

# Modelling Platinum(II) Interactions With The Amyloid- $\beta$ Peptide

**Matthew Turner**

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



School of Chemistry  
Cardiff University  
March 14, 2018

## DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed: ..... (Candidate)

Date: .....

## STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.

Signed: ..... (Candidate)

Date: .....

## STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed: ..... (Candidate)

Date: .....

## STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and inter-library loan, and for the title and summary to be made available to outside organisations.

Signed: ..... (Candidate)

Date: .....

# ACKNOWLEDGEMENTS

First of all, I would like to express my gratitude to my supervisor, Dr James A. Platts. His consistent support and guidance have played an important role in my research as well as in my personal scientific development. I could not have asked for a better supervisor.

Thanks should also go to the other members of academic staff in the Theoretical Chemistry Group, particularly Dr David J. Willock for his help and mentorship during my post-graduate studies. Further thanks should go to Prof. Robert J. Deeth for his technical advice and experience with the LFMM work contained in this thesis.

I would also like to thank all of the PhD and Post-Doctoral students, past and present, that I have worked alongside. Particular notice should be given to Dr Liam Thomas, Dr Christian Reece and Dr Shaun Mutter, who have provided a great deal of advice and camaraderie during my time at 55 Park Place. Without them, my PhD would have been not only more difficult, but significantly less enjoyable. In addition, honourable mentions should be given to the undergraduate students I have had the pleasure of helping through the final stages of their degrees, especially Luke Perrott, Alice Schmitz, Roddy Stark and Alice Minto.

Further thanks should go to my family, particularly my parents and grandparents, for their support and encouragement of my academic endeavours, and to whom this thesis is dedicated.

Finally, I would like to thank Cardiff University for the scholarship which has funded my PhD studies and granted me the freedom to pursue my own ideas. Thanks also to ARCCA for access to the Raven supercomputing cluster, without which this project might never have been completed.

*For MJT*

## PUBLICATIONS

Work related to this thesis has been published in:

- (1) M. Turner, J. A. Platts and R. J. Deeth, *Journal of Chemical Theory and Computation*, 2016, **12**, 1385–1392.
- (2) M. Turner, R. J. Deeth and J. A. Platts, *Journal of Inorganic Biochemistry*, 2017, **173**, 44–51.
- (3) S. T. Mutter, R. J. Deeth, M. Turner and J. A. Platts, *Journal of Biomolecular Structure and Dynamics*, 2017, **36**, 1–9.
- (4) M. Turner, S. Mutter, R. J. Deeth and J. A. Platts, *Plos One*, In press.

# ABSTRACT

The A $\beta$  peptide is a key molecule in the development of Alzheimer’s Disease - A $\beta$  peptides form toxic aggregates in the brain. Density functional theory (DFT), Parametric Model 7 (PM7) and Ligand-Field Molecular Mechanics (LFMM) methods have been used to model the interactions of a series of potential therapeutic Pt<sup>II</sup>(ligand) complexes with various fragments of the A $\beta$  peptide. LFMM calculations with the AMBER forcefield were used to generate conformations of Pt<sup>II</sup>-A $\beta$ 6-14 via LowMode MD and results validated against BHandH. While LFMM showed insufficient agreement with DFT, the semi-empirical PM7 method displayed strong geometric and energetic agreement and was used to predict coordination preference and stable conformations of Pt<sup>II</sup>(bipy) and Pt<sup>II</sup>(phen) complexes. These species are shown to restrict the conformational freedom of A $\beta$  and the coordination of Pt<sup>II</sup>(phen) is shown to be in agreement with experimental data. Studies of additional Pt<sup>II</sup>(ligand) complexes in this manner revealed distinct preferences in metal binding mode for each of the complexes studied, with varied His N $\delta$  and His N $\epsilon$  coordination observed. Analysis of peptide conformations using Ramachandran plots and the STRIDE algorithm indicate that coordination of the Pt<sup>II</sup>(ligand) complexes disrupts existing peptide structure in a ligand-specific fashion, interrupting or translating the turn structure, suggesting that controlling peptide structure and behaviour may be achieved via ligand design. Ligand-field molecular dynamics (LFMD) simulations of Pt<sup>II</sup>(phen) -A $\beta$ 16 and -A $\beta$ 42 are compared to those of the metal-free peptides to investigate the influence of the Pt<sup>II</sup> complex on the structure and properties of the peptide. Simulations of A $\beta$ 16 and Pt<sup>II</sup>(phen)-A $\beta$ 16 revealed that Pt<sup>II</sup> coordination does not drastically alter peptide size, but increases the occurrence of 3,10-helical conformations while disrupting hydrogen bond and salt bridge networks. Simulation data also highlights the prevalence of  $\pi$ - $\pi$  stacking interactions between residues Phe4 and His13 with the phenanthroline ligand. Similarly, microsecond timescale simulations of Pt<sup>II</sup>(phen)-A $\beta$ 42 and A $\beta$ 42 illustrate profound effects of Pt<sup>II</sup> coordination on peptide structure; while A $\beta$  is shown to adopt collapsed conformations, Pt<sup>II</sup>-A $\beta$ 42 systems assume extended structures. Pt<sup>II</sup> coordination also induces large changes in peptide secondary structure, particularly an increase in helical character throughout the central hydrophobic region of the peptide, considered a potential route to preventing formation of A $\beta$  fibrils. The results detailed here provide insight into the coordination of these complexes to the peptide and present a new understanding of the effects of Pt<sup>II</sup> complexes on A $\beta$ .

# Contents

	Page
<b>1 Introduction</b>	<b>1</b>
1.1 Overview . . . . .	1
1.2 Alzheimer’s Disease . . . . .	1
1.3 A $\beta$ and the Amyloid Cascade Hypothesis . . . . .	2
1.4 Metal ions in AD . . . . .	9
1.4.1 Natural Metals . . . . .	9
1.4.2 Transition Metal Therapeutics . . . . .	11
1.5 Modelling A $\beta$ and Related Biomolecular Systems . . . . .	20
1.5.1 Quantum Mechanical Methods . . . . .	20
1.5.2 Semi-Empirical Methods . . . . .	23
1.5.3 Molecular Mechanics and Molecular Dynamics . . . . .	26
1.6 Modelling Metal-Biomolecule Systems . . . . .	33
1.6.1 Quantum Mechanical Methods . . . . .	33
1.6.2 Semi-Empirical Methods . . . . .	36
1.6.3 Molecular Mechanics-Based Methods . . . . .	39
<b>2 Theory</b>	<b>57</b>
2.1 Introduction . . . . .	57
2.2 Hartree-Fock Theory . . . . .	57
2.2.1 The Schrodinger Equation . . . . .	57
2.2.2 Born-Oppenheimer Approximation . . . . .	58
2.2.3 Molecular Orbital Approximation . . . . .	59
2.2.4 Hartree-Fock Approximation . . . . .	60
2.2.5 Linear Combination of Atomic Orbitals and the Roothaan-Hall Equations . . . . .	61
2.2.6 Self-Consistent Field Theory . . . . .	61
2.3 Density Functional Theory . . . . .	63
2.3.1 Local Density Approximation . . . . .	65

2.3.2	Generalised Gradient Approximation . . . . .	66
2.3.3	Hybrids . . . . .	66
2.3.4	Remarks . . . . .	67
2.4	Basis Sets . . . . .	69
2.4.1	Effective Core Potentials . . . . .	71
2.5	Molecular Mechanics . . . . .	72
2.5.1	Molecular Mechanics for Transition Metal Complexes . . . . .	74
2.5.2	Ligand Field Molecular Mechanics . . . . .	75
2.6	Semi-Empirical Methods . . . . .	85
2.6.1	ZDO Approximation . . . . .	85
2.6.2	NDDO Approximation . . . . .	86
2.7	Exploring Conformational Freedom . . . . .	88
2.7.1	Stochastic Search Methods . . . . .	89
2.7.2	LowMode Molecular Dynamics . . . . .	90
2.7.3	Molecular Dynamics . . . . .	92
2.7.4	Comparisons of Conformational Search Methods . . . . .	104
2.8	Practical Considerations . . . . .	106
2.8.1	Geometry Optimisation . . . . .	106
2.8.2	Solvent Models . . . . .	108
2.8.3	Protein Secondary Structure Assignment . . . . .	109
2.8.4	Programs and Methods . . . . .	112
<b>3</b>	<b>Modelling Pt<sup>II</sup>-Amyloid-<math>\beta</math> Complexes</b>	<b>120</b>
3.1	Introduction . . . . .	120
3.2	Modelling Pt <sup>II</sup> Crystal Structures . . . . .	120
3.3	Molecular Mechanics for Pt <sup>II</sup> Complexes . . . . .	128
3.3.1	Forcefield Validation . . . . .	128
3.3.2	Pt <sup>II</sup> Complexes . . . . .	131
3.4	Modelling the Metal-Binding Region of A $\beta$ . . . . .	135
3.4.1	Charge Schemes . . . . .	136
3.4.2	Exploring Conformations of Pt <sup>II</sup> -A $\beta$ 6-14 . . . . .	139
3.4.3	Application of the Modelling Protocol . . . . .	145
3.5	Conclusions . . . . .	158
<b>4</b>	<b>Predicting Ligand Effects on Platinum(L)-A<math>\beta</math> Coordination</b>	<b>163</b>
4.1	Introduction . . . . .	163
4.2	Application of the Modelling Protocol . . . . .	163
4.2.1	Computational Details . . . . .	164



4.2.2	Modelling Six Different Ligand Systems . . . . .	165
4.3	Conclusions . . . . .	183
<b>5</b>	<b>Molecular Dynamics Simulations of <math>\text{Pt}^{\text{II}}(\text{L})</math> - <math>\text{A}\beta</math> Systems</b>	<b>187</b>
5.1	Introduction . . . . .	187
5.2	$\text{A}\beta 16$ and $\text{Pt}^{\text{II}}(\text{phen})$ - $\text{A}\beta 16$ Simulations . . . . .	187
5.2.1	Computational Details . . . . .	188
5.2.2	$\text{A}\beta 16$ Equilibration . . . . .	193
5.2.3	$\text{A}\beta 16$ Production MD . . . . .	195
5.3	$\text{A}\beta 42$ and $\text{Pt}^{\text{II}}(\text{phen})$ - $\text{A}\beta 42$ Simulations . . . . .	220
5.3.1	Computational Details . . . . .	220
5.3.2	$\text{A}\beta 42$ Equilibration . . . . .	221
5.3.3	$\text{A}\beta 42$ Production MD . . . . .	224
5.4	Conclusions . . . . .	234
<b>6</b>	<b>Conclusion</b>	<b>238</b>
<b>7</b>	<b>Further Work</b>	<b>242</b>
<b>A</b>	<b>Appendix</b>	<b>243</b>
A.1	Appendices for Chapter 3 . . . . .	243
A.1.1	DFT Modelling of Xray Crystal Structures . . . . .	243
A.1.2	LFMM Parameters . . . . .	246
A.1.3	Validation of LFMM Parameters . . . . .	247
A.2	Appendices for Chapter 4 . . . . .	251
A.3	Appendices for Chapter 5 . . . . .	255
A.3.1	$\text{A}\beta 16$ and $\text{Pt}(\text{A}\beta 16)$ Data . . . . .	255
A.3.2	$\text{A}\beta 42$ and $\text{Pt}(\text{A}\beta 42)$ Data . . . . .	260
A.4	Guide for LFMD in DL-POLY . . . . .	262

# List of Figures

	Page
<b>Introduction</b>	<b>1</b>
1.1 Schematic of A $\beta$ 42 . . . . .	4
1.2 Solution phase structure of A $\beta$ 40 [51] containing a central 3,10 helical structure. . . . .	5
1.3 Experimentally determined A $\beta$ fibril structures.[58] Top: Monomer conformation in the fibril. Bottom: Association of $\beta$ -sheet monomers into large fibrils. . . . .	7
1.4 Generic fibril structure identified by Sunde <i>et al.</i> [71] . . . . .	8
1.5 Illustration of some of the the Pt <sup>II</sup> -ligand complexes discussed in this section. Top row: Pt <sup>II</sup> (1,10-phenanthroline) <b>1</b> , Pt <sup>II</sup> (4,7-diphenyl - 1,10-phenanthroline) <b>2</b> , Pt <sup>II</sup> (4,7-diphenyl-1,10-phenanthroline disulfonate) <b>3</b> . Middle row: <i>Cis</i> -platin <b>4</b> , Pt <sup>II</sup> (8-BQ) <b>5</b> , Pt <sup>II</sup> (pyridyl-benzimidazole) <b>6</b> . Bottom row: Pt <sup>II</sup> ( $\varphi$ -MePy) <b>7</b> , Pt <sup>II</sup> (2,2-bipyridine) <b>8</b> . . . . .	13
1.6 Cartoon of Pt <sup>II</sup> (phen)-A $\beta$ interaction.[90] . . . . .	14
1.7 Structures of [(bpy) <sub>2</sub> Ru <sup>II</sup> (dpp)Pt <sup>II</sup> Cl <sub>2</sub> ]Cl <sub>2</sub> ( <b>9</b> ) [98] and the bifunctional Pt <sup>II</sup> (bipyridine)(cyclen) complex ( <b>10</b> ) [99]. . . . .	17
1.8 Three dominant structures observed in A $\beta$ 1-39 simulations[139]. . . .	27
<b>Theory</b>	<b>57</b>
2.1 STO and GTO functions. STO is shown in red, GTO in black. . . . .	69
2.2 Reference angles in organic and transtion metal chemistry. Top row: Tetrahedral, trigonal planar and linear carbons. Bottom row: Octahedral and trigonal bipyramidal metal centres. . . . .	74
2.3 Overlapping ligand and coordination regions in Ligand Field Molecular Mechanics. Adapted from [22]. . . . .	75

2.4	Ligand positions in a square planar transition metal complex. These positions are used for Tables 2.1 and 2.2 below. . . . .	77
2.5	Plot showing the variation of ligand-metal $dz^2$ orbital overlap with angle, where $\theta$ is the deviation of the coordinating ligand from the z-axis.[23] . . . . .	78
2.6	$d$ -orbital splitting energies for a square planar complex. . . . .	79
2.7	Illustration of the ligand atom - $C_\alpha$ atom bond, highlighted in red. LFMM fails to model changes in this bond length upon metal coordination.[28] . . . . .	81
2.8	Illustration of the impact of timestep on MD simulation. Too small a timestep leads to limited coverage of phase space (left); Too large a timestep leads to unstable simulation (middle); An appropriate timestep allows a simulation to efficiently explore phase space (right).[3] . . . . .	98
2.9	Illustration of a Lennard-Jones potential with a potential cutoff, as used in MD simulation. . . . .	100
2.10	An example Potential Energy Surface, adapted from [83] . . . . .	107
2.11	Illustration of $\psi$ , $\phi$ and $\omega$ peptide backbone dihedral angles. . . . .	109
2.12	An example Ramachandran plot, $\Phi$ vs. $\Psi$ . . . . .	110

## Modelling $\text{Pt}^{\text{II}}$ -Amyloid- $\beta$ Complexes 120

3.1	Complexes used for DFT methods validation, where structures are taken from X-ray crystallographic data. Top row, left-right: (Dichloro-(histamine-N,N')-platinum(II) CSD ID: XUKWUV [11], chloro-(1-methyl-1H-imidazole)-(2-(pyridine-2-yl)phenyl)-platinum(II) CSD ID: OXUYUC [12], trans-Diamminebis (N-methyl-imidazole) platinum(II) chloride CSD ID: DAMIPT[13]. Bottom row, left-right: Chloro (ethylene-diamine) (1-methylimidazole)platinum(II) nitrate CSD ID: HEKJEM[14] and $\text{Pt}(\text{N-methylimidazole})_4$ CSD ID: MIMDPT[15]. . . . .	122
3.2	Atom numbering for the phenanthroline ligand, used for forcefield validation. . . . .	129
3.3	Illustration of AMBER types for $\text{Pt}^{\text{II}}$ and surrounding ligands, based on work by Scheeff [23]. . . . .	131
3.4	Example small model systems used for validation of LFMM parameters. Top row: Crystal structure MIMDPT[15], $\text{Pt}^{\text{II}}(\text{bipy})(\epsilon\epsilon \text{ Me-imidazole})_2$ . . . . .	133

3.5	Schematic of $\text{Pt}^{\text{II}}(\text{ligand})$ complexes binding to His6 and His13 (left) and His6 and His14 (right). The ligand, represented by the blue scaffold, may be either bipyridine or phenanthroline. . . . .	135
3.6	Atom names/numbers for an example $\text{Pt}^{\text{II}}(\text{phen})(\text{Me-imid})_2$ MK partial charge calculation . . . . .	138
3.7	Plot of DFT vs. PM7 relative energies at LFMM geometry . . . . .	151
3.8	Representative overlay of LFMM (blue) vs DFT (grey) geometry of $\text{Pt}^{\text{II}}(\text{bipy})\text{-A}\beta\text{6-14}$ complex. . . . .	152
3.9	Representative overlay of PM7 (yellow) vs DFT (grey) geometry of $\text{Pt}^{\text{II}}(\text{bipy})\text{-A}\beta\text{6-14}$ complex. . . . .	153
3.10	DFT optimised structure (grey) overlaid with PM7-DFT optimised structure (yellow) of $\text{Pt}^{\text{II}}(\text{bipy})\text{-A}\beta\text{6-14}$ complex. . . . .	154
3.11	Low energy conformation of $\text{Pt}^{\text{II}}(\text{bipy})\text{A}\beta\text{6-14}$ . . . . .	155
3.12	Lowest (grey), second- and third-lowest (yellow and blue, respectively) conformations of $\text{Pt}^{\text{II}}(\text{phen})\text{A}\beta\text{6-14}$ system. . . . .	156
3.13	$\pi$ - $\pi$ interaction in the $\text{Pt}^{\text{II}}(\text{bipy})\text{A}\beta\text{6-14}$ system. . . . .	158

## **Predicting Ligand Effects on Platinum(L)-A $\beta$ Coordination** **163**

4.1	Schematic of ligand systems studied. Top row, left-right: bipyridyl ( <b>1</b> ), phenanthroline ( <b>2</b> ); Middle row, left-right: diphenylphenanthroline ( <b>3</b> ), dppz ( <b>4</b> ); Bottom row, left-right: 8-(1H-benzimidazol-2-yl)quinolone [2] ( <b>5</b> ) and 2-pyridyl-benzimidazole [3] ( <b>6</b> ). . . . .	164
4.2	Low energy conformation of $\text{Pt}^{\text{II}}(\textbf{1})\text{-A}\beta\text{6-14}$ . . . . .	171
4.3	Low energy conformation of $\text{Pt}^{\text{II}}(\textbf{2})\text{-A}\beta\text{6-14}$ . . . . .	172
4.4	Low energy conformation of $\text{Pt}^{\text{II}}(\textbf{3})\text{-A}\beta\text{6-14}$ . . . . .	172
4.5	Low energy conformation of $\text{Pt}^{\text{II}}(\textbf{4})\text{-A}\beta\text{6-14}$ . . . . .	173
4.6	Low energy conformation of $\text{Pt}^{\text{II}}(\textbf{5})\text{-A}\beta\text{6-14}$ . . . . .	173
4.7	Low energy conformation of $\text{Pt}^{\text{II}}(\textbf{6})\text{-A}\beta\text{6-14}$ . . . . .	174
4.8	Ramachandran map of the generated conformations of the free A $\beta$ 6-14 fragment, plotted using JMP[16] . . . . .	176
4.9	Ramachandran map for the generated conformations of the low energy binding mode of $\text{Pt}^{\text{II}}(\textbf{1})\text{-A}\beta\text{6-14}$ . . . . .	177
4.10	Ramachandran map for the generated conformations of the low energy binding mode of $\text{Pt}^{\text{II}}(\textbf{2})\text{-A}\beta\text{6-14}$ . . . . .	178
4.11	Ramachandran map for the generated conformations of the low energy binding mode of $\text{Pt}^{\text{II}}(\textbf{3})\text{-A}\beta\text{6-14}$ . . . . .	178

4.12	Ramachandran map for the generated conformations of the low energy binding mode of Pt <sup>II</sup> ( <u>4</u> )-A $\beta$ 6-14 . . . . .	179
4.13	Ramachandran map for the generated conformations of the low energy binding mode of Pt <sup>II</sup> ( <u>5</u> )-A $\beta$ 6-14 . . . . .	179
4.14	Ramachandran map for the generated conformations of the low energy binding mode of Pt <sup>II</sup> ( <u>6</u> )-A $\beta$ 6-14 . . . . .	180
4.15	Ligand-Tyr10 distances in a) Pt <sup>II</sup> ( <u>1</u> ) His6 $\epsilon$ - His13 $\epsilon$ . . . . .	182
4.16	Ligand-Tyr10 distances in b) Pt <sup>II</sup> ( <u>1</u> ) His6 $\delta$ - His13 $\epsilon$ . . . . .	183
4.17	Ligand-Tyr10 distances in c) Pt <sup>II</sup> ( <u>2</u> ) His6 $\epsilon$ - His14 $\epsilon$ . . . . .	183
4.18	Ligand-Tyr10 distances in d) Pt <sup>II</sup> ( <u>2</u> ) His6 $\delta$ – His13 $\delta$ . . . . .	183

## **Molecular Dynamics Simulations of Pt<sup>II</sup>(L) - A $\beta$ Systems 187**

5.1	Schematic of A $\beta$ 16 . . . . .	188
5.2	Schematic of Pt <sup>II</sup> (A $\beta$ 16) . . . . .	188
5.3	Pt <sup>II</sup> (phen)(Me-imid) <sub>2</sub> region used in calculation of Merz-Kollman partial charges. . . . .	189
5.4	Superposed conformations A-E of A $\beta$ 16. . . . .	191
5.5	Superposed conformations F-J of Pt <sup>II</sup> (A $\beta$ 16). . . . .	192
5.6	RMSD ( $\text{\AA}$ ) vs time for 5 simulations of A $\beta$ 1-16 (A-E). . . . .	194
5.7	RMSD ( $\text{\AA}$ ) vs time for 5 simulations of Pt <sup>II</sup> (A $\beta$ 1-16) (F-J). . . . .	195
5.8	R <sub>g</sub> ( $\text{\AA}$ ) vs time for 5 simulations of A $\beta$ 1-16 (A-E). . . . .	197
5.9	R <sub>g</sub> ( $\text{\AA}$ ) vs time for 5 simulations of Pt <sup>II</sup> (A $\beta$ 1-16) (F-J). . . . .	199
5.10	Superposition of snapshots at time = 200 ns of simulations G and I of the Pt <sup>II</sup> (A $\beta$ 16) system. . . . .	201
5.11	RMSF per residue for simulations A-E of A $\beta$ 16 . . . . .	201
5.12	RMSF per residue for simulations F-J of Pt <sup>II</sup> (A $\beta$ 16). Data for simulation G is obscured by simulation I. . . . .	202
5.13	Top: Comparison of RMSF per residue for A $\beta$ 16 and Pt <sup>II</sup> (A $\beta$ 16) simulations. Bottom: Difference in RMSF between A $\beta$ 16 and Pt <sup>II</sup> (A $\beta$ 16) simulations . . . . .	203
5.14	Average SASA per residue of A $\beta$ 16 simulations A-E . . . . .	205
5.15	Average SASA per residue of Pt <sup>II</sup> (A $\beta$ 16) simulations F-J. . . . .	205
5.16	Percentage secondary structure by residue for A $\beta$ 16 simulation. Secondary structure assigned using STRIDE algorithm. . . . .	208
5.17	Percentage secondary structure by residue for Pt <sup>II</sup> (A $\beta$ 16) simulation. Secondary structure assigned using STRIDE algorithm. . . . .	209
5.18	Salt bridge contact map for A $\beta$ 16 simulation . . . . .	212

5.19	Salt bridge contact map for $\text{Pt}^{\text{II}}(\text{A}\beta 16)$ simulation . . . . .	213
5.20	Images of the three most prominent salt-bridges in $\text{A}\beta 16$ simulation. From top to bottom: Arg5-Asp7, Asp1-Lys16, Arg5-Glu11. Neg- atively charges residues are highlighted in blue; positively charged residues are highlighted in red. . . . .	215
5.21	Images of the three most prominent salt-bridges in $\text{Pt}^{\text{II}}(\text{A}\beta 16)$ simu- lation. From top to bottom: Lys16-Glu11, Lys16-Asp7, Arg5-Asp7. Negatively charges residues are highlighted in blue; positively charged residues are highlighted in red. . . . .	216
5.22	$\pi$ - $\pi$ contact distance for A) Phe4, B) Tyr10 and C) His13 residues with the phenanthroline ligand in $\text{Pt}^{\text{II}}(\text{A}\beta 16)$ simulation . . . . .	218
5.23	Snapshot of $\text{Pt}^{\text{II}}(\text{A}\beta 16)$ simulation F, where Tyr10 and His13 form $\pi$ - $\pi$ contacts of less than 5 Å with the phenanthroline ligand. Other residues are omitted for clarity . . . . .	219
5.24	Plot of $\pi$ - $\pi$ contact distance over production MD trajectory of simula- tion F. Phe4-ligand distance is shown in black, Tyr10-ligand distance in red and His13-ligand distance in blue. . . . .	219
5.25	Overlay of initial $\text{A}\beta 42$ conformations. Hydrogens, and non-aromatic, non-charged residues omitted for clarity . . . . .	220
5.26	Overlay of initial $\text{Pt}^{\text{II}}\text{-A}\beta 42$ conformations. $\text{Pt}^{\text{II}}(\text{phen})$ shown in black. Hydrogens and non-aromatic residues omitted for clarity. . . . .	221
5.27	RMSD (Å) vs time for 4 simulations of $\text{A}\beta 42$ . . . . .	222
5.28	RMSD (Å) vs time for 4 simulations of $\text{Pt}^{\text{II}}\text{-A}\beta 42$ . . . . .	223
5.29	$R_g$ (Å) vs time for simulations of $\text{A}\beta 42$ . . . . .	224
5.30	$R_g$ (Å) vs time for simulations of $\text{Pt}^{\text{II}}\text{-A}\beta 42$ . . . . .	225
5.31	Average RMSF per residue for simulations 1-4 of $\text{A}\beta 42$ and simu- lations 5-8 of $\text{Pt}^{\text{II}}\text{-A}\beta 42$ (top) and the difference in RMSF between $\text{A}\beta 42$ and $\text{Pt}^{\text{II}}\text{-A}\beta 42$ (bottom) . . . . .	227
5.32	Percentage secondary structure by residue for $\text{A}\beta 42$ simulation. Sec- ondary structure assigned using STRIDE algorithm . . . . .	228
5.33	Percentage secondary structure by residue for $\text{Pt}^{\text{II}}\text{-A}\beta 42$ simulation. Secondary structure assigned using STRIDE algorithm . . . . .	229
5.34	Salt bridge contact map for $\text{A}\beta 42$ simulation. . . . .	231
5.35	Salt bridge contact map for $\text{Pt}^{\text{II}}\text{-A}\beta 42$ simulation. . . . .	232

A.1	Ramachandran map for the generated conformations of the His6 N $\delta$ His13 N $\delta$ binding mode of Pt <sup>II</sup> ( <u>2</u> )-A $\beta$ 6-14 . . . . .	251
A.2	Ramachandran map for the generated conformations of the His6 N $\delta$ His13 N $\epsilon$ binding mode of Pt <sup>II</sup> ( <u>2</u> )-A $\beta$ 6-14 . . . . .	251
A.3	Ramachandran map for the generated conformations of the His6 N $\epsilon$ His13 N $\delta$ binding mode of Pt <sup>II</sup> ( <u>2</u> )-A $\beta$ 6-14 . . . . .	252
A.4	Ramachandran map for the generated conformations of the His6 N $\epsilon$ His13 N $\epsilon$ binding mode of Pt <sup>II</sup> ( <u>2</u> )-A $\beta$ 6-14 . . . . .	252
A.5	Ramachandran map for the generated conformations of the His6 N $\delta$ His14 N $\delta$ binding mode of Pt <sup>II</sup> ( <u>2</u> )-A $\beta$ 6-14 . . . . .	253
A.6	Ramachandran map for the generated conformations of the His6 N $\delta$ His14 N $\epsilon$ binding mode of Pt <sup>II</sup> ( <u>2</u> )-A $\beta$ 6-14 . . . . .	253
A.7	Ramachandran map for the generated conformations of the His6 N $\epsilon$ His14 N $\delta$ binding mode of Pt <sup>II</sup> ( <u>2</u> )-A $\beta$ 6-14 . . . . .	254
A.8	Ramachandran map for the generated conformations of the His6 N $\epsilon$ His14 N $\epsilon$ binding mode of Pt <sup>II</sup> ( <u>2</u> )-A $\beta$ 6-14 . . . . .	254
A.9	Initial conformations of A $\beta$ 16 simulations. Top row: Conformations A-C. Bottom row: Conformations D-E. . . . .	255
A.10	Initial conformations of Pt(phen)-A $\beta$ 16 simulations. Top row: Con- formations F-H Bottom row: Conformations I-J . . . . .	255
A.11	RMSF data for trajectories A-E of A $\beta$ 16 . . . . .	256
A.12	RMSF data for trajectories F-J of Pt <sup>II</sup> -A $\beta$ 16 . . . . .	256
A.13	SASA by residue for simulations A-E of A $\beta$ 16 . . . . .	257
A.14	SASA by residue for simulations F-J of Pt <sup>II</sup> -A $\beta$ 16 . . . . .	258
A.15	Plot of $\pi$ - $\pi$ contact distance over production MD trajectory of simula- tion G. Phe4-ligand distance is shown in black, Tyr10-ligand distance in red and His13-ligand distance in blue. . . . .	258
A.16	Plot of $\pi$ - $\pi$ contact distance over production MD trajectory of simula- tion H. Phe4-ligand distance is shown in black, Tyr10-ligand distance in red and His13-ligand distance in blue. . . . .	259
A.17	Plot of $\pi$ - $\pi$ contact distance over production MD trajectory of simula- tion I. Phe4-ligand distance is shown in black, Tyr10-ligand distance in red and His13-ligand distance in blue. . . . .	259
A.18	Plot of $\pi$ - $\pi$ contact distance over production MD trajectory of simula- tion J. Phe4-ligand distance is shown in black, Tyr10-ligand distance in red and His13-ligand distance in blue. . . . .	260
A.19	Average RMSF per residue for simulations 1-4 of A $\beta$ 42 . . . . .	261

A.20 Average RMSF per residue for simulations 5-8 of $\text{Pt}^{\text{II}}$ -A $\beta$ 42 . . . . .	261
--	-----



# 1 Introduction

## 1.1 Overview

This thesis aims to present research carried out on modelling the interactions between platinum<sup>II</sup>-ligand complexes and differing size fragments of the amyloid- $\beta$  ( $A\beta$ ) peptide using various quantum mechanical, semi-empirical and molecular mechanics methods. The literature review contained in this chapter concentrates on three main areas of research related to  $A\beta$  in the context of Alzheimer's disease. Firstly, the role and nature of  $A\beta$  in Alzheimer's disease will be examined. Secondly, the role of metal ions in  $A\beta$  aggregation and their potential in therapeutics will be discussed, before finally considering the computational modelling of these and related metal-biomolecule systems.

## 1.2 Alzheimer's Disease

Alzheimer's disease (AD) is a widespread neurodegenerative disease associated with progressive cognitive decline in patients and is most commonly seen in those over the age of 65. AD is the most common cause of dementia in adults - responsible for 60-70% of cases - and currently affects more than 30 million people worldwide [1, 2], with approximately a further 1.5 million additional cases each year.[3] Common symptoms of AD include short-term memory loss, mild cognitive impairment, confusion and aggression. Over time, cognitive function degrades, and control of bodily functions is lost, eventually leading to death. As a result, AD is the fourth most common cause of death in Western countries.[4]

AD is associated with damage to brain regions involved in memory and cognition, specifically the hippocampus and cerebral cortex.[5] AD brains also exhibit several abnormal structures, particularly deposits of neurofibrillary tangles, and senile

plaques.[6, 7]; formation of these structures is generally accepted to be the major pathological identifier of Alzheimer's.[8] In addition, brains show oxidative stress caused by reactive oxygen species (ROS) [9, 10] and increased concentrations of metal ions, such as copper [11], zinc [5, 12], iron [13] and calcium.[14] That these ions are present in such high concentrations suggests there is a breakdown in metal trafficking regulation in the affected brains.[15, 16]

The causes and development of AD are not well understood - the pathogenesis of AD is very complex, involving the interplay of a range of molecular, cellular and physiological processes[17], but risk factors have been reported to include various lifestyle features (*e.g.* smoking [18], alcohol use [19], diabetes [20], low physical and intellectual activity levels [21] and diet [22]) as well as genealogy.[23] However, the most important risk factor is age.[24] Despite a great deal of research and many clinical trials, at the time of writing there is no known cure for AD. Current treatments only alleviate symptoms, at best slowing the disease progression by a matter of months.[25]

## 1.3 $A\beta$ and the Amyloid Cascade Hypothesis

Many studies have suggested that the development of Alzheimer's disease is primarily associated with the formation of senile plaques and neurofibrillary tangles in the brain.[26] These plaques are dense deposits of protein that are typically seen surrounding nerve cells, while the neurofibrillary tangles – mainly consisting of hyperphosphorylated tau protein – are seen inside the nerve cells.[26] The amyloid- $\beta$  ( $A\beta$ ) peptide was first recognised as the main component of the deposited plaques seen in Alzheimer's patient brain tissue in 1985, in work by Masters *et al.* [27] which signalled the start of the modern era of research into AD. In this section, we consider the most prominent hypotheses of AD progression, and the chemistry and structure of  $A\beta$ , seemingly a key molecule in AD, is discussed.

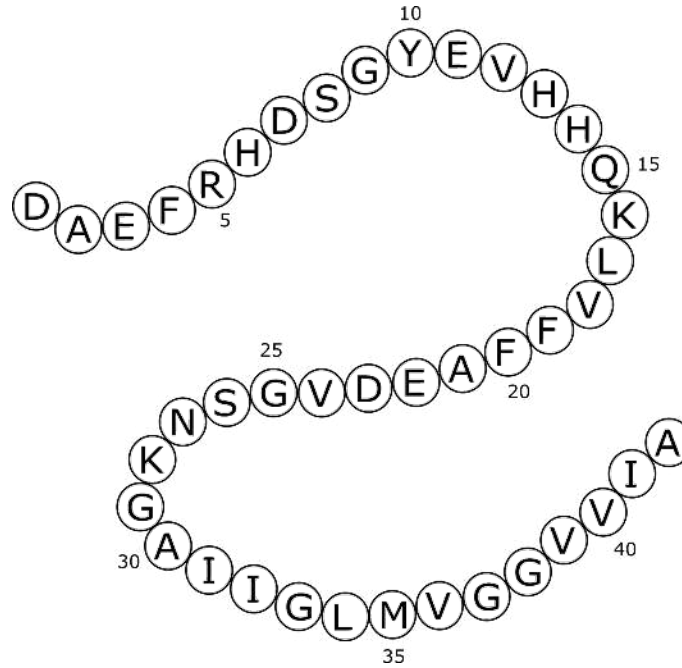
There are three main hypotheses on the causes and progression of AD: the amyloid (cascade) hypothesis, the metal ion hypothesis and the oxidative stress hypothesis. Importantly, these hypotheses are not independent - metals such as copper and zinc may act as seeds for amyloid aggregation and/or may play a role in the mechanism of oxidative stress.

The amyloid (or amyloid cascade) hypothesis is the most prominent of the current hypotheses and is arguably the most complete. Here, the main cause of AD is believed to be the build-up of  $A\beta$  in the brain, and that amyloids are the main neurotoxic substance in disease progression. Alternatively, the metal ion hypothesis states that AD is caused by a breakdown in the transport and regulation of metal ions, such as copper, zinc and iron. As a result of this, there is an imbalance between  $A\beta$  generation and clearance, leading to neurotoxicity. This understanding of the disease progression is becoming increasingly popular. Finally, the oxidative stress hypothesis suggests that the production of reactive oxygen species in the brain leads to a greater degree of oxidative stress. This stress then leads to neurological disorders and cell death.

The research in this thesis sits at the intersection of the amyloid and metal ion hypotheses, so the rest of this chapter will consider the details of these ideas, rather than the oxidative stress hypothesis.

$A\beta$  peptides, typically between 39-43 residues in length, are formed by cleavage of the Amyloid Precursor Protein (APP) - a membrane-associated protein that is believed to act as a cell surface signalling molecule - by  $\beta$ - and  $\gamma$ -secretases. Of the peptides produced,  $A\beta_{40}$  and  $A\beta_{42}$  are the dominant forms -  $A\beta_{42}$  is represented in Figure 1.1. The  $A\beta_{40}$  variant is far in excess of  $A\beta_{42}$ , but this longer peptide is both more prone to aggregation [28, 29] and more toxic to cells.[26, 30–32] The  $A\beta$  monomer is amphiphilic as it contains both a hydrophilic N-terminal region and a hydrophobic C-terminal region.[5]  $A\beta$  – regardless of its form – contains six negatively charged residues (Asp1, Glu3, Asp7, Glu11, Glu22 and Asp23) and three positively charged residues (Arg5, Lys16 and Lys28), giving a net charge of -3 on the monomer. The three histidine residues are neutral at physiological pH.[33]

Early versions of the amyloid hypothesis [34] suggested that  $A\beta$  possesses some innate toxicity and that the plaque deposits indicated an overload of  $A\beta$ , causing AD.[35, 36] This was known as the amyloid cascade hypothesis, where an imbalance of  $A\beta$  generation and clearance leads to gradual accumulation of  $A\beta$  plaques in the brain.[37–40] Thus, the formation of these plaques leads to a local inflammatory response and subsequently to neuronal cell death, leading to gradual cognitive decline.[8, 26] The rest of the known disease process, including the formation of neurofibrillary tangles containing hyperphosphorylated tau protein, would then occur from this imbalance of  $A\beta$  generation and clearance.[26, 36]

Figure 1.1: Schematic of A $\beta$ 42

While this amyloid cascade hypothesis has been highly credited and provides a general framework to explain the disease process, it lacks detail and several observations do not fit with the hypothesis, as explained below.[8, 26] As a result, the *cascade* term is now frequently set aside, and the modified understanding is known simply as the amyloid hypothesis.

Firstly, the number of amyloid deposits shows a weak correlation with the degree of cognitive impairment.[25, 35] In fact, patients without symptoms of Alzheimer's disease show significant deposits of A $\beta$  [41–43]; there seems to be a much better correlation of cognitive impairment with the concentration of soluble A $\beta$  oligomers present.[38, 40] Crucially, it has been shown that soluble oligomeric forms of A $\beta$  are more toxic than the fibrillated plaques [36, 38, 40]; these oligomeric forms of A $\beta$  are now increasingly believed to be the key toxic species on AD.[44–47] As a result, the amyloid hypothesis is now focused on these diverse oligomeric species. Finally, while A $\beta$  is created throughout the brain [16], AD initiates in specific brain regions such as the hippocampus, suggesting that there may be another disease trigger specific to these regions [48] – recent work has shown that these regions show high concentrations of transition metal ions (vide infra).

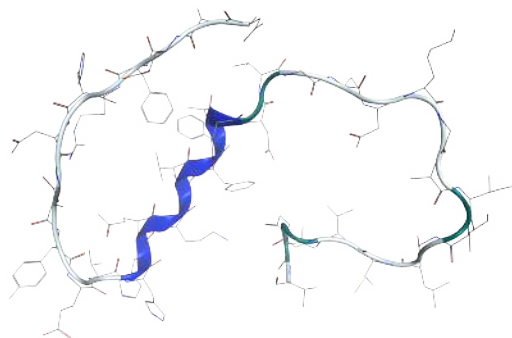


Figure 1.2: Solution phase structure of A $\beta$ 40 [51] containing a central 3,10 helical structure.

However, the fact remains that understanding the structure of these various forms of A $\beta$  may play an important role in further elucidating its role in AD progression. The structure and chemistry of A $\beta$  has been reviewed in detail previously [3, 5, 29, 49, 50], but the physical characteristics of A $\beta$  are still poorly understood:

In monomeric form, A $\beta$  is an intrinsically disordered peptide, meaning that in aqueous solution, it adopts a heterogeneous, random coil structure [51–53] – this disorder means that there is little solution-phase NMR structural data and neither A $\beta$ 40 nor A $\beta$ 42 monomers crystallise to provide X-ray crystallographic data. However, the peptide secondary structure also depends strongly on the chemical conditions (*i.e.* solvent, pH *etc.*) In water, various studies have reported the A $\beta$  monomer to contain between 5-20% helical content and 0-25%  $\beta$ -sheet content, though the peptide remained mostly “random coil” in nature, as seen in PDB models 2LFM [51], 1AML [52], 1BA4 [54], and others.[49] In general, the C-terminus of A $\beta$  is considered to form an extended helix-type structure, as shown in Figure 1.2. This structural variation poses challenges for studying A $\beta$ , as identifying relevant conformations of the monomer is critical to understanding the oligomerisation and aggregation process, as well as for providing appropriate starting structures for computational studies of A $\beta$ . [26, 29]

In the amyloid hypothesis, A $\beta$  monomers aggregate, forming A $\beta$  oligomers, which may lead to formation of the mature amyloid fibrils. To date, a broad range of oligomeric species have been reported,[49, 55] but no high-resolution atomic structure has yet been defined for any.[50] In particular, A $\beta$  dimers and trimers have attracted much research interest:

A $\beta$  trimers have been shown to be especially harmful to brain function,[56] while dimeric species have been obtained directly from AD brain tissue.[57] There is some suggestion that a dimeric structure of A $\beta$  may be the precursor to further aggregation *i.e.* that mature fibrils of A $\beta$ 40 and A $\beta$ 42 are constructed from dimers of A $\beta$  stacked into parallel  $\beta$ -sheets.[58–60]

In general, the  $\beta$ -sheet content of A $\beta$  oligomers appears to increase over time [61], as well as increasing with the degree of aggregation.[62] However, the smallest oligomers have also been shown to retain a significant degree of disorder – NMR experiments detected oligomers which displayed predominantly turn and coil type structure.[63] Alternatively, analysis of fibril-like A $\beta$  oligomers show that they possess significant  $\beta$ -sheet content.[60] Regardless, it appears that the peptide monomer must undergo some change in conformation, likely involving the formation of a hairpin near Gly25 flanked by large  $\beta$ -sheets, in order for oligomerisation and aggregation to take place.[64]

A $\beta$  amyloid fibrils are thought to be the end-point of a complex and protracted aggregation pathway. The formation of fibrils may be understood in terms of an uncontrolled protein folding process: a monomer interacts with an identical neighbour and the structure collapses, becoming more ordered. This is repeated over time, and the aggregate formed becomes increasingly compact and ordered, eventually forming the mature fibrils of stacked A $\beta$  peptide. These fibrils appear to resemble the plaques seen in the brains of AD patients.

The formation of these fibrils is generally monitored by a fluorescence assay using thioflavin T (ThT). Upon fibril formation, a large increase in fluorescence is observed, reaching saturation as complete fibril structures are obtained. ThT binds to  $\beta$ -sheet regions of A $\beta$  fibrils, so an increased signal corresponds to a greater degree of  $\beta$ -sheet content and therefore, of fibril formation.

As before, understanding the structure of these A $\beta$  fibrils is crucial for understanding their chemistry. Firstly, it appears that multiple fibril structures exist, generally based on parallel  $\beta$ -sheet structures,[65, 66] though those containing the anti-parallel type are also known. Work by Tycko *et al.* [58], illustrated in Figure 1.3, showed that fibrils of A $\beta$ 40 contain dimeric peptide units stacked into  $\beta$ -sheets, where residues on adjacent chains are aligned; fibrils of A $\beta$ 42 possess identical structure, except

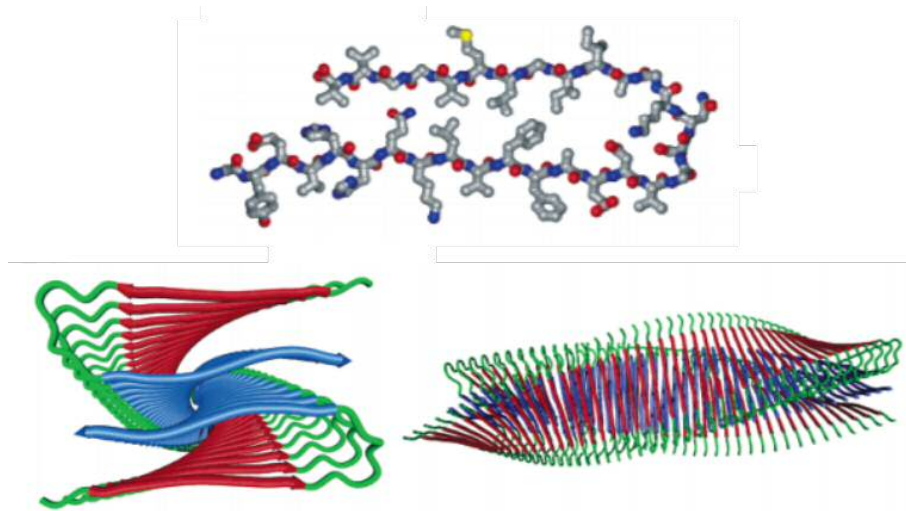


Figure 1.3: Experimentally determined A $\beta$  fibril structures.[58] Top: Monomer conformation in the fibril. Bottom: Association of  $\beta$ -sheet monomers into large fibrils.

that residues are not aligned.[34] Each monomer unit contains two  $\beta$ -sheet regions (His13-Asp23 and Lys28-Val40) separated by a hairpin turn, and are stabilised by a salt bridge interaction between Asp23 and Lys28.[58] Alternatively, a 3D structure of A $\beta$ 42 fibrils has been determined by solid-state NMR [60], showing parallel  $\beta$ -sheets involving residues Val18-Val36. At least two monomers are required to produce the repeating unit of the fibril. This appears in agreement with other work [67, 68] which suggest that the central core of A $\beta$ 42 consists mainly of  $\beta$ -sheet type structure, though the exact residues involved vary by study.

In addition, the structure of amyloid fibrils has been studied using X-ray diffraction: Kirschner *et al.* [69] characterised A $\beta$ 1-28 fibrils as tubular, with average diameter of 86 Å, where the walls consist of cross- $\beta$ -pleated sheets. Here, the peptide chains run approximately perpendicular to the fibril axis. In addition, high  $\beta$ -turn potentials were identified at residues Ser8 and Ser26, and the intervening region adopts either  $\alpha$ -helical or  $\beta$ -sheet conformation. Inouye *et al.*[70] also identified the amyloid fibril structure as a tube, with five or six 'crystallites' forming the wall of the tube, where each crystallite is made of five or fewer  $\beta$ -pleated sheets. More recently, Sunde *et al.*[71] identified A $\beta$  fibril structure as a helix of highly-ordered  $\beta$ -sheet peptides, arranged parallel to the fibril axis - see Figure 1.4.

While there are certainly weaknesses in the amyloid hypothesis, none of these provide compelling reasons to abandon this theory. However, they do point to limitations in the current understanding of Alzheimer's disease.

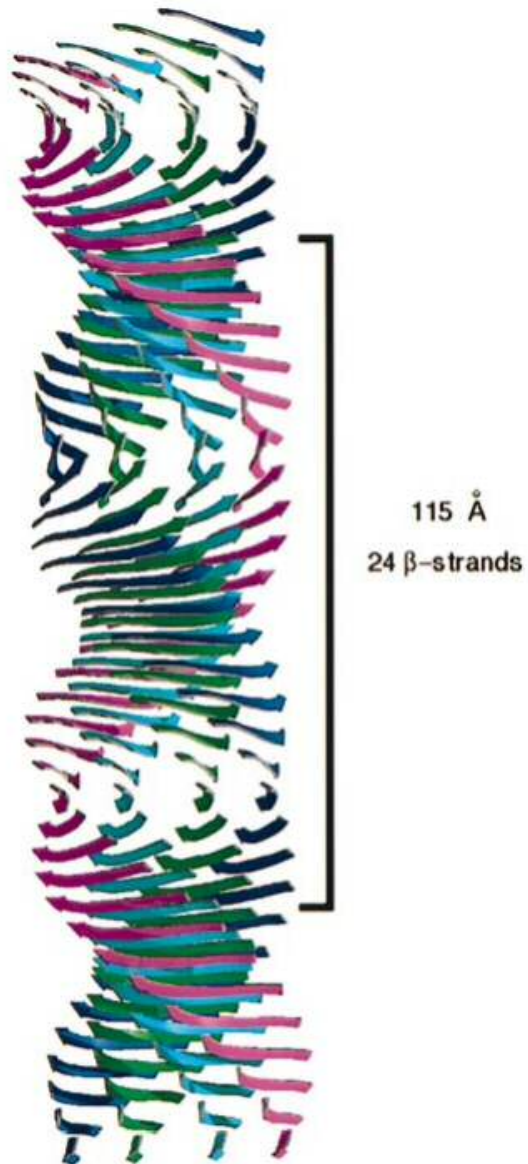


Figure 1.4: Generic fibril structure identified by Sunde *et al.*[71]



## 1.4 Metal ions in AD

### 1.4.1 Natural Metals

This section will provide an overview of the coordination chemistry of naturally occurring transition metal ions  $\text{Cu}^{\text{II}}$ ,  $\text{Zn}^{\text{II}}$  and  $\text{Fe}^{\text{II}}$  in the context of  $\text{A}\beta$ , and their effects on its aggregation properties.

The metal ion hypothesis (see Section 1.3) is founded on observations that the progression of AD correlates with the breakdown of  $\text{Cu}^{\text{II}}$  /  $\text{Zn}^{\text{II}}$  /  $\text{Fe}^{\text{II}}$  ion transport and distribution in the brain. [5, 72–74] Interestingly,  $\text{Cu}^{\text{II}}$  concentration is generally reported as decreased in AD brain tissue, while  $\text{Fe}^{\text{II}}$  concentration appears elevated.[5]  $\text{Zn}^{\text{II}}$  concentrations are more controversial, with many varied reports on the subject.[5] The key breakthrough in this area illustrated that the presence of  $\text{Cu}^{\text{II}}$  /  $\text{Zn}^{\text{II}}$  is essential for the formation and stability of amyloid aggregates such as fibrils or oligomers, [5, 75] while normal metal ion concentrations are unable to induce  $\text{A}\beta$  aggregation.

As a result, the role of the natural metal ions in AD is increasingly studied – these topics have been extensively reviewed elsewhere, and this section aims to provide a summary of the pertinent points.[5, 12, 13, 29, 49, 50]

#### 1.4.1.1 $\text{Cu}^{\text{II}}$ -, $\text{Zn}^{\text{II}}$ - and $\text{Fe}^{\text{II}}$ - $\text{A}\beta$ Coordination Chemistry

$\text{Cu}^{\text{II}}$  is a redox active, open shell  $d^9$  ion which typically interacts with  $\text{A}\beta$  in a 1:1 stoichiometry, while a second low-affinity  $\text{Cu}^{\text{II}}$  binding site has also been detected. These interactions have been well-studied via techniques such as NMR and electron paramagnetic resonance (EPR), but the nature of the  $\text{Cu}^{\text{II}}$  coordination is pH-dependent and varies depending on the peptide fragment studied.[5, 49] These different coordination modes are known as type I (dominant at low pH) and type II binding (dominant at high pH). At physiological pH, type I binding dominates.

Type I  $\text{Cu}^{\text{II}}$  coordination consists of two or more species in equilibrium, but typically involves three N-donor ligands (*e.g.* His residues or the N-terminal amide) and at

least one oxygen-donor ligand (*e.g.* the Ala2 carbonyl oxygen, or possibly the Asp1 sidechain carboxylate) giving a four- or five-coordinate  $\text{Cu}^{\text{II}}$  centre.[5, 49] For type II, various coordination modes have been proposed, though there is no consensus on the residues involved. This may be partly explained by the fact that addition of paramagnetic ions such as  $\text{Cu}^{\text{II}}$  produces line broadening and loss of signals close to the metal in NMR spectra.

$\text{Zn}^{\text{II}}$  is a  $d^{10}$  closed shell ion, and also interacts with  $\text{A}\beta$  in a 1:1 stoichiometry. NMR experiments have determined that the  $\text{Zn}^{\text{II}}$  binding site on  $\text{A}\beta$  involves the three N-terminal histidines (His6, His13 and His14) alongside Glu11, forming a distorted tetrahedral geometry at the metal.[76] However, the nature of the final (non-His) coordinating group is a matter for debate - Tyr10, the Asp1 N-terminal amine and the deprotonated amide of the Arg-5 backbone have been detected, depending on reaction conditions. In addition, some studies have shown that  $\text{Zn}^{\text{II}}$  binds to His residues in adjacent  $\text{A}\beta$  peptides, thus facilitating aggregation.[77]

Like  $\text{Cu}^{\text{II}}$ ,  $\text{Fe}^{\text{II}}$  is redox active ( $\text{Fe}^{\text{II}}/\text{Fe}^{\text{III}}$  redox couples are common in biological systems), but to date, few studies have detected or characterised iron- $\text{A}\beta$  binding, perhaps due to the easy oxidation of  $\text{Fe}^{\text{II}}$  to  $\text{Fe}^{\text{III}}$ . However, it appears that the  $\text{Fe}^{\text{II}}$  ligands on  $\text{A}\beta$  are some combination of Asp1, Glu3, His6, His13 and His14.[78]

#### 1.4.1.2 $\text{Cu}^{\text{II}}$ and $\text{Zn}^{\text{II}}$ Effects on $\text{A}\beta$ Aggregation

It is now generally accepted that the formation of  $\text{A}\beta$  aggregates in the brain is not spontaneous, but requires the presence of metal ions, such as  $\text{Cu}^{\text{II}}$  or  $\text{Zn}^{\text{II}}$ . A failure to regulate the concentration of these metal ions may lead to the formation of  $\text{A}\beta$  aggregates *i.e.* metal ions may act as initiators of  $\text{A}\beta$  aggregation. In fact, both  $\text{Cu}^{\text{II}}$  and  $\text{Zn}^{\text{II}}$  are thought to promote the aggregation of  $\text{A}\beta$ , perhaps by coordinating multiple peptide chains, as previously mentioned. The binding of these metal ions to  $\text{A}\beta$  changes the peptide structure, though the induced changes are diverse - and often contradictory - and results depend on the relative concentration of the metal ion. These interactions are described in detail elsewhere.[5, 49] (and references within.)

Firstly,  $\text{Cu}^{\text{II}}$ -amyloids are known to be neurotoxic,[79] but the form of these ag-

gregates depends on  $\text{Cu}^{\text{II}}$  concentration and reaction conditions; in general, sub-equimolar  $\text{Cu}^{\text{II}}$  binding has been shown to accelerate fibril formation,[80, 81] while higher concentrations of  $\text{Cu}^{\text{II}}$  induce formation of amorphous aggregates. Interestingly, these amorphous structures  $\text{Cu}^{\text{II}}$  aggregates have been reported to be nontoxic [79, 82], though there is some evidence that these amorphous aggregates convert into the structured fibrils on timescales longer than normally studied by experiment. Additionally, it has been reported that transition of the  $\text{A}\beta$  monomer from coil to  $\beta$ -sheet structure upon  $\text{Cu}^{\text{II}}$  binding, observed via circular dichroism, is accompanied by formation of  $\text{A}\beta$  dimers.[49]

In contrast,  $\text{Zn}^{\text{II}}$ -amyloid structures themselves are generally considered to be non-toxic, though  $\text{Zn}^{\text{II}}$  has been implicated in the formation of  $\text{A}\beta$  oligomers.[79] Like  $\text{Cu}^{\text{II}}$ ,  $\text{Zn}^{\text{II}}$  coordination may produce fibrillar or amorphous  $\text{A}\beta$  aggregates depending on conditions, though these amorphous structures may be converted to the amyloid fibrils during longer incubation times. The contradictory nature of these results is evident in studies which suggest that  $\text{Zn}^{\text{II}}$  binding increases the  $\alpha$ -helical character of  $\text{A}\beta$ , while others suggest that  $\text{Zn}^{\text{II}}$  in fact disrupts the helical nature of the  $\text{A}\beta$  monomer, aiding the  $\alpha$ -helix to  $\beta$ -sheet transition seen in the aggregation process.[5, 49]

## 1.4.2 Transition Metal Therapeutics

As we have already seen, there has been a great deal of research interest into elucidating the nature of the interactions between naturally occurring metals and  $\text{A}\beta$ . At the same time, there is an emerging field of research that is focused on metal-based imaging and therapeutic compounds, designed to target  $\text{A}\beta$ . Full discussion of the transition metal based imaging agents is beyond the scope of this thesis, but interested readers are directed to recent reviews on the topic [17, 83] and the references therein. Instead, this section will focus on the research that forms the foundation of this work – the use of transition metal based therapeutics to target  $\text{A}\beta$  in the context of AD, with a particular focus on  $\text{Pt}^{\text{II}}$  compounds.

Importantly,  $\text{Pt}^{\text{II}}$  compounds are stable and essentially redox inert when present in biological systems. The slow kinetics associated with substitution reactions at the  $\text{Pt}^{\text{II}}$  centre means that, once bound to a target, the  $\text{Pt}^{\text{II}}$  metal is difficult to displace. Furthermore, as a ‘soft’ metal,  $\text{Pt}^{\text{II}}$  has a preference for ligands with ‘soft’

donor atoms. Within the A $\beta$  sequence, potential Pt<sup>II</sup> binding sites include the ‘soft’ sulphur atom of Met35 and the imidazole nitrogen atoms of the three histidine side chains that are considered intermediate between ‘hard’ and ‘soft’ ligands – these imidazole side chains are known to be excellent ligands for a variety of metal ions, including Pt<sup>II</sup>.

The field of Pt<sup>II</sup>-based inhibitors of A $\beta$  aggregation was pioneered by Barnham *et al.*[84] who examined a series of Pt<sup>II</sup>(Ligand)Cl<sub>2</sub> complexes for their ability to bind to A $\beta$ . Here, the ligands studied were 1,10-phenanthroline (shown as **1** in Figure 1.5), 4,7-diphenyl-1,10-phenanthroline (**2** in Figure 1.5) and 4,7-diphenyl-1,10-phenanthroline disulfonate (**3**), while the classic anticancer drug cisplatin (cis-Pt(NH<sub>3</sub>)<sub>2</sub>Cl<sub>2</sub>) (**4**) was used as a control. Mass spectrometry (MS) and NMR experiments illustrated that the Pt<sup>II</sup> complexes formed adducts with A $\beta$ . NMR data showed that while cisplatin binds solely to Met35, Pt<sup>II</sup>(phenanthroline)-type complexes bind to the imidazole side chains of His6, His13, and His14. Furthermore, data in the cisplatin case was consistent with the formation of A $\beta$  oligomers and multiple other products.

Pt<sup>II</sup>(phenanthroline) complexes were shown to inhibit A $\beta$  fibril formation using ThT fluorescence, while cisplatin did not. This was reinforced by electron microscopy studies; Pt<sup>II</sup>(phenanthroline) generated amorphous A $\beta$  aggregates rather than fibrils. Pt<sup>II</sup>(phenanthroline) complexes were also able to suppress hydrogen peroxide production by Cu<sup>II</sup>-A $\beta$ 42, as monitored via fluorimetric assay[85], while cisplatin produced no observable effect. This is surprising, as cisplatin is shown to coordinate to Met35, a residue widely implicated in the redox chemistry of A $\beta$ [86, 87]. Finally, these complexes inhibit A $\beta$  toxicity in mouse neuronal cell cultures, while cisplatin proved inactive. Importantly, these Pt<sup>II</sup> complexes were not toxic at the concentrations required to produce this restorative effect.

The difference in behaviour between the Pt<sup>II</sup>(phenanthroline) complexes and cisplatin suggests that the planar aromatic ligands provide some specificity for A $\beta$  to the platinum complexes. In fact, the free ligands possess an intrinsic (weak) affinity for the A $\beta$  peptide, as they interact with the aromatic residues Phe-4, Tyr-10, and Phe-19[88], suggesting that the Pt<sup>II</sup>(phenanthroline)-A $\beta$  interactions were mediated by  $\pi$ - $\pi$  stacking arrangements.

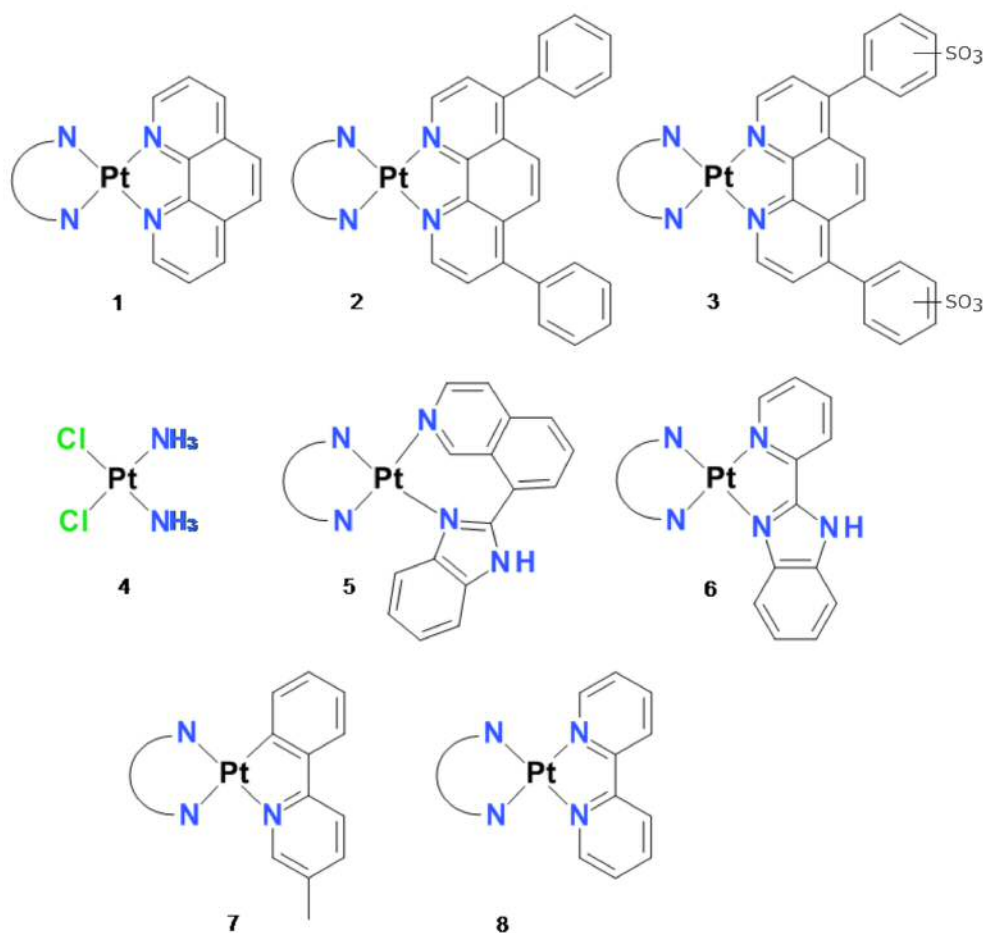


Figure 1.5: Illustration of some of the the Pt<sup>II</sup>-ligand complexes discussed in this section. Top row: Pt<sup>II</sup>(1,10-phenanthroline) **1**, Pt<sup>II</sup>(4,7-diphenyl -1,10-phenanthroline) **2**, Pt<sup>II</sup>(4,7-diphenyl-1,10-phenanthroline disulfonate) **3**. Middle row: *Cis*-platin **4**, Pt<sup>II</sup>(8-BQ) **5**, Pt<sup>II</sup>(pyridyl-benzimidazole) **6**. Bottom row: Pt<sup>II</sup>( $\phi$ -MePy) **7**, Pt<sup>II</sup>(2,2'-bipyridine) **8**.

Recently, Streltsov *et al.*[89] used extended X-ray absorption fine structure spectroscopy (EXAFS), MS, dynamic light scattering (DLS) and density functional theory (DFT) to study the interactions of cisplatin and Pt<sup>II</sup>(4,7-diphenyl -1,10-phenanthroline disulfonate)Cl<sub>2</sub> with A $\beta$ 16 and A $\beta$ 42. DLS experiments confirmed that Pt<sup>II</sup>(4,7-diphenyl-1,10-phenanthroline disulfonate)Cl<sub>2</sub> inhibits the aggregation of A $\beta$ 42. MPW1PW91 / SDD calculations performed in gas phase reproduced crystallographic bond distances and angles at Pt<sup>II</sup> to 0.14 Å or better and 3.60° or better, respectively. DFT-optimised geometries of small models of Pt<sup>II</sup>-A $\beta$  binding - containing Pt<sup>II</sup>, ligand and either imidazoles or thioethers - were used in conjunction with EXAFS data to estimate Pt-N, Pt-Cl and Pt-S distances and produce

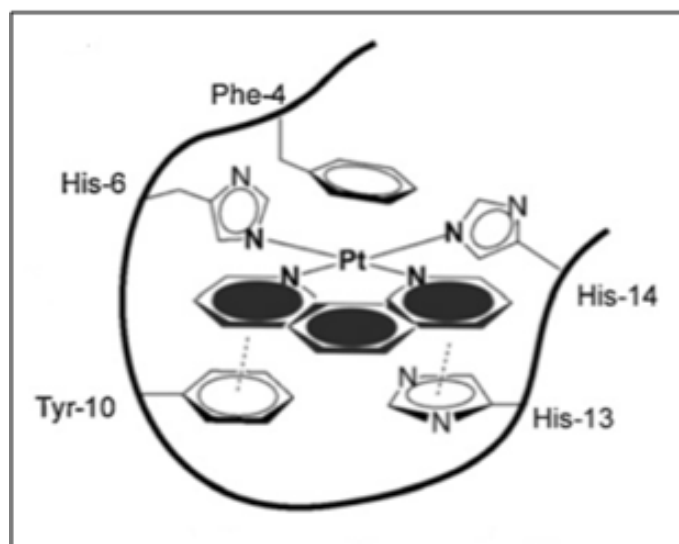


Figure 1.6: Cartoon of  $\text{Pt}^{\text{II}}(\text{phen})\text{-A}\beta$  interaction.[90]

structural models of these interactions.

Work by Ma *et al.*[90] investigated the interactions of  $\text{Pt}^{\text{II}}(\text{phenanthroline})$  with  $\text{A}\beta_{16}$  - displayed in Figure 5.1 - using HPLC, ESI-MS and NMR and detected more than ten product peaks, though a single product appeared dominant; changing the reaction conditions (temperature, pH) gave no clear improvement in product complexity. MS data showed mono- and bis-platinated products, suggesting that there are multiple coordination sites in  $\text{A}\beta_{16}$ . In contrast, cisplatin did not react with  $\text{A}\beta_{16}$ . ESI-MS/MS and NMR showed that the dominant product had platinum binding sites at His6 and His14 on  $\text{A}\beta$ , as shown in Figure 1.6, while no coordination to His13 was detected. Other major products displayed His6/Lys16 and Asp7/His14 binding, where  $\text{Pt}^{\text{II}}$ -coordination is predicted to occur via carboxylate groups. The fourth major product was identified as the bis-platination adduct, meaning that four binding sites on  $\text{A}\beta_{16}$  may be simultaneously occupied.

Later, the same authors (Ma *et al.*,[91]) studied the reaction of  $\text{Pt}^{\text{II}}(\text{phen})\text{Cl}_2$  complexes with  $\text{Cu}^{\text{II}}$ - and  $\text{Zn}^{\text{II}}$ - $\text{A}\beta_{16}$  complexes by HPLC, ESI-MS, EPR and tandem MS. Here, the platinum coordination was shown to alter the binding modes of  $\text{Cu}^{\text{II}}$  and  $\text{Zn}^{\text{II}}$ . Reaction of  $\text{Pt}^{\text{II}}(\text{phen})\text{Cl}_2$  complexes with  $\text{A}\beta_{16}$  formed a new adduct, indicating that  $\text{Cu}^{\text{II}}$ -coordination influences the products of this reaction, though His6 and His14 remain the preferred  $\text{Pt}^{\text{II}}$  binding site. The presence of  $\text{Cu}^{\text{II}}$  slows down this reaction as  $\text{Cu}^{\text{II}}$  binds to the same N-terminal residues as the  $\text{Pt}^{\text{II}}$  complex. Upon incubation with  $\text{Pt}^{\text{II}}(\text{phen})$ ,  $\text{Cu}^{\text{II}}$ - $\text{A}\beta$  EPR signals were altered, from type I

(binding via Asp1, His6, His13/His14 and the N-terminus) to type II (binding Ala2, His6, His13 and His14).

Similarly, reaction of  $\text{Zn}^{\text{II}}$ -A $\beta$ 16 with  $\text{Pt}^{\text{II}}(\text{phen})\text{Cl}_2$  confirmed the formation of mono- and bis-platinated A $\beta$ 16 species, as well as the presence of a Pt-Zn bimetallic species, though further work to elucidate the metal binding sites of these systems was not performed.

More recently, Barnham *et al.*[92] developed a platinum complex suitable for *in vivo* testing. Here an inactive  $\text{Pt}^{\text{IV}}$  species may be converted in the body to the active  $\text{Pt}^{\text{II}}$  complex by natural reductants. Importantly, these  $\text{Pt}^{\text{IV}}$  complexes possess very slow kinetics, such that in a biological context, the complexes are very stable and able to survive the acidic environment of the stomach and cross the gut membrane. Here, the 8-(1H-benzoimidazol-2-yl)-quinoline(8-BQ)[93] (**5**) ligand was studied – it provides the large aromatic surface area required to make the platinum complex target A $\beta$ , while the NH group means the ligand may be easily modified to alter the pharmacokinetic properties of the complex; in this work, an N,N-dimethylaminoethyl group was used to improve complex stability and solubility.[94]

This platinum complex was shown to inhibit the formation of amyloid fibrils by ThT fluorescence and electron microscopy experiments. MS analysis showed that the  $\text{Pt}^{\text{II}}$  complex binds to the peptide by losing one or both chloride ligands. While treatment of mouse neuronal cell cultures with A $\beta$ 42 was shown to be toxic, co-incubation with the active complex almost completely restored cell viability. Importantly, this complex was not toxic at the concentrations tested. Finally, testing the  $\text{Pt}^{\text{IV}}$  pro-drug in a mouse model of AD showed that treatment lead to a significant reduction in A $\beta$ 42 levels and plaque number.

Work by Yellol *et al.*[95] also studied a new ligand scaffold for  $\text{Pt}^{\text{II}}$  inhibitors of A $\beta$ , and also studied the corresponding  $\text{Ru}^{\text{II}}$  and  $\text{Ir}^{\text{III}}$  complexes, which are discussed later. Here, the pyridyl-benzimidazole ligand (**6**) was used, which may also be easily functionalised to tune the pharmacokinetic properties. Inhibition of A $\beta$ 42 aggregation was tested by the ThT fluorescence assay - 1  $\mu\text{M}$  concentration of the  $\text{Pt}^{\text{II}}$  species was sufficient to inhibit the amyloid aggregation process. This result was confirmed by transmission electron microscopy (TEM). While treatment of mouse neuronal cell cultures with A $\beta$  is toxic, addition of this  $\text{Pt}^{\text{II}}$  complex shows only a

marginal, insignificant decrease in A $\beta$  toxicity.

In 2012, Sasaki *et al.*[96] studied the coordination of a new cyclometallated Pt<sup>II</sup> system to the A $\beta$ 28 peptide by EPR, NMR and ESI-MS methods. In this study, [Pt<sup>II</sup>( $\varphi$ -MePy)(DMSO)Cl] , where  $\varphi$ -MePy is the 2-phenyl-5-methyl-pyridine ligand, (**7**) was compared with the Pt<sup>II</sup>(4,7-diphenyl-1,10-phenanthroline)Cl<sub>2</sub> and Pt<sup>II</sup>(2,2'-bipyridine)Cl<sub>2</sub> (**8**) systems. The cyclometallated system contains a covalent platinum – carbon bond that may impart stability in aqueous solution to the complex. In fact, the reactivity of the two labile ligands were shown to be different - the ligand trans to the Pt<sup>II</sup>-C bond is proposed to be more labile than the ligand cis to the Pt<sup>II</sup>-C bond. The authors proposed that this complex binds A $\beta$  via one His residue, where this His imidazole ring is located trans to the Pt<sup>II</sup>-C bond, in place of the most labile ligand. EPR data showed that binding of the cyclometallated complex altered the Cu<sup>II</sup> coordination, from type I to type II (see Section 1.4.1.1). ESI-MS experiments showed that these complexes primarily formed mono-platinated adducts with A $\beta$ 28.

Later, the same authors[97] investigated the effect of five Pt<sup>II</sup> compounds on Zn<sup>II</sup>-induced A $\beta$ 28 aggregation and Cu<sup>II</sup>-induced ROS production. These complexes induced no significant change in ROS production, despite these complexes being known to alter the Cu<sup>II</sup>-A $\beta$ 28 coordination mode, as studied by EPR. Addition of each of the five Pt<sup>II</sup> complexes to Zn<sup>II</sup>-A $\beta$ 28 samples decreased fibril formation; these results were confirmed by TEM. Furthermore, NMR and HPLC/MS data suggests that the cyclometallated complexes react with A $\beta$ 28 to a greater extent than the other Pt<sup>II</sup> species studied. This was further probed using HPLC coupled to MS/MS experiment, which identified His13 and Glu11 as potential platinum binding sites.

An interesting alternative to these Pt<sup>II</sup> complexes was published by Rangachari *et al.*[98] who studied a mixed-binuclear metal complex containing Pt<sup>II</sup> and Ru<sup>II</sup> metal centres: [(bpy)<sub>2</sub>Ru<sup>II</sup>(dpp)Pt<sup>II</sup>Cl<sub>2</sub>]Cl<sub>2</sub> - (**9** in Figure 1.7). Here, the bipyridyl ligands provide a large hydrophobic surface designed to target the complex to A $\beta$ .

ThT fluorescence showed that incubation with the mixed metal complex inhibited fibril formation, while NMR data suggested that this complex may bind differently to A $\beta$ 42 than the original Pt<sup>II</sup>(phen) species; monomeric A $\beta$  was detected with two Ru<sup>II</sup>-Pt<sup>II</sup> complexes bound, indicating that there must be more than one binding



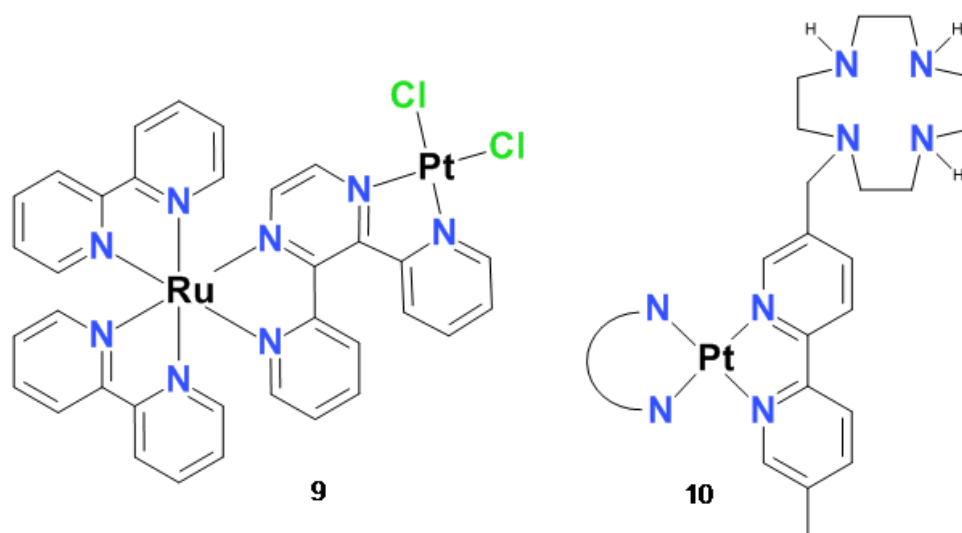


Figure 1.7: Structures of  $[(\text{bpy})_2\text{Ru}^{\text{II}}(\text{dpp})\text{Pt}^{\text{II}}\text{Cl}_2]\text{Cl}_2$  (**9**) [98] and the bifunctional  $\text{Pt}^{\text{II}}(\text{bipyridine})(\text{cyclen})$  complex (**10**) [99].

site for this compound on the peptide. Furthermore, the authors performed a series of quantum chemical calculations to study the binding energy of the bimetallic species. Calculations were performed using the BHandHLYP method with a series of basis sets, and the PCM solvation model. These data suggest that it is preferable for the  $\text{Ru}^{\text{II}}\text{-Pt}^{\text{II}}$  complex to bind to His residues rather than the N-terminus.

A different approach was taken by Wang *et al.*[99], who designed bifunctional  $\text{Pt}^{\text{II}}$ -based inhibitors of metal-induced  $\text{A}\beta_{40}$  aggregation, where  $\text{Pt}^{\text{II}}(\text{bipyridine})\text{Cl}_2$  was used as the  $\text{A}\beta$  binding unit, with cyclen to chelate the naturally-occurring metal ions,  $\text{Cu}^{\text{II}}$  and  $\text{Zn}^{\text{II}}$ , shown as **10** in Figure 1.7. NMR, MS and tandem-MS experiments showed that these complexes bind to the peptide via one or both of His13 and His14.

Inhibition of  $\text{A}\beta$  aggregation was studied through turbidimetry, TEM and a protein assay. Addition of these complexes greatly inhibited  $\text{Cu}^{\text{II}}$  or  $\text{Zn}^{\text{II}}$ -induced  $\text{A}\beta$  aggregation, while  $\text{Cu}^{\text{II}}$ - $\text{A}\beta$  ROS generation was shown to decrease by approximately 75%. TEM data showed that addition of the  $\text{Pt}^{\text{II}}$  complexes produced small ‘granule-like’ species, with different morphology to either  $\text{A}\beta$  or the  $\text{Cu}^{\text{II}}$ - or  $\text{Zn}^{\text{II}}$ -  $\text{A}\beta$  samples. Finally, these complexes were tested in mouse neuronal cell cultures and mouse models. While the  $\text{Cu}^{\text{II}}$ - or  $\text{Zn}^{\text{II}}$ - $\text{A}\beta$  samples were toxic to cells, addition of the  $\text{Pt}^{\text{II}}$  species increased cell viability significantly.

The success of these  $\text{Pt}^{\text{II}}$  agents in inhibiting the aggregation of  $\text{A}\beta$  has led to an expanding field of TM-based therapeutics, including the repurposing of existing biologically-active species to target this process. An overview of these other TM-based species is given below:

In 2010, Valensin *et al.* [100] investigated the interactions between a ruthenium<sup>II</sup> complex,  $\text{fac-}[\text{Ru}^{\text{II}}(\text{CO})_3\text{Cl}_2(\text{N1-thiazole})]$ , and the  $\text{A}\beta_{28}$  peptide. Like platinum<sup>II</sup>, Ruthenium<sup>II</sup> complexes are known to target histidine residues, and are typically less cytotoxic than their group 10 counterparts. ESI-MS and NMR studies confirmed that the  $\text{Ru}(\text{CO})_3$  unit binds to the peptide via N-terminal histidine residues, while the thiazole ligand is released. However, subsequent line-broadening in NMR spectra prevented detailed analysis of the  $\text{Ru}^{\text{II}}$ - $\text{A}\beta_{28}$  adducts formed. NMR signals for both the peptide and the  $\text{Ru}^{\text{II}}$  complex decay over time, which might be accounted for by the formation of large insoluble aggregates. In addition, circular dichroism studies illustrated that while the free  $\text{A}\beta_{28}$  adopts a random coil conformation, addition of the  $\text{Ru}^{\text{II}}$  species led to formation of insoluble aggregates.

Similarly, Messori *et al.* [101] studied three ruthenium<sup>III</sup> complexes (NAMI A, KP1019 and PMRU20) for their ability to block  $\text{A}\beta_{42}$  aggregation and toxicity using an *in vitro* model of AD. While NAMI A and KP1019 were poorly active, PMRU20 was shown to be highly active in protecting cortical neurons against  $\text{A}\beta_{42}$  toxicity. In addition, this complex displays no significant cytotoxicity at the concentrations tested. Characterisation using ESI-MS suggests that a stable adduct is formed between PMRU20 and  $\text{A}\beta_{42}$ , while ThT fluorescence assays indicate that the complex interacts with  $\text{A}\beta_{42}$  and reduces  $\text{A}\beta$  fibril formation. However, later work by Jones *et al.* [102] showed that KP1019 is able to interact with  $\text{A}\beta_{42}$  and promotes formation of high molecular weight aggregates rather than toxic  $\text{A}\beta$  oligomers. In addition, KP1019 was also shown to limit  $\text{A}\beta_{42}$  toxicity in human cell cultures.

As previously mentioned, Yellol *et al.* [95] studied  $\text{Pt}^{\text{II}}$ ,  $\text{Ru}^{\text{II}}$  and  $\text{Ir}^{\text{III}}$  complexes with the pyridyl-benzimidazole ligand. Here, 1  $\mu\text{M}$  of all three compounds was sufficient to inhibit  $\text{A}\beta$  amyloid aggregation, as confirmed by ThT fluorescence and TEM. However, while the  $\text{Ru}^{\text{II}}$  compound is toxic and therefore unable to rescue  $\text{A}\beta$  toxicity, the  $\text{Ir}^{\text{III}}$  compound exhibited a robust rescuing of  $\text{A}\beta$  neurotoxicity.

Recently, Cobalt<sup>III</sup> Schiff base complexes have been investigated by Heffern *et al.* [103] for their ability to interact with A $\beta$ 16, as well as their effects on the aggregation of A $\beta$ 42. HPLC-MS, NMR and fluorescence techniques were used alongside DFT studies to demonstrate that the Co<sup>III</sup> complexes interact with the N-terminal His residues in A $\beta$  via dissociative ligand exchange at the axial ligand positions. Coordination of the cobalt complexes to A $\beta$ 42 altered peptide structure and promoted the formation of large (> 30 kDa) soluble oligomers – this correlated with a reduced level of binding of A $\beta$  to synapses in hippocampal neurons. DFT calculations at the B3LYP/6-31+G(d) level were performed on small models of Cobalt<sup>III</sup>- A $\beta$  and determined that coordination of His6 to one axial position and either His13 / His14 to the other was the most favourable ligand arrangement.

Similarly, work by Li *et al.* [104] used triple-helical dinuclear complexes [105] to enantioselectively inhibit A $\beta$ 40 aggregation. These complexes were shown to be able to cross the blood-brain barrier, bind A $\beta$ , inhibit the aggregation process and destabilise preformed A $\beta$  fibrils. Furthermore, these complexes suppress A $\beta$ 40-induced ROS production in cell cultures. Circular dichroism spectroscopy showed that these complexes could inhibit the coil- $\beta$ -sheet structural transition seen in A $\beta$  aggregation. Molecular docking studies predicted that the complexes bind between residues 16-23. The authors also suggest that benzene rings in the complexes are able to form  $\pi$ - $\pi$  stacking interactions with Phe19 and Phe20 residues in the peptide, disrupting the natural hydrophobic and  $\pi$ - $\pi$  interactions of these residues.

Work by Man *et al.* [106] studied iridium<sup>III</sup> and rhodium<sup>III</sup> complexes as inhibitors of A $\beta$ 40 aggregation. Here, the effect of these compounds on A $\beta$  aggregation was studied by fluorescence microscopy. Of these complexes, the rhodium<sup>III</sup> compound was shown to be the most effective anti-A $\beta$  agent.

Finally, a different approach to A $\beta$  aggregation inhibition was exhibited by He *et al.* [107], who used peroxovanadium-based complexes to bind A $\beta$  and oxidise the Met35 residue; methionine oxidation has been reported to be important in amyloid peptide aggregation as the resulting methionine sulfoxide is less flexible and more polar than the natural residue. ThT fluorescence was significantly decreased after incubation with the vanadium complexes, suggesting that these complexes reduce the ability of A $\beta$  to aggregate. In addition, the vanadium complexes induce a change in A $\beta$ 42 secondary structure, as characterised by the circular dichroism spectrum.

Finally, cell viability studies showed that incubation of the vanadium complexes decreased the cytotoxic effect of A $\beta$ 42.

On this evidence, it appears that there is great potential in the area of metal- and especially platinum<sup>II</sup>-based therapeutics for the discovery of new anti-Alzheimer’s drugs, due in part to the relative lack of research that has been devoted to it. However, as with all new classes of drugs, it is vitally important to determine the toxicity of new metal-based compounds; while these compounds are known to target biologically important proteins, little is known about their neurotoxicity *in vivo*. It is especially important to consider the stability of the complex in the body and its uptake properties. Furthermore, it should be checked that the concentration of agent required to inhibit aggregation does not cause damage to other organs. If these conditions are met, it may be possible to develop the drug pharmaceutically. However, it has been suggested by Barnham [92] that simply reducing the amyloid build-up in the brain is not enough to recover cognitive function. It has long been suggested [8] that A $\beta$  deposits up to two decades before patients present clinical symptoms, meaning that anti-amyloid treatment would need to be begin early on in the disease process in order to be effective, potentially even at the pre-clinical stage of the disease if patients could be identified.

## 1.5 Modelling A $\beta$ and Related Biomolecular Systems

The flexible nature of A $\beta$  described above means that detailed structural information is difficult to obtain via experimental methods. As a result, much work has been done to elucidate the structure and chemistry of these peptides by molecular modelling, and particularly, molecular dynamics methods. In this section, a selection of this work will be discussed, with focus on the full A $\beta$  peptide.

### 1.5.1 Quantum Mechanical Methods

While popular and generally accurate, quantum mechanical approaches are not commonly applied to large biological systems; these calculations are often too expensive to use for biomolecular systems, particularly when trying to find the global minimum on a large and complex potential energy surface. Importantly, there is no published

data using quantum mechanical calculations to predict the structure of A $\beta$ . In this section, a selection of research into modelling small peptides using these methods is discussed.

Vargas *et al.* [108] used second-order Moller-Plesset (MP2, with the aug-ccpVDZ, aug-cc-pVTZ, and aug-cc-pVQZ basis sets) and DFT (B3LYP or BLYP/TZVP+) calculations to study six conformers of the alanine dipeptide. DFT accurately represents peptide structure when compared to the higher level method, though in flat regions of the potential energy surface, structural differences may be large, with errors in dihedral angles up to 37°. For the six conformers, all methods predict the same order of stability - DFT methods give energies that differ by less than 0.40 kcal mol<sup>-1</sup>, though BLYP/ TZVP+ and MP2/aug-cc-pVXZ levels, give an energy difference of 1.31 kcal mol<sup>-1</sup> *i.e.* DFT is generally able to reproduce the accuracy of higher levels of theory for conformations of small biomolecules.

Perczel *et al.* [109] studied nine gas phase conformers of the alanine dipeptide via HF and B3LYP (with 3-21G, 6-31+G(d) and 6-311++G(d,p) basis sets) and compared results to MP2, MP4D and CCSD(T) data. HF and DFT predict between five and seven minima for this system, rather than all nine. However, the geometries and relative energies of these lower level calculations show strong correlation with MP2 and CCSD(T) results. All methods were shown to predict similar potential energy surfaces for the peptide, indicating that DFT is able to suitably model peptide systems.

In 2011, Abo-Riziq *et al.* [110] used B3LYP/6-31G\*\* to study eighty conformers of Tyr-Gly-Gly peptide. Here, fewer low energy conformers are found than for smaller Tyr or Tyr-Gly systems. Calculated stable conformations agree with those identified by experiment, including absorption and vibrational spectra.

Valdes *et al.*[111] studied realistic models of aromatic-aromatic side chain interactions in proteins using a number of conformations of the phenylalanyl-glycyl-phenylalanine tripeptide. Here, the system was modelled using a range of DFT functionals, which were compared to MP2 data. In general, DFT produced accurate peptide geometries and predicted the correct energy ordering of conformations; in particular, the Truhlar functionals[112] (MUE = 0.42 kcal/mol) and BHandH[113] (MUE = 0.89 kcal/mol) proved accurate.

Work by Shields[114] and Holroyd[115] studied conformers of the Tyr-Gly dipetide using B3LYP/6-31+G(d) and MP2/6-31+G(d). By comparison with higher ab initio calculations, it was shown that DFT was able to successfully predict the most stable conformer, while MP2 could not. However, B3LYP DFT calculations produced incorrect changes in potential energy when the residues are close together, suggesting that these methods require inclusion of dispersion effects to model these systems.

Following this, van Mourik [116] assessed a series of density functionals (B3LYP, B97-1, X3LYP, BHandH, mPW2-PLYP, PWB6K, M05-2X, M06-2X, M06-L, B3LYP-D, and mPW2-PLYP-D) to handle this problem, including modelling of several low energy conformations. This work highlights the difficulty of accurately describing flexible peptides containing  $\pi$ -interactions; while most methods identify the three minima expected, only M06-2X and mPW2-PLYP-D predict the correct order of stability of these three structures.

Shubert *et al.* [117] assessed the performance of DFT for predicting the gas-phase structure of two 20-residue peptides (Ac-Ala19-Lys and Ac-Lys-Ala19), where weak interactions such as dispersion or hydrogen bonds are important. Here, the PBE and PBE0 functionals used, with several different dispersion corrections, to model 3026 conformations of these peptides. Computational results are validated against MS and IR data. These calculations predict a low number of stable conformers, while the lowest energy structure agrees with experimental characterisation.

Rauschenberg *et al.* [118] used BLYP-D3/def2-TZVP calculations along side extensive experimental data to study the interactions between two cyclic hexapeptides (Cys-His-Cys and Cys-Tyr-Cys) with a selection of small molecules in aqueous solution in both gas phase and COSMO implicit solvent. 64 conformations of each complex were optimised using this DFT method and the calculations confirm the small molecule selectivity of the peptides observed in experiment. In addition, DFT predicts the formation of multiple intermolecular hydrogen bonds in the most stable complexes.

Work by Ninkovic[119] studied the interactions of aromatic and aliphatic amino acids in amyloid formation. Analysis of model systems from the Protein Data Bank showed that B3LYP-D3/6-31G\* calculations showed good geometric and energetic agreement with CCSD(T)/CBS data; as such these DFT calculations were applied

to a series of two tetramers of aromatic and aliphatic peptides, used to represent  $\beta$ -sheet structure in adjacent amyloid peptides. These were analysed for their total interaction energy, as well contributions from sidechain-sidechain, sidechain-backbone and backbone-backbone interactions. This revealed that in amyloids containing aromatic residues, interaction of aromatic-aliphatic side chains are the most important stabilising factor, while in aliphatic amyloids, stability comes from side-chain interactions with the backbone as well as hydrogen bonds.

Barnham *et al.* [120] used DFT to propose a reaction scheme for catalytic generation of  $\text{H}_2\text{O}_2$  by  $\text{A}\beta$ . These calculations suggest that a critical role of Tyr10 in this process: a tyrosine radical leads to cross-linked  $\text{A}\beta$  chains and peptide oligomerisation. DFT data was shown to be in good agreement with experimental work studying production of  $\text{H}_2\text{O}_2$  by  $\text{A}\beta$  /  $\text{Cu}^{\text{II}}$ .

In summary, the application of quantum mechanical methods to large, flexible or unstructured biomolecules such as  $\text{A}\beta$  remains uncommon - *i.e.* where a single conformation is insufficient to describe peptide behaviour - though these methods have found a great deal of success in modelling smaller or more structured biomolecular systems.

## 1.5.2 Semi-Empirical Methods

Semi-empirical methods, particularly the newer PM6 and PM7 methods (see Section 2.6), are also relatively infrequently applied to large, flexible biological systems, in part due to reasons of cost and in part due to concerns over accuracy. In this section, a series of applications of these modelling techniques to biomolecules are discussed. To the knowledge of this author, neither PM6 (and related methods) or PM7 have previously been applied to studies of intrinsically disordered peptides such as  $\text{A}\beta$ .

Crucially, Stewart [121] assessed the suitability of PM6 for modelling proteins by studying the properties of a series of proteins obtained from the Protein Data Bank. For single amino acid residues, PM6 gave average all-atom RMSD versus B3LYP/6-31G(d) calculations of 0.225 Å, with errors associated with rotations of carboxylic acid and amino groups. This work also considered classical protein secondary structures, and compared PM6 optimised RMSD with that of peptide X-ray structures; in all cases, errors were low. For 26-residue helices, the all atom RMSD was reported

as 0.73 Å, with errors in the associated Ramachandran plot (see Section 2.8.3) of approximately 6°.  $\beta$ -sheet structures were modelled to approximately the same degree of accuracy, even in systems containing multiple peptide chains. The performance of this method for modelling salt-bridge interactions was also studied – here, reported RMSDs from X-ray data was around 0.9 Å in peptides of up to 46 residues. Protein tertiary and quaternary structure were also successfully reproduced by PM6, with RMSD (versus X-ray) values on the order of 1 Å, indicating that PM6 can be used for accurately modelling protein structures and properties.

Pichierri used semi-empirical PM5 and PM6 methods to study human erythropoietin[122] (133 residues) and charybdotoxin, a scorpion venom peptide[123], (37 residues) respectively. In both cases, calculation of the PM5/6 was not validated against another theoretical method, but calculated HOMOs correlate with reactive sites in the biomolecule, suggesting that the method suitably describes these systems.

In 2014, Temelso *et al.*[124] used semi-empirical calculations to determine low energy conformers of a threonine-proline-valine-asparagine tetrapeptide. This approach was used to screen out high-energy conformers using lower levels of theory and predict accurate peptide structures for the most stable conformers. By comparing results against MP2 data, it was shown that DFT structures gave the highest agreement (RMSD < 0.2 Å) while all semi-empirical methods (with the exception of PM3) also produced accurate peptide geometries (RMSD < 1.0 Å). In the case of the semi-empirical methods, most of this structural deviation was seen in the freely rotatable capping groups on the peptide. For calculating relative energies of these conformations, PM6-H2 and PM6-DH+ methods produce the lowest mean-square errors of all semi-empirical methods tested (0.35 and 0.90 kcal/mol) relative to MP2 data, though observed standard deviations (1.69 and 2.14 kcal/mol) are significantly greater than those found using DFT.

Faver *et al.* [125] investigated the performance of various semi-empirical methods, including PM3, PM6, and PM6-DH2 in modelling the interaction energy of ubiquitin protein fragments by comparing them to MP2/CBS and CCSD(T)/CBS values. Of these, PM6-DH2 was the most accurate, though standard deviations of the errors were still quite large.

Dobes *et al.*[126] used PM6-DH2X and AMBER to investigate complexes of a large



protein kinase, containing  $\alpha$ -helical and  $\beta$ -sheet structural elements, with a series of 9 aromatic, halogen-containing inhibitors. The semi-empirical method was shown to be able to accurately reproduce known interactions between inhibitors and the protein, both in terms of overall geometry (PM6-DH2X RMSD = 0.14 Å, AMBER RMSD = 1.14 Å) and the description of halogen bonds. In addition, binding free energies were calculated and compared to known experimental inhibition constants. These data showed strong agreement between the semi-empirical results and experiment ( $R^2 = 0.86$ ), while no agreement was observed for AMBER ( $R^2 = 0.04$ ).

In 2013, Stewart introduced the next iteration of semi-empirical methods, PM7, which was shown to reduce errors from PM6 and improve applicability of the semi-empirical methods to biomolecules and crystalline solids.[127] Of particular interest is the application of PM7 to large biological systems - structures of about 70 proteins from the PDB were modelled, with the assumption that if existing structures can be accurately modelled, then alternate conformations (*i.e.* those found in biochemical processes) would also be accurately modelled, and vice versa. Average all-atom RMSDs for these systems were reported as less than 1 Å, though results were sensitive to errors in weak, long-range interactions.

Ivan *et al.*[128] used semi-empirical PM6 and PM7 methods to study twenty-eight conformations of the relatively complex drug molecule (3S,5S,6S)-6-acetylamidopenicillanic acid. Few differences were observed in the geometries or heats of formation of these conformers with the two methods, illustrating that semi-empirical methods can be used to study many distinct conformations of organic species.

In 2015, Martin *et al.*[129] investigated the accuracy of PM7 for modelling a series of 19 *in vivo* protein structures, containing 21-276 residues. Calculations were performed both in gas phase and implicit solvent via COSMO. The range of size of proteins studied made comparison of all-atom RMSDs uninformative, but energetic changes per atom were calculated and shown to be of the order of 1 kcal mol<sup>-1</sup>/atom *i.e.* small changes in peptide geometry introduced by PM7 optimisation. For the S22 and S88 benchmark data sets, the average change per atom was reported as 0.12 kcal mol<sup>-1</sup>/atom. However, comparison with B3LYP-D3 data showed that PM7 routinely underestimated inter-residue interaction distances.

Similarly, Ryan *et al.*[130] used PM7 to model the MTH1 enzyme (containing >150

residues, with both  $\alpha$ -helix and  $\beta$ -sheet structure present) and its interactions with a small biomolecule, 8-oxo-dGTP. PM7 reproduced experimental data well - errors in bond lengths and angles are of the order of 0.04 Å and 2-3° while the calculated all-atom RMSD value versus X-ray is 1.37 Å; backbone RMSD was 0.67 Å, while sidechain RMSD was 0.93 Å, indicating that most differences originate from sidechain movements. In addition, PM7 was shown to be able to predict the existence of charged sites in the protein as well as model interactions between them. Average errors in hydrogen bond distances of approximately 0.1 Å were observed, though description of long, weak hydrogen bonds was less accurate.

Hostas *et al.*[131] assessed the performance of PM6 and its modern variants (PM6-DH+, PM6-DH2(X) and PM6-D3H4X) as well as PM7 across thirteen standard data sets; eleven contain geometries and interaction energies of molecular complexes, while the two remaining sets contain structures and relative energies of different conformations of complexes and complex molecules, including peptides. While none of the methods could be unambiguously recommended for general modelling use, PM6-D3H4X displayed the most accurate results for non-covalent interactions such as hydrogen bonding. The PM7 method produced slightly worse results but was shown to describe other molecular properties more accurately. Notably, however, PM7 produced relatively poor results for dispersion-bound species, and over-stabilised systems containing multiple hydrogen bonds, such as biomolecules.

In addition, several authors have used these semi-empirical methods in combination with quantum mechanical methods in hybrid QM/MM schemes [132] as well as improving the accuracy of molecular docking results.[133, 134]

### 1.5.3 Molecular Mechanics and Molecular Dynamics

While there is relatively little quantum mechanical work on A $\beta$ , a great deal of research has been performed using molecular mechanics and molecular dynamics techniques. Molecular dynamics has become an increasingly important tool for the study of biomolecules – several general reviews of its application may be found in [135, 136], while A $\beta$ -specific examples may be found in reviews by Kepp[5, 49] and Rauk[29], among others.[50, 137] A selection of this literature is discussed here.

One of the first applications of molecular mechanics techniques to A $\beta$  was performed

by Mager in 1998.[138] This work did not consider the flexibility of the peptide, but rather performed a simple optimisation of the peptide using the MM+ forcefield. This identified two clear regions of secondary structure – an  $\alpha$ -helical region at residues Leu17-Ala21 and a  $\beta$ -sheet at Lys28-Val40. Since then, MM and MD methods have been extensively applied to studying the free A $\beta$  peptide.

In 2008, Anand *et al.*[139] applied standard MD and replica-exchange molecular dynamics (REMD) to the A $\beta$ 39 peptide at room temperature. Here, the authors used the Generalised-Born implicit solvent model and the AMBER99 forcefield. Dynamics simulations of 90 ns were used, following 0.1 ns of equilibration. Simulation data indicated that the A $\beta$  monomer does not occupy a single conformation in water, but adopts one of several low energy structures, consistent with the idea that A $\beta$  is a disordered peptide. Analysis of peptide secondary structure using DSSP showed little secondary structure content, with a turn at Val12-Leu17 and helices at Gly15-Val18, Gln11-Glu22 or Ile32-Met35, illustrated in Figure 1.8. Notably, there was no  $\beta$ -sheet structure formed, suggesting that  $\beta$ -sheet structures are formed during the oligomerisation processes and are not found in the monomer.

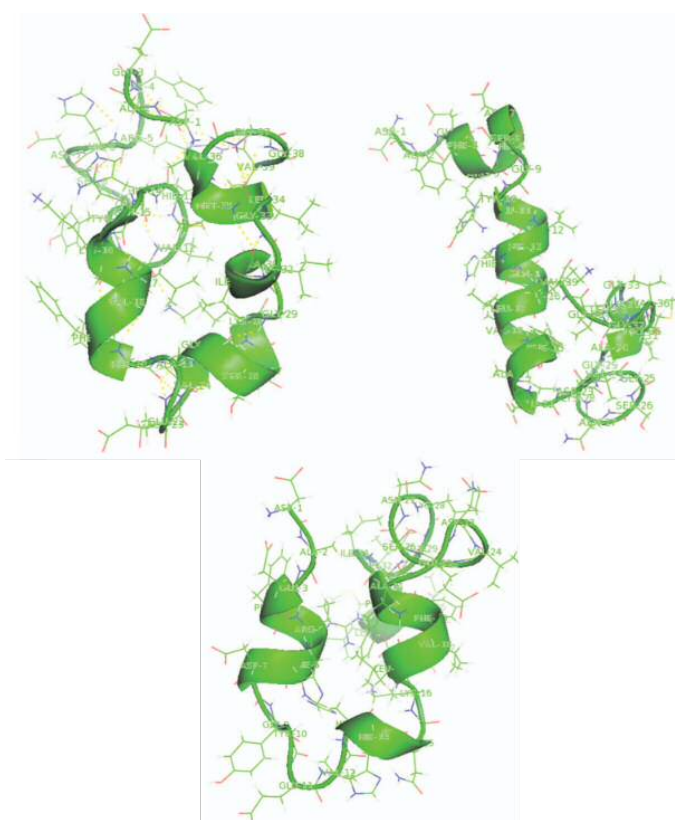


Figure 1.8: Three dominant structures observed in A $\beta$ 1-39 simulations[139].

Yang and Teplow[140] used microsecond timescale REMD simulations to study A $\beta$ 40 and A $\beta$ 42 and to construct their free energy surfaces. Here, simulations used Generalised - Born implicit solvent and the AMBER (PARM99SB) forcefield. The authors used 16 replicas, each for 110 ns, where the first 10 ns were treated as equilibration. Here, neither A $\beta$ 40 nor A $\beta$ 42 was found to be unstructured; while neither possessed high levels of  $\alpha$ -helix or  $\beta$ -sheet, each peptide has several independent folding units, containing residues Asp1–Arg5, Tyr10–His13, Leu17–Glu22, Lys28–Gly37, and Val39–Ala42, connected by four turns. Residues 6–40 were shown to exist as turn-type structure with 30–80% probability, while residues His6–Gly9 and Asp23–Asn27 were most frequently designated as turn-type. The central hydrophobic core and the C-terminal region of A $\beta$  often displayed  $\alpha$ -helical structure, while residues Lys16–Asn27 and Ala30–Val36 occupied  $\beta$ -sheet structure for 6–8% and 8–14% of simulation time. The free-energy surfaces of both peptides are characterised by two large basins, containing conformations with relatively substantial  $\alpha$ -helix or  $\beta$ -sheet content.

In 2007, Raffa and Rauk[141] performed a series of MD simulations on both A $\beta$ 42 and Cu<sup>II</sup>-A $\beta$ 42 – these metal-containing systems are discussed later. Simulations of A $\beta$ 42 were performed for 790 ns using the GROMOS forcefield with periodic boundary conditions and the simple point-charge (SPC) water model. The first 350 ns of simulation were considered as equilibration and remaining data used for analysis – the authors note that this equilibrated conformation may be an intermediate state, leading to more stable conformers at longer simulation times. The radius of gyration of A $\beta$ 42 was calculated as 9.2 Å, in line with experimental data.[142, 143] Analysis of the MD trajectories showed that the central core of A $\beta$  adopts predominantly turn-type structure, with  $\alpha$ -helical structure at residues Glu22–Asn27. Only a small degree of  $\beta$ -sheet structure was detected.

Sgourakis *et al.*[144] characterised A $\beta$ 40 and A $\beta$ 42 using REMD and NMR experiments and showed that REMD reproduced NMR data across microsecond timescale simulations. These simulations were performed using the OPLS forcefield and SPC water; 52 replicas were modelled for 100 ns each. A $\beta$ 40 simulations revealed the presence of a short 3<sub>10</sub> helix at the N-terminus and a hairpin at residues Val12–Val18. In contrast, the C-terminus of the peptide (Phe19–Val40) was shown to be conformationally flexible. In addition, the termini of the peptide were found to be in close proximity. A $\beta$ 42 data showed a disordered N-terminal tail, while residues

Ser8–Gly29 occupy a series of turn-type structure. Furthermore, the C terminus of A $\beta$ 42 is more structured than that of A $\beta$ 40, containing two short  $\beta$ -sheets. The authors speculate that this reduced C-terminal flexibility may be responsible for the higher propensity of A $\beta$ 42 to form amyloids.

Later, the same authors[145] used REMD to study conformations of A $\beta$ 42 in explicit (TIP4P) water using the AMBER99SB forcefield. Here, 52 replicas of the peptide were again used, each simulated for 225 ns. Results were shown to converge to a diverse ensemble representative of experimental data after approximately 60 ns/replica of simulation.  $\beta$ -sheet structures were frequently observed involving the C-terminus of the peptide - in particular, the authors note the formation of a  $\beta$ -sheet involving strands Phe4-His6 / Met35-Gly38 / Val18-Phe20 and Gly38-Val40. In addition, residues Ser8-Val12 and/or Phe20-Asp23 are shown to occupy  $\alpha$ -helical structure in some conformations, while a 3,10-helix was detected at both Gly29-Gly33 and Ala21-Val24 during simulation. Finally, turn-type structures appeared at Tyr10, Phe19 and Phe20.

In 2006, Luttmann and Fels[146] performed united-atom (*i.e.* where non-polar hydrogens are collapsed into their heavy atoms) simulations of A $\beta$ 40 and A $\beta$ 42. These simulations were performed for 10-20 ns at 300 K using the GROMOS forcefield, SPC water and periodic boundary conditions. While no explicit consideration of equilibration is given, only the final 10 ns of the simulations are used for analysis. While initial conformations (based on NMR data) contain fifteen to twenty residues of  $\alpha$ -helical character, this drops rapidly to between three and ten residues of  $\alpha$ -helix in conjunction with increased population of  $\pi$ -helices; it is generally accepted that transitions of  $\alpha$ -helices to 3,10- or  $\pi$ -helices are important steps in the folding and unfolding of helices.[147] In addition, the central region of the peptide was shown to adopt  $\beta$ -sheet and turn-type structure, while the first and final ten residues remain flexible.

Xu *et al.*[148] studied the conformational changes of A $\beta$ 40 via MD simulations in both aqueous solution and in a lipid environment. In this work, the 150 ns simulations used the GROMOS forcefield, SPC water and periodic boundary conditions. Six concurrent MD simulations illustrated that the initial  $\alpha$ -helical character at residues Val24-Gly37 of the peptide was rapidly converted into  $\beta$ -strand and turn-type structures, while the central hydrophobic region (residues Lys16-Ala21)

remained largely  $\alpha$ -helical, with some 3<sub>10</sub>- and  $\pi$ -helical character. Residues Asp1-His14 were reported as mostly random coil in nature. Overall, A $\beta$  was characterised as 36% helix, 12%  $\beta$ -sheet and 62% coil. The authors also suggest that residues Val24–Val40 may be the core for A $\beta$  oligomerisation.

Recently, Huy *et al.*[149] studied A $\beta$ 42 via MD. Here, the AMBER99SB forcefield was used alongside TIP3P water and simulations were carried out for 1  $\mu$ s. Interestingly, the authors define a quasi-equilibration, where the root-mean-square deviation (RMSD) of the system fluctuates around a stable value, rather than undertaking full equilibration of the disordered peptide. Of the initial structures studied, these systems reach quasi-equilibrium at different timescales. The simulations show that the peptide adopts a compact structure, with radius of gyration  $10.1 \pm 0.6$  Å, which agrees well with existing experimental data.[150] Analysis of peptide secondary structure showed that the turn- and coil-type content is high (85.8%), in line with other studies. In addition, the  $\alpha$ -helical content is very low (<5%), while the overall  $\beta$ -sheet content of A $\beta$ 42 is around 10%; both of these agree with existing data.[144, 151, 152] The authors also performed a thorough inspection of intramolecular salt-bridge interactions and frequently observed Lys28-Asp23, Lys16-Glu11 and Arg5-Asp1 contacts.

In 2013, Ball *et al.*[153] used the AMBER ff99SB forcefield with TIP4P water to study both A $\beta$ 40 and A $\beta$ 42 using REMD. While some similarities were observed, (*e.g.* a highly populated turn at Asp7-Ser8 or a helix near Ser26) the two peptides were shown to have substantially different secondary structures – in particular, A $\beta$ 40 occupied both collapsed and extended conformations. In addition, A $\beta$ 42 often showed two hairpin-type structures at Lys16-Ala21 and Gly29-Val36, as well as frequent turn-type structure at residues Ser26-Asn27.

Velez-Vega and co-workers[151] used OPLS REMD simulations with TIP3P water to model A $\beta$ 42 and mutant variations. Here, the authors used 32 replicas, starting from collapsed peptide conformations, for a total of 1.76  $\mu$ s simulation. This found that A $\beta$  is largely unstructured, with 41% coil structure, 52%  $\beta$ -bridge /  $\beta$ -turn, 4% helix and 3%  $\beta$ -sheet. A $\beta$ 42 displayed ‘a tendency to adopt a collapsed coil structure,’ including a helix in the central (residues Phe20-Ala30) peptide region. Further specific interactions included a hairpin between Gly33-Met35 and Val40-Ala42 or Ala30-Ile32 and Leu34-Val36, and N- and C-terminal interactions such as

Asp1-Ala42, Ala2-Val40, and Glu3-Val40.

Rosenmann *et al.*[154] used a combined NMR/MD method to study A $\beta$ 40 and A $\beta$ 42. Here, 52 replicas of REMD simulation performed with the OPLS forcefield and TIP3P water illustrated that simulation time on the order of ‘hundreds of ns’ is required to reach convergence. The N-terminus of both peptides was reported to be disordered, though turn-structures were observed at Phe4-His6 and His14-Lys16 in A $\beta$ 40. These N-terminal residues also appear to be in contact with the rest of the protein, suggesting a collapsed conformation is favoured. A $\beta$ 40 also possessed more  $\beta$ -sheet type character than A $\beta$ 42, particularly at Leu17-Ala21 and Ala30-Val36, while A $\beta$ 42 displayed a variety of hairpin structures near the C-terminus.

Later, these authors[155] studied the same A $\beta$  systems using three different REMD approaches: OPLS / TIP3P water, AMBER99 SB-ILDN / TIP4P water and CHARMM-22 / TIP3PSP water. Encouragingly, the resulting ensembles show ‘strong convergence in structural properties.’ In particular,  $\beta$ -sheet structures between Leu17-Ala21 and Ala30-Leu34 are commonly observed, while the central Ala21-Ala30 region rarely interacts with the rest of the peptide. However, the structure of this central region was shown to depend on the chosen forcefield. A $\beta$ 42 simulations also displayed a greater degree of hairpin-type structure near the C-terminus. In addition, OPLS simulations appear to favour especially collapsed conformations, with large numbers of intramolecular contacts and these contacts depend strongly on the forcefield used.

An interesting alternative approach was implemented by Mitternacht *et al.*[156] who used an all-atom potential where only torsional degrees of freedom were considered. Here, 10 independent Monte Carlo simulations at six different temperatures were performed in implicit solvent. This identified two free-energy basins; one corresponded to peptides with extended  $\beta$ -sheet character, while the other contained much less  $\beta$ -sheet character. The propensity of residues to form  $\alpha$ -helical structures was uniformly low, while  $\beta$ -sheet propensity was much higher, particularly at residues Leu17-Phe14. In addition, residues His14, Gly25 and Gly37 correspond to regions of turn-type structure.

Huy *et al.*[149] also used AMBER dynamics simulations on the  $\mu$ s timescale to study dimers of A $\beta$ 42. Here the peptides displayed 6.5%  $\alpha$ -helix and approximately 13%

$\beta$ -sheet structure, in agreement with other experiments. In addition, a high number of intrachain salt-bridges between N- and C- termini were observed. Similar results were obtained using coarse-grain[152] and discrete molecular dynamics simulations[157], with  $\beta$ -sheet propensity of 30% and 20%, respectively.

In addition to this, several authors have used coarse-grained molecular mechanics methods (*i.e.* where residues are represented by a small number of beads rather than their constituent atoms) to study A $\beta$ . [152, 158, 159] However, these results are not discussed in detail as the coarse-grain approach has not been used in this research.

In the context of this work, it will be important to compare the structure of the Pt<sup>II</sup>-A $\beta$  complexes with the naturally occurring peptide. However, it is important to note that the reported secondary structural data here depends on the forcefield chosen – different forcefields have different propensities to form different secondary structural elements, leading to significant structural differences and predicted behaviour.[49, 53].

Somavarapu *et al.*[53] assessed a range of 10 forcefields (of various AMBER, OPLS, GROMOS and CHARMM types) for their ability to reproduce NMR and CD structural data of the A $\beta$ 40 peptide and identified major differences in the structural ensembles. Simulations of approximately 1  $\mu$ s for each forcefield revealed that some over-stabilise helices, others stabilise  $\beta$ -strands, while OPLS variants produce highly unstructured conformations that generally display very poor agreement with the experimental data; CHARMM22\* and AMBER99sb-ILDN versions were found to most closely resemble experimental results.

Similar work by Carballo-Pacheco and Strodel[160] assessed a series of forcefields using REMD simulations with 32 replicas and 200 ns simulation each to model A $\beta$ 42. While all forcefields produced different peptide structures, they were also shown to accurately reproduce NMR data, though CHARMM22\* appeared to be most accurate.

Current MD simulations of the full A $\beta$  peptide take place on the nano- to micro-second timescale, with REMD simulations (*vide supra*) the longest of these. How-



ever, it is notable that in MD simulation work on A $\beta$  peptides, there has been little discussion of what constitutes equilibration, with the exception of work by Huy *et al.* above.[149] This must be in part due to the disordered nature of A $\beta$ , in that it is sufficiently conformationally diverse that it may never reach true equilibrium, or if it does, it does so on a timescale beyond what is feasible to study using current methods.

As discussed, Huy *et al.*[149] defined a quasi-equilibrium, where the system is stable enough to obtain useful structural information from the dynamics simulation, even if the structure is not truly at equilibrium. A $\beta$ 42 systems were shown to take on the order of hundreds of ns to reach this state. Other authors[141] explicitly state that the conformations sampled may be an intermediate folded form of the peptide, which may lead to more stable forms at longer simulation times.

Rauk[141] gives an equation for an estimate of the time scale,  $t$ , in ns, for folding a small protein:

$$\exp(0.5N^{(2/3)}) < t < \exp(1.5N^{(2/3)}) \quad (1.1)$$

where  $N$  is the number of residues. For A $\beta$ 42, this suggests a folding time of 420 ns to 74 ms. Work on the disordered nine-residue vasopressin peptide by Haensele *et al.* [161] illustrated that peptides of this size may not equilibrate within 25  $\mu$ s of metadynamics simulation, indicating that pursuing full equilibration of A $\beta$  is likely beyond current capability.

## 1.6 Modelling Metal-Biomolecule Systems

It is also important to consider modelling approaches to biomolecular systems containing metal ions, as this underpins much of the research in this thesis. This section consists of a discussion of some of these techniques and their application.

### 1.6.1 Quantum Mechanical Methods

For reasons of computational cost, quantum mechanical calculations are not usually applied to complete metal-biomolecule systems. Instead, the region near the metal is often modelled using quantum mechanical methods while the remainder of the system is handled using a cheaper method or omitted entirely. In this section,

recent applications of quantum mechanical methods to metal-biomolecule systems will be discussed. General reviews of the topic are provided in references.[162–164]

As already mentioned, Streltsov *et al.* [89] used DFT (MPW1PW91 / SDD) calculations alongside EXAFS data to derive structural models of the interactions of cisplatin and  $\text{Pt}^{\text{II}}(\text{phen})$  complexes with the histidine and methionine residues of  $\text{A}\beta$ . This level of theory was judged appropriate based on modelling small model compounds from the crystal structure database –  $\text{Pt}^{\text{II}}$  bond distances and  $\text{Pt}^{\text{II}}$  bond angles agree with the experimental data values to 0.14 Å or better and 3.60° or better, respectively.

Recently, Novato *et al.*[165] used DFT to study interactions of  $\text{Pt}^{\text{II}}(\text{phen})$  and  $\text{Pt}^{\text{II}}(\text{dpphen})$  (**1** and **2**, respectively, in Figure 1.5) complexes with Met, Lys, His, Glu, Asp, Asn and Gln residues of  $\text{A}\beta$ . Here, B3LYP calculations with LanL2DZ / 6-31G(d) basis sets were used, while reaction kinetics were calculated using LC- $\omega$ PBE and CAM-B3LYP with SDD / 6-31+G(2df) basis sets. DFT accurately reproduced experimental geometries – while bond lengths were slightly longer than those observed in the solid state, bond angles agree within 5%. Results suggest a reactivity order for these complexes of: Met  $\sim$  Lys  $\sim$  His( $\epsilon$ )  $>$  Glu  $\sim$  Asp  $\gg$  Asn  $\sim$  Gln; thermodynamics data suggests a preference for Lys or His residues, in line with other experimental data.

Zimmerman *et al.* have published a series of calculations on the interactions of cisplatin with cysteine and methionine amino acid residues, both in gas phase [166] and in implicit solvent[167, 168]. In this work, hydrated cisplatin complexes were used in a two step reaction with the amino acid – substitution of aqua ligand by the amino acid before dissociative chelate formation. DFT optimisations were performed at B3LYP / SDD / 6-31+G(d) level in implicit solvent, where the SDD ECP was applied to sulphur, chlorine and platinum. Single-point energies were then calculated at the B3LYP / SDD / 6-311++G(2df,2pd) level. Results suggest a thermodynamic preference for amino acid coordination: most favourable is binding of cysteine sulphur, followed by amino group nitrogen, methionine thioether sulphur, and carboxyl group oxygen. More specifically, in the case of Cys coordination, coordination via sulphur is strongly favoured. For Met coordination, coordination depends on the nature of the cisplatin complex –  $\text{Pt}^{\text{II}}\text{-N}(\text{Met})$  coordination is preferred in the chloro complex, but  $\text{Pt}^{\text{II}}\text{-S}(\text{Met})$  in the hydroxo complex.

Ali-Torres *et al.* [169] used homology modelling and quantum mechanics to determine three-dimensional models for  $\text{Cu}^{\text{II}}$ -A $\beta$ 16 binding. Here, both  $\epsilon$  and  $\delta$  coordination of histidines 6, 13, and 14 were considered, with Asp1, Glu3, Asp7, Glu11, and CO-Ala2 as the fourth ligand. Gas phase DFT calculations (BHandHLYP / 6-31G++G(d,p)) predicted  $\text{Cu}^{\text{II}}$  bound to Asp7<sub>O</sub>, His6<sub>N $\epsilon$</sub> , His13<sub>N $\delta$</sub> , His14<sub>N $\delta$</sub> , while inclusion of implicit solvent lead to CO-Ala2<sub>CO</sub>, His6<sub>N $\epsilon$</sub> , His13<sub>N $\delta$</sub> , His14 binding<sub>N $\epsilon$</sub> . Alternatively, hybrid QM/MM calculations (BHandHLYP: UFF) calculations suggest a different  $\text{Cu}^{\text{II}}$  binding mode: Glu3<sub>O</sub>, His6<sub>N $\delta$</sub> , His13<sub>N $\epsilon$</sub> , His14<sub>N $\epsilon$</sub> . While each method suggests a different copper coordination sphere, results indicate a preference for oxygen ligands (Glu3, Asp7) rather than backbone nitrogen atoms.

Marino *et al.*[170] aimed to determine coordination patterns of  $\text{Cu}^{\text{II}}$  and  $\text{Zn}^{\text{II}}$  to A $\beta$  via DFT calculations on truncated model systems. Here, B3LYP with 6-31G(d) and 6-31+G(d) basis sets were used, with SDD ECP for metal centres. Results indicate that while  $\text{Cu}^{\text{II}}$  prefers penta-coordination,  $\text{Zn}^{\text{II}}$  is flexible – occupying five- and four- coordinate states. Furthermore, the most stable  $\text{Cu}^{\text{II}}$  coordination was via two equatorial His residues, an axial Tyr and two waters, while for  $\text{Zn}^{\text{II}}$ , the same residues with a single water molecule appeared most stable.

Azimi and Rauk [171] used B3LYP/6-31+G(d) and MP2/6-311+(2df,2p) calculations to study  $\text{Cu}^{\text{I}}$  and  $\text{Cu}^{\text{II}}$  interactions with models of the Asp1, Ala2, His6, and His13/His14 regions of A $\beta$ . These data indicate that  $\text{Cu}^{\text{II}}$  binds Asp1, His6 and either His13 or His14 (component I), while component II  $\text{Cu}^{\text{II}}$  binds backbone oxygen in Ala2, His6, His13, and His14. Free energy calculations suggest that  $\text{Cu}^{\text{I}}$  binds these sites in A $\beta$  more strongly than  $\text{Cu}^{\text{II}}$ . Finally, the computed reduction potential for  $\text{Cu}^{\text{II}}$  closely matches experimental data.

Streltsov *et al.* [172] combined EXAFS and DFT data to study  $\text{Cu}^{\text{II}}$ -A $\beta$ 16 binding. In this case, DFT calculations at the B3LYP / LanL2DZ level were used on hexa-coordinate  $\text{Cu}^{\text{II}}$ , bound to three His residues and either Glu and one water molecule or Tyr and two water molecules. Fitting and refinement against EXAFS data displayed better agreement with the 3 His, Glu, H<sub>2</sub>O ligand arrangement than that containing Tyr.

Dudev and Lim [173] used DFT calculations to study  $\text{Zn}^{\text{II}}$ ,  $\text{Mg}^{\text{II}}$  and  $\text{Ca}^{\text{II}}$  binding affinity and selectivity for nonstandard amino acid residues. Here, the metal and first coordination sphere was optimised using the S-VWN functional with SDD / 6-31+G\* basis set combination, before free energy calculation at the B3LYP / 6-311++G(2df,2p) level. The remainder of each protein was treated as a continuum dielectric. While this is a simplistic modelling approach, results suggest that the nonstandard residues have greater metal binding capacity than their standard counterparts. These residues display preference for  $\text{Zn}^{\text{II}}$  over  $\text{Mg}^{\text{II}}$  or  $\text{Ca}^{\text{II}}$ , suggest they could be used as metal-binding sites in protein engineering.

The same authors [174] also studied the competition between protein ligands and cytoplasmic inorganic anions for  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Mg}^{2+}$  and  $\text{Ca}^{2+}$  ions via DFT/continuum methods. S-VWN / SDD / 6-31+G\* optimisation and B3LYP / 6-311++G(2df,2p) free energy calculations indicate that the metal cation is able to bind its target protein despite a high concentration of cytoplasmic anions. This is because desolvating a carboxylate is less energetically expensive than desolvating an inorganic anion, while the protein acts as a polydentate ligand, providing a chelate effect.

Similarly, the same authors studied the competition between  $\text{Li}^+$  and  $\text{Mg}^{2+}$  in metalloproteins using B3LYP / 6-31+G(3d,p) calculations.[175] Here, metal binding sites were simplified to acetate, amide and backbone groups, while the remainder of the peptide was not considered. These calculations suggest that competition between  $\text{Mg}^{2+}$  and  $\text{Li}^+$  depends on the net charge of the metal complex and the solvent exposure of the binding site; in  $\text{Mg}^{2+}$ -binding proteins, the binding site is typically solvent-inaccessible and lined by Asp/Glu residues.

As modelling transition metals often requires electronic structure methods, hybrid QM/MM approaches are often used to study metal-biomolecule systems. Of particular interest is the work on  $\text{Pt}^{\text{II}}$ -DNA systems by Gkionis and coworkers.[176–178]

## 1.6.2 Semi-Empirical Methods

In this section, the application of semi-empirical calculation methods to metal-biomolecule systems is summarised. While this is not a new approach, there is little work using this model to model many conformations of large flexible peptide

units, as seen in this research. In addition, the newer PM7 method, despite its improved parameterisation, has not yet been widely applied to metal-biomolecule systems.

In 2010, Stewart[121] demonstrated the applicability of PM6 to modelling biomolecules, including a series of metalloproteins. Here, the metal ions studied included copper, zinc, iron, cobalt, molybdenum, manganese, magnesium and potassium. In these large metallo-peptide systems, PM6 predicted geometries of both the covalently-bound and free metal ions in agreement with X-ray crystallographic data; associated coordination bond lengths agreed with experiment, even for the electronically difficult  $\text{Cu}^{\text{II}}$  systems. In terms of overall metallo-peptide geometries, PM6 also performed well – successfully reproducing peptide secondary structure while RMSD from experimental structure was typically below 1 Å.

In 2011, Xia *et al.*[179] used gas-phase PM6 and B3LYP/6-31G\* calculations to model a cyclodextrin complex with  $\text{Pt}^{\text{II}}$  trans-dichloro (dipyridine), a useful but poorly bioavailable anti-tumour agent. Multiple orientations of the host-guest interaction were considered; while both methods predicted the same  $\text{Pt}^{\text{II}}$ -complex – cyclodextrin orientation, PM6 did not accurately reproduce the difference in relative energies.

Bertoli *et al.*[180] used PM6 and DFT (B3LYP/LANL2DZ) calculations to investigate  $\text{Zn}^{\text{II}}$ -citrate isomers in collaboration with experiment. In addition, the authors calculated thermodynamic properties of these zinc complexes. PM6 was shown to predict  $\text{Zn}^{\text{II}}$  geometry with similar accuracy to DFT, and both showed good agreement with experiment. The semi-empirical method also correctly predicted energy ordering of the isomers of the complex (using DFT as a reference), though the relative energy differences were smaller than in DFT.

In similar work, the same authors [181, 182] studied  $\text{Cd}^{\text{II}}$ -,  $\text{Pb}^{\text{II}}$ -,  $\text{Cu}^{\text{II}}$ - and  $\text{Fe}^{\text{II}}$ -citrate complexes with PM6. Calculation of coordination isomers predicted stable structures that agreed with experimental data and allowed for assignment of FT-IR spectra. While PM6 correctly predicted energy ordering, the relative energies were not correct – PM6 provided less exoenergetic thermodynamic stabilities of all species.

A series of work by Keglevich and co-workers[183, 184] used PM6, PM6-DH2 and DFT to study geometries of  $\text{Pt}^{\text{II}}$  complexes with phosphine borane ligands. While DFT calculations showed the best agreement with experimental data, both semi-empirical methods also modelled these transition metal complexes accurately, including the formation of a weak  $\pi$ -stacking interaction.

Shahabadi and Heidari[185] employed PM6 calculations to characterise a  $\text{Pt}^{\text{II}}$  (met-formin) complex in the context of DNA binding. Here, the semi-empirical method was shown to correctly predict the structure of the complex, as well accurately predicting vibrational frequencies – including those around the platinum – as compared with FT-IR spectra.

Suarez *et al.*[186, 187] used short (100 - 500 ps) PM6 dynamics simulations in implicit solvent to study the structure and thermodynamics of molybdocene-cysteine and molybdocene-glutathione complexes. These semi-empirical dynamics produced a range of conformations of the metal-peptide structure; single-point DFT (B3LYP / aug-cc-pVDZ) calculations were then used to determine the relative stabilities of these conformers. In both cases, while calculated geometries were accurate, PM6 relative energies did not agree with experimental data.

Fernandes *et al.*[188] used semi-empirical PM6 calculations to characterise the metal binding sites of the heparinase II enzyme, with  $\text{Zn}^{\text{II}}$  and  $\text{Ca}^{\text{II}}$ . For the gas-phase  $\text{Zn}^{\text{II}}$  case, favourable metal coordination was shown to agree with existing experimental data; for  $\text{Ca}^{\text{II}}$ , PM6 predicted much weaker bonds to the peptide, which the authors suggest could explain the difference in enzyme activity in the presence of the different metals. Explicit inclusion of water molecules around the metal binding site did not perturb  $\text{Zn}^{\text{II}}$  binding, while  $\text{Ca}^{\text{II}}$  showed weak interactions with two solvent molecules.

In 2014, Shahabadi [189] used PM6 to study the structure of a  $\text{Cu}^{\text{II}}$ -aspartame complex. Here, PM6 correctly modelled the individual copper coordination bonds in agreement with experimental data, but also successfully described the Jahn-Teller distortion at the metal, giving a distorted octahedral geometry.

De Santis *et al.*[190] used PM6/COSMO calculations to study the interaction of a

21-membered cyclopeptoid with  $\text{Ca}^{\text{II}}$ ,  $\text{Ba}^{\text{II}}$ ,  $\text{Sr}^{\text{II}}$  and  $\text{Mg}^{\text{II}}$  metals. These calculations showed that the peptoid favoured backbone carbonyl coordination, with either 4 or 6 coordinate metals. Importantly, PM6 predictions of coordination structure were shown to agree with both NMR and CD spectra.

Caturello *et al.*[191] used PM6-D3H4X in vacuum to study supramolecular complexes, containing repeating ‘layers’ of platinum and palladium complexes with organic ligands, where the supramolecular arrangement contained  $\pi$ -stacking arrangements. The predicted structures were shown to be in good agreement with experimental results, further indicating that the semi-empirical methods are capable of modelling complex species containing transition metals. In addition, PM6-D3H4X calculations were employed to study the thermodynamics of the self-aggregation of individual complexes to the supramolecular whole, and suggested that a single supramolecular aggregate is constructed from two smaller columns of monomer units.

Notably, there is some work which uses semi-empirical approaches within molecular docking simulations of metal-biomolecules[192] but as before, as the underlying technique is not used in this work, the research is not discussed in detail.

### 1.6.3 Molecular Mechanics-Based Methods

Modelling transition metal species without resorting to computationally expensive quantum mechanical or semi-empirical methods is an attractive proposition, and has been thoroughly explored in the past. The application of these molecular mechanics based methods to transition metal-biomolecule species is less well studied however - a selection of this literature is discussed below:

In 1992, Rappe *et al.*[193] developed the Universal Force-Field (UFF), capable of modelling the complete periodic table. While the general parameterisation of this approach limits its accuracy, it was shown to produce accurate geometries of simple transition metal complexes as well as larger organic species.

A detailed review of *ad hoc* parameters applied to  $\text{Pt}^{\text{II}}$ -DNA complexes may be found in [194]. Of note is work by Scheeff *et al.*[195], who developed AMBER parameters to study cisplatin and oxaliplatin adducts with a double-helical DNA

strand, based on crystal structure data. This built on previous work by Yao *et al.*[196] These parameters produced heavy atom RMSD values of approximately 1 Å from the starting crystal data. Over a 200 ps molecular dynamics simulation, these parameters successfully predicted the platinum coordination geometry, though some deviations from expected behaviour were observed in the DNA structure; in particular, the cyclohexane ring of oxaliplatin distorted the DNA major groove.

Chan *et al.*[197] studied cadmium- and zinc-containing metallothioneins using the MM2 forcefield, augmented using metal parameters from the CACHE program, which estimates forcefield parameters for all elements. Short molecular dynamics simulations of these systems revealed that metal-geometry and overall peptide configuration show good agreement with spectroscopic data.

More recently, Xiang and Ponder[198] introduced the angular overlap model (see Section 2.5.2) to the polarisable AMOEBA forcefield to study two type I Cu<sup>II</sup> proteins. Parameters were obtained by fitting molecular mechanics data to B2PLYP-D/cc-pVDZ geometry optimised structures and MP2/cc-pVDZ energies. This produced good agreement with *ab initio* results, though AMOEBA-AOM failed to reproduce the trans effect of some ligands. Here Cu<sup>II</sup> was coordinated to two histidines, one cysteine and one methionine residue. MD simulations of the copper-proteins (2 ns simulation each) showed that the Cu<sup>II</sup> geometry was very stable, and addition of the copper parameters did not lead to disruption of the remaining peptide system.

In 2015, Duarte *et al.*[199] used a dummy atom model to study Mn<sup>2+</sup>, Zn<sup>2+</sup>, Mg<sup>2+</sup>, Ca<sup>2+</sup>, Ni<sup>2+</sup>, Co<sup>2+</sup>, and Fe<sup>2+</sup> interactions with two glyoxalase enzymes. Here, the metal centre is described by a set of cationic dummy atoms around the central metal in the specific coordination geometry to be attained – in this case octahedral. Dummy models do not contain explicit bonds from metal to ligands, allowing for ligand exchange and conversion between different coordination geometries. Molecular dynamics simulations do not show any ligand exchange, while the calculated metal-oxygen coordination bonds lie within the range of values seen in experiment, with errors of approximately 0.04 Å. In addition, metal-peptide systems were very stable, with little structural deviation observed.

In 2015, Liao *et al.*[200] introduced a dummy atom model for Zn<sup>II</sup> and Cu<sup>II</sup> that is able to model Jahn-Teller distortion. This model was implemented in GROMACS.



Zn<sup>II</sup> and Cu<sup>II</sup> parameters were tested in 100 ns MD simulations of two biomolecules - A $\beta$ 16 and the enzyme superoxide dismutase, and correctly predict distorted square planar geometry in Cu<sup>II</sup> binding sites, while reported bond lengths are in close agreement with experiment. The Zn<sup>II</sup> model performs less well for these examples, with large changes in metal coordination geometry.

Of particular relevance to this thesis is the work of Huy *et al.*[149], who developed AMBER parameters for Cu<sup>II</sup>-A $\beta$  complexes and performed microsecond timescale MD calculations using them. Here, both A $\beta$  monomers and dimers were studied, where Cu<sup>II</sup> was coordinated to His6, His13 (or His14), and Asp1 with distorted planar geometry. Cu<sup>II</sup> parameters were obtained from small model systems via UB3LYP calculations in implicit solvent, using the 6-31G(d) basis for main group elements and the SDD ECP for Cu<sup>II</sup>. Pseudo-equilibration of the metal-peptide system was reached after several hundred nanoseconds of simulation; this revealed significant changes in the peptide salt-bridge and hydrogen-bond network upon Cu<sup>II</sup> coordination, as well as changes in peptide secondary structure and radius of gyration.

Similarly, Raffa and Rauk[141] developed their own Cu<sup>II</sup> parameters to model Cu<sup>II</sup>-A $\beta$  interactions. Here, Cu<sup>II</sup> parameters were obtained from B3LYP/6-31G(d) level optimisations of small model structures, while Cu<sup>II</sup> was bound via a water ligand, His13, His14 and a backbone nitrogen / oxygen atom. In this case, the metal-peptide system equilibrated after 350 ns of simulation, and subsequent data showed changes in peptide structure dependent on the exact metal coordination mode.

While the above authors have found some success in modelling metal-peptide interactions by developing *ad hoc* metal parameters, an alternative approach, Ligand Field Molecular Mechanics, was pioneered by Burton and Deeth[201] (see Section 2.5.2), which has also been used in the field of metal-peptide interactions:

Deeth developed LFMM parameters based on DFT modelling of homoleptic model systems to study the binding of type I copper centres in proteins.[202] Notably, this was the first application of LFMM to second-row donor atoms such as sulphur; the parameters used to model these interactions successfully reproduce the trans influence of these ligands, leading to a lengthening of the pseudo-trans copper-ligand distance. In addition, this LFMM approach was used to study five complete copper-

containing proteins. LFMM optimisations of these structures were performed in a 5 Å thick shell of water molecules, and resulting structures display root-mean-square-deviations (RMSD) from experimental data of less than 0.42 Å.

Following this, in 2007 Deeth used LFMM to model more than twenty copper-containing proteins, with either four- or five-coordinate copper centres.[203] In this work, LFMM parameters for Cu-N(imidazole), Cu-S(thiolate), Cu-S(thioether), and Cu-O(carbonyl) interactions were developed on the basis of experimental and theoretical data such that LFMM calculations of model copper-peptide binding sites agree structurally and energetically with DFT (PW91 / TZP / DZP) calculations. Here, LFMM optimisations of the complete systems were completed in a 10 Å deep layer of water and results compared to existing experimental data; LFMM structures reproduced the entire peptide geometry to “well within experimental error.” Overall, this LFMM approach is able to provide accuracy equivalent to existing QM/MM calculations, only several orders of magnitude faster.

Later, Deeth applied the LFMM model to a copper- tyrosinase system using the AMBER forcefield.[204] Deviations between DFT (BP86/TZP) and LFMM bond lengths are of the order of 0.02 Å, further highlighting the accuracy of LFMM. LFMM also reproduced relative energies of different metal binding site geometries. Two short (16 ns) NVT Ligand Field Molecular Dynamics (LFMD) simulations at 300 K were then performed on the entire copper-peptide system, and the structure rapidly equilibrated. This equilibrium geometry was compared to X-ray data, giving RMSD values of all  $\alpha$ -carbons of only 1.16 Å. RMSF data in these trajectories also suggested that the coordinating His residues are less able to move than other residues, though differences are not pronounced.

Work by Mutter *et al.*[205] developed LFMM parameters for Cu<sup>II</sup>-A $\beta$ 16 systems. Parameters were validated against B3LYP-D2/6-31G(d)/LANL2DZ optimised structures on small models of the metal-A $\beta$  interaction, and LFMM bond lengths and angles were shown to agree to within 0.1 Å and 5° of DFT results. These parameters were then tested in a series of short LFMD simulations; while this is too short to model peptide behaviour, it is the first LFMM-based foray into Cu<sup>II</sup>-A $\beta$  coordination geometry.

Following the introduction of a series of LFMM parameters for Pt<sup>II</sup> complexes [206],

Tai *et al.*[207] applied the LFMM approach to study DNA distortions induced by  $\text{Pt}^{\text{II}}$  complexes. LFMM descriptions of  $\text{Pt}^{\text{II}}$ -guanine interactions were validated against experimental and DFT data before LFMD simulations were performed. Here, 6 ns of production data was obtained for the system under periodic boundary conditions and TIP3P water via DL-POLY-LF. Local metal geometry showed good agreement with experimental data, while DNA was shown to bend near the minor groove.

## Bibliography

- (1) *World Alzheimer Report 2015*, <http://www.worldalzreport2015.org/> (visited on 12/14/2017).
- (2) *WHO | Dementia*, WHO, [www.who.int/mediacentre/factsheets/fs362/en/](http://www.who.int/mediacentre/factsheets/fs362/en/) (visited on 12/14/2017).
- (3) K. P. Kepp, *Progress in Neurobiology*, 2016, **143**, 36–60.
- (4) R. Brookmeyer, E. Johnson, K. Ziegler-Graham and H. M. Arrighi, *Alzheimer's and Dementia: The Journal of the Alzheimer's Association*, 2007, **3**, 186–191.
- (5) K. P. Kepp, *Chemical Reviews*, 2012, **112**, 5193–5239.
- (6) S. Karantzoulis and J. E. Galvin, *Expert Review of Neurotherapeutics*, 2011, **11**, 1579–1591.
- (7) K. S. Kosik, C. L. Joachim and D. J. Selkoe, *Proceedings of the National Academy of Sciences of the United States of America*, 1986, **83**, 4044–4048.
- (8) J. Hardy and G. Higgins, *Science*, 1992, **256**, 184–185.
- (9) G. E. Gibson and H.-M. Huang, *Neurobiology of Aging*, 2005, **26**, 575–578.
- (10) K. Honda, G. Casadesus, R. B. Petersen, G. Perry and M. A. Smith, *Annals of the New York Academy of Sciences*, 2004, **1012**, 179–182.
- (11) L. E. Scott and C. Orvig, *Chemical Reviews*, 2009, **109**, 4885–4910.
- (12) M. A. Greenough, J. Camakaris and A. I. Bush, *Neurochemistry International*, 2013, **62**, 540–555.
- (13) H. Kozlowski, M. Luczkowski, M. Remelli and D. Valensin, *Coordination Chemistry Reviews*, 2012, **256**, 2129–2141.
- (14) S. Chakroborty and G. E. Stutzmann, *Science China Life Sciences*, 2011, **54**, 752–762.
- (15) A. I. Bush, *Trends in Neurosciences*, 2003, **26**, 207–214.
- (16) A. I. Bush and R. E. Tanzi, *Neurotherapeutics*, 2008, **5**, 421–432.
- (17) D. Valensin, C. Gabbiani and L. Messori, *Coordination Chemistry Reviews*, 2012, **256**, 2357–2366.

- (18) C. Reitz, T. den Heijer, C. van Duijn, A. Hofman and M. Breteler, *Neurology*, 2007, **69**, 998–1005.
- (19) A. Ruitenberg, J. C. v. Swieten, J. C. Witteman, K. M. Mehta, C. M. v. Duijn, A. Hofman and M. M. Breteler, *The Lancet*, 2002, **359**, 281–286.
- (20) A. Ott, R. P. Stolk, A. Hofman, F. v. Harskamp, D. E. Grobbee and M. M. B. Breteler, *Diabetologia*, 1996, **39**, 1392–1397.
- (21) A.-M. Tolppanen, A. Solomon, J. Kulmala, I. K  reholt, T. Ngandu, M. Rusanen, T. Laatikainen, H. Soininen and M. Kivipelto, *Alzheimer’s & Dementia: The Journal of the Alzheimer’s Association*, 2015, **11**, 434–443.e6.
- (22) B. Singh, A. K. Parsaik, M. M. Mielke, P. J. Erwin, D. S. Knopman, R. C. Petersen and R. O. Roberts, *Journal of Alzheimer’s Disease*, 2014, **39**, 271–282.
- (23) J.-C. Lambert, C. A. Ibrahim-Verbaas, D. Harold, A. C. Naj, R. Sims, C. Bellenguez, G. Jun, A. L. DeStefano, J. C. Bis, G. Russo, T. A. Thornton-Wells, N. Jones, A. V. Smith, V. Chouraki, C. Thomas, M. A. Ikram, D. Zelenika, B. N. Vardarajan, Y. Kamatani, C.-F. Lin, A. Gerrish, H. Schmidt, B. Kunkle, M. L. Dunstan, A. Ruiz, M.-T. Bihoreau, S.-H. Choi, C. Reitz, F. Pasquier, P. Hollingworth, A. Ramirez, O. Hanon, A. L. Fitzpatrick, J. D. Buxbaum, D. Campion, P. K. Crane, C. Baldwin, T. Becker, V. Gudnason, C. Cruchaga, D. Craig, N. Amin, C. Berr, O. L. Lopez, P. L. De Jager, V. Deramecourt, J. A. Johnston, D. Evans, S. Lovestone, L. Letenneur, F. J. Moron, D. C. Rubinsztein, G. Eiriksdottir, K. Sleegers, A. M. Goate, N. Fievet, M. J. Huentelman, M. Gill, K. Brown, M. I. Kamboh, L. Keller, P. Barberger-Gateau, B. McGuinness, E. B. Larson, R. Green, A. J. Myers, C. Dufouil, S. Todd, D. Wallon, S. Love, E. Rogaeva, J. Gallacher, P. St George-Hyslop, J. Clarimon, A. Lle  , A. Bayer, D. W. Tsuang, L. Yu, M. Tsolaki, P. Bossu, G. Spalletta, P. Proitsi, J. Collinge, S. Sorbi, F. Sanchez-Garcia, N. C. Fox, J. Hardy, M. C. D. Naranjo, P. Bosco, R. Clarke, C. Brayne, D. Galimberti, M. Mancuso, F. Matthews, S. Moebus, P. Mecocci, M. Del Zompo, W. Maier, H. Hampel, A. Pilotto, M. Bullido, F. Panza, P. Caffarra, B. Nacmias, J. R. Gilbert, M. Mayhaus, L. Lannfelt, H. Hakonarson, S. Pichler, M. M. Carrasquillo, M. Ingelsson, D. Beekly, V. Alvarez, F. Zou, O. Valladares, S. G. Younkin, E. Coto, K. L. Hamilton-Nelson, W. Gu, C. Razquin, P. Pastor, I. Mateo, M. J. Owen, K. M. Faber, P. V. Jonsson, O. Combarros, M. C. O’Donovan, L. B. Cantwell, H. Soininen, D. Blacker, S. Mead, T. H. Mosley, D. A. Bennett, T. B. Harris, L. Fratiglioni, C. Holmes, R. F. A. G. de Bruijn, P. Passmore, T. J. Montine, K. Bettens, J. I. Rotter, A.

- Brice, K. Morgan, T. M. Foroud, W. A. Kukull, D. Hannequin, J. F. Powell, M. A. Nalls, K. Ritchie, K. L. Lunetta, J. S. K. Kauwe, E. Boerwinkle, M. Riemenschneider, M. Boada, M. Hiltunen, E. R. Martin, R. Schmidt, D. Rujescu, L.-S. Wang, J.-F. Dartigues, R. Mayeux, C. Tzourio, A. Hofman, M. M. Noethen, C. Graff, B. M. Psaty, L. Jones, J. L. Haines, P. A. Holmans, M. Lathrop, M. A. Pericak-Vance, L. J. Launer, L. A. Farrer, C. M. van Duijn, C. Van Broeckhoven, V. Moskvina, S. Seshadri, J. Williams, Gerard D. Schellenberg and P. Amouyel, *Nature Genetics*, 2013, **45**, 1452–U206.
- (24) T. Lu, Y. Pan, S.-Y. Kao, C. Li, I. Kohane, J. Chan and B. A. Yankner, *Nature*, 2004, **429**, 883.
- (25) P. Sorrentino, A. Iuliano, A. Polverino, F. Jacini and G. Sorrentino, *FEBS Letters*, 2014, **588**, 641–652.
- (26) J. Hardy and D. J. Selkoe, *Science*, 2002, **297**, 353–356.
- (27) C. Masters, G. Simms, N. Weinman, G. Multhaup, B. McDonald and K. Beyreuther, *Proceedings of the National Academy of Sciences of the United States of America*, 1985, **82**, 4245–4249.
- (28) F. Chiti and C. M. Dobson, *Annual Review of Biochemistry*, 2006, **75**, 333–366.
- (29) A. Rauk, *Chemical Society Reviews*, 2009, **38**, 2698–2715.
- (30) M. K. Tiwari and K. P. Kepp, *Journal of Alzheimer’s Disease*, 2015, **47**, 215–229.
- (31) A. K. Somavarapu and K. P. Kepp, *ACS Chemical Neuroscience*, 2015, **6**, 1990–1996.
- (32) K. N. Dahlgren, A. M. Manelli, W. B. Stine, L. K. Baker, G. A. Krafft and M. J. LaDu, *Journal of Biological Chemistry*, 2002, **277**, 32046–32053.
- (33) A. L. Hansen and L. E. Kay, *Proceedings of the National Academy of Sciences of the United States of America*, 2014, **111**, E1705–E1712.
- (34) G. G. Glenner and C. W. Wong, *Biochemical and Biophysical Research Communications*, 1984, **120**, 885–890.
- (35) A. F. Teich and O. Arancio, *Biochemical Journal*, 2012, **446**, 165–177.
- (36) J. Hardy, *Journal of Alzheimer’s Disease*, 2006, **9**, 151–153.
- (37) D. Selkoe, *Neuron*, 1991, **6**, 487–498.
- (38) D. M. Walsh, I. Klyubin, J. V. Fadeeva, W. K. Cullen, R. Anwyl, M. S. Wolfe, M. J. Rowan and D. J. Selkoe, *Nature*, 2002, **416**, 535–539.
- (39) C. Haass and B. De Strooper, *Science*, 1999, **286**, 916–919.
- (40) J. P. Cleary, D. M. Walsh, J. J. Hofmeister, G. M. Shankar, M. A. Kuskowski, D. J. Selkoe and K. H. Ashe, *Nature Neuroscience*, 2005, **8**, 79–84.

- (41) H. J. Aizenstein, R. D. Nebes, J. A. Saxton, J. C. Price, C. A. Mathis, N. D. Tsopelas, S. K. Ziolk, J. A. James, B. E. Snitz, P. R. Houck, W. Bi, A. D. Cohen, B. J. Lopresti, S. T. DeKosky, E. M. Halligan and W. E. Klunk, *Archives of Neurology*, 2008, **65**, 1509–1517.
- (42) F. H. Bouwman, N. S. M. Schoonenboom, N. A. Verwey, E. J. v. Elk, A. Kok, M. A. Blankenstein, P. Scheltens and W. M. v. d. Flier, *Neurobiology of Aging*, 2009, **30**, 1895–1901.
- (43) C. R. Jack, D. S. Knopman, W. J. Jagust, L. M. Shaw, P. S. Aisen, M. W. Weiner, R. C. Petersen and J. Q. Trojanowski, *The Lancet. Neurology*, 2010, **9**, 119–128.
- (44) C. Haass and D. J. Selkoe, *Nature Reviews Molecular Cell Biology*, 2007, **8**, 101.
- (45) F. Hefti, W. F. Goure, J. Jerecic, K. S. Iverson, P. A. Walicke and G. A. Krafft, *Trends in Pharmacological Sciences*, 2013, **34**, 261–266.
- (46) C. G. Glabe, *Neurobiology of Aging*, 2006, **27**, 570–575.
- (47) L. W. Hung, G. D. Ciccotosto, E. Giannakis, D. J. Tew, K. Perez, C. L. Masters, R. Cappai, J. D. Wade and K. J. Barnham, *Journal of Neuroscience*, 2008, **28**, 11950–11958.
- (48) C. Schmitz, B. P. F. Rutten, A. Pielen, S. Schäfer, O. Wirths, G. Tremp, C. Czech, V. Blanchard, G. Multhaup, P. Rezaie, H. Korr, H. W. M. Steinbusch, L. Pradier and T. A. Bayer, *The American Journal of Pathology*, 2004, **164**, 1495–1502.
- (49) K. P. Kepp, *Coordination Chemistry Reviews*, 2017, DOI: 10.1016/j.ccr.2017.05.007.
- (50) S Warmlander, A Tiiman, A. A, L. J, J. J, S. Kl, D. J and G. A, *Chembiochem : a European journal of chemical biology*, 2013, **14**, 1692–1704.
- (51) S. Vivekanandan, J. R. Brender, S. Y. Lee and A. Ramamoorthy, *Biochemical and Biophysical Research Communications*, 2011, **411**, 312–316.
- (52) H Sticht, P Bayer, D Willbold, S Dames, C Hilbich, K Beyreuther, R. W. Frank and P Rösch, *European Journal Of Biochemistry / FEBS*, 1995, **233**, 293–298.
- (53) A. K. Somavarapu and K. P. Kepp, *Chemphyschem: A European Journal of Chemical Physics and Physical Chemistry*, 2015, **16**, 3278–3289.
- (54) M. Coles, W. Bicknell, A. A. Watson, D. P. Fairlie and D. J. Craik, *Biochemistry*, 1998, **37**, 11064–11077.
- (55) R. Roychaudhuri, M. Yang, M. M. Hoshi and D. B. Teplow, *The Journal of Biological Chemistry*, 2009, **284**, 4749–4753.

- 
- (56) M. Townsend, G. M. Shankar, T. Mehta, D. M. Walsh and D. J. Selkoe, *The Journal of Physiology*, 2006, **572**, 477–492.
- (57) G. M. Shankar, S. Li, T. H. Mehta, A. Garcia-Munoz, N. E. Shepardson, I. Smith, F. M. Brett, M. A. Farrell, M. J. Rowan, C. A. Lemere, C. M. Regan, D. M. Walsh, B. L. Sabatini and D. J. Selkoe, *Nature Medicine*, 2008, **14**, 837.
- (58) A. T. Petkova, W.-M. Yau and R. Tycko, *Biochemistry*, 2006, **45**, 498–512.
- (59) T. Sato, P. Kienlen-Campard, M. Ahmed, W. Liu, H. Li, J. I. Elliott, S. Aimoto, S. N. Constantinescu, J.-N. Octave and S. O. Smith, *Biochemistry*, 2006, **45**, 5503–5516.
- (60) T. Lührs, C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Döbeli, D. Schubert and R. Riek, *Proceedings of the National Academy of Sciences of the United States of America*, 2005, **102**, 17342–17347.
- (61) A. Abelein, B. Bolognesi, C. M. Dobson, A. Gräslund and C. Lendel, *Biochemistry*, 2012, **51**, 126–137.
- (62) K. Ono, M. M. Condron and D. B. Teplow, *Proceedings of the National Academy of Sciences*, 2009, **106**, 14745–14750.
- (63) S. A. Kotler, J. R. Brender, S. Vivekanandan, Y. Suzuki, K. Yamamoto, M. Monette, J. Krishnamoorthy, P. Walsh, M. Cauble, M. M. B. Holl, E. N. G. Marsh and A. Ramamoorthy, *Scientific Reports*, 2015, **5**, 11811.
- (64) J. D. Harper, S. S. Wong, C. M. Lieber and P. T. Lansbury, *Biochemistry*, 1999, **38**, 8972–8980.
- (65) J. J. Balbach, A. T. Petkova, N. A. Oyler, O. N. Antzutkin, D. J. Gordon, S. C. Meredith and R. Tycko, *Biophysical Journal*, 2002, **83**, 1205–1216.
- (66) D. Eisenberg and M. Jucker, *Cell*, 2012, **148**, 1188–1203.
- (67) T. L. Benzinger, D. M. Gregory, T. S. Burkoth, H. Miller-Auer, D. G. Lynn, R. E. Botto and S. C. Meredith, *Proceedings of the National Academy of Sciences of the United States of America*, 1998, **95**, 13407–13412.
- (68) I. Bertini, L. Gonnelli, C. Luchinat, J. Mao and A. Nesi, *Journal of the American Chemical Society*, 2011, **133**, 16013–16022.
- (69) D. A. Kirschner, H. Inouye, L. K. Duffy, A. Sinclair, M. Lind and D. J. Selkoe, *Proceedings of the National Academy of Sciences*, Oct. 1, 1987, **84**, 6953–6957.
- (70) H. Inouye, P. E. Fraser and D. A. Kirschner, *Biophysical Journal*, Feb. 1993, **64**, 502–519.
- (71) M. Sunde, L. C. Serpell, M. Bartlam, P. E. Fraser, M. B. Pepys and C. C. F. Blake, *Journal of Molecular Biology*, Oct. 31, 1997, **273**, 729–739.



- (72) M. Lovell, *Journal of Alzheimer's disease : JAD*, 2009, **16**, 471–483.
- (73) I. Shcherbatykh and D. O. Carpenter, *Journal of Alzheimer's disease: JAD*, 2007, **11**, 191–205.
- (74) C. J. Maynard, A. I. Bush, C. L. Masters, R. Cappai and Q. X. Li, *International Journal of Experimental Pathology*, 2005, **86**, 147–159.
- (75) S. C. Drew and K. J. Barnham, *Accounts of Chemical Research*, 2011, **44**, 1146–1155.
- (76) S. Zirah, S. A. Kozin, A. K. Mazur, A. Blond, M. Cheminant, I. Segalas-Milazzo, P. Debey and S. Rebuffat, *Journal of Biological Chemistry*, 2006, **281**, 2151–2161.
- (77) E. Gaggelli, A. Janicka-Klos, E. Jankowska, H. Kozlowski, C. Migliorini, E. Molteni, D. Valensin, G. Valensin and E. Wieczerszak, *The Journal of Physical Chemistry. B*, 2008, **112**, 100–109.
- (78) F. Bousejra-ElGarah, C. Bijani, Y. Coppel, P. Faller and C. Hureau, *Inorganic Chemistry*, 2011, **50**, 9024–9030.
- (79) V. Tōugu, A. Karafin, K. Zovo, R. S. Chung, C. Howells, A. K. West and P. Palumaa, *Journal of Neurochemistry*, 2009, **110**, 1784–1795.
- (80) J. H. Viles, *Coordination Chemistry Reviews*, 2012, **256**, 2271–2284.
- (81) C. J. Sarell, S. R. Wilkinson and J. H. Viles, *Journal of Biological Chemistry*, 2010, **285**, 41533–41540.
- (82) Y. Yoshiike, K. Tanemura, O. Murayama, T. Akagi, M. Murayama, S. Sato, X. Sun, N. Tanaka and A. Takashima, *Journal of Biological Chemistry*, 2001, **276**, 32293–32299.
- (83) A. Spinello, R. Bonsignore, G. Barone, B. K. Keppler and A. Terenzi, *Current Pharmaceutical Design*, 2016, **22**, 3996–4010.
- (84) K. J. Barnham, V. B. Kenche, G. D. Ciccotosto, D. P. Smith, D. J. Tew, X. Liu, K. Perez, G. A. Cranston, T. J. Johanssen, I. Volitakis, A. I. Bush, C. L. Masters, A. R. White, J. P. Smith, R. A. Cherny and R. Cappai, *Proceedings of the National Academy of Sciences of the United States of America*, 2008, **105**, 6813–6818.
- (85) C. Opazo, X. Huang, R. A. Cherny, R. D. Moir, A. E. Roher, A. R. White, R. Cappai, C. L. Masters, R. E. Tanzi, N. C. Inestrosa and A. I. Bush, *Journal of Biological Chemistry*, 2002, **277**, 40302–40308.
- (86) S. Varadarajan, J. Kanski, M. Aksenova, C. Lauderback and D. A. Butterfield, *Journal of the American Chemical Society*, 2001, **123**, 5625–5631.
- (87) K. J. Barnham, G. D. Ciccotosto, A. K. Tickler, F. E. Ali, D. G. Smith, N. A. Williamson, Y.-H. Lam, D. Carrington, D. Tew, G. Kocak, I. Volitakis,

- F. Separovic, C. J. Barrow, J. D. Wade, C. L. Masters, R. A. Cherny, C. C. Curtain, A. I. Bush and R. Cappai, *Journal of Biological Chemistry*, 2003, **278**, 42959–42965.
- (88) S. G. Yao, R. A. Cherny, A. I. Bush, C. L. Masters and K. J. Barnham, *Journal of Peptide Science*, 2004, **10**, 210–217.
- (89) V. A. Streltsov, V. C. Epa, S. A. James, Q. I. Churches, J. M. Caine, V. B. Kenche and K. J. Barnham, *Chemical Communications*, 2013, **49**, 11364–11366.
- (90) G. Ma, F. Huang, X. Pu, L. Jia, T. Jiang, L. Li and Y. Liu, *Chemistry-a European Journal*, 2011, **17**, 11657–11666.
- (91) G. Ma, E. Wang, H. Wei, K. Wei, P. Zhu and Y. Liu, *Metallomics*, 2013, **5**, 879.
- (92) K. J. Barnham, V. B. Kenche, L. W. Hung, K. Perez, I. Volitakes, G. Ciccosto, J. Kwok, N. Critch, N. Sherratt, M. Cortes, V. Lal, C. L. Masters, K. Murakami, R. Cappai and P. A. Adlard, *Angewandte Chemie-International Edition*, 2013, **52**, 3374–3378.
- (93) T. K. Govindan, *Proceedings of Indian Academy of Sciences, Section A*, 1956, **44**, 123.
- (94) A. Saleem, R. J. Harte, J. C. Matthews, S. Osman, F. Brady, S. K. Luthra, G. D. Brown, N. Bleehen, T. Connors, T. Jones, P. M. Price and E. O. Aboagye, *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, 2001, **19**, 1421–1429.
- (95) G. S. Yellol, J. G. Yellol, V. B. Kenche, X. M. Liu, K. J. Barnham, A. Donaire, C. Janiak and J. Ruiz, *Inorganic Chemistry*, 2015, **54**, 470–475.
- (96) I. Sasaki, C. Bijani, S. Ladeira, V. Bourdon, P. Faller and C. Hureau, *Dalton Transactions*, 2012, **41**, 6404–6407.
- (97) F. Collin, I. Sasaki, H. Eury, P. Faller and C. Hureau, *Chemical Communications*, 2013, **49**, 2130–2132.
- (98) V. Rangachari, A. Kumar, L. Moody, J. F. Olaivar, N. A. Lewis, R. L. Khade, A. A. Holder and Y. Zhang, *Acs Chemical Neuroscience*, 2010, **1**, 691–701.
- (99) X. Wang, X. Wang, C. Zhang, Y. Jiao and Z. Guo, *Chemical Science*, 2012, **3**, 1304.
- (100) D. Valensin, P. Anzini, E. Gaggelli, N. Gaggelli, G. Tamasi, R. Cini, C. Gabbiani, E. Michelucci, L. Messori, H. Kozlowski and G. Valensin, *Inorganic Chemistry*, 2010, **49**, 4720–4722.

- 
- (101) L. Messori, M. Camarri, T. Ferraro, C. Gabbiani and D. Franceschini, *ACS Medicinal Chemistry Letters*, 2013, **4**, 329–332.
- (102) M. R. Jones, C. Mu, M. C. P. Wang, M. I. Webb, C. J. Walsby and T. Storr, *Metallomics*, 2014, **7**, 129–135.
- (103) M. C. Heffern, P. T. Velasco, L. M. Matosziuk, J. L. Coomes, C. Karras, M. A. Ratner, W. L. Klein, A. L. Eckermann and T. J. Meade, *Chembiochem: A European Journal of Chemical Biology*, 2014, **15**, 1584–1589.
- (104) M. Li, S. E. Howson, K. Dong, N. Gao, J. Ren, P. Scott and X. Qu, *Journal of the American Chemical Society*, 2014, **136**, 11655–11663.
- (105) S. E. Howson, A. Bolhuis, V. Brabec, G. J. Clarkson, J. Malina, A. Rodger and P. Scott, *Nature Chemistry*, 2012, **4**, 31–36.
- (106) B. Y.-W. Man, H.-M. Chan, C.-H. Leung, D. S.-H. Chan, L.-P. Bai, Z.-H. Jiang, H.-W. Li and D.-L. Ma, *Chemical Science*, 2011, **2**, 917–921.
- (107) L. He, X. Wang, D. Zhu, C. Zhao and W. Du, *Metallomics*, 2015, **7**, 1562–1572.
- (108) R. Vargas, J. Garza, B. P. Hay and D. A. Dixon, *The Journal of Physical Chemistry A*, 2002, **106**, 3213–3218.
- (109) A. Perczel, Farkas, I. Jákl, I. A. Topol and I. G. Csizmadia, *Journal of Computational Chemistry*, 2003, **24**, 1026–1042.
- (110) A. Abo-Riziq, L. Grace, B. Crews, M. P. Callahan, T. van Mourik and M. S. d. Vries, *The Journal of Physical Chemistry A*, 2011, **115**, 6077–6087.
- (111) H. Valdes, K. Pluhackova and P. Hobza, *Journal of Chemical Theory and Computation*, 2009, **5**, 2248–2256.
- (112) Y. Zhao and D. G. Truhlar, *Theoretical Chemistry Accounts*, 2008, **120**, 215–241.
- (113) A. D. Becke, *The Journal of Chemical Physics*, 1993, **98**, 5648–5652.
- (114) A. E. Shields and T. van Mourik, *The Journal of Physical Chemistry A*, 2007, **111**, 13272–13277.
- (115) L. F. Holroyd and T. van Mourik, *Chemical Physics Letters*, 2007, **442**, 42–46.
- (116) T. v. Mourik, *Journal of Chemical Theory and Computation*, 2008, **4**, 1610–1619.
- (117) F. Schubert, M. Rossi, C. Baldauf, K. Pagel, S. Warnke, G. v. Helden, F. Filsinger, P. Kupser, G. Meijer, M. Salwiczek, B. Koks, M. Scheffler and V. Blum, *Physical Chemistry Chemical Physics*, 2015, **17**, 7373–7385.
- (118) M. Rauschenberg, S. Bandaru, M. P. Waller and B. J. Ravoo, *Chemistry – A European Journal*, 2014, **20**, 2770–2782.

- (119) D. B. Ninković, D. P. Malenov, P. V. Petrović, E. N. Brothers, S. Niu, M. B. Hall, M. R. Belić and S. D. Zarić, *Chemistry - A European Journal*, 2017, **23**, 11046–11053.
- (120) K. J. Barnham, F. Haeffner, G. D. Ciccotosto, C. C. Curtain, D. Tew, C. Mavros, K. Beyreuther, D. Carrington, C. L. Masters, R. A. Cherny, R. Cappai and A. I. Bush, *The FASEB Journal*, 2004, **18**, 1427–1429.
- (121) J. J. P. Stewart, *Journal of Molecular Modeling*, 2008, **15**, 765–805.
- (122) F. Pichierri, *Bioorganic & Medicinal Chemistry Letters*, 2006, **16**, 587–591.
- (123) F. Pichierri, *Computational and Theoretical Chemistry*, 2011, **963**, 384–393.
- (124) B. Temelso, K. A. Alser, A. Gauthier, A. K. Palmer and G. C. Shields, *The Journal of Physical Chemistry B*, 2014, **118**, 4514–4526.
- (125) J. C. Faver, M. L. Benson, X. He, B. P. Roberts, B. Wang, M. S. Marshall, C. D. Sherrill and K. M. M. Jr, *PLOS ONE*, 2011, **6**, e18868.
- (126) P. Dobeš, J. Řezáč, J. Fanfrlík, M. Otyepka and P. Hobza, *The Journal of Physical Chemistry B*, 2011, **115**, 8581–8589.
- (127) J. J. P. Stewart, *Journal of Molecular Modeling*, 2013, **19**, 1–32.
- (128) D. Ivan, S. Funar-Timofei, M. Medeleanu, M. Mracec and M. Mracec, *Rev. Roum. Chim.*, 2013, **58**, 463–472.
- (129) B. P. Martin, C. J. Brandon, J. J. P. Stewart and S. B. Braun-Sand, *Proteins: Structure, Function, and Bioinformatics*, 2015, **83**, 1427–1435.
- (130) H. Ryan, M. Carter, P. Stenmark, J. J. P. Stewart and S. B. Braun-Sand, *Journal of Molecular Modeling*, 2016, **22**, 168.
- (131) J. Hostaš, J. Řezáč and P. Hobza, *Chemical Physics Letters*, 2013, **568-569**, 161–166.
- (132) W. Hua, L. Xu, Y. Luo and S. Li, *ChemPhysChem*, 2011, **12**, 1325–1333.
- (133) J. Fanfrlík, A. K. Bronowska, J. Rezac, O. Prenosil, J. Konvalinka and P. Hobza, *The Journal of Physical Chemistry B*, 2010, **114**, 12666–12678.
- (134) P. Dobes, J. Fanfrlík, J. Rezac, M. Otyepka and P. Hobza, *Journal of Computer Aided Molecular Design*, 2011, **25**, 223–235.
- (135) M. Karplus and J. A. McCammon, *Nature Structural Biology*, 2002, **9**, 646–652.
- (136) R. Brooks, *Current opinion in structural biology*, 1998, **8**, 222–226.
- (137) J. A. Lemkul and D. R. Bevan, *ACS Chemical Neuroscience*, 2012, **3**, 845–856.
- (138) P. P. Mager, *Medicinal Research Reviews*, 1998, **18**, 403–430.
- (139) P. Anand, F. S. Nandel and U. H. E. Hansmann, *The Journal of Chemical Physics*, 2008, **128**, 165102.

- 
- (140) M. Yang and D. B. Teplow, *Journal of Molecular Biology*, 2008, **384**, 450–464.
- (141) D. F. Raffa and A. Rauk, *The Journal of Physical Chemistry B*, 2007, **111**, 3789–3799.
- (142) D. M. Walsh, A. Lomakin, G. B. Benedek, M. M. Condrón and D. B. Teplow, *The Journal of Biological Chemistry*, 1997, **272**, 22364–22372.
- (143) G. Bitan, M. D. Kirkitadze, A. Lomakin, S. S. Vollers, G. B. Benedek and D. B. Teplow, *Proceedings of the National Academy of Sciences of the United States of America*, 2003, **100**, 330–335.
- (144) N. G. Sgourakis, Y. Yan, S. A. McCallum, C. Wang and A. E. García, *Journal of Molecular Biology*, 2007, **368**, 1448–1457.
- (145) N. G. Sgourakis, M. Merced-Serrano, C. Boutsidis, P. Drineas, Z. Du, C. Wang and A. E. García, *Journal of Molecular Biology*, 2011, **405**, 570–583.
- (146) E. Luttmann and G. Fels, *Chemical Physics*, 2006, **323**, 138–147.
- (147) R. Armen, D. O. V. Alonso and V. Daggett, *Protein Science: A Publication of the Protein Society*, 2003, **12**, 1145–1157.
- (148) Y. Xu, J. Shen, X. Luo, W. Zhu, K. Chen, J. Ma and H. Jiang, *Proceedings of the National Academy of Sciences of the United States of America*, 2005, **102**, 5403–5407.
- (149) P. D. Q. Huy, Q. V. Vuong, G. La Penna, P. Faller and M. S. Li, *ACS Chemical Neuroscience*, 2016, **7**, 1348–1363.
- (150) S. Nag, B. Sarkar, A. Bandyopadhyay, B. Sahoo, V. K. A. Sreenivasan, M. Kombrabail, C. Muralidharan and S. Maiti, *Journal of Biological Chemistry*, 2011, **286**, 13827–13833.
- (151) C. Velez-Vega and F. A. Escobedo, *The Journal of Physical Chemistry B*, 2011, **115**, 4900–4910.
- (152) S. Côté, P. Derreumaux and N. Mousseau, *Journal of Chemical Theory and Computation*, 2011, **7**, 2584–2592.
- (153) K. A. Ball, A. Phillips, D. Wemmer and T. Head-Gordon, *Biophysical Journal*, 2013, **104**, 2714–2724.
- (154) D. J. Rosenman, C. R. Connors, W. Chen, C. Wang and A. E. García, *Journal of Molecular Biology*, 2013, **425**, 3338–3359.
- (155) D. J. Rosenman, C. Wang and A. E. García, *The Journal of Physical Chemistry. B*, 2016, **120**, 259–277.
- (156) S. Mitternacht, I. Staneva, T. Härd and A. Irbäck, *Proteins: Structure, Function, and Bioinformatics*, 2010, **78**, 2600–2608.

- 
- (157) B. Urbanc, M. Betnel, L. Cruz, G. Bitan and D. B. Teplow, *Journal of the American Chemical Society*, 2010, **132**, 4266–4280.
- (158) B. Urbanc, L. Cruz, F. Ding, D. Sammond, S. Khare, S. V. Buldyrev, H. E. Stanley and N. V. Dokholyan, *Biophysical Journal*, 2004, **87**, 2310–2321.
- (159) Y. Chebaro, N. Mousseau and P. Derreumaux, *The Journal of Physical Chemistry B*, 2009, **113**, 7668–7675.
- (160) M. Carballo-Pacheco and B. Strodel, *Protein Science: A Publication of the Protein Society*, 2017, **26**, 174–185.
- (161) E. Haensele, N. Saleh, C. M. Read, L. Banting, D. C. Whitley and T. Clark, *Journal of Chemical Information and Modeling*, 2016, **56**, 1798–1807.
- (162) F. Neese, *JBIC Journal of Biological Inorganic Chemistry*, 2006, **11**, 702–711.
- (163) T. A. Rokob, M. Srnec and L. Rulíšek, *Dalton Transactions*, 2012, **41**, 5754–5768.
- (164) C. J. Cramer and D. G. Truhlar, *Physical Chemistry Chemical Physics*, 2009, **11**, 10757–10816.
- (165) W. T. G. Novato, P. H. F. Stroppa, A. D. Da Silva, N. P. Botezine, F. C. Machado, L. A. S. Costa and H. F. Dos Santos, *Chemical Physics Letters*, 2017, **667**, 4–8.
- (166) T. Zimmermann, M. Zeizinger and J. V. Burda, *Journal of Inorganic Biochemistry*, 2005, **99**, 2184–2196.
- (167) T. Zimmermann, Z. Chval and J. V. Burda, *Journal of Physical Chemistry B*, 2009, **113**, 3139–3150.
- (168) T. Zimmermann and J. V. Burda, *Interdisciplinary Sciences: Computational Life Sciences*, 2010, **2**, 98–114.
- (169) J. Alí-Torres, J.-D. Maréchal, L. Rodríguez-Santiago and M. Sodupe, *Journal of the American Chemical Society*, 2011, **133**, 15008–15014.
- (170) T. Marino, N. Russo, M. Toscano and M. Pavelka, *Interdisciplinary Sciences-Computational Life Sciences*, 2010, **2**, 57–69.
- (171) S. Azimi and A. Rauk, *International Journal of Alzheimer’s Disease*, 2011, **2011**, 1–15.
- (172) V. A. Streltsov, S. J. Titmuss, V. C. Epa, K. J. Barnham, C. L. Masters and J. N. Varghese, *Biophysical Journal*, 2008, **95**, 3447–3456.
- (173) T. Dudev and C. Lim, *Annual Review of Biophysics*, 2008, **37**, 97–116.
- (174) T. Dudev and C. Lim, *Journal of the American Chemical Society*, 2006, **128**, 10541–10548.

- (175) T. Dudev and C. Lim, *Journal of the American Chemical Society*, 2011, **133**, 9506–9515.
- (176) K. Gkionis and J. A. Platts, *JBIC Journal of Biological Inorganic Chemistry*, 2009, **14**, 1165–1174.
- (177) K. Gkionis and J. A. Platts, *Computational and Theoretical Chemistry*, 2012, **993**, 60–65.
- (178) K. Gkionis, S. T. Mutter and J. A. Platts, *Rsc Advances*, 2013, **3**, 4066–4073.
- (179) Y. Xia, X. Wang, Y. Zhang and B. Luo, *Computational and Theoretical Chemistry*, 2011, **967**, 213–218.
- (180) A. C. Bertoli, R. Carvalho, M. P. Freitas, T. C. Ramalho, D. T. Mancini, M. C. Oliveira, A. de Varennes and A. Dias, *Inorganica Chimica Acta*, 2015, **425**, 164–168.
- (181) A. C. Bertoli, R. Carvalho, M. P. Freitas, T. C. Ramalho, D. T. Mancini, M. C. Oliveira, A. de Varennes and A. Dias, *Journal of Inorganic Biochemistry*, 2015, **144**, 31–37.
- (182) A. C. Bertoli, R. Carvalho, M. P. Freitas, T. C. Ramalho, D. T. Mancini, M. C. Oliveira, A. de Varennes and A. Dias, *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2015, **137**, 271–280.
- (183) K. M. Pietrusiewicz, A. Flis, V. Ujj, T. Körtvélyesi, L. Drahos, P. Pongrácz, L. Kollár and G. Keglevich, *Heteroatom Chemistry*, 2011, **22**, 730–736.
- (184) G. Keglevich, A. Kerenyi, T. Kortvelyesi, K. Ludanyi and L. Drahos, *Letters in Organic Chemistry*, 2010, **7**, 235–239.
- (185) N. Shahabadi and L. Heidari, *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2014, **128**, 377–385.
- (186) D. Suárez, N. Díaz and R. López, *Journal of Computational Chemistry*, 2014, **35**, 324–334.
- (187) D. Suárez and N. Díaz, *ChemPhysChem*, 2015, **16**, 1646–1656.
- (188) C. L. Fernandes, G. B. Escouto and H. Verli, *Journal of Biomolecular Structure and Dynamics*, 2014, **32**, 1092–1102.
- (189) N. Shahabadi, M. M. Khodaei, S. Kashanian, F. Kheiridoosh and S. M. Filli, *Molecular Biology Reports*, 2014, **41**, 3271–3278.
- (190) E. D. Santis, A. A. Edwards, B. D. Alexander, S. J. Holder, A.-S. Biesse-Martin, B. V. Nielsen, D. Mistry, L. Waters, G. Siligardi, R. Hussain, S. Faure and C. Taillefumier, *Organic & Biomolecular Chemistry*, 2016, **14**, 11371–11380.
- (191) N. A. M. S. Caturello, Z. Csók, G. Fernández and R. Q. Albuquerque, *Chemistry – A European Journal*, 2016, **22**, 17681–17689.

- 
- (192) K. Karami, M. Rafiee, Z. M. Lighvan, M. Zakariazadeh, A. Y. Faal, S.-A. Esmaeili and A. A. Momtazi-Borojeni, *Journal of Molecular Structure*, 2018, **1154**, 480–495.
- (193) A. Rappe, C. Casewit, K. Colwell, W. Goddard and W. Skiff, *Journal of the American Chemical Society*, 1992, **114**, 10024–10035.
- (194) T. W. Hambley and A. R. Jones, *Coordination Chemistry Reviews*, 2001, **212**, 35–59.
- (195) E. D. Scheeff, J. M. Briggs and S. B. Howell, *Molecular Pharmacology*, 1999, **56**, 633–643.
- (196) S. Yao, J. P. Plastaras and L. G. Marzilli, *Inorganic Chemistry*, 1994, **33**, 6061–6077.
- (197) J. Chan, Z. Huang, M. E. Merrifield, M. T. Salgado and M. J. Stillman, *Coordination Chemistry Reviews*, 2002, **233–234**, 319–339.
- (198) J. Y. Xiang and J. W. Ponder, *Journal of Chemical Theory and Computation*, 2014, **10**, 298–311.
- (199) F. Duarte, P. Bauer, A. Barrozo, B. A. Amrein, M. Purg, J. Åqvist and S. C. L. Kamerlin, *The Journal of Physical Chemistry B*, 2014, **118**, 4351–4362.
- (200) Q. Liao, S. C. L. Kamerlin and B. Strodel, *The Journal of Physical Chemistry Letters*, 2015, **6**, 2657–2662.
- (201) V. Burton, R. Deeth, C. Kemp and P. Gilbert, *Journal of the American Chemical Society*, 1995, **117**, 8407–8415.
- (202) R. J. Deeth, *Chemical Communications*, 2006, **0**, 2551–2553.
- (203) R. J. Deeth, *Inorganic Chemistry*, 2007, **46**, 4492–4503.
- (204) R. J. Deeth and C. Diedrich, *JBIC Journal of Biological Inorganic Chemistry*, 2010, **15**, 117–129.
- (205) S. T. Mutter, R. J. Deeth, M. Turner and J. A. Platts, *Journal of Biomolecular Structure and Dynamics*, 2017, **36**, 1–9.
- (206) A. E. Anastasi and R. J. Deeth, *Journal of Chemical Theory and Computation*, 2009, **5**, 2339–2352.
- (207) H.-C. Tai, R. Brodbeck, J. Kasparkova, N. J. Farrer, V. Brabec, P. J. Sadler and R. J. Deeth, *Inorganic Chemistry*, 2012, **51**, 6830–6841.



## 2 Theory

### 2.1 Introduction

In this section, the theoretical concepts behind the computational methods used in this report will be explained. The basis of this chapter comes from a selection of the standard computational chemistry textbooks.[1–3]

### 2.2 Hartree-Fock Theory

#### 2.2.1 The Schrodinger Equation

The Schrödinger equation (Equation 2.1) forms the basis of modern theoretical chemistry. Expressed here in the time-independent form, the Schrödinger equation uses the wavefunction of a given particle to predict the steady-state properties of a system.

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

Where  $\hat{H}$  is the Hamiltonian operator,  $\Psi$  is the wavefunction and  $E$  is the energy eigenvalue of the system. Expansion of this equation yields a second order partial differential equation. The Hamiltonian may be expressed in terms of the nuclei and electron positions ( $R$  and  $r$ , respectively), as illustrated in Equation 2.2:

$$\hat{H} = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 - \sum_{k=1}^M \frac{1}{2M_k} \nabla_k^2 - \sum_{i=1}^N \sum_{k=1}^M \frac{Z_k}{r_{ik}} + \sum_{i=1}^N \sum_{j>1}^N \frac{1}{r_{ij}} + \sum_{k=1}^M \sum_{l>k}^M \frac{Z_k Z_l}{R_{kl}} \quad (2.2)$$

Where  $N$  is the number of electrons,  $M$  is the number of nuclei,  $i$  and  $j$  are electrons,  $k$  and  $l$  are nuclei,  $Z$  is the nuclear atomic number,  $M_k$  is the ratio of the

mass of nucleus  $k$  to an electron and  $\nabla^2$  is the Laplacian operator. In this form, the first and second terms are equal to the kinetic energy operators for the electrons and nuclei respectively, while the remaining terms represent the Coulombic interactions of electron-nucleus attraction, electron-electron repulsion and nucleus-nucleus repulsion, respectively.

### 2.2.2 Born-Oppenheimer Approximation

Exact solutions to the Schrödinger equation allow for the calculation of electron distribution within a system. Unfortunately, it is impossible to solve the Schrödinger equation exactly for any system containing three or more particles. This issue arises from the correlated motion of particles and is known as the three-body problem. Therefore, in order to carry out calculations using systems with more than two particles, approximations must be made.

The first of these is the Born-Oppenheimer approximation, which decouples the motion of particles by ignoring nuclear kinetic energy and assuming that the inter-nuclei repulsion remains constant. This approximation can be made as the nuclei are many times more massive than the electrons and therefore any change in the position of the electrons may be considered instantaneous in relation to the nuclei. More simply, we consider the nuclei to be fixed with respect to electronic motion. As a result, the full Hamiltonian above may be split into nuclear and electronic components, as shown in Equation 2.3.

$$\hat{H}_{tot} = \hat{H}_{elec} + \hat{H}_{nuc} \quad (2.3)$$

$$\hat{H}_{elec} = -\frac{1}{2} \sum_{i=1}^N \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_k}{r_{ik}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} \quad (2.4)$$

Where the symbols retain their meaning from Equation 2.2 above. The electronic Hamiltonian therefore contains the first, third and fourth terms from the full Hamiltonian shown previously. The nuclear-nuclear potential energy term is retained, but in the Born-Oppenheimer approximation, this is constant for any nuclear arrangement. The Schrodinger equation is then solved where electrons move in the electrostatic field of the nuclei.

The wavefunctions expressed in the above equations have no clear physical meaning on their own. However, they may be interpreted as proposed by Born, which considers the wavefunction as a probability density (see Equations 2.5-2.6).

$$|\Psi|^2 = \Psi^* \Psi \quad (2.5)$$

$$\int \Psi^* \Psi \delta\tau = N \quad (2.6)$$

The calculated value  $|\Psi|^2$  multiplied by some volume,  $\delta\tau$ , gives the probability of finding an electron in that volume of space. Furthermore, integration of this probability density over possible space will equal the total number of electrons,  $N$ , present in the system.

### 2.2.3 Molecular Orbital Approximation

It is not possible to find exact solutions to the Schrodinger equation for many-electron systems, so additional approximations are required. The molecular orbital approximation rewrites the total wavefunction as the product of one electron wavefunctions, or molecular orbitals. As with the Born-Oppenheimer approximation, this approximation is virtually ubiquitous in modern theoretical chemistry. Here, we consider the electron spin orientation in addition to its spatial coordinates, where the spin ('up' or 'down') is described by functions  $\alpha$  or  $\beta$ . Each electron is therefore described by four coordinates – three spatial and one spin. The product of the spatial and spin functions for a given electron yields spin orbitals, denoted  $\chi$ . Many-electron systems are then represented as a product of these spin orbitals, displayed in Equation 2.7.

$$\Psi^{HP}(1, 2 \dots N) = \chi_1(1)\chi_2(2) \dots \chi_N(N) \quad (2.7)$$

Exchange of any two electrons must also give a change in the sign of the wavefunction. The above Hartree product approach does not satisfy this anti-symmetry principle. However, a linear combination of these products does satisfy this principle; the resulting wavefunctions may be expressed as Slater determinants, as shown in Equation 2.8.

$$\Psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(1) & \chi_2(1) & \cdots & \chi_N(1) \\ \chi_1(2) & \chi_2(2) & \cdots & \chi_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(N) & \chi_2(N) & \cdots & \chi_N(N) \end{vmatrix} \quad (2.8)$$

### 2.2.4 Hartree-Fock Approximation

Since the three-body problem cannot be solved exactly, approximate solutions to the Schrodinger equation are needed. The Hartree-Fock (HF) method utilises the one-electron Fock operator,  $\hat{f}$ , to simplify the many-electron problem to a one-electron problem. As Equation 2.9 shows, this operator enables the electron-electron repulsion to be expressed as the average potential,  $V^{HF}$ , that any given electron experiences from all other electrons in the system.

$$\hat{f} = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^M \frac{Z_A}{r_{iA}} + V^{HF} \quad (2.9)$$

$$V^{HF} = \sum_j J_j - K_j \quad (2.10)$$

The first two terms of the Fock operator represent the one-electron Hamiltonian of an electron moving in space in a constant nuclear field, as described previously. The average electron-electron repulsion (Equation 2.10) potential contains both Coulombic terms,  $J$ , for the electrostatic repulsion between two negatively charged electrons as well as exchange interactions,  $K$ , representing the spin correlation. The Hartree-Fock equation, where the Fock operator included is found in Equation 2.11.

$$\hat{f}\chi_i = \varepsilon\chi_i \quad (2.11)$$

Here, the Fock operator acts on a molecular orbital,  $\chi_i$ , giving the molecular orbital multiplied by an eigenvalue,  $\varepsilon$ . A set of orbitals that are a solution to this equation are known as self-consistent orbitals. The Fock operator depends on the coordinates of every electron present; a solution for one electron therefore affects the other electrons, due to the presence of the average potential term  $V^{HF}$ . Thus, the Hartree-Fock method must be solved iteratively, using the self-consistent field approach (see Section 2.2.6).

### 2.2.5 Linear Combination of Atomic Orbitals and the Roothaan-Hall Equations

Directly solving the above Hartree-Fock equations for molecular systems is impractical; an alternative is to write each spin orbital,  $\chi$ , as a linear combination of atomic orbitals (LCAO), as seen in Equation 2.12

$$\chi_i = \sum_{\nu}^K C_{\nu i} \phi_{\nu} \quad (2.12)$$

Where  $K$  is the number of atomic orbitals,  $C_{\nu i}$  are the coefficients and  $\phi_{\nu}$  are atomic orbitals. The Hartree-Fock equations therefore take the form given in Equation 2.13. Using a specific set of these atomic orbitals, termed a basis, produces the Roothaan-Hall equations, 2.14.

$$f \sum_{\nu}^K C_{\nu i} \phi_{\nu} = \varepsilon_i \sum_{\nu}^K C_{\nu i} \phi_{\nu} \quad (2.13)$$

$$FC = SC\varepsilon \quad (2.14)$$

Where  $F$  is the Fock matrix,  $C$  is the matrix of coefficients,  $S$  is the overlap matrix and  $\varepsilon$  is a diagonal matrix of orbital energies, representing the energy of an electron in a spin orbital.

### 2.2.6 Self-Consistent Field Theory

One of the biggest challenges in computational chemistry is the accurate description of indistinguishable interacting particles, such as electrons. This issue is most commonly encountered when attempting to find solutions to the Schrodinger equation, as we have already seen. The interaction of these particles must be contained within the Hamiltonian, so the position of these particles must be known in order to find the Hamiltonian. However, the positions of these particles can only be determined by finding the Hamiltonian.

To solve this problem, an iterative, self-consistent method is applied. To begin, approximate wavefunctions,  $\Psi$ , are chosen for all occupied molecular orbitals and the one-electron operators,  $\hat{h}$ , are constructed. The Schrödinger equation is then solved

using these Hamiltonians to afford a new set of wavefunctions, presumably more accurate than those before. From these new  $\Psi$ , a new set of one-electron Hamiltonians are formed and the Schrödinger equation is then solved continuously until the resulting values converge.

This must be an iterative process, since the effective potential of each electron depends on all other electrons in the system; reliable results will only be obtained when the values are self-consistent. All values contained within this report were calculated according to this method, summed up below:

- (i) Guess approximate starting orbitals
- (ii) Calculate effective potential ( $V_{HF}$ ) for these orbitals
- (iii) Adjust orbitals to respond to new  $V_{HF}$
- (iv) Compare original and new orbitals. If they are the same, proceed to Step (v); if they are different, return to Step (ii)
- (v) Self Consistent Field achieved

## 2.3 Density Functional Theory

One of the most significant flaws in Hartree-Fock theory is that it averages inter-electronic repulsion, as we have seen. This fundamental assumption of Hartree-Fock theory allows for calculation of the molecular Hamiltonian, but can cause significant errors in calculating accurate wavefunctions. The energy error associated with the Hartree-Fock approximation is known as electron correlation,  $E_{corr}$ , which can be expressed as shown in Equation 2.15.

$$E_{corr} = E_{exact} - E_{HF} \quad (2.15)$$

The electron correlation energy is important for chemical systems containing many electrons as well as those containing dispersion interactions. This means that Hartree-Fock methods would be inappropriate for modelling the systems reported here.

Density functional theory (DFT), introduced by Hohenberg and Kohn in 1964 [4], is one of the most successful methods for the investigation of ground state electronic structure in multi-electron systems. This is as a result of the accuracy of DFT, due to accounting for the above electron correlation effects, and its low computational expense when compared to more complex *ab initio* methods. DFT uses functionals in order to calculate structural and electronic properties of a system, where functionals are defined as the function of a function; here the functional is electron density ( $\rho$ ), which in turn is a function of position in 3D space. In DFT, the ground state electronic energy is determined completely by the electron density.

As already mentioned, one major advantage of DFT over wavefunction-based methods may be illustrated by considering the computational resources required:

The wavefunction for a system containing  $N$  electrons has  $4N$  variables – three spatial coordinates ( $x, y, z$ ) and one spin coordinate for each electron. The electron density can be considered as the square of the wavefunction, integrated over  $N-1$  electron coordinates. When using orbital-free DFT, each spin density only depends on three spatial coordinates, regardless of the number of electrons. This means that while the complexity of a wavefunction increases drastically as the number of electrons in the system increases - as does the computational requirement to complete

the associated calculations - the electron density has the same number of variables, independent of the size of the system.

Hohenberg and Kohn [4] are responsible for two theorems that highlight the importance of electron density to electronic structure calculations. The existence theorem states that the ground state energy and properties of a given system are governed purely by the electron density. Thus, the total ground state energy is a unique functional of the electron density, as seen in Equation 2.16.

$$E[\rho(r)] = \int V_{ext}(r)\rho(r)dr + F[\rho(r)] \quad (2.16)$$

Here, the first term represents the interaction of the electrons with an external potential,  $V_{ext}(r)$ , exerted by the nuclei. The second term,  $F[\rho(r)]$ , is the sum of electronic kinetic energy and interelectronic interactions.

The variational theorem states the ground state energy may be obtained variationally; a given trial electron density will produce an energy that is greater than or equal to the true energy. This means that the electron density that minimises the absolute energy of the system must equal the ground state electron density. However, this orbital-free density functional theory is flawed, as it does not predict the binding of molecules from atoms. In addition, the exact form of the second term,  $F[\rho(r)]$ , is not known; in turn, this means that the dependence of the total energy on  $\rho(r)$  is not known either.

Subsequently, this method was adapted by Kohn and Sham [5] and has become the most common DFT formalism used in computational chemistry. In Kohn-Sham DFT, a set of non-interacting orbitals described by single determinant wavefunctions are used to reproduce the electron density, shown in Equation 2.17.

$$\rho(r) = \sum_{i=1}^N |\Psi_i(r)|^2 \quad (2.17)$$

Since Kohn-Sham [5] DFT re-introduces orbitals, the computational complexity increases from 3 to 3N variables. The Kohn-Sham approach aims to correct the poor description of the interelectronic interactions and kinetic energy in the  $F[\rho(r)]$  term from Hohenberg-Kohn DFT by approximating the term as a sum of different contributions to the functional - see Equation 2.18.



$$F[\rho(r)] = E_{KE}[\rho(r)] + E_H[\rho(r)] + E_{XC}[\rho(r)] \quad (2.18)$$

Where  $E_{KE}[\rho(r)]$  is the kinetic energy,  $E_H[\rho(r)]$  is the electron-electron Coulombic repulsion and  $E_{XC}[\rho(r)]$  is the exchange-correlation function, which typically contributes approximately 1% to the total energy of the system. The exchange-correlation energy is the only unknown functional in the above expression. Using these terms, the full expression for the energy in Kohn-Sham DFT may be written as shown in Equation 2.19.

$$\begin{aligned} [E\rho(r)] = & \sum_{i=0}^N \int \psi_i(r) \left(-\frac{\nabla^2}{2}\right) \psi_i(r) dr + \frac{1}{2} \int \int \frac{\rho(r_1)\rho(r_2)}{|r_1 - r_2|} dr_1 dr_2 \\ & + E_{XC}[\rho(r)] - \sum_{k=1}^M \int \frac{Z_k}{|r - R_k|} \rho(r) dr \end{aligned} \quad (2.19)$$

### 2.3.1 Local Density Approximation

The key to accurate DFT calculations is the accurate approximation of the exchange-correlation functional, yet even simple approximations provide reasonably accurate models for computational calculations. The distinction between different DFT methods is the description of the unknown exchange-correlation term,  $E_{XC}[\rho(r)]$ .

The simplest approximation of the exchange-correlation term is local density approximation, LDA, where the electron density is assumed to be slowly varying such that the exchange-correlation energy is identical to a uniform electron gas of equal density. Integration over all space gives the total exchange-correlation energy, displayed in Equation 2.20.

$$E_{XC}[\rho(r)] = \int \rho(r) \varepsilon_{XC}[\rho(r)] dr \quad (2.20)$$

Where  $\varepsilon_{XC}[\rho(r)]$  is the exchange-correlation energy of a uniform electron gas. This LDA approach assumes that electron spin is equal to zero at all points; in instances where  $\alpha$  and  $\beta$  spins are not equal, LDA produces significant errors. One alternative method for these cases is the more general local spin density approximation

(LSDA) approach. These LDA methods fail where the gradient of the electron density changes rapidly, such as in molecules, but still find use in the study of metallic systems.

### 2.3.2 Generalised Gradient Approximation

The ideal model of a uniform electron gas introduces errors to the DFT approach. One solution to this issue was the introduction of the Generalised Gradient Approximation (GGA) method, where the exchange-correlation energy becomes dependent on the gradient of the electron density *i.e.* the derivatives of  $\rho(r)$ .

Popular examples of GGA functionals are Becke's B88[6] and the functional of Lee, Yang and Parr.[7] These GGA methods may be expanded to include higher-order derivatives of the electron density, leading to meta-GGA methods. While these GGAs do not yield a significant improvement over LDA methods for all systems, they are known to be more accurate when calculating molecular structures and modelling inter-molecular bonds.[8]

### 2.3.3 Hybrids

One development in DFT methods was the introduction of hybrid functionals. These hybrid methods involve the combination of various correlation and exchange functionals from standard DFT models. Typically, hybrid functionals contain some contribution from both the exact exchange energy from Hartree-Fock theory and the exchange and correlation energies from LSDA or gradient-corrected methods.

These hybrid methods provide a simple way of accurately calculating a broad variety of molecular properties and therefore are very popular in quantum chemistry. Two popular hybrid functionals are discussed below.

#### 2.3.3.1 B3LYP

The most widely used hybrid functional is B3LYP, introduced following work by Becke [9] and Lee, Yang and Parr [7]. This method uses a combination of exact, LSDA and Becke's B88 exchange energy alongside LSDA and LYP correlation en-

ergy. These terms are weighted by three parameters, as fitted to experimental data. The exchange correlation functional may be written as Equation 2.21.

$$E_{XC} = E_{XC}LDA + \alpha_0(E_XHF - E_XLDA) + \alpha_X(E_XGGA - E_XLDA) + \alpha_C(E_CGGA - E_CLDA) \quad (2.21)$$

Where the parameters  $\alpha_0 = 0.20$ ,  $\alpha_X = 0.72$  and  $\alpha_C = 0.81$ . The generalised gradient approximations,  $E_XGGA$  and  $E_CGGA$  correspond to the Becke 88 exchange functional and the LYP correlation functional respectively. The local density approximation to the correlation functional is written  $E_CLDA$ .

### 2.3.3.2 BHandH

An alternate method is the BHandH functional, also referred to as Becke's 'half and half' functional [9]. This hybrid functional uses a 1:1 mixture of exact and LDA exchange energies *i.e.* half of the Hartree-Fock exchange energy and half of the local density approximation exchange energy, allied to the LYP correlation energy [9], shown in Equation 2.22

$$E_{XC} = 0.5E_XHF + 0.5E_XLDA + E_CLYP \quad (2.22)$$

### 2.3.4 Remarks

Despite its success, density functional theory is not without limitations. Firstly, it is important to note that unlike Hartree-Fock, DFT is not a variational method *i.e.* a lower energy value does not necessarily correspond to a better, more accurate description of a given system. Secondly, there is no way to know whether a given method will be suitable for a particular system without performing calculations. To compound this issue, there is no systematic way to improve the functional if it is inadequate for the system in question.

In addition, dispersion is not included in the construction of traditional Kohn-Sham DFT, leading to inaccurate modelling of dispersion forces, such as van der Waals interactions. However, modern DFT methods provide several possible solutions to this problem; one such example is modifying the exchange-correlation functional.

The modified functionals, such as the Minnesota functionals of Truhlar, [10] may be parameterised to account for dispersion interactions. An alternative solution is the inclusion of an explicit dispersion correction, known as DFT-D, proposed by Grimme [11], which takes the form of an atom pairwise sum of  $r^{-6}$  potentials. The most common form of DFT-D is known as a “D2-type” correction, displayed in Equations 2.23-2.25.

$$E_{DFT-D} = E_{KS-DFT} + E_{DISP} \quad (2.23)$$

$$E_{DISP} = -S_6 \sum_{i=1}^{N_{at}-1} \sum_{j=i+1}^{N_{at}} \frac{C_6^{ij}}{R_{ij}^6} f_{dmp}(R_{ij}) \quad (2.24)$$

$$f_{dmp}(R_{ij}) = \frac{1}{1 + e^{-d(R_{ij}/R_r-1)}} \quad (2.25)$$

Here,  $s$  is a global scaling factor,  $N_{at}$  is the number of atoms present,  $C_6^{ij}$  is the dispersion coefficient for atoms  $i$  and  $j$ ,  $R_{ij}^6$  is the distance between atoms  $i$  and  $j$  to the power six and  $f_{dmp}$  is a damping function, shown in Equation 2.25.

More recently, the “D3-type” dispersion correction (Equation 2.26) has been introduced[12], which has both greater flexibility and increased accuracy over its predecessor, due to its inclusion of  $r^{-8}$  terms in addition to the existing  $r^{-6}$  terms.

$$E_{DISP}^{D3} = -\frac{1}{2} \sum_{j \neq i} (S_6 \frac{C_6^{ij}}{R_{ij}^6 + [f_{dmp}(R_{ij})]^6} + S_8 \frac{C_8^{ij}}{R_{ij}^8 + [f_{dmp}(R_{ij})]^8}) \quad (2.26)$$

## 2.4 Basis Sets

The linear combination of atomic orbitals (LCAO) approach, illustrated in Equation 2.12, Section 2.2.5, is commonly used for calculating molecular orbital energies and coefficients. As alluded to earlier, the molecular orbitals are expressed as one-electron functions, centred on atomic nuclei. These are known as basis functions; a collection of these functions used to describe a chemical system is known as a basis set. When performing calculations, all quantum mechanical methods used required a basis set to be specified. The individual basis functions are constructed from Slater-type (STO) or Gaussian-type orbitals (GTO), shown in Figure 2.1.

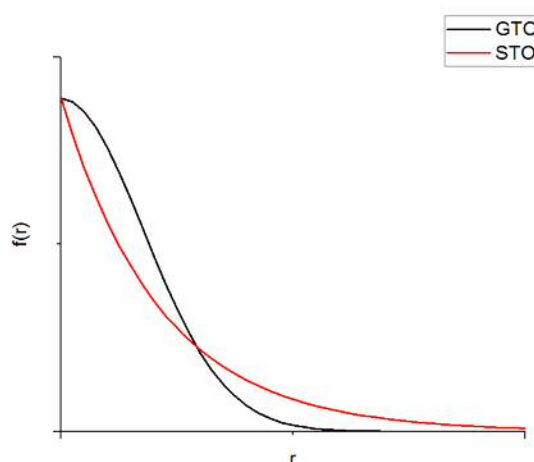


Figure 2.1: STO and GTO functions. STO is shown in red, GTO in black.

STOs provide better descriptions than GTOs both at  $r = 0$  and at large  $r$  values; STOs correctly describe the cusp at the nucleus ( $r = 0$ ), given by the finite slope and display the expected exponential decay over large  $r$ . In contrast, GTOs exhibit a zero slope at  $r = 0$  and a more rapid decay over large  $r$ .

While STOs are more accurate, computing the three- or four-centre two-electron integrals seen in the HF equations is computationally intensive, largely limiting their use to small chemical systems. GTOs, though less accurate, are less computationally demanding than STOs and may be used to express atomic orbitals as linear combinations of the Gaussian basis functions as an approximation to a STO. These orbitals are known as contracted Gaussian functions (CGF). Despite the increase in the number of basis functions using CGF, calculation of the required two-electron integrals is simpler, meaning that GTOs are commonly used.

The smallest collection of these functions/orbitals that describes all electrons present in the ground-state of the free atoms that constitute the chemical system is known as the minimal basis set. Here, each occupied orbital is described by a single basis function. The traditional example of a minimal basis set is STO-3G, where each orbital is described by three CGFs.

Unfortunately, such simple basis sets are often inadequate for describing complex chemical systems. These basis sets may be improved by increasing the number of basis functions used. The number of basis functions used per atomic orbital is an indicator of the reliability of a given basis set; more reliable basis sets tend to use multiple basis functions per atomic orbital. One common alteration to the minimal basis set is to double or triple the number of basis functions, resulting in double- $\zeta$  (DZ) and triple- $\zeta$  (TZ) basis sets. Alternatively, it is common to use one basis function per atomic orbital for the core electrons and multiple such functions for the valence electrons. This is known as a split valence basis set. This linear combination of basis functions affords the valence electrons more flexibility than the core electrons, leading to a more accurate description of these outer electrons.

Further improvement can be obtained by the inclusion of polarisation and/or diffuse functions. Polarisation functions, usually denoted by an asterisk (\*), consist of the additional of basis functions with greater angular momentum than is required in the free atom. This may include the addition of p-functions on hydrogen and *d*-functions on first-row elements. These polarisation functions allow for more accurate descriptions of electrons in the molecular environment. Diffuse functions, denoted by a plus sign (+) involve the addition of shallow Gaussian functions to model the ‘tail’ of atomic orbitals *i.e.* the region far from the nucleus. These are especially effective for ‘soft’ anionic systems. As a consequence of the additional functions, these more complex basis functions require more computational time and resources than their simpler counterparts.

One split valence basis set used in this work was 6-31G\*. This is a standard basis set, where the core electrons are modelled by six contracted Gaussian functions and the valence electrons are then described by a double- $\zeta$  basis set of three Gaussians and one Gaussian. This basis set also utilises the first set of polarisation functions, corresponding to the addition of *d*-orbitals on all non-hydrogen atoms.

### 2.4.1 Effective Core Potentials

Quantum mechanical modelling of elements in the lower rows of the periodic table poses two significant challenges. First, these elements possess a large number of core electrons. Modelling these electrons would require a large number of basis functions to allow for expansion of the corresponding orbitals. As such, the computational cost associated with explicit modelling of core electrons in heavy atoms is relatively prohibitive. Secondly, these elements often display relativistic effects; as the core electrons interact with the increased nuclear charge of these elements their velocities reach non-negligible fractions of the speed of light. These effects are often significant for many chemical properties, but the standard Hamiltonian (see section 2.2) is incapable of accounting for these effects.

These issues may be solved by assuming that the core electrons are fixed and can be modelled using a suitable analytical function, a solution first proposed by Hellman in 1935.[13] Only the valence electrons are treated explicitly. The function used to model the core electrons is known as an Effective Core Potential (ECP) or Pseudopotential (PP).

These ECP methods may be used to obtain accurate results at a fraction of the computational cost required to model all electrons within these heavy atoms. Used in *ab initio* theory, ECPs are able to account for all electrostatic and quantum mechanical interactions between the valence electrons and the core *i.e.* Coulombic attraction and Pauli repulsion effects, in addition to accounting for the relativistic effects seen in heavy atoms.

The ECP used for large sections of this research was a Stuttgart-Dresden (SDD) [14, 15] effective core potential. In the case of  $\text{Pt}^{\text{II}}$ , this replaces 60 core electrons, or up to and including those with principal quantum number ( $n$ ) = 4. This means that for the  $\text{Pt}^{\text{II}}$  ion ( $[\text{Xe}]4f^{14}5d^8$ ), only the  $5s^25p^66s^25d^8$  electrons are modelled explicitly.

## 2.5 Molecular Mechanics

Large molecules can often be modelled without resorting to quantum mechanical methods; the most well-known of these alternative methods is molecular mechanics (MM). Here, atoms are treated as single particles with an associated mass and charge *i.e.* electrons are not considered as individual particles, so electronic motion is completely neglected. This means that finding solutions to the electronic Schrödinger equation is avoided; instead, energy is calculated as a function of nuclear positions and bonding information must be provided explicitly.

In MM, molecules are described by the ‘ball and spring’ model, where the atoms (balls) are assigned types according to their mass, charge and connectivity, and the chemical bonds (springs) possess different force constants and an equilibrium bond length equal to the experimental or calculated bond length. This means that the dynamics of the system can be approximated by Newtonian mechanics.

The energy of a given system is described by a series of inter-atomic potentials, which represent the different interactions present, as presented in Equation 2.27.

$$E_{Tot} = E_{stretch} + E_{bend} + E_{torsion} + E_{VDW} + E_{el} \quad (2.27)$$

Here,  $E_{stretch}$  is the energy function for stretching a given bond,  $E_{bend}$  accounts for the energy required to bend an angle,  $E_{torsion}$  corresponds to the energy for rotation about a bond,  $E_{VDW}$  and  $E_{el}$  represent the non-bonding van der Waals and electrostatic interactions. The above equation may be expanded to give a more detailed breakdown of the constituent terms, shown in Equation 2.28.

$$E_{Tot} = \frac{1}{2} \sum_{bonds} k_i (l_i - l_{i,0})^2 + \frac{1}{2} \sum_{angles} k_i (\theta - \theta_{i,0})^2 + \frac{1}{2} \sum_{torsions} V_n (1 + \cos(n\omega - \gamma)) \\ + \sum_{i=1}^N \sum_{j=i+1}^N (4\varepsilon_{ij} [(\frac{\sigma_{ij}}{r_{ij}})^{12} - (\frac{\sigma_{ij}}{r_{ij}})^6] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}) \quad (2.28)$$

The first term is a sum of the interactions between bonded pairs of atoms, modelled using a harmonic potential. This potential gives an increase in energy as the bond length,  $l_i$ , deviates from the reference bond length,  $l_{i,0}$ . The second term is a sum



over every intramolecular bond angle and is also modelled using a harmonic potential. The third term is a periodic torsional potential used to represent the energy change involved in bond rotation. The final term combines the van der Waals and electrostatic interactions between all pairs of atoms ( $i$  and  $j$ ) *i.e.* those in different molecules or those in the same molecule but separated by at least three bonds. The van der Waals term is usually expressed as a Lennard-Jones potential, where the attractive forces fall as  $r^{-6}$ , while repulsive forces rise as  $r^{-12}$  and  $r$  is the inter-atomic distance. The electrostatic terms are modelled using Coulombic potentials, where the energy falls as  $r^{-1}$ .

From this energy function of atomic coordinates, molecular geometries and relative energies can be found by optimising the structures to a minimum of  $E_{Tot}$  *i.e.* a minimum on the potential energy surface.

The collection of atom type definitions and bonding parameters, including equilibrium bond lengths, force constants, bond angles *etc.* used to describe a system is known as the force-field (FF). The exact implementation of the terms in equation 2.28 defines the different FFs, and more complex FFs include additional terms, such as the out-of-plane energy,  $E_{oop}$ , the stretch-bend energy,  $E_{stb}$ , and the cross-coupling energy,  $E_{cross}$ , which describes coupling between  $E_{stretch}$ ,  $E_{bend}$  and  $E_{torsion}$ .

There are a variety of different FFs available for studying chemical systems, including AMBER [16], MMFF[17], CHARMM[18], GROMOS[19], OPLS[20] and UFF[21], each designed for modelling a general class of molecules. Selection of the most appropriate FF is therefore important in producing accurate calculations.

Arguably the biggest advantage of MM methods is that calculations can be performed rapidly, allowing very large systems to be modelled at a fraction of the computational cost of QM methods. In the limit of large molecules, the computational time for calculating the forcefield energy increases as approximately the square of the number of atoms, which compares favourably with the  $N^4$  variables ( $N$  = number of electrons) that determine computational time in wavefunction-based methods (see section 2.3).

One major problem with MM methods is the parameters used to define a force-field;

if a molecule has not been crystallised or modelled previously, then it is likely that only poor parameters exist, if any, meaning that accurate modelling of these systems using MM is difficult. In addition, generation of new parameter sets is an arduous task. Despite this, MM is still widely used in computational chemistry, particularly for modelling biological systems.

### 2.5.1 Molecular Mechanics for Transition Metal Complexes

The extension of MM methods to transition metal (TM) systems is complicated by two factors:

Firstly, the MM approach cannot consider the effects of *d*-orbital electrons on complex structures, since the electrons are handled implicitly in the calculations. Secondly, the typical MM formalism (see Equation 2.28) relies upon a single reference angle for a given ligand-metal-ligand (L-M-L) atom set.

In traditional organic chemistry, carbon atoms typically adopt one of three standard geometries, each with a single reference angle *i.e.* tetrahedral (bond angle  $109.5^\circ$ ), trigonal planar ( $120^\circ$ ) or linear ( $180^\circ$ ). In contrast, many metal complexes require multiple reference angles. For example, a homoleptic octahedral complex would require two unique L-M-L reference angles ( $90^\circ$  and  $180^\circ$ ), see Figure 2.2, lower row, left side) while a trigonal bipyramidal species would require three reference angles ( $90^\circ$ ,  $120^\circ$  and  $180^\circ$ , see Figure 2.2, lower row, right side), meaning that the traditional MM formalism is inadequate for describing TM complexes.

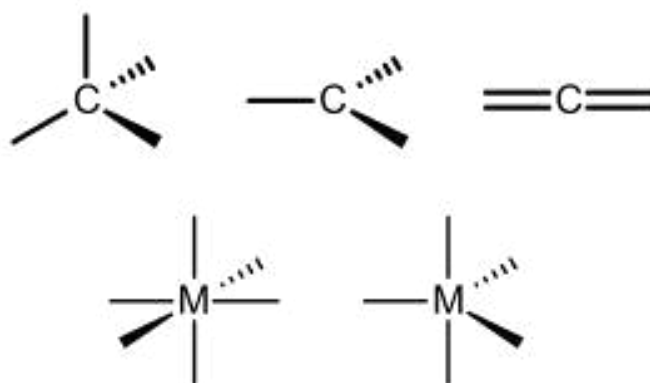


Figure 2.2: Reference angles in organic and transition metal chemistry. Top row: Tetrahedral, trigonal planar and linear carbons. Bottom row: Octahedral and trigonal bipyramidal metal centres.

### 2.5.2 Ligand Field Molecular Mechanics

Modelling transition metals using MM necessitates an extension to the traditional MM equation (see Equation 2.28) by including additional potential energy terms to describe the Ligand Field Stabilisation Energy (LFSE) and the influence of  $d$ -shell electrons (see Equation 2.29). One method that implements these factors is Ligand Field Molecular Mechanics (LFMM).[22]

$$E_{LFMM} = E_{stretch} + E_{bend} + E_{torsion} + E_{VDW} + E_{el} + LFSE \quad (2.29)$$

This idea may be thought of as an amalgamation of Ligand Field Theory and conventional MM, using MM for the ‘organic’ regions of a transition metal complex (*i.e.* the ligands) and an angular overlap model (AOM) to account for the LFSE at the metal centre. Here, the system to be studied is split into two overlapping regions:

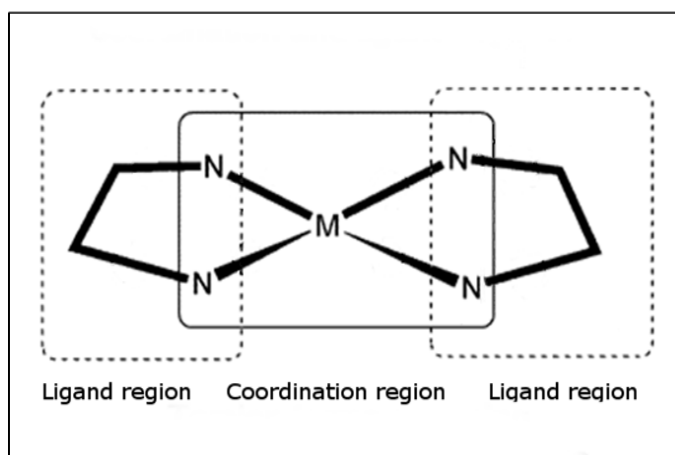


Figure 2.3: Overlapping ligand and coordination regions in Ligand Field Molecular Mechanics. Adapted from [22].

The larger ligand region contains all atoms except for the metal centre(s), while the coordination region(s) contain only the metal(s) and any immediately bonded atoms.[22] (See Figure 2.3.) The LFMM procedure is primarily concerned with the coordination region(s) and is responsible for the LFSE, metal-ligand bond stretching and ligand-metal-ligand angle bending.

The LFSE term is implemented in LFMM via the angular overlap model (AOM). The AOM calculates the strength of the interaction between individual ligand orbitals and metal  $d$ -orbitals based on the overlap between them; the angular overlap

name stems from the fact that the degree of overlap depends on the angle between the metal d-orbital and the approaching ligand orbital.

In the AOM, the total ligand field potential is a sum of the individual, local M-L contributions. This means that a given complex can be considered as a series of simple diatomic molecules. Subsequently, these diatomics are separated into  $\sigma$  and  $\pi$  contributions, represented by AOM parameters  $e_\sigma$ ,  $e_{\pi x}$  and  $e_{\pi y}$ . The AOM is a bond-centred approach, meaning the model intuitively corresponds to traditional bonding concepts *i.e.* the  $e_\sigma$  and  $e_\pi$  parameters relate to the degree of  $\sigma$ - and  $\pi$ -bonding in a given M-L bond. This means that the energies of the individual  $d$ -orbitals are a function of these AOM parameters for all ligands present.[22] The  $d$ -orbital energies are derived from the 5 x 5 ligand field potential matrix,  $V_{LF}$ . The matrix elements are given by Equation 2.30.

$$\langle d_i | V_{LF} | d_j \rangle = N \sum_l^N \sum_k^{symm} F_{ik}^l F_{kj}^l e_k^l \quad (2.30)$$

where  $d_i$  and  $d_j$  are  $d$ -orbitals, which sum over  $N$  ligands and three symmetry modes,  $e_\sigma$ ,  $e_{\pi x}$  and  $e_{\pi y}$ , which represent M-L  $\sigma$ -bonding and M-L  $\pi$ -bonding in two perpendicular directions, respectively.[22, 23]

The contribution of a particular ligand to a given  $d$ -orbital energy depends on the square of the orbital overlap, denoted  $S^2$ , between the  $d$ -orbital of interest and a ligand-based orbital. First, consider  $\sigma$ -donor interactions. The strongest  $\sigma$  interaction occurs when a ligand is located in an axial position, pointing directly at the  $z$ -axis lobe of the metal  $d_{z^2}$  orbital. Here, the orbital overlap is at a maximum *i.e.*  $S^2 = 1$  and the full  $e_\sigma$  amount of destabilisation is applied to the  $d_{z^2}$  orbital. The strength of other  $\sigma$  interactions is determined relative to this reference point. For example, if the ligand is moved to an equatorial position, the overlap is greatly reduced due to the smaller ‘doughnut’ band of the  $d_{z^2}$  orbital. Here  $S^2 = 0.25$ , so the  $d_{z^2}$  orbital is only destabilised by  $0.25e_\sigma$ . Similar changes in orbital energy arise from other  $\sigma$ -donor interactions between metal d-orbitals and ligand orbitals, and the magnitude of the energy change depends on the ligand location and the d-orbital in question (see Figure 2.4 and Table 2.1).

Many ligands also interact with metal  $d$ -orbitals in a  $\pi$ -bonding manner, either acting as  $\pi$ -acceptors or  $\pi$ -donors. The strongest  $\pi$ -interaction takes place where

a ligand  $\pi^*$  orbital overlaps with a metal  $d_{xz}$  orbital. The  $d$ -orbital energy is then perturbed by the full  $e_\pi$  amount. As before, other  $\pi$  interactions are weaker than this interaction, with the magnitude dependent on the degree of overlap between the orbitals (see Table 2.2). However, the sign of this perturbation (*i.e.* whether the  $d$ -orbital is stabilised or destabilised) depends on the  $\pi$ -bonding characteristics of the ligand. For  $\pi$ -acceptor ligands, the  $d$ -orbitals are stabilised by this interaction, while for  $\pi$ -donors, the  $d$ -orbitals are destabilised. Importantly, the orbital overlap in  $\pi$  interactions is smaller than the overlap for the  $\sigma$  interactions above, so the AOM parameters  $e_\pi < e_\sigma$ .

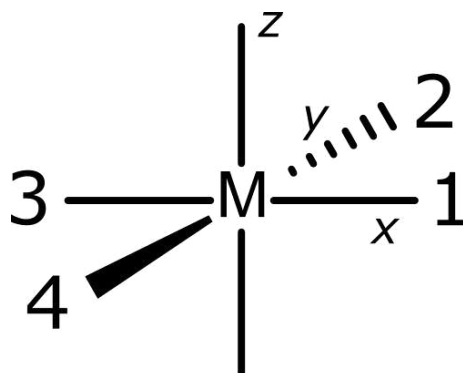


Figure 2.4: Ligand positions in a square planar transition metal complex. These positions are used for Tables 2.1 and 2.2 below.

Ligand Position	$d_{z^2}$	$d_{x^2-y^2}$	$d_{xy}$	$d_{xz}$	$d_{yz}$
1	0.25	0.75	0	0	0
2	0.25	0.75	0	0	0
3	0.25	0.75	0	0	0
4	0.25	0.75	0	0	0

Table 2.1: Sigma interactions (in units of  $e_\sigma$ ) for square planar ligands with metal  $d$ -orbitals. Ligand positions are given in Figure 2.4. Adapted from [24]

Ligand Position	$d_{z^2}$	$d_{x^2-y^2}$	$d_{xy}$	$d_{xz}$	$d_{yz}$
1	0	0	1	1	0
2	0	0	1	0	1
3	0	0	1	1	0
4	0	0	1	0	1

Table 2.2: Pi interactions (in units of  $e_\pi$ ) for square planar ligands with metal  $d$ -orbitals. Ligand positions are given in Figure 2.4. Adapted from [24]

The exact relationship between an AOM parameter and the  $d$ -orbital energy is given by the F-factors in Equation 2.30 above, which are trigonometric functions that depend on the angular coordinates of the ligands [22, 23] (see Figure 2.5). The  $d_{z^2}$  example is given in Equation 2.31:

$$E_{d_{z^2}} = F_\sigma^2(d_{z^2})e_\sigma = \frac{1}{16}(1 + 3\cos 2\theta)^2 e_\sigma \quad (2.31)$$

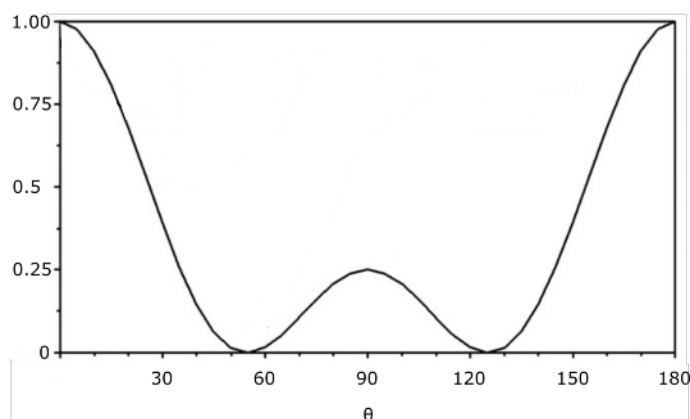


Figure 2.5: Plot showing the variation of ligand-metal  $dz^2$  orbital overlap with angle, where  $\theta$  is the deviation of the coordinating ligand from the  $z$ -axis.[23]

In the case of square-planar complexes, such as platinum(II), the resulting  $d$ -orbital energies may be written as shown in Equation 2.32. This produces the familiar  $d$ -orbital splitting diagram, shown in Figure 2.6.

$$\begin{aligned}
 d_{z^2} &= e_\sigma - 4e_{ds} \\
 d_{x^2-y^2} &= 3e_\sigma \\
 d_{xy} &= 4e_\pi \\
 d_{xz} &= 2e_\pi \\
 d_{yz} &= 2e_\pi
 \end{aligned} \quad (2.32)$$

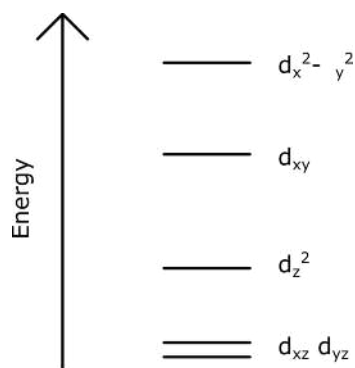


Figure 2.6:  $d$ -orbital splitting energies for a square planar complex.

The AOM parameters  $e_\sigma$ ,  $e_{\pi x}$  and  $e_{\pi y}$  are given as a function of the bond length,  $r$ :

$$E_\lambda = a_0 + a_1 r + a_2 r^{-2} + a_3 r^{-3} + a_4 r^{-4} + a_5 r^{-5} + a_6 r^{-6} \quad (2.33)$$

where  $E_\lambda$  represents either  $e_\sigma$ ,  $e_{\pi x}$  or  $e_{\pi y}$  and the  $a_i$  terms are empirically determined parameters.[22] The many terms present in this equation allow a great deal of flexibility, but usually only one term is used, while the others are set to zero. While all ligands require an  $e_\sigma$  parameter, not all ligands need  $e_\pi$  parameters; many ligands do not form  $\pi$  interactions with the metal due to the nature of the bonding within the ligand itself.[22] As the parameters  $e_{\pi x}$  and  $e_{\pi y}$  suggest, LFMM is able to handle asymmetric  $\pi$ -bonding by distinguishing between the  $e_{\pi x}$  and  $e_{\pi y}$  contributions.

In addition, there is a  $d-s$  mixing parameter,  $e_{ds}$ , which describes the configuration interaction between the metal valence  $s$ -orbital and a  $d$ -orbital with matching symmetry; this is important in square planar complexes.[22, 25] While the  $e_{ds}$  term has little influence on the metal geometry, it has been shown to be important in predicting the relative energies of metal complexes.[25] This  $e_{ds}$  term is able to produce correct  $d$ -orbital energies in these cases, while the original  $\sigma$ - and  $\pi$ - only construction of LFMM failed. The typical example of this is the  $d_{z^2}$  orbital in planar systems with molecular point group  $D_{4h}$  *e.g.*  $[\text{CuCl}_4]^{2-}$ , where the copper(II)  $4s$  and  $3d_{z^2}$  orbitals both have  $A_{1g}$  symmetry. In the absence of any  $d-s$  mixing, LFMM predicts the  $d_{z^2}$  orbital to be approximately  $6000 \text{ cm}^{-1}$  higher than it should be; addition of the  $d-s$  mixing term modifies this by  $-4e_{ds}$  (*i.e.*  $e_{ds} = 1500 \text{ cm}^{-1}$ ), giving the correct  $d_{z^2}$  orbital energy.[22]

Finally, the  $e_{\text{pair}}$  parameter facilitates accurate descriptions of different spin state energies. The LFSE describes the one-electron contribution to the ligand field and

therefore always favours the low-spin state electron configuration. Conversely, the two-electron electrostatic terms (*i.e.*  $d-d$  interelectronic repulsion) favour the high-spin state, while the true spin state is a balance between these opposing factors. As discussed above, traditional MM methods require distinct parameter sets for each possible spin state; the  $e_{pair}$  term is used to model the effects of interelectronic repulsion, allowing LFMM to use a single parameter set for all possible spin states.[22]

In LFMM, the standard forcefield file must be slightly modified to include terms that span the ligand and coordination regions. With reference to Figure 2.3, the terms required for LFMM to function correctly are threefold:

- (i) M-N-C bond angle
- (ii) M-N-C-C torsion angle and
- (iii) N-M-N-C torsion angle

However, the implementation of these terms within the forcefield is not uniform. For the M-N-C and M-N-C-C terms, the magnitude of the forcefield parameters is derived from existing comparable terms. In contrast, the force constants for any N-M-N-C torsions must be set to zero. While there would be no issue for *cis*-N ligands, *trans*-N ligands would place the N-M-N triad in a straight line, where the twist angle cannot be defined.[22]

As already mentioned, the forcefield also contains definitions for all available atoms in the relevant forcefield. Thus, atom types for metals must be defined and added to the chosen forcefield.[22] In this work, we are primarily interested in platinum (Pt(II)) systems, denoted by a new atom type “PT+2”. In addition, care should be taken regarding the ligand donor atom types. To avoid tedious definition and parameterisation of new atom types, the donor atom types are carried over from the free ligand and used when the ligand is bound to a metal centre.

Both metal and ligand atom types must also be assigned correctly, which may be achieved automatically using the smiles arbitrary target specification (*SMARTS*) [26] extension to the simplified molecular input line entry system (*SMILES*) [27] notation within the forcefield file. In terms of *SMARTS*, this means that for the ligand donor atoms, the general transition metal symbol [#T] must be added to the existing ligand atom type definition, while PT+2 may be matched by any atom



with atomic number 78.[22]

While this approach ensures that the treatment of metal-free systems remains unchanged, it also means that any changes in the ligand which occur when binding to a metal centre cannot be captured by LFMM. However, as pointed out by Bol and Comba,[28] the coordination of organic ligands to a transition metal leads to ‘*small but significant*’ changes in the ligand structure, with bond length changes between donor atom- $C_\alpha$  of approximately 1% - the bond in question is illustrated in Figure 2.7.

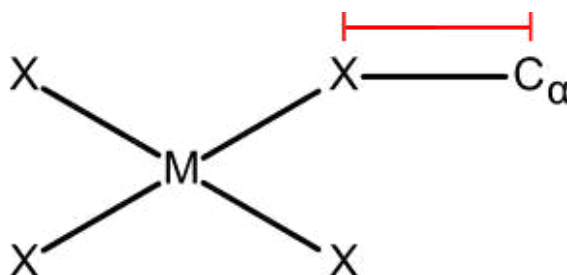


Figure 2.7: Illustration of the ligand atom -  $C_\alpha$  atom bond, highlighted in red. LFMM fails to model changes in this bond length upon metal coordination.[28]

While the FF file handles the ligand region (in addition to the terms specified above), LFMM uses a separate parameter (.PAR) file for all metal-ligand interaction terms, including AOM parameters. Such separation of the parameters for the two regions allows greater control over the metal-ligand bond stretching term. For example, this permits the use of Morse functions for metal-ligand bond stretches even in cases where the original forcefield uses simple harmonic functions for all bond stretch terms *e.g.* AMBER. Each metal-ligand bond therefore requires dissociation energy ( $D$ ), reference bond length ( $r_0$ ) and curvature parameter ( $\alpha$ ) in the .PAR file (see Equation 2.34)[22]

$$E_{Morse} = D(1 - e^{-\alpha(r-r_0)})^2 \quad (2.34)$$

LFMM does not require L-M-L angle bend terms, and thus avoids the aforementioned issue regarding multiple reference angles for a single L-M-L atom set. In place of these terms, the Valence-Shell Electron Pair Repulsion (VSEPR) method is used to generate the approximate coordination geometry [22] *i.e.* ligands around a metal centre will arrange themselves so as to obtain the greatest possible separation between them. In addition, LFMM uses a purely repulsive 1,3 ligand-ligand term

in the form of a Lennard-Jones potential (see Equation 2.28), allied to the L-M-L angular component of the LFSE, allowing LFMM to model ‘*any number of ligands and any symmetry*’.[22]

### 2.5.2.1 Technical Asides

The LFMM model was incorporated into an existing piece of molecular modelling software, Molecular Operating Environment (MOE)[29], by the Deeth Inorganic Computational Chemistry Group at Warwick University. A copy of this program, *d*-orbital molecular mechanics in MOE (DommiMOE),[22] was obtained from Professor Deeth and used extensively throughout this work.

In addition, the LFMM model was recently ported to the DL-POLY Classic molecular dynamics software package.[30] This DL-POLY-LF allows for long timescale molecular dynamics simulations to be carried out on high-performance computing resources using the LFMM technique. However, DommiMOE must still be used to generate the DL-POLY-LF input files.

One of the advantages of MOE for traditional MM work is that it automatically generates parameters for any interaction that is not defined within the FF file. However, in the case of transition metals, these generated parameters are typically poor.[22] DommiMOE must therefore be able to prevent the use of these parameters, which is achieved by performing an additional MM calculation on the central coordination region (see Figure 2.3). The potential energy and first derivatives from this calculation are subtracted from the total energy and derivatives, which eliminates the unreliable MOE-assigned parameters from the calculations.[22]

Solvation must also be handled with care. The interaction of the metal and ligand-donor atom charges with the solvent dielectric medium is preserved by performing the coordination region calculation specified above with the solvation terms disabled. In this way, the energy of solvating the metal coordination region is retained.[22]

### 2.5.2.2 Alternative Approaches to the LFMM Problem

It should be noted that LFMM is by no means the only method for handling transition metals within molecular mechanics. A brief overview of a selection of these alternate methods is given below:

As mentioned above, one of the foremost problems with using MM to model transition metals species is the issue of multiple L-M-L reference angles. One accessible solution to this problem is focused on the valence shell electron pair repulsion (VSEPR) model. This ‘points on a sphere’ method was expanded to consider ligand-ligand repulsions and was used to successfully model the structures of several thousand compounds by Kepert. [31] While this approach is able to produce regular coordination structures (*i.e.* octahedral, tetrahedral), it cannot generate more complex geometries such as square planar structures, unless this is enforced by the surrounding ligand structure.

This design principle has also seen a great deal of use by the Comba group in their MOMEK package.[32] As before, the L-M-L bond angles are handled using ligand-ligand repulsion terms. However, MOMEK requires different parameter sets for axial and equatorial bonds.

The forcefield method VALBOND - based on valence bond theory - is also a popular alternative for modelling transition metal species.[33–36] VALBOND is based on the description of atoms, bonds and angles in terms of their hybridisation; here, transition metals are considered as hypervalent centres. The L-M-L angles are described by the overlap between  $sd^n$  hybridised metal-ligand orbitals. The VALBOND-TRANS variation is also able to model trans effects within complexes.[37]

The valence bond theory approach has also been implemented into the atomic multipole optimised energetics for biomolecular applications (AMOEBA) method for Cu(II) and Zn(II) ions, though this AMOEBA-VB method failed to reproduce the Jahn-Teller distortions of hexa-aqua Cu(II) complexes. [38] This lead directly to the incorporation of the AOM method into AMOEBA by the same authors.[39].

In addition, Carlsson and Zapata [40], have derived an AOM-type description of the

$d$ -orbital effects using analytical expressions for the angular potential at the metal, where the functions depend on the  $\sigma$  or  $\pi$  nature of the ligands. This is essentially a simplified version of the LFMM method described above.

Lastly, the full LFMM procedure has recently been implemented into the popular and freely available Tinker [41] modelling software and numerical results validated against DommiMOE.[42]

### 2.5.2.3 Final Remarks

Despite its strengths, LFMM remains a relatively niche technique. However, in recent years it has been successfully applied to a range of transition metal species, with particular focus given to copper, nickel and platinum systems.[30, 43–47] These applications are discussed in Chapter 1.

One of the problems with LFMM is due to the parameters required to reliably model transition metal complexes; a large number of these parameters are needed, but generating the parameters to accurately reproduce experimental or DFT structures is a very time-consuming process. The wider application of LFMM in this area of research is in direct competition with hybrid QM/MM methods, which require no such parameterisation, but come with their own complications.

In addition, the primary implementation of LFMM is within MOE, a powerful but expensively licensed commercial software package. This clearly limits the uptake of the LFMM approach in the wider modelling community, an issue that may be partially remedied by its inclusion in alternative packages outlined above, as surmised by Foscatto *et al.*: ‘*broader availability of LFMM could open up a range of new and interesting applications to large systems, various computer simulation techniques and large-scale design studies*’.[42]

## 2.6 Semi-Empirical Methods

The computational cost of performing Hartree-Fock calculations increases as  $N^4$ , where  $N$  is the number of basis functions. The greatest contributor to this phenomenon is the number of two-electron integrals required to complete the calculation *i.e.* the J and K integrals appearing in the Fock matrix elements. Semi-empirical methods are molecular orbital theories based on the Hartree-Fock method, but reduce the computational demand by reducing the number of two-electron integrals required. One approach would be to estimate the value of these integrals accurately in an *a priori* fashion, such that no numerical integration must be performed. This forms the basis of the so-called semi-empirical modelling techniques.

The first step in this process is to consider only the valence electrons explicitly. The core electrons are then dealt with by introducing function(s) to model the repulsion due to the nuclei and core electrons, or by simply reducing the nuclear charge. The valence electrons are then described by a minimal set of basis functions *i.e.* using only the minimum number of basis functions required to consider the electrons in the free atom (see section 2.4)

### 2.6.1 ZDO Approximation

The Zero Differential Overlap (ZDO) approximation is key to semi-empirical methodologies. This approximation neglects the products of basis functions that depend on two electrons located on different atoms. The Coulomb integrals in a Hartree-Fock calculation consider the interaction of two electrons in regions defined by the basis functions; if these basis functions are on different atoms, then the value of the integral will be almost zero and may be neglected. In large systems, it is possible to discard many of these two electron integrals, meaning that the required number of calculations is greatly diminished. This process should still provide an answer relatively close to that obtained from a full calculation.

The ZDO approximation has several consequences: Firstly, it simplifies the Roothaan-Hall equation (Equation 2.14) to  $FC = C\varepsilon$  *i.e.* the overlap matrix is now a unit matrix. Secondly, all three and four-centre, two electron integrals are neglected. Finally, all three-centre, one electron integrals are set to zero. The remaining integrals

are assigned values based on experimental data or accurate computational calculations. The number of integrals neglected by the ZDO approximation, in addition to the method for assignment of parameter values, defines the different semi-empirical methods.

The first semi-empirical method was Complete Neglect of Differential Overlap (CNDO), where parameters were developed from *ab initio* calculations.[48–51] Following this, several other variants have been developed, including Intermediate Neglect of Differential Overlap (INDO), Modified Intermediate Neglect of Differential Overlap (MINDO) and Neglect of Diatomic Differential Overlap (NDDO).

## 2.6.2 NDDO Approximation

One of the more recent advances in semi-empirical methodology is the Neglect of Diatomic Differential Overlap (NDDO) approximation. Here, no further simplifications (*i.e.* beyond the ZDO approximation) are required. The semi-empirical methods employed in this research all utilise the NDDO approximation.

### 2.6.2.1 Parameterised Model 3

Published in 1989, Stewart’s Parameterized Model 3 (PM3) [52] proved to be a robust semi-empirical model. In contrast to existing methods at the time, all parameters were simultaneously derived via an automated procedure. In addition, these parameters were tested using a much larger data set than previous efforts.

PM3 initially provided parameters for just 12 atoms (H, C, N, O, F, Al, Si, P, S, Cl, Br and I) but subsequent work has enabled parameterisation of all main-group elements.[53] One of the major problems with earlier semi-empirical techniques was in predicting hydrogen bond lengths and geometries. To correct this, Dewar *et al.*, in their AM1 method [54], and later, Stewart, in PM3, modified the nuclear repulsion term by ‘*adding a stabilisation Gaussian function to the MNDO core-core interaction.*’[54, 55] PM3 uses two Gaussian functions for this purpose, as opposed to the four in AM1.

### 2.6.2.2 Parametric Model 6

The relatively recent Parametric Model 6 (PM6) method was developed by Stewart and contains parameters for 70 atoms. [55] PM6 incorporates the core-core diatomic interaction parameters of Voityuk [56] as well as Thiel’s *d*-orbital approximation[57, 58], which enables NDDO methods to model the transition metal elements. Subsequent expansion of this work has lead to variants which implement dispersion, hydrogen- and halogen-bonding corrections.[12, 59–64] These corrections typically require some re-parameterisation of the original method.

### 2.6.2.3 Parametric Model 7

More recently, Parametric Model 7 (PM7) has been published,[65] containing parameters for 70 elements in total. The PM7 method includes a “D2”-type correction for dispersion forces , though only elements H, He, B, C, N, O, F, Ne, P, S, Cl, Ar, Br and Kr are currently supported.[65] Other elements are handled using core-core attractive Gaussian terms to mimic dispersion effects.[66] The dispersion and hydrogen bond corrections were implemented before parameter optimisation was performed, meaning that PM7 is designed to reproduce both intermolecular interaction energies and heats of formation.[65]

## 2.7 Exploring Conformational Freedom

The physical, chemical and biological properties of a molecule often depend on the conformations it can adopt. Conformational analysis – the study of conformations of a molecule and the influence on its properties – rose to prominence in the mid-20<sup>th</sup> century, and allied to emerging techniques such as NMR and X-ray crystallography, became an important tool for modern chemists. As such, conformational analysis is widely used in drug discovery and (biological) catalyst design.[3, 67]

The conformational search – identifying the possible conformations of a molecule – is a critical part of conformational analysis. Conformational searching is used to locate minima on the molecular potential energy surface; ideally this would produce the molecule’s preferred conformations – those which determine its properties and behaviour.

Generally, the aim is to identify every minimum energy conformer. Often, identifying these minimum energy positions is difficult due to the exponential increase in the number of minima with the number of variables, such as rotatable bonds. This is known as the combinatorial explosion problem. In these cases, it is often more useful to sample a representative set of local minima.

The classic example which demonstrates this problem is exploration of the conformations of linear alkanes. Here, there are 3 possible conformations around each C-C-C-C bond. Thus for any linear alkane with  $n$  rotatable bonds, there are  $3^n$  possible conformations. This is illustrated in Table 2.3, for  $CH_3(CH_2)_{n+1}CH_3$  systems:

n	Number of possible conformations ( $3^n$ )	Time (1 conformation = 1 second)
1	3	3 seconds
5	243	4 minutes
10	59049	16 hours
15	14348907	166 days
20	3486784401	40356 days

Table 2.3: Combinatorial explosion as illustrated by linear alkanes.[1]



For larger molecules, this systematic approach is clearly not possible. However, there are several alternate methods which are applicable to large molecular systems. In this section, these computational methods for conformational searching will be introduced.

### 2.7.1 Stochastic Search Methods

The first conformational search method to be discussed is the stochastic or ‘random’ search. In the stochastic approach, a random change is made to the current conformation at every iteration. This change is usually achieved in one of two ways:

The first of these is the Cartesian method. Here, the new conformation is generated by adding a random amount to the (x, y, z) coordinates of all atoms in the molecule. The perturbed geometry produced is then refined using energy minimisation.

The second method applies a random change to the rotatable bonds in a system. In this approach, the torsion angles are randomly selected and rotated by a random value. Once again, the structure produced is energy minimised. Stochastic conformational searches performed in the course of this research used MOE [29]/Dommi-MOE [22] which employs this dihedral approach.

In both cases, if the generated conformer has not been identified previously - as determined by a user-defined root mean square deviation (RMSD) cut off - and providing it meets any additional user-specified criteria, such as maximum energy cut-offs - it is saved. Over time, a database of unique conformations is produced. A new perturbed structure is then generated from one of these conformers and the cycle continues.

The procedure is terminated once a given number of iterations have been performed or no new conformations can be found. In stochastic search methods, there is no natural end point, since it is almost impossible to be certain that all minimum energy conformations have been found. In practice, the search is usually performed until no new conformations are found. This means that in the process of a conformational search, each region of the conformational space is explored many times.

The stochastic method has been used as a point of comparison for more recent conformational search methods in a number of studies focused on flexible molecules, similar to those investigated here.[67–70] However, it was widely out-performed by the more modern methods, proving its lack of suitability for this research (see Section 2.7.4).

## 2.7.2 LowMode Molecular Dynamics

In 1996, the ‘Low Mode Search’ (LMOD) conformational search method was proposed by Kolossváry and Guida[71]. This method focused on following low-frequency vibrations in the structure; for every iteration, the molecular vibrations are calculated and the system is systematically perturbed along low-frequency vibrational modes. The conformation produced is then optimised.

The underlying principle of LMOD is attractive for conformational analysis *i.e.* the use of low-frequency vibrational modes to “hop” between low-energy conformations. LMOD also assumes that these transitions between states are able to completely explore the ensemble of low energy conformations.[67] LMOD is therefore able to explore conformational space well, and is also able to account for non-bonded interactions, meaning that this approach is suitable for complex molecular systems. Unfortunately, the vibrational mode analysis conducted at the beginning of each iteration comes at a significant computational cost, particularly for large systems. To counteract this, the LLMOD variation was developed,[72] where the vibrational analysis is performed at the beginning and the vibrational modes produced are used for the duration of the conformational search.

LowMode MD (LMMD), proposed by Labute in 2010 [67], is a recent addition to the array of conformational search methods available. Briefly, this may be considered to be a type of quenched molecular dynamics; a short (1 ps) molecular dynamics simulation - with random initial velocities - and an energy minimisation, followed by collating the visited conformations.[69] The success of LMMD stems from its design; it screens out high-frequency vibrational modes and concentrates kinetic energy along low frequency modes in molecular dynamics trajectories.[67] LMMD is able to analyse complex molecular systems such as protein loops and macrocycles - taking into account complex non-bonded interactions - making it ideal for this research. It has been shown to be able to locate low-energy conformations for a

variety of structures in a computationally efficient manner.[67–69] It is notable that the LMMD approach may be applied using any forcefield, provided that a gradient is available to perform the relevant energy minimisations.[67]

The LMMD protocol is outlined below:

1. [*Start*] Energy minimise the system and set the conformation list to empty.
2. [*Velocities*] Generate a random set of atomic velocities from the Maxwell-Boltzmann distribution and analytically remove rigid body rotational and translational velocity, if appropriate.
3. Filter the velocities so that there is little kinetic energy along high-frequency vibrational modes.
4. Scale the resulting velocities to 300 K.
5. Perform a constant temperature molecular dynamics simulation for 1 ps.
6. [*Minimisation*] Energy minimise the conformation produced.
7. [*Save*] If the conformation is not in the conformation list (taking symmetry into account), then add the conformation to the list. When saving conformations, any two conformers are deemed to be duplicates if the root-mean-square deviation between them is less than a user-defined threshold. Conformations may also be rejected on strain energy grounds.
8. [*Terminate*] If sufficient conformations have been generated or a user-specified iteration limit is exceeded then terminate the search; otherwise, go to Step 2.

In practice, the user must specify the total number of iterations of the procedure *i.e.* the total number of attempts to find new conformations. In addition, a strain-energy cut-off and a root-mean-square-distance cut-off must be defined. Finally, the search will terminate if a certain number of successive failures to generate new conformations occurs; the user must specify this number of iterations.

The LMMD approach is built on the same foundations as the preceding (L)LMOD methods, in that it involves perturbing a conformation by concentrating the kinetic energy present in a system along low-frequency vibrational modes and using these transitions to explore the conformational space. However, in LMMD this is achieved without the need to calculate vibrational frequencies, second derivatives

of the Hessian matrix or eigenvectors and without the need to specify the low-frequency threshold value. In LMMD, the undesirable high-frequency vibrations are dampened, which effectively removes these vibrations from the calculation of velocities.[67]

While LMMD is similar to (L)LMOD, differences can be identified in the perturbations employed by each. In the (L)LMOD method, the vibrational analysis is performed once and the resulting low-frequency modes are used throughout the search, as previously mentioned. This assumes that the low-frequency vibrational modes identified in the original structure are applicable to all subsequent conformers. This appears implausible, especially if there is a significant change in molecular conformation. Conversely, the perturbations in LMMD are governed by the molecular forces of that particular conformation; in principle, this means that the perturbations in LMMD will be better than the previously discussed approaches.

In addition to this, the methods used to perturb coordinates differs significantly between (L)LMOD and LMMD; LMMD uses NVT molecular dynamics while (L)LMOD uses a linear, systematic perturbation along each low-frequency mode identified. This means that (L)LMOD may produce large conformational distortions, but LMMD should not, because all atoms obey the equations of motion throughout a conformational change.[67]

## 2.7.3 Molecular Dynamics

Molecular Dynamics (MD) simulations enable the calculation of structural and thermodynamic properties and phenomena, as well as the study of time-dependent behaviour of atomic or molecular systems, such as conformational changes. As a result, MD has become a powerful tool in chemistry, biology, physics and materials science.

### 2.7.3.1 Statistical Mechanics

MD simulations produce data at the microscopic level, such as atomic positions and velocities. The conversion of this microscopic data to macroscopic properties is the realm of statistical mechanics. Here, a brief introduction to the relevant statistical

mechanical concepts is given.

In a classical system, the Hamiltonian,  $H$ , depends on both particle positions,  $\mathbf{r}$ , and momenta,  $\mathbf{p}$ , and may be expressed in terms of kinetic energy,  $K(\mathbf{p})$ , and potential energy,  $U(\mathbf{r})$ , as shown in Equation 2.35.

$$H = H(\mathbf{r}, \mathbf{p}) = K(\mathbf{p}) + U(\mathbf{r}) = \sum_i \left[ \frac{p_i^2}{2m_i} + U(\mathbf{r}) \right] \quad (2.35)$$

Where  $m_i$  and  $p_i$  are the mass of particle  $i$  and the momentum of particle  $i$ , respectively. A single microscopic state of the system may be characterised by values  $\mathbf{r}, \mathbf{p}$ .

In a system of  $N$  atoms,  $6N$  values are required to define a state of the system - three coordinates per atom and three components of the momentum (in the x, y and z directions). Every combination of position and momentum defines a point in  $6N$ -dimensional phase space. An ensemble is a collection of points in this phase space. In MD, the system of interest moves through phase space, generating a series of points connected by time - for example, a molecular dynamics simulation performed using the NVE ensemble (*i.e.* constant system volume and energy) will sample the phase space with constant energy.

If the number of particles, the volume and the temperature of a system are fixed, the probability density,  $\rho$ , is given by the Boltzmann distribution function (Equation 2.36).

$$\rho(\mathbf{r}, \mathbf{p}) = \exp(-H(\mathbf{r}, \mathbf{p})/k_B T) / Z \quad (2.36)$$

Where  $Z$  is the canonical partition function - the integral over all phase space of the Boltzmann fractions given in the numerator.  $k_B$  is the Boltzmann constant.

This partition function plays a vital role in statistical mechanics - it allows for calculation of the thermodynamic average of any dynamic variable,  $A$ , such as kinetic energy, or atomic positions. The average of this property,  $A$ , across all possible states of the system is called the ensemble average and is written in Equation 2.37.

$$\langle A \rangle = \int \int d\mathbf{p}^N d\mathbf{r}^N A(\mathbf{p}^N, \mathbf{r}^N) \rho(\mathbf{p}^N, \mathbf{r}^N) \quad (2.37)$$

Here, there should be  $6N$  integral signs, for the  $6N$  positions and momenta of the particles, though just two are written for convenience. The final term,  $\rho(\mathbf{p}^N, \mathbf{r}^N)$ , is the probability density of the ensemble, indicating the probability of occupying a configuration of the system with momenta  $\mathbf{p}^N$  and positions  $\mathbf{r}^N$ . The ensemble average is then found by integrating over all possible configurations. During a molecular dynamics simulation, the instantaneous value of the variable  $A$  fluctuates due to interactions between particles. The value of  $A$  measured by simulation is an average of  $A$  over the simulation time, and is called the time average, shown in Equation 2.38.

$$A_{ave} = \lim_{\tau \rightarrow \infty} 1/\tau \int_{t=0}^{\tau} A(\mathbf{p}^N(t), \mathbf{r}^N(t)) dt \quad (2.38)$$

Where  $\tau$  is the simulation time.

As the simulation time is increased, the value of the time average approaches the true value, the ensemble average. According to the ergodic hypothesis, for an infinitely long simulation, the time average is equal to the ensemble average, provided the system reaches equilibrium *i.e.* when the simulation is sufficiently long, the entire phase space of the system will have been explored, so values of  $A$  obtained via time- and ensemble- averaging become equal.

One major problem with estimating thermodynamic quantities from MD simulations is that we assume that the generated configurations form a representative set of all possible configurations. In reality, this is impossible to verify. If all points in the  $6N$  dimensional phase space were sampled, it would be possible to calculate the partition function and the resulting MD trajectory would be termed ergodic *i.e.* the results are independent of the initial configuration. Unfortunately, such sampling is impractical even for very simple systems, meaning that simulations can only provide estimates of the true energy (or other dynamic variable) of the system.

A typical MD simulation time is of the order of tens to hundreds of nanoseconds, which is far too short a timespan to sample all phase space. This means that there is a risk that a simulation may become trapped in a small region of phase space for the entire trajectory, providing an unrepresentative sampling. To mitigate this, and test the sensitivity of the simulation results, several simulations may be performed with different starting conditions and/or configurations.

### 2.7.3.2 Newtonian MD

MD simulations produce configurations of a system connected by time via integration of Newton's equations of motion. Newton's second law (Equation 2.39) states that the force acting on a particle equals the rate of change of momentum.

$$F_i = m_i a_i = m_i \frac{\delta^2 r_i}{\delta t^2} \quad (2.39)$$

Where  $F_i$  is the force on particle  $i$ ,  $m_i$  is the mass of particle  $i$  and  $a_i$  is the acceleration of particle  $i$  and  $r_i$  is the coordinate.

Early MD simulations were carried out using a hard-body approach, where atoms act as solid spheres and do not interact before they collide in a purely elastic manner. This is not a realistic model of atomic interactions, and so more modern MD simulations use a Lennard-Jones type potential for the interaction between spheres *i.e.* the potential force between 2 atoms changes continuously with separation. In this case, the force on each particle changes if the particle changes its position, or if any of the other particles that it interacts with changes position. This many-body problem cannot be solved analytically, so the equations of motion must be integrated using a finite difference method.

Integrating the equations of motion breaks them down into many small steps, separated by a time difference,  $\delta t$ . The total force on each particle at time  $t$  may be calculated as a vector sum of all interactions with other particles. Subsequently, particle accelerations may be found; these values are combined with particle positions and velocities at time  $t$  in order to obtain new particle positions and velocities at time  $t + \delta t$ . It is assumed that the forces acting on each particle remains constant during the timestep (from  $t \rightarrow \delta t$ ). Forces acting on the particles must then be calculated at the new positions, which produce the next set of particle positions and velocities at  $t + 2\delta t$ . This procedure is repeated for the duration of the simulation.

There are several common algorithms for implementing this integration method, which are typically found in most modern MD simulation software. Each of these

methods assumes that the particle positions, velocities and accelerations can be approximated using a series of Taylor expansions (Equations 2.40-2.43).

$$r(t + \delta t) = r(t) + \delta t v(t) + \frac{1}{2} \delta t^2 a(t) + \frac{1}{6} \delta t^3 b(t) + \frac{1}{24} \delta t^4 c(t) + \dots \quad (2.40)$$

$$v(t + \delta t) = v(t) + \delta t a(t) + \frac{1}{2} \delta t^2 b(t) + \frac{1}{6} \delta t^3 c(t) + \dots \quad (2.41)$$

$$a(t + \delta t) = a(t) + \delta t b(t) + \frac{1}{2} \delta t^2 c(t) + \dots \quad (2.42)$$

$$b(t + \delta t) = b(t) + \delta t c(t) + \dots \quad (2.43)$$

Here,  $r$  is the particle position,  $v$  is the first derivative of the position with respect to time, corresponding to the velocity,  $a$  is the second derivative, the acceleration,  $b$  is the third derivative,  $c$  is the fourth derivative and so on.

The first – and most simple – finite difference method is the Verlet algorithm.[73] In this case, particle positions and accelerations at time  $t$  and particle positions at time  $t - \delta t$  are used to calculate new particle positions at time  $t + \delta t$ , shown in Equations 2.44-2.46.

$$r(t + \delta t) = r(t) + \delta t v(t) + \frac{1}{2} \delta t^2 a(t) + \dots \quad (2.44)$$

$$r(t - \delta t) = r(t) - \delta t v(t) + \frac{1}{2} \delta t^2 a(t) - \dots \quad (2.45)$$

Adding these equations together gives an expression for  $r(t + \delta t)$ , shown in Equation 2.46.

$$r(t + \delta t) = 2r(t) - r(t - \delta t) + \delta t^2 a(t) \quad (2.46)$$

While the Verlet algorithm is relatively easy to implement and has low memory requirements – it requires just two sets of particle positions and one set of particle accelerations – it is not without problems. New particle positions are found by adding a small  $\delta t^2 a t$  term to two much larger terms (see Equation 2.46), leading to lack of precision. In addition, particle velocities are not explicitly used in the calculations, and velocities at time  $t$  are not calculated until time  $t + \delta t$ . Lastly, the Verlet method is not self-starting – as already mentioned, new particle positions are calculated from current and preceding sets of positions. However, at time  $t = 0$ , there are no previous particle positions. As such, the Verlet method requires some other way of generating particle positions at time  $t - \delta t$ . This may be achieved using



a truncated Taylor series.

The second integration algorithm is the Leapfrog method, proposed by Hockney[74] and illustrated in Equations 2.47-2.49 which is similar to the above Verlet scheme.

$$v(t + \frac{1}{2}\delta t) = v(t - \frac{1}{2}\delta t) + \delta t a(t) \quad (2.47)$$

$$r(t + \delta t) = r(t) + \delta t v(t + \frac{1}{2}\delta t) \quad (2.48)$$

$$v(t) = \frac{1}{2}[v(t + \frac{1}{2}\delta t) + v(t - \frac{1}{2}\delta t)] \quad (2.49)$$

Here, particle velocities at time  $t - 1/2\delta t$  and accelerations at time  $t$  are used to calculate velocities at time  $t + 1/2\delta t$ , according to Equation 2.47. Positions at time  $t + \delta t$  can be calculated from particle velocities and positions at time  $t$ , according to Equation 2.48. This means that the calculated velocities ‘leapfrog’ over the particle positions, giving velocities at time  $t + 1/2\delta t$ . In turn, positions ‘leapfrog’ over the velocities, producing particle positions at time  $t + \delta t$ . This cycle repeats throughout the simulation.

In the leapfrog algorithm, particle velocities are explicitly included in the calculations, though velocities and positions are out of synchronisation by half a timestep. However, this explicit use of particle velocities allows the system to be coupled to an external heat bath for controlling simulation conditions. In addition, this algorithm does not require calculation of the difference between large numbers – as such, the numerical accuracy here is better than for the Verlet method.

The non-synchronous calculation of particle positions and velocities is addressed in the Velocity Verlet method.[75] Here, particle positions, velocities and accelerations are calculated simultaneously, without the loss of precision described in the original Verlet method.

$$r(t + \delta t) = r(t) + \delta t v(t) + \frac{1}{2}\delta t^2 a(t) \quad (2.50)$$

$$v(t + \delta t) = v(t) + \frac{1}{2}\delta t[a(t) + a(t + \delta t)] \quad (2.51)$$

$$v(t + \frac{1}{2}\delta t) = v(t) + \frac{1}{2}\delta t a(t) \quad (2.52)$$

$$v(t + \delta t) = v(t + \frac{1}{2}\delta t) + \frac{1}{2}\delta t^2 a(t + \delta t) \quad (2.53)$$

In this case, the particle positions at time  $t + \delta t$  are calculated as shown in Equation 2.50. However, calculation of particle velocities at time  $t + \delta t$  requires accelerations at both time  $t$  and  $t + \delta t$  (Equation 2.51). The velocities at time  $t + \frac{1}{2}\delta t$  are calculated using Equation 2.52. New forces acting on the particles are then calculated using the current positions, giving the acceleration term,  $a(t + \delta t)$ . Finally, velocities at time  $t + \delta t$  are found using Equation 2.53.

### 2.7.3.3 Selecting Integration Timestep

Selecting the integration timestep is an important consideration in MD simulations; the size of the timestep has a profound effect on the errors produced by the integration algorithms, where the total error is roughly correlated with the size of the timestep. If a small timestep is used, a better integration quality is obtained, but a simulation will only cover a limited region of the system’s phase space, while too large a timestep may lead to unstable simulations caused by high-energy interactions between particles. These scenarios are illustrated in Figure 2.8; clearly there is a balance between simulating the “correct” trajectory and adequate sampling of phase space.

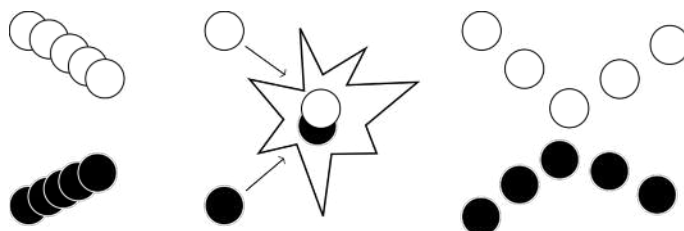


Figure 2.8: Illustration of the impact of timestep on MD simulation. Too small a timestep leads to limited coverage of phase space (left); Too large a timestep leads to unstable simulation (middle); An appropriate timestep allows a simulation to efficiently explore phase space (right).[3]

As an approximate guide, the timestep for simulating flexible molecules should be approximately an order of magnitude smaller than the fastest motion in the system. In flexible molecules, the highest frequency vibrations are usually C-H bond stretches, with a repeat period 10 fs. These bond stretches are often not interesting as they exert little influence on molecular properties or system behaviour; MD simulations therefore expend significant effort describing relatively unimportant bond

stretching motions.

One solution to this problem is to freeze out these high-frequency vibrations by constraining the bonds to their equilibrium bond length, while still allowing the remaining degrees of freedom in the system to vary. Eliminating these high-frequency vibrations allows a larger timestep to be used.

#### 2.7.3.4 SHAKE algorithm

The SHAKE algorithm[76] is the most common method used to constrain hydrogen coordinates in MD simulations. In SHAKE, the equations of motion are solved while simultaneously satisfying any imposed constraints. First, all atoms are allowed to move under the influence of the forces present, assuming an absence of constraint forces. Next, the deviation of each bond containing hydrogen from its reference bond length is used to calculate the constraint force that “corrects” the bond length. Once this correction has been applied, every relevant bond length is checked; if the deviation is greater than the specified tolerance, the correction calculation is repeated until all bond lengths satisfy the criteria. This means that SHAKE uses an iterative procedure to adjust hydrogen atom coordinates so that the deviation from the reference bond length falls within a given tolerance for all bonds containing hydrogen, as shown in Equation 2.54.

$$S_k = \frac{d_k(t)^2 - d_{k0}^2}{d_{k0}} < \varepsilon \quad (2.54)$$

Where  $d_k(t)$  is current bond length of interest,  $d_{k0}$  is the reference bond length and  $\varepsilon$  is the tolerance value, often  $10^{-4} - 10^{-8}$  Angstroms.

#### 2.7.3.5 Long-Range Forces

The most computationally expensive part of a MD simulation is the calculation of non-bonded energies and forces – while the number of bond, angle and torsion terms is proportional to the number of particles ( $N$ ) in a system, the number of non-bonded interactions increases as a factor of  $N^2$ . Ideally these non-bonded interactions would be calculated for all particles interacting with all others, but this becomes impractical for even moderately sized systems. In addition, calculating these interactions is often not justified. Non-bonded interactions are modelled using a Lennard-Jones potential, which accurately represents the  $r^{-6}$  distance dependence

of these interactions. As Figure 2.9 shows, the potential decays rapidly with distance, such that the value at  $2.5r$  falls to 1% of its value at  $r$ .

One practical way to handle the excessively large number of non-bonded interaction terms is to employ a cut-off in the potential. Here, the interaction energy between any two particles with a separation of greater than the cut-off is set to zero.

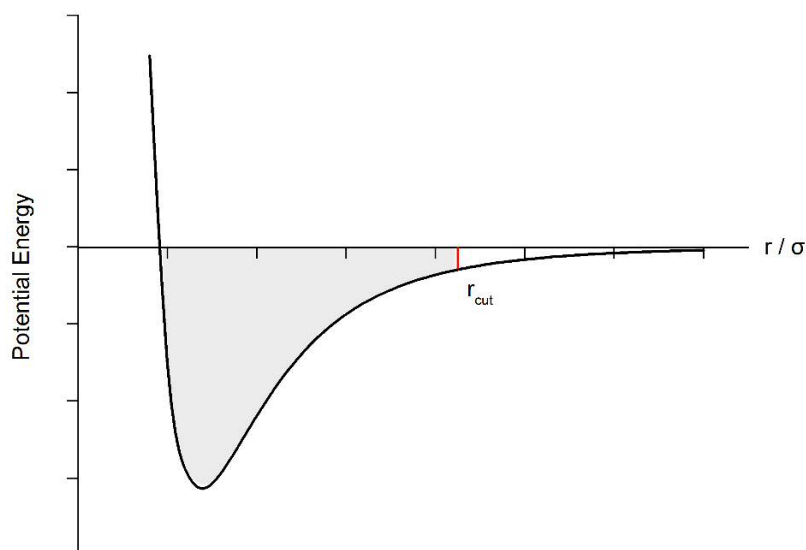


Figure 2.9: Illustration of a Lennard-Jones potential with a potential cutoff, as used in MD simulation.

Similarly, the charge-charge interaction - which decays as  $r^{-1}$  - is problematic in molecular dynamics simulations. In this case, a larger cut-off must be applied due to the slower decay of the interaction energy with distance.

The use of cut-offs for non-bonded and charge-charge interactions inevitably leads to errors in the simulation and deviations from the true potential, but suitably chosen cut-offs are able to minimise this.

### 2.7.3.6 Periodic Boundary Conditions

In MD simulations, it is common to employ boundary conditions, such that the microscopic system experiences forces and interactions as if it were in a macroscopic system. This may be achieved using periodic boundary conditions. Here, a three-dimensional cell (typically a cube) is replicated in all directions – a central cube is

surrounded by 26 identical cubes (known as images), which are in turn surrounded by 96 identical cubes and so on, giving a periodic array of cubes. The positions of the particles in these surrounding periodic images may be found by adding or subtracting integral multiples of the length of the side of the periodic cell.

If, during the simulation, a particle leaves the original cell, it is replaced by a particle from the image cell on the opposite side *i.e.* if a molecule exits through the right wall of the original cell, its image will enter from the left wall from the adjacent periodic cell, keeping the number of particles present in the cell constant.

Using periodic boundary conditions also affects the way the long-range forces mentioned above are handled. In the case of non-bonded interactions, the minimum image convention is required. Here, a given particle experiences only one instance of every other particle in the system and the interaction or force is calculated for the nearest atom or periodic image. This means that the applied cut-off should not allow for a given particle to interact with its own image or the same particle twice. In effect, this limits the cut-off distance to half the length of the cubic simulation cell.

Modelling charge-charge interactions in periodic systems is also challenging as the interaction extends beyond the length of a simulation cell. One common way to handle this is to use the Ewald summation method[77], where a given particle interacts with all other particles in the original simulation cell as well as all of their periodic images in the periodic array. The Ewald sum then considers the charge-charge interaction as a sum of a short range and long range interactions.

### 2.7.3.7 Molecular Dynamics Ensembles

MD simulations are typically performed in the NVE (micro-canonical) ensemble, where the system moves through phase space along a contour of constant energy while the temperature and pressure fluctuate. The total energy is a sum of the kinetic energy and the potential energy, calculated as shown in Equation 2.55.

$$E_{total} = \sum^N \frac{1}{2} m_i v_i^2 + V(r) \quad (2.55)$$

Where  $m$  is the mass of a particle,  $v$  is the velocity and  $r$  is the position.

However, there are several alternative ensembles that may be sampled, which provide results more relevant to experimental conditions. Simulating these other ensembles necessitates some way to constrain an instantaneous observable property, such as temperature or pressure. This may be achieved either by directly constraining the property to a specified macroscopic value at all timesteps or by allowing the instantaneous value to fluctuate while ensuring that the average remains equal to the desired macroscopic value. Here, an overview of these ensembles and the methods used to achieve them is covered.

### (i) Constant Temperature Simulations

In constant temperature simulations, we control the simulation cell volume and the system temperature – this is the canonical (NVT) ensemble. The temperature of this system is related to the average kinetic energy, as shown in Equation 2.56.

$$E_{<kinetic>} = \frac{1}{2}(3N_{atoms} - N_{constraint})k_B T \quad (2.56)$$

The simplest way to control the temperature of the system is to scale the velocities at every timestep, altering the kinetic energy. If the temperature at time  $t$  is  $T(t)$  and the desired temperature is  $T_0$ , the velocities may be scaled by a factor of  $(T_0/T(t))^{1/2}$ . [78]

Alternatively, the system may be coupled to an external heat bath that is fixed at the desired temperature and supplies or removes heat from the system as required. This is known as a thermostat. The kinetic energy of the system is controlled by scaling the particle velocities, as before, but the rate of change of temperature depends on the difference between the desired and actual temperature, as illustrated in Equation 2.57. The velocity scaling factor is then shown in Equation 2.58.

$$\frac{dT(t)}{dt} = \frac{1}{\tau}(T_{bath} - T(t)) \quad (2.57)$$

Where  $\tau$  is a coupling parameter which determines how strongly the heat bath is coupled to the system. If  $\tau$  is large, the coupling is weak; if  $\tau$  is small, the coupling is large.

$$\lambda = [1 + \frac{dt}{\tau}(\frac{T_0}{T(t)} - 1)]^{1/2} \quad (2.58)$$

However, both of these methods affect the dynamics such that simulations do not generate exact canonical ensembles. One final method for performing NVT dynamics are the extended system methods proposed by Nose and Hoover.[79, 80] Here, the thermal bath is considered as part of the system, represented by an additional degree of freedom,  $s$ . In this case, each state of the extended system that is generated corresponds to a unique state of the real system. The particle velocities in the real system may be calculated using Equation 2.59, while the timestep is given by Equation 2.60 – both depend on the degree of freedom of the extended system.

$$v_i = s \frac{dr_i}{dt} \quad (2.59)$$

$$\delta t = s \delta t' \quad (2.60)$$

The heat bath has kinetic and potential energy given by Equations 2.61 and 2.62, respectively. Here,  $f$  is the number of degrees of freedom (normally  $3N-3$ ),  $T$  is the desired temperature and  $Q$  is the thermal inertia parameter, which determines the coupling between the heat bath and the real system. If  $Q$  is large, the energy flow is slow; if  $Q$  is small, rapid energy flow leads to unstable, oscillating temperature in the real system.

$$E_{bath\,potential} = (f + 1)k_B \ln(s) \quad (2.61)$$

$$E_{bath\,kinetic} = (\frac{Q}{2})(\frac{ds}{dt})^2 \quad (2.62)$$

## (ii) Constant Pressure Simulations

In constant pressure simulations, we control the pressure and the temperature - this is the isothermal-isobaric (NPT) ensemble. In this case, the pressure is controlled by changing the volume of the simulation cell. However, this NPT method does not feature in the work included in this thesis, so discussion of NPT simulation methods is brief:

Many of the methods of controlling pressure are analogous to the temperature controls discussed above *i.e.* the pressure can be maintained by scaling the volume at all timesteps. Alternatively, the system may be coupled to a pressure bath in a similar fashion to the heat bath mentioned earlier. This is known as a barostat. In the extended system methods, the pressure bath is considered to be part of the system to be studied. An additional degree of freedom,  $V$ , is added, which corresponds to the volume of the simulation box. This may be thought of as a piston acting to compress the real system. In effect, pressure may be maintained by scaling particle positions at every timestep.

### 2.7.4 Comparisons of Conformational Search Methods

Clearly, the relative computational cost of different conformational search methods, and the efficiency with which they sample conformations, is an important factor in their selection for a given research problem. Below, several popular conformational search methods are compared.

According to work by Labute[67], the biggest problem with molecular dynamics for exploring conformations is that it must wait for kinetic energy to localise in certain low-frequency vibrational modes in order to cause large conformational changes. As the system increases in size, this wait increases, so the large changes in conformation are less common. In contrast, LMMD works to direct the kinetic energy to the important low-frequency modes, inducing conformational change much more rapidly. Recent work[67] illustrated that LowMode MD is far more efficient than traditional MD for exploring molecular conformations.

The most time-consuming aspect of LMMD is usually the number of force-field evaluations needed to perform the energy minimisation, though detection of duplicate conformers may also be significant, especially when modelling large, flexible molecules.[67] While LMMD is the most computationally demanding of the methods outlined here (per iteration), this increased expense is counterbalanced by the efficiency with which it identifies low energy conformations:

For a set of simple cyclic alkanes, a LMMD simulation and a stochastic conformational search (both performed in MOE[29]) find comparable minimum energy structures.[67] While the stochastic search is less computationally demanding, LMMD



finds many more low-energy conformations as the alkane ring size increases. This suggests that as the conformational space grows, the stochastic search method becomes increasingly less likely to generate a low-energy conformer.

Secondly, LMMD has been compared to twelve alternate conformational search methods employed by Bonnet[70] to study a range of macrocyclic peptides and sugars. Originally, Bonnet determined that a self-organising superimposition (SOS2) variant [81] of the stochastic proximity embedding (SPE) [82] conformational search was the optimum method for modelling these macrocyclic compounds, though LMMD was later proven to identify equivalent or lower global minimum energy conformations in almost all cases.[67] In addition, LMMD was able to find a greater number of unique low energy conformations, with a similar or greater degree of diversity (measured by the difference between the minimum and maximum values of the radius of gyration ( $R_{gyrmin.}$  and  $R_{gyrmax.}$ ) of the macrocycle of interest. Lastly, the LMMD search was also found to be approximately ten times more efficient than the SOS2 method preferred by Bonnet [70], where efficiency is calculated as the number of iterations performed per unique conformer generated. This clearly suggests that LMMD samples the low energy conformations of these systems more thoroughly than SOS2; by implication, LMMD is better than the alternate methods analysed by Bonnet, making it a leading option for the conformational analysis of complex molecules.

In addition, the performance of many conformational sampling protocols can be influenced by a range of tuneable settings, including the number of search cycles, the energy cut-off and the degree of similarity permitted between the retained conformers, as alluded to when discussing the LMMD process in Section 2.7.2. This fine-tuning of conformational search methods has been explored by Chen and Foloppe,[68] who found that adjusting the search protocol to the problem significantly improved the performance of the search method.

The aforementioned work studied a range of drug-like, flexible and macrocyclic compounds and assessed the relative performance of a selection of different conformational search methods and the effects of details of each protocol. These authors identified LMMD as the best search method for these complex molecules - LMMD produces larger conformational ensembles (*i.e.* greater conformational coverage), more likely to contain a representative of the bioactive structure, while locating the

global energy minimum more frequently than other available search methods.

Alternatively, research by Watts *et al.*[69] illustrated that the LMMD approach was outperformed by the macrocycle baseline search within MacroModel (MMBS) for a greater range of molecules than included in the research outlined above. Notably, Watts states that MMBS performs less well for macrocyclic rings containing more than 30 atoms; importantly for the context of this research, the macrocyclic ring formed by the binding of Pt(II) to histidine residues in A $\beta$  forms a ring 30+ atoms in size, so the MMBS approach would be less appropriate.

Clearly, the question of which conformational search method is the best is divisive and appears to depend on the nature of the molecule being studied. This means that several methods should be tested before selecting a particular modelling approach. However, LMMD has been shown to be a generally favourable method to thoroughly explore the conformational space of complex molecules. In addition, and perhaps most importantly, LMMD – unlike many of the alternative methods discussed above – allows for direct implementation of the LFMM procedure (Section 2.5.2) via DommiMOE.

## 2.8 Practical Considerations

### 2.8.1 Geometry Optimisation

Optimising the geometry of chemical systems plays an important role in much computational work and a discussion of the approaches to this is best framed in the context of the potential energy surface (PES) of the system.

The PES (see Figure 2.10) is a function of the nuclear coordinates; if there are  $N$  nuclei present, there are  $3N - 6$  dimensions to describe the nuclear arrangement. Visualisation of the PES therefore requires derivatives of the energy at points on the surface. Of particular interest for chemical systems are the stationary points on the surface (*i.e.* where the first derivative, or gradient, of the energy equals zero) – these correspond to energy minima, maxima and saddle points.

The goal of the optimisation process is usually to identify the global minimum

energy point, corresponding to the most stable conformation of the system of interest. Chemical systems may move between minima – provided energy is added – by moving over the surrounding energy barriers via saddle points, also known as transition states. In order to distinguish between minima and saddle points, where the first derivative is zero, the second derivative is calculated. Here, negative values correspond to the transition states and positive values correspond to energy minima.

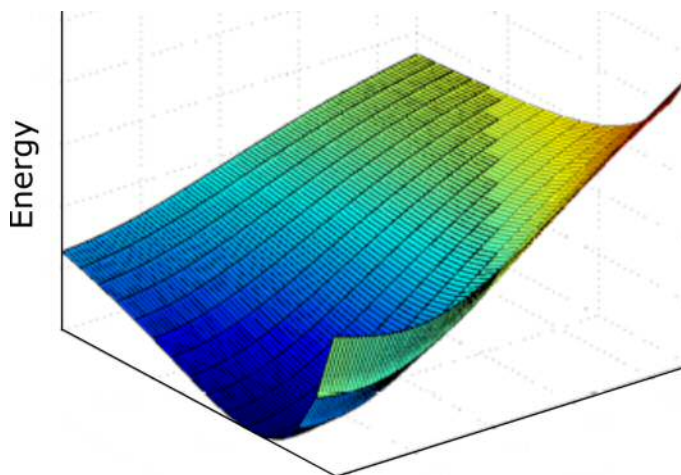


Figure 2.10: An example Potential Energy Surface, adapted from [83]

Modern computational chemistry packages contain algorithms for this optimisation procedure; one of the most common is the Newton-Raphson approach. In this approach, a function  $f(x)$  is constructed as a Taylor expansion around point  $x_0$ . The first derivative of  $f(x)$ ,  $f'(x)$ , may be written as shown in Equation 2.63.

$$f'(x) = x f'(x_0) + (x - x_0) f''(x_0) \quad (2.63)$$

If  $f(x)$  takes the form of a quadratic, the second derivative,  $f''(x)$ , is equal at all points *i.e.* the second derivative of  $x$  is equal to the second derivative of  $x_0$ . Thus at the minimum, where  $f'(x) = 0$ ,  $x$  may be calculated using Equation 2.64.

$$x = x_0 - \frac{f'(x_0)}{f''(x_0)} \quad (2.64)$$

If  $x_0$  is one solution to  $f'(x) = 0$ ,  $x$  calculated above is a new solution. The above equation is therefore used to provide an iterative solution. This Newton-Raphson method is relatively computationally demanding as each iteration requires calculation and inversion of the Hessian matrix; this lead to the development of quasi-Newton optimisation methods, which only require first derivatives of the energy and

construct the inverted Hessian matrix over successive iterations.

## 2.8.2 Solvent Models

Accurate simulations and modelling of chemical systems necessitates the inclusion of solvent effects. This is especially true for biomolecular systems. Typically, the solvent may be accounted for in one of two ways: explicitly or implicitly.

In explicit solvent calculations, each solvent molecule must be specified in the calculation input file. However, adequate solvation usually requires many hundreds or thousands of solvent molecules; this addition of extra molecules may render a calculation impractical.

Alternatively, in implicit solvent calculations, the solvent is described as a uniform polarisable medium with some dielectric constant,  $\epsilon$ . A cavity is created in the solvent medium into which the molecule or system of interest is placed. This is known as the self-consistent reaction field (SCRF) method. Here, interaction with the solvent perturbs the molecule, which in turn results in a perturbation of the reaction field. This may be accounted for in the Hamiltonian, displayed in Equation 2.65.

$$H_{total} = H_0 = \mu^T \frac{2(\epsilon - 1)}{2(\epsilon - 1)\alpha^3} < \Psi | \mu | \Psi > \quad (2.65)$$

Here,  $H_0$  is the original unperturbed Hamiltonian,  $\mu$  is the dipole moment and  $\alpha$  is the radius of the cavity. However, there is an energy penalty associated with creating the cavity in the solvent, written as  $\Delta G_{cavity}$ . This is mitigated by stabilising electrostatic and van der Waals interactions between the solute and solvent, written as  $\Delta G_{elec}$  and  $\Delta G_{disp}$  respectively. The total free energy change for solvation,  $\Delta G_{solvation}$ , may be expressed as Equation 2.66.

$$\Delta G_{solvation} = \Delta G_{cavity} + \Delta G_{elec} + \Delta G_{disp} \quad (2.66)$$

Three solvent models have been used during the course of this work. First, the polarisable continuum model (PCM)[84, 85] was used for DFT calculations. Here, the cavity is modelled using a van der Waals surface. Secondly, the COSMO [86] method,

which acts as a conductor-like screening method was used for semi-empirical calculations. In COSMO, the cavity is formed within a conductor with infinite dielectric constant. Small corrections to this conductor then provide close approximations of solvents with high dielectric constants, such as water. Lastly, the reaction-field solvation model was used for calculations in MOE and DL-POLY. Here, the solute is surrounded by a spherical cavity of finite radius. Inside the cavity, electrostatic interactions are calculated explicitly; outside, the system is treated as a dielectric continuum with a specified dielectric constant. A net dipole in the cavity then induces a polarisation in the dielectric, which in turn interacts with the given molecule.

### 2.8.3 Protein Secondary Structure Assignment

Peptide dihedral angles -  $\varphi$ ,  $\psi$  and  $\omega$  (see Figure 2.11) - determine peptide conformations and are responsible for essentially all variation in peptide backbone structure. By convention, backbone dihedral angles are defined by four atoms in the main peptide chain. When viewed along the central bond, the dihedral angle is defined as the angle between the end atom nearest the viewer and the end atom furthest away. The dihedral angle is positive if this angle is clockwise, and negative if anti-clockwise.

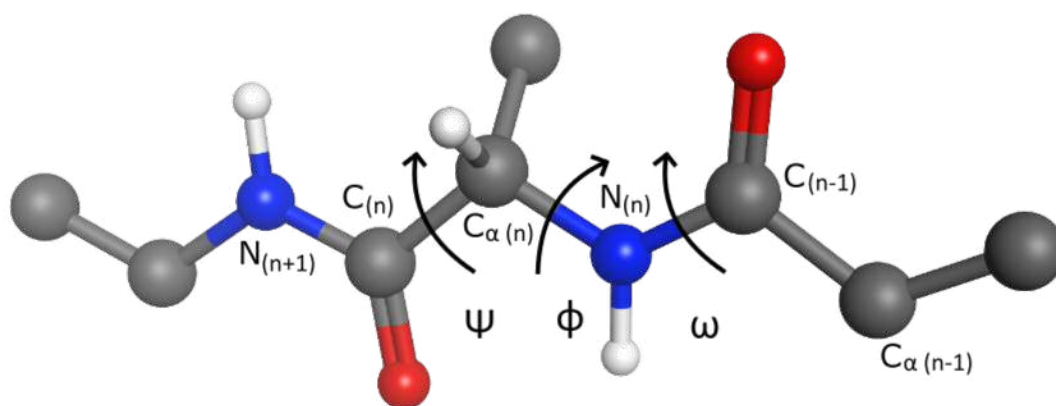


Figure 2.11: Illustration of  $\psi$ ,  $\varphi$  and  $\omega$  peptide backbone dihedral angles.

This means that  $\varphi$  dihedrals are defined by  $C_{\text{carbonyl}(n-1)}$ ,  $N_{(n)}$ ,  $C_{\alpha(n)}$  and  $C_{\text{carbonyl}(n)}$  atoms,  $\psi$  dihedrals by  $N_{(n)}$ ,  $C_{\alpha(n)}$ ,  $C_{\text{carbonyl}(n)}$  and  $N_{(n+1)}$  atoms and  $\omega$  dihedrals by  $C_{\alpha(n-1)}$ ,  $C_{\text{carbonyl}(n-1)}$ ,  $N_{(n)}$ , and  $C_{\alpha(n)}$  atoms. Of these, the  $\omega$  dihedral angle is the least important, since it is close to  $180^\circ$ , producing a trans, planar peptide where neighbouring  $\alpha$ -carbons and the peptide N, H, C, and O atoms lie in a single

plane.

The  $\varphi$  and  $\psi$  angles in a protein may be visualised and analysed using a Ramachandran plot - a graph of  $\varphi$  vs  $\psi$  angles, first proposed by Ramachandran *et al.* [87, 88] Ramachandran plots enable the user to identify energetically allowed regions of conformational space as well as plot distinct data points within a single structure.

In a Ramachandran plot (see Figure 2.12), there are regions of conformational space that correspond to the standard secondary structure types (*i.e.*  $\alpha$ -helix and  $\beta$ -sheet), allowing the user to predict secondary structure using only dihedral angle data.

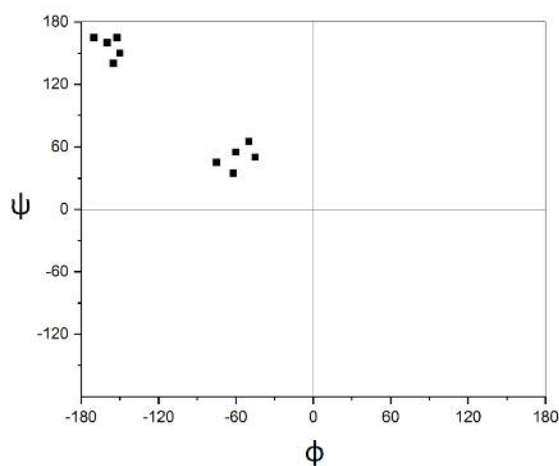


Figure 2.12: An example Ramachandran plot,  $\Phi$  vs.  $\Psi$ .

### 2.8.3.1 STRIDE

STRIDE (STRucture IDentification)[89] is an automatic algorithm for peptide secondary structure assignment based on backbone torsion angles and hydrogen bonding patterns, given only atomic coordinates and protein residue information. It allows for classification of secondary structure elements even if the structure does not fit the ideal secondary structure geometries of  $\alpha$ -helix or  $\beta$ -sheet.

Assignments of secondary structure are made based on a weighted product of hydrogen-bond energies and propensities of amino acids with specific  $\varphi$ ,  $\psi$  angles to form part of a particular secondary structure element. In STRIDE, the hydrogen bond energy depends on the distance of the interaction in question, as well as its angular geometry. This means that residues with non-standard backbone torsion angles may still

be accepted by STRIDE provided they form strong hydrogen bonds and *vice versa*. STRIDE was originally developed using large test set of protein structures from the PDB database, where secondary structure was determined by protein crystallography.

The Ramachandran plot is divided into  $20^\circ$  by  $20^\circ$  regions ( $i$ ). STRIDE then calculates the probability that backbone dihedral angles of residues assigned as  $\alpha$ -helix/ $\beta$ -sheet in the test set were found in the  $i^{th}$  region of the Ramachandran, illustrated in Equations 2.67-2.68.

$$P_i^\alpha = \frac{N_i^\alpha}{N_i^{total}} \quad (2.67)$$

$$\text{if } -180^\circ < \psi < 10^\circ \text{ and } -120^\circ < \phi < 45^\circ$$

$$P_i^\beta = \frac{N_i^\beta}{N_i^{total}} \quad (2.68)$$

$$\text{if } -180^\circ < \psi < 0^\circ \text{ and } -180^\circ < \phi < -120^\circ \text{ or } 45^\circ < \phi < 180^\circ$$

Where  $N_i^\alpha$  and  $N_i^\beta$  are the numbers of residues in the  $\phi, \psi$  zone assigned as either  $\alpha$ -helix or  $\beta$ -sheet, respectively, and  $N_i^{total}$  is the total number of residues with torsion angles in the  $i^{th}$  zone in the test set. These probability values are set equal to zero outside of the accepted  $\alpha$ -helical and  $\beta$ -sheet regions of the Ramachandran plot.

In STRIDE, a minimal  $\alpha$ -helix includes at least 2 consecutive hydrogen bonds between residues  $k$  and  $k + 4$ , with a hydrogen bond energy greater than a cutoff. A  $\beta$ -sheet is defined as two consecutive hydrogen bonded bridges – residues in the  $\beta$ -sheet must either participate in two hydrogen bonds (via backbone carbonyl and peptide hydrogen atoms) or be flanked by residues each forming a single hydrogen bond.

Characterisation of other secondary structure elements such as  $3_{10}$ -helices,  $\pi$ -helices and turns is complicated by their relatively irregular and sparse  $\phi, \psi$  data. In STRIDE,  $3_{10}$ -helices and  $\pi$ -helices were generally defined using the same rules as used in the DSSP secondary structure program by Kabsch and Sander.[90]

In addition, turns were defined according to work by Richardson [91], and further

refined by Wilmot and Thornton.[92] Structures are assigned as turns if the  $\alpha$ -carbon atoms in residues  $i$  and  $i + 4$  are less than 7 Å apart, with a hydrogen bond across from  $i - i + 4$ . Hydrogen bonds are assigned if the electronegative atoms (nitrogen and oxygen) are < 3.3 Å apart and the angles between the hydrogen atom, electronegative atom and the next main-chain atom are > 90°.

Each of the possible turn types are defined by the dihedral angles,  $\varphi$ ,  $\psi$ , in Table 2.4. STRIDE does not distinguish between each of these turn-types, instead generalising any qualifying structure as a turn; further classification is possible according to the values in Table 2.4.

Turn-type	Res. $i + 1$ $\varphi$ / °	Res. $i + 1$ $\psi$ / °	Res. $i + 2$ $\varphi$ / °	Res. $i + 2$ $\psi$ / °
<i>I</i>	-60	-30	-90	0
<i>I'</i>	60	30	90	0
<i>II</i>	-60	120	80	0
<i>II'</i>	60	-120	-80	0
<i>VI<sub>a</sub></i>	-60	120	-90	0
<i>VI<sub>b</sub></i>	-120	120	-60	0
<i>VIII</i>	-60	-30	-120	120

Table 2.4: Dihedral angles (°) of residues  $i + 1$  and  $i + 2$  in different turn-types, according to Wilmot and Thornton [92]

Many alternative algorithms and programs are available for assigning protein secondary structure, such as DSSP, but there is no consensus on the most accurate tool. A thorough examination of several popular methods may be found in work by Zhang.[93]

## 2.8.4 Programs and Methods

The vast majority of DFT calculations have been performed using Gaussian 09.[94] A small remainder were performed using NWCHEM[95] and these are noted in the text. Unless specified otherwise, the BHandH functional [9] has been used with the 6-31G(d) basis set in conjunction with the Stuttgart-Dresden (SDD) ECP [14, 15, 96] for transition metals. Where appropriate, the PCM solvent model [84, 85] has been used for aqueous solvation.



Semi-empirical calculations using the PM6, [55] PM6-DH2, [59, 61] and PM7[65] methods were carried out using MOPAC.[66] Where appropriate, the COSMO [86] solvent model has been used.

All LFMM calculations contained in Chapters 3 and 4 were performed within the DommiMOE [22] extension to MOE.[29] MD and LFMD calculations in Chapter 5 were performed using DL-POLY CLASSIC [97] and DL-POLY-LF, respectively.[30]

For simulations of the free peptide, DL-POLY input files (CONFIG, FIELD and CONTROL) were generated using DL-FIELD.[98] For simulations containing platinum, DL-POLY-LF input files (CONFIG, FIELD, CONTROL and lfse.in) were generated in DommiMOE using scripts written by the Deeth group.

For all LFMM data in Chapters 3 and 4, the original *AMBER94\_TM* [22] force-field was used alongside the reaction-field solvent model with default parameters. All MM calculations were carried out in MOE. LFMD calculations (Chapter 5) were performed using the corrected version of the *AMBER94\_TM* forcefield. Analysis of LFMD trajectories was performed using Visual Molecular Dynamics (VMD) [99], while secondary structure assignment was achieved using the STRIDE program. [89]

Geometrical data was obtained using either Molden [100] or Gaussview [101] visualisation software. Overlay images were obtained using Chimera[102] or MOE. The RMSD values for static structures contained within Chapters 3-5 were calculated using a python script obtained via GitHub [103, 104] and checked using Chimera. Finally, Ramachandran plots in Chapter 4 were obtained using JMP.[105]

## Bibliography

- (1) F. Jensen, *Introduction to Computational Chemistry*, Wiley, Chichester, UK, 2nd Ed., 2007.
- (2) C. J. Cramer, *Essentials of Computational Chemistry*, Wiley, Chichester, UK, 1st Ed., 2002.
- (3) A. R. Leach, *Molecular modelling : principles and applications*, Prentice Hall, Harlow ; New York, 2nd ed., 2001, xxiv+744; 16.
- (4) P. Hohenberg and W. Kohn, *Physical Review*, 1964, **136**, B864–B871.
- (5) W. Kohn and L. J. Sham, *Physical Review*, 1965, **140**, A1133–A1138.
- (6) A. D. Becke, *Physical Review A*, 1988, **38**, 3098–3100.
- (7) C. Lee, W. Yang and R. Parr, *Physical Review B*, 1988, **37**, 785–789.
- (8) Y. Juan and E. Kaxiras, *Physical Review. B, Condensed Matter*, 1993, **48**, 14944–14952.
- (9) A. D. Becke, *The Journal of Chemical Physics*, 1993, **98**, 5648–5652.
- (10) Y. Zhao and D. G. Truhlar, *Theoretical Chemistry Accounts*, 2008, **120**, 215–241.
- (11) S. Grimme, *Journal of Computational Chemistry*, 2006, **27**, 1787–1799.
- (12) S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *The Journal of Chemical Physics*, 2010, **132**, 154104.
- (13) H. Hellmann, *The Journal of Chemical Physics*, 1935, **3**, 61–61.
- (14) P. Fuentealba, H. Preuss, H. Stoll and L. Vonszentpaly, *Chemical Physics Letters*, 1982, **89**, 418–422.
- (15) D. Andrae, U. Haussermann, M. Dolg, H. Stoll and H. Preuss, *Theoretica Chimica Acta*, 1990, **77**, 123–141.
- (16) W. Cornell, P. Cieplak, C. Bayly, I. Gould, K. Merz, D. Ferguson, D. Spellmeyer, T. Fox, J. Caldwell and P. Kollman, *Journal of the American Chemical Society*, 1995, **117**, 5179–5197.
- (17) T. A. Halgren, *Journal of Computational Chemistry*, 1996, **17**, 490–519.
- (18) B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *Journal of Computational Chemistry*, 1983, **4**, 187–217.

- 
- (19) W. R. P. Scott, P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber, P. Krüger and W. F. van Gunsteren, *The Journal of Physical Chemistry A*, 1999, **103**, 3596–3607.
- (20) W. L. Jorgensen and J. Tirado-Rives, *Journal of the American Chemical Society*, 1988, **110**, 1657–1666.
- (21) A. Rappe, C. Casewit, K. Colwell, W. Goddard and W. Skiff, *Journal of the American Chemical Society*, 1992, **114**, 10024–10035.
- (22) R. J. Deeth, N. Fey and B. Williams–Hubbard, *Journal of Computational Chemistry*, 2005, **26**, 123–130.
- (23) R. J. Deeth, A. Anastasi, C. Diedrich and K. Randell, *Coordination Chemistry Reviews*, 2009, **253**, 795–816.
- (24) G. L. Meissler and D. A. Tarr, *Inorganic Chemistry, 2nd Edition (Miessler, Gary L.; Tarr, Donald A.) p362-366*, Pearson, New Jersey, 3rd edn., 2004.
- (25) R. J. Deeth and D. L. Foulis, *Physical Chemistry Chemical Physics*, 2002, **4**, 4292–4297.
- (26) <http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html>.
- (27) <http://www.daylight.com/dayhtml/doc/theory/theory.smiles.html>.
- (28) J. E. Bol, C. Buning, P. Comba, J. Reedijk and M. Ströhle, *Journal of Computational Chemistry*, 1998, **19**, 512–523.
- (29) *Molecular Operating Environment (MOE)*, (2013.08), version 2013.08, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2013.
- (30) H.-C. Tai, R. Brodbeck, J. Kasparkova, N. J. Farrer, V. Brabec, P. J. Sadler and R. J. Deeth, *Inorganic Chemistry*, 2012, **51**, 6830–6841.
- (31) D. Kepert, *Inorganic Stereochemistry*, Springer-Verlag, New York, 1st edn., 1982, vol. 6.
- (32) P. Comba, T. W. Hambley and B. Martin, *Molecular Modeling of Inorganic Compounds*, John Wiley & Sons, 2009, 347 pp.
- (33) D. M. Root, C. R. Landis and T. Cleveland, *Journal of the American Chemical Society*, 1993, **115**, 4201–4209.
- (34) T. Cleveland and C. R. Landis, *Journal of the American Chemical Society*, 1996, **118**, 6020–6030.
- (35) C. R. Landis, T. Cleveland and T. K. Firman, *Journal of the American Chemical Society*, 1998, **120**, 2641–2649.
- (36) T. K. Firman and C. R. Landis, *Journal of the American Chemical Society*, 2001, **123**, 11728–11742.
- (37) I. Tubert-Brohman, M. Schmid and M. Meuwly, *Journal of Chemical Theory and Computation*, 2009, **5**, 530–539.

- 
- (38) J. Y. Xiang and J. W. Ponder, *Journal of Computational Chemistry*, 2013, **34**, 739–749.
- (39) J. Y. Xiang and J. W. Ponder, *Journal of Chemical Theory and Computation*, 2014, **10**, 298–311.
- (40) A. E. Carlsson and S. Zapata, *Biophysical Journal*, 2001, **81**, 1–10.
- (41) J. W. Ponder and F. M. Richards, *Journal of Computational Chemistry*, 1987, **8**, 1016–1024.
- (42) M. Foscato, R. J. Deeth and V. R. Jensen, *Journal of Chemical Information and Modeling*, 2015, **55**, 1282–1290.
- (43) R. J. Deeth and L. J. A. Hearnshaw, *Dalton Transactions*, 2006, 1092–1100.
- (44) R. J. Deeth, *Chemical Communications*, 2006, **0**, 2551–2553.
- (45) R. J. Deeth, *Inorganic Chemistry*, 2007, **46**, 4492–4503.
- (46) R. J. Deeth and K. Randell, *Inorganic Chemistry*, 2008, **47**, 7377–7388.
- (47) A. E. Anastasi and R. J. Deeth, *Journal of Chemical Theory and Computation*, 2009, **5**, 2339–2352.
- (48) J. A. Pople, D. P. Santry and G. A. Segal, *The Journal of Chemical Physics*, 1965, **43**, S129–S135.
- (49) J. A. Pople and G. A. Segal, *The Journal of Chemical Physics*, 1965, **43**, S136–S151.
- (50) J. A. Pople and G. A. Segal, *The Journal of Chemical Physics*, 1966, **44**, 3289–3296.
- (51) D. P. Santry and G. A. Segal, *The Journal of Chemical Physics*, 1967, **47**, 158–174.
- (52) J. J. P. Stewart, *Journal of Computational Chemistry*, 1989, **10**, 209–220.
- (53) J. J. P. Stewart, *Journal of Molecular Modeling*, 2004, **10**, 155–164.
- (54) M. J. S. Dewar, E. G. Zoebisch, E. F. Healy and J. J. P. Stewart, *Journal of the American Chemical Society*, 1985, **107**, 3902–3909.
- (55) J. J. P. Stewart, *Journal of Molecular Modeling*, 2007, **13**, 1173–1213.
- (56) A. A. Voityuk and N. Rösch, *The Journal of Physical Chemistry A*, 2000, **104**, 4089–4094.
- (57) W. Thiel and A. A. Voityuk, *Theoretica chimica acta*, 1992, **81**, 391–404.
- (58) W. Thiel and A. A. Voityuk, *The Journal of Physical Chemistry*, 1996, **100**, 616–626.
- (59) J. Rezac, J. Fanfrlik, D. Salahub and P. Hobza, *Journal of Chemical Theory and Computation*, 2009, **5**, 1749–1760.
- (60) M. Korth, *Journal of Chemical Theory and Computation*, 2010, **6**, 3808–3816.

- 
- (61) M. Korth, M. Pitonak, J. Rezac and P. Hobza, *Journal of Chemical Theory and Computation*, 2010, **6**, 344–352.
- (62) J. Řezáč and P. Hobza, *Chemical Physics Letters*, 2011, **506**, 286–289.
- (63) J. Řezáč and P. Hobza, *Journal of Chemical Theory and Computation*, 2012, **8**, 141–151.
- (64) P. S. Brahmshatriya, P. Dobeš, J. Fanfrlik, J. Rezac, K. Paruch, A. Bronowska, M. Lepšík and P. Hobza, *Current Computer-Aided Drug Design*, 2013, **9**, 118–129.
- (65) J. J. P. Stewart, *Journal of Molecular Modeling*, 2013, **19**, 1–32.
- (66) J. J. Stewart, <http://openmopac.net/index.html>.
- (67) P. Labute, *Journal of Chemical Information and Modeling*, 2010, **50**, 792–800.
- (68) I.-J. Chen and N. Foloppe, *Bioorganic & Medicinal Chemistry*, 2013, **21**, 7898–7920.
- (69) K. S. Watts, P. Dalal, A. J. Tebben, D. L. Cheney and J. C. Shelley, *Journal of Chemical Information and Modeling*, 2014, **54**, 2680–2696.
- (70) P. Bonnet, D. K. Agrafiotis, F. Zhu and E. Martin, *Journal of Chemical Information and Modeling*, 2009, **49**, 2242–2259.
- (71) I. Kolossvary and W. C. Guida, *Journal of the American Chemical Society*, 1996, **118**, 5011–5019.
- (72) I. Kolossvary and G. M. Keseru, *Journal of Computational Chemistry*, 2001, **22**, 21–30.
- (73) L. Verlet, *Physical Review*, 1967, **159**, 98–103.
- (74) R. W. Hockney, *Methods in Computational Physics*, 1970, **9**, 135–211.
- (75) W. C. Swope, H. C. Andersen, P. H. Berens and K. R. Wilson, *The Journal of Chemical Physics*, 1982, **76**, 637–649.
- (76) J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *Journal of Computational Physics*, 1977, **23**, 327–341.
- (77) P. P. Ewald, *Annalen der Physik*, 1921, **369**, 253–287.
- (78) L. Woodcock, *Chemical Physics Letters*, 1971, **10**, 257–261.
- (79) S. Nosé, *The Journal of Chemical Physics*, 1984, **81**, 511–519.
- (80) W. G. Hoover, *Physical Review A*, 1985, **31**, 1695–1697.
- (81) F. Zhu and D. K. Agrafiotis, *Journal of Computational Chemistry*, 2007, **28**, 1234–1239.
- (82) H. Xu, S. Izrailev and D. K. Agrafiotis, *Journal of Chemical Information and Computer Sciences*, 2003, **43**, 1186–1191.
- (83) J. Scott and J. Vanderkooi, *Water*, 2010, **2**, 14–28.

- 
- (84) S. Miertus, E. Scrocco and J. Tomasi, *Chemical Physics*, 1981, **55**, 117–129.
- (85) R. Cammi and J. Tomasi, *Journal of Computational Chemistry*, 1995, **16**, 1449–1458.
- (86) A. Klamt and G. Schüürmann, *Journal of the Chemical Society, Perkin Transactions 2*, 1993, 799–805.
- (87) G. N. Ramachandran, C. Ramakrishnan and V. Sasisekharan, *Journal of Molecular Biology*, 1963, **7**, 95–99.
- (88) G. N. Ramachandran and V. Sasisekharan, *Advances in Protein Chemistry*, 1968, **23**, 283–438.
- (89) D. Frishman and P. Argos, *Proteins-Structure Function and Genetics*, 1995, **23**, 566–579.
- (90) W. Kabsch and C. Sander, *Biopolymers*, 1983, **22**, 2577–2637.
- (91) J. S. Richardson, *Advances in Protein Chemistry*, 1981, **34**, 167–339.
- (92) C. M. Wilmot and J. M. Thornton, *Protein Engineering*, 1990, **3**, 479–493.
- (93) Y. Zhang and C. Sagui, *Journal of Molecular Graphics and Modelling*, 2015, **55**, 72–84.
- (94) M. Frisch and G. Trucks et al., *Gaussian09*, Wallingford, CT, 2009.
- (95) M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus and W. A. de Jong, *Computer Physics Communications*, 2010, **181**, 1477–1489.
- (96) D. Andrae, U. Häußermann, M. Dolg, H. Stoll and H. Preuß, *Theoretica chimica acta*, 1991, **78**, 247–266.
- (97) W. Smith, C. W. Yong and P. M. Rodger, *Molecular Simulation*, 2002, **28**, 385–471.
- (98) C. W. Yong, *Journal of Chemical Information and Modeling*, 2016, **56**, 1405–1409.
- (99) W. Humphrey, A. Dalke and K. Schulten, *Journal of Molecular Graphics*, 1996, **14**, 33–38, 27–28.
- (100) G. Schaftenaar and J. Noordik, *J. Comput.-Aided Mol. Design*, 2000, **14**, 123–124.
- (101) R. Dennington, T. Keith and J. Millam, *Gaussview*, version 5, Shawnee Mission KS, 2009.
- (102) E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *Journal of Computational Chemistry*, 2004, **25**, 1605–1612.
- (103) J. Charnley, *GitHub: Calculate RMSD for two XYZ structures*, <https://github.com/charnley/rmsd>.

- (104) W. Kabsch, *Acta Crystallographica Section A*, 1976, **32**, 922–923.
- (105) *JMP®*, Version 12.0.1 SAS Institute Inc., Cary, NC, 1989-2007.

## 3 Modelling Pt<sup>II</sup>-Amyloid- $\beta$ Complexes

### 3.1 Introduction

The work in this chapter aims to test the suitability of various modelling approaches, including DFT and semi-empirical methods, for calculating structures and energies of platinum complexes bound to fragments of the A $\beta$  peptide. Small model systems have been studied to validate these methods before studying different platinum-A $\beta$  binding modes as well as exploring conformations of the N-terminal region of the peptide.

Research contained in this chapter comes from work by our group, published in reference [1]. All of the work contained here from this publication was carried out by myself.

### 3.2 Modelling Pt<sup>II</sup> Crystal Structures

Pt<sup>II</sup> complexes are known to interact with biomolecules, particularly via nitrogen or sulphur donor atoms, such as histidine residues or DNA bases (see Chapter 1). Recently, Pt<sup>II</sup> complexes have been found to bind to the A $\beta$  peptide and inhibit its aggregation, thought to be a key part of the development of Alzheimer’s disease (see Section 1.4.2). Before work concerning the binding of platinum to large A $\beta$  peptides could begin, it was important to conduct preliminary experiments on small model systems using different modelling techniques in order to determine which computational methods produced the most reliable results, with the aim of identifying which method and basis set reproduced existing experimental data most accurately.

The choice of density functionals and basis set for this work was investigated by calculating the optimised, gas-phase geometries for several high-resolution crystal



structures of simple  $\text{Pt}^{\text{II}}$  complexes containing histidine or methyl-imidazole groups - as models of the key histidine residues in  $\text{A}\beta$  (see Chapter 1) - from the Cambridge Structural Database. Cartesian coordinates were obtained from the original references for each system and modelling was performed in Gaussian09[2] using a range of DFT methods: BHandH[3], B97D[4], B3LYP[3, 5, 6], PBE1PBE[7, 8],  $\omega$ B97xD[9] and LC- $\omega$ PBE[10]. In addition, each of these functionals was tested using thirteen different basis sets, given below, leading to 78 combinations of functional and basis set.

- |                 |                    |                    |
|-----------------|--------------------|--------------------|
| • lan2dz        | • SDD/6-31G(dp)    | • SDD/6-31+G(2dfp) |
| • cep-121G      | • SDD/6-31+G(d)    | • SDD/cc-pVDZ      |
| • SDD           | • SDD/6-31+G(dp)   | • SDD/TZV          |
| • SDD/6-31G(d)  | • SDD/6-31+G(2dp)  |                    |
| • SDD/6-311G(d) | • SDD/6-31++G(2dp) |                    |

The selected crystal structures are displayed in Figure 3.1.

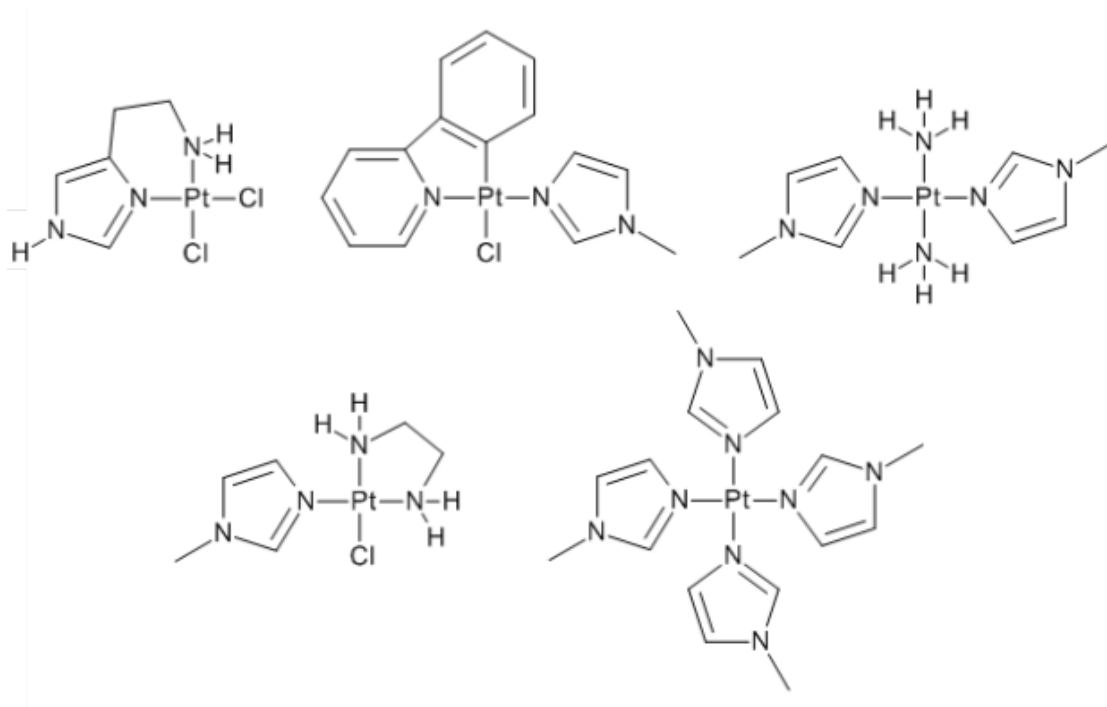


Figure 3.1: Complexes used for DFT methods validation, where structures are taken from X-ray crystallographic data. Top row, left-right: (Dichloro-(histamine-N,N'))-platinum(II) CSD ID: XUKWUV [11], chloro-(1-methyl-1H-imidazole)-(2-(pyridine-2-yl)phenyl)-platinum(II) CSD ID: OXUYUC [12], trans-Diamminebis (N-methylimidazole) platinum(II) chloride CSD ID: DAMIPT[13]. Bottom row, left-right: Chloro (ethylenediamine) (1-methylimidazole)platinum(II) nitrate CSD ID: HEK-JEM[14] and Pt(N-methylimidazole)<sub>4</sub> CSD ID: MIMDPT[15].

For much of this early work, we are primarily interested in the bond distances and angles of platinum and its coordinating ligands, as these are the most relevant to the binding of a platinum complex to A $\beta$ . Selected geometric data for the five crystal structures are shown in Tables 3.1 - 3.5, using the SDD/6-31G(d) basis set *i.e.* using a SDD ECP[16][17] for Pt<sup>II</sup> and the 6-31G(d) basis for all other atoms.

The calculated geometries of these structures show that the bond distances and angles show excellent agreement with the experimental (crystal) values. However, BHandH clearly outperforms the other DFT methods for these complexes. In particular, the platinum bond distances and angles agree to within 0.028 Å or better and 3.2° or better, respectively. For comparison, the best non-BHandH approach used the LC- $\omega$ PBE functional and the SDD/6-31++G(2dp) basis, where values agreed to 0.044 Å or better and 3.4° or better (see Table A.2). This is in contrast to work from Malik and Michalska[18], who found that the LC- $\omega$ PBE functional performed

best for cis-platin type complexes, though BHandH was not studied during that work.

CSD I.D.	Bond (Å)/Angle (°)	Expt. Value	BHandH	B97D	B3LYP	PBE1PBE	wB97xD	LC-wPBE
XUKWUV	Pt-N(ring)	1.994	2.013	2.064	2.063	2.035	2.054	2.036
	Pt-NH <sub>2</sub>	2.046	2.050	2.135	2.111	2.082	2.099	2.074
	Pt-Cl(1)	2.299	2.289	2.343	2.342	2.314	2.325	2.303
	Pt-Cl(2)	2.309	2.292	2.342	2.343	2.315	2.326	2.303
	Cl-Pt-Cl	91.7	93.6	93.2	93.5	93.3	93.6	93.4
	N(ring)-Pt-Cl(2)	90.2	89.9	90.7	90.4	90.2	90.2	90.1
	NH <sub>2</sub> -Pt-Cl(2)	86.9	83.7	84.0	83.8	83.7	84.0	83.9

Table 3.1: Selected geometric data for crystal structure XUKWUV, where data was obtained by experimental and various DFT methods using the SDD/6-31G(d) basis.

CSD I.D.	Bond (Å)/Angle (°)	Expt. Value	BHandH	B97D	B3LYP	PBE1PBE	wB97xD	LC-wPBE
OXUYUC	Pt-N(His)	2.026	2.008	2.053	2.057	2.032	2.045	2.033
	Pt-N(ppy)	2.017	2.012	2.055	2.051	2.030	2.046	2.033
	Pt-C(ppy)	1.978	1.975	2.003	2.007	1.987	1.992	1.985
	Pt-Cl	2.411	2.396	2.476	2.470	2.432	2.444	2.411
	C(ppy)-Pt-N(ppy)	81.3	91.1	80.9	80.8	81.0	80.9	91.0
	N(His)-Pt-C(ppy)	94.8	96.4	96.0	97.1	96.7	96.3	96.4
	N(His)-Pt-Cl	87.5	87.4	87.3	86.9	86.9	87.2	87.1
	N(ppy)-Pt-Cl	96.6	95.2	95.9	95.4	95.5	95.7	95.7

Table 3.2: Selected geometric data for crystal structure OXUYUC, where data was obtained by experimental and various DFT methods using the SDD/6-31G(d) basis.

CSD I.D.	Bond (Å)/Angle (°)	Expt. Value	BHandH	B97D	B3LYP	PBE1PBE	wB97xD	LC-wPBE
DAMIPT	Pt-N(imid)(1)	1.998	2.011	2.057	2.059	2.035	2.045	2.036
	Pt-N(imid)(2)	1.998	2.012	2.059	2.061	2.037	2.046	2.035
	Pt-NH <sub>3</sub> (1)	2.067	2.040	2.106	2.093	2.068	2.082	2.065
	Pt-NH <sub>3</sub> (2)	2.067	2.040	2.107	2.093	2.068	2.082	2.065
	NH <sub>3</sub> (1)-Pt-N(imid)(1)	90.4	89.9	90.0	90.0	90.0	89.9	89.9
	NH <sub>3</sub> (1)-Pt-N(imid)(1)	89.6	90.2	90.1	90.1	90.2	90.2	90.2

Table 3.3: Selected geometric data for crystal structure DAMIPT, where data was obtained by experimental and various DFT methods using the SDD/6-31G(d) basis.

CSD I.D.	Bond (Å)/Angle (°)	Expt. Value	BHandH	B97D	B3LYP	PBE1PBE	wB97xD	LC-wPBE
HEKJEM	Pt-N(imid)	2.026	2.002	2.040	2.046	2.023	2.033	2.024
	Pt-NH <sub>2</sub> (1)	2.038	2.036	2.103	2.086	2.062	2.078	2.057
	Pt-NH <sub>2</sub> (2)	2.040	2.062	2.144	2.125	2.094	2.108	2.085
	NH <sub>2</sub> (2)-Pt-N(imid)	93.2	95.8	95.9	96.7	96.4	96.0	96.2
	NH <sub>2</sub> (1)-Pt-NH <sub>2</sub> (2)	83.4	83.3	82.4	82.6	83.0	82.7	83.2
	Cl-Pt-N(imid)	91.8	91.4	91.4	90.8	90.7	91.0	90.8

Table 3.4: Selected geometric data for crystal structure HEKJEM, where data was obtained by experimental and various DFT methods using the SDD/6-31G(d) basis.

CSD I.D.	Bond (Å)/Angle (°)	Expt. Value	BHandH	B97D	B3LYP	PBE1PBE	wB97xD	LC-wPBE
MIMDPT	Pt-N(1)	1.985	2.013	2.053	2.057	2.034	2.043	2.034
	Pt-N(2)	2.012	2.009	2.061	2.059	2.039	2.049	2.037
	Pt-N(3)	1.985	2.013	2.053	2.057	2.034	2.043	2.034
	Pt-N(4)	2.012	2.009	2.061	2.059	2.039	2.049	2.037
	N(1)-Pt-N(2)	92.6	89.7	90.2	90.0	90.2	90.2	90.2
	N(1)-Pt-N(4)	87.4	90.3	89.8	90.0	89.8	89.8	89.8

Table 3.5: Selected geometric data for crystal structure MIMDPT, where data was obtained by experimental and various DFT methods using the SDD/6-31G(d) basis.

Data for additional basis sets may be found in the Appendix (Tables A.1 and A.2). In general, these data show that BHandH is the optimum functional for complexes of this type. In addition, these data suggest that for the basis sets studied here, the choice of basis has relatively little influence on the results.

### 3.3 Molecular Mechanics for Pt<sup>II</sup> Complexes

While DFT calculations are in principle able to model any system, and are shown to accurately model Pt<sup>II</sup> crystal structures (Section 3.2), their high computational expense means that they are impractical for the study of large, flexible peptides such as A $\beta$ . For this, forcefield-based methods are required, though these are typically not parameterised for transition metal elements (see Section 2.5.1).

#### 3.3.1 Forcefield Validation

The deciding factor in the accuracy of molecular mechanics calculations is the choice of forcefield. As discussed in Section 2.5, different forcefields are parameterised to best model particular classes of molecule - the most common family of forcefields for modelling proteins is AMBER.

The performance of the AMBER forcefield - explicitly designed for peptide and DNA molecules - was also investigated for the small organic ligands that are bound to the Pt<sup>II</sup> centre in these potential therapeutic agents. Here, for the sake of brevity, only data for the phenanthroline ligand is included. DFT calculations were carried out in G09 using the BHandH functional with a 6-31G(d) basis set for all atoms. Solvation was handled using the polarised continuum model.[19, 20] AMBER calculations were performed in MOE[21], using AMBER94[22] parameters for all atoms and the reaction-field solvation model. Resulting bond distances and angles are shown in Table 3.6, where atom numbering corresponds to Figure 3.2.



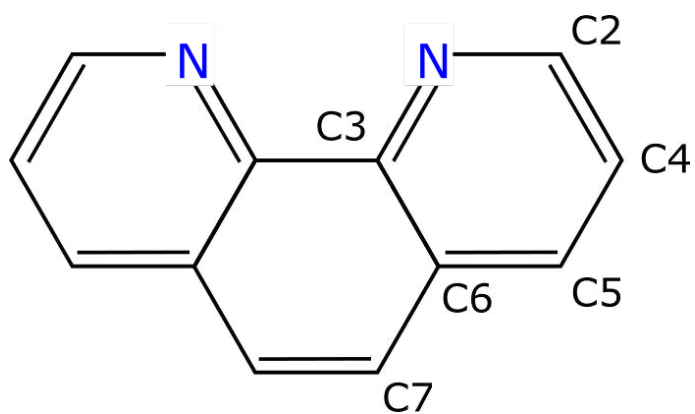


Figure 3.2: Atom numbering for the phenanthroline ligand, used for forcefield validation.

Bond	Ave. DFT Bond Length / Å	Ave. AMBER Bond Length / Å	$\Delta$ AMBER-DFT Length / Å
N-C2	1.305	1.350	0.045
N-C3	1.335	1.363	0.028
C3-C3	1.441	1.426	-0.015
C2-C4	1.395	1.406	0.011
C4-C5	1.363	1.405	0.043
C5-C6	1.397	1.407	0.011
C6-C7	1.421	1.406	-0.015
C7-C7	1.344	1.403	0.060
C2-H2	1.085	1.084	-0.001
C4-H4	1.081	1.082	0.001
C5-H5	1.083	1.082	-0.001
C7-H7	1.083	1.083	-0.000
MUE			0.019
Angle	Ave. DFT Bond Angle / °	Ave. AMBER Bond Angle / °	$\Delta$ AMBER-DFT Angle / °
C3-N-C2	118.46	122.37	3.91
N-C2-C4	123.90	120.23	-3.67
C2-C4-C5	118.18	119.77	1.59
C4-C5-C6	119.39	118.66	-0.74
C5-C6-C3	117.77	119.89	2.12
C6-C3-N	122.30	119.09	-3.21
C6-C7-C7	120.69	120.07	-0.62
C7-C6-C3	120.48	120.71	0.24
C6-C3-C3	118.84	119.22	0.39
MUE			1.83

Table 3.6: Geometric data for the phenanthroline ligand, calculated using BHandH/SDD/6-31G(d) and AMBER94.

This data shows that AMBER models this small organic ligand well; all bond distances agree with DFT calculations to within 0.06 Å and all internal angles of the ligand agree to within 4°. This accuracy in performance illustrates that there is little problem in using the AMBER forcefield to handle molecules of this type, such as other small aromatic ligands will be investigated later.

### 3.3.2 Pt<sup>II</sup> Complexes

As we have seen, molecular mechanics methods are generally not able to model transition metal complexes without modification. Modelling the Pt<sup>II</sup> complexes in this work requires AMBER parameters for the region of the peptide that is coordinating to platinum, as well as platinum itself. Fortunately, this is not a new problem.

Approximate AMBER parameters were obtained from work by Scheeff[23] - originally used for modelling cisplatin-DNA adducts - and then implemented into the forcefield files. In order for the central platinum to remain square planar, several new atom types were defined in the forcefield and identified using SMILES[24] and SMARTS.[25] These atom types - the histidine nitrogen atoms coordinating to Pt<sup>II</sup> - were given identical properties to pre-existing atom types, but given type names which allowed distinction between those in different residues. These new AMBER types - NX, NZ and NA1 (see Figure 3.3) - were necessary to control the platinum geometry using proper torsion potential energy terms. Proper torsions were used because the more convenient out-of-plane and improper torsion potential energy terms are not present in the construction of the AMBER forcefield. Without these additional torsion terms, the Platinum<sup>II</sup> centre adopted a tetrahedral geometry.

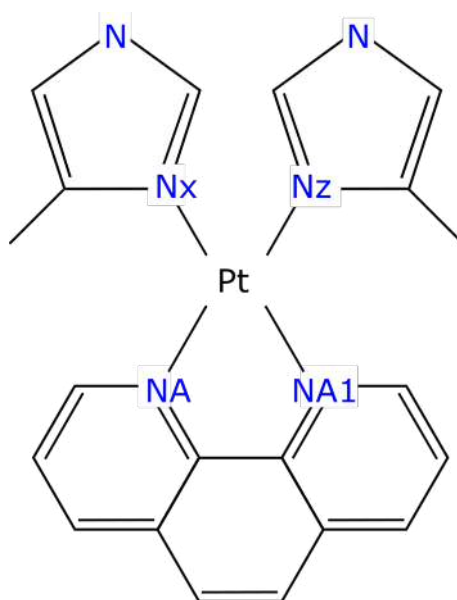


Figure 3.3: Illustration of AMBER types for Pt<sup>II</sup> and surrounding ligands, based on work by Scheeff [23].

However, this approach requires validation of the literature parameters as well as accurate parameterisation of these new torsion terms, achieved by fitting to exper-

imental or high-level computational data. For this reason, this pure-MM approach was not utilised beyond this point.

Alternatively, Ligand Field Molecular Mechanics (LFMM)[26–28] is a powerful tool for modelling transition metal complexes (see Section 2.5.2). As the AMBER forcefield is the most appropriate for modelling peptide systems and is also capable of accurately modelling the aromatic ligands involved in this work, DommiMOE[27] was configured to use this forcefield type. While recent iterations of the AMBER family of forcefields are typically available (*e.g.* AMBER99 and those using Extended Huckel Theory), the choice of forcefield in this work is dictated by what is available in the DommiMOE program. At this stage, AMBER94 is the only iteration available. Importantly, however, there is very little difference in the parameterisation of the newer versions compared to AMBER94 for the peptide atoms (*i.e.* carbon, nitrogen, oxygen, sulphur and hydrogen) and the bonds between them, so this choice will have little effect on the generation and modelling of conformations of these systems.

Even though LFMM is an established technique for complexes of the type studied here, it is important to validate the parameters and check that LFMM produces accurate structures of  $\text{Pt}^{\text{II}}$  complexes.

To this end, LFMM optimisations were performed for appropriate small model systems from Section 3.2 and geometries compared to experimental and calculated data. Unfortunately, crystal structures XUKWUV, OXUYUC, DAMIPT and HEKJEM (see Figure 3.1) are not suitable for this purpose, as there are no forcefield parameters for some of the atom types *e.g.* coordinated  $\text{NH}_2\text{R}$  and Cl groups and no corresponding LFMM parameters. In lieu of additional crystal structures, a series of complexes of the type  $\text{Pt}^{\text{II}}(\text{ligand})(\text{Me-imidazole})_2$  were constructed and optimised using DFT (BHandH/SDD/6-31G(d)) in Gaussian09, where the ligand may be bipyridine, phenanthroline, diphenyl-phenanthroline or dipyridophenazine (DPPZ), as shown in Figure 3.4. Each of these additional ligand systems was modelled with all combinations of Me-imidazole binding via  $\text{N}\delta$  and  $\text{N}\epsilon$  atoms, giving twelve combinations of ligand and Me-imidazole binding mode.

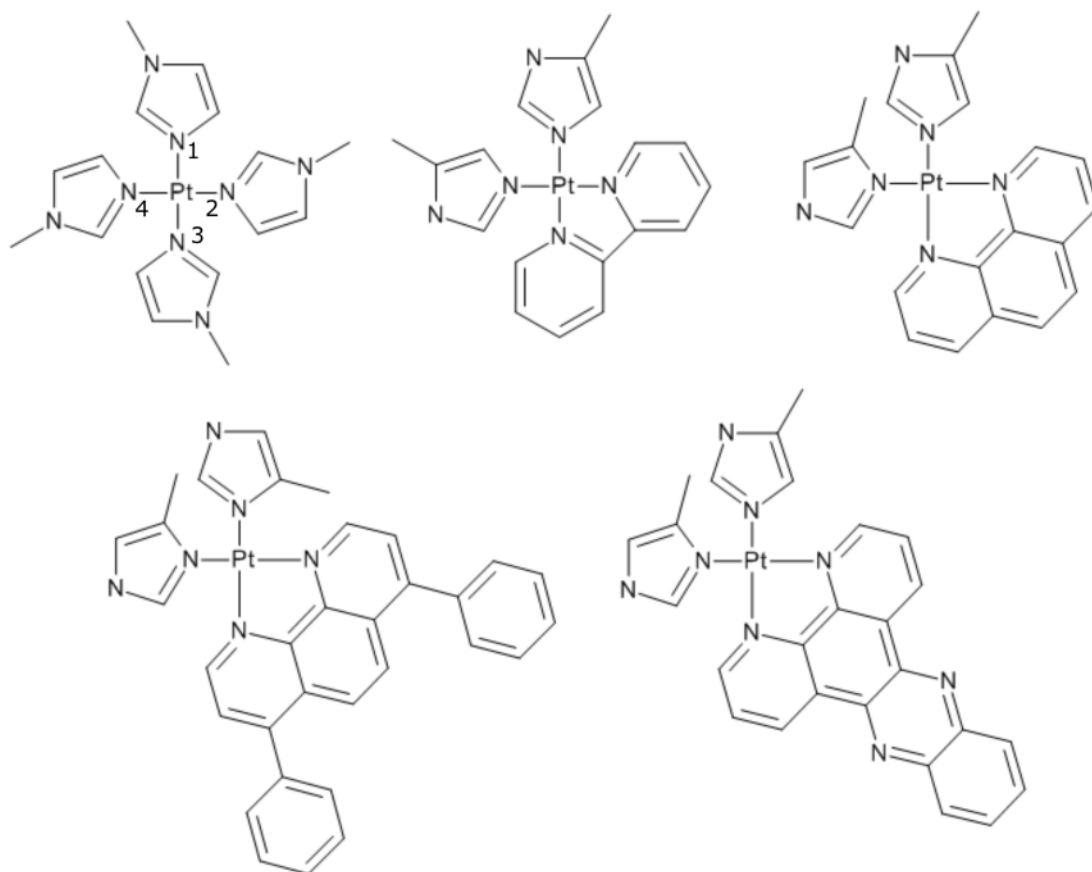


Figure 3.4: Example small model systems used for validation of LFMM parameters. Top row: Crystal structure MIMDPT[15],  $\text{Pt}^{\text{II}}(\text{bipy})(\epsilon\epsilon\text{-Me-imidazole})_2$  and  $\text{Pt}^{\text{II}}(\text{phen})(\delta\delta\text{-Me-imidazole})_2$ . Bottom row:  $\text{Pt}^{\text{II}}(\text{Dpphen})(\delta\delta\text{-Me-imidazole})_2$  and  $\text{Pt}^{\text{II}}(\text{Dppz})(\delta\epsilon\text{-Me-imidazole})_2$ .

CSD I.D.	Bond (Å)/Angle (°)	Expt.	BHandH	LFMM
MIMDPT	Pt-N(1)	1.9585	2.013	1.971
	Pt-N(2)	2.0119	2.009	1.970
	Pt-N(3)	1.9585	2.013	1.971
	Pt-N(4)	2.0119	2.009	1.970
	N(1)-Pt-N(2)	92.6	89.7	90.5
	N(1)-Pt-N(4)	87.4	90.3	89.8
MUE (Bond)			0.026	0.015
MUE (Angle)			0.0	0.2

Table 3.7: Selected geometric data for crystal structure MIMDPT, where data was obtained by experimental, DFT and LFMM methods.

Bipy N $\delta$ N $\delta$	Bond ( $\text{\AA}$ )/Angle ( $^\circ$ )	BHandH	LFMM
	Ave. Pt-N <sub>Lig</sub>	2.011	1.960
	Ave. Pt-N <sub>Me-imid</sub>	2.005	1.952
	N <sub>Me-imid</sub> -N <sub>Me-imid</sub>	87.9	89.3
	N <sub>Lig</sub> -N <sub>Lig</sub>	80.5	92.7
	N <sub>Me-imid</sub> -N <sub>Lig</sub>	95.8	89.5
MUE (Bond)			0.052
MUE (Angle)			2.4

Table 3.8: Selected geometric data for the small model system  $\text{Pt}^{\text{II}}(\text{bipy})(\text{Me-imid})_2$  where Me-imid is bound by N $\delta$  atoms, as noted in the top left. Data obtained by DFT and LFMM methods.

For the sake of brevity, only one of the twelve data tables for the  $\text{Pt}^{\text{II}}(\text{ligand})(\text{Me-imidazole})_2$  complexes are placed here - the remainder may be found in the Appendix (Tables A.3 - A.13). The LFMM-optimised geometries of these structures show good agreement with experimental data and those calculated using BHandH. Here,  $\text{Pt}^{\text{II}}$ -containing bond lengths and angles agree with experiment to within 0.042  $\text{\AA}$  and  $2.4^\circ$ , respectively. In addition, platinum bond lengths agree with DFT data to within 0.060  $\text{\AA}$  (largest deviation 0.058  $\text{\AA}$ ). Coordination angles fare less well, but are still reasonable - these agree to within  $13.7^\circ$  of DFT results. This relatively large deviation from calculated DFT values appears only for the angle between N<sub>lig</sub>-Pt-N<sub>lig</sub>: DFT predicts that this angle is compressed, with values 80.4-81.2 $^\circ$  observed, while LFMM predicts values 92.2-94.5 $^\circ$ . For all other coordination angles, LFMM data agree with DFT to within  $7.3^\circ$ .

This data illustrates that LFMM - with the current parameters - is able to reproduce experimental and DFT data in both gas and condensed phases to a good degree of accuracy, though in general, the LFMM geometries are too rigidly square planar, with less variation in coordination geometry than predicted by DFT. In principle, it is possible to improve the fit of the LFMM data to the crystal and calculated data, but this is at the risk of decreasing the transferability of the parameters.

### 3.4 Modelling the Metal-Binding Region of A $\beta$

In this section, a protocol for modelling Pt<sup>II</sup>(ligand) complexes bound to the A $\beta$ 6-14 fragment of the A $\beta$  peptide is established and is validated for the two most simple Pt<sup>II</sup>(ligand) complexes examined in this work.

The peptide sequence His6-Asp-Ser-Gly-Tyr-Glu-Val-His14, taken from the N-terminal domain of A $\beta$ , was built in an extended conformation in MOE.[21] Pt was added in eight distinct coordination modes (*i.e.* to His6 and either His13 or His14 through N $\delta$  or N $\epsilon$ ), and the bipyridine or phenanthroline ligand manually constructed, as shown in Figure 3.5 and Table 3.9.

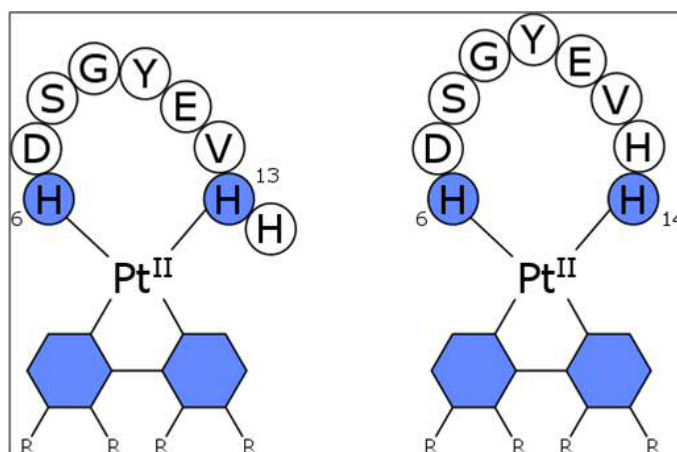


Figure 3.5: Schematic of Pt<sup>II</sup>(ligand) complexes binding to His6 and His13 (left) and His6 and His14 (right). The ligand, represented by the blue scaffold, may be either bipyridine or phenanthroline.

Binding Mode	His-6	His-13	His-14
1	$N\delta$	$N\delta$	-
2	$N\delta$	$N\varepsilon$	-
3	$N\varepsilon$	$N\delta$	-
4	$N\varepsilon$	$N\varepsilon$	-
5	$N\delta$	-	$N\delta$
6	$N\delta$	-	$N\varepsilon$
7	$N\varepsilon$	-	$N\delta$
8	$N\varepsilon$	-	$N\varepsilon$

Table 3.9: Possible binding modes of platinum complexes to the histidine residues of  $A\beta$

### 3.4.1 Charge Schemes

An important factor in molecular mechanics-type approaches is the method used to determine atomic partial charges. Several possible methods for assigning charges were explored in this research, as detailed below.

Firstly, it is possible to use the AMBER forcefield partial charges for all atoms, including the  $\text{Pt}^{\text{II}}(\text{ligand})$  complex. In DommiMOE, this calculation is instantaneous, so forcefield charges are very easy to apply. However, using a forcefield designed for organic structures to calculate the partial charge of a transition metal ion and its coordinating ligands leads to peculiar atomic charges around the platinum centre *e.g.* 4-valent nitrogen atoms are given negative atomic partial charges. This is likely due to the fact that the MOE software attempts to automate parameters and charges for unknown structures - in the case of transition metal systems, these parameters are often poor. Despite its ease of use, this method was not used beyond initial tests in this research.

It is important that the method used to define partial charges is consistent with the underlying AMBER forcefield approach. To this end, calculation of Merz-Kollman partial charges for the entire system (approx. 100-200 atoms) was performed by calculating the electrostatic potential at the HF/6-31G(d)/SDD level in Gaussian09, with Pt given a van der Waals radius of 2.0 Å *i.e.* calculate the atomic partial charges that would reproduce the electrostatic potential (ESP) of the molecular



surface. This relatively fast calculation must be done outside of DommiMOE, but the charges are input manually. For systems of this size, this is a very slow process. These Merz-Kollman charges show good agreement with those calculated by AMBER for the peptide, with the exception of the region close to the platinum. However, this approach produces asymmetric charges for chemically identical atoms such as N-donor atoms in the 2,2-bipyridyl ligand. Additionally, these charges were found to be highly dependent on the platinum binding mode and the peptide conformation. This means that one set of these charges would be inadequate for describing all conformations of these systems.

Alternatively, it is possible to calculate Merz-Kollman charges in the same manner for only the region close the metal centre *i.e.* the ligand, the methyl imidazole regions of the coordinating histidine residues and the metal itself (see Figure 3.6). These small structures were optimised to obtain molecular symmetry prior to Merz-Kollman charge calculations. The remaining peptide region – relatively distant from the metal centre – was modelled using standard AMBER forcefield charges. This method produced chemically reasonable and essentially symmetric partial charges, unlike the approaches detailed above. Since a Merz-Kollman calculation of this size is rapid (approx. 50-60 atoms, depending on the ligand) manual input of these charges is not excessively time-consuming. Furthermore, since only a fast Merz-Kollman charge calculation is required, this approach is applicable to peptide systems of any size. As such, for practical as well as chemical reasons, this “MK-FF” approach was deemed to be both the most reliable and most applicable for both this and future work, and was used to model all  $\text{Pt}^{\text{II}}(\text{ligand})$  systems studied. Calculated charges are contained within Tables 3.10-3.11.

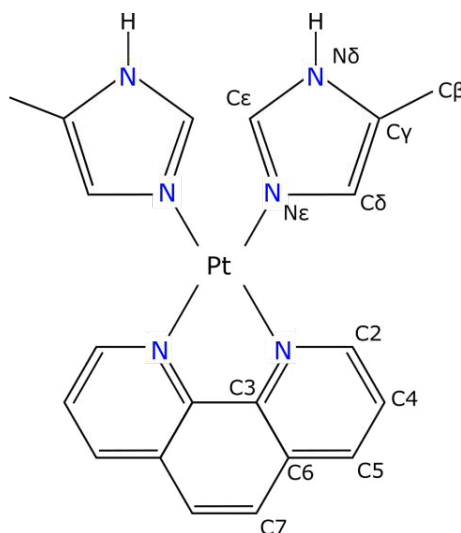


Figure 3.6: Atom names/numbers for an example  $\text{Pt}^{\text{II}}(\text{phen})(\text{Me-imid})_2$  MK partial charge calculation

Metal		His		Bipy	
Atom Name	Charge / $e$	Atom Name	Charge / $e$	Atom Name	Charge / $e$
Pt	0.560	N $\epsilon$	0.324	N1	0.093
		C $\epsilon$	-0.140	C2	-0.050
		N $\delta$	-0.350	C3	0.189
		C $\delta$	-0.448	C4	-0.243
		C $\gamma$	0.293	C5	0.069
		C $\beta$	-0.423	C6	-0.294
		H $\delta$ N	0.425	H2	0.195
		H $\delta$ C	0.267	H4	0.213
		H $\beta$ 1	0.160	H5	0.180
		H $\beta$ 2	0.162	H6	0.218

Table 3.10: Merz-Kollman partial charges for  $\text{Pt}^{\text{II}}(\text{bipy})(\text{Me-imid})_2$  subsystem of A $\beta$  peptide. Atom names are given in Figure 3.6.

Metal		His		Phen	
Atom Name	Charge / $e$	Atom Name	Charge / $e$	Atom Name	Charge / $e$
Pt	0.560	N $\epsilon$	0.324	N1	0.025
		C $\epsilon$	-0.140	C2	0.001
		N $\delta$	-0.350	C3	0.134
		C $\delta$	-0.448	C4	-0.262
		C $\gamma$	0.293	C5	-0.017
		C $\beta$	-0.423	C6	0.077
		H $\epsilon$	0.248	C7	-0.219
		H $\delta$ N	0.425	H2	0.190
		H $\delta$ C	0.267	H4	0.217
		H $\beta$ 1	0.160	H5	0.192
		H $\beta$ 2	0.162	H7	0.216

Table 3.11: Merz-Kollman partial charges for Pt<sup>II</sup>(phen)(Me-imid)<sub>2</sub> subsystem of A $\beta$  peptide. Atom names are given in Figure 3.6.

### 3.4.2 Exploring Conformations of Pt<sup>II</sup>-A $\beta$ 6-14

It is also important to establish a method for exploring the conformational freedom of the disordered A $\beta$  peptide, since there is little or no experimental data that provides a complete structure of the peptide or the platinated adducts.

Firstly, the MOE stochastic conformational search was used to model Pt<sup>II</sup>(ligand)-A $\beta$ 6-14, with the MK-FF charge scheme as discussed above. However, the random nature of this method lead to the inversion of amino acid chiral centres and generation of implausible chemical structures, where the peptide backbone passes through the centre of aromatic rings, giving a conformation that is tying itself in knots. The stochastic method was unable to escape these unlikely structures, leading to termination of the conformational search and insufficient exploration of possible conformers. This approach was not used beyond initial testing.

As explained in Sections 2.7.2 - 2.7.4, molecular dynamics does not provide fast enough exploration of possible conformations to be suitable for producing a large number of different structures in a computationally tractable time-frame. However, the LowMode MD method [29] in MOE is considered a rapid, accurate method

for generating low energy conformations of complex molecules (see Sections 2.7.2 - 2.7.4) and was used to explore the flexibility of this peptide fragment.

Several technical factors were investigated in order to establish the optimum Low-Mode MD protocol: the number of iterations in the search, the failure limit, the energy window and the RMSD.

- (i) The number of iterations is the maximum number of attempts to identify possible conformations. Clearly, a large number of iterations are required in order to adequately explore the conformational space and produce a representative sample of conformations of a given structure. Since this is simply a maximum number and we want to thoroughly investigate the possible conformations, an arbitrarily high iteration limit of 10,000 was selected throughout this work. The termination of the search was controlled by the other three factors:
- (ii) The failure limit is the number of successive iterations that fail to generate new conformations before the search terminates.
- (iii) The RMSD cut-off controls the detection of duplicate conformations *i.e.* if two conformations have a root mean square deviation (of atomic coordinates) less than a specified value then they are considered duplicates and the second conformer is not saved. Altering the RMSD cut-off allows us to control the number of conformations that are retained by changing the threshold for what is considered a duplicate *i.e.* large RMSD values restrict the number of conformers saved, while small RMSD values do not.
- (iv) The final factor in LowMode MD is the energy window. This is the value (in kcal mol<sup>-1</sup> - all other energy values in these results are reported in kJ mol<sup>-1</sup>) of the maximum allowable energy relative to the most stable conformation. Otherwise feasible conformations which are found by LFMM to be ‘high energy’ (*i.e.* outside this energy window) are discarded. Typically, structures with relative energies above 10 kcal mol<sup>-1</sup> are not significantly occupied at physiological temperature.

Tuning these settings to the research problem at hand affects the performance of the method. Here, these settings were altered systematically - akin to work by Chen and Foloppe [30] - to determine optimum LowMode parameters. Each set of parameters were used three times in order to negate the inherent random nature of the method and the average number of conformations found are reported in Tables 3.12-3.13.

During a LowMode MD conformational search a large portion of the kinetic energy was localised on the four Pt-N coordination bonds. These bonds are not important for exploring conformations since their final position is determined by the LFSE contribution to LFMM (*i.e.*  $\text{Pt}^{\text{II}}$  is always square planar). Often, this led to the LowMode search becoming ‘stuck’ and unable to escape a certain conformation, resulting in the premature termination of the LowMode procedure. This issue was resolved by setting the platinum centre at a fixed potential. This has the effect of making the mass of the fixed atom effectively infinite, meaning that when LowMode calculates vibrational frequencies, the bonds containing the fixed atom have a zero frequency. These bonds are then screened in the LowMode procedure, so the kinetic energy in the MD run is then concentrated on bonds that are more involved in peptide conformations.

System	Rejection Limit	RMSD / Å	E cutoff / kcal mol <sup>-1</sup>	Ave. No. Confs
A $\beta$ 6-14	100	0.25	7	9
	200	0.25	7	8
	500	0.25	7	16
	1000	0.25	7	18
	2000	0.25	7	11
	10000	0.25	7	14
	100	0.25	20	216
	200	0.25	20	257
	500	0.25	20	695
	1000	0.25	20	552
	2000	0.25	20	412
	10000	0.25	20	865
	100	0.25	50	8014
	200	0.25	50	7455
	500	0.25	50	7941
	1000	0.25	50	8093
	2000	0.25	50	7838
	1000	0.25	50	8007
	100	0.5	7	5
	200	0.5	7	7
	500	0.5	7	5
	1000	0.5	7	12
	2000	0.5	7	42
	10000	0.5	7	8
	100	0.5	20	112
	200	0.5	20	393
	500	0.5	20	617
	1000	0.5	20	854
	2000	0.5	20	614
	10000	0.5	20	545
	100	0.5	50	7325
	200	0.5	50	7362
	500	0.5	50	8092
	1000	0.5	50	7592
	2000	0.5	50	7730
	10000	0.5	50	7186

Table 3.12: Testing LMMD parameters with the free A $\beta$  peptide

These data show that for conformers of the free peptide, the energy cut-off is the most important factor - as this restriction is loosened, the number of conformations increases drastically. In addition, with a small energy cut-off (note that these values are in kcal mol<sup>-1</sup>), the number of structures found remains small, regardless of the failure limit or RMSD cut-off *i.e.* increasing the rejection limit or RMSD cut-off does not result in an appreciable increase in the number of identified conformations. Furthermore, increasing the rejection limit greatly extends the time taken for the conformational search, but does not produce significantly more structures.

System	Rejection Limit	RMSD / Å	E cutoff / kcal mol <sup>-1</sup>	Ave. No. Confs
Pt(Aβ)	100	0.25	20	12
	200	0.25	20	34
	500	0.25	20	21
	100	0.25	50	36
	200	0.25	50	57
	500	0.25	50	23
	100	0.25	100	104
	200	0.25	100	83
	500	0.25	100	97
	100	0.25	1000	221
	200	0.25	1000	156
	500	0.25	1000	253
	100	0.25	10000	672
	200	0.25	10000	689
	500	0.25	10000	681
	100	0.50	20	9
	200	0.50	20	25
	500	0.50	20	31
	100	0.50	50	33
	200	0.50	50	41
	500	0.50	50	42
	100	0.50	100	52
	200	0.50	100	71
	500	0.50	100	78
	100	0.50	1000	112
	200	0.50	1000	124
	500	0.50	1000	104
	100	0.50	10000	172
	200	0.50	10000	186
	500	0.50	10000	205
	100	1.00	20	5
	200	1.00	20	17
	500	1.00	20	44
	100	1.00	50	49
	200	1.00	50	22
	500	1.00	50	45



(cont.)

System	Rejection Limit	RMSD / Å	E cutoff / kcal mol <sup>-1</sup>	Ave. No. Confs
Pt(Aβ)	100	1.00	100	42
	200	1.00	100	19
	500	1.00	100	26
	100	1.00	1000	6
	200	1.00	1000	10
	500	1.00	1000	41
	100	1.00	10000	37
	200	1.00	10000	12
	500	1.00	10000	38

Table 3.13: Testing LMMD parameters with the Pt(Aβ) system

Equivalent data for the platinated system shows that this system is much more restricted than the free peptide with fewer conformations identified, regardless of settings. However, the energy cut-off is again key - increasing this lead to greater sampling of possible conformations. Importantly, these LFMM energies are not yet validated, so selecting a protocol with a large energy cut-off to generate many structures without pre-emptively screening them on an energetic basis was necessary. In addition, data for the RMSD cut-off indicates that a small value is required - this maximises the number of conformations and does not screen out those that are geometrically similar without being identical. A larger RMSD cut-off (1 Å) reduced the number of conformations drastically, leading to too few peptide conformations for a system of this size. On the basis of the above data, the final Low-Mode protocol was set to use 10,000 search iterations, 100 iteration failure limit, RMSD cut-off 0.25 Å and 10,000 kcal mol<sup>-1</sup> energy cut-off, as the LFMM energies are not validated. With the exception of the energy cut-off, this is identical to the optimum Low-Mode protocol determined by Chen and Foloppe[30].

### 3.4.3 Application of the Modelling Protocol

Following the above testing of methods, databases of conformations of Pt<sup>II</sup>(ligand)-Aβ6-14 were generated using the optimum parameters determined above: Complexes were described using molecular mechanics via a combination of ligand field molecular mechanics (LFMM) for Pt<sup>II</sup>[31, 32] and AMBER94 parameters for all other

atoms, as implemented in the DommiMOE extension to MOE. Partial charges were calculated as detailed in Section 3.4.1. Solvation effects were modelled using the reaction-field model with default parameters.

Conformational freedom was explored through the LowMode MD method in MOE. LowMode MD searches were configured to terminate after 100 successive failures to generate a new molecular conformation, up to a maximum of 10,000 iterations. A large energy cut-off (10,000 kcal mol<sup>-1</sup>) was used alongside an RMSD cut-off of 0.25 Å for retention of all new conformations.

Using this conformational search protocol, searches were performed on the eight platinum binding modes of interest for each ligand system. This produced between 400 and 1600 conformers for each platinum binding mode (Table 3.14).

Pt <sup>II</sup> Binding Mode	Bipy	Phen
His-6 Nδ His-13 Nδ	893	425
His-6 Nδ His-13 Nε	426	495
His-6 Nε His-13 Nδ	484	468
His-6 Nε His-13 Nε	995	405
His-6 Nδ His-14 Nδ	574	411
His-6 Nδ His-14 Nε	1570	454
His-6 Nε His-14 Nδ	429	618
His-6 Nε His-14 Nε	1135	450

Table 3.14: Number of conformations identified for each possible binding mode of two platinum-ligand species with  $A\beta$ .

There is appreciable variation between sets which may be at least partially due to the stochastic nature of the LowMode MD method as well as the inherent flexibility of the different complexes. For comparison, an identical simulation of the peptide fragment without Pt<sup>II</sup>-coordination results in 9,962 conformations from a total of 10,000 starting iterations. It is therefore apparent from Table 3.14 that Pt<sup>II</sup> binding to two His residues allows the peptide chain significant flexibility, no matter which coordination mode is considered, but that this conformational freedom is significantly reduced compared to the free peptide.

While the LMMD method used to generate data in Table 3.14 has been successfully applied to many similar problems, it is important to validate the results it provides against other theoretical approaches. In particular, we note that LFMM has long-standing use for metal-peptide binding, but that  $\text{Pt}^{\text{II}}$  is not yet among the metals used. However,  $\text{Pt}^{\text{II}}$  parameters are available, and have been used to model Pt-DNA complexes[32]. In the current case, the relatively flexible peptide fragment contrasts with more structured protein and DNA for which LFMM has typically been used. While LFMM produces accurate structures for small model  $\text{Pt}^{\text{II}}$  complexes (see Section 3.3.2), structures and relative energies of larger systems are yet to be tested.

Unfortunately, it is impractical to perform quantum-level calculations on these entire databases, so a subset of 158 unique low-energy conformers, encompassing several different platinum-binding modes, was selected in order to test further computational methods. DFT calculations were performed via Gaussian09, using the BHandH functional with a 6-31G(d) basis set on all atoms in conjunction with a Stuttgart-Dresden (SDD) effective core potential for platinum<sup>II</sup>, within a polarisable continuum model (PCM) model of aqueous solvation[19, 20], as tested above. In addition, two more modern functionals, B97D [4] and M06-2X[33], specifically designed to account for dispersion, were tested using the same SDD/6-31G(d)/PCM approach.

We first consider the relative energies of LFMM-generated structures, calculated without further geometry relaxation. We find little correlation between LFMM and BHandH relative energies at LFMM geometry ( $R^2 = 0.14$ ), suggesting that the combined LFMM/AMBER approach does not accurately predict the relative energies of different conformers of  $\text{Pt}^{\text{II}}$ -A $\beta$ 6-14 complexes.

In order to check the reliability of this DFT method used as a benchmark, a set of non-platinated conformations was constructed by removing the  $\text{Pt}^{\text{II}}$ (ligand) species from conformations located as discussed above, and their energies calculated using DFT and AMBER. The three different DFT methods (BHandH, B97D and M06-2X, all using the 6-31G(d) basis) show excellent agreement with one another ( $R^2$  BHandH vs. B97D = 0.94, BHandH vs. M06-2X 0.93, B97D vs. M06-2X 0.95), supporting the conclusion that BHandH is a suitable method for these structures.

Furthermore, these DFT methods show little agreement with the AMBER-TM en-

ergies for these non-platinated systems ( $R^2$  BHandH vs. AMBER =  $-0.02$ ). This suggests that the poor agreement between DFT and LFMM/AMBER energies described above stems from the use of AMBER94-TM for the peptide, rather than from use of LFMM for the metal centre.<sup>1</sup>

Due to the high computational expense of DFT calculations, it is not feasible to optimise all structures in the large conformational databases produced, as optimisation of a single conformer takes approximately three days on the computing resources available to us. This means that a faster computational method was required to identify conformers of interest.

To this end, the semi-empirical methods outlined in Section 2.6 may be used at a fraction of the computational cost. Several common semi-empirical methods (PM6, PM6-D3, PM6-DH+ and PM7) were used to model the X-ray crystal structures in Section 3.2 using the MOPAC [34] package and the COSMO model of aqueous solvation.[35] In the case of PM7, full relaxation was achieved through use of keywords LET and DDMIN=0. Results were compared to the original experimental data as well as the optimum DFT method determined above, BHandH/SDD/6-31G(d). Selected geometric data is contained in Tables 3.15 - 3.19.

Bond (Å) / Angle (°)	Expt.	BHandH	PM6	PM6-D3	PM6-DH+	PM7
Pt-N	1.994	2.013	1.932	1.935	1.932	1.981
Pt-NH <sub>2</sub>	2.046	2.050	2.054	2.057	2.053	2.086
Pt-Cl(1)	2.299	2.289	2.332	2.332	2.332	2.264
Pt-Cl(2)	2.309	2.292	2.318	2.318	2.319	2.264
Cl-Pt-Cl	91.7	93.6	96.1	95.8	96.4	94.5
N-Pt-Cl(2)	90.2	89.9	86.5	86.8	86.2	88.5
NH <sub>2</sub> -Pt-Cl(2)	86.9	83.7	79.8	80.0	79.7	83.5
MUE (Bond)		0.001	0.003	0.002	0.003	0.013
MUE (Angle)		0.5	2.1	2.1	2.2	0.8

Table 3.15: Selected geometric data for crystal structure XUKWUV, where data was obtained by experimental, DFT and semi-empirical methods.

<sup>1</sup>This issue of inaccurate AMBER-TM energies was later resolved with the help of Chemical Computing Group, the developers of the MOE software. As such, data within Chapter 5 uses the corrected AMBER-TM forcefield.

Bond (Å) / Angle (°)	Expt.	BHandH	PM6	PM6-D3	PM6-DH+	PM7
Pt-N(His)	2.026	2.008	1.949	1.948	1.944	1.954
Pt-N(ppy)	2.017	2.012	1.977	1.976	1.978	1.980
Pt-C(ppy)	1.978	1.975	1.997	1.996	1.997	1.952
Pt-Cl	2.411	2.396	2.431	2.430	2.428	2.356
C(ppy)-Pt-N(ppy)	81.3	81.1	82.2	82.3	82.2	81.2
N(His)-Pt-C(ppy)	94.8	96.4	99.2	99.1	96.6	100.1
N(His)-Pt-Cl	87.5	87.4	82.0	82.1	82.1	82.0
N(ppy)-Pt-Cl	96.6	95.2	96.7	96.6	97.1	96.7
MUE (Bond)		0.010	0.019	0.020	0.021	0.048
MUE (Angle)		0.0	0.0	0.0	0.1	0.1

Table 3.16: Selected geometric data for crystal structure OXUYUC, where data was obtained by experimental, DFT and semi-empirical methods.

Bond (Å) / Angle (°)	Expt.	BHandH	PM6	PM6-D3	PM6-DH+	PM7
Pt-N(imid)(1)	1.998	2.011	1.979	1.977	1.977	1.984
Pt-N(imid)(2)	1.998	2.012	1.979	1.977	1.977	1.982
Pt-NH <sub>3</sub> (1)	2.067	2.040	2.059	2.059	2.059	2.042
Pt-NH <sub>3</sub> (2)	2.067	2.040	2.059	2.059	2.059	2.042
NH <sub>3</sub> (1)-Pt-N(imid)(1)	90.4	89.9	89.7	89.5	90.0	90.2
NH <sub>3</sub> (1)-Pt-N(imid)(1)	89.6	90.2	90.0	89.8	90.2	89.8
MUE (Bond)		0.007	0.014	0.015	0.015	0.020
MUE (Angle)		0.1	0.2	0.4	0.1	0.0

Table 3.17: Selected geometric data for crystal structure DAMIPT, where data was obtained by experimental, DFT and semi-empirical methods.

Bond (Å) / Angle (°)	Expt.	BHandH	PM6	PM6-D3	PM6-DH+	PM7
Pt-N(imid)	2.026	2.002	1.963	1.962	1.963	1.973
Pt-NH <sub>2</sub> (1)	2.038	2.036	2.030	2.033	2.030	2.029
Pt-NH <sub>2</sub> (2)	2.040	2.062	2.067	2.070	2.066	2.044
NH <sub>2</sub> (2)-Pt-N(imid)	93.2	95.8	97.4	96.7	97.5	98.5
NH <sub>2</sub> (1)-Pt-NH <sub>2</sub> (2)	83.4	83.3	87.3	87.3	87.3	82.4
Cl-Pt-N(imid)	91.8	91.4	87.9	88.7	87.7	91.4
MUE (Bond)		0.001	0.015	0.013	0.015	0.019
MUE (Angle)		0.7	1.4	1.4	1.4	1.3

Table 3.18: Selected geometric data for crystal structure HEKJEM, where data was obtained by experimental, DFT and semi-empirical methods.

Bond (Å) / Angle (°)	Expt.	BHandH	PM6	PM6-D3	PM6-DH+	PM7
Pt-N(1)	1.985	2.013	1.976	1.975	1.975	1.971
Pt-N(2)	2.012	2.009	1.977	1.975	1.975	1.970
Pt-N(3)	1.985	2.013	1.976	1.975	1.975	1.971
Pt-N(4)	2.012	2.009	1.977	1.975	1.975	1.970
N(1)-Pt-N(2)	92.6	89.7	90.3	90.7	90.7	90.5
N(1)-Pt-N(4)	87.4	90.3	89.7	89.4	89.4	89.8
MUE (Bond)		0.012	0.022	0.024	0.024	0.028
MUE (Angle)		0.0	0.0	0.1	0.1	0.2

Table 3.19: Selected geometric data for crystal structure MIMDPT, where data was obtained by experimental, DFT and semi-empirical methods.

Here, there is relatively little difference in calculated geometries for each of the semi-empirical methods tested: PM6 agrees with experimental data to within 0.077 Å and 7.1°; PM6-D3 agrees within 0.078 Å and 6.9°; PM6-DH+ agrees within 0.082 Å and 7.2° and PM7 agrees within 0.073 Å and 5.5°. These data illustrate that semi-empirical methods are able to reproduce the geometry of these crystal structures to approximately the same degree of accuracy as DFT, giving confidence that they may perform well for Pt<sup>II</sup>A<sub>36-14</sub> systems.

While the semi-empirical PM6 method and its variants (PM6-D3 and PM6-DH+) are able to reproduce the experimental data seen above, the predicted relative en-

ergies of  $\text{Pt}^{\text{II}}\text{A}\beta 6\text{-}14$  conformations show little agreement with DFT - calculated  $R^2$  values for relative energies of these conformers against DFT are: PM6  $R^2 = 0.56$ ; PM6-D3  $R^2 = 0.61$ ; PM6-DH+  $R^2 = 0.66$ . However, PM7 and BHandH energies of  $\text{Pt}^{\text{II}}\text{-A}\beta 6\text{-}14$  conformations at MM geometries are in good agreement with one another ( $R^2 = 0.78$ ) - see Figure 3.7. PM7 therefore reproduces DFT relative energies at a much lower computational cost. This - along with data above for modelling crystal structures - suggests that PM7 is a suitable theoretical method for structures of this type. This assertion is reinforced by the modelling of the non-platinated conformations: BHandH and PM7 calculations show strong agreement ( $R^2 = 0.77$ ), further showing that PM7 is an accurate and reliable alternative to DFT.

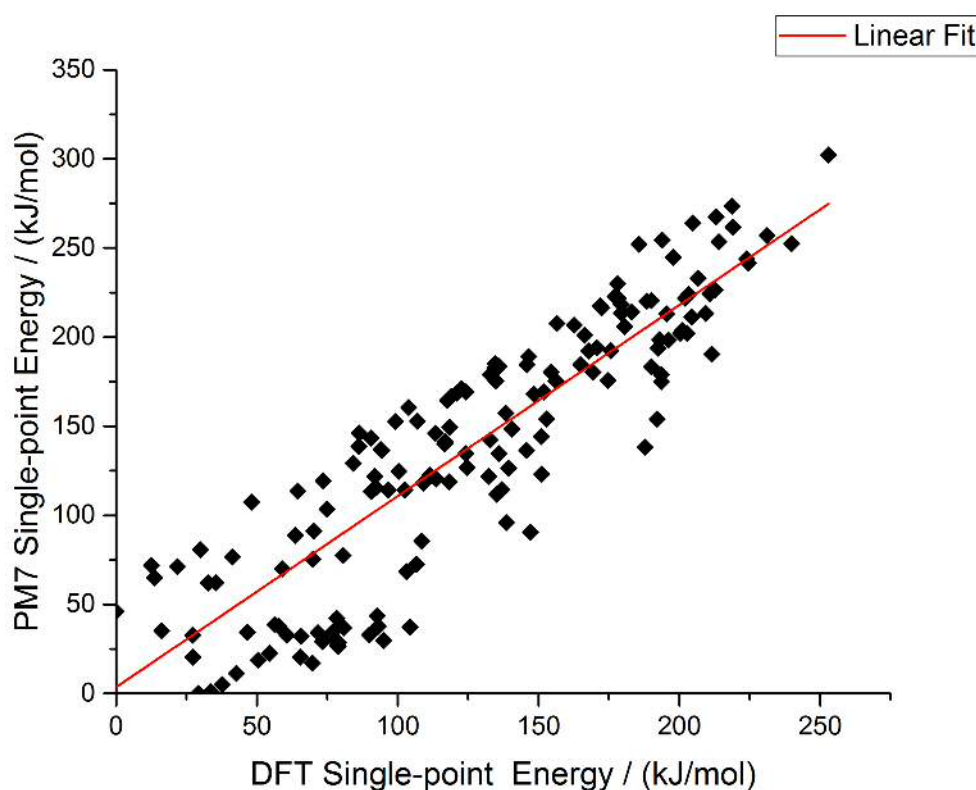


Figure 3.7: Plot of DFT vs. PM7 relative energies at LFMM geometry

The other key consideration of the LFMM results is whether the geometries produced are reliable. We therefore carried out DFT geometry optimisation of this subset of 158 platinated conformers, and evaluated the all-atom RMSD of the resulting structures from those obtained from LFMM energy minimisation. Here, overlay figures were obtained using Chimera imaging software.[36] RMSD values were calculated using a python script obtained via GitHub.[37][38] Table 3.21 reports the results of such comparison, and demonstrates that LFMM and DFT are in good general

-	LFMM	PM7	DFT
LFMM	1.00	0.38	0.14
PM7	-	1.00	0.78
DFT	-	-	1.00

Table 3.20: Correlation between relative energies of 158 conformers of  $\text{Pt}^{\text{II}}\text{A}\beta\text{6-14}$ , modelled using three different methods.

agreement *i.e.* DFT optimisation of a selected structure does not significantly alter the overall geometry, indicating that LFMM produces robust geometries (see Figures 3.8 - 3.9).

RMSD Values	LFMM vs. DFT	PM7 vs. DFT
Mean	0.670	0.817
SD	0.191	0.307
Min.	0.336	0.339
Max.	1.293	1.982

Table 3.21: Summary of geometry comparison between LFMM, PM7 and DFT.

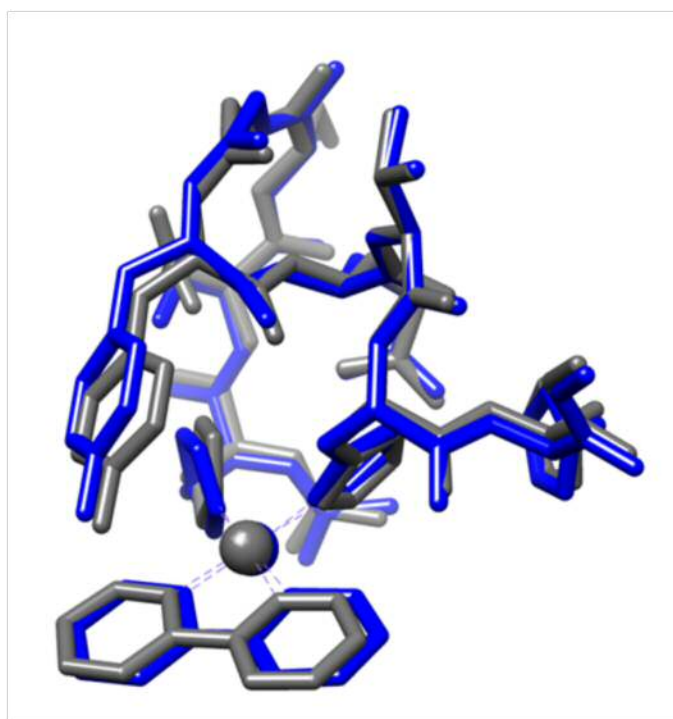


Figure 3.8: Representative overlay of LFMM (blue) vs DFT (grey) geometry of  $\text{Pt}^{\text{II}}(\text{bipy})\text{-A}\beta\text{6-14}$  complex.



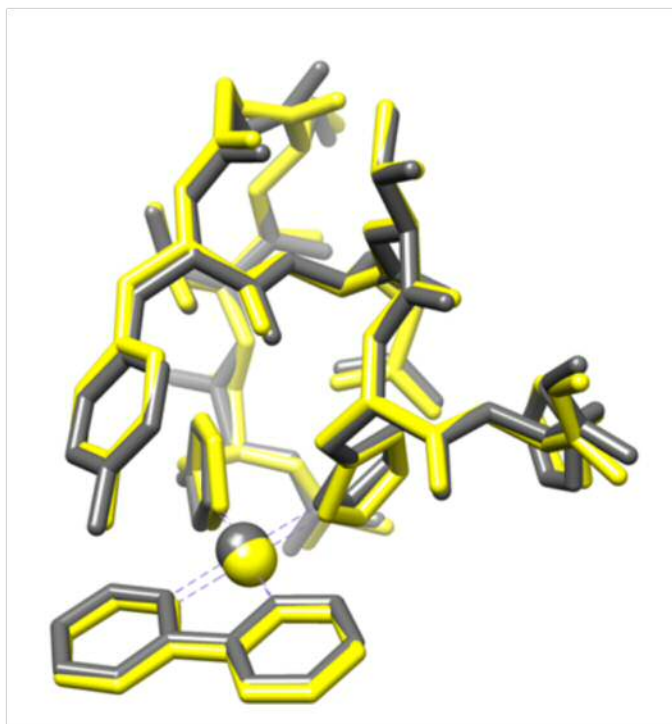


Figure 3.9: Representative overlay of PM7 (yellow) vs DFT (grey) geometry of  $\text{Pt}^{\text{II}}(\text{bipy})\text{-A}\beta 6\text{-14}$  complex.

We also investigated whether the much faster PM7 approach might be suitable for this purpose. Table 3.21 shows that, on average, PM7-optimised structures are slightly further from DFT ones than LFMM structures (mean RMSD 0.82 vs. 0.67 Å). These values include data from 9 conformers for which PM7 predicts large ( $> 1.3$  Å) RMSD from DFT-optimised geometry. For the remaining 149 structures, the mean RMSD is 0.77 Å. Interestingly, further DFT optimisation for the 9 complexes with large RMSD, starting from the PM7 endpoint produced significant reduction in energy, indicating that in some cases PM7 geometry optimisation in MOPAC is able to escape from the local minimum located by LFMM conformational search. One such case is illustrated in Figure 3.10 below, in which DFT optimisation following initial PM7 refinement leads to a structure that is  $62.3 \text{ kJ mol}^{-1}$  more stable than that reached by optimisation directly from the LFMM structure.

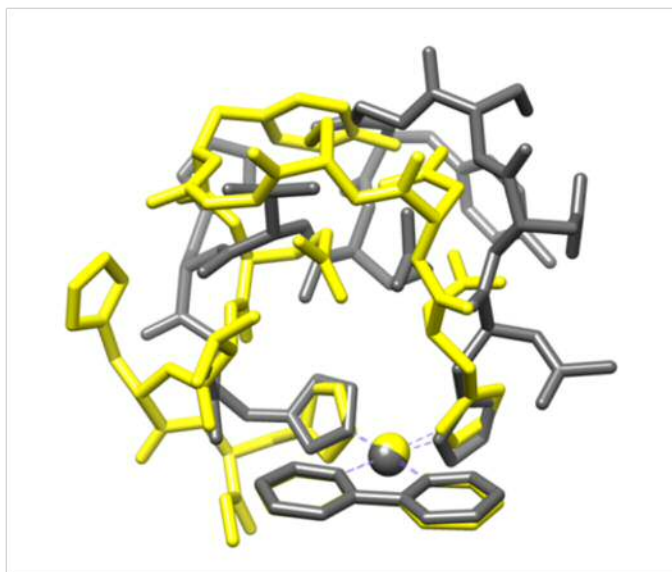


Figure 3.10: DFT optimised structure (grey) overlaid with PM7-DFT optimised structure (yellow) of  $\text{Pt}^{\text{II}}(\text{bipy})\text{-A}\beta 6\text{-14}$  complex.

We therefore surmise that PM7/COSMO optimisation following initial LFMM conformational search is a useful step in obtaining reliable energy and geometry data. Although the extra step requires significant additional computational effort (PM7 optimisation takes approximately 20 minutes per structure on a single compute node), the evidence above indicates that it provides a more balanced set of data than either LFMM alone or PM7 at the LFMM geometry.

We therefore performed this step for all conformations located for both ligands, a total of 6506 for 2,2-bipyridine (ligand 1) and 3726 for 4,10-phenanthroline (ligand 2). The resulting relative energies were collated and used in Boltzmann weighting to determine the binding mode(s) and conformation(s) that contribute significantly to the overall ensemble at 310 K, the results of which are reported in Table 3.22.

On the basis of these results, we predict that  $\text{Pt}^{\text{II}}$  binds to histidine residues via  $\text{N}\epsilon$  rather than  $\text{N}\delta$ . The first conformations that feature binding via  $\text{N}\delta$  of any His residue are found over 22 and 15  $\text{kJ mol}^{-1}$  for bipyridyl and phenanthroline ligands, respectively, and so contribute almost nothing to the Boltzmann-weighted ensemble of structures. This is an inherent property of the histidyl residue: calculations on a model system  $\text{Pt}(\text{bipy})(\text{Cl})\text{His}$  gives a preference for  $\text{N}\epsilon$  over  $\text{N}\delta$  of 5  $\text{kJ mol}^{-1}$  using DFT, and 3  $\text{kJ mol}^{-1}$  using PM7. The data in Table 3.22 indicate that, out of the several thousand conformations considered, just a small number contribute significantly to the Boltzmann-weighted ensemble. A single conformation accounts

Ligand	Coordination Mode	Rel. E / kJmol <sup>-1</sup>	Boltzmann Fraction
Bipy	His-6 NεHis-13 Nε	0.00	0.79
	”	6.27	0.07
	”	7.07	0.05
	”	7.51	0.04
Phen	His-6 NεHis-14 Nε	0.00	0.38
	”	0.77	0.28
	”	1.96	0.18
	”	4.14	0.08
	”	4.38	0.07

Table 3.22: Relative energy and Boltzmann fractions for low energy conformations of platinum complexes with ligands 1 & 2.

for almost 80% of the ensemble of bipy complexes, and four conformations for over 95%. Three low energy conformations account for over 80% of the ensemble of phen complexes, and five account for 99%. This shows that coordination of Pt massively reduces the conformational freedom of the peptide by “pinning” two histidines to the metal centre. The low-energy structures identified in this process are shown in Figures 3.11 and 3.12.

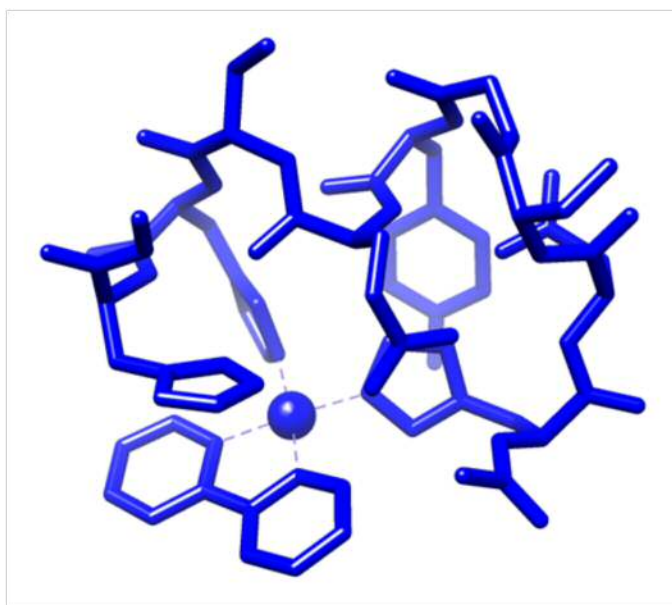


Figure 3.11: Low energy conformation of Pt<sup>II</sup>(bipy)Aβ6-14.

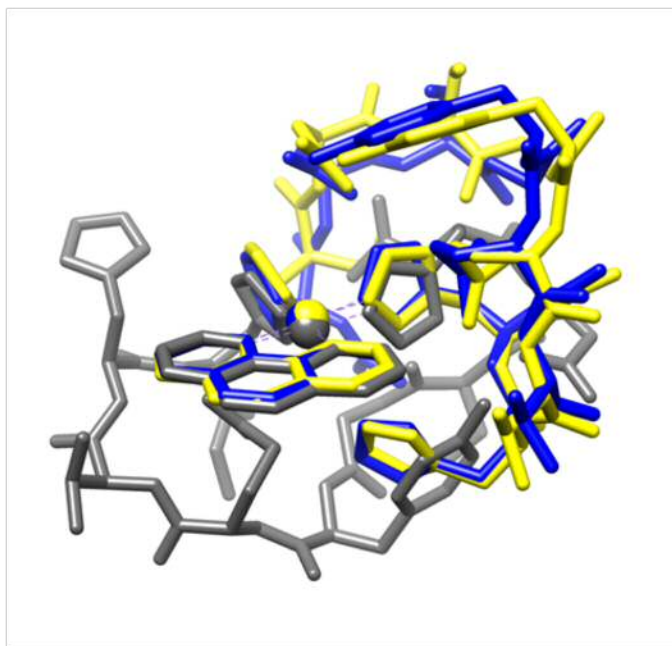


Figure 3.12: Lowest (grey), second- and third-lowest (yellow and blue, respectively) conformations of  $\text{Pt}^{\text{II}}(\text{phen})\text{A}\beta 6\text{-}14$  system.

There is also a clear preference in coordination mode for each ligand; metal binding occurs via  $\text{N}\epsilon$  of His-6 and His-13 for the bipy systems and via  $\text{N}\epsilon$  of His-6 and His-14 for the phenanthroline systems. This result is particularly promising for the phenanthroline case, since the preferred metal-binding mode identified here ( $6\epsilon - 14\epsilon$ ) was also determined to be the binding mode of this ligand system in experimental work carried out by Ma *et al.*[39] The fact that a relatively small change in ligand from bipy to phen alters the preferred coordination mode is somewhat surprising, and suggests that binding mode can be controlled through suitable choice of ligand.

To probe this change of binding mode in more detail, Table 3.23 reports selected details of the DFT-optimised structures of the low energy conformation(s) for each ligand system. These data show that there is almost no variation in  $\text{Pt}-\text{N}$  bond lengths following change in conformation and/or coordinated ligand. In the bipy complex, the bonds to the N-donor ligands are slightly longer than those to the histidine residues of  $\text{A}\beta$  (mean values  $1.963 \pm 0.004$  vs.  $1.986 \pm 0.000$  Å). Moreover,  $\text{Pt}-\text{N}_{\text{His}}$  distances are slightly shorter in the low energy bipy conformation than the average over all conformations located for the same binding mode (mean =  $1.966 \pm 0.005$  Å), whereas  $\text{Pt}-\text{N}_{\text{Lig}}$  distances are slightly larger (mean =  $1.981 \pm 0.004$  Å). However, differences are typically within 1 or 2 standard deviations, and so not

statistically significant. One apparently significant difference lies in the  $N_{\text{His}}\text{-Pt-}N_{\text{His}}$  angle that is on average  $84.5 (\pm 1.75)^\circ$  for the collection of  $6\epsilon - 13\epsilon$  conformations but  $88.5^\circ$  in the single low-energy conformation. It is notable that these  $\text{Pt}-N$  distances are in good agreement with those reported by Streltsov *et al.*  $\text{Pt}-N_{\text{imidazole}}$  and  $\text{Pt}-N_{\text{phen}}$  bond distances of  $1.99(1) \text{ \AA}$  and  $1.993(5) \text{ \AA}$ , respectively.[40]

Ligand	Structure	$\text{Pt-}N_{\text{His}}$ Length / $\text{\AA}$	$\text{Pt-}N_{\text{Lig}}$ Length / $\text{\AA}$	$N_{\text{His}}\text{-Pt-}N_{\text{His}}$ Angle / $^\circ$	$N_{\text{His}}\text{-Pt-}N_{\text{His}}$ Angle / $^\circ$
Bipy	#1	1.960, 1.966	1.986, 1.986	88.5	80.9
Phen	#1	1.958, 1.961	1.994, 1.989	85.2	82.1
	#2	1.964, 1.966	1.994, 1.993	83.1	81.8
	#3	1.961, 1.976	1.992, 1.995	82.8	81.6

Table 3.23: Selected geometrical details of the single low energy conformation of  $\text{Pt}^{\text{II}}(\text{bipy})$  complex and three low energy conformations of  $\text{Pt}^{\text{II}}(\text{phen})$  complex.

Ligand	Structure	No. H-bonds	No. $\pi$ - $\pi$ Interactions
Bipy	#1	14	1
Phen	#1	12	0
	#2	7	0
	#3	6	0

Table 3.24: Number of hydrogen bonds and  $\pi$ - $\pi$  interactions in the single low energy conformation of Pt-1 and three low energy conformations of Pt-2.

As expected for biomolecular systems, a large number of intramolecular hydrogen bonds are present in the low energy conformations, including backbone-backbone, backbone-side chain and side chain-side chain interactions, as shown in Table 3.24. The low energy  $\text{Pt-bipy}$  complex contains 14 H-bonds, as judged by geometrical criteria. Of these, five are concentrated around Glu11 and a further three around Asp7. Four of these H-bonds are relatively long at over  $2 \text{ \AA}$ , but a number of strong interactions were also identified, including those between backbone and residue ( $\text{Asp7}_{\text{sidechain } C=O} - \text{Tyr10}_{\text{backbone } N-H}$  at  $1.659 \text{ \AA}$  and  $\text{Glu11}_{\text{sidechain } C=O} - \text{N-terminal}_{N-H}$  at  $1.794 \text{ \AA}$ ) and residue-residue ( $\text{Glu11}_{\text{sidechain } C-O} - \text{His14}_{N\epsilon-H}$  at  $1.691 \text{ \AA}$  and  $\text{Asp7}_{\text{sidechain } C=O} - \text{His13}_{N\delta-H}$  at  $1.775 \text{ \AA}$ ) which may be responsible for the stability of this conformation. In addition, four backbone-backbone H-bonds were identified in residues Gly9-His14:  $\text{Gly9}_{C=O} - \text{Val12}_{N-H}$ ,  $\text{Tyr10}_{C=O} -$

His13<sub>N-H</sub>, Glu11<sub>C=O</sub> - His14<sub>N-H</sub> and Val12<sub>C=O</sub> - C-terminal<sub>N-H</sub>. The low energy bipy conformation also contains a possible  $\pi$ - $\pi$  interaction between bipy and the imidazole side chain of His14. The two ring systems lie in an approximately parallel orientation (angle between mean planes = 29.9°) with an inter-centroid distance of approximately 4 Å, as shown in Figure 3.13.

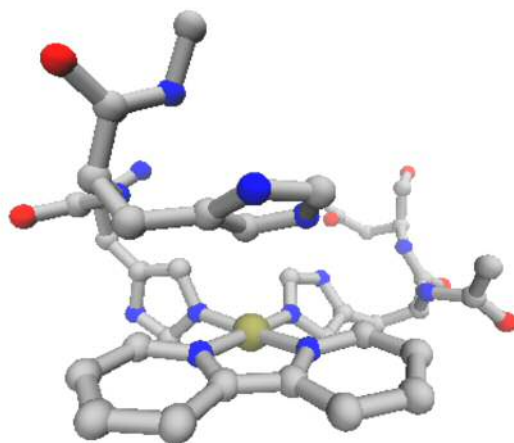


Figure 3.13:  $\pi$ -  $\pi$  interaction in the Pt<sup>II</sup>(bipy)A $\beta$ 6-14 system.

The Pt-phen complexes identified exhibit markedly fewer H-bonds than the bipy complex, and also no evidence for  $\pi$  -  $\pi$  stacking interactions is found in these low energy conformations (see Figure 3.12). As before, the H-bonds are generally found around residues Glu11 (Glu11<sub>sidechain C=O</sub> - Tyr10<sub>backbone N-H</sub>, Glu11<sub>sidechain C=O</sub> - His13<sub>backbone N-H</sub>, Glu11<sub>sidechain C=O</sub> - His14<sub>backbone N-H</sub>) and Asp7, including Asp7<sub>sidechain C=O</sub> - His6<sub>N $\delta$ -H</sub> and a strong interaction (average length 1.526 Å over the three phen conformations) between the hydroxyl group of Tyr10 and Asp7<sub>sidechain C=O</sub>.

### 3.5 Conclusions

LFMM has previously been demonstrated to be a powerful tool in predicting geometries and conformations of biomolecules bound to TM species. In the present work, we aim to apply this methodology to the study of platinum-complexes of the type studied by Barnham[41, 42] bound to model A $\beta$  fragments in order to determine favourable metal-binding modes and peptide conformations. Thorough exploration

of the conformational space of these biomolecules was achieved using LowMode MD in conjunction with AMBER molecular mechanics parameters. Comparison of calculated relative energies of a subset of these conformations indicates that LFMM fails to reproduce DFT results ( $R^2 = 0.14$ ). However, QM evaluation of structures of this size is computationally expensive, so the less expensive semi-empirical PM7 method was utilised, which agrees well with QM results ( $R^2 = 0.78$ ); PM7 therefore reproduces DFT energies at a much lower computational cost. In addition, PM7 optimisation generally retains the DFT structure, with average all-atom RMSD of  $0.817 \text{ \AA}$  over 158 conformers. This illustrates that PM7 accurately reproduces DFT geometries as well as energies. The present results indicate that PM7 is an appropriate and computationally manageable approach to modelling these metal-biomolecule systems.

This semi-empirical approach was subsequently extended to all conformations identified by LFMM. Boltzmann weighting was applied to the generated databases in order to generate relative populations of each conformer at 310K. It was found that platinum coordination occurs via the  $\text{N}\epsilon$  atom in His residues for both the bipy and phen ligand systems investigated. Favourable platinum-coordination modes were identified for each ligand:  $6\epsilon$ - $13\epsilon$  and  $6\epsilon$ - $14\epsilon$  for bipy and phen complexes, respectively. Furthermore, this change in binding mode for different ligands suggests that the binding mode can be controlled by the choice of ligand.

In the bipy systems, the Boltzmann population is dominated by a single conformer that accounts for almost 80% of the weighted ensemble containing a large number of intramolecular H-bonds and a possible  $\pi$ -stacking interaction between His 14 and the bipy ligand. In the phenanthroline case, three conformations make significant contributions to the ensemble. Each of these conformations contains fewer intramolecular H-bonds than the dominant bipy conformation and no evidence of  $\pi$ - $\pi$  stacking interactions was found.

## Bibliography

- (1) M. Turner, J. A. Platts and R. J. Deeth, *Journal of Chemical Theory and Computation*, 2016, **12**, 1385–1392.
- (2) M. Frisch and G. Trucks et al., *Gaussian09*, Wallingford, CT, 2009.
- (3) A. D. Becke, *The Journal of Chemical Physics*, 1993, **98**, 5648–5652.
- (4) S. Grimme, *Journal of Computational Chemistry*, 2006, **27**, 1787–1799.
- (5) C. Lee, W. Yang and R. Parr, *Physical Review B*, 1988, **37**, 785–789.
- (6) P. Stephens, F. Devlin, C. Chabalowski and M. Frisch, *Journal of Physical Chemistry*, 1994, **98**, 11623–11627.
- (7) J. Perdew, K. Burke and M. Ernzerhof, *Physical review letters*, 1996, **77**, 3865–3868.
- (8) J. P. Perdew, K. Burke and M. Ernzerhof, *Physical Review Letters*, 1997, **78**, 1396–1396.
- (9) J.-D. Chai and M. Head-Gordon, *Physical chemistry chemical physics: PCCP*, 2008, **10**, 6615–6620.
- (10) O. A. Vydrov and G. E. Scuseria, *The Journal of Chemical Physics*, 2006, **125**, 234109.
- (11) P. Garnuszek, J. K. Maurin, J. Witowska-Jarosz and B. Ptasiewicz-Bak, *Inorganica Chimica Acta*, 2002, **338**, 119 –126.
- (12) A. Esmaeilbeig, H. Samouei, S. Abedanzadeh and Z. Amirghofran, *Journal of Organometallic Chemistry*, 2011, **696**, 3135 –3142.
- (13) Carmicha.jw, N. Chan, D. Johnson, A. Cordes and C. Fair, *Inorganic Chemistry*, 1972, **11**, 1117–&.
- (14) G. Arvanitis, D. Gibson, T. Emge and H. Berman, *Acta Crystallographica Section C-Crystal Structure Communications*, 1994, **50**, 1217–1220.
- (15) M. B. Cingi, A. M. M. Lanfredi, A. Tiripicchio, C. G. V. Kralingen and J. Reedijk, *Inorganica Chimica Acta*, 1980, **39**, 265 –270.
- (16) P. Fuentealba, H. Preuss, H. Stoll and L. Vonszentpaly, *Chemical Physics Letters*, 1982, **89**, 418–422.



- 
- (17) D. Andrae, U. Haussermann, M. Dolg, H. Stoll and H. Preuss, *Theoretica Chimica Acta*, 1990, **77**, 123–141.
- (18) M. Malik and D. Michalska, *Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy*, 2014, **125**, 431–439.
- (19) S. Miertus, E. Scrocco and J. Tomasi, *Chemical Physics*, 1981, **55**, 117–129.
- (20) R. Cammi and J. Tomasi, *Journal of Computational Chemistry*, 1995, **16**, 1449–1458.
- (21) *Molecular Operating Environment (MOE)*, (2013.08), version 2013.08, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2013.
- (22) W. Cornell, P. Cieplak, C. Bayly, I. Gould, K. Merz, D. Ferguson, D. Spellmeyer, T. Fox, J. Caldwell and P. Kollman, *Journal of the American Chemical Society*, 1995, **117**, 5179–5197.
- (23) E. D. Scheeff, J. M. Briggs and S. B. Howell, *Molecular Pharmacology*, 1999, **56**, 633–643.
- (24) <http://www.daylight.com/dayhtml/doc/theory/theory.smiles.html>.
- (25) <http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html>.
- (26) V. J. Burton and R. J. Deeth, *Journal of the Chemical Society, Chemical Communications*, 1995, 573–574.
- (27) R. J. Deeth, N. Fey and B. Williams–Hubbard, *Journal of Computational Chemistry*, 2005, **26**, 123–130.
- (28) R. J. Deeth, A. Anastasi, C. Diedrich and K. Randell, *Coordination Chemistry Reviews*, 2009, **253**, 795–816.
- (29) P. Labute, *Journal of Chemical Information and Modeling*, 2010, **50**, 792–800.
- (30) I.-J. Chen and N. Foloppe, *Bioorganic & Medicinal Chemistry*, 2013, **21**, 7898–7920.
- (31) A. E. Anastasi and R. J. Deeth, *Journal of Chemical Theory and Computation*, 2009, **5**, 2339–2352.
- (32) H.-C. Tai, R. Brodbeck, J. Kasparkova, N. J. Farrer, V. Brabec, P. J. Sadler and R. J. Deeth, *Inorganic Chemistry*, 2012, **51**, 6830–6841.
- (33) Y. Zhao and D. G. Truhlar, *Theoretical Chemistry Accounts*, 2008, **120**, 215–241.
- (34) J. J. Stewart, <http://openmopac.net/index.html>.
- (35) A. Klamt and G. Schüürmann, *Journal of the Chemical Society, Perkin Transactions 2*, 1993, 799–805.

- 
- (36) E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *Journal of Computational Chemistry*, 2004, **25**, 1605–1612.
- (37) W. Kabsch, *Acta Crystallographica Section A*, 1976, **32**, 922–923.
- (38) J. Charnley, *GitHub: Calculate RMSD for two XYZ structures*, <https://github.com/charnley/rmsd>.
- (39) G. Ma, F. Huang, X. Pu, L. Jia, T. Jiang, L. Li and Y. Liu, *Chemistry-a European Journal*, 2011, **17**, 11657–11666.
- (40) V. A. Streltsov, V. C. Epa, S. A. James, Q. I. Churches, J. M. Caine, V. B. Kenche and K. J. Barnham, *Chemical Communications*, 2013, **49**, 11364–11366.
- (41) K. J. Barnham, V. B. Kenche, G. D. Ciccotosto, D. P. Smith, D. J. Tew, X. Liu, K. Perez, G. A. Cranston, T. J. Johanssen, I. Volitakis, A. I. Bush, C. L. Masters, A. R. White, J. P. Smith, R. A. Cherny and R. Cappai, *Proceedings of the National Academy of Sciences of the United States of America*, 2008, **105**, 6813–6818.
- (42) K. J. Barnham, V. B. Kenche, L. W. Hung, K. Perez, I. Volitakes, G. Ciccotosto, J. Kwok, N. Critch, N. Sherratt, M. Cortes, V. Lal, C. L. Masters, K. Murakami, R. Cappai and P. A. Adlard, *Angewandte Chemie-International Edition*, 2013, **52**, 3374–3378.

## 4 Predicting Ligand Effects on Platinum(L)-A $\beta$ Coordination

### 4.1 Introduction

The work in this chapter aims to utilise the modelling protocol established in Chapter 3 and apply it to a series of Pt<sup>II</sup>(ligand) systems that may be used as inhibitors of the A $\beta$  aggregation process. The choice of ligands studied is determined by recent literature on the subject (see Section 1.4.2). In particular, this chapter focuses on the differences in metal binding and peptide conformation induced by coordination of the different ligand systems.

Research contained in this chapter comes from work by our group, published in reference [1]. All of the work contained here from this publication was carried out by myself.

### 4.2 Application of the Modelling Protocol

Ligand field molecular mechanics (LFMM) and semi empirical Parametric Model 7 (PM7) methods are applied to a series of six Pt<sup>II</sup>-Ligand systems (shown in Figure 4.1) binding to the N-terminal domain of the A $\beta$  peptide (residues 6-14, see Figure 3.5), as described above. This peptide contains the N-terminal residues necessary to study transition metal binding, and is sufficiently small to allow calibration of findings through use of density functional theory (see Chapter 3). Molecular dynamics using a combined LFMM/Assisted Model Building with Energy Refinement (AMBER) approach is used to explore the conformational freedom of the peptide fragment, and identify favourable platinum binding modes and peptide conformations for each ligand investigated.

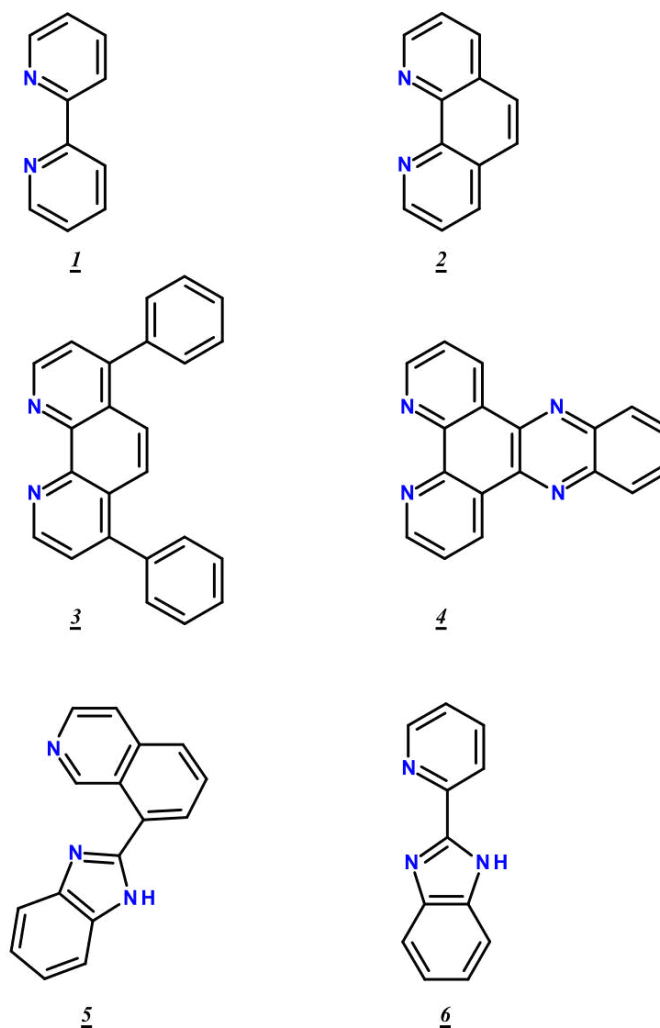


Figure 4.1: Schematic of ligand systems studied. Top row, left-right: bipyridyl (1), phenanthroline (2); Middle row, left-right: diphenylphenanthroline (3), dppz (4); Bottom row, left-right: 8-(1H-benzimidazol-2-yl)quinoline [2] (5) and 2-pyridyl-benzimidazole [3] (6).

### 4.2.1 Computational Details

The peptide sequence His6-Asp-Ser-Gly-Tyr-Glu-Val-His-His14 was built in an extended conformation in Molecular Operating Environment (MOE)[4], and protonation states at pH of 7.4 assigned using the Protonate3D module of this package.  $\text{Pt}^{\text{II}}$ (ligand) complexes were manually constructed and bound to the peptide in eight distinct binding modes from all combinations of  $\text{N}\delta$  and  $\text{N}\epsilon$  of His6 and His13/His14.

Complexes were described using a combination of ligand field molecular mechanics (LFMM) for  $\text{Pt}^{\text{II}}$  [5, 6] and Assisted Model Building with Energy Refinement 94

(AMBER94)[7] parameters for all other atoms, as implemented in the d-orbital extension to MOE (DommiMOE).[8] Partial charges were calculated for model  $\text{Pt}^{\text{II}}\text{-(imidazole)}_2\text{(ligand)}$  systems using the Merz-Kollman scheme from HF/6-31G(d)/SDD electrostatic potential in Gaussian09[9], with  $\text{Pt}^{\text{II}}$  given a van der Waals radius of 2.0 Å. The remaining peptide atoms were assigned AMBER94 charges as calculated by MOE. Solvation effects were modelled using the reaction-field model with default parameters.

Conformational freedom was explored via the LowMode MD (LMMD)[10] method in MOE.[4] LMMD searches were configured to terminate after 100 successive failures to generate a new conformation up to a maximum of 10,000 iterations, with a large energy cut-off (10,000 kcal/mol) and an Root-Mean-Square Deviation (RMSD) cut-off of 0.25 Å used for removal of any duplicate conformations. During the LMMD search, the platinum centre was set at a fixed potential to avoid premature termination of the conformational search.[11]

Semi-empirical calculations were performed using the Molecular Orbital PACkage (MOPAC) in its 2012 version[12], with the Parametric Model 7 (PM7) method[13] and the CONductor-like Screening MOdel (COSMO) model of aqueous solvation.[14] Overlay images were obtained using Chimera imaging software.[15] Ramachandran maps were plotted using the JMP statistical package.[16]

### 4.2.2 Modelling Six Different Ligand Systems

In this section, we generate and analyse conformations of a series of  $\text{Pt}^{\text{II}}$  species (see Figure 4.1) bound to a fragment of A $\beta$  peptide, with the aim of identifying favourable metal-binding modes and peptide conformations. Using the protocol outlined above, conformational searches were performed on each platinum binding mode for all ligand systems - 8 for the symmetric ligands **1** - **4**; 16 for the asymmetric ligands **5** and **6**, where the ligand 5-membered ring may be cis- or trans-to His6. This produced a large number of possible conformers for each platinum binding mode (Table 4.1).

Coordination Mode	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	Ratio <sup>a</sup>
His6 $\delta$ -His13 $\delta$	893	425	477	970	2531	1595	1.00
His6 $\delta$ -His13 $\epsilon$	426	495	1114	423	1640	3254	1.07
His6 $\epsilon$ -His13 $\delta$	484	468	469	680	1736	1707	0.80
His6 $\epsilon$ -His13 $\epsilon$	995	405	434	670	1709	1556	0.84
His6 $\delta$ -His14 $\delta$	574	411	1407	508	1204	4627	1.27
His6 $\delta$ -His14 $\epsilon$	1570	454	513	626	3146	1152	1.08
His6 $\epsilon$ -His14 $\delta$	429	618	404	403	1430	2118	0.78
His6 $\epsilon$ -His14 $\epsilon$	1135	450	768	298	1226	1350	0.76
Ratio <sup>b</sup>	1.00	0.57	0.86	0.70	2.67	2.25	

Table 4.1: Number of conformers found for each platinum coordination mode using LowMode MD for the six ligands studied.

<sup>a</sup>: Ratio of the total number of conformations of a coordination mode, relative to His6  $\delta$ - His13  $\delta$ .

<sup>b</sup>: Ratio of the total number of conformations of a ligand, relative to ligand 1

There is a great deal of variation in the number of conformations found between these six different ligands and coordination modes. As noted previously, an identical simulation using the equivalent free peptide fragment produced 9962 conformations, emphasising the fact that coordination of these platinum complexes greatly reduces the conformational freedom of A $\beta$ . [11]

It is particularly notable that both asymmetric ligands (5 and 6) produced significantly greater numbers of possible structures than their symmetric counterparts (1-4). While some of this difference may be attributed to the increased number of conformational searches required for these systems, *i.e.* two different simulations for the two orientations of the ligands, with the ligand 5-membered ring *cis*- or *trans*- to His6, this does not explain the large numbers of structures identified for certain coordination modes. This greater number of conformations suggests that these ligands do not restrict the flexibility of A $\beta$  to the same extent as the first four ligands.

It is also notable that the number of conformations for a given coordination mode are similar (Ratio<sup>a</sup> = 0.76-1.27, Table 4.1), but vary dramatically across the ligand series (Ratio<sup>b</sup> = 0.57-2.67). This suggests that the nature of the ligand affects the

flexibility of A $\beta$  more than the coordination mode.

As we have shown previously (work contained in Chapter 3 and reference [11]), PM7/COSMO optimisation of each conformation is able to provide both accurate geometries and relative energies of these systems. We therefore performed this step for all conformations identified above, a total of over 52,000 conformations. The resulting relative energies were collated and used in Boltzmann weighting calculations to determine the binding mode(s) and conformation(s) that contribute significantly to the overall ensemble at 310 K, the results of which are reported in Table 4.2.

Ligand	Coordination	Rel. E / kJ/mol	Boltzmann Fraction
<u><b>1</b></u>	His6 $\epsilon$ -His13 $\epsilon$	0.00	0.79
	His6 $\epsilon$ -His13 $\epsilon$	6.27	0.07
	His6 $\epsilon$ -His13 $\epsilon$	7.07	0.05
	His6 $\epsilon$ -His13 $\epsilon$	7.51	0.04
<u><b>2</b></u>	His6 $\epsilon$ -His14 $\epsilon$	0.00	0.38
	His6 $\epsilon$ -His14 $\epsilon$	0.77	0.28
	His6 $\epsilon$ -His14 $\epsilon$	1.96	0.18
	His6 $\epsilon$ -His14 $\epsilon$	4.14	0.08
	His6 $\epsilon$ -His14 $\epsilon$	4.38	0.07
<u><b>3</b></u>	His6 $\delta$ -His13 $\epsilon$	0.00	0.67
	His6 $\delta$ -His13 $\epsilon$	5.80	0.07
	His6 $\delta$ -His13 $\epsilon$	6.30	0.06
	His6 $\epsilon$ -His13 $\epsilon$	6.45	0.05
	His6 $\delta$ -His14 $\epsilon$	7.54	0.04
<u><b>4</b></u>	His6 $\delta$ -His14 $\delta$	0.00	0.99
	His6 $\epsilon$ -His13 $\delta$	12.69	0.01
	His6 $\epsilon$ -His13 $\delta$	13.22	0.01
<u><b>5</b></u>	<sup>a</sup> His6 $\delta$ -His14 $\epsilon$	0.00	0.39
	<sup>a</sup> His6 $\delta$ -His14 $\epsilon$	0.31	0.34
	<sup>a</sup> His6 $\delta$ -His14 $\epsilon$	3.00	0.12
	<sup>a</sup> His6 $\delta$ -His14 $\epsilon$	5.42	0.05
<u><b>6</b></u>	<sup>a</sup> His6 $\epsilon$ -His14 $\epsilon$	0.00	0.47
	<sup>a</sup> His6 $\epsilon$ -His14 $\epsilon$	1.79	0.23
	<sup>a</sup> His6 $\epsilon$ -His14 $\epsilon$	2.28	0.19
	<sup>a</sup> His6 $\delta$ -His14 $\delta$	3.74	0.11

Table 4.2: Relative energy and Boltzmann factors at 310K for low energy conformations.

<sup>a</sup>: His6 bound trans to ligand 5-membered ring

While low energy conformations are the main focus of this work, we note that the



range of relative energy of conformations found in this process is large; the maximum PM7 energy for each ligand system is as follows: Pt<sup>II</sup>(**1**): 583, Pt<sup>II</sup>(**2**): 403, Pt<sup>II</sup>(**3**): 233, Pt<sup>II</sup>(**4**): 372, Pt<sup>II</sup>(**5**): 657, Pt<sup>II</sup>(**6**): 726 kJ mol<sup>-1</sup>. That such high-energy conformations survive LFMM followed by PM7 optimisation is slightly surprising, but gives us reassurance that the conformational freedom of the peptide has been sampled. We also note that PM7 optimisation resulted in between 10 and 20 (1 to 2%) of conformations reported in Table 4.1 no longer being unique, as judged by peptide dihedral angles - a small value that attests to the quality of the geometries that result from LFMM and LowMode MD.

On the basis of these results, we predict that the binding of Pt<sup>II</sup> to histidine residues is strongly dependent on the identity of the ligand, though these systems exhibit some preference for Pt<sup>II</sup>-binding via N $\epsilon$  rather than N $\delta$  in His residues. This appears to be an inherent property of the histidyl residue: calculations on a series of model Pt<sup>II</sup>(Ligand)(Cl)His complexes using PM7 and DFT (performed at the BHandH/SDD/6-31G(d) level, as discussed in Chapter 3) gives preference for N $\epsilon$  binding in His residues, except in the case of ligands **4** and **6**, where His binding via N $\delta$  is strongly favoured (see Table 4.3). However, the relative stability of N $\epsilon$  vs. N $\delta$  binding in these model systems is not always reflected in the full peptide systems (see Table 4.2, with particular emphasis on ligand **6**) *i.e.* this preference can be overcome by stabilising other binding modes with intramolecular interactions, such as hydrogen-bonding or  $\pi$ - $\pi$  stacking arrangements, and/or low strain peptide conformations.

Ligand	N $\delta$ / N $\epsilon$ Binding	PM7 Rel. E / kJ/mol	DFT Rel. E / kJ/mol
<u><b>1</b></u>	$\delta$	6.79	17.46
	$\epsilon$	0.00	0.00
<u><b>2</b></u>	$\delta$	7.84	9.01
	$\epsilon$	0.00	0.00
<u><b>3</b></u>	$\delta$	1.09	10.68
	$\epsilon$	0.00	0.00
<u><b>4</b></u>	$\delta$	0.00	0.00
	$\epsilon$	19.50	21.84
<u><b>5</b></u>	$\delta 1^a$	4.08	2.75
	$\delta 2^b$	1.77	6.88
	$\epsilon 1^a$	4.26	1.85
	$\epsilon 2^b$	0.00	0.00
<u><b>6</b></u>	$\delta 1^a$	0.00	0.00
	$\delta 2^b$	25.59	12.26
	$\epsilon 1^a$	18.57	28.36
	$\epsilon 2^b$	17.22	24.20

Table 4.3: PM7 and DFT (BHandH/SDD/6-31G(d)) relative energies of different His binding modes in Pt<sup>II</sup>(ligand)(Cl)His model systems.

<sup>a</sup>: His bound trans to ligand 5-membered ring

<sup>b</sup>: His bound trans to ligand 6-membered ring

In the model systems using asymmetric ligands **5** and **6**, the orientation of the ligands appears to be important; the coordination of ligand **5** when His coordinates trans- to the 6-membered ring is significantly more stable than when His coordinates cis- to the 6-membered ring (see Figure 4.1 for reference). Interestingly, while N $\epsilon$  binding is favoured by approximately the same amount in the first two ligand systems, there is very little energy difference between the N $\delta$  / N $\epsilon$  binding for the ligand **3**. This may go some way to explaining the mixed N $\delta$  / N $\epsilon$  coordination mode seen in the low energy conformations (Table 4.2).

Furthermore, the orientation of the ligand remains important in the full peptide systems: in the case of ligand **5**, the lowest energy conformation where His6 coordinates trans- to the 6-membered ring rather than the 5-membered ring is located 46 kJ/mol above the most stable conformer. Similarly, in the case of ligand **6**, the first conformation identified where His6 coordinates trans- to the 6-membered ring appears 49 kJ/mol above the minimum energy conformer.

The data in Table 4.2 also indicate that a very small number of conformations contribute significantly to the Boltzmann-weighted ensemble at 310 K. For ligands **1**, **3** and **4**, a single conformer dominates the ensembles, whereas for **2**, **5** and **6** several conformers are predicted to co-exist with significant probability. The alternative low energy conformations of these complexes are discussed in more detail below. These results are consistent with our previous findings and reinforce the assertion that coordination of these Pt<sup>II</sup> complexes greatly reduces conformational freedom of A $\beta$ .<sup>[11]</sup> The low energy conformations identified in Table 4.2 are shown in Figures 4.2-4.7.

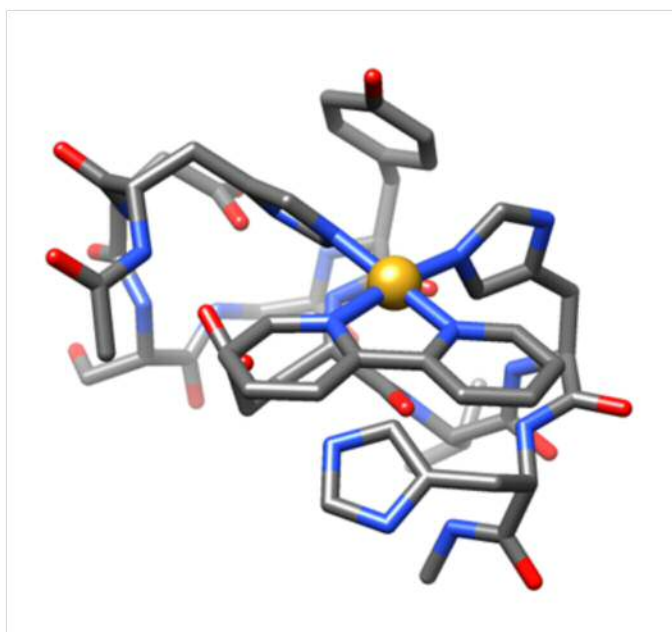


Figure 4.2: Low energy conformation of Pt<sup>II</sup>(**1**)-A $\beta$ 6-14

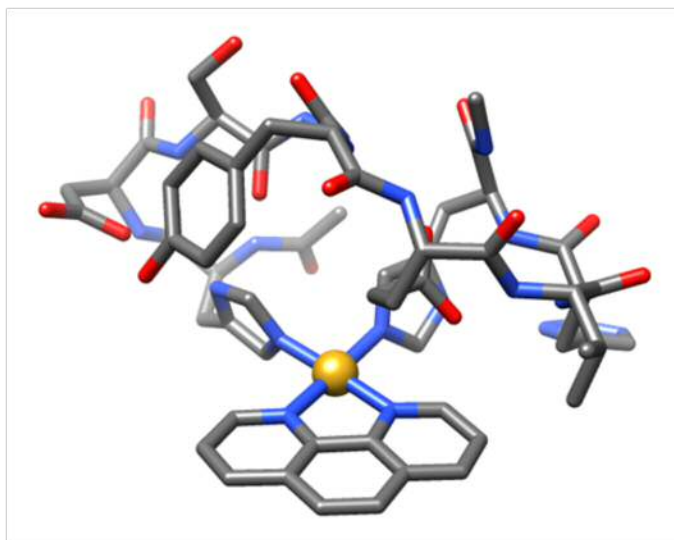


Figure 4.3: Low energy conformation of Pt<sup>II</sup>(**2**)-Aβ<sub>6-14</sub>

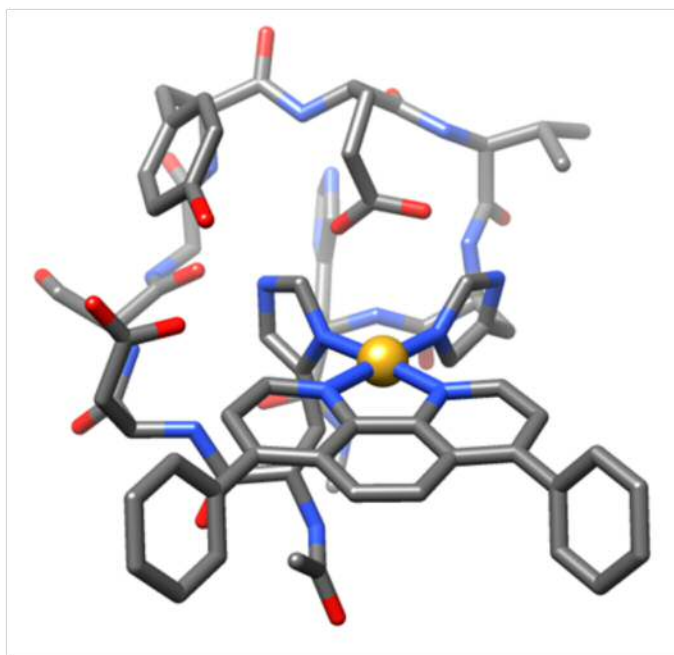


Figure 4.4: Low energy conformation of Pt<sup>II</sup>(**3**)-Aβ<sub>6-14</sub>

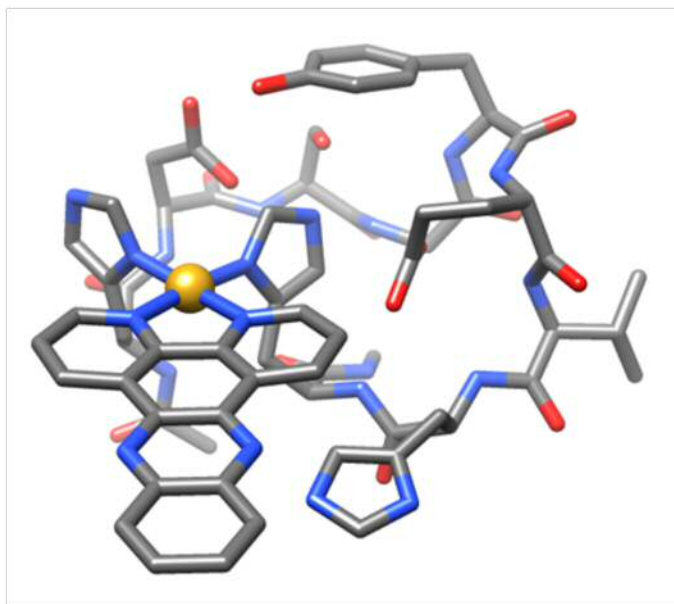


Figure 4.5: Low energy conformation of Pt<sup>II</sup>(4)-Aβ6-14

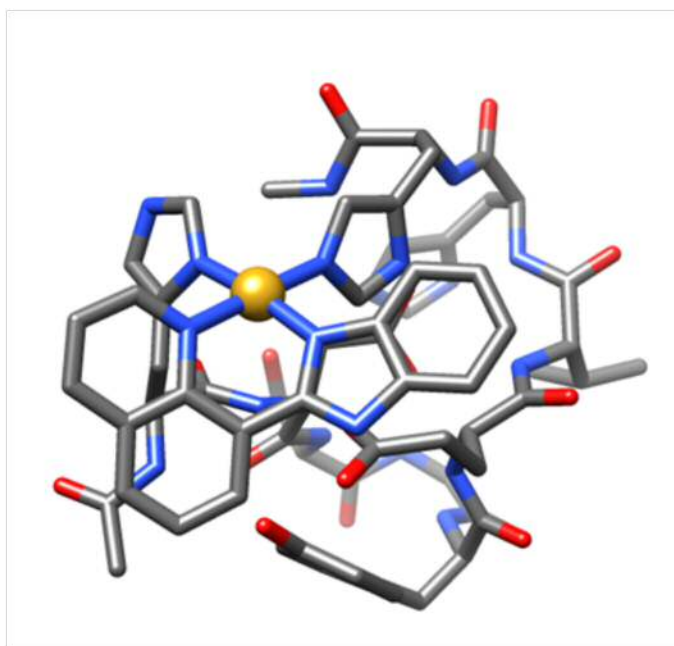


Figure 4.6: Low energy conformation of Pt<sup>II</sup>(5)-Aβ6-14

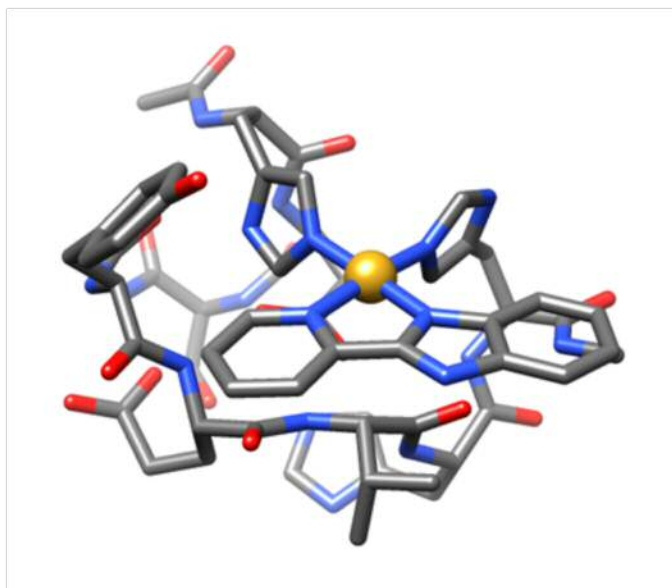
Figure 4.7: Low energy conformation of  $\text{Pt}^{\text{II}}(\underline{\mathbf{6}})\text{-A}\beta\mathbf{6-14}$ 

Table 4.2 also shows there is distinct preference in coordination mode for each of the ligands studied; this coordination is summarised in Table 4.4 below. This also provides further evidence that changing the ligand bound to platinum is able to influence the preferred  $\text{Pt}^{\text{II}}$ -binding mode and thus, that the binding mode may be controlled through choice and design of the ligand.

Ligand	Coordination Mode
<u><b>1</b></u>	His6 $\epsilon$ - His13 $\epsilon$
<u><b>2</b></u>	His6 $\epsilon$ - His14 $\epsilon$
<u><b>3</b></u>	His6 $\delta$ - His13 $\epsilon$
<u><b>4</b></u>	His6 $\delta$ - His14 $\delta$
<u><b>5</b></u>	<sup>a</sup> His6 $\delta$ - His14 $\epsilon$
<u><b>6</b></u>	<sup>a</sup> His6 $\epsilon$ - His14 $\epsilon$

Table 4.4: Favoured  $\text{Pt}^{\text{II}}$ -coordination mode for each ligand system.

<sup>a</sup>: His6 bound trans to ligand 5-membered ring

To investigate these results further, we focus on two structural features: i)  $\text{Pt}^{\text{II}}$  coordination and ii) peptide backbone geometry, in order to assess the range of geometries sampled in the conformational searching process and to identify any features of low energy conformations.

The calculated bond lengths and angles around the central platinum in each ligand system are summarised in Table 4.5. These data show that there is little variation in  $\text{Pt}^{\text{II}}\text{--N}(\text{His})$  bond lengths following change in conformation and/or coordinated ligand, but that there is notable variation of the  $\text{Pt}^{\text{II}}\text{--N}(\text{Lig})$  bond distances across the range of ligands. Additionally, these data show that there is a systematic difference in length of the  $\text{Pt}^{\text{II}}\text{--N}(\text{Lig})$  and  $\text{Pt}^{\text{II}}\text{--N}(\text{His})$  bonds – specifically the average  $\text{Pt}^{\text{II}}\text{--N}(\text{His})$  bond distances are slightly shorter than the corresponding  $\text{Pt}^{\text{II}}\text{--N}(\text{Lig})$  bond distances. However, comparison with the geometrical data for the low energy conformations (Table 4.5, every second row) clearly shows that these stable structures are not significantly different from the average conformation in each case, suggesting that the platinum geometry is not especially important in determining complex stability. It is also notable that these  $\text{Pt}^{\text{II}}\text{--N}$  distances are in good agreement with those of Streltsov et al.[17] for ligand **2**, who reported  $\text{Pt}^{\text{II}}\text{--N}(\text{imidazole})$  and  $\text{Pt}^{\text{II}}\text{--N}(\text{phen})$  bond distances of 2.03(1) Å and 1.993(5) Å, respectively, as well as confirming the square planar coordination of this ligand to amino acids.

Ligand	Pt-N <sub>Lig</sub> Bond Length / Å	Pt-N <sub>His</sub> Bond Length / Å	His-Pt-His Bond Angle / °
<b><u>1</u></b>	1.982 ± 0.005	1.966 ± 0.004	84.5 ± 1.75
	1.986 ± 0.000	1.963 ± 0.003	88.5
<b><u>2</u></b>	1.993 ± 0.059	1.967 ± 0.050	85.7 ± 3.65
	1.991 ± 0.003	1.960 ± 0.001	85.2
<b><u>3</u></b>	1.986 ± 0.005	1.966 ± 0.006	85.5 ± 2.10
	1.984 ± 0.000	1.960 ± 0.001	85.1
<b><u>4</u></b>	1.986 ± 0.031	1.967 ± 0.018	86.2 ± 2.15
	1.983 ± 0.001	1.964 ± 0.003	87.4
<b><u>5</u></b>	1.976 ± 0.037	1.968 ± 0.057	87.0 ± 2.65
	1.975 ± 0.013	1.963 ± 0.006	88.1
<b><u>6</u></b>	1.984 ± 0.040	1.965 ± 0.037	86.3 ± 2.38
	1.983 ± 0.014	1.961 ± 0.002	82.5

Table 4.5: Selected geometrical parameters over all conformations for each ligand (Å and °). First line: average and standard deviation; second line: value from lowest energy conformer

Backbone dihedral angles for all conformations were also calculated, allowing Ramachandran plots to be constructed. This allows us to identify any secondary struc-

ture elements formed within the usually random coil N-terminal region of A $\beta$ . The free A $\beta$ 6-14 fragment was modelled using an identical protocol, which identified 9962 possible conformations; backbone dihedrals for these conformations were calculated and used to produce a Ramachandran map using JMP[16] (see Figure 4.8).

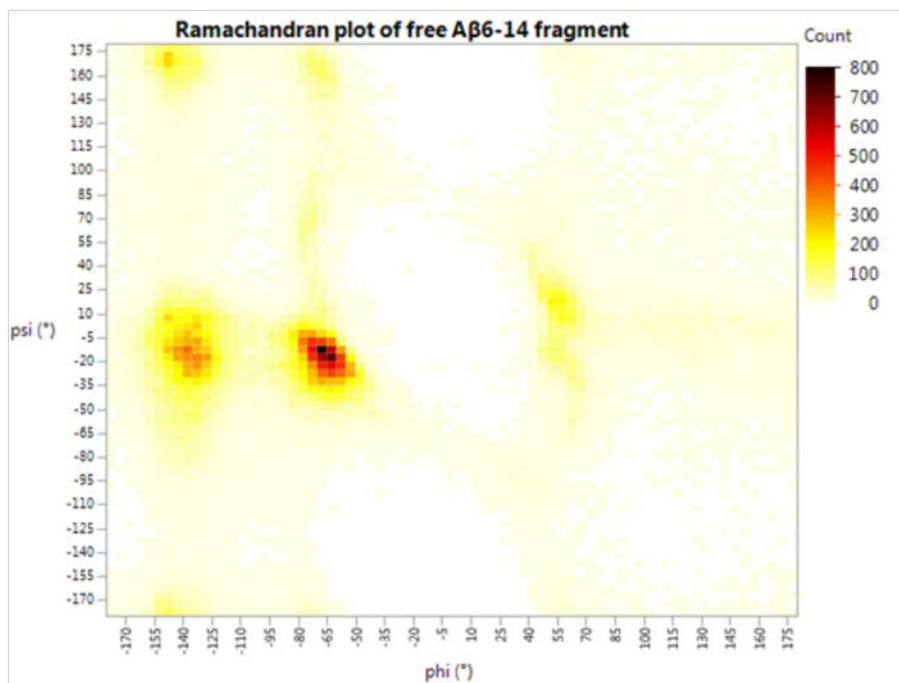


Figure 4.8: Ramachandran map of the generated conformations of the free A $\beta$ 6-14 fragment, plotted using JMP[16]

In agreement with previous findings, this fragment of the A $\beta$  peptide is highly flexible, illustrated by the wide coverage of the Ramachandran plot.[18–21] Conformations are concentrated around two distinct right-handed helical regions of conformational space, with smaller contributions from  $\beta$ -sheet and turn regions. Values of  $\phi$ ,  $\psi$  for the lowest energy conformation located are reported in Table 4.6, and secondary structure assigned by STRIDE[22] in Table 4.7. Despite occupying the region of  $\phi, \psi$  space expected for helices, A $\beta$ 6-14 residues are best described as turn or coil due to lack of H-bonding expected in  $\alpha$ - and  $3_{10}$ -helices, in agreement with the findings of Yang and Teplow.[21] Those authors also identified ‘highly populated turn structures, centred at [residues] 6-9’ of the free A $\beta$  peptide, and highlighted the importance of such turns for aggregation.

Ramachandran plots were then generated for each Pt<sup>II</sup>(ligand) system studied. For example, Figure 4.10 shows that for the favoured His6  $\epsilon$  - His14 $\epsilon$  coordination of the platinum complex with ligand 2, conformations cover a similar range to that observed



for the free peptide, but the concentration of residues in the right-handed helical region is reduced, with much greater scatter of  $\varphi$ ,  $\psi$  angles in the area expected for helical peptides, including a significant contribution at negative  $\varphi$ / positive  $\psi$ . Moreover, the Pt(2) complex exhibits a greater propensity for turn-like structures centred around  $+70$  /  $-60^\circ$ .

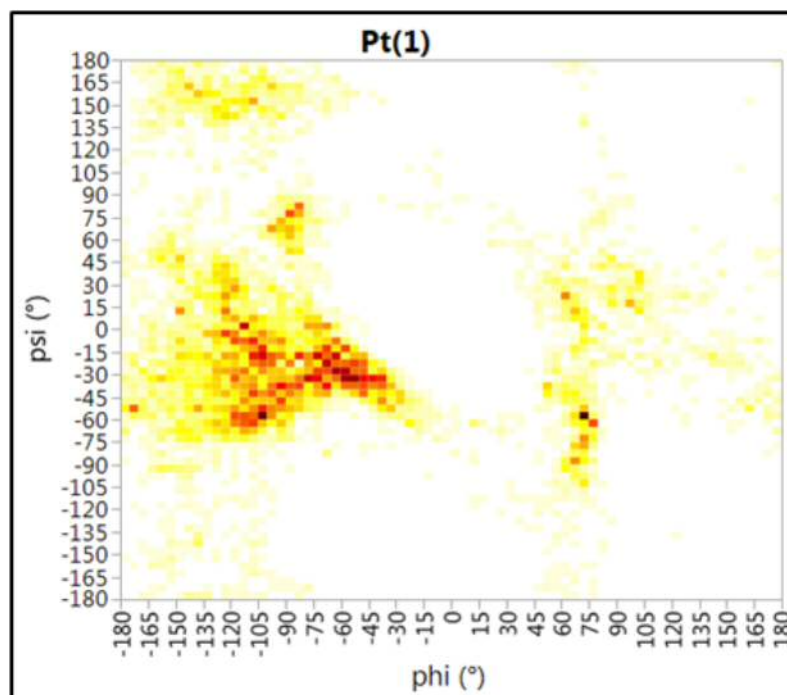


Figure 4.9: Ramachandran map for the generated conformations of the low energy binding mode of Pt<sup>II</sup>(1)-A $\beta$ 6-14

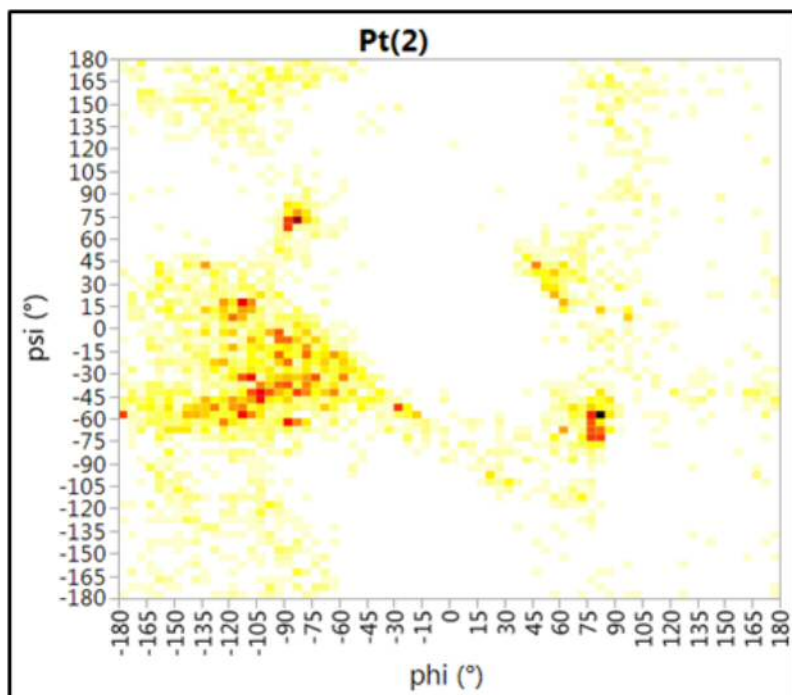


Figure 4.10: Ramachandran map for the generated conformations of the low energy binding mode of  $\text{Pt}^{\text{II}}(\underline{\mathbf{2}})\text{-A}\beta\mathbf{6}\text{-14}$

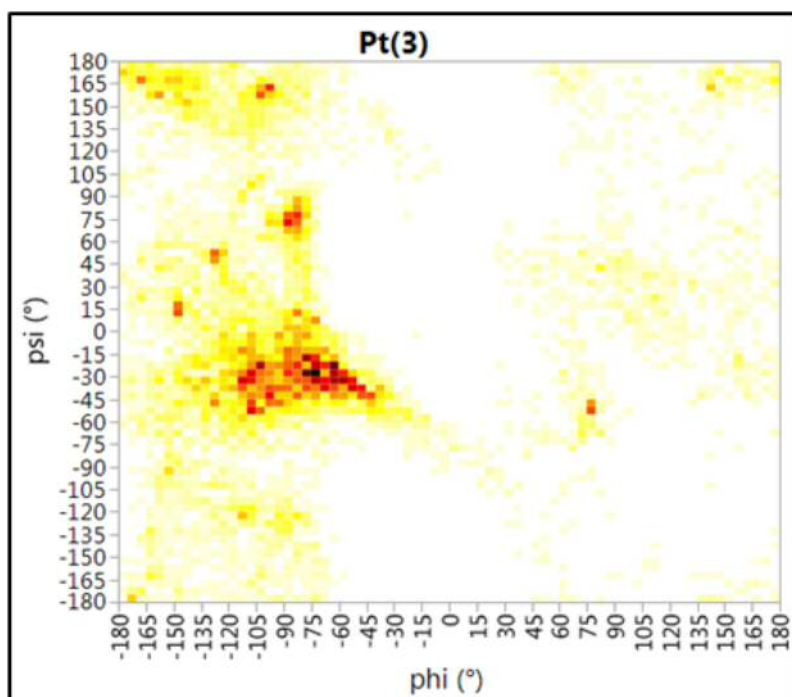


Figure 4.11: Ramachandran map for the generated conformations of the low energy binding mode of  $\text{Pt}^{\text{II}}(\underline{\mathbf{3}})\text{-A}\beta\mathbf{6}\text{-14}$

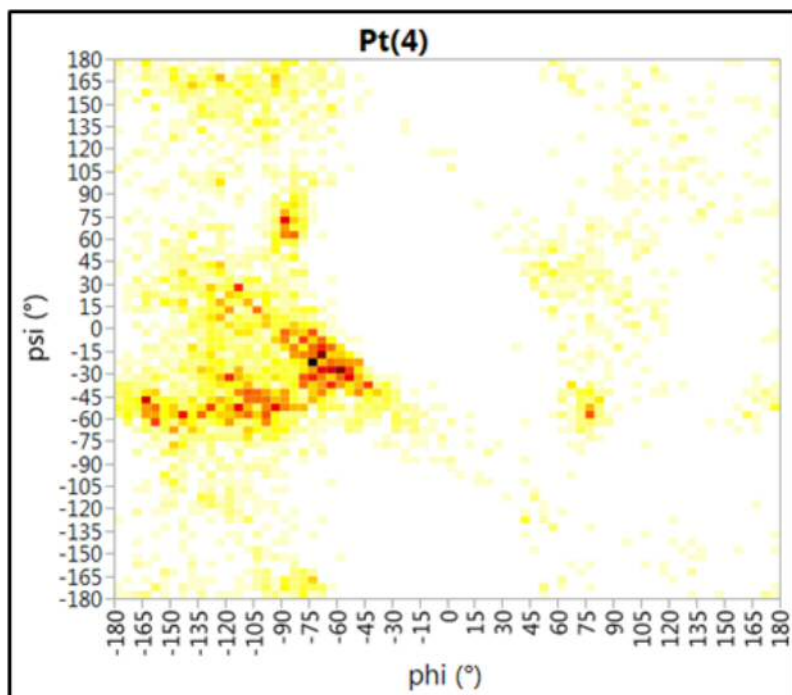


Figure 4.12: Ramachandran map for the generated conformations of the low energy binding mode of  $\text{Pt}^{\text{II}}(\underline{4})\text{-A}\beta 6\text{-14}$

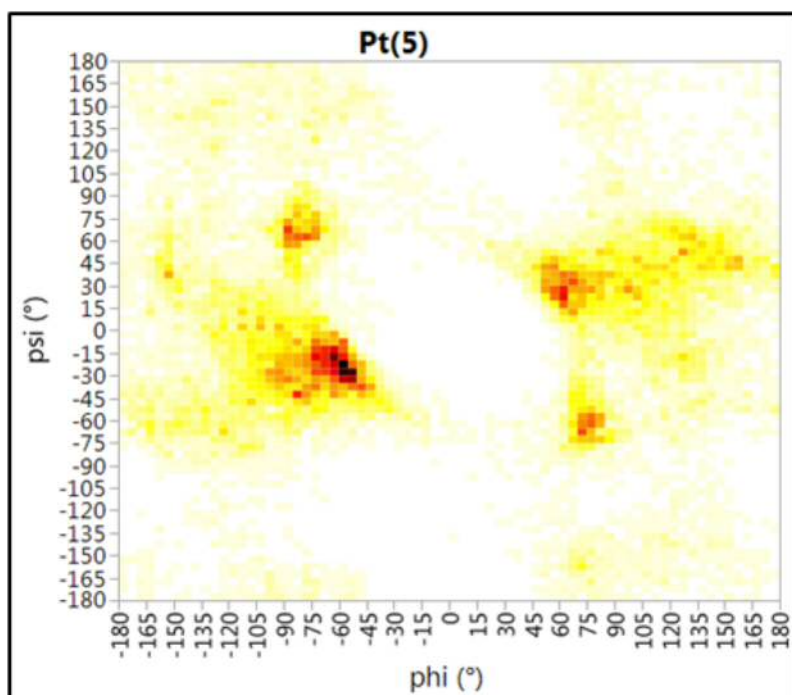


Figure 4.13: Ramachandran map for the generated conformations of the low energy binding mode of  $\text{Pt}^{\text{II}}(\underline{5})\text{-A}\beta 6\text{-14}$

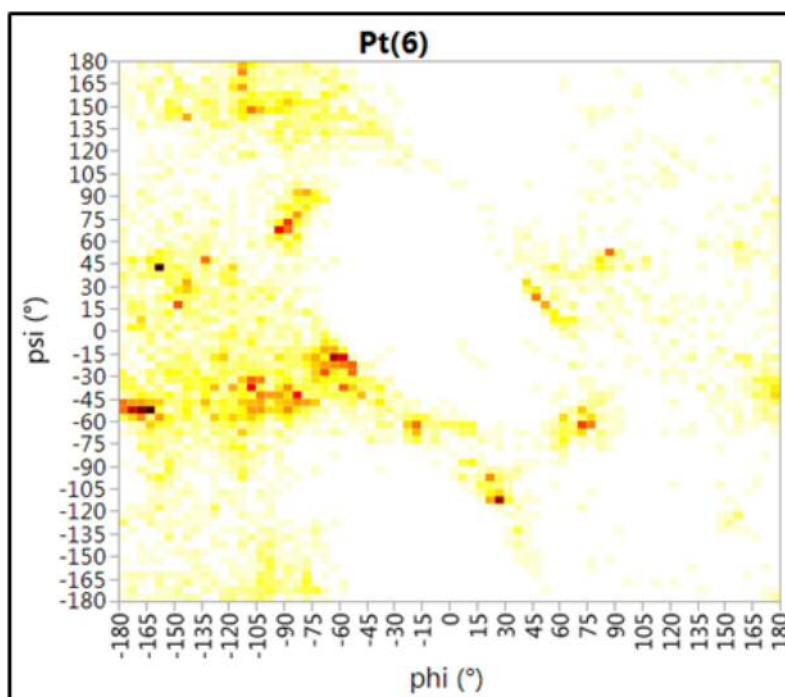


Figure 4.14: Ramachandran map for the generated conformations of the low energy binding mode of  $\text{Pt}^{\text{II}}(\underline{\mathbf{6}})\text{-A}\beta\mathbf{6}\text{-14}$

Analogous plots for each coordination mode with ligand **2** (see Appendix, Figures A.1 - A.8) illustrate Ramachandran maps are broadly similar regardless of platinum binding mode. In all cases, the second lowest energy conformation has similar backbone conformation to the lowest energy one, and in  $\text{Pt}(\underline{\mathbf{6}})$  especially the third conformation differs only in one angle of one residue (Tyr10). Thus, even in cases where multiple conformations are identified the overall flexibility of the peptide backbone is still very restricted. As noted above, three complexes exhibit multiple low-energy conformations.

However, as illustrated in Table 4.2, many of the generated conformations are of relatively high energy, so are of little biological relevance. The peptide geometries within these low energy conformations are summarised in Table 4.6.

Angle / °	Free A $\beta$	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>
$\varphi_{\text{His6}}$	-77.6	-94.8	-130.9	-81.2	-115.9	-125.7	-88.7
$\psi_{\text{His6}}$	-21.7	-49.4	-91.6	172.4	-64.0	-52.2	-27.5
$\varphi_{\text{Asp7}}$	-131.1	-143.0	-67.1	-101.0	-130.5	-52.1	-149.6
$\psi_{\text{Asp7}}$	-50.7	-17.3	-46.5	7.2	-34.6	177.6	-67.1
$\varphi_{\text{Ser8}}$	-162.7	76.1	-152.9	-149.0	-154.5	-149.1	-99.5
$\psi_{\text{Ser8}}$	6.3	-5.8	174.9	162.7	-67.2	46.3	-76.3
$\varphi_{\text{Gly9}}$	82.5	-110.3	-87.6	-80.3	101.2	-81.5	-112.6
$\psi_{\text{Gly9}}$	-43.0	-68.8	159.7	81.2	11.2	4.8	-95.0
$\varphi_{\text{Tyr10}}$	-86.3	-66.3	173.9	-131.7	-59.1	114.8	-126.4
$\psi_{\text{Tyr10}}$	-9.0	-12.5	38.9	-21.1	-25.4	32.6	73.1
$\varphi_{\text{Glu11}}$	70.3	-28.0	-157.3	-113.3	-99.1	77.7	-100.0
$\psi_{\text{Glu11}}$	11.1	-43.4	-108.8	-124.6	-10.4	55.9	-81.5
$\varphi_{\text{Val12}}$	-130.6	-55.5	-128.0	-101.4	74.0	-128.2	168.3
$\psi_{\text{Val12}}$	66.9	-5.8	-22.5	-27.6	37.3	-22.6	-30.7
$\varphi_{\text{His13}}$	-122.7	-67.6	-102.8	-109.8	-132.7	-116.0	-61.4
$\psi_{\text{His13}}$	-33.0	-19.6	-47.6	-34.6	-53.2	-10.1	-15.4
$\varphi_{\text{His14}}$	56.0	-111.1	-106.3	76.4	-96.1	103.1	-51.3
$\psi_{\text{His14}}$	17.9	-45.7	-0.2	-50.1	19.7	33.9	-30.7

Table 4.6: Dihedral angles (°) of the lowest energy conformations for free A $\beta$  in addition to each Pt<sup>II</sup>(ligand) complex studied.

Analysis of these conformations using STRIDE to identify any secondary structure elements was performed and results shown in Table 4.7. Data for the free peptide are in agreement with previous findings[23] of a turn in residues 5 – 9, with coil-like structure in residues 10-14. In general, a clear picture of disruption of the free peptide’s turn between residues 5 and 10 emerges from this analysis. The detailed effect of platination on this pattern varies with ligand: in most cases, a mix of turn and coil values are found, but the location of the turn residues changes. Pt(3) is probably the closest to the free peptide but truncates the observed turn by two residues, whereas Pt(2), Pt(4) and Pt(6) move the turn later in the sequence. Interestingly, STRIDE identifies a short stretch of  $3_{10}$ -helix in the low energy Pt(1) structure, while Pt(5) is assigned as almost completely turn in nature.

Residue	A $\beta$	Pt( <u>1</u> )	Pt( <u>2</u> )	Pt( <u>3</u> )	Pt( <u>4</u> )	Pt( <u>5</u> )	Pt( <u>6</u> )
His6	Turn	Coil	Coil	Turn	Coil	Coil	Coil
Asp7	Turn	Turn	Coil	Turn	Coil	Turn	Coil
Ser8	Turn	Turn	Turn	Turn	Coil	Turn	Coil
Gly9	Turn	Turn	Turn	Turn	Turn	Turn	Coil
Tyr10	Turn	Turn	Turn	Coil	Turn	Turn	Turn
Glu11	Turn	3,10	Turn	Coil	Turn	Turn	Turn
Val12	Coil	3,10	Turn	Coil	Turn	Turn	Turn
His13	Coil	3,10	Turn	Coil	Coil	Turn	Turn
His14	Coil	Coil	Coil	Coil	Coil	Turn	Coil

Table 4.7: Secondary structure analysis of the lowest energy conformations of free A $\beta$  and each Pt(ligand) system studied using STRIDE[22].

Conformations were also analysed for close contacts between peptide and ligand, defined as the distance between  $C_\gamma$  in a residue and a  $C$  atom in the centre of the ligand. This revealed that Tyr10 in particular forms numerous contacts with ligands, and that such contacts are more prevalent in the favoured coordination modes than in alternative forms. Figures 4.15 - 4.18 compare such contacts between the favoured coordination mode and the next most stable for Pt(1) and Pt(2). In both cases, the favoured coordination mode exhibits a distribution skewed towards short distances between ligand and Tyr10. The situation is less clear cut for other ligands, but in general the favoured coordination mode has on average one of the shortest contact distances. We speculate, therefore, that formation of such contacts are one of the main factors in determining how a particular ligand system alters how platinum binds to A $\beta$ .

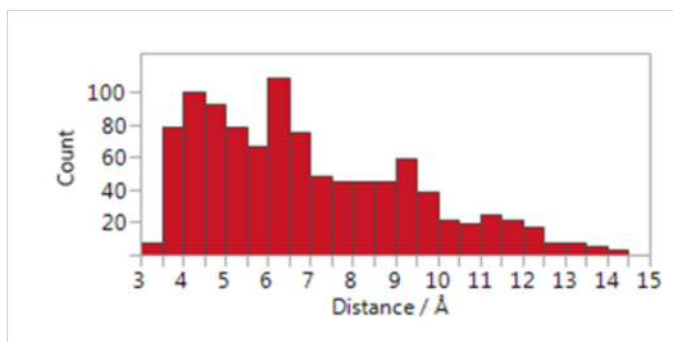


Figure 4.15: Ligand-Tyr10 distances in a) Pt<sup>II</sup>(1) His6  $\epsilon$  - His13  $\epsilon$

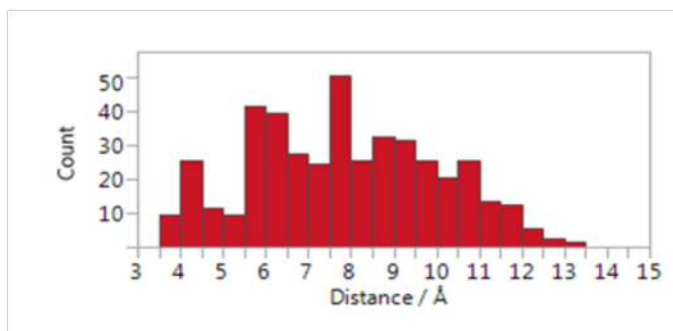


Figure 4.16: Ligand-Tyr10 distances in b)  $\text{Pt}^{\text{II}}(\underline{1})$  His6  $\delta$  - His13  $\epsilon$

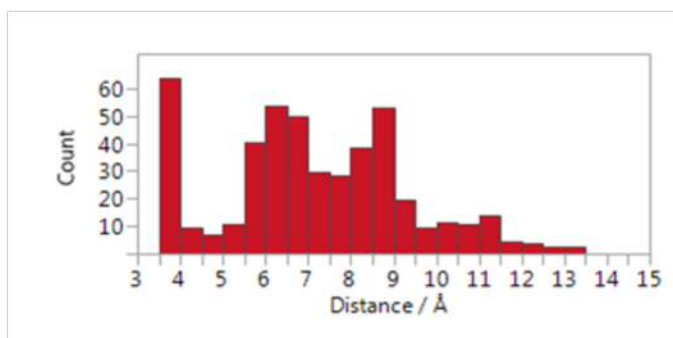


Figure 4.17: Ligand-Tyr10 distances in c)  $\text{Pt}^{\text{II}}(\underline{2})$  His6  $\epsilon$  - His14  $\epsilon$

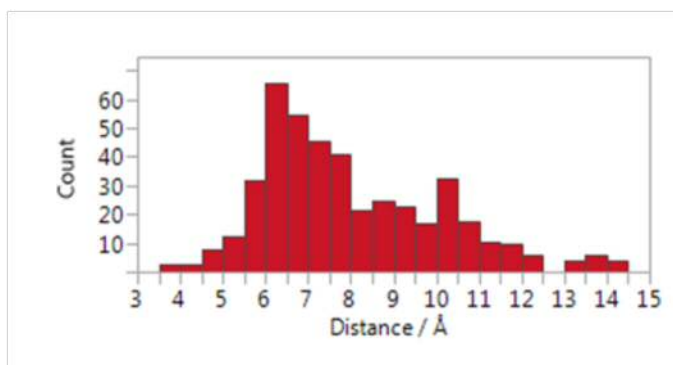


Figure 4.18: Ligand-Tyr10 distances in d)  $\text{Pt}^{\text{II}}(\underline{2})$  His6  $\delta$  - His13  $\delta$ .

### 4.3 Conclusions

LFMM is a powerful modelling tool for the study of transition metal complexes and has previously found success in predicting their interactions with biomolecules. In this work, we apply the LFMM approach to a series of six  $\text{Pt}^{\text{II}}$  complexes - as studied by Barnham[2, 3, 24] and others[25–27] - and their interaction with model fragments of the A $\beta$  peptide in order to determine favourable metal binding modes

as well as their influence on peptide secondary structure. Conformational space of A $\beta$  was explored using LowMode MD with AMBER molecular mechanics parameters, followed by further optimisation using the semi-empirical PM7 method.

Boltzmann populations for each conformation were calculated at 310 K and favourable platinum binding modes were identified for each complex studied. The preferred binding mode was found to be dependent on the nature of the ligand, with varied N $\delta$  or N $\epsilon$  coordination across the N-terminal histidine residues of A $\beta$ . These changes in binding mode suggest that the platinum coordination may be controlled via the choice of the ligand. Furthermore, the Boltzmann populations for these complexes indicate that a small number of conformations contribute significantly to the ensemble at 310 K, suggesting that the flexibility of the peptide is severely reduced after coordination of these platinum complexes.

Analysis of the platinum coordination geometry illustrated that the low energy conformations were not significantly different from the average across databases, meaning that their stability is likely conferred by the peptide conformation rather than the metal geometry, and in particular close contacts between residue side chains the planar ligands. It is surprising that such ligands, which differ only slightly in their steric and electronic nature, effect such large changes on coordination and conformation. The available peptide conformations were plotted as Ramachandran maps, allowing for identification of secondary structure elements within the peptide. While the N-terminal region of A $\beta$  is known to be disordered, the most stable conformation of the free peptide found here was shown to exhibit turn-type secondary structure in residues His6-Glu11. Coordination of each platinum-ligand complex disrupts this secondary structure differently, including interrupting or translating the existing turn-type structure in a ligand-specific fashion, as well as forming a short 3,10-helix. Restricting the conformational freedom and disrupting the secondary structure of A $\beta$  may have consequences for the effect of these complexes on limiting aggregation, for instance by hindering the formation of  $\beta$ -sheet structures known to be important in plaques.



## Bibliography

- (1) M. Turner, R. J. Deeth and J. A. Platts, *Journal of Inorganic Biochemistry*, 2017, **173**, 44–51.
- (2) K. J. Barnham, V. B. Kenche, L. W. Hung, K. Perez, I. Volitakes, G. Ciccosto, J. Kwok, N. Critch, N. Sherratt, M. Cortes, V. Lal, C. L. Masters, K. Murakami, R. Cappai and P. A. Adlard, *Angewandte Chemie-International Edition*, 2013, **52**, 3374–3378.
- (3) G. S. Yellol, J. G. Yellol, V. B. Kenche, X. M. Liu, K. J. Barnham, A. Donaire, C. Janiak and J. Ruiz, *Inorganic Chemistry*, 2015, **54**, 470–475.
- (4) *Molecular Operating Environment (MOE)*, (2013.08), version 2013.08, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2013.
- (5) A. E. Anastasi and R. J. Deeth, *Journal of Chemical Theory and Computation*, 2009, **5**, 2339–2352.
- (6) H.-C. Tai, R. Brodbeck, J. Kasparkova, N. J. Farrer, V. Brabec, P. J. Sadler and R. J. Deeth, *Inorganic Chemistry*, 2012, **51**, 6830–6841.
- (7) W. Cornell, P. Cieplak, C. Bayly, I. Gould, K. Merz, D. Ferguson, D. Spellmeyer, T. Fox, J. Caldwell and P. Kollman, *Journal of the American Chemical Society*, 1995, **117**, 5179–5197.
- (8) R. J. Deeth, N. Fey and B. Williams–Hubbard, *Journal of Computational Chemistry*, 2005, **26**, 123–130.
- (9) M. Frisch and G. Trucks et al., *Gaussian09*, Wallingford, CT, 2009.
- (10) P. Labute, *Journal of Chemical Information and Modeling*, 2010, **50**, 792–800.
- (11) M. Turner, J. A. Platts and R. J. Deeth, *Journal of Chemical Theory and Computation*, 2016, **12**, 1385–1392.
- (12) J. J. Stewart, <http://openmopac.net/index.html>.
- (13) J. J. P. Stewart, *Journal of Molecular Modeling*, 2013, **19**, 1–32.
- (14) A. Klamt and G. Schüürmann, *Journal of the Chemical Society, Perkin Transactions 2*, 1993, 799–805.

- 
- (15) E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *Journal of Computational Chemistry*, 2004, **25**, 1605–1612.
- (16) JMP®, Version 12.0.1 SAS Institute Inc., Cary, NC, 1989-2007.
- (17) V. A. Streltsov, V. C. Epa, S. A. James, Q. I. Churches, J. M. Caine, V. B. Kenche and K. J. Barnham, *Chemical Communications*, 2013, **49**, 11364–11366.
- (18) F. Chiti and C. M. Dobson, *Annual Review of Biochemistry*, 2006, **75**, 333–366.
- (19) N. G. Sgourakis, Y. Yan, S. A. McCallum, C. Wang and A. E. Garcia, *Journal of Molecular Biology*, 2007, **368**, 1448–1457.
- (20) N. G. Sgourakis, M. Merced-Serrano, C. Boutsidis, P. Drineas, Z. Du, C. Wang and A. E. Garcia, *Journal of Molecular Biology*, 2011, **405**, 570–583.
- (21) M. Yang and D. B. Teplow, *Journal of Molecular Biology*, 2008, **384**, 450–464.
- (22) D. Frishman and P. Argos, *Proteins-Structure Function and Genetics*, 1995, **23**, 566–579.
- (23) D. B. Teplow, N. D. Lazo, G. Bitan, S. Bernstein, T. Wyttenbach, M. T. Bowers, A. Baumketner, J.-E. Shea, B. Urbanc, L. Cruz, J. Borreguero and H. E. Stanley, *Accounts of Chemical Research*, 2006, **39**, 635–645.
- (24) K. J. Barnham, V. B. Kenche, G. D. Ciccotosto, D. P. Smith, D. J. Tew, X. Liu, K. Perez, G. A. Cranston, T. J. Johanssen, I. Volitakis, A. I. Bush, C. L. Masters, A. R. White, J. P. Smith, R. A. Cherny and R. Cappai, *Proceedings of the National Academy of Sciences of the United States of America*, 2008, **105**, 6813–6818.
- (25) G. Ma, F. Huang, X. Pu, L. Jia, T. Jiang, L. Li and Y. Liu, *Chemistry-a European Journal*, 2011, **17**, 11657–11666.
- (26) F. Collin, I. Sasaki, H. Eury, P. Faller and C. Hureau, *Chemical Communications*, 2013, **49**, 2130–2132.
- (27) G. Ma, E. Wang, H. Wei, K. Wei, P. Zhu and Y. Liu, *Metallomics*, 2013, **5**, 879.

## 5 Molecular Dynamics Simulations of $\text{Pt}^{\text{II}}(\text{L})$ - $\text{A}\beta$ Systems

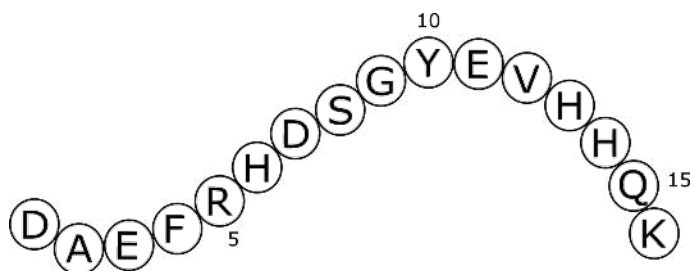
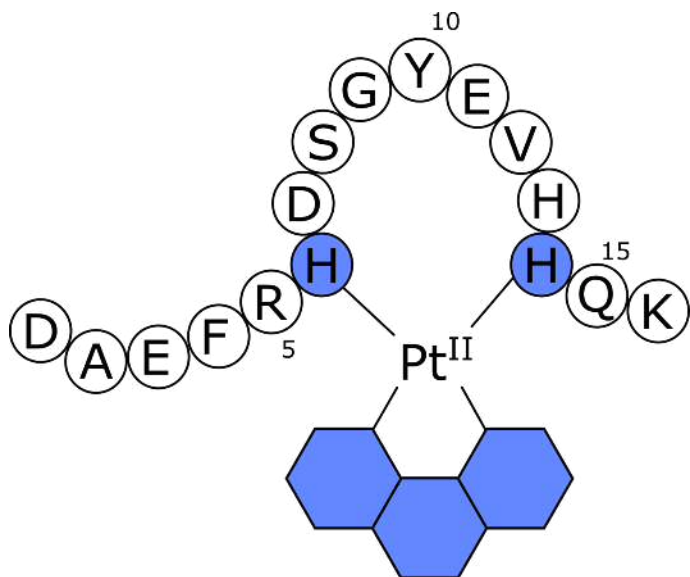
### 5.1 Introduction

In this chapter, molecular dynamics simulations will be used to explore the behaviour of  $\text{A}\beta$  over time, and the effect of  $\text{Pt}^{\text{II}}(\text{phenanthroline})$  coordination on the structure, interactions and dynamics of the peptide. Here, simulations of the free  $\text{A}\beta_{16}$  and  $\text{A}\beta_{42}$  peptides will be used to illustrate the effects of  $\text{Pt}^{\text{II}}$ -coordination.

Research in this chapter comes from work by our group, contained in reference [1]. All of the work contained here from this publication was carried out by myself.

### 5.2 $\text{A}\beta_{16}$ and $\text{Pt}^{\text{II}}(\text{phen})$ - $\text{A}\beta_{16}$ Simulations

The sixteen-residue N-terminal fragment of  $\text{A}\beta$  is considered a useful model of the full peptide - it is often used in experimental work as it has better handling properties and does not aggregate. Importantly, this peptide fragment contains the N-terminal metal-binding residues (*e.g.* histidines) and has been used in studies of the coordination and effects of  $\text{Pt}^{\text{II}}$  complexes (see 1.4.2). Illustrations of this sixteen-residue peptide fragment with and without the  $\text{Pt}^{\text{II}}(\text{phen})$  unit attached are shown in Figures 5.1 and 5.2, respectively.

Figure 5.1: Schematic of A $\beta$ 16Figure 5.2: Schematic of Pt<sup>II</sup>(A $\beta$ 16)

### 5.2.1 Computational Details

Molecular Dynamics (MD) and Ligand-Field Molecular Dynamics (LFMD) simulations were carried out using a serial version of DL-POLY Classic[2] and DL-POLY-LF.[3] Pt<sup>II</sup>(A $\beta$ ) complexes were described using a combination of ligand field molecular mechanics (LFMM) for Pt<sup>II</sup> [3, 4] and Assisted Model Building with Energy Refinement 94 (AMBER94)[5] parameters for all other atoms. The A $\beta$ 1-16 peptide was built in an extended conformation in MOE[6] and protonation states at pH of 7.4 assigned using the Protonate3D module of this package. Pt<sup>II</sup>(phenanthroline) complexes were bound to the peptide via His6-N $\epsilon$  and His14-N $\epsilon$ , as identified in previous chapters.[7, 8] Initial peptide conformations for MD and LFMD simulation were selected from a LowMode Molecular Dynamics[9] simulation in MOE. DL-POLY input files were generated using DL-FIELD[10] and the DommiMOE[11] extension to MOE.

For all simulations of the free A $\beta$  peptide, AMBER partial charges were used. For Pt<sup>II</sup>(phen)-A $\beta$  simulations, Merz-Kollman charges were calculated for model Pt(phen)(Me-imidazole)<sub>2</sub> systems from HF/6-31G(d)/SDD electrostatic potential in Gaussian09[12], with Pt<sup>II</sup> given a van der Waals radius of 2.0 Å. The remaining peptide atoms were assigned AMBER94 charges as calculated by MOE. Pt<sup>II</sup>(phen)-A $\beta$  MK charges are summarised in Table 5.1.

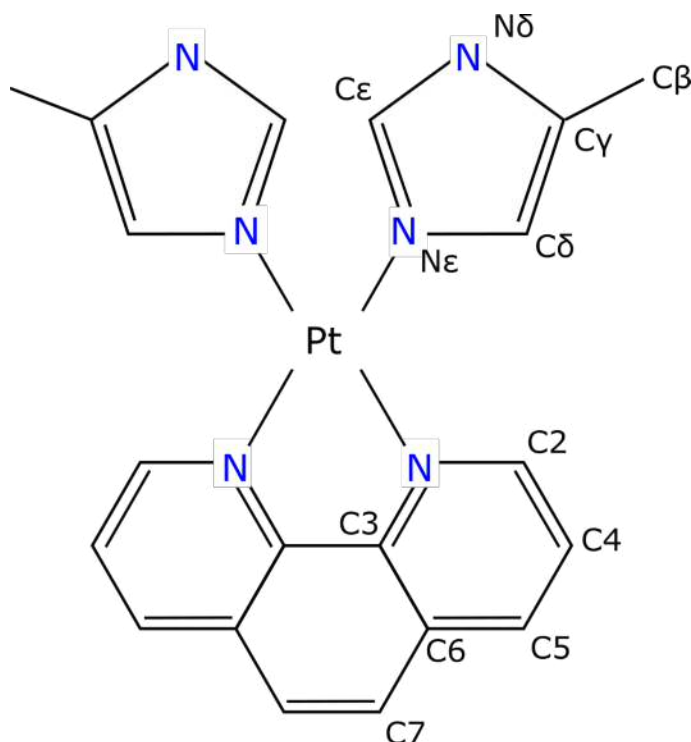


Figure 5.3: Pt<sup>II</sup>(phen)(Me-imid)<sub>2</sub> region used in calculation of Merz-Kollman partial charges.

Metal		His		Phen	
Atom	Charge / $e$	Atom	Charge / $e$	Atom	Charge / $e$
Pt	0.560	N $\epsilon$	0.324	N1	0.025
		C $\epsilon$	-0.140	C2	0.001
		N $\delta$	-0.350	C3	0.134
		C $\delta$	-0.448	C4	-0.262
		C $\gamma$	0.293	C5	-0.017
		C $\beta$	-0.423	C6	0.077
		H $\epsilon$	0.248	C7	-0.219
		H $\delta$ N	0.425	H2	0.190
		H $\delta$ C	0.267	H4	0.217
		H $\beta$ 1	0.160	H5	0.192
		H $\beta$ 2	0.162	H7	0.216

Table 5.1: Merz-Kollman partial charges for Pt<sup>II</sup>(phen)(Me-imid)<sub>2</sub> subsystem of Pt<sup>II</sup>(phen)-A $\beta$ . Atom names are given in Figure 5.3.

In all MD simulations, complexes were modelled as isolated systems, with reaction field solvation and dielectric constant 78.4. Simulations were performed in the NVT ensemble, where temperature was controlled at 310 K using the Nosé-Hoover thermostat[13, 14] with relaxation constant 0.5 ps. Equations of motion were integrated using a Verlet Leapfrog algorithm, with a timestep of 1 fs. The SHAKE algorithm[15] with tolerance  $10^{-8}$  Å was used to constrain bonds containing hydrogen. The vdW forces were calculated with a cutoff of 1 nm, while 2.1 nm was used as a cutoff for electrostatics. A series of other electrostatics cut-offs were tested, but showed little influence on results. In each molecular dynamics trajectory, atomic positions and velocities were recorded every 500 fs and used for subsequent analysis.

Five simulations (labelled A-E) of the free A $\beta$ 16 fragment were carried out, each for 200 ns. An overlay image of initial conformations A-E is shown in Figure 5.4, where the conformations are superposed in such a way as to minimise the RMSD between them. Individual images of these initial conformations are shown in the Appendix.

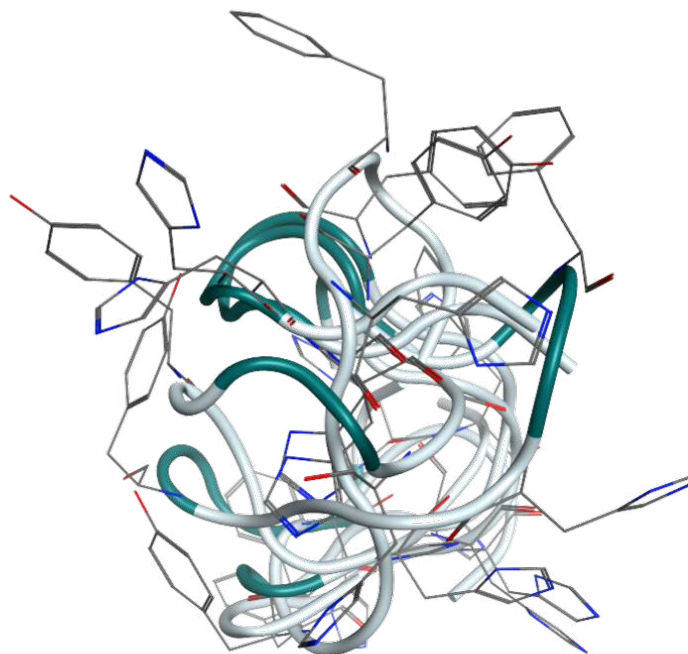


Figure 5.4: Superposed conformations A-E of A $\beta$ 16.

Table 5.2 shows an RMSD matrix of each A $\beta$ 16 starting configuration against all others, illustrating that all initial conformations are significantly different from one another, ensuring more efficient sampling of the molecular phase space during simulations. The most similar conformations are A and B, with RMSD of 2.140 Å, but visual inspection of these structures (Figures 5.4 and A.9) shows that the peptide adopts a notably different structure, with different secondary structure elements present; while conformation A is assigned as purely random coil in nature, B has several turn-type residues, as determined by STRIDE[16], illustrated by the green sections of the peptide ribbon.

RMSD / (Å)	A	B	C	D	E
A	-	2.140	4.420	6.459	8.213
B	-	-	4.310	6.362	8.259
C	-	-	-	5.488	7.261
D	-	-	-	-	5.763
E	-	-	-	-	-

Table 5.2: RMSD matrix of conformations A-E

In addition to this, five identical simulations (labelled F-J) were performed for the

Pt<sup>II</sup>(A $\beta$ 16) system. As for the free A $\beta$ 16 system, an RMSD matrix of these initial conformations was constructed, and is shown in Table 5.3. Here, the conformations with the lowest RMSD are G and I, at 3.301 Å. As before, an overlay image of these conformations, shown in Figure 5.5, illustrates these RMSD values and the different structures being studied. Individual images of the initial configurations are shown in the Appendix.

RMSD / (Å)	F	G	H	I	J
F	-	7.605	6.344	7.485	6.323
G	-	-	6.400	3.301	5.949
H	-	-	-	5.744	5.807
I	-	-	-	-	5.797
J	-	-	-	-	-

Table 5.3: RMSD matrix of conformations F-J.

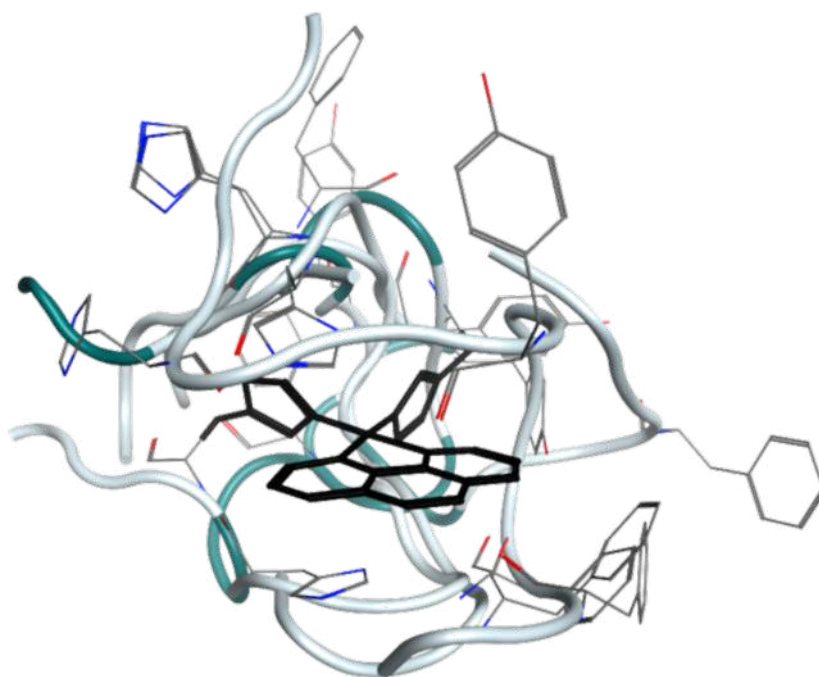


Figure 5.5: Superposed conformations F-J of Pt<sup>II</sup>(A $\beta$ 16).



### 5.2.2 A $\beta$ 16 Equilibration

In all MD simulations, it is important to properly equilibrate the system of interest. However, equilibrating an intrinsically disordered peptide such as A $\beta$  is challenging.

In DL-POLY and DL-POLY-LF, equilibration is controlled by the ‘equilibration’ directive, where a number of timesteps must be specified, in which the system will be brought to the desired physical conditions (temperature, pressure etc.). In this work, we do not consider equilibration to be a stabilisation of the physical conditions; rather, we are interested in equilibration in terms of molecular structure *i.e.* the system reaches a stable configuration. Therefore, if the system is deemed to equilibrate during this ‘equilibration’ phase of the simulation, measures of molecular shape and size (*e.g.* Rg, RMSD) are expected to stabilise. However, this approach proved insufficient to equilibrate these flexible systems; various tests using 10 ps to 10 ns of ‘equilibration’ were run, and while system properties (*e.g.* temperature) stabilised, the peptide remained structurally unstable. Thus, equilibration must be monitored over the course of the simulation.

Work by Huy *et al.*[17] on the A $\beta$ 42 system utilised the concept of pseudo-equilibration *i.e.* stabilisation of molecular RMSD over time indicates that the system MD trajectory has stabilised sufficiently for useful data to be extracted from production MD. Similarly, in this work, systems were considered to be equilibrated when the RMSD fluctuates around a central point by approximately 1 Å, as used by Dong *et al.*[18] for simulations of the A $\beta$ 40 peptide.

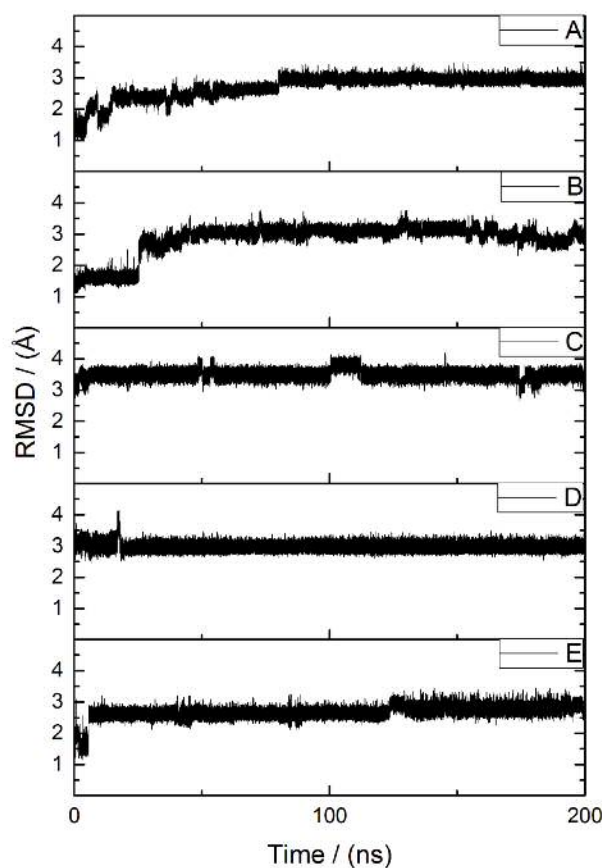


Figure 5.6: RMSD ( $\text{\AA}$ ) vs time for 5 simulations of A $\beta$ 1-16 (A-E).

Here, RMSD values were calculated against the starting structure, generated by LMMD conformational searches in DommiMOE. As Figure 5.6 shows, the A $\beta$ 16 systems reach stable RMSD values relatively quickly, indicating that LMMD produced stable conformations of the peptide as starting points for MD simulations. A $\beta$ 16 conformations reached equilibration at 80 ns, 40 ns, 10 ns, 20 ns and 10 ns for conformations A-E, respectively. Simulation data beyond these points was used for analysis. These plots may be directly compared to those for the Pt<sup>II</sup>(phen)-A $\beta$ 16 system.

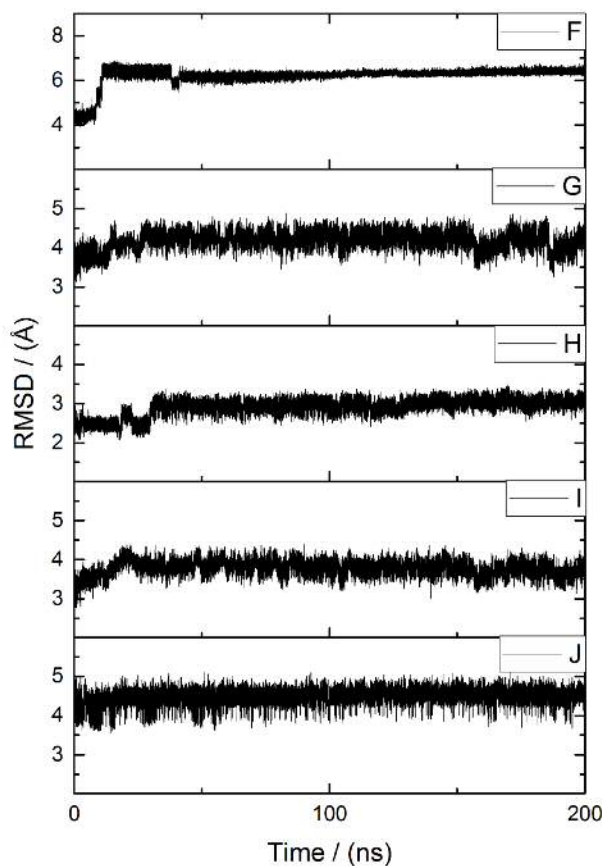


Figure 5.7: RMSD ( $\text{\AA}$ ) vs time for 5 simulations of  $\text{Pt}^{\text{II}}(\text{A}\beta 1\text{-}16)$  (F-J).

Figure 5.7 shows the RMSD of  $\text{Pt}^{\text{II}}(\text{A}\beta 16)$  systems over the course of 200 ns LFMD simulation. Similarly to the free  $\text{A}\beta 16$  systems above, the  $\text{Pt}^{\text{II}}(\text{A}\beta 16)$  systems reach stable RMSD values quickly, and are then considered to be equilibrated. Conformations F-J reach equilibration at 50 ns, 25 ns, 30 ns, 25 ns and 25 ns respectively. As for the  $\text{A}\beta 16$  simulations, trajectories after these points were used for further analysis.

### 5.2.3 $\text{A}\beta 16$ Production MD

Resulting MD trajectories were analysed using VMD[19], with RMSD, radius of gyration, peptide secondary structure, hydrogen bonds, salt bridges, solvent accessible surface area and RMSF data being recorded.

Table 5.4 summarises RMSD data for the five A $\beta$ 16 production runs (A-E), where RMSD values remain relative to the initial configuration. Each of these simulations shows an RMSD value around 3 Å, though conformations C and E are significantly larger and smaller than this value, respectively (*i.e.* more than two standard deviations from the other data points). Based on this standard deviation data, conformation B shows the greatest degree of variance in structure over the course of the simulation, while each of these conformers shows a similar range of calculated RMSD values.

A $\beta$ 16	Ave. RMSD / (Å)	SD	Min.	Max.
A	2.946	0.083	2.535	3.514
B	3.043	0.162	2.299	3.774
C	3.503	0.140	2.671	4.182
D	2.969	0.088	2.549	3.485
E	2.686	0.149	2.090	3.449

Table 5.4: Average, standard deviation, min. and max. RMSD of production MD for A $\beta$ 16 simulations A-E.

The corresponding data for the Pt<sup>II</sup>(A $\beta$ 16) simulations is shown in Table 5.5, where RMSD values are once again measured versus the initial molecular configuration. Here, RMSD values tend to be higher than for the free peptide, indicating that the structures sampled during Pt<sup>II</sup>(A $\beta$ 16) simulations deviate more from the initial conformations produced by LMMD than the corresponding free peptide simulations. The standard deviation data for the Pt<sup>II</sup>(A $\beta$ 16) simulations is broadly similar to that of the free peptide, suggesting that these structures fluctuate to a similar degree. In addition, Pt<sup>II</sup>(A $\beta$ 16) simulations also display similar range of RMSD data as the free peptide.

Pt <sup>II</sup> (A $\beta$ 16)	Ave. RMSD / ( $\text{\AA}$ )	SD	Min.	Max.
F	6.292	0.101	5.680	6.777
G	4.225	0.192	3.251	4.887
H	2.984	0.135	2.413	3.471
I	3.794	0.166	3.013	4.382
J	4.501	0.176	3.628	5.092

Table 5.5: Average, standard deviation, min. and max. RMSD of production MD for Pt<sup>II</sup>(A $\beta$ 16) simulations F-J.

While RMSD plots illustrate the overall degree of change in structure over time, molecular radius of gyration provides more information regarding molecular shape over time. Figure 5.8 shows radius of gyration (Rg) data for the A $\beta$ 16 simulations. These Rg values are summarised in Table 5.6.

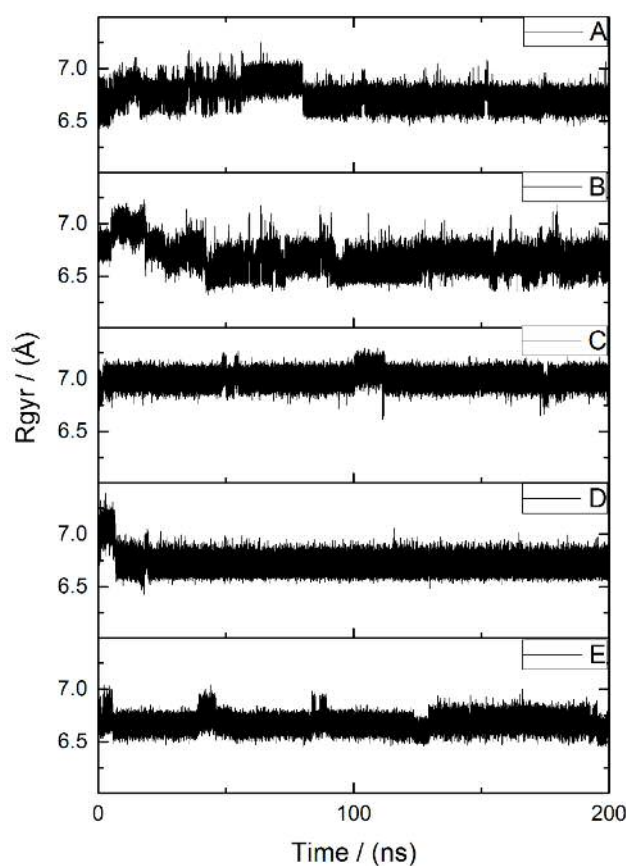


Figure 5.8: Rg ( $\text{\AA}$ ) vs time for 5 simulations of A $\beta$ 1-16 (A-E).

A $\beta$ 16	Ave. R <sub>g</sub> / (Å)	SD	Min.	Max.
A	6.695	0.060	6.456	7.084
B	6.627	0.095	6.324	7.177
C	6.998	0.057	6.609	7.292
D	6.706	0.051	6.484	7.060
E	6.675	0.058	6.450	7.043

Table 5.6: Average, standard deviation, min. and max. R<sub>g</sub> of production MD for A $\beta$ 16 simulations A-E.

These data show that the free A $\beta$ 16 fragment adopts a relatively compact structure, with R<sub>g</sub> values typically less than 7 Å. For comparison, the A $\beta$ 16 peptide in an extended conformation has R<sub>g</sub> of 16.99 Å, and for an  $\alpha$ -helical structure has an R<sub>g</sub> of 9.16 Å. The standard deviation within this data is also small, indicating that there is little change in the compactness of the peptide structure during simulation, suggesting that it is more favourable for the peptide to form stabilising intra-molecular interactions than to interact with the reaction-field solvent. Further discussion of these intra-molecular interactions is included later in the text.

In general, these systems display a similar range of R<sub>g</sub> values; this suggests that while the peptide as a whole remains compact, flexible peptide side-chains may be responsible for the similar degree of variation observed. However, simulation C shows a R<sub>g</sub> value of  $6.998 \pm 0.06$  Å that is significantly different to the other simulations (*i.e.* more than two standard deviations from other data). Interestingly, simulation C also displayed the greatest average RMSD value of the A $\beta$ 16 systems studied (see Table 5.4). The significantly different R<sub>g</sub> suggests that this conformer occupies a noticeably different configuration than A, B, D and E during the course of the simulation. Equivalent R<sub>g</sub> data for the Pt<sup>II</sup>(A $\beta$ 16) simulations is shown in Figure 5.9 and summarised in Table 5.7.

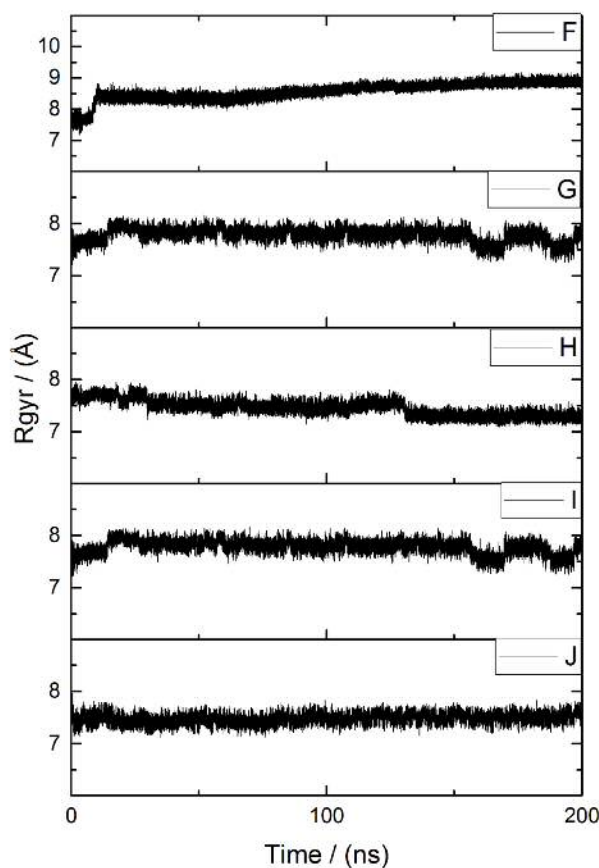


Figure 5.9:  $R_g$  (Å) vs time for 5 simulations of  $\text{Pt}^{\text{II}}(\text{A}\beta 1-16)$  (F-J).

$\text{Pt}^{\text{II}}(\text{A}\beta 16)$	Ave. $R_g$ / (Å)	SD	Min.	Max.
F	8.659	0.196	7.932	9.173
G	7.786	0.125	7.252	8.165
H	7.417	0.126	7.085	7.812
I	7.786	0.125	7.252	8.156
J	7.481	0.095	7.121	7.831

Table 5.7: Average, standard deviation, min. and max.  $R_g$  of production MD for  $\text{Pt}^{\text{II}}(\text{A}\beta 16)$  simulations F-J.

In contrast, the  $\text{Pt}^{\text{II}}(\text{A}\beta 16)$  simulations display significantly larger  $R_g$  values than their free peptide counterparts - with values generally exceeding 7.4 Å- illustrating that coordination of the large, sterically demanding  $\text{Pt}^{\text{II}}(\text{phenanthroline})$  system to the peptide forces the structure to adopt different, less compact conformations. Standard deviation data is consistent across these simulations, though slightly larger

than for the free peptide systems.

The data in Table 5.7 show that the  $\text{Pt}^{\text{II}}(\text{A}\beta 16)$  simulations are centred around 3 distinct  $R_g$  values, approximately 7.45 Å, 7.78 Å and 8.66 Å. Simulation F displays the largest  $R_g$  of all  $\text{Pt}^{\text{II}}(\text{A}\beta 16)$  systems ( $8.66 \pm 0.2$  Å), significantly larger than all other simulations examined. The  $\text{Pt}^{\text{II}}(\text{phen})$  unit itself has  $R_g$  of 2.29 Å, so the  $\text{Pt}^{\text{II}}(\text{A}\beta 16)$  adducts are smaller than the sum of  $\text{Pt}^{\text{II}}(\text{phen})$  plus peptide. Interestingly, simulation F displays a high prevalence of stabilising  $\pi$ - $\pi$  interactions, which may cause the peptide backbone to extend and re-organise in order to accommodate these stacking arrangements. These  $\pi$ - $\pi$  interactions are discussed in detail later.

Interestingly, data in Figure 5.9 and Table 5.7 illustrate that conformations G and I display incredibly similar  $R_g$  values. Here, average  $R_g$  values are identical to 4 decimal places, with approximately the same degree of similarity in minimum and standard deviation of  $R_g$  values. Comparison of snapshots of the simulated structures at time = 200 ns indicates that not only is the  $R_g$  identical, but that the structures are near-identical as well, as judged by RMSD calculations. These RMSD values are shown in Table 5.8 and snapshots are overlaid in Figure 5.10. This means that despite the simulations starting from distinctly different conformations, the simulations have converged to a common structure within the equilibration period.

Conformations compared	RMSD at 0 ns / (Å)	RMSD at 200 ns / (Å)
G vs. I	3.301	0.106

Table 5.8: Comparison of RMSD for simulations G and I of the  $\text{Pt}^{\text{II}}(\text{A}\beta 16)$  system at time = 0 ns and time = 200 ns.



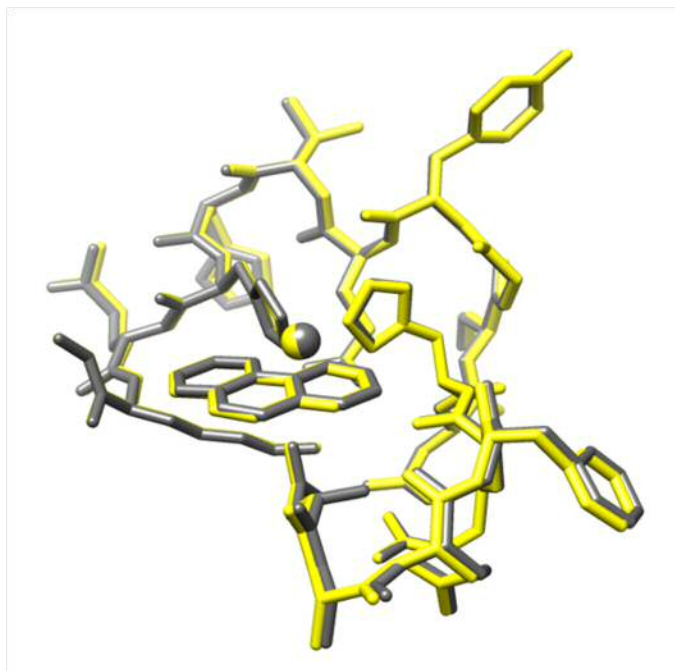


Figure 5.10: Superposition of snapshots at time = 200 ns of simulations G and I of the  $\text{Pt}^{\text{II}}(\text{A}\beta 16)$  system.

Next, the root mean square fluctuation (RMSF) of each residue was measured during simulations using VMD. Figure 5.11 and Figure 5.12 show the RMSF values by residue of trajectories A-E and F-J, respectively. In addition, plots of RMSF values by residue, averaged across simulations for the  $\text{A}\beta 16$  and  $\text{Pt}^{\text{II}}(\text{A}\beta 16)$  simulations are shown in the Appendix.

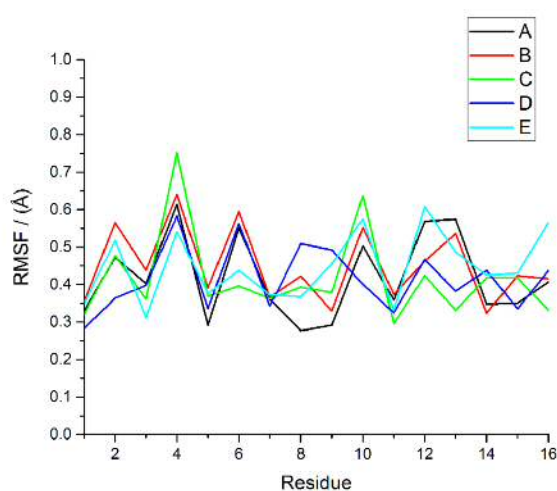


Figure 5.11: RMSF per residue for simulations A-E of  $\text{A}\beta 16$

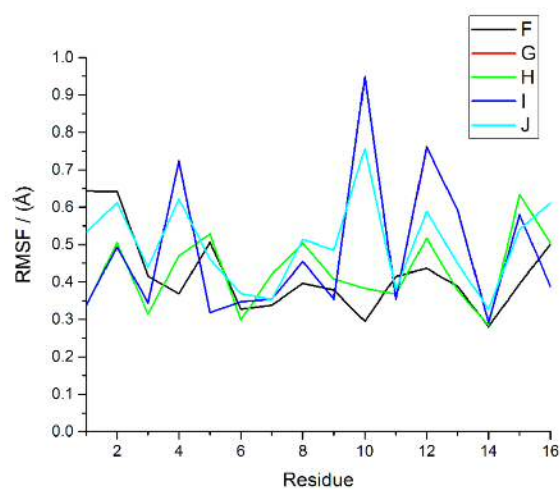


Figure 5.12: RMSF per residue for simulations F-J of  $\text{Pt}^{\text{II}}(\text{A}\beta_{16})$ . Data for simulation G is obscured by simulation I.

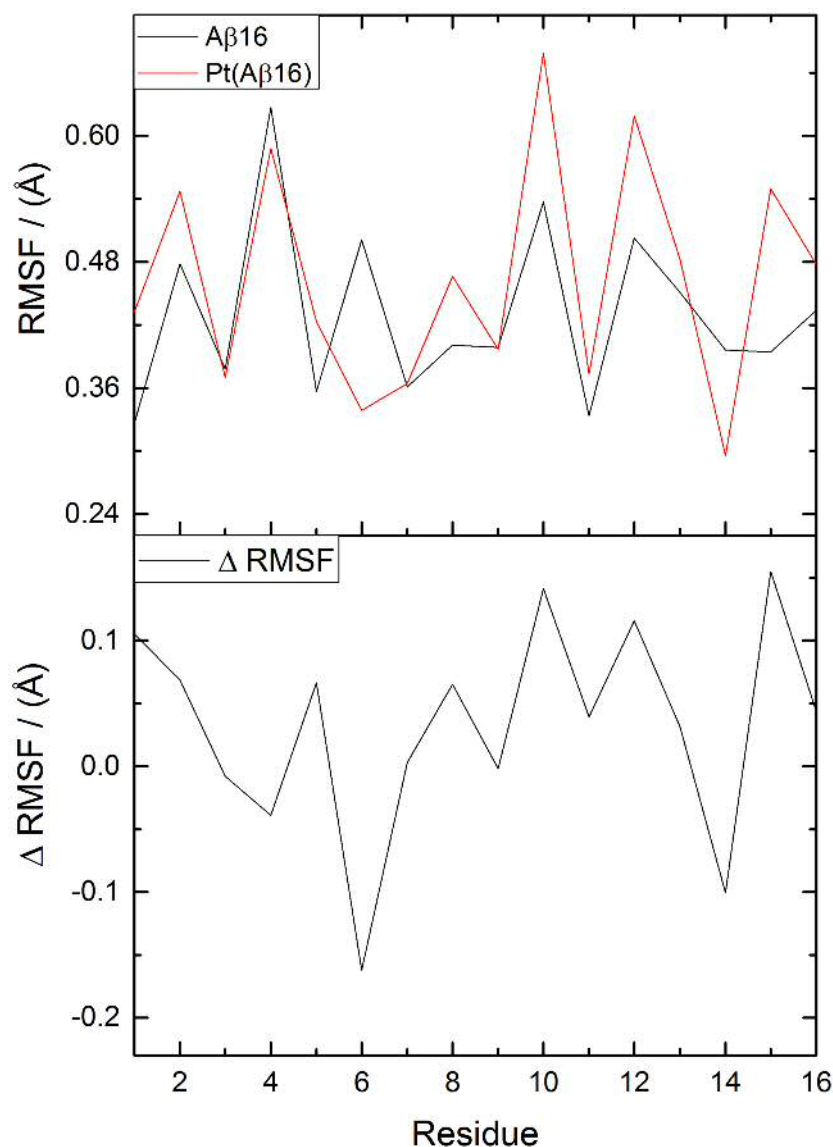


Figure 5.13: Top: Comparison of RMSF per residue for Aβ16 and Pt<sup>II</sup>(Aβ16) simulations. Bottom: Difference in RMSF between Aβ16 and Pt<sup>II</sup>(Aβ16) simulations

Figure 5.13 shows a comparison plot of RMSF data for the Aβ16 and Pt<sup>II</sup>(Aβ16) simulations, as well as a plot of ΔRMSF *i.e.* the difference in peptide RMSF caused by coordination of the Pt<sup>II</sup>(phenanthroline) complex. Here, negative values correspond to the Pt<sup>II</sup>(Aβ16) system displaying lower RMSF than Aβ16. For the free peptide, large RMSF values are observed for Phe4, His6, Tyr10 and Val12, and small values for Asp1, Glu3, Arg5, and Glu11. Coordination of Pt<sup>II</sup>(phen) at histidines 6 and 14 unsurprisingly reduces their RMSF, with Phe4 also moving less than in the

metal-free case, whereas values for Tyr10, Val12 and Lys16 are on average larger after platinum coordination. These data therefore suggests that coordination of  $\text{Pt}^{\text{II}}(\text{phen})$  affects the peptide in more subtle ways than might first be thought, in particular promoting more flexibility in residues lying between coordination sites. However, differences between free and  $\text{Pt}^{\text{II}}(\text{phen})$  trajectories are of similar magnitude to those between repeat simulations. RMSF for Tyr10, for instance, is relatively constant at  $0.53 \pm 0.09 \text{ \AA}$  over five simulations of the free peptide, but varies rather more ( $0.67 \pm 0.31 \text{ \AA}$ ,  $\text{max} = 0.95 \text{ \AA}$ ,  $\text{min} = 0.29 \text{ \AA}$ ) for the equivalent  $\text{Pt}^{\text{II}}(\text{phen})$  trajectories. Only three residues (His6, His14 and Gln15) exhibit differences in RMSF that exceed the sum of standard deviations - full details can be found in the Appendix.

The clear increase in the  $R_g$  of the peptide when bound to  $\text{Pt}^{\text{II}}(\text{phenanthroline})$  is consistent with the increased solvent accessible surface area (SASA) observed, relative to the free peptide. SASA is calculated by a Monte Carlo sampling procedure, using 50,000 sample points to estimate the accessible surface area.

Average SASA per residue for simulations A-E of A $\beta$ 16 is shown in Figure 5.14, while Figure 5.15 shows the corresponding data for conformations F-J of the  $\text{Pt}^{\text{II}}(\text{A}\beta 16)$  systems. Data for each individual simulation is shown in the Appendix. Each of the simulations A-E show similar values and trends in SASA data by residue, indicating that while each residue has a different solvent accessible surface area, the peptide SASA does not vary drastically with conformation; A similar pattern is seen for conformations F-J.

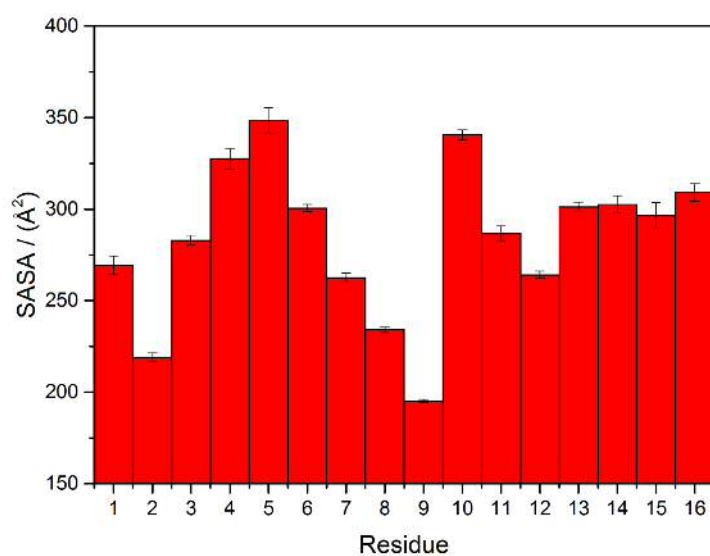
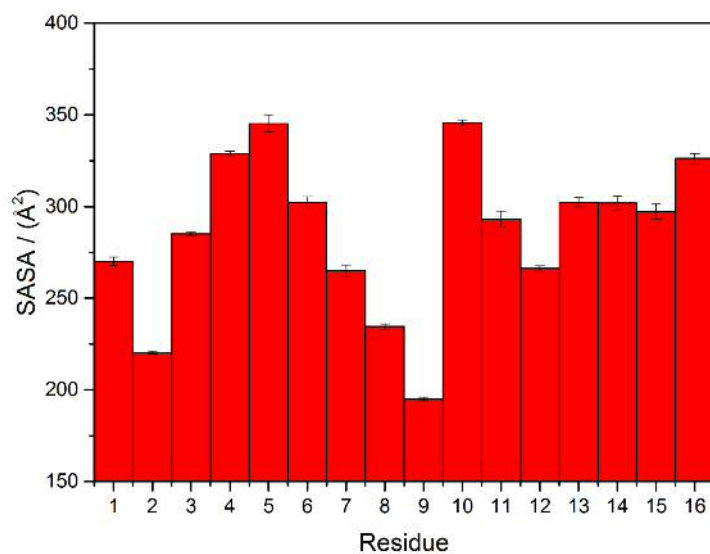


Figure 5.14: Average SASA per residue of Aβ16 simulations A-E

Figure 5.15: Average SASA per residue of Pt<sup>II</sup>(Aβ16) simulations F-J.

The SASA data corresponding to Figures 5.14-5.15 are reported in Table 5.9. In particular, this data shows that the most solvent-exposed residues in the Aβ16 systems are Arg5, Tyr10 and Phe4 - the largest residues - while the least solvent-exposed residues are Ser8, Ala2 and Gly9. For the Pt<sup>II</sup>(Aβ16) simulations, residues Arg5, Tyr10 and Phe4 remain the most solvent exposed while Ser8, Ala2 and Gly9 remain the least solvent exposed. The Pt<sup>II</sup>(phen) fragment itself has a SASA of  $240.3 \pm$

9.6 Å<sup>2</sup>. Data in Table 5.9 illustrates that for the sixteen-residue peptide, SASA is governed more by the two-dimensional chemistry rather than the three-dimensional peptide conformation *i.e.* peptide primary structure is more important than peptide secondary structure.

Residue	Aβ16 SASA / (Å <sup>2</sup> )	Pt <sup>II</sup> (Aβ16) SASA / (Å <sup>2</sup> )
1	269.5 ± 5.18	270.1 ± 2.35
2	219.0 ± 2.46	220.3 ± 0.86
3	283.1 ± 2.67	285.3 ± 0.89
4	327.6 ± 5.70	328.9 ± 1.38
5	348.4 ± 7.03	345.4 ± 4.80
6	300.6 ± 2.03	302.4 ± 3.21
7	262.6 ± 2.61	265.1 ± 3.06
8	234.1 ± 1.52	234.5 ± 1.49
9	195.3 ± 0.77	195.1 ± 0.94
10	340.7 ± 2.81	345.7 ± 1.58
11	286.8 ± 3.92	293.3 ± 4.17
12	264.2 ± 1.94	266.3 ± 1.51
13	301.4 ± 2.40	302.3 ± 2.46
14	302.6 ± 4.87	302.0 ± 4.07
15	296.5 ± 7.08	297.3 ± 3.99
16	309.2 ± 4.78	326.3 ± 2.82

Table 5.9: SASA data per residue for Aβ16 and Pt<sup>II</sup>(Aβ16) simulations

Total average SASA values for the whole Aβ16 and Pt<sup>II</sup>(Aβ16) systems are summarised in Table 5.10. These data show that while the Pt<sup>II</sup>(Aβ16) systems display greater total SASA than the free Aβ16 systems, the difference between them is largely accounted for by the large Pt<sup>II</sup>(phenanthroline) complex, rather than any major changes in peptide conformation upon metal binding. Here, the difference between Aβ16 and Pt<sup>II</sup>(Aβ16) SASA ranges from 265-307 Å<sup>2</sup>, meaning that there is a small increase in overall SASA of the peptide.

A $\beta$ 16	A	B	C	D	E
SASA / ( $\text{\AA}^2$ )	4544.6	4524.3	4537.0	4550.5	4550.0
Pt <sup>II</sup> (A $\beta$ 16)	F	G	H	I	J
SASA / ( $\text{\AA}^2$ )	4831.3	4831.6	4810.12	4831.53	4815.52

Table 5.10: Total SASA for A $\beta$ 16 and Pt<sup>II</sup>(A $\beta$ 16) simulations A-E and F-J, respectively

In addition, the effect of Pt<sup>II</sup>(phenanthroline) coordination on the peptide secondary structure was considered. While A $\beta$  is widely considered to be an intrinsically disordered peptide, little work has been done to determine the effects of Pt<sup>II</sup> coordination on this structure, though it is thought that platinum coordination promotes amorphous aggregation, rather than fibril formation.[20] The STRIDE algorithm[16], as implemented in VMD, was used for secondary structure calculation. This characterised each residue as either Turn,  $\beta$ -sheet,  $\beta$ -bridge,  $\alpha$ -helix, 3,10-helix,  $\pi$ -helix or Coil- type structure.

Total percentage secondary structure for simulations A-J is shown in Table 5.11. In addition, Figure 5.16 and Figure 5.17 show plots of the peptide secondary structure, per residue, over the course of simulations i) A-E and ii) F-J for the A $\beta$ 16 and Pt<sup>II</sup>(A $\beta$ 16) systems, respectively.

		Turn	$\beta$ -sheet	Bridge	$\alpha$ -helix	3,10-helix	$\pi$ -helix	Coil
A $\beta$ 16	A	73.52	0.00	0.07	0.00	4.69	0.00	21.72
	B	67.77	0.00	2.57	0.00	0.73	0.03	28.90
	C	77.21	0.00	0.00	0.00	0.00	0.00	22.79
	D	66.20	0.00	0.00	0.00	0.00	0.00	33.80
	E	59.09	0.00	0.00	0.00	0.00	0.00	41.91
Pt <sup>II</sup> (A $\beta$ 16)	F	41.62	0.00	0.00	0.00	16.35	0.00	42.03
	G	82.66	0.00	0.00	0.00	0.00	0.00	17.34
	H	55.90	0.00	0.00	0.00	0.00	0.00	44.10
	I	82.68	0.00	0.00	0.00	0.00	0.00	17.32
	J	53.27	0.00	0.00	0.00	0.00	0.00	46.73

Table 5.11: Total percentage secondary structure for A $\beta$ 16 and Pt<sup>II</sup>(A $\beta$ 16) simulations

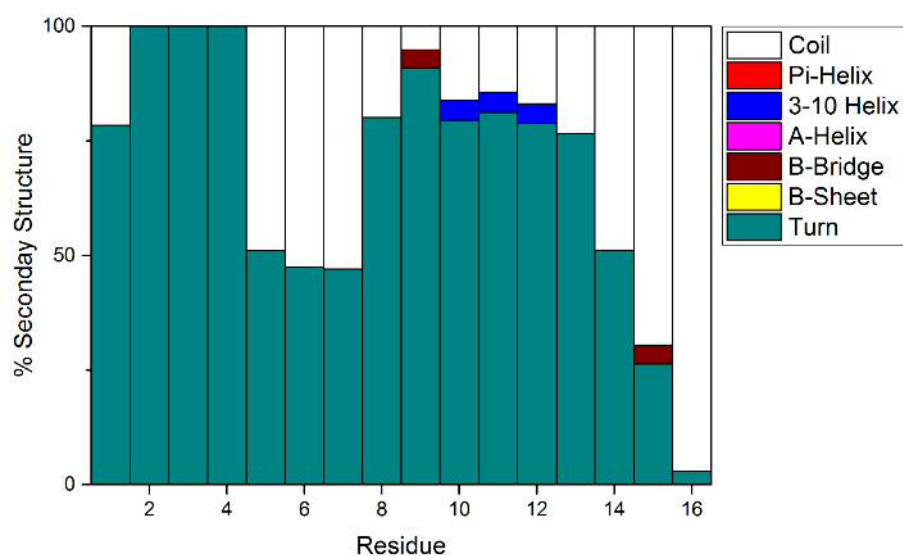


Figure 5.16: Percentage secondary structure by residue for A $\beta$ 16 simulation. Secondary structure assigned using STRIDE algorithm.



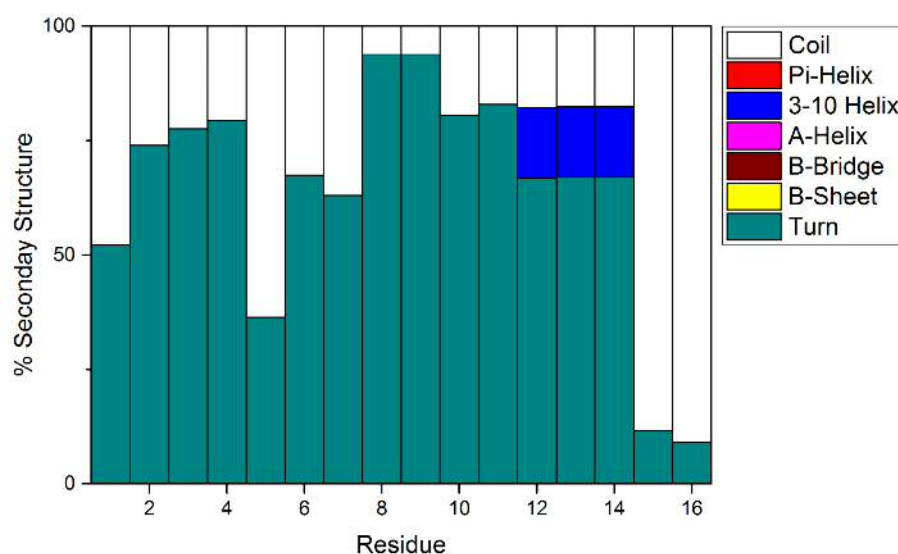


Figure 5.17: Percentage secondary structure by residue for Pt<sup>II</sup>(Aβ<sub>16</sub>) simulation. Secondary structure assigned using STRIDE algorithm.

As expected for an intrinsically disordered peptide, the majority of the residues throughout trajectories A-E exhibit coil or turn conformations. Interestingly, there is a low propensity (4% of total simulation time) for residues Tyr10-Glu11-Val12 to form a minimal (3 residue) 3,10 helix, while residues Gly9 and Gln15 infrequently (4% of simulation) adopt β-bridge (*i.e.* hairpin) structures. The prevalence of turn and coil structures in these simulations is in good general agreement with most data for Aβ, though the propensity of residues to adopt turn structures is surprisingly high. Other authors have noted that STRIDE has a greater tendency to assign turn structure than other secondary structure programs[21], but observed that most residues in Aβ exist as turn structure with approximately 30-80% probability. In addition, their simulations of Aβ<sub>42</sub>[21] displayed 3,10-helical structure for residues Tyr10-Val12 of approximately 3-5%, in agreement with this data, though they observed similar propensity throughout the peptide.

In the Pt<sup>II</sup>(Aβ<sub>16</sub>) simulations, secondary structure is again predominantly assigned as either turn- or coil-type conformations, but the percentage of each secondary structure element is different from the free peptide system. This indicates that while coordination of the Pt<sup>II</sup>(phenanthroline) complex does not drastically change the secondary structure of this N-terminal fragment of the Aβ peptide, it induces subtle effects on peptide backbone structure. Here, residues Val12-His13-His14 adopt a

3,10 helical structure for 15% of total simulation time; coordination of the platinum complex not only increases the incidence of this well-defined secondary structure, but also shifts the helix towards the C-terminus relative to the free peptide. Strikingly, residues involved in metal binding (His6, His14) show an increase in defined secondary structure (*i.e.* turn or helix) in the Pt<sup>II</sup>(A $\beta$ 16) simulations compared to the natural peptide, from approximately 50% in A $\beta$ 16 to 60-80% in the platinated system.

Furthermore, the formation of hydrogen bond networks was also studied using VMD, as hydrogen bonds not only link in to the secondary structure data above, but provide further insight into the differences in intramolecular interactions between the free A $\beta$ 16 peptide and Pt<sup>II</sup>(A $\beta$ 16) simulations.

The total number of hydrogen bonds present per frame of the A $\beta$ 16 and Pt<sup>II</sup>(A $\beta$ 16) simulations was calculated and data presented in Table 5.12 and Table 5.13. In general, the Pt<sup>II</sup>(A $\beta$ 16) systems consistently show fewer hydrogen bonds than the free A $\beta$ 16 simulations, suggesting that the Pt<sup>II</sup>(phenanthroline) system may interfere with the natural hydrogen bonding patterns of A $\beta$ 16. Further, the maximum number of hydrogen bonds in the Pt<sup>II</sup>(A $\beta$ 16) systems never reaches the maximum observed in the free A $\beta$ 16 simulations. However, this difference was not significant, as the large standard deviation values (*ca.* 1.9-2.0) suggest that the hydrogen bonds are transient. Surprisingly, this data shows that some frames of all A $\beta$ 16 and Pt<sup>II</sup>(A $\beta$ 16) simulations display no hydrogen bonds. In addition, there is no significant difference in the number of hydrogen bonds between all of the free A $\beta$ 16 systems or all of the Pt<sup>II</sup>(A $\beta$ 16) systems, suggesting that peptide conformation has little effect on the ability of the peptide to form these intramolecular interaction networks.

A $\beta$ 16	Ave. # HB	SD	Min.	Max.
A	6.4	2.0	0	16
B	5.6	1.9	0	14
C	7.1	1.9	0	17
D	7.1	2.1	0	17
E	5.8	1.8	0	15

Table 5.12: Average, SD, min. and max. values of the number of hydrogen bonds present in A $\beta$ 16 simulation.

Pt <sup>II</sup> (A $\beta$ 16)	Ave. # HB	SD	Min.	Max.
F	5.7	1.7	0	13
G	5.2	1.7	0	13
H	3.9	1.6	0	12
I	5.2	1.7	0	12
J	6.0	1.8	0	13

Table 5.13: Average, SD, min. and max. values of the number of hydrogen bonds present in Pt<sup>II</sup>(A $\beta$ 16) simulation.

Finally, intramolecular salt bridge interactions were monitored during the course of the simulations. Salt bridges are known to be important in the formation of fibrils, so any disruption of this network by the Pt<sup>II</sup>(L) system may provide an explanation as to how these agents disrupt the A $\beta$  aggregation process. As the A $\beta$ 16 peptide fragment has two positively charged and four negatively charged residues, there are a total of eight possible salt bridge interactions. These salt bridge interactions were measured in VMD, and are defined as any contact of less than 3.2 Å between O-N atoms in charged residues.

The salt bridge data for the A $\beta$ 16 simulations is summarised in Table 5.14 and presented graphically in Figure 5.18, while equivalent data for the Pt<sup>II</sup>(A $\beta$ 16) simulations may be found in Table 5.15 and Figure 5.19.

A $\beta$ 16	Asp1	Glu3	Asp7	Glu11
Arg5	20.99	25.93	85.64	55.88
Lys16	55.90	21.39	0.00	28.57

Table 5.14: Combined salt-bridge data (% presence) for A $\beta$ 16 simulations

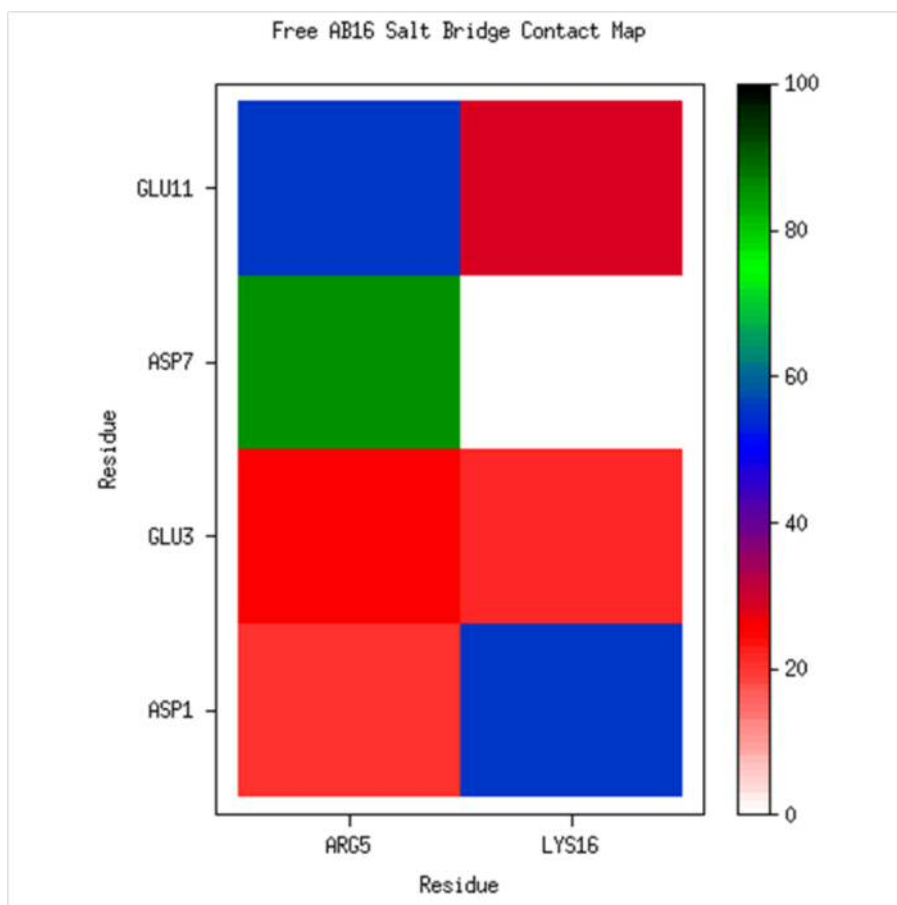


Figure 5.18: Salt bridge contact map for Aβ16 simulation

Pt <sup>II</sup> (Aβ16)	Asp1	Glu3	Asp7	Glu11
Arg5	77.62	79.71	61.26	0.00
Lys16	39.57	0.00	61.91	99.44

Table 5.15: Combined salt-bridge data (% presence) for Pt<sup>II</sup>(Aβ16) simulations

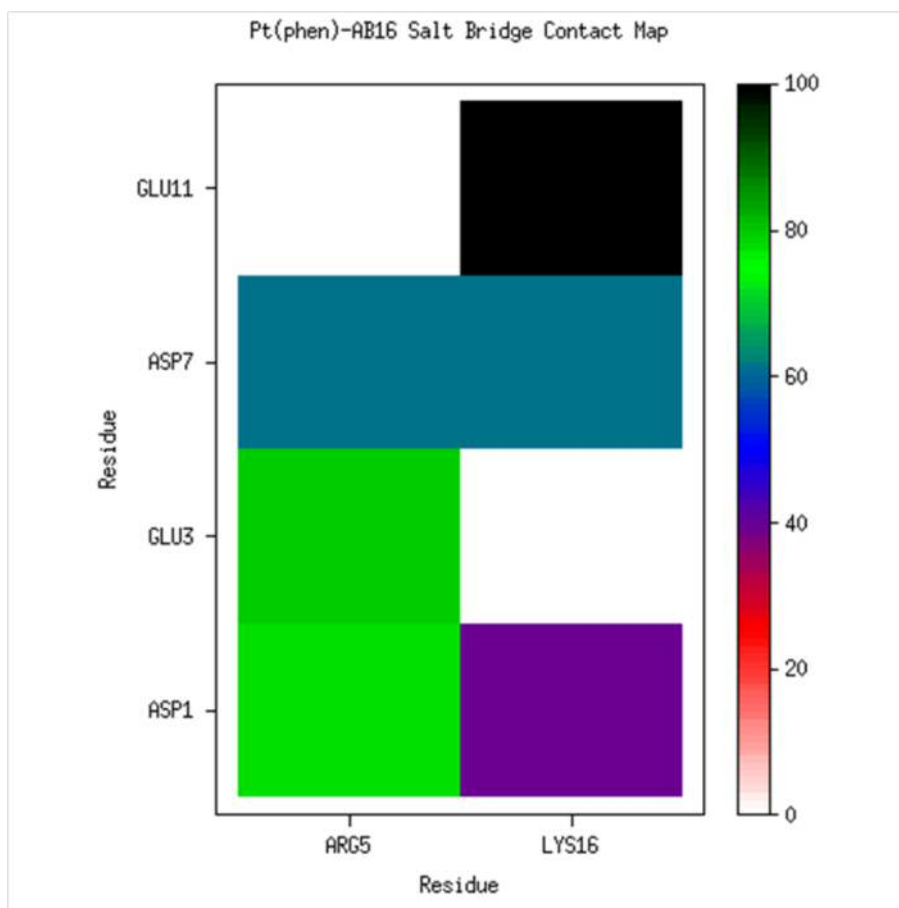


Figure 5.19: Salt bridge contact map for Pt<sup>II</sup>(Aβ16) simulation

In the free Aβ16 systems, there are a large number of salt bridges present for only low fractions of the total simulation time. Here, Arg5 predominantly interacts with Asp7 (86% of frames), but also forms frequent interactions with Glu11 (56%), towards the C-terminus of this peptide fragment. This Arg residue also forms salt bridges less frequently with residues closer to the N-terminus, such as Glu3 and Asp1. Similarly, Lys16, found at the C-terminus of this fragment, forms salt bridges most frequently with Asp1 (56%), at the N-terminus, though it also interacts with Glu3 and Glu11 for 20-30% of total simulation time. In all Aβ16 simulations, Lys16 does not form salt bridge interactions with Asp7. The three most common salt-bridges in Aβ16 simulation are shown in Figure 5.20.

Binding the Pt<sup>II</sup>(phenanthroline) to this peptide fragment induces clear changes in salt bridge structure, as shown in Figure 5.19. In the Pt<sup>II</sup>(Aβ16) system, Arg5 still interacts with Asp7 but with reduced frequency (61%), and instead primarily interacts with Glu3 and Asp1 (*ca.* 80%). In contrast to the free peptide, Pt<sup>II</sup>(Aβ16) simulations show no formation of Arg5-Glu11 salt bridges. In addition, Lys16 forms

a near-constant salt bridge with Glu11, as well as frequent interactions with Asp7 (62% of simulation) and Asp1 (40%). In contrast to the free peptide, Lys16 does not form a salt bridge with Glu3 in the  $\text{Pt}^{\text{II}}(\text{A}\beta 16)$  simulations. The three most common salt-bridges in  $\text{Pt}(\text{A}\beta 16)$  simulation are shown in Figure 5.21.

In general, coordination of the  $\text{Pt}^{\text{II}}(\text{phenanthroline})$  unit to  $\text{A}\beta 16$  causes Arg5 salt bridges to switch between primarily interacting with Asp7 and Glu11 in the free peptide to Glu3 and Asp1, while Lys16 salt bridges are formed with Glu11 and Asp7 instead of Asp1. It is interesting that the combined frequency of the observed salt bridge interactions exceeds 100%, indicating that residues are close enough to their charged partners to be able to form multiple salt bridge-type interactions simultaneously.

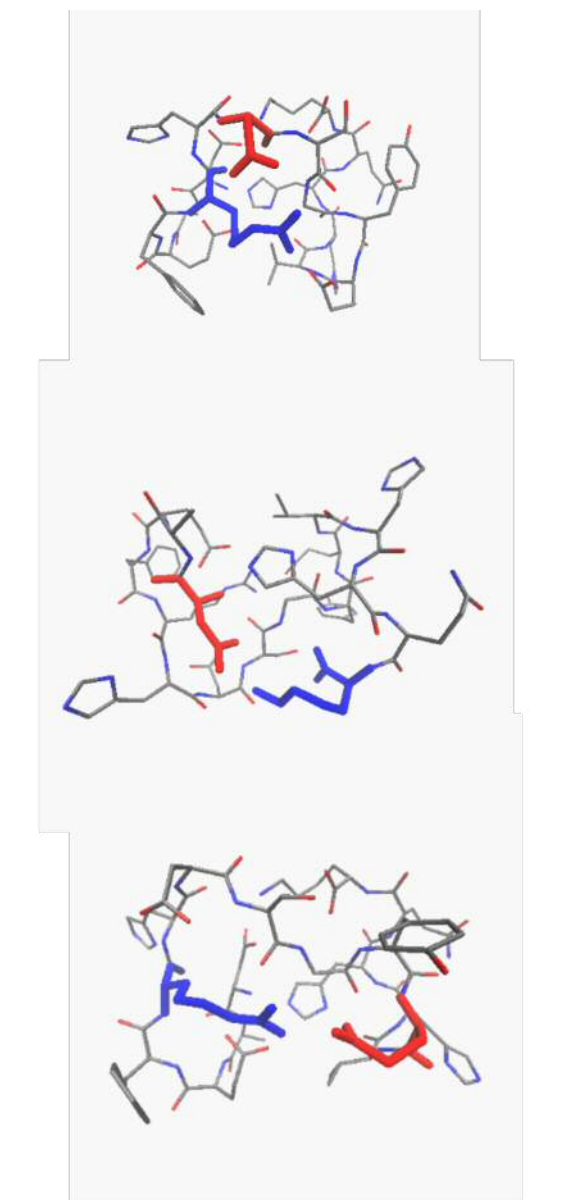


Figure 5.20: Images of the three most prominent salt-bridges in A $\beta$ 16 simulation. From top to bottom: Arg5-Asp7, Asp1-Lys16, Arg5-Glu11. Negatively charged residues are highlighted in blue; positively charged residues are highlighted in red.

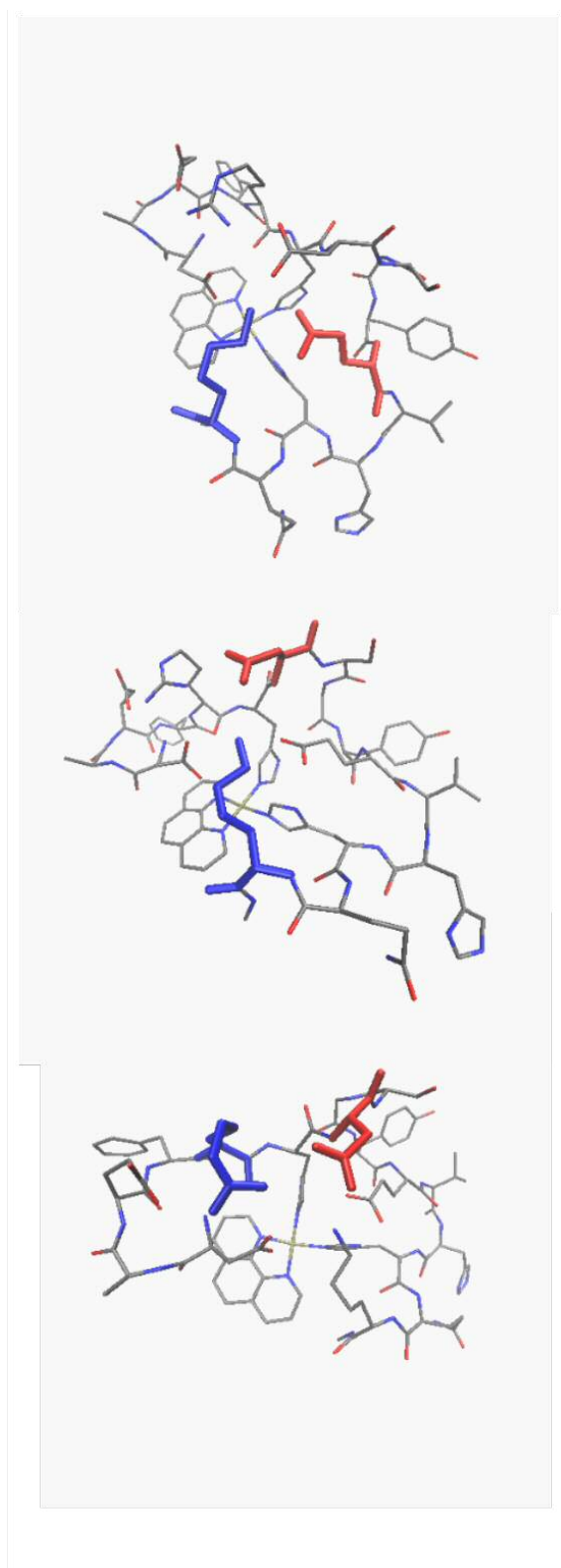


Figure 5.21: Images of the three most prominent salt-bridges in  $\text{Pt}^{\text{II}}(\text{A}\beta_{16})$  simulation. From top to bottom: Lys16-Glu11, Lys16-Asp7, Arg5-Asp7. Negatively charged residues are highlighted in blue; positively charged residues are highlighted in red.



The formation of  $\pi$ - $\pi$  stacking interactions between the ligand and aromatic residues in A $\beta$  are thought to be important in targeting the Pt<sup>II</sup>(L) drugs to the N-terminus of A $\beta$ , as well as stabilising the Pt<sup>II</sup>(L)(A $\beta$ ) adducts.

Here, the contact distance between the phenanthroline ligand and each of the aromatic residues in A $\beta$ 16 that could form these stacking interactions (*i.e.* Phe4, Tyr10 and His13) was monitored during Pt<sup>II</sup>(A $\beta$ 16) simulations. This contact was defined as the distance between C $\gamma$  in the aromatic residue and a carbon atom in the centre of the ligand. Histograms showing the distribution of these inter-plane contacts are shown in Figure 5.22.

This revealed that Phe4 in particular forms frequent  $\pi$ - $\pi$  interactions with the phenanthroline ligand, with a high distribution of contacts  $< 5$  Å (see plot A in Figure 5.22). Tyr10 forms almost no  $\pi$ - $\pi$  interactions with the phenanthroline ligand, with only a small distribution of frames at distances of approximately 5-6 Å, while His13 shows moderate distribution of states where the  $\pi$ - $\pi$  inter-plane distance is less than 5 Å, sufficient for weak  $\pi$ - $\pi$  interaction (see plots B and C in Figure 5.22, respectively). At short  $\pi$ - $\pi$  distances, the aromatic planes are essentially parallel - see Figure 5.23 - while larger distances ( $> 6$  Å) are not important for  $\pi$ - $\pi$  interactions, and so the angles between planes is not measured.

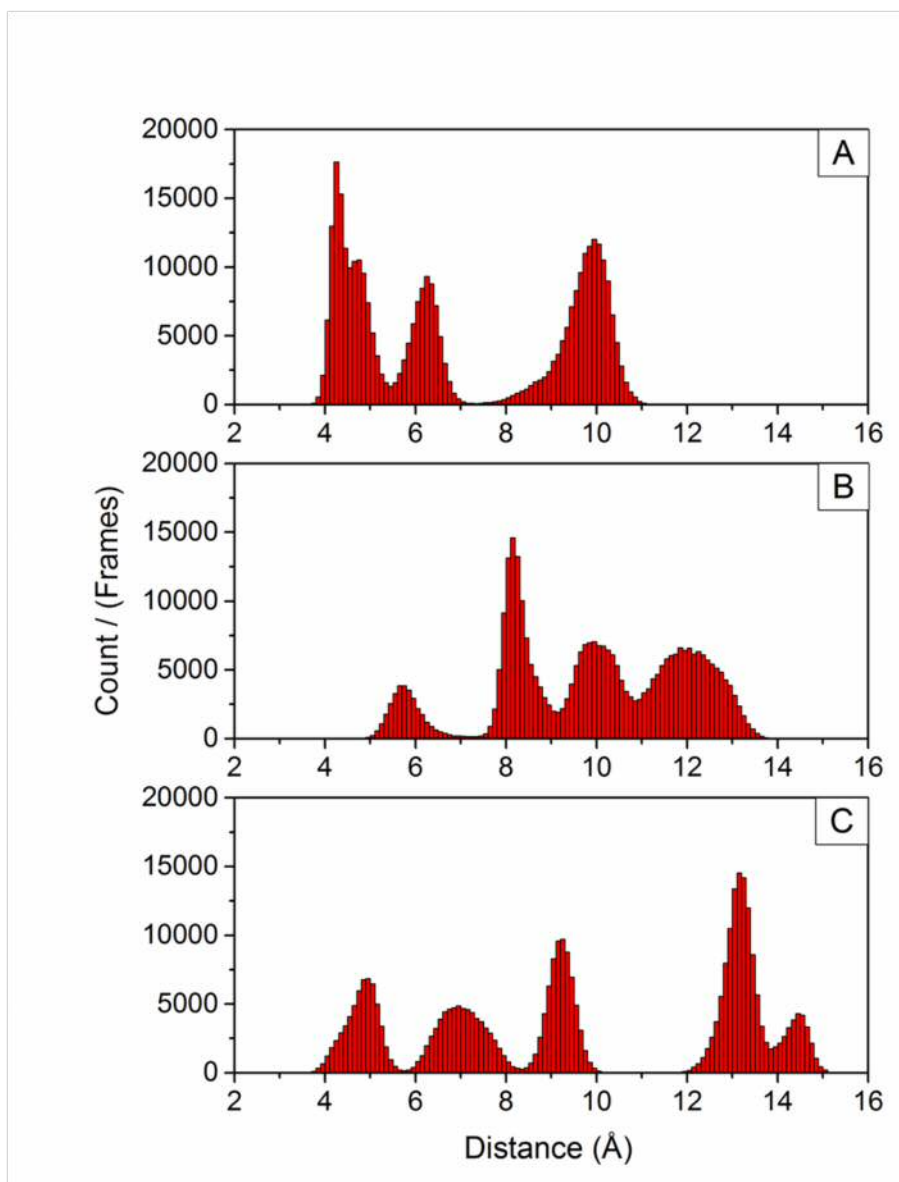


Figure 5.22:  $\pi$ - $\pi$  contact distance for A) Phe4, B) Tyr10 and C) His13 residues with the phenanthroline ligand in  $\text{Pt}^{\text{II}}(\text{A}\beta 16)$  simulation

Of particular interest is simulation F, where both Phe4 and His13 simultaneously form  $\pi$ - $\pi$  stacking arrangements with the ligand for approximately 150 ns of the simulation, with one residue above and below the ligand plane, as shown in Figure 5.23.

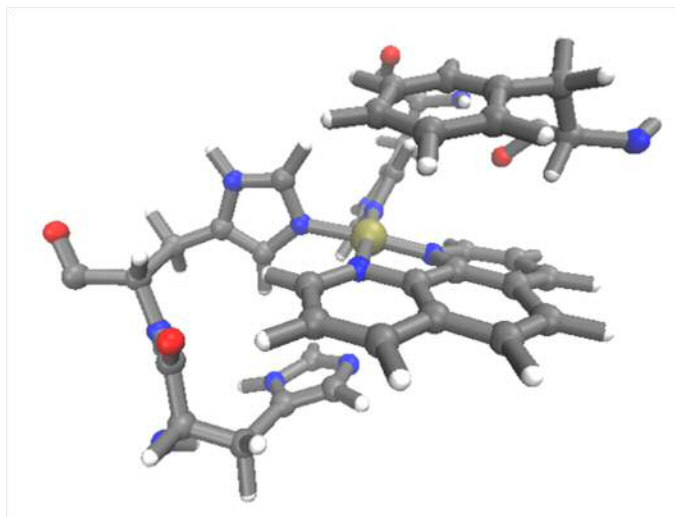


Figure 5.23: Snapshot of  $\text{Pt}^{\text{II}}(\text{A}\beta 16)$  simulation F, where Tyr10 and His13 form  $\pi$ - $\pi$  contacts of less than 5 Å with the phenanthroline ligand. Other residues are omitted for clarity

In addition, the  $\pi$ - $\pi$  contact distance may be plotted as a function of time. Figure 5.24 shows the contact distance for simulation F, where both Phe4 and His13 form  $\pi$ - $\pi$  interactions with the phenanthroline ligand with contact distance of 4-5 Å. Corresponding data for simulations G-J is shown in the Appendix.

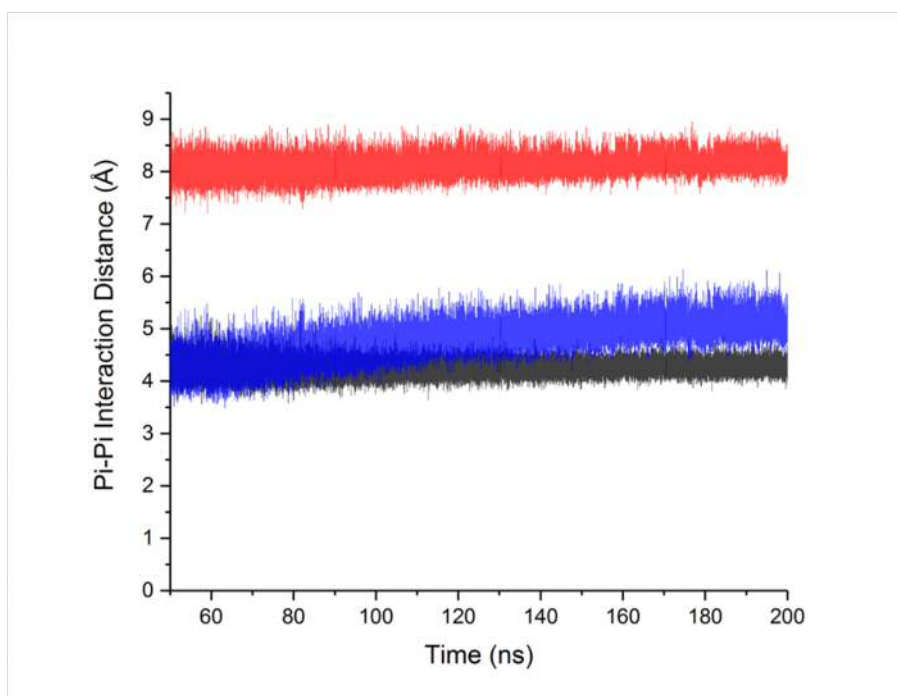


Figure 5.24: Plot of  $\pi$ - $\pi$  contact distance over production MD trajectory of simulation F. Phe4-ligand distance is shown in black, Tyr10-ligand distance in red and His13-ligand distance in blue.

## 5.3 A $\beta$ 42 and Pt<sup>II</sup>(phen)-A $\beta$ 42 Simulations

While the A $\beta$ 16 fragment provides a useful model of Pt<sup>II</sup>-A $\beta$  interaction and has been characterised experimentally (see Chapter 1), greater insight into the metal-peptide interactions would be gained by modelling the full 42-residue A $\beta$  peptide. In this section, four further molecular dynamics simulations are performed on each Pt<sup>II</sup>(phen)-A $\beta$ 42 and A $\beta$ 42 systems, along with preliminary analysis of the trajectories, towards determining the effects of the Pt<sup>II</sup>(phen) complex on peptide size, secondary structure, and formation of hydrogen bonding and salt bridges.

### 5.3.1 Computational Details

Here, MD and LFMD calculations were performed as described in Section 5.2.1, although for this work a parallel version of DL-POLY and DL-POLY-LF were used.

Conformations of A $\beta$ 42 and Pt<sup>II</sup>(phen)-A $\beta$ 42 were generated using LowMode MD, as described previously.[9] Conformations were selected to give a range of both collapsed and extended conformations. Initial conformers of the A $\beta$  and Pt<sup>II</sup>(phen)-A $\beta$ 42 systems used for simulation are shown in Figures 5.25 - 5.26.

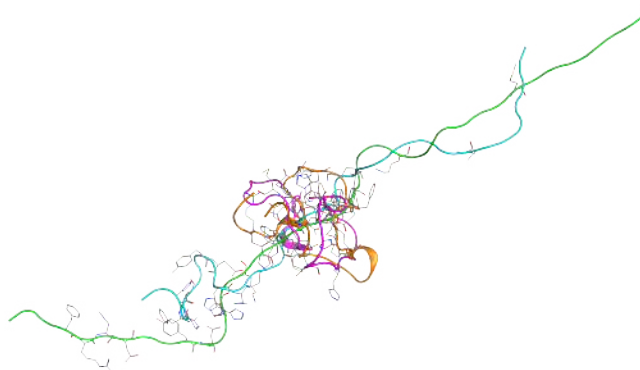


Figure 5.25: Overlay of initial A $\beta$ 42 conformations. Hydrogens, and non-aromatic, non-charged residues omitted for clarity

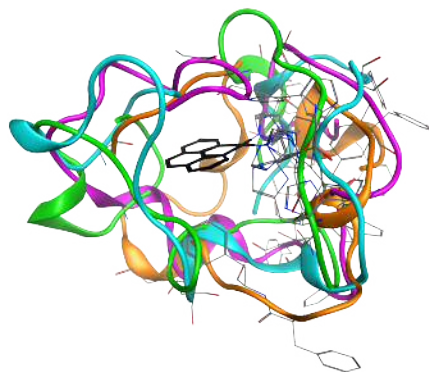


Figure 5.26: Overlay of initial  $\text{Pt}^{\text{II}}$ - $\text{A}\beta 42$  conformations.  $\text{Pt}^{\text{II}}(\text{phen})$  shown in black. Hydrogens and non-aromatic residues omitted for clarity.

### 5.3.2 $\text{A}\beta 42$ Equilibration

True equilibration of these large disordered peptides is beyond current capability; as before, systems were judged to be pseudo-equilibrated when the RMSD stabilised and fluctuated around a central value.[17]. RMSD plots of  $\text{A}\beta 42$  and  $\text{Pt}^{\text{II}}(\text{phen})$ - $\text{A}\beta 42$  trajectories are shown in Figures 5.27 and 5.28:

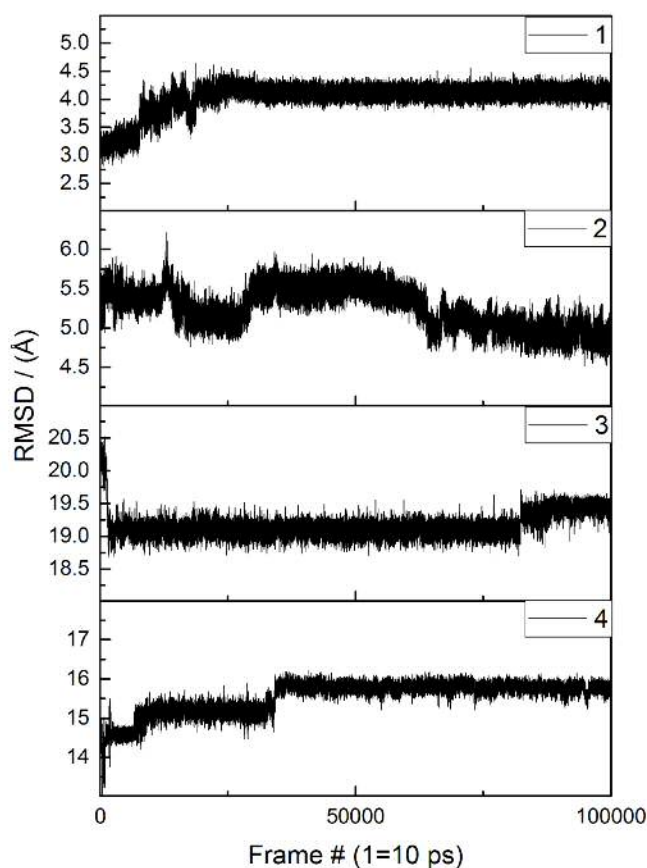


Figure 5.27: RMSD ( $\text{\AA}$ ) vs time for 4 simulations of A $\beta$ 42

As such, the systems were considered to be pseudo-equilibrated at 300, 700, 100 and 400 ns for simulations of A $\beta$ 42. While simulation 2 occupies a meta-stable state through 300-550 ns, this data was not used for analysis. Similarly, there is a jump in RMSD for simulation 3 at approximately 800ns after a long period of stable simulation. Data from 100-800 ns was used, but data after 800 ns was not. That simulations require hundreds of nanoseconds to reach stable structures is also in agreement with existing work.[17, 22]

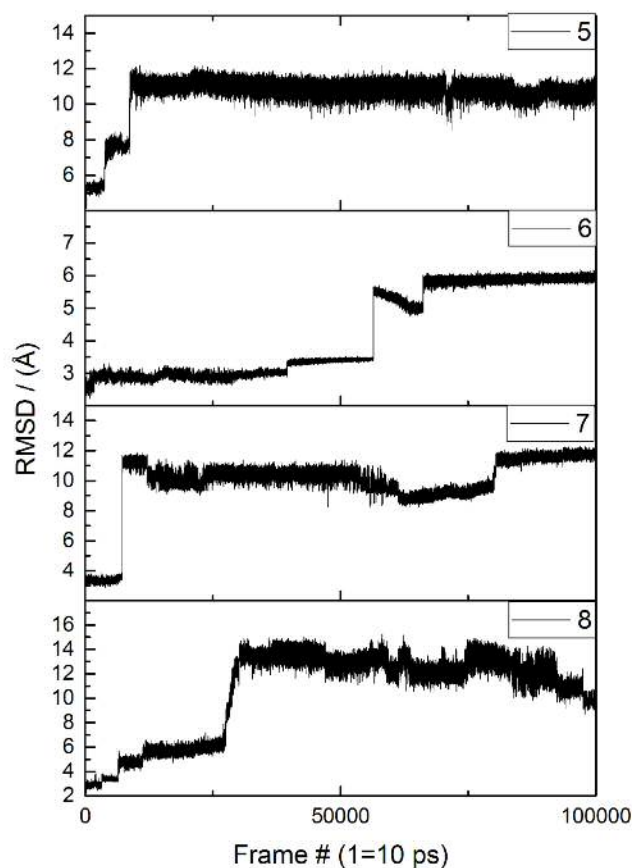


Figure 5.28: RMSD (Å) vs time for 4 simulations of  $\text{Pt}^{\text{II}}$ -A $\beta$ 42

Similarly, pseudo-equilibration was reached at 250, 700, 250 and 400 ns for  $\text{Pt}^{\text{II}}$ -A $\beta$ 42 simulations; RMSD data for the pseudo-equilibrated trajectories is shown in Tables A.14 and A.15 in the Appendix. Trajectories beyond these points were used for further analysis.

Notably, the reported RMSD values are much larger for A $\beta$ 42 systems than their A $\beta$ 16 counterparts. While this may be partly explained by their increased size and flexibility, it is also important that a range of starting conformations was used, including both extended and collapsed structures. Significant conformational changes of either of these extremes during equilibration would give rise to a large RMSD.

### 5.3.3 A $\beta$ 42 Production MD

As before, simulations were analysed using VMD and associated tools.[19] Radius of gyration data for A $\beta$ 42 simulations is shown in Figure 5.29 and summarised in Table 5.16, while the corresponding data for Pt<sup>II</sup>(phen)-A $\beta$ 42 is shown in Figure 5.30 and Table 5.17.

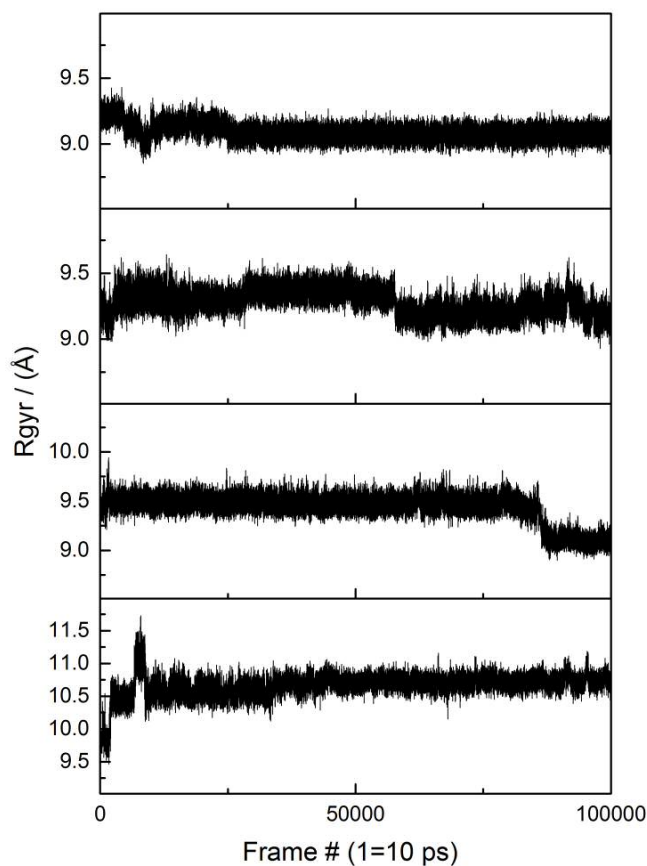


Figure 5.29:  $R_g$  ( $\text{\AA}$ ) vs time for simulations of A $\beta$ 42

A $\beta$ 42	Ave. $R_g$ / ( $\text{\AA}$ )	SD	Min.	Max.
1	9.070	0.044	8.897	9.927
2	9.217	0.076	8.928	9.618
3	9.482	0.066	9.225	9.835
4	10.725	0.089	10.152	11.187

Table 5.16: Average, standard deviation, min. and max.  $R_g$  of production MD for A $\beta$ 42 simulations.



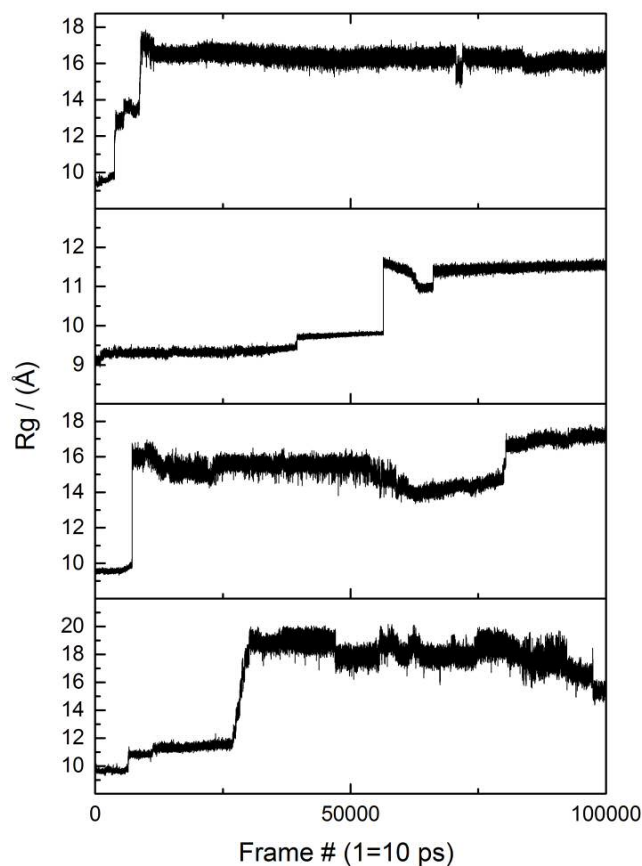


Figure 5.30:  $R_g$  (Å) vs time for simulations of  $\text{Pt}^{\text{II}}\text{-A}\beta 42$

$\text{Pt}^{\text{II}}(\text{A}\beta 42)$	Ave. $R_g$ / (Å)	SD	Min.	Max.
5	16.277	0.274	14.647	17.290
6	11.488	0.061	11.199	11.732
7	15.517	1.046	13.347	17.817
8	17.971	0.946	14.356	20.168

Table 5.17: Average, standard deviation, min. and max.  $R_g$  of production MD for  $\text{Pt}^{\text{II}}(\text{A}\beta 42)$  simulations.

The  $R_g$  data in Figure 5.29 for the  $\text{A}\beta 42$  systems are generally in agreement with other simulations (see Section 1.5.3), as well as experimental work, which report values of  $9 \pm 1$  Å and  $10.1 \pm 0.6$  Å (see [17] and references within). The fourth simulation shows much larger  $R_g$  than the others, but is still within results referred to above. While the first two simulations are not significantly different to one another, the third and fourth display  $R_g$  values that are significantly different to the others -

indicating that we are sampling multiple stable conformations of A $\beta$ 42, as expected for an intrinsically disordered peptide.

In contrast, the Pt<sup>II</sup>-A $\beta$ 42 systems display much larger Rg values than their non-platinated counterparts. To date, there is no experimental data to compare this against. As mentioned above, the Pt<sup>II</sup>(phen) fragment has an Rg of 2.29 Å, so the addition of this fragment alone cannot account for the difference observed *i.e.* the Pt<sup>II</sup>(phen) unit therefore induces a significant structural change in the A $\beta$ 42 peptide, from collapsed (small Rg) to extended (larger Rg). Simulations 5 and 7 are not significantly different to one another, judged on SD data. By the same assessment, simulation 8 is significantly different to simulation 7 but not simulation 5. Simulation 6 is significantly different to all other Pt<sup>II</sup>-A $\beta$ 42 simulations. Here, there is a greater range of Rg values observed, suggesting that with the Pt<sup>II</sup>(phen) unit attached, the peptide becomes more flexible, perhaps due to disruption of other intramolecular interactions (*vide infra*); Rg standard deviations are also much larger for the Pt<sup>II</sup>(phen) simulations, supporting the suggestion of greater overall flexibility.

In addition, RMSF per residue data was collected for these trajectories. Plots of average RMSF by residue for simulations 1-4 of A $\beta$ 42 and simulations 5-8 of Pt<sup>II</sup>-A $\beta$ 42 as well as a comparison, is shown in Figure 5.31. As before, negative values indicate that the Pt<sup>II</sup>-A $\beta$ 42 system has lower RMSF than A $\beta$  alone. Plots of only A $\beta$ 42 and Pt<sup>II</sup>-A $\beta$ 42 are shown in Figures A.19 and A.20 in the Appendix.

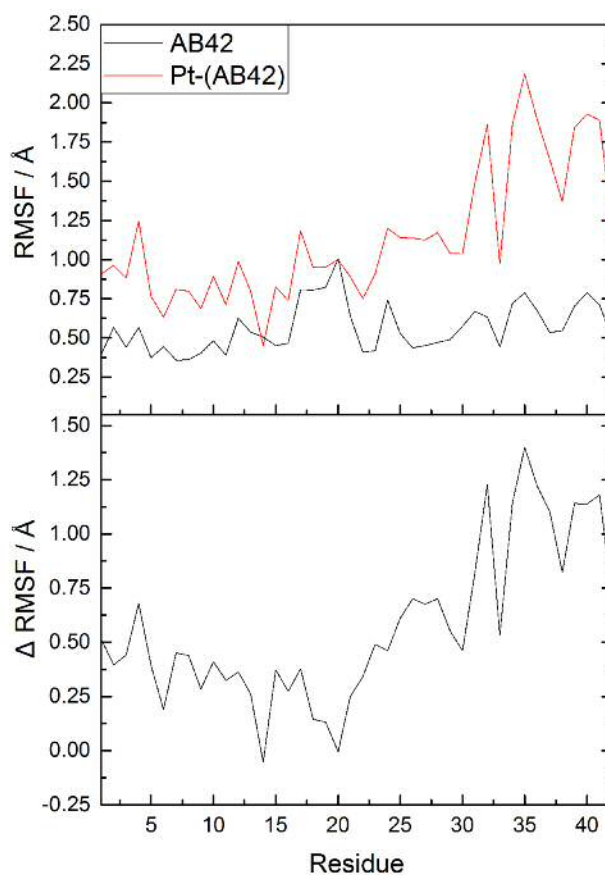


Figure 5.31: Average RMSF per residue for simulations 1-4 of A $\beta$ 42 and simulations 5-8 of Pt<sup>II</sup>-A $\beta$ 42 (top) and the difference in RMSF between A $\beta$ 42 and Pt<sup>II</sup>-A $\beta$ 42 (bottom)

For A $\beta$ 42, relatively large RMSF values are observed for the central region of the peptide, particularly residues Leu17-Phe20. Other notably large RMSF values occur at Val24, Leu34-Val36 and Val40. On the other hand, the peptide N-terminus displays surprisingly low RMSF, with values typically below 0.65 Å. In Pt<sup>II</sup>-A $\beta$ 42, RMSF values are consistently larger than for the free peptide - only His14 and Phe20 display lower RMSF than in A $\beta$ 42. Of these, His14 is one of the Pt<sup>II</sup> binding sites, so it is not surprising that RMSF values decrease. However, His6 does not show this behaviour. Of particular note are the very large RMSF values observed for residues in the C-terminal region of Pt<sup>II</sup>-A $\beta$ 42. Here, RMSF generally remains above 1 Å, with Ile32, Leu34 and Met35-Ala42 displaying values up to approximately 2.2 Å. In this peptide segment, Gly33 shows a clear decrease in RMSF relative to adjacent residues that is reflected in the free peptide; while possibly due to its small size, the small RMSF values are not seen to the same extent for Gly37-38. In the

$\Delta$ RMSF plot, the greater fluctuations of  $\text{Pt}^{\text{II}}$ -A $\beta$ 42 over A $\beta$ 42 are clearly illustrated. However, these highly flexible systems provide large standard deviations in RMSF values, even for these stable regions of simulation. As a result, none of the differences in RMSF described here were significant (*i.e.* no differences in RMSF exceed two standard deviations).

As before, peptide secondary structure was calculated using STRIDE in VMD. Overall secondary structure plots for A $\beta$ 42 and  $\text{Pt}^{\text{II}}$ -A $\beta$ 42 simulations are shown in Figures 5.32 and 5.33, respectively.

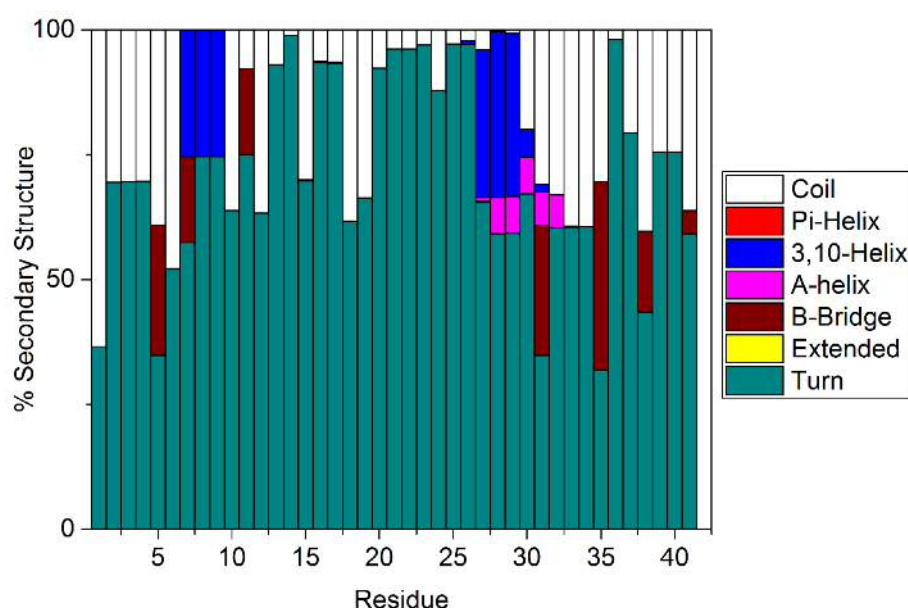


Figure 5.32: Percentage secondary structure by residue for A $\beta$ 42 simulation. Secondary structure assigned using STRIDE algorithm

As for the shorter A $\beta$ 16 fragment, the free A $\beta$ 42 adopts largely turn-/coil- type secondary structure (69% and 22%, respectively). This is in agreement with many other modelling studies, reviewed in Chapter 1. Encouragingly however, these A $\beta$ 42 simulations show considerable amounts of 3,10-helical structure, at residues Asp7-Gly9 (25%) and Ser26-Ile31 (up to 35%), unlike many other simulations (see Chapter 1). The appearance of this structure is in agreement with experimentally determined secondary structure of A $\beta$  in water (PDB code: 2LFM) though the location is different - in that case, 3,10-helical structure was observed at residues Val12-Val24. In addition, residues Arg5 (26%), Asp7 (17%), Glu11 (17%), Ile31 (26%), Met35 (38%), Gly38 (16%) and Ile41 (5%) occupy  $\beta$ -bridge conformations for significant periods of

simulation time. Furthermore, the A $\beta$  simulations also display a small amount (< 8%) of  $\alpha$ -helical structure in residues Asn27-Ile32 of the central hydrophobic region of the peptide.

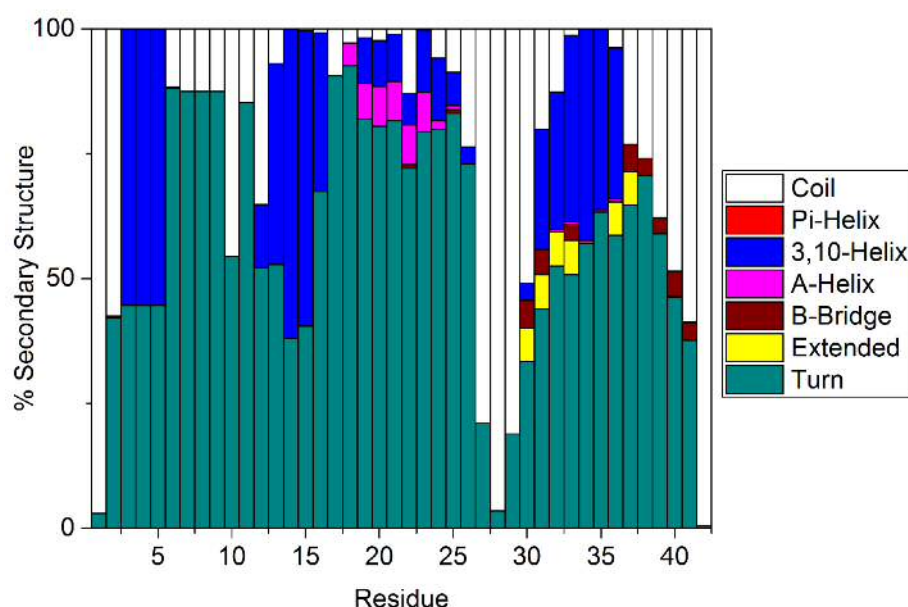


Figure 5.33: Percentage secondary structure by residue for Pt<sup>II</sup>-A $\beta$ 42 simulation. Secondary structure assigned using STRIDE algorithm

In the Pt<sup>II</sup>-A $\beta$ 42 simulations, residues are overwhelmingly assigned as turn- or coil-type secondary structure (58% and 24%) respectively. Among these data, the high probability of turn-structure within the N-terminus is in agreement with work discussed in previous chapters, though there is also a considerable amount of 3,10 helix found in residues Arg5-Asp7 (55% of total simulation time). Interestingly, the degree of helical character is significantly increased in the Pt<sup>II</sup>-A $\beta$ 42 systems compared to the free A $\beta$ 42 peptide; there are a succession of short 3,10 helices in Pt<sup>II</sup>-A $\beta$ 42 at Val12-Lys16 (up to 62%), Phe19-Ser (up to 13%) Ala30-Val36 (up to 42%). There is also low propensity for the central region of the peptide Val18-Gly25 to adopt an  $\alpha$ -helical conformation (< 8%). Finally, the Pt<sup>II</sup>-A $\beta$ 42 simulations also show decreased prevalence of  $\beta$ -bridge structures relative to the free peptide, though residues Ala30-Ile41 all show some degree of  $\beta$ -bridge character. Residues Gly29-Val40 also display extended conformations for small periods of the simulation (< 7%). This illustrates that not only does coordination of the Pt<sup>II</sup>(phen) unit to A $\beta$  alter the shape and size of the peptide, it also disrupts the existing peptide secondary structure. Despite binding near the N-terminus, these effects on structure are observed over the entire

peptide chain.

Interestingly, the helices in these  $\text{Pt}^{\text{II}}$ -A $\beta$ 42 simulations are observed in the central hydrophobic region of the peptide, which is known to be important in amyloid aggregation. In MD simulations of A $\beta$ , Yang and Teplow [21] highlighted the importance of helices for peptide aggregation and suggested that agents that stabilise helical conformations (or destabilise  $\beta$ -sheets) would block the formation of A $\beta$  oligomers. The increased prevalence of helical structures in our  $\text{Pt}^{\text{II}}$ -A $\beta$ 42 simulations suggest that  $\text{Pt}^{\text{II}}$ (phen) stabilises these structures, thereby inhibiting the formation of toxic oligomers.

The peptide salt bridge network was also calculated using VMD. In the full 42-residue peptide there are three positively charged and six negatively charged amino acid residues, giving a total of eighteen possible salt bridge interactions. As before, salt bridges are defined as any contact distance of less than 3.2 Å between O-N atoms in charged residues. Salt bridge contact maps for A $\beta$ 42 and  $\text{Pt}^{\text{II}}$ -A $\beta$ 42 are shown in Figures 5.34 and 5.35 and the underlying data are contained in Tables A.16 and A.17.

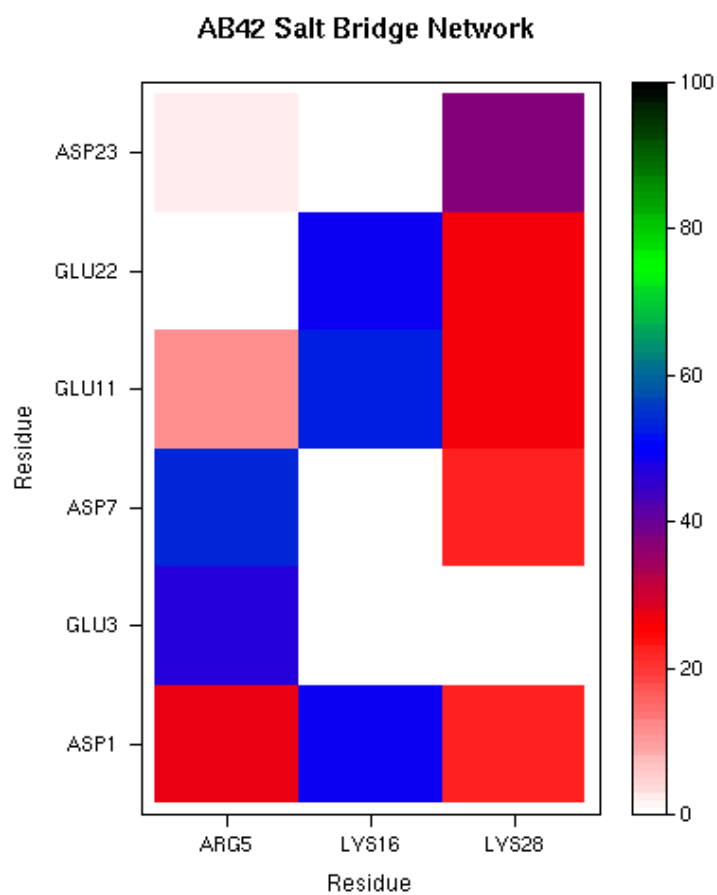


Figure 5.34: Salt bridge contact map for A $\beta$ 42 simulation.

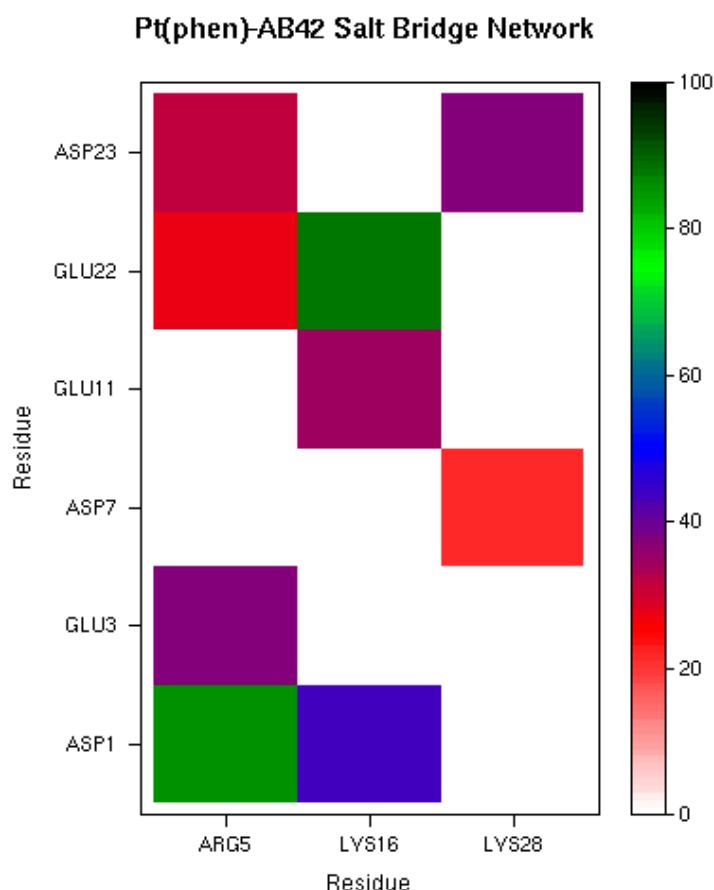


Figure 5.35: Salt bridge contact map for Pt<sup>II</sup>-Aβ42 simulation.

In the free Aβ42 system, thirteen of the possible salt bridges are occupied during simulation, though most are only formed for a low proportion of time, similar to the Aβ16 results above. However, the identity of these interactions are different. In Aβ42, Arg5 predominantly interacts with Asp7 (54%) and Glu3 (47%) and forms sporadic interactions with Asp1 (27%) and Glu11 (11%). Lys16 frequently forms salt bridges with Glu11 (53%), Glu22 (49%) and Asp1 (49%) - this high incidence of Lys16 salt bridge interactions is possibly due to its location between two collapsed regions of the peptide (*vide supra*). In contrast to the Aβ16 peptide, Lys16 does not interact with Glu3 in our simulations. Finally, Lys28 most commonly interacts with Asp23 (38%) - a common salt bridge in Aβ- while forming less abundant salt bridges with Glu22 (26%), Glu11 (26%), Asp1 (23%) and Asp7 (22%). As before, the combined frequency of the observed salt bridge interactions exceeds 100%, indicating that residues are close enough to form multiple salt bridge-type interactions simultaneously.

In a similar fashion to the Aβ16 systems, coordination of the Pt<sup>II</sup>(phen) unit to



the peptide causes significant disruption to the salt bridge pattern. Here, only nine of the possible eighteen interactions are observed during simulation. In Pt<sup>II</sup>-A $\beta$ 42, Arg5 frequently interacts with Asp1 (85%), and forms less common salt bridges with Glu3 (37%), Asp23 (31%) and Glu22 (27%). Interestingly, Arg5 does not interact with Asp7 - these residues are either side of the Pt<sup>II</sup>(phen) binding site, suggesting that coordination of the metal complex prevents this salt bridge from forming. Lys16 predominantly forms a salt bridge with Glu22 (87%), much more frequently than observed for the free peptide; this residue still interacts with Glu11, but with reduced frequency compared to A $\beta$ 42 (34%). Lys28 forms notably fewer salt bridges in Pt<sup>II</sup>-A $\beta$ 42, only interacting with Asp23 (38%) and Asp7 (21%).

The Lys28-Asp23 salt bridge is noted for its importance in stabilising  $\beta$ -sheet structures and subsequent fibril formation. As such, agents which decrease the prevalence of this interaction are thought to reduce the likelihood of A $\beta$  fibril formation (see Chapter 1).[17] Interestingly, the Pt<sup>II</sup>(phen) system studied here has no effect on this salt bridge, though it decisively impacts others. Given that Pt<sup>II</sup>(phen) is known to prevent fibril formation (see Chapter 1, [20]), this suggests that its effects on peptide (secondary) structure and other interactions are more important than its effect on Lys28-Asp23 interaction, precluding  $\beta$ -sheet and fibril formation.

Finally, hydrogen bond networks were analysed. The total number of hydrogen bonds per simulation frame for A $\beta$ 42 and Pt<sup>II</sup>(A $\beta$ 42) were calculated and data shown in Tables 5.18 and 5.19, respectively. Similar to the results for the 16-residue peptide studied above, the Pt<sup>II</sup>-A $\beta$ 42 systems display fewer hydrogen bonds than their non-platinated A $\beta$ 42 counterparts, though the difference is generally not significant due to the large SD values. This again suggests that hydrogen bonds in these systems are highly transient. In contrast to A $\beta$ 16, the A $\beta$ 42 systems never adopt conformations with no hydrogen bonds. In addition, the Pt<sup>II</sup>-A $\beta$ 42 systems do not form the maximum observed number of hydrogen bonds seen in A $\beta$ 42 in three of the four simulations. It is probable that the exact identity of the hydrogen bonds will be different in the Pt<sup>II</sup>-A $\beta$ 42 and A $\beta$ 42 simulations, especially as the prevailing secondary structures in these systems are different. Further work could analyse the changes in hydrogen bonding patterns in more detail (*i.e.* on an atom-by-atom basis) but was not performed for this thesis due to time constraints.

A $\beta$ 42	Ave. # HB	SD	Min.	Max.
1	14.0	2.8	2	26
2	13.7	2.8	2	25
3	10.5	2.4	2	22
4	13.3	2.7	3	24

Table 5.18: Average, SD, min. and max. values of the number of hydrogen bonds present in A $\beta$ 42 simulation.

Pt <sup>II</sup> (A $\beta$ 42)	Ave. # HB	SD	Min.	Max.
5	11.6	2.5	2	23
6	7.6	2.0	1	15
7	11.5	2.4	2	22
8	9.0	2.4	0	19

Table 5.19: Average, SD, min. and max. values of the number of hydrogen bonds present in Pt<sup>II</sup>(A $\beta$ 42) simulation.

## 5.4 Conclusions

Molecular dynamics, using LFMM description of metal coordination coupled with AMBER description of peptide, elaborates details of the structure and properties of the complex formed between Pt<sup>II</sup>(phenanthroline) and both the metal binding N-terminal fragment and the complete A $\beta$  peptide.

Modelling the Pt<sup>II</sup>(phenanthroline)-A $\beta$ 16 fragments was achieved using five distinct starting structures, along with analogous simulations of the metal-free peptide. Here, simulations reach equilibration within a few tens of nanoseconds. A total of over 800 ns of equilibrated data for metal-free and platinated peptides allow for detailed comparison of size, secondary structure, and formation of hydrogen bonding and salt bridges. Small changes in overall size are observed on Pt<sup>II</sup> binding, but rather larger differences in the mobility of individual residues, measured by root mean square fluctuation, occur. Changes in secondary structure, hydrogen bonding and salt-bridges on complexation of Pt<sup>II</sup> are also observed: in general, His6 and His14 that are bound to Pt<sup>II</sup> are less mobile and more structured than their Pt<sup>II</sup>-

free counterparts. Residues between these are slightly more mobile when bound, and exhibit slightly greater propensity to adopt turn and 3,10-helical structures. Hydrogen bonding is reduced by complexation, but salt bridges are more likely to form in the presence of  $\text{Pt}^{\text{II}}$ , while close contacts between phenanthroline ligand and aromatic residues Phe4, Tyr10 and His13 are present in at least some of the trajectories, with one structural motif of a sandwich of phenanthroline between Phe4 and His13 observed in a significant proportion of simulation time. In summary, microsecond timescale molecular dynamics quantifies the subtle but definite changes induced by coordination of  $\text{Pt}^{\text{II}}(\text{phen})$  to the N-terminal portion of the A $\beta$  peptide.

For the  $\text{Pt}^{\text{II}}(\text{phenanthroline})\text{-A}\beta 42$  and A $\beta 42$  systems, four simulations of each were performed using the same modelling protocol. Each of the eight simulations was run for 1  $\mu\text{s}$  and equilibration was reached over the course of several hundred nanoseconds. Equilibrated data was again used to study peptide size, secondary structure and intramolecular interactions such as hydrogen bonds and salt bridges. In this complete peptide system, pronounced changes are observed upon  $\text{Pt}^{\text{II}}$  coordination, particularly for shape, size and secondary structure;  $\text{Pt}^{\text{II}}\text{-A}\beta 42$  systems adopt more extended conformations than their free peptide counterparts, while the prevalence of helical structure was also greatly increased. Such stabilisation of helices is thought to prevent  $\beta$ -sheet formation, a critical component of amyloid fibril structure.[21] This, along with the range of conformations adopted by  $\text{Pt}^{\text{II}}\text{-A}\beta 42$ , supports the suggestion that these  $\text{Pt}^{\text{II}}$  agents promote amorphous aggregation rather than fibril formation.[20] Further effects of  $\text{Pt}^{\text{II}}$  coordination are seen in the salt bridge network, where  $\text{Pt}^{\text{II}}$  decreases the number of interactions, likely as a result of its influence on the overall peptide structure. These results provide the first insight into the precise intramolecular interactions of  $\text{Pt}^{\text{II}}(\text{phen})\text{-A}\beta 42$  systems and provide a plausible explanation for the experimentally observed effects of  $\text{Pt}^{\text{II}}(\text{phen})$  on A $\beta 42$ .

## Bibliography

- (1) M. Turner, S. Mutter, R. J. Deeth and J. A. Platts, *Plos One*, In press.
- (2) W. Smith, C. W. Yong and P. M. Rodger, *Molecular Simulation*, 2002, **28**, 385–471.
- (3) H.-C. Tai, R. Brodbeck, J. Kasparkova, N. J. Farrer, V. Brabec, P. J. Sadler and R. J. Deeth, *Inorganic Chemistry*, 2012, **51**, 6830–6841.
- (4) A. E. Anastasi and R. J. Deeth, *Journal of Chemical Theory and Computation*, 2009, **5**, 2339–2352.
- (5) W. Cornell, P. Cieplak, C. Bayly, I. Gould, K. Merz, D. Ferguson, D. Spellmeyer, T. Fox, J. Caldwell and P. Kollman, *Journal of the American Chemical Society*, 1995, **117**, 5179–5197.
- (6) *Molecular Operating Environment (MOE)*, (2013.08), version 2013.08, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2013.
- (7) M. Turner, J. A. Platts and R. J. Deeth, *Journal of Chemical Theory and Computation*, 2016, **12**, 1385–1392.
- (8) M. Turner, R. J. Deeth and J. A. Platts, *Journal of Inorganic Biochemistry*, 2017, **173**, 44–51.
- (9) P. Labute, *Journal of Chemical Information and Modeling*, 2010, **50**, 792–800.
- (10) C. W. Yong, *Journal of Chemical Information and Modeling*, 2016, **56**, 1405–1409.
- (11) R. J. Deeth, N. Fey and B. Williams–Hubbard, *Journal of Computational Chemistry*, 2005, **26**, 123–130.
- (12) M. Frisch and G. Trucks et al., *Gaussian09*, Wallingford, CT, 2009.
- (13) S. Nosé, *The Journal of Chemical Physics*, 1984, **81**, 511–519.
- (14) W. G. Hoover, *Physical Review A*, 1985, **31**, 1695–1697.
- (15) J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *Journal of Computational Physics*, 1977, **23**, 327–341.
- (16) D. Frishman and P. Argos, *Proteins-Structure Function and Genetics*, 1995, **23**, 566–579.

- (17) P. D. Q. Huy, Q. V. Vuong, G. La Penna, P. Faller and M. S. Li, *ACS Chemical Neuroscience*, 2016, **7**, 1348–1363.
- (18) M. Dong, H. Li, D. Hu, W. Zhao, X. Zhu and H. Ai, *ACS Chemical Neuroscience*, 2016, **7**, 599–614.
- (19) W. Humphrey, A. Dalke and K. Schulten, *Journal of Molecular Graphics*, 1996, **14**, 33–38, 27–28.
- (20) K. J. Barnham, V. B. Kenche, G. D. Ciccotosto, D. P. Smith, D. J. Tew, X. Liu, K. Perez, G. A. Cranston, T. J. Johanssen, I. Volitakis, A. I. Bush, C. L. Masters, A. R. White, J. P. Smith, R. A. Cherny and R. Cappai, *Proceedings of the National Academy of Sciences of the United States of America*, 2008, **105**, 6813–6818.
- (21) M. Yang and D. B. Teplow, *Journal of Molecular Biology*, 2008, **384**, 450–464.
- (22) D. F. Raffa and A. Rauk, *The Journal of Physical Chemistry B*, 2007, **111**, 3789–3799.

## 6 Conclusion

The work presented in this thesis aims to summarise investigations into modelling the interactions of a series of platinum<sup>II</sup> complexes and the amyloid- $\beta$  peptide, believed to play a critical role in the development of Alzheimer’s disease. These complexes have been shown experimentally to interact with the A $\beta$  peptide and prevent the formation of amyloid fibrils, but there has been limited modelling work performed on these systems. Here, this was achieved using a range of DFT, semi-empirical and molecular mechanics-based approaches.

Though many Pt<sup>II</sup> agents have previously been studied, only the coordination mode of the original Pt<sup>II</sup>(phen) complex has been elucidated by experiment. Coordination of these Pt<sup>II</sup> agents to single amino acid residues to model their basic coordination preference is not representative of the true Pt<sup>II</sup>-A $\beta$  complexes, and so larger peptide models were pursued in this work. Initially, the metal binding region (His6 – His14) was studied before extension to the 16-residue N-terminal fragment (A $\beta$ 16) and the complete A $\beta$ 42 peptide.

While accurate, DFT methods are too expensive to be used to generate and model large numbers of conformations of flexible molecules such as A $\beta$ . As an alternative, semi-empirical and molecular mechanics-based methods are considered. Traditionally, molecular mechanics approaches fail to model the *d*-electron effects of transition metals, though parameters to handle Pt<sup>II</sup> in biological systems have previously been published. However, implementation of this purely classical technique required additional parameterisation and may not provide the desired flexibility in the description of Pt<sup>II</sup> coordination. As a result, the LFMM model - which directly introduces the influence of *d*-electrons to the classical forcefield - was explored as a computationally tractable alternative to DFT. Additionally, this LFMM approach requires no additional parameterisation.

A $\beta$  is an intrinsically disordered peptide – it contains little or no defined structure – and so the consideration of peptide flexibility and the sampling of available peptide conformations is key to this work. While systematic and stochastic conformational searches are impractical for systems of the size studied here, the recent LowMode MD search method is known to efficiently identify low energy conformations of flexible macromolecules, making it an attractive option for studies of A $\beta$ . In addition, while LowMode MD is able to produce distinct conformations of the peptide, further insight into the effects of Pt<sup>II</sup> coordination on A $\beta$  are obtained using molecular dynamics techniques.

Following modelling of various Pt<sup>II</sup> crystal structures using a range of DFT methods, the hybrid BHandH functional was shown to accurately model Pt<sup>II</sup> coordination geometry and has been used for similar investigations previously. Studies of two Pt<sup>II</sup> agents, namely Pt<sup>II</sup>(bipy) and Pt<sup>II</sup>(phen), with A $\beta$ 6-14 were used to validate the LFMM and LowMode MD methods against BHandH results; while LowMode MD produces accurate peptide geometries, LFMM was shown to accurately and reliably represent Pt<sup>II</sup> coordination, though relative energies of conformations were initially poorly reproduced. Investigation of semi-empirical methods showed that PM7 was able to reproduce both the geometries of these crystal structures and the relative energies and geometries of Pt<sup>II</sup>-A $\beta$  complexes at the DFT level at much lower computational expense.

With the suitability of the LFMM / PM7 approach ascertained, this protocol was used to model a series of six Pt<sup>II</sup>(ligand) complexes, consisting mainly of those taken directly from current chemical literature, and study their effects on A $\beta$ 6-14 peptide structure in detail. In particular, the preferred coordination sites of these complexes and resulting peptide secondary structure were investigated. Importantly, the coordination of these Pt<sup>II</sup> complexes to A $\beta$  is found to greatly restrict the conformational freedom of the peptide fragment, regardless of the ligand present. In addition, Boltzmann weighting of the energies of these conformations shows that Pt<sup>II</sup> complexes possess only a small number of accessible conformations at 310 K. However, each of the different complexes show a clear A $\beta$  coordination preference, with varied binding to His6, His13 and His14 as well as N $\delta$  and N $\epsilon$  atoms within those residues. Among these, the Pt<sup>II</sup>(phen) complex is shown to bind to His6 N $\epsilon$  His14 N $\epsilon$ , in agreement with experimental data.

Furthermore, analysis of the peptide structure in the  $\text{Pt}^{\text{II}}\text{-A}\beta$  adducts using Ramachandran plots and the STRIDE algorithm revealed that complexes largely adopt conformations found in  $\alpha$ -helical areas of conformational space regardless of the ligand, though wide coverage of the Ramachandran plot was observed in all cases. Notably, the non-planar ligand (**5**) promotes conformations found on the right hand side of the Ramachandran plot – these conformations are observed much more frequently than for any of the other ligand systems studied, suggesting some importance of this ligand structure. Analysis of peptide secondary structure using STRIDE showed that the free peptide adopts turn-type secondary structure in residues His6-Glu11; coordination of the  $\text{Pt}^{\text{II}}$  complexes disrupts this secondary structure in a ligand-specific manner, interrupting or translating the turn structure or forming a 3<sub>10</sub>-helix. These results suggest that  $\text{Pt}^{\text{II}}$  coordination preference depends on the nature of the ligand in the complex, and that this also strongly affects the secondary structure of the N-terminal region of  $\text{A}\beta$ . As such, the metal coordination and subsequent alteration of peptide structure may be controlled via ligand design.

Further studies using ligand-field molecular dynamics on  $\text{Pt}^{\text{II}}(\text{phen})\text{-A}\beta_{16}$  and  $\text{A}\beta_{42}$  have been used to investigate the influence of the  $\text{Pt}^{\text{II}}$  complex on the structure and properties of  $\text{A}\beta$ . True structural equilibration of this intrinsically disordered peptide is beyond current simulation capability, but both  $\text{A}\beta$  and  $\text{Pt}^{\text{II}}\text{-A}\beta$  systems were pseudo-equilibrated in timescales up to several hundred nanoseconds. Analysis of dynamics trajectories allowed for comparisons of size, secondary structure and intramolecular interactions (hydrogen bonds and salt bridges) of the  $\text{A}\beta$  and  $\text{Pt}^{\text{II}}(\text{phen})\text{-A}\beta$  systems.

MD results for free  $\text{A}\beta$  monomers are shown to be in general agreement with existing experimental and simulation data. Coordination of the  $\text{Pt}^{\text{II}}(\text{phen})$  unit to  $\text{A}\beta_{16}$  does not induce large changes in peptide size - all structures remain in collapsed conformations - but this is not the case for  $\text{A}\beta_{42}$ , where  $\text{Pt}^{\text{II}}$  coordination produces clear changes in peptide size and structure, particularly through the peptide's central and C-terminal regions. Furthermore, while secondary structural elements in both free and platinated peptides are generally consistent across  $\text{A}\beta_{16}$  and  $\text{A}\beta_{42}$  systems, notable deviations, such as additional helices in  $\text{Pt}^{\text{II}}\text{-A}\beta_{42}$ , are observed in the complete peptide model. In addition, important differences are also observed in salt bridge patterns within the N-terminus of  $\text{A}\beta_{42}$  systems compared to  $\text{A}\beta_{16}$ . This suggests that while  $\text{A}\beta_{16}$  is a useful model when considering the N-terminal metal



binding sites, peptide models larger than the 16-residue N-terminus are required to capture the full extent of the effects of  $\text{Pt}^{\text{II}}$  coordination on  $\text{A}\beta$ . Importantly, the  $\text{Pt}^{\text{II}}$ - $\text{A}\beta_{42}$  structures display a clear increase in  $\alpha$ -helical and  $3_{10}$ -helical character throughout the central hydrophobic region of the peptide, as well as the N-terminus compared to free  $\text{A}\beta_{42}$ ; stabilisation of helical structure in  $\text{A}\beta$  has been previously suggested as a route to preventing  $\beta$ -sheet and amyloid fibril formation. As a result, these simulations suggest a rationale for how these complexes affect  $\text{A}\beta$  structure and its ability to form amyloid fibrils, a phenomenon that was previously unexplained.

The results presented in this thesis provide the first computational approaches to study the interactions of a series of  $\text{Pt}^{\text{II}}$ -based agents - purported to be useful in the treatment of Alzheimer's disease - and the amyloid- $\beta$  peptide. Notably, this work presents both the first dynamics simulations of these anti- $\text{A}\beta$   $\text{Pt}^{\text{II}}$  complexes and the first long-timescale results of LFMD. These investigations offer insight into the nature of the coordination of these complexes to the peptide, their effects on peptide structure and dynamics, and provide a reasonable explanation for their experimentally observed effects on  $\text{A}\beta$ .

## 7 Further Work

There are several research avenues that follow directly from the work in this thesis:

First, the molecular dynamics simulations in Chapter 5 could be expanded in both simulation length and scope. In particular, exploration of enhanced sampling techniques such as metadynamics or replica-exchange molecular dynamics may provide further insight into the structure and behaviour of these peptides over time. Unfortunately, the DL-POLY-LF program is not capable of REMD simulations, so the ligand-field procedure would need to be implemented in another simulation package. In addition to this, molecular dynamics simulations of  $\text{Pt}^{\text{II}}$ -A $\beta$  dimer species may provide more detailed information into how these agents disrupt A $\beta$  aggregation. Furthermore, the inclusion of explicit solvent in these simulations may influence the available peptide conformations and the transitions between them.

Secondly, it would be interesting to model the  $\varphi$ -MePy ligands proposed by Fallor and co-workers (see Section 1.4.2). These complexes contain a  $\text{Pt}^{\text{II}}$ -C bond that is not parameterised in DommiMOE, so modelling would require thorough parameterisation and validation of any new work. However, the formally anionic C is isoelectronic with nitrogen, for which LFMM parameters do exist; these current parameters would provide a useful starting point. Once this was complete these  $\text{Pt}^{\text{II}}$  complexes could be modelled as described in this thesis and any similarities or differences in metal coordination and induced peptide behaviour noted.

Finally, LFMM parameters could be developed to study the  $\text{Ru}^{\text{II}}$  and  $\text{Co}^{\text{III}}$  anti-A $\beta$ -aggregation agents described in Section 1.4.2. While the  $\text{Pt}^{\text{II}}$  parameters used in this work could serve as a useful starting point, accurate parameterisation of two new metals and all associated ligands would likely be a long process. However, once complete, the procedures in this thesis could be replicated for these different classes of compound, and comparisons drawn.

## A Appendix

### A.1 Appendices for Chapter 3

#### A.1.1 DFT Modelling of Xray Crystal Structures

Basis	B3LYP		BHandH		PBE	
	Err. Length (Å)	Err. Angle (°)	Err. Length (Å)	Err. Angle (°)	Err. Length (Å)	Err. Angle (°)
LANL2DZ	0.123	4.4	0.065	4.5	0.090	4.4
cep-121G	0.105	4.1	0.048	4.3	0.071	4.2
SDD	0.102	4.0	0.042	4.1	0.070	4.0
SDD / 6-31G(d)	0.099	3.5	0.027	3.2	0.075	3.3
SDD / 6-311G(d)	0.103	3.6	0.052	3.3	0.077	3.3
SDD / 6-31G(d,p)	0.099	3.5	0.055	3.2	0.075	3.2
SDD / 6-31+G(d)	0.096	3.8	0.050	3.4	0.075	3.3
SDD / 6-31+G(d,p)	0.090	3.7	0.050	3.4	0.075	3.3
SDD / 6-31+G(2d,p)	0.096	3.5	0.046	3.7	0.071	3.6
SDD / 6-31++G(2d,p)	0.096	3.5	0.046	3.7	0.071	3.6
SDD / 6-31+G(2d,f,p)	0.087	3.5	0.044	3.6	0.068	3.4
SDD / cc-pVDZ	0.100	3.6	0.052	3.1	0.077	3.4
SDD / TZV	0.108	4.0	0.048	4.2	0.068	4.1

Table A.1: Max errors in crystal structures XUKWUV, OXUYUC, DAMIPT, HEKJEM, and MIMDPT using different DFT functionals and basis sets

Basis	B97-D		$\omega$ B97xD		LC- $\omega$ PBE	
	Err. Length ( $\text{\AA}$ )	Err. Angle ( $^\circ$ )	Err. Length ( $\text{\AA}$ )	Err. Angle ( $^\circ$ )	Err. Length ( $\text{\AA}$ )	Err. Angle ( $^\circ$ )
LANL2DZ	0.135	4.1	0.103	4.5	0.079	4.3
cep-121G	0.120	3.8	0.083	4.0	0.072	4.0
SDD	0.112	3.6	0.082	3.9	0.068	3.9
SDD / 6-31G(d)	0.095	2.9	0.084	2.9	0.075	3.0
SDD / 6-311G(d)	0.106	2.8	0.086	2.9	0.078	3.1
SDD / 6-31G(d,p)	0.108	2.8	0.084	2.9	0.075	3.0
SDD / 6-31+G(d)	0.106	3.0	0.084	3.1	0.075	3.1
SDD / 6-31+G(d,p)	0.109	2.9	0.083	3.0	0.075	3.1
SDD / 6-31+G(2d,p)	0.104	3.1	0.068	3.4	0.044	3.4
SDD / 6-31++G(2d,p)	0.104	3.1	0.068	3.4	0.044	3.4
SDD / 6-31+G(2d,f,p)	0.106	3.0	0.078	3.2	0.069	3.3
SDD / cc-pVDZ	0.095	2.8	0.076	2.9	0.052	2.8
SDD / TZV	0.118	3.6	0.084	3.9	0.055	4.0

Table A.2: Max errors in crystal structures XUKWUV, OXUYUC, DAMIPT, HEKJEM and MIMDPT using different DFT functionals and basis sets

### A.1.2 LFMM Parameters

The relevant LFMM.par parameters are shown here in MOE format:

[Morse]

```
#T1 T2 dist D a
M+2 * 2.000 200.000 1.000
* * 2.000 200.000 1.000
Pt+2 NA 2.075 60.0 1.70 0.0 0.0
Pt+2 NB 2.075 60.0 1.70 0.0 0.0
```

[vdW]

```
#M L A B
* * 10000 500
```

[ll]

```
#M L rep n
* * 1000 6
Pt+2 NA 3000 6
Pt+2 NB 2500 6
```

[esig]

```
#M L esig0 esig1 esig2 esig3 esig4 esig5 esig6
* * 0 0 0 0 0 0 0
Pt+2 NA 0 0 0 0 105000 0 0
Pt+2 NB 0 0 0 0 102000 0 0
```

[epix]

```
#M L epix0 epix1 epix2 epix3 epix4 epix5 epix6
* * 0 0 0 0 0 0 0
Pt+2 NA 0 0 0 0 0 0 0
Pt+2 NB 0 0 0 0 0 0 0
```

[epiy]

```
#M L epiy0 epiy1 epiy2 epiy3 epiy4 epiy5 epiy6
* * 0 0 0 0 0 0 0
Pt+2 NA 0 0 0 0 20000 0 0
Pt+2 NB 0 0 0 0 16000 0 0
```

```

[exds]
#M L exds0 exds1 exds2 exds3 exds4 exds5 exds6
* * 0 0 0 0 0 0
Pt+2 NA 0 0 0 0 0 0 147000
Pt+2 NB 0 0 0 0 0 0 137000

[pair]
#M L P0 P1 P2 P3 P4 P5 P6
* * 0 0 0 0 0 0

```

### A.1.3 Validation of LFMM Parameters

Bipy $\epsilon\epsilon$	Bond ( $\text{\AA}$ ) / Angle ( $^\circ$ )	BHandH	LFMM
	Ave. Pt-N <sub>Lig</sub>	2.011	1.979
	Ave. Pt-N <sub>Me-imid</sub>	2.004	1.997
	N <sub>Me-imid</sub> -N <sub>Me-imid</sub>	87.3	83.9
	N <sub>Lig</sub> -N <sub>Lig</sub>	80.4	92.2
	N <sub>Me-imid</sub> -N <sub>Lig</sub>	96.2	92.02

Table A.3: Selected geometric data for the small model system  $\text{Pt}^{\text{II}}(\text{bipy})(\text{Me-imid})_2$  where Me-imid is bound by N $\epsilon$  atoms. Data obtained DFT and LFMM methods.

Bipy $\delta\epsilon$	Bond ( $\text{\AA}$ ) / Angle ( $^\circ$ )	BHandH	LFMM
	Ave. Pt-N <sub>Lig</sub>	2.011	1.960
	Ave. Pt-N <sub>Me-imid</sub>	2.007	1.952
	N <sub>Me-imid</sub> -N <sub>Me-imid</sub>	87.7	90.3
	N <sub>Lig</sub> -N <sub>Lig</sub>	80.4	92.7
	N <sub>Me-imid</sub> -N <sub>Lig</sub>	96.0	88.9

Table A.4: Selected geometric data for the small model system  $\text{Pt}^{\text{II}}(\text{bipy})(\text{Me-imid})_2$  where Me-imid is bound by N $\delta$  and N $\epsilon$  atoms. Data obtained DFT and LFMM methods.

Phen $\delta\delta$	Bond ( $\text{\AA}$ ) / Angle ( $^\circ$ )	BHandH	LFMM
	Ave. Pt-N <sub>Lig</sub>	2.011	1.960
	Ave. Pt-N <sub>Me-imid</sub>	2.007	1.952
	N <sub>Me-imid</sub> -N <sub>Me-imid</sub>	88.5	89.6
	N <sub>Lig</sub> -N <sub>Lig</sub>	81.2	94.8
	N <sub>Me-imid</sub> -N <sub>Lig</sub>	95.3	88.3

Table A.5: Selected geometric data for the small model system  $\text{Pt}^{\text{II}}(\text{phen})(\text{Me-imid})_2$  where Me-imid is bound by N $\delta$  atoms. Data obtained DFT and LFMM methods.

Phen $\epsilon\epsilon$	Bond ( $\text{\AA}$ ) / Angle ( $^\circ$ )	BHandH	LFMM
	Ave. Pt-N <sub>Lig</sub>	2.019	1.982
	Ave. Pt-N <sub>Me-imid</sub>	2.001	1.998
	N <sub>Me-imid</sub> -N <sub>Me-imid</sub>	88.1	84.0
	N <sub>Lig</sub> -N <sub>Lig</sub>	81.2	93.9
	N <sub>Me-imid</sub> -N <sub>Lig</sub>	95.4	91.2

Table A.6: Selected geometric data for the small model system  $\text{Pt}^{\text{II}}(\text{phen})(\text{Me-imid})_2$  where Me-imid is bound by N $\epsilon$  atoms. Data obtained DFT and LFMM methods.

Phen $\delta\epsilon$	Bond ( $\text{\AA}$ ) / Angle ( $^\circ$ )	BHandH	LFMM
	Ave. Pt-N <sub>Lig</sub>	2.019	1.972
	Ave. Pt-N <sub>Me-imid</sub>	2.003	1.971
	N <sub>Me-imid</sub> -N <sub>Me-imid</sub>	88.4	85.2
	N <sub>Lig</sub> -N <sub>Lig</sub>	81.2	94.5
	N <sub>Me-imid</sub> -N <sub>Lig</sub>	95.2	90.4

Table A.7: Selected geometric data for the small model system  $\text{Pt}^{\text{II}}(\text{phen})(\text{Me-imid})_2$  where Me-imid is bound by N $\delta$  and N $\epsilon$  atoms. Data obtained DFT and LFMM methods.



DpPhen $\delta\delta$	Bond ( $\text{\AA}$ ) / Angle ( $^\circ$ )	BHandH	LFMM
	Ave. Pt-N <sub>Lig</sub>	2.014	1.958
	Ave. Pt-N <sub>Me-imid</sub>	2.008	1.951
	N <sub>Me-imid</sub> -N <sub>Me-imid</sub>	88.4	89.6
	N <sub>Lig</sub> -N <sub>Lig</sub>	80.8	93.0
	N <sub>Me-imid</sub> -N <sub>Lig</sub>	95.5	89.3

Table A.8: Selected geometric data for the small model system  $\text{Pt}^{\text{II}}(\text{phen})(\text{Me-imid})_2$  where Me-imid is bound by N $\delta$  atoms. Data obtained DFT and LFMM methods.

DpPhen $\varepsilon\varepsilon$	Bond ( $\text{\AA}$ ) / Angle ( $^\circ$ )	BHandH	LFMM
	Ave. Pt-N <sub>Lig</sub>	2.013	1.958
	Ave. Pt-N <sub>Me-imid</sub>	2.003	1.950
	N <sub>Me-imid</sub> -N <sub>Me-imid</sub>	87.9	90.5
	N <sub>Lig</sub> -N <sub>Lig</sub>	80.7	92.8
	N <sub>Me-imid</sub> -N <sub>Lig</sub>	95.8	88.68

Table A.9: Selected geometric data for the small model system  $\text{Pt}^{\text{II}}(\text{phen})(\text{Me-imid})_2$  where Me-imid is bound by N $\varepsilon$  atoms. Data obtained DFT and LFMM methods.

DpPhen $\delta\varepsilon$	Bond ( $\text{\AA}$ ) / Angle ( $^\circ$ )	BHandH	LFMM
	Ave. Pt-N <sub>Lig</sub>	2.013	1.958
	Ave. Pt-N <sub>Me-imid</sub>	2.004	1.951
	N <sub>Me-imid</sub> -N <sub>Me-imid</sub>	88.4	90.5
	N <sub>Lig</sub> -N <sub>Lig</sub>	80.8	92.9
	N <sub>Me-imid</sub> -N <sub>Lig</sub>	95.4	88.8

Table A.10: Selected geometric data for the small model system  $\text{Pt}^{\text{II}}(\text{phen})(\text{Me-imid})_2$  where Me-imid is bound by N $\delta$  and N $\varepsilon$  atoms. Data obtained DFT and LFMM methods.

DPPZ $\delta\delta$	Bond ( $\text{\AA}$ ) / Angle ( $^\circ$ )	BHandH	LFMM
	Ave. Pt-N <sub>Lig</sub>	2.020	1.962
	Ave. Pt-N <sub>Me-imid</sub>	2.006	1.950
	N <sub>Me-imid</sub> -N <sub>Me-imid</sub>	88.4	91.2
	N <sub>Lig</sub> -N <sub>Lig</sub>	81.0	93.4
	N <sub>Me-imid</sub> -N <sub>Lig</sub>	95.3	88.0

Table A.11: Selected geometric data for the small model system Pt<sup>II</sup>(phen)(Me-imid)<sub>2</sub> where Me-imid is bound by N $\delta$  atoms. Data obtained DFT and LFMM methods.

DPPZ $\varepsilon\varepsilon$	Bond ( $\text{\AA}$ ) / Angle ( $^\circ$ )	BHandH	LFMM
	Ave. Pt-N <sub>Lig</sub>	2.018	1.954
	Ave. Pt-N <sub>Me-imid</sub>	2.001	1.950
	N <sub>Me-imid</sub> -N <sub>Me-imid</sub>	88.0	90.9
	N <sub>Lig</sub> -N <sub>Lig</sub>	81.1	93.5
	N <sub>Me-imid</sub> -N <sub>Lig</sub>	95.5	88.2

Table A.12: Selected geometric data for the small model system Pt<sup>II</sup>(phen)(Me-imid)<sub>2</sub> where Me-imid is bound by N $\varepsilon$  atoms. Data obtained DFT and LFMM methods.

DPPZ $\delta\varepsilon$	Bond ( $\text{\AA}$ ) / Angle ( $^\circ$ )	BHandH	LFMM
	Ave. Pt-N <sub>Lig</sub>	2.019	1.962
	Ave. Pt-N <sub>Me-imid</sub>	2.001	1.944
	N <sub>Me-imid</sub> -N <sub>Me-imid</sub>	88.5	89.2
	N <sub>Lig</sub> -N <sub>Lig</sub>	81.1	93.5
	N <sub>Me-imid</sub> -N <sub>Lig</sub>	95.3	89.1

Table A.13: Selected geometric data for the small model system Pt<sup>II</sup>(phen)(Me-imid)<sub>2</sub> where Me-imid is bound by N $\delta$  and N $\varepsilon$  atoms. Data obtained DFT and LFMM methods.

## A.2 Appendices for Chapter 4

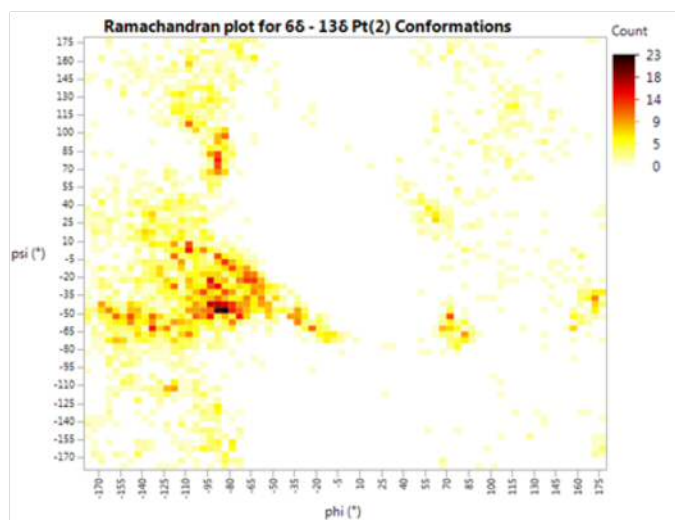


Figure A.1: Ramachandran map for the generated conformations of the His6 N $\delta$  His13 N $\delta$  binding mode of Pt<sup>II</sup>(2)-A $\beta$ 6-14

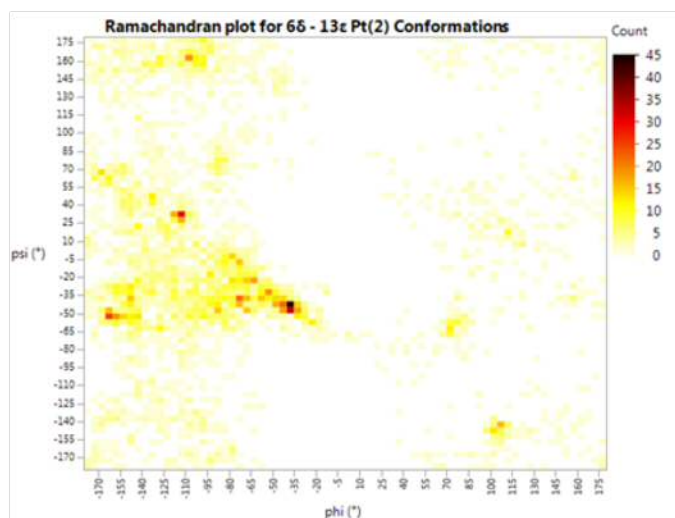


Figure A.2: Ramachandran map for the generated conformations of the His6 N $\delta$  His13 N $\epsilon$  binding mode of Pt<sup>II</sup>(2)-A $\beta$ 6-14

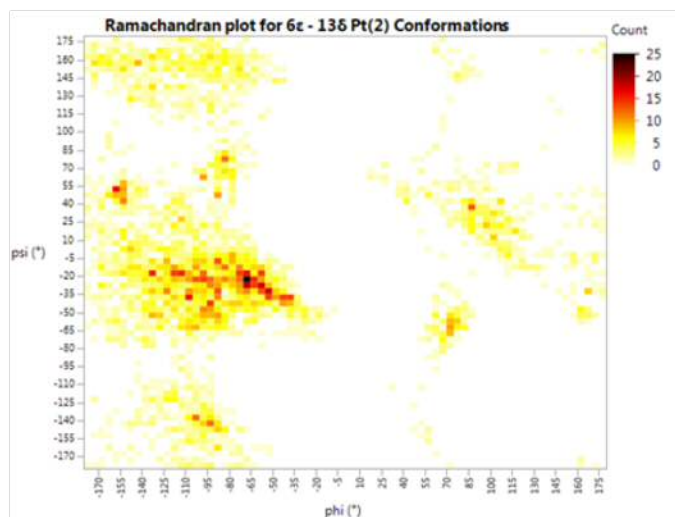


Figure A.3: Ramachandran map for the generated conformations of the His6 Nε His13 Nδ binding mode of Pt<sup>II</sup>(2)-Aβ6-14

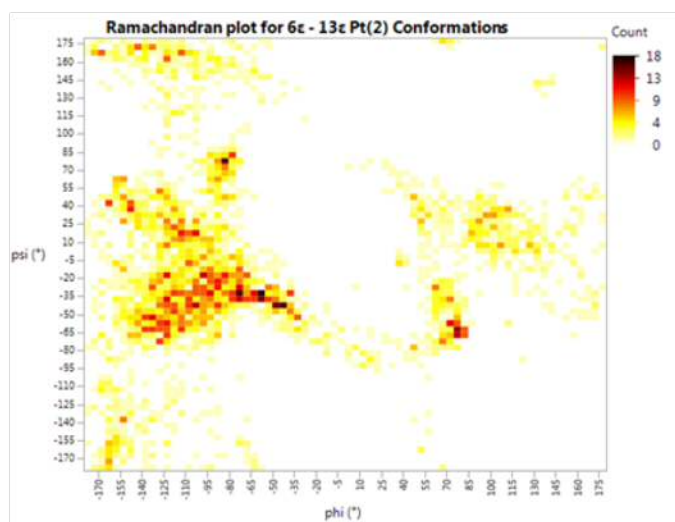


Figure A.4: Ramachandran map for the generated conformations of the His6 Nε His13 Nε binding mode of Pt<sup>II</sup>(2)-Aβ6-14

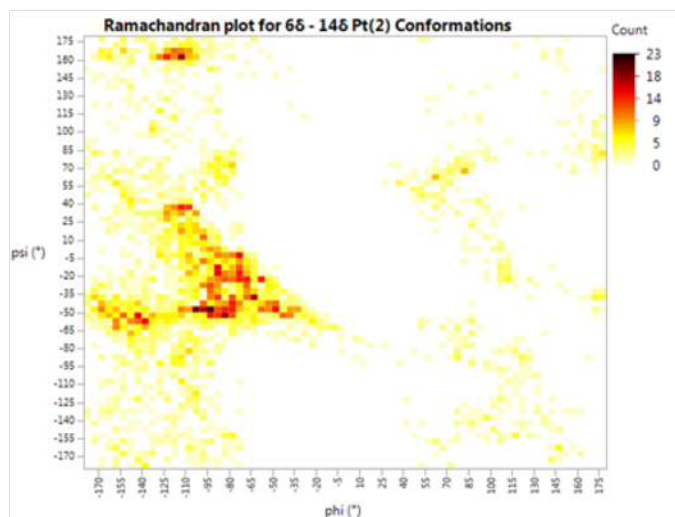


Figure A.5: Ramachandran map for the generated conformations of the His6 N $\delta$  His14 N $\delta$  binding mode of Pt<sup>II</sup>(2)-A $\beta$ 6-14

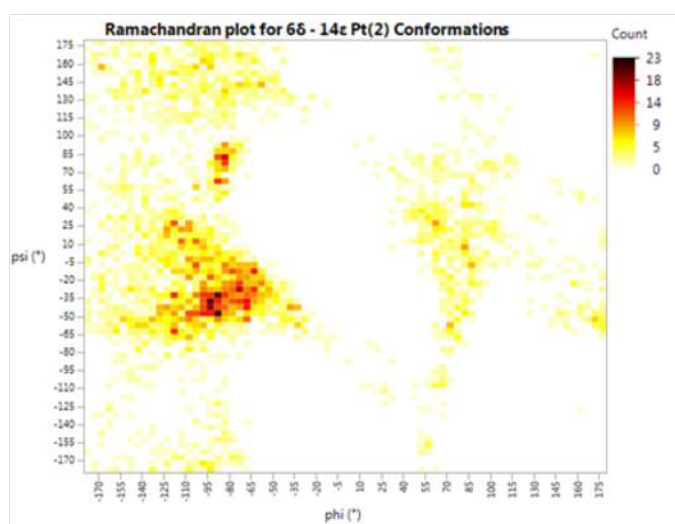


Figure A.6: Ramachandran map for the generated conformations of the His6 N $\delta$  His14 N $\epsilon$  binding mode of Pt<sup>II</sup>(2)-A $\beta$ 6-14

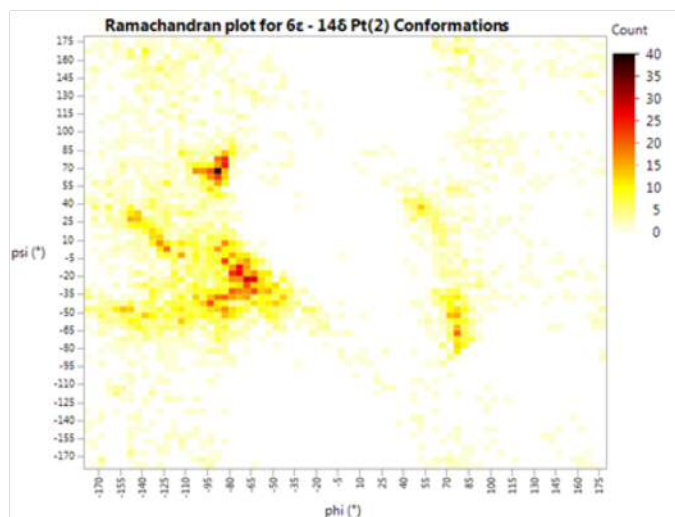


Figure A.7: Ramachandran map for the generated conformations of the His6 Nε His14 Nδ binding mode of Pt<sup>II</sup>(2)-Aβ6-14

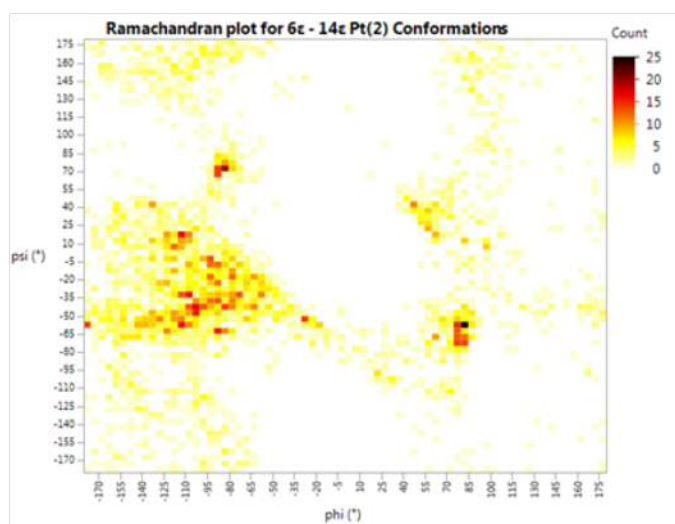


Figure A.8: Ramachandran map for the generated conformations of the His6 Nε His14 Nε binding mode of Pt<sup>II</sup>(2)-Aβ6-14

## A.3 Appendices for Chapter 5

### A.3.1 A $\beta$ 16 and Pt(A $\beta$ 16) Data

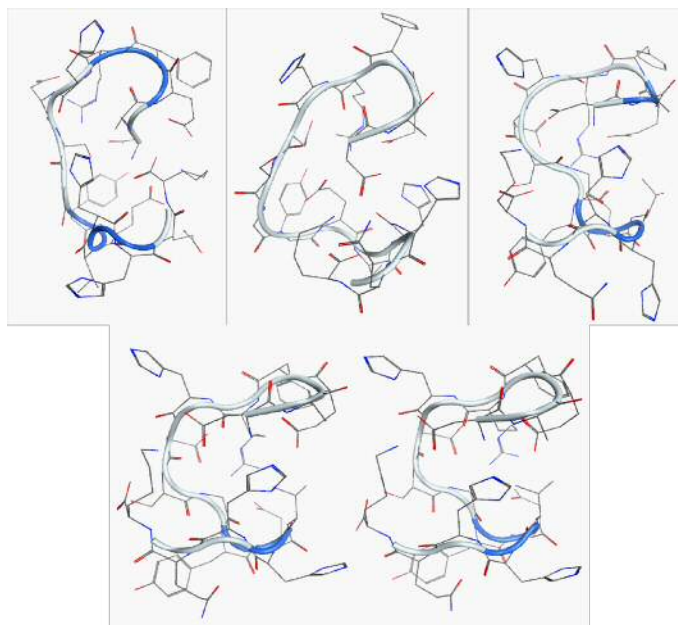


Figure A.9: Initial conformations of A $\beta$ 16 simulations. Top row: Conformations A-C. Bottom row: Conformations D-E.

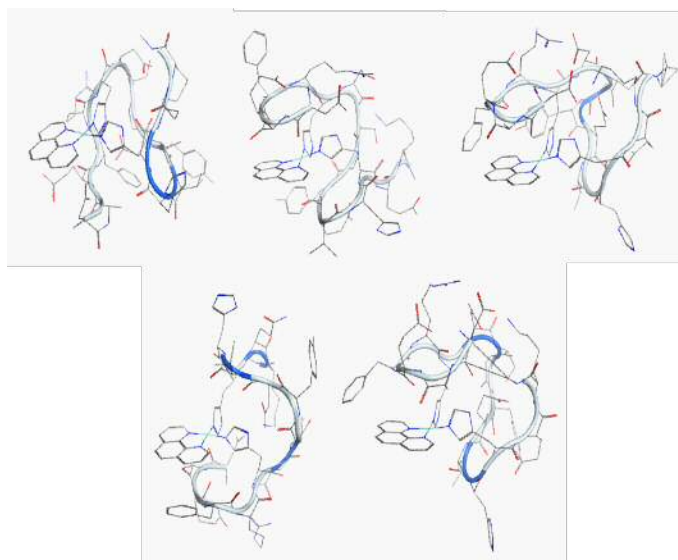


Figure A.10: Initial conformations of Pt(phen)-A $\beta$ 16 simulations. Top row: Conformations F-H Bottom row: Conformations I-J

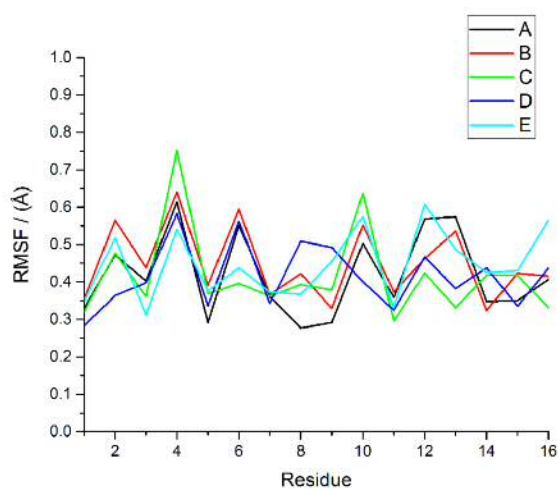
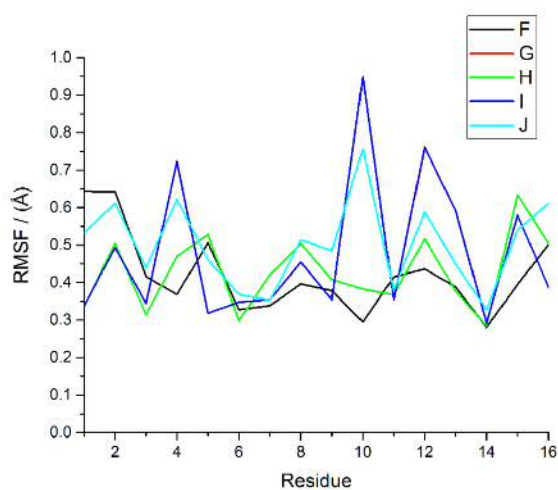


Figure A.11: RMSF data for trajectories A-E of Aβ16

Figure A.12: RMSF data for trajectories F-J of Pt<sup>II</sup>-Aβ16



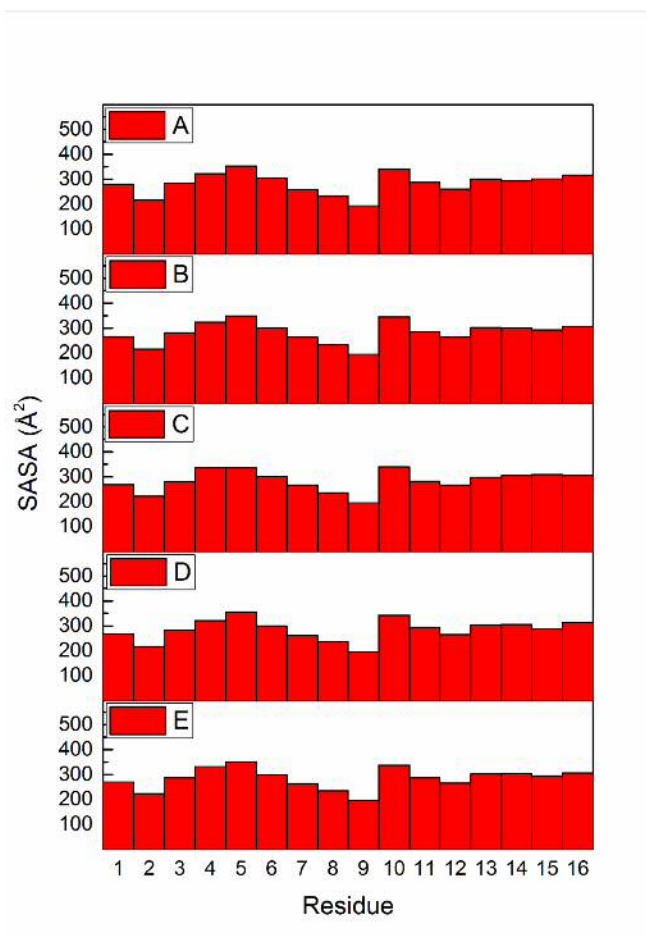


Figure A.13: SASA by residue for simulations A-E of Aβ16

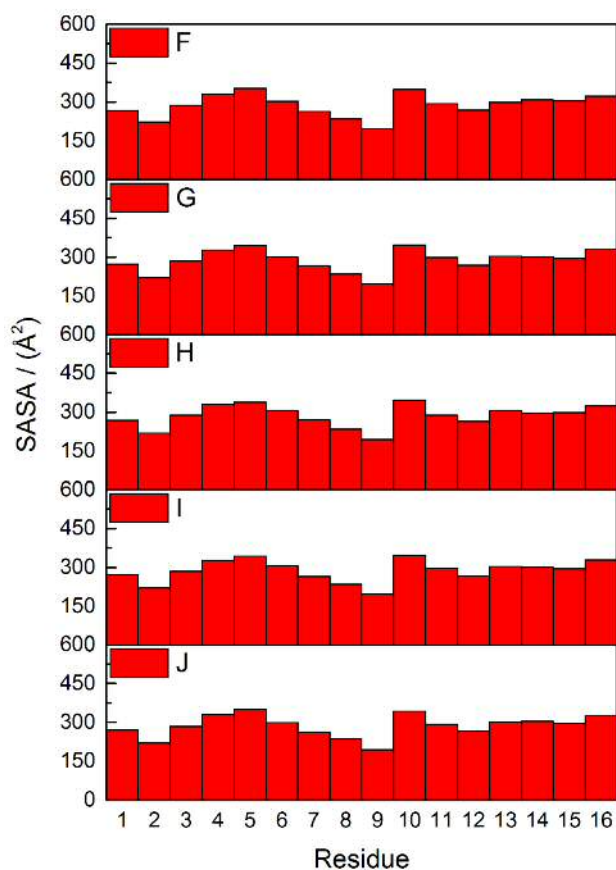


Figure A.14: SASA by residue for simulations F-J of  $\text{Pt}^{\text{II}}$ -A $\beta$ 16

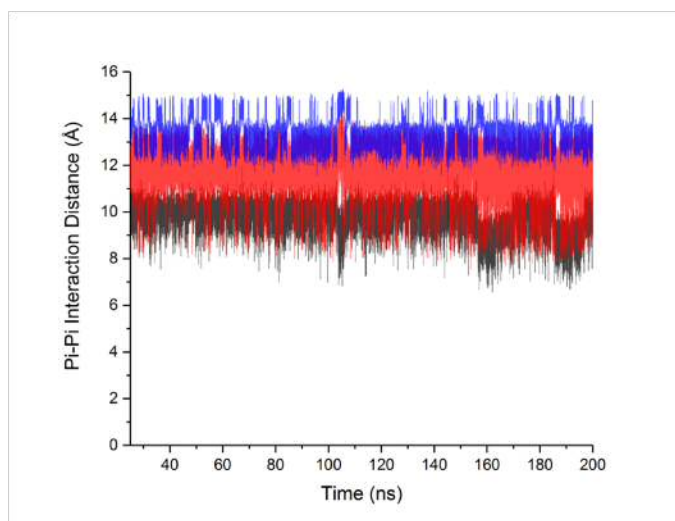


Figure A.15: Plot of  $\pi$ - $\pi$  contact distance over production MD trajectory of simulation G. Phe4-ligand distance is shown in black, Tyr10-ligand distance in red and His13-ligand distance in blue.

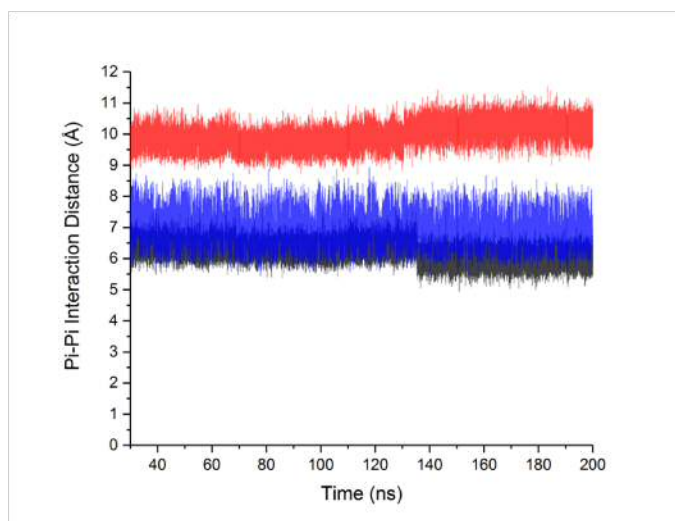


Figure A.16: Plot of  $\pi$ - $\pi$  contact distance over production MD trajectory of simulation H. Phe4-ligand distance is shown in black, Tyr10-ligand distance in red and His13-ligand distance in blue.

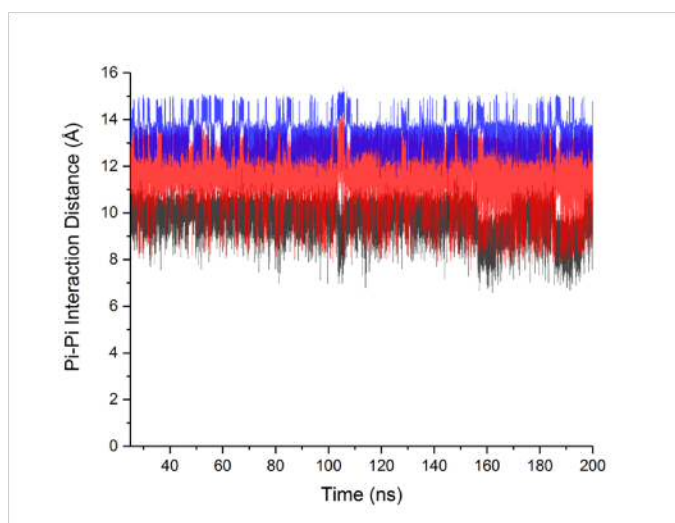


Figure A.17: Plot of  $\pi$ - $\pi$  contact distance over production MD trajectory of simulation I. Phe4-ligand distance is shown in black, Tyr10-ligand distance in red and His13-ligand distance in blue.

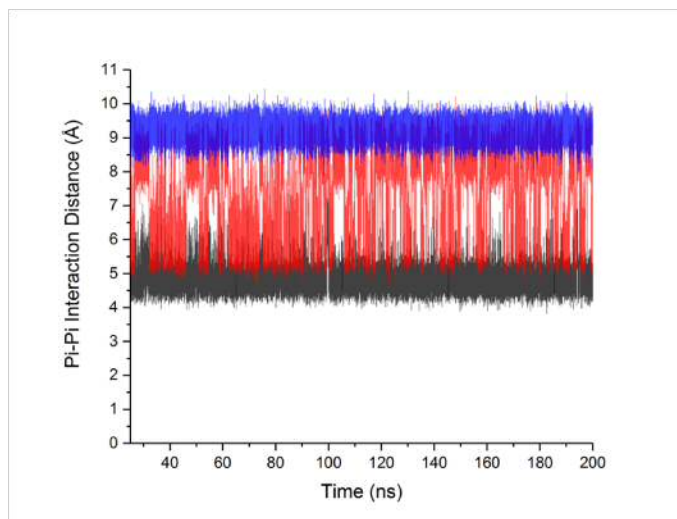


Figure A.18: Plot of  $\pi$ - $\pi$  contact distance over production MD trajectory of simulation J. Phe4-ligand distance is shown in black, Tyr10-ligand distance in red and His13-ligand distance in blue.

### A.3.2 A $\beta$ 42 and Pt(A $\beta$ 42) Data

A $\beta$ 42	Ave. RMSD / ( $\text{\AA}$ )	SD	Min.	Max.
1	4.120	0.082	3.754	4.561
2	4.973	0.134	4.513	5.428
3	19.087	0.090	18.707	19.631
4	15.786	0.111	15.201	16.221

Table A.14: Average, standard deviation, min. and max. RMSD of production MD for A $\beta$ 42 simulations.

A $\beta$ 42	Ave. RMSD / ( $\text{\AA}$ )	SD	Min.	Max.
5	10.860	0.342	8.551	12.180
6	5.891	0.073	5.498	6.179
7	10.333	0.944	8.241	12.304
8	12.555	1.087	8.704	15.230

Table A.15: Average, standard deviation, min. and max. RMSD of production MD for Pt<sup>II</sup>-A $\beta$ 42 simulations.

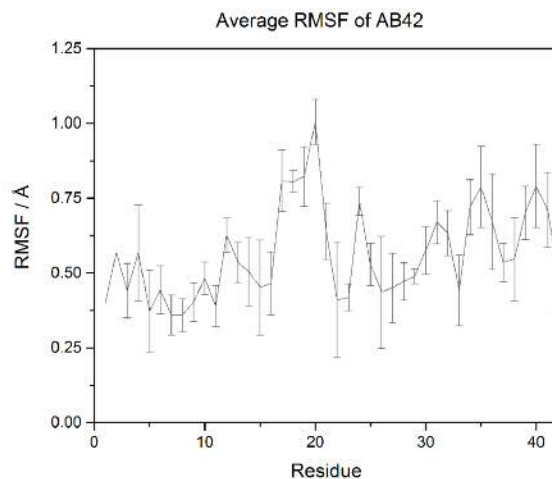


Figure A.19: Average RMSF per residue for simulations 1-4 of A $\beta$ 42

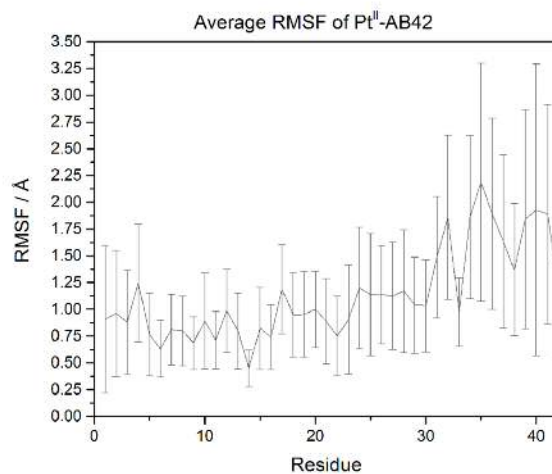


Figure A.20: Average RMSF per residue for simulations 5-8 of Pt<sup>II</sup>-A $\beta$ 42

A $\beta$ 42	Asp1	Glu3	Asp7	Glu11	Glu22	Asp23
Arg5	27.25	46.69	53.51	11.37	0.00	2.23
Lys16	48.56	0.00	0.00	52.78	48.62	0.00
Lys28	22.63	0.00	22.48	26.38	26.43	37.66

Table A.16: Combined salt-bridge data (% presence) for A $\beta$ 42 simulations

Pt <sup>II</sup> (A $\beta$ 42)	Asp1	Glu3	Asp7	Glu11	Glu22	Asp23
Arg5	85.22	37.30	0.00	0.00	27.09	31.25
Lys16	43.71	0.00	0.00	34.26	97.47	0.00
Lys28	0.00	0.00	21.26	0.00	0.00	37.50

Table A.17: Combined salt-bridge data (% presence) for Pt<sup>II</sup>-A $\beta$ 42 simulations

## A.4 Guide for LFMD in DL-POLY

This guide aims to cover how the user should set up DL-POLY Ligand Field Molecular Dynamics simulations using MOE.

- (i) Construct your system using MOE, ensuring that all metal and ligand donor atoms appear in the same chain. This can be checked using the atom manager, or via the SEQ button in the top right of the MOE main window. In most cases, this may be achieved by adding an arbitrary atom to one of the coordinating residues, then using the atom manager to convert this new atom to the metal of interest.
- (ii) Save structure as .MOE and .PDB formats, then close the system.
- (iii) Re-open the PDB file. Check atom types of metals and ligand donor atoms and correct if necessary.
- (iv) Calculate forcefield charges for all peptide atoms.
- (v) Run MT/STM *adjustcharge.svl* script to add Merz-Kollman charges to metal atoms and coordinated ligands. For novel systems, this script must be modified to include the new coordination types and appropriate Merz-Kollman charges.
- (vi) Save as .MOE file.
- (vii) Open the SVL window via the button in the top right of the MOE main window.
- (viii) In the SVL window, run *LFMM\_main*[], selecting the appropriate number of d-electrons and spin-state.
- (ix) In the SVL window, run *LFMM\_export\_lfse\_input*[].
- (x) In the SVL window, run *LFMM\_write\_dlpoly\_inp*[]

You should now have the DL-POLY FIELD file, lfse.in file and the DL-POLY CONFIG file and PDB file, with atoms in the same order, required for post-simulation analysis.

Post-simulation analysis and visualisation is possible using the Visual Molecular Dynamics (VMD) package, version 1.9.2:

- (i) In the VMD main window, click *File* → *New Molecule* and open the PDB file generated earlier in MOE using *Load*.
- (ii) Once loaded, click *File* → *New Molecule* in the main window again. In the drop-down for *Load Files For* : select the PDB file. Open the DL-POLY HISTORY file, making sure to specify the DL-POLY v2 HISTORY file type. Press *Load*. Your simulation will now open, and will include residue information such that the analysis tools in VMD may be used.