## *Structure and Activity of Antimicrobial Peptoids*

VANESSA JANE WOODHOUSE

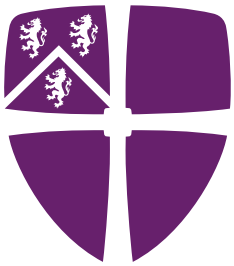
Durham University

# Structure and Activity of Antimicrobial Peptoids

Vanessa J. Woodhouse

A thesis presented for the degree of
*Doctor of Philosophy*

Supervised by Dr Elizabeth H. C. Bromley and Professor Mark R. Wilson



Soft Matter and Biological Physics
Department of Physics and Astronomy

Stockton Road
Durham
DH1 3LE

September 14, 2020

# Abstract

This thesis concerns complementary experimental and computational investigations into the relationship between the primary sequence and secondary structure of peptoids. Peptoids are a class of peptide mimetic molecules with applications as novel antimicrobial agents. The antimicrobial properties of peptoids are linked to their interactions with lipid bilayers in cell membranes, which in turn are linked to their helical secondary structure, making understanding sequence to structure relationships crucial to the design of functional sequences. Here we investigate a library of linear, cationic peptoid sequences with structural variations in the proportion and positioning of helix inducing residues and the chemical nature of the cationic side chains. We use circular dichroism spectroscopy to characterise the peptoids in aqueous and organic solvent and also to investigate structural changes upon binding to lipid bilayers designed to mimic mammalian and bacterial membranes. We present a new set of force field parameters, derived from GAFF and quantum mechanical calculations, that accurately capture the backbone torsional preferences of peptoids. Subsequently we use the modified force field to perform atomistic MD simulations of our library of peptoid sequences, using Hamiltonian replica exchange to improve sampling at less computational expense than traditional replica exchange methods.

The CD spectra reveal that the peptoids adopt characteristically helical secondary structures with variations depending on primary sequence. The intensity of helical features increases upon increasing the proportion of helix inducing residues, switching from an aqueous to an organic environment and as extra methylene groups are added to the cationic side chains, increasing their length. The length and proportion of cationic side chains also influences the folded hydrophobicity of the peptoids, though this does not correlate to their antimicrobial activity. Modelling the binding of the peptoids to lipids as a two state system enables us to estimate, in some cases, the free energy of transfer into the bilayer, where the length of the cationic side chain is also influential. MD simulations do not reveal a clear distinction in peptoid backbone conformation depending on cationic side chain length however it is clear that the peptoid backbone is more flexible and deviates more from a perfect helical conformation in aqueous than organic solvent. Ultimately these findings may aid in the rational design of new sequences.

# Acknowledgements

First and foremost my greatest thanks go to my supervisors Dr Beth Bromley and Prof. Mark Wilson for their support and guidance throughout this process. I have been extremely lucky to have two supervisors who were so generous with their time and created such an enjoyable working atmosphere.

I thank Prof. Stephen Cobb and members of his group, particularly Dr Hannah Bolt and Dr Diana Giminez, for synthesising the peptoids used in this project and also for useful discussions of results. Thanks also to Dr Martin Walker and Dr Lara Small for their help with various computational and experimental techniques as well as their general guidance and encouragement.

Thanks to all my co-inhabitants of rooms 200X, 117A and OC125, past and present, for their support and invaluable distractions at the best and worst of times. Thanks also to Will Foster and Hugo Woodhouse for their excellent proof-reading. Finally, thanks to my family for being awesome.

# Dedication

This thesis is dedicated to the memory of Dr Hilary Moorby.

# Declaration

The work presented in this thesis has been carried out under the supervision of Dr. Elizabeth H. C. Bromley of the Department of Physics and Astronomy at Durham University and Professor Mark R. Wilson of the Department of Chemistry at Durham University. All text and figures are the work of the author, unless otherwise stated. The specific contributions of other researchers to the work presented is detailed in the next section. No part of this thesis has been presented for any other degree or qualification.

# Contributions

Aside from that specified here, the work presented in this thesis is that of the author, Vanessa J. Woodhouse. The contributions of others listed here are also reiterated at the appropriate point in the main body of the text.

**Chapter 3**

- The PBS-octanol partitioning experiments to determine Log $D$ for motif 1 peptoids were carried out by Kirsten Allen, under the supervision of the author, Vanessa J. Woodhouse.
- TEM of peptoid RM5 was carried out by Dr Budhika Mendis.

**Chapter 4**

- Quantum mechanical calculations to obtain rotational energy profiles for a model peptoid were carried out by Professor Mark R. Wilson.
- Coarse graining of peptoid RM3 and investigation into its membrane activity was carried out by Ewan South.

**Chapter 5**

- HPLC times for the repeat motif sequences were collected by Hannah L. Bolt.
- HPLC times for the scrambled sequences were collected by Wataru Ichinose.
- Fluorescence anisotropy data was collected by Jamie B. Taylor.

# List of Abbreviations

**AMP** ..... Antimicrobial Peptide
**CAMP** ..... Collection of Anti-Microbial Peptides
**NMR** ..... Nuclear Magnetic Resonance
**ROS** ..... Reactive Oxygen Species
**DNA** ..... Deoxyribonucleic Acid
**GUV** ..... Giant Unilamellar Vesicles
**2D** ..... 2 Dimensional
**CD** ..... Circular Dichroism
**LD** ..... Linear Dichroism
**3D** ..... 3 Dimensional
**UV** ..... Ultra Violet
**PBS** ..... Phosphate Buffered Saline
**LUV** ..... Large Unilamellar Vesicles
**SUV** ..... Small Unilamellar Vesicles
**MRE** ..... Mean Residue Ellipticity
**HT** ..... High Tension
**DLS** ..... Dynamic Light Scattering
**DOPC** ..... 1,2-Dioleoyl-sn-Glycero-3-Phosphocholine
**DOPS** ..... 1,2-Dioleoyl-sn-Glycero-3-Phospho-L-Serine
**UV-Vis** ..... Ultra Violet-Visible
**MD** ..... Molecular Dynamics
**GROMACS** ..... Groningen Machine for Chemical Simulations
**SD** ..... Stochastic Dynamics
**PME** ..... Particle Mesh Ewald
**REMD** ..... Replica Exchange Molecular Dynamics
**HREX** ..... Hamiltonian Replica Exchange
**PLUMED** ..... PLUgin for MolEcular Dynamics
**RB** ..... Ryckaert-Bellemans
**GAFF** ..... General Amber Force Field
**QC** ..... Quantum Chemical

**DFT** ..... Density Functional Theory

**PP** ..... Polyproline

**H-bond** ..... Hydrogen Bond

**IR** ..... Infra Red

**AFM** ..... Atomic Force Microscopy

**TEM** ..... Transmission Electron Microscopy

**RM** ..... Repeat Motif

**SS** ..... Scrambled Sequence

**EDX** ..... Energy Dispersive X-Ray

**CHARMM** ..... Chemistry at Harvard Macromolecular Mechanics

**OPLS** ..... Optimized Potential for Liquid Simulations

**ACPYPE** ..... AnteChamber PYthon Parser interfacE

**RESP** ..... Restrained Electrostatic Potential

**BCC** ..... Bond Charge Corrections

**QM** ..... Quantum Mechanical

**G*N*AFF** ..... General *N*-substituted Amber Force Field

**SPC** ..... Simple Point Charge

**DMA** ..... Dimethylacetamide

**DPPC** ..... 1,2-Dipalmitoyl-sn-Glycero-3-Phosphocholine

**POPG** ..... 1-Palmitoyl-2-Oleoyl-sn-Glycero-3-Phosphoglycerol

**MIC** ..... Minimum Inhibitory Concentration

**$ED_{50}$** ..... Median Effective Dose

**$IC_{50}$** ..... Half Maximal Inhibitory Concentration

**HPLC** ..... High Performance Liquid Chromatography

**LPS** ..... Lipopolysaccharide

# Contents

# 1.0 Introduction

The work presented in this thesis is an investigation into the structural characteristics and antimicrobial activity of a particular class of molecule known as peptoids. This introductory chapter encompasses a general overview of the history and properties of peptoids and a brief summary of their many applications. Particular attention is paid to the potential use of peptoids as novel antimicrobial compounds to combat the rise of antimicrobial resistance. Antimicrobial peptoids are believed to be structurally and mechanistically analogous to naturally occurring antimicrobial peptides (AMPs). An overview of the threat of antimicrobial resistance and the potential for AMPs and similar molecules to help address the problem is given as context for the data presented in the latter chapters of this thesis.

## 1.1 Peptoids

### *1.1.1 Origins, Synthesis and Structure*

Poly-N-substituted glycines, known as peptoids, are a class of synthetic, bio-mimetic polymers. Peptoids originated in California in the late 1980s as part of a drug discovery program [1]. In the following years it became apparent that their great potential for conformational and chemical diversity would make them promising candidates not only in drug discovery but also in a wide range of other applications in medicine and biotechnology [2, 3]. Peptoids are peptidomimetics, meaning that they are designed to mimic certain aspects of the structure and function of the ubiquitous biopolymers, peptides. Peptides are sequences of amino acids linked by amide bonds between adjacent carboxyl and amino groups. Each sequence begins with a free amino group, referred to as the N-terminus and ends with a free carboxyl group, referred to as the C-terminus. Either or both of these termini may be functionalised. The particular ordering of amino acids contained within a polypeptide chain is referred to as the primary structure or sequence.

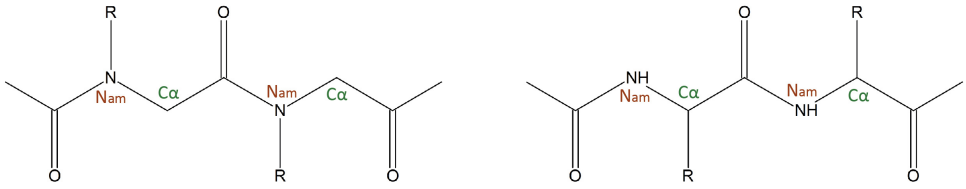



*Fig. 1.1:* Peptoid (left) vs peptide (right) backbone structure. Backbone atoms referred to as the alpha carbon and amide nitrogen are denoted as C$\alpha$ and $N_{am}$ respectively. R refers to the side chain group which can be any of the 21 natural amino acid groups in peptides and any synthetically accessible amine in peptoids.

The primary structure of peptoids is identical to that of peptides with two exceptions. Firstly, the side chain group in each peptoid unit is connected to the backbone amide nitrogen atom, as opposed to the alpha-carbon atom in peptides (Figure 1.1). Secondly, the chemical composition of the side chains in peptoids is not limited to that found in natural amino acids: peptoids can be synthesised from any synthetically accessible amine [4]. The substitution of the side chain onto the peptoid amide nitrogen results in their possessing unique structural characteristics. Notably, the absence of the amide hydrogen in peptoids prevents the formation of intra-backbone hydrogen bonding networks that are found in peptides. The peptoid backbone is additionally more flexible that its peptide counterpart, increasing the potential for conformation diversity [1, 5]. These characteristics and their consequences are discussed in detail in Chapters 3 and 4 of this thesis, which concern the structural characterisation of peptoids through a combination of experimental and computational methods.

The vast chemical diversity accessible in peptoids is possible due to the sub-monomer synthesis approach developed by Ronald Zuckermann shortly after their invention. This is a two step method where, rather than treating peptoids as sequences of N-substituted glycine units, they are treated as copolymers of alternating acetate and amine units which can be linked via successive acylation and displacement reactions [4]. This method has since been widely adopted as it holds several advantages over the alternative solid phase or solution polymerisation methods. The success of the sub-monomer method has made combinatorial library generation for peptoids extremely accessible. As a result, somewhat unusually, the field is not limited by the ability of synthetic chemists to explore the almost infinite sequence space, but rather by our understanding of the relationships between primary sequence, structure and functionality [1]. In recent years therefore, there has been considerable interest in elucidating these relationships in order to facilitate the rational design of useful peptoids [6].

Much of the functionality of peptides and proteins comes from their ability to fold into distinctive three dimensional configurations, known as secondary structures. Secondary structures can be irregular or take the form of distinctive structural elements, among which the $\alpha$-helix and $\beta$-sheet are the most common [7]. In turn, higher order structures may

form in proteins due to interactions between the secondary structural elements. Peptoids too have been observed to adopt distinctive three dimensional conformations that bear some structural similarities to those adopted by peptides. The recent interest in peptoid sequence-structure relationships has led to the discovery of several distinct helical and sheet-like structural elements which are tunable through sequence and side chain chemistry [1]. These are described and discussed in detail in the introduction to Chapter 3.

### *1.1.2 Applications*

While the rules of peptoid structure-function relationships are still being unravelled, combinatorial library screening has already led to the identification of sequences which have shown promise in a remarkable number of diverse applications. In particular there has been considerable interest in peptoid based therapeutics. Researchers have begun to explore the efficacy of peptoids in a variety of therapeutic contexts including as cell receptors, endothelial growth factors and lung surfactant mimics [8, 9]. In addition, peptoids have found application in other fields, showing promise as novel bio-compatible metal complexing and chelating agents [10–12], pH sensors [13] and building blocks for nanotechnology [14, 15]. In the latter case the unique structural diversity and tunability of peptoids means that they have the potential to play an increasingly important role in this field in the future [15].

The most notable application of peptoids in the context of this thesis is as novel antimicrobial compounds that may help tackle the rapidly growing problem of antimicrobial resistance. Peptoids can be designed to mimic naturally occurring AMPs, which themselves are of great interest, having shown considerable promise as a new class of antimicrobials [16]. The structural characteristics and mechanisms of AMPs are discussed subsequently in this introduction and the antimicrobial activity and mechanisms of peptoids are reviewed in the introduction to Chapter 5.

## 1.2 Antimicrobial Resistance

Antimicrobial resistance is considered to be one of the greatest threats to global health of the present day. The progress made over the last century in infectious disease control underpins the efficacy of all modern medicine but is currently jeopardized by the rise of resistance to conventional drugs and treatments among bacteria, viruses and parasites [17, 18]. Development of drug resistance occurs naturally and randomly through genetic mutation but has been hugely exacerbated by the widespread overuse and misuse of antibiotics and other antimicrobial compounds [19]. In some cases, particularly among gram-negative bacteria [20], this has occurred to such a great extent that multi-drug resistant microbes,

or "superbugs" have emerged as the cause of effectively untreatable infections. This results in millions of death per year worldwide [19, 21]. As well as representing an immediate threat to global health, antimicrobial resistance carries an economic and social impact as increased difficulty in fighting routine infections results in higher medical costs, increased duration of hospital stays and diminished access to effective treatments, particularly in less economically developed areas [22, 23].

The rise of antimicrobial resistance has been extremely rapid owing to the extraordinary rate at which biological evolution occurs among pathogenic species and the diversity of mechanisms by which they have developed resistance [19]. Developing the resources and practices to effectively counter this presents a significant challenge to researchers and clinicians. Currently the mechanisms known to be employed by microbes to develop drug resistance include, but are not limited to, the alteration of drug target sites, enzymatic drug inactivation, metabolic pathway modification and cell membrane alterations which reduce permeability to drug molecules [19, 24, 25]. Given the diversity and specificity of these resistance mechanisms, the search for new antimicrobials is often directed towards identifying novel broad-spectrum compounds with non-specific killing mechanisms. The aforementioned AMPs are among the molecules currently of interest due to their broad-spectrum antimicrobial activity at relatively low doses [16].

### *1.2.1 Antimicrobial Peptides*

Ubiquitous in nature, AMPs are small molecules produced by all multicellular organisms as the first line of innate immune defence [16, 26]. For this reason AMPs are also known as host-defence peptides. The properties of AMPs isolated from a wide range of organisms have been extensively studied to elucidate the physical and chemical origins of their biological activity [27, 28]. Due to the widespread interest in their properties, resources have been developed to document the findings concerning AMPs. These include the Antimicrobial Peptide Database [29] and the Collection of Anti-Microbial Peptides (CAMP) [30], which contain classification and structural data for over 2500 molecules. While AMPs are clearly a vital and effective component of the innate immune response system and have also shown potency against a wide variety of pathogens in *in vitro* studies, very few AMP based therapeutics have been approved for clinical use to date [31]. This is due to their unfortunate susceptibility to *in vivo* proteolysis, which affords them poor bio-availability and biological half life and renders them ineffective as drugs [28].

#### *Structural Characteristics of AMPs*

Though extremely diverse in origin, AMPs tend to share common structural features. AMPs are typically short peptides containing between 10 and 50 amino acids that often

include a significant proportion of lysine and/or arginine, which are positively charged at neutral pH [16]. This net cationicity is believed to play an important role in contributing to the selectivity of AMPs towards pathogenic cells [32, 33]. Many AMPs have been found to adopt amphipathic secondary structures including $\alpha$-helices, $\beta$-sheets, loops and extended structures and are often classified in this way [34]. Structure is thought to be highly influential in determining the AMPs mechanism of action [35]. Some examples highlighting the structural diversity of AMPs are shown in Figure 1.2.

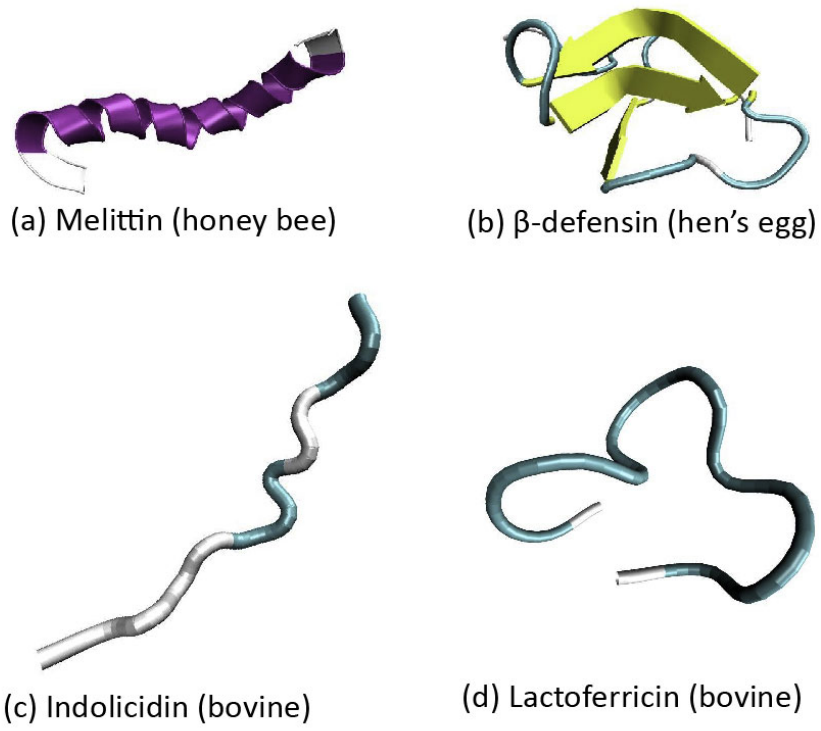



*Fig. 1.2:* Examples of the secondary structures adopted by certain AMPs in membrane environments. (a) Melittin isolated from honey bee venom adopts an $\alpha$-helical conformation [36]. (b) Hen's egg $\beta$-defensin peptides adopt a structure with a high proportion of residues in $\beta$-sheet conformation [37]. (c) Bovine indolicidin has an extended conformation [38]. (d) Bovine lactoferricin adopts a loop structure [39]. Figures show representative structures that were generated from solution NMR structures downloaded from the Protein Data Bank [40]. The PDB codes for the structures are (a) 6DST (b) 2MJK (c) 5ZVF and (d) 1Y58.

AMPs can be broadly classified according to the mechanisms by which they induce cell death, which are either membrane disruptive or non-membrane disruptive by nature. Membrane-disruptive mechanisms are particularly relevant to the work presented in this thesis, which in Chapter 5 concerns the membrane activity of antimicrobial peptoids. The

mechanisms of these peptoids are potentially analogous to those observed in studies of AMPs. These are therefore discussed in detail subsequently. Non-membrane disruptive mechanisms have been reported for AMPs against both gram-positive and gram-negative bacteria, with apoptosis (cell-death) induction after cell penetration via membrane translocation by the AMP [41]. Studies have indicated that these non-disruptive mechanisms may rely on processes such as mitochondrial dysfunction, induced by the release of reactive oxygen species (ROS) by the peptides; DNA binding and fragmentation and chromatic condensation [41–43]. Experimental techniques for elucidating these non-disruptive mechanisms include confocal and fluorescence microscopy and flow cytometry for detecting ROS accumulation and cell death *in situ* [42, 44]. In some cases non-membrane disruptive AMP activity has been shown to be cellular energy and salt concentration dependent, due to cellular uptake of the peptide being affected by these conditions [44].

### *Membrane-Active Mechanisms of Action*

The cause of cell death by membrane disruption is chiefly attributed to the leakage of the cell contents. This results in loss of control of ion concentration gradients, homeostatic imbalance and osmotic lysis [45]. The exact mechanisms of membrane disruption by AMPs have been extensively researched [46, 47] but are still somewhat debated within the field. Several distinct mechanisms have been proposed to be associated with particular AMPs. Broadly speaking these mechanisms can be divided into pore-formation mechanisms (barrel-stave and toroidal pore models) and general, non-specific membrane perturbation (the carpet model). Schematic representations of these models are shown in Figure 1.3. The overall mechanism of a particular AMP may include a combination of different effects. Understanding the balance of structural characteristics which induce these different membrane-disruptive effects is crucial to designing functional new antimicrobial molecules [47].

The barrel-stave model, which has been proposed as the mechanism of the $\alpha$-helical fungal peptide alamethicin among others, involves the formation of stable, trans-membrane pores [48]. To form these pores the AMPs bind to the membrane surface, assembling at low surface density with their hydrophilic surfaces facing each other. Individual peptides then insert into the bilayer interior and arrange with their hydrophobic face in contact with the lipid alkyl chains and their hydrophilic face facing into the centre of the pore, thus lining the pore to prevent contact between the aqueous medium and the hydrophobic lipid tail groups. It has been proposed that this requires an electrostatic interaction between the AMP and the lipid head groups, as well as the hydrophobic interaction with the lipid tails, which drives the insertion of the AMPs into the membrane interior [45, 49].

The electrostatic interaction between the AMPs and the lipid head groups is thought to be crucial in the other main proposed pore formation mechanism: the toroidal pore model.

Toroidal pores are transient pores which form when the electrostatics are sufficient to allow the AMPs to interact with the membrane such as to rearrange the structure, pulling the lipid head groups around to line pores with the AMPs and reducing the chance of contact between the aqueous medium and bilayer interior. This type of pore is proposed to relieve stress caused by the AMP insertion by inducing positive membrane curvature. Among those AMPs believed to form toroidal pores are melittin, a toxin isolated from honey bee venom [50], and magainins [51], which originate in the african clawed frog.

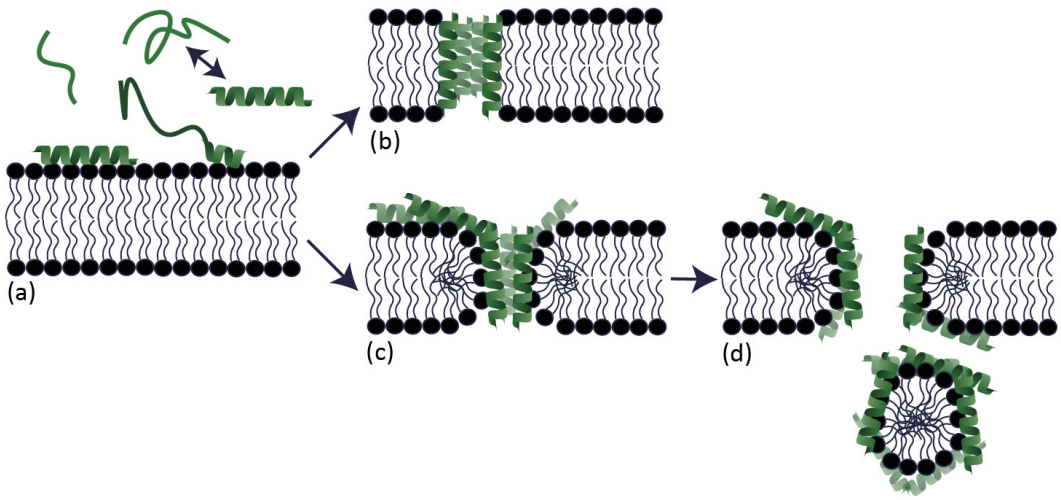



*Fig. 1.3:* Schematic representation of the membrane active mechanisms of AMPS. (a) Folding of AMPs into helical configurations and initial binding to membrane surface. (b) Formation of barrel-stave pores. (c) Formation of transient, toroidal pores. (c) Micellisation of membrane.

Pore formation requires a cooperative effect, with the AMPs reaching a critical concentration before the pores can form and the cell leakage begins to occur. Barrel-stave pores can form at relatively low concentration whereas toroidal pores often form following the mass accumulation of peptoids at the membrane surface. The accumulation of AMPs at the membrane surface is also associated with the non-specific membrane perturbation referred to as the carpet model [27]. In this mechanism the AMPs insert between lipid head groups, parallel to the plane of the membrane and cause mechanical stress, which results in some cases in the formation of toroidal pores and in others, small transient openings and even detergent-like micellisation. In this model electrostatic interactions result in the inital binding of the peptides to the membrane surface. The peptides then rearrange to orient their hydrophobic residues towards the membrane interior [27, 52].

The disruption mechanisms of membranes by AMPs are often identified by dye leakage assays using giant unilamellar vesicles (GUVs) that contain fluorescent dye as model cell systems. The rate of dye leakage from the GUVs, which is observed via fluorescence microscopy, can be analysed to determine the mechanism of action [44, 53]. Constant slow

leakage indicates non-cooperative carpet model type membrane disruption while sporadic leakage in quick bursts indicates pore formation [54]. The application of additional analytical techniques to probe the orientation of individual AMPs relative to the plane of the lipid bilayer can be used to distinguish between the formation of transient pores or stable, lined pores. For the latter to occur, AMP monomers must either be large enough to span the bilayer or interact in a way as which forms a complex to span the bilayer.

Most membrane active AMPs adopt $\alpha$-helical, amphipathic secondary structures in membrane environments. Structure-function studies aiming to elucidate the mechanisms of these AMPs have widely used lipid bilayers as model membrane systems. Methods for probing the structural changes which occur during interactions include circular dichroism (CD) spectroscopy [32, 55] and 2D proton NMR [52, 55]. Techniques such as linear dichroism (LD) spectroscopy [56] and oriented circular dichroism [48] are useful for studying the orientation of peptides in bilayers. Langmuir-Blodgett films have also been used to study AMP-membrane interactions under controlled conditions [57]. In recent years the advances made in multi-scale simulation have enabled the interactions between peptides and membranes to be studied and visualised at the molecular level [58]. This has enhanced understanding and visualisation of the mechanisms of activity and toxicity and, as computational resources continue to improve, is becoming a useful predictive tool for AMPs [59–62].

## 1.3 Statement of Aims

In this thesis we present an experimental and computational investigation into the secondary structures adopted by a select library of antimicrobial peptoids. The particular peptoids we have chosen to study were selected from an extensive combinatorial library that was designed and synthesised by members of the Cobb group at Durham University. Biological screening of this library resulted in the identification of several promising antimicrobial candidates and also revealed that seemingly minor structural rearrangements and substitutions can have a dramatic impact on the activity of a particular sequence.

The peptoids concerned in this investigation were, in some cases, selected for their notable potency against particular pathogens and in others for their unusual primary sequences containing motifs not yet discussed in detail in the literature. Through this investigation we hope to enhance our current understanding of how peptoid primary sequence relates to secondary structure and in turn how secondary structure is responsible for antimicrobial activity.

Chapter 2 concerns the experimental and computational techniques that were used to gather the data presented in this thesis. A brief discussion of the methodology is included. In subsequent chapters the results of the experimental and computational investigations

are presented. In Chapter 3 the library of peptoids upon which this work is based are introduced and characterised by circular dichroism spectroscopy. Chapter 4 comprises the results of work towards developing new simulation parameters to computationally predict the structures of peptoids in our library using molecular dynamics. Chapter 5 concerns the application of biophysical experimental techniques to studying the interactions between peptoids and model cell membranes, with the aim of expounding structure-function relationships. Relevant literature is reviewed at the beginning of each chapter in order to present the results obtained in the context of the wider field of peptoid research. Finally, the conclusions and outlook from this work are summarised in Chapter 6.

# References: Chapter 1


[1] R. N. Zuckermann, "Peptoid origins.", *Biopolymers* **2011**, *96*, 545–55.

[2] R. J. Simon, R. S. Kania, R. N. Zuckermann, V. D. Huebner, D. A. Jewell, S. Banville, S. Ng, L. Wang, S. Rosenberg, C. K. Marlowe, D. C. Spellmeyer, R. Tan, A. D. Frankel, D. V. Santi, F. E. Cohen, P. A. Bartlett, "Peptoids: A modular approach to drug discovery", *Proceedings of the National Academy of Sciences of the United States of America* **1992**, *89*, 9367–9371.

[3] P. A. Modular, I. R. J. Simon, S. Francisco, A. Bartlett, V. Daniel, S. Francisco, "United States Patent [ 19 ]", **1998**, *114*.

[4] R. N. Zuckermann, J. M. Kerr, W. H. Moosf, S. B. Kent, "Efficient Method for the Preparation of Peptoids [Oligo(N-substituted glycines)] by Submonomer Solid-Phase Synthesis", *Journal of the American Chemical Society* **1992**, *114*, 10646–10647.

[5] J. Sun, R. N. Zuckermann, "Peptoid polymers: A highly designable bioinspired material", *ACS Nano* **2013**, *7*, 4715–4732.

[6] S. A. Fowler, H. E. Blackwell, "Structure-function relationships in peptoids: recent advances toward deciphering the structural requirements for biological function.", *Organic & Biomolecular Chemistry* **2009**, *7*, 1508–1524.

[7] H. Lodish, A. Berk, S. Zipursky in *Molecular Cell Biology*, New York: W. H. Freeman, **2000**, Section 3.1, Hierarchical Structure of Proteins.

[8] N. J. Brown, J. Johansson, A. E. Barron, "Biomimicry of surfactant protein C", *Accounts of Chemical Research* **2008**, *41*, 1409–1417.

[9] M. T. Dohm, S. L. Seurynck-Servoss, J. Seo, R. N. Zuckermann, A. E. Barron, "Close mimicry of lung surfactant protein B by "clicked" dimers of helical, cationic peptoids.", *Biopolymers* **2009**, *92*, 538–553.

[10] M. Baskin, H. Zhu, Z. W. Qu, J. H. Chill, S. Grimme, G. Maayan, "Folding of unstructured peptoids and formation of hetero-bimetallic peptoid complexes upon side-chain-to-metal coordination", *Chemical Science* **2019**, *10*, 620–632.

[11] G. Maayan, M. D. Ward, K. Kirshenbaum, "Metallopeptoids", *Chemical Communications* **2009**, 56–58.

[12] M. Baskin, G. Maayan, "Water-soluble chiral metallopeptoids", *Biopolymers* **2015**, *104*, 577–584.

[13] A. A. Fuller, C. A. Holmes, F. J. Seidl, "A fluorescent peptoid pH-sensor", *Biopolymers* **2013**, *100*, 380–386.

[14] J. Sun, X. Jiang, R. Lund, K. H. Downing, N. P. Balsara, R. N. Zuckermann, "Self-assembly of crystalline nanotubes from monodisperse amphiphilic diblock copolypeptoid tiles", *Proceedings of the National Academy of Sciences of the United States of America* **2016**, *113*, 3954–3959.

[15] J. Sun, R. N. Zuckermann, "Peptoid polymers: A highly designable bioinspired material", *ACS Nano* **2013**, *7*, 4715–4732.

[16] R. E. W. Hancock, T. J. Falla, "Antimicrobial peptides: Broad-spectrum antibiotics from nature", *Clinical Microbiology and Infection* **1996**, *1*, 226–229.

[17] WHO, "Antimicrobial resistance.", *Bulletin of the World Health Organization* **2014**, *61*, 383–94.

[18] W. Wang, M. I. Arshad, M. Khurshid, M. H. Rasool, M. A. Nisar, M. A. Aslam, M. U. Qamar, "Antibiotic resistance : a rundown of a global crisis", *Infection and Drug Resistance* **2018**, 1645–1658.

[19] J. Davies, D. Davies, "Origins and evolution of antibiotic resistance", *Microbiology and Molecular Biology Reviews* **2010**, *74*, 417–433.

[20] E. M. Eichenberger, J. T. Thaden, "Epidemiology and mechanisms of resistance of extensively drug resistant gram-negative bacteria", *Antibiotics* **2019**, *8*.

[21] G. Ippolito, S. Leone, F. N. Lauria, E. Nicastri, R. P. Wenzel, "Methicillin-resistant Staphylococcus aureus: the superbug", *International Journal of Infectious Diseases* **2010**, *14*, 7–11.

[22] J. E. McGowan, "Economic impact of antimicrobial resistance.", *Emerging infectious diseases* **2001**, *7*, 286–92.

[23] M. Ahmad, A. U. Khan, "Global economic impact of antibiotic resistance: A review", *Journal of Global Antimicrobial Resistance* **2019**, *19*, 313–316.

[24] J. M. Munita, C. A. Arias, "Mechanisms of Antibiotic Resistance", *Microbiology Spectrum* **2016**, *4*, 1–37.

[25] A. H. Holmes, L. S. Moore, A. Sundsfjord, M. Steinbakk, S. Regmi, A. Karkey, P. J. Guerin, L. J. Piddock, "Understanding the mechanisms and drivers of antimicrobial resistance", *The Lancet* **2016**, *387*, 176–187.

[26] M. Zasloff, "Antimicrobial peptides of multicellular organisms", *Nature* **2002**, *415*, 389–395.

[27] R. M. Epand, H. J. Vogel, "Diversity of antimicrobial peptides and their mechanisms of action", *Biochimica et Biophysica Acta - Biomembranes* **1999**, *1462*, 11–28.

[28] Y. J. Gordon, E. G. Romanowski, A. M. McDermott, "A review of antimicrobial peptides and their therapeutic potential as anti-infective drugs.", *Current Eye Research* **2005**, *30*, 505–15.

[29] G. Wang, X. Li, Z. Wang, "APD3: The antimicrobial peptide database as a tool for research and education", *Nucleic Acids Research* **2016**, *44*, 1087–1093.

[30] F. H. Waghu, L. Gopi, R. S. Barai, P. Ramteke, B. Nizami, S. Idicula-Thomas, "CAMP: Collection of sequences and structures of antimicrobial peptides", *Nucleic Acids Research* **2014**, *42*, 1154–1158.

[31] J. Lei, L. C. Sun, S. Huang, C. Zhu, P. Li, J. He, V. Mackey, D. H. Coy, Q. Y. He, "The antimicrobial peptides and their potential clinical applications", *American Journal of Translational Research* **2019**, *11*, 3919–3931.

[32] E. Glukhov, M. Stark, L. L. Burrows, C. M. Deber, "Basis for selectivity of cationic antimicrobial peptides for bacterial versus mammalian membranes", *Journal of Biological Chemistry* **2005**, *280*, 33960–33967.

[33] K. Matsuzaki, "Control of cell selectivity of antimicrobial peptides", *Biochimica et Biophysica Acta - Biomembranes* **2009**, *1788*, 1687–1692.

[34] J. Koehbach, D. J. Craik, "The Vast Structural Diversity of Antimicrobial Peptides", *Trends in Pharmacological Sciences* **2019**, *40*, 517–528.

[35] T.-H. Lee, K. N. Hall, M.-I. Aguilar, "Antimicrobial Peptide Structure and Mechanism of Action: A Focus on the Role of Membrane Structure", *Current Topics in Medicinal Chemistry* **2015**, *16*, 25–39.

[36] L. S. Ramirez, J. Pande, A. Shekhtman, "Helical Structure of Recombinant Melittin", *Journal of Physical Chemistry B* **2019**, *123*, 356–368.

[37] V. Hervé, H. Meudal, V. Labas, S. Réhault-Godbert, J. Gautron, M. Berges, N. Guyot, A. F. Delmas, Y. Nys, C. Landon, "Three-dimensionalnmrstructure of hen egg gallin (chicken ovodefensin) reveals a new variation of the $\beta$-defensin fold", *Journal of Biological Chemistry* **2014**, *289*, 7211–7220.

[38] R. Dwivedi, P. Aggarwal, N. S. Bhavesh, K. J. Kaur, "Design of therapeutically improved analogue of the antimicrobial peptide, indolicidin, using a glycosylation strategy", *Amino Acids* **2019**, *51*, 14443–1460.

[39] L. T. Nguyen, D. J. Schibli, H. J. Vogel, "Structural studies and model membrane interactions of two peptides derived from bovine lactoferricin", *Journal of Peptide Science* **2004**, *11*, 379–389.

[40] H. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. Shindyalov, P. Bourne, "The Protein Data Bank", *Oxford University Press Nucleic Acids Research* **2000**, *28*, 235–242.

[41] J. Lee, D. G. Lee, "Antimicrobial peptides (AMPs) with dual mechanisms: Membrane disruption and apoptosis", *Journal of Microbiology and Biotechnology* **2014**, *25*, 759–764.

[42] B. Hwang, J. S. Hwang, J. Lee, J. K. Kim, S. R. Kim, Y. Kim, D. G. Lee, "Induction of yeast apoptosis by an antimicrobial peptide, Papiliocin", *Biochemical and Biophysical Research Communications* **2011**, *408*, 89–93.

[43] J. H. Hwang, I. S. Hwang, Q. H. Liu, E. R. Woo, D. G. Lee, "(+)-Medioresinol leads to intracellular ROS accumulation and mitochondria-mediated apoptotic cell death in Candida albicans", *Biochimie* **2012**, *94*, 1784–1793.

[44] J. Lee, J. S. Hwang, I. S. Hwang, J. Cho, E. Lee, Y. Kim, D. G. Lee, "Coprisin-induced antifungal effects in Candida albicans correlate with apoptotic mechanisms", *Free Radical Biology and Medicine* **2012**, *52*, 2302–2311.

[45] Y. Shai, "Mode of action of membrane active antimicrobial peptides", *Biopolymers - Peptide Science Section* **2002**, *66*, 236–248.

[46] B. Bechinger, S. U. Gorr, "Antimicrobial Peptides: Mechanisms of Action and Resistance", *Journal of Dental Research* **2017**, *96*, 254–260.

[47] J. Li, J. J. Koh, S. Liu, R. Lakshminarayanan, C. S. Verma, R. W. Beuerman, "Membrane active antimicrobial peptides: Translating mechanistic insights to design", *Frontiers in Neuroscience* **2017**, *11*, 1–18.

[48] S. Qian, C. Wang, L. Yang, H. W. Huang, "Structure of the alamethicin pore reconstructed by x-ray diffraction analysis", *Biophysical Journal* **2008**, *94*, 3512–3522.

[49] Y. Shai, "Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by alpha-helical antimicrobial and cell non-selective membrane-lytic peptides", *Biochimica et Biophysica Acta - Biomembranes* **1999**, *1462*, 55–70.

[50] K. Matsuzaki, S. Yoneyama, K. Miyajima, "Pore formation and translocation of melittin", *Biophysical Journal* **1997**, *73*, 831–838.

[51] K. Matsuzaki, K. I. Sugishita, N. Ishibe, M. Ueha, S. Nakata, K. Miyajima, R. M. Epand, "Relationship of membrane curvature to the formation of pores by magainin 2", *Biochemistry* **1998**, *37*, 11856–11863.

[52] J. P. S. Powers, R. E. W. Hancock, "The relationship between peptide structure and antibacterial activity", *Peptides* **2003**, *24*, 1681–1691.

[53] N. B. Leite, A. Aufderhorst-Roberts, M. S. Palma, S. D. Connell, J. R. Neto, P. A. Beales, "PE and PS Lipids Synergistically Enhance Membrane Poration by a Peptide with Anticancer Properties", *Biophysical Journal* **2015**, *109*, 936–947.

[54] A. S. Ladokhin, W. C. Wimley, S. H. White, "Leakage of membrane vesicle contents: determination of mechanism using fluorescence requenching.", *Biophysical journal* **1995**, *69*, 1964–71.

[55] S. Rex, "Pore formation induced by the peptide melittin in different lipid vesicle membranes", *Biophysical Chemistry* **1996**, *58*, 75–85.

[56] C. E. Caesar, E. K. Esbjörner, P. Lincoln, B. Nordén, "Membrane interactions of cell-penetrating peptides probed by tryptophan fluorescence and dichroism techniques: Correlations of structure to cellular uptake", *Biochemistry* **2006**, *45*, 7682–7692.

[57] S. Li, M. Micic, J. Orbulescu, J. D. Whyte, R. M. Leblanc, "Human islet amyloid polypeptide at the air-aqueous interface: A Langmuir monolayer approach", *Journal of the Royal Society Interface* **2012**, *9*, 3118–3128.

[58] R. Pokhrel, N. Bhattarai, P. Baral, B. S. Gerstman, J. H. Park, M. Handfield, P. P. Chapagain, "Molecular mechanisms of pore formation and membrane disruption by the antimicrobial lantibiotic peptide Mutacin 1140", *Physical Chemistry Chemical Physics* **2019**, *21*, 12530–12539.

[59] B. A. Hall, A. P. Chetwynd, M. S. Sansom, "Exploring peptide-membrane interactions with coarse-grained MD simulations", *Biophysical Journal* **2011**, *100*, 1940–1948.

[60] Z. L. Li, H. M. Ding, Y. Q. Ma, "Interaction of peptides with cell membranes: Insights from molecular modeling", *Journal of Physics Condensed Matter* **2016**, *28*.

[61] S. Liu, J. Bao, X. Lao, H. Zheng, "Novel 3D Structure Based Model for Activity Prediction and Design of Antimicrobial Peptides", *Scientific Reports* **2018**, *8*, 1–12.

[62] A. Catte, M. R. Wilson, M. Walker, V. S. Oganesyan, "Antimicrobial action of the cationic peptide, chrysophsin-3: A coarse-grained molecular dynamics study", *Soft Matter* **2018**, *14*, 2796–2807.

# 2.0 Methodology and Technical Details

In this chapter we describe the experimental and computational techniques used to collect the data presented in this thesis. The theoretical basis of each technique is briefly described, with particular detail given where relevant to the results obtained. The technical details associated with the particular experiments and simulations described in latter chapters are then outlined.

## 2.1 Experimental Methods

### *2.1.1 Circular Dichroism Spectroscopy*

Circular dichroism (CD) is an optical phenomenon which arises due to the differential absorption of circular polarized light by systems with intrinsic or extrinsic chirality.

Light is said to be circularly polarized when its electric field vector rotates around the propagation direction of the beam while maintaining constant magnitude, tracing out a full circle over the course of one period of the wave. Since the light is propagating in a particular direction the electric field vector ultimately traces out a helix in the direction of propagation. If an absorbing, optically active sample is placed in the path of the circularly polarized light it will absorb the left and right handed components to different extents. To be optically active and produce this effect a sample must be chiral, i.e. non superimposable on its own mirror image, and therefore have a handedness.

CD spectroscopy is an analytical technique in which the differential absorption, $\Delta A$, of the left ($A_{\mathrm{L}}$) and right ($A_{\mathrm{R}}$) handed components of circularly polarised light by a chiral system is measured.

$$CD = \Delta A = A_{\mathrm{L}} - A_{\mathrm{R}} \tag{2.1}$$

Practically, a beam of circularly polarised light may be radiated into the sample of interest via alternate pulses of left and right handed polarized light of equal magnitude. The sample, if chiral, will absorb the two polarizations to different extents over a range of wavelengths, yielding an output of light which is now elliptically polarized. This is recorded as the CD spectrum. Prior to collecting a CD spectrum, purging of the system

with nitrogen is necessary to prevent the distortion of spectra and formation of ozone by oxygen absorbing light at wavelengths of around 200 nm. Care must also be taken in selecting solvents and buffers for these measurements to avoid those containing absorbing species which will distort the spectra.

### *Application of CD Spectroscopy to the Study of Biomolecules*

CD spectroscopy is widely applied to the study of secondary structure in biomolecules such as proteins, peptides and DNA. All natural amino acids, with the exception of glycine, are chiral and therefore the majority of peptides are chiral. However, in the far-UV region (180-240 nm) the folding of peptides into certain 3D structures results in an additional chirality that is conformational rather than configurational, meaning the CD spectrum varies due the arrangement of the atoms in 3D space rather than just the presence of certain atoms and bonds. The peptide carbonyl bond is the main chromophore in the far-UV region and the electronic transitions associated with it are depicted in Figure 2.1. Conformational changes affect the transition energies and therefore the shape and intensity of peptide CD spectra. Given that the peptoid backbone is structurally very similar to the peptide backbone, peptoids may also exhibit a conformation dependent CD signal in the far-UV.

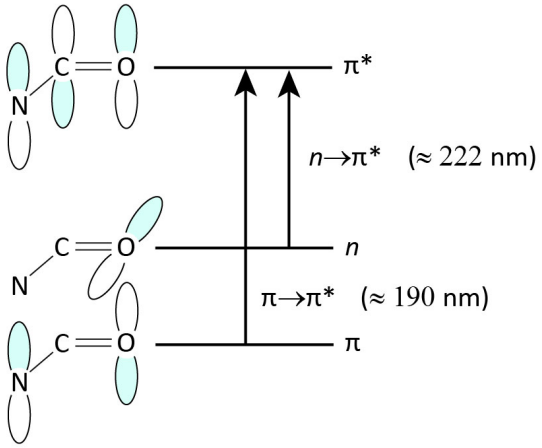



*Fig. 2.1:* Electronic transitions of the carbonyl bond in a peptide/peptoid backbone amide.

Since the CD spectrum and transition energies are sensitive to the particular conformation of the peptide backbone, particular secondary structural motifs have distinctive CD spectral features associated with them, examples of which are depicted in Figure 2.2. The CD spectra of proteins are often analysed by comparison to standard reference spectra [1, 2]. There are commercially available spectral deconvolution tools which are designed to elucidate the overall percentage of the sample adopting particular structural conformations by fitting the experimental spectra as combinations of reference spectra assigned to each particular structural motif [3–5]. This type of analysis has been made increasingly

accurate for peptides and proteins but is potentially misleading as it is not necessarily true that the spectrum corresponding to a particular conformation would be identical for different peptide sequences [6]. In the case of peptoids it is particularly unlikely that the reference spectra could be meaningfully related to any measured spectra and therefore this style of analysis is generally avoided in peptoid studies, including this thesis.

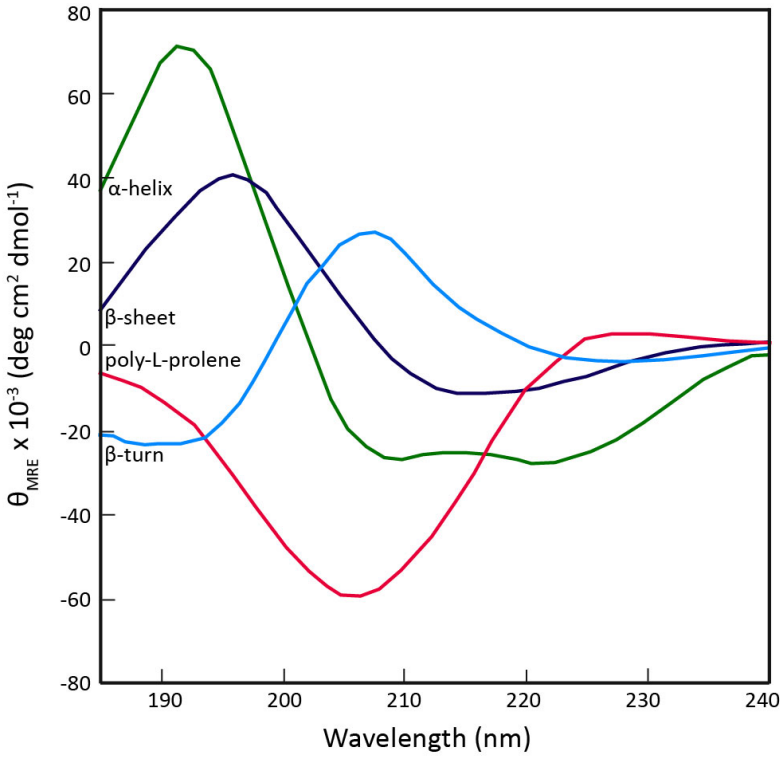



*Fig. 2.2:* Example CD spectra for polypeptides in different secondary structural conformations. Typical spectrum for $\alpha$-helix shown in green, $\beta$-sheet in dark blue, $\beta$-turn in light blue and poly-L-proline in red. Spectra redrawn with permission from Brahms and Brahms [7] and Greenfield [2].

CD is still extremely useful in the study of peptoid structure as it can be used to detect conformational differences between sequences and monitor conformational changes in different solvent environments, at variable temperature and upon binding to lipid vesicles, among other examples. In Chapter 3 we use CD spectroscopy to characterise a library of peptoids by directly comparing the spectra of different sequences and in Chapter 5 we are able to investigate the binding of the same peptoids to lipid vesicles, as the CD is sensitive to the conformational changes which the peptoids undergo during this process. One particularly beneficial aspect of CD is that it allows us to probe the structure of molecules in solution, unlike X-ray crystallography, which is also readily used to study protein structure [8]. Additionally, CD is not limited to small molecular fragments like other solution characterisation techniques. The use of NMR for peptoid structural characterisation is more complex than the equivalent analysis of peptides, due to the lack of amide protons in the peptoid backbone [9]. Attempts to characterise peptoids by solution NMR have

often been hampered by poor peak dispersion and considerable spectral overlap [10].

### *Experimental Details*

The CD spectra presented in this thesis were collected using a Jasco J-1500 spectrophotometer. Peptoid samples were dissolved in 0.01 M phosphate buffered saline (PBS) solution or octanol, at a concentration of 25 uM unless specified otherwise. The samples were analysed in 1 mm path length cuvettes made of high quality quartz suprasil and supplied by Hellma UK Ltd.

Spectra were collected across a range of wavelengths from 190-260 nm at a scanning speed of 50 nm per minute. The bandwidth was 3 nm and digital integration time (the time over which photons were collected by the detector at each wavelength before processing the signal) was set at 8 s to maximise the signal to noise ratio. In the majority of cases the displayed spectra are the average of 15 accumulations, with the exception of the spectra collected during thermal melts which are the average of 3 accumulations.

Unless otherwise stated, measurements were conducted at 20 °C. The temperature of the sample cell was controlled by a Peltier circulating water bath system. In the case of peptoid melts, where spectra were collected over a range of temperatures, the temperature of the system was increased at a rate of 1 °C per minute and the temperature ramping was halted during the collection of spectra.

The Jasco J-1500 reports the CD signal as ellipticity in units of degrees. As is convention, the spectra displayed in this thesis have been converted into units of mean residue ellipticity ($\Theta_{\mathrm{MRE}}$), which normalises the CD signal for both peptoid concentration and sequence length according to equation 2.2, where $\Theta$ is the ellipticity in degrees, $c$ is the peptoid concentration, $l$ is the path length and $n_{\mathrm{R}}$ is the number of residues in the peptoid sequence.

$$\Theta_{\mathrm{MRE}} = \frac{100\Theta}{cl(n_{\mathrm{R}} - 1)} \tag{2.2}$$

### *Discussion of Errors Associated with CD Spectra of Peptoids*

Error bars are rarely displayed on CD spectra as they may obscure figures. However, as with any experimental technique, there is error associated with each measurement. In order to minimise instrumentation error the scan parameters discussed previously were selected to maximise the signal to noise ratio. These parameters were used consistently throughout the work so the signal to noise relating to operator choice should remain constant for all spectra collected, though it may vary due to other factors such as solvent conditions. Error in sample concentration is generally reported to be the most significant source of error in CD spectra [11].

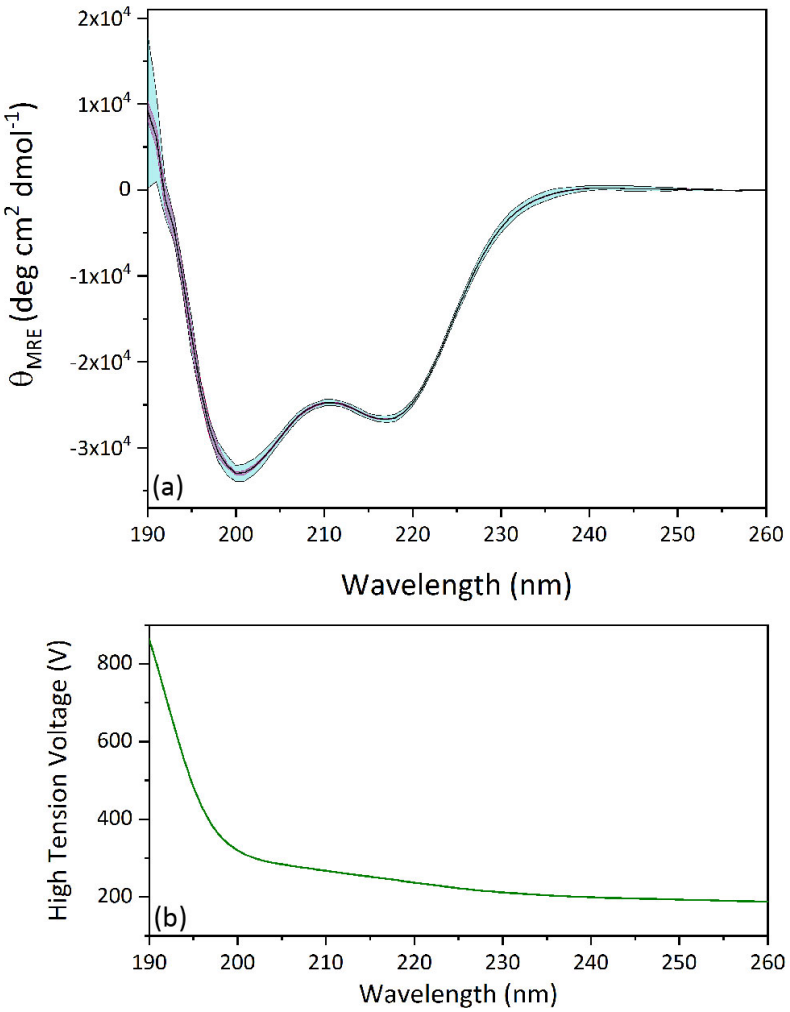



*Fig. 2.3:* Representation of error associated with CD spectra for peptoids. The spectrum shown in (a) is for $(N\text{ae}N\text{spe})_6$ at 25 $\mu$M concentration in 0.01 M PBS. Average spectrum shown as black line. Standard deviation for accumulated spectra from a single sample shown as pink band. Standard deviation for spectra from multiple samples shown as blue band. (b) Shows the average high tension voltage trace for the spectra shown in (a).

Figure 2.3a demonstrates the typical magnitude of error associated with the CD spectra collected from the peptoids discussed in this thesis. The spectrum shown as a black line is the average for one of the peptoids from our library, $(N\text{ae}N\text{spe})_6$, at a concentration of 25 $\mu$M in 0.01 M PBS. Our peptoid library is introduced in full in Chapter 3. The coloured bands extending from the spectrum in Figure 2.3 represent the magnitude of error typically associated with our measurements. The pink band shows the error associated with collecting 15 spectral accumulations from the same sample. This band extends 1 standard deviation away from the mean spectrum in both directions. The blue band is representative of the error associated with preparing samples at different concentrations (or separate samples at the same concentration) and taking spectra of each of these samples. This band extends 1 standard deviation away from the average of 4 spectra. The

standard deviations from the mean spectrum for the separate samples are small and for the accumulations, even smaller. This indicates that these spectra can reliably be collected in a repeatable fashion.

The larger deviations from the mean at low wavelengths are associated with the absorbance of light by chloride ions present in the PBS and for the sake of analysis the spectrum is not considered over this very low wavelength range in the work described in this thesis (typically, $\lambda < 200$ nm is discarded). Decisions to discard data in certain wavelength ranges are are based on the high tension (HT) voltage. Figure 2.3b shows the average HT voltage trace for the spectra shown in Figure 2.3a. The HT voltage varies during a measurement in order to amplify the detector sensitivity. As the wavelength and the absorbance of light by the sample change, the HT voltage is adjusted to ensure enough photons reach the detector to produce a strong signal relative to the noise generated by stray light. The HT voltage is therefore a useful parameter to monitor to ensure the collection of high quality spectra. Generally data is discarded where the HT voltage is above 600 V.

The behaviour shown in Figure 2.3 is representative of all the peptoids investigated experimentally in this thesis, with the exception of one sequence which is discussed in detail in Chapter 3.

### *CD Studies Using Vesicles as Model Membranes*

We carried out a series of experiments using CD to monitor the changes in peptoid secondary structure upon binding to lipid vesicles. The results of these experiments are presented and discussed in Chapter 5. We used large unilamellar vesicles (LUVs), approximately 100 nm in diameter, as the model membrane systems. In early CD studies of lipid bound peptides, tip-sonicated small unilamellar vesicles (SUVs) [12] (diameter $\approx$ 30 nm) were initially used for preference, in order to minimise the amount of light scattering due to their presence and hence minimise the distortion of the true peptide signal. However, SUVs are metastable structures and not suited to thermodynamic studies [13, 14]. In addition their high curvature has been shown to cause unreliable results in peptide partitioning experiments [15, 16]. Ladhokin *et al.* demonstrated that LUVs (diameter $\approx$ 100 nm) do not cause significant scattering so as to distort the spectra of peptides bound to them and thus are suitable for use in CD experiments at concentrations of up to 3 mM at wavelengths longer than 200 nm [17].

In this work we used LUVs at a range of lipid concentrations from 25-2500 $\mu$M. The HT traces corresponding to spectra for samples with LUVs at a range of concentrations are shown in Figure 2.4. As the lipid concentration increases the HT voltage rises above the acceptable signal to noise threshold of 600 V at longer wavelengths. In order to collect spectra across a useful range of wavelengths we used a maximum lipid concentration of

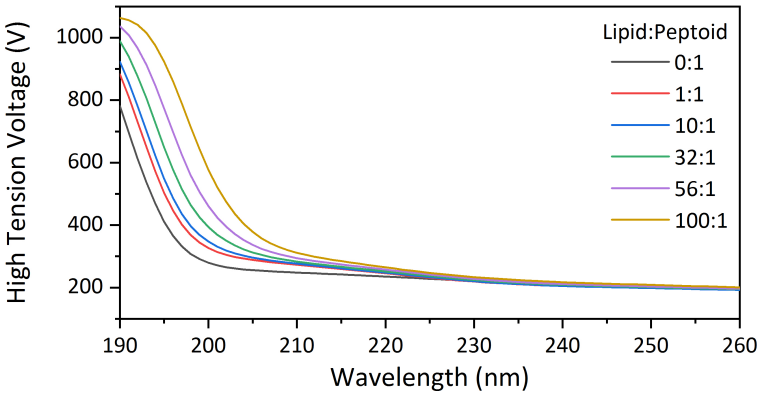



*Fig. 2.4:* HT voltage traces corresponding to CD spectra for peptoids with LUVs at increasing concentration. Peptoid concentration was 15 $\mu$M in each case and the lipid to peptoid ratio varies from 0:1 to 100:1 as detailed in the legend.

2500 $\mu$M. At this lipid concentration the HT voltage rose above 600 V at approximately 200 nm wavelength.

### *LUV Production Protocol*

We produced the LUVs for our study by the extrusion method [18]. Lipids dissolved in chloroform were obtained from Avanti Polar Lipids. To produce each batch of LUVs a calculated amount of lipid solution was dried under nitrogen and then under vacuum for a minimum of 3 hours to ensure all traces of chloroform had evaporated. The dry lipid sample was then rehydrated in 0.01 M PBS, to the desired concentration. The sample was bath sonicated for 10 minutes to promote the formation of a polydisperse solution of multilamellar vesicles. This solution was then extruded through a polycarbonate membrane, with pore size 50 nm, a minimum of 17 times until LUVs were formed. This was apparent as the previously cloudy solution became clear. This is a well established method for the production of LUVs [18, 19]. The size and polydispersity of vesicles produced in this way varies depending on the concentration and composition of the lipid solution, the pore size in the membrane, applied pressure and the number of times the solution is extruded [20–22]. In order to make different batches of vesicles as uniform as possible we kept each of these factors constant. Each batch of vesicles produced was characterised by dynamic light scattering.

## *2.1.2 Dynamic Light Scattering*

Dynamic Light Scattering (DLS) was used to characterise the LUVs. This technique can be used to indirectly measure the average size of particles in solution by monitoring their movement over a period of time. A laser beam is shone through the sample and scattered by the suspended particles. The scattered intensity fluctuates in time as the particles in the

sample move around. Time shifted intensity traces can be used to generate a correlation function which can be fitted to determine a diffusion coefficient for the particles. Assuming the particles are spherical and undergoing Brownian motion, the size distribution can then be calculated via Stoke's Law. Sedimentation of the sample is therefore detrimental to these measurements.

*Experimental Details*

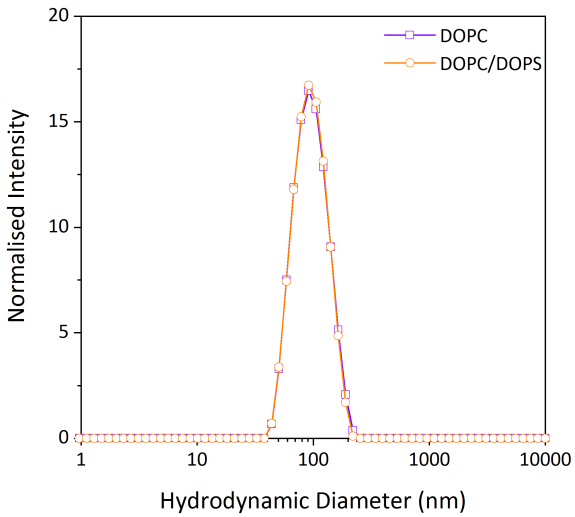



*Fig. 2.5:* Example DLS data for DOPC (purple) and DOPC/DOPS (8:2) (orange) vesicles in 0.01 M PBS. The lipid concentration was approximately 1 mg/ml in each case.

DLS experiments were carried out using a Malvern Zetasizer Nano ZSP instrument, using disposable polystyrene cuvettes. The vesicle samples were analysed at 1 mg/ml concentration in PBS and their average size and polydispersity was characterised according to the intensity distribution. An example of the DLS data obtained for pure DOPC and mixed DOPC/DOPS vesicles is shown in Figure 2.5. Data was collected at 20 °C. For each sample a total of 3 measurements was made with each measurement being the average of 11 individual data collections of 10 seconds duration.

*Discussion of Experimental Error*

DLS is highly sensitive to contamination by small particles such as dust and therefore utmost care was taken to ensure the cleanliness of all samples. The measurement procedure allows for the dismissal of measurements which deviate significantly from the overall average due to the presence of contaminants. The presence of a single, narrow peak in the majority of our DLS measurements indicates good sample purity. In Chapter 5 the accessible lipid area is used in a model for the interaction between peptoids and lipid mem-

branes. The error in this value was determined on a case by case basis from the standard deviation in the result of triplicate analysis of the vesicle size distribution.

### *2.1.3 UV-Visible Spectroscopy*

Ultra Violet-Visible (UV-Vis) Spectroscopy is absorption spectroscopy in the UV and visible regions of light. This technique works on the principle that, across the wavelength region concerned, electrons from bonding and non-bonding orbitals in the molecules of interest absorb light and are promoted to excited states. Small energy gaps between the highest occupied and lowest unoccupied molecular orbitals correspond to absorption at longer wavelengths. Measuring the absorption across a range of wavelengths results in spectral peaks which can be correlated to the bonds present in certain molecules and thus the technique can be used to quantitatively determine the presence of certain analytes in solution. These include biological molecules, metal ions and conjugated organic molecules.

The concentration of a given analyte in solution can be determined from the Beer-Lambert Law (equation 2.3), where $A$ is the absorbance, $\epsilon$ is the molar absorptivity, $l$ is the path length and $c$ is the concentration. The molar absorptivity is an intrinsic property of chemical species which is a measure of how strongly they attenuate light at a given wavelength.

$$A = \varepsilon cl \tag{2.3}$$

In this work UV-Vis is used exclusively to determine the concentration of peptoids in solution. This is possible given that every peptoid studied in this work contains a certain fraction of *N*spe residues, with side chains containing an aromatic group like that in the amino acid phenylalanine (illustrated in Figure 2.6). This has an absorption band centred around 257 nm. The concentration measured by this method is the concentration of *N*spe residues which can then be divided by the number of *N*spe residues in each peptoid sequence to give the peptoid concentration of the solution.

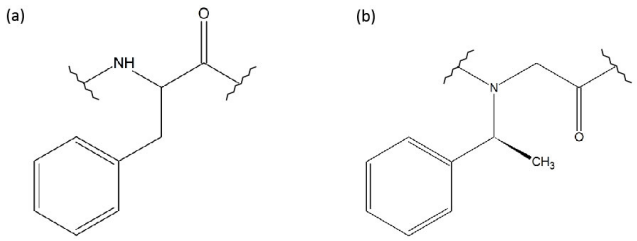



*Fig. 2.6:* Structure of (a) phenylalanine residue in peptides and (b) *N*spe residue in peptoids. Both contain the aromatic group responsible for a peak at 257 nm in the UV-Vis spectrum.

This method was used to measure the concentration of stock solutions of peptoids

which were then further diluted for the collection of the CD spectra presented in Chapters 3 and 5. Additionally, UV-Vis was used to calculate peptoid concentrations in partitioning experiments designed to measure Log $D$, the PBS-octanol distribution coefficients, which are presented in Chapter 5.

### *Experimental Details*

In order to make peptoid solutions of the desired concentration for CD measurements, a high concentration stock solution was made up for each peptoid. The concentration of each stock solution was determined by UV-Vis spectroscopy. UV-Vis measurements were carried out using a ThermoFisher Scientific Nanodrop 1 microvolume UV-Vis spectrophotometer. Measurements were made over the wavelength range 200-350 nm using 3 $\mu$L samples. Though measurements are performed on microvolume samples the nanodrop normalizes the path length to 1 cm. The spectrum for each peptoid solution was averaged over 5 measurements and a baseline was fitted and subtracted to account for scattering. The concentration of peptoid was then calculated according to the Beer-Lambert Law (given in equation 2.3), with $l = 1$ cm, $\epsilon = 195$ cm$^{-1}$ M$^{-1}$ and the measured absorbance at 257 nm.

### *PBS-octanol Partitioning Protocol*

The nanodrop was used to determine the relative concentrations of peptoid in both phases in the PBS-octanol partitioning experiments reported in Chapter 5. These experiments were carried out with a protocol adapted from that presented by Bolt *et al.* in their 2017 study of peptoid partitioning [23]. The key change made to the protocol was that instead of preparing large volumes of sample and measuring the peptoid concentration via bulk UV-Vis measurements, microvolume samples were prepared and measured according to the following protocol.

Powdered peptoid was dissolved in PBS at concentration in the region of 1 mM. The concentration was measured as detailed above, using the nanodrop. 50 $\mu$L of peptoid solution was deposited in an eppendorf tube and 50 $\mu$L of octanol was then deposited in the same tube. The sample was then gently agitated for a period of 24 hours following which it was allowed to rest for 1 week to allow the peptoid to partition between the two solvents and the system to equilibrate. The concentration of peptoid in each phase was then measured (again using the nanodrop) by taking a 3 $\mu$L aliquot from each solvent phase. This protocol was carried out in triplicate for each peptoid.

Great care was necessary in extracting the PBS aliquot from the mixed phase sample. As the pipette tip has to pass through the octanol to reach the PBS phase and small amounts of octanol tend to cling to the plastic tip this can result in the PBS samples

being contaminated by octanol, which distorts the spectrum. To circumvent this, the pipette plunger was not depressed until the end of the tip was submerged in the PBS phase, expelling the air and any octanol. The PBS aliquot was then drawn up into the tip and the tip removed from the sample. The tip was rested briefly on the clean inner side of the sample tube to let any external octanol run off before the sample itself was deposited on the nanodrop stage. It was clear from the signal to noise in the collected spectra whether a clean PBS sample had been collected successfully and in any cases where contamination was evident the data was discarded and the process repeated.

*Discussion of Experimental Errors Associated with Concentration Measurement by UV-Vis*

Inconsistency in sample concentration is one of the main sources of experimental error in CD spectra, therefore accurately measuring the concentration of each peptoid sample is important. Each peptoid stock solution was prepared at least 12 hours prior to measuring the concentration to ensure all powder was fully dissolved. For each sample 5 UV-Vis spectra were collected and averaged before the calculation of the peptoid concentration. The concentration of the peptoids in solution is particularly important in the calculation of the Log $D$ values reported in Chapter 5. Errors in Log $D$ were calculated from the standard deviation in results from repeat measurements. The errors associated with particular peptoids that partitioned strongly into one phase (PBS or octanol) are much larger than those where the peptoid was more evenly distributed between the two phases. This is due to the difficulty of obtaining accurate concentration measurements of low concentration solutions using micro-volume samples on the nanodrop.

## 2.2 Computational Methods

The main computational technique concerned in this work is molecular dynamics (MD) simulations. MD has become a very popular technique for the simulation of biomolecules such as proteins and nucleic acids. The ultimate aim of such techniques is to use atomic level details to predict the behaviour and properties of a system on a macroscopic scale. MD is an extremely well developed technique which has been refined over decades but only more recently has the potential to apply this technique to the simulation of peptoids begun to be explored in detail. A review of the computational studies of peptoids to date and details of the derivation of specific peptoid force field parameters are given at the beginning of Chapter 4. The more general principles underlying MD are described in the next sections in this chapter.

### 2.2.1 Molecular Dynamics

The simulations presented in this thesis have been carried out using the Groningen Machine for Chemical Simulations (GROMACS) [24, 25] version 2016.5. The GROMACS engine is an all in one package ideal for carrying out MD simulations and with a large range of inbuilt analysis tools. The principles of molecular dynamics as implemented in GROMACS are extremely well documented in the GROMACS manual [26], among other sources, and therefore only a brief overview of the theory and methods for implementation of the technique are presented here.

MD is based on a purely classical treatment of a system of atoms. Such a system is constructed and allowed to evolve over time by solving Newton's equations of motion via numerical integration. The atoms are allowed to move for a fixed duration, giving a representation of the dynamical evolution of the system. In equation 2.4, $F_i$ is the force on the $i$th particle, $m_i$ is the mass of and $r_i$ is the position of the $i$th particle, for a system of $N$ interacting atoms.

$$m_i \frac{\delta^2 r_i}{\delta t^2} = F_i, i = 1...N \tag{2.4}$$

The forces are given by equation 2.5 and are the negative derivatives of a potential function, $V(r_1, r_2, ..., r_N)$

$$F_i = -\frac{\delta V}{\delta r_i} \tag{2.5}$$

For any particular system, MD is carried out by first computing the forces on each atom in the system and then updating the molecular configuration by solving Newton's equations of motion, based on the initial velocities and positions, to give the movement of the atoms. This can be implemented by one of several algorithms designed to integrate the equations of motion.

#### *Integration Algorithms*

To carry out an MD simulation the forces on each atom must be calculated using an integration algorithm. The default integration algorithm in GROMACS is the leap-frog algorithm [27] which updates the atomic positions at time, $t$, and velocities at time $t - \frac{1}{2}\Delta t$ according to the forces, $F(t)$, and the following relations:

$$r(t + \Delta t) = r(t) + \Delta t v(t + \frac{1}{2}\Delta t) \tag{2.6}$$

$$v(t + \frac{1}{2}\Delta t) = v(t - \frac{1}{2}\Delta t) + \frac{\Delta t}{m} F(t) \tag{2.7}$$

This offers increased numerical precision on a previous algorithm, the Verlet algorithm, despite being algebraically equivalent and producing identical trajectories [28]. The choice of time step, $\Delta t$ is very important. The time step must be short enough to capture interactions that occur on even the shortest time scales in the system, however the use of longer time steps results in lower computational expense. Ultimately the time step must be approximately 10 times shorter than the highest frequency vibrations, which in most systems are the bond vibrations, resulting in time steps of the order of 1 fs [29].

### *Stochastic Dynamics*

Simulations may also be performed in GROMACS using the stochastic dynamics (SD) integrator. In SD a friction term is added to the Newtonian equations of motion so that they take the form of Langevin equations. SD can be implemented using an integration algorithm that is essentially equivalent to the normal MD leap-frog algorithm in both implementation and accuracy [30]. SD in a vacuum allows a system to explore conformational space much faster than MD. The dynamics of the system will vary between the two methods but the sampling should be correct in either case, provided the simulations are run over a sufficiently long period of time.

### *Control of Thermodynamic Conditions*

It is important to control the thermodynamic conditions of a system during a simulation in order for the simulation to have physical meaning. The two main conditions to be controlled are the temperature and pressure. MD with no attempt to control either temperature or pressure results in an $NVE$ ensemble, where the number of molecules, $N$, volume of the simulation box, $V$ and energy, $E$ are constant. However, it is almost always more useful to create an $NVT$ (canonical) or $NPT$ ensemble, where $T$ and $P$ refer to constant temperature and pressure respectively. In an $NVT$ ensemble the volume of the box is fixed whereas in an $NPT$ ensemble it is allowed to fluctuate to enable the pressure to remain constant. In order to generate such ensembles the thermodynamic properties of the system must be controlled. In the case of the number of particles and volume of the box this is naturally straight forward and occurs without intervention. Temperature and pressure however must be controlled through the use of thermostats and barostats.

Given that the temperature of a system is calculated from its kinetic energy, it is therefore proportional to the velocities of its constituent particles. These relationships are defined in equations 2.8 and 2.9 where $E_\mathrm{k}$ is the kinetic energy of the system, $N$ is the number of atoms, $k_\mathrm{B}$ is Boltzmann's constant and $m_i$ and $v_i$ the mass and velocity of the $i$th particle respectively.

$$T = \frac{2}{3}\frac{E_{\mathrm{k}}}{Nk_{\mathrm{B}}} \tag{2.8}$$

$$E_k = \frac{1}{2}\sum_{i=1}^{N} m_i v_i^2 \tag{2.9}$$

Computational thermostats work by scaling the velocities of the particles in a system to maintain a constant temperature. Different thermostats achieve this via different mechanisms. In the simulations performed to gather the data presented in this thesis we utilise the Berendsen and Nosé-Hoover thermostats, the former for equilibration of systems and the latter for production runs.

The Berendsen thermostat scales the velocities of all particles in a system by the same time dependent factor, $\lambda$, which is given by equation 2.10 where $T$ is temperature, $T_0$ is the reference temperature and $\tau_T$ is the temperature coupling constant.

$$\lambda = \left[1 + \frac{\Delta t}{\tau t}\left(\frac{T_0}{T} - 1\right)\right]^{\frac{1}{2}} \tag{2.10}$$

This is a method of temperature coupling has the merit of simplicity and is computationally efficient however it suppresses fluctuations in the kinetic energy and hence does not generate a Boltzmann distribution of velocities. This is somewhat inaccurate and hence used only for equilibration runs. The Nosé-Hoover thermostat [31, 32] can correctly define an $NVT$ ensemble as it does produce a Boltzmann distribution of velocities by introducing a thermal reservoir and a friction term into the equations of motion, as in equation 2.11.

$$\frac{d^2 r_i}{dt_2} = \frac{F_i}{m_i} - \zeta \upsilon_i \tag{2.11}$$

The friction term, $\zeta$, is shown in equation 2.12, where $Q$ is a constant that determines the strength of the coupling to the thermal reservoir and $X$ is the number of degrees of freedom in the system.

$$\frac{d\zeta(t)}{dt} = \frac{1}{Q}\left[\sum_{i}^{N} m_i v_i^2 - (X+1)k_{\mathrm{B}}T_0\right] \tag{2.12}$$

Barostats are used to control the pressure during the course of an MD simulation by scaling the size of the simulation box. A Berendsen barostat which operates in a similar way to the Berendsen thermostat is useful for system equilibration but results in large pressure fluctuations over the course of the simulation. The alternative, Parinello-Rahman pressure coupling, is therefore used for production runs. The Parinello-Rahman barostat operates in a similar way to the Nosé-Hoover thermostat and the two combined

should product a true $NPT$ ensemble [33].

### 2.2.2 Force Fields

The potential function in equation 2.5 and associated parameters are known as a force field. Force fields can be created to describe a system at various levels of detail. Atomistic force fields model each atom in the system as an individual entity, whereas coarse grained force fields model small groups of atoms as individual "pseudo-atoms" or "beads". Atomistic force fields comprise a complete set of parameters and potentials to describe all of the interactions between all of the atoms in the system. Many different force fields have been designed to describe all manner of systems. Key differences in these force fields arise from the method by which the parameters are derived and variations in the potentials used. It is therefore vital to choose an appropriate force field that correctly reproduces known properties of the system of interest [34].

The overall potential in any force field is given by the sum of the bonded and non-bonded interaction potentials (equation 2.13).

$$E_{\text{total}} = \sum E_{\text{bonded}} + \sum E_{\text{non-bonded}} \tag{2.13}$$

The bonded interactions are given by equation 2.14 and the non bonded interactions are given by equation 2.15.

$$E_{\text{bonded}} = \sum E_{\text{bonds}} + \sum E_{\text{angles}} + \sum E_{\text{dihedrals}} \tag{2.14}$$

$$E_{\text{non-bonded}} = \sum E_{\text{electrostatic}} + \sum E_{\text{van der waals}} \tag{2.15}$$

#### *Bonded Interactions*

The bonded interactions in equation 2.14 are the method by which the force field imposes the correct molecular geometry on the system. These are important to implement correctly as the structural characteristics of individual molecules can be influential in the overall behaviour of the system. $E_{\text{bonds}}$ is the bond stretching potential, which acts between two atoms directly connected by a bond and is given by a harmonic potential (equation 2.16) where $r_{ij}$ is the distance between two atoms $i$ and $j$ (the bond length) and $r_0$ is the equilibrium bond length. The minimum of this potential corresponds to the equilibrium bond length.

$$E_{\text{bonds}}(r_{ij}) = \frac{1}{2}K(r_{ij} - r_0)^2 \tag{2.16}$$

The second bonded interaction, $E_{\text{angles}}$ is the angle bending potential, which controls

the behaviour of the angles formed between three adjacent bonded atoms ($\Theta_{ijk}$). These are represented by another harmonic potential (equation 2.17) where $K_\Theta$ is a force constant and $\Theta_0$ is the equilibrium bond angle.

$$E_{\text{angles}}(\Theta_{ijk}) = \frac{1}{2}K_\Theta(\Theta_{ijk} - \Theta_0)^2 \tag{2.17}$$

For every 4 adjacently bonded atoms in a molecule there is a pair of intersecting planes, each defined by 3 atoms, with 2 atoms in common. The angle between these 2 planes is referred to as the dihedral or torsion angle. In a force field this is known as a proper dihedral and in GROMACS these are commonly defined by the Ryckaert-Bellemans function (equation 2.18) where $\phi_{ijkl}$ is the dihedral angle between the planes defined by atoms $ijk$ and $jkl$.

$$E_{\text{dihedrals}}(\phi_{ijkl}) = \sum_{n=0}^{5} C_n(\cos(\phi_{ijkl}))^n \tag{2.18}$$

This is known as a proper dihedral and defines the potential energy for rotation around a chemical bond (the bond between atoms $j$ and $k$) which essentially controls how flexible the molecule is. An additional contribution to the overall dihedral term is given by equation 2.19 which is known as the improper dihedral potential. Improper dihedrals are dihedrals where the 4 atoms defining the planes are not all explicitly linked by bonds and are often used to enforce properties such as the planarity of aromatic rings. In the potential form given in equation 2.19, $\varphi_{ijkl}$ is the angle between the $ijk$ and $jkl$ planes, $n$ is the periodicity and $\varphi_0$ is the angle corresponding to the potential maximum. $K$ is once again a force constant.

$$E_{\text{improper}}(\varphi_{ijkl}) = K(1 + \cos(n\varphi_{ijkl} - \varphi_0)) \tag{2.19}$$

The dihedral term is of particular importance in the context of this thesis as we focus on re-parametrizing this aspect of the force field to better capture peptoid backbone torsions. This is discussed in greater detail in Chapter 4.

### *Non-Bonded Interactions*

In the majority of force fields there are two major components describing the non-bonded interactions in the system. These are the electrostatic and Van der Waals interactions. The electrostatic potential is necessary to define the interactions between the partial charges assigned to each atom in the system. These can be described in a quick and simple manner by the Coulomb interaction (equation 2.20) which is implemented within a specified cut-off distance, beyond which the interaction potential is set to 0. However, this treatment neglects the effect of long range electrostatics, as in reality $E_{\text{electrostatic}}$ tends to 0 very

slowly as $r$ increases and therefore reasonable cut-off distances may exclude interactions at larger $r$ that may be important in capturing the correct behaviour of the system.

$$E_{\text{electrostatic}} = \frac{q_i q_j}{4\pi\epsilon_0 r} \tag{2.20}$$

In equation 2.20 $q_i$ and $q_j$ are the charges on atoms $i$ and $j$ respectively, $r$ is the distance between the 2 atoms and $\epsilon_0$ is the permittivity of free space.

There are several methods designed to improve upon the use of the Coulomb potential and include long-range electrostatics in force fields but perhaps the most popular (and the one implemented in the work in this thesis) is the particle-mesh Ewald (PME) method. In this method, the total electrostatic interactions are split into short-range and long-range components according to a cut-off distance. The short range interactions are described by the Coulomb potential which converges quickly in real space and the long-range interactions are calculated by a summation in Fourier space [35].

The Van der Waals interactions are normally described in atomistic force fields as a function of distance, $r$, by the Lennard-Jones potential (equation 2.21) where $\varepsilon$ is the potential well depth and $\sigma$ is the distance at which the potential is equal to 0.

$$E_{\text{Lennard-Jones}} = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{2.21}$$

This potential plays several important roles in ensuring the physical behaviour of the system. The presence of the repulsive part of the potential prevents steric clashes and atoms overlapping. The attractive well defines the strength of the interaction between atoms at a particular distance. At large distances this attraction tends to 0 and consequently there is no interaction between atoms separated by such distances.

### *2.2.3 Enhanced Sampling in MD: Hamiltonian Replica Exchange*

In some systems the energy barriers between different states are high enough that certain behaviours are inaccessible or too computationally expensive to observe over normal atomistic MD time scales. Examples of such behaviours are the folding of peptides and proteins with rugged energy landscapes and high barriers. Some of the behaviours that we expect to observe in peptoids, such as rotations of the backbone amide, are inaccessible, or extremely computationally expensive on normal MD time scales and at ambient temperatures [36]. Therefore an accelerated MD technique is required for adequate sampling of structural conformations. A wide range of biasing and accelerating techniques are available to overcome such barriers [37]. These include umbrella sampling [38], metadynamics [39], parallel tempering [40], accelerated MD [41] and replica exchange MD [42], all of which have their own advantages and disadvantages, depending on the system in

question. Here we have chosen to use a particular form of replica exchange MD, known as Hamiltonian replica exchange to enhance the rate of sampling across the high barrier to rotation around the peptoid amide bond. Replica exchange molecular dynamics (REMD) is an enhanced sampling technique where a set of identical systems (replicas) are simulated in parallel at different temperatures [37]. Systematic attempts to exchange coordinates between replicas via Monte Carlo moves are attempted at regular intervals. The number of replicas required to span a chosen temperature range scales with the size of the system. This method has previously been employed to enhance the sampling of peptoid backbone conformations [36, 43] but remains computationally expensive.

An alternative to traditional replica exchange in temperature is Hamiltonian Replica Exchange (HREX). In HREX rather than simulating and exchanging between replicas with different temperatures we simulate and exchange between replicas with different Hamiltonians. Two advantages of HREX over temperature REMD are that each replica is at the same temperature so velocity rescaling is not required after a coordinate swap and the replica exchange can be applied to a specific part of the potential, rather than the system as a whole, minimising computational expense. HREX is supported in GROMACS through the PLUgin for MolEcular Dynamics (PLUMED) [44] which enables the parallel running of simulations that have different topologies but are otherwise identical. This has previously been used successfully for rapid conformational sampling of small peptides [45].

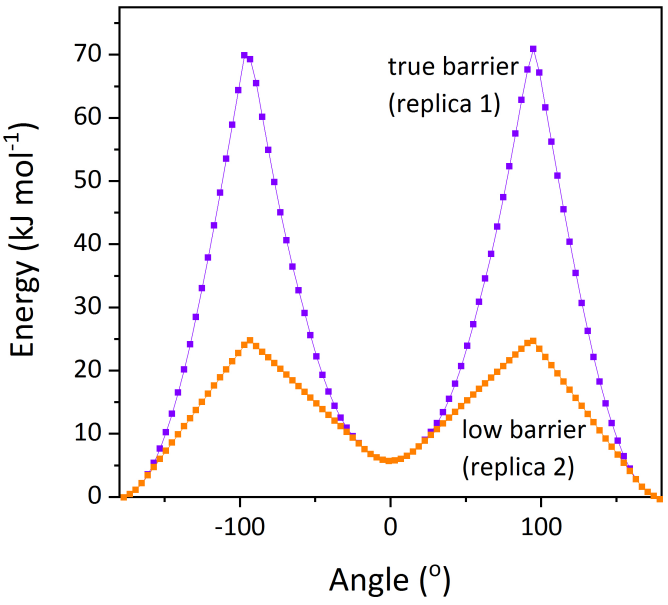



*Fig. 2.7:* Rotational energy profiles around $\omega$ backbone dihedral in peptoids. True force field profile shown in purple, false low barrier to enable crossing in second replica shown in orange.

In this work we attempt to use HREX to overcome barriers to peptoid conformational

sampling by running parallel simulations of just 2 replicas. The topology for the first replica includes the normal peptoid force field parameters. The topology for the second replica is identical with the exception that a new set of Ryckaert-Bellemans (RB) coefficients are derived to lower the barrier to rotation around one peptoid backbone dihedral ($\omega$, which is discussed in detail in Chapter 4). The standard and modified energy profiles for rotation around the backbone dihedral are shown in Figure 2.7. The new RB coefficients to produce the lower energy barrier were derived in the same way as modified GAFF parameters were derived. This is described in Chapter 4.

### *2.2.4 Quantum Chemical Calculations*

Within the results presented in this thesis are several quantum chemical (QC) calculations, carried out to investigate the rotational energy profiles of the peptoid backbone and subsequently used in the re-parametrisation of an atomistic force field.

QC calculations can be broadly broken down into 2 categories: *ab initio* methods and density functional theory (DFT). In true *ab initio* methods useful information, such as the energy and electron density of a system, can be acquired by finding solutions to the time-independent Schrödinger equation without the use of any derived parameters or empirical approximations. Only physical constants (Planck's constant, the speed of light in a vacuum and electronic and nuclear charge) are used in the calculations. This can be highly accurate but computationally expensive and is limited to only the smallest of systems. The most common *ab initio* formalism is the Hartree-Fock algorithm where the variational principle is used to express the total wave function as a product of single electron wave functions. This inherently includes the Born-Oppenheimer approximation, that the motion of electrons and atomic nuclei can be treated separately.

DFT provides a simplified approach to QC calculations by allowing the properties of the system to be determined only from the electron density. This offers greater computational efficiency as the electron density can be expressed as a function of only 3 variables, $r = (x, y, z)$. A general expression for energy in DFT is given in equation 2.22 which is known as the Kohn-Sham model [46].

$$E[\rho(r)] = V_{\mathrm{ne}}[\rho(r)] + V_{\mathrm{ee}}[\rho(r)] + T_{\mathrm{ni}}[\rho(r)] + E_{\mathrm{xc}}[\rho(r)] \tag{2.22}$$

This model includes contributions from the attraction between the nuclei, $V_{\mathrm{ne}}$, the classical Coulomb interaction between pairs of electrons, $V_{\mathrm{ee}}$, kinetic energy between non-interacting electrons, $T_{\mathrm{ni}}$ and the exchange-correlation energy, $E_{\mathrm{xc}}$, all of which are expressed as functions of 3 variables, $r = (x, y, z)$.

The underlying physics pertaining to the details of the exact methods for performing such calculations and the effect on the accuracy of results are not greatly relevant to the re-

sults presented in this thesis and therefore will not be elaborated on here but may be found elsewhere [47]. The major consideration when interpreting the results of QC calculations is the level of theory and choice of basis set [48]. This refers to the specific mathematical formulation according to which the Schrödinger equation is solved, the choice of which will ultimately determine the accuracy of the result.

# References: Chapter 2


[1] N. Sreerama, R. W. Woody, "Estimation of Protein Secondary Structure from Circular Dichroism Spectra : Comparison of CONTIN , SELCON , and CDSSTR Methods with an Expanded Reference Set", *Analytical Biochemistry* **2000**, *260*, 252–260.

[2] N. J. Greenfield, "Using circular dichroism spectra to estimate protein secondary structure", *Nature Protocols* **2006**, *1*, 2876–2890.

[3] L. Whitmore, B. A. Wallace, "DICHROWEB , an online server for protein secondary structure analyses from circular dichroism spectroscopic data", *Nucleic Acids Research* **2004**, *32*, 668–673.

[4] L. E. E. Whitmore, R. W. Janes, B. A. Wallace, "Protein Circular Dichroism Data Bank ( PCDDB ): Data Bank and Website Design", *Chirality* **2006**, *429*, 426–429.

[5] B. M. Bulheller, A. Rodger, J. D. Hirst, "Circular and linear dichroism of proteins.", *Physical Chemistry Chemical Physics* **2007**, *9*, 2020–2035.

[6] M. C. Manning, "Underlying assumptions in the estimation of secondary structure content in proteins by circular dichroism spectroscopy - a critical review", *Journal of Pharmaceutical and Biomedical Analysis* **1989**, *7*, 1103–1119.

[7] S. Brahms, J. Brahms, "Quantitative Characterization of Protein Conformation in Solution By Vacuum Ultraviolet Circular Dichroism", *Biomolecular Structure Conformation Function and Evolution* **1981**, 31–43.

[8] M. W. Parker, "Protein Structure from X-Ray Diffraction", *Journal of Biological Physics* **2003**, *29*, 341–362.

[9] L. T. Roe, J. G. Pelton, J. R. Edison, G. L. Butterfoss, B. W. Tresca, B. A. LaFaye, S. Whitelam, D. E. Wemmer, R. N. Zuckermann, "Unconstrained peptoid tetramer exhibits a predominant conformation in aqueous solution", *Biopolymers* **2019**.

[10] P. Armand, K. Kirshenbaum, R. A. Goldsmith, S. Farr-Jones, A. Barron, K. T. V. Truong, K. A. Dill, D. F. Mierke, F. E. Cohen, R. N. Zuckermann, E. K. Bradley, "NMR determination of the major solution conformation of a peptoid pentamer with chiral side chains", *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 4309–4314.

[11] J. P. Hennessey, W. C. Johnson, "Experimental errors and their effect on analyzing circular dichroism spectra of proteins", *Analytical Biochemistry* **1982**, *125*, 177–188.

[12] C. H. Huang, "Studies on Phosphatidylcholine Vesicles. Formation and Physical Characteristics", *Biochemistry* **1969**, *8*, 344–352.

[13] S. E. Schullery, C. F. Schmidt, P. Felgner, T. W. Tillack, T. E. Thompson, "Fusion of Dipalmitoylphosphatidylcholine Vesicles", *Biochemistry* **1980**, *19*, 3919–3923.

[14] C. F. Schmidt, D. Lichtenberg, T. E. Thompson, "Vesicle-Vesicle Interactions in Sonicated Dispersions of Dipalmitoylphosphatidylcholine", *Biochemistry* **1981**, *20*, 4792–4797.

[15] S. F. Greenhut, V. R. Bourgeois, M. A. Roseman, "Distribution of cytochrome b5 between small and large unilamellar phospholipid vesicles.", *The Journal of Biological Chemistry* **1986**, *261*, 3670–3675.

[16] A. S. Ladokhin, S. Jayasinghe, S. H. White, "How to measure and analyze tryptophan fluorescence in membranes properly, and why bother?", *Analytical Biochemistry* **2000**, *258*, 235–245.

[17] A. S. Ladokhin, M. Fernández-Vidal, S. H. White, "CD spectroscopy of peptides and proteins bound to large unilamellar vesicles", *Journal of Membrane Biology* **2010**, *236*, 247–253.

[18] F. Olson, C. A. Hunt, F. C. Szoka, W. J. Vail, D. Papahadjopoulos, "Preparation of liposomes of defined size distribution by extrusion through polycarbonate membranes", *BBA - Biomembranes* **1979**, *557*, 9–23.

[19] R. C. MacDonald, R. I. MacDonald, B. P. M. Menco, K. Takeshita, N. K. Subbarao, L. rong Hu, "Small-volume extrusion apparatus for preparation of large, unilamellar vesicles", *BBA - Biomembranes* **1991**, *1061*, 297–303.

[20] P. J. Patty, B. J. Frisken, "The pressure-dependence of the size of extruded vesicles", *Biophysical Journal* **2003**, *85*, 996–1004.

[21] D. G. Hunter, B. J. Frisken, "Effect of extrusion pressure and lipid properties on the size and polydispersity of lipid vesicles", *Biophysical Journal* **1998**, *74*, 2996–3002.

[22] A. Hinna, F. Steiniger, S. Hupfeld, P. Stein, J. Kuntsche, M. Brandl, "Filter-extruded liposomes revisited: a study into size distributions and morphologies in relation to lipid-composition and process parameters", *Journal of Liposome Research* **2016**, *26*, 11–20.

[23] H. L. Bolt, C. E. Williams, R. V. Brooks, R. N. Zuckermann, S. L. Cobb, E. H. Bromley, "Log D versus HPLC derived hydrophobicity: The development of predictive tools to aid in the rational design of bioactive peptoids", *Biopolymers* **2017**, *108*, 1–7.

[24] H. J. Berendsen, D. van der Spoel, R. van Drunen, "GROMACS: A message-passing parallel molecular dynamics implementation", *Computer Physics Communications* **1995**, *91*, 43–56.

[25] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindah, "Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers", *SoftwareX* **2015**, *1-2*, 19–25.

[26] M. J. Abraham, D. Van Der Spoel, E. Lindahl, B. Hess, T. G. development Team, GROMACS User Manual version 2016.5, tech. rep., **2018**.

[27] R. W. Hockney, S. P. Goel, J. W. Eastwood, "Quiet high-resolution computer models of a plasma", *Journal of Computational Physics* **1974**, *14*, 148–158.

[28] H. J. C. Berendsen, W. F. van Gunsteren in Enrico Fermi Summer School, Varenna, Italy, **1986**, pp. 43–65.

[29] D. Fincham, "Choice of timestep in molecular dynamics simulation", *Computer Physics Communications* **1986**, *40*, 263–269.

[30] N. Goga, A. J. Rzepiela, A. H. De Vries, S. J. Marrink, H. J. Berendsen, "Efficient algorithms for langevin and DPD dynamics", *Journal of Chemical Theory and Computation* **2012**, *8*, 3637–3649.

[31] S. Nosé, "A molecular dynamics method for simulations in the canonical ensemble", *Molecular Physics* **1984**, *52*, 255–268.

[32] W. G. Hoover, "Canonical dynamics: Equilibrium phase-space distributions", *PHYSICAL REVIEW A* **1985**, *31*, 1695–1697.

[33] M. Parrinello, A. Rahman, "Polymorphic transitions in single crystals: A new molecular dynamics method", *Journal of Applied Physics* **1981**, *52*, 7182–7190.

[34] V. Hornak, R. Abel, A. Okur, B. Strockbine, R. Adrian, C. Simmerling, "Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters", *PROTEINS: Structure Function and Bioinformatics* **2006**, *65*, 712–725.

[35] T. Darden, D. York, L. Pedersen, "Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems", *The Journal of Chemical Physics* **1993**, *98*, 10089–10092.

[36] V. A. Voelz, K. A. Dill, I. Chorny, "Peptoid conformational free energy landscapes from implicit-solvent molecular simulations in AMBER", *Biopolymers* **2011**, *96*, 639–650.

[37] R. Zhou, "Replica exchange molecular dynamics method for protein folding simulation.", *Methods in molecular biology* **2007**, *350*, 205–223.

[38] J. Kästner, "Umbrella sampling", *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 932–942.

[39] A. Laio, F. L. Gervasio, "Metadynamics : a method to simulate rare events and reconstruct the free energy in biophysics , chemistry and material science", *Reports on Progress in Physics* **2008**, *71*.

[40] D. J. Earl, M. W. Deem, "Parallel tempering: Theory, applications, and new perspectives", *Physical Chemistry Chemical Physics* **2005**, *7*, 3910–3916.

[41] D. Hamelberg, J. Mongan, J. A. Mccammon, "Accelerated molecular dynamics : A promising and efficient simulation method for biomolecules", *Journal of Chemical Physics* **2004**, *120*, 11919–11929.

[42] H. Nymeyer, "How Efficient Is Replica Exchange Molecular Dynamics ? An Analytic Approach", *Journal of Chemical Theory and Computation* **2008**, *4*, 626–636.

[43] S. Mukherjee, G. Zhou, C. Michel, V. A. Voelz, "Insights into Peptoid Helix Folding Cooperativity from an Improved Backbone Potential", *Journal of Physical Chemistry B* **2015**, *119*, 15407–15417.

[44] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia, M. Parrinello, "PLUMED: A portable plugin for free-energy calculations with molecular dynamics", *Computer Physics Communications* **2009**, *180*, 1961–1972.

[45] G. Bussi, "Hamiltonian replica exchange in GROMACS: A flexible implementation", *Molecular Physics* **2014**, *112*, 379–384.

[46] C. J. Cramer, *Essentialls of Computational Chemistry, Theories and Models*, Wiley, **2004**, Chapter 8.

[47] E. G. Lewars, *Computational chemistry: Introduction to the theory and applications of molecular and quantum mechanics: Third Edition 2016*, 3rd ed., Springer, **2016**, pp. 1–728.

[48] E. R. Davidson, D. Feller, "Basis Set Selection for Molecular Calculations", *Chemical Reviews* **1986**, *86*, 681–696.

# 3.0 Experimental Secondary Structural Characterisation

The work presented in this chapter concerns the secondary structural characteristics of peptoids. We use circular dichroism spectroscopy to probe the effect of rearrangements in peptoid primary structure on secondary structure in aqueous and organic solutions. A review of relevant literature and background information is presented first, followed by data gathered for the experimental characterisation of a library of peptoids. The results are discussed in the context of the literature and the other work presented later in this thesis.

## 3.1 Introduction and Review of Literature

The folding of peptoids into stable secondary structures similar to the $\alpha$-helices and $\beta$-sheets that occur in proteins and peptides is of great interest, as better understanding of the process will enable rational design of useful structures. Peptide and protein structures have been extensively investigated and characterised, while in the case of peptoids the field is still in its infancy [1].

Secondary structure is formally defined as the pattern of hydrogen bonds between the oxygens and amide hydrogens along the folded peptide backbone. Peptoids lack the intra-backbone hydrogen bonding ability of peptides and the peptoid backbone does not contain any chiral centres to impart handedness to a structure [2]. In addition the peptoid backbone is more flexible than the peptide backbone. This imparts considerable potential for conformational diversity and flexibility. Many peptoid sequences have nevertheless been shown to adopt stable configurations akin to peptide secondary structures in solution [3–6]. Like many other synthetic foldamers, these structures are predominantly stabilised through non-covalent interactions such as sterics and electrostatics and have been found to bear some resemblance to the sheets and helices formed by natural biopolymers. The structure of the side chains incorporated into the peptoid primary sequence can therefore be highly influential over the secondary structure and has been a focus of many studies aiming to elucidate sequence to structure relationships [7].

The peptoids that we have investigated in this work almost exclusively exhibit CD spectra that indicate that they adopt helical secondary structures in aqueous and organic solution. Therefore the bulk of this review will concern the structural similarities and differences between the various forms of helical secondary structure known to exist in peptides and proteins and those that have so far been recorded for peptoids. Particular attention will be paid to structural characterisation by CD spectroscopy, as this is the technique by which we have characterised the sequences in our peptoid library.

### 3.1.1 Helices in Peptides

#### *Structural Characteristics*

Polypeptides may adopt various different types of helical conformations which can be defined and distinguished by geometric properties such as the pitch, radius and handedness of the helix.

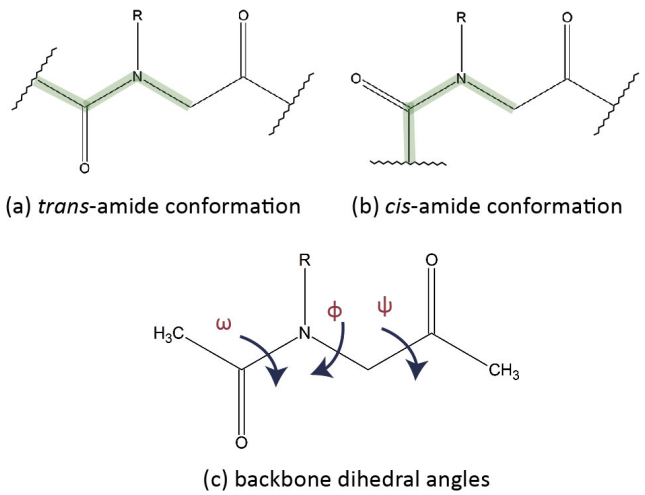



*Fig. 3.1:* Representative structures of the peptoid backbone amide in the (a) *trans* and (b) *cis* conformations. The same positioning of the heavy backbone atoms in these two conformations applies in peptides. (c) Peptoid backbone dihedral angles, defined as the angles between the 2 intersecting planes defined by 4 consecutive atoms. For example, $\omega$ represents the rotation around the C-N bond and is the angle between the C-C-N and the C-N-C planes.

The helical character of a peptide can also be characterised by the dihedral angles, $\omega$, $\phi$ and $\psi$, adopted by the peptide backbone in a particular conformation. The peptoid/peptide backbone dihedral angles, which are defined as the angles between two planes between 4 consecutive backbone atoms, are shown in Figure 3.1c. The $\omega$ angle represents the conformation of the peptide bond, which is both planar and 180° in almost all cases. This is referred to as the *trans* conformation (Figure 3.1a). In rare cases the peptide bond exists in the *cis* conformation (Figure 3.1b) with $\omega = 0°$. In peptides $\phi$ and $\psi$ may take a

range of values that correspond to particular secondary structural conformations. These can be visualised through Ramachandran plots, a form of 2-D histogram that shows the conformational energy or probability for each pair of $\phi/\psi$ dihedral angles in the peptide backbone [8].

Ramachandran plots are a useful medium for visualising peptoid secondary structural preferences, given their structural similarities to peptides. By convention peptide/protein Ramachandran plots are normally plotted over the range -180° - +180° with $\phi$ on the x-axis and $\psi$ on the y-axis. The conventions for peptoid Ramachandran plots are slightly different to account for the potential *cis-trans* isomerisation of the peptoid bond. As such, they are often plotted in pairs, one for the $\phi$-$\psi$ energies for the *cis* conformation and another for the *trans* conformation. These are normally plotted with axes for the angles from 0° - 360°. This results in centering of regions commonly populated by peptoid structures.

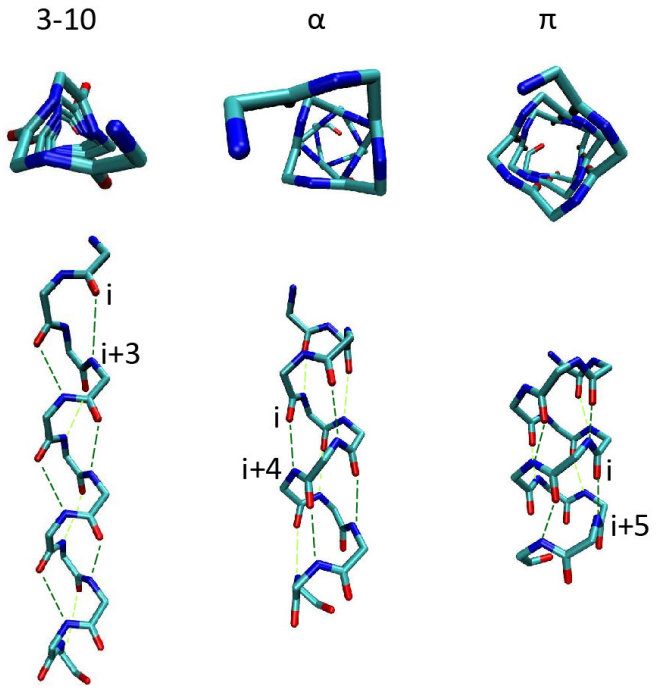



*Fig. 3.2:* Representative structures for helical backbones in the 3-10, $\alpha$ and $\pi$ configurations, viewed from above, looking down the helical axis and from the side, looking along the helical axis. The characteristic hydrogen bonding networks for each structure are shown in green. Backbone structures were generated using Pro-Builder online platform [9].

The secondary structures found in peptides and proteins occupy certain regions of the Ramachandran Plot corresponding to the backbone dihedral angles associated with them. The $\alpha$-helix is the most common secondary structure in nature. It is a right handed helix and has 3.6 residues per turn with a translation of 1.5 Å per amino acid, resulting in a pitch of 5.4 Å. The $\alpha$-helix is stabilised by a distinctive pattern of hydrogen bonds between the amide hydrogen and the carbonyl group four amino acid residues earlier in

the sequence. This is denoted as (i+4 $\longrightarrow$ i) hydrogen bonding [10]. $\alpha$-helices can also be identified by their backbone ($\phi$, $\psi$) dihedral angles, where adjacent $\phi$ and $\psi$ angles sum to approximately -105°, with typical helices adopting the angles ($\phi$, $\psi$) = (-60°, -45°) [11]. $\alpha$-helices are found in peptides which vary greatly in length, from sequences as short as 4 amino acids [12] up to 200 amino acids and above [13]. This conformation is often adopted by peptides in hydrophobic environments such as the lipid bilayer interior and is rarely stable in aqueous environments due to the relative weakness of the hydrogen bond network, which does not compensate sufficiently for the entropic cost of folding [14–16].

The second most commonly observed peptide helix, the 3-10 helix, is a 3 residue per turn, right handed helix with a pitch of 6 Å. These helices have an (i+3 $\longrightarrow$ i) hydrogen bonding pattern and exist most commonly as very short sections of the peptide chain (normally a maximum of 4 amino acids). The backbone dihedrals angles typically sum to approximately -75°, though there is considerable variation given that these structures tend to exist only in very short sections. The 3-10 helix is slightly more elongated and tightly wound than the $\alpha$-helix with a translation of 2 Å along the helical axis per amino acid, compared to the $\alpha$-helix 1.5 Å [17, 18].

A third right handed helical structure, the $\pi$-helix is also found in natural peptides. This a 4.1 residue per turn helix with an (i+5 $\longrightarrow$ i) hydrogen bonding pattern and a pitch of 4.7 Å, making it considerably more compact than either the $\alpha$ or 3-10 helix [19]. $\pi$-helices do not adopt specific dihedral angles and are usually only 7 amino acids in length [20]. In all of the helices discussed so far, the omega backbone dihedral, which describes the geometry of the peptide bond is approximately 180° and thus each amide is fixed in the *trans* conformation. Visual representations of the 3-10, $\alpha$- and $\pi$-helices are shown in Figure 3.2.

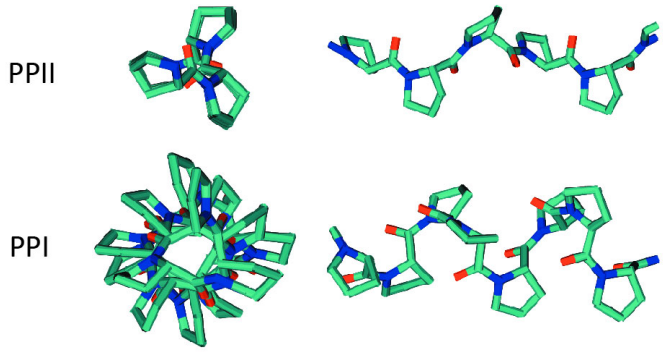



*Fig. 3.3:* Representative structures for polyproline helices, viewed from above, looking down the helical axis and from the side, looking along the helical axis. Figure adapted with permission from [21].

An additional pair of helical conformations exist in peptide sequences containing proline residues (Figure 3.3). These are therefore known as polyproline helices. Polyproline I

(PPI) helices are characterised by dihedral angles of ($\phi$, $\psi$) = (-75°, 160°) with *cis* amide bonds. These are right handed helices with approximately 3 residues per turn. Proline is the only natural amino acid which has been observed to adopt this conformation. Polyproline II (PPII) helices on the other hand have been observed in sequences containing not only proline but other amino acids too. The PPII helix is more compact than PPI and is a 3 residue per turn, left handed structure containing *trans* amides with average backbone dihedrals ($\phi$, $\psi$) = (-75°, 145°) [22]. Neither of the PP helices are hydrogen bond stabilised as the amide nitrogen is involved in the unique proline side chain structure and therefore cannot function as the H-bond donor. As a result, polyproline peptides may bear some structural similarities to peptoids.

### *Experimental Characterisation of Helices*

The presence of different helical conformations in proteins and polypeptides can be determined by many different experimental techniques including x-ray crystallography, NMR, Infrared (IR) and CD spectroscopy, among others. Each of these has unique advantages and disadvantages. This chapter concerns the characterisation of helical structures by CD spectroscopy and this will therefore be the focus of review here. As discussed in Chapter 2, CD has long been used to determine the relative fractions of different secondary structural motifs present in proteins and polypeptides in solution. This technique is particularly sensitive to, and apt for observing structural changes in biomolecules due to environmental changes. However, the ability to distinguish between the subtly different helical conformations that exist in peptides is still debated.

The presences of $\alpha$-helices is widely reported to produce distinctive features in a protein or peptide CD spectrum. These arise due to the electron transitions associated with the peptide bond in this particular conformation. Transitions from the amide non-bonding $\pi$ orbital to the anti-bonding $\pi^*$ orbital result in a strong positive band centred around 190 nm and a negative band centred around 208 nm. An additional negative band occurs at wavelengths of around 220 nm due to n-$\pi^*$ transition [23]. The subtle backbone rearrangement from $\alpha$ to 3-10 helix has been said by some to translate into minor differences in CD spectra, though others contest that the two structures are indistinguishable. Toniolo *et al.* suggested that the intensity of the negative band at approximately 208 nm is enhanced in the 3-10 helix relative to the $\alpha$-helix and the intensity of the positive band at approximately 195 nm is diminished. The ratio of second to first minimum is therefore different in 3-10 helices than $\alpha$-helices. The theoretical prediction for the ratio of the intensities of the two spectral minima for a 3-10 helix is 0.4, which is consistent with published experimental CD spectra [24, 25]. There is overlap of the allowed regions of the Ramachandran plot for the 3-10 and $\alpha$-helix so a molecule may transition from one structure to the other without losing the characteristics of a helix at any point [25]. However,

further studies have failed to confirm that the 3-10 helix may present a characteristic CD signature [26]. Polyproline helices have CD spectra which are quite distinct from those of $\alpha$-helices, as seen in Figure 2.2 (in Chapter 2), exhibiting a single, intense minimum at wavelengths in the region of 205 nm and a slight positive peak beyond 220 nm.

### *3.1.2 Secondary Structural Characteristics of Peptoids*

Much like peptides, peptoids adopt a variety of different secondary structural conformations, with increasing reports of new motifs as the range of possible primary sequences is explored. Some of these potential conformations have been examined in detail through computational studies. These are discussed in a review of the computational modelling of peptoids in Chapter 4. CD spectroscopy has also been a useful tool in identifying stable secondary structural motifs in peptoids and in the following sections we investigate the CD spectral characteristics associated with particular peptoid secondary structural motifs.

#### *The Peptoid Helix*

A variety of peptoid sequences have been found to have CD spectra that indicate that they adopt helical secondary structural conformations. In particular, these have been widely observed to exhibit similar spectroscopic characteristics to those associated with $\alpha$-helices, despite the lack of H-bonding ability of the peptoid backbone. The *N*spe peptoid helix is the most extensively studied and well characterised peptoid secondary structural element thus far and was first reported in 1997 by Armand *et al.* [5].

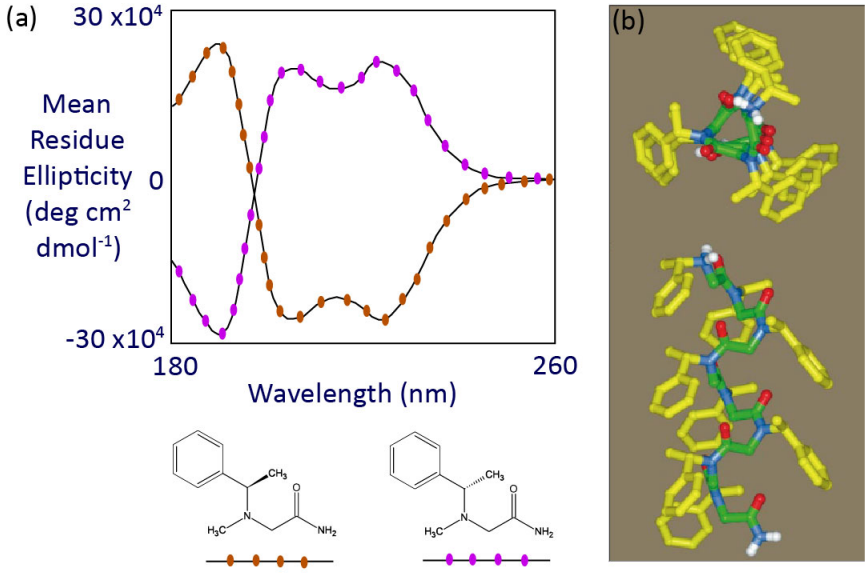



*Fig. 3.4:* (a) CD spectra representative of peptoid helices with opposite handed chiral side chains *N*spe and *N*rpe. (b) Structure of the poly-*N*spe helix proposed by Armand *et al.* viewed from the top and the side. Both adapted with permission from [5].

Poly-*N*spe has been predicted to form a three residue per turn, right handed helix with a pitch of approximately 6.7 Å and all *cis* amides [5]. This structure, shown in Figure 3.4b, was predicted computationally through a combination of *ab initio* and atomistic force field methods (which are discussed in Chapter 4). The *cis*-amide conformation is thought to be enforced by the bulky nature of the *N*spe side chain groups, the structure of which is shown in Figure 3.4a. The backbone dihedrals, $\phi$ and $\psi$, were predicted to adopt average angles of approximately -75° and 170° respectively. These angles are very similar to those seen in the PPII helix, however poly-*N*spe's CD spectrum, shown in 3.4a, more closely resembles that of a peptide $\alpha$-helix, or the less common but structurally similar 3-10 helix [4].

The predicted backbone dihedral angles and all-*cis* amides for poly-*N*spe match those obtained by NMR for the major conformation of $(N\text{spe})_5$ in methanol [27]. However, this work also predicted the existence of minor conformations with the presence of some *trans* amides, indicating that slow transitions between different conformations occur alongside more rapid, small fluctuations in the major conformation. This results in considerable conformational flexibility [27]. As yet, no attempts have been made to predict a theoretical CD spectrum from simulations using this computational model for poly-*N*spe and therefore we ultimately do not yet know if this model is correct or how changes in the model structure affect the CD spectrum. Application of techniques such as TD-DFT to directly generate a theoretical spectrum from a simulated trajectory of computational structures could prove extremely useful in further validating the model and the efficacy of methods such as molecular dynamics in capturing the structural fluctuations.

### *Helicity Modulation*

The proposed 3 residue per turn helical conformation embodied by poly-*N*spe is not exclusive to pure *N*spe helices but has been shown to be induced by other aromatic side chain groups with chiral centres that have similar CD spectra [28]. Researchers are increasingly identifying relationships between sequence and structure that can be used as design parameters to create sequences with certain structural properties. Stringer *et al.* suggested that to form stable peptoid helices (in organic solvents such as acetonitrile) reminiscent of the poly-*N*spe helix, sequences should contain at least 50% chiral aromatic residues and these should be positioned in the sequence to maximise the the $\pi$-$\pi^*$ interactions between side chain aromatic groups [3]. In addition, the placement of chiral residues at and close to the C-terminus also increases helical character in sequences composed of combinations of chiral and achiral residues.

However, these requirements are not particularly conducive to solubility in aqueous solution, due to the hydrophobic nature of the aromatic groups. In order to increase the water solubility of helical sequences the systematic incorporation of charged residues has

proved a good strategy [29]. Alternatively, Darapeneni *et al.* have demonstrated that functionalisation of the C and N terminus with piperazine groups increases the water solubility of peptoids without the need to modify the primary sequence [30].

The exact conformation of peptoid helices is sequence length dependent, with a minimum length of 5 residues generally agreed to be required for helix formation and a maximum helical configuration being reached by 12 residues. Beyond this the helices become sequence length independent [31]. This is illustrated in Figure 3.5 which shows that sequences of *N*rpe have quite different characteristic CD signals depending on their length.

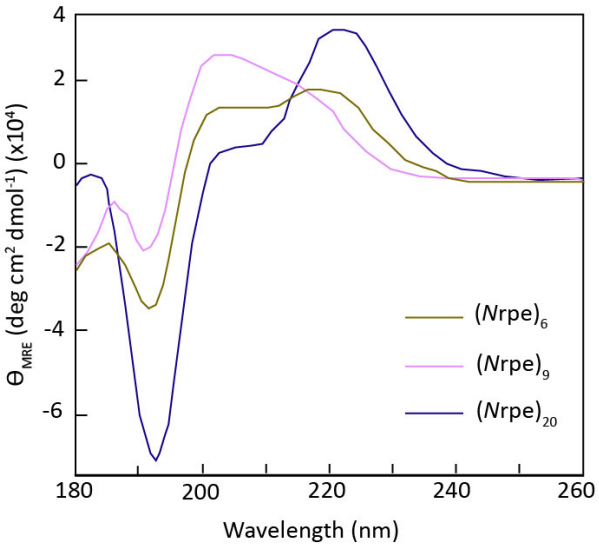



*Fig. 3.5:* CD spectra of *N*rpe polypeptoid sequences of varying length (6, 9 and 20 residues) in acetonitrile. Spectra redrawn with permission from [31].

The spectra for sequences 5-8 residues in length closely resembled the shape of that of the 6-mer. Somewhat unusually the intensity of these spectra decreases with sequence length [31]. Likewise sequences 11 residues and longer were very similar in shape to the 20-mer, but these show an increase in intensity with sequence length. This possibly indicates a switch between two different conformations, with the first being most prevalent in the short sequences and the second increasing in stability as the sequence length increases beyond 11 residues [31].

The handedness of a peptoid helix can be controlled through the choice of side chain. Substituting *N*spe for its enantiomer, *N*rpe, inverts the CD spectrum and the sign of the backbone dihedrals, indicating a switch to helices of the opposite handedness [28], as seen in the spectra in Figure 3.4a. The CD spectrum for the *N*spe helix was found to retain its characteristic features in a variety of aqueous and organic solvents including acetonitrile and methanol, though solvent dependent shifting of the positions of features was observed [28].

The overall nature of the helices formed by peptoids can primarily be controlled

through two mechanisms; altering the length of the sequence or number of helix inducing residues, (*N*spe or others), and the positioning of the residues at particular sites in the primary sequence. The latter was investigated by Shin *et al.* in a study of 30 peptoid sequences with strategically placed chiral, helix inducing *N*spe residues and achiral *N*pm residues that alone do not produce a CD signal [32]. The structure of *N*pm is almost identical to *N*spe (shown in Figure 3.4), but the former lacks the methyl group which imparts chirality to the latter.

Analysis of the CD spectra of these sequences revealed that certain positions in the sequence have greater influence over the overall secondary structure than others. The CD spectra for heptamer sequences containing only 1 *N*spe residue were found to vary depending on the positioning of said residue. Sequences with the chiral residue placed near the C and N termini had CD spectra with greater intensity and helical character than those where the chiral residue was somewhere in the middle of the sequence. The first position from the C-terminus was identified as particularly influential in producing the strong peak at approximately 198 nm, characteristic of the carbonyl $\pi$-$\pi^*$ transition. In sequences containing only two *N*spe residues, the overall peptoid helicity (measured by CD signal) was greater when the *N*spe residues were placed in neighbouring positions. This was found to be a synergistic enhancement of helicity, in contrast to when *N*spe residues were placed in positions separated by the achiral *N*pm, where the increase in helicity going from one to two *N*spe residues was merely additive [32].

The proportion of chiral and achiral residues was also reaffirmed by this study as important for inducing a helical structure. It was demonstrated that peptoid sequences containing as little as 14% *N*spe could produce helical CD spectra in acetonitrile and that the intensity of the spectra increased with increasing proportion of *N*spe residues. Though the authors assert that the spectra resemble that of a PPI type helix [32], they have more in common with the traditional definition of an $\alpha$-helix. This is corroborated by other studies of similar *N*spe containing peptoids of which the CD spectra are described as resembling most closely that of the peptide $\alpha$-helix [2, 31].

Peptoid sequences containing chiral, aliphatic residues have also been shown, by X-ray crystallography, to form poly-proline type helices [4]. The helices adopt similar backbone conformations to the *N*spe helix as well as the PPI helices observed in peptides. For the first time, an X-ray crystal structure was collected for a peptoid and presented in this study which confirmed that, unlike *N*spe and other aromatic group induced helices, the helices formed by peptoids with aliphatic side chain groups exhibit CD spectra which, consistent with the x-ray structure, indicate a poly-proline type conformation. This indicates that the aromatic side chains such as *N*spe may contribute to the typical helical CD spectrum.

Methods to create structures where the backbone amides adopt an all *trans* conformation and resemble the peptide PPII helix have also been reported. Stringer *et al.* created

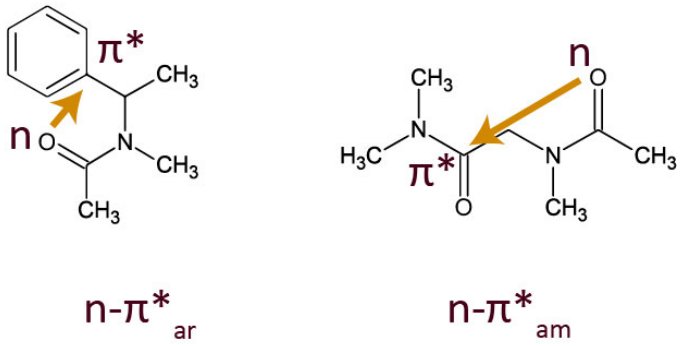



*Fig. 3.6:* Possible n-$\pi^*$ interactions in peptoids which promote the *cis* and *trans* amide conformations in sequences with particular side chain groups [33].

a turn structure by the systematic placement of N-aryl side chains which form structure-stabilising hydrogen bonds with the carbonyl oxygen in the peptoid backbone [34]. The *cis/trans* ratio of the peptoid backbone amides can be controlled via the selection of side chain groups which influence the relative energies of the n-$\pi^*$ interactions. The n-$\pi^*$ interactions in peptoids include the n-$\pi^*$ amide (n-$\pi^*_{am}$) interactions between the carbonyl oxygen and the preceding carbonyl carbon and the n-$\pi^*$ aryl (n-$\pi^*_{ar}$) interactions between the carbonyl oxygen and aromatic side chain groups, if any are present. These are illustrated in Figure 3.6. Generally, the presence of electron deficient aromatic rings, such as those in the *N*spe side chain, promotes a *cis* amide conformation by increasing the strength of the n-$\pi^*_{ar}$ interactions. In contrast electron rich side chain groups demote the n-$\pi^*_{ar}$ interactions and promote n-$\pi^*_{am}$ interactions due to the donation of electron density from amide nitrogen to the carbonyl carbon [33, 35].

### *The Threaded Loop*

A uniquely folded structure arises in *N*spe nonamers solvated in acetonitrile that deviates from the standard *N*spe helical conformation. This structure, shown in Figure 3.7a, is known as the threaded loop and was discovered by Huang *et al.* [36]. It is an approximately planar loop-like structure stabilised by three hydrogen bonds between the N-terminal amine and the 5th, 7th and 9th residue (from the N-terminus) carbonyl groups and a fourth between the carbonyl group of the 2nd residue and the C-terminal primary amide (Figure 3.7a). Half of the threaded loop backbone residues are in the *cis* conformation and half in the *trans* conformation, while $\phi$ and $\psi$ fluctuate around values of $\pm$ 75 °and $\pm$ 180 °. The side chain groups radiate outwards from the looped backbone, resulting in full exposure of the aromatic groups to the solvent and shielding of the backbone.

The structure was first identified upon examination of the CD spectrum which revealed a single negative band at approximately 203 nm, which is somewhat broad and features a small shoulder at approximately 220 nm (Figure 3.7b). Further investigation via solu-

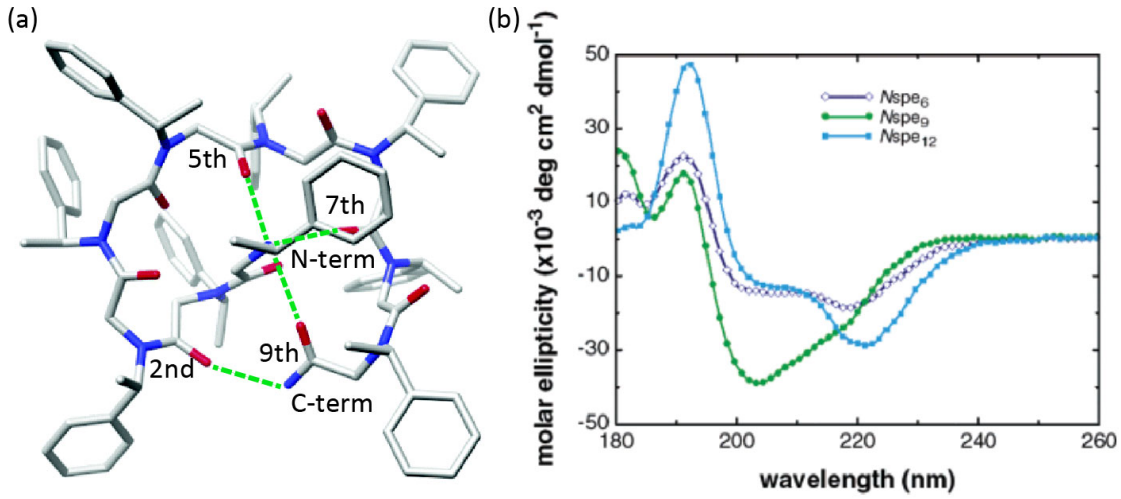



*Fig. 3.7:* (a) Representative solution NMR structure of the threaded loop conformation adopted by *Nspe*$_9$. Hydrogen bonds are shown in lime green and the residues involved in the hydrogen bonding are labelled, along with the C and N termini. (b) CD spectra for *N*spe$_9$ (green), *N*spe$_6$ (purple) and *N*spe$_{12}$ (blue) in acetonitrile. The spectra for latter two sequences are characteristic of the peptoid helix whereas the former is unique to the threaded loop among peptoid sequences. Figure adapted with permission from [36].

tion NMR resulted in the elucidation of the particular structural characteristics described above. As the threaded loop is stabilised by hydrogen bonds it is easily denatured via changes in solvent conditions and its CD spectrum tends towards that of a peptoid helix at high pH.

### *The Peptoid Ribbon*

Further investigation into methods for manipulating the *cis*/*trans* preferences of the peptoid backbone, via careful selection of monomers, lead to the discovery of a new peptoid structural motif by Crapster *et al.*: the peptoid ribbon [37]. The primary sequences investigated in this study consist of alternating *N*s1npe residues and *N*-aryl monomers which were found to favour the *cis* and *trans* amide conformations respectively. Consequently, the peptoid ribbon backbone comprises alternating *cis* and *trans* amides and like many peptoid structures, lacks an intrabackbone hydrogen bonding network. The $\phi$ and $\psi$ values for peptoids adopting a ribbon-like structure were found to be in the region of ($\phi$, $\psi$ = ± 55-64 °, ± 160 °).

Though the detailed structural characterisation was done by solution NMR, Crapster *et al* also collected CD spectra for the peptoid ribbon forming sequences in various solvent conditions. The CD spectra for the ribbon forming peptoids in acetonitrile, a mixture of acetonitrile and water and additionally in methanol are shown in Figure 3.8c. In each solvent the peptoid ribbon CD spectrum has a very broad positive band centred in the region of 205 nm and a negative band in the region of approximately 225 nm. Additionally,

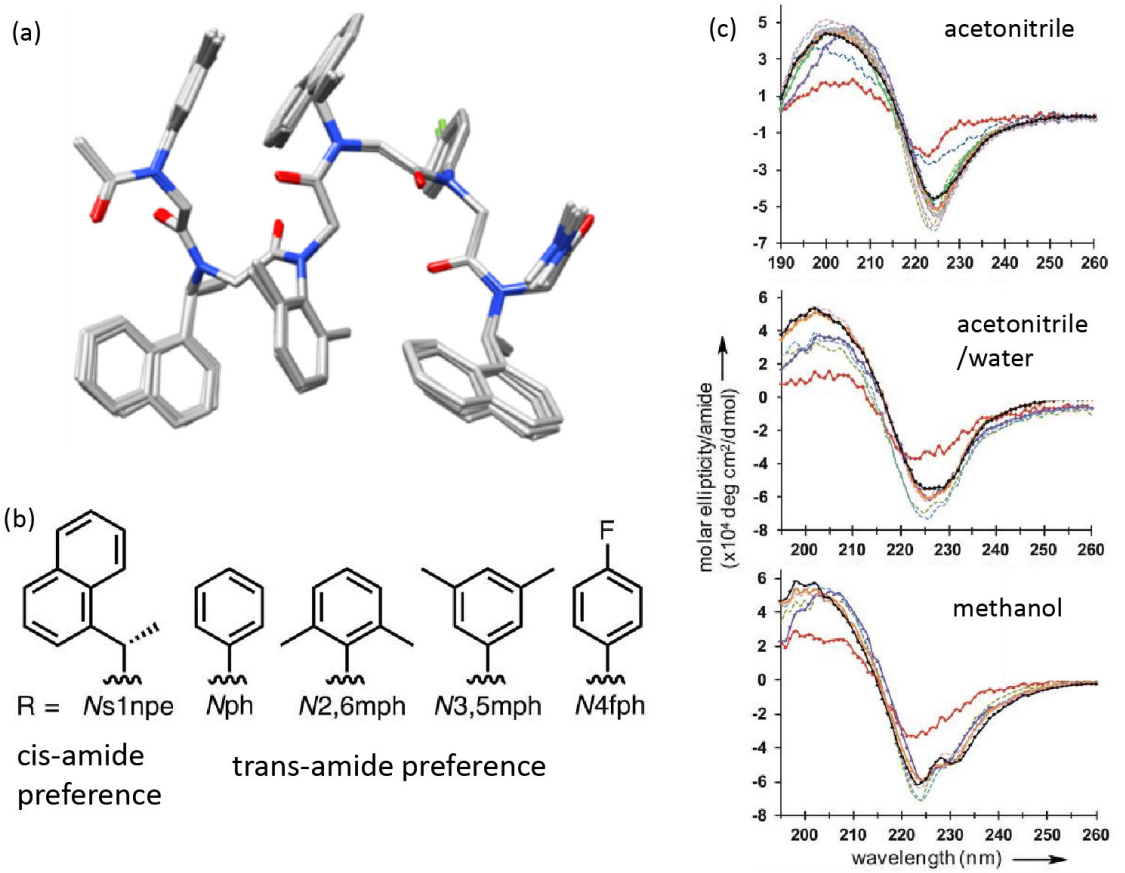



*Fig. 3.8:* (a) Representative solution NMR structure of the peptoid ribbon conformation adopted by Ac-*N*2,6mph-*N*s1npe-*N*ph-*N*s1npe-dma. (b) Structures of the *cis* and *trans* amide inducing side chains used to design the peptoid ribbon. (c) CD spectra for peptoid sequences investigated by Crapster *et al* in acetonitrile, acetonitrile/water and methanol. Figure adapted with permission from [37].

there is a shoulder at approximately 231 nm in the negative band for the peptoids in methanol. By investigating the concentration dependence of the normalised (MRE) CD spectrum, Crapster *et al.* concluded that the peptoid ribbon is a monomeric structure.

### *The Square Helix*

Much work has been directed at controlling the *cis/trans* preferences of the peptoid backbone. Gorske *et al.* demonstrated that the two additional backbone dihedral angles in peptoids can also be influenced by the careful selection of side chains and described a new secondary structure, termed the $\eta$-helix due to its similarities to the peptide $\pi$ helix [38]. This is a square helix with alternating *cis* and *trans* amides. The backbone dihedrals are similar to those associated with the peptoid ribbon but with the sign of the $\phi$ dihedral alternating due to the systematic incorporation of both enantiomers of a 1-naphthylethyl residue. Within this structure the $\psi$ angles are clustered around 180° and $\phi$ alternates

between +60° and -60°, according to the stereoconfiguration of the side chain groups. This results in a square helical structure in octamers, with characteristic 90° turns in the backbone.

### *Peptoid Turn Structures in Cyclic Sequences*

Thus far we have discussed only linear peptoid sequences. Cyclisation naturally limits the conformational freedom of peptoids and has been found to induce unique structural characteristics. Head to tail macro-cyclisation of achiral peptoids has been shown by the Kirshenbaum group to result in a turn structure with alternating *cis* and *trans* backbone amides [39]. This closely resembles $\beta$-turns seen in peptides. The secondary structure of the peptoids was determined by X-ray crystallography in this case. The $\phi$ and $\psi$ dihedrals observed in this structure show very little variation with sequence length and average at $\pm$ 80° and 180° respectively. The amide bond however, shows significant deviation from planarity, with a standard deviation of approximately 12° and 14° for the *cis* and *trans* conformation respectively; more than double that reported in cyclic peptide turn structures [39].

Turn structures may also be induced in linear peptoid sequences by the incorporation of specific side chain groups with turn-inducing powers. Triazole rings have been used to this purpose [40] along with various *N*-aryl groups. The latter can stabilise turn structures through hydrogen bonding between the side chains and backbone [34].

### *Higher Order Structures*

There are several distinct higher order structures that have been reported in peptoids. The self-assembly of nanosheets at the air-water interface from peptoids with two fold periodicity in the primary sequence have been of particular interest. These sheets have been characterised by atomic force microscopy (AFM), X-ray diffraction and transmission electron microscopy (TEM) [41, 42]. These structures are robust to changes in pH and resist chemical denaturation, increasingly as the length of the primary sequence of the constituent peptoids increases. Computational studies of peptoid nanosheets using molecular dynamics and an atomistic force field revealed that the formation of nanosheets is linked to the propensity of particular peptoid sequences to adopt a novel secondary structural motif in which the backbone remains linear [6]. The nanosheets are proposed to be heterogeneous and possibly porous to both water and ions.

The formation of hollow, crystalline nanotubes from diblock polypeptoids has also been reported [43]. Though the mechanisms of peptoid self-assembly have not been fully elucidated, there are several reports of investigations into their properties and the use of these structures in biomedical applications [44–46].

### 3.1.3 Robust Nature of Peptoid Secondary Structures

The various different peptoid secondary structures all appear to maintain folded conformations over a wide range of temperatures, as observed by investigating the CD spectra. The peptoid ribbon showed a linear decrease in spectral intensity with increasing temperature between 10° and 75°, maintaining the same spectral shape throughout [37]. Likewise the threaded loop was also found to maintain its unique CD signal between 25°C and 65°C [36]. The peptoid helix equally appears very robust to thermal changes, with the CD spectrum virtually unchanged over a wide range of temperatures. The helices are also resistant to chemical [2] and biological denaturation [47].

## 3.2 Peptoid Library

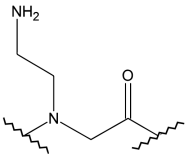
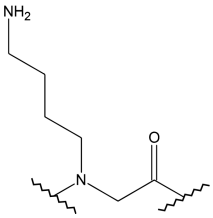
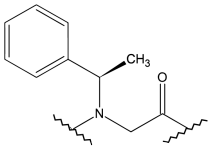
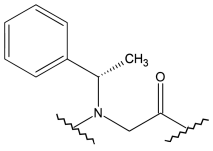
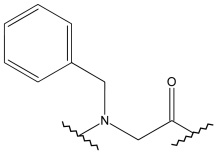
The peptoids investigated in this thesis were synthesised by members of the Cobb Group at Durham University. From their extensive peptoid library a small subset was selected to examine in detail the relationships between sequence, structure and activity. The selected peptoids were chosen for either their interesting structural rearrangements or notable biological activity.

### 3.2.1 Side Chain Groups

In this section we introduce a library of peptoid sequences containing a small selection of different aromatic and charged side chain groups. While a significant proportion of aromatic side chains containing a chiral center are required to induce the formation of a helical secondary structure, inclusion of charged side chains can also promote water solubility. The main aromatic residue considered in this work is *N*spe, which has been discussed widely in the peptoid literature and is known to induce helical secondary structure. We also briefly consider the substitution of *N*spe for the opposite handed equivalent, *N*rpe and a similar aromatic but achiral side chain, *N*phe. Two types of charged side chains are considered, *N*ae and *N*Lys. The *N*Lys residue is the peptoid equivalent of the naturally occurring amino acid lysine. *N*ae differs from *N*Lys only in that the side chain contains two fewer carbon atoms and is therefore referred to as "shorter" in discussion. The structures and full chemical names of each of these side chains are shown in Table 3.1.

### 3.2.2 Primary Sequences

In this work we exclusively consider linear peptoid sequences, all of which are 12 residues in length and thus relatively short in comparison to many antimicrobial peptides. We

| Side Chain | Structure | Full Name |
|---|---|---|
| *N*ae |  | *N*-(2-aminoethyl)glycine |
| *N*Lys |  | *N*-(4-aminobutyl)glycine |
| *N*spe |  | *N*-(S-phenylethyl)glycine |
| *N*rpe |  | *N*-(R-phenylethyl)glycine |
| *N*phe |  | *N*-(benzyl)glycine |

*Tab. 3.1:* Chemical structure, full chemical name and shorthand name of each peptoid monomer discussed in the experimental and computational results within this thesis.

divide our complete library of peptoids into two sub libraries: the repeat motif sequences and the scrambled sequences. The repeat motif peptoids were selected to investigate due to their biological activity against a range of bacteria and parasites. The details of this activity are discussed in Chapter 5. The scrambled sequence library contains a series of rearrangements of the primary sequence of two of the repeat motif peptoids. Every sequence in the library has an unmodified N-terminal group and an amidated C-terminal group.

### *Repeat Motif Sequences*

The repeat motif peptoids contain mixtures of the aforementioned charged and aromatic side chains in varying proportions, arranged in three different repeat motif sequences which are shown in Figure 3.9, where blue circles represent a charged residue (*N*ae or *N*Lys) and yellow circles represent an aromatic residue (*N*spe, *N*rpe or *N*phe). Throughout the thesis we use this colour scheme to represent charged and aromatic residues non-specifically. In

some cases we also use pink circles to specifically represent *N*ae residues and teal circles to specifically represent *N*Lys residues. The different repeat motif peptoid sequences

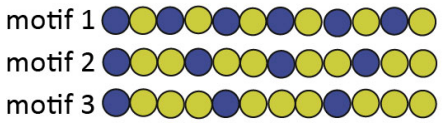



*Fig. 3.9:* Arrangements of charged (blue) and aromatic (yellow) residues in repeat motif peptoid sequences.

have different net charges in aqueous solution due to varying proportions of charged and aromatic residues between motifs. Motif 1 has a total charge of +6 and aromatic content of 50%, motif 2 a charge of +4 and aromatic content of 67% and motif 3 a charge of +3 and aromatic content of 75%. In this chapter we examine how the relative ratios and positioning of charges affect peptoid secondary structure in different solvent conditions. Later, in Chapter 5 we investigate how these variations affect the antimicrobial activity of the peptoids.

### *Scrambled Sequences*

The scrambled sequences are peptoids with primary structures that are a series of rearrangements of the motif 2 peptoids. The motif 2 peptoids were initially selected as interesting molecules to study due to their significant antimicrobial properties. The investigation of the scrambled sequences allows insight into whether the distinct patterning of charged and aromatic residues is necessary for the peptoids to maintain the biophysical properties to which their biological activity is at least partially attributed. The scram-

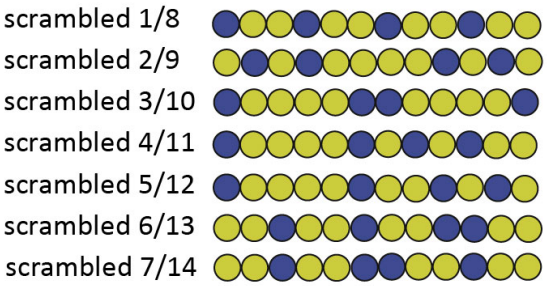



*Fig. 3.10:* Arrangements of charged (blue) and aromatic (yellow) residues in scrambled peptoid sequences. Numbers 1-7 refer to sequences where the charged side chain is *N*ae and 8-14 refer to sequences where the charged side chain is *N*Lys.

bled sequences have the same proportions of charged and aromatic residues as the motif 2 peptoids (a charge of +4 and 67% aromatic content). The key difference between the peptoids in this library is therefore the distribution of charged and aromatic residues along the peptoid backbone. The charges are distributed fairly evenly along the backbone and

none of the sequences contain more than 2 neighbouring charged residues. Some of the sequences do contain regions with up to 4 neighbouring aromatic residues.

*Summary of Complete Peptoid Library and Naming Convention*

Table 3.2 lists the full library of primary sequences considered in this chapter and throughout the thesis. The following naming convention is adopted for all future references to individual peptoids: each peptoid is allocated a number prefaced with letters referring to the sub-library to which it belongs. The first of the repeat motif peptoid sequences is therefore referred to as RM1 and the first of the scrambled sequences as S1. Note

| Peptoid Reference Name | Primary Sequence |
|---|---|
| RM1 | $(NaeNspe)_6$ |
| RM2 | $(NLysNspe)_6$ |
| RM3 | $(NaeNspeNspe)_4$ |
| RM4 | $(NLysNspeNspe)_4$ |
| RM5 | $(NaeNspeNspeNspe)_3$ |
| RM6 | $(NLysNspeNspeNspe)_3$ |
| RM7 | $(NaeNrpeNrpe)_4$ |
| RM8 | $(NLysNpheNphe)_4$ |
| S1 | $(NaeNspeNspe)_4$ |
| S2 | $(NspeNae)_2(Nspe)_4(NaeNspe)_2$ |
| S3 | $Nae(Nspe)_4(Nae)_2(Nspe)_4Nae$ |
| S4 | $Nae(Nspe)_4(NaeNspe)_3Nspe$ |
| S5 | $Nae(Nspe)_4Nae(Nspe)_2(NaeNspe)_2$ |
| S6 | $(NspeNspeNae)_3Nae(Nspe)_2$ |
| S7 | $(NspeNspeNae)_2(NaeNspeNspe)_2$ |
| S8 | $(NLysNspeNspe)_4$ |
| S9 | $(NspeNLys)_2(Nspe)_4(NLysNspe)_2$ |
| S10 | $NLys(Nspe)_4(NLys)_2(Nspe)_4NLys$ |
| S11 | $NLys(Nspe)_4(NLysNspe)_3Nspe$ |
| S12 | $NLys(Nspe)_4NLys(Nspe)_2(NLysNspe)_2$ |
| S13 | $(NspeNspeNLys)_3NLys(Nspe)_2$ |
| S14 | $(NspeNspeNLys)_2(NLysNspeNspe)_2$ |

*Tab. 3.2:* Primary Sequences of each peptoid investigated in this thesis. For clarity, *N*ae residues are shown in pink, *N*Lys residues in teal, *N*spe residues in black, *N*rpe residues in gold and *N*phe residues in grey.

that under this naming system, two peptoids actually appear twice: $(NaeNspeNspe)_4$ and $(NLysNspeNspe)_4$ appear as RM3 and S1 and RM4 and S8 respectively (due to their belonging to both the repeat motif and scrambled libraries).

### 3.2.3 Experimental Aims

By investigating this library of peptoids we aim to address two key questions about the relationship between sequence and structure that have not yet been fully explored in the literature. Firstly, the repeat motif sequences allow us to investigate how the relative proportion of different charged and aromatic residues affects the peptoid structure and whether the nature of the cationic side chain is influential. Secondly, we use the scrambled sequences to examine how the secondary structure varies in sequences with identical proportions of charged and aromatic residues but arranged in different orders, hence giving insight into how the charge distribution along the chain affects the secondary structure of the peptoid.

## 3.3 Results and Discussion

### 3.3.1 Repeat Motif Peptoid CD Spectra in PBS and Octanol

The CD spectrum for each peptoid at 25 $\mu$ M concentration was measured in PBS (0.01 M) at 20 °C. PBS was selected as a suitable buffer as its osmolarity and ion concentrations match those of human blood plasma, allowing for a discussion of the results in a biological context later in this thesis. The repeat motif peptoids were also dissolved in octanol and their CD spectra measured, allowing for a comparison of the conformations adopted in organic and aqueous environments. Though the peptoids do not vary in length the CD data has been converted to mean residue ellipticity (MRE) for reasons of both convention and to enable comparison to any future work on sequences of different length.

#### *Spectral Characteristics in PBS*

The CD spectra for RM1-6 in PBS are shown in Figure 3.11a. Most of the spectra exhibit features similar to those that are characteristic of right handed, helical secondary structures in peptides and proteins. For $\alpha$-helical peptides these features include a positive band at approximately 193 nm and two negative bands (minima) in the regions of 208 nm ($\lambda_1$) and 222 nm ($\lambda_2$). The repeat motif peptoid CD spectra contain broadly similar features, although some blue shifting of the minima to shorter wavelengths is observed, to different extents depending on the primary sequence. These features are also very similar to those observed in the spectra of pure *N*spe sequences in both aqueous and organics solvents [2, 28, 31]. The spectrum for RM4 (shown as the dashed green line in Figure 3.11a) is also qualitatively similar to that published by Chongsiriwatana *et al.* for the same sequence in 10 mM Tris buffer [29], though the signal intensity is greater.

Each of the spectra shown in Figure 3.11a have helical features. However, there is

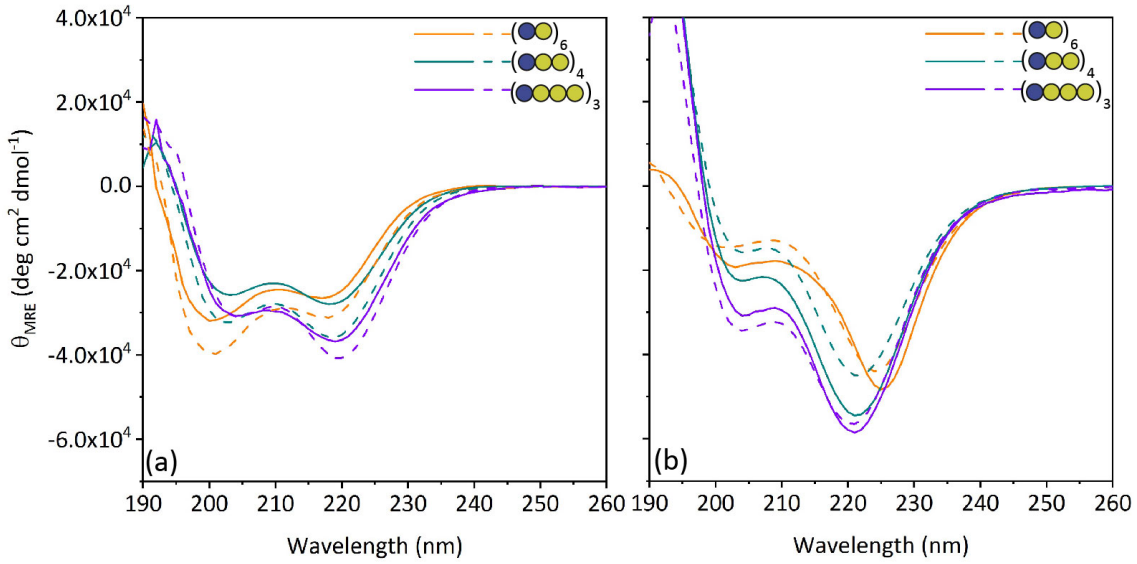



*Fig. 3.11:* CD spectra of RM peptoids 1-6 in (a) 0.01 M PBS and (b) octanol. The Y-scale is the same for both sets of spectra. Motif 1 sequences are shown in orange, motif 2 in teal and motif 3 in purple. The *N*ae side chain versions are shown as solid lines and the *N*Lys side chain versions are shown as dashed lines. The concentration of each peptoid was 25 $\mu$M and spectra were collected at 20 °C using the scan parameters detailed in Chapter 2.

variation in the spectral shape, intensity and the position of features between the different sequences. The spectra for the motif 1 peptoids in particular appear quite different from the others. In the case of the motif 2 and 3 sequences $\lambda_2$ is slightly more intense than $\lambda_1$, which is consistent with the published spectra of many $\alpha$-helical peptides and proteins, and also the peptoid *N*spe helix [5, 31]. The spectra for the motif 1 peptoids on the other hand have minima at $\lambda_1$ that are considerably more intense than $\lambda_2$. These two spectra share this characteristic with the spectrum corresponding to the threaded loop conformation [36]. However, in the latter case the difference relative to the helical spectrum is more extreme, with $\lambda_2$ for the threaded loop being considerably less well defined, to the point it is often described as a shoulder rather than a distinct minimum [48].

### *Positions of Minima*

The positions of the spectral minima vary slightly between sequences. The relative proportion of charged and aromatic residues appears to influence this more strongly than the particular structure of the charged side chain. Substituting *N*ae for *N*Lys has a negligible effect on the positions of the minima across all three motifs. The negative bands are centred around longer wavelengths for sequences where the percentage of *N*spe residues is higher and the overall peptoid charge is lower. As with the relative intensity of the minima, the difference is more pronounced between the motif 1 and 2 peptoids than the motif 2 and 3 peptoids. These spectral differences are illustrated in Figure 3.12 where

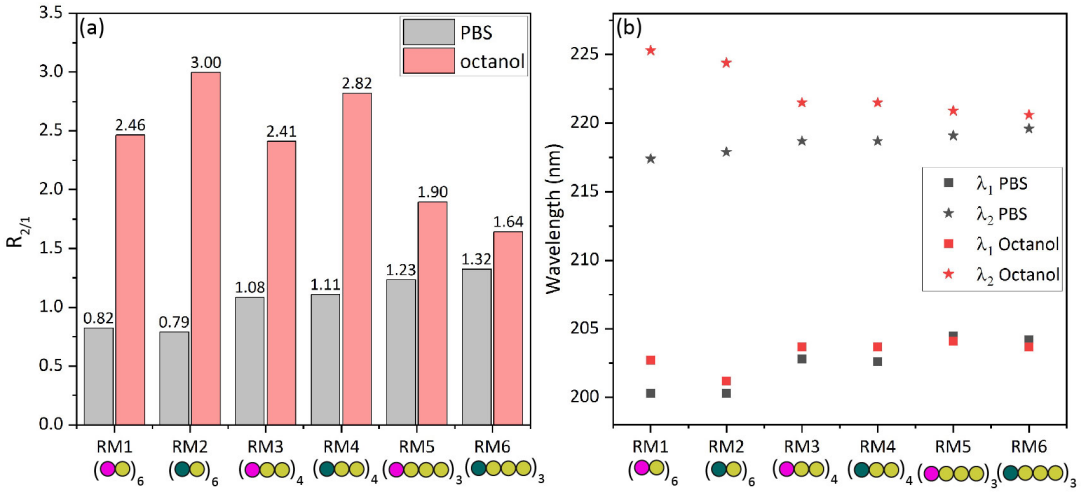



*Fig. 3.12:* (a) Ratio of the spectral intensity at the second minimum ($\lambda_2$) to intensity at the first minimum ($\lambda_1$), $R_{2/1}$, for repeat motif peptoids 1-6 in 0.01 M PBS (grey) and octanol (pink). (b) Positions of $\lambda_2$ (square markers) and $\lambda_1$ (star markers) from the CD spectra of each of repeat motif peptoids 1-6 in 0.01 M PBS (black) and octanol (red). The motif patterns are displayed as ball graphics where magenta represents an *N*ae residue, teal represents an *N*Lys residue and yellow represents an *N*spe residue.

part a shows the ratio of the mean residue ellipticity at $\lambda_2$ to $\lambda_1$ ($R_{2/1}$) and part b shows the positions of $\lambda_1$ and $\lambda_2$ for each peptoid sequence in PBS and also 1-octanol, which is discussed subsequently.

### *Intensity of Spectral Features*

Another notable quality of the spectra shown in Figure 3.11a is that the positive peak at approximately 195 nm is quite diminished relative to that typical for an $\alpha$-helical peptide, where it is more intense than either of the minima. The reasons for this are unclear. The band is poorly resolved in our peptoid spectra due to the strong absorbance of the chloride ions in the PBS at wavelengths below 200 nm but this alone should not necessarily be expected to cause such a considerable and consistent reduction in intensity. In addition it seems likely that there is a conformational cause of the low intensity of this feature as it is consistent with many other published spectra for helical peptoids in various solvent conditions [5, 31, 32], though rarely commented upon. The origin of this feature in peptides is traditionally attributed to the $\pi$-$\pi^*$ transition. The altered nature of the backbone amide in peptoids relative to peptides could potentially lead to a conformational change altering the transition energy and hence the intensity of the spectral band.

The spectra shown in Figure 3.11a vary considerably in overall intensity. The magnitudes and positions of the minima for each sequence are listed in Table 3.3. In peptides the magnitude of the CD signal (MRE) is typically thought to indicate to some extent the

overall helical content of each sample, with more intense signal corresponding to greater helical character. Where spectral deconvolution to determine overall helicity by fitting the entire spectrum to reference data sets is not appropriate, the MRE at $\lambda_2$ can be used for quantitative comparison between different sequences. $\lambda_2$ is used in preference to $\lambda_1$ for various reasons. In proteins $\lambda_2$ has been shown, as a first approximation, to vary linearly with the helical content of a protein [49]. This is a somewhat crude analysis, as, if multiple helical configurations are present and the observed spectrum is a combination of different helical states, such a simple relationship is unlikely to hold true. However in the case of peptoids, where currently the lack of reference spectra and deep understanding of the relationships between structure and CD spectrum makes more in-depth analysis impossible, comparing the intensity of the features between different sequences can be a useful first approximation.

$\lambda_2$ is better defined than $\lambda_1$ in the peptoid spectra presented here, with a lower degree of variation between spectral accumulations for single samples and between samples (see Figure 3.13). In addition, $\lambda_1$ is generally less well defined in peptoids than peptides in PBS as it is shifted to shorter wavelengths, closer to the region where the absorbance by chloride ions in the buffer reduces the signal to noise ratio. Vaz and Brunsveld suggest that, in peptoid CD spectra, $\lambda_2$ is directly proportional to the helical content and $\lambda_1$ inversely proportional, with the ratio between them, $R_{2/1}$, therefore increasing as the helical character increases [50]. This parameter can therefore be used as an alternative measure of sequence helicity.

| Solvent | Peptoid | $\lambda_{min1}$ (nm) | $\lambda_{min2}$ (nm) | $\theta_{MRE,min1}$ | $\theta_{MRE,min2}$ |
|---|---|---|---|---|---|
| PBS | RM1 | 200.3 | 217.4 | -32238 | -26541 |
| | RM2 | 200.3 | 217.9 | -39707 | -31369 |
| | RM3 | 202.8 | 218.7 | -25864 | -28053 |
| | RM4 | 202.6 | 218.7 | -32464 | -36198 |
| | RM5 | 204.5 | 219.1 | -30918 | -36971 |
| | RM6 | 204.2 | 219.6 | -31079 | -40576 |
| Octanol | RM1 | 202.7 | 225.3 | -19649 | -48402 |
| | RM2 | 201.2 | 224.4 | -14694 | -44021 |
| | RM3 | 203.7 | 221.5 | -22581 | -54451 |
| | RM4 | 203.7 | 221.5 | -15981 | -45083 |
| | RM5 | 204.1 | 220.9 | -30854 | -58475 |
| | RM6 | 203.7 | 220.6 | -34460 | -56575 |

*Tab. 3.3:* Wavelengths of $\lambda_1$ and $\lambda_2$ with corresponding MRE values (in units of deg cm$^2$ dmol$^{-1}$) for repeat motif peptoids 1-6.

Two general trends are observed in the spectral intensity, and thus helicity by these measures, of the repeat motif peptoids in PBS. A higher proportion of the helix inducing *N*spe residues results in a more intense helical signal at $\lambda_2$. Substituting the shorter

charged side chain (*N*ae) for the longer (*N*Lys) also results in a more intense helical signal at $\lambda_2$. The overall magnitude of the signal for each peptoid therefore appears to depend on a balance of these two factors. This is illustrated by comparing the spectra of RM2 and RM3 where the former is more intense at $\lambda_2$ than the latter, despite the sequence containing fewer *N*spe residues. The enhanced helical character of the motif 3 sequences (and correspondingly lesser helical character of the motif 1 sequences) could be a result of both the relative proportions of the stabilising *N*spe residues and the increased number of these residues in the C-terminal region of the motif 3 sequence. The latter is proposed by Wu *et al.* to help counteract the enhanced flexibility of the carboxamide group at the C-terminal end of the sequence relative to the N-terminal end [31].

The proportion and positioning of the charged residues may also contribute to the moderation of peptoid helicity. Both the length of the charged side chains and their positioning within the sequence have been found to affect helix formation in peptides [51, 52]. In peptide sequences substitution of lysine, the amino acid equivalent to *N*Lys, for equivalent units with shorter side chain groups, such as ornithine and diaminobutyric acid (Dab) (the peptide analogue of *N*ae), reduces the intensity of the helical CD signal by destabilising $\alpha$-helices. Lysine is known as a helix promoting residue but the shorter side chain equivalents destabilise helices increasingly as the side chain length is reduced [51]. This is thought to occur because, in sequences with the shorter cationic side chains, the charged NH group on the side chain will be in closer proximity to the backbone, which it can hydrogen bond with, disrupting the characteristic H-bond network that would otherwise stabilise the helix.

We observe a similar effect in the peptoid repeat motif sequences as substituting *N*Lys for *N*ae reduces the helicity of each sequence. However, the origins of this effect must be different in peptoids than peptides as the former do not have the capacity to stabilise helices via an intra-backbone H-bond network and therefore would not be disrupted by side chain to backbone hydrogen bonds. Helix formation in peptoids will likely result in a loss of conformational entropy of the side chain groups [52], to a greater extent in the shorter groups (*N*ae) than the longer (*N*Lys) which would result in the observed reduction in helical character of the *N*ae sequences as the entropic penalty of folding would be greater.

### *Spectral Variations with Solvent Conditions*

CD spectra were collected for each of the repeat motif peptoids dissolved in 1-octanol. These are shown in Figure 3.11b. The positions of the $\lambda_1$ and $\lambda_2$ minima and the corresponding MRE values for each of the repeat motif peptoids in both PBS and octanol are summarised in Table 3.3. The ratios of the two minima for each spectrum are shown in Figure 3.12a. There are clear differences between the spectra for each peptoid sequence in PBS and in octanol. Most notably an increase in the intensity of $\lambda_2$ relative to $\lambda_1$ renders

$R_{2/1}$ significantly greater in the octanol spectra than the PBS spectra. These changes somewhat resemble the sequence length dependent changes in CD spectrum observed by Wu *et al.* in pure *N*rpe helices and Huang *et al.* in equivalent *N*spe helices in acetonitrile. These authors concluded that the magnitude of change in the spectral shape and intensity was indicative of a distinct conformational change and therefore this may be the case for the repeat motif peptoids too.

### *Shift in Position of Spectral Features*

A red shift in the positions of $\lambda_1$ and $\lambda_2$ to longer wavelengths is observed going from PBS to octanol. This is observed to a greater extent in $\lambda_2$ than $\lambda_1$, resulting in a slight broadening of the overall spectrum in each case. The motif 1 peptoids display the largest positional shift of approximately 7 nm, while the motif 3 peptoids have a much smaller 1-2 nm shift, as can be seen in Figure 3.12b.

Differences in the CD spectra of biomolecules between solvents are common and are often a result of structural rearrangement of the chromophore groups in response to the solvent conditions. This alters the transition energies and hence the positions and intensities of spectral features. This is known as a primary effect. CD spectra may also be subject to secondary effects where interactions between the solvent and chromophore group cause an additional alteration in the choromophore electron density distribution, altering the transition energies and hence the appearance of spectral features [53, 54]. Reductions in solvent polarity such as that corresponding to a switch from PBS to octanol have been shown to correspond to red shifting of spectral features for peptides, including the AMP alamethicin [53]. Secondary effects may contribute to the observed differences in the peptoid spectra in PBS and octanol.

It is possible to deconvolute primary and secondary effects on CD spectra due to solvent conditions where the structure of the analyte is known in both solvents by some other technique. However, we do not have this knowledge for the peptoid sequences considered here. The overall magnitude of the observed difference is large enough that it seems likely that there is a rearrangement of the amide bond structure which results in the observed spectral shaped changes, though secondary effects may also contribute to the overall change.

### *Variations in Intensity of Spectral Features*

Changes in spectral shape can be quantified by examining the ratio of the spectral intensity at the second minimum to that at the first minimum, $R_{2/1}$. It is debatable whether physical meaning can reliably be derived from $R_{2/1}$ in terms of how it correlates to changes in secondary structure, particularly between different types of helix. In peptide studies this

ratio has been used as an indicator of what type of helix is present[24] and whether a sample contains single isolated helices or coordinated bundles of helices, such as coiled coils. For alpha helices, $R_{2/1} < 1.1$ has been reported to indicate isolated helices. $R_{2/1} < 0.4$ may indicate the presence of 3-10, rather than $\alpha$-helices. $R_{2/1} > 1.1$ may indicate the presence of coiled coil-type bundles [55]. However, there is little consistency in the literature and no definitive answer as to whether a value for $R_{2/1}$ in a particular range corresponds to a particular isolated helical conformation or bundle of helices [56]. Even if this was the case it would not necessarily follow that the same would be true for peptoids.

However, in some studies $R_{2/1}$ has been used as a semi-quantitative indicator of the helical content of peptoids [50]. Therefore we have calculated $R_{2/1}$ for each peptoid in our library. $R_{2/1}$ for the repeat motif peptoids is depicted in Figure 3.12a and varies considerably between the different sequences and solvent conditions. In PBS the values of $R_{2/1}$ (grey bars) indicate that all the peptoids are likely to exist as isolated helices, except RM5 where $R_{2/1} = 1.23$. The motif 1 peptoids have the lowest $R_{2/1}$ values, they are the only repeat motif sequences where the first minimum is more intense than the second ($R_{2/1} < 1$). This could indicate that these peptoids are adopting a conformation closer to the 3-10 helix than the other peptoids whose spectra appear more $\alpha$-helical in character. This could be due to the higher charge density on the motif 1 peptoids forcing them to adopt a helix with a longer translation along the helical axis, to minimise electrostatic repulsion between the charged residues. However, this is merely speculative and the secondary structural motifs of peptides may not be directly applicable to peptoids at all.

$R_{2/1}$ is much larger in octanol than PBS for each of the peptoids. This is a strong indication that the peptoids adopt distinctly different structures in the different solvents. It is possible that the increase in $R_{2/1}$ in octanol could also correspond to some formation of helical bundles to shield certain residues from exposure to the organic solvent, however no concentration dependent behaviour was observed in either PBS or octanol for any of the sequences except RM5 (which is discussed in detail in Section 3.3.2), indicating that the peptoids are likely to be monomeric. The change in $R_{2/1}$ going from PBS to octanol decreases in magnitude going from motif 1 to motif 3. This suggests that the conformational change could be charge mediated as the sequences with higher net positive charge vary more in spectral shape between solvents than those with lower net charge.

Interestingly, the signal intensity at $\lambda_2$ of the *N*ae side chain variant of each sequence is greater than that of their *N*Lys counterparts in octanol, whereas the opposite is true in PBS. This could imply that the increased length of the cationic side chain in the *N*Lys sequences promotes a stable helical conformation in PBS, as discussed previously, but in octanol has the opposite effect and destabilises the helix. In PBS, for both cationic side chain variants $R_{2/1}$ increases almost linearly as the number of charged residues in the sequence decreases and number of *N*spe residues increases. In octanol $R_{2/1}$ decreases under

the same conditions, though not linearly. The spectra and hence secondary structures therefore appear to change in a predictable manner within the repeat motif sequences. These changes correlate to changes in the primary sequences, though their underlying physical origins remain unclear.

### *3.3.2 Anomalous Behaviour of RM5*

#### *Monomeric Nature of Peptoids*

The CD spectra presented so far indicate that the peptoids in question are monomeric in solution and that no bulk aggregation is occurring at any point during the experiments. In the majority of cases we believe this to be true for a number of reasons. CD spectra for each peptoid were collected over a range of different peptoid concentrations from 10-100 $\mu$M. Higher concentrations resulted in increased raw signal intensity, but each spectrum normalised to the same intensities in the mean residue ellipticity (MRE), within a small margin of error. This consistency between spectra collected over a tenfold increase in concentrations indicated that it is unlikely that the peptoids form any stable intermolecular association, consistent with several reports for similar sequences in the literature [31]. In addition the UV-vis spectra of the stock peptoid solutions (with concentrations typically in the region of 200-500 $\mu$M) did not show a level of background scattering indicative of aggregation. The exception to these observations, the first in particular, was RM5, $((N\text{ae}N\text{spe}N\text{spe}N\text{spe})_3)$ which proved difficult to work with due to its apparent tendency to aggregate in PBS.

#### *Spectral Error Associated with RM5*

The difficulties associated with RM5 are immediately obvious when examining the spectra collected for a series of different samples of this peptoid, as opposed to the others in the library. Figure 3.13 shows the average PBS spectrum for RM1 $((N\text{ae}N\text{spe})_6)$, as shown and described in Chapter 2, along with the equivalent spectrum and associated error for RM5. The pink bands show the error associated with collecting 15 spectral accumulations from the same sample (the peptoid concentration for the sample shown was 25 $\mu$M). The blue band is representative of the error associated with preparing samples at different concentrations (or separate samples at the same concentration) and taking spectra of each of these samples.

The typical error associated with RM5 spectra in PBS is shown in Figure 3.13b and is considerably larger than that associated with any of the other peptoids. In addition to the CD spectrum for RM5 being less reproducible than that of the other peptoids, RM5 showed visible signs of aggregation and appeared to be considerably less soluble in PBS than any of the other peptoids in this study. Dissolving the powdered form of RM5

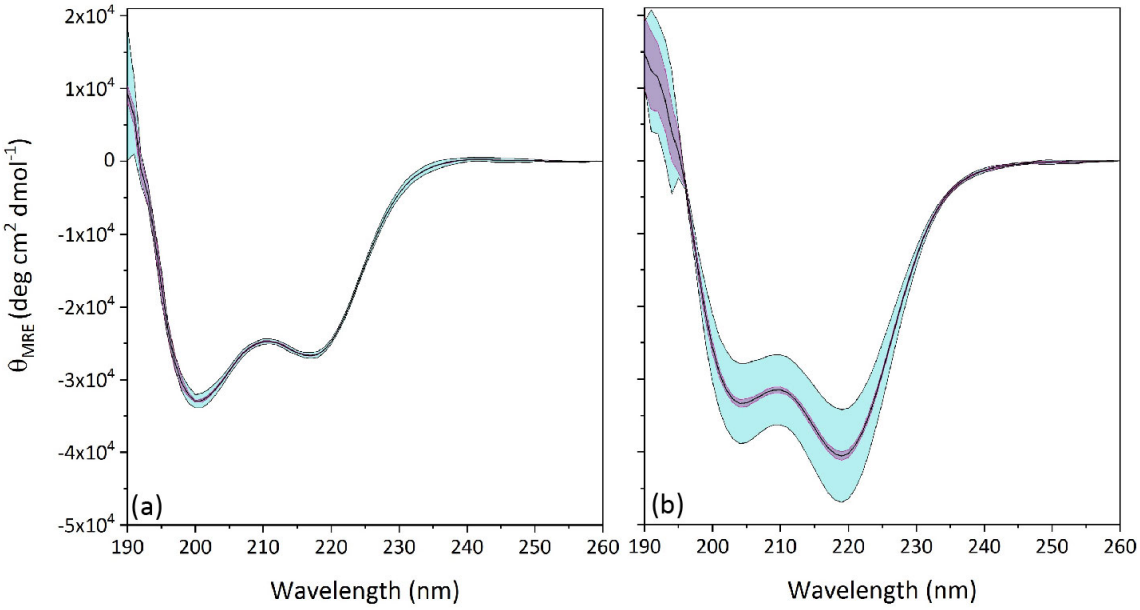



*Fig. 3.13:* (a) Representation of error associated with CD spectra for RM1 at 25 $\mu$M in PBS. Actual displayed spectrum is RM1. (b) Representation of error associated with the CD spectrum for RM5 at 25 $\mu$M in PBS. In both (a) and (b) the average spectrum is shown as a black line. The standard deviation from the average spectrum for multiple measurements of the same sample is shown as the pink bands. The standard deviation from the average spectrum value over a series of samples at various concentrations is shown as the blue bands.

(as supplied) in PBS proved difficult and was achieved by first dissolving in ultra pure water and then adding a concentrated PBS solution to the achieve correct dilution (0.01 M) of the buffer. Upon the addition of concentrated PBS the peptoid solution became visibly cloudy, strongly indicating aggregation and precipitation of the peptoid. When a sample was left undisturbed over a period of several weeks an amount of the peptoid had visibly sedimented and settled in the bottom of the sample tube. By contrast the other peptoids in the library were easily dissolved from powder into 0.01 M PBS and the solutions appeared clear upon initial mixing, remaining so over long periods of time. The spectrum for RM5 in PBS shown in Figure 3.11a is the best estimation of the true spectrum that could be obtained, given the experimental difficulties. This was obtained following an investigation into the behaviour of RM5 and an adapted sample preparation protocol for obtaining spectra which is described subsequently.

### *Investigation of Concentration Dependent Behaviour*

The aggregation of RM5 samples is one of a number of factors that may contribute to the inconsistency in the spectra collected for this peptoid. Aliquots taken from the stock solution could vary in concentration, causing the observed spectral mismatches in Figure 3.13b. However, in addition to being difficult to reproduce individually, the CD spectrum

for RM5 also appears to be concentration dependent as single samples which were serially diluted, with spectra taken at different concentrations, resulted in varied intensity in the MRE. Ideally, the MRE CD spectrum of a monomeric peptide, protein or peptoid should not vary with sample concentration and this behaviour is rarely reported or discussed in the literature. However the MRE could vary due to concentration dependent intermolecular interactions or solubility issues. In the case of RM5, either or both of these factors could be at play.

Initial investigations of the concentration dependence of the RM5 CD spectrum were inconclusive as different behaviour was observed across samples. This may be due to the difficulty of reliably preparing samples of a particular concentration, since initially all CD samples were prepared individually from the aggregated stock solution. In some cases the intensity of the spectrum appeared to vary almost linearly with concentration, with minimal change in spectral shape (Figure 3.14b). In other cases both the intensity and shape of spectra varied with concentration in an unpredictable manner (Figure 3.14a).

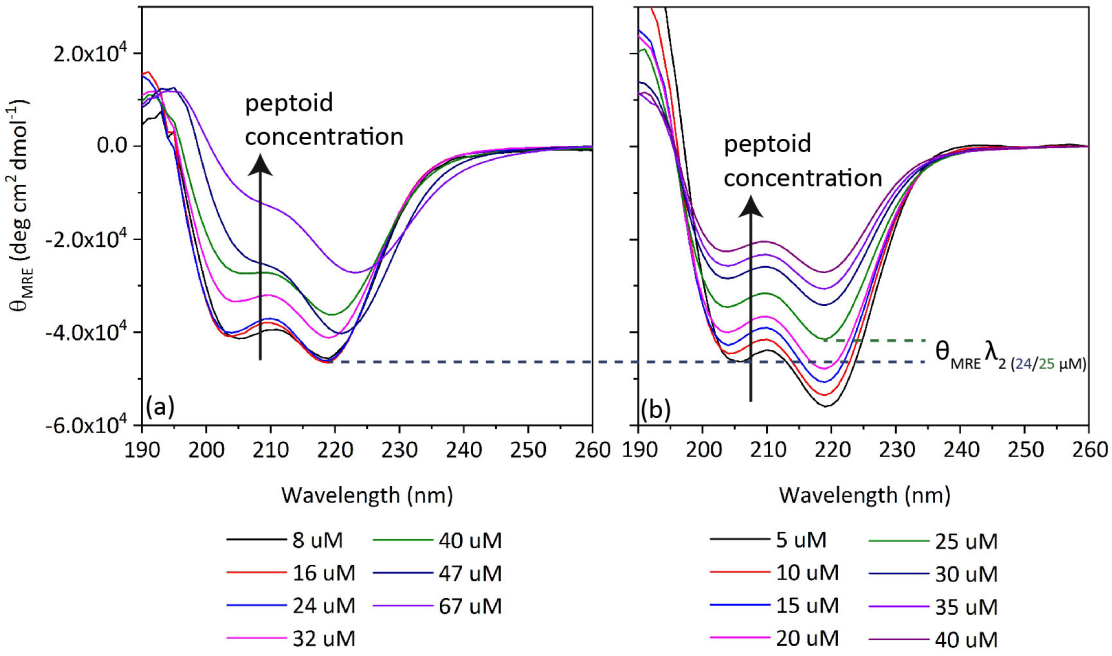



*Fig. 3.14:* CD spectra for peptoid RM5 at various concentrations in 0.01 M PBS. The estimated peptoid concentration for each spectrum is shown underneath the collected spectra. The samples for (a) and (b) were prepared from separate stock solutions and each individual sample was prepared by diluting an aliquot from the stock solution. Dashed lines across (a) and (b) highlight the different intensity of the signal for the supposed 24 $\mu$M and 25 $\mu$M samples.

We confirmed by TEM that RM5 does not aggregate into higher order structures but appears to form amorphous aggregates with considerable variation in size and shape in both water and PBS. TEM images were collected by Dr Budhika Mendis. Energy-dispersive X-ray analysis (EDX) was used to confirm that the observed aggregates (Figure

3.15) were peptoids.

Figure 3.14 clearly shows the variable behaviour of RM5 in PBS. In 3.14a the spectrum at high concentration is distinctly different in shape to the low concentration equivalent. The changes in intensity and shape also appear to plateau at the lowest concentrations. In 3.14b there is no plateau at low concentrations and the shape of the spectra appears less variable. The spectra shown in Figure 3.14 were taken from samples prepared under conditions which were identical as far as possible. The only notable difference was the time between preparing samples and collecting spectra which was not closely documented as it did not affect the spectra of any other sequences. However all stock solutions were allowed to equilibrate for a minimum of 24 hours before samples at different dilutions were prepared and spectra collected that same day.

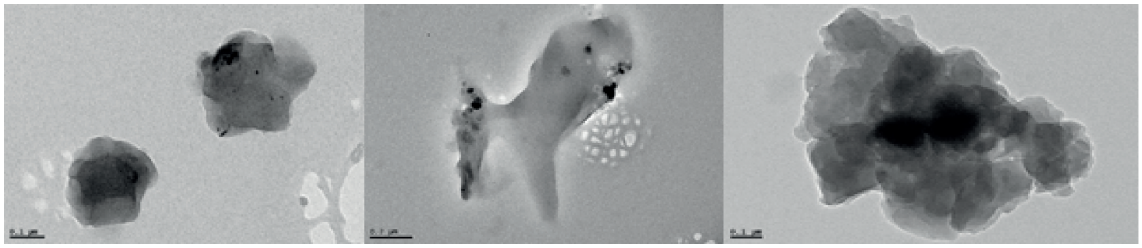


*Fig. 3.15:* TEM images of RM5 peptoid aggregates in PBS.

The solubility of RM5 was investigated via the collection of CD spectra of a 100 $\mu$M sample at intervals (T) over a period of 20 hours following initial preparation from the stock solution. Over this period of time the spectrum varied in shape and increased considerably in intensity (Figure 3.16a). Given this result, it is probable that the various behaviours observed in Figure 3.14 could be an artefact of some samples not reaching the solubility equilibrium within the time between sample preparation and spectra collection.

Following the 20 hour period the same sample used to collect the data shown in Figure 3.16a was heated to 90 °C while the CD signal was monitored at $\lambda_2$ and full spectra collected at 10 °C intervals. The melt spectra are shown in Figure 3.16b and MRE at $\lambda_2$ against temperature is shown in Figure 3.16c. A gradual loss of signal intensity was seen over the lower temperatures until approximately 45 °C where there appears to be a sudden increase in signal intensity, corresponding to presumably an increase in dissolved peptoid forming helical structures. The intensity then decreases again over the course of the rest of the experiment at a faster rate than it had prior to the sudden increase. The spectra shown in Figure 3.16b include a spectrum taken at 20 °C after the sample had been heated then returned to room temperature. This spectrum is shown as the black dashed line and shows the CD signal is considerably more intense after completion of the melt than prior to it (solid black line spectrum), indicating that a greater proportion of the peptoid was dissolved and adopting a helical structure post-melt.

This response to temperature is unusual. While peptides tend to exhibit characteristic

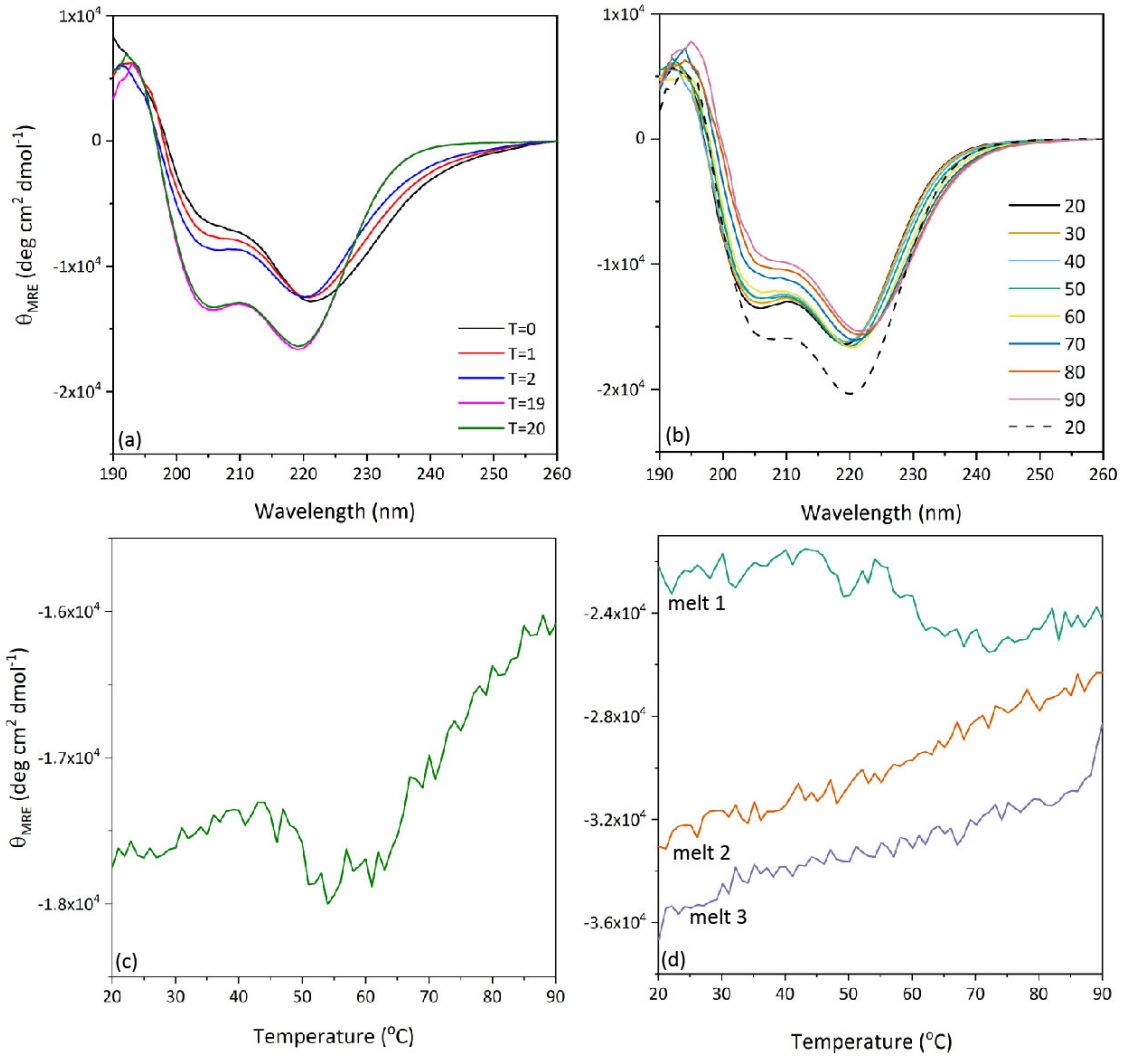



*Fig. 3.16:* (a) CD spectra for $(NaeNspeNspeNspe)_4$ at 100 $\mu$M concentration in PBS (0.01 M). Spectra were collected over a period of 20 hours at time intervals, T. The first spectrum was collected immediately following the dilution of the sample from the stock solution. (b) Spectra collected at 10°C intervals as peptoid was heated from 20-90°C. (c) Melt curve showing variation of CD intensity at $\lambda_2$ over the course of the melt of 100 $\mu$M sample. (d) Curves for 3 successive melts of the same 75 $\mu$M sample of RM5 in PBS.

changes over a similar temperature range, there is considerable evidence that peptoids are extremely thermally robust [3, 4, 28, 31] and would therefore exhibit little CD dependence with temperature, with any changes that do occur likely to be reversible. The experiments carried out to collect the data shown in Figure 3.16a, 3.16b and 3.16c were repeated with separately prepared samples of RM5 at various concentration in the range of 50 to 100 $\mu$M. Qualitatively similar behaviour was observed in each case though the behaviour was exaggerated at higher concentrations.

Figure 3.16d shows three melt curves for RM5 at 75 $\mu$M carried out sequentially on the same sample. The strange characteristic behaviour was only observed in the first melt, though after the second melt the spectrum at 20°C again increased in intensity relative to

what it was prior to the melt. Only after the third melt did the post-melt 20°C intensity of the spectrum closely resemble the spectrum at 20°C prior to that melt. Again this behaviour was observed in several separately prepared samples. This behaviour indicates that solubility of RM5 is sensitive to both concentration and temperature.

In order to establish whether the concentration dependence observed in the CD spectrum is a result of this variable solubility we carried out melts on a single sample which was serially diluted between measurements. These are shown in Figure 3.17a. In a similar manner to the melt curves for repeated melts on the same 100 $\mu$M sample, the melt curves for the serially diluted sample only show the unusual behaviour in the first melt. The curves shown are normalised to the 20°C MRE to highlight the differences in shape, particularly of the 100 $\mu$M curve, that were somewhat smoothed out when they were all displayed on the same axes as MRE.

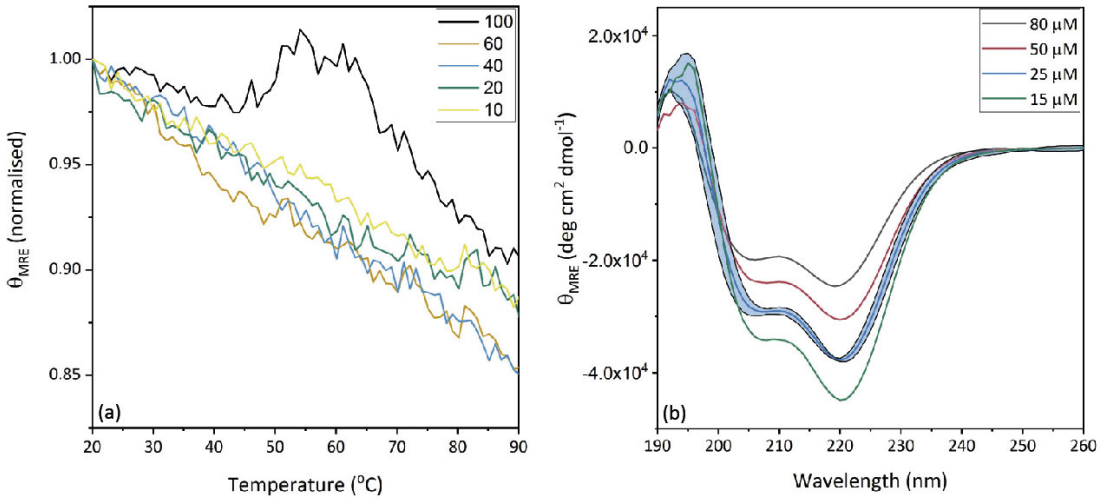



*Fig. 3.17:* (a) Melt curves for serially diluted sample of RM5 at various concentrations (shown in legend in °C). Mean residue ellipticity of the curves is normalised to the 20°C value as the variation in MRE between samples otherwise made the change in intensity with temperature difficult to distinguish with all 5 curves on the same axes. (b) CD spectra of RM5 at varying concentration taken from a single sample that was serially diluted following three melt and cool cycles of the stock samples. Standard deviation of the spectrum for sepearate 25 $\mu$M samples prepared in this manner is shown for the 25 $\mu$M spectrum as the blue band.

In addition we carried out a serial dilution of a sample which had been subjected to 3 melt cycles at high temperature, collecting the spectra without reheating the sample at each dilution. Examples of spectra obtained in this way are shown Figure 3.17b. Clearly there is still a dependence of the MRE on sample concentration, despite the melts somewhat helping to obtain a stock solution as fully dissolved as possible. The standard deviation of several 25 $\mu$M samples prepared in this manner is shown as the blue band enclosing the 25 $\mu$M spectrum. This is considerably smaller than the equivalent error shown for RM5 in Figure 3.13b, indicating that the melting and cooling of the stock

solution followed by dilution improves how accurately we can collect the CD spectrum for this peptoid in PBS, though it is still not as repeatable as the others in the library. Pre-melts were therefore carried out on all RM5 samples in PBS used in subsequent investigations.

The work presented here indicates that the self-association of RM5 in PBS is non-specific. Though we have developed a protocol to minimise the effect of this behaviour in our overall study of the repeat motif sequences we have not elucidated the underlying physical cause of the observed aggregation. Non-specific self association is scarcely discussed in the peptoid literature and never in sequences containing the two side chain groups present in RM5. Fuller *et al.* describe a self-associating, water soluble peptoid sequence composed of several different charged and aromatic side chains (including *N*ae and *N*spe) [57]. The behaviour of this peptoid is moderated by pH and the presence of solvent additives. Future investigation into the behaviour of RM5 in different solvent conditions may help reveal the origins of its unusual behaviour in PBS.

### *Effect of Side Chain Chirality*

The influence of the handedness of the chiral, helix inducing side chain on overall helix handedness has been well documented in sequences composed purely of such residues [28]. Here we confirm a similar effect in sequences composed of a mixture of these residues and those with cationic, non-chiral side chain groups. Peptoids RM7 and RM8 are both motif 2 sequences but with substitutions of the *N*spe residues for *N*rpe and *N*phe to modify the chirality. The *N*spe residue induces an overall right handed helical structure in the peptoids RM1-6. Substituting *N*spe for its opposite handed counterpart *N*rpe inverts the CD spectrum. This is observed in both PBS and octanol, as seen in Figure 3.18a, which shows the PBS and octanol spectra for RM7, $(N\text{ae}N\text{rpe}N\text{rpe})_4$.

While peptides in random coil conformations have a distinctive CD spectrum it is highly unlikely that a peptoid with no overall secondary structural character would have any CD spectrum associated with it at all due to the achiral nature of the peptoid subunits. This is seen in the literature, where sequences which do not contain any side chains with chiral centers have zero CD signal associated with them [32]. The CD spectrum for RM8, an achiral analogue of RM4, was collected in PBS to demonstrate that substitution of the chiral *N*spe side chain for achiral *N*phe results in the loss of overall chirality in the folded peptoid. This does not necessarily mean that the peptoid does not adopt a helical conformation but that helices of both handedness are equally likely. This results in zero CD signal arising from this peptoid in PBS, as seen in Figure 3.18b.

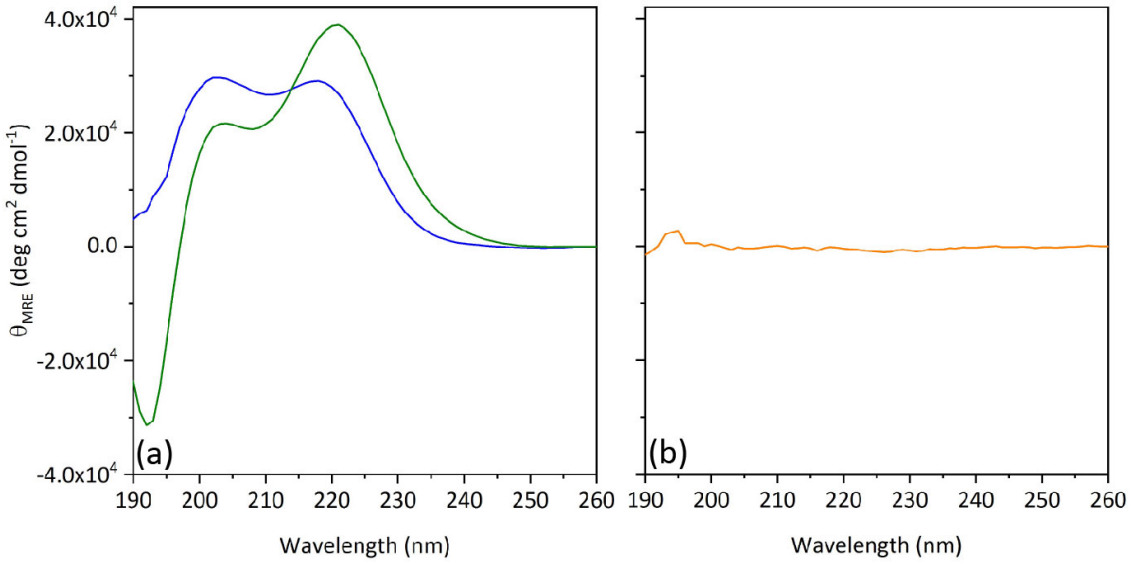



*Fig. 3.18:* (a) CD spectra of peptoid RM7, $(N\text{ae}N\text{rpe}N\text{rpe})_4$, in 0.01 M PBS (blue spectrum) and octanol (green spectrum). (b) CD spectrum of peptoid RM8, $(N\text{Lys}N\text{phe}N\text{phe})_4$ in 0.01 M PBS. This peptoid contains only achiral side chains and therefore exhibits no CD signal. The scale of the Y-axis is the same for both (a) and (b).

### *3.3.3 CD Spectra of Scrambled Sequences*

Peptoids reported in the literature to form stable helices generally contain a regular pattern of helix inducing residues in the primary sequence. The repeat motif peptoid CD spectra demonstrate that the proportion of *N*spe in the primary sequence affects the secondary structure but the positioning of these residues has also been reported to be influential in sequences containing a mixture of chiral and non-chiral residues [32]. In addition, the positioning of charged residues such as lysine and histidine has been shown to influence the overall helical character of peptide sequences [58] so the distribution of *N*ae and *N*Lys may be of consequence in the peptoid sequences. To investigate this we collected CD spectra for the scrambled sequence peptoids, with both charged side chain variants, in PBS at 20°C. We have subdivided the scrambled sequences into two groups based on the number of adjacent *N*spe residues in each sequence. The spectra for the first subset of sequences, which contain a maximum of two adjacent *N*spe, are shown in Figure 3.19a. The spectra for the second subset of sequences, each of which contain up to 4 adjacent *N*spe are shown in Figure 3.19b.

Each of the scrambled peptoids has a spectrum that indicates helical secondary structure. The positions and intensities of the spectral minima are listed in Table 3.4. Trends observed in the CD spectra of the repeat motif peptoids were also observed in the scrambled sequences. There is very little variation in the positions of the minima across the scrambled sequences with both charged side chains. However, there is considerable variation in spectral intensity. Contrary to expectation, the repeat motif variants (S1 and S8)

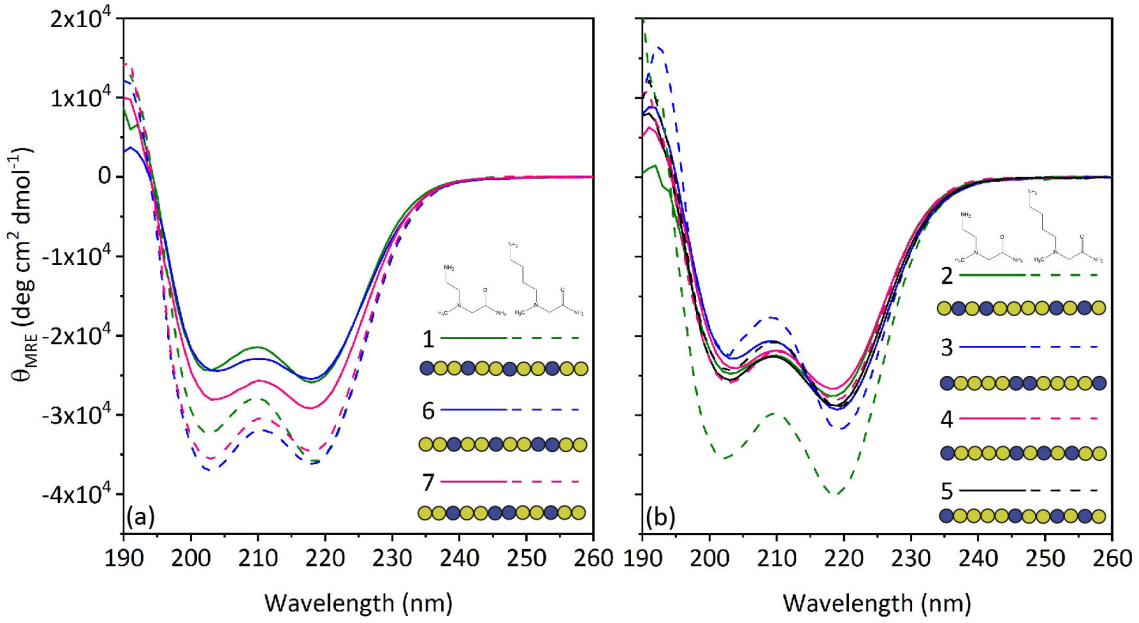
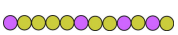
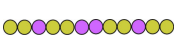
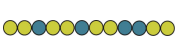
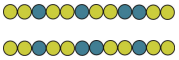



*Fig. 3.19:* CD spectra of scrambled sequences peptoids in 0.01 M PBS. (a) Shows the spectra for sequences 1, 6 and 7 and (b) shows the spectra of sequences 2, 3, 4 and 5. The *N*ae side chain sequences are shown as solid lines and the *N*ae side chain sequences as dashed lines in the same colours. Peptoid concentration is 25 $\mu$M in each case.

were not the most helical of the library by the measure of the spectral intensity at $\lambda_2$ or the ratio of the minima, $R_{2/1}$. Substituting *N*Lys for *N*ae results in a more helical peptoid for every sequence variation. This was particularly clear in the case of the sequences with up to 2 adjacent *N*spe residues (Figure 3.19a) where all of the *N*Lys sequences have a considerably greater signal intensity that any of the *N*ae sequences. There is a less obvious distinction between the *N*Lys and *N*ae variants of the sequences containing up to 4 adjacent *N*spe residues, although in each case the *N*Lys variant does have a greater overall helicity, as measured by the intensity of $\lambda_2$. There was not a clear variation in spectral intensity between the two subgroups, indicating that increasing the number of adjacent *N*spe residues without regard to the specific positioning along the chain does not moderate the overall helical content or conformation where measured by spectral intensity. This is contrary to the findings of Shin *et al.* who propose adjacent *N*spe residues synergistically enhance the helical character of sequences in acetonitrile [32]. However, the Shin *et al.* study only considered heptamers with 27% *N*spe (2 residues) whereas the scrambled sequences discussed here have 67% *N*spe and therefore the synergistic effect of the adjacent residues may be diminished in sequences with a higher overall percentage of helix inducing residues.

There is some variation in the ratio of the two minima and hence the spectral shape of the scrambled sequences. $R_{2/1}$ for the scrambled sequence are depicted in Figure 3.20a where the magenta bars represent the *N*ae variants and the teal bars the *N*Lys variants.

| Side Chain | Sequence | $\lambda_{min1}$ (nm) | $\lambda_{min2}$ (nm) | $\theta_{MRE,min1}$ | $\theta_{MRE,min2}$ |
|---|---|---|---|---|---|
| *N*ae | S1  | 202.8 | 218.7 | -25877 | -27998 |
| | S2  | 203.0 | 218.3 | -24816 | -27683 |
| | S3  | 203.3 | 219.1 | -22942 | -29395 |
| | S4  | 203.7 | 218.6 | -24113 | -26674 |
| | S5  | 203.1 | 218.8 | -25664 | -28854 |
| | S6  | 203.7 | 218.0 | -24494 | -25430 |
| | S7  | 203.5 | 217.6 | -28151 | -29191 |
| *N*Lys | S8  | 202.3 | 218.7 | -32320 | -36100 |
| | S9  | 202.1 | 218.9 | -35647 | -40175 |
| | S10  | 202.8 | 219.4 | -22382 | -32026 |
| | S11  | 202.8 | 218.9 | -26140 | -28245 |
| | S12  | 202.8 | 219.2 | -24329 | -29015 |
| | S13  | 202.8 | 218.2 | -37006 | -36236 |
| | S14  | 202.9 | 218.2 | -35648 | -34742 |

*Tab. 3.4:* Wavelengths of the 2 minima and corresponding MRE values (in units of deg cm$^2$ dmol$^{-1}$) for scrambled sequences 1-14. In the graphic representations of the sequences yellow circles represent *N*spe residues, pink circles represent *N*ae residues and teal circles represent *N*Lys residues.

Neither charged side chain consistently results in a greater value of $R_{2/1}$ than the other in the equivalent sequence. In 3 out of the 7 sequences the *N*ae variant has a higher value of $R_{2/1}$, while in the other 4 $R_{2/1}$ is greater for the *N*Lys variant. Overall $R_{2/1}$ varies to a greater extent with the sequence pattern than the structure of the charged side chain, as was also the case for the repeat motif sequences. Sequence 3 in particular stands out as it has a particularly high value of $R_{2/1}$ relative to the rest of the scrambled and repeat motif sequences. This sequence is unique among the library for having a string of 4 adjacent *N*spe residues close in the C-terminal region of the sequence, though it does not have one in the C-terminal position itself. In addition, it is the only sequence with 2 sets of 4 adjacent *N*spe. Potentially this high density of stabilising aromatic residues could help to stabilise the otherwise flexible C-terminal end of the sequence, resulting in the enhanced helical character of the CD spectrum relative to the others in the library. Alternatively this higher ratio could indicate that this sequence is not monomeric but instead forms helical bundles, though there is no other evidence that this is the case.

The scrambled sequences with *N*Lys residues vary more in both $R_{2/1}$ and overall intensity than their *N*ae counterparts. This indicates that the positioning of *N*Lys residues in different places among the *N*spe causes a greater disruption to the helical structure (or on the other hand promotes folding to a greater extent) than the equivalent positioning of *N*ae. The standard deviation in MRE at $\lambda_2$ is approximately 1335 deg cm$^2$ dmol$^{-1}$ for the *N*ae sequences and approximately 4275 deg cm$^2$ dmol$^{-1}$ for the *N*Lys sequences. The standard deviation in $R_{2/1}$ was 0.08 for the *N*ae sequences and 0.16 for the *N*Lys

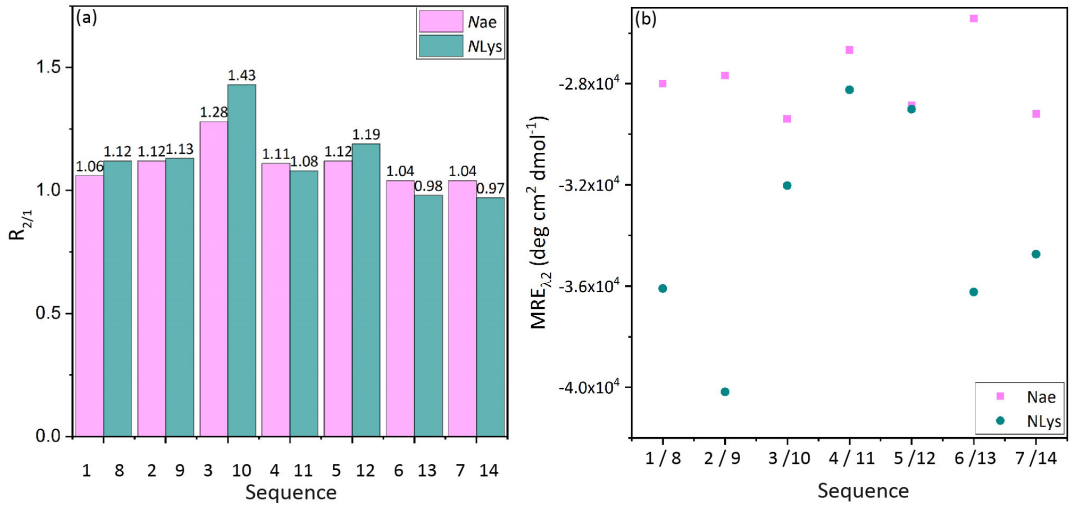



*Fig. 3.20:* (a) Ratio of mean residue ellipticity at $\lambda_2$ to mean residue ellipticity at $\lambda_1$ ($R_{2/1}$) for the scrambled sequence library with data for *N*ae side chains shown in magenta and *N*Lys side chains in teal. (b) MRE at $\lambda_2$ for each scrambled sequence with *N*ae variants shown in magenta and *N*Lys in teal.

sequences.

It is possible that the increased length of the *N*Lys side chain relative to *N*ae allows the peptoids containing the former greater flexibility to minimise electrostatic repulsions between the charged groups and this in turn alters the backbone conformation. The *N*ae side chains, with less flexibility to rearrange, could induce strain in the peptoid backbone as the entropic penalty associated with rearranging the side chains to minimise electrostatic repulsions would be higher.

The particular positioning of the charged residues may also affect the scrambled sequence spectra. Rules relating to charges in end positions stabilising helices in peptides may be transferable to peptoids as they do not concern the hydrogen bonding network but rather the helix macrodipole. The C=O dipoles in the peptoid helix described by Armand *et al.* [5] are almost parallel to the helical axis. This results in a helix macrodipole, similar but opposite in direction to macrodipoles observed in peptide $\alpha$-helices, since the C=O point towards the N-terminus in the peptoid and the C-terminus in the peptide. The knowledge that charged amino acids near the C-terminus stabilise helices in peptides led Armand *et al.* to predict that charged residues near the N-terminus may stabilise peptoid helices. It could be argued that this is what we observe in the scrambled sequences as 1, 3, 4 and 5, which all have a charged residue at the N terminus, appear to have more intensely helical CD spectra than the others in the library.

Overall it is very difficult to de-convolute the effects that different factors may have on the CD spectra for the scrambled sequences. Switching the charged side chain in otherwise equivalent sequences has the most distinctive effect on the helicity, with the longer *N*Lys resulting in more helical sequences that the shorter *N*ae. The effect of the particular

placement of residues is more difficult to identify and quantify as the positioning of both the aromatic *N*spe residues and the charged residues could influence the peptoid secondary structure. In addition the total helical content could be crudely quantified by either the MRE at $\lambda_2$ or $R_{2/1}$ but different scrambled sequences appear more helical than others depending on which of these is chosen.

### *3.3.4 Analysis of Helical Hydrophobic Moment*

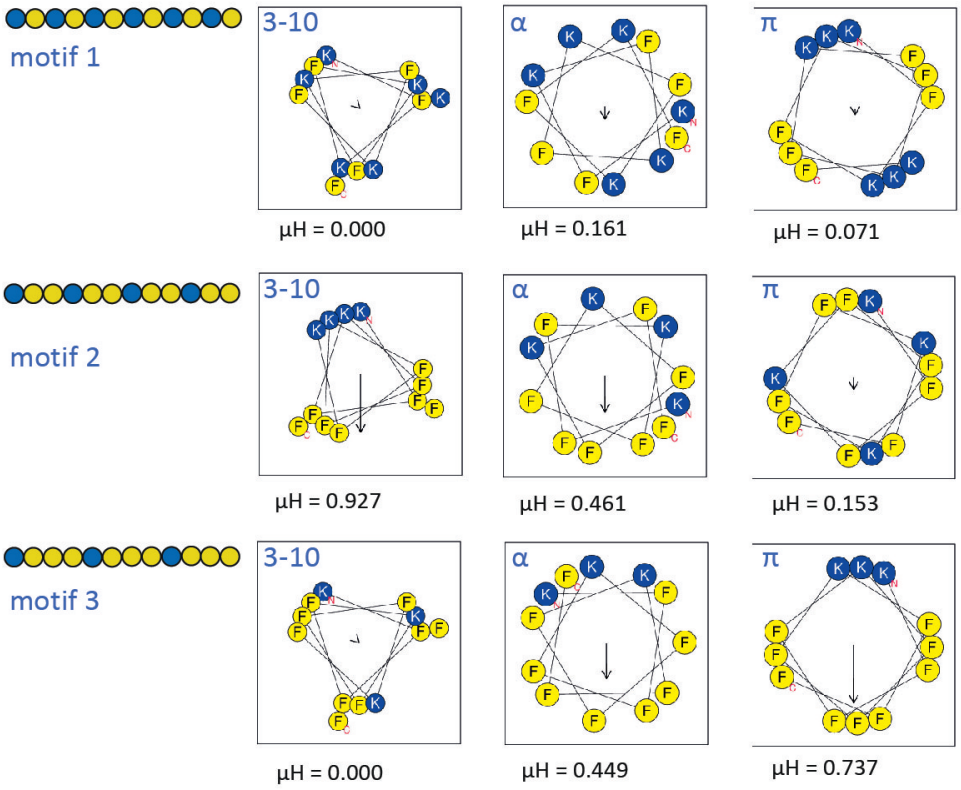



*Fig. 3.21:* Helical wheel diagrams for the repeat motif sequences generated according to model backbone structures corresponding to 3-10, $\alpha$ and $\pi$ helices. For each sequence and model backbone the helical hydrophobic moment, $\mu$H, is displayed beneath the wheel diagram. Diagrams were generated and $\mu$H was calculated using the HeliQuest web server [59]. Each helix has been rotated so that the $\mu$H vectors are aligned vertically. Charged residues (*N*ae/*N*Lys) are represented by blue circles and aromatic residues (*N*spe) by yellow circles. The C and N termini are marked in red letters.

For each of our peptoid sequences we used the Heliquest server [59] to generate helical wheel diagrams to visualise the amphipathicity and charge distribution around the helical axis, additionally calculating the mean helical hydrophobic moment, $\mu$H. $\mu$H is used to quantify the amphiphilicity of peptide helices and is the vector sum of the hydrophobicities of each of the individual side chains in a peptide (or peptoid) sequence, arranged in a particular helical conformation. $\mu$H has been found to correlate to peptide helicity (mea-

sured by CD) and biological activity for certain AMPs [60]. We therefore investigated the possibility of a correlation between $\mu$H and helicity in our peptoid library. In particular we were interested in trying to gain insight into the spectral variations observed between the different scrambled sequences as it is difficult to correlate the primary sequences with the changes in CD spectra. A large value of $\mu$H (close to 1) indicates that the helix is

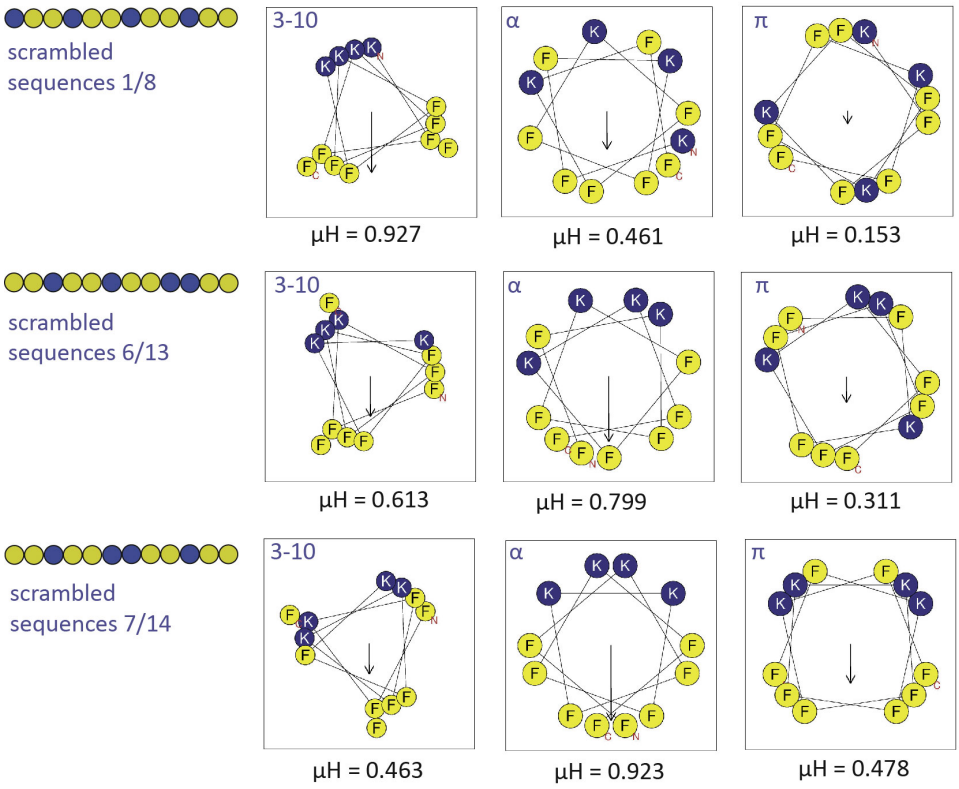



*Fig. 3.22:* Helical wheel diagrams for the scrambled sequences containing a maximum of 2 adjacent *N*spe residues, generated according to model backbone structures corresponding to $3-10$, $\alpha$ and $\pi$ helices. For each sequence and model backbone the helical hydrophobic moment, $\mu$H, is displayed beneath the wheel diagram. Diagrams were generated and $\mu$H was calculated using the HeliQuest web server [59]. Each helix has been rotated so that the $\mu$H vectors are aligned vertically. Charged residues (*N*ae/*N*Lys) are represented by blue circles and aromatic residues (*N*spe) by yellow circles. The C and N termini are marked in red letters.

amphipathic perpendicular to its helical axis, with values greater than 0.6 generally indicating a highly amphipathic helix. To our knowledge there is no published data concerning the hydrophobicities of individual peptoid side chains, including those within our peptoid library, which somewhat limits this analysis. Nevertheless we can estimate $\mu$H for different peptoid sequences by using the hydrophobicity of phenylalanine as an approximation to *N*spe and lysine as an approximation to *N*Lys and *N*ae. This allows us to distinguish between $\mu$H for approximations to the different primary sequence variations in our library, but not between the two different charged side chain variants. Heliquest performs calculations using the provided sequence of amino acids and one of three standard helical

conformations: the $\alpha$, 3-10 or $\pi$ helix. The likelihood that the peptoids adopt a structure different to those found in peptides (and available in Heliquest) is an additional limiting factor on the accuracy of this analysis. Figure 3.21 shows the helical wheel diagrams and $\mu$H values for the repeat motif peptoid sequences in each of the three aforementioned helical configurations.

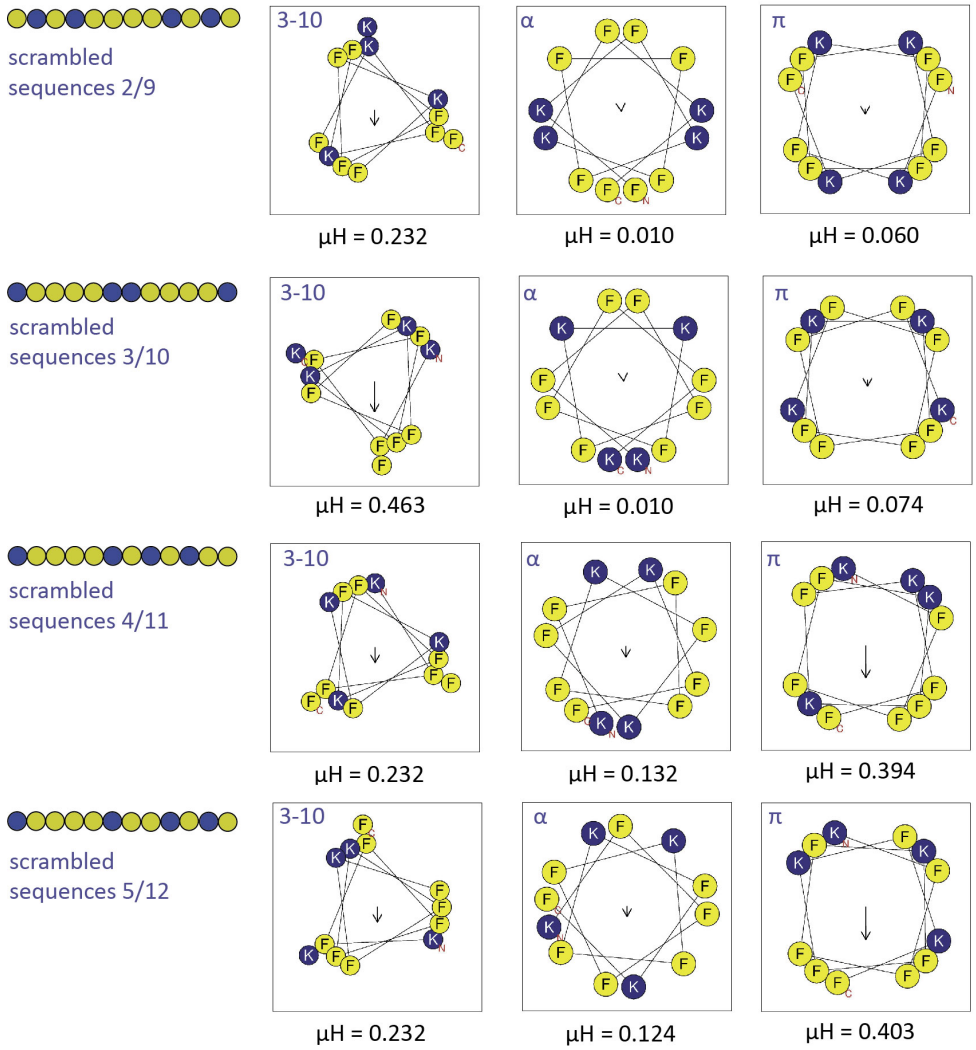



*Fig. 3.23:* Helical wheel diagrams for the scrambled sequences containing at least of 4 adjacent *N*spe residues, generated according to model backbone structures corresponding to $3-10$, $\alpha$ and $\pi$ helices. For each sequence and model backbone the helical hydrophobic moment, $\mu$H, is displayed beneath the wheel diagram. Diagrams were generated and $\mu$H was calculated using the HeliQuest web server [59]. Each helix has been rotated so that the $\mu$H vectors are aligned vertically. Charged residues (*N*ae/*N*Lys equivalent) are represented as blue circles and aromatic residues (*N*spe equivalent) as yellow circles. The C and N termini are marked in red letters.

The helical wheel diagrams for the repeat motif sequences reveal that the motif 2 and motif 3 peptoids have the potential to form highly amphipathic, monomeric helices and the motif 1 peptoids do not. The motif 1 peptoids have a $\mu$H value no higher than 0.161

in any of three helical conformations tested, with the $\alpha$ conformation being the most amphipathic for this sequence and the 3-10 conformation being the least with $\mu$H = 0. The motif 2 peptoids would be highly amphipathic if they were to adopt a 3-10 conformation and moderately amphipathic in an $\alpha$ conformation. The motif 3 peptoids would also be moderately amphipathic in an $\alpha$ conformation and would be highly amphipathic in a $\pi$ conformation. This can be understood intuitively given the relative number of residues in a single motif and the number of residues per turn in each helical structure. It should be noted however that none of the peptide helical configurations could be a good representation of the true structures adopted by the peptoids and therefore this analysis is merely to provide a speculative representation of how the charged and aromatic residues may be distributed around a helical axis in each peptoid sequence. There does not appear to be any correlation between $\mu$H and spectral intensity for these sequences.

Helical wheel diagrams and $\mu$H values for the scrambled sequences are shown in Figures 3.22 (for sequences with maximum 2 adjacent *N*spe) and 3.23 (for sequences containing up to 4 adjacent *N*spe). In general the sequences containing larger numbers of adjacent *N*spe residues result in less amphipathic structures over all three helical templates. This makes sense given that each of sequences 2,3,4 and 5 contain at least 4 adjacent *N*spe residues and therefore, as none of the helices have more than 4 residues per turn, there is guaranteed to be a distribution of the hydrophobic residues all around the helical axis. This is not the case for sequences 1,6 and 7 which only contain a maximum of 2 adjacent *N*spe residues and these sequences result in moderately and highly amphipathic structures across all the helical templates.

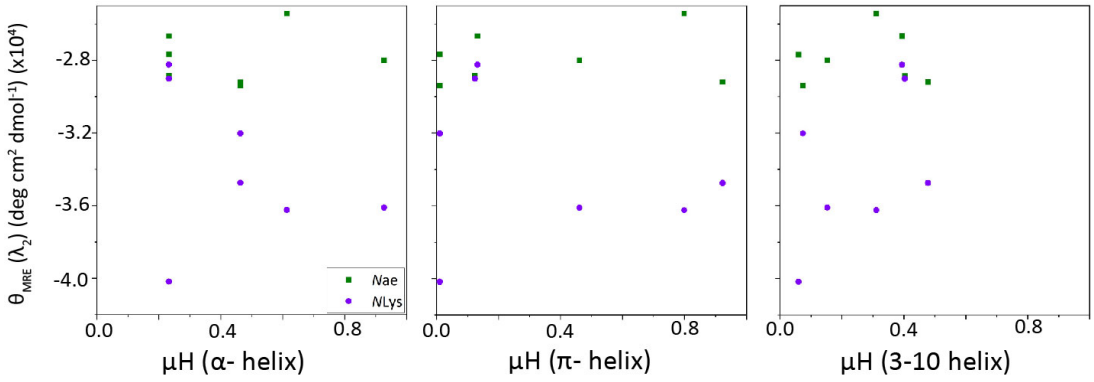



*Fig. 3.24:* Helical hydrophobic moment ($\mu$H) for each scrambled sequence as calculated by Heliquest, based on standard 3-10, $\alpha$ and $\pi$ -helical backbones, plotted against the MRE values for each of the *N*ae (green) and *N*Lys (purple) variants at $\lambda_2$ in PBS.

Figure 3.24 shows the hydrophobic moment ($\mu$H) for each scrambled sequence plotted against the intensity at $\lambda_2$ in the CD spectrum. There does not appear to be any correlation between $\mu$H and the spectral intensity for any of the standard peptide helical backbone conformations. This is not surprising and further indicates the the peptoids adopt helical

conformations distinctly different to those observed in peptides.

### 3.3.5 *Secondary Structural Response to Temperature*

Thermally induced unfolding is often observed in peptides via changes in CD spectral shape over a range of temperatures. In the simplest case the structural transition observed is from a folded state at low temperatures to an unfolded state at higher temperatures. The midpoint of the transition represents the melting temperature of the peptoid. The temperature dependent CD spectra for a given peptide (or peptoid) can be used to calculate thermodynamic properties including the free energy of unfolding. The melt curves for peptides that show cooperative thermal unfolding are characteristically sigmoidal in shape [61].

Peptoids have been widely reported to exhibit remarkable thermal stability, maintaining their folded conformations at much higher temperatures than many antimicrobial peptides [2, 36, 62]. We investigated the thermal unfolding of the repeat motif and scrambled sequence peptoids in PBS by measuring ellipticity at the wavelength of the second spectral minimum ($\lambda_2$) as a function of temperature, between 20 °C and 90 °C. Additionally, full spectra were collected at 10 °C intervals to give an indication of how the spectral shape varied with temperature.

The melt curves for the repeat motif peptoids are shown in Figure 3.25a. The change in intensity of $\lambda_2$ between 20 °C and 90 °C is small, providing very little evidence of thermally induced unfolding of any of the sequences. Each peptoid shows a slight decrease in signal intensity at $\lambda_2$ with increasing temperature. The total loss of spectral intensity was in the region of 10-15% for each sequence over the temperature range studied.

During the course of the peptoid melts there was very little change in spectral shape, which can be seen in the spectra collected at 10 °C intervals. The melt spectra for RM3 are shown in Figure 3.25b as an example of this behaviour and are generally representative of the behaviour of all the sequences. From the full spectra it is evident that the positions of the minima gradually shift to longer wavelengths over the course of the melt. This shift is exaggerated in the case of $\lambda_1$ as the loss of spectral resolution occurs at longer wavelengths as temperature increases. For this reason the intensity of $\lambda_2$ was chosen over $\lambda_1$ for monitoring.

The two-state analysis often applied to peptides could not be applied to this data due to the very slow rate of change in the CD spectra over the measured temperature range. Though each peptoid is clearly in a folded state at 20°C there is no convergence of the spectra over certain temperature ranges (within those studied). It is possible that lower temperatures could induce further folding and therefore we do not know what the spectra of the maximally folded and minimally folded states would look like and hence

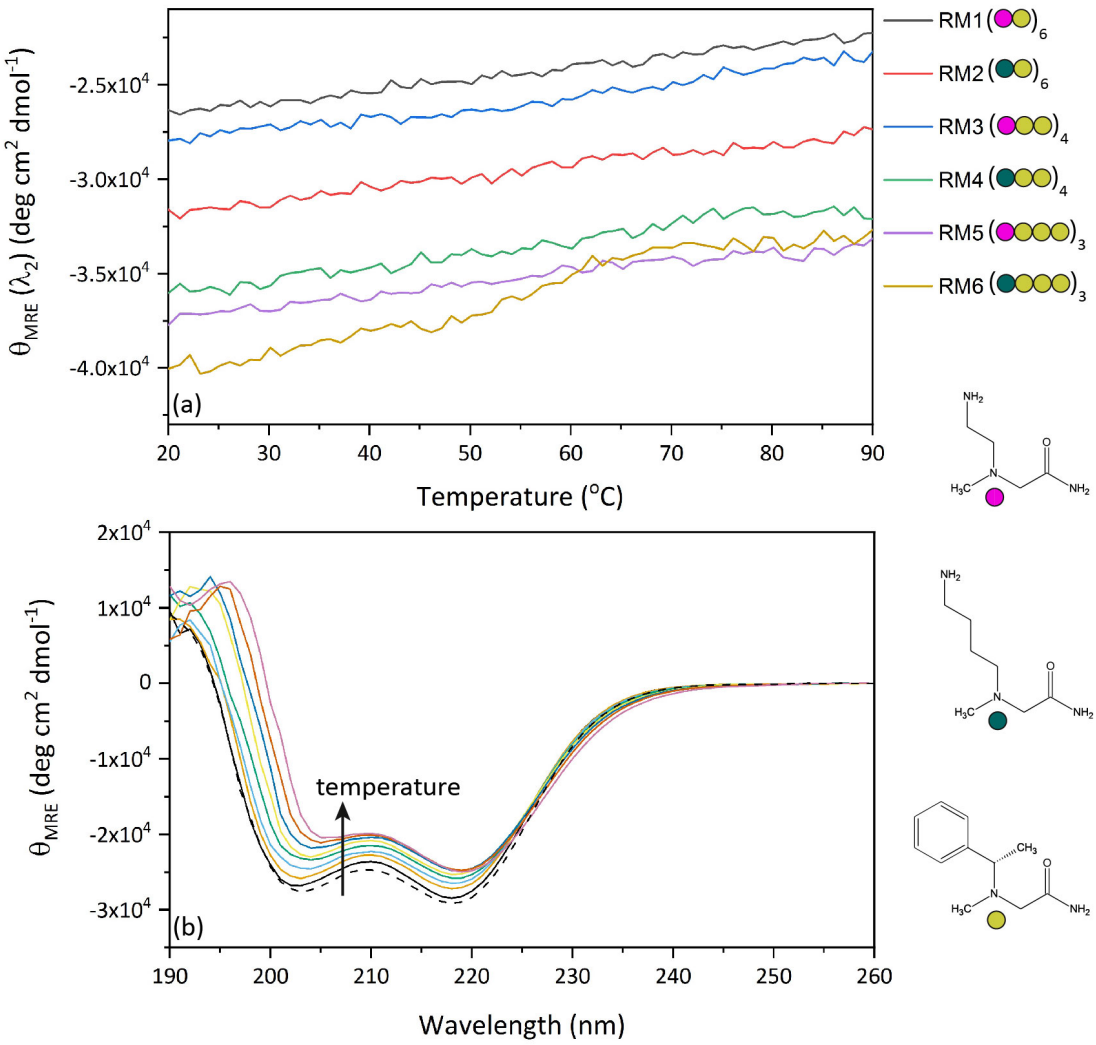



*Fig. 3.25:* (a) Melt curves for peptoids RM1-6 in 0.01 M PBS. Peptoid concentration is 25 $\mu$M in each case. (b) Full spectra for RM3 in PBS collected every 10 °C over the course of a melt, from 20 °C (black line) to 90 °C (purple line). On completion of the melt the sample temperature was returned to 20 °C and an addition spectrum collected which is shown as the black dashed line.

cannot meaningfully fit a two state model two the data. Given the many reports in the literature asserting that peptoid secondary structures are extremely robust, particularly in response to temperature, the melt curves obtained for the repeat motif sequences were not unexpected. Given however, that the sequences reported in the literature all contain a regular repeating motif assumed to be important in stabilising the secondary structure, it is interesting to consider how the scrambling of sequences affects this and whether disruption to the regular patterning of *N*spe residues is influential.

The melt curves for the scrambled sequences in our peptoid library are shown in Figure 3.26a and 3.26b with the library again divided into two by the number of adjacent *N*spe residues in each sequence. Like the repeat motifs, the scrambled sequences showed very little change in spectral shape and intensity over the range of temperatures studied. The

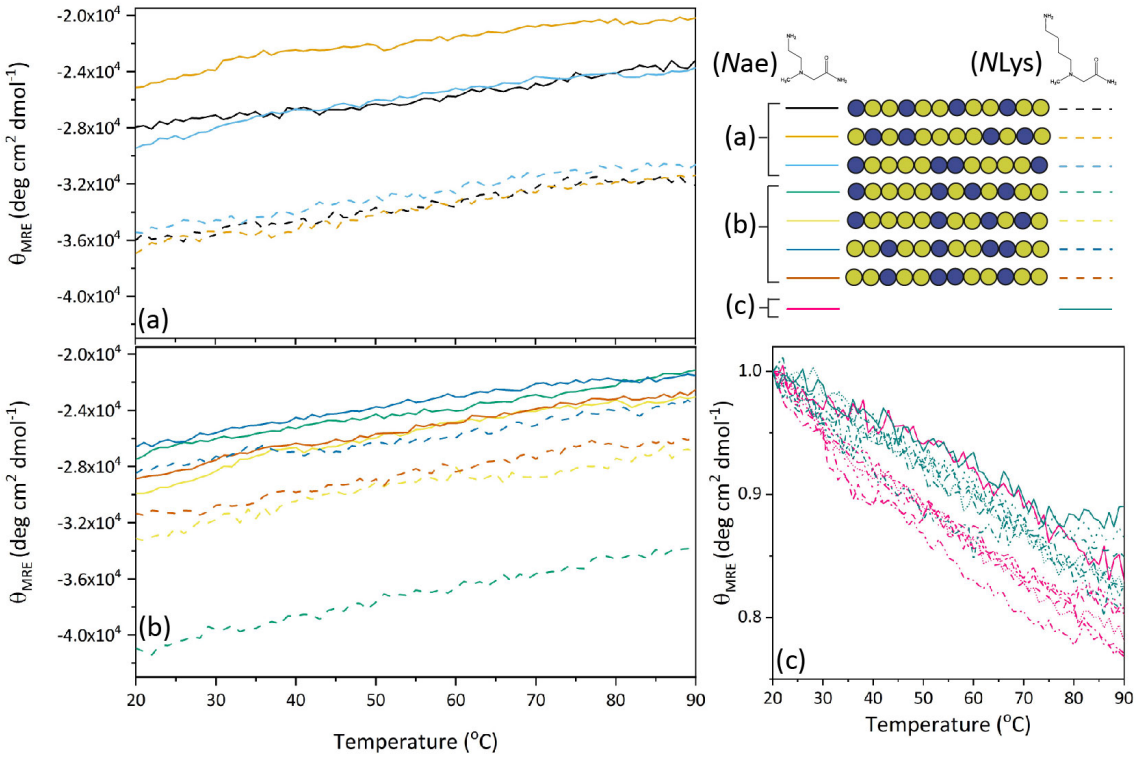



*Fig. 3.26:* (a) Melt curves for the scrambled sequences with a maximum of 2 adjacent *N*spe residues in 0.01 M PBS. (b) Melt curves for the scrambled sequences with a maximum of 4 adjacent *N*spe residues in 0.01 M PBS. Peptoid concentration is 25 $\mu$M in each case. (c) Melt curves for every scrambled sequence with data normalised to the mean residue ellipticity at 20 °C. The sequences with *N*ae charged side chains are shown in magenta and the sequences with *N*Lys charged side chains are shown in teal with sequences 1 and 8 (the repeat motif 2 sequences) shown as solid lines and the scrambled variants shown in dotted and dashed lines.

melt curves normalised to the spectral intensity at 20 °C are shown in Figure 3.26c. The sequences with the *N*ae charged side chain are shown in pink and the *N*Lys charged side chain shown in teal. Peptoids S1 and S8 (which are the repeat motif sequences RM3 anf RM4) are shown as solid lines while the scrambled versions are dotted and dashed. The normalised melt curves reveal that, overall, the sequences with the *N*ae charged side chains show a greater loss of spectral intensity over the course of the melts than the sequences with the *N*Lys charged side chains. The exception to this is S1, the repeat motif *N*ae sequence. The repeat motif *N*Lys sequence also shows the smallest overall loss of intensity of any of the sequences, indicating that for all the sequences containing 4 cationic and 8 aromatic residues, the two containing the repeat motif were the least sensitive to temperature change within the range investigated.

It seems intuitive that RM3 and RM4 exhibit the least thermal unfolding behaviour of the scrambled sequences. These have the most regular distribution of charge along the backbone and therefore also the most regular distribution of hydrophobic residues. The sequences with a maximum of 4 adjacent *N*spe residues showed a greater loss of intensity

at $\lambda_2$ over the course of the melt than those with a maximum of 2 adjacent *N*spe residues. The sequences with *N*ae side chains and a maximum of 2 adjacent *N*spe residues showed a 17-20% loss of intensity where their *N*Lys counterparts had a 10-15% loss of intensity. The sequences with *N*ae side chains a maximum of 4 adjacent *N*spe residues showed a 19-23% loss of intensity and their *N*Lys counterparts had a 17-19% loss of intensity. It therefore appears that a regular distribution of charge along the peptoid stabilises the sequences to a small extent at increasing temperatures.

## 3.4 Conclusions

### *3.4.1 Summary*

In this chapter we set out to investigate the relationship between the primary sequence and secondary structure of a small library of linear peptoids. The key conclusions drawn from our investigation are summarized in the following bullet points:

- The repeat motif and scrambled peptoid sequences have CD spectra that indicate that they adopt monomeric, helical, folded conformations in PBS.
- The handedness of the peptoid helix is determined by the handedness of the chiral side chains.
- The peptoids maintain a folded conformation over a wide range of temperatures (20°C - 90°C).
- One sequence, RM5, self-associates in PBS in a non-specific manner that is sensitive to both temperature and concentration.
- The subset of the library with *N*Lys side chains are more varied in spectral shape and intensity than the equivalent subset of sequences with *N*ae side chains.
- Each sequence with *N*Lys side chains is more helical than its *N*ae side chain counterpart.
- The spectral characteristics of the repeat motif sequences differ in octanol and PBS, indicating that they adopt different structures in aqueous and organic environments. Features associated with helical structure are exaggerated in the octanol spectra relative to PBS.

### *3.4.2 Outlook*

The characteristic features associated with $\alpha$-helical structure in peptides are present in the CD spectra for the peptoids, despite the lack of amide hydrogen prohibiting the

formation of inter-backbone hydrogen bond stabilised structures in peptoids. The spectra we collected also resemble the spectra of pure *N*spe peptoid helices and variations thereof published in the literature [5, 31, 36, 63]. It therefore seems likely that the inclusion of *N*spe residues in the primary sequences that we have studied is responsible for inducing their helical secondary structure. This is clearly demonstrated in the case of substituting *N*spe for its opposite handed counterpart *N*rpe, or the similar but achiral *N*phe. The former inverts the CD spectrum and the latter results in a the total loss of CD signal. Additionally, within the repeat motif sequences the total percentage of chiral *N*spe residues varies and the CD spectra increase in intensity as the overall proportion of *N*spe increases, indicating that the sequences become more helical. These results demonstrate the power of the choice of chiral monomer in controlling the overall secondary structure, including its handedness.

While the increase in helical character with increasing proportion of *N*spe was perhaps predictable, less so was the influence of the structure of the charged side chain group. Peptoids with *N*ae side chains appear to be less helical than their identical sequence counterparts with the longer *N*Lys side chains. This was observed to be the case, without exception, for every sequence explored within this chapter. The physical reasons for this cannot be inferred from CD spectra alone, however it could result from the increased length and therefore flexibility of *N*Lys relative to *N*ae allowing sequences to rearrange, minimising electrostatic repulsions and allowing the peptoid backbone to adopt a folded conformation at lesser cost.

In addition, across the library as a whole, the *N*Lys sequences were considerably more varied in spectral intensity and shape than their *N*ae counterparts. This indicates that the former side chain is a better choice for the development of combinatorial libraries with conformational diversity than the latter. It would be interesting to investigate the structure of additional peptoids with the same primary sequences as those studied here but with charged side chains of varying length. The peptoid analogue of ornithine would be a good starting point for such an investigation as it is intermediate in length between *N*ae and *N*Lys but otherwise structurally analogous. Further investigations into the effect of altering the length of the charged side chain group in these sequences may also provide insight into the self-association which is observed in $(N\text{ae}N\text{spe}N\text{spe})_3$ but not $(N\text{Lys}N\text{spe}N\text{spe})_3$.

The conformational diversity within the scrambled sequence spectra is very difficult to correlate to any particular property of the the primary sequences. It seems likely that the positioning of both the charged and the aromatic residues affect the overall secondary structure and therefore it is difficult to de-convolute the effect of each of these on any individual spectrum. Further investigation into a new library of sequences with systematic substitution of residues in particular positions may help rationalise our results here in the

future. Nevertheless, we have clearly demonstrated that sequences that do not contain a standard repeating motif of residues do still adopt stable helical conformations in aqueous solution. These structures are thermally robust over a wide range of temperatures, as are the repeat motif sequences. This is a strong indication that the helices formed in peptoids are stabilised by non-covalent interactions.

The spectral intensity for the peptoids is increased in octanol relative to PBS, with the intensity of the second minima greatly enhanced relative to the first. This indicates that the peptoids adopt a distinctly different structural conformations in the two solvent conditions, evident in the change in overall spectral shape. This provides the basis for the work in the following chapters, where we investigate these structural differences computationally and then use them to examine the nature of interactions between peptoids and model biological membranes, via CD spectroscopy.

# References: Chapter 3


[1] R. N. Zuckermann, "Peptoid origins.", *Biopolymers* **2011**, *96*, 545–55.

[2] T. J. Sanborn, C. W. Wu, R. N. Zuckermann, A. E. Barron, "Extreme stability of helices formed by water-soluble poly-N-substituted glycines (polypeptoids) with alpha-chiral side chains", *Biopolymers* **2002**, *63*, 12–20.

[3] J. R. Stringer, J. A. Crapster, I. A. Guzei, H. E. Blackwell, "Extraordinarily Robust Polyproline Type I Peptoid Helices Generated via the Incorporation of $\alpha$-Chiral Aromatic N-1- Naphthylethyl Side Chains", *Journal of the American Chemical Society* **2012**, *100*, 130–134.

[4] C. W. Wu, K. Kirshenbaum, T. J. Sanborn, J. a. Patch, K. Huang, K. a. Dill, R. N. Zuckermann, A. E. Barron, "Structural and spectroscopic studies of peptoid oligomers with alpha-chiral aliphatic side chains.", *Journal of the American Chemical Society* **2003**, *125*, 13525–13530.

[5] P. Armand, K. Kirshenbaum, A. Falicov, R. L. Dunbrack, K. A. Dill, R. N. Zuckermann, F. E. Cohen, "Chiral N-substituted glycines can form stable helical conformations", *Folding and Design* **1997**, *2*, 369–375.

[6] R. V. Mannige, T. K. Haxton, C. Proulx, E. J. Robertson, A. Battigelli, G. L. Butterfoss, R. N. Zuckermann, S. Whitelam, "Peptoid nanosheets exhibit a new secondary-structure motif", *Nature* **2015**, *526*, 415–420.

[7] S. A. Fowler, H. E. Blackwell, "Structure-function relationships in peptoids: recent advances toward deciphering the structural requirements for biological function.", *Organic & Biomolecular Chemistry* **2009**, *7*, 1508–1524.

[8] G. N. Ramachandran, C. Ramakrishnan, V. Sasisekharan, "Stereochemistry of Polypeptide Chain Configurations", *Journal of Molecular Biology* **1963**, *7*, 96–99.

[9] ProBuilder Online.

[10] L. Pauling, R. Corey, H. Branson, "The structure of proteins: Two hydrogen bonded configurations of the polypeptide chain", *Proceedings of the National Academy of Sciences of the United States of America* **1951**, *37*, 205–211.

[11] B. Haimov, S. Srebnik, "A closer look into the $\alpha$ -helix basin", *Scientific Reports* **2016**, *6*, 1–12.

[12] R. S. Harrison, N. E. Shepherd, H. N. Hoang, G. Ruiz-gómez, T. A. Hill, R. W. Driver, V. S. Desai, P. R. Young, G. Abbenante, D. P. Fairlie, "Downsizing human , bacterial , and viral proteins to short water-stable alpha helices that maintain biological potency", *Proceedings of the National Academy of Sciences of the United States of America* **2010**, 1–6.

[13] M. Wolny, M. Batchelor, G. J. Bartlett, E. G. Baker, M. Kurzawa, P. J. Knight, L. Dougan, D. N. Woolfson, E. Paci, M. Peckham, "Characterization of long and stable de novo single alpha-helix domains provides novel insight into their stability", *Scientific Reports* **2017**, 1–14.

[14] B. H. Zimm, J. K. Bragg, "Theory of the Phase Transition between Helix and Random Coil in Polypeptide", *Journal of Chemical Physics* **1959**, *31*, 526–535.

[15] R. Aurora, T. P. Creamer, G. D. Rose, "Local Interactions in Protein Folding : Lessons from the alpha-Helix", *Journal of Biological Chemistry* **1997**, *272*, 1413–1416.

[16] Y. Levy, J. Jortner, O. M. Becker, "Solvent effects on the energy landscapes and folding kinetics of polyalanine", *Proceedings of the National Academy of Sciences* **2001**, *98*, 2188–2193.

[17] J. D. Dunitz, "Pauling's Left-Handed a-Helix", *Angew. Chem. Int. Ed.* **2001**, *22*, 4167–4173.

[18] M. E. Karpen, P. L. de Haseth, K. E. Neet, "Differences in the amino acid distributions of 310 -helixes and alpha- helixes", *Protein Science* **1992**, *1*, 1333–1342.

[19] R. B. Baybutt, "The $\pi$ helix-a hydrogen bonded configuration of the polypeptide chain", *Journal of the American Chemical Society* **1952**, *74*, 5806–5807.

[20] R. B. Cooleya, D. J. Arp, P. A. Karplusa, "Evolutionary origin of a secondary structure: $\pi$-helices as cryptic but widespread insertional variations of $\alpha$-helices enhancing protein functionality", *Journal of Molecular Biology* **2010**, *404*, 232–246.

[21] M. Moradi, V. Babin, C. Roland, C. Sagui, "A classical molecular dynamics investigation of the free energy and structure of short polyproline conformers", *Journal of Chemical Physics* **2010**, *133*.

[22] M. Moradi, V. Babin, C. Roland, T. A. Darden, C. Sagui, "Conformations and free energy landscapes of polyproline peptides", *Proceedings of the National Academy of Sciences of the United States of America* **2009**, *106*, 20746–20751.

[23] B. M. Bulheller, A. Rodger, J. D. Hirst, "Circular and linear dichroism of proteins.", *Physical Chemistry Chemical Physics* **2007**, *9*, 2020–2035.

[24] C. Toniolo, A. Polese, F. Formaggio, M. Crisma, J. Kamphuis, " Circular Dichroism Spectrum of a Peptide 3 10 -Helix ", *Journal of the American Chemical Society* **2002**, *118*, 2744–2745.

[25] C. Tonlolo, E. Benedetti, "The polypeptide 310-helix", *Trends in Biochemical Sciences* **1991**, *16*, 350–353.

[26] N. H. Andersen, Z. Liu, K. S. Prickett, "Efforts toward deriving the CD spectrum of a 310 helix in aqueous medium", *FEBS Letters* **1996**, *399*, 47–52.

[27] P. Armand, K. Kirshenbaum, R. A. Goldsmith, S. Farr-Jones, A. Barron, K. T. V. Truong, K. A. Dill, D. F. Mierke, F. E. Cohen, R. N. Zuckermann, E. K. Bradley, "NMR determination of the major solution conformation of a peptoid pentamer with chiral side chains", *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 4309–4314.

[28] K. Kirshenbaum, A. E. Barron, R. A. Goldsmith, P. Armand, E. K. Bradley, K. T. V. Truong, K. A. Dill, F. E. Cohen, R. N. Zuckermann, "Sequence-specific polypeptoids: a diverse family of heteropolymers with stable secondary structure", *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 4303–4308.

[29] N. P. Chongsiriwatana, J. A. Patch, A. M. Czyzewski, M. T. Dohm, A. Ivankin, D. Gidalevitz, R. N. Zuckermann, A. E. Barron, "Peptoids that mimic the structure, function, and mechanism of helical antimicrobial peptides.", *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105*, 2794–2799.

[30] C. M. Darapaneni, P. J. Kaniraj, G. Maayan, "Water soluble hydrophobic peptoids via a minor backbone modification", *Organic and Biomolecular Chemistry* **2018**, *16*, 1480–1488.

[31] C. W. Wu, T. J. Sanborn, R. N. Zuckermann, A. E. Barron, "Peptoid oligomers with alpha-chiral, aromatic side chains: Effects of chain length on secondary structure", *Journal of the American Chemical Society* **2001**, *123*, 2958–2963.

[32] H. M. Shin, C. M. Kang, M. H. Yoon, J. Seo, "Peptoid helicity modulation: precise control of peptoid secondary structures via position-specific placement of chiral monomers", *Chemical Communications* **2014**, *50*, 4465–4468.

[33] B. C. Gorske, B. L. Bastian, G. D. Geske, H. E. Blackwell, "Local and tunable n$\rightarrow\pi^*$ interactions regulate amide isomerism in the peptoid backbone", *Journal of the American Chemical Society* **2007**, *129*, 8928–8929.

[34] J. R. Stringer, J. A. Crapster, I. A. Guzei, H. E. Blackwell, "Construction of peptoids with all trans -amide backbones and peptoid reverse turns via the tactical incorporation of N -aryl side chains capable of hydrogen bonding", *Journal of Organic Chemistry* **2010**, *75*, 6068–6078.

[35] B. C. Gorske, J. R. Stringer, B. L. Bastian, S. A. Fowler, H. E. Blackwell, "New strategies for the design of folded peptoids revealed by a survey of noncovalent interactions in model systems", *Journal of the American Chemical Society* **2009**, *131*, 16555–16567.

[36] K. Huang, C. W. Wu, T. J. Sanborn, J. A. Patch, K. Kirshenbaum, R. N. Zuckermann, A. E. Barron, I. Radhakrishnan, "A threaded loop conformation adopted by a family of peptoid nonamers", *Journal of the American Chemical Society* **2006**, *128*, 1733–1738.

[37] J. A. Crapster, I. A. Guzei, H. E. Blackwell, "A Peptoid Ribbon Secondary Structure", *AngewandteChemie* **2014**, *52*, 5079–5084.

[38] B. C. Gorske, E. M. Mumford, C. G. Gerrity, I. Ko, "A Peptoid Square Helix via Synergistic Control of Backbone Dihedral Angles", *Journal of the American Chemical Society* **2017**, *139*, 8070–8073.

[39] S. B. Y. Shin, B. Yoo, L. J. Todaro, K. Kirshenbaum, "Cyclic peptoids", *Journal of the American Chemical Society* **2007**, *129*, 3218–3225.

[40] J. K. Pokorski, L. M. Miller Jenkins, H. Feng, S. R. Durell, Y. Bai, D. H. Appella, "Introduction of a triazole amino acid into a peptoid oligomer induces turn formation in aqueous solution", *Organic Letters* **2007**, *9*, 2381–2383.

[41] E. J. Robertson, G. K. Olivier, M. Qian, C. Proulx, R. N. Zuckermann, G. L. Richmond, E. J. Robertson, G. K. Olivier, M. Qian, C. Proulx, R. N. Zuckermann, G. L. Richmond, "Assembly and molecular order of two-dimensional peptoid nanosheets through the oil–water interface", *Proceedings of the National Academy of Sciences* **2014**, *111*, 13284–13289.

[42] E. J. Robertson, A. Battigelli, C. Proulx, R. V. Mannige, T. K. Haxton, L. Yun, S. Whitelam, R. N. Zuckermann, "Design, Synthesis, Assembly, and Engineering of Peptoid Nanosheets", *Accounts of Chemical Research* **2016**, *49*, 379–389.

[43] J. Sun, X. Jiang, R. Lund, K. H. Downing, N. P. Balsara, R. N. Zuckermann, "Self-assembly of crystalline nanotubes from monodisperse amphiphilic diblock copolypeptoid tiles", *Proceedings of the National Academy of Sciences of the United States of America* **2016**, *113*, 3954–3959.

[44] S. B. Vollrath, C. Hu, S. Bräse, K. Kirshenbaum, "Peptoid nanotubes: An oligomer macrocycle that reversibly sequesters water via single-crystal-to-single-crystal transformations", *Chemical Communications* **2013**, *49*, 2317–2319.

[45] Y. Luo, Y. Song, M. Wang, T. Jian, S. Ding, P. Mu, Z. Liao, Q. Shi, X. Cai, H. Jin, D. Du, W. Dong, C. Chen, Y. Lin, "Peptoid Nanotubes: Bioinspired Peptoid Nanotubes for Targeted Tumor Cell Imaging and Chemo-Photodynamic Therapy", *Small* **2019**, *15*.

[46] H. Jin, Y. H. Ding, M. Wang, Y. Song, Z. Liao, C. J. Newcomb, X. Wu, X. Q. Tang, Z. Li, Y. Lin, F. Yan, T. Jian, P. Mu, C. L. Chen, "Designable and dynamic single-walled stiff nanotubes assembled from sequence-defined peptoids", *Nature Communications* **2018**, *9*, 1–11.

[47] S. M. Miller, R. J. Simon, S. Ng, R. N. Zuckermann, J. M. Kerr, W. H. Moos, "Proteolytic studies of homologous peptide and N-substituted glycine peptoid oligomers", *Bioorganic and Medicinal Chemistry Letters* **1994**, *4*, 2657–2662.

[48] H. W. Huang, "Molecular mechanism of antimicrobial peptides: The origin of cooperativity", *Biochimica et Biophysica Acta - Biomembranes* **2006**, *1758*, 1292–1302.

[49] Y. H. Chen, J. T. Yang, H. M. Martinez, "Determination of the Secondary Structures of Proteins by Circular Dichroism and Optical Rotatory Dispersion", *Biochemistry* **1972**, *11*, 4120–4131.

[50] B. Vaz, L. Brunsveld, "Stable helical peptoids via covalent side chain to side chain cyclization", *Organic and Biomolecular Chemistry* **2008**, *6*, 2988–2994.

[51] S. Padmanabhan, E. J. York, J. M. Stewart, R. L. Baldwin, "Helix propensities of basic amino acids increase with the length of the side-chain", *Journal of Molecular Biology* **1996**, *257*, 726–734.

[52] T. P. Creamer, G. D. Rose, "Alpha-helix forming propensities in peptides and proteins", *PROTEINS: Structure Function and Bioinformatics* **1994**, *19*, 85–97.

[53] Y. Chen, B. A. Wallace, "Secondary solvent effects on the circular dichroism spectra of polypeptides in non-aqueous environments: Influence of polarisation effects on the far ultraviolet spectra of alamethicin", *Biophysical Chemistry* **1997**, *65*, 65–74.

[54] B. A. Wallace, J. G. Lees, A. J. W. Orry, A. Lobely, R. W. Janes, "Analyses of circular dichroism spectra of membrane proteins.", *Protein Science* **2003**, *12*, 875–884.

[55] N. E. Zhou, C. M. Kay, R. S. Hodgess, "Synthetic Model Proteins: Positional effects of interchain hydrophobic interactions on stability of two-stranded a-helical coiled-coils", *Journal of Biological Chemistry* **1992**, *4*, 2664–2670.

[56] M. G. Wuo, A. B. Mahon, P. S. Arora, "An Effective Strategy for Stabilizing Minimal Coiled Coil Mimetics", *Journal of the American Chemical Society* **2015**, *137*, 11618–11621.

[57] A. A. Fuller, J. Huber, C. J. Jimenez, K. M. Dowell, S. Hough, A. Ortega, K. N. McComas, J. Kunkel, P. Asuri, "Solution effects on the self-association of a water-soluble peptoid", *Biopolymers* **2018**.

[58] K. M. Armstrong, R. L. Baldwin, "Charged histidine affects $\alpha$-helix stability at all positions in the helix by interacting with the backbone charges", *Proceedings of the National Academy of Sciences of the United States of America* **1993**, *90*, 11337–11340.

[59] R. Gautier, D. Douguet, B. Antonny, G. Drin, "HELIQUEST: A web server to screen sequences with specific $\alpha$-helical properties", *Bioinformatics* **2008**, *24*, 2101–2102.

[60] Y. Chen, M. T. Guarnieri, A. I. Vasil, M. L. Vasil, C. T. Mant, R. S. Hodges, "Role of peptide hydrophobicity in the mechanism of action of $\alpha$-helical antimicrobial peptides", *Antimicrobial Agents and Chemotherapy* **2007**, *51*, 1398–1406.

[61] N. J. Greenfield, "Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions", *Nature Protocols* **2006**, *1*, 2527–2535.

[62] J. Sun, R. N. Zuckermann, "Peptoid polymers: A highly designable bioinspired material", *ACS Nano* **2013**, *7*, 4715–4732.

[63] H. L. Bolt, C. E. Williams, R. V. Brooks, R. N. Zuckermann, S. L. Cobb, E. H. Bromley, "Log D versus HPLC derived hydrophobicity: The development of predictive tools to aid in the rational design of bioactive peptoids", *Biopolymers* **2017**, *108*, 1–7.

# 4.0 Computational Structure Prediction

In this chapter we use molecular dynamics (MD) to study peptoid secondary structure. The results presented here concern the adaptation of the General Amber Force Field for peptoid simulations and subsequent investigation of the structural preferences of the repeat motif sequences in water and octanol. Results are discussed in the context of the experimental data presented previously and the literature on peptoid simulation, which is reviewed at the beginning of this chapter.

## 4.1 Introduction and Review of Literature

Application of computational techniques to the study of peptoids could be extremely useful in understanding relationships between primary sequence and secondary structure. In particular, such methods can reveal molecular level details which are inaccessible by experimental techniques. Since the first attempts at computational peptoid studies, these have been enhancing our understanding of the peptoid folding and stabilisation process by enabling researchers to examine in detail the conformational rearrangements of peptoids in various conditions. A review of the published efforts to elucidate precise peptoid secondary structural characteristics is presented in the following sections. This concerns the application of quantum chemical methods to predict the properties of small model systems and subsequently attempts to accurately capture peptoid behaviour through the use of MD with atomistic force fields.

### *4.1.1 Strategies for Peptoid Simulation*

The computational approaches to peptoid structure prediction have largely followed the same routes as those applied to protein folding. Unlike protein and peptide simulation however, the field of peptoid simulation is still relatively new and there is considerable scope for development [1]. *Ab initio* methods, which are too computationally expensive to apply to large systems, have been carried out on peptoid monomers and fragments [2–4]. The data obtained from these calculations has been utilised in building atomistic models and coarse grained models in turn [1]. Due to the highly involved and time-consuming

nature of developing a force field from scratch, efforts to develop atomistic models for peptoids have so far been limited to adapting existing protein force fields to account for the unique structural preferences of the peptoid backbone [1]. These force fields must be carefully validated to ensure they capture the behaviour of peptoids accurately.

### *Ab Initio Studies of Peptoids*

The first efforts to computationally compare the structures of peptoids and peptides were published in 1996 in a study by Moehle and Hofmann [2]. This was an *ab initio* study in which the sarcosine dipeptoid was compared to its peptide analogue, the alanine dipeptide. The structure of each of these molecules can be seen in Figure 4.6. *Ab initio* simulations can be used to perform highly accurate structural energy minimisation by solving the electronic Schrödinger equation [5]. This is advantageous for peptoid simulation as it is purely quantum mechanical and requires only physical constants rather than specific force field parameters tuned to experimental data, which is sparse for peptoids. This comes at high computational expense so these studies are exclusively limited to small fragments or peptoid monomers.

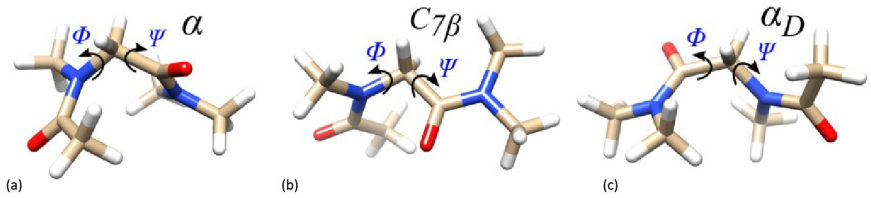



*Fig. 4.1:* Representative structures of sarcosine dipeptoid in (a) $\alpha$ conformation, (b) $C_{7\beta}$ conformation and (c) $\alpha_D$ conformation. Adapted with permission from [6].

In their 1996 study [2] Moehle and Hofmann identified a global minimum energy conformation for the sarcosine dipeptoid in the vacuum-phase. This was termed the $\alpha_D$ region with average dihedral angles ($\phi \sim 75°$, $\psi \sim 170°$) and the peptoid bond in the *cis* conformation ($\omega \sim 0°$). In addition they identified local minima in the region of angles populated by alpha helical peptides ($\phi \sim 55°$, $\psi \sim 55°$), the *trans* $\omega$ $\alpha_D$ and another region unique to peptoids, termed $C_{7\beta}$ ($\phi \sim 130°$, $\psi \sim 78°$). These structures are illustrated in Figure 4.1. This study was the first to demonstrate that both the *cis* and *trans* conformations are equally stable for peptoids in polar solvent conditions, unlike peptides where there is a strong bias towards *trans* conformations. Additionally it was shown that solvation stabilised the *cis* $\alpha_D$ conformation in the sarcosine dipeptoid.

Many of the results of Moehle and Hofmann have been corroborated in further *ab initio* studies of peptoid structure. Baldauf *et al.* identified the same conformational minima in hexamers of sarcosine in their investigation into the foldamer properties of $\alpha$ and $\beta$ peptoids [3]. The results of their calculations, which involved a systematic variation of the backbone dihedrals followed by *ab initio* geometry optimisation at the Hartree-Fock level

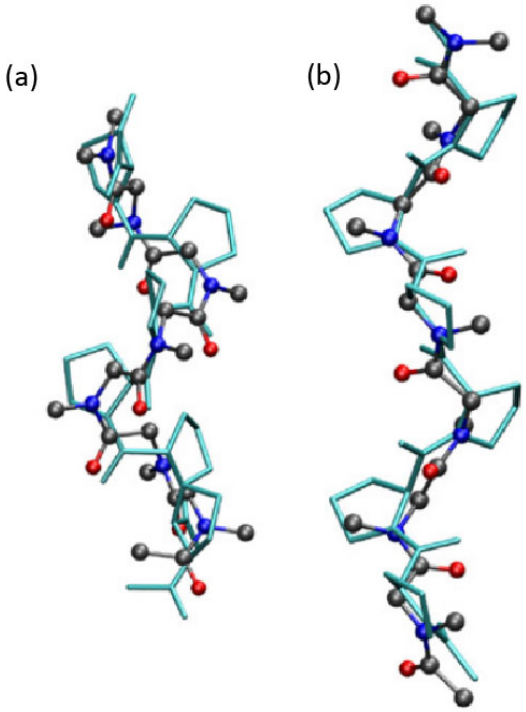



*Fig. 4.2:* Structures of sarcosine hexamers minimised in the (a) *cis* and (b) *trans* $\alpha_D$ conformations by Baldauf *et al.* The optimised structures of polyproline helices ((a) PPI and (b) PPII) are superimposed over the peptoid helices to highlight the similarities between these structures. Adapted with permission from [3].

of theory and using the 6-31G* basis set, show the presence of the previously identified conformational minima in longer sequences and also identify a new local minimum referred to as $\alpha/\alpha_D$. The $\alpha$ conformation with *cis* amides was found to be unstable relative to this novel $\alpha/\alpha_D$ conformation which has *cis* amides and $\phi$ and $\psi$ angles alternating between those of the $\alpha$ and $\alpha_D$ minima. The structural similarities between the *cis* and *trans* $\alpha_D$ conformations and polyproline helices were noted and illustrated by superimposing the latter over the peptoid helices. These are shown in Figure 4.2 where the peptoid helices are represented as ball and stick molecules with carbon atoms coloured in grey, nitrogen atoms in dark blue and oxygen atoms in red. The polyproline helices are overlaid in cyan.

Expanding the scope of *ab initio* studies of peptoids to consider the influence of side chain structure, Butterfoss *et al.* performed quantum mechanical calculations on a series of short peptoid oligomers, including the sarcosine dipeptoid. The backbone conformations of these oligomers, which include 7 cyclic and 3 linear sequences, were characterised experimentally by X-ray diffraction and solution NMR [4]. The experimentally obtained dihedral angles for peptoids with a variety of side chains were localised to specific regions corresponding to minima of the predicted Ramachandran energy landscape for the sarcosine dipeptoid. This is a promising indication that parameters for sarcosine are likely to be transferable to other peptoid structures and therefore parametrisation schemes for

atomistic force fields using sarcosine as a model system may be successful.

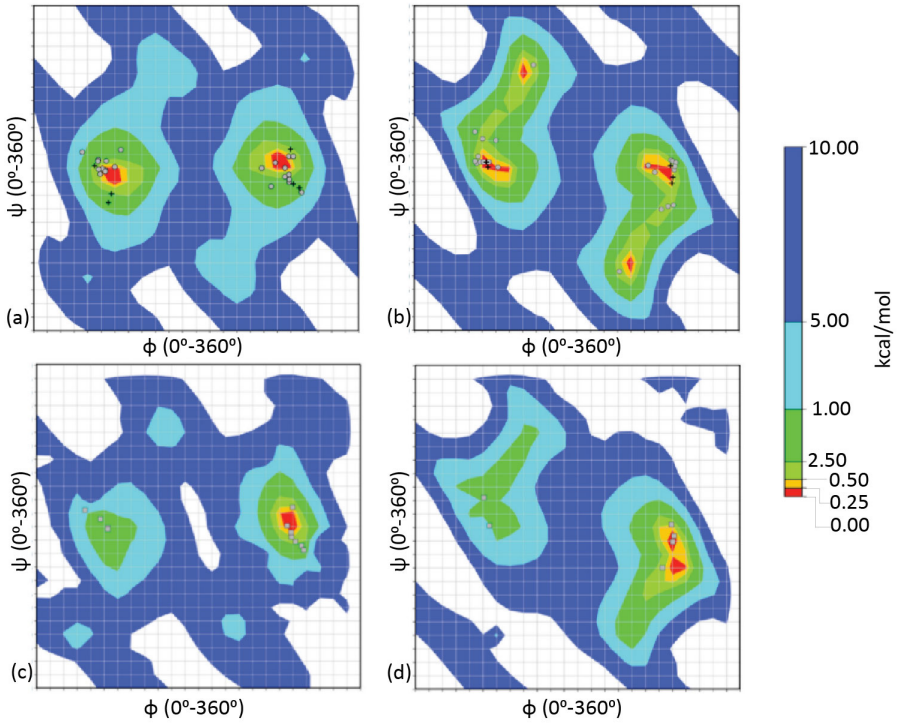



*Fig. 4.3:* Ramachandran-style energy landscapes generated by Butterfoss *et al.* using *ab initio* calculations. The energy landscape showing $\phi/\psi$ preferences for sarcosine are shown with (a) *cis* amide conformation and (b) *trans* amide conformation. Equivalent energy landscapes for a model peptoid with *N*spe side chain shown in (c) with *cis* amide conformation and (d) with *trans* amide conformation. The crosses and beige circles overlayed on the energy landscapes represent experimentally determined dihedral angles obtained for a wide range of side chains (a and b) and *N*spe side chains only (c and d). Adapted with permission from [4].

The energy landscapes (Figure 4.3) for peptoids with various side chains appear broadly similar, though with increased preference for the $\alpha_D$ conformation than the $C_{7\beta}$ relative to the sarcosine dipeptoid. Notably the *N*spe residue appears to impart handedness to structures by introducing a preference for the $\phi$ dihedral to adopt a negative angle, thus stabilising the *cis*-$\alpha_D$ minimum. This is consistent with experimental observations that *N*spe containing sequences tend to form right handed helices whereas equivalent sequences with *N*rpe form left handed structures [7, 8]. The side chain dihedral angle ($\chi_1$) preferences were also considered in this study. $\chi_1$ was found to have a broad minimum around $\pm$ 90° for almost all side chains studied, with a slight preference for the negative angle. Higher barriers to rotation around $\chi_1$ for *N*spe than other side chain groups imply that these side chains are restrictive and likely to stabilise particular conformations, as seen in the *N*spe helix. Prior to the Butterfoss study this particular helix was characterised by Armand *et al.* for an *N*spe octamer $(N\text{spe})_8$ [9]. Ramachandran plots for an *N*spe dipeptoid were obtained by mixed *ab initio* and all atom methods, using an all atom Amber force field

and the AMSOL force field for semi-empirical quantum mechanical calculations. The same distinctive asymmetry was observed in the *N*spe energy landscape as that shown in Figure 4.3c and 4.3d, which were subsequently generated at a higher level of theory by Butterfoss *et al.* [4].

Armand *et al.* describe the global minimum conformation for $(N\text{spe})_8$ as a right handed helix with all *cis* amides, $\phi$ of approximately -75° and $\psi$ of approximately 170°, as described in the introduction to Chapter 3. The authors noted that the barriers to conformational changes were small, particularly for the side chains to adopt different rotamers [9].

An alternative method for the computational prediction of peptoid secondary structure, with a focus on side chain dynamics, is to utilise the Rosetta framework [10]. This is not strictly a molecular dynamics approach as Rosetta utilises a conformational search algorithm based on the Metropolis Monte Carlo method to identify minimum energy conformations, refining identified structures based on a pre-calculated "rotamer" library of energetically allowed side chain conformations. Rosetta was designed for protein and peptide structures but parameters have been added by Drew *et al.* to enable peptoid conformational searches by modifying the search process to include *cis-trans* isomerisation and adjusting the backbone dihedral potentials [11]. Subsequently Renfrew *et al.* generated a large library of peptoid rotamers for use in Rosetta and validated Drew's work against known peptoid structures [12].

### *Atomistic Peptoid Force Fields*

Many decades of work has gone into developing, testing and refining the range of force fields that currently exist for atomistic simulations of biomolecules such as peptides and proteins [13–15]. Adapting and re-parametrizing a force field is therefore not a trivial task and starting from scratch would be an enormous undertaking. In particular, the lack of experimental structural data (especially X-ray crystal structures) available for peptoids adds to the general difficulties. Various strategies for producing atomistic force fields which accurately describe peptoids have been reported in the literature [1]. Each of these cases is based on adapting existing force fields which were originally parametrized for peptides, which seem the most suitable starting point for peptoid force fields.

Moehle and Hofmann were the first to simulate sarcosine monomers at the atomistic level and found good agreement with their results obtained using the CHARMM22 force field and their *ab initio* results, with each of the 3 main minima being reproduced by the force field [2]. However, the atomistic model has a considerably lower energy barrier between the $\alpha$-helical minimum and the other minima than the *ab initio* results suggested and was found to over-stabilise the *cis* $\alpha$ conformation. This suggests that force fields designed for peptides may potentially bias peptoids towards $\alpha$-helical configurations [2].

The Optimized Potential for Liquid Simulations (OPLS) force field has also been used

without modification to simulate sarcosine polypeptoids, as well as similar sequences with methoxyethyl ($CH_3OC_2H_5$-) side chains [16]. Force field parameters were assigned to the peptoid molecules in this study by referring to existing parameters for amino acids and tertiary amides. The dihedral torsions for peptides were applied to the peptoid systems enforcing the same level of planarity on the peptoid backbone amide as the peptide. This study included some of the longest peptoid sequences thus far examined by simulation, up to 50 residues in length [16].

It was found that *trans* $\alpha_D$ conformation was favoured in sarcosine polypeptoids ranging in length from 5 to 50 residues. The Ramachandran plot for the sarcosine 50mer closely resembled that of the energy landscape of the dimer produced using *ab initio* methods by Butterfoss *et al.*, indicating that the folding properties of the dimer unit are applicable to longer sequences. The authors assert that this indicates that peptoid folding is driven by interactions at the local level rather than long range intra-molecular interactions. This is somewhat contradictory to the evidence of Shin *et al.* (described in Chapter 3), which indicates that incorporation of small numbers of helix inducing residues can enforce a significant degree of helicity in the peptoid backbone [17].

In the OPLS study the substitution of the sarcosine side chain to a methoxyethyl side chain had a sequence length dependent effect on the backbone properties, with the Ramachandran plot only resembling that of sarcosine in longer sequences. The rate of *cis-trans* isomerisation and total *cis* amide population during simulations was also found to vary with the sequence length of sarcosine polypeptoids, reaching a maximum of approximately 20% *cis* amides in the 50mer. The sarcosine polypeptoids were also found to favour left handed helical structures, in contradiction to the symmetric nature of the *ab initio* sarcosine Ramachandran plot [16].

Given the apparent biasing effect of certain force fields towards peptide specific conformations, attempts to adapt force field parameters for peptoids have been made. Mirijanian *et al.* [6] parametrised CHARMM22 to improve the representation of peptoid conformational preferences in this force field. To construct this adapted force field, which was named MFTOID, parameters for peptoid backbone atoms were transferred from similar atoms in the peptide backbone. Additionally, certain parameters were re-fitted to quantum mechanical data using the sarcosine dipeptoid as a model system. These included the Lennard-Jones parameters for the carboxyl carbon and the partial charges on the amide nitrogen and carboxyl carbon. The potential terms for the $\phi$ and $\psi$ dihedral were not adjusted but the $\omega$ term was refitted to allow *cis/trans* isomerisation [6].

Mannige *et al.* [18] used the MFTOID parameters to investigate peptoids as a potential building block for higher order structures, focussing particularly on nanosheets. They demonstrated that the formation of stable supramolecular assemblies from peptoids with combinations of aromatic, anionic and cationic side chains is both possible and dependent

on the length of the peptoid sequences, corroborating their computational results with microscopy and X-ray scattering data [18]. As a result of this work, a novel secondary structural motif was discovered which allows the peptoids to distribute their charged and aromatic side chains in a way which stabilises the nanosheet conformations. Termed the $\Sigma$ strand, this is a linear structure with a distinctive zigzag conformation which results from all *trans* amides and alternating pairs of $\phi$ and $\psi$ of opposite handedness.

Significant work has been done by the Voelz group to evaluate and improve Amber force fields for peptoid simulations. In one study Voelz *et al.* [19] simulated a selection of peptoid sequences previously studied in the *ab initio* work by Butterfoss *et al* [4], to evaluate how well certain force fields capture the backbone dihedral preferences of peptoids. The force fields considered were OPLS, AMBER ff96, AMBER ff99, AMBER ff03 and GAFF. Simulations were performed with both explicit and implicit water models using the AMBER MD engine and Replica Exchange MD (REMD) was used to enhance conformational sampling.

Ramachandran free energy landscapes for the different force field and solvation method combinations showed that each reproduced the *ab initio* results to some extent though GAFF and implicit solvation with the GBSA model was identified as one of the most promising. The latter however failed to reproduce the negative $\phi$ preference associated with *N*spe. Further work by Mukherjee *et al.* sought to correct this by tuning GAFF for peptoids and specifically the *N*spe helix [20]. This new version of GAFF was termed GAFF-$\phi$ as it was adapted by the addition of an *ad hoc* harmonic biasing potential which enforced a preference for negative $\phi$ angles of approximately 1 kcal/mol. Using these new parameters Mukherjee *et al.* carried out a detailed analysis of the folding of *N*spe helices. By adapting the Lifson-Roif helix-coil theory, which is often used to descibe helix formation in peptides, Mukherjee *et al.* calculated the enthalpies and entropies of *N*spe helix formation, demonstrating that this is entropically favourable and most likely occurs due to the steric interactions between the side chain groups.

In each of the above studies normal MD was not sufficient to capture the full range of behaviours of the peptoid backbone due to the high barrier associated with amide *cis/trans* isomerism [21]. The barrier to rotation around tertiary amides is particularly high [22] and as such isomerisation in simulations is a rare event that can occur on very slow time scales, on the order of microseconds [23], in a normal room temperature MD simulation. Therefore enhanced sampling techniques were required. Mirijanian *et al.* used umbrella sampling to this end whereas the Voelz and Mukherjee studies used temperature REMD to overcome the barriers between conformational minima [19, 20]. In future atomistic studies balancing the computational expense associated with these techniques and the need to achieve sufficient sampling will be of importance. This remains somewhat of a challenge which we attempt to address in this work through the use of Hamiltonian replica

exchange.

## 4.2 Optimization of the General Amber Force Field to Describe Peptoids

### *4.2.1 GAFF as a Starting Point for a Peptoid Force Field*

The initial choice of a force field to adapt is likely to affect the results obtained. The General Amber Force Field (GAFF) was selected as a starting point for this work. This force field was first described by Wang *et al.* in 2004 as a general Amber force field for organic molecules [24]. As the name suggests, it is designed to be very general, describe a wide range of organic and pharmaceutical molecules and be compatible with the existing Amber protein and peptide force fields. Voelz *at al.* [19] carried out a study where they tested the accuracy of various force field and solvation method combinations and GAFF emerged as a promising candidate which was particularly accurate when combined with implicit solvation methods. Here we aim to modify GAFF to allow for accurate peptoid simulations using explicit solvation. Ultimately, we aim to use the adapted GAFF to simulate the repeat motif peptoids sequences, which we have previously characterised by experimental methods, explicitly solvated in water and octanol.

GAFF employs a simple harmonic functional form shown in equation 4.1.

$$E_{\text{pair}} = \sum_{\text{bonds}} K_r(r - r_{\text{eq}})^2 + \sum_{\text{angles}} K_\theta(\theta - \theta_{\text{eq}})^2 + \sum_{\text{dihedrals}} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)] + \sum_{i<j} \left[ \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\epsilon R_{ij}} \right] \tag{4.1}$$

In this work we focus on refitting the torsional parameters that differ between peptides and peptoids. This was approached in a manner that required a minimal number of parameters to be changed. Therefore we focus only on the dihedral term of the GAFF functional form and aim to modify the coefficients, $V_n$, to capture peptoid backbone torsional preferences better than those in the native force field. This is done by introducing a new set of $V_n$ coefficients, such that the energy when rotating round each angle matches the quantum mechanical energy rotating around the same angles. We focused on these torsional parameters as they are likely to be non-transferable between peptides and peptoids as a result of their subtly different backbone architecture.

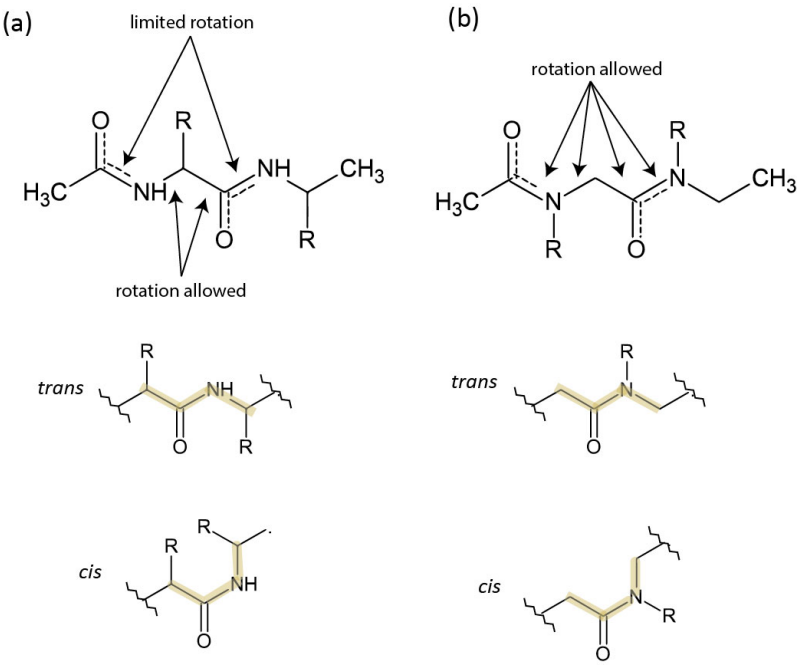



*Fig. 4.4:* Comparison of peptide and peptoid backbones. (a) shows the peptide backbone with fragments showing the steric clash between side chains when the peptide secondary amide is in the *cis* conformation. (b) shows the peptoid backbone with the fragments demonstrating the steric equivalence between the *cis* and *trans* conformations in the peptoid tertiary amide.

### *Atomistic Structural Differences Between Peptides and Peptoids*

The substitution of the side chain from the $\alpha$-carbon onto the amide nitrogen introduces subtle differences into the torsional properties of the peptoid backbone, as compared to the peptide backbone. Given that the existing torsional parameters in GAFF have been parametrized for the peptide backbone, adjustments must be made for the simulation of peptoids. Several of the torsions present in peptoids may also be found in peptides and therefore we assume that these particular parameters are transferable.

The peptide bond has partial double bond character which renders it rigid and planar (Figure 4.4). This results in a high barrier to rotation around the bond and *cis-trans* isomerism occurring extremely rarely. The *trans* conformation is favoured in the overwhelming majority of cases [25] due to the steric clash between side chains when the amide is in the *cis* conformation making the *trans* the lower energy conformation. Polyproline peptides are the exception to this as the cyclisation of the side chain which includes the amide nitrogen results in the *cis* and *trans* states being considerably closer in energy and therefore both exist in natural structures.

Similarly in peptoids, the substitution of the side chain onto the amide nitrogen results in *cis* and *trans* isomers that are close in energy and thus *cis-trans* isomerisation occurs much more readily than it does in peptides. The peptoid backbone is also set apart from the peptide backbone in that it is achiral and does not form hydrogen bonds equivalent to

those which stabilise the distinctive helical conformations seen in peptides. This, combined with allowable rotations around all three backbone dihedral angles renders the peptoid backbone considerably dynamic and flexible.

### 4.2.2 *Creating Topology Files for Peptoids*

We used the Amber Tools suite of programs [26] to obtain GAFF parameters as a starting point for our peptoid force field. The protocol followed for preparing files for peptoid simulations is summarised in Figure 4.5. Crystal structures are not available for the 12-mer peptoids that we wished to study, therefore we generated a starting conformation for the peptoid backbone using the ProBuilder online platform [27]. This service is designed to construct 3D peptides from the amino acid sequence using user defined backbone dihedral angles to specify the secondary structure. In order to construct a peptoid backbone we used the program to create coordinates for a *cis*-amide poly-glycine helix with backbone dihedral angles $\phi$ = -75° and $\psi$ = 170°. These angles were selected as they have been predicted for an idealised *N*spe helix [9]. We then used the Avogadro molecular editor [28] to add side chains to the peptoid backbone. The side chain geometry was optimised in Avogadro using the steepest descent algorithm and the GAFF force field. During this optimization the backbone atoms were restrained.

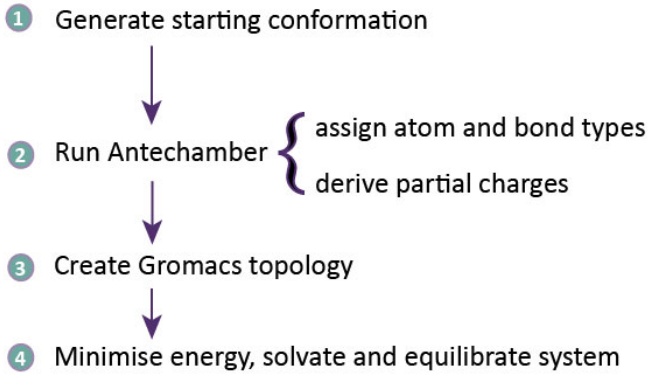



*Fig. 4.5:* Preparation of input files for peptoid simulations in GROMACS.

Using the Avogadro starting structures, Amber's Antechamber package was used to generate the force field parameters for each peptoid topology file. This is a suite of programs that allows automated assignation of atom types and bond types and partial charge derivation. Parmchk was used to fill in any missing parameters [29]. Antechamber outputs topology information in the Amber format. These were converted to GROMACS format using the AnteChamber PYthon Parser interfacE (ACPYPE). ACPYPE is a wrapper script designed to facilitate the use of Antechamber for non-Amber users by converting the outputs into other formats [30].

This same protocol was followed to produce topology and coordinate files for the sarcosine dipeptoid, which was used as a model molecule to re-parametrise GAFF. However, due to its small size, this molecule was built in an extended conformation and minimised in Avogadro, without the use of Pro-Builder to generate a helical backbone. Two versions of this peptide were generated, one with the backbone amide in the *cis* conformation and the other in the *trans* conformation.

### *Derivation of Partial Charges*

Differing atomic electro-negativities result in a non-uniform distribution of electron density within molecules. This dictates how the total charge is distributed across the molecule, which in turn affects how it behaves during simulations [31]. For example, in the carbonyl groups present in both peptide and peptoid backbones, the oxygen atom will have a small negative electrostatic charge associated with it and the carbon atom to which it is bonded will have a small positive charge. These are known as partial charges.

It is necessary to assign partial charges to atoms for MD simulations so that the electrostatic interactions can be treated with a simple Coulombic model within the force field (described in Chapter 2). Given the structural differences between peptoids and peptides the distribution of partial charges between the two classes of molecule are likely to vary. In addition, peptoids can contain side chain groups that differ from the native amino acids present in peptides and therefore charges will need to be derived for these too. Unlike parameters such as bond length, bond angle and Van der Waals coefficients, which should be transferable from peptide models, the partial charge on a particular atom will vary depending on its connectivity and local environment and therefore should be derived specifically for the peptoid backbone and non-native side chains.

In this work we use Antechamber to derive the partial charges for our atomistic peptoid model. Within Antechamber there are several different methods to do this. Two distinct quantum mechanical methods are frequently used to derive charges on small molecules: restrained electrostatic potential (RESP) fitting and the AM1-BCC model. In RESP fitting the charges assigned to each atom are fitted to reproduce a Hartree-Fock (H-F) quantum mechanically derived grid of electrostatic potential (ESP) points [32, 33]. The AM1-BCC method is a semi-empirical method which was designed to reproduce ESP results but with reduced computation time required. This is a two step method where initially the formal charges and electron distributions are captured at the Austin Model 1 (AM1) level [34]. This alone does not reproduce the quantum mechanical H-F ESP at a satisfactory level, so the second step is the application of simple bond charge corrections (BCC) to achieve this. AM1-BCC methods have been globally parametrised to produce a set of BCCs applicable to any small organic molecule [35]. In this work we have used the AM1-BCC method to derive the partial charges for our peptoid topology files.

Each peptoid was run through Antechamber twice, to generate two topology files with subtly different partial charges to account for the fact that the cationic residues, *N*ae and *N*Lys would be protonated in water/PBS, giving the peptoid an overall positive charge, whereas the same residues would not be protonated in octanol, thus resulting in the peptoid remaining neutral overall.

## 4.3 Results and Discussion

### *4.3.1 Derived Partial Charges for Peptoids*

We calculated the partial charges on the atoms in each peptoid in Antechamber according to the AM1-BCC method. The two key atoms where the partial charges are expected to be different in peptoids and peptides are the amide nitrogen and alpha carbon, due to the altered connectivity of these atoms by the side chain substitution. It is also important to verify that the backbone atoms in the longer peptoid sequences are assigned similar partial charges by Antechamber to the backbone atoms in sarcosine, which was used to re-parametrise the force field.

The amide nitrogen atoms in the sarcosine dipeptoid were assigned an average partial charge of -0.4643 e with a standard deviation of 0.0163 e. The amide nitrogen atoms in the longer peptoids had an average partial charge of -0.45562 e with a standard deviation of 0.11585 e. There is therefore good consistency between the model peptoid and longer sequences. This indicates that, although the peptoid 12-mers (approximately 260 atoms) are at the upper limit of system size which can be processed in Antechamber, the charge derivation appears to be consistent.

### *4.3.2 Optimization of Peptoid Backbone Torsional Potentials*

We selected the sarcosine dipeptoid as a small molecule to use as a model for parametrising GAFF for peptoid simulations. The sarcosine dipeptoid is the peptoid analogue of the alanine dipeptide and a structural comparison between the two is shown in Figure 4.6. The latter was used extensively in the development and parametrisation of protein and peptide force fields [36] including Amber force fields [37]. We focussed on re-parametrisation of the peptoid backbone torsions. The alanine dipeptide is generally considered to be a good model for representing protein backbone torsional preferences, thus by analogy the sarcosine dipeptoid should represent the peptoid backbone well, as it contains all the relevant backbone features which may occur in peptoid sequences. Previously reported work has suggested that there should be a good level of transferability for parameters from the sarcosine dipeptoid to other peptoid structures [2]. Both the quantum mechanical

studies of Butterfoss *et al.* [4] and the Voelz group's work on GAFF [19] have demonstrated the dominance of sterics in determining peptoid secondary structure and that the sarcosine backbone dihedral minima are similar to those where a range of different side chains are present.

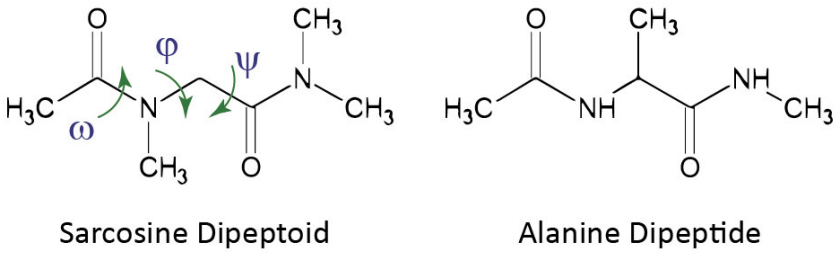



*Fig. 4.6:* Sarcosine dipeptoid primary structure (left) compared to alanine dipeptide (right). The backbone dihedral/torsional angles, $\phi$, $\psi$ and $\omega$ are marked on the sarcosine structure.

The peptoid backbone torsional angles $\omega$, $\phi$ and $\psi$ are marked on the disarcosine structure shown in Figure 4.6. These are represented in the force field by the proper dihedrals. In order to assess the accuracy of the GAFF parameters in capturing the dihedral preferences in sarcosine we investigated the minimum energy conformation of the molecule at set dihedral angles. Quantum mechanical energy minimisations were performed by Mark R.Wilson for the structure rotated around each of the three backbone dihedrals in turn and the outcome compared to the equivalent force field energies. The quantum mechanical calculations were performed using Gaussian09 [38] at the B3LYP level of theory. The dihedrals were scanned at 6° intervals through the full range of 0° to 360°. At each point, the dihedral was fixed while the rest of the structure was minimised and the energy calculated. The energies were then plotted against angle, giving an energy profile for rotating around each dihedral. The same process was carried out in GROMACS using a program written by Mark R. Wilson to calculate the GAFF rotational energy profiles, scanning the full range of dihedrals at 3° intervals. Interpolation between data points was carried out in order to directly compare the QM and force field energies at 1° intervals. The rotational energy profiles were not always symmetric due to the occurrence of nitrogen inversion, therefore the calculations were carried out starting at different angles and cycling through the full range both clockwise and anticlockwise. The minimum energy at each angle was extracted in order to construct the profiles shown in Figure 4.7.

### 4.3.3 Rotational Energy Profiles and Optimised Backbone Dihedral Parameters

The GAFF rotational energy profiles and the quantum mechanical (QM) energy profiles that we obtained are broadly similar. These are shown for all three backbone dihedrals in Figure 4.7. This suggests that GAFF in its unmodified state represents the peptoid backbone torsions reasonably well and is therefore a good starting point for a modified force field. The differences between the profiles could result in some inaccuracy or bias towards certain conformational minima if the unmodified force field was used.

The energy minimum at 0° (*cis* conformation) for the $\omega$ dihedral was less deep for the force field in comparison to the quantum mechanical profile, indicating a slight bias towards a *trans* backbone conformation in unmodified GAFF. For the $\phi$ dihedral, GAFF produced two minima of equal depth around ±120° with a barrier of approximately 25 kJ/mol. These GAFF minima were broader than their quantum mechanical counterparts with an approximately 8 kJ/mol lower barrier between the positive and negative minima. The $\psi$ dihedral GAFF and QM rotational energy profiles were the least well matched of all three backbone dihedrals. Here, the GAFF and QM energy profiles were similar in shape but the positions of the minima and height of the barriers were mismatched.

In order to correct these differences new parameters were obtained to map the force field energies onto the quantum mechanical energies. This was achieved by fitting a new set of parameters to minimise the squared difference between the the GAFF and the QM energies over the full range of angles, $\phi_i$, according to equation 4.2. We refer to the version of GAFF with these modified parameters as the General *N*-substituted Amber Force Field (G*N*AFF).

$$\chi^2 = \sum_{i=0}^{i=360} [E_{\mathrm{QM}}(\phi_i) - E_{\mathrm{MD}}(\phi_i)]^2 \tag{4.2}$$

The parameters fitted correspond to the $V_n$ coefficients of the dihedral potential in the GAFF force field (equation 4.3). However, this dihedral potential is slightly different to the one used in the GROMACS topology files generated by ACPYPE, which is known as the Ryckaert-Bellemans function and shown in equation 4.4, where $\psi = \phi - 180°$.

$$E_{\mathrm{dihedral}} = \frac{V_n}{2}[1 + \cos(n\phi - \gamma)] \tag{4.3}$$

$$E_{\mathrm{RB}}(\phi_{ijkl}) = \sum_{n=0}^{5} C_n(\cos(\psi))^n \tag{4.4}$$

To modify GAFF for parameters to use in GROMACS we therefore fitted the Ryckaert-Bellemans function for the parameters $C_n$, $n$=0-5. This is the format in which the script

ACPYPE produces the dihedral potential when converting the Antechamber output for GROMACS. The relationship between the Ryckaert-Bellemans dihedral coefficients and the Amber dihedral coefficients can be shown by expressing the Amber dihedral term as a series of Fourier dihedrals, shown in equation 4.5.

$$E_{\text{dihedral}} = \frac{1}{2}[V_1(1+\cos(\phi)) + V_2(1-\cos(2\phi)) + V_3(1+\cos(3\phi)) + V_4(1-\cos(4\phi))] \quad (4.5)$$

Substitutions using the multiple angle formulae for cosine allow the $C_n$ coefficients to be expressed in terms of the $V_n$ coefficients. These are listed in equations 4.6-4.11.

$$C_0 = V_2 + \frac{1}{2}[V_1 + V_3] \quad (4.6)$$

$$C_1 = \frac{1}{2}[V_1 - 3V_3] \quad (4.7)$$

$$C_2 = 4V_4 - V_2 \quad (4.8)$$

$$C_3 = 2V_2 \quad (4.9)$$

$$C_4 = -4V_4 \quad (4.10)$$

$$C_5 = 0 \quad (4.11)$$

The rotational energy profiles for the three backbone dihedrals using the modified G*N*AFF parameters are shown in Figure 4.7 parts ai, bi and ci, along with the QM and unmodified GAFF profiles for comparison in a, b and c.

The force field modifications correct the mismatches between GAFF and the QM profiles in the positions of the maxima and minima and the height of barriers between them. Notably, the energy difference between the *cis* and *trans* conformations is smaller in G*N*AFF than GAFF. G*N*AFF also has a slightly more pronounced additional minimum in the $\phi$ profile than GAFF which resembles a small shoulder at $\pm 75°$ and corresponds to the $\alpha_D$ minimum. In the $\psi$ energy profile for G*N*AFF the positions of the minima and the barrier heights closely match the QM versions, correcting for the presence of minima corresponding to the peptide $\alpha$-helical conformation at $\pm 55°$ seen in the GAFF profile, which could bias peptoids towards a typical peptide conformation.

The original GAFF RB parameters for the peptoid dihedrals derived by ANTECHAMBER and the modified G*N*AFF parameters are summarised in Table 4.1. Several sets of atoms can contribute to the same rotation about any particular dihedral angle. The RB coefficients reported in Table 4.1 are those which only directly concern the heavy atoms in the peptoid backbone (carbons and nitrogens). Thus the $\omega$ coefficients listed in Table 4.1 correspond to the rotation around the C-N bond, defined by the four backbone atoms C-C-N-C. Likewise the $\phi$ coefficients correspond to rotation around N-C, defined by C-N-

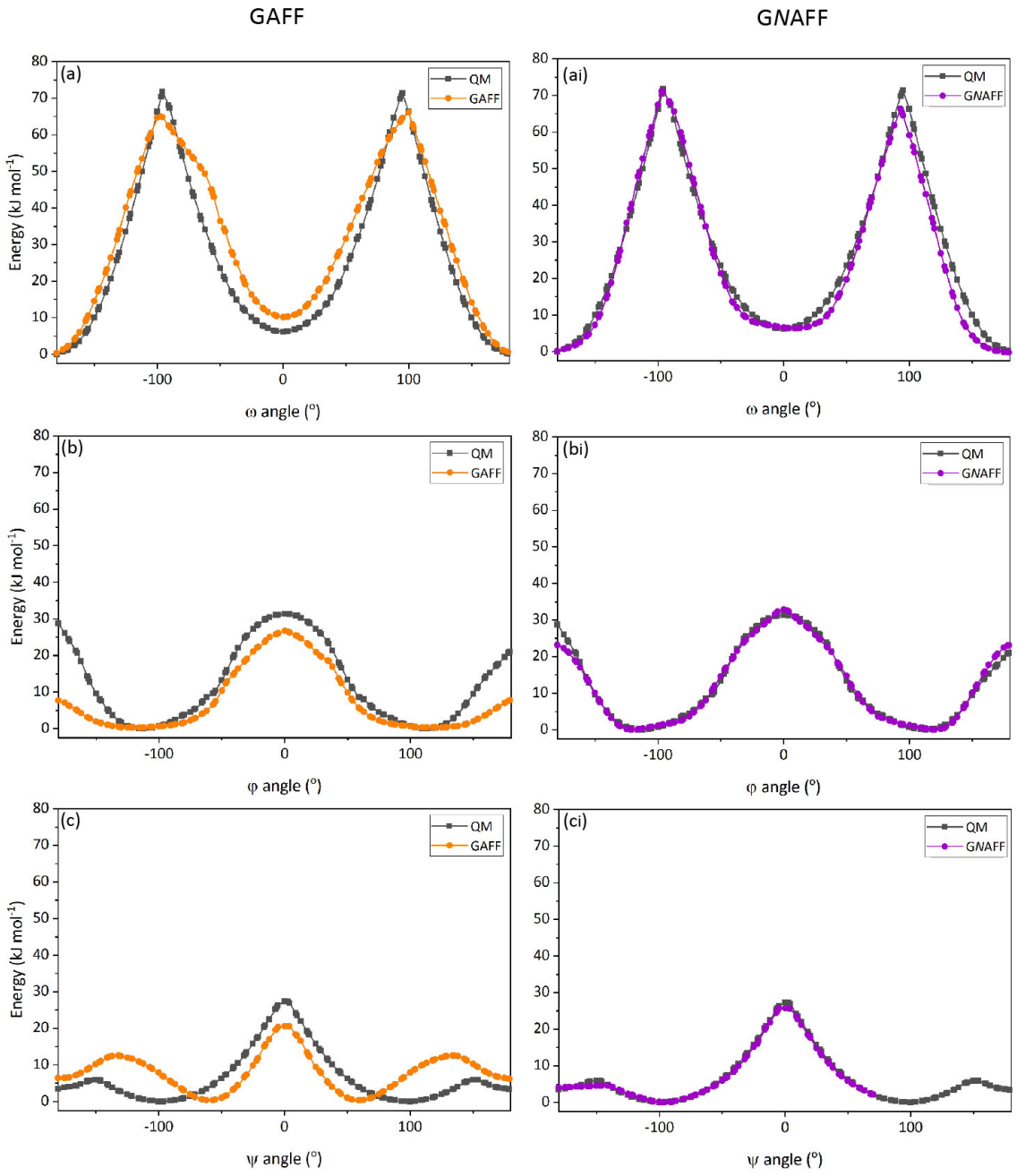



*Fig. 4.7:* Comparison of QM and modified GAFF rotational energy profiles for (a) $\omega$, (b) $\phi$ and (c) $\psi$ dihedrals in the sarcosine dipeptoid backbone. Quantum mechanical profiles shown in black, MD profile with original GAFF coefficients shown in orange, new RB coefficients for G*N*AFF force field shown in purple.

C-C and the $\psi$ coefficients correspond to rotation around C-C, defined by N-C-C-N. In the case of each dihedral there are additional sets of four atoms which also contribute to the same rotations, for example C-N-C-H contributes to $\phi$, as illustrated in Figure 4.8. These additional contributing sets of atoms are not considered, as the modification to the heavy atom contributions appears to sufficiently adjust the force field to capture the quantum mechanical behaviour correctly.

| | | Ryckeart-Bellemans Coefficients | | | | | |
|---|---|---|---|---|---|---|---|
| Force Field | Angle | C0 | C1 | C2 | C3 | C4 | C5 |
| GAFF | $\omega$ | 20.92000 | 0.00000 | -20.92000 | 0.00000 | 0.00000 | 0.00000 |
| | $\phi$ | 10.46000 | -3.34720 | -7.11280 | 0.00000 | 0.00000 | 0.00000 |
| | $\psi$ | 23.84880 | 7.11280 | -16.73600 | 0.00000 | 0.00000 | 0.00000 |
| G*N*AFF | $\omega$ | 20.78616 | 11.16128 | -57.13083 | -23.30766 | 34.97895 | 13.83058 |
| | $\phi$ | 9.96705 | -3.64772 | -10.80460 | 1.11549 | 14.03927 | 5.67685 |
| | $\psi$ | 18.55031 | -11.26660 | -8.10125 | 20.20707 | -1.43807 | -5.85075 |

*Tab. 4.1:* Table of Ryckaert-Bellemans ($C_n$) coefficients for peptoid backbone dihedrals, derived for GAFF in Antechamber and modified by fitting to QM data. The dihedral $\omega$, $\phi$ and $\psi$ refer to only the heavy, backbone atom contributions to these rotations. The atoms considered in $\omega$ are C-C-N-C. In $\phi$ the atoms considered are C-N-C-C and in $\psi$ the atoms considered are N-C-C-N.

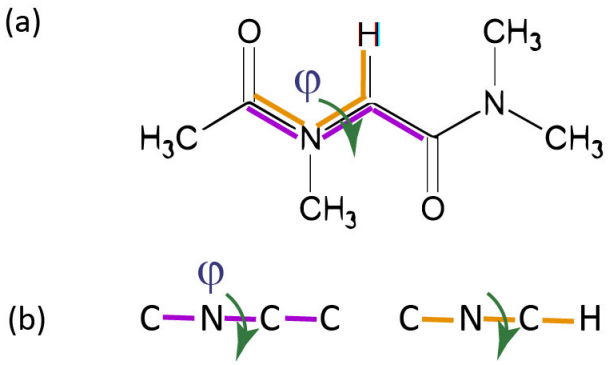



*Fig. 4.8:* (a) Sarcosine dipeptoid with $\phi$ dihedral and two separate contributing sets of atoms, C-N-C-C and C-N-C-H, highlighted in magenta and orange respectively. (b) Linear representations of the C-N-C-C ($\phi$) (magenta) and C-N-C-H (orange) contributions to the rotation around the N-C bond.

Overall, modifying the heavy, backbone atom $C_n$ coefficients in the RB potential appears to have improved the representation of the peptoid backbone torsions in GAFF, though further efforts at force field validation are presented subsequently, where we will also consider the effect of solvation on conformational preferences. It is important to consider that, while the new G*N*AFF parameters produce energy profiles that closely match our QM energy profiles, the QM data will vary depending on the level of theory at which the calculations were carried out. Moehle and Hofmann give an extensive report of the relative energies of the backbone dihedral conformations in the sarcosine dipeptoid, calculated quantum mechanically at various levels of theory [2]. The size of barriers and position of minima for G*N*AFF, which are summarised in Table 4.2, are similar to those obtained by other groups and presented in the literature [6, 19]. In their work developing a CHARMM based peptoid force field Mirijanian *et al* report a quantum mechanical barrier to rotation around $\phi$ in the sarcosine dipeptoid of approximately 25 kJ/mol, from DF-MP2/aug-cc-pVQZ//MP2/6-311G* single point energy calculations. They also report a *cis-trans* barrier for solvated sarcosine of approximately 49 kJ/mol [6], with the *trans*

conformation being approximately 5 kJ/mol lower in energy than the *cis* conformation, using their force field parameters.

| | Minima | | Maxima | |
|---|---|---|---|---|
| Angle | Position (°) | Energy (kJ/mol) | Position (°) | Energy (kJ/mol) |
| $\omega$ | ± 180 | 0.0 | ± 96 | 71.8 |
| | 0 | 6.6 | - | - |
| $\phi$ | ± 120 | 0 | ± 180 | 30 |
| | ± 75 | 3.7 | 0 | 32.8 |
| $\psi$ | ± 96 | 0 | 0.0 | 25.7 |
| | ± 180 | 3.4 | 146 | 4.5 |

*Tab. 4.2:* Table of relative energies of peptoid torsional mimima and maxima in G*N*AFF force field. These values are taken from the 2D rotational energy profiles shown in Figure 4.7 where the energy values are relative to the global minimum, set to zero.

*Parameters Derived for Modified Potential for Hamiltonian Replica Exchange*

In order to implement Hamiltonian Replica Exchange (HREX) to improve conformational sampling in our simulations we derived an additional set of RB parameters ($C_n$) to create a modified potential for the $\omega$ dihedral, with a lower barrier between the *cis* and *trans* conformations. The low barrier $\omega$ rotational energy profile is shown in Chapter 2 in Figure 2.7, in direct comparison to the standard G*N*AFF energy profile (which can also be seen in Figure 4.7 ai). RB parameters to describe this modified potential within the force field were derived in the same way as the G*N*AFF parameters in the previous section. New parameters were derived to minimise the squared difference between the low barrier potential and the G*N*AFF potential according to equation 4.2. The new RB parameters obtained are given in Table 4.3.

| | Ryckeart-Bellemans Coefficients for $\omega$ dihedral | | | | | |
|---|---|---|---|---|---|---|
| Barrier | C0 | C1 | C2 | C3 | C4 | C5 |
| Standard | 20.78616 | 11.16128 | -57.13083 | -23.30766 | 34.97895 | 13.83058 |
| Low | -18.58464 | -9.16799 | 33.56821 | 23.91972 | -17.22789 | -13.48187 |

*Tab. 4.3:* Table of Ryckeart-Bellemens ($C_n$) coefficients for peptoid backbone C-C-N-C $\omega$ dihedral derived for G*N*AFF (as shown previously in Table 4.1) and the modified form of G*N*AFF for a low barrier potential for HREX simulations.

The energy barriers to rotation around the other backbone dihedrals, $\phi$ and $\psi$, are considerably lower than the $\omega$ barrier and therefore the RB parameters pertaining to these angles were left unchanged for the second replica in HREX simulations.

### 4.3.4 *Validation of GNAFF for Small Model Peptoid Systems*

*Dimethylacetamide*

Though not strictly a peptoid, dimethylacetamide (DMA) is a small molecule which is peptoid-like in that, being a single tertiary amide, it contains the set of atoms which make up a single unit of the peptoid backbone. The structure of DMA is shown in Figure 4.9a. We selected DMA as a model to test the *cis-trans* sampling of the peptoid amide bond with the newly derived G*N*AFF parameters. The *cis/trans* barrier of DMA has been determined experimentally to be in the region of 71-73 kJ mol$^{-1}$ [39, 40].

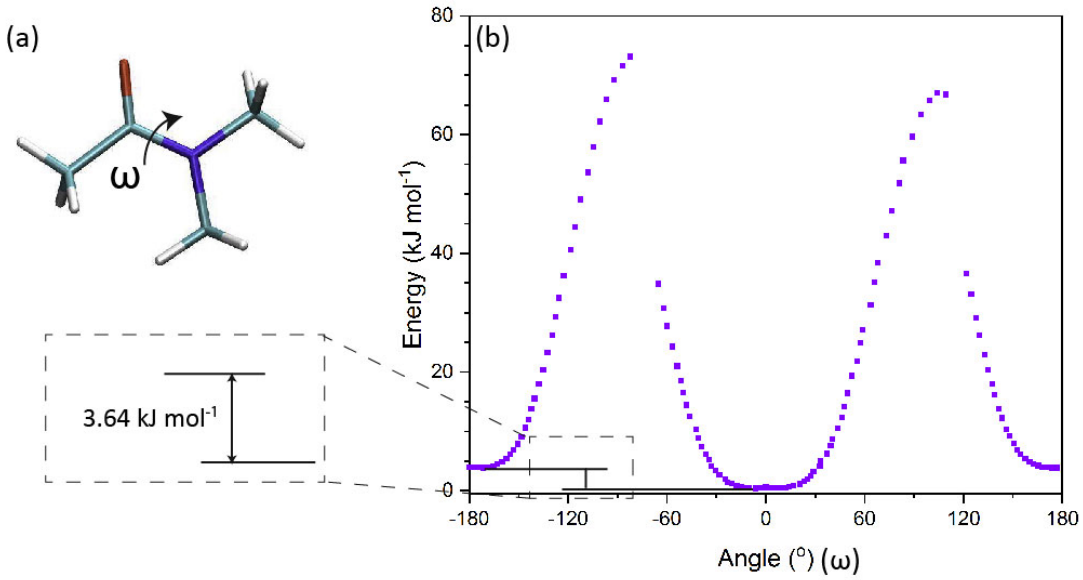



*Fig. 4.9:* (a) Representative structure of DMA. Carbon represented in aquamarine, nitrogen in royal blue, hydrogen in white and oxygen in red. (b) Rotational energy profile for DMA with G*N*AFF parameters.

Figure 4.9b shows the rotational energy profile for the $\omega$ dihedral in DMA (the only peptoid relevant backbone dihedral in this molecule) with G*N*AFF force field parameters. The *cis/trans* barrier here is approximately 75.5 kJ mol$^{-1}$ which is close to the experimental value. The depth of the 0° (*cis* conformation) minimum is approximately 3.64 kJ mol$^{-1}$ lower than the 180° (*trans* conformation). This slight favouring of the *cis* conformation is not observed quantum mechanically (in this case the minima are equal in depth) but is a result of using the G*N*AFF force field parameters which has imposed the slight increase in the depth of the *cis* minimum that is observed in sarcosine relative to the unmodified GAFF force field.

We tested whether modification of the force field parameters affected the heat of vaporization ($\Delta_{\text{vap}}H$) of DMA. $\Delta_{\text{vap}}H$ was calculated according to equation 4.12, where $E_{\text{pot}}(g)$ is the intramolecular energy in the ideal gas phase and $E_{\text{pot}}(l)$ is the intermolecular energy

in the liquid phase. These two quantities were obtained via simulation.

$$\Delta_{\text{vap}}H = (E_{\text{pot}}(g) + RT) - E_{\text{pot}}(l) \tag{4.12}$$

The ideal gas phase simulation of a single DMA molecule was carried out using the stochastic dynamics (SD) integrator in GROMACS. The liquid phase MD simulation of a cubic periodic box containing 257 DMA molecules was carried out at 298 K and a pressure of 1 bar, which were maintained during the production run by a Nosé-Hoover thermostat and a Parrinello-Rahman barostat. Prior to the production run the box was compressed at pressure of 100 bar and then allowed to relax to to the correct liquid state density. The density of the DMA over a 100 ps equilibration run was 945.96 kg/m$^3$, which is very close to the experimental value of 940 kg/m$^3$ [41, 42]. The initial equilibration runs were carried out with a Berendsen barostat before switching to Parrinello-Rahman for the production run.

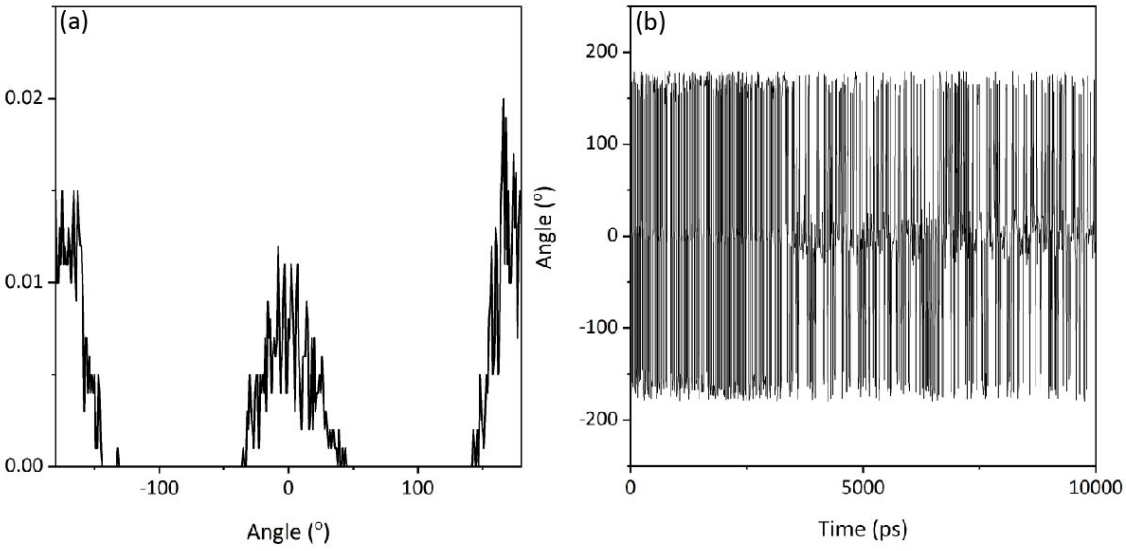



*Fig. 4.10:* (a) Distribution of $\omega$ angles for DMA solvated in SPC water over 10 ns HREX simulation. (b) *cis-trans* isomerism of $\omega$ in DMA over the 10 ns HREX simulation. Data is from the standard G*N*AFF force field replica.

The modifications to the force field parameters made only a tiny difference to the heat of vaporization of the DMA calculated by this method. $\Delta_{\text{vap}}H$ calculated from simulations with the original GAFF parameters was 55.52 kJ mol$^{-1}$ and with the modified, G*N*AFF parameters was 55.49 kJ mol$^{-1}$.

We also investigated the *cis-trans* isomerism in DMA as a preliminary test of the efficacy of HREX methods for sampling over the high barriers to rotation around the peptoid backbone. In a normal MD simulation of a single molecule of DMA solvated in explicit water (simple point charge model) there was no observed *cis-trans* isomerisation. In an equivalent simulation with exchange between two Hamiltonian replicas we observed

rapid exchange between the *cis* and *trans* conformations (Figure 4.10b). The *cis* and *trans* conformations were both well populated (Figure 4.10a) in the HREX simulation with normal distributions of the angles around 0° and 180°, consistent with that obtained from temperature replica exchange simulations of DMA by Voelz *et al.* [19] and for small peptoid structures studied experimentally by Butterfoss *et al.* [4]. The consistency between these results is encouraging as it demonstrates that adequate conformational sampling is achieved using HREX between just two replica systems, at lesser computational expense than other methods where large sets of parallel replicas are required.

### *Sarcosine Dipeptoid*

After tuning the GAFF parameters for the sarcosine dipeptoid to the quantum mechanical energies we ran a series of simulations of sarcosine to check the conformational sampling in vacuum and in explicit solvent using both the original GAFF parameters and the modified G*N*AFF parameters. Implicit (GBSA) water solvation has been found to more accurately reproduce peptoid conformational minima than explicit water solvation [19]. However, the improved performance of implicit solvation relative to explicit may result from the lesser computational expense of the former allowing access to faster conformational sampling and longer time scales within equivalent simulations [43]. For organic solvents implicit solvent methods have been generally found to agree poorly with both experiment and explicit methods, which perform better [44]. Therefore in this work we present a direct comparison of the structural preferences of the peptoids explicitly solvated in water and octanol.

### *Vacuum Phase*

Initially we investigated the vacuum conformations of the sarcosine dipeptoid with both GAFF and G*N*AFF force field parameters. A single disarcosine molecule was simulated over 100 ns using the SD integrator. Figure 4.11 consists of Ramachandran style plots which show the $\phi$ and $\psi$ dihedral sampling with the amide bond in the *trans* conformation (a and b) and the *cis* conformation (c and d) with unmodified GAFF force field parameters (a and c) and the modified G*N*AFF parameters (b and d).

The modification of the force field results in a change in the dihedral preferences, seen in the Ramachandran plots. In both cases, where the amide bond is in the *trans* conformation it remains so throughout the simulation. Likewise, when the amide bond is in the *cis* conformation it remains so. Hence the simulations were run in duplicate to explore the sampling from each amide conformation. There is a clear dependence of $\phi$ and $\psi$ on the preceding $\omega$ that can be seen in the variations between the *cis* and *trans* Ramachandran plots. This coupling between $\omega$ and $\phi/\psi$ is stronger for *cis* $\omega$ than *trans*

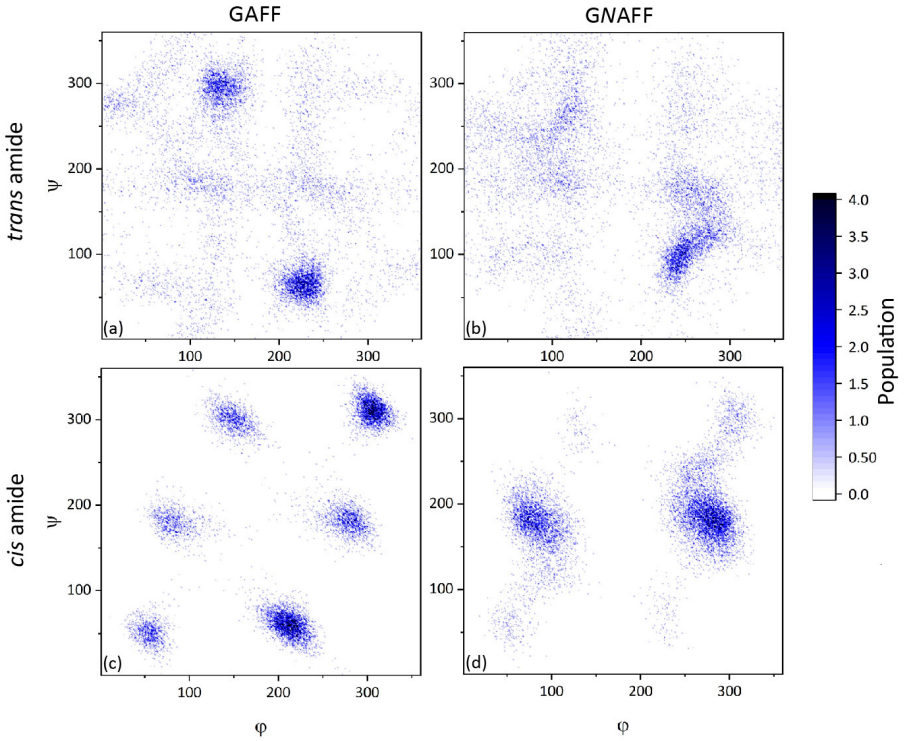



*Fig. 4.11:* Ramachandran style plots showing the sampling of backbone dihedral angles by sarcosine dipeptoid over 100 ns stochastic dynamics simulation with (a) GAFF force field parameters and *trans* amide, (b) G*N*AFF parameters and *trans* amide, (c) GAFF parameters and *cis*amide and (d) G*N*AFF parameters and *cis* amide.

$\omega$. With the amide group in the *trans* conformation $\phi$ and $\psi$ sample a much broader distribution of angles than when the amide is in the *cis* conformation. With GAFF parameters the peptoid favours regions of the Ramachandran plot corresponding to the $C_{7\beta}$ and the $\alpha$-helix conformations, whereas the G*N*AFF parameters enhance the preference for the global minimum $\alpha_D$ conformation, while additionally sampling the *trans*-$C_{7\beta}$ and, to a lesser extent, $\alpha$-helical conformations. The G*N*AFF plots correspond more closely to the B3LYP/6-31+1G(2d,p)// HF/6-31G* energy landscape generated by Butterfoss *et al.* [4] than GAFF, indicating an improvement in the modelling of the peptoid backbone torsions.

### *Solvated Behaviour of the Sarcosine Dipeptoid*

Solvation has been shown to affect the energy landscape of peptoids [19]. We therefore investigated the sampling of different structural conformations by a single sarcosine molecule solvated in explicit simple point charge (SPC) model water. SPC is a 3 point model for water with a tetrahedral geometry which has generally be found to be computationally efficient and to reproduce the properties of bulk water well [45]. The computational efficiency of SPC relative to other water models is advantageous for studying large systems such as those containing longer peptoid sequences and many solvent atoms. The Ramachandran

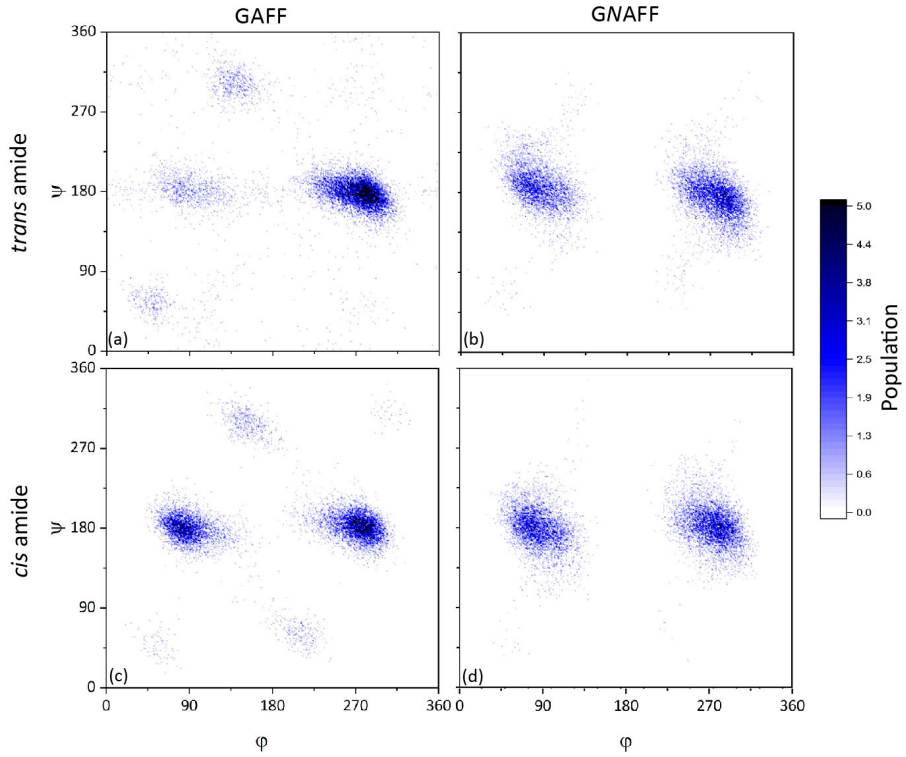



*Fig. 4.12:* Ramachandran plots showing the dihedral sampling of sarcosine dipeptoid solvated in SPC water by normal MD. Figure 4.11 consists of Ramachandran plots which show the $\phi$ and $\psi$ dihedral sampling with the amide bond in the *trans* conformation (a and b) and the *cis* conformation (c and d) with unmodified GAFF force field parameters (a and c) and the modified G*N*AFF parameters (b and d).

plots showing the distributions of the dihedral angles over the course of the solvated sarcosine simulations using GAFF and G*N*AFF parameters are shown in Figure 4.12. As with the SD simulations discussed in the previous section, two identical production runs were performed with the amide in the *cis* conformation for one and the *trans* conformation for the other.

Solvation clearly stabilises the $\alpha_D$ minima in sarcosine, particularly for G*N*AFF, as the highest density of angles sampled are collected in the regions of ($\phi$,$\psi$) = ($\pm$ 75° , $\pm$ 175°). The runs with the GAFF parameters indicate that the unmodified version of the force field imposes a preference for negative $\phi$. This is somewhat contradictory to previous QM studies which reveal the energy landscape for small peptoids with methyl side chain groups to be symmetric [4, 6, 19]. The favouring of angles corresponding to the $\alpha_D$ minima is consistent with the clustering of experimentally determined angles in this region reported by Butterfoss *et al.* [4]. The *cis* and *trans* plots for G*N*AFF are reasonably similar, though the minima outside of $\alpha_D$ are sampled more frequently when the amide is in the *trans* conformation.

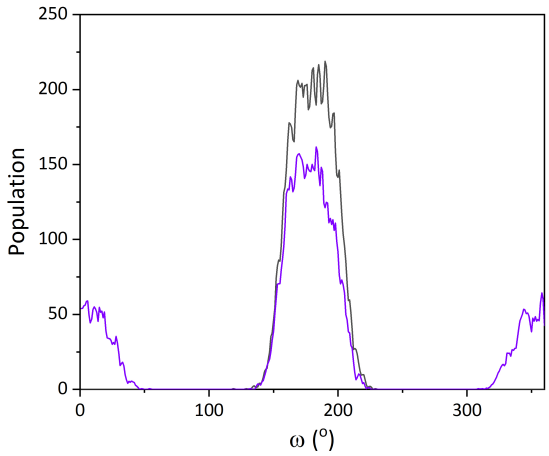



*Fig. 4.13:* Distribution of $\omega$ angle in the sarcosine dipeptoid solvated in SPC water during 100 ns MD and HREX simulation. Simulations performed using G*N*AFF parameters. Normal MD distribution shown in grey and HREX distribution shown in purple.

### *HREX Simulations of Disarcosine in Water*

The barrier between the *cis* and *trans* amide conformations in the sarcosine dipeptoid is above 70 kJ mol$^{-1}$, which is high enough to prohibit crossing in a normal MD simulation. Indeed we see no *cis/trans* isomerism occurring in the SD and normal MD simulations with either the GAFF or G*N*AFF force field parameters. We therefore carried out HREX simulations with 2 replicas (as described in Chapter 2) and G*N*AFF parameters to improve the sampling and remove the need for multiple starting conformations that would be prohibitive to simulation of longer sequences. Using HREX we observe *cis/trans* isomerisation in the sarcosine dipeptoid solvated in water.

The distributions of $\omega$ over the course of two 100 ns simulations are shown in Figure 4.13, with the normal MD distribution (from the *trans* conformation) shown in grey and the HREX distribution in purple. This shows that $\omega$ is distributed only around 180° for the normal MD but is distributed around 180° and 0° for the HREX simulation, with the amide adopting the *cis* conformation approximately 28% of the time. In an experimental study Sui *et al.* established *via* NMR spectroscopy that in small N-acetylated peptoid monomers the amide bond populated the *cis* conformation 28.6% of the time [46], consistent with the results we obtained here. This ratio of *cis/trans* is also close to that which has been observed in temperature REMD simulations [16, 19].

The Ramachandran style $\phi/\psi$ plots from the HREX trajectory for the standard G*N*AFF parameters is shown in Figure 4.14a and, as expected, closely resembles the combined plot from the normal MD simulations with *cis* and *trans* amide starting conformations. Additional insight into the relationship between the amide bond conformation and the proceeding $\phi$ and $\psi$ angles can be gained by plotting Ramachandran style plots of $\omega$ against

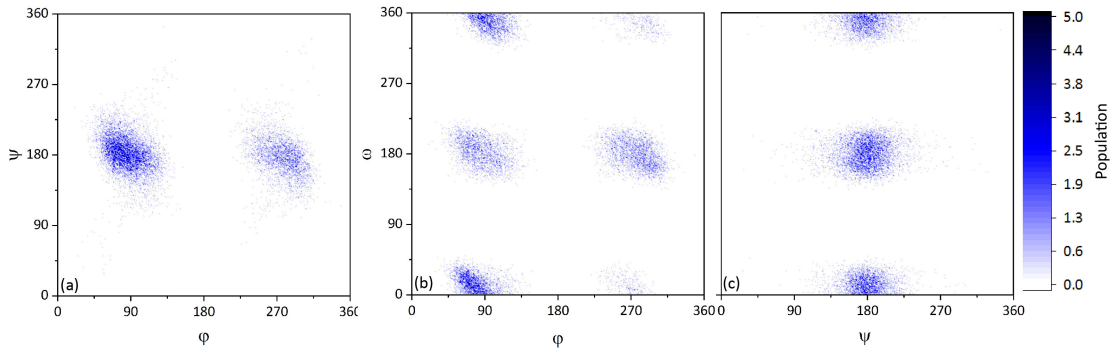



*Fig. 4.14:* (a) Combined *cis/trans* Ramachandran plot for disarcosine in water simulated with G*N*AFF parameters using HREX over 100 ns. (b) $\phi$-$\omega$ plot showing distribution of $\omega$ and $\phi$ angles from the same simulation. (c) $\psi$-$\omega$ plot showing distribution of $\omega$ and $\phi$ angles from the same simulation.

both $\phi$ and $\psi$, which are shown in Figure 4.14b and Figure 4.14c. In the plot of $\omega$ against $\phi$ there are 4 distinct populated regions corresponding to the *cis* and *trans* $\alpha_D$ and $C_{7\beta}$ minima. Contrastingly, in the $\omega$ against $\psi$ plot there are only 2 strongly populated regions, indicating that for both the *cis* and *trans* amide conformation $\psi$ is restricted to be distributed around 180° when the peptoid is solvated in explicit water.

### *4.3.5 Structural Characterisation of Repeat Motif Peptoids in Water*

The aim of modifying GAFF to better capture peptoid behaviour was to be able to simulate longer peptoid sequences, such as those in our repeat motif library, and to identify structural differences between these sequences in different solvent conditions. As such, we created structure and topology files for each of the repeat motif peptoids and ran MD simulations in both water and octanol.

One consideration for simulating longer peptoid sequences, such as the 12-mers in our repeat motif library, was the starting conformation of the peptoid. Normally a crystal structure would be available as a starting point to simulate a protein or peptide but this is not the case for peptoids. *De novo* structure prediction from an extended conformation can be expensive and remains difficult to achieve. Each repeat motif sequence contains a minimum of 50% *N*spe residues and the *N*spe helix has been well characterised in several independent studies [9, 20]. Additionally, the CD spectra we obtained for these sequences (Chapter 3) in many ways resemble those reported in the literature for pure *N*spe helices. Therefore we constructed starting conformations for the repeat motif peptoids based on the average backbone dihedrals of the *N*spe helix [9], following the protocol summarised in Figure 4.5. Topology and coordinate files for sequences RM1-RM6 were constructed in this way.

*Repeat Motif Peptoids Solvated in SPC water*

A single molecule of each peptoid sequence was placed in a periodic box and solvated with SPC water. Sodium and chloride ions were added at 0.15 M concentration. Due to the presence of the cationic peptoid an excess of the negative chloride ions was required to neutralise the system. Following energy minimisation, each system was equilibrated and simulated using HREX for a period of 200 ns with a Nose-Hoover thermostat and Parinello-Rahman barostat. The backbone dihedral sampling during the second half (100 ns) of the simulation was analysed and Ramachandran plots showing the $\phi/\psi$ distributions in each peptoid sequence are shown in Figure 4.15.

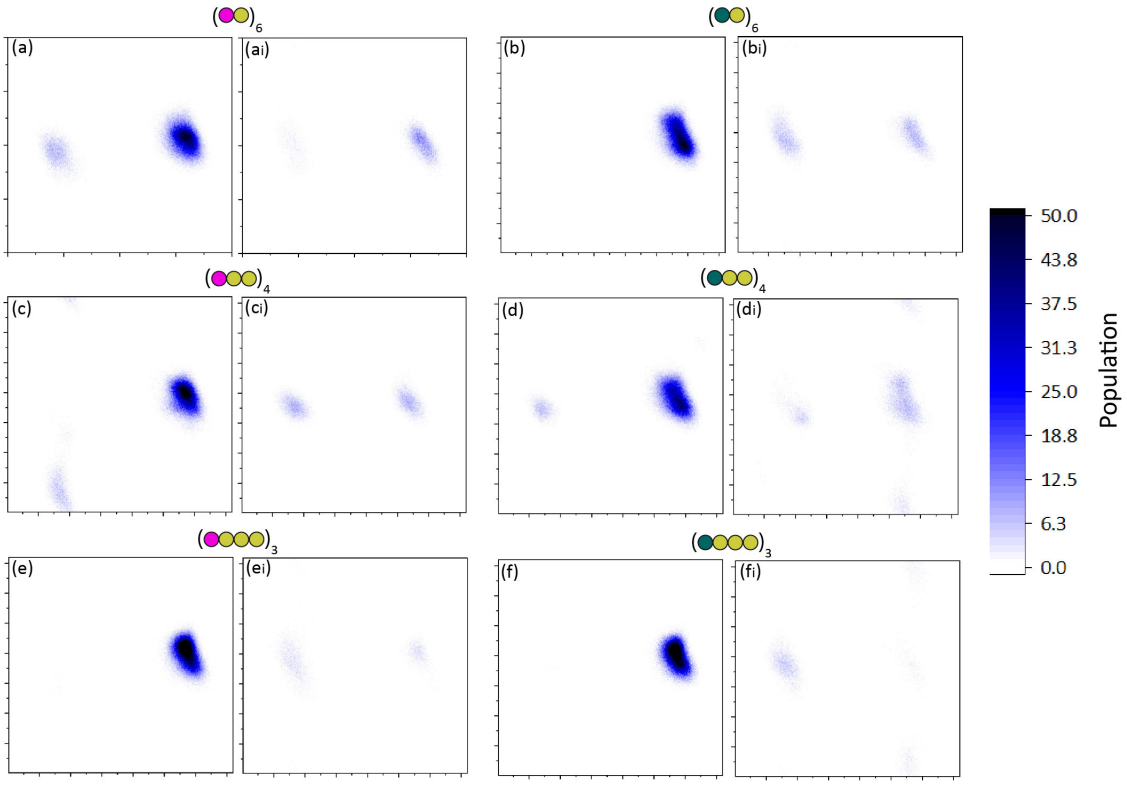



*Fig. 4.15:* Standard $\phi$-$\psi$ Ramachandran plots for RM1-RM6 solvated in SPC water. Data taken from 100 ns HREX simulation. Axes are $\phi$ and $\psi$ in x and y respectively, over the range of angles 0°- 360°. Plots for peptoids RM1-RM6 labelled a-f with the *cis* plot labelled a-f and the *trans* plot labelled ai-fi with ball graphics to represent the peptoid sequence corresponding to each plot (Magenta represents *N*ae, teal represents *N*Lys and yellow represents *N*spe). Plots are shown over the range of angles 0°-360°.

For each peptoid the most populated conformation is that of the *cis* amide $\alpha_D$ helix. As the number of *N*spe residues in the sequence increases the population of this region increases, indicating that this conformation is stabilised by the *N*spe residues. In particular, the *trans* Ramachandran plot is barely populated for the motif 3 sequences, but is more so for the motif 1 and 2 sequences, indicating greater favouring of helical dihedral angles as the proportion of *N*spe residues in the sequence increases. The preference for a right handed structure is clear, as the plots indicate a strong favouring of $\phi$ in the region

of 285° (which is equivalent to -75° ).

The similarity in the Ramachandran plots for these peptoids is unsurprising, given the similarities in their CD spectra and the lack of diversity in the dihedral angles of structures presented within the literature. In Chapter 3 we observed that the CD spectra of RM1 and RM2 somewhat resemble that of the unique threaded loop conformation described by Huang *et al.* [47]. The NMR determined $\phi$ and $\psi$ angles associated with the threaded loop are ±70° and 180° and therefore distinctions between this structure and the *N*spe helix are not clear on a classic Ramachandran plot, perhaps accounting for the similarities observed in the plots in Figure 4.15, despite the different characters of the CD spectra.

### *Cis/Trans Isomerisation of Backbone*

While the Ramachandran plots for each of the peptoids in water are very similar, different rates of *cis/trans* isomerism were observed in the different sequences. Additionally, individual residues within each sequence had very different *cis/trans* distributions. This leads to significant disruption to the overall helical structure of the peptoids despite the distribution of backbone dihedrals around the $\alpha_D$ minimum. Simulation snapshots of RM1 in water are shown in Figure 4.16.

The starting conformation for RM1 is shown in Figure 4.16a. All of the repeat motif sequences were minimised and equilibrated from this conformation prior to the HREX production runs. Within the simulation snapshots for RM1 in water, which were randomly selected from the trajectory, it appears that the structure is considerably dynamic and samples many different macroscopic conformations despite the fact that $\phi$ and $\psi$ generally remain distributed around ±75° and 180° for the duration of the simulation. Determination of the average backbone structure over the course of the production run was attempted using the GROMACS gmx covar program, which generates a mass weighted covariance matrix. Only heavy atoms were considered due to the inherent flexibility of the hydrogens. However, the backbone itself was sufficiently flexible that the analysis did not produce a single defined average structure. The motif 2 and 3 peptoids also deviated from a completely helical structure over the course of the simulations, though to a lesser extent as the proportion of *N*spe residues increased. This is consistent with our experimental findings in Chapter 3 in which the CD spectra for sequences with a high proportion of *N*spe residues had more intense helical bands than those with a lower proportion of *N*spe residues.

In Chapter 3 we also observed that the CD spectra of the *N*Lys sequences were more intense than their *N*ae counterparts, indicating greater helical character. The ratio of *cis:trans* amides in equivalent sequences with *N*ae and *N*Lys residues was similar, though for each sequence the *N*Lys variant had slightly greater sampling of *trans* amide conformations than the *N*ae variant, indicating a slightly greater deviation from the perfect helical

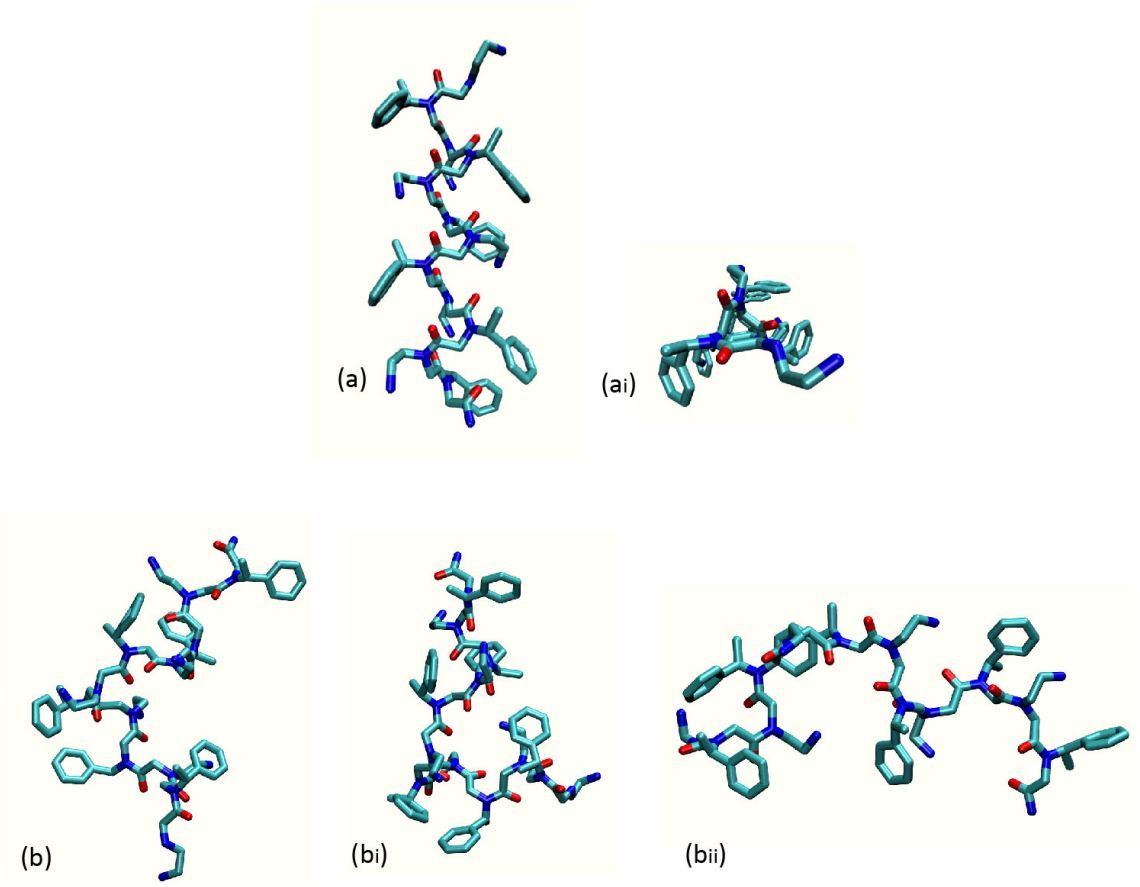



*Fig. 4.16:* Simulation snapshots showing: (a) Starting conformation of RM1 as ideal *N*spe helix, side view. (ai) Starting conformation of RM1 as ideal *N*spe helix, top view. (b), (bi), (bii) Different conformations sampled during the HREX simulation.

structure in the former than the latter. This was unexpected and appears to demonstrate that the helical features observed in the CD spectra may not directly correlate to the peptoid adopting the perfect helical conformation, but other minor conformations may contribute to this.

### *Helical Handedness*

It has been well documented that the *N*spe residues induce a preference for negative $\phi$ angles (or angles larger than 180° on the 0°-360° scale that is commonly used for peptoid Ramachandran plots) that impart a right handed turn to the peptoid helix. GAFF in its native form has been reported to fail to capture this behaviour for peptoids with implicit water models [19]. Our results so far indicate that G*N*AFF does capture the propensity of *N*spe to induce helical handedness as we see a strong preference for negative $\phi$ angles in the Ramachandran plots for the repeat motif peptoids in water. The occurrence of *cis/trans* isomerisation in these simulations indicates that the peptoid is sampling conformations

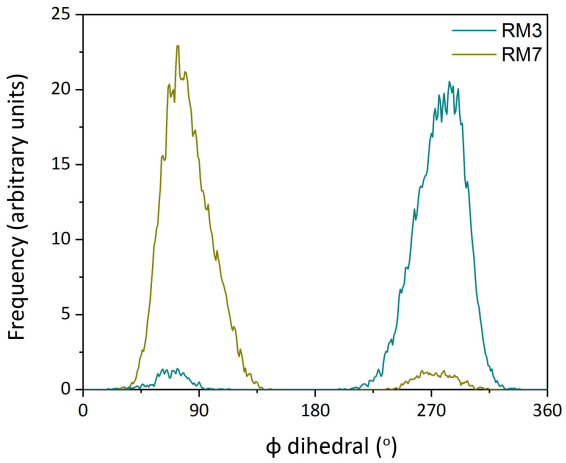



*Fig. 4.17:* Distribution of $\phi$ over 100 ns HREX simulation of $(N\text{ae}N\text{spe}N\text{spe})_4$ (shown in teal) and $(N\text{ae}N\text{rpe}N\text{rpe})_4$ (shown in gold).

effectively and, since the barrier to rotation around $\phi$ is smaller than $\omega$, there should not be any sampling bias towards negative $\phi$, despite this being the starting conformation. Nevertheless we checked the sampling of $\phi$ in RM3 and RM7, which are identical sequences with $N$spe and the opposite handed $N$rpe respectively, from a fully extended starting conformation. The distributions of $\phi$ for both sequences over the course of 100 ns HREX simulation are shown in Figure 4.17.

These angle distributions show that G$N$AFF clearly captures the preference for negative $\phi$ imparted by the $N$spe side chain and the corresponding preference for positive $\phi$ which is induced by the enantiomer, $N$rpe.

### 4.3.6 *Structural Characterisation of Peptoids in Octanol*

#### *Octanol Solvent Box Parameters*

An equilibrated box of octanol was generated with which to solvate peptoid systems. GAFF parameters for a single octanol molecule were generated using Antechamber and ACPYPE in the same way that parameters were generated for the peptoid topologies. A cubic simulation box was populated with octanol molecules and equilibrated, first at high pressure, then at a standard pressure of 1 bar, to allow the octanol to first compress from a low density and then relax to its equilibrium density. To confirm that GAFF models octanol at a reasonable level the bulk liquid properties were investigated. The density of a box of 248 molecules averaged at 823.21 $\pm$ 0.2 kg/m$^3$ over 100 ns standard atmospheric pressure $NPT$ simulation. The heat of vaporization was calculated according to equation 4.12 to be 76.99 kJ/mol. These values are compared to equivalent experimental values from the literature in table 4.4 and indicate that GAFF provides a reasonable model for

octanol.

| Property | GAFF | Experiment | % Difference |
|---|---|---|---|
| Density (kg/m$^3$) | 823.21 | 827 | -0.46 % |
| $\Delta H_{\text{vap}}$ (kJ/mol) | 76.99 | 71.20 | +8.13 % |

*Tab. 4.4:* Table of properties of GAFF octanol model compared to published experimental values [48].

### *Sarcosine Dipeptoid Solvated in Octanol*

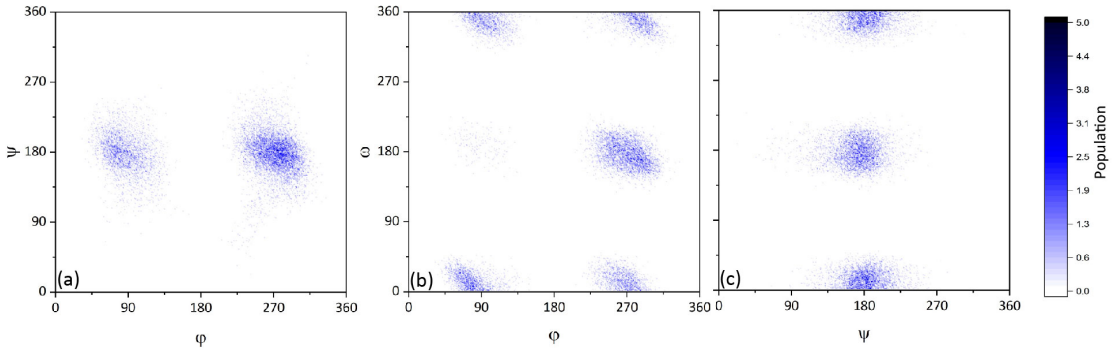



*Fig. 4.18:* (a) Combined *cis/trans* Ramachandran plots showing dihedral sampling by sarcosine dipeptoid in octanol, simulated using HREX over 100 ns. (b) $\phi$-$\omega$ plot showing distribution of amide bond ($\omega$) and $\phi$ angles from the same simulation. (c) $\psi$-$\omega$ plot showing distribution of amide bond ($\omega$) and $\phi$ angles from the same simulation.

To our knowledge there have been no simulations of peptoids in octanol reported in the literature. Therefore we initially investigated the structural preferences of the sarcosine dipeptoid in octanol as a small model system. A single sarcosine dipeptoid molecule was solvated in our explicit octanol model and simulated with HREX, using the parameters previously discussed, to allow *cis/trans* isomerisation. The total simulation time was 200 ns and the data analysed and displayed in the figures is taken from the latter 100 ns of the production run. Figure 4.18 shows the standard, combined $\phi$-$\psi$ Ramachandran plot and the $\phi$-$\omega$ and $\psi$-$\omega$ plots for disarcosine in octanol. The standard Ramachandran plot for the sarcosine dipeptoid in octanol is very similar to that obtained for the sarcosine dipeptoid in water, with the $\alpha_D$ minimum appearing to be stabilised in both solvents. Likewise the $\psi$-$\omega$ plots are very similar for both solvents, with the two distinct populated regions indicating that $\psi$ is restricted to angles distributed around 180° when the amide bond is in both the *cis* and *trans* conformations. The $\phi$-$\omega$ plot shown in Figure 4.18c indicates that solvation in octanol restricts the population of the *trans* $\alpha_D$ minimum far more than solvation in SPC water.

Solvation in octanol also promotes a *cis* amide conformation in the sarcosine dipeptoid. The distribution of $\omega$ over the course of 100 ns HREX simulation is shown in Figure 4.19a.

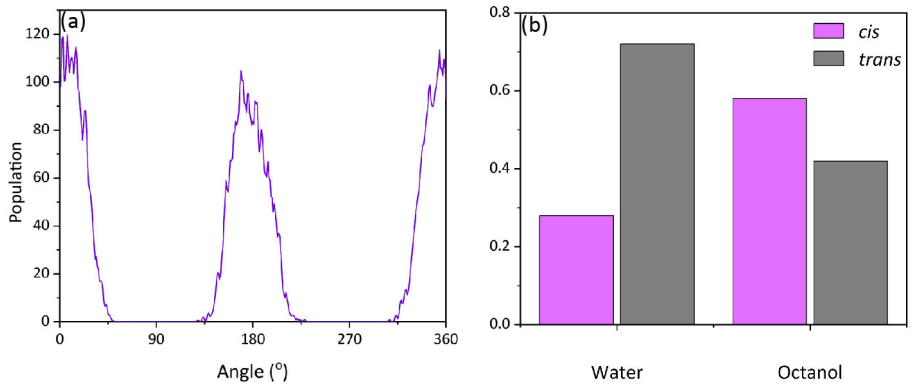



*Fig. 4.19:* (a) Distribution of $\omega$ angles sampled over 100 ns HREX simulation of disarcosine solvated in octanol. (b) Comparison of *cis*/*trans* sampling by disarcosine in water and octanol.

There is a slight preference for the *cis* conformation which is adopted approximately 58 % of the time when the peptoid is solvated in octanol. The *cis*/*trans* populations during the water and octanol simulations are shown in Figure 4.19. Solvation in octanol clearly promotes a *cis* amide conformation relative to solvation in water.

### *Dihedral Sampling of Repeat Motif Sequences in Octanol*

The Ramachandran plots for RM1-6 solvated in octanol are shown in Figure 4.20. The data shown is taken from the second half (100 ns) of a 200 ns HREX production run. In this case only the *cis* conformation plots are shown as *cis-trans* isomerisation occurred only in the C-terminal residues while the majority of the peptoid backbone never sampled the *trans* amide conformation. The solvation in octanol therefore appears to strongly stabilise the *cis* amide conformation in the longer sequences, as it did in the sarcosine dipeptoid. The peptoids deviated very little from the ideal *N*spe helix conformation during these simulations, despite the use of HREX to enhance conformational sampling.

The lack of *cis-trans* isomerisation and distribution of angles around $\phi$,$\psi$ = (-75°, 180°) indicate that the ideal *N*spe helix conformation is more favourable in octanol than water. Each of the 6 sequences has a very similar Ramachandran plot in octanol. The preference for negative $\phi$, induced by the handedness of the *N*spe residues, also appears to be enhanced in octanol relative to water which also contributes to the stabilisation of the *cis* $\alpha_D$ minimum. Viewing the simulation trajectories of the repeat motif peptoids in octanol also shows that the overall structure is considerably less dynamic in octanol than water, particularly due to the lack of *cis-trans* isomerisation.

## *4.3.7 Side-Chain-Backbone Hydrogen Bonding*

Though peptoids are unable to form the intra-backbone hydrogen bond networks known to stabilise secondary structures in peptides and proteins, it remains possible for hydrogen

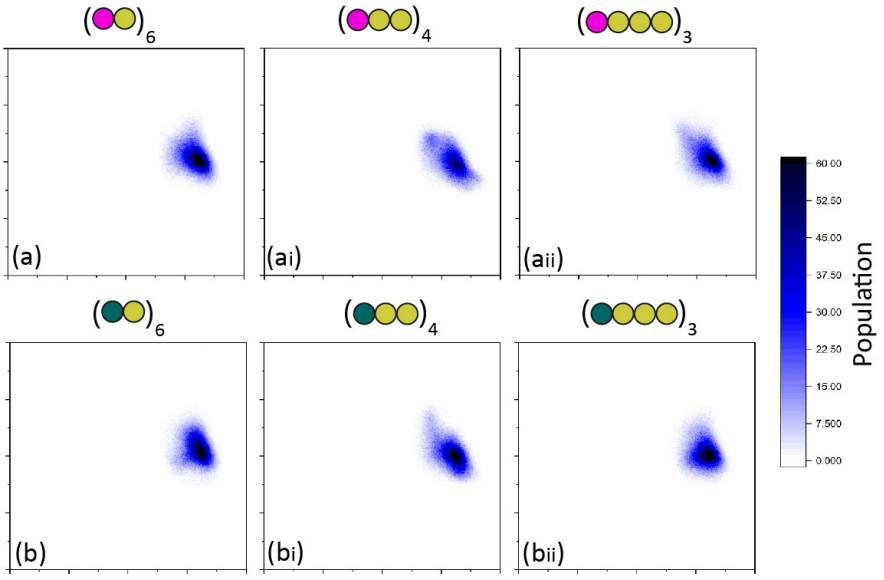



*Fig. 4.20:* Ramachandran plots for repeat motif peptoid sequences in octanol showing dihedral sampling over 100 ns HREX simulation. Axes are $\phi$ and $\psi$ in x and y respectively, over the range of angles 0°- 360°. *Cis* plots only shown due to scarce population of *trans* plots. (a) RM1 (ai) RM3 (aii) RM5 (b) RM2 (bi) RM4 (bii) RM6.

bonds to form between certain peptoid side chains and the backbone [49, 50]. Notably, sequences containing *N*ae residues have been found to form hydrogen bonds where the positively charged ammonia group on the side chain is the donor and the carbonyl oxygen is the acceptor [51].

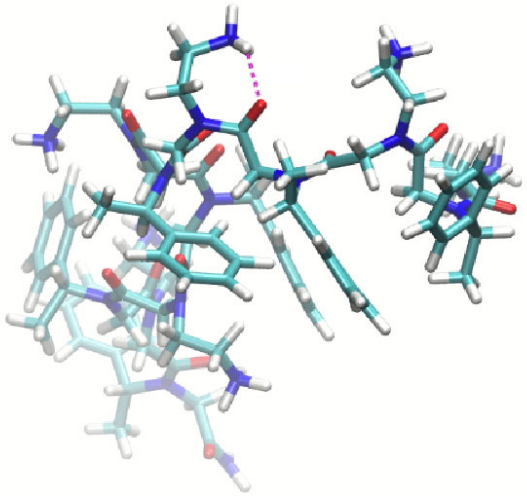


*Fig. 4.21:* Example snapshot of *N*ae side chain oriented to form a hydrogen bond, shown as purple dashed line, with an adjacent carbonyl oxygen.

There is potential for such hydrogen bonds to form in our peptoid sequences containing *N*ae side chains and potentially *N*Lys side chains too, as they also include the hydrogen

bond donating ammonia group. However the increased length of *N*Lys relative to *N*ae may make hydrogen bond formation less likely due to steric effects. We investigated the formation of hydrogen bonds within the repeat motif peptoid HREX trajectories by using the GROMACS program gmx hbond to calculate the average number of hydrogen bonds in each simulation snapshot. The calculations were performed with a cut off radius of 0.35 nm and cut off angle of 30°. An example snapshot with a *N*ae side chain oriented to hydrogen bond with a neighbouring carbonyl oxygen is shown in Figure 4.21. The average numbers of intra-peptoid hydrogen bonds and hydrogen bonds between the peptoid and surrounding solvent (for both water and octanol) are shown in Table 4.5.

| | Number of H-bonds | | | |
|---|---|---|---|---|
| Peptoid | Water | | Octanol | |
| | Peptoid-Peptoid | Peptoid-Solvent | Peptoid-Peptoid | Peptoid-Solvent |
| RM1 | 1.299 | 32.698 | 1.985 | 24.268 |
| RM2 | 0.596 | 35.883 | 0.173 | 23.774 |
| RM3 | 0.925 | 28.055 | 3.775 | 18.244 |
| RM4 | 0.293 | 30.907 | 0.384 | 20.596 |
| RM5 | 0.587 | 26.866 | 2.328 | 17.809 |
| RM6 | 0.265 | 28.155 | 0.120 | 18.687 |

*Tab. 4.5:* Average number of hydrogen bonds per trajectory snapshot from HREX simulations in water and octanol.

In water, sequences with *N*ae side chains appear to form approximately twice as many intra-peptoid hydrogen bonds as their *N*Lys counterparts. The *N*Lys sequences on the other hand form a larger number of hydrogen bonds between the peptoid and surrounding solvent. The average number of hydrogen bonds per snapshot also decreased, as might be expected, as the number of cationic residues decreased, going from motif 1 to 3. This may contribute to the observed differences in CD spectra between sequences that are discussed in Chapter 3. In Chapter 3 we theorised that the *N*Lys sequences are more characteristically helical than their *N*ae counterparts due to the reduced length of *N*ae resulting steric effects that induce strain in the backbone and increase the entropic cost of folding. The arrangement of *N*ae side chains in a position to hydrogen bond with neighbouring carbonyl oxygen atoms would serve to exacerbate this.

We also observed in Chapter 3 that among the scrambled sequences, the *N*Lys versions were considerably more diverse in spectral shape and intensity in PBS than the *N*ae versions. Though we have not simulated the scrambled sequences, it could be expected that they would form a similar number of intra-peptoid hydrogen bonds as the motif 2 sequences, being structurally analogous, and *N*ae sequences would therefore be conformationally restricted while *N*Lys sequences could fold more freely to maximise shielding of the hydrophobic *N*spe residues and expose the cationic residues to the polar solvent.

Roe *et al.* observed through NMR and MD, that the presence of *N*ae residues in peptoid tetramers in aqueous solution promoted a *trans-cis-trans* backbone, resulting in a dominant conformation rather than the expected flexibility [51]. The hydrogen bonding between *N*ae and the backbone could similarly promote a dominant conformation in the scrambled sequences to a greater extent than *N*Lys. In the CD melts presented in Chapter 3 the *N*ae scrambled sequences exhibited greater thermal unfolding than their *N*Lys counterparts, with loss of spectral intensity in the region of 17-23% and 10-19% respectively. This perhaps indicates that the helical structure of the *N*ae sequences is stabilised to a very small extent more by hydrogen bonds than that of the *N*Lys sequences, as these would be more sensitive to changes in temperature than sequences stabilised by sterics alone.

In octanol the *N*ae sequences form considerably more peptoid-peptoid hydrogen bonds than in water, whereas the *N*Lys sequences form fewer (with the exception of RM4). There are fewer available hydrogen bond accepting sites in octanol than water and therefore competition between peptoid-peptoid and peptoid-solvent bonds will be less in the former. In the *N*ae sequences this appears to promote peptoid-peptoid bonds whereas in the *N*Lys sequences it does not, perhaps due to the additional steric restrictions of the extra methylene groups in the latter side chain. All of the sequences have dramatically enhanced helical characteristics in the octanol CD spectra than PBS and the *N*ae sequences are more intense than their *N*Lys counterparts in octanol. The Ramachandran plots also indicate that all of the sequences deviate from the perfect *N*spe dihedrals considerably less in octanol than water. It therefore seems that the number of intra-peptoid hydrogen bonds does not strongly influence the overall structure of the peptoids, which appears to be determined by a balance of a number of factors including the proportion and distribution of *N*spe, as well as the distribution of charged residues.

### *4.3.8 Structure and Behaviour of Peptoids in Membranes*

We selected octanol as a solvent to study the peptoid secondary structural preferences due to its widespread use as a model for the interior environment of biological membranes and additionally because we used octanol as a solvent in our experimental investigation of the structures of the repeat motif sequences in Chapter 3. Computational techniques also afford us the ability to study peptoids inserted into model lipid membranes. Generally, simulations of lipid bilayers are computationally expensive as a large number of atoms are required to model a bilayer properly and the time scales associated with peptide/peptoid insertion into the membrane are long. Therefore coarse grained force fields are frequently used for membrane systems. The MARTINI coarse grained force fields that were originally published in 2007 by Marrink *et al.* have become a popular choice for the simulation of

biomolecular systems, with computation times reduced by 2-3 orders of magnitude relative to atomistic alternatives. Notably they have recently been used to shown pore formation mechanisms by antimicrobial peptides in bilayers of various lipid compositions, revealing that charge has an important role in moderating the peptide-lipid interactions [52]. It would therefore be useful to develop an accurate coarse grained peptoid force field suitable for similar investigations.

### *MARTINI Coarse Grained Peptoid Force Field*

Work has been carried out in conjunction with that presented in this thesis to develop a coarse grained force field for peptoids that is compatible with the MARTINI force field [53, 54] for lipids. This work was carried out by Ewan P. South under the supervision of Professor Mark R. Wilson and with additional guidance in coarse graining methods from Dr Thomas Potter and Dr Martin Walker. A coarse grained model for RM3, shown in Figure 4.22, was developed by following a 4:1 mapping scheme of atoms into coarse grained beads and using a combination of top-down and bottom-up methods, consistent with those used to develop the original MARTINI [55]. The coarse grained topology and co-ordinate files for RM3 were created using an adapted version of the martinize.py python script [56], using the energy minimised *N*spe helix atomistic structure from section 4.3.5 as a reference. Martinize builds coarse grained models according to reference structures found in the Protein Databank and was adapted by Ewan South to optimise the representation of the atomistic peptoid structure in the output while also reproducing experimental free energies of partitioning between water and octanol.

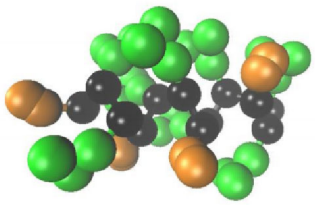


*Fig. 4.22:* Coarse grained representation of the structure of RM3 developed by Ewan South. Backbone beads are shown in black, *N*ae side chain beads shown in orange and *N*spe side chain beads shown in green. Figure by Ewan P. South.

In antimicrobial mechanism studies the starting position of peptides in lipid bilayer simulations has been found to be important. Several studies have shown that MARTINI appears to over-stabilise peptides bound to the head groups at the bilayer-water interface. Peptides initially placed at this interface do not permeate into the membrane interior, even over very long time scales (of the order of hundreds of microseconds) [52, 57]. In the studies referenced, peptides initially placed in a trans-membrane position displayed membrane disruptive behaviour including pore formation. The initial orientation of the peptide within the membrane appears to be arbitrary. Given this information and the

investigations of Catte *et al.* into optimised lipid:peptoid ratios for simulation of pore forming events [52], systems were set up with peptoids evenly distributed in transmembrane positions throughout model bilayers composed of pure DPPC and a lipid mixture based on the membrane of gram-negative bacteria *Helicobacter Pylori* (a combination of DPPC, POPG and cardiolipin 2), at a lipid:peptoid ratio of 18:1. These simulations revealed that the peptoids migrate to the bilayer surface and aggregate, forming disordered toroidal pores, such as those shown in Figure 4.23, stabilised by the surface aggregate.

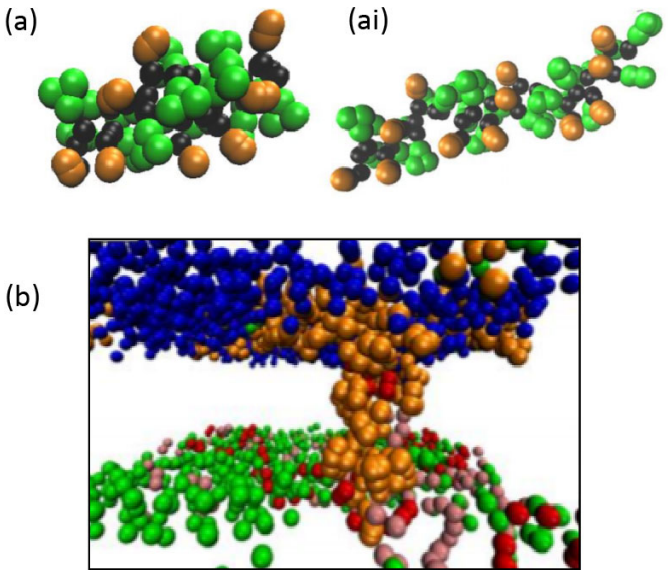



*Fig. 4.23:* (a) Structure of 2 peptoid membrane aggregate formed by coarse grained model of RM3 in DPPC and *H.pylori* Martini bilayers. (ai) Three peptoid aggregate. (b) Cross section of *H.pylori* bilayer showing disordered toroidal pore formed by RM3. Peptoids shown in orange, lipids in blue, green, pink and red. Water and ions omitted for clarity. Figures by Ewan P. South.

The 2 and 3 peptoid aggregates shown in Figure 4.23a and 4.23ai respectively were frequently observed forming at the bilayer surface during simulations, with the dimeric aggregate appearing particularly stable. Upon the migration of further peptoids into the aggregates (up to 11 in total) the aggregates were observed sinking into the bilayer interior and forming transient pores, an example of which is shown in Figure 4.23b. The presence of lipid head groups in the pore channel in the bilayer interior is consistent with reports of disordered toroidal pores [58, 59].

The model bilayer for *H.pylori* used is this work is considerably more complex than the model membranes used in the experimental investigation of peptoid-lipid interactions described subsequently in Chapter 5. There is great potential to continue the work described here to investigate the interactions between the RM peptoids and simpler model membranes in order to directly compare the experimental results with computation investigations, to elucidate the mechanism of action of these peptoids.

## 4.4 Conclusions

### *4.4.1 Summary*

In this chapter we described modifications to GAFF to describe the peptoid backbone and subsquent study of the structural preferences of the repeat motif peptoids in different solvent conditions. The key conclusions from the work described in this chapter are summarised in the bullet points below.

- The native GAFF force field captures the QM torsional minima of peptoids reasonably well but also features peptide $\alpha$-helical minima not seen in the QM profiles.
- Our modified peptoid force field, G*N*AFF, captures the minima in the QM rotational energy profiles accurately.
- Solvation of the sarcosine dipeptoid in both SPC water and octanol stabilises the $\alpha_D$ minimum conformation.
- Solvation in octanol increases the preference for the *cis* amide conformation in the sarcosine dipeptoid, relative to water.
- Each of the RM peptoids adopts backbone dihedral angles distributed around $\phi$,$\psi$ = (-75°, ±180°) in water and octanol, consistent with those reported in the literature for pure *N*spe helices.
- G*N*AFF captures the propensity of *N*spe residues to induce right handed helical structure via a preference for negative $\phi$ angles.
- The global helical structure of the RM peptoids in water is disrupted by *cis/trans* isomerisation of certain residues, resulting in highly dynamic structures.
- *Cis-trans* isomerisation occurs only in the C-terminal residue of RM peptoids solvated in octanol.
- Hydrogen bonds can form between the cationic side chain ammonia groups and the backbone carbonyl atoms and do so to a greater extent in *N*ae sequences than in *N*Lys sequences, in both water and octanol.

### *4.4.2 Outlook*

The work we have done here has shown that we can capture the behaviours of the peptoid backbone, previously observed in QM studies, with our modified all atom force field, G*N*AFF. We modified GAFF by fitting the rotational energy profiles to QM equivalents

which appears to improve the modelling of the peptoid backbone torsions in small model systems.

Generally the results presented here and within the literature suggest that peptoids with a diverse range of side chains and structural motifs appear to share common backbone structural preferences. This is consistent with our experimental results presented in Chapter 3, where the peptoids exhibit CD spectra with broadly similar features. This is also what we observe in the Ramachandran plots obtained from our simulations of the repeat motif peptoids in water and octanol. In both solvents the backbone dihedrals are distributed around the angles associated with the pure *N*spe helix, which has been identified in previous studies and corresponds to the $\alpha_D$ minimum. When solvated in octanol the overall structure of our peptoid sequences deviate much less from this structure than in water. In water each peptoid sequence appears to have a dynamically changing structure and additionally there is more structural variation between different sequences in water than octanol. The structures of RM peptoids in water appear to be increasingly dynamic as the proportion of *N*spe residues is reduced, corroborating our experimental observation that the RM1 sequences have CD spectra in PBS that are less characteristically helical than the RM2 and RM3 sequences.

Peptoid sequences are not stabilised by an extensive intra-backbone hydrogen bond network in the manner of peptide helices. Nonetheless the observation of side chain-backbone hydrogen bonds in our simulations and the discrepancies between different sequences could to some extent explain our observations in Chapter 3, where *N*ae sequences were found to be more structurally sensitive to temperature than their *N*Lys counterparts. The effect of side chain-backbone hydrogen bonding may also contribute to the differences observed in the CD spectra for each of the sequences. It is difficult to make distinctions between the conformations favoured by different peptoid sequences simply by inspecting the Ramachandran plot. Further work is required to fully identify the minimum energy conformations associated with each sequence in water due to their dynamic nature in solution. This could be achieved by performing QM energy minimisations of structures sampled in the literature and identifying the minimum conformation sampled. However, due to the size of the sequences in question, this would be extremely computationally expensive.

A useful avenue of research to pursue in future would be the calculation of CD spectra from the conformations sampled during an MD trajectory. This would allow a direct comparison between the simulated peptoids and the experimentally obtained CD spectra which would shed light on both the accuracy of the force field and the sampling using the HREX method that we have employed in this work. This could also increase our understanding of the electronic level structural details which correspond to the observation of the characteristic bands observed in the CD spectra of peptoids. Such calculations have previously been performed for peptides, though generally with only qualitative agreement

found between calculated and experimental spectra, indicating that further refinement of the methods are required.

Coarse graining methods have been successfully applied to peptoids to produce force field parameters compatible with the simulation of lipid bilayers using the MARTINI model. This allows the exploration of the membrane-disruptive antimicrobial activity of these molecules, which we probe using experimental techniques in the next chapter.

# References: Chapter 4


[1] L. J. Weiser, E. E. Santiso, "Molecular modeling studies of peptoid polymers", *AIMS Materials Science* **2017**, *4*, 1029–1051.

[2] K. Moehle, H.-J. Hofmann, "Peptides and Peptoids-A Quantum Chemical Structure Comparison", *Biopolymers* **1996**, *38*, 781–790.

[3] C. Baldauf, R. Günther, H. J. Hofmann, "Helices in peptoids of $\alpha$- and $\beta$-peptides", *Physical Biology* **2006**, *3*.

[4] G. L. Butterfoss, P. D. Renfrew, B. Kuhlman, K. Kirshenbaum, R. Bonneau, "A preliminary survey of the peptoid folding landscape", *Journal of the American Chemical Society* **2009**, *131*, 16798–16807.

[5] R. A. Friesner, "Ab initio quantum chemistry: Methodology and applications", *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 6648–6653.

[6] D. T. Mirijanian, R. V. Mannige, R. N. Zuckermann, S. Whitelam, "Development and use of an atomistic CHARMM-based forcefield for peptoid simulation", *Journal of Computational Chemistry* **2014**, *35*, 360–370.

[7] K. Kirshenbaum, A. E. Barron, R. A. Goldsmith, P. Armand, E. K. Bradley, K. T. V. Truong, K. A. Dill, F. E. Cohen, R. N. Zuckermann, "Sequence-specific polypeptoids: a diverse family of heteropolymers with stable secondary structure", *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 4303–4308.

[8] C. W. Wu, T. J. Sanborn, R. N. Zuckermann, A. E. Barron, "Peptoid oligomers with alpha-chiral, aromatic side chains: Effects of chain length on secondary structure", *Journal of the American Chemical Society* **2001**, *123*, 2958–2963.

[9] P. Armand, K. Kirshenbaum, A. Falicov, R. L. Dunbrack, K. A. Dill, R. N. Zuckermann, F. E. Cohen, "Chiral N-substituted glycines can form stable helical conformations", *Folding and Design* **1997**, *2*, 369–375.

[10] R. Bonneau, J. Tsai, I. Ruczinski, D. Chivian, C. Rohl, C. E. M. Strauss, D. Baker, "Rosetta in CASP4: Progress in ab initio protein structure prediction", *Proteins: Structure Function and Genetics* **2001**, *45*, 119–126.

[11] K. Drew, P. D. Renfrew, T. W. Craven, G. L. Butterfoss, F. C. Chou, S. Lyskov, B. N. Bullock, A. Watkins, J. W. Labonte, M. Pacella, K. P. Kilambi, A. Leaver-Fay, B. Kuhlman, J. J. Gray, P. Bradley, K. Kirshenbaum, P. S. Arora, R. Das, R. Bonneau, "Adding Diverse Noncanonical Backbones to Rosetta: Enabling Peptidomimetic Design", *PLoS ONE* **2013**, *8*.

[12] P. D. Renfrew, T. W. Craven, G. L. Butterfoss, K. Kirshenbaum, R. Bonneau, "A rotamer library to enable modeling and design of peptoid foldamers", *Journal of the American Chemical Society* **2014**, *136*, 8772–8782.

[13] A. D. Mackerell, M. Feig, C. L. Brooks, "Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulation", *Journal of Computational Chemistry* **2004**, *25*, 1400–1415.

[14] K. Lindorff-Larsen, P. Maragakis, S. Piana, M. P. Eastwood, R. O. Dror, D. E. Shaw, "Systematic validation of protein force fields against experimental data", *PLoS ONE* **2012**, *7*, 1–6.

[15] P. S. Nerenberg, T. Head-Gordon, "New developments in force fields for biomolecular simulations", *Current Opinion in Structural Biology* **2018**, *49*, 129–138.

[16] S. H. Park, I. Szleifer, "Structural and Dynamical Characteristics of Peptoid Oligomers with Achiral Aliphatic Side Chains studied by Molecular Dynamics Simulation", *Journal of Physical Chemistry* **2011**, *15*, 10967–10975.

[17] H. M. Shin, C. M. Kang, M. H. Yoon, J. Seo, "Peptoid helicity modulation: precise control of peptoid secondary structures via position-specific placement of chiral monomers", *Chemical Communications* **2014**, *50*, 4465–4468.

[18] R. V. Mannige, T. K. Haxton, C. Proulx, E. J. Robertson, A. Battigelli, G. L. Butterfoss, R. N. Zuckermann, S. Whitelam, "Peptoid nanosheets exhibit a new secondary-structure motif", *Nature* **2015**, *526*, 415–420.

[19] V. A. Voelz, K. A. Dill, I. Chorny, "Peptoid conformational free energy landscapes from implicit-solvent molecular simulations in AMBER", *Biopolymers* **2011**, *96*, 639–650.

[20] S. Mukherjee, G. Zhou, C. Michel, V. A. Voelz, "Insights into Peptoid Helix Folding Cooperativity from an Improved Backbone Potential", *Journal of Physical Chemistry B* **2015**, *119*, 15407–15417.

[21] M. Feigel, "Rotation barriers of amides in the gas phase", *Journal of Physical Chemistry* **1983**, *87*, 3054–3058.

[22] E. M. Duffy, D. L. Severance, W. L. Jorgensen, "Solvent Effects on the Barrier to Isomerization for a Tertiary Amide from ab Initio and Monte Carlo Calculations", *Journal of the American Chemical Society* **1992**, *114*, 7535–7542.

[23] K. Nguyen, M. Iskandar, D. L. Rabenstein, "Kinetics and equilibria of cis/trans isomerization of secondary amide peptide bonds in linear and cyclic peptides", *Journal of Physical Chemistry B* **2010**, *114*, 3387–3392.

[24] J. Wang, R. Wolf, J. Caldwell, P. A. Kollman, "Development and Testing of a General AMBER Force Field", *Journal of Computational Chemistry* **2004**, *25*, 1157–1174.

[25] A. E. Tonelli, "On the Stability of Cis and Trans Amide Bond Conformations in Polypeptides", *Journal of the American Chemical Society* **1971**, *93*, 7153–7155.

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

[27] ProBuilder Online.

[28] M. Hanwell, D. Lonie, T. Vandermeersch, E. Zurek, G. Hutchison, D. E. Curtis, "Avogadro: An Advanced Semantic Chemical Editor, Visualisation, and Analysis Platform", *Journal of Chemical Informatics* **2012**, *4.17*, 1–17.

[29] J. Wang, W. Wang, P. A. Kollman, D. A. Case, "Automatic atom type and bond type perception in molecular mechanical calculations", *Journal of Molecular Graphics and Modelling* **2006**, *25*, 247–260.

[30] A. W. Sousa da Silva, W. W. F. Vranken, J. Wang, W. Wang, P. Kollman, D. Case, V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, H. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. Bhat, H. Weissig, I. Shindyalov, P. Bourne, J. Klepeis, K. Lindorff-Larsen, R. Dror, D. Shaw, A. Brunger, C. Schwieters, J. Kuszewski, N. Tjandra, G. Clore, C. Schwieters, J. Kuszewski, G. Clore, G. Kleywegt, K. Henrick, E. Dodson, D. van Aalten, A. Schüttelkopf, D. van Aalten, A. Ribeiro, B. Horta, R. de Alencastro, Y. Duan, C. Wu, S. Chowdhury, M. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, P. Kollman, W. Jorgensen, D. Maxwell, J. Tirado-Rives, A. Malde, L. Zuo, M. Breeze, M. Stroet, D. Poger, P. Nair, C. Oostenbrink, A. Mark, C. Oostenbrink, A. Villa, A. Mark, W. V. Gunsteren, J. Ponder, D. Case, J. Wang, R. Wolf, J. Caldwell, P. Kollman, D. Case, K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A. M. Jr, B. Brooks, C. B. III, A. M. Jr, L. Nilsson, R. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. Pastor, C. Post, J. Pu, M. Schaefer, B. Tidor, R. Venable, H. Woodcock, X. Wu, W. Yang, D. York, M. Karplus, E. Sorin, V. Pande, J. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. Skeel, L. Kalé, K. Schulten, R. Guha, M. Howard, G.

Hutchison, P. Murray-Rust, H. Rzepa, C. Steinbeck, J. Wegner, E. Willighagen, J. Gasteiger, M. Marsili, A. Pérez, I. Marchán, D. Svozil, J. Sponer, T. C. III, C. Laughton, M. Orozco, R. Walker, M. Crowley, D. Case, F. Dupradeau, C. Cézard, R. Lelong, É. Stanislawiak, J. Pêcher, J. Delepine, P. Cieplak, D. Mobley, J. Chodera, K. Dill, J. Ryckaert, A. Bellemans, W. DeLano, H. Berman, K. Henrick, H. Nakamura, J. Markley, J. Sadowski, J. Gasteiger, G. Klebe, W. W. F. Vranken, W. Boucher, T. Stevens, R. Fogh, A. Pajon, M. L. M, E. Ulrich, J. Markley, J. Ionides, E. Laue, S. Velankar, C. Best, B. Beuth, C. Boutselakis, N. Cobley, A. S. da Silva, D. Dimitropoulos, A. Golovin, M. Hirshberg, M. John, E. Krissinel, R. Newman, T. Oldfield, A. Pajon, C. Penkett, J. Pineda-Castillo, G. Sahni, S. Sen, R. Slowley, A. Suarez-Uruena, J. Swaminathan, G. van Ginkel, W. W. F. Vranken, K. Henrick, G. Kleywegt, S. Velankar, Y. Alhroub, A. Alili, C. Best, H. Boutselakis, S. Caboche, M. Conroy, J. Dana, G. van Ginkel, A. Golovin, S. Gore, A. Gutmanas, P. Haslam, M. Hirshberg, M. John, I. Lagerstedt, S. Mir, L. Newman, T. Oldfield, C. Penkett, J. Pineda-Castillo, L. Rinaldi, G. Sahni, G. Sawka, S. Sen, R. Slowley, A. S. da Silva, A. Suarez-Uruena, G. Swaminathan, M. Symmons, W. W. F. Vranken, M. Wainwright, G. Kleywegt, A. Nederveen, J. Doreleijers, W. W. F. Vranken, Z. Miller, C. Spronk, S. Nabuurs, P. Güntert, M. Livny, J. Markley, M. Nilges, E. Ulrich, R. Kaptein, A. Bonvin, R. Engh, R. Huber, J. Doreleijers, W. W. F. Vranken, C. Schulte, J. Lin, J. Wedell, C. Penkett, G. Vuister, G. Vriend, J. Markley, E. Ulrich, W. Humphrey, A. Dalke, K. Schulten, E. Krieger, G. Koraimann, G. Vriend, L. Schuler, X. Daura, W. V. Gunsteren, R. Fogh, W. Boucher, W. W. F. Vranken, A. Pajon, T. Stevens, T. Bhat, J. Westbrook, J. Ionides, E. Laue, W. Rieping, M. Habeck, B. Bardiaux, A. Bernard, T. Malliavin, M. Nilges, M. Yoneya, D. van der Spoel, P. van Maaren, C. Caleman, B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, "ACPYPE - AnteChamber PYthon Parser interfacE.", *BMC research notes* **2012**, *5*, 367.

[31] J. P. Jcambeck, F. Mocci, A. P. Lyubartsev, A. Laaksonen, "Partial atomic charges and their impact on the free energy of solvation", *Journal of Computational Chemistry* **2013**, *34*, 187–197.

[32] C. I. Bayly, P. Cieplak, W. D. Cornell, P. A. Kollman, "A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving", *Journal of Physical Chemistry* **1993**, *97*, 10269–10280.

[33] W. D. Cornell, P. Cieplak, C. I. Bayly, P. A. Kollman, "Application of RESP Charges To Calculate Conformational Energies, Hydrogen Bond Energies, and Free Energies of Solvation", *Journal of the American Chemical Society* **1993**, *115*, 9620–9631.

[34] M. J. Dewar, E. G. Zoebisch, E. F. Healy, J. J. Stewart, "AM1: A New General Purpose Quantum Mechanical Molecular Model1", *Journal of the American Chemical Society* **1985**, *107*, 3902–3909.

[35] A. Jakalian, D. B. Jack, C. I. Bayly, "Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation", *Journal of Computational Chemistry* **2002**, *23*, 1623–1641.

[36] M. Feig, "Is alanine dipeptide a good model for representing the torsional preferences of protein backbones?", *Journal of Chemical Theory and Computation* **2008**, *4*, 1555–1564.

[37] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, "Comparison of multiple amber force fields and development of improved protein backbone parameters", *Proteins: Structure Function and Genetics* **2006**, *65*, 712–725.

[38] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ort, D. J. Fox, Gaussian 9, **2016**.

[39] F. P. Gasparro, "NMR determination of the rotational barrier in n,n-dimethylacetamide a physical chemistry experiment", *Journal of Chemical Education* **1977**, *54*, 258–261.

[40] R. L. Jarek, R. J. Flesher, S. K. Shin, "Kinetics of internal rotation of N,N-dimethylacetamide: A spin-saturation transfer experiment: An undergraduate physical chemistry experiment using FT-NMR to determine an internal rotational barrier", *Journal of Chemical Education* **1997**, *74*, 978–982.

[41] N. C. f. B. Information., N,N-Dimethylacetamide.

[42] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang, E. E. Bolton, "PubChem 2019 update: Improved access to chemical data", *Nucleic Acids Research* **2019**, *47*, D1102–D1109.

[43] R. Anandakrishnan, A. Drozdetski, R. C. Walker, A. V. Onufriev, "Speed of conformational change: Comparing explicit and implicit solvent molecular dynamics simulations", *Biophysical Journal* **2015**, *108*, 1153–1164.

[44] J. Zhang, H. Zhang, T. Wu, Q. Wang, D. Van Der Spoel, "Comparison of Implicit and Explicit Solvent Models for the Calculation of Solvation Free Energy in Organic Solvents", *Journal of Chemical Theory and Computation* **2017**, *13*, 1034–1043.

[45] J. Zielkiewicz, "Structural properties of water: Comparison of the SPC, SPCE, TIP4P, and TIP5P models of water", *Journal of Chemical Physics* **2005**, *123*.

[46] Q. Sui, D. Borchardt, D. L. Rabenstein, "Kinetics and Equilibria of Cis/Trans Isomerization of Backbone Amide Bonds in Peptoids", *Journal of the American Chemical Society* **2007**, 12042–12048.

[47] H. W. Huang, "Molecular mechanism of antimicrobial peptides: The origin of cooperativity", *Biochimica et Biophysica Acta - Biomembranes* **2006**, *1758*, 1292–1302.

[48] W. E. J. Acree, J. S. Chickos, H. Y. Afeefy, J. F. Liebman, S. E. Stein, D. R. J. Burgess in *NIST Chemistry WebBook- NIST Standard Reference Database Number 69*.

[49] K. Huang, C. W. Wu, T. J. Sanborn, J. A. Patch, K. Kirshenbaum, R. N. Zuckermann, A. E. Barron, I. Radhakrishnan, "A threaded loop conformation adopted by a family of peptoid nonamers", *Journal of the American Chemical Society* **2006**, *128*, 1733–1738.

[50] J. R. Stringer, J. A. Crapster, I. A. Guzei, H. E. Blackwell, "Construction of peptoids with all trans -amide backbones and peptoid reverse turns via the tactical incorporation of N -aryl side chains capable of hydrogen bonding", *Journal of Organic Chemistry* **2010**, *75*, 6068–6078.

[51] L. T. Roe, J. G. Pelton, J. R. Edison, G. L. Butterfoss, B. W. Tresca, B. A. LaFaye, S. Whitelam, D. E. Wemmer, R. N. Zuckermann, "Unconstrained peptoid tetramer exhibits a predominant conformation in aqueous solution", *Biopolymers* **2019**.

[52] A. Catte, M. R. Wilson, M. Walker, V. S. Oganesyan, "Antimicrobial action of the cationic peptide, chrysophsin-3: A coarse-grained molecular dynamics study", *Soft Matter* **2018**, *14*, 2796–2807.

[53] S. J. Marrink, A. H. De Vries, A. E. Mark, "Coarse Grained Model for Semiquantitative Lipid Simulations", *Journal of Physical Chemistry B* **2004**, *108*, 750–760.

[54] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, A. H. De Vries, "The MARTINI force field: Coarse grained model for biomolecular simulations", *Journal of Physical Chemistry B* **2007**, *111*, 7812–7824.

[55] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, A. H. De Vries, "The MARTINI force field: Coarse grained model for biomolecular simulations", *Journal of Physical Chemistry B* **2007**, *111*, 7812–7824.

[56] D. H. De Jong, G. Singh, W. F. Bennett, C. Arnarez, T. A. Wassenaar, L. V. Schäfer, X. Periole, D. P. Tieleman, S. J. Marrink, "Improved parameters for the martini coarse-grained protein force field", *Journal of Chemical Theory and Computation* **2013**, *9*, 687–697.

[57] K. P. Santo, M. L. Berkowitz, "Difference between magainin-2 and melittin assemblies in phosphatidylcholine bilayers: Results from coarse-grained simulations", *Journal of Physical Chemistry B* **2012**, *116*, 3021–3030.

[58] H. Leontiadou, A. E. Mark, S. J. Marrink, "Antimicrobial Peptides in Action clawed frog Xenopus lae V is 1 that exhibit a wide range of", *Journal of the American Chemical Society* **2006**, 12156–12161.

[59] P. Maculatin, P. J. Bond, D. L. Parton, J. F. Clark, M. S. P. Sansom, "Coarse-Grained Simulations of the Membrane-Active Antimicrobial", *Biophysical Journal* **2008**, *95*, 3802–3815.

# 5.0 Biophysical Investigation of Antimicrobial Activity

In this chapter we investigate the biophysical basis for peptoid antimicrobial activity. We investigate the hydrophobicity and membrane affinity of our library of peptoids and discuss the findings in the context of their biological activity and relevant literature, which is reviewed at the beginning of this chapter.

## 5.1 Introduction and Review of Literature

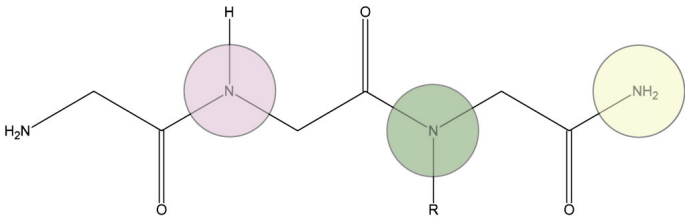



*Fig. 5.1:* Illustration of representative structure containing a primary (yellow shade), secondary (pink shade) and tertiary amide (green shade). Primary amides are those where the nitrogen is directly bonded to 1 carbon atom. In secondary and tertiary amides the nitrogen is directly bonded to 2 and 3 carbon (or other non-hydrogen) atoms respectively.

The use of peptoids as antimicrobial agents was identified as a promising application soon after their development [1]. Peptoids are attractive candidates for investigation in this field due to their potential for conformational and chemical diversity. They are also resistant to proteolytic degradation, which makes them considerably more bio-stable than AMPs. This bio-stability was first highlighted by Miller *et al.*, who characterised a library of peptoids and their peptide counterparts, assessing their in-vivo properties and concluding that the peptoids are not subject to the enzymatic degradation to which AMPs are notoriously vulnerable [2]. The enzymatic resistance of peptoids has been attributed to the side chain substitution from the alpha carbon to the amide nitrogen, so that peptoid backbone is composed of tertiary amides that are not recognised by enzymes in the way that peptide secondary amides are (Figure 5.1 illustrates the difference between primary, secondary and tertiary amides). This is a property common to all peptoids, independent

of sequence or side chain chemistry, which has been corroborated by several further studies [3, 4].

Charge, hydrophobicity and amphipathicity have been identified as structural features which affect the antimicrobial activity of peptides and peptoids. The balance between these physio-chemical properties is important in ensuring that the antimicrobial agent is both sufficiently active and selective [5]. Selectivity is crucial for a successful drug, which must effectively target pathogenic cells and cause minimal damage to the cells of the host organism. It is therefore important to investigate the mechanisms of peptoid toxicity to mammalian type cells, as well as toxicity to pathogenic cells, in order to design optimum sequences. Thus far the process of identifying active sequences has mainly been based upon the synthesis, characterisation and biological screening of large combinatorial peptoid libraries, with connections between structure and function being identified from the results of these large scale experiments [1, 5–7]. Understanding these structure-function relationships and the mechanisms behind them will facilitate *de novo* design of useful sequences in the future.

### *5.1.1 Stucture-Activity Relationships*

An extensive selection of peptoids designed to mimic and improve upon the antimicrobial properties of natural AMPs have been reported in the literature but so far there is relatively little published data proposing definitive links between peptoid structure and antimicrobial activity. Even for AMPs, progress into developing mechanistic knowledge to facilitate rational sequence design has been slow [8]. Large peptoid libraries have been screened for antimicrobial activity in biological assays and found, in many instances, to show promising activity levels against a range of gram positive and gram negative bacteria [9–11], including bacterial biofilms [3, 12]. Peptoids have also been found to be active against fungal species [13, 14] and various parasites that are responsible for high profile infectious diseases including malaria and the lesser-known, neglected tropical disease Leishmaniasis [15, 16]. The range and level of peptoid antimicrobial activity indicates that they represent a promising candidate for a new class of broad spectrum therapeutic compounds which may help to address the issue of antimicrobial resistance in the future.

The screening of peptoid libraries has begun to help in identifying key structural features that relate to activity against particular pathogens, with the ultimate aim of discovering rules to facilitate rational design. Some studies have focussed on the direct translation of naturally occurring antimicrobial peptides into their peptoid analogues [6, 17–19]. However, the folded conformation of these AMPs is often crucial to their activity and it does not necessarily follow that the analogous peptoids would adopt the same conformations and therefore it can be difficult to reliably obtain active peptoids in this manner

[5]. Several groups have outlined systematic protocols for the direct translation of AMP sequences in to active petoids, with varying levels of success [20]. However, given that a vast selection of amines are available to incorporate into peptoid sequences, but do not appear within the range of natural amino acids, there is potential for the creation of active peptoid sequences far beyond the scope of direct AMP mimics. Therefore it is important to understand the relationships between peptoid sequence, structure and activity. Studies have so far revealed that the activity of peptoids may depend upon physical factors such as the length of the sequence, side chain chemistry and patterning of different residues [5]. Each of these factors contributes to the overall hydrophobicity, amphipathicity and charge of the peptoids and therefore its antimicrobial activity.

### *Effects of Sequence Length*

Several groups have independently suggested that peptoid activity varies with sequence length [21–23]. A study by Bolt *et al.* which explored both the activity of a library of peptoids against gram positive and gram negative bacteria and toxicity towards mammalian cell lines included 6-, 9- and 12-mer sequences, with each sequence containing an increasing number of repeats of the same ABB patterned trimer of residues, where A is a cationic residue and B an aromatic or chiral residue [23]. The results indicated that the 6-mers were generally inactive, the 9-mers active in some instances and the 12-mers almost all active, concluding that 12 residues may be the optimum sequence length for antimicrobial peptoids. A study by the Barron group [22] offers the same conclusion with regards to 12 residues being the optimum sequence length, demonstrating for another peptoid library based around repeated trimer sequences, that shorter sequences are not as potent and longer ones show increased levels of haemolytic activity such that they are not selective enough to be considered for therapeutic use. Though this is a clear conclusion for repeating trimer sequences, whether a similar trend would be observed in scrambled versions of these peptoids has not been addressed.

### *Effects of Side Chain Chemistry: Charge, Hydrophobicity and Amphipathicity*

#### *Hydrophobicity*

The peptoid library described in the aforementioned Barron group study [22] was based around variations on the $(N\text{Lys}N\text{spe}N\text{spe})_4$ (motif 2) peptoid sequence. By making substitutions for different monomers, shown in Figure 5.2, the group investigated the effect of varying the peptoid charge and hydrophobicity on the antimicrobial activity. To modify the peptoid hydrophobicity independently of sequence length $N$smb and $N$sna were substituted into sequences in varying proportions in place of $N$spe. These substitutions resulted in peptoids which were more active with higher hydrophobicity. Unfortunately

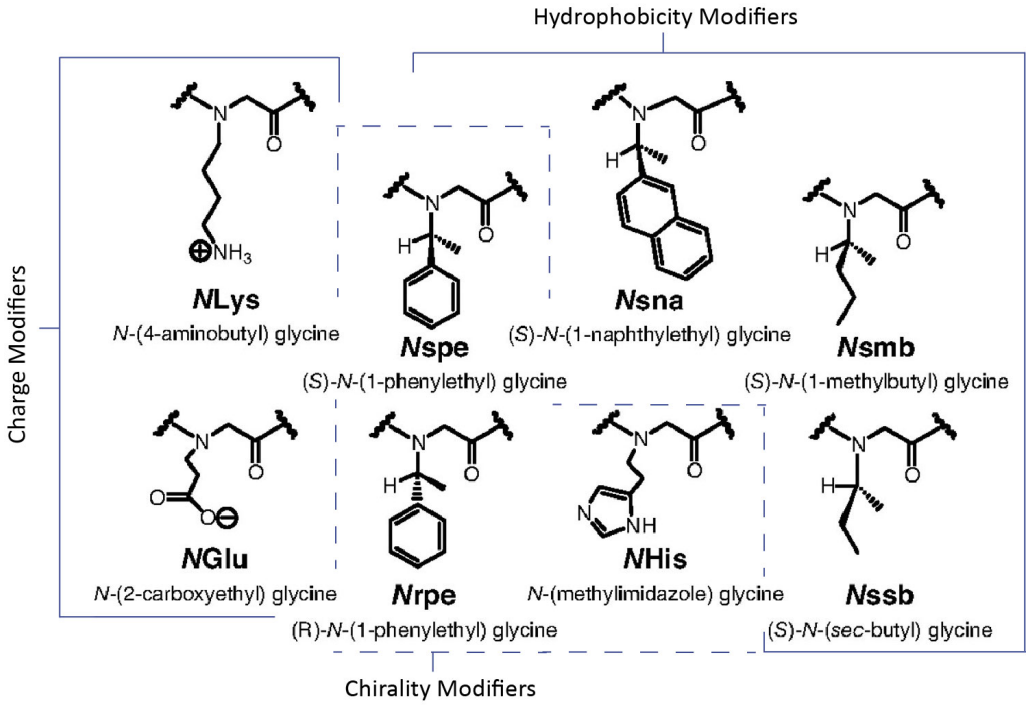



*Fig. 5.2:* Side chain groups used in the Barron group study, with indication of which monomers were used to modify peptoid charge, hydrophobicity and chiral handedness. Adapted from [22].

increased hydrophobicity also correlated to increased levels of cytotoxicity and therefore worse selectivity ratios.

Hydrophobicity modification by substituting *N*Lys side chains for *N*ae has been considered in peptoids as the equivalent substitutions in antimicrobial peptides (lysine for ornithine and diaminobutyric acid) have been found to modulate activity and selectivity [24]. Thus far *N*ae to *N*Lys substitution in peptoids has been shown to have only a limited effect on the sequence hydrophobicity, as measured by HPLC retention time. However, such substitutions do appear to affect activity and selectivity. *N*ae containing sequences demonstrate blood cell-lysing ability at lower concentrations than their *N*Lys counterparts [18]. Additionally this substitution moderates peptoid activity against bacterial biofilms, with the *N*ae side chain imparting enhanced activity relative to *N*Lys [3]. Substitution of the even longer equivalent side chain, *N*ah, which has a 6 carbon atom chain, resulted in further reduction in activity levels [3].

### *Charge*

Charge influences peptoid activity and selectivity [3, 18, 23, 25, 26]. Charge moderation by the partial substitution of cationic *N*Lys for zwitterionic *N*Glu was detrimental to the peptoid antibacterial properties in the Barron group study [22]. Lower charged sequences exhibited reduced potency towards bacteria but maintained similar levels of cytotoxicity to

other sequences in the library. This was attributed to the loss of electrostatics promoting interactions between the peptoid and anionic bacterial membranes. Complete substitution of *N*Lys for *N*Glu was found to render sequences completely inactive against all cell types tested [22]. There is some evidence that simple net charge may not be an ideal parameter to predict peptoid activity but more sophisticated measurements of the charge distribution over the structure of the peptoid as a whole can be useful [6], as it has been in quantitative structure activity relationship studies of AMPs [27, 28].

*Amphipathicity*

Modifying the amphipathicity of a peptoid sequence is not a trivial task as it requires accurate prediction of the folded state of the peptoid. Within their study the Barron group circumvent this difficulty by creating distinctly terminally amphipathic sequences, rather than facially amphipathic ones [22]. It is still difficult to rule out the effect that unpredicted folding could have on the amphipathicicty of these sequences however. A terminally amphipathic analogue of $(N\text{Lys}N\text{spe}N\text{spe})_4$, showed slightly reduced antibacterial and haemolytic activity than the standard sequence with identical monomer composition. This is a subtle indication that the repeating motif, or at least some regular distribution of charges and hydrophobic groups along the backbone, may be important in creating active sequences. Within the same study a scrambled peptoid sequence, thought to be neither terminally or facially amphipathic, was found to have comparable antimicrobial activity but reduced haemolytic activity relative to its repeat motif analogue. This indicates that careful control of the amphipathicity via sequence scrambling could lead to the design of optimally selective sequences. As many of the most potent peptoids discovered so far, including $(N\text{Lys}N\text{spe}N\text{spe})_4$, are predicted to be facially amphipathic, variation in this respect still warrants further investigation.

### *5.1.2 Peptoid Antimicrobial Mechanism Studies*

Further studies have been described to analyse the mechanisms by which peptoids kill pathogens. It is believed that due to their structural similarities and observed correlations with charge and hydrophobicity, peptoid mechanisms of action are somewhat similar to those of natural AMPs and that membrane activity in particular is an important factor. Nevertheless there are many different proposed mechanisms by which AMPs disrupt membranes (described in Chapter 1), which remain disputed. The proposed similarities between AMP and peptoid membrane disruption mechanisms are supported by reports that the two classes of molecule can act in a synergistic manner [29]. Similarly, AMPs from different species have also been shown to act synergistically [30].

Chongsiriwatana *et al.* demonstrated that synergistic AMP-peptoid pairs could be de-

signed by selecting a pair of one highly selective compound with moderate hydrophobicity and one less selective compound with higher hydrophobicity [29]. The peptoids studied in this work were a series of sequence variations of $(N\text{Lys}N\text{spe}N\text{spe})_4$. For the best pairings this produced more than an 8-fold increase in activity against *E coli* compared to each of the compounds alone. The strength of the mechanistic analogy between the two classes of molecule was demonstrated by showing that the substitution of a non-selective AMP with a similarly non selective peptoid or vice versa did not reduce the synergistic effect of pairing compounds. It was also shown that by careful selection of moderately hydrophobic pairings the antimicrobial activity could be enhanced and the haemolytic activity minimised, potentially providing a new avenue to explore in terms of the therapeutic potential of these molecules [29]. Though this study demonstrated the potential analogy between peptoid and peptide mechanisms, the peptoid mechanisms were not explicitly investigated.

Various techniques previously used to investigate AMP mechanisms are increasingly being used in peptoid studies, including microscopy methods to investigate visual changes in membrane structure, intracellular targeting and membrane pore formation [31–34]. Dye leakage assays are particularly useful in evaluating AMP induced pore formation. In dye leakage assays using giant unilamellar vesicles (GUVs) the rate of leakage over time is used to determine whether the peptides form stable or transient pores, or generally act by non-specific permeabilisation (carpet model) [31, 33]. This technique is increasingly being applied in peptoid mechanistic studies, often using a peptide with a well-defined membrane disruption method as a reference [25]. Mojsoska *et al.* compared the calcein leakage rate from GUVs due to two peptoids and melittin (an AMP found naturally in honey bee venom), demonstrating that the leakage induced by the peptoids was considerably slower than the melittin. Nevertheless the peptoid induced leakage appeared to be concentration dependent, indicating that the interaction could occur via a two state mechanism analogous to AMPs [35]. Scanning Electron Microscopy showed membrane damage occurred in *E. coli* due to the action of the peptoids. A mismatch in the observed levels of membrane disruption and the bacterial killing kinetics in this study indicates that potentially the peptoids act via a dual mechanism, with some crossing the membrane and binding to internal targets.

### *Interactions with Model Membranes*

A large number of mechanistic studies of AMPs are based on the interaction between peptides and model membranes. In the majority of cases, the model membrane systems are either lipid vesicles or supported lipid bilayers of varying complexity. Similar studies are increasingly being used to probe the interactions of peptoids with model membrane systems [29]. Some of these studies focus on evaluating structural changes in the membranes as

a result of peptoid activity. Andreev *et al.* used AFM to demonstrate that peptoids aggregate at the surface of supported lipid bilayers causing morphological changes in the bilayer [36]. Cyclisation has been shown to enhance these membrane disruptive tendencies [37].

The adsorption of linear sequences to bilayers has been shown to depend on sequence length and charge, which is intuitive given the correlation between these factors and peptoid antimicrobial activity. Landry *et al.* demonstrate that peptoid adsorption mechanisms vary with sequence length. However, in contrast to biological studies, Landry *et al.* found that trimers with high net cationicity and hydrophobicity adsorb more strongly to DOPC bilayers than longer sequences [38]. This indicates that predicting activity may require investigation of factors other than simple membrane affinity, potentially because model membranes do not capture the complexity of biological membranes. This poses a problem in investigations such as these. It is clearly important to find a compromise between simplicity in model systems that allows for systematic investigation of the physical principles governing the biological activity, but also the model systems must represent the true system closely enough for these rules to apply to both.

It has also been established that in many cases the peptoids themselves also undergo structural changes in response to interactions with model membranes [22]. This is another feature of peptoids which is analogous, though not identical to AMPs, many of which undergo a transition from a random coil to $\alpha$-helical structure upon binding to lipid membranes [39]. Peptoids on the other hand appear to transition more subtly from one folded state to another [22]. CD spectroscopy has been used to evaluate structural changes in AMPs upon binding to different lipid membranes [40]. CD has also been used to analyse the structure in peptoids bound to lipid vesicles [29]. Chongsiriwatana *et al.* demonstrated using CD spectroscopy that peptoids, including $(N\text{Lys}N\text{spe}N\text{spe})_4$, adopt different conformations when interacting with lipid vesicles than when free in solution. The spectral differences which indicate this change include an increase in intensity of the second characteristic helical minimum ($\lambda_2$) and corresponding reduction in intensity of the first minimum ($\lambda_1$) [29]. These are qualitatively similar to the differences we observed in Chapter 2 between the CD spectra of our repeat motif peptoids in PBS and octanol. Within their study Chongsiriwatana *et al.* found that the helicity of the membrane-bound peptoids did not correlate strongly with antimicrobial activity but did correlate to haemolytic activity, concluding that the amphipathicity that a helical structure imparts upon a sequence contributes strongly the interaction with the mammalian-type membranes.

## 5.2 Summary of Biological Data for Repeat Motif Peptoid Library

Some of the peptoids investigated in this thesis have been subject to a range of *in-vitro* studies investigating their activity against different bacteria, parasites and mammalian cell lines. The biological data is reported as values for the minimum inhibitory concentration (MIC), the median effective dose ($ED_{50}$) and the half maximal inhibitory concentration ($IC_{50}$). These are subtly different measures used to report the *in-vitro* efficacy of a drug compound against a particular pathogen. The MIC is the minimum drug concentration which can prevent the visible growth of the target bacteria. $IC_{50}$ is the drug concentration necessary to inhibit a particular biological process in the target. $ED_{50}$ is the drug concentration which elicits the desired response in 50% of the target. In the cases discussed here the desired response is the death of the pathogenic cells, or the mammalian cell in the case of the cytotoxicity assays. An ideal peptoid drug candidate will display low $IC_{50}$, $ED_{50}$ and MIC values against pathogenic cell lines and high values against mammalian cell lines, indicating that it is both antimicrobially active and selective. This is often quantified as a selectivity ratio, which is the ratio of the affinity for (or efficacy against) the target (pathogenic) cells to the off-target cells. Therefore a high selectivity ratio is indicative of a good drug. In the following sections we present a summary of biological activity data for the repeat motif peptoid sequences investigated in this thesis.

### *5.2.1 Antibacterial Activity*

Bacteria can be classified as gram-positive or gram-negative depending on the properties of their cell wall, which are illustrated in Figure 5.3. The cell walls of gram-positive bacteria are composed of a single lipid bilayer with a thick (20-80 nm) outer layer of peptidoglycan. Gram-negative bacteria cell walls are made up of a double bilayer, with a thin (2-3 nm) peptidoglycan layer between the two bilayers and an additional covering of lipopolysaccharides on the outer membrane. This variation in cell wall structure in pathogenic bacteria results in different responses to therapeutics, with gram-positive bacteria often being more receptive to membrane active compounds, due to the lack of outer lipid membrane [8].

Activity data for a sub-set of the motif 2 peptoids, which have been screened against two strains of gram positive bacteria, *Staphylococcus epidermidis* and *Staphylococcus aureus*, is presented in Table 5.1. *S. epidermidis* is a bacterium frequently found in the normal human skin flora. It is not normally pathogenic but can present a threat to immune-compromised individuals and can also form detrimental biofilms on surgical implants such as catheters and prostheses [41]. Several strains of *S. epidermidis* have developed resistance to conventional antibiotics [42, 43]. *S. aureus* is also often present as part

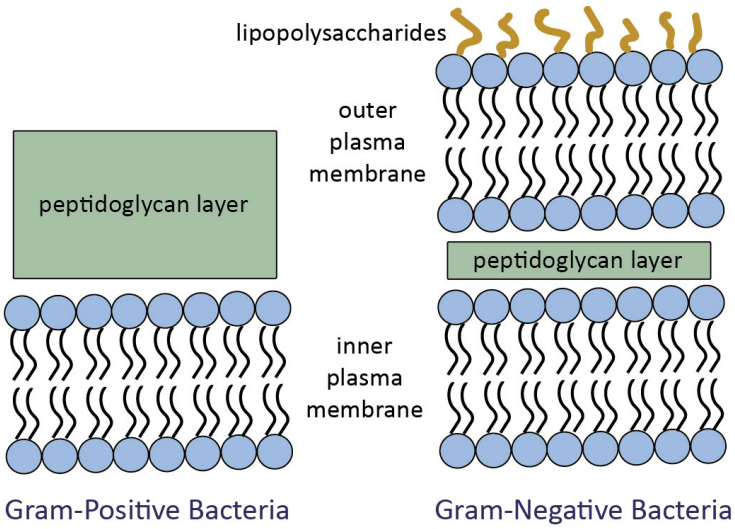



*Fig. 5.3:* Schematic diagram showing the different compositions of the gram positive and gram negative bacterial cell wall. This is a highly simplified representation of complex biological membranes which shown only the key features of interest and is not to scale.

of the normal skin flora and can be innocuous but also develop pathogenic effects in a wide variety of contexts including skin, bone and joint infections,[44] and food poisoning [45]. Certain strains of *S. aureus* have developed extreme multiple drug resistance and are now known as the superbug MRSA [46, 47].

The activity of of these two peptoids was also screened against two gram-negative strains, *Escherichia coli* and *Pseudomonas aeruginosa*. *E. coli* is a ubiquitous and largely innocuous bacteria found in the human digestive tract. Some strains are notable for causing serious food poisoning [48]. *P. aeruginosa* can cause opportunistic infections which are often chronic or extremely serious, such as pneumonia and septic shock and displays advanced levels of drug resistance [49]. It has a particularly wide and worrisome variety of intrinsic and mutated mechanisms of resistance that are partially attributed to its tendency towards biofilm formation, but also the emission of antibiotic degrading enzymes [50, 51]. Peptoids therefore represent a promising new line of defence against this pathogen due to their proteolytic resistance.

In addition to these bacteria, the peptoids were screened against two mammalian cell lines, HaCaT and $HepG_2$, to provide a measure of their cytotoxicity and selectivity. These are derived from human skin and liver cell respectively and are widely used in research.

The motif 2 peptoids are considerably more active against the gram positive than the gram negative bacteria, with promising MIC values of 1-2 $\mu$M against the former. Generally the *N*Lys variant of the sequence is more active against the pathogenic cells, but also against the human cell lines.

| Peptoid | MIC ($\mu$M) | | | | ED$_{50}$ ($\mu$M) | |
|---|---|---|---|---|---|---|
| | Gram Positive | | Gram Negative | | Human Cell Lines | |
| | *S. aureus* | *S. epidermis* | *P. aeruginosa* | *E. coli* | HaCaT | HepG$_2$ |
| $(N\text{ae}N\text{spe}N\text{spe})_4$ | 2 | 2 | 50 | 100 | 26 | 41 |
| $(N\text{Lys}N\text{spe}N\text{spe})_4$ | 2 | 1 | 50 | 25 | 20 | 29 |

*Tab. 5.1:* Activity of the motif 2 peptoids against gram positive, gram negative and human cell lines. This data is from work by Bolt *et al.* [52].

### 5.2.2 Anti-Parasitic Activity

The repeat motif 2 peptoids $(N\text{ae}N\text{spe}N\text{spe})_4$ and $(N\text{Lys}N\text{spe}N\text{spe})_4$ have been screened against several protozoan parasites of the *Leishmania* genus. These parasites are transmitted to humans via sandfly bite and can cause the disease Leishamiasis, of which there estimated to be up to two million new cases each year, with many resulting in death or lifelong disability [53]. Current treatments for leishmaniasis are largely ineffective, with pentavalent antimonials remaining the only truly effective option. This requires hospitalisation and intravenous administration, or else extremely painful intramuscular administration, making treatment especially difficult in the tropical and subtropical regions where the disease is most prevalent [54].

Eggimann *et al.* screened a library of peptoids with selected cationic and aromatic side chains arranged as motif 2, including $(N\text{ae}N\text{spe}N\text{spe})_4$ and $(N\text{Lys}N\text{spe}N\text{spe})_4$. These peptoids varied in length and were either 3-, 6-, 9- or 12-mers. The 12-mers were found to be the most active against different forms of the *Leishmania* parasite, with $N$ae side chains showing enhanced efficacy relative to $N$Lys side chains in otherwise identical sequences [15]. The high activity levels of RM3 and RM4 against these parasites contributed to their selection for investigation in this work.

| Peptoid | MIC ($\mu$M) *L. mexicana* | |
|---|---|---|
| | Promastigotes | Amastigotes |
| $(N\text{ae}N\text{spe}N\text{spe})_4$ | 7 | 17 |
| $(N\text{Lys}N\text{spe}N\text{spe})_4$ | 8 | >100 |

*Tab. 5.2:* Activity of the motif 2 peptoids against the promastigote and amastigote forms of *Leishmania Mexicana* This data is from the 2015 study by Eggimann *et al.* [15].

Both the motif 2 and the motif 1 peptoids have been screened against *P. falcipicarum* a common causative agent of malaria. Each of these peptoids displayed high levels of activity

against this parasite, with the motif 1 peptoids also showing high selectivity (Table 5.3).

| Peptoid | $IC_50$ *P. falcipicarum* | Selectivity Index |
|---|---|---|
| $(N\text{ae}N\text{spe})_6$ | 0.68 | 49.2 |
| $(N\text{Lys}N\text{spe})_6$ | 2.6 | 10.2 |
| $(N\text{ae}N\text{spe}N\text{spe})_4$ | 0.68 | 6.3 |
| $(N\text{Lys}N\text{spe}N\text{spe})_4$ | 0.56 | 6.5 |

*Tab. 5.3:* Activity of the motif 1 and 2 peptoids against *P. falcipicarum*, with selectivity index calculated using activity against L-6 rat myoblasts. This data is from the 2015 study by Bolt *et al.* [23].

The activity of both motifs against the parasite is comparable but the high selectivity index for both motif 1 sequences, the *N*ae variant in particular indicates that these are much less cytotoxic than the motif 2 sequences. This could be an example of the decreased amphipathicity of motif 1, relative to motif 2, moderating the interaction of these sequences with mammalian-type membranes.

## 5.3 Results and Discussion

### *5.3.1 PBS-Octanol Partitioning*

Partitioning between immiscible aqueous and organic solvents is used routinely in pre-clinical drug discovery to indicate how hydrophobic a compound is and therefore how it will distribute to different regions in the body. Hydrophobic compounds localise to areas such as cell membranes, where the lipid bilayer interior provides a hydrophobic environment. Hydrophilic compounds localise to aqueous environments such as blood plasma. In the partitioning experiments reported in this work we use PBS as an aqueous phase as its neutral pH (7.4), osmolarity and ion concentrations make it isotonic with human blood plasma. We use octanol as the organic phase due to its widespread use in pharmaceutical partitioning studies. Partition coefficients are often calculated using octanol as the organic phase as it is thought to be a reasonable model for a bilayer interior due to its polar hydroxyl group and long alkyl chain [55].

The distribution coefficient, $D$, is the equilibrium ratio of the concentrations of a compound (in this case a peptoid) in each of the two phases which it has been allowed to partition between. The logarithm of the ratio, Log $D$, gives an indication of the hydrophobicity of the compound where one solvent is aqueous and the other organic. In this work Log $D$ is defined as follows:

$$\text{Log}D = \log\left(\frac{[\text{peptoid}]_{\text{PBS}}}{[\text{peptoid}]_{\text{octanol}}}\right) \tag{5.1}$$

Where $[\text{peptoid}]_{\text{PBS}}$ is the concentration of peptoid in the PBS phase and $[\text{peptoid}]_{\text{octanol}}$ is the concentration, in the octanol phase in the equilibrated system. By this definition a positive value for Log $D$ indicates that the compound partitions preferentially into the PBS and a negative value indicates a preference for the octanol. The former case implies a less hydrophobic peptoid than the latter.

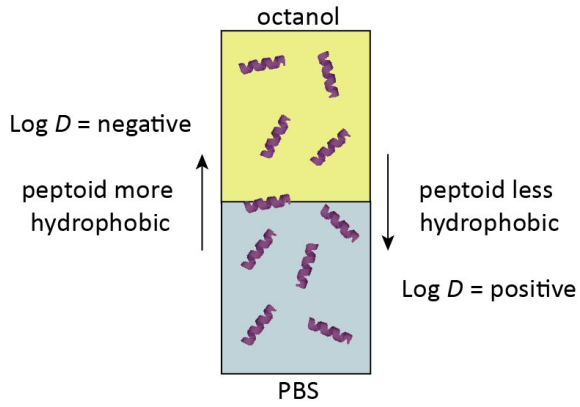



*Fig. 5.4:* Schematic diagram representing the experimental set up for the partitioning of peptoids between PBS and octanol to calculate Log $D$ as a measure of hydrophobicity.

The first use of partitioning experiments to calculate Log $D$ values for peptoids was reported by Bolt *et al.* in 2017. The authors experimentally determined Log $D$ for a small library of peptoids and compared the results to reverse-phase HPLC and biological activity data. The results demonstrated that hydrophobicity measured by Log $D$ can differ significantly to when measured by HPLC and that the former could be used as a new predictive tool for membrane active peptoids, as the folded hydrophobicity of a peptoid can differ from its net hydrophobicity [52]. Bolt *et al.* report values for Log $D$ for two of the peptoids in our repeat motif library: $(N\text{ae}N\text{spe}N\text{spe})_4$ and $(N\text{Lys}N\text{spe}N\text{spe})_4$. The values reported are -1.85 and 1.21 respectively, indicating that choice of charged side chain is influential in this case, as the $N$ae variant partitions into the octanol phase but the $N$Lys variant does not. The folded conformation of the former is therefore more hydrophobic than the latter. In contrast, the two peptoids have similar HPLC retention times (16.7 and 14.9 minutes respectively), indicating similar net hydrophobicity. Therefore in the case of these particular peptoids the folded hydrophobicity may be a more useful and physiologically relevant characteristic to predict membrane activity.

### *Experimentally Determined Log D Values for Repeat Motif and Scrambled Sequences*

We carried out PBS-octanol partitioning experiments for the repeat motif and scrambled sequence peptoids. The experimental protocol for partitioning is detailed fully in Chapter 2 and summarised in Figure 5.4. This protocol is based on that described by Bolt *et al.*

[52] but using micro-volume samples rather than bulk UV-Vis measurements. The Log $D$ values for the repeat motif peptoids are shown in Figure 5.5. The data for the motif 1 peptoids was collected by Kirsten S. Allan. The repeat motif sequences have different overall charges and percentages of hydrophobic $N$spe residues depending on sequence and therefore the Log $D$ (and hydrophobicity) is expected to vary between sequences. However, for motifs 1 and 2 our experimentally determined Log $D$ values are very similar, with the exception that the $N$ae variant of motif 2 alone partitions into the octanol phase (negative Log $D$ value), indicating that this peptoid is significantly more hydrophobic than the other three, which preferentially partition into the PBS phase (positive Log $D$ values). Significant experimental difficulty was encountered in gathering the data to calculate these values for the motif 1 peptoids, which resulted in a much larger estimate for the error in these Log $D$ values.

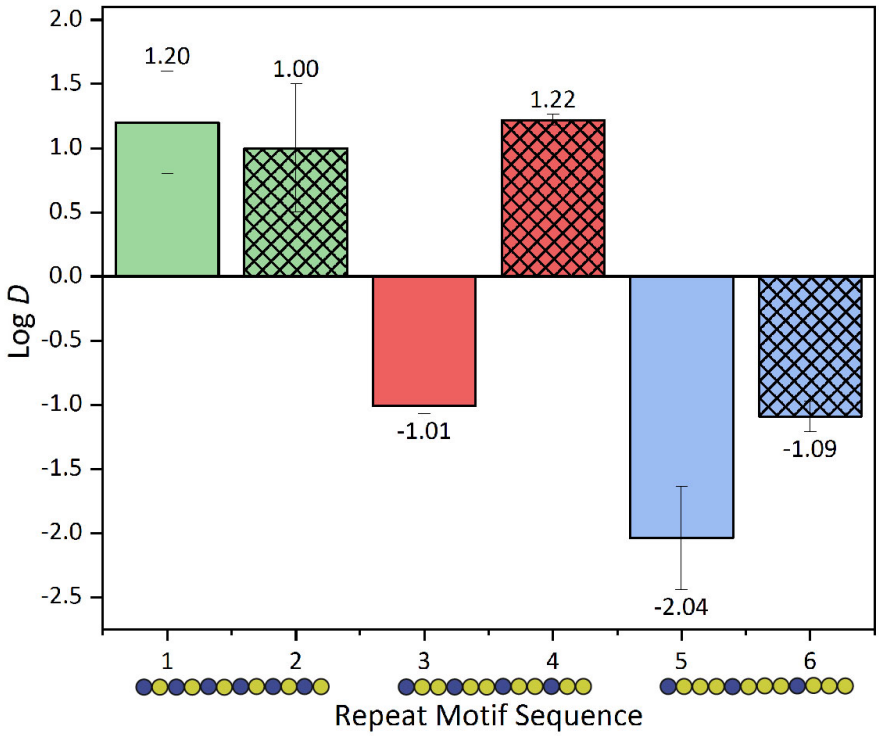



*Fig. 5.5:* Log D values for the repeat motif peptoids partitioning between PBS and octanol. Motif 1 sequences shown in green, motif 2 sequences in red and motif 3 sequences in blue with $N$ae side chain variants in solid colour and $N$Lys with diamonds. Data for motif 1 peptoids was collected by Kirsten S. Allan.

Agitation of the motif 1 peptoid partitioning samples resulted in the formation of an emulsion of octanol droplets stabilised in the PBS phase (verified by microscopy). Therefore an additional step of centrifugation, followed by further equilibration for 24 hours to restore complete phase separation in the system was carried out before measuring the peptoid concentration in each phase. The emulsion formation suggests that these particular peptoids are very surface active and potentially the additional steps taken to

account for this are not sufficient and this resulted in large discrepancies in the value of Log $D$ calculated for each repeat of the partitioning experiment (and the large error values). In the future, repeat measurements using larger volume samples and a bulk UV-vis measurement rather than the micro-volume protocol followed here, could provide further insight into the partitioning of these peptoids.

The results for the repeat motif peptoids indicate that there is a correlation between the charge of the peptoid and the overall folded hydrophobicity (as measured by Log $D$). Both motif 1 sequences (+6 charge) preferentially remain in the PBS phase while both motif 3 sequences (+3 charge) preferentially partition into the octanol phase indicating that as the charge on the peptoid decreases and the number of hydrophobic aromatic groups increases, the folded peptoid becomes more hydrophobic. Interestingly in the case of the motif 2 peptoids (+4 charge) the preferred phase depends on the structure of the charged side chain, with the shorter $N$ae variant partitioning into the octanol phase and the longer $N$Lys variant preferentially remaining in the PBS. This same behaviour was reported by Bolt *et al.* for the motif 2 peptoids [52].

| Sequence | HPLC retention time (minutes) | Log $D$ |
|---|---|---|
| $(N\text{ae}N\text{spe})_6$ | 13.7 | 1.2 ± 0.4 |
| $(N\text{Lys}N\text{spe})_6$ | 14.1 | 1.0 ± 0.5 |
| $(N\text{ae}N\text{spe}N\text{spe})_4$ | 16.7 | -1.01 ± 0.06 |
| $(N\text{Lys}N\text{spe}N\text{spe})_4$ | 14.9 | 1.22 ± 0.04 |
| $(N\text{ae}N\text{spe}N\text{spe}N\text{spe})_3$ | 20.8 | -2.04 ± 0.41 |
| $(N\text{Lys}N\text{spe}N\text{spe}N\text{spe})_3$ | 20.0 | -1.09 ± 0.12 |

*Tab. 5.4:* Log $D$ and HPLC retention times for the repeat motif peptoid sequences. HPLC data was collected by Hannah L. Bolt.

The values for Log $D$ for the repeat motif sequences are also reported in Table 5.4, alongside the HPLC values for these peptoids. The HPLC values indicate a general trend of increasing hydrophobicity with increasing number of $N$spe residues. Log $D$ for the $N$ae side chain sequences appears to vary linearly with the net charge of the peptoid (Figure 5.6a). Contrastingly, the $N$Lys sequences appear to plateau for charges above 4. Similar trends are observed in the HPLC time, shown in Figure 5.6b, though as there are only 3 data points for each subset of peptoids this has not been explored fully enough to confidently assert this conclusion.

Since the motif 2 peptoids have Log $D$ values of opposite sign depending on the structure of the charged side chain group we investigated the scrambled sequences to see if this relationship holds true for other peptoids with the same overall charge (+4) and number of hydrophobic $N$spe residues. The variations in the scrambled sequence CD spectra indicate that each of these peptoids may adopt slightly different helical configurations in PBS which could affect their overall folded hydrophobicity by shielding hydrophobic

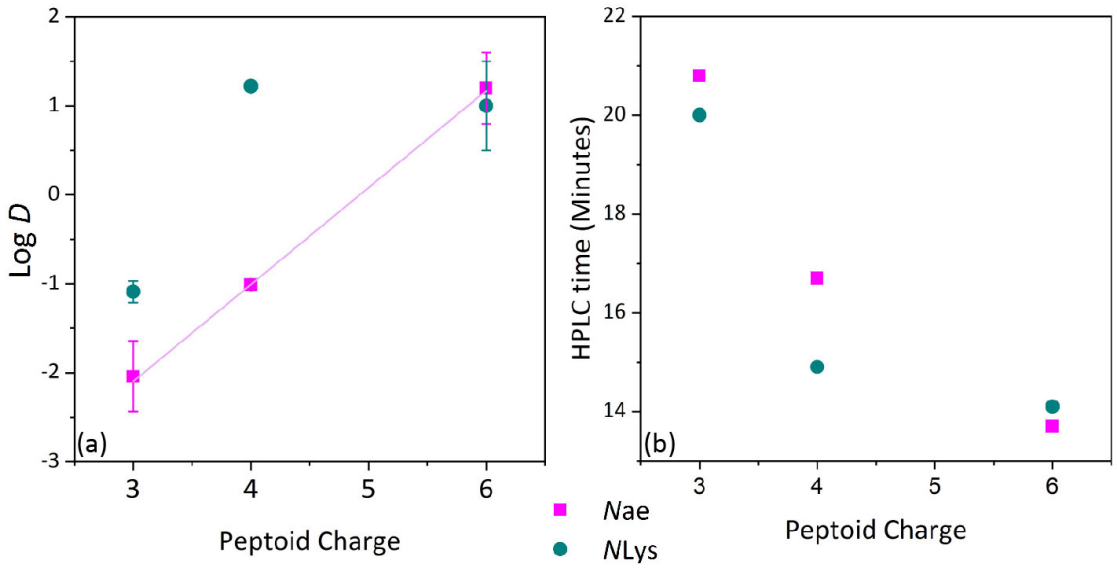



*Fig. 5.6:* (a) Log $D$ values for the repeat motif peptoids plotted against peptoid charge. The peptoids with $N$ae side chains are shown in magenta and $N$Lys side chains in teal. (b) HPLC retention times (collected by Hannah L. Bolt) for the repeat motif peptoids plotted against peptoid charge.

residues and exposing polar ones to different extents. The Log $D$ values for the scrambled sequences are shown in Figure 5.7.

The most striking result of the scrambled sequence partitioning experiments was that the observation regarding the nature of the charged side chain in the motif 2 sequences holds true. In all cases scrambled sequences containing the $N$ae monomer as the charged side chain partitioned into the octanol phase (yielding a negative value for Log $D$), whereas identical sequences with $N$Lys as the charged monomer partitioned into the PBS phase (yielding a positive value for Log $D$). This is somewhat counter-intuitive as it would be expected that the presence of the additional 2 methylene groups in the $N$Lys side chain would make it more hydrophobic than $N$ae. Therefore it seems likely that enhanced hydrophobicity of the $N$ae sequences relative to their $N$Lys counterparts must arise from variations in their secondary structure.

Figure 5.8 shows the Log $D$ for the scrambled sequences with $N$ae side chains plotted against Log $D$ for the equivalent sequences with $N$Lys side chains. These are clearly not correlated and show that there is substantially more variation in the Log $D$ values for the $N$ae side chain sequences than for the $N$Lys side chain sequences. Their HPLC times, which were collected by Wataru Ichinose, on the other hand are very similar and all fall within a small range (14.7-16.7 minutes). The physical origins of this behaviour are unclear, particularly as the variations in Log $D$ do not obviously correlate to any features of the primary sequences such as the number of adjacent $N$spe residues or particular positioning of positive residues. The CD spectra of the $N$Lys side chain sequences in PBS indicate that there is more conformational diversity in this subset of the scrambled

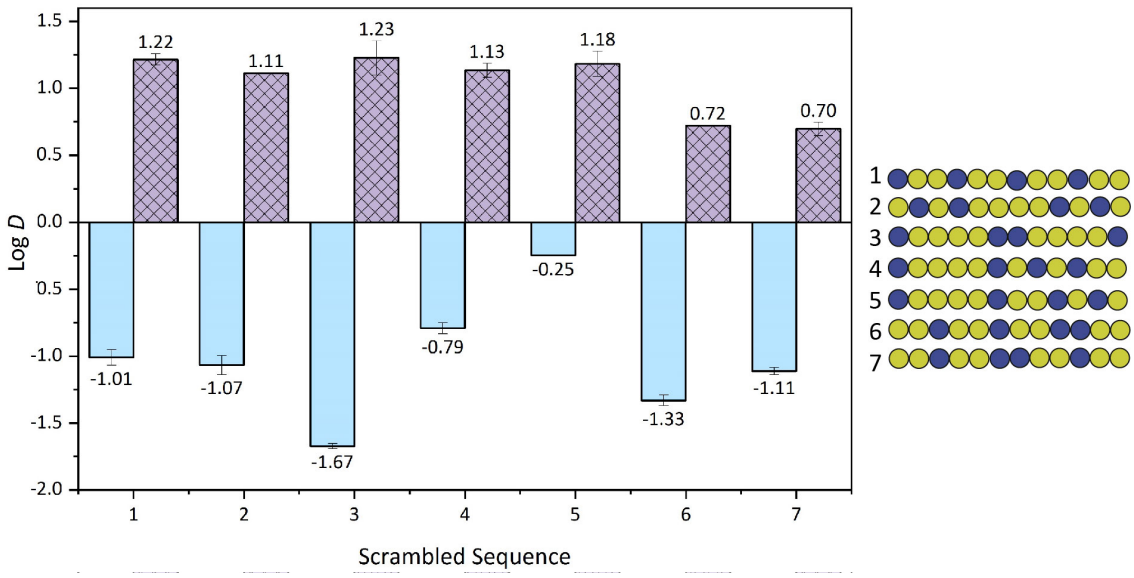



*Fig. 5.7:* Log $D$ values for the scrambled sequence peptoids partitioning between PBS and octanol. $N$ae side chain variants shown in blue, $N$Lys side chain variants shown in lilac.

sequence peptoids than in their $N$ae side chain counterparts. This could be a consequence of backbone rearrangement of the $N$Lys sequences allowing for better shielding of the hydrophobic regions by the side chain groups, resulting in the more consistent values for hydrophobicity. The $N$ae side chain sequences on the other hand may be more conformationally limited due to the reduced side chain length and therefore have less freedom to arrange the charged groups in a way that effectively shields hydrophobic regions.

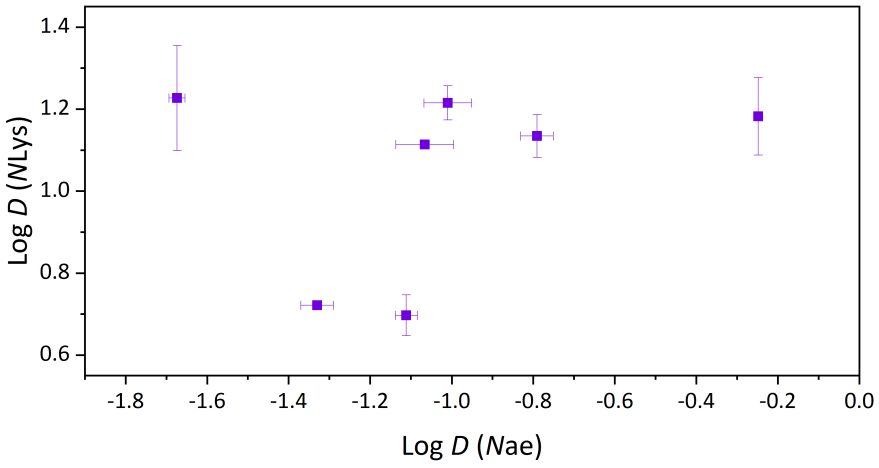



*Fig. 5.8:* Log $D$ values for the scrambled sequence peptoids, with the $N$ae side chain values plotted against the $N$Lys side chain values.

Despite the lack of clarity as to why we observe these differences in Log $D$ for the scrambled sequences, there is nevertheless a clear design rule which has emerged from this study. The folded hydrophobicity of a sequence can be moderated by the relative

proportion of charged and aromatic residues, as observed in the repeat motif sequences. There also appears to be a critical charge:aromatic ratio, which in the case of these linear, 12 residue sequences is 1:2, where the folded hydrophobicity can be moderated by the nature of the charged side chain. The increased variation in Log D for the *N*ae sequences also suggests that there is potential to build more combinatorial libraries with diverse hydrophobicity using this particular side chain rather than *N*Lys. Given the clear distinction in behaviour between the *N*ae and *N*Lys sequences, further sequences incorporating analogous charged side chains with different numbers of methylene groups in the side chain would merit investigation in the future. Additionally, systematic investigation of different side chains could help to reveal further sequence to structure relationships.

### *5.3.2 Interactions with Model Membranes*

Peptoid antimicrobial activity is believed to be attributed at least partially to membrane activity [5]. Even in cases where this is not believed to be the primary cause of cell death, peptoids must be able to traverse membranes in order to access intracellular targets. Therefore it is useful to be able to understand how strongly and in what manner different peptoid sequences interact with membranes. Systems such as lipid vesicles and supported lipid bilayers are used extensively as model membranes in biophysical studies.

In this section we use circular dichroism (CD) spectroscopy to study changes in peptoid secondary structure upon binding to model membranes. We use large unilamellar vesicles (LUVs) as model membranes, which due to their small size (approximate diameter 100 nm), are an ideal model membrane system for use in CD studies as the scattering they cause is not sufficient to distort the peptoid signal and they are equilibrium structures, unlike small unilamellar vesicles (SUVs), which cause less scattering but are only metastable [40]. The size and polydispersity of the LUVs used in this work were characterised by DLS and an example of this data is shown in the discussion of methods in Chapter 2. The changes observed in peptoid secondary structure upon binding to lipid vesicles are more subtle than the conformational switching (from extended to $\alpha$-helix) observed in many AMPs but may nevertheless give insight into the membrane activity of peptoids.

Real biological membranes are extremely complex and contain many different lipids as well as membrane proteins and other features, depending on the organisms in which they are found. One of the key differentiations between mammalian and bacterial membranes is that the former contain mixtures of zwitterionic lipids and have no net charge on the outer leaflet, whereas the latter are often negatively charged. We designed two simple LUV model membranes with net charge as the differentiating factor between them.

We used LUVs composed of pure DOPC as a mammalian-type membrane mimic. DOPC is a phospolipid with a choline head group which is found in all plant and mammal

cells but absent in most bacterial membranes. We used LUVs composed of a mixture of DOPC and DOPS at an 8:2 molar ratio as a bacterial membrane mimic. The DOPS was incorporated in this proportion to mimic the fact that most bacterial membranes contain approximately 20-25% negatively charged lipids with the rest of their composition being zwitterionic lipids and other molecules [56, 57]. It should be noted that real biological membranes are exceedingly complex and we have not attempted to accurately reproduce their compositions here. There are many different lipids which could be combined to create LUVs which more accurately represent the true composition of a mammalian or bacterial type membrane. However for the purposes of this experiment we wish to investigate only the effect of membrane charge on the peptoid-membrane interactions and have therefore selected the described compositions as a very simple way to do this without the complications of varying other membrane properties such as the fluidity, sterol content or curvature. Both DOPC and DOPS are in the fluid phase at 20 °C, at which temperature all of the experiments presented here were carried out. It is assumed that the distribution of DOPC and DOPS across the inner and outer leaflets is even.

By collecting the CD spectra for each of the repeat motif peptoids in PBS, octanol and the two different composition LUVs at a range of concentrations, we aimed to establish whether the peptoids interact with the lipid membranes in a way that causes them to alter their helical conformation. Additionally, this experiment provides insight into whether the peptoids adopt the same or similar conformations in octanol and different lipid composition bilayer environments. This may confirm or oppose the hypothesis that octanol is a good model for the bilayer interior.

Figures 5.9, 5.10 and 5.11 show that the repeat motif peptoids adopt different conformations in the presence of DOPC and DOPC/DOPS vesicles and that these are different again from the PBS and octanol spectra. The lipid concentration for the displayed LUV-bound spectra in these figures varies, as the spectra selected were either for the highest lipid to peptoid ratio investigated or at the lipid concentration above which no further evolution was observed in the case of a spectral plateau. In each of these cases the displayed spectrum represents the maximally bound peptoid conformation within the limits of the experimental set-up but it is important to note that this will not be a true representation of the case where every peptoid monomer is bound to or interacting with the membrane in some respect. Comparison between the overall intensity of the spectra for a particular peptoid binding to different composition LUVs is therefore limited, as the percentage of the total peptoid in the bound conformation may vary between cases, depending on the strength of the interaction.

The motif 1 peptoids appeared to change conformation in the presence of the vesicles to different extents depending on the charged side chain (Figure 5.9). The change in spectral shape between pure PBS and PBS with LUVs was much more pronounced for the

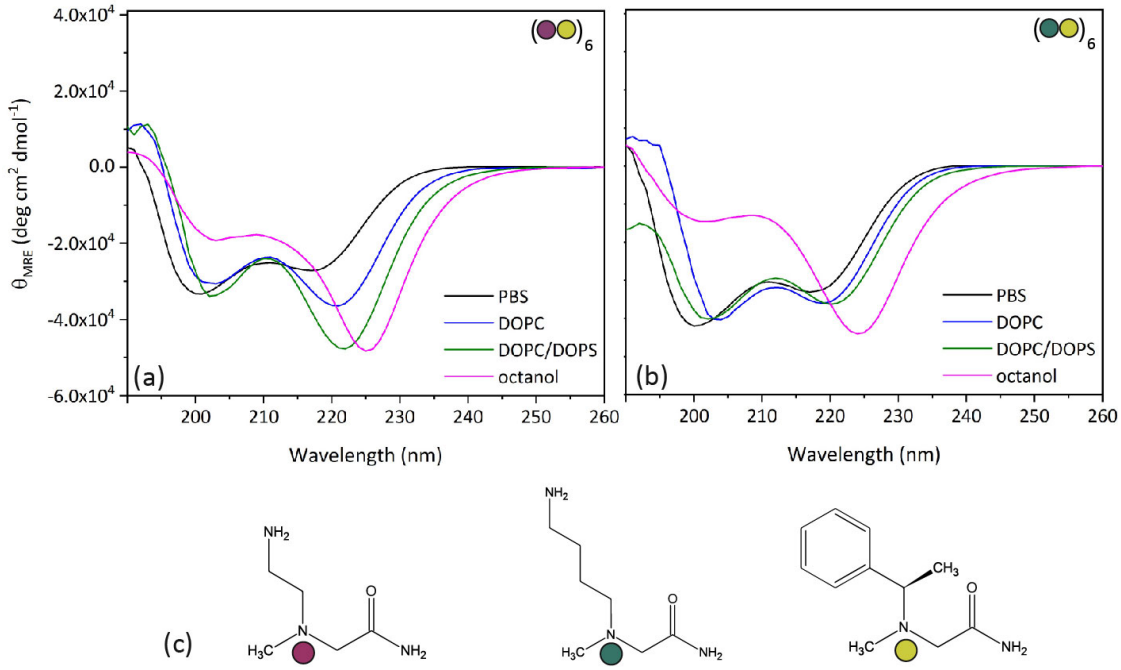



*Fig. 5.9:* CD spectra for motif 1 peptoids in PBS and octanol and in PBS with DOPC and DOPC/DOPS (8:2) LUVs. The peptoid concentration is 25 $\mu$M and spectra were collected at 20 °C. The Y-scale is the same for both peptoids and (a) shows the spectra for RM1 and (b) for RM2. The lipid concentrations for each peptoid with vesicles spectrum is 2500 $\mu$M.

*N*ae variant (RM1) than the *N*Lys variant (RM2). In both cases the observed changes were qualitatively similar to those observed going from PBS to octanol, with a reduction in the intensity of the shorter wavelength minimum with respect to the longer wavelength minimum and a shift in the positions of both minima to longer wavelengths. The magnitude of change in spectral shape and intensity for RM2 was very small, indicating that this peptoid does not interact strongly with either type of LUV or if it does interact, there is no significant structural rearrangement associated with the interaction and therefore it is not detectable by CD. RM1 on the other hand shows considerable conformational change, with the similarities of the LUV spectra to the octanol spectra indicating possible exposure of the peptoid to the hydrophobic environment of the bilayer interior. The extent of the spectral change at equal lipid concentration is greater for the DOPC/DOPS LUVs than the pure DOPC LUVs, indicating that there is a stronger interaction with the former than with the latter, which given their differing lipid compositions, is likely to be charge mediated. This could also indicate that the peptoid adopts subtly different conformations when interacting with the two different vesicles types. The spectra for RM2 with DOPC and DOPC/DOPS LUVs are very similar in shape and magnitude, indicating a similar extent of interaction between the two vesicle types, though this is very small in both cases. This may additionally indicate that this peptoid adopts a very similar conformation when interacting with both types of vesicles.

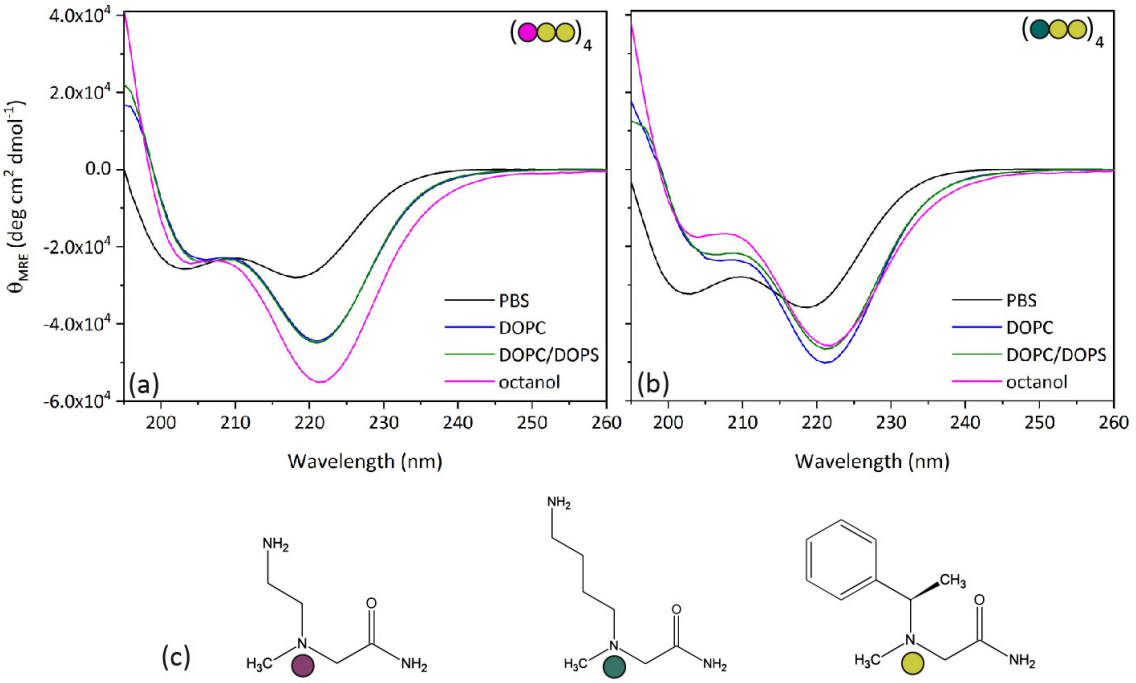



*Fig. 5.10:* CD spectra for motif 2 peptoids in PBS and octanol and in PBS with DOPC and DOPC/DOPS (8:2) LUVs. The peptoid concentration is 25 $\mu$M and spectra were collected at 20 °C. The Y-scale is the same for both peptoids and (a) shows the spectra for RM3 and (b) for RM4. Lipid concentrations for RM3 with vesicles are 2500 $\mu$M and 250 $\mu$M for DOPC and DOPC/DOPS respectively. Lipid concentration for RM4 with both vesicle types is 1400 $\mu$M.

The motif 2 peptoids both showed considerable spectral change between PBS and PBS with LUVs of both varieties (Figure 5.10). The changes in spectral shape were again reminiscent of the changes observed going from PBS to octanol, with the change in the relative intensities and positions of the minima occurring for each peptoid and LUV combination. This indicates that both of these peptoids interact with or bind to both vesicle types. The spectra for RM3 with both LUV types are more intense relative to the PBS spectrum than the equivalent RM4 spectra. This indicates that there is either a greater conformational difference between the PBS and lipid bound-states for RM3 than RM4 or that the relative strength of the interaction is stronger and a larger proportion of RM3 adopts the bound conformation that RM4. The octanol spectrum is also considerably more intense relative to the PBS spectrum for RM3 than RM4. The overall change in spectral shape is quantitatively similar for RM3 and RM4 but the spectra evolve in different ways. In RM3 the increase in intensity of $\lambda_2$ is much greater than the equivalent increase for RM4. The decrease in intensity of $\lambda_1$ in RM3 however is much smaller than that observed for RM4. The combination of these changes result in a similar overall spectral change for both peptoids. Notably, the lipid bound spectra appear very similar for both vesicle compositions (particularly for RM3), indicating that the peptoids may be adopting the same folded conformation in both LUV types.

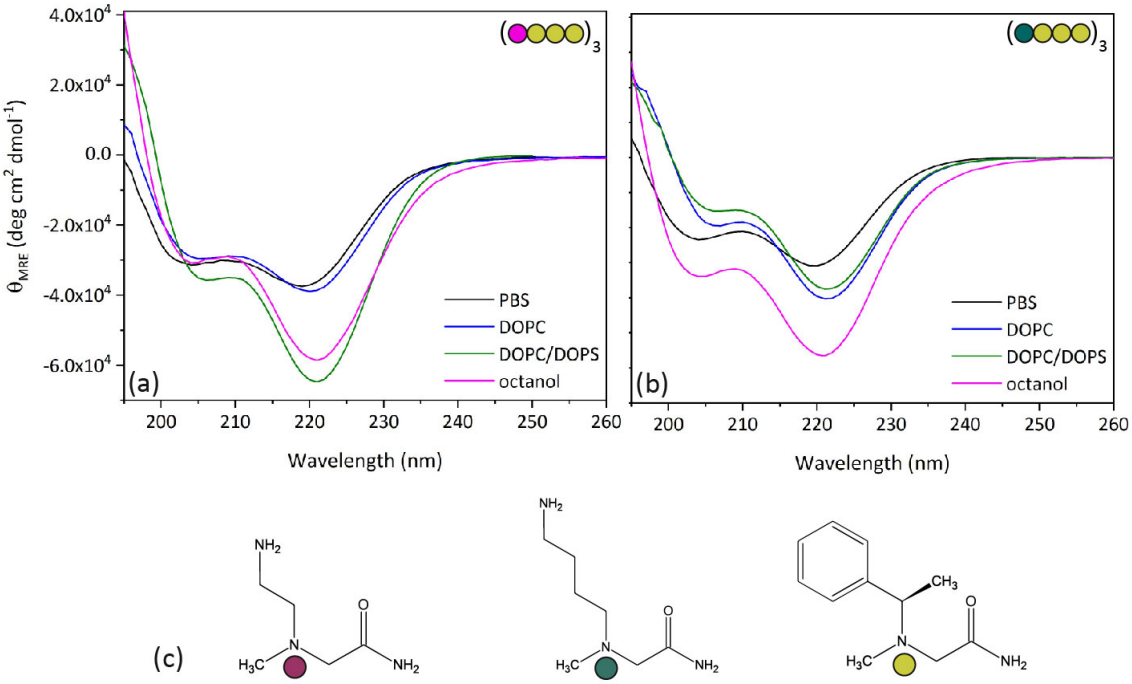



*Fig. 5.11:* CD spectra for motif 3 peptoids in PBS and octanol and in PBS with DOPC and DOPC/DOPS (8:2) LUVs. The peptoid concentration is 25 $\mu$M and spectra were collected at 20 °C. The Y-scale is the same for both peptoids and (a) shows the spectra for RM5 and (b) for RM6. Lipid concentrations for RM5 with vesicles are 2500 $\mu$M and 1400 $\mu$M for DOPC and DOPC/DOPC respectively. Lipid concentration for RM6 with both vesicle types is 2500 $\mu$M.

The motif 3 peptoids did not both interact with each vesicle type (Figure 5.11). The RM6 spectra changed in a similar way to the motif 2 peptoids, though the intensity of these spectra with the LUVs was considerably less than the octanol spectrum. RM5 appeared to (detectably) interact very little with the DOPC LUVs but showed considerable change in spectral shape and intensity in the presence of the DOPC/DOPS LUVs, with the most intense CD signal at the second minimum (approximately 222 nm) of any of the peptoid/lipid combinations investigated. The self-association of this peptoid discussed in Chapter 2 could potentially be responsible for its lack of interaction with the neutral vesicles, while the additional electrostatic interaction with the negatively charged vesicles could help to overcome this and produce the large magnitude of spectral change observed in Figure 5.11.

The spectra shown in Figures 5.9, 5.10 and 5.11 indicate that the majority of the repeat motif peptoids interact with the LUVs in a way that is detectable via CD. However, they do not give any indication of the nature of the interactions and comparison between spectra is also limited by the lack of knowledge of the extent of the interaction in each case. In order to gain additional insight into these interactions the spectral evolution of the peptoids with increasing lipid concentration is examined in the following sections.

### 5.3.3 *Peptoid CD Spectra with Neutral Membranes*

A series of samples of each of the repeat motif peptoids at 25 $\mu$M in PBS were prepared with different concentrations of DOPC vesicles and hence lipid:peptoid ratios. A CD spectrum was collected for each lipid to peptoid ratio. Changes observed in the shape and intensity of the spectra as the peptoids alter their secondary structure with increasing lipid concentration could provide insight into the strength and nature of the interaction between the peptoid and membrane. All of the repeat motif peptoids showed some extent of spectral evolution in the presence of the LUVs, though some far more than others, indicating that these peptoids have varying affinity for the DOPC LUVs.

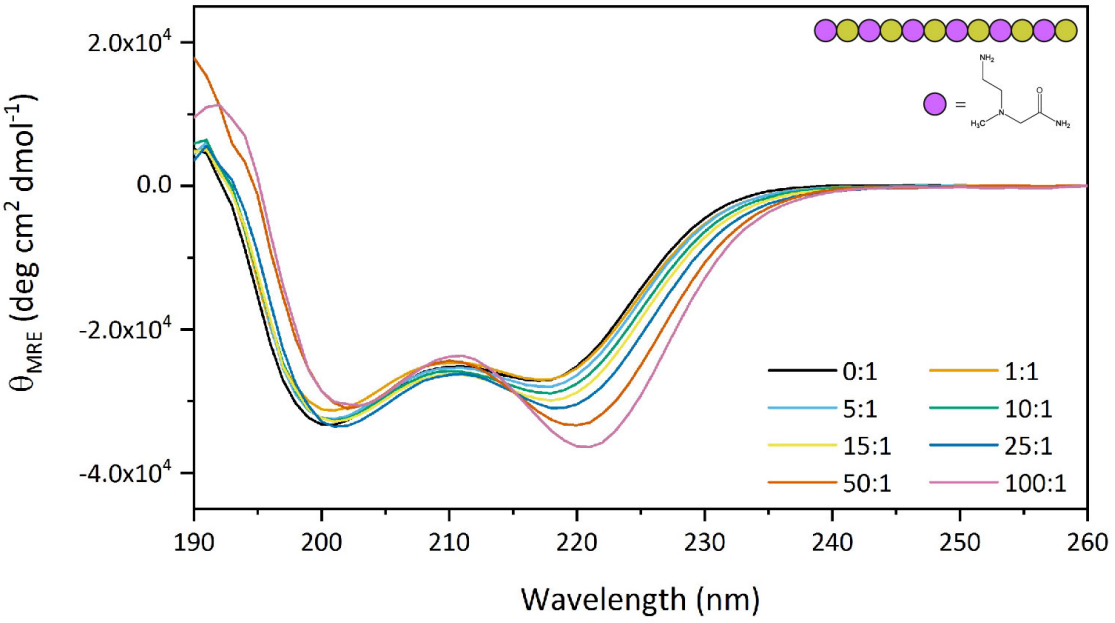



*Fig. 5.12:* CD spectra for peptoid RM1 ((*N*ae*N*spe)$_6$) in 0.01 M PBS with DOPC LUVs at various lipid to peptoid ratios. Peptoid concentration is 25 $\mu$M and lipid concentrations vary according to the lipid:peptoid ratios, displayed in the legend. Spectra were collected at 20˚C.

The spectral evolution of the motif 1 peptoids in the presence of DOPC vesicles was very different for each of the two charged side chain variants. RM1, with the shorter charged side chain (*N*ae) appeared to interact with the vesicles in solution (Figure 5.12). The change in the spectral shape going from the PBS spectrum to the high lipid concentration spectra was reminiscent of that observed going from PBS to octanol. The intensity of the second, longer wavelength, minimum ($\lambda_2$) increased as the lipid concentration increased and the position of this minimum red-shifted by approximately 3 nm in total. The intensity and position of the first, shorter wavelength, minimum ($\lambda_1$) remained relatively constant at low lipid to peptoid ratios then appeared to decrease in intensity at L:P = 50:1 and L:P = 100:1. This was accompanied by a red-shift of approximately 3 nm. The

ratio of the two minima, $R_{2/1}$ (as defined in Chapter 2) therefore increased considerably with increasing lipid concentration. These structural changes indicate that the peptoid interacts with the lipids in a way that causes it to alter its helical configuration. The similarity between the high lipid concentration spectra and the octanol spectrum indicate that the peptoid is at least partially exposed to the hydrophobic environment of the lipid bilayer interior. This implies a degree of binding or insertion. The spectrum continues to change right up to the highest lipid concentration studied indicating that investigating the same system with higher lipid concentrations may reveal further spectral change. This was not pursued further here due to the increased optical density of higher lipid concentration samples affecting the quality of the spectra.

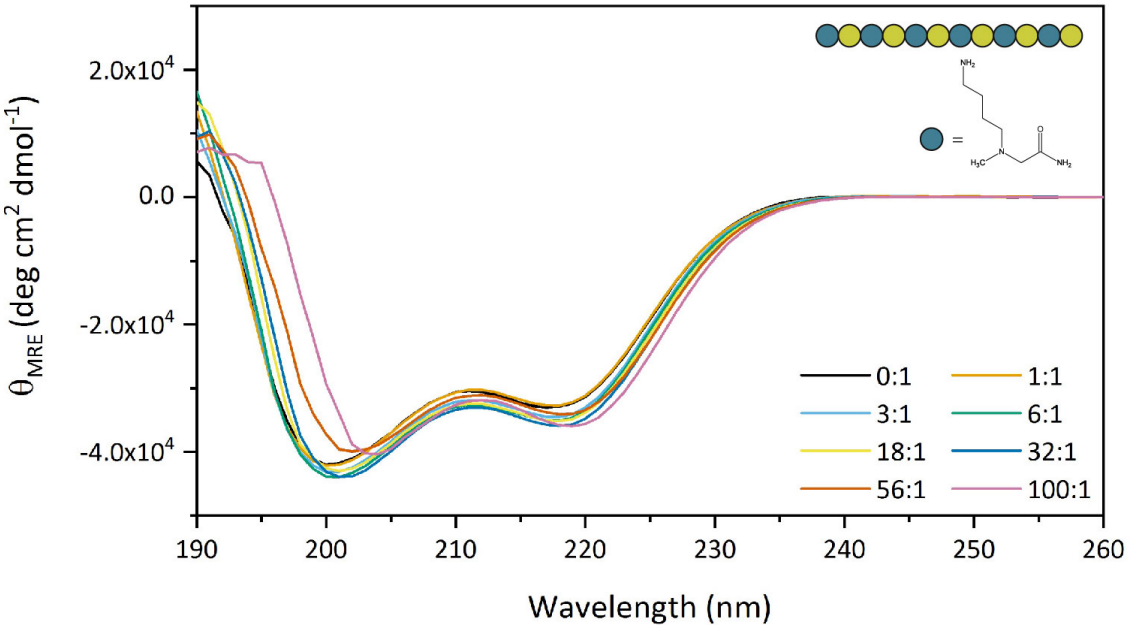



*Fig. 5.13:* CD spectra for peptoid RM2 ($(N\text{Lys}N\text{spe})_6$) in 0.01 M PBS with DOPC LUVs at various lipid to peptoid ratios. Peptoid concentration is 25 $\mu$M and lipid concentrations vary according to the lipid:peptoid ratios, displayed in the legend. Spectra were collected at 20°C.

In contrast, the spectra for RM2, the equivalent sequence with the longer charged side chain variant ($N$Lys), do not show a distinctive or consistent change in shape with increasing lipid concentration (Figure 5.13). There is also very little change in the intensity of the spectra or the positions of the two minima. The spectrum at the highest lipid to peptoid ratio does not resemble the octanol spectrum for this peptoid but maintains a shape characteristic of the motif 1 peptoids in PBS, with the shorter wavelength minimum more intense than the longer wavelength minimum. Small spectral changes do occur for RM2 with these LUVs, though any trends are difficult to distinguish due to the overlapping of many of the spectra. Inspection of each spectrum individually indicated that at the lower lipid to peptoid ratios (1-32) there is a vary small increase in intensity of both

minima as the lipid concentration increases. At the two highest L:P (56:1 and 100:1) the intensity of $\lambda_1$ is reduced and the position red-shifted, though $\lambda_1$ remains less intense than $\lambda_2$ even at these high lipid concentrations.

Though RM2 does not appear to undergo a significant secondary structural rearrangement in the presence of the DOPC LUVs this does not exclude the possibility that the peptoid interacts with the vesicles in some way. The lack of detectable interaction between RM2 and the DOPC LUVs could be attributed to its low hydrophobicity, as its positive Log $D$ value (reported in section 5.3.1) indicate that it would preferentially partition into PBS rather than octanol or presumably, the DOPC bilayer. However, a positive Log $D$ value was measured for RM1, which did interact with the DOPC SUVs, indicating that there is some other driving force for the peptoid-lipid interaction in this instance. If the mechanism of RM2 is purely membrane-disruptive, the lack of interaction with DOPC LUVs could be indicative of low cytotoxicity, although considerable work would be needed to validate pure DOPC vesicles as a reliable model mammalian membrane to confirm this.

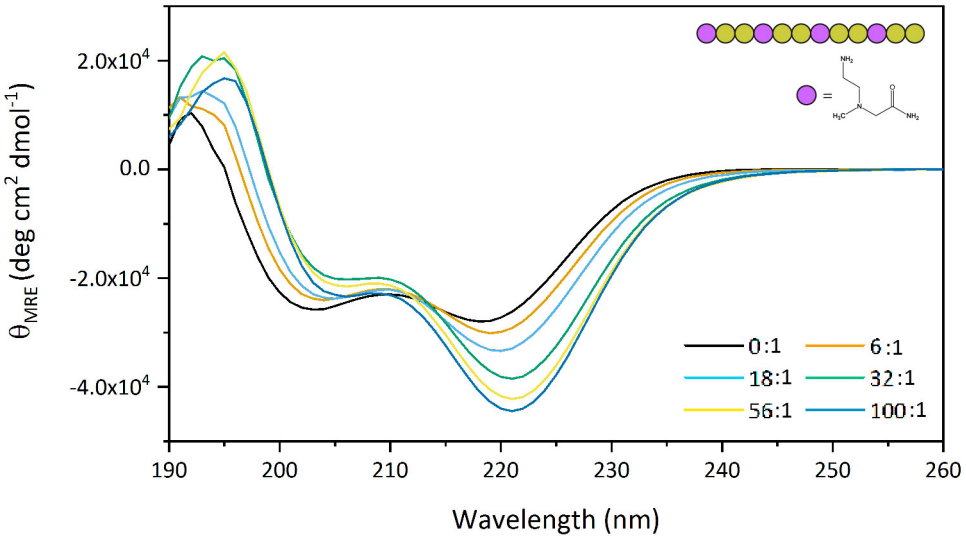



*Fig. 5.14:* CD spectra for peptoid RM3 ($(N\text{ae}N\text{sp}eN\text{sp}e)_4$) in 0.01 M PBS with DOPC LUVs at various lipid to peptoid ratios. Peptoid concentration is 25 $\mu$M for each spectrum and lipid concentrations vary according to the lipid:peptoid ratios, displayed in the legend. Spectra were collected at 20°C.

The motif 2 peptoids both display distinctive changes in CD spectrum and hence their secondary structure at increasing lipid concentrations with DOPC LUVs (Figures 5.14 and 5.15). This indicates that these peptoids both interact with the LUVs in some manner, with the similarities between the lipid-bound spectra and the octanol spectra indicating exposure to the hydrophobic environment of the bilayer interior.

In both cases, the intensity of $\lambda_1$ decreases in the presence of increasing LUV concentration and the intensity of $\lambda_2$ increases in a largely monotonic way. This is accompanied

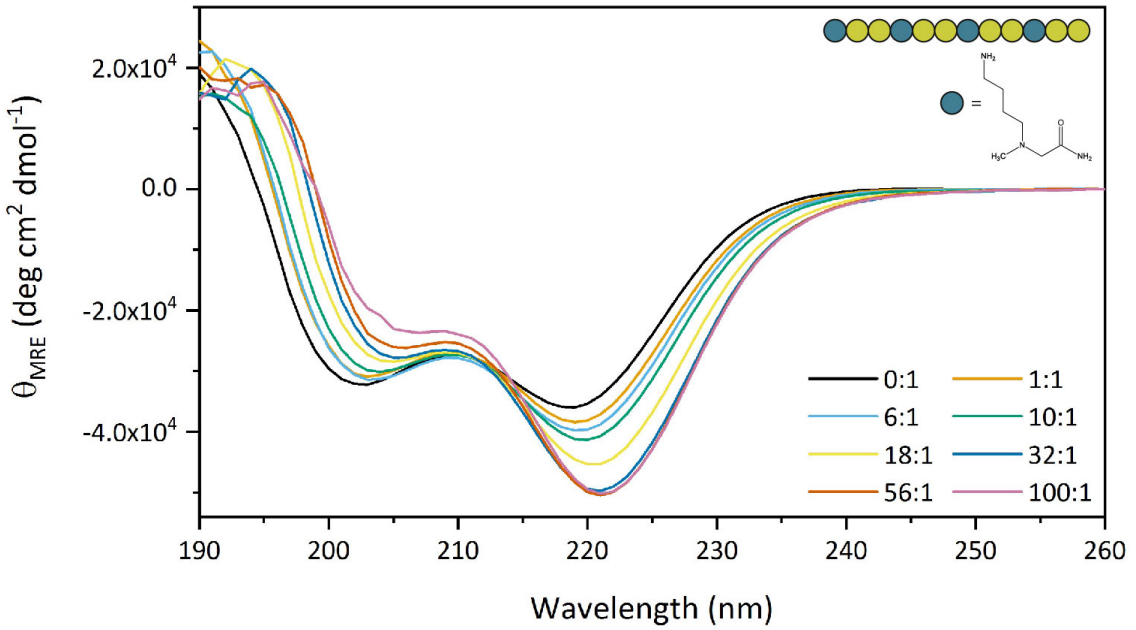



*Fig. 5.15:* CD spectra for peptoid RM4 ((*N*Lys*N*spe*N*spe)$_4$) in 0.01 M PBS with DOPC LUVs at various lipid to peptoid ratios. Peptoid concentration is 25 $\mu$M for each spectrum and lipid concentrations vary according to the lipid:peptoid ratios, displayed in the legend. Spectra were collected at 20°C.

by a gradual red-shifting of the wavelengths of both minima to values approximately the same as those of the octanol spectrum for each of the two peptoids. There is also an isodichroic point present in both these sets of spectra, though it is not as clearly defined for RM3 as for RM4. The presence of an isodichroic point indicates that the interaction of these peptoids with the DOPC LUVs could be modelled as a two state system, where each spectrum is a linear combination of the PBS spectrum and the spectrum for a fully bound peptoid. This implies that the binding of the peptoid induces a conformational switch, from the free state to the bound state and there is no intermediate aggregation stage or conformation, such as have been observed in several AMPs including melittin [58]. CD spectra for melittin and increasing concentration of DOPC LUVs are shown in Figure 5.16. Comparison between the spectra for the motif 2 peptoids and the melittin peptide with DOPC LUVs indicates how the spectral evolution is less dramatic and complex for the former than the latter. This indicates that simple binding models may be applicable to the peptoid spectra.

The possibility of a two-state binding process between the peptoids and LUVs is investigated further in section 5.3.5, where a two-state model is fitted to the data in an attempt to estimate values for free energy of transfer into the membrane and qualitatively compare the affinities of the different sequences with the DOPC LUVs. One notable difference between the two motif 2 peptoids' spectra with DOPC LUVs is that the intensity of $\lambda_2$ for RM3 continues to increase at a steady rate with L:P over the full range studied,

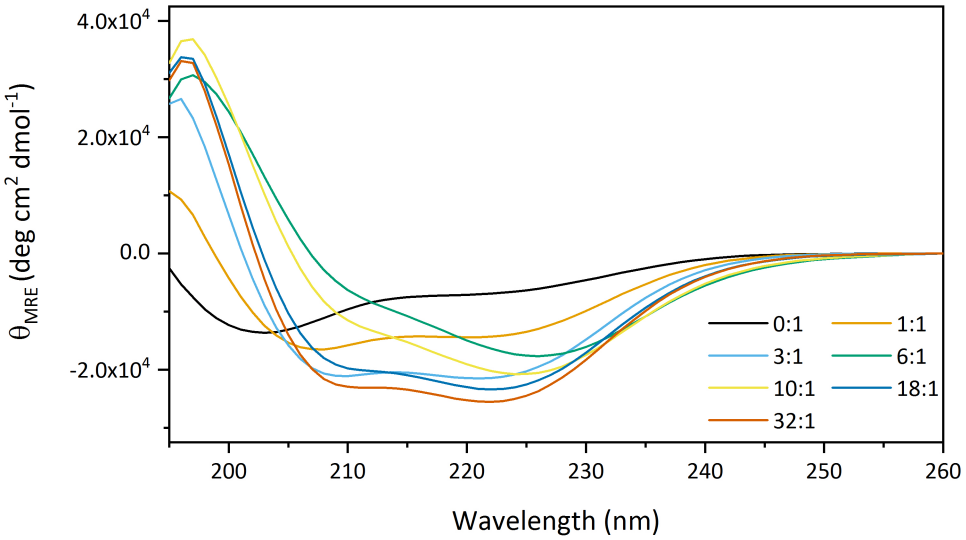



*Fig. 5.16:* CD spectra for the AMP melittin, derived from honey bee venom, in 0.01 M PBS with DOPC LUVs at various lipid to peptoid ratios. The concentration of melittin is 25 $\mu$M for each spectrum and lipid concentrations vary according to the lipid:peptoid ratios, displayed in the legend. Spectra were collected at 20°C.

while $\lambda_1$ plateaus slightly in the higher L:P spectra. The opposite is observed for RM4 where $\lambda_2$ appears to plateau at higher lipid concentrations and $\lambda_1$ continues to decrease in intensity.

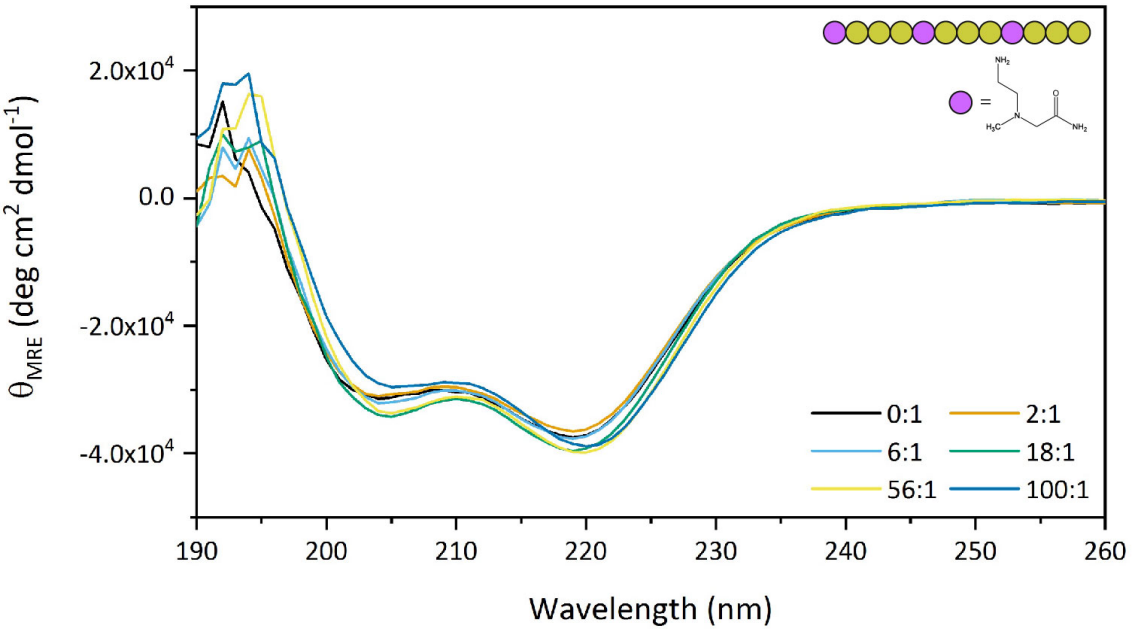



*Fig. 5.17:* CD spectra for peptoid RM5 ((*N*ae*N*spe*N*speNspe)$_3$) in 0.01 M PBS with DOPC LUVs at various lipid to peptoid ratios. Peptoid concentration is 25 $\mu$M for each spectrum and lipid concentrations vary according to the lipid:peptoid ratios, displayed in the legend. Spectra were collected at 20°C.

The spectra for RM5 (Figure 5.17) show very little evolution with increasing lipid concentration and the overlapping of the spectra makes differences difficult to distinguish. However, the intensity of the spectra across all wavelengths (above 200 nm, where the signal to noise is acceptable) does appear to increase marginally and this is reminiscent of the behaviour of RM2, in that at the highest L:P there is a sudden decrease in the intensity of $\lambda_1$ but no corresponding increase in the intensity of $\lambda_2$. Excluding the possibility that the peptoid could interact with the vesicles without undergoing a conformational change, this indicates that RM5 has a very low affinity for the DOPC LUVs, preferring to remain in the PBS environment. The self-association of this peptoid (discussed in Chapter 3) could be preventing interactions with the lipids in this case.

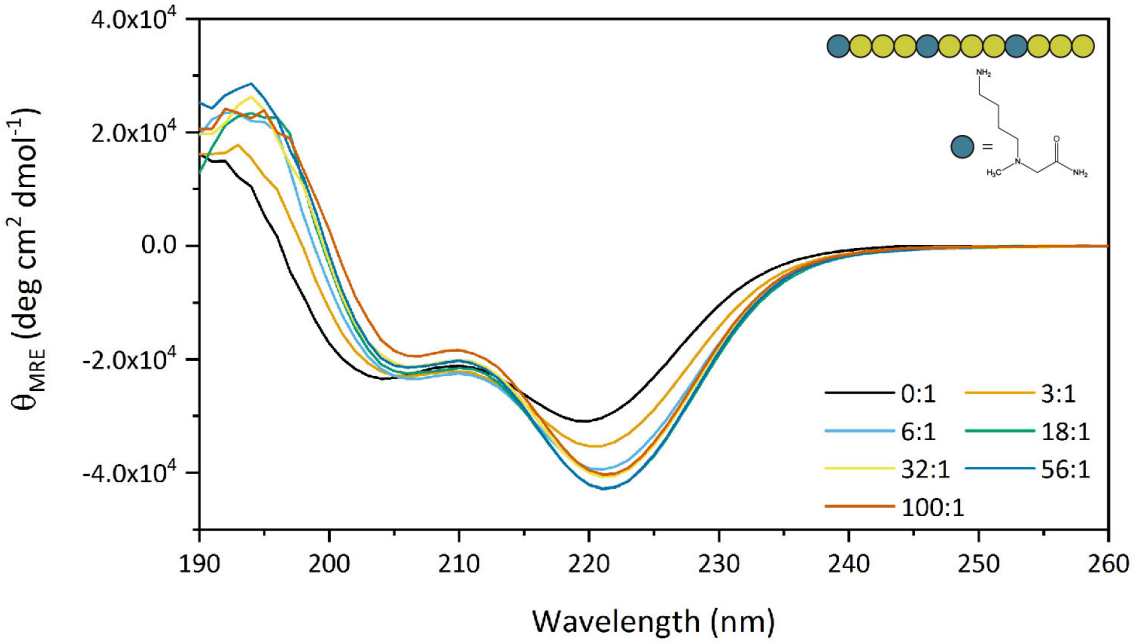



*Fig. 5.18:* CD spectra for peptoid RM6 ((*N*Lys*N*spe*N*speNspe)$_3$) in 0.01 M PBS with DOPC LUVs at various lipid to peptoid ratios. Peptoid concentration is 25 $\mu$M for each spectrum and lipid concentrations vary according to the lipid:peptoid ratios, displayed in the legend. Spectra were collected at 20°C.

Contrastingly, the spectra for RM6 indicate that it interacts with the DOPC LUVs in a way that causes a distinct change in its secondary structure. The spectra for this peptoid, shown in Figure 5.18, evolve in a way that is qualitatively different to that observed for the other repeat motif peptoids. An immediate shift in the position of $\lambda_1$ occurs at the lowest measured L:P (3:1) for this combination, but the position of this minimum then remains roughly constant at higher lipid to peptoid ratios. The intensity of $\lambda_1$ decreases very slowly, but regularly as the lipid concentration increases. The intensity of $\lambda_2$ increases very rapidly over the range of L:P 0-6:1, more gradually over the range 6-16:1 and then is reduced again at the highest L:P measured (100:1). $R_{2/1}$ for this peptoid therefore increases rapidly over the lower range of L:P and less rapidly over the higher range, though it does not appear

to plateau over the full range investigated.

$R_{2/1}$ provides a measure by which to compare the interactions of each peptoid with the LUVs where it is not possible to fit a sensible binding model to the data. The values for $R_{2/1}$ are plotted against L:P for every peptoid/LUV combination in Figure 5.26 and subsequently discussed in greater detail.

### 5.3.4 *Peptoid CD Spectra with Anionic Membranes*

In the following experiments we investigated the interaction of the repeat motif peptoids with the bacterial model LUVs, composed of a mixture of zwitterionic DOPC and negatively charged DOPC at 8:2 molar ratio. CD spectra were collected for each of the peptoids with increasing lipid to peptoid ratios following the same protocol as for the pure DOPC membranes.

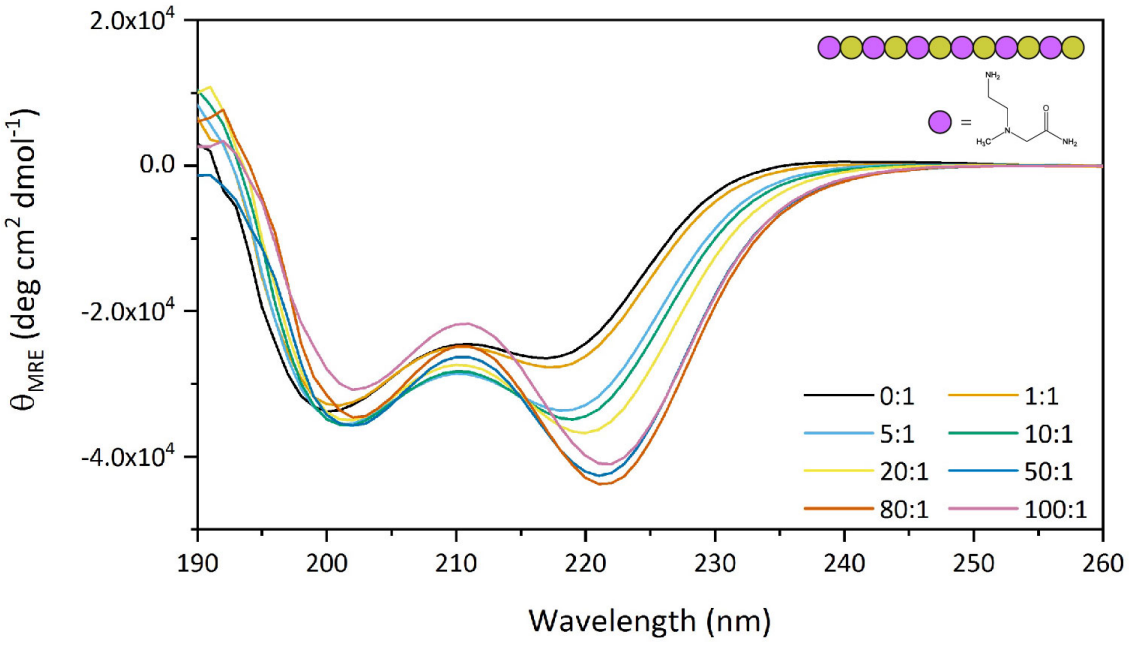



*Fig. 5.19:* CD spectra for peptoid RM1 ($(N\text{ae}N\text{spe})_6$) in 0.01 M PBS with DOPC/DOPS (8:2 molar ratio) LUVs at various lipid to peptoid ratios. Peptoid concentration is 25 $\mu$M and lipid concentrations vary according to the lipid:peptoid ratios, displayed in the legend. Spectra were collected at 20°C.

The spectra for the motif 1 peptoids indicate that these interact more strongly with the charged membranes than the neutral ones, although in the case of the sequence with *N*Lys side chains there is very minimal interaction. The spectrum for peptoid RM1 ($(N\text{ae}N\text{spe})_6$) (Figure 5.19) appears to change shape in a qualitatively similar way for both the DOPC and DOPC/DOPS vesicles. However in the case of the charged vesicles the spectral intensity increases more rapidly with increasing lipid concentration than it does in the case of the neutral vesicles, so ultimately there is a larger overall change across

the same range of lipid concentrations. This increased affinity for the charged vesicles could be due to the contribution of electrostatic interactions between the vesicles and the cationic peptoids to the overall affinity.

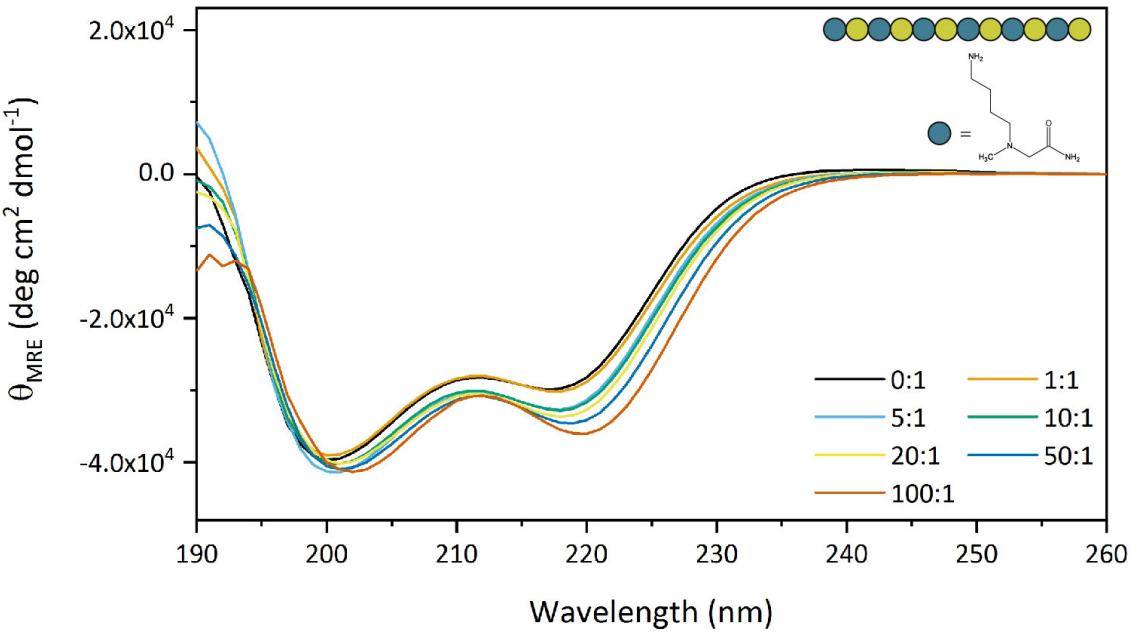



*Fig. 5.20:* CD spectra for peptoid RM2 ((*N*Lys*N*spe)$_6$) in 0.01 M PBS with DOPC/DOPS (8:2 molar ratio) LUVs at various lipid to peptoid ratios. Peptoid concentration is 25 $\mu$M and lipid concentrations vary according to the lipid:peptoid ratios, displayed in the legend. Spectra were collected at 20°C.

There is more spectral evolution for RM2 for the charged vesicles (Figure 5.20) than for the neutral, but the overall change in spectral shape and intensity is still very small and the spectrum at the highest lipid to peptoid ratio does not resemble the octanol spectrum for this peptoid. Instead it retains the characteristic of having the more intense signal at the first minimum than at the second and therefore $R_{2/1} < 0$. The spectral change that does occur for this peptoid-lipid combination over the full range of L:P (1-100) investigated, qualitatively resembles the change that occurs for RM1 with these LUVs, but over only the lower end of the range of L:P (1-20). Within the spectra for RM1 there appears to be two regions of behaviour. Over the lower range of L:P the observed $\lambda_1$ undergoes very little change in intensity or position, however at higher L:P there is a sudden shift to lower intensity and longer wavelengths (for both LUV types).

The motif 2 peptoids show qualitatively similar spectral evolution in the presence of DOPC/DOPS LUVs as they do with the pure DOPC LUVs. Again there is a largely monotonic decrease in $\lambda_1$ and increase in $\lambda_2$ with increasing L:P. The spectra for RM3 (Figure 5.21) with the charged vesicles converge at relatively low L:P, with each spectrum from L:P 10:1 and higher overlaying almost exactly, particularly over the range of wavelengths above 200 nm, where the signal to noise ratio is best. This peptoid appears to

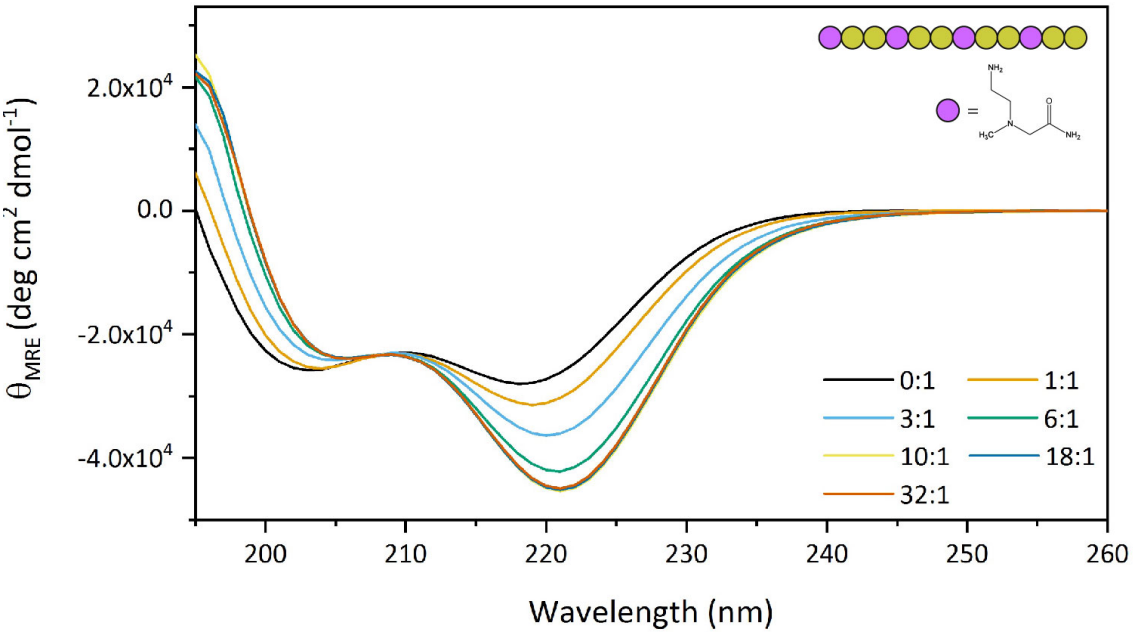



*Fig. 5.21:* CD spectra for peptoid RM3 ((*N*ae*N*spe*N*spe)$_4$) in 0.01 M PBS with DOPC/DOPS (8:2 molar ratio) LUVs at various lipid to peptoid ratios. Peptoid concentration is 25 $\mu$M for each spectrum and lipid concentrations vary according to the lipid:peptoid ratios, displayed in the legend. Spectra were collected at 20°C.

have a much greater affinity for the charged vesicles than the neutral, since a comparable level of spectral change is observed at much lower L:P for the former (10:1), than the latter (100:1). The spectral plateau is unexpected at such low lipid concentrations however.

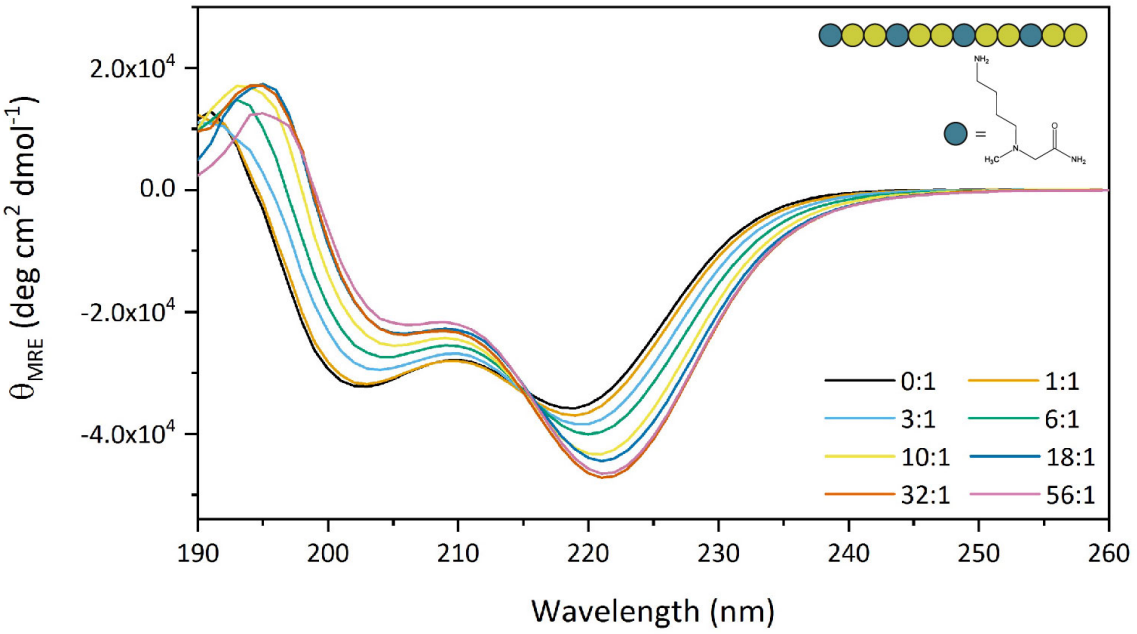



*Fig. 5.22:* CD spectra for peptoid RM4 ((*N*Lys*N*spe*N*spe)$_4$) in 0.01 M PBS with DOPC/DOPS (8:2 molar ratio) LUVs at various lipid to peptoid ratios. Peptoid concentration is 25 $\mu$M for each spectrum and lipid concentrations vary according to the lipid:peptoid ratios, displayed in the legend. Spectra were collected at 20°C.

The spectral evolution for RM4 with increasing concentration of DOPC/DOPS LUVs (Figure 5.22) is also qualitatively similar to that observed for this peptoid with the pure DOPC LUVs. In this instance however, the spectra do not converge fully, as they do for RM3. Though the increase in intensity of $\lambda_2$ between the highest to L:P (32-56) for RM4 is very small, there is still a significant decrease in the intensity of $\lambda_1$ between these two spectra. Given these differences it could be inferred that RM4 has a greater affinity for the charged vesicles than for the neutral vesicles, though not as great an affinity as RM3 has for the charged vesicles. This observation is consistent with the Log $D$ results, which indicate that the $N$ae side chain version of motif 2 has a greater folded hydrophobicity than the $N$Lys version and would therefore be more likely to partition into the bilayer. A positive value for Log $D$ for the latter was reported both earlier in this work and in the literature [52]. This indicates a preference for the PBS phase over the octanol but, despite this, RM4 appears to partition into the LUVs of both compositions to some extent. This indicates that something beyond just the hydrophobicity is driving the interaction between the peptoid and the membrane. In the case of the DOPC/DOPS SUVs this is likely to be an electrostatic interaction.

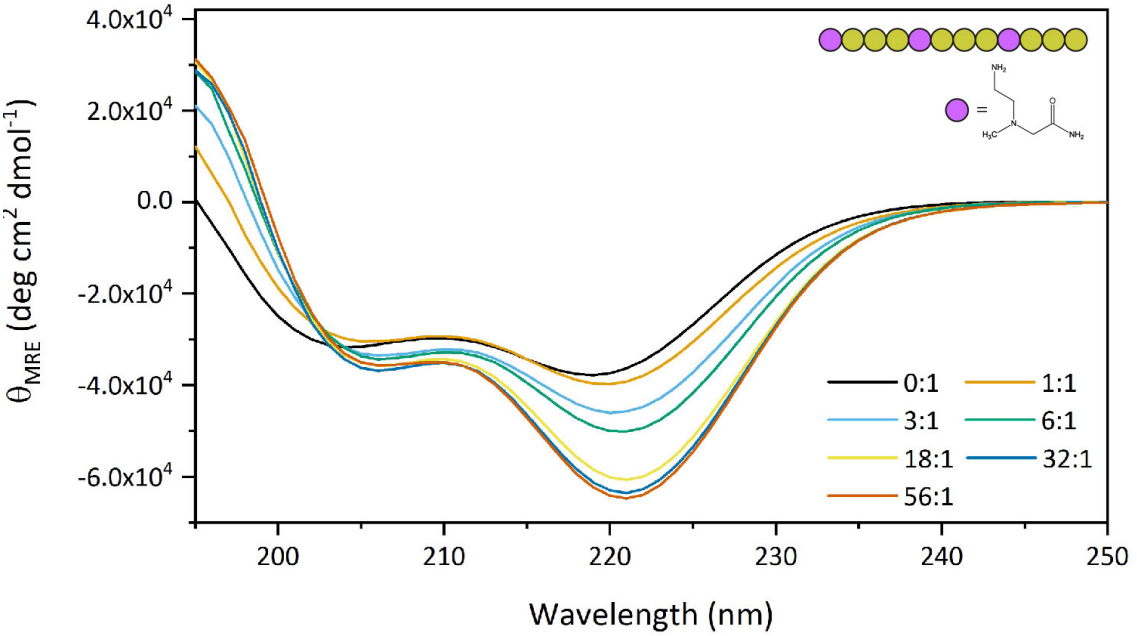



*Fig. 5.23:* CD spectra for peptoid RM5 ((*N*ae*N*spe*N*speNspe)$_3$) in 0.01 M PBS with DOPC/DOPS (8:2 molar ratio) LUVs at various lipid to peptoid ratios. Peptoid concentration is 25 $\mu$M for each spectrum and lipid concentrations vary according to the lipid:peptoid ratios, displayed in the legend. Spectra were collected at 20°C.

The spectra for RM5 and the negatively charged LUVs are shown in Figure 5.23. The change in overall shape and intensity for this peptoid is greater than that observed for any of the other peptoid-LUV combinations. The relative intensities of the two minima in this case do not change in a predictable way over the whole range of L:P investigated.

The differences between the PBS spectrum and the L:P = 0:1 spectrum for this peptoid are similar to those observed for the motif 2 peptoid and lipid combinations, with corresponding increase in $\lambda_2$ and decrease in $\lambda_1$. However, at L:P = 3:1 there is a significant increase in the overall intensity of the spectrum for RM5. At each subsequent higher L:P the intensities of both minima continue to increase. Over the lower L:P range for this peptoid $\lambda_2$ increases faster than $\lambda_1$, so the ratio of the two minima increases dramatically, but this somewhat plateaus at L:P of 18:1 and above. The initial increase in $R_{2/1}$ suggests that the peptoids interacting with the membrane undergo a conformational switch between a free and bound state but the additional increase in intensity of the spectra with increasing lipid concentration could potentially imply that the presence of the LUVs is also promoting the folding of the peptoid.

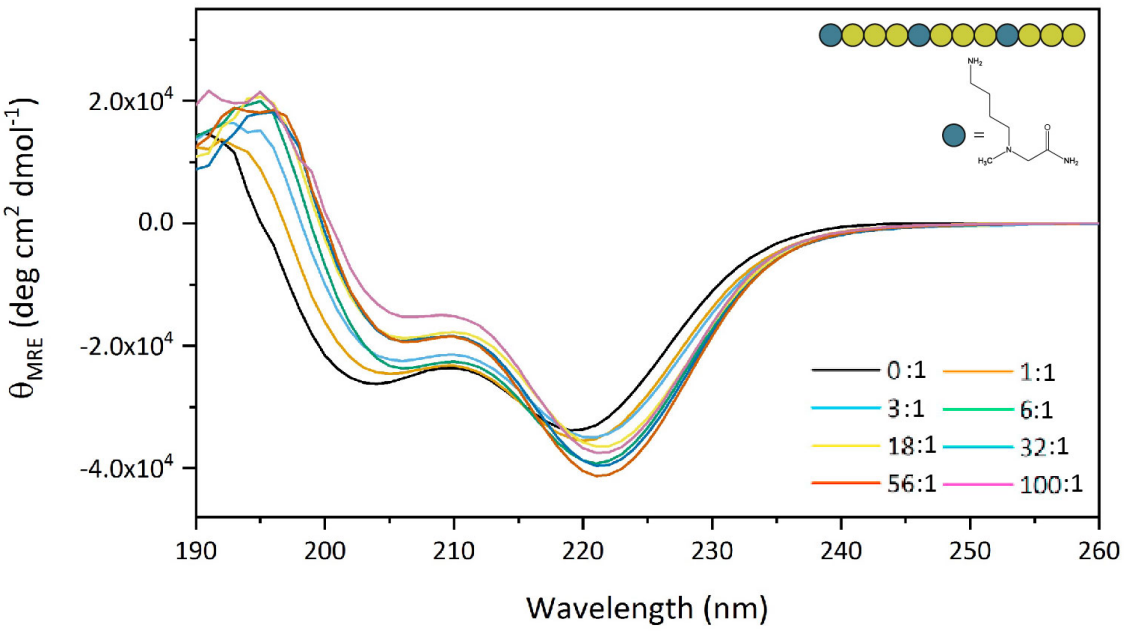



*Fig. 5.24:* CD spectra for peptoid RM6 ((*N*Lys*N*spe*N*speNspe)$_3$) in 0.01 M PBS with DOPC/DOPS (8:2 molar ratio) LUVs at various lipid to peptoid ratios. Peptoid concentration is 25 $\mu$M for each spectrum and lipid concentrations vary according to the lipid:peptoid ratios, displayed in the legend. Spectra were collected at 20°C.

One of the key difficulties associated with the analysis of peptoid CD spectra, given their lack of spectral change over a measurable temperature range, is that it is impossible to determine what fraction of the monomers is in a folded state in equilibrium at room temperature. Given the added difficulties associated with the concentration dependent mean residue ellipticity for this particular peptoid in PBS, which were discussed in Chapter 2, making it even more difficult to determine the normal folded state in PBS, it is very difficult to draw meaningful conclusions from the lipid binding data in this case. It is clear however, that there is some interaction occurring with the DOPC/DOPS LUVs. Given the stark contrast to the data for RM5 with the pure DOPC, where very little spectral

change was observed, it could be speculated that RM5 has a strong enough affinity for the charged vesicles to overcome its self-association, but that this is not true in the case of the neutral vesicles.

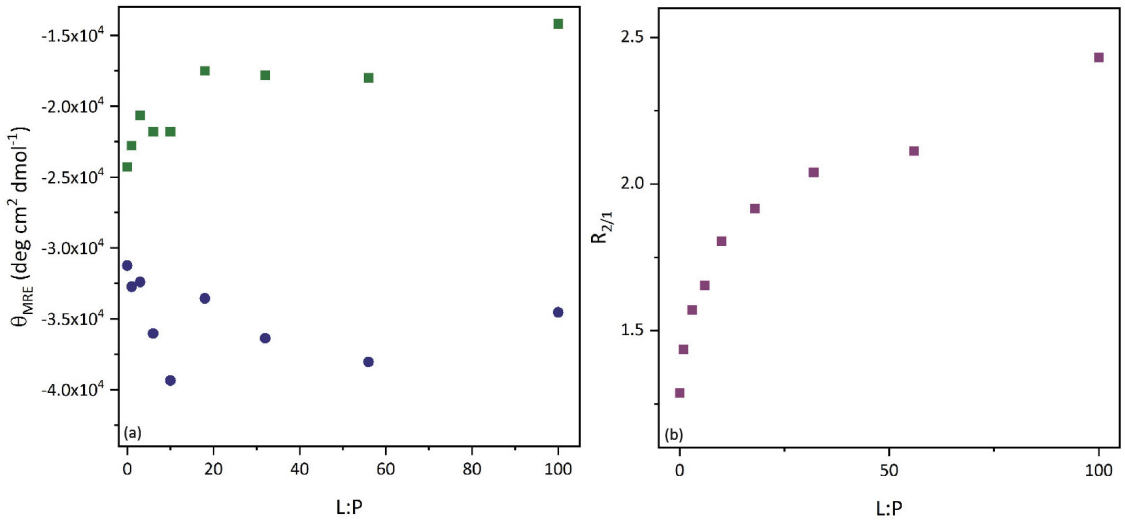



*Fig. 5.25:* (a) The intensity of the first (green squares) and second (blue circles) minima for RM6 against L:P for the peptoid with DOPC/DOPS vesicles. (b) The ratio of the two minima plotted against L:P

The spectra for RM6 and DOPC/DOPS LUVs are shown in Figure 5.24. For this peptoid the intensity of the spectrum for each L:P appears somewhat random, but despite this, $R_{2/1}$ increases consistently with increasing lipid concentration. This is due in part to a much more substantial reduction in the intensity of $\lambda_1$ than increase in the intensity of $\lambda_2$. Due to the overlapping of many of the spectra this data is easier to visualise and examine when plotted as the intensity of the signal at each minimum against L:P and $R_{2/1}$ against L:P. These are both shown in Figure 5.25

In Figure 5.25a it can be seen that the intensity of $\lambda_1$ (green squares) initially increases, then decreases slightly before increasing again at the highest L:P. Meanwhile $\lambda_2$ (blue circles) increases strongly at the lower L:P before a sudden decrease at L:P = 18:1, after which it slowly increases again until finally it increases at the highest L:P. These variations in the intensities of the minima seem too large to attribute to random errors in peptoid concentration due to sample preparation.

The ratio of the two minima at different lipid to peptoid ratios provides an interesting point of comparison between each peptoid and LUV combination. Ignoring any overall intensity effects, $R_{2/1}$, gives an indication of the strength of the interaction of each peptoid with each vesicle type, assuming that the interactions are characterised by a single conformational switch between the PBS state and the lipid bound state (the validity of this assumption is discussed in detail subsequently). $R_{2/1}$ across the full set of peptoids and LUVs is shown in Figure 5.26. It is immediately obvious that without exception, there is

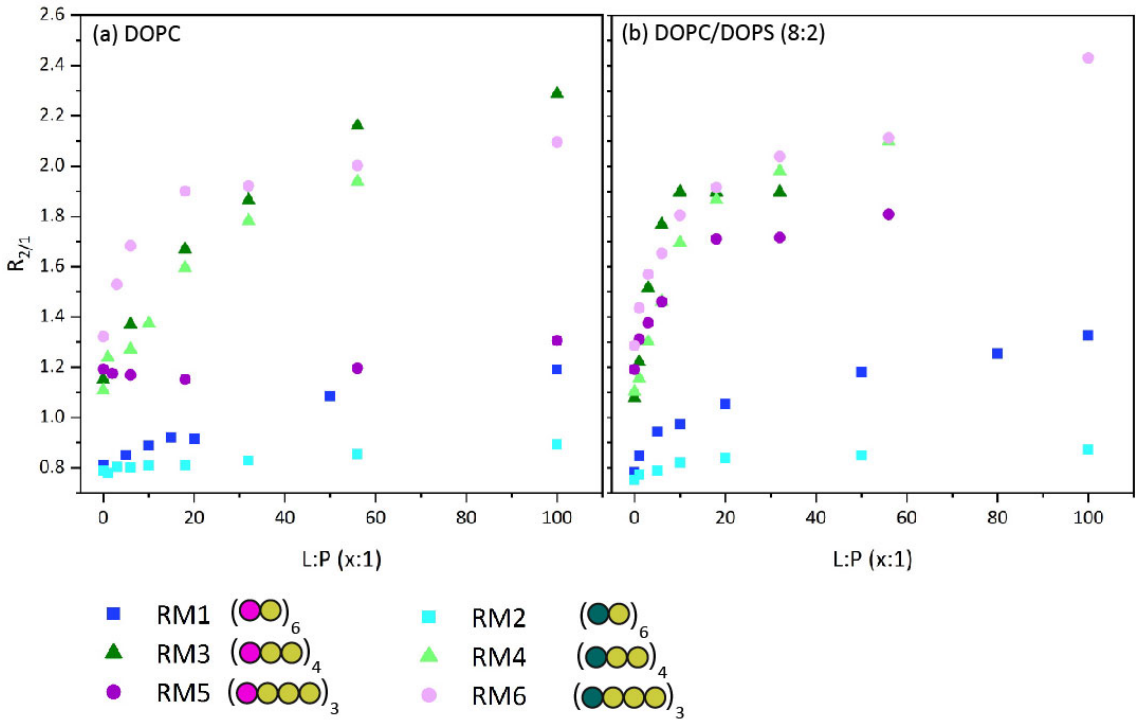



*Fig. 5.26:* R2/1 for each of the repeat motif peptoids, plotted against the lipid to peptoid ratio for (a) DOPC LUVs and (b) DOPC/DOPS (8:2) LUVs. Motif 1 peptoid data is shown as blue shades, motif 2 in green shades and motif 3 in pink shades. In each case spectra were collected at 20°C and the peptoid concentration was 25 $\mu$M, with lipid concentration varying according to the lipid to peptoid ratio, L:P.

a more rapid increase in $R_{2/1}$ at lower L:P for each peptoid with the charged vesicles than with the neutral vesicles, indicating that there is a positive contribution to the interaction due to electrostatics.

### 5.3.5 *Two State Model to Describe Peptoid-Membrane Binding*

The spectral changes observed in these experiments could, in the simplest case scenario, represent the fact that the peptoids adopt a different conformation when bound to or interacting with the lipid membrane, than while free in PBS. Therefore, the spectral evolution with increasing lipid concentration would reflect the fact that as the volume of accessible lipid membrane increases a larger percentage of the peptoid monomers are able to bind to, or interact with the membrane in such a way that they switch from the PBS conformation to the bound conformation. It would be useful to fit a binding model to the CD spectra to quantify the strength of the interactions between the different peptoids and LUV compositions and potentially identify how these relate to the peptoid primary sequence.

Having examined the spectra on a case by case basis however, it is clear that attempting to fit a single model to each peptoid-lipid pairing would not be appropriate or useful,

given the qualitative differences observed in the interaction behaviour. However, several of the sets of spectra do contain an isodichroic point which is a good indication that the interaction could be described by a two state model and therefore attempts to fit such a model, with varying levels of complexity, have been made where suitable.

*Model Derivation*

The peptoid-lipid binding system can be described by a two state model where each peptoid monomer adopts either a free PBS helical structure or an alternative helical structure when interacting with the lipid bilayer. In order to fit this model to the experimental data, the CD spectrum at each lipid to peptoid ratio can be represented as a linear combination of the PBS and lipid-bound spectra, which will vary depending on the proportion of peptoid monomers in each conformation. This relationship can be described by equation 5.2, where $\Upsilon$ represents the mixed state spectrum, $\Theta_{PBS}$ and $\Theta_b$ represent the PBS spectrum and fully lipid-bound spectrum respectively and $\alpha$ is the fraction of peptoid monomers in the PBS state.

$$\Upsilon = \alpha\Theta_{\mathrm{PBS}} + (1-\alpha)\Theta_{\mathrm{b}} \tag{5.2}$$

The accuracy of the model spectrum for each lipid to peptoid ratio using this model will vary depending on the quality of the spectrum used as $\Theta_{\mathrm{b}}$, that is, how closely it matches the true spectrum for the fully bound peptoid. Given that each sample is at equilibrium, it is not possible to measure a spectrum where all of the peptoid is in the true lipid-bound state due to the finite energy difference between that and the PBS state. Therefore $\Theta_{\mathrm{b}}$ must be estimated in some way. The octanol spectrum could potentially be used to this purpose, as an approximation for the lipid-bound spectrum. However, it is not necessarily true that a peptoid should adopt the exact same conformation in octanol and in the lipid environment therefore $\Theta_{\mathrm{octanol}}$ may not represent the best approximation to $\Theta_{\mathrm{b}}$.

Alternatively, the spectrum collected at the highest lipid to peptoid ratio could be used to approximate $\Theta_{\mathrm{b}}$. In the case of most peptoid-lipid combinations, this would be $\Upsilon_{100}$. Since $\Upsilon_{100}$ in reality will itself be a combination of $\Theta_{\mathrm{PBS}}$ and $\Theta_{\mathrm{b}}$ (the true fully bound spectrum,) an additional free parameter, $\beta$, must be introduced to the model to account for this.

The spectrum for $\Upsilon_{100}$ can be expressed as:

$$\Upsilon_{100} = \alpha_{100}\Theta_{\mathrm{PBS}} + (1-\alpha_{100})\Theta_{\mathrm{b}} \tag{5.3}$$

Where $\alpha_{100}$ represents the number of monomers in the PBS state at the 100:1 lipid to peptoid ratio.

Each model spectrum can be expressed as:

$$\Upsilon = \beta\Theta_{\mathrm{PBS}} + (1-\beta)\Upsilon_{100} \tag{5.4}$$

Therefore the model spectra fitted with $\Upsilon_{100}$ as the approximate fully bound state can be expressed in terms of $\beta$ and as $\alpha_{100}$ follows:

$$\Upsilon = (1-(1-\beta)(1-\alpha_{100}))\Theta_{\mathrm{PBS}} + (1-\beta)(1-\alpha_{100})\Theta_{\mathrm{b}} \tag{5.5}$$

As $\Upsilon_{100}$ itself will be a linear combination of $\Theta_{\mathrm{PBS}}$ and $\Theta_{\mathrm{b}}$, as seen in equation 5.4.

Additionally, a linear scaling parameter, $\tau$, can be incorporated into the model as follows where necessary. This allows the model to scale the intensity of each spectrum without altering the overall shape, which accounts for random error in peptoid concentration between samples. This could also account for an additional "invisible" state where the peptoid has interacted with the membrane in some manner that does not result in a conformational change detectable in the CD spectra.

$$\Upsilon = \tau(\beta\Theta_{\mathrm{PBS}} + (1-\beta)\Theta_{\mathrm{b}}) \tag{5.6}$$

This model can be fitted to the experimental spectra for each lipid:peptoid ratio and the quality of the fit tested by calculating the reduced chi squared ($\chi^2$) statistic in each case. Due to the poor signal to noise ratio below 200 nm in our experimental spectra, fitting was only carried out over the range 200-260 nm. Following the fitting of the model to the experimental data, the obtained values of $\alpha$ or $\beta$ (depending on the exact model) can be used to calculate an estimate for the free energy of transfer into the membrane, $\Delta G$. Assuming each peptoid monomer can transfer spontaneously between states a dissociation constant can be defined as follows:

$$\Theta_{\mathrm{PBS}} \leftrightharpoons \Theta_{\mathrm{b}}, K_{\mathrm{d}} = \frac{[\Theta_{\mathrm{b}}]}{[\Theta_{\mathrm{PBS}}]} = \frac{1-\alpha}{\alpha}\frac{V_{\mathrm{PBS}}}{V_{\mathrm{M}}} \tag{5.7}$$

Where $[\Theta_{\mathrm{b}}]$ is the concentration of peptoid monomers in the bound state and $[\Theta_{\mathrm{PBS}}]$ is the concentration in the PBS state. $V_{\mathrm{PBS}}$ is the accessible volume of PBS and $V_{\mathrm{M}}$ the accessible volume of membrane. A simple geometric approximation allows the PBS and membrane volumes to be estimated for each lipid to peptoid ratio and $K_{\mathrm{d}}$ expressed in terms of $\beta$ and $\alpha_{100}$ as follows:

$$K_{\mathrm{d}} = \frac{(1-\beta)(1-\alpha_{100})}{1-((1-\beta)(1-\alpha_{100})}\frac{1}{N_{\mathrm{A}}axC_{\mathrm{L}}} \tag{5.8}$$

Where $N_{\mathrm{A}}$ is Avogadro's number, $a$ is the area per lipid head group in a fluid bilayer, $x$ is the ratio of lipid volume to surface area and $C_{\mathrm{L}}$ is the lipid concentration. From

equation 5.8 $K_d$ and $\alpha_{100}$ can be determined from the values of $\beta$ over the range of lipid concentrations, obtained from fitting the two state model to the experimental spectra. The free energy of transfer is then given by:

$$\Delta G = -\mathrm{R}T\ln K_d \tag{5.9}$$

*Model Fitting to DOPC LUV Data*

*Motif 2 Peptoids*

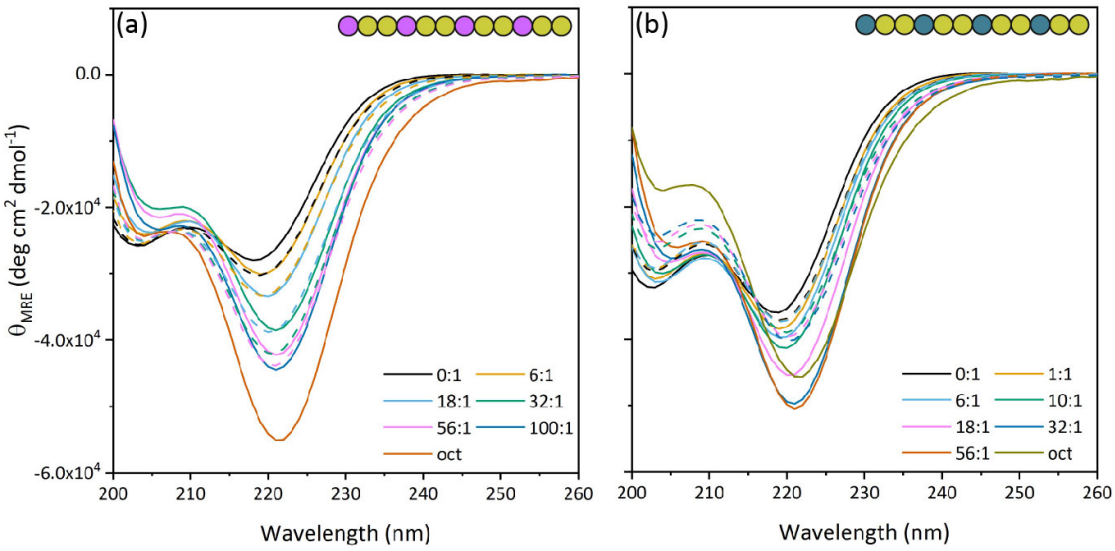



*Fig. 5.27:* Experimental (solid lines) and model (dashed lines) spectra for (a) RM3 and (b) RM4 with DOPC LUVs at various lipid to peptoid ratios (shown in legend). For these model spectra the octanol spectrum was used as a representative fully membrane bound spectrum for each peptoid. The scale on the Y-axis is the same for both sets of spectra. The model spectra are shown in different colours to the equivalent L:P experimental spectra.

We applied the two-state model to the data collected for the motif 2 peptoids as these particular spectra displayed isodichroic points, which are a characteristic feature of a two state system. Initially, we tested whether using the octanol spectrum to represent the fully bound peptoid state for each of the motif 2 peptoids binding to the DOPC LUVs could provide a good fit with this model. In this case we fitted the model spectra to the experimental data according to equation 5.3, with the PBS spectrum as $\Theta_{PBS}$ and the octanol spectrum as $\Theta_b$. The experimental and model spectra are shown overlaid in Figure 5.27. Clearly, the fit is very poor in the case of RM4 (Figure 5.27b). The reduced intensity of the octanol spectrum relative to the higher L:P spectra results in model spectra much less intense than their experimental equivalents. The fit for RM3 (Figure 5.27a) is better but still poor as the the intensity of $\lambda_1$ is much greater in the octanol spectrum than the experimental LUV spectra. It is clear from Figure 5.27 that the octanol spectra for the motif 2 peptoids are not a good representation of the fully DOPC bound spectra for

these peptoids, indicating that the peptoids may adopt distinctly different conformations in these two environments, which are different again from their PBS conformations.

There are several possible physical reasons for this spectral mismatch between octanol and the DOPC environment. The two minima in the octanol spectra for both peptoids are centred around slightly shorter wavelengths than the spectra at high lipid concentrations. This red-shift observed in the membrane environment could be a consequence of the longer alkyl chains in the DOPC tail groups than in octanol, resulting in a lower solvent dielectric constant in the membrane environment, altering the transition energies. This could potentially be corrected for in the model, however it is unlikely to be the sole cause of the poor fit using the octanol spectrum. The octanol spectrum may not be a good substitute for the fully lipid-bound spectrum as the peptoids may not be fully immersed in the hydrophobic alkyl chain region of the lipids, but additionally interact with the lipid head groups. The peptoids are small enough to fully immerse in the bilayer interior but their amphipathic nature may make an inter-facial orientation preferable.

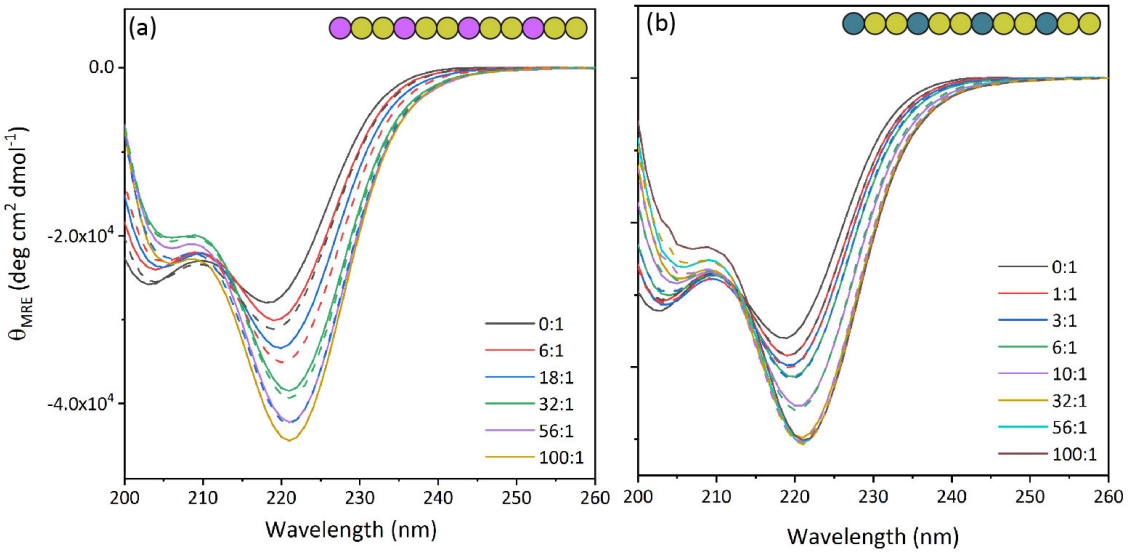



*Fig. 5.28:* Experimental and model spectra for (a) RM3 and (b) RM4 in PBS with DOPC LUVs where the highest L:P measured spectrum was used to approximate the fully bound state spectrum. The model spectra shown are those where a scaling parameter, $\tau$, is included, as this provided the best fit to the experimental spectra.

To achieve a better fit of the model to the experimental spectra we used the adapted model as described by equation 5.4 to enable the use of the highest L:P measured spectrum to represent $\Theta_b$. The model was fitted in both this form and in its scaled form (equation 5.6), which allows adjustment of the intensity of each model spectrum to account for any small experimental errors in peptoid concentration in each case. In a case where the model fits with a large scaling factor, this could also represent the presence of a third conformational state. This scaling provided a better fit in cases where the isodichroic point in the spectra was not particularly well defined.

We applied this model to the spectra for both RM3 and RM4, using the 100:1 spectrum as the representative fully bound spectrum in each case. Figure 5.28 shows the model spectra along with the experimental spectra for these peptoids and DOPC LUVs and Table 5.5 shows the $\beta$ and $\tau$ values from the fitting. This form of the model provides a much better fit to the data than the form using the octanol spectrum as $\Theta_b$. The scaled model also provides a better fit than the standard form. In order to calculate $K_d$ and $\alpha_{100}$, equation 5.7 was rearranged and fitted to the values of $\beta$ obtained from fitting the scaled two-state model, with $K_d$ and $\alpha_{100}$ as free parameters.

For RM3 this resulted in a value of $\alpha_{100}$ of 0.41 $\pm$ 0.08, which indicates that 41% of the peptoid monomers remained in the PBS state at L:P 100:1. This is consistent with the lack of convergence of the spectra over the range of L:P studied, indicating the potential for further spectral evolution at higher lipid concentrations. The fitted $K_d$ was used to estimate the free energy of transfer into the membrane for each peptoid according to equation 5.9, yielding a value of -17.45 $\pm$ 0.27 kJ/mol. These results for RM3 appear to be somewhat anomalous as the fitted values for $\beta$ deviate from the Boltzmann distribution that would be expected to describe the distribution of the peptoids between the two states. This is likely due to a combination of the poor definition of the isodichroic point in the spectra for RM3 and the deviation from a monotonic decrease in intensity of $\lambda_1$ as $\lambda_2$ increases with increasing lipid concentration. Together these indicate that while the two state model can be fitted to these spectra, the interaction may be more complex and not fully captured by this model.

The values of $\beta$ for RM4 fit much better to the expected distribution and $\alpha_{100}$ for RM4 was calculated at 0.14 $\pm$ 0.05 indicating that approximately 14% of the peptoid monomers remained in the PBS conformation at L:P 100:1. The free energy of transfer into the membrane was calculated to be -18.97 $\pm$ 0.19 kJ/mol for RM4. These results indicate that RM interacts more strongly with the DOPC LUVs than RM3. Additionally, the poor fit of the $\beta$ values for RM3 indicate that in this instance the interaction is not described well by the model and may be more complex that a two state system.

### *Motif 1 and 3 Peptoids*

Given that the spectra for the motif 1 and 3 peptoids interacting with DOPC LUVs did not show the distinctive monotonic changes in intensity of the minima and isodichroic points observed in the motif 2 peptoids, the analysis of these sequences was not straight forward. The spectra for RM2, the *N*Lys variant of motif 1 were overlapping and showed no clear trend in spectral evolution such that it was not deemed sensible to attempt to model the behaviour of this peptoid interaction with the DOPC membrane. RM1, the *N*ae version of this motif sequence, showed significant spectral evolution with increasing concentration of DOPC. However, the lack of a single, well defined isodichroic point in

| L:P | $\beta$ (normal) | $\chi^2_\nu$ (normal) | $\beta$ (scaled) | $\chi^2_\nu$ (scaled) | $\tau$ |
|---|---|---|---|---|---|
| **RM3** | | | | | |
| 6:1 | 0.901 ± 0.008 | 15.6 | 0.839 ± 0.008 | 13.4 | 0.97353 ± 0.00236 |
| 18:1 | 0.778 ± 0.011 | 26.7 | 0.63697 ± 0.011 | 14.9 | 0.98376 ± 0.00248 |
| 32:1 | 0.452 ± 0.024 | 14.3 | 0.3028 ± 0.010 | 12.0 | 0.88414 ± 0.00155 |
| 56:1 | 0.389 ± 0.018 | 18.3 | 0.10378 ± 0.006 | 4.9 | 0.95126 ± 0.00452 |
| **RM4** | | | | | |
| 1:1 | 0.837 ± 0.004 | 12.9 | 0.845 ± 0.001 | 3.1 | 1.014 ± 0.001 |
| 6:1 | 0.778 ± 0.013 | 29.4 | 0.806 ± 0.001 | 5.3 | 1.047 ± 0.001 |
| 10:1 | 0.653 ± 0.012 | 28.4 | 0.684 ± 0.003 | 3.4 | 1.044 ± 0.002 |
| 18:1 | 0.381 ± 0.018 | 17.4 | 0.441 ± 0.005 | 11.3 | 1.062 ± 0.002 |
| 32:1 | 0.140 ± 0.024 | 23.3 | 0.241 ± 0.006 | 11.9 | 1.086 ± 0.003 |
| 56:1 | 0.059 ± 0.016 | 24.1 | 0.128 ± 0.006 | 9.7 | 1.054 ± 0.002 |

*Tab. 5.5:* Values of $\beta$ and $\chi^2_\nu$ for normal and scaled model fitted to experimental spectra for RM3 $(N\text{ae}N\text{sp}e N\text{sp}e)_4$ and RM4, $(N\text{Lys}N\text{sp}e N\text{sp}e)_4$, with DOPC LUVs.

these spectra indicates that this may not be a clear case of the peptoid transitioning between two distinct states. Nevertheless, fitting of the two state model, using the 100:1 L:P spectrum as the representative bound state spectrum was attempted, in both the normal and scaled form. Neither form of the model provided a good fit to the data with values of $\chi^2_\upsilon$ being in the region of 16-41 across the range of L:P. Even so, the scaled model $\beta$ values were fitted for $K_\text{d}$ and $\alpha_{100}$ and $\Delta$G was calculated at -16.38 ± 0.31 kJ/mol.

Similar to RM2, the *N*ae variant of motif 3, RM5, showed little systematic signs of spectral evolution with increasing lipid concentration. RM6 however, showed a rapid increase in spectral intensity at $\lambda_2$ at relatively low lipid concentrations with DOPC LUVs. The spectra for this peptoid also exhibit what could potentially be an isodichroic point, though it is poorly resolved and the region of spectral overlap spans approximately 4 nm. Therefore this particular lipid and peptoid combination might merit further analysis via fitting of the two state model. However, selecting a spectrum to represent the fully membrane-bound state proved particularly problematic in this instance due to the unusual behaviour at higher L:P. The highest L:P spectrum may prove a poor representation of the fully bound spectrum as, though the spectra continue to change in shape (increasing $R_{2/1}$) with increasing lipid concentration, there is a sudden reduction in spectral intensity for L:P 25:1 and above. The octanol spectrum for this peptoid is considerably larger in magnitude over the full wavelength range than any of the lipid-bound spectra collected, but the high L:P ratio spectra exhibit a red-shift in the positions of the two minima to longer wavelengths than those of the octanol spectrum. The octanol spectrum would therefore represent a poor model for the fully DOPC-bound spectrum.

The normal model using the L:P 100:1 spectrum provided a very poor fit to the data as it was unable to capture the increased magnitude of the 18:1 and 25:1 spectra relative

to the representative maximum bound spectrum. While the scaled model provided a somewhat improved fit, with $\chi^2_\upsilon$ values in the range of 15-39, it is clear that the binding behaviour in this instance is more complex than can be captured by this model. Indeed the values of $\beta$ from fitting vary non-systematically with the lipid concentration and equation 5.8 provides a very poor fit to these, with a very large error in the value of $K_\mathrm{d}$ and $\alpha_{100}$ = 0. It seems unlikely that the reduced signal intensity at higher lipid concentrations could be attributed merely to errors in peptoid concentration during sample preparation as this behaviour was observed over repeats of the experiment. The variations in intensity could indicate that there is perhaps some threshold concentration for the peptoids in the membrane above which they undergo an additional structural change. In either case, understanding this more complex process is beyond the scope of the two state model.

### *Model Fitting to DOPC/DOPS LUV Data*

#### *Motif 2 Peptoids*

The model was tested with the data collected for the motif 2 peptoids with DOPC/DOPS LUVs due to the presence of isodichroic points in these spectra. Again, using the octanol spectrum as a representative fully bound spectrum provided a poor fit to the data and therefore the model was applied in the form shown in equation 5.4 and the spectra fitted for values of $\beta$. The experimental and model spectra for RM3 and RM4 are shown in Figure 5.29, in parts (a) and (b) respectively. The $\beta, \tau$ and $\chi^2_\upsilon$ values are given in Table 5.6.

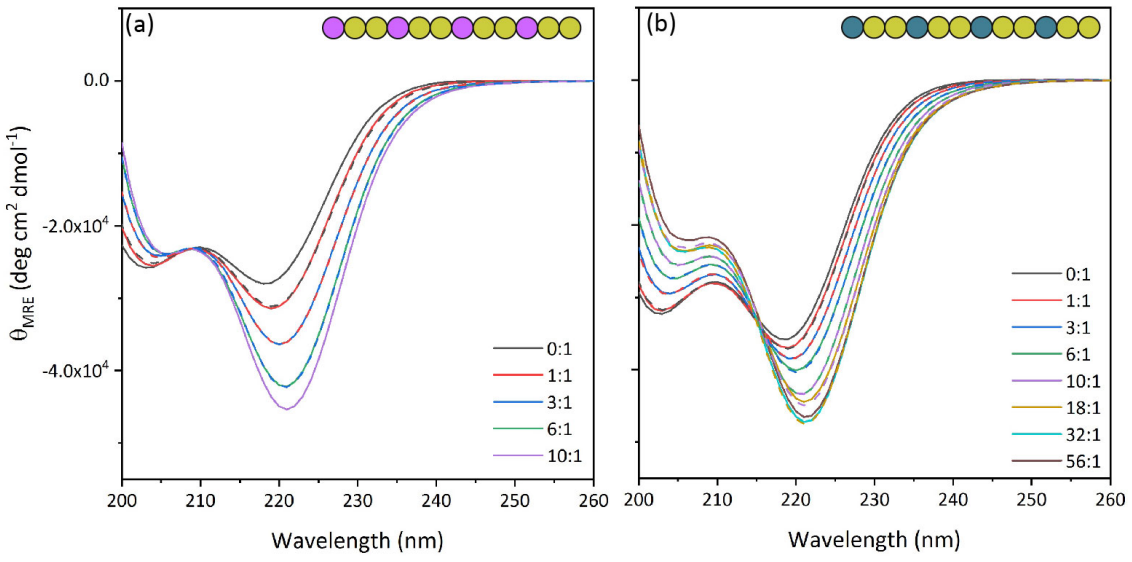



*Fig. 5.29:* Experimental and model spectra for (a) RM3 and (b) RM4 in PBS with DOPC/DOPS LUVs. Model spectra are shown as dashed lines in different colours to the experimental spectra (solid lines) for clarity. L:P for each experimental spectrum is shown in the legend.

Given the spectral plateau above L:P 10:1 for RM3 with the DOPC/DOPS vesicles, there was nothing to gain from fitting the final two highest L:P spectra and therefore the 10:1 spectrum was used as the representative fully bound spectrum. This resulted in only 3 fitted values for $\beta$ but these were nevertheless used to fit $K_{\mathrm{d}}$ and $\alpha_{10}$ according to equation 5.7. The value calculated for $\Delta$G from this analysis was -20.29 $\pm$ 0.01 kJ/mol. However the fitted value of $\alpha_{100}$ was 0.58 $\pm$ 0.003 which indicates that only approximately 42% of the peptoid monomers were in the bound state, which given the strength of the interaction and spectral plateau, seems unrealistically low.

There are several physical factors which may contribute to a simple two state model being less appropriate for modelling the peptoid binding to the DOPC/DOPS LUVs than the pure DOPC LUVs. Electrostatic interactions may result in the cationic peptoids interacting exclusively with the negative lipid head groups while maintaining their PBS conformation. There may also be crowding effects at low lipid concentrations due to the enhanced affinity of the peptoids for the DOPS lipids over the DOPC lipids due to the electrostatic interaction. In this scenario, at low L:P ratios the electrostatic interaction allows some of the peptoid to bind interfacially to the small amount of available DOPS, remaining in its PBS conformation, while the hydrophobic effect would drive an additional fraction of the total peptoid population to insert into the bilayer interior. At higher L:P ratios the accessible area of negatively charged lipid head groups will be greater and therefore a larger proportion of the peptoids may preferentially remain in the PBS conformation while interacting with the DOPS head groups. This would explain to some extent the observed spectral plateau. Unfortunately, if true, this complicates the analysis of the binding affinity as the interfacial binding conformation would represent a third state which is not accounted for in the two state model and is also impossible to distinguish from the PBS spectrum, if indeed these two conformations are very alike. This is very difficult to confirm from the data presented here. However, the use of additional experimental techniques such as linear dichroism or oriented circular dichroism (which can both provide information about the orientation of the peptoids relative to the plane of the lipid bilayer) may provide a promising line of investigation in future work.

The scaled two-state model provided a reasonable fit to the spectra for peptoid RM4 binding to the DOPC/DOPS LUVs. This fit is shown in Figure 5.29b. Fitting equation 5.7 to the obtained $\beta$ for the scaled fit resulted in a free energy of transfer estimate of -21.90 kJ/mol, corroborating the conclusion from the visual interpretation of the spectra that this peptoid has a stronger affinity for the negatively charged LUVs than the neutral LUVs. The $\alpha_{56}$ value obtained from the fit was 0.10 $\pm$ 0.04, indicating that approximately 89% of the peptoid monomers had adopted the bound conformation at the highest measured lipid concentration.

These results indicates that RM4, the *N*Lys variant of motif 2 had a greater affinity

| L:P | $\beta$ (normal) | $\chi^2_\nu$ (normal) | $\beta$ (scaled) | $\chi^2_\nu$ (scaled) | $\tau$ |
|---|---|---|---|---|---|
| **RM3** | | | | | |
| 1:1 | 0.833 ± 0.009 | 2.8 | 0.796 ± 0.008 | 0.25 | 1.002 ± 0.007 |
| 3:1 | 0.490 ± 0.005 | 1.4 | 0.482 ± 0.004 | 0.72 | 0.998 ± 0.004 |
| 6:1 | 0.168 ± 0.005 | 0.6 | 0.150 ± 0.003 | 0.35 | 0.994 ± 0.003 |
| **RM4** | | | | | |
| 1:1 | 0.911 ± 0.006 | 8.1 | 0.916 ± 0.005 | 5.8 | 1.017 ± 0.004 |
| 3:1 | 0.744 ± 0.006 | 5.4 | 0.750 ± 0.005 | 3.7 | 1.014 ± 0.003 |
| 6:1 | 0.550 ± 0.007 | 6.9 | 0.555 ± 0.006 | 5.8 | 1.011 ± 0.003 |
| 10:1 | 0.321 ± 0.011 | 16.8 | 0.330 ± 0.009 | 11.5 | 1.023 ± 0.004 |
| 18:1 | 0.127 ± 0.004 | 2.3 | 0.127 ± 0.004 | 2.3 | 1.000 ± 0.002 |
| 32:1 | 0.058 ± 0.012 | 18.5 | 0.745 ± 0.006 | 4.3 | 1.041 ± 0.003 |

*Tab. 5.6:* Values of $\beta$ and $\chi^2_\nu$ for two state model fitted to experimental spectra for RM3, $(N\text{ae}N\text{spe}N\text{spe})_4$ and RM4 $(N\text{Lys}N\text{spe}N\text{spe})_4$, with DOPC:DOPS (8:2) LUVs.

for the DOPC/DOPS LUVs than its $N$ae counterpart, RM3. This is the opposite to what was concluded from the qualitative investigation of the spectra. Given that the Log $D$ values for the motif 2 peptoids are of opposite sign and indicate a much higher folded hydrophobicity for RM3 than RM4, which did not partition preferentially into the octanol phase, this enhanced affinity for the lipid environment is somewhat unexpected. However, it is possible that the longer butyl group of the $N$Lys side chain relative to $N$ae's ethyl group (which has two less carbons in the chain) allow for a structural arrangement that maximises the exposure of the charged side chain groups and therefore there may be a stronger electrostatic contribution to the interaction with negative LUVs for RM4 than RM3. However, if this was the case, a spectral plateau might be expected but is not observed.

### *Motif 1 and 3 Peptoids*

The spectra for RM1 with the DOPC/DOPS LUVs indicate that the interaction between this peptoid and lipid composition is more complex than a simple conformational switch between two distinct states. Within the mid range of L:P (10:1-80:1) there appears to be a reasonably well defined isodichroic point at approximately 214 nm and these spectra all roughly overlap across an additional wavelength region from 200 nm - 205 nm. However the L:P 0:1, 1:1 and 100:1 spectra do not overlap with the rest in these regions. The two-state model did not fit well to either RM1 or RM2 and therefore further analysis of the binding affinities of these two peptoids with the DOPC/DOPS has not been carried out.

The spectral evolution of RM6 with DOPC/DOPS LUVs is particularly complex and exhibits significant unpredictable intensity variation as well as shape variation observed via the changing ratio of the minima. The octanol spectrum appears unsuitable as a

model fully-bound spectrum and the intensity of the highest lipid concentration spectrum is much lower than many of the lower lipid concentration spectra. Attempting to fit the model with the highest L:P spectrum as the representative bound spectrum resulted in a fit that did not converge, leading to the conclusion that the behaviour of this peptoid interacting with DOPC/DOPS is more complex than can be described by a simple model.

The two-state model was applied with some success to the spectra for RM5 with DOPC/DOPS LUVs. The normal model provided a poor fit to the experimental data but the fit for the scaled model was reasonable (but not excellent at each L:P) with $\chi^2_\nu$ values in the range of 3-25. The $\beta$ values scaled with lipid concentration such that the form of $\beta$ given by equation 5.8 provides a reasonable fit, resulting in a $\Delta G$ estimation of -23.99 $\pm$ 0.03 kJ/mol, the strongest calculated for any of the peptoid-lipid combinations. Given the particularly complex behaviour of this peptoid in PBS, the poor fit of the model to the experimental data could be a consequence of larger than normal fluctuations in the peptoid concentration between each sample. There is the added complication that variation in this peptoid's concentration induces a change in the CD spectrum independent of that resulting from any interaction with the lipid. Therefore the analysis of this interaction is somewhat compromised by the inability to separately resolve these effects.

### *Discussion of Results in Relation to Biological Activity*

A summary of the $\Delta G$ for the peptoids with the different LUVs is given in Table 5.7. Unfortunately, given the increased complexity of the interaction between some of the peptoids and the LUVs it was not feasible to sensibly estimate $\Delta G$ for every peptoid-lipid combination. However, where it was possible, the values can give us some insight into the relationship between peptoid structure and lipid affinity. Additionally, comparing this data to biological studies of the antimicrobial and antiparasitic activity could show whether these biophysical experiments may be a useful predictive tool for these.

| Peptoid | $\Delta G$ (kJ/mol) DOPC | $\Delta G$ (kJ/mol) DOPC/DOPS |
|---|---|---|
| $(N\text{ae}N\text{spe})_6$ | -16.38 $\pm$ 0.31 | - |
| $(N\text{ae}N\text{spe}N\text{spe})_4$ | -17.45 $\pm$ 0.27 | -20.29 $\pm$ 0.01 |
| $(N\text{Lys}N\text{spe}N\text{spe})_4$ | -18.79 $\pm$ 0.19 | -21.90 $\pm$ 0.24 |
| $(N\text{ae}N\text{spe}N\text{spe}N\text{spe})_3$ | - | -23.99 $\pm$ 0.03 |

*Tab. 5.7:* $\Delta G$ calculated for repeat motif peptoids and mammalian/bacterial model membranes where two-state model for binding was applicable.

In the case of the motif 2 peptoids it was possible to calculate $\Delta G$ for both cationic side chain variants with both LUV compositions. The $N$Lys variant (RM4) appeared to have a greater affinity for both the bacterial and mammalian model membranes than its $N$ae counterpart (RM3). Correspondingly, RM4 had lower MIC and $ED_{50}$ values against

the bacterial and mammalian cell lines discussed in section 5.2.1 than RM3. The greater affinity of both peptoids for the negatively charged LUVs than the neutral LUVs also correlates with the peptoids' enhanced activity against bacterial cells versus mammalian cells [52]. In this respect it appears that the LUV binding experiments were a successful predictor of the peptoids' antimicrobial activity.

However, the activity of these sequences does vary considerably against different pathogenic species. Neither of the motif 2 peptoids were strongly active against gram negative bacteria, indicating that a more complex model membrane may be required to study this from a biophysical perspective. One of the defining features of the gram negative cell wall is the external lipopolysaccharide (LPS) layer. An investigation into the binding of peptoids to LPS free in solution could provide insight into whether this affects the ability of the peptoids to permeate the lipid bilayer. LPS has previously been found to induce folding in AMPs which are active against gram negative bacteria [59]. The correlation between the values of $\Delta G$ and the activity of the motif 2 peptoids against gram positive bacteria suggests that this type of bacterial membrane may be modelled better by simple bilayers than gram negative bacteria or parasites. The low $IC_{50}$ for the motif 1 sequences against the parasite *P.falcipicarum* contrasting with their very weak interaction with the LUVs also indicates that this may be the case.

The motif 1 data also indicates that some peptoids may have antimicrobial properties stemming from non-membrane disruptive mechanisms *in vivo*. In the case of $(N\text{Lys}N\text{spe})_6$ the peptoid appears to interact minimally with the LUVs but has been shown to be biologically active against *P. falciparum*. Therefore this peptoid may have a non-membrane active mechanism and target intracellular features. Further investigation into the interactions between peptoids and intracellular targets may help to confirm this.

It would be interesting to consider whether the biological activity of the peptoids correlates to how typically helical their lipid-bound CD spectra are. Chongsiriwatana *et al* [29] noted in their study of a series of linear sequences, including $(N\text{Lys}N\text{spe}N\text{spe})_4$, that the helical intensity of CD signal was poorly correlated to antibacterial activity but does correlate with haemolytic activity. Each of the repeat motif peptoids investigated here, with the exception of RM2, have lipid-bound CD spectra strongly reminiscent of a helical structure that is distinguishable from their structure in PBS. However, the lipid-bound spectra only represent a combination of the fully bound spectrum and the PBS spectrum. We see from our analysis of the spectral evolution with increasing lipid concentration that the strength and nature of the interactions between different peptoids and LUVs varies considerably and therefore comparing lipid-bound spectra like-for-like is not sensible, as they do not represent the true bound spectra but only the best experimentally obtainable estimate in this instance. Therefore any correlations observed between helicity and activity could be incidental.

### 5.3.6 *Peptoid-DNA Interactions*

A preliminary investigation of whether structural changes can be observed via CD on the binding of peptoids to DNA is presented in this section. Investigations into antimicrobial mechanisms have suggested that certain peptoids can cross bacterial membranes and target intracellular features such as DNA. Some peptoid sequences have also successfully been employed as transfection agents, meaning that they interact with DNA in such a way as to condense it into small aggregate particles which are protected from enzymatic degradation and can pass through cell membranes [60, 61]. This activity has been shown to depend strongly on the peptoid primary sequence and therefore, likely the secondary structure. In particular, longer peptoid sequences in the region of 36 residues have proved efficient transfection agents [60].

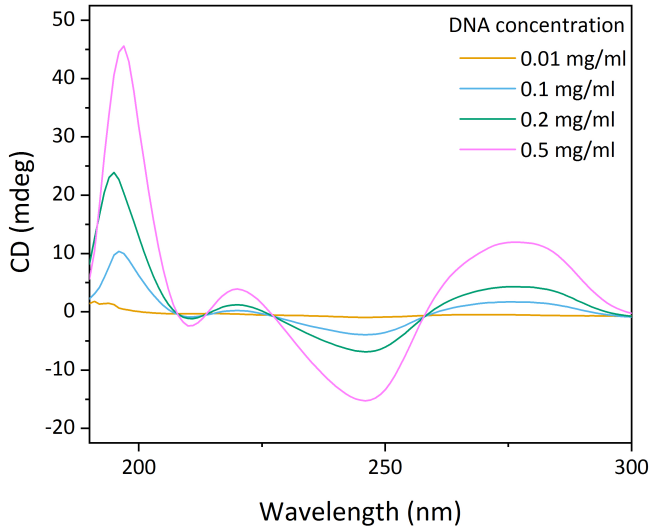



*Fig. 5.30:* CD spectra for calf thymus DNA at various concentrations in 0.01 M PBS. Data shown are the raw spectra in units of millidegrees.

Biological investigations (as yet unpublished) by Jamie Taylor in the Cobb group at Durham University included fluorescence anisotropy to determine DNA binding by different peptoid sequences and high resolution microscopy with fluorescent tagging of peptoids, which revealed that some sequences localise to the DNA inside cells. In this investigation we consider the interaction between both the *N*ae and *N*Lys motif 2 peptoids and DNA. The *N*Lys version of motif 2 (RM4) showed an increase in fluorescence anisotropy of 0.002 upon binding to DNA. The *N*ae version (RM3) showed an increase in anisotropy of 0.0295, more than tenfold higher than the *N*Lys sequence. A larger increase in anisotropy indicates a greater degree of binding of the peptoid to the DNA, therefore this result indicates that substitution of *N*Lys for *N*ae greatly enhanced the affinity of the peptoid for the DNA.

Having demonstrated the potential to probe peptoid-lipid interactions with CD spectroscopy via structural changes in the peptoid upon binding, we tested the application of this technique to peptoid-DNA interactions. DNA itself has a characteristic CD signal, so the analysis of spectra is potentially complicated by the need to de-convolute spectral changes due to changes in peptoid structure and spectral changes due to changes in DNA structure. As with the lipid vesicle binding experiments, we collected spectra for a series of samples with constant peptoid concentration and varying DNA concentration in PBS. For each concentration of DNA investigated we also collected the spectrum of an equivalent sample of pure DNA in PBS, which are shown in Figure 5.30. The motif 2 peptoids were selected due to their diverse anisotropy results and additionally, their being the sequences that interacted in the most predictable manner with the LUVs. The DNA used was double stranded calf-thymus DNA with molecular weight $>$ 20 kpb, obtained from Sigma Aaldrich in lyophilized powder form.

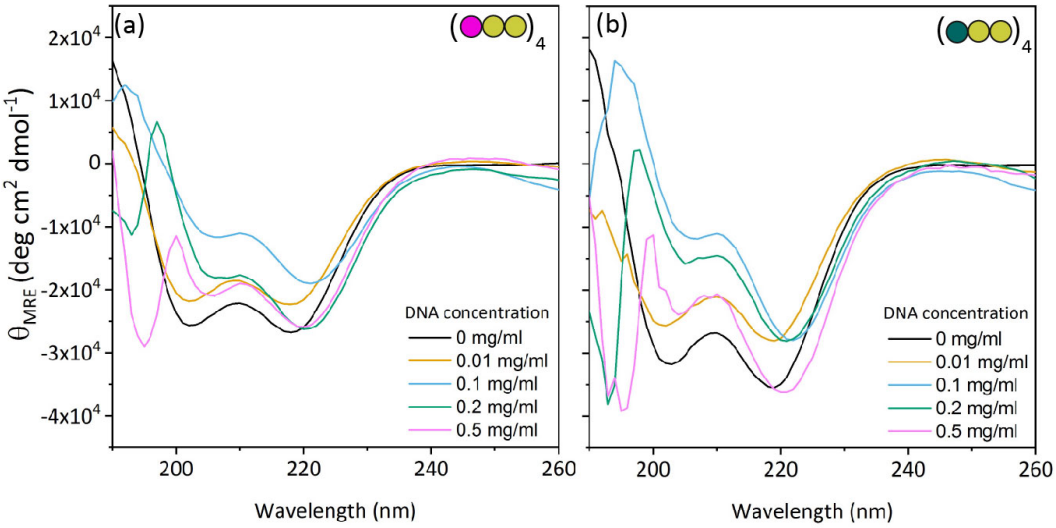



*Fig. 5.31:* CD spectra for (a) ($N$ae$N$sp$e N$sp$e$)$_4$ and (b) ($N$Lys$N$sp$e N$sp$e$)$_4$, both at 25 $\mu$M in 0.01 PBS with DNA at varying concentrations.

The CD spectra for the DNA at varying concentration are shown in Figure 5.30. Each spectrum has 5 distinct bands, 3 positive and 2 negative. These features are characteristic of the double stranded, helical B conformation [62]. Normalising each spectrum for concentration indicated that there are no concentration effects and that only the magnitude of the signal increases as the concentration increases. The spectra collected from samples containing both DNA and peptoids show clearly that some interaction between the two species occurs as they do not merely resemble the combination of the spectra of the two species separately. The mixed peptoid-DNA spectra with the equivalent pure DNA spectra subtracted are shown for both peptoids in Figure 5.31.

The presence of the DNA appears to alter the peptoid CD spectrum in a manner that

indicates that interaction between the two species initiates a conformational change in the peptoid. These changes are qualitatively similar to those observed in the peptoids upon binding to PBS, with the ratio of the two minima changing considerably as the concentration of DNA increases. In contrast to the peptoid-lipid binding spectra, the intensity of the peptoid-DNA binding spectra decreases as the DNA concentration increases. Notably, the changes observed in the peptoid CD spectra are similar for both sequences investigated, indicating that both interact with the DNA in the same manner and perhaps to a similar extent. This is in contrast to the fluorescence anisotropy results which indicate RM3 binds much more strongly to the DNA than RM4. Therefore further investigation to investigate the cause of this discrepancy could be useful.

## 5.4 Conclusions

### *5.4.1 Summary*

In this chapter we have investigated the biophysical basis for the antimicrobial activity of our library of peptoids. The key conclusions from this work are summarised below:

- The folded hydrophobicity of peptoids, quantified as Log *D*, varies between sequences, even where the net hydrophobicity (HPLC retention time) is very similar.
- Log D varies with the net charge on the peptoid, with higher positive charge correlating to lesser hydrophobicity.
- Peptoid sequences with a net charge of +4 partition according to the length of the cationic side chain alkyl group. Motif 2 sequences with *N*ae (shorter) side chains favour the organic phase where identical sequences with *N*Lys (longer) side chains favour the aqueous phase.
- The Log *D* values for the *N*ae versions of the scrambled sequences are considerably more diverse than for the *N*Lys versions.
- The interaction between peptoids and model membranes can be probed by monitoring secondary structural changes with CD spectroscopy.
- Many of the repeat motif sequences interact with both neutral and charged LUVs.
- In some cases the interaction between peptoids and LUVs may be modelled as a conformational switch between two distinct states, though in other cases the behaviour is more complex.

- In the case of the motif 2 sequences, $\Delta G$ for the *N*Lys variant is more negative than for the *N*ae variant, which correlates to the enhanced antimicrobial activity and cyctotoxicity of the former relative to the latter.

- The motif 2 peptoids undergo a structural rearrangement in the presence of DNA, with changes somewhat reminiscent of those upon interaction with lipid bilayers.

### *5.4.2 Outlook*

The PBS-octanol partitioning experiments revealed some relationships between peptoid primary structure and folded hydrophobicity (as measured by Log *D*). This is particularly interesting given the lack of diversity in the net hydrophobicity of the sequences within the library. There was experimental difficulty associated with collecting the data to calculate Log *D* for some of the repeat motif peptoids. The motif 1 peptoids were found to be extremely surface active and the complex self-association behaviour of the *N*ae variant of motif 3 in PBS was also problematic. Nevertheless the results indicate that increasing the number of charged residues in the sequence results in greater folded hydrophilicity. This can also be interpreted in the sense that increasing the number of hydrophobic *N*spe residues increases the folded hydrophobicity.

The most striking result in the Log *D* calculations for the scrambled sequences was that, without exception, the *N*ae sequences preferentially partitioned into the octanol phase while the *N*Lys sequences partitioned into the PBS phase, indicating a greater folded hydophobicity for the former than the latter. The calculated Log *D* values for these peptoids indicate that there is considerably greater variation in the folded hydrophobicity of different sequences than there is in the hydrophobicity as measured by HPLC retention time. In particular, sequences containing *N*ae residues were more diverse in Log *D* values than equivalent sequences containing *N*Lys residues. In Chapter 4 we found that in MD simulations the *N*ae sequences form approximately twice as many side-chain-backbone hydrogen bonds than the equivalent *N*Lys sequences. The positioning of the cationic side chain group in close proximity to the backbone carbonyl oxygen in order to form such bonds could limit its interactions with the surrounding solvent and result in greater shielding of the charge by neighbouring aromatic residues. This to some extent could contribute to the observed discrepancy in Log *D* between *N*ae and *N*Lys sequences and also the variation in Log *D* for the *N*ae sequences. The positioning of the side chains to form the hydrogen bonds, combined with the lesser conformational diversity observed in the CD spectra of these sequences would result less conformational freedom to shield aromatic residues and expose charged residues in polar solvent.

This presents a promising new area for investigation as the majority of studies thus far have utilised *N*Lys as a charge modifying residue [6, 12, 18, 63]. Further investigation into

the precise nature of the relationship between the proportion and position of *N*ae residues may result in improvements in the ability to precisely tune the hydrophobicity of peptoid sequences. Investigation of equivalent sequences with differing lengths of alkyl chain in the charged side chain, such as a peptoid analogue of ornithine, may also add insight into the origins of our observations here. In addition, it would be interesting and perhaps useful to further investigate the relationship between Log $D$ and the antimicrobial activity of the scrambled sequences, to establish whether experimentally determined values of Log $D$ represent a useful predictive tool for activity.

In this chapter we also investigated whether the interaction between peptoids and model membrane systems can be used to predict biological activity. We have demonstrated that the interactions between peptoids and lipid membranes can be probed by CD spectroscopy where there is a secondary structural change associated with the interaction. The particular focus was on investigating the relative strength of the repeat motif peptoid interactions with neutral and negatively charged vesicles. In most cases the peptoids did appear to interact with both LUV types and the structural changes that occurred were reminiscent of those features and changes observed for peptoids in octanol relative to PBS. These generally included an increase in intensity of the second minimum with respect to the first, plus a shift of position of the features to longer wavelengths. However, close comparison of the spectra and the fitting of a two-state binding model to the spectra indicated that the octanol spectrum is not a good match for the (experimentally unobtainable) fully lipid-bound spectrum and that therefore the peptoids may adopt subtly different conformations in these different environments.

The two-state binding model was not applicable to all the peptoids investigated as there appeared to be different behaviour across the range of peptoid-lipid combinations. It is therefore difficult to draw conclusions as to what extent the interactions are mediated by charge or hydrophobicity. Where it was possible to calculate $\Delta G$ the results indicated that the peptoids have stronger affinity for the DOPC/PS vesicles than the pure DOPC vesicles. This was expected, due to the addition of an electrostatic interaction to the hydrophobic effect. It seems very likely that charge did have a positive contribution to the interactions, as several peptoids did not appear to interact with the neutral vesicles, but all except one had a detectable interaction with the charged vesicles. In the case of the motif 2 peptoids, substitution of the shorter cationic chain from *N*ae to *N*Lys increased the affinity of the sequence for both LUV types. This is consistent with the biological data since the MIC and $ED_{50}$ values for these peptoids show the *N*Lys sequence (RM4) to be more active against bacterial and mammalian cell lines than its *N*ae counterpart (RM3). However, the Log $D$ value for RM4 was positive, due a preference to remain in the PBS phase rather than partition into octanol, which indicates that this peptoid is less hydrophobic than RM3, for which Log $D$ was negative. In this instance therefore, greater

folded hydrophobicity (as measured by Log $D$) did not correlate to enhanced membrane activity.

Across the library of peptoids we measured several with positive Log $D$ values which still appeared to partition into the membrane environment so it appears that Log $D$ is not necessarily an accurate predictor of membrane activity. However this was an extremely small library of peptoids and $\Delta G$ was calculated for only a subset of these, so a systematic investigation of a more extensive library of sequences would be useful to try and identify whether there is a correlation under certain conditions. Nevertheless, the results here are encouraging as they demonstrate that we are able to measure the affinity between peptoids and model membranes via biophysical experiments. In the future this could be built upon by developing more complex model membranes to mimic particular pathogenic species.

# References: Chapter 5


[1] R. N. Zuckermann, "Peptoid origins.", *Biopolymers* **2011**, *96*, 545–55.

[2] S. M. Miller, R. J. Simon, S. Ng, R. N. Zuckermann, J. M. Kerr, W. H. Moos, "Proteolytic studies of homologous peptide and N-substituted glycine peptoid oligomers", *Bioorganic and Medicinal Chemistry Letters* **1994**, *4*, 2657–2662.

[3] Y. Luo, H. L. Bolt, G. A. Eggimann, D. F. McAuley, R. McMullan, T. Curran, M. Zhou, P. C. A. Jahoda, S. L. Cobb, F. T. Lundy, "Peptoid Efficacy against Polymicrobial Biofilms Determined by Using Propidium Monoazide-Modified Quantitative PCR", *ChemBioChem* **2017**, *18*, 111–118.

[4] M.-D. Seo, H.-S. Won, J.-H. Kim, T. Mishig-Ochir, B.-J. Lee, "Antimicrobial peptides for therapeutic applications: a review.", *Molecules* **2012**, *17*, 12276–86.

[5] S. A. Fowler, H. E. Blackwell, "Structure-function relationships in peptoids: recent advances toward deciphering the structural requirements for biological function.", *Organic & Biomolecular Chemistry* **2009**, *7*, 1508–1524.

[6] A. M. Czyzewski, H. Jenssen, C. D. Fjell, M. Waldbrook, N. P. Chongsiriwatana, E. Yuen, R. E. W. Hancock, A. E. Barron, "In Vivo, In Vitro, and In Silico Characterization of Peptoids as Antimicrobial Agents.", *PloS one* **2016**, *11*.

[7] N. Molchanova, P. R. Hansen, H. Franzyk, "Advances in development of antimicrobial peptidomimetics as potential drugs", *Molecules* **2017**, *22*.

[8] J. Li, J. J. Koh, S. Liu, R. Lakshminarayanan, C. S. Verma, R. W. Beuerman, "Membrane active antimicrobial peptides: Translating mechanistic insights to design", *Frontiers in Neuroscience* **2017**, *11*, 1–18.

[9] J. A. Patch, A. E. Barron, "Helical peptoid mimics of magainin-2 amide", *Journal of the American Chemical Society* **2003**, *125*, 12092–12093.

[10] R. Kapoor, P. R. Eimerman, J. W. Hardy, J. D. Cirillo, C. H. Contag, A. E. Barron, "Efficacy of antimicrobial peptoids against Mycobacterium tuberculosis", *Antimicrobial Agents and Chemotherapy* **2011**, *55*, 3058–3062.

[11] C. A. Olsen, H. L. Ziegler, H. M. Nielsen, N. Frimodt-Møller, J. W. Jaroszewski, H. Franzyk, "Antimicrobial, hemolytic, and cytotoxic activities of $\beta$-peptoid- peptide hybrid oligomers: Improved properties compared to natural AMPs", *ChemBioChem* **2010**, *11*, 1356–1360.

[12] R. G. Von Borowski, S. C. B. Gnoatto, A. J. Macedo, R. Gillet, "Promising antibiofilm activity of peptidomimetics", *Frontiers in Microbiology* **2018**, *9*.

[13] A. E. Corson, S. A. Armstrong, M. E. Wright, E. E. McClelland, K. L. Bicker, "Discovery and Characterization of a Peptoid with Antifungal Activity against Cryptococcus neoformans", *ACS Medicinal Chemistry Letters* **2016**, *7*, 1139–1144.

[14] S. K. Spicer, A. Subramani, A. L. Aguila, R. M. Green, E. E. McClelland, K. L. Bicker, "Toward a clinical antifungal peptoid: Investigations into the therapeutic potential of AEC5", *Biopolymers* **2019**, *110*.

[15] G. A. Eggimann, H. L. Bolt, P. W. Denny, S. L. Cobb, "Investigating the anti-leishmanial effects of linear peptoids", *ChemMedChem* **2015**, *10*, 233–237.

[16] H. L. Bolt, G. A. Eggimann, P. W. Denny, S. L. Cobb, "Enlarging the chemical space of anti-leishmanials: A structure-activity relationship study of peptoids against: Leishmania mexicana, a causative agent of cutaneous leishmaniasis", *Medicinal Chemistry Communications* **2016**, *7*, 799–805.

[17] J. K. Bang, Y. H. Nan, E. K. Lee, S. Y. Shin, "A novel Trp-rich model antimicrobial peptoid with increased protease stability", *Bulletin of the Korean Chemical Society* **2010**, *31*, 2509–2513.

[18] B. Mojsoska, R. N. Zuckermann, H. Jenssen, "Structure-activity relationship study of novel peptoids that mimic the structure of antimicrobial peptides", *Antimicrobial Agents and Chemotherapy* **2015**, *59*, 4112–4120.

[19] R. Dwivedi, P. Aggarwal, N. S. Bhavesh, K. J. Kaur, "Design of therapeutically improved analogue of the antimicrobial peptide, indolicidin, using a glycosylation strategy", *Amino Acids* **2019**, *51*, 14443–1460.

[20] B. Hoffmann, T. Ast, T. Polakowski, U. Reineke, R. Volkmer, "Transformation of a Biologically Active Peptide into Peptoid Analogs While Retaining Biological Activity", *Protein & Peptide Letters* **2006**, *13*, 829–833.

[21] J. A. Patch, A. E. Barron, "Helical peptoid mimics of magainin-2 amide", *Journal of the American Chemical Society* **2003**, *125*, 12092–12093.

[22] N. P. Chongsiriwatana, J. A. Patch, A. M. Czyzewski, M. T. Dohm, A. Ivankin, D. Gidalevitz, R. N. Zuckermann, A. E. Barron, "Peptoids that mimic the structure, function, and mechanism of helical antimicrobial peptides.", *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105*, 2794–2799.

[23] H. L. Bolt, G. A. Eggimann, C. A. B. Jahoda, R. N. Zuckermann, G. J. Sharples, S. L. Cobb, "Exploring the links between peptoid antibacterial activity and toxicity", *Medicinal Chemistry Communications* **2017**, *8*, 886–896.

[24] I. Zelezetsky, A. Tossi, "Alpha-helical antimicrobial peptides-Using a sequence template to guide structure-activity relationship studies", *Biochimica et Biophysica Acta - Biomembranes* **2006**, *1758*, 1436–1449.

[25] B. Mojsoska, G. Carretero, S. Larsen, R. V. Mateiu, "Peptoids successfully inhibit the growth of gram negative E . coli causing substantial membrane damage", *Nature Publishing Group* **2017**, 1–12.

[26] R. Kapoor, P. R. Eimerman, J. W. Hardy, J. D. Cirillo, C. H. Contag, A. E. Barron, "Efficacy of antimicrobial peptoids against Mycobacterium tuberculosis", *Antimicrobial Agents and Chemotherapy* **2011**, *55*, 3058–3062.

[27] H. Jenssen, T. Lejon, K. Hilpert, C. D. Fjell, A. Cherkasov, R. E. Hancock, "Evaluating different descriptors for model design of antimicrobial peptides with enhanced activity toward P. aeruginosa", *Chemical Biology and Drug Design* **2007**, *70*, 134–142.

[28] C. D. Fjell, H. Jenssen, K. Hilpert, W. A. Cheung, N. Panté, R. E. Hancock, A. Cherkasov, "Identification of novel antibacterial peptides by chemoinformatics and machine learning", *Journal of Medicinal Chemistry* **2009**, *52*, 2006–2015.

[29] N. P. Chongsiriwatana, M. Wetzler, A. E. Barron, "Functional synergy between antimicrobial peptoids and peptides against gram-negative bacteria", *Antimicrobial Agents and Chemotherapy* **2011**, *55*, 5399–5402.

[30] H. Yan, R. E. W. Hancock, "Synergistic interactions between mammalian antimicrobial defense peptides", *Antimicrobial Agents and Chemotherapy* **2001**, *45*, 1558–1560.

[31] S. Rex, "Pore formation induced by the peptide melittin in different lipid vesicle membranes", *Biophysical Chemistry* **1996**, *58*, 75–85.

[32] N. B. Leite, A. Aufderhorst-Roberts, M. S. Palma, S. D. Connell, J. R. Neto, P. A. Beales, "PE and PS Lipids Synergistically Enhance Membrane Poration by a Peptide with Anticancer Properties", *Biophysical Journal* **2015**, *109*, 936–947.

[33] A. S. Ladokhin, M. E. Selsted, S. H. White, "Sizing membrane pores in lipid vesicles by leakage of co-encapsulated markers: pore formation by melittin", *Biophysical Journal* **1997**, *72*, 1762–1766.

[34] C. E. Caesar, E. K. Esbjörner, P. Lincoln, B. Nordén, "Membrane interactions of cell-penetrating peptides probed by tryptophan fluorescence and dichroism techniques: Correlations of structure to cellular uptake", *Biochemistry* **2006**, *45*, 7682–7692.

[35] H. W. Huang, "Molecular mechanism of antimicrobial peptides: The origin of cooperativity", *Biochimica et Biophysica Acta - Biomembranes* **2006**, *1758*, 1292–1302.

[36] K. Andreev, M. W. Martynowycz, A. Ivankin, M. L. Huang, I. Kuzmenko, M. Meron, B. Lin, K. Kirshenbaum, D. Gidalevitz, "Cyclization Improves Membrane Permeation by Antimicrobial Peptoids", *Langmuir* **2016**, *32*, 12905–12913.

[37] K. Andreev, M. W. Martynowycz, M. L. Huang, I. Kuzmenko, W. Bu, K. Kirshenbaum, D. Gidalevitz, "Hydrophobic interactions modulate antimicrobial peptoid selectivity towards anionic lipid membranes", *Biochimica et Biophysica Acta - Biomembranes* **2018**, *1860*, 1414–1423.

[38] M. R. Landry, J. L. Rangel, V. P. Dao, M. A. Mackenzie, F. L. Gutierrez, K. M. Dowell, A. L. Calkins, A. A. Fuller, G. Y. Stokes, "Length and Charge of Water-Soluble Peptoids Impact Binding to Phospholipid Membranes", *Journal of Physical Chemistry B* **2019**, *123*, 5822–5831.

[39] L. R. Brown, W. Braun, A. Kumar, K. Wüthrich, "High resolution nuclear magnetic resonance studies of the conformation and orientation of melittin bound to a lipid-water interface.", *Biophysical Journal* **1982**, *37*, 319–28.

[40] A. S. Ladokhin, M. Fernández-Vidal, S. H. White, "CD spectroscopy of peptides and proteins bound to large unilamellar vesicles", *Journal of Membrane Biology* **2010**, *236*, 247–253.

[41] P. D. Fey, M. E. Olson, "Current concepts in biofilm formation of Staphylococcus epidermidis", *Future Microbiol.* **2011**, *5*, 917–933.

[42] M. Otto, "Staphylococcus epidermidis – the "accidental" pathogen", *Nature Reviews Microbiology* **2009**, *7*, 555–567.

[43] M. Miragaia, I. Couto, S. F. F. Pereira, K. G. Kristinsson, H. Westh, J. Almeida, I. Santos-sanches, "Molecular Characterization of Methicillin-Resistant", *Journal of Clinical Microbiology* **2002**, *40*, 430–438.

[44] S. Y. Tong, J. S. Davis, E. Eichenberger, T. L. Holland, V. G. Fowler, "Staphylococcus aureus infections: Epidemiology, pathophysiology, clinical manifestations, and management", *Clinical Microbiology Reviews* **2015**, *28*, 603–661.

[45] M. Á. Argudín, M. C. Mendoza, M. R. Rodicio, "Food Poisoning and Staphylococcus aureus Enterotoxins", *Toxins* **2010**, *2*, 1751–1773.

[46] M. Barber, "Methicillin-resistant staphylococci", *Journal of Clinical Pathology* **1961**, *14*, 385–393.

[47] G. Ippolito, S. Leone, F. N. Lauria, E. Nicastri, R. P. Wenzel, "Methicillin-resistant Staphylococcus aureus: the superbug", *International Journal of Infectious Diseases* **2010**, *14*, 7–11.

[48] S. C. Yang, C. H. Lin, I. A. Aljuffali, J. Y. Fang, "Current pathogenic Escherichia coli foodborne outbreak cases and therapy development", *Archives of Microbiology* **2017**, *199*, 811–825.

[49] S. L. Gellatly, R. E. Hancock, "Pseudomonas aeruginosa: New insights into pathogenesis and host defenses", *Pathogens and Disease* **2013**, *67*, 159–173.

[50] P. D. Lister, D. J. Wolter, N. D. Hanson, "Antibacterial-resistant Pseudomonas aeruginosa: Clinical impact and complex regulation of chromosomally encoded resistance mechanisms", *Clinical Microbiology Reviews* **2009**, *22*, 582–610.

[51] Z. Pang, R. Raudonis, B. R. Glick, T. J. Lin, Z. Cheng, "Antibiotic resistance in Pseudomonas aeruginosa: mechanisms and alternative therapeutic strategies", *Biotechnology Advances* **2019**, *37*, 177–192.

[52] H. L. Bolt, C. E. Williams, R. V. Brooks, R. N. Zuckermann, S. L. Cobb, E. H. Bromley, "Log D versus HPLC derived hydrophobicity: The development of predictive tools to aid in the rational design of bioactive peptoids", *Biopolymers* **2017**, *108*, 1–7.

[53] A. Tabbabi, "Review of leishmaniasis in the middle east and North Africa", *African Health Sciences* **2019**, *19*, 1329–1337.

[54] R. Arenas, E. Torres-Guerrero, M. R. Quintanilla-Cedillo, J. Ruiz-Esmenjaud, "Leishmaniasis: A review", *F1000Research* **2017**, *6*, 1–15.

[55] A. Pagliara, M. Reist, S. Geinoz, P.-A. Carrupt, B. Testa, "Evaluation and Prediction of Drug Permeation", *Journal of Pharmacy and Pharmacology* **1999**, *51*, 1339–1357.

[56] R. H. Raetzs, M. Sharp, "Biosynthesis and Function of Phospholipids in Escherichia coli", *Journal of Biological Chemistry* **1990**, *265*, 1235–1238.

[57] C. Sohlenkamp, O. Geiger, "Bacterial membrane lipids : diversity in structures and pathways", *FEMS Microbiology Reviews* **2016**, *40*, 133–159.

[58] Y. Shai, "Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by alpha-helical antimicrobial and cell non-selective membrane-lytic peptides", *Biochimica et Biophysica Acta - Biomembranes* **1999**, *1462*, 55–70.

[59] C. Avitabile, L. D. D'Andrea, A. Romanelli, "Circular Dichroism studies on the interactions of antimicrobial peptides with bacterial cells", *Scientific Reports* **2014**, *4*.

[60] J. E. Murphy, T. Uno, J. D. Hamer, F. E. Cohen, V. Dwarki, R. N. Zuckermann, "A combinatorial approach to the discovery of efficient cationic peptoid reagents for gene delivery", *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 1517–1522.

[61] B. A. Lobo, J. A. Vetro, D. M. Suich, R. N. Zuckermann, C. R. Middaugh, "Structure / Function Analysis of Peptoid / Lipitoid : DNA Complexes", *Journal of Pharmaceutical Sciences* **2003**, *92*, 1905–1918.

[62] J. Kypr, I. Kejnovska, D. Renciuk, M. Vorlıckova, "Circular dichroism and conformational polymorphism of DNA", *Nucleic Acids Research* **2009**, *37*, 1713–1725.

[63] J. K. Kim, S. A. Lee, S. Shin, J. Y. Lee, K. W. Jeong, Y. H. Nan, Y. S. Park, S. Y. Shin, Y. Kim, "Structural flexibility and the positive charges are the key factors in bacterial cell selectivity and membrane penetration of peptoid-substituted analog of Piscidin 1", *Biochimica et Biophysica Acta - Biomembranes* **2010**, *1798*, 1913–1925.

# 6.0 Summary and Conclusions

Within this thesis we have explored the relationships between the primary sequence and the secondary structures of peptoids and subsequently investigated the biophysical origins of peptoid antimicrobial activity. Peptoids represent a uniquely tunable class of biomimetic polymers with exciting applications in many fields, not least as future antimicrobial molecules that may help tackle the global problem of antimicrobial resistance. The near limitless sequence space available to synthetic chemists has resulted in rapid generation of combinatorial libraries from which biologically active sequences have been identified. The details pertaining to the physical origins of this activity however, remain unclear [1].

Membrane disruptive antimicrobial activity in AMPs and peptoids is thought to be inextricably linked to their ability to adopt certain secondary structural conformations. We therefore aimed to explore how rearrangements in peptoid primary sequence affect the secondary structure adopted in both aqueous and organic environments. We used PBS and octanol as solvents for their similarities to blood plasma and the cell membrane interior. While peptoids in aqueous solvent have been investigated extensively, ours are among the first reports of the structural preferences of peptoids in octanol. In Chapter 3 we used CD spectroscopy to investigate structural rearrangements in a small library of peptoids including both repeat motif and scrambled sequences. The investigation was focussed on two key questions: the influence of the positioning of aromatic and cationic residues along the peptoid backbone and the influence of the chemical nature of the cationic side chains.

The sequences we investigated adopt helical secondary structures in PBS and distinct but also characteristically helical structures in octanol. This was anticipated due to the inclusion of a number of *N*spe residues, known to induce a helical backbone, in each sequence. The shape and intensity of spectra varied with both the positioning of residues and the choice of cationic side chain. As expected, the intensity of helical character was found to increase as the proportion of helix inducing residues increased. Less predictably, the helical character was also enhanced, invariably, in sequences where the cationic side chain was *N*Lys, compared to equivalent sequences where it was *N*ae. The only difference between these two side chains is the addition of two extra methylene groups in *N*Lys, making the side chain longer and therefore, we concluded, reducing strain in the peptoid

backbone and allowing easier distribution of charge, resulting in the observed differences between the *N*Lys and *N*ae spectra. We also observed that peptoids are thermally robust and minimal unfolding occurs between 20°C and 90°C, though it does to a greater extent in *N*ae sequences than *N*Lys sequences.

In Chapter 4 we aimed to explore our library of peptoids computationally. Much of the data presented in this chapter pertained to the development of new parameters to adapt a force field, GAFF, to correctly capture the backbone torsional preferences of peptoids and enable us to perform atomistic MD simulations. This was achieved by fitting rotational energy profiles for the peptoid backbone obtained from GAFF in its native form, to the equivalent quantum mechanical energy profiles. The success of this approach was validated through simulations of small peptoid molecules and fragments which have been characterised, both computationally and experimentally, in the literature [2–4]. We also, for the first time, reported the use of Hamiltonian replica exchange between just two replicas, to enhance the sampling in our simulations at lesser computational cost than alternative methods such as traditional temperature replica exchange.

Using our adapted force field parameters we simulated the repeat motif peptoid sequences in water and octanol, principally analysing the trajectories through Ramachandran plots, which illustrate the sampling of the backbone dihedral angles during the simulation. The repeat motif sequences generally sampled angles distributed around those associated with the peptoid *N*spe helix and the global peptoid minimum [5], with more variation in water than octanol, indicating a more dynamically changing structure in the former than the latter. Additionally, *cis-trans* isomerisation of the amide bond was observed in water but rarely in octanol. These results are consistent with both literature reports of a lack of diversity in peptoid backbone dihedral angles [6] and our CD spectra from Chapter 3, which indicate a greater extent of helical character in octanol than PBS.

There was little in the Ramachandran plots to differentiate between the structural preferences of the *N*ae and *N*Lys sequences, highlighting the drawbacks of these plots as a method for comparing sequences. However, further analysis revealed that the *N*ae sequences formed side chain to backbone hydrogen bonds approximately twice as readily as equivalent *N*Lys sequences. This shed new light on the experimental observations made in Chapter 3. Though we did not simulate the scrambled sequences, the consistent preference for the *N*ae side chain to hydrogen bond with the backbone across all the repeat motif sequences indicates that it may constrain the sequences to certain conformations, resulting in the lack of diversity in CD spectra for *N*ae sequences, relative to *N*Lys sequences. The former also display a greater overall loss of spectral intensity at high temperatures than the latter, which may be a direct result of partial stabilisation of the structure by hydrogen bonds, rather than steric interactions.

We also gave a brief overview in Chapter 4 of the first steps towards developing a coarse

grained Martini-compatible force field for these peptoid sequences. Further development in this area will be a valuable next step as it will enable simulations which can access longer time scales and larger systems, such as the interactions between peptoids and lipid bilayers. The first results in this area indicate that the formation of transient, disordered toroidal pores may contribute to the antimicrobial activity of motif 2 sequences. Continuing with this work will help further elucidate the nature of peptoid-membrane interactions and provide mechanistic insight to complement the experimental results presented in Chapter 5.

Having confirmed that the peptoids have distinctly different CD spectra in aqueous and organic environments, we used this to probe their interactions with lipid vesicles in Chapter 5. The repeat motif sequences underwent structural changes upon interaction with neutral and negatively charged LUVs, designed to mimic mammalian and bacterial cell membranes respectively. In some instances we were able to fit a two state model for binding to the data and calculate the free energy of transfer into the membrane. In other cases, the behaviour was sufficiently complex to prohibit this kind of analysis. The spectra of the motif 2 sequences evolved in a more systematic manner with increasing lipid concentration than the other sequences. Among the motif 2 sequences the *N*Lys sequence had a greater affinity for both LUV types than the *N*ae sequence which correlated to their activity against gram positive bacteria and mammalian cell lines. More generally, the simple comparison between neutral and anionic LUVs enabled us to highlight that electrostatics contribute to the interaction between peptoid and membrane.

Finally, within Chapter 5 we also investigated peptoid folded hydrophobicity by PBS-octanol partitioning. Very clear patterns emerged in the results, including the folded hydrophobicity increasing with an increasing proportion of *N*spe residues and decreasing net positive charge in the sequences. Most intriguingly the sequences with a net charge of +4 (motif 2 and scrambled) partitioned according to length of charged side chain, with *N*ae imparting a preference for the organic phase and *N*Lys a preference for the aqueous phase.

This contrast between *N*ae and *N*Lys sequences would perhaps be the most interesting outcome of this work to pursue imminently, as there were clear results pertaining to this seemingly minor substitution in every aspect of the investigations presented in thesis. The shorter *N*ae produces sequences that are uniform in CD spectrum but diverse in folded hydrophobicity. For the longer *N*Lys, we observe the opposite. This knowledge will be useful in guiding the design of combinatorial libraries in future where diversity in one or the other of these properties is required. While this link has been clearly identified, it is much harder to elucidate its physical origin from the data presented here, though the MD simulations revealed a greater tendency of *N*ae to form side chain-backbone hydrogen bonds than *N*Lys, which may contribute to the experimentally observed behaviours.

Investigation of additional cationic side chain structures, such as a peptoid analogue of ornithine, a mid length analogue to *N*ae and *N*Lys, may help to elucidate the origins of these behaviours. Understanding this fully would enable better control and tunability of structural properties in the design of new sequences in the future.

Log D was not found to be a strong indicator of the activity level of a particular sequence, among those explored here, as it was not found to directly correlate to either biological activity or the interaction with model membrane systems. However this in itself sheds light on the nature of the interactions between the peptoids and vesicles as it reveals that in some cases, peptoids which preferentially remain in the aqueous phase in partitioning experiments will nonetheless interact favourably with lipid membranes. Computational investigations comparing the interactions of *N*ae and *N*Lys sequences with lipid membranes could be a useful next step in investigating why this might be the case and ultimately will be important for understanding the origins of peptoid antimicrobial activity and utilising them in rational sequence design.

# References: Chapter 6


[1] R. N. Zuckermann, "Peptoid origins.", *Biopolymers* **2011**, *96*, 545–55.

[2] D. T. Mirijanian, R. V. Mannige, R. N. Zuckermann, S. Whitelam, "Development and use of an atomistic CHARMM-based forcefield for peptoid simulation", *Journal of Computational Chemistry* **2014**, *35*, 360–370.

[3] V. A. Voelz, K. A. Dill, I. Chorny, "Peptoid conformational free energy landscapes from implicit-solvent molecular simulations in AMBER", *Biopolymers* **2011**, *96*, 639–650.

[4] G. L. Butterfoss, P. D. Renfrew, B. Kuhlman, K. Kirshenbaum, R. Bonneau, "A preliminary survey of the peptoid folding landscape", *Journal of the American Chemical Society* **2009**, *131*, 16798–16807.

[5] P. Armand, K. Kirshenbaum, A. Falicov, R. L. Dunbrack, K. A. Dill, R. N. Zuckermann, F. E. Cohen, "Chiral N-substituted glycines can form stable helical conformations", *Folding and Design* **1997**, *2*, 369–375.

[6] L. J. Weiser, E. E. Santiso, "Molecular modeling studies of peptoid polymers", *AIMS Materials Science* **2017**, *4*, 1029–1051.