A COMPUTATIONAL STUDY OF INTERACTIONS BETWEEN AMYLOID-β AND TRANSITION METALS VIA MOLECULAR DYNAMICS

Oliver Douglas Kennedy-Britten

A thesis submitted to

Cardiff University

in accordance with the requirements for the degree of

Doctor of Philosophy

School of Chemistry

Cardiff University

July 2022

For Doug and Dunny

Acknowledgments

I would like to take this opportunity to thank my supervisor, Dr Jamie Platts, for allowing me this opportunity, the funding and his constant advice and support throughout the course of my research. Your guidance and feedback have been fundamental in allowing me to reach this point and I extend my utmost thanks for allowing me to achieve something I never imagined I could do. To Nadiya Al-Shammari and Loizos Savva, I thank you for supporting me as both peers and friends. You have helped me through times it felt difficult to see the wood for the trees and I appreciate the kind words and for always pushing me when I felt I'd reached my limit. I class you as good friends as well as colleagues who have made this journey so enjoyable. A huge thank you to Dr Shaun Mutter and Dr Matthew Turner who helped me when I first arrived at Cardiff University whose knowledge and skills allowed me to get started on this endeavour.

I thank and acknowledge all the fantastic minds who I have had the pleasure of working with at both 55 Park Place and room 2.65B as well as the wider research groups within CCI including the doctors and professors whose insight was always appreciated. It would be remiss of me to not also thanks the fantastic teachers and lecturers I've had who helped me nurture my love for chemistry, with particular thanks to Dr Natasha Galea and Mrs Tina Murray. I also thank Cardiff University for allowing me to pursue this chance to express my ideas in the form of my research and my thesis.

My mother, Alex, and my aunt, Kath, who always listened to my ideas and have only ever wanted the best for me. You have shaped my whole path to where I am today, and I cannot thank you enough. To my partner, Josh, who has only ever known me whilst I've been a student, thank you. It has been tough, there have been times we both felt the strain, but you have supported and encouraged me every step of the way. My closest friends: Sophie, Hannah, Megan, Eden and Emma, you have never allowed me to give up and I thank you for all your support in not only this, but everything I do. You give me courage and "strenath" every single day.

Publications

Sections of this thesis have been published in:

1- Kennedy-Britten O D, Al-Shammari N and Platts J A, Molecular dynamics simulations of copper binding to N-terminus mutants of amyloid-β, *J Biomol Stuct Dyn*, 2020, **39** (6), 2003–13.

2- Kennedy-Britten O D, Alshammari N, Platts and J A, Accelerated Molecular Dynamics to Explore the Binding of Transition Metals to Amyloid-β, *ACS Chem Neurosci*, 2021, **12** (21), 4065-75

Abstract

The aggregational properties of Aβ are crucial in understanding the causes of the neurodegenerative Alzheimer's disease. Coordination of transition metal centres to these peptides have been shown to have differing effects on the mechanism and rate of aggregation of Aβ into characteristic neurotoxic deposits. Within this work, the interaction of Aβ alloforms and various metal ions are investigated computationally via use of molecular dynamics. Initially, genetic mutations of truncated N-terminus Aβ peptides were bound to Cu(II) to replicate effects of metal coordination on the full-length structure compared to wild-type unaltered Aβ. This study showed effects of these variants were marked and varied affecting secondary structure, stability and conformations adapted. Some mutants showed more consistent compact conformations whereas some formed more flexible structures. Contrasts between comparable mutations at similar sites, such as A2T/A2V and D7H/D7N, show the location as well as the type of mutation have effects on protein structure. Notable changes in peptide structure at residues remote to the site of substitution showed these mutations influence the entirety of Aβ. Effects on secondary structure differ between mutations, most notably a change in incidence of β-strand, which has been linked to enhanced aggregational properties for the peptide. Next, accelerated molecular dynamics (aMD) simulations of four different lengths of Aβ and their complexes when bound to Cu(II), Fe(II), or Zn(II) were reported. The presence of a metal ion leads to reduced size and decreased mobility relative to the free peptide due to the anchoring effect of the ions. The reduced mobility was shown largely to be due to the restricted movement in N-terminal residues, most notably Asp1 and His6 that are involved in the metal-ion coordination in all cases. Similarities were noted between results for Zn(II) and Fe(II), whereas results for Cu(II) are more comparable to that of the free peptides. Finally, dimers of full-length $A\beta_{42}$ were simulated via aMD, with free structures compared to those connected via a Zn(II) bridge. The zinc-bound structures adopted more compact configurations shown via Rg, SASA and cluster data compared to the free Aβ dimers. Differences in secondary structures were observed with free dimers forming higher frequencies of helical structures, compared to zinc-bound dimers. The metal-ion bridge between monomers allowed greater amounts of intermolecular interactions than those seen in the free Aβ, meaning inferences can be made on its propensity for enhanced aggregation.

Contents

1 Introduction 14

1.1 Overview..... 14

1.2 Dementia and Alzheimer's disease 14

1.3 The role of amyloid-β in Alzheimer's disease onset..... 15

1.4 Amyloid cascade hypothesis and aggregation 16

1.5 The role of metal ions in AD 18

1.6 Dimerisation of Aβ 20

1.7 Aβ mutations 22

1.7.1 A2T/A2V 22

1.7.2 H6R 23

1.7.3 D7H/D7N 24

1.7.4 E11K 24

1.7.5 K16N..... 25

1.8 Computational modelling of Aβ and biomolecular systems..... 25

1.8.1 Quantum mechanics (QM) methods 26

1.8.2 Hybrid QM/molecular mechanics (MM) methods..... 26

1.8.3 Molecular dynamics (MD) methods 28

1.9 Computational modelling of metal-Aβ systems..... 33

1.9.1 QM methods..... 33

1.9.2 Hybrid QM/MM methods..... 34

1.9.3 MM methods..... 36

1.9.4 Research aims..... 37

1.10 References..... 39

2 Computational theory..... 49

2.1 Overview..... 49

2.2 Born-Oppenheimer approximation..... 49

2.3 Molecular mechanics (MM)..... 50

2.3.1 General concepts..... 50

2.3.2 Forcefield energy terms..... 51

2.3.3 Solvation of a system 55
2.3.4 Advantages and disadvantages of MM 57
2.3.5 Modelling transition-metal complexes using MM 58
2.3.6 Metal-centre parameter builder (MCPB) 59
2.4 Molecular dynamics (MD) 60
2.4.1 General concepts 60
2.4.2 Evolution of time-series and conformational ensembles 61
2.4.3 Advantages and disadvantages of MD 63
2.4.4 Accelerated molecular dynamics (aMD) 63
2.5 References 66
3 Molecular dynamics simulations of copper binding to N-terminus mutants of amyloid-β 69
3.1 Introduction 69
3.2 Computational methods 70
3.3 Results 72
3.4 Discussion 95
3.5 Conclusions 98
3.6 References 100
4 Accelerated Molecular Dynamics to Explore the Binding of Transition Metals to Amyloid-β 104
4.1 Introduction 104
4.2 Computational methods 105
4.3 Results and discussion 108
4.4 Conclusion 142
4.5 References 144
5 Effects of Zn(II) binding on full-length $A\beta_{42}$ dimers using accelerated molecular dynamics (aMD) 148
5.1 Introduction 148
5.2 Computational methods 150
5.3 Results 154
5.4 Discussion 167

5.5 Conclusion 168
5.6 Further work 169
5.7 References 170
6 Conclusions and further work 174
A Appendix..... 178

List of figures

1 Introduction

1.1 Amino acid sequence of $A\beta_{42}$ highlighting N-terminus residues..... 15

1.2 A representative diagram of the cleavage-dependent processing pathways of APP.. 17

1.3 Coordination modes of Cu(II)-Aβ complexes 19

1.4 Coordination modes of Zn(II)-Aβ complexes 19

1.5 Proposed binding mode of an Fe(II)-Aβ complex 20

1.6 Bridging coordination mode of dimeric Zn-Aβ complexes 21

1.7 The amino acid sequence for the first 16 residues of the unaltered Aβ protein 22

1.8 Illustration of $A\beta_{42}$ fibrils formed between Aβ monomers in β-strand formation..... 33

2 Computational theory

2.1 Representation of force field energy terms 51

2.2 Simplified Newman projection of potential dihedral angles (Ω) 53

2.3 TIP3P and TIP4P solvation models of water 56

2.4 Ligand field splitting of octahedral orbitals and associated electronic energies 58

2.5 aMD application of boost potential to energy wells 64

3 Molecular dynamics simulations of copper binding to N-terminus mutants of amyloid-β

3.1 Amino acid sequence for N-terminal of $A\beta_{wt}$, highlighting 7 known mutations 69

3.2 Simulated coordination mode of Cu(II) to Aβ 71

3.3 RMSD plots for each simulation across 500 ns (Å)..... 75-76

3.4 R_g plots for each simulation across 500 ns (Å)..... 77-78

3.5 RMSF plot per residue (Å) 80

3.6 Highest populated clusters for each simulation 82

3.7 Comparison of C_α on peptide backbones of cluster structures showing highest levels of similarity from RMSD data 83

3.8 Salt bridge plots by percentage for equilibrated simulations 84-85

3.9 View of H6R, with sidechains Glu3, Arg5 and Arg6, along with metal binding site... 86

3.10 Contact map of average distance between C_α (Å) 88

3.11 Full Ramachandran plots for equilibrated trajectories including angles and levels of incidence 89-90

3.12 Secondary structure percentages by residue 92-93

3.13 Binding energies with standard deviations for each mutant 95

4 Accelerated Molecular Dynamics to Explore the Binding of Transition Metals to Aβ

4.1 Binding modes of Aβ with Cu(II), Zn(II) and Fe(II) 105

4.2 RMSD (Å) plots for 3 x 250 ns simulations (50 ns cMD, 200 ns aMD) of all 16 Aβ structures analysed 109-110

4.3 R_g (Å) plots for 3 x 200 ns aMD simulations of all 16 Aβ structures analysed 111-112

4.4 RMSF data (Å) based on individual residues 114

4.5 Salt bridge plots between charged residues by percentage incidence 117-120

4.6 Secondary structure incidence plots as a percentage of trajectory frames by amino acid residue 123-126

4.7 Hydrogen bond incidence plots of equilibrated aMD trajectories 128-129

4.8 R_g against PMF as a function of free energy 132-135

4.9 Number of hydrogen bonds against PMF as a function of free energy 136-139

4.10 2D plots of end-to-end distance against R_g as a function of free energy..... 140-141

5 Effects of Zn(II) binding on Aβ dimers using accelerated molecular dynamics (aMD)

5.1 Dimeric starting structures of Aβ..... 149

5.2 Binding mode of Zn(II) ions showing cross-linkage between monomers..... 150

5.3 Individual RMSD plots of Run 2 for free Aβ 153

5.4 RMSD (Å) relative to starting minimised structures for combined MD and aMD trajectories for all four Aβ dimer structures 154

5.5 Graphical representation of R_g data (Å) for aMD trajectories on free-Aβ and zinc-bound dimers..... 155

5.6 Salt-bridge incidence plots by residue for free-Aβ and Aβ-Zn 157

5.7 Secondary structure plots for free-Aβ and Aβ-Zn dimers 159

5.8 RMSF (Å) of four individual aMD trajectories comparing runs within free and zinc dimers 160

5.9 RMSF (Å) of four combined aMD trajectories comparing combined free and zinc trajectories 161

5.10 Hydrogen bond plots for free and zinc-dimers 162

5.11 Representative structures of highest populated clusters 163

5.12 Solvent-accessible surface area ($Å^2$) of Aβ trajectories 164

5.13 Contact maps (Å) for free and zinc dimers 166

List of tables

3 Molecular dynamics simulations of copper binding to N-terminus mutants of amyloid-β

3.1 RMSD equilibration times..... 72

3.2 Statistical analysis of RMSD data (Å)..... 74

3.3 Post equilibration R_g data (Å)..... 74

3.4 RMSF values per residue (Å)..... 81

3.5 Cluster analysis on equilibrated trajectories..... 82

3.6 Percentage of residues classified as helical, strand, or other..... 94

4 Accelerated Molecular Dynamics to Explore the Binding of Transition Metals to Aβ

4.1 aMD boost parameter equations and values used for simulations..... 107

4.2 Statistical analysis of radius of gyration (R_g) data (Å)..... 113

4.3 Cluster analysis data for equilibrated trajectories..... 115

4.4 Percentage of secondary structure types by residue classified as helical, strand or other 121

4.5 Statistical data for hydrogen bond of equilibrated aMD trajectories 127

5 Effects of Zn(II) binding on Aβ dimers using accelerated molecular dynamics (aMD)

5.1 aMD boost potential values derived from initial MD simulations 152

5.2 RMSD (Å) statistics relative to initial minimised structure for combined aMD trajectories 155

5.3 Statistical data for R_g (Å) of dimeric Aβ aMD simulations 155

5.4 Secondary structure of equilibrated aMD trajectories by percentage 158

5.5 Statistical data for hydrogen bond plots 162

5.6 Percentage of highest populated clusters 163

5.7 Statistical analysis of solvent-accessible surface area ($Å^2$) 165

List of appendices

A1 Appendices for Chapter 4

A1.1 Hydrogen bond output for incidence > 5% for Aβ16-Free aMD trajectory data..... 178

A1.2 Hydrogen bond output for incidence > 5% for Aβ16-Cu aMD trajectory data..... 179

A1.3 Hydrogen bond output for incidence > 5% for Aβ16-Zn aMD trajectory data..... 179

A1.4 Hydrogen bond output for incidence > 5% for Aβ16-Fe aMD trajectory data..... 180

A1.5 Hydrogen bond output for incidence > 5% for Aβ28-Free aMD trajectory data..... 180

A1.6 Hydrogen bond output for incidence > 5% for Aβ28-Cu aMD trajectory data..... 181

A1.7 Hydrogen bond output for incidence > 5% for Aβ28-Zn aMD trajectory data..... 182

A1.8 Hydrogen bond output for incidence > 5% for Aβ28-Fe aMD trajectory data..... 183

A1.9 Hydrogen bond output for incidence > 5% for Aβ40-Free aMD trajectory data..... 184

A1.10 Hydrogen bond output for incidence > 5% for Aβ40-Cu aMD trajectory data..... 185

A1.11 Hydrogen bond output for incidence > 5% for Aβ40-Zn aMD trajectory data..... 186

A1.12 Hydrogen bond output for incidence > 5% for Aβ40-Fe aMD trajectory data..... 187

A1.13 Hydrogen bond output for incidence > 5% for Aβ42-Free aMD trajectory data..... 188

A1.14 Hydrogen bond output for incidence > 5% for Aβ42-Cu aMD trajectory data..... 190

A1.15 Hydrogen bond output for incidence > 5% for Aβ42-Zn aMD trajectory data..... 191

A1.16 Hydrogen bond output for incidence > 5% for Aβ42-Fe aMD trajectory data..... 192

A2 Appendices for Chapter 5

A2.1 Hydrogen bond output for incidence > 5% for dimeric Aβ42-Free aMD trajectory data..... 193

A2.2 Hydrogen bond output for incidence > 5% for dimeric Aβ42-Zn aMD trajectory data..... 196

1 Introduction

1.1 Overview

The aim of this research was to use computational methods to model the interactions of selected transition metal ions with various alloforms, genetic variants and dimers of amyloid-β (Aβ) peptides. The main area of focus was analysis and comparison of how coordination to a metal centre can affect dynamics and structures to make inferences on the ability of Aβ to aggregate into potentially harmful aggregates and neurotoxic species. Copper(II), iron(II) and zinc(II) were selected as the ions of interest to investigate here.

The role of Aβ in the onset of Alzheimer's disease (AD) is extensively documented in literature; there are fewer publications available also involving metal ion binding studies conducted via computational methods. Specifically, our studies used both conventional and accelerated molecular dynamics (MD), complemented by density functional theory (DFT) and semi-empirical methods where appropriate, to highlight notable differences in structure attributed to the presence and location of amino acid mutations, the type of metal centre present and the length of the peptide.

1.2 Dementia and Alzheimer's disease

Dementia encompasses several neurological conditions in which typically patients' memory and personality are gradually affected by neuronal cell death[1,2]. Dementia cases are typically seen in the elderly and with a global increase seen in average life expectancy, growing case numbers of these diseases are becoming more of a financial and emotional strain on patients, families, and the healthcare system[3,4].

Globally, there are approximately 50 million people diagnosed with dementia[5], with estimated cases at over 1 million in the UK alone in 2021[6]. The most common form of dementia is Alzheimer's disease (AD), an ailment that affects over two-thirds of dementia patients[7-9]. AD is a progressive neurodegenerative disorder causing a variety of symptoms in those afflicted, such as behavioural changes, a decline in motor or cognitive skills, and eventual death[10]. It was first characterised by Alois Alzheimer in a report from 1907[11,12], with further work published in 1910[13]; these studies involved patients displaying symptoms

we commonly associate today with dementia. Masters *et al.*[14] were the first to note the significant presence of a protein called amyloid-β with neuronal plaques within AD patients, which has led to subsequent analysis and discoveries regarding the aetiology of this disease.

Whilst age appears to be the main risk factor of AD, due to lifetime accumulation of Aβ deposits dependent on formation of cytotoxic forms of the peptide[15], many other factors can be attributed to aetiology of AD. Other risk factors shown to contribute towards AD include obesity[16,17], high blood pressure[18,19], smoking[18,20], cholesterol[21,22], and notable concentrations of metal ions found in the brain such as aluminium[23], lead[24] and mercury[25]. Additionally, rare instances of AD can be accounted for by genetic mutations in APP/Aβ[26].

1.3 The role of amyloid-β in Alzheimer's disease onset

Characteristic Aβ deposits appear to originate within the hippocampus of patients afflicted with AD[27] before spreading through the rest of the cerebral regions of the brain over time[28-30]. The presence of these plaques and neurofibrillary tangles causes fatal damage to neuronal cells. Communication via neurotransmitters is diminished leading to a gradual shrinkage of brain tissue, presenting itself as the degradation in cognition and fine motor skills observed in patients diagnosed with AD[31,32].

Aβ itself is around 39-43 residues in length and is generated as a result of cleavage of the amyloid precursor protein (APP), the dominant forms being either 40 or 42 amino acids long[33,34]. The full amino acid sequence of $A\beta_{42}$ is displayed in Figure 1.1.

Figure 1.1- Amino acid sequence of $A\beta_{42}$ highlighting N-terminus residues

Aβ can be defined by distinct regions of importance. The first 16-residues of the peptide are called the N-terminus, where this hydrophilic region is the site of coordination for metal ions such as iron, copper, and zinc (amongst others)[35-37]. The N-terminal residues will be discussed at greater length later in this Chapter and as a focus of study within Chapters 3

and 4. The central region of residues 21-30 have been shown extensively in studies to be linked with increased fibril formation of Aβ, in comparison to interactions between N-termini of monomers leading to formation of oligomers instead[38-40]. Key amino acid interactions, such as those between Asp23 and Lys28, have been shown to be a driver for fibrillogenesis[41,42]. The main full-length alloforms of Aβ tend to be either 40 or 42 residues long, and despite a difference of only 2 residues at the C-terminus, $A\beta_{40}$ appear to exist at higher concentrations but $A\beta_{42}$ has a higher propensity to aggregate with other Aβ peptides[7,43,44].

1.4 Amyloid cascade hypothesis and aggregation

The amyloid cascade hypothesis suggests disruption in the homeostatic rate of production and clearance of Aβ gives rise to the build-up of these characteristic protein deposits[45]. The accumulation of these peptides over time was originally hypothesised to be the main catalyst for the acceleration of AD development[46,47]. Aβ itself is a naturally occurring peptide, cleaved from the amyloid precursor protein (APP)[48,49] and found in all humans and other animals. So why isn't AD something that happens to us all as a standard part of aging? Notable concentrations of Aβ have been identified posthumously in individuals asymptomatic of any traits associated with AD or dementia. The hypothesis on the role of Aβ as a potential contributor to the causation of AD has developed and evolved, studies indicate the form Aβ takes is crucial[50,51]. There is no exact correlation between Aβ levels and incidence of AD; soluble oligomers have proven to be more toxic and harmful and causative to AD onset alongside other factors[52,53].

Monomeric Aβ forms soluble oligomers and potentially aggregate further into larger fibrillar structures. The rate at which these naturally disorganised oligomers form more highly ordered fibrillar structures appears to affect accumulation of Aβ deposits[54], where intermediary "protofibrils" interacting with smaller Aβ oligomers can branch and stimulate further fibril growth[55,56]. These smaller oligomeric species appear to have enhanced neurotoxicity in comparison to fibrillar Aβ structures, suggesting it is the form adopted by the protein, rather than the concentration of Aβ, that is the key component of understanding the mechanism of AD onset[57]. Full-length Aβ is generated via cleavage at the β and γ sites of the amyloid precursor protein (APP)[58,59]. Cleavage at the alpha (α) and γ-

sites yields the p3-peptide which is not prone to aggregation or oligomerisation (Figure 1.2)[60,61].

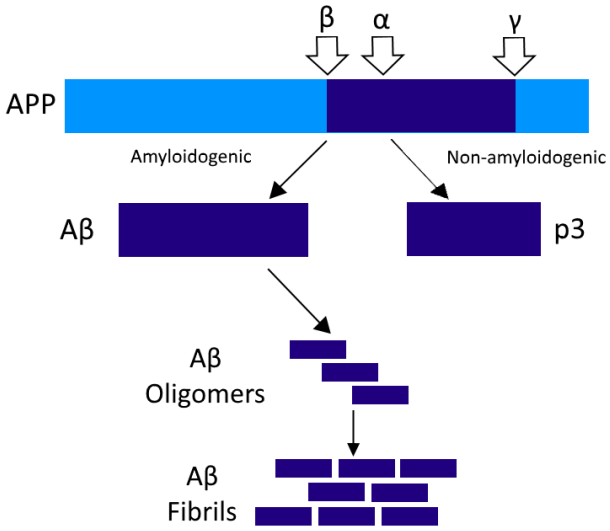


Figure 1.2- A representative diagram of the cleavage-dependent processing pathways of APP

The dominant forms of Aβ are either 40 or 42 amino acids in length ($Aβ_{40}$ and $Aβ_{42}$ respectively); the former existing at higher concentrations *in vivo* whilst the latter form possesses a greater propensity for aggregation[7] where misfolding leads to eventual formation of the plaques indicative of AD diagnosis[8,30,62]. Several studies have used truncated alloforms, such as solely using the N-terminus of Aβ, to successfully model full-length Aβ peptides and make inferences on aggregational properties of the complete monomers[63,64].

Secondary structure of Aβ has been linked to aggregational properties. Free-Aβ adopts intrinsically disorganised coil structures[65,66]. Accumulation of Aβ in this form creates amorphous oligomeric species. Alternatively, presence of increased levels of Aβ forming β-strand secondary structures appears to accelerate the aggregation process by organising themselves into fibril structures comprised of parallel β-sheets, which can potentially misfold into plaques characteristic of AD[67-69].

1.5 The role of metal ions in AD

One of the criticisms of the amyloid cascade hypothesis is the lack of a clear correlation between production of Aβ and onset of AD. It is hypothesised instead that the presence of increased concentrations of metal ions within the brain has been shown to be a potential contributing factor in enhanced formation and aggregation of aforementioned neurotoxic forms of Aβ[26,70]. Metal-Aβ complexes have been shown to prevent transition of oligomers into relatively more organised fibrils and instead accumulate into amorphous, soluble aggregates from these oligomeric species[71]. Additionally, oxidative stress appears to be an early warning sign of AD, exacerbated by dyshomeostasis of transition metals and subsequent coordination to Aβ leading to an increased production of reactive oxygen species (ROS).[72-75]

Within the cores of senile Aβ plaques, patients afflicted with AD were also found to possess notable concentrations of transition metal ions such as iron, copper, and zinc[76-79]. Experimental evidence shows these metal ions bind readily to the peptide between the β-cleavage and α-cleavage sites of APP, or within the N-terminus when there is a dysregulation of either metals or Aβ[37,80,81].

Cu(II) binds to this region via a slightly distorted, square-planar complex due to its d^9 electronic configuration and Jahn-Teller distortions[36]. pH-dependent binding modes of Cu-Aβ complexes can vary but typically involve coordination via three N-donor atoms and at least one oxygen atom (Figure 1.3)[35,82]. At physiological and lower pH, donor atom contributions are observed from the carbonyl oxygen and backbone nitrogen of Asp1 alongside two of either the δ or ε nitrogen atoms of two histidine residues from His6 and either position His13 or His14 (i.e component I)[83,84]. Deprotonation of a nitrogen within residue Ala2 at higher pH levels give rise to an alternative binding mode (i.e component II)[85]. The deprotonated backbone nitrogen and carbonyl oxygen of Ala2, (alongside binding to one of the histidine residues at positions 6, 13 or 14), can then act as a ligand in place of contributions from the carbonyl of Asp1. Cu(II) ions and Aβ form metalloprotein complexes in a 1:1 stoichiometric ratio and additional evidence has also suggested the presence of an additional axial interaction from oxygen in water or the terminal carboxylate of Asp1[86].

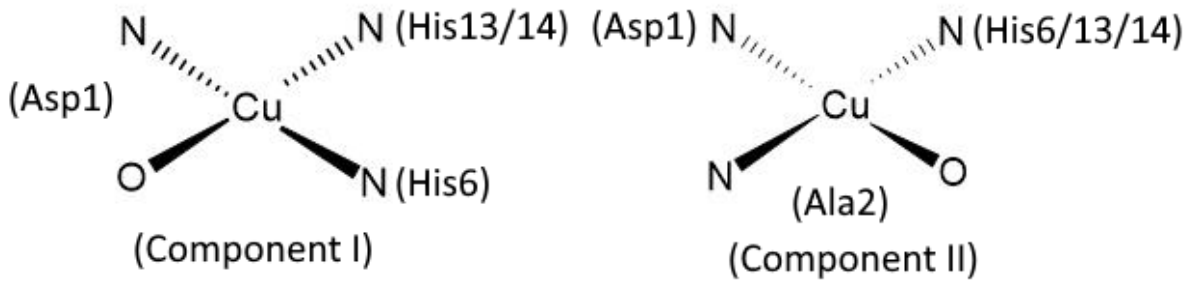


Figure 1.3- Coordination modes of Cu(II)-Aβ complexes

Zn(II) ions can't be studied via Electron paramagnetic resonance (EPR) due to their diamagnetic nature from their full d^{10} orbitals, instead NMR experimentation was used to deduce potential binding modes[24]. Zn(II) forms Aβ-complexes in a 1:1 ratio[87] and various coordination modes have been proposed (Figure 1.4). One suggested model is a tetrahedral geometry involving three histidine residues (His6, 13 and 14) and either Asp1 or Glu11 (component I)[88-90]. An alternative mode of binding was elucidated to involve only two of the three histidine residues as well as oxygen donors from both Asp1 and Glu11 (component II)[80,91].

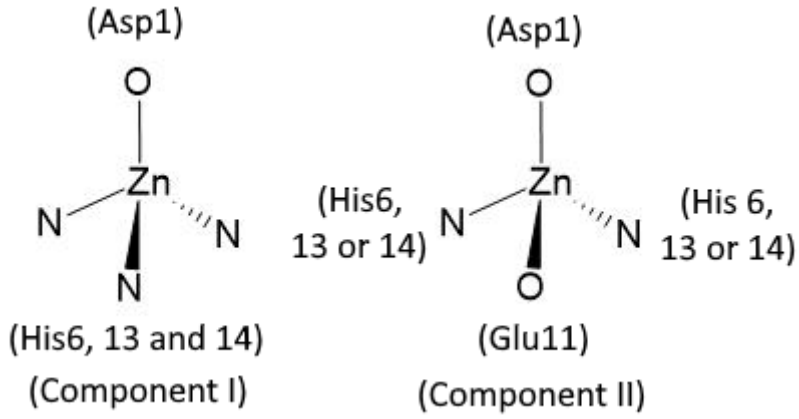


Figure 1.4- Coordination modes of Zn(II)-Aβ complexes

The role of iron in the aetiology of AD is still slightly ambiguous due to its effects being more closely associated with apparent secondary effects of iron on generation or Aβ from APP, an enhanced production of reactive oxygen species and being causative to death of neuronal

receptors via a different mechanism to that seen by Aβ[80,92,93]. Studies on naturally-occurring Fe(II)-Aβ have provided limited information compared to those containing copper and zinc, yet have been able to show the importance of these metalloproteins in contributing to aggregation[79]. Aβ can act as a pentadentate ligand with Fe(II), contributing potential donor atoms from an oxygen and nitrogen from Asp1 or Glu3 and either His6, 13 or 14 alongside a further nitrogen from one of the remaining histidines (Figure 1.5)[36,80].

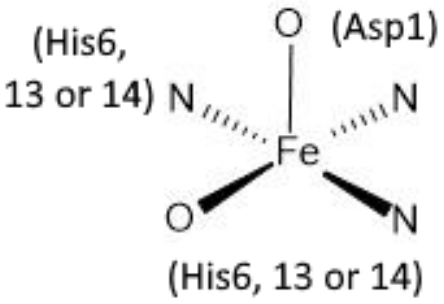


Figure 1.5- Proposed binding mode of an Fe(II)-Aβ complex

As per the Irving-Williams series[94] defining binding affinity of transition metal complexes, Cu(II) binds more strongly to Aβ than Zn(II), whilst Fe(II) shows weaker binding than both of these ions potentially due to its capability of readily oxidising to Fe(III)[36,95]. Zinc-Aβ monomeric complexes appear to accelerate aggregation of oligomers at a more rapid rate and at lower neuronal concentrations than Cu(II)[96]. Whilst metalloprotein complexes of Aβ and Fe(II) have been observed and characterised[97,98], the low binding affinity of iron may suggest a lower prevalence of these structures. This suggests that effects of *in vivo* iron-Aβ structures may be supplementary to the greater production of ROS and alteration of other neurologically important proteins such as ferritin caused by Fe(II)[99]. Transition metals can also accelerate aggregation via formation of cross-links between monomeric units of Aβ[100].

1.6 Dimerisation of Aβ

Whilst the discussion Aβ toxicity usually extends to fibrils and oligomeric species, enhanced synaptotoxicity has also been observed within even dimeric structures of Aβ[101,102]. Experimental data for various dimeric forms of Aβ that have been identified are limited due to their propensity to aggregate further into oligomers[103,104]. Mature fibrils appear to form from Aβ dimers in parallel β-strand structures, suggesting formation of dimeric alloforms

could potentially be an explanation of their capability as an accelerant to aggregation[66,105,106]. Stable dimers of Aβ appear to form readily in equilibrium with neuronal oligomers and fibrils[107,108] with evidence of some β-sheet formation experimentally and in-vivo[109-111].

Coordination between metal ions and Aβ monomers in a 1:2 ratio respectively gives rise to cross-links between peptides to form bridged dimers. Metal bridging in dimers appears to reduce formation of β-sheet secondary structures[112] and, whilst suggesting a decrease in subsequent formation of more organised fibrils, could indeed propagate possible formation of more amorphous oligomeric aggregates[113]. The primary coordination site of zinc for metal-dependent dimerisation of Aβ has been deduced as Glu11 and His13 or 14 from each of the monomer contributors (Figure 1.6)[114].

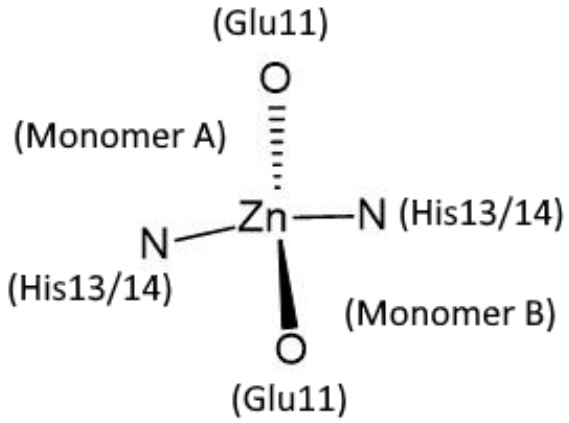


Figure 1.6- Bridging coordination mode of dimeric Zn-Aβ complexes

The central region of Aβ, between residues 21-30, has been directly linked to accelerated aggregation and changes to secondary structure of these amino acids upon dimerisation could be a potentially causative factor to further oligomerisation[115]; other studies suggest an alternative mechanism of aggregation, where extended structures form due to increased inter-monomeric contacts leading to β-sheet formation and subsequent fibrilisation[116]. As such, it can be deduced that due to the multiple pathways of gradual accumulation stemming from formation of dimers, these species appear to play a crucial role in the aggregation process and onset of AD.

1.7 Aβ mutations

As discussed above, cases of dementia are commonly inferred to be a standard occurrence as part of the ageing process. Yet symptoms and diagnosis have also been observed in younger individuals (< 65 years old[117]) due to genetic alterations in the amino acid sequence of Aβ[100]. These instances of AD are referred to as early-onset familial Alzheimer's disease (EOFAD), and account for 5% of AD diagnoses[7,118]. Within the N-terminus of human Aβ there are 7 known mutations (Figure 1.7)[119].

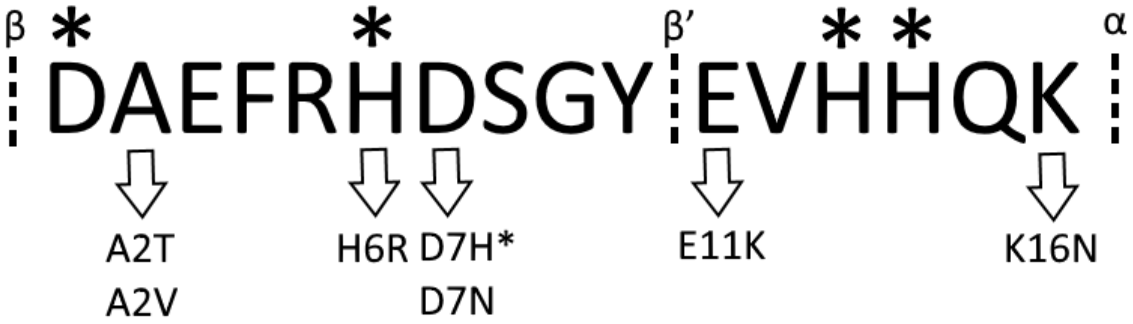


Figure 1.7- The amino acid sequence for the first 16 residues of the unaltered Aβ protein ($Aβ_{wt}$), highlighting known mutations, cleavage sites and residues involved in coordination of Cu(II) (highlighted with an asterisk)

1.7.1 A2T/A2V

The alanine residue at position 2 of Aβ (adjacent to the N-terminus) can be substituted by either polar threonine or the hydrophobic valine amino acid[120]. The A2T mutation was found at higher levels than the naturally occurring Aβ in a control group of Icelandic individuals. Those possessing the altered protein were categorised as displaying no symptoms of AD in a control group compared to those diagnosed with EOFAD[121]. The protective nature of this particular mutation occurs due to inhibition of the BACE1 enzyme at the β-site of the protein[122]. Due to the proximity of the mutated residue and the resulting effect on the enzyme responsible for cleavage, experimental data has shown a 20% decrease in production of Aβ from APP[123].

In contrast, the A2V mutation, observed in a study on an Italian population[124], has potential to increase the rate of aggregation in homozygous carriers. It is recessive in nature as

heterozygous carriers possessing both wild type and the altered form of APP displayed no symptoms or increased deposits of Aβ plaques[125]. The presence of this mutation in homozygotic cases was shown to increase aggregation and fibril formation both experimentally and *in vitro*[126].

Comparing the two, it is observed that the alanine to threonine mutation generates no increase in amyloid production or aggregation when compared to the unaltered peptide. This is also true for heterozygous carriers of the A2V mutation, whilst recessive inheritance of this mutation actually accelerates oligomerisation[127]. Both mutants exist at position 2 of Aβ but have different effects. It can be reasoned that the similarities in hydrophobicity of alanine and valine account for the aggregational properties in wild-type Aβ and the A2V substitution, compared to the polar residue present in the A2T form[128].

1.7.2 H6R

The histidine to arginine amino acid substitution that was observed in a population of English patients occurs at position 6 of Aβ. This novel mutation is interesting in the sense that it does not affect the cleavage of APP and there is no increase in production of Aβ. Instead, this particular mutation affects the aggregational properties of Aβ, leading to more rapid formation of toxic oligomers[129]. These oligomers form fibrils at a much faster rate than wild-type Aβ as the intermediate protofibrils formed are short lived intermediates that convert readily[130]. The secondary structure of $A\beta_{40}$ remains relatively unchanged; however, $A\beta_{42}$ experiences a decrease in turns and an increase in coils of secondary structure. Coupled with this, it is seen that dimer/trimerization is observed more readily as a decrease in overall charge of the molecule and an increase in hydrophobic properties drives aggregation[131].

Histidine-6 was identified as a key residue involved in coordination with metal ion centres including Cu(II), Zn(II) and Fe(II). The substitution of that residue would be expected to cause a change in the kinetics of ligand binding. This was not the case and in fact, despite the His6 residue providing stability of the complex, its absence does not actually disrupt the rate of initial coordination of the metal centre[132]. It can be hypothesised that the histidine-13 or -14 residues would act as replacement binding sites intramolecularly, or dimerization could occur with other molecules of Aβ[68,133,134].

1.7.3 D7H/D7N

The aspartic acid residue at position 7 can undergo a change into either a histidine residue (D7H, Taiwanese) or asparagine (D7N, Japanese/Tottori)[135]. As seen with other mutants, an increase in Aβ production is observed due to beta-cleavage in APP. This promotes the amyloidogenic pathway leading to an increase in the formation of fibrils in $A\beta_{40}$ and oligomers in $A\beta_{42}$[6]. In the D7H mutation of both $A\beta_{40}$ and $A\beta_{42}$, there is a notable decrease in the formation of protofibrils compared to the wild type Aβ. This is due to stability of the oligomers and potentially increased neurotoxicity in $A\beta_{42}$ as well as a rapid conversion of short-lived intermediates into fibrous structures in $A\beta_{40}$[136]. Due to the addition of another histidine residue, there is an increased propensity for coordination with Cu(II) ions (and other pertinent metal ions) leading to an increased aggregation propensity[137].

Conversely, the D7N mutation shows no increased production of Aβ and instead affects the formation of secondary structures and subsequent oligomerization[138]. In this case, there is an increased concentration of stable oligomers and, similar to the Taiwanese D7H mutation, the levels of protofibrils are lower than the wild-type Aβ due to accelerated fibril synthesis from rapid conversion of the intermediate species[139]. Changes in secondary structure follow a very similar pattern to those of the English H6R mutant and an increased lifespan of salt bridges within both the N- and C-termini lead to accelerated fibril formation[140].

1.7.4 E11K

This mutation was noted in a Belgian patient and was seen to be the cause of an increase in total Aβ produced. It was observed however that there was a decrease in $A\beta_{40}$ and an increase in $A\beta_{42}$ when compared to the ratio of $A\beta_{40/42}$ of the unaltered Aβ[141]. An abundance of the BACE1 enzyme leads to cleavage at an alternative β'-site located between tyrosine-10 and glutamic acid-11 at the N-terminus of Aβ[142] (as shown previously in Figure 2). The mutation of glutamic acid to lysine at position 11 of the Aβ peptide pushes cleavage of APP to the β-site between methionine-671 and aspartic acid-672 of APP (position 1 of Aβ) even with a high concentration of BACE1. This leads to increased cleavage via the amyloidogenic pathway forming full length molecules of Aβ with no effect on the α-cleavage site near the C-terminus[143].

It cannot be said for definite that this mutation shows an increase in pathogenicity as β'-site cleavage only occurs in an overexpression of BACE1, so β-site cleavage is not unusual. It has been proposed that if cleavage can be shifted to favour the β'-site somehow, inhibiting production of full-length Aβ, this could have potentially beneficial effects on pathogenicity and onset of Alzheimer's disease. At present, it is unclear whether this approach is possible, or how it may be achieved[59].

1.7.5 K16N

So far, the mutations have shown to affect cleavage involving the β-sites, near the N-terminus of Aβ. Unlike the other mutations discussed here, K16N substitutes lysine for asparagine at the α-cleavage site, located between lysine-16 and leucine-17[144]. Cleavage at this site encourages a non-amyloidogenic pathway. K16N mutation causes a reduction of cleavage at this α-site by approximately 50%[145]. This leads to an increase of full length $A\beta_{40/42}$ being generated as it follows the amyloidogenic pathway outlined previously[6].

This mutant strain was not pathogenic on recessive carriers, but an increase in toxicity was found when expressed in those possessing heterozygous alleles for this mutation. The presence of the mutant peptide alongside unaltered Aβ in heterozygous carriers shows the dominant nature of this mutation. Heterozygous carriers produced equimolar concentrations of both standard human Aβ and the mutant Aβ. These formed stable oligomers due to increased hydrogen bonding between K16 of Aβ and N16 of the mutant type[146]. Additionally, the mutant peptide caused disruption to neprilysin, the enzyme which is involved in disposal of Aβ thus leading to increased aggregation[147].

1.8 Computational modelling of Aβ and biomolecular systems

Aβ is a flexible and disorganised peptide which makes structure determination difficult to achieve via conventional experimental methods. Its propensity to transition between conformations or oligomerise makes it difficult to examine experimentally especially at the monomeric level. Computational methods have thus been employed to determine structures and chemistry of this peptide. A selection of work is discussed in this section, highlighting some key discoveries using these methods as well as any limitations.

1.8.1 Quantum mechanics (QM) methods

Quantum mechanics (QM) is a branch of physics that can be applied computationally to predict probability of physical properties of atoms and subatomic particles. QM is a popular method due to its highly accurate results based on *ab initio* and semi-empirical values. Despite this, no data currently exists modelling full-length Aβ utilising solely QM methods due to the sheer size of the system as well as the flexible nature of Aβ. Density functional theory (DFT) is a QM method used extensively in the study of electronic structures of multi-body systems. A study by Morgado *et al*.[148] was able to successfully highlight the accuracy of DFT as well as an augmented method, DFT-D (DFT implementing an empirical dispersion term), when modelling key noncovalent interactions typically found within biomolecular systems. Schubert *et al.*[149] built upon this, highlighting the importance and effects of intermolecular interactions in gas-phase 20-residue peptides (Ac-Ala_{19}-Lys and Ac-Lys-Ala_{19}) using DFT with PBE and PBE0 hybrid functionals, generating data for conformers showing good agreement with mass-spec and infrared experimental data. Work by van Mourik[150] set out to assess a variety of density functionals by modelling minima conformers of a Tyr-Gly dipeptide via DFT. The author noted the difficulty in using computational models for structures inherently affected by electrostatic and dispersion forces in the presence of π-electron clouds. Most of the methods tested here were able to successfully identify the structures expected for this particular dipeptide showing the high-level of accuracy this method can achieve.

1.8.2 Hybrid QM/molecular mechanics (MM) methods

Molecular mechanics (MM) is a classical method that empirically accounts for electronic effects of a system within its parameters by treating atoms as particles using Newtonian descriptors (this will be explored in greater detail in Section 2.2). These empirical models make for more rapid simulations on larger systems whilst also allowing for greater conformational sampling at less computational cost. As both electronic and nuclear interactions are implicitly considered to be included within these particles, this means that more approximations are made compared to QM methods. As such, this leads to potential differences in accuracy levels between these methods. Due to the extensive computational studies and method refinement, bond parameters and approximations are typically transferable between systems containing comparable bonding components. In an effort to

reduce computational expense, hybrid methods using both QM and MM can be implemented for large biomolecular systems by partitioning larger structures into separate models based on size. QM can be used to simulate smaller, targeted sites of interest due to its limitation in the number of atoms it can simulate successfully. MM can then generate a suitable model for the rest of the system and when used in conjunction, these methods can accurately predict structures and energies within a more feasible timescale than using QM alone. The energy of the system can be defined via Equations 1.1 and 1.2 as either additive or subtractive methods respectively. In the additive method, total energy equals the sum of the energies for both the QM and MM models plus the interactions between the QM system and the MM environment whereas the subtractive (or extrapolative) model incorporates the MM energy of the real system minus that of the MM model[151].

$$E_{QM/MM} = E_{QM} + E_{MM} + E_{QM-MM} \quad (1.1)$$

$$E_{QM/MM} = E_{QM,model} + E_{MM,real} - E_{MM,model} \quad (1.2)$$

A review by Senn & Thiel[152] provides greater insight into the application of these hybrid techniques as well as extensive comparison of the combinations of methods that can be used. Selected studies using these hybrid methods have been noted here showing their validity in modelling Aβ and related structures.

Boopathi and Kolandaivel[153] utilised the subtractive ONIOM method[151] to model interactions of the Asp23-Lys28 saltbridge of Aβ, a notable interaction which has been shown to play a crucial role in aggregation of the peptide[105]. This study involved modelling truncated Aβ structures of only these two amino acids and those four linking these residues (Asp23-Val24-Gly26-Ser27-Lys28) as a representation of full-length Aβ dimers. Asp23 and Lys28 were modelled using QM whereas the other four residues utilised MM descriptors. Small, potentially therapeutic molecules (TPT, AQ and morin) were then modelled within this QM region to examine interactions with Aβ in order to assess viability for disruption of this region noted for its role in aggregation. The study highlighted this region's propensity to forming stacked β-strand structures between Aβ monomers and that the drug molecules

interacting with this site were able to disrupt the formation of these conformations linked to oligomerisation. Interactions between truncated Aβ and the antibody bapineuzumab were examined using ONIOM2 and QTAIM and reported by Gutierrez *et al.*[154]. In this study, isoforms of the first six residues of Aβ (Asp1-His6) and water were modelled via QM whereas the remainder of the system was modelled using MM due to bapineuzumab being considerably larger than the Aβ peptide studied. Binding energies obtained showed good agreement with experimental results as well as highlighting coordination modes between Aβ and the antibody. They also offered potential explanations and improvements that could be made to increase the viability of this particular drug as a therapeutic candidate against AD. Guisasola *et* al.[155] carried out multiscale modelling on monomeric Aβ bound to an inhibitor peptide, DZK. MM-GBSA was implemented to highlight key residues involved in DZK-Aβ binding from representative structures of a prior molecular dynamics (MD) simulation of these complexes. These key residues then underwent optimisation via QM/MM on the DZK peptide (MM) bound to these selected amino acids (QM). From this, it was deduced coordination to this inhibitor peptide disfavours formation of neurotoxic oligomers.

1.8.3 Molecular dynamics (MD) methods

Molecular dynamics (MD) is a method used for assessing the movement of atoms over a time-dependent series which commonly employs (but is not limited to) classical MM methods to produce a series of snapshots of structures over a timeseries. These can subsequently be plotted across a potential energy surface (PES) for use in identifying local minima and preferred conformations. As per the Born-Oppenheimer approximation, this method employs empirical descriptors for particles, implicitly accounting for electronic descriptors and fixed nuclear values. Use of classical descriptors allows for rapid simulations even on structures as large as full-length Aβ. The work included within this thesis mainly considers only MM/MD methods and a selection of studies pertinent to our research are reported here.

MM/MD studies can be carried out on considerably larger biomolecular systems such as dimers and oligomers of full-length Aβ as in studies by Urbanc *et al.*[156,157] and Mehrazma & Rauk[158]. Both of these studies are evidence of these method's capabilities of computing

simulation data on systems of a larger scale than QM methods. The former looks at full-length Aβ dimers and the latter study reports 9.5 μs worth of data (both of which on systems of approximately 1200 atoms in size), which would potentially be infeasible to simulate for systems of these sizes using QM due to the comparably greater computational cost compared to MM. Additionally, a system of this size could not be simulated using QM alone due to this method being limited to hundreds of atoms. The experiments mentioned were successful in providing conformational analysis and secondary structure data on β-strand formation of Aβ dimers, inferring potential links to oligomerisation. Whilst it is possible to simulate systems of this size in full, Cecchini *et al.*[159] instead focused on overlapping six-residue truncated Aβ peptides spanning the full peptide length, in order to focus on each region's contribution towards aggregational properties and subsequent oligomer formation. As in previous studies, they deduced a notable contribution of β-strand secondary structure within the central hydrophobic region, but also showed some slight contributions within the N-terminus and markedly less from the C-terminus. Despite this, the results suggested that the C-terminal residues still contributed towards accumulation of Aβ but via a potentially different mechanism.

Conventional MD has proven itself a viable technique for modelling biomolecular systems comprised of several thousand atoms and of a particular flexible nature in many studies. Like every method there are some opportunities for improvement; some simulations observe prolonged periods of time sampling structures that are similar in energy, due to being unable to transition between minima due to high energy barriers on the potential energy surface. Some advanced sampling methods have been developed building upon the concepts and theory of conventional MD in order to provide better sampling of conformational space for flexible systems. Replica-exchange MD (REMD) allows systems to sample structures of similar potential energies but at different temperatures. Typically, a range of temperatures are input during the set-up phase in order to direct the simulation to allow it to overcome potentially high energy barriers which may limit structures not observed via conventional MD. This method works by exchanging temperature values between parallel MD simulations of like structures (replicas) which can then potentially lead onto further conformational change dependent on temperature[160,161].

Baumketner and Shea[162] utilised this method to examine secondary structure of a 25-residue fragment of full-length Aβ in explicit solvent. This study highlighted the disordered state of Aβ peptides forming globular structures that rapidly interchange between one another. The Asp23-Lys28 salt-bridge is again noted for its importance here and its potentially causative manner of leading to more organised secondary structure and subsequent aggregation. Nguyen *et al.*[163] used REMD on monomeric, dimeric and trimeric Aβ fragments (Lys16-Glu22) to compare atomistic forcefields, some of which support Baumketner and Shea's findings. Simulations using OPLS[164] showed a range of structures, some of which supported the disordered nature of Aβ observed in other works but also showed some conformations adopting more organised assemblies. AMBER99[165] simulations were shown to adopt helical structures whilst not displaying any β-strand secondary structure, whereas GROMOS96[166] predicted strong formation of antiparallel β-sheets in dimeric and trimeric Aβ indicative of enhanced aggregation upon formation of links between monomers. Building upon these studies, Ngo *et al.*[167] implemented REMD on trimeric Aβ in explicit solvent, of which, limited information was available prior. High levels of β-sheet secondary structure were observed with an emphasis placed upon the central hydrophobic region. Interestingly, the Asp23-Lys28 salt-bridge that has been so heavily implied to be causative of presence of β-strand was actually replaced here by intermolecular interactions between Asp23 and residues between Val24-Gly29. Even without notable presence of this salt-bridge, these interactions still led to formation of a loop region that allowed for antiparallel β-sheet formation.

Due to the nature of MD, it isn't uncommon for a structure to return to previously sampled structures. Metadynamics (also referred to as bias-exchange metadynamics) attempts to dissuade the system from returning to previously sampled points on the PES by filling the energy surface along collective variables (CVs) such as bond lengths, dihedral angles *etc*. Metadynamics applies a bias to the system in order to encourage the simulation to sample new parts of the CVs in order to explore greater conformational space and satisfy the ergodic hypothesis. Jong *et al.*[168] conducted MD simulations on a zwitterionic C-terminal fragment of Aβ (Ala30-Met35) within water using metadynamics to further explore the free-energy surface of these simulations. Here, metadynamics was successful in identifying minima structures as well as the potential of these C-terminal residues in disrupting

conformational structures that might be linked to amyloid aggregation. However, it was noted that their method was unable to study the dynamical properties of this system as it was not considered a typical time series observed in conventional MD, and instead was only able to provide potential conformations that could be sampled. Zerze *et al.*[169] compared both temperature-based REMD and bias-exchange metadynamics on intrinsically disordered proteins (IDPs). Metadynamics increases in cost exponentially as system size and simulation time increase due to the complexity of the calculations on addition of CVs. Both of these methods proved to increase exploration of conformational space compared to the "brute-force" method of conventional MD and it was also shown the results between these two methods showed good agreement with one another.

Another advanced sampling method that has proved successful for biomolecular systems such as Aβ is accelerated molecular dynamics (aMD). The theory behind this method will be explained in greater detail in Section 2.4.4, but the general concept is an application of a boost in potential energy to the system in order to overcome potentially insurmountable energy barriers. This, once again, allows for easier transitions between minima structures and greater exploration of conformational space. Similar to REMD, this achieves better sampling than conventional MD but does so at a single temperature without requiring exchange of structures or multiple parallel trajectories. Jose et al.[170] modelled full length dimers comprised of Aβ42 and alpha synuclein (αSyn95) in aqueous solution using aMD. These peptides were found to be present in diagnosis of the neurological disorders, AD and Parkinson's disease, respectively. Results showed that, similar to Aβ dimers, the central hydrophobic region played a key role as an interface between these peptides. Over 1 μs worth of simulation data was generated showing aMD's capacity to generate sufficient datasets within a reasonable timeframe for a larger system of this size. The question of ergodicity in biomolecular simulation is long-standing and still open. In principle, a conventional MD simulation will visit all accessible conformations eventually, but in practice finite simulation lengths mean that this may or may not be the case depending on the specifics of the system. One study found that suitable sampling to ensure ergodicity for a small model system, namely the alanine dipeptide, required 0.4 ms of conventional MD[171]. This is several hundred times longer than we are able to access with current computational resources for peptides of the size of interest here. REMD improves sampling by including

configurations that can only be accessed at higher temperatures than the target, but does so at the expense of requiring many (typically 16 or 32) parallel simulations and hence significantly larger computational resource. Enhanced sampling techniques such as metadynamics or aMD allow more configurations to be visited within single trajectories by boosting the potential energy to lower barriers between kinetically isolated structures. However, these methods formally do not follow the ergodic hypothesis, since a modified potential energy surface is explored. It is possible to recover the correct ensemble through reweighting, effectively subtracting the effects of the boost potential. It was shown that through such reweighting with suitable choice of boost potential, proper sampling of alanine dipeptide could be achieved, but with too "aggressive" boost sampling was degraded. One of the first applications of aMD[172] to a biomolecule compared 1 ms of conventional MD with 500 ns of boosted simulation, and showed that "the same conformational space is sampled by both approaches." We therefore proceed with aMD for some of our simulations despite its formal lack of ergodicity, and stress that conventional MD would suffer from similar problems albeit from a different source, since we cannot hope to reach the timescales required for exhaustive simulation of systems of the size and flexibility of interest here. Aβ monomers and dimers were further studied by Zhang et al.[173] using a multiscale approach, first simulating this system via conventional MD followed by subsequent aMD and REMD simulations. Here, dimeric Aβ did not display the level of β-strand character typically expected in fibrillar Aβ, in this instance, the central hydrophobic region and C-terminus appeared to actually stabilise the structure via hydrogen bonding. Supporting this, Huang et al.[174] also studied dimeric Aβ42 using aMD over a range of temperatures. This study generated 500 ns worth of data and actually showed Aβ dimers displayed enhanced levels of α-helical secondary structure and suggests β-strand formation could be temperature dependent.

Further development to these methods are still being created, such as Gaussian aMD (GaMD)[175], which applies the boost potential observed in aMD but as a "Gaussian approximation" of free energy calculations without the need for CVs, whilst generating a smoother potential energy surface via use of a harmonic boost potential. A comprehensive review by Tran and Ha-Duong[176] compared literature on various MD methods on full-length $A\beta_{40}$ and $A\beta_{42}$. These considered classical MD using a range of forcefields as well as

enhanced and simplified sampling methods such as REMD, coarse-grain models and discrete MD (DMD). Results were generally consistent, giving insight into the structure being generally disordered within the N-terminus and a propensity to form β-strand secondary structures within the central hydrophobic core and C-terminus.

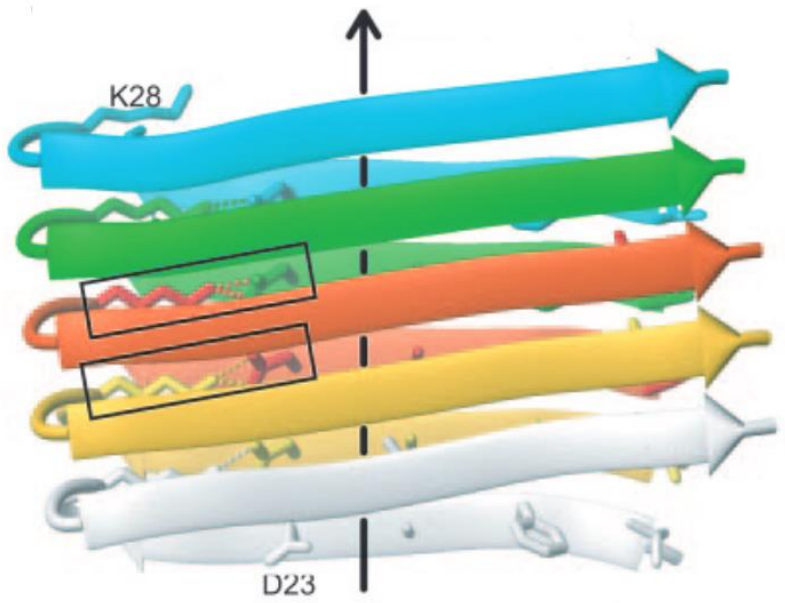


Figure 1.8- Illustration of the $A\beta_{42}$ fibrils formed as a result of intermolecular salt-bridge formation (outlined in the rectangles) between Aβ monomers in β-strand formations[69]

1.9 Computational modelling of metal-Aβ systems

There are still questions remaining on the exact effects of metal coordination to Aβ and some of these are currently highly theoretical. Experimental analyses such as nuclear magnetic resonance (NMR), X-ray absorption spectroscopy (XAS) and electron paramagnetic resonance (EPR) have been unable to provide definitive identification and effects of metal-Aβ structures[177-183]. Computational methods are often employed to simulate potential structures of metal-coordination sites and examine a representative of the dynamics of Aβ. A review by Strodel and Coskuner-Weber[184] extensively describes the limitations and benefits of using computational methods to model metal-Aβ complexes as well as identifying previous studies successfully conducted upon these structures. Addition of metal to these systems adds another layer of complexity such as increased modes of binding, as well as varying electronic effects and oxidation states. As before, a selection of pertinent studies has been chosen and key findings reported below.

1.9.1 QM methods

QM is particularly effective in modelling metal-bound systems due to considering electronic effects explicitly. The limitations of the number of atoms QM can simulate usually restrict this method to only being able to model atoms and groups directly bound to metals or small molecules. DFT is especially pertinent for describing the nature of interactions between Aβ (as well as other biopeptides) with transition metals with varying oxidation states. This method can focus on binding sites for potential metal coordination. DFT is a means of solving electronic structure using the density, ρ, which means important electron correlation effects can be included for relatively low computational cost. DFT has been utilised for many studies of metal-Aβ complexes, the main observations of a select few are reported below. A DFT study conducted by Marino *et al.*[185] proposed multiple coordination modes for Cu(II) and Zn(II) to Aβ involving combinations of Tyr10, Glu11, His6, 13, and 14, as well as interactions between the metal centre and water within the solvent. Both copper and zinc binding displayed propensity to form pentacoordinate geometries whilst Zn(II) also formed four-coordinate structures due to its enhanced flexibility. Reactive oxygen species discussed previously, such as the highly toxic and reactive hydroxyl radical ($OH^{\cdot}$), have also been associated with the onset of AD. Prosdocimi *et al.*[186] used DFT to calculate the energies associated with the reduction of Cu(II) to Cu(I) whilst bound to Aβ, leading to the dissociation of peroxide molecules (H_2O_2). This study is one of many that show how both the form of Aβ and metal association affect AD pathology as opposed to solely being reliant on concentration of the Aβ peptide. Another study looking at Cu(II)-$A\beta_{16}$ conducted by Alí-Torres *et al.*[187] used DFT compared binding modes of copper within the N-terminus. This article identified several low-energy coordination modes comparing both residues and atom types, with the most stable containing Cu(II) bound to Ala2, His6, 13, and 14. Dudev and Lim[188] studied interactions and affinities of Zn(II), Mg(II) and Ca(II) with various nonstandard amino acids and functional groups using DFT methods. Affinity for Zn(II) was notably increased in comparison to the other metals analysed, showing its potential capacity as a preferred choice for metal binding in biomolecular systems.

1.9.2 Hybrid QM/MM methods

Whilst QM is effective in modelling electronic effects on metal-bound biomolecules, the

computational expense and long simulation times make them less preferable for larger bioinorganic structures such as Aβ. Typically, smaller, or truncated models utilise QM whereas the majority of studies on full-length Aβ, alongside other biomolecules, use hybrid QM/MM methods. As discussed previously, QM models can be constructed to only consider any metal centres or residues directly bound to them and utilise MM to model the remainder of the system.

Maiorana *et al.*[189] took structures of Zn(II)-$A\beta_{16}$ determined from their own MD simulations and optimised them using QM/MM. The simulations were conducted in the non-physiological phase, with no solvent, to identify solvent effects on zinc coordination to Aβ, and it was found that the presence of water as solvent does not affect the inner coordination sphere of Zn(II) with Aβ. This was shown by comparing results of computational experiments utilising solvent against the QM/MM minimised values of these systems, which showed good agreement with one another. A novel binding mode of the N-terminus of Aβ was studied by Kulikova *et al.*[190] who focused on zinc-coordination including a phosphorylated version of the residue Ser8 (pSer8). The QM model focused on the tetradentate metal-binding residues; His6, Asp7, and pSer8 whereas the other residues in the N-terminus were modelled via MM. This suggested mode of binding was found to affect the geometries typically adopted surrounding zinc metal centres upon coordination with Aβ, and was even found to induce dimerisation of monomeric species studied via interactions in residues between Glu11 and His14.

Certain semi-empirical approximations applied to DFT are relatively computationally expensive and by nature can only provide a certain level of accuracy compared to experimental values[191]. From this, QM methods (especially DFT) are limited in terms of defining accurate properties of ground-state systems and electrons[192]. When compared to molecular mechanical (MM) calculations or even hybrid QM/MM methods, MM can be seen as relatively inexpensive in comparison to QM on similarly sized metal-bound systems[193]. Truncated models are generally used due to this expense and due to the disorganised nature of these Aβ peptides first-principles methods and approximations can sometimes be unable to give a truly accurate representations of structure and dynamics.

1.9.3 MM methods

Molecular modelling can provide a greater insight into interactions at an atomistic level, identifying low-energy conformers, and allowing inferences to be made regarding aggregational properties on larger or greater numbers of these biomolecules. This is especially notable for metal-peptide interactions, as in MM, atoms are treated as discrete particles and a series of approximations are made during parameterisation of the system. When compared to QM, computational cost is lower for similarly sized systems using MM. Due to negating electronic effects and use of harmonic bond potentials, most MM methods are unable to model bond breaking and formation and thus cannot be used to simulate chemical reactions. As such, MM has been extensively used for metal-Aβ studies on supplied geometries and coordination modes and is typically unable to generate and identify potential binding sites of its own. Molecular dynamics (MD) in particular is beneficial in conformational sampling of flexible systems such as Aβ. MD utilises MM to sample the evolution of structures over a trajectory in time in order to identify accessible conformations at biologically relevant temperature.

Azam *et al.*[194] investigated docking of non-steroidal anti-inflammatory drugs (NSAIDs) optimised with DFT using Becke's three-parameter hybrid model, Lee-Yang-Parr (B3LYP)[195], a common electron-density functional employed with hybrid-DFT. MD was then conducted on minimised Aβ fibril structures docked using these NSAIDs. This study highlighted the affinity of the Aβ fibrils within residues Leu17-Ala21 for docked monomeric Aβ and suggested this to be a potential therapeutic valency site. Boopathi & Kolandaivel[196] simulated Fe(II)-Aβ complexes and showed the exact modes of iron binding to Aβ is still elucidated naturally in part. This experiment shows the validity of using MD as a way of testing proposed binding sites for potential further studies (either experimentally or computationally).

Studies carried out within the Platts research group by Al-Shammari *et al.*[197] modelled interactions of various Zn(II) binding modes within the N-terminus of Aβ. As mentioned previously, MM is unable to generate bonds and must be provided with coordination modes. This experiment is a good example though of MD's capacity to simulate metal-peptide structures to a high level of accuracy even using empirical data and approximations. These simulations in implicit solvent show good agreement with experimentally determined

structures, especially when using the ff14SB forcefield. Additionally, within the same research group, Turner *et al.*[198] report REMD data for Pt(II)-phenanthroline coordinated with both truncated $A\beta_{16}$ and full-length $A\beta_{42}$. Secondary structure values were shown to be comparable with that in literature and the ff14SB forcefield used here (as well as in the previous study mentioned) showed enhanced sampling compared to previous versions and enhanced secondary structure characterisation relative to NMR data for this structure.

1.9.4 Research aims

The role of Aβ in the onset of AD is interesting and as discussed, there are studies demonstrating that the form of the peptide, presence of metal ions, and genetic mutations can affect its aggregational properties and formation of structures potentially harmful to cognitive ability. As stated previously, classical methods have been shown as an effective means to model structures such as Aβ which usually is hundreds of atoms in size (though can be thousands when looking at dimers, trimers, and larger structures) at a reasonable computational expense. The main focus within this body of work is utilising MD and enhanced sampling techniques as a means of exploring metal binding with alloforms of Aβ. This structure and dynamics of these complexes can be studied to make inferences on aggregational properties. Studying mutations of Aβ in conjunction with effects of metal binding can provide greater insight into the noted protective or neurotoxic forms of these genetic variants, whilst also investigating the proposed importance of the presence of enhanced concentrations of transition metal ions observed within the brains of patients affected by AD. The presence of transition metals has been suggested to be partially causative in the dyshomeostasis of excess Aβ peptides and therefore, by simulating these metal-peptide complexes, inferences can be made on how flexibility of these structures are affected and therefore their propensity to aggregate when in excess. Furthermore, studies here in various chain lengths of Aβ allow for direct comparison between binding of different metals as well as in the absence of any metal centre. Finally, dimers of Aβ have been linked extensively to being a potential pathway for further evolution into toxic oligomers. By comparing free dimers and metal-bridged structures, proposals can be made on the effects of these metal-peptide complexes as well as illustrating aggregational properties for even larger Aβ forms. Overall, the overarching aim of the work reported here is to generate

evidence on the role of these notable transition metal ions when bound to Aβ and how this can be potentially linked to future therapeutic work on understanding the mechanism for onset of AD.

1.10 References

1 NIH: The National Institute on Aging, https://www.nia.nih.gov/health/what-is-dementia, (Accessed June 30thth 2022)

2 D. Geldmacher, P. Whitehouse, *N Engl J Med*, 1996, **335**, 330-336

3 G. Livingston, A. Sommerlad, V. Orgeta, S. G. Costafreda, J. Huntley, D. Ames, C. Ballard, S. Banerjee, A. Burns, J. Cohen-Mansfield, C. Cooper, N. Fox, L. N. Gitlin, R. Howard, H. C. Kales, E. B. Larson, K. Ritchie, K. Rockwood, E. L. Sampson, Q. Samus, L. S. Schneider, G. Selbaek, L. Teri, N. Mukadam, J. Cohen-Mansfield and N. Gitlin, *Lancet*, 2017, **390**, 2673–2734.

4 Cleusa P Ferri, Martin Prince, Carol Brayne, Henry Brodaty, Laura Fratiglioni, Mary Ganguli, Kathleen Hall, Kazuo Hasegawa, Hugh Hendrie, Yueqin Huang, Anthony Jorm, Colin Mathers, Paulo R Menezes, Elizabeth Rimmer and Marcia Scazufca, *Lancet*, 2005, **366**, 2112–17.

5 World Health Organization, Dementia, http://www.who.int/news-room/fact-sheets/detail/dementia, (accessed June 7thth 2022).

6 W. T. Chen, C. J. Hong, Y. T. Lin, W. H. Chang, H. T. Huang, J. Y. Liao, Y. J. Chang, Y. F. Hsieh, C. Y. Cheng, H. C. Liu, Y. R. Chen and I. H. Cheng, *PLoS ONE*, 2012, **7**, e35807.

7 C. Ballard, S. Gauthier, A. Corbett, C. Brayne, D. Aarsland and E. Jones, *Lancet*, 2011, **377**, 1019–1031.

8 J. Hardy, *Curr Alzheimer Res*, 2006, **3**, 71–73.

9 What is Alzheimer's | Alzheimer's Association, https://alz.org/alzheimers-dementia/what-is-alzheimers, (accessed June 30th 2022).

10 G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price and E. Stadlan, *Neurology*, 1984, **34**, 939–944.

11 A. Alzheimer, Über eine eigenartige Erkrankung der Hirnrinde. Allgemeine Zeitschrift für Psychiatrie und psychisch-gerichtliche Medizin, https://www.alzforum.org/papers/uber-eine-eigenartige-erkrankung-der-hirnrinde, (accessed April 20th 2022).

12 R. A. Stelzmann, H. N. Schnitzlein and F. R. Murtagh, *Clin Anat*, 1995, **8**, 429–43.

13 B. A. Carabello, W. J. Paulus, P. Blase and A. Carabello, *Lancet*, 2009, **373**, 956–966.

14 C. L. Masters, G. Simms, N. A. Weinman, G. Multhaupt, B. L. Mcdonald and K. Beyreuthert, *Proc Natl Acad* Sci, 1985, **82**, 4245–4249.

15 K. Blennow, M. J. de Leon and H. Zetterberg, *Lancet*, 2006, **368**, 387–403.

16 D. Gustafson, E. Rothenberg, K. Blennow, B. Steen and I. Skoog, *Arch Intern Med*, 2003, **163**, 1524-1528

17 R. Whitmer, E. Gunderson, C. Quesenberry, J. Zhou and K. Yaffe, *Curr Alzheimer Res*, 2007, **4**, 103–109.

18 L. J. Launer, G. W. Ross, H. Petrovitch, K. Masaki, D. Foley, L. R. White and R. J. Havlik, *Neurobiol Aging*, 2000, **21**, 49-55

19 D. E. Barnes and K. Yaffe, *Lancet Neurol*, 2011, **10**, 819-828

20 C. Reitz, T. den Heijer, C. van Duijn, A. Hofman and M. M. B. Breteler, *Neurology*, 2007, **69**, 998–1005.

21 I.-L. Notkola, R. Sulkava, J. Pekkanen, T. Erkinjuntti, C. Ehnholm, P. Kivinen, J. Tuomilehto and A. Nissinen, *Neuroepidemiology*, 1998, **17**, 14-20

22 H. D. Soares and D. L. Sparks, *Alzheimer's Disease: Topics in Medicinal Chemistry, vol2*, Springer, Berlin, Heidelberg, 53–80. 2007

23 L. C. Neri and D. Hewitt, *Lancet*, 1991, **338**, 390.

24 M. Loef, L. F. Mendoza and H. Walach, *Toxin Rev*, 2011, **30**, 103-114

25 J. Mutter, A. Curth, J. Naumann, R. Deth and H. Walach, *J Alzheimers Dis*, 2010, **22**, 357–374.

26 K. P. Kepp, *Chem Rev*, 2012, **112**, 5193-5239

27 Alzheimer's disease facts and figures, https://www.nia.nih.gov/health/alzheimers-disease-fact-sheet, (accessed June 30th 2022).

28 M. Yamada, *Neuropathology*, 2000, **20**, 8–22.

29 J. F. Norfray and J. M. Provenzale, *Am J Roentgenol*, 2004, **182**, 3–13.

30 S. Karantzoulis and J. E. Galvin, *Expert Rev Neurother*, 2011, **11**, 1579–1591.

31 C. Haass and D.J. Selkoe, *Nat Rev Mol Cell Biol*, 2007, **8**, 101-12

32 R. Kandimalla, P. H. Reddy and P. Hemachandra Reddy, *J Alzheimers Dis*, 2017, **57**, 1049–1069.

33 S. G. Younkin, *Ann Neurol*, 1995, **37**, 287–288.

34 D. M. Walsh, D. M. Hartley, Y. Kusumoto, Y. Fezoui, M. M. Condron, A. Lomakin, G. B. Benedek, D. J. Selkoe and D. B. Teplow, *J Biol Chem*, 1999, **274**, 25945–52.

35 C. Hureau and P. Faller, *Biochimie*, 2009, **91**, 1212–1217.

36 K. P. Kepp, *Coordin Chem Rev*, 2017, **351**, 127–159.

37 E. Atrián-Blasco, M. del Barrio, P. Faller and C. Hureau, *Anal Chem*, 2018, **90**, 5909–5915.

38 A. Baumketner, S. L. Bernstein, T. Wyttenbach, N. D. Lazo, D. B. Teplow, M. T. Bowers and J.-E. Shea, *Protein Sci*, 2006, **15**, 1239–1247.

39 M. G. Krone, A. Baumketner, S. L. Bernstein, T. Wyttenbach, N. D. Lazo, D. B. Teplow, M. T. Bowers and J.-E. Shea, *J Mol Biol*, 2008, **381**, 221–228.

40 T. Takeda and D. K. Klimov, *J Phys Chem B*, 2009, **113**, 6692-6702

41 K. L. Sciarretta, D. J. Gordon, A. T. Petkova, R. Tycko and S. C. Meredith, *Biochemistry-US*, 2005, **44**, 6003-6014

42 S. T. Mutter, M. Turner, R. J. Deeth and J. A. Platts, *ACS Chem Neurosci*, 2018, **9**, 2795-2806

43 G. D. Schellenberg, T. D. Bird, E. M. Wijsman, H. T. Orr, L. Anderson, E. Nemens, J. A. White, L. Bonnycastle, J. L. Weber, M. E. Alonso, H. Potter, L. L. Heston and G. M. Martin, *Science*, 1992, **258**, 668-671

44 E. Levy-Lahad, W. Wasco, P. Poorkaj, D. M. Romano, J. Oshima, W. H. Pettingell, C. E. Yu, P. D. Jondro, S. D. Schmidt, K. Wang, A. C. Crowley, Y. Fu, S. Y. Guenette, D. Galas, E. Nemens, E. M. Wijsman, T. D. Bird, G. D. Schellenberg and R. E. Tanzi, *Science*, 1995, **269**, 973–77

45 J. Hardy and D. Allsop, *Trends Pharmacol Sci*, 1991, **12**, 383–388.

46 D. J. Selkoe, *Neuron*, 1991, **6**, 487–498.

47 John A. Hardy and Gerald A. Higgin, *Science*, 1992, **256**, 184–185.

48 Y. H. Hung, A. I. Bush and R. A. Cherny, *J Biol Inorg Chem*, 2010, **15**, 61–76.

49 A. I. Bush and R. E. Tanzi, *Neurotherapeutics*, 2008, **5**, 421-432

50 W. L. Klein, *Neurochem Intl*, 2002, **41**, 345–352.

51 C. B. Eckman and E. A. Eckman, *Neurol Clin*, 2007, **25**, 669–679.

52 R. Ricciarelli and E. Fedele, *Curr Neuropharmacol*, 2017, **15**, 926–935.

53 S. Lovestone, *Dialogues in Clinical Neurosciences*, 2000, **2**, 101-110

54 L. Nagel-Steger, M. C. Owen and B. Strodel, *ChemBioChem*, 2016, **17**, 657–676.

55 I. v Baskakov, *FEBS J*, 2007, **274**, 3756-3765

56 A. Šarić, Y. C. Chebaro, T. P. J. Knowles and D. Frenkel, *Proc Natl Acad Sci USA*, 2014. **111**, 17869-17874

57 J. Cleary, D. M. Walsh, J. J. Hofmeister, G. M. Shankar, M. A. Kuskowski, D. J. Selko and K. H. Ashe, *Nat Neurosci*, **8**, 79–84

58 V. Tõugu, A. Karafin, K. Zovo, R. S. Chung, C. Howells, A. K. West and P. Palumaa, *J Neurochem*, 2009, **110**, 1784–1795

59 J. Hardy and R. Guerreiro, *EMBO Mol Med*, 2011, **3**, 247–248.

60 P. Sorrentino, A. Iuliano, A. Polverino, F. Jacini and G. Sorrentino, *FEBS Lett*, 2014, **588**, 641–652.

61 A. Vandersteen, E. Hubin, R. Sarroukh, G. de Baets, J. Schymkowitz, F. Rousseau, V. Subramaniam, V. Raussens, H. Wenschuh, D. Wildemann and K. Broersen, *FEBS Lett*, 2012, **586**, 4088–4093.

62 Y. Yan and C. Wang, *J Mol Biol*, 2006, **364**, 853-862

63 S. A. Kozin, A. A. Kulikova, A. N. Istrate, P. O. Tsvetkov, S. S. Zhokhov, Y. v Mezentsev, O. I. Kechko, A. S. Ivanov, V. I. Polshakov and A. A. Makarov, *Metallomics*, 2015, **7**, 422.

64 D. Greene, T. Po, J. Pan, T. Tabibian and R. Luo, *J Phys Chem B*, 2018, **122**, 4521–4536

65 S. Vivekanandan, J. R. Brender, S. Y. Lee and A. Ramamoorthy, *Biochem Bioph Res Co*, 2011, **411**, 312–316.

66 A. K. Somavarapu and K. P. Kepp, *ChemPhysChem*, 2015, **16**, 3278–3289.

67 K. Halverson, P. E. Fraser, D. A. Kirschner and P. T. Lansbury, *Biochemistry-US*, 1990, **29**, 2639–2644.

68 A. Schmechel, H. Zentgraf, S. Scheuermann, G. Fritz, R. Pipkorn, J. Reed, K. Beyreuther, T. A. Bayer and G. Multhaup, *J Biol Chem*, 2003, **278**, 35317–35324.

69 T. Lührs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Döbeli, D. Schubert, R. Riek and F. Hoffmann-La Roche, *ADV Behav Biol*, 2005, **102**, 17342-17347

70 A. I. Bush, *Trends Neurosci*, 2003, **26**, 207–214

71 A. S. Pithadia, M. H. Lim, F. Armstrong and L. Que, *Curr Opin Chem Biol*, 2012, **16**, 67–73.

72 K. Honda, G. Casadesus, R. B. Petersen, G. Perry and M. A. Smith, *Ann NY Acad Sci*, 2004, **1012**, 179–182.

73 G. E. Gibson and H. M. Huang, *Neurobiol Aging*, 2005, **26**, 575–578.

74 D. S. Kalinowski, C. Stefani, S. Toyokuni, T. Ganz, G. J. Anderson, N. v. Subramaniam, D. Trinder, J. K. Olynyk, A. Chua, P. J. Jansson, S. Sahni, D. J. R. Lane, A. M. Merlot, Z. Kovacevic, M. L. H. Huang, C. S. Lee and D. R. Richardson, *BBA - Mol Cell Res*, 2016, **1863**, 727–748.

75 C. Cheignon, M. Jones, E. Atrián-Blasco, I. Kieffer, P. Faller, F. Collin and C. Hureau, *Chem Sci*, 2017, **8**, 5107–5118.

76 D. J. Ewing, I. W. Campbell, B. ; Clarke, B. F. Clarke, I. ; Campbell, E. P. Sharpey-Shafer, P. ; Taylor, R. H. Lloyd-Mostyn, P. ; Watkins, M. N. Levy and P. J. Martin, *Lancet*, 1981, **317**, 186–188.

77 M. A. Lovell, J. D. Robertson, W. J. Teesdale, J. L. Campbell and W. R. Markesbery, *J Neurol Sci*, 1998, **158**, 47–52.

78 B. Alies, H. Eury, C. Bijani, L. Rechignat, P. Faller and C. Hureau, *Inorg Chem*, 2011, **50**, 11192–11201.

79 H. Kozlowski, M. Luczkowski, M. Remelli and D. Valensin, *Coordin Chem Rev*, 2012, **256**, 2129–2141.

80 F. Bousejra-Elgarah, C. Bijani, Y. Coppel, P. Faller and C. Hureau, *Inorg Chem*, 2011, **50**, 9024–9030.

81 S. Boopathi, P. Dinh Quoc Huy, W. Gonzalez, P. E. Theodorakis and M. S. Li, *Proteins*, 2020, **88**, 1285–1302.

82 C. D. Syme, R. C. Nadal, S. E. J. Rigby and J. H. Viles, *J Biol Chem*, 2004, **279**, 18169-18177

83 T. Kowalik-Jankowska, M. Ruta, K. Wisniewska and L. Lankiewicz, *J Inorg Biochem*, 2003, **95**, 270–282

84 S. C. Drew, C. J. Noble, C. L. Masters, G. R. Hanson and K. J. Barnham, *J Am Chem Soc*, 2009, **131**, 1195–1207

85 C. Hureau, *Coordin Chem Rev*, 2012, **256**, 2164–2174.

86 V. A. Streltsov, S. J. Titmuss, V. C. Epa, K. J. Barnham, C. L. Masters and J. N. Varghese, *Biophys J*, 2008, **95**, 3447–3456.

87 V. Tõugu, A. Karafin and P. Palumaa, *J Neurochem*, 2008, **104**, 1249-1259

88 Y. Mekmouche, Y. Coppel, K. Hochgräfe, L. Guilloreau, C. Talmard, H. Mazarguil and P. Faller, *ChemBioChem*, 2005, **6**, 1663–1671.

89 S. Zirah, S. A. Kozin, A. K. Mazur, A. Blond, M. Cheminant, I. Ségalas-Milazzo, P. Debey and S. Rebuffat, *J Biol Chem*, 2006, **281**, 2151–2161.

90 E. Gaggelli, A. Janicka-Klos, E. Jankowska, H. Kozlowski, C. Migliorini, E. Molteni, D. Valensin, G. Valensin and E. Wieczerzak, *J Phys Chem B*, 2008, **112**, 100–109.

91 J. Nasica-Labouze, P. H. Nguyen, F. Sterpone, O. Berthoumieu, N. V. Buchete, S. Coté, A. de Simone, A. J. Doig, P. Faller, A. Garcia, A. Laio, M. S. Li, S. Melchionna, N. Mousseau, Y. Mu, A. Paravastu, S. Pasquali, D. J. Rosenman, B. Strodel, B. Tarus, J. H. Viles, T. Zhang, C. Wang and P. Derreumaux, *Chem Rev*, 2015, **115**, 3518–3563

92 J. H. Cheah, S. F. Kim, L. D. Hester, K. W. Clancy, S. E. Patterson, V. Papadopoulos and S. H. Snyder, *Neuron*, 2006, **51**, 431–440.

93 D. Valensin, C. Migliorini, G. Valensin, E. Gaggelli, G. la Penna, H. Kozlowski, C. Gabbiani and L. Messori, *Inorg Chem*, 2011, **50**, 54.

94 H. Irving and R. J. P. Williams, *J Chem Soc*, 1953, 3192-3210

95 A. Clements, D. Allsop, D. M. Walsh and C. H. Williams, *J Neurochem*, 1996, **66**, 740–747.

96 E. Atrián-Blasco, A. Conte-Daban and C. Hureau, *Dalton T*, 2017, **46**, 12750.

97 M. Brüne, P. Palanza, S. Parmigiani and A. Troisi, *J Nerv Ment Dis*, 2022, **210**, 235–245.

98 G. M. Bishop, S. R. Robinson, Q. Liu, G. Perry, C. S. Atwood, M. A. Smith and M. Smith, *Dev Neurosci*, 2002, **24**, 184–187.

99 J. A. Duce and A. I. Bush, *Prog Neurobiol*, 2010, **92**, 1–18.

100 B. Jiao, B. Tang, X. Liu, J. Xu, Y. Wang, L. Zhou, F. Zhang, X. Yan, Y. Zhou and L. Shen, *Neurobiol Aging*, 2014, **35**, 1957.e1-1957.e6.

101 G. M. Shankar, S. Li, T. H. Mehta, A. Garcia-Munoz, N. E. Shepardson, I. Smith, F. M. Brett, M. A. Farrell, M. J. Rowan, C. A. Lemere, C. M. Regan, D. M. Walsh, B. L. Sabatini and D. J. Selkoe, *Nat Med*, 2008, **14**, 837-842

102 S. L. Bernstein, N. F. Dupuis, N. D. Lazo, T. Wyttenbach, M. M. Condron, G. Bitan, D. B. Teplow, J.-E. Shea, B. T. Ruotolo, C. v Robinson and M. T. Bowers, *Nat Chem*, 2009, **1**, 326–331.

103 G. Bitan, M. D. Kirkitadze, A. Lomakin, S. S. Vollers, G. B. Benedek and D. B. Teplow, *Proc Natl Acad Sci USA*, 2003, **100**, 330-335

104 M. D. Kirkitadze, M. M. Condron and D. B. Teplow, *J Mol Biol*, 2001, **312**, 1103–1119.

105 A. T. Petkova, W.-M. Yau and R. Tycko, *Biochemistry-US*, 2006, **45**, 498–512

106 T. Sato, P. Kienlen-Campard, M. Ahmed, W. Liu, H. Li, J. I. Elliott, S. Aimoto, S. N. Constantinescu, J.-N. Octave and S. O. Smith, *Biochemistry-US*, 2006, **45**, 5503-5516

107 B. Soreghan, J. Kosmoski and C. Glabes, *J Biol Chem*, 1994, **269**, 28551–28554

108 W. Garzon-Rodriguez, M. Sepulveda-Becerra, S. Milton and C. G. Glabe, *J Biol Chem*, 1997, **272**, 21037–21044.

109 D. M. Walsh, I. Klyubin, J. v Fadeeva, W. K. Cullen, R. Anwyl, M. S. Wolfe, M. J. Rowan and D. J. Selkoe, *Nature*, 2002, **416**, 535-539

110 W. Hwang, S. Zhang, R. D. Kamm and M. Karplus, *Proc Natl Acad Sci USA*, 2004, **101**, 12916-12921

111 X. Zhu, R. P. Bora, A. Barman, R. Singh and R. Prabhakar, *J Phys Chem B*, 2012, **116**, 4405–4416.

112 P. D. Q. Huy, Q. van Vuong, G. la Penna, P. Faller and M. S. Li, *ACS Chem Neurosci*, 2016, **7**, 1348–1363.

113 P. O. Tsvetkov, I. A. Popov, E. N. Nikolaev, A. I. Archakov, A. A. Makarov and S. A. Kozin, *ChemBioChem*, 2008, **9**, 1564–1567.

114 S. A. Kozin, Y. v Mezentsev, A. A. Kulikova, A. v Golovin, A. S. Ivanov, P. O. Tsvetkov and A. A. Makarov, *Roy Soc Ch*, 2011, **7**, 1053–1055.

115 S. Côté, R. Laghaei, P. Derreumaux and N. Mousseau, *J Phys Chem B*, 2012, **116**, 4043–4055.

116 E. Zerovnik and S. Giannini, *Eur J Biochem*, 2002, **269**, 3362-3371

117 A. F. Teich and O. Arancio, *Biochem J*, 2012, **446**, 165–177.

118 C. Cheignon, M. Tomas, D. Bonnefont-Rousselot, P. Faller, C. Hureau and F. Collin, *Redox Biol*, 2018, **14**, 450–464.

119 APP | ALZFORUM, https://www.alzforum.org/mutations/app, (accessed June 30th 2022).

120 M. L. Peacock, J. T. Warren, A. D. Roses and J. K. Fink, *Neurology*, 1993, **43**, 1254LP – 1254.

121 T. Jonsson, J. K. Atwal, S. Steinberg, J. Snaedal, P. v. Jonsson, S. Bjornsson, H. Stefansson, P. Sulem, D. Gudbjartsson, J. Maloney, K. Hoyte, A. Gustafson, Y. Liu, Y. Lu, T. Bhangale, R. R. Graham, J. Huttenlocher, G. Bjornsdottir, O. A. Andreassen, E. G. Jonsson, A. Palotie, T. W. Behrens, O. T. Magnusson, A. Kong, U. Thorsteinsdottir, R. J. Watts and K. Stefansson, *Nature*, 2012, **488**, 96–99.

122 J. A. Maloney, T. Bainbridge, A. Gustafson, S. Zhang, R. Kyauk, P. Steiner, M. van der Brug, Y. Liu, J. A. Ernst, R. J. Watts and J. K. Atwal, *J Biol Chem*, 2014, **289**, 30990–31000.

123 B. Murray, M. Sorci, J. Rosenthal, J. Lippens, D. Isaacson, P. Das, D. Fabris, S. Li and G. Belfort, *Proteins*, 2016, **84**, 488–500.

124 G. Giaccone, M. Morbin, F. Moda, M. Botta, G. Mazzoleni, A. Uggetti, M. Catania, M. L. Moro, V. Redaelli, A. Spagnoli, R. S. Rossi, M. Salmona, G. di Fede and F. Tagliavini, *Acta Neuropathol*, 2010, **120**, 803–812.

125 M. Messa, L. Colombo, E. del Favero, L. Cantu, T. Stoilova, A. Cagnotto, A. Rossi, M. Morbin, G. de Fede, F. Tagliavini and M. Salmona, *J Biol Chem*, 2014, **289**, 24143–24152.

126 I. Benilova, R. Gallardo, A. A. Ungureanu, V. C. Cano, A. Snellinx, M. Ramakers, C. Bartic, F. Rousseau, J. Schymkowitz and B. de Strooper, *J Biol Chem*, 2014, **289**, 30977–30989.

127 X. Zheng, D. Liu, R. Roychaudhuri, D. B. Teplow and M. T. Bowers, *ACS Chem Neurosci*, 2015, **6**, 1732–1740.

128 G. di Fede, M. Catania, M. Morbin, G. Rossi, S. Suardi, G. Mazzoleni, M. Merlin, A. R. Giovagnoli, S. Prioni, A. Erbetta, C. Falcone, M. Gobbi, L. Colombo, A. Bastone, M. Beeg, C. Manzoni, B. Francescucci, A. Spagnoli, L. Cantu, E. del Favero, E. Levy, M. Salmona and F. Tagliavini, *Science*, 2009, **323**, 1473–1477.

129 K. Ono, M. M. Condron and D. B. Teplow, *J Biol Chem*, 2010, **285**, 23186–23197.

130 Y. Hori, T. Hashimoto, Y. Wakutani, K. Urakami, K. Nakashima, M. M. Condron, S. Tsubuki, T. C. Saido, D. B. Teplow and T. Iwatsubo, *J Biol Chem*, 2007, **282**, 4916–4923.

131 M. H. Viet, P. H. Nguyen, P. Derreumaux and M. S. Li, *ACS Chem Neurosci*, 2014, **5**, 646–657.

132 P. Girvan, T. Miyake, X. Teng, T. Branch and L. Ying, *ChemBioChem*, 2016, **17**, 1732–1737.

133 S. A. Khmeleva, S. P. Radko, S. A. Kozin, Y. Y. Kiseleva, Y. v. Mezentsev, V. A. Mitkevich, L. K. Kurbatov, A. S. Ivanov and A. A. Makarov, *J Alzheimers Dis*, 2016, **54**, 809–819.

134 C. Hureau and P. Dorlet, *Coordin Chem Rev*, 2012, **256**, 2175–2187.

135 P. M. Truong, M. H. Viet, P. H. Nguyen, C. K. Hu and M. S. Li, *J Phys Chem B*, 2014, **118**, 8972–8981.

136 L. Xu, Y. Chen and X. Wang, *Proteins*, 2014, **82**, 3286–3297.

137 V. I. Polshakov, A. B. Mantsyzov, S. A. Kozin, A. A. Adzhubei, S. S. Zhokhov, W. van Beek, A. A. Kulikova, M. I. Indeykina, V. A. Mitkevich and A. A. Makarov, *Angew Chem Int Edit*, 2017, **56**, 11734–11739.

138 Y. Wakutani, K. Watanabe, Y. Adachi, K. Wada-Isoe, K. Urakami, H. Ninomiya, T. C. Saido, T. Hashimoto, T. Iwatsubo and K. Nakashima, *J Neurol Neurosur Ps*, 2004, **75**, 1039–1042.

139 M. M. Gessel, S. Bernstein, M. Kemper, D. B. Teplow and M. T. Bowers, *ACS Chem Neurosci*, 2012, **3**, 909–918.

140 M. H. Viet, P. H. Nguyen, S. T. Ngo, M. S. Li and P. Derreumaux, *ACS Chem Neurosci*, 2013, **4**, 1446–1457.

141 M. K. Tiwari and K. P. Kepp, *Alzheimers Dement*, 2016, **12**, 184–194.

142 R. Vassar, B. D. Bennett, S. Babu-khan, S. Kahn, E. A. Mendiaz, P. Denis, D. B. Teplow, S. Ross, P. Amarante, R. Loeloff, Y. Luo, S. Fisher, J. Fuller, S. Edenson, J. Lile, M. A. Jarosinski, A. L. Biere, E. Curran, T. Burgess, J.-C. Louis, F. Collins, J. Treanor, G. Rogers and M. Citron, *Science*, 1999, **286**, 735–741.

143 L. Zhou, N. Brouwers, I. Benilova, A. Vandersteen, M. Mercken, K. van Laere, P. van Damme, D. Demedts, F. van Leuven, K. Sleegers, K. Broersen, C. van Broeckhoven, R. Vandenberghe and B. de Strooper, *EMBO Mol Med*, 2011, **3**, 291–302.

144 Y. W. Zhang, R. Thompson, H. Zhang and H. Xu, *Mol Brain*, 2011, **4**, 1-13

145 A. Hatami, S. Monjazeb, S. Milton and C. G. Glabe, *J Biol Chem*, 2017, **292**, 3172–3185.

146 D. Kaden, A. Harmeier, C. Weise, L. M. Munter, V. Althoff, B. R. Rost, P. W. Hildebrand, D. Schmitz, M. Schaefer, R. Lurz, S. Skodda, R. Yamamoto, S. Arlt, U. Finckh and G. Multhaup, *EMBO Mol Med*, 2012, **4**, 647–659.

147 Y. Tian, Z. Wang, Z. Ding and H. Zhang, *Admet & Dmpk*, 2015, **3**, 203–215.

148 C. Morgado, M. A. Vincent, I. H. Hillier and X. Shan, *Phys Chem Chem Phys*, 2007, **9**, 448-451

149 F. Schubert, M. Rossi, C. Baldauf, K. Pagel, S. Warnke, G. von Helden, F. Filsinger, P. Kupser, G. Meijer, M. Salwiczek, B. Koksch, M. Scheffler and V. Blum, *Phys Chem Chem Phys*, 2015, **17**, 7373.

150 T. van Mourik, *J Chem Theory Comput*, 2008, **4**, 1610–1619.

151 L. Wa Chung, W. M. C Sameera, R. Ramozzi, A. J. Page, M. Hatanaka, G. P. Petrova, T. v Harris, X. Li, Z. Ke, F. Liu, H.-B. Li, L. Ding and K. Morokuma, *Chem Rev*, 2015, **115**, 5678-5796

152 H. M. Senn and W. Thiel, *Angew Chem Int Edit*, 2009, **48**, 1198–1229.

153 S. Boopathi and P. Kolandaivel, *Mol Bio Syst*, 2015, **11**, 2031.

154 L. J. Gutierrez, E. Barrera Guisasola, N. Peruchena and R. D. Enriz, *Mol Simulat*, 2016, **42**, 196-207

155 E. E. Barrera Guisasola, L. J. Gutiérrez, R. E. Salcedo, F. M. Garibotto, S. A. Andujar, R. D. Enriz and A. M. Rodríguez, *Comput Theor Chem*, 2016, **1080**, 56-65

156 B. Urbanc, L. Cruz, F. Ding, D. Sammond, S. Khare, S. v. Buldyrev, H. E. Stanley and N. v. Dokholyan, *Biophys J*, 2004, **87**, 2310–2321.

157 B. Urbanc, M. Betnel, L. Cruz, G. Bitan and D. B. Teplow, *J Am Chem Soc*, 2010, **132**, 4266-4280

158 B. Mehrazma and A. Rauk, *J Phys Chem A*, 2019, **123**, 4658–4670.

159 M. Cecchini, R. Curcio, M. Pappalardo, R. Melki and A. Caflisch, *J Mol Biol*, 2006, **357**, 1306–1321.

160 Y. Sugita and Y. Okamoto, *Chem Phys Lett*, 1999, **314**, 141-151

161 R. Qi, G. Wei, B. Ma and R. Nussinov, *Methods Mol Biol*, 2018, **1777**, 101-119

162 A. Baumketner and J.-E. Shea, *J Mol Biol*, 2007, **366**, 275-285

163 P. H. Nguyen, A. Mai, S. Li and P. Derreumaux, *Phys Chem Chem Phys* 2011, **13**, 9778–9788.

164 W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J Am Chem Soc*, 1996, **118**, 11225-11236

165 A. A. Chen and R. v Pappu, *J Chem Phys B*, 2007, **111**, 11884-11887

166 W. R. P. Scott, P. H. Hu, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Kru and W. F. van Gunsteren, *J Phys Chem A*, 1999, **103**, 3596-3607

167 S. T. Ngo, A. Huynh, M. Hung, C. Duc, T. Truong and M. T. Nguyen, *Phys Chem Chem Phys*, 2017, **19**, 1909–1919.

168 K. Jong, L. Grisanti and A. Hassanali, *J Chem Inf Model*, 2017, **57**, 1548-1562

169 H. Zerze, C. M. Miller, D. Granata and J. Mittal, *J Chem Theory Comput*, 2015, **11**, 2776-2782

170 J. C. Jose, P. Chatterjee and N. Sengupta, *PLoS ONE*, 2014, **9**, e106883

171 T. Shen and D. Hamelberg, *J Chem Phys*, 2008, **129**, 034103:1-9

172 L. C. T. Pierce, R. Salomon-Ferrer, C. A. F. de Oliveira, J. A. McCammon and R. C. Walker, *J Chem Theory Comput*, 2012, **8**, 2997-3002

173 Y. Zhang, M. Hashemi, Z. Lv and Y. L. Lyubchenko, *Nanoscale*, 2016, **8**, 18928.

174 K.-F. Huang, S.-H. Dong, S.-S. Zhong, H. Li and L.-L. Duan, *Commun Theor Phys*, 2019, **71**, 1121.

175 J. Wang, P. R. Arantes, A. Bhattarai, R. v. Hsu, S. Pawnikar, Y. ming M. Huang, G. Palermo and Y. Miao, *WIREs Comput Mol Sci*, 2021, **11**, e1521

176 L. Tran and T. Ha-Duong, *Peptides*, 2015, **69**, 86–91

177 L. Hou and M. G. Zagorski, *J Am Chem Soc*, 2006, **128**, 9260-9261

178 C. D. Syme and J. H. Viles, *Biochim Biophys Acta*, 2006, **1764**, 246-256

179 J. Danielsson, R. Pierattelli, L. Banci and A. Gräslund, FEBS J, 2007, **274**, 46–59.

180 N. G. Nair, G. Perry, M. A. Smith and V. P. Reddy, *J Alzheimers Dis*, 2010, **20**, 57–66.

181 J. Shearer, P. E. Callan, T. Tran and V. A. Szalaiz, *Roy Soc Ch*, 2010, **46**, 9137-9139

182 C. D. Syme, R. C. Nadal, S. E. J. Rigby and J. H. Viles, *J Biol Chem*, 2004, **279**, 18169-18177

183 P. Dorlet, S. Gambarelli, P. Faller and C. Hureau, *Angew Chem Int Edit*, 2009, **48**, 9273–9276.

184 B. Strodel and O. Coskuner-Weber, *J Chem Inf Model*, 2019, **59**, 1782–1805.

185 T. Marino, N. Russo, M. Toscano and M. Pavelka, *Interdiscip Sci*, 2010, **2**, 57–69.

186 T. Prosdocimi, L. de Gioia, G. Zampella and L. Bertini, *J Biol Inorg Chem*, 2016, **21**, 197–212.

187 J. Alí-Torres, J. D. Maréchal, L. Rodríguez-Santiago and M. Sodupe, *J Am Chem Soc*, 2011, **133**, 15008–15014.

188 T. Dudev and C. Lim, *Annu Rev Biophys*, 2008, **37**, 97-116

189 A. Maiorana, T. Marino, V. Minicozzi, S. Morante and N. Russo, *Biophys Chem*, 2013 **182**, 86-93

190 A. A. Kulikova, P. O. Tsvetkov, M. I. Indeykina, I. A. Popov, S. S. Zhokhov, A. v Golovin, V. I. Polshakov, S. A. Kozin, E. Nudler and A. A. Makarov, *Mol BioSyst*, 2014, **10**, 2590.

191 A. J. Cohen, P. Mori-Sánchez and W. Yang, *Science*, 2008, **321**, 792–794.

192 N. Schuch and F. Verstraete, *Nat Phys*, 2009, **5**, 732-735

193 T. Vreven, K. Morokuma, O. Farkas, H. B. Schlegel and M. J. Frisch, *J Comput Chem*, 2003, **24**, 760-9

194 F. Azam, N. H. Alabdullah, M. Ehmedat, A. R. Abulifa, I. Taban and S. Upadhyayula, *J Biomol Struct Dyn*, 2018, **36**, 2099-2117

195 C. Lee, W. Yang and R. G. Parr, *Phys Rev B*, 1988, **37**, 785-789

196 S. Boopathi and P. Kolandaivel, *Proteins*, 2016, **84**, 1257–1274.

197 N. Al-Shammari, L. Savva, O. Kennedy-Britten and J. A. Platts, *Comput Biol Chem*, 2021, **93**, 107540.

198 M. Turner, S. T. Mutter, O. D. Kennedy-Britten and J. A. Platts, *PLoS ONE*, 2019, **14**, 1–14.

2 Computational theory

2.1 Overview

This section will seek to explain the theory behind the computational methods employed in this report, mainly focusing on molecular mechanics and molecular dynamics. The main information on concepts within this chapter come from a range of computational chemistry textbooks[1-4].

2.2 Born-Oppenheimer approximation

The Schrödinger equation is used to calculate distribution of electrons within a system. Solving for this equation exactly is impossible when considering multi-body systems of three or more particles. Certain approximations must therefore be made for systems of these sizes. Physicists Born and Oppenheimer[5] remarked that nuclei are heavier than electrons and as such move on a timescale of around two orders of magnitude longer. It was therefore inferred that when solving the Schrödinger equation, values for the nuclei could be fixed when solving for electronic motions. The full Hamiltonian can be divided into both nuclear and electronic descriptors (Equation 2.1).

$$\hat{\mathrm{H}}(\mathbf{Total}) = \hat{\mathrm{H}}(\mathbf{Electronic}) + \hat{\mathrm{H}}(\mathbf{Nuclear}) \qquad (\mathbf{2.1})$$

Nuclei are heavy enough that quantum effects are negligible and nuclear descriptors can act as a good approximation as classical models. Quantum mechanics solves the Schrödinger equation, and nuclear motions ($\phi(R)$) are effectively ignored and molecules treat these implicitly within descriptors of electronic motions ($\psi(r)$) using the Born-Oppenheimer approximation to describe the electronic energy of the system in terms of nuclear positions only (Equation 2.2).

$$\boldsymbol{\Psi(R,r) \approx \varphi(R)\psi(r)} \qquad (\mathbf{2.2})$$

2.3 Molecular Mechanics (MM)

2.3.1 General concepts

As molecules increase in the number of atoms, so do the degrees of freedom of the system. This makes some commonly used computational methods potentially unsuitable for conformational searching on certain larger structures or biomolecules. Quantum mechanical methods such as *ab initio* or semi-empirical calculations commonly used for smaller systems may not be able to complete extensive sampling of geometries within a reasonable timeframe and, if even possible, may be extremely computationally expensive. Molecular mechanics (MM) is a relatively cheap empirical computational method that calculates energy of a system based on positioning of nuclei by considering atoms as discrete particles based on their mass, charge, polarizability and bonds with adjacent atoms.

Characteristics of atoms and bond types transfer reasonably well between systems due to the way they attempt to replicate "ideal" values and remain consistent between like bonds/functional groups, meaning that accurate comparisons can be drawn by application of tested methods on structures containing similar atom interactions. By using classical or Newtonian "ball and spring" mathematical models, energies of the system are calculated from bond lengths from strains (the "spring") acting on atoms (the "ball") observed in different geometries. Molecules will attempt to replicate ideal bond length and angles by adopting geometries and structures as similar as possible to these values. These models can allow for transferability of structurally similar units between molecules, such as treating C-H bonds as approximately constant unless adjacent atoms and functional groups affect them[6,7]. Ideal values can be obtained from calculated and experimental sources.

Geometry optimisation of a given conformation involves modifying the structure gradually over iterative increments until a low-energy minimum is identified. These local minima may not be indicative of the global minima, such that this further conformational searching is conducted to sample a range of structures to give a potential energy surface (PES) showing several minima structures being sampled.

Particle interactions can be described as either force (F) or potential (V) which are equivalent to one another due to force being the derivative of the potential in respect to position (r). MD is described as a deterministic method, meaning position of particles can be

predicted at a given time (t) using Newton's second law over a timeseries. Knowing that force equals mass (m) times acceleration (a), in classical mechanics at low velocities, force can be expressed to solve for second derivatives of positions with respect to time, as seen in (Equation 2.3).

$$\boldsymbol{F = ma}$$

$$\boldsymbol{F = -\left(\frac{\partial v}{\partial r}\right) = m\left(\frac{\partial^2 r}{\partial t^2}\right)} \qquad \textbf{(2.3)}$$

2.3.2 Forcefield energy terms

For MM calculations, a forcefield is applied to the system, which is series of mathematical representations and constants encompassing key components crucial to the simulation. These include descriptors for atoms and their behaviour, for example differentiating carbon atoms possessing different properties and interactions depending on the functional group(s) associated to them. MM utilises a series of equations empirically derived from the Born-Oppenheimer PES to give a set of functions based on the Westheimer method which provides values for structures and energies for bonds and atoms[8]. These parameters can be adjusted to be optimised for each system to generate properties that replicate those seen experimentally. Total potential energy of the system based on only nuclear descriptors can be characterised by bonded and non-bonded terms (Figure 2.1 and Equation 2.4).

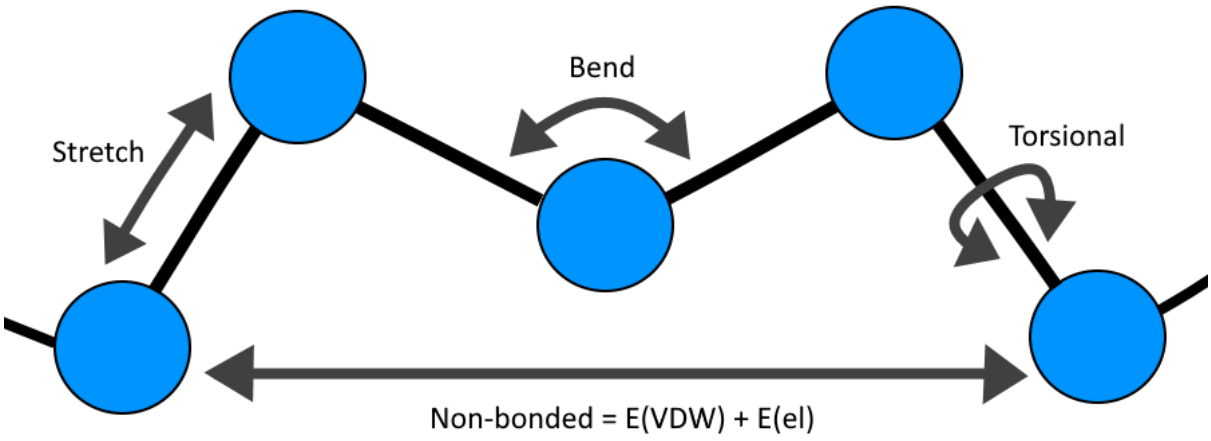


Figure 2.1- Representation of force field energy terms

$$E(Total) = \underbrace{E(Stretch) + E(Bend) + E(Torsion)}_{Bonded\ Terms} + \underbrace{E(VDW) + E(el)}_{Nonbonded\ Terms} \quad (2.4)$$

The bonded terms are energies associated with a bond as it stretches, bends at an angle, and rotates around a dihedral. Bond stretching descriptors (Equation 2.5) model the vibrational motions between bonded atoms using a harmonic potential (where k = the spring constant, r = the bond distance and r_o = the bond distance at equilibrium). All units of energy are in kcal/mol.

$$E(Stretch) = \sum_{bonds} k(r - r_0)^2) \quad (2.5)$$

The sum of all bond angles between sets of three interconnected atoms (also modelled using a harmonic potential) gives the second term of the equation for total potential energy (Equation 2.6) (k_ϑ = the angle constant, ϑ = the bond angle in radians and ϑ_0 = the bond angle, in radians, at equilibrium).

$$E(Bend) = \sum_{angles}(k_\theta(\theta - \theta_0)^2) \quad (2.6)$$

Torsional angles around four interconnected atoms (e.g., A-B-C-D) are described via characteristics of the first 3 atoms in the series (e.g., A-B-C) and the last 3 atoms (e.g., B-C-D) for all dihedrals within a system. In this instance, k is a multiplicative constant (not the same value of k as in Equation 2.6), n is a non-negative constant representing periodicity and φ is the phase shift angle in radians. Ω is the angle (in radians) between both the A-B-C and B-C-D planes of the dihedral (Figure 2.2 and Equation 2.7).

$$E(Torsion) = \sum_{torsions}(k(1 + cos(n\Omega + \phi)) \quad (2.7)$$

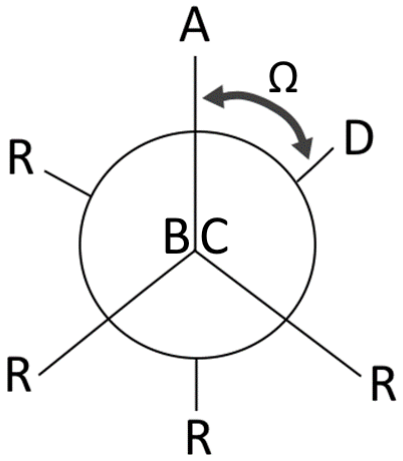


Figure 2.2- Simplified Newman projection of potential dihedral angles (*Ω*) between atoms A and D viewing down the plan of bond between atoms B and C, R represents any functional group attached to B and C

The two non-bonded terms are accounted for by Van der Waals forces and electrostatic interactions and usually require the bulk of the computational power over the course of dynamics simulations to calculate. A 12-6 Lennard-Jones potential[2] (Equation 2.8) is used as a representation of VDW forces. This equation uses expressions of *r* which is the distance between two nonbonded atoms (*A* and *B*), *ε* which represents the VDW well depth and *σ* which is the distance when *E(VDW)* = 0. With this, inter-atomic attractive and repulsive forces decay quickly and are negligible beyond around 10 Å in typical systems. This allows for use of a cut-off that accelerates calculations on larger systems, negating VDW interactions with zero energy beyond these distances. In some instances, a discontinuity is observed in potential forces due to truncation at the cut-off value. Typically, constant values are added below the cut-off to uniformly shift the potential of the system to meet that of the original potential. Alternatively, a switching function modifies the shape of the potential function by applying a function such as a polynomial to ensure a smoother curve of the non-bonded potential at the cut-off point.

$$\boldsymbol{E(VDW_{LJ})} = \sum_{A=1}^{N}\sum_{B=A+1}^{N} \mathbf{4\varepsilon}\left[\left(\frac{\sigma_{AB}}{r_{AB}}\right)^{12} - \left(\frac{\sigma_{AB}}{r_{AB}}\right)^{6}\right] \qquad (2.8)$$

An uneven distribution of polarity between non-bonded atoms can be modelled via a charge potential within the forcefield of MM simulations (Equation 2.9). q_Aq_B signify partial atomic charges between nonbonded atoms (again represented by *A* and *B*), *r* once again represents

distance between atoms and ε_o is vacuum permittivity which is a measure of an electric field's capacity to permeate a vacuum (also known as permittivity of free space).

$$\boldsymbol{E(el)} = \sum_{A=1}^{N} \sum_{B=A+1}^{N} \frac{q_A q_B}{4\pi\varepsilon_0 r} \qquad (2.9)$$

These five functions of potential energy are incorporated into the forcefield setup, and allow extremely good approximations to be made, comparable to experimental structures and energies, in a rapid calculation time. As MM simulations progress, geometry of structures sampled can be plotted based on nuclear positions as a function of their associated energies on what is known as a potential energy surface (PES). By sampling points on this PES that vary in bond lengths, angles, and torsional angles over the course of a simulation, MM attempts to locate low-energy conformations that replicate those seen *in vivo* and *in vitro*. This optimisation of structures is effectively the means of determining the identity of molecules positioned at stationary points, especially minima. These are calculated from the first derivative of *E(Total)* based off the individual terms shown in Equation 2.4 and forces from Equation 2.3.

An effective forcefield should incorporate appropriate parameters for variables such as force constants, bond lengths and angles of the system for all the atoms of interest. There are multiple forcefields available for MM calculations that offer bespoke parameterisation for different sets of atoms. For example, some are utilised mainly for proteins and macromolecules; AMBER[9-11], CHARMM[12], GROMOS[13-15] are just a few examples of these. AMBER (Assisted Model Building with Energy Refinement)[16] is a molecular simulation computational package encompassing the set of AMBER forcefields mentioned previously, which was designed to simulate biomolecular systems via molecular dynamics (MD). The computer program AMBER incorporates a collection of tools that allows for parameterisation and preparation of structures such as proteins and nucleic acids for analysis. In addition, calculations and post-completion analysis of trajectories are all offered within the same program. AMBER forcefields are not limited to being used within the computer package of the same name and additionally, AMBER software offers utilisation of many other forcefields.

AMBER's ff14SB forcefield[17] builds upon and refines parameters derived in previous iterations (ff94[18], ff99[19] and ff99SB[20]). ff94 applied more generic torsional parameters which

were improved in ff99 on a greater range of small molecules such as amino acids. ff99SB showed great improvement in modelling backbone dihedrals to its predecessors by refining the bond rotational energy profiles as well as 1-4 non-bonded interactions in dihedrals. Despite this, ff99SB still showed slightly inaccurate rotamer preferences due to being mainly based on ff99 dihedral parameters. The introduction of ff14SB showed further developments on dihedral parameterisation by using QM benchmarks and empirical corrections to backbone dihedral energy profiles. ff14SB is still one of the recommended parameter sets for protein modelling by AMBER.

2.3.3 Solvation of a system

In order to best replicate naturally existing conditions of biomolecules, a solvation model can be implemented into simulations to mimic the behaviour of molecular structure of interest within a solvent. Conformational searches carried out in a vacuum (as a representation of gas-phase) can potentially produce infeasible structures due to interactions that would usually occur between the simulated molecule and the solvent, instead occurring intramolecularly.

One method of solvating a system is enclosing the solute being investigated in a sphere or cube of an explicitly defined solvent. Rigid water models use electrostatic interactions based on Coulomb's Law, as well as repulsion and dispersion forces from the Lennard-Jones potential mentioned earlier, to model water as a solvent for the system. The TIP3P[21] and TIP4P[22] (transferable intermolecular potential with 3 and 4 points respectively) can be employed as explicit solvation models used in the simulation of biomolecules.

The two methods of solvation for simulating water mentioned here uses two different models of water shown in Figure 2.3. Both models use the three atoms of H_2O with associated point charges and the Lennard-Jones parameters on each oxygen atom. The TIP4P model also includes a negatively charged fourth "dummy" atom, which improves distribution of electrostatic charge around the solvent. This does not, however, accurately replicate certain properties of water in bulk, such as density, and thus requires further development before use on a wider range of systems[23-25].

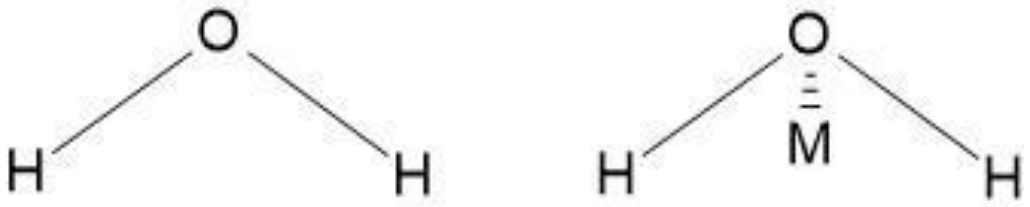


Figure 2.3- TIP3P and TIP4P solvation models of water (left and right respectively)

The benefits of running MM and MD calculations in explicit solvent can be seen from results replicating those seen experimentally even for systems with many degrees of freedom[26]. The caveat to this is a greater computational cost to run these types of simulations that increases dependent on the level of solvation. Additionally, explicit solvation has been shown to affect conformers of extended peptides studied via MD[27].

Calculations involve implementation of a solvation environment with a choice of several solvent models, depending on the system being studied. So-called implicit solvent models often involve the solute being placed in a cavity surrounded by the chosen solvent model. The free energy of solvation of a given system can be expressed in terms of free energy changes (Equation 2.10), including the energies associated with creating a cavity within the solvent, as well as electronic, dispersion and exchange interactions between solvent and solute.

$$\Delta \boldsymbol{G}_{\boldsymbol{solvation}} = \Delta \boldsymbol{G}_{\boldsymbol{cavity}} + \Delta \boldsymbol{G}_{\boldsymbol{electronic}} + \Delta \boldsymbol{G}_{\boldsymbol{dispersion}} + \Delta \boldsymbol{G}_{\boldsymbol{exchange}} \qquad \textbf{(2.10)}$$

To counteract the expense observed in explicitly solvated systems, an implicit (or continuum) solvation model can be implemented which is use of a physical representation replicating the characteristics of water molecules (or indeed other solvents or lipids) observed in explicit solvation. This simplified model allows for rapid calculations at a fraction of the cost. The Generalized Born Surface Area (GBSA)[28-30] approach is an algorithm designed to provide empirical approximations of the $\Delta G_{\text{electronic}}$ descriptor in the equation for solvation energy (Equation 2.10) based on Born radii of atoms, bond distances and atomic volume[31]**.** In this dielectric, implicitly solvated system, interactions between explicit solvent molecules and the solute structure being modelled are negated to allow for faster calculation speeds. It is a good choice for biomolecules as greater conformation sampling can occur without the necessity to equilibrate the solvent system beforehand.

Alternatives to GBSA include the conductor-like screening model (COSMO)[32] solvation model that generates a cavity for the solute based on the solvent accessible surface and VDW radii, calculating approximate values using a scaling factor for the surface of the solvent in sections. The polarizable continuum model (PCM)[33-34] utilises Hartree-Fock and DFT levels of theory to provide QM descriptors making it more suitable for *ab initio* studies compared to COSMO.

2.3.4 Advantages and disadvantages of MM

Molecular mechanics is a strong candidate for calculations involving large bioinorganic systems. As mentioned previously, the exclusion of electronic motion makes MM faster in comparison to QM simulations, lowering computational expense. In addition, the transferability of accurate parameters and mathematical representations of atomic behaviours between systems mentioned earlier mean that previously modelled structures generate a good starting point for studies on molecules where further research is required.

However, the converse to this is MM suffers in cases where experimental or high-level theoretical data is lacking. There is also some potential evidence of bias towards user input of geometries, i.e., if the starting structure is built in a square planar geometry, it may adopt conformations that replicate this whereas experimental data may show preference for tetrahedral geometry. As mentioned prior, parameterisation means MM calculations require a forcefield best suited to the type of molecule being simulated, but no one forcefield fits all types and sizes.

In addition, a quantum approach is typically used in reactions involving bond formation and breaking due to the nature of only considering nuclear descriptors of ground-state systems. The system usually is explicitly programmed to contain any formal bonds during setup. MM methods do exist to allow approximations of reactive interactions involving the breaking and formation of bonds, such as the ReaxFF[35] and Empirical Valence Bond[36] approaches, but QM is usually still the preferred model for interactions of this type.

The removal of electronic terms in forcefields of MM simulations means that any structures containing transition metals will affect calculations; d-orbital interactions are not generally

considered in most methods discussed above, and add another level of complexity to the Newtonian model (this is explored further in Section 2.3.5-2.3.6).

2.3.5 Modelling transition-metal complexes using MM

As discussed previously, modelling transition metal-complexes computationally present novel problems when using MM. Whilst QM implicitly considers d-orbital effects, treating particles based on nuclear positions in MM makes it difficult to consider electronic interactions. d-orbitals (and even f-orbitals) have much more complicated shapes, leading to more complex modes of bonding and geometries, especially when considering their variable oxidation states. Electronic spin can be collectively described as either high- or low-spin states with different properties but at relatively similar energies. Jahn-Teller distortions add to the difficulty in modelling these structures using classical methods, and these factors make the transferability of bond parameters less feasible than for organic structures[37].

Certain methods can be implemented to account for the electronic effects discussed previously. Ligand field molecular mechanics[38] builds on the equation for total potential energy (Equation 2.4) and applies a ligand field stabilisation energy (LFSE) to explicitly account for d-orbital and electronic effects of any metal centres.

LFSE (or crystal field theory, CFT) is a description of the orbital arrangement and bonding in metal-ligand/protein complexes and models the geometry and coordination of metal-ligand structures. Based on molecular orbital splitting profiles for an octahedral environment (Figure 2.4) defining electrons as either high or low spin states, LFSE can be calculated via equation (Equation 2.11) where Δ_o represents the energy difference based on octahedral geometry which is then added to the total potential energy calculation (Equation 2.12).

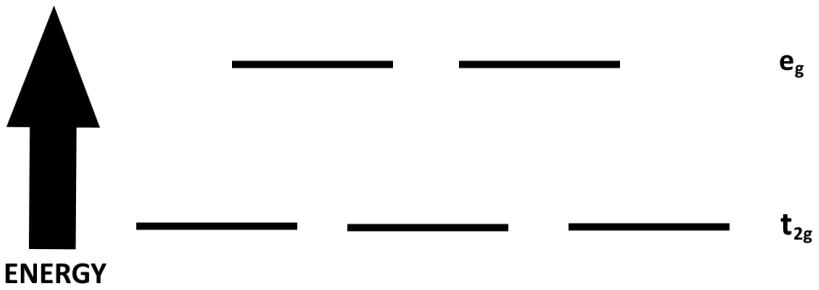


Figure 2.4- Ligand field splitting of octahedral orbitals and associated electronic energies

$$\boldsymbol{LFSE} = \left[\left(\mathbf{0.6} \times \boldsymbol{Ne_g electrons}\right) - \left(\mathbf{0.4} \times \boldsymbol{Nt_{2g} electrons}\right)\right]\Delta_o \qquad (2.11)$$

$$\boldsymbol{E(Total)} = \boldsymbol{E(Stretch)} + \boldsymbol{E(Bend)} + \boldsymbol{E(Torsion)} + \boldsymbol{E(VDW)} + \boldsymbol{E(el)} + \boldsymbol{LFSE} \quad (2.12)$$

This method provides MM parameters representative of experimental data for systems containing transition metals; which allows MM to accurately simulate d-orbital effects on the system. This is an additional computational demand of the generally more simplistic MM methods, but still well within the scope of its ability to accurately simulate these types of structures for large molecular systems in realistic timescales. Geometries around a metal centre in LFMM come from the balance between ligand-ligand repulsion and the LFSE that determines the d-orbital splitting.

2.3.6 Metal-centre parameter builder (MCPB)

Modelling parameters for metalloprotein systems can be calculated via use of a computational package titled Metal Center Parameter Builder (MCPB)[39]. This method takes metalloproteins or organometallic systems and prepares them for MM/MD calculations. MCPB will first determine primary and secondary ligands associated with a transition metal centre using descriptors from studies by Harding[40]. A coordination sphere is thus modelled, which can be used for *ab initio* calculations to produce forcefield parameters compatible with AMBER-style forcefields for organic/biological systems. Metals are labelled by element and bound atoms as well as being assigned descriptors based on *ab initio* or DFT ground state properties, such as mass, charge, and d-orbital configurations. With this method, two different models are utilised, where parameters including force constants and equilibrium bond lengths/angles are generated on a "small" model; non-bonded parameters, such as charges, are generated from a larger model including more of the coordination sphere of the metal ion in question.

Using these models, MCPB is able to parameterise metal centres and bound residues with a reasonably high degree of both accuracy and speed. The small model uses the metal centre as well as CH_3R groups to replicate side chains and terminal ACE or NME residues typically

used to cap the end of peptides but in this instance represent the remaining peptide backbone. The large model will cap metal-bound amino acids using ACE or NME again, whereas and coordinating residues within less than 5 amino acids away from one another will maintain these residues in between but they will instead by replaced by glycine (GLY) residues[39,41]. These models are then able to empirically generate parameters specifically for the metal centre to be used in subsequent MM or molecular dynamics (MD) calculations.

2.4 Molecular dynamics (MD)

2.4.1 General concepts

As discussed previously, MM models atoms use classical, or Newtonian, methods to estimate energy, and hence force. MD uses these forces to propagate a system over a time-dependent series to find accessible structures at a given temperature. As well as this, it is a viable theoretical method of simulating large-scale biological interactions (some MD simulations have been performed on systems of millions of atoms) such as dissociation/association and protein folding, in addition to local interactions and movements of individual atoms or sidechains. The scope to simulate varying sized structures via MD allows study of macroscopic systems at a molecular level.

The first recorded MD simulations were performed by Alder & Wainwright[42,43] modelling interactions of systems of hard spheres. These experiments formed the basis of future simulations on biomolecules and materials. Further development and refinement of this as a viable method of computational study led to the first realistic application of MD when Stillinger & Rahman[44] studied a system of liquid water and McCammon *et al.*[45] conducted the first protein MD simulation on Bovine Pancreatic Trypsin Inhibitor (BPTI).

MD is specifically utilised for studying and making predictions on structural changes over a time-dependent series. A key aspect of drug design is conformational analysis of biomolecules to assess viability as potential candidates for further stages of development. Due to the minor differences on an atomistic level, computational studies are sometimes preferable as an initial method to deduce suitable structures for experimental treatments. By utilising MM as a means of energy evaluation, properties of biomolecules such as protein

folding, or enzyme functionality can be carried out using MD at less of a computational expense and on larger systems when compared to QM methods. However, with an increase in the number of bonds and atoms present within a system, the number of degrees of freedom increases rapidly as there are 3*N*-6 degrees of freedom in non-linear molecules (where *N* is the number of atoms). In conformation searching, the cartesian method of sampling changes the coordinates of atoms randomly whereas the dihedral approach rotates torsional angles randomly.

2.4.2 Evolution of time-series and conformational ensembles

Phase space describes a multi-dimensional representation of the momentum and positions of particles within a system. All possible system states can be represented by all possible values of these variables. At any single point of time during simulation, individual atoms can be defined by their positions relative to other atoms, known as μ-space. The sum of μ-spaces can then be used to phase space for the system as a whole (also known as gamma or Γ-space). This can be used to describe the state of the system at a given moment in time. In order to ensure sufficient exploration of the conformational space, a suitable timestep integration must be selected during setup. The integration timestep of a simulation is selected based on mobility of a system and vibrational energies of flexible structures. Selecting a smaller timestep will improve the quality of integration however it will limit the sampling ability of the simulation due to limited exploration of the phase space. Using a larger timestep will resolve this but may also yield unstable simulation data due to high-energy interactions between particles due to high-frequency vibrations.

In statistical mechanics, the Maxwell-Boltzmann distribution can be used to describe the probability for the population of states of energy and thus speed of particles in a given system based off of the temperature, *T*, and mass of the particles. Random sampling of velocities within this distribution can be applied to atoms in the starting structure of an MD simulation and the evolution of structure and energy can be recorded over time as a result of the application of these velocities. These velocities should be random to ensure notable changes in structures between repeated simulations. If not, the system will evolve in a similar way each time due to the deterministic treatment of nuclear descriptors. Application

of random, initial velocities is the only degree of randomness applied using MD; subsequent evolution over time is deterministic.

As time progresses, an equilibration point is reached where any artefacts from the starting point at t = 0 ns are removed. The endpoint of this initial simulation can be used as the start point of subsequent MD simulations to give a better representation of the system using only data from the equilibrated structure. Alternatively, post-trajectory analysis can be used to determine the equilibration point and any data from prior to this is discarded allowing for subsequent analysis to be performed only on the equilibrated system

The series of structures sampled during MD that satisfy the thermodynamical properties of the system can be classified as being within the same ensemble. Canonical ensembles contain a fixed number of atoms and maintain constant volume and temperature (also known as the NVT ensemble). In contrast, a micro-canonical ensemble (or NVE ensemble) will retain a constant energy of the system whilst allowing temperature of the system to change accordingly as the structure evolves over time.

The Boltzmann distribution links the probability of the system under study occupying a particular state to the energy of that state, according to the fundamental relation in Equation 2.13, where p_i is probability distribution and Z is the partition function.

$$p_i = \frac{exp^{\varepsilon_i/T}}{Z} \tag{2.13}$$

This shows that lower energy states are more likely to be populated than higher ones. In all the MD studies used here, temperature T is kept constant at 310 K, so the Boltzmann distribution determines the probability of the simulation reaching high energy states such as intermediates and transition states at this temperature. The ergodic hypothesis states that if a system evolves in time indefinitely, then all potential states of a system are equiprobable of being sampled[46,47]. Experimental ensemble averages can be calculated for all functions of momenta (p) and positions (r) but in MD, points within the ensemble are calculated in sequential iterations over time. Additionally, simulations are typically only performed on individual, or a small number, of structures as opposed to experimental ensembles which are a collection of molecules, often as many as 10^{23}. This time average ($\langle A \rangle_{time}$), is calculated based on simulation time and number of time steps. Ideally, sufficient simulation

length will lead to adequate representative conformations being sampled in order to satisfy the hypothesis that:

$$\langle A \rangle_{ensemble} = \langle A \rangle_{time} \quad (2.14)$$

2.4.3 Advantages and disadvantages of MD

Sufficient sampling of conformational ensembles satisfies the ergodic hypothesis (Equation 2.13), allowing for resulting structures and energies to be compared with those observed experimentally, but there is debate of what classifies as "sufficient sampling"[48]. By using MM as a means of energy evaluation, this allows for generally rapid calculations and the implementation of parameterisation for metal centres means it can be used as a less computationally demanding method on larger systems compared to QM. As system size and simulation time increase, however, the computational cost of MD can still be high. Additionally, the use of approximations and empirical forcefields in MD means results are often not as accurate as QM methods which can lead to potential issues when used for things such as drug design, but this is a trade-off for potentially enhanced efficiency[49]. As time-dependent evolution of the system samples conformations along the PES, high energy barriers may limit transitions between minima structures and become stuck in energy basins producing similar structures for a substantial duration of simulation time.

2.4.4 Accelerated molecular dynamics (aMD)

To overcome the issue of potentially insurmountable energy barriers and increase conformational sampling of flexible structures, an algorithm called accelerated molecular dynamics can be applied to the system. This method utilises a boost potential to decrease the potential energy required to overcome the energy barrier between basins on the PES. This is achieved by raising the potential of minima structures in energy wells, *V(r)*, to an increased potential, *V*(r)* more suitable for transitioning between structures potentially unobtainable in conventional MD[50,51] (figure 2.5).

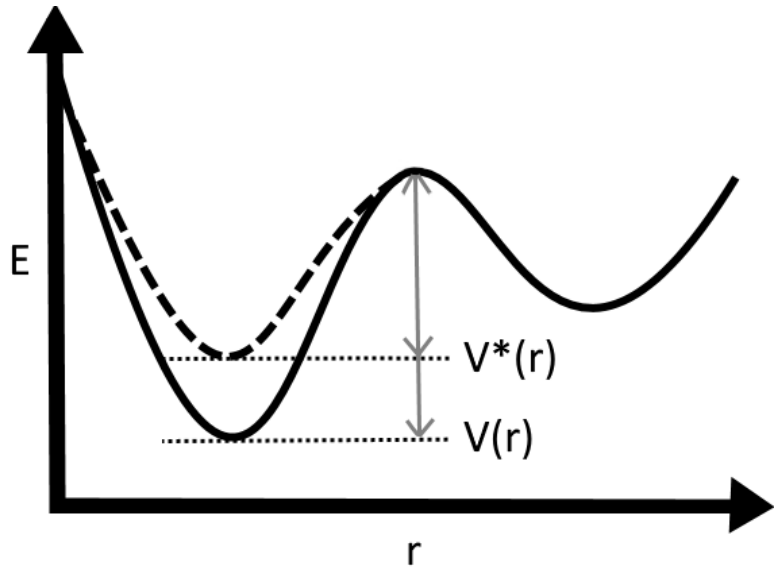


Figure 2.5- aMD application of boost potential to energy wells (solid curve, *V(r)*), when energy threshold between minima is too high, with modified potential (dashed curve, *V*(r)*)

The boost in potential energy applied to the system, Δ*V(r)*, can be calculated in Equation 2.14 when *V(r)* < *E* and *α* is the acceleration factor. The value of *α* can be amended to alter the trajectory; when the value decreases, more transitions occur between low-energy states by generating a flatter PES[52].

$$\Delta \boldsymbol{V}(\boldsymbol{r}) = \frac{\left(\boldsymbol{E}+\boldsymbol{V}(\boldsymbol{r})\right)^{2}}{\boldsymbol{\alpha}+\boldsymbol{E}-\boldsymbol{V}(\boldsymbol{r})} \qquad (2.15)$$

There are two different versions of aMD that can apply this energy bias to the system: to all dihedrals (known as "dihedral-boost", Equation 2.16) or to all atoms as well as the dihedral-boost (known as the "dual-boost, Equation 2.17)[53] and are calculated different to E(Total) and E(Torsion) in Equation 2.7. N_{res} defines the number of amino acid residues present and N_{atoms} the number of atoms. Here, α_{dihed} describes the acceleration factor with respect to all dihedrals in a system and α_{total} is with regards to all atoms in addition to the dihedral boost.

$$\boldsymbol{E}_{\boldsymbol{dihed}=} \boldsymbol{V}_{\boldsymbol{dihed\_avg}} + \boldsymbol{3.5N}_{\boldsymbol{res}},\ \boldsymbol{\alpha}_{\boldsymbol{dihed}} = \frac{\boldsymbol{3.5N}_{\boldsymbol{res}}}{\boldsymbol{5}} \qquad (2.16)$$

$$\boldsymbol{E}_{\boldsymbol{total}} = \boldsymbol{V}_{\boldsymbol{total\_avg}} + \boldsymbol{0.175N}_{\boldsymbol{atoms}},\ \boldsymbol{\alpha}_{\boldsymbol{total}} = \boldsymbol{0.175N}_{\boldsymbol{atoms}} \qquad (2.17)$$

Subsequently, reweighting of the PES post-completion of the MD simulation can provide the ensemble average, <*A*>, of the trajectory to consider the energy of the system without the bias potential[54] (Equation 2.18). In this equation, β equals $(k_BT)^{-1}$, where k_B is the Boltzmann constant and *T* is temperature; *A(r)* refers to the experimental ensemble average with respect to position, *r*.

$$\langle \boldsymbol{A} \rangle = \frac{\langle \boldsymbol{A}(\boldsymbol{r}) \mathbf{exp}\ (\boldsymbol{\beta}\Delta \boldsymbol{V}(\boldsymbol{r})) \rangle^*}{\langle \mathbf{exp}\ (\boldsymbol{\beta}\Delta \boldsymbol{V}(\boldsymbol{r})) \rangle^*} \qquad (2.18)$$

Standard practices use a short conventional MD simulation to allow for equilibration, and to generate dihedral and potential energies of the system that can be used to derive aMD parameters and values for the boost potential. The endpoint of the initial MD simulation can then be used as a starting point for subsequent aMD calculations.

2.5 References

1 T. Gund, in *Guidebook on Molecular Modelling in Drug Design*, ed. N. C. Cohen, Academic Press, Queensland, Australia, 1996, Chapter 3, 55-92

2 C. J. Cramer, *Essentials of Computational Chemistry: Theories and Models*, John Wiley & Sons, Chichester, 2002

3 E. G. Lewars, *Computational Chemistry: Introduction to the Theory and Applications of Molecular and Quantum Mechanics*, Springer, New York, 2003

4 F. Jensen, *Introduction to Computational Chemistry*, John Wiley & Sons, Chichester, 3rd edn, 2017

5 M. Born and J. R. Oppenheimer, *Annalen der Physik*, **389**, 457-484

6 M. Levitt, *J Mol Biol*, 1983, **168**, 595-620

7 M. Karplus and G. A. Petsko, *Nature*, 1990, **347**, 631-639

8 C. Altona and D. H. Faber, *Top Curr Chem*, 1974, **45** 1-38

9 D. A. Case, R. M. Betz, D. S. Cerutti, T. E. Cheatham, T.A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. M. Merz, G. Monard, H. Nguyen, H. T. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. Roitberg, C. Sagui, C. L. Simmerling, W. M. Botello-Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X .Wu, L. Xiao and P. A. Kollman. (2016), AMBER 2016, University of California, San Fransisco

10 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J Comput Chem*, 2004, **25**, 1157-74

11 B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S Swaminathan and M. Karplus, *J Comput Chem,* 1983, **4**, 187-217

12 K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov and A. D. Mackerell, *J Comput Chem*, 2010, **31**, 671–690.

13 W. F. van Gunsteren and H. J. C. Berendsen, Groningen Molecular Simulation (GROMOS) Library Manual, Biomos, Groningen, The Netherlands, 1987

14 W. R. P. Scott, P. H. Hu, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Kru and W. F. van Gunsteren, *J Phys Chem*, 1999, **103**, 3596-3607

15 C. Oostenbrink, A. Villa, A. E. Mark and W. F. van Gunsteren, *J Comput Chem*, 2004, **25**, 1656–1676.

16 D. A. Case, T. E. Cheatham III, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, *J Comput Chem*, 2005, **26**, 1668-88

17 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput*, 2015, **11**, 48.

18 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J Am Chem Soc*, 1995, **117**, 5179–5197.

19 J. Wang, P. Cieplak and P. A. Kollman, *J Comput Chem*, 2000, **21**, 1049-74

20 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins*, 2006, **65**, 712-25

21 P. Mark and L. Nilsson, *J Phys Chem A*, 2001, **105**, 9954-60

22 J. L. F. Abascal and C. Vega, *J Chem Phys*, 2005, **123**, 234505

23 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J Chem Phys*, 1983, **79**, 926–935.

24 J. D. Bernal and R. H. Fowler, *J Chem Phys*, 1933, **1**, 515-48

25 H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura and T. Head-Gordon, *J Chem Phys*, 2004, **120**, 9665-78

26 J. P. Linge, M. A. Williams, C. A. E. M. Spronk, A. M. J. J. Bonvin and M. Nilges, *Proteins*, 2003, **50**, 496-506

27 P. Florová, P. S. Sklenovsky´, P. Banáš and M. Otyepka, *J Chem Theor Comp*, 2010, **6**, 3569-79

28 W. Clark Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J Am Chem Soc*, 1990, **112**, 6127–9

29 G. D. Hawkins, C. J. Cramer and D. G. Truhlar, *Chem Phys Lett*, 1995, **246**, 122–129

30 A. v Onufriev and D. A. Case, *Ann Rev Biophys*, 2019, **48**, 275–296

31 D. Qiu, P. S. Shenkin, F. P. Hollinger and W. C. Still, *J Phys Chem*, 1996, **101**, 3005-14

32 A. Klamt and G. Schüürmann, *J Chem Soc*, 1993, **2**, 799–805

33 S. Miertuš, E. Scrocco and J. Tomasi, *Chem Phys*, 1981, **55**, 117–129

34 B. Mennucci, *Wires Comput Mol Sci*, 2012, **2**, 386–404

35 T. P. Senftle, S. Hong, M. Islam, S. B. Kylasa, Y. Zheng, Y. K. Shin, C. Junkermeier, R. Engel-Herbert, M. J. Janik, H. Metin Aktulga, T. Verstraelen, A. Grama and A. C. T. Van Duin, *Comp Mater Sci*, 2016, 15011

36 Y. Kulkarni, S. Caroline and L. Kamerlin, in *Advances in Physical Organic Chemistry*, ed I. H. Williams and N. H. Williams, Academic Press, London, 2019, Chapter 3, 69-104

37 P. Li and K. M. Merz Jr, *Chem Rev*, 2017, **117**, 1564-1686

38 R. J. Deeth, N. Fey and B. Williams-Hubbard, *J Comput Chem*, 2005, **26**, 123–130

39 P. Li and K. M. Merz, *J Chem Inf Mod*, 2016, **56**, 599–604

40 M. M. Harding, *Acta Crystallogr D*, 2006, **62**, 678-82

41 M. B. Peters, Y. Yang, B. Wang, L. Füsti-Molnár, M. N. Weaver and K. M. Merz Jr, *J Chem Theor Comput*, 2010, **6**, 2935-47

42 B. J. Alder and T. E. Wainwright, *J Chem Phys*, 1957, **27**, 1208-9

43 B. J. Alder and T. E. Wainwright, *J Chem Phys*, 1959, **31**, 459-66

44 F. H. Stillinger and A. Rahman, *J Chem Phys*, 1974, **60**, 1545-74

45 J. A. Mccammon, B. R. Gel and M. Karplus, *Nature*, 1977, **267**, 585-90

46 J. von Plato, *Arch Hist Exact Sci*, 1991**, 42**, 71-89

47 T. E. Cheatham and M. A. Young, *Biopolymers*, 2001, **56**, 232-256

48 N. Simányi, *Invent Math*, 2003, **154**, 123–178

49 P. Mo, C. Li, D. Zhao, Y. Zhang, M. Shi, J. Li and J. Liu, *NPJ Compu Mats*, 2002, **107**, 1-15

50 M. Bhandarkar, A. Bhatele, E. Bohm, R. Brunner, F. Buelens, C. Chipot, A. Dalke, S. Dixit, G. Fiorin, P. Freddolino, P. Grayson, J. Gullingsrud, A. Gursoy, D, Hardy, C. Harrison, J. Hénin, W. Humphrey, D. Hurwitz, N. Krawetz, S. Kumar, D. Kunzman, J. Lai, C. Lee, R. McGreevy, C. Mei, M. Nelson, J. Phillips, O. Sarood, A. Shinozaki, D. Tanner, D. Wells, G. Zheng and F. Zhu, https://www.ks.uiuc.edu/Research/namd/2.9/ug/node63.htm[l#], (Accessed June 30th[th] 2022)

51 D. Hamelberg, J. Mongan and J. A. McCammon, *J Chem Phys,* 2004, **120**, 11919–29

52 Y. Miao, S. E. Nichols and J. A. Mccammon, *Phys Chem Chem Phys*, 2014, **16**, 6398-6406

53 Y. Miao, W. Sinko, L. Pierce, D. Bucher, R. C. Walker and J. A. Mccammon, *J Chem Theory Comput*, 2014, **10**, 2677-89

54 W. Sinko, Y. Miao, C. A. F. de Oliveira and J. A. McCammon, *J Phys Chem B*, 2013, **117**, 12759-68

3 Molecular dynamics simulations of copper binding to N-terminus mutants of amyloid-β

Author's note: This chapter is a slightly modified version of the paper "*Molecular dynamics simulations of copper binding to N-terminus mutants of amyloid-β"* published in Journal of Biomolecular Structure and Dynamics (2020), DOI 10.1080/07391102.2020.1745692. This paper was originally written by Mr Oliver Kennedy-Britten, with contributions from Dr Nadiya Al-Shammari who assisted with carrying out analysis on mutant structures, with Dr Jamie Platts assisting with the design of the study and reviewing the final draft of the manuscript.

3.1 Introduction

In Chapter 1, it was discussed how cases of dementia are commonly inferred to be a standard occurrence as part of the ageing process; yet symptoms and diagnosis have also been observed in younger individuals (< 65 years old[1]) due to genetic alterations in the amino acid sequence of Aβ[2]. These instances of AD are referred to as early-onset familial Alzheimer's disease (EOFAD), and account for 5% of AD diagnoses[3].

This segment focuses on 7 known mutations within the N-terminus of Aβ[4,5] (Figure 3.1). Some mutants cause an increase in production of the peptide from its precursor protein[6] (APP) such as E11K[7]. Increased pathogenicity has also been observed in A2V[8] and K16N[9] carriers via recessive and dominant-heterozygous genotypes respectively. Conversely, protective variants show an overall decrease in amyloidogenesis, such as in the case of A2T[10]. Unmutated, wild-type Aβ is hereby referred to as WT.

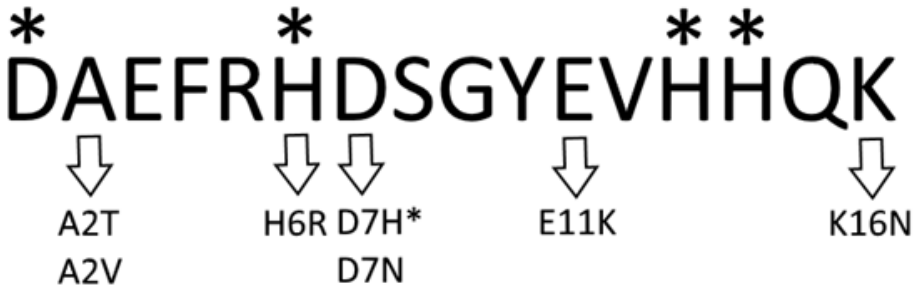


Figure 3.1- Amino acid sequence for N-terminal of $A\beta_{wt}$, highlighting 7 known mutations; residues associated with coordination of Cu(II) are highlighted with an asterisk.

Based off previous literature, the fact the system contains hundreds to thousands of atoms and the flexible nature of Aβ, molecular dynamics (MD) seems a suitable choice for investigating the interactions and structures of this protein when bound with Cu(II), which has additionally been linked with increased formation of oligomeric and fibrous Aβ. In Chapter 1, it was noted that Cu(II) binds within the first 16 residues of Aβ so this N-terminus can be used as a model for the full-length peptide. Similar computational studies have shown N-termini of Aβ peptides to be effective models to make inferences on interactions and structures of full-length Aβ[11,12].

In order to study effects of these mutants when coordinated to Cu(II), and to compare to the unaltered WT, results of molecular dynamics simulations on Cu(II)-bound, truncated mutant and WT peptides noted previously in Figure 3.1 were reported. This allows for comparisons to be drawn on secondary structure and stability. From this, conformations and energies of each mutant system can be sampled, and in doing so, inferences can be made on aggregation behaviour compared against literature.

3.2 Computational methods

Aβ1–16 was constructed in an extended conformation in the program Molecular Orbital Environment, MOE[13], with appropriate protonation states for physiological pH. Selected N-terminal amino acids were then substituted based on the type of mutation being modelled to generate 7 further starting structures for study. Cu was coordinated to the peptides as shown in Figure 3.2 via Asp1, His6, and His13, *i.e.* component I[14]. His14 could have been used in place of His13 but the decision was made to choose one of these residues for consistency across all simulations. The exception to this mode of binding was H6R, which was bound via Asp1, His13 and His14[15] in the absence of His6.

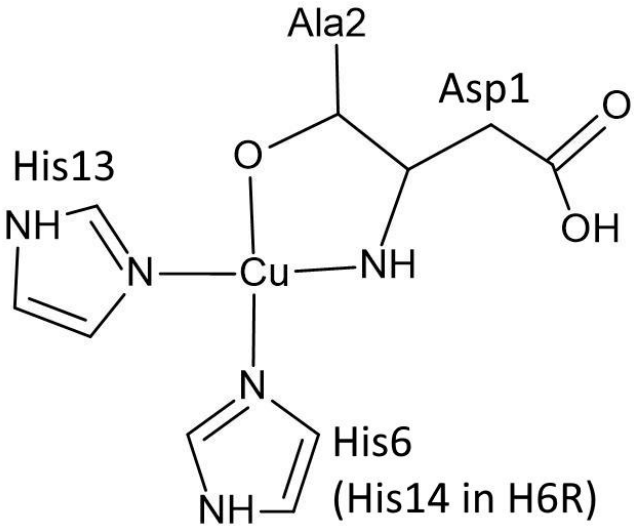


Figure 3.2 Simulated coordination mode of Cu(II) to Aβ modelled across all simulations

All constructed peptides were subjected to brief LowMode[16] conformational searching to obtain starting structures which is a method using atomic velocities and focusing on low-frequency vibrations to sample structures via an MD trajectory. Minimisation of all 8 structures was performed using AMBER94[17] as well as the d-orbital extension of MOE, DommiMOE[18]. MD simulations were performed using the AMBER16[19] package. The AMBER ff14SB[20] forcefield parameter set was used to model all standard amino acid residues, while parameters for the metal and bound residues were obtained using the MCPB.py program[21]. Here, parameters are obtained from B3LYP/6-31G(d), and RESP charges for the metal-coordinating regions were obtained at the same level of theory using Gaussian09[22]. Semi-empirical calculations used the GFN2-XTB method within Grimme's xtb package[23,24], a quantum-mechanical method that was used to deduce binding energies of copper to mutant structures.

The geometry of each system was optimised using 1000 steps of steepest descent to reduce the high energy rapidly and 1000 steps of conjugate gradient methods to get closer to minimum energy once this has fallen. MD simulations were carried out in the NVT ensemble, using a Langevin thermostat[25] to control the temperature at 310 K. Three separate 500 ns MD simulations of each Cu-mutant complex were carried out, starting from the same minimised structure but with different initial velocities, randomly sampled from the Maxwell-Boltzmann distribution at 310 K. Electrostatic interactions were neglected beyond a cut-off of 12 Å, and the Generalised Born solvation model used to solvate all

systems[26-30]: this approach has been shown to enhance conformational sampling of flexible systems in implicit solvent with results comparable to that in explicit solvent[31]. During all simulations, the SHAKE algorithm[32] was used to constrain bonds to hydrogen. Simulations were performed using a 2 fs integration timestep. This was chosen to remove high-frequency bond vibrations and works well with the SHAKE algorithm which constrains the hydrogen bonds. Equilibration times were taken from RMSD data for all simulations, all pre-equilibrated data from the three 500 ns runs were excluded and the rest was combined to form full individual trajectories to be analysed for all eight systems. This led to around 1.4 μs of data collected for each peptide. Cut-off values for each simulation are provided in Table 3.1. Analysis of the trajectories was performed using CPPTRAJ v16.16[33] and VMD 1.9.3[34]. Ramachandran maps were made using MDplot[35] with nomenclature used to describe there from Hollingsworth & Karplus[36].

	A	**B**	**C**
WT	5	10	10
A2T	25	10	10
A2V	35	45	20
H6R	25	25	30
D7H	10	5	70
D7N	5	20	20
E11K	5	15	15
K16N	5	15	5

Table 3.1- RMSD equilibration times (ns), only MD data after these points were analysed

3.3 Results

Root mean square displacement (RMSD) of all backbone atoms relative to starting structure was used as the primary measure of equilibration. Plots of backbone RMSD for each run show that simulations reach stable values after between 5 and 70 ns (shown previously in Table 3.1)

All analysis reported is taken from data extracted from frames after these equilibration points[37], averaged over three separate runs. Once these frames were combined, this led to over 1.4 μs of simulation data collected and analysed for each system. Run C for D7H took

the longest amount of time to equilibrate out of all simulations, before reaching a conformational ensemble similar to the other two runs. Standard deviations (which shows levels of variance in values) and averages of RMSD collected over frames after the selected equilibration point (displayed in Table 3.2) confirm equilibration: averages are in the range 2 to 5 Å with standard deviations between 0.3 and 1.2 Å. K16N stands out in this data as being particularly immobile, having the smallest maximum, mean and standard deviation (sd) from the starting point. H6R also has small standard deviation which shows lower levels of variance compared to other results, although maximum and mean values are larger than for WT and K16N. Most other mutants exhibit similar properties; A2V is the only simulation with a larger standard deviation value than WT, indicating greater flexibility within this mutant.

Table 3.3 reports post-equilibration radius of gyration (R_g) data which is used as a measure of size and can represent how compact or diffuse a molecule is. These results show that on average most mutants are smaller than the wild-type peptide, even in cases where the mutated residue is larger than the one it replaces, such as A2V. E11K has an average value similar to WT: given the larger size of Lys over Glu, this also indicates a more compact set of conformations. K16N is particularly small, in accord with low RMSD values noted above, although some of this change may stem from the smaller size of Gln compared to Lys. Standard deviations are small for all cases, further demonstrating the equilibration of the relevant trajectories. D7N exhibits the most variability as well as the largest average size and sd, but amongst mutants there is no obvious relationship between R_g and RMSD data. For instance, in the average R_g values for A2V suggesting a decrease in size compared to WT despite its RMSD data indicating higher flexibility but a more compact structure being adopted. The result suggests that the relatively large RMSD value for A2V corresponds to motions that do not affect the overall size of the peptide. Full plots for RMSD and R_g data can be seen in Figures 3.3 and 3.4, respectively.

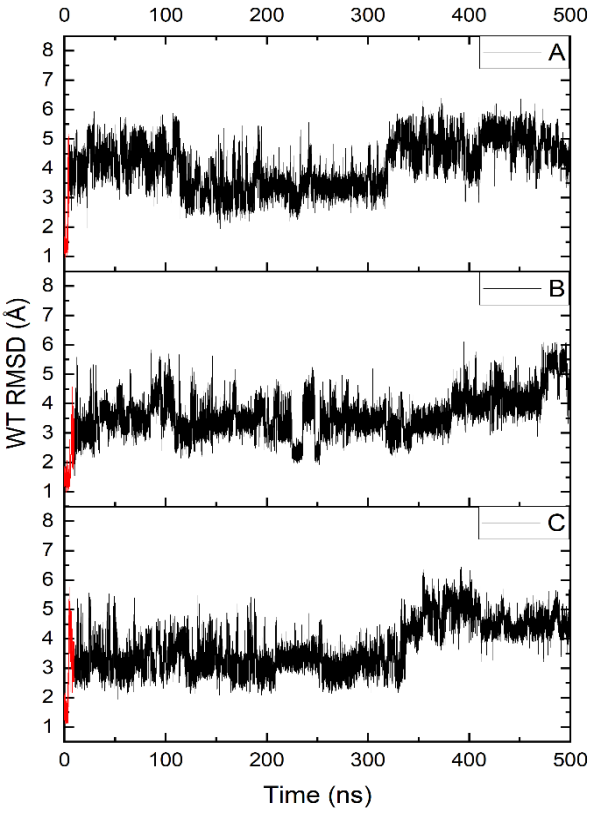
	Avg	Min	Max	SD
WT	3.84	1.54	6.43	0.80
A2T	3.92	2.04	6.17	0.53
A2V	5.09	2.65	7.18	1.23
H6R	4.57	2.81	6.85	0.29
D7H	3.11	1.77	4.02	0.40
D7N	3.87	2.05	6.37	0.70
E11K	4.08	2.23	5.26	0.52
K16N	2.39	1.30	3.42	0.28

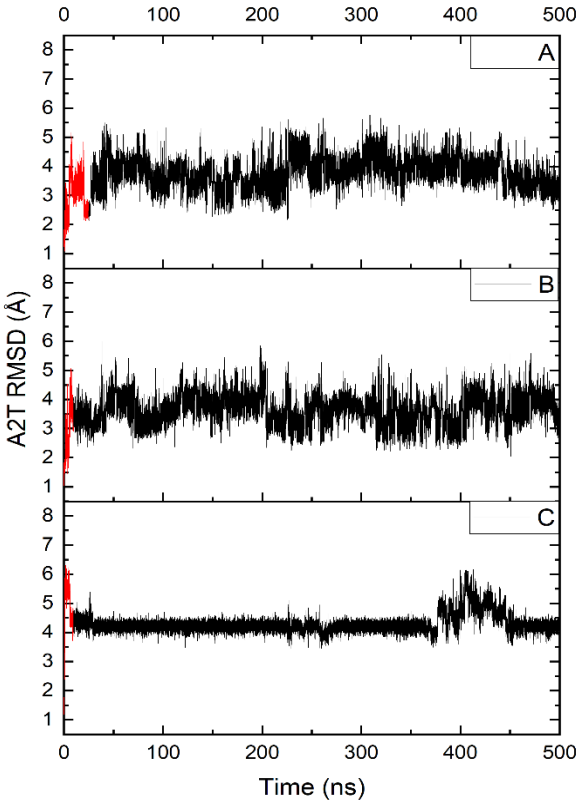
Table 3.2- Statistical analysis of RMSD data (Å)

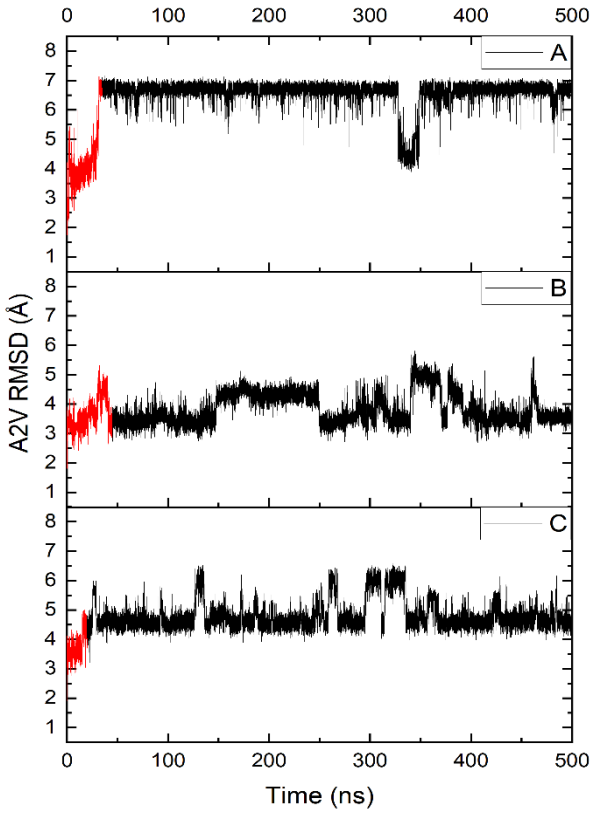
	Avg	Min	Max	SD
WT	7.90	6.77	9.68	0.44
A2T	7.62	6.70	9.66	0.44
A2V	7.37	6.74	8.66	0.24
H6R	7.26	6.75	8.65	0.20
D7H	7.29	6.72	8.21	0.23
D7N	8.14	6.78	9.98	0.49
E11K	7.90	6.90	9.51	0.32
K16N	7.04	6.40	8.74	0.23

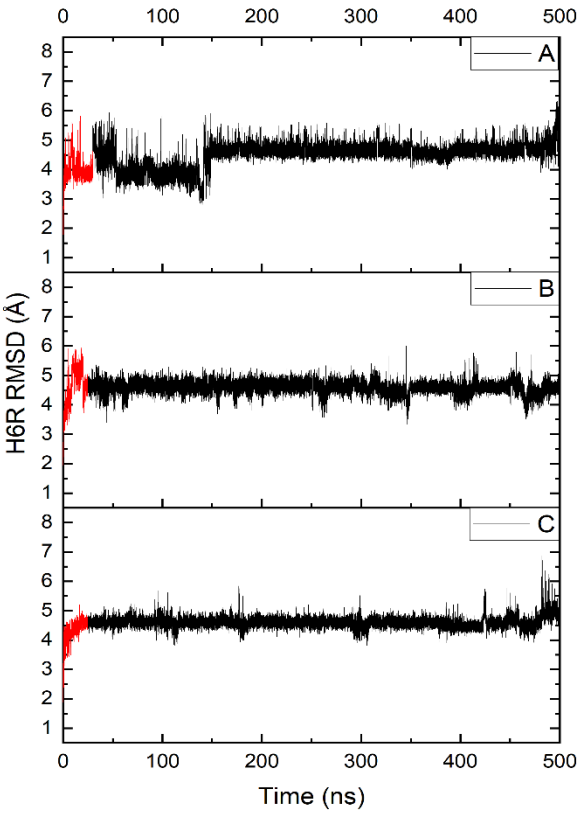
Table 3.3- Post equilibration R_g data (Å)

Solvent accessible surface area (SASA) per residue, calculated over post-equilibration trajectories, are almost identical for analogous residues between mutants, indicating all residue are fully solvent exposed over the course of their trajectories.


WT RMSD (Å)
A
B
C
Time (ns)


A2T RMSD (Å)
A
B
C
Time (ns)


A2V RMSD (Å)
A
B
C
Time (ns)


H6R RMSD (Å)
A
B
C
Time (ns)

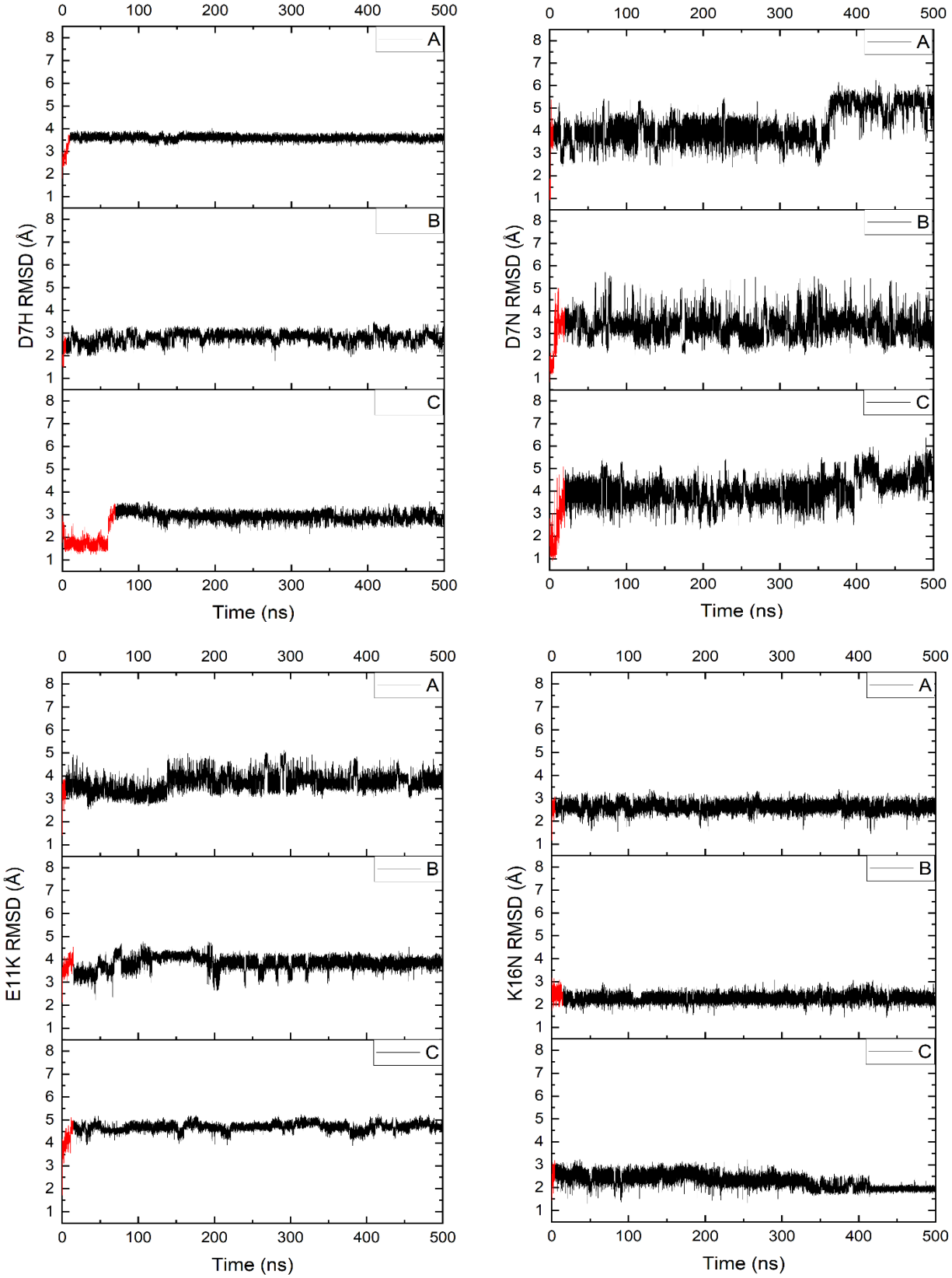
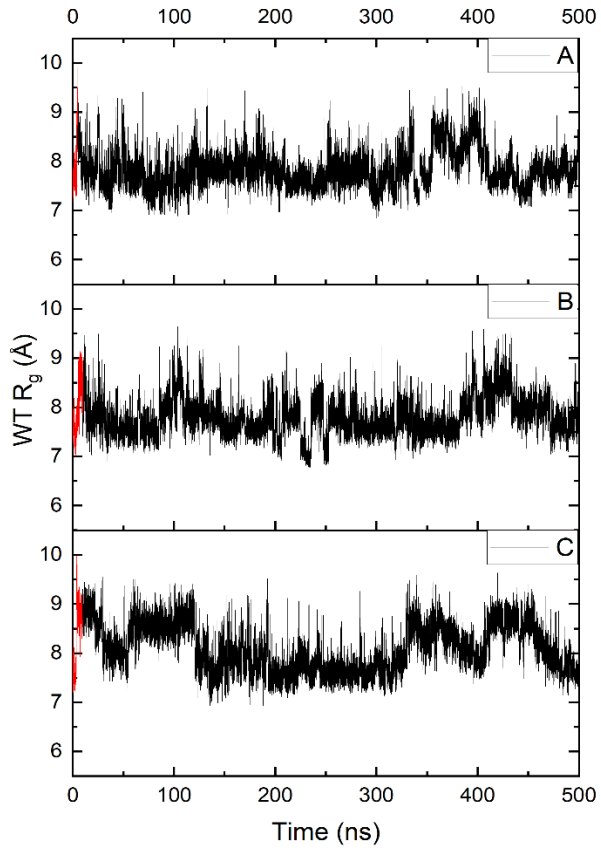
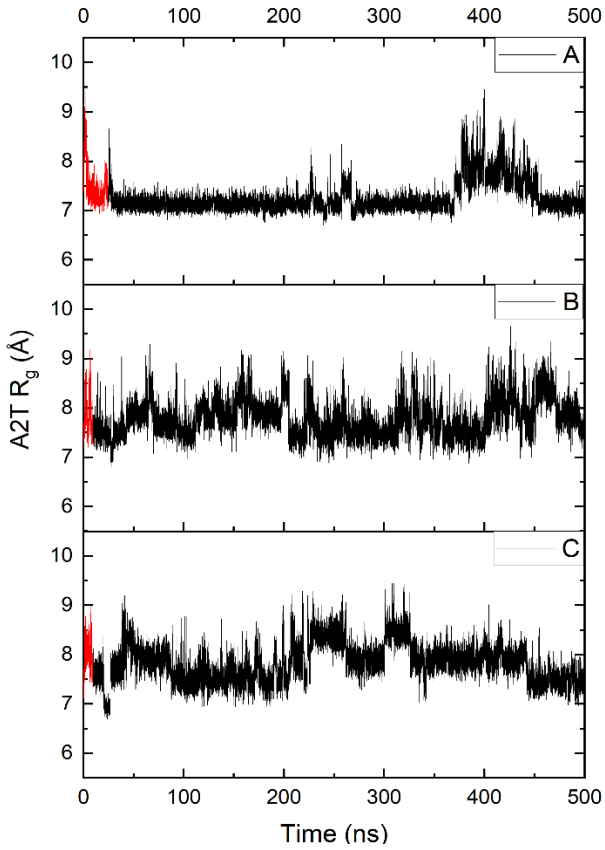
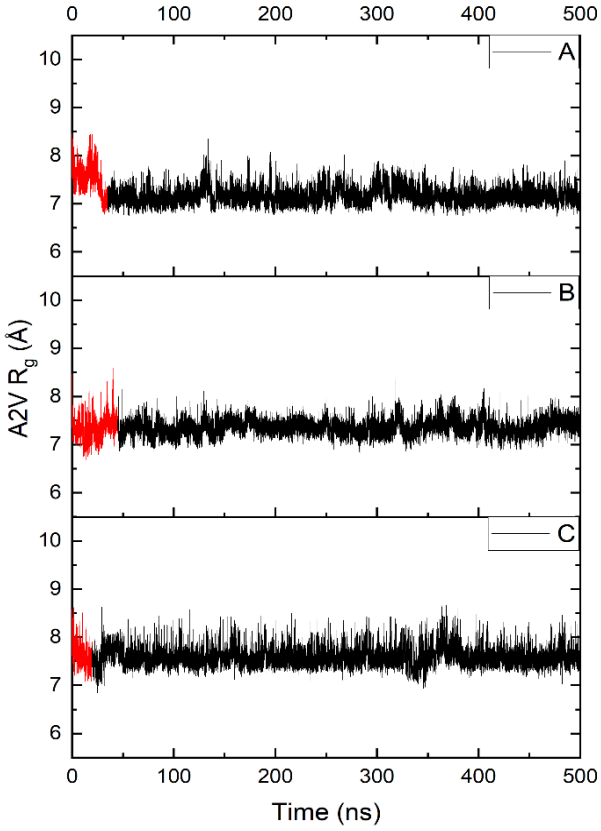
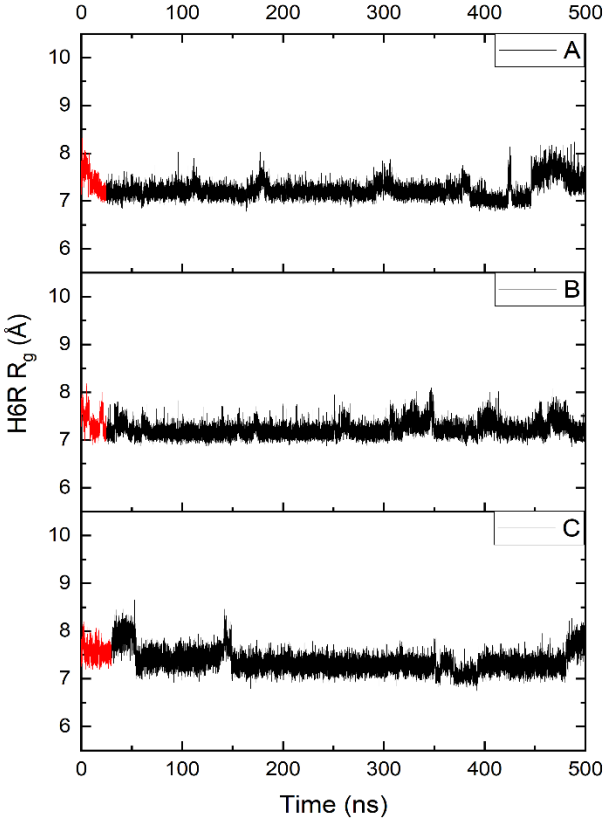


Figure 3.3- RMSD plots for each of the three simulations for all 8 structures across 500 ns. Red denotes pre-equilibration data that was discarded and black is equilibrated data used for subsequent analysis.


WT R_g (Å)
A
B
C
Time (ns)


A2T R_g (Å)
A
B
C
Time (ns)


A2V R_g (Å)
A
B
C
Time (ns)


H6R R_g (Å)
A
B
C
Time (ns)

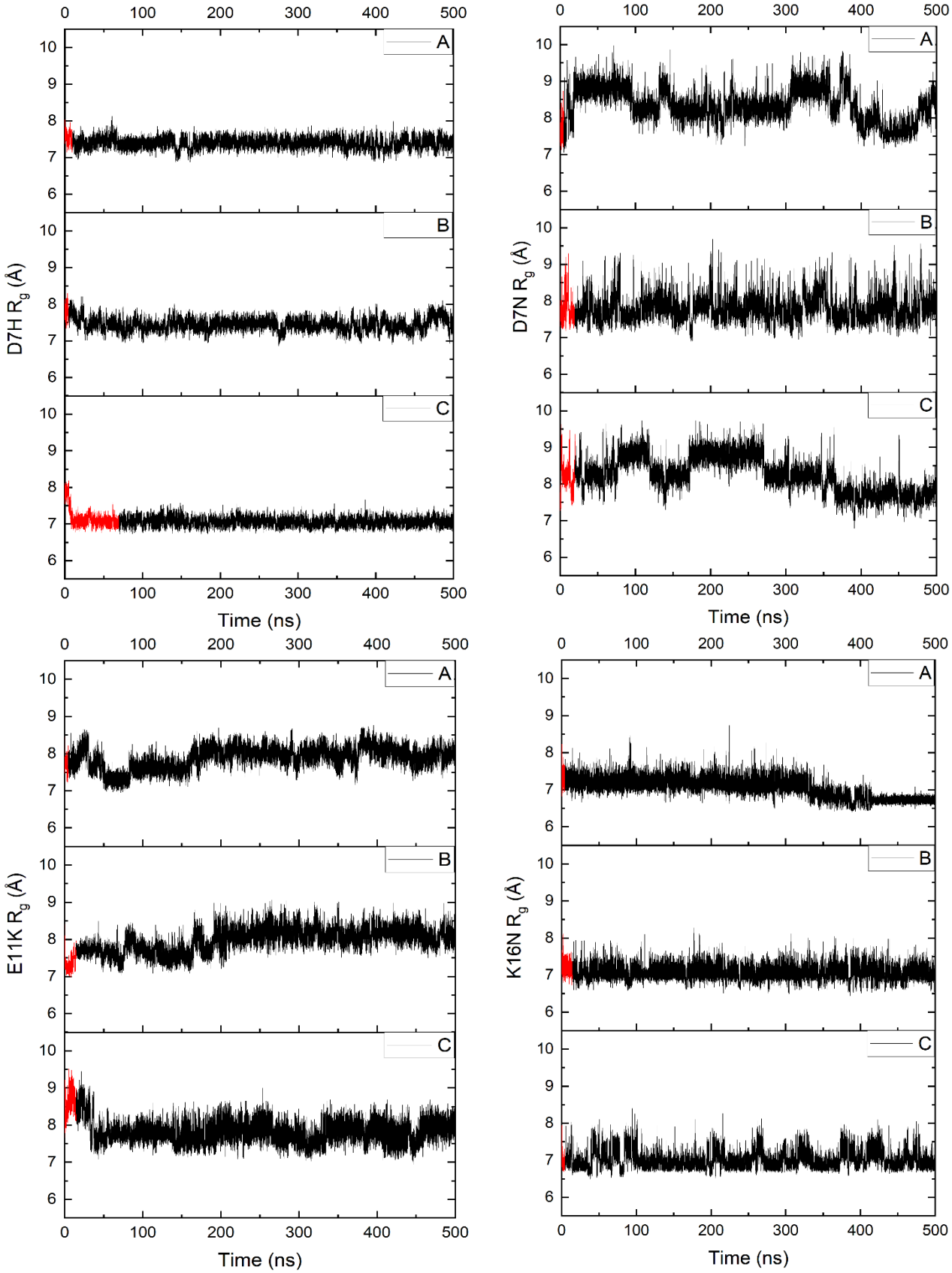
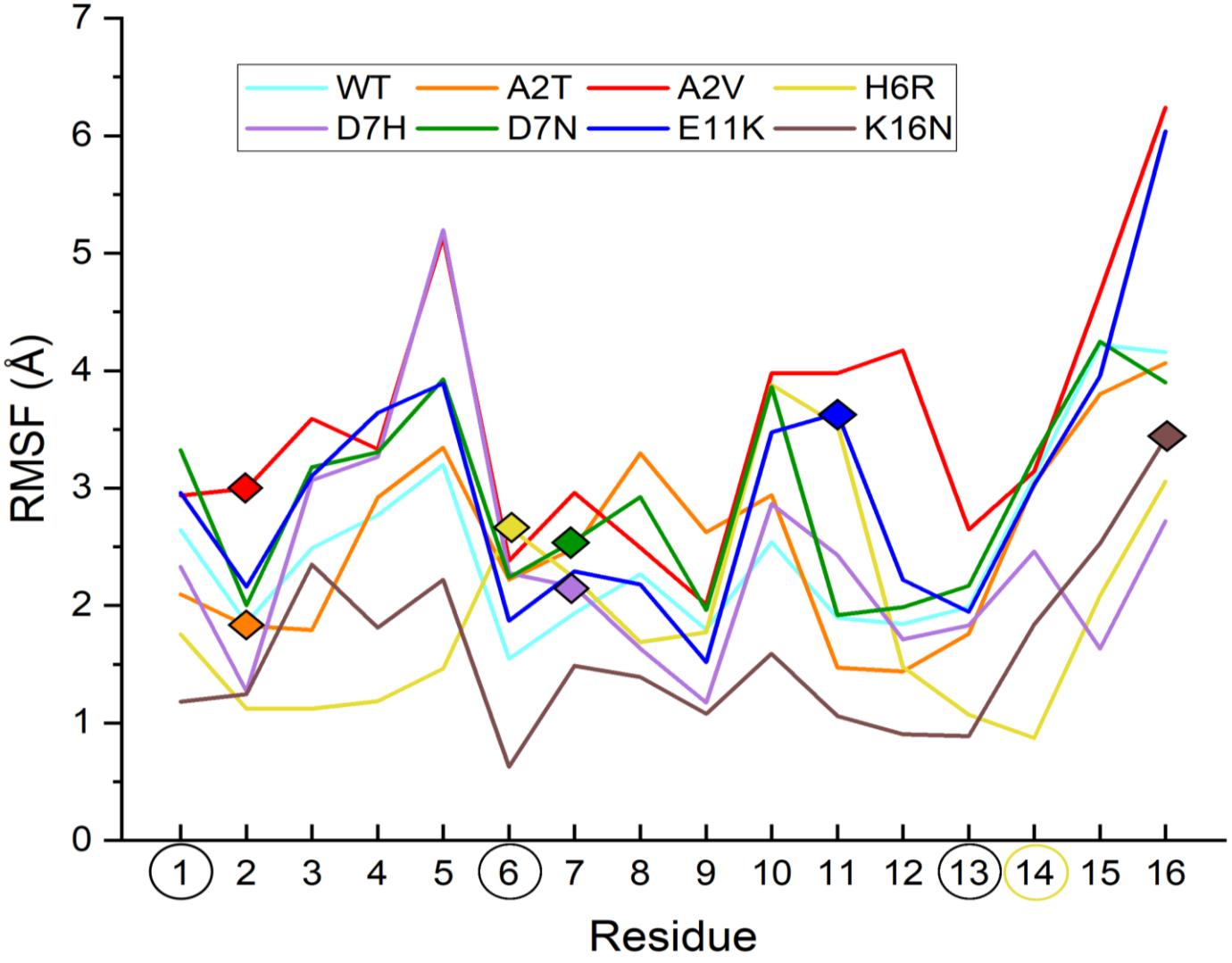


Figure 3.4- Radius of gyration (R_g) plots of the three simulations for all 8 structures across 500 ns. Red denotes pre-equilibration data that was discarded and black is equilibrated data used for subsequent analysis.

Root mean square fluctuation (RMSF) of each residue for all trajectories in each system are reported and illustrated in Figures 3.5 and Table 3.4. Although there is substantial scatter in the data, some trends are apparent. The mutated residues themselves do not stand out as having unusual properties: residue 11 in E11K is flexible, but values are high for residue 11 in other systems. Residues towards the N-terminal are typically less mobile than those closer to the C-terminal, with residue 16 being particularly flexible in all cases. In agreement with RMSD data, K16N has low RMSF values for all residues. Interestingly, H6R values are also rather low, with the exception of residues 10 & 11 and as mentioned before, residue 16. This agrees with low R_g and RMSD data despite H6R having one of the highest average RMSD values. In contrast, numerous residues in A2V have high RMSF values, but these are not located at or even near the mutation; instead, largest values are centred on residues 10-12. Copper-binding residues (1, 6 and 13 for most systems, 1, 13 and 14 for H6R) are among the least mobile, indicating that the metal acts as an "anchor" to bound amino acids. This is especially notable for H6R, suggesting that metal binding to adjacent residues reduces flexibility more than to those that are separated. Bound residues are also notably more rigid in the K16N mutant compared to all other simulations.

It is interesting to note the relative differences in RMSF values between peptides containing mutations at similar positions such as A2T/A2V and D7H/D7N. These proteins with relatively similar amino acid sequences would be expected to display similar RMSF figures, yet there is contrasting data shown between systems for bound residues such as Asp1, which showed variance in RMSF values of 2.33 Å for D7H and 3.33 Å for D7N. As well as this, values differ for both the site of mutation as well as throughout the whole peptide structure displayed in RMSF data for the mutated Ala2 residues of A2T & A2V of 1.83 and 2.99 Å, respectively. In addition, we report a difference of 2.73 Å between values for Val12 and 2.51 Å for Glu11 in these two systems. Reduced incidence of salt-bridges at position Glu11 in A2V allow for increased mobility of residues around this point, as shown from the differing values between the two mutant proteins.



Copper-bound residues are denoted with a black circle (with the exception of H6R which replaces one of its coordinating residues with His14 in the absence of His6 which is highlighted by the gold circle). Mutated residues are marked by a diamond in the corresponding colour.

Figure 3.5 RMSF (Å) plot per residue of the combined trajectories for all 8 structures across approximately 1.4 µs each.

Residue	WT	A2T	A2V	H6R	D7H	D7N	E11K	K16N
1	2.64	2.10	2.94	1.76	2.33	3.33	2.96	1.18
2	1.85	**1.83**	**2.99**	1.12	1.27	2.00	2.16	1.25
3	2.49	1.79	3.59	1.12	3.07	3.18	3.10	2.35
4	2.77	2.92	3.33	1.19	3.27	3.31	3.64	1.81
5	3.20	3.35	5.15	1.46	5.20	3.93	3.89	2.22
6	1.55	2.22	2.38	**2.67**	2.27	2.24	1.87	0.62
7	1.93	2.49	2.96	2.23	**2.16**	**2.55**	2.29	1.49
8	2.27	3.30	2.49	1.69	1.64	2.92	2.18	1.39
9	1.80	2.62	2.01	1.77	1.17	1.96	1.52	1.08
10	2.54	2.94	3.98	3.88	2.86	3.86	3.47	1.59
11	1.90	1.47	3.98	3.54	2.43	1.92	**3.63**	1.06
12	1.84	1.44	4.17	1.48	1.71	1.99	2.22	0.90
13	1.99	1.76	2.65	1.07	1.83	2.17	1.95	0.89
14	3.10	3.04	3.14	0.87	2.46	3.26	3.03	1.84
15	4.22	3.80	4.67	2.08	1.63	4.25	3.96	2.53
16	4.16	4.06	6.24	3.06	2.72	3.90	6.04	**3.42**

Table 3.4- RMSF per residue (Å). Mutated residues in **bold**, Cu-bound residues are underlined.

Clustering further highlights the trends in stability/mobility between mutants: Table 3.5 reports the number of clusters, and the percentage population of the most and second-most prevalent ones calculated from the DBSCAN clustering algorithm in CPPTRAJ[38]. This shows that K16N in particular, but also H6R, fall into a single dominant cluster, reflecting the lack of flexibility and variation in RMSD and R_g discussed above. WT and D7H form a relatively highly populated single cluster, albeit with lower prevalence, while A2T, A2V, D7N and E11K fall into several clusters with smaller populations. Views of a representative snapshot of the most populated cluster for each mutant are shown in Figure 3.6, indicating the change of peptide structure relative to the metal-binding site and site of mutation for all simulations. These snapshots show i) the consistency of the metal binding site and ii) the variability and overall lack of defined secondary structure in any given snapshot. The latter is explored in more detail below.

Mutant	# Clusters	Most populated (%)	Second Most Populated (%)
WT	10	64.1	1.5
A2T	11	45.9	30.2
A2V	9	32.6	31.5
H6R	7	89.9	5.9
D7H	4	63.8	34.6
D7N	12	49.3	24.1
E11K	16	30.3	28.8
K16N	1	99.8	N/A

Table 3.5- Cluster analysis on equilibrated trajectories

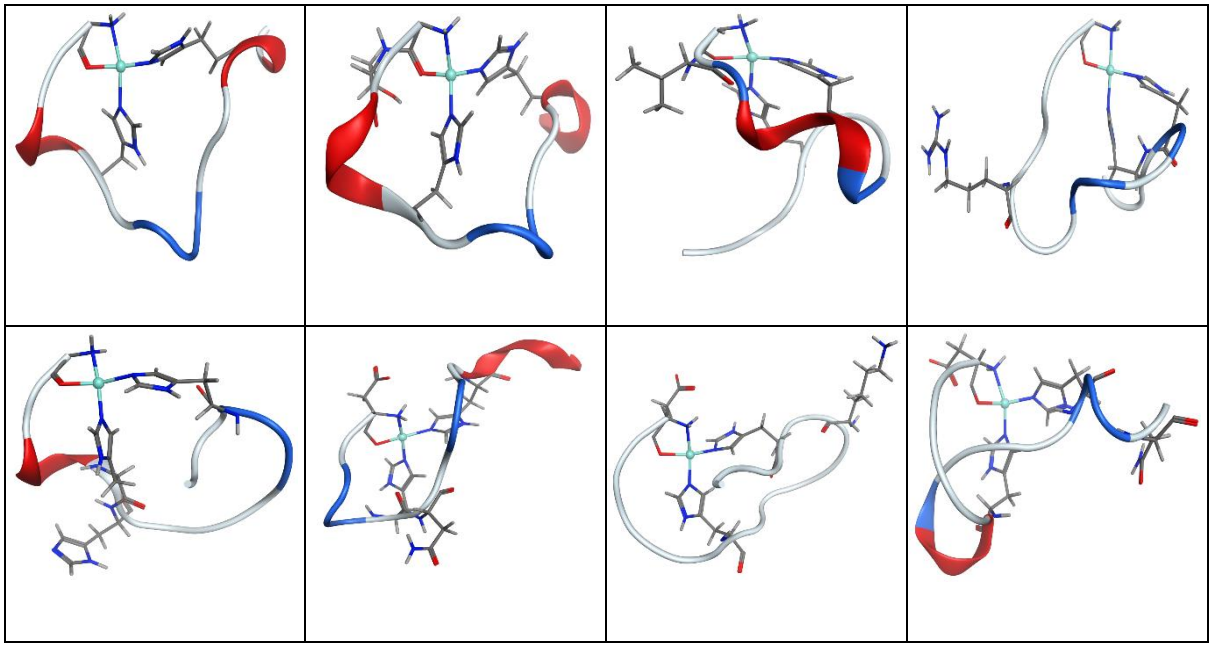


Figure 3.6- Highest populated clusters for each simulation. Top row (L-R) WT, A2T, A2V, H6R; Bottom Row (L-R) D7H, D7N, E11K, K16N. Cu is represented as the teal ball, relevant atoms on coordinated sites as well as mutated residues are shown as wireframe. Protein back bone is characterised by its secondary structure: red = α-helix, blue = turn & white = random coil.

Structural comparisons were made via alpha-Carbons (C_α) of the backbones of the two most-prevalent clusters for all peptides using the UCSF Chimera[39] software tool. The closest RMSD values occurred between the highest populated cluster for WT and A2T, which differ by C_alpha RMSD by 0.970Å. These two clusters also showed high similarity to that of the second most-populated cluster for D7N with a difference in RMSD of 0.374 Å and 0.905 Å respectively; the structures of all 3 are compared in Figure 3.7.

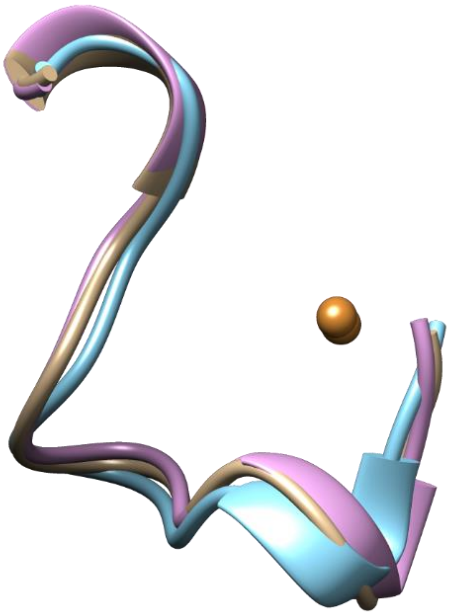
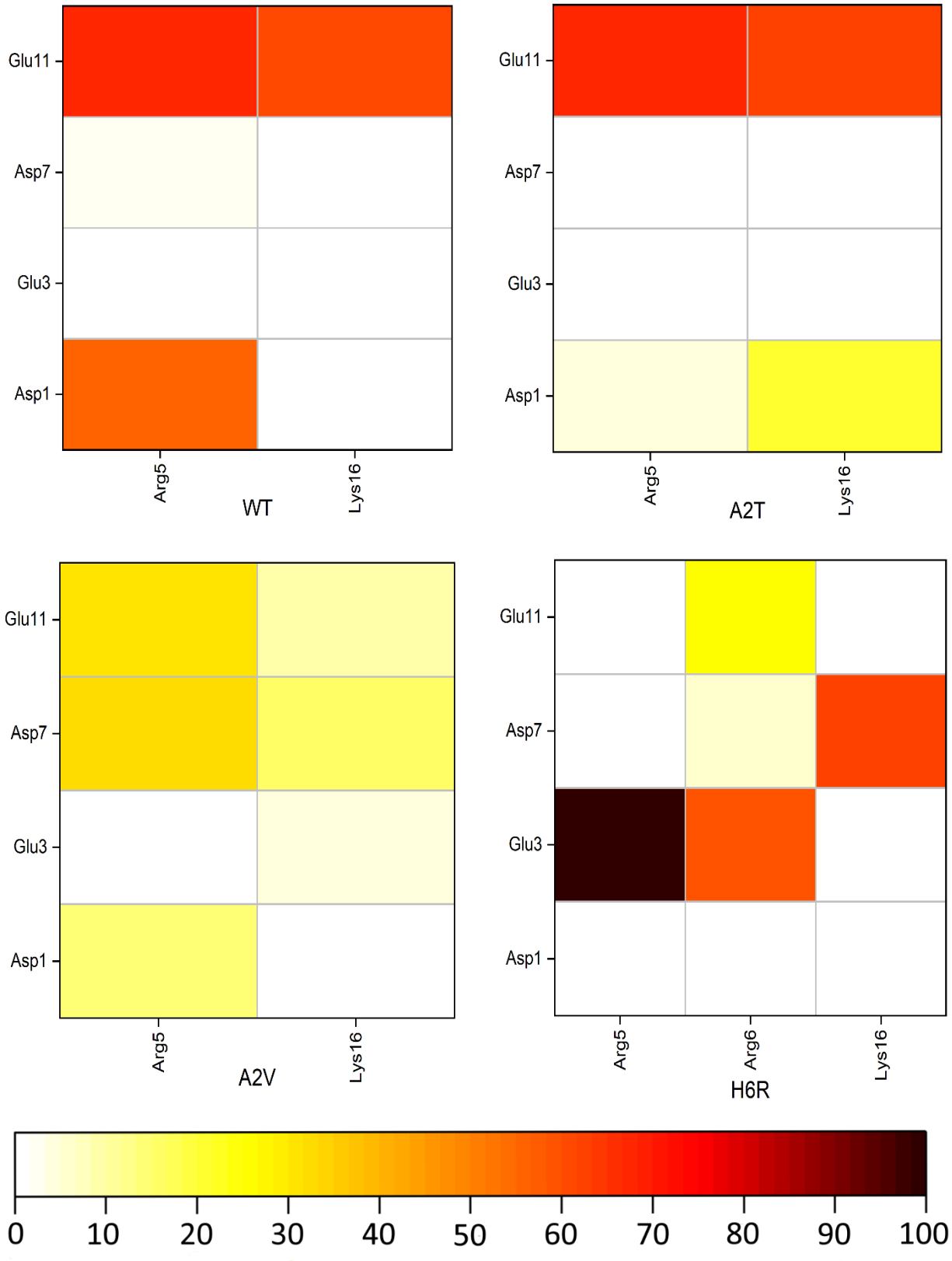


Figure 3.7- Comparison of C_α on peptide backbones of cluster structures showing highest levels of similarity from RMSD data. **Tan**- WT, **Blue**- A2T, **Purple**- D7N, **Orange Sphere-** Cu

Salt-bridges play an important role in peptide structure: the percentage populations of all possible combinations of oppositely charged residues across equilibrated trajectories are displayed in Figure 3.8. The data shows that all mutations have a strong effect on the number and distribution of salt bridges. Compared to WT, the two mutations that leave the number of charged residues unchanged, A2T and A2V, reduce the frequency of Asp1-Arg5 and increase that of Asp7-Lys16 in A2V, while the incidence of Glu11-Arg5 and Glu11-Lys16 is also diminished in A2V but remains consistent in A2T. Salt-bridge profiles between these two mutants show contrasts in types of interaction and frequency, with Glu3-Lys16 and Asp7-Lys16 present in A2V but absent in A2T, whilst Asp1-Lys16 interactions appear only in A2T.


Glu11
Asp7
Glu3
Asp1
Arg5
Lys16
WT
A2T
A2V
Arg6
H6R
0
10
20
30
40
50
60
70
80
90
100

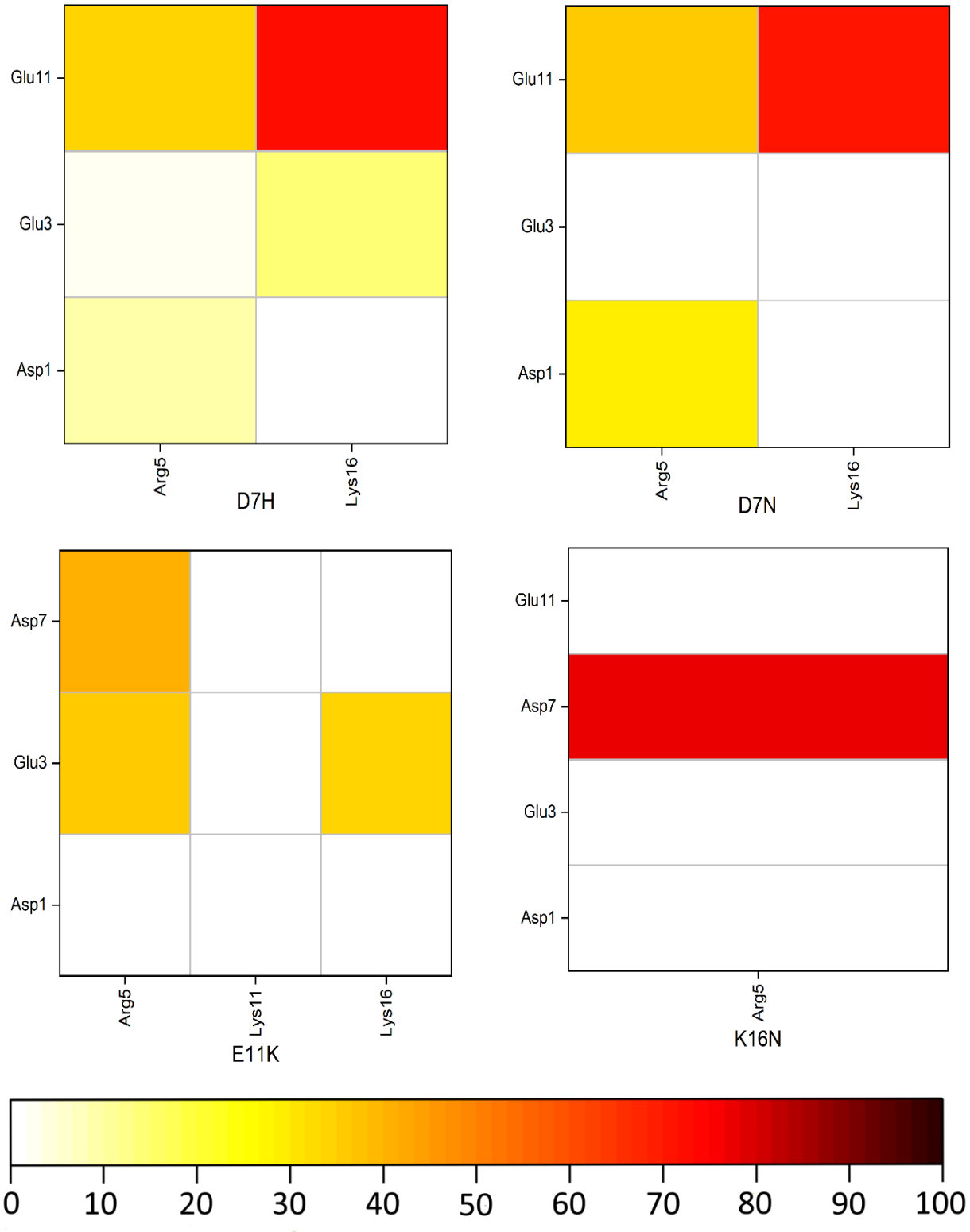


Figure 3.8- Salt bridge plots by percentage of incidence (%) for of the combined trajectories for all 8 structures

H6R introduces an extra positively charged residue, which interacts most commonly with Glu3, but also Glu11 and occasionally Asp7. Glu3 is also found in contact with Arg5 for almost every recorded frame: the proximity of these residues is illustrated in Figure 3.9, showing that Glu3 bridges between the two adjacent positive residues. The presence of Arg6 also acts to remove completely the interactions of Arg5 with both Asp1 and Glu11, and also the Glu11-Lys16 link, that were prevalent in WT.

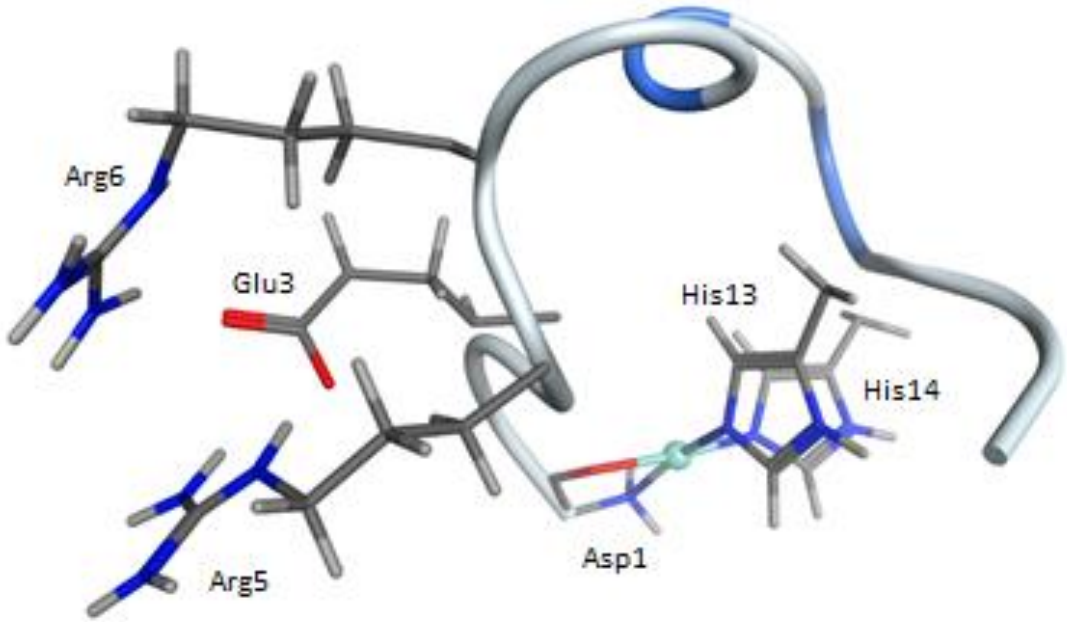


Figure 3.9- View of H6R, with sidechains of Glu3, Arg5 and Arg6, along with metal binding site, shown as wireframe, and the backbone of the remaining peptide as a ribbon

In contrast, D7N and D7H remove a negatively charged residue; however, Asp7 is not heavily involved in salt bridges in WT, such that the pattern of salt bridge population is closest to WT for these mutants. Some changes are still evident, such as a reduction in the interactions of Asp1, with concomitant increase in contacts to Glu3 in D7H. The two mutations of Asp7 show a decrease in the occurrence of Glu11-Arg5 and an increase in Glu11-Lys16 salt-bridges compared to WT, but remain consistent with each other at relatively similar levels of incidence.

E11K swaps the sign of residue 11; the introduced positive sidechain does not engage in any significant interactions. The loss of Glu11 leads to changes in interactions of Arg5 and Lys16, especially with Glu3, and also to the complete loss of interactions of Asp1. K16N removes a

positive residue, leaving only Arg5, which forms a highly populated bridge to Asp7, but no other significantly populated interactions. Across all the simulations, the most commonly observed salt bridge is that between Glu11-Lys16, at 46% of all possible frames, while Asp1-Lys16 is observed for only 3.2% of the full set of equilibrated trajectories.

Contact maps show average distances (Å) between C_α within the peptide backbone per residue (Figure 3.10). These maps reflect the patterns in flexibility noted above; for instance, the least mobile peptides K16N and D7H display large areas of short contact (blue in Figure 3.10), whereas the most flexible ones (W2T, A2T and D7N) exhibit large areas of longer average contacts (orange/red in Figure 3.10). However, the precise pattern of contacts varies: short contacts between Ala2-His6 and Tyr10-Glu11 are present in K16N; while the closest contacts in D7H are between Gln15-Lys16and His6-Tyr10. High incidences of salt-bridges formed, such as with Glu3-Arg5 in E11K, are also seen as short distances contact maps.

Ramachandran plots for all post-equilibration frames, for all mutants, are shown in Figure 3.11. In WT, the highest incidence is found within the α-region, centred on ϕ-ψ = -63°, -43°, followed by P_{II} (-105°, 100° to -30°, 200°) and β (-180°, 90° to -105°, 190°), as well as some δ' character (35°, 60° to 100°, -25°), at similar levels to those found in the β region. P_{II} exists in the region similar to the β-region but describes residues that typically don't adopt β-sheets. A2T, A2V, and H6R have similar Ramachandran maps, showing increased population of P_{II} and reduced of α, whilst maintaining relatively similar levels of β character to WT. D7H differs from the others, as both the P_{II} and α regions are equally populated, and also as the only plot to possess a significant amount of character within the δ region (-30°, -65° to -135°, 40°).

D7N differs from D7H showing less P_{II} character than its Asp7 counterpart as well as an increase in δ' making it the most comparable plot to WT. E11K has similar level of P_{II} and α, with less β character than others considered. This plot also has the most incidence of conformations with positive ϕ, which for non-glycine residues is usually an indicator of steric hindrance. K16N is broadly similar to WT, with most residues located within the α-region, but this mutant lacks any P_{II} character, with greater population of the β-region, albeit spread out over a broader distribution then seen in the other plots.

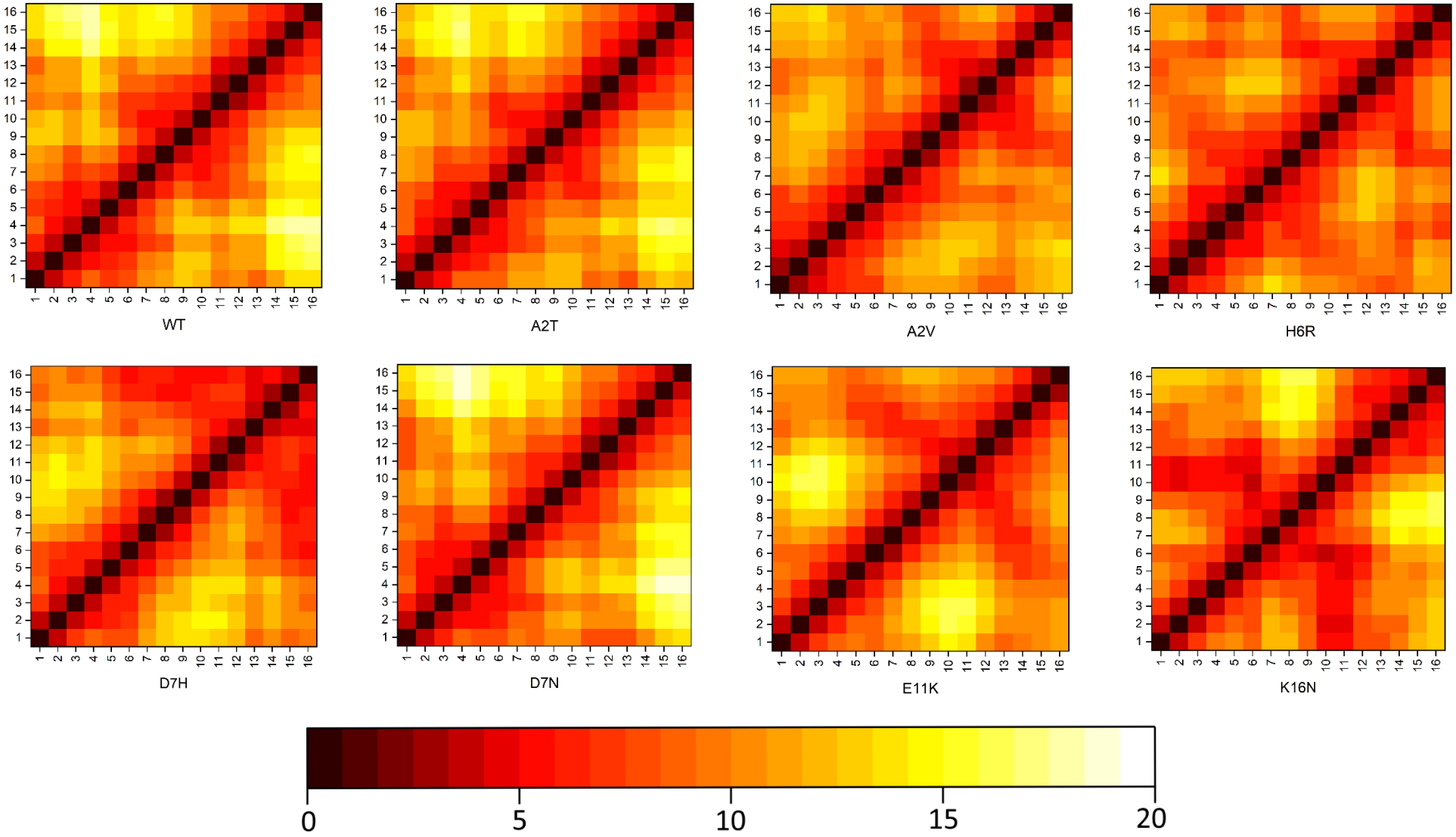
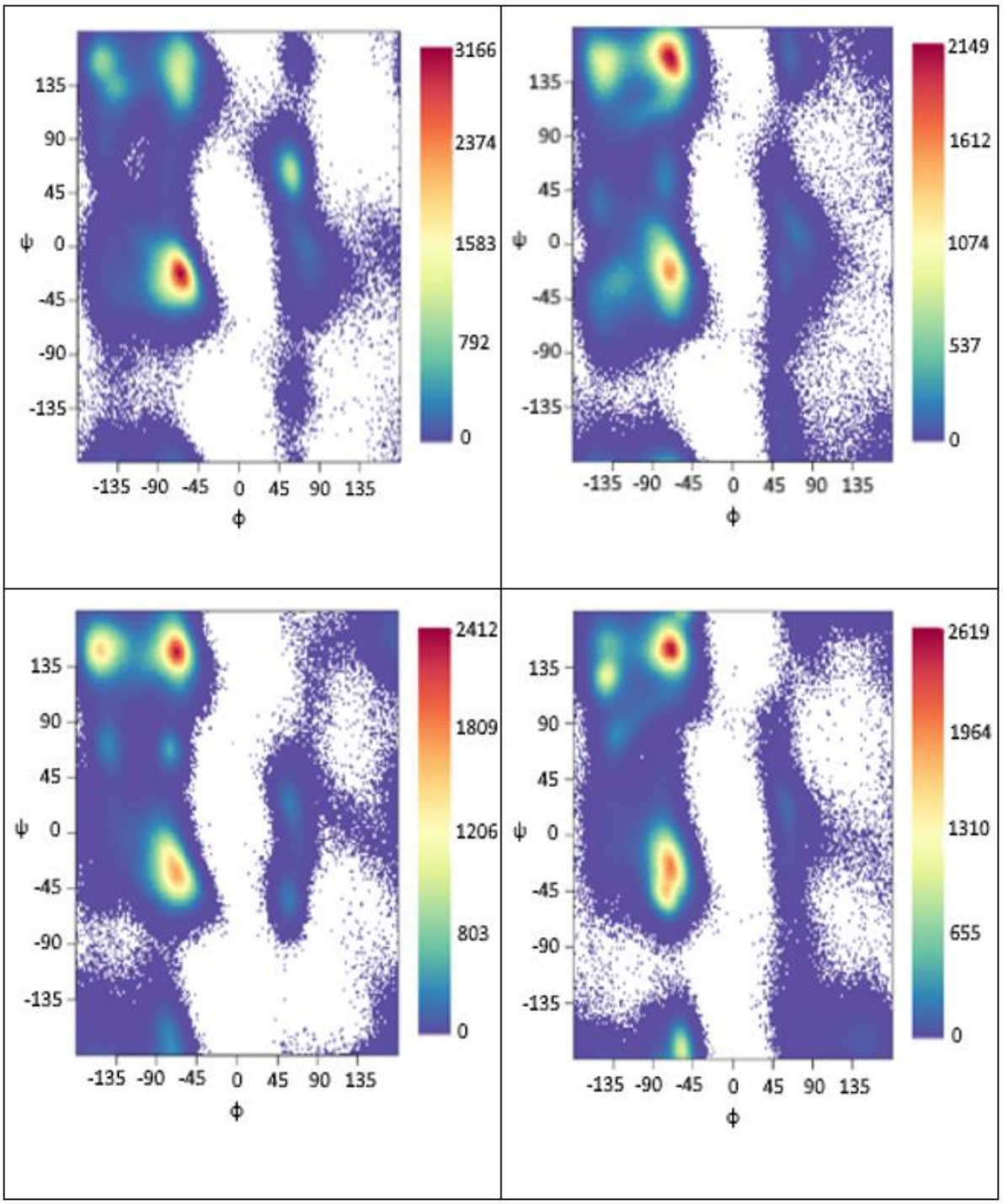


Figure 3.10- Contact map of average distance between C_α(Å) for of the combined trajectories for all 8 structures.


135
90
45
ψ 0
-45
-90
-135
-135 -90 -45 0 45 90 135
ϕ
3166
2374
1583
792
0
2149
1612
1074
537
0
2412
1809
1206
803
0
2619
1964
1310
655
0

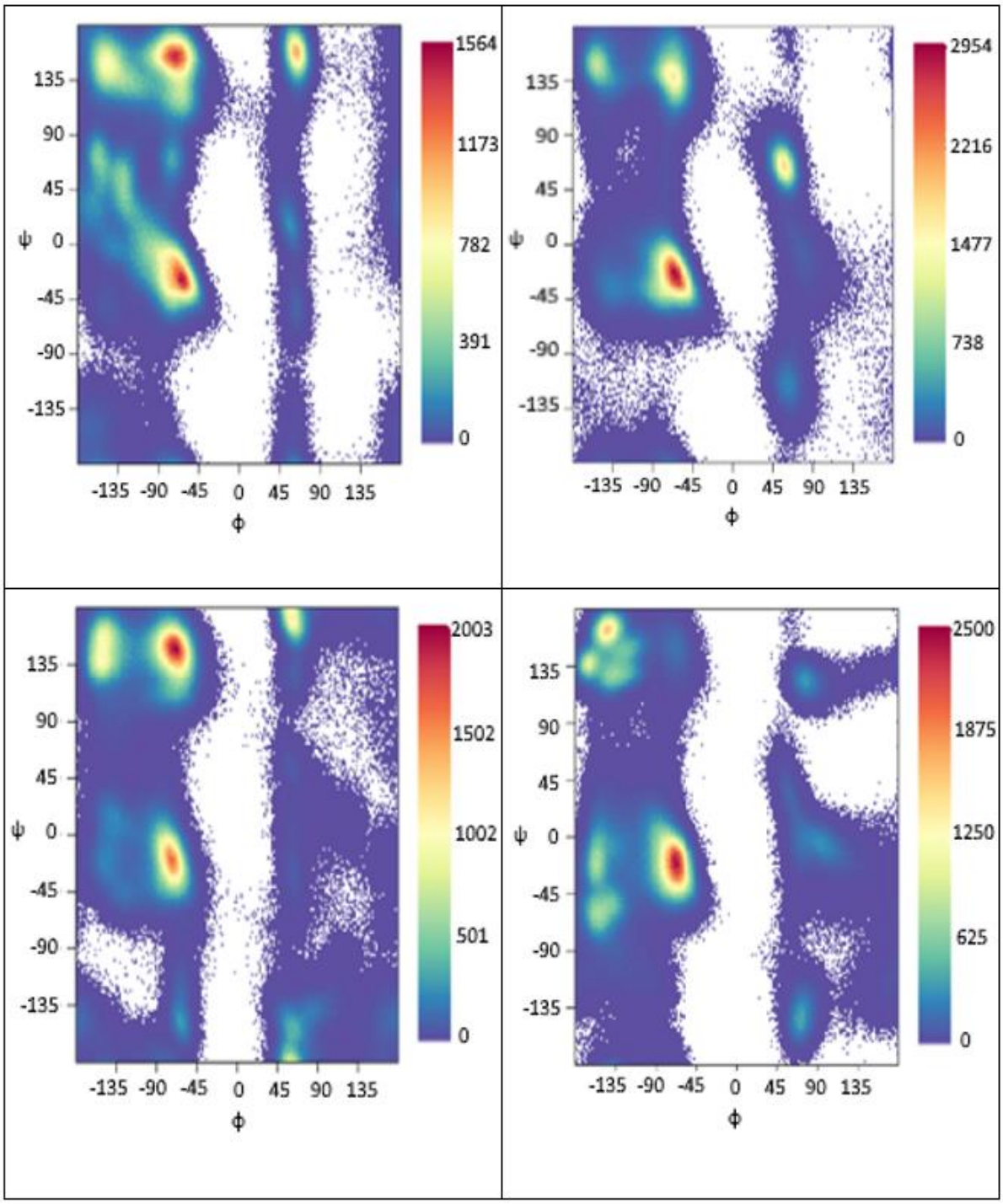


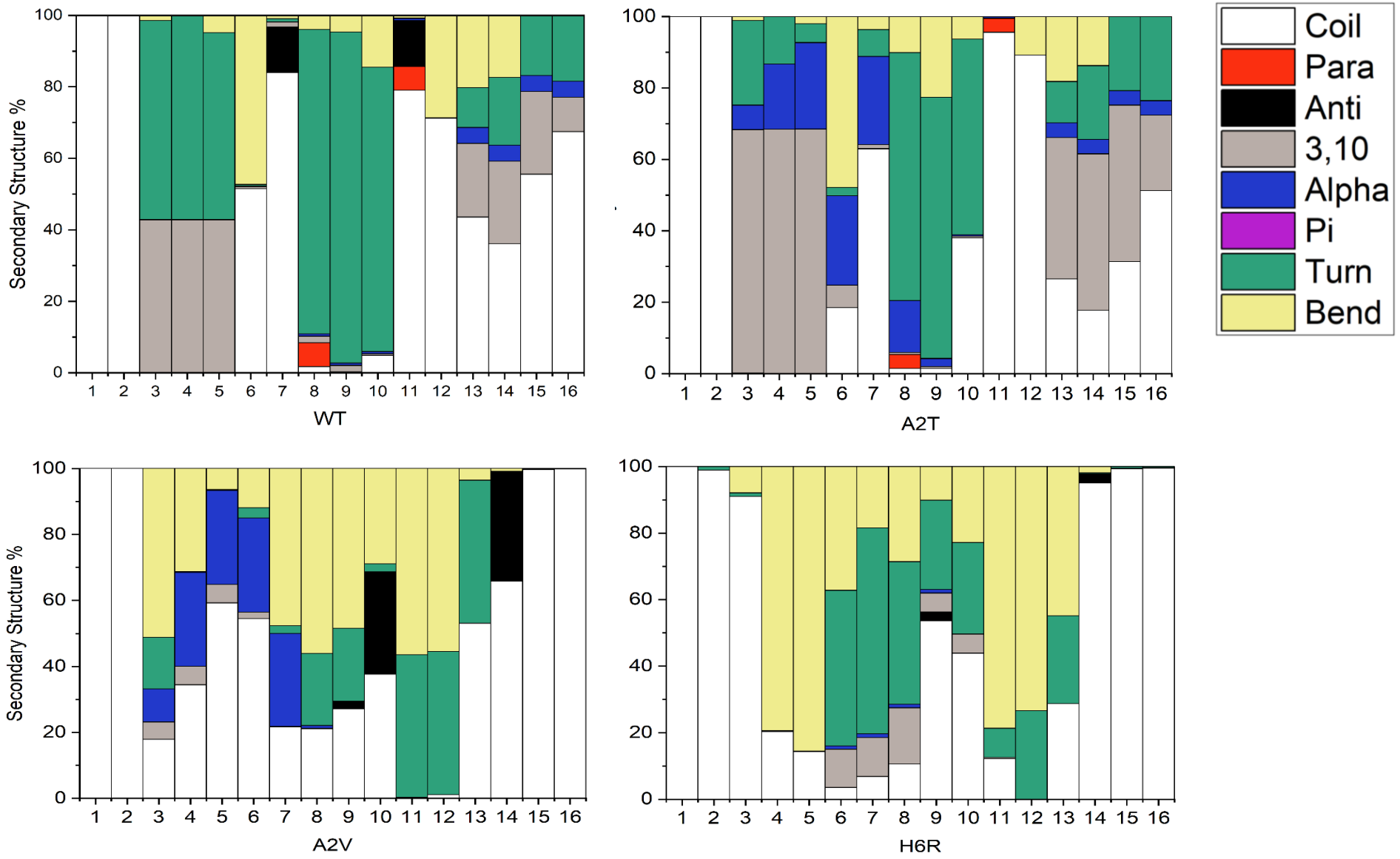
Figure 3.11- Full Ramachandran plots for equilibrated trajectories including angles (°) and levels of incidence by frame count. **1st row (L-R);** WT, A2T. **2nd row (L-R);** A2V, H6R. **3rd row (L-R);** D7H, D7N. **4th row (L-R);** E11K, K16N

The effect of mutations on secondary structure are marked and varied, as shown in Figure 3.12 and Table 3.6. WT is characterised by a large amount of coil (at termini) and turn (residues 3-5 and 8-10), along with 3,10 helices and some β-strand, but almost no α-helix. This is apparently at odds with the Ramachandran plot for WT above, which shows high concentration of frames lying in the region associated with α-helical structure (which as explains before surround the apex of this region around ϕ-ψ = -63°, -43°). It is, however, in accord with Hollingsworth and Karplus's finding that residues that are not classified as helical are still found in this region of ϕ-ψ space and are perhaps better thought of as belonging to "an extended δ-region".

A2T shows the greatest increase in helical character over the whole sequence, whilst losing some of its β-character, whereas A2V only displays α-helix across residues 3-7 as well as an increase in β-strand content. H6R exhibits very little helix or strand structure, being dominated by coil/turn/bend structure with only small elements of 3,10-helix and strand located mainly between residues 6-10.

D7H shares similarities to WT, albeit with greater proportions of β-strand. D7N and D7H are closest in resemblance to each other in terms of α-content despite notable variances in salt-bridge profiles suggesting a difference in structures. In D7H, the presence of helical geometry is limited to residues 3-6 whilst D7N displays this between 3-5 but also towards the C-termini between residues 13-16. The percentage of strand character differs between these two mutant systems with D7H displaying β-characteristics over a larger range of residues than D7N.

E11K displays almost no helix content and predominantly forms coil/turn/bend formations making up the predominant character of its secondary structure. K16N leads to strand content closer to the N-terminus and helices at the C-terminus at incidence levels comparable to that of WT. No clear pattern of changes in the mutated residues themselves is found, such that changes to secondary structure are global rather than local.


Secondary Structure %
100
80
60
40
20
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
WT
A2T
A2V
H6R
Coil
Para
Anti
3,10
Alpha
Pi
Turn
Bend

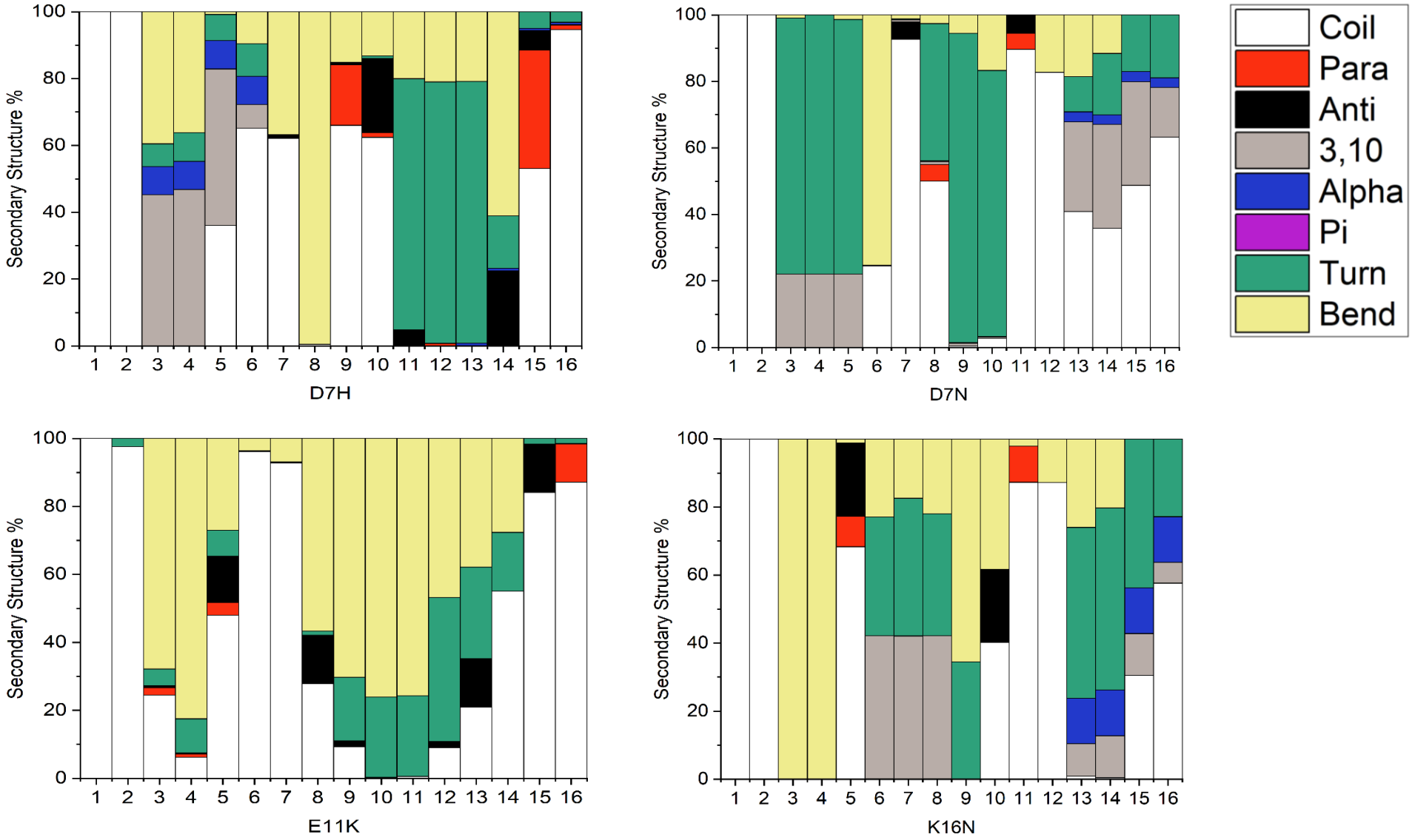


Figure 3.12- Secondary structure percentages by residue (%) for of the combined trajectories for all 8 structures

	Helix	Strand	Other
WT	14.5	2.4	83.1
A2T	31.0	0.5	68.5
A2V	8.9	4.2	86.9
H6R	3.6	0.3	96.1
D7H	11.4	7.2	81.4
D7N	11.6	1.3	87.1
E11K	<0.1%	4.9	95.0
K16N	13.8	3.9	82.3

Table 3.6- Percentage of residues classified as helical, strand, or other

Average binding energies of Cu(II) to each peptide were calculated using the semi-empirical GFN2-xTB method. Structures were taken every 100 ns from all 8 equilibrated trajectories of approximately 1.4 μs and minimised to account for any vibrations potentially occurring as a structure is sampled as is standard procedure for this method. Calculation of the total energy of Cu-peptide complex and 4 H_2O to replicate the binding of Aβ were compared to the free peptide and $[Cu(H_2O)4]^{2+}$, all in implicit model of aqueous solvent. No major conformational changes were observed following optimisation, as shown by comparable RMSD values of approximately 2 Å difference for all systems between structures generated in AMBER and minimised structures from xTB.

Most Cu binding energies are typically in the range of -70 to -110 kJ mol^{-1}, indicating that most peptides considered have similar affinities for binding with copper ions. H6R is one of the closest in binding energy to WT, despite possessing an extra positively charged residue and a different mode of bonding from all other simulations. However, Cu(II) is more strongly bound to WT as per the greatest binding energy value of all structures analysed, whereas binding to K16N is markedly weaker than all other systems with a binding energy of -30 kJ mol^{-1}. All mutant peptides had a lower average difference in binding energy compared to the WT, showing weaker binding to the metal centre. Relatively high standard deviations indicate a wide range of values for binding energies for all mutant simulations, indicating a

high level of variability across trajectories. Average binding energies are displayed with standard deviations in Figure 3.13.

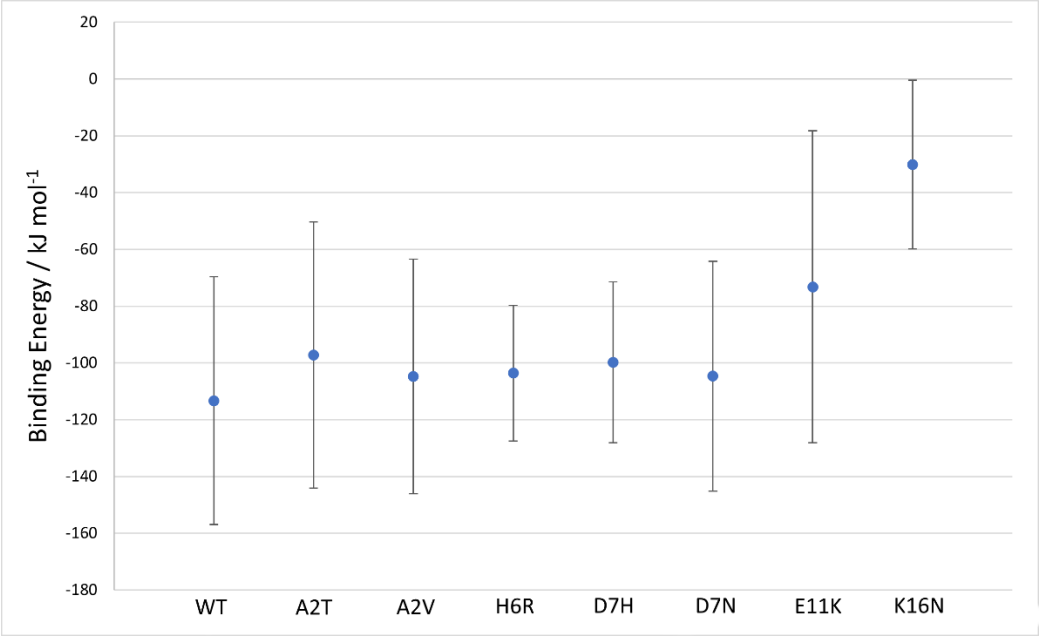


Figure 3.13- α-binding energies with standard deviations for each mutant

3.4 Discussion

MD simulations of Cu-complexes of N-terminal mutants yielded evidence that the effects of point mutations of Aβ vary significantly, and depend on the site of mutation as well as the specific amino acids involved. Marked differences were observed between mutants in secondary structure, conformations adopted, and flexibility/stability.

Within the mutations that do not alter charge state, *i.e.* A2V and A2T, greater flexibility was observed in A2V compared to WT, despite it adopting more compact conformations, as shown by R_g data. In conjunction with RMSF results, the greatest contribution to this mobility occurs after Tyr10, but increased movement was noted across the entire structure. A2T, however, displayed less flexibility than both A2V and WT, in addition to adopting conformations comparable in size to WT, shown via similar R_g values. Additionally, A2T showed less mobility in its RMSF values after Tyr10, in direct contrast to A2V, due to a notable decrease in formation of salt-bridges with Glu11 in the latter. Trajectories for these mutants fall into a similar number of clusters as the WT, albeit with more evenly distributed

population, further confirming their dynamic nature. As charge is unchanged and sequences relatively similar, similar salt-bridge profiles were expected to be generated. Instead, the incidence and combinations of these attractive forces displayed notable differences from the WT protein and each other, with only Glu11 interactions remaining consistent between WT and A2T. Several salt-bridges formed in the A2T and A2V simulations were rather transient, further demonstrating the constant fluctuation of atoms and residues in these systems over the course of the simulations.

H6R and K16N are similar in terms of structure and stability. RMSD and R_g data show they adopt more rigid and compact conformations compared to WT, with lower RMSF values per residue, demonstrating the stability of these mutated peptides. Further evidence of the rigidity and size is exhibited in cluster analysis; both adopt a single prevalent structure over the course of their simulations. Addition of a positive residue in the H6R peptide was expected to yield an increase in salt-bridge formation at Arg6, which was indeed found; however, some salt-bridges present in WT are lost completely, while new ones are observed remotely from the site of mutation. In contrast, K16N loses a positive residue and thus has reduced potential for salt-bridges, evident in the fact that only one such interaction forms for an appreciable time, suggesting that the stability of this mutant cannot be accounted for by these forces. The Ramachandran plot lacks P_{II} character, present in all other simulations, but shows increased presence of organised α and β-character, which may be the origin of the relative stability.

D7H and D7N mutants have contrasting properties, unanticipated for two different mutations at the same site, including RMSD and R_g compared to WT, indicating they possess different structures; D7N is more comparable to WT than D7H, which has a much lower RMSD and R_g, indicating it possesses a more rigid, compact set of conformations. RMSF data clearly shows the differences between these peptides: values for residues in D7N are consistently higher than WT across the whole peptide; whereas D7H has values generally similar to or much lower than WT, especially in residues towards the C-terminus. These differences are also evident in clustering data. D7H populates fewer conformations than D7N, which is more comparable to the A2T/A2V systems, occupying a greater number of more sparsely populated clusters. These differences between systems are also seen in Ramachandran plots: D7H has an even distribution between P_{II} and α regions, as well as

significant γ-character, whereas D7N has a Ramachandran plot comparable to WT with similar population of P_{II}, α, and even δ', consistent with the unmutated peptide. Mutation of Asp7 was not expected to strongly affect salt-bridges, as this residue is barely involved in these interactions in WT; however, significant differences in nature and incidence of interactions were observed. Glu11 interactions were consistent with one another but at different levels to WT; additionally, no other salt-bridges formed at significant levels in these two mutant systems.

E11K adopts a less rigid conformation than the other peptides simulated. Average R_g for this system is the same as WT, but the increased size of Lys means that E11K adopts a more compact structure than the WT overall. Despite this, the RMSF data is greater for all residues in E11K compared to WT, indicating greater flexibility. Additionally, this mutant forms the highest number of clusters. This particular variant substitutes a positive residue for a negative one increasing the possible number and types of salt-bridge combinations possible, but the new Lys forms no significant salt-bridge interactions, and also eliminates all salt-bridges of Asp1 that were present in WT. The Ramachandran plot for this structure displays a significant population in the positive ϕ side, usually indicative of steric hindrance, which could be the origin of the flexibility of this mutant.

Overall, a decrease in helical character was observed in all simulations compared to WT except A2T: some mutants lose nearly all helical structure, such as E11K, whereas formation of helices increased in A2T by more than double. Increased β-character was recorded in most simulations, with the exceptions of A2T, H6R, and D7N which saw a decrease or similar levels of β-strand structures compared to WT. Previous studies have shown a link between β-character and enhanced aggregational properties[40]. The lack of β-strands in simulation data for A2T provides some evidence for the protective nature of this mutant[41]. H6R, D7H, and D7N have been shown in similar MD simulations[42] to form an increased level of β-character which held true for the D7H simulation data we report. It can be expected that the monomeric forms of mutant species reported here may not generate results indicative of those observed in oligomeric species associated with enhanced neurotoxicity and aggregation[43], however, the results generated for the truncated species seem to be in agreement with similar MD experiments on larger Aβ peptides showing these are good models for full-length Aβ and even species containing multiple monomers.

D7H and D7N showed differences in their secondary structure profiles, in keeping with their contrasting results in other analyses despite being so similar in terms of amino acid sequence. K16N was seen to be most similar to the unaltered WT peptide in terms of percentage incidence of secondary structure type despite differences seen in other analyses such as salt-bridge profiles and cluster data. It was observed that the effects of mutations on peptide structure is global as opposed to local, as evident in from the varied salt-bridge profiles for each mutant, with changes in interactions and structures remote to the site of mutation in all analysis. Ramachandran plots and secondary structure analysis show distinct differences and similarities between systems and highlight the contrasts in structure between comparable mutations at a similar location such as A2T/A2V and D7H/D7N.

3.5 Conclusions

In this study, MD was used to explore the effects of N-terminus genetic mutations on the truncated Aβ 1-16 peptide when bound with Cu(II), with the aim of finding differences between mutants and drawing comparisons of these variants with the WT peptide. Literature data indicates varying effects on pathogenicity and structure between mutants. All mutants varied in terms of rigidity and size, as seen in RMSD and R_g data comparable to WT, as well as one another, showing these mutations have differing effects on morphology of the Aβ peptide. Some notable conformational changes were observed between each system from this data, in conjunction with cluster analysis showing varying degrees of mobility.

Results for different mutations at similar locations were independent of each other, showing some similarities as well as distinctions between systems, such as D7H compared to D7N. From this it can be ascertained that both the location and type of mutation that alters the structure of the peptide. Salt-bridge data was markedly varied between simulations and in conjunction with Ramachandran plots showing different profiles for each mutant, showed that the changes are not local to the site of mutation, and that overall structure should be considered when comparing such systems.

In nearly all cases, the levels of helical character decreased in comparison to WT, forming either more β-character or coil/turn/bend. Previous studies have shown an increase in β-

sheets have been the driving force for aggregation. It is interesting to note that the decrease in β-character observed in A2T agrees with the previously reported protective nature of this mutation. E11K and A2V display an increase in β-strand like structures, which was expected due to their reported pathogenicity. Despite simulations only being performed on the truncated, monomeric peptides, the MD results reported indicate that these are effective models of the full-length Aβ and inferences of the effects of these genetic mutations can still be made from this data. The impact of mutations on aggregation properties could be explored further by using MD to further model peptides dimers or oligomers in systems that closer replicate those *in vivo*.

3.6 References

1 A. F. Teich and O. Arancio, *Biochem J*, 2012, **446**, 165–177.

2 B. Jiao, B. Tang, X. Liu, J. Xu, Y. Wang, L. Zhou, F. Zhang, X. Yan, Y. Zhou and L. Shen, *Neurobiol Aging*, 2014, **35**, 1957.e1-1957.e6.

3 C. Ballard, S. Gauthier, A. Corbett, C. Brayne, D. Aarsland and E. Jones, *Lancet*, 2011, **377**, 1019–1031.

4 S. Weggen and D. Beher, *Alzheimers Res Ther*, 2012, **4**, 1-14

5 K. P. Kepp, *Coordin Chem Rev*, 2017, **351**, 127–159.

6 D. C. Ryman, N. Acosta-Baena, P. S. Aisen, T. Bird, A. Danek, N. C. Fox, A. Goate, P. Frommelt, B. Ghetti, J. B. S. Langbaum, F. Lopera, R. Martins, C. L. Masters, R. P. Mayeux, E. McDade, S. Moreno, E. M. Reiman, J. M. Ringman, S. Salloway, P. R. Schofield, R. Sperling, P. N. Tariot, C. Xiong, J. C. Morris and R. J. Bateman, *Neurology*, 2014, **83**, 253-260

7 L. Zhou, N. Brouwers, I. Benilova, A. Vandersteen, M. Mercken, K. Van Laere, P. Van Damme, D. Demedts, F. Van Leuven, K. Sleegers, K. Broersen, C. Van Broeckhoven, R. Vandenberghe and B. De Strooper, *EMBO Mol Med*, 2011, **3**, 291–302.

8 G. Di Fede, M. Catania, M. Morbin, G. Rossi, S. Suardi, G. Mazzoleni, M. Merlin, A. R. Giovagnoli, S. Prioni, A. Erbetta, C. Falcone, M. Gobbi, L. Colombo, A. Bastone, M. Beeg, C. Manzoni, B. Francescucci, A. Spagnoli, L. Cantu, E. Del Favero, E. Levy, M. Salmona and F. Tagliavini, *Science*, 2009, **323**, 1473–1477.

9 D. Kaden, A. Harmeier, C. Weise, L. M. Munter, V. Althoff, B. R. Rost, P. W. Hildebrand, D. Schmitz, M. Schaefer, R. Lurz, S. Skodda, R. Yamamoto, S. Arlt, U. Finckh and G. Multhaup, *EMBO Mol Med*, 2012, **4**, 647–659.

10 B. Murray, M. Sorci, J. Rosenthal, J. Lippens, D. Isaacson, P. Das, D. Fabris, S. Li and G. Belfort, *Proteins*, 2016, **84**, 488–500.

11 J. Alí-Torres, J. D. Maréchal, L. Rodríguez-Santiago and M. Sodupe, *J Am Chem Soc*, 2011, **133**, 15008–15014.

12 S. Furlan, C. Hureau, P. Faller and G. La Penna, *J Phys Chem B*, 2010, **114**, 15119–15133.

13 *Molecular Operating Environment (MOE),* 2013, 1010 Sherbrooke St. West, Suite #910, Chemical Computing Group ULC, Montreal, QC, Canada, H3A 2R7

14 E. Atrián-Blasco, M. Del Barrio, P. Faller and C. Hureau, *Anal Chem*, 2018, **90**, 5909–5915.

15 B. Alies, H. Eury, C. Bijani, L. Rechignat, P. Faller and C. Hureau, *Inorg Chem*, 2011, **50**, 11192–11201.

16 P. Labute, *J Chem Inf Model*, 2010, **50**, 792–800.

17 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J Am Chem Soc*, 1995, **117**, 5179–5197.

18 R. J. Deeth, N. Fey and B. Williams-Hubbard, *J Comput Chem*, 2005, **26**, 123–130.

19 D.A. Case, R.M. Betz, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao and P.A. Kollman, AMBER 2016, 2016, University of California, San Francisco.

20 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J Chem Theory Comput*, 2015, **11**, 3696–3713.

21 P. Li and K. M. Merz, *J Chem Inf Model*, 2016, **56**, 599–604.

22 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G.

Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian Inc., 2016, Wallingford CT

23 C. Bannwarth, S. Ehlert and S. Grimme, *J Chem Theory Comput*, 2019, **15**, 1652–1671.

24 GitHub (2019), from https://github.com/grimme-lab/xtb/, (Accessed June 30th 2022)

25 J. A. Izaguirre, D. P. Catarello, J. M. Wozniak and R. D. Skeel, *J Chem Phys*, 2001, **114**, 2090–2098.

26 R. Constanciel and R. Contreras, *Theor Chim Acta*, 1984, **65**, 1–11.

27 W. Clark Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J Am Chem Soc*, 1990, **112**, 6127–9.

28 M. Schaefer and M. Karplus, *J Phys Chem*, 1996, **100**, 1578–1599.

29 A. Okur, L. Wickstrom, M. Layten, R. Geney, K. Song, V. Hornak and C. Simmerling, *J Chem Theory Comput*, 2006, **2**, 420–433.

30 R. Geney, M. Layten, R. Gomperts, V. Hornak and C. Simmerling, *J Chem Theory Comput*, 2006, **2**, 115–127.

31 J. Nasica-Labouze, P. H. Nguyen, F. Sterpone, O. Berthoumieu, N. V. Buchete, S. Coté, A. De Simone, A. J. Doig, P. Faller, A. Garcia, A. Laio, M. S. Li, S. Melchionna, N. Mousseau, Y. Mu, A. Paravastu, S. Pasquali, D. J. Rosenman, B. Strodel, B. Tarus, J. H. Viles, T. Zhang, C. Wang and P. Derreumaux, *Chem Rev*, 2015, **115**, 3518–3563.

32 J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J Comput Phys*, 1977, **23**, 327–341.

33 D. R. Roe and T. E. Cheatham, *J Chem Theory Comput*, 2013, **9**, 3084–3095.

34 W. Humphrey, A. Dalke and K. Schuten, *J Mol Graphics*, 1996, **14**, 33–38.

35 C. Margreitter and C. Oostenbrink*, Europe PMC Funders Group*, 2017, **9**, 164–186.

36 S. A. Hollingsworth and P. A. Karplus, *Biomol Concepts*, 2010, **1**, 271–283.

37 P. D. Q. Huy, Q. Van Vuong, G. La Penna, P. Faller and M. S. Li, *ACS Chem. Neurosci*, 2016, **7**, 1348–1363.

38 M. Ester, H. P. Kriegel, J. Sander· and X. Xu, *Proc 2nd Int Conf Knowl Discov Data Min*, 1996, **7**, 226–231.

39 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J Comput Chem*, 2004, **25**, 1605–1612.

40 G. G. Glenner and C. W. Wong, *Biochem Bioph Res Co*, 1984, **120**, 885–890.

41 J. A. Maloney, T. Bainbridge, A. Gustafson, S. Zhang, R. Kyauk, P. Steiner, M. Van Der Brug, Y. Liu, J. A. Ernst, R. J. Watts and J. K. Atwal, *J Biol Chem*, 2014, **289**, 30990–31000.

42 L. Xu, Y. Chen and X. Wang, *Proteins*, 2014, **82**, 3286–3297.

43 K. Ono, M. M. Condron and D. B. Teplow, *J Biol Chem*, 2010, **285**, 23186–23197

4 Accelerated Molecular Dynamics to Explore the Binding of Transition Metals to Aβ

Author's note: This chapter is a slightly modified version of the paper written by Mr Oliver Kennedy-Britten; "*Accelerated Molecular Dynamics to Explore the Binding of Transition Metals to Amyloid-β"* published in ACS Chemical Neuroscience (2021), DOI 10.1021/acschemneuro.1c00466. This paper was originally written with contributions from Dr Nadiya Al-Shammari who ran simulations and carried out analysis on iron-bound structures with contributions from Dr Jamie Platts in designing the study and reviewing the final draft of the manuscript.

4.1 Introduction

In this work, accelerated molecular dynamics (aMD) simulations on metal bound Aβ peptides of varying lengths were reported. aMD was selected for enhanced conformational sampling as well as being an appropriate method to simulate relatively large biomolecules, in this instance full-length $A\beta_{42}$ containing over 600 atoms [1]. In addition, investigation via use of implicit solvent derived from the Generalised Born solvation model[2,3] allows better sampling and transitions between minima structures in comparison to conventional MD. Via the use of a boost potential, from tailored parameters based on dihedral and potential energy of the system under investigation (Table 4.1), aMD overcomes energy barriers that would limit conversion between minima that may be inaccessible by conventional MD[4]. Subsequent analysis performed on equilibrated trajectories yield insight and comparative analysis into secondary structure, size, and intramolecular interactions.

Four chain lengths were selected; Aβ16 (to simulate effects solely within the metal binding region), Aβ40 & 42 (to examine properties of the full-length peptide and the aggregational properties of them mentioned previously), and Aβ28 (an intermediate length chain that has been used extensively experimentally and computationally[5-8]). For all four of these sequences, Cu(II), Fe(II), or Zn(II) were bound within the N-terminus. As well as these, free peptides were modelled with no metal centre. Thus, 16 structures were analysed in total.

4.2 Computational methods

Four chain lengths (16, 28, 40, & 42 residues) of Aβ were manually constructed in extended conformations in MOE[9] in protonation states appropriate for physiological pH. All four peptide lengths were modelled as both a free peptide as well as versions bound to copper, zinc, or iron, creating sixteen structures in total. Cu(II) ions were coordinated via Asp1, His6 and His13; Zn(II) ions were bonded to Asp1, His6, Glu11 and His14; whereas Fe(II) were bound to Asp1, His6 and His14 based off of noted binding modes in literature for these particular metal ions[10-13] (Figure 4.1) . All structures were minimised with combination of AMBER94[14] and LFMM in DommiMOE[15], resulting in structures best characterised as random coil. From this point onwards, specific peptides with a metal centre will be referred to as Aβ[Chain Length]-Metal and any free peptides will be described as Aβ[Chain Length]-Free.

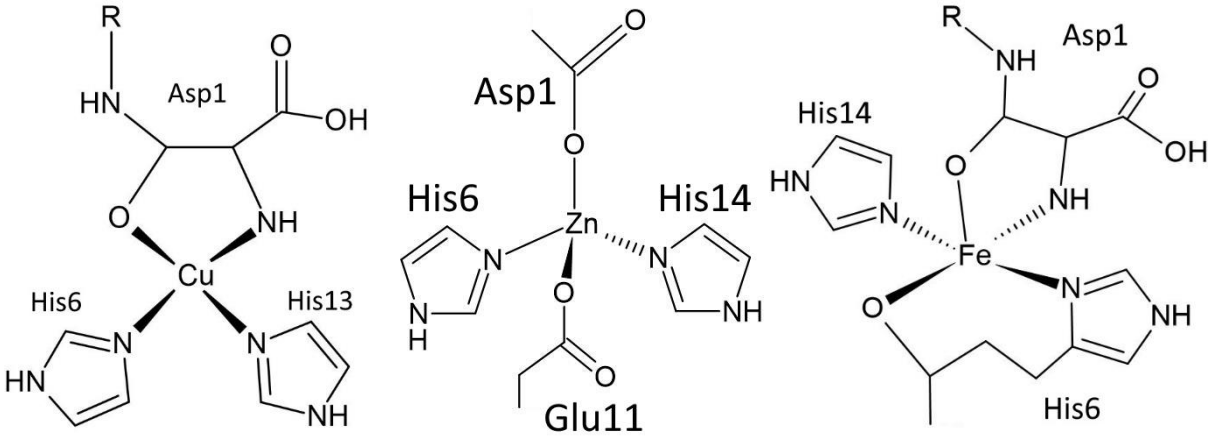


Figure 4.1- Binding modes of Aβ with Cu(II), Zn(II) and Fe(II)

All conventional and accelerated molecular dynamics (aMD) simulations were carried out using the AMBER16[16] package. Standard amino acids were modelled using the AMBER ff14SB[17] forcefield parameter set and the LEaP[18] module of AMBER, whereas parameters for metal-centres and any bound residues were generated via MCPB.py[19]. Bonded and non-bonded parameters for metal binding were extracted from B3LYP/6-31G(d) calculations performed using Gaussian09[20]. Harmonic bond length, angle and torsion parameters were extracted from the DFT Hessian matrix for this molecule using the Seminario[21] method, while non-bonded parameters were drawn from DFT electrostatic potential using the RESP

method[22] which assigns partial charges throughout the molecule. All simulations were performed in implicit solvent using the Generalised Born solvation model[23-25]. Previous work within the research group shows comparable results between explicit and implicit solvent simulations on Aβ-Zn(II) structures[26], with no improvement in results with explicit solvent. Implicit solvent was therefore used as it allows for more rapid simulations and conversion between conformations. Structures were first minimised by 1000 steps of steepest descent, followed by 1000 conjugate gradient steps. Implicit solvent has been shown to be a suitable solvent model for peptides of this nature without requirement for explicit solvent[27,28] and use of Generalised Born solvation model has been shown to enhance conformational sampling of mobile systems[29].

All simulations were performed in the NVT ensemble, in which temperature was regulated to remain at 310 K via use of a Langevin thermostat[30]. For all sixteen structures, three 50 ns conventional MD simulations were performed: each started from the same minimised structure but assigned a different set of random initial velocities sampled from the Maxwell-Boltzmann distribution. Subsequently, three independent 200 ns of aMD were performed, using the 16 final structures of each MD equilibration simulation as a starting point; thus generating 48 individual simulations. Parameters for the boost potential in aMD were based on the size of the systems and the associated potential & dihedral energy from the prior MD calculations[31,32]. These are explained further via calculations and values for the boost used in aMD calculations in Table 4.1. From this, a bias potential was applied to boost the whole system simultaneously at points in the Potential Enegy Surface (PES) where the energy barrier would be too high to overcome using conventional MD.

The SHAKE algorithm[33] was implemented to ensure any bonds to hydrogens were maintained through the course of the simulation, and electrostatic interactions beyond 12 Å were disregarded. The initial 50 ns was removed to account for equilibration and the three 200 ns aMD data was combined for each structure (from the 48 aMD simulations in total) to form 600 ns trajectories for each of the individual sixteen systems. Analysis of these trajectories were carried out using CPPTRAJ v16.16[34] and VMD 1.9.3.[35] Free energy plots were constructed by re-weighting the boost potential from aMD following the procedures in literature[36,37].

Table 4.1- aMD boost parameter equations and values used for simulations

[EPTot] = Average total potential energy of conventional MD system / **[Dihed]** = Average dihedral energy of conventional MD system

iamd = 1 (Implements boost across the entire potential at once)

AlphaP = Inverse strength boost factor for total potential energy = **0.16 x [Number of Atoms]**

AlphaD = Inverse strength boost factor for dihedral energy of system = **0.16 x (3.5 x [Number of Residues])**

EthreshP = Average total potential energy threshold = **[EPTot] + [AlphaP]**

EthreshD = Average dihedral energy threshold = **[Dihed] + (3.5 x [Number of Residues])**

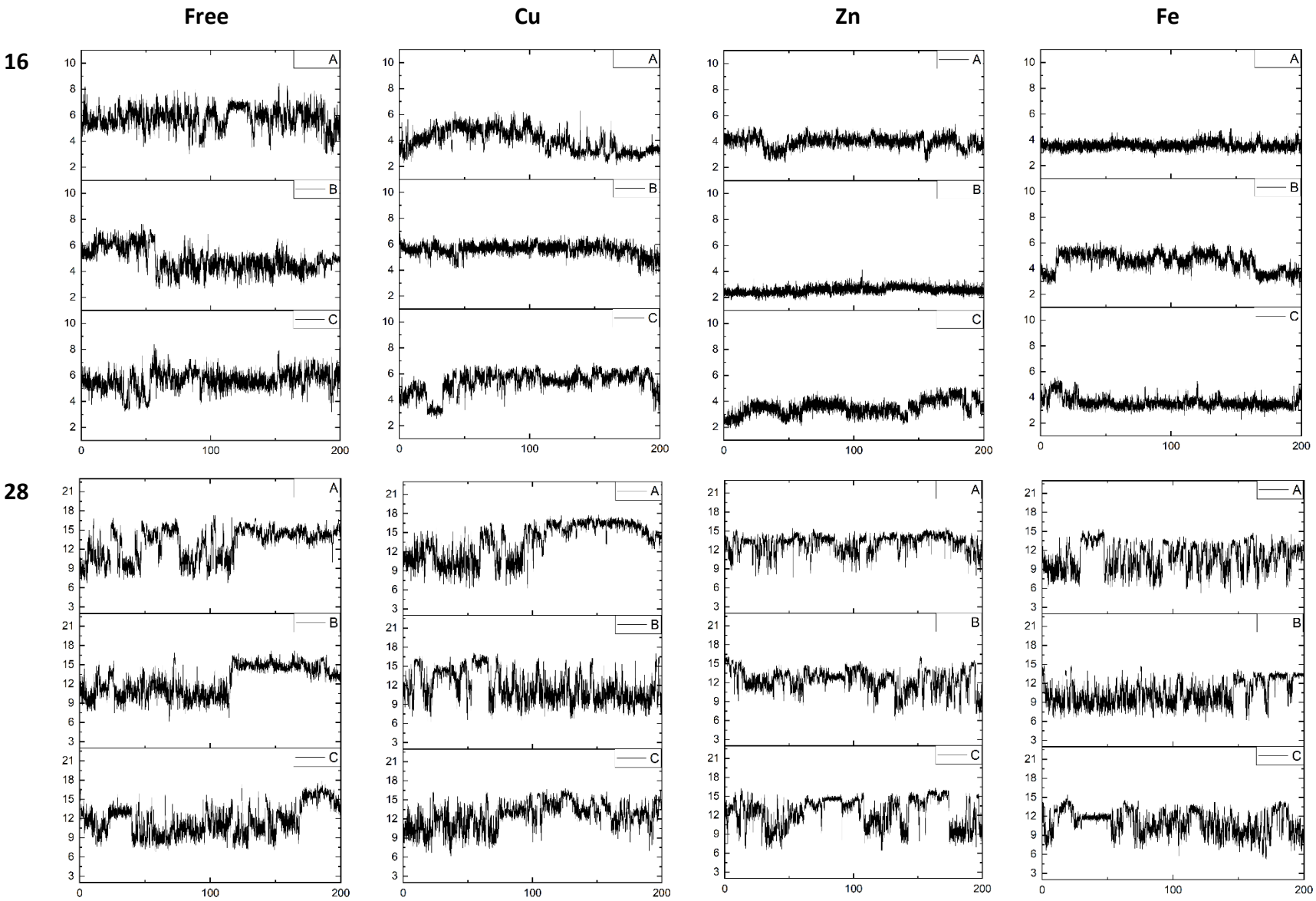
Coordination	Chain Length	AlphaP	AlphaD	EthreshP	EthreshD
Free Aβ	*16*	42	9	-436	254
	28	70	16	-603	436
	40	95	22	-584	597
	42	100	24	-569	621
Cu	*16*	52	11	-362	252
	28	89	20	-473	442
	40	121	28	-448	603
	42	126	29	-418	625
Zn	*16*	42	9	-285	245
	28	71	16	-445	432
	40	97	22	-432	585
	42	101	24	-405	618
Fe	*16*	52	12	-278	264
	28	89	20	-428	447
	40	121	29	-436	609
	42	162	30	-395	633

4.3 Results and discussion

For all 16 structures, 50 ns of conventional MD (cMD) data was generated followed by 3 parallel 200 ns aMD calculations. These aMD trajectories were then combined creating 16 600 ns aMD trajectories, one for each structure. Each system has reached a quasi-equilibration point[38] during the initial 50 ns conventional MD for each run, as observed from RMSD data used as a measure of equilibration. RMSD plots for individual simulations can be observed in Figure 4.2, giving clear indication that the simulations have reached a suitably equilibrated state, as seen from relatively consistent RMSD values. All data reported is taken from combined aMD data after this period of equilibration.

Radius of gyration (R_g) measurements were used as a measure of peptide size. Statistics for combined 600 ns aMD trajectories, forming full graphs for individual trajectories, are displayed in Figure 4.3 and Table 4.2. Simulations for Aβ bound to zinc and iron generally showed the smallest averages suggesting these sampled more compact structures relative to the other simulations of comparable chain lengths. This could be attributed to iron structures here being pentacoordinate which could lead to reduced flexibility of residues (supported by evidence discussed later in this Chapter compared to copper) whilst zinc-bound structures adopt more tetrahedral geometries compared to square-planar observed in structures containing copper. This difference in observed geometries can give rise to more restraint on the Zn-complex and account for the more compact structures observed. For 16 residue peptides, mean R_g is almost identical for free peptide and all three metals, but the maximum extent and standard deviation is lower when bound to metal, indicating that coordination to a metal centre within the N-terminus limits mobility and flexibility of these residues. For larger peptides, Zn and Fe binding restricts size and flexibility, as shown by relatively small maximum values, whereas Cu binding leads to similar or even greater size compared to free peptide. The increase in size and flexibility as the sequence is extended is also evident, although 40- and 42-residue peptides are not significantly different in mean or sd of R_g with no change in coordination.

Similarities are evident in R_g statistics between zinc & iron-bound peptides and copper & metal-free structures, most notably in the standard deviation, suggesting similar dynamics and interactions within these two pairs of structures. Maximum R_g values for free peptides provide evidence of enhanced mobility as a result of an absence of a metal centre and


Free
Cu
Zn
Fe
16
28
A
B
C

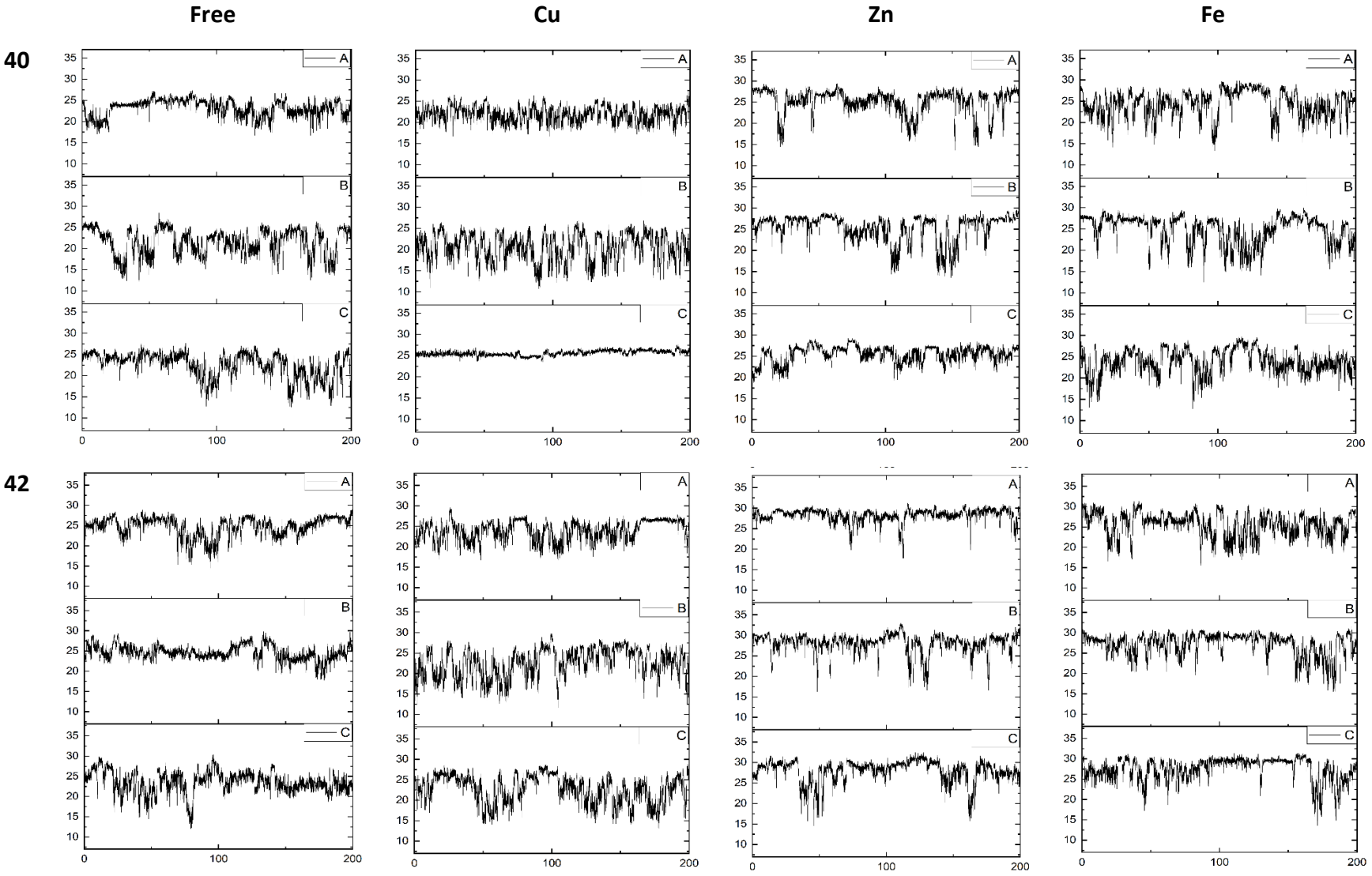
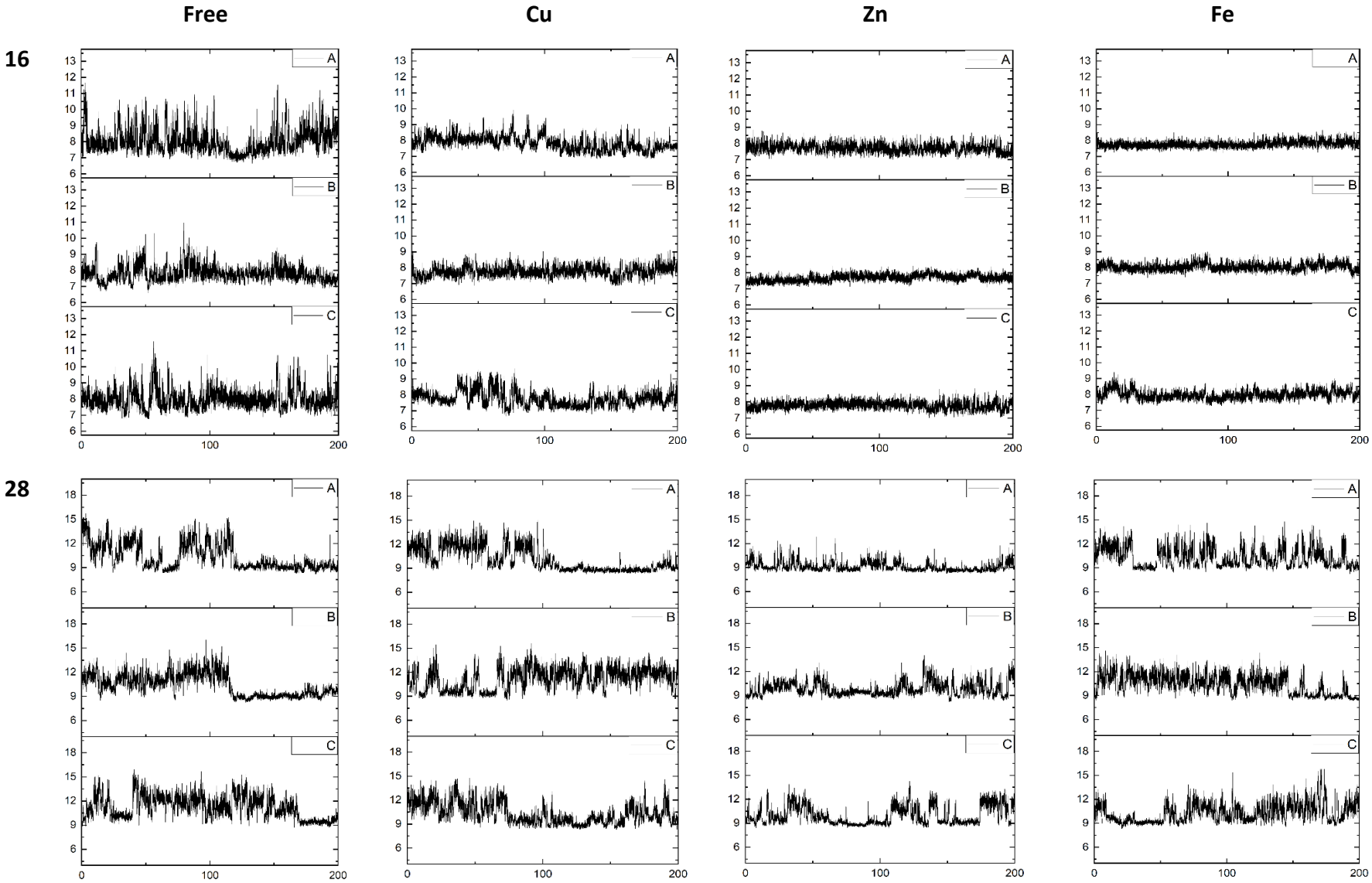


Figure 4.2- RMSD (Å, y-axis) plots for 3 x 200 ns aMD simulations of all 16 Aβ structures against time (ns, x-axis)


Free
Cu
Zn
Fe
16
28
A
B
C

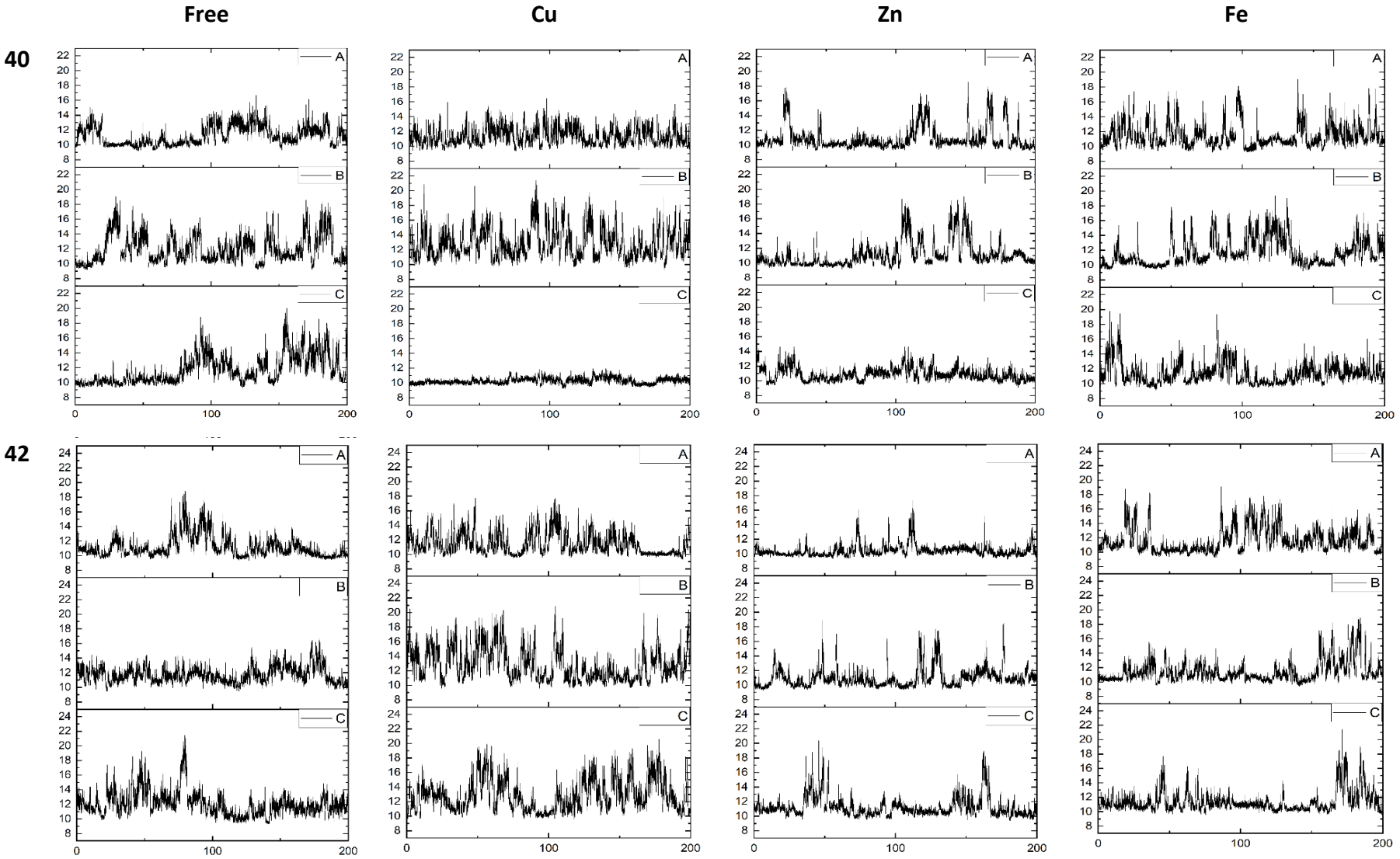


Figure 4.3- R_g (Å, y-axis) plots for 3 x 200 ns aMD simulations of all 16 Aβ structures analysed against time (ns, x-axis)

highlight the potential for restriction in movement upon metal binding, such as seen for zinc and iron containing structures. R_g data shows varying levels of mobility between simulations, with longer chain-lengths of Aβ-Cu and free Aβ showing relatively larger, more diffuse structures in comparison to peptides of the same length for zinc and iron.

	Average	Min	Max	Standard Deviation
16 residue chain length				
Cu	7.78	6.59	9.94	0.41
Zn	7.72	6.94	8.86	0.24
Fe	7.94	7.19	9.40	0.26
Free	7.93	6.60	11.66	0.61
28 residue chain length				
Cu	10.53	8.23	15.63	1.44
Zn	9.69	8.16	14.27	0.96
Fe	10.35	8.24	14.78	1.19
Free	10.65	8.14	16.29	1.50
40 residue chain length				
Cu	11.49	9.19	21.36	1.72
Zn	10.95	9.07	19.05	1.40
Fe	11.48	9.16	20.37	1.56
Free	11.63	9.14	19.99	1.70
42 chain length				
Cu	12.38	9.29	21.35	2.07
Zn	10.86	9.20	20.37	1.22
Fe	11.46	9.26	19.08	1.50
Free	11.66	9.27	21.59	1.46

Table 4.2- Statistical analysis of radius of gyration (R_g) data (Å)

Root mean square fluctuation (RMSF) data provides information on the level of mobility of individual residues within the simulations reported, and is shown graphically in Figure 4.4. Asp1 and His6 are coordinated to a metal-centre in all bound peptides. All free-peptide simulations have the highest RMSF values for these residues, demonstrating the anchoring effect of the metal centres within the N-terminus and the enhanced flexibility observed here in the absence of metal coordination. Values for metal binding residues in copper simulations are still relatively high compared to other metal-bound structures, which could again provide further evidence of similarities in dynamics between these and free peptides. It could be suggested that binding of copper to Aβ provides less restriction on these structures as opposed to other metal centres, giving rise to these enhanced RMSF values.

Arg5 also gives some of the highest RMSF values in most Cu(II) simulations despite metal binding via the adjacent residue, His6. Binding of metal ions to Aβ peptides occurs within the N-terminus of all structures, yet results showed that the effect of metal coordination on these molecules happened in both this region of the peptide as well as in residues remote to the coordination sites, towards the C-terminus.

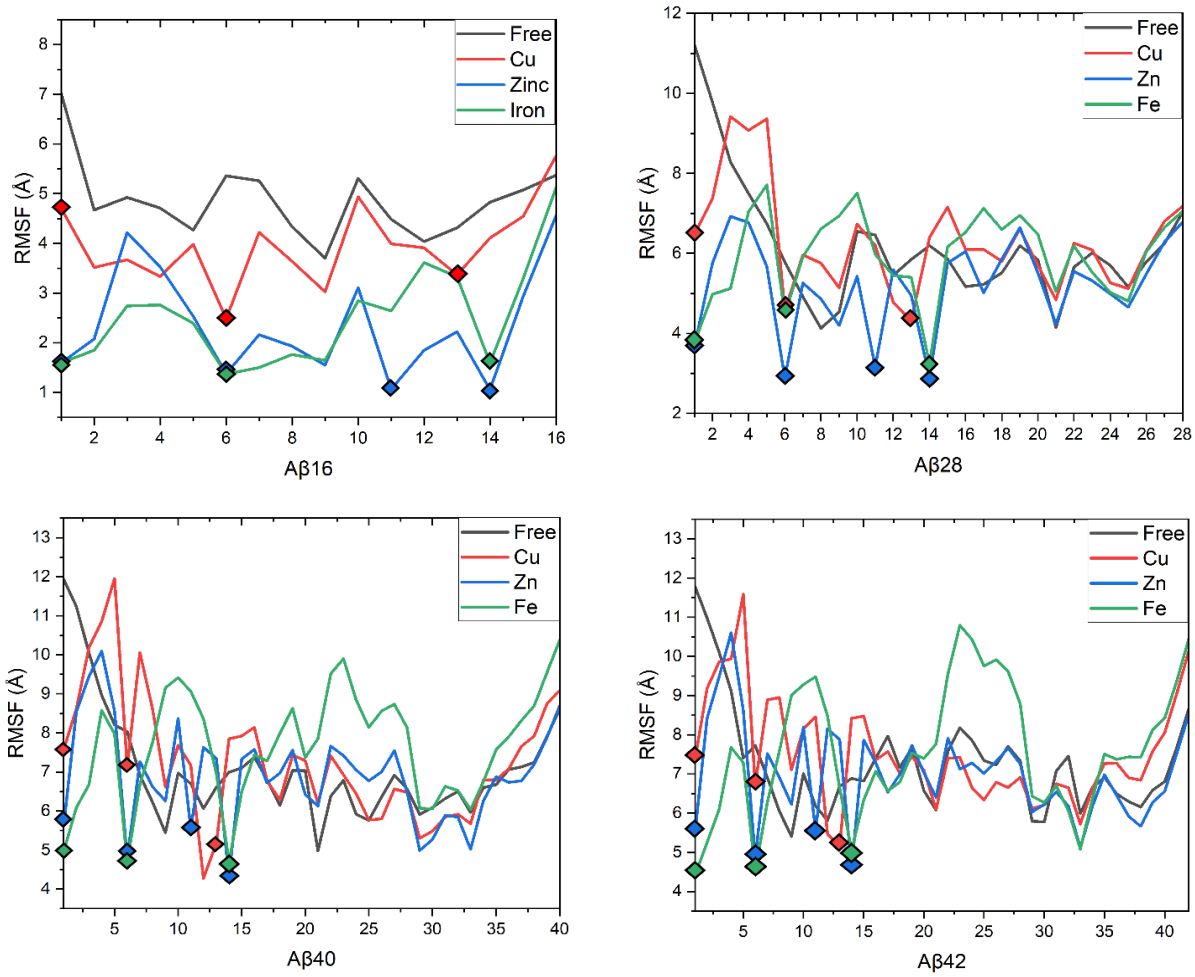


Figure 4.4- RMSF data (Å) based on individual residues. (Metal-bound residues highlighted in points of corresponding colours).

In the free peptides, the greatest contribution to flexibility arises from the Asp1 and His6 residues, which show largest increases in RMSF compared to all metal-bound peptides. The data indicates large peaks in mobility for residues towards the C-terminus for Fe(II)-bound structures, which confirms the effect of metal binding being a global effect on the peptide's structure, as opposed to solely at the binding site. The mobility observed in the R_g data for Aβ40-Fe and Aβ42-Fe is also apparent from its RMSF data, most notably between residues Glu22-Lys28, which give the greatest contribution to the peptide's overall movement. This

may be due to disruption of salt-bridges in this section of peptide induced by binding of Fe (*vide infra)*.

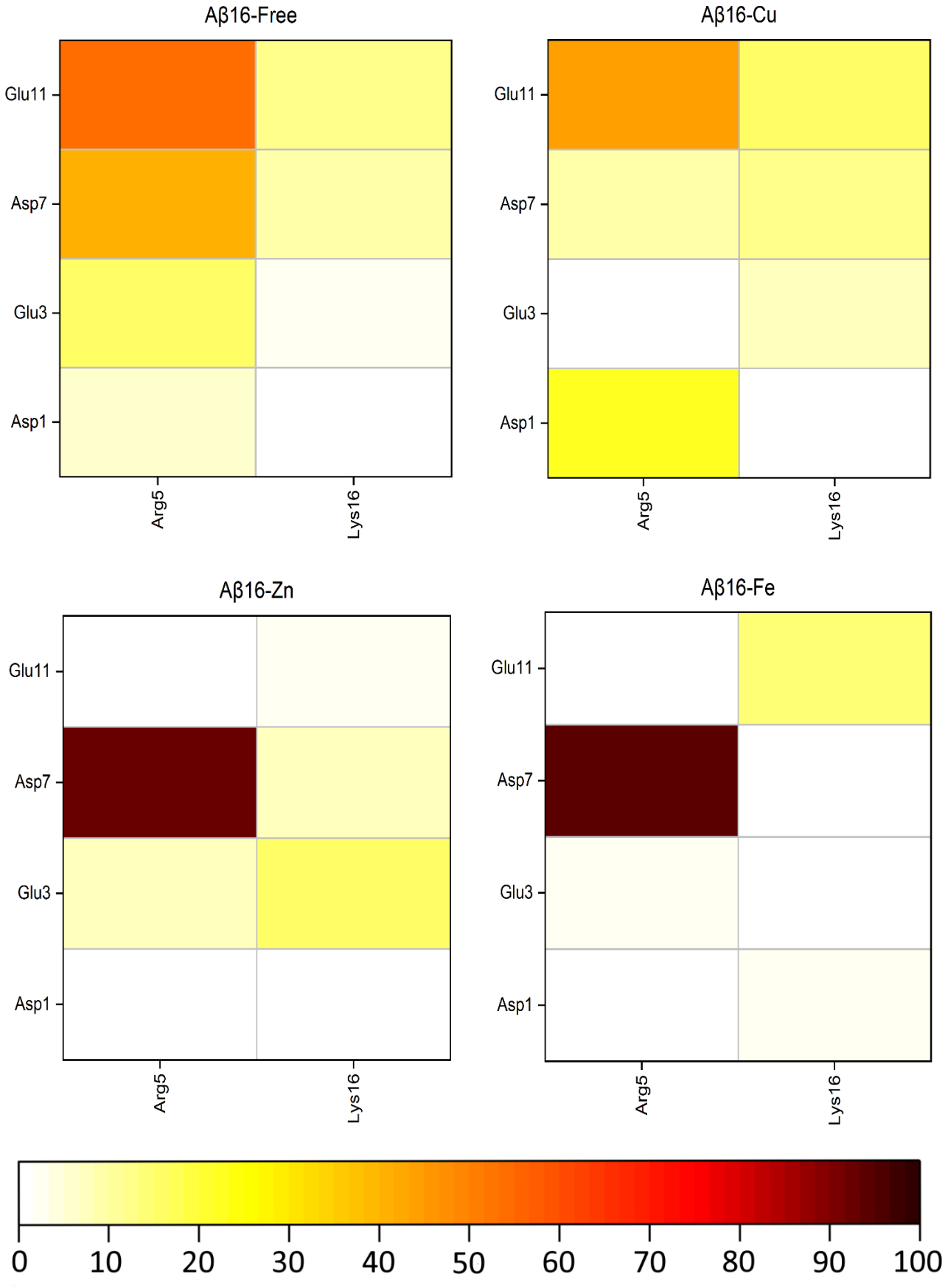
Cluster analysis identified and grouped sets of similar structures within a trajectory, based on RMSD data of the peptide backbone, via the DBSCAN[39] clustering algorithm within CPPTRAJ grouping structures in clusters with a difference in RMSD of 0.8 Å from one another relative to the start point. The low amount of cluster data is evidence of our enhanced sampling of these highly flexible peptides. Some trajectories sampled such large amounts of structures across the PES that the results couldn't be categorised into similar conformational ensembles, thus yielding low numbers or short-lived clusters. The simulations that did yield cluster data are highlighted in Table 4.3, below. All simulations of 16 residue peptides were found to form some clusters: those for the free peptide and Aβ16-Cu have rather low population, whereas for Aβ16-Zn and Aβ16-Fe over 70% of frames fall into just two clusters. Larger structures did not yield populated clusters due to greater degrees of freedom; metal complexes of Aβ28 exhibit just 3 clusters, with population of no more than 3%. A single cluster was found for Aβ42-Fe, but this was present for 0.9% of simulation time (approximately 540 frames out of 60,000).

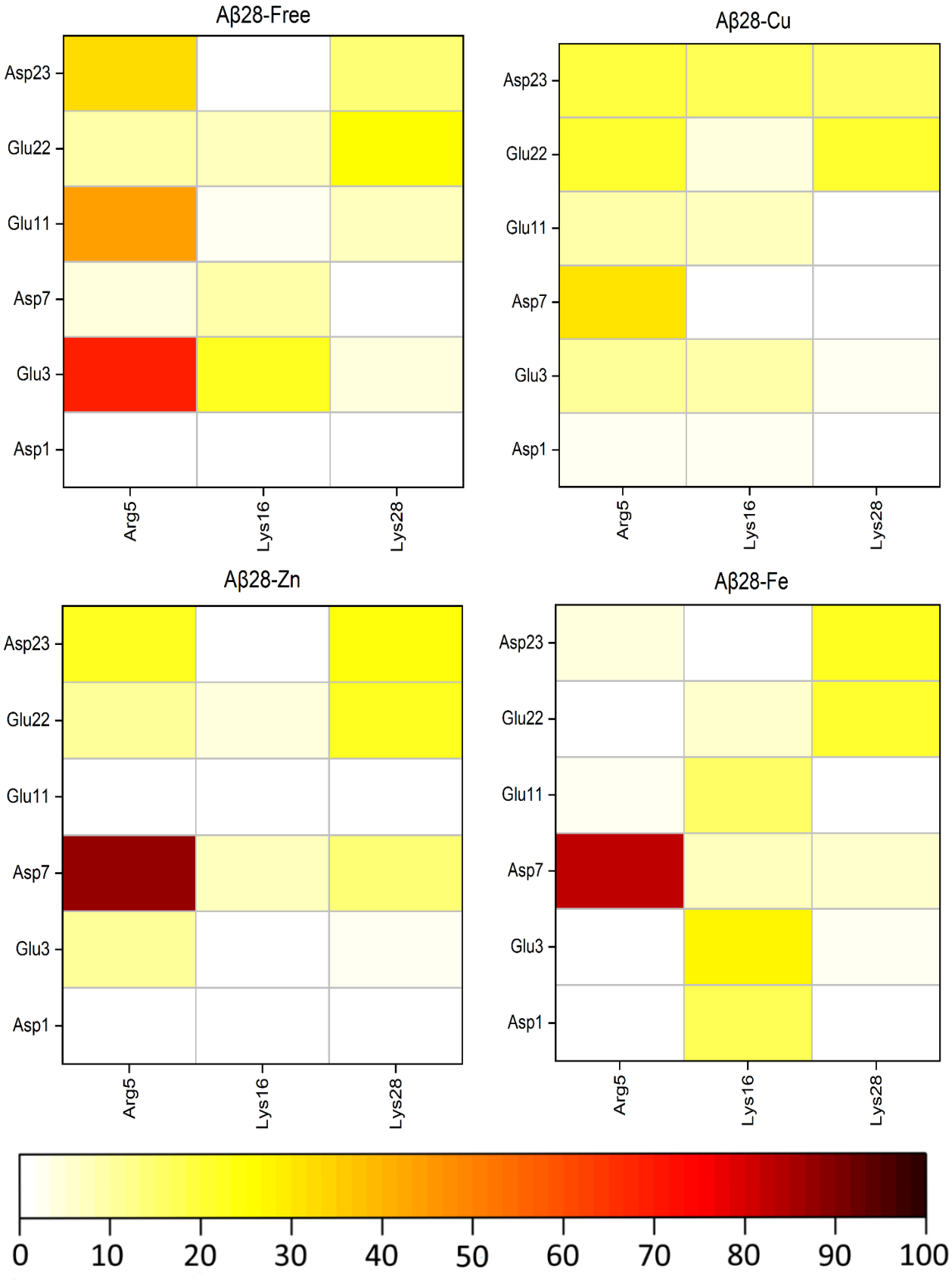
Structure	# Clusters	Most populated (%/time)	2nd most populated (%/time)
Aβ16- Free	6	8.8	1.6
Aβ16- Cu	36	11.9	5.2
Aβ16- Zn	12	51.8	21.8
Aβ16- Fe	7	61.9	10.7
Aβ28- Free	0	-	-
Aβ28- Cu	3	0.7	0.7
Aβ28- Zn	3	1.3	0.5
Aβ28- Fe	3	2.8	0.5

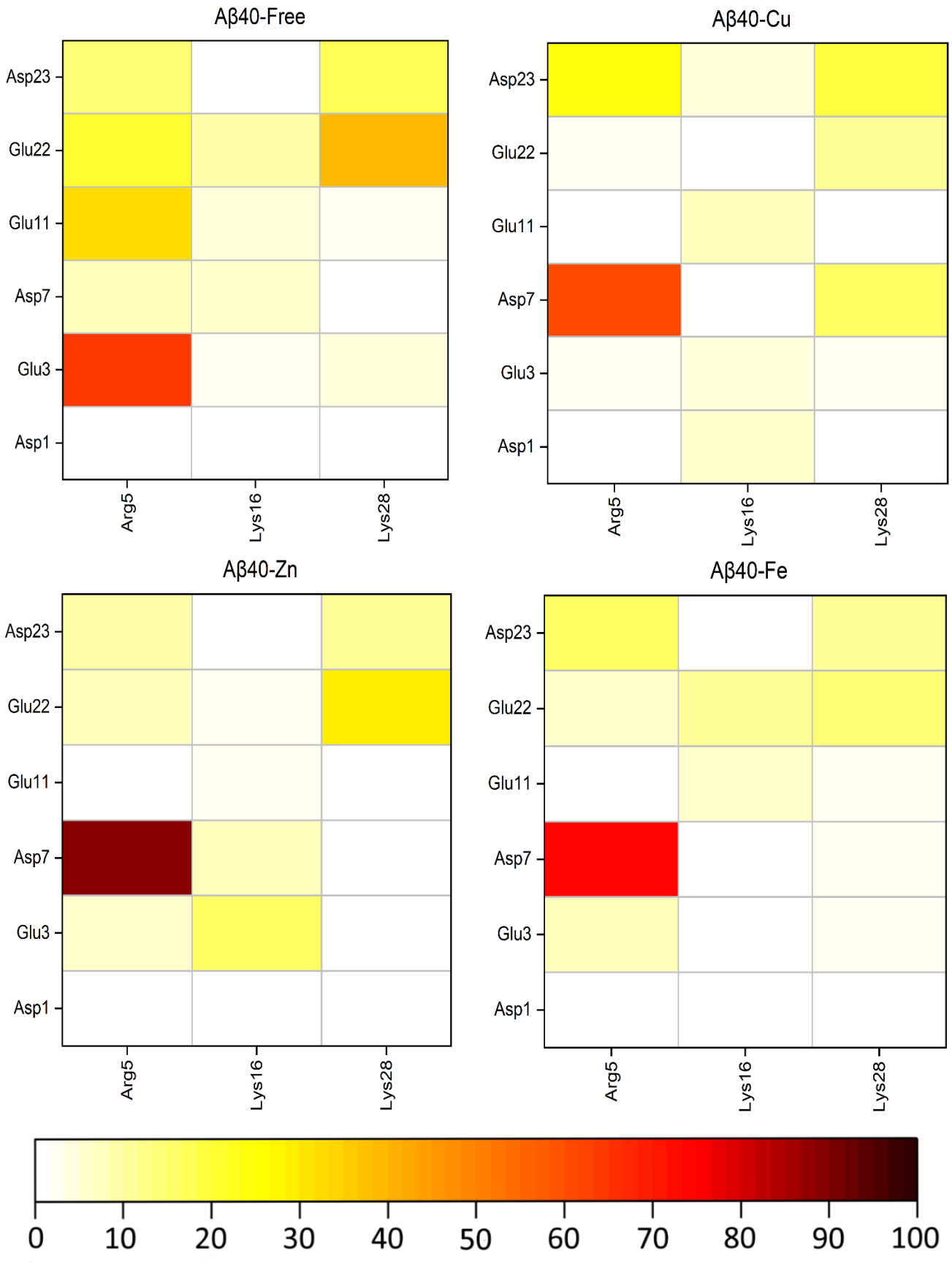
Table 4.3- Notable cluster analysis data for equilibrated trajectories, $A\beta_{40/42}$ not included as no significant clusters were found

Figure 4.5 displays the prevalence of salt-bridge contacts between charged residues, for all trajectories, by percentage of total frames. Aβ16 has the lowest amount of potential salt-bridge combinations that can be formed, due to the limited number of charged residues at the N-terminus. Types and occupation of salt-bridges formed in all four versions of the shortest peptide vary markedly, especially those involving Asp1 and Asp7, which are close to metal binding sites. Binding of Cu(II) reduces the number and persistence of some previously more common interactions. In the shortest peptide, copper binding leads to almost complete loss of Arg5-Asp7, presumably due to metal binding at His6, while Arg5-Glu11 is common at similar incidence levels in both the copper-bound and free peptide. In Aβ16-Zn and Aβ16-Fe simulations, however, the percentage incidence of the interaction between Arg5-Asp7 increases notably, whereas Arg5-Glu11 contacts are almost non-existent. The overall salt-bridge profile for these two structures is comparable in both types of interactions as well as frequency, further reinforcing the pattern noted above that Zn and Fe complexes behave similarly, while Cu complex shows behaviour closer to the free peptide. For example, Arg5-Asp7 maintains consistent incidence across all results for zinc and iron bound Aβ peptides, while copper and free peptide simulations show similarity.

The introduction of Glu22, Asp23 and Lys28 present several different combinations of potential salt-bridges that are not available in the shortest peptides. In the longer peptides, notable differences in salt-bridge formation are reported between the metallopeptide structures and the trajectories of those lacking in a metal centre. For instance, Arg5-Glu3 and Arg5-Glu11 contacts are common in longer metal-free peptides, but almost completely absent in metal-bound systems. All interactions involving Lys28 occur at similar levels of incidence across all simulations where this residue is present. Interestingly, Lys16 and Asp23 appear to never be in contact in any trajectory. The most common salt-bridge interaction by percentage in all potential frames was between Arg5-Asp7, occurring at 55.5% of all simulation data.


Aβ16-Free
Glu11
Asp7
Glu3
Asp1
Arg5
Lys16
Aβ16-Cu
Glu11
Asp7
Glu3
Asp1
Arg5
Lys16
Aβ16-Zn
Glu11
Asp7
Glu3
Asp1
Arg5
Lys16
Aβ16-Fe
Glu11
Asp7
Glu3
Asp1
Arg5
Lys16
0
10
20
30
40
50
60
70
80
90
100


Aβ28-Free
Asp23
Glu22
Glu11
Asp7
Glu3
Asp1
Arg5
Lys16
Lys28
Aβ28-Cu
Asp23
Glu22
Glu11
Asp7
Glu3
Asp1
Arg5
Lys16
Lys28
Aβ28-Zn
Asp23
Glu22
Glu11
Asp7
Glu3
Asp1
Arg5
Lys16
Lys28
Aβ28-Fe
Asp23
Glu22
Glu11
Asp7
Glu3
Asp1
Arg5
Lys16
Lys28
0
10
20
30
40
50
60
70
80
90
100


Aβ40-Free
Asp23
Glu22
Glu11
Asp7
Glu3
Asp1
Arg5
Lys16
Lys28
Aβ40-Cu
Asp23
Glu22
Glu11
Asp7
Glu3
Asp1
Arg5
Lys16
Lys28
Aβ40-Zn
Asp23
Glu22
Glu11
Asp7
Glu3
Asp1
Arg5
Lys16
Lys28
Aβ40-Fe
Asp23
Glu22
Glu11
Asp7
Glu3
Asp1
Arg5
Lys16
Lys28
0
10
20
30
40
50
60
70
80
90
100

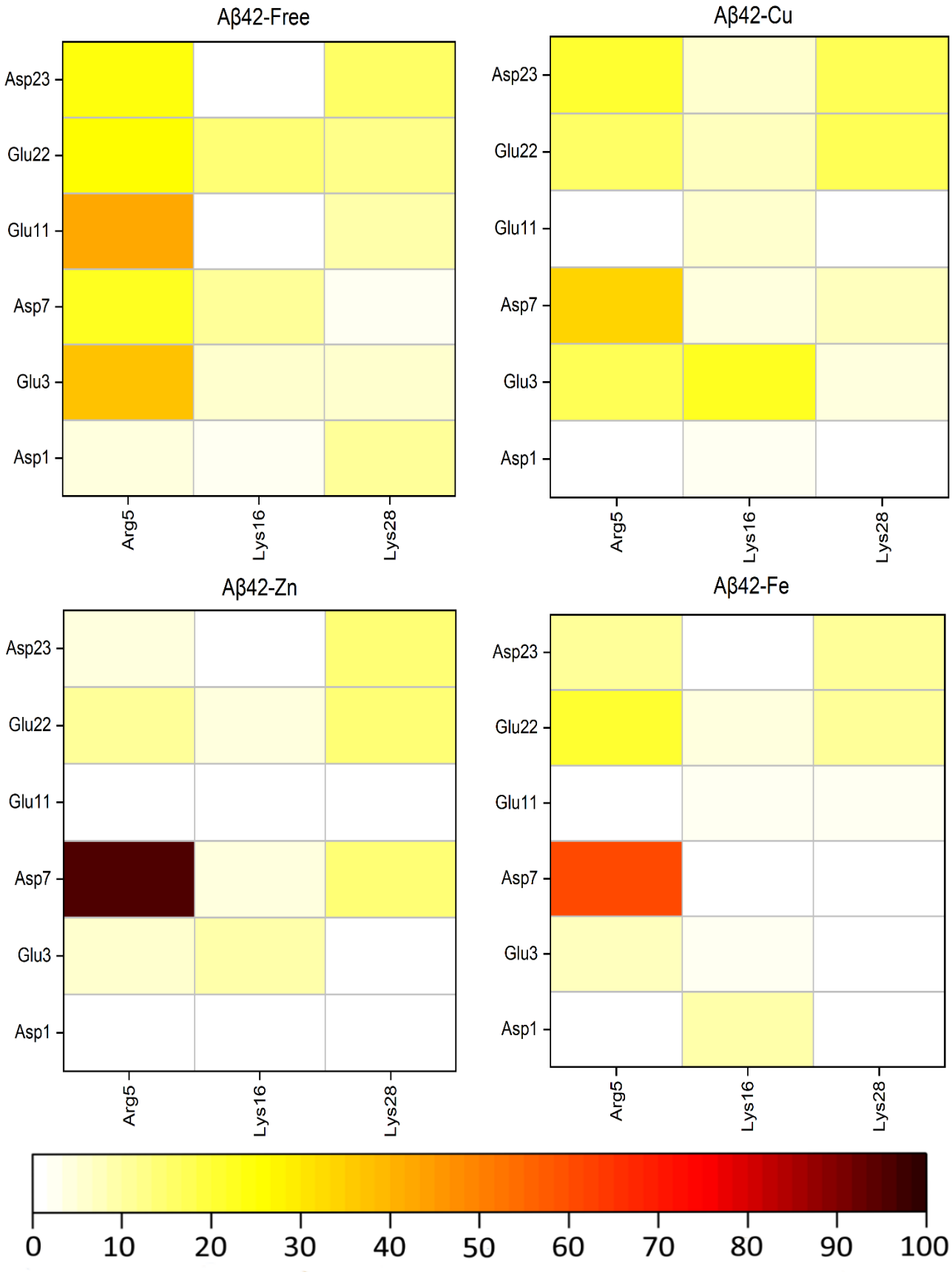


Figure 4.5-Salt-bridge plots between oppositely charged residues by percentage (%)

Aβ is intrinsically disordered by nature yet, secondary structure has commonly been associated with aggregational properties of Aβ. Misfolding of proteins and enhanced levels of β-sheet structures can give rise to increased propensity for formation of potentially neurotoxic species[40,41]. Incidence plots illustrate secondary structure type per residue across entire trajectories; these can be further classified as either helical, strand, or other. Data for these plots is in Table 4.4 and displayed alternatively in Figure 4.6. From the data the levels of helical secondary structure in metal-bound structures increases with chain length of the peptide. The highest levels of secondary structure in all peptides are found as helices, which make up 30 to 40% of the larger peptides; this is particularly evident in Cu-bound and free peptide, for which almost all lengths of peptide adopt helical character for at least 30% of residues. For Zn and Fe complexes, helical content increases with peptide length but remains significantly lower than Cu or free. Iron and Zinc simulations for 16-residue structures show considerably lower incidence levels of helix structure compared to all other data.

	Fe			**Zn**			**Cu**			**Free**		
	Helix	**Strand**	**Other**	**Helix**	**Strand**	**Other**	**Helix**	**Strand**	**Other**	**Helix**	**Strand**	**Other**
AB16	1.1	< 0.1	98.8	2.2	1.2	96.6	30.0	1.1	68.9	22.1	1.7	76.2
AB28	18.6	0.2	81.2	24.8	1.1	74.1	29.5	0.5	70.0	45.0	0.8	54.2
AB40	23.3	1.4	75.3	25.8	1.9	72.3	31.5	1.1	67.4	42.4	1.2	56.4
AB42	28.3	0.8	70.9	30.7	1.7	67.6	33.8	1.4	64.8	38.1	3.4	58.5

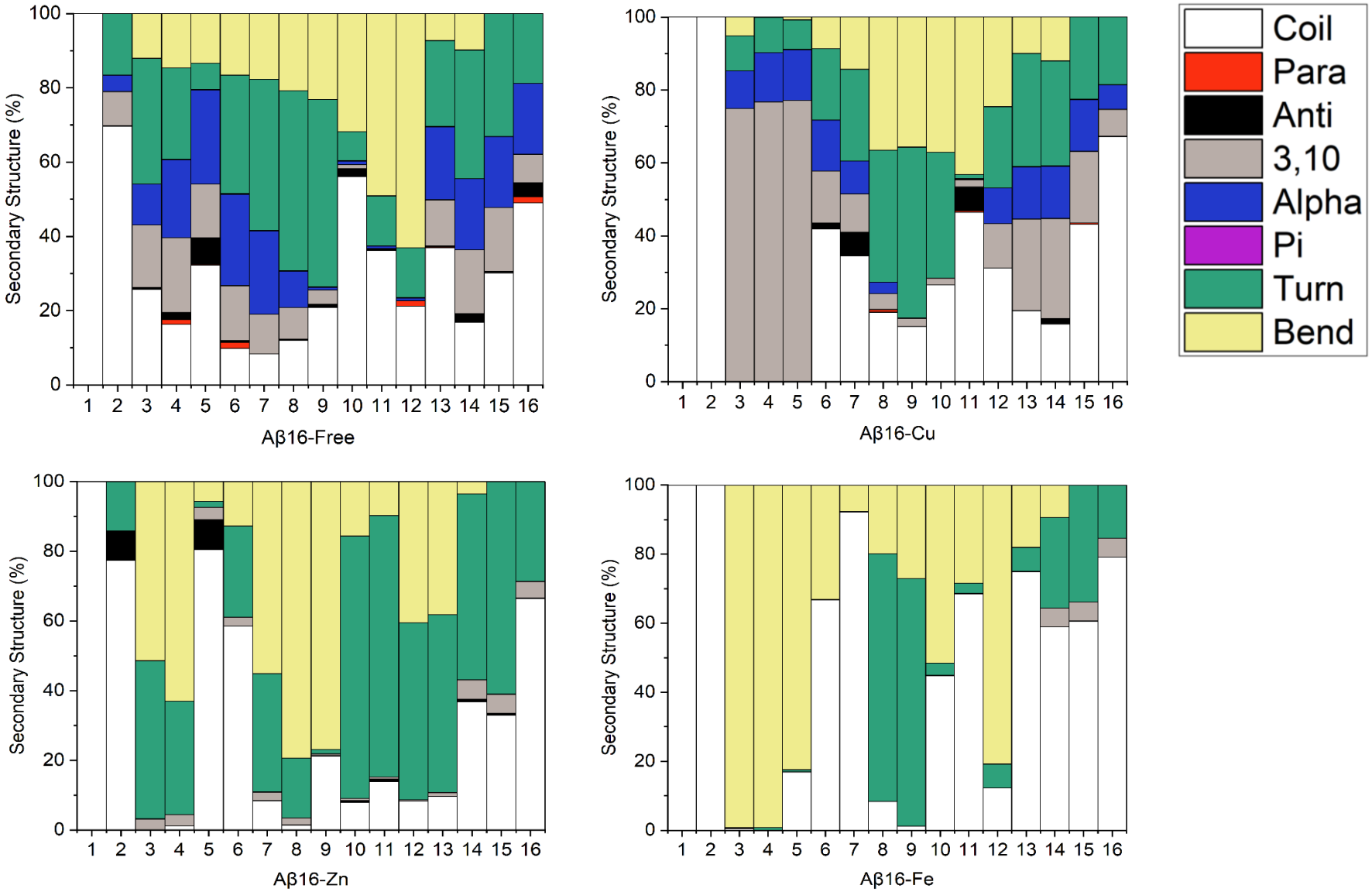
Table 4.4- Percentage time of secondary structure types by residue classified as helical, strand, or other

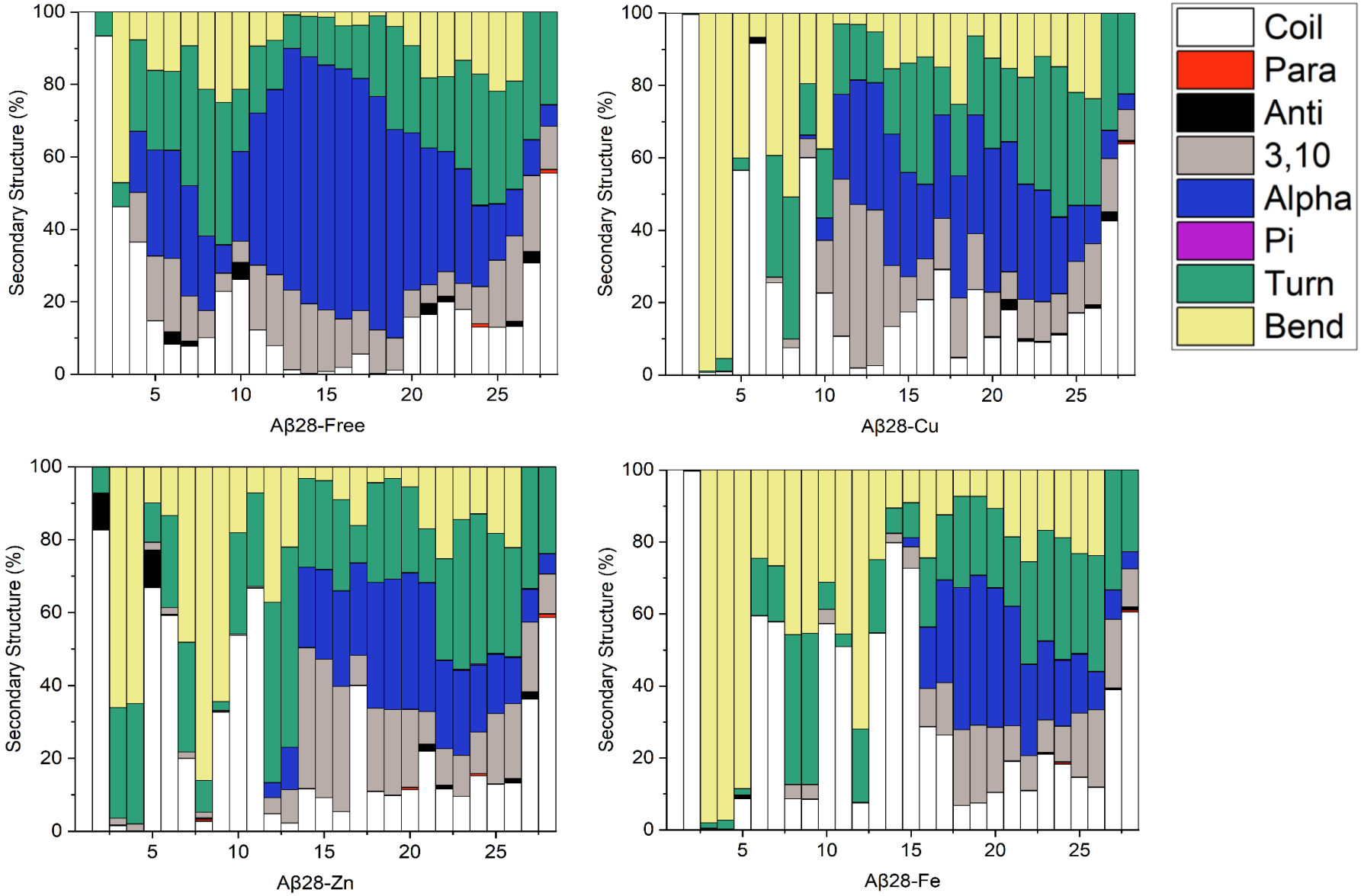
For most simulations, the level of β-strand is slightly larger in the free peptide than in the metal-bound complexes, with the level broadly increasing in longer peptides. Aβ42-free shows the highest percentage incidence of β-strand, at 3.4%, which is double the next most prevalent, namely Aβ16 and Aβ42-Zn at 1.7%. β-strand secondary structure has commonly

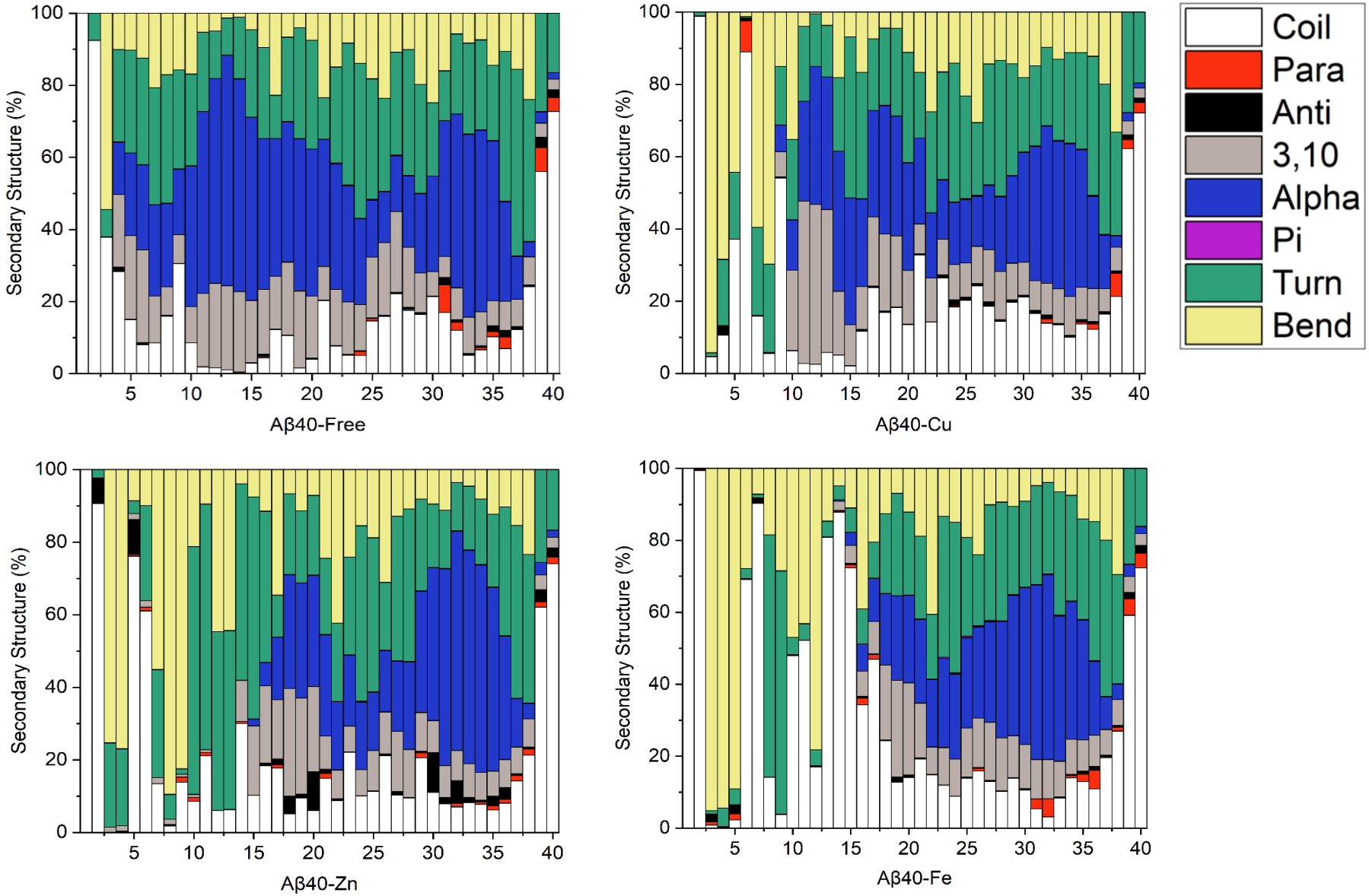
been associated with increased formation of neurotoxic species even at relatively low levels[42]. The longest form of the peptide, Aβ42, has the greatest propensity for self-aggregation and in this experiment, Aβ42 free has the greatest value for levels of beta secondary structure, which on the surface supports previous findings on the increased proclivity for aggregation possessed by this structure. However, the percentage incidence of strand/sheet secondary structure reported in Table 4.4, is uncertain with respect to enhanced proclivity for aggregation into potentially neurotoxic species. A 2003 study conducted by Schmechel *et al.*[43] showed levels of β-sheet secondary structure at a higher incidence in comparison to the simulation data presented here, which shows the levels of β-strand within the experimental results in this chapter do not replicate the levels of those observed *in vivo* to definitively prove increased Aβ toxicity[44].

Helical and strand secondary structure are generally absent or at very low levels within residues at the N-terminus for all metallopeptide simulations. All free peptide trajectories show an increase in the distribution of these forms of secondary structure across the whole peptide, mainly within residues towards the N-terminus mentioned previously. The result further highlights the enhanced rigidity shown within this region due to coordination to a metal centre. As with the salt-bridge profiles, there is also a direct comparison that can be observed showing similarities in secondary structures levels and content between iron and zinc simulations, as well as comparable data shown for copper-bound and free peptides.

Secondary structure characterisation also demonstrates the effect of these ions within the metal-binding region of the N-terminus. There is a notable absence of any helical or strand secondary structure within the first 10 residues of all metallopeptides analysed. They mainly arrange themselves into coil/turn/bend assemblies, and helical or strand secondary structures only form within this region when no metal centre is present. For all free peptides, organised secondary structure is present consistently throughout the whole peptide highlighting the relative rigidity caused by binding of a metal ion.


Secondary Structure (%)
100
80
60
40
20
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Aβ16-Free
Aβ16-Cu
Aβ16-Zn
Aβ16-Fe
Coil
Para
Anti
3,10
Alpha
Pi
Turn
Bend


Secondary Structure (%)
100
80
60
40
20
0
5
10
15
20
25
Aβ28-Free
Aβ28-Cu
Aβ28-Zn
Aβ28-Fe
Coil
Para
Anti
3,10
Alpha
Pi
Turn
Bend


Secondary Structure (%)
Aβ40-Free
Aβ40-Cu
Aβ40-Zn
Aβ40-Fe
Coil
Para
Anti
3,10
Alpha
Pi
Turn
Bend

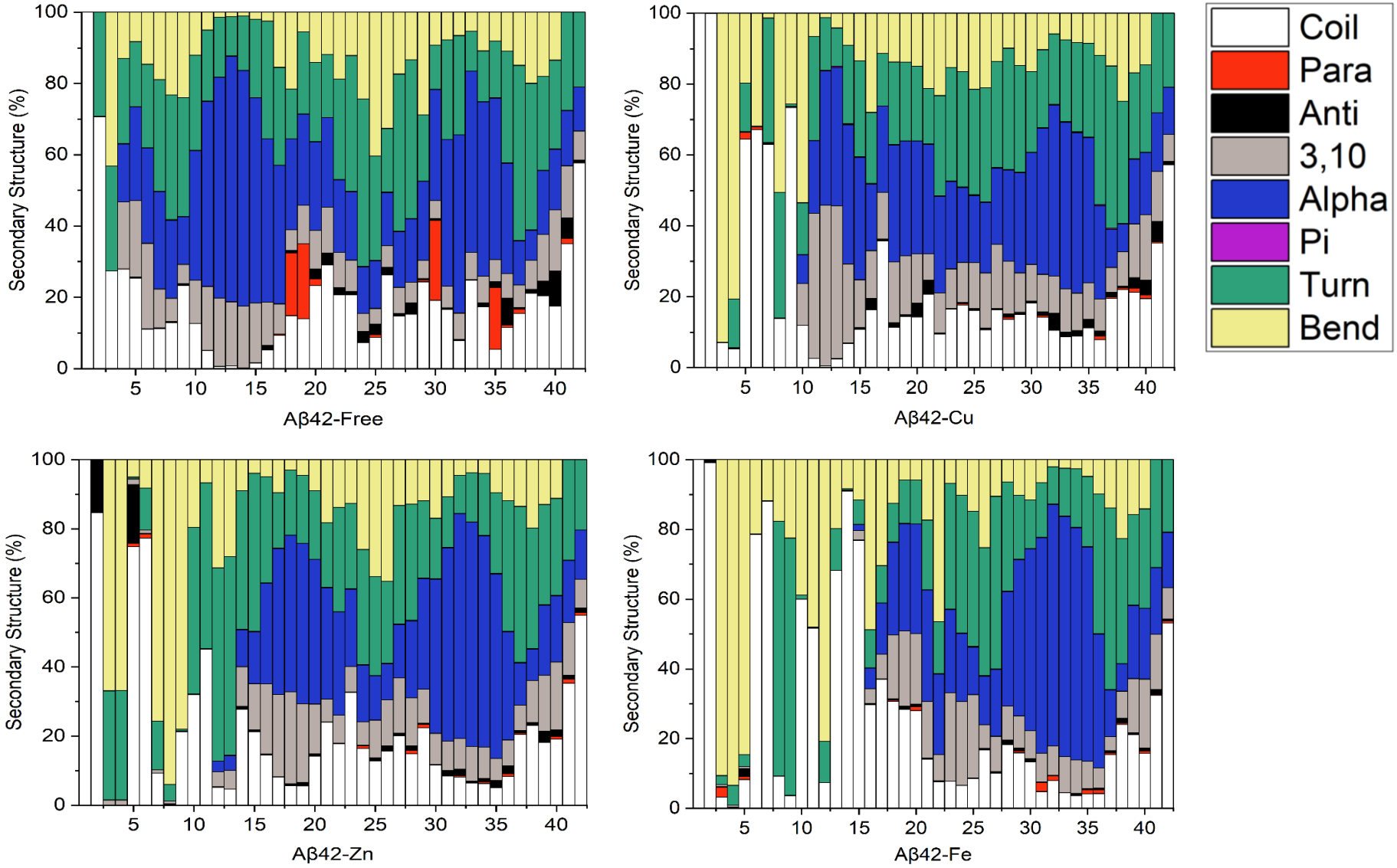


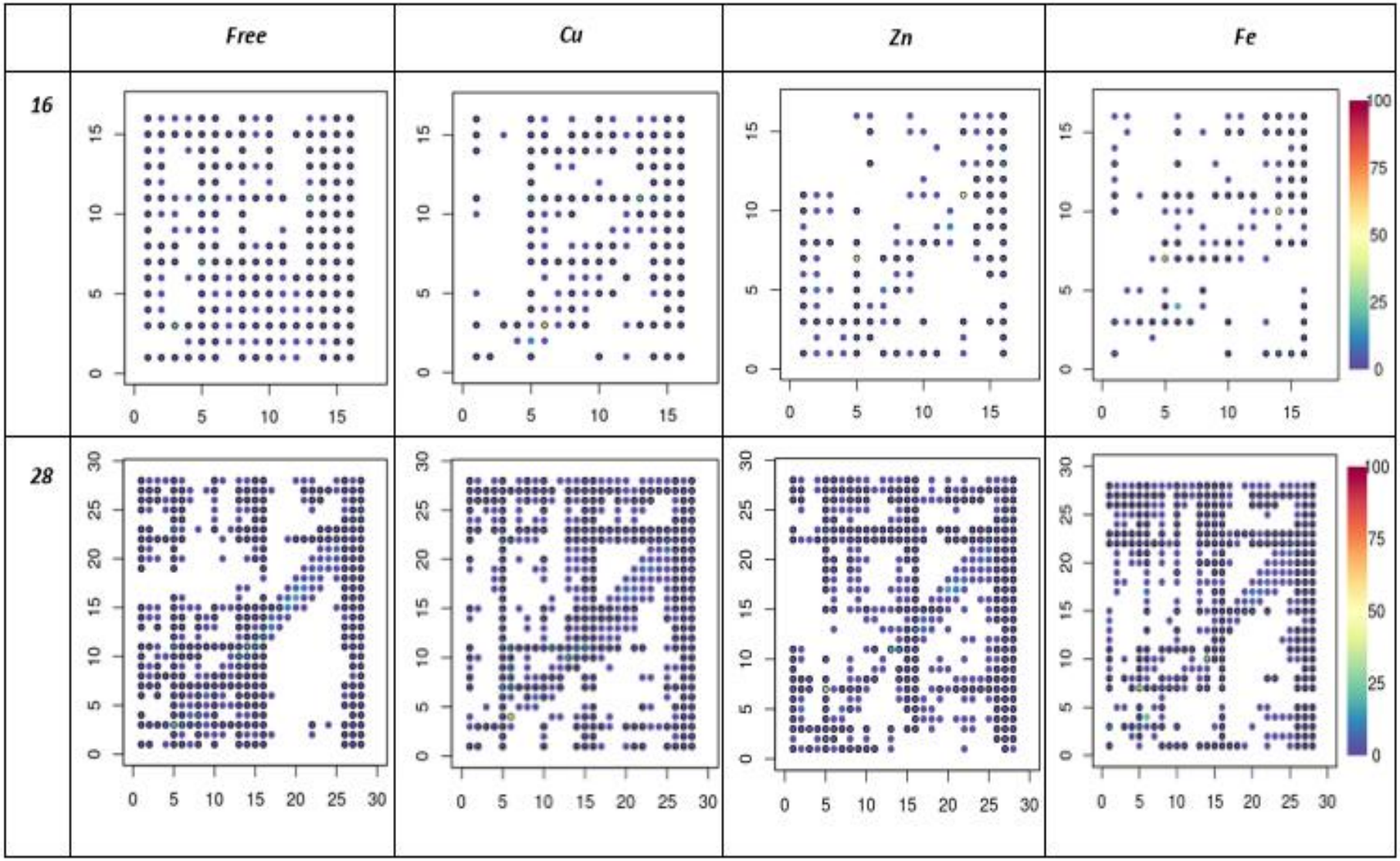
Figure 4.6- Secondary structure incident plots as a percentage of trajectory frames by amino acid for all 16 structures

Hydrogen bond values in Table 4.5 and Figure 4.7 illustrate interactions between appropriate atoms per residue, based on their capacity as a donor or an acceptor. Statistics taken from these graphs allow for comparisons on the formation of hydrogen bonds and thus the transient nature of conformations being sampled. As an implicit solvent is used in all simulations, any potential hydrogen bonding between peptides and solvents are not present in the model. Hydrogen bond input files for interactions greater than 5% of simulation timeare included in the Appendix (A1.1-A1.16).

	Average	**Min**	**Max**	**Standard Deviation**
16 residue chain length				
Cu	4.93	0	13	1.81
Zn	4.16	0	13	1.73
Fe	4.04	0	11	1.70
Free	6.46	0	16	2.04
28 residue chain length				
Cu	7.96	1	19	2.37
Zn	7.26	0	21	2.42
Fe	7.23	0	20	2.41
Free	9.38	0	22	2.81
40 residue chain length				
Cu	10.17	1	23	2.64
Zn	9.38	0	22	2.75
Fe	9.11	0	25	2.64
Free	11.07	1	24	3.00
42 residue chain length				
Cu	10.11	0	22	2.75
Zn	9.72	0	24	2.94
Fe	9.91	0	23	2.74
Free	12.10	1	27	3.15

Table 4.5- Statistical data for hydrogen bonds of equilibrated aMD trajectories. Hydrogen bonds are characterised by maximum default angles of 135^{o} and distances of 3 Å

Most simulations possess at least one structure with zero hydrogen bonds, whilst the maximum amounts of hydrogen bonds within simulations at a given snapshot range from 13-27. This shows the mobile nature of the peptides analysed, with folding and unfolding occurring over the course of the trajectories observed. Free peptides show the highest maximum, standard deviation, and average values, indicating the particularly dynamic


Free
Cu
Zn
Fe
16
28
0
5
10
15
20
25
30
100
75
50
25
0

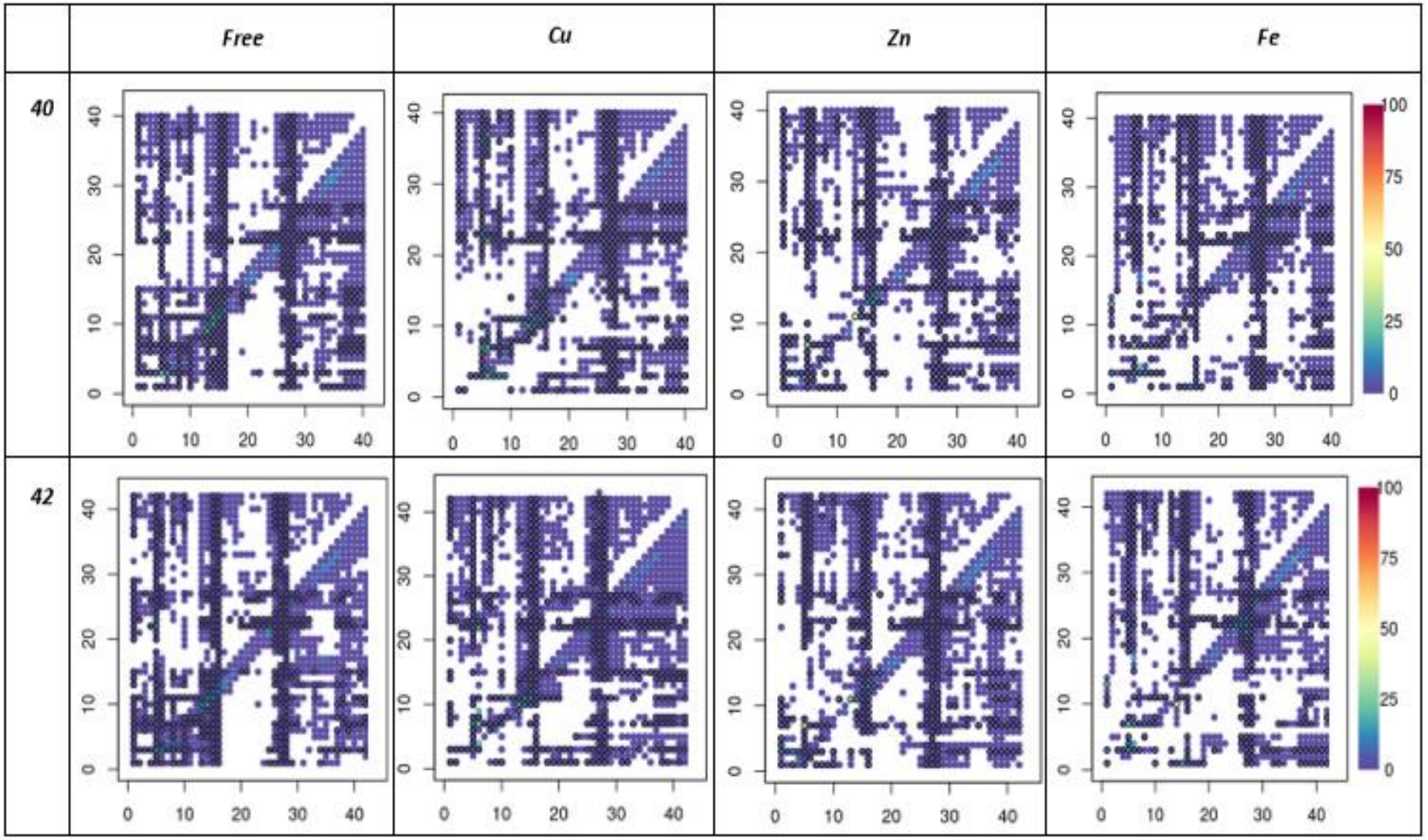


Figure 4.7- Hydrogen bond incident plots of equilibrated aMD trajectories (X-axis: Donor residues, Y-axis: Acceptor residues)

properties of Aβ in the absence of a metal centre. In comparison, zinc possessed the lowest average values of all types of structures analysed. The result was surprising as previously zinc structures were shown to have the lowest R_g values, indicating it adopted potentially more compact conformations. This in turn, would suggest shorter contact distances between residues, which would be expected to encourage greater incidence of hydrogen bond interactions between residues. This information suggests that potentially the structures were particularly transient, which is supported by the increased number of different clusters in Aβ16-Zn, relatively lower RMSF values, and short-lived salt-bridge contacts for example. Once again, similarities were observed in values between zinc and iron-containing systems, supporting previous conclusions of similarities in dynamics between Aβ peptides bound to these ions.

Taken together, the data reveals a picture of metal-induced disruption of peptide flexibility, secondary structure, and salt-bridge patterns. While only monomers were studied here, some potential implications for aggregation can be drawn. The salt-bridge Asp23-Lys28 is known to play an important role in the kinetics of fibrillogenesis[45,46]: we find that this interaction is moderately populated in free and copper-bound peptides (16 to 19% of frames), but that this is significantly reduced in iron and zinc complexes. One can envisage, therefore, that metals might disrupt the β-turn structure required for the hydrophobic interactions that stabilise fibrils. Disruption of secondary structure is also evident; much of this is in the N-terminal region, but variance is seen across the entire structure. Changes were also notes in regions such as the central hydrophobic cluster ($L_{17}VFFA_{21}$). In free Aβ42 especially, this sequence exhibits some β-strand character which is almost completely absent in all metal adducts. A similar effect is apparent in the hydrophobic residues in the C-terminal region, especially residues 35-40. Such β-strands have been proposed as a likely seed for aggregation[47], such that results suggest metal ions may disrupt the canonical aggregation, perhaps leading to altered oligomers and aggregates and/or kinetics of association. Although these conclusions are highly speculative, it is clear that despite metal binding occurring exclusively in the N-terminal region, its structural effects are felt across the whole peptide, including those regions known to affect aggregation.

Reweighting of the bias potential applied in accelerated MD recovers the free energy surface of the unbiased system, expressed as a function of properties of the system, known

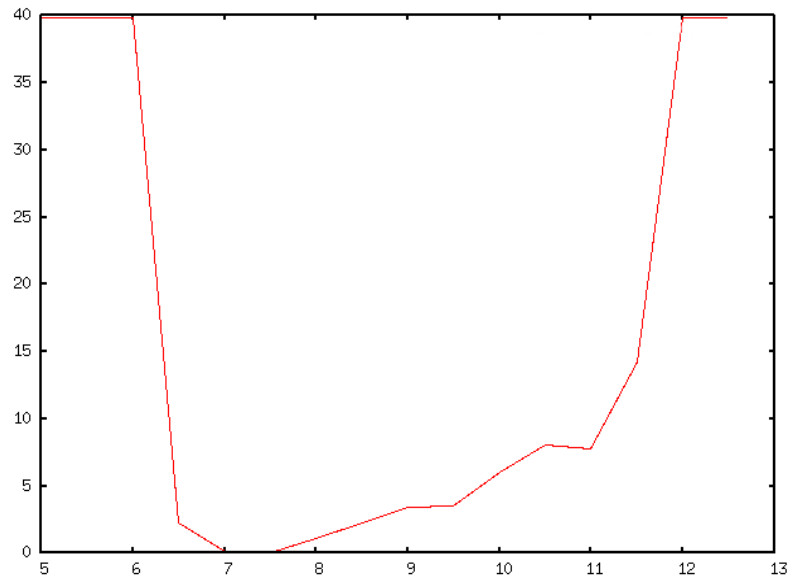
as potential of mean force (PMF). Plots of PMF as a function of R_g, hydrogen bond count, and end-to-end distance were used to examine energetic properties of structures sampled within the trajectories generated via aMD. PMF vs R_g is plotted in Figure 4.8. R_g values associated with minima structures for every trajectory were used as data points plot against PMF. For the smaller peptides (16 & 28 residues long), the metal-free minima possess a wider range of R_g within thermally accessible free energy, whilst Zn and Fe-bound peptides possessed the smallest range of R_g for minima values in all simulations. For the larger chain lengths analysed (40 & 42 residues), the results for all trajectories were comparable to one another, occupying similar ranges of R_g between all trajectories.

Hydrogen bond count PMF plots showed that, in general, peptides containing an iron or zinc metal centre possessed the lowest number and narrowest range of hydrogen bonds in low free energy structures. Free peptides generally exhibit more (and wider range) of hydrogen bonds, due to enhanced freedom of movement and interactions across the structure allowed by lack of hindrance that stems from coordination to a metal centre. Figure 4.9 shows these plots for all simulations as a representation of the trends observed across all structures.

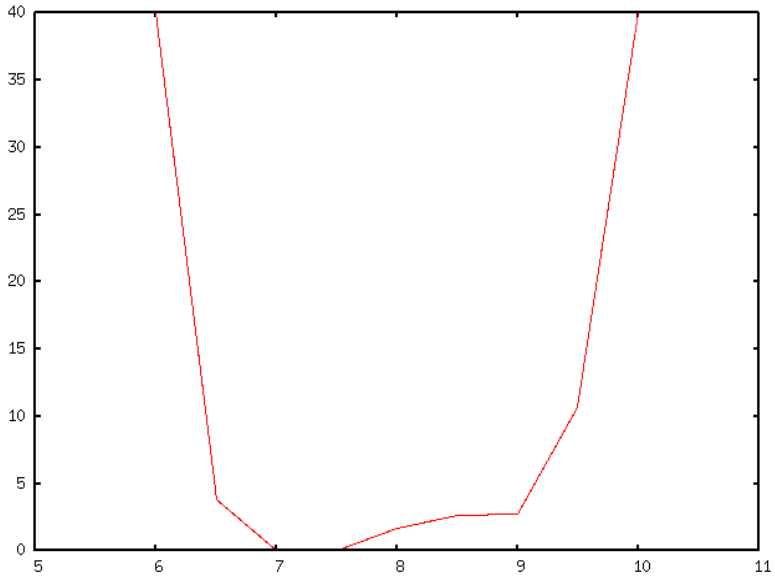
Two-dimensional PMF employing R_g and end-to-end distance (from Asp1 to the final residue in the sequence) were plotted for all sixteen trajectories (figure 4.10). R_g values for minima followed the same trend noted above, of similar ranges between zinc & iron structures and copper and free peptides. Generally, across all peptide lengths, smaller ranges of R_g and end-to-end distance were observed within simulations for Aβ bound to iron and zinc compared to the other two sets of trajectory data. These lower values indicate that structures containing iron or zinc adopt more compact conformations and exhibit lower mobility as copper and free peptide, in accord with data from other results, as well as other experimental data for similar structures[48]. The values observed are supported by RMSF data indicating an enhanced level of lability, especially within residue Asp1, when not bound to a metal centre, but also at relatively increased levels in Aβ-Cu peptides compared to other transition metals investigated.

Accelerated MD was chosen as a method for this study due to its enhanced conformational sampling, which is required for intrinsically disordered peptides of this size. Compared to previous studies conducted within the research group[49], increased conformational sampling

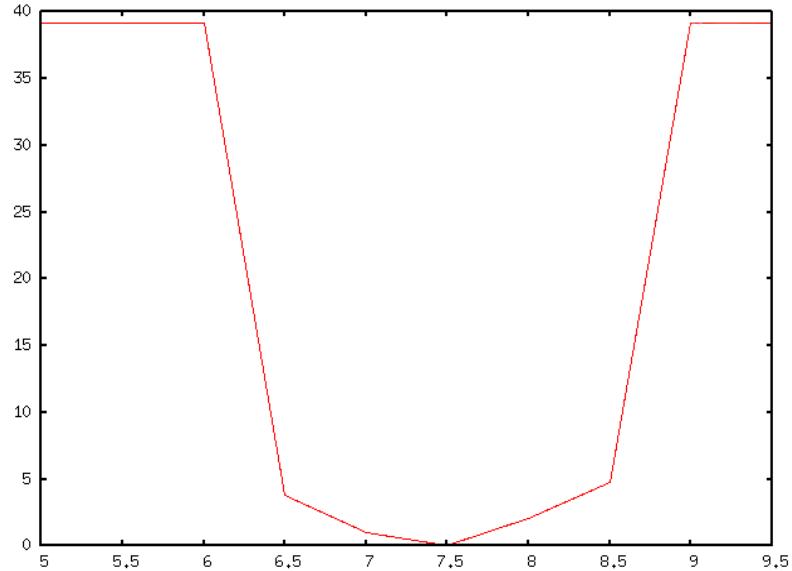
Aβ16-Free



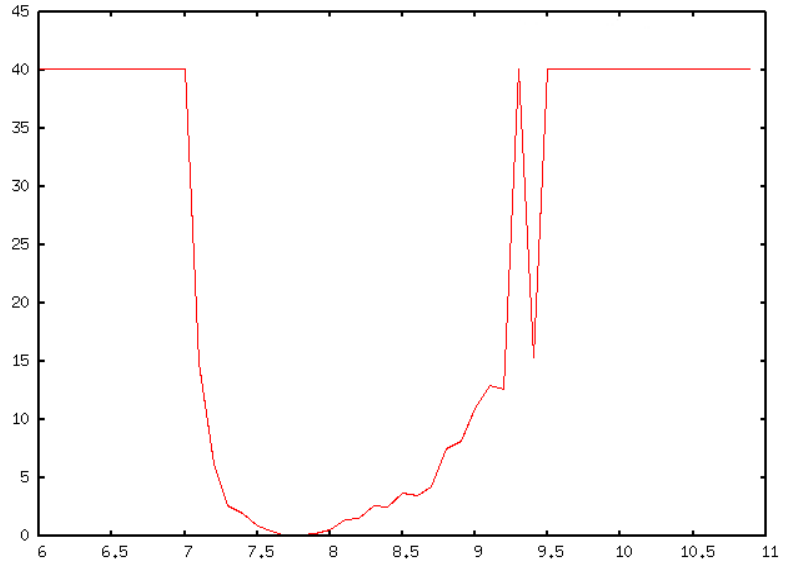
Aβ16-Cu



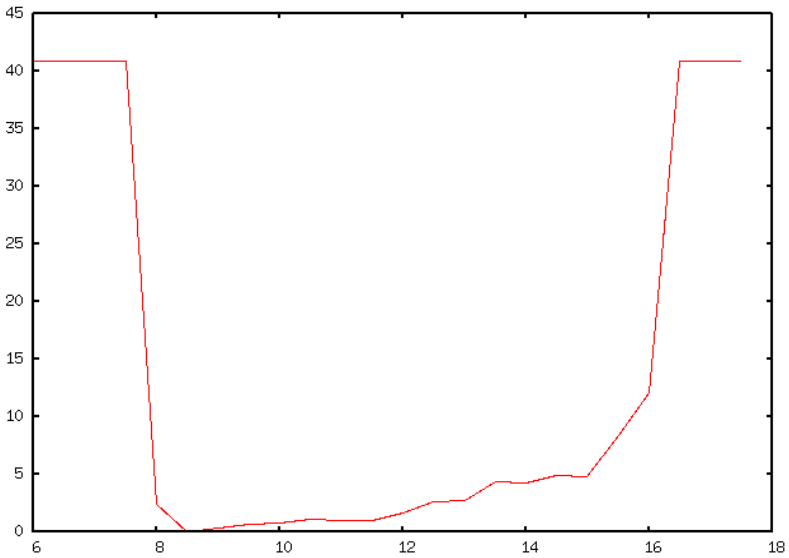
Aβ16-Zn



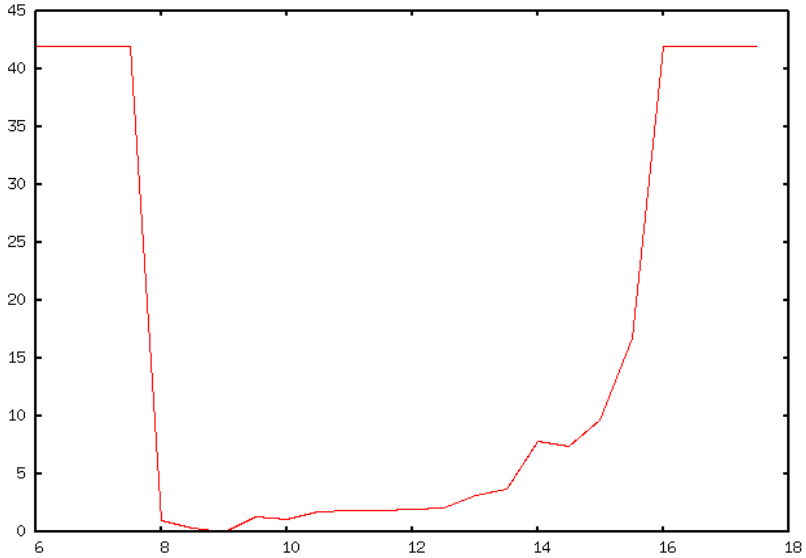
Aβ16-Fe



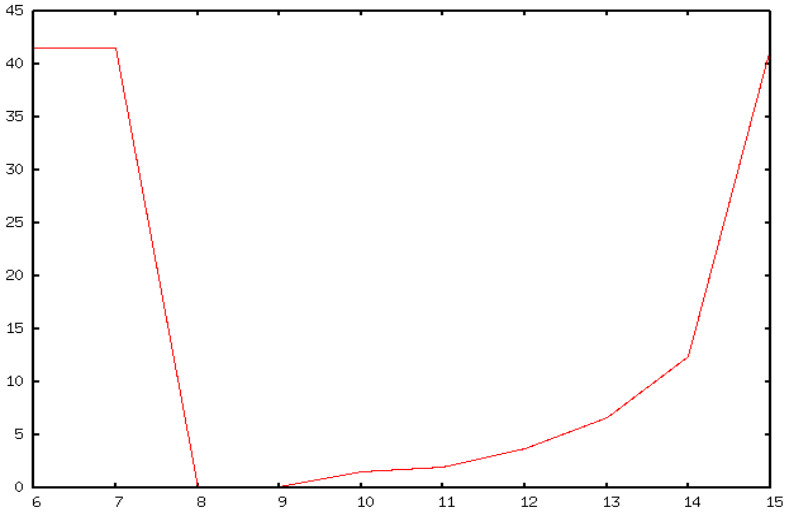
Aβ28-Free



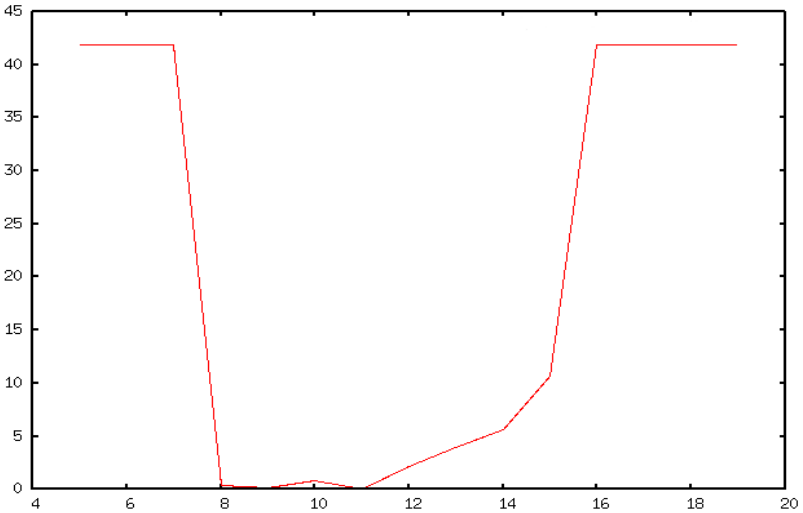
Aβ28-Cu



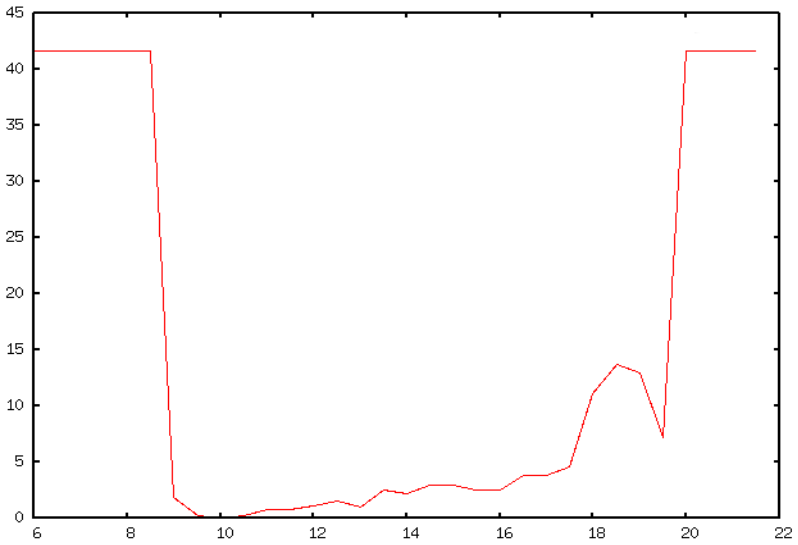
Aβ28-Zn



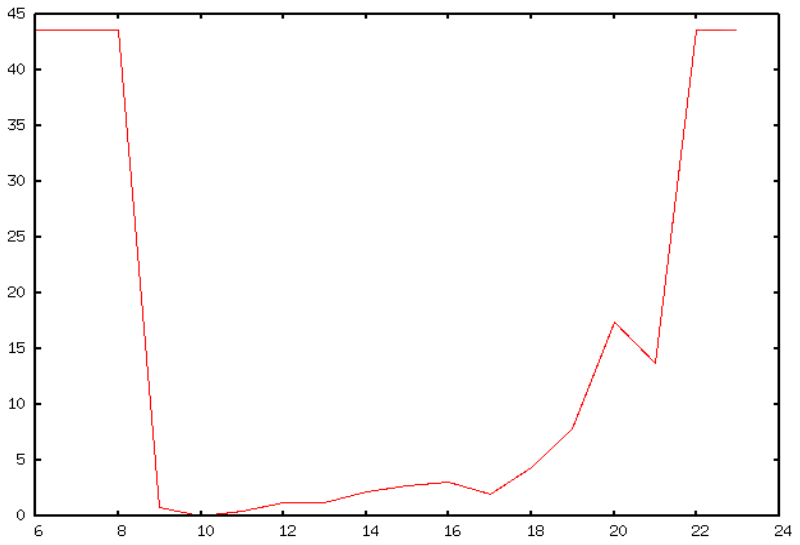
Aβ28-Fe



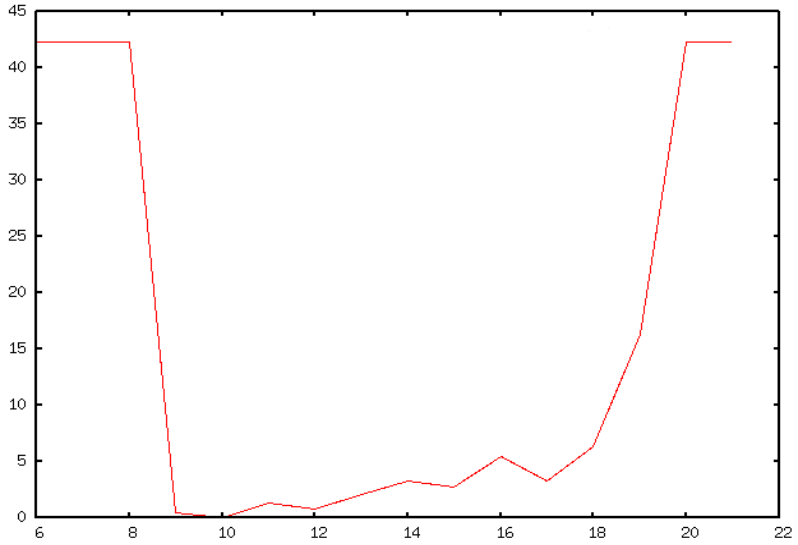
Aβ40-Free



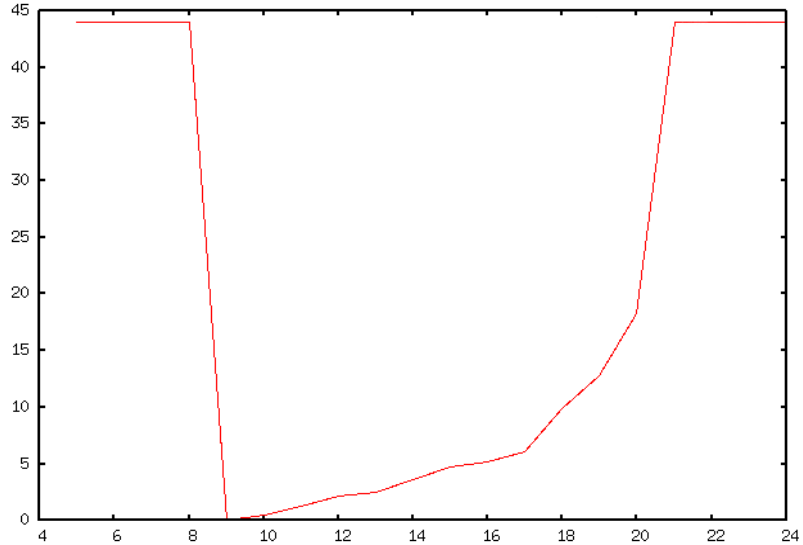
Aβ40-Cu



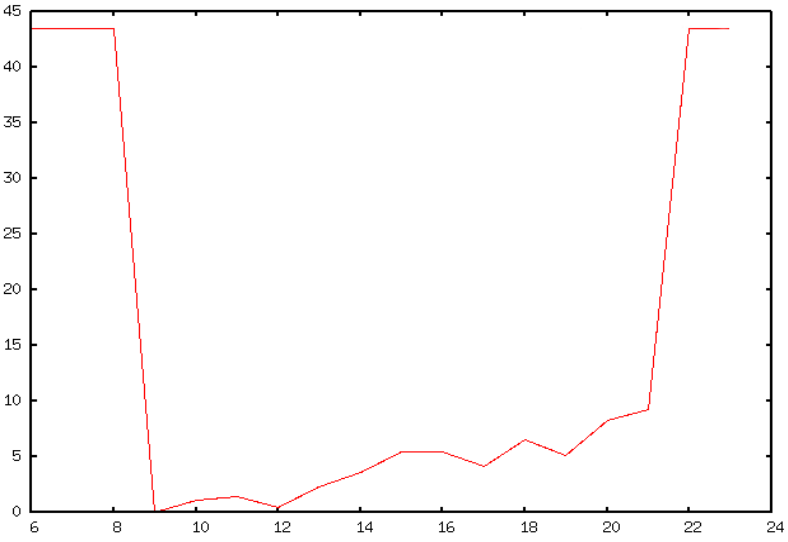
Aβ40-Zn



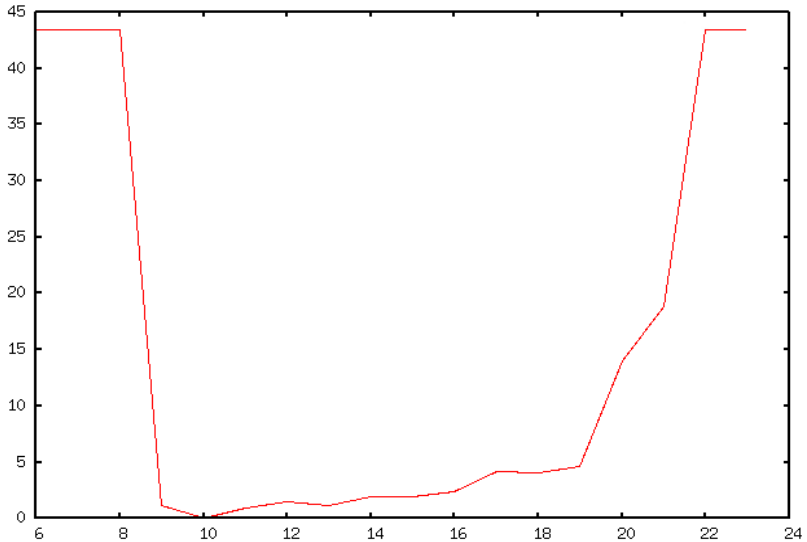
Aβ40-Fe



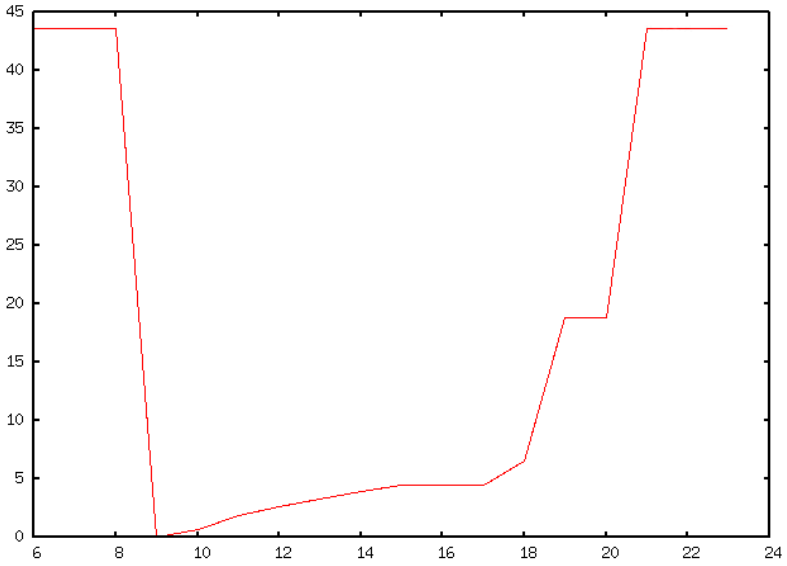
Aβ42-Free



Aβ42-Cu



Aβ42-Zn



Aβ42-Fe

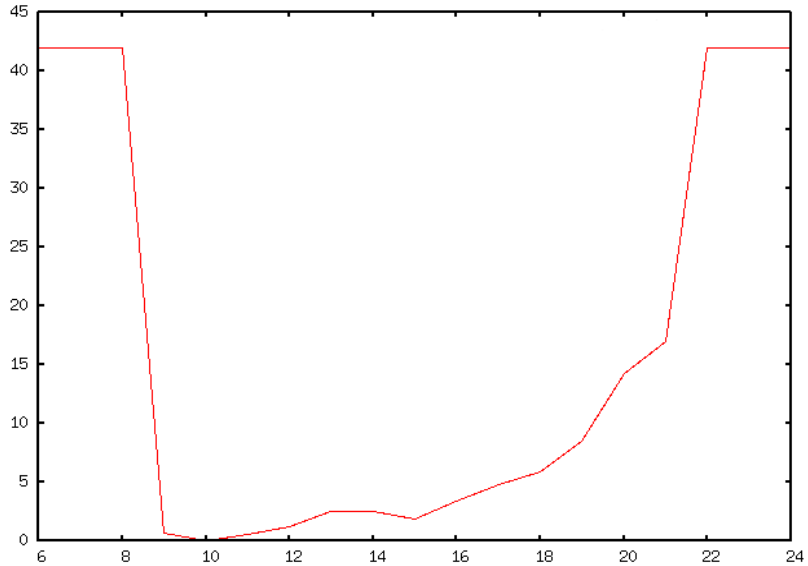
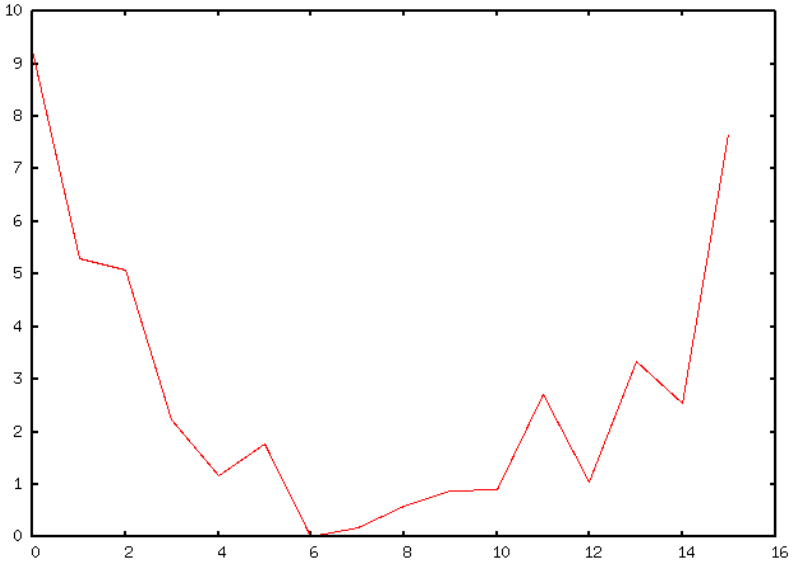
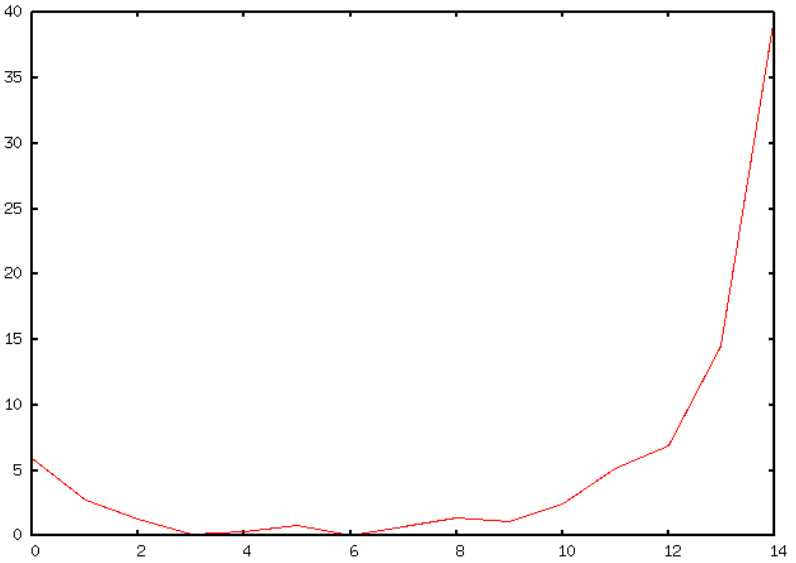


Figure 4.8- Graphs or R_g (X-axis) against potential mean force (PMF) (Y-axis) as a function of free energy

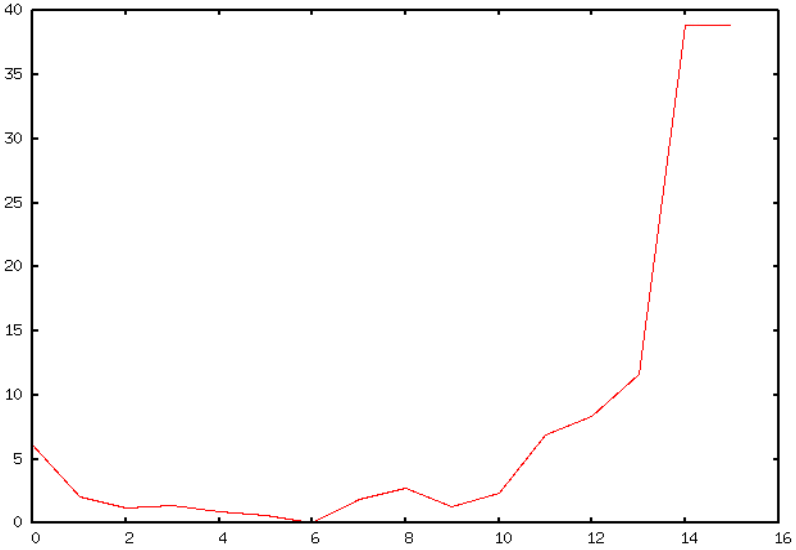
Aβ16-Free



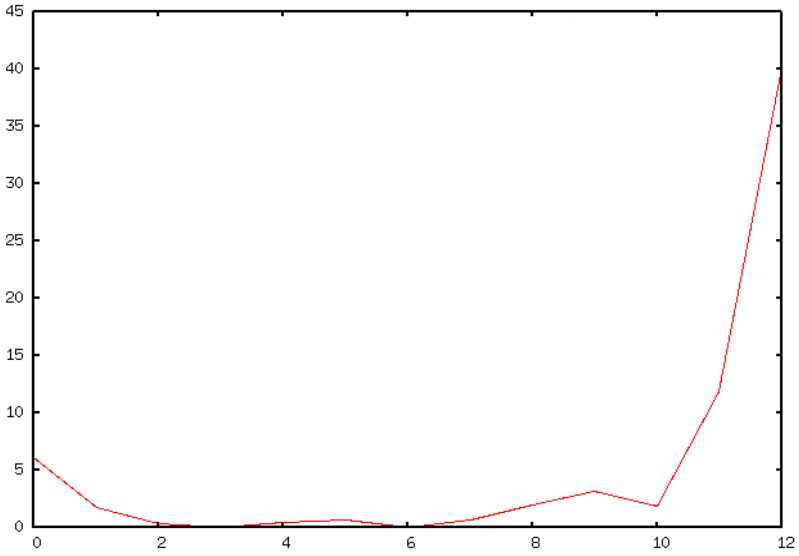
Aβ16-Cu



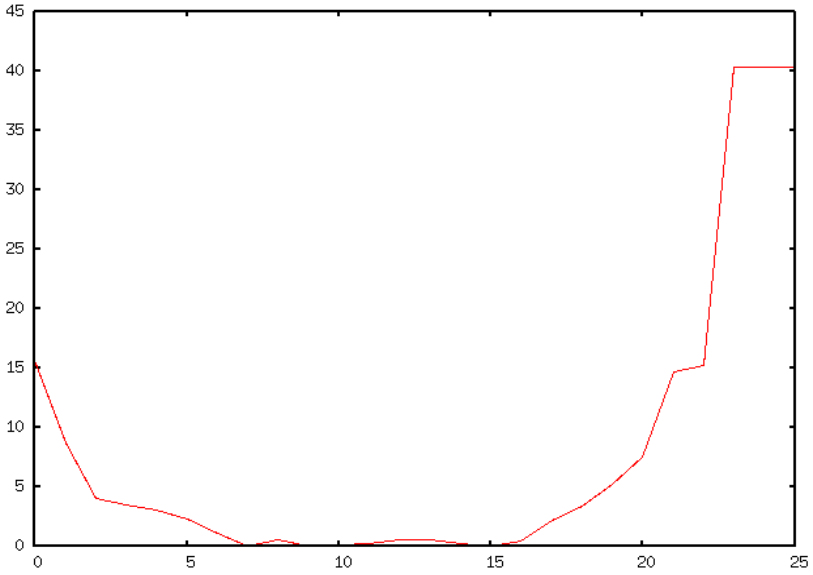
Aβ16-Zn



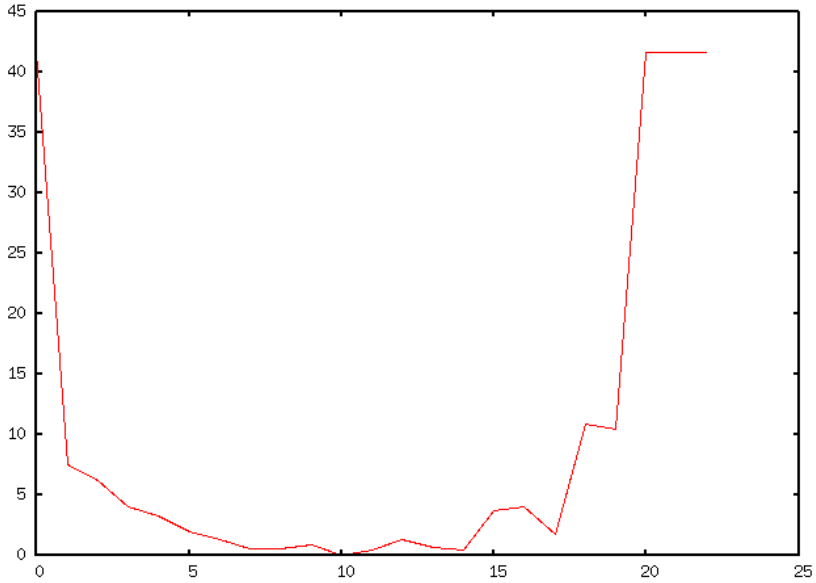
Aβ16-Fe



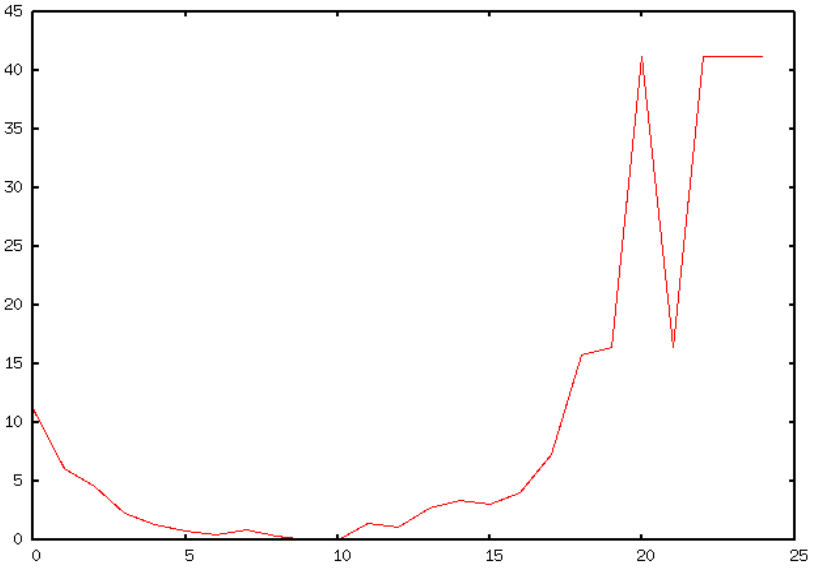
Aβ28-Free



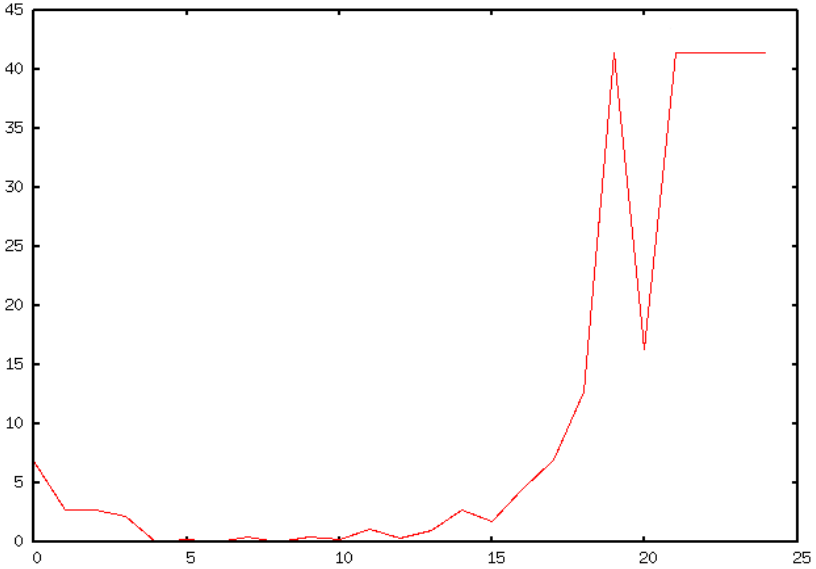
Aβ28-Cu



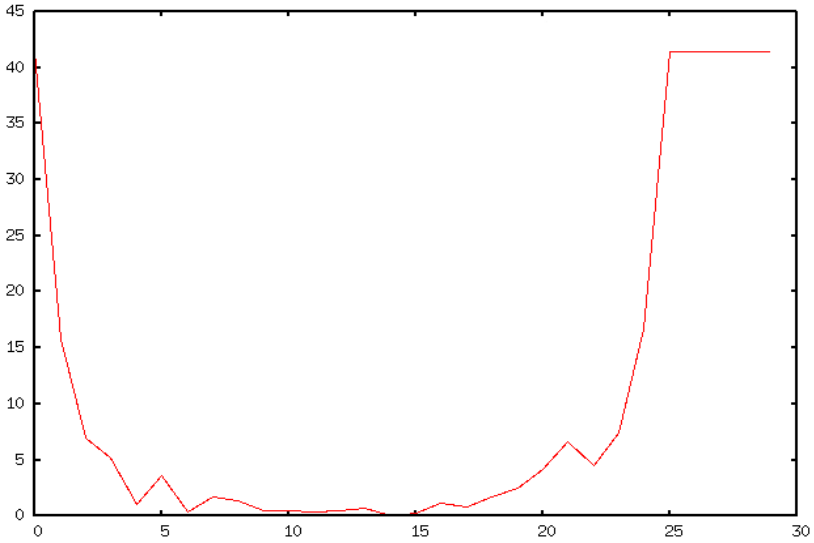
Aβ28-Zn



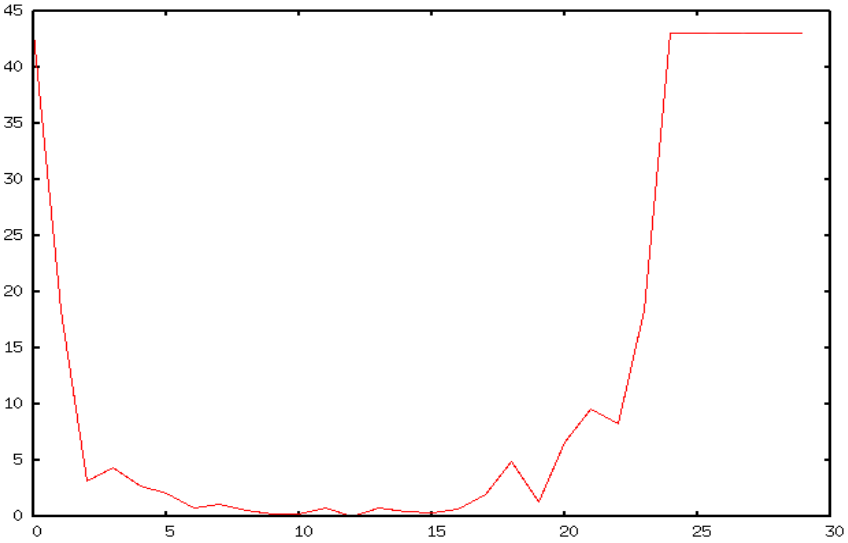
Aβ28-Fe



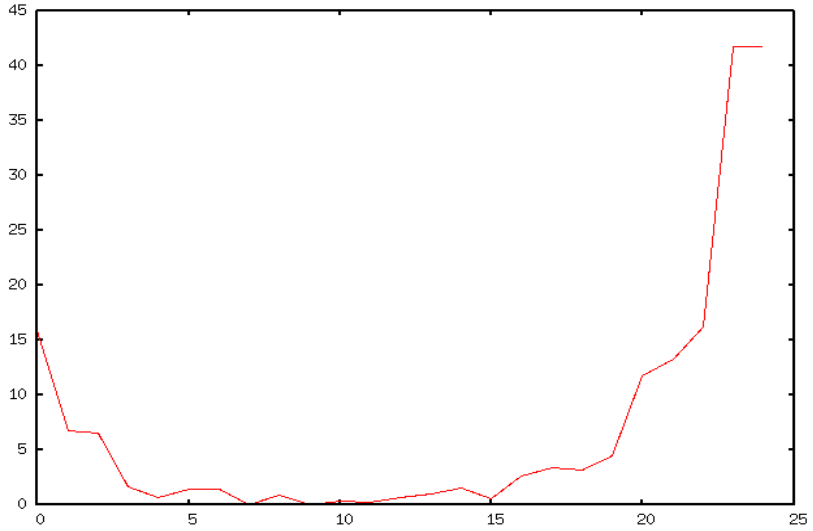
Aβ40-Free



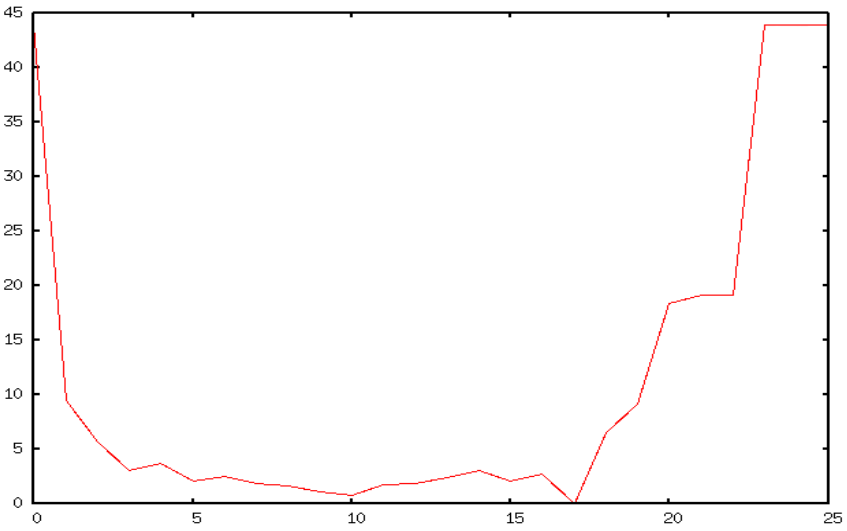
Aβ40-Cu



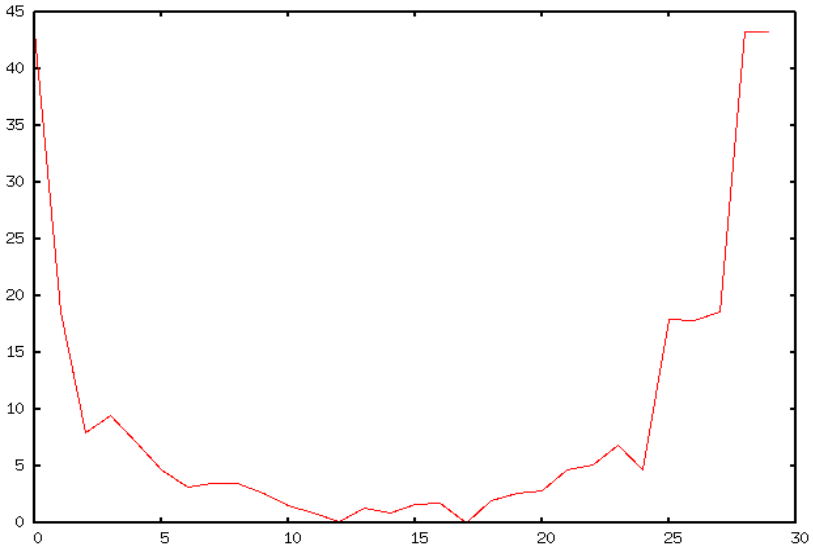
Aβ40-Zn



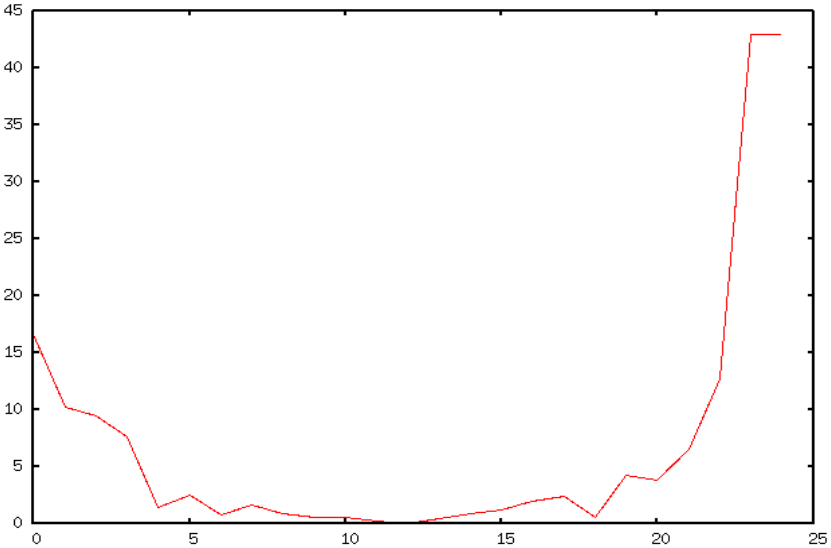
Aβ40-Fe



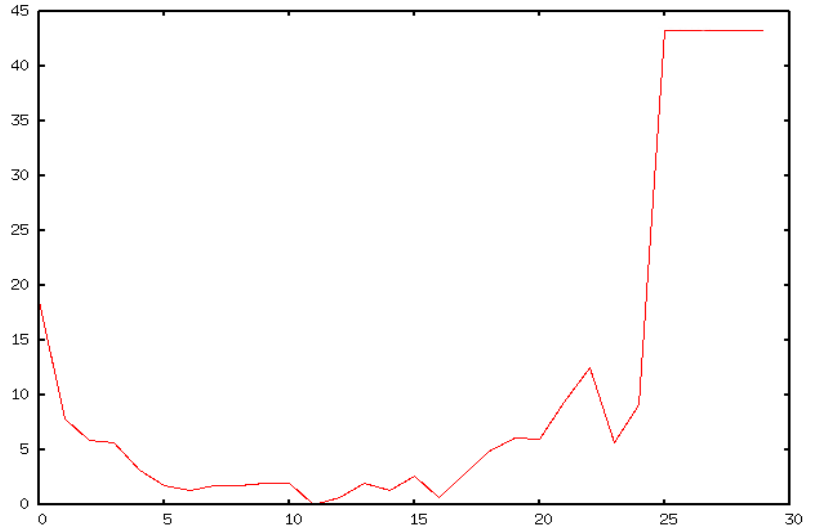
Aβ42-Free



Aβ42-Cu



Aβ42-Zn



Aβ42-Fe

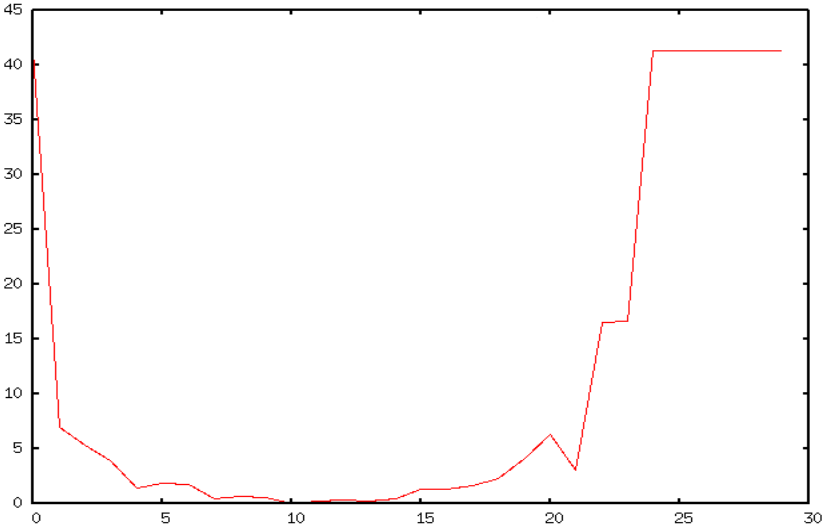
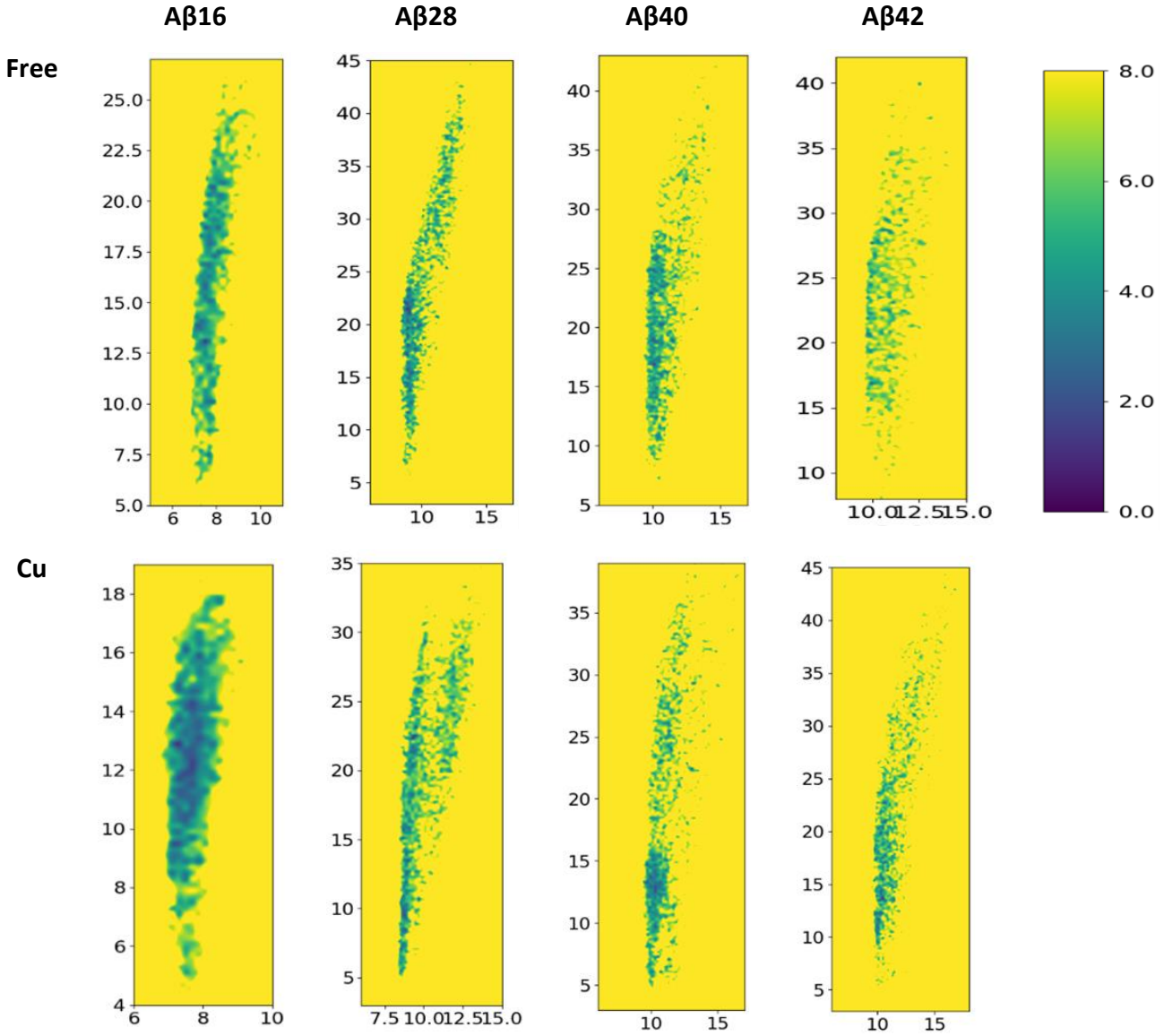


Figure 4.9- Graphs of number of hydrogen bonds (X-axis) against potential mean force (PMF) (Y-axis) as a function of free energy


Aβ16
Aβ28
Aβ40
Aβ42
Free
Cu

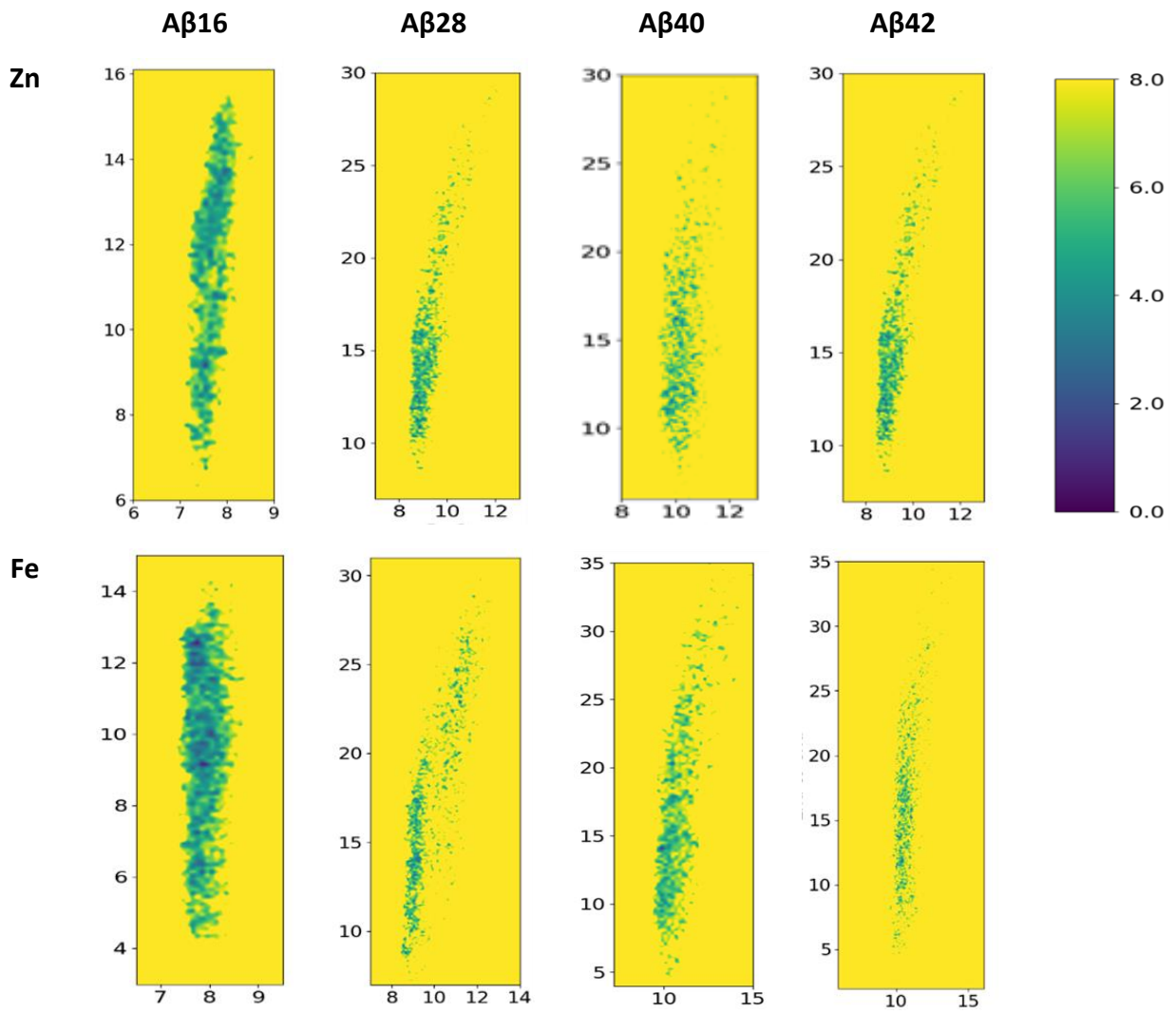


Figure 4.10- 2D plots of R_g (x-axis) against end-to-end distance (y-axis) as a function of free energy shown in the key (kcal/mol).

was observed as displayed from the range of R_g values reported. The number of discrete minima structures displayed within the 2D free-energy plots confirms the need for the boost potential provided via aMD, since conventional MD could be expected to get trapped in low energy minima. Subsequent reweighting was effective at further highlighting the comparisons between structures as a function of free energy when comparing R_g, hydrogen bonds, and end-to-end distances. From this data another example is provided of the comparative results between zinc/iron and copper/free peptide.

4.4 Conclusions

From the results presented here, clear patterns emerge for the similarities in dynamics between two pairs of structures. Zinc and iron-bound peptides showed characteristically similar salt-bridge interactions and secondary structure profiles to one another, in addition to showing similar ranges of data values occupied for minima structures displayed within 2-dimensional free energy plots. Likewise, free peptides and structures bound to Cu(II) displayed comparable results for salt-bridge incidence plots and R_g statistics.

Free peptides were more mobile compared to metal bound ones, due to the lack of anchoring effect on bound residues. RMSF data clearly shows the greatest contribution to this reduced movement is from residues Asp1 and His6, which are bound in all metal simulations; the highest level of mobility is in the unbound Aβ when compared to all other residues. In addition to the R_g data for these metal-free structures, the free peptides showed the highest levels of helical secondary structure as chain length increased. Levels of helical and strand secondary structure were reduced within the N-terminus of all peptides when bound to metal ions, demonstrating the anchoring effect of metal centres in this region. Effects of metal binding were also shown to have a global effect on secondary structure at areas of the peptide remote from those used for coordination.

Since the first proposal of the amyloid cascade hypothesis, it was discovered that metal ions play an important role in the aggregation and formation of characteristic Aβ deposits. Specifically, that due to the high affinity of the transition metals Fe(II), Cu(II) and Zn(II), these ions bind readily to the peptide which leads to an increase in production of reactive oxygen species[50] as discussed in Chapter 1. These in turn encourage formation of

intermolecular bonds between monomeric Aβ peptides, hindering the ability to dispose of any Aβ that has been overproduced, as well as leading to this influx of ROS[51]. Studies suggest Cu(II) competes with Zn(II) ions for similar binding sites of Aβ[52]. Zinc-Aβ complexes appear to aggregate more readily in comparison to these copper-containing metallopeptides, yet at high enough concentrations the rate of aggregation for Aβ-Cu increases to similar levels.

All three metals were analysed having different binding interactions with Aβ to one another, though this does not appear to have caused the simulations to possess motions completely independent of each other. Zn(II) and Fe(II), for example, form 4- and 5-coordinate structures, respectively, but still produce results indicative of similar structures, motions and interactions being sampled over the course of the trajectories.

Accelerated MD proved to be an effective computational method for investigating and comparing the structure and dynamics of peptides of this nature, reporting better conformational sampling compared to conventional MD. Further work could include simulating other peptide structures, different metal centres, binding modes or revisiting mutant peptide structures simulated previously for further comparative studies in order to draw further conclusions on metal-bound structures previously investigated using only conventional MD as well as further highlighting the improvements brought about from use of aMD.

4.5 References

1 P. R. L. Markwick and J. A. McCammon, *Phys Chem Chem Phys*, 2011, **13**, 20053–20065.

2 A. Okur, L. Wickstrom, M. Layten, R. Geney, K. Song, V. Hornak and C. Simmerling, *J Chem Theory Comput*, 2006, **2**, 420–433.

3 R. Geney, M. Layten, R. Gomperts, V. Hornak and C. Simmerling, *J Chem Theory Comput*, 2006, **2**, 115–127.

4 D. Hamelberg, J. Mongan and J. A. McCammon, *J Chem Phys*, 2004, **120**, 11919–11929.

5 J. Talafous, K. J. Marcinowski, G. Klopman and M. G. Zagorski, *Biochemistry-US*, 1994, **33**, 7788–7796.

6 K. Kirshenbaum and V. Daggett, *Biochemistry-US*, 1995, **34**, 7629–7639.

7 T. Kowalik-Jankowska, M. Ruta, K. Wisniewska and L. Lankiewicz, *J Inorg Biochem*, 2003, **95**, 270–282.

8 N. Ntarakas, I. Ermilova and A. P. Lyubartsev, *Eur Biophys J*, 2019, **48**, 813–824.

9 *Molecular Operating Environment (MOE),* 2013, 1010 Sherbrooke St. West, Suite #910, Chemical Computing Group ULC, Montreal, QC, Canada, H3A 2R7

10 E. Atrián-Blasco, M. Del Barrio, P. Faller and C. Hureau, *Anal Chem*, 2018, **90**, 5909–5915.

11 F. Bousejra-Elgarah, C. Bijani, Y. Coppel, P. Faller and C. Hureau, *Inorg Chem*, 2011, **50**, 9024–9030.

12 J. Nasica-Labouze, P. H. Nguyen, F. Sterpone, O. Berthoumieu, N. V. Buchete, S. Coté, A. De Simone, A. J. Doig, P. Faller, A. Garcia, A. Laio, M. S. Li, S. Melchionna, N. Mousseau, Y. Mu, A. Paravastu, S. Pasquali, D. J. Rosenman, B. Strodel, B. Tarus, J. H. Viles, T. Zhang, C. Wang and P. Derreumaux, *Chem Rev*, 2015, **115**, 3518–3563.

13 H. Kozlowski, M. Luczkowski, M. Remelli and D. Valensin, *Coord Chem Rev*, 2012, **256**, 2129–2141.

14 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J Am Chem Soc*, 1995, **117**, 5179–5197.

15 R. J. Deeth, N. Fey and B. Williams-Hubbard, *J Comput Chem*, 2005, **26**, 123–130.

16 D.A. Case, R.M. Betz, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao and P.A. Kollman, 2016, AMBER 2016, University of California, San Francisco.

17 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J Chem Theory Comput*, 2015, **11**, 3696–3713.

18 Zhang, W, Hou, T, Schafmeister, C, Ross, W. S, Case, D. A, 2010, LEaP and gleap: AmberTools Users' Manual Version 1.4,

19 P. Li and K. M. Merz, *J Chem Inf Model*, 2016, **56**, 599–604.

20 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, 2016, Gaussian, Inc., Wallingford CT

21 J. M. Seminario, *Int J Quantum Chem*, 1996, **60**, 1271–1277.

22 W. Clark Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J Am Chem Soc*, 1990, **112**, 6127–9.

23 R. Constanciel and R. Contreras, *Theor Chim Acta*, 1984, **65**, 1–11.

24 M. Schaefer and M. Karplus, *J Phys Chem*, 1996, **100**, 1578–1599.

25 N. Al-Shammari, L. Savva and O. Kennedy-Britten, *Comput Biol Chem,* 2021, **93**, 107540

26 M. Cecchini, R. Curcio, M. Pappalardo, R. Melki and A. Caflisch, *J Mol Biol*, 2006, **357**, 1306–1321.

27 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J Chem Phys*, 1983, **79**, 926–935.

28 A. V Onufriev and D. A. Case, *Annu Rev Biophys*, 2019, **48**, 275–296.

29 R. Anandakrishnan, A. Drozdetski, R. C. Walker and A. V. Onufriev, *Biophys J*, 2015, **108**, 1153–1164.

30 J. A. Izaguirre, D. P. Catarello, J. M. Wozniak and R. D. Skeel, *J Chem Phys*, 2001, **114**, 2090–2098.

31 S. V Lushchekina, E. D. Kots, D. A. Novichkova, K. A. Petrov and P. Masson, *Bionanoscience*, 2017, **7**, 396–402.

32 B. Zhao, M. A. Cohen Stuart and C. K. Hall, *PLoS Comput Biol*, 2017, **13**, 1-22

33 J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J Comput Phys*, 1977, **23**, 327–341.

34 D. R. Roe and T. E. Cheatham, *J Chem Theory Comput*, 2013, **9**, 3084–3095.

35 W. Humphrey, A. Dalke and K. Schuten, *J Mol Graphics*, 1996, **14**, 33–38.

36 W. Sinko, Y. Miao, C. A. F. De Oliveira and J. A. McCammon, *J Phys Chem B*, 2013, **117**, 12759–12768.

37 Y. Miao, W. Sinko, L. Pierce, D. Bucher, R. C. Walker and J. A. McCammon, *J Chem Theory Comput*, 2014, **10**, 2677–2689.

38 P. D. Q. Huy, Q. Van Vuong, G. La Penna, P. Faller and M. S. Li, *ACS Chem Neurosci*,

2016, **7**, 1348–1363.

39 M. Ester, H. P. Kriegel, S. J. and X. Xu, *Proc. 2nd Int Conf Knowl Discov Data Min*, 1996, **7**, 226–231.

40 B. Strodel and O. Coskuner-Weber, *J Chem Inf Model*, 2019, **59**, 1782–1805.

41 S. Vivekanandan, J. R. Brender, S. Y. Lee and A. Ramamoorthy, *Biochem Biophys Res Commun*, 2011, **411**, 312–316.

42 R. Aleksis, F. Oleskovs, K. Jaudzems, J. Pahnke and H. Biverstal, *Biochimie*, 2017, **140**, 176-192

43 A. Schmechel, H. Zentgraf, S. Scheuermann, G. Fritz, R. Pipkorn, J. Reed, K. Beyreuther, T. A. Bayer and G. Multhaup, *J Biol Chem*, 2003, **278**, 35317–35324.

44 K. L. Sciarretta, D. J. Gordon, A. T. Petkova, R. Tycko and S. C. Meredith, *Biochemistry-US*, 2005, **44**, 6003–6014.

45 S. T. Mutter, M. Turner, R. J. Deeth and J. A. Platts, *ACS Chem Neurosci*, 2018, **9**, 2795-2806.

46 X. Zheng, D. Liu, R. Roychaudhuri, D. B. Teplow and M. T. Bowers, *ACS Chem Neurosci*, 2015, **6**, 1732–1740.

47 C. Wurth, N. K. Guimard and M. H. Hecht, *J Mol Biol*, 2002, **319**, 1279–1290.

48 C. Talmard, L. Guilloreau, Y. Coppel, H. Mazarguil and P. Faller, *ChemBioChem*, 2007, **8**, 163–165.

49 O. D. Kennedy-Britten, N. Al-Shammari and J. A. Platts, *J Biomol Struct Dyn*, 2020, **39**, 2003-2013

50 C. Cheignon, M. Jones, E. Atrián-Blasco, I. Kieffer, P. Faller, F. Collin and C. Hureau, *Chem Sci*, 2017, **8**, 5107–5118.

51 J. Alí-Torres, A. Mirats, J. D. Maréchal, L. Rodríguez-Santiago and M. Sodupe, *J Phys Chem B*, 2014, **118**, 4840–4850.

52 K. Suzuki, T. Miura and H. Takeuchi, *Biochem Biophys Res Commun*, 2001, **285**, 991–6

5 Effects of Zn(II) binding on full-length $A\beta_{42}$ dimers using accelerated molecular dynamics (aMD)

5.1 Introduction

Studies reported previously have shown a progression in size of the system investigated. Experiments were originally conducted using conventional MD on truncated Aβ N-termini, focusing on copper-binding within the first 16 residues as well as genetic variants leading to substitutions within this region. Following this, Aβ peptides of 28, 40 and 42 residues in length were modelled using aMD to enhance conformational sampling whilst also comparing differences caused by metal binding of copper, iron and zinc ions.

All previous work had been conducted on individual monomeric Aβ peptides only, so the next logical step would be to investigate multiple structures. Hypothesis on the aetiology of Alzheimer's disease suggests disruption of the homeostatic rate of neuronal Aβ is the causation of gradual accumulation and formation of characteristic plaques[1–4]**.** Interactions and bonds between individual peptides can increase difficulty of clearing excess Aβ produced from the amyloid precursor protein (APP) with increased solubility proving to show correlation with toxicity[5–7].

Alongside the presence of these neuronal Aβ deposits, enhanced levels of transition metals were found within Alzheimer's patients[8,9]. The variable oxidation state and multiple available residues for metal binding leads to a high affinity to form metalloprotein complexes[10,11] further complicating the homeostatic rate of clearance of Aβ and these metals[8,12,13]. Metal ions can coordinate to monomeric units or even form cross-links between Aβ peptides, forming relatively stable dimers leading to greater propensity for aggregation[14–16]. Dimers of Aβ exist readily alongside larger oligomers and fibrils in equilibrium making experimental analysis more difficult[17].

Within this Chapter, results are reported of a comparative computational investigation into the effects of zinc binding on full-length dimeric Aβ against those absent of a metal centre. Similar to previous work conducted on full-length Aβ monomers, aMD was selected in order to explore enhanced conformational sampling, in comparison to conventional MD, as well as being shown in literature as a suitable method for modelling biomolecules of this nature

and size[18,19]. Due to the flexible nature of these proteins, use of an implicit solvent (derived from the Generalised Born Solvation model) was chosen as further means of allowing better transitioning between structures and conformations when compared to similar experiments using explicit solvent[20,21]. Analysis on combined aMD trajectories focused on comparing mobility as well as interactions formed between and within monomeric units of the systems.

Four Aβ starting structures (Figure 5.1) were constructed and used for simulations and subsequent analysis; two coordinated with zinc, and two without. Zinc ions were bound within the N-terminus (as gathered from literature[22]). All four starting structures were simulated for 600 ns each with an aim to combine these to generate 1.2 μs worth of data for analysis for both the zinc and free dimers.

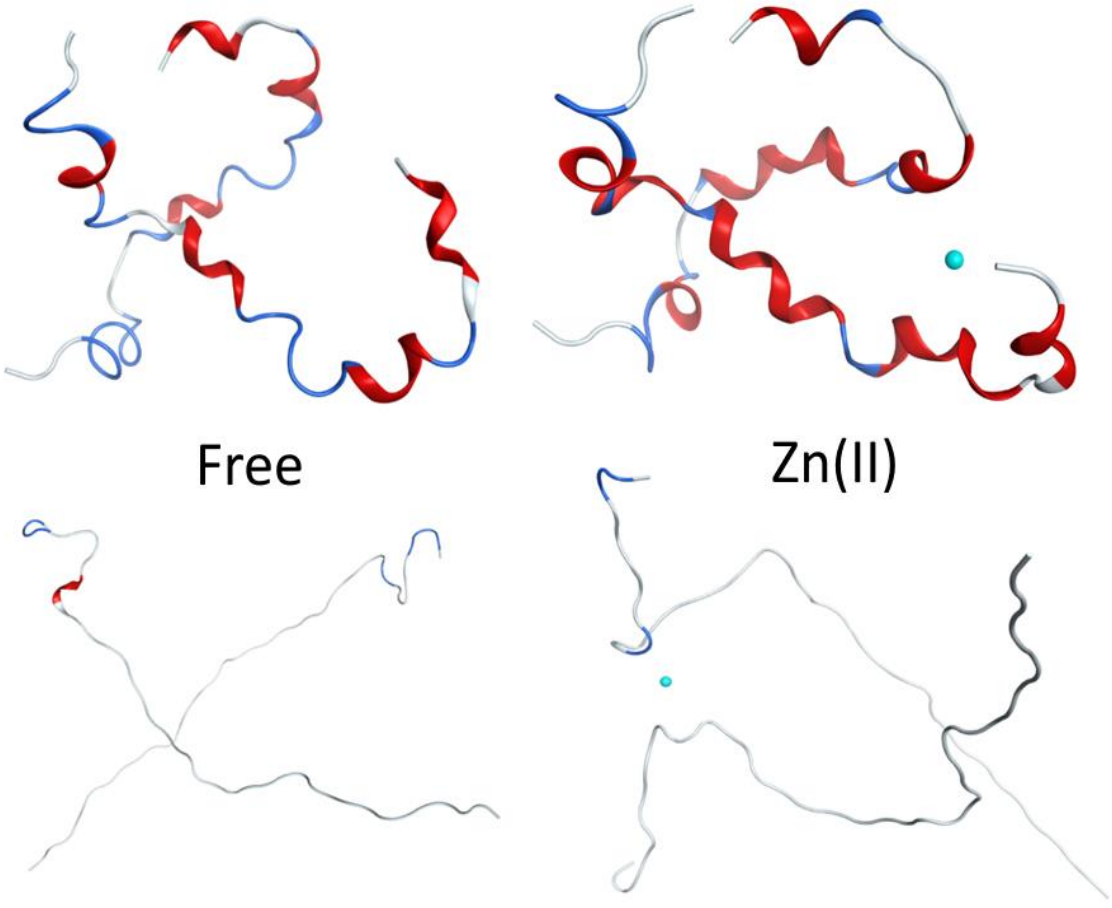


Figure 5.1- Dimeric starting structures of Aβ, helical character defined by areas in red, strand in blue, coil, turn and bend secondary structure in white, and copper ions denoted by a turquoise sphere.

5.2 Computational methods

Two full-length Aβ peptides (42 residues in length) were first manually constructed using MOE[23], in extended conformations, with one pair possessing greater levels of helical character throughout the length of the peptide compared to the other pair when minimised. Separately, these were then duplicated in Packmol[24] to generate two dimeric Aβ structures. From herein, amino acids within chain A of each dimer will be labelled numbers A1-42 and the corresponding residues in chain B will be denoted as B1-42.

To both dimers, a single zinc (II) ion was bound via the corresponding glutamic acid and histidine residues of each monomer (11 & 14 from chain A and 53 & 56 from chain B) to form crosslinks (Figure 5.2), as per the suggested binding method proposed by Kozin *et al.*[22]. The result was four starting structures: two free dimers and two coordinated with zinc. Structures were minimised using both AMBER94[25] and LFMM within DommiMOE[26] to generate starting structures for all subsequent simulations.

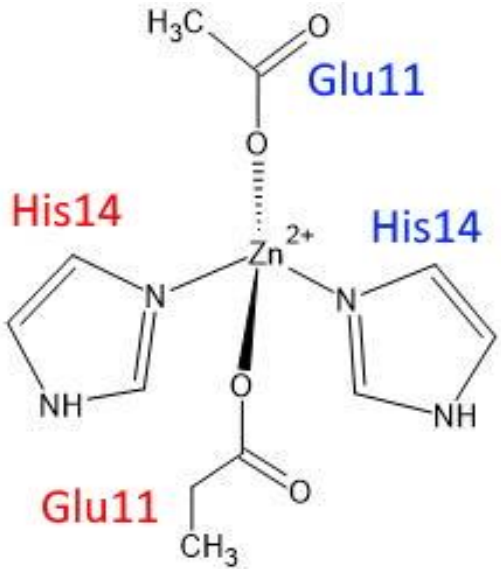


Figure 5.2- Binding mode of Zn(II) ions to chain 1 (blue) and chain 2 (red) showing cross-linkage between the monomers

The AMBER ff14SB[27] forcefield parameters and the LEaP[28] module in AMBER were used to model amino acids, whilst the MCPB.py[29] program was used to generate parameters for zinc ions and associated bound residues. Any bonded and non-bonded parameters for

coordination of the metal centres were taken from B3LYP/6-31G(d) calculations that were carried out using Gaussian09[30]**.** Parameters for angles, torsion and harmonic bond lengths were drawn from the Seminario[31] method whereas non-bonded parameters were extracted from the DFT electrostatic potential from using the RESP[32,33] method.

All MD simulations (both conventional and aMD) were performed using the AMBER16[34] package in implicit solvent utilising the Generalised Born solvation model[35,36]. Previous studies conducted within our research groups show good agreement in MD simulations of Aβ peptides in implicit and explicit solvent[37]. The choice was made to conduct experiments in implicit solvent, to allow shorter simulation time with enhanced conformational sampling for systems of this size and mobility.

Aβ dimers were minimised via 1000 steps of steepest descent succeeded by 1000 conjugate gradient steps. A Langevin thermostat[38] was used to regulate the temperature at 310K, meaning all simulations were performed within the NVT ensemble. Electrostatic interactions beyond 12 Å were disregarded and the SHAKE algorithm[39] was used to maintain bonds to hydrogen atoms within the systems.

50 ns of conventional MD was simulated for all four minimised starting structures, in order to allow the system to reach a point of pseudo-equilibration and to derive bespoke parameters for subsequent aMD calculations, based on the average potential and dihedral energies of each of the four systems. The final structures of all four MD simulations were used as the starting point for subsequent aMD calculations.

For each of the four systems, three aMD simulations of 200 ns in length were performed, each using a different random initial velocity sampled from the Maxwell-Boltzmann distribution, generating twelve sets of aMD trajectories: three trajectories for each of the four structures analysed (calculations used in Chapter 4 (Table 4.1), values for the boost potential used can be found in Table 5.1). The bias potential values only differed slightly between systems due to their similarities in structure, and were applied at points in the Potential Energy Surface (PES) in order to overcome potentially high energy barriers that may have prevented structure sampling via conventional MD.[40]

	EthreshD	EthreshP	alphaD	alphaP
Free Aβ (Run 1)	1267	-1347	47	201
Free Aβ (Run 2)	1246	-1147	48	202
Zn Aβ (Run 3)	1257	-1106	49	202
Zn Aβ (Run 4)	1268	-1098	49	202

Table 5.1- aMD boost potential values as per methods discussed in Chapter 2 derived from initial MD simulations

The three 200 ns aMD simulations were combined for each of the four simulations; the resulting four 600 ns trajectories were then combined into a single dataset based on the presence or absence of a metal centre, meaning the aim was to have two 1.2 μs worth of aMD data; one for the zinc-bound dimers and one for the free-Aβ dimers. This was successful for the zinc dimer but, due to the lack of the bridging ion in the free Aβ structures, one aMD simulation of the metal-free peptides resulted in monomers moving apart. This behaviour was tentatively assigned to the boost potential in aMD biasing the simulation away from close contact between monomers. The 600 ns from run 1 for free Aβ was combined with 380 ns of usable data from run 2 of the free peptide to yield 980 ns of trajectory data for the free dimer. By disregarding the trajectory after this separation occurred in run 2, the 380 ns worth of frames were obtained from a combination of 100 ns, 200 ns and 80 ns from the 3 appropriate trajectories (shown in Figure 5.3) which had no impact on the results or statistics. The initial 50 ns conventional MD data was not used for any further analysis of the systems, all data reported was collected from combined aMD trajectories. CPPTRAJ v16.16[41] and VMD 1.9.3.[42] was used to analyse combined trajectories, focusing on structure and dynamics of the peptides simulated and the results are reported here.

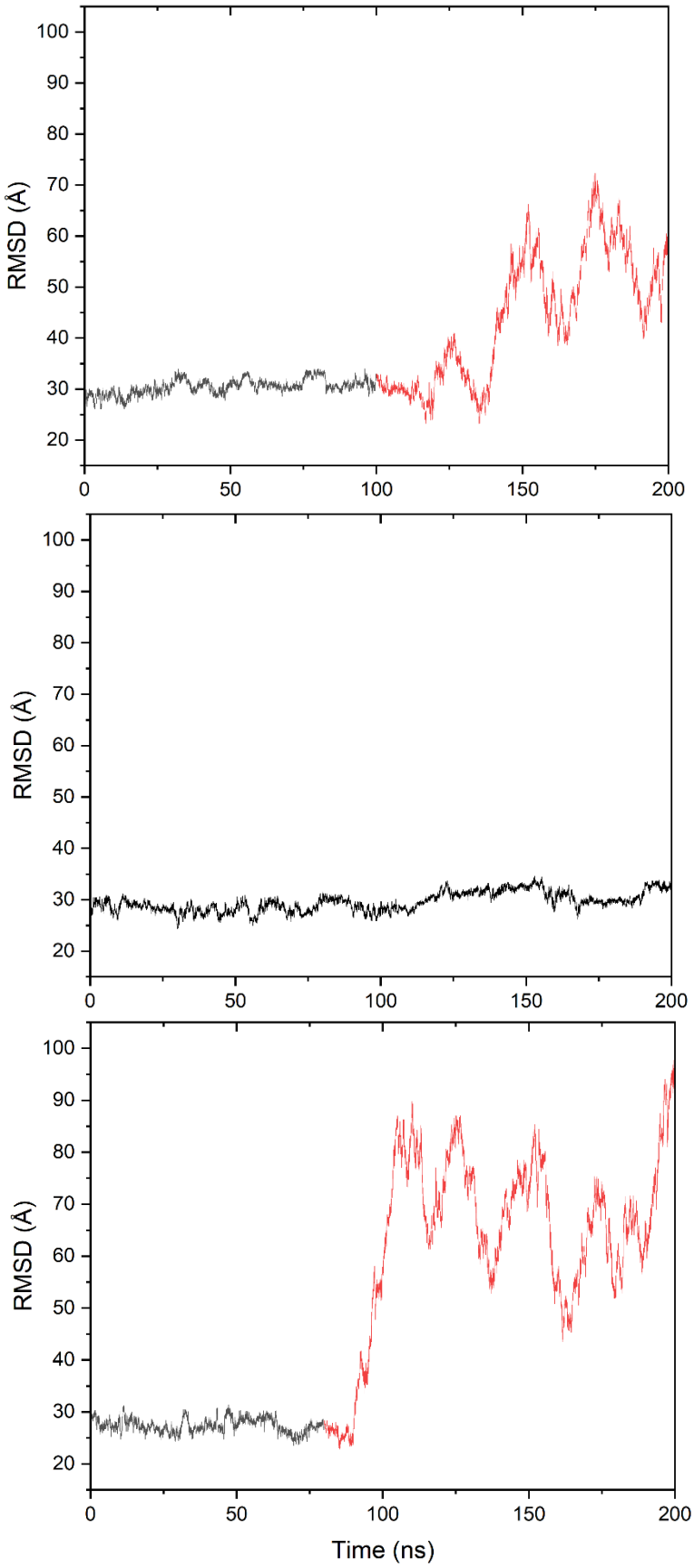


Figure 5.3- Individual RMSD plots of Run 2 for free AB (data in red was discarded and not used for further analysis)

5.3 Results

RMSD values for aMD trajectories (displayed in Figure 5.4 using black) are shown to sample a wider set of conformations compared to conventional MD (in the same figure but in red) without major changes in RMSD compared to the start point, demonstrating the ability of aMD to sample such systems.

RMSD values were relatively consistent, as seen via the small standard deviation values, while the much larger difference between maximum and minimum values shows that sampling during this time was sufficient. The range of RMSD values observed in figure 5.5 show that different conformational ensembles are being sampled for each of the four individual simulations in the combined trajectories, potentially due to the different starting structures for all four trajectories and differing initial velocities.

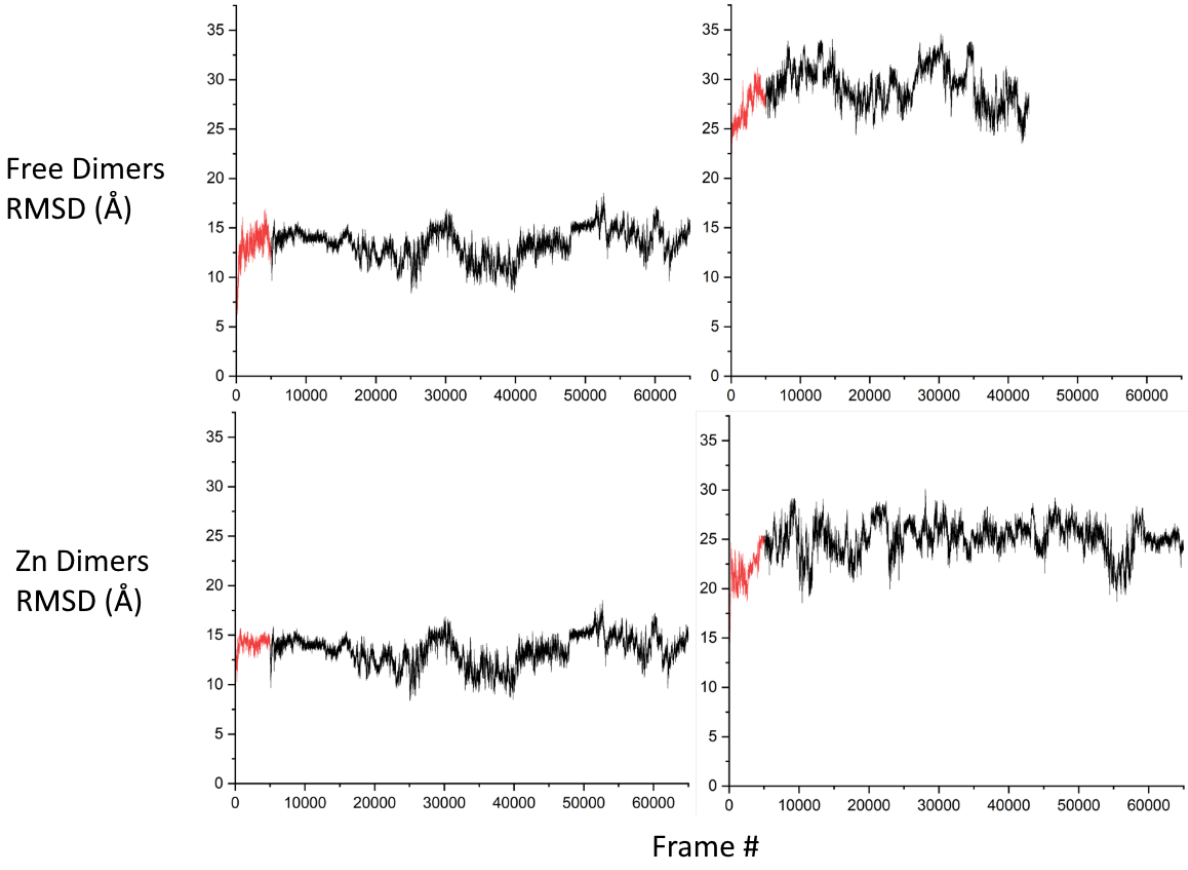


Figure 5.4- RMSD (Å) relative to starting minimised structures for combined MD (red) and aMD (black) trajectories, for all four Aβ dimer structures. Frame numbers are plotted on x-axis as start point was the same for each respective combined simulation and is not one continuous time-series.

RMSD (Å)		Avg	Min	Max	SD
Free Dimers	**Run 1 (600 ns)**	13.99	8.84	22.28	1.56
	Run 2 (380 ns)	29.34	23.50	34.60	1.97
Zn Dimers	**Run 3 (600 ns)**	13.41	8.40	18.54	1.49
	Run 4 (600 ns)	25.15	18.55	30.12	1.61

Table 5.2- RMSD (Å) statistics relative to initial minimised structure for combined aMD trajectories

Radius of gyration (R_g) values (Table 5.3) show statistics for combined aMD trajectories, with 980 ns for free dimers and 1200 ns worth of data for zinc-bound structures. Statistical data for R_g on combined aMD simulations (Figure 5.5) shows zinc-bound dimers have lower R_g values than their free-peptide counterparts, highlighting the effect of metal-coordination on the size of these structures. It was anticipated that the binding of cross-linking zinc ions, within the N-terminus of the two monomers in each system, would limit the flexibility and mobility of the dimers compared to the metal-free peptides. It is also apparent that the standard deviation of R_g is smaller for the Zn-bound dimers, as is the maximum value. In contrast, the minimum value of R_g is similar between systems.

R_g (Å)	***Avg***	***Min***	***Max***	***SD***
Free-Aβ dimer (980 ns)	15.8	12.2	24.5	1.8
Zn-Aβ dimer (1200 ns)	14.9	11.9	20.8	1.3

Table 5.3- Statistical data for R_g (Å) of dimeric Aβ aMD simulations

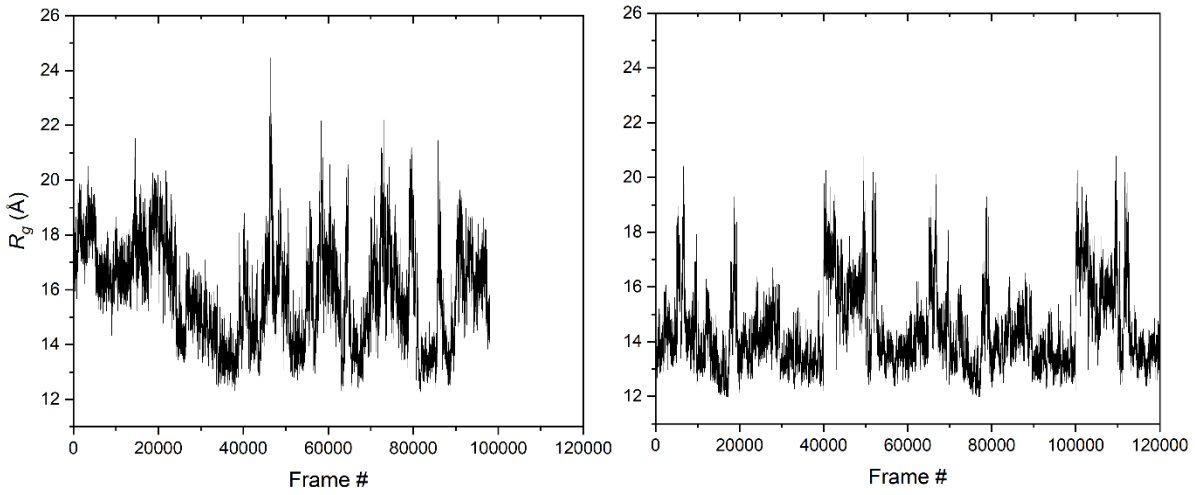


Figure 5.5- Graphical representation of R_g data (Å) for aMD trajectories on free-Aβ and zinc-bound dimers (left and right, respectively)

Salt-bridge profiles (Figure 5.6) show residues within proximity to one another and thus allow a better insight to conformations adopted, as well as representing interactions within and between monomeric chains of the dimers. Similarities were observed between the salt-bridge profiles of the zinc and the metal-free dimers for the types of interactions sampled, but also differences observed in percentage incidence of contacts that are common between systems.

Out of 72 possible types of salt-bridge interactions that could form for the simulated peptides, the most common salt-bridge interactions in the free dimers were monomeric between A-Glu11Arg5 (53% of simulation time) and B-Glu11Arg5 (45% of simulation time). Interestingly, these residues are in the same position of their respective monomer highlighting this as a key interaction within the Aβ amino acid sequence.

For the metal-bound species, A-Glu3Arg5 is the most prevalent salt-bridge contact (observed for 42% of frames). The equivalent residues within chain B of the same dimer, B-Glu3Arg5, were observed to be in proximity of one another for 33% of simulation time. Interestingly, the most frequent salt-bridge contact observed in the Aβ-Zn dimer was only observed for 26% of simulation time in the free-Aβ trajectory. This also held true for the most common interactions in the free peptide mentioned previously, which only existed for 24% and 0.3% (for A-Glu3Arg5 and B-Glu3Arg5, respectively) of simulation time for the zinc-containing trajectory. The presence of a zinc metal centre appears to disrupt interactions between residues within the N-terminus of Aβ; this was most apparent with residues Glu11 and Glu53, which are directly involved in metal coordination and are seen to form a high percentage of salt-bridges in the free Aβ structures. Three combinations of residues are never in contact for any amount of simulation time for both structures: Asp23Lys16; Asp23Lys58; and Glu53Lys70.

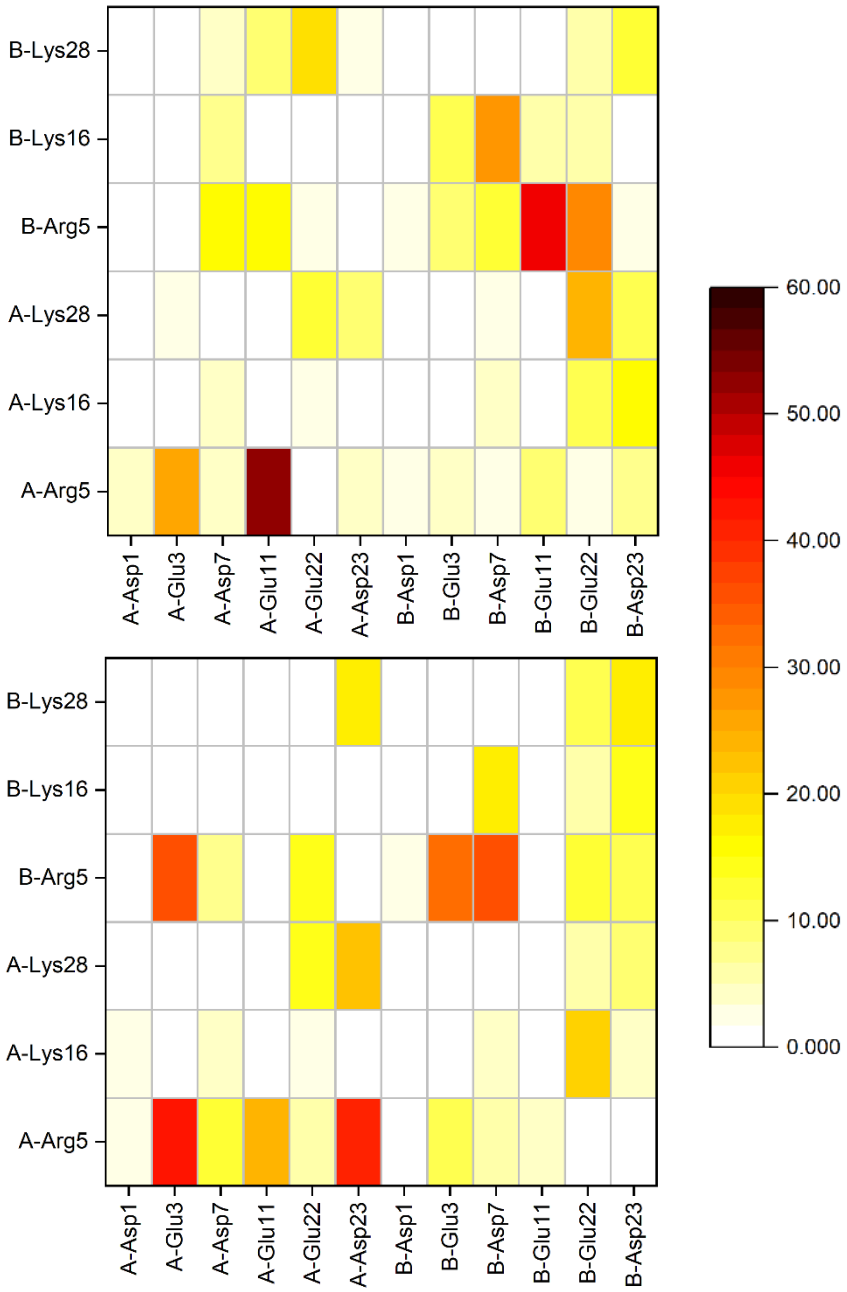


Figure 5.6- Salt-bridge incidence plots (% time) by residue for free-Aβ and Aβ-Zn (Top and bottom respectively)

Overall, a greater incidence of salt-bridge formation was observed intramolecularly; within monomeric units, as opposed to between peptide chains of each dimer. Between the two systems though, generally a greater incidence of intermolecular contacts occurs in the zinc-containing structures; this is displayed by interactions such as between A-Glu3-B-Arg5, which rises in incidence by 36% in the zinc-dimer compared to the metal-free structures.

Glu11 for monomer chains A and B see a decrease in the frequency of salt-bridges formed overall in the zinc simulations, due to involvement directly in the coordination to the metal centre. The anchoring effect of the bound zinc limits its mobility, and hinders interactions with a greater number of residues within the N-terminus, where the metal-binding region exists.

Enhanced propensity for formation of potentially neurotoxic alloforms of Aβ has been attributed in part to the secondary structure adopted by this peptide. Heightened levels of insoluble, strand-like structures have been associated with aggregation into oligomers and neuronal plaques associated with diagnosis of AD. Secondary structure incidence values (Table 5.4) and plots (Figure 5.7) display secondary structure as a percentage of simulation time by residue categorised by either “strand”, “helix”, or “other”.

% of simulation time	***Free Dimers (980 ns)***		
	Full Dimer	***Chain A***	***Chain B***
Helix	46.2	45.7	46.9
Strand	3.1	3.2	3.1
Other	50.7	51.1	50.0
	Zinc Dimers (1200 ns)		
	Full Dimer	***Chain A***	***Chain B***
Helix	32.0	30.6	33.4
Strand	2.4	2.4	2.4
Other	65.6	67.0	64.2

Table 5.4- Secondary structure of equilibrated aMD trajectories by percentage (%)

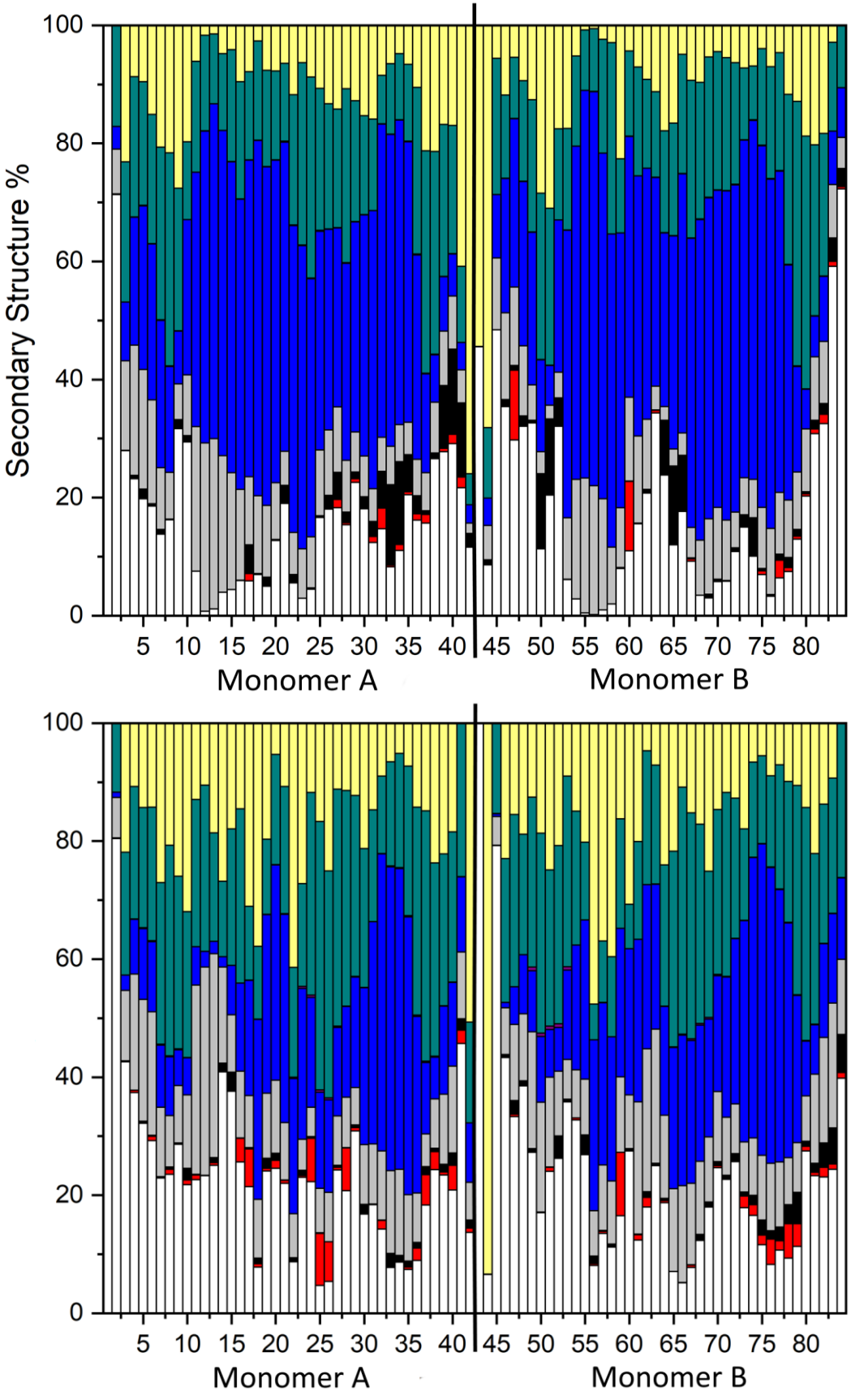


Figure 5.7- Secondary structure plots for free-Aβ and Aβ-Zn dimers (Top and Bottom respectively)

Coordination to a zinc ion leads to a decrease in the levels of helical and strand secondary structures observed via aMD, which could be due to limited mobility upon the presence of a metal centre, limiting certain conformations. A helical structure is generally observed throughout the entirety of the Aβ dimers, but lower levels are observed within residues located at the N-terminus (residues 1-16), where binding to the metal centres occurs.

Monomeric units tend to adopt similar secondary structure types within simulations, showing that chains within dimers generally sample similar geometries and conformations as each other.

Root-mean square fluctuation values are used as a measure of mobility for each residue. Comparing RMSF values (Å) between trajectories shows generally comparative trends in levels of flexibility per residue across the length of the dimers (Figure 5.8); both free dimers and zinc-bound structures have similar RMSF values for both of their respective simulations. This is a good indication of sufficient sampling in terms of both the length of the trajectories and also the number of starting structures used.

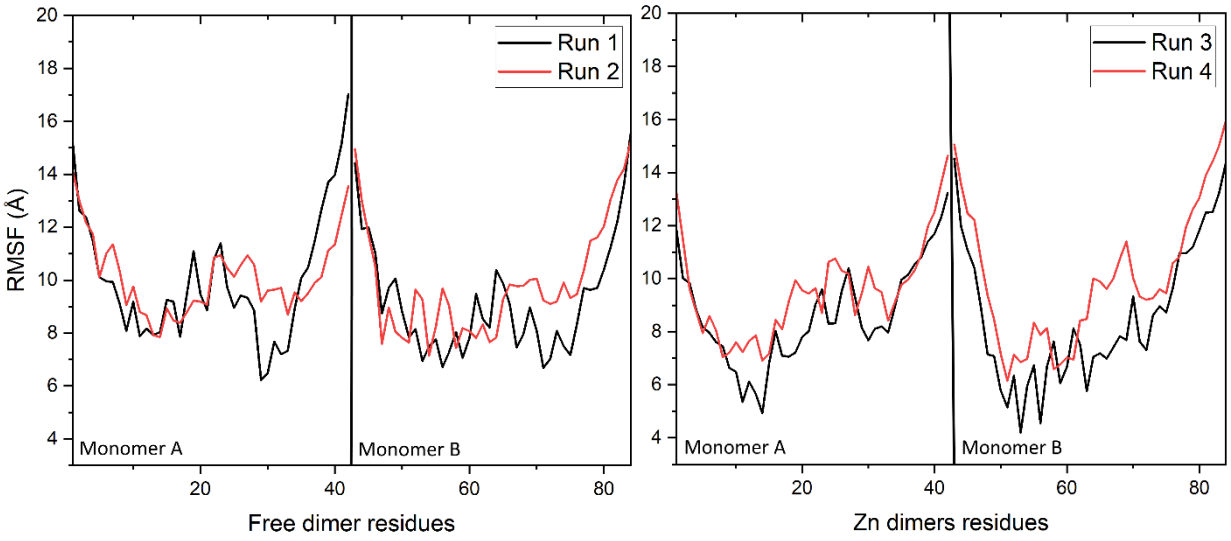


Figure 5.8- RMSF (Å) of four individual aMD trajectories comparing runs within free and zinc dimers

Figure 5.9 shows combined RMSF values comparing both trajectories for free Aβ against those containing a zinc ion. Generally, residues in free Aβ structures displayed higher RMSF values than their metal-bound counterparts. The highest RMSF values were observed at the

terminal residues (Asp1 and Ala42 for both monomer chains), whilst the least mobile residues were observed between around Glu11-Leu17 of each monomer. This is significant, especially for the zinc-bound dimers, due to this being the location for the coordination site for the metal-centre (residues Glu11 and His14 of each respective chain). The decreased flexibility within this region is attributed to the presence of a metal centre, which further evidences disruption of dynamics within the N-terminus usually observed within free Aβ peptides. RMSF plots for all four individual trajectories highlight the similarities in mobility between the two aMD simulations containing a zinc ion, and shows comparable values in those absent of metal. These similarities are good evidence to suggest that trajectories were ran for a sufficient amount of time to show comparable dynamics and interactions.

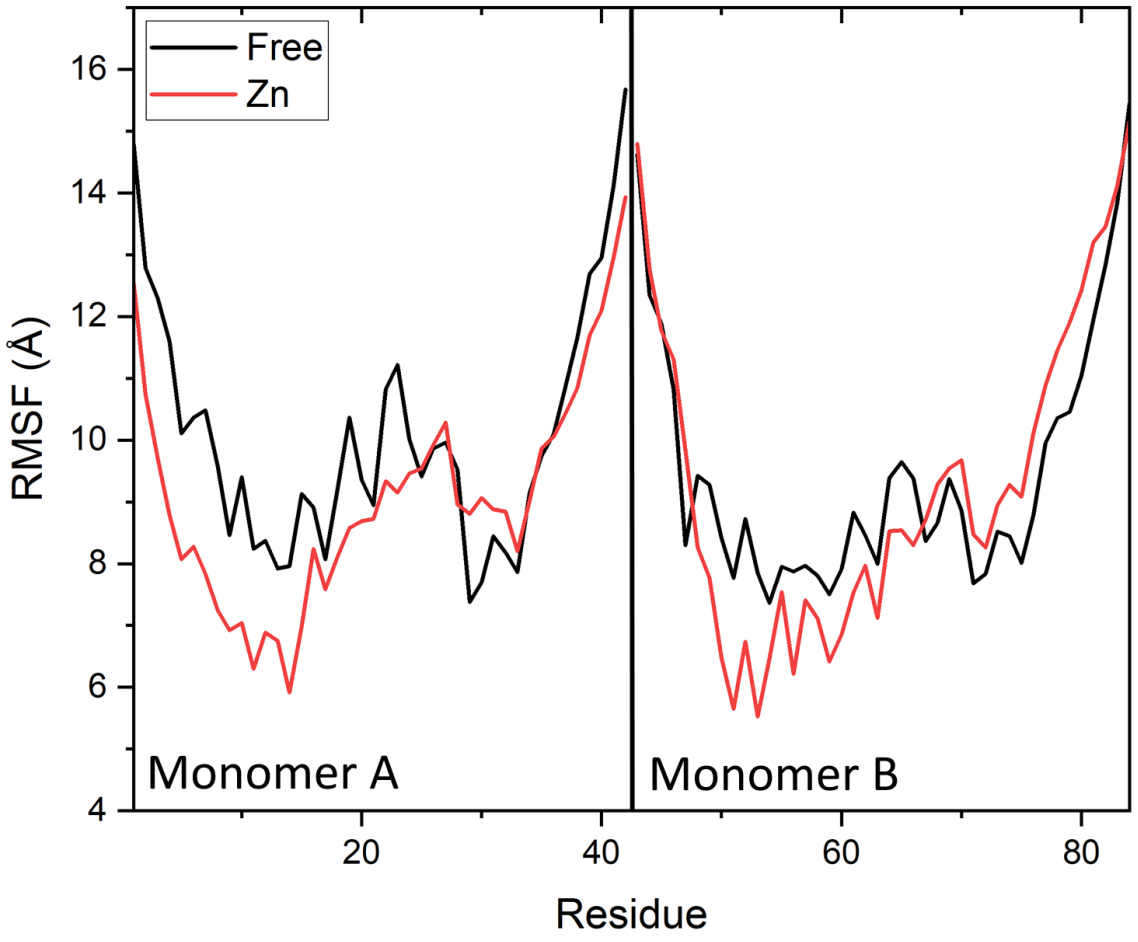


Figure 5.9- RMSF (Å) of four combined aMD trajectories comparing combined free and zinc trajectories (with Zinc being bound to Glu11, His14 in both chains A and B)

Hydrogen bond plots add further evidence of contacts and interactions between residues based on atomistic properties as either a donor or an acceptor. These plots indicate that these systems are statistically similar in terms of the quantity and frequency of hydrogen bonds occurring within these dimers (Table 5.5 and Figure 5.10 with original hydrogen bond output values in Appendix A2.). Looking at potential combinations of atoms within these Aβ peptides, the free dimer forms a greater number of intermolecular interactions between monomers, and the converse is true for the zinc-bound peptides, forming more hydrogen bonds within each individual monomer. This could be potentially due to the nature of the metal centre anchoring the N-terminus, restricting movement and the ability to interact with its monomeric counterpart.

Hydrogen bonds	Avg	Min	Max	Standard Deviation
Free Dimers	25.3	9.0	45.0	4.4
Zn Dimers	24.3	8.0	45.0	4.4

Table 5.5- Statistical data for hydrogen bond plots

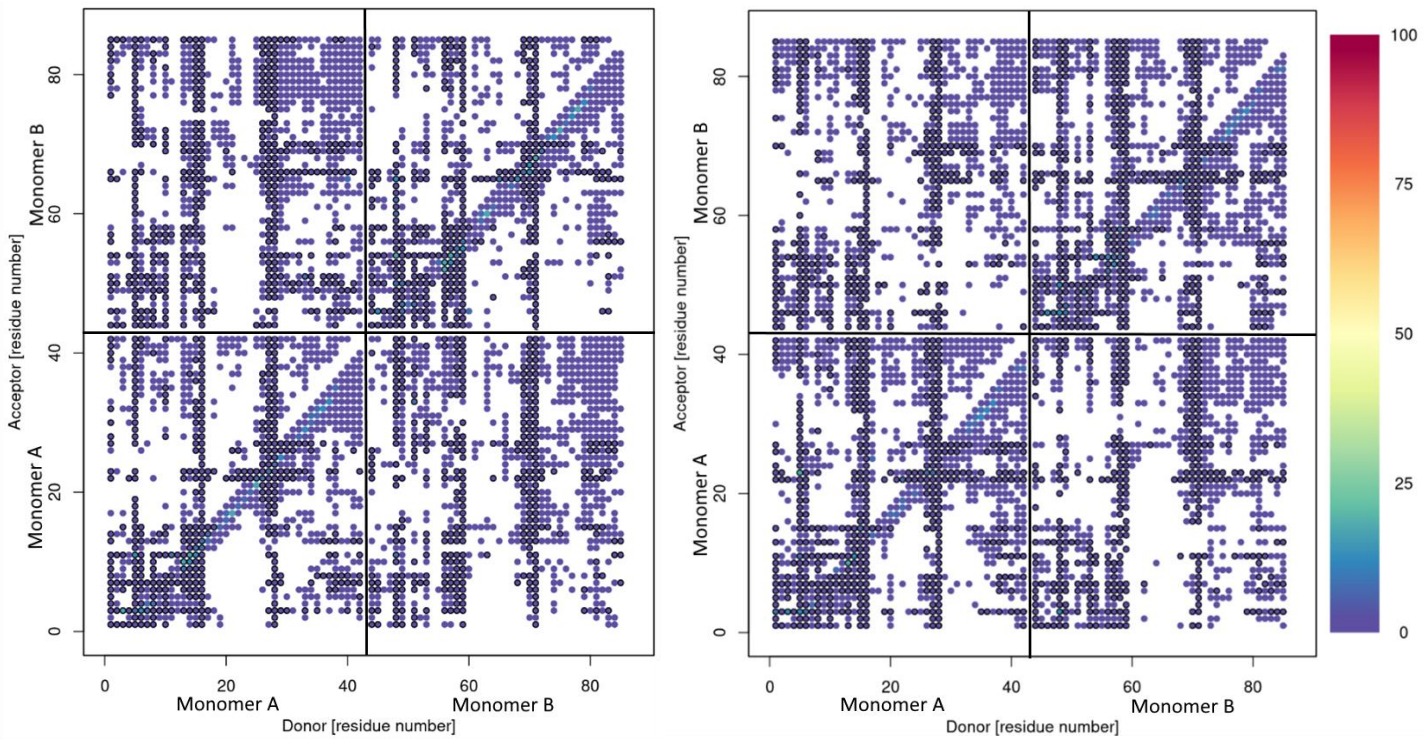


Figure 5.10- Hydrogen bond plots for free and zinc-dimers (L-R respectively)

Using the DBSCAN clustering algorithm[43] with a cut-off distance of 4 Å, structures characterised as a cluster were defined. Structures were grouped based on RMSD values within 4 Å of each other, based on backbone atoms, to identify similar conformations being sampled over the course of the trajectories. The highest populated clusters are represented in Figure 5.11 and Table 5.6.

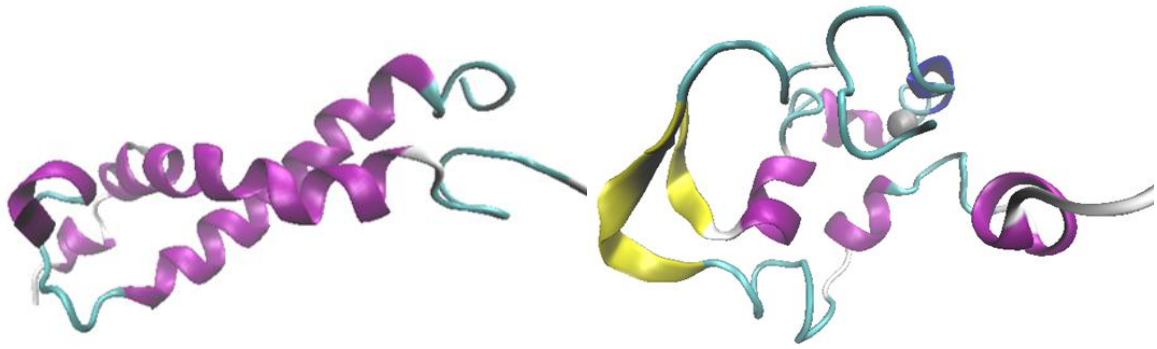


Figure 5.11- Representative structures of highest populated clusters for free and zinc Aβ dimers (L-R, respectively).

	Primary populated cluster (%)	***Secondary populated cluster (%)***	***Total number of notable clusters***
Free Dimers (980 ns)	22.7	12.0	11
Zn Dimers (1200 ns)	19.8	10.1	15

Table 5.6- Percentage of highest populated clusters for entire combined simulations

A cut-off of clusters that are present for at least 1% of frames was selected for all datasets; 11 notable clusters were thus identified for the free dimers, whereas 15 were sampled for the Aβ-Zn complexes. Free dimers possess fewer clusters but these exist for a greater length of simulation time in comparison to their zinc-bound counterparts.

Interestingly, the most common cluster of the free dimer, which existed for nearly a quarter of simulation time, showed each monomer forming helices throughout most of the structure, but predominantly within the hydrophobic central region. The primary cluster for the zinc-bound dimer also showed helical character but at much lower levels, as well as some β-strand, as supported by the previous secondary structure incidence plots.

Solvent accessible surface area (SASA) analysis can be used both in conjunction with RMSD and RMSF data as a measure of pseudo-equilibration, but also gives insight to atoms and residues which are interfacing between monomers, which can lead to inferences on key interactions linked to aggregation. Values for the free peptides compared to the zinc-bound dimer are displayed in Figure 5.12 and Table 5.7. Generally, chains in the free dimers have comparable SASA values, suggesting similar dynamics and motions; similarities are also observed between the two Aβ monomers in the zinc dimers. The results can be compared to RMSF data, which provides further evidence that the simulations were run to a point that sufficient sampling was achieved.

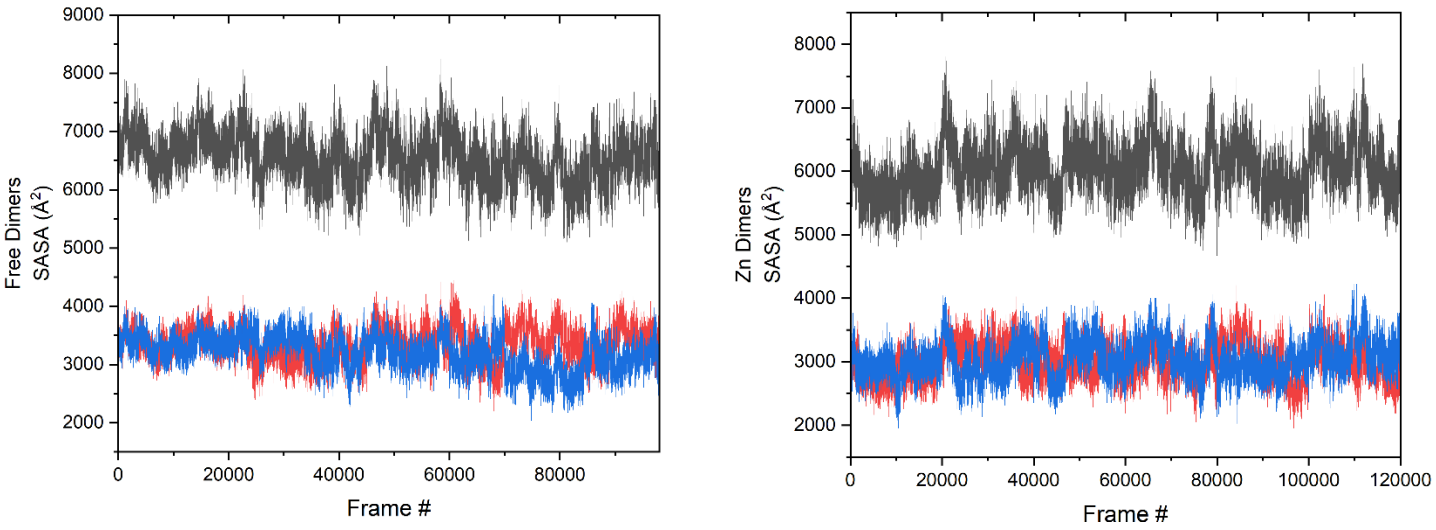


Figure 5.12- Solvent-accessible surface area (SASA, Å^2) of free and zinc-bound Aβ dimer trajectories (Left and right respectively) representing monomer chain A (red), chain B (blue) and the entire dimer structure (black)

Slightly lower values were observed for the dimers containing zinc and, when used alongside the clustering data, clear evidence emerges of the flexible nature of these systems. Primary populated clusters for free Aβ form helices in more extended conformers when compared to the zinc-bound peptides, which show more compact structures. Thus, effects of reduced solvent accessibility to some atoms within the SASA results were clear, particularly those at the interface between dimers; however, differences between chains or

between free and Zn-bound dimers are relatively small, and in all cases are within one standard deviation of each other.

SASA (Å²)		*Avg*	*Min*	*Max*	*SD*
Free Dimers (980 ns)	**All**	6540	5106	8252	404
	Chain A	3347	2201	4446	275
	Chain B	3193	2041	4217	292
Zn Dimers (1200 ns)	**All**	6023	4666	7805	399
	Chain A	2995	1958	4204	273
	Chain B	3027	1950	4225	281

Figure 5.7- Statistical analysis of solvent-accessible surface area (Å^2).

Contact maps visualising distances between residues show clear differences in mobility and dynamics of the structures as plotted in Figure 5.13. On average, free dimers displayed a greater range of distances, showing the enhanced lability and flexibility of this system in the absence of a metal centre. As expected, the greatest distances were observed between the N- and C-terminal residues of each chain in the free Aβ, highlighting the mobile nature of these regions, adopting more extended structures as opposed to folding or any direct contact between terminal residues. Taking cluster conformations into account as well, it appears that the helices formed throughout the free structures didn't allow for much interaction between these terminal residues. The same terminal residues in the zinc-Aβ complexes (Asp1 and Ala42) show comparable values to the rest of the amino acid sequence, showing potentially more compact conformations being adopted which is supported by lower R_g for zinc-dimers; due to effects of the metal centre on the whole length of the peptide, including areas remote to the coordination site. Overall, the dimers containing zinc showed consistently lower distances which, in conjunction with R_g data, indicates more compact structures being adopted.

Regions of close contact were observed approximately between residues Glu11 of chains A and B within the zinc dimers, due to the bridging nature of the metal ion centre;

interactions of a similar percentage incidence were also seen around residues Met35 in chain A and Gly33 in chain B of free-Aβ. Lack of a cross-linking ion here suggests residues A-Met35 and B-Gly33 could be linked to aggregational properties of Aβ. This potentially indicates the zinc-bound and free peptides demonstrate different mechanisms of binding and aggregation of monomeric units to one another.

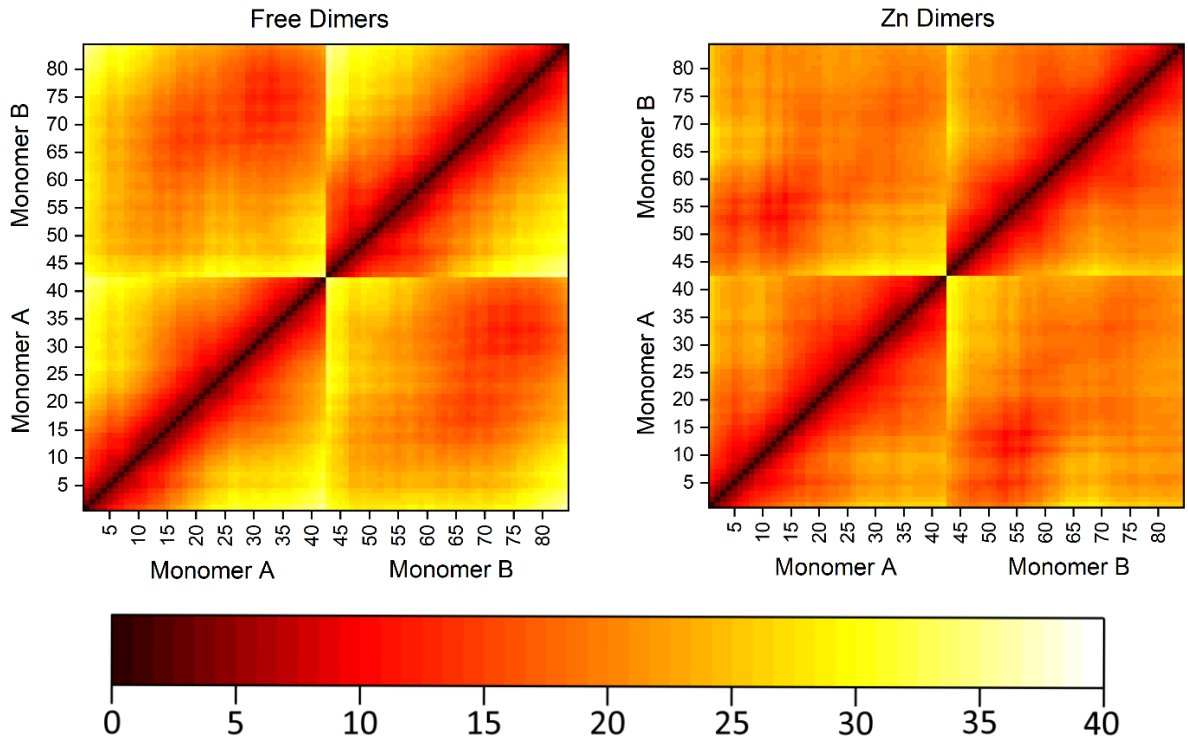


Figure 5.13- Contact maps (Å) for free and zinc dimers (L-R respectively)

Interactions between monomers differ between the two types of dimers studied here. For the zinc-bound peptides the lowest intermolecular contacts occur within N-terminus residues, due to the metal centre restricting movement within this area. The opposite is true for free Aβ, where closer contacts are observed between residues in the central hydrophobic region and the C-terminus, in agreement with data that suggests the presence of an increased helical secondary structure in clusters and secondary structure plots for these amino acids.

5.4 Discussion

Results reported within this work show notable differences between free-Aβ dimers and those bound to a zinc(II) ion when investigated using aMD. Structures possessing a metal centre were generally found to sample more compact conformations, as displayed within smaller values for R_g. Cluster data showed that the most populated clusters for metal-free peptides generally have helical secondary structure that formed more extended conformations. In contrast, the most populated zinc-bound clusters sampled conformers that were shown to be more disorganised in terms of secondary structure, as well as showing a decrease in solvent accessibility for these dimers. Increased levels of inter-molecular interactions were seen between chains in salt-bridge data for zinc-bound Aβ. In addition, notably shorter contact distances between separate N-termini, compared to the free peptides, were observed due to the cross-linkage of the metal centre appearing to bridge these individual Aβ monomers.

Zinc-binding within the N-terminus of Aβ shows clear evidence of being responsible for differences in dynamics across the whole amino acid sequence. RMSF statistics show that peptides containing a zinc-centre had lower average mobility compared to dimers absent of a metal. The greatest decrease in RMSF was observed directly within the metal-binding region of the N-terminus, but effects were seen remote to the coordination site. Salt-bridge incidence plots showed that the presence of a metal centre affecteds the types of contacts forming and alters the levels at which interactions were observed when compared to free peptides. Examples include notable decreases in salt-bridges formed between A-Glu11Arg5 and corresponding residues in chain B, which were directly involved in metal-coordination. Greater levels of intermolecular salt-bridge formation were reported within zinc-Aβ trajectories, further exhibiting the effect of the cross-linkage between monomers caused by the metal ion present.

Most notably, the comparison of secondary structure between dimers shows the the presence of a metal-centre affects conformations sampled over the course of the trajectories. Levels of helical and strand-like secondary structure were higher overall in free dimers compared to those bound to zinc. The N-terminus appeared to undergo the greatest change in secondary structure, but levels of structures characterised as coil, turn, and bend were adopted at a higher incidence over the whole amino acid sequence in the presence of

a central zinc ion. The highest populated clusters give further representation of these differences in secondary structure. As discussed previously, the free dimers show some propensity to form helices across the length of each monomer, whilst this helical character is present within fewer residues. This can be attributed to the zinc ion being coordinated within N-terminus residues, and thus affecting the ability to form these helices observed in free Aβ.

Prior computational and experimental studies of both monomeric and dimeric Aβ, discussed previously, suggests coordination of a metal ion leads to enhanced propensity for aggregation. Literature states that the presence of a metal centre (including zinc) can cause misfolding of proteins into these neurotoxic oligomeric structures. Bridging of Aβ monomers by binding to a metal ion has been shown to increase stability of the system and thus disrupt the homeostatic clearance of these peptides. The results presented within this work show good agreement with these studies, mainly the effects on secondary structure reported in comparison to free Aβ. The decrease in helical secondary structure and the increased proximity of monomeric units suggest a potential pathway to misfolding, as well as an inability to clear Aβ due to the bridging effects seen within our results.

5.5 Conclusion

Use of aMD to simulate Aβ dimers was successful in allowing comparisons to be drawn between zinc-bound peptides and free dimers. RMSD and RMSF values showed sampling was sufficient in exploring conformational space. R_g, cluster data and SASA values show that zinc dimers adopt more compact structures. The presence of the metal centre affects the dynamics of the peptide remote to the coordination site, as well as within the N-terminus, within the secondary structure plots and salt-bridge contact maps.

The zinc ion appears to affect secondary structure and interactions between monomers. The higher frequency of helices observed within the free dimers is not seen within zinc Aβ structures, due to the ridigity seen within the N-terminus as a result of coordination to the central zinc ion. This also leads to monomers within the zinc dimer appearing to be closer in proximity and forming a greater level of intermolecular interactions opposed to the free peptides. Results presented here show the importance of metal binding in aggregation and

formation of structures, such as dimers, that may potentially be important in neurotoxic alloforms of Aβ.

5.6 Further Work

Due to time constraints towards the end of this research program, there are further studies that could be conducted into the effects of metal binding on these Aβ dimers. Similar to previous studies conducted within the research group, a comparative analysis of binding modes and different transition metals such as copper or iron could be useful in seeing whether these changes in metal binding reflects results seen within the monomeric form of this peptide. Previously, dynamics were reported on the truncated N-termini of Aβ containing mutations in the amino acid sequence; these same mutants (or others beyond the first 16 amino acids) could be applied here and simulated to study effects of these variants. As with most MD studies, longer simulation times or more parallel trajectories could be generated to confirm more confidently that sufficient sampling has taken place. It would also be relevant to investigate any potential effect of explicit solvent on the Aβ dimers investigated, particularly at the interface between chains as well as other enhanced sampling methods to further understand the key interactions occurring between monomer chains such as salt-bridges and hydrogen bonds, which can then give a better insight to aggregation of dimeric-Aβ.

5.7 References

1 G. G. Glenner and C. W. Wong, *Biochem Bioph Res Co*, 1984, **120**, 885–890.

2 C. L. Masters, G. Simms, N. A. Weinman, G. Multhaupt, B. L. McDonald and K. Beyreuthert, *P Natl Acad Sci USA*, 1985, **82**, 4245-4249

3 B. Alies, H. Eury, C. Bijani, L. Rechignat, P. Faller and C. Hureau, *Inorg Chem*, 2011, **50**, 11192–11201.

4 S. S. Yoon and S. A. Jo, *Biomol Ther*, 2012, **20**, 245–255.

5 W. L. Klein, *Neurochem Int*, 2002, **41**, 345–352.

6 D. J. Selkoe, *Behav Brain Res*, 2008, **192**, 106–113.

7 W. I. Rosenblum, *Neurobiol Aging*, 2014, **35**, 969–974.

8 P. A. Adlard and A. I. Bush, *J Alzheimer's Dis*, 2018, **62**, 1369–1379.

9 P. Wang and Z. Y. Wang, *Ageing Res Rev*, 2017, **35**, 265–290.

10 D. S. Kalinowski, C. Stefani, S. Toyokuni, T. Ganz, G. J. Anderson, N. v. Subramaniam, D. Trinder, J. K. Olynyk, A. Chua, P. J. Jansson, S. Sahni, D. J. R. Lane, A. M. Merlot, Z. Kovacevic, M. L. H. Huang, C. S. Lee and D. R. Richardson, *Biochim Biophys Acta*, 2016, **1863**, 727–748.

11 C. Cheignon, M. Jones, E. Atrián-Blasco, I. Kieffer, P. Faller, F. Collin and C. Hureau, *Chem Sci*, 2017, **8**, 5107–5118.

12 M. A. Lovell, J. D. Robertson, W. J. Teesdale, J. L. Campbell and W. R. Markesbery, *J Neurol Sci*, 1998, **158**, 47–52.

13 S. Wärmländer, A. Tiiman, A. Abelein, J. Luo, J. Jarvet, K. L. Söderberg, J. Danielsson and A. Gräslund, *ChemBioChem*, 2013, **14**, 1692–1704.

14 P. O. Tsvetkov, I. A. Popov, E. N. Nikolaev, A. I. Archakov, A. A. Makarov and S. A. Kozin, *ChemBioChem*, 2008, **9**, 1564–1567.

15 S. Côté, R. Laghaei, P. Derreumaux and N. Mousseau, *J Phys Chem B*, 2012, **116**, 4043–4055.

16 P. D. Q. Huy, Q. Van Vuong, G. La Penna, P. Faller and M. S. Li, *ACS Chem Neurosci*, 2016, **7**, 1348–1363.

17 X. Zhu, R. P. Bora, A. Barman, R. Singh and R. Prabhakar, *J Phys Chem B*, 2012, **116**, 4405–4416.

18 P. R. L. Markwick and J. A. McCammon, *Phys Chem Chem Phys*, 2011, **13**, 20053–20065.

19 B. Sharma and K. A. Dill., *Commun Biol*, 2021, **4**, 1-11

20 H. Kozlowski, M. Luczkowski, M. Remelli and D. Valensin, *Coordin Chem Rev*, 2012, **256**, 2129–2141.

21 J. Nasica-Labouze, P. H. Nguyen, F. Sterpone, O. Berthoumieu, N. V. Buchete, S. Coté, A. De Simone, A. J. Doig, P. Faller, A. Garcia, A. Laio, M. S. Li, S. Melchionna, N. Mousseau, Y. Mu, A. Paravastu, S. Pasquali, D. J. Rosenman, B. Strodel, B. Tarus, J. H. Viles, T. Zhang, C. Wang and P. Derreumaux, *Chem Rev*, 2015, **115**, 3518–3563.

22 S. A. Kozin, Y. v Mezentsev, A. A. Kulikova, A. v Golovin, A. S. Ivanov, P. O. Tsvetkov and A. A. Makarov, *Roy Soy Ch*, 2011, **7**, 1053–1055.

23 *Molecular Operating Environment (MOE),* 2013, Chemical Computing Group ULC, 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7

24 L. Martínez, R. Andrade, E. G. Birgin and J. M. Martínez, *J Comput Chem*, 2009, **30**, 2157-2164.

25 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J Am Chem Soc*, 1995, **117**, 5179–5197.

26 R. J. Deeth, N. Fey and B. Williams-Hubbard, *J Comput Chem*, 2005, **26**, 123–130.

27 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J Chem Theory Comput*, 2015, **11**, 3696–3713.

28 Zhang, W, Hou, T, Schafmeister, C, Ross, W. S, Case, D. A., 2010, LEaP and gleap: AmberTools Users' Manual Version 1.4

29 P. Li and K. M. Merz, *J Chem Inf Model*, 2016, **56**, 599–604.

30 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R.

Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, 2016, Gaussian, Inc., Wallingford CT

31 J. M. Seminario, *Int J Quantum Chem*, 1996, **60**, 1271–1277.

32 C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *J Phys Chem*, 1996, **97**, 10269– 10280.

33 P. Cieplak, W. D. Cornell, C. Bayly and P. A. Kollman, *J Comput Chem*, 1995, **16**, 1357–1377.

34 D.A. Case, R.M. Betz, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao and P.A. Kollman, 2016, AMBER 2016, University of California, San Francisco.

35 A. Okur, L. Wickstrom, M. Layten, R. Geney, K. Song, V. Hornak and C. Simmerling, *J Chem Theory Comput*, 2006, **2**, 420–433.

36 R. Geney, M. Layten, R. Gomperts, V. Hornak and C. Simmerling, *J Chem Theory Comput*, 2006, **2**, 115–127.

37 N. Al-Shammari, L. Savva, O. Kennedy-Britten and J. A. Platts, *Comput Biol Chem*, 2021, **93**, 1476–9271.

38 J. A. Izaguirre, D. P. Catarello, J. M. Wozniak and R. D. Skeel, *J Chem Phys*, 2001, **114**, 2090–2098.

39 J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J Comput Phys*, 1977, **23**, 327–341.

40 B. Zhao, M. A. Cohen Stuart and C. K. Hall, *PLoS Comput Biol*, 2017, **13**, 1-22

41 D. R. Roe and T. E. Cheatham. (2013) PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data, *J. Chem. Theory Comput.*, 9 (7), 3084–3095.

42 W. Humphrey, A. Dalke and K. Schuten, *J Mol Graphics*, 1996, **14**, 33–38.

43 M. Ester, H. P. Kriegel, J. Sander and X. Xu, *Proc 2nd Int Conf Knowl Discov Data Min*, 1996, **7**, 226–231.

6 Conclusions and Future Work

The aim of the work presented here was to model and investigate structures and dynamics of Aβ alloforms bound to various transition metals in order to deduce potential inferences on aggregational properties linked with the onset of Alzheimer's disease, as well as effects of metal coordination. Aβ peptides have been studied previously but have provided novel challenges when trying to investigate them due to their flexible nature. Computational methods have been applied extensively to free Aβ but less has been published on Aβ metalloproteins. Further to the primary objective, the aim here was to highlight the validity of certain computational methods in the study of these structures. Both conventional and accelerated MD were selected to simulate the trajectory of motion in space of these transient peptides over time due to their capability of modelling structures of this size and nature.

In Chapter 3, seven copper-bound genetic variants of truncated N-terminus Aβ structures were first simulated using conventional MD to compare the effects of amino acid substitution and copper-binding to non-mutated, wild-type Aβ. In addition, this study was able to confirm whether MD was a suitable method for further studies on larger Aβ models. All structures analysed showed variance in size and levels of flexibility. Cluster data and R_g showed notable differences in conformations sampled between mutant structures and wild-type Aβ, highlighting the effects of these alterations in the amino acid sequence within the N-terminus. Salt-bridge incidence data and Ramachandran plots showed interesting differences in structure between experiments, further highlighting the effect of these slight modifications in residues. Comparison of amino acid alterations at similar positions in the amino acid chain, such as A2T/A2V and D7H/D7N, showed that both the site of mutation and the type of substitution impacted the structure and dynamics of the peptide as a whole, not just the amino acid being altered. Secondary structure data supported evidence seen in literature of previous studies. The protective qualities of the A2T mutation were supported by the overall decrease in β-strand secondary structure, whereas the heightened levels of this type of secondary structure were observed in the A2V and E11K variants, complementing evidence of their increased neurotoxic properties and increased propensity of aggregation. Results obtained showed clear evidence that MD is a suitable method for

use on larger Aβ structures as the data collected for these truncated peptides can be used as a good model for full-length Aβ.

As mentioned previously, conventional MD generated good results for the N-terminus structures but in order to successfully replicate this success on the more flexible full-length Aβ, accelerated MD was selected. Chapter 4 reported results on aMD studies of varying lengths of Aβ bound to a selection of transition metals as well as free Aβ. Enhanced mobility was observed in the free peptides, as expected, shown by RMSF values being highest in the metal-free simulations but especially residues that were coordinated to a metal centre in other trajectories. Thus, this highlighted the anchorage effect these metal ions have on the structure as a whole despite metal binding occurring within the N-terminal region only. This was also displayed by levels of helical and strand-like secondary structure decreasing within the N-terminus of any peptides containing a metal-centre. Despite metal ions possessing different binding modes from one another, similarities in motions and dynamics were observed between Zn(II) and Fe(II)-bound peptides. Furthermore, secondary structure and free energy data showed further comparisons between these metalloproteins, whereas comparable data for R_g and salt-bridge profiles was observed between free Aβ and species containing Cu(II). The decision to use aMD for these simulations was successful in enhancing the exploration of conformational space compared to structures of a similar size previously simulated using conventional MD.

Finally, Chapter 5 describes aMD computational studies comparing full-length Aβ dimers in the presence of a Zn(II) bridging ion against free Aβ dimeric structures. This study presented novel challenges as for one of the free-Aβ structures, the two monomeric units moved away from one another to distances they could no longer be classed as dimeric. This issue was not observed in the zinc-bound simulations nor the other free dimer. This could potentially be attributed to the boost potential from aMD causing a bias towards structures where the monomers were no long in close contact with one another. To counteract this, of the 600 ns aMD data generated for this specific structure in run 2, 380 ns of usable data of the peptides still in dimeric form were taken for further analysis in conjunction with the full 600 ns of run 1 for the free peptide. When analysed against results for the other free peptide, where the full 1200 ns of data were successfully generated from 2 runs, these simulations displayed comparable values, showing little impact on results. Both RMSF and RMSD values suggested

sufficient sampling had occurred, meaning these results were still valid for further study. Higher levels of helical secondary structures were found in free dimers, due to the absence of a metal centre limiting mobility within the N-terminus. Zinc-bound dimers were shown to adopt more compact structures due to the bridging metal ion holding the monomers in close proximity to one another, displayed by clustering analysis, surface-area solvent accessibility and R_g values. As a result of this, Zn(II) dimers showed higher incidence of intermolecular interactions between monomers compared to free Aβ, which in turn formed higher levels of intramolecular interactions, evidenced by salt-bridge data. This supports theories that the presence of zinc and Aβ in its dimeric form can lead to further aggregation into potentially neurotoxic forms.

Having witnessed the capability of methods used within this research, there are multiple directions these studies could further explore the nature of metal ion interactions with Aβ in future. Work presented here only looked at N-terminus mutations of Aβ whereas there are many more towards the C-terminus, which could show differing effects on structures of Aβ. Knowing it is possible for modelling of both these mutants and full length Aβ structures, it is definitely feasible that these variants could be explored in further detail. In addition, studies reported here only looked at copper binding to generic variants of Aβ, but other metals such as zinc and iron could yield different results in conjunction with these alterations in the amino acid sequence. Further to this, there are a number of other metals shown to bind to Aβ that could be studied, for example aluminium, manganese, and platinum where there is limited computational investigations into complexes containing these metal centres. These ions could also be applied in the further studies of the dimeric species of Aβ. Different binding modes of metals examined within this body of work could also be investigated to study the effects of coordination to different amino acids. Typically, with most computational studies, longer simulation times and even more parallel trajectories could provide greater confirmation that sufficient sampling has been performed. All studies reported here were performed in implicit solvent, use of an explicit solvation model could highlight any effects of solvent upon Aβ to give even better replication of results observed *in vivo*, especially for dimeric species where any solvent effects (if any) on interfacing atoms between monomers can be observed. These results have shown good inferences can be made upon the aggregational properties of Aβ dependent on the form and presence of any

metal ions, but more studying of these could further highlight the impact these peptides have on aggregation and provide greater insight into their role in the onset of Alzheimer's disease.

A Appendix

A1 Appendices for Chapter 4

A1.1 Hydrogen bond output for incidence > 5% for Aβ16-Free aMD trajectory data

#Acceptor	DonorH	Donor	Frames	Fraction	Average Distance (Å)	Average Angle (°)
GLU_3@OE1	GLU_3@H	GLU_3@N	13287	0.2215	2.8268	151.1149
GLU_3@OE2	GLU_3@H	GLU_3@N	12995	0.2166	2.8293	151.08
GLU_11@OE2	HID_13@HD1	HID_13@ND1	12597	0.2099	2.8143	158.6036
GLU_11@OE1	HID_13@HD1	HID_13@ND1	11973	0.1996	2.8167	158.4056
ASP_7@OD2	ARG_5@HE	ARG_5@NE	9269	0.1545	2.8322	157.8629
ASP_7@OD1	ARG_5@HH22	ARG_5@NH2	8836	0.1473	2.8249	156.6234
ASP_7@OD2	ARG_5@HH22	ARG_5@NH2	8353	0.1392	2.8255	155.6447
ASP_7@OD1	ARG_5@HE	ARG_5@NE	7173	0.1196	2.8365	157.7028
GLU_11@OE2	HID_13@H	HID_13@N	6619	0.1103	2.8735	159.4293
GLU_11@OE1	ARG_5@HH11	ARG_5@NH1	5942	0.099	2.8153	155.8511
GLU_11@OE1	HID_13@H	HID_13@N	5817	0.0969	2.8772	158.9741
GLU_11@OE1	ARG_5@HH21	ARG_5@NH2	5499	0.0916	2.8157	155.0461
GLU_11@OE2	ARG_5@HH22	ARG_5@NH2	5488	0.0915	2.8031	154.8816
GLU_11@OE1	ARG_5@HH22	ARG_5@NH2	5302	0.0884	2.8014	155.0233
GLU_11@OE2	ARG_5@HH11	ARG_5@NH1	5128	0.0855	2.8172	155.9327
VAL_12@O	GLN_15@H	GLN_15@N	5114	0.0852	2.8988	153.4124
GLU_11@OE2	ARG_5@HH21	ARG_5@NH2	4991	0.0832	2.8135	155.4628
VAL_12@O	LYS_16@H	LYS_16@N	4327	0.0721	2.8934	160.7332
ASP_7@OD1	HID_13@HD1	HID_13@ND1	4314	0.0719	2.8249	158.145
GLU_3@OE1	ASP_1@H2	ASP_1@N	4299	0.0717	2.8121	154.5664
HID_13@O	NME_17@H	NME_17@N	4241	0.0707	2.8813	153.7349
GLU_3@OE1	ASP_1@H1	ASP_1@N	4223	0.0704	2.8112	154.9202
PHE_4@O	SER_8@H	SER_8@N	4211	0.0702	2.8875	157.7614
GLU_3@OE2	ASP_1@H2	ASP_1@N	4190	0.0698	2.8125	154.8478
GLU_3@OE1	ASP_1@H3	ASP_1@N	4176	0.0696	2.8131	154.8955
GLU_3@OE2	ASP_1@H1	ASP_1@N	4141	0.069	2.8089	154.6743
GLU_11@OE2	ARG_5@HE	ARG_5@NE	4117	0.0686	2.8398	153.4267
GLU_3@OE2	ASP_1@H3	ASP_1@N	4018	0.067	2.8122	154.7632
GLU_11@OE1	ARG_5@HE	ARG_5@NE	3924	0.0654	2.8391	153.362
GLU_3@O	HID_6@HD1	HID_6@ND1	3839	0.064	2.8473	157.3098
HID_13@O	LYS_16@H	LYS_16@N	3519	0.0587	2.8967	153.9527
ASP_7@OD2	HID_13@HD1	HID_13@ND1	3134	0.0522	2.835	157.9279

A1.2 Hydrogen bond output for incidence > 5% for Aβ16-Cu aMD trajectory data

#Acceptor	DonorH	Donor	Frames	Fraction	Avgerage Distance (Å)	Avgerage Angles (º)
GLU_3@O	HD1_6@H	HD1_6@N	37995	0.6332	2.8257	157.9848
GLU_11@OE1	HD2_13@HD1	HD2_13@ND1	11743	0.1957	2.7941	156.1114
GLU_11@OE2	HD2_13@HD1	HD2_13@ND1	11058	0.1843	2.7958	156.1639
GLU_3@OE2	HD1_6@HD1	HD1_6@ND1	9159	0.1527	2.8204	153.4395
GLU_3@OE1	HD1_6@HD1	HD1_6@ND1	9001	0.15	2.8204	153.5
ALA_2@O	ARG_5@H	ARG_5@N	8524	0.1421	2.9041	152.8548
GLU_11@OE2	ARG_5@HH22	ARG_5@NH2	7679	0.128	2.7998	156.5467
GLU_11@OE1	ARG_5@HH22	ARG_5@NH2	7538	0.1256	2.8019	156.613
GLU_11@OE1	HD2_13@H	HD2_13@N	7071	0.1178	2.8371	160.8988
GLU_11@OE2	HD2_13@H	HD2_13@N	6995	0.1166	2.8367	160.3983
GLU_11@O	HID_14@HD1	HID_14@ND1	6103	0.1017	2.8252	156.2879
GLU_11@O	HID_14@H	HID_14@N	5908	0.0985	2.8876	155.2486
GLU_11@OE2	ARG_5@HE	ARG_5@NE	4946	0.0824	2.8467	155.3047
GLU_11@OE1	ARG_5@HE	ARG_5@NE	4895	0.0816	2.8464	155.3442
GLN_15@OE1	GLN_15@H	GLN_15@N	3406	0.0568	2.8444	149.2834
GLU_3@OE2	GLU_3@H	GLU_3@N	3054	0.0509	2.8298	151.8926
GLU_3@OE1	GLU_3@H	GLU_3@N	3019	0.0503	2.8292	152.0149

A1.3 Hydrogen bond output for incidence > 5% for Aβ16-Zn aMD trajectory data

#Acceptor	DonorH	Donor	Frames	Fraction	Average Distance (Å)	Average Angle (º)
GU1_11@O	HIE_13@H	HIE_13@N	27140	0.4523	2.7997	146.6833
ASP_7@OD2	ARG_5@HH11	ARG_5@NH1	20678	0.3446	2.8252	157.114
ASP_7@OD1	ARG_5@HH11	ARG_5@NH1	20213	0.3369	2.8244	156.9227
ASP_7@OD1	ARG_5@HE	ARG_5@NE	17658	0.2943	2.8391	158.2646
ASP_7@OD2	ARG_5@HE	ARG_5@NE	17135	0.2856	2.8417	157.9545
GLY_9@O	VAL_12@H	VAL_12@N	6898	0.115	2.9015	154.9513
HIE_13@O	LYS_16@H	LYS_16@N	6310	0.1052	2.8969	156.7712
HIE_13@ND1	GLN_15@H	GLN_15@N	5679	0.0946	2.9288	155.9424
ARG_5@O	ASP_7@H	ASP_7@N	5651	0.0942	2.8498	148.5448
GLU_3@OE2	GLU_3@H	GLU_3@N	5196	0.0866	2.815	150.3758
GLU_3@OE1	GLU_3@H	GLU_3@N	5095	0.0849	2.8146	150.4519
ARG_5@O	ALA_2@H	ALA_2@N	3943	0.0657	2.8885	157.6516
ALA_2@O	ARG_5@H	ARG_5@N	3910	0.0652	2.8954	156.226
GU1_11@O	HE2_14@H	HE2_14@N	3487	0.0581	2.8939	161.0827
HE2_14@O	LYS_16@H	LYS_16@N	3153	0.0525	2.839	147.4044

A1.4 Hydrogen bond output for incidence > 5% for Aβ16-Fe aMD trajectory data

#Acceptor	DonorH	Donor	Frames	Fraction	Average Distance (Å)	Average Angle (°)
TYR_10@O	HE2_14@HE2	HE2_14@NE2	29797	0.4966	2.8386	158.682
ASP_7@OD1	ARG_5@HH11	ARG_5@NH1	20825	0.3471	2.8176	154.0815
ASP_7@OD2	ARG_5@HH11	ARG_5@NH1	20270	0.3378	2.8182	153.6696
ASP_7@OD2	ARG_5@HE	ARG_5@NE	15630	0.2605	2.8413	153.6114
ASP_7@OD1	ARG_5@HE	ARG_5@NE	14977	0.2496	2.8417	153.4274
PHE_4@O	HE1_6@H	HE1_6@N	9917	0.1653	2.8172	146.0115
HIE_13@O	AP1_1@H3	AP1_1@N	5620	0.0937	2.8694	152.6916
ASP_7@OD2	SER_8@H	SER_8@N	4924	0.0821	2.8331	143.9274
GLU_3@O	ARG_5@H	ARG_5@N	4741	0.079	2.8203	149.9299
ASP_7@OD1	SER_8@H	SER_8@N	4670	0.0778	2.8333	144.0791
GLU_3@OE2	AP1_1@H2	AP1_1@N	3820	0.0637	2.8661	153.6697
AP1_1@OD2	HE2_14@H	HE2_14@N	3814	0.0636	2.8495	157.3727
AP1_1@OD2	AP1_1@H3	AP1_1@N	3783	0.063	2.8145	142.1817
GLU_3@OE1	AP1_1@H2	AP1_1@N	3718	0.062	2.8682	153.8605
AP1_1@OD2	GLN_15@H	GLN_15@N	3595	0.0599	2.8759	158.4542
AP1_1@OD1	AP1_1@H3	AP1_1@N	3582	0.0597	2.8126	142.1053
AP1_1@OD1	HE2_14@H	HE2_14@N	3547	0.0591	2.8521	157.2973
AP1_1@OD1	GLN_15@H	GLN_15@N	3274	0.0546	2.8772	158.2217
GLU_3@OE2	GLU_3@H	GLU_3@N	3158	0.0526	2.81	150.1407
GLU_3@OE2	ASP_7@H	ASP_7@N	3118	0.052	2.8553	151.5785
GLU_3@OE1	ASP_7@H	ASP_7@N	3080	0.0513	2.859	151.6486
GLU_3@OE1	GLU_3@H	GLU_3@N	3024	0.0504	2.8101	150.1004

A1.5 Hydrogen bond output for incidence > 5% for Aβ28-Free aMD trajectory data

#Acceptor	DonorH	Donor	Frames	Fraction	Average Distance (Å)	Average Angle (°)
GLU_3@OE1	ARG_5@HH11	ARG_5@NH1	14186	0.2364	2.8049	158.0239
GLU_3@OE2	ARG_5@HH11	ARG_5@NH1	12174	0.2029	2.8079	157.761
GLU_3@OE2	ARG_5@HE	ARG_5@NE	11131	0.1855	2.8465	159.2539
LYS_16@O	PHE_20@H	PHE_20@N	10239	0.1706	2.8808	157.7429
GLU_3@OE1	ARG_5@HE	ARG_5@NE	9756	0.1626	2.844	158.6062
GLU_11@O	HID_14@HD1	HID_14@ND1	9579	0.1596	2.8328	157.5213
TYR_10@O	HID_13@HD1	HID_13@ND1	8203	0.1367	2.8374	157.1611
ASP_23@OD1	ARG_5@HH12	ARG_5@NH1	8108	0.1351	2.8121	157.9141
GLU_11@OE1	ARG_5@HH11	ARG_5@NH1	7987	0.1331	2.8043	154.5464
GLU_11@OE2	ARG_5@HH11	ARG_5@NH1	7933	0.1322	2.8032	154.4245
GLU_3@O	HID_6@HD1	HID_6@ND1	7872	0.1312	2.8412	156.4322
ASP_23@OD2	ARG_5@HH22	ARG_5@NH2	7755	0.1293	2.8126	158.4201
VAL_12@O	LYS_16@H	LYS_16@N	7678	0.128	2.8863	159.3625

TYR_10@O	HID_14@HD1	HID_14@ND1	7325	0.1221	2.8445	158.1701
ASP_23@OD2	ARG_5@HH12	ARG_5@NH1	7185	0.1197	2.816	157.0098
PHE_4@O	SER_8@H	SER_8@N	6612	0.1102	2.8885	158.8973
GLU_11@OE2	ARG_5@HE	ARG_5@NE	6576	0.1096	2.8284	154.7978
ASP_23@OD1	ARG_5@HH22	ARG_5@NH2	6532	0.1089	2.8133	158.3164
GLU_11@OE1	ARG_5@HE	ARG_5@NE	6489	0.1081	2.8309	154.7658
GLU_11@O	GLN_15@H	GLN_15@N	6407	0.1068	2.8865	158.5989
LEU_17@O	ALA_21@H	ALA_21@N	6175	0.1029	2.8865	157.7474
GLN_15@O	PHE_19@H	PHE_19@N	5994	0.0999	2.8899	159.6194
HID_13@O	LEU_17@H	LEU_17@N	5745	0.0958	2.8878	157.6588
GLU_3@OE2	PHE_4@H	PHE_4@N	5665	0.0944	2.8422	151.9244
VAL_12@O	GLN_15@H	GLN_15@N	5655	0.0943	2.8932	153.5835
GLU_3@OE2	ARG_5@H	ARG_5@N	5638	0.094	2.8786	161.8173
TYR_10@O	HID_14@H	HID_14@N	4985	0.0831	2.8837	158.1189
GLY_9@O	HID_13@HD1	HID_13@ND1	4776	0.0796	2.8506	157.6548
GLU_3@OE1	ARG_5@H	ARG_5@N	4593	0.0766	2.8784	161.5506
GLU_3@OE1	PHE_4@H	PHE_4@N	4232	0.0705	2.8402	150.8866
PHE_4@O	ASP_7@H	ASP_7@N	4062	0.0677	2.8941	151.8092
GLU_3@O	HID_6@H	HID_6@N	3990	0.0665	2.904	157.7685
GLU_22@O	GLY_25@H	GLY_25@N	3883	0.0647	2.8901	152.1771
LYS_16@O	PHE_19@H	PHE_19@N	3738	0.0623	2.8974	152.2466
GLN_15@O	VAL_18@H	VAL_18@N	3477	0.058	2.9017	151.6061
LEU_17@O	PHE_20@H	PHE_20@N	3412	0.0569	2.8954	152.194
PHE_19@O	ASP_23@H	ASP_23@N	3398	0.0566	2.8909	158.1351
ALA_21@O	GLY_25@H	GLY_25@N	3332	0.0555	2.8725	152.7403
HID_6@O	GLY_9@H	GLY_9@N	3305	0.0551	2.8948	152.107
TYR_10@O	HID_13@H	HID_13@N	3267	0.0544	2.901	153.8412
HID_13@O	LYS_16@H	LYS_16@N	3155	0.0526	2.8991	152.9195
HID_14@O	LEU_17@H	LEU_17@N	3135	0.0522	2.9005	152.2338
PHE_4@O	SER_8@HG	SER_8@OG	3081	0.0513	2.8305	158.1345
VAL_18@O	GLU_22@H	GLU_22@N	3075	0.0512	2.8904	157.051
PHE_20@O	VAL_24@H	VAL_24@N	3027	0.0505	2.8942	159.8675

A1.6 Hydrogen bond output for incidence > 5% for Aβ28-Cu aMD trajectory data

#Acceptor	DonorH	Donor	Frames	Fraction	Average Distance (Å)	Average Angle (°)
PHE_4@O	HD1_6@H	HD1_6@N	23062	0.3844	2.7316	151.0203
TYR_10@O	HD2_13@H	HD2_13@N	11341	0.189	2.8738	153.9272
GLY_9@O	HD1_6@HD1	HD1_6@ND1	8559	0.1426	2.8266	151.0323
GLU_11@OE2	HD1_6@H	HD1_6@N	8457	0.1409	2.6875	160.2849
GLU_11@O	HID_14@H	HID_14@N	8026	0.1338	2.893	154.0165
GLU_11@O	GLN_15@H	GLN_15@N	7946	0.1324	2.8805	160.3515
ASP_7@OD1	HD1_6@H	HD1_6@N	7895	0.1316	2.6863	158.334

ASP_7@O	HD1_6@HD1	HD1_6@ND1	7730	0.1288	2.7702	163.7451
GLU_11@OE1	HD1_6@HD1	HD1_6@ND1	7644	0.1274	2.6749	159.6137
GLU_11@O	HID_14@HD1	HID_14@ND1	7572	0.1262	2.8286	157.0959
ASP_7@OD2	HD1_6@H	HD1_6@N	7307	0.1218	2.6862	158.52
GLU_11@OE1	GLU_11@H	GLU_11@N	7111	0.1185	2.819	152.5156
GLU_11@OE2	GLU_11@H	GLU_11@N	6898	0.115	2.8204	152.7618
GLU_11@OE1	HD1_6@H	HD1_6@N	6343	0.1057	2.7022	159.3489
ASP_7@OD2	ARG_5@HH11	ARG_5@NH1	5909	0.0985	2.81	155.2511
ASP_7@OD1	ARG_5@HH11	ARG_5@NH1	5603	0.0934	2.8093	155.4087
GLU_11@OE2	HD1_6@HD1	HD1_6@ND1	5057	0.0843	2.6791	158.6712
LEU_17@O	ALA_21@H	ALA_21@N	4809	0.0801	2.8852	158.0212
GLU_22@OE2	HD1_6@HD1	HD1_6@ND1	4561	0.076	2.7033	163.1152
GLU_22@OE1	HD1_6@HD1	HD1_6@ND1	4410	0.0735	2.7041	162.7587
ASP_7@OD2	ARG_5@HE	ARG_5@NE	4304	0.0717	2.8557	153.1075
LEU_17@O	PHE_20@H	PHE_20@N	4163	0.0694	2.8969	153.698
GLU_22@OE1	ARG_5@HH11	ARG_5@NH1	4155	0.0693	2.7962	157.8516
GLU_22@O	GLY_25@H	GLY_25@N	4042	0.0674	2.8886	152.3804
ASP_7@OD1	ARG_5@HE	ARG_5@NE	4013	0.0669	2.8545	152.965
ALA_21@O	GLY_25@H	GLY_25@N	3936	0.0656	2.8716	153.8655
TYR_10@O	HD2_13@HD1	HD2_13@ND1	3894	0.0649	2.8366	149.8867
GLU_22@OE2	ARG_5@HH11	ARG_5@NH1	3770	0.0628	2.793	156.5927
PHE_19@O	GLU_22@H	GLU_22@N	3763	0.0627	2.8964	153.3656
TYR_10@O	HID_14@HD1	HID_14@ND1	3725	0.0621	2.854	157.4845
SER_8@O	HD1_6@HD1	HD1_6@ND1	3581	0.0597	2.7795	158.8429
PHE_19@O	ASP_23@H	ASP_23@N	3458	0.0576	2.8878	157.6351
HD1_6@O	GLY_9@H	GLY_9@N	3355	0.0559	2.9021	156.3918
LYS_16@O	PHE_20@H	PHE_20@N	3188	0.0531	2.8854	158.1815

A1.7 Hydrogen bond output for incidence > 5% for Aβ28-Zn aMD trajectory data

#Acceptor	DonorH	Donor	Frames	Fraction	Average Distance (Å)	Average Angle (°)
ASP_7@OD1	ARG_5@HH11	ARG_5@NH1	19326	0.3221	2.8202	156.0837
ASP_7@OD2	ARG_5@HH11	ARG_5@NH1	19080	0.318	2.8203	156.1986
ASP_7@OD1	ARG_5@HE	ARG_5@NE	15963	0.2661	2.8392	157.2781
ASP_7@OD2	ARG_5@HE	ARG_5@NE	15738	0.2623	2.8412	157.3038
GU1_11@O	HIE_13@H	HIE_13@N	9961	0.166	2.7976	146.2938
LEU_17@O	PHE_20@H	PHE_20@N	9464	0.1577	2.8917	155.0586
HIE_13@O	LYS_16@H	LYS_16@N	8337	0.1389	2.8936	153.2106
GU1_11@O	HE2_14@H	HE2_14@N	8058	0.1343	2.8897	158.2132
LEU_17@O	ALA_21@H	ALA_21@N	6632	0.1105	2.888	158.8102
ASP_23@OD2	ARG_5@HH12	ARG_5@NH1	5132	0.0855	2.8159	157.4304
ASP_23@OD1	ARG_5@HH22	ARG_5@NH2	4934	0.0822	2.8161	158.0523
GLU_3@OE1	GLU_3@H	GLU_3@N	4756	0.0793	2.8189	151.2739

GLU_22@O	GLY_25@H	GLY_25@N	4668	0.0778	2.8905	152.5828
GLU_3@OE2	GLU_3@H	GLU_3@N	4594	0.0766	2.8191	151.3416
HE2_14@O	LEU_17@H	LEU_17@N	4441	0.074	2.9031	156.1212
ASP_23@OD1	ARG_5@HH12	ARG_5@NH1	4438	0.074	2.8187	156.6671
ARG_5@O	ALA_2@H	ALA_2@N	4111	0.0685	2.8875	158.0432
ASP_23@OD2	ARG_5@HH22	ARG_5@NH2	3972	0.0662	2.8212	157.3495
ARG_5@O	ASP_7@H	ASP_7@N	3791	0.0632	2.8541	146.7995
VAL_18@O	GLU_22@H	GLU_22@N	3653	0.0609	2.8926	157.3836
HIE_13@O	LEU_17@H	LEU_17@N	3569	0.0595	2.8907	160.1548
VAL_24@O	ASN_27@HD22	ASN_27@ND2	3347	0.0558	2.8638	159.3264
ALA_21@O	GLY_25@H	GLY_25@N	3218	0.0536	2.8715	154.2491
ALA_2@O	ARG_5@H	ARG_5@N	3131	0.0522	2.8936	156.2466

A1.8 Hydrogen bond output for incidence > 5% for Aβ28-Fe aMD trajectory data

#Acceptor	DonorH	Donor	Frames	Fraction	Average Distance (Å)	Average Angle (°)
ASP_7@OD1	ARG_5@HH11	ARG_5@NH1	19421	0.3237	2.8128	156.0809
ASP_7@OD2	ARG_5@HH11	ARG_5@NH1	19406	0.3234	2.8126	156.4072
TYR_10@O	HE2_14@HE2	HE2_14@NE2	14739	0.2457	2.8361	157.3511
ASP_7@OD1	ARG_5@HE	ARG_5@NE	12706	0.2118	2.8349	155.2346
PHE_4@O	HE1_6@H	HE1_6@N	12340	0.2057	2.8365	145.171
ASP_7@OD2	ARG_5@HE	ARG_5@NE	11532	0.1922	2.8367	155.0701
GLU_3@O	ARG_5@H	ARG_5@N	7062	0.1177	2.8127	150.1011
GLU_3@OE1	AP1_1@H2	AP1_1@N	6927	0.1154	2.8406	157.1462
GLU_3@OE2	AP1_1@H2	AP1_1@N	6440	0.1073	2.8425	157.3155
LEU_17@O	ALA_21@H	ALA_21@N	5324	0.0887	2.8836	157.9266
LEU_17@O	PHE_20@H	PHE_20@N	5104	0.0851	2.894	153.6828
HE2_14@O	AP1_1@H3	AP1_1@N	4651	0.0775	2.885	154.9922
GLN_15@O	PHE_19@H	PHE_19@N	4139	0.069	2.8827	160.0143
HIE_13@O	GLN_15@H	GLN_15@N	4008	0.0668	2.8333	146.5271
LYS_16@O	PHE_20@H	PHE_20@N	3882	0.0647	2.8845	158.2731
GLN_15@O	VAL_18@H	VAL_18@N	3831	0.0639	2.8939	154.1253
ALA_21@O	GLY_25@H	GLY_25@N	3549	0.0592	2.8693	154.5022
VAL_18@O	GLU_22@H	GLU_22@N	3350	0.0558	2.8879	156.4749
ASP_7@OD1	SER_8@H	SER_8@N	3344	0.0557	2.8352	145.2592
GLU_22@O	GLY_25@H	GLY_25@N	3225	0.0537	2.89	152.2151
ASP_7@OD2	SER_8@H	SER_8@N	3116	0.0519	2.8379	145.4375
LEU_17@O	HE1_6@HE2	HE1_6@NE2	3092	0.0515	2.8395	155.3317
VAL_24@O	ASN_27@HD22	ASN_27@ND2	3089	0.0515	2.8614	158.9857

A1.9 Hydrogen bond output for incidence > 5% for Aβ40-Free aMD trajectory data

#Acceptor	DonorH	Donor	Frames	Fraction	Average DIstance (Å)	Average Angle (°)
GLU_3@OE1	ARG_5@HH11	ARG_5@NH1	14031	0.2339	2.803	157.6745
GLU_3@OE2	ARG_5@HH11	ARG_5@NH1	13254	0.2209	2.8047	157.8581
TYR_10@O	HID_14@HD1	HID_14@ND1	11187	0.1865	2.8425	157.719
GLU_3@OE1	ARG_5@HE	ARG_5@NE	10256	0.1709	2.8483	157.5466
GLU_3@OE2	ARG_5@HE	ARG_5@NE	10156	0.1693	2.8489	157.9505
GLU_11@O	HID_14@HD1	HID_14@ND1	8883	0.148	2.8355	157.4289
GLU_3@O	HID_6@HD1	HID_6@ND1	8819	0.147	2.8383	156.344
GLY_9@O	HID_13@HD1	HID_13@ND1	8676	0.1446	2.849	157.3599
TYR_10@O	HID_13@HD1	HID_13@ND1	8111	0.1352	2.8354	156.321
LEU_17@O	ALA_21@H	ALA_21@N	7512	0.1252	2.8841	157.3257
GLU_11@O	GLN_15@H	GLN_15@N	7495	0.1249	2.8858	158.9851
ILE_31@O	MET_35@H	MET_35@N	7239	0.1206	2.8904	157.3501
ILE_32@O	VAL_36@H	VAL_36@N	6338	0.1056	2.8907	160.3285
VAL_12@O	LYS_16@H	LYS_16@N	6296	0.1049	2.8863	157.4471
ALA_21@O	GLY_25@H	GLY_25@N	6193	0.1032	2.8672	153.489
ILE_31@O	LEU_34@H	LEU_34@N	5987	0.0998	2.8953	152.3941
TYR_10@O	HID_14@H	HID_14@N	5629	0.0938	2.8857	157.7153
PHE_4@O	SER_8@H	SER_8@N	5302	0.0884	2.8858	157.3685
GLU_11@OE1	ARG_5@HH11	ARG_5@NH1	5277	0.088	2.8137	154.2936
GLU_11@OE2	ARG_5@HH11	ARG_5@NH1	5187	0.0864	2.8161	154.2587
GLU_3@OE2	ARG_5@H	ARG_5@N	5085	0.0848	2.8761	160.8771
GLU_3@OE1	ARG_5@H	ARG_5@N	5067	0.0844	2.8752	160.693
LEU_17@O	PHE_20@H	PHE_20@N	5053	0.0842	2.8969	153.5772
GLY_33@O	GLY_37@H	GLY_37@N	4994	0.0832	2.8775	153.3837
MET_35@O	GLY_38@H	GLY_38@N	4937	0.0823	2.8941	150.8918
GLU_11@OE1	ARG_5@HE	ARG_5@NE	4803	0.08	2.8193	155.537
VAL_12@O	GLN_15@H	GLN_15@N	4789	0.0798	2.8952	153.8146
GLU_3@O	HID_6@H	HID_6@N	4754	0.0792	2.8987	157.7902
GLU_11@OE2	ARG_5@HE	ARG_5@NE	4588	0.0765	2.8205	155.6472
HID_13@O	LEU_17@H	LEU_17@N	4588	0.0765	2.8896	158.6965
LYS_16@O	PHE_20@H	PHE_20@N	4373	0.0729	2.886	158.2404
GLU_22@OE2	ARG_5@HH22	ARG_5@NH2	4350	0.0725	2.798	158.4445
GLU_3@OE1	PHE_4@H	PHE_4@N	4283	0.0714	2.8434	151.85
GLU_22@O	GLY_25@H	GLY_25@N	4276	0.0713	2.894	152.5293
GLU_3@OE2	PHE_4@H	PHE_4@N	4266	0.0711	2.8393	151.7621
PHE_4@O	ASP_7@H	ASP_7@N	4166	0.0694	2.8945	152.8617
GLU_22@OE1	ARG_5@HH12	ARG_5@NH1	4027	0.0671	2.822	156.568
GLU_22@OE1	ARG_5@HH22	ARG_5@NH2	3993	0.0665	2.8008	158.0192
HID_13@O	LYS_16@H	LYS_16@N	3779	0.063	2.8976	153.604
GLU_11@OE2	GLN_15@HE22	GLN_15@NE2	3743	0.0624	2.8313	162.0288
GLU_22@OE2	ARG_5@HH12	ARG_5@NH1	3714	0.0619	2.8252	156.237

GLY_29@O	GLY_33@H	GLY_33@N	3587	0.0598	2.8813	155.4614
VAL_24@O	ASN_27@HD22	ASN_27@ND2	3510	0.0585	2.8646	159.6977
GLU_11@OE1	GLN_15@HE22	GLN_15@NE2	3457	0.0576	2.8343	162.1802
TYR_10@O	HID_13@H	HID_13@N	3433	0.0572	2.9042	153.5863
GLY_9@O	HID_13@H	HID_13@N	3353	0.0559	2.8921	158.4949
ASP_7@OD1	HID_6@HD1	HID_6@ND1	3271	0.0545	2.8365	160.4153
ILE_32@O	MET_35@H	MET_35@N	3189	0.0532	2.8933	150.5429
ASP_7@OD2	HID_6@HD1	HID_6@ND1	3188	0.0531	2.8368	160.5189
ALA_30@O	LEU_34@H	LEU_34@N	3164	0.0527	2.8886	156.5419
GLU_11@O	HID_14@H	HID_14@N	3120	0.052	2.8971	151.5579
GLN_15@O	PHE_19@H	PHE_19@N	3113	0.0519	2.8875	159.0379
GLU_22@OE1	LYS_28@HZ1	LYS_28@NZ	3060	0.051	2.7984	157.2034
ASP_23@OD1	ARG_5@HH22	ARG_5@NH2	3046	0.0508	2.8031	158.6686
VAL_18@O	ALA_21@H	ALA_21@N	3000	0.05	2.9032	154.7284

A1.10 Hydrogen bond output for incidence > 5% for Aβ40-Cu aMD trajectory data

#Acceptor	DonorH	Donor	Frames	Fraction	Average Distance (Å)	Average Angle (°)
ASP_7@O	HD1_6@HD1	HD1_6@ND1	12424	0.2071	2.7846	162.942
GLU_22@OE2	HD1_6@H	HD1_6@N	12403	0.2067	2.6534	160.4818
GLU_3@OE1	HD1_6@H	HD1_6@N	11787	0.1965	2.6449	155.3642
TYR_10@O	HD2_13@H	HD2_13@N	11409	0.1902	2.8739	153.4343
GLY_37@O	HD1_6@H	HD1_6@N	10954	0.1826	2.7426	158.1561
ASP_7@OD2	ARG_5@HE	ARG_5@NE	10626	0.1771	2.8404	156.4897
ASP_7@OD1	ARG_5@HE	ARG_5@NE	10482	0.1747	2.8387	156.5302
ASP_7@OD1	ARG_5@HH11	ARG_5@NH1	10378	0.173	2.8091	156.6611
ASP_7@OD2	ARG_5@HH11	ARG_5@NH1	10142	0.169	2.8101	156.3654
GLU_22@OE1	HD1_6@HD1	HD1_6@ND1	8841	0.1474	2.6848	159.0727
GLU_22@OE1	HD1_6@H	HD1_6@N	7520	0.1253	2.6538	159.9664
GLU_11@O	HID_14@H	HID_14@N	7333	0.1222	2.8954	153.5316
LEU_17@O	ALA_21@H	ALA_21@N	7260	0.121	2.8826	158.3145
GLU_11@O	HID_14@HD1	HID_14@ND1	7256	0.1209	2.8368	157.1333
GLU_11@O	GLN_15@H	GLN_15@N	7051	0.1175	2.8781	160.6239
GLU_3@OE2	GLY_9@H	GLY_9@N	6999	0.1167	2.8336	161.0511
GLU_3@OE2	HD1_6@H	HD1_6@N	6903	0.115	2.6448	156.214
LEU_17@O	PHE_20@H	PHE_20@N	6845	0.1141	2.8917	153.5744
PHE_4@O	HD1_6@HD1	HD1_6@ND1	6823	0.1137	2.8066	149.5325
GLY_9@O	HD1_6@HD1	HD1_6@ND1	6291	0.1048	2.793	157.663
GLU_11@OE1	GLU_11@H	GLU_11@N	6097	0.1016	2.8146	153.4849
GLU_11@OE2	GLU_11@H	GLU_11@N	6007	0.1001	2.8147	153.4038
GLY_33@O	GLY_37@H	GLY_37@N	5900	0.0983	2.869	153.5895
GLU_3@OE2	SER_8@H	SER_8@N	5891	0.0982	2.8633	161.8408
GLY_29@O	GLY_33@H	GLY_33@N	4613	0.0769	2.8728	154.3935

ILE_31@O	MET_35@H	MET_35@N	4529	0.0755	2.8854	157.3766
ILE_32@O	VAL_36@H	VAL_36@N	4528	0.0755	2.8903	159.7573
GLU_22@OE2	HD1_6@HD1	HD1_6@ND1	4425	0.0737	2.6871	159.4767
HID_14@O	LEU_17@H	LEU_17@N	4355	0.0726	2.8921	153.3399
GLN_15@OE1	GLN_15@H	GLN_15@N	4004	0.0667	2.8445	150.7322
TYR_10@O	HD2_13@HD1	HD2_13@ND1	4003	0.0667	2.8448	150.2474
MET_35@O	HD1_6@H	HD1_6@N	3715	0.0619	2.7723	153.6112
SER_26@O	GLY_29@H	GLY_29@N	3672	0.0612	2.891	152.4074
ASP_23@OD2	ARG_5@HH11	ARG_5@NH1	3500	0.0583	2.8188	156.9493
ASP_23@OD1	ARG_5@HH11	ARG_5@NH1	3448	0.0575	2.8201	156.0294
GLU_3@OE2	ASP_7@H	ASP_7@N	3401	0.0567	2.8362	149.7955
ASP_23@OD1	ARG_5@HH22	ARG_5@NH2	3379	0.0563	2.8106	156.9404
ILE_31@O	LEU_34@H	LEU_34@N	3361	0.056	2.8968	152.2334
ASP_7@OD1	ARG_5@HH12	ARG_5@NH1	3318	0.0553	2.8196	155.8316
ALA_21@O	GLY_25@H	GLY_25@N	3227	0.0538	2.8731	155.6599
ASP_7@OD2	ARG_5@HH12	ARG_5@NH1	3223	0.0537	2.8217	155.8984
VAL_12@O	LYS_16@H	LYS_16@N	3213	0.0536	2.8856	157.0184
TYR_10@O	HID_14@HD1	HID_14@ND1	3135	0.0522	2.8547	157.1583
ASP_23@OD2	ARG_5@HH22	ARG_5@NH2	3105	0.0517	2.8097	157.7124
ALA_30@O	GLY_33@H	GLY_33@N	3069	0.0512	2.8845	150.903
LYS_16@O	PHE_20@H	PHE_20@N	3066	0.0511	2.8835	158.2127
VAL_36@O	ARG_5@H	ARG_5@N	3025	0.0504	2.873	161.0311

A1.11 Hydrogen bond output for incidence > 5% for Aβ40-Zn aMD trajectory data

#Acceptor	DonorH	Donor	Frames	Fraction	Average Distance (Å)	Average Angle (º)
GU1_11@O	HIE_13@H	HIE_13@N	24812	0.4135	2.7956	146.5061
ASP_7@OD2	ARG_5@HH11	ARG_5@NH1	19684	0.3281	2.8206	156.6519
ASP_7@OD1	ARG_5@HH11	ARG_5@NH1	19479	0.3246	2.8208	156.56
ASP_7@OD1	ARG_5@HE	ARG_5@NE	16709	0.2785	2.8368	157.6811
ASP_7@OD2	ARG_5@HE	ARG_5@NE	16392	0.2732	2.8396	157.5511
ILE_32@O	VAL_36@H	VAL_36@N	7866	0.1311	2.8915	160.1114
GLU_3@OE2	GLU_3@H	GLU_3@N	7279	0.1213	2.8141	151.3837
LEU_17@O	PHE_20@H	PHE_20@N	7265	0.1211	2.8973	154.398
ILE_31@O	MET_35@H	MET_35@N	7073	0.1179	2.8856	157.2223
GLU_3@OE1	GLU_3@H	GLU_3@N	6971	0.1162	2.8157	151.3452
GLY_29@O	GLY_33@H	GLY_33@N	6969	0.1162	2.8765	154.9423
GLY_33@O	GLY_37@H	GLY_37@N	6667	0.1111	2.8732	153.4573
HE2_14@O	LYS_16@H	LYS_16@N	6319	0.1053	2.8291	147.9535
HIE_13@O	LYS_16@H	LYS_16@N	6000	0.1	2.8932	155.4518
GLY_9@O	VAL_12@H	VAL_12@N	5861	0.0977	2.9026	154.9332
HIE_13@ND1	GLN_15@H	GLN_15@N	5797	0.0966	2.9248	154.9536
LEU_17@O	ALA_21@H	ALA_21@N	5673	0.0945	2.8832	157.6987

ARG_5@O	ASP_7@H	ASP_7@N	5598	0.0933	2.8426	148.0884
ALA_2@O	PHE_4@H	PHE_4@N	5136	0.0856	2.8173	148.2727
SER_26@O	GLY_29@H	GLY_29@N	4518	0.0753	2.895	153.1151
ALA_30@O	LEU_34@H	LEU_34@N	4166	0.0694	2.8888	156.6044
ILE_31@O	LEU_34@H	LEU_34@N	4009	0.0668	2.8974	151.113
HE2_14@O	LEU_17@H	LEU_17@N	3946	0.0658	2.8973	158.1004
LYS_28@O	ILE_32@H	ILE_32@N	3916	0.0653	2.8891	161.9704
GLU_22@O	GLY_25@H	GLY_25@N	3617	0.0603	2.8883	152.641
ARG_5@O	ALA_2@H	ALA_2@N	3554	0.0592	2.8882	157.8265
ILE_32@O	MET_35@H	MET_35@N	3549	0.0592	2.8961	150.8861
ALA_30@O	GLY_33@H	GLY_33@N	3456	0.0576	2.888	150.0124
MET_35@O	GLY_38@H	GLY_38@N	3431	0.0572	2.8953	150.3413
GU1_11@O	HE2_14@H	HE2_14@N	3340	0.0557	2.8947	161.8259
ASP_23@O	SER_26@H	SER_26@N	3014	0.0502	2.9051	154.47

A1.12 Hydrogen bond output for incidence > 5% for Aβ40-Fe aMD trajectory data

#Acceptor	DonorH	Donor	Frames	Fraction	Average Distance (Å)	Average Angle (°)
TYR_10@O	HE2_14@HE2	HE2_14@NE2	27380	0.4563	2.8433	159.1685
ASP_7@OD1	ARG_5@HH11	ARG_5@NH1	17053	0.2842	2.8161	153.8105
ASP_7@OD2	ARG_5@HH11	ARG_5@NH1	15905	0.2651	2.8168	154.1988
HIE_13@O	AP1_1@H3	AP1_1@N	14052	0.2342	2.8649	155.3187
PHE_4@O	HE1_6@H	HE1_6@N	12709	0.2118	2.8277	146.4025
ASP_7@OD1	ARG_5@HE	ARG_5@NE	12008	0.2001	2.842	153.1664
ASP_7@OD2	ARG_5@HE	ARG_5@NE	10719	0.1787	2.8437	153.2024
AP1_1@OD2	GLN_15@H	GLN_15@N	8373	0.1396	2.8754	158.6079
GLY_29@O	GLY_33@H	GLY_33@N	7700	0.1283	2.8722	154.2346
AP1_1@OD1	GLN_15@H	GLN_15@N	6677	0.1113	2.8803	158.0775
LEU_17@O	HE1_6@HE2	HE1_6@NE2	6300	0.105	2.8443	155.7805
GLU_22@O	GLY_25@H	GLY_25@N	6136	0.1023	2.8874	152.3449
ILE_31@O	MET_35@H	MET_35@N	5177	0.0863	2.8865	157.4804
GLU_3@OE2	AP1_1@H2	AP1_1@N	5049	0.0842	2.8656	153.4479
GLU_3@OE1	AP1_1@H2	AP1_1@N	4929	0.0822	2.863	153.6108
ASP_7@OD1	SER_8@H	SER_8@N	4799	0.08	2.8255	144.1628
GLU_3@OE2	ASP_7@H	ASP_7@N	4686	0.0781	2.8571	150.5127
GLU_3@OE1	ASP_7@H	ASP_7@N	4539	0.0756	2.8573	150.3981
VAL_18@O	ALA_21@H	ALA_21@N	4497	0.075	2.8983	154.574
LYS_28@O	ILE_32@H	ILE_32@N	4415	0.0736	2.892	161.7722
LEU_17@O	PHE_20@H	PHE_20@N	4375	0.0729	2.8976	154.9983
ILE_32@O	VAL_36@H	VAL_36@N	4339	0.0723	2.8905	159.6797
GLU_22@O	SER_26@H	SER_26@N	4167	0.0694	2.8773	157.1532
ASP_7@OD2	SER_8@H	SER_8@N	4154	0.0692	2.8277	144.1874
GLY_33@O	GLY_37@H	GLY_37@N	3939	0.0657	2.8699	153.4351

GLY_25@O	GLY_29@H	GLY_29@N	3841	0.064	2.8694	154.1364
ILE_31@O	LEU_34@H	LEU_34@N	3826	0.0638	2.8992	151.8781
GLU_11@OE2	GLU_11@H	GLU_11@N	3751	0.0625	2.8309	151.2284
ASP_7@OD1	ARG_5@H	ARG_5@N	3703	0.0617	2.799	157.6424
ASP_7@OD1	HE1_6@H	HE1_6@N	3659	0.061	2.8214	163.249
GLU_11@OE1	GLU_11@H	GLU_11@N	3578	0.0596	2.8333	151.3189
GLU_3@OE2	HE1_6@H	HE1_6@N	3574	0.0596	2.8158	154.8806
SER_26@O	GLY_29@H	GLY_29@N	3560	0.0593	2.8918	152.5124
GLU_3@OE2	ARG_5@H	ARG_5@N	3557	0.0593	2.8312	150.0133
LYS_28@O	ILE_31@H	ILE_31@N	3479	0.058	2.8948	151.5135
ASP_7@OD2	HE1_6@H	HE1_6@N	3413	0.0569	2.818	163.3015
ASP_7@OD2	ARG_5@H	ARG_5@N	3191	0.0532	2.804	156.5055
MET_35@O	GLY_38@H	GLY_38@N	3182	0.053	2.8927	151.0724
ALA_21@O	GLY_25@H	GLY_25@N	3113	0.0519	2.8683	153.9749
ALA_30@O	GLY_33@H	GLY_33@N	3075	0.0512	2.8882	150.8545

A1.13 Hydrogen bond output for incidence > 5% for Aβ42-Free aMD trajectory data

#Acceptor	DonorH	Donor	Frames	Fraction	Average Distance (Å)	Average Angle (º)
TYR_10@O	HID_14@HD1	HID_14@ND1	10109	0.1685	2.8438	157.9923
GLU_11@O	HID_14@HD1	HID_14@ND1	8252	0.1375	2.8292	157.1548
GLU_3@OE2	ARG_5@HH11	ARG_5@NH1	8239	0.1373	2.8057	157.0983
GLU_3@O	HID_6@HD1	HID_6@ND1	8026	0.1338	2.841	156.6513
GLU_3@OE1	ARG_5@HH11	ARG_5@NH1	7549	0.1258	2.8104	156.5454
TYR_10@O	HID_13@HD1	HID_13@ND1	7544	0.1257	2.8384	156.8063
SER_26@O	GLY_29@H	GLY_29@N	7288	0.1215	2.883	153.7938
GLU_11@O	GLN_15@H	GLN_15@N	7168	0.1195	2.8867	159.0884
GLY_9@O	HID_13@HD1	HID_13@ND1	6887	0.1148	2.851	157.5478
GLU_22@OE2	ARG_5@HH22	ARG_5@NH2	6725	0.1121	2.7849	159.4738
GLU_11@OE1	ARG_5@HH11	ARG_5@NH1	6721	0.112	2.8132	154.8184
GLU_22@OE1	ARG_5@HH22	ARG_5@NH2	6581	0.1097	2.7837	159.073
GLU_3@OE1	ARG_5@HE	ARG_5@NE	6544	0.1091	2.8454	157.5563
GLU_11@OE2	ARG_5@HH11	ARG_5@NH1	6542	0.109	2.812	155.4768
GLU_11@OE1	ARG_5@HE	ARG_5@NE	6492	0.1082	2.8169	156.3667
ALA_21@O	GLY_25@H	GLY_25@N	6370	0.1062	2.8595	152.5282
TYR_10@O	HID_14@H	HID_14@N	6305	0.1051	2.8828	158.3126
ILE_32@O	VAL_36@H	VAL_36@N	6179	0.103	2.8906	160.4436
HID_13@O	LEU_17@H	LEU_17@N	6166	0.1028	2.8852	157.7204
PHE_4@O	SER_8@H	SER_8@N	6138	0.1023	2.8883	158.9613
GLU_3@OE2	ARG_5@HE	ARG_5@NE	5934	0.0989	2.8453	156.9766
ILE_31@O	MET_35@H	MET_35@N	5844	0.0974	2.8867	156.9945
GLU_22@OE1	ARG_5@HH12	ARG_5@NH1	5758	0.096	2.8252	156.959
GLU_11@OE2	ARG_5@HE	ARG_5@NE	5747	0.0958	2.8157	156.4535

VAL_12@O	LYS_16@H	LYS_16@N	5729	0.0955	2.8902	157.8168
GLY_33@O	GLY_37@H	GLY_37@N	5724	0.0954	2.8735	153.4377
VAL_18@O	ALA_30@H	ALA_30@N	5606	0.0934	2.8845	156.9545
GLU_22@OE2	ARG_5@HH12	ARG_5@NH1	5585	0.0931	2.8268	156.8714
MET_35@O	PHE_19@H	PHE_19@N	5418	0.0903	2.8622	161.7928
GLY_29@O	GLY_33@H	GLY_33@N	4834	0.0806	2.8768	155.259
GLU_3@O	HID_6@H	HID_6@N	4792	0.0799	2.8986	157.0659
LEU_17@O	LEU_34@H	LEU_34@N	4734	0.0789	2.8536	155.5415
ALA_30@O	GLY_33@H	GLY_33@N	4657	0.0776	2.8957	151.7469
HID_14@O	VAL_18@H	VAL_18@N	4572	0.0762	2.8879	159.6212
PHE_4@O	ASP_7@H	ASP_7@N	4536	0.0756	2.8958	153.5651
HID_13@O	LYS_16@H	LYS_16@N	4335	0.0722	2.8968	152.5904
ILE_31@O	LEU_34@H	LEU_34@N	4322	0.072	2.8944	151.9185
ASP_7@OD1	ARG_5@HH11	ARG_5@NH1	3897	0.0649	2.8245	155.1493
GLY_38@O	ALA_42@H	ALA_42@N	3872	0.0645	2.8838	159.5575
GLU_11@OE1	GLU_11@H	GLU_11@N	3820	0.0637	2.8311	150.9258
ASP_23@OD1	ARG_5@HH21	ARG_5@NH2	3819	0.0636	2.8019	153.9703
ASP_7@OD2	ARG_5@HH11	ARG_5@NH1	3810	0.0635	2.8281	155.432
VAL_18@O	ALA_21@H	ALA_21@N	3804	0.0634	2.8938	154.3305
GLU_11@O	HID_14@H	HID_14@N	3803	0.0634	2.8944	151.655
TYR_10@O	HID_13@H	HID_13@N	3772	0.0629	2.8967	152.4362
ALA_30@O	LEU_34@H	LEU_34@N	3763	0.0627	2.8878	156.7371
GLY_9@O	HID_13@H	HID_13@N	3720	0.062	2.8884	159.4348
HID_14@O	LEU_17@H	LEU_17@N	3660	0.061	2.9004	152.3687
GLU_11@OE2	GLU_11@H	GLU_11@N	3642	0.0607	2.8279	151.6416
GLU_11@OE2	ARG_5@HH12	ARG_5@NH1	3602	0.06	2.8079	156.541
ASP_23@OD1	ARG_5@HH22	ARG_5@NH2	3526	0.0588	2.8078	159.6668
GLU_11@OE2	ARG_5@HH22	ARG_5@NH2	3447	0.0575	2.8178	155.8447
ASP_23@OD2	ARG_5@HH12	ARG_5@NH1	3417	0.057	2.8045	159.1548
LEU_17@O	MET_35@H	MET_35@N	3413	0.0569	2.891	159.9422
ASP_7@OD2	ARG_5@HE	ARG_5@NE	3387	0.0565	2.8356	157.5432
VAL_12@O	GLN_15@H	GLN_15@N	3349	0.0558	2.8975	152.0923
MET_35@O	GLY_38@H	GLY_38@N	3314	0.0552	2.8944	151.037
GLU_11@OE2	GLN_15@HE22	GLN_15@NE2	3260	0.0543	2.8324	161.6676
PHE_4@O	SER_8@HG	SER_8@OG	3248	0.0541	2.8159	158.5473
ASP_7@OD1	ARG_5@HE	ARG_5@NE	3222	0.0537	2.8301	157.5971
GLN_15@O	PHE_19@H	PHE_19@N	3108	0.0518	2.8795	157.0485
GLY_38@O	ILE_41@H	ILE_41@N	3012	0.0502	2.901	154.1325

A1.14 Hydrogen bond output for incidence > 5% for Aβ42-Cu aMD trajectory data

#Acceptor	DonorH	Donor	Frames	Fraction	Average Distance (Å)	Average Angle (°)
ASP_7@O	HD1_6@HD1	HD1_6@ND1	17012	0.2835	2.7716	164.0919
PHE_4@O	HD1_6@H	HD1_6@N	13040	0.2173	2.751	146.9534
GLY_9@O	HD1_6@HD1	HD1_6@ND1	12959	0.216	2.8137	150.463
ASP_7@OD2	HD1_6@H	HD1_6@N	11204	0.1867	2.6809	157.9705
GLU_22@OE2	HD1_6@HD1	HD1_6@ND1	10798	0.18	2.7242	159.1441
GLU_11@OE2	GLU_11@H	GLU_11@N	10745	0.1791	2.8197	152.6119
GLU_11@OE1	GLU_11@H	GLU_11@N	10419	0.1736	2.8201	152.782
TYR_10@O	HD2_13@H	HD2_13@N	9801	0.1633	2.8771	153.7164
GLU_22@OE1	HD1_6@HD1	HD1_6@ND1	8590	0.1432	2.7206	156.3081
GLU_11@O	GLN_15@H	GLN_15@N	8366	0.1394	2.8773	159.229
ASP_7@OD1	ARG_5@HH11	ARG_5@NH1	8185	0.1364	2.7987	155.4155
ASP_7@OD1	HD1_6@H	HD1_6@N	8147	0.1358	2.6805	157.6492
GLU_22@OE1	HD1_6@H	HD1_6@N	7958	0.1326	2.6757	155.624
GLU_11@O	HID_14@HD1	HID_14@ND1	7792	0.1299	2.8348	157.5372
GLU_11@O	HID_14@H	HID_14@N	7660	0.1277	2.8943	154.3254
ASP_7@OD2	ARG_5@HH11	ARG_5@NH1	6521	0.1087	2.7999	155.8489
TYR_10@O	HD2_13@HD1	HD2_13@ND1	6011	0.1002	2.8498	150.7508
GLY_33@O	GLY_37@H	GLY_37@N	5728	0.0955	2.8745	154.23
ILE_32@O	VAL_36@H	VAL_36@N	5036	0.0839	2.8918	160.0795
ILE_31@O	MET_35@H	MET_35@N	4986	0.0831	2.8885	157.3612
TYR_10@O	HID_14@HD1	HID_14@ND1	4839	0.0806	2.8527	157.07
GLY_9@O	HD2_13@HD1	HD2_13@ND1	4776	0.0796	2.8517	154.8467
ASP_7@OD1	ARG_5@HE	ARG_5@NE	4763	0.0794	2.8545	151.1632
LEU_17@O	ALA_21@H	ALA_21@N	4639	0.0773	2.8856	157.8784
GLU_22@OE2	HD1_6@H	HD1_6@N	4629	0.0771	2.6776	156.1029
GLY_29@O	GLY_33@H	GLY_33@N	4624	0.0771	2.8756	154.7749
LEU_17@O	PHE_20@H	PHE_20@N	4613	0.0769	2.8946	154.0111
ALA_21@O	GLY_25@H	GLY_25@N	4604	0.0767	2.8675	153.8336
ILE_31@O	LEU_34@H	LEU_34@N	4192	0.0699	2.8996	151.6914
GLY_38@O	ALA_42@H	ALA_42@N	4022	0.067	2.8821	159.9079
HD1_6@O	GLY_9@H	GLY_9@N	3925	0.0654	2.9041	157.4778
VAL_12@O	LYS_16@H	LYS_16@N	3867	0.0644	2.8855	157.7604
ASP_7@OD2	ARG_5@HE	ARG_5@NE	3730	0.0622	2.8581	151.7441
ALA_30@O	GLY_33@H	GLY_33@N	3406	0.0568	2.8877	150.8245
GLY_38@O	ILE_41@H	ILE_41@N	3369	0.0561	2.8946	154.085
LYS_28@O	ILE_32@H	ILE_32@N	3287	0.0548	2.8893	161.1104
ASP_23@OD2	ARG_5@HH12	ARG_5@NH1	3285	0.0548	2.8107	156.5789
ASP_23@OD1	ARG_5@HH12	ARG_5@NH1	3243	0.0541	2.8083	156.5402
GLY_25@O	GLY_29@H	GLY_29@N	3163	0.0527	2.8756	154.1473
GLU_3@OE1	ARG_5@HH11	ARG_5@NH1	3149	0.0525	2.8086	155.9085
ASP_23@OD2	ARG_5@HH22	ARG_5@NH2	3148	0.0525	2.8124	156.3999

ASP_23@OD1	ARG_5@HH22	ARG_5@NH2	3124	0.0521	2.8123	156.6182
GLU_22@O	GLY_25@H	GLY_25@N	3031	0.0505	2.8917	152.1999
ILE_32@O	MET_35@H	MET_35@N	3028	0.0505	2.8989	151.4952

A1.15 Hydrogen bond output for incidence > 5% for Aβ42-Zn aMD trajectory data

#Acceptor	DonorH	Donor	Frames	Fraction	Average DIstance (Å)	Average Angle (°)
ASP_7@OD1	ARG_5@HH11	ARG_5@NH1	22140	0.369	2.8217	156.6456
ASP_7@OD2	ARG_5@HH11	ARG_5@NH1	21519	0.3587	2.8197	156.5122
ASP_7@OD2	ARG_5@HE	ARG_5@NE	18161	0.3027	2.8426	157.6384
ASP_7@OD1	ARG_5@HE	ARG_5@NE	16565	0.2761	2.8437	157.6695
GU1_11@O	HIE_13@H	HIE_13@N	16302	0.2717	2.792	146.0797
ILE_32@O	VAL_36@H	VAL_36@N	8325	0.1388	2.8893	159.9665
GLU_3@OE1	GLU_3@H	GLU_3@N	7828	0.1305	2.823	152.1127
GLU_3@OE2	GLU_3@H	GLU_3@N	7699	0.1283	2.821	152.0458
GU1_11@O	HE2_14@H	HE2_14@N	7404	0.1234	2.8935	158.6457
ILE_31@O	MET_35@H	MET_35@N	6995	0.1166	2.8854	157.3489
GLY_29@O	GLY_33@H	GLY_33@N	6785	0.1131	2.8758	155.2799
LEU_17@O	ALA_21@H	ALA_21@N	6487	0.1081	2.8844	157.7723
GLY_33@O	GLY_37@H	GLY_37@N	5744	0.0957	2.8734	153.3268
HIE_13@O	LYS_16@H	LYS_16@N	5720	0.0953	2.8987	156.532
ARG_5@O	ALA_2@H	ALA_2@N	5676	0.0946	2.887	157.7901
LYS_16@O	PHE_20@H	PHE_20@N	5534	0.0922	2.8836	158.2175
ALA_21@O	GLY_25@H	GLY_25@N	5109	0.0852	2.8665	154.2739
ALA_30@O	LEU_34@H	LEU_34@N	4971	0.0828	2.8889	156.3232
LEU_17@O	PHE_20@H	PHE_20@N	4908	0.0818	2.8986	153.3081
GLY_38@O	ALA_42@H	ALA_42@N	4857	0.0809	2.8826	159.6142
GLY_9@O	VAL_12@H	VAL_12@N	4436	0.0739	2.903	155.6667
ILE_31@O	LEU_34@H	LEU_34@N	4370	0.0728	2.8984	151.8738
ILE_32@O	MET_35@H	MET_35@N	4150	0.0692	2.8949	150.7106
GLY_38@O	ILE_41@H	ILE_41@N	3958	0.066	2.9	154.3708
HIE_13@ND1	GLN_15@H	GLN_15@N	3941	0.0657	2.9258	155.1747
LYS_16@O	PHE_19@H	PHE_19@N	3930	0.0655	2.8974	153.1231
ALA_2@O	ARG_5@H	ARG_5@N	3875	0.0646	2.8955	156.8203
LYS_28@O	ILE_32@H	ILE_32@N	3687	0.0614	2.8912	161.7382
SER_26@O	GLY_29@H	GLY_29@N	3586	0.0598	2.8925	153.1863
MET_35@O	GLY_38@H	GLY_38@N	3558	0.0593	2.8926	150.5306
GLU_3@O	AP1_1@H2	AP1_1@N	3347	0.0558	2.8416	154.5333
ALA_2@O	PHE_4@H	PHE_4@N	3243	0.0541	2.8187	147.5372
ASP_23@OD2	ASN_27@HD22	ASN_27@ND2	3159	0.0527	2.8337	162.4905
GLN_15@O	VAL_18@H	VAL_18@N	3137	0.0523	2.9	153.9415
ASP_23@OD1	ASN_27@HD22	ASN_27@ND2	3109	0.0518	2.8349	162.6219
HE2_14@O	LYS_16@H	LYS_16@N	3078	0.0513	2.8331	147.7686

ALA_30@O	GLY_33@H	GLY_33@N	3015	0.0503	2.8888	149.9972
VAL_39@O	NME_43@H	NME_43@N	3009	0.0502	2.8803	153.958

A1.16 Hydrogen bond output for incidence > 5% for Aβ42-Fe aMD trajectory data

#Acceptor	DonorH	Donor	Frames	Fraction	Average Distance (Å)	Average Angle (°)
TYR_10@O	HE2_14@HE2	HE2_14@NE2	33897	0.5649	2.8433	158.8743
ASP_7@OD1	ARG_5@HH11	ARG_5@NH1	14101	0.235	2.8161	154.1098
ASP_7@OD2	ARG_5@HH11	ARG_5@NH1	13319	0.222	2.8182	154.379
HIE_13@O	AP1_1@H3	AP1_1@N	12511	0.2085	2.8749	153.3709
AP1_1@OD1	GLN_15@H	GLN_15@N	11182	0.1864	2.8714	159.24
ASP_7@OD1	ARG_5@HE	ARG_5@NE	9817	0.1636	2.8422	153.364
ASP_7@OD2	ARG_5@HE	ARG_5@NE	9156	0.1526	2.8439	153.6909
LEU_17@O	HE1_6@HE2	HE1_6@NE2	8761	0.146	2.8291	154.4395
ILE_32@O	VAL_36@H	VAL_36@N	8630	0.1438	2.8902	160.0056
PHE_4@O	HE1_6@H	HE1_6@N	8612	0.1435	2.8391	146.0863
ASP_7@OD2	ARG_5@H	ARG_5@N	8600	0.1433	2.8188	154.5579
ASP_7@OD2	HE1_6@H	HE1_6@N	8568	0.1428	2.8283	163.2466
GLY_29@O	GLY_33@H	GLY_33@N	8067	0.1344	2.8767	154.4859
ILE_31@O	MET_35@H	MET_35@N	7426	0.1238	2.8878	157.7985
AP1_1@OD2	GLN_15@H	GLN_15@N	7387	0.1231	2.8769	158.8035
ASP_7@OD1	ARG_5@H	ARG_5@N	7352	0.1225	2.8138	155.0593
ASP_7@OD1	HE1_6@H	HE1_6@N	7168	0.1195	2.833	163.0951
GLY_33@O	GLY_37@H	GLY_37@N	7152	0.1192	2.8729	153.3955
GLU_22@O	GLY_25@H	GLY_25@N	6861	0.1143	2.8901	151.0305
GLU_3@O	ARG_5@H	ARG_5@N	5915	0.0986	2.8099	150.0073
ALA_21@O	ASN_27@HD22	ASN_27@ND2	5813	0.0969	2.8626	157.8248
LYS_28@O	ILE_32@H	ILE_32@N	5690	0.0948	2.8903	161.8944
GLU_3@OE2	AP1_1@H2	AP1_1@N	5376	0.0896	2.8383	153.6761
GLU_3@OE1	AP1_1@H2	AP1_1@N	4963	0.0827	2.8415	153.457
GLY_38@O	ALA_42@H	ALA_42@N	4887	0.0814	2.8823	159.7451
GLU_22@OE1	ARG_5@HH12	ARG_5@NH1	4732	0.0789	2.8041	159.3689
ALA_30@O	LEU_34@H	LEU_34@N	4600	0.0767	2.887	156.3937
GLU_22@OE2	ARG_5@HH12	ARG_5@NH1	4584	0.0764	2.8031	159.3604
GLU_22@OE2	ARG_5@HH22	ARG_5@NH2	4561	0.076	2.8081	159.3473
GLU_22@OE1	ARG_5@HH22	ARG_5@NH2	4493	0.0749	2.8067	159.0598
PHE_20@O	GLY_29@H	GLY_29@N	4391	0.0732	2.8734	156.0427
ASP_23@OD1	HE1_6@HE2	HE1_6@NE2	4278	0.0713	2.7922	160.6651
ILE_31@O	LEU_34@H	LEU_34@N	4252	0.0709	2.8984	151.8037
ILE_32@O	MET_35@H	MET_35@N	4158	0.0693	2.896	150.1498
GLN_15@OE1	HE2_14@H	HE2_14@N	3918	0.0653	2.8525	158.3805
GLY_38@O	ILE_41@H	ILE_41@N	3882	0.0647	2.8983	153.9961
PHE_4@O	ARG_5@HE	ARG_5@NE	3808	0.0635	2.8576	154.6362

ASP_23@O	SER_26@H	SER_26@N	3702	0.0617	2.896	150.4482
GLY_33@O	GLY_38@H	GLY_38@N	3619	0.0603	2.8691	152.9283
GLU_22@OE1	GLU_22@H	GLU_22@N	3549	0.0592	2.8158	152.6585
VAL_18@O	ALA_21@H	ALA_21@N	3529	0.0588	2.8969	154.1661
GLU_3@OE2	ARG_5@H	ARG_5@N	3465	0.0578	2.8141	152.0603
GLU_11@OE2	GLU_11@H	GLU_11@N	3451	0.0575	2.8285	150.6567
GLU_3@OE1	ARG_5@H	ARG_5@N	3406	0.0568	2.8074	152.2267
VAL_39@O	NME_43@H	NME_43@N	3397	0.0566	2.8772	153.7417
ASN_27@O	ALA_30@H	ALA_30@N	3347	0.0558	2.8998	154.2749
ASP_23@OD2	ASN_27@HD21	ASN_27@ND2	3327	0.0554	2.8464	163.4276
GLU_11@OE1	GLU_11@H	GLU_11@N	3301	0.055	2.8262	151.0593
GLU_22@OE2	GLU_22@H	GLU_22@N	3297	0.0549	2.8173	152.024
VAL_18@O	GLU_22@H	GLU_22@N	3268	0.0545	2.89	157.8239
LYS_28@O	ILE_31@H	ILE_31@N	3193	0.0532	2.8956	150.6043
LEU_17@O	ALA_21@H	ALA_21@N	3054	0.0509	2.8888	156.8351
SER_26@O	GLY_29@H	GLY_29@N	3048	0.0508	2.8956	153.7806
LEU_17@O	PHE_20@H	PHE_20@N	3046	0.0508	2.9004	153.2004

A2 Appendices for Chapter 5

A2.1 Hydrogen bond output for incidence > 5% for dimeric Aβ42-Free aMD trajectory data

#Acceptor	DonorH	Donor	Frames	Fraction	Average Distance (Å)	Average Angle (°)
GLY_68@H	LYS_71@HZ2	LYS_71@HZ1	17313	0.1767	2.8663	154.7775
TYR_53@CE2	HID_56@CD2	HID_56@NE2	16700	0.1704	2.858	157.6563
GLY_72@H	ILE_75@HD12	ILE_75@HD11	16469	0.1681	2.8737	154.6085
GLU_11@O	HID_14@HD1	HID_14@ND1	16028	0.1636	2.8288	157.7863
GLU_54@HG3	HID_57@CD2	HID_57@NE2	15350	0.1566	2.864	158.2599
LEU_60@CD2	PHE_63@CD2	PHE_63@HE2	14931	0.1524	2.8815	157.178
GLU_3@OE2	GLU_3@H	GLU_3@N	14874	0.1518	2.8278	151.1571
ALA_21@O	GLY_25@H	GLY_25@N	14778	0.1508	2.8695	154.0419
TYR_10@O	HID_13@HD1	HID_13@ND1	14638	0.1494	2.8367	157.091
LEU_17@O	ALA_21@H	ALA_21@N	14617	0.1492	2.8863	158.4287
GLU_54@CG	ARG_48@CZ	ARG_48@HE	14464	0.1476	2.8085	154.2625
GLU_3@OE1	GLU_3@H	GLU_3@N	14162	0.1445	2.8283	150.9885
GLY_52@H	HID_56@CB	HID_56@HA	14037	0.1432	2.8584	157.7194
ILE_75@CD1	MET_78@HE2	MET_78@HE1	14000	0.1429	2.8879	159.294
GLU_54@HB3	ARG_48@CZ	ARG_48@HE	13822	0.141	2.8106	153.6729
GLU_46@CG	ALA_45@HB2	ALA_45@HB1	13652	0.1393	2.8254	151.0927
GLU_46@HB3	ALA_45@HB2	ALA_45@HB1	13630	0.1391	2.8252	150.9553
VAL_67@CG2	ASN_70@HD21	ASN_70@ND2	13423	0.137	2.879	160.5184
GLY_29@O	GLY_33@H	GLY_33@N	13204	0.1347	2.8712	154.8648
GLU_65@CG	ARG_48@HH21	ARG_48@NH2	12800	0.1306	2.7992	158.872

ASP_66@HB3	ASN_70@HB2	ASN_70@CB	12722	0.1298	2.8534	157.9328
GLU_65@OE1	ARG_48@HH12	ARG_48@NH1	12168	0.1242	2.8074	157.4433
LYS_71@NZ	ILE_74@HD12	ILE_74@HD11	11962	0.1221	2.8889	161.6763
ILE_32@O	VAL_36@H	VAL_36@N	11890	0.1213	2.8879	160.2889
GLU_54@HB3	ARG_48@HG3	ARG_48@HG2	11843	0.1208	2.8211	154.8541
ILE_74@CD1	LEU_77@HD22	LEU_77@HD21	11708	0.1195	2.8904	156.1571
VAL_55@CG2	LYS_59@HD3	LYS_59@HD2	11622	0.1186	2.8756	158.7733
GLY_25@O	GLY_29@H	GLY_29@N	11429	0.1166	2.865	154.1249
GLU_54@CG	ARG_48@HG3	ARG_48@HG2	11378	0.1161	2.8227	154.8471
GLY_33@O	GLY_37@H	GLY_37@N	11337	0.1157	2.872	154.0286
TYR_10@O	HID_14@HD1	HID_14@ND1	11174	0.114	2.8422	157.3165
GLU_11@O	GLN_15@H	GLN_15@N	11091	0.1132	2.8841	159.2943
GLU_65@HB3	ARG_48@HH21	ARG_48@NH2	10935	0.1116	2.8011	158.4413
GLU_65@OE2	ARG_48@HH12	ARG_48@NH1	10808	0.1103	2.8102	157.1855
TYR_53@CE2	HID_57@CB	HID_57@HA	10781	0.11	2.8478	157.8209
GLU_11@OE1	ARG_5@HE	ARG_5@NE	10723	0.1094	2.8286	155.1284
GLU_11@OE2	ARG_5@HE	ARG_5@NE	10674	0.1089	2.8278	154.8751
GLU_65@HG3	GLY_68@HA2	GLY_68@CA	10263	0.1047	2.8796	156.0562
TYR_10@O	HID_14@H	HID_14@N	9829	0.1003	2.8825	158.7561
GLU_11@OE2	ARG_5@HH21	ARG_5@NH2	9779	0.0998	2.8076	154.3933
GLY_76@H	VAL_79@HG22	VAL_79@HG21	9686	0.0988	2.8764	153.4706
GLU_3@O	HID_6@HD1	HID_6@ND1	9616	0.0981	2.8408	157.0097
GLU_11@OE1	ARG_5@HH21	ARG_5@NH2	9232	0.0942	2.8082	154.4172
LYS_16@O	PHE_20@H	PHE_20@N	9049	0.0923	2.8869	159.4797
PHE_47@CE2	ASP_50@OD1	ASP_50@CG	8980	0.0916	2.8867	158.3027
ALA_64@CB	VAL_67@HG22	VAL_67@HG21	8587	0.0876	2.8731	154.6195
GLY_9@O	HID_13@HD1	HID_13@ND1	8579	0.0875	2.8492	157.4027
VAL_24@O	LYS_28@H	LYS_28@N	8545	0.0872	2.882	160.3645
GLU_11@OE2	GLU_11@H	GLU_11@N	8460	0.0863	2.8213	151.579
GLU_22@O	SER_26@H	SER_26@N	8442	0.0861	2.8778	155.6634
GLU_11@OE1	GLU_11@H	GLU_11@N	8206	0.0837	2.8196	151.7387
SER_69@HB2	GLY_72@HA2	GLY_72@CA	8159	0.0833	2.889	155.4792
ALA_73@CB	GLY_76@HA2	GLY_76@CA	8104	0.0827	2.8861	155.7454
LYS_28@O	ILE_32@H	ILE_32@N	8058	0.0822	2.891	162.5009
PHE_62@O	ARG_48@HH21	ARG_48@NH2	8024	0.0819	2.8515	155.029
PHE_47@CE2	HID_49@CD2	HID_49@NE2	7940	0.081	2.8754	154.9611
PHE_4@O	SER_8@H	SER_8@N	7936	0.081	2.8856	158.1879
ASP_23@O	ASN_27@HD21	ASN_27@ND2	7735	0.0789	2.8485	158.9523
LYS_59@NZ	PHE_62@CD2	PHE_62@HE2	7706	0.0786	2.8883	158.3025
ILE_31@O	MET_35@H	MET_35@N	7603	0.0776	2.8895	157.7621
GLN_15@O	PHE_19@H	PHE_19@N	7493	0.0765	2.8849	159.0137
HID_56@HE1	LYS_59@HZ2	LYS_59@HZ1	7391	0.0754	2.8856	157.8531
ASP_66@HB3	SER_69@OG	SER_69@HB3	7270	0.0742	2.8877	158.0033
PHE_20@O	VAL_24@H	VAL_24@N	7258	0.0741	2.8933	160.4077

GLU_11@O	HID_14@H	HID_14@N	7176	0.0732	2.8951	153.1703
LEU_17@O	PHE_20@H	PHE_20@N	6900	0.0704	2.8967	152.9631
VAL_12@O	LYS_16@H	LYS_16@N	6869	0.0701	2.8898	158.433
TYR_10@O	HID_13@H	HID_13@N	6831	0.0697	2.8982	153.3967
GLU_3@OE1	ARG_5@HH11	ARG_5@NH1	6649	0.0678	2.8055	157.54
VAL_55@CG2	GLN_58@HE21	GLN_58@NE2	6632	0.0677	2.8879	158.698
GLY_33@O	SER_51@HA	SER_51@CA	6614	0.0675	2.8668	157.2345
VAL_55@CG2	GLN_58@HG2	GLN_58@CG	6590	0.0672	2.8886	154.0803
GLU_11@OE1	ARG_48@NH1	ARG_48@HE	6551	0.0668	2.7976	158.4889
GLU_11@OE1	ARG_5@HH22	ARG_5@NH2	6506	0.0664	2.8061	156.5742
GLU_11@OE2	ARG_48@NH1	ARG_48@HE	6450	0.0658	2.7951	158.8848
HID_13@O	LEU_17@H	LEU_17@N	6412	0.0654	2.8899	158.1828
GLU_11@OE2	ARG_5@HH22	ARG_5@NH2	6368	0.065	2.8066	156.1448
VAL_18@O	GLU_22@H	GLU_22@N	6246	0.0637	2.8913	157.0253
GLU_3@OE2	ARG_5@HE	ARG_5@NE	6205	0.0633	2.8473	157.9364
GLU_11@OE2	ARG_5@HH12	ARG_5@NH1	6167	0.0629	2.8107	156.1419
GLU_3@O	HID_6@H	HID_6@N	6138	0.0626	2.8975	154.9731
TYR_53@CE2	HID_56@CB	HID_56@HA	6107	0.0623	2.8559	156.4241
HID_13@O	LYS_16@H	LYS_16@N	6105	0.0623	2.8985	153.6688
ALA_30@O	LEU_34@H	LEU_34@N	6091	0.0622	2.8878	156.3607
ASP_50@O	ARG_48@HH11	ARG_48@NH1	6089	0.0621	2.8238	157.1805
GLU_11@OE1	ARG_48@NE	ARG_48@HD2	6038	0.0616	2.8127	158.0772
GLU_54@HG3	GLN_58@HG2	GLN_58@CG	5952	0.0607	2.8744	158.7666
GLU_3@OE1	ARG_5@HE	ARG_5@NE	5923	0.0604	2.8447	157.2748
GLU_11@OE1	ARG_5@HH12	ARG_5@NH1	5916	0.0604	2.8107	156.0396
LEU_60@CD2	PHE_62@CD2	PHE_62@HE2	5862	0.0598	2.8788	150.7056
GLU_3@OE2	ARG_5@HH11	ARG_5@NH1	5859	0.0598	2.8053	157.3266
GLU_46@O	LEU_60@H	LEU_60@N	5794	0.0591	2.8703	160.7261
TYR_53@CE2	HID_57@CD2	HID_57@NE2	5722	0.0584	2.8442	157.309
GLU_54@CG	GLN_58@HG2	GLN_58@CG	5554	0.0567	2.8266	161.9424
GLU_54@HG3	HID_57@CB	HID_57@HA	5535	0.0565	2.8597	155.7845
GLU_22@O	GLY_25@H	GLY_25@N	5533	0.0565	2.8867	152.2221
GLU_11@OE2	ARG_48@NE	ARG_48@HD2	5508	0.0562	2.8093	158.6057
HID_14@O	LEU_17@H	LEU_17@N	5449	0.0556	2.9002	153.4212
MET_35@O	GLY_38@H	GLY_38@N	5408	0.0552	2.8925	150.7321
PHE_47@CE2	SER_51@HA	SER_51@CA	5376	0.0549	2.8145	158.171
GLU_65@O	VAL_40@H	VAL_40@N	5317	0.0543	2.87	160.8302
GLU_3@O	ASP_7@H	ASP_7@N	5304	0.0541	2.8844	159.6533
ILE_32@O	MET_35@H	MET_35@N	5085	0.0519	2.892	149.9939
GLU_65@CG	GLU_65@OE1	GLU_65@CD	5049	0.0515	2.8285	154.1234
GLU_65@HB3	GLU_65@OE1	GLU_65@CD	5022	0.0512	2.8278	153.872
LEU_60@O	ARG_48@H	ARG_48@N	4908	0.0501	2.8705	155.5746
GLU_46@HG3	HID_49@CD2	HID_49@NE2	4908	0.0501	2.8802	160.6084

A2.2 Hydrogen bond output for incidence > 5% for dimeric Aβ42-Zn aMD trajectory data

#Acceptor	DonorH	Donor	Frames	Fraction	Average DIstance (Å)	Average Angle (°)
TYR_10@O	HID_13@HD1	HID_13@ND1	21126	0.1761	2.8373	157.515
ASP_23@OD1	ARG_5@HH12	ARG_5@NH1	19536	0.1628	2.8077	157.5388
ASP_23@OD1	ARG_5@HH22	ARG_5@NH2	19230	0.1603	2.8032	158.146
ASP_23@OD2	ARG_5@HH12	ARG_5@NH1	18682	0.1557	2.8087	157.8137
ASP_23@OD2	ARG_5@HH22	ARG_5@NH2	18080	0.1507	2.8029	158.4398
GLU_3@OE1	GLU_3@H	GLU_3@N	16129	0.1344	2.8294	151.3845
GLU_46@OE2	GLU_46@H	GLU_46@N	15444	0.1287	2.8248	151.289
GLU_3@OE2	GLU_3@H	GLU_3@N	15421	0.1285	2.8297	151.35
ASP_50@OD2	ARG_48@HH21	ARG_48@NH2	15289	0.1274	2.8157	157.2885
GLU_3@OE2	ARG_5@HH11	ARG_5@NH1	14978	0.1248	2.8041	157.9086
GLU_3@OE1	ARG_5@HH11	ARG_5@NH1	14885	0.124	2.8065	157.9219
GLU_46@OE1	GLU_46@H	GLU_46@N	14758	0.123	2.8261	151.2726
ASP_50@OD1	ARG_48@HE	ARG_48@NE	14697	0.1225	2.8333	158.0092
ASP_50@OD2	ARG_48@HE	ARG_48@NE	13778	0.1148	2.8349	157.5396
TYR_53@O	HD2_57@H	HD2_57@N	13558	0.113	2.8515	161.1523
ASP_50@OD1	ARG_48@HH21	ARG_48@NH2	13537	0.1128	2.8186	156.8406
ILE_32@O	VAL_36@H	VAL_36@N	13347	0.1112	2.8896	160.2104
GLU_3@OE2	ARG_5@HE	ARG_5@NE	13048	0.1087	2.8388	158.0803
GLU_3@OE1	ARG_5@HE	ARG_5@NE	13035	0.1086	2.8382	158.1685
GLY_33@O	GLY_37@H	GLY_37@N	12945	0.1079	2.8721	153.7285
GU2_54@OE1	GU2_54@H	GU2_54@N	12768	0.1064	2.8518	146.8741
GLU_3@OE1	ARG_48@HH22	ARG_48@NH2	12648	0.1054	2.8099	156.1653
ILE_75@O	VAL_79@H	VAL_79@N	12592	0.1049	2.8868	159.4626
GLY_72@O	GLY_76@H	GLY_76@N	12379	0.1032	2.8757	155.6637
GLU_46@OE1	ARG_48@HE	ARG_48@NE	12107	0.1009	2.8271	158.7197
GU1_11@O	HD1_14@H	HD1_14@N	12069	0.1006	2.8944	160.3107
GU2_54@O	HID_56@H	HID_56@N	11895	0.0991	2.8008	148.8896
GLU_3@OE2	ARG_48@HH22	ARG_48@NH2	11558	0.0963	2.8144	155.4517
ILE_31@O	MET_35@H	MET_35@N	11508	0.0959	2.8867	157.2746
GLU_46@OE2	ARG_48@HE	ARG_48@NE	11003	0.0917	2.8348	158.5911
GLU_3@OE2	ARG_48@HH12	ARG_48@NH1	10790	0.0899	2.8021	157.6346
ILE_74@O	MET_78@H	MET_78@N	10756	0.0896	2.8868	157.9412
GLY_76@O	GLY_80@H	GLY_80@N	10734	0.0895	2.8712	153.6284
ALA_64@O	GLY_68@H	GLY_68@N	10589	0.0882	2.8636	155.9034
GLU_3@OE1	ARG_48@HH12	ARG_48@NH1	9714	0.0809	2.8082	156.8602
GU2_54@O	GLN_58@H	GLN_58@N	9676	0.0806	2.8784	160.3668
GLU_65@O	GLY_68@H	GLY_68@N	9664	0.0805	2.8872	152.5128
GLU_46@OE2	ARG_48@HH21	ARG_48@NH2	9257	0.0771	2.8166	154.8531
ILE_31@O	LEU_34@H	LEU_34@N	9071	0.0756	2.8942	152.5679
GU2_54@O	HD2_57@H	HD2_57@N	9030	0.0752	2.8629	155.3639
GLN_15@OE1	GLN_15@H	GLN_15@N	8493	0.0708	2.8403	149.0124

GLN_58@OE1	GLN_58@H	GLN_58@N	8364	0.0697	2.8453	149.5239
GLU_46@OE1	ARG_48@HH11	ARG_48@NH1	8338	0.0695	2.7992	157.8231
ILE_32@O	MET_35@H	MET_35@N	8307	0.0692	2.8937	151.4893
GLU_65@O	SER_69@H	SER_69@N	8197	0.0683	2.8786	156.6162
GLU_3@O	HID_6@HD1	HID_6@ND1	8118	0.0677	2.8373	156.996
GLY_29@O	GLY_33@H	GLY_33@N	8064	0.0672	2.8751	154.9617
GLY_81@O	ALA_85@H	ALA_85@N	7975	0.0665	2.8813	159.674
ALA_73@O	LEU_77@H	LEU_77@N	7954	0.0663	2.8863	156.3251
GLU_46@O	HID_49@HD1	HID_49@ND1	7859	0.0655	2.8333	156.7271
LEU_17@O	ALA_21@H	ALA_21@N	7717	0.0643	2.8873	159.5406
LEU_17@O	PHE_20@H	PHE_20@N	7701	0.0642	2.8907	153.5071
GLY_25@O	LEU_17@H	LEU_17@N	7691	0.0641	2.8658	159.9647
ASP_23@O	SER_26@H	SER_26@N	7624	0.0635	2.8979	154.398
GLY_9@O	GU1_11@H	GU1_11@N	7550	0.0629	2.8175	148.6045
GU1_11@OE1	ARG_5@HE	ARG_5@NE	7482	0.0624	2.8601	155.2437
LEU_60@O	ALA_64@H	ALA_64@N	7481	0.0623	2.8797	157.0939
GLY_52@O	HID_56@HD1	HID_56@ND1	7377	0.0615	2.8491	157.6902
PHE_47@O	HID_49@HD1	HID_49@ND1	7301	0.0608	2.8524	158.4355
GLY_38@O	ALA_42@H	ALA_42@N	7262	0.0605	2.8821	159.6528
MET_78@O	GLY_81@H	GLY_81@N	7204	0.06	2.8872	151.4861
ILE_75@O	MET_78@H	MET_78@N	7116	0.0593	2.8959	151.7895
LEU_60@O	PHE_63@H	PHE_63@N	7030	0.0586	2.8994	154.521
GU1_11@OE1	ARG_5@HH21	ARG_5@NH2	6972	0.0581	2.8479	153.8723
GLU_46@OE1	ARG_48@HH21	ARG_48@NH2	6960	0.058	2.8287	153.2094
ALA_73@O	GLY_76@H	GLY_76@N	6936	0.0578	2.8871	151.1771
ILE_74@O	LEU_77@H	LEU_77@N	6925	0.0577	2.8979	151.8551
GLU_46@OE2	ARG_48@HH11	ARG_48@NH1	6835	0.057	2.8013	157.1671
ALA_30@O	GLY_33@H	GLY_33@N	6700	0.0558	2.8876	151.6043
HD2_57@O	HID_56@HD1	HID_56@ND1	6665	0.0555	2.8642	156.1048
GLY_81@O	ILE_84@H	ILE_84@N	6601	0.055	2.9	155.0668
VAL_61@O	ALA_64@H	ALA_64@N	6572	0.0548	2.9002	155.6911
PHE_19@O	GLU_22@H	GLU_22@N	6493	0.0541	2.8998	154.2666
PHE_20@O	VAL_24@H	VAL_24@N	6490	0.0541	2.8912	158.9993
ASP_23@O	ASN_27@H	ASN_27@N	6393	0.0533	2.882	157.1784
GLU_22@OE2	GLU_22@H	GLU_22@N	6372	0.0531	2.8228	151.0826
GLU_3@OE2	ARG_5@H	ARG_5@N	6361	0.053	2.8821	162.2932
HD1_14@O	LEU_17@H	LEU_17@N	6350	0.0529	2.897	155.2337
MET_35@O	GLY_38@H	GLY_38@N	6323	0.0527	2.8889	151.0641
GLU_22@OE1	GLU_22@H	GLU_22@N	6314	0.0526	2.8227	150.904
VAL_18@O	GLU_22@H	GLU_22@N	6282	0.0524	2.8892	157.916
LYS_71@O	ILE_75@H	ILE_75@N	6270	0.0522	2.8905	161.7144
GLU_46@OE1	PHE_47@H	PHE_47@N	6205	0.0517	2.8457	152.5195
PHE_63@O	ASP_66@H	ASP_66@N	6178	0.0515	2.8921	152.926
HID_56@O	LEU_60@H	LEU_60@N	6110	0.0509	2.8807	159.4096

GLY_68@O	GLY_72@H	GLY_72@N	6068	0.0506	2.8722	154.1922