

**Exploration of the Effect of Metal-Ions Interaction on
Biomolecular Systems**

Nadiyah Alshammari

A thesis submitted to
Cardiff University
in accordance with the requirements for the degree of
Doctor of Philosophy



School of Chemistry
Cardiff University
February 2022

Publications

Parts of this thesis have been published in:

- 1) N. Alshammari and J. A. Platts, Theoretical study of copper binding to GHK peptide, *Comput. Biol. Chem.*, 2020, 86, 107265.
- 2) N. Al-Shammari, L. Savva, O. Kennedy-Britten and J. A. Platts, Forcefield evaluation and accelerated molecular dynamics simulation of Zn(II) binding to N-terminus of amyloid- β , *Comput. Biol. Chem.*, 2021, 107540.
- 3) N. Alshammari and J. A. Platts, Can ionic effects induce α -sheet conformation of Peptides? *Chem. Phys. Lett.*, 2021, 784, 139095.

Acknowledgements

A special thanks to the government of Saudi Arabia; Ministry of Education for sponsoring my PhD studies. Thanks for giving me this excellent opportunity to achieve my dream.

I want to express my gratitude to Dr Jamie Platts for his excellent supervision, valuable advice, and continuous support to guide me during my doctoral research. Without his knowledgeable advice and encouragement, I could not be here. I want to thank my amazingly supportive peers who helped me during my studies, Oliver Kennedy-Britten and Loizos Savva. I just wanted to say thank you for always having my back. In my heart, you are not just colleagues; you are also friends. I also wish to thank Dr Shaun T. Mutter, the first person who taught me at the beginning of my PhD. I would like to acknowledge Dr Mala Alhaji for encouraging me to overcome the obstacles as a first-year graduate student.

I must acknowledge the first sparkle of this effort by my mother, Motairah, for giving me the courage and her honest prayer. I would thank my brothers and sisters to believe in me. I must express my profound gratitude to my wonderful husband, Hani, who supported me emotionally and morally throughout my years of study until this work came to the surface. Thank you for being there for me when I thought I could not go on any longer; you gave me the strength I needed to keep moving forward. I would also like to thank my kids, Obaid, Saba, and Shouq, for their patience with me being a busy mother. Finally, specific thank to my friend, Haifa for a cherished time spent together in and out the university. Your friendship was special for me and made my studying years less stressful and lonely.

I dedicate this work to my mother and the memory of my father, Mohsen (rest in peace). His dream for me was to see me holding a high level of education, and achieving his dream was mine too. I dedicate it also to my husband, kids, and all who made this study likely.

Abstract

Herein, complexes of amyloid- β (1-16) monomer with physiologically relevant metal cations such as zinc, iron, and copper, have been analysed using density functional theory DFT, semiempirical GFN2-xTB, and molecular mechanics methods. A short peptide (GHK) bound to copper was used as a model for examination of method efficacy. The semiempirical method under review was shown to reproduce the DFT energy and geometry obtained with ligand field molecular mechanics (LFMM). This made it then appropriate for further application on three different fragmentation lengths of A β -16 bonded to Cu(II). Accelerated molecular dynamics (aMD) simulations with the AMBER14SB force field were used to simulate the free and metal-bound A β -16 peptide with Zn(II), Fe(II), and Cu(II), through different binding modes. The simulation showed all metals stabilized the peptide mobility and increased the compactness in terms of RMSD, Rg, and RMSF, compared to that in the unbound monomer. The most frequent salt bridge interaction in all metal-A β forms, was found between Arg5 and Asp7 amino acids. The aMD simulations also showed that the number of coordinating atoms, as well as the element and/or residues, influence the overall structure, size, and stability. The observation of α -sheet secondary structure in A β simulations led us to extend the study to different proteins that have been reported to have this uncommon structure, also related to A β aggregation. DFT and semiempirical GFN2-xTB calculations were used on the modelled α -sheet residues of the peptides. The α -helix is the most stable form, compared to α -sheet and β -strand conformations, in most cases. However, in the example of peptide 1E9T, the α -sheet presents the most stable form overall, and the stability increased in the existence of cationic ions influenced mainly by the ionic charge and radius of the bound ions. Mg²⁺ and Ca²⁺ have the greatest effect on the relative stability followed by K⁺, Na⁺ and Li⁺. MD was performed on full length 1E9T peptide in a range of pH and at 310 and 498 K, in explicit solvent, with and without KCl ions. The results were consistent with literature in which α -sheet structure is a transient state between α -helix and β -sheet formation.

Contents

1	Introduction	1
1.1	Overview	1
1.2	Alzheimer's disease	1
1.2.1	A β cascade hypothesis.....	2
1.2.2	The structural chemistry of A β	4
1.2.3	Metal ions-peptide interaction in AD	10
1.2.4	Treatment based on the metal ion hypothesis	15
1.3	Modelling metal-A β and related biomolecular systems.....	16
1.3.1	Quantum mechanics QM	16
1.3.2	Semiempirical Methods	20
1.3.3	Molecular Mechanics MM	24
1.4	References.....	30
2	Theory	46
2.1	Introduction.....	46
2.2	Quantum Mechanics	47
2.2.1	Schrödinger equation.....	48
2.2.2	Born–Oppenheimer approximation	50
2.2.3	Hartree-Fock (HF) Theory	50
2.2.4	Density Functional Theory	52
2.3	Semiempirical quantum mechanical methods SQM.....	60
2.3.1	GFN2-xTB	61
2.4	Molecular Mechanics (MM)	63
2.4.1	Ligand Field Molecular Mechanics (LFMM).....	65

2.4.2	Molecular Dynamics (MD)	68
2.4.3	Accelerated Molecular Dynamics	69
2.5	Solvent Models.....	70
2.6	References.....	72
3	Method validation on metal-peptide systems	78
3.1	Cu(II)-GHK.....	78
3.1.1	Introduction	78
3.1.2	Computational Method	80
3.1.3	Result and discussion.....	81
3.1.4	Electronic spectra analysis.....	91
3.1.5	Conclusion.....	97
3.2	Cu(II)-Amyloid- β -16	97
3.2.1	Introduction	97
3.2.2	Computational Method	98
3.2.3	Result and discussion.....	100
3.2.4	Conclusion.....	106
3.3	References.....	107
4	Accelerated Molecular Dynamics Simulation of metal ions II (Zn, Fe, and Cu) Binding to N-Terminus of Amyloid- β	110
4.1	General chapter introduction	110
4.2	Computational Methods	111
4.3	Accelerated Molecular Dynamics Simulation of Zn(II) Binding to N- Terminus of Amyloid- β	112
4.3.1	Introduction	112
4.3.2	Results and Discussion	113
4.3.3	Conclusion.....	140

4.4	Zn (II)with capped and uncapped Lys16 comparison	140
4.4.1	Introduction	140
4.4.2	Result and discussion.....	141
4.4.3	Conclusion.....	151
4.5	Accelerated Molecular Dynamics Simulation of Fe(II) Binding to N-Terminus of Amyloid- β	151
4.5.1	Introduction	151
4.5.2	Result and discussion.....	152
4.5.3	Conclusion.....	168
4.6	Accelerated Molecular Dynamics Simulation of Cu(II) Binding to N-Terminus of Amyloid- β	168
4.6.1	Introduction	168
4.6.2	Result and discussion.....	171
4.6.3	Conclusion.....	187
4.7	General Conclusion	187
4.8	References.....	189
5	Alpha sheets structure in A β	195
5.1	Introduction.....	195
5.2	Methodology.....	198
5.2.1	Quantum Mechanics.....	198
5.2.2	Molecular Dynamics Simulation	198
5.3	Results and discussion.....	199
5.3.1	Molecular dynamics simulation of 1E9T peptide	206
5.4	Investigation the potential of α -sheet structure in Zn(II)-A β 16	215
5.5	Conclusion	217
5.6	References.....	219

6	Conclusions and future work	221
7	Appendix	225

List of Figures

Figure 1.1: Schematic of A β 42 that was cleaved from PPA via by β and γ secretase enzymes. ⁴	4
Figure 1.2: Solution phase structure of A β 40 containing a central 3 ₁₀ helical structure of 2LFM pdb. ⁶⁸	6
Figure 1.3: Experimentally determined quaternary structures of A β ₁₋₄₀ A β fibril. Top: each A β ₁₋₄₀ molecule comprises two β -strands. The red and blue colours are present (N-terminal β -strand) and (C-terminal β -strand), respectively. Here the internal contacts are between a red and a blue β -sheet and the external contacts are between two blue β -sheets. Bottom: two peptide possible conformations. F19/L34 and F19/M35 correspond to the internal contacts between residues Phe19/Leu34 and Phe19/Met35. ⁹⁴	8
Figure 1.4: Mode of generic fibril structure characterized by Sunde et al. in 1997. ⁹⁷	9
Figure 1.5: The standard structure geometries of metal complexes.	12
Figure 2.1: The orbital diagram according to LFST theory for octahedral coordination where barycentre, Δ_0 is located between lower and higher energy levels.	66
Figure 2.2: AOM barycentre in situation of a π -donor ligand showing e_π is positive (higher than barycentre, Δ_{oct}).	68
Figure 3.1: GHK structure.....	78
Figure 3.2: N δ and N ε geometries.	81
Figure 3.3: Three cis-GHK low mode conformers.....	82
Figure 3.4: Relative energy comparative between GFN2-xTB and B3LYP-D2/def2-SVP at LFMM geometries on low mode conformers.	83
Figure 3.5: Relative energy of low mode conformers from B3LYP-D2/def2-SVP and GFN2-xTB.....	84

Figure 3.6: Relative energy for low mode conformers obtained by different methods as single point energy calculations at B3LYP-D2/def2-SVP optimised geometry.....	84
Figure 3.7: Single point calculation via DFT (B3LYP-D2/def2-SVP) of seven conformers found by stochastic search at GFN2-xTB optimized geometry	84
Figure 3.8: Comparison of relative energies GFN2-xTB and DFT (B3LYP-D2/def2-SVP) for 7 conformers from stochastic search.	85
Figure 3.9: Hydrogen bonds of lowest cis and trans conformers.....	86
Figure 3.10: RMSD (\AA) over 100 ps MD simulation for cis and trans Cu-GHK.....	87
Figure 3.11: Temperature and Energy conservation for MD.....	87
Figure 3.12: Comparison of starting and frame 1758 structure from MD simulation of Cu-cis, showing stability of Cu coordination and variation in Lys sidechain.....	88
Figure 3.13: Numbering of Cu-ligand atoms (remaining atoms omitted for clarity).	89
Figure 3.14: Cu-Ligand distances values for two trajectories.....	90
Figure 3.15: Input of GHK, H_2O , and Cu separately.	91
Figure 3.16: Most stable conformer CREST conformational search.....	91
Figure 3.17: UV-vis and CD (nm) predicted spectra of Cu(II)-GHK conformers.....	94
Figure 3.18: IR data for five conformers of GHK-Cu(II). The spectral region for wavenumbers from 1000 to 1700 cm^{-1}	96
Figure 3.19: shows the coordination structure of all four arrangements of Asp1, His6, His13 bound to the Cu(II). Note; the geometries show just the active site, with the rest hidden for clarity.....	99
Figure 3.20: The sequence of Small, Middle, and Full models of A β 16-Cu(II). A copper ion was coordinated to one of the nitrogen atom of His 13, another from the imidazole ring of His 6, NH_2 from Asp1, and to the either side-chain or backbone oxygen of Asp1.....	100

Figure 3.21: Relative energy for small size fragments of A β 16 for A, B, C, and D complexes relative to (A) via GFN2-xTB and PBEh-3C optimization and DFT (B3LYP-D2/def2-SVP) single point energy correction.	101
Figure 3.22: An optimized structure D of small-size molecule via DFT (B3LYP-D2/def2-SVP) and GFN2-xTB.	103
Figure 3.23: The full A β -16 peptide, model (A) optimized via GFN2-xTB.	105
Figure 4.1: Different proposed binding modes of A β (1-16) binding to Zn(II).	114
Figure 4.2: Backbone RMSD and R _g values of Zn(II)- Modes and free A β -16. 600 ns of aMD data is reported, the trajectory is combined of the final 200 ns from each of three independent simulations each with different initial velocities.	115
Figure 4.3: RMSF of each residue of free and Zn-bound A β -16. The metal ion itself is denoted residue 18, and the C-terminal amidated cap residue 17.	117
Figure 4.4: Ribbon diagram of the most populated clusters for all binding modes of Zn(II) and free A β	120
Figure 4.5: Secondary structure for Zn(II)-A β and free A β -16.	122
Figure 4.6: Ramachandran maps for Zn(II)-A β and free A β -16. The legends represent the number of dihedrals per bin.	124
Figure 4.7: Salt bridge maps of Zn(II)-A β all modes and free A β -16.	126
Figure 4.8: Contact map of binding modes of Zn(II)-A β and free A β -16. The legend shows the average inter-residue distance in \AA	128
Figure 4.9: Hydrogen bonding occupancy Zn(II)-A β and free A β -16. Plots only display data for hydrogen bonds present for greater than 5% of simulation. Black circles indicate more than one type of hydrogen bond between the relevant residues.	134
Figure 4.10: 1D-Free energy surface of binding modes of Zn(II)-A β and free A β -16 as a function of HB count.	137

Figure 4.11: 1D-Free energy surface of binding modes of Zn(II)-A β and free A β -16 as a function of end-to-end distance (\AA)	138
Figure 4.12: 1D-Free energy surface of binding modes of Zn(II)-A β and free A β -16 as a function of $R_g(\text{\AA})$	139
Figure 4.13: Binding mode used for comparison between capped and uncapped Lys (C-terminus amidated with -NCH ₃ group)	141
Figure 4.14: RMSD and R_g of Zn-A β 16- capped and uncapped Lys (\AA).....	142
Figure 4.15: RMSF of each residue of free and Zn-bound A β -16. The metal ion itself is denoted by residue 18 in case of capped and by residue 17 in uncapped.	143
Figure 4.16: Ribbon diagram of the most populated clusters for Zn(II)-A β capped and uncapped Lys16.	143
Figure 4.17: Secondary structure plots for Zn-A β 16 capped and uncapped Lys, respectively.	144
Figure 4.18: Contact map of Zn(II)-A β capped and uncapped C-terminus, respectively. The legend shows the average inter-residue distance in \AA	145
Figure 4.19: Salt bridge maps of Zn(II) capped and uncapped A β -16(%).	146
Figure 4.20: Hydrogen bonding occupancy Zn(II)-A β . Plots only display data for hydrogen bonds present for greater than 5% of simulation. Black circles indicate more than one type of hydrogen bond between the relevant residues.	149
Figure 4.21: 1D-Free energy surface of amidated and non-amidated Zn(II)-A β as a function of end-to-end distance (\AA).	149
Figure 4.22: 1D-Free energy surface of amidated and non-amidated Zn(II)-A β as a function of $R_g(\text{\AA})$	150
Figure 4.23: 1D-Free energy surface of amidated and non-amidated Zn(II)-A β as a function of HB count.....	151

Figure 4.24: General proposed binding modes of Fe(II) to A β -16	152
Figure 4.25: Different proposed binding modes of A β (1-16) that are available for Fe(II). ...	153
Figure 4.26: Backbone RMSD and R _g values of Fe(II)- Modes and free A β -16. 600 ns of aMD data is reported, made up of the final 200 ns from each of three independent simulations each with different initial velocities.	154
Figure 4.27: RMSF of each residue of free and Fe(II)-bound A β -16. The metal ion itself is denoted residue 18, and the C-terminal amidated cap residue 17.....	156
Figure 4.28: Ribbon diagram of the most populated clusters for Fe(II) and free A β	157
Figure 4.29: Secondary structure for Fe(II)-A β and free A β -16.....	159
Figure 4.30: Ramachandran maps for Fe(II)-A β and free A β -16. The legends represent the number of dihedrals per bin.....	161
Figure 4.31: Salt bridge maps of Fe(II)-A β all modes and free A β -16.....	162
Figure 4.32: Contact map of binding modes of Fe(II)-A β and free A β -16. The legend shows the average inter-residue distance in Å.....	164
Figure 4.33: Hydrogen bonding occupancy Fe(II)-A β and free A β -16. Plots only display data for hydrogen bonds present for greater than 5% of simulation. Black circles indicate more than one type of hydrogen bond between the relevant residues.	165
Figure 4.34: 1D-Free energy surface of binding modes of Fe(II)-A β and free A β -16 as a function of end-to-end distance (Å).	167
Figure 4.35: 1D-Free energy surface of binding modes of Fe(II)-A β and free A β -16 as a function of R _g (Å).....	168
Figure 4.36: General proposed binding modes of Cu(II)	170
Figure 4.37: Different proposed binding modes of A β (1-16) binding to Cu(II).	171

Figure 4.38: Backbone RMSD and R_g values of Cu(II)- Modes and free A β -16. 600 ns of aMD data is reported, made up of the final 200 ns from each of three independent simulations each with different initial velocities.	172
Figure 4.39: RMSF of each residue of free and Cu(II)-bound A β -16. The metal ion itself is denoted residue 18, and the C-terminal amidated cap residue 17.....	174
Figure 4.40: Ribbon diagram of the most populated clusters for Cu(II)-A β -16.....	175
Figure 4.41: Ramachandran maps for Cu(II)-A β . The legends represent the number of dihedrals per bin.	177
Figure 4.42: Salt bridge maps of Cu(II)-A β 16(%).	178
Figure 4.43: Secondary structure for Cu(II)-A β	179
Figure 4.44: Hydrogen bonding occupancy Cu(II)-A β 16. Plots only display data for hydrogen bonds present for greater than 5% of simulation. Black circles indicate the number of interaction between the relevant residues.	181
Figure 4.45: Contact map of binding modes of Cu(II)-A β . The legend shows the average inter-residue distance in Å.....	183
Figure 4.46: 1D-Free energy surface of Cu(II)-A β as a function of end-to-end distance (Å).	184
Figure 4.47: 1D-Free energy surface of Cu(II)-A β as a function of R_g (Å).	185
Figure 4.48: 1D-Free energy surface of Cu(II)-A β as a function of HB count.....	186
Figure 5.1: The α_R denotes right-handed helical region, α_L denotes left-handed helical region on Ramachandran map. The alternating dihedrals in α_R and α_L region form α -sheet. The Phi (ϕ) indicates the C-N-C α -C dihedral angle and Psi (ψ) indicates the N-C α -C-N dihedral angle.	195
Figure 5.2: α -sheet residues in 1E9T peptide that comprises VAL 19, ASP 20, CYS 21, GLY 22 residues. The side chains are hidden for the purposes of clarity.....	202

Figure 5.3: Geometry optimization of α -sheet- Na^+ (top), α -sheet- Li^+ ,(middle) and α -helix - Na^+ (bottom) bound to ions.....	204
Figure 5.4: Geometry optimization of β -strand- Mg^{2+}	204
Figure 5.5: Backbone RMSD and R_g of 1E9T simulation at neutral pH and 310/498 K.	208
Figure 5.6: Dihedral angles evaluation of the region that present proposed α -sheets of 1E9T simulation at neutral pH and 310 K during 50 ns. Note: only frames that show α -strand are presented here for clarity (3 and 5 ns). The dihedral angles of the entire trajectories can be found in the Appendix. The literature reported the α -sheet of this peptide at residues 19,20,21,22 to be VAL19 (negative ϕ and ψ), ASP20 (positive ϕ and ψ), CYS21 (negative ϕ and ψ), and lastly GLY22 (positive ϕ and ψ). ⁶	209
Figure 5.7: Backbone RMSD and R_g of 1E9T simulation at medium pH and 310/498 K.....	210
Figure 5.8: Backbone RMSD and R_g of 1E9T simulation at Low pH and 310/489 K.	211
Figure 5.9: Ramachandran map plots of the average values of phi and psi. The amino acids locate on + ϕ , + ψ region (αL) are pointed as well as the alternate residues at - ϕ , - ψ region (αR) if there any.	214
Figure 5.10: Phi and Psi angles of Zn(II)-A β 16-Mode1 simulation that correspond to the different extracted frames, that showed population at left-handed helix in the Ramachandran map (αL).....	216
Figure 5.11: Zn(II)-A β 16 structure generated from frame number 59999 of aMD simulation. Glu11, Val12, and His13 form nests as they occupied αR , γL and αR regions, respectively. His13 and His13 occupied αR and αL , suggesting it is more turn than sheet. The Zn(II) here coordinated to Asp1, His6, Glu11, and His14. Note that only backbone atoms are shown, while the rest are hidden for clarity.	217

List of Tables

Table 1.1: Reactive Oxygen species examples. ¹³⁰	15
Table 3.1: RMSD between DFT (B3LYP-D2/def2-SVP) and semi-empirical (GFN2-xTB) geometries generated from stochastic search /Å.	85
Table 3.2: RMSD statistical analysis of simulation data (Å).....	88
Table 3.3: Bond distance trajectory analysis of RMSD values (Å).	90
Table 3.4: The relative energies of small-size fragments of Cu(II)-Aβ16 molecule between B3LYP-D2/def2-SVP-Opt/SP, GFN2-xTB-OPT, and PBEh-3C.	102
Table 3.5: RMSD between DFT and GFN2-xTB geometries generated of four small structures (Å).	103
Table 3.6: The Relative energy values comparison of middle-size of Cu(II)-Aβ16 molecule between B3LYP-D2/def2-SVP-SP, GFN2-xTB-OPT, and PBEh-3C.....	104
Table 3.7: The relative energy values comparison of full-size of Cu-Aβ16 molecule between B3LYP-D2/def2-SVP -SP, GFN2-xTB-OPT, and PBEh-3C.....	105
Table 4.1: Different proposed binding modes of Aβ(1-16)binding to Zn(II).....	114
Table 4.2: Backbone RMSD and R _g of free Aβ-16 and Zn(II) binding modes (Å).	116
Table 4.3: Cluster analysis data for equilibrated trajectories.....	118
Table 4.4: Total secondary structure percentages for each binding modes of Zn-Aβ (%)....	122
Table 4.5: HB count for Zn(II)-Aβ and free Aβ-16.....	135
Table 4.6: RMSD and R _g (Å) of capped and uncapped Lys.....	141
Table 4.7: Cluster analysis data for equilibrated trajectories for amidated C-terminus and un-amidated Zn-Aβ16.	144
Table 4.8: Total secondary structure percentages for each system (%).	144

Table 4.9: HB count for Zn(II)-A β for capped and uncapped C-terminus of A β -16.....	148
Table 4.10: Different proposed binding modes of A β (1-16)binding to Fe(II).....	153
Table 4.11: Backbone RMSD and R _g of free A β -16 and Fe(II) binding modes, (\AA)	155
Table 4.12: Cluster analysis data for equilibrated trajectories.....	158
Table 4.13: Total secondary structure percentages for each Fe(II) system (%).	159
Table 4.14: HB count for Fe(II)-A β and free A β -16.	165
Table 4.15: The table of different proposed binding modes of A β (1-16) binding to Cu(II). .	171
Table 4.16: Backbone RMSD and R _g of free A β -16 and Cu(II) binding modes, (\AA).	173
Table 4.17: Cluster analysis data for equilibrated trajectories.....	174
Table 4.18: Total secondary structure percentages for each system (%).	180
Table 4.19: HB count for Cu(II)-A β	181
Table 5.1: peptides that show alternating α_L and α_R amino acids of a nonredundant version of the Protein Data Bank reported by Armen et al. ⁶	200
Table 5.2: DFT relative electronic energy of and α -helix and α -sheet to β -strand for small models obtained from peptides in CPCM water (kJ mol ⁻¹).....	201
Table 5.3: DFT optimised dihedral angles (°) of α - sheet region of 1E9T peptide and ion- peptide.....	204
Table 5.4: Relative energy of α -sheet and α -helix forms relative to β -sheet in CPCM water (kJ.mol ⁻¹).....	205

1 Introduction

1.1 Overview

This thesis aims to present research carried out on modelling of the interactions between metal ions and biochemical molecules such as amyloid- β (A β), Gly-His-Lys(GHK) and other peptides that exhibit α -sheet secondary structure, through the implementation of various quantum mechanical, semi-empirical and molecular mechanic methods. The literature review covered in this chapter focuses on three main research areas related to A β in the framework of Alzheimer's disease: primarily, the role and nature of A β in Alzheimer's disease, the effect of metal ions on its structure and aggregation propensity, and the computational modelling of related metal-biomolecular systems. Furthermore, the GHK peptide has been used as a small, alternative model for testing the GFN2-xTB method. It has been noted that this peptide is a potential metal chelator.¹

Overall, this project involves investigating the metal ions binding to peptides in the brain of neurodegenerative diseases such as Alzheimer's and examining the structural alterations induced from the different binding modes. Moreover, I study how these influence the folding and aggregation propensities of the peptide using a diverse range of quantum and molecular mechanics methods.

1.2 Alzheimer's disease

Alzheimer's disease (AD) is a common neurodegenerative disease found in elderly people, first hypothesised by neuroanatomist Dr Alois Alzheimer in 1906,^{2,3} when he noticed abnormal protein deposits in the brain histology. The disease is clinically diagnosed as a gradual degeneration of cognitive functions associated with age. The symptoms shown by AD patients include memory disturbance, physical function deficit, and cognitive decline.⁴ The number of cases suffering from the disease is large worldwide and still dramatically increasing, with approximately more than 1.5 million growth in cases each year.⁵ The pathogenesis of AD is very complicated, comprising of a range of molecular, cellular and

biological processes. The causes and progress of AD are not clearly understood,⁶ but influential factors have been reported to involve numerous lifestyle attributes such as smoking,⁷ alcohol consumption,⁸ low physical and brain exercises,⁹ diet,¹⁰ diabetes, as well as family history of developing AD,⁹ but the most significant risk factor is age.¹¹ Unfortunately, there is no effective treatment to cure AD yet, and the only pharmaceutical therapy is to control the symptoms but not delay the disease.¹²

The abnormal protein aggregates in plaques and tangles in the cerebral cortex region of the brain.¹³ The plaques contain insoluble amyloid-beta (A β) proteins, cleaved from the amyloid precursor protein (APP). APP is a transmembrane protein that has an important function in the human body, used in transition metal ion regulation.¹⁴ The formation of toxic A β and its extracellular build-up causes cell death, as described in the Amyloid hypothesis.^{15,16} There is evidence that metals such as iron (Fe),¹⁷ zinc (Zn),¹⁸ copper (Cu),^{19,20} and calcium (Ca)²¹ accelerate the precipitation of A β and release of reactive oxygen species (ROS)^{22,23} and are found in concentrated amounts in the brain, suggesting an interruption in the metal-regulation operations in affected brains.^{24,25} Normally, metal ions play a crucial role in many of the essential biological operations in the body, such as the central and parasympathetic nervous systems. However, homeostasis of trace metals is vital; otherwise, they can be toxic, leading to oxidative stress that harms lipids, nuclear DNA and cells.^{16, 19, 26–31} Some studies suggested that Zn and Cu chelation in the brain tissue can help with fibre solubilization in some samples from Alzheimer's carriers.^{32–34} To find an effective treatment for AD, first an understanding of the mechanisms of plaque and tangle aggregation in the brain needs to be developed, aiding in an understanding of how the brain function is affected. The amyloid cascade hypothesis aims to unveil the main component of plaques in the brain associated with AD.

1.2.1 A β cascade hypothesis

Numerous studies have proposed that the progress of Alzheimer's disease is related to the formation and accumulation of abnormal structures in the brain, formed in senile plaques and neurofibrillary tangles.³⁵ Clinically, AD patients' brain tissue shows a gradual loss of cortical neurons that play an important functional and structural role in the normal brain.³⁶ This

neuronal death is related to irregular deposits of proteins inside and outside the cerebral cortex neurons. The extracellular deposits consist of insoluble A β protein in plaques form, whereas the intracellular deposits involve the accumulation of hyper-phosphorylated protein tau in neurofibrillary tangles (NFT), playing a significant role in intracellular transport.¹³ Masters et al. in 1985, identified the A β peptide as the critical component of the aggregated plaques observed in the brain tissue of Alzheimer's patients,³⁷ which signified the start of the current period research into AD.

In this section, the most general hypothesis of AD evolution is outlined, and the chemistry and structure of the related A β are reviewed. The amyloid hypothesis is classified as one of the most prominent hypotheses that explain the causes of AD.^{38,39} A β substance accumulates within the brain in the form of plaques and insoluble fibrils, leading to eventual neuron death.⁴⁰ Later, the metal ion hypothesis was paired with this hypothesis by presenting metals' critical role in enhancing plaque deposition and increasing the production of toxic ROS.³⁴

The formation of A β starts from fragments cut from the amyloid precursor protein (APP) by the action of β and γ secretase enzymes. These fragments are usually between 39-43 residues. A β 40 and A β 42 are the dominant species formed in the brain, exemplified in Figure 1.1. The soluble A β 40 is much more abundant than A β 42, even though A β 42 displays a higher tendency for aggregation due to the hydrophobicity of its two terminal residues (Ile and Ala).^{41,42} Certainly, A β 42 is the primary component of amyloid plaques and has been demonstrated to be neurotoxic.⁴³⁻⁴⁶

The A β monomer include both hydrophilic and hydrophobic parts (amphiphilic); a hydrophilic part at the N-terminal region and a hydrophobic part at the C-terminal region.³⁵ The hydrophilic amino acids of the A β sequence contains six negatively charged residues: Asp1, Glu3, Asp7, Glu11, Glu22 and Asp23 and three positively charged residues: Arg5, Lys16 and Lys28; resulting in a net charge of -3 at biological pH. The three histidine residues within the A β monomer (His6/13/14) were reported as neutral at biological pH.⁴⁷ A β monomers form many groups divided into three types: insoluble amyloid fibrils, protofibrils and soluble amyloid oligomers. Pharmaceutically, the drugs that exist to treat AD are based mainly on the prevention of A β plaque formation by inhibiting the production of A β aggregation or by producing molecules that interact with A β species.

The amyloid hypothesis,³⁹ which states β -amyloid protein is the major component of the amyloid plaques, proposed that AD initiates when A β overload results to cell death. This discovery suggested that a better understanding of amyloid cascade could increase the viability of drug design.³⁹ Subsequent studies revealed that the levels of A β found in AD are indeed toxic to neurons.^{48–50} A different study,⁵¹ claimed that drugs based on this hypothesis have had deficient results, despite the evidence supporting the Amyloid Hypothesis, and called for the consideration of alternative ways to treat AD other than controlling β -amyloid itself.

The rest of the disease progression, includes the formation of insoluble neurofibrillary tangles, containing hyper-phosphorylated tau protein, in its abnormal form. The causes of neurofibrillary tangles, have been revealed to be a result of an imbalance of A β generation and its incomplete discharge from the brain.^{52,53}



Figure 1.1: Schematic of A β 42 that was cleaved from PPA via by β and γ secretase enzymes.

1.2.2 The structural chemistry of A β

Even though the amyloid cascade hypothesis has been providing a general foundation to explain the disease's progression, some of its observations are weak and do not suit the hypothesis,^{16,43} resulting in the cascade term of the hypothesis to not be frequently used and instead replaced with the amyloid hypothesis. The number of amyloid deposits demonstrates a weak correlation with the level of cognitive decline,^{51,54} as the patients who show no Alzheimer's symptoms illustrate significant deposits of A β in the brain,^{55–57} while the soluble oligomeric forms of A β demonstrate significantly better correlation of cognitive impairment in the patients.^{58,59} Remarkably, it has been indicated that soluble oligomeric forms of A β are more toxic to nerve cells than the fibrillated plaques.^{53,59,60} Thus, these oligomeric structures

of A β are progressively considered to be the main toxic species of AD.^{61–64} So, an alternative hypothesis is that amyloid oligomers rather than plaques are accountable for the development of the disease. Another reason that can weaken the amyloid cascade hypothesis is that AD originates in a specific part of the brain, specifically in the hippocampal neurons; however, A β is detected throughout the brain where high concentrations of transition metal ions are located, suggesting that this may be another disease origination, specific to these regions.^{25,65}

Even though the research work and progress in understanding the A β structure of different forms of the species aided in the understanding of its role in AD, as reviewed in numerous papers,^{5, 35, 42, 66, 67} the physical properties of A β protein are still inadequately identified. A β in the monomeric phase is a naturally disordered peptide that adopts a diverse, random coil structure in an aqueous solution.^{68–70} The structural deviation of A β presents challenges and increased complexity when studying A β . Therefore, there is little solution-phase NMR structural data, and neither A β 40 or A β 42 monomers form to give X-ray crystallographic data. However, the fibril structures of the monomer are responsive to solid-state NMR. From these, significant characteristics of the solid-state monomer form have been identified, by Tycko and co-workers,^{71,72} in addition to a multiple quantum solid-state NMR study.⁷³ However, the peptide's secondary structure is also influenced by the chemical conditions, which depends on solvent and pH. Several findings have been reported, through the study of the peptide in water, with the A β monomer shown to comprise of about 5–20% helical form and around 0–25% β -sheet content, with the remaining percentage of the peptide, mostly consisting of random coil, as reported by NMR and circular dichroism (CD) spectroscopic data, displayed in PDB ensembles 1AML,⁶⁹ 2LFM,⁶⁸ 1BA4,⁷⁴ and others.¹⁹ The central hydrophobic portion of the peptide is most likely to form an extended 3₁₀ helix-type structure, as shown in Figure 1.2. The interpretation of the chemistry correlated to A β , and the detection of related monomer conformations by experimental analysis and computational studies, is an essential and active field towards the understanding of the oligomerisation and aggregation progression, which may lead to the provision of better treatment, or even prevention the AD.^{42,43}

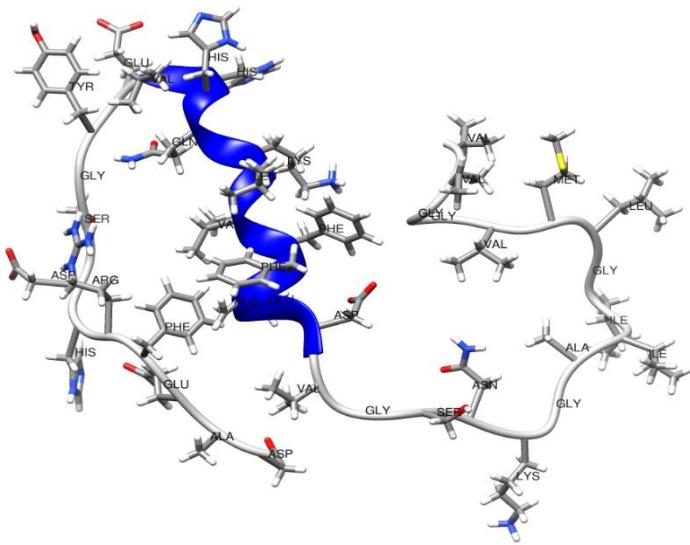


Figure 1.2: Solution phase structure of A β 40 containing a central 3₁₀ helical structure of 2LFM pdb.⁶⁸

The soluble A β monomers undergo structural modifications leading to their oligomerization or fibrillization. Numerous oligomeric ensembles have been defined,^{19,75} although no high-resolution atomic structure has been stated for an intermediate form so far.⁷⁶ Nevertheless, biophysical studies of fibrillar A β oligomers indicate that they have the same type of structure as mature amyloid fibrils.⁷⁷

It has been proven that the A β monomers can be modified to form different soluble oligomers, which may eventually promote the formation of fibrillar aggregates.⁷⁸ Therefore, these oligomer structures have been increasingly the subject of intense interest and study, as they are believed to be the leading causes of the disease.^{61–63,79} A β trimers have been demonstrated to be particularly harmful to brain cells, more so than other sizes of oligomers.⁸⁰ The study of soluble A β protein oligomers, extracted directly from the cerebral cortex of Alzheimer's disease brain, shows A β dimers actively damaging synaptic structure and functions, such as memory.⁸¹ It has been concluded that a monomer secondary structure displays mostly coil, minor portions of the helix, and very few of β -strand. However, the amount of β -strand appears to increase over time alongside the degree of the increasing oligomers.^{82–84} Conversely, the smallest oligomers have been found to preserve a great degree of structural disorder, as proven by NMR experiments which identified small oligomers displaying mainly turn and coil characteristics structure.⁸⁵ The analysis of A β 3D structure of

the fibrils consist of A β (1–42) oligomers, showing residues 18–42 to retain significant β -sheet arrangement, thus suggesting the increased toxicity, as a result of the oligomers, which may play an important role in AD.⁸⁶ In general, it seems that the monomer is involved in several alterations in conformation, including the formation of a hairpin structure near Gly25, surrounded by large β -sheets, allowing for oligomerisation and aggregation to occur.⁸⁷

During the oligomerisation process, the intermediate oligomers are then converted into amyloid fibrils which are the last step of the aggregation process in the peptide's pathway. These fibrils overload to resemble the plaques observed in the brains of AD patients. The structure of these fibrils and the correlation of their aggregation with AD pathogens, are generally studied by fluorescence using thioflavin T (ThT) assays.⁸⁸ Once fibrillation starts (elongation phase), a sharp increase in fluorescence intensity is observed which then saturates (saturation phase), when the maximum length and fully fibrillated structures are achieved. ThT mostly responds to β -sheet regions of A β fibrils, thus in turn showing an increased signal representing a greater degree of β -sheets, giving evidence of fibril formation. The fibril structures were found to consist of both of the parallel β -sheet⁸⁹ or only anti-parallel type.⁹⁰ Therefore, understanding the structure of these A β fibrils is essential towards the evaluation of their contribution in AD.

A solid-state nuclear magnetic resonance (NMR) experimental study by Tycko et al.,⁹¹ shown in Figure 1.3, proved that fibrils of A β 40 comprise of dimeric molecules packed into the four-layered β -sheets structure, where residues from neighbouring chains are interacting; fibrils of A β 42 adopt a similar structure, except the residues are not aligned.³⁹ Each A β 40 molecule has two β -strands regions seen at residues from His13 to Asp23 and from Lys28 to Val40, which are divided by a hairpin turn and are stabilised by contact between Asp23 and Lys28 residues. The study detected internal and external quaternary contacts of parallel β -sheets in the fibrils, with the internals occurring within a single molecular layer (between side chains of Leu17 and Phe19 and Ile32, Leu34, and Val36) and the external between molecular layers (side chain of Met35 and the peptide backbone at Gly33).⁹¹ A different study determined a 3D structure of A β 42 fibrils by solid-state NMR,⁸⁶ exhibiting parallel β -sheets consisting of residues Val18–Val36. A pair of monomers are essential to generate the repeat unit of the fibril. This is in agreement with another work,^{92,93} which showed that the middle region of

$\text{A}\beta_{42}$ involves mainly β -sheet content structure, but the particular residues contained is varied.

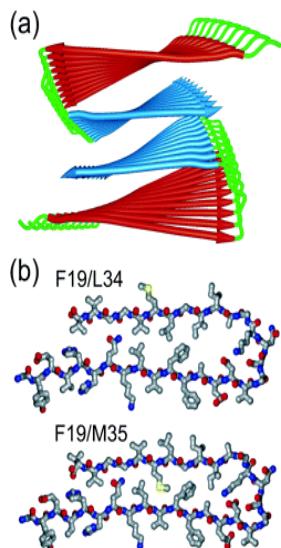


Figure 1.3: Experimentally determined quaternary structures of $\text{A}\beta_{1-40}$ $\text{A}\beta$ fibril. Top: each $\text{A}\beta_{1-40}$ molecule comprises two β -strands. The red and blue colours are present (N-terminal β -strand) and (C-terminal β -strand), respectively. Here the internal contacts are between a red and a blue β -sheet and the external contacts are between two blue β -sheets. Bottom: two peptide possible conformations. F19/L34 and F19/M35 correspond to the internal contacts between residues Phe19/Leu34 and Phe19/Met35.⁹⁴

Further study on the synthetic amyloid fibrils structure using X-ray diffraction and electron microscopy by Kirschner et al.,⁹⁵ described $\text{A}\beta_{1-28}$ fibrils as a tubular cylinder with a diameter of 86 Å, and the fibril walls comprising of a tubular cylinder of cross- β -pleated sheets with the peptide chains organised almost perpendicular to the fibril axis. Also, high β -turn potentials are concentrated at residues Ser8 and Ser26, and the intervening sequence appears at a high potential for either β -conformation or α -helix. In contrast, the N-terminal region near the first β -turn is hydrophilic and displays potential α -helical structures. Another work using X-ray diffraction by Inouye et al.⁹⁶ classified the amyloid fibril formation of multiple hollow tubes with five or six β -crystallite units. An additional study performed by Sunde et al.,⁹⁷ using high-resolution X-ray fibre diffraction detected the structure of $\text{A}\beta$ fibril to be forming a stable helical set of β -sheets running parallel to the axis, with the strands perpendicular to the fibril axis, as seen in Figure 1.4. Another study by Daggett et al.⁹⁸ detected an uncommon type of secondary structure form, termed α -sheet, in a variety of proteins that are involved in

amyloidosis. Theoretical and experimental studies suggest that the α -sheet structure in A β drives aggregation and toxicity in AD.⁹⁹ The α -sheet conformation was originally noticed by Pauling and Corey.¹⁰⁰ The α -sheet consists of around four alternating α_L and α_R amino acids(α_L , α_R , α_L , α_R). This study claimed that this type of structure could induce peptide aggregation. This will be covered in more detail in Chapter 5 of this thesis.

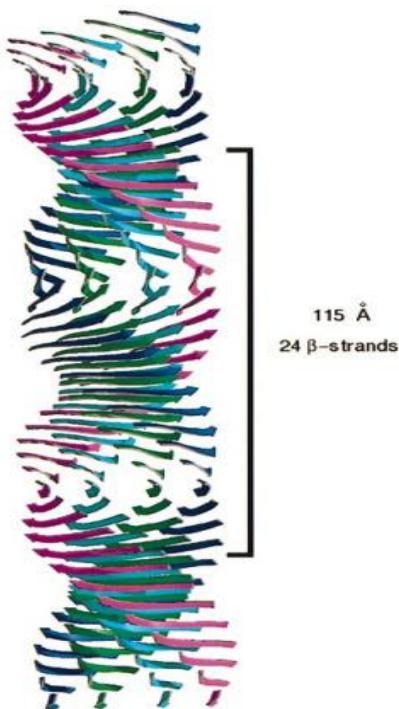


Figure 1.4: Mode of generic fibril structure characterized by Sunde et al. in 1997.⁹⁷

Even though the A β hypothesis provides an explanation for many of AD aetiology, there is definitely a weakness to this theory. All of these have been identified when all produced drugs and strategies that are based on the A β hypothesis, proved ineffective or even yielded opposing cognitive side effects in clinical trials, despite being able to decrease A β load in preclinical tests.^{101–103} That because A β is obtained in several parts of the brain, including the healthy one.⁵ AD starts in certain parts of the brain that are rich in metal ions. Therefore, there are other causes or factors for AD that could be located at this part of the brain that need to be discussed. As the age trigger of AD is probably a combination of oxidative stress and metal ions overload, which become more significant with age,¹¹ the metal ions have been linked to the A β hypothesis.¹⁹

1.2.3 Metal ions-peptide interaction in AD

1.2.3.1 Natural metals

Transition metals such as Fe, Zn, and Cu have a significant function in biological systems, particularly within proteins and enzymes. However, those trace metals exist in a regulated state naturally in the body, and any deficiency by APP regulator causes an imbalance in the concentration, potentially leading to destructive biological results. Regardless, transition metals such as copper and iron have the unique chemical property of being redox-active metals with ability to fill multiple oxidation states.¹⁰⁴ This behaviour of metals has been confirmed chemically, along with its harmful nature when metals coordinate to A β , leading to the production of toxic ROS.^{22,23,31,40}

There is evidence of the central role of naturally occurring transition metal ions (Cu, Zn, Fe) in Alzheimer's.^{17,105–108} There is a correlation between the imbalance of the levels of metals in the brain and nerve-cell damage.^{109,110} In their normal concentration, they have beneficial roles in neuron signalling, apoptosis, inflammation, oxidative stress regulation and cell proliferation.^{111,112} The first observable contribution of metal ions to AD was in 1975 in brains poisoned with lead, where it was found at tangle and senile plaques.^{113,114} Zinc was first associated with AD by Burnet in 1981.¹¹⁵ Also, it has been reported that aluminium, manganese and iron might play a role in inducing the tangle pathology.^{19,116} In this chapter, the focus is on the transition metal ions, zinc, copper, and iron, due to their valuable natural functions in the brain, interactions with A β , and the implication of metal ions in the amyloid hypotheses of AD.

1.2.3.2 Cu(II), Zn(II) and Fe(II) Coordination Chemistry

The transition metals are classified as the d block elements of the periodic table (groups 3-11) with ground state configuration $ns^{0-2}(1-n)d^{0-10}$, and partially filled orbital electronic configuration. They have physical and chemical properties that give them several features, such as multiple stable oxidation states, high melting and boiling points, variety of coloured compounds resulting from electronic transitions, flexibility, electrical conductivity after

complexation with ligands and catalytic capabilities. In transition metals' electronic configuration, the 4s sub-shell level is lower in energy than the 3d sub-shell, so the 4s valence shell will be filled first with electrons before 3d sub-shell and they will be removed from the 4s shell before the 3d shell when transition metals are ionized. Once transition metal ions are formed, these metal ions can couple to ligands to form coordinate bonds to a molecule or ion that has a lone pair of electrons in its valence orbitals, to form a metal ion complex. The ligands act as a Lewis base such as NH_3 , CO, Pyridine, Phosphine, N_2 , O_2 , OH^- , Cl^- , and $\text{C}_2\text{O}_4^{2-}$, donating electrons to transition metals, which function as a Lewis acid (acid-base interaction). The coordination numbers of metal complexes are accounted based on the number of lone pair electrons from one atom of the ligands bonded to the metals, in a monodentate, bidentate, tridentate or polydentate fashion. In addition, metal complexes can form several common structure geometries, including linear, tetrahedral, square planar, octahedral and trigonal bipyramidal structures,¹⁰⁴ Figure 1.5.

The applications of transition metal elements are essential in biochemical processes, such as Fe in haemoglobin and myoglobin, Zn, Cu and Mn and other metals as catalytic centres in enzymes. However, their oxidation state may be varied, making them redox-active, such that they exhibit toxic effects that can result to loss of function, and consequently diseases such as cancer and Alzheimer's.

N. of ligands	Geometry	Hybridization	
2	Linear	sp	—M—
4	Square planar	dsp ²	
4	Tetrahedral	sp ³	
5	Trigonal bipyramidal	dsp ³	
6	Octahedral	d ² sp ³	

Figure 1.5:The standard structure geometries of metal complexes.

1.2.3.3 Cu(II)-, Zn(II)- and Fe(II)-A β Coordination Chemistry

The structural studies of Cu(II)-A β and Zn(II)-A β described both of them as average hard–soft Lewis acids, and they have affinity toward N, S, and O ligands. Zinc contains a filled d shell where it always occurs in the symmetric closed-shell (d^{10}) state; therefore, it is redox-inactive but instead performs as a Lewis acid, able to accept a pair of electrons. On the other hand, iron and copper are redox-active as Fe(II/III), and Cu(I/II). Their redox capabilities combined with their functions are essential in biological molecules, as transition metals bound to proteins (metalloproteins) serve as reaction catalysts. Also, the metals are capable of binding to the active sites of many proteins, aiding in oxygen transport, and making them oxygen-activating enzymes, and electron-transfer proteins (biochemical reductants or oxidants).^{117,118}

Cu(II) is an open-shell ion as its electron configuration contains an unpaired electron (d^9). The coordination of the Cu(II) is pH-dependent and differs depending on the type of peptide fragment. A study using NMR and electron paramagnetic resonance (EPR) spectroscopy, detected that Cu(II) usually binds with A β in a 1:1 stoichiometry, at low pH (termed type I), with the second, low-affinity binding site, dominant at high pH (termed type II). Type I

coordination involves two or more binding modes in equilibrium, with the copper(II) bound to three N-donor ligands, such as His residues or the N-terminal amide, as well as at least one oxygen-donor ligand, such as the carbonyl oxygen of Ala2 or Asp1 sidechain carboxylate; together forming a geometry of four or five-coordinate complex.^{19,35} The difference between type I and II is the NH deprotonation nitrogen atom, from the Ala2 backbone, upon pH increase.¹¹⁹ All the different binding modes, tested in the present work for Cu(II)-A β , are presented in detail in chapter 4.

Zn(II) is a closed shell ion where d orbitals are full with 10 electrons. It also interacts with A β in a 1:1 stoichiometry. An NMR study has proposed that the Zn(II) can be bound to A β through the three N-terminal histidine amino acids (His6, His13 and His14), in addition to Glu11, giving a distorted tetrahedral geometry.¹²⁰ Residues Tyr10, Asp1, N-terminal amine and the deprotonated amide of the Arg5 backbone have also been identified as potential binding sites, depending on the reaction conditions. Also, it has been shown that Zn(II) binds to His residues from neighbouring A β peptides, which can induce aggregation through formation of cross-linkages.¹²¹

Even though the Fe (II/III) redox couples are common in biological systems, few studies are available about iron-A β binding modes, unlike copper and zinc, possibly due to the instability of the oxidation state between Fe(II) and Fe(III). The most proposed A β binding residues with Fe(II) are Asp1, Glu3, His6, His13 and His14.¹²²

1.2.3.4 The effects of Cu(II), Zn(II) and Fe(II) on A β Aggregation

The binding of metal ions to A β changes the peptide structure and its aggregation propensity.^{123,124} The metal induction effects are distinct and seemingly varying. There are three aggregation structure types: extended β -sheet fibrils, soluble oligomers, and large amorphous non-fibrillar aggregates. Equimolar amounts of Cu(II) that participate in binding has been indicated to accelerate fibril formation in dilute A β solutions.^{125,126} On the other hand, a larger ratio of Cu(II) to A β induces the production of amorphous aggregates. These two classes of aggregates have been reported to be nontoxic.^{124,127} However, the further fibrillar aggregates formed by Cu(II) are toxic.¹²⁶ The secondary structure of the A β monomer, converts from coil to β -sheet structure upon Cu(II) binding and is associated with the

formation of A β dimers, studied via circular dichroism.¹⁹ Compared to Cu(II), the Zn(II) binding to A β monomer has mostly been found to be nontoxic, as it is reported to increase the helix structure.¹²⁸ However, it has also been shown to disrupt the helix structure, upon binding to A β monomer, resulting in its aggregation changing the α -helix to β -sheet transition during the aggregation process.¹⁹

1.2.3.5 Oxidative stress and Alzheimer's disease

In a biological system, ROS are considered to be the most poisonous to macromolecules involving lipids, nucleic acids and proteins due to the free radical (two unpaired electrons in valance orbitals of oxygen), such as the superoxide anion, hydrogen peroxide and hydroxyl radical, Table 1.1. On the other hand, there is evidence that shows ROS have a beneficial side in the human body, by eliminating invading pathogens and enhancing the rate of DNA replication and cell proliferation.¹²⁹ The toxic effects of these highly reactive species on cells are more serious when reacting with lipids, nucleic acids and proteins. Once again, there is a relation between the aggregation of metals, such as iron, zinc and copper in A β and neurodegenerative disorders.³¹ Zinc and copper ions have high-affinity binding sites to the N-terminal metal-binding domains of A β , located in APP. Copper is responsible for the highly reactive hydroxyl radical (OH•) formation and contributes to the increased oxidative stress, thereby could eventually cause AD.³⁴ Moreover, it has been shown that amyloid plaques contain a high concentration of copper.^{26,107}

Table 1.1: Reactive Oxygen species examples.¹³⁰

Name	Formula	Lewis Structure
Hydrogen peroxide	O ₂ H ₂	H : 
Hydroxyl anion	OH ⁻	: 
Hydroxyl radical	•OH	• 
Peroxide	•O ₂ ²⁻	• 
Superoxide anion	•O ₂ ⁻	• 
Oxygen	O ₂	

1.2.4 Treatment based on the metal ion hypothesis

With a significant amount of the advancement of research interest into interpreting the nature of the interactions between naturally occurring metals and A β , a number of new therapeutic strategies to AD have been proposed.^{25,131,132} One is to target the metal imbalance through a metal chelating strategy, without any concern for A β .^{133,134} However, the chelator drug needs to be directed specifically to its target, which is a challenging task. A second strategy is metal-A β targeting compounds.¹³⁵ These compounds can inhibit A β -metal interactions by selectively occupying the metal ions binding site on A β (three His residues near the N-terminus in particular) which block metal coordination. A third one is based on controlling the toxicity causes, particularly by using anti-oxidant or anti-inflammatory compounds.^{136,137} Full discussion of the transition metal-based approaches is beyond the scope of this thesis, but interested readers are directed to recent reviews on the topic^{6,138} and the references within. Today, due to the massive works of researchers and pharmaceutical companies, peptide-based drugs have provided promising results in treating Alzheimer's disease.

Consequently, the last decade has observed great scientific and industrial interest in their therapeutic uses.¹³⁹ Currently, several natural and synthetic therapeutic peptides are undertaking clinical trials.¹⁴⁰ Generally, peptide-based drugs show numerous advantages over small molecule therapeutics, mainly in terms of their efficacies and fewer side effects.¹⁴¹ An attractive therapeutic approach proved that the multifunctional peptides are able to modify

A β pathologies by inhibiting A β aggregation and free metal ions.¹⁴² Neutral GHK peptide has been proposed as a potential therapeutic approach for AD by chelating the peptide to the copper due to its high affinity toward the Cu(II); this binding can lead to an interruption in the production of ROS.^{143,144} This will be discussed in detail in Chapter 3.

This thesis will focus on the research that forms the foundation of this work – the metal ion effects on the structure of A β , and in particular upon their binding in the context of AD, including A β secondary structure and using human peptide proposed chelator: Cu(II)-GHK as a model for A β computational study.

1.3 Modelling metal-A β and related biomolecular systems

The highly flexible nature, rapid aggregation, as well as solvent and paramagnetic effects of A β from experimental studies, make challenging the determination and characterization of transition metal ions bound to A β monomers and oligomers. Additionally, the impact of transition metal ion binding on the monomeric or oligomeric amyloid- β structures and dynamics is still yet to be identified. Therefore, a lot of effort has been focused on the explanation of the structure and chemistry of A β using molecular modelling approaches, to provide additional knowledge to experimental outcomes at the electronic, atomistic, and molecular levels. Density functional theory (DFT) calculations, ab initio molecular dynamics (MD) simulations, as well as hybrid molecular mechanics (QM/MM) simulations, and conventional MD simulations, as well as the inclusion of enhanced sampling, have been performed to study transition metal ions bound A β peptide monomers or oligomers in solution. There is a recent review published in 2019, that covers most of the computational methods used for modelling transition metal ions interaction with A β .¹⁴⁵ But in this section a selection of the work on modelling the A β peptide and related biomolecular systems containing metal ions, used in this study, will be reviewed.

1.3.1 Quantum mechanics QM

Even though quantum mechanical calculations are not commonly employed to large biomolecular systems that contain metals due to their computational demands, there are great many applications where they can be used. As an alternative way to reduce the cost,

these methods are used for modelling the region near the metal while a less computationally demanding method used for the remaining part of the system, or neglected completely. QM methods have shown excellent results in modelling small size biomolecular systems and more structured ones. Still, calculating one conformation of the large and flexible system via QM is insufficient to describe the behaviour of the peptide. A discussion of the recent advancements in the topic is supplied in references.^{145–147} Instead, only the related researches are considered here.

DFT is a method employed in biology, chemistry, and physics for exploring the electronic structure of many-body systems like atoms and molecules. Commonly, hybrid DFT, and in particular Becke's three-parameter exchange with Lee-Yang–Parr correlation (B3LYP)¹⁴⁸ computations have been used to analyse metal and A β interactions. The study by Streltsov et al.¹⁴⁹ utilized the DFT method with the B3LYP functional and the LANL2DZ basis set, for the optimization of structures involving Cu(II) bound A β . The coordination systems of the two models examined, contain three imidazole rings and a carboxylate side chain (Asp, Glu or Tyr) at equatorial positions and one or two water molecules at axial positions bound to the Cu(II) ion. They determined a Jahn–Teller type distorted octahedral coordination around Cu(II) and provided the bond lengths and bond angles for the optimized geometries. The extended X-ray absorption fine structure (EXAFS) spectroscopy refinement applied in the same study, detected distorted six-coordination mode (3N3O) consisting of three histidine rings, Glu and/or Asp acid, and axial water but not Tyr.

The following DFT study by Marino et al.¹⁵⁰ was performed on truncated model systems that involved 4-methylimidazole for His, phenol for Tyr, and acetate for Glu or Asp coordinated to Cu(II) and Zn(II). The calculation method used was B3LYP with 6-31G(d) and 6-31+G(d) basis sets for C, O, N, and H atoms, while Stuttgart energy-averaged effective core pseudopotentials (ECP)SDD were used for the metal centre. The investigation determines the possible coordination of the three His residues, Asp1, Tyr10, or Glu11, with the metal centre. The obtained results, show zinc preferring four- to five-coordinated atoms, and copper favouring five coordinated geometries. The coordination involving oxygen donor ligands from any of Asp or Glu and water, displayed a higher affinity than Tyr for binding to the metal ions.

Ali-Torres et al.¹⁵¹ aimed to characterize the three-dimensional coordination of Cu(II) to A β -16 via combined homology modelling (HM) techniques alongside quantum mechanics-based

approaches. The initial geometry was taken from Zn(II): A β (PDB ID: 1ZE9)¹²⁰ but substituted the zinc ion with copper. The coordination was inspected by considering both ϵ and δ N of the histidine rings (6, 13, and 14) as well as the oxygen from Asp1, Glu3, Asp7, Glu11, and CO-Ala2 as the fourth coordinate. The DFT calculation using hybrid BHandHLYP function with 6-31G++G(d,p) basis set, yields coordination involved [$O^{Asp7}, N_{\epsilon}^{His6}, N_{\delta}^{His13}, N_{\delta}^{His14}$] whereas the coordination under implicit solvent effect shows [$O_c^{Ala2}, N_{\epsilon}^{His6}, N_{\delta}^{His13}, N_{\epsilon}^{His14}$]. However, using hybrid QM/MM calculations with the BHandHLYP function: UFF calculations yield [$O^{Glu3}, N_{\delta}^{His6}, N_{\epsilon}^{His13}, N_{\epsilon}^{His14}$]. Though each method proposes a different copper coordination sphere, results show preference for oxygen ligands (Glu3, Asp7) rather than backbone nitrogen atoms.

DFT and second-order Moller-Plesset (MP2) study by Azimi and Rauk,¹⁵² where B3LYP/6-31+G(d) and MP2/6-311+(2df,2p), in implicit solvent (water) at physiological pH, were used to evaluate the binding of Cu(I/II) to N-terminus of A β , in which the coordination involved Asp1, Ala2, His6, and His13/His14. The calculation results suggest Cu(II) binds Asp1, His6 and either His13 or His14, producing the lower pH (component I). In contrast, Component II of A β and Cu(II) coordination is His6, His13, and His14, and backbone oxygen of Ala2, but Asp1 does not bind to Cu(II) when all three His residues are attached. The calculated binding affinities of Cu(II) and Cu(I) to the His13 or His14, are in remarkable agreement with experimental values, which provides confidence in the calculated free energy. The free energy calculations show that Cu(I) binds to the predicted sites in A β , more strongly than Cu(II). This study also shows the computed reduction potential for Cu(II) is in line with experimental data.

Another DFT study by Dudev and Lim,¹⁵³ looked at the binding affinity and selectivity of Zn(II), Mg(II) and Ca(II) ions toward atypical amino acid residues. The geometry optimization was performed using S-VWN functional with SDD/6-31+G* basis set on metal and first coordination sphere and then free energy calculation done by the B3LYP/ 6-311++G(2df,2p) level. The results show that the nonstandard residues have more metal binding capacity than their standard equivalents. Also, the residues exhibit greater selectivity for Zn(II) over the other ions, indicating that they could be used as metal-binding sites in protein production.

A further study by the same authors¹⁵⁴ aimed to understand the factors governing the binding competition between the small inorganic anion ligands in cytoplasmic and metal cations such as Na⁺, K⁺, Mg²⁺, and Ca²⁺, using DFT and continuum dielectric methods. The result of

optimisation by S-VWN functional coupled with the 6-31+G* basis set and free energy calculations by B3LYP functional and 6-311++G(2df,2p) basis set, found the metal cation binding to its target protein, regardless of the high concentration of cytoplasmic anions, for several reasons. One is because of desolvation of a carboxylate from Asp/Glu amino acids in a protein is less energetically expensive than the desolvation of an inorganic anion in aqueous conditions. Another, comes from the protein acting as a polydentate ligand, giving a chelate effect. A similar study was carried out by the same authors, using B3LYP/6-31+G(3d,p) calculation to analyse the factors affecting the selectivity of monovalent Li⁺ to displace divalent Mg²⁺ in proteins, such as net charge and solvent exposure of the metal-binding site.¹⁵⁵

A recent study by Coskuner-Weber,¹⁵⁶ using B3LYP functional with different basis sets; 6-31G*, 6-31G**, cc-pVDZ, ahrlrichs-vdz, ahrlrichs-vtz, def2-svpd, def2-tzvp, and lanl2dz-ecp, investigated the impact of the basis sets on Cu(II)-A β . The study also looked at the structures and thermodynamic properties of Cu(II), bound to: three His rings and Asp, three His rings and Tyr, three His rings and Glu, and three His rings and a water molecule (H₂O). The overall results show the calculated binding energies of the Asp, Glu, Tyr and H₂O bound to Cu(II), with the three Histidine rings being the preferred coordination and the order of binding energies from the most to least negative: Asp > Glu > Tyr > H₂O. The same trend was observed for Asp, Glu, and Tyr, by Mantri et al.¹⁵⁷

Sodupe, Rauk, and co-workers,¹⁵⁸ applied ab initio computational methods at the MP2/6-311+G(2df,2p) and DFT methods at B3LYP/6-31+G(d) levels, on truncated models of the Fe(II/III) binding to A β . Also, MP2 calculations utilized a continuum model for water for calculating the free energy of solvation for different coordination modes. The calculated results indicated the most stable complexes that contain His-His fragment and imidazole for both oxidation states are [Fe(O-His-His)(Imidazole)⁻(H₂O)₃]^{+2/+3} with a hexacoordinated structure. In contrast, the most stable complexes that have His-His and phenolate group of Tyr10 are the Penta-coordinated [Fe⁺²(O-His-His)(PhO)⁻(H₂O)]⁺ and [Fe⁺³(N-His-His)(PhO)⁻(H₂O)]⁺. The study also concluded that coordination of Fe(II)/(III), that includes Tyr10, His13, and His14, is thermodynamically stable in solution at physiological pH. The finding via DFT agrees with the Raman spectra experimental analysis for the most stable models.¹⁵⁹

Finally, most DFT studies mentioned in this section were performed in the gas phase or using a continuum model, but not explicit solvent. Also, the models used often are truncated, small-size models for metal–A β complexes, as a replacement for the full-length A β peptide. Furthermore, these studies did not simulate the dynamics of the disordered A β peptide and its interactions with metal ions, such as large size metallopeptidase, which still represent a significant challenge for current first-principles methods, including DFT.

1.3.2 Semiempirical Methods

Semiempirical calculation methods are applicable for biomolecular system calculations, due to reduced computational expense. The run time of calculations is roughly three orders of magnitude faster than the DFT calculation. Even though the accurate modelling of non-covalent interactions is challenging for pure semiempirical methods, where dispersion and hydrogen-bond interactions are not sufficiently accurate, several models of successful biomolecular simulations with these methods exist, and most of them include corrections.¹⁶⁰ An extended semiempirical tight-binding approach, GFN2-xTB, designed to overcome the noncovalent interaction energy problem that shows fast calculation time for structures with nearly 1000 atoms.¹⁶¹ The PM6/PM7 methods have also have been included for parameterisation and dispersion interaction, and hydrogen bonding. In general, still, there is a small number of studies using this type of theoretical method to model metal-biomolecule systems of large, disordered peptides. In this section, the studies that have been used semiempirical methods that related to the work of this thesis are summarized.

In 2009, Stewart¹⁶² presented the applicability of the PM6 method for modelling larger biochemical molecules, especially proteins, including more complicated organometallic structures. The study investigated metal ions' selectivity, including copper, zinc, iron, cobalt, molybdenum, manganese, magnesium, and potassium. The results for the large peptide systems containing metal ions, illustrate that PM6 is able to predict the X-ray crystallographic data for the geometries of both the covalently bound, and free metal ions. The calculated data displayed agreement with the associated coordination bond lengths from an experiment, which included the electronically complicated Cu(II) systems. In addition, PM6

mimicked the metallo-peptide's secondary structure, as the RMSD difference from the experimental form was 0.82 Å.

Later in 2011, Xia et al.¹⁶³ examined the interactions of β-cyclodextrin (β-CD) with *Trans*-dichloro(dipyridine) platinum(II) (DDP), in terms of structural stability. The geometry optimization of structures determined by PM6 and B3LYP/6-31G* and LanL2DZ basis sets in the gas phase. The calculated results supported the experimental results.

In 2015, Bertoli et al.¹⁶⁴ investigated complex citrate structures of zinc (II) isomers by using PM6 and DFT (B3LYP/LANL2DZ) calculations, in conjunction with experimental spectroscopy, ESI-MS. The study also included findings on the thermodynamic stability of the zinc complex. PM6 and DFT predicted the Zn(II) geometry with high accuracy, when compared with the experimental observations. Furthermore, the PM6 method accurately predicted the energy order of the isomers of the complex, when benchmarked against DFT calculations. However, the relative energy differences were smaller than in DFT. In a similar work, the same authors also used PM6, combined with electrospray ionisation mass spectrometry (ESI-MS), to study metal-citrate systems of complexes composed from Cd(II), Pb(II), Cu(II) and Fe(II).^{165,166} The calculated coordination of isomers gave stable structures, that agreed with experimental data and permitted the assignment of FT-IR spectra. PM6 predicted the energy order accurately, although the relative energies were not sufficiently accurate, due to the thermodynamic stability calculated by PM6, being less exoenergetic than DFT.

Keglevich and co-workers^{167,168} carried out PM6, PM6-DH2 and DFT-B3LYP/(6-31G and LANL2DZ), to evaluate the geometries of Pt(II) complexes with phosphine borane bidentate ligands. Semiempirical and DFT methods calculated the transition metal complexes correctly, with the best agreement to experimental data coming from the DFT calculations. Also, both of the methods detected the formation of weak π-stacking interactions.

In 2014, Shahabadi and Heidari¹⁶⁹ used the semiempirical PM6 method to characterise a newly synthesized Pt(II) complex, containing the drug metformin, bound to DNA. The results revealed a correct prediction of the structure and vibrational frequencies of the complex, using PM6 when compared to FT-IR spectra. Another study by the same author, implemented the PM6 method on Cu(II)-aspartame structure.¹⁷⁰ The method correctly modelled the coordination bonds to the copper, coming in close agreement to the experimental data, and

correctly illustrated the Jahn-Teller distortion geometry around the copper as distorted octahedral.

A combined semiempirical/PM6 and DFT computational protocol for studying bio-organometallic complexes, was developed by Suarez et al.^{171,172} The novel method, was applied to structural and thermodynamic studies looking at molybdocene-cysteine and molybdocene-glutathione complexes. The investigation included a protocol for exploring the conformational space, employing molecular dynamics simulations via the PM6 method in the implicit solvent model, followed by B3LYP-D3 energy calculation, which was applied on the most stable conformers, in terms of relative energy. The study claimed that the protocol used, produced reasonable molecular geometries and energies, as the stability and protonation states detected experimentally, were in good agreement with the average free energies produced from this computational protocol.

The role of the glycan and metal components, heparinase II enzyme binding to Zn(II) and Ca(II) on the protein scaffold, has been investigated by Fernandes et al.,¹⁷³ using semiempirical PM6 calculations. The semiempirical calculations proposed more favourable interactions with Zn(II) over Ca(II), in agreement with experimental observations. This result provides an explanation of protein inactivity, seen in the existence of this Ca ion, where it is able to modify the conformation of residues participating in substrate binding.

De Santis et al.¹⁷⁴ explored the geometry and conformer searches using the semiempirical method PM6 in MOPAC2009, while COSMO was used for geometry minimization, to study the selectivity of monovalent (Na^+ , K^+ , Ag^+) and divalent cations (Ca^{2+} , Ba^{2+} , Sr^{2+} and Mg^{2+}), towards the cyclic peptide hexamer (21-membered rings), composed of alternating α - and β -monomers. The calculated results, show the 21-membered cyclopeptide preferring complexation with divalent cations, over monovalent, and favoured backbone carbonyl coordination. The 6-coordinated complexes form with metals that have ionic radii around 1 Å, such as Ca^{2+} and Ba^{2+} , while 5-coordination is preferred for divalent cations with Ba^{2+} or Mg^{2+} . The coordination geometries predicted by PM6, were shown to agree with both NMR and CD spectra.

Caturello et al.¹⁷⁵ applied the dispersion-corrected PM6 method (PM6-D3H4X), in a vacuum to investigate supramolecular complexes made of Pt^{II} and Pd^{II} based on oligo(phenylene

ethynylene)-based pyridine (OPE) and tetrazolyl-pyridine ligands (TEP). PM6 has been used to describe the thermodynamics of all aggregates, geometries, and intermolecular interactions, which are essential to understand supramolecular polymerization, in close agreement with experimental results.

In 2016, Turner et al.¹⁷⁶ used the semiempirical PM7 method, along with DFT and MM, to examine the binding mode of two Pt(II)-L systems to an N-terminal fragment of the amyloid- β peptide. The investigation showed relatively good agreement between PM7 and DFT energies of platinated conformations, with a value of $R^2 = 0.78$ and the PM7-optimized structures were slightly different from DFT ($RMSD = 0.82 \text{ \AA}$); thus, the authors suggested this method (PM7) is a suitable theoretical method for this type of system, at a much lower computational cost than DFT.

Later in 2017, Mutter et al.¹⁷⁷ used DFT and the semi-empirical PM7 method for evaluation of parameters initiated by Ligand field molecular mechanics (LFMM), for copper (II) bound to the amyloid- β 1–16 peptide fragment. The results showed good agreement of LFMM geometries, compared to DFT and PM7 geometries for the Cu(II)-A β 16 peptide, as well as revealing that these parameters may be applied to the full length A β peptide.

Schmitz et al.¹⁷⁸ tested the application of the GFN2-xTB method on 70 organic proteins and 20 metalloproteins. The investigation was based on C_α -RMSD comparison, between experimental crystal structures and GFN2-xTB optimized structures. The deviations of this method's optimized structures from the experimental crystal structures, were similarly small compared to the experimental crystal structure. The analysis of the coordination geometry of the metal centres gave a good agreement with the reference structures, determined by the X-ray derived structure. The method presented a high potential for application in challenging systems, such large metalloproteins of up to 5,000 atoms.

Even though semiempirical methods have been used in several biological systems that contain metals, including molecular docking simulations of metal-biomolecules,¹⁷⁹ still there is a shortage of applications available using this method in the scientific community.

1.3.3 Molecular Mechanics MM

Even though the missing parameter problem is the main drawback for the application of classical molecular dynamics (MD) simulations to study bioinorganic systems, another weakness of classical force fields is their inability to model the coordination of ligands to transition metal ions correctly. Molecular mechanics (MM) force fields, however, still remain one of the main methods implemented for protein simulations. A range of different methods have been developed to overcome MM limitations, which are essential to conformational flexibility. A selection of the literature that have been used MM for modelling transition metal ions interaction with A β is discussed in this section.

Rappe et al.¹⁸⁰ introduced the Universal Force-Field (UFF) in which the parameters are extended to the entire periodic table, unlike the classic force fields that are limited to proteins, organics, or nucleic acids but not transition metal complexes. UFF has been shown to yield accurate geometries of transition metals, main group and organic compounds, with errors less than 0.1 Å in bond distances and 5° to 10° in angle bending.

In 2002, Chan et al.¹⁸¹ studied the interaction in metal-binding proteins that have a high cysteine content and low molecular weight, known as metallothioneins MT, using a molecular modelling technique. The proteins under MM investigation in this study, were the Zn₇-MT and Cd₇-MT rabbit liver proteins, via the CACHE system that uses the MM2 forcefield, with parameter improvements available for all elements in the Periodic Table. Short MD-MM2 simulations of these systems showed that metal-geometry and general peptide configuration displayed good agreement with spectroscopic data.

Xiang and Ponder¹⁸² proposed a polarizable transition-metal ion force field model for Cu(II) ion, obtained using the atomic multipole optimized energetics for biomolecular applications (AMOEBA) and an angular overlap model (AOM). The study of AMOEBA-AOM accurately examined both aqueous Cu(II) ion and type 1 blue copper (T1Cu) proteins. Parameters were derived through appropriate MM energies to a range of ab initio gas-phase calculations (B2PLYP-D/ccpVDZ and MP2/cc-pVDZ). The AMOEBA-AOM shows excellent agreement with QM for a broad range of calculations on aqueous Cu(II) complexes. The method also takes care of the Jahn-Teller distortion for hexa-aqua Cu(II) complex. This method is much more efficient than semiempirical or hybrid QM methods, when MD

simulations on the T1Cu system contain 48 000 atoms. Even though, the AMOEBA-AOM model is suitable for treating the ligand field effect, it falls short when it comes to strongly covalently-bonded TM systems.

Duarte et al.¹⁸³ used a dummy atom method that considers the central metal atom as being surrounded by dummy atoms, where each atom has a partial charge. The method is able to simulate the nonbonded model for several alkaline-earth and transition-metal complexes. The study improved the parameters for octahedrally coordinated Mn²⁺, Zn²⁺, Mg²⁺, and Ca²⁺ and supplied new parameters for Ni²⁺, Co²⁺, and Fe²⁺. The simulations of different dummy models in the active sites of the human and *E.coli* variants of Glyoxalase enzymes, showed that the models conserve stable coordination geometries, in agreement with experimental results, as well as being able to capture the predicted geometric changes that are predictable upon metal binding, without the need for any external bonds or constraints. In addition, both the M(II)–O distances and the experimental solvation free energies have been specified for transition metals, within the range of values seen in the experiment, with errors at approximately 0.04 Å.

Liao et al.¹⁸⁴ added a nonbonded dummy model for Cu(II), as it was not available before, to the already existing metal ion parameters. The reason associated with the unavailability of the copper model before, was the difficulty in addressing the Jahn–Teller distortion for the Cu(II). The parameters for Zn(II) were derived in this study, based on previous settings for the dummy model (ZnDum), which were developed for octahedral geometries. The validation of the developed parameters was tested, by studying the metal binding in the amyloid-β peptide and the mixed-metal enzyme superoxide dismutase, via 100 ns MD simulations. The results showed that the derived parameters can reproduce square planar Cu(II) geometries for the two systems. The Zn(II) model, showed a deviation in the geometry of metal coordination, due to the dummy model of zinc being developed for octahedral geometries, not tetrahedral. The comparison of zinc and copper affinity toward Aβ via the parameters of the model show Zn(II) has a lower binding affinity than Cu(II), which is in agreement with experimental findings. The CuDum and the adaptation of ZnDum models were implemented in GROMACS.

Huy et al.¹⁸⁵ developed a new AMBER force field for Cu(II) interaction, using both Aβ monomer and dimers. The derived parameters had been applied at a microsecond timescale MD, to examine the impact of Cu(II) binding on structures, by looking at the dynamics of

$\text{A}\beta_{42}$ monomer and dimers. The coordination geometry had a distorted planar geometry, with the Cu(II) bound to His6, His13 (or His14), and Asp1. Cu(II) parameters were taken from small optimized models systems (eight $\text{A}\beta_{16}$ monomer/dimer), via unrestricted hybrid density functional method (UB3LYP), in the SMD (Solvation Model based on Density) implicit solvent. The 6-31G(d) basis set was used for main group elements (C, N, O, and H) and the SDD ECP for Cu(II). The MD simulation shows the system reached pseudo-equilibration after several hundred nanoseconds of simulation. The data indicated significant changes in the peptide properties, such as salt-bridges, hydrogen bonding, secondary structure and radius of gyration upon Cu(II) binding. The β -sheets of the monomer decreased significantly in the presence of Cu(II), compared to the wild-type $\text{A}\beta_{42}$, which agrees with experiments which found metal ions facilitate the formation of amorphous aggregates, rather than amyloid fibrils with cross- β structures. The flexibility of the Asp23–Lys28 salt bridge, upon copper ion binding to the $\text{A}\beta$ N-terminus, plays an essential role in the β -sheet formation, supporting the experimental findings that metal ions promote amorphous aggregation.

Also, Raffa and Rauk¹⁸⁶ parameterised the force field for the Cu(II) binding site, via ab initio methods. The defined parameters were then applied for the following 790 ns MD simulation on the aqueous Cu(II)- $\text{A}\beta_{42}$ system. B3LYP/6-31G(d) level optimisations of small model structures were used to obtain Cu(II) parameters. Three different forms of coordination for Cu(II)- $\text{A}\beta$ interactions, were His13, His14, and oxygen or nitrogen from the backbone atom, and the water involved in the coordination. The water was replaced with 4-methylimidazole in one of the models. The system equilibrated after 350 ns into a collapsed coil conformation, which is in line with the structure determined by the experiment, but not always dependent on the exact metal coordination mode.

A different approach for modelling metal-peptide interactions is using Ligand Field Molecular Mechanics (LFMM), which was introduced by Burton and Deeth,¹⁸⁷ in 1995 (see Section 2.4.1). Later, the LFMM parameters were developed by Deeth,¹⁸⁸ as an alternative to QM/MM methods for modelling metalloproteins of type I blue copper proteins. The parameterisation is developed based on DFT data for simple, homoleptic models. The first application of LFMM to second-row donor atoms, such as sulfur, was included in this study, although it was restricted to small molecules with nitrogen donors. LFMM structures for the sulfur-ligated and Type I active site models, agree with data from DFT homoleptic models. LFMM

successfully reproduce the DFT structure, specifically the *trans* influence of the S_{CYS} ligand, which leads to lengthening of the pseudo-*trans* distance of Cu-N (2.07 Å), and pseudo-*cis* of Cu-N bond (1.99 Å). The LFMM method was also employed to study five complete copper-containing proteins. The LFMM optimization of the structures was performed within a shell of water molecules, 5 Å thick. The resulting systems demonstrate root-mean-square deviations (RMSD) of less than 0.42 Å, compared to experimental data.

In another study by Deeth in 2007,¹⁸⁹ the author implemented LFMM to model a vast system of proteins, comprising of oxidized Type I copper of four- or five-coordinate centres. These LFMM optimized structures were compared to experimental structures. Also, LFMM parameters for Cu–N(imidazole), Cu–S(thiolate), Cu–S(thioether), and Cu–O(carbonyl) interactions, were improved through the use of experimental and theoretical data from DFT (TZP/DZP/PW91), for homoleptic model complexes. The protein modelling via LFMM of the complete systems, was optimized in double layers of water. The results fall within an acceptable experimental error, as LFMM structures reproduced the entire peptide geometry, and gross errors seen were originally found in the PDB structures. The main feature of this LFMM approach is its ability to provide accuracy comparable to QM/MM calculations data, at orders of magnitude faster.

Tai et al.¹⁹⁰ used the LFMM approach and experimental studies to investigate the DNA changes, in platinum(II)-containing fragments derived from cisplatin, by Pt(II) complexes. LFMM successfully describes the Pt(II)-guanine interactions, when compared to experimental and DFT data. Also, ligand field molecular dynamics (LFMD) simulations were performed, and 6 ns of production data was extracted for the system, under periodic boundary conditions, using the TIP3P water model via DL-POLY-LF. The resulted geometry showed good agreement with experimental data, while DNA was found to bend near the minor groove.

Mutter et al.¹⁷⁷ used DFT (B3LYP, M06-2X, and BHandH functional, both with and without Grimme's empirical dispersion correction with 6-31G(d)/LANL2DZ basis sets), to optimise small test models of different possible Cu(II)-Aβ16 coordination modes, hence evaluating the accuracy of LFMM parameters. The resulting LFMM bond lengths and angles presented errors typically less than 0.1 Å and 5°, respectively. Then, short LFMD simulations were performed, using the derived parameters, on the copper bound Cu(II)-Aβ16 peptide. Several snapshots

were extracted from the trajectory for further optimization via DFT and PM7, which resulted in good agreement between these methods and LFMM geometries, in particular by PM7.

In 2016, Mutter et al.¹⁷⁶ also used LFMM parameters to study the binding of platinum-aryl ligands interaction with the A β peptide's N-terminal fragment. The LFMM approach has proven to be a powerful tool to study metal-peptide interactions, while also being able to predict reliable geometries and conformations, with the geometries of LFMM and DFT in good general agreement (mean RMSD = 0.67 Å). Despite that, LFMM fails to reproduce calculated DFT relative energies of conformers ($R^2 = 0.14$).

Boopathi and Kolandaivel¹⁹¹ examined the binding between Zn(II), Cu(II) and Fe(II), and A β 42, by performing 200 ns MD simulations in TIP3P explicit solvent. DFT M06-2X/6-311++G (2df,2pd) and LANL2DZ basis sets were used to define the force field of metal and histidine interactions. For optimization, the metal ions bind to three histidine residues without a fourth ligand, with the parameters derived via OPLS-AA. The MD simulations show Fe(II) binding reduces the coil structures and increases the β -sheet content, which leads to the loss of the helical structure of N-terminal residues. In contrast, the Cu²⁺ binding to the A β 42 shows β -sheet formation at the N-terminal residues of the peptide. Fe²⁺ binding transpires in the formation of Glu22-Lys28 salt-bridge, that can stabilize the turn conformation in the Phe19-Gly25 residues, suggesting Fe(II) binding promotes aggregation by enhancing the disordered peptide-peptide interaction that causes the oligomerization.

In addition, a number of groups have carried out standard and enhanced MD simulations, to study the effects of metal ion binding on the structure and dynamics of monomeric and oligomeric A β , such as through the use of replica exchange molecular dynamics (REMD) method.

Li et al. performed REMD simulations, using the Amber ff03 force field for the protein in implicit solvent, to investigate the effects of the zinc-binding on the conformational distributions of the A β 40 peptide.¹⁹² In this study, the zinc-binding parameters are nonbonded, derived from the modelling of the active site of HCAII (human carbonic anhydrase II). The coordination was formed from histidine N atoms in His6, His13, and His14, and the Glu11 carboxylate group oxygen atom bound to Zn(II), using bond constraints that were added between zinc and its coordinating atoms, the coordination was fitted to experimental

data. The results show that zinc significantly affects the conformational distribution of the monomer upon its binding to A β . The Zn-A β peptide sampling have more conformations which are more likely to aggregate, compared to the metal-free peptide. The authors suggested that the conformational modification of the A β monomer, upon metal binding, can be one of the possible mechanisms for the metal promoted aggregation of the A β peptide.

One of the approaches based on bonded models used to build parameters for metal sites in proteins, is the metal centre parameter builder (MCPB). The program was developed by Dr. Martin Peters from Merz research group in AMBER in 2010.¹⁹³ The study focused on zinc metalloproteins, due to the abundance of such systems in the protein data bank. The ab initio calculations (B3LYP/6-31G*) are performed on the first coordination sphere of the metal-protein, within the program's workflow, to generate updated AMBER-like force field parameters, such as charges, q_i , and determine bond (K_r) and angle (K_θ) forces constant. The generated force field parameters are then tested using minimization techniques, to examine their stability. Thus, MCPB is utilised to construct a zinc force field, which is compatible with the existing AMBER FFs for MD simulations, in order to study zinc-containing proteins, including metalloprotein crystal refinement. MCPB.py¹⁹⁴ is a python version of MCPB that can model more than 80 metal ions and various AMBER force fields. The program was developed by Pengfei Li from within the Merz research group. In order to derive force constants and charge parameters, steps of parametrization schemes are carried out within the program. The program evaluation via two metal-containing systems, shows its efficiency to provide reliable force fields for different metal ion containing complexes. A number of studies have used MCPB for metal ion modelling.^{195,196} Turner et al.¹⁹⁷ utilized the method in order to parameterize the transition metal, for investigation of the coordination effects on A β -16 peptide.

1.4 References

- 1 L. Pickart et al., The Human Tripeptide GHK-Cu in Prevention of Oxidative Stress and Degenerative Conditions of Aging: Implications for Cognitive Health, *Oxid. Med. Cell. Longev.*, 2012, **2012**, 1–8.
- 2 R. A. Stelzmann et al., An English translation of Alzheimer's 1907 paper, "über eine eigenartige erkankung der hirnrinde", *Clin. Anat.*, 1995, **8**, 429–431.
- 3 ALZHEIMER et al., Über einen eigenartigen schweren Erkrankungsprozess der Hirrinde, *Neurol. Cent.*, 1906, **25**, 1134.
- 4 A. Burns et al., Alzheimer's disease, *Lancet*, 2002, **360**, 163–165.
- 5 K. P. Kepp, Alzheimer's disease due to loss of function: A new synthesis of the available data, *Prog. Neurobiol.*, 2016, **143**, 36–60.
- 6 D. Valensin et al., Metal compounds as inhibitors of β -amyloid aggregation. Perspectives for an innovative metallotherapeutics on Alzheimer's disease, *Coord. Chem. Rev.*, 2012, **256**, 2357–2366.
- 7 C. Reitz et al., Relation between smoking and risk of dementia and Alzheimer disease: the Rotterdam Study, *Neurology*, 2007, **69**, 998–1005.
- 8 A. Ruitenberg et al., Alcohol consumption and risk of dementia: the Rotterdam Study, *Lancet*, 2002, **359**, 281–286.
- 9 L. Jones et al., Convergent genetic and expression data implicate immunity in Alzheimer's disease, *Alzheimer's Dement.*, 2015, **11**, 658–671.
- 10 B. Singh et al., Association of mediterranean diet with mild cognitive impairment and Alzheimer's disease: a systematic review and meta-analysis, *J. Alzheimer's Dis.*, 2014, **39**, 271–282.
- 11 T. Lu et al., Gene regulation and DNA damage in the ageing human brain, *Nature*, 2004, **429**, 883–891.
- 12 D. M. Holtzman et al., Alzheimer's disease: the challenge of the second century, *Sci. Transl. Med.*, 2011, **3**, 77sr1-77sr1.

- 13 J. F. Norfray et al., Alzheimer's Disease Neuropathologic Findings and Recent Advances in Imaging, *Am. J. Roentgenol.*, 2004, **182**, 3–13.
- 14 F. M. LaFerla et al., Intracellular amyloid- β in Alzheimer's disease, *Nat. Rev. Neurosci.*, 2007, **8**, 499–509.
- 15 H. W. Querfurth et al., Alzheimer's Disease, *N. Engl. J. Med.*, 2010, **362**, 329–344.
- 16 J. A. Hardy et al., Alzheimer's Disease: The Amyloid Cascade Hypothesis, *Science.*, 1992, **256**, 184–185.
- 17 H. Kozlowski et al., Copper, zinc and iron in neurodegenerative diseases (Alzheimer's, Parkinson's and prion diseases), *Coord. Chem. Rev.*, 2012, **256**, 2129–2141.
- 18 M. A. Greenough et al., Metal dyshomeostasis and oxidative stress in Alzheimer's disease, *Neurochem. Int.*, 2013, **62**, 540–555.
- 19 K. P. Kepp, Alzheimer's disease: How metal ions define β -amyloid function, *Coord. Chem. Rev.*, 2017, **351**, 127–159.
- 20 R. De Ricco et al., NMR investigations of metal interactions with unstructured soluble protein domains, *Coord. Chem. Rev.*, 2014, **269**, 1–12.
- 21 S. Chakraborty et al., Early calcium dysregulation in Alzheimer's disease: setting the stage for synaptic dysfunction, *Sci. China Life Sci.*, 2011, **54**, 752–762.
- 22 G. E. Gibson et al., Oxidative stress in Alzheimer's disease, *Neurobiol. Aging*, 2005, **5**, 575–578.
- 23 K. Honda et al., Oxidative stress and redox-active iron in Alzheimer's disease, *Ann. N. Y. Acad. Sci.*, 2004, **1012**, 179–182.
- 24 A. I. Bush, The metallobiology of Alzheimer's disease, *Trends Neurosci.*, 2003, **26**, 207–214.
- 25 A. I. Bush et al., Therapeutics for Alzheimer's disease based on the metal hypothesis, *Neurotherapeutics*, 2008, **5**, 421–432.
- 26 C. S. Atwood et al., Dramatic aggregation of Alzheimer abeta by Cu(II) is induced by conditions representing physiological acidosis, *J. Biol. Chem.*, 1998, **273**, 12817–12826.

- 27 A. I. Bush et al., Rapid induction of Alzheimer A beta amyloid formation by zinc, *Science.*, 1994, **265**, 1464–1467.
- 28 X. Huang et al., The A β Peptide of Alzheimer’s Disease Directly Produces Hydrogen Peroxide through Metal Ion Reduction, *Biochemistry*, 1999, **38**, 7609–7616.
- 29 X. Huang et al., Cu(II) potentiation of alzheimer abeta neurotoxicity. Correlation with cell-free hydrogen peroxide production and metal reduction, *J. Biol. Chem.*, 1999, **274**, 37111.
- 30 C. A. Rottkamp et al., Redox-active iron mediates amyloid- β toxicity, *Free Radic. Biol. Med.*, 2001, **30**, 447–450.
- 31 M. Lovell et al., Copper, iron and zinc in Alzheimer’s disease senile plaques, *J. Neurol. Sci.*, 1998, **158**, 47–52.
- 32 R. A. Cherny et al., Treatment with a Copper-Zinc Chelator Markedly and Rapidly Inhibits ; Amyloid Accumulation in Alzheimer’s Disease Transgenic Mice, *Neuron*, 2001, **30**, 665–676.
- 33 G. K. Gouras et al., Metal Chelator Decreases Alzheimer; β -Amyloid Plaques, *Neuron*, 2001, **30**, 641–642.
- 34 M. P. Cuajungco et al., in *Alzheimer’s Disease: Cellular and Molecular Aspects of Amyloid- β* , eds. J. R. Harris and F. Fahrenholz, Springer US, Boston, MA, 2005, pp. 235–254.
- 35 K. P. Kepp et al., Bioinorganic Chemistry of Alzheimer’s Disease, *Chem. Rev.*, 2012, **112**, 5193–5239.
- 36 D. M. A. A. Mann, Pyramidal nerve cell loss in Alzheimer’s disease, *Neurodegeneration*, 1996, **5**, 423–427.
- 37 C. L. Masters et al., Amyloid plaque core protein in Alzheimer disease and Down syndrome, *Proc. Natl. Acad. Sci.*, 1985, **82**, 4245–4249.
- 38 J. Hardy et al., Amyloid deposition as the central event in the aetiology of Alzheimer’s disease, *Trends Pharmacol. Sci.*, 1991, **12**, 383–388.

- 39 G. G. Glenner et al., Alzheimer's disease: Initial report of the purification and characterization of a novel cerebrovascular amyloid protein, *Biochem. Biophys. Res. Commun.*, 1984, **120**, 885–890.
- 40 R. A. Cherny et al., Aqueous dissolution of Alzheimer's disease Abeta amyloid deposits by biometal depletion, *J. Biol. Chem.*, 1999, **274**, 23223–23228.
- 41 F. Chiti et al., Protein misfolding, functional amyloid, and human disease, *Annu. Rev. Biochem.*, 2006, **75**, 333–366.
- 42 A. Rauk, The chemistry of Alzheimer's disease, *Chem. Soc. Rev.*, 2009, **38**, 2698–2715.
- 43 J. Hardy et al., The Amyloid Hypothesis of Alzheimer's Disease: Progress and Problems on the Road to Therapeutics, *Science.*, 2002, **297**, 353 LP – 356.
- 44 M. K. Tiwari et al., Modeling the aggregation propensity and toxicity of amyloid- β variants, *J. Alzheimer's Dis.*, 2015, **47**, 215–229.
- 45 A. K. Somavarapu et al., Direct correlation of cell toxicity to conformational ensembles of genetic A β Variants, *ACS Chem. Neurosci.*, 2015, **6**, 1990–1996.
- 46 K. N. Dahlgren et al., Oligomeric and fibrillar species of amyloid- β peptides differentially affect neuronal viability, *J. Biol. Chem.*, 2002, **277**, 32046–32053.
- 47 A. L. Hansen et al., Measurement of histidine pKa values and tautomer populations in invisible protein states, *Proc. Natl. Acad. Sci.*, 2014, **111**, E1705–E1712.
- 48 B. A. Yankner et al., Neurotrophic and neurotoxic effects of amyloid beta protein: reversal by tachykinin neuropeptides, *Science.*, 1990, **250**, 279–282.
- 49 S. A. Frautschy et al., Effects of injected Alzheimer beta-amyloid cores in rat brain, *Proc. Natl. Acad. Sci.*, 1991, **88**, 8362–8366.
- 50 C. J. Pike et al., In vitro aging of β -amyloid protein causes peptide aggregation and neurotoxicity, *Brain Res.*, 1991, **563**, 311–314.
- 51 A. F. Teich et al., Is the amyloid hypothesis of Alzheimer's disease therapeutically relevant? *Biochem. J.*, 2012, **446**, 165–177.
- 52 M. V. Rogacheva et al., Impact of Pyrophosphate and O-Ethyl-Substituted

- Pyrophosphate Groups on DNA Structure, *J. Phys. Chem. B*, 2006, **111**, 432–438.
- 53 J. Hardy, Alzheimer's disease: The amyloid cascade hypothesis: An update and reappraisal, *J. Alzheimers. Dis.*, 2006, **9**, 151–153.
- 54 P. Sorrentino et al., The dark sides of amyloid in Alzheimer's disease pathogenesis, *Febs Lett.*, 2014, **588**, 641–652.
- 55 H. J. Aizenstein et al., Frequent amyloid deposition without significant cognitive impairment among the elderly, *Arch. Neurol.*, 2008, **65**, 1509–1517.
- 56 F. H. Bouwman et al., CSF biomarker levels in early and late onset Alzheimer's disease, *Neurobiol. Aging*, 2009, **30**, 1895–1901.
- 57 C. R. Jack Jr et al., Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade, *Lancet Neurol.*, 2010, **9**, 119–128.
- 58 D. M. Walsh et al., Naturally secreted oligomers of amyloid β protein potently inhibit hippocampal long-term potentiation in vivo, *Nature*, 2002, **416**, 535–539.
- 59 J. P. Cleary et al., Natural oligomers of the amyloid- β protein specifically disrupt cognitive function, *Nat. Neurosci.*, 2005, **8**, 79–84.
- 60 V. J. Burton et al., Molecular Mechanics for Coordination Complexes: The Impact of Adding d-Electron Stabilization Energies, *J. Am. Chem. Soc.*, 2002, **117**, 8407–8415.
- 61 C. Haass et al., Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid β -peptide, *Nat. Rev. Mol. Cell Biol.*, 2007, **8**, 101–112.
- 62 F. Hefti et al., The case for soluble A β ; oligomers as a drug target in Alzheimer's disease, *Trends Pharmacol. Sci.*, 2013, **34**, 261–266.
- 63 C. G. Glabe, Common mechanisms of amyloid oligomer pathogenesis in degenerative disease, *Neurobiol. Aging*, 2006, **27**, 570–575.
- 64 L. W. Hung et al., Amyloid- β peptide (A β) neurotoxicity is modulated by the rate of peptide aggregation: A β dimers and trimers correlate with neurotoxicity, *J. Neurosci.*, 2008, **28**, 11950–11958.
- 65 C. Schmitz et al., Hippocampal neuron loss exceeds amyloid plaque load in a transgenic

- mouse model of Alzheimer's disease, *Am. J. Pathol.*, 2004, **164**, 1495–1502.
- 66 E. T. Sawey et al., Identification of a therapeutic strategy targeting amplified FGF19 in liver cancer by Oncogenomic screening, *Cancer Cell*, 2011, **19**, 347–358.
- 67 S. Wärmländer et al., Biophysical Studies of the Amyloid β-Peptide: Interactions with Metal Ions and Small Molecules, *ChemBioChem*, 2013, **14**, 1692–1704.
- 68 S. Vivekanandan et al., A partially folded structure of amyloid-beta (1–40) in an aqueous environment, *Biochem. Biophys. Res. Commun.*, 2011, **411**, 312–316.
- 69 H. Sticht et al., Structure of amyloid A4-(1–40)-peptide of Alzheimer's disease, *Eur. J. Biochem.*, 1995, **233**, 293–298.
- 70 A. K. Somavarapu et al., The Dependence of Amyloid-β Dynamics on Protein Force Fields and Water Models, *ChemPhysChem*, 2015, **16**, 3278–3289.
- 71 A. T. Petkova et al., Experimental Constraints on Quaternary Structure in Alzheimer's β-Amyloid Fibrils, *Biochemistry*, 2005, **45**, 498–512.
- 72 J. J. Balbach et al., Supramolecular structure in full-length Alzheimer's β-amyloid fibrils: Evidence for a parallel β-sheet organization from solid-state nuclear magnetic resonance, *Biophys. J.*, 2002, **83**, 1205–1216.
- 73 O. N. Antzutkin et al., Multiple quantum solid-state NMR indicates a parallel, not antiparallel, organization of β-sheets in Alzheimer's β-amyloid fibrils, *Proc. Natl. Acad. Sci.*, 2000, **97**, 13045–13050.
- 74 M. Coles et al., Solution structure of amyloid β-peptide (1– 40) in a water- micelle environment. Is the membrane-spanning domain where we think it is? *Biochemistry*, 1998, **37**, 11064–11077.
- 75 R. Roychaudhuri et al., Amyloid β-protein assembly and Alzheimer disease, *J. Biol. Chem.*, 2009, **284**, 4749–4753.
- 76 W. Kaim et al., *Bioinorganic Chemistry--Inorganic Elements in the Chemistry of Life: An Introduction and Guide*, John Wiley & Sons, 2013.
- 77 J. C. Stroud et al., Toxic fibrillar oligomers of amyloid-β have cross-β structure, *Proc.*

Natl. Acad. Sci., 2012, **109**, 7717–7722.

- 78 Y. Yoshiike et al., Specific compositions of amyloid-beta peptides as the determinant of toxic beta-aggregation, *J. Biol. Chem.*, 2003, **278**, 23648–23655.
- 79 S. Nag et al., Nature of the Amyloid- β ; Monomer and the Monomer-Oligomer Equilibrium, *J. Biol. Chem.*, 2011, **286**, 13827–13833.
- 80 M. Townsend et al., Effects of secreted oligomers of amyloid β -protein on hippocampal synaptic plasticity: a potent role for trimers, *J. Physiol.*, 2006, **572**, 477–492.
- 81 G. M. Shankar et al., Amyloid- β protein dimers isolated directly from Alzheimer's brains impair synaptic plasticity and memory, *Nat. Med.*, 2008, **14**, 837–842.
- 82 A. Abelein et al., Hydrophobicity and conformational change as mechanistic determinants for nonspecific modulators of amyloid β self-assembly, *Biochemistry*, 2012, **51**, 126–137.
- 83 K. Ono et al., Structure–neurotoxicity relationships of amyloid β -protein oligomers, *Proc. Natl. Acad. Sci.*, 2009, **106**, 14745 LP – 14750.
- 84 M. Ahmed et al., Structural conversion of neurotoxic amyloid- β 1–42 oligomers to fibrils, *Nat. Struct. Mol. Biol.*, 2010, **17**, 561–567.
- 85 S. A. Kotler et al., High-resolution NMR characterization of low abundance oligomers of amyloid- β without purification, *Sci. Rep.*, 2015, **5**, 11811.
- 86 T. Lührs et al., 3D structure of Alzheimer's amyloid- β (1–42) fibrils, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 17342 LP – 17347.
- 87 J. D. Harper et al., Assembly of A β amyloid protofibrils: an in vitro model for a possible early event in Alzheimer's disease, *Biochemistry*, 1999, **38**, 8972–8980.
- 88 A. Hatami et al., Familial Alzheimer's Disease Mutations within the Amyloid Precursor Protein Alter the Aggregation and Conformation of the Amyloid- β Peptide, *J. Biol. Chem.*, 2017, **292**, 3172–3185.
- 89 D. Eisenberg et al., The Amyloid State of Proteins in Human Diseases, *Cell*, 2012, **148**, 1188–1203.

- 90 W. Qiang et al., Antiparallel β -sheet architecture in Iowa-mutant β -amyloid fibrils, *Proc. Natl. Acad. Sci.*, 2012, **109**, 4443 LP – 4448.
- 91 A. T. Petkova et al., Experimental constraints on quaternary structure in Alzheimer's β -amyloid fibrils, *Biochemistry*, 2006, **45**, 498–512.
- 92 T. L. S. Benzinger et al., Propagating structure of Alzheimer's β -amyloid (10–35) is parallel β -sheet with residues in exact register, *Proc. Natl. Acad. Sci.*, 1998, **95**, 13407–13412.
- 93 I. Bertini et al., A new structural model of A β 40 fibrils, *J. Am. Chem. Soc.*, 2011, **133**, 16013–16022.
- 94 A. T. Petkova et al., Experimental constraints on quaternary structure in Alzheimer's β -amyloid fibrils, *Biochemistry*, 2006, **45**, 498–512.
- 95 D. A. Kirschner et al., Synthetic peptide homologous to beta protein from Alzheimer disease forms amyloid-like fibrils in vitro, *Proc. Natl. Acad. Sci.*, 1987, **84**, 6953–6957.
- 96 H. Inouye et al., Structure of beta-crystallite assemblies formed by Alzheimer beta-amyloid protein analogues: analysis by x-ray diffraction, *Biophys. J.*, 1993, **64**, 502–519.
- 97 M. Sunde et al., Common core structure of amyloid fibrils by synchrotron X-ray diffraction, *J. Mol. Biol.*, 1997, **273**, 729–739.
- 98 R. S. Armen et al., Anatomy of an Amyloidogenic Intermediate: Conversion of β -Sheet to α -Sheet Structure in Transthyretin at Acidic pH, *Structure*, 2004, **12**, 1847–1863.
- 99 D. Shea et al., α -Sheet secondary structure in amyloid β -peptide drives aggregation and toxicity in Alzheimer's disease, *Proc. Natl. Acad. Sci.*, 2019, **116**, 8895–8900.
- 100 R. S. Armen et al., Pauling and Corey's α -pleated sheet structure may define the prefibrillar amyloidogenic intermediate in amyloid disease, *Proc. Natl. Acad. Sci.*, 2004, **101**, 11622–11627.
- 101 M. D. Carter et al., The Development of New Therapeutics for Alzheimer's Disease, *Clin. Pharmacol. Ther.*, 2010, **88**, 475–486.
- 102 B. De Strooper et al., Learning by Failing: Ideas and Concepts to Tackle γ -Secretases in

- Alzheimer's Disease and Beyond, *Annu. Rev. Pharmacol. Toxicol.*, 2015, **55**, 419–437.
- 103 E. Karan et al., A critique of the drug discovery and phase 3 clinical programs targeting the amyloid hypothesis for Alzheimer disease, *Ann. Neurol.*, 2014, **76**, 185–205.
- 104 J. E. House, *Inorganic Chemistry*, Elsevier Science, 2012.
- 105 A. I. Bush, The Metal Theory of Alzheimer's Disease, *J. Alzheimer's Dis.*, 2013, **33**, S277–S281.
- 106 J. D. Robertson et al., Disruption of zinc homeostasis in Alzheimer's disease, *Nucl. Instruments Methods Phys. Res. Sect. B Beam Interact. with Mater. Atoms*, 2002, **189**, 454–458.
- 107 Y. H. Hung et al., Copper in the brain and Alzheimer's disease, *J. Biol. Inorg. Chem.*, 2010, **15**, 61–76.
- 108 A. S. Pithadia et al., Metal-associated amyloid- β species in Alzheimer's disease, *Curr. Opin. Chem. Biol.*, 2012, **16**, 67–73.
- 109 E. Gaggelli et al., Copper Homeostasis and Neurodegenerative Disorders (Alzheimer's, Prion, and Parkinson's Diseases and Amyotrophic Lateral Sclerosis), *Chem. Rev.*, 2006, **106**, 1995–2044.
- 110 N. Nelson, Metal ion transporters and homeostasis, *EMBO J.*, 1999, **18**, 4361–4371.
- 111 M. J. Berridge et al., Calcium-a life and death signal, *Nature*, 1998, **395**, 645–648.
- 112 F. Chimienti et al., Role of cellular zinc in programmed cell death: temporal relationship between zinc depletion, activation of caspases, and cleavage of Sp family transcription factors
Abbreviations: Chx, cycloheximide; PARP, poly(ADP-ribose) polymerase; TNF α , tumor necrosis factor, *Biochem. Pharmacol.*, 2001, **62**, 51–62.
- 113 W. J. Niklowitz et al., Neurofibrillary Changes Following Childhood Lead Encephalopathy: Case Report, *J. Neuropathol. Exp. Neurol.*, 1975, **34**, 445–455.
- 114 W. J. Niklowitz, Neurofibrillary changes after acute experimental lead poisoning, *Neurology*, 1975, **25**, 927 LP – 927.
- 115 F. M. Burnet, A possible Role of Zinc in the Pathology of Dementia, *Lancet*, 1981, **317**,

186–188.

- 116 D. W. Bonhaus et al., The effects of aluminum on microtubular integrity using in vitro and in vivo models, *Toxicol. Lett.*, 1980, **6**, 141–147.
- 117 E. I. Solomon et al., Copper Active Sites in Biology, *Chem. Rev.*, 2014, **114**, 3659–3853.
- 118 E. I. Solomon et al., Geometric and Electronic Structure/Function Correlations in Non-Heme Iron Enzymes, *Chem. Rev.*, 2000, **100**, 235–350.
- 119 C. Hureau, Coordination of redox active metal ions to the amyloid precursor protein and to amyloid- β peptides involved in Alzheimer disease. Part 1: An overview, *Coord. Chem. Rev.*, 2012, **256**, 2164–2174.
- 120 S. Zirah et al., Structural changes of region 1-16 of the Alzheimer disease amyloid beta-peptide upon zinc binding and in vitro aging., *J. Biol. Chem.*, 2006, **281**, 2151–61.
- 121 E. Gaggelli et al., NMR Studies of the Zn²⁺ Interactions with Rat and Human β -Amyloid (1-28) Peptides in Water-Micelle Environment, *J. Phys. Chem. B*, 2008, **112**, 100–109.
- 122 F. Bousejra-ElGarah et al., Iron(II) Binding to Amyloid- β , the Alzheimer's Peptide, *Inorg. Chem.*, 2011, **50**, 9024–9030.
- 123 C. S. Atwood et al., Characterization of copper interactions with Alzheimer amyloid β peptides: Identification of an attomolar-affinity copper binding site on amyloid β 1-42, *J. Neurochem.*, 2000, **75**, 1219–1233.
- 124 V. Tõugu et al., Zn(II)- and Cu(II)-induced non-fibrillar aggregates of amyloid- β (1–42) peptide are transformed to amyloid fibrils, both spontaneously and under the influence of metal chelators, *J. Neurochem.*, 2009, **110**, 1784–1795.
- 125 J. H. Viles, Metal ions and amyloid fiber formation in neurodegenerative diseases. Copper, zinc and iron in Alzheimer's, Parkinson's and prion diseases, *Coord. Chem. Rev.*, 2012, **256**, 2271–2284.
- 126 C. J. Sarell et al., Substoichiometric Levels of Cu²⁺ Ions Accelerate the Kinetics of Fiber Formation and Promote Cell Toxicity of Amyloid- β from Alzheimer Disease*, *J. Biol. Chem.*, 2010, **285**, 41533–41540.

- 127 Y. Yoshiike et al., New Insights on How Metals Disrupt Amyloid β -Aggregation and Their Effects on Amyloid- β Cytotoxicity, *J. Biol. Chem.*, 2001, **276**, 32293–32299.
- 128 N. S. A. Ghani et al., Unveiling Amyloid- β 1–42 Interaction with Zinc in Water and Mixed Hexafluoroisopropanol Solution in Alzheimer’s Disease, *Int. J. Pept. Res. Ther.*, 2017, **23**, 393–407.
- 129 G. Pizzino et al., Oxidative Stress: Harms and Benefits for Human Health, *Oxid. Med. Cell. Longev.*, 2017, **2017**, 8416763.
- 130 P. Held, An introduction to reactive oxygen species, *Tech Resour. Guid.*, 2012, **802**, 5–9.
- 131 A. S. DeToma et al., Misfolded proteins in Alzheimer’s disease and type II diabetes, *Chem. Soc. Rev.*, 2012, **41**, 608–621.
- 132 S. Bandyopadhyay et al., Novel drug targets based on metallobiochemistry of Alzheimer’s disease, *Expert Opin. Ther. Targets*, 2010, **14**, 1177–1197.
- 133 N. X. and L. Liu, *Mini-Reviews Med. Chem.*, 2014, **14**, 271–281.
- 134 J. Ceccom et al., Copper Chelator Induced Efficient Episodic Memory Recovery in a Non-Transgenic Alzheimer’s Mouse Model, *PLoS One*, 2012, **7**, e43105.
- 135 K. P. Kepp, Bioinorganic Chemistry of Alzheimer’s Disease, *Chem. Rev.*, 2012, **112**, 5193–5239.
- 136 G. Perry et al., Alzheimer Disease and Oxidative Stress, *J. Biomed. Biotechnol.*, 2002, **2**, 542340.
- 137 M. Dumont et al., Neuroprotective strategies involving ROS in Alzheimer disease, *Free Radic. Biol. Med.*, 2011, **51**, 1014–1026.
- 138 A. Spinello et al., Metal Ions and Metal Complexes in Alzheimer’s Disease, *Curr. Pharm. Des.*, 2016, **22**, 3996–4010.
- 139 P. Vlieghe et al., Synthetic therapeutic peptides: science and market, *Drug Discov. Today*, 2010, **15**, 40–56.
- 140 S. M. Mandal et al., Challenges and future prospects of antibiotic therapy: from

peptides to phages utilization, *Front. Pharmacol.*, 2014, **5**, 105.

- 141 D. J. Craik et al., The future of peptide-based drugs, *Chem. Biol. Drug Des.*, 2013, **81**, 136–147.
- 142 A. Shamloo et al., Designing a new multifunctional peptide for metal chelation and A β inhibition, *Arch. Biochem. Biophys.*, 2018, **653**, 1–9.
- 143 A. I. Bush, Metal complexing agents as therapies for Alzheimer's disease, *Neurobiol. Aging*, 2002, **23**, 1031–1038.
- 144 P. J. Crouch et al., Therapeutic treatments for Alzheimer's disease based on metal bioavailability., *Drug News Perspect.*, 2006, **19**, 469–474.
- 145 B. Strodel et al., Transition Metal Ion Interactions with Disordered Amyloid- β Peptides in the Pathogenesis of Alzheimer's Disease: Insights from Computational Chemistry Studies, *J. Chem. Inf. Model.*, 2019, **59**, 1782–1805.
- 146 F. Neese, A critical evaluation of DFT, including time-dependent DFT, applied to bioinorganic chemistry, *J. Biol. Inorg. Chem.*, 2006, **11**, 702–711.
- 147 C. J. Cramer et al., Density functional theory for transition metals and transition metal chemistry, *Phys. Chem. Chem. Phys.*, 2009, **11**, 10757–10816.
- 148 C. Lee et al., Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density, *Phys. Rev. B*, 1988, **37**, 785–789.
- 149 V. A. Streletsov et al., The structure of the amyloid-beta peptide high-affinity copper II binding site in Alzheimer disease, *Biophys. J.*, 2008, **95**, 3447–3456.
- 150 T. Marino et al., On the metal ion (Zn²⁺, Cu²⁺) coordination with beta-amyloid peptide: DFT computational study, *Interdiscip. Sci. Comput. Life Sci.*, 2010, **2**, 57–69.
- 151 J. Alí-Torres et al., Three Dimensional Models of Cu²⁺-A β (1–16) Complexes from Computational Approaches, *J. Am. Chem. Soc.*, 2011, **133**, 15008–15014.
- 152 S. Azimi et al., On the involvement of copper binding to the N-terminus of the amyloid Beta Peptide of Alzheimer's disease: a computational study on model systems, *Int. J. Alzheimer's Dis.*

- 153 T. Dudev et al., Metal binding affinity and selectivity in metalloproteins: insights from computational studies, *Annu. Rev. Biophys.*, 2008, **37**, 97–116.
- 154 T. Dudev et al., Competition between protein ligands and cytoplasmic inorganic anions for the metal cation: a DFT/CDM study, *J. Am. Chem. Soc.*, 2006, **128**, 10541–10548.
- 155 T. Dudev et al., Competition between Li⁺ and Mg²⁺ in metalloproteins. Implications for lithium therapy, *J. Am. Chem. Soc.*, 2011, **133**, 9506–9515.
- 156 O. Coskuner-Weber, Revisiting Cu (II) Bound Amyloid-β40 and Amyloid-β42 Peptides: Varying Coordination Chemistries, *J. Turkish Chem. Soc. Sect. A Chem.*, 2018, **5**, 981–1008.
- 157 Y. Mantri et al., Computational study of the binding of Cull to Alzheimer's amyloid-β peptide: Do Aβ42 and Aβ40 bind copper in identical fashion?, *JBIC J. Biol. Inorg. Chem.*, 2008, **13**, 1197–1204.
- 158 J. Ali-Torres et al., Structures and stabilities of Fe^{2+/-3+} complexes relevant to Alzheimer's disease: an ab initio study, *J. Phys. Chem. A*, 2011, **115**, 12523–12530.
- 159 T. Miura et al., Binding of iron (III) to the single tyrosine residue of amyloid β-peptide probed by Raman spectroscopy, *J. Mol. Struct.*, 2001, **598**, 79–84.
- 160 R. C. Bernardi et al., Enhanced sampling techniques in molecular dynamics simulations of biological systems, *Biochim. Biophys. Acta - Gen. Subj.*, 2015, **1850**, 872–877.
- 161 C. Bannwarth et al., GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions, *J. Chem. Theory Comput.*, 2019, **15**, 1652–1671.
- 162 J. J. P. Stewart, Application of the PM6 method to modeling proteins, *J. Mol. Model.*, 2009, **15**, 765–805.
- 163 Y. Xia et al., Theoretical Study on Interactions of β-cyclodextrin with Trans-dichloro (dipyridine) platinum (II), *Comput. Theor. Chem.*, 2011, **967**, 213–218.
- 164 A. C. Bertoli et al., Theoretical and experimental investigation of complex structures citrate of zinc (II), *Inorganica Chim. Acta*, 2015, **425**, 164–168.

- 165 A. C. Bertoli et al., Structural determination of Cu and Fe–Citrate complexes: theoretical investigation and analysis by ESI-MS, *J. Inorg. Biochem.*, 2015, **144**, 31–37.
- 166 A. C. Bertoli et al., Theoretical spectroscopic studies and identification of metal-citrate (Cd and Pb) complexes by ESI-MS in aqueous solution, *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.*, 2015, **137**, 271–280.
- 167 K. M. Pietrusiewicz et al., Synthesis and use of borane and platinum (II) complexes of 3-diphenylphosphino-1-phenylphospholane (LuPhos), *Heteroat. Chem.*, 2011, **22**, 730–736.
- 168 G. Keglevich et al., Borane and platinum complexes of benzo-1, 3, 2-dioxaphospholane derivatives, *Lett. Org. Chem.*, 2010, **7**, 235–239.
- 169 N. Shahabadi et al., Synthesis, characterization and multi-spectroscopic DNA interaction studies of a new platinum complex containing the drug metformin, *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.*, 2014, **128**, 377–385.
- 170 N. Shahabadi et al., Study on the interaction of a copper (II) complex containing the artificial sweetener aspartame with human serum albumin, *Mol. Biol. Rep.*, 2014, **41**, 3271–3278.
- 171 D. Suárez et al., A combined semiempirical and DFT computational protocol for studying bioorganometallic complexes: Application to molybdocene–cysteine complexes, *J. Comput. Chem.*, 2014, **35**, 324–334.
- 172 D. Suarez et al., Molecular Modeling of Bioorganometallic Compounds: Thermodynamic Properties of Molybdocene–Glutathione Complexes and Mechanism of Peptide Hydrolysis, *ChemPhysChem*, 2015, **16**, 1646–1656.
- 173 C. L. Fernandes et al., Structural glycobiology of heparinase II from Pedobacter heparinus, *J. Biomol. Struct. Dyn.*, 2014, **32**, 1092–1102.
- 174 E. De Santis et al., Selective complexation of divalent cations by a cyclic α , β -peptoid hexamer: a spectroscopic and computational study, *Org. Biomol. Chem.*, 2016, **14**, 11371–11380.
- 175 N. A. M. S. Caturello et al., Influence of Metal, Ligand and Solvent on Supramolecular

Polymerizations with Transition-Metal Compounds: A Theoretical Study, *Chem. – A Eur. J.*, 2016, **22**, 17681–17689.

- 176 M. Turner et al., Modeling of Platinum-Aryl Interaction with Amyloid- β Peptide, *J. Chem. Theory Comput.*, 2016, **12**, 1385–1392.
- 177 S. T. Mutter et al., Benchmarking of copper(II) LFMM parameters for studying amyloid- β peptides, *J. Biomol. Struct. Dyn.*, 2018, **36**, 1145–1153.
- 178 S. Schmitz et al., Quantum Chemical Calculation of Molecular and Periodic Peptide and Protein Structures, *J. Phys. Chem. B*, 2020, **124**, 3636–3646.
- 179 K. Karami et al., Synthesis, spectroscopic characterization and in vitro cytotoxicities of new organometallic palladium complexes with biologically active β -diketones; Biological evaluation probing of the interaction mechanism with DNA/Protein and molecular docking, *J. Mol. Struct.*, 2018, **1154**, 480–495.
- 180 A. K. Rappé et al., UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations, *J. Am. Chem. Soc.*, 1992, **114**, 10024–10035.
- 181 J. Chan et al., Studies of metal binding reactions in metallothioneins by spectroscopic, molecular biology, and molecular modeling techniques, *Coord. Chem. Rev.*, 2002, **233**, 319–339.
- 182 J. Y. Xiang et al., An angular overlap model for Cu (II) Ion in the AMOEBA polarizable force field, *J. Chem. Theory Comput.*, 2014, **10**, 298–311.
- 183 F. Duarte et al., Force field independent metal parameters using a nonbonded dummy model, *J. Phys. Chem. B*, 2014, **118**, 4351–4362.
- 184 Q. Liao et al., Development and application of a nonbonded Cu²⁺ model that includes the Jahn–Teller effect, *J. Phys. Chem. Lett.*, 2015, **6**, 2657–2662.
- 185 P. D. Q. Huy et al., Impact of Cu(II) Binding on Structures and Dynamics of A β 42Monomer and Dimer: Molecular Dynamics Study, *ACS Chem. Neurosci.*, 2016, **7**, 1348–1363.
- 186 D. F. Raffa et al., Molecular dynamics study of the beta amyloid peptide of alzheimer's disease and its divalent copper complexes, *J. Phys. Chem. B*, 2007, **111**, 3789–3799.

- 187 V. J. Burton et al., Molecular Mechanics for Coordination Complexes: The Impact of Adding d-Electron Stabilization Energies, *J. Am. Chem. Soc.*, 1995, **117**, 8407–8415.
- 188 R. J. Deeth, A test of ligand field molecular mechanics as an efficient alternative to QM/MM for modelling metalloproteins: the structures of oxidised type I copper centres, *Chem. Commun.*, 2006, 2551–2553.
- 189 R. J. Deeth, Comprehensive Molecular Mechanics Model for Oxidized Type I Copper Proteins: Active Site Structures, Strain Energies, and Entatic Bulging, *Inorg. Chem.*, 2007, **46**, 4492–4503.
- 190 H.-C. Tai et al., Combined theoretical and computational study of interstrand DNA guanine–guanine cross-linking by trans-[Pt (pyridine) 2] derived from the photoactivated prodrug trans, trans, trans-[Pt (N3) 2 (OH) 2 (pyridine) 2], *Inorg. Chem.*, 2012, **51**, 6830–6841.
- 191 Subramaniam Boopathi et al., *Proteins Struct. Funct. Bioinforma.*, 2016, **84**, 1257–1274.
- 192 W. Li et al., Effects of zinc binding on the conformational distribution of the amyloid- β peptide based on molecular dynamics simulations, *J. Phys. Chem. B*, 2007, **111**, 13814–13821.
- 193 M. B. Peters et al., Structural survey of zinc-containing proteins and development of the zinc AMBER force field (ZAFF), *J. Chem. Theory Comput.*, 2010, **6**, 2935–2947.
- 194 P. Li et al., MCPB.py: A Python Based Metal Center Parameter Builder, *J. Chem. Inf. Model.*, 2016, **56**, 599–604.
- 195 P. Li et al., in *Structural Genomics: General Applications*, eds. Y. W. Chen and C.-P. B. Yiu, Springer US, New York, NY, 2021, pp. 257–275.
- 196 X. Li et al., Recent advances in hypervalent iodine(III)-catalyzed functionalization of alkenes, *Beilstein J. Org. Chem.* 14154, 2018, **14**, 1813–1825.
- 197 M. Turner et al., Molecular dynamics simulation on the effect of transition metal binding to the N-terminal fragment of amyloid- β , *J. Biomol. Struct. Dyn.*, 2019, **37**, 4590–4600.

2 Theory

2.1 Introduction

Computational methods are powerful tools which can be used for solving chemical problems, and they are able to reproduce experimental data accurately. Over the last seven decades, a range of computational chemistry methods have significantly aided in the exploration and modelling of numerous chemical systems, ranging from small molecules to more complex systems such as metalloproteins. Less complicated molecules can use a very high level of theory to model the structure and energies accurately, whereas for monitoring systems containing a large number of atoms, and their effects at an atomic resolution, an application of more approximate methods is required. Electronic structure calculations can provide complementary information to experiments, and are an effective means for the analysis of compounds containing transition metal ions,¹ and metal-biomolecule interactions.² However, modelling flexible peptides when bound to metal ions is challenging.

Accurate quantum mechanical methods are computationally expensive, even for those methods that are relatively efficient such as density functional theory (DFT). Molecular mechanics (MM) may be applied for various problems in biochemistry, such as for exploring conformational space,^{3–7} but in normal formulations are not well suited for treatment of d-orbital effects in transition metals.^{8–11} Combined QM/MM is effective for modelling metals in biomolecules,^{12–20} but the inclusion of QM methods can still compromise the computational expense. Therefore, alternative methods that reduce computational time with reliable accuracy are of interest. Ligand field molecular mechanics (LFMM) were introduced by Deeth et al. introducing explicit d-electron energy terms for transition metals to the standard MM expression.^{21,22} This has previously been used to study small^{23,24} and large metal-biomolecular systems such as metalloproteins,^{25,26} including a range of transition metals such as Cu and Pt.^{27–32} Recently, our group showed that LFMM is suitable for predicting geometries and exploring the conformational space of transition-metals such as Cu(II) and Pt(II) when bound

to amyloid- β peptide, although it was found this method fails to reproduce the relative energies predicted by DFT with the BHandH functional and a 6-31G(d) basis set.²⁷

Semiempirical methods recently developed by Grimme, termed GFN2-xTB,^{33,34} have been proposed for calculating molecular geometries, vibrational frequencies, and non-covalent interaction energies. They have been shown to sample geometries in much shorter time than DFT with high accuracy, are applicable to heavy elements up to radon ($Z=86$), and have been tested in the literature on large metal-complex systems.^{35–38} This high performance makes it potentially suitable for our systems of interest, *i.e.* metal-peptide complexes.

In this chapter, the general theoretical background of the computational chemistry approaches for computing the systems that used in this work will be outlined, involving quantum mechanics (QM), molecular mechanics (MM), and molecular dynamics (MD) approaches.

2.2 Quantum Mechanics

The history of quantum mechanics began in 1900 when Max Planck³⁹ claimed that the radiation emitted from black bodies is quantised. In the 20th century, many scientists added on the quantisation concept, making it applicable to many other aspects of physical and chemical theories as well as characterising the energy of light. The Rutherford-Bohr model of atom⁴⁰ is one of the best known examples that is based on Max Planck's resolution. The emission spectrum of the Hydrogen atom was accurately identified, proving that the energy levels of electrons are quantised, as indicated by Planck's theory. However, classical mechanics concepts disagree with the quantisation assumption where the levels of energy are not continuous. Consequently, a way of improving new mechanics to explain microscopic systems, unlike the classical one, was necessary. Wave mechanics that describe the electrons motion based are an alternative solution and are also based on a quantised phenomenon.⁴¹ Following that, quantum mechanics theory was developed and expanded, in order to describe the properties of chemical systems, alongside the advancements in computer science over the past four decades.

Quantum mechanics (QM)⁴² describes the probability of finding the position of particles in space over time. However, the probability of finding particles' position, does not come at the same certainty level as in general relativity, where the position in space and time can be accurately defined through the use of three spatial coordinates and any measure of time. The probability in QM is obtained from the wave function squared, $P(r, t)=\Psi^2(r, t)$, where the wave function value, Ψ , is given from solving the Schrödinger equation. By solving the Schrödinger equation, the properties of a chosen system which contains small particles, such as electrons and nucleons are determined. Electrons, due to their very tiny mass, have both wave and particle characteristics.⁴¹

2.2.1 Schrödinger equation

The time-independent form of the Schrödinger equation incorporates the second derivative with respect to the particles' position, to determine the properties of the system. The wavefunction, Ψ , is fundamental to quantum mechanics. Implementing the appropriate operators to Ψ , provide prediction values of a given physical observation. The Hamiltonian operator, \hat{H} , is the most frequent operator employed in quantum mechanics that applies to Ψ to find the total energy, E , of a chemical system. Equation 2.1 illustrates the time-independent Schrödinger equation which represents a launching point for ab initio methods, utilized in computational chemistry.

$$\hat{H}\Psi = E\Psi \quad (2.1)$$

In the time-independent Schrödinger equation, \hat{H} corresponds to the Hamiltonian operator, Ψ to the wave function, and E is the numerical value which refers to the total energy of the system. For general N -particles system, the Hamiltonian operator is the sum of the kinetic (T) and potential (V) energy with respect to space variables (r), and can be expressed with the equation illustrated below:

$$\hat{H} = T(r) + V(r) \quad (2.2)$$

By expansion of the Hamiltonian equation terms, the second-order fractional derivative equation with respect to all particle coordinates, may then be constructed:

$$\hat{H} = - \sum_{i=1}^N \frac{\hbar^2}{2m_i} \nabla_i^2 + \sum_{i \neq j}^N V_{ij} \quad (2.3)$$

Where, the Laplacian operator $\nabla_i^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$, and m is the mass of the particle i , V_{ij} is the potential energy between two particles (electron-electron, nucleus-nucleus, electron-nucleus) and N refers to number of the particles in the system.

Using equations based on the Schrödinger equation from QM, it is possible to solve for one electron systems. The electromagnetic interactions between nuclei-electron and electron-electron, within an atom and at a molecular level, are termed Coulomb interactions. The nucleus and electron coordinates, termed R and r respectively, are incorporated into the Schrödinger equation and rewritten as follows:

$$\hat{H}\Psi(R, r) = E\Psi(R, r) \quad (2.4)$$

For a general N -particle system, the Hamiltonian operator includes the kinetic (T) and potential (V) energy for all particles. Therefore, it involves five terms of kinetic and potential forces, contributing to the total energy of a system: the kinetic energy of the electrons, T_e , the kinetic energy of nuclei, T_n , the electron-electron repulsion, V_{ee} , the nuclear-nuclear repulsion, V_{nn} , and the electrostatic interaction between electrons and the nuclei, V_{en} (Eq. 2.5).

$$\hat{H} = T_e + T_n + V_{ee} + V_{nn} + V_{en} \quad (2.5)$$

Nevertheless, the Schrödinger equation cannot be solved for many-body quantum systems that comprise of more than one particle, as it becomes too complicated. Thus, numerous

approximations have since been proposed with the aim of solving the Schrödinger equation by carrying out a number of mathematical operations, and they will be briefly outlined in the next section.

2.2.2 Born–Oppenheimer approximation

The *Born–Oppenheimer* approximation was created to simplify the Schrödinger equation for systems consisting of more than two particles. The approximation is based on discounting the insignificant nuclear kinetic energy, when compared to the electrons' contribution. The principle is assuming the nuclei are stationary due to the electrons being very light and mobile. As a result of the approximation, the electronic wave function consequently provides a potential energy surface upon which the nuclei move. Therefore, the kinetic energy of the nuclei, T_n in the Hamiltonian operator is negligible,⁴³ and the operator then be written as;

$$\hat{H} = T_e + V_{ee} + V_{nn} + V_{en} \quad (2.6)$$

$$\hat{H} = -\frac{\hbar^2}{2m_e} \sum_i \nabla_i^2 + \sum_{i \neq j} \left(-\frac{Ze^2}{4\pi\epsilon_0 r_{ij}} \right)$$

Where $\hbar = \frac{h}{2\pi}$ is the reduced Planck constant, m is the mass of the particle, r_{ij} represents the distance between two particles, Z is nuclear charge.

By this approximation, the system can be described as the electrons moving around the nuclei, which remain stationary. The electronic energy considers a parametric hypersurface as a function of the nuclear coordinates, and the movement of the nuclei on this surface can then be solved.

2.2.3 Hartree-Fock (HF) Theory

Hartree-Fock (HF) theory is the foundation of most of the ab initio computational chemistry methods.³⁹ It is the approximate solution for the Schrödinger equation of a quantum many-body system, and implicitly includes the Born-Oppenheimer approximation.⁴³ The

approximation procedure is based on taking the average repulsion between the electrons into account, when determining the orbitals. In this model, the energy of one electron is determined by taking the average interaction with all the other electrons into account. Electrons are indistinguishable particles having a spin of 1/2. Their fermionic nature requires the total wave function to be antisymmetric, so a single Slater determinant may be used to determine the exact N -body energy and wave function of the system of n spin orbitals as shown in the following equation:

$$\psi(1,2, \dots, n) = \left(\frac{1}{\sqrt{n!}} \right) \begin{vmatrix} \varphi_1(1) & \varphi_2(1) & \dots & \varphi_n(1) \\ \varphi_1(2) & \varphi_2(2) & \dots & \varphi_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1(n) & \varphi_2(n) & \dots & \varphi_n(n) \end{vmatrix} \quad (2.7)$$

In Eq. 2.7, the number of electrons in the system and the spin orbitals are characterized by n and φ_n respectively. Through the application of the variational method on the HF theory, the minimum energy can be reached by modifying the coefficients of the atomic orbitals (AOs) – this is achieved through the implementation of approximate wavefunctions, until the energy requirements are satisfied.⁴⁴ However, the variational principle⁴⁵ states that the energy calculated for the system at all times is more than or equal to the actual energy of the ground state wavefunction. Consequently, the energy of the actual ground state is lower than the HF energy of a computed molecule.

The Hartree-Fock method features the application of mean-field theory, where all the individual interactions of many-body systems, are approximated by considering the average of the effective interaction. That means, the instantaneous effect of electrons that come close simultaneously at some point is not taken into account, but instead is averaged. The difference in energy between Hartree-Fock and the exact energy is termed the electron correlation energy, expressed as following,

$$E_{correlation} = E_{exact} - E_{HF}$$

Therefore, the HF energy is always higher than the exact energy. This results in the HF approximation being insufficient for computing the exact electronic state; especially for transition metal systems that contain large numbers of electrons and may have low-lying

excited states, leading to a large deviation from experimental results. The Hartree–Fock model accounts for *ca.* 99% of the total energy, but the remaining correlation energy can be critical for chemical purposes. A number of methods have been developed to ameliorate the HF weakness, with the aim of these associated methods to solve the Schrödinger equation, by computing the remaining electron correction for the electron–electron interactions to the multi-electron wave function. An example of those modified approaches, can be found in the Coupled Cluster methods (CC) such as CCSD and CCSD (T), as well as the Møller–Plesset perturbation theories (MP2, MP3, MP4), where correlation energy can be considered as a perturbation of the Fock operator. Other modified methods that increase the actual multi-electron wave function in terms of a linear combination of Slater determinants for example, multi-configurational self-consistent field (MCSCF). Even though, these methods provide improvement on HF approximation; they are computationally expensive, especially for systems that involve transition metals, which contain a large number of electrons, giving rise to a major correlation energy.

In addition, as the number of electrons increases, the wave function of those systems becomes significantly more complicated, as a result of the consideration of four variables for each electron (three spatial variables and one spin variable). Alternative to HF methods, the density functional theory (DFT) is used where both exchange and correlation energies of electrons, are considered, and the determination of the electron density is independent of the number of electrons.

2.2.4 Density Functional Theory

The density functional theory approach uses molecular electron density to obtain the energy, and hence other properties of molecules such as structure and spectroscopic parameters. In contrast to HF, which neglects electron correlation precisely, DFT method is favourable for large systems that contain many atoms and electrons, such as proteins and transition metal-containing systems. DFT is better in doing so, compared to HF, due to considering the correlation of electrons via the electron density of the system. The foundation of the method was first conceptualised in 1927 by Llewellyn H. Thomas and Enrico Fermi,^{46,47} where a perfect conformity between electron density of a molecule and the wavefunction of the many

electron molecule, was theorised. This discovery demonstrated a new and simpler way of determining the properties of molecules through physical observations (electron density), instead of the determination of properties, based solely on wavefunction. They were able to determine the energy of an atom by developing a kinetic-energy functional, which when combined with the classical terms permitting its representation in terms of electron density, such the nucleus–electron and electron–electron interactions. However, their theory failed to describe molecular orbitals and bonding, traced to poor representation of the kinetic energy.

Later in 1964, Hohenberg and Kohn⁴⁸ added more specifics and presented two theorems, which relate to any system made up of electrons moving below the influence of an external potential. Their first theorem states that the ground-state properties of a many-electron system are defined by an electron density, as a functional that is subject to three spatial coordinates. It places the foundation for reducing the many-body problem of N electrons with $3N$ spatial coordinates to three spatial coordinates, regardless of the number of electrons in system, as shown in the form that follows:

$$\text{electron density} = \rho(x, y, z) \quad (2.8)$$

This is different to other quantum mechanics methods that are based upon four variables $4N$ ($3N$ spatial and N spin direction) for each electron.⁴² The second theorem developed by Hohenberg and Kohn classifies an energy functional for the system and confirms that the ground-state electron density minimizes this energy functional.⁴⁸ Through the combination of these theorems, it is possible to calculate the ground-state wavefunction, which corresponds to the ground-state density, $\rho(r)$, making the wavefunction a unique functional of $\rho(r)$ and consequently the ground-state energy of system, $E[\rho]$, a functional of the ground state density, $\rho(r)$. Kohn and Sham later proposed their own theory of representing the mathematics of electron densities and their subsequent correlations to the energies of the molecule.⁴⁹ The DFT potential is constructed as the sum of external potentials, V_{ext} , which is determined solely by the structure and the elemental composition of the system, and an effective potential, V_{eff} , which represents interelectronic interactions, as expressed by the following expression:

$$E[\rho] = T[\rho] + V_{en}[\rho] + V_{ee}[\rho] \quad (2.9)$$

Where E represents the energy of the electronic ground state as a function of the density, T represents the kinetic energy of electrons, and V defines the external potential, which is based on the elements comprising the system, V_{en} is the nucleus-electron potential energy, and V_{ee} is electron-electron potential energy, a functional of ρ function. The term V_{ee} can be divided into two parts; the Coulomb energy (electron-electron repulsion), J, and the electron-electron exchange-correlation energy, E_{xc} :

$$E_{DFT}[\rho] = T[\rho] + V_{en}[\rho] + J[\rho] + E_{xc}[\rho] \quad (2.10)$$

Even though the nucleus-electron potential energy function is defined exactly, the kinetic energy function and the exchange function, are unknown and not so easily derived.

The modern DFT approximation form is based on the Kohn-Sham system,⁴⁹ where orbitals are utilized to define the electron density. This system involves a Schrödinger-like equation (one-electron), which for the non-interacting particles, such as electrons, can produce the uniform density for any defined system of interacting particles. Hence, the electrons with equal density are considered as non-interacting electrons in the Kohn-Sham system. The wavefunction, expressed using a single Slater determinant, is composed of a set of orbitals with the lowest-energy Kohn-Sham orbital (ϕ_i), occupied with N electrons.

$$T_s[\rho] = \sum_{i=1}^N \left\langle \phi_i \left| -\frac{1}{2} \nabla^2 \right| \phi_i \right\rangle \quad (2.11)$$

By using this equation, the exact kinetic energy functional can be calculated, based on the density. Yet, the density value is not known precisely, thus, the ground state electron density can be characterized by a set of one-electron spatial orbitals, as the following equation;

$$\rho(r) = \sum_{i=1}^N |\phi_i(r)|^2 \quad (2.12)$$

This approximation is regarded as one of most valuable methods in modern quantum chemistry, even though it does not give an absolute answer. The remaining kinetic energy that is coupled with the exchange function is added into the exchange-correlation function and the Kohn-Sham DFT energy can be stated as,

$$E[\rho] = T_s[\rho] + V_n[\rho] + J[\rho] + E_{xc}[\rho] \quad (2.13)$$

The solution to the Kohn-Sham equations in a self-consistent manner, can be obtained by substituting the orbital terms, ϕ_i , for a one electron basis, to yield the ideal set of orbitals and consequently the optimal electron density,

$$\left[-\frac{1}{2} \nabla^2 + V_{en}(r_1) + \int \frac{\rho(r_2)}{|r_1 - r_2|} dr_2 + V_{xc}(r_1) \right] \phi_i(r_1) = \epsilon_i \phi_i(r_1) \quad (2.14)$$

Here, the Kohn-Sham eigenvalue equation, ϵ_i , for one electron orbitals, ϕ_i , at the electron positions r_1 and r_2 , consists of the kinetic energy of each electron plus an effective potential, where the effective potential involves the potential electron-nuclei attraction, V_{en} , the repulsion between other electrons, and exchange correlation potential, V_{xc} . The challenge of the above equation is to locate the accurate functional form of the exchange-correlation energy, which in turn would provide the exact energy from DFT. On the other hand, the E_{xc} functional value is unknown and often, but not always, estimated by using semi-empirical methods. Therefore, numerous methods have been developed to improve the calculation of the exchange-correlation functional, which will be outlined in the next sections.

2.2.4.1 The Local Density Approximation

The local density approximation (LDA)⁵⁰ aims to accurately approximate the exchange-correlation energy value, E_{xc} , by assuming an exchange-correlation density, $\epsilon_{xc}(n)$, identical to that of a homogeneous electron gas. The approach is based on the assumption that the

exchange-correlation energy in terms of homogeneous electron gas, is constructed on the density $\rho(r)$ at a particular point in the system:

$$E_{xc}^{LDA} = \int \rho(r) \varepsilon_{xc}(\rho(r)) dr \quad (2.15)$$

Where $\varepsilon_{xc}(\rho(r))$ is the exchange-correlation energy per particle of a homogeneous gas, coupled with an electron in a uniform electron gas of density $\rho(r)$. The exchange energy of a uniform electron gas is known analytically and presented by the Dirac equation:

$$\varepsilon_x^{LDA} = -\frac{3}{4} \sqrt{\frac{3\rho(r)}{\pi}} \quad (2.16)$$

However, the term for the correlation density, ε_c^{LDA} is approximate formula and the accurate values have been defined from quantum Monte Carlo⁵¹ (QMC) calculations. A commonly used correlation formula are Vosko, Wilk, and Nusair (VWN),⁵² and Perdew and Zunger(PZ).⁵³

However, the key idea of this approximation is that, for a molecule with many electrons in a gaseous state, the density changes slowly over the molecule. This is not true for molecules where the electron density is not uniform. In addition, the α and β electron spin densities are not equal, so LDA has been unrestricted and replaced by the Local spin density approximation (LSDA) which modifies LDA to include spin-polarised systems. The LDA approximation can provide reasonable results in terms of geometries and vibrational frequencies calculation. Nevertheless, it produces significant errors in energies as a result of the deficiencies in the exchange and correlation energy estimations; although this can be resolved by applying the generalized gradient approximation.

2.2.4.2 The Generalised Gradient Approximation

The generalized gradient approximation (GGA), accounts for the non-uniformity of the electron density leading to a significant improvement on the LDA and LSDA, through the combination of the electron density calculations with a gradient correction factor. In the GGA approximation the exchange and correlation terms are considered independently. Hence, the gradient-corrected exchange energy can be expressed as,

$$E_x^{GGA} = E_x^{LDA} - \sum_{\sigma} \int F(s_{\sigma}) \rho_{\sigma}^{4/3}(r) dr \quad (2.17)$$

F in equation 2.17 corresponds to the exchange functional, consisting of a range of functional terms, including empirical parameters, such as Becke's functional.⁵⁴ The s_{σ} term in the equation signifies the reduced density gradient, and can be expressed according to the equation below;

$$S_{\sigma} = \frac{|\nabla \rho_{\sigma}(r)|}{\rho_{\sigma}^{4/3}(r)} \quad (2.18)$$

Despite the improvements in the DFT methods, after considering the inclusion of GGA approximation, resulting in a significant increase in computational accuracy without the additional increase in computing time; the incorporation of some of the more suitable features from *ab initio* methods (specifically Hartree-Fock) to the exchange-correlation energy functional in DFT, can improve the calculations from DFT, even further. Such kind of approaches are known as Hybrid methods.

2.2.4.3 Hybrid DFT

Hybrid functional approaches involve approximations imposed onto the exchange-correlation energy functional in density functional theory. The methods are a combination of exact energy terms from Hartree-Fock and the exchange-correlation functional from local spin density approximation (LSDA) or GGA methods. These hybrid methods offer a simple path for accurately describing molecular properties, in which they are effectively calculated with standard functionals. Hybrid methods, such as B3LYP, are currently the most common and prevalent DFT methods used in practice.

2.2.4.3.1 B3LYP

B3LYP is a hybrid functional that is oftentimes considered to be a standard model in chemistry for many applications, as presented by Becke,⁵⁵ and Lee, Yang and Parr.⁵⁶ This functional

contains a combination of various approximations; local density approximation (LDA), Hartree-Fock (HF), Becke-1988 (B88)⁵⁴ exchange energy, Lee-Yang-Parr 1988 (LYP88)⁵⁶ correlation energy), and Vosko, Wilks, Nusair 1980 (VWN80).⁵² The B3LYP functional has the following formula;

$$E_{xc} = (1 - a_0)E_x(LDA) + a_0 E_x(HF) + a_x E_x(B88_x) + a_c E_c(LYP88_c) + (1 - a_c) E_c(VWN80_c) \quad (2.19)$$

Where, (a_0 , a_x and a_c) are three parameters that are fitted to experimental data as $a_0 = 0.20$, $a_x = 0.72$, and $a_c = 0.81$. The lowercase x represents the electron exchange determination, while the lowercase c signifies the electron correlation determination. The Becke88 exchange functional and the LYP correlation functional are generalised gradient approximations, while LDA is correlation functional type.

One of the main drawbacks of DFT methods is the challenge in selecting the most suitable model for a particular problem without performing calculations. In addition, dispersion is not constructed in conventional Kohn-Sham DFT, leading to poor description of dispersion forces like van der Waals interactions.⁵⁷ But modern DFT methods provide possible solutions such as modifying the exchange-correlation functional or including an explicit dispersion correction, well-known as DFT-D, as proposed by Grimme.⁵⁸

2.2.4.4 Basis Set

An approximation to solving the Schrödinger equation that fundamental in all *ab initio* methods is the presentation of a basis set. A basis set is a set of one electron functions, typically in the form of atomic orbitals, centred on the atomic nuclei within molecules and combined linearly to model molecular orbitals. For more accurate electronic structure calculations, extended basis sets are needed to be able to extrapolate reliably to a complete basis set (CBS) limit. The most common types of atomic orbitals are Gaussian-type orbitals (GTOs) and Slater-type orbitals (STOs). The more accurate of the two are STOs as they are solutions to the Schrödinger equation of hydrogen-like atoms also known as hydrogenic orbitals. These have the correct nuclear cusp and long-range behaviour, that decay occurs

(e^{-r}) , means they are sharp near the atomic nucleus at $r=0$ on a radial dependence graph, but are more computationally expensive. On the other hand, GTOs are less accurate where this method decay like $(e^{-\alpha r^2})$ and does not acquire the cusp of the wavefunction nor the exponential decay, which means they are flat near the atomic nucleus at $r=0$, however, they are less computationally expensive. Increasing GTOs accuracy is achieved by extending the number of atomic orbital functions used; instead of one large, instead basis function for each occupied orbital in the system, known as minimal basis set.

At minimal basis sets, only a single set of the valence functions, named single- ζ (SZ) basis, are included within the calculation. For example, a carbon atom has 2 s-functions and a single set of p-functions. However, upon inclusion of one more to a single set of valence functions, double- ζ (DZ), the carbon atom functions increase to two s-functions and 2 sets of p-functions. Adding onto the same idea for triple- ζ (TZ) basis set, the carbon would have 3 s-functions and 3 sets of p-functions and one more for s and p functions at quadruple- ζ . Some commonly used basis sets, based on the SZ, DZ and TZ basis sets, are STO-3G, 3-21G, 6-31G, D95V, and 6-311G, TZV. However, QZ basis sets are not commonly used due to the basis being that large. Instead, polarization functions are more appropriate.

Polarization functions implement higher angular momentum functions to AO basis, through the addition of p-functions on hydrogen, d-functions on first-row elements, and f-functions on metals and so on. Therefore, carbon will have an additional angular momentum number (d orbital) $3s^2p^1d$. The general labels for polarization functions is to add letter P to the standard zeta basis sets (DZ, TZ, QZ) labels, thus they are termed DZP, TZP, QZP and mostly denoted by an asterisk (*). The most well-known basis set names for polarization functions are 6-31G(d,p), 6-311G(d,p), and TZVP. Another gaussian function $(e^{-\alpha r^2})$, used for anions is the diffuse function, as the additional electron in anions is weakly bound, and consequently spend more time in a distance from the nucleus. The value of α is extremely small, leading to the atomic orbital (basis function), to decay slowly and then define the extra electron. The annotation "+" or "aug-" is added to a basis set when diffuse functions have been included, for example 3-21++G, 6-31+G(d,p), aug-TZP. However, adding more complex basis functions, requires more computational time and resources than their simpler equivalents.

For more complicated atoms, such as transition metals, an effective core potential (ECP) approximation is commonly used. An ECP eliminates the core electrons from the problem, substituting them with an effective potential for the valence electrons. Thus, basis sets for this type of heavy atoms neglect the core electrons from the basis functions, and describe the valence electrons alone. Further improvement can be obtained through the inclusion and mixing of different levels of standard zeta, polarisation and diffuse basis sets.

In this work, a split valence basis set, 6-31G*, was utilized. It is a typical basis set that uses six contracted Gaussian functions for the core electrons, a double- ζ basis set of three Gaussians and one Gaussian for the valence electrons, in addition to polarisation functions implemented through the addition of d-orbitals on all non-hydrogen light atoms (C, N, O) and f-functions on transition metals.

In general, *ab-initio* and DFT methods that mainly aim to calculate electronic energies, as well as further physical properties, as functions of the positions of the nuclei, come with limitations. These require intensive computation and hence tend to be restricted, to smaller molecules, radicals, and ions, for more reasonable computation times. Alternatively, faster approaches, such as semi empirical methods, based on HF and the implementation of parametrized values, obtained from experimental data or *ab initio*, are more convenient in certain scenarios.

2.3 Semiempirical quantum mechanical methods SQM

Semiempirical methods are built on the HF method, with parameters from experimental data used to simplify calculations, which include electron correlation effects neglected in HF. In semiempirical methods, only valence electrons are described using a minimal basis set while the core electrons are omitted from the calculations. The excluded quantities are replaced by pairing the results to experimental data or *ab initio* calculations. The methods can simulate large molecules, such as proteins with less computational demands than *ab initio* methods, due to the approximation known as zero differential overlap (ZDO),⁴² which is based on discounting certain integrals like two-electron repulsion integrals. Thus, a significant computational cost has been reduced when used less computational complexity in semiempirical methods scales as Order (N)² than that of formal Hartree Fock-method, which

it scales as Order(N)⁴, where N is the size of basis set. As a result, the calculation is simplified and the calculations are much faster than the equivalent *ab initio* methods. There are a number of semiempirical methods based on classical ZDO approximations, such as Intermediate Neglect of Differential Overlap (INDO), Complete Neglect of Differential Overlap (CNDO), Modified Intermediate Neglect of Differential Overlap (MINDO), and the most recent Neglect of Diatomic Differential Overlap (NDDO).⁴² The most frequently used semiempirical methods, for example Austin Model (AM1),⁵⁹ and Parametric Model number (PM3,⁶⁰ PM6,⁶¹ and PM7⁶²) belong to the NDDO approximation.

One development type of SQM, is the density functional-based tight binding (DFTB) methods. These are based on Kohn and Sham DFT, where the energy is extended in terms of density fluctuations, $\delta\rho$, relative to a linear combination of atomic orbital LCAO densities. The methods belong to the self-consistent charge-type approaches, in conjunction with first, second, or third-order charge fluctuation terms, where the highest-order variant is (DFTB3).⁶³ Unlike standard semiempirical tight binding approach, the parameterization is less empirical and more fundamental, as it is essentially obtained by a number of DFT calculations, making it a computationally efficient approach. The methods bring together the efficiency of the traditional ZDO-form minimal basis set methods with the higher accuracies of DFT. The extended tight binding methods (xTB) such as GFN n -xTB methods ($n = 0, 1, 2$) are developed through the derivative of DFTB3, in which the parameterization procedure is forced to describe chemical energies with high accuracy and covering most of the periodic table atoms, including the most chemically important transition metals.

2.3.1 GFN2-xTB

The GFN2-xTB approach is a fast and computationally affordable tight-binding semiempirical method, developed by Stefan Grimme and co-workers,³⁴ with iterative improvement on the previous GFN method.³³ It was built mainly for targeting the geometry, noncovalent interaction, and vibrational frequency calculations for large molecular systems, consisting of about 1000 atoms. In addition, this approach is applicable to organic, organometallic, and more complicated biomolecular systems. It is a tight binding model, but with inclusion of electrostatic interaction and exchange-correlation energy terms, up to second order density

fluctuation of atomic multipole moment approximations. Like previous GFN methods, the global and element specific parameters are strictly included within this method up to Z=86 of the elements. Furthermore, it can be used as a benchmark for a range of systems and in conjunction with other semiempirical methods, due to their high accuracy. Within this method, full-scale analytical gradients, such as nuclear forces, are performed. Alongside the high accuracy observed for calculations of targeted properties, the method benefits from excellent performance with few errors, as well as other properties, such as barrier heights and molecular dipole moments. Uniquely, it can be ideal for biomolecular systems in aqueous solution, after considerable refinements for multiple benchmark sets. As such, it can be used successfully in exploring the conformational space of molecules. The main disadvantage of both GFN methods is their restricted accuracy on thermochemical properties.

The total energy calculated in GFN2-xTB is expressed in following terms:

$$E_{GFN2-xTB} = E_{rep} + E_{dis} + E_{EHT} + E_{IES+IXC} + E_{AES} + E_{AXC} + E_{Fermi} \quad (2.20)$$

Where the abbreviations in subscripts represent the energies of (rep) the repulsion energy, (dis) the dispersion energy, (EHT) extended Hückel-Type energy, (IES) the isotropic electrostatic, (IXC) isotropic exchange-correlation, (AES) the anisotropic electrostatic, (AXC) the anisotropic exchange-correlation energy, (Fermi) the entropic contribution of an electronic free energy at finite electronic temperature (T_{el}) due to Fermi smearing. As a result of the suitable features of the GFN2-xTB method for large biological systems, as well as for those involving metal ions in their structures, this method is applied in this study.

Even though the quantum mechanical approaches provide an accurate and acceptable calculations for a wide range of systems, they are computationally demanding for dynamics simulations of large size, flexible, and bioinorganic complex systems such as metal-proteins. Thus, molecular mechanic/molecular dynamics (MM/MD) simulations offer an alternative approach for the evaluation of such systems. The next section will outline a theoretical background of MM/MD methods used in this work.

2.4 Molecular Mechanics (MM)

Beside the QM methods, Molecular Mechanics (MM) is another method that can be used for molecular modelling. MM apply classical mechanics, particularly Newton's second law. In this, the molecular energy is calculated as a function of nuclear positions and the atoms and bonds of molecules are described using the ball and spring model, that follows simple mathematical rules such as Hooke's law.⁴² At the atomic level, the interaction between particles can be defined in terms of either force (F) or potential (E) energies. Therefore, the force field comprises of the interactions applied in MM methods, for instance, stretch, bend, torsional, van der Waals, etc., interactions. The total energies gained from these inter-atomic potentials forces, are presented in Eq 2.21,

$$E_{tot} = E_{stretch} + E_{bend} + E_{torsion} + E_{VDW} + E_{electrostatic} \quad (2.21)$$

$E_{stretch}$ represents the energy function for bond stretching in the system, E_{bend} is for the energy needed to bend an angle, $E_{torsion}$ is the energy required for rotation around a bond, E_{VDW} and $E_{electrostatic}$ corresponds to the non-bonding interactions; van der Waals and electrostatic. Each energy term in the above equation, may be calculated, with the derived equation as shown below;

$$E_{tot} = \frac{1}{2} \sum_{bonds} k_i (l_i - l_{i,0})^2 + \frac{1}{2} \sum_{angles} k_i (\theta - \theta_{i,0})^2 + \frac{1}{2} \sum_{torsions} V_n (1 + \cos(n\phi - \gamma)) + \sum_{i < j}^{atoms} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \sum_{i < j}^{atoms} \frac{q_i q_j}{\epsilon r_{ij}} \quad (2.22)$$

The equation above is for the total energy of a system according to MM methods, and can be divided into four summed terms of (E_{bonds} , E_{angles} , $E_{torsions}$, and $E_{non-bonding}$); the first term corresponds to the sum of the atom-atom bond energy, using a harmonic potential, where l_i , signifies the bond length increase from the original length, $l_{i,0}$, resulting to the potential energy. The second term represents a total intramolecular bond angle that is also modelled by a harmonic potential where, k is refers to force constant. The third term is for a periodic

torsional potential, ϕ , applied to represent the energy revolution involved in bond rotation, where V_n , is the barrier to free rotation for the equilibrated bond, n is periodicity of the rotation and γ is the angle where the potential passes through its minimum value. The final two terms refer to non-bonding interactions (van der Waals and electrostatic interactions), that occur between atoms or molecules (i and j), which are defined according to the Lennard-Jones potential (LJ). In the LJ equation, r_{ij} symbolises the distance between two interacting particles, σ_{ij} refers to the point where the distance between two particles present zero potential energy, $V = 0$. The LJ potential illustrates that attractive forces decrease by r^{-6} , while repulsive forces increase by r^{-12} , where r is the inter-atomic distance. The electrostatic interactions term are expressed, based on Coulombic potentials, where the energy depends on r^{-1} , in this term q_i and q_j are the point charges on the atoms, r_{ij} is interatomic distance and ϵ is the dielectric constant. The values of these parameters can be obtained either experimentally, from spectroscopic analysis, such as infrared spectrum (IR) and nuclear magnetic resonance (NMR), or calculated computationally by QM methods.

The major advantage of MM methods is the ability to simulate large systems that involve thousands of atoms such as proteins and DNA with significantly lower computational expense compared to QM methods. Such methods can provide an excellent prediction of geometries and relative energies, of molecules containing many atoms, in a short time, if accurate parameters are available. The main challenge of using MM methods is the shortage of parameters, obtained from prior crystallised molecules or modelled with high level QM. Therefore, MM methods are helpful for predicting properties for groups of molecules that contain many atoms, only if enough of this information already exists. However, for molecules that lack such parameters, their modelling is limited by MM methods. Regardless of the limitations that come with MM methods, it is still commonly applied in computational chemistry, especially for the modelling biomolecular macromolecules.

There is an assortment of different force fields (FFs), with parameters implemented within programmes that are offered for modelling chemical systems, such as AMBER,⁶⁴ MMFF,^{65–67} CHARMM,⁴ GROMOS, OPLS,⁶⁸ and UFF.⁶⁹ However, modelling transition metals is complicated and requires a large number of parameters to achieve accurate FF parameterisation. Standard

MM methods cannot take the effects of d-orbital electrons on complex molecules into account, as the electrons are added implicitly in the calculations. Another challenge is that the main MM equation, as shown in equation 2.21, describes a single reference angle for a given ligand-metal-ligand (L-M-L) atom set for most of the normal organic molecules, such as tetrahedral, trigonal planar, or linear. However, most of the metal complexes involve multiple reference angles, according to their geometry, for example a homoleptic octahedral complex would require two unique L-M-L reference angles (90° , 180°), while trigonal bipyramidal geometries require three different reference angles (90° , 120° , 180°).

Deeth et al.⁷⁰ developed an approach that employs the ligand-ligand repulsion method, equivalent to points on a sphere (POS) for the angle bending of the coordinated compounds, as proposed by Comba et al.⁷¹ As a result, the reference θ_0 variables are not necessary. Deeth et al. incorporated this implementation of the ligand field molecular mechanics (LFMM)⁷⁰ model within the molecular operating environment (MOE), resulting to the extended version of the MOE software, DommiMOE.⁷⁰ MOE is a software package comprising of a selection of force fields, specially designed for the modelling of biomolecular molecules and drug discovery. The program can recognize the angular geometries around the metal centres precisely, and provides an accurate value for the strain energy of various transition metal complexes. Therefore, the LFMM method has been used in this work, as Zn^{+2} and Cu^{+2} were a part of the simulations.

2.4.1 Ligand Field Molecular Mechanics (LFMM)

The ligand field molecular mechanics (LFMM) method has been developed as solution to the d-orbital effect challenges. The LFMM model includes the ligand field stabilization energy (LFSE) term, to the potential energy equation of conventional MM, as shown in the following expression;

$$E_{LFMMM} = E_{stretch} + E_{bend} + E_{torsion} + E_{nb(VDW,el)} + LFSE \quad (2.23)$$

The LFSE model⁷² (also known as crystal field theory, CFT)^{73,74} is based on ligand field theory (LFT), which is considered the simplest general model to describe metal-ligand bonding. LFSE

calculations are a useful approach to gain information about d-orbital splitting diagrams and how the arrangement of d-electrons in high- and low-spin states, provide the overall LFSE energy value. Therefore, the term is used within the total energy expression, to determine d-electron effects, which have an important influence on the reactivity and geometry arrangement of transition metal complexes. The calculation of LFSE, starts by accounting for the difference between two energy levels that correspond to d-electron orbitals. In an octahedral environment, the d-orbitals consist of three lower energy level orbitals, labelled t_{2g} and two higher energy level orbitals, labelled e_g . While the average energy level is $2/5^{\text{th}}$ s of the way up from the lower energy level and $3/5^{\text{th}}$ s of the way down from the higher energy level, and it has a relative energy of zero. The difference value, between the higher and lower level, is the field splitting (Δ_0), where Δ stands for the energy difference between two levels and o signifies the octahedral coordination mode of the complex; the orbital diagram is shown in Figure 2.1.

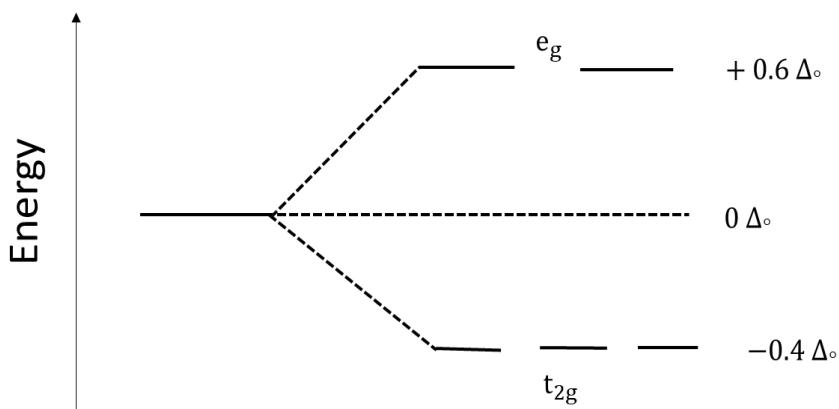


Figure 2.1: The orbital diagram according to LFST theory for octahedral coordination where barycentre, Δ_0 is located between lower and higher energy levels.

Subsequently, the LFSE value can be obtained according to the following expression;

$$LFSE = [(0.6 * \text{number of electrons in } e_g) - (0.4 * \text{number of electrons in } t_{2g})]\Delta_0 \quad (2.24)$$

The energy value of field splitting between the d-orbitals, Δ_0 , can be found from spectroscopic measurements, according to the wavenumbers obtained from the frequency or wavelength of light passing through a sample of transition metal complexes. There are some factors that influence the Δ_0 value, such as the type of metal ions, ligand identity, and repulsive energy

of pairing electrons. LFSE can accurately reproduce the experimental data for transition metal systems through the use of computational methods. QM methods can calculate d-electron effects correctly, as they are handled implicitly; while in MM methods, the LFSE values are added explicitly. Thus, Deeth et al.⁷⁰ applied the LFSE within the angular overlap model (AOM) expression in the conventional MM, instead of CFT, generating the ligand field molecular mechanics (LFMM) method.

The AOM method,⁷⁵ is used to describe the influence of transition metal–ligand interactions, by assuming the amount of overlap between the atomic orbitals of two atoms, is controlling the strength of a bond formed. The model also supposes that the sum of ligand field potential, V_{LF} , is the sum of the M-L bonds involved in the complex. The energy of each M-L bond parametrized in AOM, can be acquired from spectroscopic data of the d-d separation or theoretical analysis, defining the difference between σ and π interactions. The advantage of coupling AOM with LFMM, is the physically realistic explanation of the M-L bond. The AOM model is better than CFT, as it is able to characterize all the existing symmetry behaviour, and treat each ligand separately. For instance, the splitting of orbitals in an octahedral coordination, O_h , which has high-symmetry geometry between AOM and CFT, is different in terms of the barycentre position between orbitals. The splitting of AOM is expressed by equation 2.25 and the orbital diagram is shown in Figure 2.2.

$$\Delta_{oct} = 3_{e\sigma} - 4_{e\pi} \quad (2.25)$$

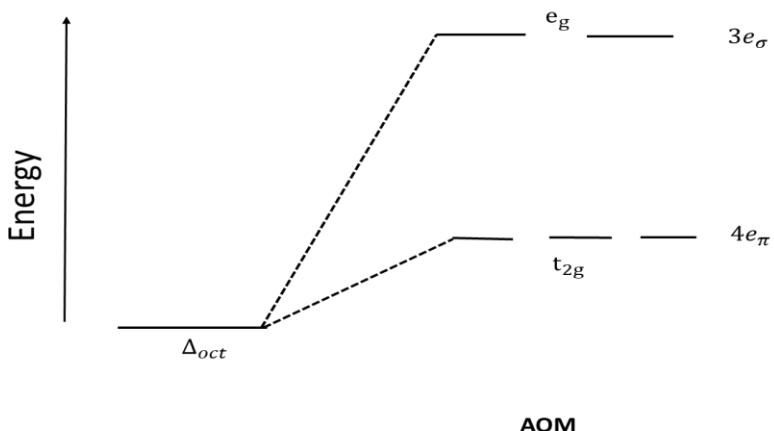


Figure 2.2: AOM barycentre in situation of a π -donor ligand showing e_π is positive (higher than barycentre, Δ_{oct}).

Moreover, the implementation of LFMM to standard MM methods demands additional parameters appropriate to transition metal complexes. Conventional MM is reliable for computing the organic section's energy, while the LFSE for the metal centre part is determined via the AOM approach. In addition to the inclusion of the LFSE term, for the computation of d-electron effects, the M-L stretching and L-M-L angle bending parameters must be included in the force field. Thus, Deeth et al. utilized the Morse function and the ligand-ligand POS terms, according to Comba et al., to describe the M-L stretching and L-M-L angle bending, respectively. In the case of the LFMM method, however, the LFSE part is incorporated directly into the MM calculations to determine the structure and energy of a system.

2.4.2 Molecular Dynamics (MD)

The equation used in molecular dynamics, defines the time progression of a system via the solution of Newton's equations of motion, equation 2.26. The motion of the atoms of mass, m , and under the force, F , is determined with Newton's law as the second order differential equation when considering the F depends on time. The integration of the equation in MD by discrete step in time using numerical methods such as Verlet algorithm to predict the new atom positions and velocities at the end of the step. Therefore, by applying the MD equation on chemical systems, the position and velocity of molecules can be predicted, at times, either before or after the relative starting point, such as the coordinates at certain time or

conformational changes. Therefore, studying the dynamical behaviour of large systems through the use of molecular dynamic (MD) method, is useful and appropriate.

$$F_i = m_i \frac{d^2 r_i(t)}{dt^2} \quad (2.26)$$

where $r_i(t) = (x_i(t), y_i(t), z_i(t))$ correspond the position vector of particle i, F_i is the force acting on particle i, m_i corresponds the mass of the particle, and t corresponds the time.

In MD, the mobility of a system with N atoms, is defined by the Newton's equations.⁴² Molecular mechanics methods are used to calculate the forces between the particles and their potential energies. The MD modelling procedure involves setting up with specific thermodynamic conditions, including statistical ensembles with constant temperature(T), pressure(P), volume(V), or energy (E). For instance, the microcanonical (NVE) ensemble samples the phase space with constant energy, particle number and volume. Nevertheless, the use of parameter sets from FF methods within MD simulations, can limit the accuracy of the method. An additional disadvantage to MD, is the difficulty to overcome barriers separating the minima, which are higher than the internal energy, defined by the simulation temperature during the conformational sampling. In principle, MD methods can sample the entire energy surface if long enough simulation times is given, but such long simulations cannot typically be reached for realistic systems. This problem can be reduced by increasing the temperature, but only the local area nearby the starting point can be sampled then. Even so, conventional MD is still an efficient method when it comes to simulating the dynamical behaviour of biomolecular systems.

2.4.3 Accelerated Molecular Dynamics

Accelerated molecular dynamics (aMD)⁷⁶ is an enhanced sampling algorithm, proposed to solve the timescale difference between classic MD methods and biological processes. The idea of this technique is to boost the conformational sampling efficiency, through the addition of a bias to the potential energy of the system leading to lower the potential. As a result, the transition between different states of the system is speeded up, so that biological systems which are trapped between two states in the potential energy landscape are released, permitting the conformational exploration to be more flexible and accessible than normal

MD.^{77,78} There are numerous studies which have implemented the aMD method to their systems, with reported success.⁷⁹⁻⁸³ Dual-boost version⁷⁷ of aMD is chosen for the acceleration of biomolecules, through the use of both a total as well as dihedral boost potential, for all atoms in the system according to the following equations:

$$V^*(r) = V(r) + \Delta V(r) \quad \text{when the system potential is } V(r) < E \quad (2.27)$$

$$\Delta V(r) = \frac{(E - V(r))^2}{\alpha + E - V(r)} \quad (2.28)$$

Where $V^*(r)$ is the modified potential, $V(r)$ is the original potential, $\Delta V(r)$ is the boost potential, α is the acceleration factor, and E is the threshold energy. The dual-boost version of aMD applies the boost potential to all atoms in the system, in addition to all the dihedrals, through the input parameters (E_{dihed} , α_{dihed} , E_{total} , α_{total}), while in the dihedral-boost version, a boost potential is applied to all dihedrals in the system alone (E_{dihed} , α_{dihed}). Therefore, the parameters which are applied to accelerated MD simulation, are given as follows:⁸⁴

$$E_{\text{dihed}} = V_{\text{dihed-avg}} + 3.5N_{\text{res}} \quad , \text{where } \alpha_{\text{dihed}} = \frac{3.5N_{\text{res}}}{5} \quad (2.29)$$

$$E_{\text{total}} = V_{\text{total-avg}} + 0.175N_{\text{atoms}} \quad , \text{where } \alpha_{\text{total}} = 0.175N_{\text{atoms}} \quad (2.30)$$

Where, N_{res} is the number of residues of protein, N_{atoms} is the total number of atoms, $V_{\text{dihed-avg}}$ and $V_{\text{total-avg}}$ are the dihedrals and total potential energy averages, generated from short conventional MD simulations.

2.5 Solvent Models

It is important to model solvation effects of biomolecular systems and this is possible through the use of solvation models. There are two main types of solvation models which differ by the way they handle solvation around a solute. The implicit models, where the solvent is considered as continuous medium (also known as continuum solvation models), are characterized by a dielectric constant, ϵ , and the solute is placed in the medium as a hole, while the second type involves explicitly modelling of water molecules around the solute molecules.⁴² The explicit solvent models usually give more accurate hydration free energy

(HFE) predictions, but require substantially higher computational resources, than implicit. In contrast, continuum models approximate the solvent structure in systems, deriving the thermodynamic parameters of solvation, without significant computational costs, in the expense of accuracy. The placing of solute as hole (cavity) in the polarized homogeneous medium with a dielectric constant, results in the release of free energy due to dispersion interaction between the solute and the solvent and repulsion interaction between the components, as well as the electric charge distribution of the solute. Thus, the solvation free energy of the system is the sum of the changes in free energy terms:

$$\Delta G_{solvation} = \Delta G_{cavity} + \Delta G_{dispersion} + \Delta G_{elec} \quad (2.31)$$

In this work, three solvent models for each specific methods have been used. For DFT calculations, the polarisable continuum model (PCM) has been applied, in which the cavity is modelled using a van der Waals surface, and a generalized Born-surface area (GB-SA) model used for semi-empirical calculation. The explicit solvent model, TIP3P,⁸⁵ was also used in Chapter 5.

2.6 References

- 1 R. Peverati et al., M11-L: A Local Density Functional That Provides Improved Accuracy for Electronic Structure Calculations in Chemistry and Physics, *J. Phys. Chem. Lett.*, 2011, **3**, 117–124.
- 2 S. Dasari et al., Cisplatin in cancer therapy: Molecular mechanisms of action, *Eur. J. Pharmacol.*, 2014, **740**, 364–378.
- 3 E. G. Atavin et al., Molecular structure of cyclododecane, C₁₂H₂₄, as determined by electron diffraction and molecular mechanics, *J. Mol. Struct.*, 1989, **212**, 87–95.
- 4 B. R. Brooks et al., CHARMM: A program for macromolecular energy, minimization, and dynamics calculations, *J. Comput. Chem.*, 1983, **4**, 187–217.
- 5 P. K. Weiner et al., AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions, *J. Comput. Chem.*, 1981, **2**, 287–303.
- 6 S. J. Weiner et al., An all atom force field for simulations of proteins and nucleic acids, *J. Comput. Chem.*, 1986, **7**, 230–252.
- 7 S. L. Mayo et al., DREIDING: a generic force field for molecular simulations, *J. Phys. Chem.*, 1990, **94**, 8897–8909.
- 8 V. S. Allured et al., SHAPES empirical force field: new treatment of angular potentials and its application to square-planar transition-metal complexes, *J. Am. Chem. Soc.*, 2002, **113**, 1–12.
- 9 R. J. Deeth, in *Coordination Chemistry*, Springer Berlin Heidelberg, Berlin, Heidelberg, 1995, pp. 1–42.
- 10 H. M. Marques et al., in *Structure and Function*, ed. P. Comba, Springer Netherlands, Dordrecht, 2010, pp. 87–109.
- 11 F. Wiesemann et al., Force Field Calculations on the Structures of Transition Metal Complexes. 1. Application to Copper(II) Complexes in Square-Planar Coordination, *Inorg. Chem.*, 1994, **33**, 1891–1898.
- 12 K. Gkionis et al., QM/MM investigation into binding of square-planar platinum complexes to DNA fragments, *J. Biol. Inorg. Chem.*, 2009, **14**, 1165–1174.
- 13 K. Spiegel et al., Cisplatin Binding to DNA Oligomers from Hybrid Car-

- Parrinello/Molecular Dynamics Simulations, *J. Phys. Chem. B*, 2004, **108**, 2699–2707.
- 14 A. Robertazzi et al., A QM/MM Study of Cisplatin–DNA Oligonucleotides: From Simple Models to Realistic Systems, *Chem. – A Eur. J.*, 2006, **12**, 5747–5756.
- 15 T. Matsui et al., Multiple Proton-Transfer Reactions in DNA Base Pairs by Coordination of Pt Complex, *J. Phys. Chem. B*, 2007, **111**, 1176–1181.
- 16 K. Spiegel et al., Modeling anticancer drug–DNA interactions via mixed QM/MM molecular dynamics simulations, *Org. Biomol. Chem.*, 2006, **4**, 2507–2517.
- 17 K. Spiegel et al., Duocarmycins Binding to DNA Investigated by Molecular Simulation, *J. Phys. Chem. B*, 2006, **110**, 3647–3660.
- 18 A. V. Vargiu et al., Anthramycin–DNA Binding Explored by Molecular Simulations, *J. Phys. Chem. B*, 2006, **110**, 24687–24695.
- 19 T. Tuttle et al., A QM/MM Study of the Bergman Reaction of Dynemicin A in the Minor Groove of DNA, *J. Phys. Chem. B*, 2007, **111**, 8321–8328.
- 20 N. Sundaresan et al., Role of Mg²⁺ and Ca²⁺ in DNA Bending: Evidence from an ONIOM-Based QM-MM Study of a DNA Fragment, *J. Phys. Chem. A*, 2006, **110**, 8826–8831.
- 21 V. J. Burton et al., Molecular Mechanics for Coordination Complexes: The Impact of Adding d-Electron Stabilization Energies, *J. Am. Chem. Soc.*, 2002, **117**, 8407–8415.
- 22 R. J. Deeth et al., Molecular modelling for transition metal complexes: Dealing with d-electron effects, *Coord. Chem. Rev.*, 2009, **253**, 795–816.
- 23 R. J. Deeth et al., Molecular modelling of Jahn–Teller distortions in Cu(ii)N₆ complexes: elongations, compressions and the pathways in between, *Dalt. Trans.*, 2006, 1092–1100.
- 24 R. J. Deeth et al., Molecular modelling for coordination compounds: Cu(ii)–amine complexes, *Dalt. Trans.*, 2005, 3638–3645.
- 25 R. J. Deeth, A test of ligand field molecular mechanics as an efficient alternative to QM/MM for modelling metalloproteins: the structures of oxidised type I copper centres, *Chem. Commun.*, 2006, 2551–2553.
- 26 H. Do et al., Computational Study of the Structure and Electronic Circular Dichroism Spectroscopy of Blue Copper Proteins, *J. Phys. Chem. B*, 2013, **117**, 8105–8112.
- 27 M. Turner et al., Modeling of Platinum-Aryl Interaction with Amyloid-β Peptide, *J. Chem. Theory Comput.*, 2016, **12**, 1385–1392.

- 28 M. Turner et al., Ligand field molecular dynamics simulation of Pt(II)-phenanthroline binding to N-terminal fragment of amyloid-beta peptide, *PLoS One*, 2018, **13**, e0193668.
- 29 S. T. Mutter et al., Metal Binding to Amyloid- β 1–42: A Ligand Field Molecular Dynamics Study, *ACS Chem. Neurosci.*, 2018, **9**, 2795–2806.
- 30 M. Ćendić et al., Molecular modeling for Cu(II)-aminopolycarboxylate complexes: Structures, conformational energies, and ligand binding affinities, *J. Comput. Chem.*, 2013, **34**, 2687–2696.
- 31 A. Bentz et al., Modeling of the Various Minima on the Potential Energy Surface of Bispidine Copper(II) Complexes: A Further Test for Ligand Field Molecular Mechanics, *Inorg. Chem.*, 2008, **47**, 9518–9527.
- 32 H.-C. Tai et al., Combined Theoretical and Computational Study of Interstrand DNA Guanine–Guanine Cross-Linking by trans-[Pt(pyridine)₂] Derived from the Photoactivated Prodrug trans,trans,trans-[Pt(N₃)₂ (OH)₂ (pyridine)₂], *Inorg. Chem.*, 2012, **51**, 6830–6841.
- 33 S. Grimme et al., A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All spd-Block Elements (Z = 1–86), *J. Chem. Theory Comput.*, 2017, **13**, 1989–2009.
- 34 C. Bannwarth et al., GFN2-xTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions, *J. Chem. Theory Comput.*, 2019, **15**, 1652–1671.
- 35 M. Bursch et al., Structure Optimisation of Large Transition-Metal Complexes with Extended Tight-Binding Methods, *Angew. Chemie Int. Ed.*, 2019, **58**, 11078–11087.
- 36 S. Grimme, Exploration of Chemical Compound, Conformer, and Reaction Space with Meta-Dynamics Simulations Based on Tight-Binding Quantum Chemical Calculations, *J. Chem. Theory Comput.*, 2019, **15**, 2847–2862.
- 37 A. Kovács et al., Benchmarking density functionals in conjunction with Grimme's dispersion correction for noble gas dimers (Ne₂, Ar₂, Kr₂, Xe₂, Rn₂), *Int. J. Quantum Chem.*, 2017, **117**, e25358.
- 38 M. Bursch et al., Fast and Reasonable Geometry Optimization of Lanthanoid Complexes with an Extended Tight Binding Quantum Chemical Method, *Inorg. Chem.*, 2017, **56**,

- 12485–12491.
- 39 C. J. Cramer, *Essentials of computational chemistry: theories and models*, John Wiley & Sons, 2013.
- 40 J. L. Heilbron, Rutherford–Bohr atom, *Am. J. Phys.*, 1981, **49**, 223–231.
- 41 E. Schrödinger, An Undulatory Theory of the Mechanics of Atoms and Molecules, *Phys. Rev.*, 1926, **28**, 1049–1070.
- 42 F. Jensen, *Introduction to computational chemistry*, Chichester, UK Hoboken, NJ : John Wiley & Sons, Third edit., 2017.
- 43 M. Born et al., Zur quantentheorie der moleküle, *Ann. Phys.*, 1927, **389**, 457–484.
- 44 J. C. Slater et al., in *Quantum Theory of Atomic Structure*, 1960.
- 45 D. H. Kobe, Variational Principle and Slater's Generalized Hartree-Fock Theory for Nuclei, *Phys. Rev.*, 1969, **188**, 1583–1589.
- 46 L. H. Thomas, The calculation of atomic fields, *Math. Proc. Cambridge Philos. Soc.*, 1927, **23**, 542–548.
- 47 E. Fermi, Statistical method to determine some properties of atoms, *Rend. Accad. Naz. Lincei*, 1927, **6**, 5.
- 48 P. Hohenberg et al., Inhomogeneous electron gas, *Phys. Rev.*, 1964, **136**, B864–B871.
- 49 W. Kohn et al., Self-consistent equations including exchange and correlation effects, *Phys. Rev.*, 1965, **140**, A1133–A1138.
- 50 R. G. Parr et al., *Density-Functional Theory of Atoms and Molecules (International Series of Monographs on Chemistry: 16)*, New York, 1989.
- 51 D. M. Ceperley et al., Ground State of the Electron Gas by a Stochastic Method, *Phys. Rev. Lett.*, 1980, **45**, 566–569.
- 52 S. H. Vosko et al., Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis, *Can. J. Phys.*, 1980, **58**, 1200–1211.
- 53 J. P. Perdew et al., Self-interaction correction to density-functional approximations for many-electron systems, *Phys. Rev. B*, 1981, **23**, 5048–5079.
- 54 A. D. Becke, Density-functional exchange-energy approximation with correct asymptotic behavior, *Phys. Rev. A*, 1988, **38**, 3098.
- 55 A. D. Becke, Density-functional thermochemistry. III. The role of exact exchange, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
- 56 C. Lee et al., Development of the Colle-Salvetti correlation-energy formula into a

- functional of the electron density, *Phys. Rev. B*, 1988, **37**, 785–789.
- 57 E. J. Meijer et al., A density-functional study of the intermolecular interactions of benzene, *J. Chem. Phys.*, 1996, **105**, 8684–8689.
- 58 S. Grimme, Semiempirical GGA-type density functional constructed with a long-range dispersion correction, *J. Comput. Chem.*, 2006, **27**, 1787–1799.
- 59 M. J. S. Dewar et al., Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model, *J. Am. Chem. Soc.*, 1985, **107**, 3902–3909.
- 60 J. J. P. Stewart, Optimization of parameters for semiempirical methods I. Method, *J. Comput. Chem.*, 1989, **10**, 209–220.
- 61 J. J. P. Stewart, Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements, *J. Mol. Model.*, 2007, **13**, 1173–1213.
- 62 J. J. P. Stewart, Optimization of parameters for semiempirical methods VI: More modifications to the NDDO approximations and re-optimization of parameters, *J. Mol. Model.*, 2013, **19**, 1–32.
- 63 M. Gaus et al., DFTB3: Extension of the self-consistent-charge density-functional tight-binding method (SCC-DFTB), *J. Chem. Theory Comput.*, 2012, **7**, 931–948.
- 64 D. A. Pearlman et al., AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules, *Comput. Phys. Commun.*, 1995, **91**, 1–41.
- 65 R. J. Deeth, Book Review: Molecular Modeling of Inorganic Compounds. By P. Comba and T. W. Hambley, *Angew. Chemie Int. Ed. English*, 1996, **35**, 1349.
- 66 T. A. Halgren, Merck molecular force field. II. MMFF94 van der Waals and electrostatic parameters for intermolecular interactions, *J. Comput. Chem.*, 1996, **17**, 520–552.
- 67 T. Halgren, Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94, *J. Comput. Chem.*, 1996, **17**, 490–519.
- 68 W. L. Jorgensen et al., The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin, *J. Am. Chem. Soc.*, 1988, **110**, 1657–1666.
- 69 A. K. Rappé et al., UFF, a full periodic table force field for molecular mechanics and

- molecular dynamics simulations, *J. Am. Chem. Soc.*, 1992, **114**, 10024–10035.
- 70 R. J. Deeth et al., DommiMOE: An implementation of ligand field molecular mechanics in the molecular operating environment, *J. Comput. Chem.*, 2005, **26**, 123–130.
- 71 P. Comba et al., in *Molecular Modeling of Inorganic Compounds*, John Wiley & Sons, Ltd, 2009, p. 1.
- 72 D. A. Johnson et al., A new analysis of the ligand field stabilization energy, *Inorg. Chem.*, 1995, **34**, 3253–3259.
- 73 J. H. Van Vleck, Theory of the Variations in Paramagnetic Anisotropy Among Different Salts of the Iron Group, *Phys. Rev.*, 1932, **41**, 208–215.
- 74 H. Bethe, Termaufspaltung in kristallen, *Ann. Phys.*, 1929, **395**, 133–208.
- 75 C. E. Schäffer et al., The angular overlap model, an attempt to revive the ligand field approaches, *Mol. Phys.*, 1965, **9**, 401–412.
- 76 D. Hamelberg et al., Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules, *J. Chem. Phys.*, 2004, **120**, 11919–11929.
- 77 D. Hamelberg et al., Sampling of slow diffusive conformational transitions with accelerated molecular dynamics, *J. Chem. Phys.*, 2007, **127**, 155102.
- 78 D. Bucher et al., On the Use of Accelerated Molecular Dynamics to Enhance Configurational Sampling in Ab Initio Simulations, *J. Chem. Theory Comput.*, 2011, **7**, 890–897.
- 79 P. R. L. Markwick et al., Studying functional dynamics in bio-molecules using accelerated molecular dynamics, *Phys. Chem. Chem. Phys.*, 2011, **13**, 20053–20065.
- 80 Y. Miao et al., Accelerated molecular dynamics simulations of protein folding, *J. Comput. Chem.*, 2015, **36**, 1536–1549.
- 81 J. Wereszczynski et al., ed. R. Baron, Springer New York, 2012, pp. 515–524.
- 82 Y. Wang et al., Accelerated Molecular Dynamics: Theory, Implementation and Applications, 2012, 165–172.
- 83 Y. Wang et al., Implementation of accelerated molecular dynamics in {NAMD}, *Comput. Sci. Discov.*, 2011, **4**, 15002.
- 84 Y. Miao et al., Improved reweighting of accelerated molecular dynamics simulations for free energy calculation, *J. Chem. Theory Comput.*, 2014, **10**, 2677–2689.
- 85 D. J. Price et al., A modified TIP3P water potential for simulation with Ewald summation, *J. Chem. Phys.*, 2004, **121**, 10096–10103.

3 Method validation on metal-peptide systems

3.1 Cu(II)-GHK

3.1.1 Introduction

GHK is a tripeptide consisting of glycine, histidine, and lysine (Gly-His-Lys) amino acids (Figure 3.1). Found naturally in blood plasma, urine and saliva,¹ it was first isolated by Pickart et al. in 1973 and identified as a liver cellular growth factor.² GHK has high binding affinity for copper and zinc cations, forming complexes whose function is believed to be mainly transport of Cu ions to cells.³ After the Cu-GHK complex was identified, it has been investigated in numerous *in vivo* and *in vitro* studies, which indicate its wide range of biological functions. The exchange dynamics and redox behaviour of GHK-Cu are stable in biological system as reported experimentally⁴ which makes it safe towards lipids and amino acids when copper is being transported to the cells.⁵⁻⁷

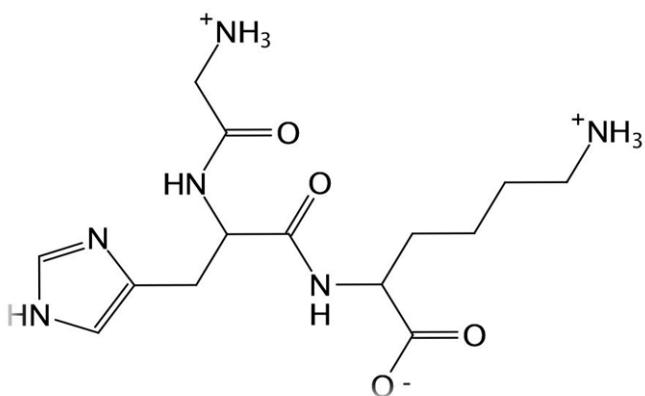


Figure 3.1: GHK structure.

The Cu-GHK complex plays significant positive roles in the body; promoting wound healing, acting as an anti-inflammatory agent, and promoting cell growth and stimulating skin cell regeneration.^{8,9} It also has been established that Cu-GHK has potential as an anticancer agent, and that it can repair damaged cells by accelerating DNA replication for cancer patients after radiation therapy.¹⁰ In addition, it shows potency toward neurodegenerative diseases such as

Alzheimer's and Prion disease by protecting neurons from amyloid- β (A β) aggregation through competitive chelation to copper instead, which in turn decreases neurotoxicity.¹¹

The well-documented chemical and physiological activity of Cu-GHK makes it interesting in its own right, while its relatively small size makes it a useful model for binding of metal ions to larger peptides. Hence, the structural and coordinating properties of Cu-GHK complex have been given much attention in the recent years. Several experimental studies have probed the structure and behaviour of Cu-GHK in different conditions, including X-ray crystallography, NMR, EPR, IR, electronic absorption and Raman spectroscopy, as well as calorimetry and titration.^{4, 6, 12–14} The structure obtained from X-ray studies shows the complex forms when Cu(II) chelates to the histidine side chain, glycine- α amino group and deprotonated nitrogen from the glycine-histidine peptide bond. Perkins et al.⁶ reported a Cu-GHK crystal structure in which three different peptides link one Cu ion through 3N2O coordination; three nitrogen atoms come from one peptide, while the other oxygens come from two separate peptides. Hureau et al.¹⁵ reported a binuclear structure, with two copper atoms in the solid state, resulting in penta-coordinated, distorted square pyramid around Cu, with 3N1O in the equatorial positions and a further oxygen occupying the apical position, in which Cu binds to three N-atoms from the same GHK, while two O atoms come from lysine of a neighbouring tripeptide.

In contrast, in solution the complex forms a mononuclear structure, indicating rupture of dimeric structure that is present in solid state. Hence, it is possible that the equatorial oxygen position, provided by carboxylate group of neighbouring lysine in the crystal, is displaced by solvent molecule in liquid state. The source of this fourth oxygen-copper bond is provided from water (solvent) or carboxyl groups of lysine from neighbouring GHK.^{5–7, 15} All studies indicate distorted square pyramidal geometry around Cu, in which four ligands are equatorial with distance of about 2.0 Å, with the remaining apical ligand slightly farther from copper at around 2.5 Å.

Hureau et al. also studied reactive oxygen species (ROS) production from Cu-GHK peptide using cyclic voltammetry, fluorescence, and EPR measurements.¹⁵ This showed the difficulty for the complex to be reduced by ascorbate, indicating redox silencing of Cu(II) when bound to GHK peptide. However, another study by Guilloréau et al.¹⁶ suggested that Cu-GHK catalyses formation of reactive species such as the hydroxyl radical, HO·.

Even with all this experimental data, there are challenges to determining the structure and conformation of Cu-GHK due to flexibility and multiple potential metal binding sites. Electronic structure calculations can provide complementary information to experiments, and are an effective means for analysis of compounds containing transition metal ions,¹⁷ and metal-biomolecule interactions.¹⁸ However, modelling flexible peptides when bound to metal ions is challenging. Accurate quantum mechanical methods are computationally expensive, even for those methods that are relatively efficient such as density functional theory (DFT). In this work, we test LFMM and GFN2-xTB approaches for describing metal coordination to GHK, using DFT methods as a benchmark, then apply these to examine the dynamical behaviour of GHK-Cu using molecular dynamics (MD).

3.1.2 Computational Method

The Cu(II)-GHK system was set up in MOE as follows. The sequence, Gly-His-Lys was constructed in extended geometry with all peptide bonds in *trans*- orientation. Cu(II) was coordinated to the neutral N-terminus of Gly, deprotonated N of Gly-His peptide bond, either N ϵ /N δ of the His imidazole ring, and to O of a single water molecule in the equatorial position to form a distorted square planar geometry, as reported by Hureau and co-workers.¹⁵ The overall charge on the complex is +1, arising from +2 on Cu and +1 from protonated Lys balanced against the deprotonated peptide N and carboxylate C-terminus. Low mode MD and stochastic^{19,20} conformational exploring were organized through MOE, using LFMM for Cu and coordinated atoms, and AMBER94²¹ force field parameters for the remainder of the peptide. Parameters for Cu-N_{his} and Cu-N_{amine} were reported previously,²² while parameters for Cu—OH₂ and Cu—N_{pept} were estimated by analogy with the chemically similar species, the former using peptide backbone O as a guide, the latter using N_{His}. Ligand field molecular mechanics (LFMM)²³ calculations were carried out by the DommiMOE extension to the Molecular Operating Environment (MOE).²⁴ Two distinct conformational searching methods were used to provide more data for benchmarking: both include geometry optimisation, such that all MM geometries and energies are, by definition, are at local minima.

Conformers generated from LFMM were used for DFT single-point calculations and geometry optimization. B3LYP-D2^{10,11} with def2-SVP^{12, 13} basis set was extensively used, since this was also used as a benchmark in previous work²² though for comparison we also examined the

performance of B3LYP-D3, PBE,^{29,30} PBE0^{31,32} also with D3³³ as well as M06-2x³⁴ and ωB97xD.³⁵ DFT calculations used the polarizable continuum model (PCM) and the conductor-like polarizable continuum model (CPCM). Geometry optimization using semiempirical GFN2-xTB methods employed a generalized Born-surface area (GB-SA) model of aqueous solvent. DFT calculations were performed using Gaussian09³⁶ and ORCA³⁷ software. Semi-empirical calculations were performed via the XTB program.³⁸ Visual molecular dynamics (VMD)³⁹ program was used for trajectory analysis.

Further conformational searches were done by conformer-rotamer ensemble sampling tool (CREST)^{40,41} approach, within the xtb suite of programs, that combines metadynamics sampling (MTD), and genetic z-matrix crossing (GC).⁴² Molecular dynamics simulations were performed using xtb within the NVT ensemble with timestep of 4.0 fs, facilitated by all bond lengths being restrained using the SHAKE algorithm⁴³ and fictitious hydrogen mass of 4 amu, at 310 K.

3.1.3 Result and discussion

The result of conformational search shows that while coordination of N ε and N δ of His imidazole ring to Cu are both possible, N δ is markedly more favorable, resulting in an approximate square pyramidal geometry, and lower energy values by (-640 kJ mol⁻¹ from LFMM/AMBER estimation) compared to N ε . The latter gives unfavorable 7-membered metallocycles, with the imidazole N-donor being located in axial position to copper, shown in Figure 3.2, leading us to choose N δ binding for subsequent calculations.

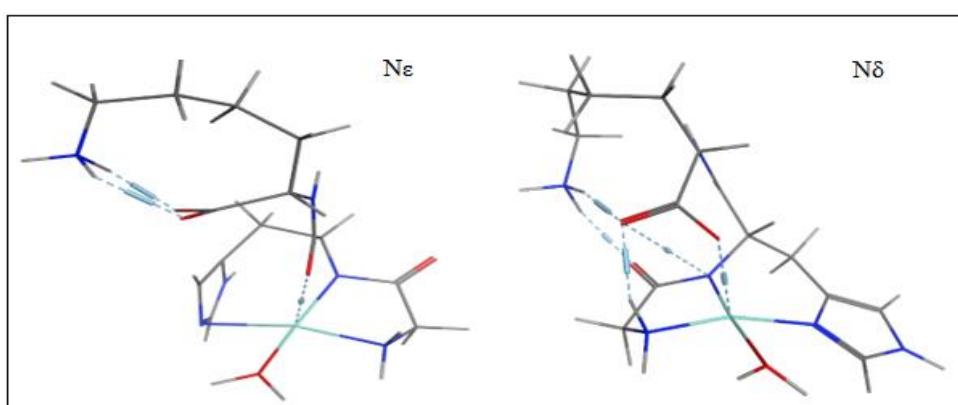


Figure 3.2: N δ and N ε geometries.

Low mode search finished with only three conformers, as shown in Figure 3.3 each of these had *cis*- orientation of peptide, indicating that this changes during the conformational search.

Superposition of conformations indicates that metal coordination remains constant, and flexibility is only found in the side-chain of Lys. For two conformations, the peptide was manually changed back to *trans*- and LFMM optimised, but for the third conformer the rotation could not be achieved due to constraints of metal coordination, leading to a total of five conformers (named as cis1, cis2, cis3, trans1, and trans3). This small number of conformations generated from low mode search is due to limited flexibility of such a small peptide linked to copper. Applying the same search for the copper-free GHK peptide produced 118 conformers, clearly demonstrating the reduction in conformational flexibility on copper binding. Stochastic conformational search found 7 conformers of Cu-GHK with mutual RMSD of at least 0.5 Å, all with *trans*- orientation of peptides, cartesian coordinates for LowMode and Stochastic conformations can be found in Table A1.

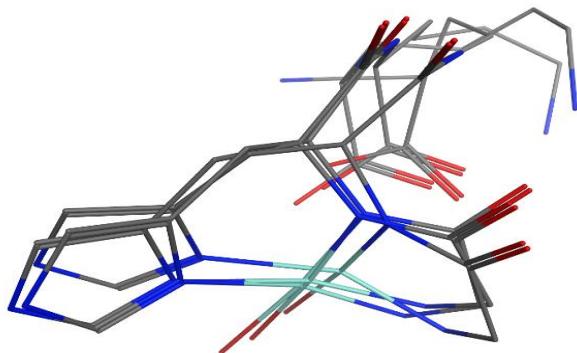


Figure 3.3: Three cis-GHK low mode conformers.

Despite the accuracy of DFT calculations, this method needs large computational expense for molecules of the size of metal-tripeptide. We therefore sought faster methods with acceptable accuracy for this system, and identified semi-empirical GFN2-xTB as a likely candidate for study of GHK-Cu complexes. However, it is important to test the results it provides. So, B3LYP-D2/def2-SVP and GFN2-xTB single point calculations were applied to the five conformer geometries obtained from Low Mode search, and relative energies were compared. The resulting energies show correlation ($R^2=0.73$) (Figure 3.4). After geometry optimization, the correlation between DFT (B3LYP-D2/def2-SVP) and GFN2-xTB relative energy becomes stronger ($R^2=0.89$), as shown in Figure 3.5, despite taking no more than 2 minutes per conformation for GFN2-xTB. Also, GFN2-xTB and the DFT functional used exhibit

similar trends as we move through the conformers; using B3LYP-D2/def2-SVP optimised geometries, including multiple functionals, *i.e.* B3LYP-D3, PBE, PBE0, M06-2x and ω B97xD3. The results show those functionals are in general agreement with B3LYP-D2/def2-SVP and GFN2-xTB energies, as shown in Figure 3.6. We also note that D2 and D3 dispersion corrections are in good agreement here, presumably due to the relatively small size of the Cu-GHK system.

In addition, GFN2-xTB geometries are in good agreement with those obtained from B3LYP-D2/def2-SVP geometry optimisation. RMSD values between GFN2-XTB and the DFT functional (B3LYP-D2) optimised geometry are in the range 0.3 to 1.1 Å (*cis*1, 2, 3 = 0.31, 0.35, 0.49 Å; *trans*1 and 3 = 1.06, 0.70 Å). Furthermore, DFT (B3LYP-D2/def2-SVP) geometries optimization and single point calculation performed on seven conformers, obtained from stochastic searching at GFN2-xTB optimized geometry, show good agreement between relative energies ($R^2=0.64$), as shown in Figure 3.7 and 3.8. The similarity between the methods, illustrated in Figures 3.4, 3.5 and 3.7, was also evaluated using rank correlation, where a statistically significant correlation was found for all three instances, as reported in Table A2. Minimal change in geometries were found from RMSD comparison as shown in Table 3.1. Therefore, we conclude that GFN2-xTB gave reasonable structures and relative energies compared to DFT (B3LYP-D2/def2-SVP), but with reduced calculation time.

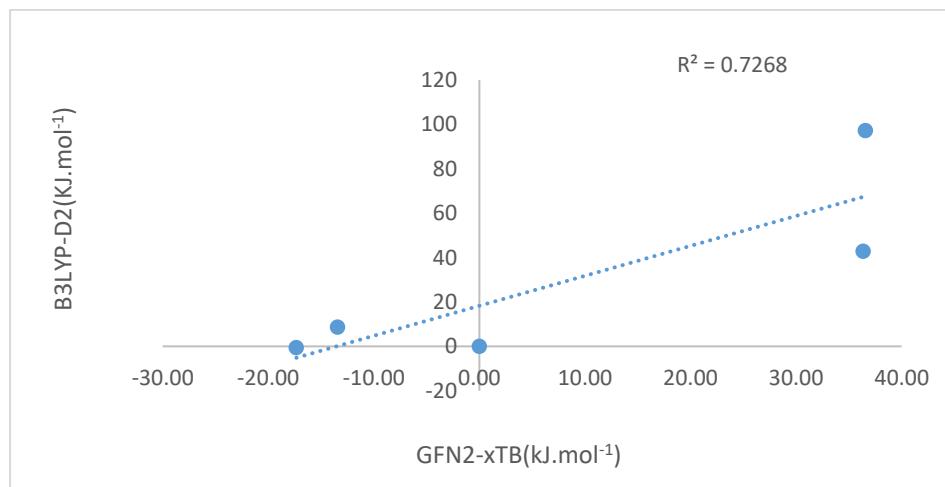


Figure 3.4: Relative energy comparative between GFN2-xTB and B3LYP-D2/def2-SVP at LFMM geometries on low mode conformers.

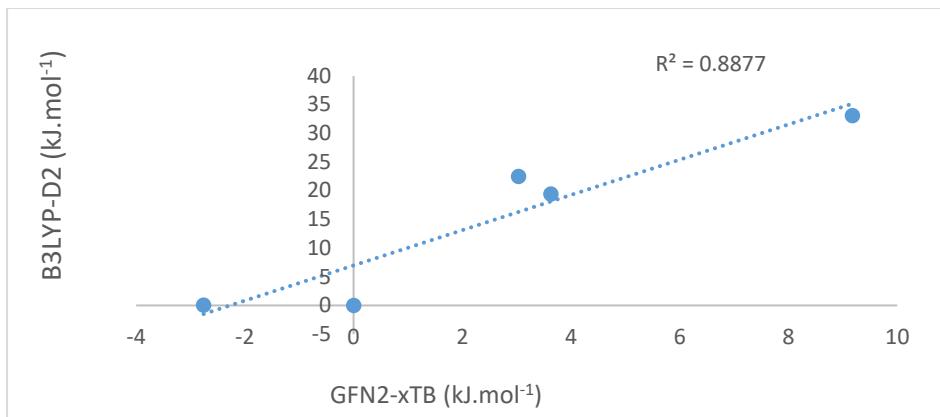


Figure 3.5: Relative energy of low mode conformers from B3LYP-D2/def2-SVP and GFN2-xTB.

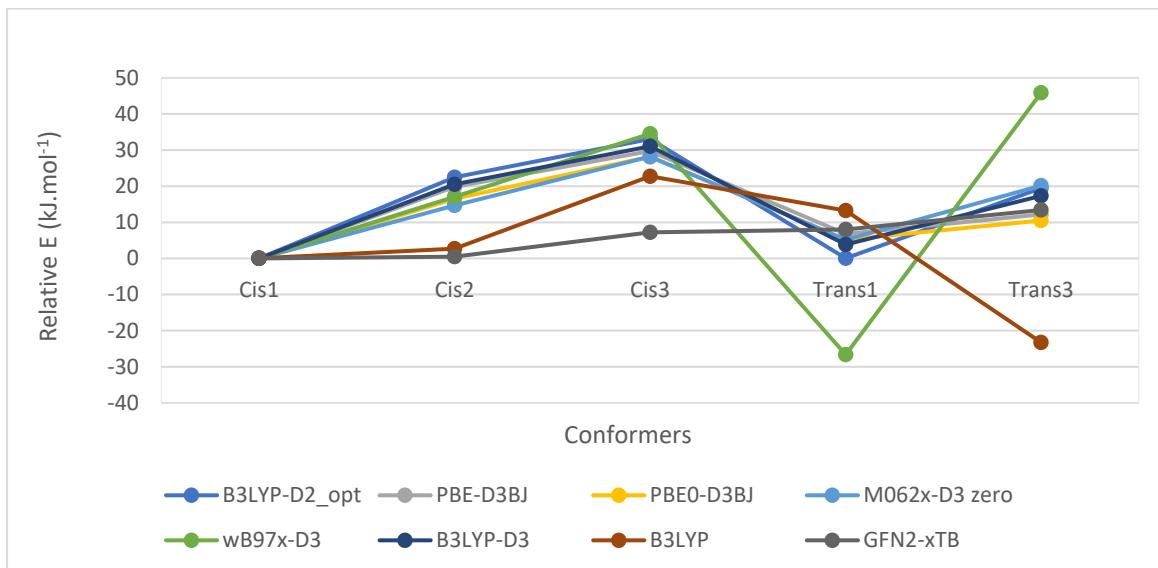


Figure 3.6: Relative energy for low mode conformers obtained by different methods as single point energy calculations at B3LYP-D2/def2-SVP optimised geometry.

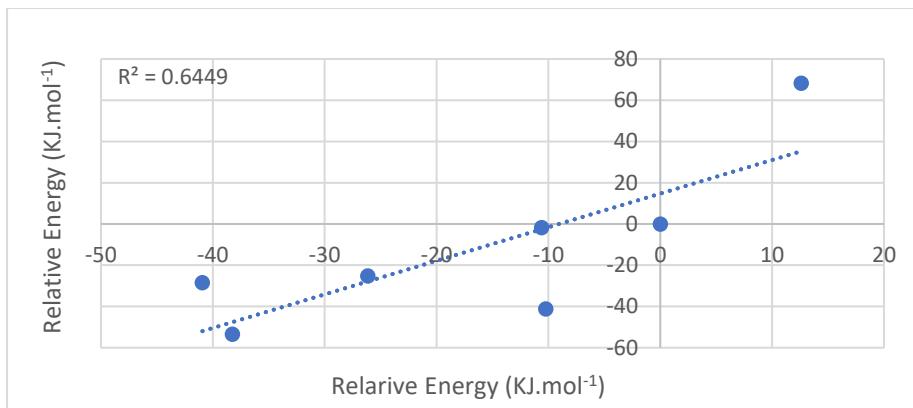


Figure 3.7: Single point calculation via DFT (B3LYP-D2/def2-SVP) of seven conformers found by stochastic search at GFN2-xTB optimized geometry.

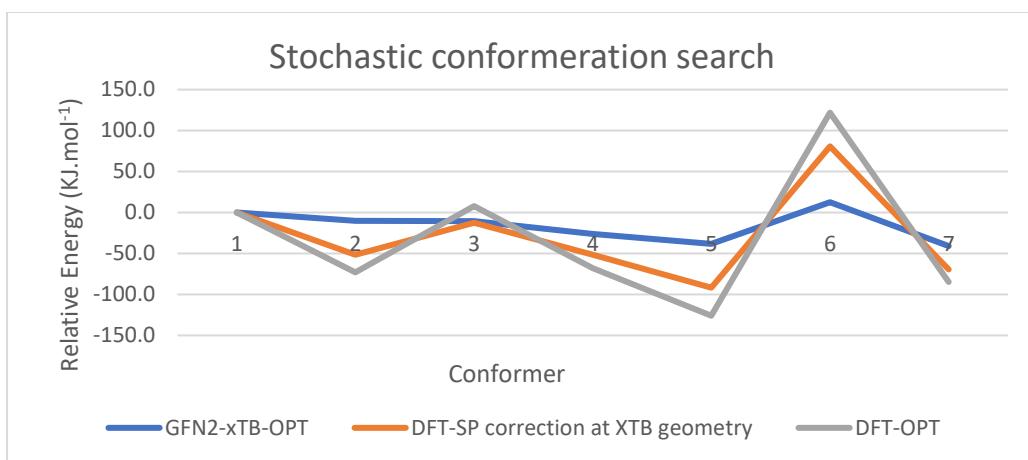


Figure 3.8: Comparison of relative energies GFN2-xTB and DFT (B3LYP-D2/def2-SVP) for 7 conformers from stochastic search.

Table 3.1: RMSD between DFT (B3LYP-D2/def2-SVP) and semi-empirical (GFN2-xTB) geometries generated from stochastic search /Å.

conformer #	RMSD
1	0.60
2	0.82
3	0.52
4	0.90
5	0.83
6	0.70
7	1.11

It is notable that most methods considered, predict the cis1 conformation to have the lowest energy. Closer inspection shows that cis1 has three hydrogen bonds compared to just two for trans1. For cis1, $-\text{NH}_3^+$ of the Lys side chain interacts with the carbonyl from the same amino acid carboxylate and also to the C-terminal carboxylate. The third hydrogen bond forms between N-terminal $-\text{NH}_2$ of Gly and Lys carboxylate group. In trans1, the same interactions to carboxylate are present but the interaction with the peptide carbonyl is lost (Figure 3.9) which may be the origin of the stability of the *cis*- form.

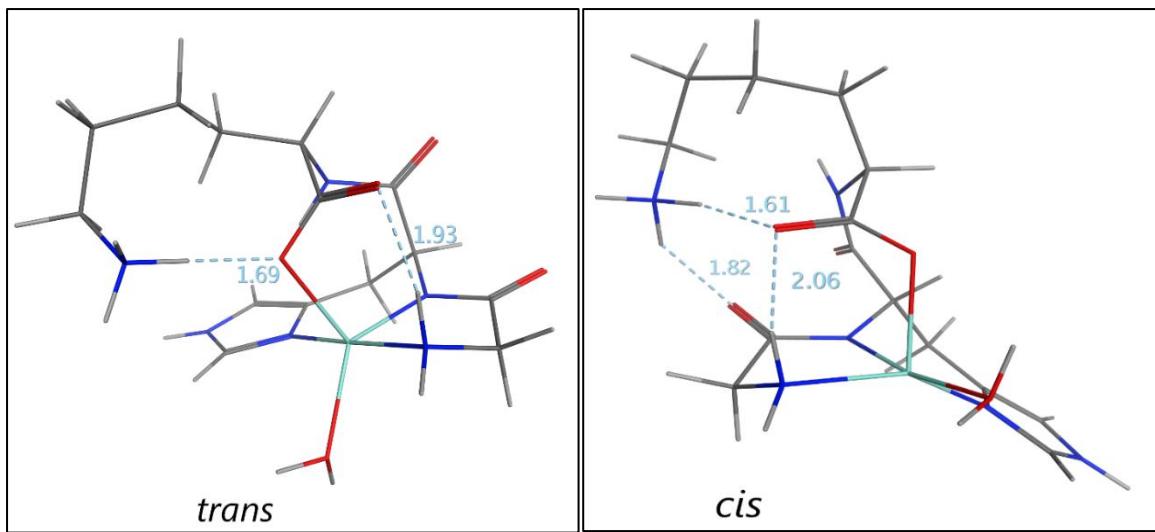


Figure 3.9: Hydrogen bonds of lowest cis and trans conformers.

The speed and accuracy of the GFN2-XTB method lends itself to dynamical simulations, so we first carried out conventional MD simulations at 310 K from low energy cis1 and trans1 conformations using GFN2-xTB. RMSD relative to starting point over a 100 ps trajectory is shown in Figure 3.10. Energy and temperature conservation over this timescale which stabilized around -84.21 au and around 300 K are shown in Figure 3.11. This shows these geometries are stable over the simulation timescale, typically remaining around 0.3 Å, although occasional increases to *ca.* 1 Å are observed in each case. Closer inspection shows that these increases do not involve the Cu-coordination sphere, but rather are due to changes in the orientation of the Lys side chain only, as shown in Figure 3.12. Table 3.2 shows the small average RMSD values and standard deviation for both trajectories. For comparison, equivalent MD simulations of the metal-free peptide were also performed (plot not shown), which show very similar RMSD statistics over this timescale.

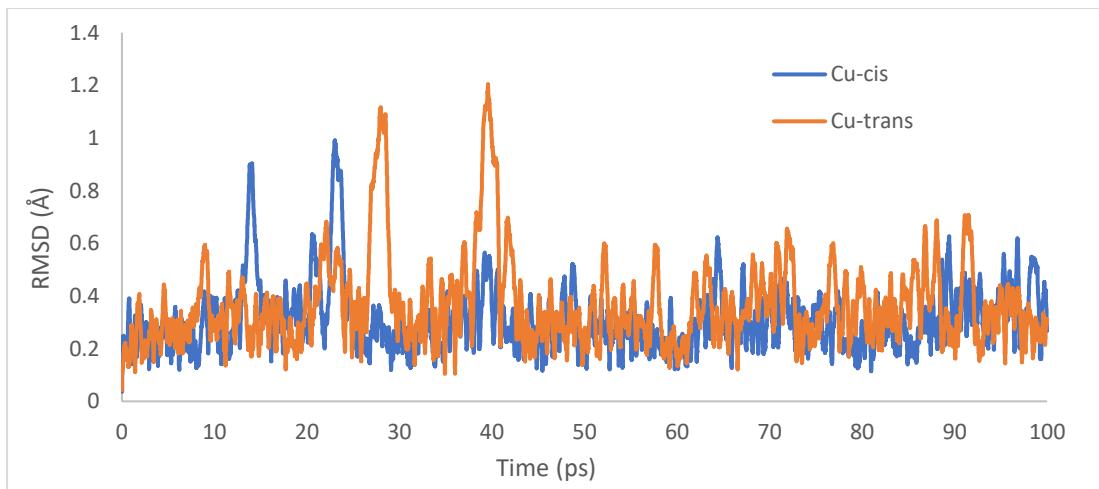


Figure 3.10: RMSD (\AA) over 100 ps MD simulation for cis and trans Cu-GHK.

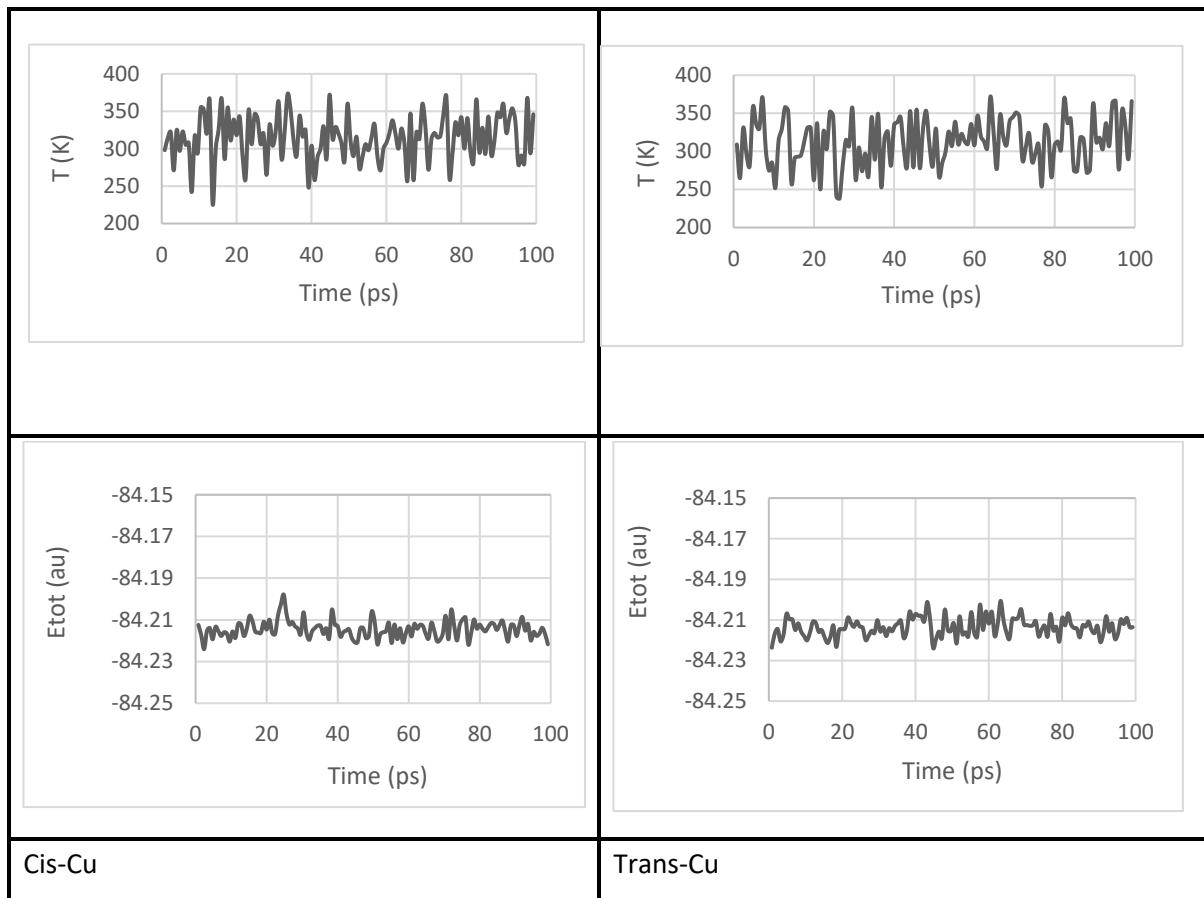


Figure 3.11: Temperature and Energy conservation for MD.

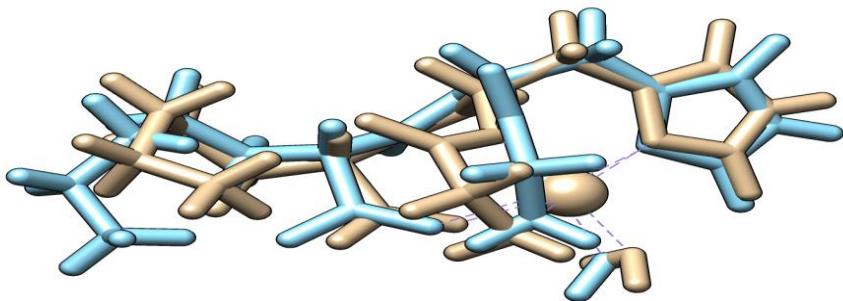


Figure 3.12: Comparison of starting and frame 1758 structure from MD simulation of Cu-cis, showing stability of Cu coordination and variation in Lys sidechain.

Table 3.2: RMSD statistical analysis of simulation data (Å).

	Ave	SD	MIN	MAX
Cu-trans	0.37	0.17	0.04	1.20
Free-trans	0.35	0.10	0.04	0.70
Cu-cis	0.31	0.12	0.04	0.99
Free-cis	0.29	0.09	0.03	0.63

Selected bond distances corresponding to the first coordination sphere of copper, as labelled in Figure 3.13, were analysed and summarized in Figure 3.14 and Table 3.3. This showed only small deviations for most bonds identified: only apical Cu—O coordination from carboxylate of Lys is found to deviate significantly from the starting value, reaching values as large as 3.72 Å, with standard deviation twice as large as found for other bonds. All four equatorial bonds to Cu are more stable, with averages close to 2.0 Å and small standard deviations, although distances as long as 2.6 Å are observed in specific frames. Taken together, this data suggests that copper binding to the peptide remains stable over the timescale of 100 ps.

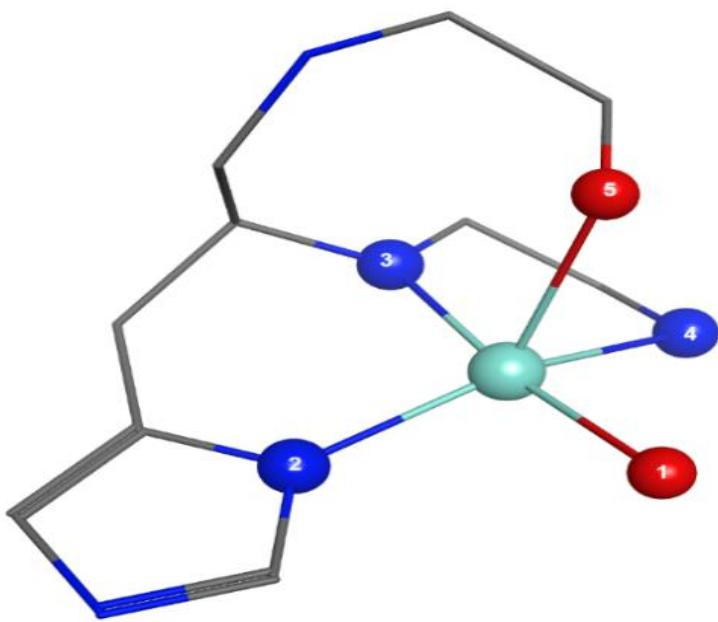
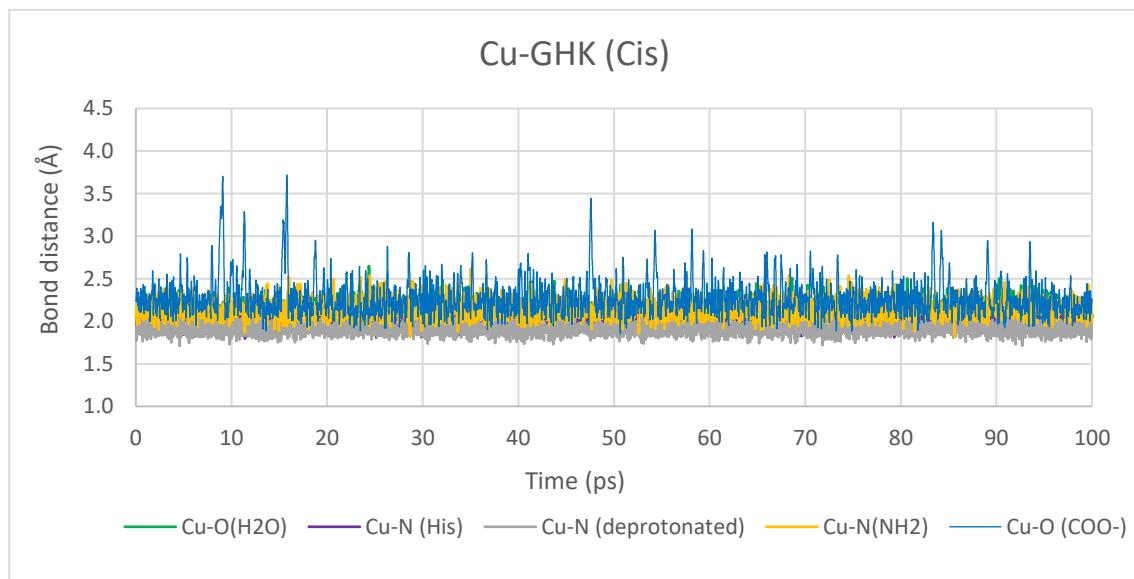


Figure 3.13: Numbering of Cu-ligand atoms (remaining atoms omitted for clarity).



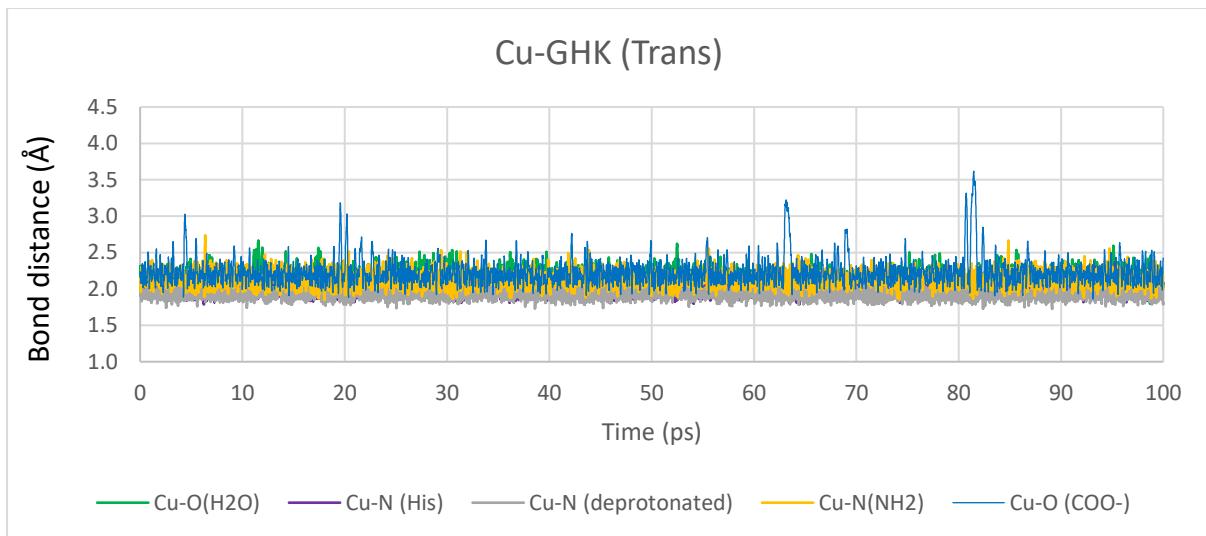


Figure 3.14: Cu-Ligand distances values for two trajectories.

Table 3.3: Bond distance trajectory analysis of RMSD values (Å).

cis1	Ave	SD	MIN	MAX
1 (Cu-H₂O)	2.09	0.10	1.80	2.65
2 (Cu-N_{His})	2.03	0.08	1.80	2.35
3 (Cu-N_{dept})	1.92	0.07	1.71	2.27
4 (Cu-N_{NH2})	2.15	0.10	1.81	2.61
5 (Cu-O_{coo})	2.28	0.21	1.85	3.72

Following the success of the GFN2-xTB approach, we explored whether this method could identify the known binding mode of Cu(II) to GHK, using the associated conformer–rotamer ensemble sampling tool (CREST). Initial input consisted of three separated fragments in one file, namely N-deprotonated GHK, Cu and H₂O as shown in Figure 3.15, with conformational search used for non-covalent mode to prevent dissociation. The most stable conformer generated, shown in Figure 3.16, is almost identical to *trans* optimized geometry that was identified from LFMM conformational search followed by GFN2-XTB (RMSD= 0.335 Å). This shows that GFN2-XTB coupled with the high efficiency of sampling in the CREST approach could be a valuable means for predicting the optimal binding mode for metals. However, we note that for this method to succeed it is necessary for the correct protonation state of the

peptide to be specified in advance, which may limit application of this approach if this is not known *a priori*.

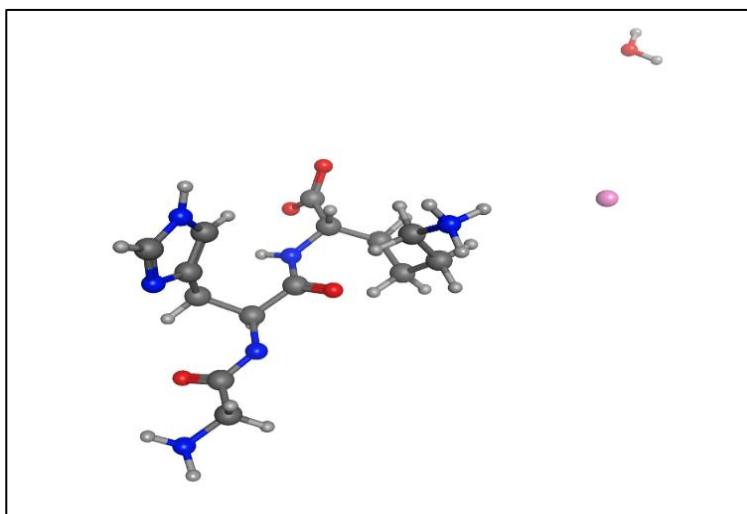


Figure 3.15: Input of GHK, H₂O, and Cu separately.

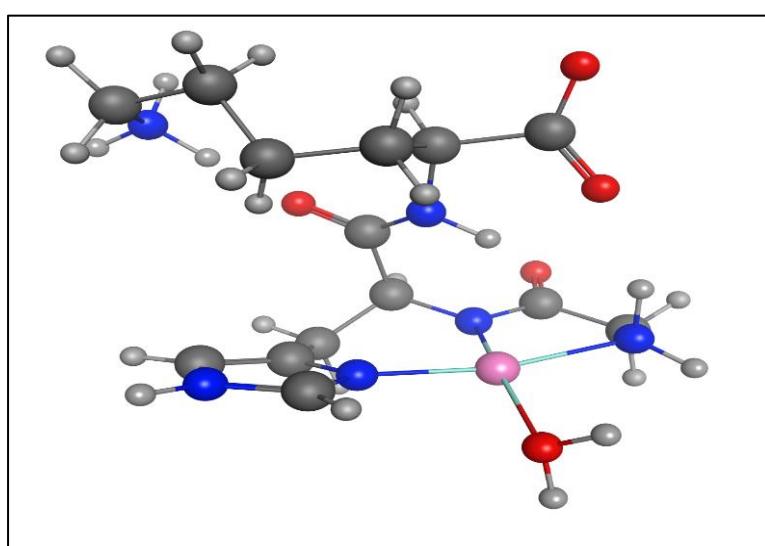


Figure 3.16: Most stable conformer CREST conformational search.

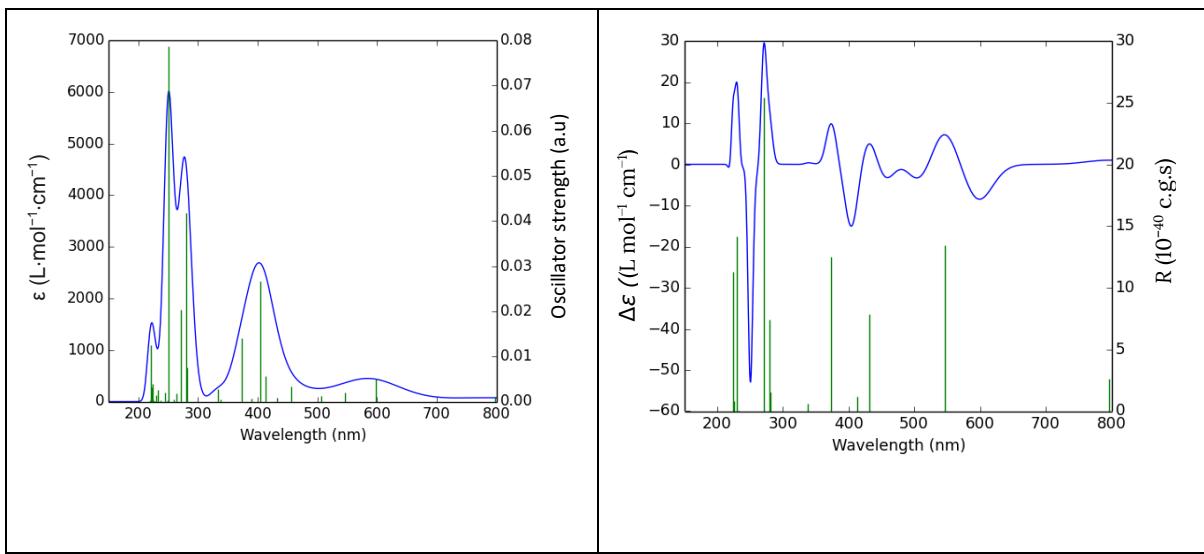
3.1.4 Electronic spectra analysis

Ultra violet spectrum (UV) and electronic circular dichroism (CD)

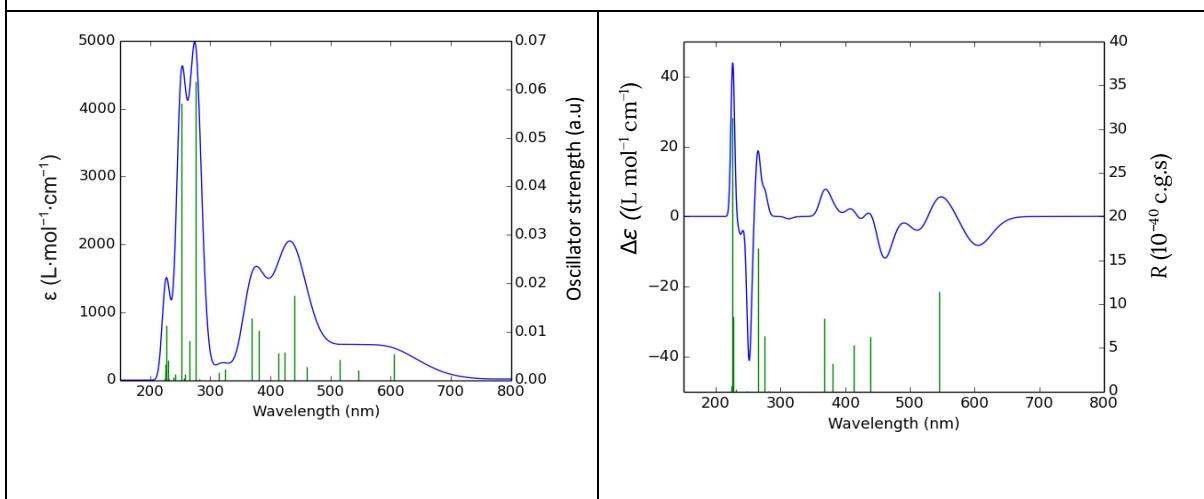
By using electronic spectra, we collect information about the structure of molecules. The ultraviolet (UV) and electronic circular dichroism (CD) prediction have been used to study

GHK-Cu absorption bands. Electronic excitation energies of the complex were calculated using time dependent TD-DFT/B3LYP method with the basis set def2SVP in solvent (water). The UV-Vis and CD spectra of the Cu (II)-GHK complex is shown in Figure 3.17.

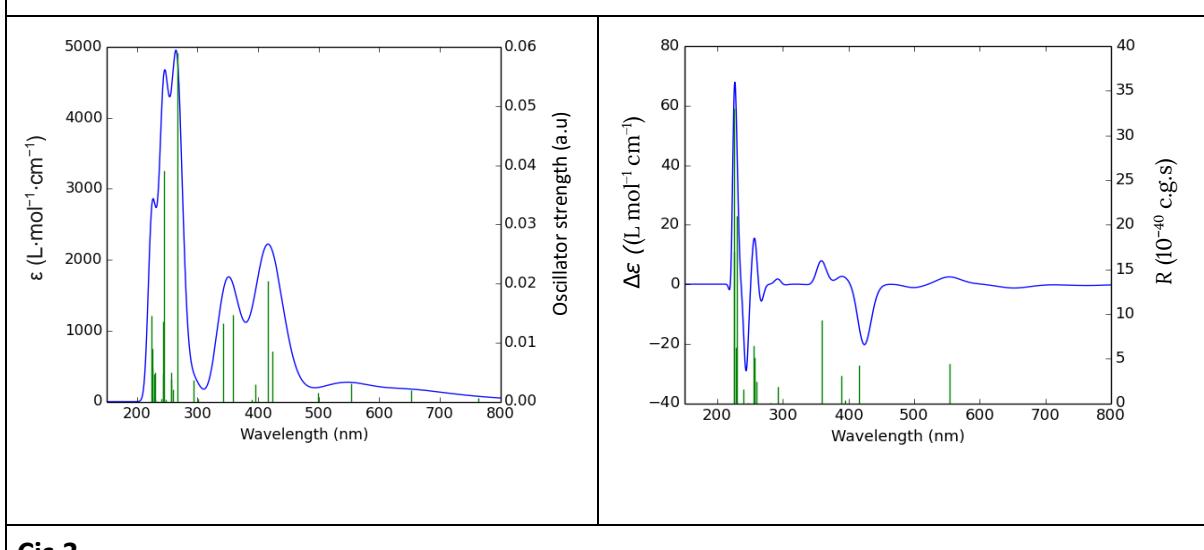
There are two distinct sets of peaks observed at the UV-vis spectrum; weak absorption bands around 500-600 nm and at a lower wavelength but at a slightly higher intensity centred around 400 nm, located at the d-d transitions region. Both are weak as they are Laporte forbidden. The second set displays high intensity, detected at 200-300 nm, in the ligand to metal charge transfer (LMCT) region. Overall, the d-d transitions are stronger in *cis* conformers than *trans*, with LMCT bands stronger in *trans*. The d-d transitions observed for cis1, cis2, cis3, trans 1, and trans 3 respectively of weak bands at ($\lambda = 598, 605, 554, 513, 574$ nm), and the stronger set bands respectively at ($\lambda = 405, 439, 416, 482, 427$ nm). In trans1 there are only two obvious sets but there is no noticeable band at 500-600 nm region where the oscillator length value here is 0.001 which indicating there is a low probability of a d-d transition occurring. The bands detected in LMCT transitions for conformers are as follows: cis1 at ($\lambda = 279, 250, 220$ nm), cis2 ($\lambda = 380, 275, 226$ nm), cis3 ($\lambda = 358, 266, 245$ nm), trans1 at ($\lambda = 279, 259, 252$ nm), trans3 at ($\lambda = 275, 267, 250$ nm) to histidine \rightarrow Cu (II), to amide \rightarrow Cu(II), and $-NH_2 \rightarrow$ Cu(II), charge transfer CT, respectively, in line with values reported in literature,⁴ where LMCT bands are reported to be detected using CD at 385, 330, and 295 nm, respectively. The same study stated the d-d transitions bands for Cu(II)-GHK at ($\lambda = 610$ and 600 nm). The similarity in wavelength between the prediction and what experiments found is expected, as the binding around the copper is the same. The results demonstrate a reasonable level of agreement with literature data,⁴ and cis2 gives the conformer most comparable to the experimental values. Therefore, the functional, basis set and solvent model chosen were able to discover the nearest exact correspondence.



Cis 1



Cis 2



Cis 3

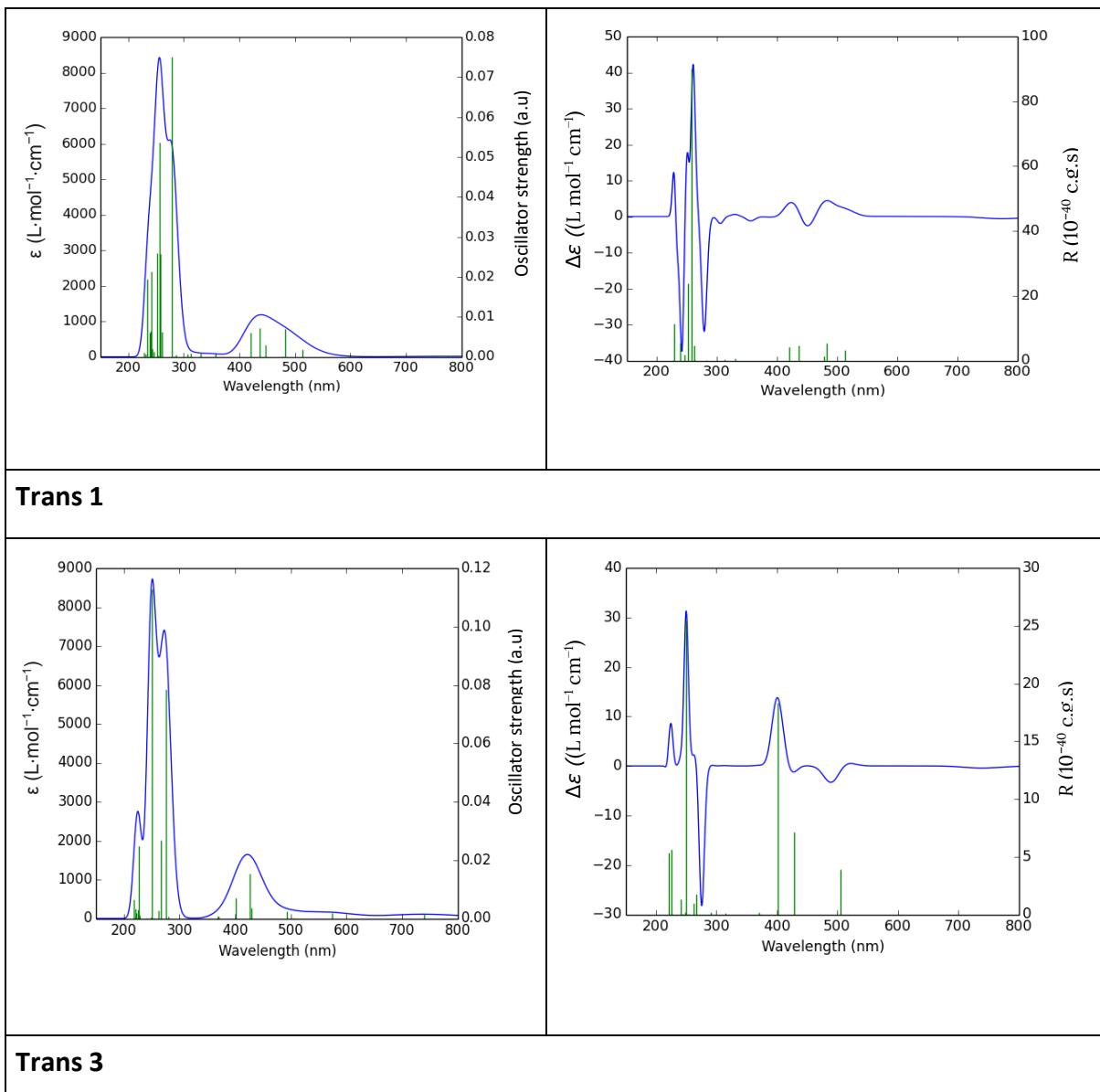
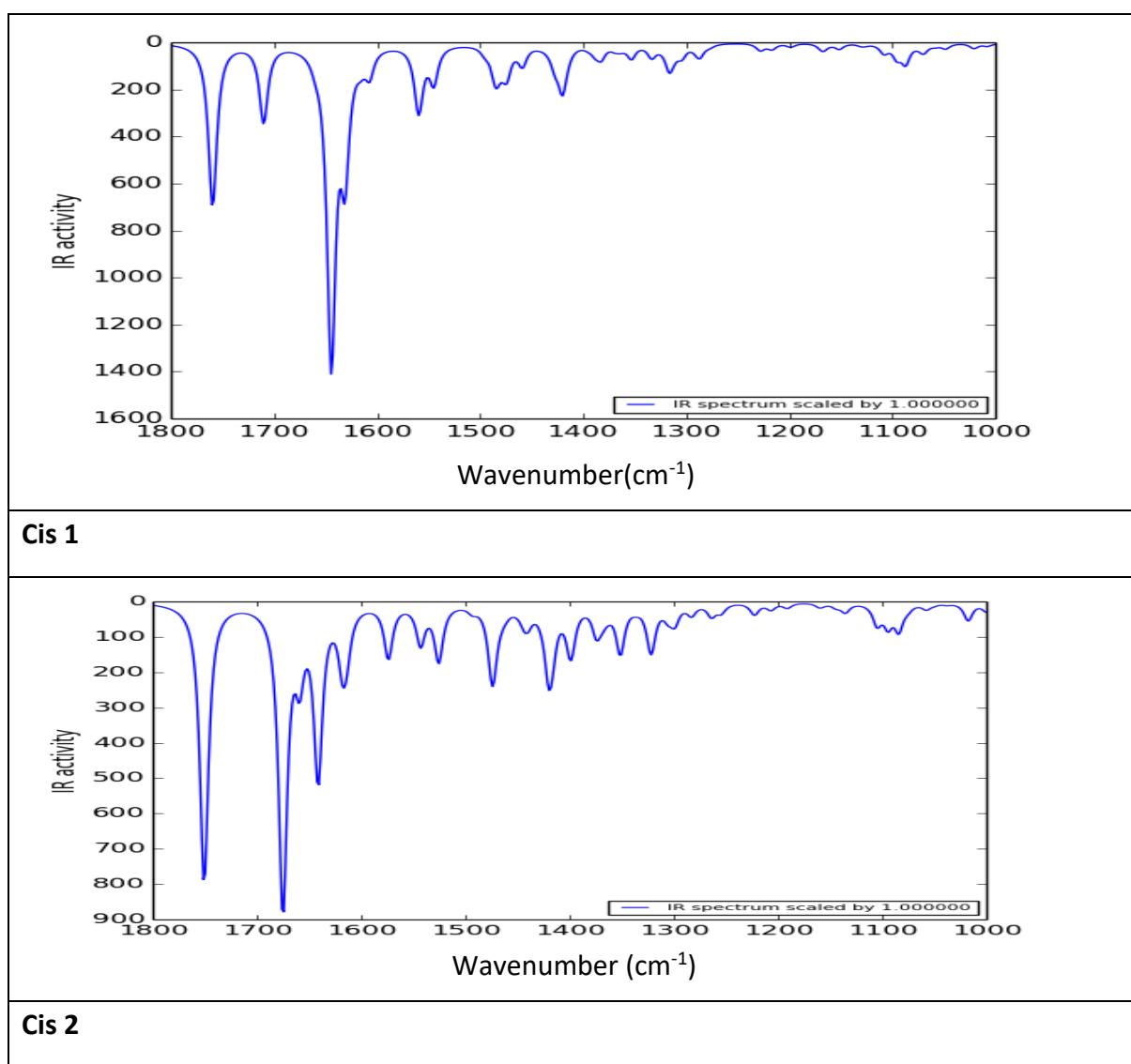


Figure 3.17: UV-vis and CD (nm) predicted spectra of Cu(II)-GHK conformers.

IR spectrum

IR spectra were calculated using Gaussian Version 09c01.³⁶ A DFT/B3LYP calculation with def2SV basis set and inclusion of empirical dispersion, was applied on all five conformers. The spectra were visualized via the GaussSum software.⁴⁴ The experimental data founded in literature for the IR band of Cu(II)-GHK was collected using dry film, so it was a concentrated structure (crystal) to avoid the water overlapping¹², however the calculation in the current study was simulated in the presence of solvent. The data obtained from IR calculations show similar IR frequency values to experimental data, that gave selected bond frequencies, such

as for C-N^{π/τ}, NH₃⁺ side chain, C=C of His ring or COO- from Lys, respectively with values 1108, 1526 and 1573 cm⁻¹ while the calculations gave 1568, 1535, 1534, 1572, and 1605 cm⁻¹ for the NH₃ group. Overall, the *trans* structures and cis 1 show slightly more vibrational intensity than cis2 and 3, otherwise spectral ranges look similar for all five conformers. Cis 1 and trans 1 present similar bands and cis 3 and trans 3 where they are originally the same starting geometry but the dihedral angles rotated manually from cis to trans. All IR predicted spectrum are shown in Figure 3.18. In addition, the computational calculations, provide more information for each of the bond modes assignment for the different conformers.



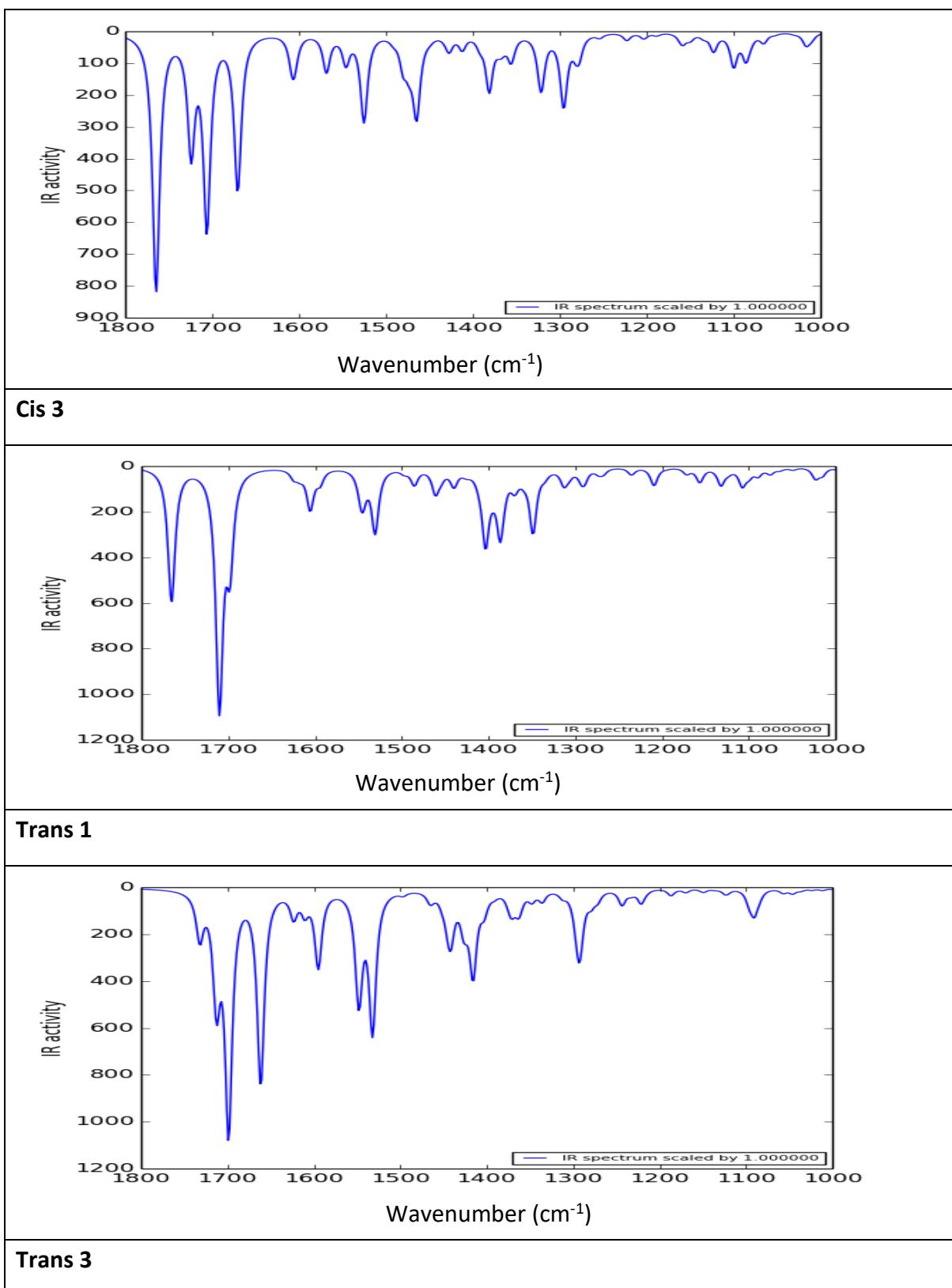


Figure 3.18: IR data for five conformers of GHK-Cu(II). The spectral region for wavenumbers from 1000 to 1700 cm⁻¹.

Electron paramagnetic resonance (EPR)

EPR is a method for studying compounds that contain unpaired electrons. It is a useful instrument for studying compounds containing metal complexes or radicals. In Cu(II)-GHK, conformers have an unpaired electron due to the Cu(II) ion having nine electrons in its outer shell. Therefore, EPR calculations are a useful tool to provide more information about this complex structure. By using UKS PBEO hybrid DFT method, the results showed a peak for copper at 141.6 MHz. The A_{\parallel} found in the literature for Cu-GHK complex to be 560 MHz and the g_{\parallel} and g_{\perp} values are 2.23, 2.05 respectively.⁴ The predicted g-factor for Cis1, Cis2, Cis3, trans1, and trans3 are 2.082, 2.076, 2.076, 2.072, and 2.073, respectively. Here, the DFT prediction shows there is not much difference in EPR parameters between conformers, as the changes in conformation are quite remote to the Cu(II) paramagnetic centre.

3.1.5 Conclusion

Cu(II)-GHK is used as a model to evaluate the semi-empirical tight binding method performance termed as GFN2-xTB, which can be then applied on larger metal-peptide systems such as amyloid- β . The main purpose of this study was to investigate the ability of this approach in predicting the geometry and energy of Cu(II)-GHK conformations. Therefore, here it is reported that the method is able to give reliable result of the lowest energy conformer and its geometry, as well as the relative energies of higher energy conformers, when compared to (B3LYP-D2/def2-SVP) data. In addition, the same method coupled with the efficient metadynamics sampling tool CREST, correctly predicted the binding site of Cu(II) to GHK. The efficiency of the GFN2-xTB method allowed the MD simulation of Cu(II)-GHK, indicating that all four equatorial bonds remain stable over 100 ps, in two different conformations, while the atypical bond to the C-terminal carboxylate is more inconsistent.

3.2 Cu(II)-Amyloid- β -16

3.2.1 Introduction

The small size of Cu(II)-GHK peptide means that it was utilised as a model to evaluate the performance of the GFN2-xTB method in the previous section of this chapter. The results

showed the ability of the method to reproduce the DFT (B3LYP-D2/def2-SVP) energy and geometry with low calculation time. Thus, the approach was also applied on the larger model of metal-peptide system; Cu(II)-A β 16. The aggregation of the peptide in the brain has been linked to Alzheimer's disease. In this section, the B3LYP-D2/def2-SVP method is used as a benchmark for the semiempirical GFN2-xTB method, with its efficiency evaluated on three fragments of A β -16 (termed: small, middle, and full – based on their size) bound to Cu(II), as illustrated in Figure 3.20. PBEh-3C geometry optimization was also explored, as this method can be a good candidate for metal-peptide system with less computational expense.⁴⁵ It has been reported that Cu(II) has a high affinity for coordination towards the N-terminal region of A β through the imidazole ring of histidine (His 6 and His 13 /14), as well as Asp1 residues.^{46–48} Thus, the four structures examined here are based on this foundation.

3.2.2 Computational Method

A β -16 peptides were constructed in MOE.²⁴ The exploration of the conformational space was conducted using the AMBER94²¹ force field parameters using the same program. The copper ion was then coordinated to the lower energy conformer, through several different sites of interaction, as suggested in the literature.^{46–48} The coordination was constructed as follows: a copper ion was coordinated to one of the nitrogen atoms of His 13, another from the imidazole ring of His 6, and NH₂ from Asp1; these three sites were kept for all geometries (labelled: A, B, C, and D), with the only difference being the fourth coordination bond, where the oxygens of Asp1 interacted from either the side-chain or backbone. In structure A, the oxygen from Asp1 backbone was bound to the central metal ion, while in structures B and C, the side-chain oxygen of Asp1 was used instead; although here, the backbone oxygen (C=O) was placed close towards Cu(II) in C, while backward opposite side of Cu(II) in B. In structure D, the coordination was similar to A, with the exception that the carboxylate ion from the side-chain of Asp1, but with different conformation where the bond angle (C α -C β -C γ) for Asp1 is around 90°, as shown in Figure 3.19.

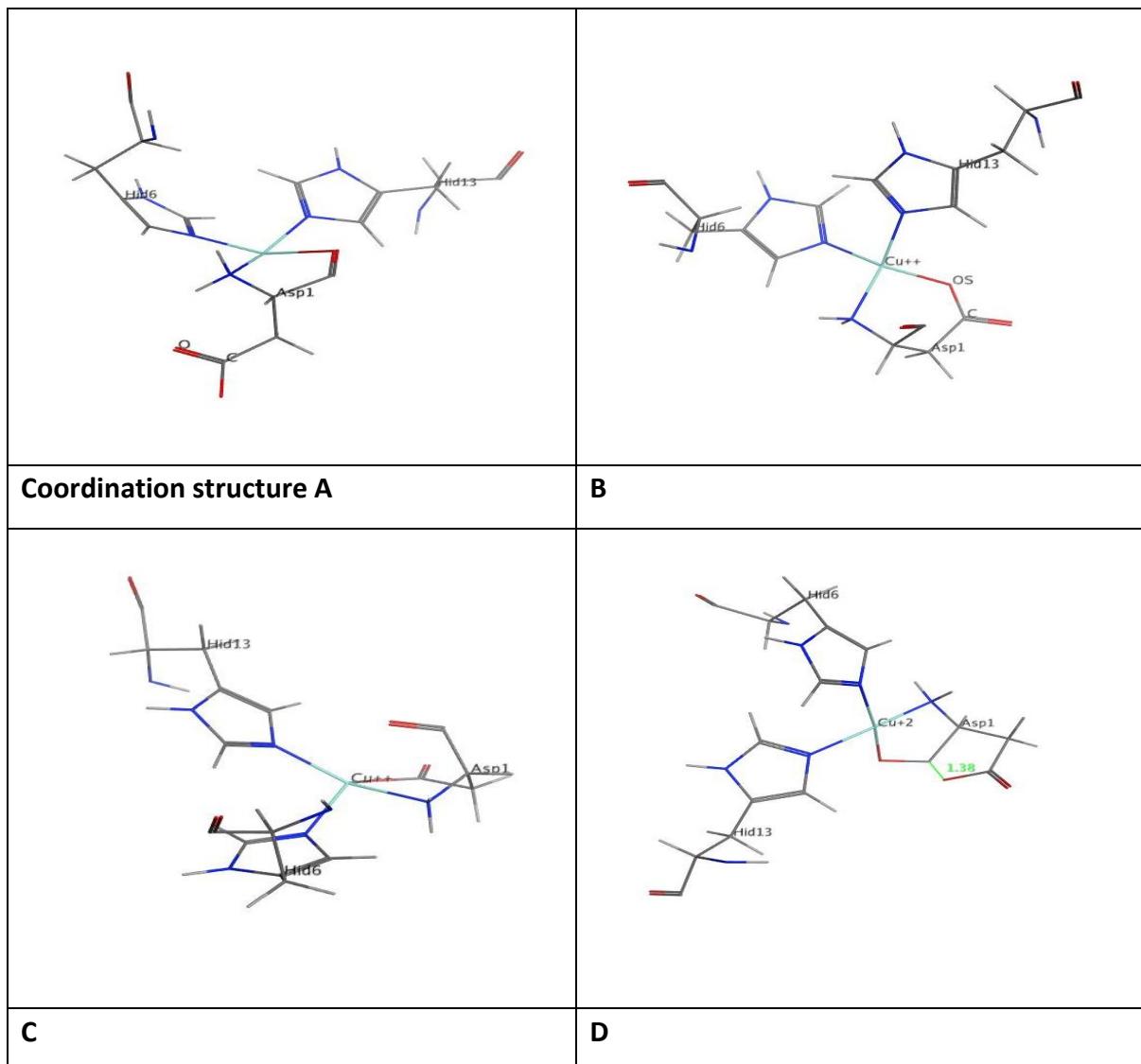


Figure 3.19: shows the coordination structure of all four arrangements of Asp1, His6, His13 bound to the Cu(II). Note; the geometries show just the active site, with the rest hidden for clarity.

The A β 16-Cu(II) complexes truncated at three length fragments (Full, middle, small) as follows: the full model consists all 16 amino acids (261 atoms), the middle has the first two coordination residues (Figure 3.20) from each side (157 atoms), and small the small model contains the first coordinated residues (Figure 3.20) bound to Cu(II) (78 atoms). All terminal ends were capped with CH₃, as shown in Figure 3.20.

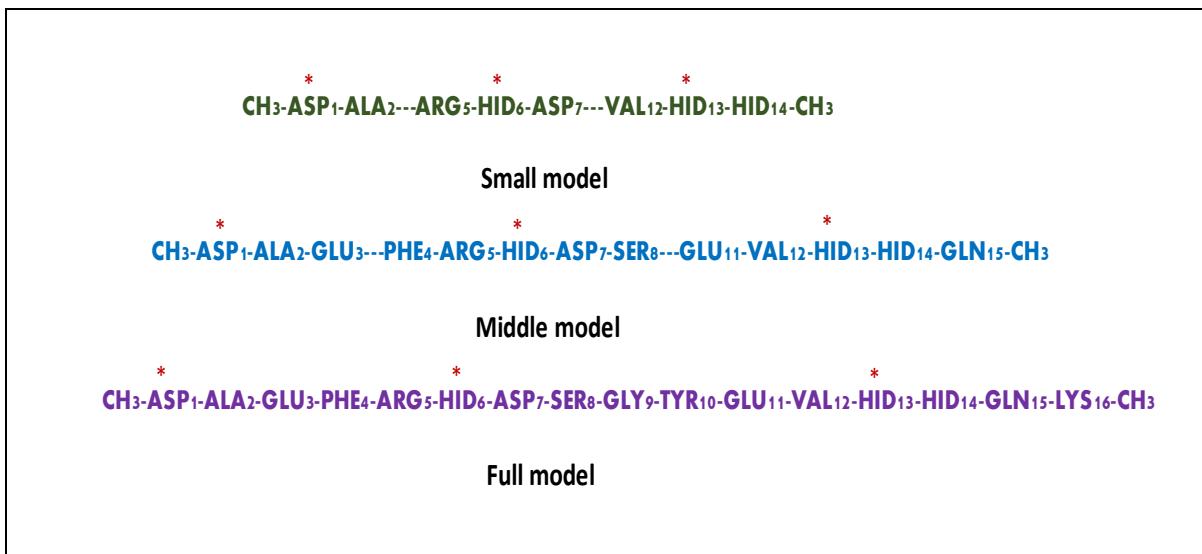


Figure 3.20: The sequence of Small, Middle, and Full models of A β 16-Cu(II). A copper ion was coordinated to one of the nitrogen atom of His 13, another from the imidazole ring of His 6, NH₂ from Asp1, and to the either side-chain or backbone oxygen of Asp1.

Having modelled the different systems, all geometries were then minimized using ligand field molecular mechanics (LFMM).²³ The small versions of A, B, C, and D were optimized at the B3LYP-D2/def2-SVP^{10,11} level and was used for benchmarking the performance of the former. The def2-SVP^{12, 13} basis set used at the conductor-like polarizable continuum model (CPCM). The geometry optimization was completed and the stationary points found, establishing that all geometries and energies were at the local minima. The A,B,C, and D structures of small, middle, and full size, were also optimized using the semi-empirical GFN2-xTB method, utilizing the generalized Born-surface area (GB-SA) implicit solvent model. Further geometry optimization was performed, using PBEh-3C, for comparative purpose with DFT and GFN2-xTB methods. B3LYP-D2/def2-SVP energy correction using CPCM implicit solvent was applied on semi-empirical optimized geometries. All DFT calculations were performed using the Gaussian09³⁶ and ORCA³⁷ software. Semi-empirical calculations were performed via the XTB program.³⁸

3.2.3 Result and discussion

The conformation search generated 9311 conformers, and the lowest energy conformer has value of -322.2 kcal.mol⁻¹. GFN2-xTB and PBEh-3C gave stable fully optimized geometries with reasonable results. The B3LYP-D2/def2-SVP approach yielded stable optimized geometries.

However, this method can be only applied on small systems, so it is not suitable for large molecules, such as proteins, especially in the presence of heavy ions. Nevertheless, we can use it for single point energy comparison. Thus, the B3LYP-D2/def2-SVP optimization was performed on small fragments only and was used as a reference for the semi-empirical methods. The optimization calculation of this small fragment took approximately 9 hours. Also, because of the geometries at the semi-empirical level being of acceptable accuracy, the energy correction using the higher level model such as DFT is desirable and can give reasonable results. The energy calculated from all the methods was compared relative to (A) energy. In general, the results of the calculations indicate that geometry A is less stable in all methods for this size of fragment, as shown in Figure 3.21 and Table 3.4.

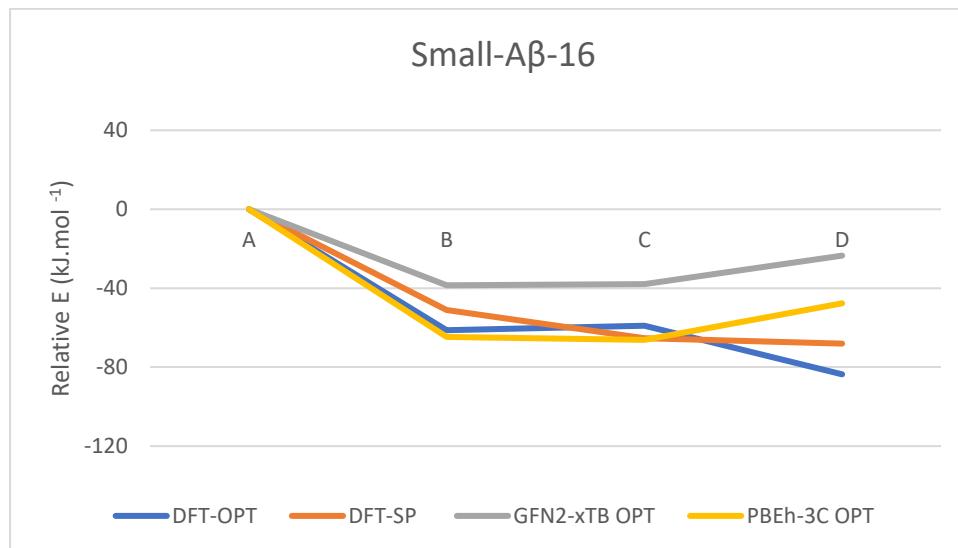


Figure 3.21: Relative energy for small size fragments of A β 16 for A, B, C, and D complexes relative to (A) via GFN2-xTB and PBEh-3C optimization and DFT (B3LYP-D2/def2-SVP) single point energy correction.

Table 3.4: The relative energies of small-size fragments of Cu(II)-A β 16 molecule between B3LYP-D2/def2-SVP-Opt/SP, GFN2-xTB-OPT, and PBEh-3C.

Small	B3LYP-D2-Opt (kJ.mol $^{-1}$)	B3LYP-D2-SP (kJ.mol $^{-1}$)	GFN2-xTB-Opt (kJ.mol $^{-1}$)	PBEh-3C-Opt (kJ.mol $^{-1}$)
A	0.00	0.00	0.00	0.00
B	-61.25	-51.09	-38.52	-64.70
C	-58.94	-65.28	-37.86	-66.19
D	-83.63	-68.07	-23.56	-47.59

The DFT optimization of small models, show the relative energies, which may be ordered from lowest to highest: D < B < C < A; with the difference between B and C being at just -2.30 kJ.mol $^{-1}$. The GFN2-xTB values follow a similar trend for A, B, and C, with the exception being complex D, which tends to have a higher energy when compared to its value, calculated using B3LYP-D2/def2-SVP.

By checking the optimization steps of structure D, the two rings, from the bond formed originally between negatively charge O atom of Asp1 side-chain and the positively polar carbon, were found to be broken, resulting in the formation of the most stable square planar geometry. Furthermore, the backbone oxygen and nitrogen of Asp1 chelate to the copper, addition to His6 and 13, equatorially, with a fifth bond coming from one of the Asp1 side-chain oxygens, binding axially, as shown in Figure 3.22. This leads to reservations, regarding whether this type of bond shown in D, between C(+ δ) and O(- δ), occurs naturally. Also, the presence of a hydrogen bond, with length 1.84 Å, between the His 6/13 backbone oxygen and NH₂ from Asp1 amino acid could result in D being a more stable structure. GFN2-xTB calculations also show breakage of this bond during optimization, with the final structure looking very similar to the one obtained from DFT (RMSD= 0.8 Å), occurring in a square planar fashion, with a fifth axial interaction from the side-chain oxygen exactly to that seen by DFT (Figure 3.22). Here, there are two hydrogen bonds (HB), one is the same as that form in DFT between backbone O and NH₂, and the second forms between CH₃ of capped terminus and COO from another terminus. The greatest difference in geometry between GFN2-xTB and DFT optimization was seen in structures B and C, after looking at their RMSD values, as shown in Table 3.5.

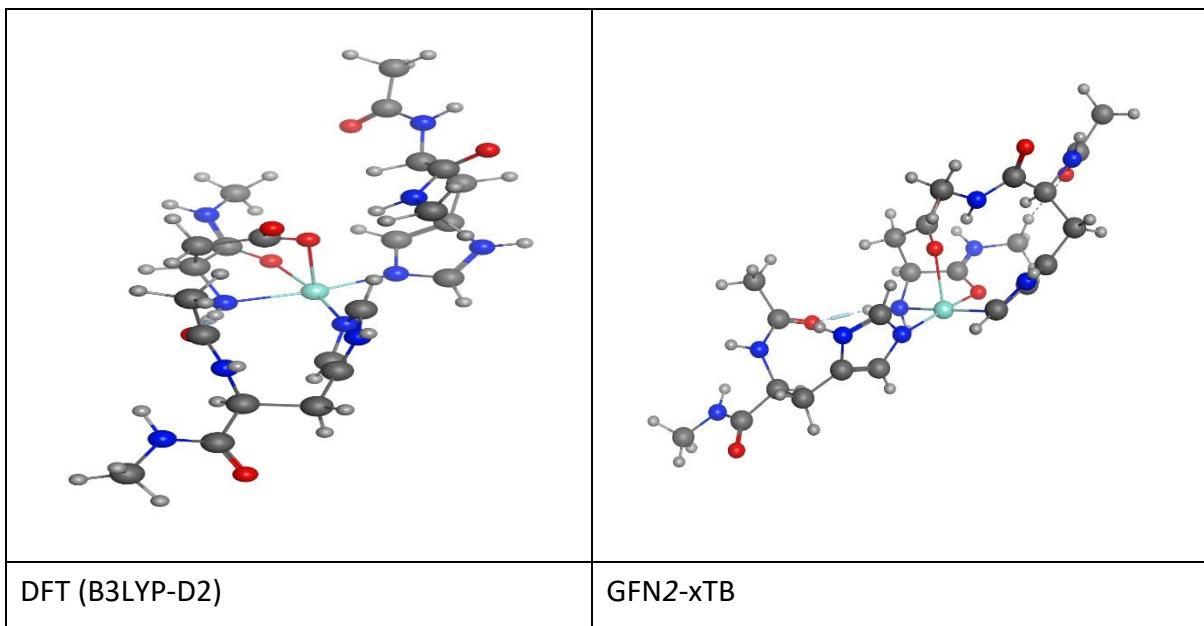


Figure 3.22: An optimized structure D of small-size molecule via DFT (B3LYP-D2/def2-SVP) and GFN2-xTB.

Table 3.5: RMSD between DFT and GFN2-xTB geometries generated of four small structures (Å).

Structure	RMSD
A	1.42
B	2.97
C	2.82
D	0.81

The DFT (B3LYP-D2/def2-SVP) correction of the GFN2-xTB geometry for the middle-size system, comes in great agreement with DFT correction, presenting similar trends of relative energy order, but showing high energy value for D ($45.75 \text{ kJ}\cdot\text{mol}^{-1}$). The GFN2-xTB optimization energy also shows acceptable agreement, when compared to the geometry of A, B, and C with B3LYP-D2/def2-SVP at PBEh-3C (Table 3.6). The GFN2-xTB optimized structure of A shows distorted square planar geometries, with additional side-chain oxygen coordination from Asp1 residues at the axial position. Geometry B shows square planar structure with two additional axial bonds (up and down) both of them are from Asp1 and 7 backbone oxygens which could lead to stabilize this geometry. However, geometry C forms trigonal bipyramidal geometry where three nitrogen from His6, 13, and Asp1 are equatorial

and two axial oxygens from Asp1 and 7 side-chains are axial. On other hand, geometry D shows only four coordination number. The relative energy for geometry D shows a deviation from the trend, by overestimating the energy from GFN2-xTB -1.83 kJ.mol⁻¹, compared to -52.36 kJ.mol⁻¹ for DFT at PBEh-3C level. The resulting geometry is forming tetrahedral, formed from four coordination number, unlike the five bonds seen in small D. This is believed to be the reason system D is the least stable conformer. The ring that forms via the carbon-oxygen bond is broken here.

Table 3.6: The Relative energy values comparison of middle-size of Cu(II)-A β 16 molecule between B3LYP-D2/def2-SVP-SP, GFN2-xTB-OPT, and PBEh-3C.

Middle	GFN2-xTB-Opt (kJ.mol ⁻¹)	B3LYP-D2-SP at xTB (kJ.mol ⁻¹)	PBEh-3C-Opt (kJ.mol ⁻¹)	B3LYP-D2-SP at PBEh- 3C (kJ.mol ⁻¹)
A	0.00	0.00	0.00	0.00
B	-43.30	-59.08	-101.97	-71.77
C	-24.81	-46.60	-21.29	-7.69
D	-1.83	45.75	-110.35	-52.36

The full 16-residue peptide that is closest in size to the real A β peptide, shows agreement between the relative energies from GFN2-xTB and DFT with an energy order from lowest to the highest: C < A < D < B for both methods, as shown in Table 3.7. The geometry optimization via GFN2-xTB of full peptide A, results in a four-bond equatorial square planar, with an extra axial bond from the carboxylate of Asp7, as shown in Figure 3.23. Geometry B shows just four bonds, with a distorted square planar geometry of 3N1O. However, geometry C shows square planar geometry but with different coordination, containing N (His13), NH₂ (Asp1), and carboxylate of Asp1 and Asp7, while His6 is not involved in the coordination anymore, after the bond distance of Cu(II)-His6 increased to 4.1 Å. Nevertheless, this does not invalidate the GFN2-xTB method, as DFT optimization was not performed on full size model to see if could give the same result specifically the relative energy looks sensible when compared with DFT-SP. There is no change in model D coordination, maintaining its square planar geometry, as well as the ring that forms between the carbon-oxygen bond, mentioned in the methodology section.

The PBEh-3C method displays comparable results in terms of the order of relative energies obtained with B3LYP-D2/def2-SVP correction for the middle and full-size geometries, but slightly overestimates the structure of D in the small model, similar to the results from GFN2-xTB (Figure 3.21). However, the calculation of this method on such a system is not as efficient, requiring a great computational time, even with multiple number of parallel processes (16), making it no different to the computational expense from DFT. Overall, the GFN2-xTB examination on small and large models provides promising results on the efficiency of the method. Therefore, energy and geometry of a large system can be measured using faster and less costly methods, with high levels of accuracy. All numerical data of the energies from the different methods, are summarised in Tables 3.5, 3.6, and 3.7.

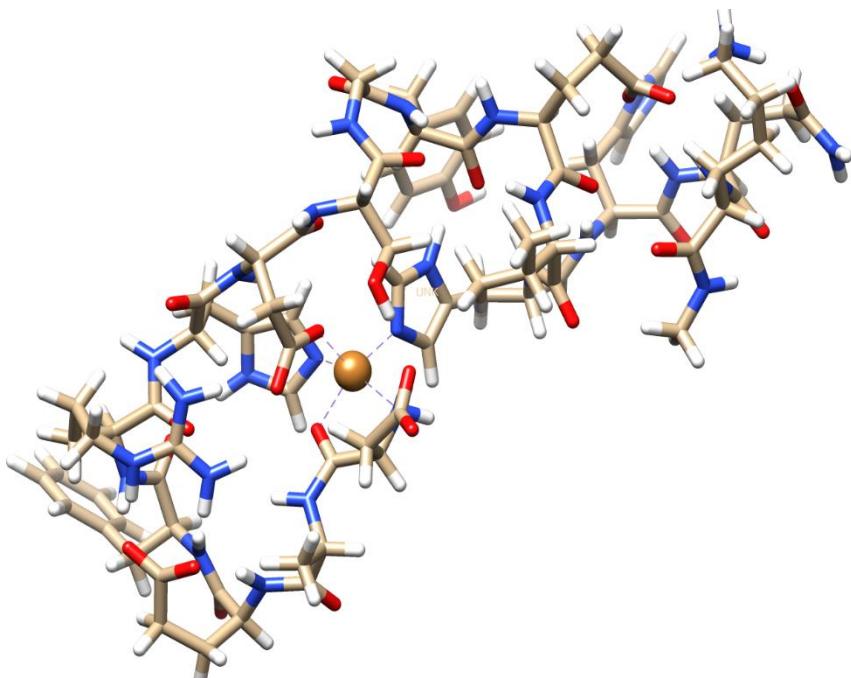


Figure 3.23: The full A β -16 peptide, model (A) optimized via GFN2-xTB.

Table 3.7: The relative energy values comparison of full-size of Cu-A β 16 molecule between B3LYP-D2/def2-SVP -SP, GFN2-xTB-OPT, and PBEh-3C.

Full	B3LYP-D2-SP (kJ.mol ⁻¹)	GFN2-xTB-Opt (kJ.mol ⁻¹)	PBEh-3C-Opt (kJ.mol ⁻¹)
A	0.00	0.00	0.00
B	144.27	20.83	79.46

Full	B3LYP-D2-SP (kJ.mol ⁻¹)	GFN2-xTB-Opt (kJ.mol ⁻¹)	PBEh-3C-Opt (kJ.mol ⁻¹)
C	-44.68	-82.03	-95.73
D	44.81	5.03	-24.50

3.2.4 Conclusion

In this section, simulation data on the copper(II)-A β binding sites has been reported, including the optimization, single point energy and geometry coordination findings of four geometries simulated with the GFN2-xTB and PBEh-3C calculation methods, and thus compared to the results from the DFT (B3LYP-D2/def2-SVP) method. The main purpose here, has been to examine the faster semi-empirical methods, which are able to predict the energy and geometry of large size molecules at a better computational efficiency, compared to DFT functional used. The GFN2-xTB method, has thus been found to be the best method to explore this system. The calculations from both GFN2-xTB and PBEh-3C methods, came in good agreement with B3LYP-D2/def2-SVP. Even though PBEh-3C proved to be a capable method of providing high quality results for Cu(II)-amyloid beta, the fast calculation time of GFN2-xTB for reproducing energy and geometry close to DFT, set this method apart as the better candidate for the simulation of the system tested here, over the other methods examined.

3.3 References

- 1 D. H. Schlesinger et al., Growth-modulating serum tripeptide is glycyl-histidyl-lysine, *Experientia*, 1977, **33**, 324–325.
- 2 L. Pickart et al., Tripeptide in human serum which prolongs survival of normal liver cells and stimulates growth in neoplastic liver, *Nat. New Biol.*, 1973, **243**, 85–7.
- 3 M. Rozga et al., Human serum albumin coordinates Cu(II) at its N-terminal binding site with 1 pM affinity, *J Biol Inorg Chem*, 2007, **12**, 913–918.
- 4 C. Hureau et al., X-ray and Solution Structures of CuIIGHK and CuIIDAHK Complexes: Influence on Their Redox Properties, *Chem. - A Eur. J.*, 2011, **17**, 10151–10160.
- 5 S. J. Lau et al., The interaction of copper(II) and glycyl-L-histidyl-L-lysine, a growth-modulating tripeptide from plasma, *Biochem. J.*, 1981, **199**, 649–656.
- 6 C. M. Perkins et al., The structure of a copper complex of the growth factor glycyl-L-histidyl-L-lysine at 1.1 Å resolution, *Inorganica Chim. Acta*, 1984, **82**, 93–99.
- 7 J. H. Freedman et al., Structure of the Glycyl-L-histidyl-L-lysine-Copper(II) Complex in Solution, *Biochemistry*, 1982, **21**, 4540–4544.
- 8 A. Siméon et al., The tripeptide-copper complex glycyl-L-histidyl-L-lysine-Cu²⁺ stimulates matrix metalloproteinase-2 expression by fibroblast cultures, *Life Sci.*, 2000, **67**, 2257–2265.
- 9 H. P. Ehrlich, in *Symposium on collagen and skin repair*, Reims, France, 1991.
- 10 M. C. McCormack et al., The Effect of Copper Tripeptide and Tretinoin on Growth Factor Production in a Serum-Free Fibroblast Model, *JAMA Facial Plast. Surg.*, 2001, **3**, 28–32.
- 11 L. Pickart et al., The Human Tripeptide GHK-Cu in Prevention of Oxidative Stress and Degenerative Conditions of Aging: Implications for Cognitive Health, *Oxid. Med. Cell. Longev.*, 2012, **2012**, 1–8.
- 12 A. Schirer et al., Similarities and differences of copper and zinc cations binding to biologically relevant peptides studied by vibrational spectroscopies, *J. Biol. Inorg. Chem.*, 2017, **22**, 581–589.
- 13 M. A. Kremennaya et al., X-ray spectroscopy study of the Cu(II)GHK peptide complex, *J. Struct. Chem.*, 2017, **58**, 1213–1219.
- 14 A. Trapaidze et al., Thermodynamic study of Cu²⁺ binding to the DAHK and GHK peptides by isothermal titration calorimetry (ITC) with the weaker competitor glycine, *J. Biol. Inorg. Chem.*, 2012, **17**, 37–47.
- 15 C. Hureau et al., X-ray and solution structures of Cu IIGHK and Cu IIDAHK complexes: Influence on their redox properties, *Chem. - A Eur. J.*, 2011, **17**, 10151–10160.
- 16 L. Guilloréau et al., Redox chemistry of copper-amyloid-beta: The generation of hydroxyl radical in the presence of ascorbate is linked to redox-potentials and aggregation state, *CHEMBIOCHEM*, 2007, **8**, 1317–1325.

- 17 R. Peverati et al., M11-L: A Local Density Functional That Provides Improved Accuracy for Electronic Structure Calculations in Chemistry and Physics, *J. Phys. Chem. Lett.*, 2011, **3**, 117–124.
- 18 S. Dasari et al., Cisplatin in cancer therapy: Molecular mechanisms of action, *Eur. J. Pharmacol.*, 2014, **740**, 364–378.
- 19 P. Labute, LowModeMD—Implicit Low-Mode Velocity Filtering Applied to Conformational Search of Macrocycles and Protein Loops, *J. Chem. Inf. Model.*, 2010, **50**, 792–800.
- 20 D. M. Ferguson et al., A new approach to probing conformational space with molecular mechanics: random incremental pulse search, *J. Am. Chem. Soc.*, 2002, **111**, 4371–4378.
- 21 W. D. Cornell et al., A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules, *J. Am. Chem. Soc.*, 1995, **117**, 5179–5197.
- 22 S. T. Mutter et al., Benchmarking of copper(II) LFMM parameters for studying amyloid-peptides, *J. Biomol. Struct. Dyn.*, 2018, **36**, 1145–1153.
- 23 R. J. Deeth et al., Molecular modelling for transition metal complexes: Dealing with d-electron effects, *Coord. Chem. Rev.*, 2009, **253**, 795–816.
- 24 M. Molecular Operating Environment (MOE) (2014) Chemical Computing Group Inc. et al., 2013.
- 25 C. Lee et al., Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density, *Phys. Rev. B*, 1988, **37**, 785–789.
- 26 S. Grimme, Semiempirical GGA-type density functional constructed with a long-range dispersion correction, *J. Comput. Chem.*, 2006, **27**, 1787–1799.
- 27 A. Schäfer et al., Fully optimized contracted Gaussian basis sets for atoms Li to Kr, *J. Chem. Phys.*, 1992, **97**, 2571–2577.
- 28 F. Weigend et al., Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy, *Phys. Chem. Chem. Phys.*, 2005, **7**, 3297–3305.
- 29 J. P. Perdew et al., Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.
- 30 J. P. Perdew et al., Generalized Gradient Approximation Made Simple [Phys. Rev. Lett. 77, 3865 (1996)], *Phys. Rev. Lett.*, 1997, **78**, 1396–1396.
- 31 J. P. Perdew et al., Rationale for mixing exact exchange with density functional approximations, *J. Chem. Phys.*, 1996, **105**, 9982–9985.
- 32 C. Adamo et al., Toward reliable density functional methods without adjustable parameters: The PBE0 model, *J. Chem. Phys.*, 1999, **110**, 6158–6170.
- 33 S. Grimme et al., A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, *J. Chem. Phys.*, 2010,

132, 154104.

- 34 Y. Zhao et al., A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions, *J. Chem. Phys.*, 2006, **125**, 194101.
- 35 J.-D. Chai et al., Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6615–6620.
- 36 M. J. Frisch et al., GAUSSIAN09, <http://www.gaussian.com/>.
- 37 F. Neese, The ORCA program system, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2012, **2**, 73–78.
- 38 C. Bannwarth et al., GFN2-xTB - An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions, *J. Chem. Theory Comput.*, 2019, **15**, 1652–1671.
- 39 W. Humphrey et al., VMD: Visual molecular dynamics, *J. Mol. Graph.*, 1996, **14**, 33–38.
- 40 P. Pracht et al., Automated exploration of the low-energy chemical space with fast quantum chemical methods, *Phys. Chem. Chem. Phys.*, 2020, **22**.
- 41 S. Grimme, Exploration of Chemical Compound, Conformer, and Reaction Space with Meta-Dynamics Simulations Based on Tight-Binding Quantum Chemical Calculations, *J. Chem. Theory Comput.*, 2019, **15**, 2847–2862.
- 42 S. Grimme et al., Fully Automated Quantum-Chemistry-Based Computation of Spin-Spin-Coupled Nuclear Magnetic Resonance Spectra, *Angew. Chemie Int. Ed.*, 2017, **56**, 14763–14769.
- 43 J.-P. Ryckaert et al., Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *J. Comput. Phys.*, 1977, **23**, 327–341.
- 44 N. M. O'boyle et al., cclib: A library for package-independent computational chemistry algorithms, *J. Comput. Chem.*, 2008, **29**, 839–845.
- 45 S. Grimme et al., Consistent structures and interactions by density functional theory with small atomic orbital basis sets, *J. Chem. Phys.*, 2015, **143**, 054107.
- 46 J. W. Karr et al., Amyloid- β Binds Cu²⁺ in a Mononuclear Metal Ion Binding Site, *J. Am. Chem. Soc.*, 2004, **126**, 13534–13538.
- 47 C. D. Syme et al., Copper Binding to the Amyloid- β (A β) Peptide Associated with Alzheimer's Disease, *J. Biol. Chem.*, 2004, **279**, 18169–18177.
- 48 C. C. Curtain et al., Alzheimer's Disease Amyloid- β Binds Copper and Zinc to Generate an Allosterically Ordered Membrane-penetrating Structure Containing Superoxide Dismutase-like Subunits, *J. Biol. Chem.*, 2001, **276**, 20466–20473.

4 Accelerated Molecular Dynamics Simulation of metal ions II (Zn, Fe, and Cu) Binding to N-Terminus of Amyloid- β

4.1 General chapter introduction

As mentioned in a chapter one, a hypothesis on Alzheimer's disease involves the role of transition metals, most markedly copper, zinc, and iron.¹⁻³ All three ions have been identified to bind to A β through the N-terminal sequence (1-16 residues) affecting the structure and folding of the peptide, which may in turn impact aggregation properties.⁴⁻¹⁰ The binding sites of these ions are highly influential alongside the type of metals involved and cannot be ignored due to their effect on structure. Properties such as hydrogen bonds and secondary structure are affected as result of these interactions in addition to the folding of the peptide which reflects the degree of stability leading to aggregation. Different experimental and theoretical studies have investigated A β -M interactions and proposed binding sites. However, a comparative study of different binding modes for all three metals via theoretical study can be beneficial and complementary.

Accelerated molecular dynamics (aMD)¹¹ is an improved sampling algorithm built to solve the timescale difference between conventional MD methods and biological processes. The idea of this method is to boost the conformational sampling efficiency through adding a bias potential to the original potential. Consequently, transitions between different states of the system are accelerated, so that biological systems stuck between two states in potential energy landscape that were previously inaccessible via MD are now available to be sampled via aMD. This method allows conformational exploration to be more efficient and accessible across the potential energy surface compared to normal MD.^{12,13} This method allows conformational exploration to be more efficient and accessible across the potential energy surface compared to normal MD, but they require proper reweighting for free energy calculation. So, the original free energy profiles of functional biomolecules can then be recovered. In this work, Maclaurin series (MC) is the reweighting algorithm which is an approximation to the exponential Boltzmann factor used to reweight the aMD trajectories for all A β complexes. There are numerous studies which have used aMD method to various biomolecular and metallopeptide systems successfully.¹⁴⁻¹⁸ In this chapter, aMD is used to

simulate the different binding modes of the N-terminal A β sequence bound to Zn, Fe and Cu (II) to probe the effect of metal coordination on structure and dynamics via different binding sites.

4.2 Computational Methods

All simulations were performed within the AMBER16 package.¹⁹ Parameters suitable for AMBER-style forcefields were calculated using the metal centre parameter builder (MCPB.py) package²⁰ from the B3LYP/6-31G(d)²¹ data generated using Gaussian09.²² Harmonic force constants compatible with AMBER simulations for metal-residue coordination were obtained from DFT optimization data via the Seminario²³ method. Atomic charges were generated from DFT electrostatic potential using the restrained electrostatic potential (RESP)^{24–26} fitting scheme. These were combined with parameters from ff14SB forcefields using the LEaP utility.²⁷ The Generalized Born surface area (GBSA) model of implicit solvent, was applied.^{28–30} The AMBER ff14SB forcefield parameter set³¹ was used to construct all 16 amino acid residues. Molecular dynamics simulations were performed using the AMBER16 package within the NVT ensemble. A 2 fs integration time-step was used for all MD simulations. The SHAKE algorithm was applied to restrain all hydrogen-containing bonds³² at 310 K. The system was solvated using the generalised Born solvation model.^{33–35} A cut-off distance of 12 Å was used for the van der Waals and electrostatic interactions. Trajectories analysis was completed using CCPTRAJ v16.16.³⁶

The extended N-terminal A β sequence was built in MOE,³⁷ amidated at C-terminus to mimic the full peptide, but the N-terminus was uncapped and therefore accessible for metal coordination. Metal ions were then added following literature proposals, as summarized by Nasica-Labouze et al.³⁸ The resulting structures were minimised using the ligand field molecular mechanics (LFMM) approach defined by Deeth et al.³⁹ These minimised structures were best described as random coil, which was then used to generate parameters using the MCPB/LEaP procedure.

Conventional molecular dynamics (cMD) were carried out after minimisation. Parameters for the aMD boost potential were taken from these simulations. Once these parameters were established, three independent 200 ns aMD simulations were carried out with random initial atomic velocities, adding the boost parameters generated from prior cMD simulations. Dual-

boost version¹² of aMD was chosen for simulations due to the suitability of it for biomolecular systems. This is based on using both a total boost potential as well as dihedral boost for all atoms in the system as shown in equation 2.27.

Following 50 ns of cMD to allow for equilibration of the system, a further three individual 200 ns aMD simulations were ran with different random seeds (ig=-1). These used the endpoint of the prior cMD as the starting structures for aMD. The trajectories then were combined to generate 600ns total data for each binding mode via Cpptraj. Analysis of root mean square deviation (RMSD), root-mean-square fluctuations (RMSF), radius of gyration (R_g), secondary structure, backbone dihedrals, salt bridges, residue contacts, and 1D free energy landscapes were used to investigate the effects of metal binding on structure and degree of flexibility and compactness of A β . The DBSCAN⁴⁰ clustering algorithm was used with 10 minimum number of points required to form a cluster and 0.8 distance cutoff between points for forming a cluster(minimum distance between clusters).

4.3 Accelerated Molecular Dynamics Simulation of Zn(II) Binding to N-Terminus of Amyloid- β

4.3.1 Introduction

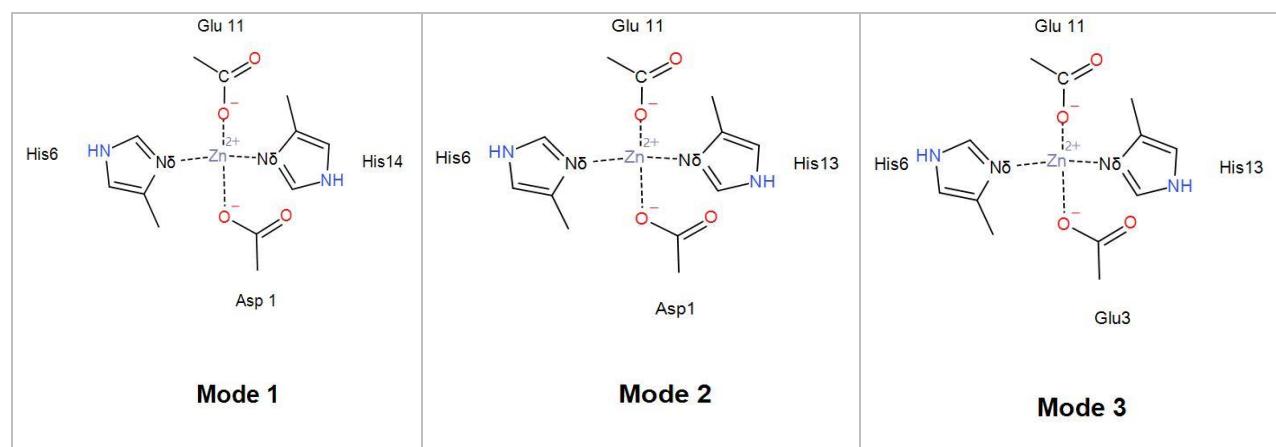
Several laboratory techniques, including scattering and diffraction, magnetic resonance, circular dichroism, etc. have demonstrated in detail the sites and specificities of metal binding and its influence on peptide secondary structure and dynamics.^{8,41–44} Different coordination modes of transition metal ions of Zn(II), Cu(II), and Fe(II) have been proposed, with some articles reviewing them.^{38,45} It has been proven Zn(II) promotes A β aggregation in a wide pH range at slightly acidic to basic condition,^{46,47} whereas Cu(II) is pH dependent and aggregates only at pH 6.0–7.0.⁴⁸ The primary metal-binding sites for Zn(II) found by Raman spectra are the imidazole ring from three amino acids (His6, His13 and His14). This Raman analysis indicated that the peptide accumulates through intermolecular His(N ε)-Zn(II)-His(N ε) bridges in a wide pH range.⁴⁹ A carboxylate of Glu11 was demonstrated to be the primary chelator to the Zn(II), in addition to minor offering by Asp1 and Glu3. Zirah et al. performed NMR to show

that Zn(II) binds to A β 16 through His6 and His14(N_{δ}) as well as His13(N_{ϵ}), and Glu11 carboxylate.¹⁰ Additional NMR studies in water-micelle solution for human A β (1-28), revealed that N_{δ} of His6 and His13 as well as N_{ϵ} of His14 participate in Zn(II) interaction in addition to Asp1 amine, and/or Glu11 COO⁻.⁵⁰

In this section, accelerated molecular dynamics simulations of Zn(II) bound through different binding sites of the N-terminus of truncated A β (1-16) were reported comparing it to the free peptide. Analysis of simulations show that Zn(II) provides considerable rigidity to the peptide, disrupts the secondary structure and pattern of salt bridges observed in the free peptide, and causes closer contact between residues. The evaluation between different binding modes of A β -16 bound to Zn(II) illustrates that variation in given atoms and residues involved in binding affects results in term of secondary structure, clusters, salt bridge, RMSF, and the compactness of A β shape and the peptide aggregation. Free energy landscapes in 1D further underscore the effect of metal coordination to the peptide stability through free energy surfaces in terms of HB count, R_g , end-to-end distance to describe the peptide structure during the dynamics.

4.3.2 Results and Discussion

Accelerated MD simulations were performed using ff14SB for the metal-free A β 16. Seven complexes with Zn(II) bound through different proposed residues/atoms shown in **Error! Reference source not found.** and Table 4.1 are taken from the literature,^{1, 10, 38, 50-55} with binding mode number 6 representing that seen in the structure proposed from NMR experiment.¹⁰



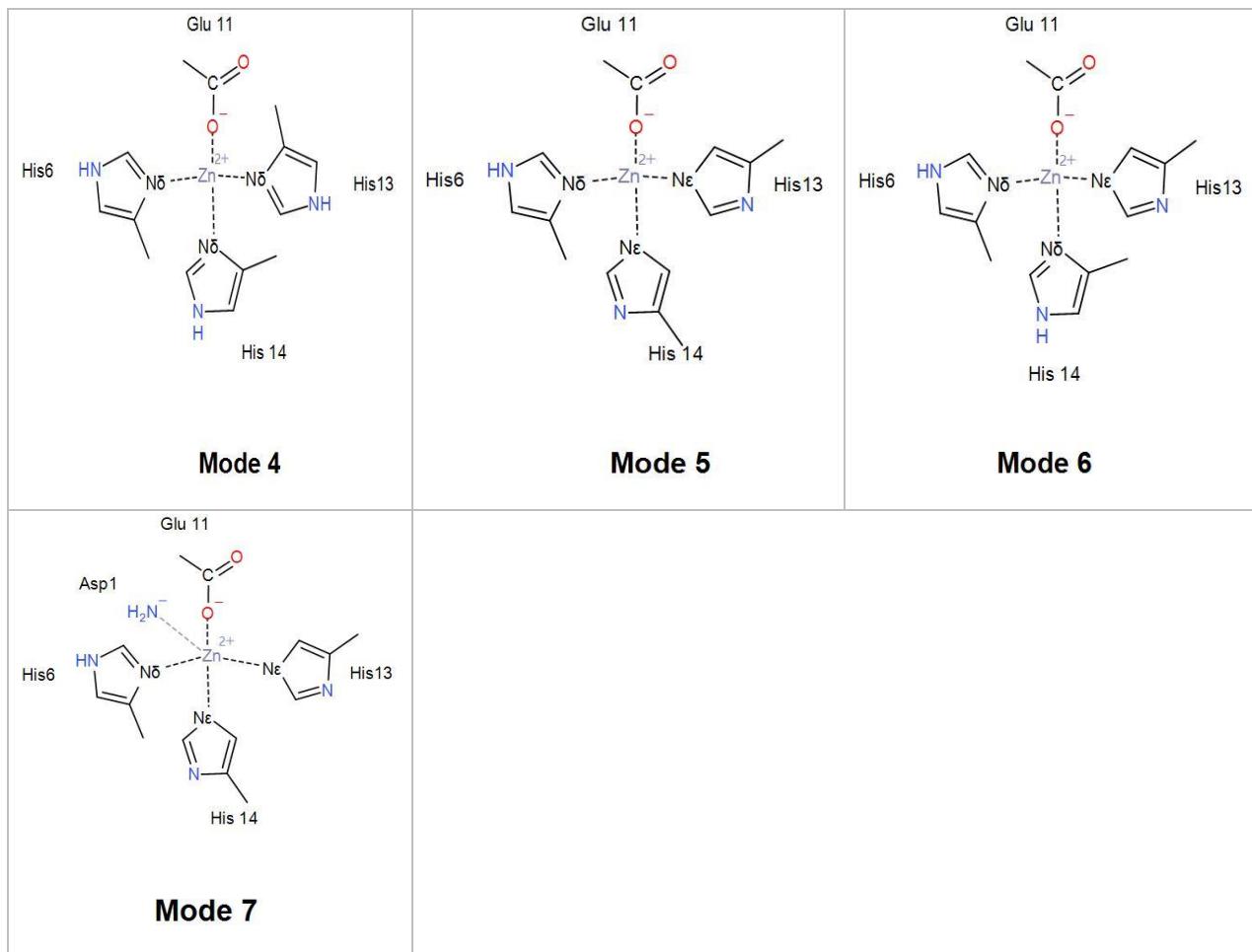


Figure 4.1: Different proposed binding modes of A β (1-16) binding to Zn(II).

Table 4.1: Different proposed binding modes of A β (1-16)binding to Zn(II).

	Coordination
Mode 1	His6 N δ , His14 N δ , Asp1 CO, Glu11 O ϵ
Mode 2	His6 N δ , His13 N δ , Asp1 CO, Glu11 O ϵ
Mode 3	His6 N δ , His13 N δ , Glu3 O ϵ , Glu11 O ϵ
Mode 4	His6 N δ , His13 N δ , His14 N δ , Glu11 O ϵ
Mode 5	His6 N δ , His13 N ϵ , His14 N ϵ , Glu11 O ϵ
Mode 6	His6 N δ , His13 N ϵ , His14 N δ , Glu11 O ϵ
Mode 7	His6 N δ , His13 N ϵ , His14 N ϵ , Glu11 O ϵ , Asp1 N α

Overall structural variations over simulation time are monitored using RMSD for stability measuring and R_g for geometries size. The result is shown in Figure 4.2: and Table 4.2. Backbone RMSD relative to an initial minimized structure over a combined 600 ns of aMD

simulation displays the metal-free peptide explored more diverse and lengthened conformations than Zn-peptides complexes, which generally stay more tightened and compact. The average RMSD of A β 16 is 10.6 Å, with sd = 1.6 Å. This is significantly larger than for any zinc binding modes under testing. Closer inspection indicates that the free peptide experiences greater and more frequent transitions in RMSD, varying between *ca.* 5 and 15 Å, while Zn(II)-A β displays smaller and less frequent fluctuations in RMSD ranged between 2.5 and 5.7 Å. Between all modes, Mode 1 shows the lowest RMSD values (Ave 2.65 Å, Sd 0.37 Å, and Max 4.63 Å), but Mode 6 demonstrates the highest average amount.

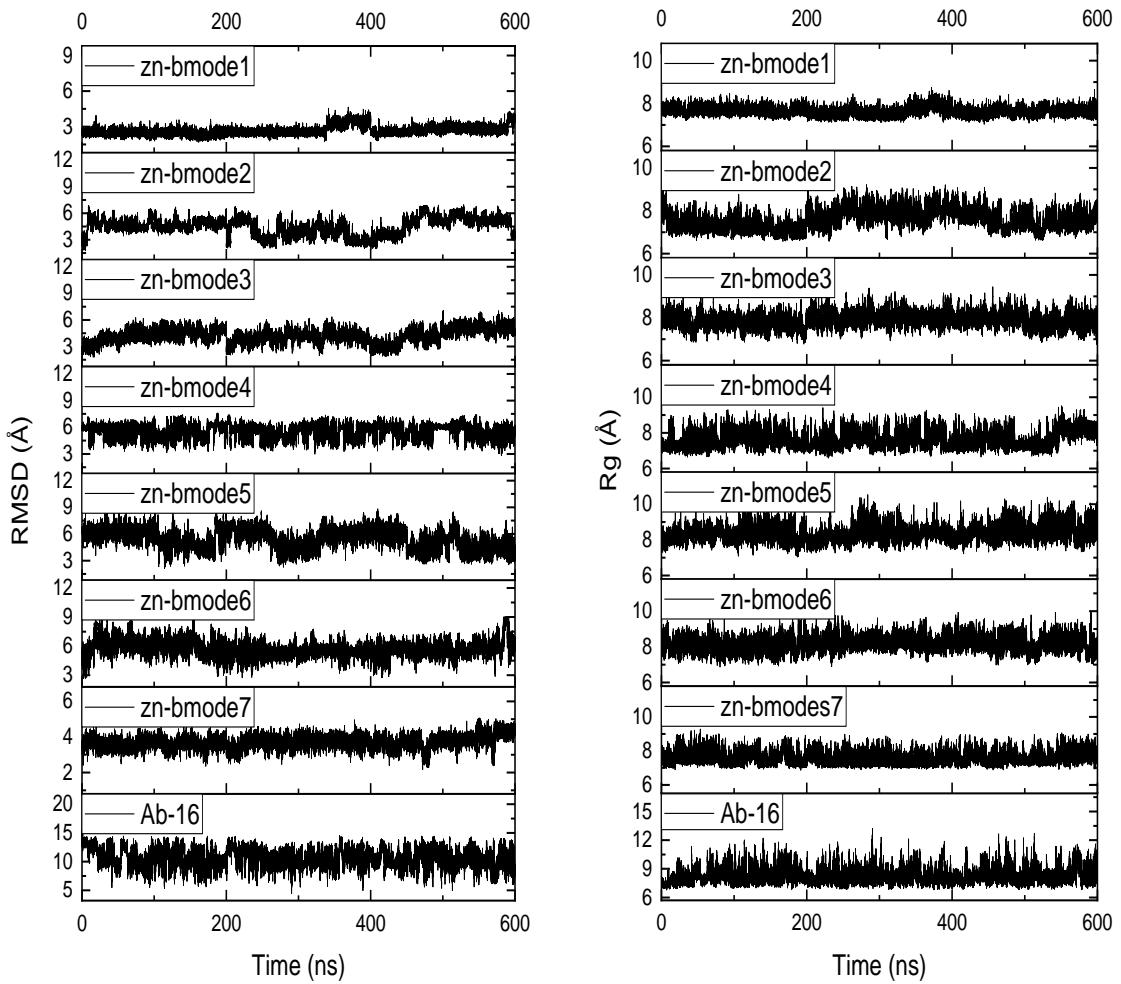


Figure 4.2: Backbone RMSD and R_g values of Zn(II)- Modes and free A β -16. 600 ns of aMD data is reported, the trajectory is combined of the final 200 ns from each of three independent simulations each with different initial velocities.

Table 4.2: Backbone RMSD and R_g of free A β -16 and Zn(II) binding modes (Å).

Zn(II)-A β 16-modes	RMSD				R_g			
	Mean	sd	Min	Max	Mean	sd	Min	Max
Mode 1	2.65	0.37	1.71	4.63	7.63	0.20	7.03	8.75
Mode 2	4.48	0.97	1.78	6.95	7.56	0.44	6.57	9.27
Mode 3	4.21	0.79	1.82	7.06	7.86	0.32	6.76	9.44
Mode 4	5.52	0.70	2.84	7.59	7.62	0.44	6.60	9.51
Mode 5	5.48	1.16	2.11	9.05	8.37	0.43	7.04	10.97
Mode 6	5.77	0.82	2.67	9.03	8.14	0.40	6.88	10.02
Mode 7	3.70	0.35	2.18	5.06	7.58	0.35	6.87	9.29
Aβ16	10.64	1.57	4.45	15.46	8.23	0.77	6.80	13.24

R_g values are also shown in Figure 4.2: and Backbone RMSD and R_g values of Zn(II)- Modes and free A β -16. 600 ns of aMD data is reported, the trajectory is combined of the final 200 ns from each of three independent simulations each with different initial velocities.

, and demonstrate similar results to those seen in RMSD where the free A β -16 has the greatest sd and maximum values (sd 0.77 Å and Max 13.24 Å), which signifies the high mobility and instability of the metal-free peptide which is expected. Conversely, binding modes 5 and 6, in which Zn is bound to residues located on the middle of the peptide chain (Glu11, His6, 13, and14), show the highest mean values that come close or exceed that observed for the metal-free peptide, albeit with lower maximum and standard deviation values. Mean R_g for binding modes 1 to 4 are significantly less than the free peptide, with binding modes 1 and 3 especially tightly packed. The measured mean for binding modes 1-3 of Zn(II)-A β (in which one of the coordination sites is supplied by Asp1 or Glu3) as well as binding mode 4 (where the Zn bound to N δ of His imidazole rings) is in line with the experimental value of 7.4 ± 0.2 Å.¹⁰ For the free peptide, a theoretical calculation for calculating the R_g ^{56,57} from its molecular weight ($M_r = 1995$ g. mol⁻¹) yields $R_g = 9.35$ Å, in sensible agreement with the simulated value of 10.6 ± 1.6 Å. A further experimental study of A β aggregation using hydrodynamic radii (R_h) by size exclusion chromatography and NMR found Zn complexes to be more compact and

structured than copper-bound structures. The R_h value of A β were smaller upon Zn(II) binding (11.4 Å),⁵⁸ which matches well with values found by aMD simulation.

Figure 4.3 shows RMSF of free A β -16 and all seven zinc binding modes. The curves follow a pattern comparable to RMSD and R_g data, with the lowest values for Zn(II) and the largest values for free peptide. This clearly demonstrates the anchoring effect the metal centre displays when comparing to the free peptide to zinc-bound structures. However, binding mode 5 has the highest values seen for Lys16 and amidated cap mobility, which may possibly clarify the high R_g value of this mode. Metal-binding residues (Asp1, His6/14/13, Glu3, and Glu11) exhibit lower RMSF values, indicating how metal coordination restricts the peptide's movement. In general, for all binding modes as well as free peptide, C-terminal (Gln15, Lys16, and amidated cap), as well as Asp7 and Tyr10 residues, generally present the greatest flexibility across all simulations. The N-terminus of free A β -16 has the highest RMSF value, but when Asp1 is bound to metal in binding modes 1 and 2 this shows a smaller level of mobility. Metals themselves, numbered residue 18 in the graph below, have minimal RMSF, with binding modes 1, 3, and 7 displaying particularly low levels of mobility. Those three modes also showed lower RMSD mean values. The lower values seen in mode 7 correspond to the extra bond coming from Asp amine.

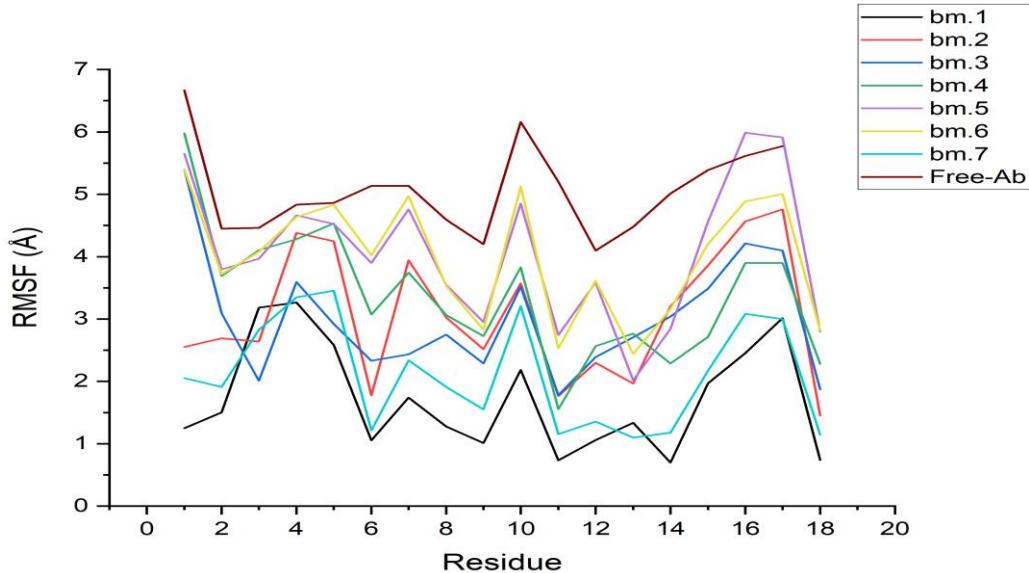
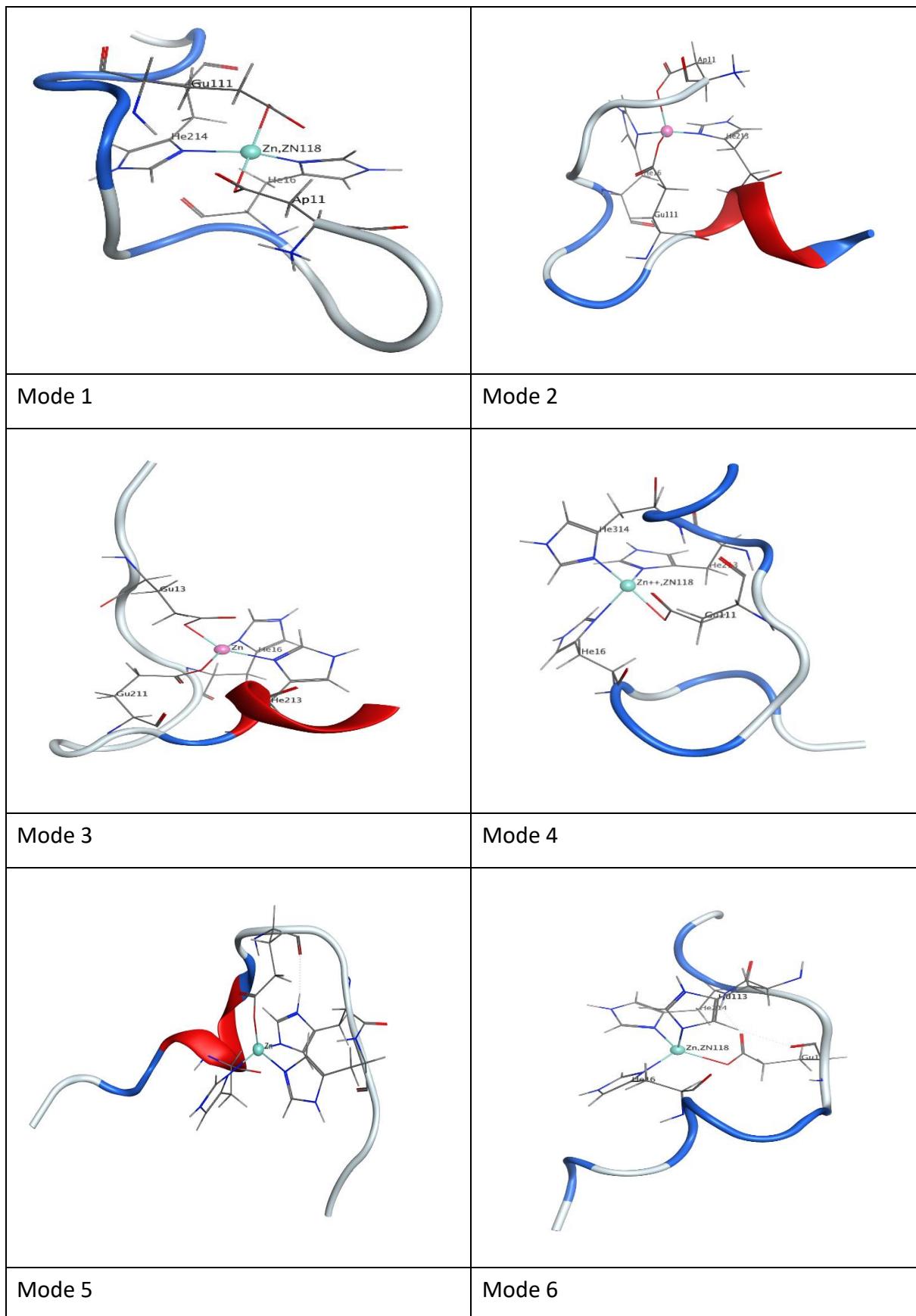


Figure 4.3: RMSF of each residue of free and Zn-bound A β -16. The metal ion itself is denoted residue 18, and the C-terminal amidated cap residue 17.

Clustering analysis of equilibrated trajectories based on backbone dihedrals was employed to identify key structural motifs; ribbon diagrams of the most demonstrative clusters are shown in Figure 4.4: and Table 4.3. Binding mode 1 of Zn(II)-A β , which exhibit the smallest R_g values, forms the lowest number of clusters with 3, of which one includes over 90% of the frames, with the remaining clusters taking no more than 2%. Binding modes 2, 3, 4, and 6 form more clusters (16, 11, 20, and 15, respectively) with populations of 14, 9, 16, and 29%, respectively. The most flexible, mode 5, forms 9 clusters with 60% of the frames in the most populated one. Mode 7 shows 6 clusters with the most populated cluster appearing for 74% of full simulation time. Free A β -16 shows no clusters for which occupancy is more than 1% of total frames, offering further evidence for the dynamic nature of this peptide under these simulation conditions. Overall, the geometries of most populated clusters show a tetrahedral shape around Zn.

Table 4.3: Cluster analysis data for equilibrated trajectories.

<i>Clusters</i>	<i># clusters</i>	<i>Most populated</i>	<i>2nd populated</i>
Mode 1	3	93%	2%
Mode 2	16	14%	12%
Mode 3	11	9%	7%
Mode 4	20	16%	8%
Mode 5	9	60%	18%
Mode 6	15	29%	17%
Mode 7	6	74%	61%



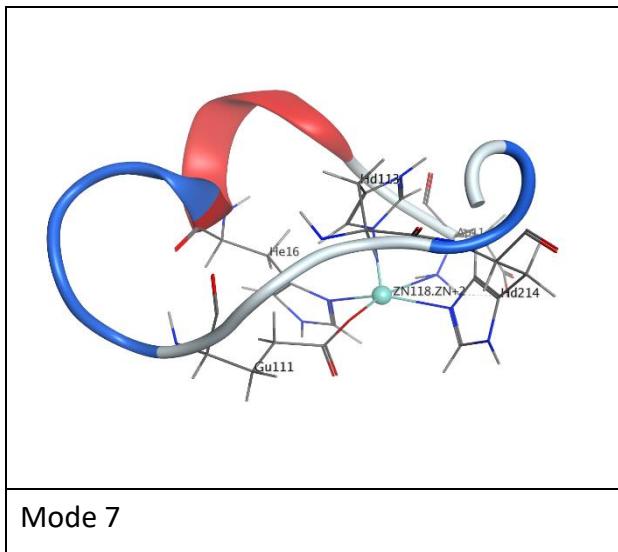
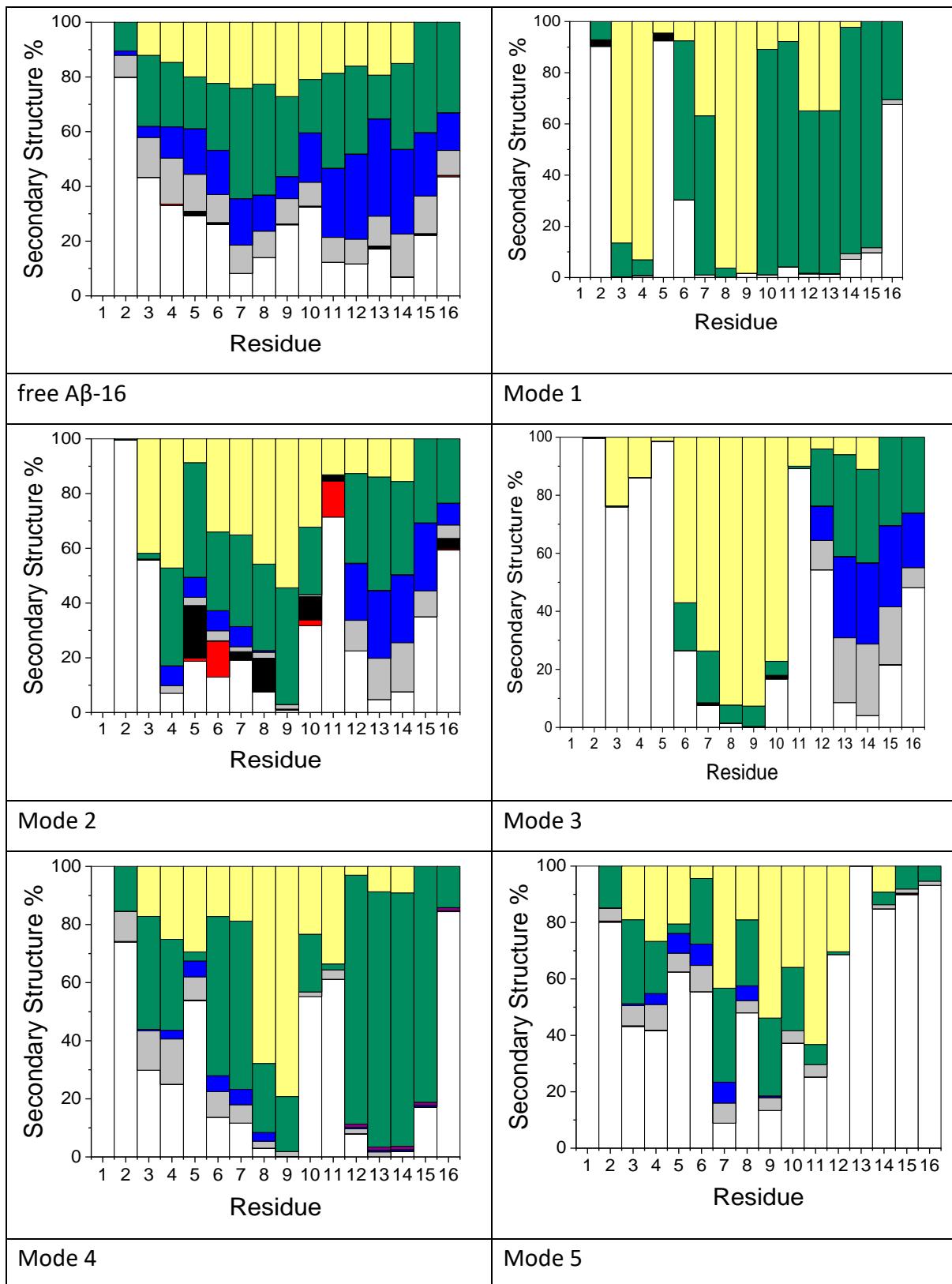


Figure 4.4: Ribbon diagram of the most populated clusters for all binding modes of Zn(II) and free A β .

To monitor the impact on secondary structure of metal coordination to A β peptides, percentage incidence plots for all aMD trajectories were generated displaying categorisation of residues by secondary structure as shown in Figure 4.5: and

. All the secondary structures in this thesis were assigned from Dictionary of Secondary Structures of Proteins (DSSP), which is implemented in cpptraj (Amber tools). This uses both ϕ/ψ angles and H-bond patterns to assign the secondary structures. The analysis shows that A β -16 forms helical (both α - and 3_{10}) conformations in 27% of frames, while Zinc binding modes adopt more bend and turn secondary structures, with a smaller amount of helical structure in most binding modes and very small amounts of β -sheet. Binding mode 1 shows small amounts of helix structures formed and the peptide occupies more turn and bend formations. However, binding mode 2 presents more variation in structure and is the only mode that illustrates 5% of β -sheet folding and higher helical structure 12%, as well as being the only mode investigated that shows parallel β -sheet contrasting to remaining modes. Modes 2, 3, 6 and 7 show heightened percentages of helical structure, but mode 7 has the greatest ratio among all other modes that seen from 1 to 10 residues at 21%. However, the secondary structure percentage by NMR suggested the Zn-A β 16 forms 33% helices, no β -sheets and 67% of bend/coil/turn structure.¹⁰ In contrast, free A β -16 is more varied indicating decreased levels of defined secondary structuring in A β , resulting from Zn binding.



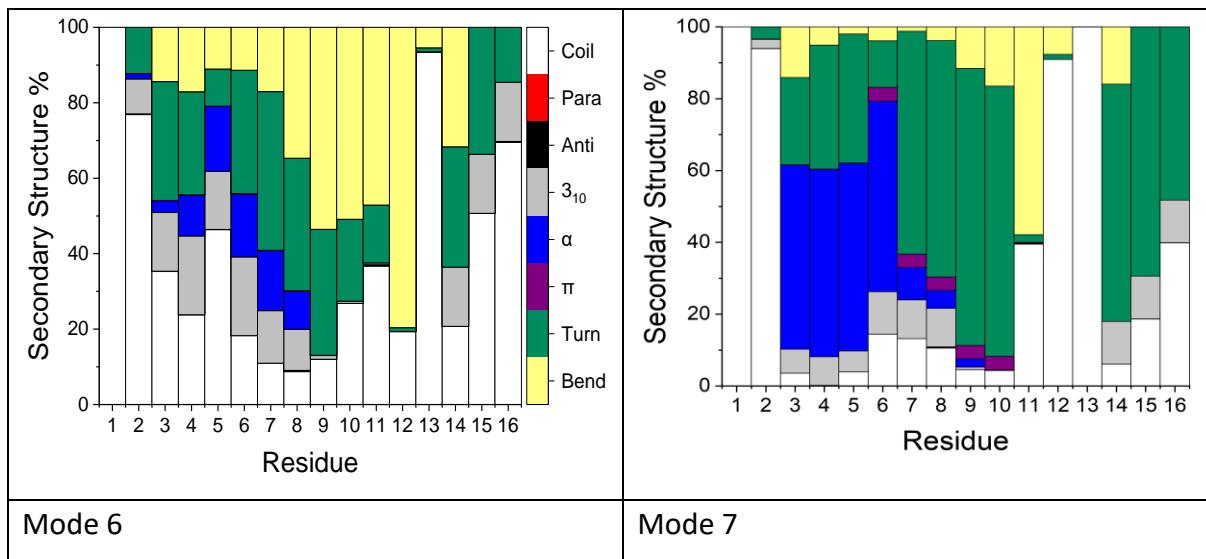


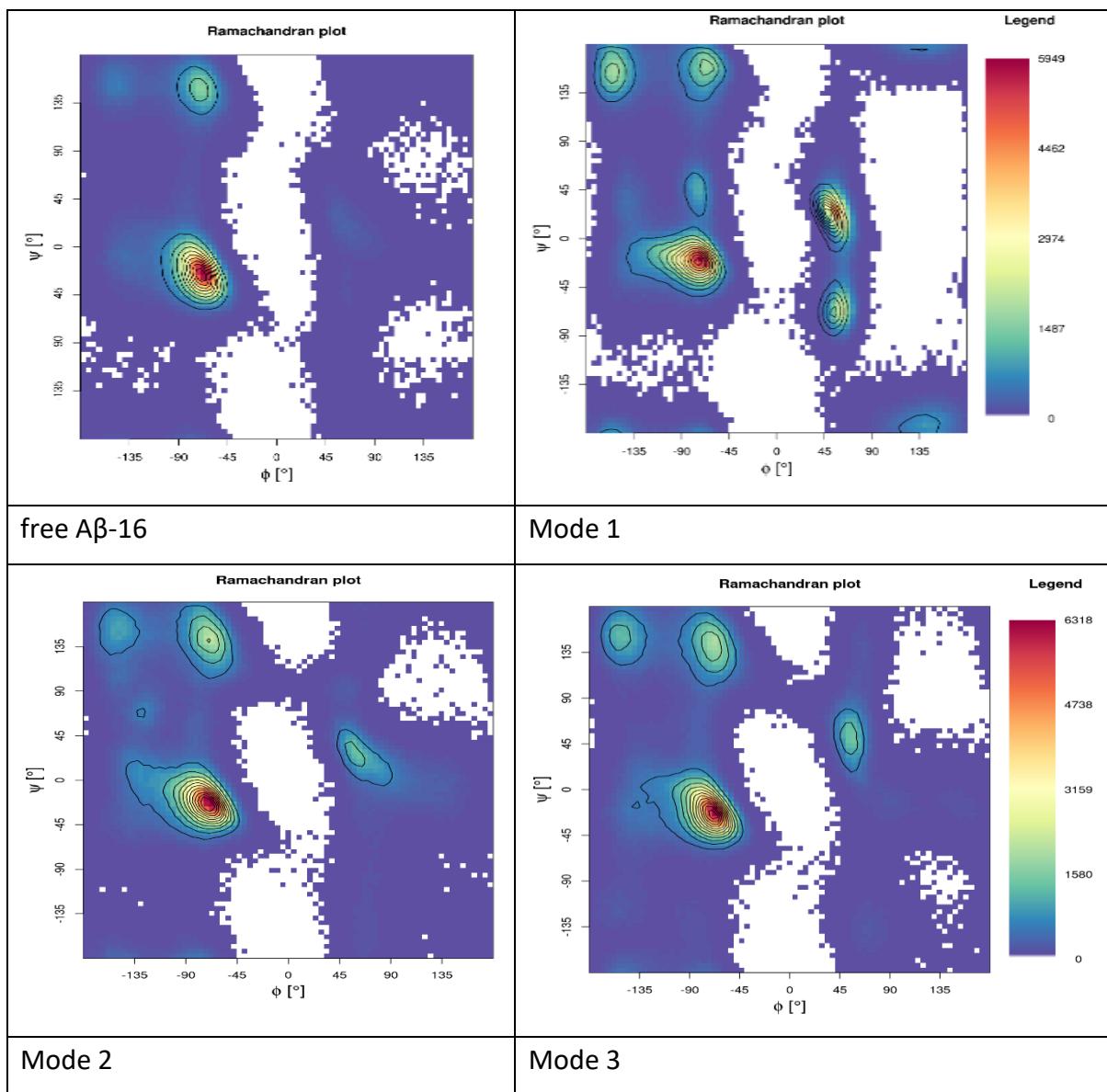
Figure 4.5: Secondary structure for Zn(II)-A β and free A β -16.

Table 4.4: Total secondary structure percentages for each binding modes of Zn-A β (%).

	Helix	Sheet	Other
Mode 1	0.5	0.3	99.2
Mode 2	13.0	5.0	82.0
Mode 3	12.5	0.2	87.4
Mode 4	6.6	0.1	93.3
Mode 5	6.2	0.1	93.7
Mode 6	14.5	0.1	85.4
Mode 7	21.1	0.1	78.9
Aβ	27.2	0.4	72.4

Ramachandran maps were utilized to study the backbone structure as shown in **Error! Reference source not found.**. All simulations show the majority of frames are in the broad general region of right-handed helix. Zn(II)-A β has a greater population of antiparallel β -sheet formed in Mode 5. Additionally, the greatest levels of left-handed helical structures are observed in binding mode number 1. Closer inspection shows that two-thirds of aMD simulations for Zn(II)-A β exist at a high density population at this region. There is an unique conformation of the polypeptide backbone proposed in the literature called the α -pleated sheet that is potentially a popular intermediate conformation in amyloidosis.^{59,60} This α -sheet

is not like the α -helix and β -sheet structures, where all amino acid residues are within a single region of dihedral angles, the α -sheet, instead is made up of alternating dihedrals in the traditional right-handed (α_R) and left-handed (α_L) helical regions of Ramachandran map in conformational space (ϕ^+, ψ^+) and (ϕ^-, ψ^-).⁶¹ It was first proposed by Pauling and Corey before β -sheet structures were characterized. This observation may explain the trend in results reported here. The possibility of α -pleated sheet in these simulations will be explored in more detail in Chapter 5.



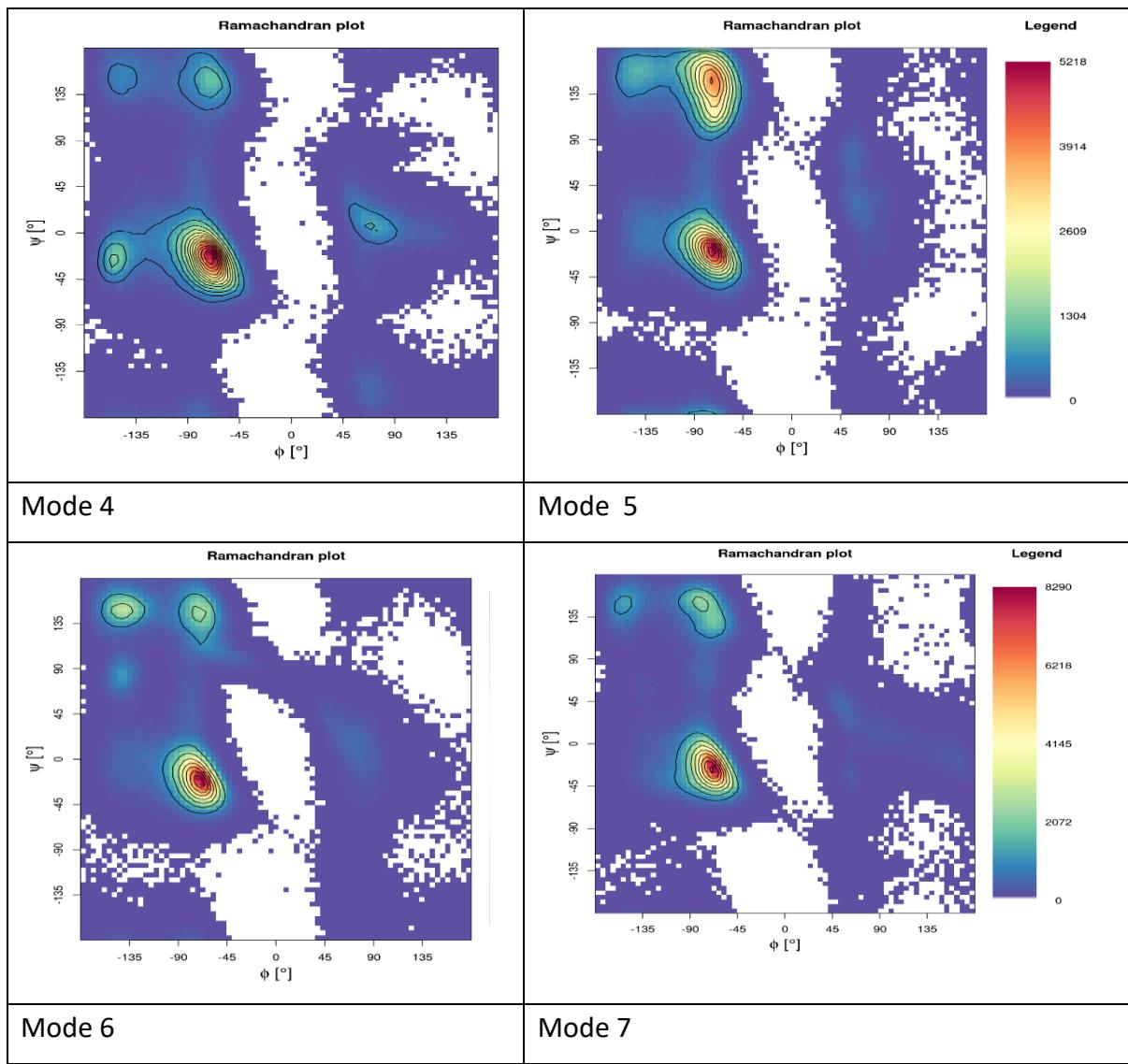
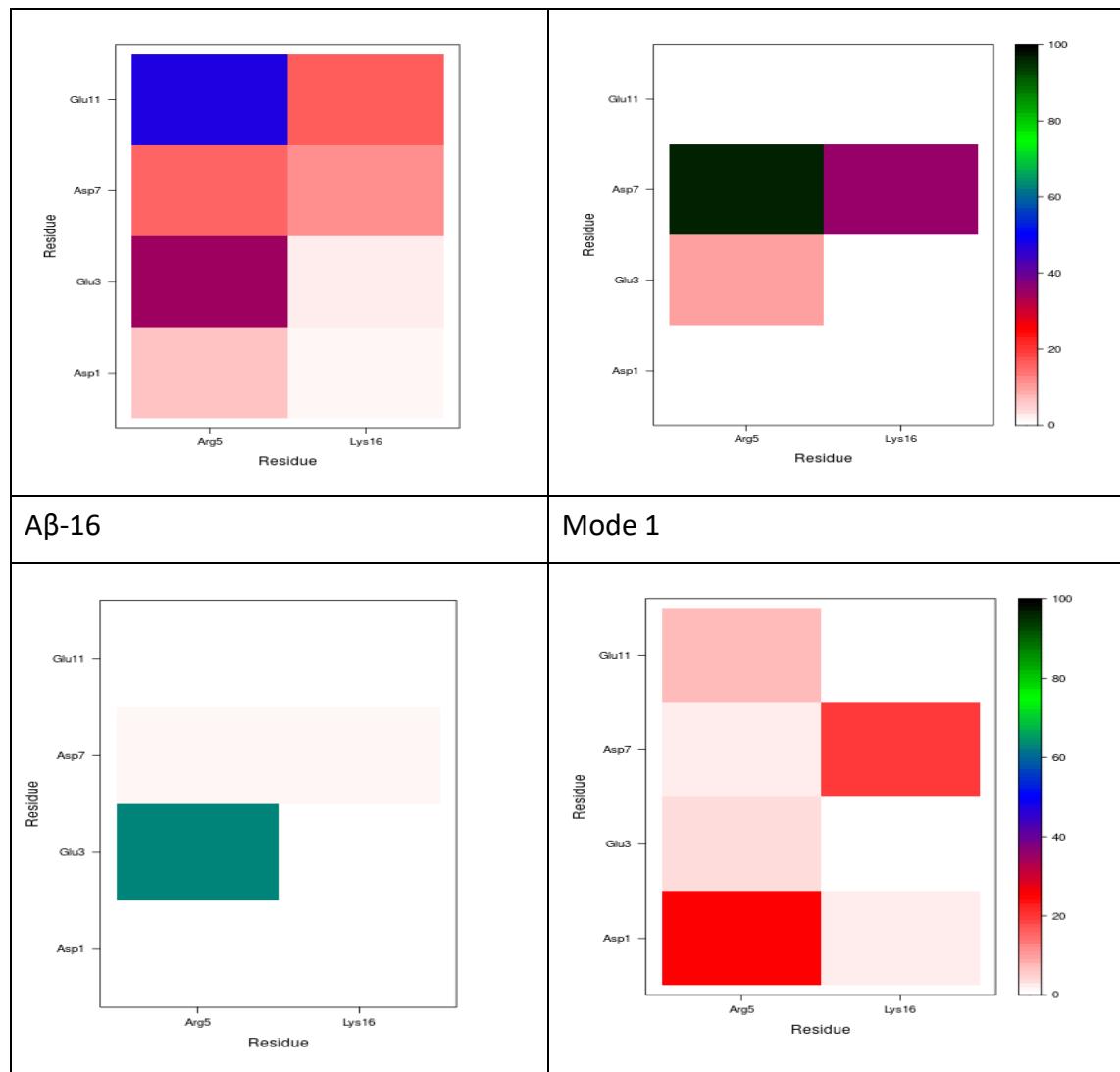


Figure 4.6: Ramachandran maps for Zn(II)-A β and free A β -16. The legends represent the number of dihedrals per bin.

The prevalence of salt-bridge contacts are shown in Figure 4.7:. Most of the binding modes, except mode 2, of Zn(II)-A β show strong contact between Asp7 and Arg5. This is especially apparent within binding mode 1 in which almost 100% of frames display this interaction. This could be the reason why mode 1 undergoes less structural transitions over the simulation, as this interaction restrains the peptide movement. This type of salt bridges between Arg5 positive residues charged and nearby residues of negatively charged Asp7 and Glu3 are presented in A β have been reported in a recent study⁶² done via high-pressure NMR combined with multi quantum chemical exchange saturation transfer (MQ-CEST) NMR as well as MD simulation. The study also suggested theses salt bridges are related with less extended

structures in the N-terminal region of A β . Another study illustrated that the brain-derived A β fibrils comprise Arg5-Glu3 salt bridges between adjacent protofibrils, proposing that the Arg5-based contacts may be a factor in increased A β aggregation.⁶³ Additionally, contacts between Asp7 and Lys16 in modes 1 (which show low RMSD and R_g) exist 40% of the simulation time. No contacts are observed at all between Arg5 and Glu11 in most of the binding modes except for a transient interaction between these two residues seen in Mode 3 at less than 5% of all frames. For mode 2, there are no notable salt-bridge contacts except a strong interaction displayed between Glu3 and Arg5. Mode 7 presents an interaction between Lys16 and Asp7 in more than 40% of trajectory data. Free A β -16 shows many more combinations of salt-bridges but with lower occupancy, with only Arg5-Glu11 is predominant for over 50% of frames showing the flexible dynamic nature of this unbound A β fragment.



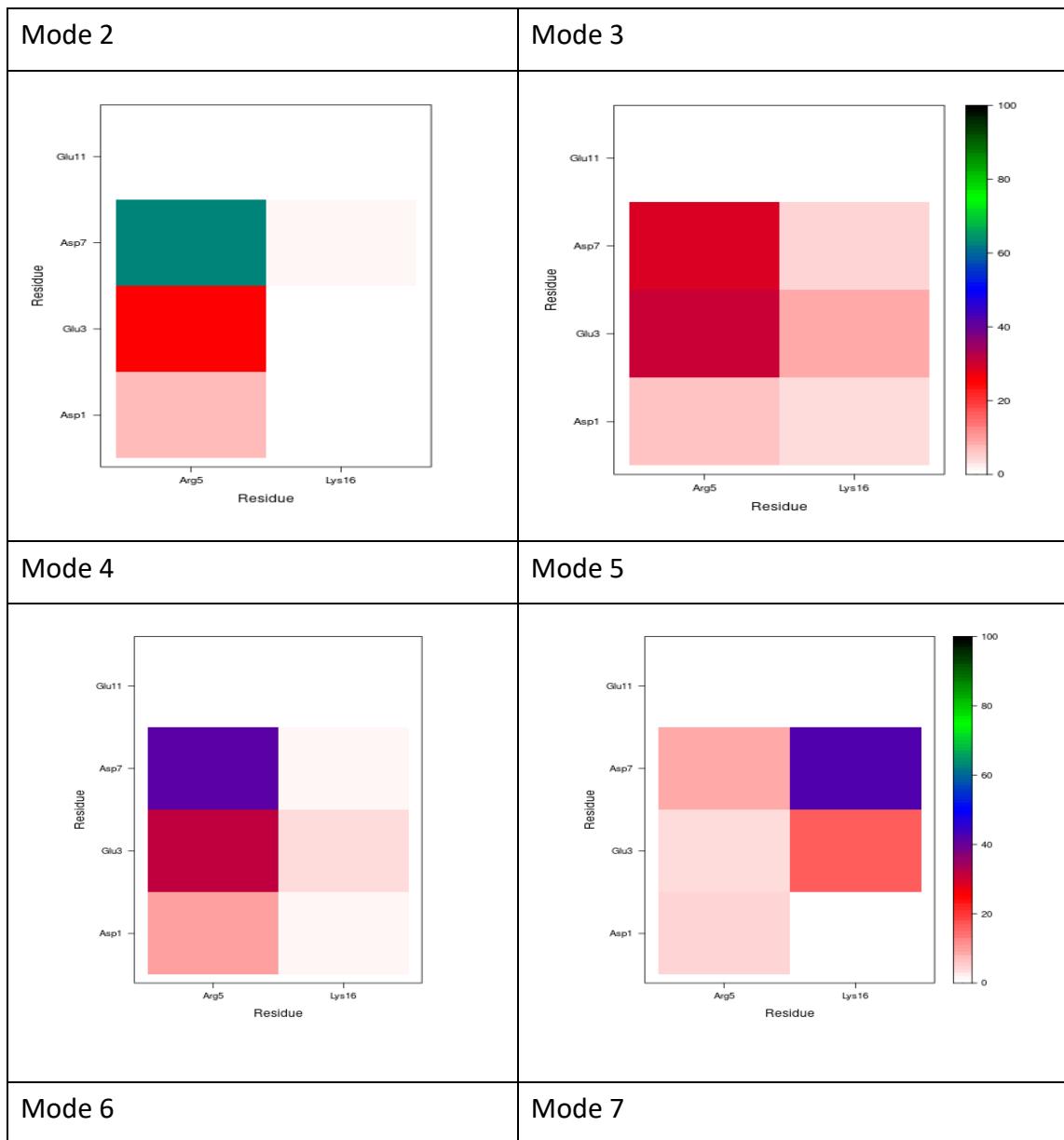
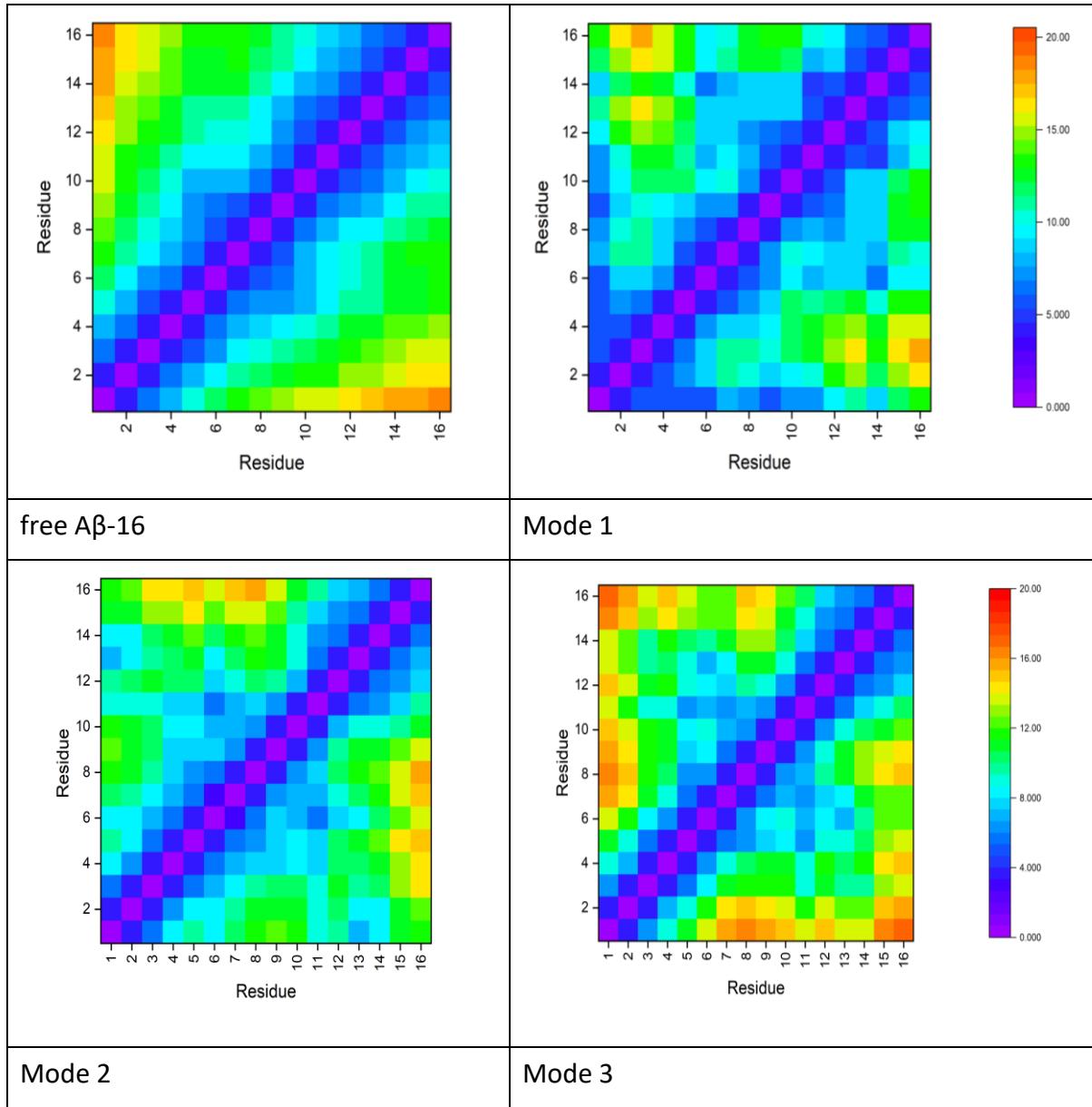


Figure 4.7: Salt bridge maps of Zn(II)-A β all modes and free A β -16.

Contact maps give the average distance between residue pairs as shown in Figure 4.8 for all eight trajectories. Free A β -16 show fewer interactions between polar residues than the Zn complexes, with long distances between N and C-termini clearly demonstrating the flexibility of the peptide in the absence of zinc ion. Zn(II)-A β structures show shorter distances between N- and C-termini, and also short contacts between other specific residues. These are mainly the residues involved in metal binding (His6-His14, His6-Glu11) but other close contacts are also evident (Asp1-Gly9 and Glu11-Gln15), reflecting the profound effect of Zn binding on the structural flexibility compared to the free peptide.



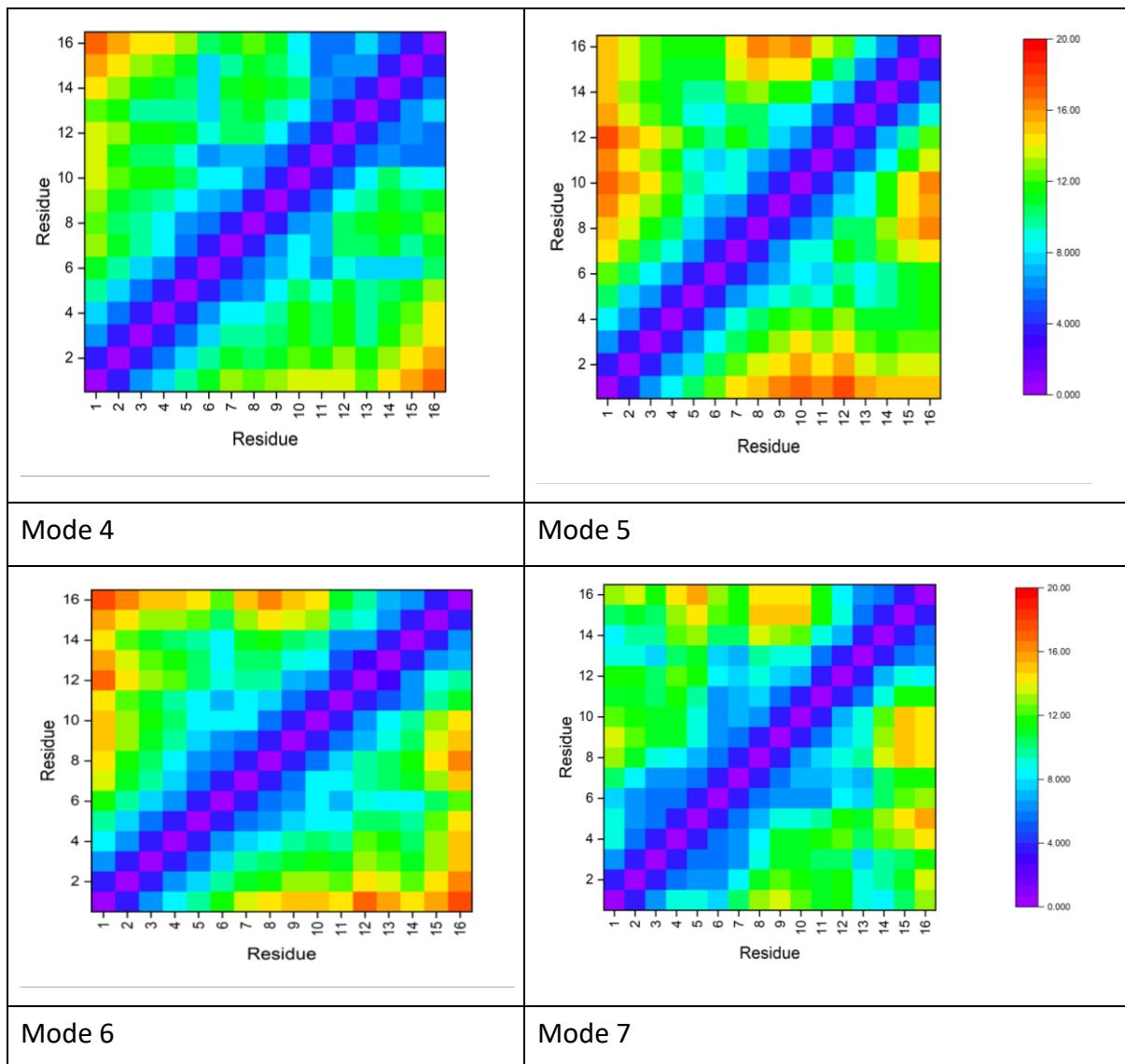
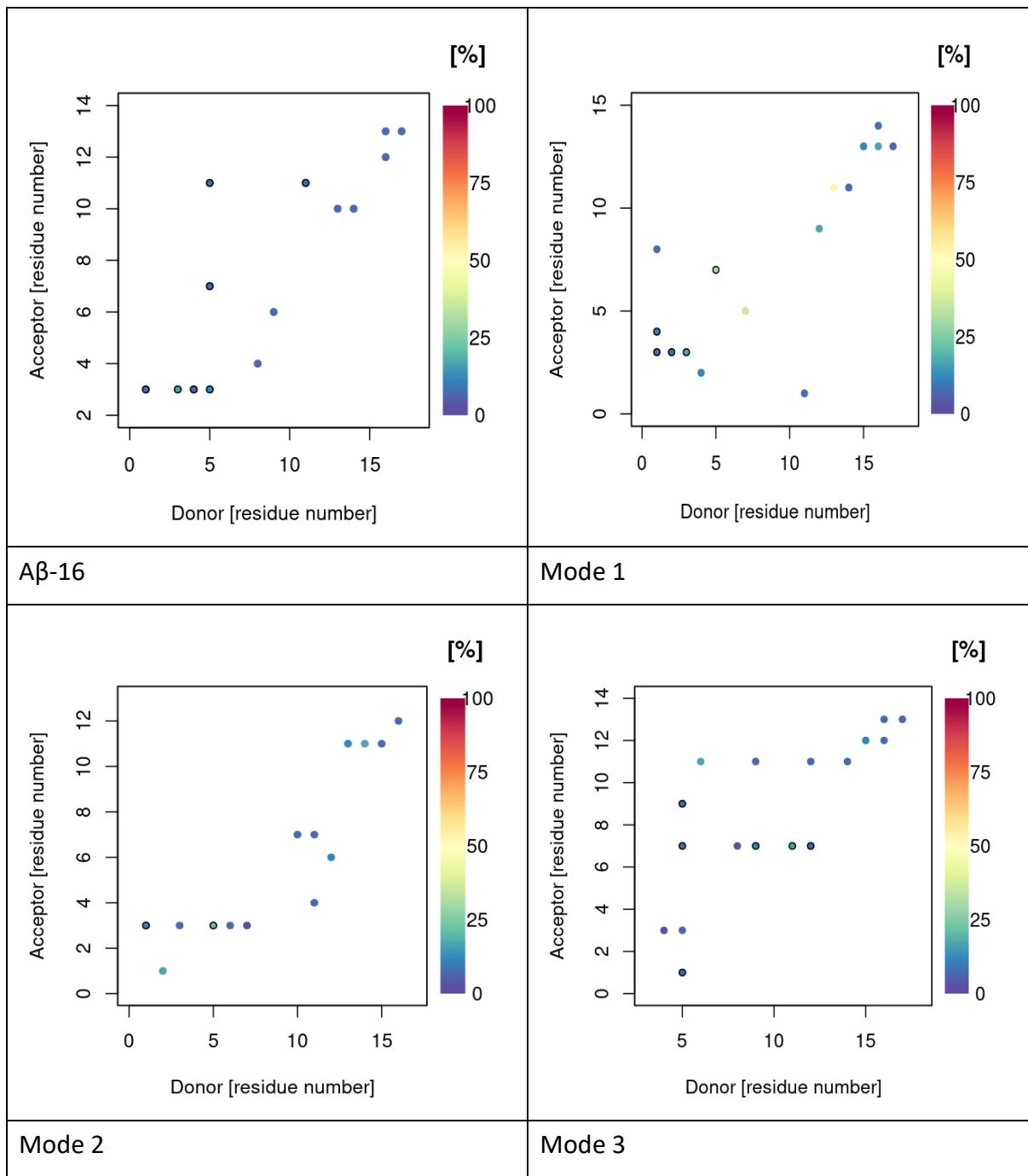


Figure 4.8: Contact map of binding modes of Zn(II)-A β and free A β -16. The legend shows the average inter-residue distance in \AA .

Hydrogen bond formation is an influential factor on the dynamics and stabilization of proteins. The number of H-bonds formed between residues are shown in Figure 4.9 and



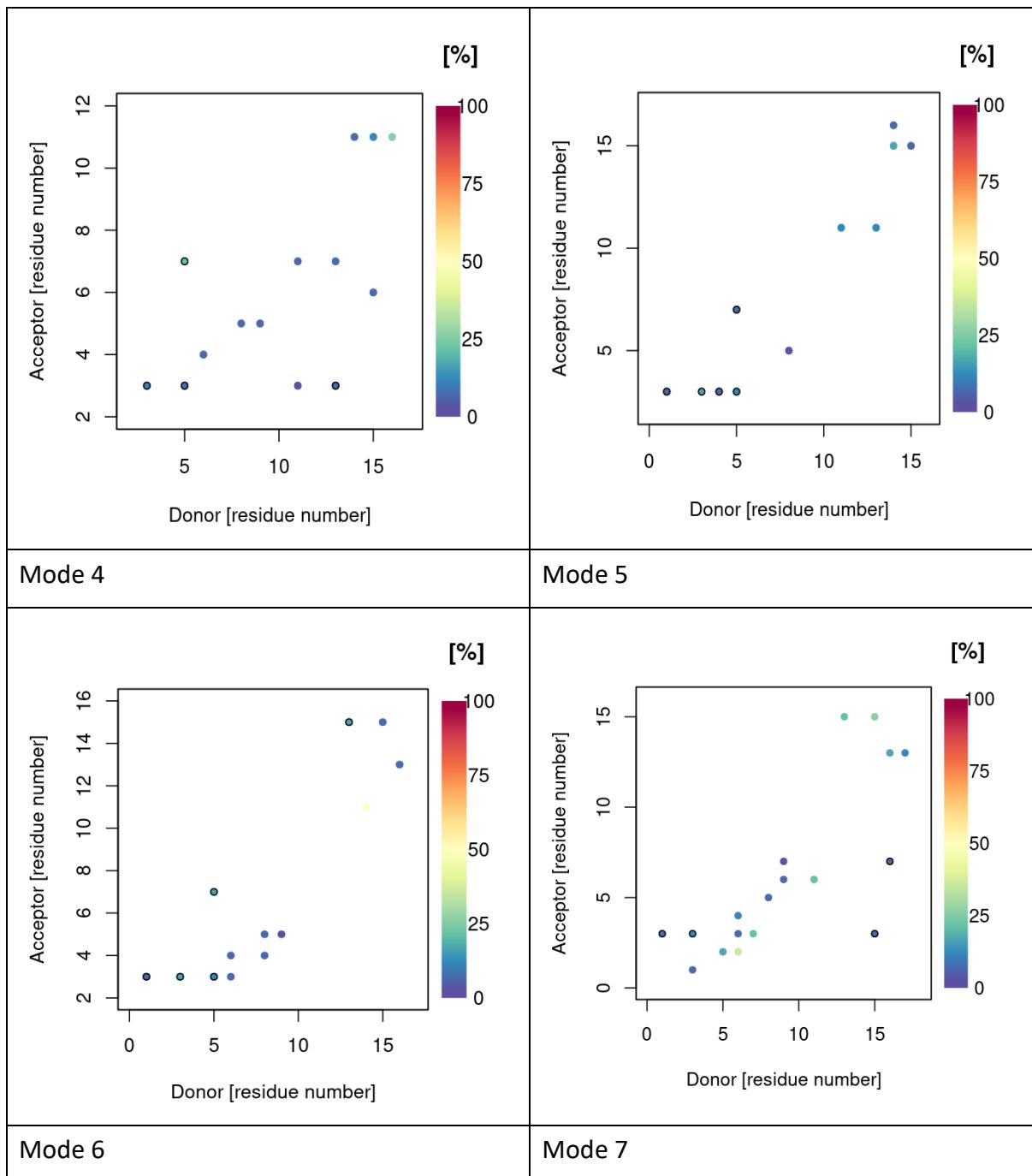
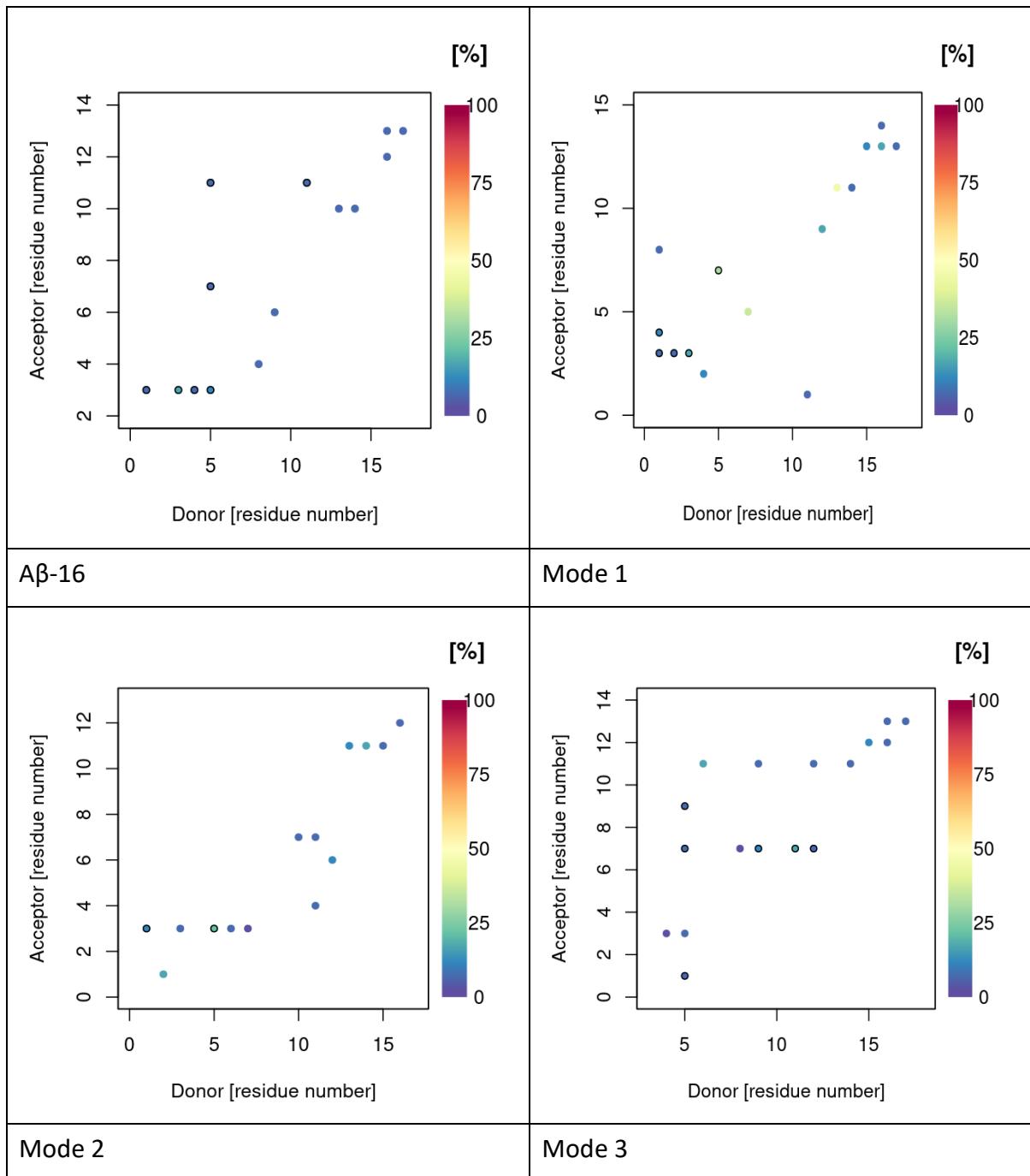


Figure 4.9: Hydrogen bonding occupancy Zn(II)-A β and free A β -16. Plots only display data for hydrogen bonds present for greater than 5% of simulation. Black circles indicate more than one type of hydrogen bond between the relevant residues.

. The data obtained from aMD simulations show that all binding modes of Zn(II)-A β studied, as well as free A β -16 exhibit very similar patterns of hydrogen bonding, each with a mean of *ca.* 5 and a maximum of 15 in some frames. The most consistent H-bonds existed for 45 and 49% of frames for binding mode 1 and 6 respectively, forming between His14 (donor) and Glu11 (acceptor), but just 16% for A β -16 intra-residue H-bond within Glu3. Mode 3 and 5

show the maximum number of hydrogen bonds within these simulations was 11, which is less than all other modes.



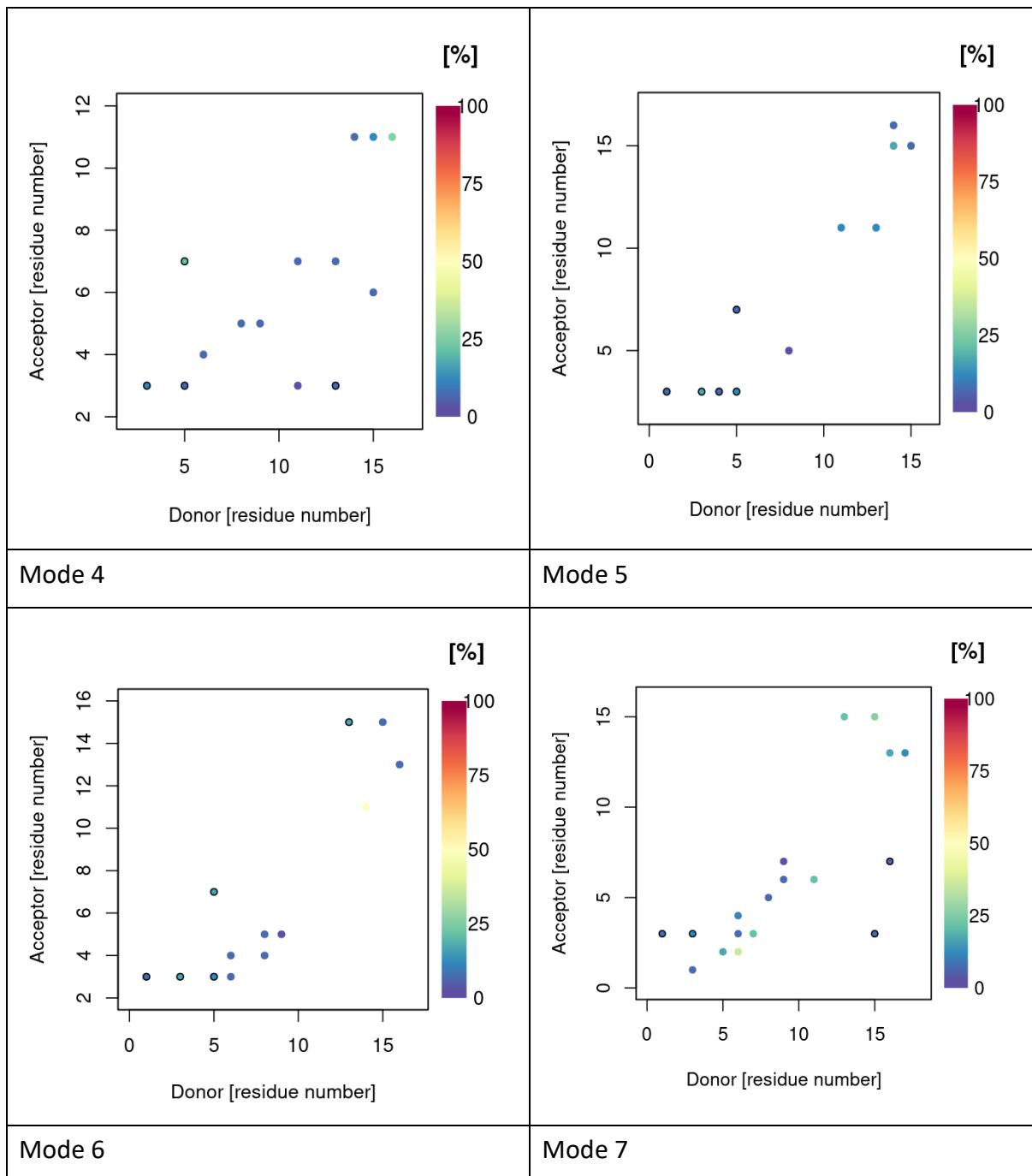
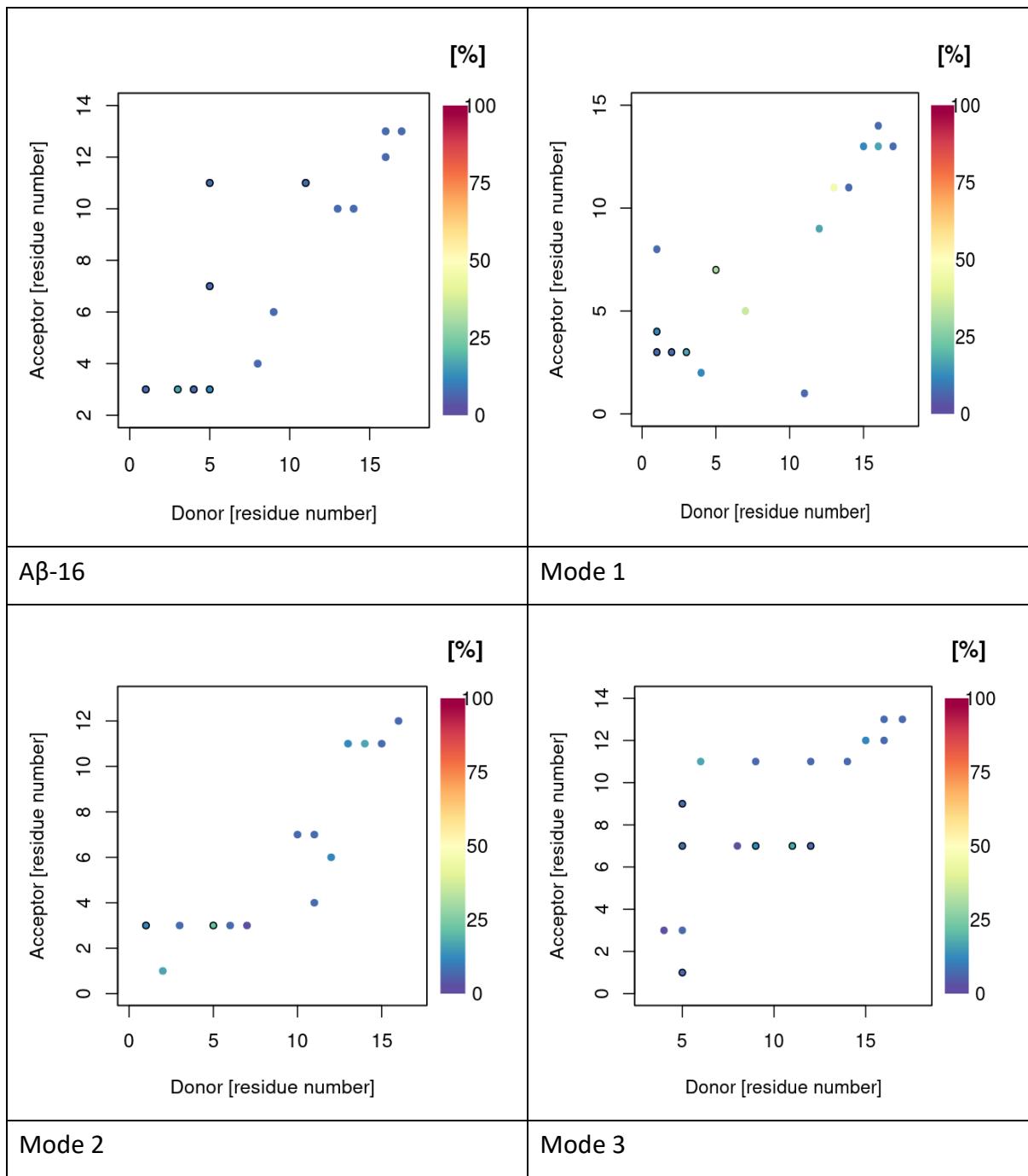


Figure 4.9: Hydrogen bonding occupancy Zn(II)-A β and free A β -16. Plots only display data for hydrogen bonds present for greater than 5% of simulation. Black circles indicate more than one type of hydrogen bond between the relevant residues.

also shows that every aMD simulation visits at least some frames in which no hydrogen bonds are present, reflecting the flexibility of the peptide.



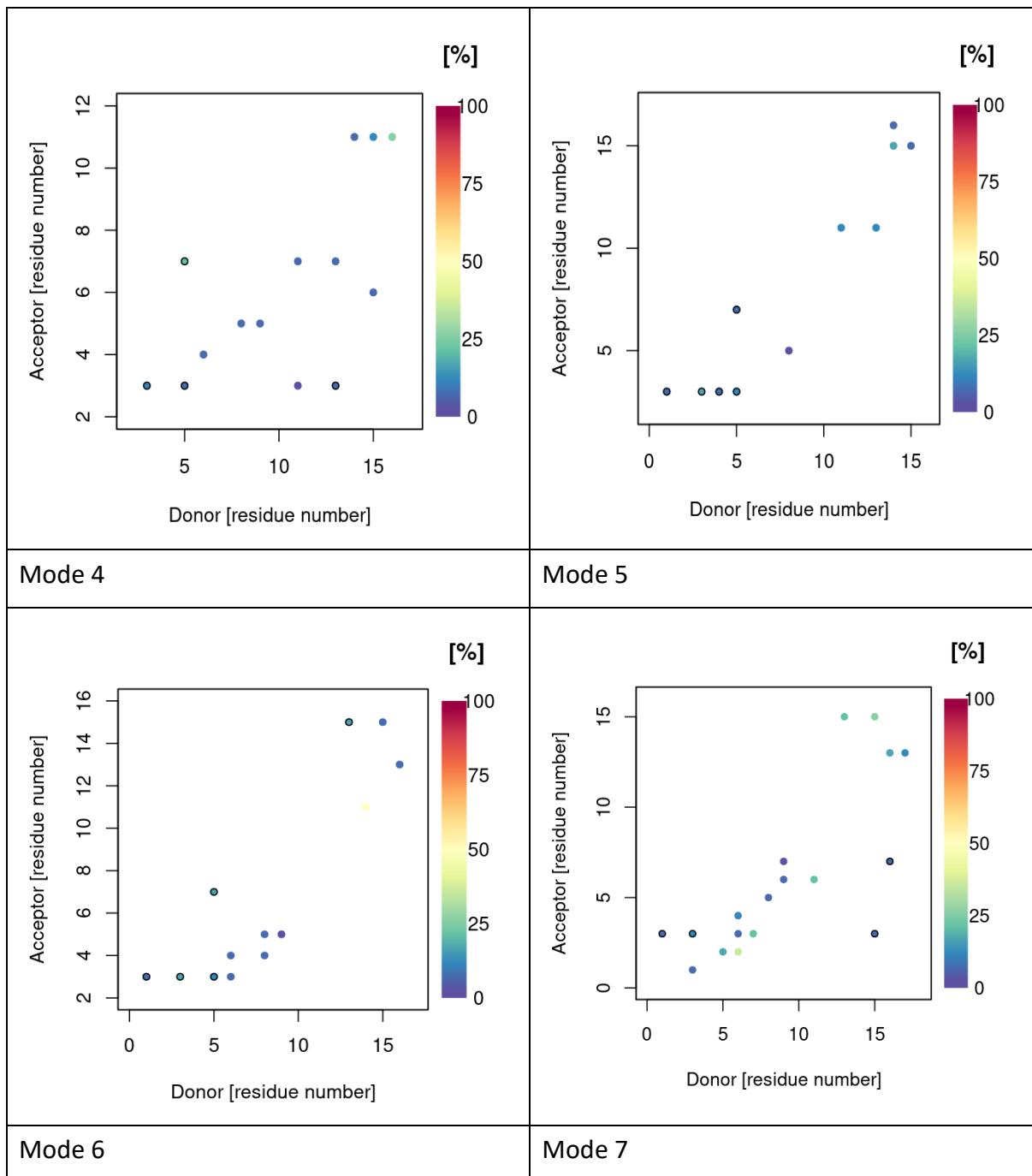


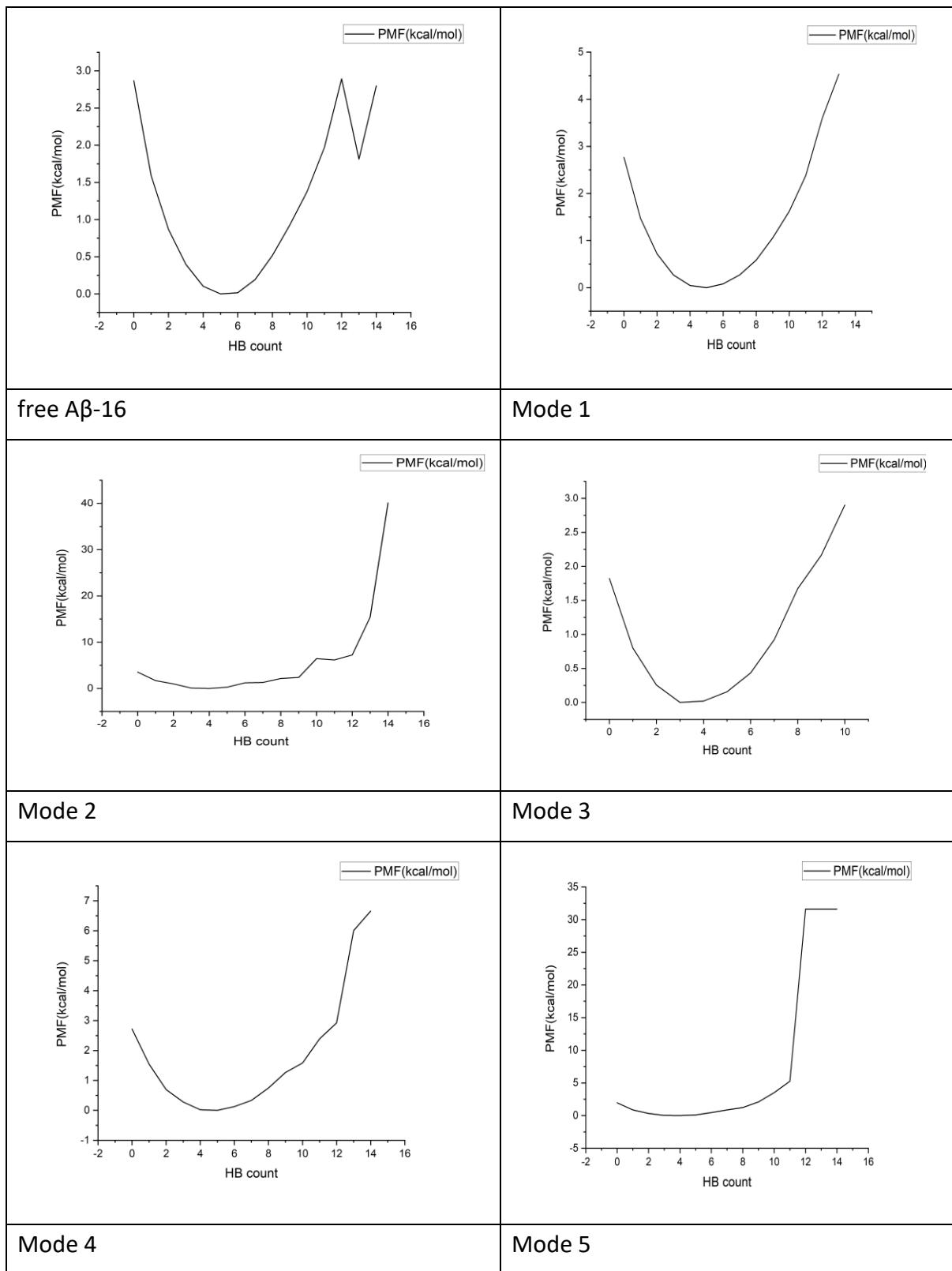
Figure 4.9: Hydrogen bonding occupancy Zn(II)-A β and free A β -16. Plots only display data for hydrogen bonds present for greater than 5% of simulation. Black circles indicate more than one type of hydrogen bond between the relevant residues.

Table 4.5: HB count for Zn(II)-A β and free A β -16.

Number of HB	Mean	sd	Max	Min
Mode 1	5.08	1.94	14	0
Mode 2	3.71	1.70	13	0
Mode 3	3.65	1.67	11	0
Mode 4	4.74	1.85	14	0
Mode 5	3.77	1.68	11	0
Mode 6	4.50	1.77	14	0
Mode 7	3.84	1.73	14	0
Aβ-16	5.14	1.98	15	0

Reweighting of the accelerated MD boost potential allows reconstruction of the free energy surface. Figures 4.10, 4.11, and 4.12 show 1D free energy as a function of hydrogen bond count, R_g , and end-to-end distance (R_{E-E} , the distance between the N-terminal and C-terminal $\text{C}\alpha$ atoms), respectively for binding modes and free peptide. These show broad minima centred on 4 or 5 H-bonds for both the Zn complexes and the free peptide. A second low free energy state is found at 13 H-bonds for free A β -16 peptide, reflecting the greater number of these interactions in this more flexible system. The R_g values of the free peptide have a higher range (reaching values up to 12 Å), compared to Zinc bound peptides. A large range of R_{E-E} is accessible at low energy: for Zn, the lowest energy lies at around 13 Å, but values of between 6 and 16 Å are within energy minima. In comparison, the metal-free peptide shows the minimum ranged within a wider distance from 5 to around 40 Å and the lowest located at *ca.*14 Å in which lies at a similar place as the minima of Zn.

Binding modes 1, 2, and 3 of Zn(II) bound to Asp1, or Glu3, reduces the R_{E-E} range dramatically due to the constraint on the N-terminus. This result has also been seen in RMSF. Whereas binding modes 4, 5, and 6, in which Zinc is bound to three His rings 6, 13 and 14 in the centre of the peptide chain, can access a wider range of R_{E-E} , nearer to those seen in free peptide values. Interestingly, this value is reduced in mode 7 due to the fifth bond provided from Asp amine to the Zinc ion.



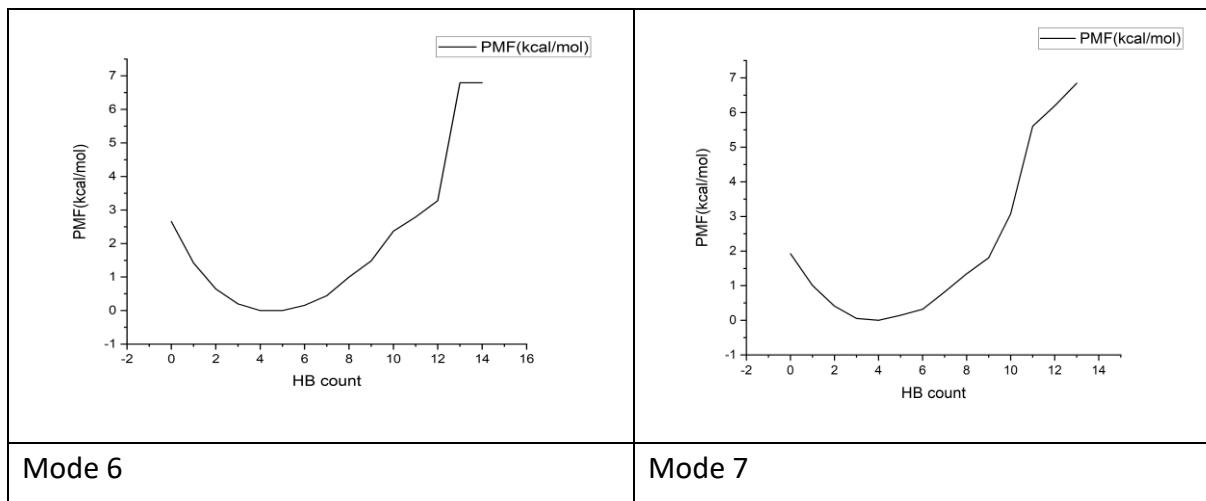
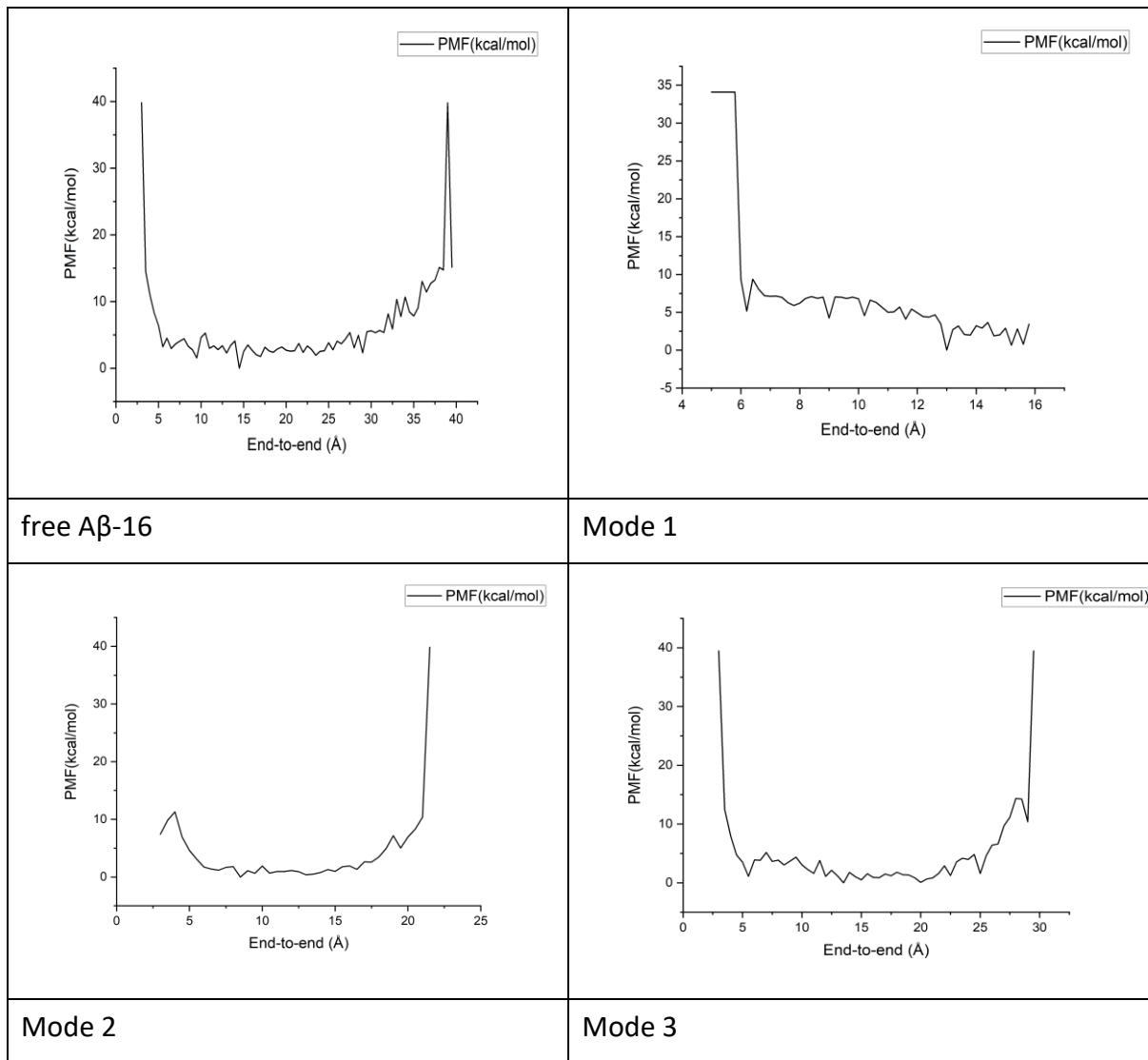


Figure 4.10: 1D-Free energy surface of binding modes of Zn(II)-A β and free A β -16 as a function of HB count.



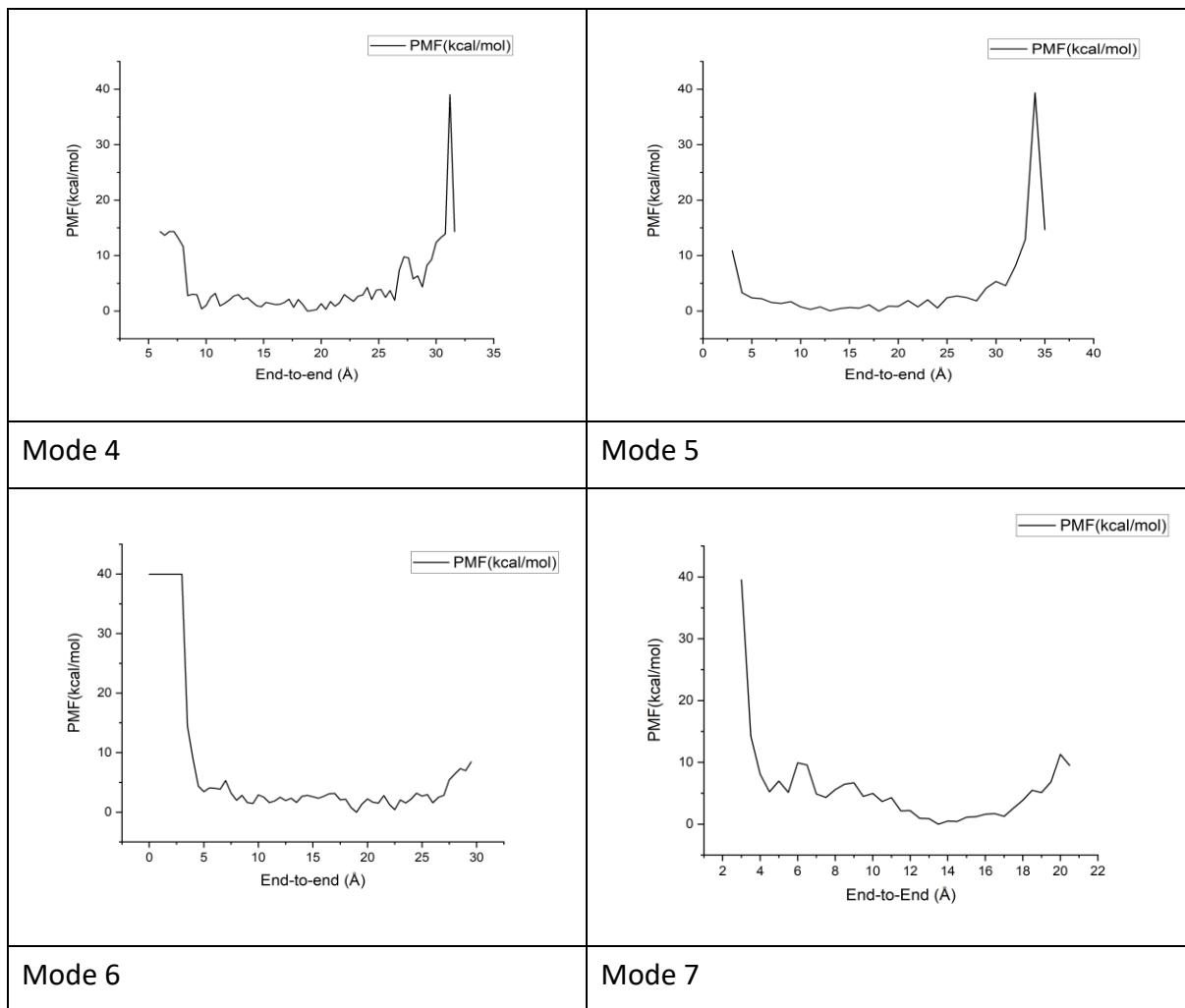
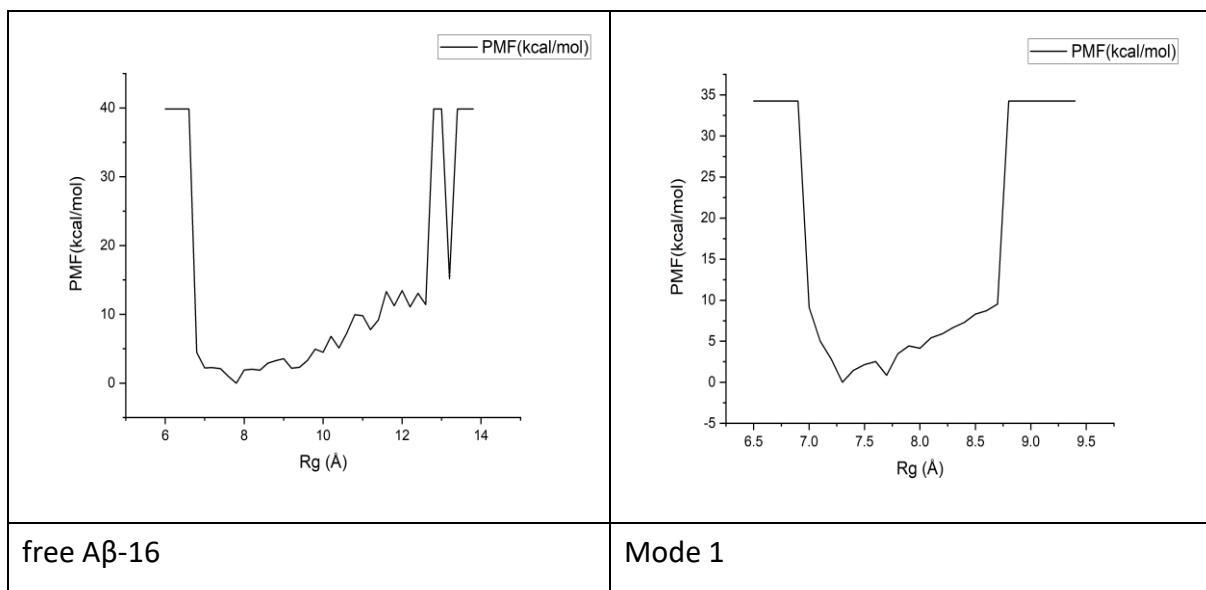


Figure 4.11: 1D-Free energy surface of binding modes of Zn(II)-A β and free A β -16 as a function of end-to-end distance (\AA).



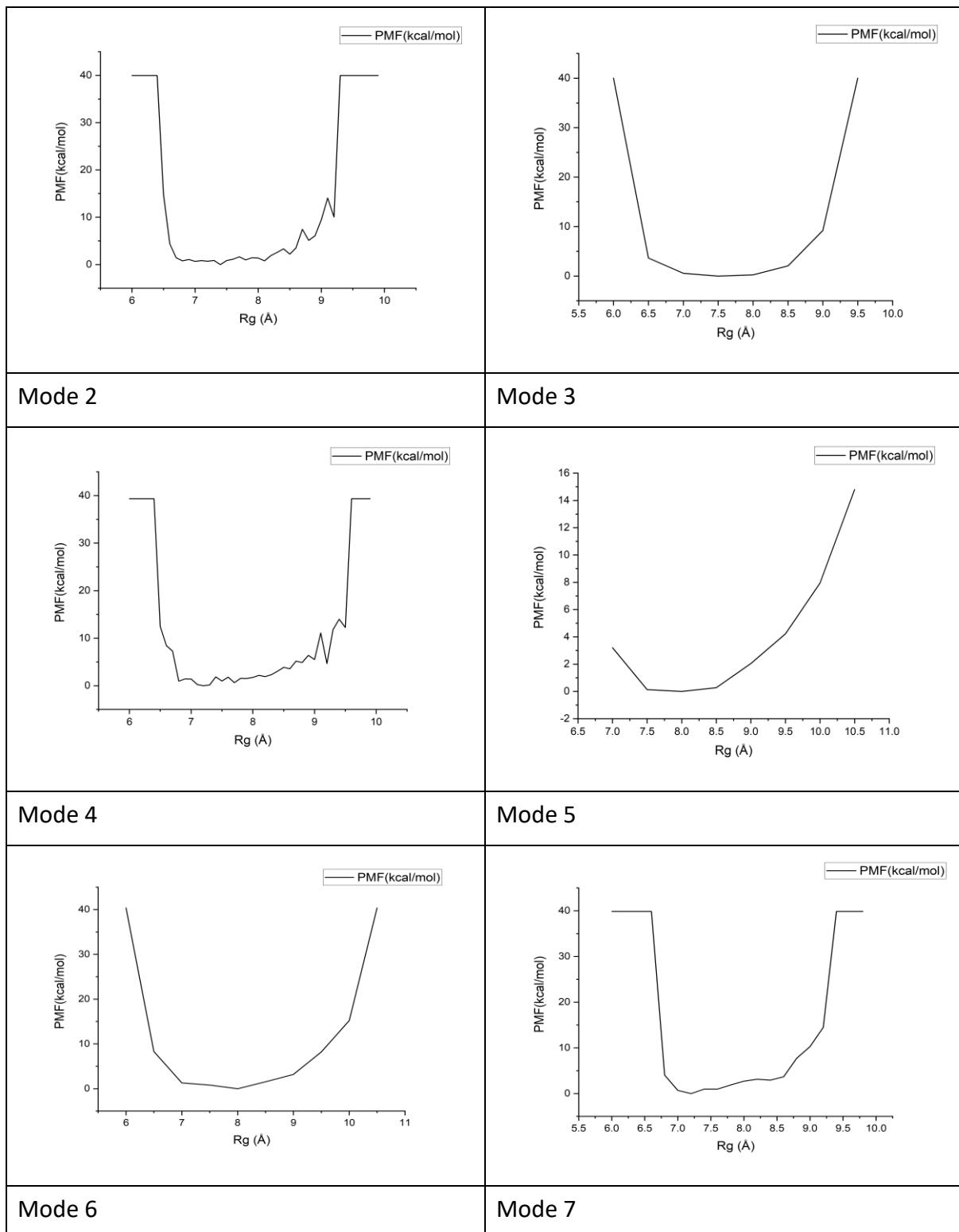


Figure 4.12: 1D-Free energy surface of binding modes of Zn(II)-A β and free A β -16 as a function of R_g (\AA).

4.3.3 Conclusion

Accelerated MD simulations were performed to explore the structures of Zn(II) bound to different binding sites of A β 1-16 as well as the free peptide with no metal centre. aMD was used due to the feature of this method, which enhances conformational sampling by overcoming insurmountable energy barriers within the potential energy surfaces which can prevent conventional MD from accessing certain states. The results show that the coordination of Zn ion within the N-terminus reduces the flexibility of the peptide, and significantly alters the pattern of salt bridges, but affects hydrogen bond interactions to a lesser extent. In addition, the exploring of different binding modes of Zn(II)-A β indicated that variation in specific atoms/residues involved in binding, affects the secondary structure, clusters, salt bridge, RMSF, and peptide size. Here, the modes where the Asp1 and Glu3 residues of N-terminal are coordinated to Zn(II) show more compact and restricted peptides due to anchorage of the metal centre. However, in the cases where the Zn(II) coordinated to two His rings and two carboxylate groups, those residues located on the middle of A β -16 chain show structures that are less constrained and present more flexible ensembles. The strong interaction between Arg5 and Asp7, that lasts almost for the entirety of the simulations, more likely restrains the peptide, affecting its flexibility. The Arg5-Asp7 salt bridge has been demonstrated in NMR analysis of free A β . The Ramachandran maps show mode 5, where the Zn(II) coordinated to N ε of two histidine rings, has population at the anti-parallel β -sheet region. It has reported in the literature, that the peptide accumulates through intermolecular His(N ε)-Zn(II)-His(N ε) bridges.⁴⁹

4.4 Zn (II)with capped and uncapped Lys16 comparison

4.4.1 Introduction

In this section, an aMD setup similar to that seen in the previous section was used to study the effect of Zn(II) on conformational ensembles of two A β 16 structures. One with amidated C-terminus peptide and one structure un-amidated (or uncapped). Lysine, which forms the 16th residue in the sequence of the N-terminus of A β , was capped with (NMe) functional group instead of a -COO⁻ ion. The simulations were performed on a different binding mode not included within the seven modes reported in previous section, avoiding repetition of the same

information, they are still within the proposed coordination residues to the Zn(II), specified in the introduction. Comparisons are drawn where appropriate between these two structures, analysed, and the free peptide shown within the previous section. This binding mode is similar to Mode 2 but Glu3 was used instead of Glu11 to coordinate to the Zn(II) in addition to Asp1, His6 and 14, as shown in Figure 4.13.

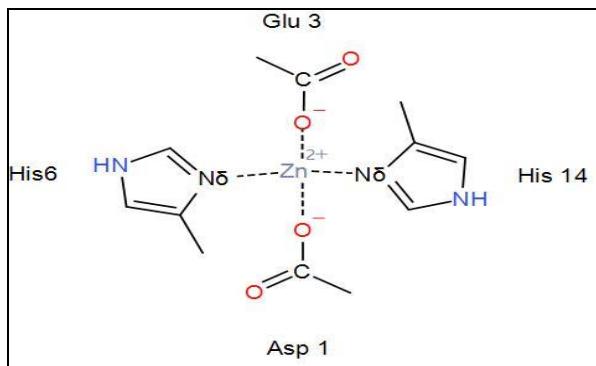


Figure 4.13: Binding mode used for comparison between capped and uncapped Lys (C-terminus amidated with -NCH₃ group).

4.4.2 Result and discussion

RMSD values of the peptides comparing backbone atoms with respect to initial structure as well as R_g data indicate that the un-capped peptide explored more diverse and more transition conformations than capped peptide, shown by larger mean and sd. The capped form remained more constant and more compact where the value stayed around 5 Å with the sd value at 0.38 Å, which reflects the small change in structure during simulation time, as shown in Figure 4.14 and Table 4.6. Unlike RMSD, the R_g values between both systems show little changes in the peptide compactness, with average values around 7.5 Å as indicated in Table 4.6 and Figure 4.14. These values are similar to those seen in Mode 2, in the previous section, and this is expected as these binding modes are similar.

Table 4.6: RMSD and R_g (Å) of capped and uncapped Lys.

Zn(II)-Aβ	RMSD				R _g			
	Mean	sd	Min	Max	Mean	sd	Min	Max
Capped Lys	4.59	0.38	2.53	6.00	7.42	0.22	6.63	8.43
Uncapped Lys	4.73	0.66	2.67	6.79	7.48	0.23	6.77	8.79

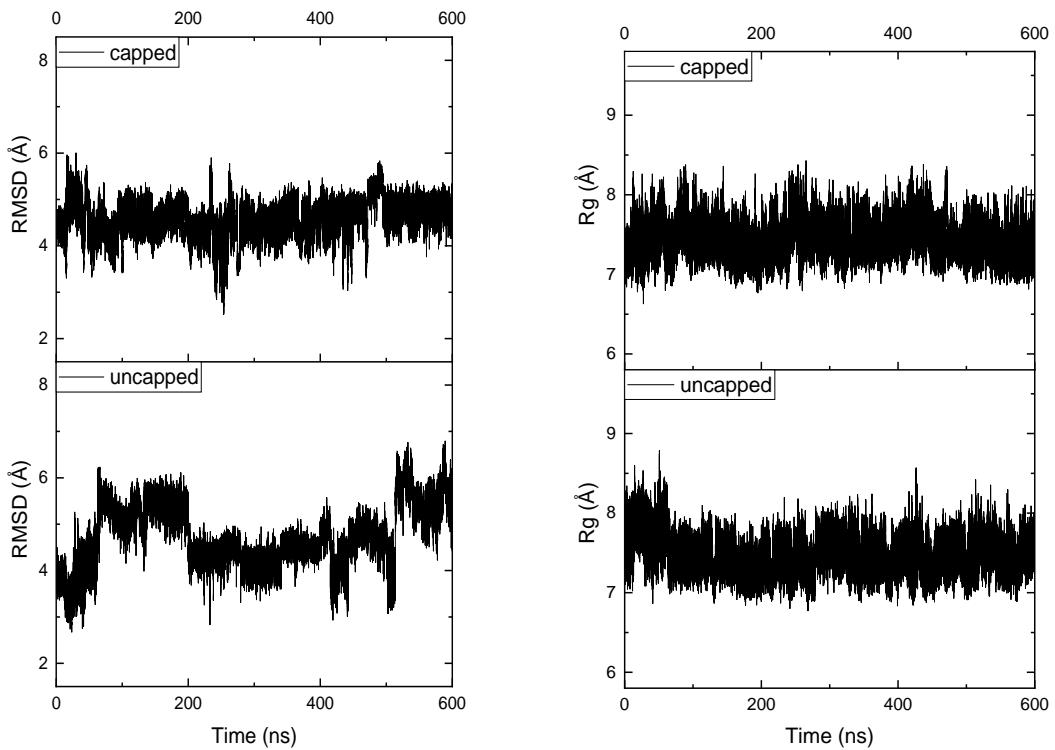


Figure 4.14: RMSD and R_g of Zn-A β 16- capped and uncapped Lys (\AA).

RMSF figures for each amino acid show similar trends generally, but the uncapped peptide shows a higher level of mobility compared to their original position. The main interest here is focusing on the C-terminal section of those two peptides (residues number 16 and 17 of uncapped and capped, respectively). Both act similarly and show comparable degrees of movement. The largest values observed are in Arg5 in both peptides, but its RMSF value is notably larger in the uncapped peptide. Asp7 and Tyr10 of the uncapped peptide show more movement which reflects their higher level of mobility compared to the amidated one as shown in Figure 4.15.

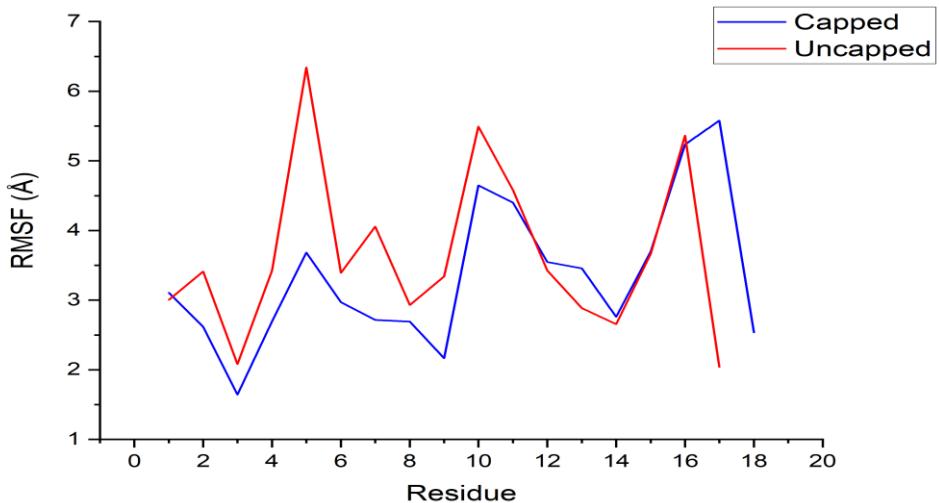


Figure 4.15: RMSF of each residue of free and Zn-bound Aβ-16. The metal ion itself is denoted by residue 18 in case of capped and by residue 17 in uncapped.

Clustering analysis of equilibrated trajectories based on backbone atom dihedrals of the two peptides show that both have high numbers of clusters with 23 and 28 groups of clusters being formed for the capped and uncapped peptides respectively. With the most populated percentages of 15% for capped and 20% for the uncapped, as shown in Table 4.7. Ribbon diagrams of the most populated clusters for both systems are provided in Figure 4.16.

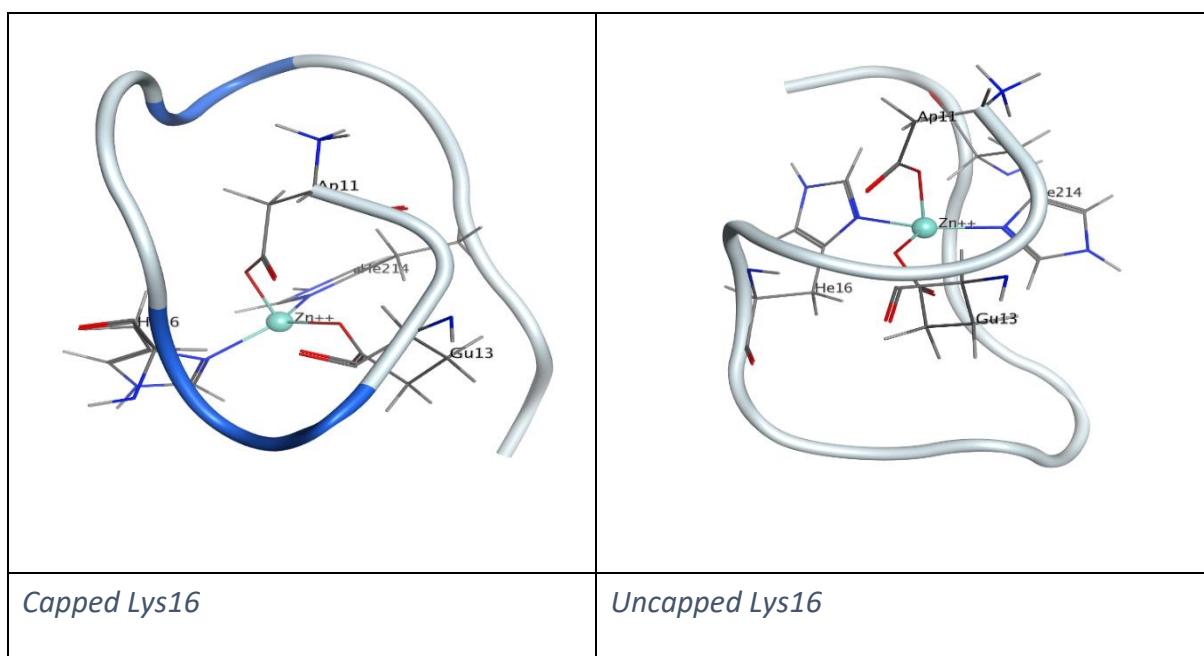


Figure 4.16: Ribbon diagram of the most populated clusters for Zn(II)-Aβ capped and uncapped Lys16.

Table 4.7: Cluster analysis data for equilibrated trajectories for amidated C-terminus and un-amidated Zn-A β 16.

Clusters	# clusters	Most populated	2nd populated
Zn-A β capped	23	15%	11%
Zn-A β uncapped	28	20%	11%

Secondary structure by residue plots provide valuable information of the metal binding effect on the structure as shown in Figure 4.17 (with numerical data presented in Table 4.8). The plots show small difference between Zn(II) with and without amidated Lys16 for the binding mode simulated. Both peptides tend to occupy bend and turn formations for the majority of their structures at levels greater than 96%. There is small amount of anti-parallel β -sheet formed in Zn(II)-A β 16 with capped Lys observed in 6, 9 and 10 residues at only 2% of simulation time. The non-amidated A β sequence induces a significantly lower propensity to form β -sheet at a negligible 0.03% of frames present. α -helix forms around 1% of both peptide structures and this percentage is seen to be lower than the metal-free A β simulated previously in which helical structure form 27% of the peptide chain.

Table 4.8: Total secondary structure percentages for each system (%).

	Helix	β-Sheet	Other
Zn-Aβ capped	1.42	2.09	96.49
Zn-Aβ uncapped	1.57	0.03	98.40
Aβ	27.2	0.4	72.4

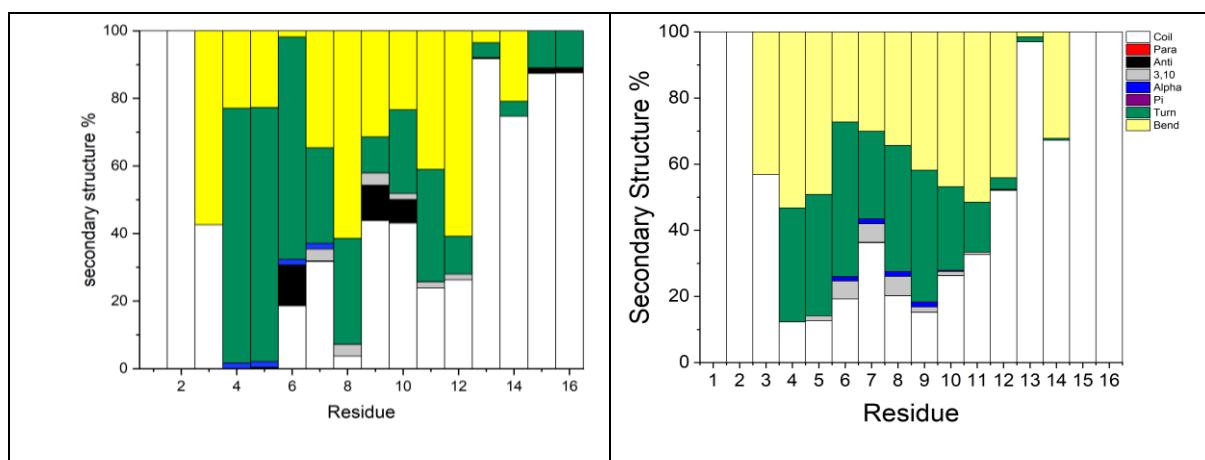


Figure 4.17: Secondary structure plots for Zn-A β 16 capped and uncapped Lys, respectively.

In order to explore the differences between the conformational ensembles of the amidated and non-amidated forms of A β coordinated with Zn(II), the average distances between the C-terminal Lys16 and all other residues were calculated over the aMD trajectories as shown in Figure 4.18. The results for our region of interest (the C-terminus) show strong interactions between the Lys16 and Gly9 residues of the uncapped terminus. This interaction still exists within the capped peptide but at a lower amount. This interaction refers to the exposed COO- atoms which make it free to interact with positively charged atoms.

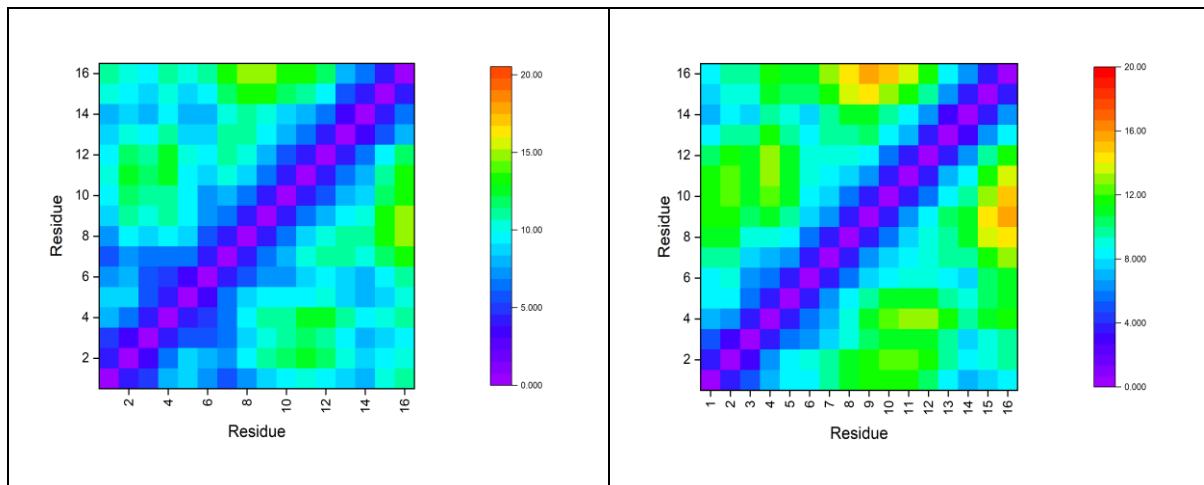


Figure 4.18: Contact map of Zn(II)-A β capped and uncapped C-terminus, respectively. The legend shows the average inter-residue distance in Å.

Additional identification of peptide structure and its stability can be obtained by using salt bridge maps which show the possibility of the contact between eight polar residues within the peptide by measuring the contact percentage between them. These results are displayed graphically in Figure 4.19. These results highlight some notable differences between capped and uncapped Lys systems of the interaction between Asp7 and Arg5. Zn capped Lys shows 0% contact between Arg5 and Asp7, in contrast uncapped shows contact by about 25% of frames. Both of two systems present no interaction between Asp1 and Arg5 nor Lys16. The later also did not form contact with Glu3 in both systems over all simulation time. Here, the

capped Lys16 residue did not show much different in term of interaction with uncapped one where both of them (Lys16) did not contact to any other residues.

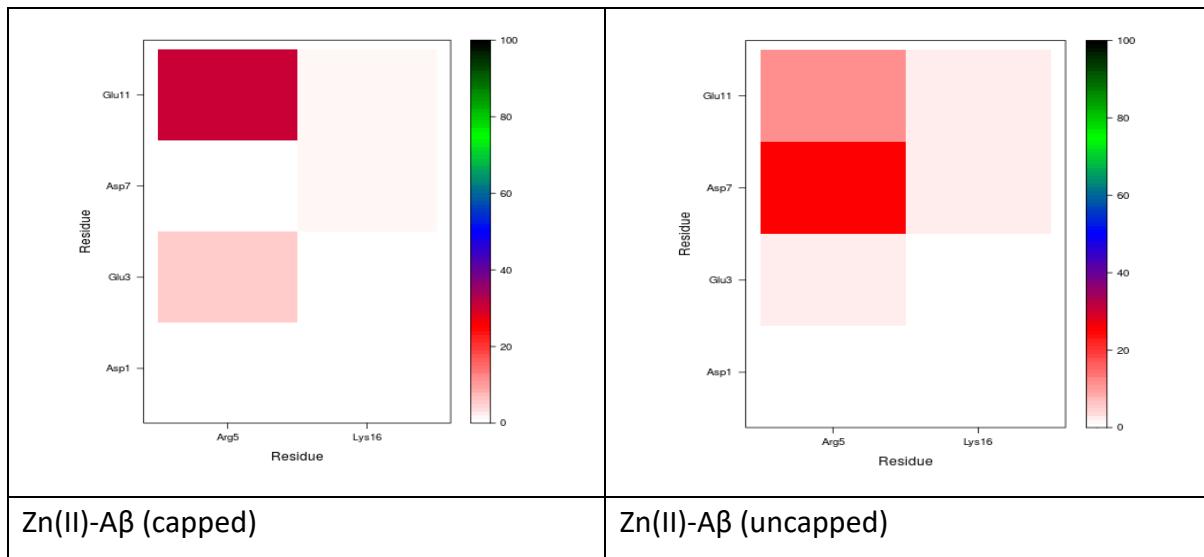
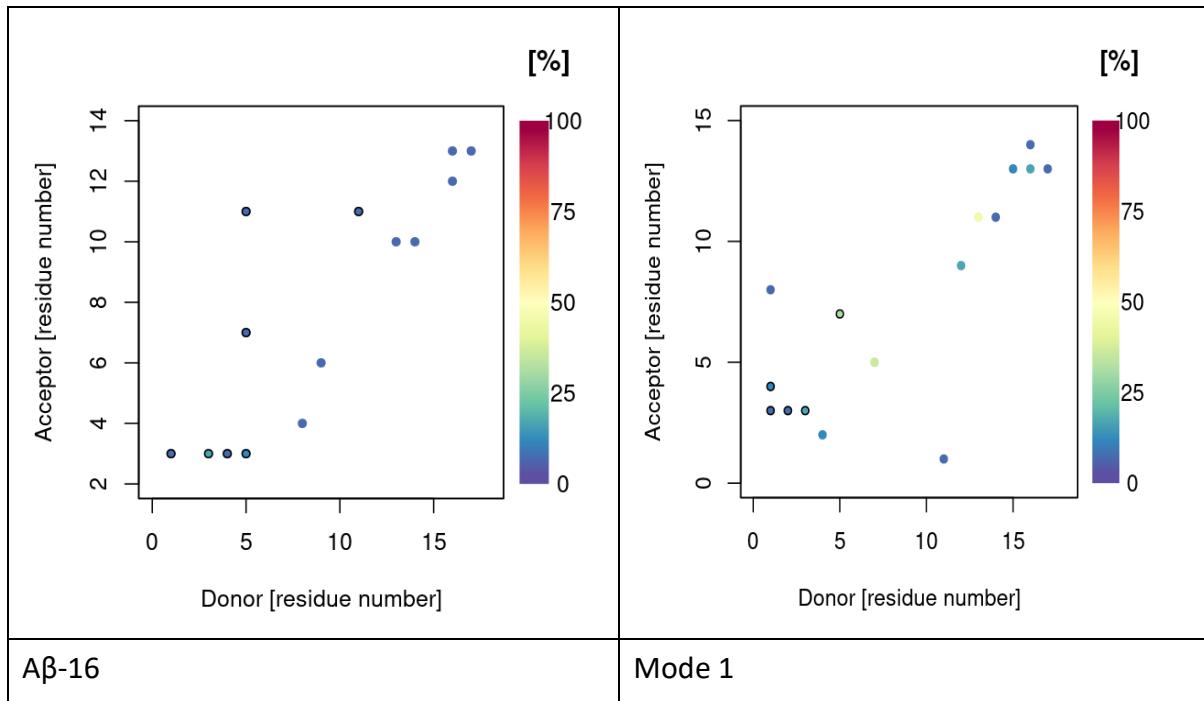
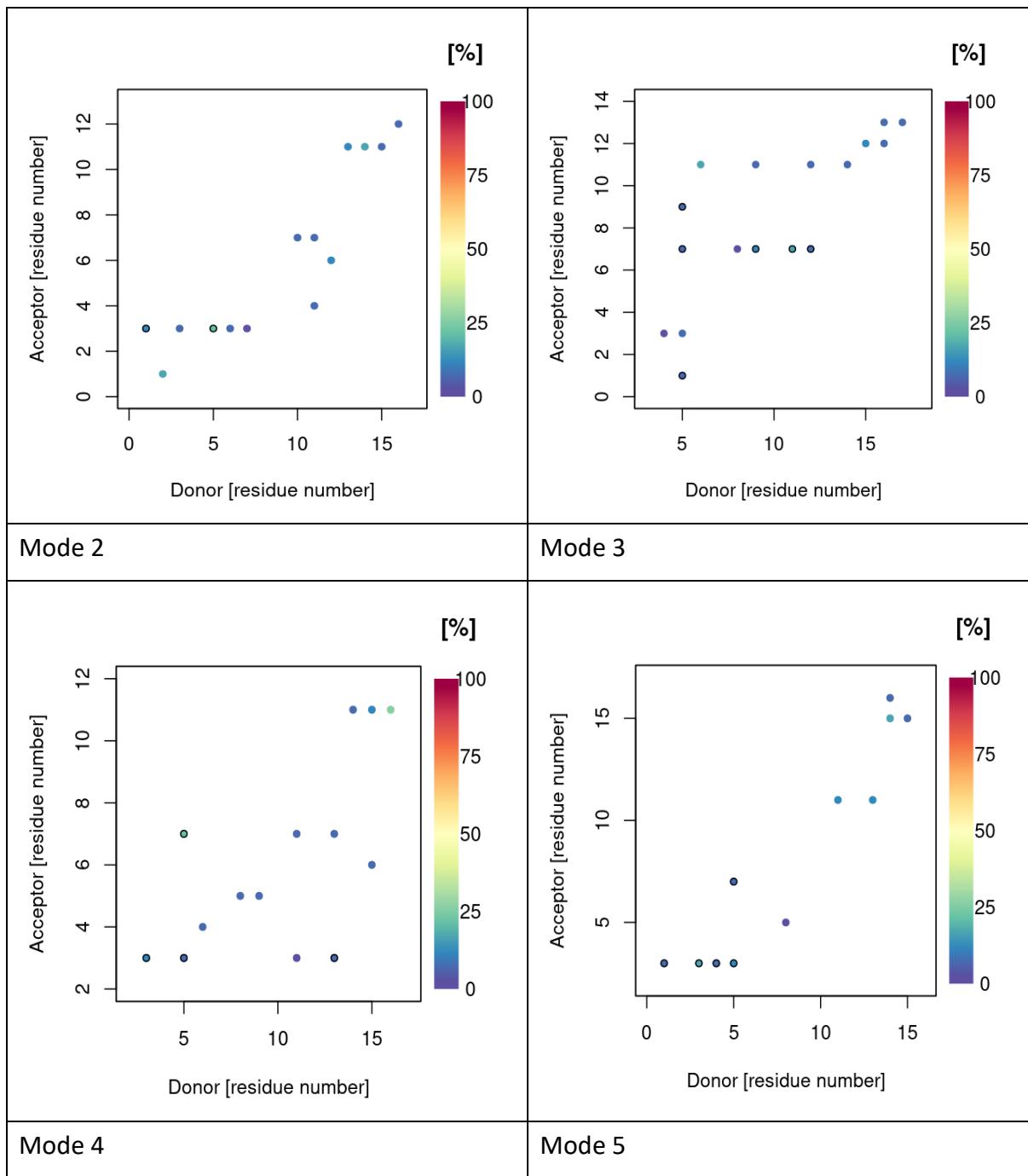


Figure 4.19: Salt bridge maps of Zn(II) capped and uncapped A β -16(%).

Hydrogen bond formation is an important factor for the dynamics and stabilization of proteins. The number of H-bonds formed between residues is shown in





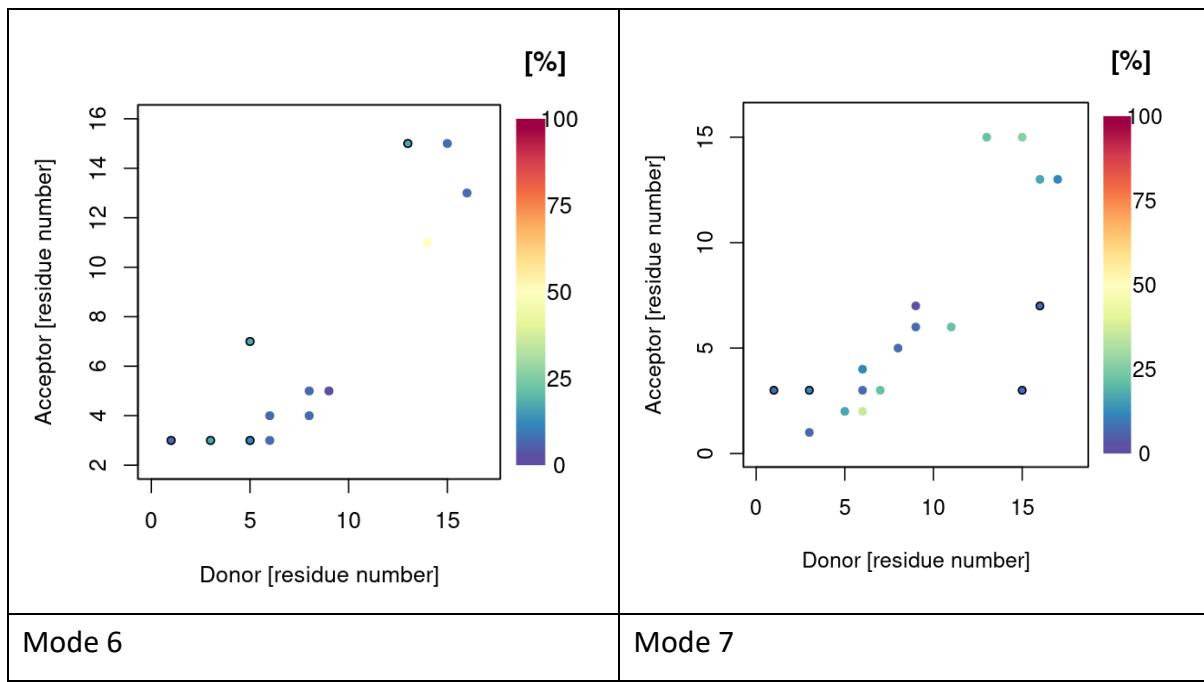


Figure 4.9: Hydrogen bonding occupancy Zn(II)-A β and free A β -16. Plots only display data for hydrogen bonds present for greater than 5% of simulation. Black circles indicate more than one type of hydrogen bond between the relevant residues.

4.9 and Figure 4.20. Both zinc-bound systems have lower Ave, Max and sd of HB number compared to free A β -16. The peptide that contains amidate group protecting the C-terminus has a negligibly lower mean number of hydrogen bonds compared to the uncapped A β fragment analysed.

Table 4.9: HB count for Zn(II)-A β for capped and uncapped C-terminus of A β -16.

Number of HB	Mean	sd	Max	Min
Zn-A β capped	3.14	1.62	13	0
Zn-A β uncapped	3.66	1.71	12	0
A β -16	5.14	1.98	15	0

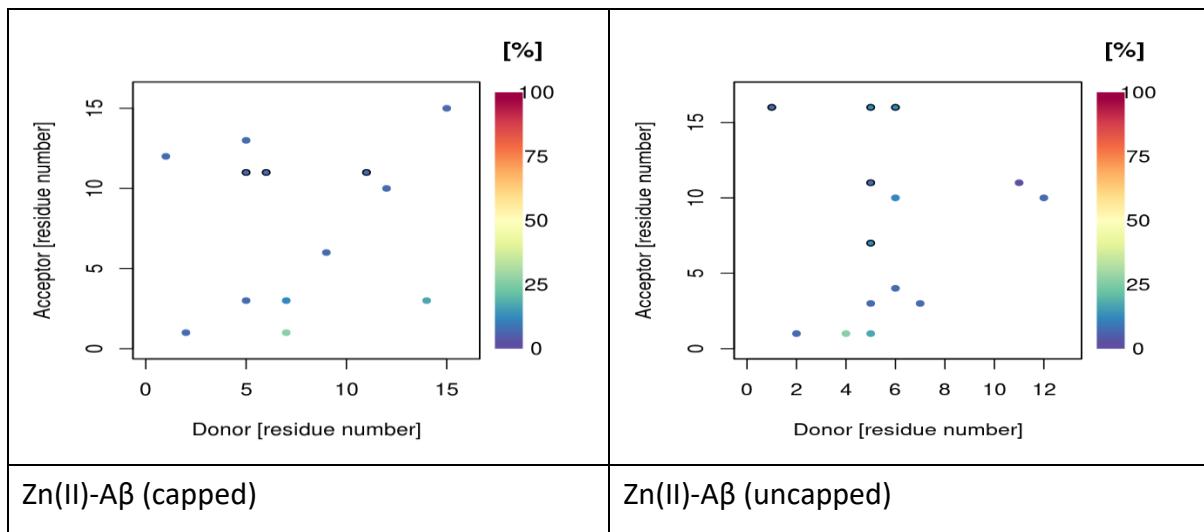


Figure 4.20: Hydrogen bonding occupancy Zn(II)-A β . Plots only display data for hydrogen bonds present for greater than 5% of simulation. Black circles indicate more than one type of hydrogen bond between the relevant residues.

1D free energy surfaces of the capped and uncapped peptides as a function of End-to-end, R_g and HB were calculated and plotted for all simulation. The end-to-end distances show the most stable conformers ($< 5 \text{ kcal}\cdot\text{mol}^{-1}$) were located at ranges from 4 to 16 Å for the capped A β peptide. However, this range is reduced in the uncapped structure, existing only to 14 Å indicating that the end to end distance decreases as charged Lys interacts with other residues as shown in Figure 4.21.

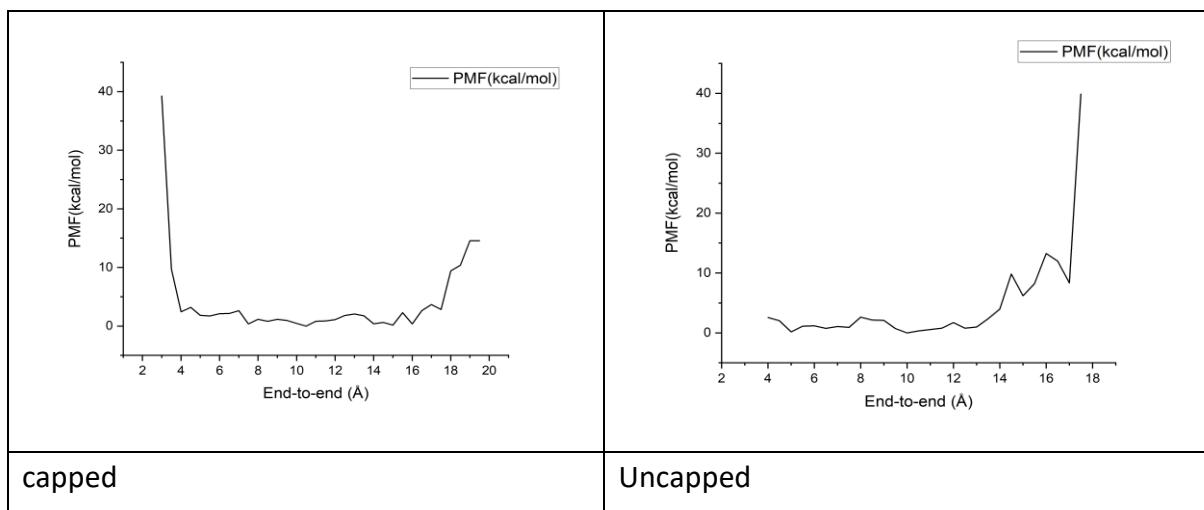


Figure 4.21: 1D-Free energy surface of amidated and non-amidated Zn(II)-A β as a function of end-to-end distance (Å).

The Free energy as function of R_g is shown in Figure 4.22, for both capped and uncapped A β 16. The capped A β 16 displays a narrower range of low energy and the minimum is concentrated at $R_g = 7.2 \text{ \AA}$, but the uncapped peptide occupied a wider range from 6.8 \AA to around 8 \AA , suggesting the uncapped peptide tends to form more extended conformations. Differences in the high energy regions are probably not statistically significant due to insufficient sampling. The most stable conformer of capped A β is concentrated at around 7-7.5 \AA . However, this is located from 7 to around 8 \AA of the uncapped one.

The free energy surface as a function of hydrogen bond count shows similar low energy conformers with values at similar regions for both capped and uncapped. As it is hard to distinguish between them as shown in Figure 4.23. This difference appears to be negligible, which supports data from the HB count values (Table 4.9) mentioned before.

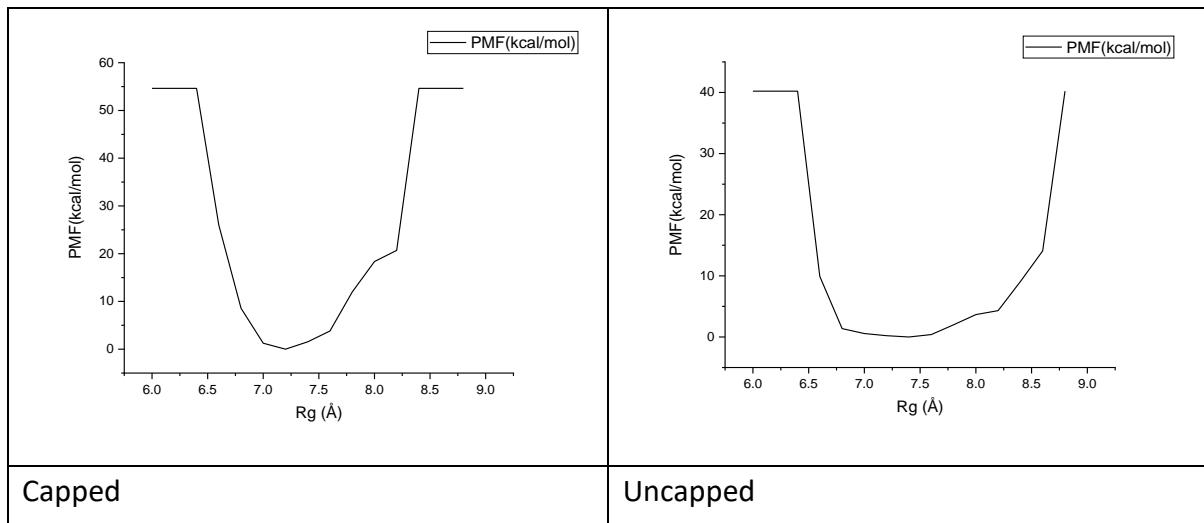


Figure 4.22: 1D-Free energy surface of amidated and non-amidated Zn(II)-A β as a function of $R_g(\text{\AA})$.

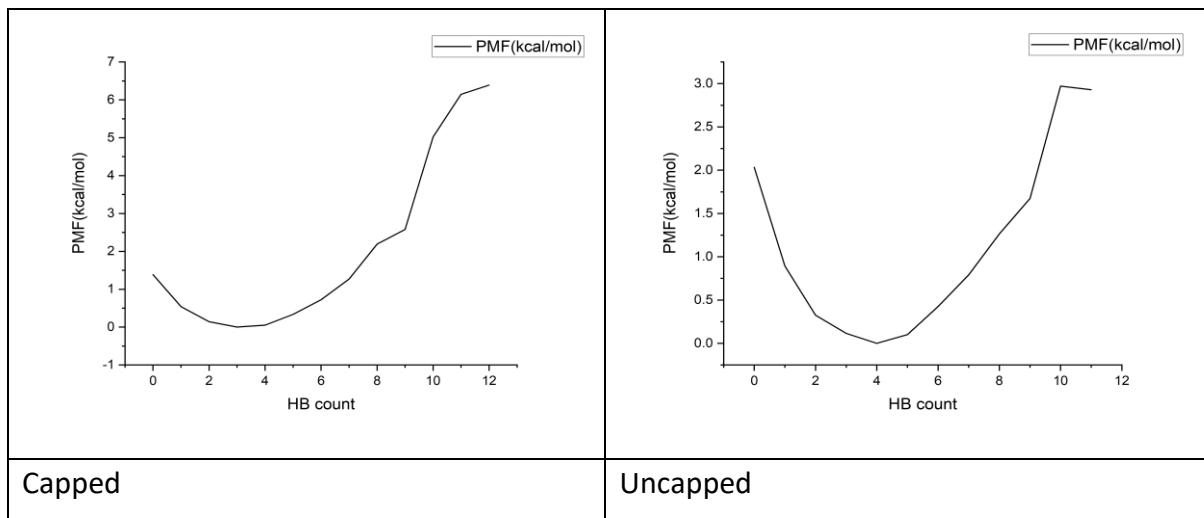


Figure 4.23: 1D-Free energy surface of amidated and non-amidated Zn(II)-A β as a function of HB count.

4.4.3 Conclusion

The overall comparison between capped and uncapped C-terminus shows the amidated peptide decreases the stability of the structure, looking at the RMSD, but not as much as expected if we consider the uncapped peptide. The negative charge is not protected at the end of the chain which makes it available for more electrostatic interaction. This protection of the C-terminus of the peptides investigated led to more β -sheet conformation than the uncapped variant. Also, the distances between Lys16 and the remaining residues are larger in the metal-free, unprotected COO $^-$ of Lys studied previously in this chapter. However, there is no notable difference or increase in the number of hydrogen bonds between them.

4.5 Accelerated Molecular Dynamics Simulation of Fe(II) Binding to N-Terminus of Amyloid- β

4.5.1 Introduction

In contrast to Cu(II) and Zn(II), there are insufficient structural studies on Fe (II) coordination. The original study of Fe(II) interaction with A β (1-16) and A β (1-42),⁵⁴ was carried out using ^1H , ^{13}C and 2D NMR experiments, controlled under anaerobic conditions and at physiological pH. The study concluded that the Fe(II) binding site is centred in the N-terminal (1-16) portion of A β -40. Asp1, Glu3 and the three histidine residues are predominantly

involved in Fe(II) coordination model but not Tyr10 or Met35. The hexacoordination of the Fe(II) ion was proposed in the literature.^{54,64,65} Despite all experimental efforts, the exact binding sites of these amino acids to metal ions have not been confirmed, and need further research. In this section, accelerated MD was performed to simulate the N-terminal A β (1-16) sequence bound to Fe(II) to probe the effect of metal coordination on structure and dynamics. This allows comparisons of the peptide's secondary structure and stability as a result of the effects of Fe(II) when bound via different proposed binding modes (Figure 4.24).

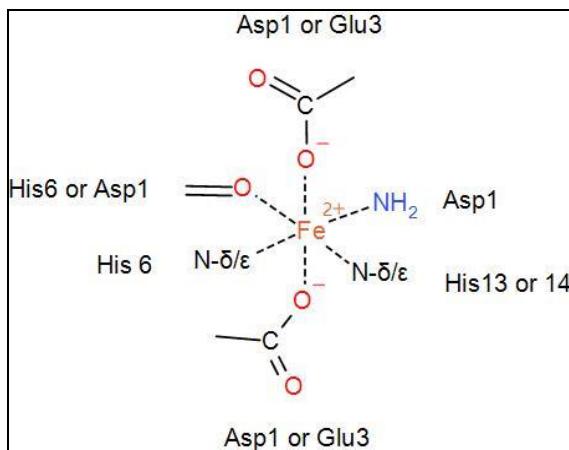


Figure 4.24: General proposed binding modes of Fe(II) to A β -16.

4.5.2 Result and discussion

Accelerated MD simulations were performed using ff14SB for the four complexes of Fe(II) bound to A β ₁₋₁₆ through different proposed residues/atoms, as shown in Figure 4.25 and Table 4.10 found in the literature. The output trajectory was analysed using cpptraj. Comparisons were also drawn between these iron-bound structures and the free A β peptide analysed earlier in this chapter.

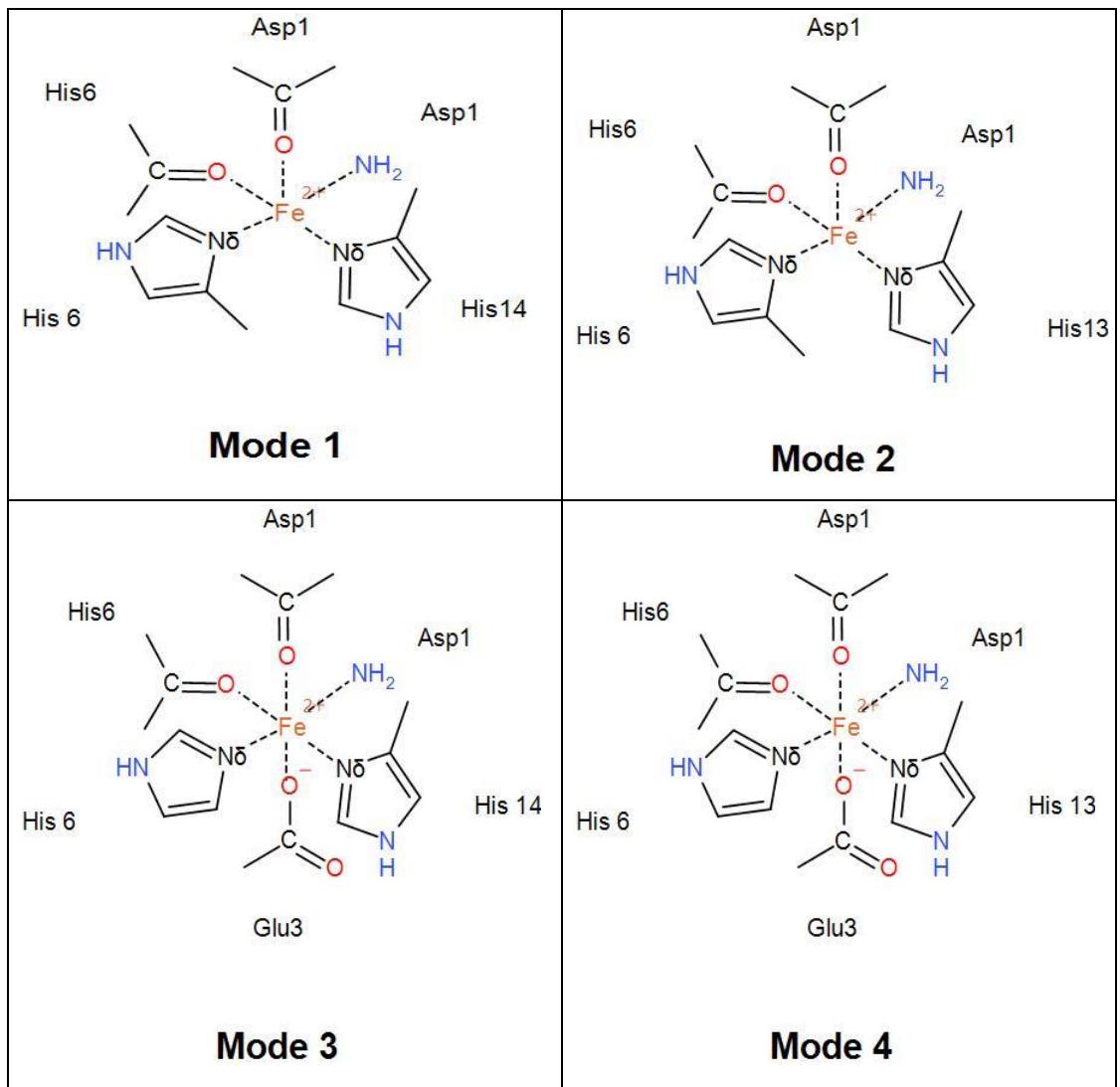


Figure 4.25: Different proposed binding modes of A β (1-16) that are available for Fe(II).

Table 4.10: Different proposed binding modes of A β (1-16) binding to Fe(II).

	Coordination
Mode 1	His6 N δ , His 6 CO, His14 N δ , Asp1 CO, Asp1 NH ₂ ,
Mode 2	His6 N δ , His 6 CO, His13 N δ , Asp1 CO, Asp1 NH ₂
Mode 3	His6 N δ , His 6 CO, His14 N δ , Asp1 CO, Asp1 NH ₂ , Glu3 O ϵ
Mode 4	His6 N δ , His 6 CO, His13 N δ , Asp1 CO, Asp1 NH ₂ , Glu3 O ϵ

Table 4.11 and Figure 4.26 illustrate the backbone RMSD and R_g comparison of Fe binding modes. All modes show significantly lower RMSD and R_g values than free A β -16 which indicates that Fe(II) decreases the mobility of A β when bound to the peptide regardless of

which mode is used. Mode 2 shows little change from its initial position as it shows the lowest RMSD value but it occupies a more extended conformational space compared to other binding modes shown from an increased mean R_g value of 8.37 Å and max at 9.97 Å. Mode 3 on the other hand illustrates the highest RMSD amongst the other iron binding modes with a mean value of 4.81 Å which reflects the high mobility. Figure 4.26 shows the high fluctuation seen in this mode, but R_g values still lie within an average comparable to the other simulations at 7.15 Å, which suggests the conformation size has changed to a similar extent when comparing to other modes investigated.

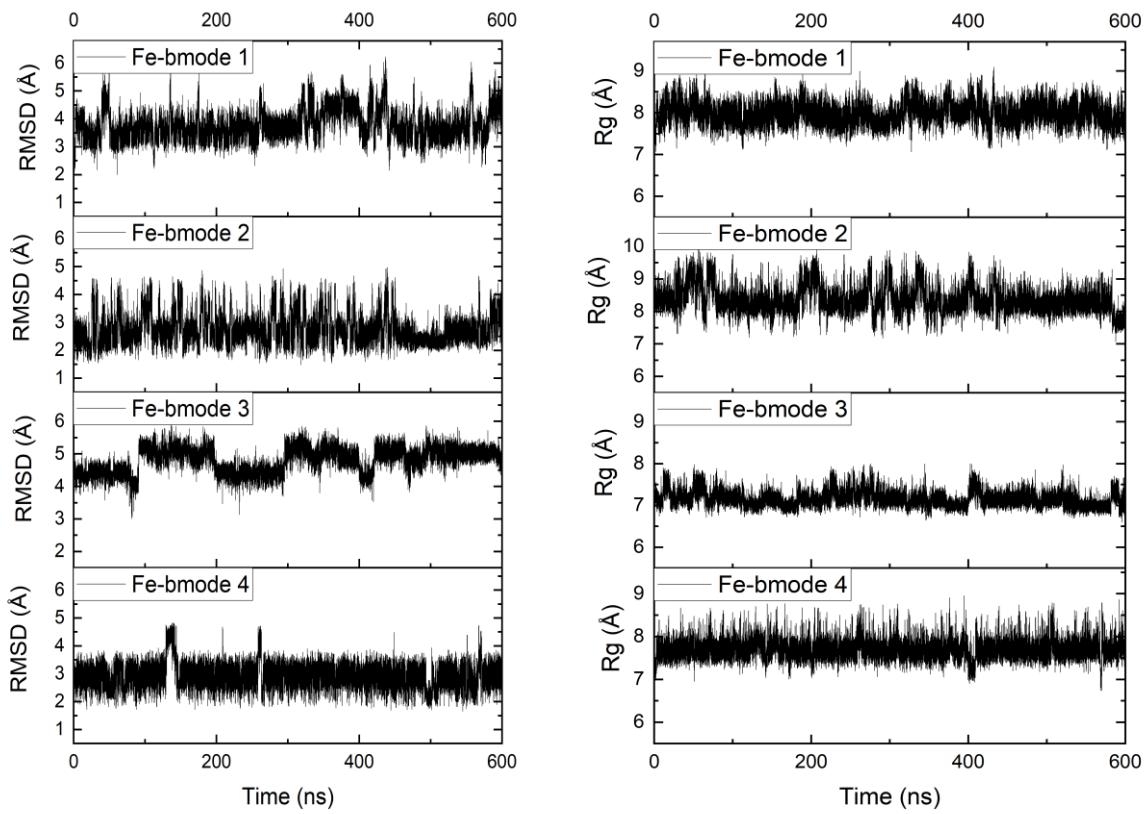


Figure 4.26: Backbone RMSD and R_g values of Fe(II)- Modes and free A β -16. 600 ns of aMD data is reported, made up of the final 200 ns from each of three independent simulations each with different initial velocities.

Table 4.11: Backbone RMSD and R_g of free A β -16 and Fe(II) binding modes, (\AA).

Fe(II)-A β 16	RMSD				R_g			
	Mean	sd	Min	Max	Mean	sd	Min	Max
Mode 1	3.92	0.68	2.10	6.19	7.94	0.26	7.19	9.40
Mode 2	2.76	0.56	1.48	4.98	8.37	0.42	7.07	9.97
Mode 3	4.81	0.38	3.01	5.92	7.15	0.18	6.60	8.00
Mode 4	2.93	0.46	1.65	4.73	7.70	0.23	6.74	8.87
Aβ16	10.4	1.57	4.45	15.46	8.23	0.77	6.80	13.24

RMSF values in Figure 4.27 signify that A β -16 has markedly higher fluctuations compared to all Fe(II) modes. This clearly shows that the 16 residues of the peptide exhibited a more mobile set of structures when a metal centre was absent which indicates the behaviour of the peptide when it is unrestricted by metal influences. In contrast, the modes of iron show much less flexibility and more compact conformations with highly similar trends between trajectories. The pattern observed across all binding modes starts with lower RMSF values that appear to increase overall along the chain to the higher values observed at the end of C-terminus. Unlike the RMSF of the residues in the Zn(II) binding modes, where Gln5, Asp7, Tyr10, and Lys16 presented more transitions from their initial positions (Figure 4.27), this flexibility is only expressed in Tyr10 and Lys16 of the iron binding modes.

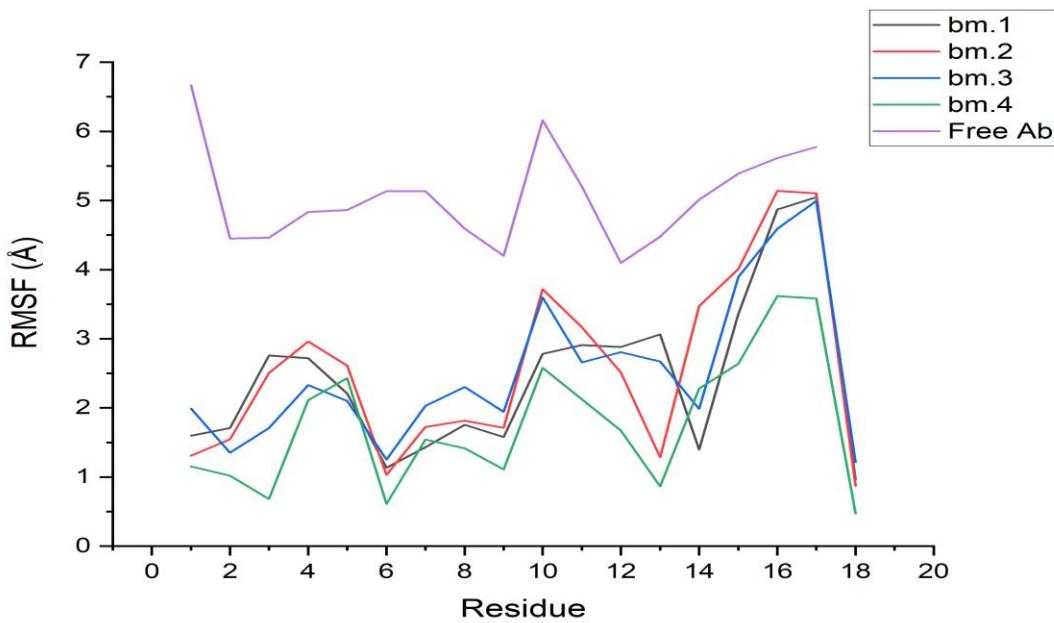


Figure 4.27: RMSF of each residue of free and Fe(II)-bound A β -16. The metal ion itself is denoted residue 18, and the C-terminal amidated cap residue 17.

The most populated percentage cluster of all trajectories was found in Mode 4 of Fe(II)-A β at 92% of the combined trajectory for that particular structure where only four clusters were found. Mode 1 and 3 possessed the next greatest populated clusters with 9 and 3 clusters at a similar occurrence of 64 and 63%, respectively. However, the trajectory for binding Mode 2 displays the highest number of clusters at 19 with the most populated cluster including only 19% of the combined trajectory. Representative structures of the highest occurring cluster for each structure can be found in Figure 4.28. The second highest clusters by populations can also be found in Table 4.12. The highest value of these was found to belong to mode 3 in which the second most populated cluster was made up of 33% of the entire frames.

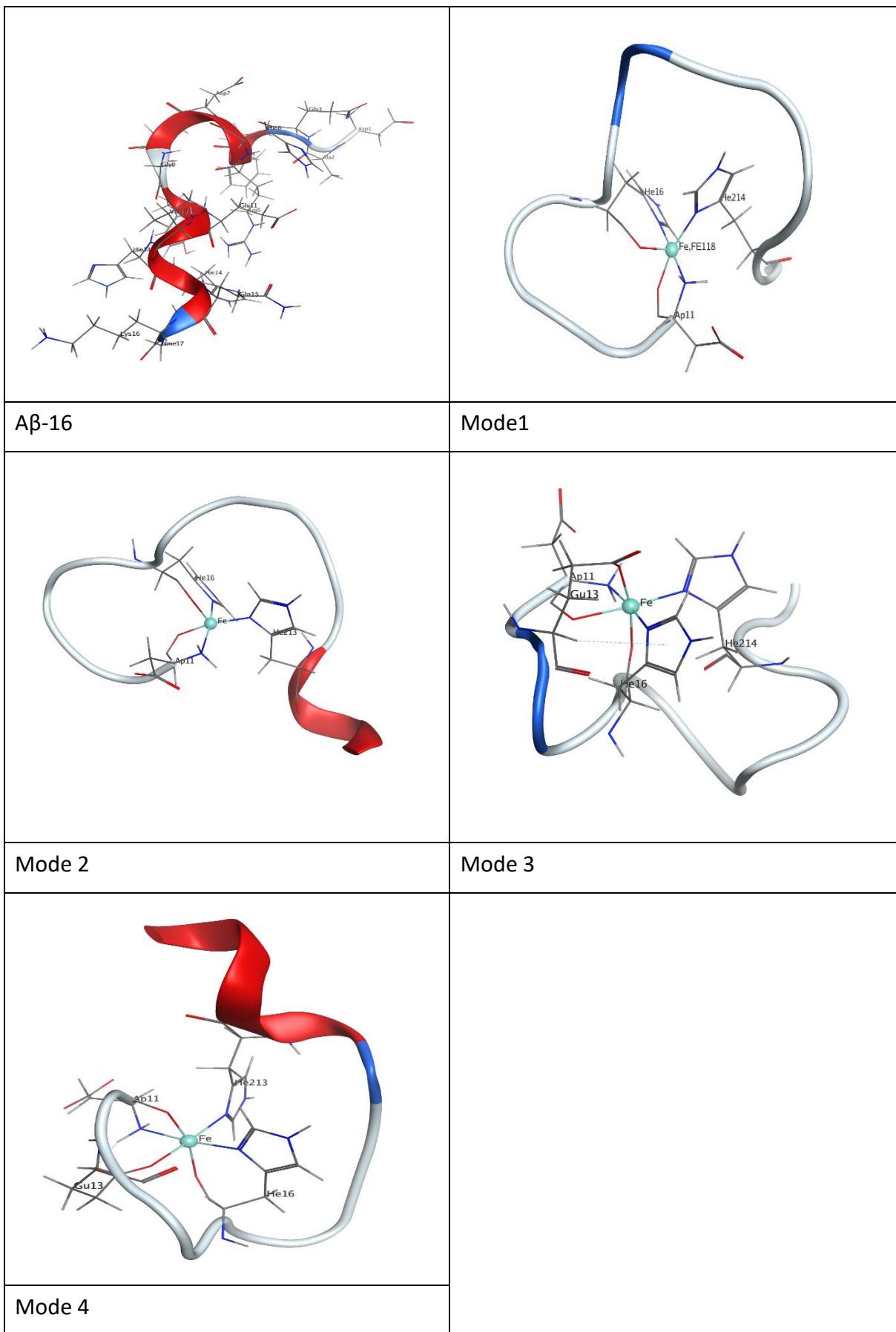
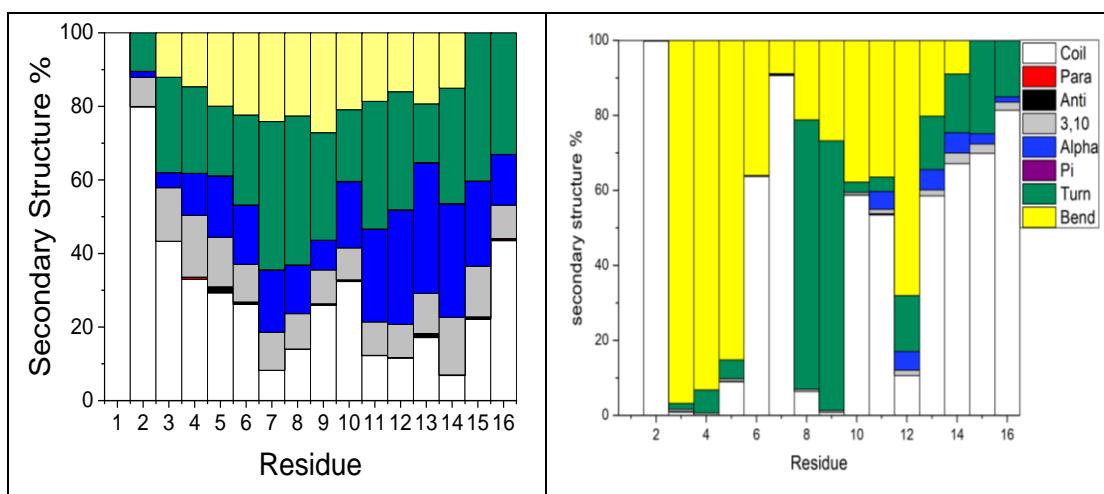


Figure 4.28: Ribbon diagram of the most populated clusters for Fe(II) and free A β .

Table 4.12: Cluster analysis data for equilibrated trajectories.

Clusters	# clusters	Most populated	2 nd most populated
Mode 1	9	64%	7%
Mode 2	19	15%	12%
Mode 3	3	63%	33%
Mode 4	4	92%	2%

Binding Mode 1 and 3 have similar coordination modes and Mode 2 and 4 look close to each other, but binding modes 3 and 4 have an extra axial bond, due to coordination to COO⁻ from Glu3, with the only difference being the binding to either His13 (modes 1 and 3) or 14 (mode 2 and 4). Thus, the secondary structures of those analogous pairs of binding modes are expected to be similar, as shown in Figure 4.29 and Table 4.13. Mode 1, 2 and 4 show similar secondary structure existing from residue 11 onwards to the metal-free peptide, which contains levels of turn, alpha and 3₁₀ helices at increasing percentages for modes 1, 2 and 4 respectively. Mode 4 has the highest helical content at *ca.* 23% of structures. Turn, Bend, and Coil configurations occupy the majority of Mode 1, 2, and 3 at levels > 91% of their structures. Mode 3 however is totally different in which there is no significant helical structure observed, though anti- and parallel-β sheets are detected. Figure 4.29 shows the secondary structure profile of Mode 4 is not the same as free Aβ-16 despite possessing similar percentages of helical character at 23.2 and 27.2% respectively. Binding mode 4 shows all helix structures from residue Glu11 onwards whilst free-Aβ has α-character across the entire structure but at lower percentages per residue.



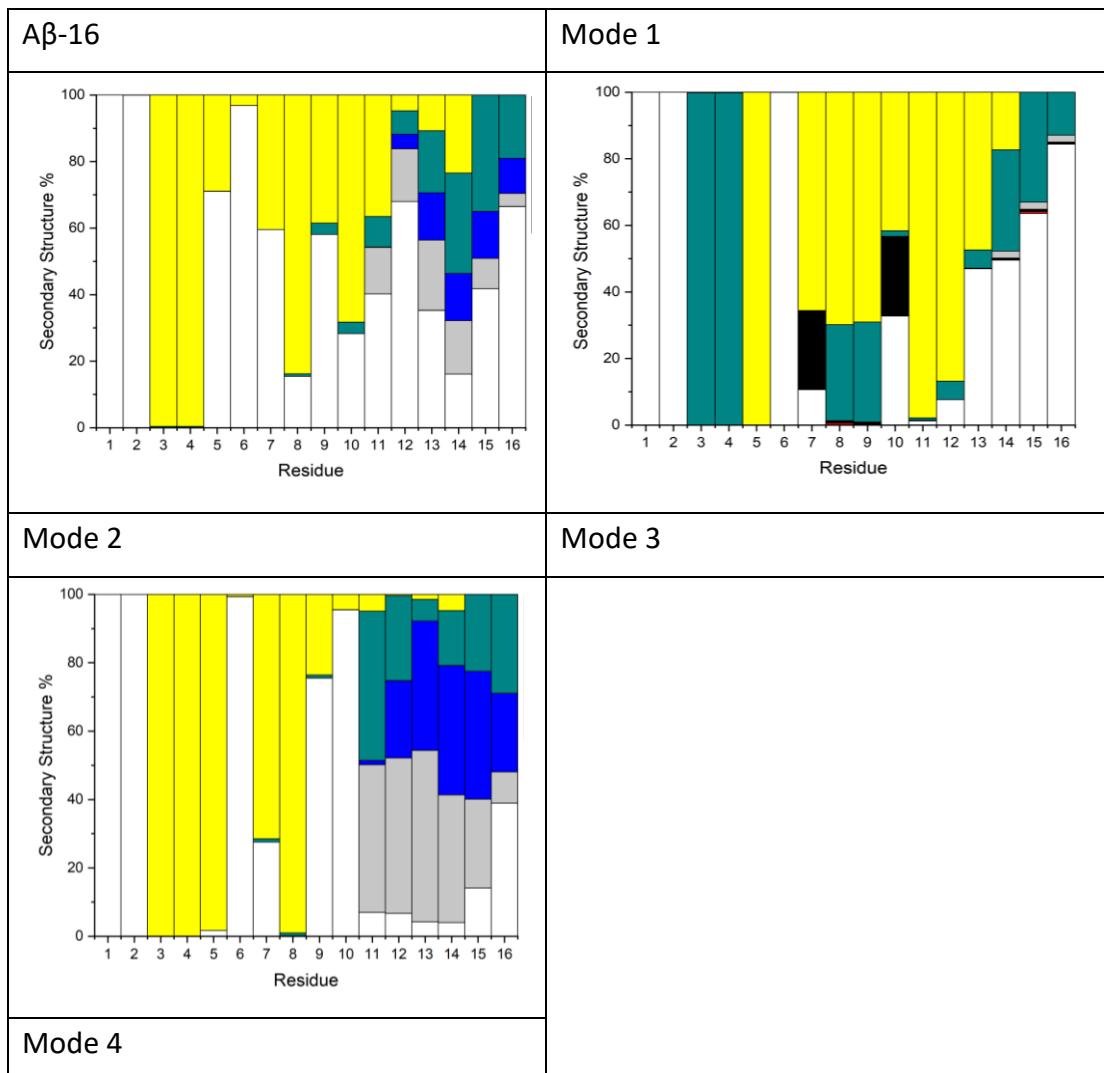


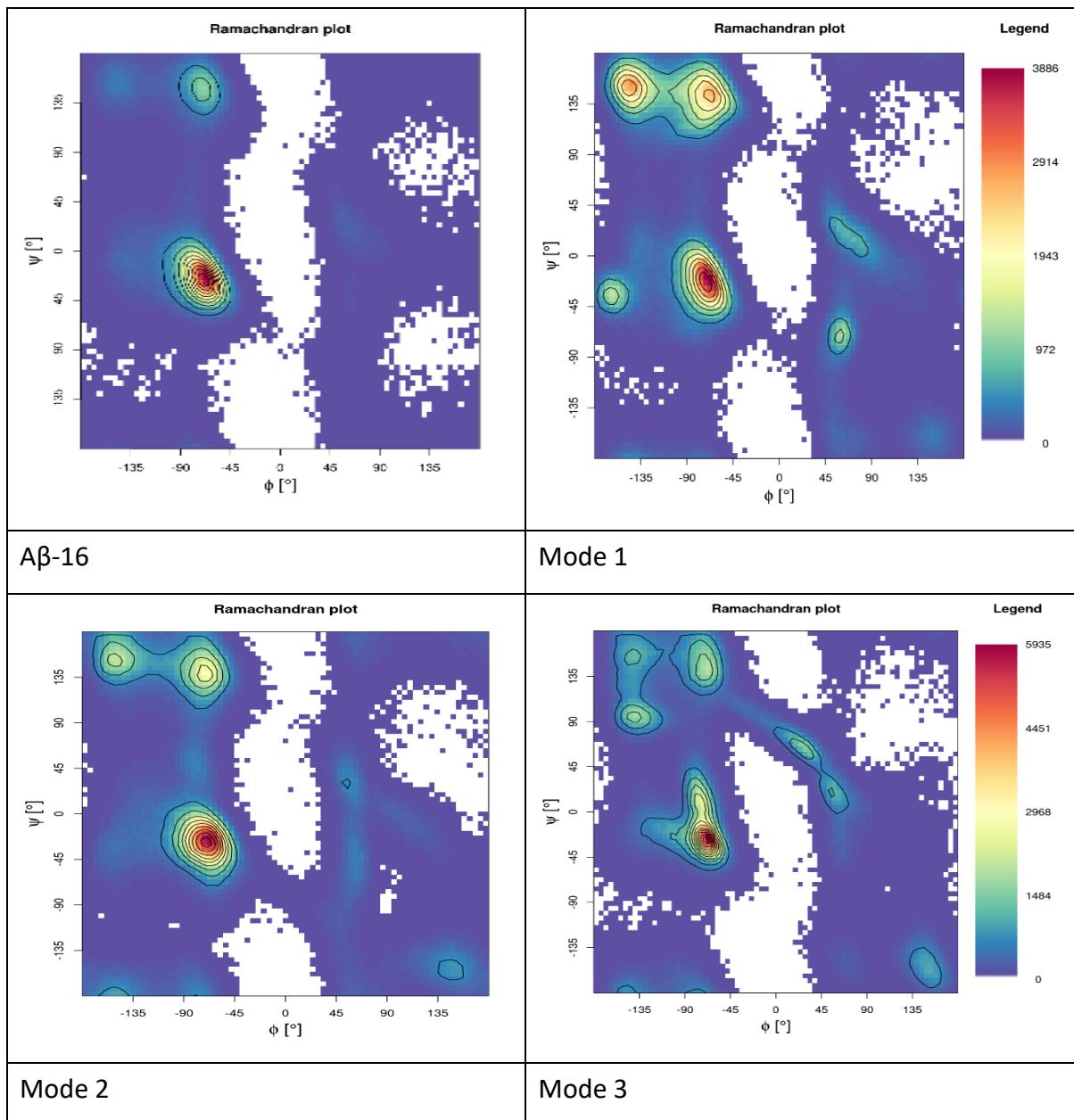
Figure 4.29: Secondary structure for Fe(II)-A β and free A β -16.

Table 4.13: Total secondary structure percentages for each Fe(II) system (%).

	Helix	Sheet	Other
Mode 1	2.50	0.09	97.41
Mode 2	8.59	0.02	91.39
Mode 3	0.40	3.27	96.33
Mode 4	23.20	0.00	76.80
Aβ	27.2	0.4	72.4

The Ramachandran maps of different modes shown in Figure 4.30 show that the right-handed helices region have the greatest population density seen in all modes. The parallel β -sheet

region is also seen to be inhabited at different concentrations but Mode 1 clearly shows the most. Mode 3 shows a small amount of anti-parallel β -sheet arrangement where the peptide backbone dihedral angles (ϕ, ψ) are about $(-135^\circ, 90^\circ)$ in antiparallel sheets.



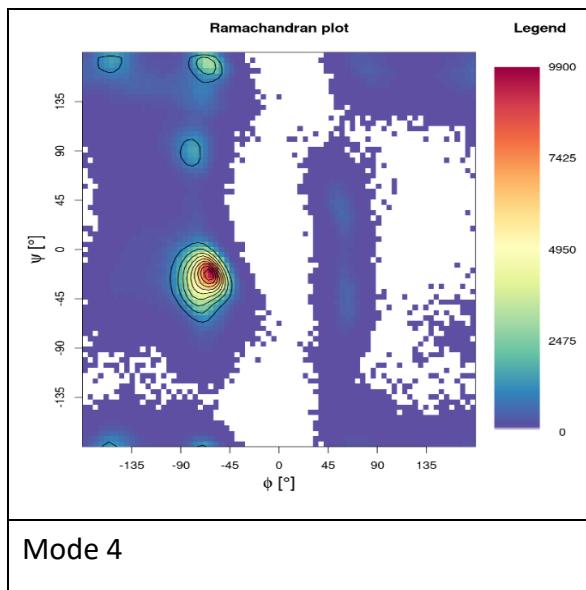


Figure 4.30: Ramachandran maps for Fe(II)-A β and free A β -16. The legends represent the number of dihedrals per bin.

Figure 4.31 shows free A β -16 peptide has different salt bridge contacts found when Fe is bound to it, and they differ between binding modes. The most notable interaction found between Arg5 and Asp7 in Mode 1 at a percentage of around 91% of the whole simulation compared to lower incidences observed for this contact occurring between 60-80% of frames in mode 2 and 4. These binding modes presented the lowest average RMSD values, indicating the effect of this interaction (Arg5-Asp7) on the peptide's structural changes, compared to the initial minimized structure. The strong salt bridge interaction seen in the free peptide between Arg5 and Glu11 is seen also in mode 3 but at a much higher occupancy than the free one. In contrast, this contact is missing or transient in all remaining modes. Also, another significant contact around 50% of frames between Asp7 and Lys16 observed in this mode. This may potentially contribute to the change observed earlier in the secondary structure of Mode 3.

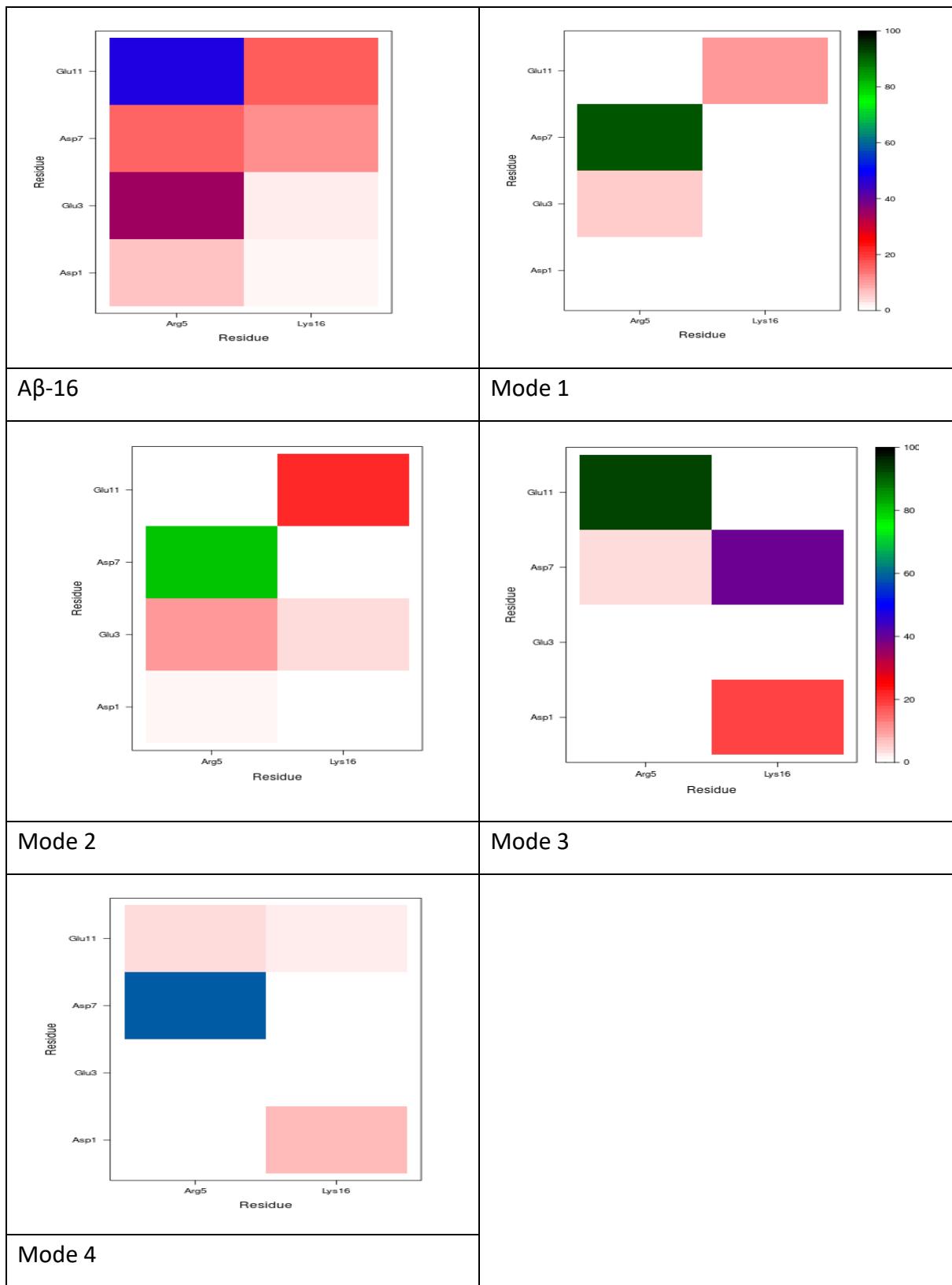
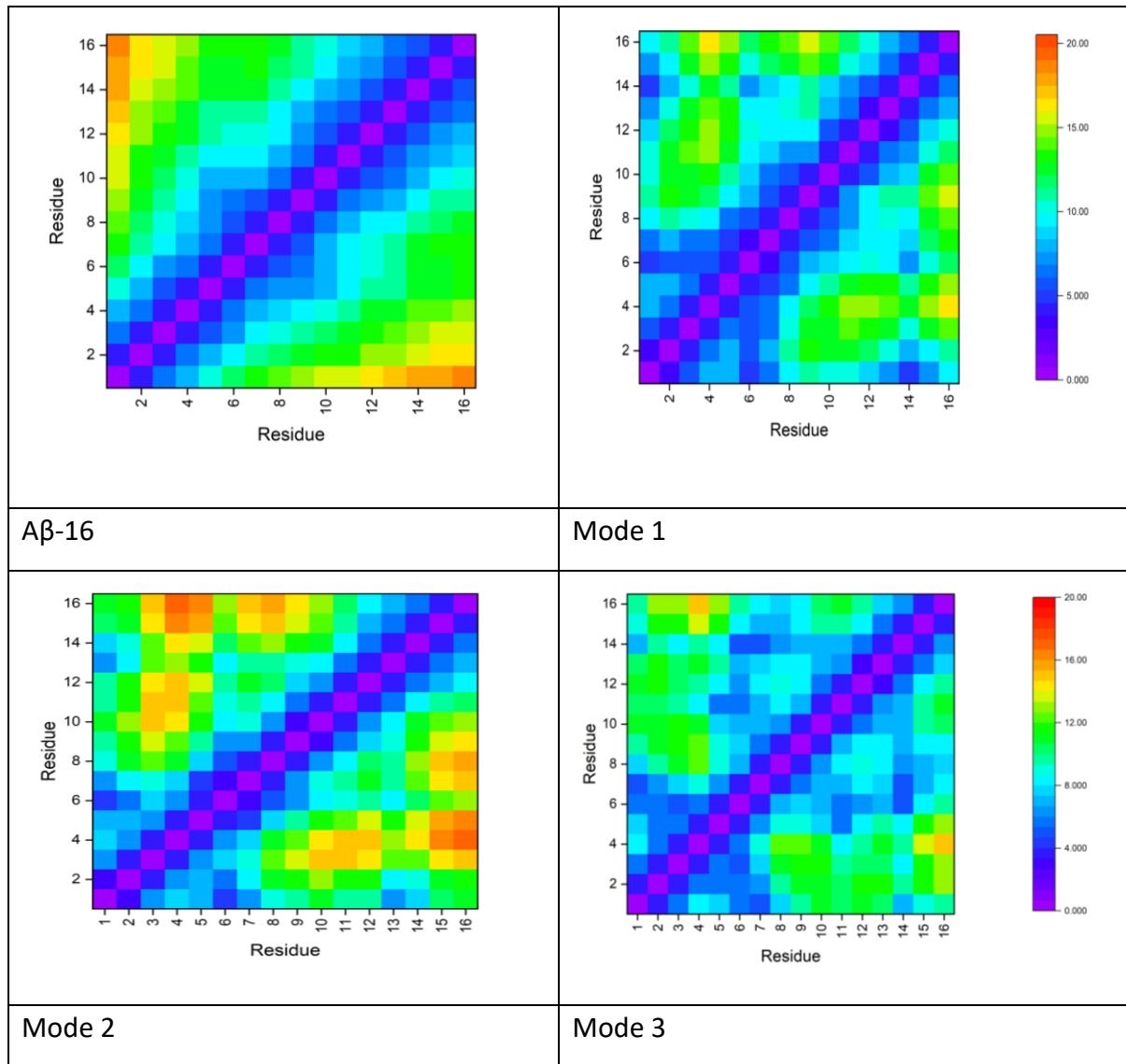
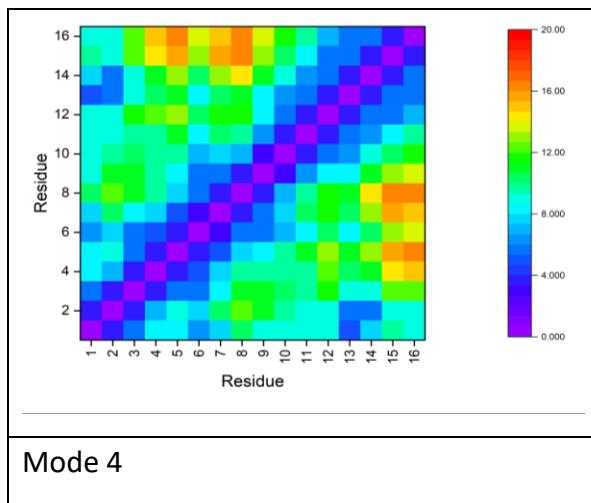


Figure 4.31: Salt bridge maps of Fe(II)-A β all modes and free A β -16.

Figure 4.32 illustrates the contact distance between two residues within the peptide for each modes of Fe-A β . From the figures it is clear that there is long distance detected between positively charged Lys16 and residues Phe4 and Ser8 in mode 2 and 4. In general the residue interactions for modes 1 and 3 are similar to each other and 2 and 4 are similar too which is expected as they have similar binding modes.

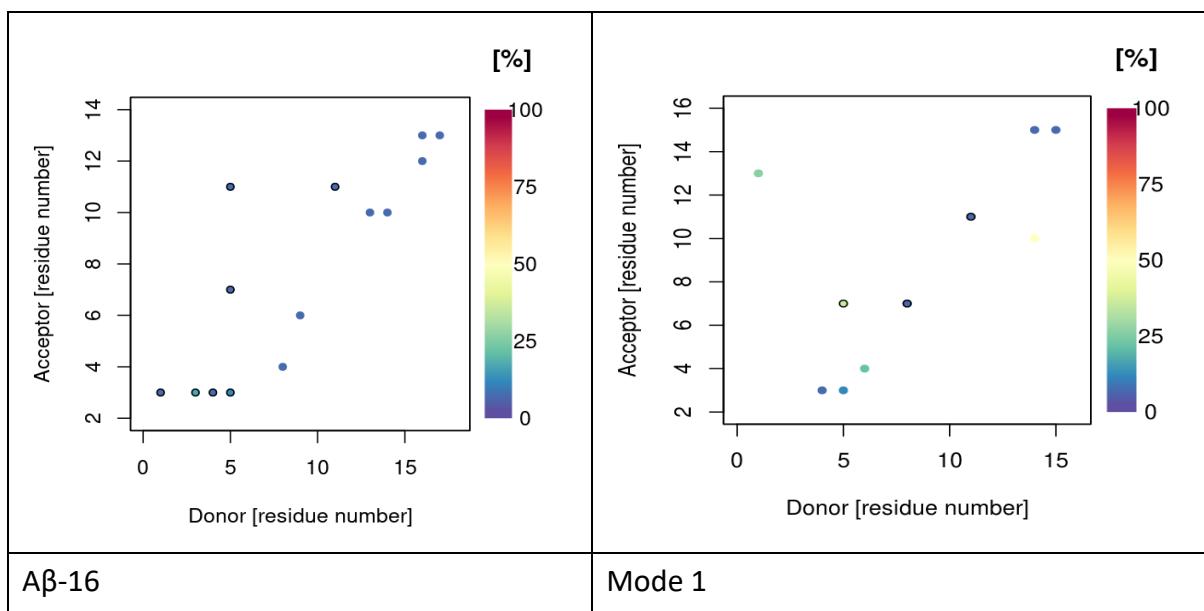




Mode 4

Figure 4.32: Contact map of binding modes of Fe(II)-A β and free A β -16. The legend shows the average inter-residue distance in \AA .

Hydrogen bond counts of the peptides can explain the spatial arrangement of the structures and their stability. The HB counts for Fe binding modes are presented in Table 4.14 and Figure 4.33. Mode 3 that presented anti-parallel β -sheets in its secondary structure plots showed the greatest average and maximum amount of HB occurred between Arg5 (donor) and Ala2 (acceptor) in 70% of frames. However, binding mode 2 shows the lowest maximum number of HB at only 10, and all the HB observed are not fixed or continuous through the simulation as all interaction last less than 25% of the combined trajectory data for that binding mode. In all binding modes, there is at least one frame that contains no H-bonds. This shows the mobile nature of this peptide (A β) with folding and unfolding occurring throughout the simulation.



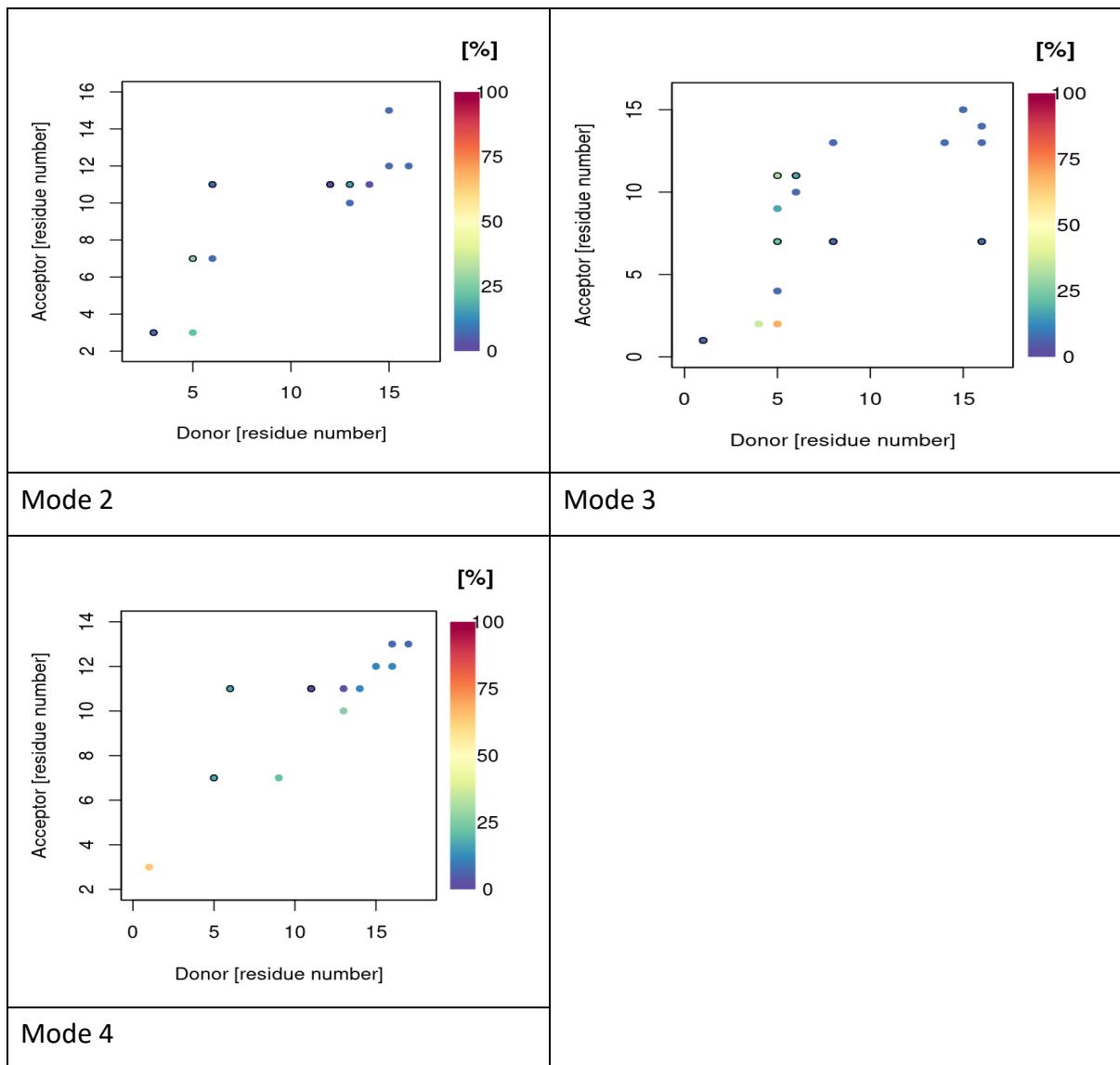
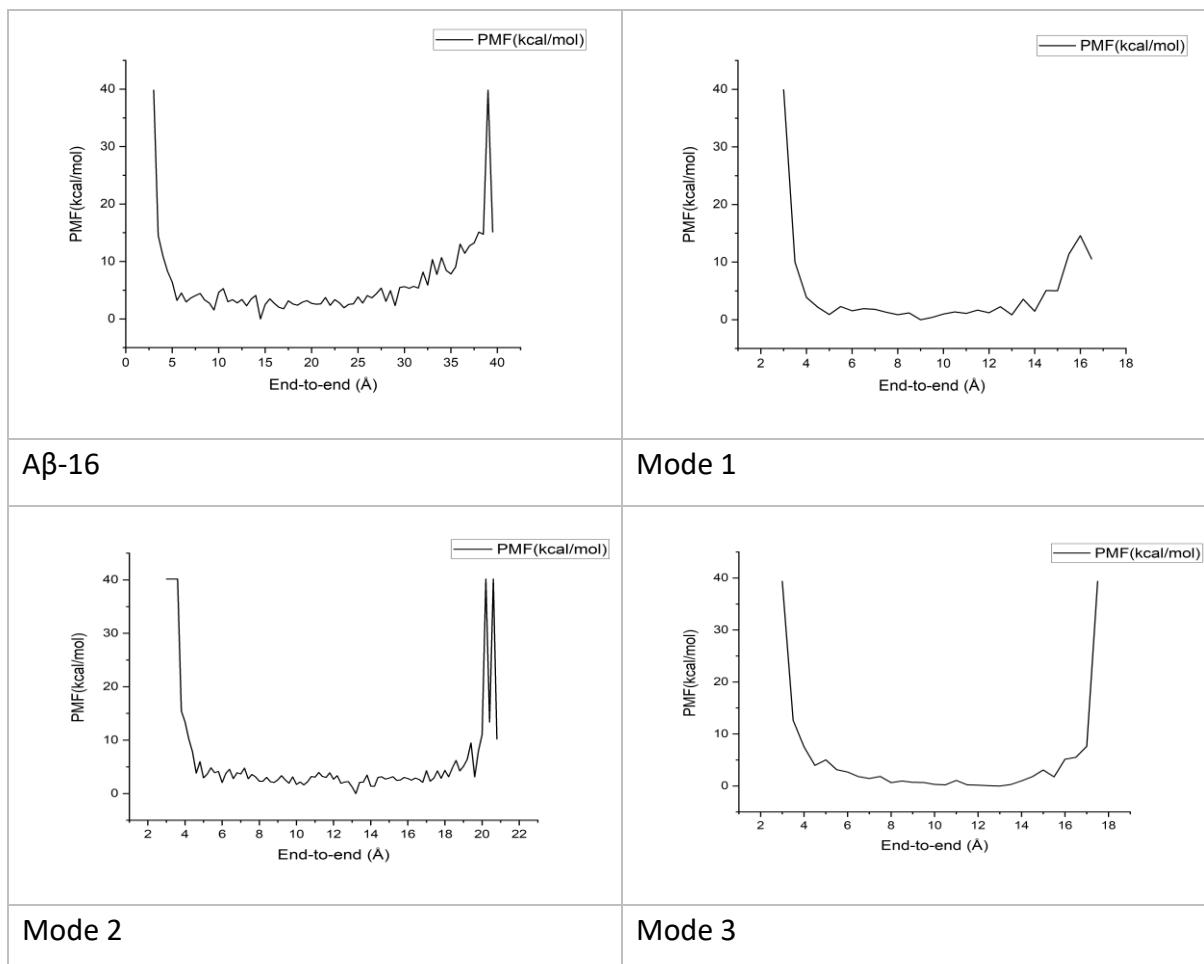


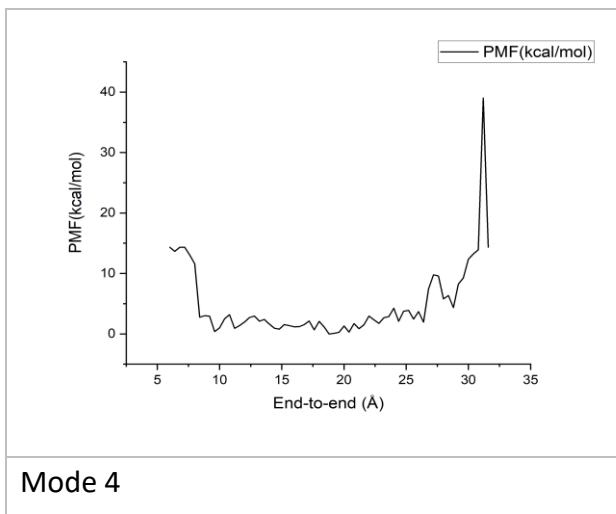
Figure 4.33: Hydrogen bonding occupancy Fe(II)-A β and free A β -16. Plots only display data for hydrogen bonds present for greater than 5% of simulation. Black circles indicate more than one type of hydrogen bond between the relevant residues.

Table 4.14: HB count for Fe(II)-A β and free A β -16.

Number of HB	Mean	sd	Max	Min
Mode 1	3.88	1.53	11	0
Mode 2	3.40	1.51	10	0
Mode 3	5.26	1.71	13	0
Mode 4	3.46	1.54	11	0

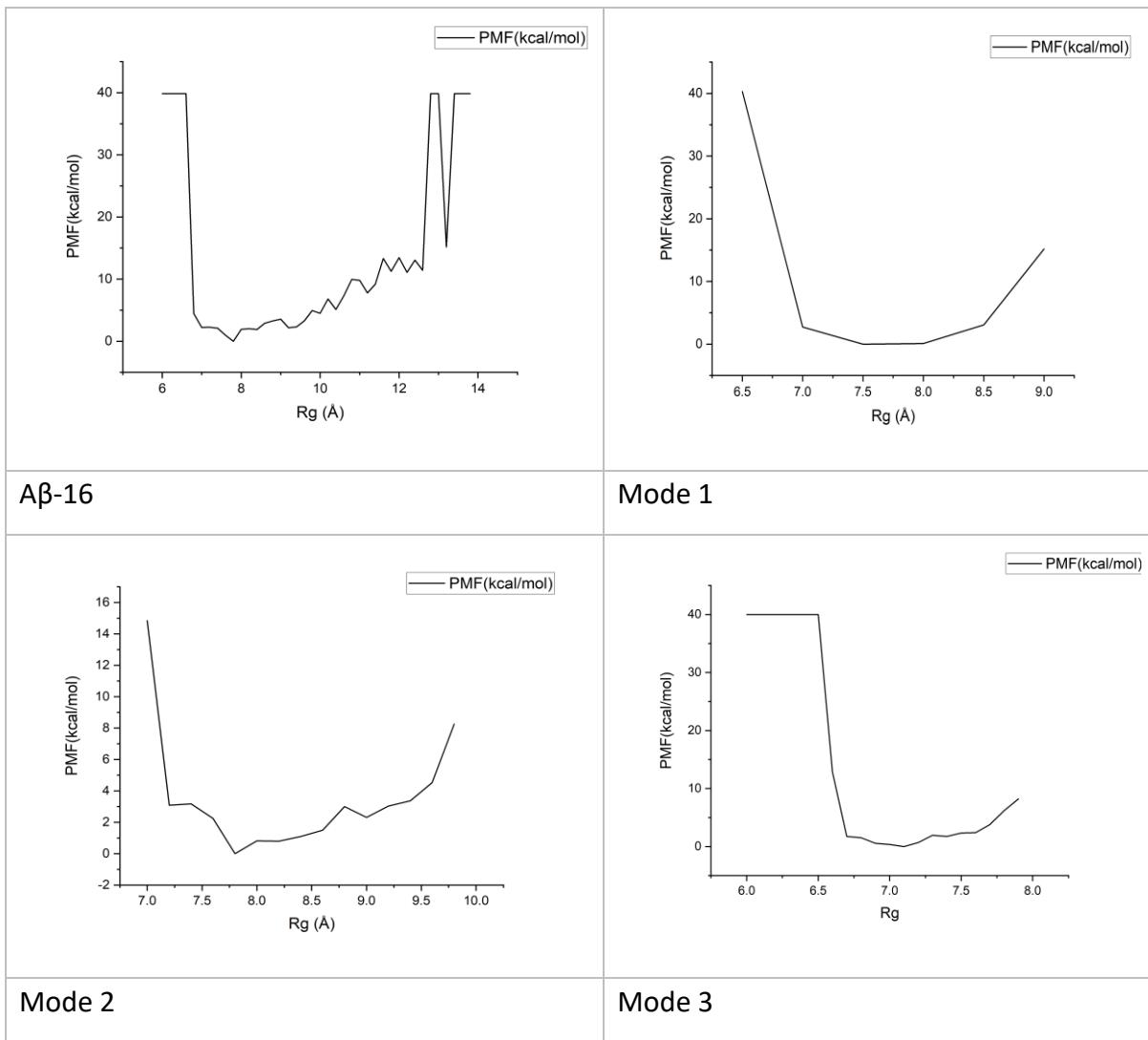
One dimensional free energy landscapes as a function of end-to-end distance (Figure 4.34) and R_g (Figure 4.35) have been generated in order to locate the most stable conformers on the free energy surfaces. End-to-end distance of the free peptide shows a very wide range of values (5-35 Å) that contain the conformers sampled of energy less than 5kcal.mol⁻¹. However, all binding modes have a lower range that most stable conformers can be found at around 4-18 Å except binding mode 4 in which the value approaches 25 Å. The most stable structures of Fe binding modes were located at ranges between 7-7.5 Å. However, the free Aβ's lowest energy value was found to be at around 8 Å.





Mode 4

Figure 4.34: 1D-Free energy surface of binding modes of Fe(II)-A β and free A β -16 as a function of end-to-end distance (\AA).



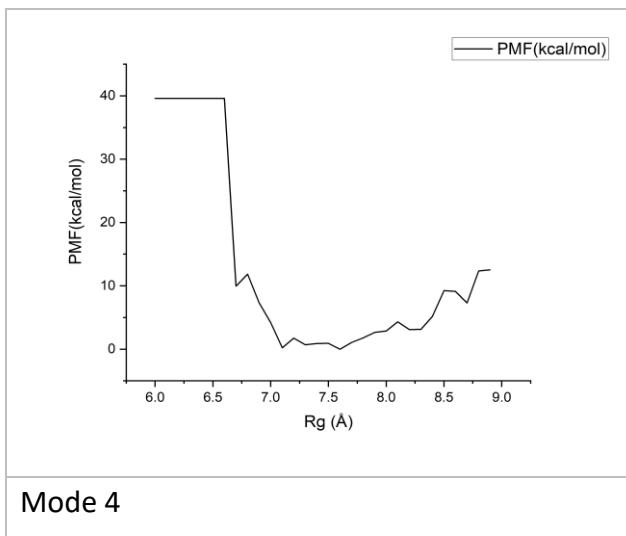


Figure 4.35: 1D-Free energy surface of binding modes of Fe(II)-A β and free A β -16 as a function of R_g (\AA).

4.5.3 Conclusion

The aMD simulations of different binding modes of Fe(II) show that the coordination number the atom/residue type can change the overall structure, size, and stability. Binding Mode 1 and 2 have a coordination number of 5, however binding mode 3 and 4 have one more at 6. Binding mode 3 shows different secondary structure and different salt-bridge interaction while the rest show high similarity. In addition, binding mode 1, 2, and 4 present lower RMSD and R_g value than Mode 3.

4.6 Accelerated Molecular Dynamics Simulation of Cu(II) Binding to N-Terminus of Amyloid- β

4.6.1 Introduction

A range of binding sites for Cu(II) to A β had been proposed in the literature.^{38,45} EPR study of Cu(II) bound to soluble and fibrillar A β (1-40) indicated that three nitrogens and one oxygen donor atoms from A β can coordinate with Cu(II). The donor atoms were sourced from three His residues (His6, His13, and His14) with an Asp1 carboxylate.⁶⁶ Typically, Cu(II) binding to A β involves three N donor atoms and at least one O donor, however it is still disputable which

of the N/O donors are engaged in the coordination. Potentiometric and spectroscopic (UV–Vis, CD, EPR) studies on Cu(II) bound to human A β (1-16) and (1-28) in aqueous solution at the pH range between 2.5 – 10.5 indicated the high-affinity Cu(II) binding sites are found in the N-terminal residues, ranging from 1 to 16 for near neutral pH. The study assigned two independent 3N1O Cu(II) coordination modes components (component I and II, respectively), in which they are found in lower and higher pH respectively. For component I, two types of binding modes (Ia and Ib) are in equilibrium at (1:1) ratio between pH 6-7. Component Ia includes the terminal NH₂ (from Asp1), the CO from the Asp1–Ala2 peptide bond, and a N atom from the imidazole ring of His6 and His13 {N_a^{D1}, O, N_{Im}^{H6}, N_{Im}^{H13}}. Alternatively, for component Ib, His13 can be exchanged with His14 {N_a^{D1}, O, N_{Im}^{H6}, N_{Im}^{H14}}. Components Ia and Ib are predominantly involved in equatorial Cu(II) binding; at pH 8.0 a coordination mode of component II was identified, that is less predominant than I,⁵⁵ with regards to the apical positions proposal, involving carboxylate groups from any one of Asp1, Asp7, Glu3, and Glu11.

A recent solid state NMR study suggested the participation of Glu side-chains and the C-terminus carboxylate group, as well as His13 and His14 through the N ε atom of the imidazole ring. In fact, in the NMR studies on the soluble peptide show all the carboxylate groups have been also affected by the presence of the Cu(II) paramagnetism.⁶⁷ Nevertheless, they were not all mutually involved in Cu(II)-binding. This was due to Cu(II) substitution between ligands that may also were affected in the case of the solid state NMR study. Fourier transform IR study pointed the effect of Cu(II) binding on the protonation state of the His residues when involved a proton transfer occur on the nitrogen atoms of the imidazole ring. The same study also the study suggested Cu (II) bound to His6 via N τ , and His13/His14 via N τ at low pH, while at high pH the Cu(II) bound to His6 atom via the N τ .⁶⁸

A multifrequency CW-EPR spectroscopy study analysed the superhyperfine interactions between Cu(II) and A β . The study also identified Ia and Ib modes of component I that were proposed by Kowalik-Jankowska and co-worker.⁹ They found component Ia composed of {N_a^{D1}, O, N ε ^{H6}, N ε ^{H13}} and component Ib is composed of {N_a^{D1}, O, N ε ^{H6}, N ε ^{H14}} between pH 6-7.³³ A recent study involving Fourier Transform IR Spectroscopy (FTIR) on component I and II of Cu(II) at 6.5 and 9.0 pH agreed with another study on Cu(II) coordination via the carboxylate side chain of the Asp1 residue at both pH values, used in a pseudo-bridging monovalent process. At a more acidic pH, His6 binds Cu²⁺ through N ε , whereas His13 and His14 are bound

through N δ . At basic pH, clear structure is given on the coordination of Cu(II) via the N ϵ atom of His6.

In this section, accelerated MD simulation was used to simulate Cu(II) coordinated via different binding modes of the N-terminal A β (1-16) sequence to investigate the effect on structure and flexibility of A β peptide in the presence of a copper centre. The binding modes examined were found from literature as shown in Figure 4.36 for a general structure and Figure 4.37 and Table 4.15 for a more detailed representation of coordination sites.

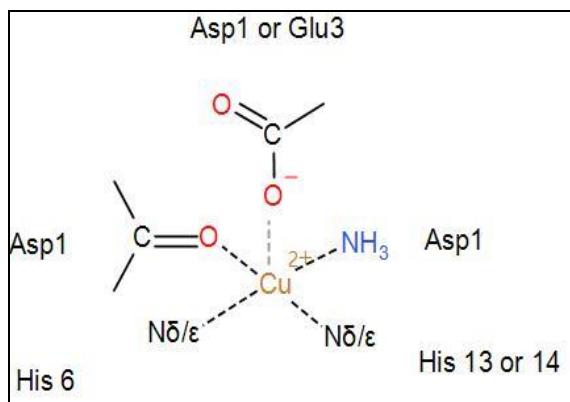
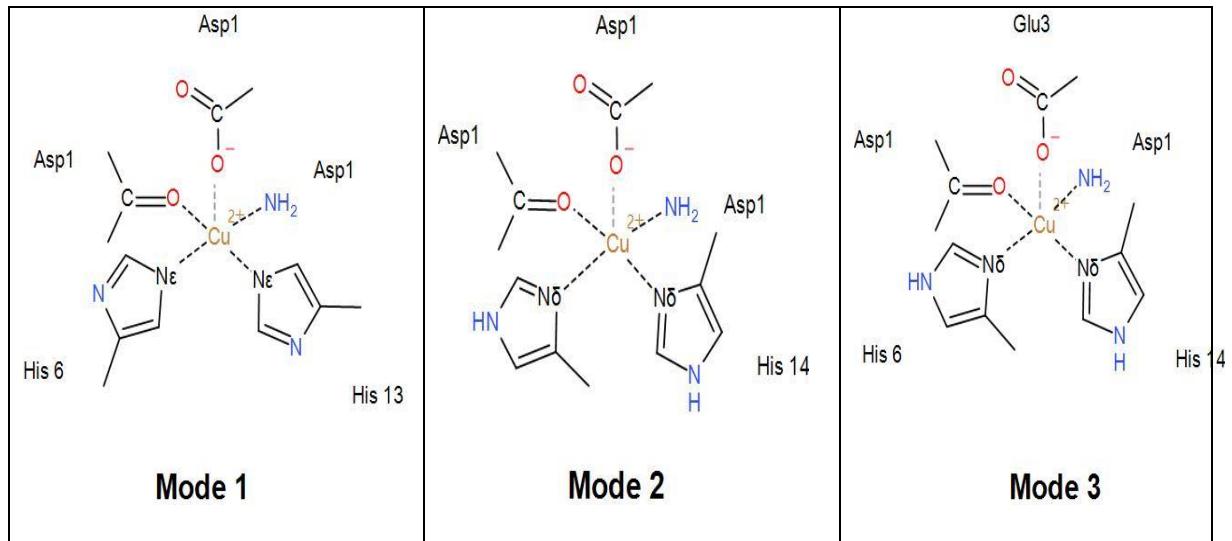


Figure 4.36: General proposed binding modes of Cu(II)



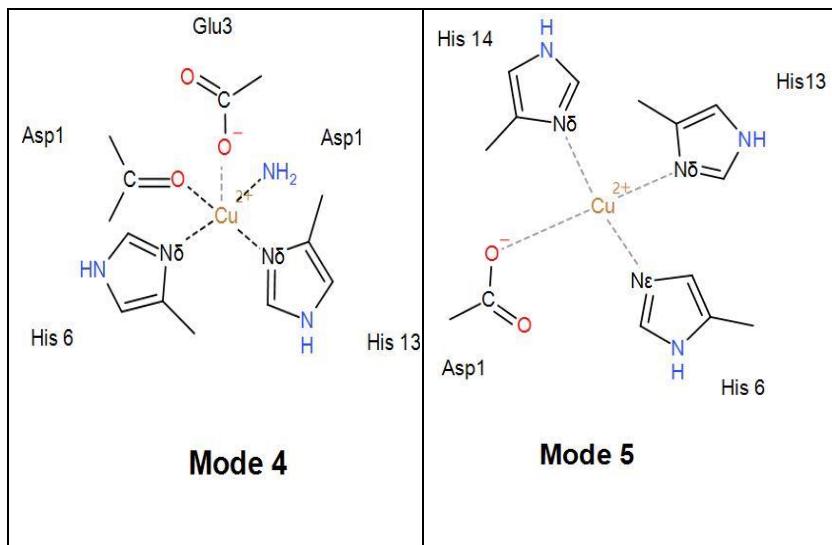


Figure 4.37: Different proposed binding modes of A β (1-16) binding to Cu(II).

Table 4.15: The table of different proposed binding modes of A β (1-16) binding to Cu(II).

Mode number	Coordination
Mode 1	His6 N δ , His13 N ϵ , Asp1 CO and NH2
Mode 2	His6 N δ , His14 N δ , Asp1 CO and NH2
Mode 3	His6 N δ , His14 N δ , Asp1 CO and NH2, Glu3 O ϵ
Mode 4	His6 N δ , His13 N δ , Asp1 CO and NH2, Glu3 O ϵ
Mode 5	His6 N ϵ , His14 N δ , His13 N δ , Asp1 CO

4.6.2 Result and discussion

As before, three 200 ns aMD simulations for each binding mode were combined and analysed and the structure of the most populated clusters were checked. Modes 1, 2, and 4 take distorted square planar more than mode 3 which shows square pyramidal coordination.

The backbone RMSD and R_g analysis in Figure 4.38 and Table 4.16 show that Cu(II) affects free peptide flexibility as seen by the reduction in the RMSD values from an average of around 11 Å to less than 5 Å after copper binding. Mode 1 shows the highest fluctuation in values during the simulation which means this binding mode is less stable proven by the greatest SD value among other modes of (1.14 Å) showing a higher variance in structures. Interestingly, this

value is still less than the SD of the RMSD for the free peptide. However, RMSD values for Mode 1 are comparable to other binding modes of Cu(II) at an average of 4.51 Å and a maximum of 6.87 Å. Binding modes 2, 3 and 5 illustrate the lowest values of both R_g and RMSD indicating the structural stability of these specific binding modes compared to the rest. On other hand, binding Mode 4 presents slightly higher RMSD and R_g averages at similar values of around 5 Å and 8.2 Å, respectively, but it is still observed to be notably lower from the free peptide values. Its structure goes through states of extending and folding which can be seen through minimum and maximum R_g values of lows around *ca.* 7 Å, which increase during the simulations to reach approximately 11-11.5 Å which in turn affect the high R_g mean value (*ca.* 8.2 Å) shown in Table 4.16.

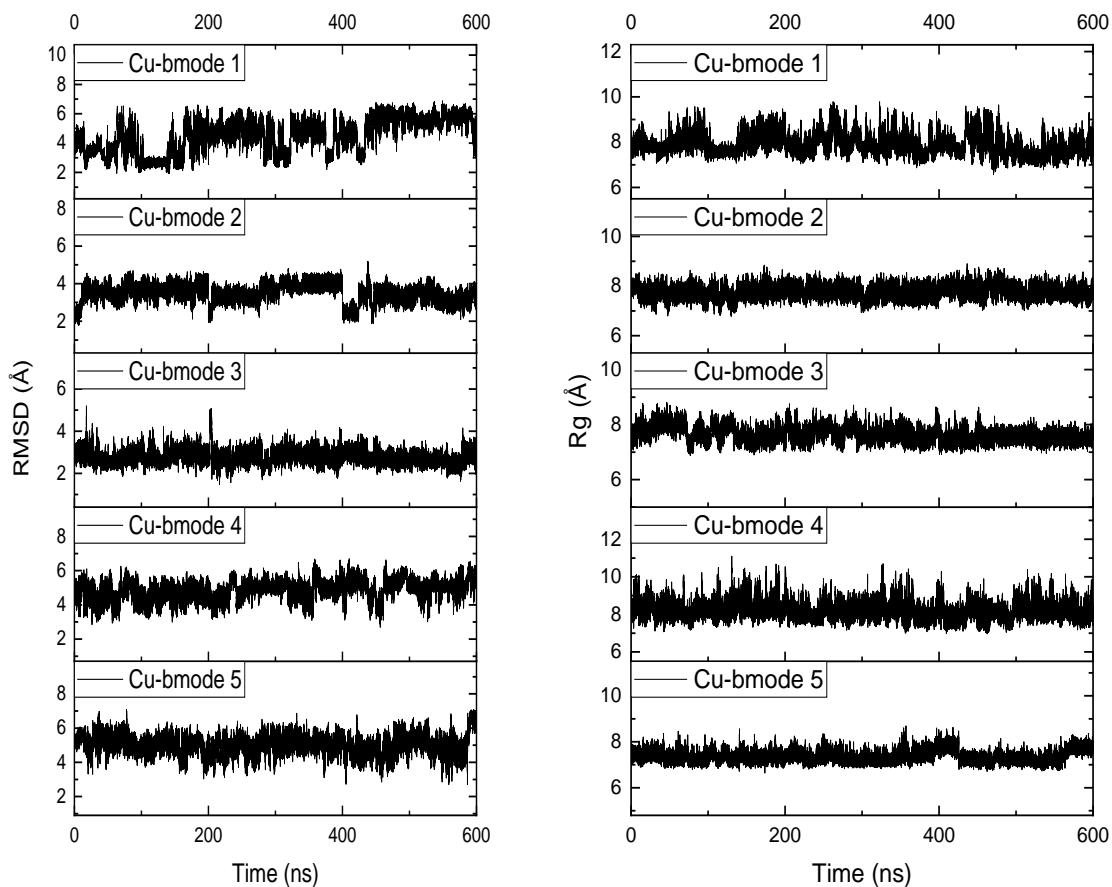


Figure 4.38: Backbone RMSD and R_g values of Cu(II)- Modes and free A β -16. 600 ns of aMD data is reported, made up of the final 200 ns from each of three independent simulations each with different initial velocities.

Table 4.16: Backbone RMSD and R_g of free A β -16 and Cu(II) binding modes, (\AA).

	RMSD				R_g			
	Mean	sd	Min	Max	Mean	sd	Min	Max
Mode 1	4.51	1.14	1.93	6.87	7.89	0.46	6.59	9.79
Mode 2	3.51	0.49	1.79	5.18	7.73	0.25	6.79	8.89
Mode 3	2.79	0.34	1.48	5.19	7.63	0.28	6.88	8.80
Mode 4	4.83	0.57	2.69	6.69	8.18	0.44	6.97	11.09
Mode 5	3.27	0.31	2.00	4.68	7.36	0.26	6.65	8.69
Aβ16	10.64	1.57	4.45	15.46	8.23	0.77	6.80	13.24

The RMSF measurements of residue's stability and their position changes during the simulations are displayed graphically in Figure 4.39. This shows the binding modes 1, 2, 3 and 5 have the lowest RMSF values on average. This indicates these modes are less flexible, and their residues show less deviation from the original starting structures including the residues involved in the metal coordination. However, the C-terminus amidated cap residue and metal ion of mode 1 shows highest values (*ca.*6 \AA) which are comparable to the free peptide. This reflects the changes in the residue's proximity to their initial positions, because Mode 1 and 4 are bound to His13 while Mode 3 is bound to His14, low RMSF values are observed for these residues. However, mode 4 shows a small increase in the RMSF of His13 compared to Mode 1. Overall, Mode 4 presents less movement in residues closer to the N-terminus. However, those residues approaching the C-terminus show increased mobility from their original positions.

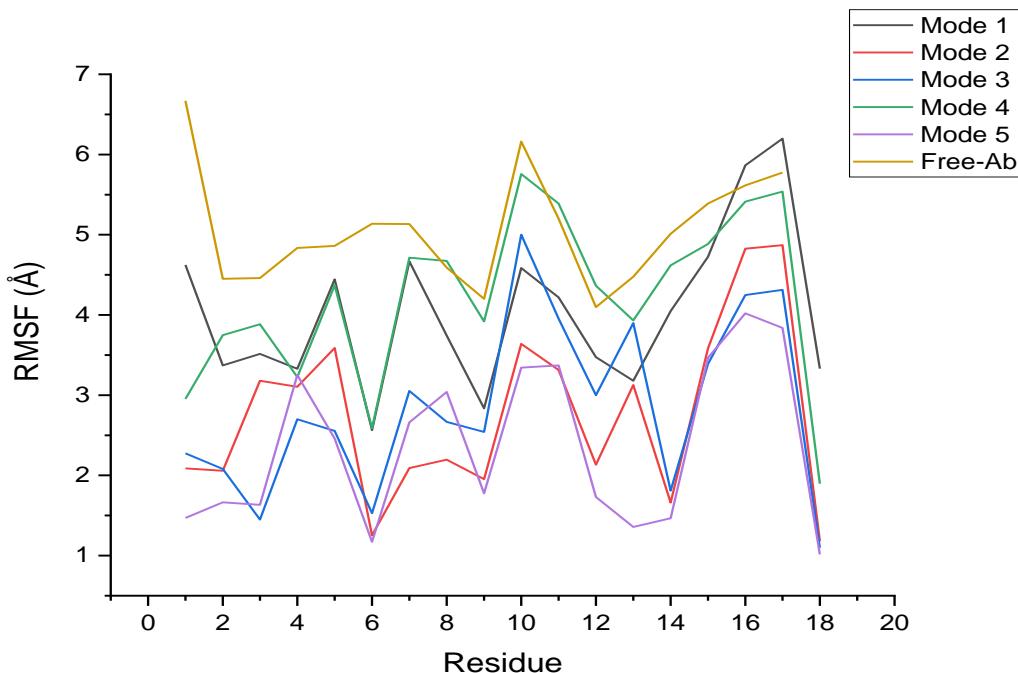


Figure 4.39: RMSF of each residue of free and Cu(II)-bound A β -16. The metal ion itself is denoted residue 18, and the C-terminal amidated cap residue 17.

Clustering analysis of equilibrated trajectories on the basis of backbone dihedrals was used to extract key structural motifs: ribbon diagrams of the most representative clusters are shown in Figure 4.40 and Table 4.17. The highest number of clusters are found in binding mode 1 with 40 notable clusters of structures. The most populated cluster among them occupying only 5% of the entire combined trajectory for this binding mode. However, the most populated cluster was present in Mode 5 at 53%. For this binding mode there were only 6 other clusters.

Table 4.17: Cluster analysis data for equilibrated trajectories.

Clusters	# clusters	Most populated	2 nd populated
Mode 1	40	5%	4%
Mode 2	23	35%	10%
Mode 3	22	19%	14%
Mode 4	25	10%	4%
Mode 5	7	53%	15%

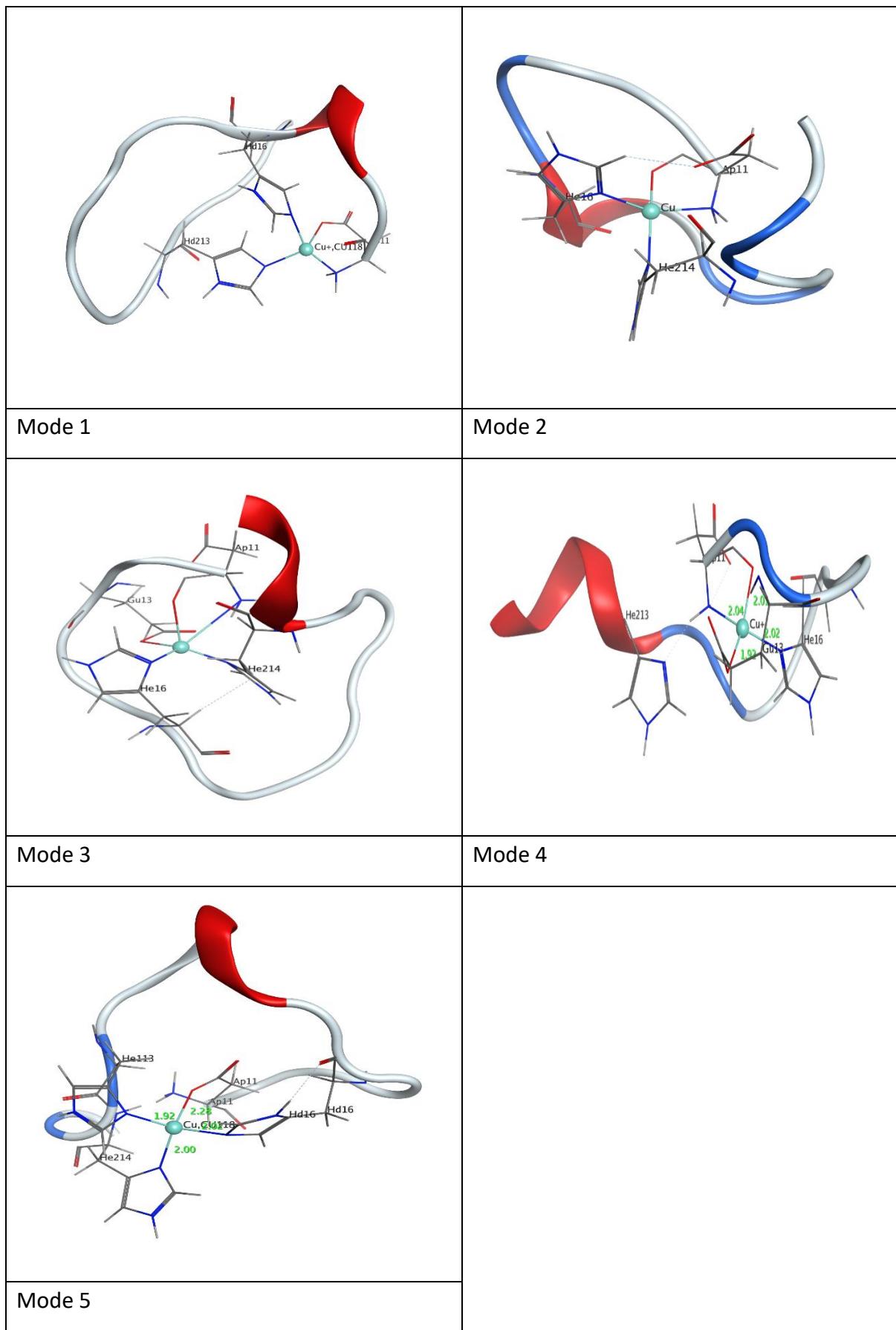
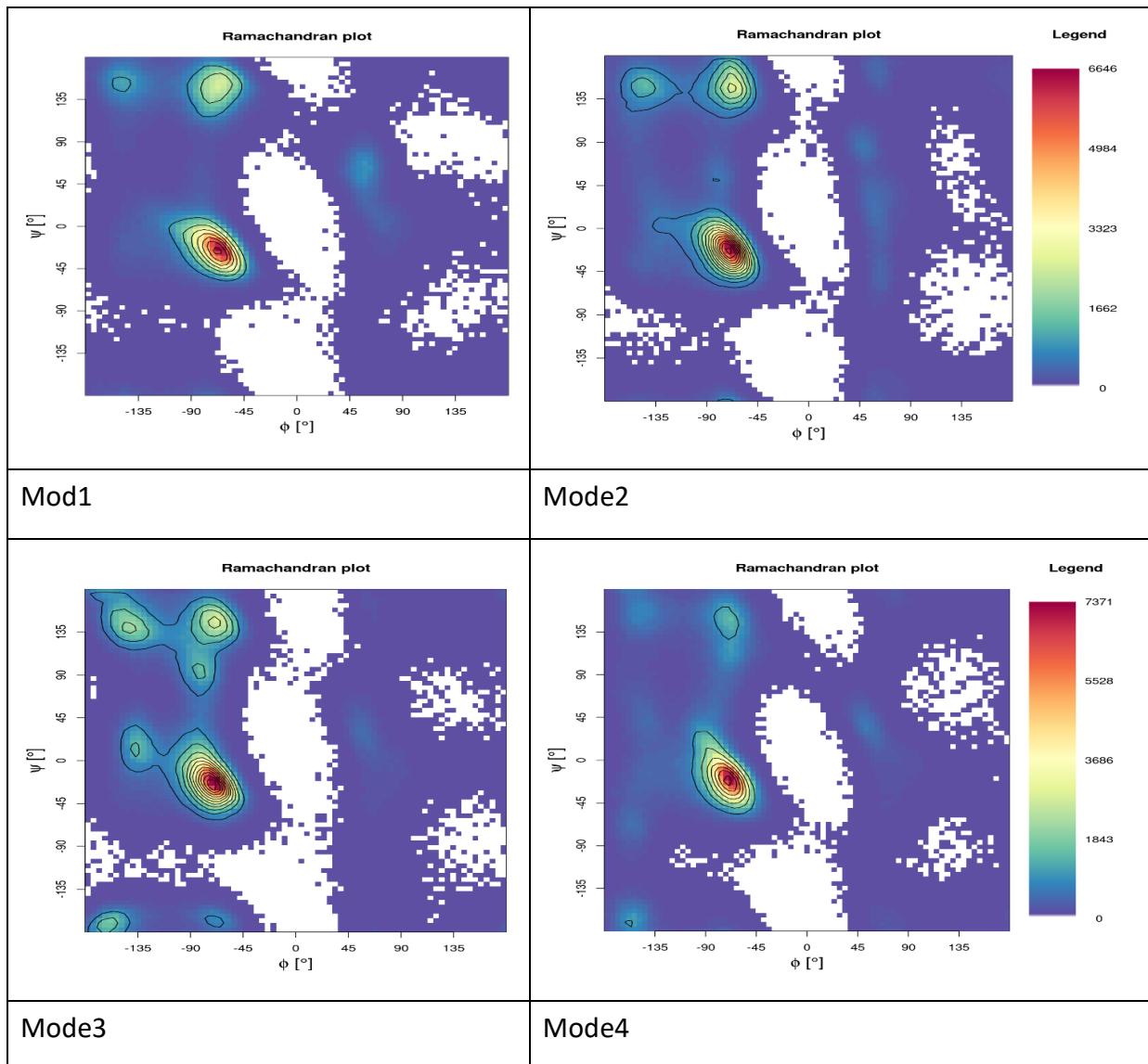


Figure 4.40: Ribbon diagram of the most populated clusters for Cu(II)-A β -16.

Ramachandran maps were used as a further analysis of secondary structure for all five bonding modes according to their backbone dihedral angles (shown in Figure 4.41). In general, the conformations of all binding modes were concentrated in the region corresponding to right-handed helices (ϕ , ψ : ca. -60° , -45°) and also for β -sheets as well (ϕ , ψ : ca. -120° , 135°) with negligible difference observed between these plots. The only other notable detail of these maps is mode 3 presented a slightly higher population density in β -sheet region compared to the all the others.



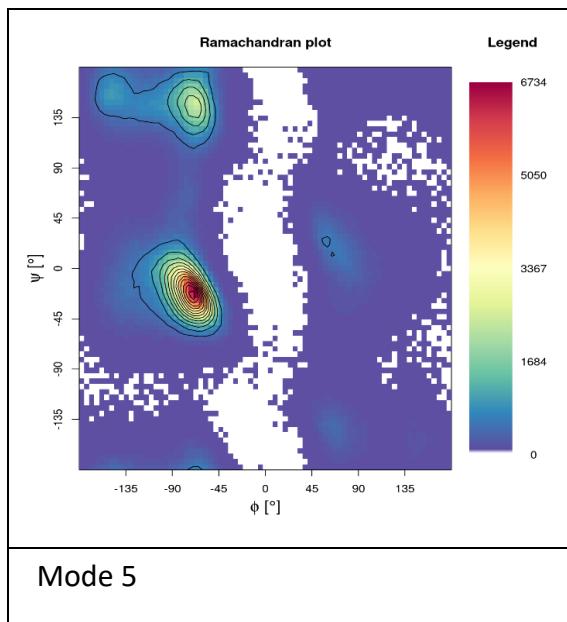
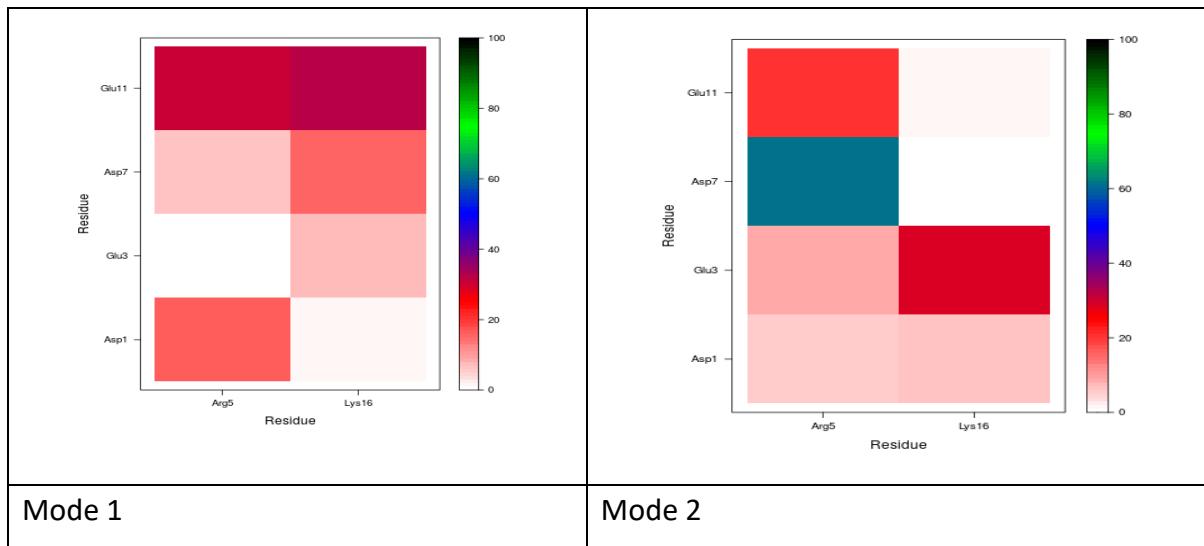


Figure 4.41: Ramachandran maps for Cu(II)-A β . The legends represent the number of dihedrals per bin.

Figure 4.42 presents the percentage of salt bridge between residues within the peptide over the simulation. The most notable interaction was between Asp7 and Arg5 in most of the binding modes by about 50-60 % except mode 1 which has a low interaction percentage (ca.10%). Also, binding modes 1 and 2 have the most amount of interactions between charged residues but the percentages are low (less than 40%). There is little or no interaction between the N- and C-termini and also between Asp7-Lys16 and Glu3-Arg5. Even though the literature⁶² reported Glu3-Arg5 and Asp7-Arg5 interactions in A β , the Cu(II) binding reduces the contact between these residues but did not affect the Asp7-Arg5 interaction.



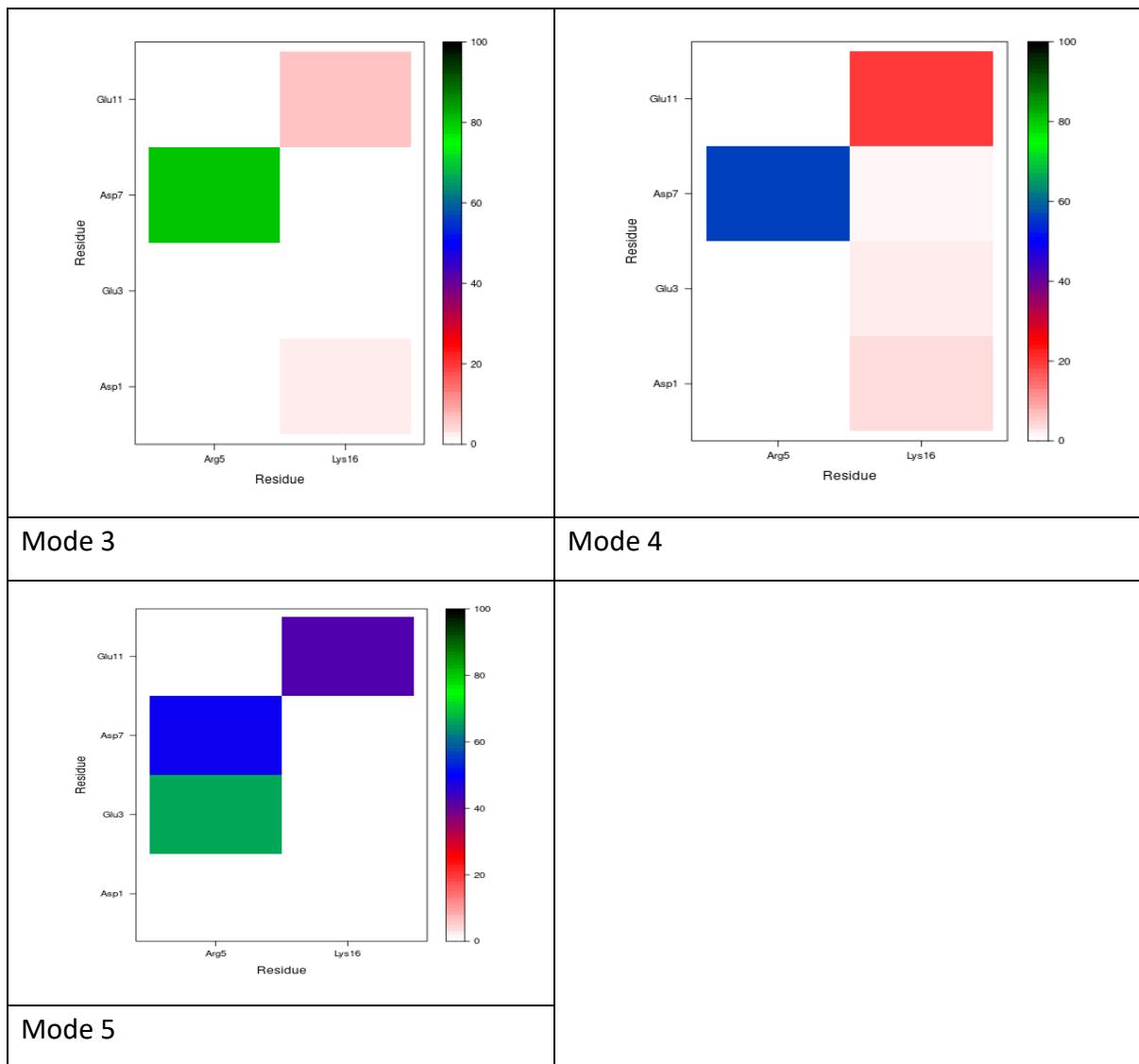


Figure 4.42: Salt bridge maps of Cu(II)-A β 16(%).

The percentage of secondary structures for Cu(II)-A β of the binding modes examined show binding mode 1 and free peptide have similar amount of structural type and ratio of secondary structure (*ca.* 27% helix, 72 % of bend, turn and coil). Mode 1 however, shows slightly more β structure upon coordination to the metal centre compared to the free A β peptide. However, the copper also has other effects on other binding modes of the peptide structure. The Cu(II) ion reduced the helical content of all other modes and increased the bend, turn, and coil formation of the chains as shown in Table 4.18 and Figure 4.43. Binding mode 5 shows more parallel β -sheet more than the other binding modes. The residues that form β -sheet in binding mode 5 are Ser8, Glu11, His13 and Lys.

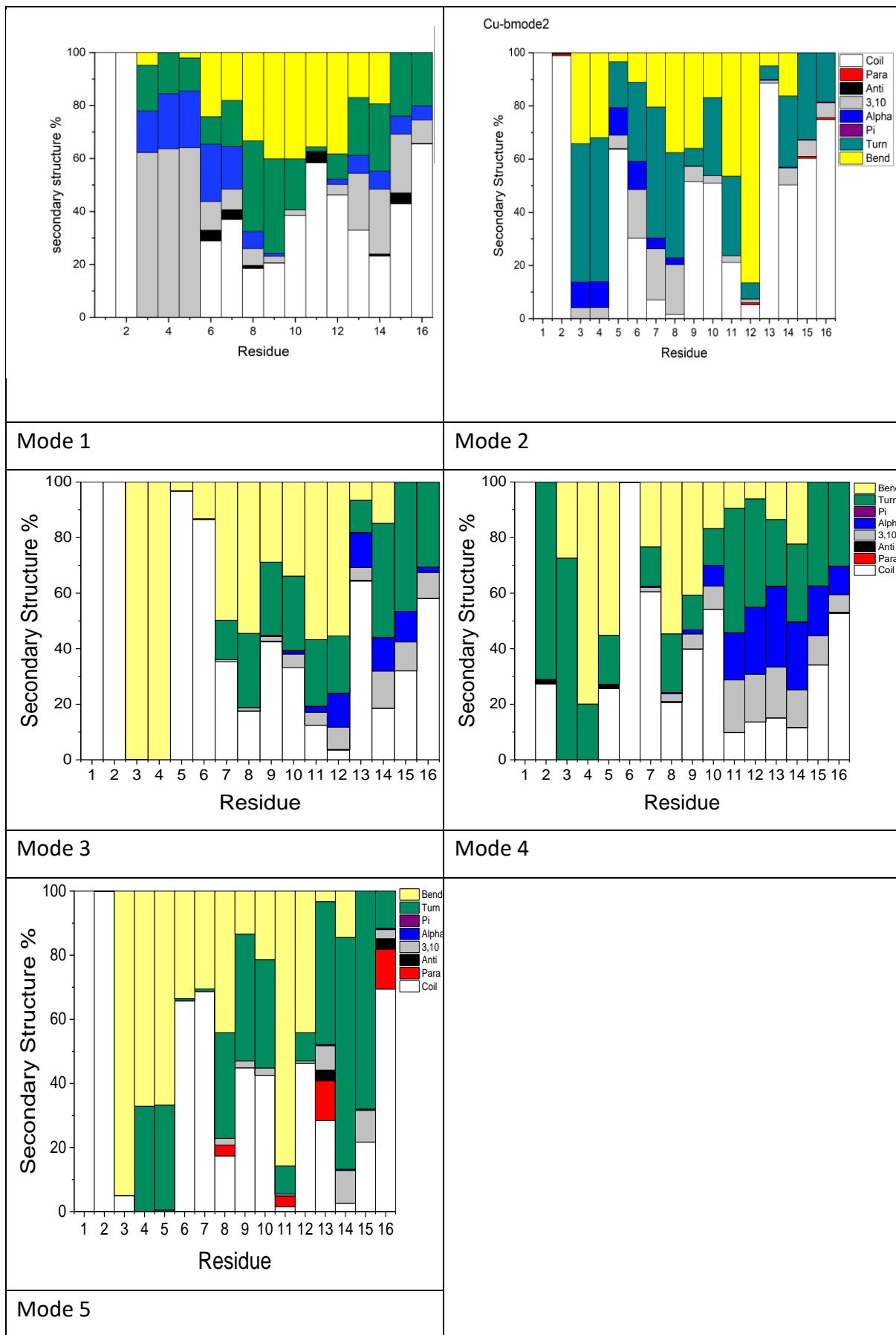
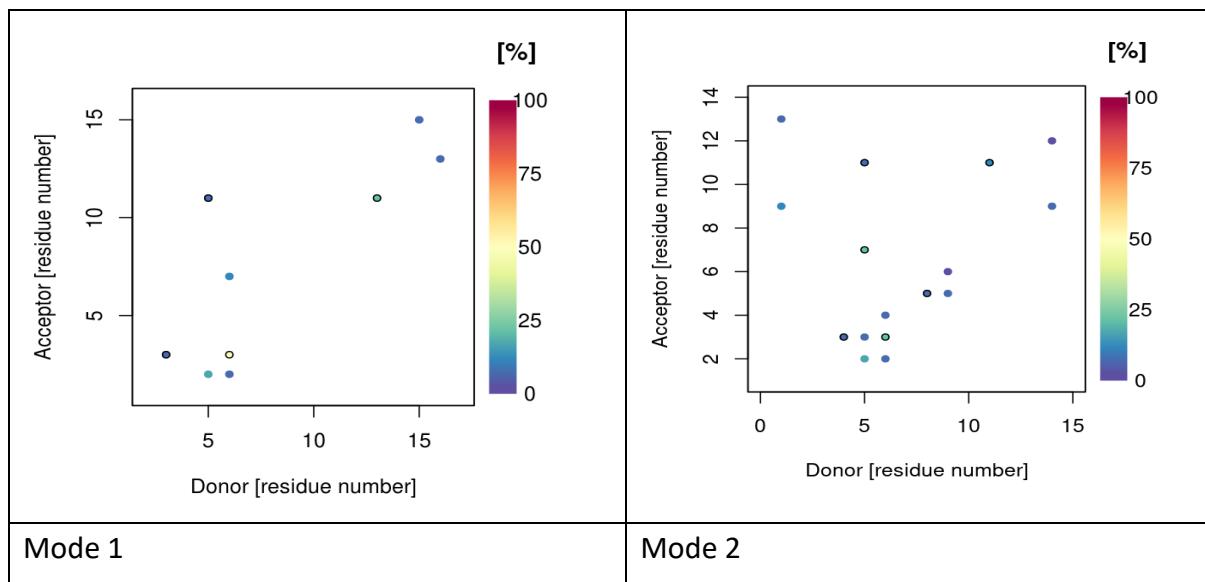


Figure 4.43: Secondary structure for Cu(II)-A β .

Table 4.18: Total secondary structure percentages for each system (%).

	Helix	Sheet	Other
Mode 1	26.94	1.16	71.90
Mode 2	9.36	0.23	90.41
Mode 3	7.05	0.08	92.87
Mode 4	14.72	0.27	85.01
Mode 5	2.22	2.14	95.64
Aβ-16	27.2	0.4	72.4

Hydrogen bonds have an important role of structural stabilization and changes in structure can be reflected here. There are plotted for all simulations of Cu(II) binding modes as shown in Figure 4.44 and Table 4.19. There are small differences in the number of HB found between each of the modes. Binding modes 1 and 2 have a maximum number of 13 HB in some frames. Binding Mode 3 and 4 show a slightly lower maximum HB number of 11. Table 4.19 also shows there were some frames over the course of all simulations in which no hydrogen bonds were present except binding mode 5 which contains minimum 2 bonds over the simulation. For further investigation, the interaction percentage between donor and acceptor residues was measured. The result shows HB that exist in 100% of the frames in mode 5 between Asp1 and His14. Another high HB ratio is found between Glu3 (H Donor) and Glu3 (O Acceptor) of mode 4 by 55%.



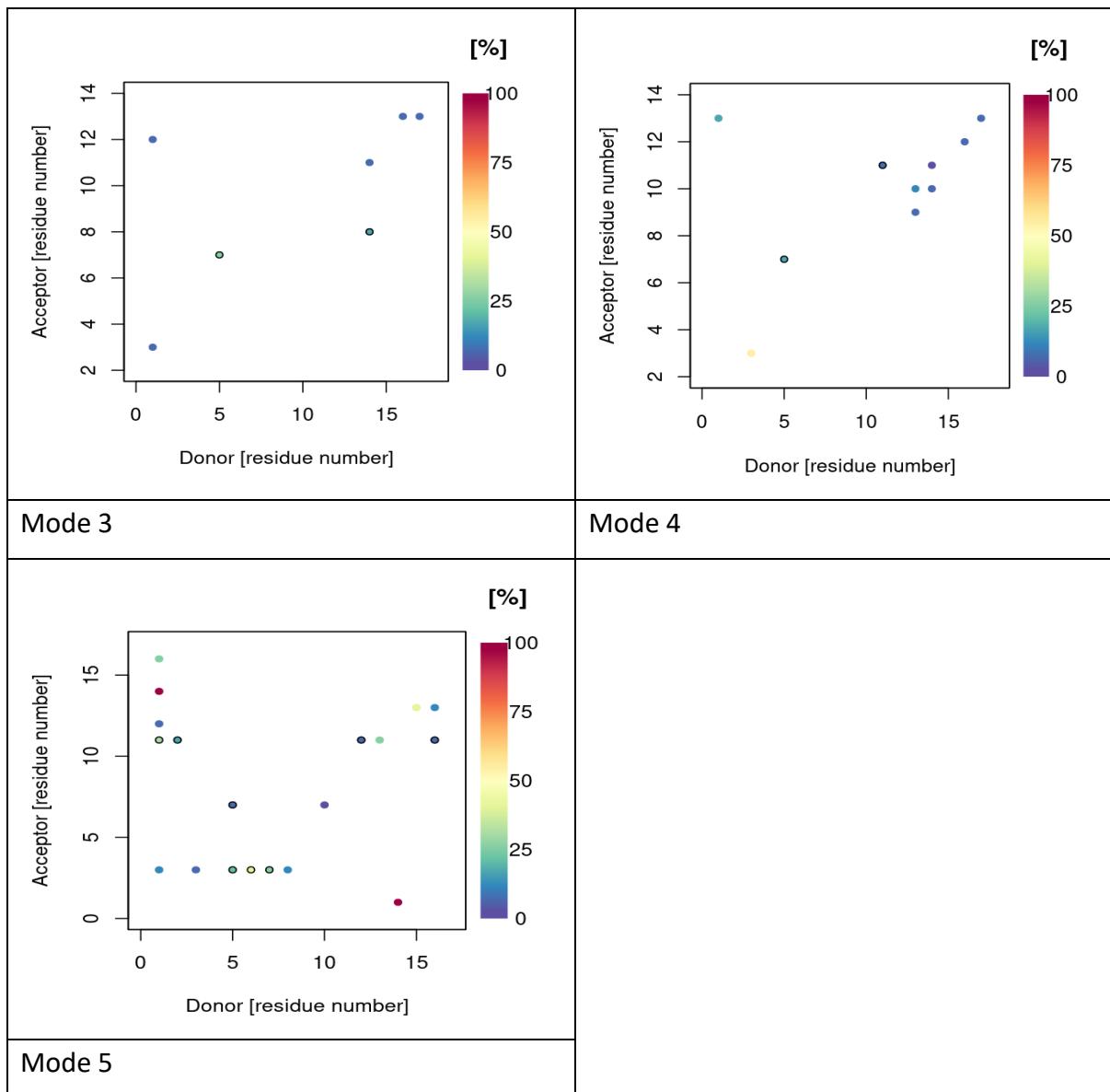
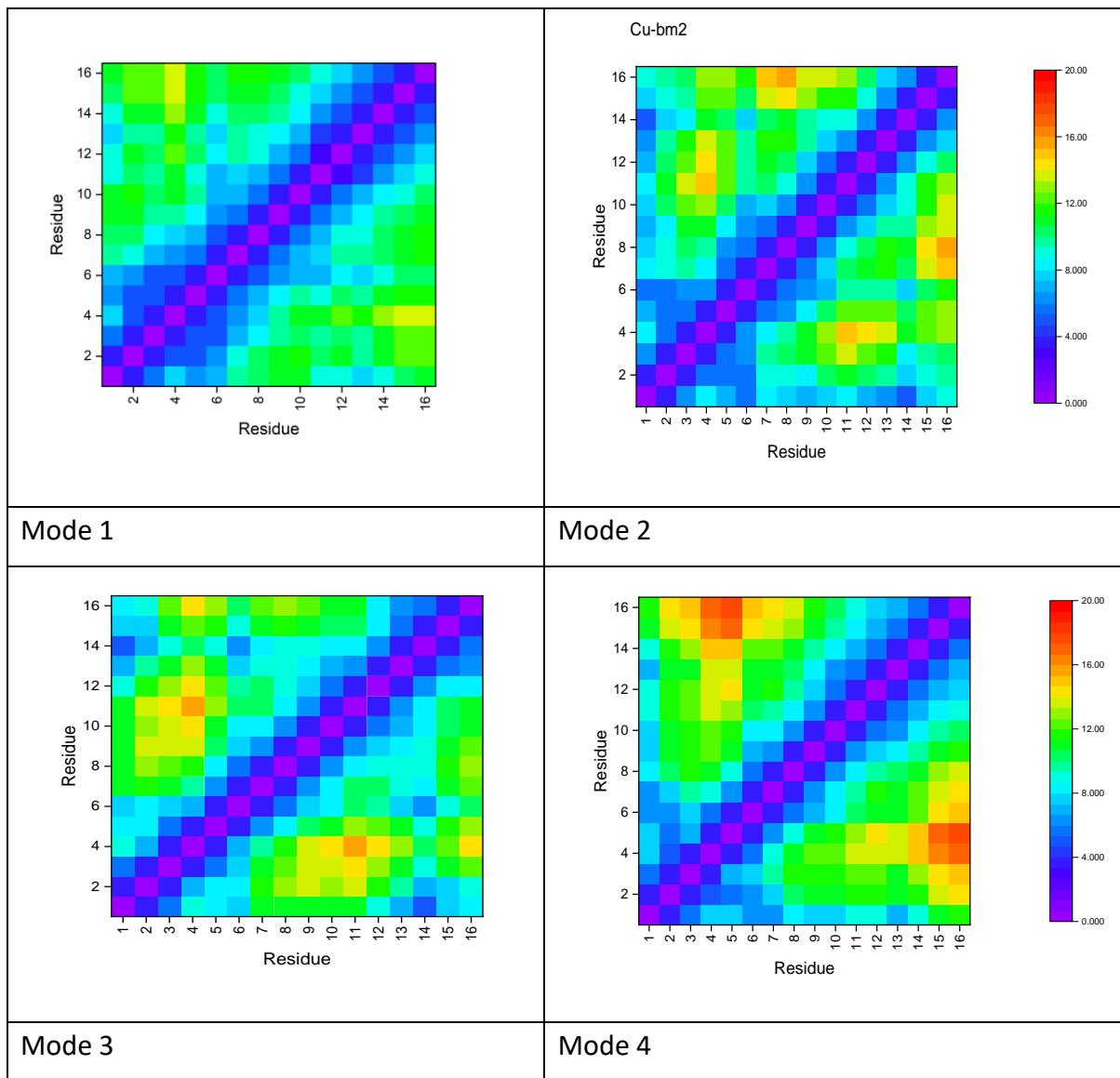


Figure 4.44: Hydrogen bonding occupancy Cu(II)-A β 16. Plots only display data for hydrogen bonds present for greater than 5% of simulation. Black circles indicate the number of interaction between the relevant residues.

Table 4.19: HB count for Cu(II)-A β .

Number of HB	Mean	sd	Max	Min
Mode 1	4.78	1.80	13	0.0
Mode 2	4.46	1.78	13	0.0
Mode 3	3.05	1.53	11	0.0
Mode 4	3.40	1.62	11	0.0
Mode 5	8.70	1.98	17	2.0

The distance between residues within the Cu(II)-A β 16 peptide were accounted for in contact maps presented through Figure 4.45. The largest distance was observed in Mode 4 between the C-terminal residue, Lys16, and both Phe4 and Arg5. There are close contacts between His14 and Asp1 seen in Mode 2 and 3 but additionally large distances recorded between Glu11 and Phe4 in the same modes. There were no notably large distances between residues in mode 1. Mode 5 also shows less overall distance between residues, and the largest distance found at this mode is between Arg5 and Gln15.



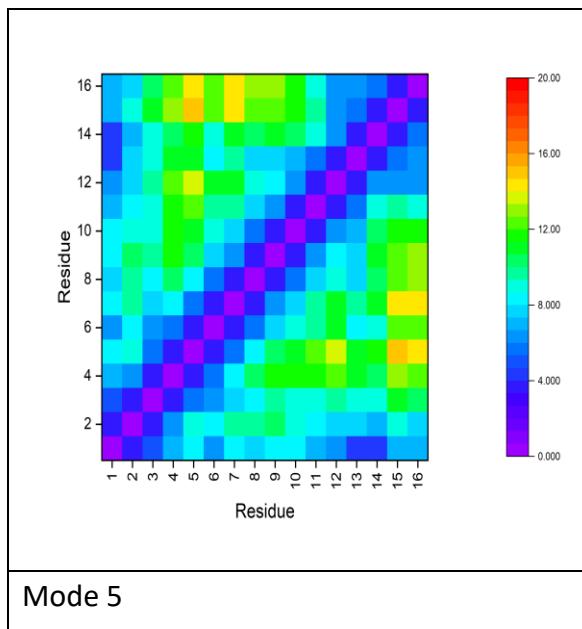


Figure 4.45: Contact map of binding modes of Cu(II)-A β . The legend shows the average inter-residue distance in \AA .

1D Landscapes of free energy surface as functions of end-to end distances (Figure 4.46) and R_g (Figure 4.47) show modes 1 and 4 have a greater range of end-to end values that low energy conformers (*ca.* 5 $\text{kcal} \cdot \text{mol}^{-1}$) can access. Modes 2 and 3 possess a smaller range that exists between 4 and 14 \AA . However, the smallest end-to end distance range is found in Mode 5 a range of 6-14 \AA . The 1D plots of free energy as a function of R_g show that the most stable conformers were seen at ranges between 7-7.5 \AA of all binding modes. Free energy plots as a function of hydrogen bond formation are displayed in Figure 4.48. The FES landscapes show the most stable conformers contained between 4-6 hydrogen bonds except Mode 5 which exhibited between 8-10 hydrogen bonds. For mode 3, the most stable basin is a region where the conformers have 3 hydrogen bonds.

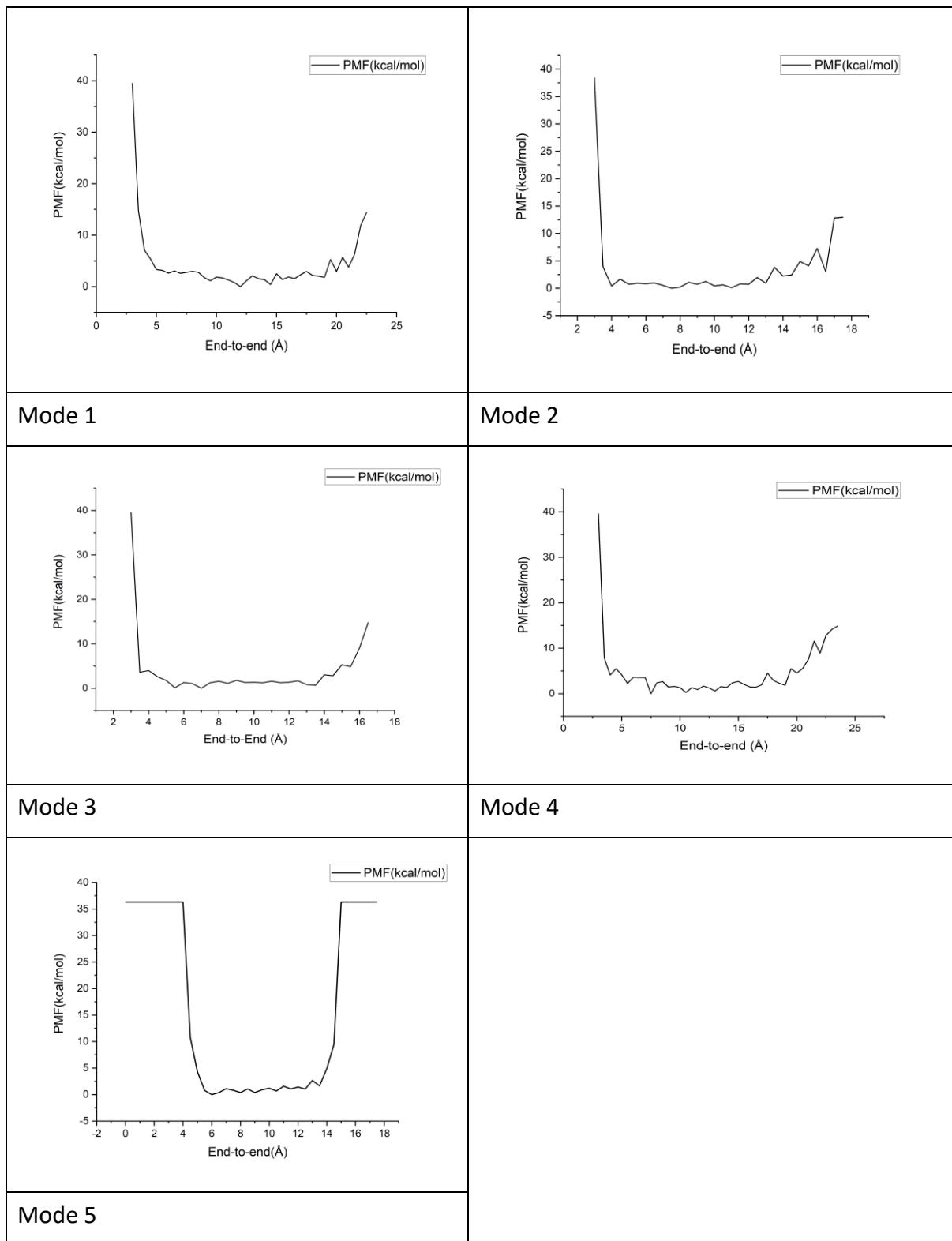


Figure 4.46: 1D-Free energy surface of Cu(II)-A β as a function of end-to-end distance (\AA).

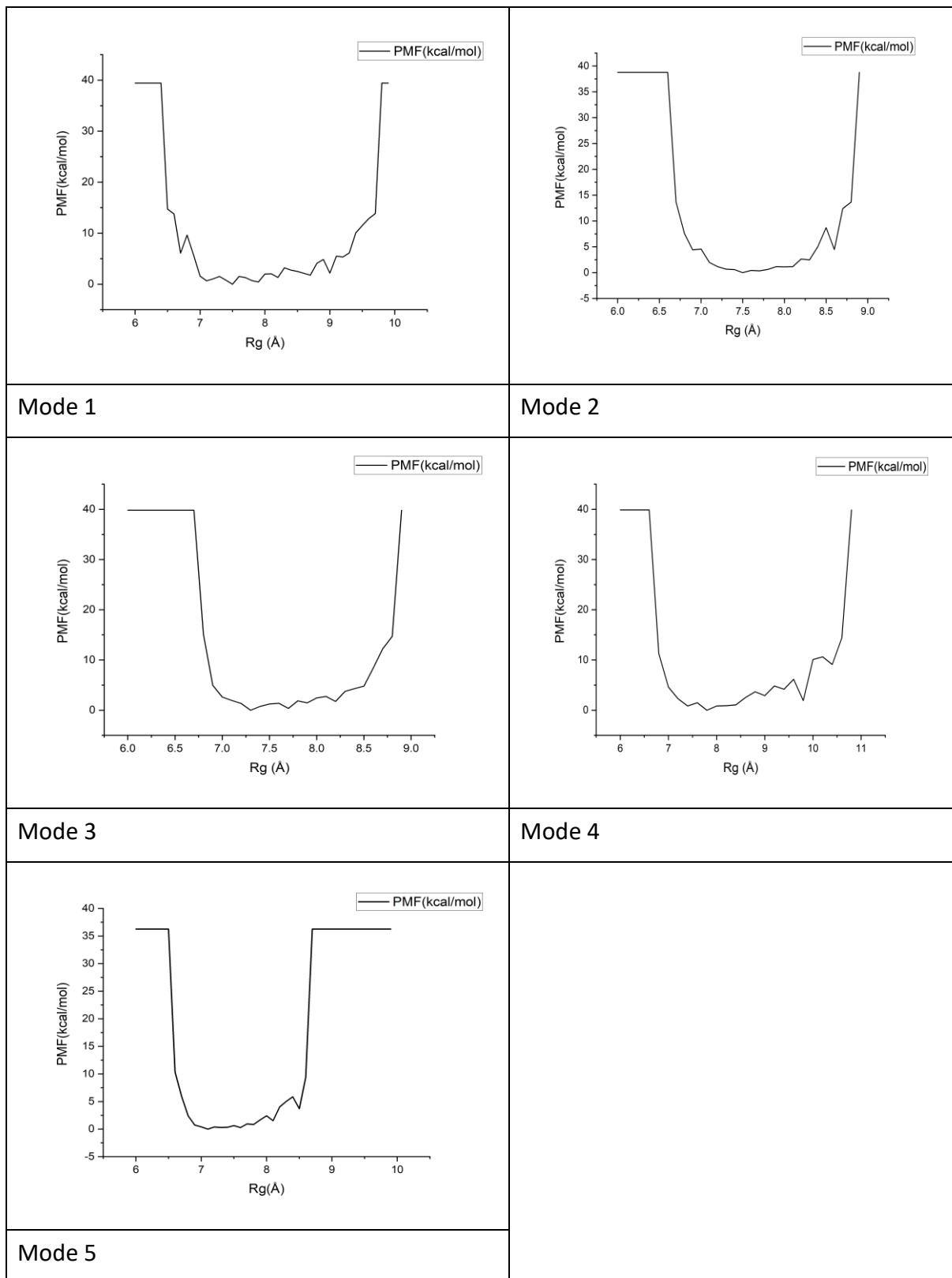


Figure 4.47: 1D-Free energy surface of Cu(II)-A β as a function of R_g(\AA).

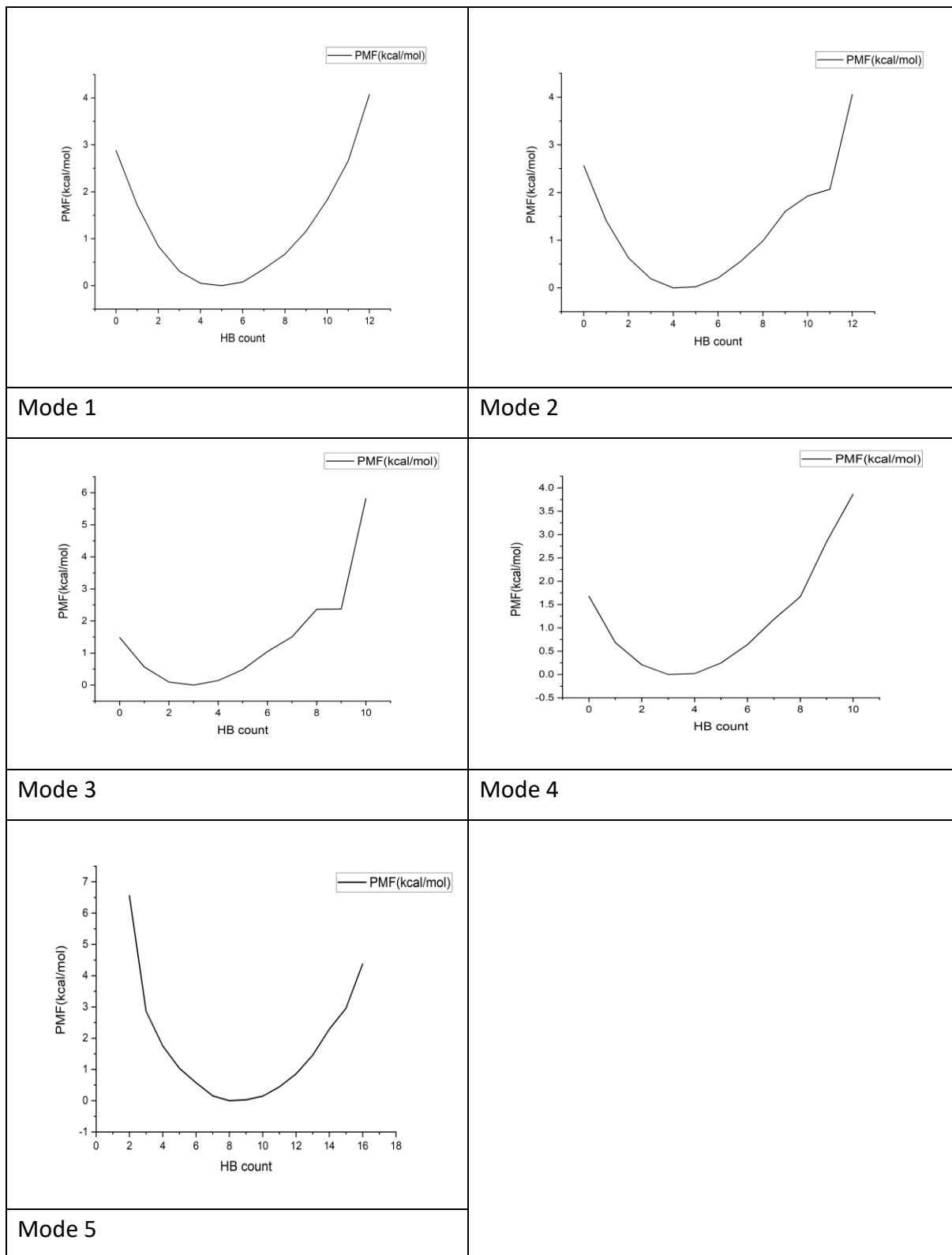


Figure 4.48: 1D-Free energy surface of Cu(II)-A β as a function of HB count.

4.6.3 Conclusion

Accelerated MD simulations of five binding modes examined for Cu(II) bound to A β 16 were performed. The comparison shows Modes 2, 3 and 5 have the lowest structural changes and flexibility in their residues. However, the remaining modes show high flexibility and increased size. This explains the high flexibility of these two modes as their R_g values are close to that of the metal-free peptide, and also indicate the reduced stability of the complex with the Cu(II) ion compared to the complexes with Fe and Zn. Also, the most persistent salt bridge found in Cu(II)-A β is like to that seen in Fe(II) and Zn(II) simulation which formed between Arg5-Asp7. Another high salt bridge percentage observed in one of the Cu(II) binding mode (Mode 5) is Glu3-Arg5 salt bridge. Both of these interactions recorded in A β by NMR according to the literature.

4.7 General Conclusion

Transition metals (such as copper, zinc, and iron) have been typically implicated in Alzheimer's disease. Their binding to A β could signal changes to the structure and folding of the peptide, resulting in promotion of A β aggregation. The binding site locations have been identified in the N-terminal region with different reported residues/atoms. In each of these binding sites, ions are highly influential in the structural transitions encountered in the peptide. This, together with the type of metals makes their presence crucial in the aetiology of the disease, and should therefore not be disregarded, due to their effect on the structure and related properties. Thus, a comparative study of different binding modes for the metal ions via theoretical research can provide valuable insight of their effects. In the present chapter, the structural differences of A β peptide that occur upon different binding modes for each of Zn(II), Fe(II), and Cu(II), have been investigated via accelerated sampling methods, so-called aMD.

The overall comparison via aMD simulations, show the metal coordination restricts the peptide's movement, indicated by RMSF values of the coordinated residues, and compared to the unbound peptide. The lowest RMSD values found amongst the coordination modes, of all three metals, correspond to the strong contact between residues, found in around 100% of the frames, which restricts the conformational changes of the peptide. The most frequent salt bridge interaction has been seen in all three metal's modes, between residues Arg5 and Asp7.

The experimental studies have mentioned this type of salt bridge between Arg5 and Asp7, and proposed that the Arg5-based contacts may be a factor in the increased A β aggregates as found in fibrils. The aMD simulations also show that the number of coordination and the atom/residue type can change the overall structure, size, and stability.

In addition, the comparison between amidated C-terminus A β -peptide, which was tested in one of the zinc binding modes, shows more conformational fluctuation during the simulation. The residues within this system show a higher level of mobility compared to the one that had not been amidated, which is expected as the end of the peptide is more free to move. Fe(II) presented the lowest RMSD values as a result of being coordinated to a higher number of atoms (five or six), indicating the number of coordination limits the flexibility of the peptide. These low RMSD values had also been seen in some Cu(II) modes, when five binding modes are bound.

4.8 References

- 1 K. I. Silva et al., ESEEM Analysis of Multi-Histidine Cu(II)-Coordination in Model Complexes, Peptides, and Amyloid- β , *J. Phys. Chem. B*, 2014, **118**, 8935–8944.
- 2 A. I. Bush, The metallobiology of Alzheimer's disease, *Trends Neurosci.*, 2003, **26**, 207–214.
- 3 M. A. Lovell et al., Copper, iron and zinc in Alzheimer's disease senile plaques, *J. Neurol. Sci.*, 1998, **158**, 47–52.
- 4 C. C. Curtain et al., Metal ions, pH, and cholesterol regulate the interactions of Alzheimer's disease amyloid- β peptide with membrane lipid, *J. Biol. Chem.*, 2003, **278**, 2977–2982.
- 5 L. Guilloréau et al., Structural and thermodynamical properties of CuII amyloid- β 16/28 complexes associated with Alzheimer's disease, *J. Biol. Inorg. Chem.*, 2006, **11**, 1024–1038.
- 6 J. W. Karr et al., N-terminal deletions modify the Cu²⁺ binding site in amyloid- β , *Biochemistry*, 2005, **44**, 5478–5487.
- 7 T. Kowalik-Jankowska et al., Coordination abilities of the 1-16 and 1-28 fragments of β -amyloid peptide towards copper(II) ions: A combined potentiometric and spectroscopic study, *J. Inorg. Biochem.*, 2003, **95**, 270–282.
- 8 S. Bolognin et al., Aluminum, copper, iron and zinc differentially alter amyloid-A β 1-42 aggregation and toxicity, *Int. J. Biochem. Cell Biol.*, 2011, **43**, 877–885.
- 9 T. Kowalik-Jankowska et al., Cu(II) interaction with N-terminal fragments of human and mouse β -amyloid peptide, *J. Inorg. Biochem.*, 2001, **86**, 535–545.
- 10 S. Zirah et al., Structural changes of region 1-16 of the Alzheimer disease amyloid β -peptide upon zinc binding and in vitro aging, *J. Biol. Chem.*, 2006, **281**, 2151–2161.
- 11 D. Hamelberg et al., Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules, *J. Chem. Phys.*, 2004, **120**, 11919–11929.

- 12 D. Hamelberg et al., Sampling of slow diffusive conformational transitions with accelerated molecular dynamics, *J. Chem. Phys.*, 2007, **127**, 155102.
- 13 D. Bucher et al., On the Use of Accelerated Molecular Dynamics to Enhance Configurational Sampling in Ab Initio Simulations, *J. Chem. Theory Comput.*, 2011, **7**, 890–897.
- 14 P. R. L. Markwick et al., Studying functional dynamics in bio-molecules using accelerated molecular dynamics, *Phys. Chem. Chem. Phys.*, 2011, **13**, 20053–20065.
- 15 Y. Miao et al., Accelerated molecular dynamics simulations of protein folding, *J. Comput. Chem.*, 2015, **36**, 1536–1549.
- 16 J. Wereszczynski et al., ed. R. Baron, Springer New York, 2012, pp. 515–524.
- 17 Y. Wang et al., Accelerated Molecular Dynamics: Theory, Implementation and Applications, 2012, 165–172.
- 18 Y. Wang et al., Implementation of accelerated molecular dynamics in NAMD, *Comput. Sci. Discov.*, 2011, **4**, 15002.
- 19 D. A. Case et al., AMBER 2016, San Francisco, 2016.
- 20 P. Li et al., MCPB.py: A Python Based Metal Center Parameter Builder, *J. Chem. Inf. Model.*, 2016, **56**, 599–604.
- 21 A. P. Scott et al., Harmonic vibrational frequencies: An evaluation of Hartree-Fock, Møller-Plesset, quadratic configuration interaction, density functional theory, and semiempirical scale factors, *J. Phys. Chem.*, 1996, **100**, 16502–16513.
- 22 M. J. Frisch et al., GAUSSIAN09, <http://www.gaussian.com/>.
- 23 J. M. Seminario, Calculation of intramolecular force fields from second-derivative tensors, *Int. J. Quantum Chem.*, 1996, **60**, 1271–1277.
- 24 C. I. Bayly et al., A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model, *J. Phys. Chem.*, 1993, **97**, 10269–10280.
- 25 B. H. Besler et al., Atomic charges derived from semiempirical methods, *J. Comput.*

- Chem.*, 1990, **11**, 431–439.
- 26 P. Cieplak et al., Application of the multimolecule and multiconformational RESP methodology to biopolymers: Charge derivation for DNA, RNA, and proteins, *J. Comput. Chem.*, 1995, **16**, 1357–1377.
- 27 D. A. Case et al., *AMBER 2019*, San Francisco, 2019.
- 28 G. D. Hawkins et al., Pairwise solute descreening of solute charges from a dielectric medium, *Chem. Phys. Lett.*, 1995, **246**, 122–129.
- 29 G. D. Hawkins et al., Parametrized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium, *J. Phys. Chem.*, 1996, **100**, 19824–19839.
- 30 H. Nguyen et al., Improved generalized born solvent model parameters for protein simulations, *J. Chem. Theory Comput.*, 2013, **9**, 2020–2034.
- 31 J. A. Maier et al., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB, *J. Chem. Theory Comput.*, 2015, **11**, 3696–3713.
- 32 J.-P. Ryckaert et al., Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *J. Comput. Phys.*, 1977, **23**, 327–341.
- 33 W. Clark Still et al., Semianalytical treatment of solvation for molecular mechanics and dynamics, *J. Am. Chem. Soc.*, 2002, **112**, 6127–6129.
- 34 R. Constanciel et al., Self consistent field theory of solvent effects representation by continuum models: Introduction of desolvation contribution, *Theor. Chim. Acta*, 1984, **65**, 1–11.
- 35 M. Schaefer et al., A Comprehensive Analytical Treatment of Continuum Electrostatics, *J. Phys. Chem.*, 1996, **100**, 1578–1599.
- 36 D. R. Roe et al., PTraj and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data, *J. Chem. Theory Comput.*, 2013, **9**, 3084–3095.
- 37 R. J. Deeth et al., DommiMOE: An implementation of ligand field molecular mechanics

- in the molecular operating environment, *J. Comput. Chem.*, 2005, **26**, 123–130.
- 38 J. Nasica-Labouze et al., Amyloid β Protein and Alzheimer's Disease: When Computer Simulations Complement Experimental Studies, *Chem. Rev.*, 2015, **115**, 3518–3563.
- 39 R. J. Deeth, The ligand field molecular mechanics model and the stereoelectronic effects of d and s electrons, *Coord. Chem. Rev.*, 2001, **212**, 11–34.
- 40 E. J. Meijer et al., A density-functional study of the intermolecular interactions of benzene, *J. Chem. Phys.*, 1996, **105**, 8684–8689.
- 41 R. Rajendran et al., A novel approach to the identification and quantitative elemental analysis of amyloid deposits-Insights into the pathology of Alzheimer's disease, *Biochem. Biophys. Res. Commun.*, 2009, **382**, 91–95.
- 42 K. Pauwels et al., Structural basis for increased toxicity of pathological A β 42:A β 40 ratios in Alzheimer disease, *J. Biol. Chem.*, 2012, **287**, 5650–5660.
- 43 S. Zirah et al., Zinc binding properties of the amyloid fragment A β (1–16) studied by electrospray-ionization mass spectrometry, *Int. J. Mass Spectrom.*, 2003, **228**, 999–1016.
- 44 S. A. Kozin et al., Zinc Binding to Alzheimer's A β (1–16) Peptide Results in Stable Soluble Complex, *Biochem. Biophys. Res. Commun.*, 2001, **285**, 959–964.
- 45 M. Rana et al., Cu and Zn interactions with A β peptides: consequence of coordination on aggregation and formation of neurotoxic soluble A β oligomers, *Metalomics*, 2019, **11**, 64–84.
- 46 A. I. Bush et al., Modulation of A beta adhesiveness and secretase site cleavage by zinc., *J. Biol. Chem.*, 1994, **269**, 12152–12158.
- 47 W. P. Esler et al., Zinc-Induced Aggregation of Human and Rat β -Amyloid Peptides In Vitro, *J. Neurochem.*, 1996, **66**, 723–732.
- 48 C. S. Atwood et al., Dramatic aggregation of Alzheimer abeta by Cu(II) is induced by conditions representing physiological acidosis, *J. Biol. Chem.*, 1998, **273**, 12817–12826.
- 49 T. Miura et al., Metal Binding Modes of Alzheimer's Amyloid β -Peptide in Insoluble

- Aggregates and Soluble Complexes, *Biochemistry*, 2000, **39**, 7024–7031.
- 50 E. Gaggelli et al., NMR Studies of the Zn²⁺ Interactions with Rat and Human β-Amyloid (1–28) Peptides in Water-Micelle Environment, *J. Phys. Chem. B*, 2008, **112**, 100–109.
- 51 S. Furlan et al., Modeling of the Zn²⁺ binding in the 1–16 region of the amyloid β peptide involved in Alzheimer’s disease, *Phys. Chem. Chem. Phys.*, 2009, **11**, 6468–6481.
- 52 B. Alies et al., Zinc(II) Binding Site to the Amyloid-β Peptide: Insights from Spectroscopic Studies with a Wide Series of Modified Peptides, *Inorg. Chem.*, 2016, **55**, 10499–10509.
- 53 M. Turner et al., Molecular dynamics simulation on the effect of transition metal binding to the N-terminal fragment of amyloid-β, *J. Biomol. Struct. Dyn.*, 2019, **37**, 4590–4600.
- 54 F. Bousejra-ElGarah et al., Iron(II) Binding to Amyloid-β, the Alzheimer’s Peptide, *Inorg. Chem.*, 2011, **50**, 9024–9030.
- 55 V. A. Streltsov et al., The structure of the amyloid-beta peptide high-affinity copper II binding site in Alzheimer disease, *Biophys. J.*, 2008, **95**, 3447–3456.
- 56 C. M. Kok et al., Relationship between the hydrodynamic radius and the radius of gyration of a polymer in solution, *Die Makromol. Chemie, Rapid Commun.*, 1981, **2**, 655–659.
- 57 J. Danielsson et al., Translational diffusion measured by PFG-NMR on full length and fragments of the Alzheimer Aβ(1–40) peptide. Determination of hydrodynamic radii of random coil peptides of varying length, *Magn. Reson. Chem.*, 2002, **40**, 89–97.
- 58 C. Talmard et al., Amyloid-Beta Peptide Forms Monomeric Complexes With Cu^{II} and Zn^{II} Prior to Aggregation, *ChemBioChem*, 2007, **8**, 163–165.
- 59 R. S. Armen et al., Pauling and Corey’s α-pleated sheet structure may define the prefibrillar amyloidogenic intermediate in amyloid disease, *Proc. Natl. Acad. Sci.*, 2004, **101**, 11622–11627.
- 60 D. Shea et al., α-Sheet secondary structure in amyloid β-peptide drives aggregation and toxicity in Alzheimer’s disease, *Proc. Natl. Acad. Sci.*, 2019, **116**, 8895–8900.

- 61 R. S. Armen et al., Anatomy of an Amyloidogenic Intermediate: Conversion of β -Sheet to α -Sheet Structure in Transthyretin at Acidic pH, *Structure*, 2004, **12**, 1847–1863.
- 62 S. P. B. Vemulapalli et al., Combined High-Pressure and Multiquantum NMR and Molecular Simulation Propose a Role for N-Terminal Salt Bridges in Amyloid-Beta, *J. Phys. Chem. Lett.*, 2021, **12**, 9933–9939.
- 63 M. Kollmer et al., Cryo-EM structure and polymorphism of A β amyloid fibrils purified from Alzheimer's brain tissue, *Nat. Commun.*, 2019, **10**, 4760.
- 64 M. del Barrio et al., in *Biometals in Neurodegenerative Diseases: Mechanisms and Therapeutics*, eds. A. R. White, M. Aschner, L. G. Costa and A. I. Bush, Academic Press, 2017, pp. 265–281.
- 65 C. Hureau, Coordination of redox active metal ions to the amyloid precursor protein and to amyloid- β peptides involved in Alzheimer disease. Part 1: An overview, *Coord. Chem. Rev.*, 2012, **256**, 2164–2174.
- 66 C. D. Syme et al., Copper binding to the amyloid-beta (A β) peptide associated with Alzheimer's disease: folding, coordination geometry, pH dependence, stoichiometry, and affinity of A β -(1-28): insights from a range of complementary spectroscopic techniques., *J. Biol. Chem.*, 2004, **279**, 18169–77.
- 67 C. Hureau et al., Deprotonation of the Asp1-Ala2 Peptide Bond Induces Modification of the Dynamic Copper(II) Environment in the Amyloid- β Peptide near Physiological pH, *Angew. Chemie Int. Ed.*, 2009, **48**, 9522–9525.
- 68 Y. El Khoury et al., New Insights into the Coordination of Cu(II) by the Amyloid-B 16 Peptide from Fourier Transform IR Spectroscopy and Isotopic Labeling, *J. Phys. Chem. B*, 2011, **115**, 14812–14821.
- 69 S. C. Drew et al., Pleomorphic Copper Coordination by Alzheimer's Disease Amyloid- β Peptide, *J. Am. Chem. Soc.*, 2009, **131**, 1195–1207.

5 Alpha sheets structure in A β

5.1 Introduction

Alzheimer's disease (AD) is described by the build-up of β -sheet-rich monomers forming insoluble A β plaques. Even though plaque load is not associated with cognitive loss in AD patients, it is correlated with the presence of toxic soluble oligomers. Recently, theoretical and experimental studies have shown that A β soluble oligomers adopt a nonstandard secondary structure, termed α -sheet.^{1,2} A recent experimental study has proved that this structure in A β might drive aggregation and has been suggested to correlate with formation of toxic oligomers.² The α -sheet secondary structure was first proposed by Pauling and Corey.³ A recent MD study, suggested a mechanism for A β aggregation that occurs when more than one adjacent α -strands interact through hydrogen bonding to form α -sheets.⁴ The α -sheet chain conformation is described by an alternation of residues in the helical, α_R and α_L forms, as represented in Figure 5.1. The peptide arrangement found in α -sheets involves the NH group on one side of the sheet and carbonyl group on the other, with hydrogen bonding (HB) forming between adjacent strands in the α -extended chain, contrasting the way α -helices and β -strands form, through repeating (ϕ , ψ) angles.

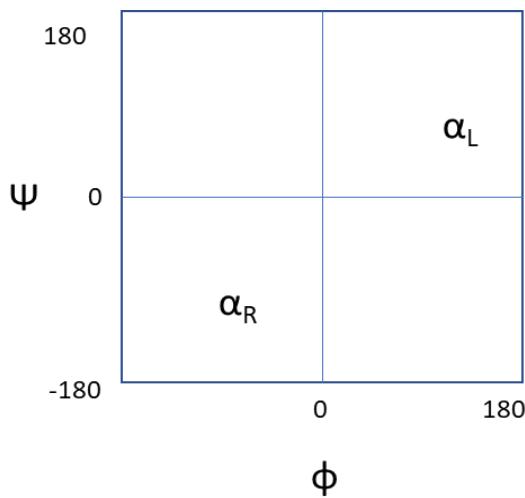


Figure 5.1: The α_R denotes right-handed helical region, α_L denotes left-handed helical region on Ramachandran map. The alternating dihedrals in α_R and α_L region form α -sheet. The Phi (ϕ) indicates the C-N-C α -C dihedral angle and Psi (ψ) indicates the N-C α -C-N dihedral angle.

The α -pleated sheet secondary structure is atypical in protein crystal structures, which is why it has not been widely studied since its proposal by Pauling and Corey. Daggett et al. recently published several papers relating to the α -sheet secondary structure.^{1,2,5,6} In these studies, four unrelated proteins were predicted to adopt this newly proposed secondary structure, under amyloidogenic conditions, using atomistic molecular dynamics simulations. This discovery suggested the α -pleated sheet structure may define the prefibrillar amyloidogenic intermediate in amyloid diseases.⁵

The study by Daggett et al. was on four proteins that are implicated in amyloid diseases; Transthyretin (TTR), β_2 -microglobulin (β_2m), lysozyme, and the prion protein (PrP).⁵ It has been proven experimentally that the formation of the amyloidogenic intermediate requires changes in the structure upon pH change. However, it is impossible to distinguish these conformational changes using experimental methods, as they do not reach atomic resolution. Thus, the study MD simulations were performed on these proteins at neutral and low pH in order to explore the conformational changes at atomic resolution. The MD simulations illustrated the presence of α -pleated sheet secondary structure formation over the strands that form at regions implicated in the amyloidogenic conversion in lysozyme and the prion protein. As a result, it was suggested that the formation of α -pleated sheet structure could be a normal conformational transition in amyloidosis as an intermediate secondary structure.

The same author then studied the monomeric TTR, at neutral and low pH, at physiological (310K) and higher (498K) temperatures.⁶ It has been reported experimentally that the native homotetramer tertiary structure of TTR protein, is split into four β -sheet monomeric amyloidogenic intermediates (two strands of both CBEF and DAGH) at pH between 3.6 and 5.2, which can then self-assemble into amyloid fibrils.⁷ Thus, the MD simulation of monomeric TTR was done at neutral pH 6–7 (all residues are at their normal protonation state), at medium pH 4–6 (all His residues are protonated), and at low pH 2–4.2 (all His, Asp, and Glu residues are protonated). The study reasoned the use of low- and medium-pH, through experimental studies which suggested amyloid fibrils forms under these conditions (pH 3.6–5.2). The simulation results at low-pH, show disturbed strands of CBEF, while α -sheet structures were observed in the DAGH strands, and deviation from the native hydrogen bonding pattern was detected. At both temperatures, the ratio of β -content was greater in medium pH than that

at neutral pH, leading to the conclusion that α -sheet can be considered as a key pathological conformation during amyloidogenesis. There is notable deviation of hydrogen bonding patterns from the native to the DAGH-sheet, depicting a conversion from β -sheet to α -sheet secondary structure.

Daggett published a review in 2006,¹ gathering information on α -sheet; such as structural properties, experimental evidence for α -sheet structure in native proteins, and the possibility of development of toxic conformers in amyloid diseases, as a result of this structure. Through the various studies, they noticed that the α -sheet structure is observed in all amyloidogenic proteins, but not in other proteins nor in amyloidogenic proteins under normal conditions. The position of α -sheet found in most amyloidogenic regions, was also determined experimentally.

While most of the clinical trials for Alzheimer's disease are targeting the monomeric or fibrillar structures of the A β peptide, the soluble oligomers are toxic and relate to the disease's progression. In 2019, Daggett synthesised and characterized α -sheet within A β 42 that was taken from the ERI Amyloid Laboratory LLC, peptides to investigate the implications of the α -sheet hypothesis in AD.² The experimental data provided evidence that A β soluble oligomers adopt α -sheet structure that forms in advance of the aggregation, also associated with toxicity. A similar conclusion was reached recently from atomistic MD simulations on A β 42.⁴ The molecular dynamics study supported that A β converts to α -strand conformation by peptide-plane flipping, then several adjacent α -strands interact via bifurcated hydrogen bonding, forming α -sheets. These α -sheets play an important role of promoting aggregation, as well as fibrillation of A β . The study aimed to understand the aggregation mechanism of Alzheimer's amyloid- β , facilitated by α -strand or α -sheet (multiple analogous α -strands). The interpretation of this structure's aggregation mechanism can provide significant structural insights for the design of effective AD medication. A recent study by Daggett on human transthyretin, showed that formation of α -sheets is destabilised by the presence of Ca²⁺.⁸

An α -strand can form an almost straight, nest-like chain, that increases the polarity of this conformation more than either α -helices or β -strands. This can create a binding site that is ready to accept an anion or cation due to adjacent main-chain carbonyl groups on one side of

the backbone, and NH groups on the other.⁹ This has been seen, for example, in a potassium channel structure (PDB entry 1BL8).¹⁰ There is a study that has distinguished the preference of amino acids, whose dihedral angles overlap at the α_L and γ_L regions on Ramachandran propensity plots, corresponding to α -sheet and β -turn forms,¹¹ respectively, reporting glycine favouring γ_L but not α_L , whereas polar residues, such as asparagine, favour α_L .¹¹

In previous chapters, we observed α_L in our simulation for A β bound to transition metal ions.¹² Thus, in this chapter, we investigate α -sheet strands in some model peptides, with regards to the relative energy stability, comparing to β -strand and α -helix in solvent, and presence of ions, via DFT, xTB and MD. The study of ionic effects on α -sheet conformation of small models of the peptides, extended to full length peptide, reported residues within the chain tend to adapt α -sheet motifs (the secreted proteins intestinal trefoil factor: 1E9T), via molecular mechanics methods in explicit solvent.

5.2 Methodology

5.2.1 Quantum Mechanics

The peptides that have been proposed to display α -sheet conformations were extracted from the PDB entries reported by Armen et al.⁶ The highlighted residues from that study were retained, and amide was capped at N- and C-termini to avoid end effects of charged groups. Canonical α -helix ($\phi = -65^\circ$, $\psi = -39^\circ$) and β -strand ($\phi = -120^\circ$, $\psi = +113^\circ$) forms, of the same sequences, were then manually constructed for comparison, including amide caps at the termini. Geometries were optimised at B3LYP-D2/def2-SVP level, in CPCM model of aqueous solvent, and at semiempirical level by GFN2-xTB method in GBSA solvent model.

5.2.2 Molecular Dynamics Simulation

The Amber package with the ff14SB force field was employed to perform MD simulations,¹³ on PDB entry 1E9T.¹⁴ These were carried out at high (498 K) and physiological (310 K) temperatures, as higher temperatures are known to accelerate protein unfolding, without changing the unfolding pathway.¹⁵ MD simulations were performed at neutral pH (His protonated at delta position), medium pH (His protonated at delta and epsilon positions), and low pH (protonated His, Asp, and Glu), to study the impact of pH on the development of α -

sheets. The protein was solvated in a cubic box of TIP3P water molecules.¹⁶ 10 Å buffering distance was added between the edges of the box and the protein, and the water molecules ranged between 4500-5300. This resulted in a cubic periodic box with side length 70 x60 x53 Å for simulation at low pH and 65x 60x 55 Å for simulation at neutral pH. Na⁺ and Cl⁻ were added to the box to neutralise the peptide based on the protonation state of the protein. An additional 20 ions of K⁺ and 20 of Cl⁻ were added, corresponding to 150 mM of buffer ions, at random positions within the box depending on the volume size. This solvent concentration was used to mimic the value used in high resolution structure of 1E9T peptide; human intestinal trefoil factor, which contains potassium phosphate and potassium chloride.¹⁴

The energy of the system was minimized using 1000 steps of the steepest descent algorithm. The structure was equilibrated in two phases: (1) initial equilibration was conducted for 50 picoseconds (ps) under an NPT ensemble (constant number of particles, pressure and temperature) – the targeted pressure was maintained at 1 bar using a Monte Carlo barostat. (2) this was followed by a second equilibration step, which was conducted for 100 ps under an NVT ensemble (constant number of particles, volume, and temperature) – temperature scaling was maintained with Langevin dynamics thermostat. The MD production was run for 50 ns at 310 and 498 K, under periodic boundary conditions and coordinate trajectories were written every 2 ps. The SHAKE algorithm was applied to restrain hydrogen atoms to their equilibrium positions.¹⁷ Short-range interactions were truncated at 10 Å and long-range interactions were treated with the particle mesh Ewald (PME) method. Trajectory analysis was done using CPPTRAJ.¹⁸

5.3 Results and discussion

The 14 peptide segments that contain three residues or more, that display the alternation of residues in (+φ, +ψ) and (-φ,-ψ) conformational space, used in this study, were selected from Armen et al.,⁶ who looked for alternating α_L and α_R amino acids in a nonredundant version of the Protein Data Bank, summarized in Table 5.1. DFT optimisation was performed to obtain reliable representation of both geometry and relative energy of conformations. GFN2-xTB was used for estimation of energies for a relatively low computational expense, especially in the case of larger peptides.

Table 5.1: peptides that show alternating α_L and α_R amino acids of a nonredundant version of the Protein Data Bank reported by Armen et al.⁶

#	PDB Code	Resolution (Å)	Residues	Amino Acids sequence	Charge
1	1B0P	2.3	673–676	Glu-Lys-Arg-Gly	+1
2	1POI	2.5	97–100	His-Lys-Ala-Asn	+1
3	1FEP	2.4	379–382	Ser-Asn-Thr-Gln	0
4	1IDM	2.1	262–265	Lys-Asn-Leu-Cys	+1
5	1BWU	2.8	76–79	Ser-His-Ser-Val	0
6	4AAH	2.4	22–25	Tyr-Ser-Gln-His	0
7	1HIU	2.6	197–200	Gly-Tyr-Ser-Phe	0
8	1GQF	2.9	341–344	Arg-Ser-Pro-Gly	+1
9	1F9B	2.7	261–264	Asp-Gly-Arg-Glu	-1
10	1A05	2.0	31–34	Alal-His-Leu-Gly	0
11	1B7T	2.5	99–102	Glu-Gln-Thr-Lys	0
12	1E9T	NMR	19–22	Val-Asp-Cys-Gly	-1
13	1QND	NMR	96–99	Gln-Gly-Lys-leu	+1
14	1BVH	NMR	66–69	His-Gly-Ile-Asn	0

The results from geometry optimization by DFT show all conformation optimizations were completed with the stationary point found. The geometry optimization at B3LYP-D2/def2-SVP/CPCM level, retains the structures of α -helix, β -strand and α -sheet: values of (ϕ, ψ) angles (Table B1, Appendix B), place the α -sheet firmly in the $\alpha_R\alpha_L$ region of the Ramachandran map. The dihedral angles outside the central region are also within canonical values and were not altered by end effects.

Since, checking the suitability of the DFT method's predictions used here is essential, the relative energy of β -strand and α -sheet conformations has been tested by our group previously on Ala4 model, using *ab initio* data on the DFT-optimised geometry, as a benchmark.¹⁹ B3LYP-D2 def2-SVP/CPCM shows the relative energy value of α -sheet relative to β -strand Ala4 is -4.80 kJ mol⁻¹ compared to -6.17 kJ mol⁻¹ predicted by CCSD(T) CBS performed at CPCM.¹⁹ The excellent performance demonstrated by B3LYP-D2 def2-SVP/CPCM comparing to *ab initio*, led to it being applied in the present study.

Here, the relative energy values predicted by B3LYP-D2/def2-SVP, of α -helix and α -sheet and comparing to β -strand, favour α -helix in most cases. However, there is a preference for to α -sheet in two fragments only (PDB:1B7T and 1E9T), for which α -sheet is more stable by -14.55 kJ.mol⁻¹ relative to β -strand and -2.06 kJ.mol⁻¹ relative to α -helix (Table 5.2). The residues that have been reported to adopt α -sheet, in these two peptides, are Glu-Gln-Thr-Lys and Val-Asp-Cys-Gly, in 1B7T and 1E9T, respectively.

Table 5.2: DFT relative electronic energy of and α -helix and α -sheet to β -strand for small models obtained from peptides in CPCM water (kJ mol⁻¹).

#	PDB code	α -sheet	α -helices
1	1B0P	1.20	75.10
2	1FEB	22.90	-12.60
3	1F9B	-24.77	-31.14
4	1A05	-5.12	-88.18
5	1GQF	65.74	-8.09
6	1QND	23.12	-10.54
7	1BVH	15.78	-21.28
8	1IDM	-23.43	-52.96
9	1P0I	-32.87	-76.57
10	1BWU	-13.16	-81.34
11	4AAH	-39.10	-104.35
12	1HIU	15.70	-24.96
13	1B7T	-105.45	-80.76
14	1E9T	-14.55	-12.49

GFN2-xTB was also used for estimation of energies, for a relatively low computational expense, allowing it to be further applied on larger, full-sized peptides. However, the semiempirical method has been found to only preserve β -strands and α -helices, close to the range expected for these forms (α -helices have both ϕ and ψ angles negative, and the residues in the β -conformation have ϕ negative and ψ positive angles), but the method fails to maintain the α -sheet dihedral angles at $\alpha_R\alpha_L$ region, in most of the peptide models (Table B2 in Appendix B). In addition, even though the relative energies predicted by GF2-xTB also

show a preference for α -helix in most cases, such as in DFT (9 out of 14 models), these values show a huge difference in predicted relative energies, when compared to B3LYP-D2 def2-SVP, as shown in Table B3 found in Appendix B. Also, in previous work,¹⁹ the semi-empirical GFN2-xTB method, has been found to not be able to reproduce DFT and *ab initio* relative energies, performed in solvent: therefore this method will not be considered to study this structure (α -sheet) in detail further.

Because DFT shows preference for α -sheet in the fragment of 1E9T peptide, as shown in Table 5.2, it was used for further investigation, to examine whether the presence of ions can change the predicted equilibrium between β -strand, α -sheet and α -helix, by manually placing Na^+ , Li^+ , Mg^{2+} , K^+ or Ca^{2+} ions in the central portion of each conformation. 1E9T is the PDB code for human intestinal trefoil cell motility factors, that is essential for the maintenance of all mucous-coated epithelial surfaces. The full size of this peptide is relatively small, having only 59 amino acids, consisting of polar residues that can place the cations or anions close to them. The location of α -sheet is on the surface, and the net charge of the tetrapeptide that forms the α -sheet, is -1 making it available for cation interaction. The full length peptide also contains disulfide bridges, which stabilize the peptide during MD simulations. The α -sheet reported in the peptide contains residues VAL, ASP, CYS, GLY at 19,20,21,22 sequence positions (α_R , α_L , α_R , α_L), respectively (Figure 5.2).

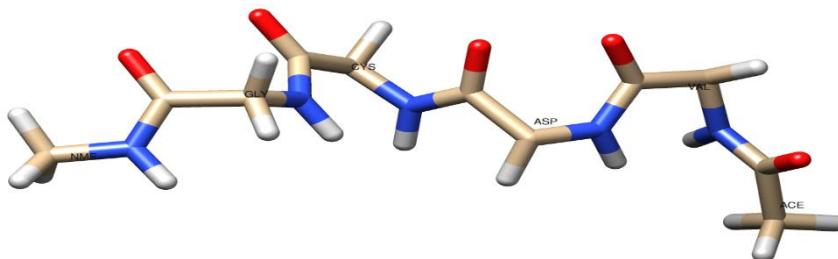
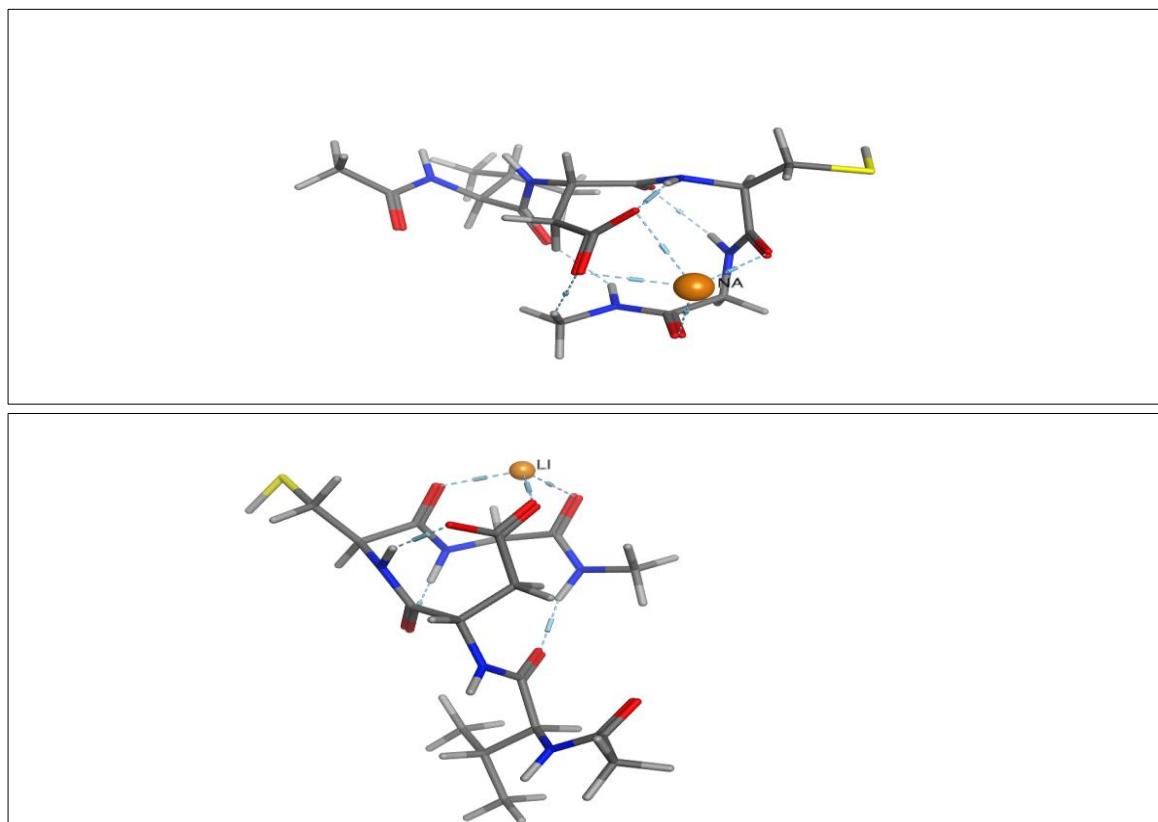


Figure 5.2: α -sheet residues in 1E9T peptide that comprises VAL 19, ASP 20, CYS 21, GLY 22 residues. The side chains are hidden for the purposes of clarity.

The initial structure for DFT calculation was chosen based on multiple starting positions, which were tested using Na^+ . This was done by first manually locating this ion in the proximity of one or more oxygen nuclei, picking the lowest final energy structure and using the position for the remaining ions. Geometry optimisation placed cations close to two oxygens from ASP

sidechain and backbone carbonyls, from both CYS and GLY amino acids, in α -sheet and β -strand, resulting from the alignment of CO and NH bonds. In the case of α -helix conformations, only two oxygens from ASP in all ions are involved in binding, suggesting that the only two contacts are available in the helix form of this fragment due to the optimization placing the ions close to oxygens of ASP side chain but far from other oxygens (Figure 5.3). Lithium ions have been found to bind to three oxygens in α -sheet form, instead of four, seen in other ions- α -sheet (one from the ASP sidechain, and backbone from both CYS, GLY), Figure 5.3. Notably, Mg^{2+} forms a tetrahedral geometry, in the case of α -sheet and β -strand, Li^+ also shows tetrahedral geometry in the case of β -strand (Figure 5.4).

The ϕ and ψ angles of tetrapeptide (VAL, ASP, CYS, and GLY) obtained from 1E9T peptide and resulting from DFT optimization show α -sheet is still there, within the sequence, after adding ions, especially at residues VAL, ASP and GLY, as seen in Table 5.3. This result is in line with a study that suggested polar residues such as ASP have a propensity to adopt the α_L conformation ($\phi,\psi)=(+,+)$ on Ramachandran map.¹¹ The result also shows GLY has favour to the same region.



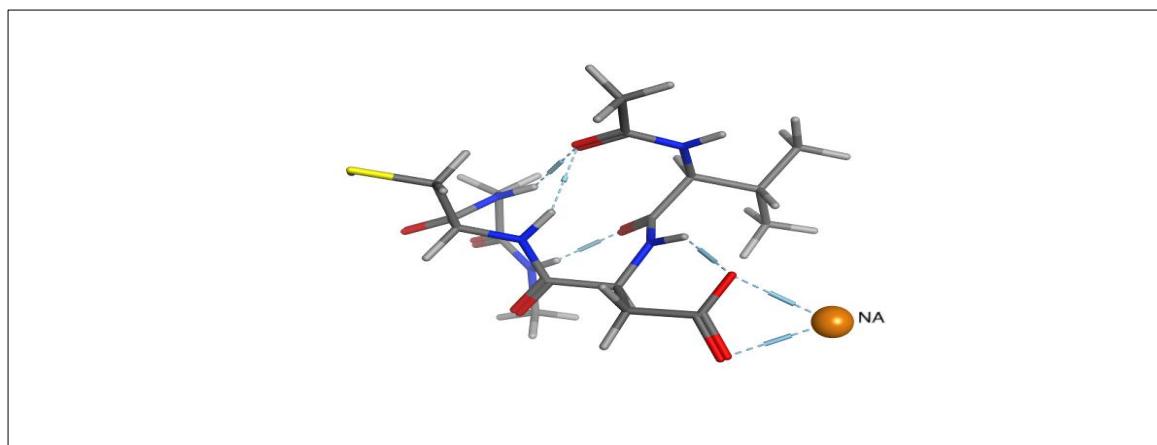


Figure 5.3: Geometry optimization of α -sheet- Na^+ (top), α -sheet- Li^+ , (middle) and α -helix - Na^+ (bottom) bound to ions.

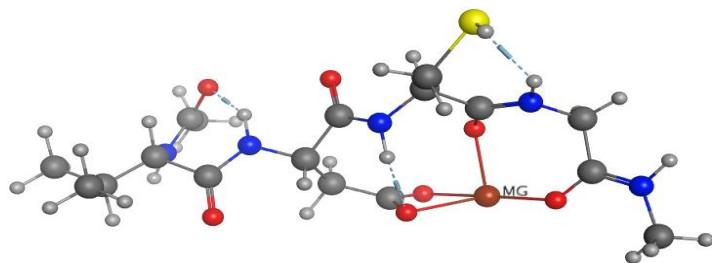


Figure 5.4: Geometry optimization of β -strand- Mg^{2+} .

Table 5.3: DFT optimised dihedral angles ($^\circ$) of α - sheet region of 1E9T peptide and ion-peptide.

Residues	Ion-free		Na^+		Mg^{2+}		Li^+		K^+		Ca^{2+}	
	ϕ	ψ	ϕ	ψ	ϕ	ψ	ϕ	ψ	ϕ	ψ	ϕ	ψ
GLY(22)	66.7	93.8	73.2	71.5	110.0	74.3	67.5	74.4	66.4	72.0	84.4	72.5
CYS(21)	-73.5	-22.7	82.2	-78.7	116.3	-89.5	-72.4	84.2	79.7	-80.1	93.6	-76.9
ASP(20)	50.9	102.5	168.3	68.6	-124	49.6	68.3	171.1	171	67.9	150	76.5
VAL(19)	-89.9	-3.9	-40.8	-60.6	-72.5	-79.7	-60.3	-41.2	-85.8,	-61.1	-24.9	-39.6

DFT/B3LYP-D2 optimization of the free four amino acid sequence, extracted from 1E9T peptide, showed a preference towards α -sheet, in terms of relative energies by $-14.55 \text{ kJ mol}^{-1}$, comparing to β -strand and α -helix (Table 5.4). The effects of cationic ions Na^+ , Li^+ ,

K^+ , Ca^{2+} , and Mg^{2+} , were also investigated to see whether or not ion binding alters the preference between these conformations.

The results show α -sheet of free sequences VDCG, obtained from 1E9T, being more stable than other forms, but by including ions, this relative stability decreased a little bit in the case of monovalent-charged ions, such as K^+ and Na^+ . In these cases, the energy difference was found to be relatively small ($0.6\text{-}9 \text{ kJ}\cdot\text{mol}^{-1}$), compared to β -strand. The stability increased in the case of Mg^{2+} by $17.3 \text{ kJ}\cdot\text{mol}^{-1}$ from the free as shown in Table 5.4, resulting from the small size of this ion and from the stable tetrahedral geometry formed. Unlike divalent Mg^{2+} , introduction of Ca^{2+} to the α -sheet decreased its stability relative to the β -strand by $8 \text{ kJ}\cdot\text{mol}^{-1}$ only.

Table 5.4: Relative energy of α -sheet and α -helix forms relative to β -sheet in CPCM water ($\text{kJ}\cdot\text{mol}^{-1}$).

1E9T	α -sheet	α -helix
Free	-14.55	-12.49
Na^+	-2.18	1.87
K^+	-8.75	1.10
Li^+	0.57	0.56
Mg^{2+}	-31.85	87.07
Ca^{2+}	8.48	40.67

In the case of α -helix, the stability decreases going from the free peptide, to the ion-bound systems, especially when divalent cations (Mg^{2+} and Ca^{2+}) were added, by 87.07 and 40.67 $\text{kJ}\cdot\text{mol}^{-1}$, with respect to β -strand. The ionic charge, as well as ionic radius, appear to be related to the degree of ionic effects and thus on the stability order of α -sheet, since K^+ , N^+ , and Li^+ have smaller effects on the relative energies, compared to the ones seen in the cases of Mg^{2+} and Ca^{2+} . Additionally to the ionic charge effects on the stability of the secondary structures, in the case of α -helix, there are fewer ion-O bonds (two for α -helix but four for β -strand and α -sheet), making this structure less stable than both β -strand and α -sheet. Finally, because the α -sheet is stabilised by bifurcated hydrogen bonding with neighbouring chains, as a result of alternating αL and αR amino acids, the more stable relative energies are expected to increase further if there is more than one chain. The size of model system required to study this is, however, too large for us to tackle with DFT.

These calculations provide some evidence for conformational preference from ionic effects, however they do not account for the dynamic processes that are important in the peptide's structure and function. In addition, using accurate DFT methods limits the size of the peptide, due to the computational cost. Likewise, there are limitations when using implicit solvent models, in specific peptide-solvent and ion-solvent interactions. Therefore, the molecular dynamics simulation of the full-length 1E9T protein was carried out in explicit solvent.

5.3.1 Molecular dynamics simulation of 1E9T peptide

The RMSD and R_g values calculated from the simulations (Figure 5.5, 5.7 and 5.8) show there is not much difference in the stability or geometry between the free and ion-bound peptide, with RMSD and R_g values at *ca.* 6 and 11.5 Å, respectively. Simulations at higher temperature showed fluctuations between the maximum and minimum values of RMSD (5 and 12 Å), with an average around 7 Å. The R_g follows a similar trend in all three different pH. An increase in RMSD values is observed towards the end of the trajectory of neutral pH at physiological temperature, from 40 to 50 ns, at the same time matching the lower value in R_g . The α -sheet structures were detected at neutral pH and 301 K in both free and with ions simulations performed in explicit solvent, but not at high temperature or low pH. The observed α -sheet structure lasts for a very short simulation time of 3 and 5 ns only, as shown in Figure 5.6. The changes in ϕ and ψ angles as a function of time, for the individual residues (VAL, ASP, CYS and GLY) can be found in Appendix C.

The introduction of ions has an effect on the longevity of the α -sheets, which appear to last for 5 ns, at 310 K, compared to the peptide without ions which only last for 3 ns. This observation, however, does not hold true for the α -sheets at low pH. Having established the transient stability of α -sheets, their appearance after equilibration of the system would be temporary.

However, the average values of ϕ and ψ of the whole trajectory (Figure 5.9), show α -sheet consisting of two alternating amino acids (–,++) of CYS21, GLY22 and ARG34, GLY35, in most of the conditions. There is also α -sheet between GLU2, TYR3 that is only seen at medium pH and 310 K. But as this form contains only two amino acids located at ϕ and ψ region of α -sheet, these may not α -sheet, and might be better characterised as a turn(α_L, γ_L), especially

since the majority of residues locate at (ϕ , ψ) = (+, +) are glycine in which favours γ_L according to the published study.¹¹

The aim of this study was to investigate the ionic effects on the protein α -sheet structure. The analysis of the results shows the presence of α -sheets only at neutral pH and 310 K, with and without ions (KCl). Also, the results show α -sheets, observed during the simulation, have transient lifetime, devolving within the first 3 ns out of the 50 ns, before turning into β -strand. These findings are in line with what was expected to develop from the MD simulations, having considered the intermediate nature of α -sheets, appearing only during the conversion of α -helices to β -sheets.^{4, 6, 20}

Also, as the literature suggested, the α -sheet has been found to form from hydrogen bonding between adjacent α -strands via peptide-plane flipping. Our future work on this structural investigation, shall include more than a single monomer, looking at the presence of α -sheet intermolecularly between dimers, using the system studied here but also different proteins. Also, our study shall involve different ions, in particular anions which have not been studied here, to see their effect on the secondary structure of the systems.

Finally, the main aim of this thesis was the investigation of the A β peptide, including metal binding and α -sheet structure, that can induce its aggregation resulting in Alzheimer's disease. The MD study reported that the VGS sequence of A β 42 (region 24–26), tends to adapt the α -strands structure observed at different pH and temperatures.⁴ In the same context, several studies have pointed out the involvement of region 23–28 in A β fibril formation.^{21–23} Therefore, investigation of the metal ion effects on the formation of this structure, especially at this region (24–26) is important, as metals have been linked to A β aggregation. The aforementioned features are to be explored in future studies.

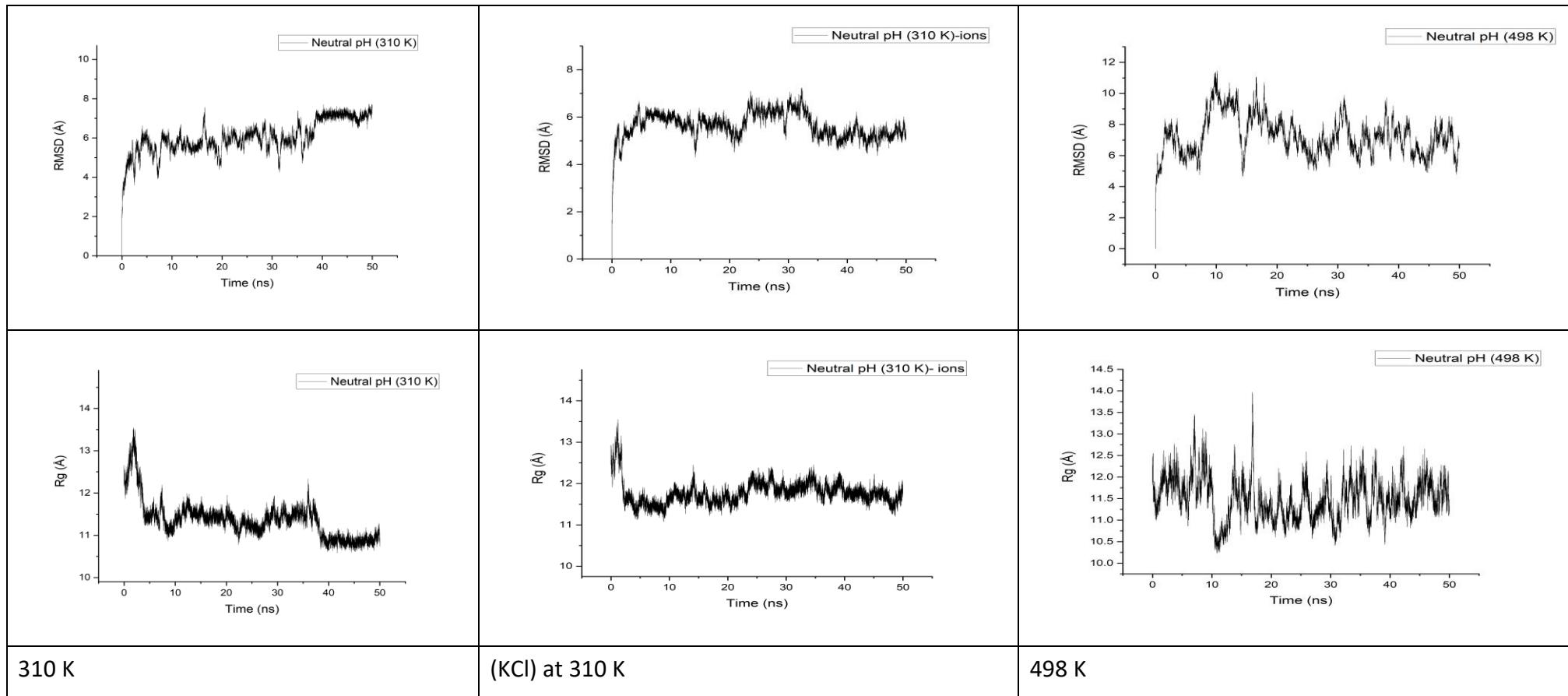


Figure 5.5: Backbone RMSD and R_g of 1E9T simulation at neutral pH and 310/498 K.

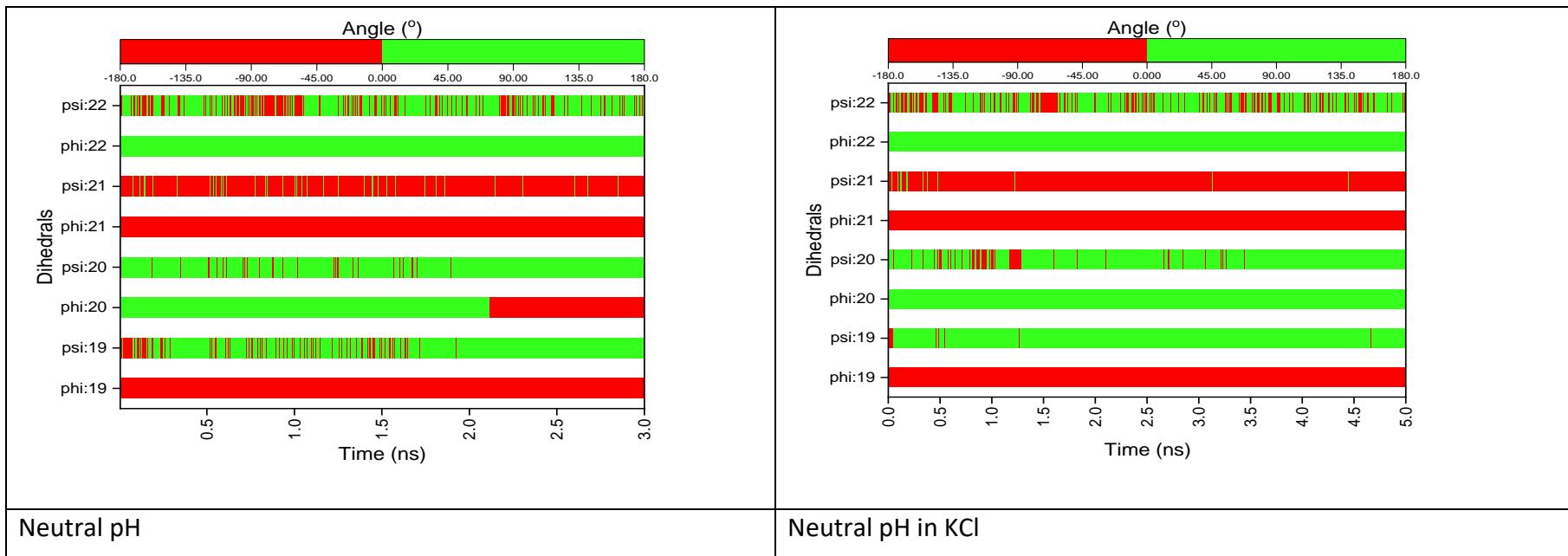


Figure 5.6: Dihedral angles evaluation of the region that present proposed α -sheets of 1E9T simulation at neutral pH and 310 K during 50 ns. Note: only frames that show α -strand are presented here for clarity (3 and 5 ns). The dihedral angles of the entire trajectories can be found in the Appendix. The literature reported the α -sheet of this peptide at residues 19,20,21,22 to be VAL19 (negative ϕ and ψ), ASP20 (positive ϕ and ψ), CYS21 (negative ϕ and ψ), and lastly GLY22 (positive ϕ and ψ).⁶

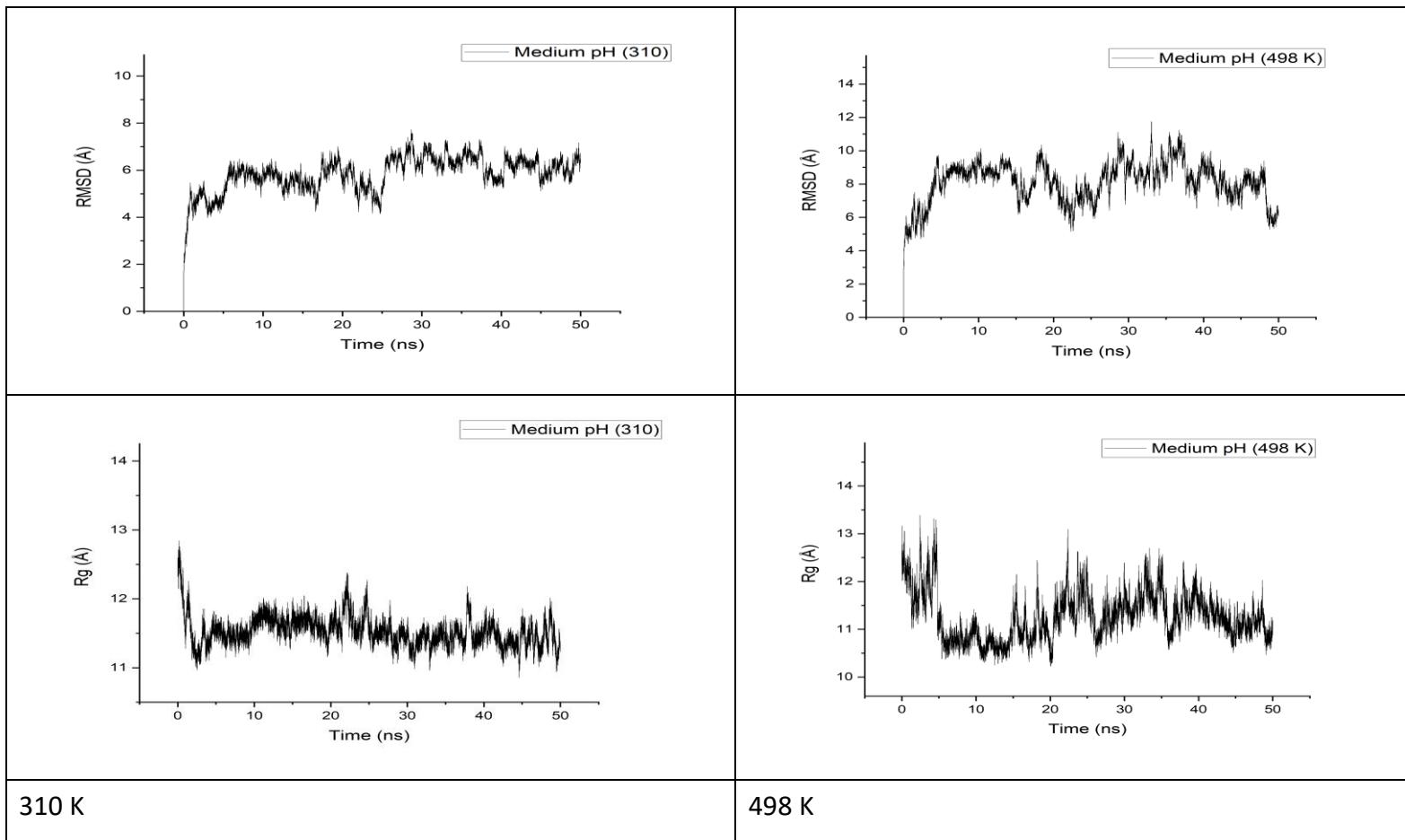


Figure 5.7: Backbone RMSD and R_g of 1E9T simulation at medium pH and 310/498 K.

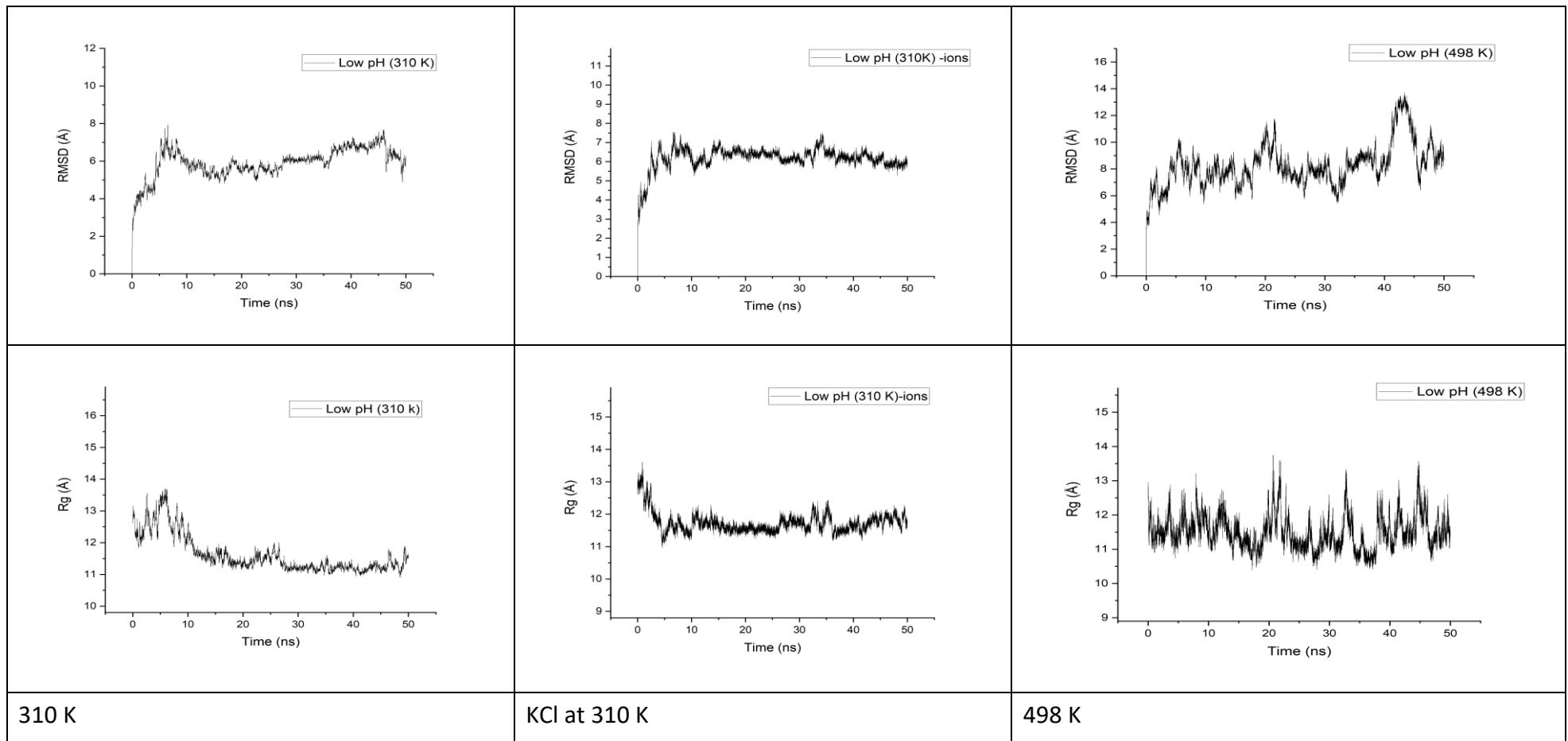
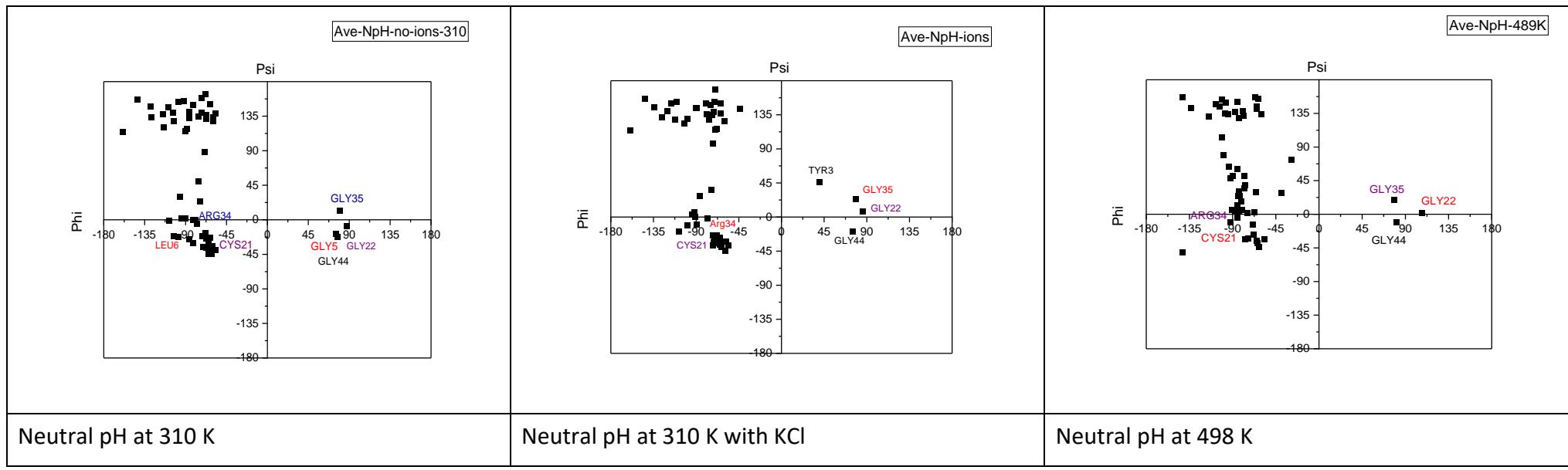
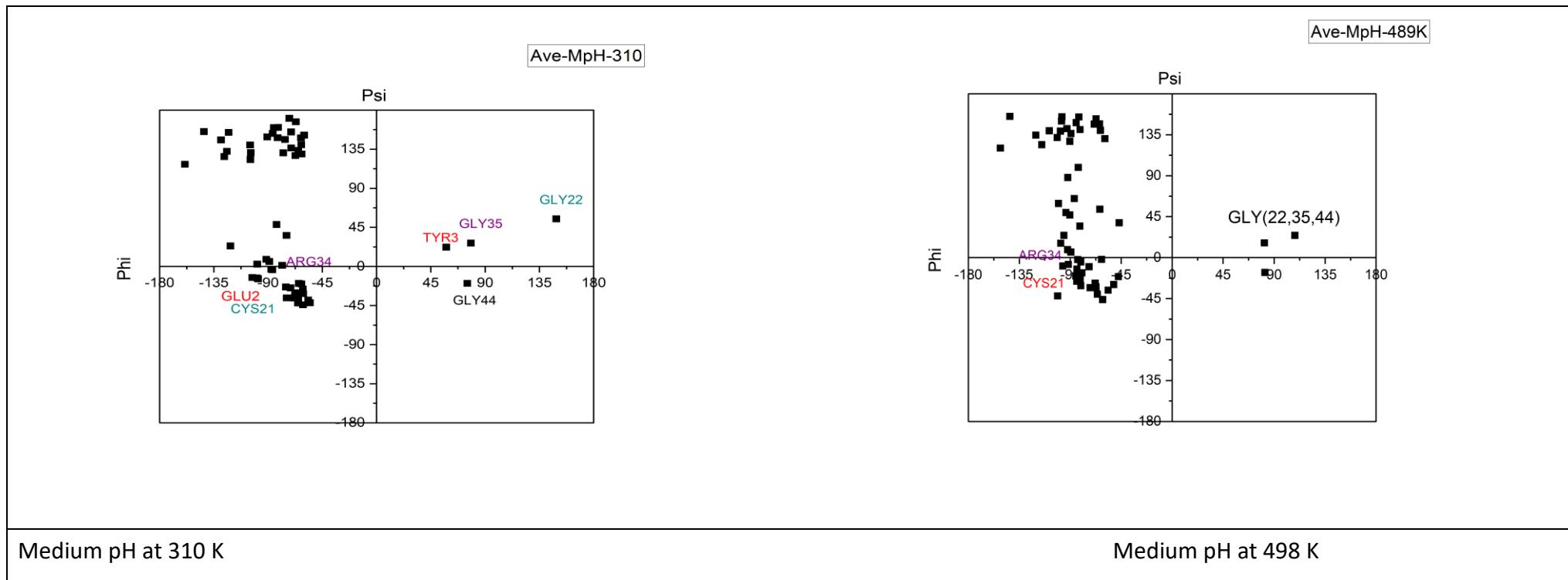


Figure 5.8: Backbone RMSD and R_g of 1E9T simulation at Low pH and 310/489 K.





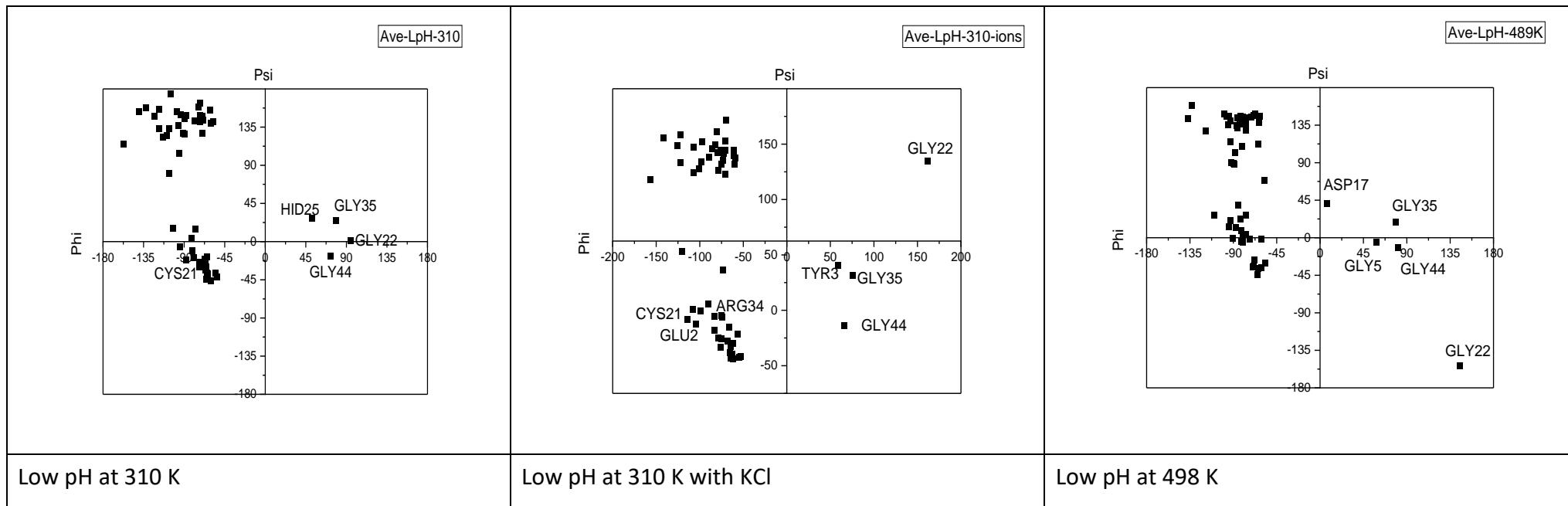
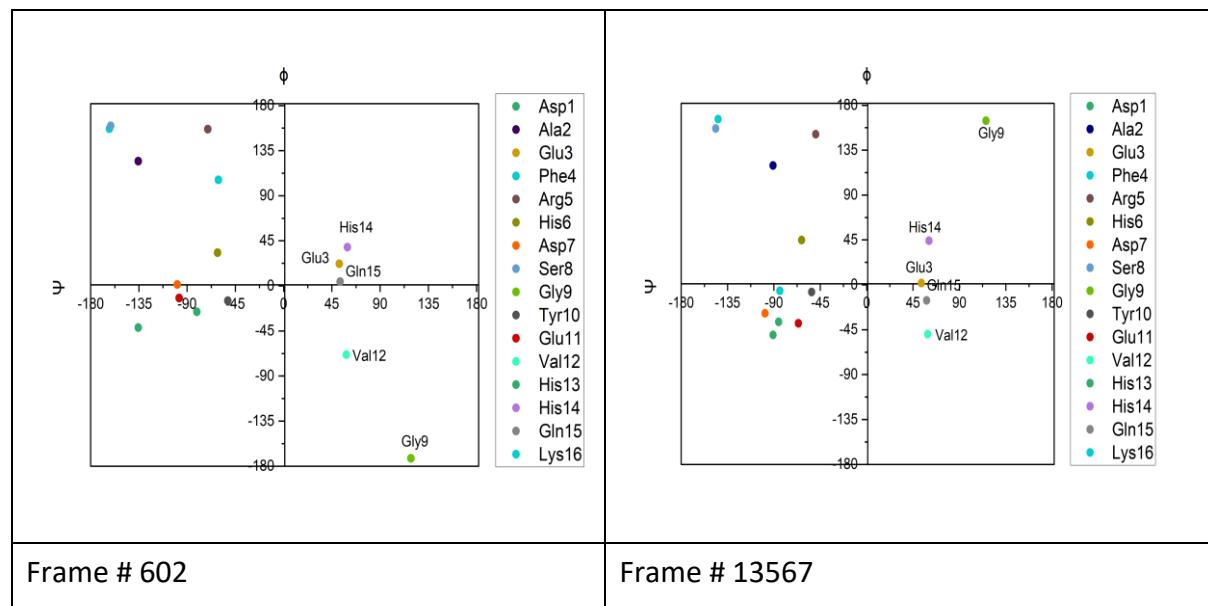


Figure 5.9: Ramachandran map plots of the average values of phi and psi. The amino acids locate on $+\phi$, $+\psi$ region (αL) are pointed as well as the alternate residues at $-\phi$, $-\psi$ region (αR) if there any.

5.4 Investigation the potential of α -sheet structure in Zn(II)-A β 16

Following the related discussion in Section 4.3.2, on the observation of left-handed helical structures (α L) in the Ramachandran map of Zn(II)-A β 16 binding mode 1; the possibility of them being expressed as a result of the presence of α -sheet is investigated here. Multiple frames, which expressed α L population were extracted individually, with the ϕ and ψ angles examined, looking for alternation of the helical α R and α L forms of these residues. However, the plots of dihedral angles (Figure 5.10) show the detected secondary structure in the α R and α L regions, was not α -sheet, according to the specified definition of this form. Even though in certain cases, residues His13, His14, and Gln15 were found at α R, α L, α R (frame: 59999), this is not seen frequently in many frames. The plots presented in Figure 5.10 show alternate residues of Glu11, Val12, and His13 adopt the α R and γ L (α R, γ L, α R) regions, suggesting a nests form. The plots also show His14, Glu3, and Gln5 amino acids are likely to be found in α L, and it has been suggested that charged and polar amino acids favour α L.¹¹ Gly9 is known to have unusual conformational freedom and is commonly found in turns, this is also seen in this mode of Zn(II)-A β (Figure 5.11). The table of ϕ , ψ values for these frames can be found in Appendix D.



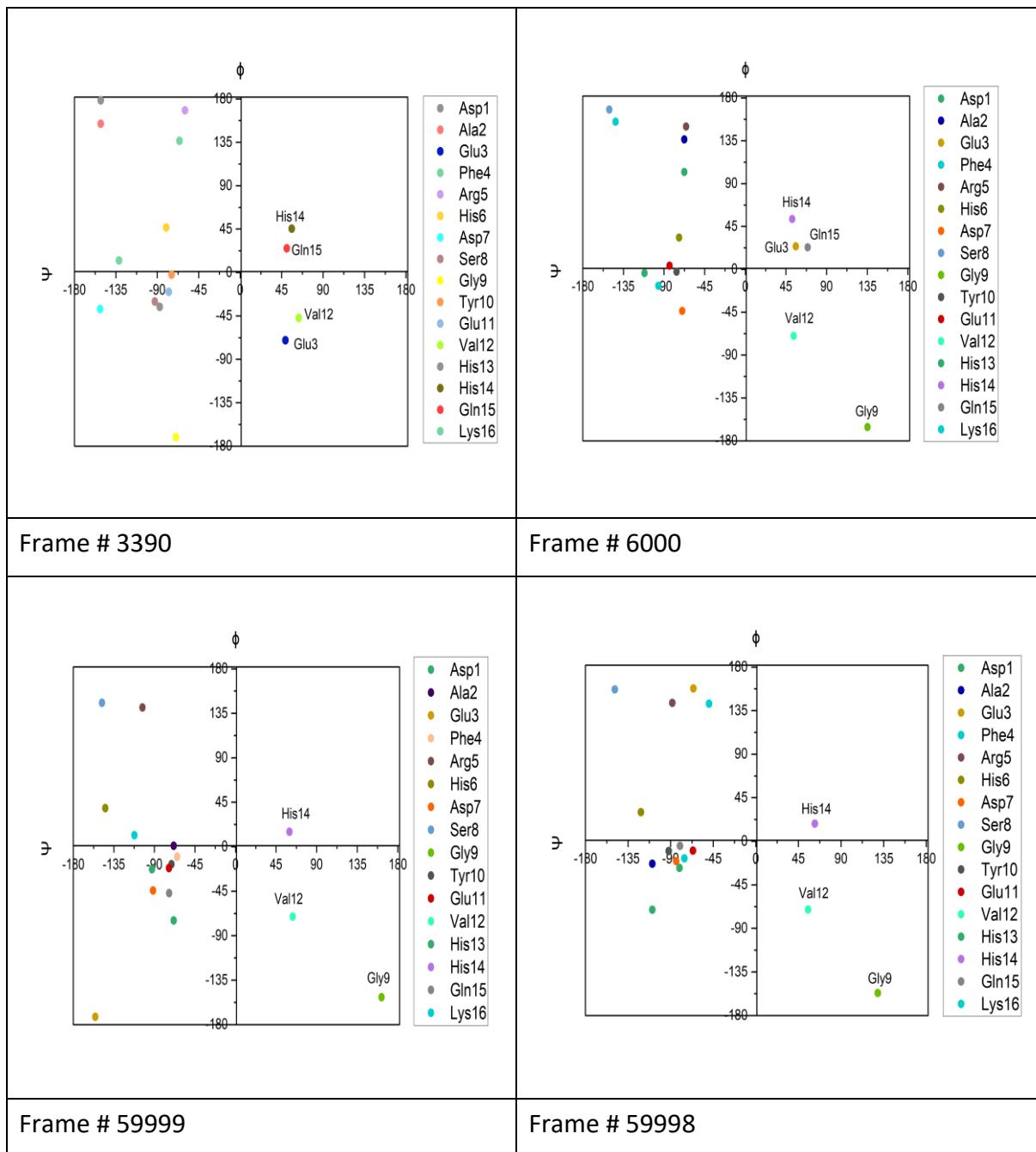


Figure 5.10: Phi and Psi angles of Zn(II)-A β 16-Mode1 simulation that correspond to the different extracted frames, that showed population at left-handed helix in the Ramachandran map (α L).

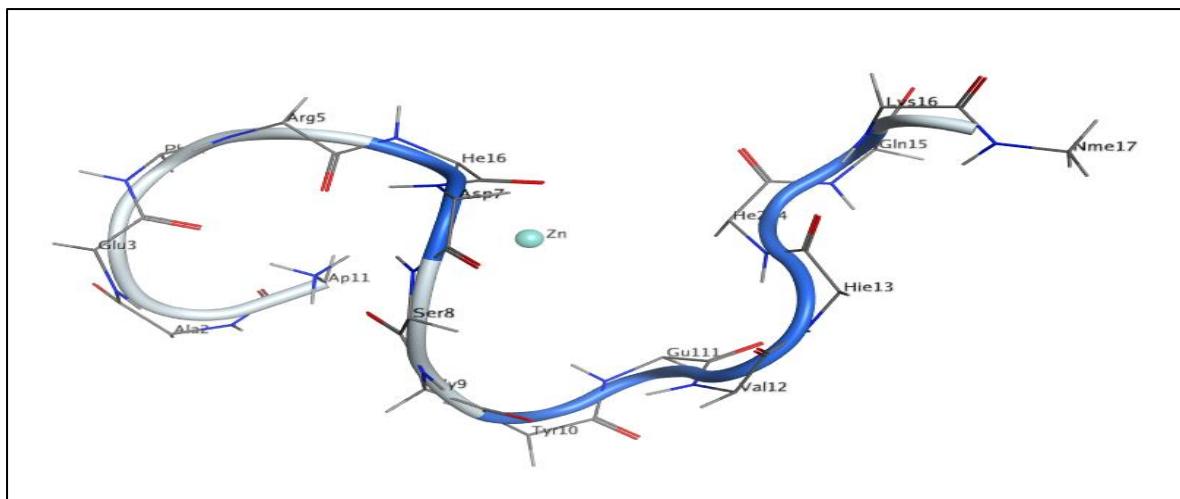


Figure 5.11: Zn(II)-A β 16 structure generated from frame number 59999 of aMD simulation. Glu11, Val12, and His13 form nests as they occupied α R, γ L and α R regions, respectively. His13 and His14 occupied α R and α L, suggesting it is more turn than sheet. The Zn(II) here coordinated to Asp1, His6, Glu11, and His14. Note that only backbone atoms are shown, while the rest are hidden for clarity.

5.5 Conclusion

A number of experimental and theoretical studies have confirmed that the α -sheet structure in A β is an intermediate form that could lead to aggregation and toxicity in AD. In this study, peptides bound to ions were tested, to examine whether ions can induce the formation of α -sheets. DFT and semiempirical GFN2-xTB calculations were used on the modelled peptides, to evaluate the relative stability of α -pleated, α -helix, and β -strand conformations. The α -helix in the fragmented models of the peptides, presents the most stable forms in most cases, of tetrapeptides, when compared to α -sheet and β -strand characters. However, in the case of peptide 1E9T, the α -sheet presents the most stable form, and the stability changes in the presence of ions. The stability order has been found to be largely influenced by the ionic charge of the bound ions. Ca $^{2+}$ and Mg $^{2+}$ have the greatest effect on the relative stability followed by K $^+$, Na $^+$ and Li $^+$, which have very little effect on the relative stability, when compared to the system without an ion. In the case of 1E9T, α -helix was predicted to have higher relative energy values, when compared to other forms, as a result of ions being bound to only two oxygen atoms, as opposed to them being bound to four in the other forms.

The MD simulations of the entire 1E9T protein, show ions do not have much effect on the α -sheet structure. Also, the simulations show α -sheets have a short lifetime during the

simulations, in agreement with literature that described this conformation as an intermediate towards a more stable secondary structure.

5.6 References

- 1 V. Daggett, α -Sheet: The Toxic Conformer in Amyloid Diseases?, *Acc. Chem. Res.*, 2006, **39**, 594–602.
- 2 D. Shea et al., α -Sheet secondary structure in amyloid β -peptide drives aggregation and toxicity in Alzheimer's disease, *Proc. Natl. Acad. Sci.*, 2019, **116**, 8895–8900.
- 3 L. Pauling et al., The Pleated Sheet, A New Layer Configuration of Polypeptide Chains, *Proc. Natl. Acad. Sci.*, 1951, **37**, 251–256.
- 4 A. Balupuri et al., Aggregation Mechanism of Alzheimer's Amyloid β -Peptide Mediated by α -Strand/ α -Sheet Structure, *Int. J. Mol. Sci.*, 2020, **21**, 1094.
- 5 R. S. Armen et al., Pauling and Corey's α -pleated sheet structure may define the prefibrillar amyloidogenic intermediate in amyloid disease, *Proc. Natl. Acad. Sci.*, 2004, **101**, 11622–11627.
- 6 R. S. Armen et al., Anatomy of an Amyloidogenic Intermediate: Conversion of β -Sheet to α -Sheet Structure in Transthyretin at Acidic pH, *Structure*, 2004, **12**, 1847–1863.
- 7 S. L. McCutchen et al., Transthyretin mutation Leu-55-Pro significantly alters tetramer stability and increases amyloidogenicity, *Biochemistry*, 1993, **32**, 12119–12127.
- 8 E. Wieczorek et al., Destabilisation of the structure of transthyretin is driven by Ca²⁺, *Int. J. Biol. Macromol.*, 2021, **166**, 409–423.
- 9 J. D. Watson et al., A novel main-chain anion-binding site in proteins: The nest. A particular combination of ϕ , ψ values in successive residues gives rise to anion-binding sites that occur commonly and are found often at functionally important regions, *J. Mol. Biol.*, 2002, **315**, 171–182.
- 10 D. A. Doyle et al., The structure of the potassium channel: molecular basis of K⁺ conduction and selectivity, *Science*, 1998, **280**, 69–77.
- 11 S. Hayward et al., Determination of amino acids that favour the α L region using Ramachandran propensity plots. Implications for α -sheet as the possible amyloid

intermediate, *J. Struct. Biol.*, 2021, **213**, 107738.

- 12 N. Al-Shammari et al., Forcefield evaluation and accelerated molecular dynamics simulation of Zn(II) binding to N-terminus of amyloid- β , *Comput. Biol. Chem.*, 2021, 107540.
- 13 J. A. Maier et al., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB, *J. Chem. Theory Comput.*, 2015, **11**, 3696–3713.
- 14 X. Lemercinier et al., High-Resolution Solution Structure of Human Intestinal Trefoil Factor and Functional Insights from Detailed Structural Comparisons with the Other Members of the Trefoil Family of Mammalian Cell Motility Factors, *Biochemistry*, 2001, **40**, 9552–9559.
- 15 R. Day et al., Increasing temperature accelerates protein unfolding without changing the pathway of unfolding, *J. Mol. Biol.*, 2002, **322**, 189–203.
- 16 D. J. Price et al., A modified TIP3P water potential for simulation with Ewald summation, *J. Chem. Phys.*, 2004, **121**, 10096–10103.
- 17 J.-P. Ryckaert et al., Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *J. Comput. Phys.*, 1977, **23**, 327–341.
- 18 D. R. Roe et al., PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data, *J. Chem. Theory Comput.*, 2013, **9**, 3084–3095.
- 19 N. Alshammari et al., Can ionic effects induce α -sheet conformation of Peptides?, *Chem. Phys. Lett.*, 2021, **784**, 139095.
- 20 A. Balupuri et al., Computational insights into the role of α -strand/sheet in aggregation of α -synuclein, *Sci. Rep.*, 2019, **9**, 59.
- 21 A. Melquiond et al., *Curr. Alzheimer Res.*, 2008, **5**, 244–250.
- 22 R. Roychaudhuri et al., Gly25-Ser26 Amyloid β -Protein Structural Isomorphs Produce Distinct A β 42 Conformational Dynamics and Assembly Characteristics, *J. Mol. Biol.*, 2014, **426**, 2422–2441.
- 23 R. Roychaudhuri et al., A Critical Role of Ser26 Hydrogen Bonding in A β 42 Assembly and Toxicity, *Biochemistry*, 2017, **56**, 6321–6324.

6 Conclusions and future work

The main aim of the work presented in this thesis was to examine the bonding effect between transition metal ions and the amyloid- β monomer that is associated with Alzheimer's disease, as a causal factor to its aggregation. Trace ions have been implicated in the pathogenesis of the Alzheimer's disease through alteration of the kinetics pathway of A β peptide. Even though the peptide has been studied experimentally, the flexible nature and transient character of the peptide are challenging, preventing assessment of the peptide's structure and dynamics where these can play a key role in functionality.

Therefore, theoretical approaches are a powerful way to complement experimental studies, in order to understand the peptide's conformational dynamics and related properties. However, modelling complicated systems such as A β peptides that also form complexes with metal ions, with ab initio methods is computationally costly. Alternatively, we considered testing less computational expensive approaches; tight binding, semiempirical method GFN2-xTB, which can give valuable information and save calculation time, for modelling similar systems involving protein-metal complexes. The initial method evaluation was performed on the small GHK-Cu(II), similar in structure to the A β -16 fragments, which were then tested. The LFMM model was used for modelling the d-electron effects of transition metals, while the ff14SB amber force field was used to model the rest of the peptide.

The accelerated MD method, in which energy barriers are artificially reduced, was used to increase the sampling space in simulations of metals binding to different binding coordination sites of A β 16 monomer, to study the effect of different binding modes and metals on the peptide's properties. In the same context, α -sheet secondary structure has been proposed as the key component of the prefibrillar intermediate during amyloid formation. Therefore, MD simulation with ff14SB was performed on protein (PDB CODE: 1E9T), reported to contain this structure, to investigate its effect on amyloid aggregation.

Chapter 3 described an extended version of xTB semiempirical methods that was used to investigate the biological-metal systems. The approach's efficiency was first assessed on the small model (GHK-Cu), before being utilised on A β 16-Cu fragments. The performance of GFN2-xTB was good and shows satisfactory result when compared to DFT(B3LYP-D2) method with less computational cost. The method looks promising for modelling more complicated

biological systems that contain metal ions such as proteins or peptides and ligands in the future.

Despite the fact that the metal-A β peptide complexes have previously been studied theoretically, the comparison between coordination modes had not been fully covered. In Chapter 4, different binding modes, obtained from experimental findings of Zn(II), Fe(II), and Cu(II) bound to A β 16, along with the peptide free of metal were simulated using advanced sampling methods. Accelerated molecular dynamics (aMD) were employed to explore the conformational space more efficiently than cMD, to seek an understanding of the conformational dynamics of A β peptide and structural changes with and without metals. The overall comparison shows the impact of metal coordination on the free A β peptide's flexibility and folding size, indicated by the reduced RMSD and R_g values, as well as RMSF of the residues' mobility, in contrast to their initial position. The most significant salt bridge interaction seen in all metal's modes, took place between residues Arg5 and Asp7. Some of the contacts between residues last over the entire simulation time and can affect the overall backbone RMSD of the peptide, restricting its conformational transitions. The number of coordinated atoms limit the flexibility of the peptide, as seen in Fe(II) and Cu(II). The result shows that the overall peptide structure, size, and stability change depending on the different binding modes. Also, C-terminal amidation of the peptide can reduce the conformational fluctuation, during the simulation when compared to the uncapped one.

In Chapter 5, a preliminary attempt to analyse the proteins that experimentally presented α -sheet, via MD simulations in explicit solvent, is presented. The goal was to inspect whether this type of secondary structure can enhance the peptide's aggregation propensity, especially in the presence of ions. The results show there is no significant effect on the α -sheet structure when ions are present or not. The detected α -sheets were only found at neutral pH and physiological temperature (310K), albeit with a short life-time, during the simulation. The dihedrals were seen to quickly convert to β -sheet angles, which is in agreement with the literature, that described this conformation as an intermediate towards a more stable secondary structure.

In this study, we modelled only one protein for MD simulations that was reported to exhibit α -sheet, although we need to use more than one protein from the PDB and also explore

dimeric and oligomeric species as they are phases within the aggregation process, but since we did not have sufficient time this will remain a task for future work.

Overall, there are some future goals based on this thesis. The most important are to extend the GFN2-xTB to other more complicated biological systems that contain metals, as this method illustrated its ability to reproduce results from DFT method. In addition, we should keep searching and exploring for appropriate computer approaches that can provide molecular-scale information on the interactions between proteins or peptides and ligands, including metal cations with high accuracy but less expense. In addition, evaluating different enhanced sampling techniques, such as metadynamics or replica-exchange molecular dynamics, could offer more insight into the structure and behaviour of disordered peptides, such as A β or different ones over time.

The peptide fragment of A β used in this study is the N-terminus part (A β 16), as this region was reported to contain the binding sites for metals. But we need to extend the study of the peptide to involve the full length, because the A β 16 fragment might not reflect the structural behaviour of the full length A β peptide (A β 40 and A β 42). The force fields chosen in simulation have effect on A β fragment structures. Therefore, in the future we will perform comparative studies between AMBER, GROMOS, CHARMM and OPLS force fields simulation on monomers, dimers, and trimers of full-length A β 40-42 combined with water models in the presence of metal ions, to study their effects on the structure of A β peptide and for better understanding of the conformational transitions from the oligomeric state to the fibrillar state, in detail and in a comparative way.

The bioavailable metal cations that will be under our investigation in the future, are mainly based on recent chemical literature, which shows that those exhibit strong binding to the A β peptide, such as Co $^{2+}$, Ni $^{2+}$, Al $^{+3}$, Mg $^{2+}$, Ca $^{2+}$, Mn $^{2+}$, Na $^+$, and K $^+$. Also, we shall seek to explore different protocols that can affect the peptide's aggregation, such as ligands' interaction, as they are directly related to drug designing, as well as peptide modulation via changing pH in the presence of metals.

In chapter 5 we examined the uncommon structure of α -sheets, that was linked to amyloid aggregation, as an intermediate state during the fibrillation process, on a full length protein that was taken directly from the literature. Our investigation focused on measuring the

dihedral angles with and without the presence of ions. However, in order to understand the aggregation pathway, we need to assess two chains and more oligomeric types (dimers and trimers), because the literature suggested the α -sheet forms through hydrogen bonding between adjacent α -strands. Also we will investigate different metal ions as they have a critical role in the induction of aggregation.

7 Appendix

Chapter 3:

Appendix A:

Table A1: Cartesian coordinates for LowMode and Stochastic conformations

Cartesian coordinates for LowMode conformations

cis1

N 3.2410 3.9190 -2.8210

H 4.1300 4.4090 -2.8560

H 3.0110 3.5220 -1.9230

C 2.1740 4.7260 -3.4290

H 2.2650 5.7640 -3.1060

H 1.1920 4.3360 -3.1600

C 2.3690 4.6570 -4.9400

O 2.0380 5.6050 -5.6370

N 2.9980 3.5660 -5.3790

Cu 3.4360 2.5120 -4.0080

C 3.5840 3.1770 -6.6660

H 4.5180 2.6500 -6.4640

C 4.0330 4.3560 -7.5290

O 3.3510 4.7770 -8.4550

C 2.6870 2.1320 -7.3270

H 1.6790 2.5290 -7.4560

C 2.6430 0.8830 -6.4740

H 3.0960 1.8770 -8.3060

N 2.9490 0.8290 -5.1110

C 2.8140 -0.4550 -4.7570

H 2.9690 -0.8400 -3.7610

C 2.3330 -0.3700 -6.9100

N 2.4460 -1.1990 -5.8140

H 2.0620 -0.6490 -7.9190

H 2.2840 -2.1970 -5.7980

N 5.2230 4.9100 -7.2440

H 5.4810 5.6270 -7.9040

C 6.2210 4.6130 -6.1880

H 6.6650 3.6510 -6.4470

C 5.6720 4.4200 -4.7620

O 5.0840 5.3550 -4.1820

O 5.5400 3.2550 -4.3040

C 7.3720 5.6420 -6.2740

H 8.1980 5.2240 -5.6970

H 7.7180 5.6960 -7.3080

C 7.1650 7.0850 -5.7540

H 6.7630 7.0720 -4.7430

H	8.1660	7.5080	-5.6560
C	6.3770	8.0930	-6.6190
H	6.5170	9.0760	-6.1670
H	6.8330	8.1310	-7.6090
C	4.8630	7.8940	-6.7990
H	4.6920	7.1160	-7.5400
H	4.4330	8.8130	-7.2010
N	4.1750	7.5330	-5.5430
H	4.1380	8.3170	-4.9120
H	4.6790	6.7650	-5.0780
H	3.2410	7.1740	-5.7380
O	4.2490	1.2250	-2.4290
H	4.6230	1.5200	-1.5990
H	4.0160	0.3090	-2.2750

cis2

N	2.8120	4.3160	-3.2460
H	3.5890	4.9470	-3.4270
H	2.6510	4.1190	-2.2700
C	1.6130	4.7250	-3.9900
H	1.5370	5.8130	-3.9990
H	0.7170	4.2930	-3.5440
C	1.7700	4.2310	-5.4230
O	1.0970	4.7200	-6.3110
N	2.6890	3.2850	-5.6020
Cu	3.3030	2.7170	-4.0510
C	3.3790	2.7710	-6.7890
H	4.4120	2.5520	-6.5140
C	3.5190	3.8200	-7.8910
O	2.8010	3.8110	-8.8830
C	2.7790	1.4180	-7.1680
H	1.7100	1.5220	-7.3600
C	3.0050	0.4270	-6.0470
H	3.2680	1.0490	-8.0710
N	3.2650	0.7650	-4.7150
C	3.4300	-0.3950	-4.0680
H	3.6380	-0.4900	-3.0130
C	3.0230	-0.9300	-6.1710
N	3.2920	-1.4280	-4.9150
H	2.8650	-1.4930	-7.0810
H	3.3770	-2.4040	-4.6630
N	4.4750	4.7520	-7.7300
H	4.4640	5.4350	-8.4720
C	5.5660	4.8810	-6.7250
H	6.2610	4.0520	-6.8630

C	5.1320	4.8430	-5.2450
O	4.5750	5.8260	-4.7110
O	5.2020	3.7630	-4.6070
C	6.3200	6.1810	-7.0560
H	6.6750	6.1410	-8.0870
H	5.6010	6.9960	-6.9940
C	7.5280	6.4530	-6.1370
H	7.3800	6.0110	-5.1530
H	8.4090	5.9750	-6.5670
C	7.8160	7.9510	-5.9400
H	8.6470	8.0480	-5.2390
H	8.1340	8.3760	-6.8930
C	6.6300	8.7790	-5.4130
H	5.9040	8.9020	-6.2200
H	6.9970	9.7700	-5.1380
N	5.9580	8.1530	-4.2510
H	6.6310	7.9460	-3.5290
H	5.5010	7.2710	-4.5310
H	5.2400	8.7620	-3.8890
O	4.3260	2.0590	-2.2140
H	4.8110	2.6200	-1.6080
H	4.6510	1.1770	-2.0360

cis3

N	1.8450	4.3810	-3.4360
H	2.2350	5.3130	-3.5300
H	1.6700	4.0930	-2.4860
C	0.7050	4.1880	-4.3430
H	0.1680	5.1290	-4.4680
H	0.0260	3.4330	-3.9460
C	1.2250	3.7370	-5.7050
O	0.5020	3.8140	-6.6800
N	2.4730	3.2790	-5.7320
Cu	3.1810	3.2650	-4.1050
C	3.3800	2.9020	-6.8180
H	4.4040	3.0110	-6.4650
C	3.3240	3.8820	-7.9890
O	2.6930	3.6380	-9.0090
C	3.2770	1.3990	-7.0500
H	2.2470	1.1180	-7.2740
C	3.7550	0.6950	-5.7930
H	3.9190	1.1180	-7.8860
N	3.8350	1.2800	-4.5190
C	4.3380	0.3370	-3.7150
H	4.5180	0.4600	-2.6570

C	4.2150	-0.5850	-5.7190
N	4.5750	-0.7920	-4.4050
H	4.2870	-1.2880	-6.5380
H	4.9530	-1.6460	-4.0160
N	4.0040	5.0360	-7.8350
H	3.8520	5.6800	-8.5980
C	5.0360	5.4160	-6.8260
H	5.9050	4.8010	-7.0540
C	4.6830	5.1260	-5.3580
O	3.6760	5.5830	-4.7860
O	5.2310	4.1650	-4.7740
C	5.4540	6.8830	-7.0460
H	5.4150	7.0940	-8.1160
H	4.7410	7.5490	-6.5550
C	6.8870	7.2150	-6.5820
H	7.5840	6.5240	-7.0590
H	7.1270	8.2100	-6.9590
C	7.1070	7.2330	-5.0550
H	7.4380	8.2350	-4.7760
H	6.1710	7.0660	-4.5260
C	8.1690	6.2440	-4.5490
H	8.2150	6.3220	-3.4610
H	9.1410	6.5390	-4.9520
N	7.8850	4.8420	-4.9310
H	7.9700	4.7420	-5.9320
H	6.9220	4.5840	-4.6840
H	8.5250	4.2080	-4.4760
O	4.2550	3.2830	-2.1320
H	4.3370	4.0330	-1.5430
H	4.9650	2.6960	-1.8730

trans1

N	1.9694	4.7632	-4.6623
H	1.0173	4.5238	-4.8909
H	2.2933	5.6043	-5.1272
C	2.2494	4.7697	-3.2253
H	1.5619	4.1068	-2.6990
H	2.1478	5.7830	-2.8340
C	3.6828	4.3006	-3.0057
O	4.2236	4.5202	-1.9366
N	4.2945	3.6819	-4.0159
Cu	3.0381	3.3942	-5.3232
C	5.7599	3.4692	-4.0597
H	6.1550	3.6393	-3.0554
C	6.4393	4.5401	-4.9218

O	7.0260	5.4596	-4.3641
C	6.1319	2.0251	-4.4329
H	7.2029	1.9809	-4.6350
C	5.3885	1.4023	-5.5887
H	5.9429	1.4024	-3.5579
N	4.1360	1.7847	-6.0536
C	3.7784	0.8649	-6.9592
H	2.8339	0.8387	-7.4810
C	5.7916	0.2948	-6.2742
N	4.7697	-0.0223	-7.1416
H	6.7090	-0.2568	-6.1146
H	4.7163	-0.8510	-7.7221
N	6.2222	4.5086	-6.2409
H	5.9552	3.6191	-6.6311
C	5.6372	5.6758	-6.9140
H	5.7456	6.5273	-6.2386
C	4.1402	5.4251	-6.9953
O	3.7475	4.2747	-7.3060
O	3.3790	6.1403	-6.3302
C	6.3588	6.1029	-8.2119
H	6.1680	7.1729	-8.3094
H	7.4367	6.0015	-8.0761
C	5.9453	5.4853	-9.5610
H	4.8831	5.2446	-9.5672
H	6.0894	6.2537	-10.321
C	6.7787	4.2749	-10.010
H	6.4668	3.9981	-11.018
H	7.8286	4.5670	-10.057
C	6.6561	3.0493	-9.1076
H	7.0862	3.2873	-8.1342
H	7.2386	2.2356	-9.5450
N	5.2487	2.6310	-8.9595
H	4.8384	2.4569	-9.8648
H	4.7058	3.3687	-8.4922
H	5.1888	1.8001	-8.3827
O	1.3195	2.8170	-6.6591
H	0.5731	3.3711	-6.8865
H	1.3391	2.1486	-7.3437

trans2

N	2.1530	4.3387	-2.6265
H	1.3182	3.9210	-2.2437
H	2.0033	5.2484	-3.0410
C	3.2995	4.3169	-1.6980
H	3.2553	3.4407	-1.0500

H	3.3082	5.2242	-1.0923
C	4.5670	4.2724	-2.5447
O	5.6322	4.6601	-2.1007
N	4.3698	3.8912	-3.7995
Cu	2.7160	3.2325	-4.0193
C	5.0860	4.2643	-5.0189
H	5.8825	4.9772	-4.7948
C	4.0370	4.9052	-5.9593
O	2.8652	4.9975	-5.5441
C	5.6664	2.9841	-5.6354
H	6.2538	3.2465	-6.5163
C	4.5908	1.9990	-6.0368
H	6.3336	2.5078	-4.9158
N	3.3047	1.9488	-5.4961
C	2.6738	0.9635	-6.1477
H	1.6519	0.6610	-5.9749
C	4.7059	1.0455	-7.0028
N	3.4894	0.4015	-7.0550
H	5.5778	0.8519	-7.6132
H	3.2379	-0.3478	-7.6855
N	4.4240	5.2724	-7.1797
H	5.3782	5.1160	-7.4887
C	3.5566	5.7137	-8.2743
H	2.5479	5.3349	-8.1165
C	4.0634	5.0845	-9.5793
O	5.2939	4.8843	-9.6691
O	3.2081	4.8070	-10.451
C	3.5130	7.2581	-8.2855
H	3.1664	7.5859	-7.3042
H	4.5174	7.6594	-8.4331
C	2.5554	7.8505	-9.3351
H	1.6775	7.2082	-9.4078
H	2.2124	8.8253	-8.9861
C	3.2219	8.0517	-10.712
H	3.6218	9.0653	-10.758
H	4.0681	7.3760	-10.829
C	2.2779	7.8413	-11.901
H	2.8503	8.0203	-12.814
H	1.4616	8.5650	-11.856
N	1.7391	6.4638	-11.930
H	0.9800	6.3740	-11.274
H	2.4532	5.8035	-11.582
H	1.4451	6.1932	-12.856
O	0.7171	2.5148	-4.2608
H	-0.0407	2.7743	-3.7372

H 0.4629 1.6795 -4.6526

Cartesian coordinates for Stochastic conformations

1

N -1.8370 1.9520 -2.3930

H -1.3630 2.5980 -1.7730

H -1.3390 1.7870 -3.2580

C -3.2360 2.3470 -2.6030

H -3.7350 1.6510 -3.2770

H -3.2970 3.3610 -2.9850

C -3.8120 2.2580 -1.2170

O -4.7900 2.9020 -0.8570

N -3.0380 1.4830 -0.4680

Cu -1.9450 0.2740 -1.4690

C -2.7560 1.6670 0.9400

H -3.2050 2.6000 1.2770

C -1.2320 1.7990 1.0100

O -0.7330 2.8980 0.7960

C -3.3460 0.5210 1.7670

H -4.4220 0.6680 1.8000

C -3.0880 -0.8570 1.2240

H -2.9740 0.5860 2.7880

N -2.5320 -1.1430 -0.0250

C -2.5310 -2.4820 -0.1000

H -2.1900 -3.0470 -0.9520

C -3.4000 -2.0180 1.8670

N -3.0340 -3.0350 1.0180

H -3.8560 -2.1170 2.8400

H -3.1590 -4.0230 1.1810

N -0.4960 0.7020 1.1890

H -0.9830 -0.1370 1.4660

C 0.8690 0.5200 0.7090

H 1.0130 1.1530 -0.1680

C 0.9970 -0.9440 0.2770

O 1.3900 -1.1520 -0.8930

O 0.6930 -1.8330 1.1050

C 1.8830 0.9490 1.7870

H 1.7210 0.3840 2.7110

H 1.7090 2.0110 1.9920

C 3.3560 0.8010 1.3380

H 3.4530 1.1740 0.3120

H 3.9690 1.4450 1.9740

C 3.9210 -0.6350 1.4350

H 3.1520 -1.3030 1.8320

H 4.7490 -0.6510 2.1480

C	4.4030	-1.2160	0.0940
H	3.6480	-1.0460	-0.6690
H	4.5510	-2.2920	0.2100
N	5.6780	-0.6240	-0.3490
H	5.9810	-1.0700	-1.2080
H	5.5580	0.3680	-0.5020
H	6.3750	-0.7730	0.3670
O	-0.9940	-1.1020	-2.5940
H	-1.3060	-1.9700	-2.8290
H	-0.1270	-1.0340	-2.9900

2

N	-0.5370	-2.6940	2.2570
H	-1.2770	-2.3450	2.8570
H	0.3870	-2.3980	2.5470
C	-0.5860	-4.1620	2.0220
H	-1.5860	-4.4700	1.7130
H	-0.2590	-4.7180	2.8990
C	0.4190	-4.3800	0.8990
O	1.0280	-5.4360	0.7390
N	0.6100	-3.2500	0.2320
Cu	-0.8100	-1.9740	0.4800
C	1.8550	-2.8800	-0.4550
H	2.6470	-3.5610	-0.1480
C	2.2890	-1.4430	-0.1010
O	3.1180	-0.8430	-0.7770
C	1.6250	-3.0630	-1.9560
H	1.4830	-4.1280	-2.1470
C	0.3990	-2.3280	-2.4450
H	2.5020	-2.7360	-2.5160
N	-0.6050	-1.7920	-1.6270
C	-1.5590	-1.3980	-2.4950
H	-2.5290	-1.0150	-2.2070
C	0.0360	-2.2320	-3.7570
N	-1.1960	-1.6290	-3.7700
H	0.5840	-2.6150	-4.6040
H	-1.7790	-1.4970	-4.5890
N	1.6980	-0.8970	0.9710
H	1.0770	-1.4880	1.4980
C	1.8340	0.4510	1.4860
H	2.8930	0.6910	1.5380
C	1.2230	0.4930	2.9000
O	-0.0220	0.3910	3.0000
O	2.0150	0.6240	3.8620
C	1.1240	1.4410	0.5440

H	0.0650	1.1950	0.4640
H	1.5690	1.3530	-0.4470
C	1.2760	2.8920	1.0270
H	2.3350	3.0940	1.1940
H	0.7390	3.0370	1.9640
C	0.7410	3.8760	-0.0170
H	-0.3530	3.8340	-0.0420
H	1.1320	3.6040	-1.0000
C	1.2040	5.2950	0.3160
H	0.7220	5.6360	1.2330
H	2.2820	5.2800	0.4680
N	0.9020	6.2290	-0.7810
H	1.3020	7.1350	-0.5750
H	-0.0980	6.3150	-0.8860
H	1.3000	5.8710	-1.6410
O	-2.2000	-0.5250	0.3780
H	-2.9200	-0.4860	-0.2500
H	-2.0810	0.3850	0.6570

3

N	-2.9160	2.1870	2.3410
H	-3.0490	1.2190	2.5940
H	-3.6350	2.7860	2.7210
C	-1.5650	2.6400	2.7010
H	-1.5530	3.7250	2.7430
H	-1.2870	2.2310	3.6730
C	-0.5900	2.1540	1.6400
O	0.6150	2.0810	1.8770
N	-1.1880	1.8440	0.4790
Cu	-3.0470	2.3040	0.4390
C	-0.6270	1.2290	-0.7330
H	-1.3070	0.4270	-1.0160
C	0.7300	0.5490	-0.5360
O	1.7780	1.0670	-0.9260
C	-0.6480	2.2180	-1.9100
H	-0.0150	1.8400	-2.7140
C	-2.0480	2.3690	-2.4460
H	-0.2680	3.1920	-1.5970
N	-3.1920	2.3840	-1.6510
C	-4.2200	2.4130	-2.5110
H	-5.2640	2.4070	-2.2290
C	-2.4090	2.4090	-3.7590
N	-3.7800	2.4330	-3.7780
H	-1.7540	2.3880	-4.6120
H	-4.3670	2.4210	-4.5970

N	0.6690	-0.6470	0.0640
H	-0.2310	-0.9720	0.3890
C	1.7640	-1.6010	0.1440
H	2.3470	-1.5200	-0.7780
C	1.1250	-2.9930	0.1870
O	0.9040	-3.5250	1.2980
O	0.8330	-3.4940	-0.9200
C	2.6980	-1.2650	1.3270
H	2.2560	-1.5740	2.2770
H	2.8120	-0.1820	1.3630
C	4.1140	-1.8510	1.1720
H	4.5010	-1.5850	0.1900
H	4.7470	-1.3430	1.9020
C	4.2770	-3.3640	1.4050
H	3.3740	-3.8090	1.8190
H	5.0500	-3.4920	2.1670
C	4.7490	-4.1510	0.1740
H	5.1660	-5.0960	0.5170
H	5.5300	-3.5840	-0.3320
N	3.6730	-4.4500	-0.7850
H	4.0340	-5.0340	-1.5270
H	2.9230	-4.9330	-0.3110
H	3.3150	-3.5890	-1.1760
O	-4.9650	1.8070	0.7400
H	-5.4200	1.0600	0.3430
H	-5.6710	2.4020	0.9960

4

N	3.9820	-0.3290	-0.7200
H	3.8690	-0.5150	0.2700
H	4.1890	-1.1640	-1.2510
C	4.9220	0.7730	-0.9380
H	5.0230	1.0030	-1.9990
H	5.8950	0.5370	-0.5050
C	4.2580	1.8730	-0.1750
O	4.8320	2.8750	0.2320
N	3.0090	1.4810	0.0630
Cu	2.2950	0.3740	-1.3070
C	2.2680	1.5990	1.3010
H	2.9620	1.7870	2.1240
C	1.5570	0.2380	1.5410
O	2.1570	-0.6580	2.1370
C	1.3180	2.7970	1.1730
H	0.6370	2.7900	2.0280
C	0.5380	2.7770	-0.1260

H	1.9070	3.7160	1.2100
N	0.8090	1.8910	-1.1840
C	-0.1700	2.1100	-2.0690
H	-0.3030	1.5410	-2.9800
C	-0.5610	3.5290	-0.4260
N	-0.9920	3.0910	-1.6540
H	-1.0260	4.2850	0.2010
H	-1.8170	3.4130	-2.1410
N	0.3150	0.0730	1.0450
H	-0.0990	0.8700	0.5810
C	-0.4170	-1.1520	0.8120
H	-0.2640	-1.8250	1.6550
C	0.1500	-1.7800	-0.4520
O	0.8260	-2.8220	-0.3050
O	-0.1020	-1.2100	-1.5390
C	-1.9240	-0.8610	0.6310
H	-2.3380	-1.6260	-0.0290
H	-2.0790	0.1060	0.1410
C	-2.6900	-0.9060	1.9650
H	-2.3630	-1.7780	2.5310
H	-2.4280	-0.0170	2.5440
C	-4.2260	-0.9730	1.8450
H	-4.5760	-0.0900	1.3140
H	-4.6500	-0.9410	2.8520
C	-4.7630	-2.2170	1.1200
H	-4.3400	-2.2450	0.1130
H	-5.8470	-2.1350	1.0370
N	-4.4330	-3.4800	1.7970
H	-3.4310	-3.5430	1.9350
H	-4.7410	-4.2610	1.2390
H	-4.8990	-3.5070	2.6950
O	2.7100	-1.3200	-2.3120
H	3.3400	-1.9890	-2.0360
H	2.1240	-1.7770	-2.9060

5

N	-2.6700	-1.7260	-1.5120
H	-2.5960	-0.7220	-1.6110
H	-2.3300	-2.2410	-2.3190
C	-4.0190	-2.1410	-1.1000
H	-4.0310	-3.2080	-0.8690
H	-4.7650	-1.8820	-1.8570
C	-4.1930	-1.3480	0.1650
O	-5.2600	-0.8500	0.5160
N	-2.9880	-1.1130	0.6840

Cu	-1.5930	-2.2370	-0.0040
C	-2.6330	0.1350	1.3690
H	-3.5040	0.7890	1.3670
C	-1.5610	0.8380	0.5040
O	-1.9700	1.5350	-0.4260
C	-2.2560	-0.0930	2.8460
H	-3.1770	-0.0530	3.4440
C	-1.6010	-1.4190	3.0940
H	-1.6050	0.7140	3.1940
N	-1.2420	-2.2880	2.0710
C	-0.8220	-3.3960	2.7040
H	-0.5160	-4.2960	2.1920
C	-1.3750	-2.0040	4.3150
N	-0.8780	-3.2650	4.0410
H	-1.5830	-1.5780	5.2950
H	-0.6680	-4.0150	4.6940
N	-0.2430	0.6470	0.7390
H	0.0440	0.0910	1.5350
C	0.8210	1.0970	-0.1710
H	0.4480	1.0210	-1.1960
C	2.0040	0.1390	-0.0130
O	2.5500	0.0650	1.1120
O	2.3380	-0.5260	-1.0210
C	1.2310	2.5710	0.0850
H	1.9810	2.6210	0.8710
H	0.3550	3.1120	0.4490
C	1.7260	3.3420	-1.1640
H	0.8780	3.4140	-1.8480
H	1.9770	4.3590	-0.8600
C	2.9100	2.7830	-1.9830
H	2.9000	3.3040	-2.9420
H	2.7540	1.7290	-2.2020
C	4.3180	2.9950	-1.4020
H	4.4750	4.0580	-1.2140
H	5.0450	2.6620	-2.1440
N	4.5570	2.2420	-0.1610
H	5.5180	2.3640	0.1310
H	3.9400	2.5850	0.5620
H	4.3770	1.2610	-0.3210
O	-0.0520	-3.3540	-0.6340
H	0.4020	-3.2630	-1.4730
H	0.5660	-3.8290	-0.0780
6			
N	0.3810	1.7680	4.1910

H	0.0130	2.4700	4.8170
H	1.2560	2.0280	3.7620
C	0.4310	0.4580	4.8420
H	-0.3870	0.3750	5.5580
H	1.3820	0.3430	5.3610
C	0.2920	-0.6180	3.7760
O	0.6140	-1.7810	4.0270
N	-0.1940	-0.1940	2.6000
Cu	-0.8850	1.5820	2.7860
C	-0.2530	-1.0470	1.3910
H	-0.8250	-1.9340	1.6590
C	1.1530	-1.5370	0.9490
O	2.1470	-1.4170	1.6670
C	-0.9680	-0.4120	0.1690
H	-0.2450	0.2150	-0.3550
C	-2.1960	0.4360	0.3510
H	-1.2460	-1.2270	-0.5020
N	-2.3420	1.4190	1.3100
C	-3.4840	2.0490	1.0250
H	-3.8800	2.8770	1.5920
C	-3.2590	0.4930	-0.5060
N	-4.0670	1.5110	-0.0590
H	-3.4240	-0.1100	-1.3870
H	-4.9240	1.8200	-0.4950
N	1.2250	-2.0480	-0.2930
H	0.3590	-2.1710	-0.7980
C	2.4380	-2.2370	-1.0840
H	2.9860	-3.0680	-0.6370
C	3.3810	-1.0170	-1.0580
O	2.9090	0.1170	-0.8140
O	4.5870	-1.2480	-1.2870
C	2.0870	-2.6520	-2.5330
H	1.8390	-3.7140	-2.5110
H	2.9810	-2.5580	-3.1500
C	0.9140	-1.9450	-3.2520
H	-0.0130	-2.2070	-2.7410
H	0.8410	-2.3660	-4.2560
C	0.9640	-0.4130	-3.3740
H	0.9050	0.0200	-2.3790
H	0.0820	-0.0760	-3.9210
C	2.2170	0.1320	-4.0720
H	2.2650	1.2100	-3.9170
H	3.1010	-0.3250	-3.6270
N	2.2060	-0.1350	-5.5230
H	1.4370	0.3620	-5.9540

H	3.0780	0.1830	-5.9260
H	2.1020	-1.1260	-5.6870
O	-1.2640	3.4780	3.2940
H	-1.8840	4.0580	2.8530
H	-0.6540	4.0750	3.7340

7

N	3.4420	-0.3350	2.0870
H	3.3430	-0.9800	1.3080
H	4.2190	0.3060	1.9930
C	3.4180	-1.0590	3.3630
H	4.2550	-1.7580	3.4420
H	3.4110	-0.3380	4.1780
C	2.0840	-1.7510	3.2640
O	1.8300	-2.8200	3.8160
N	1.3130	-1.0930	2.3950
Cu	1.8470	0.7190	2.1210
C	0.2690	-1.6700	1.5520
H	0.3390	-2.7580	1.6010
C	0.4260	-1.2400	0.0690
O	0.6250	-2.1030	-0.7820
C	-1.0930	-1.2560	2.1150
H	-1.2350	-1.7380	3.0820
C	-1.2550	0.2340	2.2910
H	-1.8760	-1.6190	1.4490
N	-0.2040	1.1590	2.3750
C	-0.8060	2.3410	2.5540
H	-0.3010	3.2910	2.6450
C	-2.4500	0.8660	2.4250
N	-2.1510	2.1960	2.5850
H	-3.4310	0.4200	2.4280
H	-2.8380	2.9390	2.6910
N	0.3500	0.0660	-0.2430
H	0.1870	0.7090	0.5140
C	0.5040	0.7340	-1.5350
H	1.2430	0.2090	-2.1410
C	1.0240	2.1540	-1.2730
O	0.3910	2.8700	-0.4600
O	2.0450	2.5090	-1.9010
C	-0.8490	0.7860	-2.2790
H	-0.9130	1.7200	-2.8430
H	-1.6660	0.7920	-1.5550
C	-1.0180	-0.3750	-3.2760
H	-0.2430	-1.1280	-3.1220
H	-0.8930	0.0240	-4.2830

C -2.3940 -1.0490 -3.1550
 H -2.4540 -1.5540 -2.1880
 H -3.1860 -0.3000 -3.2040
 C -2.6090 -2.0890 -4.2650
 H -3.4640 -2.7060 -4.0070
 H -1.7240 -2.7190 -4.3380
 N -2.8690 -1.4530 -5.5730
 H -3.7150 -0.9020 -5.5170
 H -2.0930 -0.8560 -5.8170
 H -2.9800 -2.1680 -6.2780
 O 2.7210 2.4790 1.7400
 H 3.0370 3.1130 2.3840
 H 3.1640 2.7230 0.9250

Table A2: Rank correlation of relative energies between GFN2-xTB and B3LYP-D2/def2-SVP at LFMM geometries on low mode conformers.

Methods Compared	Spearman's coefficient (r_s)	p (2-tailed)
B3LYP-D2-SP vs GFN2-xTB-SP	0.9	0.04
B3LYP-D2-OPT vs GFN2-xTB-OPT	0.8	0.1
GFN2-xTB -OPT vs B3LYP-D2-SP	0.7	0.07

Chapter 5:

Appendix B: Dihedral angles of model peptides that were optimized via DFT and xTB.

Table B1: DFT optimised dihedral angles of central peptides and terminus (°)

DFT		α -sheet	β -strand	α -helices
1.1B0P	end	-66.5, -17.1	-161.8, 171.7	-58.3, -37.0
	$(\phi, \psi)_1$	66.3, 31.3	-117.6, 110.5	-73.7, -19.9
	$(\phi, \psi)_2$	-97.9, -2.9	-161.3, 140.4	-70.4, -42.3
	end	73.9, 19.3	-60.6, 142.4	-58.6, -40.5
2. IFEB	end	-61.9, -41.3	-112.3, -162.9	-54.2, -38.6
	$(\phi, \psi)_1$	49.6, 45.6	-55.8, -28.6	-73.9, -19.9
	$(\phi, \psi)_2$	-54.3, -44.9	-95.2, 124.3	-80.9, -33.7
	end	61.2, 42.3	-161.6, 149.6	-66.8, -41.2
3.1F9B	end	-95.4, -10.8	-66.8, 145.2	-61.5, -37.3

	$(\phi, \psi)_1$	56.3, 48.7	-147.7, 102.2	-59.0, -27.9
	$(\phi, \psi)_2$	-59.8, -31.9	-80.5, 71.5	-86.5, -45.5
	end	58.9, 43.5	-161.9, 162.3	-100.3, -19.5
4.1A05	end	-69.1, -36.5	-162.5, 170.2	-68.3, -19.4
	$(\phi, \psi)_1$	55.1, 43.1	-66.0, 170.9	-63.9, -18.5
	$(\phi, \psi)_2$	-71.8, -26.3	-109.6, 169.3	-107.8, -44.9
	end	75.5, 20.0	-161.7, 172.0	-125.6, 13.3
5.1GQF	end	-49.5, -41.7	-82.8, 73.4	-65.8, -28.7
	$(\phi, \psi)_1$	55.9, 49.1	-52.2, 123.0	-69.9, -56.1
	$(\phi, \psi)_2$	-76.3, -12.2	-84.6, 51.5	-62.4, -27.6
	end	86.6, 175.1	-67.7, 152.1	-87.5, 14.9
6.1QND	end	-72.6, -26.9	-157.9, 140.2	-61.2, -30.3
	$(\phi, \psi)_1$	62.2, 40.9	-62.6, 141.7	-61.1, -32.7
	$(\phi, \psi)_2$	-52.3, -42.9	-142.1, 160.0	-75.4, -30.3
	end	62.1, 38.8	-124.5, 161.9	-63.2, -38.5
7.1BVH	end	-64.5, -34.5	-83.4, -175.0	-59.1, -39.8
	$(\phi, \psi)_1$	91.8, 170.5	-100.2, -12.1	-65.4, -27.5
	$(\phi, \psi)_2$	-63.1, -38.6	-89.2, 74.1	-68.1, -39.9
	end	62.0, 33.1	-119.2, 136.2	-62.9, -39.2
8.1IDM	end	-113.4, 8.7	-136.2, 136.6	-58.3, -38.3
	$(\phi, \psi)_1$	56.4, 37.9	-134.7, 102.2	-61.4, -38.3
	$(\phi, \psi)_2$	-72.5, -28.9	-138.0, 154.3	-67.6, -41.5
	end	60.4, 35.9	-160.7, 172.9	-58.6, -39.9
9.1POI	end	-65.2, -32.4	-82.8, 66.2	-52.3, -51.4
	$(\phi, \psi)_1$	61.9, 34.1	-142.7, 149.3	-68.3, -26.3
	$(\phi, \psi)_2$	-68.5, -28.8	-128.2, 137.5	-64.6, -38.1
	end	62.5, 36.4	-146.3, 155.2	-67.4, -39.8
10. 1BWU	end	-72.9, -27.8	-120.0, 113.0	-64.8, -19.0
	$(\phi, \psi)_1$	55.0, 36.3	-120.0, 113.0	-64.3, -14.6
	$(\phi, \psi)_2$	-65.7, -33.3	-120.0, 113.0	-98.9, 0.7
	end	46.5, 48.7	-120.0, 113.0	-115.6, -54.1
11. 4AAH	end	-61.0, -26.4	-141.3, 154.4	-61.8, -21.7
	$(\phi, \psi)_1$	62.8, 35.5	-160.3, 171.3	-55.2, -29.0
	$(\phi, \psi)_2$	-83.3, -15.5	-92.5, 64.6	-103.8, -38.0
	end	61.4, 35.8	-76.4, 164.8	-109.4, -0.2
12. 1HLU	end	66.6, 23.9	-104.3, 136.9	-104.3, 136.9
	$(\phi, \psi)_1$	-58.6, -38.9	-130.9, -124.1	-130.9, -124.1
	$(\phi, \psi)_2$	64.5, 32.5	-85.1, 64.5	-85.1, 64.5

	end	-76.9, -19.5	-176.7, 170.1	-167.7, 170.0
13. 1B7T	end	-64.3, -27.3	-160.7, 151.3	-60.9, -31.1
	$(\phi, \psi)_1$	94.0, 23.9	-83.7, 55.8	-57.1, -29.3
	$(\phi, \psi)_2$	-75.9, -6.9	145.5, 167.8	-73.7, -3.2
	end	64.1, 33.8	-135.9, 159.6	-85.2, -9.9
14. 1E9T	end	66.67, 93.83	136.5, 139.3	-55.9, -42.3
	$(\phi, \psi)_1$	-73.49, -22.66	-160.0, 158.3	-60.7, -37.4
	$(\phi, \psi)_2$	50.94, 102.5	-91.2, 67.5	-66.0, -38.2
	end	-89.99, -3.98	153.9, -165.7	-60.8, -36.9

Table B2 : GFN2-xTB optimised dihedral angles of central peptides and terminus ($^{\circ}$)

XTB		α -sheet	β -strand	α -helices
	end	-55.9, -42.2	-165.7, 162.6	-49.6, -41.7
1.1B0P	$(\phi, \psi)_1$	43.8, 54.1	-125.6, 149.7	-58.6, -34.6
	$(\phi, \psi)_2$	-57.0, -44.9	-130.3, 87.2	-84.1, -39.8
	end	77.7, 21.1	-82.8, 178.1	-52.2, -43.2
	end	-73.5, -29.9	-148.4, 166.1	-39.9, -55.4
2. IFEB	$(\phi, \psi)_1$	62.6, -172.9	-128.9, 159.1	-86.0, 3.6
	$(\phi, \psi)_2$	-157.1, -59.5	-131.1, 168.4	-92.9, -32.8
	end	37.4, 60.1	-153.0, 154.6	-84.0, -43.5
	end	-44.5, 126.8	-117.0, 149.2	-40.7, -56.6
3.1F9B	$(\phi, \psi)_1$	-80.2, 67.5	-132.0, 37.1	-58.7, -37.5
	$(\phi, \psi)_2$	20.9, 69.9	-85.1, 78.1	-86.5, -36.9
	end	-86.0, -22.9	-97.1, -150.4	-70.3, -22.5
	end	-50.9, -51.0	-130.4, 161.6	-46.0, -48.2
4.1a05	$(\phi, \psi)_1$	38.8, 59.3	-71.1, 169.5	-70.3, -20.1
	$(\phi, \psi)_2$	-72.1, -29.7	-116.9, 163.7	-80.6, -41.7
	end	85.7, -64.4	-113.5, 171.1	-64.9, -44.7
	end	-45.4, -47.0	-84.9, 78.7	-58.3, -38.7
5.1GQF	$(\phi, \psi)_1$	49.0, 53.7	-104.4, 93.4	-65.3, -62.2
	$(\phi, \psi)_2$	-56.6, -39.8	-77.2, 172.9	-57.2, -32.9
	end	-123.5, 171.5	-144.1, 175.7	-95.5, 29.5
	end	-96.3, -16.2	-128.0, 142.1	-46.2, -46.3
6.1QND	$(\phi, \psi)_1$	44.9, 51.8	-121.2, 163.7	-59.3, -19.9
	$(\phi, \psi)_2$	-51.4, 45.2	-137.9, 158.1	-101.0, 2.4
	end	57.2, 45.1	-127.2, 163.1	-113.2, -47.5

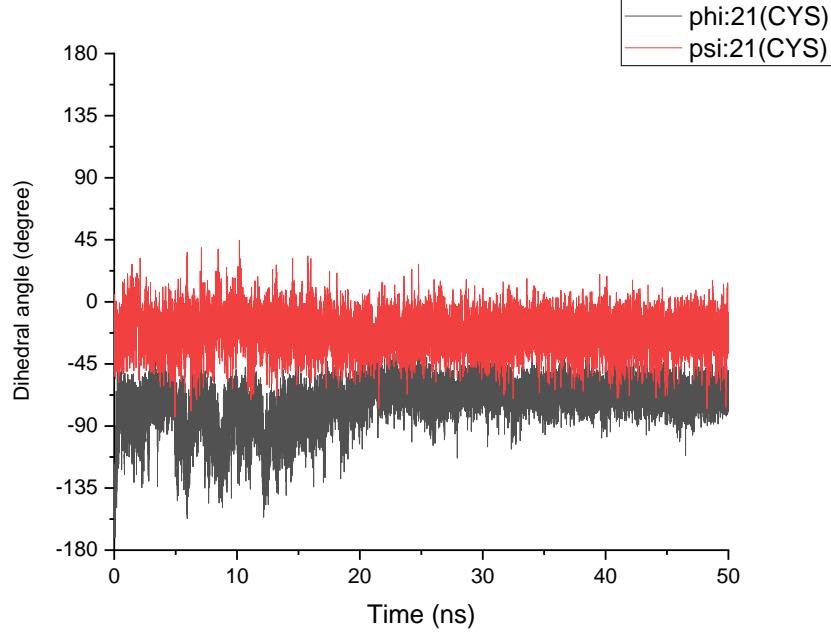
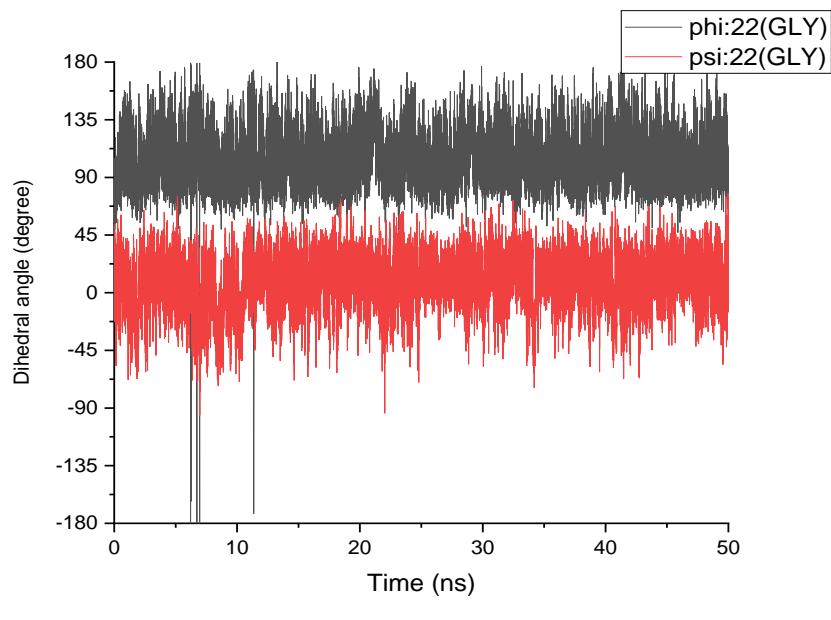
	end	-50.3, -45.9	-79.9, 65.8	-37.5, -60.4
7.1BVH	$(\phi,\psi)_1$	93.6, 10.7	-78.8, 57.8	-65.8, -12.6
	$(\phi,\psi)_2$	-35.7, -42.5	-50.7, 132.1	-88.2, -15.7
	end	48.2, 50.5	-151.5, 165.3	-103.8, -50.6
	end	-79.6, 61.7	-132.2, 157.4	-50.4, -38.0
8.1IDM	$(\phi,\psi)_1$	-74.5, 94.8	-131.6, 157.7	-45.5, -42.2
	$(\phi,\psi)_2$	-67.2, -14.1	-131.2, 165.1	-84.4, -44.5
	end	57.3, 43.0	-153.5, 170.3	-40.7, -54.1
	end	-66.6, -39.3	168.0, 168.5	-51.9, -45.8
9.1POI	$(\phi,\psi)_1$	56.5, 41.9	-132.3, 79.3	-55.8, -36.4
	$(\phi,\psi)_2$	-79.6, -25.4	-62.3, 82.5	-70.5, -6.1
	end	62.1, 42.9	-140.3, 160.6	-97.9, -54.5
	end	-94.6, -13.4	-146.4, 160.7	-61.1-31.6
10. 1BWU	$(\phi,\psi)_1$	46.3, 36.8	-92.8, 168.6	-43.7, -39.4
	$(\phi,\psi)_2$	-99.2, 61.9	-125.6, 167.4	-80.8, -19.9
	end	-108.3, 151.0	-124.4, 143.2	-82.4, -50.2
	end	-54.4, -32.6	-133.9, 158.8	-50.4, -42.4
11. 4AAH	$(\phi,\psi)_1$	39.4, 61.1	-151.8, 171.2	-48.3, -38.6
	$(\phi,\psi)_2$	-85.1, 58.7	-140.2, 151.1	-93.3, -43.6
	end	45.3, 50.9	-71.1, 165.5	-61.8, -41.6
	end	89.9, 4.4	-116.9, 163.5	-36.3, -54.0
12. 1HIU	$(\phi,\psi)_1$	-54.3, -46.9	-138.3, 82.4	-42.1, -44.0
	$(\phi,\psi)_2$	39.1, 59.5	-83.4, 71.8	-78.3, -22.8
	end	-57.8, -41.9	-85.8, 167.5	-86.7, -29.8
	end	-83.8, 53.2	-106.4, 125.4	-46.5, -42.2
13. 1B7T	$(\phi,\psi)_1$	34.1, 55.9	-131.1, 161.6	-55.9, -16.9
	$(\phi,\psi)_2$	-95.7, 0.3	-124.2, 162.2	-80.3, -13.8
	end	47.2, 51.7	-130.8, 159.9	-119.3, -47.4
	end	-49.1, -46.6	-90.9, 138.9	-41.4, -50.6
14. 1E9T	$(\phi,\psi)_1$	64.2, -156.8	-117.5, 150.6	-27.8, -51.7
	$(\phi,\psi)_2$	-108.2, -52.8	-141.8, 54.0	-77.1, -16.5
	end	114.7, -35.7	-85.6, 69.3	-92.4, -52.3

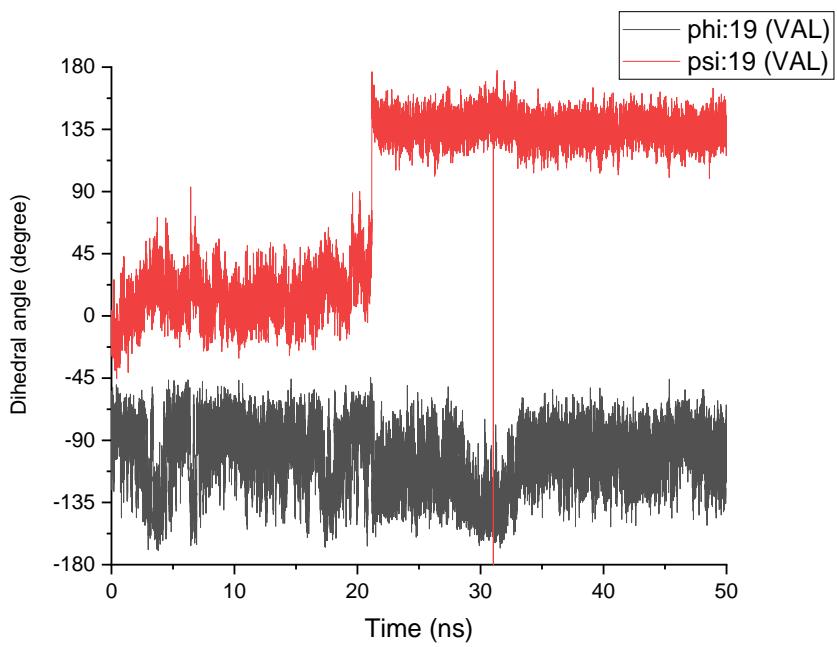
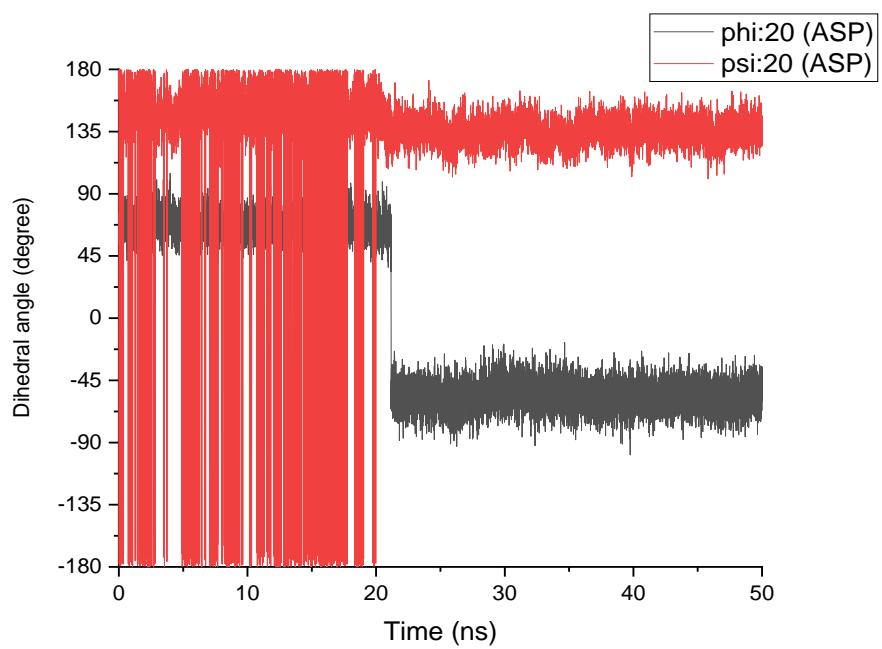
Table B3: DFT and GFN2-xTB relative electronic energy of α -helix and α -sheet to β -strand for peptides in water (kJ mol⁻¹)

#	PDB code	Method	α -sheet	α -helix
1	1B0P	XTB	52.9	-12.0
		DFT	1.2	75.1
2	1FEB	XTB	-21.7	-42.3
		DFT	22.9	-12.6
3	1F9B	XTB	-19.29	-31.12
		DFT	-24.77	-31.14
4	1A05	XTB	9.25	-33.44
		DFT	-5.12	-88.18
5	1GQF	XTB	-4.04	-31.99
		DFT	65.74	-8.09
6	1QND	XTB	33.93	6.32
		DFT	23.12	-10.54
7	1BVH	XTB	21.22	-18.89
		DFT	15.78	-21.28
8	1IDM	XTB	-23.429	-49.6
		DFT	-23.433	-53.0
9	1P0I	XTB	7.85	-14.42
		DFT	-32.87	-76.57
10	1BWU	XTB	-2.25	-30.88
		DFT	-13.16	-81.34
11	4AAH	XTB	4.20	0.43
		DFT	-39.10	-104.35
12	1HIU	XTB	15.50	-16.56
		DFT	15.70	-24.96
13	1B7T	XTB	26.90	-40.31
		DFT	-105.45	-80.7
14	1E9T	XTB	17.29	1.26
		DFT	-14.55	-12.49

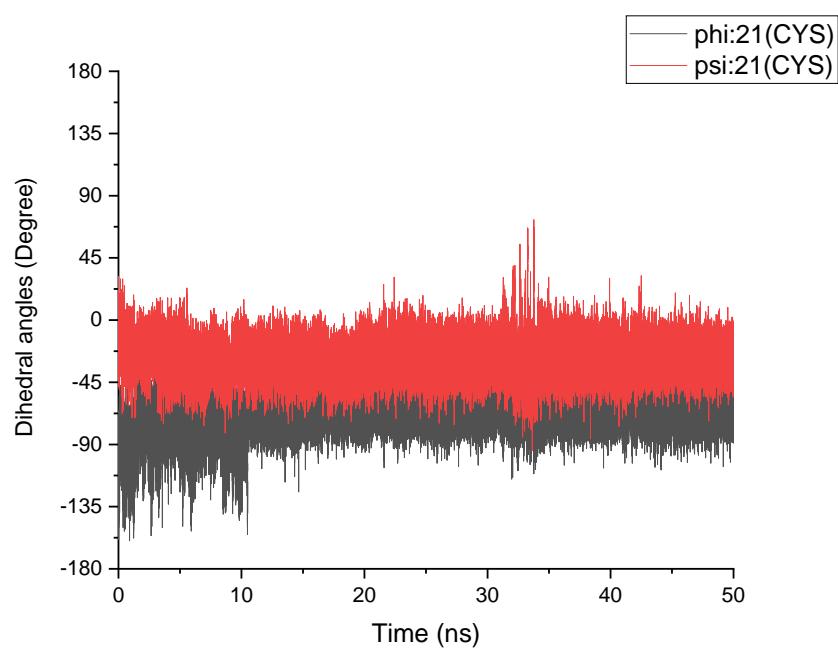
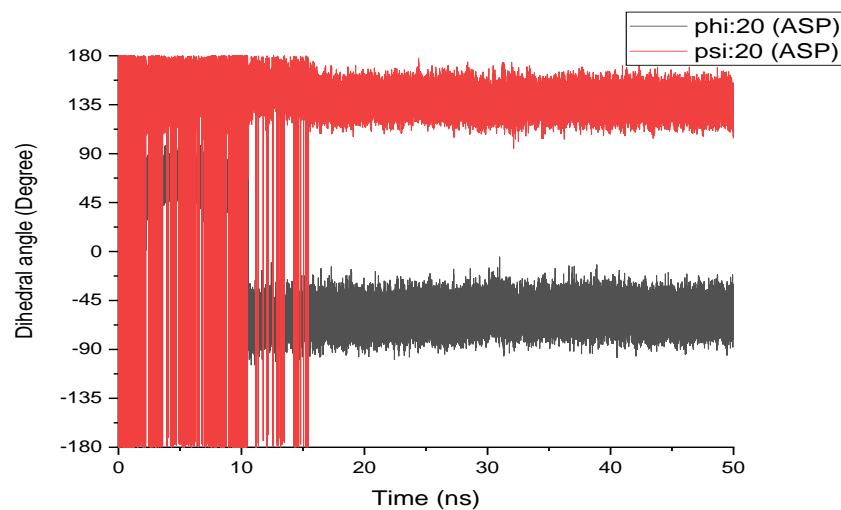
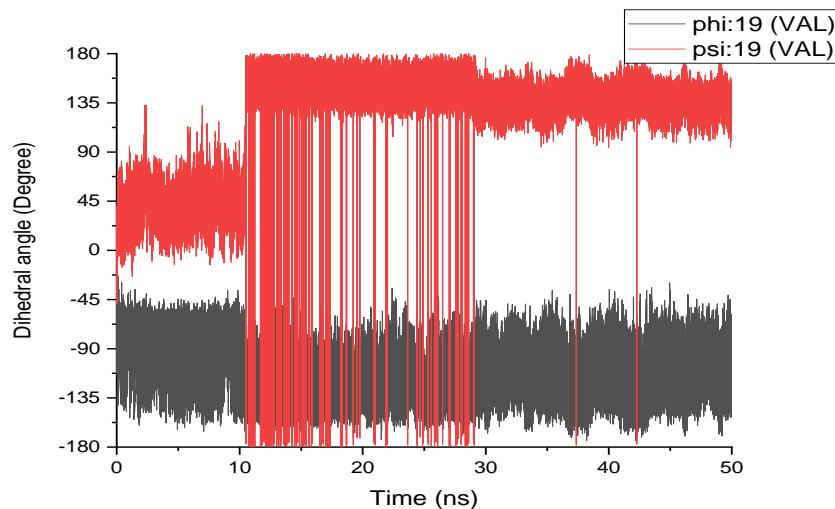
Appendix C: Dihedral angle for individual residues (VAL19, ASP20, CYS21, GLY21) at Neutral-, Medium-, Low-pH:

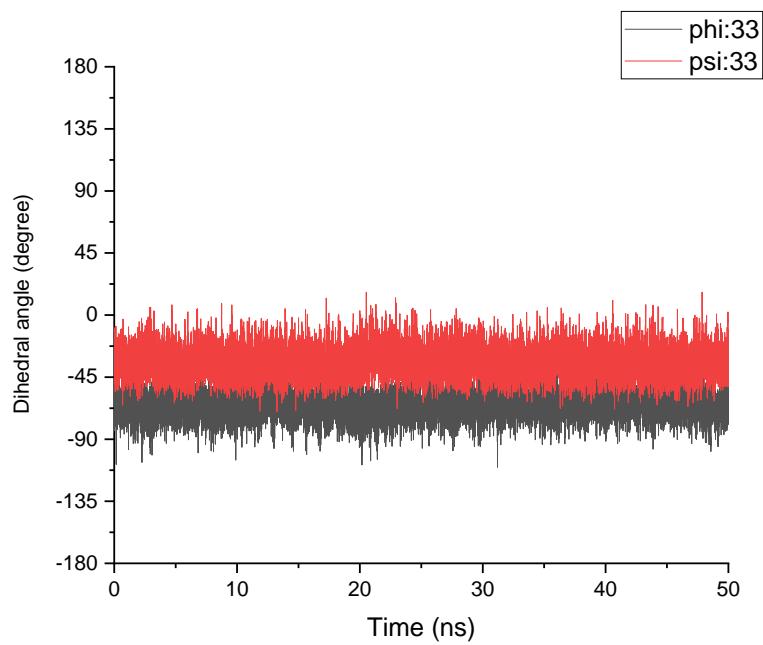
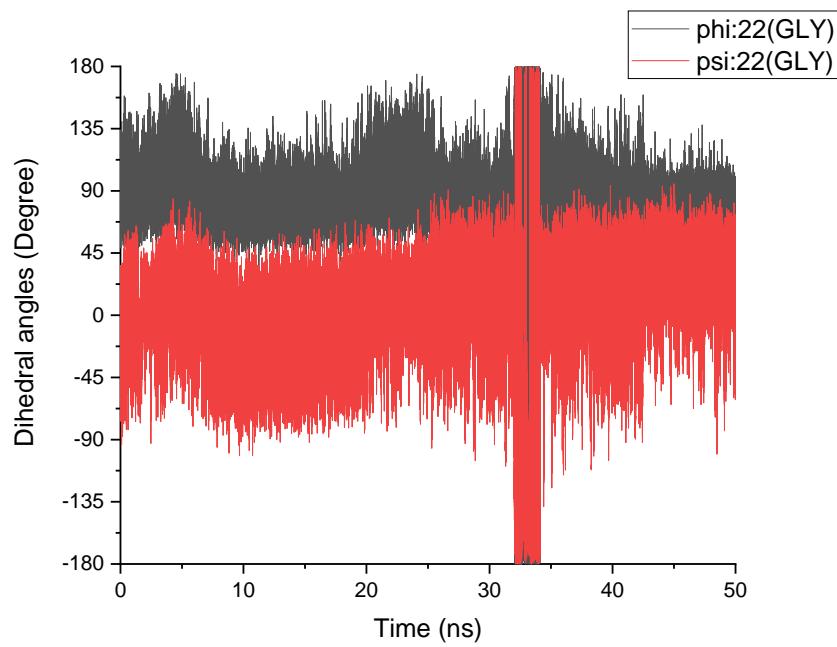
1-Neutral pH at 310 K

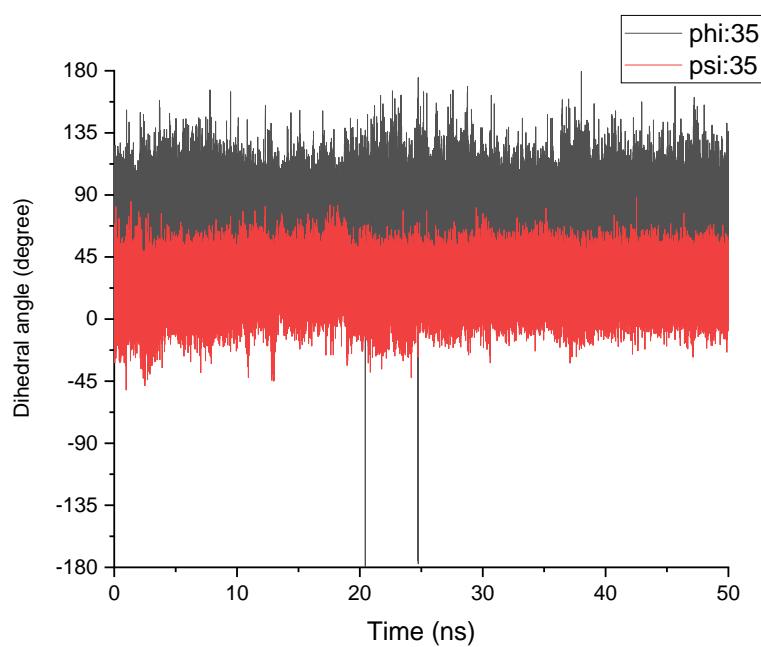
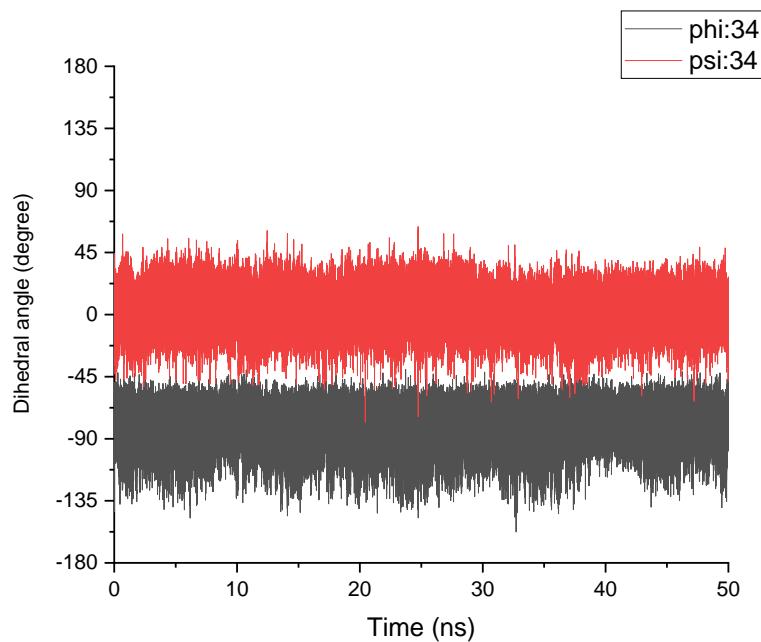




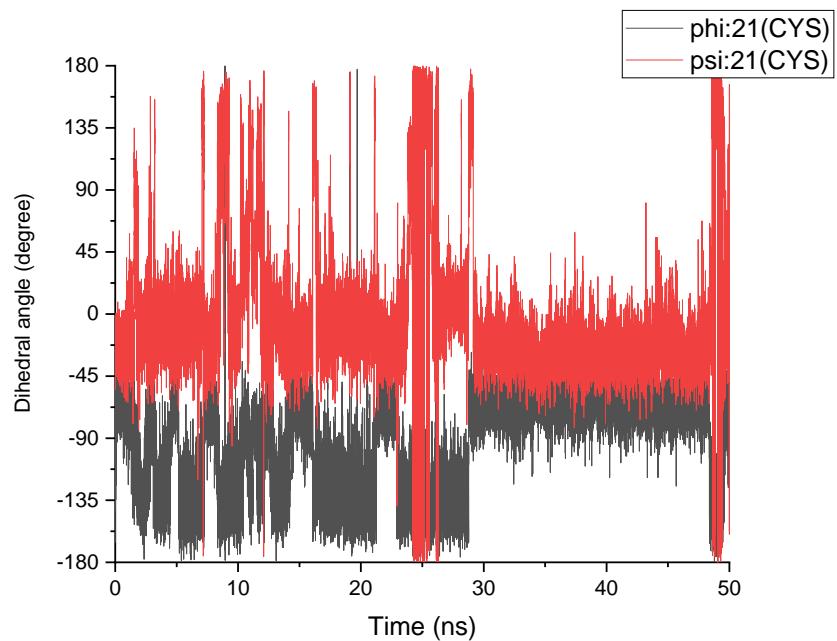
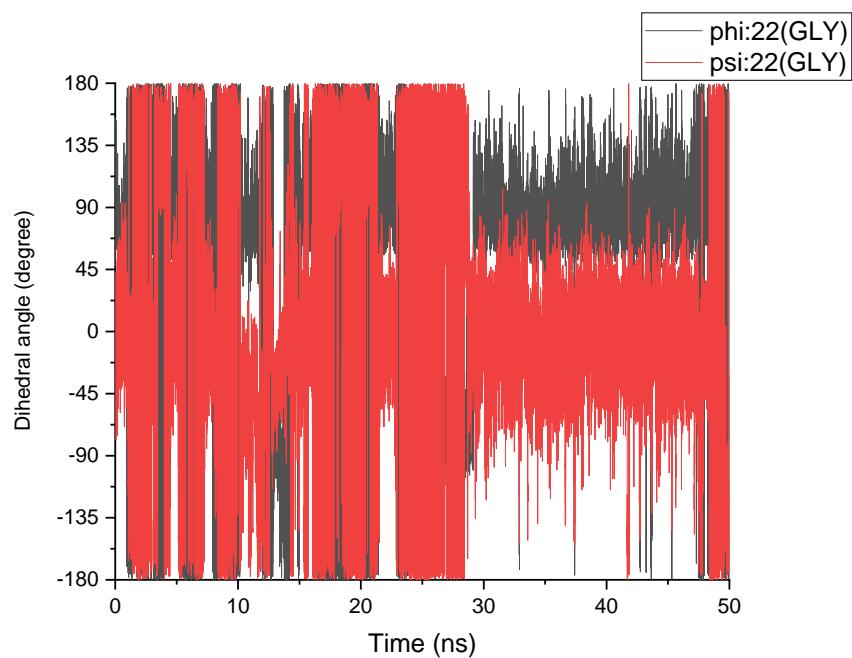
2- Neutral pH at 310k in presence of ions:

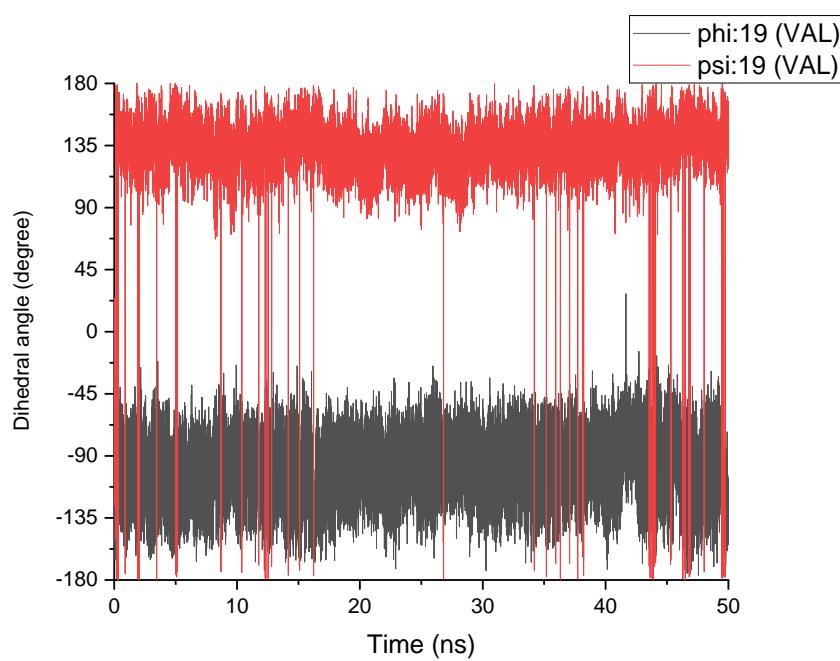
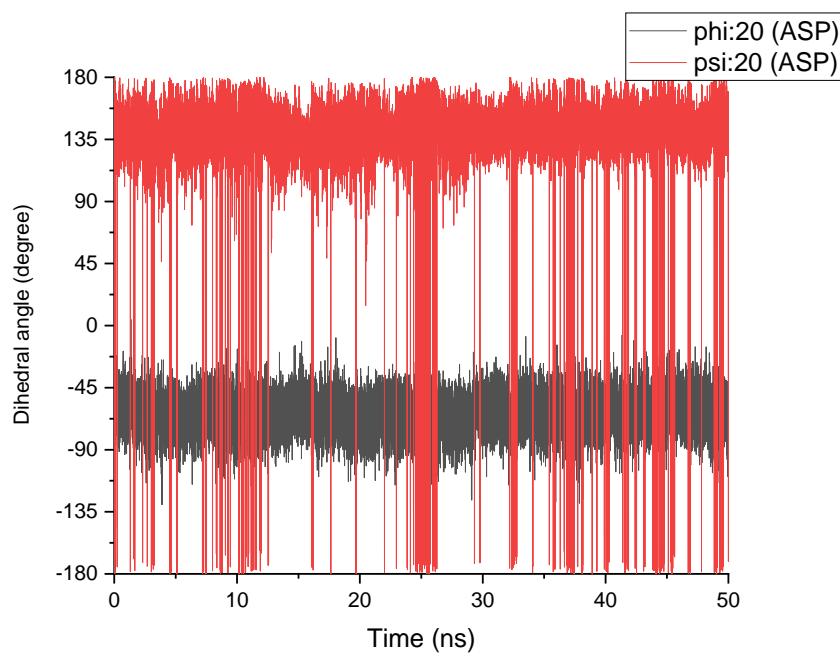




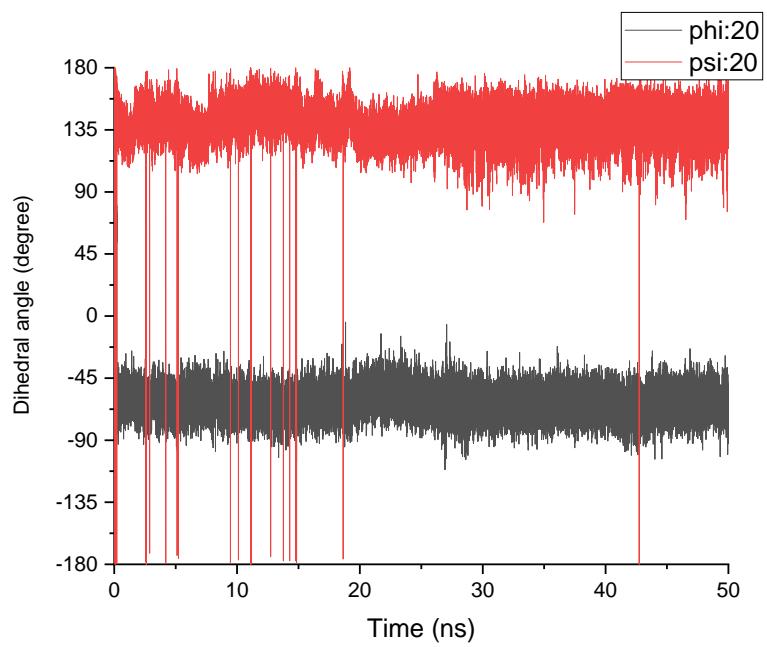
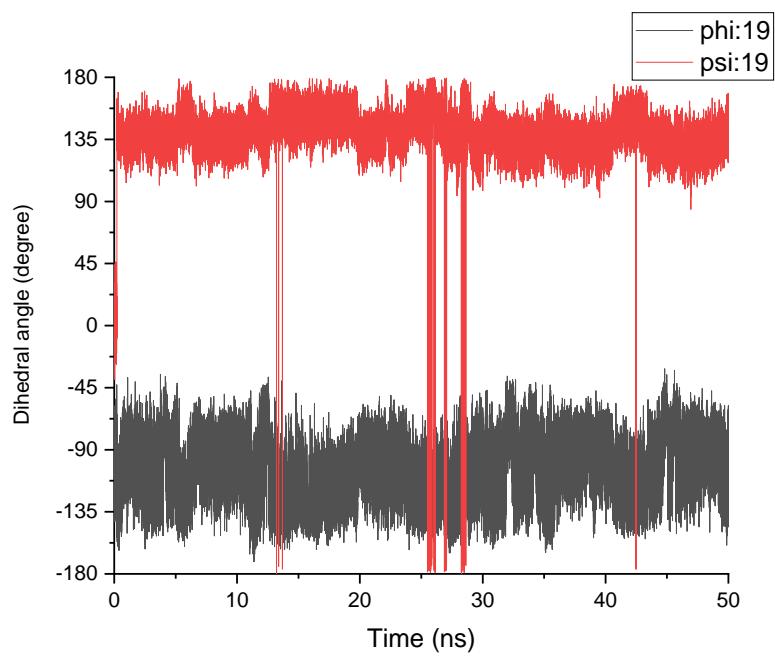


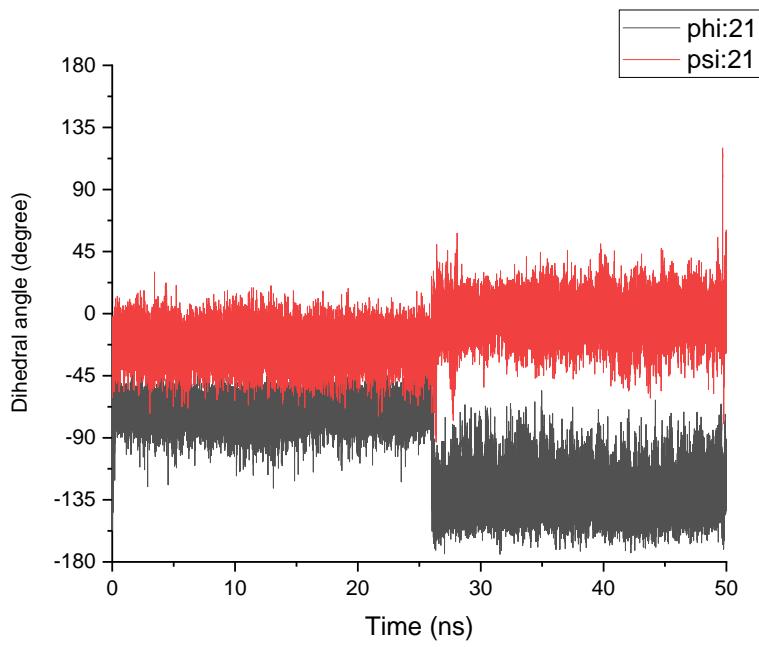
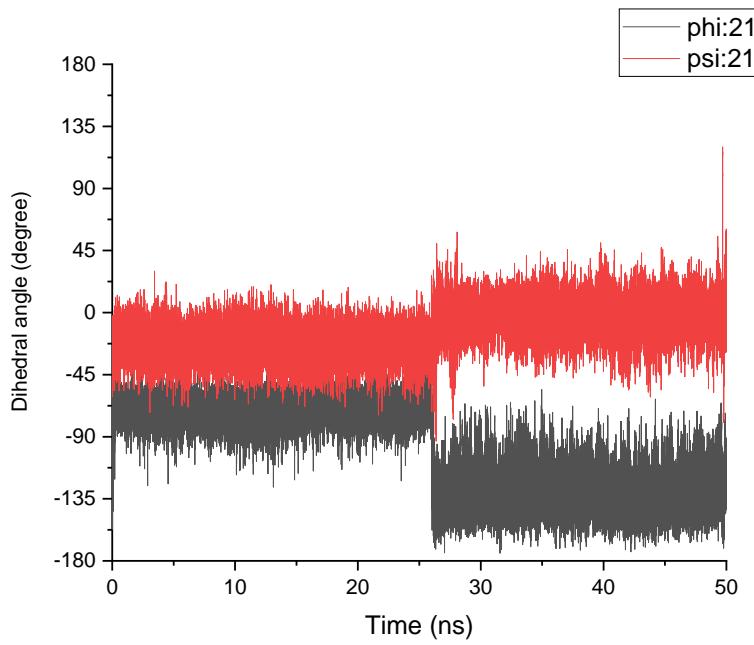
3- Neutral pH at 498K:

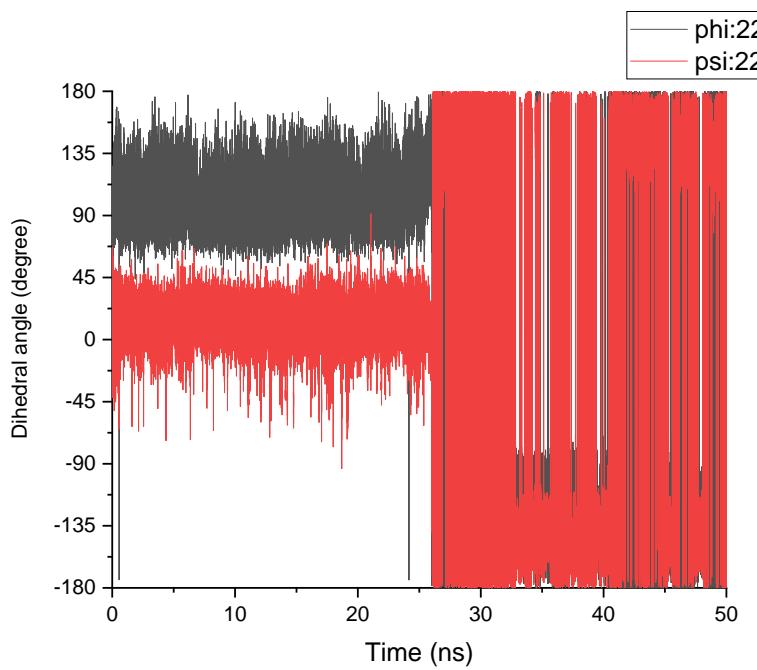




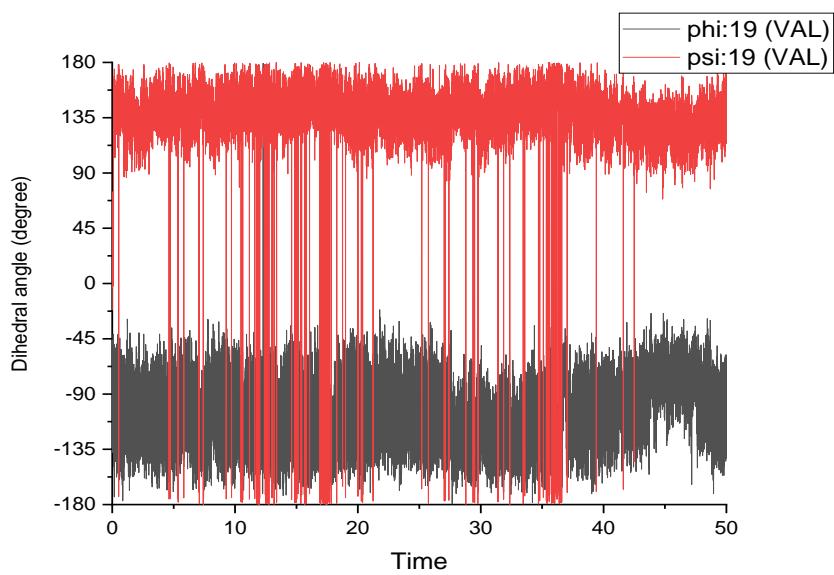
4- Medium pH at 310K:

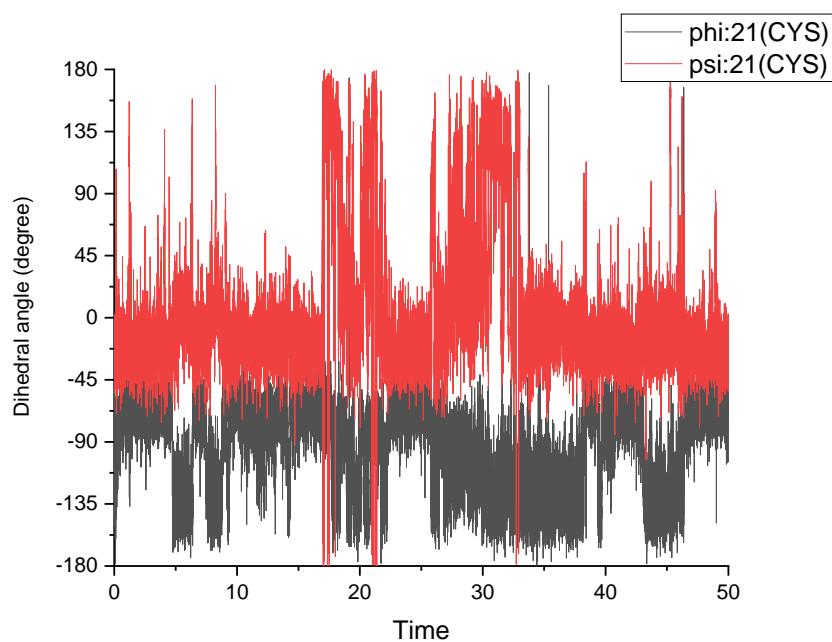
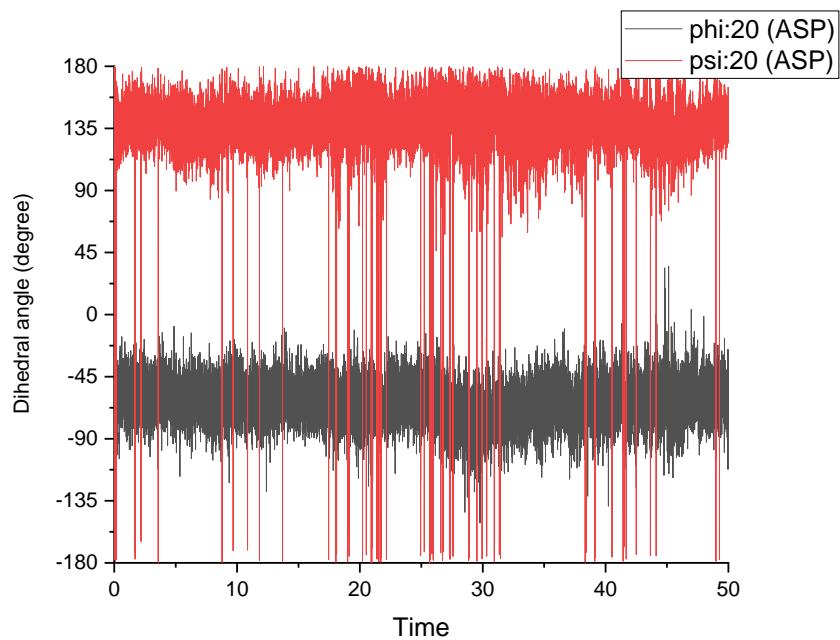


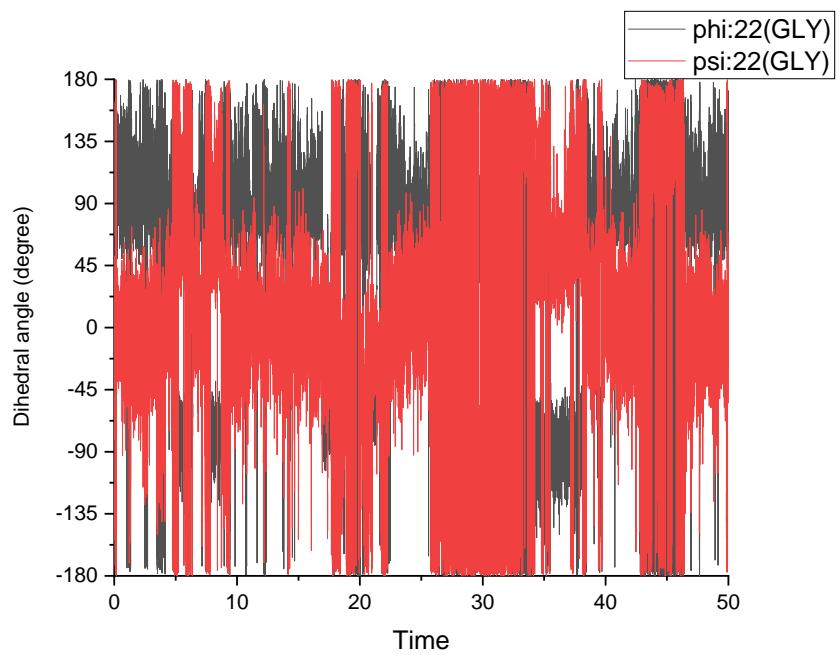




Medium pH at 498K:







Appendix D: Dihedral angle for chosen residues from selected frames of Zn(II)-A β 16-Mode1 simulation:

#Frame	Glu3 (ϕ)	Glu3(ψ)	Gly9(ϕ)	Gly9(ψ)	Glu11(ϕ)	Glu11(ψ)	Val12(ϕ)	Val12(ψ)	His13(ϕ)	His13(ψ)	His14(ϕ)	His14(ψ)	Gln15(ϕ)	Gln15(ψ)
602	52.00	22.00	118.85	-172.23	-97.38	-12.21	58.54	-68.61	-80.92	-25.79	59.50	38.64	52.74	4.18
3390	49.09	-70.16	-69.75	-170.80	-77.56	-19.96	63.68	-47.09	-87.36	-35.54	55.94	45.61	50.58	25.22
6000	55.87	24.03	135.31	-165.64	-83.85	3.81	53.60	-69.88	-111.56	-4.05	51.91	52.79	69.07	23.15
13567	53.18	1.76	115.92	164.39	-66.22	-38.56	59.16	-49.41	-85.40	-37.22	60.62	44.22	58.23	-15.67
50000	-75.24	152.13	133.96	-165.76	-94.38	0.38	52.44	-60.17	-99.98	-24.29	48.83	27.32	72.92	8.12
59998	-65.90	157.80	128.87	-156.72	-66.45	-9.60	55.19	-70.52	-80.64	-27.55	62.33	18.19	-80.06	-4.71
59999	-155.48	-172.11	162.09	-152.20	-73.83	-21.61	63.47	-70.85	-92.71	-22.74	59.92	15.02	-73.62	-47.04
60000	-77.24	179.55	147.32	-179.27	-68.21	-29.90	57.63	-50.54	-95.16	-24.69	58.66	36.63	-64.99	-25.20

