed at the ends of the $A\beta$ protofibril structure that served here as a model of the full-length $A\beta_{42}$ fibril. These “capping” complexes most frequently formed in the vicinity of the Phe19-Phe20 dyad, anchored by aromatic stacking interactions and hydrogen bonds to the peptide backbone and Asp23. By occupying the backbone hydrogen bonding groups and side chains that would otherwise be accessible to an incoming peptide, morin precludes this interaction, inhibiting fibril elongation, as we have shown in simulations IV-4, IV-5, and IV-6.

We also find that there is a finite probability of morin entering into the hydrophobic core of the $A\beta$ protofibril. Although this process was not completely observed in any of the Set II simulations, a single molecule of morin interacted strongly with the Asp23 and Lys28 residues of peptides A and B, respectively, in simulation II-2. These interactions allowed this morin molecule to stably bind at this location and partially insert into the core of the protofibril. The possibility of complete insertion, perhaps inaccessible on the computational time scale, led to the simulations conducted as members of Set III, wherein morin was placed within the hydrophobic core of the protofibril at the outset of the simulations. Over these trajectories, morin tended to orient itself such that it interacted with the Asp23-Lys28 salt bridges, adopting a position similar to that of simulation II-2, but more deeply embedded in the protofibril. The relative agreement of the morin position in simulations II-2 and those of Set III (Figure 8.21) provides strong evidence that morin may interact with the Asp23-Lys28 salt bridge, a region of the $A\beta$ structure that is accessible from the surrounding solvent, to gain entry into the hydrophobic core and lead to destabilization. Such positioning near the Asp23-Lys28 salt bridges disrupted these ionic interactions and allowed morin to compete for backbone hydrogen bonds between peptides. Upon destabilizing these native hydrogen bonds, the peptide backbone became solvated and increasingly accessible to other morin molecules in the unit cell, a phenomenon that disfavored the re-formation of these native interactions.

The tendency for morin to associate with $A\beta$ in this bend region is not entirely surprising. The bend region presents somewhat of a unique environment, a strong ionic interaction (the Asp23-Lys28 salt bridge) packed against two bulky hydrophobic residues (Ile32 and Leu34) that define a narrow channel of water that promotes the stability of the protofibril structure as a whole, as we demonstrated previously [346], in agreement with existing reports of model $A\beta$ fibrils [202, 304, 316]. Hydration of internal cavities has also been demonstrated experimentally in the case of polyglutamine [353] and folding intermediates of β 2-microglobulin fibrils [354]. By virtue of its numerous hydrogen bonding groups, morin established favorable interactions with the salt bridge in this region of $A\beta$, while simultaneously packing against the hydrophobic moieties in the β 2 region.

The effects of flavonoids on fibril formation in a number of other systems have been studied recently [351, 352, 355], highlighting the role of aromatic interactions between fibrillation inhibitors

and phenylalanine or tyrosine residues in the amyloidogenic protein. Our results are in agreement with the results of these other systems. When morin “capping” complexes form around the terminal peptides in the protofibril, some aromatic stacking interactions contribute to the stability of these capping networks, but there is little structural perturbation in these cases, as the RMSF of these residues remains comparable to control levels. Thus we do not find any interaction between morin and the phenylalanine dyad in the $L_{17}VFFAE_{22}$ sequence that leads to structural destabilization. These results contrast somewhat with the results of Convertino *et al.* [350], who modeled a trimeric assembly of the $H_{14}QKLVFF_{20}$ $A\beta$ fragment in the presence of anthracene and 9,10-anthraquinone. That study found that aromatic interactions involving the Phe19-Phe20 dyad attracted the test compounds to the peptide fragment and promoted backbone-aromatic interactions that destabilized the native hydrogen bonding structure. We observed no such behavior in our simulations, and thus in the context of morin, we do not believe that the $H_{14}QKLVFF_{20}$ sequence is the critical target for destabilizing $A\beta_{42}$ fibrils, although only a small part of this sequence was modeled in the present study. It may be of interest to consider even larger models of mature, full-length $A\beta$ fibrils, since several smaller models of this assembly have now been studied. The results of the present work indicate that morin may destabilize $A\beta$ fibrils by disrupting interactions at the ends of the assembly, an observation that would extend to larger systems, but it also remains of interest to examine the potential for morin to interrupt the interactions between fibrils within amyloid plaques.

The results of the *in vitro* work by Ono *et al.* suggest that flavonoids (including morin) may reduce the rate of fibril assembly by either promoting depolymerization or by inhibiting polymerization [160]. Our results indicate that perhaps both phenomena are possible. The results of simulations IV-4, IV-5, and IV-6 suggest that morin can preclude the attachment of an incoming peptide, thus reducing the rate of polymerization. The results of Sets II and III together also suggest that morin can destabilize existing protofibril assemblies, perhaps promoting depolymerization, although complete dissociation of any peptide from the protofibril was not observed within the timeframes explored in our simulations.

Previous studies on small molecule inhibitors of amyloid fibril formation have concluded that some compounds are capable of stabilizing non-fibrillating conformations of the amyloidogenic proteins [351, 356, 357]. In one particularly relevant example, Meng *et al.* demonstrated that flavonoids could bind and stabilize monomeric α -synuclein, preventing its incorporation into fibrils [351]. Our results do not contradict this possibility in the context of morin interactions with $A\beta$, as such information was not directly examined here. Using MD methodology on other $A\beta$ systems (monomers, dimers, small oligomers, *etc.*) in the presence of flavonoids would further increase our understanding of the molecular events of flavonoid-induced aggregation inhibition, including the possibility that flavonoids may stabilize non-aggregating or non-fibrillating conformations of $A\beta$. As such, this concept will be explored in Chapter 9.

The present work has explored the effects of morin deposition on the surface of the protofibril, the potential for morin to penetrate into the hydrophobic core and the resulting destabilization that occurs once morin enters the protofibril, and the ability of morin to prevent the attachment of an incoming $A\beta$ peptide onto the preformed protofibril. The features of morin that allow it to act in

this multifaceted capacity are its aromaticity and hydrogen bonding capacity. The hydrophobic aromatic rings drive the initial hydrophobic contacts with the $A\beta$ protofibril (Figure 8.11), promote insertion into the hydrophobic core, pack against Ile32 and Leu34 (Figure 8.28C,D), and aid in anchoring the “capping” networks to the protofibril (Figure 8.10A and 8.14). The numerous hydrogen bonding groups serve to further stabilize the “capping” networks on the ends of the protofibril, disrupt the Asp23-Lys28 salt bridges (Figures 8.28B and 8.29), and perturb native backbone hydrogen bonding after gaining access to the interior of the protofibril.

The principal mode of protofibril destabilization by morin is exerted through interactions with the Asp23-Lys28 salt bridges, which we previously proposed to be key to the stability of the mature $A\beta_{42}$ fibril [346]. The contacts formed between morin and these residues principally involve the hydrogen bonding groups (hydroxyl groups and ketone carbonyl) on the fused A and B rings of morin (Figure 8.1). These features are shared among many flavonoids and may explain why members of this class of compounds have similar efficacy in inhibiting fibril extension and destabilizing preformed fibrils despite subtle differences in the number and position of hydroxyl groups on the C ring [160]. These observations may further explain why compounds such as catechin and epicatechin have reduced efficacy in the same assays [160]; they are less planar (by virtue of chiral centers in their structures) and lack a strong hydrogen bond acceptor (the ketone carbonyl), features present in compounds like morin. Taken together, this information provides molecular insight into chemical features of small molecules that will aid in furthering drug design for the treatment of Alzheimer’s disease.

Chapter 9

Inhibiting Early A β Aggregation Events with Morin

Copyright statement: The contents of this chapter have been submitted for publication to the journal *Biochemistry* but are unpublished at this time. J. A. Lemkul and D. R. Bevan. “Morin Inhibits the Early Stages of Amyloid β -Peptide Aggregation by Altering Tertiary and Quaternary Interactions to Produce ‘Off-Pathway’ Structures.”

Attribution: J. A. L. and D. R. B. designed the research. J. A. L. conducted the simulations and performed analysis. J. A. L. and D. R. B. wrote the paper.

9.1 Introduction

Although numerous experimental studies have concluded that flavonoids can inhibit A β aggregation and destabilize A β fibrils [159, 160], the mechanism by which these molecules exert their effects remains somewhat unclear. In order to develop viable therapeutics for the treatment of Alzheimer’s disease, molecular insight into the functionally relevant features of the putative therapeutic compound and the specific target location(s) on A β is necessary. In our earlier work (Chapter 8) [358], we proposed a mechanism by which morin destabilizes A β protofibrils, and likely full-length fibrils, but the mechanism by which morin inhibits the formation of smaller, more toxic, oligomers is still unexplored. Such information is crucial to the development of effective anti-A β therapeutics.

Since experimental high-resolution structural characterization of the intermediates in the A β aggregation pathway remains challenging, computational methods play an important role in examining the specific interactions in the formation of toxic oligomers and higher-order A β assemblies. In particular, molecular dynamics (MD) simulations can provide atomic-level detail into A β unfolding and aggregation, as well as compounds that may inhibit this process. The utility of MD

simulations to investigate these types of systems has been demonstrated previously in a number of studies [201, 203, 204, 318, 359, 360]. We employ MD simulations here to explore the binding of morin to $A\beta$ monomers and dimers, with the goal of elucidating the mechanism by which this flavonoid inhibits the formation of toxic $A\beta$ oligomers.

9.2 Methods

9.2.1 System Construction

The coordinates of monomeric $A\beta_{40}$ and $A\beta_{42}$ were taken from PDB entries 1BA4 [77] and 1IYT [78], respectively. For fibril-derived $A\beta_{40}$ and $A\beta_{42}$ dimers, a pair of peptides was extracted from the coordinate files of fibril models proposed by Petkova *et al.* [361] and Lührs *et al.* [278], respectively. An additional $A\beta_{42}$ dimer structure was taken from the NMR structure determined by Yu *et al.* [276]. Each of these structures was missing between 8-16 N-terminal residues, which were built in the xLeap module of AmberTools [362], assigning a fully-extended configuration to the amino acid backbone. These coordinates were aligned with the existing residues in VMD [363] to create full-length $A\beta$ dimers.

For systems containing two $A\beta$ peptides that were not yet aggregated, the final configurations of monomer simulations (described below) were used. Peptide pairs were combined at random and positioned in an arbitrary orientation, such that the minimum distance between all atoms was at least 1.4 nm. Three systems each of $A\beta_{40}$ and $A\beta_{42}$ non-aggregated dimers were prepared in this manner.

All peptide systems were centered in rhombic dodecahedral boxes with a minimum solute-box distance of 1.0 nm. Parameters for all species were taken from the GROMOS96 53A6 force field [174]. For morin, parameters were taken from our previous work [358]. SPC water [268] was used as a solvent, and all systems contained ~ 100 mM NaCl, including counterions to compensate for the net 3 charge on each $A\beta$ peptide. For systems containing morin, these molecules were placed randomly in the unit cell prior to solvation. All morin-treated systems contained a 2:1 mole ratio of morin: $A\beta$, except certain monomer simulations (denoted with a subscript “XM,” indicating systems with excess morin) which contained a 10:1 mole ratio.

9.2.2 Simulation Protocol

All simulations were performed using the GROMACS package, version 4.0.7 [180]. All systems were energy-minimized using the steepest descent method. Equilibration was then conducted in two phases, during which position restraints were placed on all heavy protein atoms. The first phase applied an NVT ensemble for 50 ps, using the Berendsen weak coupling method [187] to maintain the temperature of the system at 310 K. The protein and solvent (including ions and morin, if

applicable) were coupled separately. The second phase of equilibration applied an NPT ensemble for 50 ps, using the Nosé-Hoover thermostat [188, 189, 190] and Parrinello-Rahman barostat [191, 192] to maintain temperature (310 K) and pressure (1 bar), respectively. Production MD was then conducted in the absence of any restraints, using the same NPT ensemble. Simulations of $A\beta$ monomers were conducted for 150 ns, while all dimer systems were simulated for at least 250 ns. Simulations were stopped after the backbone root mean square deviation (RMSD) of the $A\beta$ peptides was stable for at least 100 ns, over which time data were analyzed. As a consequence, some simulations were considerably longer than 250 ns. Simulation systems, their designators, and a tabulation of individual simulation times are given in Tables 9.1 and 9.2. The total simulation time, for all systems and all replicates, was 23.65 μ s.

Table 9.1: Monomeric simulation system designators, contents, and trajectory lengths.

Classification	System Type	Contents	Designator	System Size (atoms)	Simulation Length (ns)
Monomers	$A\beta_{40}$ monomer	1BA4	M40	24,145	150 ns \times 6
	$A\beta_{40}$ monomer + morin	M40 + 2 morin	M40 _M	24,140	150 ns \times 6
		M40 + 10 morin	M40 _{XM}	24,018	150 ns \times 6
	$A\beta_{42}$ monomer	1IYT	M42	23,546	150 ns \times 6
	$A\beta_{42}$ monomer + morin	M42 + 2 morin	M42 _M	23,535	150 ns \times 6
		M42 + 10 morin	M42 _{XM}	23,425	150 ns \times 6

All simulations employed three-dimensional periodic boundary conditions. All bond lengths were constrained using P-LINCS [270], allowing an integration timestep of 2 fs. The neighbor list was cut off at 1.4 nm and updated every 10 fs. All short-range nonbonded interactions were cut off at 1.4 nm, with dispersion correction applied to energy and pressure terms to account for the truncation of van der Waals interactions. Long-range electrostatic interactions were calculated with the smooth particle mesh Ewald method [193, 194] using cubic-spline interpolation and a Fourier grid spacing of approximately 0.12 nm. All analyses were conducted using programs in the GROMACS package. Secondary structure was assigned according to the DSSP algorithm [271].

9.3 Results

9.3.1 Simulations of $A\beta_{40}$ Monomers

During the simulations in the M40 set, much of the initial helicity of $A\beta_{40}$ was lost in favor of random coil elements and short β -strands (Figure 9.1 and Table 9.3). In two simulations (M40-5 and M40-6, see Figure 9.1), short, parallel three-stranded β -sheets formed. In three of the remaining M40 simulations (M40-1, M40-3, and M40-4), antiparallel β -hairpins developed in N-terminal and

Table 9.2: Dimeric simulation system designators, contents, and trajectory lengths.

Classification	System Type	Contents	Designator	System Size (atoms)	Simulation Length (ns)	
Dimers	A β_{40} dimers	M40-1 + M40-6	D40 ₁	44,693	450 ns \times 1 250 ns \times 1 400 ns \times 1	
		M40-2 + M40-5	D40 ₂	39,818	500 ns \times 1 550 ns \times 1 600 ns \times 1	
		M40-3 + M40-4	D40 ₃	48,212	250 ns \times 1 500 ns \times 1 450 ns \times 1	
	A β_{40} dimers + morin	D40 ₁ + 4 morin	D40 _{1M}	44,584	450 ns \times 1 300 ns \times 1 250 ns \times 1	
		D40 ₂ + 4 morin	D40 _{2M}	39,739	250 ns \times 2 500 ns \times 1	
		D40 ₃ + 4 morin	D40 _{3M}	48,178	750 ns \times 1 450 ns \times 1 250 ns \times 1	
	A β_{42} dimers	M42-1 + M42-6	D42 ₁	51,693	450 ns \times 1 250 ns \times 1 350 ns \times 1	
		M42-2 + M42-5	D42 ₂	55,057	250 ns \times 2 550 ns \times 1	
		M42-3 + M42-4	D42 ₃	61,414	250 ns \times 2 500 ns \times 1	
	A β_{42} dimers + morin	D42 ₁ + 4 morin	D42 _{1M}	51,632	250 ns \times 2 350 ns \times 1	
		D42 ₂ + 4 morin	D42 _{2M}	55,002	250 ns \times 1 300 ns \times 1 500 ns \times 1	
		D42 ₃ + 4 morin	D42 _{3M}	61,362	450 ns \times 1 400 ns \times 1 500 ns \times 1	
	Preformed Dimers	A β_{40} fibril-derived dimer	Petkova dimer	A β_{40} P40	91,797	250 ns \times 3
		A β_{40} fibril-derived dimer + morin	P40 + 4 morin	P40 _M	91,742	250 ns \times 3
		A β_{42} fibril-derived dimer	Lührs A β_{42} dimer	L42	107,583	250 ns \times 3
A β_{42} fibril-derived dimer + morin		L42 + 4 morin	L42 _M	107,507	250 ns \times 3	
A β_{42} soluble dimer		Yu A β_{42} dimer	Y42	84,526	250 ns \times 3	
A β_{42} soluble dimer + morin		Y42 + 4 morin	Y42 _M	84,474	250 ns \times 3	

central regions of A β_{40} . In all simulations, the peptide adopted a collapsed structure, converging to an average R_g of 1.00 ± 0.05 nm, while simultaneously increasing heavy atom contacts by 8% over the final 100 ns of the six replicate simulations (Table 9.3 and Figure 9.2A). Hydrophobic contacts increased by 20% from the equilibrated structures (Table 9.3 and Figure 9.2C). The greater rela-

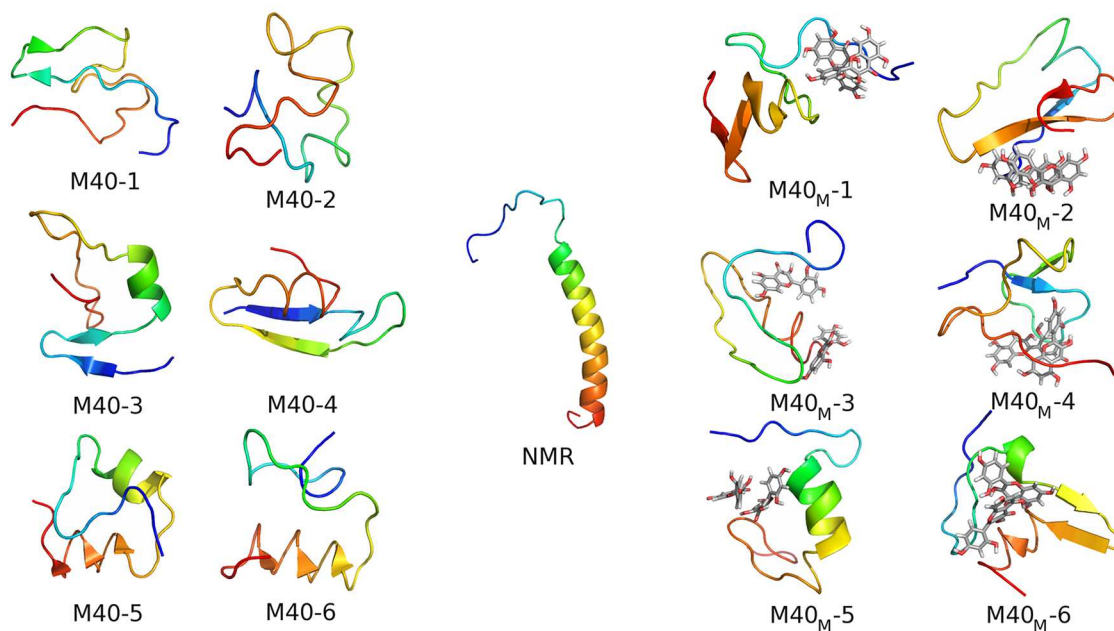


Figure 9.1: $A\beta_{40}$ structures. The starting structure (1BA4, from NMR) is centered between the final snapshots of M40 (left) and M40_M (right) systems. Peptides are rendered as cartoons and colored using a gradient from blue (N-terminus) to red (C-terminus). Morin molecules are shown as sticks and colored by element (C = gray, H = white, O = red).

tive increase in hydrophobic contacts (as compared to all heavy atom contacts) indicates that the structural changes observed in $A\beta_{40}$ were driven by the formation of a stable hydrophobic nucleus in the structure.

Hydrogen bonding patterns also changed over the course of the M40 simulations. In the initial, equilibrated structure of $A\beta_{40}$, there were approximately 24 intrapeptide hydrogen bonds, of which 19 (79.2%) involved backbone groups. Over time, as $A\beta_{40}$ collapsed and lost much of its initial helicity, 24 ± 3 total intrapeptide hydrogen bonds persisted, but only half of these involved backbone groups (Table 9.3). The native backbone hydrogen bonding pattern changed substantially, such that while the total number of intrapeptide hydrogen bonds was effectively constant, many of the initially interacting backbone groups became exposed to solvent as other, non-native hydrogen bonds developed.

The $A\beta_{40}$ peptides in simulations in the presence of morin (M40_M) behaved similarly to those of the M40 set described above. $A\beta_{40}$ had a tendency to become more collapsed than its initial configuration, with an average R_g comparable to that of the M40 set (Table 9.3). Binding of morin had only a small effect on the overall secondary structural characteristics of $A\beta_{40}$, decreasing β -strand content to $17 \pm 7\%$ and increasing total helicity and random coil content to $10 \pm 11\%$ and $39 \pm 6\%$, respectively (Table 9.3). Morin interacted equally with backbone and sidechain moieties on $A\beta_{40}$, forming 2.1 ± 0.9 hydrogen bonds with backbone groups and 2.3 ± 0.8 with

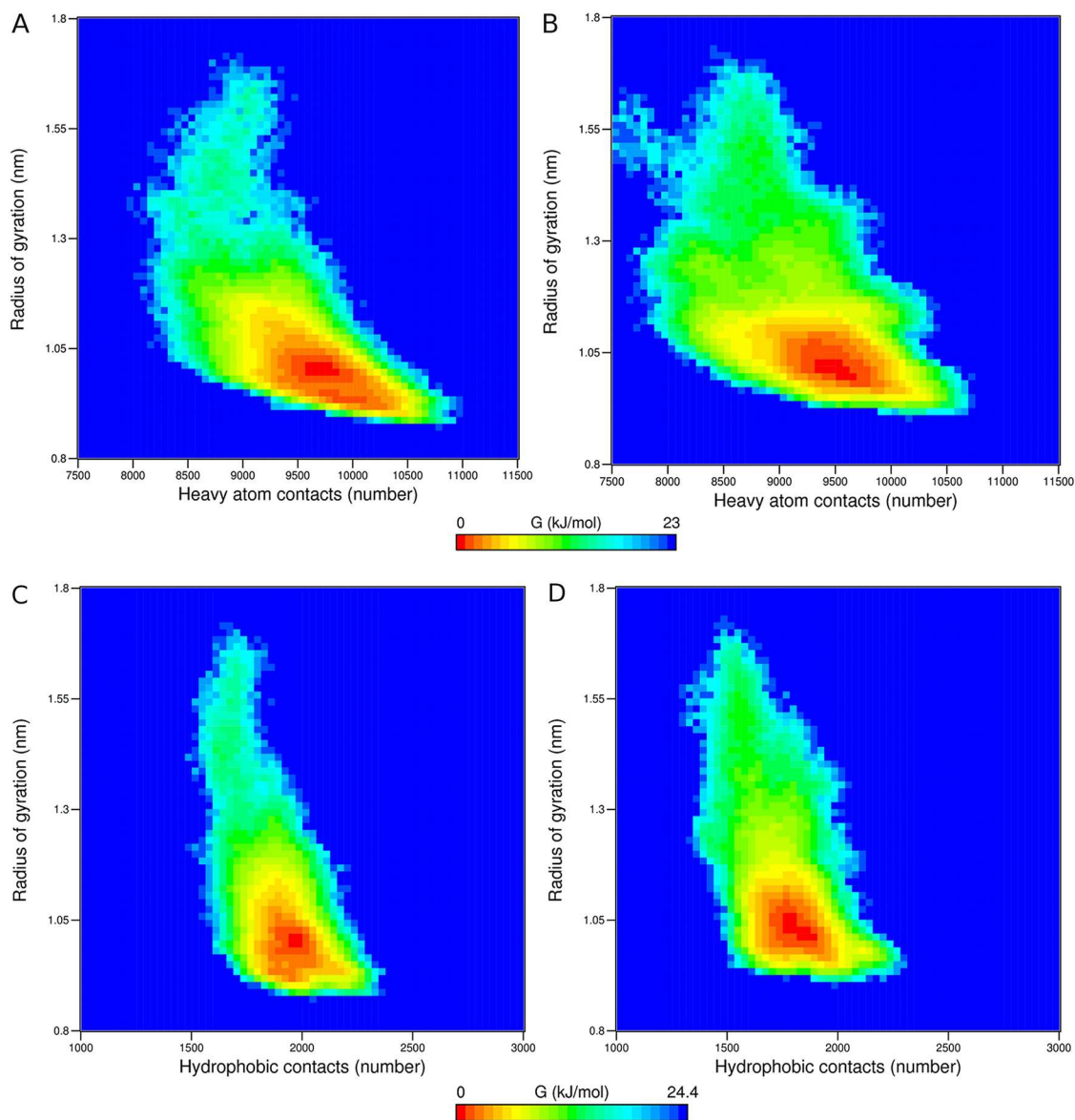


Figure 9.2: Gibbs free energy surfaces of R_g vs. heavy atom contacts for (A) M40 and (B) M40_M and R_g vs. hydrophobic contents for (C) M40 and (D) M40_M simulations. Data are a composite of all six replicate simulations in each set. A contact was defined as occurring within 0.6 nm. Note that there is a slight difference in the Gibbs free energy scale between the upper and lower panels.

Table 9.3: Structural parameters for monomeric A β_{40} peptides averaged over the last 100 ns of simulation time. All entries reflect averages with standard deviations from 6 simulations in each set. Initial values indicate the average of these parameters in all six equilibrated structures.

	Initial value	M40	M40 _M	M40 _{XM}
Total helix* (%)	45	7 \pm 8	10 \pm 11	15 \pm 9
β -strand [†] (%)	0	20 \pm 12	17 \pm 7	11 \pm 6
Bend (%)	10	24 \pm 7	22 \pm 8	16 \pm 4
Turn (%)	10	12 \pm 8	12 \pm 7	20 \pm 8
Random coil (%)	65	36 \pm 7	39 \pm 6	39 \pm 8
RMSD (nm)	N/A	1.1 \pm 0.1	1.03 \pm 0.06	0.85 \pm 0.08 [‡]
R_g (nm)	1.375 \pm 0.005	1.00 \pm 0.05	1.04 \pm 0.04	1.1 \pm 0.1 [‡]
Heavy atom contacts (count)	9013 \pm 73	9713 \pm 221	9373 \pm 317	8808 \pm 783 [‡]
Hydrophobic contacts (count)	1619 \pm 26	1950 \pm 38	1789 \pm 66 [‡]	1650 \pm 232 [‡]
Intra-peptide hydrogen bonds (count)	24 \pm 2	23 \pm 3	21 \pm 1	18 \pm 5
Backbone hydrogen bonds (count)	18.7 \pm 0.8	12 \pm 3	11 \pm 3	11 \pm 2

*Total helix is the sum of α -, 3_{10} -, and π -helical content.

[†] β -strand content reflects the sum of extended β -strand and isolated β -bridge structures.

[‡]Significantly different from control (M40) simulations, as determined by a two-tailed *t*-test with $p < 0.05$.

sidechain groups. These hydrogen bonds competed for intra-protein hydrogen bonds, reducing the total number of intra-protein hydrogen bonds relative to the M40 simulations (Table 9.3).

The most pronounced effect of morin binding to A β_{40} was a reduction in the number of hydrophobic contacts formed within A β_{40} . Fewer heavy atom contacts were formed within A β_{40} in the M40_M simulations than in the M40 simulations (Table 9.3 and Figure 9.2B), and the final value of 9373 \pm 317 represents an increase of 4% from the initial value. In addition, hydrophobic contacts were significantly reduced compared to the M40 results (Table 9.3 and Figure 9.2D). In the M40_M simulations, 1789 \pm 66 hydrophobic contacts formed (an 11% increase from the starting value of 1619 \pm 26), which was fewer than in the M40 systems (1950 \pm 38), which increased by 20% relative to the starting structure. Some broadening of the energy minimum in the case of the M40_M simulations is apparent (Figure 9.2), indicative of the ability of morin to induce greater structural heterogeneity in the final state of A β_{40} when exposed to 2:1 morin.

Having examined systems that modeled experimental conditions [159, 160], we sought to explore the effects of a large molar excess of morin on the structure of A β_{40} in a simulation set called M40_{XM}, which contained a 10:1 morin:A β mole ratio. Though this mole ratio would only be relevant when A β levels are extremely low, it was of interest to determine if there was any concentration dependence on the effects observed in the M40_M set. Such a mole ratio was also recently used in a simulation study by Liu *et al.* in the context of EGCG binding to A β [364]. In the M40_{XM} set, we found that, as with the M40_M set, the presence of even 10:1 morin had no significant effect on the secondary structure of A β_{40} (Table 9.3), however, like the simulations in the M40_M set, both heavy atom and hydrophobic contacts were significantly reduced relative to the M40 results (Table 9.3), with hydrophobic contacts increasing by just 2% from the number of

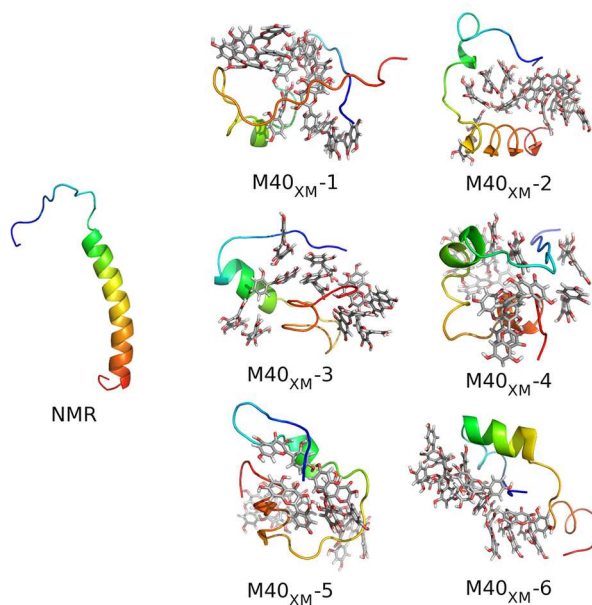


Figure 9.3: $A\beta_{40}$ structures. The starting structure (1BA4, from NMR) is shown at left, alongside the final snapshots (150 ns) of the $M40_{XM}$ (right) systems. Peptides are rendered as cartoons and colored using a gradient from blue (N-terminus) to red (C-terminus). Morin molecules are shown as sticks and colored by element (C = gray, H = white, O = red).

hydrophobic contacts initially present in the $M40_{XM}$ starting structure.

As a consequence of the fact that the high level of morin reduced the ability of $A\beta_{40}$ to form intrapeptide contacts, the peptide remained significantly more elongated than in the $M40$ set, with an R_g value of 1.1 ± 0.1 nm (Table 9.3). In the $M40_{XM}$ systems, morin formed networks through hydrogen bonding and aromatic stacking, aligning with hydrophobic regions on the peptide and forming hydrogen bonds with $A\beta_{40}$. Snapshots of the final structures of each of the $M40_{XM}$ systems are shown in Figure 9.3.

From our analysis of monomeric $A\beta_{40}$, it is clear that even a large molar excess of morin, beyond the concentrations yet explored *in vitro*, is unable to significantly alter the secondary structure content of the peptide, though the effects on tertiary structure in the $M40_{XM}$ set were augmented relative to the $M40_M$ simulations, indicating a concentration-dependent effect. At both 2:1 and 10:1 morin: $A\beta$, morin inhibited the formation of intrapeptide contacts, particularly hydrophobic contacts, thus interfering with the formation of a stable hydrophobic nucleus that may be critical for further aggregation. Thus, in the context of monomeric $A\beta_{40}$, it appears that morin affects tertiary, but not secondary, structure.

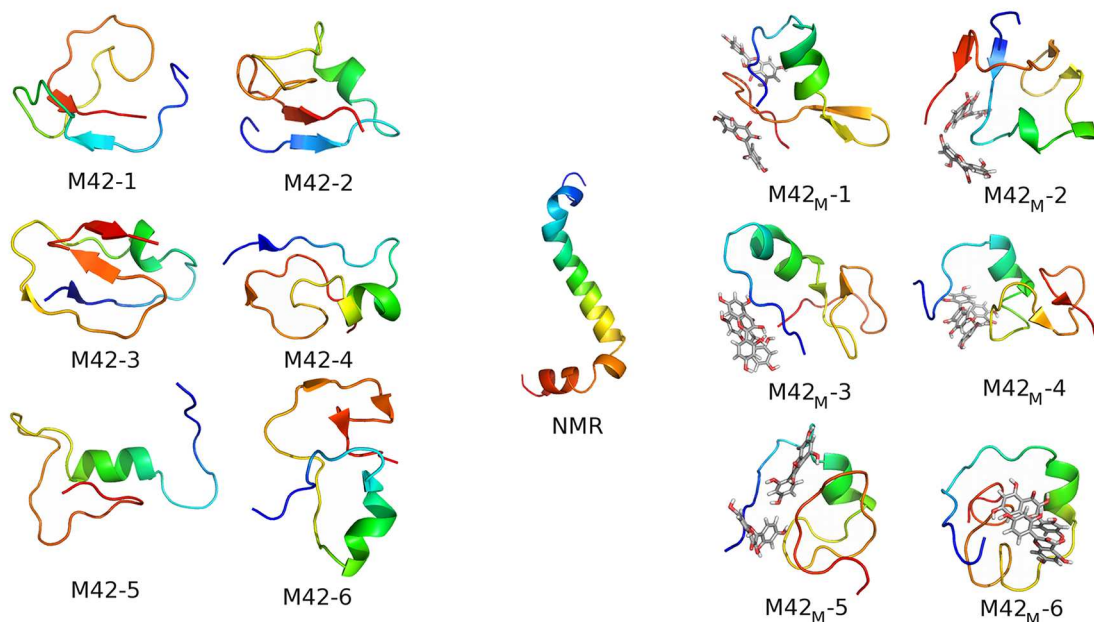


Figure 9.4: $A\beta_{42}$ structures. The starting structure (1IYT, from NMR) is centered between the final snapshots of M42 (left) and M42_M (right) systems. Peptides are rendered as cartoons and colored using a gradient from blue (N-terminus) to red (C-terminus). Morin molecules are shown as sticks and colored by element (C = gray, H = white, O = red).

9.3.2 Simulations of $A\beta_{42}$ Monomers

In the M42 simulations, as in the M40 simulation set described above, $A\beta_{42}$ tended to lose its considerable initial helicity over time and adopt a more compact structure (Table 9.4 and Figure 9.4). Random coil elements were the predominant secondary structural feature (Table 9.4), with short β -strand structures persisting in five of the six simulations (all except M42-5, see Figure 9.4). Helicity was generally retained in residues within the sequence Y₁₀EVHHQKLVFF₂₀. $A\beta_{42}$ stabilized at an average R_g of 1.00 ± 0.03 nm (Table 9.4), equivalent to that of $A\beta_{40}$. As was the case with $A\beta_{40}$, the $A\beta_{42}$ peptide also established more intrapeptide heavy atom contacts over time, increasing by 4% over the final 100 ns of each of the six replicate simulations (Table 9.4 and Figure 9.5A). Hydrophobic contacts increased by 12% over these trajectories (Table 9.4 and Figure 9.5C).

As with the M40 results, the greater relative increase in hydrophobic contacts suggests a prominent role for these contacts in driving the structural conversion of $A\beta_{42}$. The $A\beta_{42}$ peptides in the M42 set established a similar number of hydrophobic contacts as did the $A\beta_{40}$ peptides in the M40 set, though the starting value for these contacts was higher in $A\beta_{42}$ due to the kinked starting structure leading to more initial contacts in the C-terminal region. The greater number of total heavy atom contacts in M42 relative to M40 is a consequence of $A\beta_{42}$ having two additional amino acids.

Despite these subtle differences, the overall structures and level of contact within the untreated A β_{40} and A β_{42} peptides were very similar.

Table 9.4: Structural parameters for monomeric A β_{42} peptides averaged over the last 100 ns of simulation time. All entries reflect averages with standard deviations from 6 simulations in each set. Initial values indicate the average of these parameters in all six equilibrated structures.

	Initial value	M42	M42 _M	M42 _{X_M}
Total helix* (%)	69	15 ± 4	15 ± 5	24 ± 7 [‡]
β -strand [†] (%)	0	13 ± 5	11 ± 5	4 ± 5 [‡]
Bend (%)	0	26 ± 6	26 ± 3	20 ± 6
Turn (%)	23.8	8 ± 2	9 ± 4	11 ± 7
Random coil (%)	7.1	38 ± 6	39 ± 3	41 ± 6
RMSD (nm)	N/A	1.2 ± 0.1	1.2 ± 0.2	1.0 ± 0.1 [‡]
R_g (nm)	1.504 ± 0.005	1.00 ± 0.03	1.03 ± 0.06	1.20 ± 0.07 [‡]
Heavy atom contacts (count)	9665 ± 73	10029 ± 118	9979 ± 249	9116 ± 404 [‡]
Hydrophobic contacts (count)	1749 ± 25	1955 ± 93	1896 ± 75	1702 ± 112 [‡]
Intra-peptide hydrogen bonds (count)	26 ± 1	23 ± 1	23 ± 2	19 ± 3 [‡]
Backbone hydrogen bonds (count)	25 ± 1	11 ± 1	10.6 ± 0.8	11 ± 2

*Total helix is the sum of α -, 3_{10} -, and π -helical content.

[†] β -strand content reflects the sum of extended β -strand and isolated β -bridge structures.

[‡]Significantly different from control (M40) simulations, as determined by a two-tailed t -test with $p < 0.05$.

At the outset of the M42 simulations, there were approximately 26 total intra-protein hydrogen bonds, of which nearly all (96%) involved backbone groups (Table 9.4). The extensive backbone hydrogen bonding can be attributed to the considerable α -helical content present in the NMR structure, which is a result of structural determination in a nonpolar solvent [78]. As a result, over the course of our M42 simulations in water, much of the initial backbone hydrogen bonding was lost. Though 23 ± 1 total hydrogen bonds persisted in the A β_{42} structure over the six replicate simulations, only 47.8% involved backbone groups (Table 9.4), a result that is similar to the behavior of the A β_{40} peptides in the M40 simulation set discussed above.

The simulations of A β_{42} in the presence of 2:1 morin (the M42_M series) manifested behavior similar to the M40_M results. The A β_{42} peptide again tended to collapse and lose some of its initial α -helicity (Table 9.4), but its overall secondary structural features were essentially indistinguishable from those of untreated A β_{42} peptides in the M42 simulation set. Thus, as in the case of the M40_M simulations, the presence of morin had no effect on the loss of helicity or emergence of β -strand structures. In M42_M simulations, the A β_{42} peptide became more compact than its initial configuration; concomitant with this compaction was an increase in total and hydrophobic contacts (Table 9.4). Initially, there were 9678 ± 89 total heavy atom contacts within the A β_{42} peptide, a figure that increased by 3% over the course of the simulations. Hydrophobic contacts increased by 9%, a slightly smaller relative increase than was observed in the M42 simulations (12%). Thus, we conclude that the formation of hydrophobic contacts within the A β_{42} peptide was not influenced as strongly by the presence of morin as was the case for the A β_{40} peptide, though a measurable

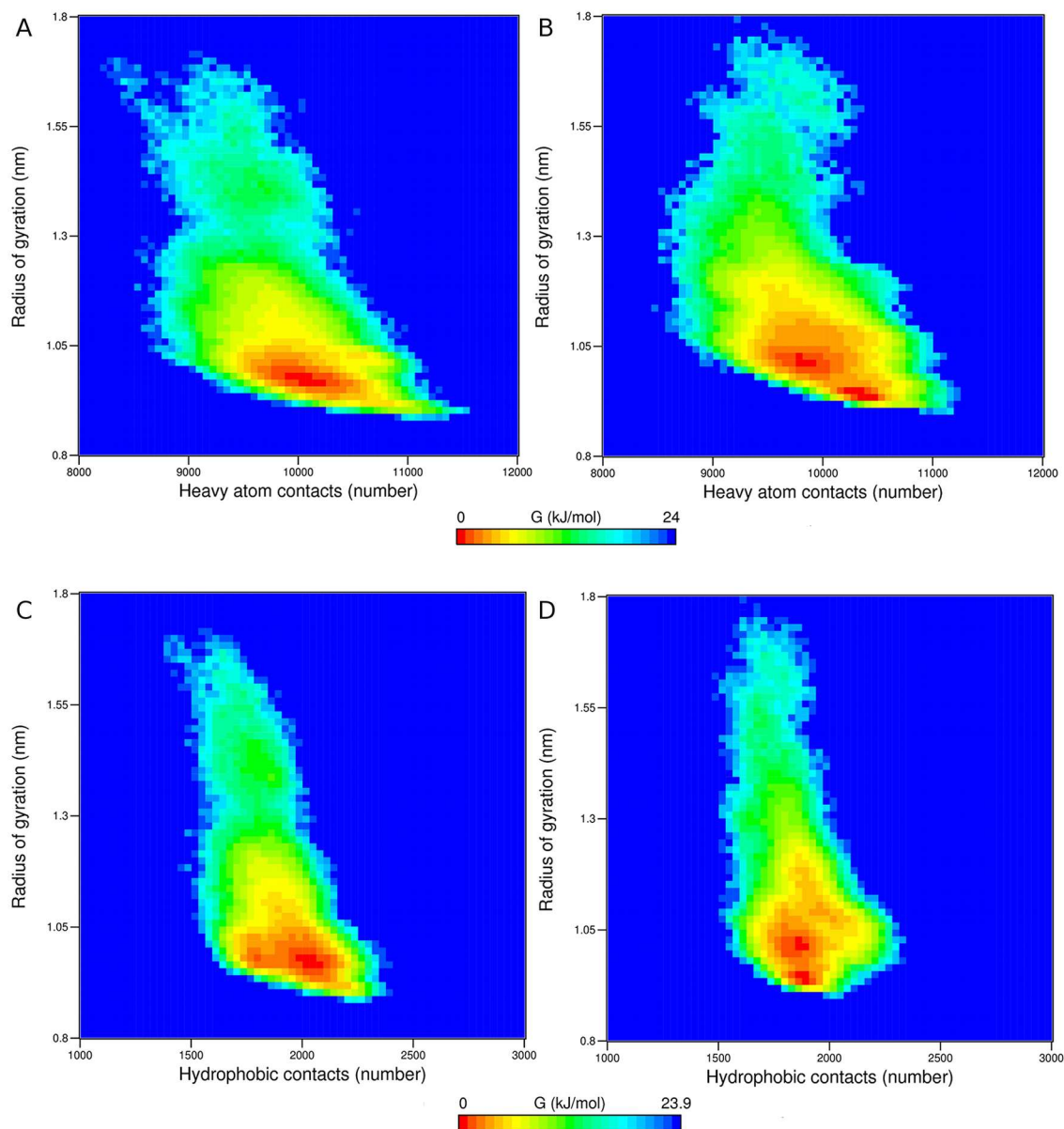


Figure 9.5: Gibbs free energy surfaces of R_g vs. heavy atom contacts for (A) M42 and (B) M42_M and R_g vs. hydrophobic contents for (C) M42 and (D) M42_M simulations. Data are a composite of all six replicate simulations in each set. A contact was defined as occurring within 0.6 nm. Note that there is a slight difference in the Gibbs free energy scale between the upper and lower panels.

difference was still observed. The importance of hydrophobic contacts in the context of A β_{42} aggregation will be described in greater detail later.

It is also important to note the appearance of two minima in the free energy surfaces plotted in Figures 9.5B and 9.5D (total and hydrophobic contacts in A β_{42} , respectively). In the absence of morin (Figures 9.5A and 9.5C), there is a single energy minimum in both free energy surfaces. The presence of dual energy minima in the M42_M simulations indicates that the presence of morin produced two clusters of configurations with comparable energy, derived from alterations to the tertiary structure of A β_{42} based on the location of morin binding. In trajectories where morin interacted principally with charged and polar residues in the N-terminal region of A β_{42} (M42_M-2, M42_M-5, and M42_M-6), close intrapeptide contacts were formed between the central hydrophobic cluster (CHC) of A β_{42} (L₁₇VFFA₂₁) and the C-terminal region, leading to more compact structures that populated the free energy minimum corresponding to the smaller value of R_g . In cases where morin bound to the CHC and other hydrophobic residues (M42_M-3 and M42_M-4), these contacts were reduced and A β_{42} was more elongated, leading to the free energy minimum at a greater value of R_g . In simulation M42_M-1, morin interacted with both polar and nonpolar regions on A β_{42} , leading to configurations that sampled both energy minima.

In the final set of simulations of A β monomers (the M42_{XM} set), the A β_{42} peptide was treated with a 10:1 mole ratio of morin. This treatment produced peptide configurations that were the most native-like out of all of the M42 simulations conducted here, having the lowest backbone RMSD (Table 9.4). The peptide retained significantly more helical content (Table 9.4) and produced significantly less β -strand content than in the control M42 simulations. The A β_{42} peptide remained more elongated than in the M42 set, with an average R_g of 1.20 ± 0.07 nm. Both heavy atom and hydrophobic contacts were significantly reduced in the M42_{XM} set compared to M42 (Table 9.4). These effects can be attributed to the network of morin molecules that acted to coat the peptide and restrict its ability to collapse and interconvert its secondary structure between helical and β -strand elements. Snapshots from the end of the six replicate simulations in the M42_{XM} set are shown in Figure 9.6.

9.3.3 Effect of Morin on A β_{40} Dimerization

To explore the process by which A β_{40} forms dimers in solution and the potential for morin to inhibit this process, we assembled three starting configurations of random dimer pairs, using the final configurations of the M40 simulation set as shown in Table 9.2. For each of the three resulting starting configurations, three independent simulations were initiated using different random velocities. This procedure resulted in nine independent trajectories in the D40 simulation set. Though there are effectively an infinite number of combinations of starting configurations and orientations that can be used to start these simulations, we feel we have established a reasonable subset here, recognizing that comprehensive sampling of all possible A β dimerization events would be a much larger, and perhaps intractable, task. For this reason, we analyze the data from the D40 simulation set as an aggregate, rather than considering the individual simulation subsets (D40₁, D40₂, and

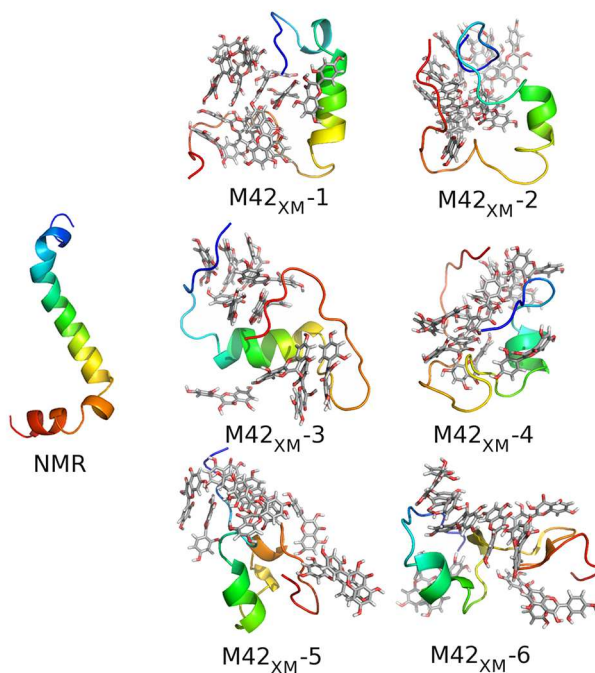


Figure 9.6: $A\beta_{42}$ structures. The starting structure (1IYT, from NMR) is shown at left, alongside the final snapshots (150 ns) of the $M42_{XM}$ (right) systems. Peptides are rendered as cartoons and colored using a gradient from blue (N-terminus) to red (C-terminus). Morin molecules are shown as sticks and colored by element (C = gray, H = white, O = red).

D40₃) as being distinct. All of their starting orientations are plausible and part of an infinitely large set of possible configurations, with no single subset being any more likely than the others. Representative examples will be highlighted throughout the discussion, as necessary, for particular simulations that are especially illustrative of specific phenomena.

In all D40 simulations, the $A\beta_{40}$ peptides (initially separated, with all interpeptide atomic distances greater than 1.4 nm) aggregated to form dimers. The average β -strand content over the final 100 ns of all simulations was $24 \pm 5\%$. Total helical content was $4 \pm 4\%$ over this same time period. These values represent a slight increase in β -strand content relative to the monomeric form of $A\beta_{40}$, concomitant with a small reduction in total helical content (Table 9.3).

The dimerization of $A\beta_{40}$ caused the total protein solvent-accessible surface area (SASA) to decrease by 2.78% (to an average area of $72.5 \pm 2.5 \text{ nm}^2$), with hydrophobic SASA decreasing by 2.5% and polar SASA by 2.86%. Thus, the dimerization of $A\beta_{40}$ was driven almost equally by burial of polar and nonpolar surfaces.

The aggregation of the $A\beta_{40}$ peptides observed in the D40 simulations produced compact structures with an average R_g of $1.25 \pm 0.05 \text{ nm}$ (Figure 9.7). The peptides formed 928 ± 288 heavy atom contacts, of which 244 ± 83 were hydrophobic (26.3%). In addition, 5 ± 3 interpeptide

hydrogen bonds formed, with 2 ± 2 (40%) involving backbone groups. Contacts formed between all regions of the A β_{40} peptides, but the most common occurrence was the establishment of contacts in the central hydrophobic cluster (CHC) of the peptide, which encompasses the sequence L₁₇VFFA₂₁, and is a region known to be critical for aggregation [365]. Close interpeptide contacts were established in this region in 7 of the 9 simulations in the D40 set. Short β -hairpin structures appeared in 5 of the 9 simulations in the D40 set (D40₁-1, D40₁-2, D40₁-3, D40₂-1, and D40₃-1), as shown in Figure 9.7.

The presence of morin in the D40_M simulation set did not prevent the aggregation of the A β_{40} peptides but did affect interpeptide interactions. Two principal binding modes emerged for morin, influencing the nature of the interpeptide interactions in different ways. The first we call “interfacial binding,” in which at least two molecules of morin were present between the two A β peptides (i.e., at the dimerization interface). The second binding mode was termed “surface binding,” which describes the cases in which the majority of the morin molecules were either clustered or scattered over the solvent-exposed surface of the A β peptides. In the D40_M simulation set, four trajectories (D40_{1M}-1, D40_{1M}-2, D40_{2M}-3, and D40_{3M}-3) manifested interfacial binding, while the remaining five trajectories (D40_{1M}-3, D40_{2M}-1, D40_{2M}-2, D40_{3M}-1, and D40_{3M}-2) showed surface binding (Figure 9.8).

Interfacial binding of morin in the D40_M set significantly ($p < 0.05$) reduced the total interpeptide heavy atom contacts (548 ± 118) and hydrophobic contacts (146 ± 40) relative to controls. That is, morin impaired the ability of the A β_{40} dimers to form stable hydrophobic nuclei that would normally be shielded from solvent. Interfacial binding similarly reduced the number of interpeptide backbone and total hydrogen bonds that formed over the trajectories to 1 ± 1 (a reduction of 50%) and 3 ± 1 (a reduction of 40%), respectively. The weakened interactions between the A β_{40} peptides in these simulations resulted in dimers with larger values of R_g (1.33 ± 0.07 nm) than those of the control D40 simulations.

Interfacial binding of morin resulted in the manifestation of less SASA (69.5 ± 2.4 nm²), a decrease of 7.6% relative to the initial configurations; hydrophobic SASA decreased by 7.4%, and polar SASA by 8.0%. Thus, with morin present between the two peptides, a greater amount peptide surface area was buried than in controls, but as in the case of the D40 controls, there was equal burial of hydrophobic and polar surface area. It is clear from the data presented above regarding interpeptide contacts that the presence of interfacial morin inhibited the formation of a stable hydrophobic nucleus within the dimers that formed, and thus, while the manner in which A β_{40} aggregated was unaffected by the presence of morin (both polar and nonpolar interactions were formed) the resulting dimers interacted more weakly than in the case of the D40 controls.

In the D40 simulations, dimerization of A β_{40} led to an increase in β -strand content relative to the M40 monomer simulations. This behavior was observed in the case of the D40_M simulations, as well, but to a lesser extent in the simulations that produced interfacial binding than in the D40 simulations. The total β -strand content in these four simulations (D40_{1M}-1, D40_{1M}-2, D40_{2M}-3, and D40_{3M}-3) was $22 \pm 7\%$. It thus appears that despite the weakened interactions between the A β_{40} peptides, the overall ability of A β_{40} to form β -strand structures was not inhibited significantly,

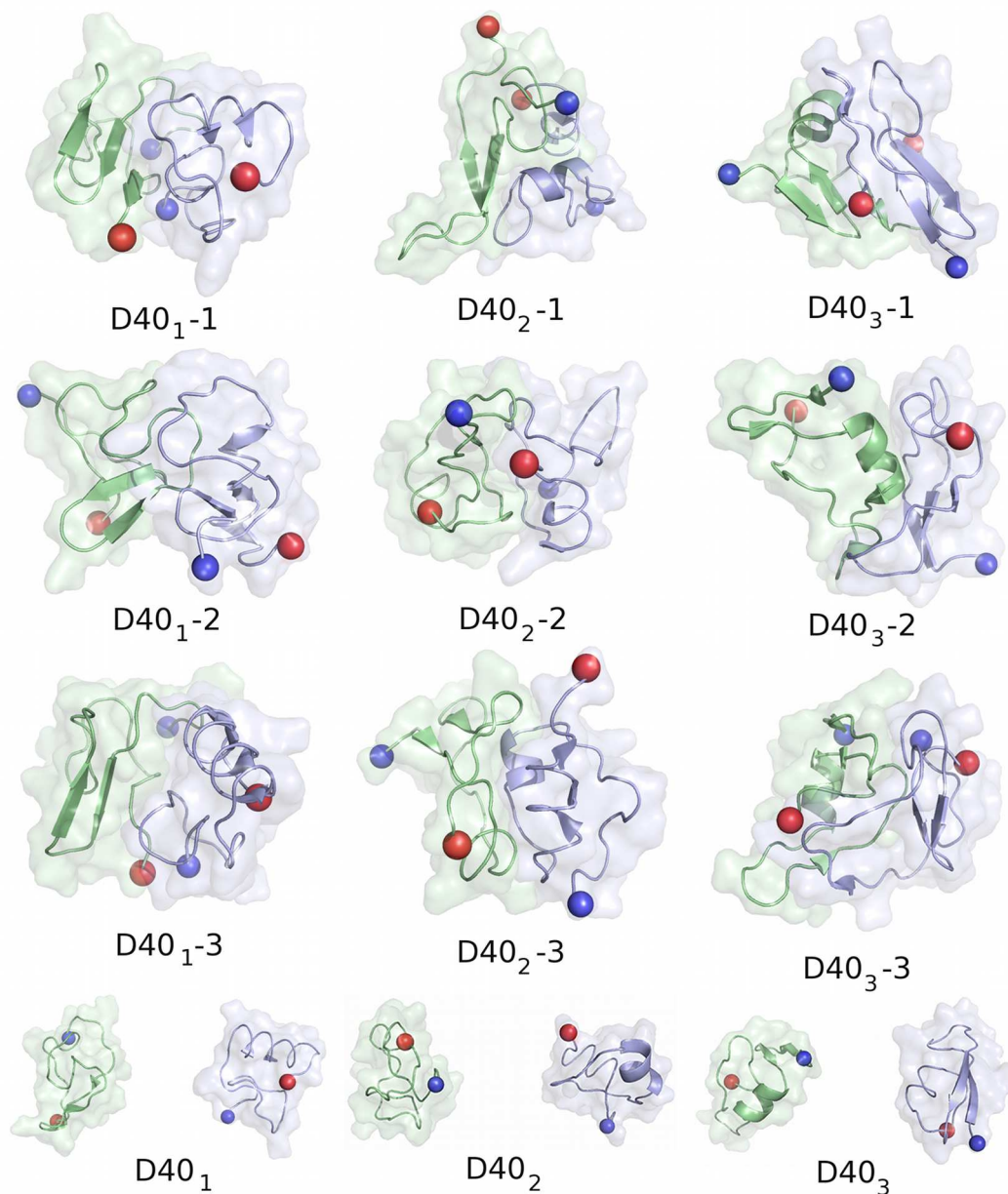


Figure 9.7: Structures of D40 dimers. Starting structures are shown below the final snapshots of D40 simulations. Individual peptides are rendered as cartoons and colored blue and green with an overlaid transparent surface to give an impression of overall topology. N- and C-termini are shown for each peptide as blue and red spheres, respectively.

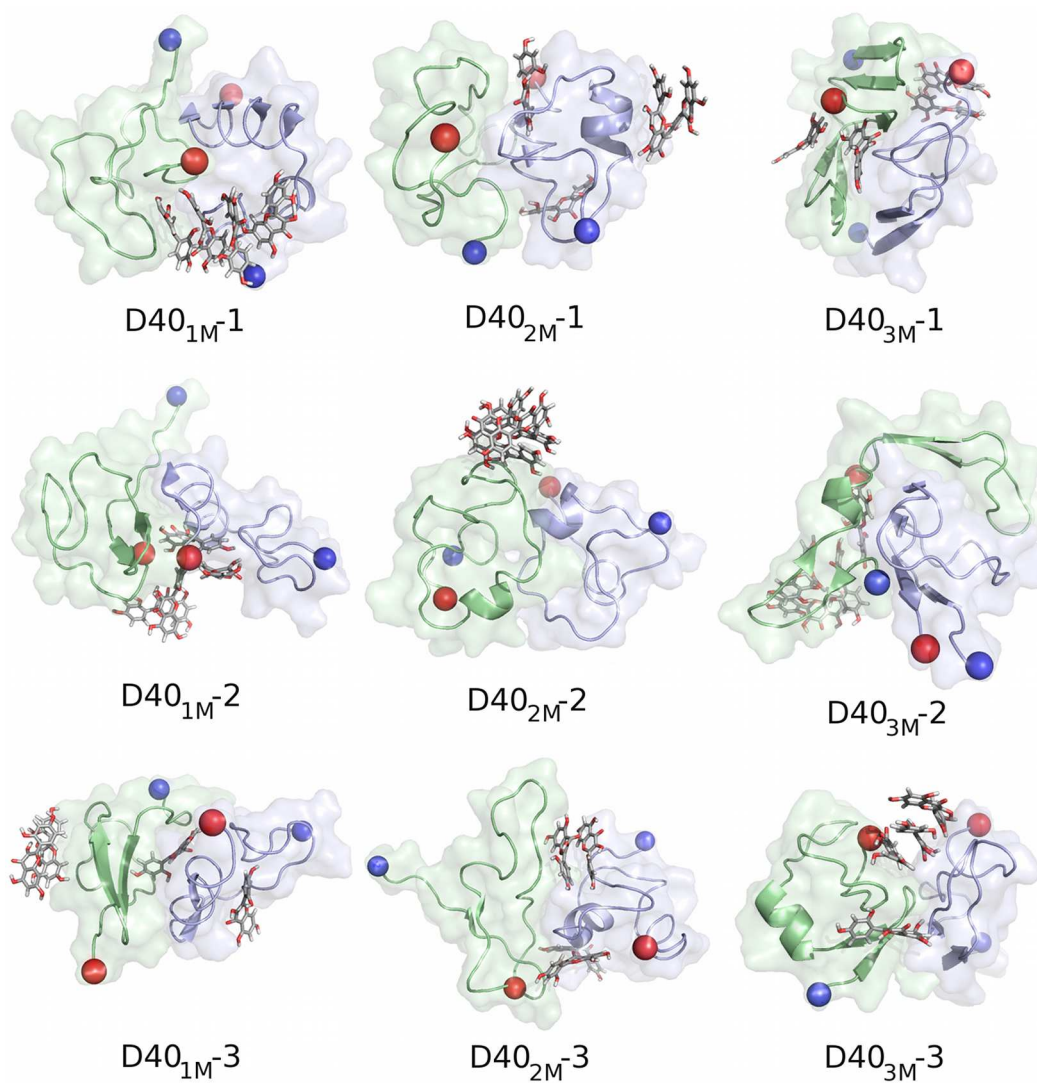


Figure 9.8: Structures of $D40_M$ dimers, taken from the final snapshots of $D40_M$ simulations. Peptides are rendered as in Figure 9.7. Morin molecules are shown as sticks and colored by element (C = gray, H = white, O = red). Starting configurations are identical to those shown in Figure 9.7, but with morin molecules randomly distributed in the simulation cell.

though some subtle changes were notable. The nature of secondary structural changes depended upon the residues with which morin interacted in these simulations (Figure 9.9). Interfacial binding of morin to the $A\beta_{40}$ dimers resulted in morin making close contact with a large number of residues throughout the $A\beta$ sequence (Figure 9.9A), whereas in the case of surface binding (Figure 9.9B), contacts were established with more discrete regions of the peptides. Surface binding of morin in the vicinity of residues 13-23 slightly increased the appearance of helical structures (Figure 9.9C) and slightly decreased the formation of β -strand structures towards the C-terminus (Figure 9.9D). Interfacial binding had no clear implications for secondary structure propensity of any region of the peptide.

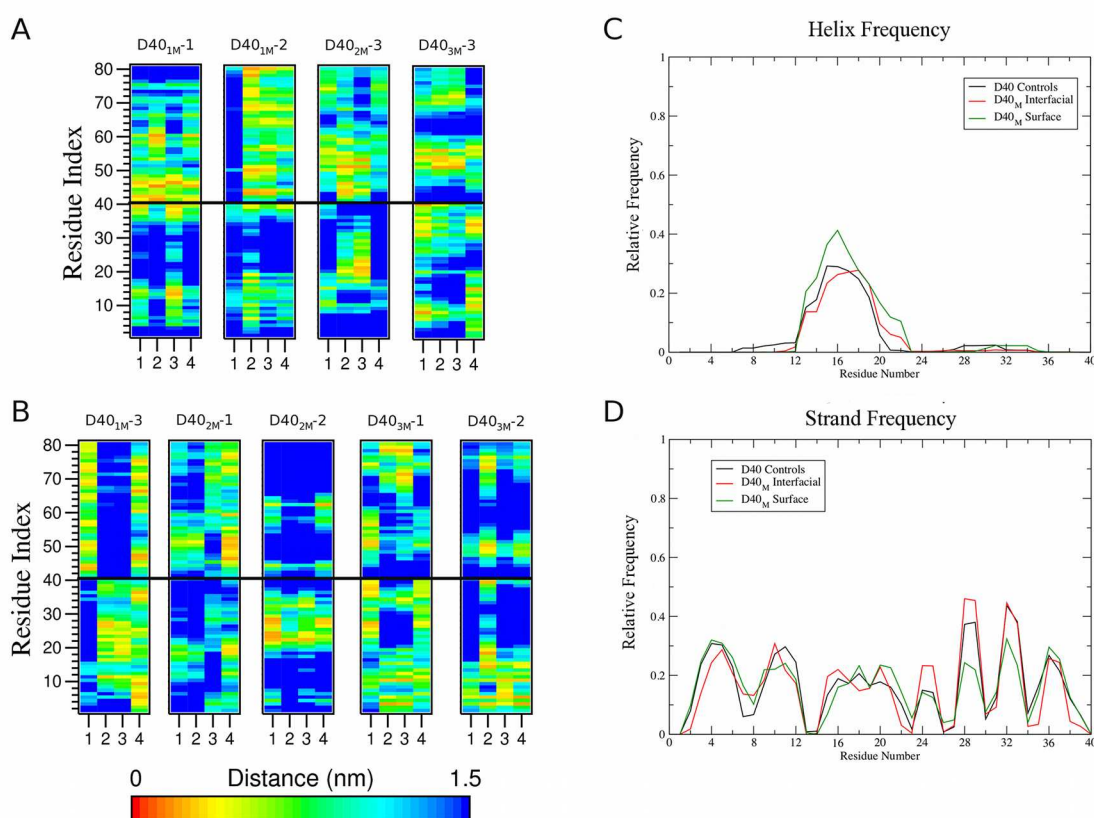


Figure 9.9: Distance matrices, averaged over the last 100 ns of each trajectory for $D40_M$ simulations in which morin bound (A) interfacially and (B) to the surface of the $A\beta_{40}$ dimers. For panels (A) and (B), the residue numbering on the y-axis corresponds to peptides 1 (1-40) and 2 (41-80). Secondary structure frequencies for (C) total helix and (D) β -strand content, also averaged over the last 100 ns of each trajectory. Frequencies were calculated from frames pooled over all simulations in the relevant sets.

Surface binding of morin to $A\beta_{40}$ in the $D40_M$ simulations led to an insignificant decrease in the number of interpeptide heavy atom contacts and hydrophobic contacts (699 ± 192 and 208 ± 59 , respectively). Interpeptide backbone hydrogen bonding was also reduced by 50% relative to

controls (1 ± 2 per timeframe), but total interpeptide hydrogen bonding was unaffected (6 ± 3 per timeframe). Overall, the size of the $A\beta_{40}$ dimers in the simulations that produced surface binding was comparable to controls, with an average R_g value of 1.27 ± 0.04 nm, indicating that these dimers were more compact than in the cases where morin bound interfacially.

As in the cases wherein morin bound interfacially, surface binding of morin to the $A\beta_{40}$ dimers led to a decrease in SASA relative to controls, with a total area of 68.3 ± 1.6 nm² observed over the final 100 ns of these simulations, a decrease of 8.6% from the SASA present in the initial configurations. Hydrophobic SASA decreased by 7.7% and polar SASA by 10.0%, both larger decreases than in the case of interfacial binding. Thus, as compared to interfacial binding of morin, surface binding resulted in slightly less surface area being exposed to solvent, which is primarily a consequence of the morin molecules coating the polar surface of the $A\beta_{40}$ dimers, a behavior manifested in the greater relative reduction in polar SASA.

Taken together with the interpeptide contact data, it can be concluded that surface binding of morin does not significantly affect the ability of $A\beta_{40}$ to dimerize, nor does it significantly affect the structural nature of the dimers. Surface binding of morin also had a minimal effect on the ability of $A\beta_{40}$ to form β -strand structures, with a total content of $25 \pm 16\%$ in these simulations, a result that was indistinguishable from controls in the D40 set. Conversely, interfacial binding of morin had a more pronounced impact on the dimer structures, resulting in significantly fewer interpeptide contacts than in the D40 controls and a larger value of R_g , indicating that the interpeptide interactions in these species were weakened by the presence of morin.

9.3.4 Effect of Morin on $A\beta_{42}$ Dimerization

Simulations from the D42 series were prepared in the same manner similar as those of the D40 set, deriving monomeric starting configurations from the final snapshots of the M42 simulations (Table 9.2). As with the D40 set, nine independent trajectories were generated (three each from three distinct starting configurations see Figure 9.10). As was the case in the D40 simulations, all of the D42 simulations produced dimers of $A\beta_{42}$ (Figure 9.10). Upon dimerization, the total β -strand content of $A\beta_{42}$ nearly doubled from $13 \pm 5\%$ (in the monomeric state, see Table 9.4) to $25 \pm 8\%$. The dimers observed over the course of these simulations had an average R_g of 1.30 ± 0.06 nm and formed an extensive network of intermolecular contacts. On average, 1040 ± 304 contacts were present between the $A\beta_{42}$ peptides, of which 245 ± 62 (23.6%) were hydrophobic. Along with the formation of these contacts, an average of 9 ± 4 total interpeptide hydrogen bonds were formed, of which 4 ± 2 involved backbone groups.

Dimerization of $A\beta_{42}$ in the D42 simulations led to an overall decrease in SASA, converging from a starting value of 80.7 ± 3.3 nm² to 75.1 ± 2.0 nm² over the final 100 ns of the simulations, a decrease of 6.94%. In contrast to the dimerization of $A\beta_{40}$ in the D40 simulations, wherein both polar and nonpolar surfaces were buried equally, the formation of $A\beta_{42}$ dimers was driven by burial of hydrophobic SASA, which decreased by 8.65%, while polar SASA decreased by 4.37%.

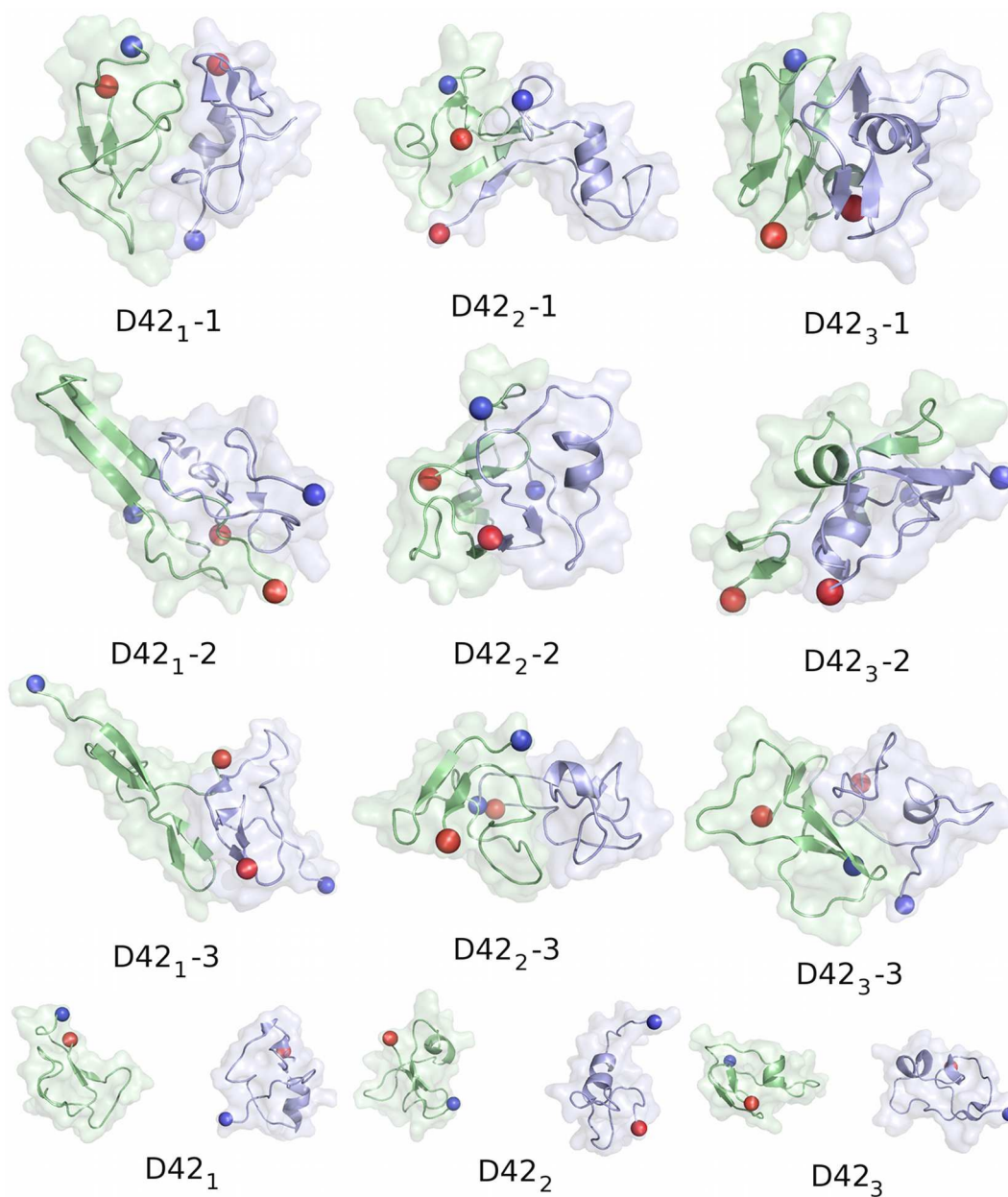


Figure 9.10: Structures of D42 dimers. Starting structures are shown below the final snapshots of D42 simulations. Individual peptides are rendered as cartoons and colored blue and green with an overlaid transparent surface to give an impression of overall topology. N- and C-termini are shown for each peptide as blue and red spheres, respectively.

Whereas dimerization of $A\beta_{40}$ in the D40 simulation set most frequently involved interactions within the CHC of the peptide, $A\beta_{42}$ dimers formed in the D42 set were principally driven by contacts within the C-terminal regions of the peptide sequence. The CHC is flanked by a sequence to its N-terminus that is largely polar, while residues C-terminal of the CHC are mostly nonpolar. This observation explains why $A\beta_{42}$ dimerization was mainly due to interactions between hydrophobic surfaces while $A\beta_{40}$ dimerization was due to both polar and hydrophobic interactions.

In several of the simulations in the D42 set, elongated β -hairpin configurations resembling those found in the mature fibril state were observed (Figure 9.10, specifically simulations D42₁-2, D42₁-3, and D42₃-3). In only one simulation in the D42_M (morin-treated) set was a similar conformation observed (D42_{3M}-3, Figure 9.11). The appearance of these structures in one-third of the D42 simulations is an expected behavior of $A\beta_{42}$, which is prone to more rapid aggregation in solution than is $A\beta_{40}$ [366, 367]. Coupled with the considerable increase in β -strand content relative to the monomeric state, it is clear that dimerization of $A\beta_{42}$ promotes the formation of oligomeric species rich in β -strands that further propagate aggregation.

As in the D40_M simulations, two binding modes for morin (interfacial and surface) were observed in the D42_M simulation set (Figure 9.11). In the instances wherein multiple morin molecules bound at the interface of the two $A\beta_{42}$ peptides (simulations D42_{1M}-2, D42_{2M}-1, D42_{3M}-2, and D42_{3M}-3), heavy atom contacts between the peptides were reduced to 694 ± 222 , of which 170 ± 70 (24.5%) were hydrophobic. Both of these values are reduced relative to the control D42 simulations by 33.3% for heavy atom contacts and 30.6% for hydrophobic contacts. Interfacial clusters of morin principally formed in the vicinity of hydrophobic residues in the $A\beta_{42}$ peptides, and as such they directly competed for peptide-peptide interactions in the hydrophobic, solvent-inaccessible region of the $A\beta_{42}$ dimer. Binding of morin in the hydrophobic core of the $A\beta_{42}$ dimer (interfacial binding) led to an overall decrease in total SASA from an initial value of 80.6 ± 2.5 nm² to 70.6 ± 2.1 nm², a reduction of 12.3%, leaving a total SASA that was 5.9% smaller than controls in the D42 simulation set.

Interpeptide hydrogen bonding was also inhibited in the presence of interfacially bound morin, with 1.4 ± 0.7 backbone hydrogen bonds maintained over the last 100 ns of each trajectory (a significant reduction relative to D42 controls). In total, 6 ± 2 interpeptide hydrogen bonds were maintained, a 33.3% reduction relative to D42 controls. The weakening of interpeptide interactions resulted in $A\beta_{42}$ dimers with a slightly expanded R_g , 1.33 ± 0.05 nm. There was no measurable difference in β -strand content in $A\beta_{42}$ when morin bound interfacially within the dimers ($25 \pm 8\%$ in the D42 simulations, and $24 \pm 8\%$ in these simulations).

When morin bound to the surface of $A\beta_{42}$ dimers (simulations D42_{1M}-1, D42_{1M}-3, D42_{2M}-2, D42_{2M}-3, and D42_{3M}-1), heavy atom and hydrophobic contacts were reduced (816 ± 240 and 224 ± 59 , respectively), although these reductions relative to controls were less dramatic than in the case of interfacially bound morin. Interpeptide hydrogen bonding was reduced, but not significantly, in these simulations (6 ± 3 total hydrogen bonds, of which 2 ± 2 involved backbone groups). In these simulations, the $A\beta_{42}$ dimers that formed were essentially as compact as those of the control D42 simulations, with an average R_g of 1.31 ± 0.05 nm. Interestingly, the binding

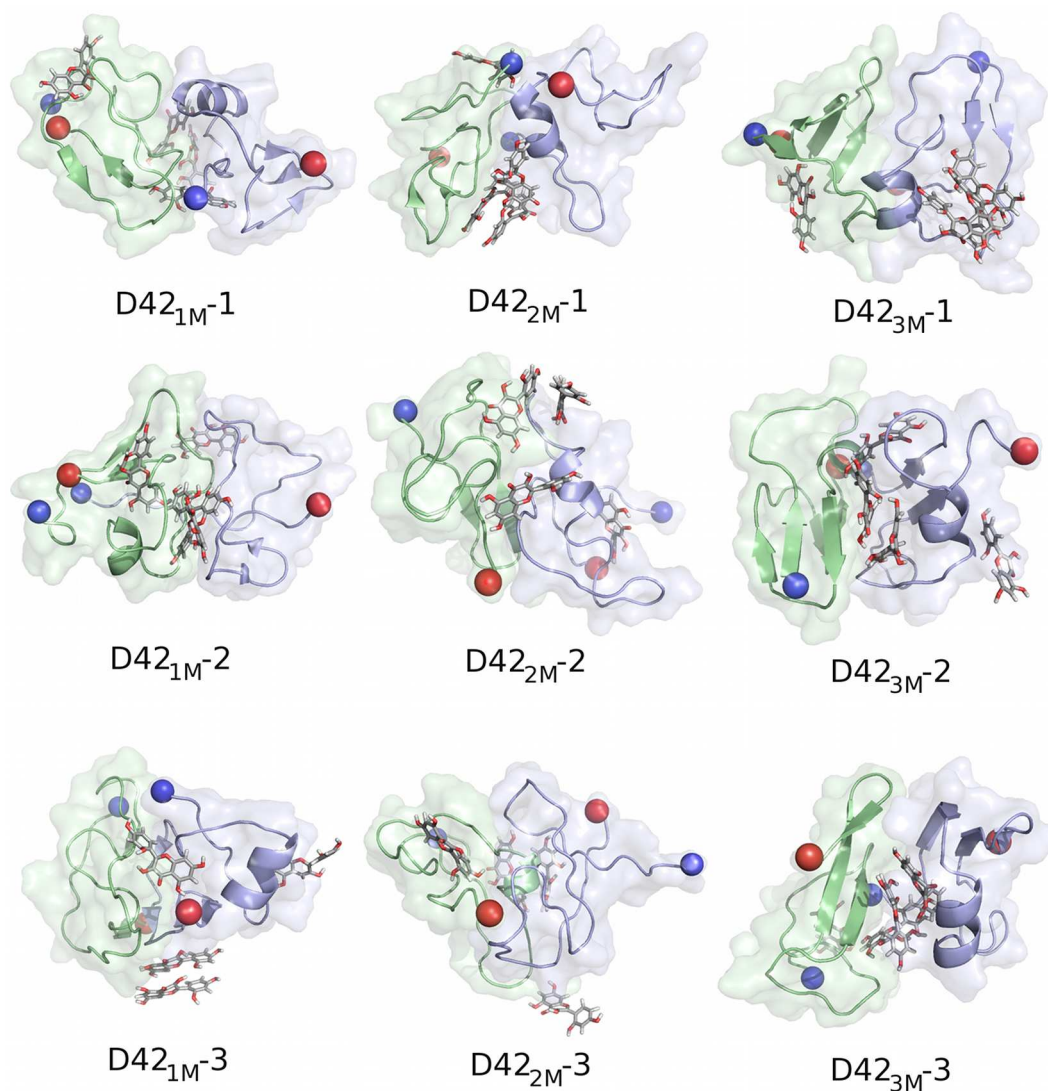


Figure 9.11: Structures of $D42_M$ dimers, taken from the final snapshots of $D42_M$ simulations. Peptides are rendered as in Figure 9.10. Morin molecules are shown as sticks and colored by element (C = gray, H = white, O = red). Starting configurations are identical to those shown in Figure 9.10, but with morin molecules randomly distributed in the simulation cell.

of morin to the surface of $A\beta_{42}$ dimers did induce a significant reduction in β -strand content ($17 \pm 5\%$) relative to control D42 simulations ($24 \pm 8\%$), a slight increase relative to the results of the M42 monomer simulations ($13 \pm 5\%$). We attribute this behavior to the locations along the $A\beta_{42}$ peptide to which morin bound. As in the case of the D40_M simulations, the specific binding sites of morin along the $A\beta$ sequence determined the frequency with which certain secondary structural features emerged in the D42_M simulations. Interfacial binding of morin often occurred at residues in the CHC or hydrophobic C-terminus (Figure 9.12A), causing an increase in total helical content (Figure 9.12C) and decrease in β -strand probability (Figure 9.12D) in the CHC and residues preceding it (H₁₄Q₁₅K₁₆), while also increasing the appearance of helical elements between residues 24-28 (Figure 9.12C). Surface binding of morin had a similar effect, but the decrease in β -strand frequency in the CHC was even more pronounced (Figure 9.12D). When morin bound to the surface of the $A\beta_{42}$ peptides, it principally associated with residues flanking the CHC rather than with the CHC directly (Figure 9.12B), inhibiting the elongation of this region into a β -strand structure.

Binding of morin to the surface of D42_M dimers reduced the overall SASA by 8.83%, from an initial value of $80.0 \pm 3.3 \text{ nm}^2$ to $73.0 \pm 2.3 \text{ nm}^2$. The subsequent reductions in hydrophobic and polar SASA were 10.4% and 6.3%. The reduction in these components of the SASA were both less than in the case of interfacial binding, indicating that even when morin bound to the surface of the $A\beta_{42}$ dimers, the N-terminal polar regions remained solvent-exposed, causing them to be disordered rather than organized in β -strands or otherwise buried such that they form close interpeptide interactions.

We conclude from these observations that the binding of morin to hydrophobic regions of $A\beta_{42}$ (interfacial binding) substantially alters the quaternary, but not secondary, structure of $A\beta_{42}$ dimers. In contrast, surface binding to principally polar residues in $A\beta_{42}$ strongly affects secondary structure, but effects on quaternary structure are diminished and thus the extent of peptide-peptide interactions (and thus aggregation) is comparable to control simulations. Thus, the manner in which morin binds to the $A\beta_{42}$ peptide is critical for determining the manner in which the structure of these dimers is affected. These findings are in contrast with the results for the dimerization of the $A\beta_{40}$ peptide, wherein the secondary structure content was largely unaffected by either interfacial or surface binding of morin. We attribute this phenomenon to the fact that $A\beta_{40}$ dimerization buried a greater amount of polar surface area than did the $A\beta_{42}$ peptide, which was driven together by hydrophobic interactions, leaving more polar area exposed. Burial of polar surface area in the case of $A\beta_{40}$ prevented the disruption of β -strand formation, while exposure to solvent in the case of the $A\beta_{42}$ dimers caused the β -strand content to decrease when these residues were exposed to morin and the aqueous solvent. The fundamental differences in aggregation of these two peptides explain how morin may affect each differently.

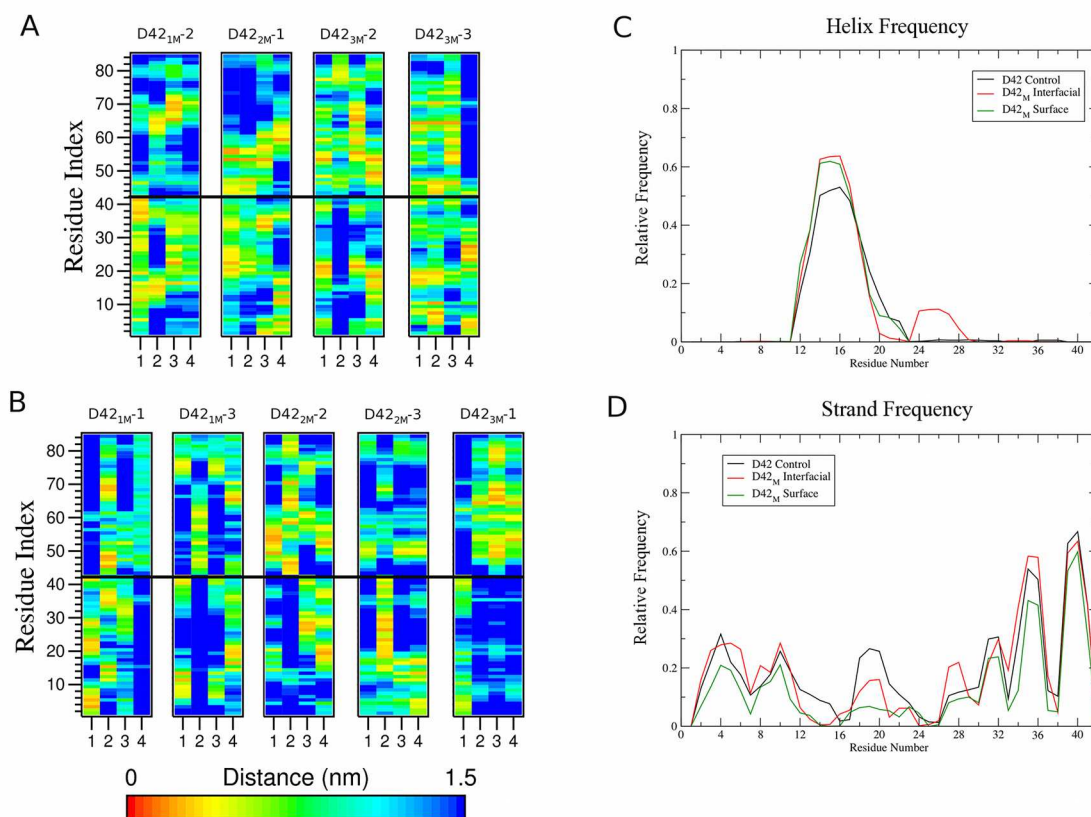


Figure 9.12: Distance matrices, averaged over the last 100 ns of each trajectory for $D42_M$ simulations in which morin bound (A) interfacially and (B) to the surface of the $A\beta_{42}$ dimers. For panels (A) and (B), the residue numbering on the y-axis corresponds to peptides 1 (1-42) and 2 (43-84). Secondary structure frequencies for (C) total helix and (D) β -strand content, also averaged over the last 100 ns of each trajectory. Frequencies were calculated from frames pooled over all simulations in the relevant sets.

9.3.5 Binding of Morin to Fibril-derived $A\beta_{40}$ and $A\beta_{42}$ Dimer

In the case of preformed $A\beta$ dimers, the effect of morin on the structural stability of these aggregates appears to be dependent upon the nature of the aggregate itself. Two models of fibril-derived dimers are considered here, one derived from the structure of the $A\beta_{40}$ fibril [361], and the other from the $A\beta_{42}$ fibril [278]. There are subtle differences between these two structures in terms of intra- and interpeptide packing and orientation [368], indicating that $A\beta_{40}$ and $A\beta_{42}$ may follow slightly different aggregation pathways to reach the mature fibril state. Since $A\beta_{42}$ is generally regarded as being more toxic than $A\beta_{40}$ due to its more rapid aggregation in solution [366, 367], understanding subtle differences in the structures of $A\beta_{40}$ and $A\beta_{42}$ at each stage of the aggregation cascade may be important in understanding how $A\beta$ aggregates and how this process can be inhibited. Examining fibril-derived dimers allows us to investigate the stability of “fibril-competent”

configurations, or end-stage aggregation.

In the context of the $A\beta_{40}$ fibril-derived dimer from Petkova *et al.* [361], herein referred to as P40, the control (untreated) systems showed that these dimers were very stable throughout the simulations, with a backbone RMSD of 0.80 ± 0.01 nm. The P40 dimers remained largely elongated, with an average R_g of 1.99 ± 0.02 nm. The two peptides twisted around one another during these simulations (Figure 9.13), a behavior that has been theorized to contribute to the stability of amyloid aggregates [202, 303, 314]. This twisting established an extensive network of interpeptide heavy atom contacts (2974 ± 117), an increase of 11.5% from the starting value of 2667 ± 24 .

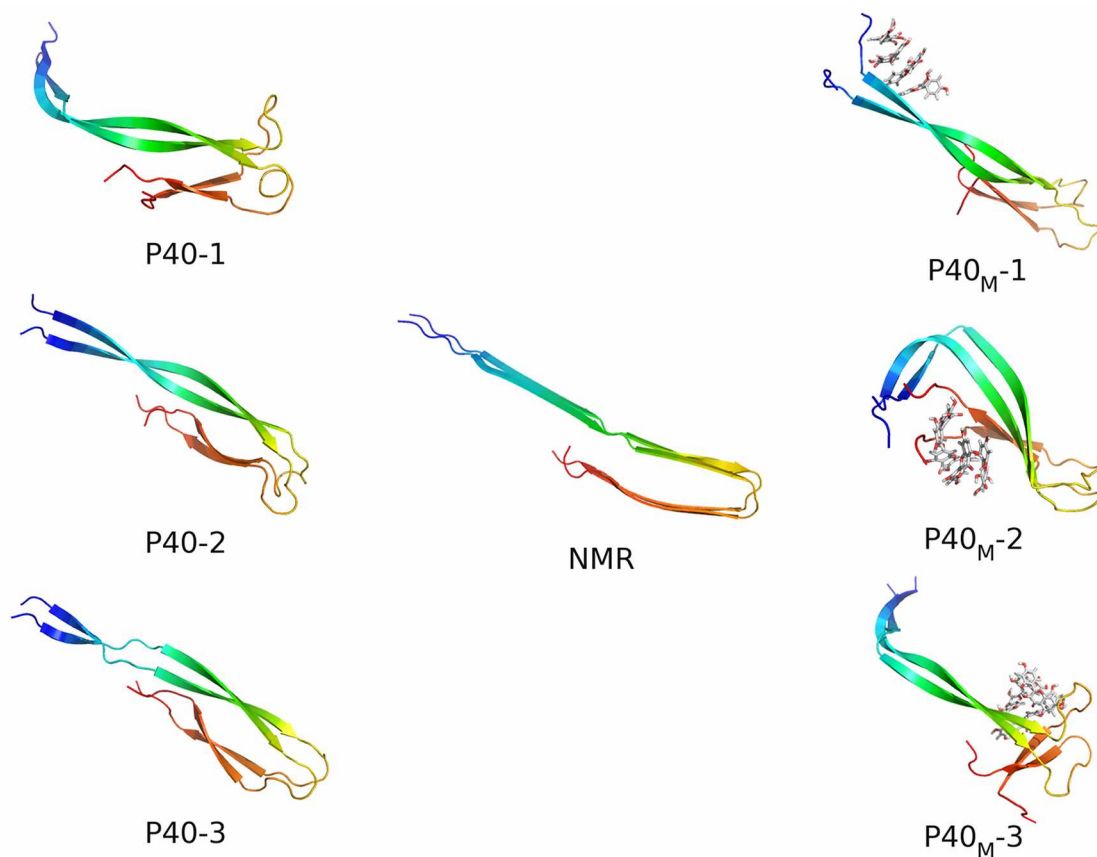


Figure 9.13: Petkova $A\beta_{40}$ dimers. The starting structure (from NMR, with the missing N-terminal 8 residues reconstructed) is centered between the final snapshots of P40 (left) and $P40_M$ (right) systems. Peptides are rendered as cartoons and colored using a gradient from blue (N-terminus) to red (C-terminus). Morin molecules are shown as sticks and colored by element (C = gray, H = white, O = red).

The effect of morin binding to the P40 dimers was dependent upon the region of $A\beta$ to which morin bound. In simulation $P40_M-1$ (Figure 9.13), morin bound to the polar N-terminal residues of the $A\beta_{40}$ peptides, resulting in no significant effect on the dimer structure, in terms of backbone RMSD (0.74 ± 0.08 nm), R_g (2.19 ± 0.05 nm), or interpeptide contacts (2949 ± 242). In this simulation,

two morin molecules bound to the P40 dimer in monomeric form, diffusing along much of the structure before forming strong interactions with the polar N-terminal region. The other two morin molecules bound to the P40 structure as a dimer, ultimately interacting with the other monomeric morin molecules to form a tetramer (Figure 9.13) that persisted for the remainder of the simulation. This cluster of morin interacted principally with residues Phe4, Arg5, His6, Asp7, and Tyr10 of the A β sequence, forming hydrogen bonds to the polar residues and π -stacking interactions with Phe4 and Tyr10.

In simulations P40_M-2 and P40_M-3, morin bound to hydrophobic residues towards the C-terminus of the P40 dimers, destabilizing the dimer structure, manifesting in increased RMSD (1.45 ± 0.03 nm), a more collapsed structure ($R_g = 1.8 \pm 0.2$ nm), and a slight decrease in interpeptide contacts (2920 ± 170). In simulation P40_M-2, one morin molecule bound to the hydrophobic C-terminal residues, while a trimeric assembly of morin molecules formed near the polar bend region. Interactions of this trimer with polar residues in the bend (most notably, Glu22, Asp23, and Ser25) led to the formation of a twisted, four-stranded β -sheet with the morin trimer spanning the CHC and bend regions. The remaining morin molecule remained tightly associated with the hydrophobic C-terminal residues, causing local perturbation of the backbone in this region. A tetrameric cluster of morin formed over time, interacting with residues in the CHC and C-terminal region (Figure 9.13). In P40_M-3, all four morin molecules aggregated in solution prior to binding to the P40 dimer structure. The morin tetramer bound to the polar bend region, and after a short time, split into two dimers, one of which dissociated from the P40 structure while the other moved to interact with residues Phe19 and Phe20 of the CHC. The morin tetramer re-formed over time, interacting with this Phe dyad and the Asp23-Lys28 salt bridge, causing the interpeptide backbone hydrogen bonds of the bend region to separate (Figure 9.13).

It thus appears that the most pronounced effects on the P40 dimer structure appeared when morin bound to hydrophobic residues in the A β sequence, particularly the CHC and C-terminal region. Upon binding to these residues, the P40 dimer adopted a more distorted conformation, a twisted β -sheet in the case of P40_M-2, and a perturbed dimer with disrupted backbone hydrogen bonding in the case of P40_M-3. Binding to polar residues towards the N-terminus of the structure (as in P40_M-1) did not affect the stability of the P40 dimer.

Simulations of the preformed dimer derived from the A β ₄₂ fibril, determined by Lührens *et al.* using NMR [278], and herein referred to as the L42 model, showed that this dimer structure was very stable, even when treated with 2:1 morin. In the control L42 simulations, the dimer structure remained elongated ($R_g = 2.0 \pm 0.2$ nm) over the course of the simulations, while simultaneously twisting (Figure 9.14) and reaching an RMSD value of 1.2 ± 0.3 nm. Due to these structural changes, the number of interpeptide heavy atom contacts stabilized at a value of 2505 ± 268 , an increase of 26.4% from the starting value of 1982 ± 32 . In the L42 simulations, the bend region encompassing the Asp23-Lys28 salt bridges was more flexible than in the P40 dimer simulations, resulting in configurations with propeller-like twists (Figure 9.14).

The association of morin with the L42 dimer showed little effect in any of the three replicate L42M simulations. The A β ₄₂ dimer remained elongated ($R_g = 2.1 \pm 0.1$ nm) and showed similar

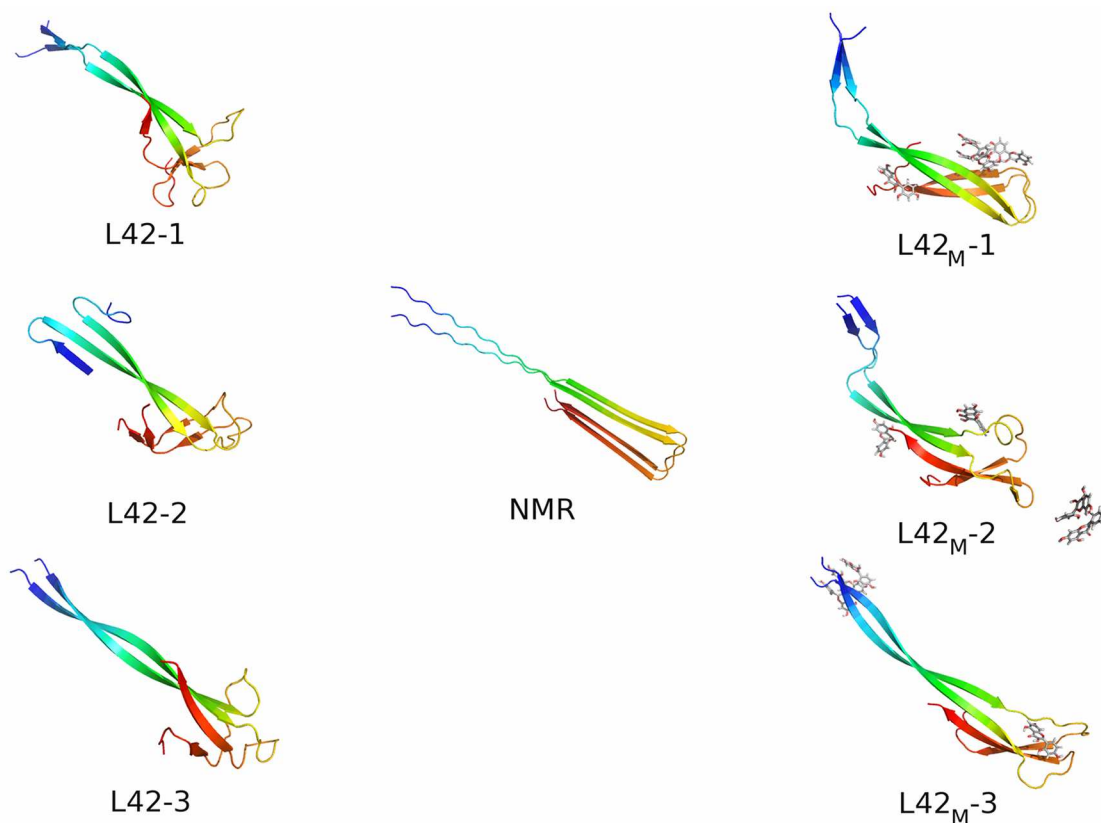


Figure 9.14: Lührs $A\beta_{42}$ dimers. The starting structure (from NMR, with the missing N-terminal 16 residues reconstructed) is centered between the final snapshots of L42 (left) and L42M (right) systems. Peptides are rendered as cartoons and colored using a gradient from blue (N-terminus) to red (C-terminus). Morin molecules are shown as sticks and colored by element (C = gray, H = white, O = red).

structural features to the controls. The average RMSD for $A\beta_{42}$ dimers in the L42_M set was 1.0 ± 0.2 nm, and the dimers formed 2581 ± 110 interpeptide contacts. Transient destabilization of several regions in the L42 structure was observed in these simulations and will be examined here.

In the L42_M-1 simulation, two morin molecules bound as monomers to the $A\beta$ dimer, one towards the N-terminus and the other at the C-terminus. Shortly thereafter, the remaining morin molecules bound as a dimer to the bend region, destabilizing the Asp23-Lys28 salt bridge and backbone hydrogen bonding. The morin dimer then moved towards the C-terminal hydrophobic region, forming a trimer with the morin molecule already there. The backbone of the bend region reformed some native hydrogen bonding, but the salt bridges did not re-form. For the remainder of the trajectory, the morin molecules exchanged dimer and trimer partners and resulted in the structure shown in Figure 9.14.

In simulation L42_M-2, one morin molecule bound to the N-terminal polar region, while the re-

maining three molecules bound as a trimer to Lys28 and Ile32. In this position, the Asp23-Lys28 salt bridge was disrupted, after which the trimer diffused towards the CHC, a phenomenon that led to distortion of the bend region. Eventually, a tetrameric morin cluster formed, covering the CHC, but this complex was short-lived, and ultimately only a dimer of morin remained bound to the L42 structure, as the remaining two molecules diffused away from $A\beta$ after interacting with Asn27.

In the final trajectory, L42_M-3, two morin molecules bound as a dimer to the N-terminal polar region, while the remaining two molecules bound as monomers to the N- and C-terminal regions. Ultimately, a trimer of morin formed around Phe4, Arg5, and Asp7, while the remaining morin monomer interacted with the Asp23-Lys28 salt bridge to disrupt it and destabilize surrounding backbone hydrogen bonding.

9.3.6 Binding of Morin to an Oligomer-derived $A\beta_{42}$ Dimer

Simulations in the Y42 set modeled a dimer of $A\beta_{42}$ that was derived from a soluble oligomer of $A\beta_{42}$ [276]. Thus, we believe this structural model to be the most representative of the early aggregated state of $A\beta_{42}$ both *in vitro* and *in vivo*. The structure is characterized by a pair of twisted, antiparallel intramolecular β -strands from residues 18 - 32 (connected by a hairpin bend from residues 23 - 28) and parallel intermolecular β -strands from residues 33 - 42 that form the dimerization interface between the two peptides. These features were well preserved during the simulations in the Y42 set (Figure 9.15). Over the course of these trajectories, N-terminal residues formed short β -strand structures but were principally disordered, in agreement with experimental findings [276].

Though there was considerable structural heterogeneity within the dimer structures in the Y42 series, many gross structural features were consistently observed among these trajectories. The principal secondary structural elements in the Y42 dimers were β -strand ($51 \pm 2\%$) and random coil ($34 \pm 1\%$), with the remaining structure described by a combination of turns and bends. Helical content was negligible in all cases. The Y42 dimers remained largely elongated, establishing an average R_g of 1.8 ± 0.2 nm, from an initial value of 2.4 nm. This measured size is in excellent agreement with results obtained by Walsh *et al.* [369]. Twisting and compaction of the N-terminal region of each $A\beta_{42}$ peptide in the dimer accounted for the majority of the change in R_g (Figure 9.15). These structural changes amounted to an average backbone RMSD for the dimer of 1.19 ± 0.08 nm.

As the Y42 dimers twisted, the $A\beta_{42}$ peptides increased both heavy atom contacts and interpeptide hydrogen bonding. Initially, there were 943 ± 40 heavy atom contacts between the peptides, but over the final half of the 250-ns trajectories, 2201 ± 187 contacts were established, an increase of 133.4%. In addition, there were initially just 7 interpeptide hydrogen bonds within the Y42 dimers (principally in the C-terminal parallel β -strands), but over time, 22 ± 2 total interpeptide hydrogen bonds were established, of which 16 ± 4 involved backbone groups.

The addition of morin in the Y42_M systems had a pronounced impact on the structure of the $A\beta_{42}$

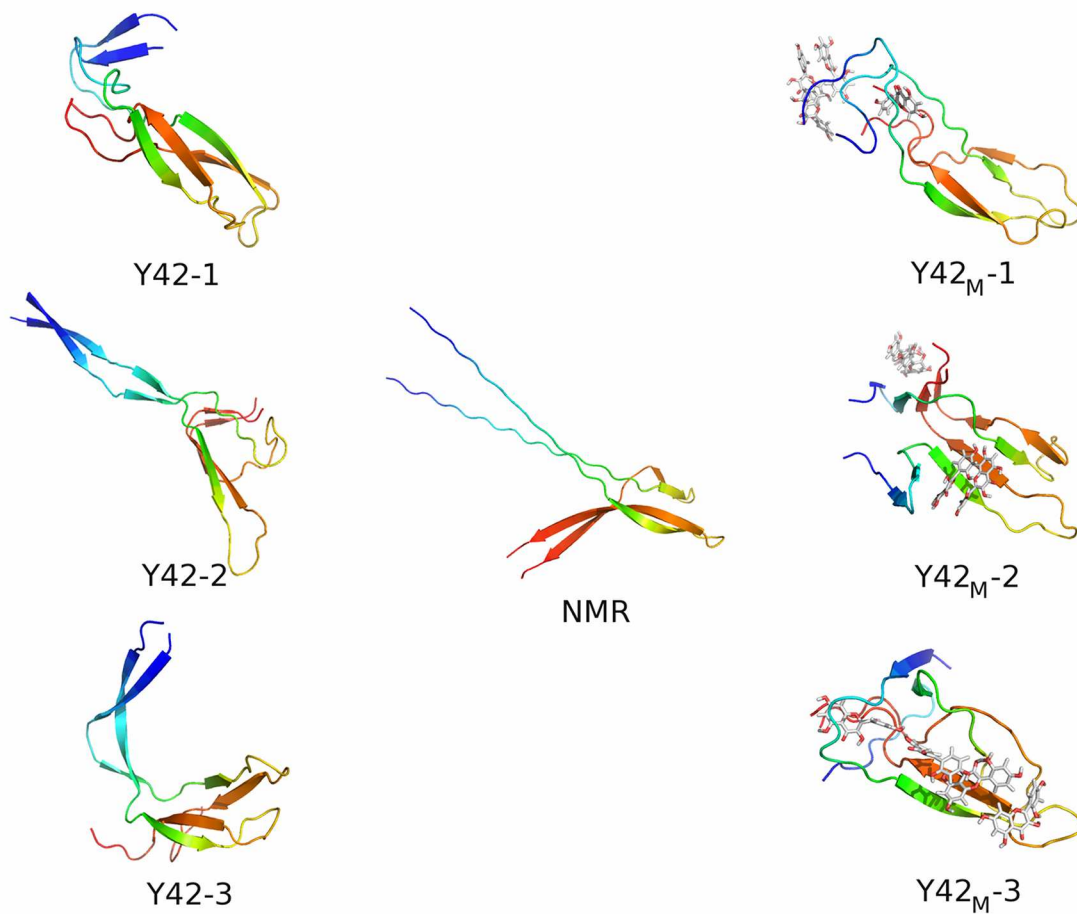


Figure 9.15: Yu $A\beta_{42}$ dimers. The starting structure (from NMR, with the missing N-terminal 15 residues reconstructed) is centered between the final snapshots of Y42 (left) and Y42M (right) systems. Peptides are rendered as cartoons and colored using a gradient from blue (N-terminus) to red (C-terminus). Morin molecules are shown as sticks and colored by element (C = gray, H = white, O = red).

dimers. Binding of morin to the dimer structure (Figure 9.15) reduced the β -strand content of the $A\beta_{42}$ peptides to $46 \pm 4\%$ and increased the random coil content $36 \pm 5\%$ relative to the control Y42 simulations. The $Y42_M$ dimers also became more compact than the Y42 dimers, with an average R_g of 1.5 ± 0.1 nm. Morin binding also reduced the strength of the interaction between the two peptides, such that only 1917 ± 293 heavy atom contacts and 17 ± 3 total hydrogen bonds (12.2 ± 0.8 of which involved backbone amides) persisted throughout the $Y42_M$ trajectories.

When morin associated with the $Y42_M$ dimers, the twisting present in the Y42 simulations was inhibited. Instead, β -strand structures were converted to random coils (as in $Y42_M$ -3 and, to a lesser degree, $Y42_M$ -1) or flattened β -sheets ($Y42_M$ -2). N-terminal residues contained almost exclusively random coil structures, in contrast to Y42 simulations, in which short β -strands evolved over time. The more dramatic changes in structure induced by the binding of morin to the $A\beta_{42}$ peptides resulted in a backbone RMSD of 1.5 ± 0.2 nm, considerably higher than the RMSD values of the Y42 dimers, 1.19 ± 0.08 nm.

In simulation $Y42_M$ -1, a single morin molecule bound to polar residues including Ser8, Tyr10, and His13 and diffused along the N-terminal polar region before associating with the C-terminus and hydrophobic residues in this region. The remaining morin molecules deposited on the Y42 dimer as a trimer, which broke and re-formed over the course of the simulation, interacting with many of the polar N-terminal residues.

For the first 75 ns of $Y42_M$ -2, monomeric morin molecules weakly associated with the Y42 dimer, binding and dissociating from residues in both the N-terminal and bend regions. A dimer of morin bound to Asp1 and Glu3 at approximately 98 ns, where it remained for the duration of the trajectory. The other two morin molecules bound as monomers to the bend region and nearby CHC, eventually forming a dimer that bridged these regions on both peptides, leading to the overall flattening of the Y42 dimer structure.

In the final simulation in this set, $Y42_M$ -3, a morin dimer bound quickly to the Asp23-Lys28 salt bridge and was further stabilized in this location through interactions with Glu22. An additional morin molecule bound to Asn27 and diffused to associate with this dimer, forming a trimer very quickly. The final morin molecule bound to residues Phe4, Ser8, and Glu11. Over time, the morin trimer moved to associate with the CHC, and as the Y42 dimer twisted in response to these interactions, a morin tetramer formed as interactions were established among all four morin molecules (Figure 9.15).

9.4 Discussion

The early stages of $A\beta$ aggregation are crucial to the progression of Alzheimer's disease, as low molecular weight oligomeric species are highly toxic to cells [84, 88]. Thus, therapeutic intervention in Alzheimer's disease is likely to be most effective during the first stages of oligomer formation. These early events in the aggregation cascade are poorly understood. Existing litera-

ture suggests that flavonoids and related compounds impede A β aggregation [159], providing the motivation for the present study.

In control simulations of untreated A β monomers, both A β_{40} and A β_{42} lost much of their initial α -helicity in favor of random coil and β -strand structures. The resulting structures collapsed over time to form configurations with radii of gyration that agree well with experimental findings [370] and previous simulation results [319, 371, 372, 373]. As a result, though each of the simulation sets for monomeric A β had only six members, we believe the configurations observed throughout the trajectories to be reasonable members of a larger conformational ensemble that A β samples *in vitro* and *in vivo*. Exhaustive sampling of structurally dynamic proteins using atomistic simulations is a difficult task given current computational limitations, though our simulations are extensive, representing 900 ns of accumulated simulation time in each simulation set of A β monomers.

Binding of morin to monomeric A β_{40} and A β_{42} principally affected the tertiary structure of the peptides, diminishing their ability to form hydrophobic contacts and thus impeding the collapse of the structure from its initial α -helical, extended configuration to a more compact, disordered structure. In the context of A β_{40} , the binding of morin to the peptide significantly reduced hydrophobic contacts and broadened the minimum of the free energy surface defined by these contacts and the radius of gyration (R_g) of the peptide. Administration of a large molar excess of morin (10:1) augmented these effects, leading to a significant expansion of the peptide structure, as manifested by the increase in R_g relative to the results of the control M40 simulations. The administration of 2:1 morin to the A β_{42} peptide had minimal effect on the structure of the peptide, indicating that it is more resilient to this treatment. The 10:1 molar excess of morin to A β_{42} caused significant secondary and tertiary structural effects. The free energy minima for A β_{42} defined by contacts and R_g showed dual energy minima, indicating two populations of energetically favorable configurations. Thus, we conclude that a greater level of therapeutic compounds would be necessary to inhibit the aggregation of A β_{42} , consistent with the results by Ono *et al.* in their fibril formation assays [160]. Specifically for morin, inhibiting A β_{42} fibril formation required nearly three times the concentration of morin than in the case of A β_{40} fibril formation. Since fibril formation proceeds through oligomeric intermediates [374, 375, 376], we expect the effects to be similar for these species.

In all of the simulations of A β dimerization, we observed that monomers that were initially separated in the simulation cell aggregated to form stable dimers. This behavior was observed regardless of the presence of morin, but important differences can be found in these simulations. It was also observed that the manner in which A β_{40} and A β_{42} dimerized was different. In the D40 simulations, A β_{40} principally associated via both polar and nonpolar (hydrophobic) contacts, forming many close interpeptide interactions between residues in the center of the peptide sequence. Specifically, contacts involving the CHC (L₁₇VFFA₂₁) and the flanking peptide sequences were common. In contrast, A β_{42} peptides were driven together by hydrophobic contacts involving C-terminal residues, leaving the more polar N-terminal region exposed to solvent. Bitan *et al.* previously concluded that the manner in which these two alloforms of A β aggregated was different, such that the addition of Ile41 and Ala42 to the C-terminus of A β (to give the longer and more hydrophobic A β_{42} alloform) led to greater stability of low-molecular weight aggregates that more rapidly gave rise to large, insoluble complexes [377]. Our simulation results predict that even at

the earliest stages of aggregation (dimer formation), the morphology and stabilizing interpeptide interactions between A β_{40} and A β_{42} are different.

Dimerization of both A β_{40} and A β_{42} led to a decrease in total helicity and an increase in β -strand content, relative to the monomeric state. In the M40 simulations, the A β_{40} peptides contained, on average, $20 \pm 12\%$ β -strand and $7 \pm 8\%$ total helix; in the dimeric state (D40 simulations), β -strand content increased slightly to $24 \pm 5\%$ and total helicity decreased to $4 \pm 4\%$. With respect to A β_{42} dimerization, β -strand content nearly doubled relative to the monomeric state ($13 \pm 5\%$ in the M42 simulations, $25 \pm 13\%$ in the D42 set), while total helicity was reduced from $15 \pm 4\%$ to $8 \pm 5\%$. While not in absolute quantitative agreement, the helical and β -strand content observed in the D40 and D42 simulations are in good agreement with the results obtained by Bitan *et al.* regarding secondary structure of A β in the dimeric state [377]. Further, Kim and Hecht [378] determined that A β_{42} has enhanced amyloidogenicity relative to A β_{40} due to the added hydrophobicity of residues Ile41 and Ala42 and their propensity to form β -strand structures. In our simulations, D42 dimers were driven together by C-terminal hydrophobic contacts. These residues showed a greater frequency of forming β -strands (in excess of 60% at Ile41, Figure 9.12D) than the C-terminal residues of A β_{40} (approximately 40%, Figure 9.9D).

Treatment of A β_{40} and A β_{42} with morin resulted in both interfacial and surface binding of morin. Interfacial binding was classified as the presence of two or more morin molecules at the dimerization interface between the two peptides. Surface binding indicated that morin bound principally on the outer surface of the dimer, such that morin was scattered over the surface. Neither of these binding modes significantly affected the secondary structure content of A β_{40} dimers, but the extent of interpeptide interaction was very different in the case of interfacial binding. Specifically, hydrophobic contacts between the two A β_{40} peptides were significantly reduced. In contrast, surface binding to A β_{40} led to an insignificant decrease in interpeptide contacts and dimer structures that were comparable to controls in terms of their level of compactness. Neither binding mode significantly altered the secondary structure content of the peptides.

Binding of morin to A β_{42} dimers produced many of the same effects as in the simulations of A β_{40} , with several notable differences. While interfacial binding did affect quaternary interactions between the two peptides (heavy atom and hydrophobic contacts), the reduction in contacts was considerable, but not statistically significant, as was the case in the D40_M simulations in which morin bound interfacially. Hydrogen bonding, however, was significantly reduced between the peptides with morin bound in this manner. When morin bound to the surface of A β_{42} dimers, β -strand content was reduced significantly, with the effect being greatest in the CHC. Binding of morin to polar residues flanking this sequence caused a considerable reduction in β -strand probability, concomitant with an increase in helicity.

Simulations of preformed A β dimers (both fibril- and oligomer-derived species) showed that the application of morin had little effect on the structural stability of these structures. That is, once formed, our simulations predict that A β dimers are very resilient when challenged with morin. Simulations of A β_{40} and A β_{42} dimers produced structures principally composed of β -strand and random coil structures, and their overall size (as measured by R_g) was comparable to experimental

results obtained by Walsh *et al.* [369].

As was the case with the $D40_M$ and $D42_M$ simulations, binding of morin to different regions of the $A\beta_{40}$ and $A\beta_{42}$ peptides had different effects on the preformed dimer structures. Binding to polar residues towards the N-terminus of preformed dimers had little effect on the stability of the peptides, but these interactions did serve to attract morin to $A\beta$ and in many cases acted as a conduit by which morin moved along the peptide sequence towards C-terminal hydrophobic regions and the polar bend region, in which the Asp23-Lys28 salt bridge is present. Interactions between morin and hydrophobic regions of $A\beta_{40}$ and $A\beta_{42}$ showed the greatest level of destabilization in the dimers. We observed considerable twisting and flattening of β -sheets, in addition to backbone destabilization in the areas with which morin interacted, in all of the simulations of preformed dimers in the presence of morin ($P40_M$, $L42_M$, and $Y42_M$).

Our previous work [358] with the pentameric form of the L42 model (a so-called “protofibril”) in the presence and absence of morin indicated that morin interacted with the Asp23-Lys28 salt bridges in the pentamer, using this region to penetrate into the hydrophobic core of the protofibril and disrupt backbone interactions. While we see some binding of morin to this same region of the $A\beta_{42}$ peptide in the $L42_M$ simulations, the destabilizing effect is not observed to the same extent. This apparent discrepancy can be largely attributed to the reduced stability of the Asp23-Lys28 salt bridges in the L42 dimers. Interpeptide salt bridges that were very stable in the protofibril were unstable in the dimeric form of the L42 structure due to the greater solvent exposure of these groups. Without a rigid binding site, morin interacted more weakly with the L42 dimer than with the protofibril form of L42, causing only transient destabilization of backbone hydrogen bonding in the $L42_M$ simulations. It has been proposed that Asp23-Lys28 salt bridges are less stable in oligomeric forms of $A\beta$ than in the mature protofibril or fibril forms [204], in which case the morin-susceptible region in the L42 structure may be absent in these structures, resulting in a weaker affinity between morin and $A\beta_{42}$ and binding occurring in multiple locations. The addition of morin to the dimeric structures led to twisting and flattening of the β -sheet structures that was largely absent in the protofibril structures. Instead, in the larger, more rigid pentameric protofibrils, a monomer of $A\beta$ partially dissociated from the core of the pentamer. The size and inherent rigidity of the protofibril led to the disruption of the Asp23-Lys28 salt bridge and backbone hydrogen bonding.

Recent work by Ladiwala *et al.* [379] concluded that the polyphenol myricetin, which is structurally very similar to morin, did not prevent the aggregation of $A\beta$, but instead was capable of remodeling prefibrillar oligomers and fibrils into unstructured, insoluble aggregates. They concluded that condensation of $A\beta$ monomers was indeed possible, even in the presence of such molecules, and our results agree with those findings. Though we found that both $A\beta_{40}$ and $A\beta_{42}$ formed dimers from initially separated monomers, we have described many differences between the control (untreated) structures and those that formed in the presence of morin. Ladiwala *et al.* concluded that myricetin and related compounds promoted nonspecific interactions between $A\beta$ monomers, a theory that is directly supported by the results obtained in the present work. Further, Ladiwala *et al.* predicted that myricetin could remodel existing soluble oligomers into disordered aggregates that would not lead to fibril formation. The structural deformations (overall structural

twisting and flattening of β -strands) seen in our preformed dimer simulations in the presence of morin, though subtle, represent possible outcomes that would divert oligomeric species from becoming fibrils or other higher-order aggregated structures.

The results we present here are also compatible with recent work by Convertino *et al.* [380], who used MD simulations to determine that the intrinsic disorder of the $A\beta(12-28)$ fragment precluded any shared binding mode among diverse compounds containing aromatic groups. Local effects on the structure of the peptide sequence were largely dependent upon specific interactions that arose between $A\beta$ and the library of test compounds. The secondary structure propensities we have determined in the present work for full-length $A\beta_{40}$ and $A\beta_{42}$ (Figures 9.9 and 9.12) agree well with their findings with respect to the structure of the peptide fragment in their study. Notably, we found that binding of morin near residues of the CHC caused a marked decrease in the appearance of β -strand structures in the $D42_M$ simulations (Figure 9.12); thus morin induced its effect via long-range interactions. Morin did not bind directly to the CHC sequence, but rather to residues flanking it. These results agree well with the prediction by Convertino *et al.* that $A\beta$ aggregation inhibitors may act at indirect or allosteric sites to influence peptide structure.

9.5 Conclusions

In the present work, we have used extensive explicit-solvent, atomistic MD simulations to describe the interactions of the flavonoid morin with monomeric and dimeric $A\beta_{40}$ and $A\beta_{42}$ in an attempt to understand the means by which this compound inhibits $A\beta$ aggregation. We found that binding of morin to monomeric $A\beta$ alters the free energy surface defined by tertiary interactions within the peptide, altering its ability to collapse and form a stable hydrophobic nucleus. Dimerization was not prevented by the administration of morin, but the structure of newly-formed $A\beta$ dimers was affected by the binding location of morin, either at the dimerization interface or the surface of the aggregate. Interpeptide contacts and secondary structural elements were differentially affected by the binding location of morin, with tertiary and quaternary interactions most consistently modulated by the presence of morin. In the context of preformed $A\beta$ dimers, application of morin had only minimal effects, which were again dependent on the location along the peptide sequence to which morin binds. The regions on the peptide that gave rise to the largest structural changes (in this case, tertiary and quaternary interactions) are the CHC and polar bend region. We thus conclude that, along with minor changes in secondary structure propensities, the main effect that morin has on the structure of $A\beta_{40}$ and $A\beta_{42}$ in the monomeric and dimeric states is to alter tertiary and quaternary interactions that give rise to structures that are measurably different from controls. These results provide molecular insight into the nature of so-called “off-pathway” aggregates that may have reduced toxicity compared to untreated, soluble $A\beta$ aggregates, the principal toxic entity in Alzheimer’s disease.

9.6 Acknowledgments

The authors thank Dr. Robert Tycko for providing the coordinates of A β ₄₀ fibrils, Dr. Liping Yu for providing the coordinates of the soluble A β ₄₂ dimer, and Advanced Research Computing at Virginia Tech for computing time on the SystemX and Athena supercomputers.

Chapter 10

GridMAT-MD: A Grid-based Membrane Analysis Tool for Use With Molecular Dynamics

Copyright statement: The contents of this chapter are reprinted with permission from W. J. Allen, J. A. Lemkul, and D. R. Bevan (2009) “GridMAT-MD: A Grid-based Membrane Analysis Tool for Use With Molecular Dynamics.” *J. Comput. Chem.* **30** (12): 1952-1958.

<http://dx.doi.org/10.1002/jcc.21172>

Attribution: W. J. A. conceived of the algorithm for, and wrote, the GridMAT-MD program. J. A. L. conducted validation systems and analysis. W. J. A., J. A. L., and D. R. B. wrote the paper.

10.1 Introduction

Molecular dynamics (MD) simulations of membranes and membrane proteins have proven useful to investigate structure-function relationships, and they provide researchers a unique perspective at the atomic level, one that cannot be easily seen experimentally [381, 382, 383, 384, 385]. Converting three-dimensional coordinate files, however, to useful figures for analysis and publication can be challenging, so concise methods of visualization are needed. GridMAT-MD is a free, open source program that has been developed to conduct the analysis of two important lipid bilayer parameters: thickness and area per lipid headgroup.

Three-dimensional coordinate files of membranes and membrane proteins can be rendered such that they produce detailed figures, but comparison among multiple images is time-consuming and difficult. A two-dimensional rendering of a three-dimensional structure also has its pitfalls. For example, trends in the thickness of the bilayer can be hard to distinguish, especially near the core of the bilayer where the view is obstructed by other lipid atoms. The area per lipid head group is

simple enough to calculate for a pure-lipid model system. It is the lateral surface area of the lipid bilayer divided by the number of lipids in a given leaflet. However, this value can be difficult to obtain when there is a protein present. The lateral area of ideal proteins such as helices or barrels can be estimated, but irregularly shaped proteins are a challenge.

GridMAT-MD was designed to facilitate analysis of lipid bilayer thickness and area per lipid headgroup. On characterizing bilayer thickness, the program generates the data for a top-down perspective of the lipid bilayer, which can be colored according to the thickness of the bilayer using an external viewer program, much like a contour map. From this view, it is easier to examine fluctuations in bilayer dimensions across the unit cell. GridMAT-MD also has the ability to not only calculate the area per lipid headgroup when proteins are present, but it will also report the specific lateral area for each lipid from a polygon-based tessellation, and the lateral area for the protein.

GridMAT-MD is designed to readily interface with coordinate files generated as part of the output of an MD simulation conducted with GROMACS [179, 180]. Current GROMACS analysis tools for lipid bilayers include those that analyze deuterium order parameters, center of mass distances, and densities, all of which are commonly used in the literature when analyzing the results of membrane simulations. The limitation in reporting center of mass distances, averaged over time, is that the calculation is often applied to a large group of atoms, such as phosphorus, limiting the level of detail that can be obtained through this analysis.

A method for analyzing bilayer thickness exists and has been successfully applied to large, coarse-grain membrane protein systems [386]. The method relies on a fitting algorithm and coordinate extraction and processing from external software, working well for the large system in that study. However, its applicability to smaller systems is limited (X. Periole and T. Huber, personal communication). Thus, we sought to design a generally-applicable tool to analyze bilayer thickness.

Some investigators have rightly noted that it is difficult to calculate the explicit lateral area of irregularly shaped proteins in a membrane [387, 388]. An essential tool that is still missing in bilayer analysis is one that can efficiently calculate the area per lipid headgroup, even in the presence of an embedded protein. To our knowledge, no such tool is freely and generally available. Existing methods for conducting such calculations have been reported on various online forums [389], but often require extensive scripting knowledge and commercial software.

GridMAT-MD is a superior tool for membrane analysis because (*i*) it is free to use and modify under the terms of the GNU public license [390], (*ii*) no installation is required, provided the Perl interpreter [391] is installed, (*iii*) it calculates an explicit area for proteins, making no assumptions of “ideal proteins,” and (*iv*) it is very fast to run and easy to operate.

10.2 Description

Only two files are needed as input for this program: a GROMACS-format coordinate file (.gro or .pdb) that contains a lipid bilayer that is oriented normal to the z -axis, and a parameter file that

contains user-specified information about the coordinate file, desired output, and other parameters, an example of which is distributed with the program. A user's guide is also distributed with the program that provides clear definitions of all of the options in the parameter file, as well as detailed descriptions on how to run specific analyses. In this section, the logic of the program and the type of data it generates will be discussed.

10.2.1 Processing the coordinate file

A flowchart of the GridMAT-MD algorithm is shown in Figure 10.1. Initially, GridMAT-MD parses through the GROMACS coordinate file in order to separate all of its components. The lipid and solvent atoms are stored in separate arrays for later processing. The ions are deleted, and anything remaining is treated as a "protein" atom, which can more broadly be interpreted as a protein, small molecule, or other macromolecule. The first two lines in the GROMACS coordinate file (the title and the number of atoms) are discarded, but the last line (box vectors) is stored for later use.

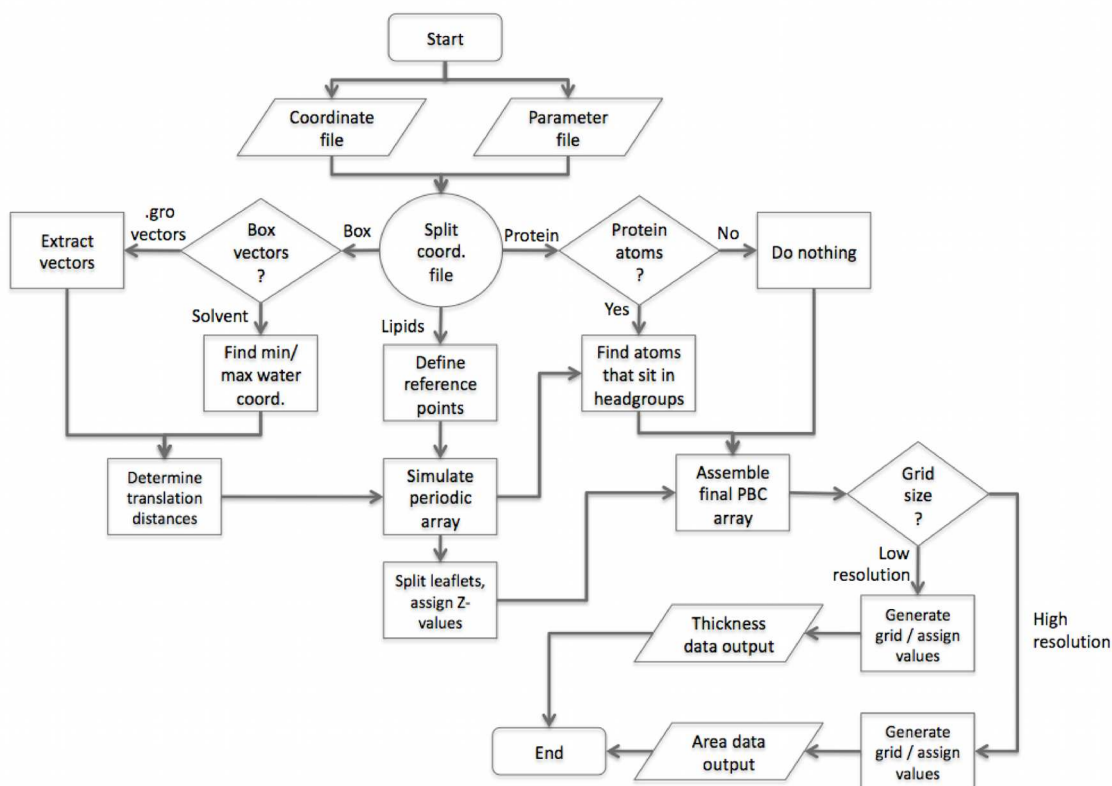


Figure 10.1: Flowchart of the GridMAT-MD logic structure.

Not all of the lipid atoms are needed to determine the thickness of the membrane, so most can be deleted. Those that the user selects are preserved, and can range from one atom per lipid (such as

phosphorus), to several (such as the glycerol backbone). In the case of a user choosing multiple atoms, the geometric center of mass is determined and the original coordinates are discarded. Ultimately, the program stores one point in space per lipid (a “reference point”), representative of the lipid to which it belongs. After that, the reference points are further divided into an upper and a lower leaflet.

The next step generates a 3×3 periodic array in the x - and y -directions of the reference points. The goal of this step is to have a central image for analysis that is immediately surrounded by eight identical images, since periodic boundary conditions are routinely applied in almost all MD simulations. The unit-cell translation distances for generating such an array can come from one of two sources, the upper and lower x - and y -limits of the box of solvent, or the last line of the GROMACS coordinate file - the box vectors. Since post-processing of coordinates may be necessary to account for effects such as periodicity, centering, and structural fitting, among others, sometimes the box vectors printed to the coordinate file may not reflect the actual dimensions of the system. Thus, we implemented a feature that allows the user to choose how the system size should be determined - from the printed box vectors, or from the dimensions of the solvent, which are often representative of a rectangular unit cell.

Once a periodic array of nine images is built, the next step is to determine which protein atoms are contained within the lateral area of the lipid headgroups (Figure 10.2). For each protein atom, the program first finds all lipid reference points within a user-defined radius of that protein atom on the $x - y$ plane. If at least one of those reference points has a greater z -coordinate value and one has a lower z -coordinate value than the protein atom, that protein atom is said to fall within the lipid headgroups. This process is repeated for each protein atom, and the total array of those protein atoms that fall within the lipid headgroups is also translated into a 3×3 periodic array in the x - and y -directions.

10.2.2 Calculating thickness

The most important concept in calculating the thickness of the bilayer is the Z-value. Each reference point in the top and the bottom leaflet is assigned a Z-value that represents the thickness (in nanometers) of the bilayer at that point. In order to determine the Z-value, the program considers one reference point at a time, starting in the top leaflet. It will search for the nearest neighbor of that reference point in the opposite leaflet, in just the x - and y -directions. Then it will take the difference of the z -coordinates between the two reference points, and that value becomes the Z-value. A Z-value is calculated for every reference point in the top leaflet, then this process is repeated for every point in the bottom leaflet. The user defines an arbitrary Z-value for the protein atoms. The goal here is not to try to represent the thickness of the bilayer with the protein atoms, but just to set them apart. So if most of the bilayer will have thickness values ranging from 3 to 4 nm, the Z-value of the protein might be set to 10. This way, protein atoms stand out in contrast against the lipid reference points.

Once the Z-values have been calculated, a grid is aligned to each leaflet in the center image of

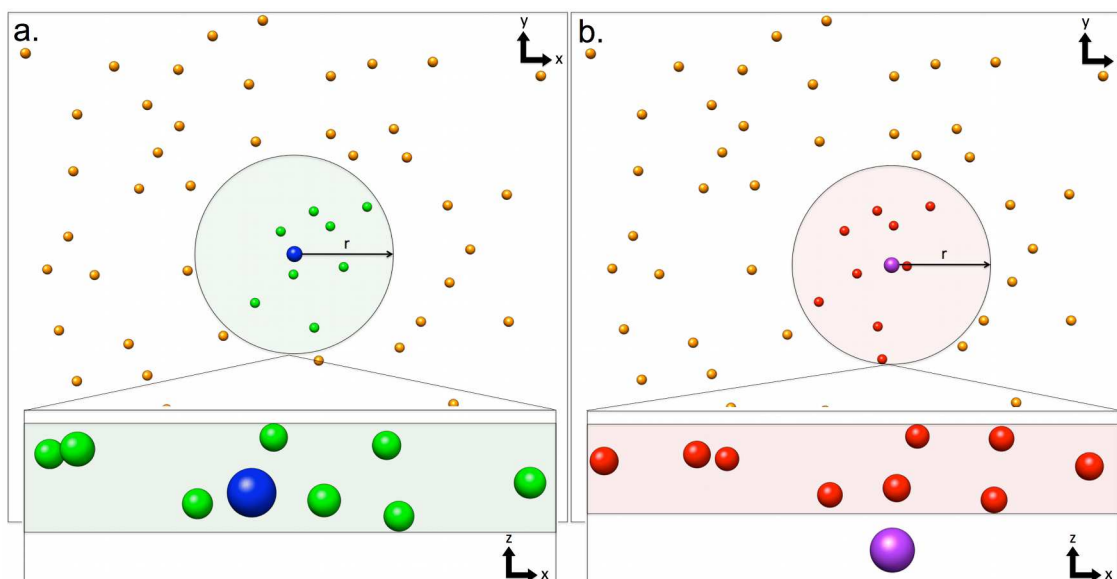


Figure 10.2: Diagram illustrating GridMAT-MD decision structure in choosing protein atoms, image rendered using USCF Chimera [392]. (a.) A top-down view of the top leaflet. The phosphorous atoms are the reference points and are shown as orange spheres. A protein atom is depicted as a blue sphere. Within a given radius, 8 reference points are considered. (insert) A side view of the top leaflet reference atoms. The protein atom falls within the upper and lower z -coordinates of the reference atoms, so its coordinates are conserved. (b.) The same view with a different protein atom depicted as a purple sphere. (insert) From the side view, one can see that the protein atom does not fall within the lipid headgroups. Its coordinates are discarded.

the periodic array. The resolution and shape of the grid are also user-defined. Each grid point in the top grid then searches for the nearest reference atom in the top leaflet, in just the x - and y -directions. Once found, that grid point is assigned the Z -value from that reference atom. This process is repeated for the bottom grid in the bottom leaflet. The periodicity of the bilayer allows reference atoms on one side of the system to be “seen” by grid points on the opposite side of the system, as would be the case in a molecular dynamics simulation using true periodic boundary conditions. The results of this process are two similar matrices of data: the thickness of the bilayer from the top leaflet down, and the thickness of the bilayer from the bottom leaflet up. Each matrix can be visualized using an external viewer by assigning a color to each grid point depending on the magnitude of the Z -value, easily distinguishing thin and thick parts of the lipid bilayer (Figure 10.3). Visually, the top and bottom matrices typically look very similar with some minor differences, so the program averages the data in order to give the best representation. For typical thickness calculations, we have generated ideal images using grids with points spaced every 3-4 Å. This corresponds to a grid of about 20×20 points for a bilayer system with 50-60 lipids per leaflet.

As described, GridMAT-MD will provide an image of bilayer thickness for a single coordinate

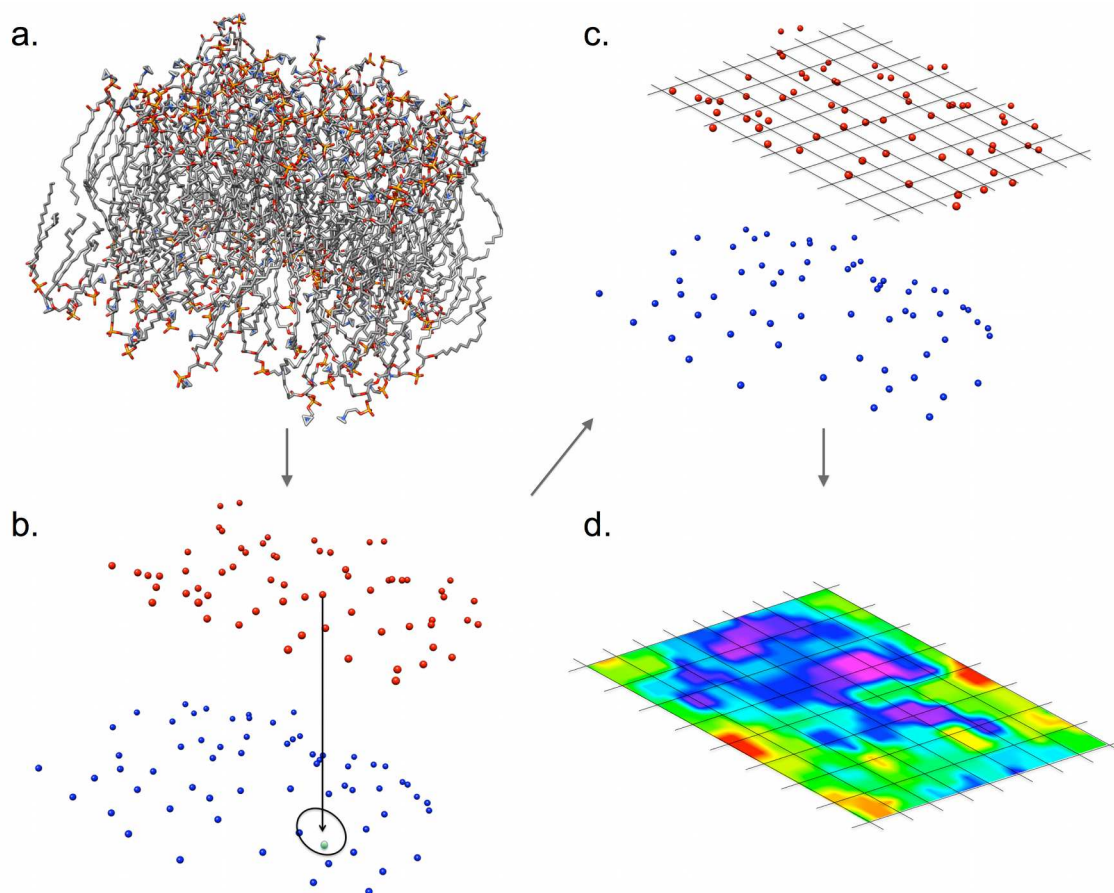


Figure 10.3: Schematic diagram of the extraction of thickness from a lipid bilayer. (a.) A rendering of a 128-lipid POPC [228] bilayer created with USCF Chimera [392]. (b.) The phosphorus atoms are chosen as reference points and each searches for its nearest neighbor in the opposite leaflet. The distance of separation becomes the Z-value of that reference point. (c.) Once all Z-values are assigned, a grid is overlaid to the top leaflet. For clarity, the grid assigned to the bottom leaflet is not shown. (d.) Every vertex in the grid is assigned the Z-value of the nearest reference point, and is colored based on the magnitude of the value.

file, however one snapshot from a long MD does not provide an accurate representation of the fluctuations of the lipid bilayer over time. A better approach is to take a snapshot every few timesteps, and average the results. The raw data format printed by GridMAT-MD makes further manipulation with scripts or spreadsheet software very straightforward.

10.2.3 Calculating area per lipid headgroup

In developing GridMAT-MD, we observed that as the resolution of the grid is increased to about 200×200 grid points, the membrane system could be divided into polygons. Further, the number of polygons and lipid molecules always occurred in a one-to-one ratio. Upon generating the polygons, calculating not only the area per lipid headgroup, but also the lateral area of a specific lipid, becomes a simple task. All of the grid points within a certain polygon are assigned the same Z-value, and correspond to the same reference point. Subsequently, the number of grid points within a given polygon, divided by the total number of grid points, is equivalent to the area of a specific lipid headgroup as a fraction of the total lateral area of the bilayer system. Further, this reasoning extends to the protein atoms that are assigned an arbitrary Z-value by the user. The protein atoms compete for grid points with the lipid reference points, so a simple count of the points with the user-defined arbitrary Z-value will yield the lateral surface area that the protein occupies in the lipid bilayer. We believe that this is a much more accurate way to estimate the lateral area of a membrane protein, rather than making assumptions or estimations for “ideal proteins,” based on ideal geometry.

In calculating bilayer thickness, it is useful to average the matrix data from the top and bottom leaflets as previously mentioned. However, in calculating area per lipid headgroup, it may only be meaningful to look at one leaflet at a time, especially in the case of an asymmetrically-oriented or surface-adsorbed peptide, protein, or other molecule. GridMAT-MD was designed in such a way that the data for the top and the bottom leaflets can be printed separately. The data output from area per lipid calculations is a generic text file containing the area of each individual lipid, printed with the lipid residue number to which it corresponds. Then it is quite easy for the user to look at trends in area for a specific lipid (or set of lipids, for example, that are within a certain distance from a protein) over the span of a trajectory.

10.3 Validation Simulations

10.3.1 Methods

To test the ability of GridMAT-MD to reproduce previous reports of bilayer thickness, we replicated the work of Kandasamy and Larson by reconstructing one of their systems, the KALP₁₅ peptide in a DPPC membrane. This system (identified as L1615-16 in their work) served the original authors as part of an examination of hydrophobic mismatch [229].

The KALP₁₅ peptide was built within the xLeap module of AmberTools, version 1.0 [393], imposing the backbone geometry of an ideal α -helix ($\phi = -60^\circ$, $\psi = -40^\circ$) on the constituent residues. The N- and C-termini were capped with acetyl and amide groups, respectively, to give uncharged termini.

The peptide-bilayer system was prepared based on the procedure originally described by Kan-

dasamy and Larson, with some small differences. The previous work employed the methodology proposed by Faraldo-Gómez *et al.* [394] for inserting the peptide into the membrane. We chose to use a more recent “InflateGRO” methodology proposed by Kandt *et al.* [267], as it does not rely on multiple programs or modifications to the GROMACS source code. By using the InflateGRO method, 4 lipids were deleted from the bilayer (2 from each leaflet), leaving 124 DPPC molecules packed around the KALP₁₅ peptide.

Parameters from the GROMOS96 53A6 parameter set were applied to the peptide, water, and ions in the system. Lipids were described by parameters derived by Berger *et al.* [266]. The equilibration procedures used were the same as those described in the original work. Briefly, an equilibration phase of 5 ns duration was conducted under an *NPT* ensemble, using the Berendsen weak coupling method for both temperature and pressure [187]. The temperature of the peptide, membrane, and solvent and ions were maintained at 323 K separately using a coupling constant of 0.1 ps. The pressure was regulated semi-isotropically at 1 bar in all directions, using a coupling constant of 1.0 ps. Position restraints were applied to the peptide backbone during the equilibration phase. After equilibration, restraints were removed from the peptide and the production phase began. The treatment of nonbonded calculations was identical to that of the original work, cutting off short-range Coulombic and van der Waals interactions at 1.2 nm, and calculating long-range electrostatics with the particle mesh Ewald (PME) method [193, 194]. Production simulations were allowed to proceed for 50 ns. All simulations were conducted with the GROMACS package, version 3.3.3, using 16 CPUs of Virginia Tech’s SystemX supercomputer [395].

10.3.2 Results

With respect to the bilayer thickness in the vicinity of the KALP₁₅ peptide, our results are in good agreement with those previously reported [229]. The hydrophobic region of the peptide is shorter than the transmembrane dimension of the bilayer, a situation called negative hydrophobic mismatch. As a result, the authors of the original work reported a total depression of bilayer dimensions of 15 Å from data collected over the last 25 ns of 50-ns simulation. Over this same period, we note a depression of 13 Å in the vicinity of the KALP₁₅ peptide. Figure 10.4 illustrates the capabilities of GridMAT-MD to display these data concisely. Snapshots are shown from configurations at the start and end of the trajectory, as well as the average over the last 25 ns of the trajectory. The peptide is shown in each panel for reference regarding its location and conformation. The average peptide structure over the last 25 ns, based on RMSD clustering analysis, is shown in the final panel. For comparison, a traditional rendering of the system (in its final configuration, at 50 ns) is shown in Figure 10.5.

GridMAT-MD produces bilayer thickness data very quickly. Each frame analyzed to produce the data in Figure 10.4 was processed in less than 5 seconds on a MacBook laptop (2.4 GHz Intel Core 2 Duo, 2 GB RAM).

GridMAT-MD also provides a concise representation of the lipid molecules in the system. Each lipid is assigned to a polygon, as discussed above. Figure 10.6 shows a representation of this

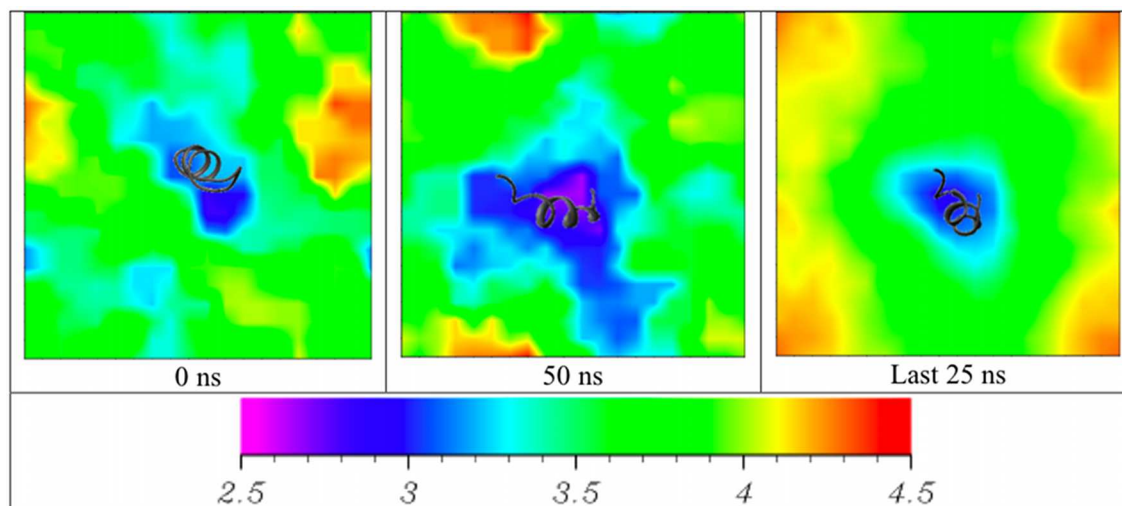


Figure 10.4: Snapshots from the KALP₁₅-DPPC trajectory. The legend shows the bilayer thickness values mapped to a rainbow color gradient, in units of nm. Data were visualized and images rendered using the Xmatrix viewer (version 2.8.0), distributed as part of the Matpack C++ Numerics and Graphics Library [396]. The peptide was rendered using UCSF Chimera [392].

output, also from the final configuration of the KALP₁₅-DPPC system. Over the last 25 ns of the 50-ns trajectory, the average area per lipid headgroup for the top and bottom leaflets of the DPPC bilayer was $62.2 \pm 1.3 \text{ \AA}^2$ and $61.8 \pm 0.6 \text{ \AA}^2$, respectively. Our own simulations of pure DPPC bilayers under the same conditions give an average area per lipid headgroup of $63.0 \pm 0.9 \text{ \AA}^2$ (Chapter 4), in agreement with previous simulation [266, 397] and experimental results [250]. Area per lipid calculations for the KALP₁₅-DPPC system were also completed quickly by GridMAT-MD, with each frame processed in, on average, 2.5 minutes using the same hardware described above.

10.4 Summary

In summary, GridMAT-MD is capable of fast calculations of membrane thickness and area per lipid headgroup with no installation necessary, making it accessible to any of the major operating systems that provide a Perl interpreter. The resulting contour plots are a convenient way to represent the thickness of a lipid bilayer, and the area per lipid calculations are the most convenient and most rigorous that we believe to be publicly available. GridMAT-MD is available for download at <http://www.bevanlab.biochem.vt.edu/GridMAT-MD>.

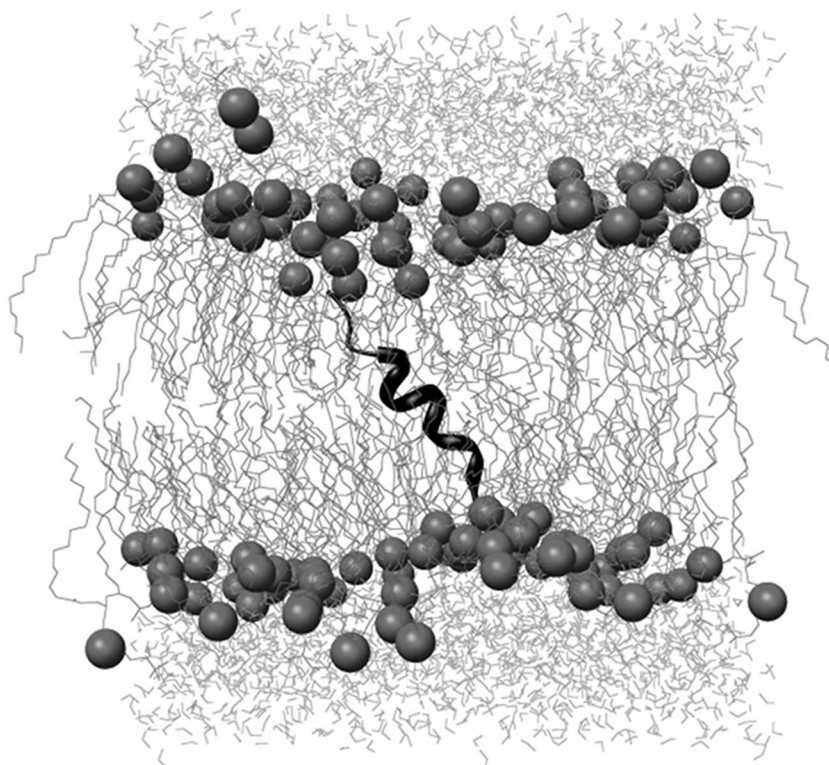


Figure 10.5: The KALP₁₅-DPPC system, at 50 ns, in a traditional rendering. The peptide is shown as a ribbon, phosphorus atoms as spheres, and lipid chains and water shown as lines. The image was rendered using UCSF Chimera [392].

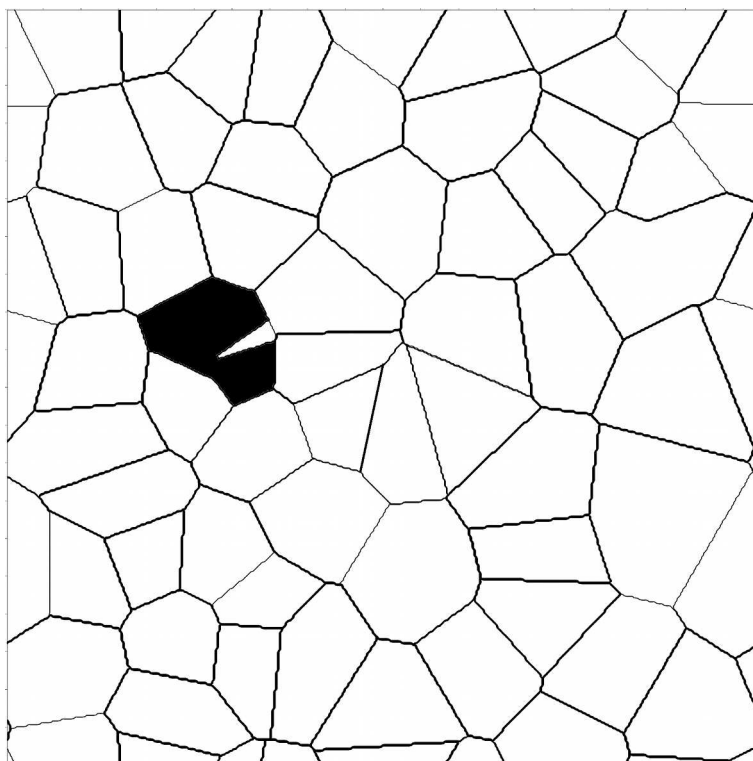


Figure 10.6: Representation of area per lipid headgroup in the top leaflet, from the final configuration of the system. Each polygon represents an individual lipid, determined using the headgroup phosphorus as a reference point. The dark shaded region corresponds to the lateral area of the KALP₁₅ peptide that extends into the lipid headgroup region. Image rendered using the Xmatrix viewer (version 2.8.0), distributed as part of the Matpack C++ Numerics and Graphics Library [396].

Chapter 11

Practical Considerations for Building GROMOS-Compatible Small Molecule Topologies

Copyright statement: The contents of this chapter are reprinted with permission from J. A. Lemkul, W. J. Allen, and D. R. Bevan (2010) “Practical Considerations for Building GROMOS-Compatible Small Molecule Topologies.” *J. Chem. Inf. Model.* **50** (12): 2221-2235. <https://acs.manuscriptcentral.com:443/AOR-7PSbSVz7MqyDKpuGwM7D> Copyright (2010) American Chemical Society.

Attribution: J. A. L., W. J. A., and D. R. B. designed the experiments. W. J. A. performed simulations of pure liquid and gas systems, the UGM-FAD complex, and conducted the Spartan '04 calculations. J. A. L. performed simulations of amino acid systems and conducted the Antechamber calculations. J. A. L., W. J. A., and D. R. B. wrote the paper.

11.1 Introduction

Preparation of small molecule topologies for use in molecular dynamics (MD) simulations is of critical importance in studies of drug-enzyme and drug-receptor interactions, partitioning of drugs into and across membranes, among others. The principal challenge in including small molecules in these simulations is the effort necessary to produce a topology that is consistent with the underlying theory of the force field, which is often based on common macromolecules such as proteins. Thus, less common functional groups present in drug molecules and synthetic compounds are difficult to parameterize. Different force fields frequently used for biomolecular simulation, such as AMBER [169, 170, 171, 186], CHARMM [172, 173, 398], OPLS [167, 168], and GROMOS [174, 182, 184, 399, 400], have different functional forms and underlying assumptions, requiring different parameterization procedures and validation protocols. While parameterization of

small molecules for AMBER and CHARMM force fields has been aided by the development of the general AMBER force field (GAFF) [401] and the CHARMM general force field [402], topology generation for OPLS and GROMOS remains more challenging. Procedures involving quantum mechanical calculations and geometry optimization are thoroughly described for the OPLS-AA parameter set [167, 168], but such detailed information is not publicly available for the GROMOS96 parameter sets. Thus, investigators seeking to use the GROMOS96 force fields must rely on empirical parameter assignment and subsequent validation through the time-consuming process of thermodynamic integration [174], a challenge that has long been recognized within the simulation field [267].

The GROMOS96 force fields offer an advantage over the AMBER, CHARMM, and OPLS parameter sets in that they are based on the united-atom approach to parameterization. That is, nonpolar hydrogens are not explicitly represented, decreasing the number of atoms in the system appreciably and speeding up simulations. These force fields are attractive for long-time simulations and large sets of simulations that may be utilized *in silico* to analyze the dynamics of small molecules and potential drugs.

Automated topology generation is possible through a number of tools, including the very popular PRODRG server [403, 404]. Topologies produced by this program have been used widely in studies of protein-peptide [405], protein-ligand [406, 407], protein-lipid [408], and small molecule-lipid interactions [409, 410], drug partitioning into lipid membranes [411], and simulations of small molecules that inhibit amyloid aggregation [359]. Careful inspection of the topologies produced by PRODRG reveals that the charges and charge groups assigned to the functional groups of small molecules are often inconsistent with the same groups present in the GROMOS96 43A1 parameter library. The effects of these inconsistencies are unknown, providing the motivation for the present study. Several methods for generating GROMOS-compatible charges for refining PRODRG topologies have been proposed in the literature [412, 413, 414, 415], but thus far, to the best of our knowledge, no standard has been set for the accuracy and applicability of any of these methods.

Here, we evaluate the quality of small-molecule topologies generated by the automated server PRODRG [403, 404], using small molecules whose functional groups are described by the GROMOS96 43A1 parameter set [182]. We explore a variety of systems, including pure liquids, amino acids at a hydrophobic-hydrophilic interface, and a cofactor bound to an enzyme. By demonstrating the implications of the inherent deficiencies in automated parameter generation and by making recommendations for proper parameter development, we hope to elaborate on the best practices in small molecule topology generation under the GROMOS force fields.

11.2 Methods

To assess the quality of small-molecule topologies in a range of scenarios, several systems were constructed: (1) condensed-phase and gas-phase systems of hexane, ethanol, and *p*-cresol, (2)

several amino acids (alanine, asparagine, aspartate, isoleucine, lysine, and serine) at the interface of water and cyclohexane, and (3) the protein UDP-galactopyranose mutase (UGM) with bound flavin adenine dinucleotide (FAD). All systems were prepared and simulations conducted using facilities present in the GROMACS package, version 4.0.7 [180]. One set of small molecule topologies was generated using the PRODRG 2.5 server [403, 404], which is designed to produce topologies consistent with the GROMOS96 43A1 parameter set. The other topologies were generated using known GROMOS96 43A1 [182] functional group charges and charge groups, as implemented in the GROMACS distribution. Bonded parameters and atom types for each molecule were kept the same between the different topologies, such that the comparisons involved only charges and charge groups assigned by PRODRG relative to those implemented in the GROMOS96 43A1 parameter set. Unless otherwise noted, all simulations employed a twin-range cutoff scheme for short-range nonbonded interactions, with the real-space contribution to Coulombic terms truncated at 0.8 nm and short-range van der Waals interactions truncated at 1.4 nm. All bond lengths were constrained using the LINCS method [270], allowing a 2-fs time step. The neighbor list was updated every 5 simulation steps (10 fs). Long-range electrostatic interactions were calculated with the smooth particle mesh Ewald (PME) method [193, 194]. Equilibration simulations utilized the Berendsen weak coupling method [187] to control temperature and/or pressure, while production simulations utilized the Nosé-Hoover thermostat [188, 189, 190] and Parrinello-Rahman barostat [191, 192] to generate a rigorous *NPT* ensemble.

11.2.1 Condensed-phase and gas-phase systems

To measure various thermodynamic and physical properties of pure liquid systems as a function of the molecular topology, systems containing hexane, ethanol, and *p*-cresol were constructed by replicating a single molecule of each compound in three dimensions to create a cubic grid of 512 molecules. For gas-phase systems, molecules were separated by 50 nm and kept in place with weak position restraints, in accordance with GROMOS parameterization methodology, and plain cutoffs were used to calculate electrostatic interactions [174]. Steepest descent minimization was performed on gas-phase and condensed-phase systems alike, followed by 100 ps of isochoric-isothermal (*NVT*) equilibration at 298 K. Further equilibration under an isothermal-isobaric (*NPT*) ensemble was performed for 300 ps at the same temperature and 1 bar of pressure. Finally, each system was simulated for another 20 ns production run under the same conditions. Charges and charge groups for these molecules are illustrated in Figure 11.1.

11.2.2 Interfacial amino acid systems

Amino acids represent some of the most common components of biomolecular force fields and thus should serve as reliable test systems. By using amino acids, we can assess the quality of PRODRG topologies against groups that have been rigorously defined. Six amino acids (alanine, asparagine, aspartate, isoleucine, lysine, and serine) were chosen to represent the twenty common

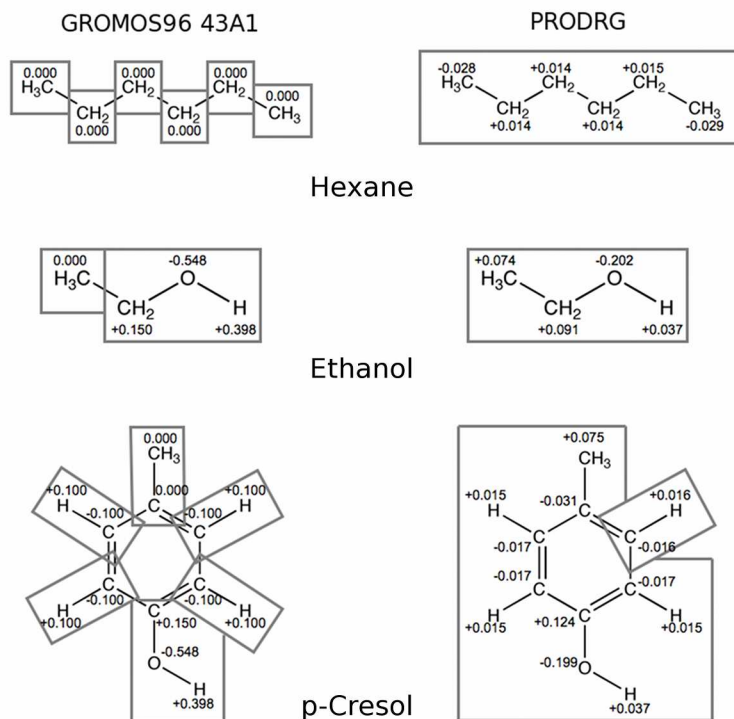


Figure 11.1: Comparison of GROMOS96 43A1 and PRODRG charges and charge groups for hexane (top), ethanol (middle), and *p*-cresol (bottom). Charges are given in units of e and charge groups are indicated by gray boxes.

amino acids. The chosen molecules encompass hydrophobic (Ala and Ile), polar uncharged (Asn and Ser), and charged (Asp and Lys) species. Each amino acid was modeled as a dipeptide, with N- and C-terminal groups capped with acetyl and N-methyl groups, respectively, to give uncharged termini. Since the goal of the GROMOS96 parameter sets is to accurately describe partitioning of amino acids between polar (water) and hydrophobic (cyclohexane) media [174], we sought to produce a biphasic system that would model partitioning behavior of the different amino acids.

A 5-nm cubic box of cyclohexane was generated, containing a total of 466 molecules. Standard bond, angle, and dihedral parameters from the GROMOS96 43A1 library were assigned in the cyclohexane topology. All atoms were represented as united-atom CH₂ with zero charge. Following steepest descent minimization, the cyclohexane box was equilibrated first under an isochoric-isothermal (*NVT*) ensemble for 100 ps (298 K), and then under an isothermal-isobaric (*NPT*) ensemble for 500 ps (298 K, 1 bar), at which point the box vectors and density had stabilized. The system was simulated for an additional 10 ns to generate the cyclohexane layer used in building the interfacial systems.

Each amino acid was placed at the center of a rectangular box such that both the plane of the amino acid backbone and the C α -C β bond were coincident with the $x - y$ plane of the unit cell

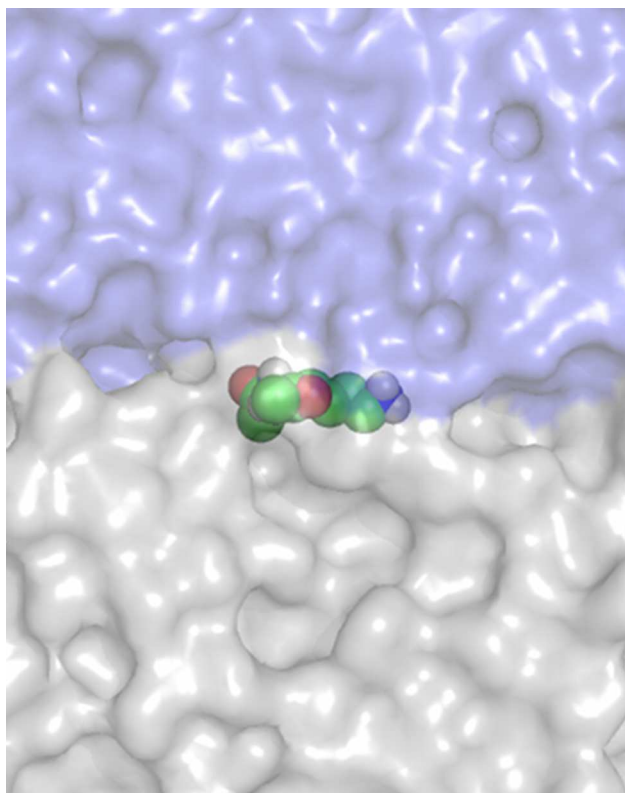


Figure 11.2: Initial position of lysine in the cyclohexane-water interfacial system. The amino acid backbone is perpendicular to the page and the sidechain is extending to the right in the image. Lysine is shown with its atoms colored by element and rendered as van der Waals spheres. The co-solvents are rendered as transparent surfaces, with water in blue and cyclohexane in gray. The image was rendered with PyMOL [272].

(Figure 11.2). The equilibrated cyclohexane box was placed in the lower half of the simulation cell, while the remainder of the volume was filled with SPC water [268]. A single Cl^- counterion was added to the aqueous phase of the lysine system to compensate for the net charge of the amino acid. In the case of aspartate, a single Na^+ ion was added. No ions were added to systems containing uncharged amino acids. Each system was energy-minimized and equilibrated in the same manner as the cyclohexane box discussed above, with position restraints applied to amino acid heavy atoms during equilibration. Three replicates of each system were generated from different random initial velocities. The data collection period lasted for 20 ns. Charges and charge groups used for the simulations of amino acids are given in Figures 11.3 and 11.4.

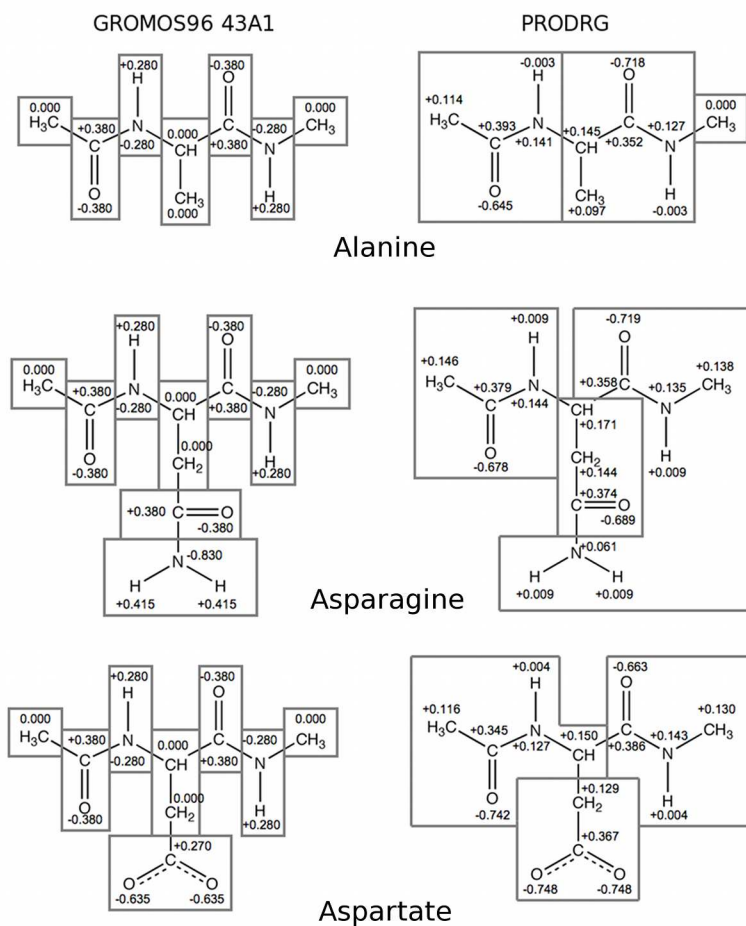


Figure 11.3: Comparison of GROMOS96 43A1 and PRODRG charges and charge groups for alanine (top), asparagine (middle), and aspartate (bottom). Charges are given in units of e and charge groups are indicated by gray boxes.

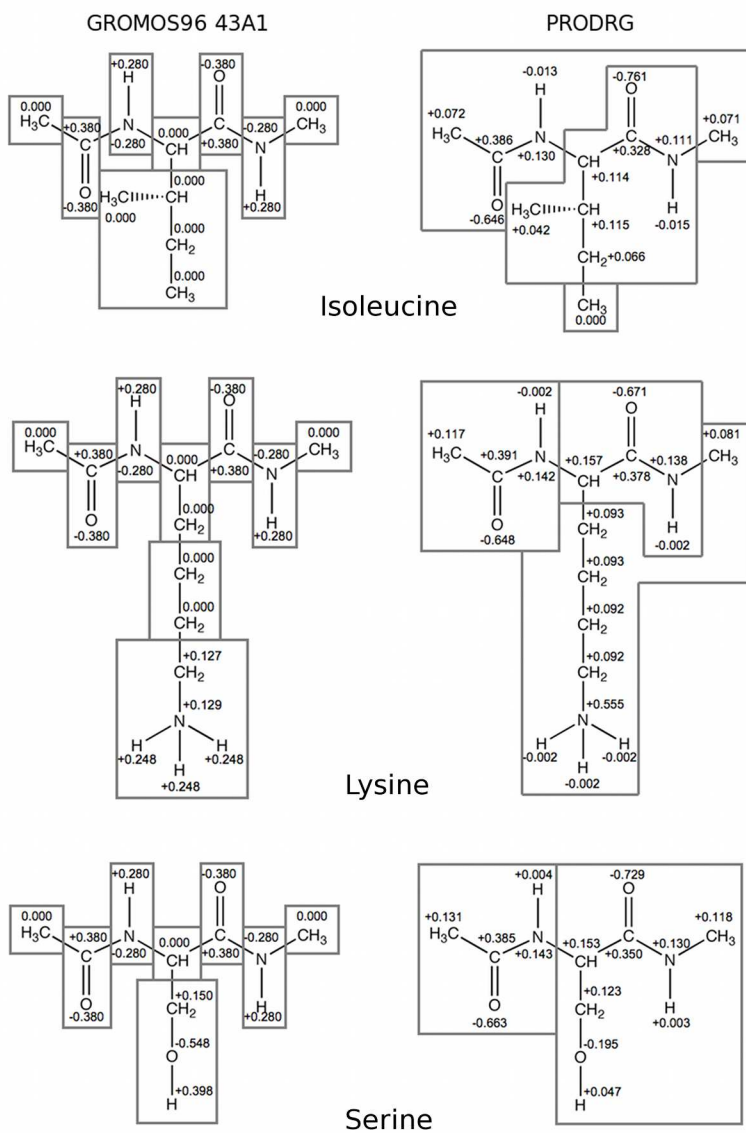


Figure 11.4: Comparison of GROMOS96 43A1 and PRODRG charges and charge groups for isoleucine (top), lysine (middle), and serine (bottom). Charges are given in units of e and charge groups are indicated by gray boxes.

11.2.3 UDP-galactopyranose mutase

To examine the stability of a cofactor bound to a protein, the structure of UDP-galactopyranose mutase (UGM) bound non-covalently with flavin adenine dinucleotide (FAD) was utilized. The starting structure for these simulations was taken from the crystal structure of the *K. pneumoniae* enzyme-cofactor complex determined by Beis *et al.* (PDB code 2BI7) [416], which contains an oxidized FAD cofactor. Parameters for the protein were assigned from the GROMOS96 43A1 parameter set, and FAD parameters were either taken from this same force field or generated by PRODRG (Figure 11.5). The protein-cofactor complex was centered in a rhombic dodecahedral unit cell that was subsequently filled with SPC water [268] and sufficient counter ions to balance the charge and bring the final ionic concentration to a near-physiological 100 mM. Equilibration procedures similar to the previous systems were implemented. Steepest descent minimization was performed, followed by 100 ps of isochoric-isothermal (*NVT*) equilibration. Next, 500 ps of equilibration under an isothermal-isobaric (*NPT*) ensemble were performed, followed by 40-ns production simulations. Position restraints were applied to the backbone of the protein during equilibration, and released for the production run. Different random initial velocities were used to generate independent trajectories.

11.3 Results and Discussion

The goals of our approach include the assessment of several physical and thermodynamic parameters in the liquid- and gas-phase systems, behavior of amino acids in the presence of hydrophobic and polar media, and the interactions of an enzyme-bound molecule with its partner protein. These scenarios represent some of the most common uses of small molecule topologies in the literature, and provide us with versatile systems in which to test the quality of these topologies. We provide an analysis of the various intermolecular interactions that give rise to the observed behavior, focusing primarily on electrostatic and van der Waals interactions, hydrogen bonding, and relative orientation of the different molecules. While these quantities are largely interdependent, a comprehensive analysis of these terms is appropriate as they relate directly to quantities that are often evaluated in MD simulations.

For all comparisons between the results of using PRODRG and GROMOS96 43A1 topologies, the behavior observed in the GROMOS96 43A1 simulations is considered the benchmark for all comparisons. Parameters for novel species must be derived in a manner consistent with the existing, validated force field, which is taken as the standard for accuracy, and thus we evaluate the results of using PRODRG parameters in this context.

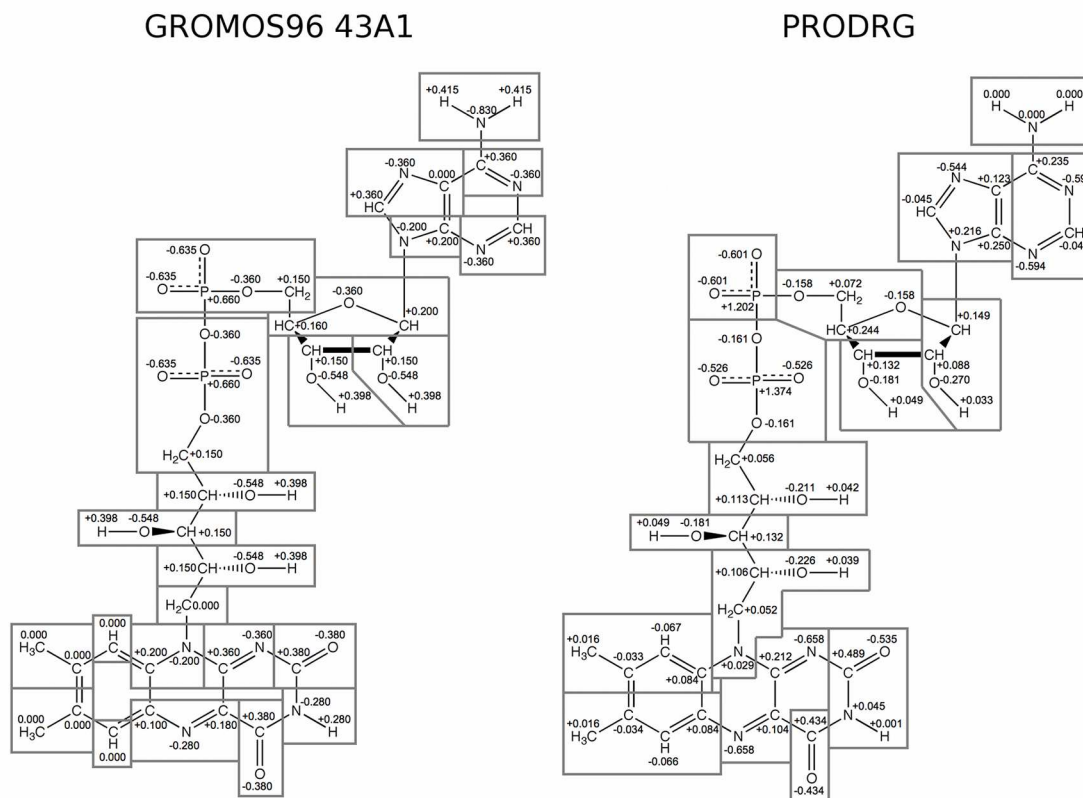


Figure 11.5: Comparison of GROMOS96 43A1 and PRODRG charges and charge groups for FAD. Charges are given in units of e and charge groups are indicated by gray boxes.

11.3.1 Condensed-phase and gas-phase systems

The parameterization of the GROMOS96 force field parameter sets relies on accurate reproduction of condensed-phase behavior and thermodynamic data. In the most recent edition of the force field (53A6) [174], the criteria used to assess the accuracy of the derived parameters included density and the heat of vaporization, ΔH_{vap} . These quantities can be calculated from MD simulations of hexane, ethanol, and *p*-cresol.

Table 11.1 summarizes the results of density and ΔH_{vap} calculations. The equation used to calculate ΔH_{vap} was:

$$\Delta H_{vap} = \langle E_{gas} \rangle - \langle E_{liquid} \rangle + RT \quad (11.1)$$

In Equation 11.1, the quantities in angle brackets denote the time average of the total energy of the gaseous and liquid systems. R is the gas constant, and T is the absolute temperature in Kelvin.

Table 11.1: Density and ΔH_{vap} for small molecules.

System	Density [g mL ⁻¹]			ΔH_{vap} [kJ mol ⁻¹]		
	43A1	PRODRG	Expt.	43A1	PRODRG	Expt.
Ethanol	0.766 ± 0.008	0.972 ± 0.007	0.789	20.3 ± 0.2	9.4 ± 0.4	42.3
Hexane	0.681 ± 0.005	0.679 ± 0.005	0.661	17.2 ± 0.3	17.2 ± 0.3	31.6
<i>p</i> -Cresol	1.097 ± 0.004	1.045 ± 0.005	1.019	48.7 ± 0.5	35 ± 2	73.1*

All experimental data taken from [417], except *[418].

We found relatively good agreement between the GROMOS96 43A1 topologies and experimental data with respect to density (Table 11.1). Densities were slightly overestimated for hexane and *p*-cresol (by 3.0% and 7.7%, respectively), and slightly underestimated for ethanol (by 3.0%). For PRODRG-produced topologies, the densities of hexane and *p*-cresol were in close agreement with both the GROMOS96 43A1 results and experimental data (overestimated by 2.7% for hexane and 2.6% for *p*-cresol), but that of ethanol was significantly overestimated (by 23.2%).

With respect to ΔH_{vap} values, neither GROMOS96 43A1 nor PRODRG parameters closely reproduced experimental data, with all simulation results producing values that were well below the experimentally-determined ΔH_{vap} . The values of ΔH_{vap} generated by GROMOS96 43A1 topologies, however, generally were in somewhat closer agreement with experimental values than were values generated using PRODRG parameters (Table 11.1).

To explain these results, we examined the structures of the liquids themselves in terms of both hydrogen bonding and radial distribution functions (RDF), which describe the density of chosen particles as a function of distance, a calculation that can be applied to individual atoms or to whole molecules. In an RDF plot, closely interacting groups are separated by shorter distances, such that the resulting density peak is higher, indicating a greater probability of interaction at a given distance.

The decreased value of ΔH_{vap} for ethanol and *p*-cresol resulting from PRODRG parameters can be attributed mainly to a significant decrease in hydrogen bonding capacity. For ethanol, the use of GROMOS96 43A1 parameters resulted in 0.93 ± 0.01 hydrogen bonds per molecule (averaged over time). For *p*-cresol, this value was 0.94 ± 0.01 . Use of PRODRG parameters gave substantially lower hydrogen bonding for these two species, 0.12 ± 0.02 hydrogen bonds per ethanol molecule and only 0.05 ± 0.01 per *p*-cresol molecule. The RDF of these two species (Figure 11.6) provides insight into the reason for this observed behavior.

PRODRG parameters caused ethanol molecules to align head-to-tail, that is, with the hydroxyl group associating most prominently with the methyl group of a nearby molecule, at a distance of approximately 0.3 nm (Figure 11.6A). Interactions between hydroxyl groups occurred at a much greater distance, nearly 0.5 nm, exceeding the optimal distance for hydrogen bonding. The combination of a relatively large charge (+0.071 *e*) on the terminal methyl group and the comparatively

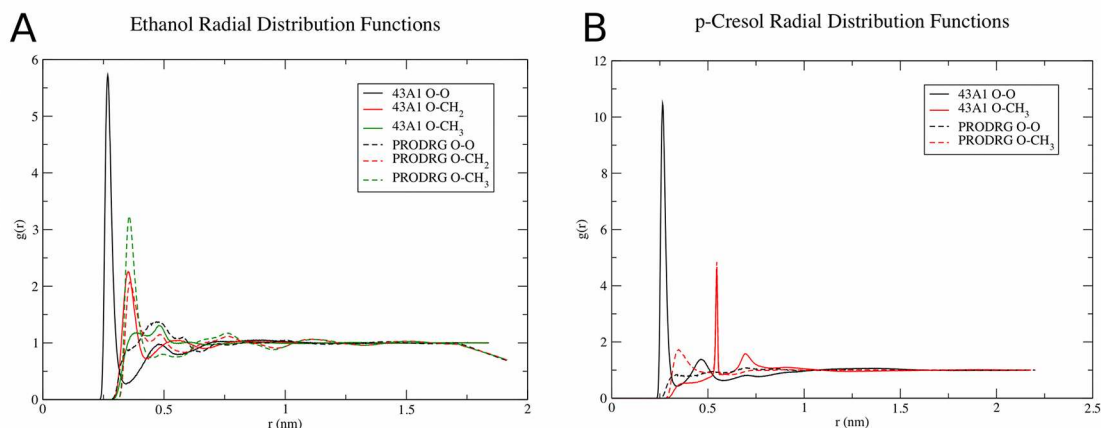


Figure 11.6: RDF for different atom pairs in (A) ethanol and (B) *p*-cresol. In (B), the intramolecular O-CH₃ peak referred to in the text appears as a single, solid line since it is present in both curves and is thus overlapping.

small charge on the hydroxyl hydrogen (+0.037 e) leads to this behavior. The association of the hydroxyl oxygen atoms with the terminal methyl groups disrupted hydrogen bonding, decreased ΔH_{vap} , and increased the density relative to the results obtained using GROMOS96 43A1 parameters. Proper hydrogen bonding was observed using GROMOS96 43A1 parameters for ethanol, with hydroxyl-hydroxyl interactions occurring at approximately 0.3 nm, indicating strong hydrogen bonding and comparing closely with the RDF of previous simulations using different parameters [419].

In the case of *p*-cresol, we also observed a reduction in hydrogen bonding when PRODRG parameters were applied, as well as an increase in hydroxyl-hydroxyl distance relative to the results obtained using GROMOS96 43A1 parameters (Figure 11.6B). Under PRODRG parameters, there was no distinct hydroxyl-hydroxyl RDF peak, but there was a probable interaction between the hydroxyl group and the *p*-methyl group occurring at approximately 0.35 nm, indicating head-to-tail orientation similar to the case of ethanol. Use of GROMOS96 43A1 parameters showed this behavior as well, but at a much larger distance, approximately 0.7 nm, while hydrogen bonding was clearly occurring due to hydroxyl-hydroxyl interactions at approximately 0.25 nm. A peak at approximately 0.55 nm occurred in the RDF produced by both PRODRG and GROMOS96 43A1 parameters; this peak is the intramolecular distance between the hydroxyl group and the *p*-methyl group.

RDF plots for hexane (data not shown) indicate that there was no distinguishable difference in the liquid structure due to the application of GROMOS96 43A1 or PRODRG parameters. The average end-to-end distance of hexane was 0.607 ± 0.001 nm in both cases, with a radius of gyration of 0.216 ± 0.003 nm. Thus, the presence of small partial charges on hexane assigned by PRODRG did not significantly influence the configurations of hexane or the structure of the bulk liquid.

11.3.2 Interfacial amino acid systems

To assess the quality of PRODRG-generated amino acid topologies relative to those present in the GROMOS96 43A1 force field library, several parameters were analyzed: the position and orientation of the amino acids in the biphasic cyclohexane-water system, the nonbonded energies (Coulombic and Lennard-Jones) between the amino acids and the two solvents, and the ability of relevant functional groups in the amino acids to hydrogen bond with water. The balance of these terms will dictate the behavior of the amino acids in these systems and thus is an important measure of the quality of the small molecule topologies.

11.3.2.1 Configurations and Trajectory Observations

The position of amino acids parameterized under the GROMOS96 43A1 force field was largely consistent over time, with all amino acids remaining at the cyclohexane-water interface. The only exception was one simulation of isoleucine, in which the molecule diffused into the cyclohexane layer to localize at the cyclohexane-water interface that is formed at the periodic boundary of the system in the z -dimension (Figure 11.7). The orientation of the sidechains varied depending on the chemical nature of the amino acid. Hydrophobic amino acids (alanine and isoleucine) were oriented such that their side chains were either coincident with the interface or buried in the cyclohexane layer. Polar uncharged amino acids (asparagine and serine) were positioned with their sidechains coincident with the interface or directed towards the aqueous solvent. Charged amino acids (aspartate and lysine) consistently oriented their sidechains towards the aqueous solvent.

The behavior of the amino acids whose parameters were assigned by the PRODRG server was less consistent (Figure 11.8). In all three simulations of alanine, the amino acid diffused freely into the water layer. Simulations of isoleucine parameterized by PRODRG performed comparably to the simulations under the GROMOS96 43A1 parameters, with isoleucine remaining at the cyclohexane-water interface and its sidechain buried in the cyclohexane layer. The position of asparagine across three simulations was inconsistent; in two cases, asparagine remained largely interfacial, occasionally diffusing as much as 3 nm into the aqueous phase before returning to the interface. In the third simulation of asparagine, the amino acid diffused to the top of the unit cell to associate with the cyclohexane-water interface at the periodic boundary. In the simulations of serine, two systems remained interfacial with their polar sidechains buried in the cyclohexane layer. In the third serine simulation, the amino acid diffused to the top of the unit cell to ultimately bury its sidechain into the periodic cyclohexane layer. In simulations of both lysine and aspartate, the amino acids diffused freely into the aqueous phase; their positions were variable over time.

11.3.2.2 Energetics of Interactions with Co-Solvents

The interactions of the amino acids with the two co-solvents of the interfacial systems were strongly influenced by the nature of the electrostatic and van der Waals interactions between the

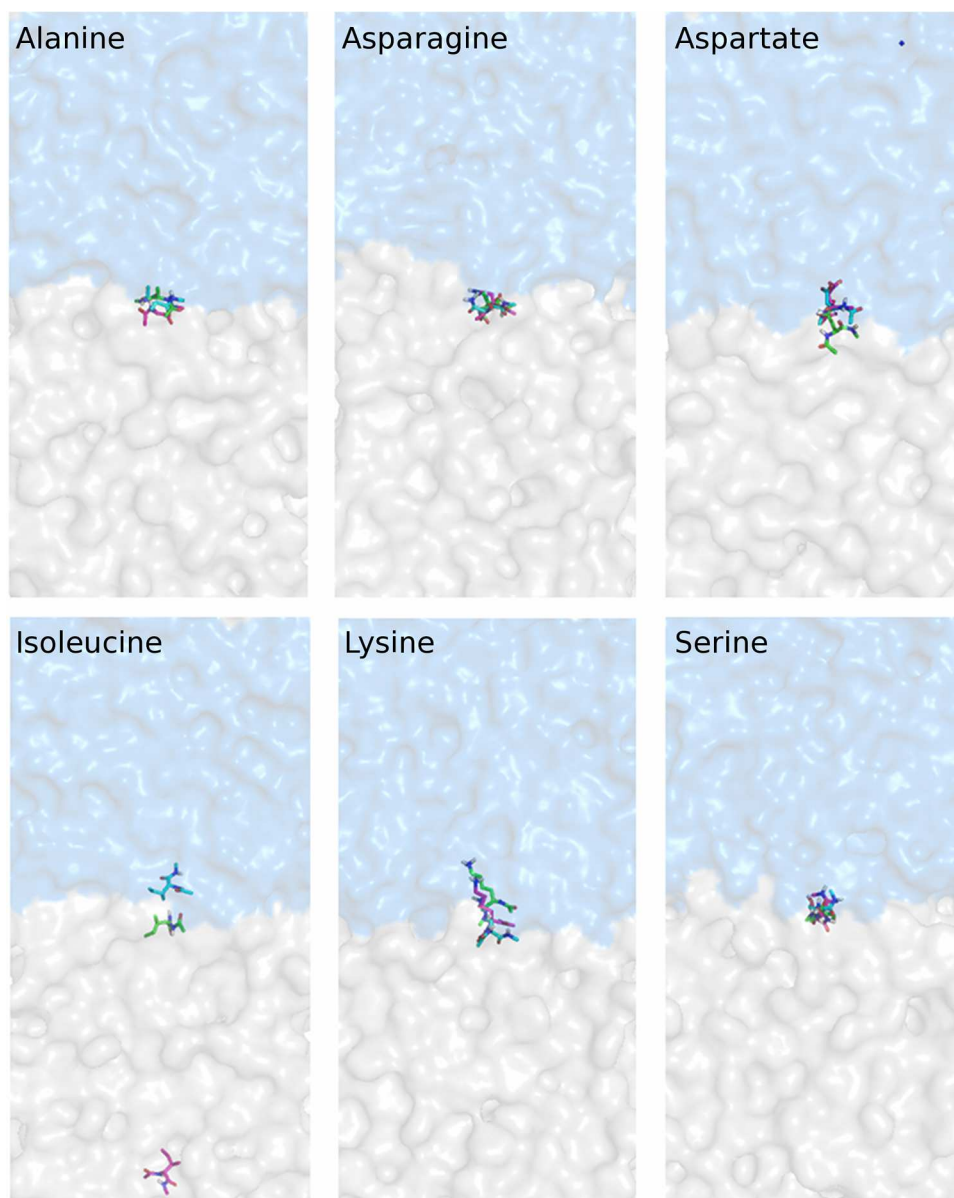


Figure 11.7: Overlays of amino acid positions generated by GROMOS96 43A1 parameters, taken from the final snapshot (20 ns) of each simulation after fitting for global $x - y$ translation. Amino acids are shown as sticks and colored separately by replicate. Water and cyclohexane are rendered as in Figure 11.2. The images were generated with PyMOL [272].

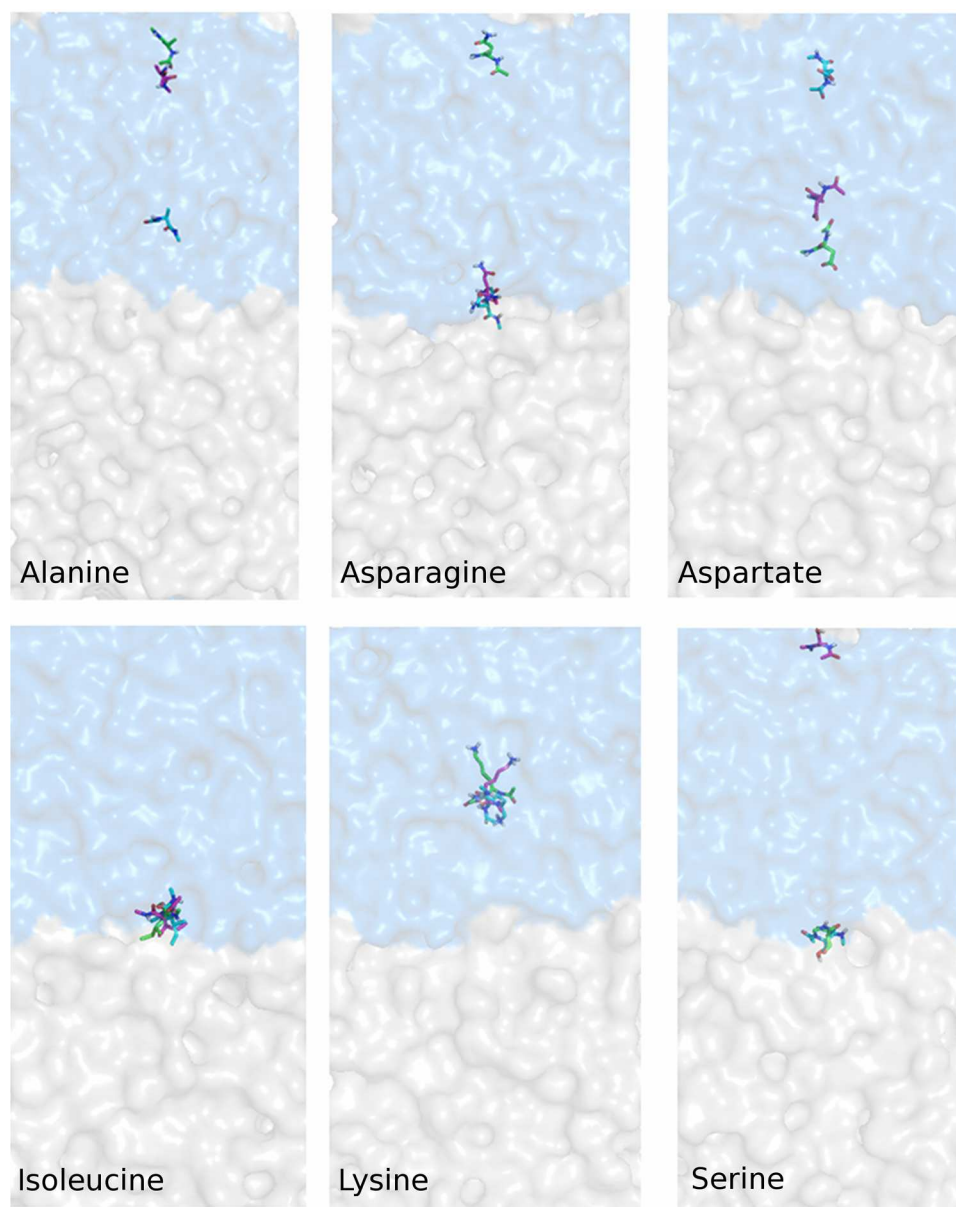


Figure 11.8: Overlays of amino acid positions generated by PRODRG parameters, taken from the final snapshot (20 ns) of each simulation after fitting for global $x - y$ translation. Amino acids are shown as sticks and colored separately by replicate. Water and cyclohexane are rendered as in Figure 11.2. The images were generated with PyMOL [272].

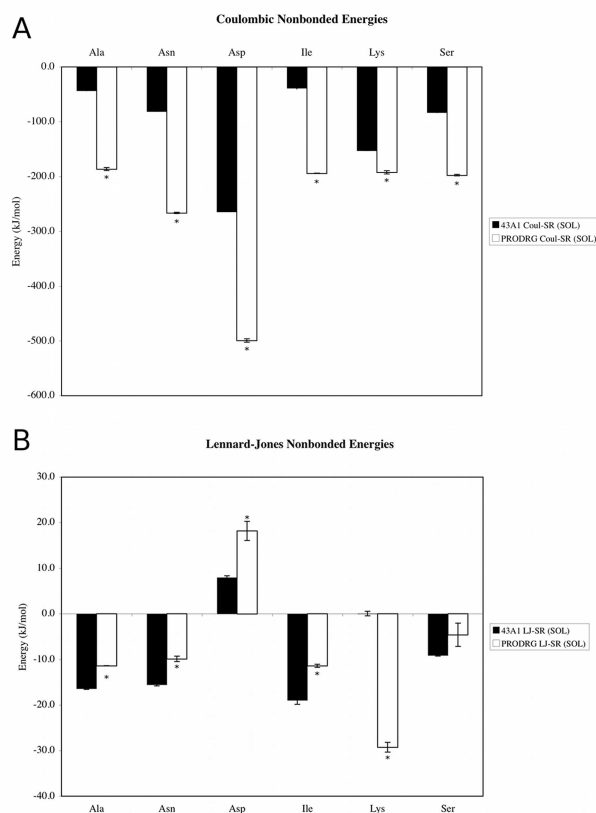


Figure 11.9: (A) Electrostatic and (B) Lennard-Jones interaction energies between the amino acids (listed on the x -axis) and water. Asterisks (*) indicate a significant difference based on a two-tailed Student's t -test ($p < 0.05$). Error bars represent standard deviations of the three simulations of each system.

small molecules and the aqueous and hydrophobic environments. The balance between these terms dictated the partitioning behavior of the amino acids. Since the cyclohexane topology included no charges on the constituent atoms, the only electrostatic interactions in these systems (aside from water-water interactions) occurred between the amino acid and water. In all cases, the charges assigned in PRODRG-generated topologies led to significantly lower-energy (more favorable) electrostatic interactions with water than the results obtained using GROMOS96 43A1 parameters (Figure 11.9A).

As a consequence of the greater interaction with the aqueous phase in the PRODRG-parameterized systems, atoms were drawn closer together by stronger electrostatic interactions. As a result, van der Waals interactions with the aqueous phase became more repulsive for all amino acids except for lysine, which experienced a greater net Lennard-Jones attraction (Figure 11.9B). RDF analysis (Figure 11.10) indicates that more water molecules interacted with the sidechain methylene groups when PRODRG charges were applied. These groups are normally uncharged in the GROMOS96 43A1 parameter set, leading to weaker solvent interactions. The oxygen atom in water interacted

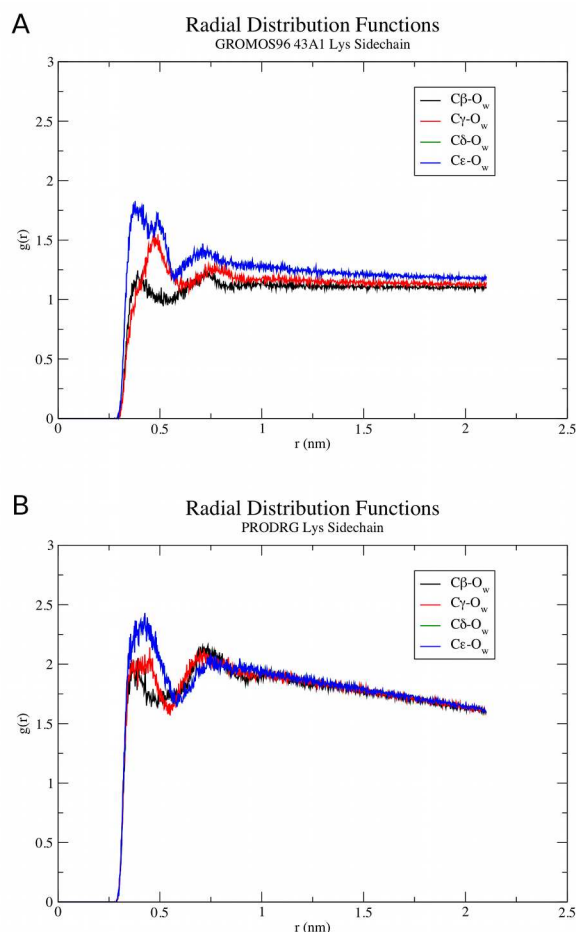


Figure 11.10: RDF for SPC oxygen (O_w) atoms with respect to lysine sidechain atoms using (A) GROMOS96 43A1 and (B) PRODRG parameters. In both panels, the curves for $C\delta$ and $C\epsilon$ atoms are overlapping.

with each of the sidechain methylene groups when small partial charges (assigned by PRODRG) were present such that favorable van der Waals interactions were established, a behavior that was promoted by favorable dipole interactions from the partial charges present.

For all the other amino acids considered here, the use of PRODRG parameters caused water molecules to be distributed around the solute molecules at either a closer distance or at the same distance but with higher probability than with GROMOS96 43A1 parameters, based on RDF analysis (data not shown). The attraction of water molecules to these species derived largely from the over-polarization of the backbone and/or sidechains of each amino acid in the PRODRG topologies. The result was a net repulsion due to an increased density of water around these amino acids, particularly at distances between atomic pairs above the energy minimum on the Lennard-Jones curve.

11.3.2.3 Hydrogen Bonding

All amino acids parameterized under GROMOS96 43A1 formed essentially the same number of backbone hydrogen bonds with water (approximately 2), averaged over all simulations and all snapshots in each simulation (Figure 11.11A). The consistency across simulations is likely due to the fact that the backbone amide groups had identical parameters, in accordance with the transferability of functional groups within the GROMOS parameter sets. That is, in the GROMOS force fields, the same functional groups have the same charges, even if they occur in different species, since charges are derived from model species that represent a small portion of an existing compound, typically an amino acid sidechain [174]. For all the amino acids examined here, the PRODRG-produced topologies led to a significant increase in the number of backbone-water hydrogen bonds relative to the results obtained with GROMOS96 43A1 parameters, nearly 6 in the case of aspartate. This effect can be attributed to the greater separation of charges in the backbone amide groups (Figures 11.3 and 11.4). These charges vary somewhat in their magnitude depending on the amino acid, but in all cases the resulting dipoles are greater than those present in the GROMOS96 43A1 topologies.

Sidechain hydrogen bonds are another important consideration when examining these systems. Illustrated in Figure 11.11B are hydrogen bonds between sidechain groups (where applicable) and water. In the cases of asparagine and aspartate, significantly more hydrogen bonds formed between sidechain groups (amide and carboxylate, respectively) in the case of PRODRG-generated topologies. The principal reason for this increase is the greater magnitude of charge separation on the amide carbonyl in asparagine and carboxylate in aspartate. The stronger dipoles in these groups led to more hydrogen bonding with nearby water. The opposite effect is observed in the case of lysine and serine systems, wherein hydrogen bonds were significantly decreased relative to the system simulated with GROMOS96 43A1 parameters. In the case of lysine, the charges on the ϵ -amino group hydrogens assigned by PRODRG are $-0.002 e$, opposite in sign and significantly reduced in magnitude compared to the charges that are assigned under GROMOS96 43A1 ($+0.248 e$). The partial negative charges assigned by PRODRG likely inhibit the ability of this ϵ -amino group to effectively serve as a hydrogen bond donor. As described above, in the simulations of serine, the PRODRG-derived charges caused the amino acid to bury its sidechain in the cyclohexane layer, rather than in water. The dipole on the sidechain hydroxyl group is reduced compared to the GROMOS96 43A1 topology, and thus hydrogen bonding to water was less of a driving force in the interactions in these systems.

The results from the interfacial amino acid simulations call into question the use of unrefined PRODRG topologies for small molecules. PRODRG was unable to replicate the partial charges of the GROMOS96 43A1 parameter set for the atoms of any of the amino acids, which are the most well-characterized species in most biomolecular force fields. Very little information is publicly available regarding the method by which PRODRG assigns charges to the input molecules, aside from a brief statement in the original paper stating that charges are referenced from a database of common functional groups [403]. In fact, it seems that, in the case of uncharged molecules, PRODRG assigns charges that are a hybrid of polar and nonpolar groups. That is, polar groups

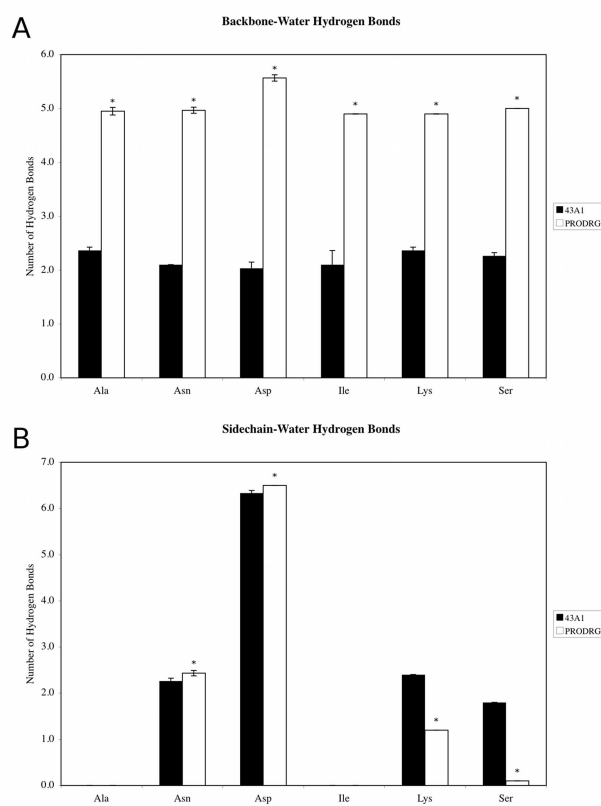


Figure 11.11: (A) Backbone and (B) sidechain hydrogen bonds to water. Asterisks (*) indicate a significant difference based on a two-tailed Student's t -test ($p < 0.05$). Error bars represent standard deviations of the three simulations of each system.

(Asn and Ser sidechains, for example) are insufficiently polar, while normally uncharged groups (Ala and Ile sidechains, methylene groups in Lys) are assigned small partial charges, rendering them more hydrophilic. Some other polar groups, such as backbone amides in all the amino acids, have much larger dipoles assigned by PRODRG than would be expected under the GROMOS96 43A1 parameter set. In the case of charged groups, such as the carboxylate group of the Asp sidechain and the amino group of the Lys sidechain, the dipole on the carboxylate group is too large, while that of the amino group is too small.

11.3.3 UDP-galactopyranose mutase

The PRODRG server was designed to produce ligand topologies for use in modeling protein-ligand structures [403, 404]. Thus, it is important to assess its applicability to this purpose. To design a test system for which ligand parameters are already present in the GROMOS96 43A1 force field library, we chose to model the UGM enzyme in complex with its FAD cofactor. Modeling other small molecules in protein binding sites would be an interesting exercise, but the lack of a clear method for generating their topologies has limited the focus of this study.

Over the course of triplicate 40-ns trajectories under both GROMOS96 43A1 parameters and those generated by PRODRG, the FAD molecule remained bound to the protein (which, in all cases, was assigned parameters under the GROMOS96 43A1 force field), with its position largely consistent over time. That is, there was no global translation of the cofactor out of its binding site, or rotation therein. However, we did observe a number of important differences with respect to the configuration of FAD and its interactions with the UGM enzyme. Structural overlays, taken from the final snapshot of each trajectory, are shown in Figure 11.12. While the position of the isoalloxazine ring was largely consistent across the simulations under both GROMOS96 43A1 and PRODRG, the location of the adenine ring was more variable in the simulations utilizing PRODRG parameters.

The configurations generated in these simulations would seem to indicate that the PRODRG parameters assigned to FAD are sufficiently accurate for simulation. However, a conclusion regarding accuracy cannot be made based on structures alone. The nature of the interactions between FAD and UGM is an important criterion for assessing the accuracy of these parameters, and is also critically important in simulations of small molecules docked to receptors. MD simulations can be used to assess the stability of docking results and can be used to calculate the energetics of the resulting interactions over time.

We found that the parameters assigned to FAD by PRODRG led the cofactor to participate in significantly fewer hydrogen bonds to UGM. In the original crystal structure, eight hydrogen bonds between FAD and UGM were detected, using a cutoff angle of 30° and a cutoff radius of 3.5 \AA . Under GROMOS96 43A1, 10.3 ± 0.4 hydrogen bonds were formed (averaged over time and over three replicate simulations). Using PRODRG parameters resulted in the formation of just 5.0 ± 0.9 hydrogen bonds. Several of these missing hydrogen bonds can likely be attributed to the assignment of zero charge to the atoms of the C6-amino group of the adenine ring (Figure 11.5). The absence of hydrogen bonding between this moiety and UGM also explains the variability of

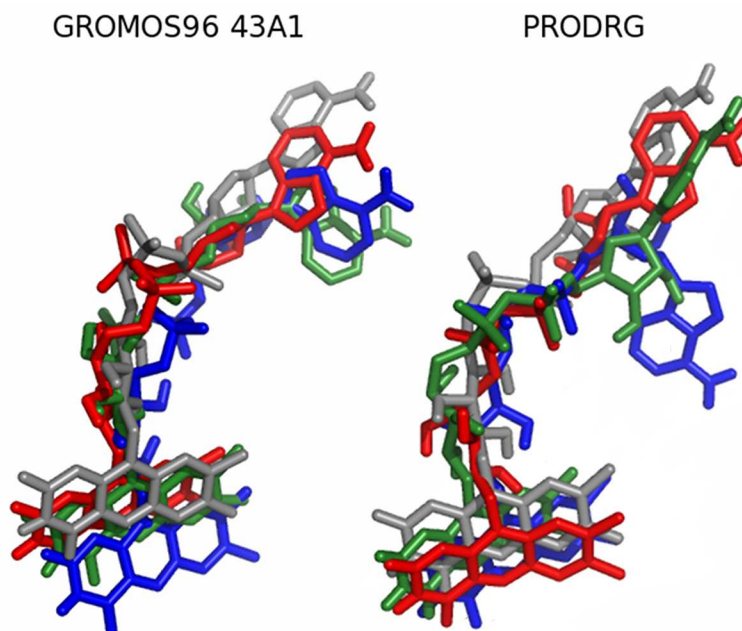


Figure 11.12: Alignment of FAD structures, with the reference position (the energy-minimized configuration of FAD) for each simulation shown in gray. The final snapshot of each 40-ns simulation is shown and is colored by simulation. Overlays were produced by fitting the structure of the protein to remove global translation and rotation and extracting the resulting FAD coordinates. The images were generated with PyMOL [272].

the position of the adenine ring across the three simulations.

MD simulations are also commonly used to predict binding energies between small molecules and their target proteins [420]. By measuring nonbonded interaction energies between FAD and UGM, the nature of the interactions between the protein and its cofactor can be assessed. When GROMOS96 43A1 charges were assigned to FAD, the average total interaction energy between UGM and FAD was $-866 \pm 52 \text{ kJ mol}^{-1}$. When PRODRG-derived charges were assigned to FAD, the average interaction energy was $-588 \pm 30 \text{ kJ mol}^{-1}$. While the Lennard-Jones energies were similar in both simulation sets, the Coulombic energies differed drastically. With GROMOS96 43A1 charges, the magnitude of the intermolecular Coulombic interactions was, on average, $-427 \pm 36 \text{ kJ mol}^{-1}$, but with PRODRG charges, this term was $-158 \pm 22 \text{ kJ mol}^{-1}$. Thus, despite the apparent similarity in the positions of FAD within the UGM enzyme, the energetics of this interaction were remarkably different, particularly with respect to the Coulombic interactions.

It is apparent that the inconsistent charge and charge group assignments generated by PRODRG had an adverse effect on the dynamics and behavior of the FAD cofactor in its binding site. However, it is even more of a concern that the PRODRG-assigned FAD charges might influence the dynamics of the protein itself. A measure of the root mean square fluctuation (RMSF) in UGM indicated that helix $\alpha 9$ and the loop following it (residues 215-224) significantly increased in fluc-

tuation ($p < 0.05$) in one simulation in which the FAD molecule was assigned PRODRG charges (Figure 11.13). The average RMSF over the last 20 ns of simulation using GROMOS96 43A1 parameters for FAD was 0.18 ± 0.04 nm, while that of the simulation using PRODRG parameters was 0.26 ± 0.08 . This loop is important in binding the adenine moiety of the FAD, and we attribute the increased fluctuation to a lack of hydrogen bonding between UGM and FAD in this region. The two other trajectories using PRODRG charges for FAD resulted in reasonable configurations for the cofactor, similar to those obtained by assigning charges from GROMOS96 43A1. Nonetheless, we have demonstrated that the energetics of these interactions deviate substantially from the expected results using GROMOS96 43A1. The implication that PRODRG topology charges extend beyond the small molecule for which they were originally intended raises concerns about the validity of any simulations that have used unrefined PRODRG topologies.

11.3.4 Strategies for charge assignment in GROMOS topologies

The empirical charge assignment applied in the GROMOS96 force field parameter sets presents one of the principal challenges in deriving small molecule topologies. There is no complete description in the literature regarding the proper derivation of charges for molecules within the GROMOS96 parameter sets, aside from empirical fitting. It is for this reason that automated methods of charge calculation and assignment are so attractive. Having concluded that the charges assigned by PRODRG did not reproduce the expected behavior for even the most well-defined molecules and common functional groups, we offer strategies for initial charge assignment that are relatively fast and easily applied to GROMOS topologies. For application to small molecules in MD simulations, these initial charges should be evaluated and refined as necessary to reproduce proper condensed-phase behavior using the methodology described by the authors of the GROMOS96 parameter sets [174, 182, 184, 399, 400] and perhaps the strategies employed here. The charge-derivation strategies we discuss here should be considered a starting point, not an ending point, for this process.

Quantum mechanical calculations are not explicitly stated as a source of GROMOS96 charges, although they present one method for deriving these parameters. We utilized the Antechamber program within AmberTools [393] (version 1.0) to calculate charges for all-atom versions of several of the small molecules studied here by applying several different semi-empirical charge calculation methods. The results of these calculations are summarized in Tables 11.2 - 11.8. Overall, the charge calculation method that best reproduced the GROMOS96 43A1 partial charges is AM1-BCC [421, 422]. Charges on O and H atoms of hydroxyl groups were in good agreement with expected force field values for ethanol, *p*-cresol, and serine. For primary alcohol functional groups in these molecules, if the AM1-BCC partial charges of the C and H atoms in the CH₂ group are combined, the resulting charge was similar to the force field value of $+0.150 e$, $+0.1677 e$ in the case of ethanol and $+0.1919 e$ in the case of serine.

Aromatic C and H atoms had charges of comparable magnitude as well, a result that was also produced by Mulliken charges [423]. The Mulliken charges for amide groups were in closest

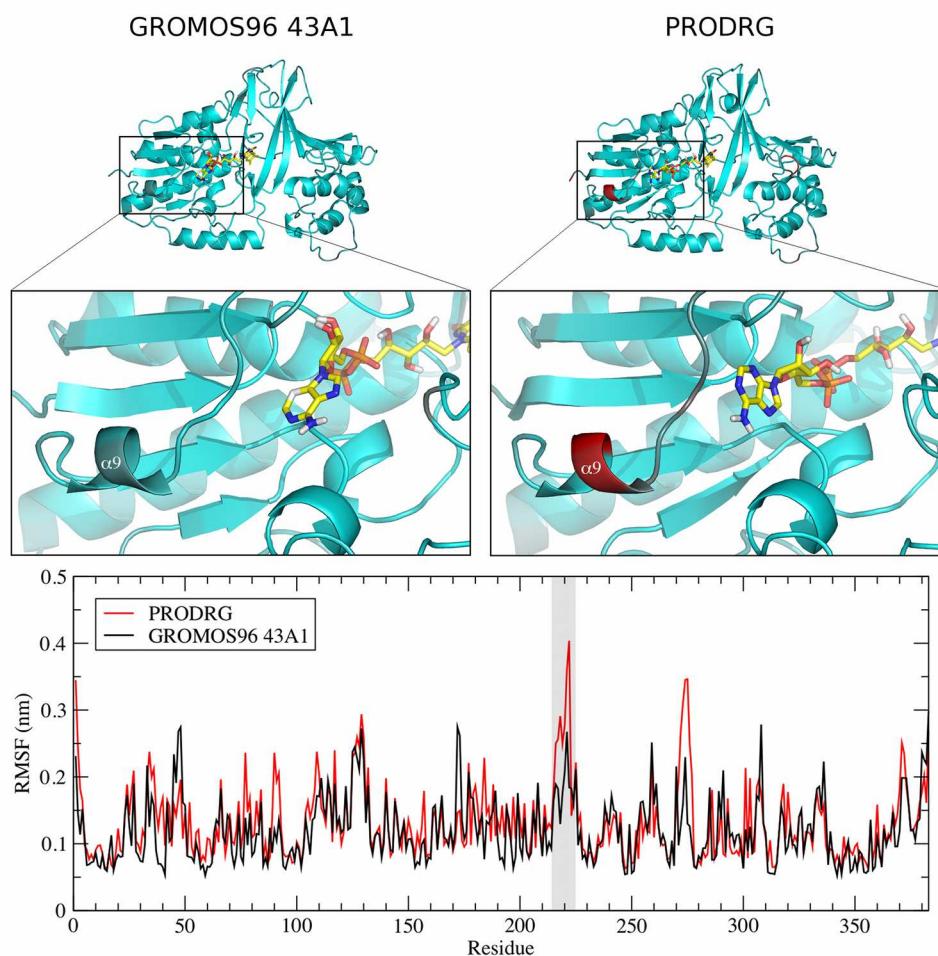


Figure 11.13: RMSF of the UGM backbone. Only the charges on the FAD molecule differ between the two models. The protein (cartoon) is colored according to the backbone RMSF, using a color gradient from blue to red, with blue regions indicating less fluctuation and red regions indicating more fluctuation. The FAD cofactor is shown as sticks and colored by element. Structures are taken from representative snapshots of each simulation set (GROMOS96 43A1 and PRODRG FAD parameters). The images were generated with PyMOL [272]. The RMSF of residues referred to in the main text is indicated in the shaded box on the RMSF plot.

Table 11.2: Small-molecule charges calculated by Antechamber semi-empirical methods.

Molecule	Group	Atom	Partial charges (e)				
			AM1-BCC	Mulliken	Gasteiger	GROMOS96	
Ethanol	CH ₃	C	-0.0971	-0.2150	-0.041839		
		H1	+0.0380	+0.0773	+0.025375		
		H2	+0.0481	+0.0874	+0.025375		
		H3	+0.0482	+0.0875	+0.025375		
		Sum	+0.0372	+0.0372	+0.034286	0.000	
	CH ₂	C	+0.1312	-0.0192	+0.040211		
		H1	+0.0183	+0.0576	+0.056075		
		H2	+0.0182	+0.0575	+0.056075		
		Sum	+0.1677	+0.0959	+0.152361	+0.150	
	OH	O	-0.6024	-0.3296	-0.396675	-0.548	
		H	+0.3976	+0.1996	+0.210027	+0.398	
	<i>p</i> -Cresol	CH ₃	C	-0.0487	-0.1739	-0.039771	
			H1	+0.0399	+0.0792	+0.027965	
H2			+0.0437	+0.0830	+0.027965		
H3			+0.0437	+0.0830	+0.027965		
Sum			+0.0786	+0.0713	+0.044214	0.000	
C		C	-0.1123	-0.1050	-0.050945	0.000	
CH*		C	-0.1381	-0.1381	-0.044274	-0.100	
		H	+0.1371	+0.1371	+0.063842	+0.100	
COH		C	+0.1187	+0.0736	+0.071255	+0.150	
		O	-0.4987	-0.2526	-0.360943	-0.548	
		H	+0.4179	+0.2169	+0.218239	+0.398	

* Averaged over all equivalent groups in the ring.

Table 11.3: Alanine charges calculated by Antechamber semi-empirical methods.

Molecule	Group	Atom	Partial charges (e)			GROMOS96
			AM1-BCC	Mulliken	Gasteiger	
Alanine	Acetyl CH ₃	C	-0.1760	-0.2439	+0.012703	
		H1	+0.0782	+0.1175	+0.032805	
		H2	+0.0657	+0.1050	+0.032805	
		H3	+0.0797	+0.1190	+0.032805	
		Sum	+0.0476	+0.0976	+0.111118	0.000
	N-terminal amide	C	+0.6647	+0.3086	+0.209132	+0.380
		O	-0.6045	-0.3654	-0.278060	-0.380
		N	-0.5889	-0.3980	-0.304237	-0.280
		H	+0.3170	+0.2305	+0.150160	+0.280
	C α H	C	+0.0409	+0.0142	+0.096349	
		H	+0.0891	+0.1284	+0.060201	
		Sum	+0.1300	+0.1426	+0.156550	0.000
	C β H ₃	C	-0.1083	-0.2262	-0.037773	
		H1	+0.0631	+0.1024	+0.025456	
		H2	+0.0547	+0.0940	+0.025456	
		H3	+0.0406	+0.0799	+0.025456	
		Sum	+0.0501	+0.0501	+0.038595	0.000
	C-terminal amide	C	+0.6266	+0.2705	+0.231571	+0.380
		O	-0.6072	-0.3681	-0.275604	-0.380
		N	-0.5638	-0.3729	-0.316918	-0.280
H		+0.3124	+0.2259	+0.149047	+0.280	
C-terminal N-CH ₃	C	+0.0823	-0.0730	-0.000397		
	H1	+0.0407	+0.0800	+0.043014		
	H2	+0.0363	+0.0756	+0.043014		
	H3	+0.0568	+0.0961	+0.043014		
	Sum	+0.2161	+0.1787	+0.128645	0.000	

Table 11.4: Isoleucine charges calculated by Antechamber semi-empirical methods.

Molecule	Group	Atom	Partial charges (e)			GROMOS96
			AM1-BCC	Mulliken	Gasteiger	
Isoleucine	Acetyl CH ₃	C	-0.1730	-0.2409	+0.012704	
		H1	+0.0778	+0.1171	+0.032805	
		H2	+0.0619	+0.1012	+0.032805	
		H3	+0.0647	+0.1040	+0.032805	
		Sum	+0.0314	+0.0814	+0.111119	0.000
	N-terminal amide	C	+0.6622	+0.3061	+0.209151	+0.380
		O	-0.6148	-0.3757	-0.278060	-0.380
		N	-0.5600	-0.3691	-0.303683	-0.280
		H	+0.3265	+0.2400	+0.150185	+0.280
	C α H	C	+0.0411	+0.0144	+0.101681	
		H	+0.0979	+0.1372	+0.060787	
		Sum	+0.1390	+0.1516	+0.162468	0.000
	C β H	C	-0.0669	-0.1062	-0.017161	
		H	+0.0494	+0.0887	+0.032411	
		Sum	-0.0175	-0.0175	+0.01525	0.000
	C γ H ₃	C	-0.0973	-0.2116	-0.060239	
		H1	+0.0434	+0.0827	+0.023402	
		H2	+0.0353	+0.0746	+0.023402	
		H3	+0.0449	+0.0842	+0.023402	
		Sum	+0.0263	+0.0299	+0.009967	0.000
	C γ H ₂	C	-0.0743	-0.1529	-0.051301	
		H1	+0.0452	+0.0854	+0.026729	
		H2	+0.0416	+0.0809	+0.026729	
		Sum	+0.0125	+0.0134	+0.002157	0.000
	C δ H ₃	C	-0.1046	-0.2225	-0.065045	
		H1	+0.0527	+0.0920	+0.023047	
		H2	+0.0360	+0.0753	+0.023047	
H3		+0.0291	+0.0684	+0.023047		
Sum		+0.0132	+0.0132	+0.004096	0.000	
C-terminal amide	C	+0.6375	+0.2814	+0.232134	+0.380	
	O	-0.6166	-0.3775	-0.275579	-0.380	
	N	-0.5736	-0.3827	-0.316889	-0.280	
	H	+0.3141	+0.2276	+0.149047	+0.280	
C-terminal N-CH ₃	C	+0.0873	-0.0680	-0.000396		
	H1	+0.0454	+0.0847	+0.043014		
	H2	+0.0394	+0.0787	+0.043014		
	H3	+0.0439	+0.0832	+0.043014		
	Sum	+0.2160	+0.1786	+0.128646	0.000	

Table 11.5: Asparagine charges calculated by Antechamber semi-empirical methods.

Molecule	Group	Atom	Partial charges (<i>e</i>)			
			AM1-BCC	Mulliken	Gasteiger	GROMOS96
Asparagine	Acetyl CH ₃	C	-0.1735	-0.2414	+0.012704	
		H1	+0.0752	+0.1145	+0.032805	
		H2	+0.0623	+0.1016	+0.032805	
		H3	+0.0614	+0.1007	+0.032805	
		Sum	+0.0254	+0.0754	+0.110399	0.000
	N-terminal amide	C	+0.6557	+0.2996	+0.209155	+0.380
		O	-0.6149	-0.3758	-0.278060	-0.380
		N	-0.5534	-0.3625	-0.303433	-0.280
		H	+0.3234	+0.2369	+0.150190	+0.280
	C _α H	C	-0.0286	-0.0553	+0.107680	
		H	+0.0998	+0.1391	+0.061037	
		Sum	+0.0712	+0.0838	+0.168717	0.000
	C _β H ₂	C	+0.0614	+0.0328	+0.051398	
		H1	+0.0590	+0.0983	+0.038395	
		H2	+0.0754	+0.1147	+0.038395	
		Sum	+0.1958	+0.2458	+0.128188	
	Sidechain amide	C	+0.5935	+0.2374	+0.211443	+0.380
		O	-0.6521	-0.4130	-0.277893	-0.380
		N	-0.6593	-0.4193	-0.329083	0.830
		H1	+0.3313	+0.2268	+0.145484	+0.415
H2		+0.3264	+0.2399	+0.145484	+0.415	
C-terminal amide	C	+0.6302	+0.2741	+0.232385	+0.380	
	O	-0.6330	-0.3939	-0.275574	-0.380	
	N	-0.5693	-0.3784	-0.316895	-0.280	
	H	+0.3390	+0.2525	+0.149047	+0.280	
C-terminal N-CH ₃	C	+0.0868	-0.0685	-0.000396		
	H1	+0.0381	+0.0774	+0.043014		
	H2	+0.0229	+0.0622	+0.043014		
	H3	+0.0603	+0.0996	+0.043014		
	Sum	+0.2081	+0.1707	+0.128646	0.000	

Table 11.6: Serine charges calculated by Antechamber semi-empirical methods.

Molecule	Group	Atom	Partial charges (e)			
			AM1-BCC	Mulliken	Gasteiger	GROMOS96
Serine	Acetyl CH ₃	C	-0.1766	-0.2445	+0.012705	
		H1	+0.0829	+0.1222	+0.032806	
		H2	+0.0783	+0.1176	+0.032806	
		H3	+0.0674	+0.1067	+0.032806	
		Sum	+0.0520	+0.1029	+0.111123	0.000
	N-terminal amide	C	+0.6657	+0.3096	+0.209208	+0.380
		O	-0.6022	-0.3631	-0.278059	-0.380
		N	-0.5854	-0.3945	-0.301933	-0.280
		H	+0.3153	+0.2288	+0.150258	+0.280
	C _α H	C	+0.0447	+0.0180	+0.122426	
		H	+0.1033	+0.1426	+0.062991	
		Sum	+0.1480	+0.1606	+0.185417	0.000
	C _β H ₂	C	+0.1238	-0.0266	+0.070315	
		H1	+0.0283	+0.0676	+0.058897	
		H2	+0.0398	+0.0791	+0.058897	
		Sum	+0.1919	+0.1201	+0.188109	+0.150
	OH	O	-0.5855	-0.3127	-0.393908	-0.548
		H	+0.4101	+0.2091	+0.210159	+0.398
	C-terminal amide	C	+0.6454	+0.2893	+0.234251	+0.380
		O	-0.6069	-0.3678	-0.275489	-0.380
N		-0.5679	-0.3770	-0.316830	-0.280	
H		+0.3207	+0.2342	+0.149048	+0.280	
C-terminal N-CH ₃	C	+0.0873	-0.0680	-0.000395		
	H1	+0.0491	+0.0884	+0.043014		
	H2	+0.0269	+0.0662	+0.043014		
	H3	+0.0357	+0.0750	+0.043014		
	Sum	+0.1990	+0.1616	+0.128647	0.000	

Table 11.7: Aspartate charges calculated by Antechamber semi-empirical methods.

Molecule	Group	Atom	Partial charges (e)			
			AM1-BCC	Mulliken	Gasteiger	GROMOS96
Aspartate	Acetyl CH ₃	C	-0.1932	-0.2619	+0.012704	
		H1	+0.0502	+0.0898	+0.032805	
		H2	+0.1394	+0.1789	+0.032805	
		H3	+0.0499	+0.0888	+0.032805	
		Sum	+0.0463	+0.0956	+0.111119	
	N-terminal amide	C	+0.6678	+0.3120	+0.209153	+0.380
		O	-0.6685	-0.4291	-0.278060	-0.380
		N	-0.5656	-0.3748	-0.303570	-0.280
		H	+0.3217	+0.2351	+0.150187	+0.280
	C α H	C	+0.0653	+0.0390	+0.104340	
		H	+0.0906	+0.1297	+0.060901	
		Sum	+0.1559	+0.1687	+0.165241	
	C β H ₂	C	-0.2249	-0.2537	+0.014249	
		H1	+0.0513	+0.0905	+0.035313	
		H2	+0.0475	+0.0866	+0.035313	
		Sum	-0.1261	-0.0766	+0.084875	
	Sidechain carboxylate	C	+0.9029	+0.3223	+0.043753	+0.270
		O1	-0.8321	-0.5667	-0.550083	-0.635
		O2	-0.8471	-0.5818	-0.550083	-0.635
	C-terminal amide	C	+0.6527	+0.2963	+0.232248	+0.380
O		-0.6593	-0.4202	-0.275577	-0.380	
N		-0.5642	-0.3734	-0.316897	-0.280	
H		+0.2999	+0.2136	+0.149047	+0.280	
C-terminal N-CH ₃	C	+0.0547	-0.0997	-0.000396		
	H1	+0.0240	+0.0635	+0.043014		
	H2	+0.0197	+0.0592	+0.043014		
	H3	+0.1173	+0.1559	+0.043014		
	Sum	+0.2157	+0.1789	+0.128646		

Table 11.8: Lysine charges calculated by Antechamber semi-empirical methods.

Molecule	Group	Atom	Partial charges (e)			GROMOS96
			AM1-BCC	Mulliken	Gasteiger	
Lysine	Acetyl CH ₃	C	-0.1762	-0.2399	+0.012705	
		H1	+0.0948	+0.1341	+0.032808	
		H2	+0.0647	+0.1040	+0.032808	
		H3	+0.0597	+0.0990	+0.032808	
		Sum	+0.0430	+0.0972	+0.111120	
	N-terminal amide	C	+0.6577	+0.3016	+0.209136	+0.380
		O	-0.5695	-0.3304	-0.278068	-0.380
		N	-0.6005	-0.4096	-0.303951	-0.280
		H	+0.3257	+0.2392	+0.150176	+0.280
	C α H	C	+0.0523	+0.0256	+0.099156	
		H	+0.0843	+0.1236	+0.060503	
	Sum	+0.1366	+0.1492	+0.159659		
	C β H ₂	C	-0.0858	-0.1644	-0.025604	
		H1	+0.0693	+0.1086	+0.029088	
		H2	+0.0622	+0.1015	+0.029088	
		Sum	+0.0457	+0.0457	+0.032572	
	C γ H ₂	C	-0.0682	-0.1468	-0.046073	
		H1	+0.0710	+0.1103	+0.026951	
		H2	+0.0542	+0.0935	+0.026951	
		Sum	+0.0570	+0.0570	+0.007829	
	C δ H ₂	C	-0.1081	-0.1867	-0.013759	
		H1	+0.0646	+0.1039	+0.032059	
		H2	+0.0877	+0.1270	+0.032059	
		Sum	+0.0442	+0.0442	+0.050359	
	C ϵ H ₂	C	+0.1022	-0.1346	-0.038079	
		H1	+0.1066	+0.1459	+0.081283	
		H2	+0.1205	+0.1598	+0.081283	
Sum		+0.3293	+0.1711	+0.124487	+0.127	
Sidechain amine	N	-0.8393	-0.0667	+0.231559	+0.129	
	H1	+0.4575	+0.2527	+0.196024	+0.248	
	H2	+0.4604	+0.2556	+0.196024	+0.248	
	H3	+0.4646	+0.2598	+0.196024	+0.248	
C-terminal amide	C	+0.6356	+0.2795	+0.231852	+0.380	
	O	-0.6717	-0.4326	-0.275598	-0.380	
	N	-0.5425	-0.3516	-0.316909	-0.280	
	H	+0.3300	+0.2435	+0.149049	+0.280	
C-terminal N-CH ₃	C	+0.0735	-0.0818	-0.000399		
	H1	+0.0634	+0.1027	+0.043016		
	H2	+0.0554	+0.0947	+0.043016		
	H3	+0.0395	+0.0788	+0.043016		
	Sum	+0.2318	+0.1944	+0.128649		

agreement with the expected values in the GROMOS96 43A1 parameter set. The AM1-BCC method tended to over-polarize amide functional groups, but in all cases the direction of the dipole was correct. That is, none of the N atoms were assigned partial positive charges, as was the case in all of the PRODRG amino acid topologies. All of the methods utilized here also had a tendency to place small partial charges on nonpolar groups, an inconsistency with the GROMOS96 43A1 force field shared by PRODRG. The Gasteiger method [424] produced charges that were in poorest agreement with expected GROMOS96 43A1 parameters.

Antechamber is freely available within the AmberTools distribution, but it is limited to semi-empirical calculations. Spartan '04 (Wavefunction, Inc., Irvine, CA) is a commercially-available quantum mechanical program that can perform some higher level of theory calculations, and we sought to determine what level of theory would be necessary to best reproduce GROMOS96 43A1 charges. We employed three methods: the semi-empirical AM1 method [425], the *ab initio* Hartree-Fock (HF) method [426] with the 6-31G** basis set, and a common *ab initio* density functional theory (DFT) hybrid method (B3LYP) [427], also with the 6-31G** basis set. Charges were derived using the CHELP algorithm [428] within Spartan, which is a type of electrostatic charge (ESP) method suitable for describing intermolecular interactions.

The charges calculated by the Spartan methods (Tables 11.9 - 11.15) were, in general, all in good agreement with those charges found in the GROMOS96 43A1 force field. In the case of ethanol, the AM1 calculations yielded a near-neutral charge on the CH₃ group ($-0.003 e$), whereas the HF and DFT derived charges were somewhat more negative ($-0.075 e$ and $-0.057 e$, respectively). The net charge in the CH₂ group calculated by the AM1 method ($+0.240 e$) was again closest to the GROMOS96 43A1 parameters ($+0.150 e$), and the HF- and DFT-derived charges were slightly greater ($+0.341 e$ and $+0.263 e$, respectively). The magnitudes of the charges on the oxygen and hydrogen of the hydroxyl groups of all Spartan methods were in close agreement with the GROMOS96 43A1 charges, with the closest approximation coming from the DFT calculations. The important result is that the dipoles derived by these methods were greater in the hydroxyl group than in the methyl group, likely resulting in a correct hydrogen bonding behavior, which was not observed in the case of PRODRG-derived charges.

In the case of *p*-cresol, the *p*-methyl group was assigned a net negative charge by all methods, which was offset by the net positive charge on the carbon in the 4-position. The four equivalent carbons (at the 2-, 3-, 5-, and 6- positions) were all slightly over-polarized with the closest approximation to the GROMOS96 43A1 parameters coming from the AM1 calculations. The dipole on the hydroxyl group, however, was similar to the magnitude of this same group in the GROMOS96 43A1 parameter set in all cases, with the closest approximation derived by the DFT method. It appears that the magnitude of the dipole on the hydroxyl group compared to that of the *p*-methyl group was large enough such that a reasonable pattern of hydrogen bonding behavior would result.

In the case of serine, many of the same patterns are observed. The net charges on the N- and C-terminal methyl groups were in close agreement with the accepted GROMOS96 43A1 charges, whereas both the N- and C-terminal amide atoms were calculated to have the correct sign, but with a slight over-polarization by all methods. The largest disparity among these methods was the

Table 11.9: Small-molecule charges calculated by Spartan '04.

Molecule	Group	Atom	Partial charges (<i>e</i>)			GROMOS96
			AM1	HF 6-31G**	DFT (B3LYP) 6-31G**	
Ethanol	CH ₃	C	-0.437	-0.502	-0.504	
		H1	+0.154	+0.158	+0.162	
		H2	+0.154	+0.156	+0.161	
		H3	+0.126	+0.113	+0.124	
		Sum	-0.003	-0.075	-0.057	0.000
	CH ₂	C	+0.135	+0.344	+0.219	
		H1	+0.052	-0.001	+0.022	
		H2	+0.053	-0.002	+0.022	
	Sum	+0.240	+0.341	+0.263	+0.150	
	OH	O	-0.518	-0.668	-0.572	-0.548
H		+0.318	+0.402	+0.367	+0.398	
<i>p</i> -Cresol	CH ₃	C	-0.547	-0.758	-0.778	
		H1	+0.163	+0.205	+0.208	
		H2	+0.166	+0.211	+0.217	
		H3	+0.167	+0.213	+0.217	
		Sum	-0.051	-0.129	-0.136	0.000
	C	C	+0.075	+0.307	+0.304	0.000
	CH [†]	C	-0.211	-0.324	-0.277	-0.100
		H	+0.144	+0.211	+0.178	+0.100
	COH	C	+0.417	+0.479	+0.386	+0.150
		O	-0.524	-0.648	-0.566	-0.548
H		+0.354	+0.442	+0.412	+0.398	

† Averaged over all equivalent groups in the ring.

Table 11.10: Alanine charges calculated by Spartan '04.

Molecule	Group	Atom	Partial charges (<i>e</i>)			GROMOS96
			AM1	HF 6-31G**	DFT (B3LYP) 6-31G**	
Alanine	Acetyl CH ₃	C	-0.538	-0.786	-0.757	
		H1	+0.156	+0.208	+0.200	
		H2	+0.159	+0.222	+0.212	
		H3	+0.170	+0.230	+0.221	
		Sum	-0.053	-0.126	-0.124	0.000
	N-terminal amide	C	+0.721	+0.973	+0.816	+0.380
		O	-0.559	-0.672	-0.575	-0.380
		N	-0.569	-0.897	-0.775	-0.280
		H	+0.301	+0.399	+0.364	+0.280
	C α H	C	-0.062	+0.360	+0.334	
		H	+0.139	+0.053	+0.052	
		Sum	+0.077	+0.413	+0.386	0.000
	C β H ₃	C	-0.373	-0.578	-0.570	
		H1	+0.121	+0.158	+0.178	
		H2	+0.133	+0.149	+0.158	
		H3	+0.138	+0.184	+0.146	
Sum		+0.019	-0.087	-0.088	0.000	
C-terminal amide	C	+0.605	+0.597	+0.437	+0.380	
	O	-0.529	-0.607	-0.507	-0.380	
	N	-0.451	-0.478	-0.345	-0.280	
	H	+0.290	+0.332	+0.296	+0.280	
C-terminal N-CH ₃	C	-0.269	-0.365	-0.415		
	H1	+0.142	+0.181	+0.186		
	H2	+0.142	+0.176	+0.181		
	H3	+0.133	+0.160	+0.165		
	Sum	+0.148	+0.152	+0.117	0.000	

Table 11.11: Isoleucine charges calculated by Spartan '04.

Molecule	Group	Atom	Partial charges (e)			GROMOS96
			AM1	HF 6-31G**	DFT (B3LYP) 6-31G**	
Isoleucine	Acetyl CH ₃	C	-0.581	-0.886	-0.844	
		H1	+0.163	+0.226	+0.214	
		H2	+0.183	+0.250	+0.206	
		H3	+0.155	+0.218	+0.238	
		Sum	-0.080	-0.192	-0.186	0.000
	N-terminal amide	C	+0.697	+0.969	+0.812	+0.380
		O	-0.561	-0.657	-0.561	-0.380
		N	-0.505	-0.709	-0.578	-0.280
		H	+0.307	+0.400	+0.366	+0.280
	C _α H	C	-0.082	-0.112	-0.159	
		H	+0.130	+0.122	+0.114	
		Sum	+0.048	+0.010	-0.045	0.000
	C _β H	C	-0.002	+0.355	+0.335	
		H	+0.088	+0.001	+0.005	
		Sum	+0.086	+0.356	+0.340	0.000
	C _γ H ₃	C	-0.499	-0.851	-0.798	
		H1	+0.145	+0.206	+0.196	
		H2	+0.146	+0.220	+0.209	
		H3	+0.164	+0.208	+0.195	
		Sum	-0.044	-0.217	-0.198	
	C _γ H ₂	C	-0.148	+0.017	0.000	
		H1	+0.094	+0.029	+0.062	
		H2	+0.104	+0.059	+0.032	
		Sum	+0.050	+0.105	+0.094	0.000
	C _δ H ₃	C	-0.483	-0.636	-0.613	
		H1	+0.171	+0.154	+0.180	
		H2	+0.136	+0.190	+0.152	
H3		+0.141	+0.154	+0.157		
Sum		-0.035	-0.138	-0.124		
C-terminal amide	C	+0.627	+0.740	+0.587	+0.380	
	O	-0.545	-0.610	-0.512	-0.380	
	N	-0.476	-0.588	-0.447	-0.280	
	H	+0.288	+0.354	+0.324	+0.280	
C-terminal N-CH ₃	C	-0.257	-0.306	-0.372		
	H1	+0.134	+0.162	+0.174		
	H2	+0.129	+0.150	+0.160		
	H3	+0.139	+0.164	+0.170		
	Sum	+0.145	+0.170	+0.132	0.000	

Table 11.12: Asparagine charges calculated by Spartan '04.

Molecule	Group	Atom	Partial charges (e)			GROMOS96
			AM1	HF 6-31G**	DFT (B3LYP) 6-31G**	
Asparagine	Acetyl CH ₃	C	-0.579	-0.841	-0.804	
		H1	+0.157	+0.212	+0.200	
		H2	+0.162	+0.220	+0.207	
		H3	+0.178	+0.240	+0.228	
		Sum	-0.082	-0.169	-0.169	0.000
	N-terminal amide	C	+0.708	+0.882	+0.732	+0.380
		O	-0.567	-0.635	-0.543	-0.380
		N	-0.477	-0.572	-0.449	-0.280
		H	+0.310	+0.367	+0.333	+0.280
	C α H	C	-0.252	-0.259	-0.301	
		H	+0.210	+0.211	+0.200	
		Sum	-0.042	-0.048	-0.101	0.000
	C β H ₂	C	-0.433	-0.493	-0.458	
		H1	+0.169	+0.227	+0.209	
		H2	+0.209	+0.171	+0.159	
		Sum	-0.055	-0.095	-0.090	0.000
	Sidechain amide	C	+0.670	+0.889	+0.726	+0.380
		O	-0.563	-0.641	-0.541	-0.380
		N	-0.662	-1.013	-0.872	-0.830
		H1	+0.327	+0.428	+0.411	+0.415
H2		+0.298	+0.450	+0.392	+0.415	
C-terminal amide	C	+0.696	+0.736	+0.600	+0.380	
	O	-0.558	-0.590	-0.500	-0.380	
	N	-0.449	-0.483	-0.354	-0.280	
	H	+0.304	+0.352	+0.318	+0.280	
C-terminal N-CH ₃	C	-0.243	-0.323	-0.383		
	H1	+0.121	+0.161	+0.174		
	H2	+0.131	+0.139	+0.149		
	H3	+0.134	+0.165	+0.167		
	Sum	+0.143	+0.142	+0.107	0.000	

Table 11.13: Serine charges calculated by Spartan '04.

Molecule	Group	Atom	Partial charges (<i>e</i>)			GROMOS96
			AM1	HF 6-31G**	DFT (B3LYP) 6-31G**	
Serine	Acetyl CH ₃	C	-0.547	-0.623	-0.600	
		H1	+0.181	+0.205	+0.173	
		H2	+0.162	+0.165	+0.199	
		H3	+0.163	+0.178	+0.150	
		Sum	-0.041	-0.075	-0.078	0.000
	N-terminal amide	C	+0.716	+0.875	+0.718	+0.380
		O	-0.556	-0.660	-0.562	-0.380
		N	-0.554	-0.852	-0.723	-0.280
		H	+0.291	+0.364	+0.331	+0.280
	C _α H	C	-0.065	+0.542	+0.500	
		H	+0.139	+0.028	+0.034	
		Sum	+0.074	+0.570	+0.534	0.000
	C _β H ₂	C	+0.042	-0.092	-0.161	
		H1	+0.078	+0.094	+0.102	
		H2	+0.088	+0.097	+0.100	
		Sum	+0.208	+0.099	+0.041	+0.150
	OH	O	-0.537	-0.689	-0.607	-0.548
		H	+0.349	+0.453	+0.422	+0.398
	C-terminal amide	C	+0.637	+0.598	+0.438	+0.380
		O	-0.553	-0.646	-0.544	-0.380
N		-0.443	-0.490	-0.360	-0.280	
H		+0.299	+0.353	+0.320	+0.280	
C-terminal N-CH ₃	C	-0.323	-0.410	-0.447		
	H1	+0.157	+0.171	+0.194		
	H2	+0.130	+0.194	+0.151		
	H3	+0.146	+0.146	+0.173		
	Sum	+0.110	+0.101	+0.071	0.000	

Table 11.14: Aspartate charges calculated by Spartan '04.

Molecule	Group	Atom	Partial charges (<i>e</i>)			
			AM1	HF 6-31G**	DFT (B3LYP) 6-31G**	GROMOS96
Aspartate	Acetyl CH ₃	C	-0.521	-0.817	-0.779	
		H1	+0.170	+0.200	+0.187	
		H2	+0.138	+0.225	+0.211	
		H3	+0.147	+0.213	+0.199	
		Sum	-0.066	-0.179	-0.182	0.000
	N-terminal amide	C	+0.695	+0.976	+0.810	+0.380
		O	-0.602	-0.705	-0.614	-0.380
		N	-0.557	-0.974	-0.817	-0.280
		H	+0.318	+0.419	+0.374	+0.280
	C α H	C	+0.164	+0.736	+0.629	
		H	+0.126	+0.039	+0.047	
		Sum	+0.290	+0.775	+0.676	0.000
	C β H ₂	C	-0.585	-0.607	-0.571	
H1		+0.168	+0.142	+0.132		
H2		+0.157	+0.146	+0.135		
Sum		-0.260	-0.319	-0.250		
Sidechain carboxylate	C	+0.723	+0.847	+0.667	+0.270	
	O1	-0.744	-0.787	-0.686	-0.635	
	O2	-0.775	-0.840	-0.731	-0.635	
C-terminal amide	C	+0.577	+0.473	+0.348	+0.380	
	O	-0.597	-0.659	-0.567	-0.380	
	N	-0.423	-0.483	-0.357	-0.280	
	H	+0.291	+0.348	+0.310	+0.280	
C-terminal N-CH ₃	C	-0.319	-0.347	-0.379		
	H1	+0.136	+0.149	+0.146		
	H2	+0.102	+0.099	+0.105		
	H3	+0.212	+0.207	+0.200		
	Sum	+0.131	+0.108	+0.072	0.000	

Table 11.15: Lysine charges calculated by Spartan '04.

Molecule	Group	Atom	Partial charges (e)			GROMOS96
			AM1	HF 6-31G**	DFT (B3LYP) 6-31G**	
Lysine	Acetyl CH ₃	C	-0.599	-0.681	-0.667	
		H1	+0.154	+0.221	+0.218	
		H2	+0.183	+0.211	+0.207	
		H3	+0.191	+0.165	+0.153	
		Sum	-0.071	-0.084	-0.089	0.000
	N-terminal amide	C	+0.703	+0.840	+0.691	+0.380
		O	-0.522	-0.626	-0.527	-0.380
		N	-0.563	-0.803	-0.684	-0.280
		H	+0.316	+0.404	+0.371	+0.280
	C _α H	C	-0.056	+0.294	+0.255	
		H	+0.112	+0.041	+0.042	
		Sum	+0.056	+0.335	+0.297	0.000
	C _β H ₂	C	-0.198	-0.252	-0.236	
		H1	+0.128	+0.116	+0.128	
		H2	+0.129	+0.133	+0.110	
		Sum	+0.059	-0.003	+0.002	0.000
	C _γ H ₂	C	-0.303	-0.322	-0.308	
		H1	+0.162	+0.124	+0.125	
		H2	+0.135	+0.123	+0.122	
		Sum	-0.006	-0.075	-0.061	0.000
	C _δ H ₂	C	-0.179	-0.118	-0.121	
		H1	+0.124	+0.088	+0.092	
		H2	+0.113	+0.104	+0.102	
		Sum	+0.058	+0.074	+0.073	0.000
	C _ε H ₂	C	-0.186	+0.173	+0.097	
		H1	+0.156	+0.083	+0.096	
		H2	+0.152	+0.098	+0.108	
Sum		+0.122	+0.354	+0.301	+0.127	
Sidechain amine	N	-0.018	-0.701	-0.588	+0.129	
	H1	+0.264	+0.422	+0.395	+0.248	
	H2	+0.270	+0.413	+0.387	+0.248	
	H3	+0.264	+0.414	+0.387	+0.248	
C-terminal amide	C	+0.631	+0.623	+0.470	+0.380	
	O	-0.551	-0.566	-0.475	-0.380	
	N	-0.484	-0.588	-0.452	-0.280	
	H	+0.305	+0.367	+0.338	+0.280	
C-terminal N-CH ₃	C	-0.241	-0.281	-0.335		
	H1	+0.144	+0.167	+0.172		
	H2	+0.139	+0.169	+0.177		
	H3	+0.127	+0.144	+0.152		
	Sum	+0.169	+0.199	+0.166	0.000	

charge calculated for the α -carbon. The AM1 method predicted a small negative charge ($-0.074 e$), while the HF and DFT methods both calculated large net positive charges ($+0.570 e$ and $+0.534 e$, respectively). The dipole on the hydroxyl group, however, was in close agreement with GROMOS96 43A1 parameters when calculated by all methods, with the closest being the AM1 method.

It is imperative to note that simple calculation of charges by any of the methods here would be insufficient for parameterizing new functional groups that are compatible with the GROMOS96 force fields. Topologies still must be validated using thermodynamic integration and an assessment of other condensed-phase criteria [174]. We do believe, however, that the application of any of the aforementioned methods, especially the semi-empirical AM1 methods discussed here, provides a reasonable starting point for deriving charges for novel functional groups, taking much of the guesswork out of the empirical derivation strategy. Assignment of known functional group charges to small molecules is appropriate in the case of GROMOS96 parameter sets, as they have been designed to contain transferable functional groups that have the same parameters (with respect to partial charges), regardless of the remainder of the molecule.

11.4 Conclusions

We have demonstrated the effects of incorrect small molecule topologies in a variety of simulated systems. We conclude that the automated PRODRG 2.5 server is a valuable tool for the preparation of small molecule topologies for molecular simulation. It is fast, versatile, and easy to use. Bonded parameters and atom types assigned by PRODRG are correctly assigned in all cases. Unfortunately, the resulting topologies often suffer from deficiencies in the charges and charge groups that are assigned. These parameters often deviate substantially from the functional groups defined in the GROMOS96 43A1 parameter set, with the same functional groups in different molecules having different charges, leading to incorrect behavior in a variety of well-characterized systems. In general, we find that hydrophobic groups parameterized by PRODRG are inappropriately assigned partial charges, making them overly hydrophilic. As a consequence, the amino acid alanine inappropriately diffused into the aqueous solvent of our biphasic water-cyclohexane systems and the aliphatic portion of the lysine sidechain experienced overly-attractive van der Waals interactions with water. Dipoles in polar uncharged and charged species were somewhat unpredictable, as in the case of asparagine and aspartate, the sidechain dipole was too large, resulting in increased hydrogen bonding, while for lysine and serine, the dipole of the relevant polar group was too small, decreasing the hydrogen bonding capacity of these molecules. Application of small partial charges to hydrophobic moieties of *p*-cresol and ethanol led to substantially reduced hydrogen bonding capacity and altered liquid structural properties.

Perhaps the most troubling result of all comes from the simulations of the UGM-FAD complex. The PRODRG topology for FAD appeared to give rise to reasonable configurations for FAD, despite some increased flexibility in the adenine moiety. However, the nature and strength of the interactions between FAD and UGM were significantly different between the GROMOS96 43A1-

generated FAD topology and the topology from the PRODRG server. Hydrogen bonds were significantly reduced, due in large part to a completely uncharged amine group, and the value of non-bonded interaction energies was drastically different. Further, the negative effects of the PRODRG charges and charge groups assignments extended to the dynamics of UGM itself, conferring greater flexibility to nearby residues relative to what should be expected with GROMOS96 43A1 parameters applied to FAD. Thus, the applicability of unrefined PRODRG topologies to high-throughput screening methods is questionable, at best.

The validity of any program that purports to produce small molecule topologies deserves some scrutiny, and the resulting topologies should always be rigorously validated before using these parameters in any simulation study. We caution all investigators to properly calculate partial charges for the molecules in their simulations, first by analogy to similar chemical entities already in the force field library, then by semi-empirical quantum mechanical calculations for any group not present in the force field library, and to always validate their small molecule topologies in accordance with the prescribed force field methodology. Topologies generated by PRODRG are increasingly used in diverse applications in the literature, but the results obtained may not be consistent with the parent force field, GROMOS96 43A1. That is, the results obtained may suffer from many of the limitations and inaccuracies we have reported here. This fact raises significant concerns about the widespread use of this automated tool in the absence of any further topology modification or validation. The original intent for the PRODRG server was to be used in conjunction with crystal structure refinement and energy minimization, and extrapolation to more complicated systems may not be appropriate in the absence of substantial topology refinement.

The group concept employed in the GROMOS force fields allows for versatile transferability of chemical moieties between molecules; that is, parameters of known functional groups can be applied to different molecules, independent of other structural differences between these species. Thus, it is reasonable to construct a topology for an arbitrary small molecule under the GROMOS force fields in a way that is consistent with the original force field derivation, with minimal effort, by piecing together existing building blocks. For functional groups that are not yet defined, fast charge calculation methods like AM1-BCC, for which extensive geometry optimization is not required, should suffice as a starting point for further topology refinement, but we emphasize that partial charges for any new chemical functional groups should be iteratively refined as necessary, and rigorously validated before use in production simulations.

11.5 Acknowledgments

We thank Michelle Oppenheimer for useful discussions regarding UGM and Prof. J. M. Tanko for allowing us the use of Spartan '04.

Chapter 12

Conclusions

The body of work assembled here sought to explore the behavior of the $A\beta$ peptide in both membrane and solution environments, to gain insight into the underlying phenomena that give rise to the toxicity observed in Alzheimer's disease. The principal goals of these studies were (1) to understand how $A\beta$ interacts with the plasma membrane and to propose a mechanism by which the peptide aggregates and is ultimately released into the extracellular environment and (2) to propose a mechanism by which the flavonoid morin destabilizes existing $A\beta$ aggregates and affects the early events in the aggregation cascade. The first goal led us to conduct simulations of one or two $A\beta_{40}$ peptides in a variety of lipid environments to assess lipid-peptide interplay. The second goal required an understanding of the structural features of $A\beta$ aggregates that give rise to their intrinsic stability, after which the interactions of morin with these species could be explored.

In Chapter 4, we determined that, in a model DPPC membrane, $A\beta_{40}$ remains inserted in the membrane but causes significant perturbation to the equilibrium lipid dynamics. We expanded upon these findings in Chapter 5, exploring numerous other model membranes. Most significantly, our findings led us to propose that ganglioside GM1 plays a specific role in facilitating the release of monomeric $A\beta_{40}$ from lipid raft environments enriched in this lipid, and that the presence of GM1 enhances the appearance of aggregation-prone β -strand structures. We propose that these findings help to resolve an inherent paradox in the amyloid hypothesis. That is, $A\beta$ exerts its toxicity by binding to, and inserting in, membranes, yet $A\beta$ is released from the plasma membrane through a previously unknown mechanism. Our results indicate that certain regions of the plasma membrane promote $A\beta$ release, while others are more susceptible to the structural perturbations it induces. Dimerization of $A\beta$, as explored in Chapter 6, occurs within the membrane and disfavors release, even in the presence of GM1. The frequency of dimerization events and the resulting structural perturbations induced on the membrane are dependent upon the lipid type, again suggesting that different regions of the plasma membrane have different susceptibilities to $A\beta$ toxicity.

In Chapter 7, we explored the various factors that promote the stability of mature $A\beta$ fibril precursors. We found that the Asp23-Lys28 salt bridge is critical to the stability of the protofibril structure, and that a narrow channel of water links these intermolecular salt bridges via a network

of water-mediated hydrogen bonds. The amount of water in the channel is limited by packing of Ile32 and Leu34, which, after deletion of their side chains, results in a large influx of water, an event that destabilizes the salt bridges. We determined in Chapter 8 that the Asp23-Lys28 salt bridge is the principal target of the flavonoid morin, and its binding to this location interferes with side chain interactions and backbone hydrogen bonding, leading to destabilization (but not complete dissociation) of the peptides in the pentameric protofibril. Further, the binding of morin on the surface of the protofibril prevents the attachment of an additional peptide to the protofibril structure, thus indicating that morin can impede linear growth of these assemblies.

Having addressed the impact of morin on higher-order (on-pathway) fibril intermediates, we sought in Chapter 9 to explore the effect on the early events in the aggregation cascade by examining the monomeric state of both $A\beta_{40}$ and $A\beta_{42}$ and their propensity to dimerize. We also assessed the stability of preformed dimers derived from fibril structures. We found that morin has a negligible effect on the secondary structure of either alloform of monomeric $A\beta$, but that tertiary contacts are significantly altered by the presence of morin, with the effect being particularly acute for hydrophobic contacts. These results suggest that morin impedes the formation of a collapsed hydrophobic nucleus in the $A\beta$ monomer. Further, morin alters the quaternary interactions between $A\beta$ peptides during dimerization, but again secondary structural elements are largely unaffected. Through this investigation, we also determined that $A\beta_{40}$ and $A\beta_{42}$ dimerize in distinct ways, with $A\beta_{40}$ being driven equally by polar and hydrophobic contacts. In contrast, $A\beta_{42}$ is driven more prominently by hydrophobic contacts in its extended C-terminal region. Preformed $A\beta$ dimers are largely resistant to the effects of morin treatment, indicating that once dimerization has occurred, it is very difficult to disrupt these structures.

The work described above required the creation of new membrane analysis software and the establishment of reliable methodology for deriving small molecule parameters. We called the new software program for membrane analysis “GridMAT-MD,” which was described in Chapter 10. This new program allows for detailed analysis of membrane properties and convenient visualization.

The use of small molecules in molecular simulations is a difficult task, requiring detailed parameterization. We conducted an analysis of a popular web server for this purpose in Chapter 11, and after finding its results inadequate, we proposed methodology for use by the simulation community in producing high-quality topologies for small molecules under the GROMOS96 family of force fields. We applied such methods in all simulations involving morin, thus giving us confidence that the results obtained are a suitable model of reality.

The simulation of biomolecular systems has advanced greatly since the first reported protein simulation some 35 years ago. It is now possible to simulate heterogeneous systems that contain hundreds of thousands, if not millions, of atoms over biologically relevant time scales. With advances in software algorithms and ever-improving hardware, the ability of simulations to complement experimental work will continue to grow. Experimental work on $A\beta$ is challenging due to the propensity of the peptide to aggregate in solution, but the use of simulations greatly augments available *in vitro* and *in vivo* data. Many questions certainly remain that can be addressed

by theoretical methods, including a more detailed understanding of A β -metal interactions and the means by which A β binds to ABAD, an event that may play a role in understanding the metabolic dysfunction observed throughout the course of Alzheimer's disease. Simulations of cross-linked, oxidized, or modified A β peptides would also present an interesting avenue of research into the progression of A β aggregation in an oxidizing environment. In the present work, we have also explored the interactions of A β with morin, though it would be of great interest to expand these studies to systems involving membranes to understand if morin can preserve membrane integrity or otherwise inhibit A β -lipid interactions. Finally, though high-resolution structural data for the γ -secretase complex is lacking, it would be interesting one day to explore the dynamic interplay between this enzyme complex and the surrounding lipid environment to understand if lipid composition may play a role in APP processing by altering the positioning or structure of the complex. All of these topics are well-suited for theoretical investigations, and it is hoped that the work presented here will serve as a useful framework for future investigators wishing to address them.

Bibliography

- [1] Rainulf A. Stelzmann, H. Norman Schnitzlein, and F. Reed Murtagh. An English Translation of Alzheimer's 1907 Paper, "Über eine eigenartige Erkankung der Hirnrinde". *Clinical Anatomy*, 8:429–431, 1995.
- [2] Jiaquan Xu, Kenneth D. Kochanek, and Betzaida Tejada-Vera. Deaths: Preliminary Data for 2007. *National Vital Statistics Reports*, 58(1):1–51, 2009.
- [3] Takaomi C. Saido and Nobuhisa Iwata. Metabolism of amyloid β peptide and pathogenesis of Alzheimer's disease: Towards presymptomatic diagnosis, prevention therapy. *Neuroscience Research*, 54:235–253, 2006.
- [4] *2011 Alzheimer's Disease Facts and Figures*. Alzheimer's Association, 2011.
- [5] Ron Brookmeyer, Elizabeth Johnson, Kathryn Ziegler-Graham, and H. Michael Arrighi. Forecasting the global burden of Alzheimer's disease. *Alzheimer's & Dementia*, 3:186–191, 2007.
- [6] Alan S. Cohen. *General introduction and a brief history of the amyloid fibril*, pages 3–19. Amyloidosis. Nijhoff, Dordrecht, 1986.
- [7] Jean D. Sipe and Alan S. Cohen. Review: History of the Amyloid Fibril. *Journal of Structural Biology*, 130:88–98, 2000.
- [8] Colin L. Masters, Gail Simms, Nicola A. Weinman, Gerd Multhaup, Brian L. McDonald, and Konrad Beyreuther. Amyloid plaque core protein in Alzheimer disease and Down syndrome. *Proceedings of the National Academy of Sciences of the United States of America*, 82:4245–4249, 1985.
- [9] Caine W. Wong, Vito Quaranta, and George G. Glenner. Neuritic Plaques and Cerebrovascular Amyloid in Alzheimer Disease are Antigenically Related. *Proceedings of the National Academy of Sciences of the United States of America*, 82:8729–8732, 1985.
- [10] John A. Hardy and Gerald A. Higgins. Alzheimer's Disease: The Amyloid Cascade Hypothesis. *Science*, 256(5054):184–185, 1992.

- [11] T.A. Bayer, R. Cappai, C.L. Masters, K. Beyreuther, and G. Multhaup. It all sticks together – the APP-related family of proteins and Alzheimer’s disease. *Molecular Psychiatry*, 4:524–528, 1999.
- [12] Gopal Thinakaran and Edward H. Koo. Amyloid Precursor Protein Trafficking, Processing, and Function. *Journal of Biological Chemistry*, 283(44):29615–29619, 2008.
- [13] Christina Priller, Thomas Bauer, Gerda Mitteregger, Bjarne Krebs, Hans A. Kretschmar, and Jochen Herms. Synapse Formation and Function Is Modulated by the Amyloid Precursor Protein. *Journal of Neuroscience*, 26(27):7212–7221, 2006.
- [14] Y.-D. Kwak, C.L. Brannen, T. Qu, H.M. Kim, X. Dong, P. Soba, A. Majumdar, A. Kaplan, K. Beyreuther, and K. Sugaya. Amyloid precursor protein regulates differentiation of human neural stem cells. *Stem Cells and Development*, 15(3):381–389, 2006.
- [15] Albert Y. Hung, Edward H. Koo, Christian Haass, and Dennis J. Selkoe. Increased expression of β -amyloid precursor protein during neuronal differentiation is not accompanied by secretory cleavage. *Proceedings of the National Academy of Sciences of the United States of America*, 89:9439–9443, 1992.
- [16] Richard J. O’Brien and Philip C. Wong. Amyloid Precursor Protein Processing and Alzheimer’s Disease. *Annual Review of Neuroscience*, 34:185–204, 2011.
- [17] Xinwei Cao and Thomas C. Südhof. A Transcriptionally Active Complex of APP with Fe65 and Histone Acetyltransferase Tip60. *Science*, 293:115–120, 2001.
- [18] Sung Hee Baek, Kenneth A. Ohgi, David W. Rose, Edward H. Koo, Christopher K. Glass, and Michael G. Rosenfeld. Exchange of N-CoR Corepressor and Tip60 Coactivator Complexes Links Gene Expression by NF- κ B and β -Amyloid Precursor Protein. *Cell*, 110(1):55–67, 2002.
- [19] Qiang Liu, Celina V. Zerbinatti, Juan Zhang, Hyang-Sook Hoe, Baiping Wang, Sarah L. Cole, Joachim Herz, Louis Muglia, and Guojun Bu. Amyloid Precursor Protein Regulates Brain Apolipoprotein E and Cholesterol Metabolism through Lipoprotein Receptor LRP1. *Neuron*, 56(1):66–78, 2007.
- [20] Yun-wu Zhang, Ruishan Wang, Qiang Liu, Francesca-Fang Liao, and Huaxi Xu. Presenilin/ γ -secretase-dependent processing of β -amyloid precursor protein regulates EGF receptor expression. *Proceedings of the National Academy of Sciences of the United States of America*, 104(25):10613–10618, 2007.
- [21] Raphaëlle Pardossi-Piquard, Agnès Petit, Toshitaka Kawarai, Claire Sunyach, Cristine Alves da Costa, Bruno Vincent, Sabine Ring, Luciano D’Adamio, Jie Shen, Ulrike Müller, Peter St. George Hyslop, and Frédéric Checler. Presenilin-Dependent Transcriptional Control of the A β -Degrading Enzyme Neprilysin by Intracellular Domains of β APP and APLP. *Neuron*, 46(4):541–554, 2005.

- [22] Andrew J. Beel, Charles K. Mobley, Hak Jun Kim, Fang Tian, Arina Hadziselimovic, Bing Jap, James H. Prestegard, and Charles R. Sanders. Structural Studies of the Transmembrane C-Terminal Domain of the Amyloid Precursor Protein (APP): Does APP Function as a Cholesterol Sensor? *Biochemistry*, 47(36):9428–9446, 2008.
- [23] Robert Vassar, Brian D. Bennett, Safura Babu-Khan, Steve Kahn, Elizabeth A. Mendiaz, Paul Denis, David B. Teplow, Sandra Ross, Patricia Amarante, Richard Loeloff, Yi Luo, Seth Fisher, Janis Fuller, Steven Edenson, Jackson Lile, Mark A. Jarosinski, Anja Leona Biere, Eileen Curran, Teresa Burgess, Jean-Claude Louis, Frank Collins, James Treanor, Gary Rogers, and Martin Citron. β -Secretase Cleavage of Alzheimer's Amyloid Precursor Protein by the Transmembrane Aspartic Protease BACE. *Science*, 286(22):735–741, 1999.
- [24] Sukanto Sinha, John P. Anderson, Robin Barbour, Guriqbal S. Basi, Russell Caccavello, David Davis, Minhtam Doan, Harry F. Dovey, Normand Frigon, Jin Hong, Kirsten Jacobson-Croak, Nancy Jewett, Pamela Keim, Jeroen Knops, Ivan Lieberburg, Michael Power, Hua Tan, Gwen Tatsuno, Jay Tung, Dale Schenk, Peter Seubert, Susanna M. Suomensaaari, Shuwen Wang, Donald Walker, Jun Zhao, Lisa McConlogue, and Varghese John. Purification and cloning of amyloid precursor protein β -secretase from human brain. *Nature*, 402:537–540, 1999.
- [25] Edward H. Koo and Sharon L. Squazzo. Evidence That Production and Release of Amyloid β -Protein Involves the Endocytic Pathway. *Journal of Biological Chemistry*, 269(26):17386–17389, 1994.
- [26] Dieter Edbauer, Edith Winkler, Joerg T. Regula, Brigitte Pesold, Harald Steiner, and Christian Haass. Reconstitution of γ -secretase activity. *Nature Cell Biology*, 5:486–488, 2003.
- [27] W. Taylor Kimberly, Matthew J. LaVoie, Beth L. Ostaszewski, Wenjuan Ye, Michael S. Wolfe, and Dennis J. Selkoe. γ -Secretase is a membrane protein complex comprised of presenilin, nicastrin, aph-1, and pen-2. *Proceedings of the National Academy of Sciences of the United States of America*, 100(11):6382–6387, 2003.
- [28] Harald Steiner, Regina Fluhner, and Christian Haass. Intramembrane Proteolysis by γ -Secretase. *Journal of Biological Chemistry*, 283(44):29627–29631, 2008.
- [29] Ji-Yeun Hur, Hedvig Welander, Homira Behbahani, Mikio Aoki, Jenny Frånberg, Bengt Winblad, Susane Frykman, and Lars O. Tjernberg. Active γ -secretase is localized to detergent-resistant membranes in human brain. *FEBS Journal*, 275:1174–1187, 2010.
- [30] Tobias M. J. Allinson, Edward T. Parkin, Anthony J. Turner, and Nigel M. Hooper. ADAMs Family Members As Amyloid Precursor Protein α -Secretases. *Journal of Neuroscience Research*, 74(3):342–352, 2003.

- [31] Tsuneya Ikezu, Bruce D. Trapp, Kenneth S. Song, Amnon Schlegel, Michael P. Lisanti, and Takashi Okamoto. Caveolae, plasma membrane microdomains for α -secretase-mediated processing of the amyloid precursor protein. *Journal of Biological Chemistry*, 273(17):10485–10495, 1998.
- [32] Asad Jan, Ozgun Gokce, Ruth Luthi-Carter, and Hilal A. Lashuel. The Ratio of Monomeric to Aggregated Forms of A β 40 and A β 42 Is an Important Determinant of Amyloid- β Aggregation, Fibrillogenesis, and Toxicity. *Journal of Biological Chemistry*, 283(42):28176–28189, 2008.
- [33] Akio Fukumori, Regina Fluhrer, Harald Steiner, and Christian Haass. Three-Amino Acid Spacing of Presenilin Endoproteolysis Suggests a General Stepwise Cleavage of γ -Secretase-Mediated Intramembrane Proteolysis. *Journal of Neuroscience*, 30(23):7853–7862, 2010.
- [34] Andreas Weidemann, Simone Eggert, Friedrich B. M. Reinhard, Markus Vogel, Krzysztof Paliga, Gottfried Baier, Colin L. Masters, Konrad Beyreuther, and Geneviève Evin. A Novel ϵ -Cleavage within the Transmembrane Domain of the Alzheimer Amyloid Precursor Protein Demonstrates Homology with Notch Processing. *Biochemistry*, 41(8):2825–2835, 2002.
- [35] Nobuto Kakuda, Satoru Funamoto, Sousuke Yagishita, Mako Takami, Satoko Osawa, Naoshi Dohmae, and Yasuo Ihara. Equimolar Production of Amyloid β -Protein and Amyloid Precursor Protein Intracellular Domain from β -Carboxyl-terminal Fragment by γ -Secretase. *Journal of Biological Chemistry*, 281(21):14776–14786, 2006.
- [36] Toru Sato, Naoshi Dohmae, Yue Qi, Nobuto Kakuda, Hiroaki Misonou, Rie Mitsumori, Hiroko Maruyama, Edward H. Koo, Christian Haass, Koji Takio, Maho Morishima-Kawashima, Shoichi Ishiura, and Yasuo Ihara. Potential Link between Amyloid β -Protein 42 and C-terminal Fragment γ 49-99 of β -Amyloid Precursor Protein. *Journal of Biological Chemistry*, 287(27):24294–20301, 2003.
- [37] Yue Qi-Takahara, Maho Morishima-Kawashima, Yu Tanimura, Georgia Dolios, Naoko Hirotsu, Yuko Horikoshi, Fuyuki Kametani, Masahiro Maeda, Takaomi C. Saido, Rong Wang, and Yasuo Ihara. Longer Forms of Amyloid β Protein: Implications for the Mechanism of Intramembrane Cleavage by γ -Secretase. *Journal of Neuroscience*, 25(2):436–445, 2005.
- [38] Satoru Funamoto, Maho Morishima-Kawashima, Yu Tanimura, Naoko Hirotsu, Takaomi C. Saido, and Yasuo Ihara. Truncated Carboxyl-Terminal Fragments of β -Amyloid Precursor Protein Are Processed to Amyloid β -Proteins 40 and 42. *Biochemistry*, 43(42):13532–13540, 2004.
- [39] Karen Duff, Chris Eckman, Cindy Zehr, Xin Yu, Cristian-Mihail Prada, Jordi Perez-tur, Mike Hutton, Luc Buee, Yasuo Harigaya, Debra Yager, David Morgan, Marcia N. Gordon, Leigh Holcomb, Lawrence Refolo, Brenda Zenk, John Hardy, and Steven Younkin.

- Increased amyloid- β 42(43) in brains of mice expressing mutant presenilin 1. *Nature*, 383(6602):710–713, 1996.
- [40] Martin Citron, David Westaway, Weiming Xia, George Carlson, Thekla Diehl, Georges Levesque, Kelly Johnson-Wood, Michael Lee, Peter Seubert, Angela Davis, Dora Kholodenko, Ruth Motter, Robin Sherrington, Billie Perry, Hong Yao, Robert Strome, Ivan Lieberburg, Johanna Rommens, Soyeon Kim, Dale Schenk, Paul Fraser, Peter St. George Hyslop, and Dennis J. Selkoe. Mutant presenilins of Alzheimer's disease increase production of 42-residue amyloid β -protein in both transfected cells and transgenic mice. *Nature Medicine*, 3(1):67–72, 1997.
- [41] D. Scheuner, C. Eckman, M. Jensen, X. Song, M. Citron, N. Suzuki, T. D. Bird, J. Hardy, M. Hutton, W. Kukull, E. Larson, E. Levy-Lahad, M. Viitanen, E. Peskin, P. Poorkaj, G. Schellenberg, R. Tanzi, W. Wasco, L. Lannfelt, D. Selkoe, and S. Younkin. Secreted amyloid β -protein similar to that in the senile plaques of Alzheimer's disease is increased *in vivo* by the presenilin 1 and 2 and *APP* mutations linked to familial Alzheimer's disease. *Nature Medicine*, 2(8):864–870, 1996.
- [42] Cynthia A. Lemere, Francisco Lopera, Kenneth S. Kosik, Corrine L. Lendon, Jorge Ossa, Takaomi C. Saido, Haruyasu Yamaguchi, Andres Ruiz, Alonso Martinez, Lucia Madrigal, Lilliana Hincapie, Juan Carlos Aragno L., Douglas C. Anthony, Edward H. Koo, Allison M. Goate, Dennis J. Selkoe, and Juan Carlos Arango V. The E280A presenilin 1 Alzheimer mutation produces increased $A\beta$ 42 deposition and severe cerebellar pathology. *Nature Medicine*, 2(10):1146–1150, 1996.
- [43] Weihong Song, Philip Nadeau, Menglan Yuan, Xudong Yang, Jie Shen, and Bruce A. Yankner. Proteolytic release and nuclear translocation of Notch-1 are induced by presenilin-1 and impaired by pathogenic presenilin-1 mutations. *Proceedings of the National Academy of Sciences of the United States of America*, 96(12):6959–6963, 1999.
- [44] Jie Shen and Raymond J. Kelleher III. The presenilin hypothesis of Alzheimer's disease: Evidence for a loss-of-function pathogenic mechanism. *Proceedings of the National Academy of Sciences of the United States of America*, 104(2):403–409, 2007.
- [45] Michael S. Wolfe. When loss is gain: reduced presenilin proteolytic function leads to increased $A\beta$ 42/ $A\beta$ 40: Talking Point on the role of presenilin mutations in Alzheimer disease. *EMBO Reports*, 8(2):136–140, 2007.
- [46] Bart De Strooper. Loss-of-function presenilin mutations in Alzheimer disease: Talking Point on the role of presenilin mutations in Alzheimer disease. *EMBO Reports*, 8(2):141–146, 2007.
- [47] Camilla Nilsberth, Anita Westlind-Danielsson, Christopher Eckman, Margaret M. Condrón, Karin Axelman, Charlotte Forsell, Charlotte Stenh, Johan Luthman, David B. Teplow, Steven G. Younkin, Jan Näslund, and Lars Lannfelt. The 'Arctic' APP mutation (E693G)

- causes Alzheimer's disease by enhanced $A\beta$ protofibril formation. *Nature Neuroscience*, 4(9):887–893, 2001.
- [48] Samir Kumar-Singh, Chris De Jonghe, Marc Cruets, Reinhold Kleinert, Rong Wang, Marc Mercken, Bart De Strooper, Hugo Vanderstichele, Ann Löfgren, Inge Vanderhoeven, Hubert Backhovens, Eugeen Vanmechelen, Peter M. Kroisel, and Christine Van Broeckhoven. Nonfibrillar diffuse amyloid deposition due to a γ_{42} -secretase site mutation points to an essential role for N-truncated $A\beta_{42}$ in Alzheimer's disease. *Human Molecular Genetics*, 9(18):2589–2598, 2000.
- [49] William E. Van Nostrand, Jerry P. Melchor, Hyn Soon Cho, Steven M. Greenberg, and G. William Rebeck. Pathogenic Effects of D23N Iowa Mutant Amyloid β -Protein. *Journal of Biological Chemistry*, 276(35):32860–32866, 2001.
- [50] Christopher B. Eckman, Nitin D. Mehta, Richard Crook, Jordi Perez-tur, Guy Prihar, Eric Pfeiffer, Neill Graff-Radford, Paul Hinder, Debra Yager, Brenda Zenk, Lawrence M. Refolo, Cristian Mihail Prada, Steven G. Younkin, Mike Hutton, and John Hardy. A new pathogenic mutation in the APP gene (I716V) increases the relative proportion of $A\beta_{42}(43)$. *Human Molecular Genetics*, 6(12):2087–2089, 1997.
- [51] Jill Murrell, Martin Farlow, Bernardino Ghetti, and Merrill D. Benson. A Mutation in the Amyloid Precursor Protein Associated with Hereditary Alzheimer's Disease. *Science*, 254(5028):97–99, 1991.
- [52] Yuichi Masuda, Kazuhiro Irie, Kazuma Murakami, Hajime Ohigashi, Ryutaro Ohashi, K. Takegoshi, Takahiko Shimizu, and Takuji Shirasawa. Verification of the turn at positions 22 and 23 of the β -amyloid fibrils with Italian mutation using solid-state NMR. *Bioorganic & Medicinal Chemistry*, 13:6803–6809, 2005.
- [53] Laura Obici, Andrea Demarchi, Giulia de Rosa, Vittorio Bellotti, Sabrina Marciano, Simona Donadei, Eloisa Arbustini, Giovanni Palladini, Marta Diegoli, Egidio Genovese, Giancarlo Ferrari, Sergio Coverlizza, and Giampaolo Merlini. A Novel $A\beta$ PP Mutation Exclusively Associated with Cerebral Amyloid Angiopathy. *Annals of Neurology*, 58(4):639–644, 2005.
- [54] Adam L. Cloe, Joseph P. R. O. Orgel, Joseph R. Sachleben, Robert Tycko, and Stephen C. Meredith. The Japanese Mutant $A\beta$ ($\Delta E22$ - $A\beta_{1-39}$) Forms Fibrils Instantaneously, with Low-Thioflavin T Fluorescence: Seeding Wild-Type $A\beta_{1-40}$ into Atypical Fibrils by $\Delta E22$ - $A\beta_{1-39}$. *Biochemistry*, 50(12):2026–2039, 2011.
- [55] Allison Goate, Marie-Christine Chartier-Harlin, Mike Mullan, Jeremy Brown, Fiona Crawford, Liana Fidani, Luis Giuffra, Andrew Haynes, Nick Irving, Louise James, Rebecca Mant, Phillippa Newton, Karen Rooke, Penelope Roques, Chris Talbot, Margaret Pericak-Vance, Allen Roses, Robert Williamson, Martin Rossor, Mike Owen, and John Hardy. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature*, 349:704–706, 1991.

- [56] Dieder Moechars, Ilse Dewachter, Kristin Lorent, Delphine Reversé, Veerle Baekelandt, Asha Naidu, Ina Tesseur, Kurt Spittaels, Chris Van Den Haute, Frédéric Checler, Emile Godaux, Barbara Cordell, and Fred Van Leuven. Early Phenotype Changes in Transgenic Mice That Overexpress Different Mutants of Amyloid Precursor Protein In Brain. *Journal of Biological Chemistry*, 274(10):6483–6492, 1999.
- [57] Cristina Guardia-Laguarta, Marta Pera, Jordi Clarimón, José Luis Molinuevo, Raquel Sánchez-Valle, Albert Lladó, Mireia Coma, Teresa Gómez-Isla, Rafael Blesa, Isidre Ferrer, and Alberto Lleó. Clinical, Neuropathologic, and Biochemical Profile of the Amyloid Precursor Protein I716F Mutation. *Journal of Neuropathology and Experimental Neurology*, 69(1):53–59, 2010.
- [58] Hidekuni Yamakawa, Sosuke Yagishita, Eugene Futai, and Shoichi Ishiura. β -Secretase Inhibitor Potency Is Decreased by Aberrant β -Cleavage Location of the “Swedish Mutant” Amyloid Precursor Protein. *Journal of Biological Chemistry*, 285(3):1634–1642, 2010.
- [59] Hugh A. Pearson and Chris Peers. Physiological roles for amyloid β peptides. *Journal of Physiology*, 575(1):5–10, 2006.
- [60] Flavio Kamenetz, Taisuke Tomita, Helen Hsieh, Guy Seabrook, David Borchelt, Takeshi Iwatsubo, Sangram Sisodia, and Roberto Malinow. APP Processing and Synaptic Function. *Neuron*, 37:925–937, 2003.
- [61] Joachim P. Steinbach, Ulrike Müller, Marcel Leist, Zhi-Wei Li, Pierluigi Nicotera, and Adriano Aguzzi. Hypersensitivity to seizures in β -amyloid precursor protein deficient mice. *Cell Death and Differentiation*, 5:858–866, 1998.
- [62] Ashley I. Bush. The metallobiology of Alzheimer’s disease. *Trends in Neurosciences*, 26(4):207–214, 2003.
- [63] Stephanie J. Soscia, James E. Kirby, Kevin J. Washicosky, Stephanie M. Tucker, Martin Ingelsson, Bradley Hyman, Mark A. Burton, Lee E. Goldstein, Scott Duong, Rudolph E. Tanzi, and Robert D. Moir. The Alzheimer’s Disease-Associated Amyloid β -Protein Is an Antimicrobial Peptide. *Public Library of Science ONE*, 5(3):e9505, 2010.
- [64] N. Iwata, H. Mizukami, K. Shirotani, Y. Takaki, S. Muramatsu, B. Lu, N. P. Gerard, C. Gerard, K. Ozawa, and T. C. Saido. Presynaptic localization of neprilysin contributes to efficient clearance of amyloid- β peptide in mouse brain. *Journal of Neuroscience*, 24:991–998, 2004.
- [65] E. A. Eckman, M. Watson, L. Marlow, K. Sambamurti, and C. B. Eckman. Alzheimer’s disease β -amyloid peptide is increased in mice deficient in endothelin-converting enzyme. *Journal of Biological Chemistry*, 278(4):2081–2084, 2003.
- [66] B. C. Miller, E. A. Eckman, K. Sambamurti, N. Dobbs, K. M. Chow, C. B. Eckman, L. B. Hersh, and D. L. Thiele. Amyloid- β peptide levels in brain are inversely correlated with

- insulysin activity levels in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 100:6221–6226, 2003.
- [67] Rudolph E. Tanzi and Lars Bertram. New Frontiers in Alzheimer's Disease Genetics. *Neuron*, 32(2):181–184, 2001.
- [68] G. Utermann, M. Hees, and A. Steinmetz. Polymorphism of apolipoprotein E and occurrence of dysbetalipoproteinaemia in man. *Nature*, 269(5629):604–607, 1977.
- [69] Robert W. Mahley and Stanley C. Rall Jr. Apolipoprotein E: Far More Than a Lipid Transport Protein. *Annual Review of Genomics and Human Genetics*, 1:507–537, 2000.
- [70] Berislav V. Zlokovic. Neurovascular pathways to neurodegeneration in Alzheimer's disease and other disorders. *Nature Reviews Neuroscience*, 12(12):723–738, 2011.
- [71] Luigi Puglielli, Rudolph E. Tanzi, and Dora M. Kovacs. Alzheimer's disease: the cholesterol connection. *Nature Reviews Neuroscience*, 6(4):345–351, 2003.
- [72] K. Fassbender, M. Simons, C. Bergmann, M. Stroick, D. Lütjohann, P. Keller, H. Runz, S. Kühl, T. Bertsch, K. von Bergmann, M. Hennerici, K. Beyreuther, and T. Hartmann. Simvastatin strongly reduces levels of Alzheimer's disease β -amyloid peptides A β 42 and A β 40 *in vitro* and *in vivo*. *Proceedings of the National Academy of Sciences of the United States of America*, 98(10):5856–5861, 2001.
- [73] Mikael Simons, Patrick Keller, Bart De Strooper, Konrad Beyreuther, Carlos G. Dotti, and Kai Simons. Cholesterol depletion inhibits the generation of β -amyloid in hippocampal neurons. *Proceedings of the National Academy of Sciences of the United States of America*, 95(11):6460–6464, 1998.
- [74] E. R. Frears, D. J. Stephens, C. E. Walters, H. Davies, and B. M. Austen. The role of cholesterol in the biosynthesis of β -amyloid. *Neuroreport*, 10(8):1699–1705, 1999.
- [75] Luigi Puglielli, Genevieve Konopka, Eunju Pack-Chung, Laura A. MacKenzie Ingano, Ok-sana Berezovska, Bradley T. Hyman, Ta Yuan Chang, Rudolph E. Tanzi, and Dora M. Kovacs. Acyl-coenzyme A: cholesterol acyltransferase modulates the generation of the amyloid β -peptide. *Nature Cell Biology*, 3(10):905–912, 2001.
- [76] Robert Ehehalt, Patrick Keller, Christian Haass, Christoph Thiele, and Kai Simons. Amyloidogenic processing of the Alzheimer β -amyloid precursor protein depends on lipid rafts. *Journal of Cell Biology*, 160(1):113–123, 2003.
- [77] Murray Coles, Wendy Bicknell, Andrew A. Watson, David P. Fairlie, and David J. Craik. Solution Structure of Amyloid β -Peptide(1-40) in a Water-Micelle Environment. Is the Membrane-Spanning Domain Where We Think It Is? *Biochemistry*, 37:11064–11077, 1998.

- [78] Orlando Crescenzi, Simona Tomaselli, Remo Guerrini, Severo Salvadori, Anna M. D'Ursi, Piero Andrea Temussi, and Delia Picone. Solution structure of the Alzheimer amyloid β -peptide (1-42) in an apolar microenvironment: Similarity with a virus fusion domain. *European Journal of Biochemistry*, 269:5642–5648, 2002.
- [79] Claudio Soto, Eduardo M. Castaño, Blas Frangione, and Nibaldo C. Inestrosa. The α -Helical to β -Strand Transition in the Amino-terminal Fragment of the Amyloid β -Peptide Modulates Amyloid Formation. *Journal of Biological Chemistry*, 270(7):3063–3067, 1995.
- [80] S. Zhang, K. Iwata, M. J. Lachenmann, J. W. Peng, E. R. Stimson, Y.-a. Lu, A. M. Felix, J. E. Maggio, and J. P. Lee. The Alzheimer's Peptide A β Adopts a Collapsed Coil Structure in Water. *Journal of Structural Biology*, 130:130–141, 2000.
- [81] Robin Roychaudhuri, Mingfeng Yang, Minako M. Hoshi, and David B. Teplow. Amyloid β -Protein Assembly and Alzheimer Disease. *Journal of Biological Chemistry*, 284(8):4749–4753, 2009.
- [82] Minako Hoshi, Michio Sato, Shinichiro Matsumoto, Akihiko Noguchi, Kaori Yasutake, Nat-suko Yoshida, and Kazuki Sato. Spherical aggregates of β -amyloid (amylospheroid) show high neurotoxicity and activate tau protein kinase I/glycogen synthase kinase-3 β . *Proceedings of the National Academy of Sciences of the United States of America*, 100(11):6370–6375, 2003.
- [83] Atul Deshpande, Erene Mina, Charles Glabe, and Jorge Busciglio. Different Conformations of Amyloid β Induce Neurotoxicity by Distinct Mechanisms in Human Cortical Neurons. *Journal of Neuroscience*, 26(22):6011–6018, 2006.
- [84] Rakez Kaye, Yuri Sokolov, Brian Edmonds, Theresa M. McIntire, Saskia C. Milton, James E. Hall, and Charles G. Glabe. Permeabilization of Lipid Bilayers Is a Common Conformation-dependent Activity of Soluble Amyloid Oligomers in Protein Misfolding Diseases. *Journal of Biological Chemistry*, 279(45):46363–46366, 2004.
- [85] Mark P. Mattson, Bin Cheng, Dave Davis, Karin Bryant, Ivan Lieberburg, and Russell E. Rydel. β -Amyloid Peptides Destabilize Calcium Homeostasis and Render Human Cortical Neurons Vulnerable to Excitotoxicity. *Journal of Neuroscience*, 12(2):376–389, 1992.
- [86] Paulius Cizas, Rima Budvytyte, Ramune Morkuniene, Radu Moldovan, Matteo Broccio, Mathias Lösche, Gediminas Niaura, Gintaras Valincius, and Vilmante Borutaite. Size-dependent neurotoxicity of β -amyloid oligomers. *Archives of Biochemistry and Biophysics*, 496:84–92, 2010.
- [87] Julia Tsai, Jaime Grutzendler, Karen Duff, and Wen-Biao Gan. Fibrillar amyloid deposition leads to local synaptic abnormalities and breakage of neuronal branches. *Nature Neuroscience*, 7(11):1181–1183, 2004.

- [88] Rakez Kaye, Elizabeth Head, Jennifer L. Thompson, Theresa M. McIntire, Saskia C. Milton, Carl W. Cotman, and Charles G. Glabe. Common Structure of Soluble Amyloid Oligomers Implies Common Mechanism of Pathogenesis. *Science*, 300:486–489, 2003.
- [89] James P. Cleary, Dominic M. Walsh, Jacki J. Hofmeister, Ganesh M. Shankar, Michael A. Kuskowski, Dennis J. Selkoe, and Karen H. Ashe. Natural oligomers of the amyloid- β protein specifically disrupt cognitive function. *Nature Neuroscience*, 8(1):79–84, 2004.
- [90] Ernesto E. Ambroggio, Dennis H. Kim, Frances Separovic, Colin J. Barrow, Kevin J. Barnham, Luis A. Bagatolli, and Gerardo D. Fidelio. Surface Behavior and Lipid Interaction of Alzheimer β -Amyloid Peptide 1-42: A Membrane-Disrupting Peptide. *Biophysical Journal*, 88(4):2706–2713, 2005.
- [91] Tong-Lay Lau, Ernesto E. Ambroggio, Deborah J. Tew, Roberto Cappai, Colin L. Masters, Gerardo D. Fidelio, Kevin J. Barnham, and Frances Separovic. Amyloid- β Peptide Disruption of Lipid Membranes and the Effect of Metal Ions. *Journal of Molecular Biology*, 356:759–770, 2006.
- [92] R. Preston Mason, Robert F. Jacob, Mary F. Walter, Pamela E. Mason, Nicolai A. Avdulov, Svetlana V. Chochina, Urule Igbavboa, and W. Gibson Wood. Distribution and Fluidizing Action of Soluble and Aggregated Amyloid β -Peptide in Rat Synaptic Plasma Membranes. *Journal of Biological Chemistry*, 274(26):18801–18807, 1999.
- [93] Dominic M. Walsh, Dean M. Hartley, Yoko Kusumoto, Youcef Fezoui, Margaret M. Condron, Aleksey Lomakin, George B. Benedek, Dennis J. Selkoe, and David B. Teplow. Amyloid β -Protein Fibrillogenesis: Structure and Biological Activity of Protofibrillar Intermediates. *Journal of Biological Chemistry*, 274(36):25945–25952, 1999.
- [94] Lan Zhang, Jian Zhong, Lixin Huang, Lijun Wang, Yuankai Hong, and Yinlin Sha. Parallel-Oriented Fibrogenesis of a β -Sheet Forming Peptide on Supported Lipid Bilayers. *Journal of Physical Chemistry B*, 112(30):8950–8954, 2008.
- [95] Robert M. Koffie, Melanie Meyer-Luehmann, Tadafumi Hashimoto, Kenneth W. Adams, Matthew L. Mielke, Monica Garcia-Alloza, Kristina D. Micheva, Stephen J. Smith, M. Leo Kim, Virginia M. Lee, Bradley T. Hyman, and Tara L. Spires-Jones. Oligomeric amyloid β associates with postsynaptic densities and correlates with excitatory synapse loss near senile plaques. *Proceedings of the National Academy of Sciences of the United States of America*, 106(10):4012–4017, 2009.
- [96] Laia Sánchez, Sergio Madurga, Tara Pukala, Marta Vilaseca, Carmen López-Iglesias, Carol V. Robinson, Ernest Giralt, and Natàlia Carulla. A β 40 and A β 42 Amyloid Fibrils Exhibit Distinct Molecular Recycling Properties. *Journal of the American Chemical Society*, 133(17):6505–6508, 2011.

- [97] Glenys A. Tennent, L. B. Lovat, and M. B. Pepys. Serum amyloid P component prevents proteolysis of the amyloid fibrils of Alzheimer disease and systemic amyloidosis. *Proceedings of the National Academy of Sciences of the United States of America*, 92:4299–4303, 1995.
- [98] Nelson Arispe, Eduardo Rojas, and Harvey B. Pollard. Alzheimer disease amyloid β protein forms calcium channels in bilayer membranes: Blockade by tromethamine and aluminum. *Proceedings of the National Academy of Sciences of the United States of America*, 90(2):567–571, 1993.
- [99] Arjan Quist, Ivo Doudevski, Hai Lin, Rushana Azimova, Douglas Ng, Blas Frangione, Bruce Kagan, Jorge Ghiso, and Ratnesh Lal. Amyloid ion channels: A common structural link for protein-misfolding disease. *Proceedings of the National Academy of Sciences of the United States of America*, 102(30):10427–10432, 2007.
- [100] Ratnesh Lal, Hai Lin, and Arjan P. Quist. Amyloid beta ion channel: 3D structure and relevance to amyloid channel paradigm. *Biochimica et Biophysica Acta*, 1768:1966–1975, 2007.
- [101] Hyunbum Jang, Fernando Teran Arce, Srinivasan Ramachandran, Ricardo Capone, Rushana Azimova, Bruce L. Kagan, Ruth Nussinov, and Ratnesh Lal. Truncated β -amyloid peptide channels provide an alternative mechanism for Alzheimer’s Disease and Down syndrome. *Proceedings of the National Academy of Sciences of the United States of America*, 107(14):6538–6543, 2010.
- [102] Hyunbum Jang, Fernando Teran Arce, Srinivasan Ramachandran, Ricardo Capone, Ratnesh Lal, and Ruth Nussinov. Structural Convergence Among Diverse, Toxic β -Sheet Ion Channels. *Journal of Physical Chemistry B*, 114(29):9445–9451, 2010.
- [103] Yuri Sokolov, J. Ashot Kozak, Rakez Kayed, Alexandr Chanturiya, Charles Glabe, and James E. Hall. Soluble Amyloid Oligomers Increase Bilayer Conductance by Altering Dielectric Structure. *Journal of General Physiology*, 128(6):637–647, 2006.
- [104] Gintaras Valincius, Frank Heinrich, Rima Budvytyte, David J. Vanderah, Duncan J. McGillivray, Yuri Sokolov, James E. Hall, and Mathias Lösche. Soluble Amyloid β -Oligomers Affect Dielectric Membrane Properties by Bilayer Insertion and Domain Formation: Implications for Cell Toxicity. *Biophysical Journal*, 95(10):4845–4861, 2008.
- [105] Ricardo Capone, Felipe Garcia Quiroz, Panchika Prangio, Inderjeet Saluja, Anna M. Sauer, Mahealani R. Bautista, Raymond S. Turner, Jerry Yang, and Michael Mayer. Amyloid- β -Induced Ion Flux in Artificial Lipid Bilayers and Neuronal Cells: Resolving a Controversy. *Neurotoxicity Research*, 16:1–13, 2009.
- [106] M. A. Lovell, J.D. Robertson, W. J. Teesdale, J. L. Campbell, and W. R. Markesbery. Copper, iron and zinc in Alzheimer’s disease senile plaques. *Journal of the Neurological Sciences*, 158(1):47–52, 1998.

- [107] Takashi Miura, Kiyoko Suzuki, Naohito Kohata, and Hideo Takeuchi. Metal Binding Modes of Alzheimer's Amyloid β -Peptide in Insoluble Aggregates and Soluble Complexes. *Biochemistry*, 39(23):7024–7031, 2000.
- [108] Victor A. Streltsov, Stephen J. Titmuss, V. Chandana Epa, Kevin J. Barnham, Colin L. Masters, and Joseph N. Varghese. The Structure of the Amyloid- β Peptide High-Affinity Copper II Binding Site in Alzheimer Disease. *Biophysical Journal*, 95(7):3447–3456, 2008.
- [109] Simon C. Drew, Colin L. Masters, and Kevin J. Barnham. Alanine-2 Carbonyl is an Oxygen Ligand in Cu^{2+} Coordination of Alzheimer's Disease Amyloid- β Peptide - Relevance to N-Terminally Truncated Forms. *Journal of the American Chemical Society*, 131(25):8760–8761, 2009.
- [110] Danielle G. Smith, Roberto Cappai, and Kevin J. Barnham. The redox chemistry of the Alzheimer's disease amyloid β peptide. *Biochimica et Biophysica Acta - Biomembranes*, 1768(8):1976–1990, 2007.
- [111] Ian V. J. Murray, Michael E. Sindoni, and Paul H. Axelsen. Promotion of Oxidative Lipid Membrane Damage by Amyloid β Proteins. *Biochemistry*, 44(37):12606–12613, 2005.
- [112] Hyoung-gon Lee, Xiongwei Zhu, Rudy J. Castellani, Akihiko Nonumura, George Perry, and Mark A. Smith. Amyloid- β in Alzheimer Disease: The Null versus the Alternate Hypotheses. *Journal of Pharmacology and Experimental Therapeutics*, 321(3):823–829, 2007.
- [113] Anatol Kontush. Alzheimer's Amyloid- β as a Preventative Antioxidant for Brain Lipoproteins. *Cellular and Molecular Neurobiology*, 21(4):299–315, 2001.
- [114] Anatol Kontush. Amyloid- β : An Antioxidant that Becomes a Pro-Oxidant and Critically Contributes to Alzheimer's Disease. *Free Radical Biology & Medicine*, 31(9):1120–1131, 2001.
- [115] Rozena Baruch-Suchodolsky and Bilha Fischer. Soluble Amyloid β_{1-28} -Copper(I)/Copper(II)/Iron(II) Complexes Are Potent Antioxidants in Cell-Free Systems. *Biochemistry*, 47(30):7796–7806, 2008.
- [116] Rebecca C. Nadal, Stephen E. J. Rigby, and John H. Viles. Amyloid β - Cu^{2+} Complexes in both Monomeric and Fibrillar Forms Do Not Generate H_2O_2 Catalytically but Quench Hydroxyl Radicals. *Biochemistry*, 47(44):11653–11664, 2008.
- [117] Rozena Baruch-Suchodolsky and Bilha Fischer. $\text{A}\beta_{40}$, either Soluble or Aggregated, Is a Remarkably Potent Antioxidant in Cell-Free Oxidative Systems. *Biochemistry*, 48(20):4354–4370, 2009.
- [118] Juliana Kardos, Ilona Kovács, Ferenc Hajós, Mihály Kálmán, and Miklós Simonyi. Nerve endings from rat brain tissue release copper upon depolarization. A possible role in regulating neuronal excitability. *Neuroscience Letters*, 103(2):139–144, 1989.

- [119] Paul S. Donnelly, Zhiguang Xiao, and Anthony G. Wedd. Copper and Alzheimer's disease. *Current Opinion in Chemical Biology*, 11(2):128–133, 2007.
- [120] Mark A. Lovell, Chengsong Xie, and William R. Markesbery. Protection against amyloid beta peptide toxicity by zinc. *Brain Research*, 823:88–95, 1999.
- [121] S. M. Cardoso, A. C. Rego, C. Pereira, and C. R. Oliveira. Protective Effect of Zinc on Amyloid- β 25-35 and 1-40 Mediated Toxicity. *Neurotoxicity Research*, 7(4):273–282, 2005.
- [122] Xudong Huang, Craig S. Atwood, Robert D. Moir, Mariana A. Hartshorn, Jean-Paul Vonsattel, Rudolph E. Tanzi, and Ashley I. Bush. Zinc-induced Alzheimer's A β 1-40 Aggregation Is Mediated by Conformational Factors. *Journal of Biological Chemistry*, 272(42):26464–26470, 1997.
- [123] Nasrollah Rezaei-Ghaleh, Karin Giller, Stefan Becker, and Markus Zweckstetter. Effect of Zinc Binding on β -Amyloid Structure and Dynamics: Implications for A β Aggregation. *Biophysical Journal*, 101(5):1202–1211, 2011.
- [124] Marian Valko, Dieter Leibfritz, Jan Moncol, Mark T. D. Cronin, Milan Mazur, and Joshua Telser. Free radicals and antioxidants in normal physiological functions and human disease. *The International Journal of Biochemistry & Cell Biology*, 39(1):44–84, 2007.
- [125] Math P. Cuajungco and Gordon J. Lees. Nitric oxide generators produce accumulation of chelatable zinc in hippocampal neuronal perikarya. *Brain Research*, 799:118–129, 1998.
- [126] Math P. Cuajungco and Kyle Y. Fagét. Zinc takes the center stage: its paradoxical role in Alzheimer's disease. *Brain Research Reviews*, 41:44–56, 2003.
- [127] J. R. Connor, S. L. Menzies, S. M. St. Martin, and E. J. Mufson. Cellular Distribution of Transferrin, Ferritin, and Iron in Normal and Aged Human Brains. *Journal of Neuroscience Research*, 7(4):595–611, 1990.
- [128] Luigi Zecca, Antonella Stroppolo, Alberto Gatti, Davide Tampellini, Marco Toscani, Mario Gallorini, Giuseppe Giaveri, Paolo Arosio, Paolo Santambrogio, Ruggero G. Fariello, Erdem Karatekin, Mark H. Kleinman, Nicholas Turro, Oleh Hornykiewicz, and Fabio A. Zucca. The role of iron and copper molecules in the neuronal vulnerability of locus coeruleus and substantia nigra during aging. *Proceedings of the National Academy of Sciences of the United States of America*, 101(26):9843–9848, 2004.
- [129] Beinan Liu, Aileen Moloney, Sarah Meehan, Kyle Morris, Sally E. Thomas, Louise C. Serpell, Robert Hider, Stefan J. Marciniak, David A. Lomas, and Damian C. Crowther. Iron Promotes the Toxicity of Amyloid β Peptide by Impeding Its Ordered Aggregation. *Journal of Biological Chemistry*, 286(6):4248–4256, 2011.
- [130] Xi Chen and Shi Du Yan. Mitochondrial A β : A Potential Cause of Metabolic Dysfunction in Alzheimer's Disease. *IUBMB Life*, 58(12):686–694, 2006.

- [131] P. Bubber, V. Haroutunian, G. Fisch, J. P. Blass, and G. E. Gibson. Mitochondrial abnormalities in Alzheimer brain: Mechanistic implications. *Annals of Neurology*, 57(5):695–703, 2005.
- [132] H. K. Anandatheerthavarada, G. Biswas, M. A. Robin, and N. G. Avadhani. Mitochondrial targeting and a novel transmembrane arrest of Alzheimer's precursor protein impairs mitochondrial function in neuronal cells. *Journal of Cell Biology*, 161(1):41–54, 2003.
- [133] Joyce W. Lustbader, Maurizio Cirilli, Chang Lin, Hong Wei Xu, Kazuhiro Takuma, Ning Wang, Casper Caspersen, Xi Chen, Susan Pollak, Michael Chaney, Fabrizio Trinchese, Shumin Liu, Frank Gunn-Moore, Lih-Fen Lue, Douglas G. Walker, Periannan Kuppusamy, Zay L. Zewier, Ottavio Arancio, David Stern, Shirley ShiDu Yan, and Hao Wu. ABAD Directly Links A β to Mitochondrial Toxicity in Alzheimer's Disease. *Science*, 304(5669):448–452, 2004.
- [134] D. Allan Butterfield, Tanea Reed, Shelley F. Newman, and Rukhsana Sultana. Roles of amyloid β -peptide-associated oxidative stress and brain protein modifications in the pathogenesis of Alzheimer's disease and mild cognitive impairment. *Free Radical Biology and Medicine*, 43:658–677, 2007.
- [135] Jan Näslund, Angelika Schierhorn, Ulf Hellman, Lars Lannfelt, Allen D. Roses, Lars O. Tjernberg, Jerzy Silberring, Samuel Gandy, Bengt Winblad, Paul Greengard, Christer Nordstedt, and Lars Terenius. Relative abundance of Alzheimer A β amyloid peptide variants in Alzheimer disease and normal aging. *Proceedings of the National Academy of Sciences of the United States of America*, 91(18):8378–8382, 1994.
- [136] Yu-Min Kuo, Tyler A. Kokjohn, Thomas G. Beach, Lucia I. Sue, Daniel Brune, John C. Lopez, Walter M. Kalback, Dorothee Abramowski, Christine Sturchler-Pierrat, Matthias Staufenbiel, and Alex E. Roher. Comparative Analysis of Amyloid- β Chemical Structure and Amyloid Plaque Morphology of Transgenic Mouse and Alzheimer's Disease Brains. *Journal of Biological Chemistry*, 276(16):12991–12998, 2001.
- [137] Jian Dong, Craig S. Atwood, Vernon E. Anderson, Sandra L. Siedlak, Mark A. Smith, George Perry, and Paul R. Carey. Metal Binding and Oxidation of Amyloid- β within Isolated Senile Plaque Cores: Raman Microscopic Evidence. *Biochemistry*, 42(10):2768–2773, 2003.
- [138] Angela M. Boutte, Randall L. Woltjer, Lisa J. Zimmerman, Sheryl L. Stamer, Kathleen S. Montine, Michael V. Manno, Patrick J. Cimino, Daniel C. Liebler, and Thomas J. Montine. Selectively increased oxidative modifications mapped to detergent-insoluble forms of A β and β -III tubulin in Alzheimer's disease. *The FASEB Journal*, 20(9):1473–1483, 2006.
- [139] Andrew A. Watson, David P. Fairlie, and David J. Craik. Solution Structure of Methionine-Oxidized Amyloid β -Peptide (1-40). Does Oxidation Affect Conformational Switching? *Biochemistry*, 37(37):12700–12706, 1998.

- [140] Bernd Seilheimer, Bernd Bohrmann, Luca Bondolfi, Franci Müller, Dietrich Stüber, and Heinz Döbeli. The Toxicity of the Alzheimer's β -Amyloid Peptide Correlates with a Distinct Fiber Morphology. *Journal of Structural Biology*, 119(1):59–71, 1997.
- [141] Seth W. Snyder, Uri S. Lador, Warren S. Wade, Gary T. Wang, Leo W. Barrett, Edmund D. Matayoshi, Helen Jan Huffaker, Grant A. Krafft, and Thomas F. Holzman. Amyloid- β Aggregation: Selective Inhibition of Aggregation in Mixtures of Amyloid with Different Chain Lengths. *Biophysical Journal*, 67(3):1216–1228, 1994.
- [142] Daniel S. Auld, Tom J. Kornecook, Stéphane Bastianetto, and Rémi Quirion. Alzheimer's disease and the basal forebrain cholinergic system: relations to β -amyloid peptides, cognition, and treatment strategies. *Progress in Neurobiology*, 68:209–245, 2002.
- [143] C. H. Rojas-Fernandez. Successful use of donepezil for the treatment of dementia with Lewy bodies. *Annals of Pharmacotherapy*, 35(2):202–205, 2001.
- [144] G. Xiong and P. M. Doraiswamy. Combination drug therapy for Alzheimer's disease: what is evidence-based, and what is not? *Geriatrics*, 60(6):22–26, 2005.
- [145] Bengt Winblad, Lena Kilander, Sture Eriksson, Lennart Minthon, Stellan Båtsman, Anna-Lena Wetterholm, Catarina Jansson-Blixt, and Anders Haglund. Donepezil in patients with severe Alzheimer's disease: double-blind, parallel-group, placebo-controlled study. *Lancet*, 367(9516):1057–1065, 2006.
- [146] <http://www.aricept.com/about.html>.
- [147] M. A. Rogawski and G. L. Wenk. The neuropharmacological basis for the use of memantine in the treatment of Alzheimer's disease. *CNS Drug Reviews*, 9(3):275–308, 2003.
- [148] R. Yan, M. J. Bienkowski, M. E. Shuck, H. Miao, M. C. Tory, A. M. Pauley, J. R. Brashier, N. C. Stratman, W. R. Mathews, and A. E. Buhl. Membrane-anchored aspartyl protease with Alzheimer's disease β -secretase activity. *Nature*, 402:533–537, 1999.
- [149] Shinobu Kitazume, Yuriko Tachida, Ritsuko Oka, Keiro Shirohara, Takaomi C. Saido, and Yasuhiro Hashimoto. Alzheimer's β -secretase, β -site amyloid precursor protein-cleaving enzyme, is responsible for cleavage secretion of a Golgi-resident sialyltransferase. *Proceedings of the National Academy of Sciences of the United States of America*, 98(24):13554–13559, 2001.
- [150] Stefan F. Lichtenthaler, Diana-ines Dominguez, Gil G. Westmeyer, Karina Reiss, Christian Haass, Paul Saftig, Bart De Strooper, and Brian Seed. The Cell Adhesion Protein P-selectin Glycoprotein Ligand-1 Is a Substrate for the Aspartyl Protease BACE1. *Journal of Biological Chemistry*, 278(49):48713–48719, 2003.
- [151] Qiming Li and Thomas C. Südhof. Cleavage of Amyloid- β Precursor Protein and Amyloid- β Precursor-like Protein by BACE 1. *Journal of Biological Chemistry*, 279(11):10542–10550, 2004.

- [152] Huaibin Cai, Yanshu Wang, Diane McCarthy, Hongjin Wen, David R. Borchelt, Donald L. Price, and Philip C. Wong. BACE1 is the major β -secretase for generation of A β peptides by neurons. *Nature Neuroscience*, 4(3):233–234, 2004.
- [153] Diana Dominguez, Jos Tournoy, Dieter Hartmann, Tobias Huth, Kim Cryns, Siska Deforce, Lutgarde Serneels, Ira Espuny Camacho, Els Marjaux, Katleen Craessarts, Anton J. M. Roebroek, Michael Schwake, Rudy D’Hooge, Patricia Bach, Ulrich Kalinke, Dieder Moechars, Christian Alzheimer, Karina Reiss, Paul Saftig, and Bart De Strooper. Phenotypic and Biochemical Analyses of BACE1- and BACE2-deficient Mice. *Journal of Biological Chemistry*, 280(35):30797–30806, 2005.
- [154] C. A. Saura, S. Y. Choi, V. Beglopoulos, S. Malkani, D. Zhang, B. S. Shankaranarayana Rao, S. Chattarji, R. J. Kelleher III, E. R. Kandel, K. Duff, A. Kirkwood, and J. Shen. Loss of presenilin function causes impairments of memory and synaptic plasticity followed by age-dependent neurodegeneration. *Neuron*, 31:713–726, 2004.
- [155] Clare E. Hunt and Anthony J. Turner. Cell biology, regulation and inhibition of β -secretase (BACE-1). *FEBS Journal*, 276(7):1845–1859, 2009.
- [156] Jason E. Gestewicki, Gerald R. Crabtree, and Isabella A. Graef. Harnessing Chaperones to Generate Small-Molecule Inhibitors of Amyloid β Aggregation. *Science*, 306:865–868, 2004.
- [157] Paul S. Aisen, Serge Gauthier, Bruno Vellas, Richard Briand, Daniel Saumier, Julie Laurin, and Denis Garceau. Alzhemed: A Potential Treatment for Alzheimer’s Disease. *Current Alzheimer Research*, 4:473–478, 2007.
- [158] Ismael Santa-Maria, Félix Hernández, Joaquín Del Rio, Francisco J. Moreno, and Jesús Avila. Tramiprosate, a drug of potential interest for the treatment of Alzheimer’s disease, promotes an abnormal aggregation of tau. *Molecular Neurodegeneration*, 2(1):17–28, 2007.
- [159] Kenjiro Ono, Tsuyoshi Hamaguchi, Hironobu Naiki, and Masahito Yamada. Anti-amyloidogenic effects of antioxidants: Implications for the prevention and therapeutics of Alzheimer’s disease. *Biochimica et Biophysica Acta*, 1762:575–586, 2006.
- [160] Kenjiro Ono, Yuji Yoshiike, Akihiko Takashima, Kazuhiro Hasegawa, Hironobu Naiki, and Masahito Yamada. Potent anti-amyloidogenic and fibril-destabilizing effects of polyphenols in vitro: implications for the prevention and therapeutics of Alzheimer’s disease. *Journal of Neurochemistry*, 87:172–181, 2003.
- [161] Kenjiro Ono, Margaret M. Condrón, Lap Ho, Jun Wang, Wei Zhao, Giulio M. Pasinetti, and David B. Teplow. Effects of Grape Seed-derived Polyphenols on Amyloid β -Protein Self-assembly and Cytotoxicity. *Journal of Biological Chemistry*, 283(47):32176–32187, 2008.

- [162] Lap Ho, Ling Hong Chen, Jun Wang, Wei Zhao, Stephen T. Talcott, Kenjiro Ono, David Teplow, Nelson Humala, Alice Cheng, Susan S. Percival, Mario Ferruzzi, Elsa Janle, Dara L. Dickstein, and Giulio Maria Pasinetti. Heterogeneity in Red Wine Polyphenolic Contents Differentially Influences Alzheimer's Disease-type Neuropathology and Cognitive Deterioration. *Journal of Alzheimer's Disease*, 16:59–72, 2009.
- [163] Bent H. Havsteen. The biochemistry and medical significance of the flavonoids. *Pharmacology & Therapeutics*, 96:67–202, 2002.
- [164] M. Born and J. R. Oppenheimer. On the Quantum Theory of Molecules. *Annals of Physics*, 84:457–484, 1927.
- [165] Loup Verlet. Computer “Experiments” on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Physical Review*, 159(1):98–103, 1967.
- [166] William C. Swope, Hans C. Andersen, Peter H. Berens, and Kent R. Wilson. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *Journal of Chemical Physics*, 76(1):637–649, 1982.
- [167] William L. Jorgensen, David S. Maxwell, and Julian Tirado-Rives. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *Journal of the American Chemical Society*, 118(45):11225–11236, 1996.
- [168] George A. Kaminski, Richard A. Friesner, Julian Tirado-Rives, and William L. Jorgensen. Evaluation and Reparameterization of the OPLS-AA Force Field for Proteins via Comparison with Accurate Quantum Chemical Calculations on Peptides. *Journal of Physical Chemistry B*, 105(28):6474–6487, 2001.
- [169] Wendy D. Cornell, Piotr Cieplak, Christopher I. Bayly, Ian R. Gould, Kenneth R. Merz, Jr., David M. Ferguson, David C. Spellmeyer, Thomas Fox, James W. Caldwell, and Peter A. Kollman. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society*, 117(19):5179–5197, 1995.
- [170] Junmei Wang, Piotr Cieplak, and Peter A. Kollman. How Well Does a Restrained Electrostatic Potential (RESP) Model Perform in Calculating Conformational Energies of Organic and Biological Molecules? *Journal of Computational Chemistry*, 21(12):1049–1074, 2000.
- [171] Yong Duan, Chun Wu, Shibasish Chowdhury, Mathew C. Lee, Guoming Xiong, Wei Zhang, Rong Yang, Piotr Cieplak, Ray Luo, Taisung Lee, James Caldwell, Junmei Wang, and Peter Kollman. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *Journal of Computational Chemistry*, 24(16):1999–2012, 2003.

- [172] A. D. MacKerell, Jr., D. Bashford, M. Bellot, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *Journal of Physical Chemistry B*, 102(18):3586–3616, 1998.
- [173] Alexander D. MacKerell, Jr., Michael Feig, and Charles L. Brooks III. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *Journal of Computational Chemistry*, 25(11):1400–1415, 2004.
- [174] Chris Oostenbrink, Alessandra Villa, Alan E. Mark, and Wilfred F. van Gunsteren. 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*, 25(13):1656–1676, 2004.
- [175] Siewert-Jan Marrink, H. Jelger Risselada, Serge Yefimov, D. Peter Tieleman, and Alex H. de Vries. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *Journal of Physical Chemistry B*, 111(27):7812–7824, 2007.
- [176] Luca Monticelli, Senthil K. Kandasamy, Xavier Periole, Ronald G. Larson, D. Peter Tieleman, and Siewert-Jan Marrink. The MARTINI Coarse-Grained Force Field: Extension to Proteins. *Journal of Chemical Theory and Computation*, 4(5):819–834, 2008.
- [177] J. Andrew McCammon, Bruce R. Gelin, and Martin Karplus. Dynamics of folded proteins. *Nature*, 267:585–590, 1977.
- [178] Oliver F. Lange, David van der Spoel, and Bert L. de Groot. Scrutinizing Molecular Mechanics Force Fields on the Submicrosecond Timescale with NMR Data. *Biophysical Journal*, 99(2):647–655, 2010.
- [179] David van der Spoel, Erik Lindahl, Berk Hess, Gerrit Groenhof, Alan E. Mark, and Herman J. C. Berendsen. GROMACS: Fast, Flexible and Free. *Journal of Computational Chemistry*, 26:1701–1718, 2005.
- [180] Berk Hess, Carsten Kutzner, David van der Spoel, and Erik Lindahl. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *Journal of Chemical Theory and Computation*, 4(3):435–447, 2008.
- [181] Wilfred F. van Gunsteren and Herman J. C. Berendsen. *Groningen Molecular Simulation (GROMOS) Library Manual*. BIOMOS b.v., Groningen, 1987.

- [182] Wilfred F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott, and I. G. Tironi. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*. vdf Hochschulverlag AG an der ETH Zürich and BIOMOS b.v., Zürich, Groningen, 1996.
- [183] Lukas D. Schuler and Wilfred F. van Gunsteren. On the choice of dihedral angle potential energy functions for n-alkanes. *Molecular Simulation*, 25(5):301–319, 2000.
- [184] Lukas D. Schuler, Xavier Daura, and Wilfred F. van Gunsteren. An Improved GROMOS96 Force Field for Aliphatic Hydrocarbons in the Condensed Phase. *Journal of Computational Chemistry*, 22(11):1205–1218, 2001.
- [185] Pär Bjelkmar, Per Larsson, Michel A. Cuendet, Berk Hess, and Erik Lindahl. Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models. *Journal of Chemical Theory and Computation*, 6(2):459–466, 2010.
- [186] Eric J. Sorin and Vijay S. Pande. Exploring the Helix-Coil Transition via All-Atom Equilibrium Ensemble Simulations. *Biophysical Journal*, 88(4):2472–2493, 2005.
- [187] H. J. C. Berendsen, J. P. M. Postma, Wilfred F. van Gunsteren, A. DiNola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *Journal of Chemical Physics*, 81(8):3684–3690, 1984.
- [188] Shuichi Nosé. A unified formulation of the constant temperature molecular dynamics methods. *Journal of Chemical Physics*, 81(1):511–519, 1984.
- [189] S. Nosé. A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics*, 52:255–268, 1984.
- [190] William G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A: Atomic, Molecular, and Optical Physics*, 31(3):1695–1697, 1985.
- [191] M. Parrinello and A. Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, 52(12):7182–7190, 1981.
- [192] Shuichi Nosé and M. L. Klein. Constant pressure molecular dynamics for molecular systems. *Molecular Physics*, 50(5):1055–1076, 1983.
- [193] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *Journal of Chemical Physics*, 98(12):10089–10092, 1993.
- [194] Ulrich Essmann, Lalith Perera, Max L. Berkowitz, Tom Darden, Hsing Lee, and Lee G. Pedersen. A smooth particle mesh Ewald method. *Journal of Chemical Physics*, 103(19):8577–8593, 1995.

- [195] Victor Hornak, Robert Abel, Asim Okur, Bentley Strockbine, Adrian Roitberg, and Carlos Simmerling. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins: Structure, Function, & Bioinformatics*, 65:712–725, 2006.
- [196] Dirk Matthes and Bert L. de Groot. Secondary Structure Propensities in Peptide Folding Simulations: A Systematic Comparison of Molecular Mechanics Interaction Schemes. *Biophysical Journal*, 97(2):599–608, 2009.
- [197] Robert B. Best, Nicolae-Viorel Buchete, and Gerhard Hummer. Are Current Molecular Dynamics Force Fields too Helical? *Biophysical Journal*, 95(1):L07–L09, 2008.
- [198] R. Preston Mason, Janet D. Estermyer, Jeremiah F. Kelly, and Pamela E. Mason. Alzheimer’s Disease Amyloid β Peptide 25-35 Is Localized in the Membrane Hydrocarbon Core: X-Ray Diffraction Analysis. *Biochemical and Biophysical Research Communications*, 222(1):78–82, 1996.
- [199] Yechun Xu, Jianhua Shen, Xiaomin Luo, Weiliang Zhu, Kaixian Chen, Jianpeng Ma, and Hualiang Jiang. Conformational transition of amyloid β -peptide. *Proceedings of the National Academy of Sciences of the United States of America*, 102(15):5403–5407, 2005.
- [200] Isabella Daidone, Fabio Simona, Danilo Roccatano, Ricardo A. Broglia, Guido Tiana, Giorgio Colombo, and Alfred Di Nola. β -Hairpin Conformation of Fibrillogenic Peptides: Structure and $\alpha - \beta$ Transition Mechanism Revealed by Molecular Dynamics Simulations. *Proteins: Structure, Function, & Bioinformatics*, 57:198–204, 2004.
- [201] Edgar Luttmann and Gregor Fels. All-atom molecular dynamics studies of the full-length β -amyloid peptides. *Chemical Physics*, 323:138–147, 2006.
- [202] Nicolae-Viorel Buchete, Robert Tycko, and Gerhard Hummer. Molecular Dynamics Simulations of Alzheimer’s β -Amyloid Protofibrils. *Journal of Molecular Biology*, 353:804–821, 2006.
- [203] Dagmar Flöck, Stefano Colacino, Giorgio Colombo, and Alfred Di Nola. Misfolding of the Amyloid β -protein: A Molecular Dynamics Study. *Proteins: Structure, Function & Bioinformatics*, 62:183–192, 2006.
- [204] Bogdan Tarus, John E. Straub, and D. Thirumalai. Dynamics of Asp23-Lys28 Salt-Bridge Formation in $A\beta_{10-35}$ Monomers. *Journal of the American Chemical Society*, 128(50):16159–16168, 2006.
- [205] David L. Mobley, Daniel L. Cox, Rajiv R. P. Singh, Michael W. Maddox, and Marjorie L. Longo. Modeling Amyloid β -Peptide Insertion into Lipid Bilayers. *Biophysical Journal*, 86(6):3585–3597, 2004.

- [206] Edmund Tischer and Barbara Cordell. β -Amyloid Precursor Protein: Location of Transmembrane Domain and Specificity of γ -Secretase Cleavage. *Journal of Biological Chemistry*, 271(36):21914–21919, 1996.
- [207] Cyril C. Curtain, Fedá E. Alia, Danielle G. Smith, Ashley I. Bush, Colin L. Masters, and Kevin J. Barnham. Metal Ions, pH, and Cholesterol Regulate the Interactions of Alzheimer's Disease Amyloid- β Peptide with Membrane Lipid. *Journal of Biological Chemistry*, 278(5):2977–2982, 2003.
- [208] Christopher M. Yip and JoAnne McLaurin. Amyloid- β Peptide Assembly: A Critical Step in Fibrillogenesis and Membrane Disruption. *Biophysical Journal*, 80:1359–1371, 2001.
- [209] David Eliezer. Amyloid Ion Channels: A Porous Argument or a Thin Excuse? *Journal of General Physiology*, 128(6):631–633, 2006.
- [210] Angelo Demuro, Erene Mina, Rakez Kaye, Saskia C. Milton, Ian Parker, and Charles G. Glabe. Calcium Dysregulation and Membrane Disruption as a Ubiquitous Neurotoxic Mechanism of Soluble Amyloid Oligomers. *Journal of Biological Chemistry*, 280(17):17294–17300, 2005.
- [211] Luis Cruz, Brigita Urbanc, Jose M. Borreguero, Noel D. Lazo, David B. Teplow, and H. Eugene Stanley. Solvent and mutation effects on the nucleation of amyloid β -protein folding. *Proceedings of the National Academy of Sciences of the United States of America*, 102(51):18258–18263, 2005.
- [212] Rainer A. Böckmann, Agnieszka Hac, Thomas Heimburg, and Helmut Grubmüller. Effects of Sodium Chloride on a Lipid Bilayer. *Biophysical Journal*, 85(3):1647–1655, 2003.
- [213] Kan Ma, Erin L. Clancy, Yongbo Zhang, Dale G. Ray, Kurt Wollenberg, and Michael G. Zagorski. Residue-Specific pK_a Measurements of the β -Peptide and Mechanism of pH-Induced Amyloid Formation. *Journal of the American Chemical Society*, 121(38):8698–8706, 1999.
- [214] <http://biophysics.cs.vt.edu/H++/index.php>.
- [215] John C. Gordon, Jonathan B. Myers, Timothy Folta, Valia Shoja, Lenwood S. Heath, and Alexey Onufriev. H++: a server for estimating pK_a s and adding missing hydrogens to macromolecules. *Nucleic Acids Research*, 33(Web Server issue):W368–W371, 2005.
- [216] Tomoshi Kameda and Shoji Takada. Secondary structure provides a template for the folding of nearby polypeptides. *Proceedings of the National Academy of Sciences of the United States of America*, 103(47):17765–17770, 2006.
- [217] Harry A. Stern and Scott E. Feller. Calculation of the dielectric permittivity profile for a nonuniform system: Application to a lipid bilayer simulation. *Journal of Chemical Physics*, 118(7):3401–3412, 2003.

- [218] Maria Francesca Iozzi, Maurizio Cossi, Roberto Improta, Nadia Rega, and Vincenzo Barone. A polarizable continuum approach for the study of heterogeneous dielectric environments. *Journal of Chemical Physics*, 124(184103):1–9, 2006.
- [219] http://moose.bio.ucalgary.ca/index.php?page=Structures_and_Topologies.
- [220] Michal Bachar and Oren M. Becker. Melittin at a membrane/water interface: Effects on water orientation and water penetration. *Journal of Chemical Physics*, 111(18):8672–8685, 1999.
- [221] Peter J. Bond and Mark S. P. Sansom. Bilayer deformation by the Kv channel voltage sensor domain revealed by self-assembly simulations. *Proceedings of the National Academy of Sciences of the United States of America*, 104(8):2631–2636, 2007.
- [222] Sudha Dorairaj and Toby W. Allen. On the thermodynamic stability of a charged arginine side chain in a transmembrane helix. *Proceedings of the National Academy of Sciences of the United States of America*, 104(12):4943–4948, 2007.
- [223] Anna C. V. Johansson and Erik Lindahl. Amino-Acid Solvation Structure in Transmembrane Helices from Molecular Dynamics Simulations. *Biophysical Journal*, 91(12):4450–4463, 2006.
- [224] Christian Haass and Dennis J. Selkoe. Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer’s amyloid β -peptide. *Nature Reviews Molecular Cell Biology*, 8:101–112, 2007.
- [225] Pravat K. Mandal and Jay W. Pettegrew. Alzheimer’s Disease: Soluble Oligomeric A β (1-40) Peptide in Membrane Mimic Environment from Solution NMR and Circular Dichroism Studies. *Neurochemical Research*, 29(12):2267–2272, 2004.
- [226] Maria Maddalena Sperotto, Sylvio May, and Artur Baumgaertner. Modelling of proteins in membranes. *Chemistry and Physics of Lipids*, 141:2–29, 2006.
- [227] Michal Bachar and Oren M. Becker. Protein-Induced Membrane Disorder: A Molecular Dynamics Study of Melittin in a Dipalmitoylphosphatidylcholine Bilayer. *Biophysical Journal*, 78:1359–1375, 2000.
- [228] D. Peter Tieleman, Mark S. P. Sansom, and Herman J. C. Berendsen. Alamethicin Helices in a Bilayer and in Solution: Molecular Dynamics Simulations. *Biophysical Journal*, 76:40–49, 1999.
- [229] Senthil K. Kandasamy and Ronald G. Larson. Molecular Dynamics Simulations of Model Trans-Membrane Peptides in Lipid Bilayers: A Systematic Investigation of Hydrophobic Mismatch. *Biophysical Journal*, 90:2326–2343, 2006.

- [230] Wonpil Im and Charles L. Brooks III. Interfacial folding and membrane insertion of designed peptides studied by molecular dynamics simulations. *Proceedings of the National Academy of Sciences of the United States of America*, 102(19):6771–6776, 2005.
- [231] Hugh Nymeyer, Thomas B. Woolf, and Angel E. Garcia. Folding Is Not Required for Bilayer Insertion: Replica Exchange Simulations of an α -Helical Peptide with an Explicit Lipid Bilayer. *Proteins: Structure, Function & Bioinformatics*, 59:783–790, 2005.
- [232] Buyong Ma and Ruth Nussinov. Stabilities and conformations of Alzheimer’s β -amyloid peptide oligomers ($A\beta_{16-22}$, $A\beta_{16-35}$, and $A\beta_{10-35}$): Sequence effects. *Proceedings of the National Academy of Sciences of the United States of America*, 99(22):14126–14131, 2002.
- [233] Sundeep S. Deol, Peter J. Bond, Carmen Domene, and Mark S. P. Sansom. Lipid-Protein Interactions of Integral Membrane Proteins: A Comparative Simulation Study. *Biophysical Journal*, 87:3737–3749, 2004.
- [234] Marcus Bokvist, Fredrick Lindström, Anthony Watts, and Gerhard Gröbner. Two Types of Alzheimer’s β -Amyloid (1-40) Peptide Membrane Interactions: Aggregation Preventing Transmembrane Anchoring *Versus* Accelerated Surface Fibril Formation. *Journal of Molecular Biology*, 335:1039–1049, 2004.
- [235] Atsuko Kakio, Sei-ichi Nishimoto, Katsuhiko Yanagisawa, Yasunori Kozutsumi, and Katsumi Matsuzaki. Interactions of Amyloid β -Protein with Various Gangliosides in Raft-Like Membranes: Importance of GM1 Ganglioside-Bound Form as an Endogenous Seed for Alzheimer Amyloid. *Biochemistry*, 41(23):7385–7390, 2002.
- [236] JoAnne McLaurin, Trudy Franklin, Xiaoqiong Zhang, Jianping Deng, and Paul E. Fraser. Interactions of Alzheimer amyloid- β peptides with glycosaminoglycans: Effects on fibril nucleation and growth. *European Journal of Biochemistry*, 266:1101–1110, 1999.
- [237] T. C. Terwilliger and D. Eisenberg. The structure of melittin. I. Structure determination and partial refinement. *Journal of Biological Chemistry*, 257:6010–6015, 1982.
- [238] Simon Bernèche, Mafalda Nina, and Benoît Roux. Molecular Dynamics Simulation of Melittin in a Dimyristoylphosphatidylcholine Bilayer Membrane. *Biophysical Journal*, 75:1603–1618, 1998.
- [239] William J. Allen, Justin A. Lemkul, and David R. Bevan. GridMAT-MD: A Grid-Based Membrane Analysis Tool for Use With Molecular Dynamics. *Journal of Computational Chemistry*, 30(12):1952–1958, 2009.
- [240] T. D. Bradrick, A. Philippetis, and S. Georghiou. Stopped-flow fluorometric study of the interaction of melittin with phospholipid bilayers: importance of the physical state of the bilayer and the acyl chain length. *Biophysical Journal*, 69:1999–2010, 1995.

- [241] J. P. Bradshaw, C. E. Dempsey, and A. Watts. A Combined X-Ray and Neutron Diffraction Study of Selectively Deuterated Melittin in Phospholipid Bilayers: Effect of pH. *Molecular Membrane Biology*, 11:79–86, 1994.
- [242] M. Monette and M. Lafleur. Modulation of melittin-induced lysis by surface charge density of membranes. *Biophysical Journal*, 68:187–195, 1995.
- [243] S. Ohki, E. Marcus, D. K. Sukumaran, and K. Arnold. Interaction of melittin with lipid membranes. *Biochimica et Biophysica Acta*, 1112:1–6, 1994.
- [244] D. Peter Tieleman, Lucy R. Forrest, Mark S. P. Sansom, and Herman J. C. Berendsen. Lipid Properties and the Orientation of Aromatic Residues in OmpF, Influenza M2, and Alamethicin Systems: Molecular Dynamics Simulations. *Biochemistry*, 37:17554–17561, 1998.
- [245] Jean-Paul Douliez, Alain Léonard, and Erick J. Dufourc. Restatement of Order Parameters in Biomembranes: Calculation of C-C Bond Order Parameters from C-D Quadrupolar Splittings. *Biophysical Journal*, 68:1727–1739, 1995.
- [246] D. P. Tieleman, S. J. Marrink, and H. J. C. Berendsen. A computer perspective of membranes: molecular dynamics studies of lipid bilayer systems. *Biochimica et Biophysica Acta*, 1331:235–270, 1997.
- [247] Justin A. Lemkul and David R. Bevan. A comparative molecular dynamics analysis of the amyloid β -peptide in a lipid bilayer. *Archives of Biochemistry and Biophysics*, 470(1):54–63, 2008.
- [248] Barbara A. Lewis and Donald M. Engelman. Lipid bilayer thickness varies linearly with acyl chain length in fluid phosphatidylcholine vesicles. *Journal of Molecular Biology*, 166(2):211–217, 1983.
- [249] J. F. Nagle. Area/lipid of bilayers from NMR. *Biophysical Journal*, 64:1476–1481, 1993.
- [250] John F. Nagle and Stephanie Tristram-Nagle. Structure of lipid bilayers. *Biochimica et Biophysica Acta*, 1469:159–195, 2000.
- [251] Sagar A. Pandit, David Bostick, and Max L. Berkowitz. Molecular Dynamics Simulation of a Dipalmitoylphosphatidylcholine Bilayer with NaCl. *Biophysical Journal*, 84:3743–3750, 2003.
- [252] Horia I. Petrache, Kechuan Tu, and John F. Nagle. Analysis of Simulated NMR Order Parameters for Lipid Bilayer Structure Determination. *Biophysical Journal*, 76:2479–2487, 1999.
- [253] Dennis J. Selkoe. Translating cell biology into therapeutic advances in Alzheimer’s disease. *Nature*, 399(Supp):A23–A31, 1999.

- [254] Mika Ikeda, Akio Kihara, and Yasuyuki Igarashi. Lipid Asymmetry of the Eukaryotic Plasma Membrane: Functions and Related Enzymes. *Biological & Pharmaceutical Bulletin*, 29(8):1542–1546, 2006.
- [255] Philippe F. Devaux. Static and dynamic lipid asymmetry in cell membranes. *Biochemistry*, 30(5):1163–1173, 1991.
- [256] Sandro Sonnino, Laura Mauri, Vann Chigorno, and Alessandro Prinetti. Gangliosides as components of lipid membrane domains. *Glycobiology*, 17(1):1R–13R, 2006.
- [257] Katsuhiko Yanagisawa, Asano Odaka, Nobuhiro Suzuki, and Yasuo Ihara. GM1 ganglioside-bound amyloid β -protein ($A\beta$): A possible form of preamyloid in Alzheimer's disease. *Nature Medicine*, 1(10):1062–1066, 1995.
- [258] Atsuko Kakio, Sei-ichi Nishimoto, Katsuhiko Yanagisawa, Yasunori Kozutsumi, and Katsumi Matsuzaki. Cholesterol-dependent Formation of GM1 Ganglioside-bound Amyloid β -Protein, an Endogenous Seed for Alzheimer Amyloid. *Journal of Biological Chemistry*, 276(27):24985–24990, 2001.
- [259] Katsumi Matsuzaki and Chiho Horikiri. Interactions of Amyloid β -Peptide (1-40) with Ganglioside-Containing Membranes. *Biochemistry*, 38(13):4137–4142, 1999.
- [260] I Mikhalyov, A. Olofsson, G. Gröbner, and L. B.-Å. Johansson. Designed Fluorescent Probes Reveal Interactions between Amyloid- β (1-40) Peptides and GM1 Gangliosides in Micelles and Lipid Vesicles. *Biophysical Journal*, 99(5):1510–1519, 2010.
- [261] Maho Utsumi, Yoshiki Yamaguchi, Hiroaki Sasakawa, Naoki Yamamoto, Katsuhiko Yanagisawa, and Koichi Kato. Up-and-down topological mode of amyloid β -peptide lying on hydrophilic/hydrophobic interface of ganglioside clusters. *Glycoconjugate Journal*, 26:999–1006, 2009.
- [262] Maho Yagi-Utsumi, Tomoshi Kameda, Yoshiki Yamaguchi, and Koichi Kato. NMR characterization of the interactions between lyso-GM1 aqueous micelles and amyloid β . *FEBS Letters*, 584:831–836, 2010.
- [263] Parag Mukhopadhyay, Luca Monticelli, and D. Peter Tieleman. Molecular Dynamics Simulation of a Palmitoyl-Oleoyl Phosphatidylserine Bilayer with Na^+ Counterions and NaCl. *Biophysical Journal*, 86(3):1601–1609, 2004.
- [264] Sukit Leekumjorn, Yifei Wu, Amadeu K. Sum, and Christina Chan. Experimental and Computational Studies Investigating Trehalose Protection of HepG2 Cells from Palmitate-Induced Toxicity. *Biophysical Journal*, 94(7):2869–2883, 2008.
- [265] Perttu S. Niemelä, Samuli Ollila, Marja T. Hyvönen, Mikko Karttunen, and Ilpo Vattulainen. Assessing the Nature of Lipid Raft Membranes. *Public Library of Science Computational Biology*, 3(2):0304–0312, 2007.

- [266] Oliver Berger, Olle Edholm, and Fritz Jähnig. Molecular Dynamics Simulations of a Fluid Bilayer of Dipalmitoylphosphatidylcholine at Full Hydration, Constant Pressure, and Constant Temperature. *Biophysical Journal*, 72:2002–2013, 1997.
- [267] Christian Kandt, Walter L. Ash, and D. Peter Tieleman. Setting up and running molecular dynamics simulations of membrane proteins. *Methods*, 41:475–488, 2007.
- [268] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans. Interaction models for water in relation to protein hydration. In B. Pullman, editor, *Intermolecular Forces*, pages 331–342. Reidel, Dordrecht, 1981.
- [269] Justin A. Lemkul and David R. Bevan. Perturbation of membranes by the amyloid β -peptide - a molecular dynamics study. *FEBS Journal*, 276(11):3060–3075, 2009.
- [270] Berk Hess. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *Journal of Chemical Theory and Computation*, 4(1):116–122, 2008.
- [271] W. Kabsch and C. Sander. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22:2577–2637, 1983.
- [272] Warren L. DeLano. The PyMOL Molecular Graphics System, 2009.
- [273] Eva Y. Chi, Canay Ege, Amy Winans, Jaroslaw Majewski, Guohui Wu, Kristian Kjaer, and Ka Yee C. Lee. Lipid membrane templates the ordering and induces the fibrillogenesis of Alzheimer's disease amyloid- β peptide. *Proteins: Structure, Function & Bioinformatics*, 72(1):1–24, 2008.
- [274] Canay Ege and Ka Yee C. Lee. Insertion of Alzheimer's A β 40 Peptide into Lipid Monolayers. *Biophysical Journal*, 87:1732–1740, 2004.
- [275] Marie-Paule Mingeot-Leclercq, Laurence Lins, Mariam Bensliman, Françoise Van Bambeke, Patrick Van Der Smissen, Jacques Peuvot, André Schanck, and Robert Brasseur. Membrane destabilization induced by β -amyloid peptide 29-42: Importance of the amino-terminus. *Chemistry and Physics of Lipids*, 120:57–74, 2002.
- [276] Liping Yu, Rohinton Edalji, John E. Harlan, Thomas F. Holzman, Ana Pereda Lopez, Boris Labkovsky, Heinz Hillen, Stefan Barghorn, Ulrich Ebert, Paul L. Richardson, Laura Miesbauer, Larry Solomon, Diane Bartley, Karl Walter, Robert W. Johnson, Philip J. Hajduk, and Edward T. Olejniczak. Structural Characterization of a Soluble Amyloid β -Peptide Oligomer. *Biochemistry*, 48(9):1870–1877, 2009.
- [277] Robert Tycko. Insights into the Amyloid Folding Problem from Solid-State NMR. *Biochemistry*, 42(11):3151–3159, 2003.

- [278] Thorsten Lührs, Christiane Ritter, Marc Adrian, Dominique Riek-Loher, Bernd Bohrmann, Heinz Döbeli, David Schubert, and Roland Riek. 3D structure of Alzheimer's amyloid- β (1-42) fibrils. *Proceedings of the National Academy of Sciences of the United States of America*, 102(48):17342–17347, 2005.
- [279] Elke Müller, Andrea Giehl, Günter Schwarzmann, Konrad Sandhoff, and Alfred Blume. Oriented 1,2-Dimyristoyl-sn-glycerol-3-phosphorylcholine/Ganglioside Membranes: A Fourier Transform Infrared Attenuated Total Reflection Spectroscopic Study. Band Assignments; Orientational, Hydrational, and Phase Behavior; and Effects of Ca^{2+} Binding. *Biophysical Journal*, 71:1400–1421, 1996.
- [280] Evelyne Terzi, Günter Hölzemann, and Joachim Seelig. Self-association of β -Amyloid Peptide (1-40) in Solution and Binding to Lipid Membranes. *Journal of Molecular Biology*, 252(5):633–642, 1995.
- [281] Toshio Ariga, Kiyooki Kobayashi, Akira Hasegawa, Makoto Kiso, Hideharu Ishida, and Tadashi Miyatake. Characterization of High-Affinity Binding between Gangliosides and Amyloid β -Protein. *Archives of Biochemistry and Biophysics*, 388(2):225–230, 2001.
- [282] Jacques Fantini and Nouara Yahi. Molecular Basis for the Glycosphingolipid-Binding Specificity of α -Synuclein: Key Role of Tyrosine-39 in Membrane Insertion. *Journal of Molecular Biology*, 408:654–669, 2011.
- [283] Stewart R. Durell, H. Robert Guy, Nelson Arispe, Eduardo Rojas, and Harvey B. Pollard. Theoretical Models of the Ion Channel Structure of Amyloid β -Protein. *Biophysical Journal*, 67:2137–2145, 1994.
- [284] JoAnne McLaurin and Avijit Chakrabarty. Characterization of the interactions of Alzheimer β -amyloid peptides with phospholipid membranes. *European Journal of Biochemistry*, 245:355–363, 1997.
- [285] Paul E. Fraser, Jack T. Nguyen, David T. Chin, and Daniel A. Kirschner. Effects of Sulfate Ions on Alzheimer β /A4 Peptide Assemblies: Implications for Amyloid Fibril-Proteoglycan Interactions. *Journal of Neurochemistry*, 59(4):1531–1540, 1992.
- [286] Kurt R. Brunden, Nancy J. Richter-Cook, KNishith Chaturvedi, and Robert C. A. Frederickson. pH-Dependent Binding of Synthetic β -Amyloid Peptides to Glycosaminoglycans. *Journal of Neurochemistry*, 61(6):2147–2154, 1993.
- [287] Allam S. Reddy, Aslin Izmitli, and J. J. de Pablo. Effect of trehalose on amyloid β (29-40)-membrane interaction. *Journal of Chemical Physics*, 131:085101, 2009.
- [288] Stéphane Ravault, Olivier Soubias, Olivier Saurel, Annick Thomas, Robert Bresseur, and Alain Milon. Fusogenic Alzheimer's peptide fragment A β (29-42) in interaction with lipid bilayers: Secondary structure, dynamics, and specific interaction with phosphatidyl

- ethanolamine polar heads as revealed by solid-state NMR. *Protein Science*, 14(5):1181–1189, 2005.
- [289] Charles H. Davis and Max L. Berkowitz. Interaction Between Amyloid- β (1-42) Peptide and Phospholipid Bilayers: A Molecular Dynamics Study. *Biophysical Journal*, 96(3):785–797, 2009.
- [290] Charles H. Davis and Max L. Berkowitz. Structure of the Amyloid- β (1-42) Monomer Absorbed To Model Phospholipid Bilayers: A Molecular Dynamics Study. *Journal of Physical Chemistry B*, 113(43):14480–1486, 2009.
- [291] Smita Mukherjee, Pramit Chowdhury, and Feng Gai. Effect of Dehydration on the Aggregation Kinetics of Two Amyloid Peptides. *Journal of Physical Chemistry B*, 113(2):531–535, 2009.
- [292] G. W. Roberts, S. M. Gentleman, A. Lynch, and D. I. Graham. β A4 amyloid protein deposition in brain after head trauma. *Lancet*, 338:1422–1423, 1991.
- [293] Giuseppina Tesco, Young Ho Koh, Eugene L. Kang, Andrew N. Cameron, Shinjita Das, Miguel Sena-Esteves, Mikko Hiltunen, Shao-Hua Yang, Zhenyu Zhong, Yong Shen, James W. Simpkins, and Rudolph E. Tanzi. Depletion of GGA3 Stabilizes BACE and Enhances β -Secretase Activity. *Neuron*, 54(5):721–737, 2007.
- [294] Justin A. Lemkul and David R. Bevan. Lipid Composition Influences the Release of Alzheimer’s Amyloid β -Peptide from Membranes. *Protein Science*, 20(9):1530–1545, 2011.
- [295] Li Na Zhao, See-Wing Chiu, Jérôme Benoit, Lock Yue Chew, and Yuguang Mu. Amyloid β Peptides Aggregation in a Mixed Membrane Bilayer: A Molecular Dynamics Study. *Journal of Physical Chemistry B*, 115(42):12247–12256, 2011.
- [296] Yi-Jiong Zhang, Jing-Ming Shi, Cai-Juan Bai, Han Wang, Hai-Yun Li, and Shang-Rong Ji. Intra-membrane Oligomerization and Extra-membrane Oligomerization of Amyloid- β Peptide Are Competing Processes as a Result of Distinct Patterns of Motif Interplay. *Journal of Biological Chemistry*, 287(1):748–756, 2012.
- [297] Heinrich Sticht, Peter Bayer, Dieter Willbold, Sonja Dames, Caroline Hilbich, Konrad Beyreuther, Rainer W. Frank, and Paul Rösch. Structure of amyloid A4-(1-40)-peptide of Alzheimer’s disease. *European Journal of Biochemistry*, 233:293–298, 1995.
- [298] Robert B. Best and Gerhard Hummer. Optimized Molecular Dynamics Force Fields Applied to the Helix-Coil Transition of Polypeptides. *Journal of Physical Chemistry B*, 113(26):9004–9015, 2009.

- [299] Liming Qiu, Creighton Buie, Andrew Reay, Mark W. Vaughn, and Kwan Hon Cheng. Molecular Dynamics Simulations Reveal the Protective Role of Cholesterol in β -Amyloid Protein-Induced Membrane Disruptions in Neuronal Membrane Mimics. *Journal of Physical Chemistry B*, 115(32):9795–9812, 2011.
- [300] Berk Hess, Henk Bekker, Herman J. C. Berendsen, and Johannes G. E. M. Fraaije. LINCS: A Linear Constraint Solver for Molecular Simulations. *Journal of Computational Chemistry*, 18(12):1463–1472, 1997.
- [301] C. Jarzynski. Nonequilibrium Equality for Free Energy Differences. *Physical Review Letters*, 78(14):2690–2693, 1997.
- [302] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman. The weighted histogram analysis method for free-energy calculations on biomolecules 1. The method. *Journal of Computational Chemistry*, 13(8):1011–1021, 1992.
- [303] Marcelo F. Masman, Ulrich L. M. Eisel, Imre G. Csizmadia, Botond Penke, Ricardo D. Enriz, Siewert Jan Marrink, and Paul G. M. Luiten. *In Silico* Study of Full-Length Amyloid β 1-42 Tri- and Penta-Oligomers in Solution. *Journal of Physical Chemistry B*, 113(34):11710–11719, 2009.
- [304] Yifat Miller, Buyong Ma, and Ruth Nussinov. Polymorphism of Alzheimer’s $A\beta_{17-42}$ (p3) Oligomers: The Importance of the Turn Location and Its Conformation. *Biophysical Journal*, 97(4):1168–1177, 2009.
- [305] Adrien Melquiond, Xiao Dong, Normand Mousseau, and Philippe Derreumaux. Role of the Region 23-28 in $A\beta$ Fibril Formation: Insights from Simulations of the Monomers and Dimers of Alzheimer’s Peptides $A\beta_{40}$ and $A\beta_{42}$. *Current Alzheimer Research*, 5(3):244–250, 2008.
- [306] <http://ambermd.org>.
- [307] Takako Takeda and Dmitri K. Klimov. Probing Energetics of $A\beta$ Fibril Elongation by Molecular Dynamics Simulations. *Biophysical Journal*, 96(11):4428–4437, 2009.
- [308] Takako Takeda and Dmitri K. Klimov. Replica Exchange Simulations of the Thermodynamics of $A\beta$ Fibril Growth. *Biophysical Journal*, 96(2):442–452, 2009.
- [309] Takako Takeda and Dmitri K. Klimov. Side Chain Interactions Can Impede Amyloid Fibril Growth: Replica Exchange Simulations of $A\beta$ Peptide Mutant. *Journal of Physical Chemistry B*, 113(35):11848–11857, 2009.
- [310] G. N. Patey and J. P. Valleau. Free-energy of spheres with dipoles - Monte-Carlo with multistage sampling. *Chemical Physics Letters*, 21:297–300, 1973.

- [311] G. M. Torrie and J. P. Valleau. Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid. *Chemical Physics Letters*, 28:578–581, 1974.
- [312] G. M. Torrie and J. P. Valleau. Non-physical sampling distributions in Monte-Carlo free-energy estimation - umbrella sampling. *Journal of Computational Physics*, 23:187–199, 1977.
- [313] <http://sbc.bioch.ox.ac.uk/oliver/software/>.
- [314] Nicolas Lux Fawzi, Kevin L. Kohlstedt, Yuka Okabe, and Teresa Head-Gordon. Protofibril Assemblies of the Arctic, Dutch, and Flemish Mutants of the Alzheimer's $A\beta_{1-40}$ Peptide. *Biophysical Journal*, 94(6):2007–2016, 2008.
- [315] Sergei Izrailev, Sergey Stpaniants, Barry Isralewitz, Dorina Kosztin, Hui Lu, Ferenc Molnar, Willy Wriggers, and Klaus Schulten. *Steered Molecular Dynamics*, volume 4 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, Berlin, 1998.
- [316] Jie Zheng, Hyunbum Jang, Buyong Ma, Chung-Jun Tsai, and Ruth Nussinov. Modeling the Alzheimer $A\beta_{17-42}$ Fibril Architecture: Tight Intermolecular Sheet-Sheet Association and Intramolecular Hydrated Cavities. *Biophysical Journal*, 93(9):3046–3057, 2007.
- [317] Jochen S. Hub and Bert L. de Groot. Does CO_2 Permeate through Aquaporin-1? *Biophysical Journal*, 91(3):842–848, 2006.
- [318] Bogdan Tarus, John E. Straub, and D. Thirumalai. Structures and Free-Energy Landscapes of the Wild Type and Mutants of the $A\beta_{21-30}$ Peptide Are Determined by an Interplay between Intra-peptide Electrostatic and Hydrophobic Interactions. *Journal of Molecular Biology*, 379:815–829, 2008.
- [319] Andrij Baumketner, Summer L. Bernstein, Thomas Wytttenbach, Noel D. Lazo, David B. Teplow, Michael T. Bowers, and Joan-Emma Shea. Structure of the 21-30 fragment of amyloid β -protein. *Protein Science*, 15:1239–1247, 2006.
- [320] Megan M. Murray, Mary Griffin Krone, Summer L. Bernstein, Andrij Baumketner, Margaret M. Condrón, Noel D. Lazo, David B. Teplow, Thomas Wytttenbach, Joan-Emma Shea, and Michael T. Bowers. Amyloid β -Protein: Experiment and Theory on the 21-30 Fragment. *Journal of Physical Chemistry B*, 113(17):6041–6046, 2009.
- [321] Nikolaos G. Sgourakis, Yilin Yan, Scot A. McCallum, Chunyu Wang, and Angel E. Garcia. The Alzheimer's Peptides $A\beta_{40}$ and $A\beta_{42}$ Adopt Distinct Conformations in Water: A Combined MD / NMR Study. *Journal of Molecular Biology*, 368:1448–1457, 2007.
- [322] Mingfeng Yang and David B. Teplow. Amyloid β -Protein Monomer Folding: Free-Energy Surfaces Reveal Alloform-Specific Differences. *Journal of Molecular Biology*, 384:450–464, 2008.

- [323] Mark S. Searle, Samuel R. Griffiths-Jones, and Henry Skinner-Smith. Energetics of Weak Interactions in a β -hairpin Peptide: Electrostatic and Hydrophobic Contributions to Stability from Lysine Salt Bridges. *Journal of the American Chemical Society*, 121(50):11615–11620, 1999.
- [324] Chad D. Tatko and Marcey L. Waters. Comparison of C-H \cdots π and Hydrophobic Interactions in a β -Hairpin Peptide: Impact on Stability and Specificity. *Journal of the American Chemical Society*, 126(7):2028–2034, 2004.
- [325] Sylvain Lesné, Ming Teng Koh, Linda Kotilinek, Rakez Kaye, Charles G. Glabe, Austin Yang, Michela Gallagher, and Karen H. Ashe. A specific amyloid- β protein assembly in the brain impairs memory. *Nature*, 440:352–357, 2006.
- [326] Pontus Wasling, Jonny Daborg, Ilse Riebe, My Andersson, Erik Portelius, Kaj Blennow, Eric Hanse, and Henrik Zetterberg. Synaptic Retrogenesis and Amyloid- β in Alzheimer's Disease. *Journal of Alzheimer's Disease*, 16(1):1–14, 2009.
- [327] Pasquale Picone, Rita Carrotta, Giovanna Montana, Maria Rita Nobile, Pier Luigi San Biagio, and Marta Di Carlo. A β Oligomers and Fibrillar Aggregates Induce Different Apoptotic Pathways in LAN5 Neuroblastoma Cell Cultures. *Biophysical Journal*, 96(10):4200–4211, 2009.
- [328] Aneta T. Petkova, Richard D. Leapman, Zhihong Guo, Wai-Ming Yau, Mark P. Mattson, and Robert Tycko. Self-Propagating, Molecular-Level Polymorphism in Alzheimer's β -Amyloid Fibrils. *Science*, 262:262–265, 2005.
- [329] Anant K. Paravastu, Richard D. Leapman, Wai-Ming Yau, and Robert Tycko. Molecular structural basis for polymorphism in Alzheimer's β -amyloid fibrils. *Proceedings of the National Academy of Sciences of the United States of America*, 105(47):18349–18354, 2008.
- [330] Anant K. Paravastu, Isam Qahwash, Richard D. Leapman, Stephen C. Meredith, and Robert Tycko. Seeded growth of β -amyloid fibrils from Alzheimer's brain-derived fibrils produces a distinct fibril structure. *Proceedings of the National Academy of Sciences of the United States of America*, 106(18):7443–7448, 2009.
- [331] B. Urbanc, L. Cruz, S. Yun, S. V. Buldyrev, G. Bitan, D. B. Teplow, and H. E. Stanley. *In silico* study of amyloid β -protein folding and oligomerization. *Proceedings of the National Academy of Sciences of the United States of America*, 101(50):17345–17350, 2004.
- [332] S. Yun, B. Urbanc, L. Cruz, G. Bitan, D. B. Teplow, and H. E. Stanley. Role of Electrostatic Interactions in Amyloid β -Protein (A β) Oligomer Formation: A Discrete Molecular Dynamics Study. *Biophysical Journal*, 92(11):4064–4077, 2007.
- [333] Anna Radzicka and Richard Wolfenden. Comparing the Polarities of the Amino Acids: Side-Chain Distribution Coefficients between the Vapor Phase, Cyclohexane, 1-Octanol, and Neutral Aqueous Solution. *Biochemistry*, 27(5):1664–1670, 1988.

- [334] Shuji Kitagawa, Hiromi Sakamoto, and Hiromi Tano. Inhibitory Effects of Flavonoids on Free Radical-Induced Hemolysis and Their Oxidative Effects on Hemoglobin. *Chemical & Pharmaceutical Bulletin*, 52(8):999–1001, 2004.
- [335] Shuji Kitagawa, Tomohiro Nabekura, Thomoharu Takahashi, Yutaka Nakamura, Hiromi Sakamoto, Hiromi Tano, Midori Hirai, and Go Tsukahara. Structure-Activity Relationships of the Inhibitory Effects of Flavonoids on P-Glycoprotein-Mediated Transport in KB-C2 Cells. *Biological & Pharmaceutical Bulletin*, 28(12):2274–2278, 2005.
- [336] Joseph A. Rothwell, Andrea J. Day, and Michael R. A. Morgan. Experimental Determination of Octanol-Water Partition Coefficients of Quercetin and Related Flavonoids. *Journal of Agricultural and Food Chemistry*, 53(11):4355–4360, 2005.
- [337] Stephen E. DeBolt and Peter A. Kollman. Investigation of Structure, Dynamics, and Solvation in 1-Octanol and Its Water-Saturated Solution: Molecular Dynamics and Free-Energy Perturbation Studies. *Journal of the American Chemical Society*, 117(19):5316–5340, 1995.
- [338] P. Sassi, M. Paolantoni, R. S. Cataliotti, F. Palombo, and A. Morresi. Water-alcohol mixtures: A spectroscopic study of the water-saturated 1-octanol solution. *Journal of Physical Chemistry B*, 108:19557–19565, 2004.
- [339] T. C. Beutler, A. E. Mark, R. C. van Schaik, P. R. Greber, and Wilfred F. van Gunsteren. Avoiding singularities and numerical instabilities in free energy calculations based on molecular simulations. *Chemical Physics Letters*, 222:529–539, 1994.
- [340] Berk Hess and Nico F. A. van der Vegt. Hydration Thermodynamic Properties of Amino Acid Analogues: A Systematic Comparison of Biomolecular Force Fields and Water Models. *Journal of Physical Chemistry B*, 110(35):17616–17626, 2006.
- [341] Wilfred F. van Gunsteren and H. J. C. Berendsen. A leap-frog algorithm for stochastic dynamics. *Molecular Simulation*, 1:173–185, 1988.
- [342] Lucia Rackova, Silvai Firakova, Daniela Kostalova, Milan Stefek, Ernest Sturdik, and Magdalena Majekova. Oxidation of liposomal membrane suppressed by flavonoids: Quantitative structure-activity relationship. *Bioorganic & Medicinal Chemistry*, 13:6477–6484, 2005.
- [343] A. V. Pogrebnyak and É. F. Stepanova. Quantum-Chemical Modeling of the Process of Drug Adsorption and the Immobilization of Phytocomponents in Medicinal Forms. *Pharmaceutical Chemistry Journal*, 37(7):36–40, 2003.
- [344] S. L. da Silva, A. da Silva, K. M. Honório, S. Marangoni, M. H. Toyama, and A. B. F. da Silva. The influence of electronic, steric and hydrophobic properties of flavonoid compounds in the inhibition of the xanthine oxidase. *Journal of Molecular Structure (Theochem)*, 684:1–7, 2004.

- [345] Chris Oostenbrink, Daniel Juchli, and Wilfred F. van Gunsteren. Amine Hydration: A United-Atom Force-Field Solution. *ChemPhysChem*, 6:1800–1804, 2005.
- [346] Justin A. Lemkul and David R. Bevan. Assessing the Stability of Alzheimer’s Amyloid Protofibrils Using Molecular Dynamics. *Journal of Physical Chemistry B*, 114(4):1652–1660, 2010.
- [347] E. Prabhu Raman, Takako Takeda, and Dmitri K. Klimov. Molecular Dynamics Simulations of Ibuprofen Binding to A β Peptides. *Biophysical Journal*, 97(7):2070–2079, 2009.
- [348] C line Riviere, Jean-Claude Delaunay, Fran oise Immel, Christophe Cullin, and Jean-Pierre Monti. The Polyphenol Piceid Destabilizes Preformed Amyloid Fibrils and Oligomers In Vitro: Hypothesis on Possible Molecular Mechanisms. *Neurochemical Research*, 34:1120–1128, 2009.
- [349] Jun Wang, Lap Ho, Wei Zhao, Kenjiro Ono, Clark Rosensweig, Linghong Chen, Nelson Humala, David B. Teplow, and Giulio M. Pasinetti. Grape-Derived Polyphenolics Prevent A β Oligomerization and Attenuate Cognitive Deterioration in a Mouse Model of Alzheimer’s Disease. *Journal of Neuroscience*, 28(25):6388–6392, 2008.
- [350] Marino Convertino, Riccardo Pellarin, Marco Catto, Angelo Carotti, and Amedeo Caflisch. 9,10-Anthraquinone hinders β -aggregation: How does a small molecule interfere with A β -peptide amyloid fibrillation. *Protein Science*, 18:792–800, 2009.
- [351] Xiaoyun Meng, Larissa A. Munishkina, Anthony L. Fink, and Vladimir N. Uversky. Molecular Mechanisms Underlying the Flavonoid-Induced Inhibition of α -Synuclein Fibrillation. *Biochemistry*, 48(34):8206–8224, 2009.
- [352] Yair Porat, Adel Abramowitz, and Ehud Gazit. Inhibition of Amyloid Fibril Formation by Polyphenols: Structural Similarity and Aromatic Interactions as a Common Inhibition Mechanism. *Chemical Biology & Drug Design*, 67:27–37, 2006.
- [353] M. F. Perutz, J. T. Finch, J. Berriman, and A. Lesk. Amyloid fibers are water-filled nanotubes. *Proceedings of the National Academy of Sciences of the United States of America*, 99(8):5591–5595, 2002.
- [354] Young-Ho Lee, Eri Chatani, Kenji Sasahara, Hironobu Naiki, and Yuji Goto. A Comprehensive Model for Packing and Hydration for Amyloid Fibrils of β_2 -Microglobulin. *Journal of Biological Chemistry*, 284(4):2169–2175, 2009.
- [355] Gonzalo R. Lamberto, Andr s Binolfi, Mar a L. Orcellet, Carlos W. Bertoncini, Markus Zweckstetter, Christian Griesinger, and Claudio O. Fern ndez. Structural and mechanistic basis behind the inhibitory interaction of PcTS on α -synuclein amyloid fibril formation. *Proceedings of the National Academy of Sciences of the United States of America*, 106(50):21057–21062, 2009.

- [356] Dongxiang Liu, Yechun Xu, Yu Feng, Hong Liu, Xu Shen, Kaixian Chen, Jianpeng Ma, and Hualiang Jiang. Inhibitor Discovery Targeting the Intermediate Structure of β -Amyloid Peptide on the Conformational Transition Pathway: Implications in the Aggregation Mechanism of β -Amyloid Peptide. *Biochemistry*, 45(36):10963–10972, 2006.
- [357] Fusheng Yang, Giselle P. Lim, Aynum N. Begum, Oliver J. Ubeda, Mychica R. Simmons, Surendra S. Ambegaokar, Pingping Chen, Rakez Kaye, Charles G. Glabe, Salley A. Frautschy, and Gregory M. Cole. Curcumin Inhibits Formation of Amyloid β Oligomers and Fibrils, Binds Plaques, and Reduces Amyloid in Vivo. *Journal of Biological Chemistry*, 280(7):5892–5901, 2005.
- [358] Justin A. Lemkul and David R. Bevan. Destabilizing Alzheimer’s $A\beta_{42}$ Protofibrils with Morin: Insights from Molecular Dynamics Simulations. *Biochemistry*, 49(18):3935–3946, 2010.
- [359] Fu-Feng Liu, Luo Ji, Xiao-Yan Dong, and Yan Sun. Molecular Insight into the Inhibition Effect of Trehalose on the Nucleation and Elongation of Amyloid β -Peptide Oligomers. *Journal of Physical Chemistry B*, 113(32):11320–11329, 2009.
- [360] Chewook Lee and Sihyun Ham. Characterizing Amyloid-Beta Protein Misfolding From Molecular Dynamics Simulations With Explicit Water. *Journal of Computational Chemistry*, 32(2):349–355, 2011.
- [361] Aneta T. Petkova, Wai-Ming Yau, and Robert Tycko. Experimental Constraints on Quaternary Structure in Alzheimer’s β -Amyloid Fibrils. *Biochemistry*, 45(2):498–512, 2006.
- [362] W. Zhang, T. Hou, C. Schafmeister, W. S. Ross, and D. A. Case. *LEaP and gleap*.
- [363] William Humphrey, Andrew Dalke, and Klaus Schulten. VMD: Visual Molecular Dynamics. *Journal of Molecular Graphics*, 14:33–38, 1996.
- [364] Fu-Feng Liu, Xiao-Yan Dong, Lizhong He, Anton P. J. Middelberg, and Yan Sun. Molecular Insight into Conformational Transition of Amyloid β -Peptide 42 Inhibited by (–)-Epigallocatechin-3-gallate Probed by Molecular Simulations. *Journal of Physical Chemistry B*, 115(41):11879–11887, 2011.
- [365] Lars O. Tjernberg, Jan Näslund, Fredrik Lindqvist, Jan Johansson, Anders R. Karlström, Johan Thyberg, Lars Terenius, and Christer Nordstedt. Arrest of β -Amyloid Fibril Formation by a Pentapeptide Ligand. *Journal of Biological Chemistry*, 271(15):8545–8548, 1996.
- [366] Joseph T. Jarrett and Peter T. Lansbury. Seeding “One-Dimensional Crystallization” of Amyloid: A Pathogenic Mechanism in Alzheimer’s Disease and Scrapie? *Cell*, 73:1055–1058, 1993.
- [367] James D. Harper and Peter T. Lansbury. Models of Amyloid Seeding in Alzheimer’s Disease and Scrapie: Mechanistic Truths and Physiological Consequences of the Time-Dependent Solubility of Amyloid Proteins. *Annual Review of Biochemistry*, 66:385–407, 1997.

- [368] Buyong Ma and Ruth Nussinov. Simulations as analytical tools to understand protein aggregation and predict amyloid conformation. *Current Opinion in Chemical Biology*, 10:445–452, 2006.
- [369] Dominic M. Walsh, Aleksey Lomakin, George B. Benedek, Margaret M. Condrón, and David B. Teplow. Amyloid β -Protein Fibrillogenesis: Detection of a Protofibrillar Intermediate. *Journal of Biological Chemistry*, 272(35):22364–22372, 1997.
- [370] Suman Nag, Bidyut Sarkar, Arkarup Bandyopadhyay, Bankanidhi Sahoo, Varun K. A. Sreenivasan, Mamata Kombrabail, Chandrakesan Muralidharan, and Sudipta Maiti. Nature of the Amyloid- β Monomer and the Monomer-Oligomer Equilibrium. *Journal of Biological Chemistry*, 286(16):13827–13833, 2011.
- [371] Francesca Massi, Jeff W. Peng, Jonathan P. Lee, and John E. Straub. Simulation Study of the Structure and Dynamics of the Alzheimer’s Amyloid Peptide Congener in Solution. *Biophysical Journal*, 80(1):31–44, 2001.
- [372] Duilio F. Raffa and Rauk Arvi. Molecular Dynamics Study of the Beta Amyloid Peptide of Alzheimer’s Disease and Its Divalent Copper Complexes. *Journal of Physical Chemistry B*, 111(14):3789–3799, 2007.
- [373] Andrij Baumketner and Joan-Emma Shea. Folding Landscapes of the Alzheimer Amyloid- β (12-28) Peptide. *Journal of Molecular Biology*, 362:567–579, 2006.
- [374] Aleksey Lomakin, Doo Soo Chung, George B. Benedek, Daniel A. Kirschner, and David B. Teplow. On the nucleation and growth of amyloid β -protein fibrils: Detection of nuclei and quantitation of rate constants. *Proceedings of the National Academy of Sciences of the United States of America*, 93:1125–1129, 1996.
- [375] Riccardo Pellarin and Amedeo Caffisch. Interpreting the Aggregation Kinetics of Amyloid Peptides. *Journal of Molecular Biology*, 360:882–892, 2006.
- [376] Nicolas Lux Fawzi, Yuka Okabe, Eng-Hui Yap, and Teresa Head-Gordon. Determining the Critical Nucleus and Mechanism of Fibril Elongation of the Alzheimer’s $A\beta_{1-40}$ Peptide. *Journal of Molecular Biology*, 365:535–550, 2007.
- [377] Gal Bitan, Marina D. Kirkitadze, Aleksey Lomakin, Sabrina S. Vollers, George B. Benedek, and David B. Teplow. Amyloid β -protein ($A\beta$) assembly: $A\beta_{40}$ and $A\beta_{42}$ oligomerize through distinct pathways. *Proceedings of the National Academy of Sciences of the United States of America*, 100(1):330–335, 2003.
- [378] Woojin Kim and Michael H. Hecht. Sequence Determinants of Enhanced Amyloidogenicity of Alzheimer $A\beta_{42}$ Peptide Relative to $A\beta_{40}$. *Journal of Biological Chemistry*, 280(41):35069–35076, 2005.

- [379] Ali Reza A. Ladiwala, Jonathan S. Dordick, and Peter M. Tessier. Aromatic Small Molecules Remodel Toxic Soluble Oligomers of Amyloid β through Three Independent Pathways. *Journal of Biological Chemistry*, 286(5):3209–3218, 2011.
- [380] Marino Convertino, Andreas Vitalis, and Amedeo Caffisch. Disordered Binding of Small Molecules to A β (12-28). *Journal of Biological Chemistry*, 286(48):41578–41588, 2011.
- [381] Jeffery B. Klauda, Norbert Kučerka, Bernard R. Brooks, Richard W. Pastor, and John F. Nagle. Simulation-Based Methods for Interpreting X-Ray Data from Lipid Bilayers. *Biophysical Journal*, 90(8):2796–2807, 2006.
- [382] Teemu Murtola, Mikko Kupiainen, Emma Falck, and Iipo Vattulainen. Conformational analysis of lipid molecules by self-organizing maps. *Journal of Chemical Physics*, 126(5):054707, 2007.
- [383] Wataru Shinoda and Susumu Okazaki. A Voronoi analysis of lipid area fluctuation in a bilayer. *Journal of Chemical Physics*, 109(4):1517–1521, 1998.
- [384] Bong June Sung and Arun Yethiraj. Lateral Diffusion of Proteins in the Plasma Membrane: Spatial Tessellation and Percolation Theory. *Journal of Physical Chemistry B*, 112(1):143–149, 2008.
- [385] Santi Esteban-Martín and Jesús Salgado. The Dynamic Orientation of Membrane-Bound Peptides: Bridging Simulations and Experiments. *Biophysical Journal*, 93(12):4278–4288, 2007.
- [386] Xavier Periole, Thomas Huber, Siewert-Jan Marrink, and Thomas P. Sakmar. G Protein-Coupled Receptors Self-Assemble in Dynamics Simulations of Model Bilayers. *Journal of the American Chemical Society*, 129(33):10126–10132, 2007.
- [387] D. Peter Tieleman, Justin L. MacCallum, Walter L. Ash, Christian Kandt, Zhitao Xu, and Luca Monticelli. Membrane protein simulations with a united-atom lipid and all-atom protein model: lipid-protein interactions, side chain transfer free energies and model proteins. *Journal of Physics: Condensed Matter*, 18:S1221–S1234, 2006.
- [388] Craig M. Shepherd, Hans J. Vogel, and D. Peter Tieleman. Interactions of the designed antimicrobial peptide MB21 and truncated dermaseptin S3 with lipid bilayers: molecular-dynamics simulations. *Biochemical Journal*, 370:233–243, 2003.
- [389] http://www.ks.uiuc.edu/Research/vmd/ mailing_list/.
- [390] <http://www.gnu.org/copyleft/gpl.html>.
- [391] <http://www.perl.com/>.

- [392] Eric F. Pettersen, Thomas D. Goddard, Conrad C. Huang, Gregory S. Couch, Daniel M. Greenblatt, Elaine C. Meng, and Thomas E. Ferrin. UCSF Chimera – A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13):1605–1612, 2004.
- [393] David A. Case, Thomas E. Cheatham, III, Tom Darden, Holger Gohlke, Ray Luo, Kenneth R. Merz, Jr., Alexey Onufriev, Carlos Simmerling, Bing Wang, and Robert J. Woods. The Amber Biomolecular Simulation Programs. *Journal of Computational Chemistry*, 26(16):1668–1688, 2005.
- [394] José D. Faraldo-Gómez, Graham R. Smith, and Mark S. P. Sansom. Setting up and optimization of membrane protein simulations. *European Biophysical Journal*, 31:217–227, 2002.
- [395] <http://www.arc.vt.edu/arc/SystemX/index.php>.
- [396] <http://www.matpack.de>.
- [397] D. P. Tieleman and H. J. C. Berendsen. Molecular dynamics simulations of a fully hydrated dipalmitoylphosphatidylcholine bilayer with different macroscopic boundary conditions and parameters. *Journal of Chemical Physics*, 105(11):4871–4880, 1996.
- [398] Eyal Neria, Stefan Fischer, and Martin Karplus. Simulation of activation free energies in molecular systems. *Journal of Chemical Physics*, 105(5):1902, 1996.
- [399] Xavier Daura, Alan E. Mark, and Wilfred F. van Gunsteren. Parametrization of Aliphatic CH_n United Atoms of GROMOS96 Force Field. *Journal of Computational Chemistry*, 19(5):535–547, 1997.
- [400] Walter R. P. Scott, Philippe H. Hünenberger, Ilario G. Tironi, Alan E. Mark, Salomon R. Billeter, Jens Fennen, Andrew E. Torda, Thomas Huber, Peter Krüger, and Wilfred F. van Gunsteren. The GROMOS Biomolecular Simulation Program Package. *Journal of Physical Chemistry A*, 103(19):3596–3607, 1999.
- [401] Junmei Wang, R. M. Wolf, J. W. Caldwell, Peter A. Kollman, and David A. Case. Development and testing of a general Amber force field. *Journal of Computational Chemistry*, 25(9):1157–1174, 2004.
- [402] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A. D. MacKerell Jr. CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *Journal of Computational Chemistry*, 31(4):671–690, 2010.

- [403] D. M. F. van Aalten, R. Bywater, J. B. C. Findlay, M. Hendlich, and R. W. W. Hooft. PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. *Journal of Computer-Aided Molecular Design*, 10:255–265, 1996.
- [404] A. W. Schüttelkopf and D. M. F. van Aalten. PRODRG - a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallographica Section D*, 60(8):1355–1363, 2004.
- [405] Cao Yang, Xiaolei Zhu, Jinyu Li, and Rongwei Shi. Exploration of the mechanism for LPFFD inhibiting the formation of β -sheet conformation of A β (1-42) in water. *Journal of Molecular Modeling*, 16:813–821, 2010.
- [406] Jinyu Li, Xiaolei Zhu, Cao Yang, and Rongwei Shi. Characterization of the binding of angiotensin II receptor blockers to human serum albumin using docking and molecular dynamics simulation. *Journal of Molecular Modeling*, 16:789–798, 2010.
- [407] Monika Sharma, Smriti Khanna, Gopalakrishnan Bulusu, and Abhijit Mitra. Comparative modeling of thioredoxin glutathione reductase from *Schistosoma mansoni*: A multifunctional target for antischistosomal therapy. *Journal of Molecular Graphics and Modelling*, 27:665–675, 2009.
- [408] Lihie Ben-Avraham Levin, Assaf Ganoth, Shay Amram, Esther Nachliel, Menachem Gutman, and Yossi Tsfadia. Insight into the interaction sites between fatty acid binding proteins and their ligands. *Journal of Molecular Modeling*, 16(5):929–938, 2010.
- [409] Mohan Babu Boggara, Antonio Faraone, and Ramanan Krishnamoorti. Effect of pH and Ibuprofen on the Phospholipid Bilayer Bending Modulus. *Journal of Physical Chemistry B*, 114(24):8061–8066, 2010.
- [410] Arnau Cordoní, Jesús Prades, Juan Frau, Oliver Vögler, Sérgio S. Funari, Juan J. Perez, Pablo V. Esribá, and Francisca Barceló. Interactions of fatty acids with phosphatidylethanolamine membranes: X-ray diffraction and molecular dynamics studies. *Journal of Lipid Research*, 51:1113–1124, 2010.
- [411] Mohan Babu Boggara and Ramanan Krishnamoorti. Partitioning of Nonsteroidal Antiinflammatory Drugs in Lipid Membranes: A Molecular Dynamics Simulation Study. *Biophysical Journal*, 98(4):586–595, 2010.
- [412] David C. Turner, Fuchang Yin, James T. Kindt, and Hailing Zhang. Molecular Dynamics Simulations of Glycocholate-Oleic Acid Mixed Micelle Assembly. *Langmuir*, 26(7):4687–4692, 2010.
- [413] G. Cilpa, M. T. Hyvönen, A. Koivuniemi, and M.-L. Riekkola. Atomistic Insight into Chondroitin-6-Sulfate Glycosaminoglycan Chain Through Quantum Mechanics Calculations and Molecular Dynamics Simulations. *Journal of Computational Chemistry*, 31(8):1670–1680, 2010.

- [414] Neha S. Gandhi and Ricardo L. Mancera. Can current force fields reproduce ring puckering in 2-O-sulfo- α -L-iduronic acid? A molecular dynamics simulation study. *Carbohydrate Research*, 345:689–695, 2010.
- [415] Rangana Sanjeeva and Samantha Weerasinghe. Development of a molecular mechanics force field for caffeine to investigate the interactions of caffeine in different solvent media. *Journal of Molecular Structure (Theochem)*, 944:116–123, 2010.
- [416] Konstantinos Beis, Velupillai Srikannathasan, Huanting Liu, Stephen W. B. Fullerton, Vicki A. Bamford, David A. R. Sanders, Chris Whitfield, Mike R. McNeil, and James H. Naismith. Crystal Structures of *Mycobacteria tuberculosis* and *Klebsiella pneumoniae* UDP-Galactopyranose Mutase in the Oxidised State and *Klebsiella pneumoniae* UDP-Galactopyranose Mutase in the (Active) Reduced State. *Journal of Molecular Biology*, 348:971–982, 2005.
- [417] David R. Lide, editor. *CRC Handbook of Chemistry and Physics*. CRC Press LLC, Boca Raton, 84th edition, 2004.
- [418] Laurence S. Richard, Carlos E. S. Bernardes, Hermínio P. Diogo, João P. Leal, and Manuel E. Minas da Piedade. Energetics of Cresols and of Methylphenoxy Radicals. *Journal of Physical Chemistry A*, 111:8741–8748, 2007.
- [419] Cuijuan Zhang and Xiaoning Yang. Molecular dynamics simulation of ethanol/water mixtures for structure and diffusion properties. *Fluid Phase Equilibria*, 231:1–10, 2005.
- [420] Johan Åqvist, Victor B. Luzhkov, and Bjørn O. Brandsal. Ligand Binding Affinities from MD Simulations. *Accounts of Chemical Research*, 35(6):358–365, 2002.
- [421] Araz Jakalian, Bruce L. Bush, David B. Jack, and Christopher I. Bayly. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: I. Method. *Journal of Computational Chemistry*, 21(2):132–146, 2000.
- [422] Araz Jakalian, David B. Jack, and Christopher I. Bayly. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: II. Parameterization and Validation. *Journal of Computational Chemistry*, 23(16):1623–1641, 2002.
- [423] Robert S. Mulliken. A New Electroaffinity Scale; Together with Data on Valence States and on Valence Ionization Potentials and Electron Affinities. *Journal of Chemical Physics*, 2(11):782–792, 1934.
- [424] Johann Gasteiger and Mario Marsili. Iterative partial equalization of orbital electronegativity - a rapid access to atomic charges. *Tetrahedron*, 36(22):3219–3228, 1980.
- [425] Michael J. S. Dewar, Eve G. Zoebisch, Eamonn F. Healy, and James J. P. Stewart. AM1: A New General Purpose Quantum Mechanical Molecular Model. *Journal of the American Chemical Society*, 107:3902–3909, 1985.

- [426] Attila Szabo and Neil S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, New York, 1996.
- [427] Axel D. Becke. Density-Functional Thermochemistry. III. The Role of Exact Change. *Journal of Chemical Physics*, 98(7):5648–5652, 1993.
- [428] L. E. Chirlian and M. M. Francl. Atomic charges derived from electrostatic potentials: a detailed study. *Journal of Computational Chemistry*, 8:894–905, 1987.

Appendix A

Cholesterol Topology

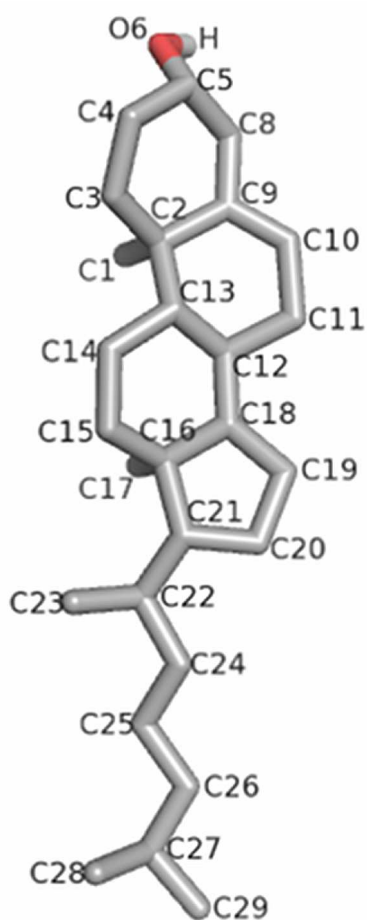


Figure A.1: The structure of cholesterol, with atom naming scheme shown.

```

[ moleculetype ]
; name      nrexcl
CHOL        3

[ atoms ]
;  nr      type  resnr  residu  atom  cgnr  charge  mass      ; total charge
   1      CH3    1      CHOL    C1    1     0       15.035
   2       C     1      CHOL    C2    2     0       12.011
   3     CH2R    1      CHOL    C3    3     0       14.027
   4     CH2R    1      CHOL    C4    4     0       14.027
   5     CH1     1      CHOL    C5    5     0.266   13.019 ; charges adopted
   6     OA     1      CHOL    O6    5    -0.674  15.9994 ; from SER in
   7      H     1      CHOL    H     5     0.408   1.008 ; ffG53a6.rtp
   8     CH2R    1      CHOL    C8    6     0       14.027
   9       C     1      CHOL    C9    7     0       12.011
  10     CR1    1      CHOL    C10   8     0       13.019
  11     CH2R    1      CHOL    C11   9     0       14.027
  12     CH1     1      CHOL    C12  10     0       13.019
  13     CH1     1      CHOL    C13  11     0       13.019
  14     CH2     1      CHOL    C14  12     0       14.027
  15     CH2     1      CHOL    C15  13     0       14.027
  16       C     1      CHOL    C16  14     0       12.011
  17     CH3     1      CHOL    C17  15     0       15.035
  18     CH1     1      CHOL    C18  16     0       13.019
  19     CH2R    1      CHOL    C19  17     0       14.027
  20     CH2R    1      CHOL    C20  18     0       14.027
  21     CH1     1      CHOL    C21  19     0       13.019
  22     CH1     1      CHOL    C22  20     0       13.019
  23     CH3     1      CHOL    C23  21     0       15.035
  24     CH2     1      CHOL    C24  22     0       14.027
  25     CH2     1      CHOL    C25  23     0       14.027
  26     CH2     1      CHOL    C26  24     0       14.027
  27     CH1     1      CHOL    C27  25     0       13.019
  28     CH3     1      CHOL    C28  26     0       15.035
  29     CH3     1      CHOL    C29  27     0       15.035

;
;
;
;
; All values were taken from the closest corresponding parameters in ffgmxnb.itp
; and ffgmxbon.itp
;
[ bonds ]
; ai  aj funct      c0      c1
   1   2     1      gb_27

```


2	3	1	gb_27
3	4	1	gb_27
4	5	1	gb_27
5	6	1	gb_5
6	7	1	gb_1
5	8	1	gb_27
8	9	1	gb_27
9	2	1	gb_27
9	10	1	gb_16
10	11	1	gb_27
11	12	1	gb_27
12	13	1	gb_27
2	13	1	gb_27
13	14	1	gb_27
14	15	1	gb_27
15	16	1	gb_27
16	17	1	gb_27
18	16	1	gb_27
18	12	1	gb_27
18	19	1	gb_27
19	20	1	gb_27
20	21	1	gb_27
16	21	1	gb_27
21	22	1	gb_27
22	23	1	gb_27
22	24	1	gb_27
24	25	1	gb_27
25	26	1	gb_27
26	27	1	gb_27
27	28	1	gb_27
27	29	1	gb_27

[pairs]

;	ai	aj	funct	c0	c1
	1	4	1		
	1	8	1		
	1	10	1		
	1	12	1		
	1	14	1		
	2	11	1		
	2	15	1		
	2	18	1		
	3	6	1		
	3	10	1		
	3	12	1		

3	14	1			
4	7	1			
4	13	1			
5	10	1			
6	9	1			
7	8	1			
8	11	1			
8	13	1			
9	12	1			
9	14	1			
10	18	1			
11	14	1			
11	16	1			
11	19	1			
12	15	1			
12	17	1			
12	20	1			
12	21	1			
13	16	1			
13	19	1			
14	17	1			
14	21	1			
15	19	1			
15	20	1			
15	22	1			
16	23	1			
16	24	1			
17	19	1			
17	20	1			
17	22	1			
18	22	1			
19	22	1			
20	23	1			
20	24	1			
21	25	1			
22	26	1			
23	25	1			
24	27	1			
25	28	1			
25	29	1			
[angles]					
; ai	aj	ak	funct	c0	c1
9	2	13	1	111.	460.240
9	2	1	1	109.5	460.240 ; according to crystal data

9	2	3	1	111.	460.240
13	2	1	1	111.	460.240
13	2	3	1	111.	460.240
1	2	3	1	111.	460.240
2	3	4	1	ga_8	
3	4	5	1	ga_8	
4	5	8	1	ga_8	
4	5	6	1	ga_13	
8	5	6	1	ga_13	
5	6	7	1	ga_12	
9	8	5	1	ga_8	
8	9	10	1	ga_27	
8	9	2	1	120.	460.240
10	9	2	1	ga_27	
11	10	9	1	ga_27	
12	11	10	1	ga_8	
13	12	18	1	ga_8	
13	12	11	1	ga_8	
18	12	11	1	ga_8	
14	13	12	1	ga_8	
14	13	2	1	111.	460.240
12	13	2	1	111.	460.240
15	14	13	1	ga_8	
16	15	14	1	ga_8	
21	16	17	1	111.	460.240
21	16	18	1	111.	460.240
21	16	15	1	111.	460.240
17	16	18	1	111.	460.240
17	16	15	1	111.	460.240
18	16	15	1	111.	460.240
19	18	16	1	111.	460.240
19	18	12	1	ga_8	
16	18	12	1	111.	460.240
18	19	20	1	ga_7	
19	20	21	1	ga_7	
20	21	22	1	ga_7	
20	21	16	1	111.	460.240
22	21	16	1	111.	460.240
23	22	24	1	ga_15	
23	22	21	1	ga_15	
24	22	21	1	ga_15	
25	24	22	1	ga_15	
26	25	24	1	ga_15	
27	26	25	1	ga_15	
28	27	29	1	ga_15	

```

28  27  26  1    ga_15
29  27  26  1    ga_15

```

```
[ dihedrals ]
```

```

; ai  aj  ak  al funct
  4   5   6   7   1  gd_23
  4   5   8   9   1  gd_10
  8   5   4   3   1  gd_10
  5   4   3   2   1  gd_10
  4   3   2   9   1  gd_10
  9   2  13  14   1   0.000   0.418   6
 13   2   9   8   1 180.000  41.840   2
  2   9   8   5   1  gd_10
  9  10  11  12   1 180.000  41.840   2
 10  11  12  13   1  gd_10
 13  12  18  19   1  gd_10
 18  12  13  14   1  gd_10
 12  13  14  15   1  gd_10
 13  14  15  16   1  gd_10
 14  15  16  21   1  gd_10
 15  16  18  19   1   0.000   0.418   6
 15  16  21  22   1   0.000   0.418   6
 20  21  22  23   1  gd_10
 19  20  21  22   1  gd_10
 21  20  19  18   1  gd_10
 20  19  18  12   1  gd_10
 21  22  24  25   1  gd_34
 22  24  25  26   1  gd_34
 24  25  26  27   1  gd_34
 25  26  27  28   1  gd_34

```

```
[ dihedrals ]
```

```

; ai  aj  ak  al funct
  2  13   3   1   2  gi_2
  5   8   4   6   2  gi_2
 12  13  18  11   2  gi_2
 13  14   2  12   2  gi_2
 16  15  21  18   2  35.264  334.720
 18  12  19  16   2  35.264  334.720
 21  22  16  20   2  gi_2
 22  24  21  23   2  gi_2
  2   9  10  11   2  gi_1
  9   2   8  10   2  gi_1
  9  10   2   8   2  gi_1

```

```
#ifdef POSRES_LIPID
#include "chol_posre.itp"
#endif
```

Appendix B

Palmitoylsphingomyelin Topology

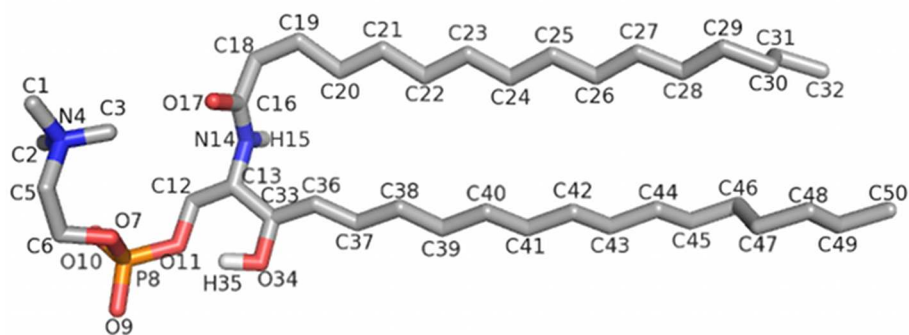


Figure B.1: The structure of palmitoylsphingomyelin, with atom naming scheme shown.

```
[ moleculetype ]
```

```
; name      nrexcl  
SM          3
```

```
[ atoms ]
```

```
;  nr  type  resnr  residu  atom  cgnr  charge  mass  ; total charge  
  1   LC3    1     SM    C1    0    0.400  15.035 ; 0.400  
  2   LC3    1     SM    C2    0    0.400  15.035 ; 0.800  
  3   LC3    1     SM    C3    0    0.400  15.035 ; 1.200  
  4   LNL    1     SM    N4    0   -0.500  14.0067 ; 0.700  
  5   LH2    1     SM    C5    0    0.300  14.027  ; 1.000  
  6   LH2    1     SM    C6    1    0.400  14.027  ; 1.400
```


7	LOS	1	SM	O7	1	-0.800	15.9994	; 0.600
8	LP	1	SM	P8	1	1.700	30.9738	; 2.300
9	LOM	1	SM	O9	1	-0.800	15.9994	; 1.500
10	LOM	1	SM	O10	1	-0.800	15.9994	; 0.700
11	LOS	1	SM	O11	1	-0.700	15.9994	; 0.000
12	LC2	1	SM	C12	2	0.000	14.027	; 0.000
13	LH1	1	SM	C13	3	0.000	13.019	; 0.000
14	N	1	SM	N14	4	-0.310	14.0067	; -0.310
15	H	1	SM	H15	4	0.310	1.008	; 0.000
16	LC	1	SM	C16	4	0.380	12.011	; 0.380
17	LO	1	SM	O17	4	-0.380	15.9994	; 0.000
18	LP2	1	SM	C18	5	0.000	14.027	; 0.000
19	LP2	1	SM	C19	6	0.000	14.027	; 0.000
20	LP2	1	SM	C20	7	0.000	14.027	; 0.000
21	LP2	1	SM	C21	8	0.000	14.027	; 0.000
22	LP2	1	SM	C22	9	0.000	14.027	; 0.000
23	LP2	1	SM	C23	10	0.000	14.027	; 0.000
24	LP2	1	SM	C24	11	0.000	14.027	; 0.000
25	LP2	1	SM	C25	12	0.000	14.027	; 0.000
26	LP2	1	SM	C26	13	0.000	14.027	; 0.000
27	LP2	1	SM	C27	14	0.000	14.027	; 0.000
28	LP2	1	SM	C28	15	0.000	14.027	; 0.000
29	LP2	1	SM	C29	16	0.000	14.027	; 0.000
30	LP2	1	SM	C30	17	0.000	14.027	; 0.000
31	LP2	1	SM	C31	18	0.000	14.027	; 0.000
32	LP3	1	SM	C32	19	0.000	15.035	; 0.000
33	LH1	1	SM	C33	20	0.266	13.019	; 0.266
34	OA	1	SM	O34	20	-0.674	15.9994	; -0.408
35	H	1	SM	H35	20	0.408	1.008	; 0.000
36	LH1	1	SM	C36	21	0.000	13.019	; 0.000
37	LH1	1	SM	C37	22	0.000	13.019	; 0.000
38	LP2	1	SM	C38	23	0.000	14.027	; 0.000
39	LP2	1	SM	C39	24	0.000	14.027	; 0.000
40	LP2	1	SM	C40	25	0.000	14.027	; 0.000
41	LP2	1	SM	C41	26	0.000	14.027	; 0.000
42	LP2	1	SM	C42	27	0.000	14.027	; 0.000
43	LP2	1	SM	C43	28	0.000	14.027	; 0.000
44	LP2	1	SM	C44	29	0.000	14.027	; 0.000
45	LP2	1	SM	C45	30	0.000	14.027	; 0.000
46	LP2	1	SM	C46	31	0.000	14.027	; 0.000
47	LP2	1	SM	C47	32	0.000	14.027	; 0.000
48	LP2	1	SM	C48	33	0.000	14.027	; 0.000
49	LP2	1	SM	C49	34	0.000	14.027	; 0.000
50	LP3	1	SM	C50	35	0.000	15.035	; 0.000

[bonds]

```

; *****
; All values were taken from the closest corresponding parameters in
; ffgmxbon.itp. The C=C double bond is from Tielemann's popc.itp
; *****

;   ai    aj    funct          c0          c1    ; bond type (in ffgmxbon)
;   1     4     1          0.14700    376560. ; NL-CH3 (used: NL-CH2)
;   2     4     1          0.14700    376560. ; NL-CH3
;   3     4     1          0.14700    376560. ; NL-CH3
;   4     5     1          0.14700    376560. ; NL-CH2
;   5     6     1          0.15300    334720. ; CH2-CH2
;   6     7     1          0.14300    251040. ; CH2-OS
;   7     8     1          0.16100    251040. ; OS-P
;   8     9     1          0.14800    376560. ; P-OM
;   8    10     1          0.14800    376560. ; P-OM
;   8    11     1          0.16100    251040. ; P-OS
;  11    12     1          0.14300    251040. ; OS-CH2
;  12    13     1          0.15300    334720. ; CH2-CH1
;  13    14     1          0.14700    376560. ; CH1-N
;  14    15     1          0.10000    374468. ; N-H
;  14    16     1          0.13300    418400. ; N-C
;  16    17     1          0.12300    502080. ; C=O
;  16    18     1          0.15300    334720. ; C-CH2
;  18    19     1          0.15300    334720. ; CH2-CH2
;  19    20     1          0.15300    334720. ; CH2-CH2
;  20    21     1          0.15300    334720. ; CH2-CH2
;  21    22     1          0.15300    334720. ; CH2-CH2
;  22    23     1          0.15300    334720. ; CH2-CH2
;  23    24     1          0.15300    334720. ; CH2-CH2
;  24    25     1          0.15300    334720. ; CH2-CH2
;  25    26     1          0.15300    334720. ; CH2-CH2
;  26    27     1          0.15300    334720. ; CH2-CH2
;  27    28     1          0.15300    334720. ; CH2-CH2
;  28    29     1          0.15300    334720. ; CH2-CH2
;  29    30     1          0.15300    334720. ; CH2-CH2
;  30    31     1          0.15300    334720. ; CH2-CH2
;  31    32     1          0.15300    334720. ; CH2-CH3
;  13    33     1          0.15300    334720. ; CH1-CH1
;  33    34     1          0.14300    334720. ; CH1-OA
;  34    35     1          0.10000    313800. ; OA-HO
;  33    36     1          0.15300    334720. ; CH1-CH1
;  36    37     1          0.13900    418400. ; CH1=CH1 (popc.itp)

```

37	38	1	0.15300	334720.	; CH1-CH2
38	39	1	0.15300	334720.	; CH2-CH2
39	40	1	0.15300	334720.	; CH2-CH2
40	41	1	0.15300	334720.	; CH2-CH2
41	42	1	0.15300	334720.	; CH2-CH2
42	43	1	0.15300	334720.	; CH2-CH2
43	44	1	0.15300	334720.	; CH2-CH2
44	45	1	0.15300	334720.	; CH2-CH2
45	46	1	0.15300	334720.	; CH2-CH2
46	47	1	0.15300	334720.	; CH2-CH2
47	48	1	0.15300	334720.	; CH2-CH2
48	49	1	0.15300	334720.	; CH2-CH2
49	50	1	0.15300	334720.	; CH2-CH3

[pairs]

```
; *****
; Most of the parameters (empty ones) will be taken from lipid.itp, the
; rest (filled in here) were adapted from ffgmxnb.itp.
; *****
```

; ai	aj	funct	c0	c1	; pair type in ffgmxnb.itp
1	6	1			
2	6	1			
3	6	1			
4	7	1			
5	8	1			
6	9	1			
6	10	1			
6	11	1			
7	12	1			
8	13	1			
9	12	1			
10	12	1			
11	14	1	0.23473E-02	0.11203E-05	; OS-N
11	33	1			
12	15	1	0.00000E+00	0.00000E+00	; CH2-H
12	16	1			
12	34	1	0.32685E-02	0.22969E-05	; CH2-OA
12	36	1			
13	17	1			
13	18	1			
13	35	1	0.00000E+00	0.00000E+00	; CH1-HO
13	37	1			

```

14 19 1 0.33923E-02 0.34700E-05 ; N-CH2
14 34 1
14 36 1 0.26633E-02 0.25147E-05 ; N-CH1
15 17 1 0.00000E+00 0.00000E+00 ; H-O
15 33 1 0.00000E+00 0.00000E+00 ; H-CH2
15 18 1 0.00000E+00 0.00000E+00 ; H-CH2
16 20 1
16 33 1
17 19 1
; 18 21 1 ; These pairs removed because of the
; 19 22 1 ; type 3 dihedrals
; 20 23 1
; 21 24 1
; 22 25 1
; 23 26 1
; 24 27 1
; 25 28 1
; 26 29 1
; 27 30 1
; 28 31 1
; 29 32 1
33 38 1
34 37 1 0.25662E-02 0.16646E-05 ; OA-CH1
35 36 1 0.00000E+00 0.00000E+00 ; HO-CH1
36 39 1
; 37 40 1
; 38 41 1
; 39 42 1
; 40 43 1
; 41 44 1
; 42 45 1
; 43 46 1
; 44 47 1
; 45 48 1
; 46 49 1
; 47 50 1

[ angles ]

; *****
; All values were taken from the closest corresponding parameters in
; ffgmxbon.itp (except for C=C from popc.itp).
; *****

; ai aj ak funct th0 cth ; angletype (ffgm)

```

1	4	2	1	109.500	334.720	; used: H-NL-H
1	4	3	1	109.500	334.720	; used: H-NL-H
1	4	5	1	109.500	376.560	; used: H-NL-CH2
2	4	3	1	109.500	334.720	; used: H-NL-H
2	4	5	1	109.500	376.560	; used: H-NL-CH2
3	4	5	1	109.500	376.560	; used: H-NL-CH2
4	5	6	1	109.500	460.240	; NL-CH2-CH2
5	6	7	1	107.900	460.240	; CH2-CH2-OS
6	7	8	1	120.000	397.480	; CH2-OS-P
7	8	9	1	109.600	397.480	; OS-P-OM
7	8	10	1	109.600	397.480	; OS-P-OM
7	8	11	1	103.000	397.480	; OS-P-OS
8	11	12	1	120.000	397.480	; P-OS-CH2
9	8	10	1	120.000	585.760	; OM-P-OM
9	8	11	1	109.600	397.480	; OM-P-OS
10	8	11	1	109.600	397.480	; OM-P-OS
11	12	13	1	111.000	460.240	; OS-CH2-CH1
12	13	14	1	109.500	460.240	; CH2-CH1-N
12	13	33	1	109.500	460.240	; CH2-CH1-CH2
13	14	15	1	115.000	376.560	; CH1-N-H
13	14	16	1	122.000	502.080	; CH1-N-C
13	33	34	1	109.500	460.240	; CH1-CH1-OA
13	33	36	1	111.000	460.240	; CH1-CH1-CH1
14	16	17	1	124.000	502.080	; N-C=O
14	16	18	1	115.000	502.080	; N-C-CH2
14	13	33	1	109.500	460.240	; N-CH1-CH2
15	14	16	1	123.000	292.880	; H-N-C
16	18	19	1	111.000	460.240	; C-CH2-CH2
17	16	18	1	121.000	502.080	; O=C-CH2
18	19	20	1	111.000	460.240	; CH2-CH2-CH2
19	20	21	1	111.000	460.240	; CH2-CH2-CH2
20	21	22	1	111.000	460.240	; CH2-CH2-CH2
21	22	23	1	111.000	460.240	; CH2-CH2-CH2
22	23	24	1	111.000	460.240	; CH2-CH2-CH2
23	24	25	1	111.000	460.240	; CH2-CH2-CH2
24	25	26	1	111.000	460.240	; CH2-CH2-CH2
25	26	27	1	111.000	460.240	; CH2-CH2-CH2
26	27	28	1	111.000	460.240	; CH2-CH2-CH2
27	28	29	1	111.000	460.240	; CH2-CH2-CH2
28	29	30	1	111.000	460.240	; CH2-CH2-CH2
29	30	31	1	111.000	460.240	; CH2-CH2-CH2
30	31	32	1	111.000	460.240	; CH2-CH2-CH3
33	34	35	1	109.500	397.480	; CH2-OA-HO
33	36	37	1	120.000	502.080	; CH1-CH1=CH1 (popc.itp)
34	33	36	1	109.500	460.240	; OA-CH2-CH1

```

36  37  38  1      120.000  502.080 ; CH1=CH1-CH2 (popc.itp)
37  38  39  1      111.000  460.240 ; CH1-CH2-CH2
38  39  40  1      111.000  460.240 ; CH2-CH2-CH2
39  40  41  1      111.000  460.240 ; CH2-CH2-CH2
40  41  42  1      111.000  460.240 ; CH2-CH2-CH2
41  42  43  1      111.000  460.240 ; CH2-CH2-CH2
42  43  44  1      111.000  460.240 ; CH2-CH2-CH2
43  44  45  1      111.000  460.240 ; CH2-CH2-CH2
44  45  46  1      111.000  460.240 ; CH2-CH2-CH2
45  46  47  1      111.000  460.240 ; CH2-CH2-CH2
46  47  48  1      111.000  460.240 ; CH2-CH2-CH2
47  48  49  1      111.000  460.240 ; CH2-CH2-CH2
48  49  50  1      111.000  460.240 ; CH2-CH2-CH2

[ dihedrals ] ; propers

; *****
; - Most dihedral values taken from ffgmxbon.itp
; - For the chains Ryckaert-Belleman's type "3" dihedral used (parameters
;   from lipid.itp)
; *****

; ai  aj  ak  al funct  phi0      cp  mult ; type in ffgmx
;  1  4  5  6   1  0.000   3.766   3 ; NL-CH2
;  4  5  6  7   1  0.000   5.858   3 ; CH2-CH2
;  5  6  7  8   1  0.000   3.766   3 ; CH2-OS
;  6  7  8  9   1  0.000   1.046   3 ; OS-P
;  6  7  8  9   1  0.000   3.138   2 ; OS-P
;  7  8 11 12   1  0.000   1.046   3 ; P-OS
;  7  8 11 12   1  0.000   3.138   2 ; P-OS
;  8 11 12 13   1  0.000   3.766   3 ; OS-CH2
; 11 12 13 14   1  0.000   5.858   3 ; CH2-CH1
; 12 13 14 15   1 180.000   0.418   6 ; CH1-N
; 13 14 16 18   1 180.000  33.472   2 ; N-C
; 14 16 18 19   1  0.000   0.418   6 ; C-CH2
; 16 18 19 20   3                ; CH2-CH2
; 18 19 20 21   3                ; CH2-CH2
; 19 20 21 22   3                ; CH2-CH2
; 20 21 22 23   3                ; CH2-CH2
; 21 22 23 24   3                ; CH2-CH2
; 22 23 24 25   3                ; CH2-CH2
; 23 24 25 26   3                ; CH2-CH2
; 24 25 26 27   3                ; CH2-CH2
; 25 26 27 28   3                ; CH2-CH2
; 26 27 28 29   3                ; CH2-CH2

```



```

27  28  29  30  3      ; CH2-CH2
28  29  30  31  3      ; CH2-CH2
29  30  31  32  3      ; CH2-CH2
13  33  34  35  1  0.000  1.255  3  ; CH2-OA
13  33  36  37  1  0.000  5.858  3  ; CH2-CH1
36  37  38  39  1  0.000  5.858  3  ; CH1-CH2
37  38  39  40  3      ; CH2-CH2
38  39  40  41  3      ; CH2-CH2
39  40  41  42  3      ; CH2-CH2
40  41  42  43  3      ; CH2-CH2
41  42  43  44  3      ; CH2-CH2
42  43  44  45  3      ; CH2-CH2
43  44  45  46  3      ; CH2-CH2
44  45  46  47  3      ; CH2-CH2
45  46  47  48  3      ; CH2-CH2
46  47  48  49  3      ; CH2-CH2
47  48  49  50  3      ; CH2-CH2

[ dihedrals ] ; impropers

; *****
; - The values for peptide bond from ffgmx.rtp
; - Carbons 13 and 33 forced in tetraedric conformation and the bond
;   between them forced in D-erythro conformation.
; - The C=C double bond was adapted from
;   http://moose.bio.ucalgary.ca/files/popc.itp
; *****

; ai  aj  ak  al  funct  q0  cq  ;
16  18  14  17  2  0.000  167.360  ; peptide bond
14  16  13  15  2  0.000  167.360  ; peptide bond
13  12  14  33  2  35.264  334.720  ; tetra-C (from ALA)
33  13  34  36  2  35.264  334.720  ; tetra-C (from 1-lauroyl-glyc)
14  13  34  33  2  0.000  167.360  ; D-erythro
33  38  36  37  2  0.000  167.360  ; CH1=CH1 (trans)

#ifdef POSRES_LIPID
#include "sm_posre.itp"
#endif

```

Appendix C

Ganglioside GM1 Topology

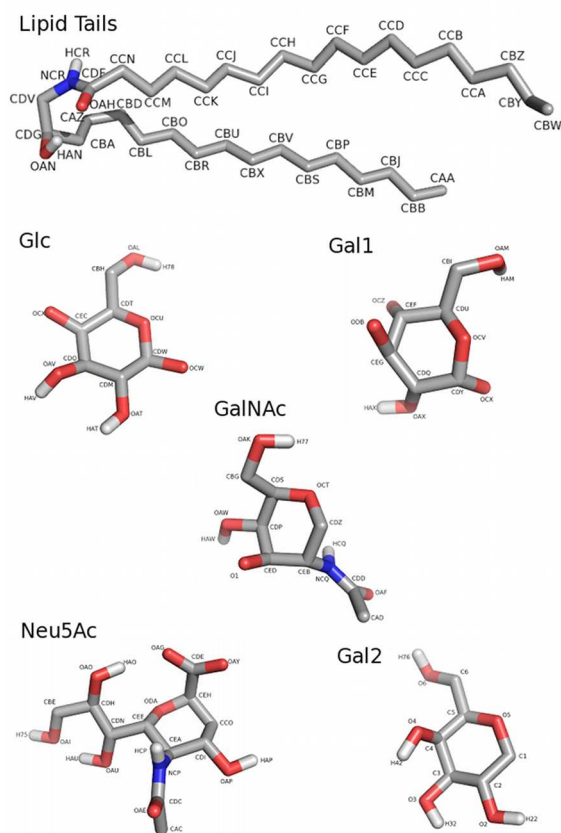


Figure C.1: The structure of ganglioside GM1, separated by functional groups, with atom naming scheme shown.

[moleculetype]

; Name nrexcl

GM1 3

[atoms]

;	nr	type	resnr	resid	atom	cgnr	charge	mass	
	1	LP3	1	GM1	CBW	7	0.000	15.0350	
	2	LP2	1	GM1	CBY	8	0.000	14.0270	
	3	LP2	1	GM1	CBZ	9	0.000	14.0270	
	4	LP2	1	GM1	CCA	10	0.000	14.0270	
	5	LP2	1	GM1	CCB	11	0.000	14.0270	
	6	LP2	1	GM1	CCC	12	0.000	14.0270	
	7	LP2	1	GM1	CCD	13	0.000	14.0270	
	8	LP2	1	GM1	CCE	14	0.000	14.0270	
	9	LP2	1	GM1	CCF	15	0.000	14.0270	
	10	LP2	1	GM1	CCG	16	0.000	14.0270	
	11	LP2	1	GM1	CCH	17	0.000	14.0270	
	12	LP2	1	GM1	CCI	18	0.000	14.0270	
	13	LP2	1	GM1	CCJ	19	0.000	14.0270	
	14	LP2	1	GM1	CCK	20	0.000	14.0270	
	15	LP2	1	GM1	CCL	21	0.000	14.0270	
	16	LP2	1	GM1	CCM	22	0.000	14.0270	
	17	LP2	1	GM1	CCN	23	0.000	14.0270	
	18	LC	1	GM1	CDF	24	0.450	12.0110	; 53a6 amide
	19	LO	1	GM1	DAH	24	-0.450	15.9994	; 53a6 amide
	20	N	1	GM1	NCR	25	-0.310	14.0067	; 53a6 amide
	21	H	1	GM1	HCR	25	0.310	1.0080	; 53a6 amide
	22	CH1	1	GM1	CDV	25	0.000	13.0190	
	23	CH1	1	GM1	CDG	26	0.266	13.0190	; 53a6 CH2OH (Ser)
	24	OA	1	GM1	DAN	26	-0.674	15.9994	; 53a6 CH2OH (Ser)
	25	H	1	GM1	HAN	26	0.408	1.0080	; 53a6 CH2OH (Ser)
	26	LH1	1	GM1	CBA	27	0.000	13.0190	; Tieleman C=C
	27	LH1	1	GM1	CAZ	28	0.000	13.0190	; Tieleman C=C
	28	LP2	1	GM1	CBD	29	0.000	14.0270	
	29	LP2	1	GM1	CBL	30	0.000	14.0270	
	30	LP2	1	GM1	CBO	31	0.000	14.0270	
	31	LP2	1	GM1	CBR	32	0.000	14.0270	
	32	LP2	1	GM1	CBU	33	0.000	14.0270	
	33	LP2	1	GM1	CBX	34	0.000	14.0270	
	34	LP2	1	GM1	CBV	35	0.000	14.0270	
	35	LP2	1	GM1	CBS	36	0.000	14.0270	
	36	LP2	1	GM1	CBP	37	0.000	14.0270	
	37	LP2	1	GM1	CBM	38	0.000	14.0270	
	38	LP2	1	GM1	CBJ	39	0.000	14.0270	
	39	LP2	1	GM1	CBB	40	0.000	14.0270	
	40	LP3	1	GM1	CAA	41	0.000	15.0350	

41	OA	1	GM1	OCW	42	-0.360	15.9994	; O1 - 53a6 Glc
42	CH1	1	GM1	CDW	42	0.232	13.0190	; C1 - 53a6 Glc
43	CH1	1	GM1	CDM	43	0.232	13.0190	; C2 - 53a6 Glc
44	OA	1	GM1	OAT	43	-0.642	15.9994	; O2 - 53a6 Glc
45	H	1	GM1	HAT	43	0.410	1.0080	; HO2- 53a6 Glc
46	CH1	1	GM1	CDO	44	0.232	13.0190	; C3 - 53a6 Glc
47	OA	1	GM1	OAV	44	-0.642	15.9994	; O3 - 53a6 Glc
48	H	1	GM1	HAV	44	0.410	1.0080	; HO3- 53a6 Glc
49	OA	1	GM1	OCU	42	-0.480	15.9994	; O5 - 53a6 Glc
50	CH1	1	GM1	CDT	42	0.376	13.0190	; C5 - 53a6 Glc
51	CH2	1	GM1	CBH	45	0.232	14.0270	; C6 - 53a6 Glc
52	OA	1	GM1	OAL	45	-0.642	15.9994	; O6 - 53a6 Glc
53	H	1	GM1	H78	45	0.410	1.0080	; HO6- 53a6 Glc
54	CH1	1	GM1	CEC	46	0.232	13.0190	; C4 - 53a6 Glc
55	OA	1	GM1	OCX	47	-0.360	15.9994	; O1 - 53a6 Gal
56	CH1	1	GM1	CDY	47	0.232	13.0190	; C1 - 53a6 Gal
57	CH1	1	GM1	CDQ	48	0.232	13.0190	; C2 - 53a6 Gal
58	OA	1	GM1	OAX	48	-0.642	15.9994	; O2 - 53a6 Gal
59	H	1	GM1	HAX	48	0.410	1.0080	; HO2- 53a6 Gal
60	CH1	1	GM1	CEG	49	0.116	13.0190	; C3 - 53a6 Gal (mod)
61	OA	1	GM1	ODB	50	-0.360	15.9994	; O1 - Sialic acid
62	CH1	1	GM1	CEH	50	0.232	12.0110	; C1 - Sialic acid
63	C	1	GM1	CDE	51	0.270	12.0110	; 53a6 C00- (Asp)
64	OM	1	GM1	OAY	51	-0.635	15.9994	; 53a6 C00- (Asp)
65	OM	1	GM1	OAG	51	-0.635	15.9994	; 53a6 C00- (Asp)
66	OA	1	GM1	ODA	50	-0.480	15.9994	; O5 - Sialic acid
67	CH1	1	GM1	CCO	52	0.000	14.0270	; C2 - Sialic acid
68	CH1	1	GM1	CDI	53	0.232	13.0190	; C3 - Sialic acid
69	OA	1	GM1	OAP	53	-0.642	15.9994	; O3 - Sialic acid
70	H	1	GM1	HAP	53	0.410	1.0080	; HO3- Sialic acid
71	CH1	1	GM1	CEA	54	0.232	13.0190	; C4 - Sialic acid
72	N	1	GM1	NCP	55	-0.310	14.0067	; amide (53a6)
73	H	1	GM1	HCP	55	0.310	1.0080	; amide (53a6)
74	C	1	GM1	CDC	56	0.450	12.0110	; acetyl (53a6)
75	O	1	GM1	OAE	56	-0.450	15.9994	; acetyl (53a6)
76	CH3	1	GM1	CAC	56	0.000	15.0350	; acetyl (53a6)
77	CH1	1	GM1	CEE	50	0.376	13.0190	; C5 - Sialic acid
78	CH1	1	GM1	CDN	57	0.266	13.0190	; 53a6 CH2OH (Ser)
79	OA	1	GM1	OAU	57	-0.674	15.9994	; 53a6 CH2OH (Ser)
80	H	1	GM1	HAU	57	0.408	1.0080	; 53a6 CH2OH (Ser)
81	CH1	1	GM1	CDH	58	0.266	13.0190	; 53a6 CH2OH (Ser)
82	OA	1	GM1	OAO	58	-0.674	15.9994	; 53a6 CH2OH (Ser)
83	H	1	GM1	HAO	58	0.408	1.0080	; 53a6 CH2OH (Ser)
84	CH2	1	GM1	CBE	59	0.266	14.0270	; 53a6 CH2OH (Ser)
85	OA	1	GM1	OAI	59	-0.674	15.9994	; 53a6 CH2OH (Ser)

86	H	1	GM1	H75	59	0.408	1.0080	; 53a6 CH2OH (Ser)
87	OA	1	GM1	OCV	50	-0.480	15.9994	; O5 - 53a6 Gal
88	CH1	1	GM1	CDU	50	0.376	13.0190	; C5 - 53a6 Gal
89	CH2	1	GM1	CBI	60	0.232	14.0270	; C6 - 53a6 Gal
90	OA	1	GM1	OAM	60	-0.642	15.9994	; O6 - 53a6 Gal
91	H	1	GM1	HAM	60	0.410	1.0080	; H06- 53a6 Gal
92	CH1	1	GM1	CEF	49	0.116	13.0190	; C4 - 53a6 Gal (mod)
93	OA	1	GM1	OCZ	61	-0.360	15.9994	; O1 - GalNAc
94	CH1	1	GM1	CDZ	61	0.232	13.0190	; C1 - GalNAc
95	CH1	1	GM1	CEB	62	0.232	13.0190	; C2 - GalNAc
96	N	1	GM1	NCQ	63	-0.310	14.0067	; amide (53a6)
97	H	1	GM1	HCQ	63	0.310	1.0080	; amide (53a6)
98	C	1	GM1	CDD	64	0.450	12.0110	; acetyl (53a6)
99	O	1	GM1	OAF	64	-0.450	15.9994	; acetyl (53a6)
100	CH3	1	GM1	CAD	64	0.000	15.0350	; acetyl (53a6)
101	OA	1	GM1	OCT	61	-0.480	15.9994	; O5 - GalNAc
102	CH1	1	GM1	CDS	61	0.376	13.0190	; C5 - GalNAc
103	CH2	1	GM1	CBG	65	0.232	14.0270	; C6 - GalNAc
104	OA	1	GM1	OAK	65	-0.642	15.9994	; O6 - GalNAc
105	H	1	GM1	H77	65	0.410	1.0080	; H06- GalNAc
106	CH1	1	GM1	CDP	66	0.232	13.0190	; C4 - GalNAc
107	OA	1	GM1	OAW	66	-0.642	15.9994	; O4 - GalNAc
108	H	1	GM1	HAW	66	0.410	1.0080	; H04- GalNAc
109	CH1	1	GM1	CED	62	0.232	13.0190	; C3 - GalNAc
110	OA	1	GM1	O1	67	-0.360	15.9994	; 53a6 Gal
111	CH1	1	GM1	C1	67	0.232	13.0190	; 53a6 Gal
112	OA	1	GM1	O5	67	-0.480	15.9994	; 53a6 Gal
113	CH1	1	GM1	C5	67	0.376	13.0190	; 53a6 Gal
114	CH2	1	GM1	C6	68	0.232	14.0270	; 53a6 Gal
115	OA	1	GM1	O6	68	-0.642	15.9994	; 53a6 Gal
116	H	1	GM1	H76	68	0.410	1.0080	; H06- 53a6 Gal
117	CH1	1	GM1	C4	69	0.232	13.0190	; 53a6 Gal
118	OA	1	GM1	O4	69	-0.642	15.9994	; 53a6 Gal
119	H	1	GM1	H42	69	0.410	1.0080	; H04- 53a6 Gal
120	CH1	1	GM1	C3	70	0.232	13.0190	; 53a6 Gal
121	OA	1	GM1	O3	70	-0.642	15.9994	; 53a6 Gal
122	H	1	GM1	H32	70	0.410	1.0080	; H03- 53a6 Gal
123	CH1	1	GM1	C2	71	0.232	13.0190	; 53a6 Gal
124	OA	1	GM1	O2	71	-0.642	15.9994	; 53a6 Gal
125	H	1	GM1	H22	71	0.410	1.0080	; H02- 53a6 Gal

[bonds]

```

; ai aj fu c0, c1, ...
  1  2  2  0.153  7150000.0  0.153  7150000.0 ; CBW  CBY
  2  3  2  0.153  7150000.0  0.153  7150000.0 ; CBY  CBZ

```

3	4	2	0.153	7150000.0	0.153	7150000.0 ;	CBZ	CCA
4	5	2	0.153	7150000.0	0.153	7150000.0 ;	CCA	CCB
5	6	2	0.153	7150000.0	0.153	7150000.0 ;	CCB	CCC
6	7	2	0.153	7150000.0	0.153	7150000.0 ;	CCC	CCD
7	8	2	0.153	7150000.0	0.153	7150000.0 ;	CCD	CCE
8	9	2	0.153	7150000.0	0.153	7150000.0 ;	CCE	CCF
9	10	2	0.153	7150000.0	0.153	7150000.0 ;	CCF	CCG
10	11	2	0.153	7150000.0	0.153	7150000.0 ;	CCG	CCH
11	12	2	0.153	7150000.0	0.153	7150000.0 ;	CCH	CCI
12	13	2	0.153	7150000.0	0.153	7150000.0 ;	CCI	CCJ
13	14	2	0.153	7150000.0	0.153	7150000.0 ;	CCJ	CCK
14	15	2	0.153	7150000.0	0.153	7150000.0 ;	CCK	CCL
15	16	2	0.153	7150000.0	0.153	7150000.0 ;	CCL	CCM
16	17	2	0.153	7150000.0	0.153	7150000.0 ;	CCM	CCN
17	18	2	0.153	7150000.0	0.153	7150000.0 ;	CCN	CDF
18	19	2	0.123	16600000.0	0.123	16600000.0 ;	CDF	OAH
18	20	2	0.134	10500000.0	0.134	10500000.0 ;	CDF	NCR
20	21	2	0.100	18700000.0	0.100	18700000.0 ;	NCR	HCR
20	22	2	0.147	8710000.0	0.147	8710000.0 ;	NCR	CDV
22	23	2	0.153	7150000.0	0.153	7150000.0 ;	CDV	CDG
22	41	2	0.144	6100000.0	0.144	6100000.0 ;	CDV	OCW
23	24	2	0.143	8180000.0	0.143	8180000.0 ;	CDG	OAN
23	26	2	0.153	7150000.0	0.153	7150000.0 ;	CDG	CBA
24	25	2	0.100	15700000.0	0.100	15700000.0 ;	OAN	HAN
26	27	2	0.153	7150000.0	0.153	7150000.0 ;	CBA	CAZ
27	28	2	0.153	7150000.0	0.153	7150000.0 ;	CAZ	CBD
28	29	2	0.153	7150000.0	0.153	7150000.0 ;	CBD	CBL
29	30	2	0.153	7150000.0	0.153	7150000.0 ;	CBL	CBO
30	31	2	0.153	7150000.0	0.153	7150000.0 ;	CBO	CBR
31	32	2	0.153	7150000.0	0.153	7150000.0 ;	CBR	CBU
32	33	2	0.153	7150000.0	0.153	7150000.0 ;	CBU	CBX
33	34	2	0.153	7150000.0	0.153	7150000.0 ;	CBX	CBV
34	35	2	0.153	7150000.0	0.153	7150000.0 ;	CBV	CBS
35	36	2	0.153	7150000.0	0.153	7150000.0 ;	CBS	CBP
36	37	2	0.153	7150000.0	0.153	7150000.0 ;	CBP	CBM
37	38	2	0.153	7150000.0	0.153	7150000.0 ;	CBM	CBJ
38	39	2	0.153	7150000.0	0.153	7150000.0 ;	CBJ	CBB
39	40	2	0.153	7150000.0	0.153	7150000.0 ;	CBB	CAA
41	42	2	0.144	6100000.0	0.144	6100000.0 ;	OCW	CDW
42	43	2	0.152	5430000.0	0.152	5430000.0 ;	CDW	CDM
42	49	2	0.144	6100000.0	0.144	6100000.0 ;	CDW	OCU
43	44	2	0.144	6100000.0	0.144	6100000.0 ;	CDM	OAT
43	46	2	0.152	5430000.0	0.152	5430000.0 ;	CDM	CDO
44	45	2	0.100	15700000.0	0.100	15700000.0 ;	OAT	HAT
46	47	2	0.144	6100000.0	0.144	6100000.0 ;	CDO	OAV

46	54	2	0.152	5430000.0	0.152	5430000.0 ;	CDO	CEC
47	48	2	0.100	15700000.0	0.100	15700000.0 ;	OAV	HAV
49	50	2	0.144	6100000.0	0.144	6100000.0 ;	OCU	CDT
50	51	2	0.152	5430000.0	0.152	5430000.0 ;	CDT	CBH
50	54	2	0.152	5430000.0	0.152	5430000.0 ;	CDT	CEC
51	52	2	0.143	8180000.0	0.143	8180000.0 ;	CBH	OAL
52	53	2	0.100	15700000.0	0.100	15700000.0 ;	OAL	H78
54	55	2	0.144	6100000.0	0.144	6100000.0 ;	CEC	OCX
55	56	2	0.144	6100000.0	0.144	6100000.0 ;	OCX	CDY
56	57	2	0.152	5430000.0	0.152	5430000.0 ;	CDY	CDQ
56	87	2	0.144	6100000.0	0.144	6100000.0 ;	CDY	OCV
57	58	2	0.144	6100000.0	0.144	6100000.0 ;	CDQ	OAX
57	60	2	0.152	5430000.0	0.152	5430000.0 ;	CDQ	CEG
58	59	2	0.100	15700000.0	0.100	15700000.0 ;	OAX	HAX
60	61	2	0.144	6100000.0	0.144	6100000.0 ;	CEG	ODB
60	92	2	0.152	5430000.0	0.152	5430000.0 ;	CEG	CEF
61	62	2	0.144	6100000.0	0.144	6100000.0 ;	ODB	CEH
62	63	2	0.153	7150000.0	0.153	7150000.0 ;	CEH	CDE
62	66	2	0.144	6100000.0	0.144	6100000.0 ;	CEH	ODA
62	67	2	0.152	5430000.0	0.152	5430000.0 ;	CEH	CCO
63	64	2	0.125	13400000.0	0.125	13400000.0 ;	CDE	OAY
63	65	2	0.125	13400000.0	0.125	13400000.0 ;	CDE	OAG
66	77	2	0.144	6100000.0	0.144	6100000.0 ;	ODA	CEE
67	68	2	0.152	5430000.0	0.152	5430000.0 ;	CCO	CDI
68	69	2	0.144	6100000.0	0.144	6100000.0 ;	CDI	OAP
68	71	2	0.152	5430000.0	0.152	5430000.0 ;	CDI	CEA
69	70	2	0.100	15700000.0	0.100	15700000.0 ;	OAP	HAP
71	72	2	0.147	8710000.0	0.147	8710000.0 ;	CEA	NCP
71	77	2	0.152	5430000.0	0.152	5430000.0 ;	CEA	CEE
72	73	2	0.100	18700000.0	0.100	18700000.0 ;	NCP	HCP
72	74	2	0.134	10500000.0	0.134	10500000.0 ;	NCP	CDC
74	75	2	0.123	16600000.0	0.123	16600000.0 ;	CDC	OAE
74	76	2	0.153	7150000.0	0.153	7150000.0 ;	CDC	CAC
77	78	2	0.152	5430000.0	0.152	5430000.0 ;	CEE	CDN
78	79	2	0.143	8180000.0	0.143	8180000.0 ;	CDN	OAU
78	81	2	0.153	7150000.0	0.153	7150000.0 ;	CDN	CDH
79	80	2	0.100	15700000.0	0.100	15700000.0 ;	OAU	HAU
81	82	2	0.143	8180000.0	0.143	8180000.0 ;	CDH	OAO
81	84	2	0.153	7150000.0	0.153	7150000.0 ;	CDH	CBE
82	83	2	0.100	15700000.0	0.100	15700000.0 ;	OAO	HAO
84	85	2	0.143	8180000.0	0.143	8180000.0 ;	CBE	OAI
85	86	2	0.100	15700000.0	0.100	15700000.0 ;	OAI	H75
87	88	2	0.144	6100000.0	0.144	6100000.0 ;	OCV	CDU
88	89	2	0.152	5430000.0	0.152	5430000.0 ;	CDU	CBI
88	92	2	0.152	5430000.0	0.152	5430000.0 ;	CDU	CEF

89	90	2	0.143	8180000.0	0.143	8180000.0	; CBI	OAM
90	91	2	0.100	15700000.0	0.100	15700000.0	; OAM	HAM
92	93	2	0.144	6100000.0	0.144	6100000.0	; CEF	OCZ
93	94	2	0.144	6100000.0	0.144	6100000.0	; OCZ	CDZ
94	95	2	0.152	5430000.0	0.152	5430000.0	; CDZ	CEB
94	101	2	0.144	6100000.0	0.144	6100000.0	; CDZ	OCT
95	96	2	0.147	8710000.0	0.147	8710000.0	; CEB	NCQ
95	109	2	0.152	5430000.0	0.152	5430000.0	; CEB	CED
96	97	2	0.100	18700000.0	0.100	18700000.0	; NCQ	HCQ
96	98	2	0.134	10500000.0	0.134	10500000.0	; NCQ	CDD
98	99	2	0.123	16600000.0	0.123	16600000.0	; CDD	OAF
98	100	2	0.153	7150000.0	0.153	7150000.0	; CDD	CAD
101	102	2	0.144	6100000.0	0.144	6100000.0	; OCT	CDS
102	103	2	0.152	5430000.0	0.152	5430000.0	; CDS	CBG
102	106	2	0.152	5430000.0	0.152	5430000.0	; CDS	CDP
103	104	2	0.143	8180000.0	0.143	8180000.0	; CBG	OAK
104	105	2	0.100	15700000.0	0.100	15700000.0	; OAK	H77
106	107	2	0.144	6100000.0	0.144	6100000.0	; CDP	OAW
106	109	2	0.152	5430000.0	0.152	5430000.0	; CDP	CED
107	108	2	0.100	15700000.0	0.100	15700000.0	; OAW	HAW
109	110	2	0.144	6100000.0	0.144	6100000.0	; CED	O1
110	111	2	0.144	6100000.0	0.144	6100000.0	; O1	C1
111	112	2	0.144	6100000.0	0.144	6100000.0	; C1	O5
111	123	2	0.152	5430000.0	0.152	5430000.0	; C1	C2
112	113	2	0.144	6100000.0	0.144	6100000.0	; O5	C5
113	114	2	0.152	5430000.0	0.152	5430000.0	; C5	C6
113	117	2	0.152	5430000.0	0.152	5430000.0	; C5	C4
114	115	2	0.143	8180000.0	0.143	8180000.0	; C6	O6
115	116	2	0.100	15700000.0	0.100	15700000.0	; O6	H76
117	118	2	0.144	6100000.0	0.144	6100000.0	; C4	O4
117	120	2	0.152	5430000.0	0.152	5430000.0	; C4	C3
118	119	2	0.100	15700000.0	0.100	15700000.0	; O4	H42
120	121	2	0.144	6100000.0	0.144	6100000.0	; C3	O3
120	123	2	0.152	5430000.0	0.152	5430000.0	; C3	C2
121	122	2	0.100	15700000.0	0.100	15700000.0	; O3	H32
123	124	2	0.144	6100000.0	0.144	6100000.0	; C2	O2
124	125	2	0.100	15700000.0	0.100	15700000.0	; O2	H22

[pairs]

; ai aj fu c0, c1, ...

1	4	1					; CBW	CCA
2	5	1					; CBY	CCB
3	6	1					; CBZ	CCC
4	7	1					; CCA	CCD
5	8	1					; CCB	CCE

6	9	1			; CCC	CCF
7	10	1			; CCD	CCG
8	11	1			; CCE	CCH
9	12	1			; CCF	CCI
10	13	1			; CCG	CCJ
11	14	1			; CCH	CCK
12	15	1			; CCI	CCL
13	16	1			; CCJ	CCM
14	17	1			; CCK	CCN
15	18	1			; CCL	CDF
16	19	1			; CCM	OAH
20	24	1			; NCR	OAN
20	42	1			; NCR	CDW
21	23	1			; HCR	CDG
21	41	1			; HCR	OCW
22	25	1			; CDV	HAN
22	43	1			; CDV	CDM
22	49	1			; CDV	OCU
23	42	1			; CDG	CDW
24	41	1			; OAN	OCW
26	29	1			; CBA	CBL
27	30	1			; CAZ	CBO
28	31	1			; CBD	CBR
29	32	1			; CBL	CBU
30	33	1			; CBO	CBX
31	34	1			; CBR	CBV
32	35	1			; CBU	CBS
33	36	1			; CBX	CBP
34	37	1			; CBV	CBM
35	38	1			; CBS	CBJ
36	39	1			; CBP	CBB
37	40	1			; CBM	CAA
41	44	1			; OCW	OAT
41	46	1			; OCW	CDO
41	50	1			; OCW	CDT
42	45	1			; CDW	HAT
42	47	1			; CDW	OAV
42	51	1			; CDW	CBH
42	54	1			; CDW	CEC
43	48	1			; CDM	HAV
43	50	1			; CDM	CDT
43	55	1			; CDM	OCX
44	47	1			; OAT	OAV
44	49	1			; OAT	OCU
44	54	1			; OAT	CEC

45	46	1	;	HAT	CDO
46	49	1	;	CDO	OCU
46	51	1	;	CDO	CBH
46	56	1	;	CDO	CDY
47	50	1	;	OAV	CDT
47	55	1	;	OAV	OCX
48	54	1	;	HAV	CEC
49	52	1	;	OCU	OAL
49	55	1	;	OCU	OCX
50	53	1	;	CDT	H78
50	56	1	;	CDT	CDY
51	55	1	;	CBH	OCX
52	54	1	;	OAL	CEC
54	57	1	;	CEC	CDQ
54	87	1	;	CEC	OCV
55	58	1	;	OCX	OAX
55	60	1	;	OCX	CEG
55	88	1	;	OCX	CDU
56	59	1	;	CDY	HAX
56	61	1	;	CDY	ODB
56	89	1	;	CDY	CBI
56	92	1	;	CDY	CEF
57	62	1	;	CDQ	CEH
57	88	1	;	CDQ	CDU
57	93	1	;	CDQ	OCZ
58	61	1	;	OAX	ODB
58	87	1	;	OAX	OCV
58	92	1	;	OAX	CEF
59	60	1	;	HAX	CEG
60	63	1	;	CEG	CDE
60	66	1	;	CEG	ODA
60	67	1	;	CEG	CCO
60	87	1	;	CEG	OCV
60	89	1	;	CEG	CBI
60	94	1	;	CEG	CDZ
61	64	1	;	ODB	OAY
61	65	1	;	ODB	OAG
61	68	1	;	ODB	CDI
61	77	1	;	ODB	CEE
61	88	1	;	ODB	CDU
61	93	1	;	ODB	OCZ
62	69	1	;	CEH	OAP
62	71	1	;	CEH	CEA
62	78	1	;	CEH	CDN
62	92	1	;	CEH	CEF

63	68	1	;	CDE	CDI
63	77	1	;	CDE	CEE
64	66	1	;	OAY	ODA
64	67	1	;	OAY	CCO
65	66	1	;	OAG	ODA
65	67	1	;	OAG	CCO
66	68	1	;	ODA	CDI
66	72	1	;	ODA	NCP
66	79	1	;	ODA	OAU
66	81	1	;	ODA	CDH
67	70	1	;	CCO	HAP
67	72	1	;	CCO	NCP
67	77	1	;	CCO	CEE
68	73	1	;	CDI	HCP
68	74	1	;	CDI	CDC
68	78	1	;	CDI	CDN
69	72	1	;	OAP	NCP
69	77	1	;	OAP	CEE
70	71	1	;	HAP	CEA
71	75	1	;	CEA	OAE
71	76	1	;	CEA	CAC
71	79	1	;	CEA	OAU
71	81	1	;	CEA	CDH
72	78	1	;	NCP	CDN
73	75	1	;	HCP	OAE
73	76	1	;	HCP	CAC
73	77	1	;	HCP	CEE
74	77	1	;	CDC	CEE
77	80	1	;	CEE	HAU
77	82	1	;	CEE	OAQ
77	84	1	;	CEE	CBE
78	83	1	;	CDN	HAQ
78	85	1	;	CDN	OAI
79	82	1	;	OAU	OAQ
79	84	1	;	OAU	CBE
80	81	1	;	HAU	CDH
81	86	1	;	CDH	H75
82	85	1	;	OAQ	OAI
83	84	1	;	HAQ	CBE
87	90	1	;	OCV	OAM
87	93	1	;	OCV	OCZ
88	91	1	;	CDU	HAM
88	94	1	;	CDU	CDZ
89	93	1	;	CBI	OCZ
90	92	1	;	OAM	CEF

92	95	1	;	CEF	CEB
92	101	1	;	CEF	OCT
93	96	1	;	OCZ	NCQ
93	102	1	;	OCZ	CDS
93	109	1	;	OCZ	CED
94	97	1	;	CDZ	HCQ
94	98	1	;	CDZ	CDD
94	103	1	;	CDZ	CBG
94	106	1	;	CDZ	CDP
94	110	1	;	CDZ	01
95	99	1	;	CEB	OAF
95	100	1	;	CEB	CAD
95	102	1	;	CEB	CDS
95	107	1	;	CEB	OAW
95	111	1	;	CEB	C1
96	101	1	;	NCQ	OCT
96	106	1	;	NCQ	CDP
96	110	1	;	NCQ	01
97	99	1	;	HCQ	OAF
97	100	1	;	HCQ	CAD
97	109	1	;	HCQ	CED
98	109	1	;	CDD	CED
101	104	1	;	OCT	OAK
101	107	1	;	OCT	OAW
101	109	1	;	OCT	CED
102	105	1	;	CDS	H77
102	108	1	;	CDS	HAW
102	110	1	;	CDS	01
103	107	1	;	CBG	OAW
103	109	1	;	CBG	CED
104	106	1	;	OAK	CDP
106	111	1	;	CDP	C1
107	110	1	;	OAW	01
108	109	1	;	HAW	CED
109	112	1	;	CED	05
109	123	1	;	CED	C2
110	113	1	;	01	C5
110	120	1	;	01	C3
110	124	1	;	01	02
111	114	1	;	C1	C6
111	117	1	;	C1	C4
111	121	1	;	C1	03
111	125	1	;	C1	H22
112	115	1	;	05	06
112	118	1	;	05	04

22	23	24	2	109.5	520.0	109.5	520.0 ;	CDV	CDG	OAN
22	23	26	2	109.5	520.0	109.5	520.0 ;	CDV	CDG	CBA
24	23	26	2	109.5	520.0	109.5	520.0 ;	OAN	CDG	CBA
23	24	25	2	109.5	450.0	109.5	450.0 ;	CDG	OAN	HAN
23	26	27	2	115.0	610.0	115.0	610.0 ;	CDG	CBA	CAZ
26	27	28	2	115.0	610.0	115.0	610.0 ;	CBA	CAZ	CBD
27	28	29	2	109.5	520.0	109.5	520.0 ;	CAZ	CBD	CBL
28	29	30	2	109.5	520.0	109.5	520.0 ;	CBD	CBL	CBO
29	30	31	2	109.5	520.0	109.5	520.0 ;	CBL	CBO	CBR
30	31	32	2	109.5	520.0	109.5	520.0 ;	CBO	CBR	CBU
31	32	33	2	109.5	520.0	109.5	520.0 ;	CBR	CBU	CBX
32	33	34	2	109.5	520.0	109.5	520.0 ;	CBU	CBX	CBV
33	34	35	2	109.5	520.0	109.5	520.0 ;	CBX	CBV	CBS
34	35	36	2	109.5	520.0	109.5	520.0 ;	CBV	CBS	CBP
35	36	37	2	109.5	520.0	109.5	520.0 ;	CBS	CBP	CBM
36	37	38	2	109.5	520.0	109.5	520.0 ;	CBP	CBM	CBJ
37	38	39	2	109.5	520.0	109.5	520.0 ;	CBM	CBJ	CBB
38	39	40	2	109.5	520.0	109.5	520.0 ;	CBJ	CBB	CAA
22	41	42	2	109.5	380.0	109.5	380.0 ;	CDV	OCW	CDW
41	42	43	2	109.5	520.0	109.5	520.0 ;	OCW	CDW	CDM
41	42	49	2	109.5	320.0	109.5	320.0 ;	OCW	CDW	OCU
43	42	49	2	109.5	520.0	109.5	520.0 ;	CDM	CDW	OCU
42	43	44	2	109.5	520.0	109.5	520.0 ;	CDW	CDM	OAT
42	43	46	2	109.5	520.0	109.5	520.0 ;	CDW	CDM	CDO
44	43	46	2	109.5	520.0	109.5	520.0 ;	OAT	CDM	CDO
43	44	45	2	109.5	450.0	109.5	450.0 ;	CDM	OAT	HAT
43	46	47	2	109.5	520.0	109.5	520.0 ;	CDM	CDO	OAV
43	46	54	2	109.5	520.0	109.5	520.0 ;	CDM	CDO	CEC
47	46	54	2	109.5	520.0	109.5	520.0 ;	OAV	CDO	CEC
46	47	48	2	109.5	450.0	109.5	450.0 ;	CDO	OAV	HAV
42	49	50	2	109.5	380.0	109.5	380.0 ;	CDW	OCU	CDT
49	50	51	2	109.5	320.0	109.5	320.0 ;	OCU	CDT	CBH
49	50	54	2	109.5	520.0	109.5	520.0 ;	OCU	CDT	CEC
51	50	54	2	109.5	520.0	109.5	520.0 ;	CBH	CDT	CEC
50	51	52	2	109.5	520.0	109.5	520.0 ;	CDT	CBH	OAL
51	52	53	2	109.5	450.0	109.5	450.0 ;	CBH	OAL	H78
46	54	50	2	109.5	520.0	109.5	520.0 ;	CDO	CEC	CDT
46	54	55	2	109.5	520.0	109.5	520.0 ;	CDO	CEC	OCX
50	54	55	2	109.5	520.0	109.5	520.0 ;	CDT	CEC	OCX
54	55	56	2	109.5	380.0	109.5	380.0 ;	CEC	OCX	CDY
55	56	57	2	109.5	520.0	109.5	520.0 ;	OCX	CDY	CDQ
55	56	87	2	109.5	320.0	109.5	320.0 ;	OCX	CDY	OCV
57	56	87	2	109.5	520.0	109.5	520.0 ;	CDQ	CDY	OCV
56	57	58	2	109.5	520.0	109.5	520.0 ;	CDY	CDQ	OAX
56	57	60	2	109.5	520.0	109.5	520.0 ;	CDY	CDQ	CEG

58	57	60	2	109.5	520.0	109.5	520.0 ;	OAX	CDQ	CEG
57	58	59	2	109.5	450.0	109.5	450.0 ;	CDQ	OAX	HAX
57	60	61	2	109.5	520.0	109.5	520.0 ;	CDQ	CEG	ODB
57	60	92	2	109.5	520.0	109.5	520.0 ;	CDQ	CEG	CEF
61	60	92	2	109.5	520.0	109.5	520.0 ;	ODB	CEG	CEF
60	61	62	2	109.5	380.0	109.5	380.0 ;	CEG	ODB	CEH
61	62	63	2	109.5	520.0	109.5	520.0 ;	ODB	CEH	CDE
61	62	66	2	109.5	320.0	109.5	320.0 ;	ODB	CEH	ODA
61	62	67	2	109.5	520.0	109.5	520.0 ;	ODB	CEH	CCO
63	62	66	2	109.5	520.0	109.5	520.0 ;	CDE	CEH	ODA
63	62	67	2	109.5	520.0	109.5	520.0 ;	CDE	CEH	CCO
66	62	67	2	109.5	520.0	109.5	520.0 ;	ODA	CEH	CCO
62	63	64	2	117.0	635.0	117.0	635.0 ;	CEH	CDE	OAY
62	63	65	2	117.0	635.0	117.0	635.0 ;	CEH	CDE	OAG
64	63	65	2	126.0	770.0	126.0	770.0 ;	OAY	CDE	OAG
62	66	77	2	109.5	380.0	109.5	380.0 ;	CEH	ODA	CEE
62	67	68	2	109.5	520.0	109.5	520.0 ;	CEH	CCO	CDI
67	68	69	2	109.5	520.0	109.5	520.0 ;	CCO	CDI	OAP
67	68	71	2	109.5	520.0	109.5	520.0 ;	CCO	CDI	CEA
69	68	71	2	109.5	520.0	109.5	520.0 ;	OAP	CDI	CEA
68	69	70	2	109.5	450.0	109.5	450.0 ;	CDI	OAP	HAP
68	71	72	2	109.5	520.0	109.5	520.0 ;	CDI	CEA	NCP
68	71	77	2	109.5	520.0	109.5	520.0 ;	CDI	CEA	CEE
72	71	77	2	109.5	520.0	109.5	520.0 ;	NCP	CEA	CEE
71	72	73	2	120.0	505.0	120.0	505.0 ;	CEA	NCP	HCP
71	72	74	2	122.0	700.0	122.0	700.0 ;	CEA	NCP	CDC
73	72	74	2	123.0	415.0	123.0	415.0 ;	HCP	NCP	CDC
72	74	75	2	124.0	730.0	124.0	730.0 ;	NCP	CDC	OAE
72	74	76	2	115.0	610.0	115.0	610.0 ;	NCP	CDC	CAC
75	74	76	2	121.0	685.0	121.0	685.0 ;	OAE	CDC	CAC
66	77	71	2	109.5	520.0	109.5	520.0 ;	ODA	CEE	CEA
66	77	78	2	109.5	520.0	109.5	520.0 ;	ODA	CEE	CDN
71	77	78	2	109.5	520.0	109.5	520.0 ;	CEA	CEE	CDN
77	78	79	2	109.5	520.0	109.5	520.0 ;	CEE	CDN	OAU
77	78	81	2	109.5	520.0	109.5	520.0 ;	CEE	CDN	CDH
79	78	81	2	109.5	520.0	109.5	520.0 ;	OAU	CDN	CDH
78	79	80	2	109.5	450.0	109.5	450.0 ;	CDN	OAU	HAU
78	81	82	2	109.5	520.0	109.5	520.0 ;	CDN	CDH	OAO
78	81	84	2	109.5	520.0	109.5	520.0 ;	CDN	CDH	CBE
82	81	84	2	109.5	520.0	109.5	520.0 ;	OAO	CDH	CBE
81	82	83	2	109.5	450.0	109.5	450.0 ;	CDH	OAO	HAO
81	84	85	2	109.5	520.0	109.5	520.0 ;	CDH	CBE	OAI
84	85	86	2	109.5	450.0	109.5	450.0 ;	CBE	OAI	H75
56	87	88	2	109.5	380.0	109.5	380.0 ;	CDY	OCV	CDU
87	88	89	2	109.5	320.0	109.5	320.0 ;	OCV	CDU	CBI

87	88	92	2	109.5	520.0	109.5	520.0 ;	OCV	CDU	CEF
89	88	92	2	109.5	520.0	109.5	520.0 ;	CBI	CDU	CEF
88	89	90	2	109.5	520.0	109.5	520.0 ;	CDU	CBI	OAM
89	90	91	2	109.5	450.0	109.5	450.0 ;	CBI	OAM	HAM
60	92	88	2	109.5	520.0	109.5	520.0 ;	CEG	CEF	CDU
60	92	93	2	109.5	520.0	109.5	520.0 ;	CEG	CEF	OCZ
88	92	93	2	109.5	520.0	109.5	520.0 ;	CDU	CEF	OCZ
92	93	94	2	109.5	380.0	109.5	380.0 ;	CEF	OCZ	CDZ
93	94	95	2	109.5	520.0	109.5	520.0 ;	OCZ	CDZ	CEB
93	94	101	2	109.5	320.0	109.5	320.0 ;	OCZ	CDZ	OCT
95	94	101	2	109.5	520.0	109.5	520.0 ;	CEB	CDZ	OCT
94	95	96	2	109.5	520.0	109.5	520.0 ;	CDZ	CEB	NCQ
94	95	109	2	109.5	520.0	109.5	520.0 ;	CDZ	CEB	CED
96	95	109	2	109.5	520.0	109.5	520.0 ;	NCQ	CEB	CED
95	96	97	2	120.0	505.0	120.0	505.0 ;	CEB	NCQ	HCQ
95	96	98	2	122.0	700.0	122.0	700.0 ;	CEB	NCQ	CDD
97	96	98	2	123.0	415.0	123.0	415.0 ;	HCQ	NCQ	CDD
96	98	99	2	124.0	730.0	124.0	730.0 ;	NCQ	CDD	OAF
96	98	100	2	115.0	610.0	115.0	610.0 ;	NCQ	CDD	CAD
99	98	100	2	121.0	685.0	121.0	685.0 ;	OAF	CDD	CAD
94	101	102	2	109.5	380.0	109.5	380.0 ;	CDZ	OCT	CDS
101	102	103	2	109.5	320.0	109.5	320.0 ;	OCT	CDS	CBG
101	102	106	2	109.5	520.0	109.5	520.0 ;	OCT	CDS	CDP
103	102	106	2	109.5	520.0	109.5	520.0 ;	CBG	CDS	CDP
102	103	104	2	109.5	520.0	109.5	520.0 ;	CDS	CBG	OAK
103	104	105	2	109.5	450.0	109.5	450.0 ;	CBG	OAK	H77
102	106	107	2	109.5	520.0	109.5	520.0 ;	CDS	CDP	OAW
102	106	109	2	109.5	520.0	109.5	520.0 ;	CDS	CDP	CED
107	106	109	2	109.5	520.0	109.5	520.0 ;	OAW	CDP	CED
106	107	108	2	109.5	450.0	109.5	450.0 ;	CDP	OAW	HAW
95	109	106	2	109.5	520.0	109.5	520.0 ;	CEB	CED	CDP
95	109	110	2	109.5	520.0	109.5	520.0 ;	CEB	CED	01
106	109	110	2	109.5	520.0	109.5	520.0 ;	CDP	CED	01
109	110	111	2	109.5	380.0	109.5	380.0 ;	CED	01	C1
110	111	112	2	109.5	320.0	109.5	320.0 ;	01	C1	05
110	111	123	2	109.5	520.0	109.5	520.0 ;	01	C1	C2
112	111	123	2	109.5	520.0	109.5	520.0 ;	05	C1	C2
111	112	113	2	109.5	380.0	109.5	380.0 ;	C1	05	C5
112	113	114	2	109.5	320.0	109.5	320.0 ;	05	C5	C6
112	113	117	2	109.5	520.0	109.5	520.0 ;	05	C5	C4
114	113	117	2	109.5	520.0	109.5	520.0 ;	C6	C5	C4
113	114	115	2	109.5	520.0	109.5	520.0 ;	C5	C6	06
114	115	116	2	109.5	450.0	109.5	450.0 ;	C6	06	H76
113	117	118	2	109.5	520.0	109.5	520.0 ;	C5	C4	04
113	117	120	2	109.5	520.0	109.5	520.0 ;	C5	C4	C3

118	117	120	2	109.5	520.0	109.5	520.0	;	O4	C4	C3
117	118	119	2	109.5	450.0	109.5	450.0	;	C4	O4	H42
117	120	121	2	109.5	520.0	109.5	520.0	;	C4	C3	O3
117	120	123	2	109.5	520.0	109.5	520.0	;	C4	C3	C2
121	120	123	2	109.5	520.0	109.5	520.0	;	O3	C3	C2
120	121	122	2	109.5	450.0	109.5	450.0	;	C3	O3	H32
111	123	120	2	109.5	520.0	109.5	520.0	;	C1	C2	C3
111	123	124	2	109.5	520.0	109.5	520.0	;	C1	C2	O2
120	123	124	2	109.5	520.0	109.5	520.0	;	C3	C2	O2
123	124	125	2	109.5	450.0	109.5	450.0	;	C2	O2	H22

[dihedrals]

```

; ai aj ak al fu c0, c1, m, ...
18 17 19 20 2 0.0 167.4 0.0 167.4 ; imp CDF CCN OAH NCR
20 18 21 22 2 0.0 167.4 0.0 167.4 ; imp NCR CDF HCR CDV
22 20 23 41 2 35.3 334.8 35.3 334.8 ; imp CDV NCR CDG OCW
23 22 24 26 2 35.3 334.8 35.3 334.8 ; imp CDG CDV OAN CBA
42 41 49 43 2 35.3 334.8 35.3 334.8 ; imp CDW OCW OCU CDM
43 42 44 46 2 35.3 334.8 35.3 334.8 ; imp CDM CDW OAT CDO
46 43 54 47 2 35.3 334.8 35.3 334.8 ; imp CDO CDM CEC OAV
50 49 51 54 2 35.3 334.8 35.3 334.8 ; imp CDT OCU CBH CEC
54 55 46 50 2 35.3 334.8 35.3 334.8 ; imp CEC OCX CDO CDT
56 55 57 87 2 35.3 334.8 35.3 334.8 ; imp CDY OCX CDQ OCV
57 56 58 60 2 35.3 334.8 35.3 334.8 ; imp CDQ CDY OAX CEG
60 57 92 61 2 35.3 334.8 35.3 334.8 ; imp CEG CDQ CEF ODB
62 61 66 63 2 35.3 334.8 35.3 334.8 ; imp CEH ODB ODA CDE
63 62 64 65 2 0.0 167.4 0.0 167.4 ; imp CDE CEH OAY OAG
68 67 69 71 2 35.3 334.8 35.3 334.8 ; imp CDI CCO OAP CEA
71 68 77 72 2 35.3 334.8 35.3 334.8 ; imp CEA CDI CEE NCP
72 71 73 74 2 0.0 167.4 0.0 167.4 ; imp NCP CEA HCP CDC
74 72 75 76 2 0.0 167.4 0.0 167.4 ; imp CDC NCP OAE CAC
77 78 66 71 2 35.3 334.8 35.3 334.8 ; imp CEE CDN ODA CEA
78 77 79 81 2 35.3 334.8 35.3 334.8 ; imp CDN CEE OAU CDH
81 78 82 84 2 35.3 334.8 35.3 334.8 ; imp CDH CDN OAO CBE
88 87 89 92 2 35.3 334.8 35.3 334.8 ; imp CDU OCV CBI CEF
92 93 88 60 2 35.3 334.8 35.3 334.8 ; imp CEF OCZ CDU CEG
94 93 95 101 2 35.3 334.8 35.3 334.8 ; imp CDZ OCZ CEB OCT
95 94 96 109 2 35.3 334.8 35.3 334.8 ; imp CEB CDZ NCQ CED
96 95 97 98 2 0.0 167.4 0.0 167.4 ; imp NCQ CEB HCQ CDD
98 96 99 100 2 0.0 167.4 0.0 167.4 ; imp CDD NCQ OAF CAD
102 101 103 106 2 35.3 334.8 35.3 334.8 ; imp CDS OCT CBG CDP
106 102 107 109 2 35.3 334.8 35.3 334.8 ; imp CDP CDS OAW CED
109 110 95 106 2 35.3 334.8 35.3 334.8 ; imp CED O1 CEB CDP
111 110 112 123 2 35.3 334.8 35.3 334.8 ; imp C1 O1 O5 C2
113 112 114 117 2 35.3 334.8 35.3 334.8 ; imp C5 O5 C6 C4

```

117	113	118	120	2	35.3	334.8	35.3	334.8	; imp	C4	C5	O4	C3
120	117	121	123	2	35.3	334.8	35.3	334.8	; imp	C3	C4	O3	C2
123	124	120	111	2	35.3	334.8	35.3	334.8	; imp	C2	O2	C3	C1
1	6	5	4	1	0.0	5.9 3	0.0	5.9 3	; dih	CBW	CBT	CBQ	CBN
2	1	6	5	1	0.0	5.9 3	0.0	5.9 3	; dih	CBY	CBW	CBT	CBQ
3	2	1	6	1	0.0	5.9 3	0.0	5.9 3	; dih	CBZ	CBY	CBW	CBT
4	3	2	1	1	0.0	5.9 3	0.0	5.9 3	; dih	CCA	CBZ	CBY	CBW
5	4	3	2	1	0.0	5.9 3	0.0	5.9 3	; dih	CCB	CCA	CBZ	CBY
6	5	4	3	1	0.0	5.9 3	0.0	5.9 3	; dih	CCC	CCB	CCA	CBZ
7	6	5	4	1	0.0	5.9 3	0.0	5.9 3	; dih	CCD	CCC	CCB	CCA
8	7	6	5	1	0.0	5.9 3	0.0	5.9 3	; dih	CCE	CCD	CCC	CCB
9	8	7	6	1	0.0	5.9 3	0.0	5.9 3	; dih	CCF	CCE	CCD	CCC
10	9	8	7	1	0.0	5.9 3	0.0	5.9 3	; dih	CCG	CCF	CCE	CCD
11	10	9	8	1	0.0	5.9 3	0.0	5.9 3	; dih	CCH	CCG	CCF	CCE
12	11	10	9	1	0.0	5.9 3	0.0	5.9 3	; dih	CCI	CCH	CCG	CCF
13	12	11	10	1	0.0	5.9 3	0.0	5.9 3	; dih	CCJ	CCI	CCH	CCG
14	13	12	11	1	0.0	5.9 3	0.0	5.9 3	; dih	CCK	CCJ	CCI	CCH
15	14	13	12	1	0.0	5.9 3	0.0	5.9 3	; dih	CCL	CCK	CCJ	CCI
16	15	14	13	1	0.0	5.9 3	0.0	5.9 3	; dih	CCM	CCL	CCK	CCJ
17	16	15	14	1	0.0	5.9 3	0.0	5.9 3	; dih	CCN	CCM	CCL	CCK
18	17	16	15	1	0.0	5.9 3	0.0	5.9 3	; dih	CCD	CCN	CCM	CCL
16	17	18	20	1	0.0	1.0 6	0.0	1.0 6	; dih	CCM	CCN	CCD	NCR
17	18	20	22	1	180.0	33.5 2	180.0	33.5 2	; dih	CCN	CCD	NCR	CDV
41	22	20	18	1	180.0	1.0 6	180.0	1.0 6	; dih	OCW	CDV	NCR	CDF
26	23	22	20	1	0.0	5.9 3	0.0	5.9 3	; dih	CBA	CDG	CDV	NCR
20	22	41	42	1	0.0	3.8 3	0.0	3.8 3	; dih	NCR	CDV	OCW	CDW
22	23	24	25	1	0.0	1.3 3	0.0	1.3 3	; dih	CDV	CDG	OAN	HAN
22	23	26	27	1	0.0	1.0 6	0.0	1.0 6	; dih	CDV	CDG	CBA	CAZ
28	27	26	23	1	180.0	5.9 2	180.0	5.9 2	; dih	CBD	CAZ	CBA	CDG
29	28	27	26	1	0.0	1.0 6	0.0	1.0 6	; dih	CBL	CBD	CAZ	CBA
30	29	28	27	1	0.0	5.9 3	0.0	5.9 3	; dih	CBO	CBL	CBD	CAZ
31	30	29	28	1	0.0	5.9 3	0.0	5.9 3	; dih	CBR	CBO	CBL	CBD
32	31	30	29	1	0.0	5.9 3	0.0	5.9 3	; dih	CBU	CBR	CBO	CBL
33	32	31	30	1	0.0	5.9 3	0.0	5.9 3	; dih	CBX	CBU	CBR	CBO
34	33	32	31	1	0.0	5.9 3	0.0	5.9 3	; dih	CBV	CBX	CBU	CBR
35	34	33	32	1	0.0	5.9 3	0.0	5.9 3	; dih	CBS	CBV	CBX	CBU
36	35	34	33	1	0.0	5.9 3	0.0	5.9 3	; dih	CBP	CBS	CBV	CBX
37	36	35	34	1	0.0	5.9 3	0.0	5.9 3	; dih	CBM	CBP	CBS	CBV
38	37	36	35	1	0.0	5.9 3	0.0	5.9 3	; dih	CBJ	CBM	CBP	CBS
39	38	37	36	1	0.0	5.9 3	0.0	5.9 3	; dih	CBB	CBJ	CBM	CBP
40	39	38	37	1	0.0	5.9 3	0.0	5.9 3	; dih	CAA	CBB	CBJ	CBM
49	42	41	22	1	0.0	3.8 3	0.0	3.8 3	; dih	OCU	CDW	OCW	CDV
46	43	42	41	1	0.0	5.9 3	0.0	5.9 3	; dih	CDO	CDM	CDW	OCW
41	42	49	50	1	0.0	3.8 3	0.0	3.8 3	; dih	OCW	CDW	OCU	CDT
42	43	44	45	1	0.0	1.3 3	0.0	1.3 3	; dih	CDW	CDM	OAT	HAT

54	46	43	42	1	0.0	5.9	3	0.0	5.9	3 ; dih	CEC	CDO	CDM	CDW
43	46	47	48	1	0.0	1.3	3	0.0	1.3	3 ; dih	CDM	CDO	OAV	HAV
55	54	46	43	1	0.0	5.9	3	0.0	5.9	3 ; dih	OCX	CEC	CDO	CDM
54	50	49	42	1	0.0	3.8	3	0.0	3.8	3 ; dih	CEC	CDT	OCU	CDW
52	51	50	49	1	0.0	5.9	3	0.0	5.9	3 ; dih	OAL	CBH	CDT	OCU
55	54	50	49	1	0.0	5.9	3	0.0	5.9	3 ; dih	OCX	CEC	CDT	OCU
50	51	52	53	1	0.0	1.3	3	0.0	1.3	3 ; dih	CDT	CBH	OAL	H78
46	54	55	56	1	0.0	3.8	3	0.0	3.8	3 ; dih	CDO	CEC	OCX	CDY
87	56	55	54	1	0.0	3.8	3	0.0	3.8	3 ; dih	OCV	CDY	OCX	CEC
60	57	56	55	1	0.0	5.9	3	0.0	5.9	3 ; dih	CEG	CDQ	CDY	OCX
55	56	87	88	1	0.0	3.8	3	0.0	3.8	3 ; dih	OCX	CDY	OCV	CDU
56	57	58	59	1	0.0	1.3	3	0.0	1.3	3 ; dih	CDY	CDQ	OAX	HAX
92	60	57	56	1	0.0	5.9	3	0.0	5.9	3 ; dih	CEF	CEG	CDQ	CDY
57	60	61	62	1	0.0	3.8	3	0.0	3.8	3 ; dih	CDQ	CEG	ODB	CEH
93	92	60	57	1	0.0	5.9	3	0.0	5.9	3 ; dih	OCZ	CEF	CEG	CDQ
67	62	61	60	1	0.0	3.8	3	0.0	3.8	3 ; dih	CCO	CEH	ODB	CEG
61	62	63	65	1	0.0	1.0	6	0.0	1.0	6 ; dih	ODB	CEH	CDE	OAG
61	62	66	77	1	0.0	3.8	3	0.0	3.8	3 ; dih	ODB	CEH	ODA	CEE
68	67	62	61	1	0.0	5.9	3	0.0	5.9	3 ; dih	CDI	CCO	CEH	ODB
78	77	66	62	1	0.0	3.8	3	0.0	3.8	3 ; dih	CDN	CEE	ODA	CEH
71	68	67	62	1	0.0	5.9	3	0.0	5.9	3 ; dih	CEA	CDI	CCO	CEH
67	68	69	70	1	0.0	1.3	3	0.0	1.3	3 ; dih	CCO	CDI	OAP	HAP
77	71	68	67	1	0.0	5.9	3	0.0	5.9	3 ; dih	CEE	CEA	CDI	CCO
68	71	72	74	1	0.0	3.8	3	0.0	3.8	3 ; dih	CDI	CEA	NCP	CDC
78	77	71	68	1	0.0	5.9	3	0.0	5.9	3 ; dih	CDN	CEE	CEA	CDI
76	74	72	71	1	180.0	33.5	2	180.0	33.5	2 ; dih	CAC	CDC	NCP	CEA
81	78	77	66	1	0.0	5.9	3	0.0	5.9	3 ; dih	CDH	CDN	CEE	ODA
77	78	79	80	1	0.0	1.3	3	0.0	1.3	3 ; dih	CEE	CDN	OAU	HAU
84	81	78	77	1	0.0	5.9	3	0.0	5.9	3 ; dih	CBE	CDH	CDN	CEE
78	81	82	83	1	0.0	1.3	3	0.0	1.3	3 ; dih	CDN	CDH	OAO	HAO
85	84	81	78	1	0.0	5.9	3	0.0	5.9	3 ; dih	OAI	CBE	CDH	CDN
81	84	85	86	1	0.0	1.3	3	0.0	1.3	3 ; dih	CDH	CBE	OAI	H75
92	88	87	56	1	0.0	3.8	3	0.0	3.8	3 ; dih	CEF	CDU	OCV	CDY
90	89	88	87	1	0.0	5.9	3	0.0	5.9	3 ; dih	OAM	CBI	CDU	OCV
93	92	88	87	1	0.0	5.9	3	0.0	5.9	3 ; dih	OCZ	CEF	CDU	OCV
88	89	90	91	1	0.0	1.3	3	0.0	1.3	3 ; dih	CDU	CBI	OAM	HAM
60	92	93	94	1	0.0	3.8	3	0.0	3.8	3 ; dih	CEG	CEF	OCZ	CDZ
101	94	93	92	1	0.0	3.8	3	0.0	3.8	3 ; dih	OCT	CDZ	OCZ	CEF
109	95	94	93	1	0.0	5.9	3	0.0	5.9	3 ; dih	CED	CEB	CDZ	OCZ
93	94	101	102	1	0.0	3.8	3	0.0	3.8	3 ; dih	OCZ	CDZ	OCT	CDS
94	95	96	98	1	0.0	3.8	3	0.0	3.8	3 ; dih	CDZ	CEB	NCQ	CDD
110	109	95	94	1	0.0	5.9	3	0.0	5.9	3 ; dih	O1	CED	CEB	CDZ
100	98	96	95	1	180.0	33.5	2	180.0	33.5	2 ; dih	CAD	CDD	NCQ	CEB
106	102	101	94	1	0.0	3.8	3	0.0	3.8	3 ; dih	CDP	CDS	OCT	CDZ
104	103	102	101	1	0.0	5.9	3	0.0	5.9	3 ; dih	OAK	CBG	CDS	OCT

```
109 106 102 101 1 0.0 5.9 3 0.0 5.9 3 ; dih CED CDP CDS OCT
102 103 104 105 1 0.0 1.3 3 0.0 1.3 3 ; dih CDS CBG OAK H77
102 106 107 108 1 0.0 1.3 3 0.0 1.3 3 ; dih CDS CDP OAW HAW
110 109 106 102 1 0.0 5.9 3 0.0 5.9 3 ; dih O1 CED CDP CDS
 95 109 110 111 1 0.0 3.8 3 0.0 3.8 3 ; dih CEB CED O1 C1
123 111 110 109 1 0.0 3.8 3 0.0 3.8 3 ; dih C2 C1 O1 CED
110 111 112 113 1 0.0 3.8 3 0.0 3.8 3 ; dih O1 C1 O5 C5
124 123 111 110 1 0.0 5.9 3 0.0 5.9 3 ; dih O2 C2 C1 O1
117 113 112 111 1 0.0 3.8 3 0.0 3.8 3 ; dih C4 C5 O5 C1
115 114 113 112 1 0.0 5.9 3 0.0 5.9 3 ; dih O6 C6 C5 O5
120 117 113 112 1 0.0 5.9 3 0.0 5.9 3 ; dih C3 C4 C5 O5
113 114 115 116 1 0.0 1.3 3 0.0 1.3 3 ; dih C5 C6 O6 H76
113 117 118 119 1 0.0 1.3 3 0.0 1.3 3 ; dih C5 C4 O4 H42
123 120 117 113 1 0.0 5.9 3 0.0 5.9 3 ; dih C2 C3 C4 C5
117 120 121 122 1 0.0 1.3 3 0.0 1.3 3 ; dih C4 C3 O3 H32
124 123 120 117 1 0.0 5.9 3 0.0 5.9 3 ; dih O2 C2 C3 C4
111 123 124 125 1 0.0 1.3 3 0.0 1.3 3 ; dih C1 C2 O2 H22
```

```
; Position restraints for GM1
#ifdef POSRES_LIPID
#include "posre_gm1.itp"
#endif
```

Appendix D

Example .mdp Files

Several general-use .mdp files are listed in this Appendix. These files provide the framework for almost all the simulations conducted in this body of work.

D.1 Generic twin-range cutoff NPT simulation (GROMOS96)

```
title                = MD simulation
; Run parameters
integrator           = md
dt                   = 0.002
nsteps               = 25000000 ; 50000 ps (50 ns)
nstcomm              = 1
; Output parameters
nstxout              = 0          ; suppress .trr output
nstvout              = 0
nstfout              = 0
nstxtcout            = 5000       ; every 10 ps
nstenergy            = 5000       ; every 10 ps
nstlog               = 5000       ; every 10 ps
; Bond parameters
constraint_algorithm = lincs
constraints          = all-bonds
continuation         = yes        ; continuing from NPT
; Twin-range cutoff scheme, parameters for Gromos96
nstlist              = 5
ns_type              = grid
rlist                = 0.9
```

```
rcoulomb          = 0.9
rvdw              = 1.4
; PME electrostatics parameters
coulombtype       = PME
fourierspacing    = 0.12
fourier_nx        = 0
fourier_ny        = 0
fourier_nz        = 0
pme_order         = 4
ewald_rtol        = 1e-5
optimize_fft      = yes
; Berendsen temperature coupling is on in two groups
Tcoupl           = Nose-Hoover
tc_grps          = Protein Non-Protein
tau_t            = 1.0    1.0
ref_t            = 310    310
; Pressure coupling is on
Pcoupl           = Parrinello-Rahman
pcoupltype       = isotropic
tau_p            = 1.0
compressibility   = 4.5e-5
ref_p            = 1.0
; Generate velocities is off
gen_vel          = no
; Periodic boundary conditions are on in all directions
pbc              = xyz
; Long-range dispersion correction
DispCorr         = EnerPres
```

D.2 Simulated annealing simulation for a membrane protein (GROMOS96)

```
title                = NPT annealing for a membrane protein
define               = -DPOSRES
; Run parameters
integrator           = md
dt                   = 0.002
nsteps               = 250000          ; 500 ps
nstcomm              = 1
; Output parameters
nstxout              = 0
nstvout              = 0
nstfout              = 0
nstlog               = 1000
nstenergy            = 1000
; Bond parameters
constraint_algorithm = lincs
constraints           = all-bonds
continuation         = yes           ; continuing from NVT
; Single interaction range, hybrid Gromos96/Berger
nstlist              = 5
ns_type              = grid
rlist                 = 1.2
rcoulomb              = 1.2
rvdw                  = 1.2
; PME electrostatics parameters
coulombtype          = PME
fourierspacing       = 0.12
fourier_nx            = 0
fourier_ny            = 0
fourier_nz            = 0
pme_order             = 4
ewald_rtol            = 1e-5
optimize_fft         = yes
; Berendsen temperature coupling is on in three groups
Tcoupl               = Berendsen
tc_grps               = Protein POPC SOL_NA+_CL-
tau_t                 = 0.1      0.1  0.1
ref_t                 = 100      100  100
; Pressure coupling is on
Pcoupl               = Berendsen
pcoupltype            = semiisotropic
tau_p                 = 2.0
```

```
compressibility      = 4.5e-5 4.5e-5
ref_p                = 1.0 1.0
; Generate velocities is off
gen_vel              = no
; Periodic boundary conditions are on in all directions
pbc                  = xyz
; Long-range dispersion correction
DispCorr              = EnerPres
; Simulated annealing
annealing             = single   single   single
annealing_npoints     = 2        2        2
annealing_time        = 0   500  0   500  0   500
annealing_temp         = 100 310 100 310 100 310
```


D.3 NPT simulation involving morin (GROMOS96)

```
title                = MD simulation
; Run parameters
integrator           = md
dt                   = 0.002
nsteps               = 25000000 ; 50000 ps (50 ns)
nstcomm              = 1
; Output parameters
nstxout              = 0          ; suppress .trr output
nstvout              = 0
nstfout              = 0
nstxtcout            = 5000       ; every 10 ps
nstenergy            = 5000       ; every 10 ps
nstlog               = 5000       ; every 10 ps
; Bond parameters
constraint_algorithm = lincs
constraints          = all-bonds
continuation         = yes        ; continuing from NPT
; Single interaction range, parameters for Gromos96
nstlist              = 5
ns_type              = grid
rlist                = 1.4
rcoulomb             = 1.4
rvdw                 = 1.4
; PME electrostatics parameters
coulombtype          = PME
fourierspacing       = 0.12
fourier_nx           = 0
fourier_ny           = 0
fourier_nz           = 0
pme_order            = 4
ewald_rtol           = 1e-5
optimize_fft         = yes
; Berendsen temperature coupling is on in two groups
Tcoupl               = Nose-Hoover
tc_grps              = Protein Non-Protein
tau_t                = 1.0      1.0
ref_t                = 310      310
; Pressure coupling is on
Pcoupl               = Parrinello-Rahman
pcoupltype           = isotropic
tau_p                = 1.0
compressibility       = 4.5e-5
```

```
ref_p                = 1.0
; Generate velocities is off
gen_vel              = no
; Periodic boundary conditions are on in all directions
pbc                  = xyz
; Long-range dispersion correction
DispCorr              = EnerPres
; Dihedral restraints for morin
dihre                 = yes
dihre_fc              = 100
dihre_tau             = 1.0
nstdihreout          = 50
```

D.4 Steered MD simulation

```
title                = Harmonic pulling simulation (SMD)
define               = -DPOSRES_B
; Run parameters
integrator           = md
dt                   = 0.002
tinit                = 0
nsteps               = 250000      ; 500 ps
nstcomm              = 10
; Output parameters
nstxout              = 5000         ; every 10 ps
nstvout              = 5000
nstfout              = 500
nstxtcout            = 500         ; every 1 ps
nstenergy            = 500
; Bond parameters
constraint_algorithm = lincs
constraints           = all-bonds
continuation         = yes         ; continuing from NPT
; Single interaction range
nstlist              = 5
ns_type               = grid
rlist                 = 1.4
rcoulomb              = 1.4
rvdw                  = 1.4
; PME electrostatics parameters
coulombtype           = PME
fourierspacing        = 0.12
fourier_nx            = 0
fourier_ny            = 0
fourier_nz            = 0
pme_order             = 4
ewald_rtol            = 1e-5
optimize_fft          = yes
; Berendsen temperature coupling is on in two groups
Tcoupl                = Nose-Hoover
tc_grps               = Protein    Non-Protein
tau_t                 = 0.5        0.5
ref_t                 = 310        310
; Pressure coupling is on
Pcoupl                = Parrinello-Rahman
pcoupltype            = isotropic
tau_p                 = 1.0
```

```
compressibility      = 4.5e-5
ref_p                = 1.0
; Generate velocities is off
gen_vel              = no
; Periodic boundary conditions are on in all directions
pbc                  = xyz
; Long-range dispersion correction
DispCorr             = EnerPres
; Pull code
pull                 = umbrella
pull_geometry        = distance ; simple distance increase
pull_dim             = N N Y    ; pull along z-axis only
pull_start           = yes      ; define initial COM distance > 0
pull_ngroups         = 1
pull_group0          = Chain_B  ; reference group
pull_group1          = Chain_A  ; group that moves
pull_rate1           = 0.01     ; 0.01 nm per ps = 10 nm per ns
pull_k1              = 1000     ; kJ mol-1 nm-2
```

D.5 Umbrella sampling simulation

```
title                = Umbrella sampling simulation
define              = -DPOSRES_B
; Run parameters
integrator          = md
dt                  = 0.002
tinit               = 0
nsteps              = 250000      ; 500 ps
nstcomm             = 10
; Output parameters
nstxout             = 5000        ; every 10 ps
nstvout             = 5000
nstfout             = 500
nstxtcout           = 500        ; every 1 ps
nstenergy           = 500
; Bond parameters
constraint_algorithm = lincs
constraints          = all-bonds
continuation         = yes        ; continuing from NPT
; Single interaction range
nstlist             = 5
ns_type             = grid
rlist               = 1.4
rcoulomb            = 1.4
rvdw                = 1.4
; PME electrostatics parameters
coulombtype         = PME
fourierspacing      = 0.12
fourier_nx          = 0
fourier_ny          = 0
fourier_nz          = 0
pme_order           = 4
ewald_rtol          = 1e-5
optimize_fft        = yes
; Berendsen temperature coupling is on in two groups
Tcoupl              = Nose-Hoover
tc_grps             = Protein    Non-Protein
tau_t               = 0.5        0.5
ref_t               = 310        310
; Pressure coupling is on
Pcoupl              = Parrinello-Rahman
pcoupltype          = isotropic
tau_p               = 1.0
```

```
compressibility      = 4.5e-5
ref_p                = 1.0
; Generate velocities is off
gen_vel              = no
; Periodic boundary conditions are on in all directions
pbc                  = xyz
; Long-range dispersion correction
DispCorr             = EnerPres
; Pull code
pull                  = umbrella
pull_geometry        = distance
pull_dim              = N N Y      ; restrain along z-axis only
pull_start            = yes         ; define initial COM distance > 0
pull_ngroups         = 1
pull_group0           = Chain_B    ; reference group
pull_group1           = Chain_A    ; group that is restrained
pull_rate1            = 0.0         ; restrain in place
pull_k1               = 1000        ; kJ mol-1 nm-2
pull_nstxout          = 1000        ; every 2 ps
pull_nstfout          = 1000        ; every 2 ps
```


Appendix E

Perl Scripts

In this Appendix, several Perl scripts are provided. These scripts have been used in data collection and analysis in the work presented here.

E.1 Average multiple data files

```
#!/usr/bin/perl -w

use strict;

# Opens multiple .dat/.xvg files and averages the results of the equivalent lines
# An equivalent line is defined as one in which the first (x) value is the same
# The first two columns are used, any additional entries are ignored
# Input is expected to be some generic text (data) file, the number of inputs is
# determined from the size of @ARGV
#
# Data files are assumed to be text only, with XmGrace-style headers (# and @)

unless(@ARGV) {
    die "Usage: perl $0 input1.dat/.xvg input2.dat/.xvg ...\n";
    exit;
}

# determine number of inputs
my $n = scalar(@ARGV);

my @big_array;
```

```
# open a series of filehandles to place each text file into an array
for (my $i=0; $i<$n; $i++) {
    open(IN, "<$ARGV[$i]") or die "Cannot open $ARGV[$i]: $!\n";
    my @array = <IN>;
    close(IN);

    # concatenate all inputs into one big array
    foreach $_ (@array) {
        unless ($_ =~ /^[#@]/) {
            push(@big_array, $_);
        }
    }
}

# splice out duplicate x-values
my @x_values;

foreach $_ (@big_array) {
    my @line = split(" ", $_);
    push(@x_values, $line[0]);
}

for (my $j=0; $j<scalar(@x_values); $j++) {
    for (my $k=$j+1; $k<scalar(@x_values); $k++) {
        if ($x_values[$k] == $x_values[$j]) {
            splice(@x_values, $k, 1);
        }
    }
}

# remove duplicate entry at the end
pop(@x_values);

# Now that there's a list of x-values, loop thru @big_array again, collecting
# the y-values that correspond to those x-values

foreach $_ (@x_values) {

    my @y_values;

    for (my $m=0; $m<scalar(@big_array); $m++) {
        my @line = split(" ", $big_array[$m]);
        if ($line[0] == $_) {
```

```
        push(@y_values, $line[1]);
    }
}

my $sum = 0;

foreach $_ (@y_values) {
    $sum += $_;
}

# determine an average
my $average = $sum / (scalar(@y_values));

# determine the standard deviation:
# 1. for each value, subtract the average and square this difference
# 2. sum these differences, divide by number of values
# 3. take square root

my $sd_sum = 0;

foreach my $value (@y_values) {
    $sd_sum += ($value - $average)**2;
}

my $sd_dividend = $sd_sum / (scalar(@y_values));

my $std_dev = sqrt($sd_dividend);

# open output file
open(OUT, ">>output.dat") or die "Cannot open output: $!\n";
printf OUT "%d %8.3f %8.3f\n", $_, $average, $std_dev;
close(OUT);
}

exit;
```

E.2 Calculate x-y distances

```
#!/usr/bin/perl

use strict;

# Program reads an .xvg file produced by g_dist and calculates the x-y
# distance rather than the full x-y-z distance. This program is mainly
# useful for measuring distances between objects within the same
# plane, i.e. a membrane

unless (@ARGV) {
    die "Cannot open input file! Usage: perl $0 input.xvg\n";
}

my $file = $ARGV[0];

open(IN, "<$file") || die "Cannot open input: $!\n";
my @in = <IN>;
close(IN);

my @times;
my @x;
my @y;

my @clean;

# remove header
foreach $_ (@in) {
    unless ($_ =~ /^[#@]/) {
        push(@clean, $_);
    }
}

# open up the file, ignore comment lines, parse out
# useful stuff (x, y dist components)
for (my $i=0; $i<(scalar(@clean)); $i++) {

    unless ($clean[$i] =~ /^[#@]/) {
        my @line = split(" ", $clean[$i]);
        $times[$i] = $line[0];
        $x[$i] = $line[2];
        $y[$i] = $line[3];
    }
}
```

```
}

# calculate distances
my @distances;

for (my $j=0; $j<(scalar(@clean)); $j++) {
    $distances[$j] = sqrt($x[$j]**2 + $y[$j]**2);
}

open(OUT, ">>xy_dist.svg") || die "Cannot open output: $!\n";
for (my $k=0; $k<(scalar(@clean)); $k++) {
    printf OUT "%8.3f%8.4f\n", $times[$k], $distances[$k];
}

exit;
```

E.3 Parse a data set

```
#!/usr/bin/perl -w

use strict;

# Open a data file with multiple data sets, parse out a single set of data
# The data set is specified by the user

unless(@ARGV) {
    die "Usage: perl $0 -i data.xvg -o out.xvg -set X\n";
}

my %args = @ARGV;

my $input;
my $output;
my $nset;

if (exists($args{"-i"})) {
    $input = $args{"-i"};
} else {
    die "No input specified!\n";
}

if (exists($args{"-o"})) {
    $output = $args{"-o"};
} else {
    print "No output filename specified. Using default \"out.xvg.\n";
}

if (exists($args{"-set"})) {
    $nset = $args{"-set"};
} else {
    die "No set number specified!\n";
}

open(IN, "<$input") || die "Cannot open input!\n";
my @in = <IN>;
close(IN);

open(OUT, ">>$output") || die "Cannot open output!\n";

foreach $_ (@in) {
```



```
if ($_ =~ /^#/) {
    print OUT $_;
} elsif ($_ =~ /^@/) {
    if ($_ =~ /^@\s+title/ || $_ =~ /^@\s+xaxis/ || $_ =~ /^@\s+yaxis/
        || $_ =~ /^@\@TYPE/) {
        print OUT $_;
    }
} else {
    my @line = split(" ", $_);
    my $time = $line[0];
    my $data = $line[$nset];

    printf OUT "%14.7f\t%14.7f\n", $time, $data;
}
}

exit;
```

E.4 Plot the contents of hbmap.xpm

```
#!/usr/bin/perl

#
# plot_hbmap.pl - plot the probability of finding a particular hydrogen bond
# based on several input files:
#   1. coordinate file (for atom naming) - MUST be a .pdb file with
#       NO CHAIN IDENTIFIERS
#   2. hbmap.xpm
#   3. hbond.ndx (modified to contain only the atom numbers in the [hbonds...]
#       section, nothing else)
#

use strict;

unless(@ARGV) {
    die "Usage: perl $0 -s structure.pdb -map hbmap.xpm -index hbond.ndx\n";
}

# define input hash
my %args = @ARGV;

# input variables
my $map_in;
my $struct;
my $ndx;

if (exists($args{"-map"})) {
    $map_in = $args{"-map"};
} else {
    die "No -map specified!\n";
}

if (exists($args{"-s"})) {
    $struct = $args{"-s"};
} else {
    die "No -s specified!\n";
}

if (exists($args{"-index"})) {
    $ndx = $args{"-index"};
} else {
    die "No -index specified!\n";
}
```

```
}

# open the input
open(MAP, "<$map_in") || die "Cannot open input map file!\n";
my @map = <MAP>;
close(MAP);

open(STRUCT, "<$struct") || die "Cannot open input coordinate file!\n";
my @coord = <STRUCT>;
close(STRUCT);

open(NDX, "<$ndx") || die "Cannot open input index file!\n";
my @index = <NDX>;
close(NDX);

# determine number of HB indices and frames
my $nres = 0;
my $nframes = 0;
for (my $i=0; $i<scalar(@map); $i++) {
    if ($map[$i] =~ /static char/) {
        my $res_line = $map[$i+1];
        my @info = split(" ", $res_line);

        $nframes = $info[0];
        my @nframes = split(' ', $nframes);
        shift(@nframes); # get rid of the "
        $nframes = join(' ', @nframes);

        $nres = $info[1];
    }
}

print "Processing the map file...\n";
print "There are $nres HB indices.\n";
print "There are $nframes frames.\n";

# initialize hashes for later output writing
# counter $a holds the HB index from hbond.ndx
my %hbonds;
for (my $a=0; $a<$nres; $a++) {
    $hbonds{$a+1} = 0;
}

# donor/acceptor hashes for bookkeeping purposes
my %donors;
```

```
for (my $b=1; $b<=$nres; $b++) {
    $donors{$b} = 0;
}

my %acceptors;
for (my $c=1; $c<=$nres; $c++) {
    $acceptors{$c} = 0;
}

# clean up the output - up to 18 lines of comments, etc.
splice(@map, 0, 18);

# remove any "x-axis" or "y-axis" lines
for (my $n=0; $n<scalar(@map); $n++) {
    if (($map[$n] =~ /x-axis/) || ($map[$n] =~ /y-axis/)) {
        shift(@map);
        $n--;
    }
}

# There should now be $nres lines left in the file
# The HB map for the last index is written first (top-down in .xpm file)
# * Element 0 is index $nres, element 1 is $nres-1, etc.

for (my $i=$nres; $i>=1; $i--) {
    # There will be $nframes+2 elements in @line (extra two are " at beginning
    # and end of the line)
    # Establish a conversion factor and split the input lines
    my $j = $nres - $i;
    my @line = split('', $map[$j]);

    # for each index, write to hash
    for (my $k=1; $k<=($nframes+1); $k++) {
        if ($line[$k] =~ /o/) {
            $hbonds{$i}++;
        }
    }
}

print "Processing the index file...\n";

# Open up the index file and work with it
for (my $n=0; $n<$nres; $n++) {
    my @line = split(" ", $index[$n]);
    $donors{$n+1} = $line[0];
}
```

```
    $acceptors{$n+1} = $line[2];
}

# some arrays for donor and acceptor atom names
my @donor_names;
my @donor_resn;
my @acceptor_names;
my @acceptor_resn;

# Open up the structure file and work with it
print "Processing coordinate file...\n";
foreach $_ (@coord) {
    my @line = split(" ", $_);
    my $natom = $line[1];
    my $name = $line[2];
    my $resn = $line[3];
    my $resnum = $line[4];

    if ($line[0] =~ /ATOM/) {
        # uncomment to remove H-bonds involving water
        # unless ($resn =~ /SOL/) {
            for (my $z=1; $z<=$nres; $z++) {
                if ($donors{$z} == $natom) {
                    $donor_names[$z] = $name;
                    $donor_resn[$z] = join('', $resn, $resnum);
                } elsif ($acceptors{$z} == $natom) {
                    $acceptor_names[$z] = $name;
                    $acceptor_resn[$z] = join('', $resn, $resnum);
                }
            }
        }
    }
}

# }

}

# open a single output file for writing
open(OUT, ">>summary_HBmap.dat") || die "Cannot open output file!\n";
printf(OUT "%10s\t%10s\t%10s\t%10s\t%10s\n", "# Donor", " ", "Acceptor", " ",
        " % Exist.");

for (my $o=1; $o<=$nres; $o++) {
    printf(OUT "%10s\t%10s\t%10s\t%10s\t%10.3f\n", $donor_resn[$o],
        $donor_names[$o], $acceptor_resn[$o], $acceptor_names[$o],
        (($hbonds{$o}/$nframes)*100));
}
}
```

```
close(OUT);
```

```
exit;
```


E.5 Plot secondary structure frequencies from ss.xpm

```
#!/usr/bin/perl

#
# plot_ss_xpm.pl - a program that reads in an ss.xpm file to generate
# individual plots of secondary structure propensity per residue
#

use strict;

# input defined as first command-line argument
my $input = $ARGV[0];

# open the input
open(IN, "<$input");
my @in = <IN>;
close(IN);

# determine number of residues and frames
my $nres = 0;
my $nframes = 0;
for (my $i=0; $i<scalar(@in); $i++) {
    if ($in[$i] =~ /static char/) {
        my $res_line = $in[$i+1];
        my @info = split(" ", $res_line);

        $nframes = $info[0];
        my @nframes = split(',', $nframes);
        shift(@nframes); # get rid of the "
        $nframes = join(',', @nframes);

        $nres = $info[1];
    }
}

print "I found $nres residues.\n";
print "There are $nframes frames.\n";

# initialize hashes for later output writing
my %a_helix;
for (my $a=1; $a<=$nres; $a++) {
    $a_helix{$a} = 0;
}
```

```
my %three_helix;
for (my $th=1; $th<=$nres; $th++) {
    $three_helix{$th} = 0;
}

my %pi_helix;
for (my $p=1; $p<=$nres; $p++) {
    $pi_helix{$p} = 0;
}

my %total_helix;
for (my $h=1; $h<=$nres; $h++) {
    $total_helix{$h} = 0;
}

my %sheet;
for (my $s=1; $s<=$nres; $s++) {
    $sheet{$s} = 0;
}

my %coil;
for (my $c=1; $c<=$nres; $c++) {
    $coil{$c} = 0;
}

my %turn;
for (my $t=1; $t<=$nres; $t++) {
    $turn{$t} = 0;
}

my %bridge;
for (my $b=1; $b<=$nres; $b++) {
    $bridge{$b} = 0;
}

my %bend;
for (my $b=1; $b<=$nres; $b++) {
    $bend{$b} = 0;
}

# The last nres lines have the secondary structure information
# So, if there are 40 residues that were analyzed, there are 40 lines
# of continuous secondary structure codes
```

```
# clean up the output - up to 18 lines of comments, etc.
splice(@in, 0, 18);

# remove any "x-axis" or "y-axis" lines
for (my $n=0; $n<scalar(@in); $n++) {
    if (($in[$n] =~ /x-axis/) || ($in[$n] =~ /y-axis/)) {
        shift(@in);
        $n--;
    }
}

# There should now be $nres lines left in the file
# The SS codes for the last residue are written first (top-down in .xpm file)
# * Element 0 is residue $nres, element 1 is $nres-1, etc.

for (my $i=$nres; $i>=1; $i--) {
    # There will be $nframes+2 elements in @line (extra two are " at beginning
    # and end of the line)
    # Establish a conversion factor and split the input lines
    my $j = $nres - $i;
    my @line = split(' ', $in[$j]);

    # for each type of ss, write to hashes

    for (my $k=1; $k<=($nframes+1); $k++) {
        if ($line[$k] =~ /H/) {
            $a_helix{$i}++;
            $total_helix{$i}++;
        } elsif ($line[$k] =~ /G/) {
            $three_helix{$i}++;
            $total_helix{$i}++;
        } elsif ($line[$k] =~ /I/) {
            $pi_helix{$i}++;
            $total_helix{$i}++;
        } elsif ($line[$k] =~ /E/) {
            $sheet{$i}++;
        } elsif ($line[$k] =~ /\~/) {
            $coil{$i}++;
        } elsif ($line[$k] =~ /T/) {
            $turn{$i}++;
        } elsif ($line[$k] =~ /B/) {
            $bridge{$i}++;
        } elsif ($line[$k] =~ /S/) {
            $bend{$i}++;
        }
    }
}
```

```
    }  
}  
  
# open a single output file for writing  
# each secondary structure type is a separate data series  
  
open(OUT, ">>summary_SS.xvg") or die "Cannot open \"summary_SS.dat\"\n";  
printf OUT "# Probability of various secondary structure elements, by residue.\n";  
printf OUT "@\ttitle\t\"Secondary Structure Content\"\n";  
printf OUT "@\txaxis\tlabel \"Residue\"\n";  
printf OUT "@\tyaxis\tlabel \"Probability\"\n";  
printf OUT "\@TYPE xy\n";  
printf OUT "@ s0 legend \"\\f{Symbol}a\\f{Times}-Helix\"\n";  
printf OUT "@ s1 legend \"\\f{Symbol}p\\f{Times}-Helix\"\n";  
printf OUT "@ s2 legend \"3\\s10\\N-Helix\"\n";  
printf OUT "@ s3 legend \"\\f{Symbol}b\\f{Times}-Strand\"\n";  
printf OUT "@ s4 legend \"\\f{Symbol}b\\f{Times}-Bend\"\n";  
printf OUT "@ s5 legend \"\\f{Symbol}b\\f{Times}-Turn\"\n";  
printf OUT "@ s6 legend \"\\f{Symbol}b\\f{Times}-Bridge\"\n";  
printf OUT "@ s7 legend \"Random Coil\"\n";  
  
for (my $o=1; $o<=$nres; $o++) {  
    printf OUT "%d\t%f\t%f\t%f\t%f\t%f\t%f\t%f\t%f\t%f\n", $o, ($a_helix{$o}/$nframes),  
        ($pi_helix{$o}/$nframes), ($three_helix{$o}/$nframes), ($sheet{$o}/$nframes),  
        ($bend{$o}/$nframes), ($turn{$o}/$nframes), ($bridge{$o}/$nframes),  
        ($coil{$o}/$nframes);  
}  
  
close(OUT);  
  
exit;
```

E.6 Create g_sham input

```
#!/usr/bin/perl -w

# This program reads in two .xvg files (specified on the command line) and prints
# out a concatenated output file suitable for input into g_sham for generating
# free energy landscapes.
# The output line contains, for example: time value1 value2
#
# Caveat 1: (should be obvious): both input files must contain the same
#           number of data points
# Caveat 2: the desired quantity should be the second item on the line
#           in the .xvg file (after time)

use strict;

unless (@ARGV) {
    die "Usage: perl sham.pl -i1 input1.xvg -i2 input2.xvg -data1 1 -data2 1
        -o graph.xvg\n";
}

my %args = @ARGV;

# define input and output files
my $input1;
my $input2;
my $output;

if (exists($args{"-i1"})) {
    $input1 = $args{"-i1"};
} else {
    die "No -i1 specified!\n";
}

if (exists($args{"-i2"})) {
    $input2 = $args{"-i2"};
} else {
    die "No -i2 specified!\n";
}

if (exists($args{"-o"})) {
    $output = $args{"-o"};
} else {
    $output = "graph.xvg"; # default output name
}
```

```
}

# define input data sets
my $d1;
my $d2;

if (exists($args{"-data1"})) {
    $d1 = $args{"-data1"};
} else {
    $d1 = 1;    # default to first non-time data column
}

if (exists($args{"-data2"})) {
    $d2 = $args{"-data2"};
} else {
    $d2 = 1;
}

open(IN1, "<$input1") || die "$!: Cannot open $input1";
my @in1 = <IN1>;
close(IN1);

open(IN2, "<$input2") || die "$!: Cannot open $input2";
my @in2 = <IN2>;
close(IN2);

# need to chop off the headers first, since they may cause files with the
# same number of data points to have different lengths (one file has
# more header info than another)

my $size1 = scalar(@in1);

for (my $i=0; $i<$size1; $i++) {
    if ($in1[$i] =~ /^[#@]/) {
        splice(@in1, $i, 1);
        $i--;
        $size1--;
    }
}

my $size2 = scalar(@in2);

for (my $j=0; $j<$size2; $j++) {
    if ($in2[$j] =~ /^[#@]/) {
        splice(@in2, $j, 1);
    }
}
```



```
        $j--;
        $size2--;
    }
}

open(OUT, ">>$output") || die "$!: Cannot open $output\n";
print OUT "# graph.xvg for g_sham\n";

for (my $i=0; $i<scalar(@in1); $i++) {
    my $line1 = $in1[$i];
    my $line2 = $in2[$i];

    if (($line1 =~ /^#[#@]/) || ($line2 =~ /^#[#@]/)) {
        # do nothing, it's a comment - probably not necessary after above cleanup
    } else {
        my @data1 = split(" ", $line1);
        my @data2 = split(" ", $line2);

        my $time = $data1[0];
        my $val1 = $data1[$d1];
        my $val2 = $data2[$d2];

        printf OUT "%8.3f\t%f\t%f\n", $time, $val1, $val2;
    }
}

close(OUT);

exit;
```

E.7 Create index groups of per-leaflet reference atoms

```
#!/usr/bin/perl

use strict;

# Program to determine "top" and "bottom" lipids of a lipid bilayer
# Z-values are specified by a chosen atom on the command line
# If z(reference) > z(middle) -> top, and vice versa

# Collect stuff from command line arguments into a hash
unless (@ARGV) {
    print "Example usage: perl bilayer_separator.pl -in bilayer.gro -ref O11
        -middle C50 -v [yes/no]\n";
}

my %args = @ARGV;
my $input;
my $ref_atom;
my $middle_atom;
my $nwater;
my $verbose;

if (exists($args{"-in"})) {
    $input = $args{"-in"};
} else {
    die "No input specified!\n";
}

# collect all the input lines into an array
open(IN, $input);
my @in = <IN>;
close(IN);

# grab some stuff from the input (non-atom lines)
my $header = shift(@in);
my $orig_natoms = shift(@in);
chomp($orig_natoms);
my $box = pop(@in);

if (exists($args{"-ref"})) {
    $ref_atom = $args{"-ref"};
} else {
    die "No reference atom specified!\n";
}
```

```
}

if (exists($args{"-middle"})) {
    $middle_atom = $args{"-middle"};
} else {
    die "No middle atom specified!\n";
}

# define verbosity
if (exists($args{"-v"})) {
    $verbose = 1;
} else {
    $verbose = 0;    # default to non-verbose mode
}

# process input file to strip out all lines corresponding to
# reference or middle atoms
my @ref_atoms;
my @middle_atoms;

foreach $_ (@in) {
    my @data = split(" ", $_);
    my $atom = $data[1];

    if ($atom =~ /$ref_atom/) {
        push(@ref_atoms, $_);
    } elsif ($atom =~ /$middle_atom/) {
        push(@middle_atoms, $_);
    }
}

# determine the middle of the bilayer from the middle atoms
# middle of bilayer is defined as the average z-coordinate of the middle atoms
my $total_z;
my $nmiddle = scalar(@middle_atoms);

foreach $_ (@middle_atoms) {
    my @data = split(" ", $_);
    my $z = $data[5];

    $total_z += $z;
}

my $middle_z = $total_z / $nmiddle;
```

```
print "Defining the middle of the bilayer as z = $middle_z.\n";

# now split the reference atoms into top and bottom
my @top_ref;
my @bottom_ref;

foreach $_ (@ref_atoms) {
    my @data = split(" ", $_);
    my $z = $data[5];

    if ($z > $middle_z) {
        push(@top_ref, $_);
    } elsif ($z < $middle_z) {
        push(@bottom_ref, $_);
    } else {
        print "Weird z-coordinate found!\n";
        print "$_\n";
    }
}

open(INDEX_OUT, ">>${ref_atom}_leaflets.ndx");

# Now, parse out the meaningful information - top
print INDEX_OUT "[ Top_${ref_atom} ]\n";

my $natom = 0;

foreach $_ (@top_ref) {
    if ($_ =~ /^ +(\d+)(\S+) +(\S+) +(\d+) +([0-9.-]+) +([0-9.-]+)
        +([0-9.-]+)/ ) {
        $natom++;
        # my $res = $1;
        my $atom = $4;
        # print TOP_OUT "$res\n";
        printf INDEX_OUT "%6d", $atom;
        if ($natom % 10 == 0) {
            print INDEX_OUT "\n";
        }
    }
}

# extract bottom information
print INDEX_OUT "\n[ Bottom_${ref_atom} ]\n";

$natom = 0;
```

```
foreach $_ (@bottom_ref) {
    if ($_ =~ /^ +(\d+)(\S+) +(\S+) +(\d+) +([0-9.-]+) +([0-9.-]+)
        +([0-9.-]+)/ ) {
        $natom++;
        # my $res = $1;
        my $atom = $4;
        # print BOT_OUT "$res\n";
        printf INDEX_OUT "%6d", $atom;
        if ($natom % 10 == 0) {
            print INDEX_OUT "\n";
        }
    }
}
close(INDEX_OUT);

exit;
```

E.8 Measure effective chain length and chain tilt angle

```
#!/usr/bin/perl -w

use strict;
use Math::Trig;

# Script reads in a .gro file and calculates effective chain length and tilt angle
# Input on command line is coordinate file, last residue number, and some
# information that is used to name the output files very explicitly
#
# The script is written for DPPC specifically, but can easily be modified

unless (scalar(@ARGV) == 4) {
    print "usage: perl $0 <input> <max res. number> <simulation ID> <time in ns>\n";
    exit;
}

my $input = $ARGV[0];

open (IN, $input) or die "No input file specified.\n";
my @input_file = (<IN>);
close (IN);

# define a max residue number
# script will stop parsing lines at this residue
my $max_res = $ARGV[1];

# parse out everything that doesn't match "DPPC"
my @input_short = grep /DPPC/, @input_file;

# declare new array to store even shorter input
my @input_parsed;
my $line;

do {
    # remove pertinent line from short input
    # when done, short input will contain the other leaflet
    $line = shift(@input_short);

    # place line into parsed input
    push(@input_parsed, $line);
} until ($line =~ /\s+${max_res}DPPC\s+C50/);
```



```
# Parse out the relevant lines now from @input_parsed
# Need to do a few things
# 1. Delete "DPPC"
# 2. grep out only C17 and C31 (sn-2 chain), C36 and C50 (sn-1 chain) atoms
# 3. Splice out relevant info - res. #, (x, y, z)
#

# Delete "DPPC"
foreach (@input_parsed) {
    $_ =~ s/DPPC//g;
}

foreach (@input_short) {
    $_ =~ s/DPPC//g;
}

# Parse out relevant atoms, store in separate arrays
my @c17_c31_top;
my @c36_c50_top;

foreach (@input_parsed) {

    if ($_ =~ /C17/ || $_ =~ /C31/) {
        push(@c17_c31_top, $_);
    } elsif ($_ =~ /C36/ || $_ =~ /C50/) {
        push(@c36_c50_top, $_);
    }

}

my @c17_c31_bottom;
my @c36_c50_bottom;

foreach (@input_short) {

    if ($_ =~ /C17/ || $_ =~ /C31/) {
        push (@c17_c31_bottom, $_);
    } elsif ($_ =~ /C36/ || $_ =~ /C50/) {
        push(@c36_c50_bottom, $_);
    }

}

# Splice out the stuff we want to keep
```

```
# Line format: <res #> <atom ID> <atom #> <x> <y> <z> <vel_x> <vel_y> <vel_z>
# Velocities not always present, all terms separate by whitespace
# Read each line as an array split by whitespace
# If length is such that it includes velocities, pop them off
# Splice out index 2 (atom #)
# Push each line into new array

# C17, C31 array
my @c17_c31_top_fixed;

foreach (@c17_c31_top) {

    my @array_mod = split(" ", $_);
    my $array = scalar(@array_mod);

    # get rid of velocities, if present
    if ($array == 9) {
        pop(@array_mod);
        pop(@array_mod);
        pop(@array_mod);
    }

    # splice out atom number
    splice(@array_mod, 2, 1);

    # re-create a string
    my $line_fixed = join(" ", @array_mod)."\n";

    # re-assemble an array
    push(@c17_c31_top_fixed, $line_fixed);

}

my @c17_c31_bottom_fixed;

foreach (@c17_c31_bottom) {

    my @array_mod = split(" ", $_);
    my $array = scalar(@array_mod);

    # get rid of velocities, if present
    if ($array == 9) {
        pop(@array_mod);
        pop(@array_mod);
        pop(@array_mod);
    }
}
```

```
    }

    # splice out atom number
    splice(@array_mod, 2, 1);

    # re-create a string
    my $line_fixed = join(" ", @array_mod)."\n";

    # re-assemble an array
    push(@c17_c31_bottom_fixed, $line_fixed);
}

# C36, C50 array
my @c36_c50_top_fixed;

foreach (@c36_c50_top) {

    my @array_mod = split(" ", $_);
    my $array = scalar(@array_mod);

    # get rid of velocities, if present
    if ($array == 9) {
        pop(@array_mod);
        pop(@array_mod);
        pop(@array_mod);
    }

    # splice out atom number
    splice(@array_mod, 2, 1);

    # re-create a string
    my $line_fixed = join(" ", @array_mod)."\n";

    # re-assemble an array
    push(@c36_c50_top_fixed, $line_fixed);
}

my @c36_c50_bottom_fixed;

foreach (@c36_c50_bottom) {

    my @array_mod = split(" ", $_);
    my $array = scalar(@array_mod);
```

```
# get rid of velocities, if present
if ($array == 9) {
    pop(@array_mod);
    pop(@array_mod);
    pop(@array_mod);
}

# splice out atom number
splice(@array_mod, 2, 1);

# re-create a string
my $line_fixed = join(" ", @array_mod)."\n";

# re-assemble an array
push(@c36_c50_bottom_fixed, $line_fixed);
}

# Join the lines corresponding to C17 & C31 of the same lipid,
# with only the (x, y, z) coordinates and residue number to produce
# a single line, i.e.: "41 x y z x y z"
# Pass this new array to subroutine, which can split the terms of
# each line to run calculations

my @assembled_array_c17_c31_top;
my $i=0;
my $c17_c31_top_length = scalar(@c17_c31_top_fixed);

while ($i<$c17_c31_top_length) {

    chomp($c17_c31_top_fixed[$i]);
    my $assembled_line = $c17_c31_top_fixed[$i].$c17_c31_top_fixed[$i+1];

    # split line into array to splice out unnecessary positions:
    # atom ID's (indices 1 & 4)
    my @split = split(" ", $assembled_line);
    splice(@split, 1, 1);
    splice(@split, 4, 1);

    # re-assemble the line
    my $join = join(" ", @split)."\n";

    push(@assembled_array_c17_c31_top, $join);
}
```

```
    $i+=2;
}

my @assembled_array_c17_c31_bottom;
$i=0;
my $c17_c31_bottom_length = scalar(@c17_c31_bottom_fixed);

while ($i<$c17_c31_bottom_length) {

    chomp($c17_c31_bottom_fixed[$i]);
    my $assembled_line = $c17_c31_bottom_fixed[$i].$c17_c31_bottom_fixed[$i+1];

    # split line into array to splice out unnecessary positions:
    # atom ID's (indices 1 & 4)
    my @split = split(" ", $assembled_line);
    splice(@split, 1, 1);
    splice(@split, 4, 1);

    # re-assemble the line
    my $join = join(" ", @split)."\n";

    push(@assembled_array_c17_c31_bottom, $join);

    $i+=2;
}

my @assembled_array_c36_c50_top;
$i=0;
my $c36_c50_top_length = scalar(@c36_c50_top_fixed);

while ($i<$c36_c50_top_length) {

    chomp($c36_c50_top_fixed[$i]);
    my $assembled_line = $c36_c50_top_fixed[$i].$c36_c50_top_fixed[$i+1];

    # split line into array to splice out unnecessary positions:
    # atom ID's (indices 1 & 4)
    my @split = split(" ", $assembled_line);
    splice(@split, 1, 1);
    splice(@split, 4, 1);

    # re-assemble the line
    my $join = join(" ", @split)."\n";
```

```
        push(@assembled_array_c36_c50_top, $join);

        $i+=2;
    }

my @assembled_array_c36_c50_bottom;
$i=0;
my $c36_c50_bottom_length = scalar(@c36_c50_bottom_fixed);

while ($i<$c36_c50_bottom_length) {

    chomp($c36_c50_bottom_fixed[$i]);
    my $assembled_line = $c36_c50_bottom_fixed[$i].$c36_c50_bottom_fixed[$i+1];

    # split line into array to splice out unnecessary positions:
    # atom ID's (indices 1 & 4)
    my @split = split(" ", $assembled_line);
    splice(@split, 1, 1);
    splice(@split, 4, 1);

    # re-assemble the line
    my $join = join(" ", @split)."\n";

    push(@assembled_array_c36_c50_bottom, $join);

    $i+=2;
}

# Now, pass the assembled arrays to the angle-calculating subroutine

my @calculated_angles_c17_c31_top = &angle(@assembled_array_c17_c31_top);

open (ANGLE_SN2_TOP, ">>angles_sn-2_${ARGV[2]}_top_${ARGV[3]}.dat");
print ANGLE_SN2_TOP @calculated_angles_c17_c31_top;
close ANGLE_SN2_TOP;

my @calculated_angles_c17_c31_bottom = &angle(@assembled_array_c17_c31_bottom);

open (ANGLE_SN2_BOTTOM, ">>angles_sn-2_${ARGV[2]}_bottom_${ARGV[3]}.dat");
print ANGLE_SN2_BOTTOM @calculated_angles_c17_c31_bottom;
close ANGLE_SN2_BOTTOM;
```



```
my @calculated_angles_c36_c50_top = &angle(@assembled_array_c36_c50_top);

open (ANGLE_SN1_TOP, ">>angles_sn-1_${ARGV[2]}_top_${ARGV[3]}.dat");
print ANGLE_SN1_TOP @calculated_angles_c36_c50_top;
close ANGLE_SN1_TOP;

my @calculated_angles_c36_c50_bottom = &angle(@assembled_array_c36_c50_bottom);

open (ANGLE_SN1_BOTTOM, ">>angles_sn-1_${ARGV[2]}_bottom_${ARGV[3]}.dat");
print ANGLE_SN1_BOTTOM @calculated_angles_c36_c50_bottom;
close ANGLE_SN1_BOTTOM;

# Now, the effective length calculations

my @calculated_lengths_c17_c31_top = &length(@assembled_array_c17_c31_top);

open (LENGTH_SN2_TOP, ">>lengths_sn-2_${ARGV[2]}_top_${ARGV[3]}.dat");
print LENGTH_SN2_TOP @calculated_lengths_c17_c31_top;
close LENGTH_SN2_TOP;

my @calculated_lengths_c17_c31_bottom = &length(@assembled_array_c17_c31_bottom);

open (LENGTH_SN2_BOTTOM, ">>lengths_sn-2_${ARGV[2]}_bottom_${ARGV[3]}.dat");
print LENGTH_SN2_BOTTOM @calculated_lengths_c17_c31_bottom;
close LENGTH_SN2_BOTTOM;

my @calculated_lengths_c36_c50_top = &length(@assembled_array_c36_c50_top);

open (LENGTH_SN1_TOP, ">>lengths_sn-1_${ARGV[2]}_top_${ARGV[3]}.dat");
print LENGTH_SN1_TOP @calculated_lengths_c36_c50_top;
close LENGTH_SN1_TOP;

my @calculated_lengths_c36_c50_bottom = &length(@assembled_array_c36_c50_bottom);

open (LENGTH_SN1_BOTTOM, ">>lengths_sn-1_${ARGV[2]}_bottom_${ARGV[3]}.dat");
print LENGTH_SN1_BOTTOM @calculated_lengths_c36_c50_bottom;
close LENGTH_SN1_BOTTOM;

#####
# Subroutines #
#####

sub angle {
```

```
# Need to feed in coordinates from array, passed into subroutine

my $angle;
my $result;
my @result_array;

foreach (@_) {
    my @values = split(" ", $_);

    my $residue = $values[0];

    # Define the coordinates
    my $x_1 = $values[1];
    my $y_1 = $values[2];
    my $z_1 = $values[3];

    # print "X1: $x_1, Y1: $y_1, Z1: $z_1\n";

    my $x_2 = $values[4];
    my $y_2 = $values[5];
    my $z_2 = $values[6];

    # print "X2: $x_2, Y2: $y_2, Z2: $z_2\n";

    # Calculate vector length
    my $dist = sqrt( ($x_2 - $x_1)**2 + ($y_2 - $y_1)**2 + ($z_2 - $z_1)**2 );

    # Establish the reference, normal to the x-z plane
    my $ref_x = $x_1;
    my $ref_y = $y_2;
    my $ref_z = $z_2;

    # Calculate reference vector length
    my $ref_dist = sqrt ( ($ref_x - $x_1)**2 + ($ref_y - $y_1)**2 +
        ($ref_z - $z_2)**2 );

    # print "The vector length is $dist\n";
    # print "The reference length is $ref_dist\n";

    # Calculate the angle and convert to degrees
    my $rad_angle = acos(($ref_dist/$dist));
    my $angle = $rad_angle * 57.2957795;

    $result = $residue." ".$angle."\n";
}
```

```
        # push result to an array
        push(@result_array, $result);
    }

return @result_array;

}

sub length {

# Measure effective length here! Simply the difference between y values...
# Same atoms (C17, C31) and (C36, C50)

my $length;
my $result;
my @result_array;

foreach (@_) {
    my @values = split(" ", $_);

    my $residue = $values[0];

    # Define the coordinates
    my $y_1 = $values[2];
    my $y_2 = $values[5];

    $length = abs($y_2 - $y_1);

    $result = $residue." ".$length."\n";

    push(@result_array, $result);
}

return @result_array;

}

exit;
```

E.9 Calculate the orientation of a lipid headgroup

```
#!/usr/bin/perl -w

use strict;
use Math::Trig;

# Script calculates the vector between the P and N atoms of lipid headgroups
# The process is general enough that a user can specify any atoms he/she
# wishes, but the common application would be for headgroup atoms
#
# The script is written specifically for DPPC, but can easily be made more general

unless (scalar(@ARGV) == 7) {
    print "usage: perl $0 <input> <atom1> <atom2> <lipid name> <max res. #>
    <simulation ID> <time in ns>\n";
    exit;
}

my $input = $ARGV[0];

open (IN, $input) or die "No input file specified.\n";
my @input_file = (<IN>);
close (IN);

# define the lipid atoms to be analyzed
my $atom_1 = $ARGV[1];
my $atom_2 = $ARGV[2];

# define the name of the lipid to be analyzed
my $lipid_name = $ARGV[3];

# define the break point for dividing the leaflets
# (largest residue number in top leaflet)
my $max_res = $ARGV[4];

# parse out everything that doesn't match the desired lipid name
my @input_short = grep /${lipid_name}/, @input_file;

# declare new array to store even shorter input
my @input_parsed;
my $line;
do {
    # remove pertinent line from short input
```

```
# when done, short input will contain the other leaflet
$line = shift(@input_short);

# place line into parsed input
push(@input_parsed, $line);

} until ($line =~ /^\\s+${max_res}DPPC\\s+C50/);

#
# Parse out the relevant lines now from @input_parsed
# Need to do a few things
# 1. Delete the lipid name
# 2. grep out only the atoms we care about
# 3. Splice out relevant info - res. #, (x, y, z)
#

# Delete lipid name
foreach (@input_parsed) {
    $_ =~ s/${lipid_name}//g;
}

foreach (@input_short) {
    $_ =~ s/${lipid_name}//g;
}

# Parse out relevant atoms, store in separate arrays
my @top;

foreach (@input_parsed) {

    if ($_ =~ /${atom_1}/ || $_ =~ /${atom_2}/) {
        push(@top, $_);
    }

}

my @bottom;

foreach (@input_short) {

    if ($_ =~ /${atom_1}/ || $_ =~ /${atom_2}/) {
        push (@bottom, $_);
    }

}

}
```

```
# Splice out the stuff we want to keep
# Line format: <res #> <atom ID> <atom #> <x> <y> <z> <vel_x> <vel_y> <vel_z>
# Velocities not always present, all terms separate by whitespace
# Read each line as an array split by whitespace
# If length is such that it includes velocities, pop them off
# Splice out index 2 (atom #)
# Push each line into new array

# Top leaflet array
my @top_fixed;

foreach (@top) {

    my @array_mod = split(" ", $_);
    my $array = scalar(@array_mod);

    # get rid of velocities, if present
    if ($array == 9) {
        pop(@array_mod);
        pop(@array_mod);
        pop(@array_mod);
    }

    # splice out atom number
    splice(@array_mod, 2, 1);

    # re-create a string
    my $line_fixed = join(" ", @array_mod)."\n";

    # re-assemble an array
    push(@top_fixed, $line_fixed);

}

my @bottom_fixed;

foreach (@bottom) {

    my @array_mod = split(" ", $_);
    my $array = scalar(@array_mod);

    # get rid of velocities, if present
    if ($array == 9) {
        pop(@array_mod);
```

```
        pop(@array_mod);
        pop(@array_mod);
    }

    # splice out atom number
    splice(@array_mod, 2, 1);

    # re-create a string
    my $line_fixed = join(" ", @array_mod)."\n";

    # re-assemble an array
    push(@bottom_fixed, $line_fixed);
}

# Join the lines corresponding to $atom_1 & $atom_2 of the same lipid,
# with only the (x, y, z) coordinates and residue number,
# i.e.: "41 x y z x y z" would be a line
# Pass this new array to subroutine, which can split the terms of each
# line to run calculations

my @assembled_array_top;
my $i=0;
my $top_length = scalar(@top_fixed);

while ($i<$top_length) {

    chomp($top_fixed[$i]);
    my $assembled_line = $top_fixed[$i].$top_fixed[$i+1];

    # split line into array to splice out unnecessary positions:
    # atom ID's (indices 1 & 4)
    my @split = split(" ", $assembled_line);
    splice(@split, 1, 1);
    splice(@split, 4, 1);

    # re-assemble the line
    my $join = join(" ", @split)."\n";

    push(@assembled_array_top, $join);

    $i+=2;
}
}
```



```
my @assembled_array_bottom;
$i=0;
my $bottom_length = scalar(@bottom_fixed);

while ($i<$bottom_length) {

    chomp($bottom_fixed[$i]);
    my $assembled_line = $bottom_fixed[$i].$bottom_fixed[$i+1];

    # split line into array to splice out unnecessary positions:
    # atom ID's (indices 1 & 4)
    my @split = split(" ", $assembled_line);
    splice(@split, 1, 1);
    splice(@split, 4, 1);

    # re-assemble the line
    my $join = join(" ", @split)."\n";

    push(@assembled_array_bottom, $join);

    $i+=2;

}

# Now, pass the assembled arrays to the angle-calculating subroutine

my @calculated_angles_top = &angle(@assembled_array_top);

open (ANGLE_TOP, ">>angles_${ARGV[5]}_top_${ARGV[6]}.xvg");
print ANGLE_TOP "\# Residue, Angle in degrees\n";
print ANGLE_TOP @calculated_angles_top;
close ANGLE_TOP;

my @calculated_angles_bottom = &angle(@assembled_array_bottom);

open (ANGLE_BOTTOM, ">>angles_${ARGV[5]}_bottom_${ARGV[6]}.xvg");
print ANGLE_BOTTOM "\# Residue, Angle in degrees\n";
print ANGLE_BOTTOM @calculated_angles_bottom;
close ANGLE_BOTTOM;

#####
# Subroutine #
#####
```

```
sub angle {

# Need to feed in coordinates from array, passed into subroutine

my $angle;
my $result;
my @result_array;

foreach (@_) {
    my @values = split(" ", $_);

    my $residue = $values[0];

    # Define the coordinates
    my $x_1 = $values[1];
    my $y_1 = $values[2];
    my $z_1 = $values[3];

    # print "X1: $x_1, Y1: $y_1, Z1: $z_1\n";

    my $x_2 = $values[4];
    my $y_2 = $values[5];
    my $z_2 = $values[6];

    # print "X2: $x_2, Y2: $y_2, Z2: $z_2\n";

    # Calculate vector length
    my $dist = sqrt( ($x_2 - $x_1)**2 + ($y_2 - $y_1)**2 + ($z_2 - $z_1)**2);

    # Establish the reference, normal to the x-z plane
    my $ref_x = $x_1;
    my $ref_y = $y_2;
    my $ref_z = $z_2;

    # Calculate reference vector length
    my $ref_dist = sqrt ( ($ref_x - $x_1)**2 + ($ref_y - $y_1)**2 +
        ($ref_z - $z_2)**2 );

    # print "The vector length is $dist\n";
    # print "The reference length is $ref_dist\n";

    # Calculate the angle and convert to degrees
    my $rad_angle = acos(($ref_dist/$dist));
    my $angle = $rad_angle * 57.2957795;
}
```

```
$result = $residue." ".$angle."\n";

# push result to an array
push(@result_array, $result);

}

return @result_array;

}
```