



# Characterization and Exploitation of Peptoid Secondary Structures: A Synergistic Experimental and Computational Approach

*Author:*  
Sébastien HOYAS

*Supervisor:*  
Prof. Jérôme CORNIL  
*Co-supervisor:*  
Prof. Pascal GERBAUX

*A thesis submitted in fulfillment of the requirements  
for the degree of Doctor of Philosophy in Science*

*in the*

Laboratory for Chemistry of Novel Materials  
Organic Synthesis and Mass Spectrometry Laboratory

ACADEMIC YEAR 2020-2021





---

# Characterization and Exploitation of Peptoid Secondary Structures: A Synergistic Experimental and Computational Approach

---

*A thesis submitted in fulfillment of the requirements  
for the degree of Doctor of Philosophy in Science*

ACADEMIC YEAR 2020-2021

**Sébastien HOYAS**

*Supervisor:*

Prof. Jérôme CORNIL

*Co-supervisor:*

Prof. Pascal GERBAUX

*Jury committee:*

Prof. Mathieu SURIN (Université de Mons, Belgique) - President

Prof. Sylvain GABRIELE (Université de Mons, Belgique) - Secretary

Prof. Ronald ZUCKERMANN (Molecular Foundry, U.S.A.)

Prof. Luca MUCCIOLI (University of Bologna, Italy)

Prof. Philippe DUGOURD (Université Claude Bernard Lyon 1, France)





---

# ABSTRACT

---

Peptoids are peptide regioisomers with attractive structural tunability in terms of sequence and three-dimensional arrangement. The conformers of peptoids are dependent on the choice of the side chains, and hence the sequence, and can form well-defined structures such as ribbons, loops, helices, etc. Peptoids are thus foreseen to have a great potential for many diverse applications. It was recently demonstrated that peptoids bearing (*S*)-1-phenylethyl (*N*<sub>spe</sub>) side chains are efficient chiral selectors when grafted on silica in chromatography columns. The enantioselectivity appears to be related to their conformation, which emphasizes the need to thoroughly characterize their secondary structure. Typically, nuclear magnetic resonance and circular dichroism spectroscopies are used to decipher their conformation. However, the information is averaged over every conformer.

Ion mobility coupled to mass spectrometry (IM-MS) is an appealing technique to provide information about the three-dimensional arrangement of gaseous ions. Using this technique, the different conformers may be resolved. Typically, the attribution of the gas-phase conformers is performed by confronting the experimental results to candidate geometries generated by molecular dynamics simulations using a properly parametrized force field.

The first part of this thesis is dedicated to the development of force field parameters to adequately describe the potential energy surface of peptoids. In the second part, we focus our attention on peptoids bearing two specific side chains: the *N*<sub>sar</sub> (sarcosine) and *N*<sub>spe</sub> side chains. Although characterized by different conformations in solution, these two peptoids adopt a loop-like conformation in gas phase because of the need to stabilize the supported charge, mandatory for mass spectrometry analyses. Compared to peptides, peptoids lack the ability to form hydrogen bonds. When turning our attention to the (*S*)-*N*-(1-carboxy-2-phenylethyl) side chain (*N*<sub>scp</sub>), the story is now different. The peptoid ions adopt a helical conformation in gas phase due to the formation of a hydrogen bond network along the peptoid backbone.

In the last part, we looked back at solution properties and investigated the conformation of *N*<sub>spe</sub> peptoids in interaction with chiral guest molecules (2,2'-bihydroxy-1,1'-binaphthyl) by molecular dynamics simulations. In particular, we simulated a section of a chiral chromatography silica bead pore, on which *N*<sub>spe</sub> peptoids are grafted. The elution of the guest molecules (*R* or *S*) was performed by pulling them in a given direction. The enantioselectivity was characterized by monitoring their average velocities that were qualitatively compared to experimental data reported in the literature.



---

## ACKNOWLEDGEMENTS

---

I would like to sincerely express my gratitude to Prof. Mathieu Surin, Prof. Sylvain Gabriele, Prof. Ronald Zuckermann, Prof. Luca Muccioli and Prof. Philippe Dugourd for accepting to read and evaluate this thesis. I feel honored to share this experience with you.

A l'heure où j'écris ces lignes, je suis presque à la fin de cette aventure, réalisée grâce aux Professeurs Jérôme Cornil et Pascal Gerbaux. C'est en grande partie grâce à eux que je peux aujourd'hui conclure et présenter ce travail. Je les remercie pour m'avoir permis de bénéficier de leur synergie. Merci aussi pour leur enthousiasme, disponibilité et toutes les discussions (scientifiques ou non 😊) partagées durant cette aventure. Ils m'ont permis d'explorer la chimie théorique et expérimentale, de découvrir différents domaines au sein de la chimie, ce qui a largement contribué à mon épanouissement. Grâce à eux, je me sens comme un poisson dans l'eau 🐟 Je les remercie aussi pour leur confiance placée dans mes idées et projets. Au-delà de l'aspect scientifique, ce fût aussi un plaisir d'apprendre à les connaître personnellement. Et bien sûr, de découvrir que l'on partage la même passion pour l'humour décalé (et parfois graveleux...). Merci.

Master Jedi, c'est (presque) comme cela que m'a été présenté Vincent Lemaure (il y a déjà... 6 ans !). Avec le temps, son rang a évolué à « Masta » Jedi au sein de notre petite communauté de théoriciens-praticiens, pour finalement atteindre le rang ultime de « Matcha » Jedi. Très vite, le contact est bien passé, et c'est un véritable plaisir d'avoir eu la chance de travailler avec lui. Ce fût aussi (et ça l'est toujours !) un plaisir de le connaître personnellement en le côtoyant au sein du même bureau, mais aussi en bureau virtuel. Heureusement d'ailleurs, que nous avons continué à partager un bureau, même virtuel, qui m'a aussi permis de faire connaissance avec une partie de Luna. 🐱 Au cours de ces années, son intérêt pour les subtilités ne m'est pas sortie de la tête : découvrira-t-il celles cachées dans ce texte ? 😊 Dans tous les cas: merci pour tout.

Ce travail met en valeur la synergie entre la chimie théorique et expérimentale pour l'étude de peptoïdes, mais n'aurait jamais pu voir le jour sans les « créatrices » de peptoïdes, Emilie Halin et Perrine Weber, et leur encadrant fou Julien De Winter. Merci à vous d'avoir synthétisés et (co)-analysés les dizaines de peptoïdes que nous avons étudiés au cours de ces dernières années. Merci également à Peptorentin pour

son travail lors du mémoire. Ce fût un plaisir de travailler avec lui, et j'espère pouvoir terminer ce que nous avons commencé.

Je remercie également le Professeur Mathieu Surin pour sa collaboration et discussions dans nos projets, entre autre à propos du dichroïsme circulaire des peptoïdes. Merci aussi à Marie Trévisan et Maxime Leclercq de m'avoir initié à la méthode.

Lors de cette thèse, j'ai eu la chance de non seulement, collaborer avec le Professeur Luca Muccioli, mais aussi d'effectuer un séjour de recherche au sein de son laboratoire à Bologne. Merci pour sa disponibilité, son partage de connaissances et le temps passé sur notre projet. Merci également aux membres du laboratoire, plus particulièrement Otello, Matteo, Andrea et Ariana pour votre accueil.

Comme indiqué précédemment, j'ai eu l'occasion de travailler de près ou de loin sur d'autres systèmes, notamment avec Quentin, alias "*bite*" (n'y voyez-là aucune insulte ou agressivité, mais bien une marque d'affection courante dans notre région). En plus d'avoir été un collègue pendant plusieurs années, Quentin est aussi et surtout un ami d'enfance avec qui j'ai passé de formidables moments, tant au labo qu'en dehors. Après mon retour de Bologne, la tradition du "*tcho' café*" matinal avec Thomas nous a mené dans des discussions et situations plutôt cocasses, pour ne pas dire improbables ! Je vous remercie pour ces moments et que tout cela continue malgré la distance.

Merci aussi Fabrice, alias "Mr. Saintmont", mon comparse d'études, de thèse, conférences et ami. Merci d'avoir partagé cette expérience pas piquée des hannetons avec moi, et ce malgré la virtualisation de notre environnement de travail.

Cette thèse étant théorique et expérimentale, j'ai eu l'occasion de partager un bureau au sein des deux services CMN et S<sup>2</sup>MOs. Merci notamment à Eric, Antoine et aux mémorants qui ont été de passage durant ces années. Merci aussi au B157, plus particulièrement à Laura, Pauline et Matcha Jedi, notamment pour les "discussions potins" et la découverte de cette infâme poudre verte. #Sebdoesnotlovematcha 😊  
Merci aussi aux amis du B157, dont certains aiment laisser des mots doux dans mes carnets (n'est-ce pas Maxime ?).

Bien évidemment, la thèse de doctorat ne se réalise pas que dans le bureau. Elle se déroule surtout au sein du labo, en l'occurrence, des labos dans mon cas. Je ne peux citer tout le monde au risque d'en oublier certains, mais vous avez contribué de près ou de loin à cette bonne humeur qui rend le travail agréable.

Quand on fait des calculs, on a besoin d'un bon ordinateur, mais surtout d'un bon informaticien. Merci également à Sébastien Kozlowskyj pour sa rapidité à résoudre mes problèmes, et ce, en tout temps.

Un jour, on m'a demandé pourquoi je faisais une thèse de doctorat en chimie. Tout compte fait, j'ai toujours aimé la Science ainsi que découper des cheveux en quatre, donc pourquoi pas ? Mais en y réfléchissant un peu plus, c'est aussi grâce au Professeur Roberto Lazzaroni. Son enseignement, du début à la fin du cursus de chimie à Mons, a vraiment été une source d'inspiration. Et bien entendu, merci à lui de m'avoir accueilli au sein de son laboratoire.

Merci aussi aux amis, déjà cités ou non, en Belgique comme à l'étranger, pour le partage de cette expérience tant à l'université qu'en dehors.

Finalement, j'aimerais remercier ma famille, surtout mes parents pour avoir eu la bonne idée de me laisser faire ce qui me plaît, ainsi que Marion pour son soutien quotidien.



---

# Contents

---

<b>Abstract</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>v</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Peptoid, a misspelling for peptide? . . . . .	1
1.2 Solid-phase synthesis: the golden standard . . . . .	3
1.3 Conformational control . . . . .	5
1.3.1 <i>cis/trans</i> isomerism . . . . .	7
1.3.2 Chirality . . . . .	8
1.4 Peptoid characterization . . . . .	10
1.4.1 Circular dichroism . . . . .	10
1.4.2 X-ray diffraction . . . . .	12
1.4.3 Nuclear magnetic resonance . . . . .	13
1.4.4 Mass spectrometry . . . . .	15
1.4.5 Molecular modeling . . . . .	16
1.5 Secondary structure . . . . .	18
1.5.1 The helical dominance . . . . .	18
1.5.2 Loops . . . . .	22
1.5.3 Ribbons . . . . .	23
1.5.4 Higher order structure . . . . .	24
1.6 Applications of peptoids . . . . .	25
1.6.1 Biomedical applications . . . . .	25
1.6.2 Materials applications . . . . .	28
<b>2 Objectives of the thesis</b>	<b>31</b>
<b>3 Methodology</b>	<b>33</b>
3.1 Synthesis . . . . .	33
3.2 Ion mobility mass spectrometry . . . . .	34

3.3	Molecular modeling . . . . .	38
3.3.1	Quantum mechanics . . . . .	38
3.3.2	Molecular mechanics . . . . .	40
3.3.3	Molecular dynamics . . . . .	44
3.3.4	Collision cross section calculation . . . . .	45
	<b>Results and Discussion</b>	<b>48</b>
<b>4</b>	<b>PEPDROID: a DREIDING-based force field for peptoids</b>	<b>49</b>
4.1	Motivations . . . . .	51
4.2	DREIDING reparametrization methodology . . . . .	52
4.3	Force field validation . . . . .	59
<b>5</b>	<b>Ion Mobility Mass Spectrometry of Peptoids</b>	<b>65</b>
5.1	How to interpret data from IM-MS of oligo- and polymeric compounds?	67
5.2	Methodology . . . . .	75
5.3	Results and Discussion . . . . .	78
5.3.1	The simplest case: sarcosine peptoids . . . . .	78
5.3.2	Chirality and bulkiness: the ingredients of gas phase helical folding? . . . . .	84
5.3.3	Back to basics: inspiration from peptides . . . . .	97
5.4	Conclusions . . . . .	108
<b>6</b>	<b>Peptoids as chiral selector molecules</b>	<b>111</b>
6.1	The importance of enantioselectivity . . . . .	113
6.2	Methodology . . . . .	116
6.3	Conformational analysis . . . . .	121
6.4	Results and Discussion . . . . .	123
6.4.1	Chiral peptoids . . . . .	123
6.4.2	Achiral peptoids: validation of the model . . . . .	134
6.4.3	Conclusions . . . . .	139
<b>7</b>	<b>Summary and perspectives</b>	<b>141</b>
<b>A</b>	<b>Additional Material of Chapter 4</b>	<b>147</b>
A.1	Dihedral profiles of peptoid side chains . . . . .	147
A.2	Ramachandran-like plots of peptoid side chains . . . . .	148
<b>B</b>	<b>Additional Material of Chapter 5</b>	<b>151</b>
B.1	Dihedral profiles of <i>N</i> scp and <i>N</i> s <sub>cm</sub> side chains . . . . .	151

<b>C</b>	<b>Additional Material of Chapter 6</b>	<b>153</b>
C.1	Conformations description . . . . .	153
C.2	<i>Nspe</i> peptoids - ( <i>S</i> )-BINOL: displacements analysis . . . . .	155
C.3	<i>Nspe</i> peptoids - ( <i>R</i> )-BINOL: displacements analysis . . . . .	156
C.4	<i>Nsar</i> peptoids - ( <i>S</i> )-BINOL: displacements analysis . . . . .	157
C.5	<i>Nsar</i> peptoids - ( <i>R</i> )-BINOL: displacements analysis . . . . .	158
C.6	<i>Nspe</i> peptoids - ( <i>S</i> )-BINOL: contact maps . . . . .	159
C.7	<i>Nspe</i> peptoids - ( <i>R</i> )-BINOL: contact maps . . . . .	163
C.8	<i>Nsar</i> peptoids - ( <i>S</i> )-BINOL: contact maps . . . . .	168
C.9	<i>Nsar</i> peptoids - ( <i>R</i> )-BINOL: contact maps . . . . .	173
	<b>Bibliography</b>	<b>178</b>
	<b>Publications</b>	<b>201</b>



---

## List of Abbreviations

---

<b>2D</b>	<b>two-dimensional</b>
<b>3D</b>	<b>three-dimensional</b>
<b>BINOL</b>	2,2'-bihydroxy-1,1'-binaphthyl
<b>CCS</b>	Collision Cross Section
<b>CSP</b>	Chiral Stationary Phase
<b>CD</b>	Circular Dichroism
<b>DFT</b>	Density Functional Theory
<b>DP</b>	Degree of Polymerization
<b>EHSS</b>	Exact Hard Sphere Scattering
<b>ESI</b>	Electrospray Ionization
<b>H-bond</b>	Hydrogen-bond
<b>HF</b>	Hartree-Fock
<b>IMMS</b>	Ion Mobility Mass Spectrometry
<b>LCAO</b>	Linear Combination of Atomic Orbitals
<b>MS</b>	Mass Spectrometry
<b>MD</b>	Molecular Dynamics
<b>MM</b>	Molecular Mechanics
<b>MP2</b>	Møller-Plesset 2
<b>NMR</b>	Nuclear Magnetic Resonance
<i>Nae</i>	<i>N</i> -2-aminoethyl
<i>Nce</i>	<i>N</i> -2-carboxyethyl
<i>Nnpr</i>	<i>N</i> - <i>n</i> -propyl
<i>Nph</i>	<i>N</i> -phenyl or <i>N</i> -aryl
<i>Npm</i> or <i>Npe</i>	<i>N</i> -phenylmethyl or <i>N</i> -benzyl
<i>Nrce</i>	( <i>R</i> )- <i>N</i> -(1-carboxy-2-ethyl)
<i>Nr1npe</i>	( <i>R</i> )- <i>N</i> -1-naphthylethyl
<i>Nrpe</i>	( <i>R</i> )- <i>N</i> -1-phenylethyl
<i>Nsar</i>	<i>N</i> -methylglycine
<i>Nscp</i>	( <i>S</i> )- <i>N</i> -(1-carboxy-2-phenylethyl)
<i>Ns1npe</i>	( <i>S</i> )- <i>N</i> -1-naphthylethyl
<i>Nspe</i>	( <i>S</i> )- <i>N</i> -1-phenylethyl
<i>NtBu</i>	<i>N</i> - <i>tert</i> -butyl
<b>PA</b>	Projection Approximation
<b>PBC</b>	Periodic Boundary Conditions
<b>PES</b>	Potential Energy Surface
<b>PME</b>	Particle Mesh Ewald

<b>QM</b>	<b>Quantum Mechanics</b>
$t_D$	<b>Drift Time</b>
<b>RMSD</b>	<b>Root Mean Square Deviation</b>
<b>TM</b>	<b>Trajectory Method</b>
<b>TWIG</b>	<b>Travelling Wave Ion Guide</b>
<b>TWIM</b>	<b>Travelling Wave Ion Mobility</b>
<b>vdW</b>	<b>van der Waals</b>

---

# Introduction

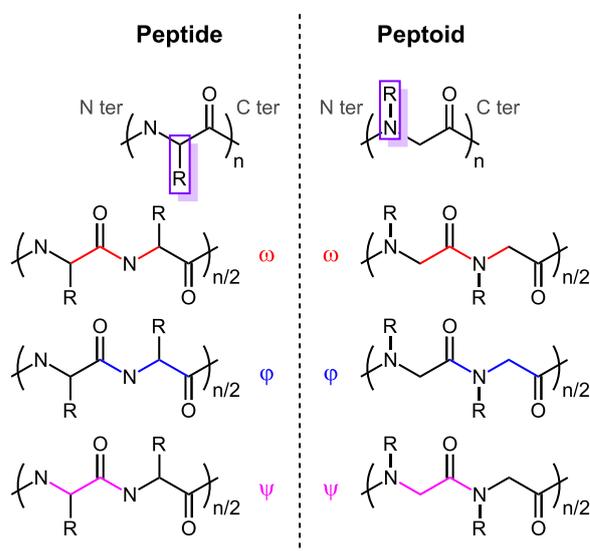
---

## 1.1 Peptoid, a misspelling for peptide?

Proteins are biomacromolecules present in every living organism.<sup>1</sup> Although the chemical composition of a protein is quite simple, the structural diversity is very rich, thanks to the numerous possibilities to link, by amide bonds, the 20 natural amino acids (often called *residues*), which differ by the nature of the side chain appended on the  $\alpha$ -carbon (**Scheme 1.1**). Proteins play a key role in biological systems by controlling a collection of cellular functions, for example via protein-protein interactions. The functionality of most proteins arises from their ability to fold into specific three-dimensional structures, according to the amino acid sequence, and to display chemical environments (or “hot-spots”) capable of interacting specifically with other molecular species at their surface.<sup>2</sup> The availability of high-quality structural data from X-ray diffraction (XRD) and nuclear magnetic resonance (NMR) of proteins enables the exploration of structure-property relationships.<sup>3</sup> New potential drug targets are identified at fast-moving pace and considerable efforts are devoted to the development and design of compounds capable of targeting the protein “hot-spots” involved in diseases. It became also apparent that combinatorial chemistry is mandatory to screen many different compounds to identify potential new drugs.<sup>4</sup> Obviously, investigating smaller protein segments to mimic the binding part of a protein comes directly to mind.<sup>5</sup> In the literature, peptides are commonly referred to as amino acid sequences composed of less than 50 residues. These compounds have the advantage to share the same chemical nature as that of the targeted proteins and are thus well-tolerated by the organism. Moreover, like proteins, peptides may form complex structures that can bind selectively to targets and mediate the protein-protein interactions, while many small molecules cannot.<sup>5</sup> Nevertheless, peptide-based therapeutics suffer from multiple drawbacks. The most common is related to their low metabolic stability since they are quickly degraded by proteolytic enzymes which cleave the amide bonds.<sup>6,7</sup>

There are multiple strategies to circumvent this issue. One way is to modify the structure of the peptide (for example by stapling or cyclizing). Another way is to develop molecules that mimic the natural peptides.<sup>6</sup> There are several molecules belonging to this peptidomimetic class, such as the  $\beta$ -peptides, arylamides or oligoureas,<sup>8</sup>

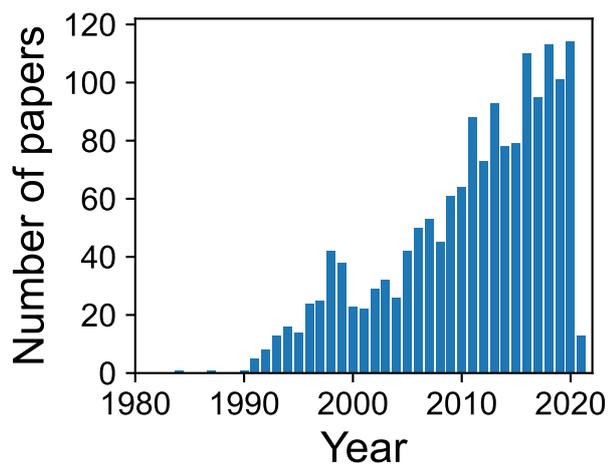
developed to mimic the nature in its whole: mimic the structures of natural peptides by combining a large variety of small building blocks that can be linked together easily and efficiently. Since short peptides are known to bind efficiently to different targets, some research groups focused their attention on molecules close to peptides featuring a large chemical diversity, a polar backbone, an improved resistance to proteolysis and amenable to automation.<sup>9</sup> In the late 1980s, the company Protos (later *Chiron Corporation*) developed a research on building blocks presenting an amide bond, insuring the resemblance to amino acids.<sup>4</sup> The candidate that stood out the crowd was a regioisomer of amino acid, where they only “changed” the side-chain position which moved from the  $\alpha$ -carbon to the amide nitrogen (*i.e.*, *N*-substituted glycines, **Scheme 1.1**).<sup>10</sup>



**Scheme 1.1:** Primary structure of peptides and peptoids as well as the definition of their backbone dihedral angles.

The structure resulting from the coupling of these *N*-substituted glycines was termed “*peptoid*” at that time, although alkyl *N*-substituted glycines were already reported in the literature.<sup>11–13</sup> Over the years, the term “peptoid” has spread in the literature due to their synthetic accessibility, structural tunability and their resistance to proteolysis for therapeutic applications (**Figure 1.1**).<sup>14,15</sup>

The side-chain displacement on the amide nitrogen greatly improves the resistance towards cleavage from protease enzymes compared to their peptide counterpart, which makes them excellent substitutes where peptides failed.<sup>17</sup> The presence of the side chain on the amide nitrogen has also other implications. First, the peptoid backbone is deprived of any stereocenter. Moreover, while the *cis/trans* equilibrium



**Figure 1.1:** Number of papers dealing with peptoids reported as a function of the year (Source: <sup>16</sup> data gathered on [www.scopus.com](http://www.scopus.com) using the “\*peptoid\*” keyword).

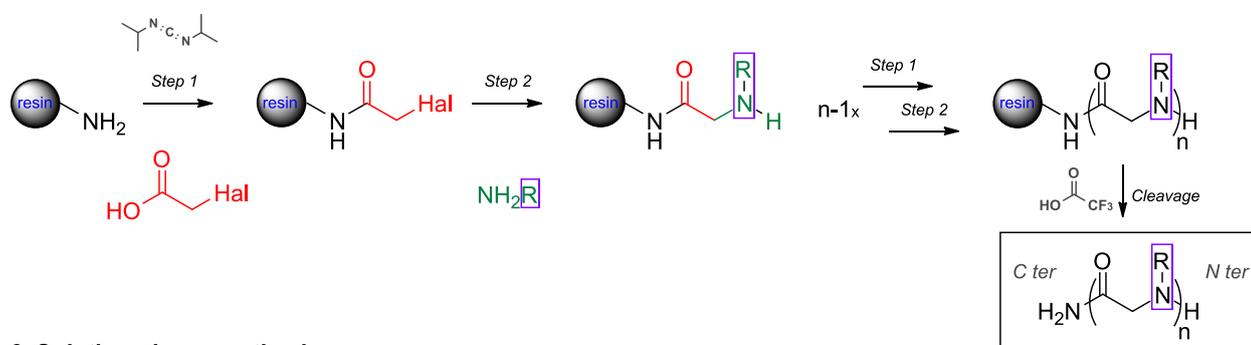
of the amide conformation in peptides favors the *trans* form, this preference is lost in peptoids due to the tertiary character of the amide bond and the absence of stereocenters. Consequently, the *cis/trans* isomerization is likely to occur although the barrier between is quite high, around 15 to 20 kcal/mol.<sup>18,19</sup> Another consequence of the side-chain displacement is the absence of hydrogen bond donor, preventing peptoids to achieve structural stabilization through this means. Despite being achiral, it is possible to reintroduce chirality in the peptoid through the side chains by carefully selecting the primary amine during the synthesis. It is also possible to reintroduce hydrogen bonding donors by design.

## 1.2 Solid-phase synthesis: the golden standard

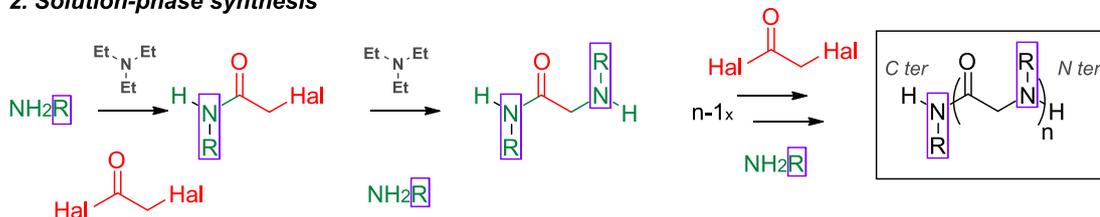
The main synthesis protocol to produce peptoids is also inspired from the peptide synthesis, and consists in growing the peptoid chain on a solid support (*i.e.*, a polymeric resin bead),<sup>13</sup> although other methods are currently used to produce polypeptoids, as it will be discussed later.<sup>20,21</sup> The most efficient method to obtain oligomers called “*sub-monomer solid-phase synthesis*” was developed by the group of **R. Zuckermann** in the early 1990s, which produces peptoids from C to N terminus.<sup>13</sup> Indeed, like peptides, peptoids are also characterized by C and N extremities. By convention, their residues sequence is always given from N to C. The synthesis generally requires the Rink-amide resin, which first needs to be deprotected, and then involves  $n$  cycles of two steps: (*i*) an acylation, for the first cycle on the primary amine bound to the resin, and for the following cycles on the secondary amine of the growing peptoid chain, with a haloacetic acid (typically bromoacetic acid); and (*ii*) a nucleophilic displacement of the halogen by a primary amine that carries the side chain represented

by “R” (Figure 1.2). The number of cycles  $n$  is chosen to match the targeted length of the desired peptoid (up to 60).<sup>22</sup> Once the desired chain is reached, it is cleaved from the resin. Each submonomer is labelled by an abbreviation derived from its chemical name. For example, *N*-methylglycine (also called *sarcosine*) can be used as the simplest sub-monomer and is abbreviated *Nsar*. If one performs 10 cycles of the two steps detailed above and cleaves the chain, one will obtain a *Nsar* decamer, also noted *Nsar*<sub>10</sub>. This method is powerful since it leads to monodisperse oligo/polymers (*i.e.*, every chain has the same number of units). Moreover, the primary amine used in the second step can be changed in each cycle. Consequently, the chemical and structural diversity of peptoids is tremendous since there are hundreds of amines commercially available.<sup>4</sup>

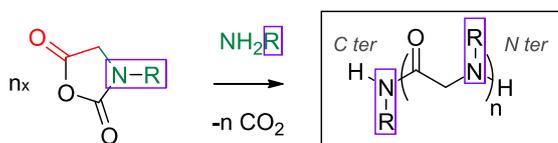
### 1. Solid-phase synthesis



### 2. Solution-phase synthesis



### 3. Ring-opening polymerization



**Figure 1.2:** Different synthetic protocols to produce peptoids. *R* is the side chain borne by the amine and later on the nitrogen amide in the peptoid backbone.

Other synthesis methods have been developed, such as the solution-phase synthesis,<sup>23,24</sup> which does not require a resin. Unlike the solid-phase synthesis, it only requires a single cycle of two steps to synthesize the peptoids. The first step consists in forming a sub-monomer, during which the acylation of a primary amine (or of the amine at the *N*-terminus extremity of the growing peptoid) is achieved with a haloacetyl halide, typically bromoacetyl bromide. It is followed by a nucleophilic displacement of the halide by another primary amine to lead to the formation of the

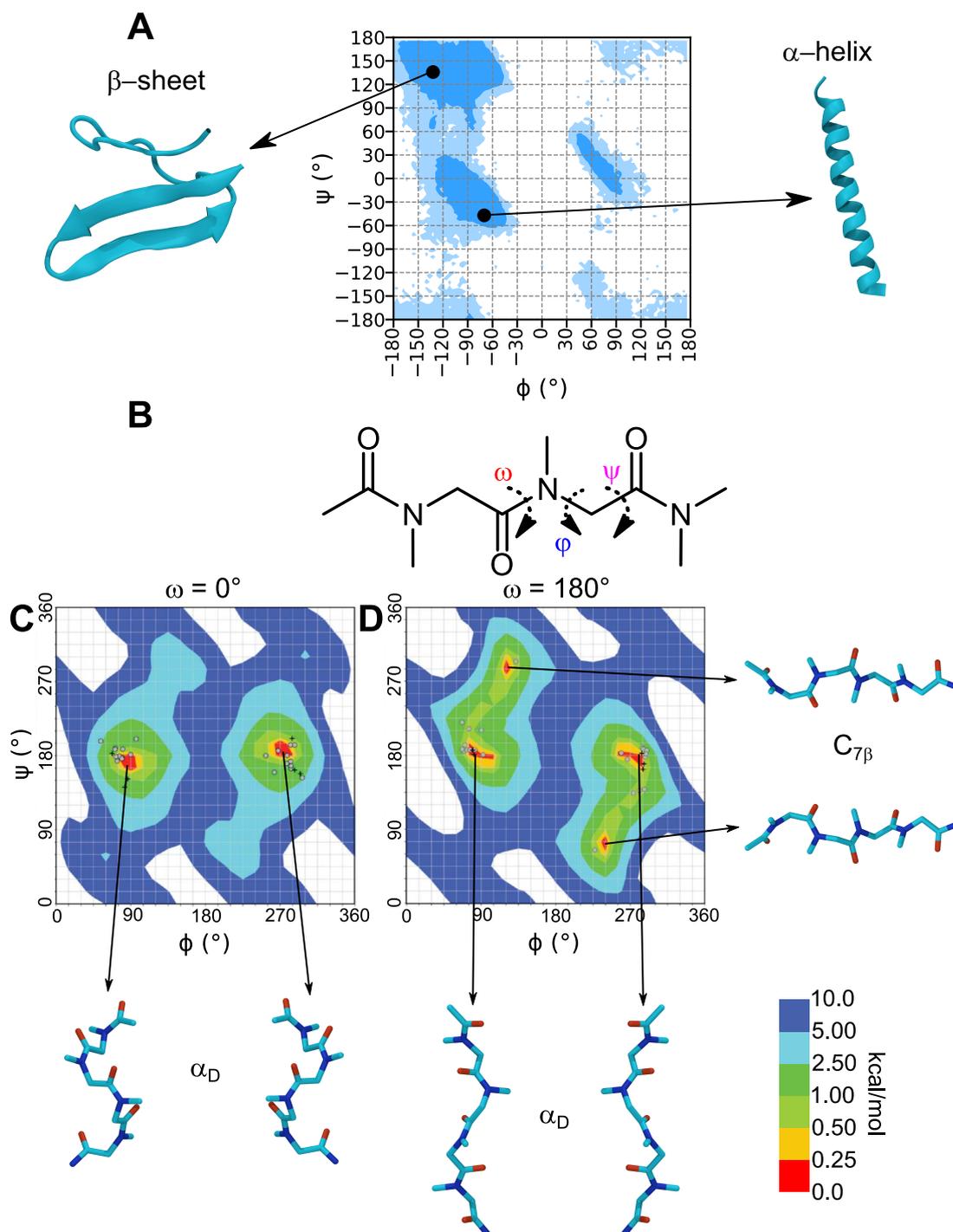
first peptoid unit or to keep growing the chain length (**Figure 1.2**). Compared to solid-phase synthesis, the final peptoid has an additional side chain on the C terminus amide. While this method allows producing larger quantities of targeted peptoids than the solid-phase synthesis, it suffers from multiple drawbacks, such as a poor control of the peptoid sequence, due to unwanted couplings of growing oligomers, and is often limited to shorter oligomers ( $< 10$  residues). Therefore, to obtain a low dispersity and a good control of the sequence, a chromatographic purification is required between the first and second steps of each cycle to remove any primary amine in excess (optional in solid-phase synthesis), which is at the end equivalent to performing  $n$  cycles to generate a peptoid with  $n$  residues.<sup>23</sup>

The last method relies on the ring-opening polymerization (ROP) of  $N$ -substituted glycine  $N$ -carboxyanhydrides (NNCAs, **Figure 1.2**) and is best suited for the synthesis of very long chains. Indeed, long chains are currently inaccessible via the previous methods.<sup>25,26</sup> Moreover, it provides access to various architectures such as star-shaped polymers,<sup>25</sup> at the cost of a higher chain length dispersity than with the solid-phase synthesis (although it can be relatively low, *i.e.*, 1.05-1.25).<sup>26</sup> Using this protocol, the final chain also has an extra side chain on the C terminus amide.

### 1.3 Conformational control

Peptides have been extensively studied and are known for their folding into specific secondary structures depending on the combination of the 3 dihedral angles defining each residue. These dihedrals,  $\omega$ ,  $\phi$  and  $\psi$ , are depicted in **Scheme 1.1**, and are similarly defined in peptoids. The *trans* form of the  $\omega$  dihedral is generally preferred in most peptides, while the dihedrals  $\phi$  and  $\psi$  can adopt different combinations, leading for example to the well-known  $\alpha$ -helices ( $\omega = \sim 180^\circ$ ,  $\phi = -60^\circ$ ,  $\psi = -45^\circ$ ) or  $\beta$ -sheets ( $\omega = \sim 180^\circ$ ,  $\phi = -135^\circ$ ,  $\psi = 135^\circ$ ). The conformational landscape of peptides was widely studied on short peptide model (one to two residues) since the pioneering work of Ramachandran,<sup>32</sup> who established a two-dimensional map of the  $\phi$  and  $\psi$  dihedrals to assess secondary structures, such as the  $\alpha$ -helices or the  $\beta$ -sheets (**Figure 1.3**).<sup>33,34</sup>

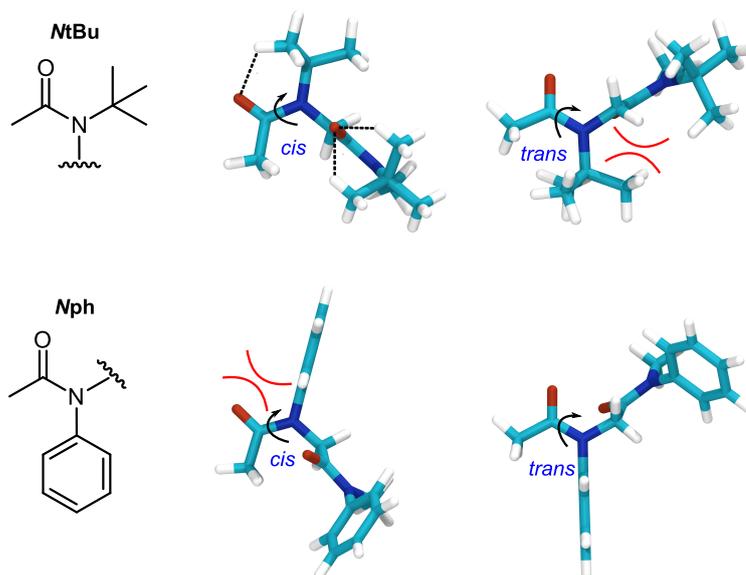
Despite the structural similarities between peptides and peptoids, their conformational landscapes are different, as it was first reported by Moehle *et al.* who studied the low-lying energy conformers of a small model peptoid ( $N,N$ -dimethyl-2-( $N$ -methylacetamido)acetamide), whose conformational dynamics was monitored using the Ramachandran plots, as for peptides.<sup>35</sup> Peptide Ramachandran plots are generally represented with the dihedral  $\phi$  varying along the X axis and the dihedral  $\psi$  along



**Figure 1.3:** (A) Peptide Ramachandran plot generated with data from Ref. 27 using the Python MDAnalysis package.<sup>28,29</sup> The amide bond is mainly in *trans* conformation.<sup>a</sup> The deep blue areas correspond to the most frequent dihedral combinations encountered in proteins, while the light blue areas correspond to the less frequent. (B) Peptoid model with  $N$ ser residues used by Butterfoss *et al.* to generate the Ramachandran-like plots (C) and (D) (B3LYP/6-31+1G(2d,p)// HF/6-31G\*).<sup>30</sup> The dihedrals involved in the plots are defined in the generic structure. Since peptoids sample both *cis* and *trans* amide bond conformation, a Ramachandran-like plot is generated for both, with each local minimum highlighted and illustrated by a peptoid sequence of 4 residues. The grey dots and black crosses in the plots correspond to experimental data obtained from NMR and XRD.<sup>30</sup> Hydrogens were omitted for clarity. Adapted with permission from Ref. 30. Copyright 2009 American Chemical Society.

<sup>a</sup>The  $\alpha$ -helix is rendered using VMD<sup>31</sup> from one of three chains of the “1COS” PDB (<https://www.rcsb.org/structure/1COS>). The  $\beta$ -sheet is rendered using VMD<sup>31</sup> from the last entry of “1ICL” PDB (<https://www.rcsb.org/structure/1ICL>).

the Y-axis, both ranging from  $-180^\circ$  to  $180^\circ$ . In contrast, peptoid Ramachandran-like plots are generally represented from  $0$  to  $360^\circ$  (**Figure 1.3**). Compared to peptides, the amide bonds of peptoids is tertiary and can thus sample both *cis* and *trans* conformations.<sup>36</sup> This unique feature greatly enhances their structural diversity. Moreover, peptoids display new structures at energy minima, such as “ $\alpha_D$ ” ( $\omega = \sim 0^\circ$  or  $\sim 180^\circ$ ,  $\phi = \pm \sim 90^\circ$ ,  $\psi = \sim 180^\circ$ ) or “ $C_{7\beta}$ ” ( $\omega = \sim 0^\circ$  or  $\sim 180^\circ$ ,  $\phi = \pm \sim 90^\circ$ ,  $\psi = \pm \sim 75^\circ$ ), which are not present in the peptide energy landscape (**Figure 1.3**).<sup>10,35</sup> It should be emphasized that the  $\alpha_D$  conformation is present on both sides of a given Ramachandran-like plot ( $\phi$  with positive and negative values while  $\omega$ , in either *cis* or *trans*, and  $\psi$  are equivalent) because of the achiral nature of the model peptoid, which renders the “right” and “left” forms degenerate. The “right” form will be abbreviated as  $\alpha_D$  and the “left” form as  $\alpha_{D+}$  for the rest of the manuscript. Although the conformational landscape is rich, it is often desirable to narrow it to a single minimum, which can generate a well-defined repeated pattern and thus a highly ordered conformation. However, the conformational landscape of the model peptoid is highly symmetric (**Figure 1.3**). Each conformation is found in both “right” and “left” forms, which prevents the formation of a well-defined sequence.



**Figure 1.4:** Illustration of the amide conformational preference for the *NtBu* and *Nph* side chains in simple models ( $\phi = -90^\circ$ ,  $\psi = 180^\circ$ ). Steric clashes are highlighted by curved red lines while favorable interactions are highlighted by dashed black lines.

### 1.3.1 *cis/trans* isomerism

The perfect control of the *cis/trans* content is particularly interesting to promote order and create specific secondary structures. However, only a few side chains can

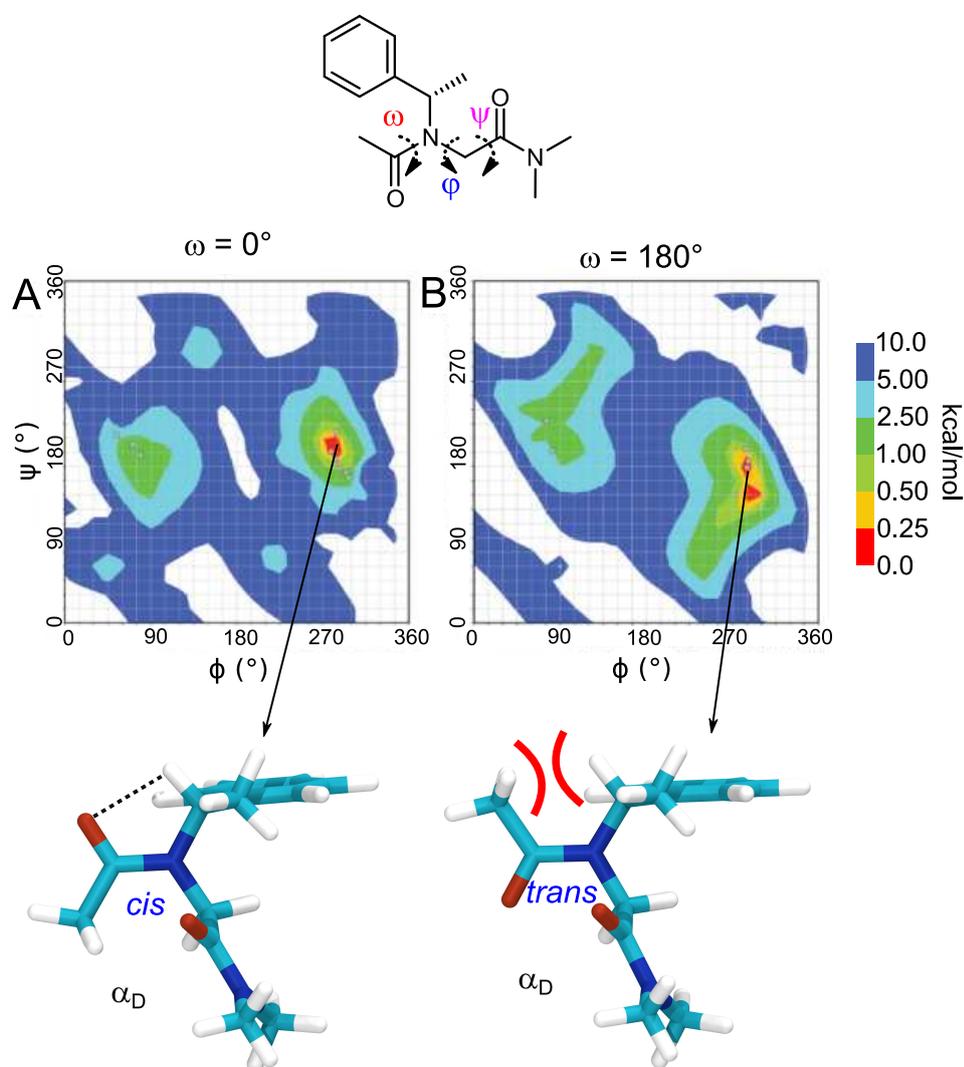
achieve such a fine control. For example, the *NtBu* (*N-tert-butyl glycine*) is an achiral bulky side chain that promotes exclusively the *cis* amide conformation in peptoids whatever the solvent used.<sup>37</sup> The isomerism control mainly arises from steric hindrance that promotes a high steric clash in the *trans* form (**Figure 1.4**). On the other hand, the *Nph* (*N-aryl glycine*) side chain promotes exclusively the *trans* amide, because of electronic repulsion between  $\pi$ -electrons from the aryl moiety and carbonyl lone pairs in the *cis* form (**Figure 1.4**).<sup>38</sup> In general, most side chains will lead to a mixture of *cis* and *trans* isomers, and the ratio will also depend on the nature of the solvent.<sup>39,40</sup>

### 1.3.2 Chirality

Whereas a nearly perfect control of the *cis/trans* isomerism is possible, there is still a symmetrical conformational landscape with degenerate conformational forms. Introducing chirality is a way to break down the symmetry while simultaneously conserving a good control over the *cis/trans* isomerism. This is a particularly challenging task, even with the expertise gathered from the last two decades.

One of the most prototypical and of earliest studied chiral side chain is the (*S*)-*N*-(1-phenylethyl)glycine (abbreviated *Nspe*), which consists in a bulky aromatic  $\alpha$ -chiral disubstituted side chain.<sup>36,41,42</sup> It was suggested by early computational predictions that such a bulky chiral disubstituted side-chain has two impacts on the conformational landscape.<sup>41</sup> First, as discussed previously, the side chain dictates the amide bond conformation. In this case, the amide bonds tend to be preferentially in their *cis* conformation, to prevent steric clashes between the backbone methylene and the methine from the side chain in the *trans* conformation (**Figure 1.5**).<sup>41</sup> Second, the chirality of the side chain acts on the conformational landscape and governs the handedness of the conformation which results in a single  $\alpha_D$  minimum in the Ramachandran-like plot (**Figure 1.5**). For such a side chain, the (*S*) enantiomer will promote the  $\alpha_D$  conformation (right-handed) while the opposite (*R*) enantiomer will promote the other ( $\alpha_{D+}$ , left-handed). A careful selection of the side chain can thus help to achieve a total control over the conformational landscape and promote the formation of highly ordered conformations.

Accordingly, peptoids are associated to the foldamer family, which is defined by Gellman as any synthetic polymer/oligomer that has a strong tendency to fold into ordered conformations, generally mimicking those formed by natural biopolymers such as proteins, nucleic acids or polysaccharides.<sup>43</sup>



**Figure 1.5:** Ramachandran-like plots of (A) *cis* and (B) *trans* conformer of the peptoid model with a single *N*spe side chain generated by Butterfoss *et al.* (B3LYP/6-31+1G(2d,p)// HF/6-31G\*).<sup>30</sup> The dihedrals involved in the plots are highlighted in the primary sequence on top. The *trans* conformer is slightly less stable than the *cis* because of steric hindrance between the methyl and the methine hydrogen (curved red lines), while a favorable CH...O is present in the *cis* conformer (dashed black line). The grey dots and black crosses in the plots correspond to experimental data obtained from NMR and XRD.<sup>30</sup> Hydrogens were omitted for clarity. Adapted with permission from Ref. 30. Copyright 2009 American Chemical Society.

## 1.4 Peptoid characterization

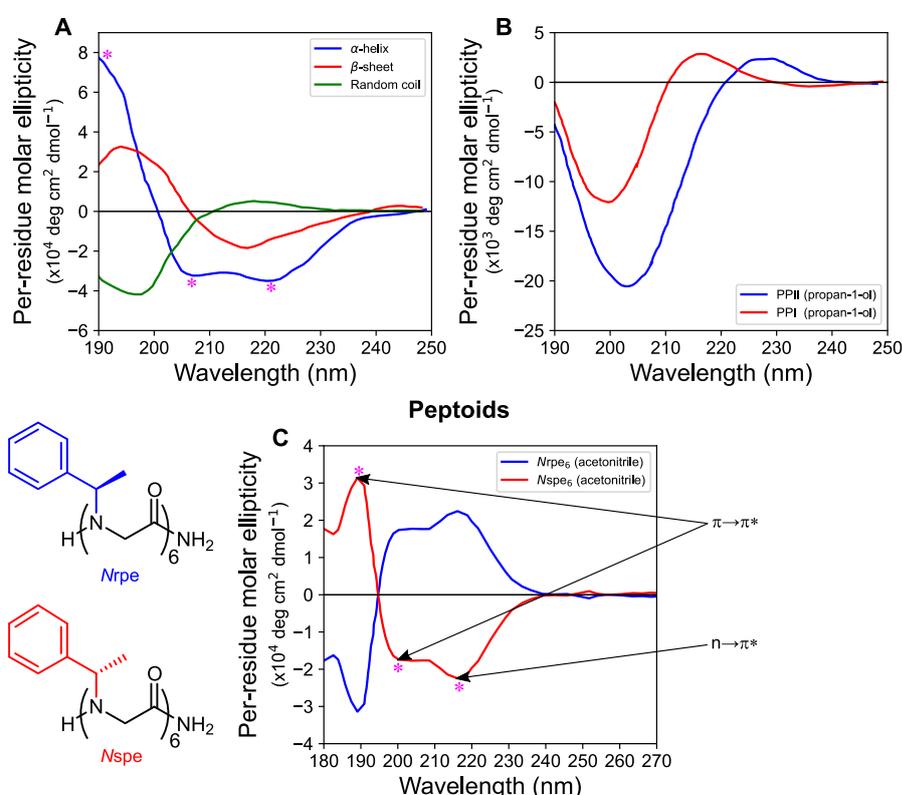
Even though peptoids have a simple backbone, the versatility in the choice of the side chain strongly complexifies their behavior. Therefore, the combination of multiple characterization methods, inspired from peptides investigations, is required to shed light on the different factors driving their folding.

### 1.4.1 Circular dichroism

With the rapid introduction of chiral side chains to restrain the conformational landscape of peptoids, circular dichroism spectroscopy (CD) became the reference technique to quickly assess the presence of chiral secondary structures. Indeed, unlike most other polarized spectroscopies that rely on the measurement of the absorption with linearly polarized light, circular dichroism spectroscopy tracks the difference of absorption between two incident light components, one left-hand circularly polarized and the other right-hand circularly polarized.<sup>44</sup> Only chiral optically active species will give rise to a non-null absorption difference between these two polarizations, commonly called “CD signal”. We should emphasize that the chirality of the sample can of course originate from stereocenters, but also from conformational chirality. For example, a helix is conformationally chiral, as it can be left- or right-handed, and can also contain additional stereocenters. In this regard, CD is particularly useful in the field of proteins, as their tertiary structures are actually formed by the association of particular secondary structures, such as  $\alpha$ -helices or  $\beta$ -sheets, in which chirality arises from the stereocenters of the constituting amino acids and from the conformation adopted by the sequence of amino acids.<sup>45</sup> Each of these secondary structures display its own CD signature, typically in the absorption range between 190 and 300 nm, which allows to first identify them and then quantify their content in a given protein.<sup>45</sup>

The first published chiral peptoids studied by CD spectroscopy were made of *Nspe* (or *Nrpe*) side chains, and displayed a CD signature reminiscent of the peptide  $\alpha$ -helix (**Figure 1.6A**),<sup>46</sup> that is characterized by three bands: two minima at  $\sim$  220 and 210 nm, followed by a maximum at 190 nm for *Nspe*, and exactly the opposite signal for *Nrpe* (**Figure 1.6C**).<sup>41,47,48</sup> On that basis, it was strongly suggested that these peptoids adopt a helical conformation, which nicely correlates with the early computational predictions.<sup>41</sup> However, the predicted helix does not correspond to an  $\alpha$ -helix geometry but rather to a poly-proline type I (PPI) geometry, although this peptoid signature has barely nothing in common with the PPI signature (**Figure 1.6B-C**).<sup>41–43,45–47,49</sup> Actually, the CD signature of peptide is believed to mainly arise

from the regular arrangement of the amide bonds that differs between each secondary structure.<sup>50</sup> Despite these differences, the transitions associated to the bands were directly assimilated to those of peptides, *i.e.*, the bands around 190 nm and 200 nm are assigned to the high and low wavelength components of  $\pi \rightarrow \pi^*$  transitions, while the band around 220 nm is assigned to  $n \rightarrow \pi^*$  transitions.<sup>42</sup> However, it cannot be excluded that these bands have additional contributions and contain a mixing between the amide and aromatic transitions, as observed with certain peptides containing aromatic amino acids,<sup>42</sup> such as L-phenylalanine or L-tyrosine.<sup>51–53</sup> Until now, there are no reports of theoretical data to support this interpretation, although it would validate these assumptions and indicate which factors do affect the shape of the CD signature.



**Figure 1.6:** Circular dichroism spectra of the main secondary structures encountered in (A) peptides<sup>a</sup> 45 and (B) proline-rich peptides<sup>b</sup> 49 compared to the circular dichroism spectra of (C) *Nrpe* and *Nspe* hexameric peptoids.<sup>c</sup> 42 The main transitions are indicated in the spectrum of *Nspe*<sub>6</sub> at 190, 200 and 220 nm and are similar to those encountered with the  $\alpha$ -helix (A, magenta stars).

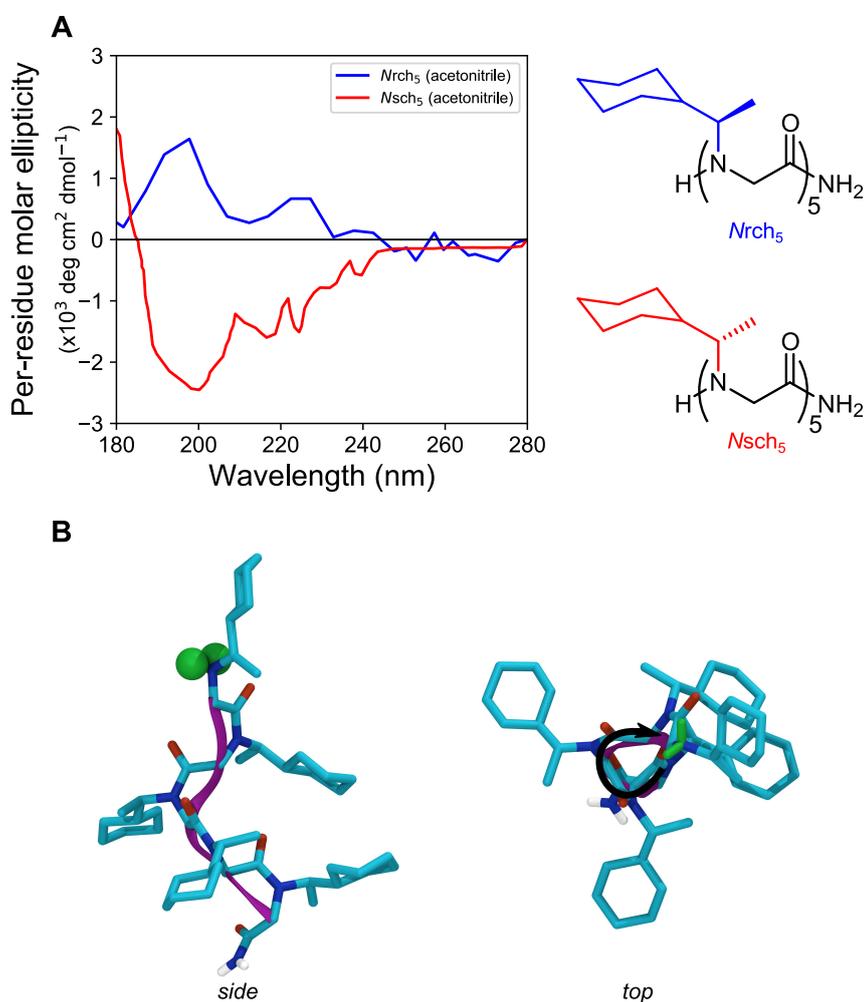
<sup>a</sup>The circular dichroism spectra were adapted with permission from 45. Copyright 2007 Springer Nature. The  $\alpha$ -helix and  $\beta$ -sheet are formed by poly-L-lysine at pH 11.1 while the random coil is formed at pH 5.7.

<sup>b</sup>The circular dichroism spectra were adapted with permission from 49. Copyright 2020 John Wiley and Sons. The PPII is formed by the P6GP6 peptide in propan-1-ol after 2 minutes, while the PPI is formed after 14 days (both at 5°C).

<sup>c</sup>The circular dichroism spectra were adapted from Ref. 42.

### 1.4.2 X-ray diffraction

While CD cannot provide atomic details about peptoids but rather an average ensemble signature, X-ray diffraction (XRD) is the predilection technique to obtain such information when crystals can be obtained.<sup>54</sup> However, only a few X-ray diffraction data are currently available in the literature. Indeed, it was already reported in the early 00s that it is difficult to grow crystal of peptoid oligomers.<sup>42</sup> It is suspected that the amide bond *cis/trans* isomerism grants a large backbone flexibility which prevents peptoid crystallization.<sup>55</sup> According to the PeptoidDataBank,<sup>56</sup> only 46 crystalline structures are reported up to early 2021. Among these, only 24 are related to linear rather than cyclic peptoids (the latter are cyclized between the N and C termini).<sup>57</sup>

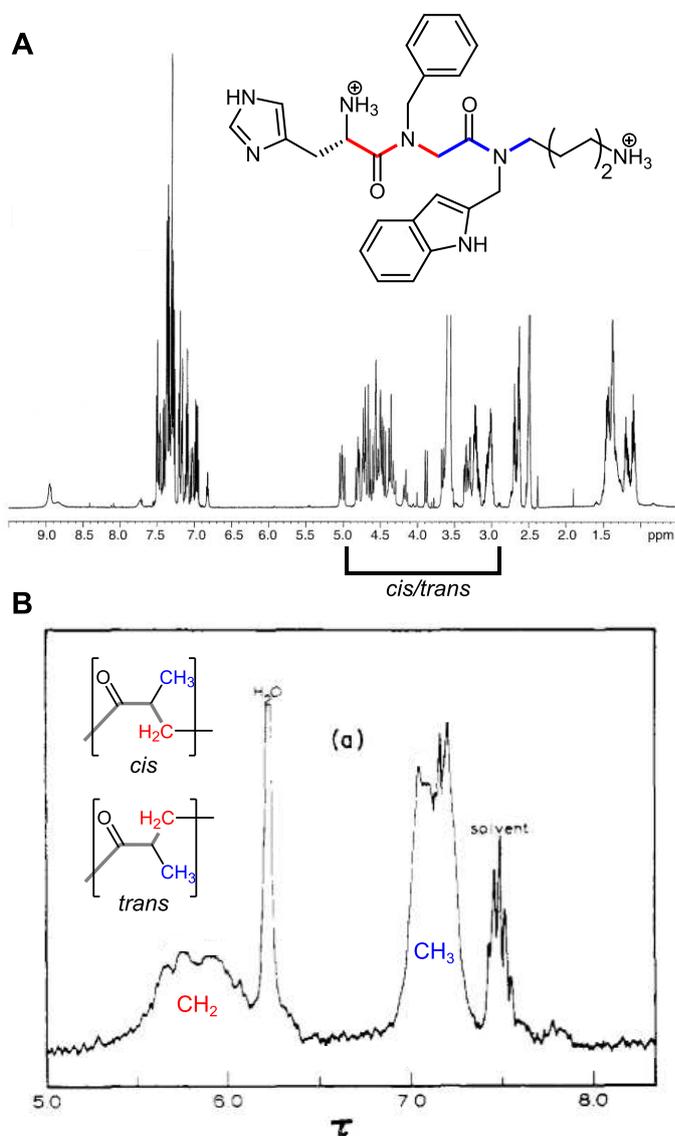


**Figure 1.7:** (A) Circular dichroism spectra of *Nrch*<sub>5</sub> and *Nsch*<sub>5</sub> peptoids.<sup>58</sup> Their signature is similar to the *Nspe/Nrpe*-type peptoids. (B) Crystalline structure of *Nrch*<sub>5</sub> peptoids from side and top view.<sup>58</sup> The helix is left-handed, as indicated by the black arrow in the top view representation. The hydrogens are omitted for clarity except for the polar hydrogens at the C terminus and the protons at the N terminus (green balls/sticks). Adapted with permission from Ref. 58. Copyright 2003 American Chemical Society.

Moreover, most of the peptoids able to crystallize are only composed of maximum two to three residues. Nonetheless, it has been possible to crystallize a pentamer bearing *N*rch side chains ((*R*)-*N*-(1-cyclohexylethyl)glycine),<sup>58</sup> which is very similar to the *N*spe side chain, except that the phenyl is replaced by a cyclohexyl moiety (**Figure 1.7**). Such side chains confer to the peptoid a helical conformation similar to the polyproline type I helix (left-handed),<sup>58</sup> as suggested by the early computational predictions.<sup>41</sup> Interestingly, the backbone dihedral angle values obtained by XRD are always in the same range, irrelevant of the linear or cyclic nature of the peptoids.<sup>59</sup> Unfortunately, as already pointed out, XRD remains quite an anecdotic technique because of the inability of linear peptoids to crystallize.

### 1.4.3 Nuclear magnetic resonance

NMR spectroscopy is a relevant alternative to XRD as it can also provide detailed atomistic information.<sup>60</sup> NMR spectroscopy provides specific information about the chemical environment of the active nuclei, typically <sup>1</sup>H and <sup>13</sup>C, and hence affords crucial data on the primary and secondary structures of the sampled molecules. 1-D NMR has been used extensively to elucidate the structure of “small” molecules, but is quite complex for large molecules, such as proteins.<sup>61</sup> Indeed, the large number of atoms often give rise to overlaps of the signals and thus prevents a correct interpretation. Moreover, contrary to peptides in which a single amide conformation is sampled, the two isomeric forms are typically probed for peptoids. Since *cis* and *trans* amide groups have a different chemical environment, they give rise to multiple transitions, that can already be observed in model compounds such as *N,N*-dimethylacetamide.<sup>62</sup> Therefore, if we consider even simple peptoid oligomers, which are constituted by only a few amide bonds in both the *cis* and *trans* conformations, the NMR spectra rapidly become difficult to interpret and it gets even worse for longer chains (**Figure 1.8A-B**).<sup>63</sup> Nonetheless, this technique is still relevant to analyze short oligomers, especially to determine the ratio between the *cis* and *trans* population (and hence the equilibrium constant  $K_{cis/trans}$ ) for a given side chain in a given solvent.<sup>64</sup> To gain further insights into the peptoid conformation, it is possible to perform 2-D NMR studies which provide complementary information, for example about the spatial coupling between spins of different nuclei.<sup>47</sup> Interestingly, the analysis of multiple 2-D NMR experiments allowed to identify that a peptoid pentamer bearing *N*spe side chain derivatives adopts mainly a right-handed helix with a PPI-like geometry (about 45% of the population), while the remaining conformations are rather disordered due to the *cis/trans* equilibrium.<sup>47</sup>



**Figure 1.8:** (A) 1-D  $^1\text{H}$  NMR spectrum of a peptoid oligomer.<sup>65</sup> The region between 3 and 5 ppm is dense because of the 4 possible rotamers accessible from the *cis/trans* isomerization. Adapted with permission from Ref. 65. Copyright 2005 Elsevier. (B) 1-D  $^1\text{H}$  NMR spectrum of a polysarcosine peptoid (94 residues). The distinction of the *cis/trans* populations is difficult due to the overlap of the different combinations in the  $\text{CH}_2$  and  $\text{CH}_3$  regions.<sup>63</sup> Adapted with permission from Ref. 63. Copyright 1968 American Chemical Society.

#### 1.4.4 Mass spectrometry

Over the last two decades, most experimental characterization techniques shed light on peptoid conformations and highlighted factors favoring their formation. As a prerequisite, the primary structure must first be properly characterized to confirm that the peptoid has the desired sequence. Mass spectrometry (MS) is a method of choice in this respect, as it provides information about the mass-to-charge ratio ( $m/z$ ) of gaseous ions.<sup>66</sup> Indeed, the  $m/z$  directly reflects the composition of the ion. To completely elucidate the primary structure, the ion can be sequenced into fragment ions *via* collision-induced dissociation (CID). This method provides information about the connectivity of the different fragments, and thus allows us to unravel the sequence of residues inside the peptoid.<sup>67-69</sup>

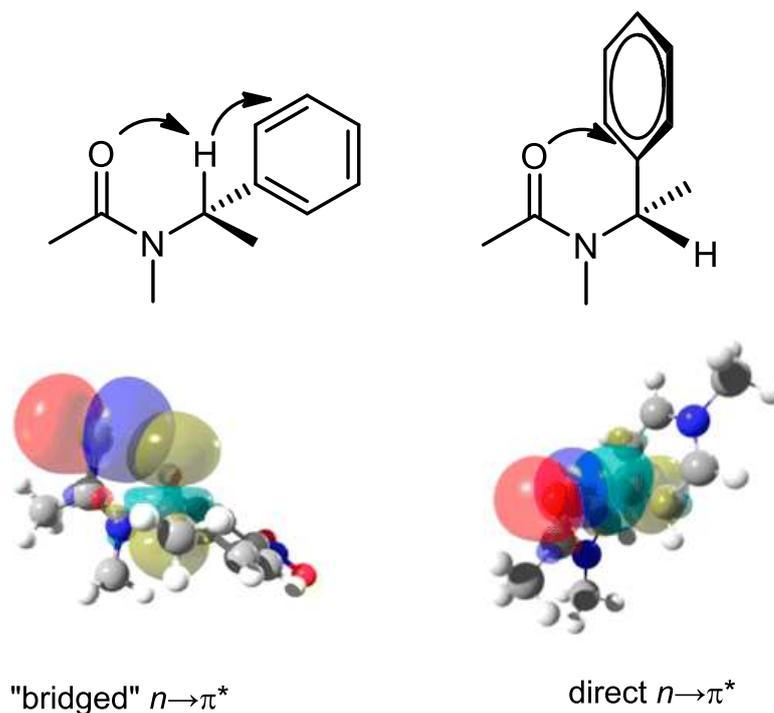
Moreover, MS can be coupled to a technique called “ion mobility spectrometry” (IM-MS) which provides information about the tridimensional shape of the ions, and thus about their conformation in gas phase.<sup>70</sup> Briefly, let us consider two ions with identical  $m/z$  but with different shapes. These ions are injected simultaneously into a mobility cell, which is filled with a gas and on which an electric field is applied. Both ions will travel through the cell and undergo collisions with the gas (that do not lead to their fragmentation), whose frequency will depend on how extended/compacted the ions are. As a result, the more extended the ion, the longer time required to travel the mobility cell. The drift time ( $t_D$ ) inside the mobility cell is directly related to the mobility ( $K$ ) of the ions. Moreover, a relation can be established between the mobility and the collision cross section ( $\Omega$ , more precisely defined as the rotationally averaged collision cross section), which is directly related to the size and shape of the ion and hence allows to derive structural information. In first approximation, the collision cross section can be seen as the surface of the ion effectively colliding with the gas. IM-MS is rarely used as a standalone technique for the elucidation of ion conformations. Most of the time, molecular modeling goes hand in hand with IM-MS to provide candidate three-dimensional structures matching the experimental results.<sup>71</sup> The IM-MS technique is not yet widespread in the peptoid world, although it can bring relevant elements about their structure. Since ions are in a vacuum-like environment during (IM)-MS measurements, the different external factors (in particular the solvent) influencing their conformational landscape are not affecting the measurements. IM-MS can thus help understanding the intrinsic factors responsible for their folding, as it has been done with peptides.<sup>72-75</sup> Moreover, IM-MS may allow to distinguish each individual conformers, provided their collision cross sections differ.

### 1.4.5 Molecular modeling

Computational chemistry was already mentioned a couple of times in the previous sections either as a prediction method or as a complementary tool to understand experimental data and support their interpretation. Molecular modelling has played an important role in the understanding of the peptoid conformational space during the early stages of peptoid research, as theory predicted that the *N*spe side chain would promote the formation of a right-handed helical conformation, and thus that peptoids are foldamers. Computational modelling is increasingly used in peptoid studies and will continue to play a crucial role in the field. Indeed, due to the large variety of primary amines available (and thus side chains) to build peptoids, the number of possible combinations is tremendous, even for short oligomers. Therefore, computational approaches are attractive since they permit to screen *in silico* a large variety of compounds and help to predict their conformations.

Peptoid folding can be approached theoretically at different levels of complexity. Quantum-chemical calculations, based in particular on the Density Functional Theory (DFT), are mainly used to determine the conformational preferences of small oligomers (typically made of one or two residues) by identifying the low-lying energy conformers.<sup>30,76,77</sup> The first quantum-mechanical studies were performed in 1996 on model peptoids of sarcosine and identified conformations such as  $C_{7\beta}$  and  $\alpha_D$  (**Figure 1.3**), that are not observed in peptides.<sup>35</sup> Later on, calculations on more complex peptoid side chains, such as *N*spe,<sup>30</sup> were carried out to build Ramachandran-like plots to identify the low-lying energy conformers based on different combinations of the backbone dihedrals. The dihedral combinations in the most stable conformers turned out to be in good agreement with the few high-quality experimental data obtained by X-ray diffraction and NMR studies, which thus demonstrates the relevance to put efforts in the development of molecular modelling techniques for the study of peptoids. Moreover, quantum-mechanical methods can shed light on the nature of the stabilizing intramolecular interactions, such as the “bridged”  $n \rightarrow \pi^*$  interaction that stabilizes the *cis* conformation of an amide bond substituted by a *N*- $\alpha$  chiral side chain, such as *N*spe (**Figure 1.9**).<sup>78</sup> This interaction mode operates concurrently with direct  $n \rightarrow \pi^*$  interactions and explains the preferential orientation of the methine of such *N*- $\alpha$  chiral side chains.

Such non-covalent interactions are only present in *N*- $\alpha$  side chains bearing a methine moiety, as found in *N*spe. The methine acts as an intermediate through its  $\sigma^*$  orbital to transfer electron density from the amide oxygen lone pair to the aromatic ring of the side chain. Quantum-chemical calculations are generally highly accurate but suffer from a high computational cost, even for relatively small oligomers.



**Figure 1.9:** Illustration of the two modes of stabilization the *cis* conformation in presence of *N*- $\alpha$  chiral side chains (here, *N*spe). Adapted with permission from <sup>78</sup>. Copyright 2013 American Chemical Society.

For systems too large to be treated by quantum-mechanical methods, it is still possible to model their structural behavior at a lower level of complexity by using molecular mechanics (MM) and molecular dynamics (MD) approaches that treat electrons implicitly.<sup>79</sup> In this context, the potential energy surface of the system is described by a force field, that is a mathematical expression expressing the potential energy as a function of the atomic positions, based on a set of parameters for bonded and non-bonded interactions. Force field simulations performed on peptoids generally rely on parameters developed for peptides. However, to provide accurate results, parameter adaptation is often recommended,<sup>80</sup> especially to correctly describe the *cis/trans* difference compared to peptides. Some recent force fields were tuned for peptoids, based on experimental data or highly accurate quantum-chemical inputs to correctly describe their conformations.<sup>81,82</sup> Molecular dynamics simulations are typically used to explore the conformational space by resolving Newton's equations of motion to study the preferential folding of peptoids.<sup>79</sup> Moreover, it can also be used to predict the stable conformations of peptoids and hence is a great asset to design new sequences and guide the synthesis towards potentially interesting conformations.<sup>83</sup> Indeed, one key feature of foldamer development consists in understanding and predicting the structural properties at the monomer level to transpose them at the chain size.<sup>43</sup>

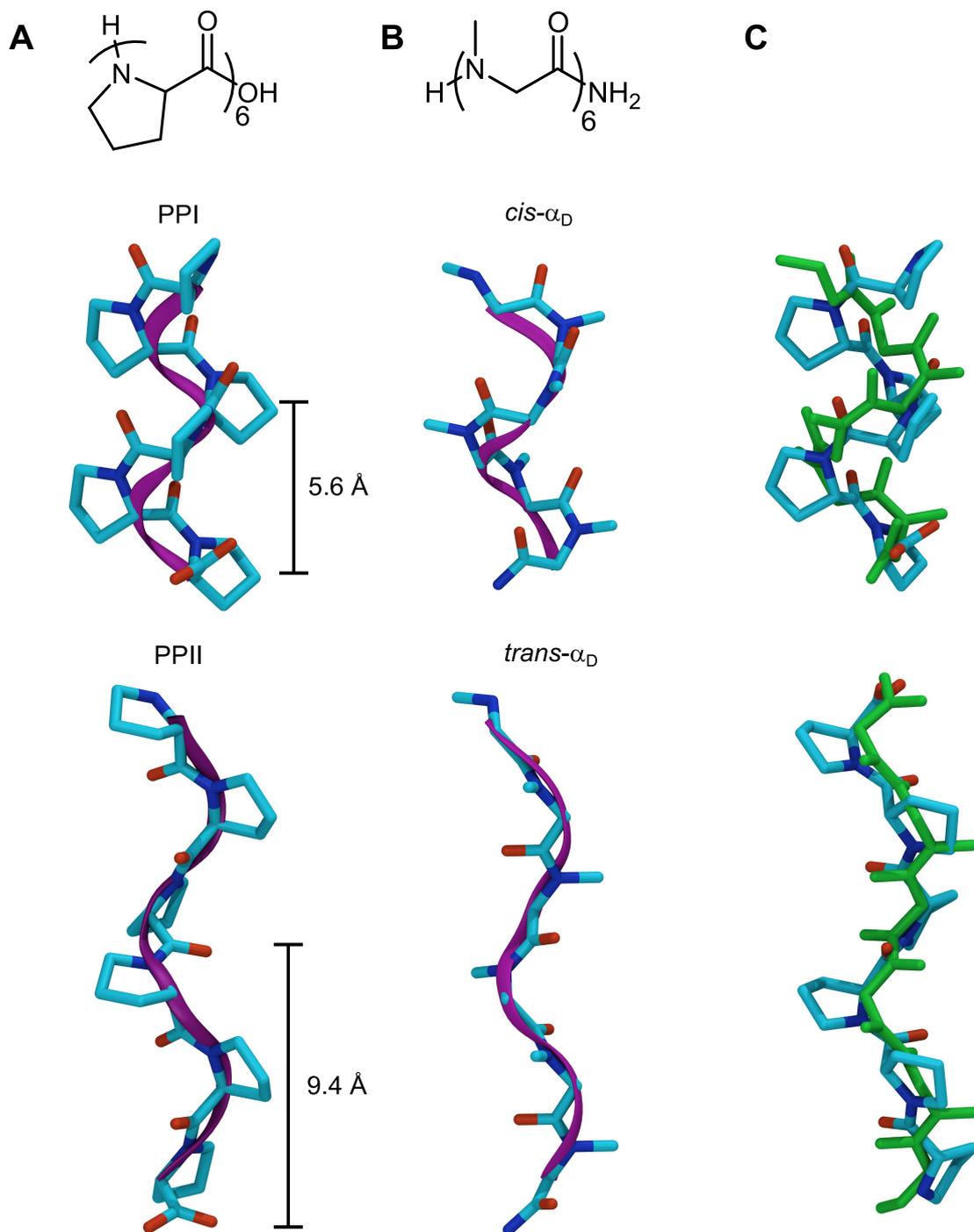
## 1.5 Secondary structure

Peptoids were originally developed to mimic the behavior of small peptides in therapeutic applications but were rapidly envisaged as potential peptide secondary structure mimics as well.<sup>84</sup> Although their primary structure is richer than peptides, only a few secondary structures were identified to date. Overall, the various peptoid secondary structures discussed below appear to remain in their folded state over a wide range of conditions (temperature, solvent, pH), indicating that peptoids are robust foldamers.<sup>85,86</sup>

### 1.5.1 The helical dominance

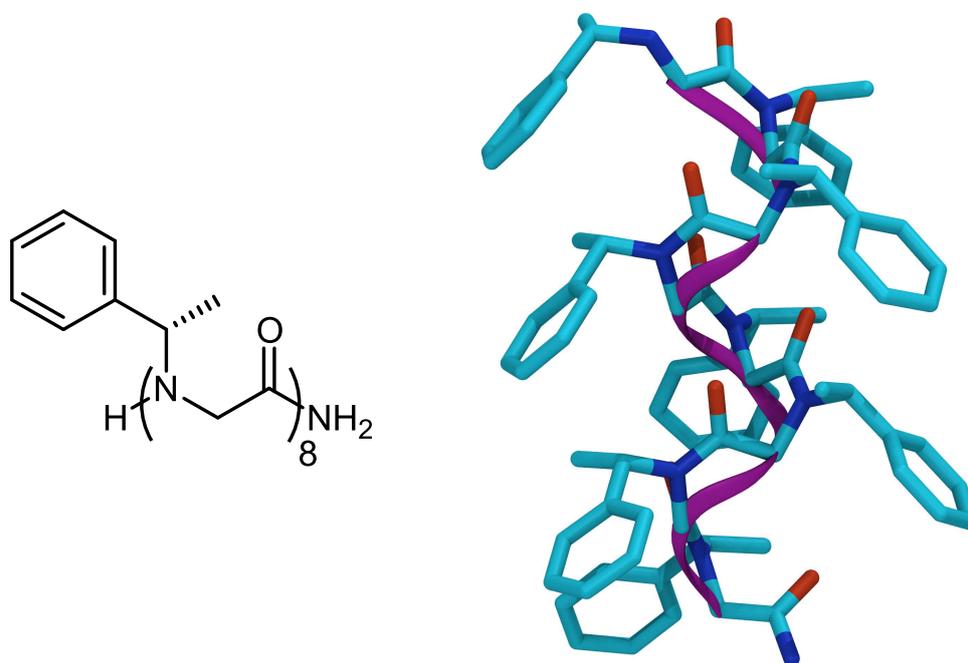
The most abundant secondary structure of peptoids is certainly the helix. This type of structure was predicted by computational calculations almost at the onset of the peptoid development.<sup>41</sup> Helices are of course abundant in peptides, but some of them are particularly relevant in the context of peptoids. Indeed, L-proline is the only natural amino acid that is *N*-substituted, and thus that impacts the *cis/trans* amide proportion by increasing considerably the *cis/trans* ratio (1:3 vs.  $\sim$  1:1000), as generally observed in peptoids (for example, *Nsar*).<sup>87</sup> Poly-proline peptides can adopt a type I or II conformation, depending on the repeating pattern of their amide conformation. The poly-proline type I (PPI) has a full *cis* amide sequence, while the poly-proline type II (PPII) has a full *trans* amide sequence. Both PPI and PPII have their  $\phi$  and  $\psi$  dihedral angles around  $-75^\circ$  and  $150^\circ$ , respectively (**Figure 1.10 A**). Due to these differences in the amide geometry, the PPII helix is much more extended than the PPI, with a helical pitch of 9.5 Å (3.3 residues per turn) and 6 Å (3 residues per turn) respectively.<sup>88</sup> Interestingly, the helices adopted by peptoids strongly resemble the poly-proline helices (**Figure 1.10B**). More precisely, the  $\phi$  dihedral lies either around  $+90^\circ$  or  $-90^\circ$  (left- and right-hand screw sense, respectively) and the  $\psi$  dihedral lies around  $180^\circ$ , both defining the  $\alpha_D$  conformation (**Figure 1.3**). Baldauf *et al.* built a helical model containing 6 sarcosine residues whose dihedrals are based on those obtained from the lowest energy conformer at the monomer level.<sup>89</sup> The all-*cis* and all-*trans* versions of these  $\alpha_D$  conformers superimpose well on the PPI and PPII conformations, and the resulting helix pitches are also very similar. The similarities in the secondary structures between proline chains and peptoids is particularly interesting since peptoids could be used as alternatives for specific applications (**Figure 1.10C**).

Nonetheless, even though sarcosine is relevant to provide insights into the peptoid folding, it is not suited for the formation of well-defined helices. Indeed, due to the chiral nature of a helix, only chiral side chains can induce a preference for a screw sense over the other. Pioneering works on the helical conformation were carried out



**Figure 1.10:** (A) Proline hexamer in its PPI (all-*cis* amide,  $\phi = -75^\circ$ ,  $\psi = 160^\circ$ ) and PPII (all-*trans* amide,  $\phi = -75^\circ$ ,  $\psi = 150^\circ$ ) conformation. (B) Helical model of a peptoid sarcosine hexamer in its *cis*- $\alpha_D$  and *trans*- $\alpha_D$  conformation ( $\phi = -85^\circ$ ,  $\psi = 180^\circ$ ). (C) Structural overlap between PPI and *cis*- $\alpha_D$  (in green) as well as PPII and *trans*- $\alpha_D$  (in green). All hydrogens were omitted for clarity.

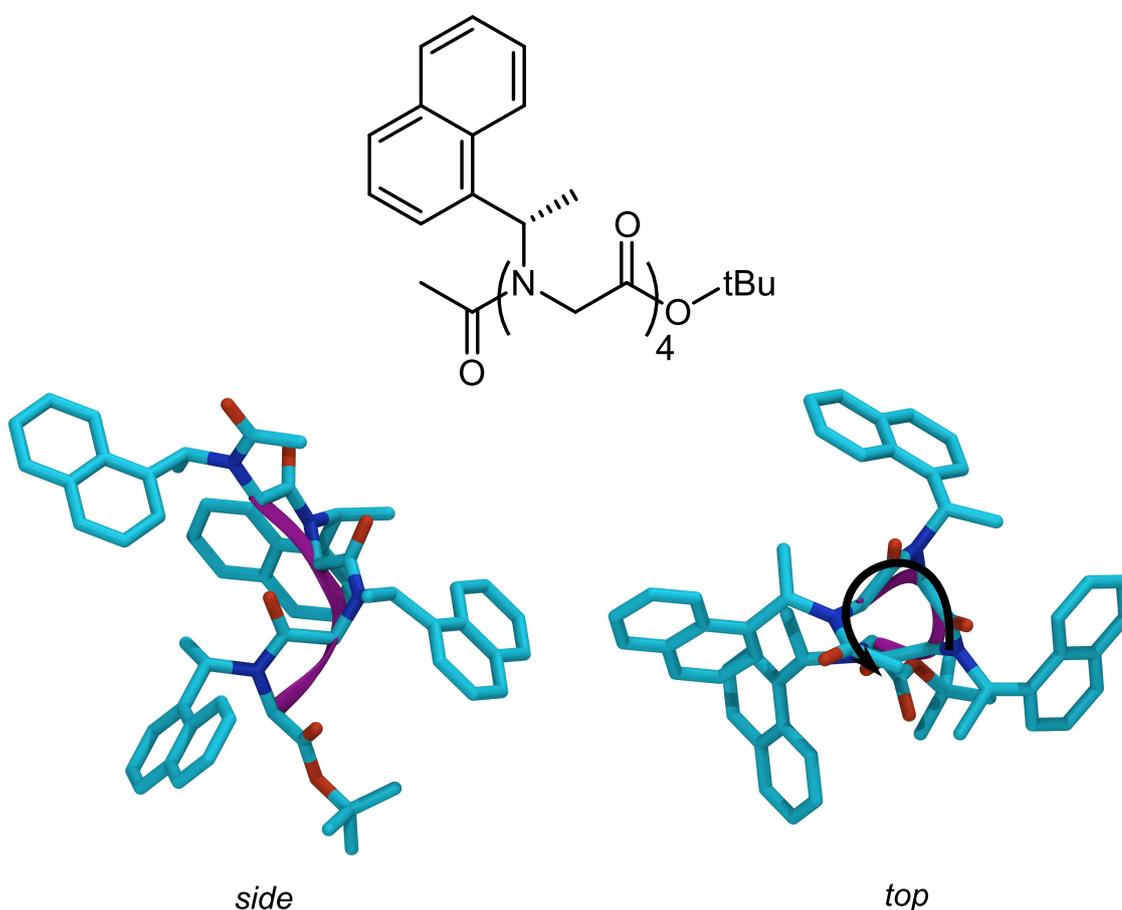
on bulky aromatic  $\alpha$ -chiral, namely  $N_{spe}$  and  $N_{rpe}$ , side chains. By combining CD and NMR spectroscopy experiments, it came out that sequences as short as five repeating units are already sufficient to induce the helix in a homo-peptoid, and could even appear with only 3 residues (**Figure 1.11**)<sup>41,42,47</sup>. Interestingly, hetero-peptoids bearing a mix of bulky aromatic  $\alpha$ -chiral and achiral side chains still remain in a stable helical form, even when the chiral content drops as low as down to 50%.<sup>48</sup> Nonetheless, the positioning of the chiral side chains has an important impact on the secondary structure formation.<sup>90</sup> For example, the second position from the N terminus extremity has the most important impact on the structure stabilization, while the positions in the middle of the chain have barely no influence.



**Figure 1.11:** Helical (right-handed) structure formed by a peptoid octamer bearing  $N_{spe}$  side chains as predicted by Armand *et al.* (all-*cis* amide,  $\phi = -70^\circ$ ,  $\psi = 165^\circ$ ) and confirmed by NMR experiments.<sup>41,47</sup>

The  $N_{spe}$  or  $N_{rpe}$  side chains presented above are “textbook” examples of side chains able to fold in an all-*cis* conformation while restricting at the same time restricting the  $\phi/\psi$  dihedrals to an  $\alpha_D$  conformation.<sup>47</sup> However, helical peptoids (as attested by the characteristic CD signature) bearing a mixture of  $N_{rpe}$  and  $N_{me}$  ( $N$ -(2-methoxyethyl)glycine) side chains display a short persistence length of about 5 to 10 Å, which is in the range of the helical pitch.<sup>87</sup> These results agree with NMR data and clearly indicate that these peptoids can sample different conformations and that only a portion of the conformers are fully helical.<sup>47</sup> This behavior is related to

the *cis/trans* isomerism that introduces a large mixing of amide configurations. Indeed, the equilibrium constant between the *cis* and *trans* forms,  $K_{cis/trans}$ , of amide bearing *Nspe* side chain is not that much in favor of the *cis* form ( $K_{cis/trans} = 1.5$  in acetonitrile).<sup>40,47,91</sup> The *cis* amide conformation is only slightly more favorable when using *Nspe* (or *Nrpe*) side chain, because of slightly less steric hindrance and favorable weak local non-covalent interactions such as “bridged”  $n \rightarrow \pi^*$  aromatic (Figure 1.9).<sup>78</sup> Recently, a modification of the *Nspe* side chain into *Ns1npe* (*N*-(*S*)-(1-naphthylethyl)glycine) was found to improve significantly the  $K_{cis/trans}$  ratio up to >19 (Figure 1.12), directly competing with the fully *cis*-inducing *NtBu* side chain, at the cost of reducing its solubility in highly polar solvent such as water.<sup>86</sup> Nonetheless, short oligomers were crystallized successfully, demonstrating that these peptoids fold into similar right-handed helices as formed by *Nspe* peptoids (Figure 1.11).

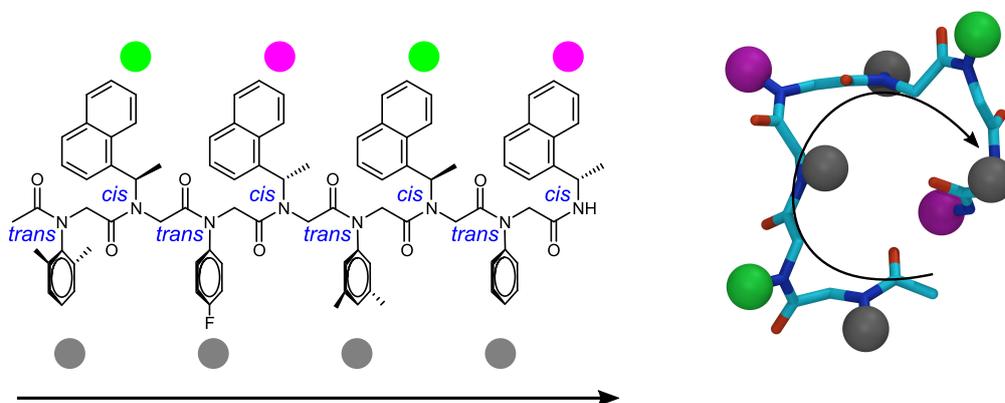


**Figure 1.12:** Crystalline structure of a peptoid tetramer bearing the strongly *cis*-inducing *Ns1npe* side chains.<sup>86</sup> It forms a right-handed structure very similar to the peptoids bearing *Nspe* side chains (curved black arrow in the top view representation).

While peptoids in a PPI helical-like conformation are often described in the literature, PPII-like structures are lacking. It was suggested that such helix could be formed by side chains favoring almost exclusively *trans* amide conformation, such

as aryl side chains.<sup>38</sup> However, to promote a helix with a well-defined screw sense, a chiral derivative of such side chains should be envisaged.

Thanks to the identification of representative side chains to control the *cis/trans* isomerism, new motifs have been designed, such as the square helix. This structure arises from the synergetic control of the  $\omega$  and  $\phi$  dihedrals. The secondary structure is formed by a sequence of alternating *trans-cis* amides, which are controlled by the strongly *trans* inducing *Nph* and the strongly *cis* inducing *Ns1npe* and *Nr1npe* side chains (**Figure 1.13**). As to the  $\phi$  dihedrals, their control is achieved through the chirality of the side chains. The repeating motif is *Nph* – *Nr1npe* – *Nph* – *Ns1npe*, which triggers alternation between positive (*Nr1npe*) and negative (*Ns1npe*)  $\phi$  values. *In fine*, it leads to a wider helix radius compared to the PPI type helix.

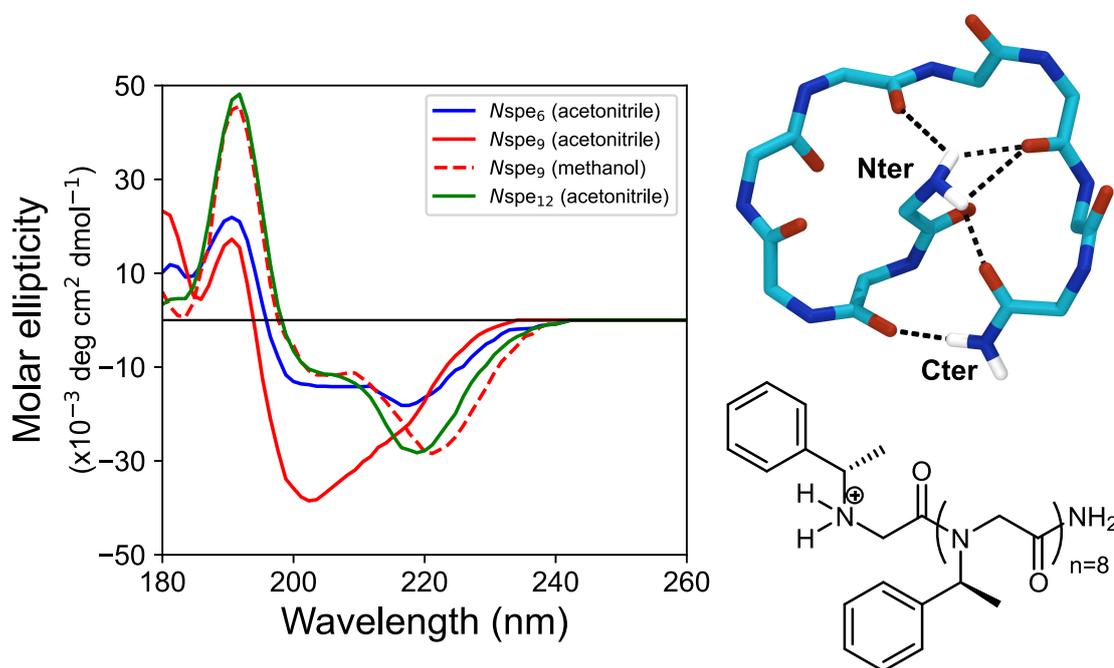


**Figure 1.13:** Primary structure of the peptoid square helix featuring an alternated sequence of *trans* (*Nph* derivatives, gray beads) and *cis* (*Nr1npe*, magenta beads and *Ns1npe* green beads) amide inducing side chains together with the associated tridimensional model structure. Hydrogens are omitted for clarity. Image adapted with permission from Ref. 92 based on their PDB file of compound 5. Copyright (2017) American Chemical Society.

## 1.5.2 Loops

Interestingly, the *Nspe* side chain is also responsible for another type of secondary structure. The analysis of a series of *Nspe* homo-peptoids by CD spectroscopy revealed a different signature only for *Nspe*<sub>9</sub> in aprotic polar solvents, such as acetonitrile, which was attributed to a “threaded loop” conformation (**Figure 1.14**).<sup>93</sup> This secondary structure relies on the formation of intramolecular hydrogen bonds between the protonated terminal amine (N terminus) and the carbonyl oxygens, as well as with the terminal amide (C terminus). Compared to the PPI helix where the amide bonds are all-*cis*, the “threaded loop” is formed by both *cis* and *trans* amide bonds (~ 50% *cis*/50% *trans*). This loop structure can be converted into a helical structure if the solvent is switched to a polar protic one, such as methanol, which will disrupt

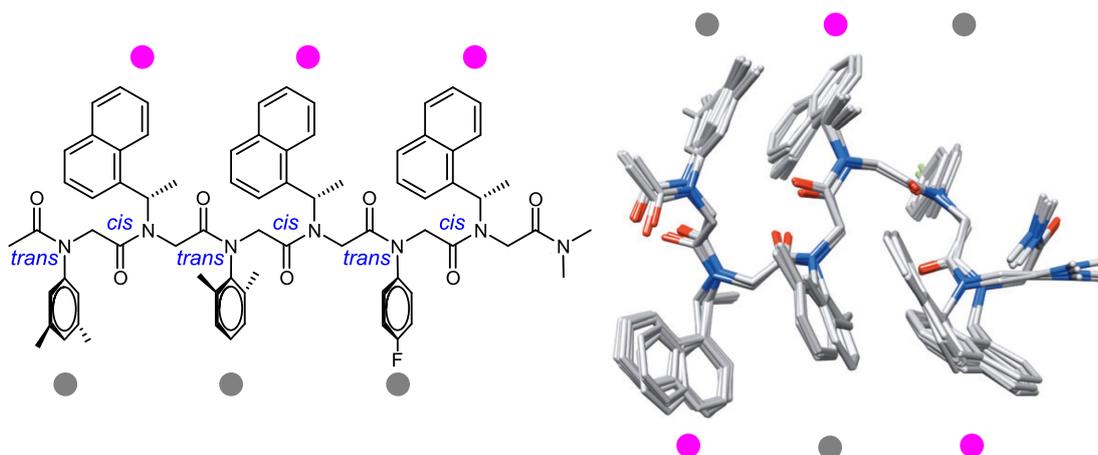
the intramolecular hydrogen bonds. So far, this type of structure was only reported for this particular chain length.



**Figure 1.14:** Circular dichroism signatures of different homo-peptoids bearing *Nspe* side chains. The signature of the nonamer drastically differs in polar protic (dashed red line) and polar aprotic solvent (plain red line, which was attributed to the “threaded loop” conformation represented on the right side). Side chains have been omitted for clarity, as well as hydrogens, except for those involved in intramolecular hydrogen bonds (dashed black lines). Image adapted with permission from Ref. 93 based on their **FIGURE 5**. The 3D structure was built based on the dihedral angle values from their **TABLE 2**. Copyright (2006) American Chemical Society.

### 1.5.3 Ribbons

The strategy to obtain peptoid ribbons is similar to that for square helices.<sup>92</sup> This structure also comprises an alternation of *cis-trans* amide bonds, generated by the alternation of strongly *cis* and *trans* inducing side chains, such as *Ns1npe/Nr1npe* and *Nph* (**Figure 1.15**).<sup>94</sup> Unlike the square-helix in which the chirality of *cis* inducing side chains alternate between (*R*) and (*S*), the side chain chirality remains the same in the ribbon conformation. It prevents the  $\phi$  dihedral to change sign. Again, the chirality of the side chains governs the handedness of the structure; in this case, *Ns1npe* side chains create an overall left-handed spiral configuration with the dihedral combination  $\phi \sim \pm 60^\circ$  and  $\psi \sim \pm 160^\circ$ . The peptoid ribbon is similar to the peptide  $\beta$ -ribbon, although it is not stabilized by intramolecular hydrogen bonds.

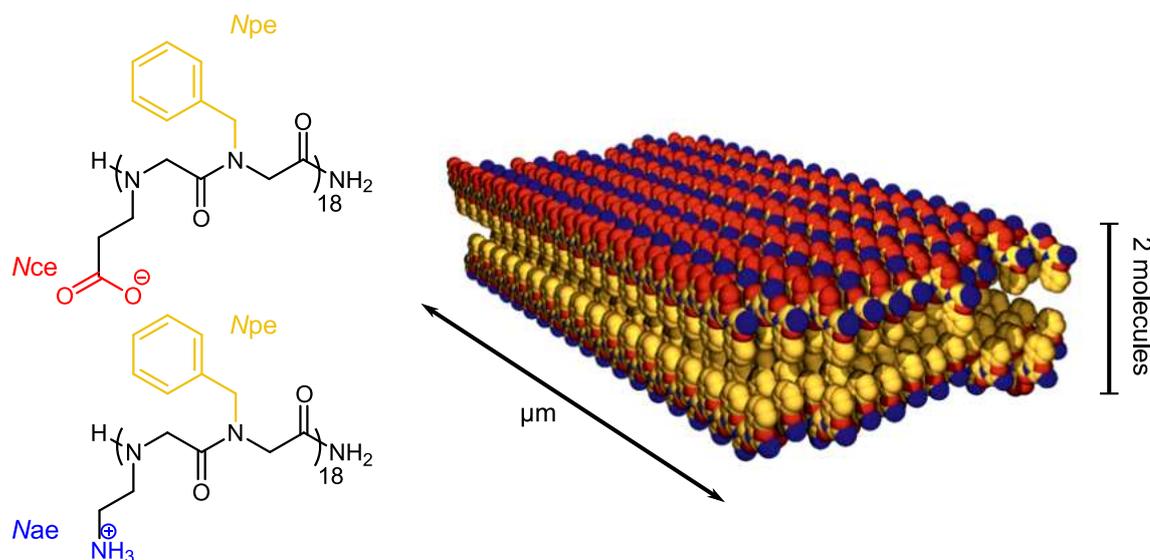


**Figure 1.15:** Peptoid ribbon formed by alternation of *cis* and *trans* amide bonds, as determined by NMR. Hydrogens are omitted for clarity. Image adapted with permission from Ref. 94, Copyright 2013, from John Wiley and Sons.<sup>94</sup>

#### 1.5.4 Higher order structure

Over the past 10 years, new structures have been reported for peptoids, in particular sheets formed by the self-assembly of peptoid chains.<sup>95,96</sup> This architecture resembles the typical lipid bilayers and consists in a free-floating planar assembly which is “two-molecule” thick and whose length can extend over several micrometers. The first peptoid nanosheet was designed by Zuckermann *et al.*<sup>95</sup> The aim was to synthesize pairs of complementary peptoid chains: one whose side chains alternate between aromatic (*Npe*, *N*-(2-phenethyl)glycine) and polar charged (*Nae*, *N*-(2-aminoethyl)glycine) and the other between *Npe* and *Nce* (*N*-(2-carboxyethyl)glycine): (*Nae-Npe*)<sub>18</sub> and (*Nce-Npe*)<sub>18</sub>.<sup>95</sup> The polar side chains are actually in their ionic form, *i.e.*, ammonium form for *Nae* and carboxylate form for *Nce*. When mixed together, these two chains self-assemble in an aqueous buffer, with the apolar side chains facing each other inside the sheet, and the polar charged side chains exposed to the solvent (or to air) (**Figure 1.16**).

Interestingly, the individual peptoid chains composing the nanosheet are completely linear, later referred to “ $\Sigma$ -strand”,<sup>97</sup> which was never observed with a single isolated peptoid chain.<sup>97</sup> This linear motif arises from a particular combination of backbone dihedrals, where the amide bonds are all-*trans* and the values of the couple ( $\phi$ ,  $\psi$ ) alternate between ( $120^\circ$ ,  $-75^\circ$ ) and the opposite values, ( $75^\circ$ ,  $-120^\circ$ ), generating the  $C_{7\beta}$  conformation (**Figure 1.3**). These observations stress once again that the conformational space of peptoids is wide, although not so many individual conformations were discovered up to date. Indeed, new structural motifs could be hidden in other supramolecular assemblies to be discovered in the future, especially in peptoids with alternating side chains (**Figure 1.17**).



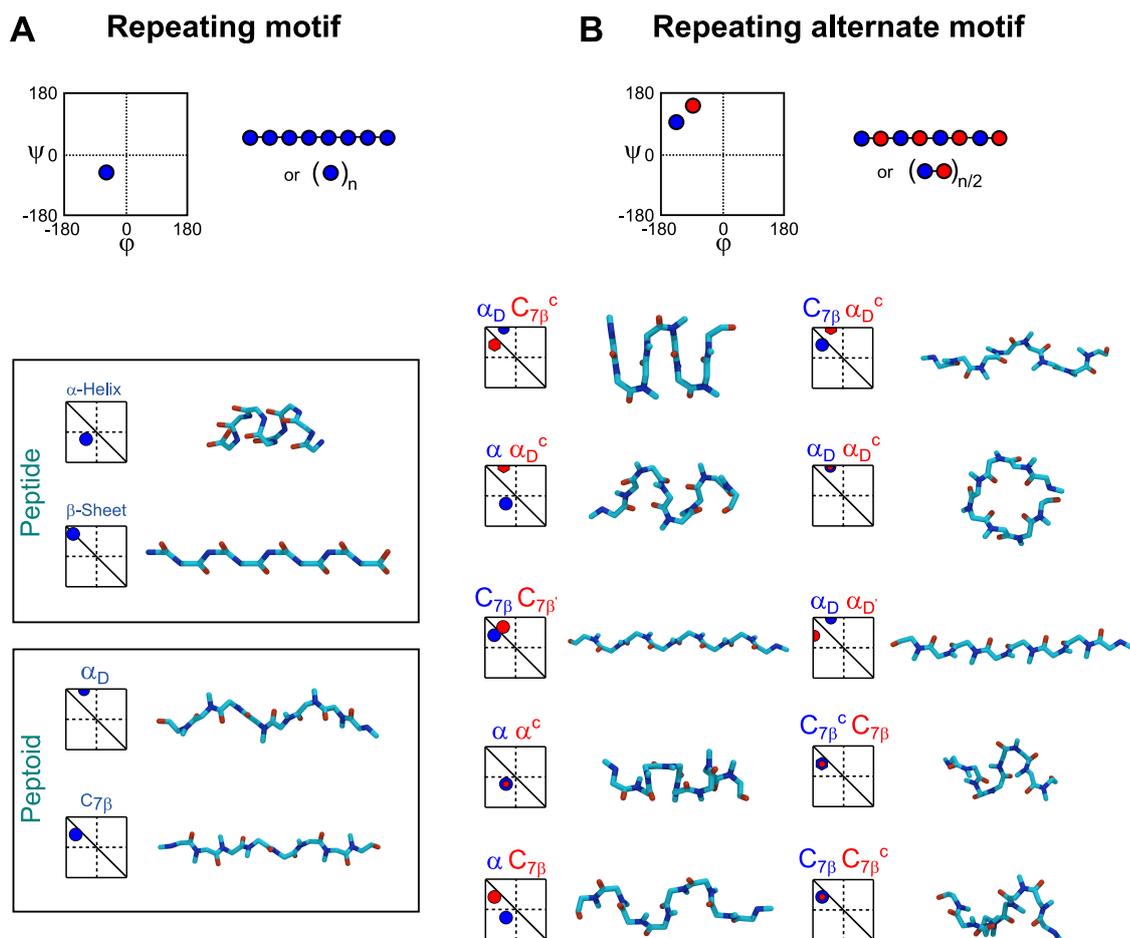
**Figure 1.16:** Primary structure of the original alternating-peptoids,  $(Nce-Npe)_{18}$  and  $(Nae-Npe)_{18}$ , that were designed to form nanosheets.<sup>95</sup> Adapted with permission from Ref. 95. Copyright 2010 Nature Publishing Group.

## 1.6 Applications of peptoids

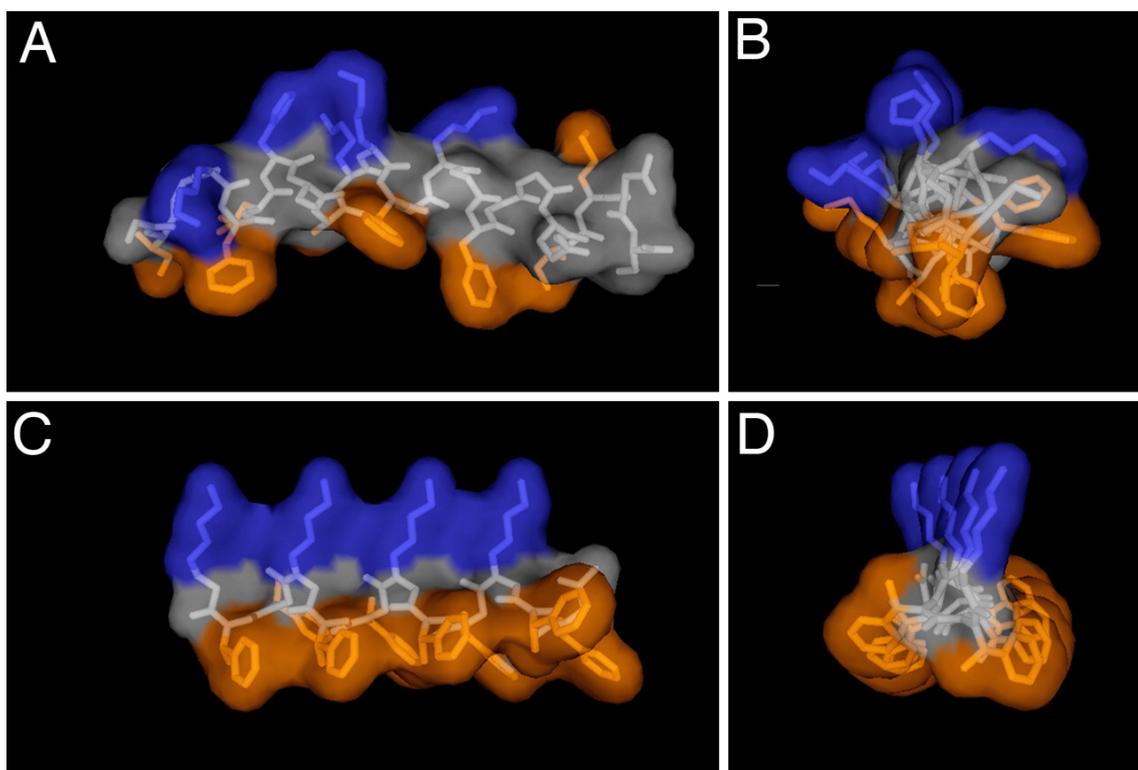
The range of applications in which peptoids could be used reflects the large versatility of primary structures. Indeed, these peptidomimetics are increasingly developed for biomedical applications but also as new materials as well as in nanotechnology. An non-exhaustive overview of these different fields is given hereafter.

### 1.6.1 Biomedical applications

Although originally designed to compete with peptides as therapeutics thanks to the possibility to create a large library of compounds, peptoids were rapidly spotted for many other applications. One of them is to use peptoids as substitutes of antimicrobial peptides (AMPs).<sup>84,98</sup> AMPs enter efficiently into the cytoplasmic membrane of bacteria and in fine cause their death, but suffer from short half-lives and poor pharmacokinetics.<sup>99</sup> These peptides must have an amphipathic structure in which polar and apolar residues are separated, as it is the case for magainin peptides that adopt an  $\alpha$ -helical conformation (**Figure 1.18A-B**). Peptoids that mimic the magainins were successfully designed and provided similar results as their peptide counterpart, except that a much lower toxicity was observed (**Figure 1.18C-D**). Moreover, peptoids with specific structures were clearly more efficient than others.



**Figure 1.17:** (A) Example of a Ramachandran plot based on a single repeating motif, as found in peptides ( $\alpha$ -helix,  $\beta$ -sheet) or in peptoids ( $\alpha_D$ ,  $C_{7\beta}$  in their all-*trans* form).<sup>97</sup> (B) Ramachandran-like plots of alternating peptoid sequences that can be envisaged for the design of new structural motifs.<sup>97</sup> Adapted with permission from Ref. 97. Copyright 2015 Nature Publishing Group, a division of Macmillan Publishers Limited.

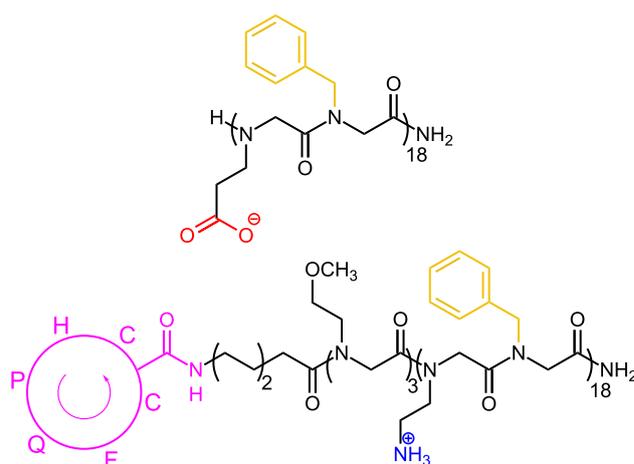


**Figure 1.18:** (A) and (B) are the side and top views of the magainin-2, as determined by NMR. (C) and (D) are the side and top views of the peptoid designed to mimic magainin-2. The blue residues are hydrophobic while the orange ones are cationic.<sup>98</sup> Image reprinted with permission from the National Academy of Sciences (U.S.).<sup>98</sup>

## 1.6.2 Materials applications

### Nanosheets

The ability of peptoids to form higher order structures such as nanosheets enable their use in a wide variety of applications ranging from molecular recognition, catalysis, templating effects, and even more.<sup>96</sup> The key asset is once again their ease of chemical tuning that allows introducing basically any kind of functional moieties to fulfill a desired role. Currently, many of these applications are only in their infancy. A proof of concept was reported in the molecular recognition domain.<sup>95</sup> One of the peptoid chain that originally formed a nanosheet,  $(Nae-Npe)_{18}$ , was chemically modified (at the N terminus) to accommodate a peptide sequence (cyclo-[CHPQFC]) which is a receptor of the streptavidin protein (**Figure 1.19**). The addition of this peptide does not prevent the nanosheet formation and provides an effective binding to the target protein. The advantage of the nanosheet is to form an organized layer that display the molecular receptor on its surface, which would not occur with a single peptoid chain as it would collapse into a globular shape.<sup>100</sup>

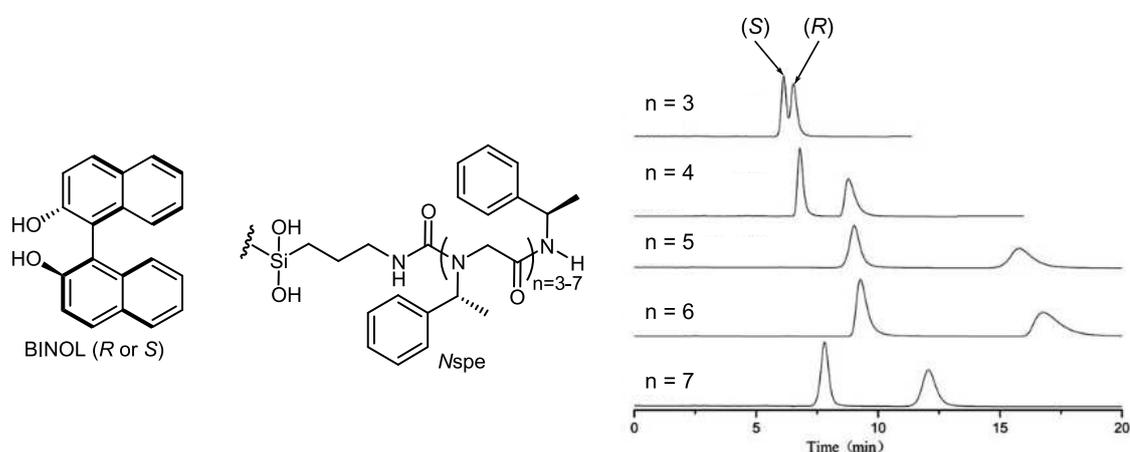


**Figure 1.19:** Original sequence of the  $(Nce-Npe)_{18}$  peptoid and the modified version of the  $(Nae-Npe)_{18}$  peptoid.<sup>95</sup> linker is added between the peptoid chain and the peptide sequence (magenta). The amino acids are represented by their letter code. The cyclization is ensured by a disulfide bond.

### Molecular recognition

Molecular recognition occurs between two (or more) molecules that exhibit a molecular complementarity, which can be pictured as a “*lock and key*”. Such recognition processes are fundamental, especially when chiral molecules are involved such as in living organisms. Pharmaceutical industries are deeply dependent on the chiral recognition, as most newly developed drugs are chiral, whose stereoisomers, including enantiomers can exhibit different behavior (potentially harmful<sup>101</sup>) in biological

environments<sup>102</sup> Therefore, chiral drugs should be either produced enantiomerically pure,<sup>103</sup> or separated after their production.<sup>104,105</sup> This second method is often more suitable, as enantioselective synthesis is difficult to set up. Chromatography, in particular chiral chromatography, appears to be an elegant way to achieve the resolution of racemic mixture.<sup>105</sup> A chiral stationary phase (CSP) replaces the traditional achiral stationary phase, for example, C18 that contains alkyl chains of 18 carbons grafted on silica. Peptide-based CSPs have proven to be effective thanks to their ability to form chiral secondary structures, such as  $\alpha$ -helices, and thus interact preferentially with one of the enantiomers.<sup>106–108</sup> However, the peptide chemistry shows its limitations because “only” 20 natural amino acids are available, compared to peptoids that can afford a huge variety of chemical moieties. In this context, peptoids grafted on amorphous silica beads proved to be very efficient in the resolution of racemic mixture of model compounds, such as [1,1'-binaphthalene]-2,2'-diol (BINOL).<sup>24</sup> The peptoids involved are of course chiral and bear the extensively studied *N*spe side chain (**Figure 1.20**). The resolution efficiency, characterized by the separation factor  $\alpha^a$ , increases from 1.1 up to 2 as the peptoid length increases from 3 to 7 chiral side chains, in consistency with the formation of a chiral secondary structure, most likely the *cis*- $\alpha_D$  helix. However, there are currently no clear evidence about the involvement of a helix nor a clear rationale about the reasons for which one enantiomer interacts preferentially with the CSP.



**Figure 1.20:** Primary structure of the peptoid sequence grafted on the amorphous silica beads and chemical structure of BINOL. Each chromatogram was obtained with an increasing number of chiral side chains, from  $n = 3$  to 7. The first eluted enantiomer is always the (*R*) which interacts less with the peptoids.<sup>24</sup> Reproduced by permission of The Royal Society of Chemistry.

<sup>a</sup>The separation factor is defined as the ratio between the effective elution times of both species:  $\alpha = \frac{t_{(R)} - t_0}{t_{(S)} - t_0}$ ,  $t_0$  being the retention time of non-retained compounds and  $t_{(R)or(S)}$  the elution time of the chiral species.



---

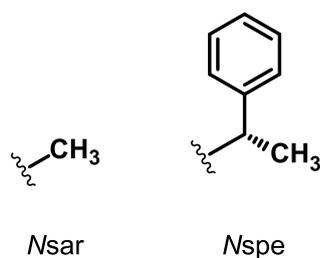
## Objectives of the thesis

---

Despite the relative youth of the field, peptoids are of increasing interest in the scientific community.<sup>99,109–111</sup> Their synthetic ease,<sup>13,112</sup> their large structural diversity through the nature of the side chains and their ability to fold into well-defined conformations make them particularly remarkable compounds that belong to the foldamer family.<sup>41,86,93,94</sup> The foldamer components can be analogously considered as LEGO® bricks: when the shape, connection points and color are known (the primary structure), one can predict and design the architecture into which it folds (secondary structure) under particular conditions for targeted application.<sup>113</sup> Currently, the role of the peptoid components, *i.e.*, the effects of the side chains, is almost exclusively characterized in solution, using NMR and CD spectroscopies (for the latter, provided the peptoids are chiral),<sup>42,47,48,58,114</sup> since crystalline data are barely available for most peptoids.<sup>48,55</sup> Gas phase information can be relevant to identify the intrinsic factors favoring folding (in absence of solvent) and may help the design of new architectures.<sup>74,75,115,116</sup> These gas phase data can be readily available through mass spectrometry, although to date, MS methods are rather used to identify the primary structure of peptoids and hence their sequence. In this regard, we want here to introduce mass spectrometry as a new method to monitor the folding behavior of peptoids, in particular using mass spectrometry coupled to ion mobility spectrometry (IM-MS).<sup>71</sup> Indeed, IM-MS provides structural information about ion conformations through their collision cross section, that depends on their tridimensional shape.

The first part of this thesis will concern the determination of the conformation of gaseous peptoid ions and identify the factors responsible for their folding in gas phase. We will first compare two different side chains: sarcosine and (*S*)-1-phenylethyl (*N*sar and *N*spe) since the latter is known to induce a folding into helical conformation in solution,<sup>90</sup> whereas sarcosine does not.<sup>117</sup> Therefore, we are expecting different behaviors for those two types of peptoid in gas phase. IM-MS characterization will be supported by molecular modeling to provide candidate structures whose theoretical collision cross sections match the experimental values to identify their conformation. We choose to describe the conformational dynamics of peptoids by using molecular dynamics simulations (MD). Proper parametrization of the force field is a mandatory step in any reliable MD work. Therefore, based on the DREIDING force field, we

will proceed to tweaking of the parameters to adequately describe the peptoid potential energy surfaces. The strength of our approach arises from the synergy between the IM-MS method and MD simulations, as both data are obtained in a vacuum-like environment.



**Scheme 2.1:** Chemical structures of the targeted side chains *Nsar* and *Nspe*.

As discussed in the [Introduction](#), peptoids are considered as selector due to their ability to achieve molecular recognition, in particular toward chiral compounds such as binaphthyl derivatives.<sup>24</sup> Enantiorecognition is a crucial topic, especially in the context of the pharmaceutical industry, since drug enantiomers may provoke dramatically different effects in our body.<sup>118,119</sup> Among the available methods to separate enantiomers, we will focus our attention on the chromatography process, that relies on a chiral stationary phase (CSP). In this context, peptoids bearing (*S*)-1-phenylethyl side chains grafted on silica beads were reoriented to separate binaphthyl derivative enantiomers. For the last part of the thesis, we will collaborate with Prof. Luca Muccioli to take advantage of “steered molecular dynamics” simulations to mimic the elution process occurring in a chromatography column.<sup>120</sup> These simulations will help us to get insights into the structural parameters that govern the recognition process. Identifying the factors responsible for the enantioselective properties could be combined to the synthetic ease of peptoids and design new peptoid sequences that can separate a broad range of chemically diverse chiral compounds.

---

# Methodology

---

## 3.1 Synthesis

The peptoids studied in this thesis were synthesized by Dr. Emilie Halin and Perrine Weber. All reactants and solvents are commercially available (VWR chemicals) and are used without any supplementary purification. The solid-phase synthesis protocol was based on the procedure described by Zuckermann *et al.*<sup>13,112,114,121</sup> The same protocol was used for each side chain type (Table 3.1). The syntheses were performed in a double-walled glass vessel equipped with a 3-way stopcock. The vessel is connected to: (i) a nitrogen supply for solution mixing and (ii) a water heating circulator.

Several initial procedures are necessary to prepare the resin to allow the peptoids to grow onto. The glass vessel is heated to 80 °C. 1 g of Rink amide (loading 0.06 mmol, Fmoc protected) resin is added into the vessel. The resin is then swelled using 5 mL of *N*-methyl-2-pyrrolidone (NMP) and submitted to agitation by nitrogen bubbling for 10 minutes. The solution is removed and 5 mL of 20% 4-methylpiperidine in NMP (v/v) is added to remove the Fmoc protecting group. The resin is agitated for another 10 minutes and the solution is drained afterwards. The resin is rinsed using 5 mL of NMP, mixed with the resin for 15 seconds and drained (repeated three times). Then, the peptoid synthesis can be initiated, with the cycle of two steps described in the Introduction (Figure 1.2), namely (i) the bromoacetylation; and (ii) the nucleophilic displacement. During the bromoacetylation step (5 min), 2.5 mL of bromoacetic acid (5 mmol) and *N,N'*-diisopropylcarbodiimide (1 equivalent, 5 mmol) in NMP are added to the resin. In the second step, 5 mL of 1.5 M of the primary amine in NMP is added (Table 3.1), and the solution is mixed for 5 to 10 minutes, then drained and rinsed three times using 5 mL of NMP. Each cycle of these two steps introduces a peptoid monomer unit inside the growing chain. Therefore, both steps are repeated until the desired chain length is reached. After the final nucleophilic displacement, the resin is rinsed three times with NMP then three times with dichloromethane. The resin is finally isolated.

A small amount of resin is placed in a glass vial and 200  $\mu$ L of 95:5 (v/v) trifluoroacetic acid:water solution is added and then shaken for 10 minutes. 10  $\mu$ L of this solution is taken off and mixed with 190  $\mu$ L of acetonitrile:water 50:50 (v/v) solution.

Finally, the sample is diluted using the same solvent mixture to reach the appropriate concentration for MS analysis ( $\sim \mu\text{g mL}^{-1}$ ). The peptoids are prepared without further purification because they are only analyzed by mass spectrometry which isolates the targeted ions in the gas phase. The primary structure of the peptoids is validated by measuring their mass to charge ratio that is compared to the expected one.

**Table 3.1:** Primary amines used in this work for the synthesis of peptoids.

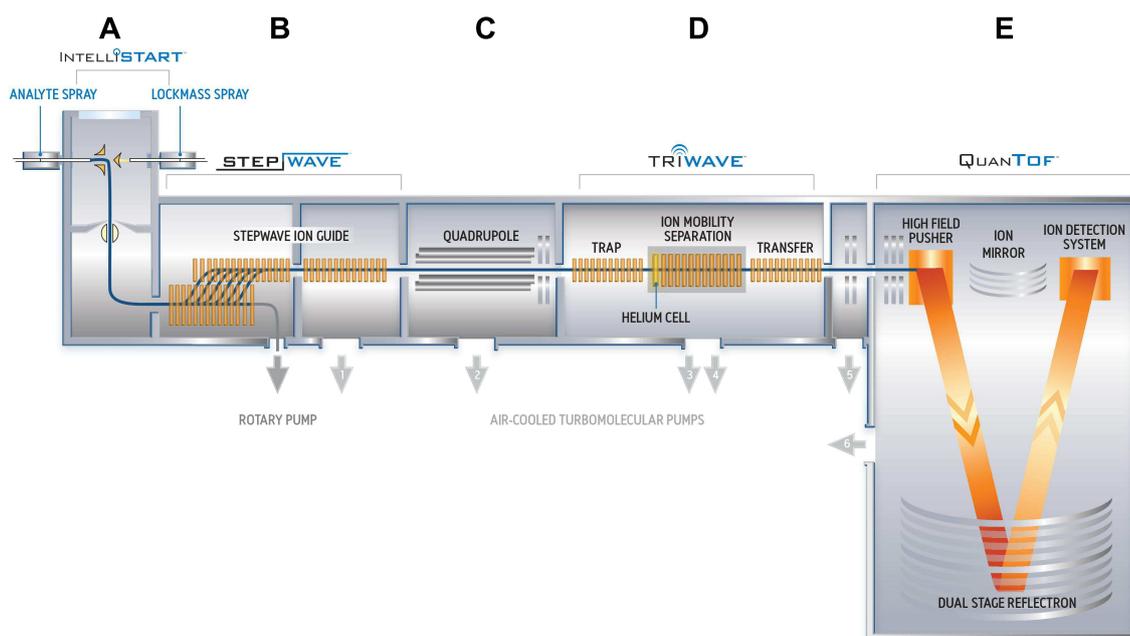
Primary amine	Abbreviation
Methylamine <sup>a</sup>	<i>Nsar</i>
( <i>S</i> )-phenylethylamine	<i>Nspe</i>
( <i>S</i> )- <i>N</i> -(1-carboxy-2-phenylethyl)glycine <sup>b</sup>	<i>Nscp</i>

<sup>a</sup>The temperature is lowered to 20 °C, and the concentration is decreased to 1 M.

<sup>b</sup>The carboxylic acid moiety is protected by esterification. At 0 °C, thionyl chloride (11 mL, 0.15 mol) is slowly added to methanol (150 mL). (*S*)-*N*-(1-carboxy-2-phenylethyl)glycine is then added at room temperature and the solution is stirred for 48 h. After the synthesis, the carboxylic moieties of chains shorter than 9 residues are deprotected by hydrolysis with NaOH 5 M (3 mL) and methanol (4 mL) at 55 °C for 5 h and >10 h for the longer chains.

## 3.2 Ion mobility mass spectrometry

The mass spectrometer used for the experiments is the Synapt G2-Si (Waters, UK). This instrument is equipped with the following elements: an Electrospray ionization source, a quadrupole, a mobility cell which employs the “travelling wave” technology and an orthogonal acceleration Time of Flight (oa-ToF) mass analyzer.



**Figure 3.1:** Components of the Waters Synapt G2-Si mass spectrometer.

The Electrospray ionization source abbreviated ESI, continuously produces ions (**Figure 3.1A**).<sup>122</sup> The solution containing the analyte(s) passes through a conductive capillary on which is applied a voltage (typically 3.1 kV) to induce an electric field. In every experiment, we applied a positive potential which causes the separation of the negative (attracted toward the capillary wall and neutralized) and the positive charges (repelled by the capillary wall), that eventually increases the positive charge density in the solution throughout its infusion in the capillary. The accumulation of positive charges and the electric field deform the jet at the tip of the capillary, which in turn produces a “Taylor cone” constituted by solution droplets. Heated N<sub>2</sub>, at typically 120 °C, is used to evaporate the solvent molecules from the droplets. The droplets thus shrink due to the evaporation until it reaches the so-called “Rayleigh limit”  $q = 8\pi r \sqrt{\varepsilon_0 \gamma r}$ , where  $q$  is the droplet charge,  $r$  is the droplet radius,  $\gamma$  is the surface tension and  $\varepsilon_0$  is the vacuum relative permittivity. This equation describes the maximum charges that a droplet can carry before the electrostatic repulsion overcomes the surface tension.<sup>122</sup> When this limit is reached, the initial droplet explodes into smaller droplets that will eventually be subjected to the same phenomenon until it only remains the ionized analyte in the gas phase. After their production, the ions are transferred into the mass spectrometer via the so-called sample cone and conducted by the Stepwave, that is an ion guide (**Figure 3.1B**) to the quadrupole analyzer (**Figure 3.1C**), which can be used either as a mass selector or as an ion guide to transfer ions to the next part.

The next part of the instrument is related to the ion mobility experiments and consists in the original TriWave setup. The complete working principle of the TriWave, including the large set of available experiments, is described elsewhere.<sup>123,124</sup> Briefly, the TriWave consists in three consecutive Travelling Wave Ion Guides (TWIGs) called the “trap cell”, the “ion mobility cell” and the “transfer cell” (**Figure 3.1D**). TWIGs are made of a series of ring-shaped electrodes on which are applied direct and alternative current (DC and AC) whose combination radially confines the ions. The AC contribution is mostly responsible for the radial focusing of the ions while the DC is applied by pulse to push the ions packets. Globally, the applied voltage combination resembles a wave that is defined as “*Traveling Wave*”.

TWIGs are filled with different gasses depending on their functions (generally Ar or N<sub>2</sub>). The first cell is the trap cell filled with Ar. It is used to trap the ions prior to their transfer into the mobility cell where packets of ions are separated. Indeed, packets of ions are injected into the mobility cell which is filled with N<sub>2</sub> gas at a higher pressure. As the ions travel through the cell, they undergo collisions with the

gas which will slow them down and will be separated according to their mass, charge and shape (and also the experimental conditions, such as temperature, pressure and of course, the nature of the buffer gas, more details below).<sup>125</sup> After the mobility separation, the ions enter the “transfer cell” which will lead them to the QuanToF mass analyzer and detector (**Figure 3.1E**).

The ion mobility  $K$  defines the ion displacement against a gas under the action of a homogeneous electric field  $\bar{E}$ . This physical parameter can be derived from the time spent by the ions to fly through the cell, the so called drift time ( $t_D$ ):

$$K \propto \frac{1}{t_D} \quad (3.1)$$

The ion mobility must ultimately be also dependent on the collisions occurring between the ion and the buffer gas. Using the collision theory, Mason and Schamp were able to derive the expression of the mobility and obtain structural information about the ion (with a given buffer gas),<sup>126</sup> namely the *collision cross section*  $\Omega$ . This physical quantity represents the size and shape of the ion that is determined by the ion - buffer gas interaction. The collision cross section can thus aid to the identification of gas phase ion conformation, provided it can be compared to theoretical values obtained on candidate geometries (see 3.3.4).

Unlike instruments that apply a homogeneous electric field affording a linear relationship between the drift time and the collision cross section, the electric field is not constant in TWIMs instruments.<sup>127</sup> The Mason-Schamp equation is thus adapted such as:<sup>123,124</sup>

$$\Omega = A t_D^B \quad (3.2)$$

where  $t_D$  is the time spent by an ion inside the mobility cell,  $A$  and  $B$  are empirical parameters determined through a calibration using reference compounds of known collision cross sections. These parameters depend on the experimental conditions, that is the temperature, pressure, nature of the gas, etc. For a given mass-to-charge ratio, ions whose shape is more extended will spend more time inside the mobility cell than compact ions and will therefore be characterized by a larger collision cross section.

### Collision cross section calibration

The  $A$  and  $B$  parameters from Eq. 3.2 are determined by associating the measured drift time of reference compounds (also called *calibrants*) to their (absolute)  $\Omega$  that

was measured in a Drift Tube instrument (using helium as drift gas). The calibrants used in this work are polymer compounds (poly(ethylene glycol) - PEG,  $\alpha$ -methyl,  $\omega$ -hydroxy poly(lactide) - PLA) that present many advantages,<sup>124</sup> such as the robustness of their gas-phase structure,<sup>128,129</sup> their intrinsic dispersity in terms of size (chain length) and charge states.<sup>124</sup> The procedure works as follows:

Firstly, the drift time of the calibrants are recorded (charge states 1+, 2+) and are then corrected ( $t'_D$ ) to discard the “dead time” arising from the crossing of the transfer cell and the ToF. Since the drift gas in our TWIM instrument is different (N<sub>2</sub> vs. He), the reduced mass of the interacting system is consequently different and must be corrected accordingly, as well as to remove the charge state dependency:

$$\Omega'_{ref} = \frac{\Omega_{ref}}{z} \sqrt{\mu} = A t'_D{}^B \quad (3.3)$$

where  $z$  is the charge and  $\mu$  is the reduced mass. The parameters A and B are obtained by transforming the equation into a linear relationship using the natural logarithm:

$$\log \Omega'_{ref} = \log A + B \log t'_D \quad (3.4)$$

The data (corrected drift times vs.  $\Omega'_{ref}$ ) are fitted using Equation 3.4 from which A and B are obtained. The collision cross section of the analytes are determined by introducing their corrected drift times and the A and B parameters in Equation 3.5:<sup>a</sup>

$$\Omega_{analyte} = A \frac{z}{\sqrt{\mu}} t'_{D,analyte} \quad (3.5)$$

The following parameters were used in the experiments, except further annotations in the main text. The ionization source parameters were: solution flow, 5  $\mu\text{Lmin}^{-1}$ ; capillary voltage, +3.1 kV; sampling cone, 30 V; source offset, 50-80 V; source temperature, 80 °C and desolvation temperature, 120 °C. Dry nitrogen was used as the ESI gas with a flow rate 500 L h<sup>-1</sup> for the desolvation gas. The IM-MS parameters were: Wave Height 40 V; Wave Velocity 500 to 600 m s<sup>-1</sup>; Mass to charge range  $m/z$  50 to 4000, Argon Trap flow 2 mL min<sup>-1</sup>; N<sub>2</sub> IMS flow 70 mL min<sup>-1</sup>; Helium Cell Gas flow 180 mL min<sup>-1</sup>; Trap bias 40 V. Most IM-MS analyses were performed together with Dr. E. Halin and P. Weber.

<sup>a</sup>IM measurements do not provide a single  $t_D$  value for each  $m/z$  but rather an *arrival time distribution* (ATD). In the calibration procedure, the arrival time distributions of every ions can be approximated by a Gaussian function (monomodal), from which the apex is extracted and used as the value of  $t_D$ . If ATDs are not monomodal, a deconvolution into multiple Gaussian functions is applied.

### 3.3 Molecular modeling

In this work, both classical and quantum approaches are employed to describe the peptoid structures, although the quantum approach only serves as the benchmark for the classical approach. In this section, we will not detail the complete computational methods but rather touch on the concepts of both approaches, starting with the quantum-mechanical approach.

#### 3.3.1 Quantum mechanics

The complete description of the methods is well documented in many textbooks (see 130,131). Before diving into the computational model, let us remind that one of the primary goals of quantum mechanics is to find the solutions of the time-independent Schrödinger equation, most likely within the *Born-Oppenheimer* approximation for a  $N_e$ -electron molecular system: <sup>132</sup>

$$H_{el}\Psi_{el}(r_i) = E\Psi_{el}(r_i) \quad (3.6)$$

where  $H_{el}$  is the electronic *Hamiltonian* of a molecular system,  $\Psi_{el}(r_i)$  are the wavefunctions that depend on the electronic coordinates  $r_i$ . The electronic Hamiltonian is expressed as:

$$H_{el} = -\frac{1}{2} \sum_{i=1}^{N_e} \nabla_i^2 - \sum_i^{N_e} \sum_I^{N_N} \frac{Z_I}{r_{iI}} + \sum_{i \neq j}^{N_e} \frac{1}{r_{ij}} = T_e + V_{eN} + V_{ee} \quad (3.7)$$

The first term,  $T_e$ , describes the kinetic energy of the electrons, the second,  $V_{eN}$ , the attractive potential between electrons  $i$  and nuclei  $I$ , and the last term,  $V_{ee}$ , accounts for the repulsive potential between electrons.

The repulsion between electrons is a difficult problem to solve as the motion of each electron depends on the position of the other electrons. At the atomic level, an elegant manner to deal with this problem is the *Hartree-Fock* method, that treats the interaction between electrons in an averaged manner. <sup>133</sup> Each electron will interact with the average electrostatic field created by the other electrons. This approach relies on the *variation principle* which states that only the exact wavefunction can lead to the exact energy and that any approximate trial wavefunction injected in the Schrödinger equation will lead to a higher energy than the expected one:

$$E_{app} = \frac{\langle \Psi_{trial} | H | \Psi_{trial} \rangle}{\langle \Psi_{trial} | \Psi_{trial} \rangle} > E_{exact} \quad (3.8)$$

By using this principle and describing the wavefunction as a *Slater determinant*<sup>134</sup> that promotes to a anti-symmetric wavefunction total and satisfies the *Pauli principle*,<sup>135</sup> we obtain the Hartree-Fock operator:

$$f_i = h_i + \sum_{m'}^{occ. MO} \{2J_{m'}(r_i) - K_{m'}(r_i)\} \quad (3.9)$$

The first term of Eq. 3.9 is a one-electron Hamiltonian that accounts for the kinetic energy and the attractive potential between the electron and nuclei. The last two terms are the Coulomb operator  $J_m$  and the exchange operator  $K_m$ . The Coulomb operator describes the repulsion between electron  $i$  present in orbital  $m'$  with the averaged spatial distribution of electron  $j$ , while the exchange operator prevents two electrons of same spin to lie at the same spatial position and is purely related to quantum effects.

Because the orbitals sought for also define the Coulomb and exchange operators, HF equations must be solved in an iterative manner and the iteration stops when the solutions are self-consistent to a pre-defined criterion (*Self Consistent Field method*).

The Hartree-Fock approach can be further extended to molecular systems by describing the molecular orbitals as a **L**inear **C**ombination of **A**tomical **O**rbitals (LCAO approximation):<sup>136,137</sup>

$$\phi_m = \sum_{o=1}^{N_0} c_{om} \chi_o \quad (3.10)$$

where  $C_{om}$  are coefficients that balance the weight of the  $o$ -th atomic orbital  $\chi_o$  in the  $m$ -th molecular orbital  $\phi_m$ . The introduction of these molecular orbitals in the Hartree-Fock method leads to the *Roothaan-Hall* equations that can be represented as a matrix:

$$F c = \varepsilon S c \quad (3.11)$$

where  $c$  is an  $N_0 \times N_0$  matrix that contains the  $c_{om}$  coefficients and  $\varepsilon$  is a  $N_0 \times N_0$  diagonal matrix made of the orbitals energies  $\varepsilon_m$ .  $S$  is the overlap matrix between the atomic orbitals ( $S_{o'o} = \int \chi_{o'}^*(r_i) \chi_o(r_i) dr_i$ ) and  $F$  is the Fock operator. The Roothaan-Hall equations also require to be solved iteratively to obtain the best LCAO coefficients.

Although the Hartree-Fock and Roothaan-Hall method introduce a correction to the energy due to exchange effects, these methods do not consider the electronic correlation effects, *i.e.*, the mutual influence of the electrons, whose account can further

improve the computational results. In this work, we use the Møller-Plesset approach (at the second order, MP2), which applies a correction to the Hartree-Fock/Roothan-Hall energy in a perturbative way to account for the electronic correlation.<sup>138</sup>

### 3.3.2 Molecular mechanics

While quantum-mechanical calculations provide highly accurate data about the geometry and electronic properties of molecular systems, these calculations are very computationally expensive as the system size increases. In this study, we consider growing peptoid chains that can be constituted by up to 400 atoms ( $\sim 15$  residues). Each residue is characterized by three main dihedrals,  $\omega$ ,  $\phi$  and  $\psi$ ,<sup>35</sup> each of which should be sampled, thus making the task quickly prohibitive. Molecular mechanics (MM) appears as an appropriate alternative to describe such systems.

Unlike quantum mechanics, molecular mechanics is based on classical physics and therefore does not account explicitly for the electrons. Atoms are treated as charged beads (to implicitly account for the electrons) and are linked by springs to form molecules.<sup>79</sup> The energy of the system is determined by a *force field*, which is a mathematical expression that relates the coordinates of the system to the corresponding potential energy. A force field contains a sum of different terms to describe bonded interactions, typically referred to as bond stretching, angle bending and dihedrals, as well as non-bonded interactions to account for the electrostatic and van der Waals interactions. The force field parameters can be derived from experiments or from high-accuracy quantum-mechanical calculations carried out on (small) model systems to reproduce several properties (densities, vaporization enthalpies, electrostatic potentials, etc.).<sup>139-143</sup> We will describe the different terms constituting a force field, in particular for the DREIDING force field that was selected for this work.<sup>139</sup> The mathematical expressions are those implemented inside the software Materials Studio 18.0 used during this work.<sup>144</sup>

#### Bond stretching

The bonding energy is described by a harmonic potential, based on Hooke's law:

$$E_{bond} = 1/2 k_0(R - R_0)^2 \quad (3.12)$$

where  $k_0$  is the force constant,  $R$  is the distance between the bonded atoms and  $R_0$  is the equilibrium distance.

### Angle bending

The angle bending is also expressed in the form of a harmonic potential:

$$E_{angle} = 1/2 k_0(\theta - \theta_0)^2 \quad (3.13)$$

where  $k_0$  is the force constant,  $\theta$  is the angle formed by three consecutively bonded atoms and  $\theta_0$  is the equilibrium angle between these three atoms.

### Dihedral

Dihedrals are formed by four consecutively bonded atoms. In DREIDING, the torsional potentials are described by a sum of cosine functions:

$$E_{dihedral} = \sum_{n=1}^6 1/2 B_n [1 - d \cos(n \phi)] \quad (3.14)$$

where  $B_n$  is the rotation barrier,  $\phi$  is the dihedral angle formed by the four atoms,  $d$  is the phase factor, that can take two values (-1 or 1) and  $n$  is the periodicity, which typically can take up to 6 values (only integers). Unlike bond stretching and angle bending that are considered as “hard” bonded terms, since they slightly vary around their equilibrium values, dihedral angles are considered as “soft” bonded terms. Indeed, the energy change to trigger conformational changes, *e.g.*, between the *cis* and *trans* amide conformation, is much lower than the required energy to deviate from stretching and bending equilibrium. The potential energy surface is thus highly dependent on the dihedrals (and the non-bonded interactions) that govern the barriers between the local minima. The DREIDING force field was parametrized to describe a wide range of compounds. Nevertheless, due to the importance of the dihedrals in the description of the conformational space, care must be taken when using a generic force field. The Chapter 4 is devoted to the validation and adaptation of the DREIDING force field to describe peptoids.

### Improper dihedral

The improper dihedral term also involves four atoms, but not consecutively bonded. It involves one central atom bonded to three others and addresses the deviation from planarity of the central atoms from the plane formed by the three others. In DREIDING, only chemical elements with  $sp^2$  hybridization have an improper dihedral term.

$$E_{improper} = \sum_{n=1}^3 k_0(1 - \cos(\omega_n)) \quad (3.15)$$

where  $k_0$  is the force constant and  $\omega_n$  is the angle formed between one (out of three) out-of-plane atom and the plane formed by the remaining 3 atoms (the central atom is included). The total improper dihedral energy is thus the sum of 3 terms, one for each out-of-plane atom.

### van der Waals interactions

The van der Waals (vdW) interactions between all pairs of atoms  $i$  and  $j$  are described by using a 12-6 Lennard-Jones potential:

$$E_{vdW} = D_0 \left[ \left( \frac{R}{R_0} \right)^{-12} - \left( \frac{R}{R_0} \right)^{-6} \right] \quad (3.16)$$

where  $D_0$  is the well depth and  $R_0$  is the equilibrium distance. The first term of the equation describes the repulsive interactions (at short distances) while the second describes the attractive interactions (at long distances). Each atom type possesses their own  $D_0$  and  $R_0$  parameters, that are combined using the Lorentz-Berthelot mixing rules in case of heteronuclear atom pairs:

$$\begin{aligned} R_0^{ij} &= 1/2 (R_0^i + R_0^j) \\ D_0^{ij} &= \sqrt{D_0^i D_0^j} \end{aligned} \quad (3.17)$$

The DREIDING vdW parameters were derived from a large data base of compounds for which experimental data were available (crystal structure and sublimation energies). However, several studies point out that the parameters of hydrogen atoms may not be suitable.<sup>145,146</sup> The investigation of reliable hydrogen parameters will also be a subject of the first chapter of the Results and Discussion.

### Electrostatic interactions

The electrostatic interactions are described by a Coulomb potential:<sup>b</sup>

$$E_{elect.} = 322.0637 \frac{Q_i Q_j}{\epsilon R_{ij}} \quad (3.18)$$

where  $Q_i$  and  $Q_j$  are the partial charges corresponding to a pair of atoms separated by a distance  $R_{ij}$  in an environment of relative permittivity  $\epsilon_r$ .<sup>c</sup> In DREIDING, it is set to 1 by default. The atomic charges are attributed by using the Gasteiger method,

<sup>b</sup>By default, the DREIDING force field implemented in Materials Studio package (6.0 and 18.0) uses a “shielded” version of this potential:  $E_{elect.} = 322.0637 \frac{Q_i Q_j}{\epsilon_r R_{ij}^2}$ . In this work, we switched to the pure Coulomb version of Equation 3.18.

<sup>c</sup>The relative permittivity is expressed as  $\epsilon_r = \epsilon/\epsilon_0$ , where  $\epsilon$  is the permittivity of the environment and  $\epsilon_0$  is the permittivity of vacuum. In the present case,  $\epsilon = \epsilon_0$ .

that relies on the difference of electronegativity between bonded atoms and their surroundings.<sup>147</sup>

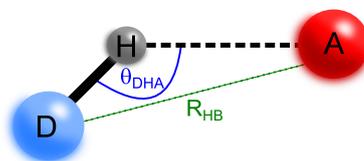
In DREIDING, the van der Waals and electrostatic interactions between pairs of atoms connected by two pairs of bonds are not computed (*i.e.*, if atom A is bonded to atom B and atom B is bonded to atom C, then the atom pairs A-B, B-C, A-C would be excluded).

### Hydrogen bonds

The DREIDING force field is quite unique, as it explicitly describes hydrogen bond interactions, unlike many other force fields:

$$E_{H.bond} = D_{hb} \left[ 5 \left( \frac{R_{hb}}{R_{DA}} \right)^{12} - 6 \left( \frac{R_{hb}}{R_{DA}} \right)^{10} \right] \cos(\theta_{DHA})^4 \quad (3.19)$$

where  $D_{hb}$  is the well depth of the hydrogen bond,  $R_{hb}$  is the equilibrium distance between the donor and acceptor atoms,  $R_{DA}$  is the actual distance between them and  $\theta_{DHA}$  is the angle between the donor - hydrogen - acceptor. The preferential orientations of the hydrogen bonds are depicted by a cosine term. This term is particularly relevant for the study of peptoids since some conformations, such as the “*threaded loop*”, rely on the formation of intra-molecular hydrogen bonds.



**Figure 3.2:** Geometry of a hydrogen bond with the Equation 3.19 parameters  $\theta_{DHA}$  and  $R_{HB}$  highlighted.

Using these equations, one can access the energy of a molecule for a given set of atomic coordinates on its potential energy surface (PES). The energy and thus the “*conformation*” of the molecule can be refined by using local algorithms that will minimize the energy by adjusting the coordinates. In this study, we used the “*Conjugate Gradient*” algorithm. To do so, the energy minimization will move towards the nearest local minimum on the PES but will not allow the crossing of energy barriers. Since the energy determines the accessible conformational space of the molecule, it is desirable to explore the PES to sample the different conformers.

### 3.3.3 Molecular dynamics

Molecular dynamics (MD) is a method to explore the PES of a system and that give access to time-dependent properties. From initial coordinates, successive configurations of the system are generated by solving Newton's law of motion in the Lagrange formalism (Eq. 3.20):<sup>79</sup>

$$F_i = \frac{\delta}{\delta r_i} L = m_i a_i \quad (3.20)$$

where  $F_i$  is the force exerted on an atom,  $\dot{r}_i$  is the atom velocity and  $L$  is the Lagrangian<sup>d</sup>.

The trajectory, *i.e.*, the variation of the atomic positions over time on the PES, is obtained by numerically solving the equations of motion; the potential energy (and the forces) is defined by the same force field as described in the previous section. The time-dependence is obtained by introducing kinetic energy to the system (through a temperature distribution applied on the atoms) that will make the atoms move. Then, the equations of motion are solved for each degree of freedom. In this work, we used the *Verlet velocity algorithm* to integrate the Newton's equations. The basic principle is that, if knowing the positions and velocities at a given instant  $t$ , one can obtain the positions (3.22) and velocities (3.23) at the instant  $t + \delta t$ . The equations are thus solved step-by-step and depend on the choice of  $\delta t$ , often referred to as the *time step* which is typically 1 fs. The Verlet algorithm works as follows:

The initial atomic coordinates  $r(t)$  are known and the velocities are initialized using a *Maxwell-Boltzmann* distribution:

$$p(v_i)dv_i = \left(\frac{m_i}{2\pi k_B T}\right)^{3/2} e^{-\frac{m_i v_i^2}{2k_B T}} 4\pi v_i^2 dv_i \quad (3.21)$$

where  $p(v_i)$  corresponds to the probability for an atom  $i$  of mass  $m_i$  to have an initial velocity  $v_i$ . Then the forces deduced from the force field potential are evaluated and the new atomic positions are updated at a new time increased by  $\delta t$ :

$$r(t + \delta t) = r(t) + v(t)\delta t + \frac{F(t)}{2m}\delta t^2 \quad (3.22)$$

The forces are then updated at  $t + \delta t$  for the new atomic positions  $r(t + \delta t)$  and are used to compute the new atomic velocities:

$$v(t + \delta t) = v(t) + \frac{1}{2} \frac{[F(t) + F(t + \delta t)]}{2m} \delta t \quad (3.23)$$

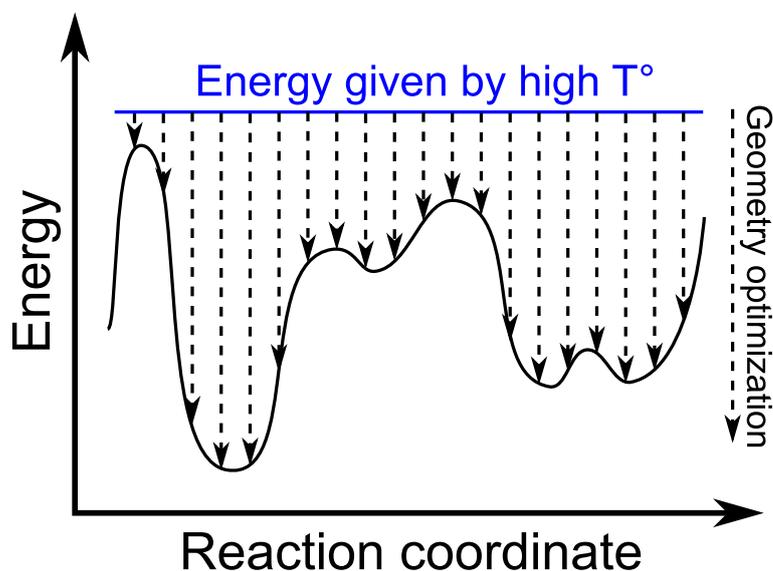
<sup>d</sup>The Lagrangian is expressed as the difference between the kinetic energy  $T$  and the potential energy  $V$ :  $L = T - V$

The process is repeated  $n$  times and leads to the trajectory that contains a series of *snapshots* or *frames* of the atomic positions.

The relevant MD parameters will be given in the chapters concerned.

### Quenched molecular dynamics

In this work, we use quenched molecular dynamics to sample the conformational space. It consists in classical molecular dynamics simulations during which geometries (“snapshots”) are saved at regular intervals (typically every 0.1 to 0.01 ns) and then optimized. Generally, a high temperature MD is performed to cross energy barriers, and the geometry optimization allows to slide down the potential energy surface and find local minima. This process is repeated multiple times on different geometries and at different temperatures to refine the sampling and to eventually obtain the low lying-energy conformers.



**Figure 3.3:** Schematics of the quenched molecular dynamic process. The high temperature allows to cross the barriers (blue line), and geometries are optimized at regular intervals to slide down the potential energy surface (black dashed arrows).

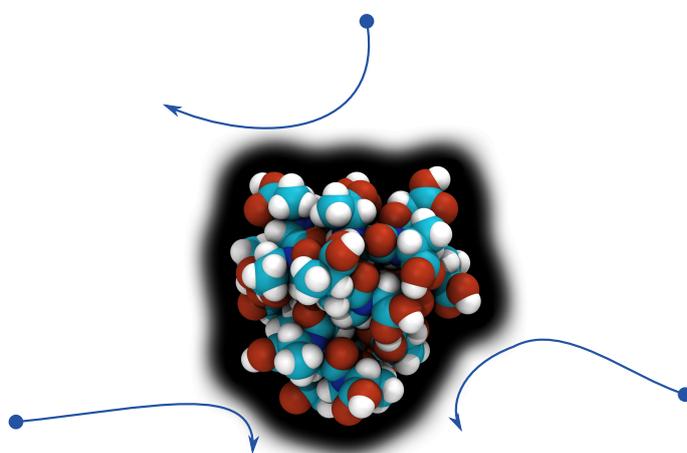
### 3.3.4 Collision cross section calculation

Ion mobility mass spectrometry experiments provide the collision cross section of the ions, which can be interpreted qualitatively, *i.e.*, two ions characterized by the same  $m/z$  have a different shape. Molecular modeling can help to provide structural insights behind the  $\Omega$  values by generating candidate structures, for example by molecular dynamics. Theoretical collision cross sections can be computed based on the geometry of the candidates and compared to the experimental values. The proposed conformers are validated when the collision cross section difference between

experiment and theory is typically within 5%.<sup>148</sup>

There are three main algorithms to compute the collision cross section: the Projection Approximation (PA),<sup>149</sup> the Exact Hard Sphere Scattering (EHSS)<sup>150</sup> and the Trajectory Method (TM).<sup>151</sup> Their complete description can be found in Ref. 152. In this work, we use the Trajectory Method, which is considered as the golden standard due to its high accuracy though it is also the most expensive computationally.

In this algorithm, collisions are simulated explicitly by firing a gas molecule toward the ion and integrate the gas molecule trajectory that is affected by the presence of the ion. Since the ion is likely randomly oriented in the experiment, the process is repeated for all possible orientation of the ion (**Figure 3.4**). Moreover, the energy distribution and spatial distribution of the gas is also taken into account by firing the gas molecule from different location around the ion (impact parameter) and using a distribution of initial velocities. The particularity of the Trajectory Method is the consideration of long-range interactions that affect the scattering of the gas molecules and thereby the trajectories. These effects are introduced by adding partial charges to account for the charge-charge interactions, as well as van der Waals interactions through a Lennard-Jones 12-6 potential.<sup>151</sup>



**Figure 3.4:** Representation of the Trajectory Method. The blue arrows represent examples of helium trajectories toward the ion. The black shaded contour schematize the interaction potential between the helium atom and the ion.

The theoretical collision cross sections were computed with helium as collision gas on candidate geometries using the Trajectory Method implemented in the MOB-CAL and Collidoscope softwares, the latter being preferred because it benefits from parallelization over many CPU cores.<sup>151,153</sup>

### Collision cross section notation

According to a recent review from the great minds of the IM-MS community,<sup>70</sup> the report of the collision cross sections should be as followed:

- Collision cross section should be symbolized by “CCS” or “ $\Omega$ ”.  $\Omega$  will be used throughout this work.
- A superscript “X” (prior to the  $\Omega$  symbol) is used to specify the experimental/theoretical method used to determined  $\Omega$ . In our work, “X” can be TW (“Travelling Wave”) for the experimental values or “TM” (“Trajectory Method”) for the theoretical values.
- A subscript “Y” (after the  $\Omega$  symbol) is used to specify the gas in which the experiment/calculation is performed. In our work, “Y” can be “ $N_2 \rightarrow He$ ” for the experimental values, which translates by “*measurement carried out in  $N_2$  and converted using values obtained in He*”, and “He” for the theoretical values.

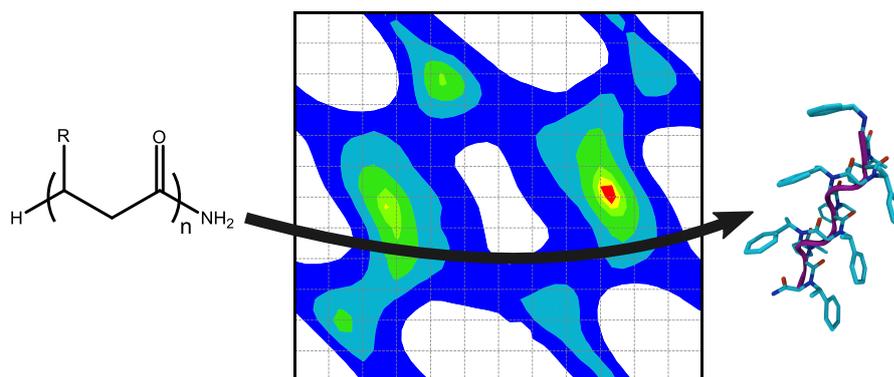
*In fine*, when the “collision cross section” is used in a generic way,  $\Omega$  is used, while  $^{TW}\Omega_{N_2 \rightarrow He}$  and  $^{TM}\Omega_{He}$  is involved to describe experimental and theoretical collision cross sections, respectively.



---

# PEPDROID: a DREIDING-based force field for peptoids

---



Based on

*Advanced Theory and Simulations*, 1, 12, **2018**, 1800089



## 4.1 Motivations

Computational chemistry or molecular modeling is nowadays essential in most research fields to rationalize experimental results and also predict the properties of new compounds/materials.<sup>55,154</sup> The research carried out in the peptoid field is no exception to this rule. Modeling played and continues to play an important role in the understanding of peptoid folding and in developing their applications.<sup>30,41,55,76,80</sup> Currently, “only” a few secondary structures are known and are mainly part of the helical family. However, given the tremendous amount of chemically diverse primary amines available that can be used as side chain carrier in the peptoid backbone, it is likely that new secondary structure motifs are still to be discovered. However, the number of possible side chain combinations is too large to be addressed efficiently experimentally. Molecular modeling can thus step in and help to establish the relationship between the primary structure, *i.e.*, the sequence of peptoid residues, and their secondary structure (if any).

Over the past 20 years, many papers about peptoids report the use of molecular modeling to investigate their secondary structure.<sup>30,35,40,41,76,81,83</sup> Quantum-mechanical methods were often employed to investigate model peptoids (typically made of 1 to 3 residues) and shed light on the backbone dihedral preferences. Although it provides very accurate data, the study of larger oligomers is not feasible since the number of degrees of freedom dramatically increases. Moreover, the dynamical aspects are limited using QM methods. In contrast, molecular mechanics and dynamics (MM/MD) are particularly adapted methods to investigate large systems over a long time scale and hence to explore a large conformational space<sup>76,81,100,155,156</sup> MM/MD requires the choice of a proper force field to describe the system under consideration. At the start of this work, most of the studies involving force fields to study peptoids were based on peptide force fields (CHARMM,<sup>81</sup> AMBER<sup>155</sup>). They typically refined some parameters, such as those linked to the  $\omega$  dihedral parameters, based on QM calculations performed on small model molecules and adjusted partial charges and vdW parameters for some atoms to reproduce experimental parameters (dipole moment, heat of vaporization).<sup>81,155</sup> Moreover, only simple side chains (*i.e.*, methyl also called *N*sar) were adequately reparametrized.

This has motivated the development of our own force field parameters that would be suitable for many different types of side chains. We chose the DREIDING force field as starting point, which is a robust force field used in many different areas, such as small molecules,<sup>157</sup> peptides<sup>158,159</sup> and polymers.<sup>145,160–163</sup>

## 4.2 DREIDING reparametrization methodology

The adaptation of the DREIDING parameters to describe the conformational space of peptoids is divided in three parts.

1. Several reports from the literature mention that the van der Waals parameters (well depth and/or equilibrium distance from Eq. 3.16 of hydrogen atoms borne by aliphatic carbons in DREIDING) might require adjustments to correctly reproduce certain properties,<sup>145,146</sup> such as the density.
2. In general, most “hard” bonded parameters of force fields adequately reproduce data from QM calculations. This assumption will be verified for a selected peptoid geometry.
3. Dihedrals are often the most crucial parameters to adjust in order to accurately describe conformers. Indeed, as the barriers rarely exceed a few kcal mol<sup>-1</sup>, a small change in energy can dramatically affect the molecular shape. Therefore, the dihedral parameters must be carefully validated to obtain reasonably accurate results.

### Hydrogen van der Waals parameters

We first start by checking the validity of the hydrogen van der Waals parameters since they will also affect the rotational barriers. We should recall that vdW parameters ( $R_0$ ,  $D_0$ ) are defined for homonuclear pairs in DREIDING. For hydrogen atom, the equilibrium distance and well depth are defined for the H – H bond, while parameters for interactions with other atoms is defined by the Lorentz-Berthelot rules (3.17).

The methodology to optimize the equilibrium distance ( $R_0$ ) for hydrogen has been achieved by comparing experimental collision cross sections obtained by IM-MS and corresponding theoretical<sup>a</sup> estimates in gas phase.

---

<sup>a</sup>The theoretical  $^{TM}\Omega_{He}$  were calculated using the Trajectory Method implemented in MOB-CAL<sup>153</sup> on candidate geometries of PEG, PLA and PCL generated with a reparametrized version of the DREIDING force field<sup>162</sup> in Materials Studio 6.0.<sup>164</sup>

We considered three polymers: (poly(ethylene glycol) - PEG,  $\alpha$ -methyl,  $\omega$ -hydroxy poly(lactide) – PLA and poly- $\epsilon$ -caprolactone - PCL) cationized by up to 2 Na<sup>+</sup>.<sup>b</sup>

To best parameterize the van der Waals parameters, we first used a fixed value of well depth (default value: 0.0152 kcal/mol) while varying the equilibrium distance between 2.75 and 3.195 Å, the latter being the default value. Each polymer ion was submitted to two consecutive quenched molecular dynamics (MD) at 600 K and 200 K for 10 ns each. Next, two consecutive MD at 298 K for 10 ns each were performed on the most stable structures provided by the last quenched MD run at 200K.  $T^M \Omega_{He}$  were computed on 100 geometries (saved every 0.1 ns) from the second MD and were finally averaged. The lowest RSMD between the experimental and theoretical data was obtained for an equilibrium distance of 2.83 Å (**Figure 4.1**).

To assess the reliability of the well depth parameter introduced in DREIDING for hydrogen, we selected different solvents (propane,<sup>165</sup> *n*-butane,<sup>166</sup> *n*-pentane,<sup>167</sup> *n*-hexane,<sup>168</sup> cyclohexane<sup>169</sup> and ethyl acetate<sup>170</sup>) and computed their density as well as their enthalpy of vaporization:<sup>c</sup>

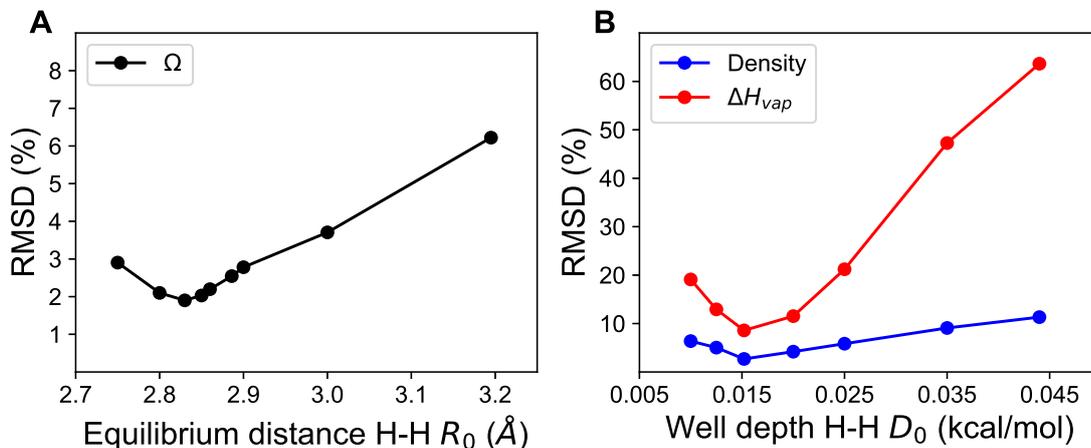
$$\Delta H_{vap} = \langle E_{intra} \rangle - \langle E_{total} \rangle + RT \quad (4.1)$$

where  $E_{total}$  is the total energy of the box and  $E_{intra}$  is the total energy of the isolated solvent molecules. We then compared the results to the experimental values obtained around room temperature available in the NIST<sup>165–170</sup> database. The calculations were performed using the optimum  $R_0$  equilibrium distance of 2.83 Å, while

<sup>b</sup>The polymer solutions were prepared at a concentration of 15  $\mu$ M in acetonitrile and cationized by 10  $\mu$ L of sodium iodide solution (13mM). The structures bear one or two positive charge(s) and have a degree of polymerization ranging from 3 to 70. Polymers were analyzed with a Waters Synapt G2-Si (see Chapter 3). All solutions were directly infused in the Electrospray ionization source (ESI) with a flow rate of 5  $\mu$ L/min, a capillary voltage of 3.1 kV, a source temperature of 100 °C and a desolvation gas temperature of 200 °C. IM spectrometry was carried out with nitrogen as the drift gas at a pressure of 2.89 mbar, an IM wave velocity of 800 m.s<sup>-1</sup> and wave height of 40 V. Data were analyzed in the Waters MassLynx program, and an arrival time distribution (ATD) was extracted for each polymer ion composition.<sup>124</sup> ATDs were then converted into collision cross sections ( $\Omega$ ) through Equation 3.2.<sup>124</sup>

<sup>c</sup>Boxes containing  $\sim$  1000 atoms were built for each solvent in the Materials Studio 6.0 package,<sup>164</sup> whose initial cell parameters were set according to their experimental densities. Atomic partial charges were assigned using the COMPASS force field.<sup>171</sup> The solvent boxes were first subjected to a high-temperature MD (NVT, 750K, 50 ps). The systems were then submitted to another MD at 298 K for 100 ps. Before fully relaxing the systems, a MD in the NPT ensemble was performed at high pressure (3 GPa, 298 K) for 10 ns, followed by another MD at ambient pressure until energy is converged. When the systems are converged, a last MD run (NPT, Nosé-Hoover thermostat at 298 K, Berendsen barostat at ambient pressure, 300 ps) is performed and the calculation of vaporization enthalpies performed through the Forcite module implemented in Materials Studio 6.0.<sup>164</sup> The Ewald summation method was used to describe the electrostatic and van der Waals interactions.

the well-depth is varied between 0.01 and 0.044 kcal/mol. We selected the best well-depth value according to the lowest RMSD, which is actually the default DREIDING value of 0.0152 kcal/mol (**Figure 4.1B**).



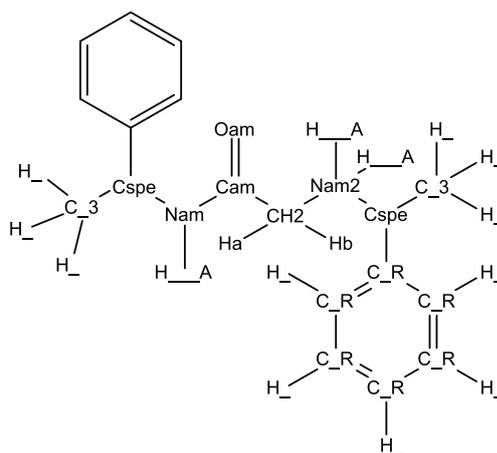
**Figure 4.1:** (A) Root mean square deviation (RMSD) between  $^{TW}\Omega_{N_2 \rightarrow He}$  and  $^{TM}\Omega_{He}$  for all polymer ions as a function of the van der Waals equilibrium distance  $R_0$  for aliphatic hydrogens. The optimum distance is 2.83 Å. (B) Root mean square deviation (RMSD) between experimental and theoretical enthalpies of vaporization ( $\Delta H_{vap}$ , red curve) and densities (blue curve) for all solvents as a function of the equilibrium well depth  $D_0$  obtained using the optimum  $R_0$ . The optimum well-depth is 0.0152 kcal/mol.

*In fine*, the refinement of the hydrogen vdW parameters leads to a very nice agreement between the experimental and theoretical results, either from gas phase (collision cross sections) or solution phase (density, enthalpy of vaporization).

### “Hard” bonded parameters

To assess the validity of the “hard” bonded parameters from DREIDING, namely bond stretching and angle bending, we compared the bond lengths and angles obtained in geometries that were optimized at the quantum mechanical and molecular mechanics levels, respectively. We chose the Møller-Plesset method at the second order with a cc-pVDZ basis set (MP2/cc-pVDZ) for the QM calculations. The MM calculation was performed using the default parameters of DREIDING, except for the vdW parameters of the aliphatic hydrogens (see previous section). The model system is a peptoid fragment bearing  $N_{spe}$  side chains (**Table 4.1**). We found a good quantitative agreement between the MP2 and original DREIDING parameters for the bond lengths and angles, leading to a root-mean-square deviation (RMSD) lower than 3% (**Table 4.1**). Since these terms are not expected to govern predominantly the three-dimensional structure of peptoids, such small deviations compared to the benchmark calculations fully validate the DREIDING default parameters.

**Table 4.1:** Comparison of selected bond lengths and angles for a model peptoid bearing two *N*spe side chains obtained after geometry optimization at the MP2/cc-pVDZ level and using the DREIDING force field (default parameters). The starting geometries are the same. The associated RMSD is displayed at the bottom of each table. The primary structure is represented using the force field types. A single aromatic ring is exemplified for sake of clarity.

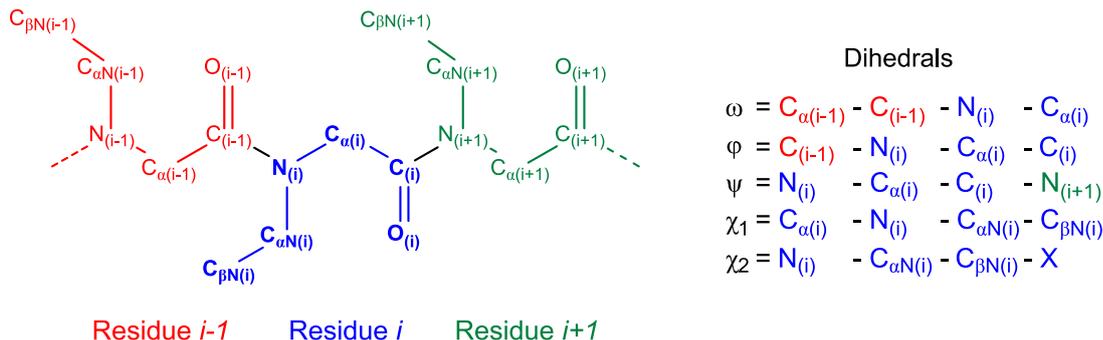


Bond pair	MP2 (Å)	DREIDING (Å)
Cspe - Nam	1.474	1.431
Nam - H_A	1.013	0.973
Nam - Cam	1.338	1.347
Cam - Oam	1.235	1.249
Cam - CH2	1.539	1.476
CH2 - Nam2	1.491	1.485
Cspe - H_	1.097	1.093
Cspe - C_3	1.532	1.540
C_3 - H_	1.091	1.091
Cspe - C_R	1.523	1.486
C_R - C_R	1.398	1.413
C_R - H_	1.087	1.021
CH2 - Ha/Hb	1.093	1.091
Nam2 - Cspe	1.538	1.488
Nam2 - H_A	1.021	1.026
<b>RMSD</b>		<b>2.56%</b>

Angle	MP2 (°)	DREIDING (°)
H_A - Nam - Cspe	118.362	117.229
H_A - Nam - Cam	114.680	116.473
Nam - Cam - Oam	124.755	116.759
Oam - Cam - CH2	117.609	118.435
Cam - CH2 - Nam	104.266	112.611
Cam - CH2 - Ha/Hb	110.212	107.035
Cspe - Nam - Cam	126.546	126.290
Nam - Cspe - H_	108.030	109.052
Nam - Cspe - C_3	108.083	108.597
Cspe - C_3 - H_	111.043	109.656
Nam - Cspe - C_R	112.355	111.742
H_ - Cspe - C_R	108.703	109.770
C_3 - Cspe - C_R	111.512	110.437
Cspe - C_R - C_R	119.495	119.579
C_R - C_R - H_	119.734	120.330
C_R - C_R - C_R	120.500	120.683
Ha/Hb - CH2 - Nam2	109.621	108.278
CH2 - Nam2 - Cspe	114.130	115.562
Nam2 - Cspe - H_	103.673	109.679
Nam2 - Cspe - C_R	108.645	110.580
Nam2 - Cspe - C_3	108.039	109.241
<b>RMSD</b>		<b>2.70%</b>

### Dihedral parameters refinement

As described earlier, the torsional profile is a very critical energy term in the description of three-dimensional structures. Since we intend to come up with a generic force field, we decomposed the peptoid structure into two components: the backbone vs. the side chains. The backbone is characterized by three distinct torsion angles ( $\omega$ ,  $\phi$  and  $\psi$ ) and the connectivity between the backbone and the side chains by up to 2 torsion angles ( $\chi_1$  and  $\chi_2$ ) (**Scheme 4.1**).



**Scheme 4.1:** Primary structure of a peptoid with its residue and dihedral definitions. “X” is presented as a wildcard that can be any atom bonded to the  $C_{\beta N(i)}$ .

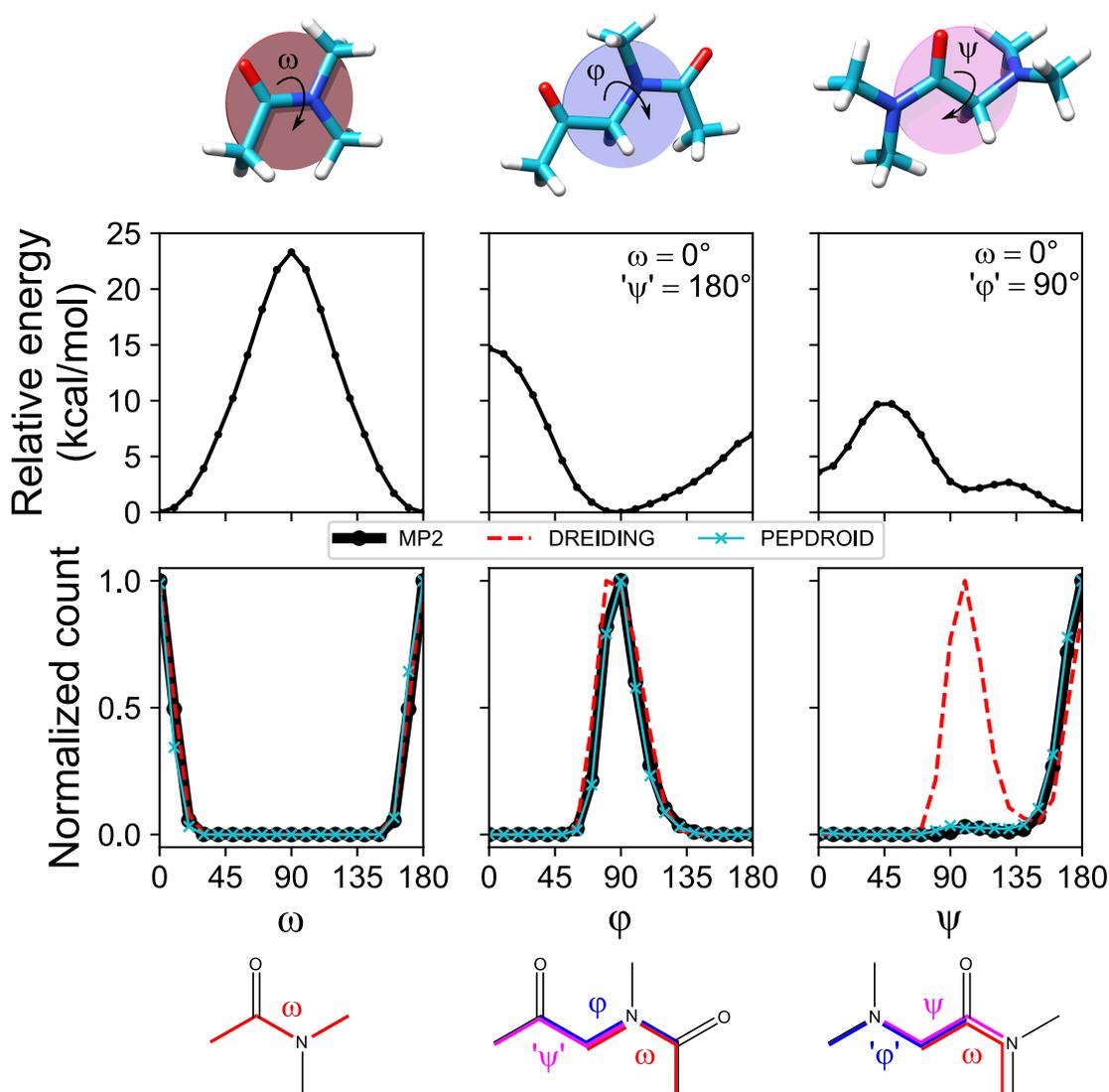
In order to validate whether DREIDING is adapted to accurately describe the dihedral angles of peptoids, we generated each torsion energy profile for peptoid fragments containing a dihedral of interest at the MP2 and DREIDING level<sup>d</sup>. From a technical point of view, whatever the level of the calculations (QM or MM), for a given set of dihedrals ( $\omega$ ,  $\phi$ ,  $\psi$ ), we optimize all geometrical parameters except the three angles when building torsion profiles associated to one of these dihedrals. For example, to build the  $\phi$  angle torsion profile,  $\phi$  is scanned by steps of  $10^\circ$  while the two other angles are constrained around their MP2 equilibrium values and all other parameters are free to relax. The constrains are indicated in the plots of the relative energy of each backbone dihedrals (**Figure 4.2**). These energy profiles were then converted into population profiles by using the Boltzmann equation at 298 K and normalized by the largest value:

$$N_i = \frac{e^{-E_i/k_B T}}{\sum_i e^{-E_i/k_B T}} \quad (4.2)$$

where  $N_i$  is the population associated to the dihedral value  $i$  with energy  $E_i$  and  $k_B$  is the Boltzmann constant. When we compare the shape of the profiles, we notice large deviations between QM and the DREIDING default values, thus preventing the use of DREIDING to accurately describe peptoid three-dimensional conformations.

<sup>d</sup>MP2 level refers to the QM level whereas DREIDING refers to the MM level.

As a result, we adjusted the DREIDING dihedral parameters from Eq. 3.14, namely the barrier height  $B_n$ , the phase factor  $d$  and the periodicity  $n$  to reproduce the MP2 populations profiles and called the new set of parameters “PEPDROID”. These parameters were systematically adjusted to yield a low RMSD between the QM and MM population profiles.

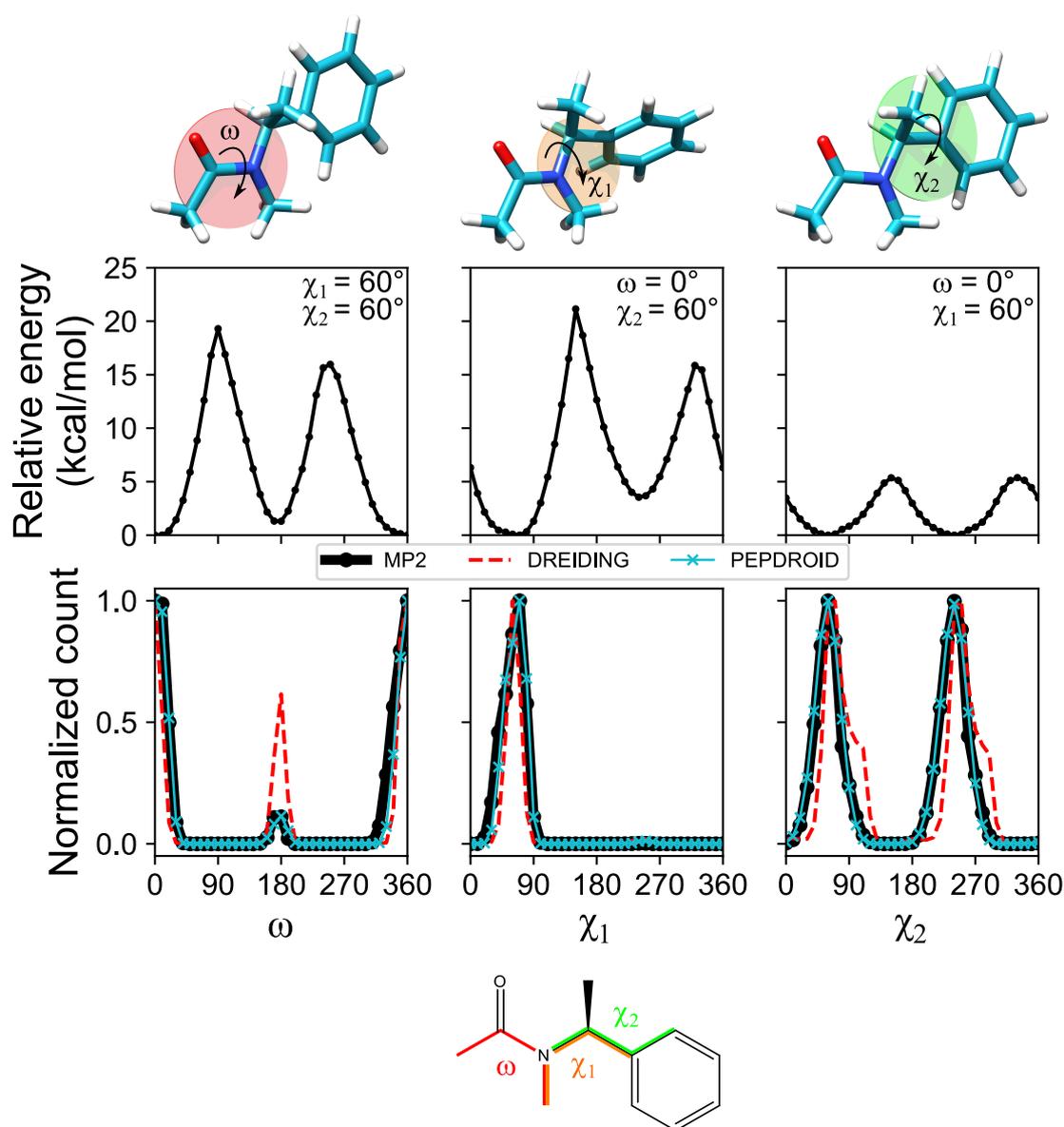


**Figure 4.2:** Energy profiles of the backbone dihedrals obtained at MP2/cc-pVDZ level (top) and normalized population count (at 298 K) obtained at MP2/cc-pVDZ (black), default DREIDING (red) and using the new sets of parameters PEPDROID (blue) (bottom). When applied, the restraints are displayed in the energy profiles. The primary structures of the models are displayed below the plots with the definition of the dihedrals.

The optimized parameters are only valid for the simplest peptoid side chain, namely sarcosine. In this research work, we want to study other side chains, typically the  $N_{spe}$  side chain. Our approach thus consists in decoupling the backbone and the side chains in order to add any new side chains in the parametrization. Three

different side chains were initially selected and used in a small peptoid model: (*S*)-*N*-phenylethyl (*N*spe), *N*-*n*-propyl (*N*npr), and *N*-benzyl (*N*pm).

In the same manner as for the backbone, we compared dihedral profiles of the side chain dihedrals ( $\chi_1$ , and  $\chi_2$ ) as well as  $\omega$  obtained at the MP2 level and at the MM level. We confirmed that the default parameters need to be adapted and followed the same procedure as previously described. For sake of clarity, we only present here the dihedral profiles for the *N*spe side chain (**Figure 4.3**), while the other profiles can be found in Appendix A.



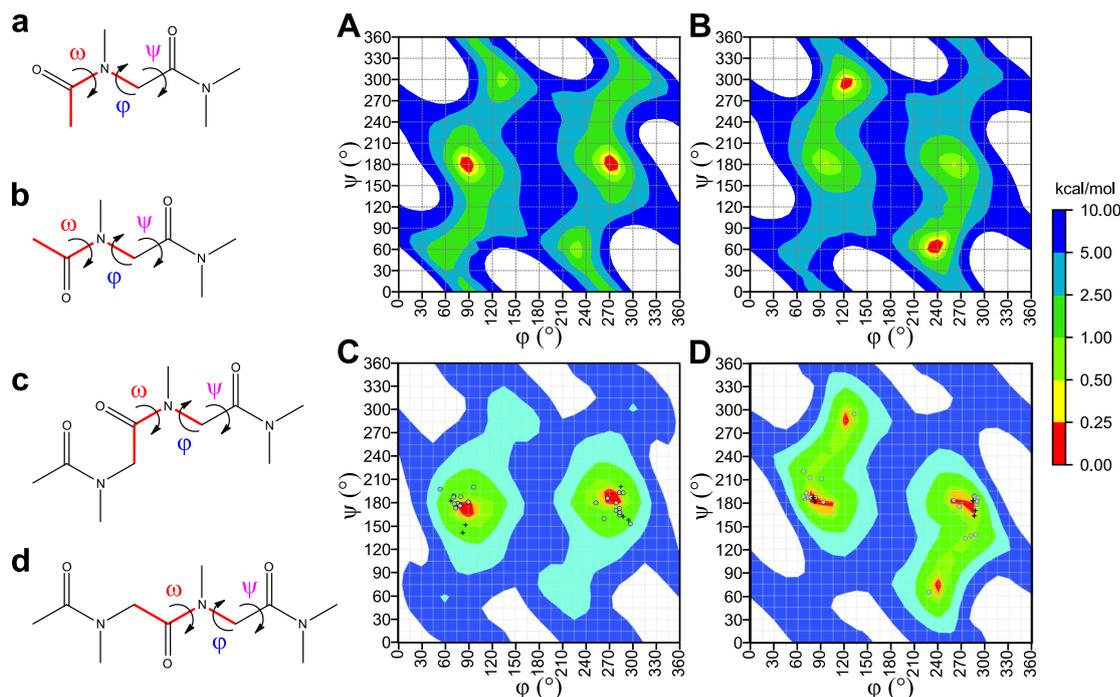
**Figure 4.3:** Energy profiles of the *N*spe side chain dihedrals obtained at MP2/cc-pVDZ level (top) and normalized population count (at 298 K) obtained at MP2/cc-pVDZ (black), default DREIDING (red) and using the new sets of parameters PEPDROID (blue) (bottom). When applied, the restraints are displayed in the energy profiles. The primary structures of the models are displayed below the plots with the definition of the dihedrals.

We also checked the validity of the *improper* dihedral parameters (Eq. 3.15). Only atoms of  $sp^2$  hybridization are concerned by the improper dihedral, that is the amide carbon and nitrogen for the backbone and the aromatic carbons present in *Nspe* and *Npm* side chains. The default value in DREIDING for the equilibrium angle is set to  $0^\circ$ , implying that the central  $sp^2$  atoms and the three connected atoms are in the same molecular plane. The equilibrium inversion angle obtained at the QM level and with the DREIDING default parameters for the model peptoid from **Table 4.1** is estimated to be  $4.3^\circ$  and  $0.6^\circ$ , respectively, thus giving confidence on the default parameter of DREIDING.

### 4.3 Force field validation

In order to validate the new set of parameters implemented in PEPDROID, we generated Ramachandran-like plots of model peptoid structures. These 2D diagrams are often used with proteins to rationalize their secondary structures.<sup>32</sup> These plots span the full range of two dihedral angles and provide a landscape of the PES; they have been marginally applied so far to peptoids.<sup>30,76,81</sup> Since the amide bond can adopt two conformations, namely *cis* and *trans*, both need to be considered to effectively get an insight into the entire potential energy surface of peptoids. We thus generated Ramachandran-like plots for the backbone dihedrals ( $\phi$ ,  $\psi$ ) of the smallest model peptoid containing every backbone dihedral, either in *cis* or *trans* conformation ( $\omega = 0^\circ$  or  $180^\circ$  respectively) (**Figure 4.4A-B**). For each data point of the 2D plots, we optimized all geometrical parameters of the molecules except for the two scanned angles ( $\phi$  and  $\psi$ ) and the  $\omega$  angle constrained to  $0^\circ$  or  $180^\circ$  for the *cis* and *trans* conformations, respectively. These plots were compared to those generated at the B3LYP/6-311+G(2d,p)//HF/6-31G\* level by Butterfoss *et al.* for a similar model peptoid (**Figure 4.4C-D**).<sup>30</sup>

The PES generated with our PEPDROID force field for the *cis*-model is in good quantitative agreement with that generated at the DFT level (**Figure 4.4A-C**). Two global minima, corresponding to the so-called *cis*- $\alpha_D$  conformations, are found at dihedral angles around  $\pm 90^\circ$  and  $180^\circ$  for  $\phi$  and  $\psi$ , respectively, in agreement with previous QM results.<sup>30,35</sup> The Ramachandran-like plot of the *trans*-peptoid model displays a very similar landscape when compared to that obtained by Butterfoss *et al.* (**Figure 4.4B-D**),<sup>30</sup> although some minor discrepancies are found owing to the fact that their peptoid contains an additional backbone unit which makes the analysis more complex.

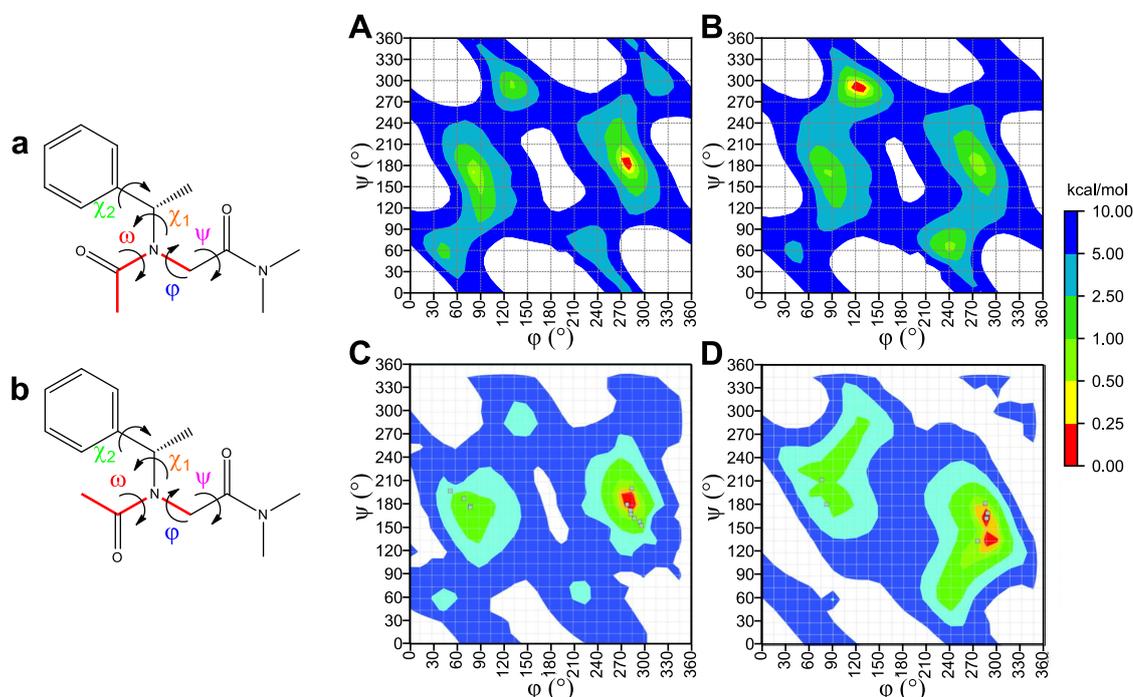


**Figure 4.4:** Ramachandran-like plots of a model peptoid with methyl side chains representing the backbone dihedrals ( $\phi$ ,  $\psi$ ) for the (A) *cis*- ( $\omega = 0^\circ$ , structure a) and (B) *trans*-amide conformations ( $\omega = 180^\circ$ , structure b) as generated with the PEPDROID parameters. Ramachandran-like plots of (C) structure c in *cis*- ( $\omega = 0^\circ$ ) and (D) structure d in *trans*-amide conformations generated at the B3LYP/6-311+G(2d,p)//HF/6-31G\* level by Butterfoss *et al.*<sup>30</sup> Adapted with permission from Ref. 30. Copyright 2009 American Chemical Society. The energy range spans from 0 to 10 kcal mol<sup>-1</sup>. The lowest energy structures (red) are set to 0 kcal mol<sup>-1</sup> in each plot, while the highest energy structures (up to 10 kcal mol<sup>-1</sup>) are displayed in blue. Structures with relative energies higher than 10 kcal mol<sup>-1</sup> correspond to the white color.

In particular, the global minimum for our model peptoid in the *trans* conformation corresponds to the *trans*-C<sub>7β</sub> conformation, as already identified by high level QM calculations by Moelhe *et al.*<sup>35</sup> for the same model peptoid. We have calculated at the MP2/cc-pVDZ level the relative energies of the *trans*-α<sub>D</sub> and C<sub>7β</sub> structures for our model peptoid and found that the fully optimized C<sub>7β</sub> structure is slightly more stable by 0.78 kcal mol<sup>-1</sup> compared to *trans*-α<sub>D</sub>; this in agreement with the 0.97 kcal mol<sup>-1</sup> energy difference obtained with PEPDROID. Although such subtle effects are not expected to be always fully reproduced with PEPDROID, the key aspect here is that PEPDROID can locate the most stable structures in the peptoid conformational space.

We next investigated the influence of the *N*spe side chain on the PES of the backbone dihedrals ( $\phi$ ,  $\psi$ ) and hence on the secondary structure preference of the peptoids. We used the same approach as previously described for the backbone, namely we generated Ramachandran-like plots for the *cis*- and *trans*-amide conformations for a model peptoid bearing a single *N*spe side chain. In contrast to *N*sar side chain, we

now observe an asymmetric PES, in both amide conformation, indicating that a chiral sterically hindered side chain restrains the available conformational space (**Figure 4.5**).

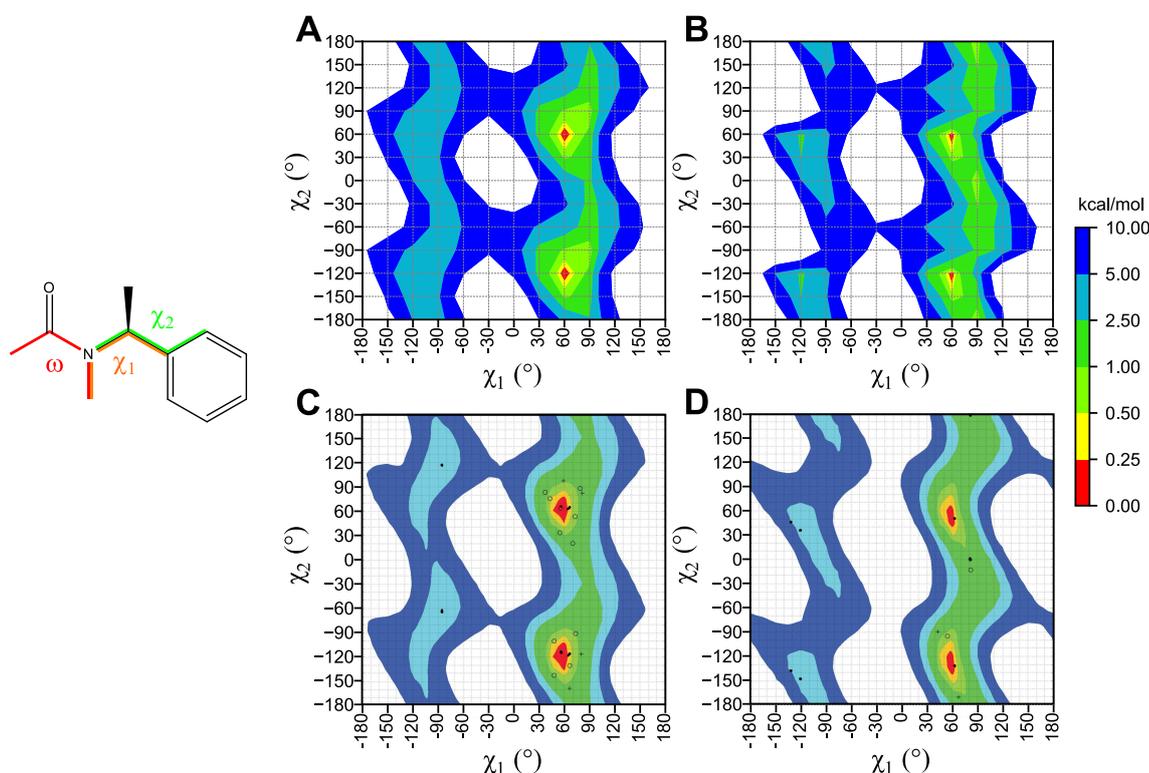


**Figure 4.5:** Ramachandran-like plots of a model peptoid with (*S*)-1-phenylethyl side chains associated to the backbone dihedrals ( $\phi$ ,  $\psi$ ) for the (A) *cis*- ( $\omega = 0^\circ$ , structure **a**) and (B) *trans*-amide ( $\omega = 180^\circ$ , structure **b**) conformations, as generated with the PEPDROID parameters. Ramachandran-like plots of (C) structure **a** in *cis*- ( $\omega = 0^\circ$ ) and (D) structure **b** in *trans*-amide conformations generated at the B3LYP/6-311+G(2d,p)//HF/6-31G\* DFT level by Butterfoss *et al.*<sup>30</sup> Adapted with permission from Ref. 30. Copyright 2009 American Chemical Society. The energy color code is the same as in **Figure 4.4**.

For the *cis* conformer, we once again obtain a good quantitative agreement with the DFT data reported by Butterfoss *et al.*,<sup>30</sup> with the global minimum corresponding to the  $\alpha_D$  conformation ( $\phi \sim -80^\circ$ ,  $\psi \sim 180^\circ$ ). For the *trans* form, our Ramachandran-like plot points to the  $C_{7\beta}$  conformation as the global minimum (see **Figure 1.3** for a representation of these conformers). Our calculations scanning the full range of each dihedral angle are not readily comparable to the QM results reported by Butterfoss *et al.* since they restricted their analysis to two specific values of  $\chi_1$  and were led to the conclusion that the *trans*- $\alpha_D$  conformation is the global minimum.<sup>30</sup> When performing MP2 geometry optimization on the global minimum extracted from PEPDROID and from Ref. 30, we find the *trans*- $C_{7\beta}$  conformation to be slightly more stable than the *trans*- $\alpha_D$  by  $0.41 \text{ kcal mol}^{-1}$  ( $0.51 \text{ kcal mol}^{-1}$  with PEPDROID), thus conforming the robustness of the description of stable secondary structures of peptoids provided by PEPDROID. Moreover, it also validates our reparametrization method since we

accurately describe the PES either in *cis* or *trans* although the side chain dihedrals were initially reparametrized in the *cis* conformation.

We went a step further in the validation of our parameters by comparing the Ramachandran-like plots of the side chain dihedrals to those generated by Renfrew *et al.*<sup>76</sup> We scanned the side chain dihedrals ( $\chi_1, \chi_2$ ) by step of  $15^\circ$  while keeping the  $\omega$  fixed at  $0^\circ$  or  $180^\circ$  (*cis* and *trans*), and optimized the geometry of each conformer. Our results match very well the data of Renfrew *et al.*<sup>76</sup> (Figure 4.6<sup>e</sup>) thus giving us supplementary evidence that the new set of parameters of PEPDROID is reliable.

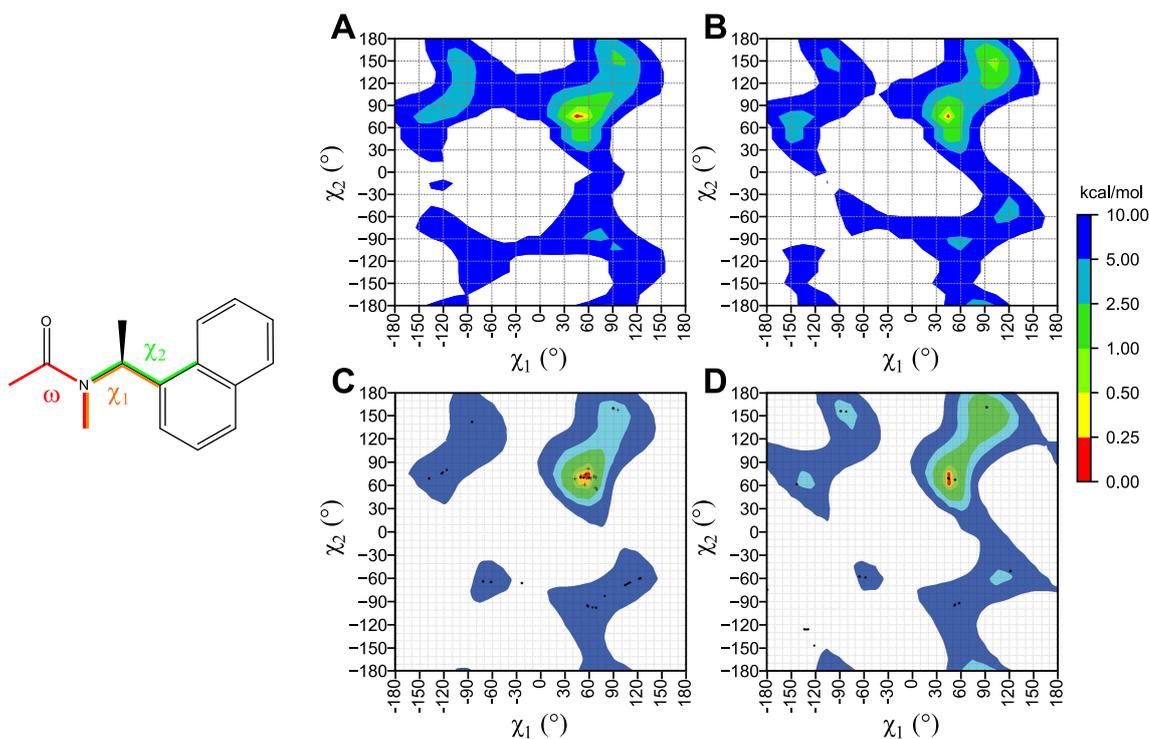


**Figure 4.6:** Ramachandran-like plots of side chain dihedrals ( $\chi_1, \chi_2$ ) for *Nspe* in the (A) *cis*- and (B) *trans*-amide conformation obtained with PEPDROID. Ramachandran-like plots in (C) *cis*- and (D) *trans*-amide conformation obtained at the B3LYP/6-311+G(d,p) DFT level by Renfrew *et al.*<sup>76</sup> Adapted with permission from 76. Copyright 2014 American Chemical Society. The energy color code is the same as in Figure 4.4.

We carried out the same methodology as explained above for two other side chains, *Nnpr* and *Npm* and also obtained a good agreement with the Ramachandran-like plots from Renfrew *et al.*<sup>76</sup> The dihedral energy profiles and Ramachandran-like plots can be found in Appendix A (Figures A.1, A.2, A.3, A.4).

<sup>e</sup>Note that the scale of the plots is different than those presented previously ( $0^\circ$  to  $360^\circ$  vs.  $-180^\circ$  to  $180^\circ$ ).

We also want to highlight another aspect of our new parameters which is related to their “transferability” for similar side chains. Indeed, without any further refinement we still accurately describe the PES of a model peptoid bearing (*S*)-*N*-(1-naphthylethyl) (*Ns1npe*) when we apply the dihedral parameters developed for the *Nspe* chain (**Figure 4.7**).



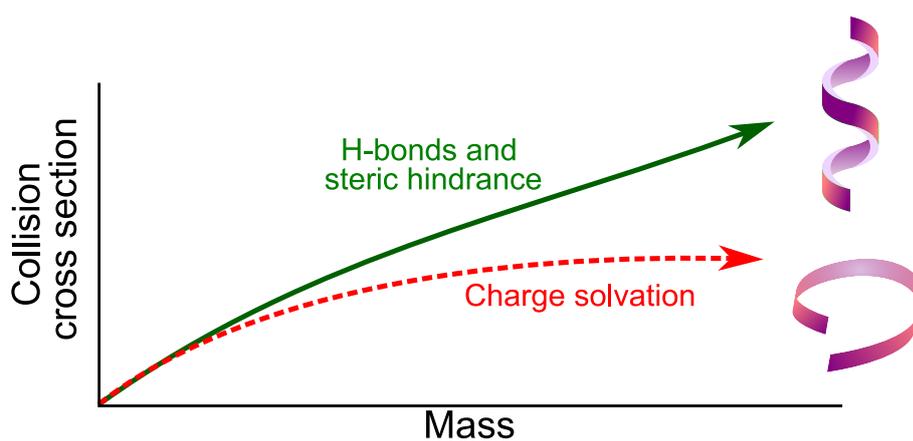
**Figure 4.7:** Ramachandran-like plots of side chain dihedrals ( $\chi_1$ ,  $\chi_2$ ) for *Ns1npe* in the (A) *cis*- and (B) *trans*-amide conformation obtained with PEPDROID. Ramachandran-like plots in (C) *cis*- and (D) *trans*-amide conformation obtained at the B3LYP/6-311+G(d,p) DFT level by Renfrew *et al.*<sup>76</sup> Adapted with permission from 76. Copyright 2014 American Chemical Society. The energy color code is the same as in **Figure 4.4**.

## Conclusions

We developed a new set of parameters based on the DREIDING force field to describe the conformational behavior of peptoids and called it ‘PEPDROID’. The key advantage is the decoupling of the backbone and side chain dihedral angles which permits to integrate any new desired side chain. We obtain a very nice agreement between the Ramachandran-like plots of dihedral pairs generated in other studies (see Ref. 30,76 and ours), revealing PEPDROID to be adequate to accurately describe the potential energy surface of peptoids. This force field will be used in the following chapters.



# Ion Mobility Mass Spectrometry of Peptoids

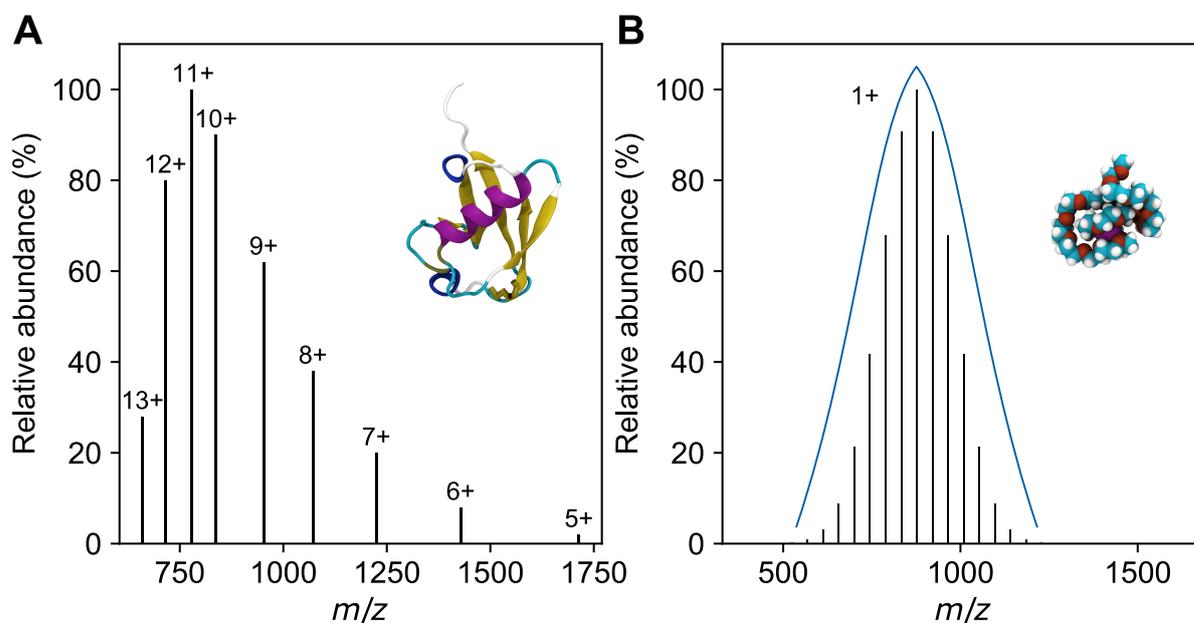


Based on  
*Biomacromolecules*, 21, 2, 2020, 903  
&  
*Biomacromolecules*, 2021, 10.1021/acs.biomac.1c00623



## 5.1 How to interpret data from IM-MS of oligo- and polymeric compounds?

In this research project, we will introduce ion mobility coupled to mass spectrometry (IM-MS) a new characterization method for peptoids. Peptoids are actually oligo/polymer compounds and are quite unique for the IM-MS analysis. Indeed, “traditional” compounds, such as natural molecules or even proteins, are characterized by unique compositions.<sup>123</sup> Depending on their nature and the ionization process, these compounds will eventually be detected as ions characterized by a distribution of charge states (**Figure 5.1A**<sup>a</sup>). Polymers are not only characterized by a charge state distribution, but also by a composition distribution due to their intrinsic chain length distribution (**Figure 5.1B**).<sup>124,162,172</sup>



**Figure 5.1:** (A) Mass spectrum of protonated ubiquitin<sup>173</sup> and (B) Mass spectrum of a chain distribution of singly charged poly(ethylene glycol): charge state vs. chain distribution. Adapted with permission from Ref. 173.

When subjected to ion mobility separation, polymer ions are characterized by their own arrival time distribution, leading to a range of collision cross sections  $\Omega$  depending on their chain length. The evolution of the collision cross section with the chain length, and hence the mass (or  $m/z$ ), may be used to determine the gas phase conformation of polymers. This method is generally referred to as “*trend line analysis*” and consists in fitting the evolution of  $\Omega$  as a function of the mass using the following equation:

<sup>a</sup>The ubiquitin is rendered using VMD<sup>31</sup> from the “1UBQ” PDB (<https://www.rcsb.org/structure/1UBQ>)

$$\Omega = A.M^B \quad (5.1)$$

where  $A$  and  $B$  are fitting parameters and  $M$  is the molecular mass. The parameter  $A$  is related to the density of the ion<sup>174</sup> and  $B$  reflects the general shape of the ions. Most singly charged polymer ions (of different nature), for instance poly(ethylene glycol)<sup>124,175,176</sup> or  $\alpha$ -methyl,  $\omega$ -hydroxy poly(lactide),<sup>124</sup> adopt a globular conformation in gas phase. For such conformation,  $B$  is systematically close to  $2/3$ , which is typical for a spherical system of increasing size. Indeed, the collision cross section of a sphere can be approximated by its area:

$$\Omega_{sphere} = \pi R^2 \quad (5.2)$$

where  $R$  is the sphere radius. The sphere volume is thus equal to:

$$V_{sphere} = \frac{4}{3}\pi R^3 = \frac{M}{\rho} \quad (5.3)$$

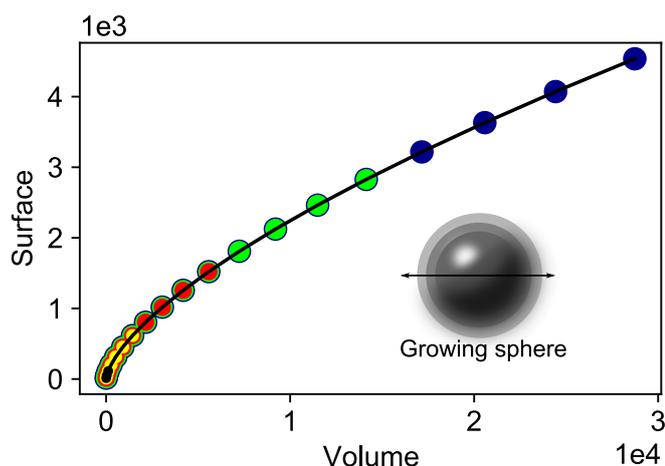
where  $M$  is the mass and  $\rho$  is the density. By combining Equations 5.2 and 5.3, we obtain the following relationship where the parameter  $B$  from Equation 5.1 is equal to  $2/3$ :

$$\Omega_{sphere} = \left( \frac{3\sqrt{\pi}}{4\rho} \right)^{2/3} M^{2/3} \quad (5.4)$$

Since  $M$  and  $\rho$  are directly related, the collision cross section is also directly proportional to  $V^{2/3}$ . In a purely geometric way, Equation 5.4 can be approximated by the evolution of the surface of a sphere as a function of its volume (**Figure 5.2**).

This value of  $2/3$  is the reference in polymer analysis by IM-MS. When  $B$  is lower than  $2/3$ , the conformation of the ions throughout the distribution is more compact than a sphere, while when  $B$  is higher than  $2/3$ , the conformation is more extended. Therefore, a value close to 1 is expected to be characteristic of fully extended structures growing linearly with the number of monomer units (or the mass), typically helices.

Since peptoids may adopt a helical conformation, we want to rationalize this hypothesis made to avoid drawing precocious conclusions. With this aim in mind, we generated theoretical trend lines of linearly growing (hypothetical) helices and computed the value of the parameter  $B$ . We chose two different class of compounds, peptides and peptoids, that can form helices of different geometries. Their geometry is directly related to the value of their backbone dihedral angles ( $\omega$ ,  $\phi$  and  $\psi$ ) and hence to the nature of their side chain. The L-alanine amino acid is known to strongly enforce the  $\alpha$ -helical conformation in peptides ( $\omega = 180^\circ$ ,  $\phi = -60^\circ$  and  $\psi = -40^\circ$ ),<sup>27,177</sup> while the  $N_{spe}$  side chain is known to induce a helical geometry close to the poly(proline)



**Figure 5.2:** Evolution of the surface of a sphere as a function of its volume (using arbitrary units). The different ranges considered (black, yellow, red, green, dark blue) lead to the same  $B$  parameter, which is constant and equal to  $2/3$ . This evolution follows exactly the same trend as the evolution of the collision cross section as a function of the mass of globular ions.

type I in peptoids ( $\omega = 0^\circ$ ,  $\varphi = -70^\circ$  and  $\psi = 180^\circ$ ) that will be referred as “peptoid helix”.<sup>41,42</sup>

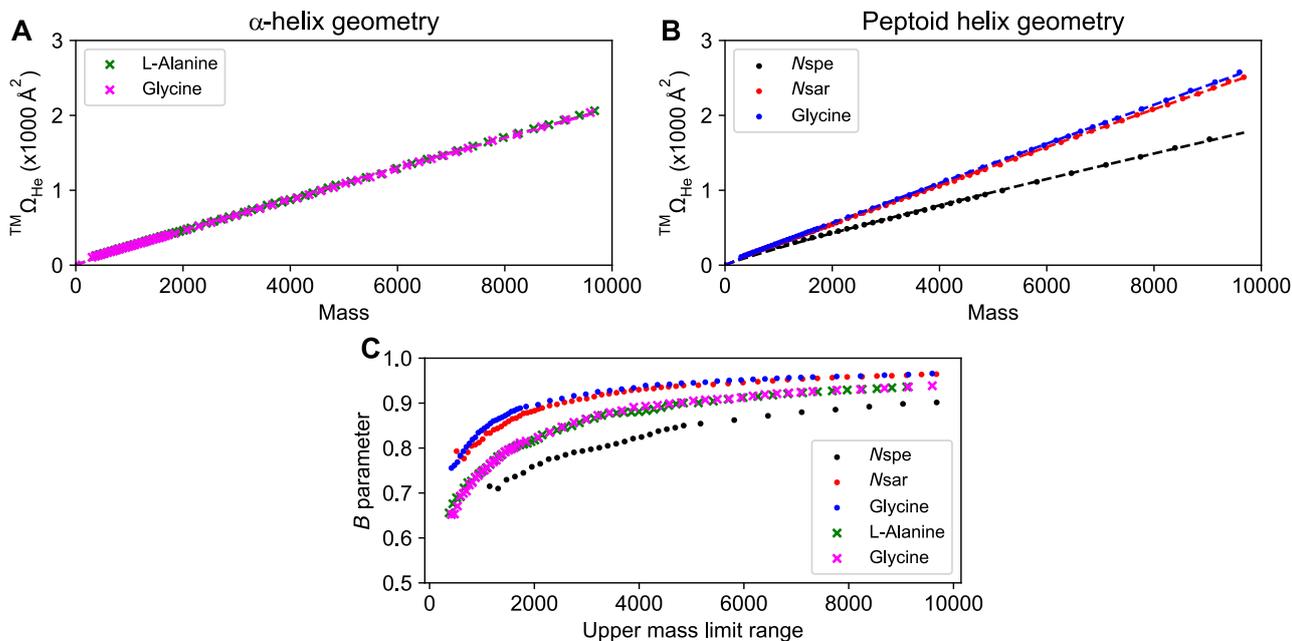
For both helical geometries, we built growing helices up to a mass of  $\sim 10,000$  Da using different side chains: hydrogen, methyl and (*S*)-1-phenylethyl<sup>b</sup>. The 3D coordinates of these helices were built in the Materials Studio 18.0 package and their geometry was not optimized. Their theoretical  ${}^{TM}\Omega_{He}$  were computed using Colli-doscope (one helix = one  $\Omega$ ) and their evolution as a function of the mass was fitted using Equation 5.1 (Figure 5.3A-B).<sup>153</sup>

<sup>b</sup>This side chain cannot be introduced into the  $\alpha$ -helical geometry due to steric clashes.

**Table 5.1:**  $A$  and  $B$  parameters obtained by fitting the  $^{TM}\Omega_{He}$  vs. mass curves from **Figure 5.2** for the  $\alpha$ - and peptoid helical geometries.

	$A$ parameter	$B$ parameter
Peptoid helix geometry		
$N_{spe}$	$0.452 \pm 0.021$	$0.901 \pm 0.0005$
$N_{sar}$	$0.358 \pm 0.007$	$0.964 \pm 0.0002$
Glycine	$0.364 \pm 0.021$	$0.966 \pm 0.0002$
$\alpha$ -helix geometry		
L-Alanine	$0.378 \pm 0.012$	$0.936 \pm 0.0004$
Glycine	$0.368 \pm 0.012$	$0.939 \pm 0.010$

After the fitting of the  $N_{spe}$  trend line, we obtain a  $B$  parameter of 0.901, which is strikingly different from the  $^{2/3}$  of the globular shape but is not equal to 1 as expected. This observation holds true for the two helical geometries and for the different side chains, although the values are not equal, as summarized in **Table 5.1**.

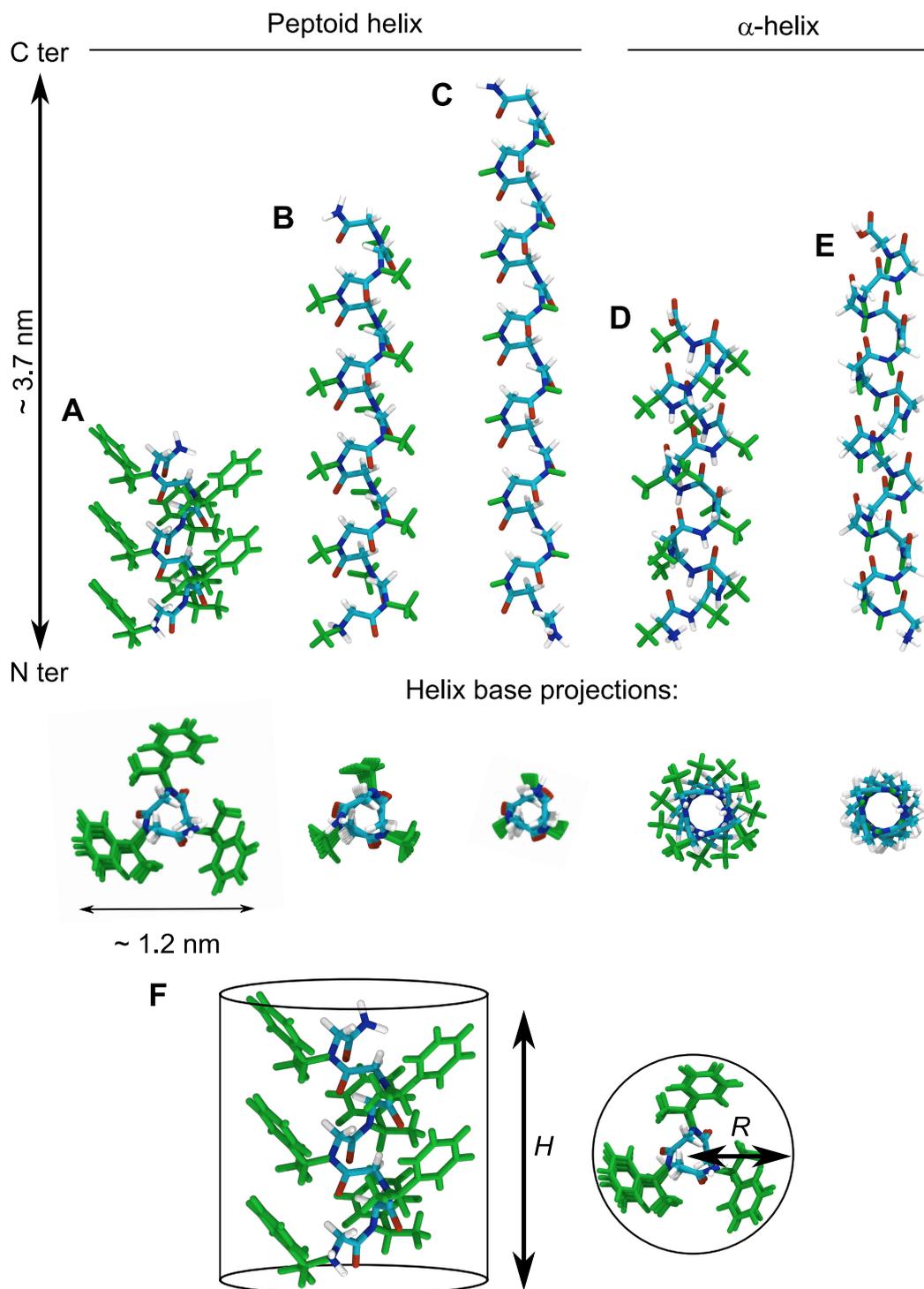


**Figure 5.3:** Theoretical trend lines of  $^{TM}\Omega_{He}$  as a function of the mass for (A) the  $\alpha$ -helix geometry with hydrogen (glycine) and methyl (alanine) side chains and for (B) the peptoid helix geometry with hydrogen (glycine), methyl (alanine) and (*S*)-1-phenylethyl ( $N_{spe}$ ). (C) Evolution of the  $B$  parameter with the mass range for the  $\alpha$ -helix geometry (crosses) and the peptoid helix geometry (dots), obtained by fitting the curves from (A) and (B) by increasing the upper limit mass range.

Although the upper mass limit, defined as the highest mass of the considered mass range, is roughly the same for every helices, the  $B$  parameter value varies between the helical geometries as well as with the side chain. Moreover, when examining the evolution of the  $B$  parameter for a given helical geometry and side chain over different mass range (always starting from the minimum), we observe an increase of

this parameter as the mass range increases. To calculate this evolution, we proceed as follows: starting at the origin, we determine the  $B$  parameter when increasing the mass range and plot the  $B$  value against the highest mass of the considered mass range (**Figure 5.3C**). However, the value of  $B$  increases at a different “rate” depending on the geometry as well as on the side chain. Indeed, when comparing the  $\alpha$ - and peptoid helices bearing methyl side chains that have the same mass, the former is more “compact” due to its specific dihedral angle combination (**Figure 5.3B,D**). Moreover, the helix base, *i.e.*, the 2D helix projection, of  $\alpha$ -helices has a circular shape with the hydrogen atoms or methyl moieties tilted toward the inner core of the helix, while the peptoid helices feature a triangular base shape, with the side chains pointing out at the edges of the triangle (**Figure 5.4A,C**).

From these observations, we can rationalize the evolution of the  $B$  parameter in a purely geometric way. Indeed, the helix can be approximated by a cylinder in which the height corresponds to the ‘number of monomer units’ (and equivalently, the mass) and the radius to the distance between the helix center and the atom lying the furthest away (**Figure 5.4**).



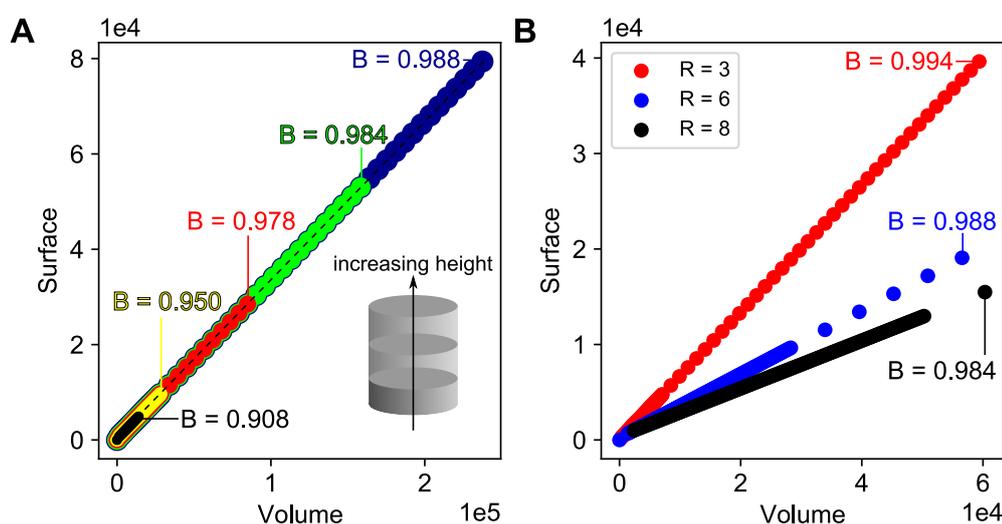
**Figure 5.4:** Ideal helical structures (viewed sidewise on top and along the helix axis from the C terminus at the bottom) for peptoids and peptides having about the same mass. (A), (B), and (C) are ideal peptoid helices ( $N_{spe}$ ,  $N_{sar}$ , and glycine, respectively); (D) and (E) are ideal  $\alpha$ -helices (L-alanine and glycine, respectively). Side chains are highlighted in green. (F) The helix is approximated by a cylinder (here, peptoid helix with  $N_{spe}$  side chains), in which the height of the cylinder  $H$  corresponds to the distance between the N and C termini and the radius  $R$  of the basis corresponds to the distance between the helix center and the outermost atom of the side chain.

For increasing helix size, the helix base (or its radius) is constant, while its height keeps increasing. For a cylinder of radius  $R$  and height  $H$ , the total surface  $S$  and volume  $V$  are expressed as:

$$S = 2\pi R(R + H) \quad (5.5)$$

$$V = \pi R^2 H \quad (5.6)$$

As demonstrated for the evolution of the surface of a growing sphere, we can approximate the collision cross section by the surface of the object. By analogy to the growing helices from **Figure 5.3**, we consider a cylinder of constant radius and increasing height, and plot the evolution of the volume as a function of the surface (**Figure 5.5A**). We estimated the radius of the peptoid helix with  $N_{\text{spe}}$  side chains to be  $\sim 6$  (Å) and the ‘height’ of a monomer unit to  $\sim 2$  (Å) and used these values to generate the cylinder. We fitted the evolution of the surface as a function of the volume using Equation 5.1 over a growing range of volume and observed the same trend as for the evolution of the  $\Omega$  as a function of the mass (see the different colors of **Figure 5.5A**): the  $B$  parameter is not equal to 1 but rather increases with the mass range under consideration and tends to 1.



**Figure 5.5:** (A) Evolution of the surface as a function of the volume of a cylinder (with arbitrary values of  $R = 6$  and  $H$  varying from 12 to 2100 and arbitrary units). The different colors represent different ranges of  $H$  values selected to perform the power fit ( $S = A.V^B$ ). Increasing the range of volumes makes the  $B$  parameter increasing, as observed for collision cross sections as a function of the mass. (B) Evolution of the surface as a function of the volume of a cylinder for different radius values ( $R = 3, 6$  or  $8$  and  $H$  varying from 12 to 2100).

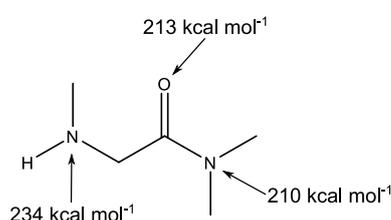
If we now vary the radius with the same height values, we indeed observe that the associated  $B$  parameter is also close to 1 (**Figure 5.5B**) but reaches more rapidly this value when the radius is small ( $R = 3$ ); this fully matches the results obtained for the two helical geometries. The peptoid helix has a triangular base shape which is smaller than the one from  $\alpha$ -helices which is circular for the hydrogen and methyl moieties. The  $B$  parameter grows more rapidly toward 1 than the larger  $\alpha$ -helix and even more than the peptoid helix with  $N_{\text{spe}}$  side chains. Therefore, there is an intimate relationship between the helix length and the “radius” of its base.

In view of these results, the trend line analysis of compounds that do not adopt globular conformation in gas phase is very ambiguous since it depends on the range of data that we consider. In the examples above, we purposely considered extreme mass range, and observed that  $B$  only tends to 1 when the mass tends to infinity. However, if we consider mass range experimentally accessible, *i.e.*, up to 3000-4000 Da, the  $B$  parameter is larger than  $2/3$  and drastically varies depending on the helix geometry (from 0.79 to 0.9, **Figure 5.3C**). Therefore, such trend line analyses may be misleading, and much care should be taken in structural interpretation of non-globular compounds. This conclusion strengthens the requirement of molecular modeling to fully interpret the experimental ion mobility data and investigate the three-dimensional conformation of gaseous ions.

In the next section, we will describe the general computational methodology that was used to generate candidate geometries for the collision cross section calculations to be compared to the experimental data obtained by ion mobility mass spectrometry experiments.

## 5.2 Methodology

Simulations were performed with Materials Studio 18.0 using the PEPDROID force field parameter set described in Chapter 4.<sup>144</sup> In the following sections, several side chains will be discussed. Moreover, we will exclusively discuss singly charged peptoids, more precisely singly protonated. Generally, peptoids have a secondary amine moiety at their N terminus extremity which is the most basic site. We verified this hypothesis by computing the proton affinity (PA)<sup>c</sup> of a model molecule bearing a secondary amine and an amide moiety, whose protonation is disfavored by about 20 kcal/mol compared to the amine (Figure 5.6).



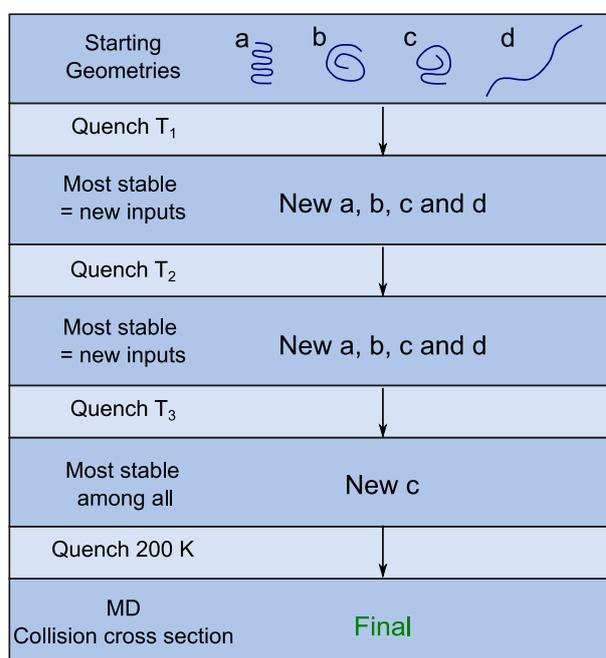
**Figure 5.6:** Model molecule that mimics the peptoid backbone with a secondary amine and an amide moiety. The proton affinity of the amine and amide are displayed (MP2/cc-pVDZ).

For any type of side chain, and for each polymerization degree, multiple starting geometries, carrying a proton on the terminal amine, were built (random coil, loop, helical, etc.). Partial charges are defined with the Gasteiger method based on the electronegativity of the bonded elements.<sup>147</sup> Each starting geometry is first optimized at the molecular mechanics level by using the Conjugate Gradient algorithm with a 200 Å cutoff value for the non-bonded interactions so that none of them are neglected. Each optimized structure is then used as the starting point of a conformational search. To sample the conformers, we used quenched molecular dynamics simulations (that will be referred to as “quenched dynamics” for the rest of the manuscript). It consists in classical molecular dynamics simulations during which geometries (“snapshots”) are saved at regular intervals (typically every 0.1 to 0.01 ns) and then optimized. The control of the temperature is ensured by the Nosé-Hoover thermostat. Generally, a high temperature MD is performed to cross energy barriers, and the geometry optimization allows to slide down the potential energy surface and find local minima (Figure 3.3).

Our conformational sampling workflow works as follows: first, we perform at least 3 consecutive quenched dynamics for 10 to 50 ns on each starting geometry

<sup>c</sup>The PA is defined as the opposite of the protonation reaction enthalpy ( $-\Delta H_{\text{protonation}} = -(H_{\text{protonated}} - H_{\text{neutral}} - H_{\text{proton}})$ ). Reported PAs are obtained from thermal enthalpies calculated at 298.15 K at the MP2/cc-pVDZ quantum mechanical level.

(NVT<sup>d</sup>: N = number of atoms, V = volume, T = temperature), typically using a temperature range of:  $T_1 = 800$  K,  $T_2 = 600$  K and  $T_3 = 400$  K. Such high temperatures are required to cross the energy barriers between the different conformers, especially between the *cis* and *trans* amide bonds. Indeed, the barrier height is about 20 kcal mol<sup>-1</sup>, which is too large to be crossed at room temperature in a reasonable time scale of simulation. The effect of the temperature on the *cis-trans* transition is demonstrated in **Figure 5.8**. Transitions start occurring at 600 K, but the conformers are still trapped for several tens of ns, while more frequent interconversions appear at 700 and 800 K, which will most likely better sample the potential energy surface. Moreover, as the peptoid chains grow, steric hindrance will increase and affect these transitions.



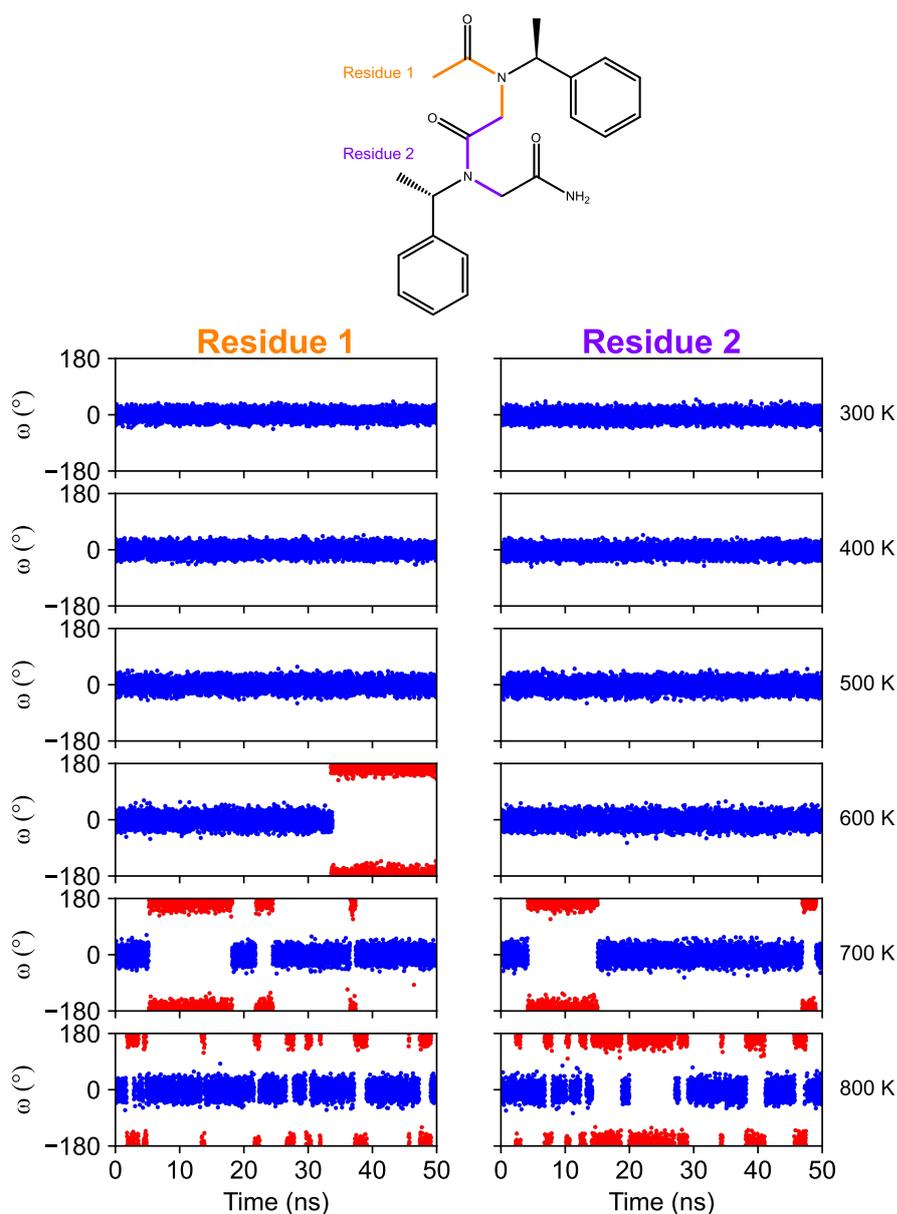
**Figure 5.7:** Workflow for the sampling of the conformational space of peptoids.

For each starting geometry, the most stable conformation among those generated when carrying out the first quenched dynamics at  $T_1$  was used as the starting point of the second quenched dynamics at  $T_2$ ; the most stable conformation from this second run was next used as a new starting point for the quenched dynamics at  $T_3$ . Finally, a last quenched MD at lower temperature is performed on the most stable conformation among each “initial” geometries at  $T_3$  (NVT;  $T = 200$  K, frames optimized every 0.1 to 0.01 ns). The workflow is summarized in **Figure 5.7**. This method of combining multiple quenched MD at different temperatures ( $T_1$ ,  $T_2$  &  $T_3$ ) has already proven to be efficient to explore the potential energy surface.<sup>157,160,178,179</sup> Performing a last quenched dynamics at low temperature allows us to focus on the minima close to the equilibrium structure. Finally, two successive molecular dynamics (NVT;  $T = 298$  K,

<sup>d</sup>During these simulations, no periodic boundary conditions are used.

10 then up to 50 ns) were performed on the most stable conformation obtained during the last step. The first is used to equilibrate the system and the second to generate from 100 up to 500 conformations (frames saved every 0.5 or 0.1 ps). Lastly, these geometries are subjected to the collision cross section calculation using the Collidoscope software.<sup>153</sup> The evolution over time is monitored and their average value is compared to the experimental values.

Experimental collision cross sections are obtained by converting the arrival time distributions into  $\Omega$  using the calibration procedure described in Chapter 3.2.

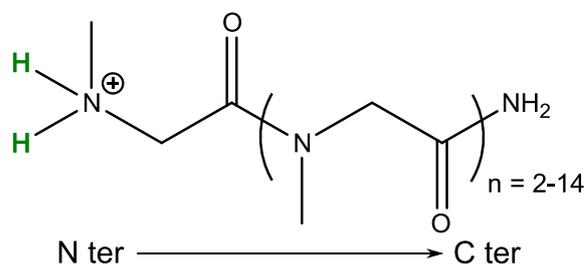


**Figure 5.8:** Model peptoid bearing two *N*spe side chains with the two highlighted  $\omega$  dihedrals (orange and purple) on each residue. Evolution of the value of the amide  $\omega$  dihedral angle over 50 ns in molecular dynamics for temperature ranging from 300 to 800 K. The *cis* conformation is represented in blue, while the *trans* is represented in red. Transition between *cis* and *trans* only starts occurring at higher temperature in these simulation time scales.

## 5.3 Results and Discussion

### 5.3.1 The simplest case: sarcosine peptoids

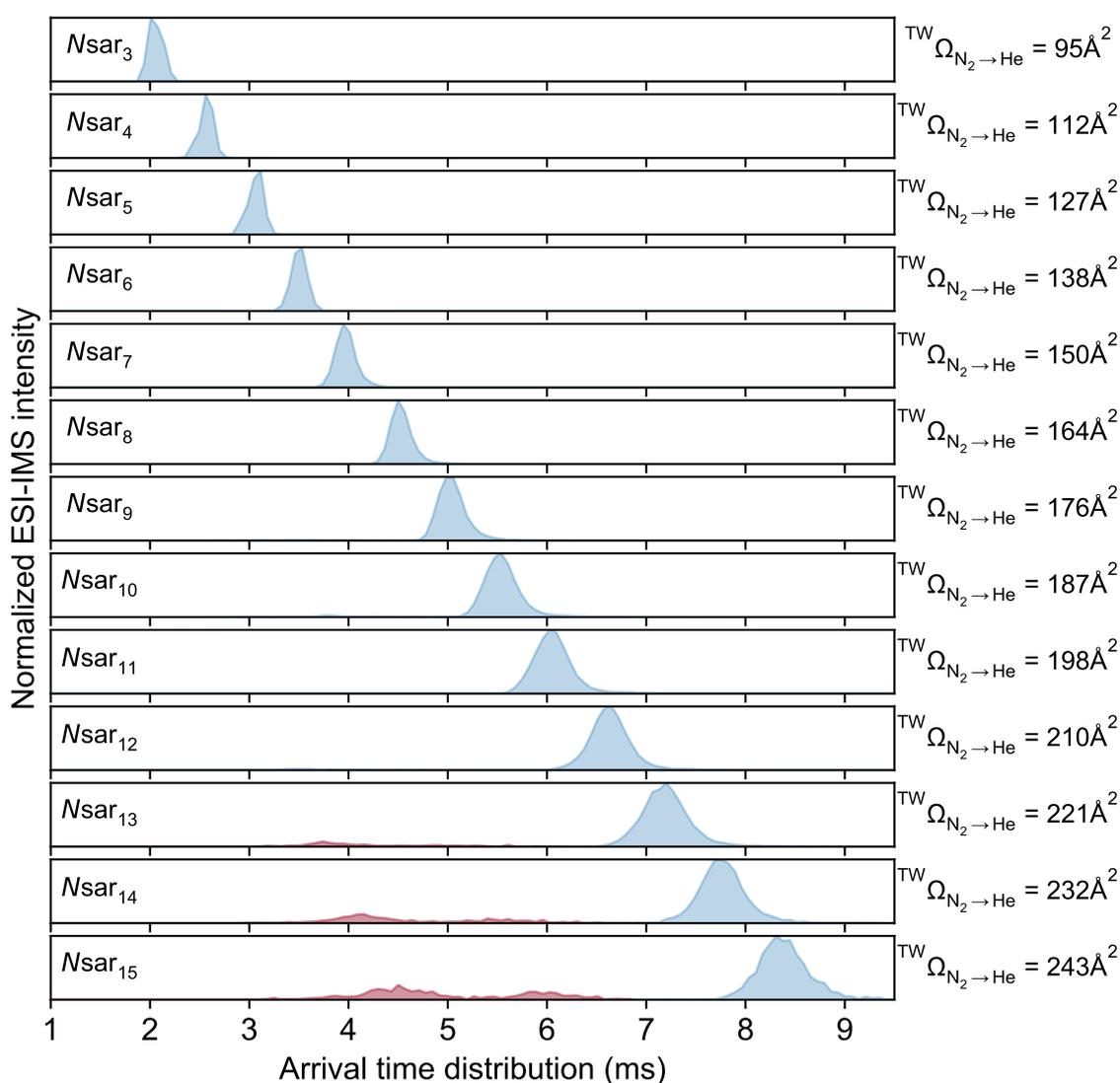
Among the possible side chains that can be carried by a peptoid, the *N*-methyl (abbreviated *Nsar*) is by far the most simple. This peptoid behaves as a random coil in solution, with the amide bonds being nearly equally present either in *cis* or *trans* conformation.<sup>11,63,180</sup> Consequently, we do not expect a particular behavior of these peptoids in gas phase.



**Figure 5.9:** Primary structure of the poly(sarcosine) peptoid protonated at the N terminus.

*Nsar* peptoids were synthesized from 3 to 15 units ( $Nsar_{3-15}$ , **Figure 5.9**). Compared to classical polymers whose sample is intrinsically polydisperse, peptoids synthesized using the submonomer method are monodisperse.<sup>112</sup> Consequently, the trend line analysis of the evolution of their collision cross section as a function of their mass is performed after the analysis of each sample containing a single chain length. After their desolvation by the Electrospray ionization source, *Nsar* peptoids are mainly detected as singly protonated species, whose proton location is most likely on the secondary amine.

For each chain length, we obtained a symmetric unimodal arrival time distribution after the ion mobility analysis, indicating that the ions are either in a single stable conformation or rapidly interconverting between different “extreme” conformations, *i.e.*, very compact and extended (**Figure 5.10**).<sup>181</sup>

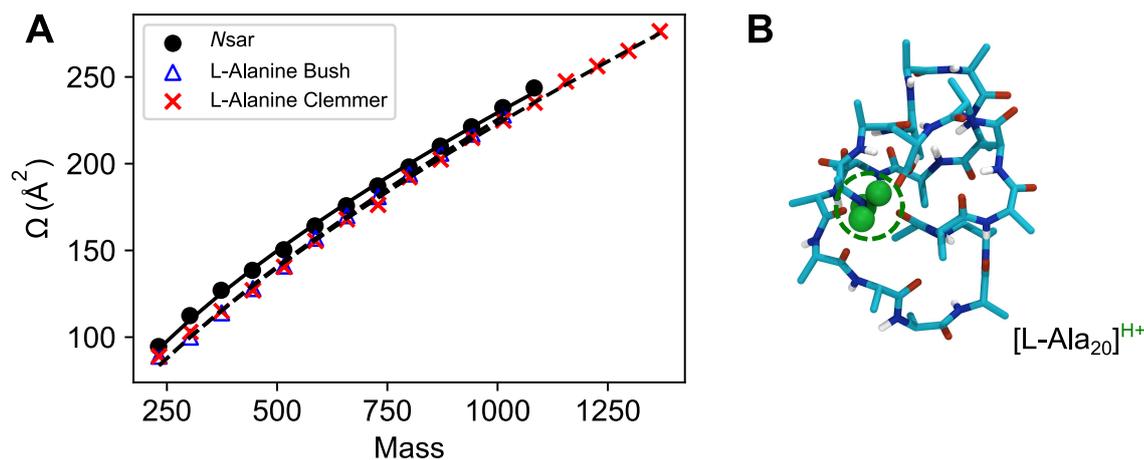


**Figure 5.10:** Arrival time distributions for each  $N_{sar}$  peptoid chain length (light blue). For longer chain length (13 to 15), background ions are present (light red).

Interestingly, the collision cross sections obtained by conversion of the ATDs (see Chapter 3.2) are very similar to those<sup>e</sup> obtained by Clemmer *et al.* and Bush *et al.* for singly protonated poly(L-alanine) ions, which are almost exact regioisomers of  $N_{sar}$  peptoids (the terminal group at the C terminus extremity is a carboxylic acid instead of an amide).<sup>75,182</sup> Moreover, the evolution of their  $\Omega$  as a function of the mass is also very similar. When both peptoid and peptide curves are fitted using the equation  $\Omega = A.M^B$ , the resulting  $B$  parameter equals 0.615 and 0.671/0.680 for the peptoids and peptides, respectively, strongly suggesting that their conformation can be associated to a globule. Hudgins *et al.* demonstrated by molecular modeling that

<sup>e</sup>The collision cross sections from Bush *et al.* are available at <https://biophysicalms.org/ccsdatabase/> (Figure 5.11A). The collision cross sections were obtained on a Drift-Tube instrument using He as buffer gas.<sup>75,182</sup>

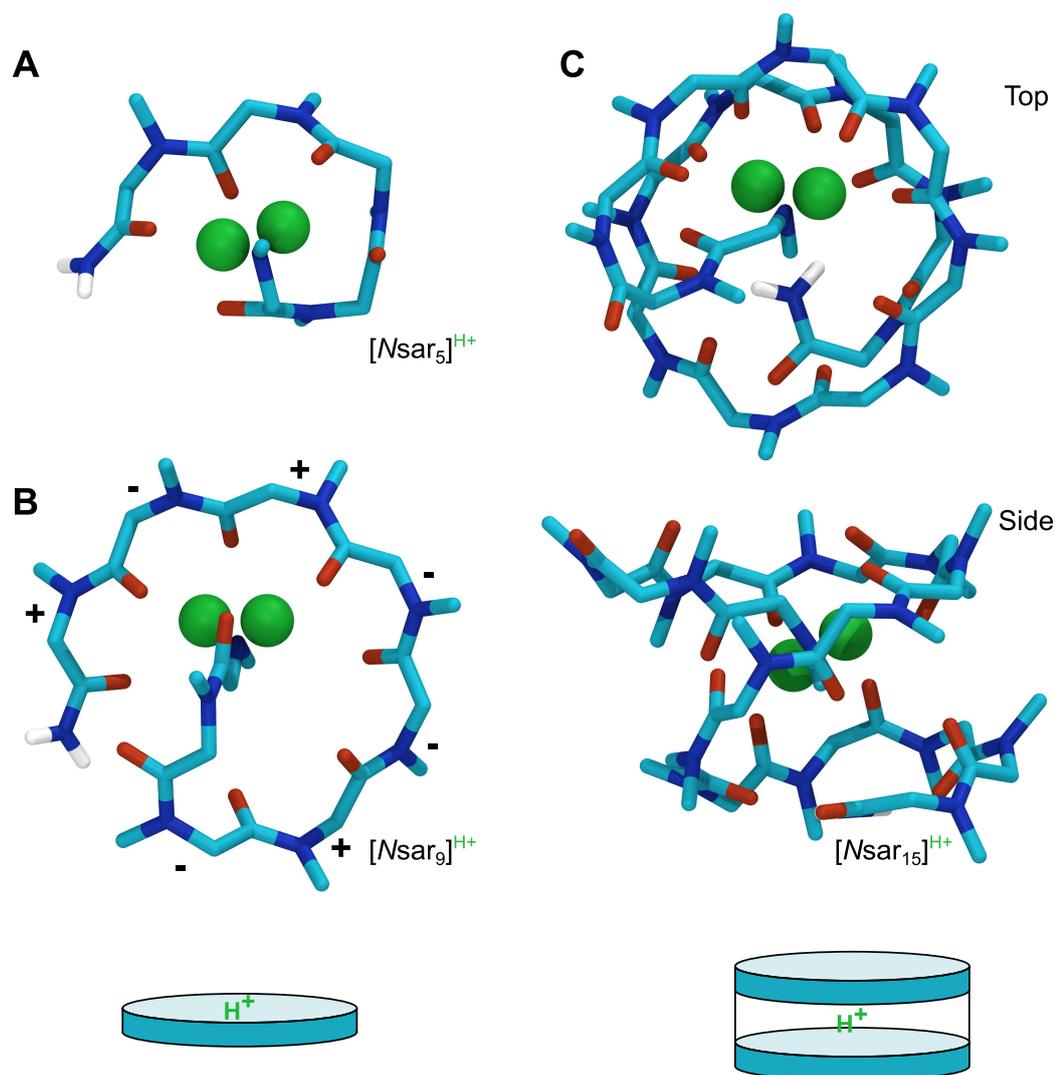
the gas phase conformation of poly(L-alanine) peptides is indeed globular (**Figure 5.11B**). For the rest of this chapter, the hydrogen atoms will only be displayed when carried by a nitrogen atom for clarity. The proton will be represented as a green atom bead. However, since the proton binds to an amine that already carries a hydrogen, the added proton is not distinguishable, and are therefore both represented as green beads.



**Figure 5.11:** (A) Evolution of the collision cross section  $\Omega$  as a function of the mass of  $N$ sar peptoids and L-alanine peptides.<sup>75,182</sup> Black (plain and dashed) curves correspond to the fitting of the data using Equation 5.1. (B) Lowest energy conformation of singly protonated (green dashed circle) L-alanine (20 units) obtained by molecular dynamics.<sup>75</sup> Adapted from Ref. 75. Copyright © 1999 The Biophysical Society. Published by Elsevier Inc. All rights reserved.

Although the general trends are very similar, the  $B$  value for peptoids is smaller than the reference  $2/3$  of globular ions, suggesting that the gas phase conformation of peptoids may be different and more compact than a globule. However, as discussed above, trend line analysis should be better considered as qualitative. We thus performed conformational sampling by quenched dynamics on the whole range of chain length to shed light on the conformation of gaseous peptoids.

For the range of 3 to 9 residues, the lowest energy conformers are folded into loop-like conformations, where all carbonyls from the amide point toward the center of the loop where the proton is located (**Figure 5.12A,B**). In these loops, the first amide residue (from N terminus) is systematically in *cis*, while every other amides are in *trans* such as in the L-alanine case.



**Figure 5.12:** Lowest energy conformers (protonated at the N terminus) obtained after several quenched molecular dynamics for (A)  $Nsar_5$ , (B)  $Nsar_9$  and (C)  $Nsar_{15}$ . (A) and (B) are in a loop-like conformation, while (C) forms two superposing loops, with the proton located in-between, as represented by the schematics. The sign of the  $\phi$  dihedrals are displayed on the  $Nsar_9$ .

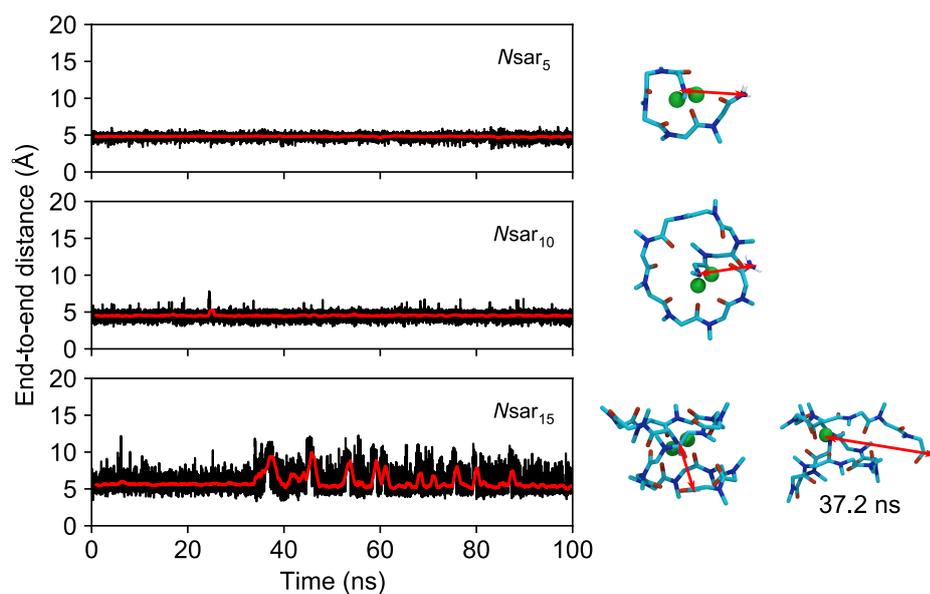
The absence of hydrogen bonds in the peptoid backbone grants a higher flexibility and thus allows the backbone to fully “wrap around” the charge. The wrapping is also possible thanks to the alternance of the sign of the dihedral  $\phi$ , which varies around  $\pm 80^\circ$  as indicated by the “+” and “-” signs in **Figure 5.11B**, while the dihedral  $\psi$  remains around  $180^\circ$ . Interestingly, these dihedral combinations correspond to the low-lying energy conformers from the Ramachandran-like plots (**Figure 4.4D**).<sup>30</sup> This type of conformation is strongly reminiscent of the “*threaded loop*” conformation, exclusively reported for peptoid nonamers in solution using polar aprotic solvents.<sup>93,183,184</sup>

As observed in the  $Nsar_9$ , the loop is ‘closed’, *i.e.*, the C terminus extremity is close to the N terminus extremity. When additional monomer units are added, the peptoid backbone starts forming a second loop superimposed to the one previously discussed, where the protonated amine is shared between the two loops (**Figure 5.12C**). The amide bonds are in *trans* conformation as well, except for the second residue (starting from N terminus) that is in *cis* conformation, which allows the orientation of the charge inside the loops. We then assessed the pertinence of these conformations by performing MD simulations for 50 ns on these lowest-energy conformers and computed their collision cross sections. The theoretical  $\Omega$  nicely agree with the experimental values (**Table 5.2**), which allows us to attribute the gaseous conformation of  $Nsar$  peptoids to a loop-like shape.

**Table 5.2:** Comparison of the experimental ( $^{TW}\Omega_{N_2 \rightarrow He}$ ) and theoretical ( $^{TM}\Omega_{He}$ ) collision cross sections of  $Nsar$  peptoids. The theoretical values correspond to the average of 500 values obtained from MD simulations at 300 K for 50 ns.

	$^{TW}\Omega_{N_2 \rightarrow He} (\text{\AA}^2)$	$^{TM}\Omega_{He} (\text{\AA}^2)$
$Nsar_3$	95	91
$Nsar_4$	112	108
$Nsar_5$	127	119
$Nsar_6$	138	135
$Nsar_7$	150	156
$Nsar_8$	164	170
$Nsar_9$	176	182
$Nsar_{10}$	187	191
$Nsar_{11}$	198	206
$Nsar_{12}$	210	218
$Nsar_{13}$	221	231
$Nsar_{14}$	232	237
$Nsar_{15}$	243	245

As stated above, since the experimental ATDs are monomodal, we are either in presence of rapidly interconverting conformers or in presence of a single conformer family. To probe further those possibilities, we extended the simulation time to 100 ns for three distinct peptoid chain length:  $Nsar_5$ ,  $Nsar_{10}$  and  $Nsar_{15}$ , and measured the end-to-end distance over time, *i.e.*, the distance between nitrogen atoms at the N and C termini (**Figure 5.13**). Interestingly, the end-to-end distances are very alike between each chain length ( $\sim 5 \text{\AA}$ ), although some variations are present for the longer  $Nsar_{15}$ . The end-to-end distance sometimes spikes at  $\sim 10 \text{\AA}$ ; this distance corresponds to a slight opening of the loop formed at the C terminus extremity, although it does not produce significant variations of the collision cross section ( $245 \text{\AA}^2$  vs.  $248 \text{\AA}^2$ ).



**Figure 5.13:** Evolution of the end-to-end distance defined between the nitrogen atoms at both N and C termini extremities (black) and the evolution of the average value (red) for three different  $Nsar$  peptoid chain lengths. The average conformation is represented next to the plots.

Based on these results, we can confidently attribute the gaseous conformation of  $Nsar$  peptoids to a loop-like shape in these conditions. Their conformation is thus indeed different from  $L$ -alanine peptides as suggested by the  $B$  parameter, and further evidenced by the detailed structure provided by although molecular modeling.

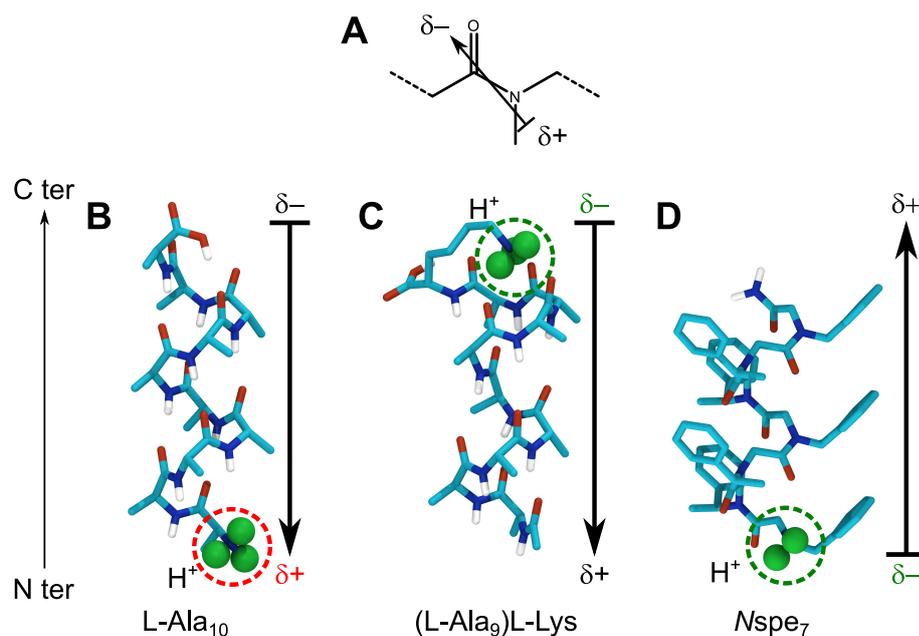
### 5.3.2 Chirality and bulkiness: the ingredients of gas phase helical folding?

In the Introduction, we discussed about the influence of the side chain on the peptoid secondary structure in solution. Several reports demonstrate that chiral bulky side chains are required (*i*) to induce a preference for the *cis* or *trans* amide conformation and (*ii*) to obtain a “enantiomerically pure” secondary structure, for example, right-handed helices.<sup>41,42,48,91,185</sup> The (*S*)-*N*-1-phenylethyl side chain (*N*<sub>spe</sub>) is one of the most studied chiral bulky side chains and was shown to induce the folding of the peptoid backbone into a right-handed helix that is similar to the poly(proline) type I secondary structure.<sup>41</sup>

Until now, every studies about *N*<sub>spe</sub> peptoids were carried out in solution. In this work, we are interested in the assessment of their secondary structure in gas phase and especially whether the helical conformation is conserved. This is a reasonable hypothesis considering that helical peptides have already been identified in gas phase by IM-MS.<sup>115,186–188</sup> In the previous section on *N*<sub>sar</sub> peptoids, we discussed about the gas phase conformation of L-alanine peptides that is globular although its solution phase conformation is an  $\alpha$ -helix. The helical disruption was attributed to the charge carried by the amine at the N terminus that interacts with the helix macrodipole.<sup>115,189</sup> Indeed, every amide carbonyl moiety constituting the  $\alpha$ -helix are oriented from the N to the C terminus, and their individual dipoles add up to form a macrodipole along the helical axis whose positive end points toward the N terminus (**Figure 5.14B**).

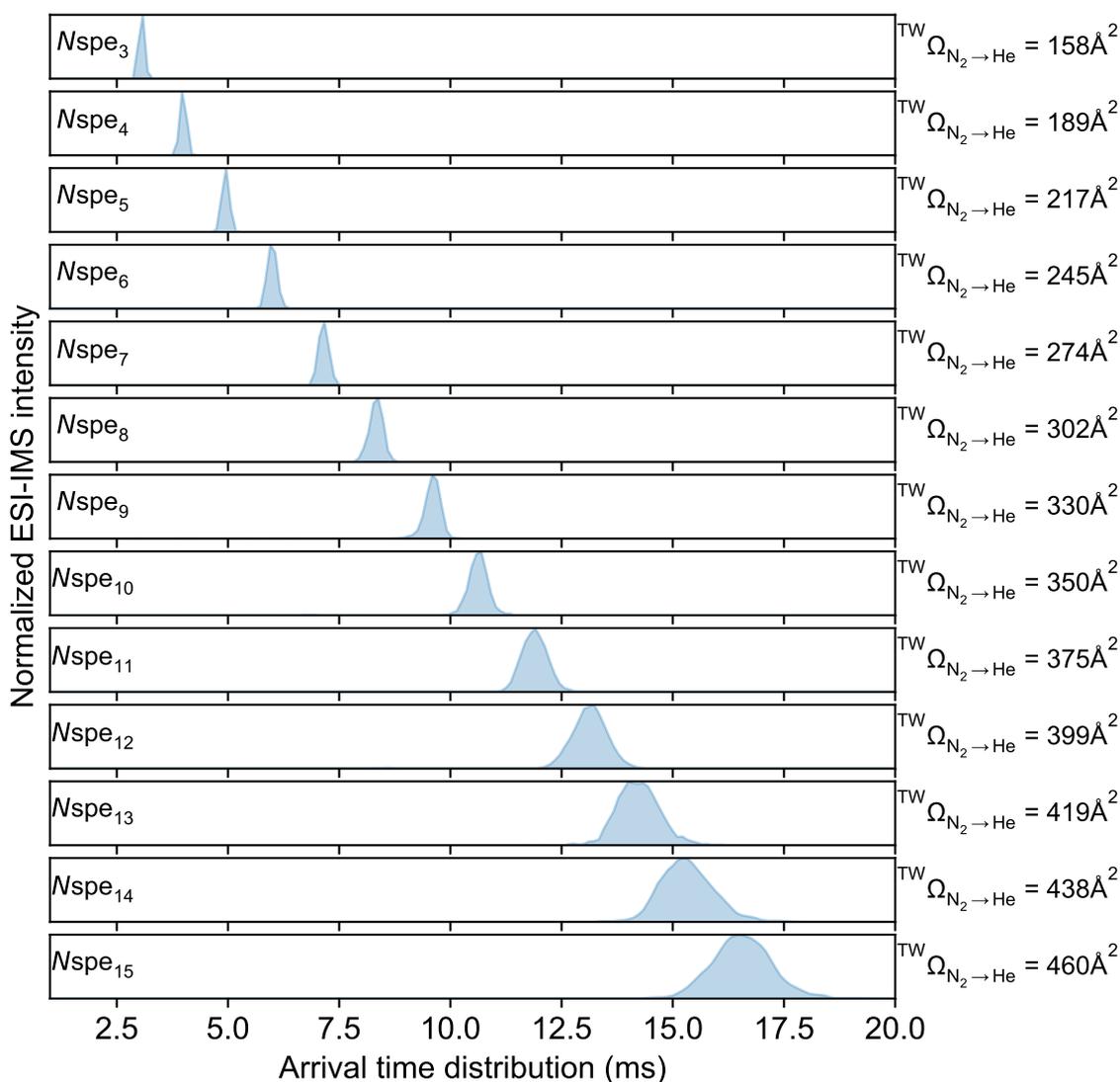
Consequently, a repulsion between the positive end of the macrodipole and the positive charge destabilizes the helix, which folds into a globular shape. Hudgins *et al.* elegantly tackled this issue by changing the position of the charge from the N to the C terminus, *i.e.*, at the negative end of the macrodipole.<sup>115</sup> To do so, they added a lysine residue at the C terminus, which carries an amine moiety that can be protonated, and got rid of the amine at the N terminus by acetylating it. In this case, the negative end of the macrodipole now interacts favorably with the positive charge, which results in the conservation of the  $\alpha$ -helix in gas phase. This example clearly emphasizes the importance of understanding the relationship between the primary structure and the secondary structure, and that this relationship may vary depending on the environment, either in solution or gas phase.

Concerning peptoids, the amide carbonyl moieties in the helix formed by *N*<sub>spe</sub> peptoids are actually oriented from the C to the N terminus (opposite to the peptide  $\alpha$ -helix) and the negative end of macrodipole is therefore already well-oriented toward the N terminus, *i.e.*, toward the charge, which *a priori* creates a favorable interaction (**Figure 5.14D**).



**Figure 5.14:** (A) Dipole orientation in amide residue. (B)  $\alpha$ -helix protonated (L-alanine<sub>10</sub>) at the N terminus. The macrodipole is formed by the individual amide dipoles oriented from N to C terminus. The  $\delta^+$  end of the macrodipole is located at the protonated site which creates a repulsion destabilizing the helix (red dashed circle). (C)  $\alpha$ -helix protonated at the lysine located at the C terminus extremity (L-alanine<sub>9</sub>L-lysine). The negative end of the macrodipole now interacts favorably with the charge located at the C terminus (green dashed circle). (D) Peptoid helix (*Nspe*<sub>7</sub>) whose macrodipole is reversed compared to  $\alpha$ -helix.

*Nspe* peptoids were also synthesized from 3 to 15 units (*Nsar*<sub>3-15</sub>). After ion mobility separation, each *Nspe* peptoid ion is characterized by a monomodal ATD, as for *Nsar* peptoids, again implying that the ions are rapidly interconverting between multiple conformer families or that they belong to a single conformer family (Figure 5.15). We demonstrated that the *B* parameter can be misleading for ion shapes that differ from a spherical shape but it still can provide qualitative information. We thus investigate the evolution of their collision cross sections as a function of the mass by fitting the data using Equation 5.1 and obtain a *B* parameter of 0.685, which is slightly larger than the reference of  $2/3$ , but also much lower than what would be expected for a growing helix (see Figure 5.3C). For an upper mass limit of  $\sim 2500$  (*Nspe*<sub>15</sub>), the *B* parameter for a perfectly right-handed peptoid helix is about 0.78. This value might vary if the helix slightly rearranges (stretches or shrinks along the helical axis), but will always be larger than  $2/3$ . This first result qualitatively suggests that the gas phase conformation of *Nspe* peptoid ions might be different from a globule but also from a helix.



**Figure 5.15:** Arrival time distributions for each  $N_{spe}$  peptoid chain length (light blue).

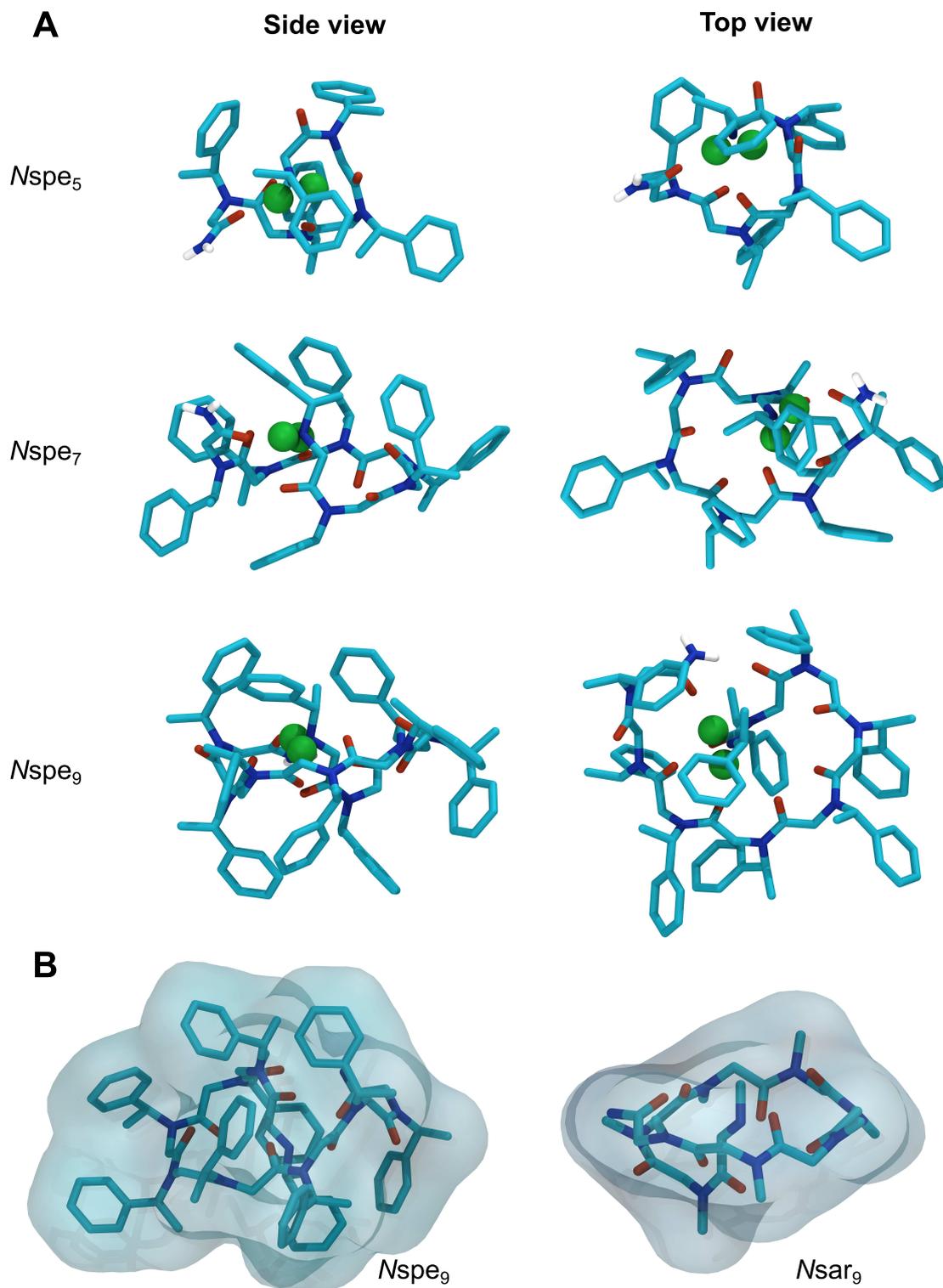
To rationalize this result, we first performed a conformational search on singly protonated  $N_{spe}$  peptoids containing from 3 to 9 residues. Against all odds, the low-lying energy conformers are all in a loop-like conformation with the amide carbonyls pointing toward the ammonium at the center, very similar to that adopted by  $N_{sar}$  peptoids. The major difference lies in the sequence of amide bond conformations: while only a single residue is in the *cis* amide conformation in  $N_{sar}$  peptoids (first or second residue), the first two residues of  $N_{spe}$  peptoids are systematically in *cis* conformation but (most of) the remaining residues are in *trans* (Table 5.3). The proportion of *cis* amide should increase when going from a methyl side chain to a (*S*)-1-phenylethyl side chain thanks to the favorable “bridged  $n \rightarrow \pi^*$ ” interaction.<sup>78</sup> The slight preference for the *cis* conformation is observed in most solvents, but also *in vacuo*, as attested by the amide rotational barrier profile (the *trans* conformation is

1.3 kcal mol<sup>-1</sup> higher than the *cis*, see **Figure 4.3**). This preference is likely affected by the ammonium, whose stabilization is increased when the amide carbonyls point toward it, especially in the *trans* amide conformation because of favorable interactions between the amide dipoles and the charge.

**Table 5.3:** Sequences of amide conformation for the lowest energy conformers of singly protonated *Nspe* peptoids. “C” stands for *cis* and “T” for *trans*. The letters in purple correspond to the helical domain formed after the loop-like shape.

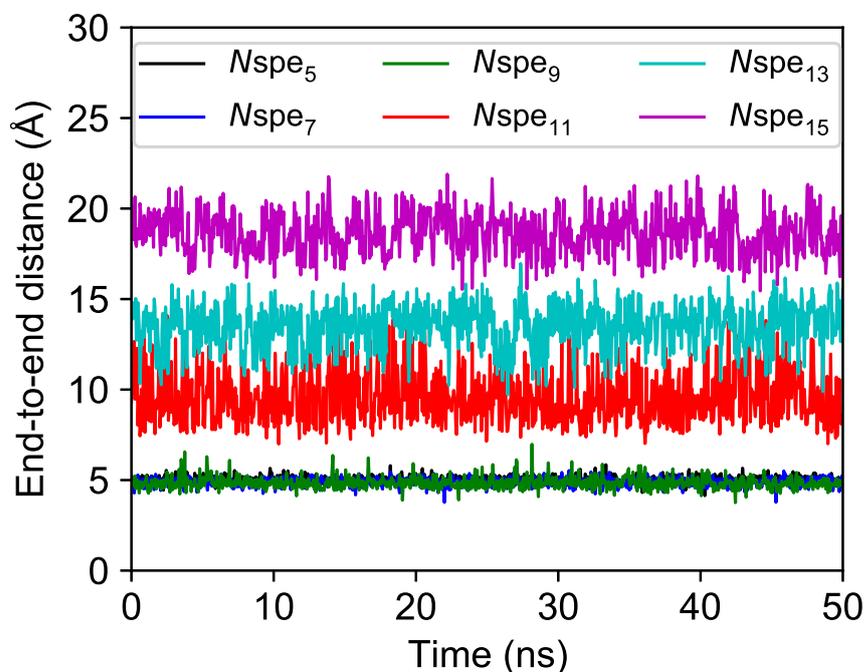
Peptoid	Residue													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
<i>Nspe</i> <sub>3</sub>	C	T												
<i>Nspe</i> <sub>4</sub>	C	C	T											
<i>Nspe</i> <sub>5</sub>	C	C	T	C										
<i>Nspe</i> <sub>6</sub>	C	C	T	T	T									
<i>Nspe</i> <sub>7</sub>	C	C	T	T	T	T								
<i>Nspe</i> <sub>8</sub>	C	C	T	T	T	T	T							
<i>Nspe</i> <sub>9</sub>	C	C	T	T	T	T	T	C						
<i>Nspe</i> <sub>10</sub>	C	C	T	T	T	T	T	C	C					
<i>Nspe</i> <sub>11</sub>	C	C	T	T	T	T	C	C	C	T				
<i>Nspe</i> <sub>12</sub>	C	C	T	T	T	T	T	T	C	C	C			
<i>Nspe</i> <sub>13</sub>	C	C	T	T	T	T	T	T	C	C	C	C		
<i>Nspe</i> <sub>14</sub>	C	C	T	T	T	T	C	C	C	C	C	C	T	
<i>Nspe</i> <sub>15</sub>	C	C	T	T	T	T	T	T	T	T	T	C	C	C
<i>Nspe</i> <sub>15</sub>	C	C	T	T	T	T	C	C	C	C	C	C	C	C

The most stable conformations of selected chain length issued from the quenched dynamics are represented in **Figure 5.16**.



**Figure 5.16:** (A) Lowest energy conformers of *Nspe*<sub>5</sub>, *Nspe*<sub>7</sub> and *Nspe*<sub>9</sub>, viewed from the side and top. (B) Molecular surface of *Nspe*<sub>9</sub> and *Nsar*<sub>9</sub>. The total shape of *Nspe*<sub>9</sub> is more globular than that of *Nsar*<sub>9</sub>.

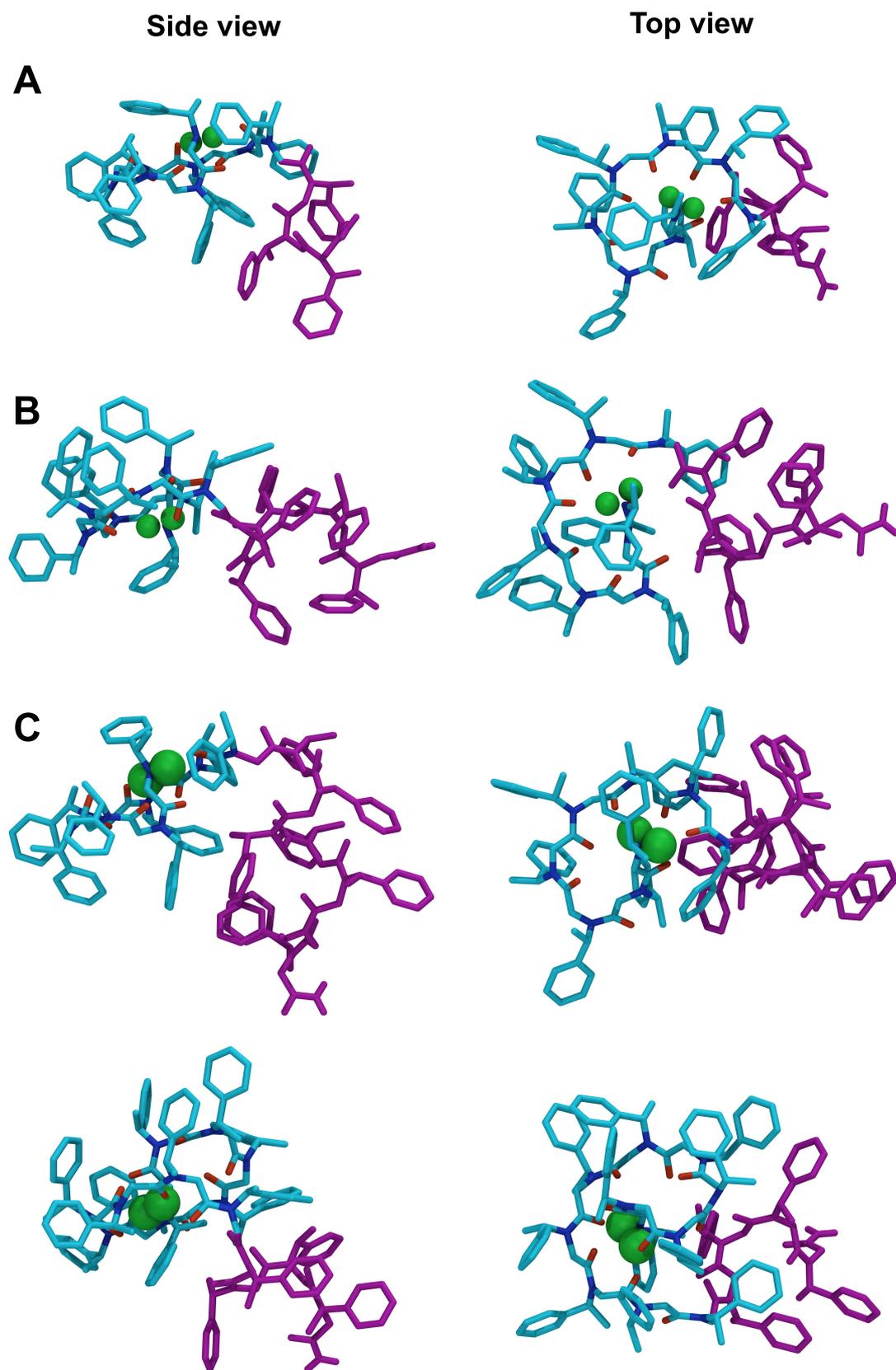
As observed with  $Nsar$  peptoids, the C terminus extremity is systematically very close to the N terminus extremity ( $\sim 5 \text{ \AA}$ ) thanks to the formation of a hydrogen bond. Consequently, the end-to-end distance does not vary much along a 50 ns MD run (**Figure 5.17**).



**Figure 5.17:** Evolution of the end-to-end distance defined between the nitrogen atoms at both N and C terminus extremities for six different  $Nspe$  peptoid chain length. The shorter chains ( $Nspe_5$ ,  $Nspe_7$  and  $Nspe_9$ ) are characterized by roughly the same end-to-end distance ( $\sim 5 \text{ \AA}$ ), while the longer chains display an increase in their end-to-end distance due to the formation of the helical domain outside the loop.

We must precise that, when we state that  $Nspe$  peptoids are in a loop-like conformation, we actually mean that the backbone is in a loop-like conformation. Indeed, if we take a closer look to  $Nspe_9$  (**Figure 5.16B**), it is clear that the “total” shape of the ion is closer to a globule than a loop. This result is consistent with the difference in the  $B$  parameter between  $Nsar$  and  $Nspe$ , which indicates that  $Nspe$  peptoids are less compact than the  $Nsar$  peptoids. As a side note, the  $Nspe_9$  conformer is once again very similar to the reported “*threaded loop*”.

For the second range of chain length, *i.e.*, from 10 to 15 residues, the low-lying energy conformers are even more surprising. While we obtained a double-loop shape with longer  $Nsar$  peptoids, longer  $Nspe$  peptoids are composed of two distinct domains: (*i*) a loop-like domain that stabilizes the ammonium in its center and (*ii*) a right-handed helix protruding out of the loop (**Figure 5.18**).



**Figure 5.18:** Lowest energy conformers of (A) *Nspe*<sub>11</sub>, (B) *Nspe*<sub>13</sub> and (C) *Nspe*<sub>15</sub>, viewed from the side and top. Two low-lying energy conformers are represented for *Nspe*<sub>15</sub>; the top one is 0.17 kcal mol<sup>-1</sup> higher in energy than the bottom one. The helical domain is highlighted in purple.

After about 9 residues, the charge thus appears to be fully stabilized. Unlike in *Nsar* peptoids, the additional residues could not form a second loop on top of the first one due to the steric hindrance caused by the *Nspe* side chains. The additional residues rather organize as a right-handed helix ( $\varphi \sim -80^\circ$ ,  $\psi \sim \pm 170^\circ$ ) with every amide in *cis* conformation, which corresponds to the *cis* <sub>$\alpha$ D</sub>-conformation (see **Figure 1.3A**). All amide carbonyls from the helical domain point to the ammonium at the center of the loop, which creates a favorable interaction between the negative end of the macrodipole and the positive charge. Given that the degrees of freedom dramatically increase with the number of residues, it is very likely that these conformers are part of a larger family with the same global shape, *i.e.*, a loop and a helical domain, with slightly different side chain orientations or amide sequences that do not significantly affect the general conformation. This hypothesis is in agreement with the broadening of the ATDs with the growing number of residues. For example, we obtain two low-lying energy conformers for *Nspe*<sub>15</sub> (difference of 0.17 kcal mol<sup>-1</sup>) whose amide sequence is slightly different in the helical domains, but leads to the same collision cross sections (**Figure 5.18C**).

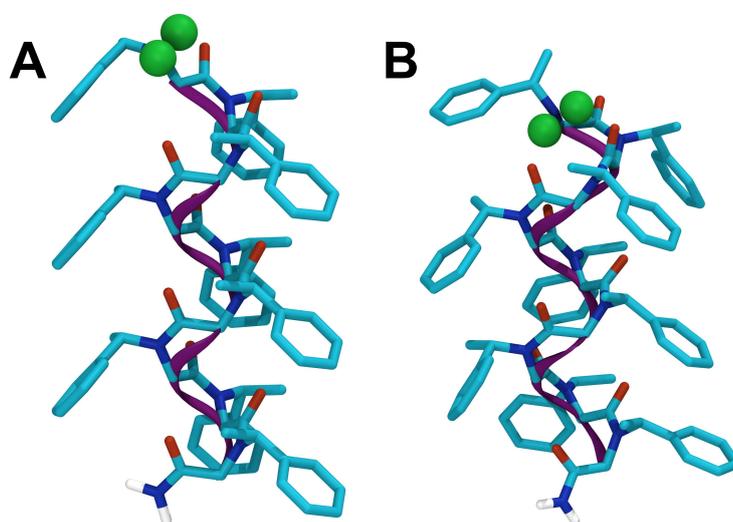
**Table 5.4:** Comparison of the experimental ( $^{TW}\Omega_{N_2 \rightarrow He}$ ) and theoretical ( $^{TM}\Omega_{He}$ ) collision cross sections of *Nspe* peptoids. The theoretical values correspond to the average of 500 values obtained during a 50 ns MD at 300 K starting from the lowest energy conformer or the right-handed helix. Their relative energy is also displayed.

Peptoid	$^{TW}\Omega_{N_2 \rightarrow He}$ (Å <sup>2</sup> )	$^{TM}\Omega_{He}$ (Å <sup>2</sup> ) (most stable conformer)	$^{TM}\Omega_{He}$ (Å <sup>2</sup> ) (relaxed helix)	Relative energy (kcal mol <sup>-1</sup> )
<i>Nspe</i> <sub>3</sub>	158	163	168	6.4
<i>Nspe</i> <sub>4</sub>	189	196	203	7.0
<i>Nspe</i> <sub>5</sub>	217	219	233	12.1
<i>Nspe</i> <sub>6</sub>	245	246	250	13.8
<i>Nspe</i> <sub>7</sub>	274	274	292	18.0
<i>Nspe</i> <sub>8</sub>	302	299	318	22.6
<i>Nspe</i> <sub>9</sub>	330	332	342	24.0
<i>Nspe</i> <sub>10</sub>	350	360	370	27.0
<i>Nspe</i> <sub>11</sub>	375	382	395	22.8
<i>Nspe</i> <sub>12</sub>	399	418	421	26.4
<i>Nspe</i> <sub>13</sub>	419	428	446	25.5
<i>Nspe</i> <sub>14</sub>	438	441	465	17.2
<i>Nspe</i> <sub>15</sub>	460	480	502	26.0

The theoretical collision cross sections associated to these conformers give us confidence about the validity of these results, although it is surprising that the *Nspe* peptoids are not helical in gas phase (**Table 5.4**). Given that the macrodipole of the helix is well-oriented, it means that the backbone is not enough stabilized by secondary interactions to avoid the disruption of the helix in the presence of a charge. Nonetheless, we built protonated right-handed helices ( $\varphi = -80^\circ$ ,  $\psi = 165^\circ$  and  $\omega =$

0°), performed a geometry optimization and subjected them to MD simulations<sup>f</sup> to obtain their collision cross sections and check whether it is possible that helices might be present in gas phase.

During the MD, the helices relax and are slightly compressed along the helical axis compared to the perfect helices described in the first part of this chapter (**Figure 5.19**), which results in smaller theoretical collision cross sections than the rigid helices discussed at the beginning of the chapter. Although close to the  $\Omega$  of the loop-like conformers, the  $\Omega$  of the relaxed helices are significantly larger. Moreover, the relative energy of the helices is tremendously higher (up to  $\sim 26$  kcal mol<sup>-1</sup>), which rules them out of the discussion.



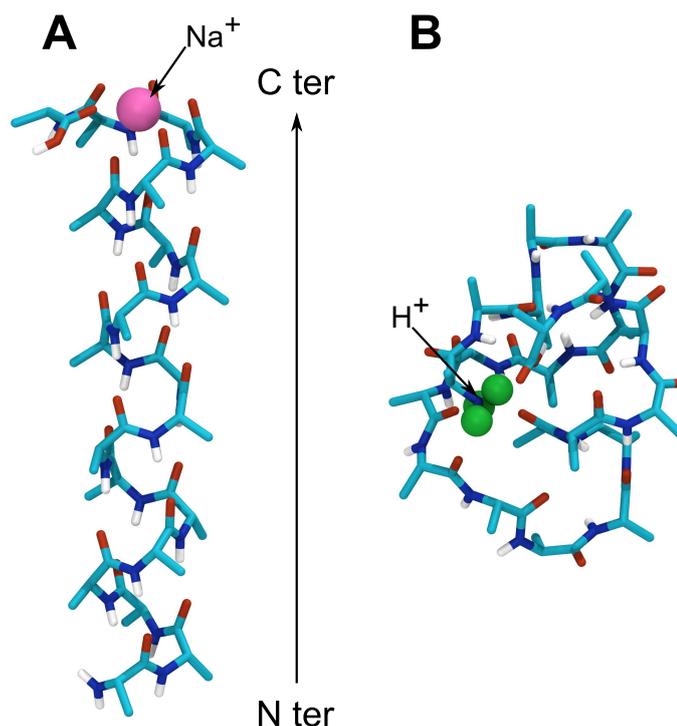
**Figure 5.19:** (A) Ideal and (B) relaxed protonated peptoid helix made of 9 *Nspe* residues. The relaxed helix is obtained by MD simulation at 300 K and is slightly compressed along the helical axis. The helical shape is highlighted by the purple ribbon.

According to these results, we face the same issue as protonated poly(L-alanine) peptides: *Nspe* peptoids do not seem to retain their helical shape in gas phase, although it is the major conformer in solution.<sup>47</sup> Moreover, despite the favorable interaction between the macrodipole and the ammonium, the non-covalent secondary interactions are not sufficient to compensate the need to stabilize the charge.

Interestingly,  $\alpha$ -helical poly(L-alanine) peptides can still be obtained in gas phase without performing a modification of their sequence. Kohtani *et al.* showed, by comparing experimental and theoretical collision cross sections, that using a metal cation as charge carrier instead of adding a proton to the amine promotes the conservation of the helix.<sup>190</sup> They obtained the best results using sodium cation (Na<sup>+</sup>). Compared

<sup>f</sup>The MD is carried out at 298 K, which prevents the amide bonds to isomerize in *trans*.

to the proton,  $\text{Na}^+$  is not strictly bound and can theoretically be located “anywhere” (at least in solution). When the solvent is evaporated during the Electrospray ionization process, the analyte and  $\text{Na}^+$  form adducts in which the charge needs to be stabilized.<sup>162,191</sup> In the case of poly(L-alanine), the  $\text{Na}^+$  is preferentially located at the C terminus with the amide carbonyls pointing toward it and where the negative end of the  $\alpha$ -helix macrodipole lies, thus leading to the conservation of the helix (**Figure 5.20**).<sup>190</sup>



**Figure 5.20:** (A) Most stable conformer of L-alanine<sub>20</sub> cationized by a single  $\text{Na}^+$  (mauve bead) at the C terminus extremity, obtained by MD simulations.<sup>190</sup> The  $\alpha$ -helical shape is conserved throughout the MD. (B) Low-lying energy conformer of singly protonated L-alanine<sub>20</sub>. The helix is disrupted by the presence of the charge at the N terminus extremity.<sup>75</sup> Adapted with permission from 190. Copyright 2004 American Chemical Society. Adapted from Ref. 75. Copyright © 1999 The Biophysical Society. Published by Elsevier Inc. All rights reserved.

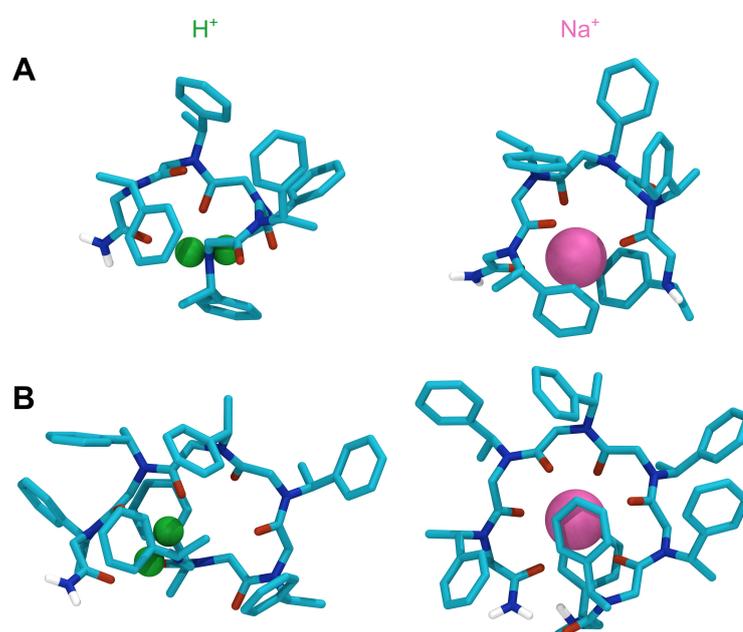
Therefore, we employed the same strategy and measured the collision cross sections of sodiated *Nspe* peptoids (from 4 to 13 residues). Because  $\text{Na}^+$  is not strictly bounded, it may interact with the N terminus extremity without distorting the helix. The difference in  $\Omega$  between protonated and sodiated species is not significant, especially after 8/9 residues (**Table 5.5**). The shorter sodiated *Nspe* peptoids have larger  $\Omega$ , which might be a hint of a more extended conformation.

We performed MD simulations on the sodiated *Nspe* to verify whether the shorter chains would remain helical with the same conformational strategy as previously described. However, since the location of the  $\text{Na}^+$  is not known, we placed the  $\text{Na}^+$  near the backbone at 3 different locations for each starting geometry. The lowest energy

**Table 5.5:** Comparison of the experimental collision cross sections ( ${}^{TW}\Omega_{N_2 \rightarrow He}$ ) of protonated and sodiated *Nspe* peptoids.

Peptoid	${}^{TW}\Omega_{N_2 \rightarrow He}$ ( $\text{\AA}^2$ )	${}^{TW}\Omega_{N_2 \rightarrow He}$ ( $\text{\AA}^2$ )
<i>Nspe</i> <sub>3</sub>	158	/
<i>Nspe</i> <sub>4</sub>	189	194
<i>Nspe</i> <sub>5</sub>	217	226
<i>Nspe</i> <sub>6</sub>	245	257
<i>Nspe</i> <sub>7</sub>	274	280
<i>Nspe</i> <sub>8</sub>	302	302
<i>Nspe</i> <sub>9</sub>	330	328
<i>Nspe</i> <sub>10</sub>	350	347
<i>Nspe</i> <sub>11</sub>	375	376
<i>Nspe</i> <sub>12</sub>	399	497
<i>Nspe</i> <sub>13</sub>	419	417
<i>Nspe</i> <sub>14</sub>	438	/
<i>Nspe</i> <sub>15</sub>	460	/

conformers of the short *Nspe* peptoids are indeed more extended, but are absolutely not helical. They rather also adopt a loop-like conformation (**Figure 5.21**). The increase in  $\Omega$  is related to the release of the constraint in the backbone caused by the proton on the terminal amine. Indeed, the loop in sodiated peptoids is larger which allows the backbone to better wrap around the charge. The difference vanishes with increasing size of the peptoid ion when the charge is fully solvated. As observed with the protonated *Nspe* peptoids, the  $\text{Na}^+$  is fully stabilized after 7 to 9 residues, and by extension, 7 to 9 carbonyls, which corresponds to the reported coordination of  $\text{Na}^+$  by several polymers in IM-MS. [162,176,192,193](#)



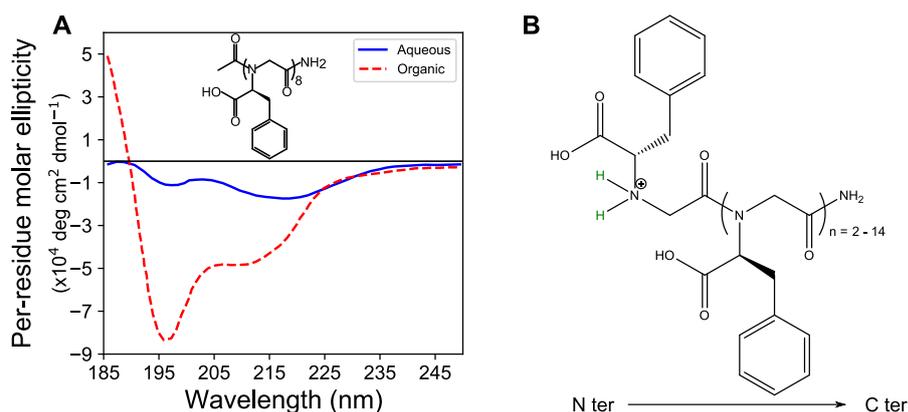
**Figure 5.21:** Most stable conformers of protonated and sodiated (A) *Nspe5* and (B) *Nspe7* (top view). The  $Na^+$  is represented as a mauve bead.

According to these results, we can conclude that  $N_{spe}$  peptoid ions do not retain their helical shape when transferred into the gas phase due to the requirement of the charge to be stabilized. It means that the “rules” developed in solution phase to obtain a peptoid helix are not sufficient, *i.e.*, that the side chain must be bulky and chiral. When the solvent is removed, the non-covalent interactions between the residues are not strong enough to keep the backbone organized in a helical fashion. This means that the gas phase folding of peptoids is governed by the charge solvation effect. We performed a conformational sampling in vacuum on neutral  $N_{spe}$  peptoids having 5, 10 and 15 residues and their lowest energy conformer are all right-handed helices ( $N_{spe_5}$  is only partially helical).

### 5.3.3 Back to basics: inspiration from peptides

From the previous discussion, we identified the key elements that may help the conservation of the helices in gas phase. First, the positive charge needs to be located on the negative end of the helix macrodipole, which will always be the case for a protonated poly(proline) type I helix. Second, strong secondary non-covalent interactions are required to maintain the backbone. In the case of L-alanine peptides, these non-covalent interactions appear in the form of a hydrogen bond network formed between the carbonyl oxygen and hydrogen atoms from the amide of residues  $i$  and  $i + 3$ .

Consequently, we envisaged another side chain that would be chiral, bulky and aromatic such as the *N*spe side chain, but that would also be able to form intramolecular hydrogen bonds. In 2007, Shin and Kirshenbaum reported a study on peptoids bearing (*S*)-*N*-(1-carboxy-2-phenylethyl) side chains (*N*scp).<sup>57</sup> The primary amine used in the synthesis is actually the L-phenylalanine, that is protected in the form of a *tert*-butyl ester and further deprotected when the peptoid chain is cleaved from the resin. In their study, they evaluate the secondary structure by circular dichroism spectroscopy. The CD intensity increases with growing chain length (from 2 to 13 residues), which suggests the presence of stable secondary structures. Moreover, the CD signature varies depending on the solvent and pH. The CD signature is reminiscent of the peptoid helix signature under neutral aqueous conditions, while in acetonitrile, the signature is reminiscent of the “threaded loop” conformation (**Figure 5.22A**).



**Figure 5.22:** (A) Circular dichroism spectra of the N terminus acetylated *N*scp<sub>8</sub> in aqueous (5mM sodium phosphate buffer, pH 7, plain blue line) and organic solvent (100% acetonitrile, dashed red curve) at 25 °C. The signature at pH < 5 is similar to the signature in organic solvent, while at pH > 5 it is similar to the signature in aqueous solvent. Adapted with permission from 57. Copyright 2007 American Chemical Society. (B) Primary structure of the *N*scp peptoids used in this study. The terminal amine is protonated to match the requirement for IM-MS analyses.

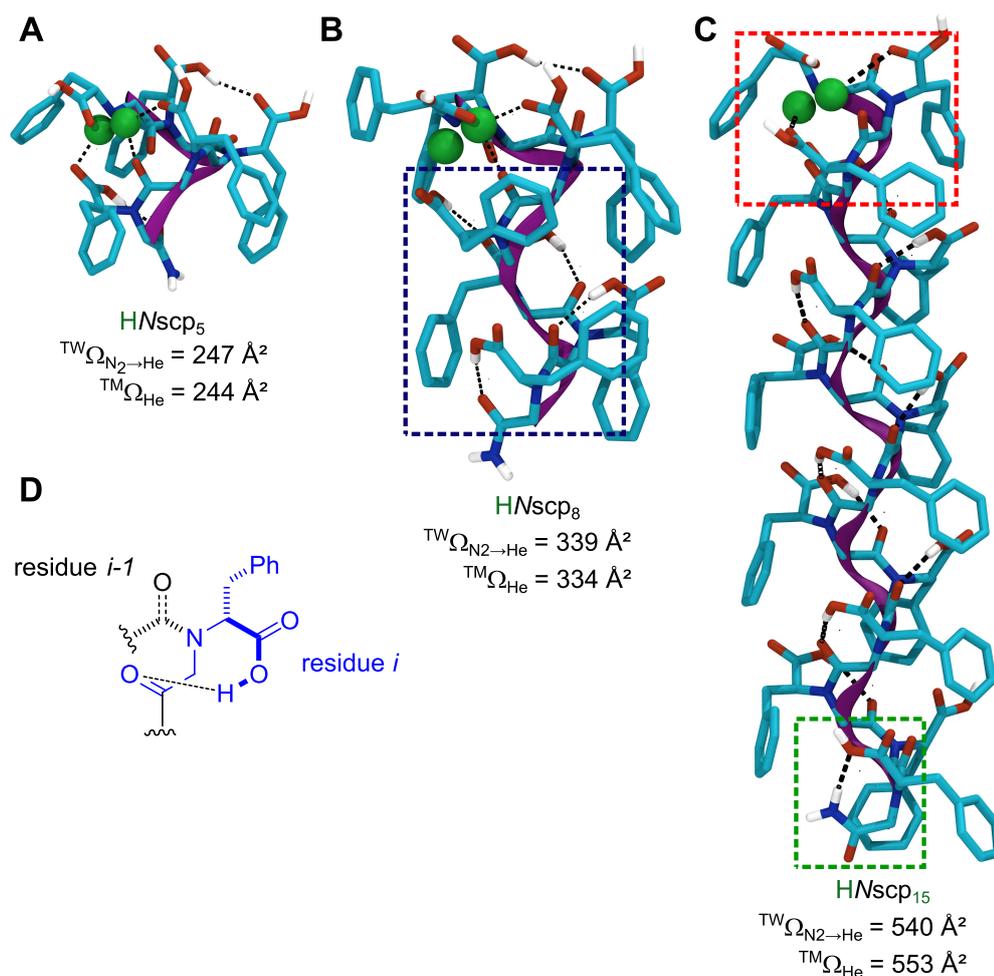
The CD signature also varies with the pH. It is believed that at  $\text{pH} > 5$ , the conformers are more extended because the carboxylic acid moieties turn into carboxylate and that electrostatic repulsion occurs. At  $\text{pH} < 5$ , the conformers are believed to be more compact. However, there are currently no other evidence about their solution phase conformation. Molecular modeling and especially Time-Dependent DFT calculations would help to shed light on the exact shapes that adopt these molecules.

We thus decided to investigate this  $N_{\text{scp}}$  side chain in the gas phase (**Figure 5.22B**). Compared to the previous sections, we first carried out the calculations to predict their gas phase conformation prior putting efforts into the synthesis. Prior to any conformational sampling, the  $N_{\text{scp}}$  side chain was reparametrized in our PEPDROID force field using the same methodology as presented in Chapter 4. The dihedral profiles can be found in Appendix B.

After the validation of the energy profiles, we carried out conformational sampling on singly protonated  $N_{\text{scp}}$  peptoids with 3 to 15 residues. Since our IM-MS analyses are performed in positive mode, the carboxylic acid moieties will remain neutral and will not turn into carboxylates. Therefore, we built the peptoids with their intact carboxylic acid moieties.

The conformation adopted by the shorter oligomers, *i.e.*,  $N_{\text{scp}_{3 \text{ to } 7}}$ , are similar to those of the short  $N_{\text{spe}}$  peptoid ions. The backbone is wrapped around the ammonium at the N terminus, with the amide carbonyls pointing toward it. The carbonyls from the carboxylic acid moieties interact with the charge at long distance, because of the constraints imposed by the backbone (**Figure 5.23A**).

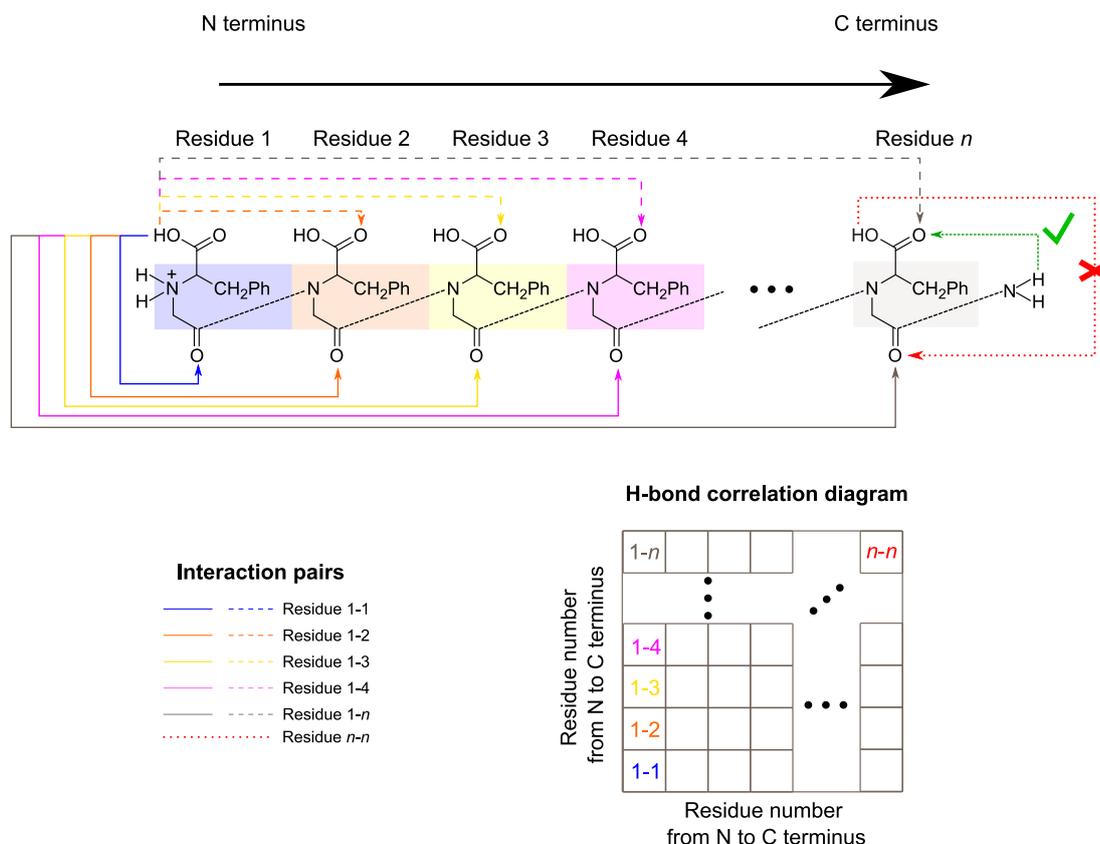
In contrast, the lowest energy conformers of the longer peptoids, *i.e.*,  $N_{\text{scp}_{8 \text{ to } 15}}$ , are more extended, with the appearance of a right-handed helix organization toward the C terminus extremity (**Figure 5.23B-C**). Although the first three/four residues lying at the N terminus extremity of the peptoid ion are in strong interaction with the ammonium group, the remaining residues get organized into a right-handed helix-like structure thanks to the creation of an intra-residue hydrogen bond network, as shown in **Figure 5.23C** for  $N_{\text{scp}_{15}}$ .



**Figure 5.23:** Most stable structures of different  $N_{scp}$  peptoid ions obtained after conformational sampling. Hydrogen bonds are represented by black dashed lines. For very short chains (A), the backbone almost completely wraps around the proton, while for longer chains, (B) and (C), the first residues close to the terminal ammonium are involved in the stabilization of the proton while the remaining part is organized in a helical fashion through a hydrogen bond network (blue dashed box). (D) Schematic of the intra-residue hydrogen bond pattern.

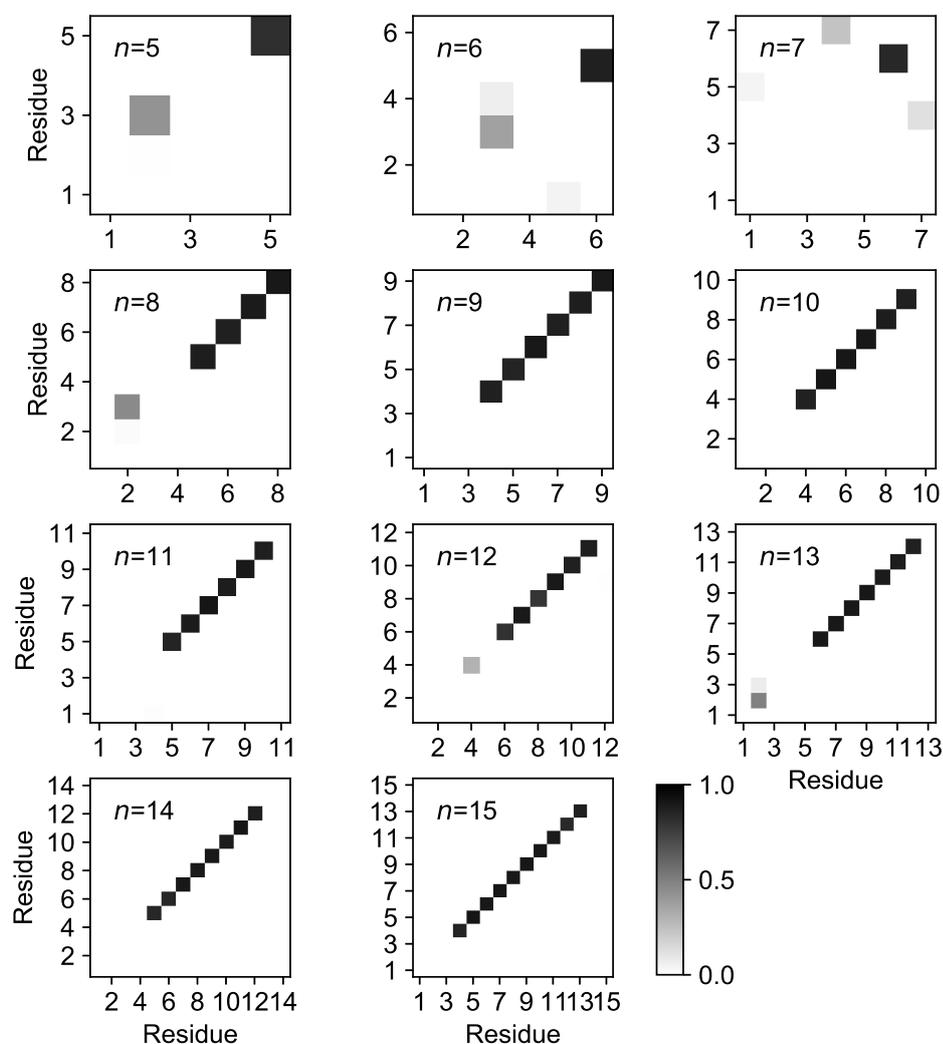
The hydrogen bond network is similar to the  $\alpha$ -helix hydrogen bond network, except that in this case, the hydrogen bond is formed inside the same residue and not between the residues in  $i$  and  $i + 3$  (Figure 5.23D). Since these conformers result from a geometry optimization, the hydrogen bond network may be disrupted when subjected to thermal fluctuations. To investigate the robustness of these conformers, we performed a first equilibration MD for 10 ns at 298 K and then a production MD at the same temperature for 50 ns. During the MD simulations, we assume that, for a given residue, the carboxylic acid group, acting as the hydrogen bond donor  $D$ , can be H-bonded to an acceptor  $A$ , which can be either an oxygen atom of the carbonyl amide (in the same residue or within another residue) or to the carboxylic acid moiety of another residue. For each donor/acceptor pair (*i.e.*, residue 1-1, residue 1-2, residue 1-3, etc.), we measure the key geometrical parameters of the hydrogen bonds

that are: (i) the distance between the carboxylic acid oxygen bearing the hydrogen atom and the amide oxygen atom or the closest oxygen from another carboxylic acid moiety, and (ii) the H-bond atomic triad angle. We consider a hydrogen bond to be effectively formed - by assigning a value of 1 - if the distance is lower than 3.5 Å and if the angle is larger than 150°; otherwise a value of 0 is attributed to the considered H-bond donor/acceptor pair.<sup>194</sup> For each donor/acceptor pair, this measurement is done in every MD snapshot (saved every 0.01 ns) and summed (leading to a H-bond correlation matrix of  $[n \times n]$ , with  $n$  the total number of residues). An illustration of the procedure is represented in **Figure 5.24**.



**Figure 5.24:** Primary structure of the peptoid chain bearing  $N_{scp}$  side chains. Each residue is highlighted using a unique color. The hydrogen bond donor used in the H-bond correlation diagram is always the hydroxyl moiety from the carboxylic acid side chains, while the possible acceptors are the oxygen from the carbonyl amide and the oxygen from the carbonyl of the carboxylic acid. For a given donor, *i.e.*, Residue 1, the H-bond is monitored in each residue pair. For residue pairs on the diagonal, *i.e.*, Residue 1-1, a H-bond can only be formed between the hydroxyl and the amide oxygen (plain color line). For every other off-diagonal pairs, a maximum of one H-bond can be formed, either with the amide oxygen or the carboxylic acid oxygen (dashed color line). In general, the last diagonal term (Residue  $n$ - $n$ ) does not form H-bond based on the donor/acceptor criteria we defined (red dotted line). It rather forms a H-bond between the hydrogen from the amide at the C terminus (acting now as the donor) and the carbonyl of the carboxylic acid (acting as the acceptor, green dotted line). However, this type of H-bond is not accounted for in the H-bond correlation diagrams.

In **Figure 5.25**, we report these values for every H-bond donor/acceptor pairs averaged over the MD simulations, *i.e.*, by dividing the total number computed for each donor-acceptor pair by the total number of generated structures. Such a representation allows to quickly visualize whether H-bonds are present as well as which residues are involved.



**Figure 5.25:** Hydrogen bond correlation matrices for  $N_{scp}$  peptoid ions ranging from 5 to 15 residues along 50 ns MD simulations. The X- and Y-axes correspond to the residue number starting from the N terminus. The hydrogen bonds are monitored simultaneously between the hydrogen of the carboxylic acid moieties and the oxygen from (i) the carbonyls of the backbone and (ii) the carbonyls of the carboxylic functions. Black squares represent robust H-bond along the MD, while lighter colors indicate that the H-bonds are more labile. The diagonal pattern observed for oligomers longer than 8-9 units indicates the conservation of the intra-molecular hydrogen bonds with the carbonyl group from the amide of the same residue over the whole simulation.

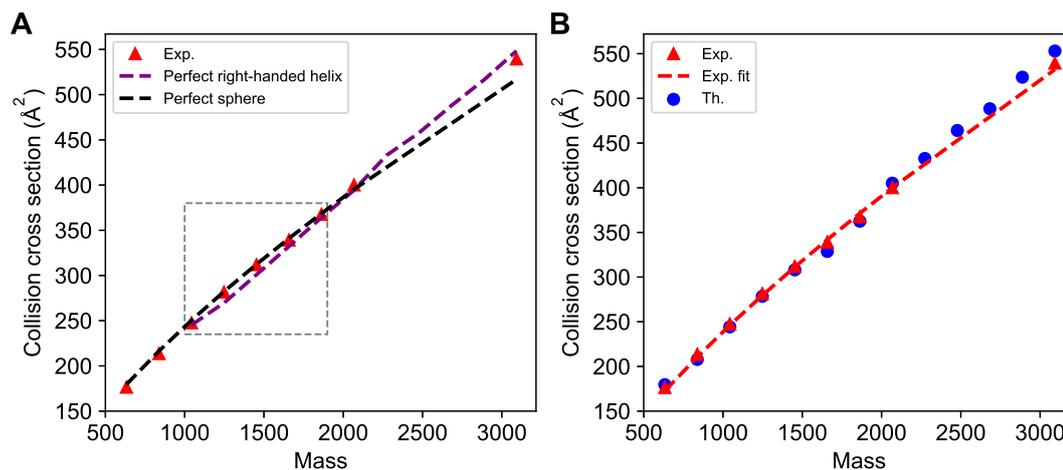
From these matrices, we mainly detect non-null data along the diagonal, indicative that intra-residue H-bonds are increasingly formed along the peptoid backbone as

the number of residues increases, except for a few residues lying at both the N and C termini, as nicely evidenced for  $N_{scp15}$ . Inspection of the lowest energy conformers reveals that the few first residues at N terminus do not form H-bonds but have rather their amide oriented towards the ammonium group (**Figure 5.23C**, red dashed box). As far as the C terminal residue is concerned, intra-residue H-bond is observed but in a different way that is not accounted in **Figure 5.25**. It associates the hydrogen from the amide (acting now as a donor) and the carbonyl from the carboxylic acid group (acting as the acceptor, **Figure 5.23C**, green dashed box and **Figure 5.24**).

Given that we finally seem to obtain a helical conformation in gas phase, the synthesis and analysis of  $N_{scp3\text{ to }10}$  and  $N_{scp15}$  were performed by Perrine Weber. Firstly, the collision cross sections of the corresponding ions were measured and compared them to the predicted  $\Omega$  of perfectly spherical ions using Equation 5.1 (**Figure 5.26A**). As indicated in the inset of the figure, the collision cross sections of the shorter oligomers (3 to 7/8 residues, mass range of 500 to 2000) are very close to those of the spherical ions. At longer chain length, the values deviate, which indicates a change in the conformation. Moreover, we also built perfectly right-handed helices starting with 5 residues,<sup>42</sup> as we did for  $N_{spe}$  peptoids using the same backbone dihedral combination and without further optimization. The associated  $^{TM}\Omega_{He}$  of the longer chain lengths nicely agree with the experimental data. This result suggests that the  $N_{scp}$  peptoid ions progressively adopt a more extended conformation that is most likely helical.

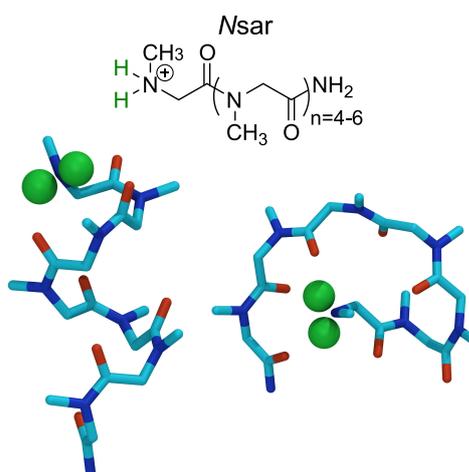
The comparison between the experimental  $\Omega$  with the theoretical  $\Omega$  generated using the coordinates from the MD simulations on the lowest energy conformers further confirm our prediction (**Figure 5.26B**).  $N_{scp}$  peptoid ions can thus form helices in gas phase. We also fitted the data using Equation 5.1 and obtained a  $B$  parameter of 0.71. This value qualitatively informs us that the general shape of  $N_{scp}$  peptoid ions is more extended than the shape of  $N_{sar}$  and  $N_{spe}$  (the  $B$  parameter is 0.615 and 0.68 respectively), in full consistency with the results obtained by molecular modeling.

The structural transition from partially helical to fully helical appears to depend on the number of hydrogen bonds that are formed. To estimate theoretically the minimum number of H-bonds required to stabilize a gaseous helix protruding out of the (charge solvation) loop, we consider a model peptoid bearing methyl groups as pending side chains ( $N_{sar}^{5-7}$ , **Table 5.6**) to remove the contributions from the side chains and exclusively focus on the stabilization provided by the backbone. These model peptoids are built in a hypothetical right-handed helix ( $\phi = -70^\circ$ ,  $\psi = 165^\circ$  and  $\omega = 0^\circ$ ).<sup>41</sup> In parallel, we consider the loop conformations of  $N_{sar}$  peptoid ions previously obtained in this Chapter (5.3.1).



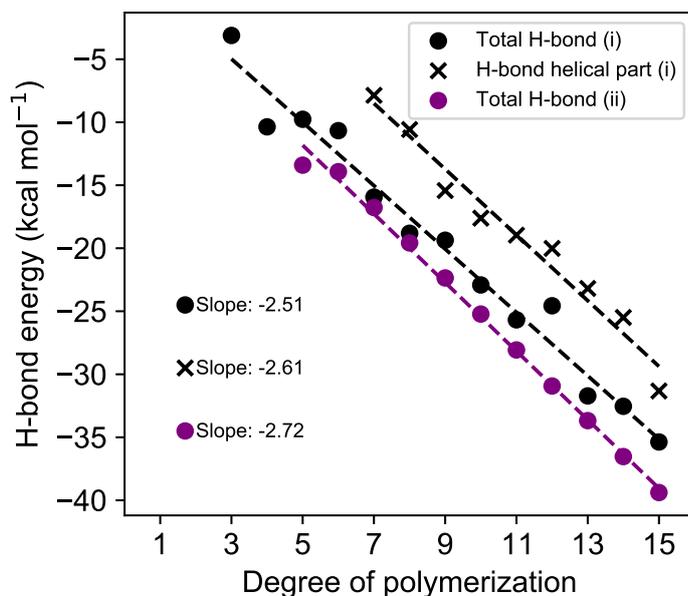
**Figure 5.26:** Experimental and theoretical collision cross sections of singly protonated  $Nscp$  peptoids. **(A)** Comparison of the evolution of the experimental  ${}^{TW}\Omega_{N_2 \rightarrow He}$  (red triangles) with the values predicted for perfectly spherical (black dashed line, using the equation  $\Omega = 2.435 \cdot M^{2/3}$ )<sup>123</sup> and perfectly helical ions (purple dashed line). The grey dashed rectangle focuses on the difference between the sphere and helix. **(B)** Comparison of the experimental  ${}^{TW}\Omega_{N_2 \rightarrow He}$  (red triangles) with the theoretical  ${}^{TM}\Omega_{He}$  values calculated on the geometries generated by MD simulations on the lowest energy conformers (blue dots). The experimental data are further fitted (red dashed line) by Equation 5.1.

**Table 5.6:** Primary structure of  $Nsar$  peptoids and the hypothetical helical structure of  $Nsar^7$  and its associated loop structure (most stable conformer) Energy contributions (kcal mol<sup>-1</sup>) extracted from the helical and loop structures from  $Nsar_{5 \text{ to } 7}$ : bonded interactions, hydrogen bonds, van der Waals (vdW) interactions, electrostatic interactions. The larger differences are observed for the electrostatic interactions between the helical and loop shapes (values in red).



$n$	Helix				Loop			
	Bonded interactions	H-bonds	vdW	Electrostatic	Bonded interactions	H-bonds	vdW	Electrostatic
5	25.4	-4.0	27.2	23.3	28.6	-4.3	24.5	11.5
6	30.7	-4.1	32.3	31.2	30.4	-5.9	32.9	16.9
7	36.0	-4.1	37.1	39.6	32.1	-3.2	37.6	25.1

Using molecular mechanics, the energy difference between different conformations can be decomposed into different energetic contributions, *i.e.*, the bonded (encompassing bond stretching, angle bending, dihedral and improper dihedrals) and the non-bonded (van der Waals, electrostatics and hydrogen bonds, as implemented in the DREIDING force field<sup>139</sup>) interactions. The bonded and non-bonded energetic contributions have been extracted for each  $N_{\text{sar}}$  model, in both the helical and loop conformations, and gathered in **Table 5.6**. For a given oligomer size, there are no major differences for the bonded energy contributions as well as for the hydrogen bond and van der Waals contributions between the helix and the corresponding loop. However, the electrostatic interactions differ significantly, with an average stabilization of about  $15 \text{ kcal mol}^{-1}$  in favor of the loop conformations. In a very naive vision, secondary strong interactions, such as H-bonds, are thus needed to compensate these  $15 \text{ kcal mol}^{-1}$  to make the helix conformations more stable than the loops.



**Figure 5.27:** Evolution of the hydrogen bond energy contribution in the  $N_{\text{scp}}$  peptoids as a function of the degree of polymerization. Contributions are calculated for two geometries: (i) the perfect right-handed helix (purple dots) and (ii) the most stable structure obtained upon conformational sampling (black dots). The H-bond contributions in the helical segments in the most stable structures (delimited by diagonal elements in **Figure 5.24**) are presented as black crosses.

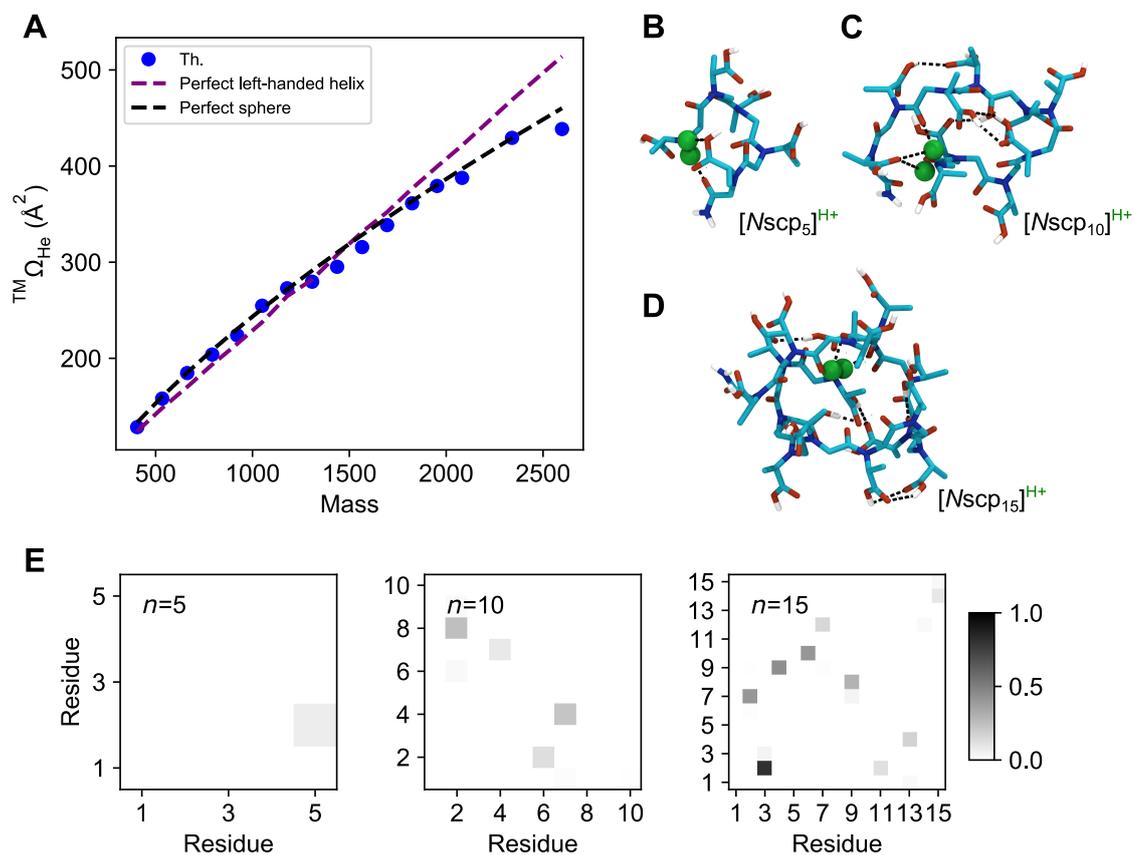
**Figure 5.27** presents the evolution of the hydrogen bond energy contribution of the  $N_{\text{scp}}$  peptoids as a function of the degree of polymerization. The contributions are calculated for two geometries: (i) the perfect right-handed helix (purple dots) and (ii) the most stable structure obtained after the conformational sampling (black dots). The total H-bond energy is very similar in both cases, (purple and black dots), indicating that the H-bond pattern found in the most stable structure is similar in magnitude to

that prevailing in the perfect helices. We have also extracted the H-bond contribution within the helical segment in the most stable structure (**Figure 5.27**, black crosses). To do so, for each chain length, we cut the most stable structure at the junction indicated by the hydrogen bond correlation matrices and performed a single point calculation on the segment that forms intra-residue hydrogen bond, *i.e.*, the helical segment, from which we extract the hydrogen bond energy contribution. Thereby we obtain three linear evolutions whose slopes afford an averaged stabilization energy per H-bond, around  $-2.5 \text{ kcal mol}^{-1}$ . This value nicely matches that of  $-2 \text{ kcal mol}^{-1}$  reported for  $\alpha$ -helix,<sup>195</sup> allowing to roughly estimate that minimum 6 to 7 inter/intra-residue H-bonds should be enough to stabilize a charged helix in the gas phase. For the partially helical  $N_{\text{scp}_8}$  (**Figure 5.23B**), we clearly count 4 H-bonds between the side chain  $i$  and the backbone carbonyl  $i$ , starting at residue 4, as shown in **Figure 5.24**. The energetic gain provided by these 4 hydrogen bonds amounts to  $-10.8 \text{ kcal mol}^{-1}$  (**Figure 5.26**), which is close to the generic helix/loop difference estimated at  $15 \text{ kcal mol}^{-1}$  from the data in **Table 5.6**.  $N_{\text{scp}_{15}}$  clearly adopts a helical conformation stabilized by 10 H-bonds, as shown in the diagonal of **Figure 5.24**, hence leading to a stabilization by  $-30 \text{ kcal mol}^{-1}$ . Here, the H-bonds clearly prevent the peptoid ion from fully wrapping around the charge to form a compact structure. Actually, the ammonium group is still solvated by the first three to four residues but this does not hamper the stabilization of a well-defined helical structure from the remaining units. Moreover, the charge is also stabilized by the macrodipole formed by the alignment of the carbonyls from the C to the N terminus and that increases with growing chain length.<sup>41</sup>

Since bulky  $\alpha$ -chiral side chains are described as a primary condition to stabilize helical peptoids in solution,<sup>42,48</sup> we further assess whether the H-bond capability of side chains is a sufficient condition to form helices in gas phase. To do so, we select the  $N_{\text{rce}}$ , '(*R*)-*N*-(1-carboxy-2-ethyl)' side chain that bears a methyl group instead of the phenyl group present in the  $N_{\text{scp}}$  side chain. In this way, the  $\alpha$ -chiral and H-bond donor/acceptor characters are conserved while partially releasing the steric hindrance via the elimination of the phenyl groups. We sample the conformational space of the *N*-protonated  $N_{\text{rce}}$  peptoids (from 3 to 20 residues). We compute the average  ${}^{TM}\Omega_{He}$  for the coordinates generated by MD simulations on lowest energy conformer and report the  $\Omega$  evolution as a function of the mass in **Figure 5.28A** (blue dots). We also build model helices for each DP (using the same dihedrals as previously) and compute the  ${}^{TM}\Omega_{He}$  for perfectly helical peptoids (**Figure 5.28A**, purple dotted line). The data of the most stable structures starts to significantly deviate from the helical trend at  $m/z$  1200 (around 7 residues). However, the data nicely agree with the globular evolution using  $\Omega = 2.435 \cdot M^{2/3}$  (black dashed line).

We have selected three representative chain lengths,  $n = 5, 10$  and  $15$ , and displayed their most stable conformations in **Figure 5.27 (B-D)**. These ions are found to be very compact, *i.e.*, they adopt a globular form, with the backbone wrapped around the charge (green beads), similar to the *N*sar peptoid and L-alanine peptide ions. When generating the hydrogen bond correlation matrices for these ions (**Figure 5.28E**), we observe that intra-residue H-bonds are barely formed, while inter-residue H-bonds tend to be increasingly favored as the chain length increases (off-diagonal squares).

These peptoids have not been synthesized in the present work, but the theoretical results clearly indicate that the *N*rce ions adopt compact structures in the gas phase, thus emphasizing that both the presence of H-bonds and the presence of bulky side chains are two key ingredients to stabilize peptoid helical ions in the gas phase.



**Figure 5.28:** (A) Theoretical  $\Omega$  of protonated *Nrce* peptoids obtained after MD simulations on the lowest energy conformers (blue dots). These data are in very nice agreement with the predicted  $\Omega$  of perfectly spherical ions (black dashed line) but not with the perfectly helical ions (purple dashed line). Optimized structures of protonated *Nrce* with (B) 5, (C) 10 and (D) 15 residues, revealing that the presence of H-bond donor/acceptor pairs along the peptoid backbone is not sufficient to stabilize helical conformations in the gas phase. The inter-residue H-bonds are represented as black dotted bonds. (E) H-bond correlation matrices of *Nrce* peptoids (B), (C) and (D) generated from 50 ns MD simulations. The X and Y axes correspond to the residue number starting from the N terminus. The hydrogen bonds are monitored between the hydrogen from the carboxylic acid moieties and the oxygen from the backbone carbonyls and the carboxylic acid carbonyls.

## 5.4 Conclusions

Altogether, this study sheds light on multiple aspects regarding ion mobility mass spectrometry. First, we demonstrated that the use of trend line method to characterize polymeric compounds can be misleading for ions whose shape is different from spherical, such as helices. Indeed, the characteristic  $B$  parameter is dependent on the range of data available, and can thus lead to the attribution of the wrong conformation. This example emphasizes the need of combining IM-MS with computational chemistry to fully interpret and characterize the gas-phase conformation of ions.

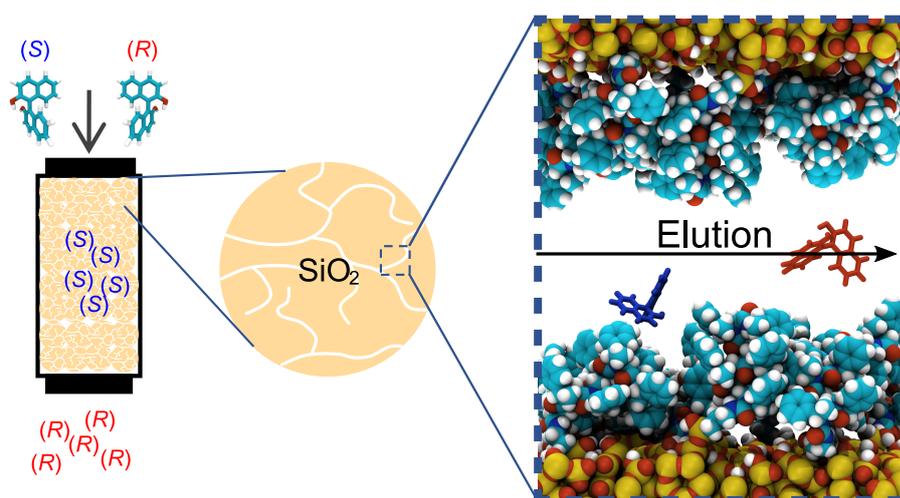
Using the IM-MS in combination with computational chemistry, we demonstrated that the gas-phase conformation of peptoids with side chains of different complexity, namely  $N_{sar}$  and  $N_{spe}$ , share common features in their gas-phase conformation. They both adopt a loop-like conformation to stabilize the charge by orienting the amide dipoles toward it, although their solution conformations are drastically different (random coil and helix for  $N_{sar}$  and  $N_{spe}$ , respectively). In the case of  $N_{spe}$ , an organization of the backbone into a helix starts appearing for long oligomers, when the charge is fully stabilized. It would be relevant here to study even longer peptoids to absolutely confirm that the protruding helix keeps increasing. However, longer peptoids would start acquiring multiple charges which will most likely affect the conformation. Nonetheless, our finding clearly evidenced that the secondary non-covalent interactions were not strong enough to maintain the backbone in its helical shape when transferred to the gas phase.

We tackled this issue by introducing another side chain,  $N_{scp}$ , that carries a carboxylic acid moiety able to form intra-molecular hydrogen bonds. Our joint experimental and theoretical investigation reveals that helical peptoid conformers can now be detected in the gas phase due primarily to the formation of an intra-residue H-bond network associating the hydrogen atom from the side chain carboxylic acid to the oxygen atom of the amide inside the same residue. Altogether, we demonstrated that obtaining stable gas-phase helical structures is conditioned by different factors: (i) the possibility to create intra-residue H-bonds; (ii) the presence of a sufficient amount of residues to energetically counterbalance the charge induced folding ( $DP > 8$ ) and; (iii) the presence of bulky side chains to generate a high degree of steric hindrance. *In fine*, these conformers arise from the compromise between maximizing charge coordination and minimizing the disruption of the hydrogen bonds.

This study thus highlights that ion mobility mass spectrometry supported by computational chemistry can be used to characterize the secondary structure of peptoids in gas phase. However, we need to bear in mind that the requirement of ionic species might alter the conformation when going from the solution to the gas phase, similarly to what may happen when going from the solid to the solution state. Ion mobility mass spectrometry still provides relevant insights about the importance of secondary non-covalent interactions to strengthen the conformation desired in solution. For instance, it is usually hoped that the analyte conserves its solution structure (also called its “native” state) upon transfer in gas phase, which commonly happens for proteins. In the present case, the solution conformation(s) are not identified yet, but CD spectroscopy indicates that these peptoids present conformational rearrangements depending on the nature of the solvent. Given that a structuring is probed by CD, it is very likely that the gas-phase structure also closely reflects that prevailing in solution.



# Peptoids as chiral selector molecules



Based on  
*Biomacromolecules*, 22, 6, 2021, 2573



## 6.1 The importance of enantioselectivity

The majority of biological molecules such as amino acids and proteins, carbohydrates, or hormones are chiral although almost exclusively one of the two enantiomers is commonly found in living organisms (left-handed amino acids and right-handed sugars).<sup>196,197</sup> In contrast, many chiral synthetic compounds, such as drugs, are obtained in racemic mixtures.<sup>102</sup> Enantiomers often display different reactivity with other compounds that are also enantiomers, as found in biological systems.<sup>101</sup> Consequently, a careful control of the enantiomeric composition is particularly critical in the pharmaceutical field to avoid medical disasters, such as the infamous thalidomide drug scandal in the late 1950s.<sup>101</sup>

Currently, three main ways may be followed to obtain enantiopure drugs, each having its pros and cons. These methods are well documented in several books/reviews (see 198) and will only be briefly described to grasp their fundamental concepts.

The first method relies on asymmetric synthesis (also called enantioselective synthesis).<sup>199</sup> Using this method, an enantiopure molecule (which can be either a catalyst or a reagent) is reacted with an achiral molecule that leads to the formation of a new enantiopure molecule.<sup>198</sup> However, the two enantiopure starting molecules are required if both enantiopure drugs are needed to assess their performance, which is not always straightforward.<sup>198</sup> Moreover, this method may be limited by the high cost of the enantiopure starting materials and often requires lengthy synthetic routes.<sup>103,200</sup>

The other two methods to obtain enantiopure drugs rather involve the derivatization and separation of the enantiomers. Chemical derivatization typically relies on the formation of a diastereoisomeric salt formed between a chiral derivatizing agent and the racemic mixture that needs to be resolved. The diastereoisomers are characterized by different physico-chemical properties and may be separated by crystallization and then treated to recover the separated enantiopure compounds.<sup>102,104,201</sup> Though efficient, this method is not universal since each system requires to find the proper chiral derivatizing agent that must be enantiomerically pure and must react equally with both enantiomers.

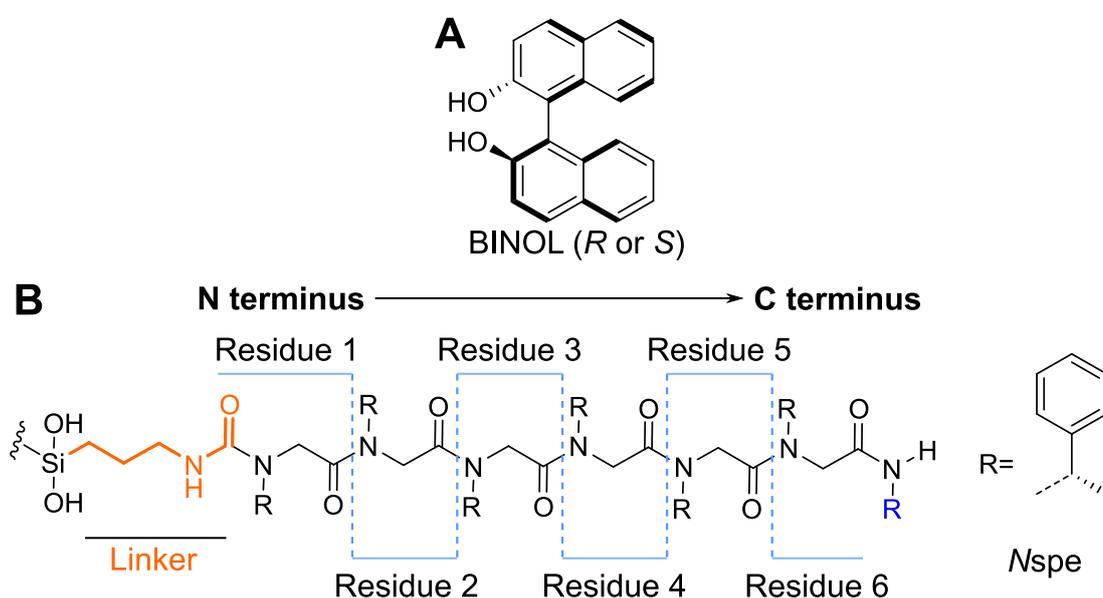
In this chapter, we will focus on the last method that relies on the preparation and direct separation of a racemate by chiral stationary phases in chromatography columns (CSPs).<sup>105</sup> Unlike usual stationary phases typically constituted of pristine silica or functionalized by covalently bound alkyl chains, chiral stationary phases consist in silica functionalized by chiral compounds. The strength and weakness of this method are closely intertwined: many chiral compounds can be employed as chiral selectors, but their efficiency will dramatically depend on the chemical nature

of the racemate that must be resolved.<sup>202</sup> Typically, cyclodextrins,<sup>203</sup> polysaccharides,<sup>204</sup> proteins or peptide oligomers<sup>106,108,205</sup> are used as chiral selectors in such chromatographic separation. In the latter case, previous studies suggested that chiral conformations can promote and enhance the enantioselectivity, such as a helical conformation.<sup>106–108,206</sup> Peptoids as chiral selector might conciliate the pros and cons of current CSPs. Indeed, as previously discussed, peptoids can be easily synthesized and many different type of side chains can be appended to their backbone, enabling a large chemical diversity.<sup>13</sup> Indeed, intrinsically, the peptoid backbone is constituted by a series of amide that can act as hydrogen bond acceptors, while the side chains could carry aromatic moieties that could form  $\pi$ - $\pi$  interactions with the racemate, if needed. Moreover, they can also exhibit conformational chirality,<sup>41</sup> for example, by forming a helix of a given screw sense. Therefore, within a single class of compounds, one could achieve the separation of virtually any racemates.

In 2011, a proof of this concept was brought by Wu *et al.* who investigated *N*spe peptoids grafted on silica as a chiral stationary phase.<sup>24</sup> They have shown that *N*spe peptoids grafted *via* their N terminus extremity on silica do exhibit enantioselective properties against chiral binaphthyl derivatives. This enantioselectivity has been linked to the fact that peptoids bearing *N*spe side chains start adopting a helical conformation at a length of about 5 residues,<sup>42,47</sup> which allows the establishment of specific interactions with a given enantiomer. There is, however, no description at the atomistic scale of the origin of these enantioselective properties. A clear understanding of this process could lead to the design of new peptoid sequences in view of optimizing the separation performances or even to broaden the chemical range of compounds that could be resolved.

This has motivated the present study which describes the process of chiral recognition of peptoids substituted by *N*spe side chains grafted on silica by means of MD simulations in collaboration with Prof. Luca Muccioli. Simulations of chiral interfaces can assist experimentalists by shedding light on the recognition mechanism and the interactions involved between the selectors and the analytes to guide synthetic efforts, as shown for the well-known Whelk-O1 chiral stationary phase,<sup>207,208</sup> or even for peptides and saccharides oligomers.<sup>209–211</sup> To best mimic the conditions reported in the experimental study, we reproduced in our simulations key relevant parameters such as the grafting density and the solvent composition.<sup>24</sup> The experimental study used a binaphthyl derivative as chiral guest compound to assess the enantioselectivity properties of peptoids, in particular the 2,2'-bihydroxy-1,1'-binaphthyl that will be abbreviated "BINOL" throughout this chapter (**Scheme 6.1A**). Compared to the peptoids discussed in the previous chapter, Wu *et al.* used the solution-phase synthesis protocol, which produces the peptoid with an additional side chain on the C

terminus amide compared to the solid-phase synthesis protocol (see [Introduction](#) for details). We considered a peptoid of six residues that consequently has seven  $N_{spe}$  side chains as host (referred to as CSP5 in Ref. 24). Using the current notation of  $N_{spe}$ , might be misleading since the number of residues and side chains do not match. Therefore, this peptoid will be referred as “peptoid hexamer” throughout the chapter for sake of clarity. The peptoid is grafted on silica beads using a linker (triethoxy-(3-isocyanatopropyl)silane) that forms a urea bond at the N terminus extremity of the peptoid chain (**Scheme 6.1A**). We selected this specific peptoid sequence since it displays one of the best separation factors towards (*R*)- and (*S*)-BINOL in the experiments carried out by Wu *et al.*<sup>24</sup>



**Scheme 6.1:** Primary structure of (A) BINOL (*R* or *S*) and (B), the peptoid bearing  $N_{spe}$  side chains, represented from N to C terminus, its linker for grafting on silica (orange). The extra side chain is highlighted in blue and has to be compared with the sequence from the solid-phase synthesis (**Figure 5.8**).

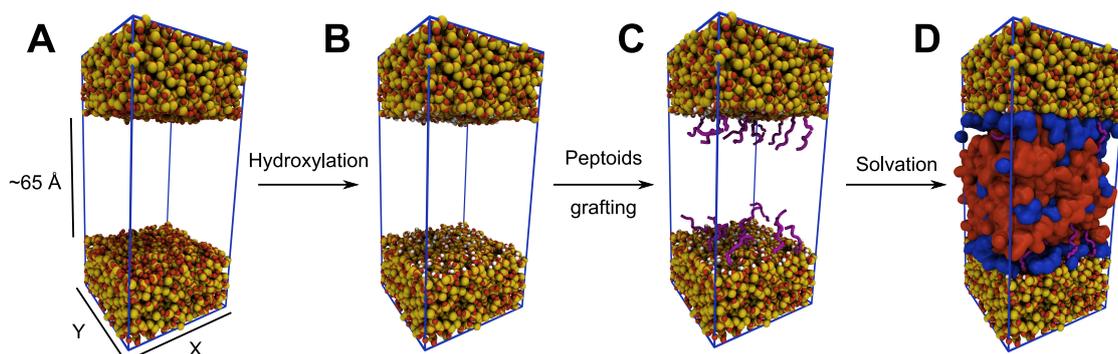
Within our approach, we simulate the elution process by pulling the guest molecules between two silica surfaces functionalized by peptoids to mimic what happens inside a silica pore.

## 6.2 Methodology

### Sample preparation

In the experimental work of Wu *et al.*,<sup>24</sup> a  $N_{\text{spe}_7}$  (CSP5) loading of  $1.3 \cdot 10^{-4}$  mol  $\text{g}^{-1}$  was reported. Since the surface area of the HPLC-grade spherical silica gel is  $300 \text{ m}^2 \text{g}^{-1}$ , this leads to a grafting density of  $4.333 \cdot 10^{-7}$  mol  $\text{m}^{-2}$ , and hence to  $\sim 0.26$  molecule  $\text{nm}^{-2}$ .

Within our simulations, the bulk sample of amorphous silica was already generated by Dr. Otello M. Roscioni in a previous study.<sup>212</sup> Briefly, a crystalline supercell of cristobalite (under periodic boundary conditions (PBC)), is heated to 4000 K and cooled to 300 K to obtain an amorphous silica glass (density of  $2.2 \text{ g cm}^3$ ).<sup>a</sup> From this structure, a slab with two free surfaces is extracted whose surface dimensions equal to  $57.1 \times 57.1 \text{ \AA}^2$ , and with a thickness of about  $60 \text{ \AA}$ . The slab was placed horizontally (in the XY plane) in the simulation box, with the two surfaces of the slab facing each other in the Z direction (Figure 6.1A).



**Figure 6.1:** Main steps of the preparation of our chromatography column model. First, the two slabs are generated from vitreous silica. Water is then added to hydroxylate the surfaces and obtain silanol moieties. Peptoids are grafted on the surfaces with a grafting density matching the experimental values from Wu *et al.* (10 chains on each surface).<sup>24</sup> Finally, a mix of 2-propanol (blue) and *n*-hexane (red) is added to mimic the elution phase (70/30 *v/v*) as well as 5 BINOL molecules.

These two surfaces form the walls of the model chromatography column, or more precisely of a single pore of the silica beads which are used in the column. Random defects were further created independently on both surfaces by removing  $\text{SiO}_2$  units in order to increase the roughness of the otherwise atomically flat surfaces. The atomic coordinates were then relaxed to allow local surface reconstruction by thermal annealing. The slab was finally energy minimized at 0 K. The surfaces were then

<sup>a</sup>The interactions between silicon and oxygen atoms are described with a Coulomb and Buckingham potential whose parameters were extracted from Ref. 213.<sup>212</sup>

hydroxylated by adding water molecules that will react with dangling silicon and oxygen atoms to obtain silanol moieties with a density of 4.9 OH groups nm<sup>-2</sup> (**Figure 6.1B**).<sup>214,215</sup>

The pore size of the silica beads used by Wu *et al.* is about 100 Å,<sup>24</sup> whereas the gap used between the two surfaces in our simulation is instead 65 Å in order to save computational time. This is a reasonable approximation because this thickness is large enough to ensure that peptoids grafted on opposite surfaces do not interact.

The surfaces were then decorated with a self-assembled monolayer of oligopeptoid chains by substituting hydroxyl moieties by the peptoid chain and its linker, leading to Si<sub>surface</sub>-OSi(OH)<sub>2</sub>-linker-peptoid. Given the surface dimension and the grafting density, 10 peptoid chains were randomly grafted on both surfaces (**Figure 6.1C**, purple ribbons). Their initial conformation is the reported right-handed helix from Armand *et al.* ( $\varphi = -75^\circ$ ,  $\psi = 165^\circ$  and  $\omega = 0^\circ$ ).<sup>41</sup>

The experimental mobile (or elution) phase is a mixture of *n*-hexane/2-propanol 70/30 (*v/v*), which translates in our simulation box into 575 and 425 molecules, respectively, to fill the gap ( $\sim 58 \times 58 \times 65 \text{ \AA}^3$ ) between the two silica surfaces (**Figure 6.1D**, red and blue molecular surfaces).

This procedure was carried out to build two samples, one with (*R*)- and the other with (*S*)-BINOL molecules.

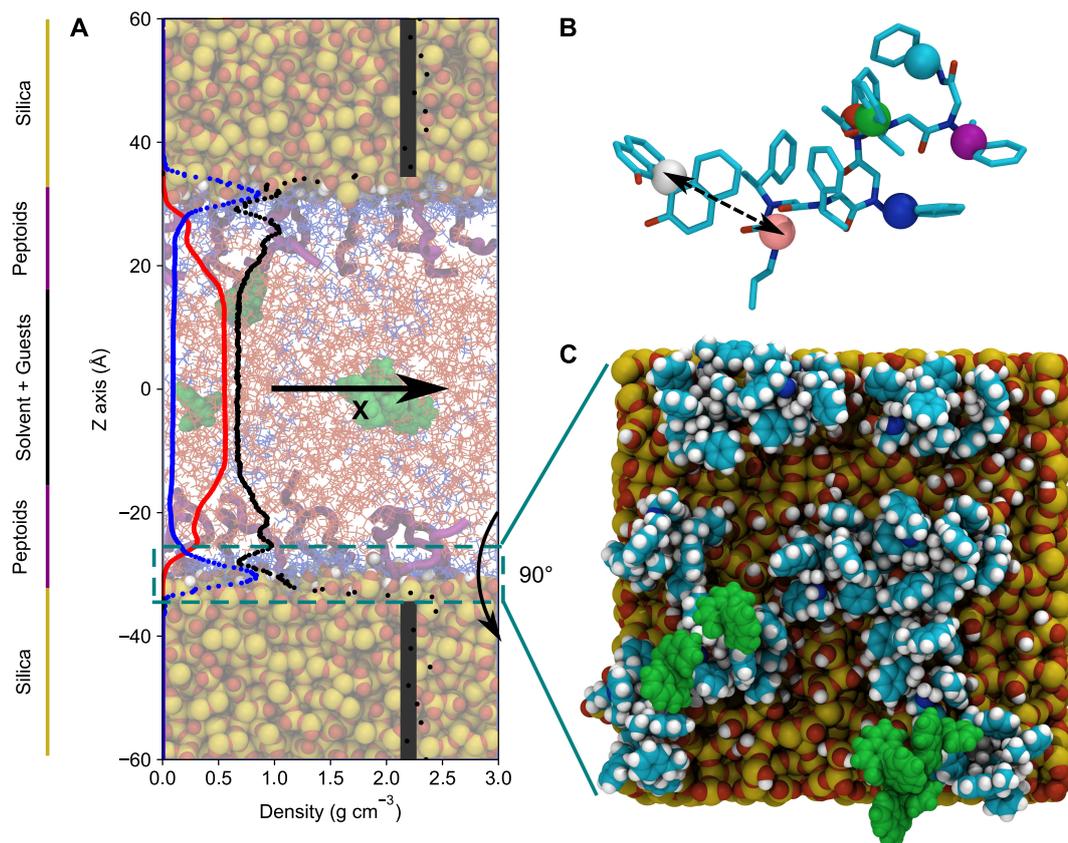
### Force fields

Our system is composed of multiple components, each being described by a suitable force field. The amorphous silica is described using the Clay force field.<sup>216</sup> The solvent mixture is described using the CGenFF force field,<sup>142</sup> which best reproduces the individual densities as well as the density of the mixture.<sup>217</sup> BINOL molecules are also described using the same force field. The partial atomic charges of the solvent and BINOL molecules were obtained on the ParamChem web server.<sup>218</sup>

In Chapter 4, we presented our reparametrization methodology of the DREIDING force field to describe peptoids. In Chapter 5, we used these force field parameters to describe the conformational behavior of different peptoids and obtained very nice agreement with the experimental results. However, we also discussed about the barrier heights between minima that are quite high, especially for amide bonds. Therefore, it is very unlikely that these barriers will be crossed at 300 K on a reasonable MD time scale and a that proper conformational sampling is made. To circumvent this issue, we reduced the barrier heights of the  $\omega$ ,  $\varphi$  and  $\psi$  dihedrals while keeping the relative stability between the existing minima intact. This strategy allows to artificially improve the conformational sampling.

### Simulation conditions

The simulations on the entire system were performed using the NAMD software using 3D periodic boundary conditions (Figure 6.2).<sup>219,220</sup> Solvent molecules (*n*-hexane and 2-propanol), peptoids, guest molecules (BINOL) and the silica atoms closer to the surfaces (within 5 Å) were subjected to thermal motion, while the core of the slab was kept frozen at its equilibrium position to save computational time.<sup>221</sup> No bond constraints were applied on the system.

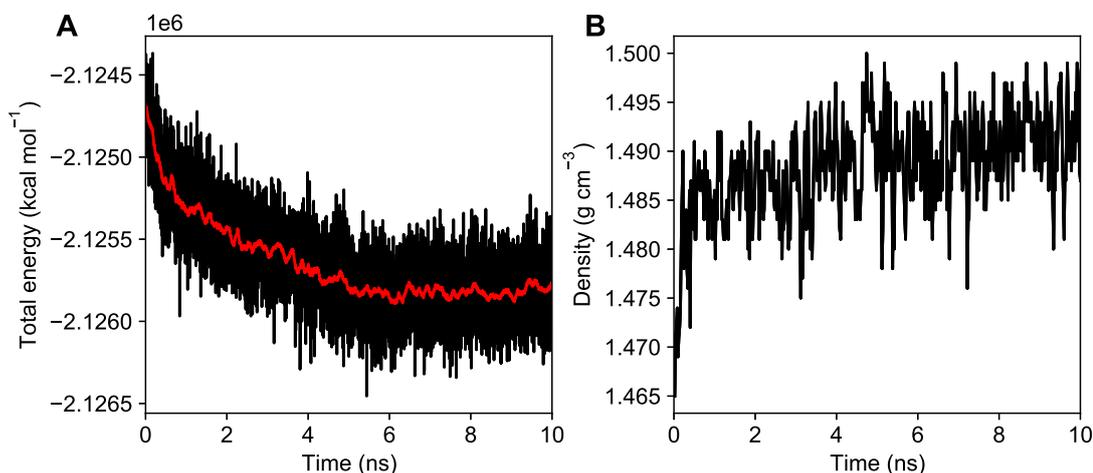


**Figure 6.2:** (A) Density profile (black dots) along the Z axis of the simulation box. The elution occurs along the X axis. The vertical gray bands correspond to the experimental density of amorphous silica. Partial phase segregation of *n*-hexane (red dots) and 2-propanol (blue dots) is observed on top of the silica surfaces. (B) Representation of the distances computed between the center of mass of BINOL and the center of mass of each peptoid residue. (C) Top view of the bottom silica surface with grafted peptoids. BINOL molecules are highlighted in green.

The simulations were systematically carried out on a system containing only a given BINOL enantiomer (*R* or *S*), since a racemic mixture would either reduce the quality of the statistical analysis made on a given enantiomer if keeping the same total number of molecules or increase too much the number of molecules in the box to guarantee that they behave independently.

Each simulation was carried out using a timestep of 1 fs. The particle mesh Ewald (PME) method was used to deal with the Coulomb interactions with a real space cutoff of 10 Å and a switching distance of 9.5 Å.<sup>222</sup> The van der Waals interactions are treated by a Lennard-Jones potential (12-6) using a 10 Å cutoff with a switching distance of 9.5 Å, and Lorentz-Berthelot mixing rules.

After optimization of its geometry, the system was first equilibrated with a Langevin thermostat<sup>219,223</sup> at 298 K and a Langevin barostat<sup>224</sup> set to 1 atm applied only in the Z direction (perpendicular to the functionalized surfaces), while the sizes of the box along the X and Y axes remain fixed. Convergence of the density and the potential energy was reached after about 10 ns for every simulations (**Figure 6.3**). The silica bulk and the solvent mixture densities reached values fairly close to the experimental values (2.23 vs. 2.2 g cm<sup>-3</sup> for silica and 0.67 vs. 0.676 g cm<sup>-3</sup> for the solvent mixture).<sup>217,225</sup>



**Figure 6.3:** Evolution of (A) the total energy (black curve) with its running average (red curved) and (B) the density of the system composed of *Nspe* peptoid chains and (*S*)-BINOL guest molecules. Equilibration is reached within 10 ns (NPT, Langevin piston 1 atm, 298 K). The same observation is made for the other simulations.

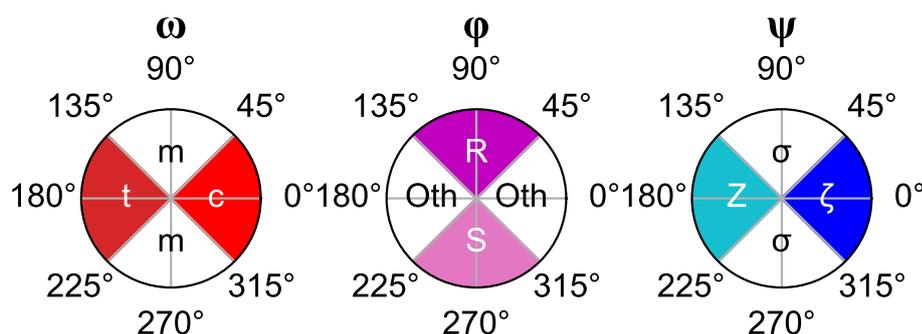
Then, we take advantage of steered molecular dynamics (SMD)<sup>226</sup> to simulate the elution process occurring in the chromatography column. SMD consists in pulling atoms either at constant velocity or constant force. We chose to apply a constant pulling force on the BINOL molecules to characterize the enantioselectivity. In this manner, we can monitor their velocities, which will be affected by their interactions with the environment and eventually slowed down when interacting favorably with peptoids. We can then compare the average velocities of both enantiomers and qualitatively match the results to the data from Wu *et al.*<sup>24</sup>

In practice, we performed simulations of 650 ns with the *Nspe* substrate in the NVT ensemble (Langevin thermostat, 300 K) for the two samples containing 5 enantiomers (*R* or *S*) using a pulling force of 1.25 kcal mol<sup>-1</sup> Å<sup>-1</sup> on a single atom (the

hydrogen attached to one of the two carbons in  $\alpha$  to the hydroxyl moiety) of each guest molecule along the X direction. We chose this value because the flow it generates leads a low deviation from the temperature imposed by the thermostat. Moreover, it keeps the elution efficient (*i.e.*, the net flow is not zero) and the computational cost reasonable, which is inversely proportional to the pulling force (the lower the force, the longer the time for the analyte to travel a certain distance). The convergence of the simulations was monitored ensuring that the moving average of the velocity of each BINOL molecule inside the simulation becomes constant at long enough simulation times.

## 6.3 Conformational analysis

Since we seek to characterize in a statistical way the conformation of the peptoids involved in the recognition process, we developed a labeling method similar to that proposed by Spencer *et al.*<sup>59</sup> We discussed multiple times about the definition of the peptoid secondary structure that depends on particular combinations of the backbone dihedrals  $\omega$ ,  $\phi$  and  $\psi$ . The typical example is a perfect right-handed helix, characterized by a periodic repetition of the pattern ( $\omega \sim 0^\circ$ ,  $\phi \sim -80^\circ$ ,  $\psi \sim 180^\circ$ ). The labeling method is based on the assignment of a given letter to a range of values for each dihedral angle. Spencer *et al.* describe  $\psi$  by the capital Greek letter  $Z$ ,  $\phi$  by R (left-handed) or S (right-handed) and  $\omega$  by c (*cis*) or t (*trans*). Accordingly, a helical conformation would be denoted as a repetition of ZSc. However, this nomenclature is too restrictive since it is mostly focused on helical-like geometries, although they are the most abundant. In this work, we extended the nomenclature to include a broader range of possible conformers and be able to account for further angular variations (**Figure 6.4**).



**Figure 6.4:** Polar plots illustrating the letter code for each dihedral of a peptoid residue with its associated range of values.

Secondary structures arise from a combination of this 3 letter code for each residue, which can rapidly become cumbersome for moderately long peptoid chains such as our peptoid hexamer. For instance, a hexamer in right-handed helical conformation would be represented by the sequence ZSc-ZSc-ZSc-ZSc-ZSc.<sup>b</sup> We decided to simplify the 3 letter code for each residue into a single letter, for the present example, “R”, because ZSc corresponds to the **R**ight-handed helix. In principle, the ideal dihedral angle  $\psi$  in a right-handed helix should be close to  $180^\circ$ , but its value can be

<sup>b</sup>Note that, although the peptoid chain has six residues, only five letters characterize its conformation because the residue prior to the N terminus has its urea moiety connecting the peptoid to the linker. Therefore, this residue is not considered in the labeling since the dihedral combinations are different from the rest of the chain (see **Scheme 4.1** for the complete dihedral definition of residues).

close to  $0^\circ$  for the last residue at C terminus. In this special case, the letter is written in lowercase (“*r*”). The combination for the Left-handed helix *ZSc*, is abbreviated by “*L*” (or “*l*” at C terminus). The combinations of  $\phi$  and  $\psi$  that yield no specific conformations are written by “*C*” if their amide is in *cis* conformation, or “*T*” if it is in *trans* conformation. A last simplification is “*M*” that corresponds to amide bonds having a value too far from *cis* or *trans* (beyond  $\pm 45^\circ$ ). The conversion from the 3 letter code residue to the simplified single letter code is summarized in **Table 6.1**.

**Table 6.1:** Table of conversion from the complete letter code of a monomer unit into its simplified version.

3 letters code	Simplified single letter code
<i>ZSc</i> (or $\zeta$ <i>Sc</i> at C terminus)	“ <i>R</i> ” (or “ <i>r</i> ”)
<i>ZRc</i> (or $\zeta$ <i>Rc</i> at C terminus)	“ <i>L</i> ” (or “ <i>l</i> ”)
<i>ZOthc</i> , $\zeta$ <i>Othc</i> , $\sigma$ <i>Othc</i> , $\zeta$ <i>Rc</i> , $\zeta$ <i>Sc</i> , $\sigma$ <i>Rc</i> or $\sigma$ <i>Sc</i>	“ <i>C</i> ”
<i>ZOht</i> , $\zeta$ <i>Oht</i> , $\sigma$ <i>Oht</i> , $\zeta$ <i>Rt</i> , $\zeta$ <i>St</i> , $\sigma$ <i>Rt</i> or $\sigma$ <i>St</i>	“ <i>T</i> ”
<i>ZOthm</i> , $\zeta$ <i>Othm</i> , $\sigma$ <i>Othm</i> , $\zeta$ <i>Rm</i> , $\zeta$ <i>Sm</i> , $\sigma$ <i>Rm</i> or $\sigma$ <i>Sm</i>	“ <i>M</i> ”

When combining these notations, we finally end up with a sequence of five letters (from N to C terminus), *i.e.*, “*RRRRR*”. We established every possible sequences that can arise from this 5 letter code in Appendix C (**Table C.1**) and summarized the most frequent obtained during our simulations in **Table 6.2**. The conformation of every peptoid chains in the simulations will be characterized using this nomenclature methodology.

**Table 6.2:** Selected conformations, simplified sequence and verbose description of their structure.

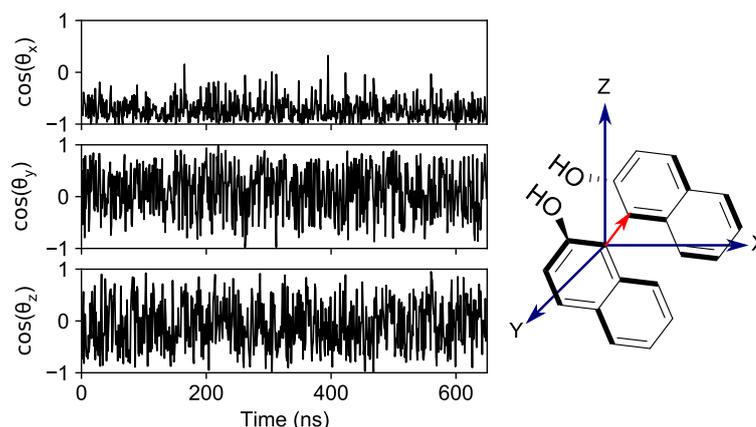
Simplified sequence	Description
“ <i>RRRRr</i> ”	Almost perfect right-handed helix, except for the C terminus residue which has its $\psi$ close to $0^\circ$
“ <i>RRRRX</i> ” with $X = L, l, M, C, T$	Partial right-handed helix having the first residue at C terminus which is not in helical conformation
“ <i>RRRXX</i> ” with $X = L, l, r, M, C, T$	Partial right-handed helix having the two residues at C terminus which are not in helical conformation
“ <i>XXXXX</i> ” with $X = R, r, L, l, M, C, T$	Random arrangement of every possible residue conformation

## 6.4 Results and Discussion

### 6.4.1 Chiral peptoids

In this section, we consider two samples with  $N_{\text{spe}}$  peptoids whose compositions are exactly the same, except the chirality of the 5 BINOL molecules. The guest molecules were randomly inserted in the solvent layer depicted in **Figure 6.2A**. During the equilibration phase, both systems display a certain phase segregation in the density profile between *n*-hexane and 2-propanol molecules, the latter forming a layer on top of the silica surface. This is explained by favorable polar interactions between silanol moieties from the surface and 2-propanol, compared to *n*-hexane. Such a segregation has also been observed by Monte Carlo simulations performed on alkane-alcohol mixtures, both in the bulk and in confined samples.<sup>209,227,228</sup>

After equilibration of the system, we carried out a SMD with the pulling force applied on a single atom of each of the 5 (*R*)- or (*S*)-BINOL molecules along the X direction. We checked that this force does not induce a specific orientation of BINOL by measuring the time evolution of the cosine of the angle  $\theta$  formed between the C-C bond linking the naphthyl moieties and the cell vectors. We report these values for a (*S*)-BINOL molecule that we labeled (*S*)-BINOL1 (**Figure 6.5**). The value of  $\cos(\theta)$  in both the Y and Z directions oscillates between -1 and 1, thus indicating that the BINOL molecules sample all possible orientations. As expected, along the X direction, the value mainly oscillates between 0 and -1 since the pulling force is applied along this direction.

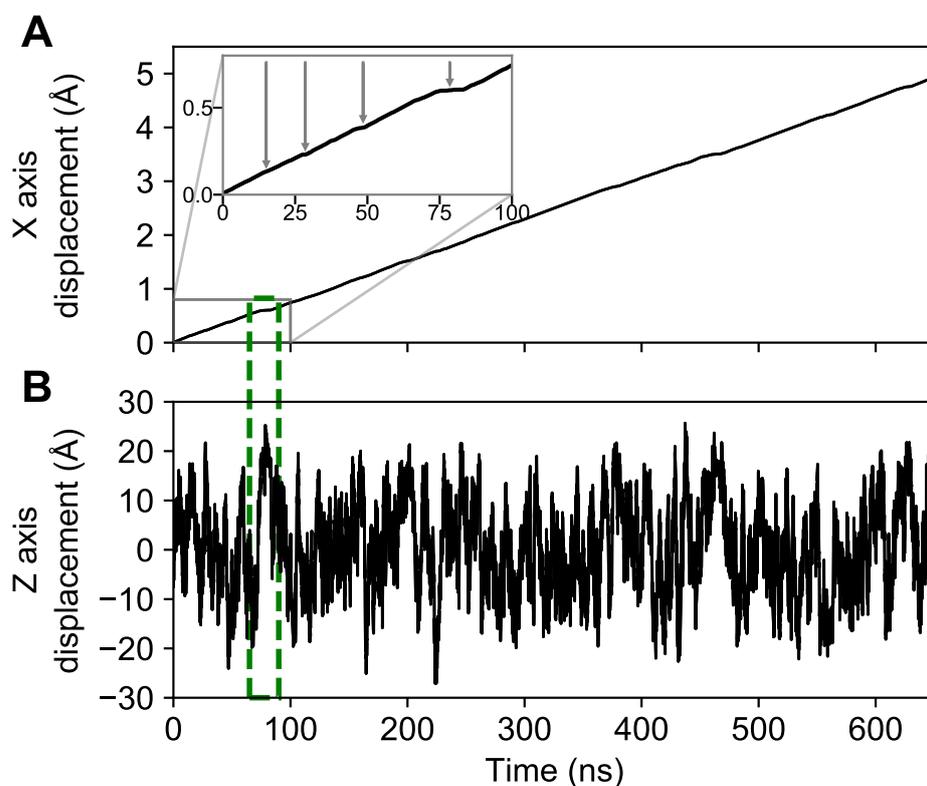


**Figure 6.5:** Time evolution of the cosine of the angle  $\theta$  between the inter-ring C-C bond (red arrow) for the (*S*)-BINOL1 molecule and the cell vectors (X, Y, and Z). The data for the other guests (either (*R*) or (*S*)) follow the same trends and are not presented to avoid redundancies.

We monitored several other parameters along the SMDs, such as the X, Y, and Z coordinates of the center of mass of each BINOL molecule, the conformation of

each peptoid chain (20 in total) using the conformational analysis described in the section 6.3. We also characterized the hydrogen bonds by measuring the distribution D-H $\cdots$ A distances and angles (starting at 2 up to 4 Å by 0.2 Å steps, and at 20° up to 180° by steps of 20°).<sup>31</sup> For sake of conciseness, we will discuss in detail the time evolution of a single (*S*)-BINOL molecule (previously labeled (*S*)-BINOL1), while other physical observables are computed as the average over the 5 molecules. The graphical analysis of the remaining BINOL molecules (*R*) and (*S*) is reported in the Appendix C.

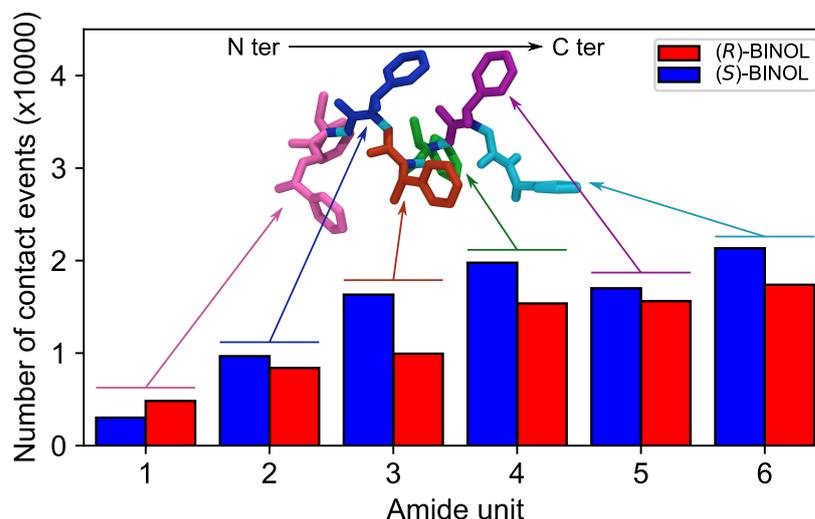
The analysis of the displacement of the center of mass of each (*S*)-BINOL along the direction X reveals that they are punctuated by several plateaus ranging from 0.1 to 10 ns, as shown in the inset of **Figure 6.6** for (*S*)-BINOL1.



**Figure 6.6:** Displacement of the (*S*)-BINOL1 molecule along the X (top) and Z axis (bottom) over 650 ns. Plateaus are observed along the X displacement as pointed out in the inset by the grey arrows, with an associated Z displacement towards  $\pm 20$ -25 Å, *i.e.*, close to the peptoids at the silica surface (highlighted for sake of illustration at  $\sim 75$  ns by a green dashed rectangle).

These plateaus correspond to periods of time during which BINOL specifically interacts with peptoid(s) near the surface on the two sides of the pore (for  $|Z| > 20$  Å), as shown by the correlation between the plateaus along the X- and Z-displacement, highlighted by the green dashed rectangle in **Figure 6.6**. To better depict this interaction, we measured the distance between the center of mass of each peptoid amide units ( $-\text{C}(\text{O})-\text{N}(\text{R})-\text{CH}_2-$ , 6 per peptoid chain, with R the *N*spe side chain) and that

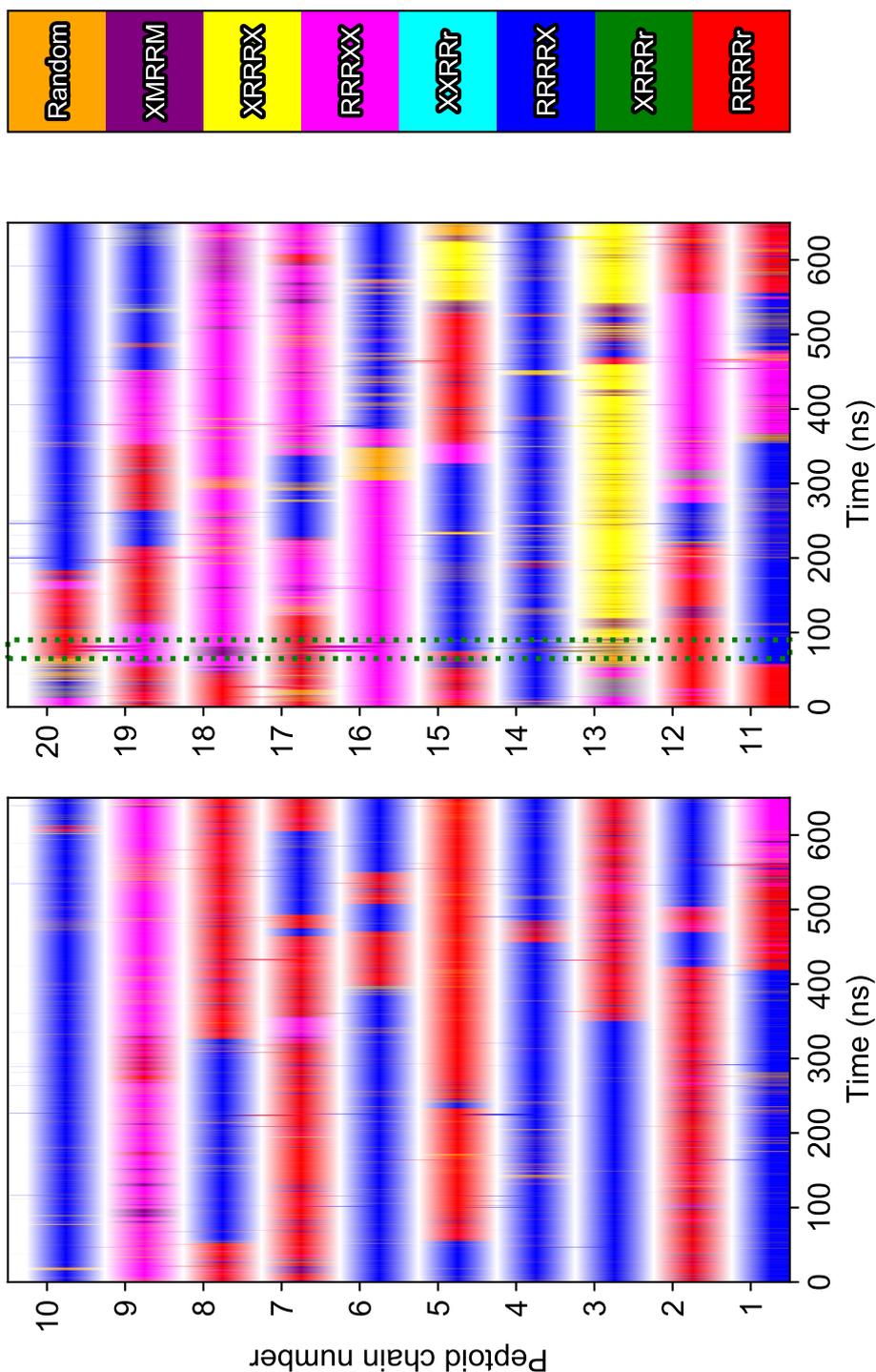
of the BINOL molecule along the trajectory, as schematized in **Figure 6.2B**. If this distance is smaller than 7 Å (the value computed for a van der Waals contact between a BINOL molecule and a peptoid residue in a right-handed helix conformation), we consider that a contact is established.<sup>c</sup> This procedure is carried out on each amide unit of every peptoid chain – BINOL pairs and averaged for each amide unit.



**Figure 6.7:** Number of contact events that occurred between (*R*)- or (*S*)-BINOL and *N*spe peptoids during the 650 ns SMD simulations. The peptoid backbone is broken down into 6 amide units from the N to C terminus. Most of the interactions occurred on the C terminus side.

This analysis indicates that: (*i*) BINOL interacts more often with the C terminus, *i.e.*, the most exposed side, than with the N terminus grafted on silica, and (*ii*) that (*R*)-BINOL molecules interact less frequently with the grafted peptoids than the (*S*)-BINOL (less contact events along the peptoid backbone). The results have been combined into conformational contact maps featuring the time-evolution of the conformation (color coded) of each individual peptoid chain, together with contact events (**Figure 6.8**). For example, the plateau along the X and Z coordinates at ~ 75 ns from **Figure 6.6** can be correlated to the interaction between (*S*)-BINOL1 and peptoid chains number 13, 16 and 19 mainly adopting a “RRRX” conformation (partial right-handed helix, with X a wildcard corresponding to any other letter of the code) during this contact event. The remaining contacts maps are available in Appendix C ((*S*)-BINOL, (*R*)-BINOL).

<sup>c</sup>Note that we count a single interaction when several residues of a given peptoid are simultaneously in contact.



**Figure 6.8:** Conformational contact maps of (*S*)-BINOL1 with each of the 20 peptoid chains. Each color corresponds to a given peptoid conformation represented by the sequences of letters in the legend, as described in **Tables 6.2** and **C.1**. For every chain number, two colored ribbons are shown, above and below the given number. The lower line depicts the evolution of the conformation along the SMD, while the upper one (mostly white) highlights the conformations occurring when the BINOL molecule is in contact with the peptoid chain. The green dotted rectangle corresponds to the plateau in the X and Z displacement from **Figure 6.6**

6.6

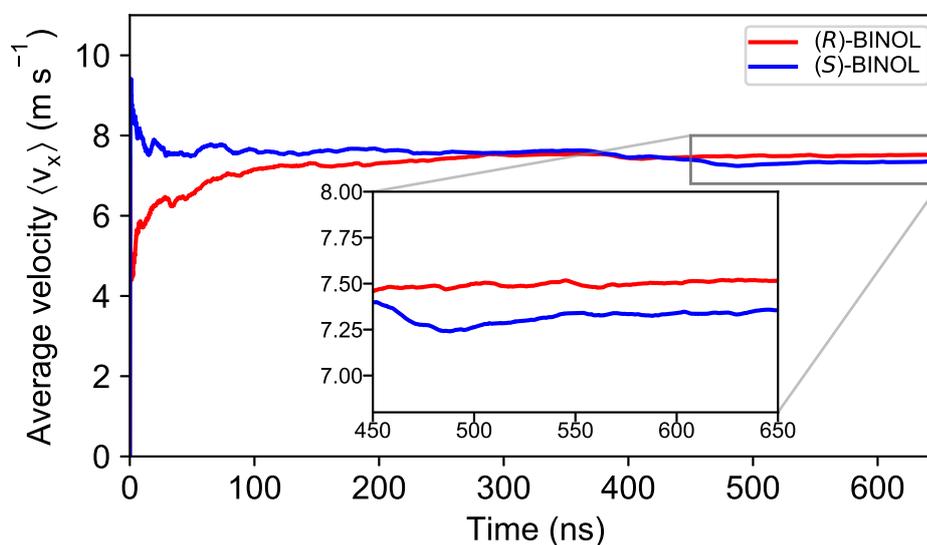
We next gathered the conformations of each peptoid chain during the contact events with BINOL molecules (for both simulations) to compute their relative abundance (Table 6.3).

**Table 6.3:** Main conformations adopted by  $N_{\text{spe}}$  peptoids only during contact events with ( $R$ )- and ( $S$ )-BINOL molecules. The “ $X$ ” character is used as wildcard and can be any of the previously defined letter ( $C, T, M, L$ ), except that in the current sequence.

Conformation	( $R$ )-BINOL (%)	( $S$ )-BINOL (%)
$RRRRX$	29	34
$RRRX$	34	25
$RRRRr$	16	25
<i>Random coil</i>	14	7

We found that the main conformations involved during contact events are all derived from a right-handed helix. In particular, the deviation from the right-handed helix mainly arises at the C terminus side (typically over the last two residues).

As stated in the methodology, we qualitatively characterize the enantioselectivity by comparing the average velocities of the BINOL enantiomers. We report the moving average velocity profiles for the 5 ( $R$ )- and ( $S$ )-BINOL molecules in Figure 6.9, that reach convergence after  $\sim 500$  ns.

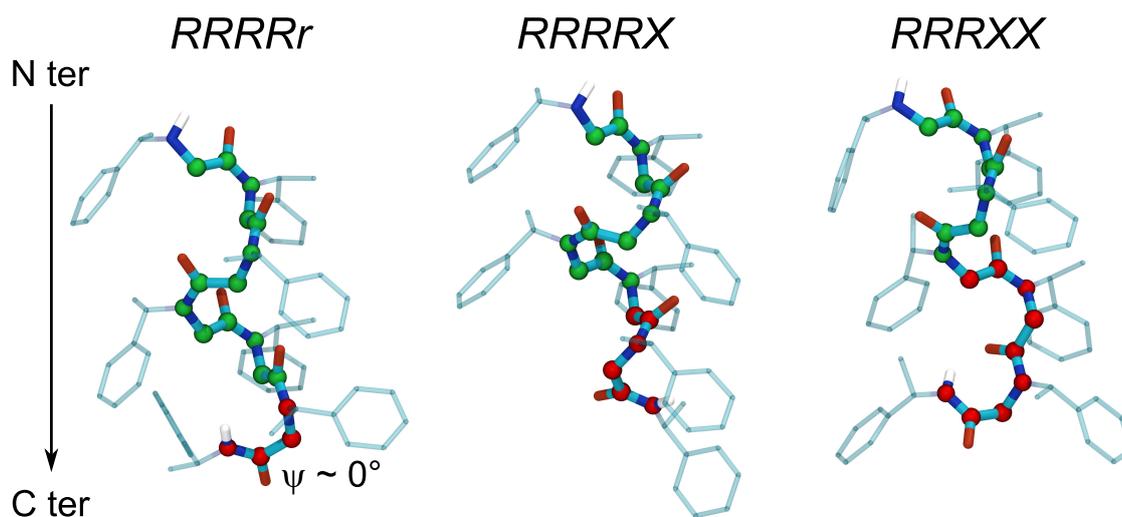


**Figure 6.9:** Evolution of the average velocity of ( $R$ )- and ( $S$ )-BINOL interacting with the  $N_{\text{spe}}$  peptoids (convergence is reached after 500 ns, as observed in the inset.)

( $R$ )-BINOL enantiomers are characterized by a higher average velocity compared to ( $S$ )-BINOL (7.515 vs. 7.354 m s<sup>-1</sup> respectively). Consequently, ( $S$ )-BINOL is characterized by a larger “elution time” in our virtual chromatographic column. This is

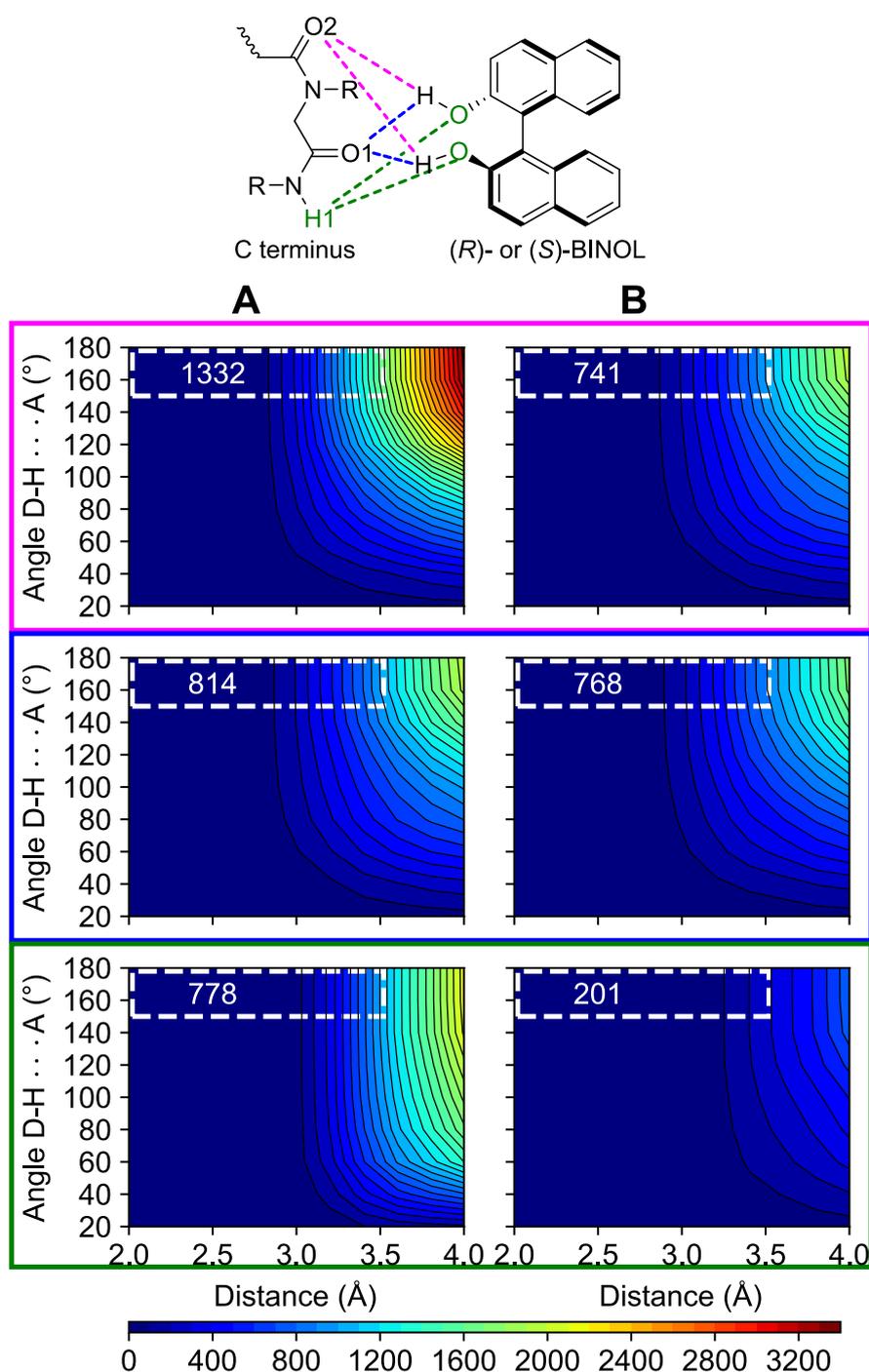
fully consistent with the experimental results obtained by Wu *et al.*<sup>24</sup> It is worth stressing that the specific values for the average velocities are tributary of the chosen pulling force and hence cannot be readily compared to experimental values, just as the experimental elution time depends on instrumental parameters such as the flux of solvent. However, the comparison of the velocity of the two enantiomers in the same theoretical (or experimental) conditions is meaningful.

The difference in the average velocities arises from a different occurrence of contact events experienced by the two enantiomers. The peptoid conformations involved during contact events are the same for both BINOL enantiomers, though their relative abundances are different (**Table 6.3**). (*R*)-BINOL interacts more frequently with helices largely deviating from the ideal right-handed helix (mainly helices of type “*RRRRXX*” and “*RRRRX*”, **Figure 6.10**) as well as with random coil conformations, while (*S*)-BINOL interacts more often with better defined helices (“*RRRRX*” and “*RRRRr*”). The structural rearrangement at the C terminus extremity may allow hydrogen bonds to develop with BINOL molecules, which is a hypothesis emitted by Wu *et al.*<sup>24</sup>



**Figure 6.10:** Examples of the most commonly encountered *N*spe peptoid conformations during contact events. The green beads represent the “*R*” residue conformation (right-handed helix), while the red beads represent the “*X*” residue conformation.

To assess this possibility, we computed the distribution of D-H···A distances and angles between donors (D, consisting in BINOL hydroxyl moieties and the –NH moieties at the peptoid C terminus) and acceptors (A, BINOL hydroxyl oxygen and peptoid amide oxygens, **Figure 6.11**).

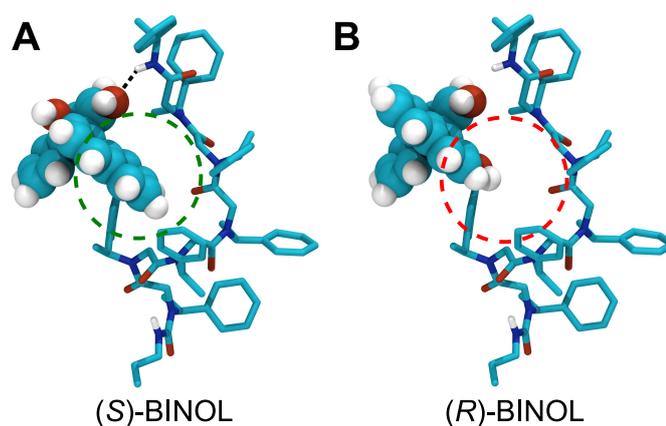


**Figure 6.11:** Distributions of angles and distances between donor (D), hydrogen (H) and acceptor (A) of BINOL-peptoid C terminus couples. The hydrogen bonds are formed between the (A) (S)- or (B) (R)-BINOL and the peptoids. Each hydrogen bond couple is represented by colored dashed bonds in the molecular representation and the corresponding heatmap is identified by its colored frame. The dashed white frame inside the distributions represent the common range used to characterize hydrogen bonds, according to Van Der Spoel *et al.* as well as the maximum number of hydrogen bonds obtained using the associated angle and distance.<sup>194</sup>

Clearly, hydrogen bonds can form for all D-H...A pairs, according to standard geometric criteria (distances lower than 3.5 Å, and angle higher than 150°<sup>194</sup>). Moreover, we do observe a clear difference when comparing the distribution of D-H...A distances and angles of the same donor-acceptor couples for (*R*)- and (*S*)-BINOL at the C terminus side of the peptoids (**Figure 6.11**). The probability of forming hydrogen bonds between (*S*)-BINOL and peptoids is much higher (in the range of 2 to 3.5 Å and angles comprised between 150° and 180°) than for (*R*)-BINOL for hydroxyl-amide O2, and N-terminus-hydroxyl oxygen bonds, while it is identical for the hydroxyl-amide O1 interaction.

The exposure of the amide hydrogen at the C terminus is therefore crucial to form hydrogen bonds, especially with (*S*)-BINOL. Moreover, the structural reorganization at the C terminus of the helix allows the good positioning of the penultimate amide oxygen to form another hydrogen bond with the (*S*)-BINOL, which appears to be the most probable one.

Besides forming hydrogen bonds, (*S*)-BINOL can interact through  $\pi - \pi$  interactions inside the cavities formed along the helix backbone, as suggested by Wu *et al.* (Figure 6.10).<sup>24</sup> In specific helical geometry of type “*RRRRX*” or “*RRRXX*”, a pocket is formed at the C terminus side of the peptoid into which the (*S*)-BINOL molecule geometry can fit adequately (Figure 6.12), while such behavior is barely observed along the trajectory for (*R*)-BINOL molecules. By swapping the (*S*)-BINOL in the complex displayed in Figure 6.12 with a (*R*)-BINOL, we observe that (*R*)-BINOL cannot fit equally well inside the cavity.



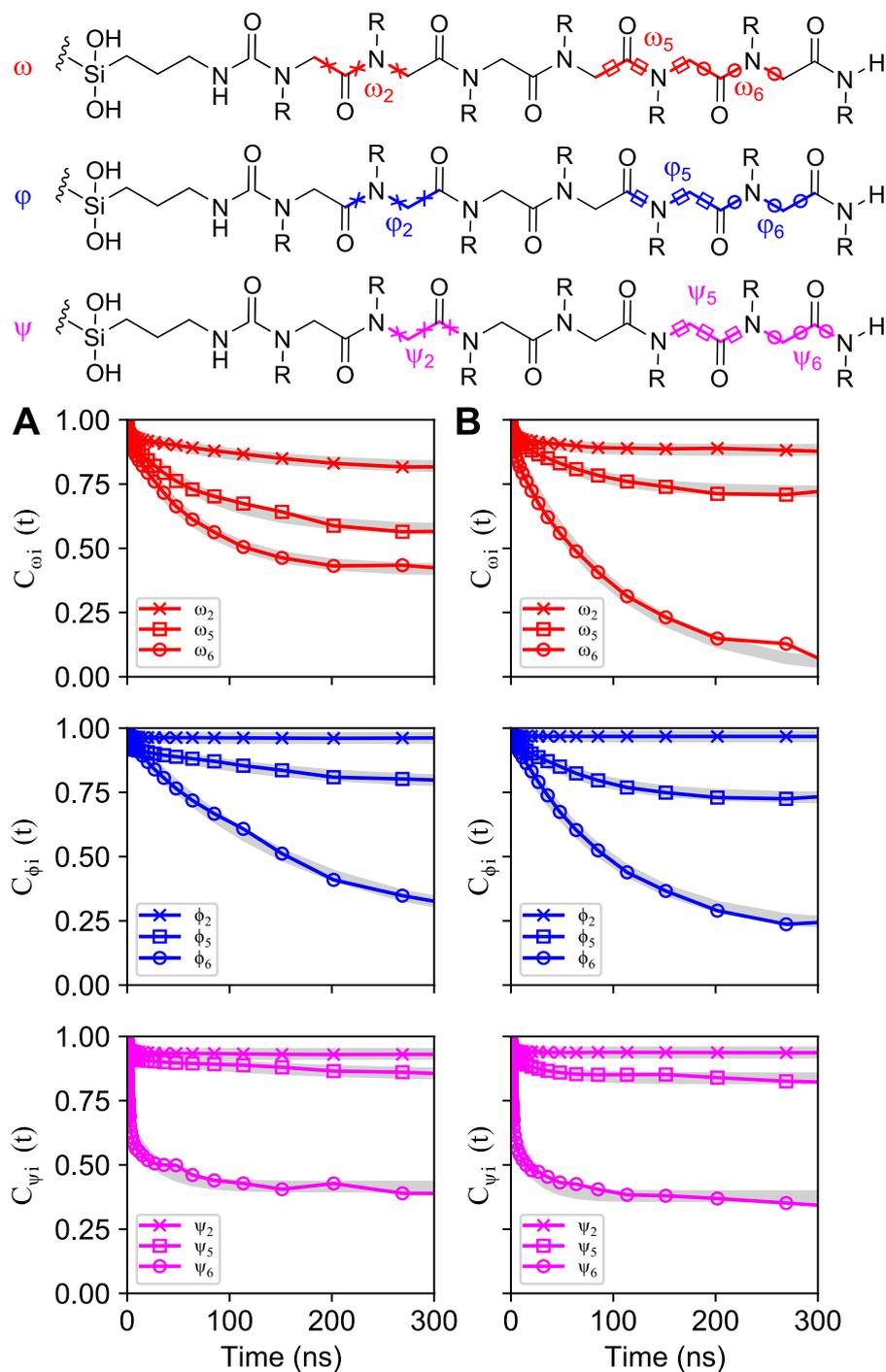
**Figure 6.12:** (A) Complex between the (*S*)-BINOL5 and the peptoid chain number 14 in a “*RRRXX*” conformation. In this configuration, a hydrogen bond is formed (black dashed line) between the amide hydrogen of the peptoid and the alcohol oxygen of (*S*)-BINOL when inserted into the pocket formed at C terminus (green dashed circle). (B) The coordinates of (*S*)-BINOL from the complex (A) were altered to change its configuration to (*R*)-BINOL. The molecule does not fit anymore into the pocket formed by the partial helix (red dashed circle).

The consequences of these complex interactions also emerge from the compared analysis of the autocorrelation function of the dihedral angles:<sup>229</sup>

$$C_{\alpha}(t) = \langle \cos \alpha(0) \cdot \cos \alpha(t) \rangle + \langle \sin \alpha(0) \cdot \sin \alpha(t) \rangle \quad (6.1)$$

where  $\alpha$  corresponds to one of the main peptoid backbone dihedrals ( $\omega$ ,  $\phi$  or  $\psi$ ).

We observe a major difference in the behavior of the  $\omega$  dihedral located at the C terminus extremity when comparing the autocorrelation functions of (*R*)- and (*S*)-BINOL simulations (Figure 6.13).



**Figure 6.13:** Autocorrelation functions related to backbone dihedrals for  $N_{spe}$  peptoids involved in the simulations with (A) (*R*)- and (B) (*S*)-BINOL. All functions were fitted by a bi-exponential decay function (grey lines). The position of the plateau depends on the location of the dihedral along the backbone (above the plots), which is higher when the dihedral is located at the N terminus, *i.e.*, in the sterically hindered region close to the silica surface. Differences between the two enantiomers appear mainly in the  $\omega$  profiles, in particular for  $\omega_6$  located at the C terminus extremity.

The autocorrelation functions can also provide the characteristic times for the fluctuations around a given conformer ( $T_1$ ) and the conformational rearrangement ( $T_2$ ). These values are obtained by fitting the autocorrelation functions using a bi-exponential decay function:

$$f(x) = A_0 + \left(\frac{A_1}{A_1 + A_2}\right)e^{-x/T_1} + \left(\frac{A_2}{A_1 + A_2}\right)e^{-x/T_2} \quad (6.2)$$

The total time is given by:

$$\left(\frac{A_1}{A_1 + A_2}\right)T_1 + \left(\frac{A_2}{A_1 + A_2}\right)T_2 \quad (6.3)$$

In the case of (*R*)-BINOL,  $\omega$  shows a much shorter correlation time for the conformational rearrangement than (*S*)-BINOL, which we attribute to a faster re-orientation of the C terminus extremity of the peptoids. Indeed, the hydrogen bonds between (*S*)-BINOL and the *Nspe* peptoids are stronger, which most likely locks the conformation and reinforces the interaction.

**Table 6.4:** Fit parameters from the autocorrelation functions of *Nspe* peptoids with either (*R*)- or (*S*)-BINOL as well as characteristic times.

Dihedral	$A_0$	$A_1$	$A_2$	$T_1$ (ps)	$T_2$ (ps)	$T_{total}$ (ps)
<i>Nspe</i> – ( <i>R</i> )-BINOL						
$\omega_2$	0.8067	0.7083	0.2916	124162	530	17027
$\omega_5$	0.5668	0.8592	0.1407	78367	591	29202
$\omega_6$	0.4125	0.8861	0.1138	65794	532	34287
$\psi_2$	0.962	0.7532	0.2467	248	10136	102
$\psi_5$	0.7758	0.3057	0.6942	568	150004	23383
$\phi_6$	0.192	0.0441	0.9558	432	168342	130012
$\psi_2$	0.9319	0.6979	0.302	327	13597	295
$\psi_5$	0.8452	0.5076	0.4923	579	152233	11643
$\psi_6$	0.4154	0.5798	0.4201	868	27562	7064
<i>Nspe</i> – ( <i>S</i> )-BINOL						
$\omega_2$	0.8826	0.51416	0.4858	40816	436	2487
$\omega_5$	0.7188	0.8003	0.1996	64987	414	14645
$\omega_6$	0	0.894	0.1059	106707	968	95494
$\phi_2$	0.9676	0.9041	0.0958	255	10284	39
$\phi_5$	0.7268	0.1557	0.8442	371	71114	16416
$\phi_6$	0.2088	0.0441	0.9558	463	97387	73653
$\psi_2$	0.9375	0.7721	0.2278	313	10004	157
$\psi_5$	0.8369	0.462	0.5379	430	33730	2990
$\psi_6$	0.3783	0.5277	0.4722	750	19907	6091

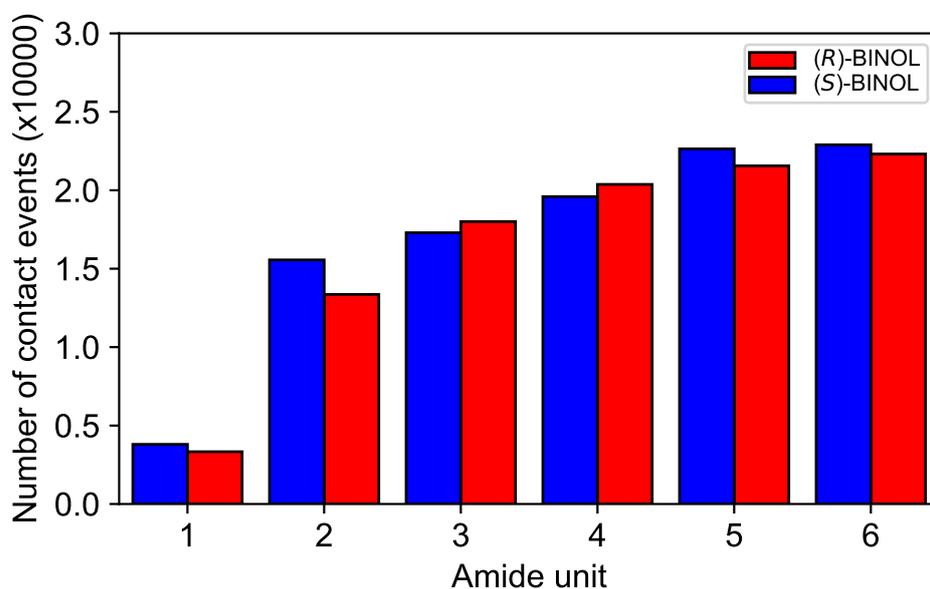
### 6.4.2 Achiral peptoids: validation of the model

In order to substantiate the results for the  $N_{spe}$  peptoid-based chiral stationary phase discussed in the previous section, we performed additional simulations using the same conditions (grafting density, peptoid oligomer length, solvent composition, etc.) on a model system made this time of achiral peptoids, which in principle should not exhibit any enantioselectivity.

We chose sarcosine, the simplest possible peptoid residue bearing a methyl side chain, as a building block for an achiral hexamer (see **Scheme 6.1**, with R replaced by  $-\text{CH}_3$ ). As discussed in Chapter 5, sarcosine peptoids are known to behave as flexible polymers and do not adopt any particular conformation in a variety of solvents (polar and apolar); the main conformation type is thus labeled as “*random coil*”.<sup>11,117</sup>

The graphical analysis of the **displacements** of the (*R*)- and (*S*)-BINOL along the X and Z axes during the simulations with the  $N_{sar}$  stationary phase, as well as the **conformational contact maps** are reported in Appendix C.

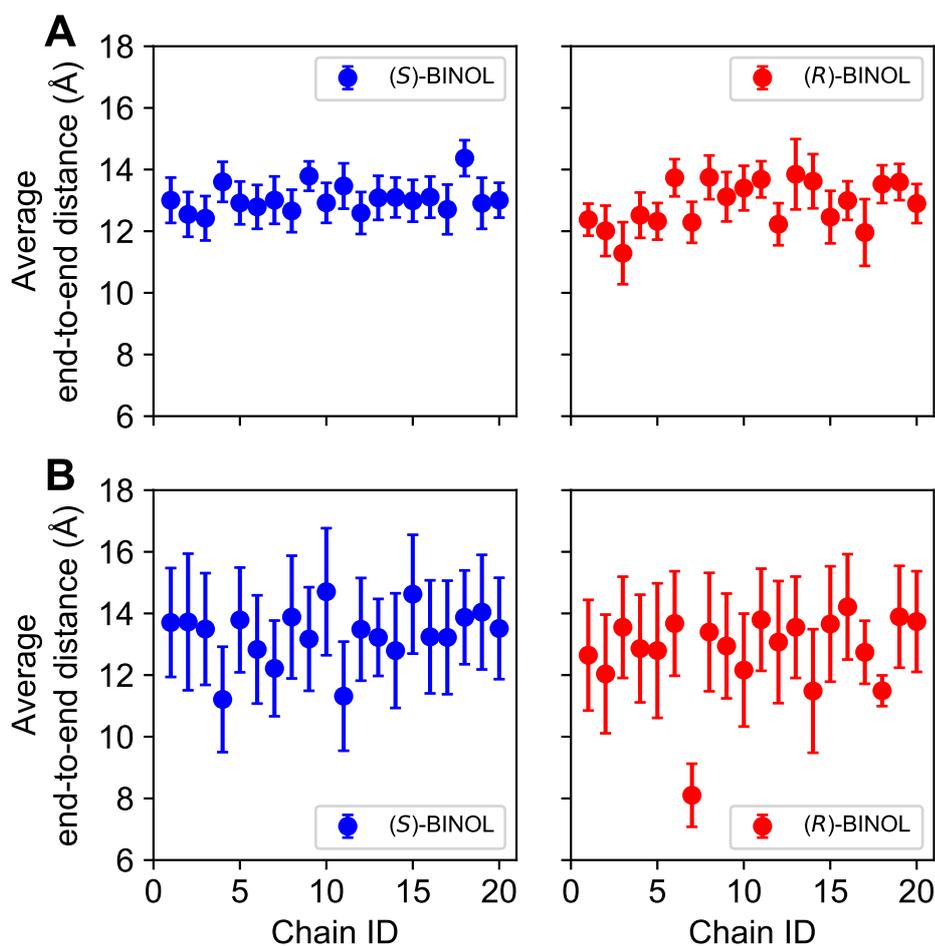
Unlike with  $N_{spe}$  hexamers, BINOL enantiomers interact more homogeneously with the residues of the sarcosine chains (except with the hardly accessible N terminus grafted on the silica substrate, **Figure 6.14**).



**Figure 6.14:** Number of contact events that occurred between (*R*)- or (*S*)-BINOL and  $N_{sar}$  peptoids during the 500 ns SMD simulations. The peptoid backbone is broken down into 6 amide units from the N to C terminus. The guest molecules interact more homogeneously with the amide units along the backbone than with  $N_{spe}$  peptoids.

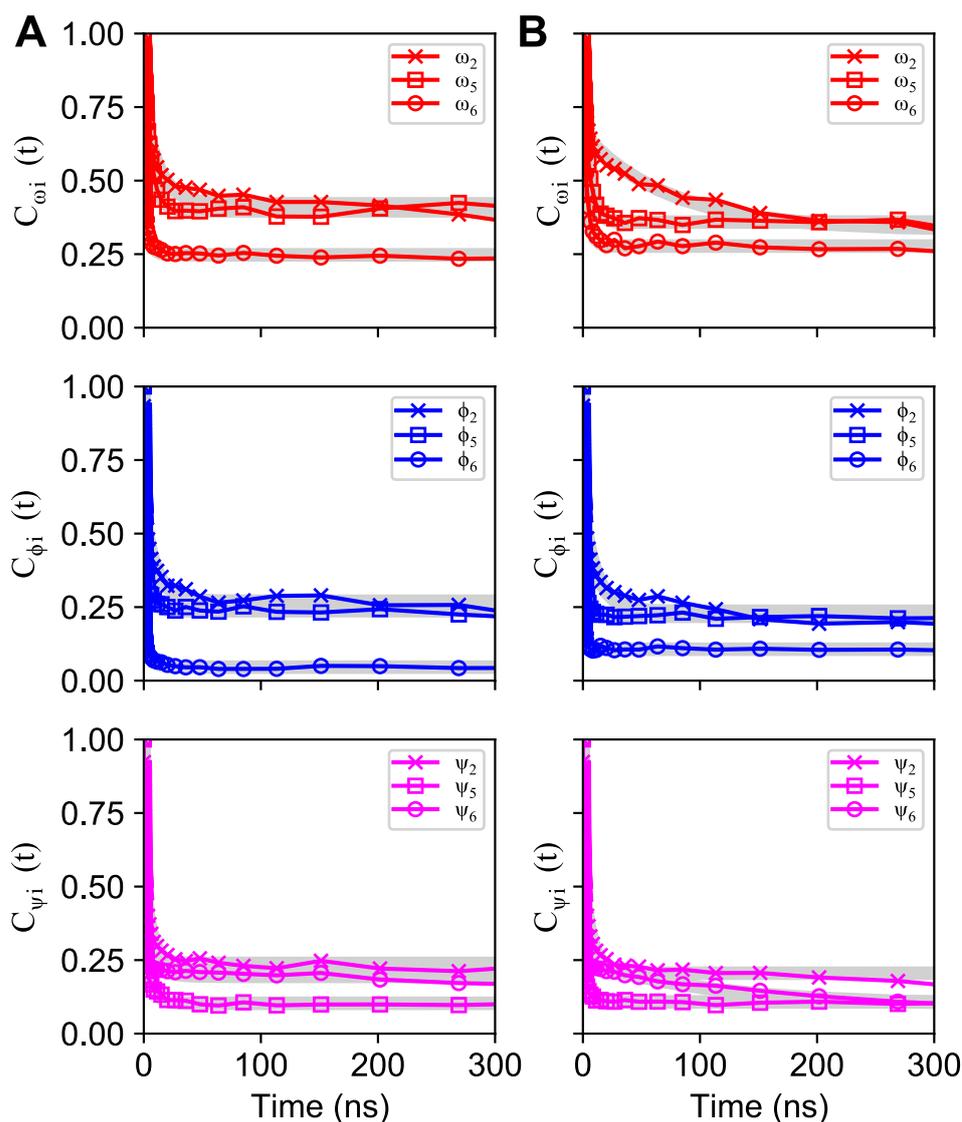
This can be rationalized by observing that, as expected, sarcosine oligomers do not assume chiral conformations. More than 95% of their conformations occurring during

contact events (but also outside these contacts) are “*randomcoil*” conformations. Sarcosine peptoids are thus far more flexible and less sterically hindered than *Nspe* peptoids, which were characterized as mainly helical-like. We assessed the flexibility of both types of peptoids by measuring the average end-to-end distances of each chain and compared them to the *Nspe* peptoids chains.



**Figure 6.15:** Average end-to-end distances (between terminal nitrogen atoms) of each peptoid chain ((A) *Nspe* and (B) *Nsar*) and their standard deviation during the SMDs with both BINOL enantiomers. Larger standard deviations (vertical bars) are observed for *Nsar*, indicating a higher flexibility.

We also computed the dihedral autocorrelation functions as previously done with the *Nspe* peptoids (Figure 6.16).



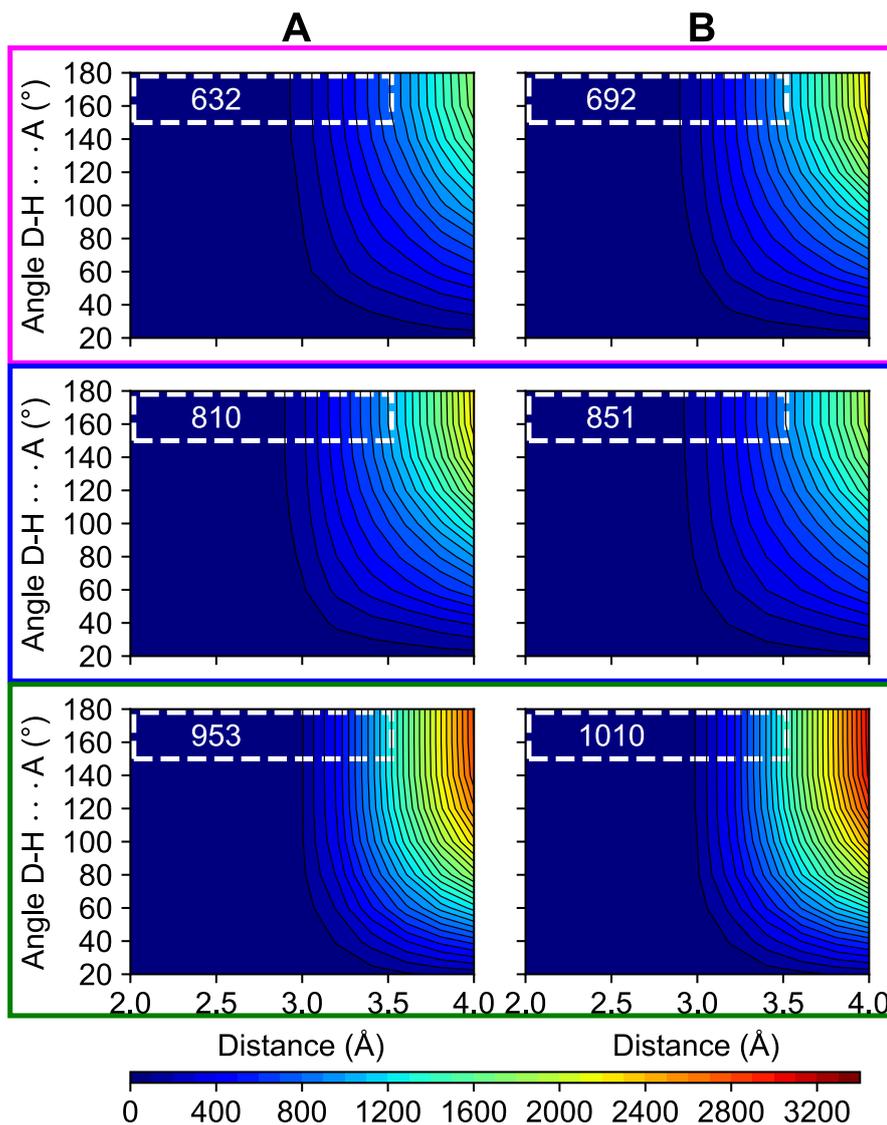
**Figure 6.16:** Autocorrelation functions related to backbone dihedrals for *Nsar* peptoids involved in the simulations with (A) (*R*)- and (B) (*S*)-BINOL. All functions were fitted by a bi-exponential decay function (grey lines).

The higher flexibility of *Nsar* is reflected both by the larger variations in the average end-to-end distances for sarcosine chains along the dynamics (**Figure 6.15**) and by the dihedral autocorrelation functions that indicate fast conformational rearrangements (characteristic timescale of about 10 ns, **Table 6.5**), compared to the *Nspe* peptoids that can reach values up to 150 ns.

**Table 6.5:** Fit parameters from the autocorrelation functions of *Nsar* peptoids with either (*R*)- or (*S*)-BINOL as well as characteristic times.

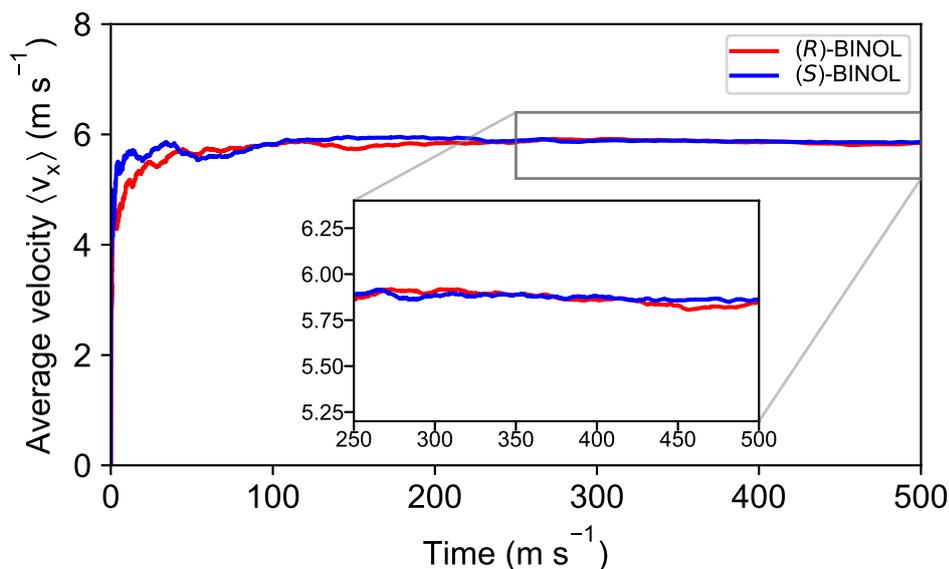
Dihedral	$A_0$	$A_1$	$A_2$	$T_1$ (ps)	$T_2$ (ps)	$T_{total}$ (ps)
<i>Nspe</i> – ( <i>R</i> )-BINOL						
$\omega_2$	0.3323	0.4911	0.5088	73900	1506	24771
$\omega_5$	0.359	0.8265	0.1734	5079	491	2745
$\omega_6$	0.2774	0.694	0.3059	3630	804	1998
$\psi_2$	0.2358	0.5401	0.4598	728	11838	4461
$\psi_5$	0.219	0.73	0.2699	444	3103	907
$\psi_6$	0.1069	0.8307	0.1692	339	1724	512
$\Psi_2$	0.2054	0.5893	0.4106	619	7718	2808
$\Psi_5$	0.1077	0.8163	0.1836	419	3281	843
$\Psi_6$	0.0969	0.8315	0.1684	295	108280	16689
<i>Nspe</i> – ( <i>S</i> )-BINOL						
$\omega_2$	0.4209	0.4364	0.5635	831	13265	4538
$\omega_5$	0.3973	0.1946	0.8053	525	6120	3032
$\omega_6$	0.2476	0.7511	0.2488	374	4117	982
$\phi_2$	0.2699	0.583	0.4169	732	10696	3567
$\phi_5$	0.2378	0.7493	0.2506	455	4699	1158
$\phi_6$	0.0459	0.8975	0.1024	273	5538	775
$\Psi_2$	0.2386	0.5873	0.4126	558	6106	2168
$\Psi_5$	0.1031	0.7986	0.2013	435	6413	1470
$\Psi_6$	0.1947	0.9181	0.0818	264	12270	1004

Although sarcosine oligomers behave as random coils, they can still form hydrogen bonds with BINOL enantiomers. However, compared to the chiral *Nspe* peptoids for which the hydrogen bond geometrical parameters are different between the two enantiomers, the probability distributions are very similar for *Nsar* with both BINOL enantiomers (**Figure 6.17**) and do not show any evidence of chiral interaction.



**Figure 6.17:** Distributions of angles and distances between donor (D), hydrogen (H) and acceptor (A) of BINOL-*N*sar peptoid C terminus couples (same definitions as in **Figure 6.11**). The hydrogen bonds are formed between the (A) (*S*)- or (B) (*R*)-BINOL and the peptoids. The dashed white frame inside the distributions represent the common range used to characterize hydrogen bonds, according to Van Der Spoel *et al.* as well as the maximum number of hydrogen bonds obtained using the associated angle and distance that are very similar between the two simulations.<sup>194</sup>

As a result, no difference is observed in the average velocities of (*S*)- versus (*R*)-BINOL, owing to the achiral nature of the grafted chains, as expected for such a system (**Figure 6.18**).



**Figure 6.18:** Evolution of the average velocity of (*R*)- and (*S*)-BINOL interacting with the *Nsar* peptoids (convergence is reached after 250 ns, as observed in the inset.)

### 6.4.3 Conclusions

Thanks to the collaboration with Prof. Luca Muccioli, we established a computational methodology, based on steered molecular dynamics simulations, for reproducing a liquid chromatography experiment. We successfully applied it to evaluate the enantioselective properties of chiral *Nspe* peptoids against BINOL enantiomers at the atomistic level.

Our results show that, during the elution process, (*S*)-BINOL is retained for a longer time than (*R*)-BINOL in our model chromatography column, due to more favorable interactions with *Nspe* peptoids in conformations derived from a right-handed helix. The geometric deviation from the perfect right-handed helix that arises on the C terminus side promotes the exposure of the amide hydrogen of *Nspe* peptoids and triggers more frequently the formation of hydrogen bonds with (*S*)-BINOL. Moreover, the C terminus can form a groove where the (*S*)-BINOL can fit and bind through  $\pi - \pi$  interactions. These observations are clearly congruent with the initial hypotheses emitted by Wu *et al.* that constituted the basis of our work. It clearly emphasizes that the C terminus extremity plays a pivotal role in the chiral recognition of binaphthyl derivatives, although the origin of the chiral selectivity lies in the right-handed conformation of *Nspe* peptoids.

In a more general perspective, our study delineates how important is the proper tuning of the chemical sequence and hence the secondary structure, especially of the

C terminus side, since it could greatly influence the selectivity of peptoids towards a variety of different chiral compounds. It also clearly demonstrates that peptoids are particularly relevant for this type of application thanks to their high tunability and the development/improvement of the computational tools to predict their conformational arrangements and their tailoring for a given application.

---

## Summary and perspectives

---

Nowadays, computational approaches are increasingly employed not only to support experimental evidences but also to predict the properties of new systems. Peptoids are no exception to the rule. Throughout this thesis, we sought to explore the conformational behavior of peptoids experimentally and rationalize the observations by taking advantage of the molecular modeling techniques, in particular molecular dynamics simulations.

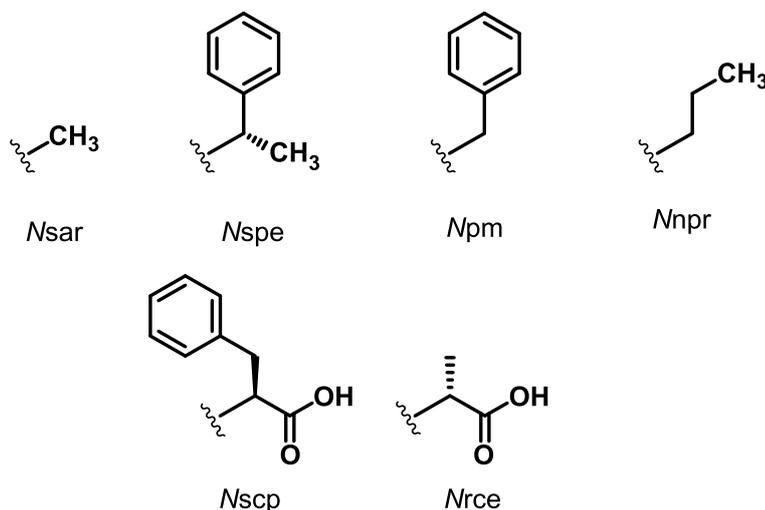
The original goal was to characterize peptoids by ion mobility mass spectrometry, a technique scarcely used to date in the field of peptoids, and to envisage whether the relationship between the primary and the secondary structures may be probed using MS. Depending on the side chain borne by the backbone amides, peptoids can adopt well-defined secondary structures, either in solution or in solid state. The studies carried out in solution mainly rely on CD spectroscopy, that points to the presence of chiral conformation(s).<sup>41,42,48</sup> NMR studies point out that not a single chiral conformer is present in solution, but rather a heterogeneous mixture of conformer families, all contributing to the CD signal.<sup>47</sup> Therefore, we were motivated to employ IM-MS to probe the different conformer families of peptoids by transferring them into gas phase. The measured collision cross sections are representative of the 3D arrangements and can be associated to computed collision cross sections from candidate geometries generated by molecular modeling.

The first part of this thesis concerned the development of a suitable computational tool to describe the conformational behavior of peptoids to be further compared to experimental results. Given the large degrees of freedom of increasingly longer peptoids, we chose to describe peptoids using molecular dynamics simulations. While many force fields exist to describe peptides, only a few were developed for peptoids given the relative youth of the field. We thus decided to develop a set of parameters for peptoids based on the DREIDING force field, that is one of the only force field in which hydrogen bonding interactions are considered explicitly.<sup>139</sup> The reparametrization of force fields is typically based on quantum-mechanical calculations of model compounds representative of the system.

Prior to developing new parameters for peptoids, we investigated the reliability of non-bonding parameters, in particular the van der Waals parameters (equilibrium

distance and well depth) of aliphatic hydrogens that were reported to be inadequate to accurately describe several physical properties such as the density.<sup>145,146</sup> The equilibrium distance was parametrized against experimental collision cross sections of different polymers, while the well-depth was parametrized against experimental vaporization enthalpies of multiple solvents.

Afterwards, we performed quantum-mechanical calculations at the HF-MP2 level on model peptoids and found out that the “hard” bonded terms, *i.e.*, bond stretching and angle bending, were well described with the default parameters. However, the default dihedral parameters did not produce the correct dihedral energy profiles. Peptoids are characterized by two types of dihedrals: (*i*) three backbone dihedrals ( $\omega$ ,  $\varphi$  and  $\psi$ ) and (*ii*) up to two dihedrals to connect the side chain to the backbone ( $\chi_1$  and  $\chi_2$ ). We reparametrized these two types of dihedral independently, allowing to add any new side chain into the force field. Initially, we introduced the *Nsar*, *Nspe*, *Npm* and *Nnpr* side chains. The new set of parameters was validated by generating Ramachandran-like plots that represent the potential energy surface resulting from the variation of two dihedral angles. Our data are in very nice agreement with those generated by Butterfoss *et al.* and Renfrew *et al.* at the quantum-mechanical level,<sup>30,76</sup> demonstrating that our strategy is well suited to describe the conformational behavior of peptoids. Moreover, new side chains can be easily integrated because of the backbone–side chain decoupling strategy that we employed.



**Scheme 7.1:** Chemical structures of the side chains studied and reparametrized in the DREIDING force field.

The second part of this thesis was related to the investigation of the gas-phase conformation of peptoids by IM-MS. Peptoids are constituted of repeating building blocks, and can *de facto* be considered as polymers. Synthetic polymers occupy a

unique place in the field of IM-MS. Due to their intrinsic chain dispersity, they offer a broad range of homologous ions with different lengths. Their gas-phase conformation can be deduced from the evolution of the collision cross section with increasing chain length (or mass), which can be fitted by an equation of the form  $\Omega = A.M^B$ , where the value  $B$  is an indicator of the shape of the ions. The reference value is  $2/3$ , and corresponds to a spherical shape. Lower values are indicative of conformations more folded than a sphere, while larger values are indicative of conformations more extended. However, values other than  $2/3$  are strictly qualitative. Indeed, in the present study, we deal with helical geometries, that can be approximated by cylinders, whose radius will depend on the nature of the side chain and the combination of the backbone dihedrals, while the height is governed by the number of repeating units. We have demonstrated in a geometrical way that the  $B$  parameter obtained by fitting the evolution of  $\Omega$  as a function of the mass is dependent on the range of repeating units. The value tends to 1 only when helices are infinitely long, which will never happen experimentally. For moderately long helices (and depending on the geometry and side chains), the values can vary a lot between  $2/3$  and 1, which prevents to confidently assign a particular conformation. Therefore, such fitting procedure should always be supported by molecular modeling, provided the system is accurately described, for instance using a reparametrized force field.

The investigation of the gas-phase conformation of singly protonated peptoids started using two distinct peptoids:  $Nsar$  and  $Nspe$ . These two peptoids behave differently in solution.  $Nsar$  is known to form random coils,<sup>11,117</sup> while  $Nspe$  is famous for mainly inducing helical conformations in a large variety of solvents, as attested by CD spectroscopy experiments.<sup>41,42,47,85</sup> Different conformers are thus present in solution, but cannot be easily discriminated. IM-MS appears to be an elegant method to tackle this problem, and it is expected that these two types of peptoid would also behave differently in gas phase. Interestingly, we obtained similar results regarding their conformations. In both cases, the most stable conformers adopt a compact loop-like shape, where most of the amides are in *trans* conformation (while they are mostly in *cis* in solution), allowing the amide dipoles to point toward the proton borne by the terminal ammonium, very similar to the reported “threaded loop” solution conformer.<sup>93</sup> For longer chains ( $n > 10$ ), the additional  $Nspe$  residues start organizing in a helical fashion after the ammonium is fully stabilized. The gas-phase conformation is thus obviously different from their solution phase conformation due to the presence of the charge that must be stabilized. Therefore, the presence of a charge on the peptoid in a vacuum-like environment dictates its conformation, preventing ion mobility mass spectrometry to identify the different conformers present in solution.

Nonetheless, IM-MS still proves relevant in this context, because it sheds light on the parameters that affect the conformational stability and rigidity of peptoids. The secondary non-covalent interactions present in *N*spe peptoids are not strong enough to prevent the backbone to wrap around the charge. To tackle this issue, we took inspiration from peptides and introduced the *N*scp side chain that can not only form hydrogen bonds but is also bulky and chiral, which are two requirements to obtain helices in solution.<sup>42</sup> Our experimental and theoretical data demonstrate that these peptoids adopt a helical conformation that arises from the compromise between maximizing charge coordination and minimizing the disruption of the hydrogen bonds. The ammonium is thus locally stabilized by the surrounding amide carbonyls (as well as the side chain carbonyls), but also by the macrodipole formed by the individual amide dipoles along the backbone that point in the same direction. If we decrease the side chain bulkiness by changing the side chain to a *N*rce side chain (from a benzyl to a methyl), the most stable conformers are rather loop-like/spherical-like, which demonstrates that bulkiness is a relevant parameter to obtain helices, either in solution or gas phase. These results are currently submitted for publication in *Biomacromolecules*. The solution conformers of *N*scp peptoids have not been yet unraveled. However, further experimental studies supported by computational chemistry, especially Time Dependent-DFT, would help to decipher their conformations based on their CD signatures. Such studies could reveal that our gas phase conformers are actually very similar to those in solution. Moreover, the group of **Valérie Gabélica** recently developed a new MS technique that provides CD spectra in gas phase.<sup>230</sup> This technique would be particularly interesting for peptoids to correlate the gas phase CD signature to that in solution.

Recently, the group of **Claude Taillefumier** investigated non-aromatic bulky chiral side chains ((*S*)-*N*-*tert*-butylethyl) that form helices in the solid state,<sup>231</sup> and also display chiral folding in solution according to their CD spectra. They demonstrate that these peptoids form C–H···O between the side chain and the amide backbone as well as London interactions between *tert*-butyl moieties of the side chains.<sup>232</sup> These new peptoids could be very interesting to investigate as their solution conformation could be conserved upon transfer in gas phase.

As a major conclusion of our studies, these gas-phase investigations really emphasize that the structural properties of peptoids can be very dependent on their environment.

The last part of this thesis was related to the application of peptoids in the context

of chiral separation. Our work is based on experimental evidences reported by Wu *et al.* who showed that *Nspe* peptoids grafted on silica in chromatography columns were able to separate enantiomers of a series of binaphthyl derivatives.<sup>24</sup> They hypothesized that the helical conformation of *Nspe* peptoids was responsible for the enantioselectivity, as well as the formation of hydrogen bonds. In this research work, we developed a methodology to provide an atomistic model of the recognition process and hence to shed light on the key parameters that rule the enantioselectivity. We focused our attention on the 2,2'-bihydroxy-1,1'-binaphthyl (also called BINOL), whose enantiomers were efficiently separated by *Nspe* oligomers.

We modeled a section of a silica bead pore on which we grafted *Nspe* peptoid hexamers (described using our force field parameters) and took advantage of the steered molecular dynamics procedure to mimic the elution process. To do so, we applied a constant pulling force on the guest chiral molecules, *i.e.*, the BINOL molecules, and monitored the evolution of their velocities. *In fine*, the (*S*)-BINOL guests were slower than the (*R*) enantiomers, which is in full consistency with the experimental results. During the simulations, we also monitored the conformations of the individual chains between and during contact events with the BINOL. Interestingly, the main conformations involved during contact events are derived from the reported right-handed helix,<sup>41</sup> which supports the hypothesis emitted in the experimental study. Moreover, hydrogen bonds are formed between the guests and the peptoids (at the C terminus extremity), and are stronger with the (*S*) enantiomer based on the actual geometry of the hydrogen bonds (distance between donor and acceptor and angle formed by the triad). This once again confirm the other hypothesis emitted by Wu *et al.*

In order to highlight the importance of the conformations adopted by the peptoids for the recognition process, we performed the same simulations by changing *Nspe* to *Nsar* chains. These peptoids mainly adopt random coil conformations during the simulations and did not induce significant difference in the average velocities of both BINOL enantiomers. Our study demonstrates once again the importance of the relationship between the peptoid sequence and their conformation. An improvement of the enantioselectivity properties of peptoids could be reached by designing a peptoid that would adopt a helical conformation and also bear other chemical moieties that could be specific toward certain analytes, for example, at the C terminus extremity.<sup>233</sup>

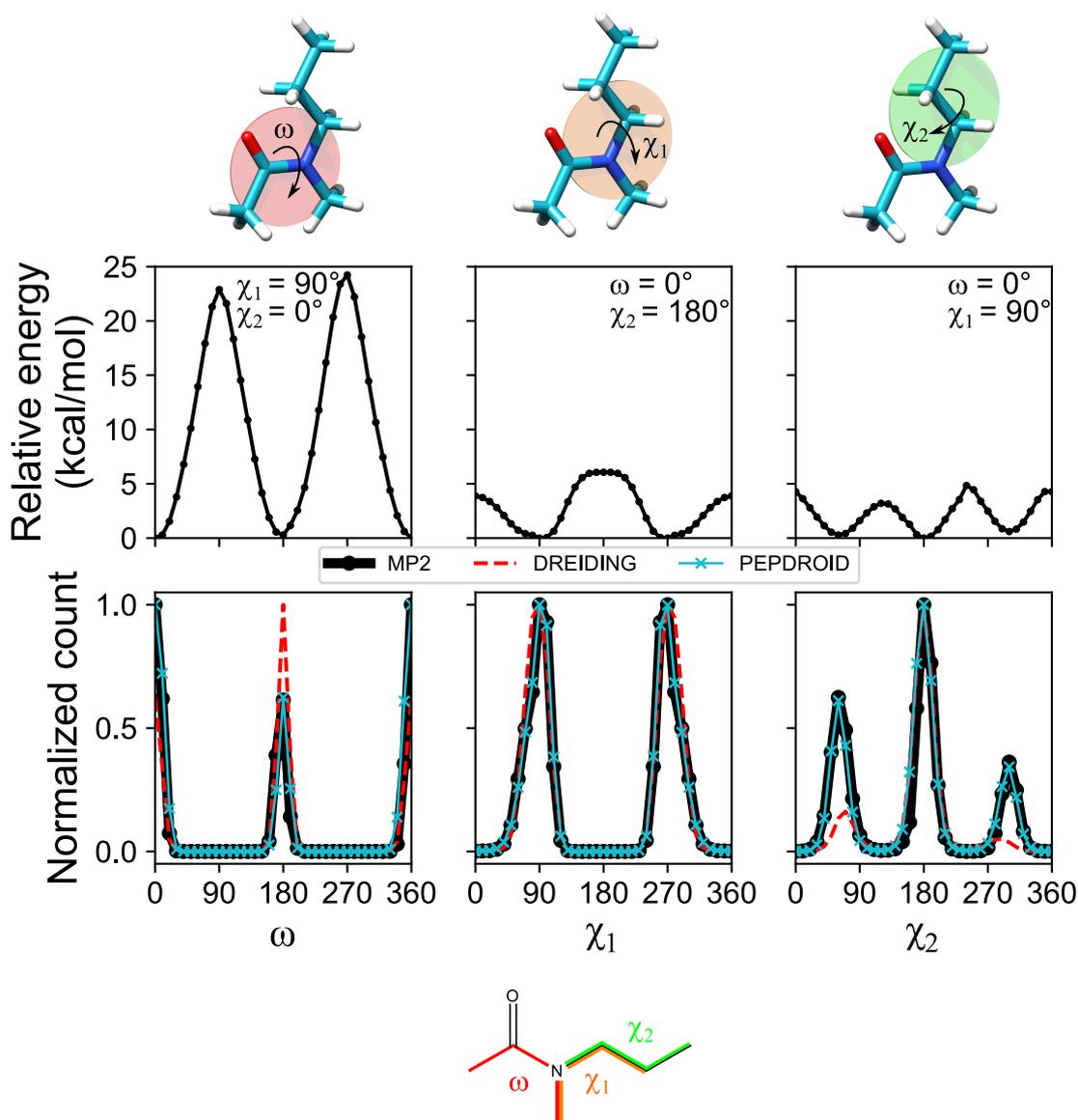
The main drawback of linear peptoid chains in the context of chiral recognition is the dependence of their conformation on the nature of their environment. Over the

past ten years, cyclic peptoids emerged thanks to the pioneering work of Kirshenbaum *et al.*<sup>57</sup> Their conformation is “simpler” than the traditional peptoids and can be controlled upon complexation with other compounds, typically cationic metals.<sup>234</sup> These peptoids could be envisaged as chiral selectors that would form inclusion complexes with guest molecules.

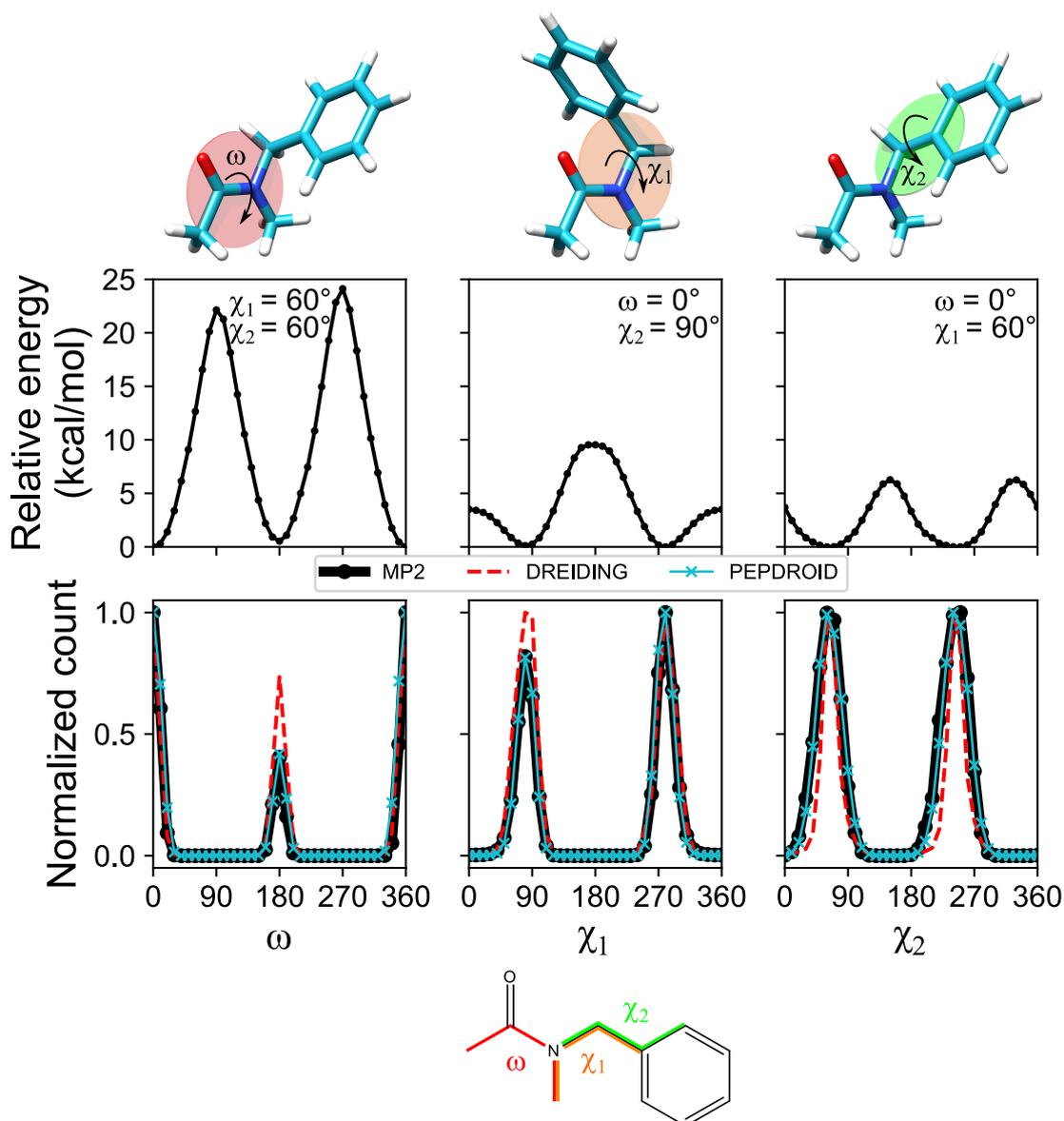
Because the combinations of side chain, sequence and architectures of peptoids are beyond what we can imagine and currently build, I strongly believe that the peptoid field will live a long life and is only at its infancy. The development of robust computational methods will most certainly help to enhance the discovery of relevant side chain combinations to promote conformers and sequences that would be significant for particular applications. However, it does not imply that currently studied side chains should be forgotten. Understanding the origin of the CD signature of chiral peptoids, such as those bearing *N*spe side chains, would directly provide insights about the secondary structure of the peptoids. To date, only a few attempts have been reported to predict and compare experimental and theoretical CD spectra using TD-DFT.<sup>235,236</sup> Application of this technique combined with molecular dynamics could prove very helpful in further validating the current accepted conformational behavior of peptoids.

## Additional Material of Chapter 4

### A.1 Dihedral profiles of peptoid side chains

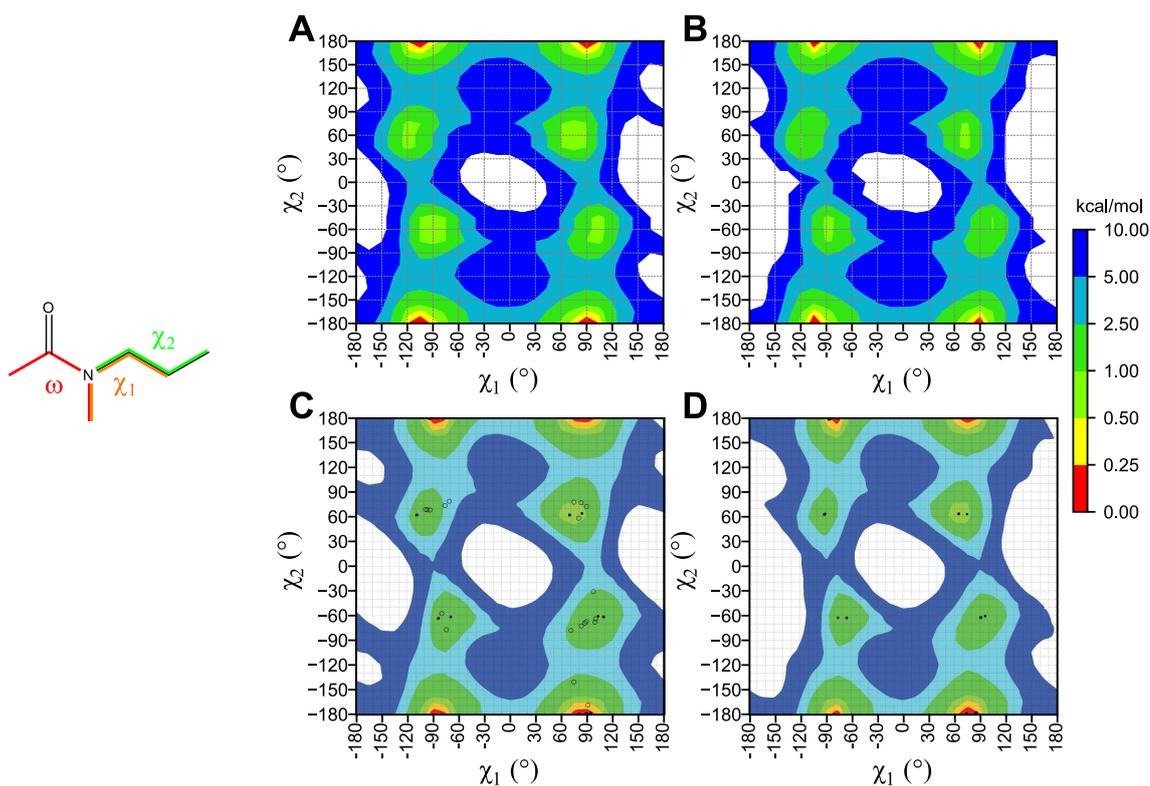


**Figure A.1:** Energy profiles of the  $Nnpr$  side chain dihedrals obtained at MP2/cc-pVDZ level (top) and normalized population count (at 298 K) obtained with MP2/cc-pVDZ (black), default DREIDING (red) and using the new sets of parameters PEPDROID (blue) (bottom). When applied, the restraints are displayed in the energy profiles. The primary structures of the models are displayed below the plots with the definition of the dihedrals.

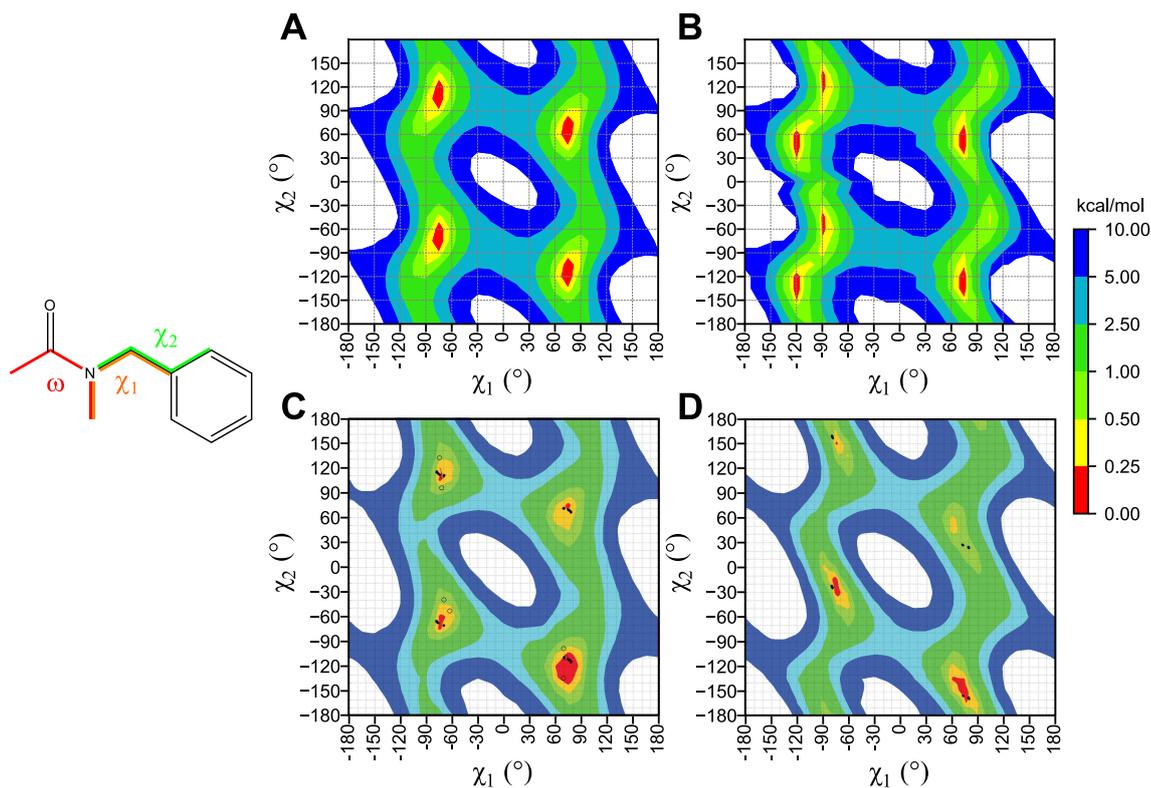


**Figure A.2:** Energy profiles of the *Npm* side chain dihedrals obtained at MP2/cc-pVDZ level (top) and normalized population count (at 298 K) obtained with MP2/cc-pVDZ (black), default DREIDING (red) and using the new sets of parameters PEPDROID (blue) (bottom). When applied, the restraints are displayed in the energy profiles. The primary structures of the models are displayed below the plots with the definition of the dihedrals.

## A.2 Ramachandran-like plots of peptoid side chains



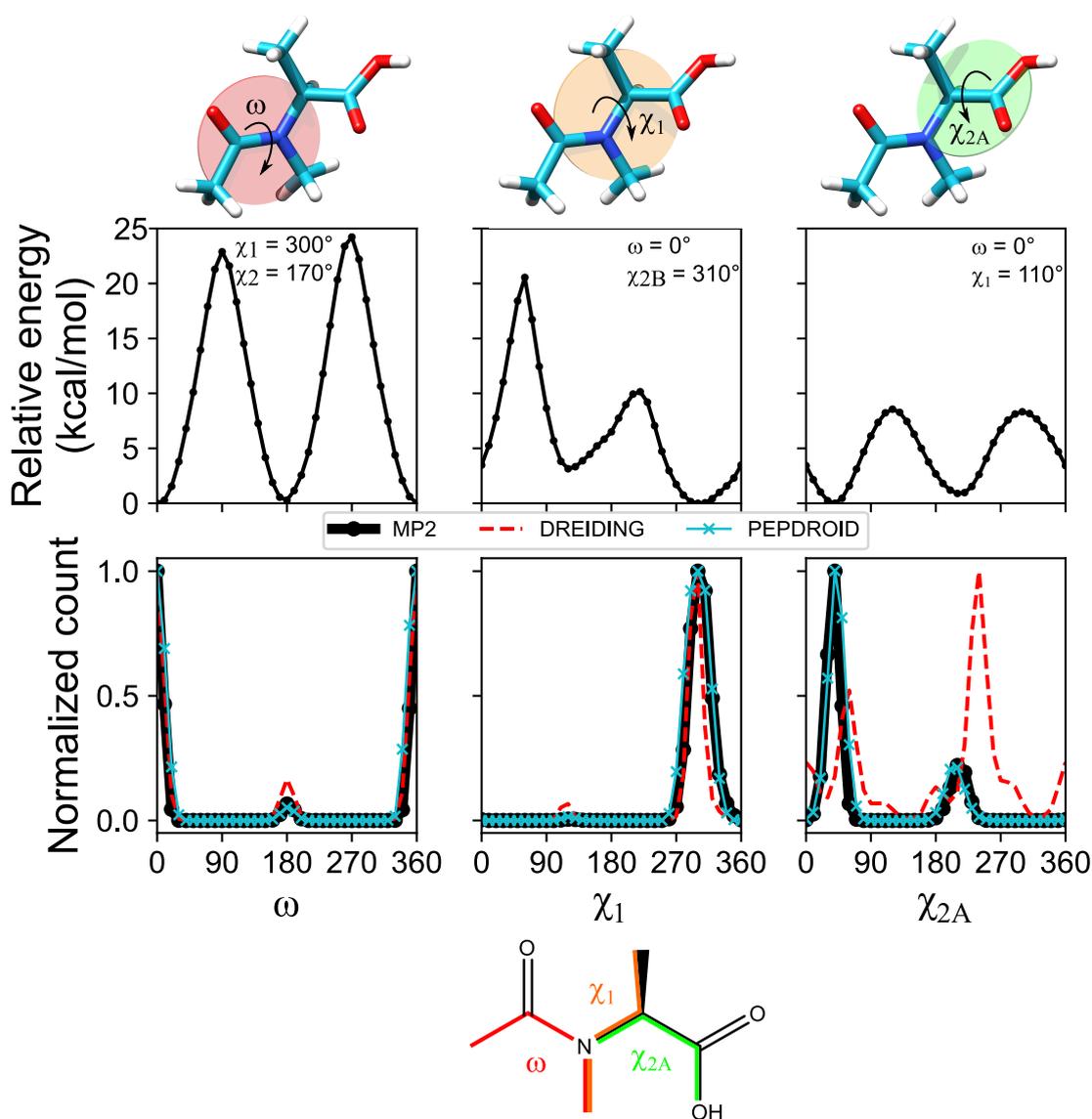
**Figure A.3:** Ramachandran-like plots of side chain dihedrals ( $\chi_1$ ,  $\chi_2$ ) for Nnpr in the (A) *cis*- and (B) *trans*-amide conformation obtained with PEPDROID. Ramachandran-like plots in (C) *cis*- and (D) *trans*-amide conformation obtained at the B3LYP/6-311+G(d,p) DFT level by Renfrew *et al.*<sup>76</sup> Adapted with permission from 76. Copyright 2014 American Chemical Society. The energy range spans from 0 to 10 kcal mol<sup>-1</sup>. The lowest energy conformers (red) are set to 0 kcal mol<sup>-1</sup> for each plot, while the highest energy conformers (up to 10 kcal mol<sup>-1</sup>) are displayed in blue. Conformers with relative energies higher than 10 kcal mol<sup>-1</sup> correspond to the white color.



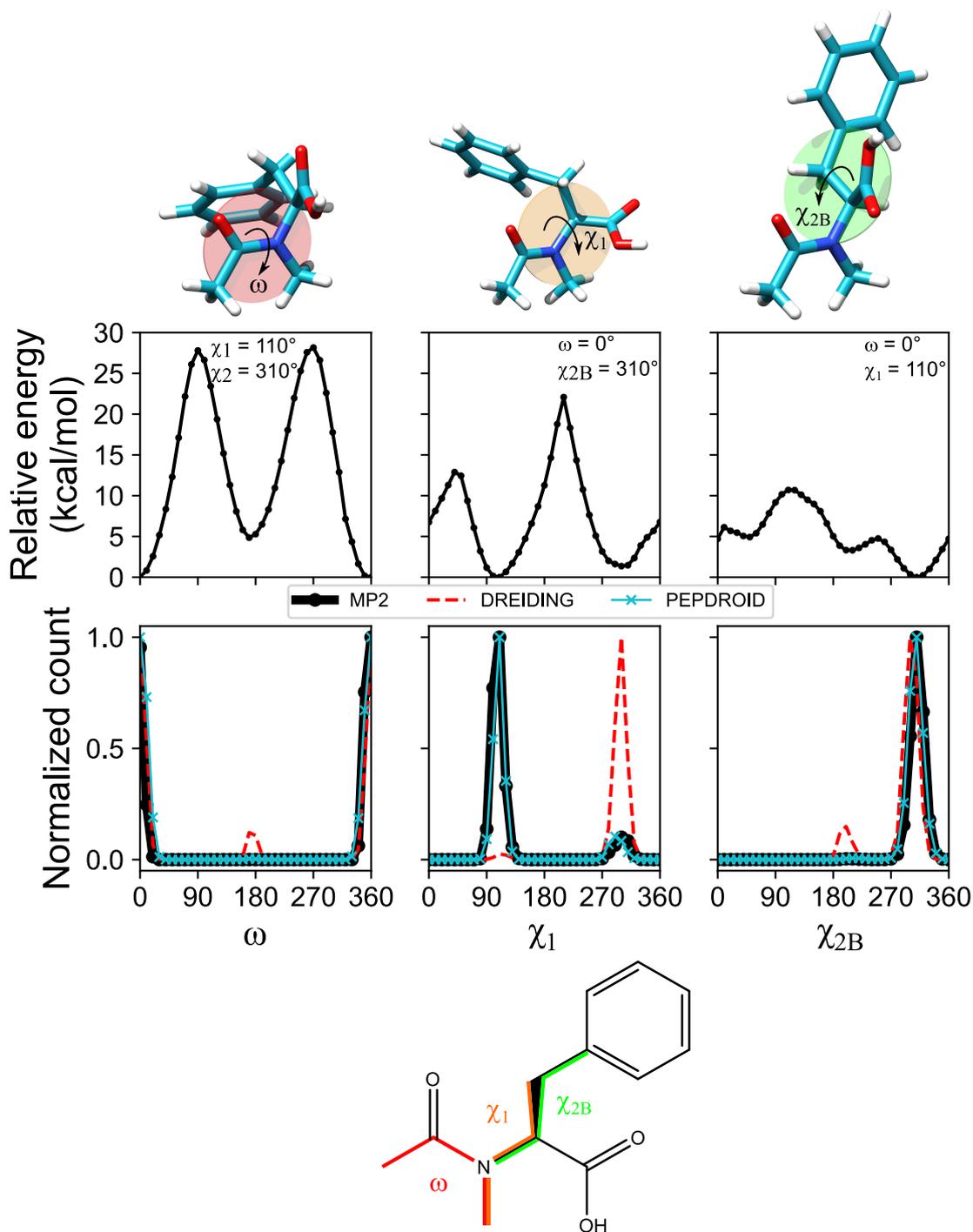
**Figure A.4:** Ramachandran-like plots of side chain dihedrals ( $\chi_1$ ,  $\chi_2$ ) for Npm in the (A) *cis*- and (B) *trans*-amide conformation obtained with PEPDROID. Ramachandran-like plots in (C) *cis*- and (D) *trans*-amide conformation obtained at the B3LYP/6-311+G(d,p) DFT level by Renfrew *et al.*<sup>76</sup> Adapted with permission from 76. Copyright 2014 American Chemical Society. The energy range spans from 0 to 10 kcal mol<sup>-1</sup>. The lowest energy conformers (red) are set to 0 kcal mol<sup>-1</sup> for each plot, while the highest energy conformers (up to 10 kcal mol<sup>-1</sup> are displayed in blue. Conformers with relative energies higher than 10 kcal mol<sup>-1</sup> correspond to the white color.

## Additional Material of Chapter 5

### B.1 Dihedral profiles of $N_{scp}$ and $N_{scm}$ side chains



**Figure B.1:** Energy profiles of the  $N_{scm}$  side chain dihedrals obtained at MP2/cc-pVDZ level (top) and normalized population count (at 298 K) obtained at MP2/cc-pVDZ (black), default DREIDING (red) and using the new sets of parameters PEPDROID (blue) (bottom). When applied, the restraints are displayed in the energy profiles. The primary structures of the models are displayed after the plots with the definition of the dihedrals.



## Additional Material of Chapter 6

### C.1 Conformations description

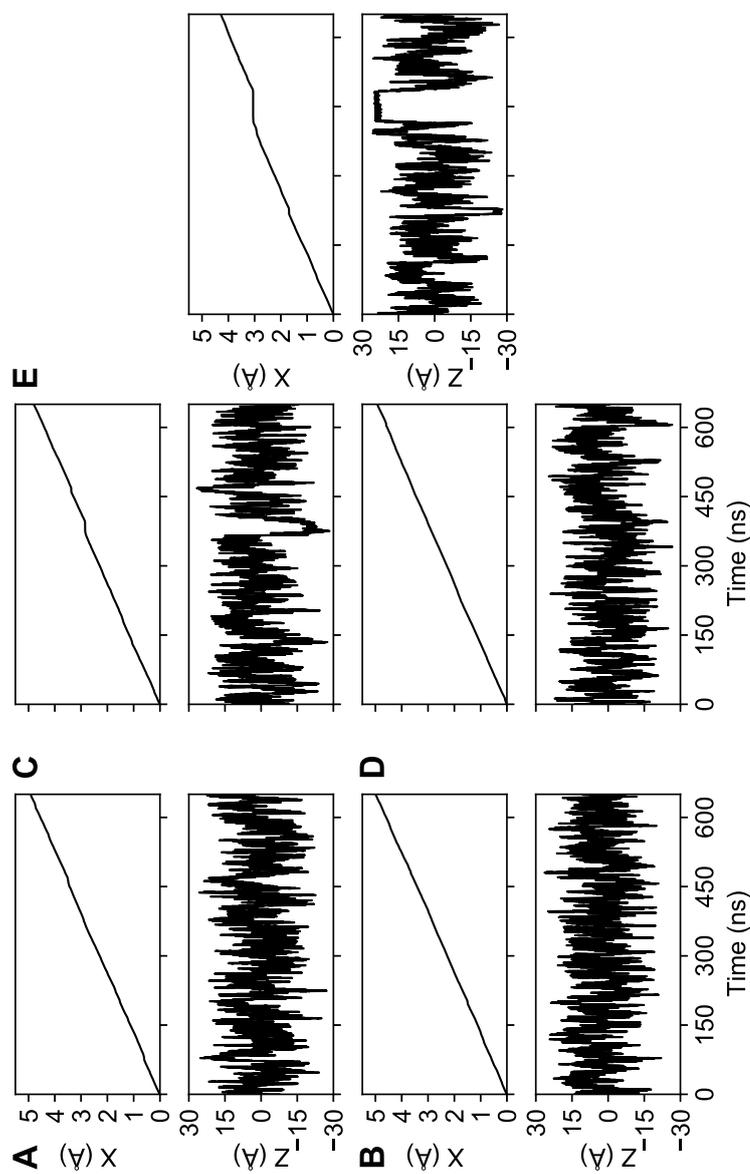
**Table C.1:** Every possible conformation with their simplified sequence and verbose description of their structure.

Simplified sequence	Description
“RRRRR”	Perfect right-handed helix having dihedrals close to those described by Armand <i>et al.</i> <sup>41</sup>
“RRRRr”	Almost perfect right-handed helix, except for the C terminus residue which has its $\psi$ close to $0^\circ$
“XRRRr” with $X = L, M, C, T$	Partial right-handed helix having the first residue at N terminus which is not in helical conformation
“XXRRr” with $X = L, M, C, T$	Partial right-handed helix having the two residues at N terminus which are not in helical conformation
“’RRRRX’ with $X = L, l, M, C, T$	Partial right-handed helix having the first residue at C terminus which is not in helical conformation
“RRRXX” with $X = L, l, r, M, C, T$	Partial right-handed helix having the two residues at C terminus which are not in helical conformation
“XRRRX” with $X = L, l, M, C, T$	Partial right-handed helix having both N and C termini not in helical conformation (first and last residues)
“XXRRX” or “XRRXX” with $X = L, l, r, M, C, T$	Proto-right-handed helix having 3 units in the center of the peptoid in helical conformation
“MRRRr” or “RRRRM”	Partial right-handed helix having its $\omega$ dihedral at N or C terminus out of <i>cis</i> or <i>trans</i>

( To be continued)

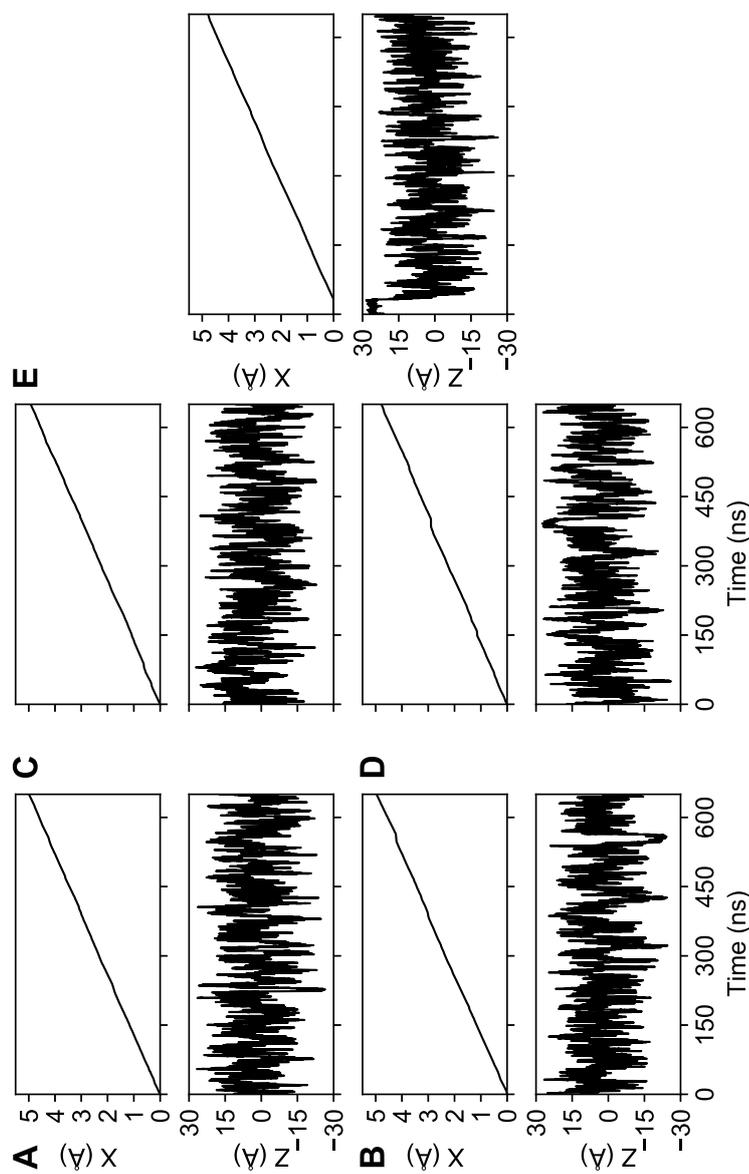
Simplified sequence	Description
“ <i>XMRRR</i> ”, “ <i>MXRRr</i> ”, “ <i>RRRXM</i> ” or “ <i>RRRMX</i> ” with $X = L, l, R, r, M, C, T$	Partial right-handed helix having its $\omega$ dihedral for two residues at N or C terminus out of <i>cis</i> or <i>trans</i>
“ <i>RLLLR</i> ”	Alternate right- and left-handed residues
“ <i>LRLRL</i> ”	Alternate left- and right-handed residues
“ <i>LLLLL</i> ”	Perfect left-handed helix having dihedrals opposite to the right-handed helix
“ <i>LLLLl</i> ”	Almost perfect left-handed helix, except for the C terminus residue which has its $\psi$ close to $0^\circ$
“ <i>XLLLL</i> ” with $X = R, M, C, T$	Partial left-handed helix having the first residue at N terminus which is not in helical conformation
“ <i>XXLLL</i> ” with $X = R, M, C, T$	Partial left-handed helix having the two residues at N terminus which are not in helical conformation
“ <i>LLLLX</i> ” with $X = R, r, M, C, T$	Partial left-handed helix having the first residue at C terminus which is not in helical conformation
“ <i>LLLXX</i> ” with $X = R, r, M, C, T$	Partial left-handed helix having the two residues at C terminus which are not in helical conformation
“ <i>XLLLX</i> ” with $X = R, r, M, C, T$	Partial left-handed helix having both N and C termini not in helical conformation
“ <i>XXLLX</i> ” or “ <i>XLLXX</i> ” with $X = L, R, r, M, C, T$	Proto-left-handed helix having 3 residues in the center of the peptoid in helical conformation
“ <i>MLLLL</i> ” or “ <i>LLLLM</i> ”	Partial left-handed helix having its $\omega$ dihedral at N or C terminus out of <i>cis</i> or <i>trans</i>
“ <i>XMLLL</i> ”, “ <i>MXLLl</i> ”, “ <i>LLLXM</i> ” or “ <i>LLLMX</i> ” with $X = R, r, L, l, M, C, T$	Partial left-handed helix having its $\omega$ dihedral for two residues at N or C terminus out of <i>cis</i> or <i>trans</i>
“ <i>*M*</i> ”	Every simplified code containing a “ <i>M</i> ”

## C.2 *Nspe* peptoids - (*S*)-BINOL: displacements analysis



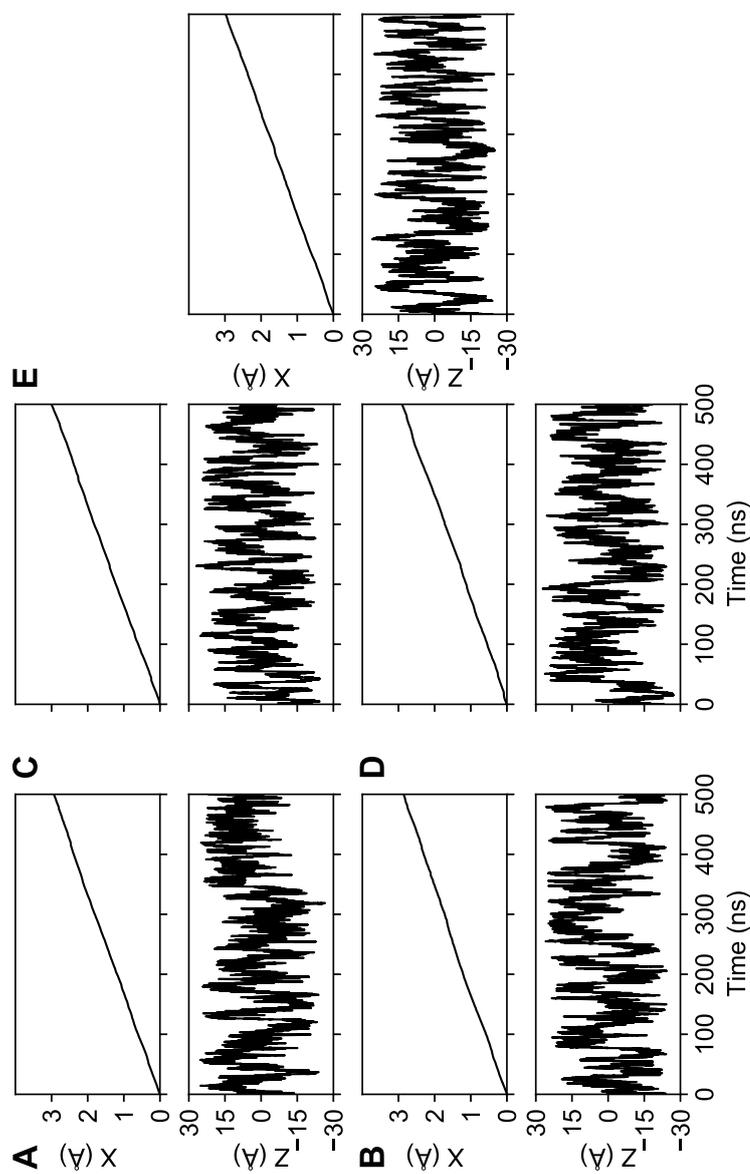
**Figure C.1:** Displacement of each (*S*)-BINOL molecule along the X and Z axis over 650 ns. A direct correlation is observed between plateaus along the X displacement and the proximity to the top or bottom surface (around  $\pm 30$  Å) in the Z direction.

### C.3 *N*spe peptoids - (*R*)-BINOL: displacements analysis



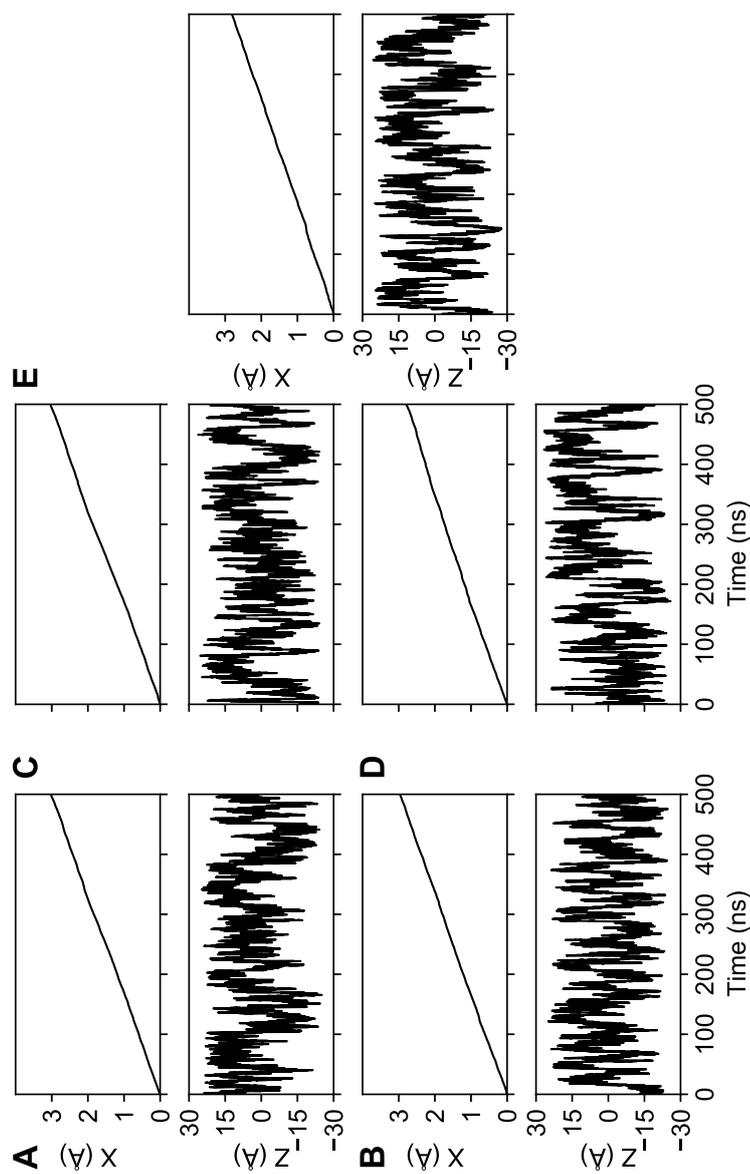
**Figure C.2:** Displacement of each (*R*)-BINOL molecule along the X and Z axis over 650 ns. A direct correlation is observed between plateaus along the X displacement and the proximity to the top or bottom surface (around  $\pm 30$  Å) in the Z direction.

## C.4 *Nsar* peptoids - (*S*)-BINOL: displacements analysis



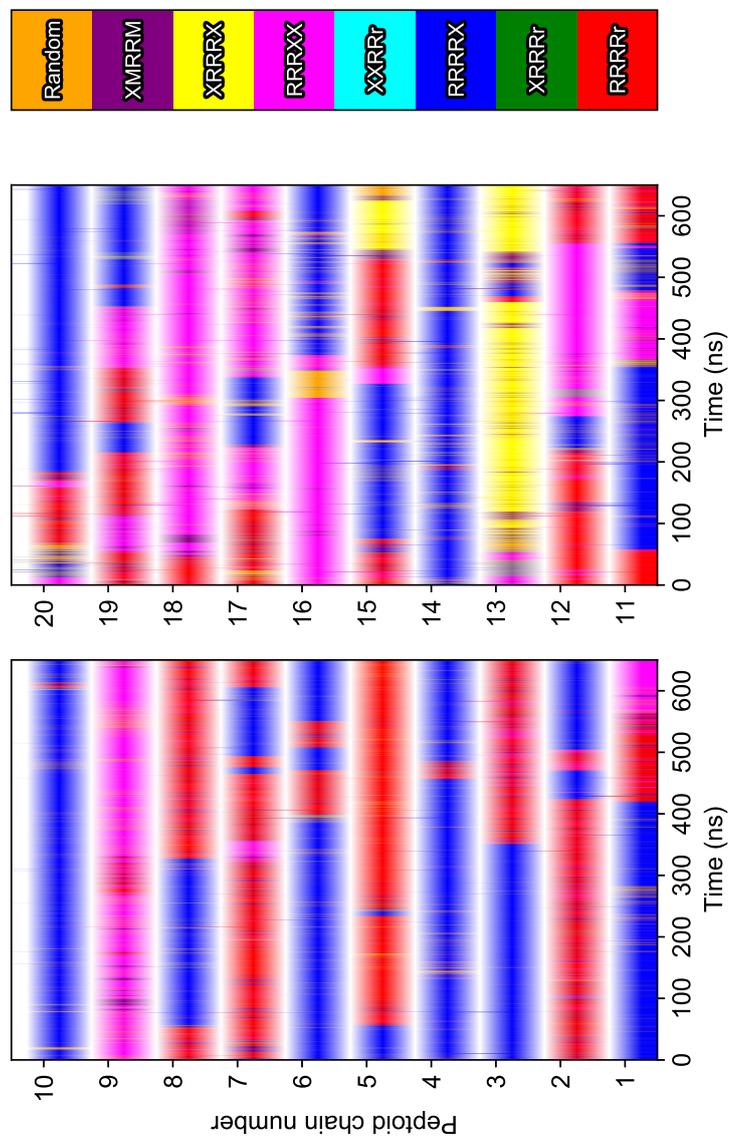
**Figure C.3:** Displacement of each (*S*)-BINOL molecule along the X and Z axis over 650 ns. A direct correlation is observed between plateaus along the X displacement and the proximity to the top or bottom surface (around  $\pm 30$  Å) in the Z direction.

## C.5 *N*sar peptoids - (*R*)-BINOL: displacements analysis



**Figure C.4:** Displacement of each (*R*)-BINOL molecule along the X and Z axis over 650 ns. A direct correlation is observed between plateaus along the X displacement and the proximity to the top or bottom surface (around  $\pm 30$  Å) in the Z direction.

## C.6 *Nspe* peptoids - (*S*)-BINOL: contact maps

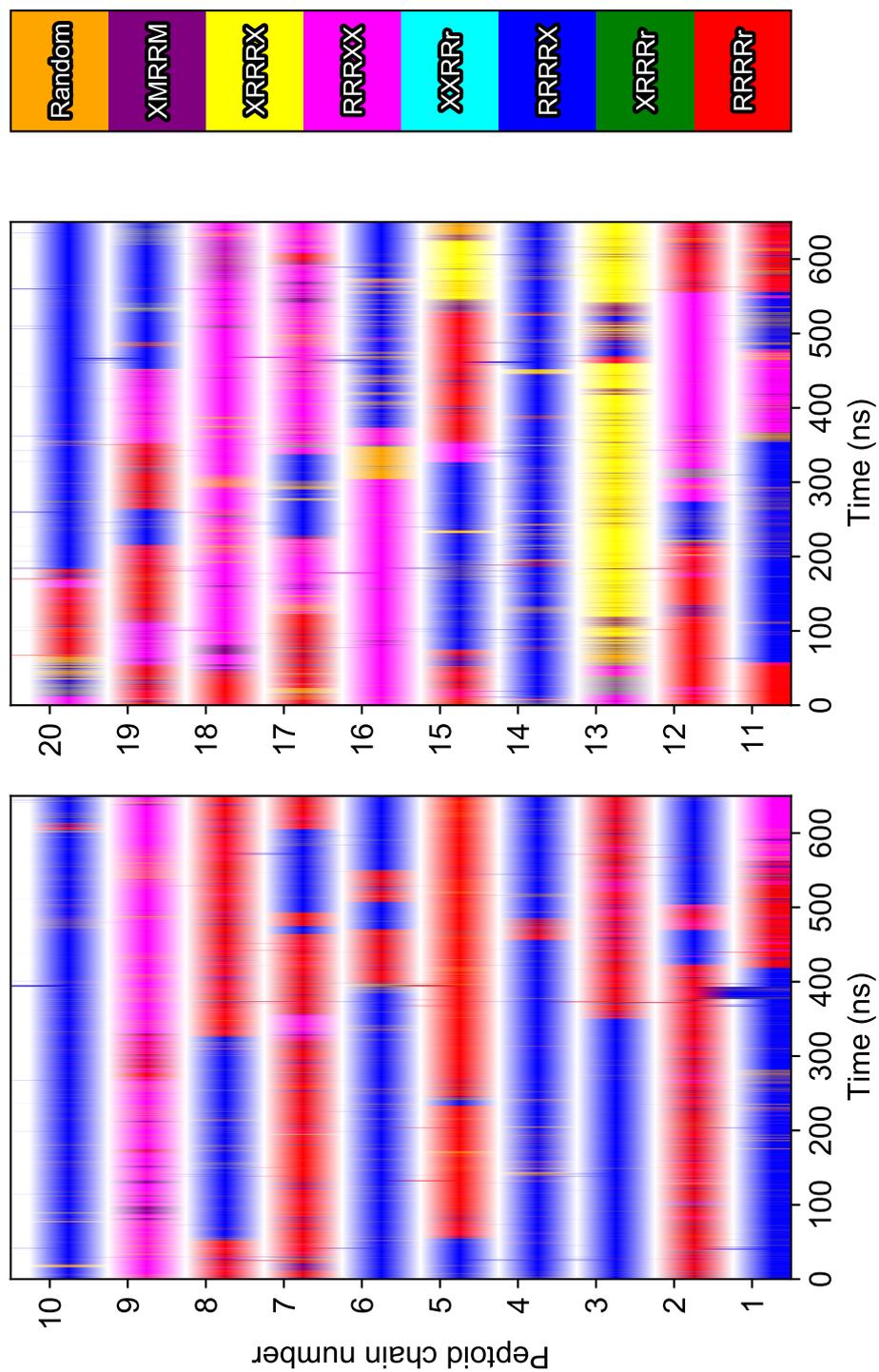


**Figure C.5:** Conformational contact maps of (*S*)-BINOL2 with each of the 20 peptoid chains.

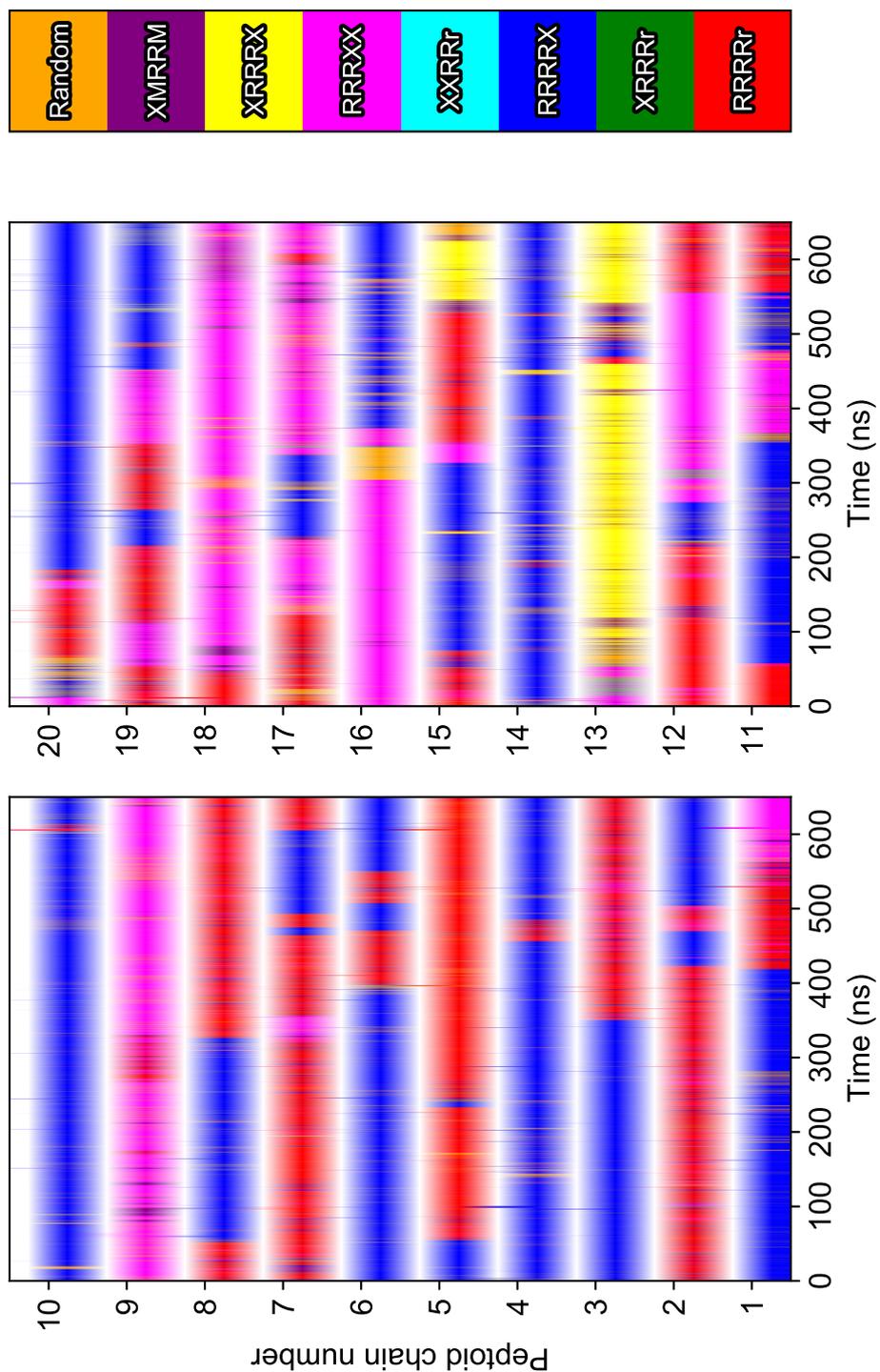
<sup>a</sup> The color code is the same as previously described in Chapter 6.

---

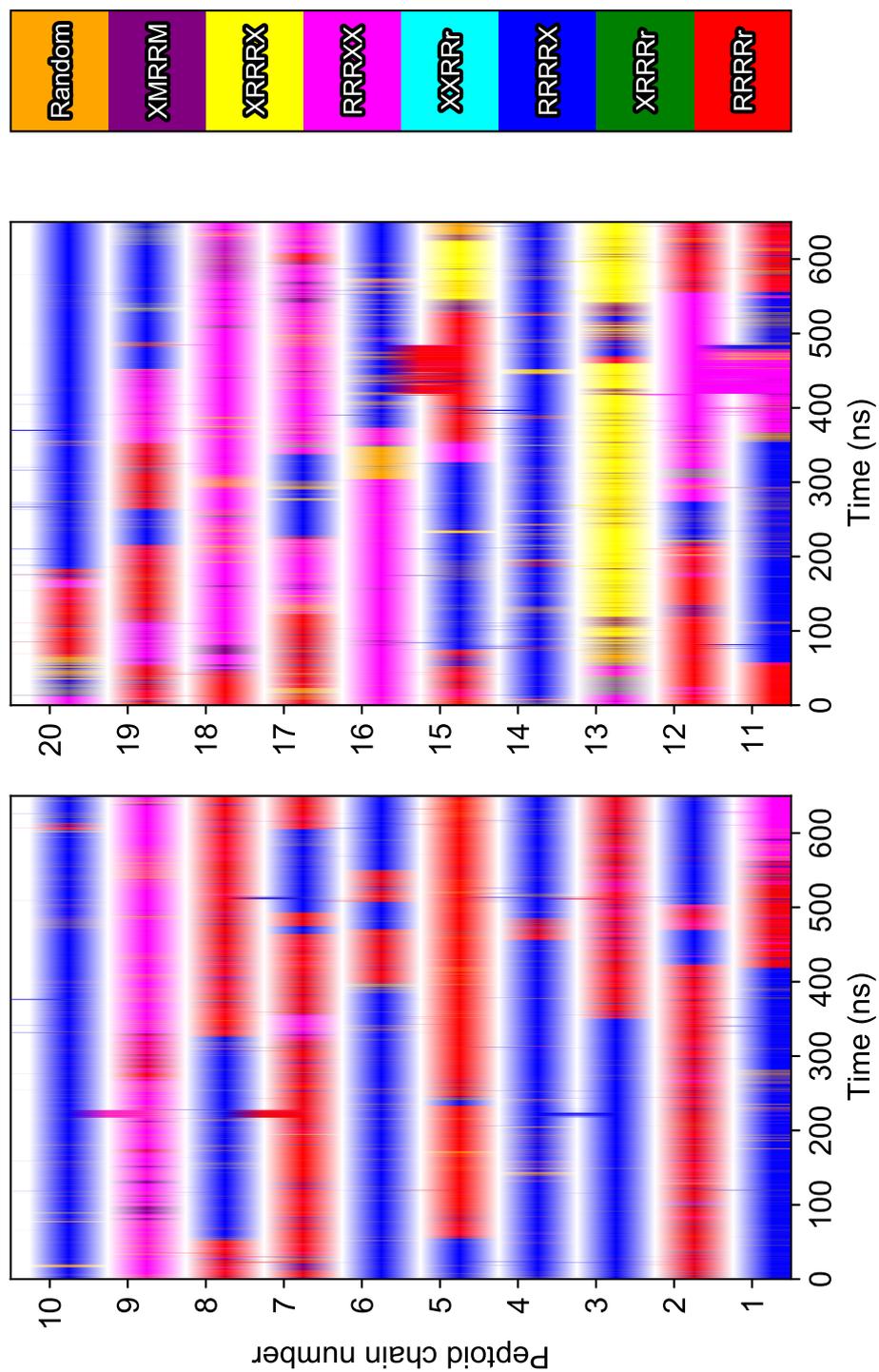
<sup>a</sup>In the PDF file, you may zoom to better visualize the contact events in the mostly white ribbons.



**Figure C.6:** Conformational contact maps of (S)-BINOL3 with each of the 20 peptoid chains. The color code is the same as previously described in Chapter 6.

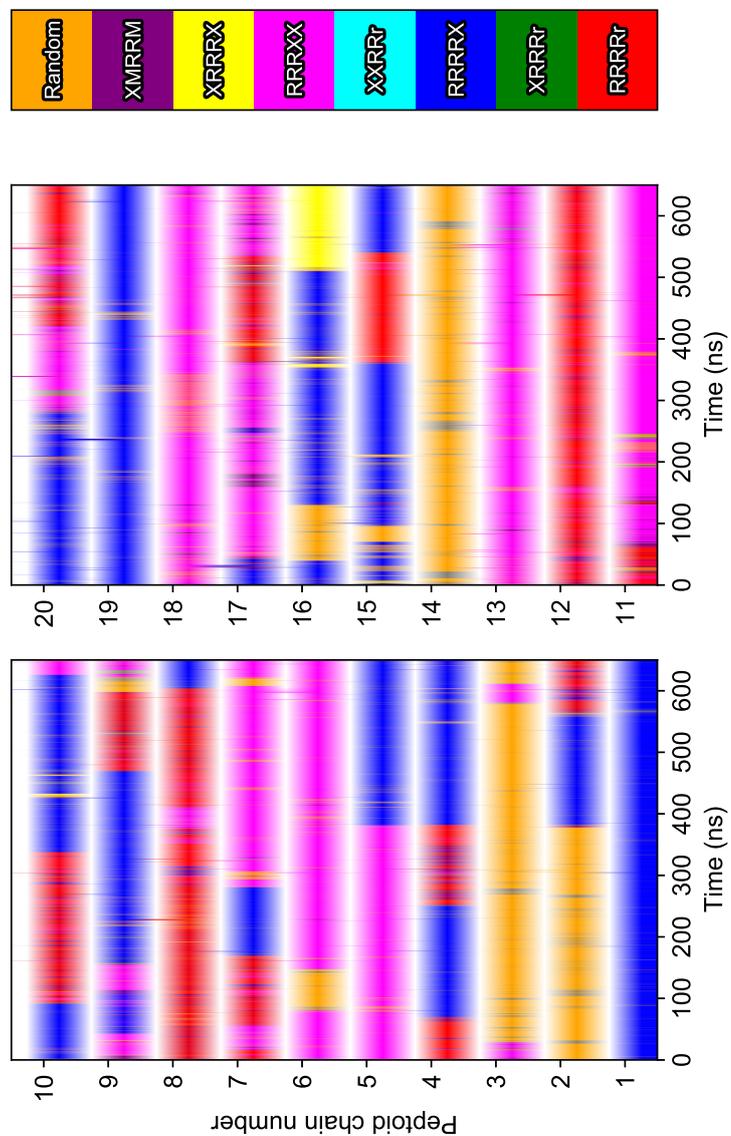


**Figure C.7:** Conformational contact maps of (*S*)-BINOL4 with each of the 20 peptoid chains. The color code is the same as previously described in Chapter 6.

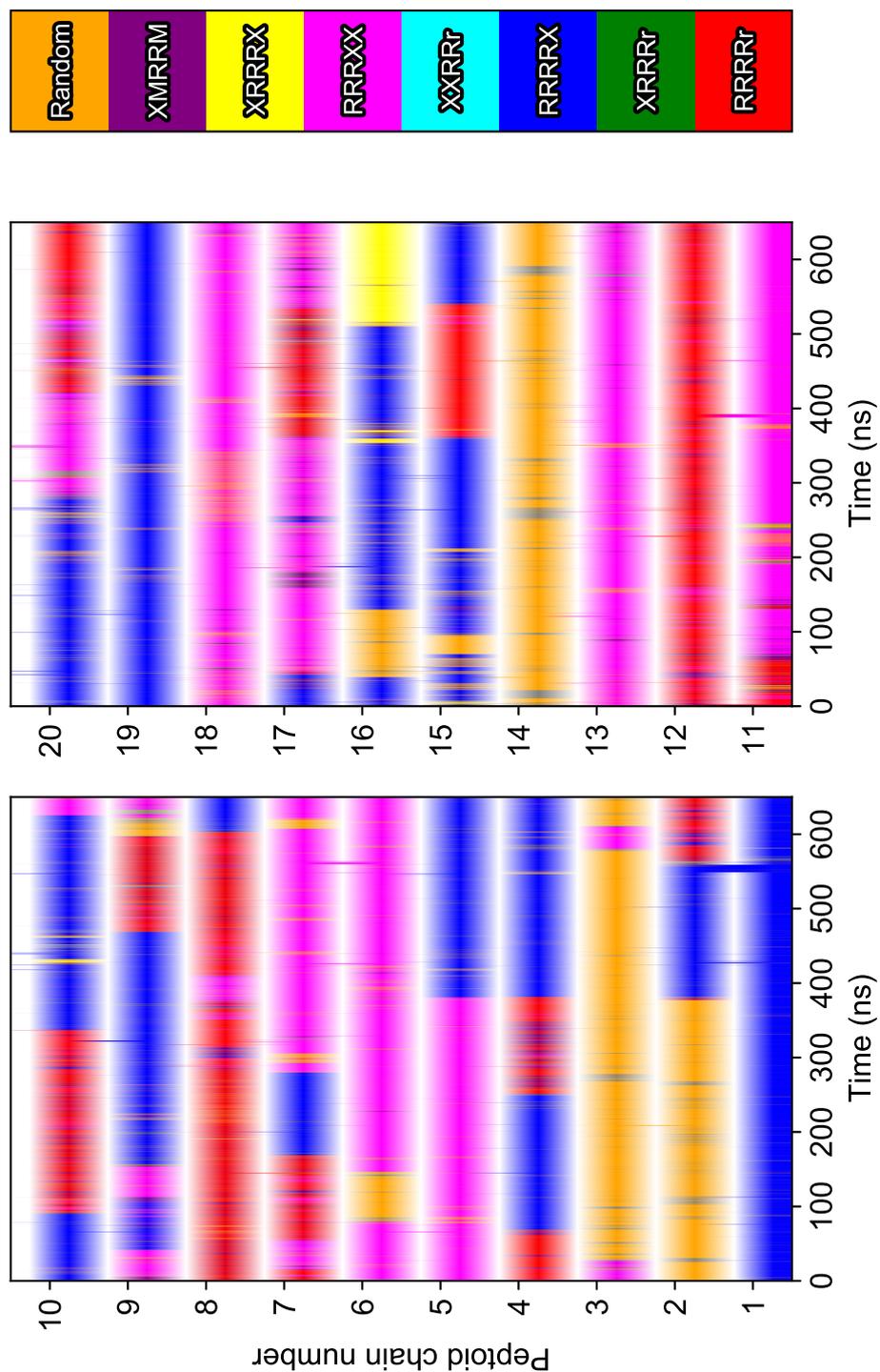


**Figure C.8:** Conformational contact maps of (S)-BINOL5 with each of the 20 peptoid chains. The color code is the same as previously described in Chapter 6.

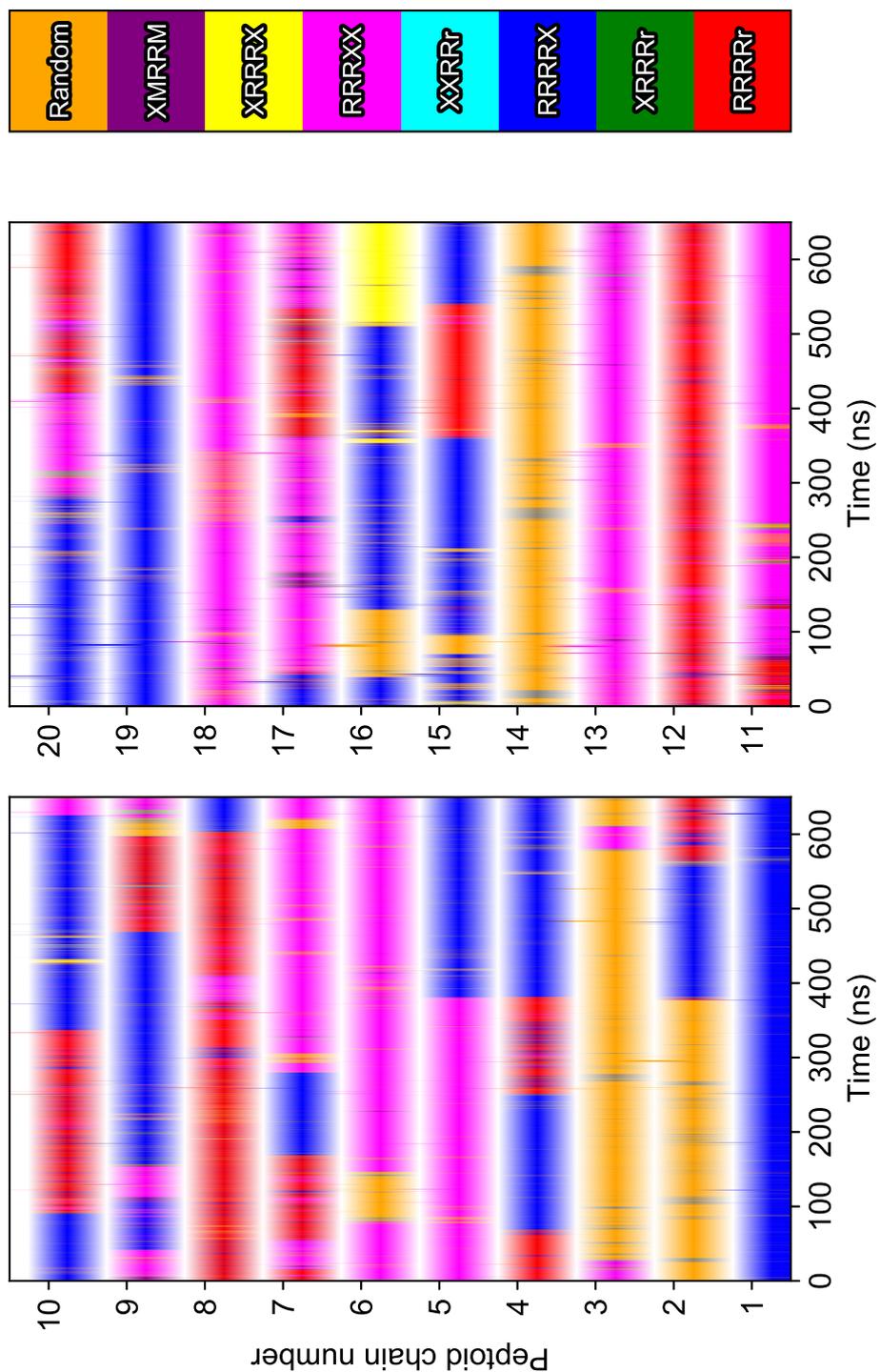
## C.7 *Nspe* peptoids - (*R*)-BINOL: contact maps



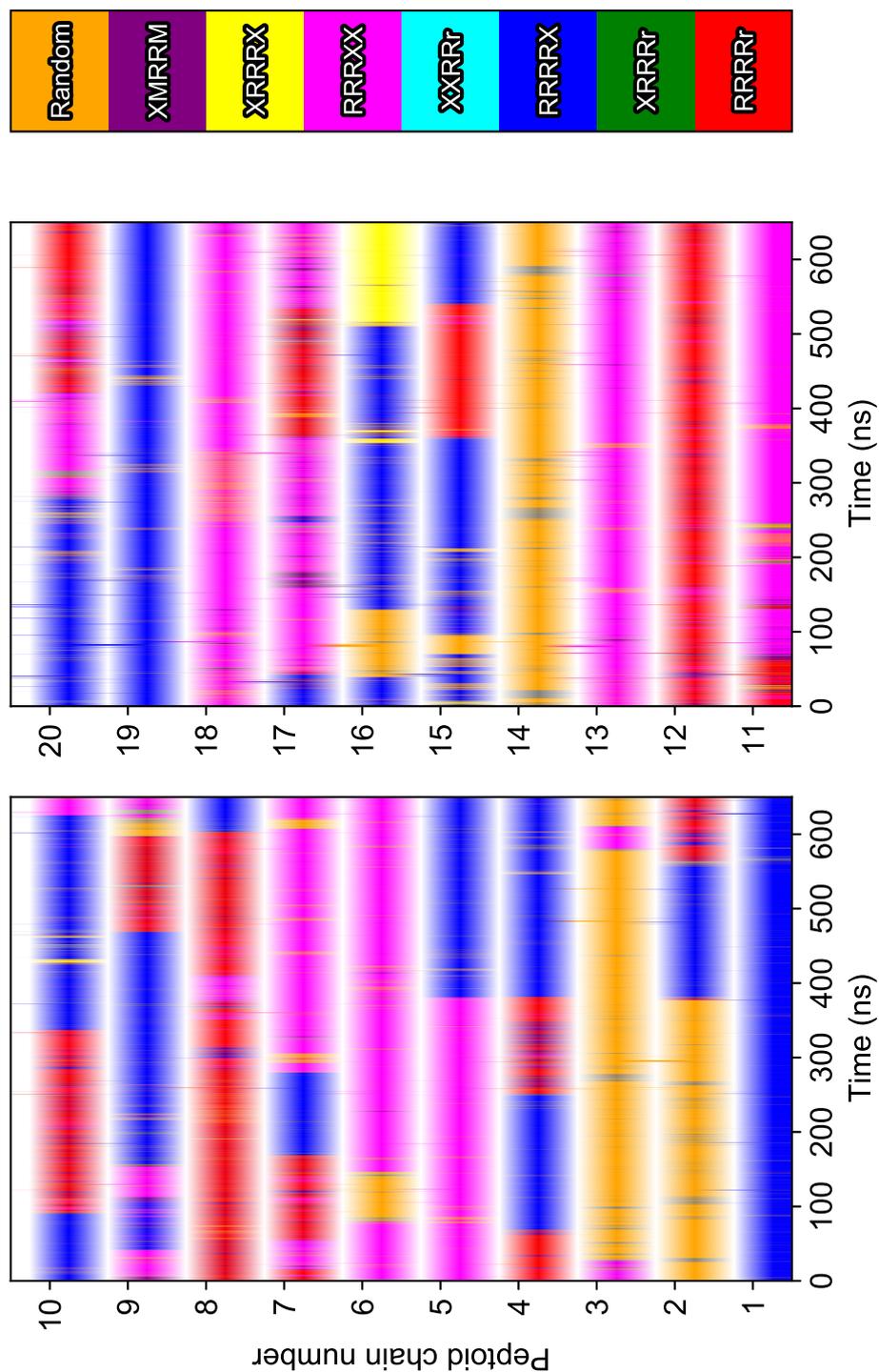
**Figure C.9:** Conformational contact maps of (*R*)-BINOL1 with each of the 20 peptoid chains. The color code is the same as previously described in Chapter 6.



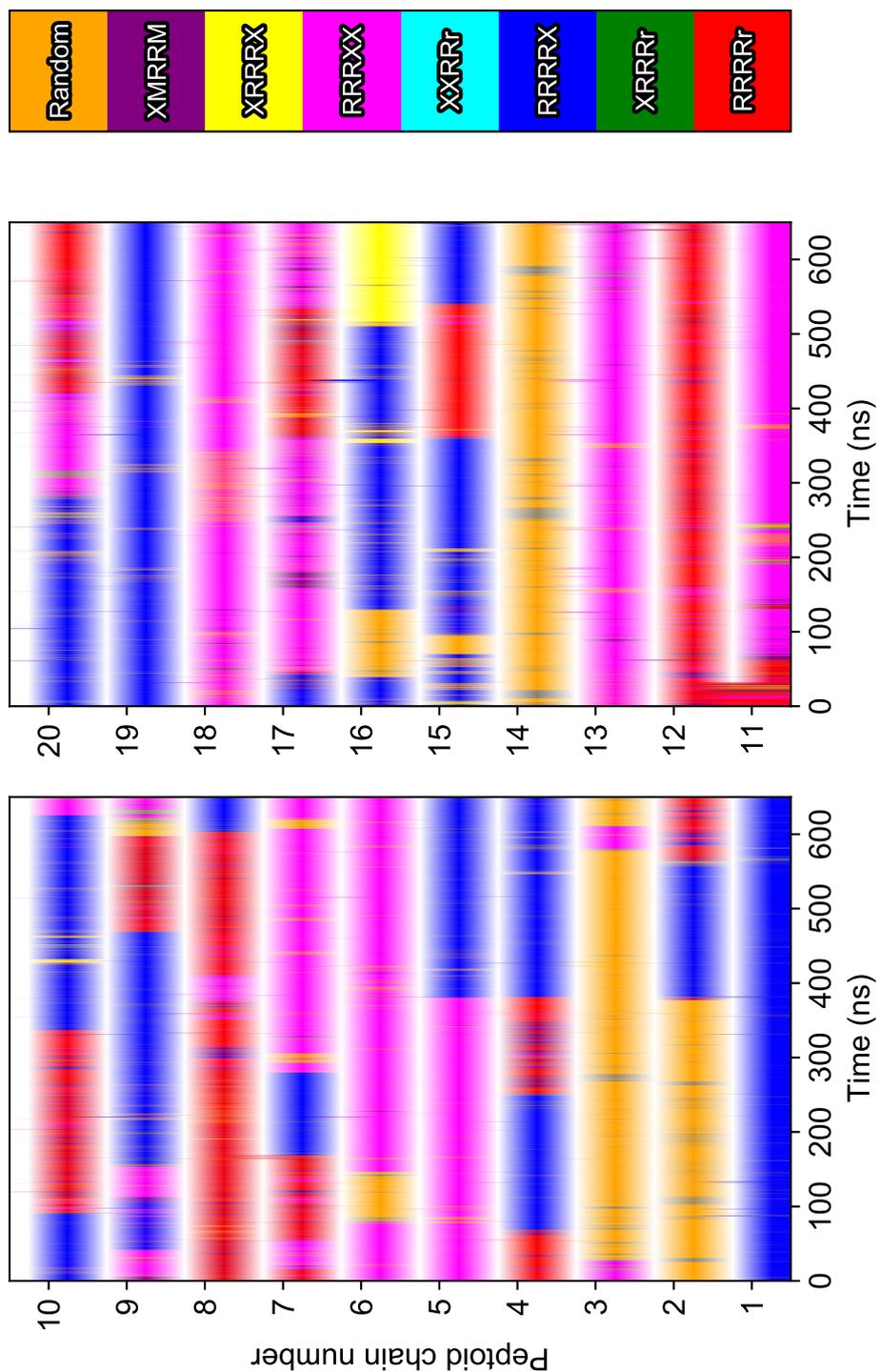
**Figure C.10:** Conformational contact maps of (R)-BINOL2 with each of the 20 peptoid chains. The color code is the same as previously described in Chapter 6.



**Figure C.11:** Conformational contact maps of (*R*)-BINOL3 with each of the 20 peptoid chains. The color code is the same as previously described in Chapter 6.

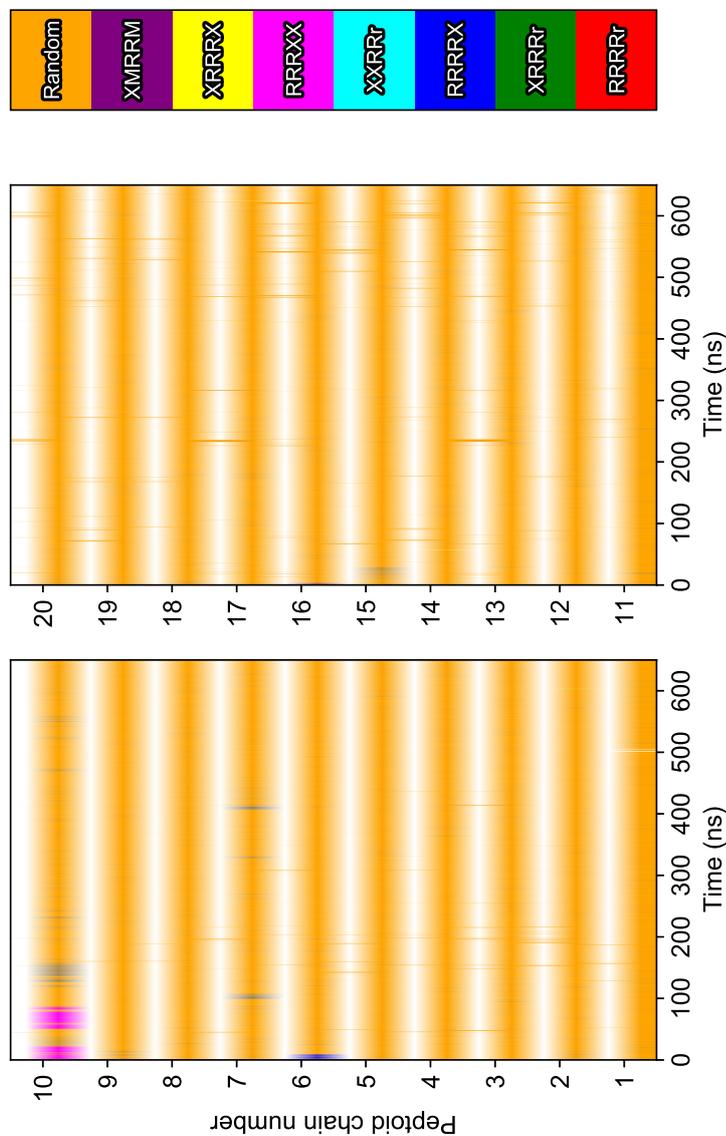


**Figure C.12:** Conformational contact maps of (R)-BINOL4 with each of the 20 peptoid chains. The color code is the same as previously described in Chapter 6.

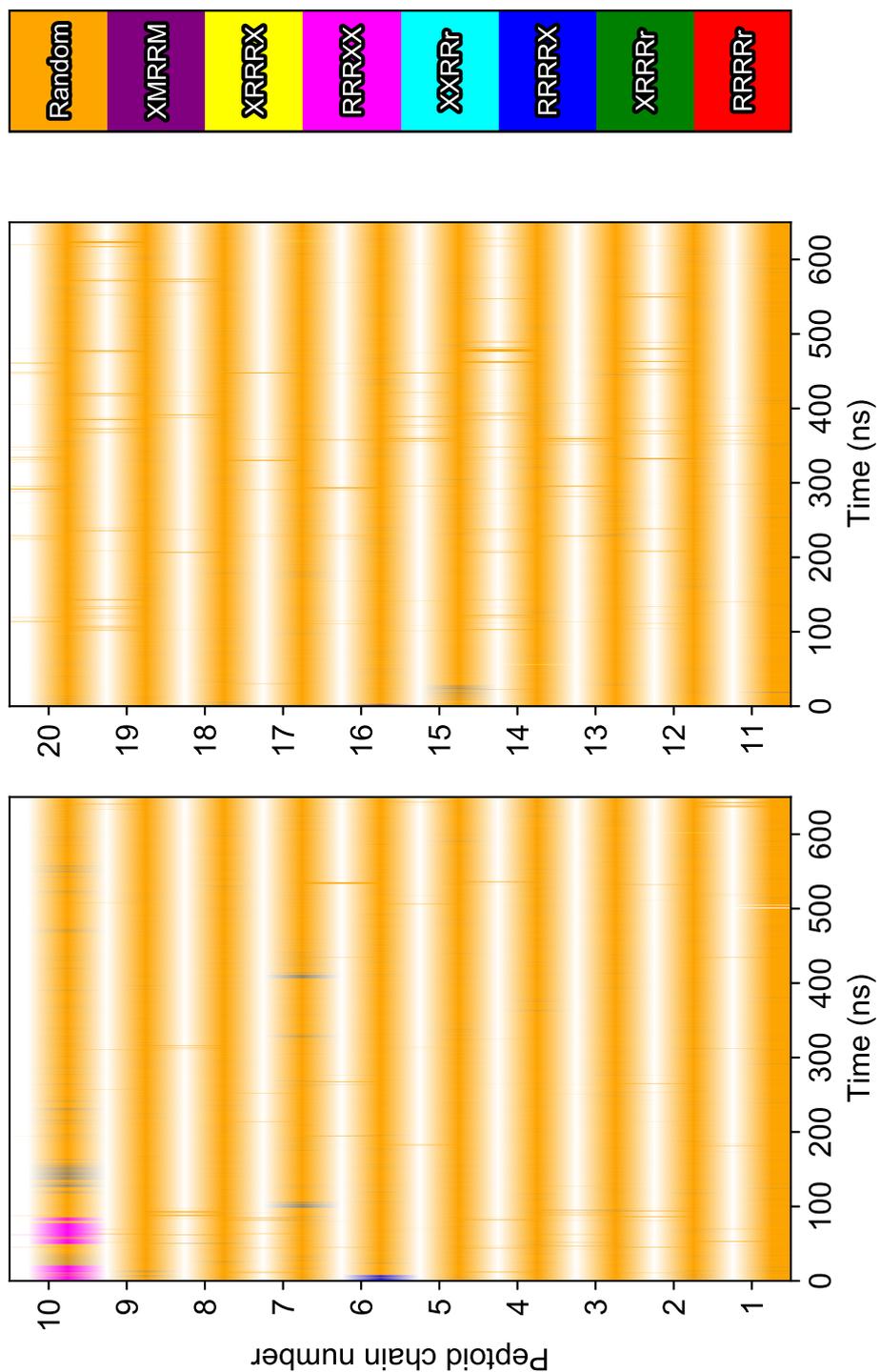


**Figure C.13:** Conformational contact maps of (*R*)-BINOL5 with each of the 20 peptoid chains. The color code is the same as previously described in Chapter 6.

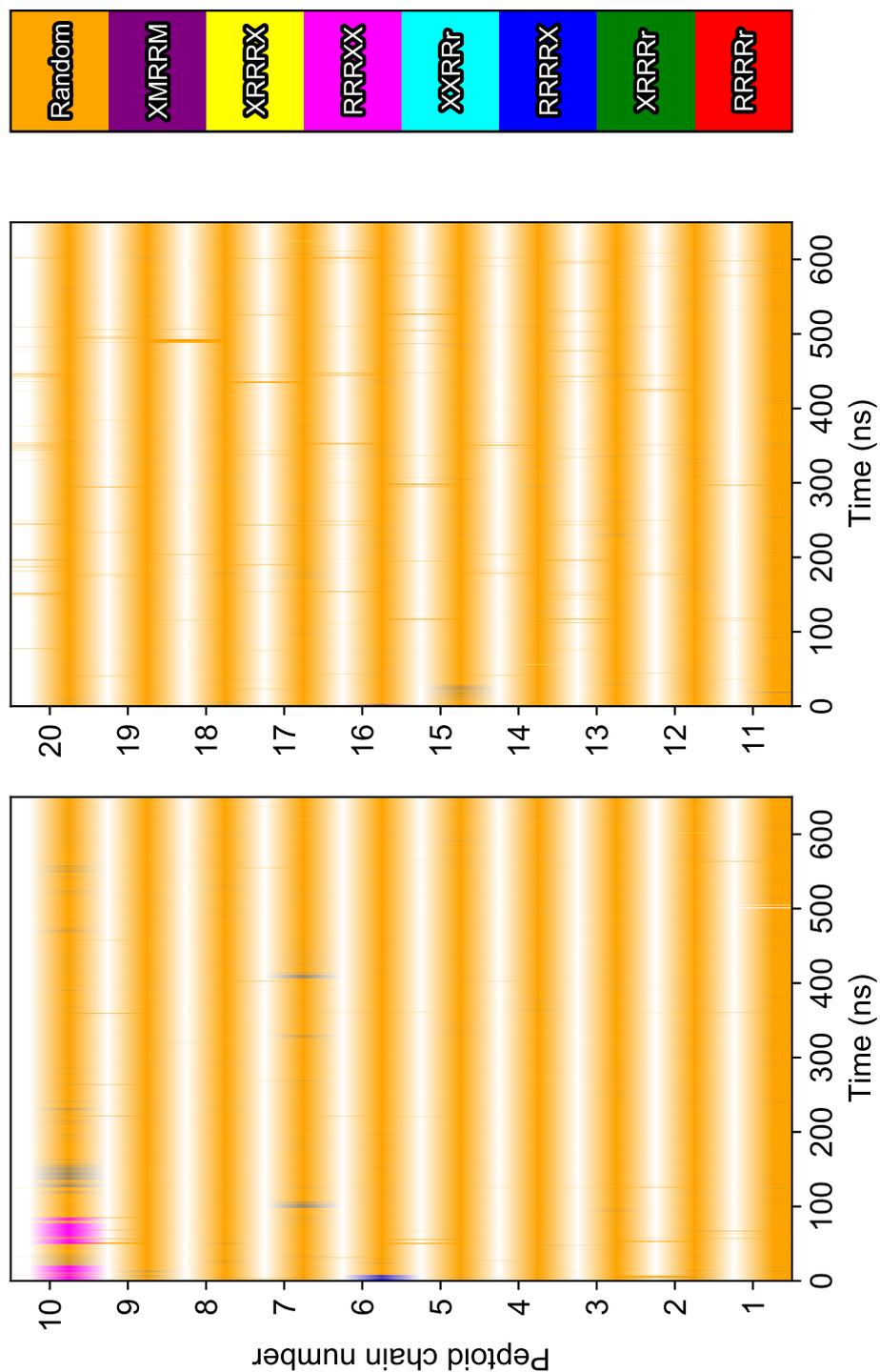
## C.8 Nsar peptoids - (S)-BINOL: contact maps



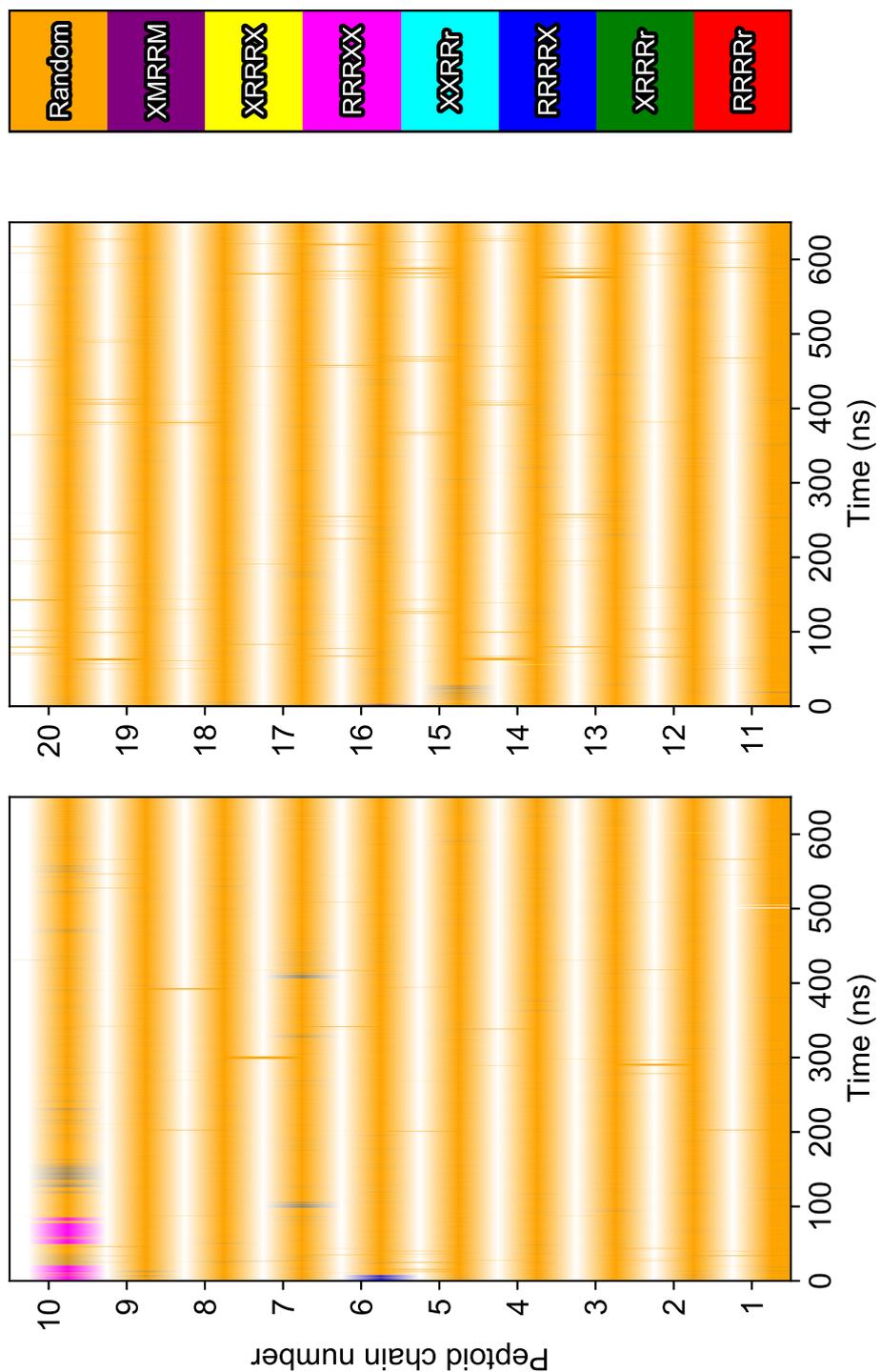
**Figure C.14:** Conformational contact maps of (S)-BINOL1 with each of the 20 peptoid chains. The color code is the same as previously described in Chapter 6.



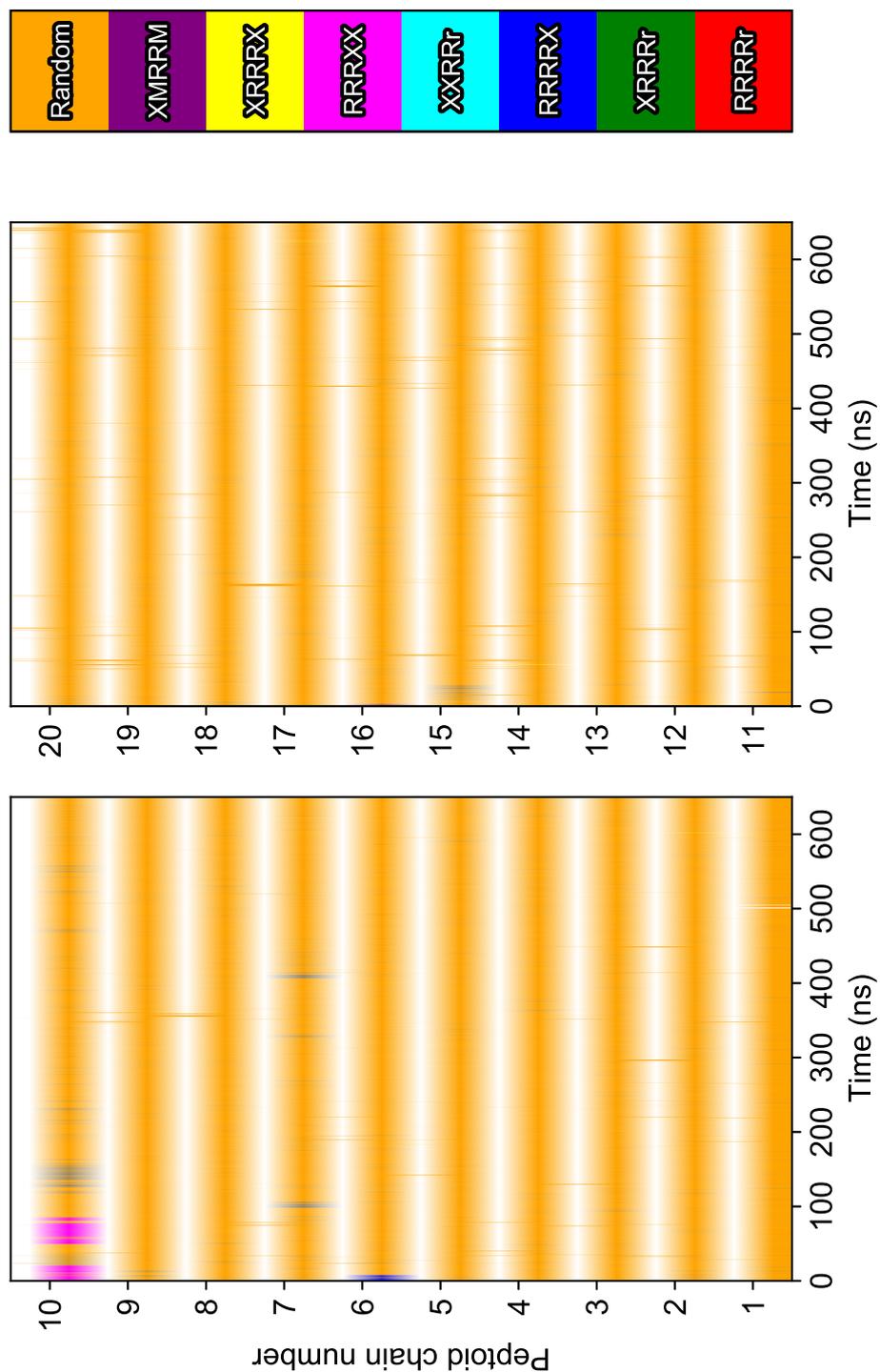
**Figure C.15:** Conformational contact maps of (*S*)-BINOL2 with each of the 20 peptoid chains. The color code is the same as previously described in Chapter 6.



**Figure C.16:** Conformational contact maps of (S)-BINOL3 with each of the 20 peptide chains. The color code is the same as previously described in Chapter 6.

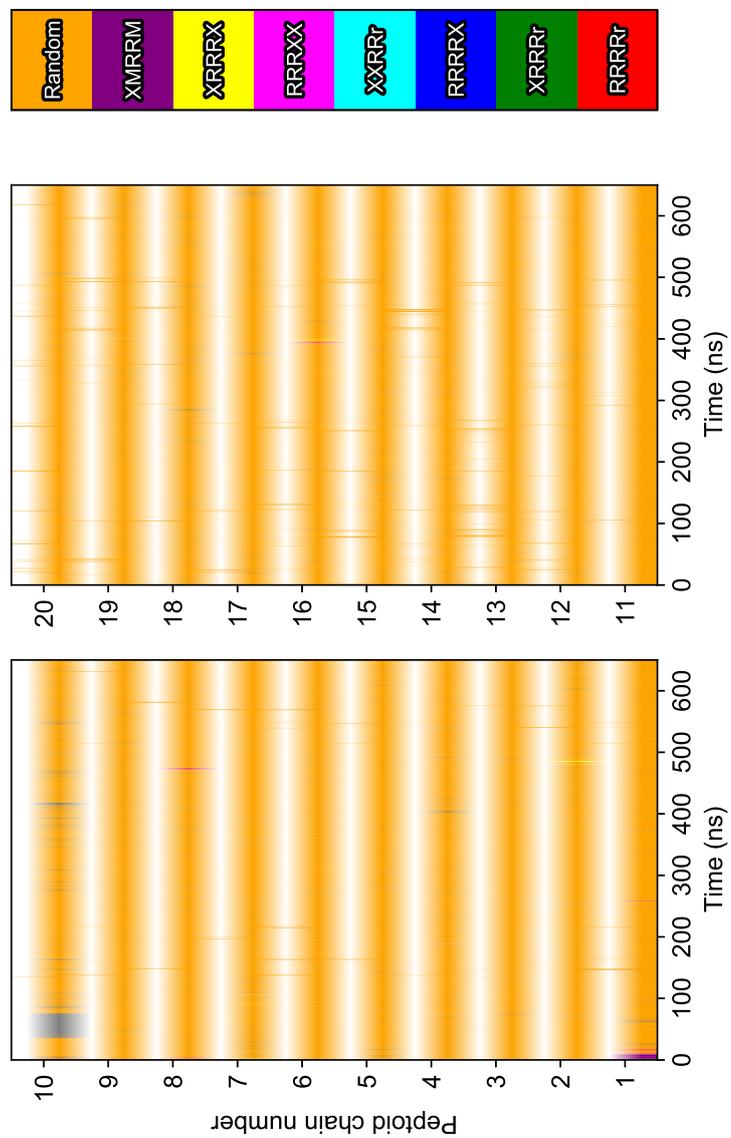


**Figure C.17:** Conformational contact maps of (S)-BINOL4 with each of the 20 peptoid chains. The color code is the same as previously described in Chapter 6.

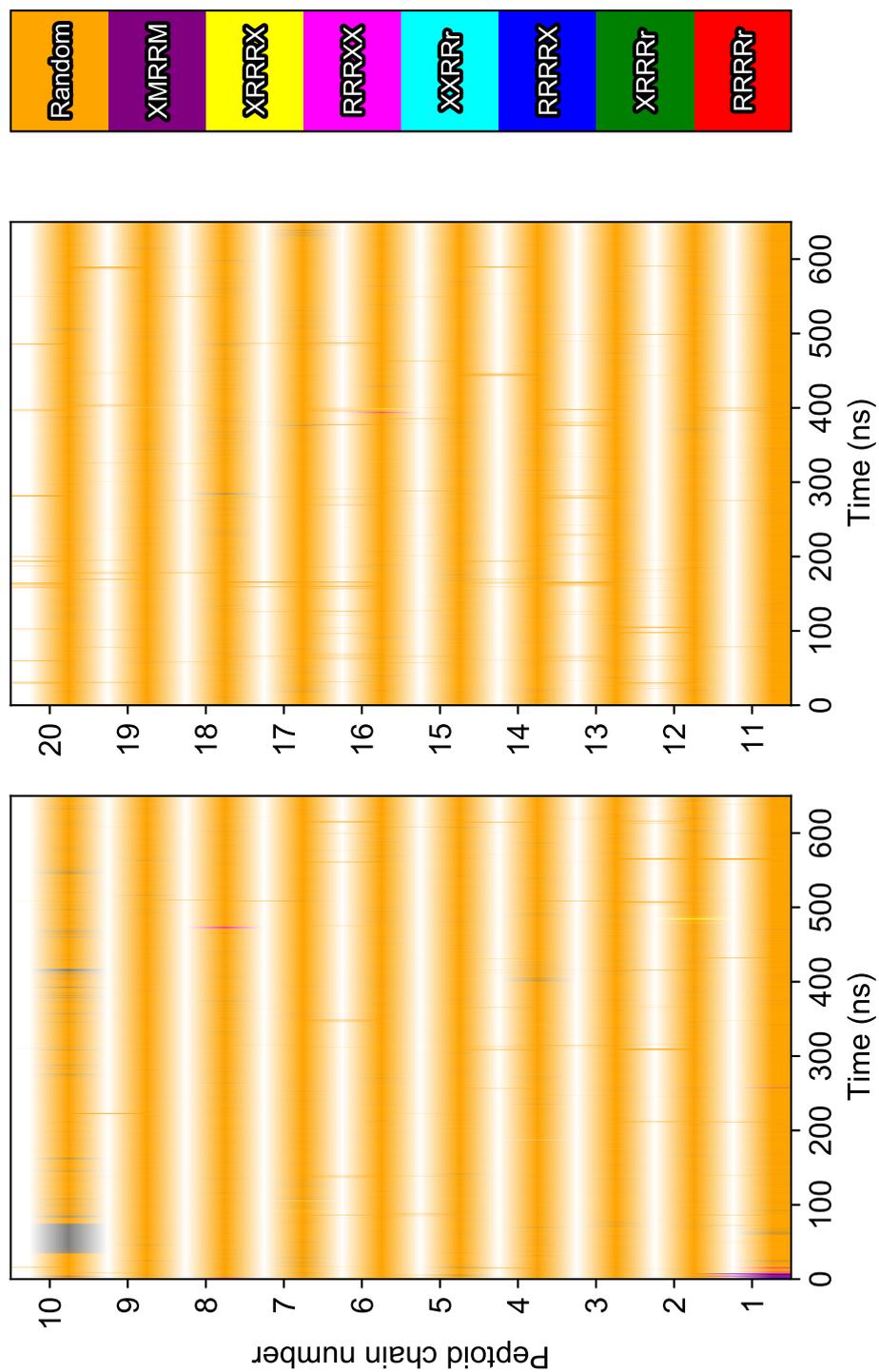


**Figure C.18:** Conformational contact maps of (S)-BINOL5 with each of the 20 peptide chains. The color code is the same as previously described in Chapter 6.

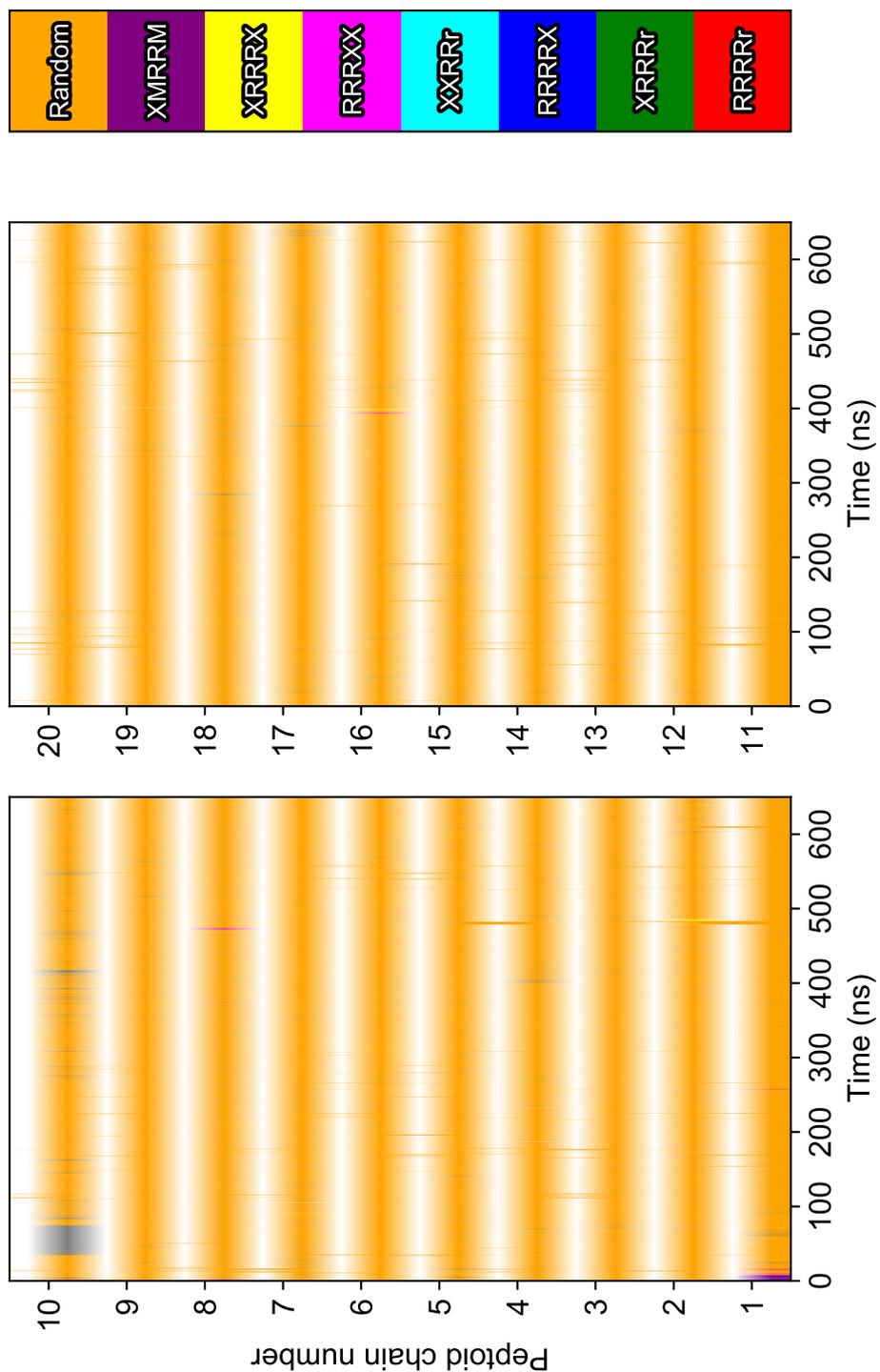
## C.9 *Nsar* peptoids - (*R*)-BINOL: contact maps



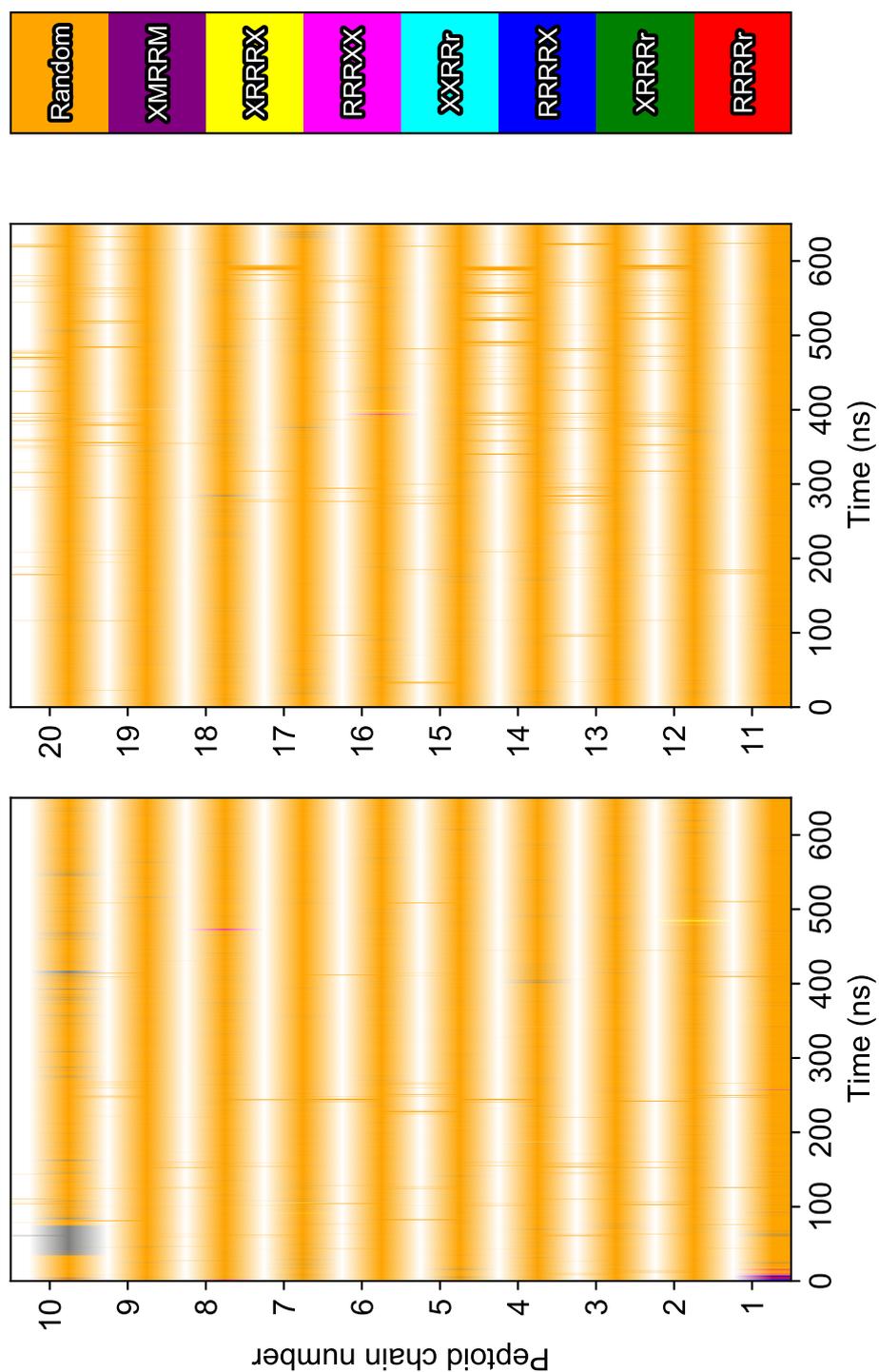
**Figure C.19:** Conformational contact maps of (*R*)-BINOL1 with each of the 20 peptoid chains. The color code is the same as previously described in Chapter 6.



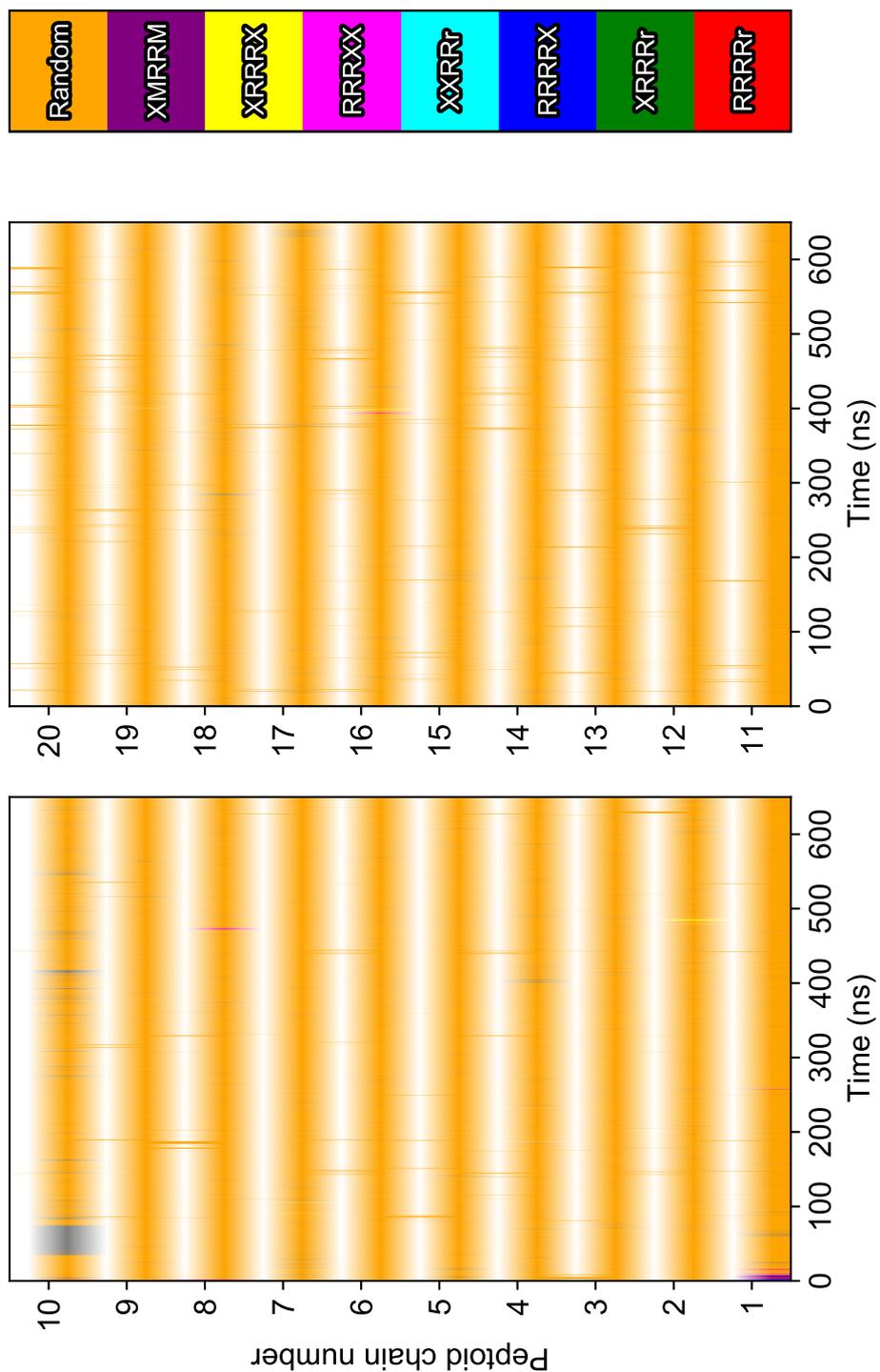
**Figure C.20:** Conformational contact maps of (R)-BINOL2 with each of the 20 peptide chains. The color code is the same as previously described in Chapter 6.



**Figure C.21:** Conformational contact maps of (R)-BINOL3 with each of the 20 peptoid chains. The color code is the same as previously described in Chapter 6.



**Figure C.22:** Conformational contact maps of (R)-BINOL4 with each of the 20 peptide chains. The color code is the same as previously described in Chapter 6.



**Figure C.23:** Conformational contact maps of (R)-BINOL5 with each of the 20 peptoid chains. The color code is the same as previously described in Chapter 6.

## Bibliography

- 1 G. Walsh, *Proteins and Proteomics*. John Wiley & Sons, Ltd, 2015.
- 2 J. A. Wells and T. Clackson, "A hot spot of binding energy in a hormone-receptor interface," *Science.*, vol. 267, no. July, pp. 383–386, 1995.
- 3 V. Krishnan and B. Rupp, "Macromolecular Structure Determination: Comparison of X-ray Crystallography and NMR Spectroscopy," in *eLS*, no. June 2012, Chichester, UK: John Wiley & Sons, Ltd, jun 2012.
- 4 R. N. Zuckermann, "Peptoid origins," *Biopolymers*, vol. 96, no. 5, pp. 545–555, 2011.
- 5 B. Leader, Q. J. Baca, and D. E. Golan, "Protein therapeutics: A summary and pharmacological classification," *Nature Reviews Drug Discovery*, vol. 7, no. 1, pp. 21–39, 2008.
- 6 A. Groß, C. Hashimoto, H. Sticht, and J. Eichler, "Synthetic Peptides as Protein Mimics," *Frontiers in Bioengineering and Biotechnology*, vol. 3, pp. 173–180, jan 2016.
- 7 B. J. Evans, A. T. King, A. Katsifis, L. Matesic, and J. F. Jamie, "Methods to Enhance the Metabolic Stability of Peptide-Based PET Radiopharmaceuticals," *Molecules*, vol. 25, p. 2314, may 2020.
- 8 C. M. Goodman, S. Choi, S. Shandler, and W. F. DeGrado, "Foldamers as versatile frameworks for the design and evolution of function," *Nature Chemical Biology*, vol. 3, no. 5, pp. 252–262, 2007.
- 9 F. Balkenhohl, C. von dem Bussche-Hünnefeld, A. Lansky, and C. Zechel, "Combinatorial Synthesis of Small Organic Molecules," *Angewandte Chemie International Edition in English*, vol. 35, pp. 2288–2337, nov 1996.
- 10 R. J. Simon, R. S. Kania, R. N. Zuckermann, V. D. Huebner, D. a. Jewell, S. Banville, S. Ng, L. Wang, S. Rosenberg, and C. K. Marlowe, "Peptoids: a modular approach to drug discovery.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, no. 20, pp. 9367–9371, 1992.
- 11 G. D. Fasman and E. R. Blout, "Copolymers of L-proline and sarcosine: Synthesis and physical-chemical studies," *Biopolymers*, vol. 1, no. 2, pp. 99–109, 1963.

- 12 M. Sisido, H. Ito, and Y. Imanishi, "Photoselective synthesis of azobenzene-containing cyclic oligosarcosine and its thermal cis  $\rightarrow$  trans isomerization in solution," *Biopolymers*, vol. 21, no. 8, pp. 1597–1611, 1982.
- 13 R. N. Zuckermann, J. M. Kerr, S. B. H. Kent, and W. H. Moos, "Efficient method for the preparation of peptoids [oligo(N-substituted glycines)] by submonomer solid-phase synthesis," *Journal of the American Chemical Society*, vol. 114, pp. 10646–10647, dec 1992.
- 14 R. N. Zuckermann and T. Kodadek, "Peptoids as potential therapeutics," *Current Opinion in Molecular Therapeutics*, vol. 11, no. 3, pp. 299–307, 2009.
- 15 M. T. Dohm, R. Kapoor, and A. E. Barron, "Peptoids: Bio-Inspired Polymers as Potential Pharmaceuticals," *Current Pharmaceutical Design*, vol. 17, pp. 2732–2747, aug 2011.
- 16 "Scopus."
- 17 S. M. Miller, R. J. Simon, S. Ng, R. N. Zuckermann, J. M. Kerr, and W. H. Moos, "Proteolytic studies of homologous peptide and N-substituted glycine peptoid oligomers," *Bioorganic and Medicinal Chemistry Letters*, vol. 4, no. 22, pp. 2657–2662, 1994.
- 18 T. Drakenberg, K. I. Dahlqvist, and S. Forsén, "The barrier to internal rotation in amides. IV. N,N-dimethylamides; substituent and solvent effects," *Journal of Physical Chemistry*, vol. 76, no. 15, pp. 2178–2183, 1972.
- 19 M. Feigel, "Rotation barriers of amides in the gas phase," *Journal of Physical Chemistry*, vol. 87, no. 16, pp. 3054–3058, 1983.
- 20 M. Sisido, Y. Imanishi, and T. Higashimura, "Molecular weight distribution of polysarcosine obtained by NCA polymerization," *Die Makromolekulare Chemie*, vol. 178, no. 11, pp. 3107–3114, 1977.
- 21 S. G. Waley and J. Watson, "The kinetics of the polymerization of sarcosine carbonic anhydride," *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, vol. 199, no. 1059, pp. 499–517, 1949.
- 22 B.-c. Lee, K. A. Dill, and R. N. Zuckermann, "Synthesis of long non-natural sequence-specific heteropolymers," *Polymer Preprints*, vol. 46, pp. 174–175, 2005.

- 23 T. Hjelmgaard, S. Faure, C. Caumes, E. De Santis, A. A. Edwards, and C. Taillefumier, "Convenient Solution-Phase Synthesis and Conformational Studies of Novel Linear and Cyclic  $\alpha,\beta$ -Alternating Peptoids," *Organic Letters*, vol. 11, pp. 4100–4103, sep 2009.
- 24 H. Wu, T. Liang, C. Yin, Y. Jin, Y. Ke, and X. Liang, "Enantioselective synthesis of peptoids with  $\alpha$ -chiral, aromatic side chains," *The Analyst*, vol. 136, no. 21, p. 4409, 2011.
- 25 C. Fetsch, A. Grossmann, L. Holz, J. F. Nawroth, and R. Luxenhofer, "Polypeptides from n-substituted glycine n-carboxyanhydrides: Hydrophilic, hydrophobic, and amphiphilic polymers with poisson distribution," *Macromolecules*, vol. 44, no. 17, pp. 6746–6758, 2011.
- 26 C. Secker, S. M. Brosnan, R. Luxenhofer, and H. Schlaad, "Poly( $\alpha$ -Peptoid)s Revisited: Synthesis, Properties, and Use as Biomaterial," *Macromolecular Bioscience*, vol. 15, pp. 881–891, jul 2015.
- 27 S. C. Lovell, I. W. Davis, W. B. Arendall, P. I. De Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, and D. C. Richardson, "Structure validation by  $C\alpha$  geometry:  $\phi,\psi$  and  $C\beta$  deviation," *Proteins: Structure, Function and Genetics*, vol. 50, no. 3, pp. 437–450, 2003.
- 28 N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, and O. Beckstein, "MDAnalysis: A toolkit for the analysis of molecular dynamics simulations," *Journal of Computational Chemistry*, vol. 32, pp. 2319–2327, jul 2011.
- 29 R. Gowers, M. Linke, J. Barnoud, T. Reddy, M. Melo, S. Seyler, J. Domański, D. Dotson, S. Buchoux, I. Kenney, and O. Beckstein, "MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations," in *Proceedings of the 15th Python in Science Conference*, no. Scipy, pp. 98–105, 2016.
- 30 G. L. Butterfoss, P. D. Renfrew, B. Kuhlman, K. Kirshenbaum, and R. Bonneau, "A preliminary survey of the peptoid folding landscape," *Journal of the American Chemical Society*, vol. 131, no. 46, pp. 16798–16807, 2009.
- 31 W. Humphrey, A. Dalke, and K. Schulten, "VMD: Visual molecular dynamics," *Journal of Molecular Graphics*, vol. 14, pp. 33–38, feb 1996.
- 32 G. N. Ramachandran and V. Sasisekharan, "Conformation of polypeptides and proteins," *Advances in protein chemistry*, vol. 23, pp. 283–438, 1968.

- 33 J. N. Scarsdale, C. Van Alsenoy, V. J. Klimkowski, L. Schäfer, and F. A. Momany, "Ab Initio Studies of Molecular Geometries. 27. Optimized Molecular Structures and Conformational Analysis of N $\alpha$ -Acetyl-N-methylalaninamide and Comparison with Peptide Crystal Data and Empirical Calculations," *Journal of the American Chemical Society*, vol. 105, no. 11, pp. 3438–3445, 1983.
- 34 T. Head-Gordon, M. Head-Gordon, M. J. Frisch, C. L. Brooks, and J. a. Pople, "Theoretical-Study Of Blocked Glycine And Alanine Peptide Analogs," *Journal Of The American Chemical Society*, vol. 113, no. 16, pp. 5989–5997, 1991.
- 35 K. Moehle and H.-J. Hofmann, "Peptides and Peptoids-A Quantum Chemical Structure Comparison," *Biopolymers*, vol. 38, pp. 781–790, 1996.
- 36 E. K. Bradley, J. M. Kerr, L. S. Richter, G. M. Figliozzi, D. A. Goff, R. N. Zuckermann, D. C. Spellmeyer, and J. M. Blaney, "NMR structural characterization of oligo-N-substituted glycine lead compounds from a combinatorial library," *Molecular diversity*, vol. 3, no. 1, pp. 1–15, 1997.
- 37 O. Roy, C. Caumes, Y. Esvan, C. Didierjean, S. Faure, and C. Taillefumier, "The tert-butyl side chain: A powerful means to lock peptoid amide bonds in the cis conformation," *Organic Letters*, vol. 15, no. 9, pp. 2246–2249, 2013.
- 38 N. H. Shah, G. L. Butterfoss, K. Nguyen, B. Yoo, R. Bonneau, D. L. Rabenstein, and K. Kirshenbaum, "(SI) Oligo (N-aryl glycines): A New Twist on Structured Peptoids Oligo (N-aryl glycines): A New Twist on Structured Peptoids," *Journal of the American Chemical Society*, pp. 16622–16632, 2008.
- 39 J. S. Laursen, J. Engel-Andreasen, P. Fristrup, P. Harris, and C. A. Olsen, "Cis – Trans Amide Bond Rotamers in  $\beta$ -Peptoids and Peptoids: Evaluation of Stereoelectronic Effects in Backbone and Side Chains," *Journal of the American Chemical Society*, vol. 135, pp. 2835–2844, feb 2013.
- 40 B. C. Gorske, J. R. Stringer, B. L. Bastian, S. A. Fowler, and 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*, vol. 131, pp. 16555–16567, nov 2009.
- 41 P. Armand, K. Kirshenbaum, A. Falicov, R. L. Dunbrack, K. a. Dill, R. N. Zuckermann, and F. E. Cohen, "Chiral N-substituted glycines can form stable helical conformations," *Folding and Design*, vol. 2, pp. 369–375, dec 1997.
- 42 C. W. Wu, T. J. Sanborn, R. N. Zuckermann, and A. E. Barron, "Peptoid Oligomers with  $\alpha$ -Chiral, Aromatic Side Chains: Effects of Chain Length on

- Secondary Structure,” *Journal of the American Chemical Society*, vol. 123, pp. 2958–2963, apr 2001.
- 43 S. H. Gellman, “Foldamers : A Manifesto,” *Accounts of Chemical Research*, vol. 31, no. 4, pp. 173–180, 1998.
- 44 G. D. Fasman, *Circular Dichroism and the Conformational Analysis of Biomolecules*. Boston, MA: Springer US, 1996.
- 45 N. J. Greenfield, “Using circular dichroism spectra to estimate protein secondary structure,” *Nature Protocols*, vol. 1, pp. 2876–2890, dec 2006.
- 46 N. J. Greenfield and G. D. Fasman, “Computed circular dichroism spectra for the evaluation of protein conformation,” *Biochemistry*, vol. 8, pp. 4108–4116, oct 1969.
- 47 P. Armand, K. Kirshenbaum, R. A. Goldsmith, S. Farr-Jones, A. E. Barron, K. T. V. Truong, K. A. Dill, D. F. Mierke, F. E. Cohen, R. N. Zuckermann, and 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*, vol. 95, pp. 4309–4314, apr 1998.
- 48 C. W. Wu, T. J. Sanborn, K. Huang, R. N. Zuckermann, and A. E. Barron, “Peptoid oligomers with  $\alpha$ -chiral, aromatic side chains: Sequence requirements for the formation of stable peptoid helices,” *Journal of the American Chemical Society*, vol. 123, no. 28, pp. 6778–6784, 2001.
- 49 S. Kakinoki, M. Kitamura, Y. Noguchi, and Y. Arichi, “Effect of residue insertion on the stability of polyproline-I and II structures: Circular dichroism spectroscopic analyses of block-type oligo-prolines (Pro)<sub>m</sub>-Gly/Ala-(Pro)<sub>n</sub>,” *Peptide Science*, vol. 112, no. 5, pp. 2–8, 2020.
- 50 W. C. Johnson, “Proteins Through Circular Dichroism,” *Ann. Rev. Biophys. Biophys. Chem.*, vol. 17, no. Cd, pp. 145–66, 1988.
- 51 P. Crabbé and W. Klyne, “Optical rotatory dispersion and circular dichroism of aromatic compounds: a general survey,” *Tetrahedron*, vol. 23, pp. 3449–3503, jan 1967.
- 52 R. W. Woody, “Aromatic side-chain contributions to the far ultraviolet circular dichroism of peptides and proteins,” *Biopolymers*, vol. 17, pp. 1451–1467, jun 1978.

- 53 R. W. Woody, “[4] Circular dichroism,” in *Methods in Enzymology*, vol. 246, pp. 34–71, 1995.
- 54 D. Eisenberg, “The discovery of the  $\alpha$ -helix and  $\beta$ -sheet, the principal structural features of proteins,” *Proceedings of the National Academy of Sciences*, vol. 100, pp. 11207–11210, sep 2003.
- 55 H. R. Reese, C. C. Shanahan, C. Proulx, and S. Menegatti, “Peptide science: A “rule model” for new generations of peptidomimetics,” *Acta Biomaterialia*, vol. 102, pp. 35–74, 2020.
- 56 K. Kirshenbaum, “PeptoidDataBank,” 2021.
- 57 S. B. Y. Shin, B. Yoo, L. J. Todaro, and K. Kirshenbaum, “Cyclic peptoids,” *Journal of the American Chemical Society*, vol. 129, no. 11, pp. 3218–3225, 2007.
- 58 C. W. Wu, K. Kirshenbaum, T. J. Sanborn, J. a. Patch, K. Huang, K. a. Dill, R. N. Zuckermann, and A. E. Barron, “Structural and Spectroscopic Studies of Peptoid Oligomers with  $\alpha$ -Chiral Aliphatic Side Chains,” *Journal of the American Chemical Society*, vol. 125, pp. 13525–13530, nov 2003.
- 59 R. K. Spencer, G. L. Butterfoss, J. R. Edison, J. R. Eastwood, S. Whitelam, K. Kirshenbaum, and R. N. Zuckermann, “Stereochemistry of polypeptoid chain configurations,” *Biopolymers*, vol. 110, pp. 1–6, jun 2019.
- 60 M. H. Werner, “Nuclear Magnetic Resonance (NMR) Spectroscopy: Structural Analysis of Proteins and Nucleic Acids,” *Encyclopedia of Life Sciences*, pp. 1–10, 2001.
- 61 M. P. Foster, C. A. McElroy, and C. D. Amero, “Solution NMR of Large Molecules and Assemblies †,” *Biochemistry*, vol. 46, pp. 331–340, jan 2007.
- 62 F. P. Gasparro and N. H. Kolodny, “NMR determination of the rotational barrier in N,N-dimethylacetamide. A physical chemistry experiment,” *Journal of Chemical Education*, vol. 54, p. 258, apr 1977.
- 63 F. A. Bovey, J. J. Ryan, and F. P. Hood, “Polymer nuclear magnetic resonance spectroscopy. XV. The conformation of polysarcosine,” *Macromolecules*, vol. 1, no. 4, pp. 305–307, 1968.
- 64 K. Nguyen, M. Iskandar, and 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*, vol. 114, no. 9, pp. 3387–3392, 2010.
- 65 F. Mutulis, I. Mutule, E. Liepinsh, A. Yahorau, M. Lapinsh, S. Kopantshuk, S. Veiksina, A. Rinken, and J. E. Wikberg, “N-alkylated dipeptide amides and related structures as imitations of the melanocortins’ active core,” *Peptides*, vol. 26, pp. 1997–2016, oct 2005.
- 66 E. de Hoffmann and V. Stroobant, *Mass spectrometry Principles and Applications*. Wiley, 3rd editio ed., 2007.
- 67 J. Ren, Y. Tian, E. Hossain, and M. D. Connolly, “Fragmentation Patterns and Mechanisms of Singly and Doubly Protonated Peptoids Studied by Collision Induced Dissociation,” *Journal of The American Society for Mass Spectrometry*, vol. 2393, pp. 646–661, 2016.
- 68 E. Halin, S. Hoyas, V. Lemaur, J. De Winter, S. Laurent, M. D. Connolly, R. N. Zuckermann, J. Cornil, and P. Gerboux, “Backbone Cleavages of Protonated Peptoids upon Collision-Induced Dissociation: Competitive and Consecutive B-Y and A1-YX Reactions,” *Journal of the American Society for Mass Spectrometry*, vol. 30, no. 12, pp. 2726–2740, 2019.
- 69 E. Halin, S. Hoyas, V. Lemaur, J. De Winter, S. Laurent, J. Cornil, J. Roithová, and P. Gerboux, “Side-chain loss reactions of collisionally activated protonated peptoids: A mechanistic insight,” *International Journal of Mass Spectrometry*, vol. 435, pp. 217–226, jan 2019.
- 70 V. Gabelica, A. A. Shvartsburg, C. Afonso, P. Barran, J. L. Benesch, C. Bleiholder, M. T. Bowers, A. Bilbao, M. F. Bush, J. L. Campbell, I. D. Campuzano, T. Causon, B. H. Clowers, C. S. Creaser, E. De Pauw, J. Far, F. Fernandez-Lima, J. C. Fjeldsted, K. Giles, M. Groessl, C. J. Hogan, S. Hann, H. I. Kim, R. T. Kurulugama, J. C. May, J. A. McLean, K. Pagel, K. Richardson, M. E. Ridgeway, F. Rosu, F. Sobott, K. Thalassinou, S. J. Valentine, and T. Wytenbach, “Recommendations for reporting ion mobility Mass Spectrometry measurements,” *Mass Spectrometry Reviews*, vol. 38, pp. 291–320, may 2019.
- 71 V. D’Atri, M. Porrini, F. Rosu, and V. Gabelica, “Linking molecular models with ion mobility experiments. Illustration with a rigid nucleic acid structure,” *Journal of Mass Spectrometry*, vol. 50, no. 5, pp. 711–726, 2015.
- 72 T. P. Creamer and G. D. Rose, “Side-chain entropy opposes alpha-helix formation but rationalizes experimentally determined helix-forming propensities.”

- Proceedings of the National Academy of Sciences*, vol. 89, pp. 5937–5941, jul 1992.
- 73 I. A. Kaltashov and C. Fenselau, “Stability of secondary structural elements in a solvent-free environment: the  $\alpha$  helix,” *Proteins: Structure, Function, and Genetics*, vol. 27, pp. 165–170, feb 1997.
- 74 R. R. Hudgins, M. A. Ratner, and M. F. Jarrold, “Design of Helices That Are Stable in Vacuo,” *Journal of the American Chemical Society*, vol. 120, pp. 12974–12975, dec 1998.
- 75 R. R. Hudgins, Y. Mao, M. a. Ratner, and M. F. Jarrold, “Conformations of GlynH<sup>+</sup> and AlanH<sup>+</sup> Peptides in the Gas Phase,” *Biophysical Journal*, vol. 76, pp. 1591–1597, mar 1999.
- 76 P. D. Renfrew, T. W. Craven, G. L. Butterfoss, K. Kirshenbaum, and R. Bonneau, “A rotamer library to enable modeling and design of peptoid foldamers,” *Journal of the American Chemical Society*, vol. 136, no. 24, pp. 8772–8782, 2014.
- 77 B. C. Gorske, B. L. Bastian, G. D. Geske, and H. E. Blackwell, “Local and Tunable  $n \rightarrow \pi^*$  Interactions Regulate Amide Isomerism in the Peptoid Backbone,” *Journal of the American Chemical Society*, vol. 129, pp. 8928–8929, jul 2007.
- 78 B. C. Gorske, R. C. Nelson, Z. S. Bowden, T. A. Kufe, and A. M. Childs, ““Bridged”  $n \rightarrow \pi^*$  Interactions Can Stabilize Peptoid Helices,” *The Journal of Organic Chemistry*, vol. 78, pp. 11172–11183, nov 2013.
- 79 A. Leach, *Molecular modelling: Principles and applications*. 2nd ed., 2001.
- 80 L. J. Weiser and E. E. Santiso, “Molecular modeling studies of peptoid polymers,” *AIMS Materials Science*, vol. 4, no. 5, pp. 1029–1051, 2017.
- 81 D. T. Mirijanian, R. V. Mannige, R. N. Zuckermann, and S. Whitlam, “Development and use of an atomistic CHARMM-based forcefield for peptoid simulation,” *Journal of Computational Chemistry*, vol. 35, pp. 360–370, feb 2014.
- 82 L. J. Weiser and E. E. Santiso, “A CGenFF-based force field for simulations of peptoids with both cis and trans peptide bonds,” *Journal of Computational Chemistry*, vol. 40, no. 22, pp. 1946–1956, 2019.
- 83 G. L. Butterfoss, B. Yoo, J. N. Jaworski, I. Chorny, K. A. Dill, R. N. Zuckermann, R. Bonneau, K. Kirshenbaum, and V. A. Voelz, “De novo structure prediction and experimental characterization of folded peptoid oligomers,” *Proceedings*

- of the National Academy of Sciences of the United States of America*, vol. 109, no. 36, pp. 14320–14325, 2012.
- 84 J. A. Patch and A. E. Barron, “Helical Peptoid Mimics of Magainin-2 Amide,” *Journal of the American Chemical Society*, vol. 125, pp. 12092–12093, oct 2003.
- 85 T. J. Sanborn, C. W. Wu, R. N. Zuckermann, and A. E. Barron, “Extreme stability of helices formed by water-soluble poly-N-substituted glycines (polypeptoids) with  $\alpha$ -chiral side chains,” *Biopolymers*, vol. 63, no. 1, pp. 12–20, 2002.
- 86 J. R. Stringer, J. A. Crapster, I. A. Guzei, and 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*, vol. 133, pp. 15559–15567, oct 2011.
- 87 A. M. Rosales, H. K. Murnen, S. R. Kline, R. N. Zuckermann, and R. a. Segalman, “Determination of the persistence length of helical and non-helical polypeptoids in solution,” *Soft Matter*, vol. 8, no. 13, p. 3673, 2012.
- 88 M. T. Ruggiero, J. Sibik, R. Orlando, J. A. Zeitler, and T. M. Korter, “Measuring the Elasticity of Poly-L-Proline Helices with Terahertz Spectroscopy,” *Angewandte Chemie*, vol. 128, pp. 6991–6995, jun 2016.
- 89 C. Baldauf, R. Günther, and H.-J. Hofmann, “Helices in peptoids of  $\alpha$ - and  $\beta$ -peptides,” *Physical Biology*, vol. 3, pp. S1–S9, feb 2006.
- 90 H.-M. Shin, C.-M. Kang, M.-H. Yoon, and J. Seo, “Peptoid helicity modulation: precise control of peptoid secondary structures via position-specific placement of chiral monomers,” *Chemical communications (Cambridge, England)*, vol. 50, no. 34, pp. 4465–8, 2014.
- 91 O. Roy, C. Caumes, Y. Esvan, C. Didierjean, S. Faure, and C. Taillefumier, “The tert-butyl side chain: A powerful means to lock peptoid amide bonds in the cis conformation,” *Organic Letters*, vol. 15, no. 9, pp. 2246–2249, 2013.
- 92 B. C. Gorske, E. M. Mumford, C. G. Gerrity, and I. Ko, “A Peptoid Square Helix via Synergistic Control of Backbone Dihedral Angles,” *Journal of the American Chemical Society*, vol. 139, pp. 8070–8073, jun 2017.
- 93 K. Huang, C. W. Wu, T. J. Sanborn, J. A. Patch, K. Kirshenbaum, R. N. Zuckermann, A. E. Barron, and I. Radhakrishnan, “A Threaded Loop Conformation Adopted by a Family of Peptoid Nonamers,” *Journal of the American Chemical Society*, vol. 128, pp. 1733–1738, feb 2006.

- 94 J. A. Crapster, I. A. Guzei, and H. E. Blackwell, "A peptoid ribbon secondary structure," *Angewandte Chemie - International Edition*, vol. 52, pp. 5079–5084, may 2013.
- 95 K. T. Nam, S. a. Shelby, P. H. Choi, A. B. Marciel, R. Chen, L. Tan, T. K. Chu, R. a. Mesch, B.-C. Lee, M. D. Connolly, C. Kisielowski, and R. N. Zuckermann, "Free-floating ultrathin two-dimensional crystals from sequence-specific peptoid polymers.," *Nature materials*, vol. 9, no. 5, pp. 454–60, 2010.
- 96 E. J. Robertson, A. Battigelli, C. Proulx, R. V. Mannige, T. K. Haxton, L. Yun, S. Whitelam, and R. N. Zuckermann, "Design, Synthesis, Assembly, and Engineering of Peptoid Nanosheets," *Accounts of Chemical Research*, vol. 49, pp. 379–389, mar 2016.
- 97 R. V. Mannige, T. K. Haxton, C. Proulx, E. J. Robertson, A. Battigelli, G. L. Butterfoss, R. N. Zuckermann, and S. Whitelam, "Peptoid nanosheets exhibit a new secondary-structure motif," *Nature*, vol. 526, no. 7573, pp. 415–420, 2015.
- 98 N. P. Chongsiriwatana, J. A. Patch, A. M. Czyzewski, M. T. Dohm, A. Ivankin, D. Gidalevitz, R. N. Zuckermann, and A. E. Barron, "Peptoids that mimic the structure, function, and mechanism of helical antimicrobial peptides," *Proceedings of the National Academy of Sciences*, vol. 105, pp. 2794–2799, feb 2008.
- 99 J. Sun and Z. Li, "Peptoid applications in biomedicine and nanotechnology," in *Peptide Applications in Biomedicine, Biotechnology and Bioengineering*, pp. 183–213, Elsevier, 2018.
- 100 T. K. Haxton, R. V. Mannige, R. N. Zuckermann, and S. Whitelam, "Modeling sequence-specific polymers using anisotropic coarse-grained sites allows quantitative comparison with experiment," *Journal of Chemical Theory and Computation*, vol. 11, no. 1, pp. 303–315, 2015.
- 101 Q. Shen, L. Wang, H. Zhou, H. D. Jiang, L. S. Yu, and S. Zeng, "Stereoselective binding of chiral drugs to plasma proteins," *Acta Pharmacologica Sinica*, vol. 34, no. 8, pp. 998–1006, 2013.
- 102 L. A. Nguyen, H. He, and C. Pham-Huy, "Chiral drugs: an overview.," *International journal of biomedical science : IJBS*, vol. 2, pp. 85–100, jun 2006.
- 103 A. Ricci, "Asymmetric Organocatalysis at the Service of Medicinal Chemistry," *ISRN Organic Chemistry*, vol. 2014, pp. 1–29, 2014.

- 104 R. M. Kellogg, J. W. Nieuwenhuijzen, K. Pouwer, T. R. Vries, Q. B. Broxterman, R. F. Grimbergen, B. Kaptein, R. M. Crois, E. de Wever, K. Zwaagstra, and A. C. van der Laan, "Dutch Resolution: Separation of Enantiomers with Families of Resolving Agents. A Status Report," *Synthesis*, no. 10, pp. 1626–1638, 2003.
- 105 E. R. Francotte, "Enantioselective chromatography as a powerful alternative for the preparation of drug enantiomers," *Journal of Chromatography A*, vol. 906, no. 1-2, pp. 379–397, 2001.
- 106 J. Huang, H. Chen, and T. Li, "Improvement of proline chiral stationary phases by varying peptide length and linker," *Journal of Chromatography A*, vol. 1113, no. 1-2, pp. 109–115, 2006.
- 107 K. Ohyama, K. Oyamada, N. Kishikawa, Y. Ohba, M. Wada, T. Maki, K. Nakashima, and N. Kuroda, "Preparation and characterization of poly(L-phenylalanine) chiral stationary phases with varying peptide length," *Journal of Chromatography A*, vol. 1208, no. 1-2, pp. 242–245, 2008.
- 108 K. Ohyama, K. Oyamada, N. Kishikawa, M. Arakawa, Y. Ohba, M. Kamino, M. Wada, K. Nakashima, and N. Kuroda, "Investigation of Novel Peptide Chiral Selectors Prepared by Solid-Phase Synthesis with a tert-Butoxycarbonyl Amino Acid," *Chromatographia*, vol. 70, no. 9-10, pp. 1501–1504, 2009.
- 109 J. Sun and R. N. Zuckermann, "Peptoid polymers: A highly designable bioinspired material," *ACS Nano*, vol. 7, no. 6, pp. 4715–4732, 2013.
- 110 S. D. Ganesh, N. Saha, O. Zandrea, R. N. Zuckermann, and P. Saha, "Peptoids and polypeptoids: biomimetic and bioinspired materials for biomedical applications," *Polymer Bulletin*, vol. 74, no. 8, pp. 3455–3466, 2017.
- 111 S. Xuan and R. N. Zuckermann, "Engineering the atomic structure of sequence-defined peptoid polymers and their assemblies," *Polymer*, vol. 202, no. April, p. 122691, 2020.
- 112 H. Tran, S. L. Gael, M. D. Connolly, and R. N. Zuckermann, "Solid-phase Submonomer Synthesis of Peptoid Polymers and their Self-Assembly into Highly-Ordered Nanosheets," *Journal of Visualized Experiments*, pp. 1–7, nov 2011.
- 113 S. Rinaldi, "The Diverse World of Foldamers: Endless Possibilities of Self-Assembly," *Molecules*, vol. 25, p. 3276, jul 2020.
- 114 K. Kirshenbaum, a. E. Barron, R. a. Goldsmith, P. Armand, E. K. Bradley, K. T. Truong, K. a. Dill, F. E. Cohen, and 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*, vol. 95, no. 8, pp. 4303–4308, 1998.
- 115 R. R. Hudgins and M. F. Jarrold, "Helix formation in unsolvated alanine-based peptides: Helical monomers and helical dimers," *Journal of the American Chemical Society*, vol. 121, no. 14, pp. 3494–3501, 1999.
- 116 S. J. Valentine, A. E. Counterman, C. S. Hoaglund-Hyzer, and D. E. Clemmer, "Intrinsic amino acid size parameters from a series of 113 lysine-terminated tryptic digest peptide ions," *Journal of Physical Chemistry B*, vol. 103, no. 8, pp. 1205–1207, 1999.
- 117 C. Toniolo, G. M. Bonora, F. C. Schilling, and F. A. Bovey, "Proton Magnetic Resonance Study of Linear Sarcosine Oligomers," *Macromolecules*, vol. 13, pp. 1381–1385, nov 1980.
- 118 D. Burley and W. Lenz, "THALIDOMIDE AND CONGENITAL ABNORMALITIES," *The Lancet*, vol. 279, pp. 271–272, feb 1962.
- 119 J. Mc.Conathy and M. J. Owens, "Stereochemistry in Drug Action," *The Primary Care Companion to The Journal of Clinical Psychiatry*, vol. 05, pp. 70–73, jun 2003.
- 120 B. Isralewitz, M. Gao, and K. Schulten, "Steered molecular dynamics and mechanical functions of proteins," *Current Opinion in Structural Biology*, vol. 11, pp. 224–230, apr 2001.
- 121 G. M. Figliozzi, R. Goldsmith, S. C. Ng, S. C. Banville, and R. N. Zuckermann, "Synthesis of N-substituted glycine peptoid libraries.," *Methods in enzymology*, vol. 267, no. 5, pp. 437–47, 1996.
- 122 M. Wilm, "Principles of electrospray ionization," *Molecular and Cellular Proteomics*, vol. 10, no. 7, pp. 1–8, 2011.
- 123 B. T. Ruotolo, J. L. P. Benesch, A. M. Sandercock, S.-J. Hyung, and C. V. Robinson, "Ion mobility–mass spectrometry analysis of large protein complexes," *Nature Protocols*, vol. 3, pp. 1139–1152, jul 2008.
- 124 Q. Duez, F. Chirot, R. Liénard, T. Josse, C. Choi, O. Coulembier, P. Dugourd, J. Cornil, P. Gerbaux, and J. De Winter, "Polymers for Traveling Wave Ion Mobility Spectrometry Calibration," *Journal of The American Society for Mass Spectrometry*, vol. 28, pp. 2483–2491, nov 2017.

- 125 A. A. Shvartsburg and R. D. Smith, "Fundamentals of traveling wave ion mobility spectrometry," *Analytical Chemistry*, vol. 80, no. 24, pp. 9689–9699, 2008.
- 126 E. A. Mason and H. W. Schamp, "Mobility of gaseous ions in weak electric fields," *Annals of Physics*, vol. 4, no. 3, pp. 233–270, 1958.
- 127 V. Gabelica and E. Marklund, "Fundamentals of ion mobility spectrometry," *Current Opinion in Chemical Biology*, vol. 42, pp. 51–59, 2018.
- 128 Y. Sun, S. Vahidi, M. A. Sowole, and L. Konermann, "Protein Structural Studies by Traveling Wave Ion Mobility Spectrometry: A Critical Look at Electrospray Sources and Calibration Issues," *Journal of The American Society for Mass Spectrometry*, vol. 27, pp. 31–40, jan 2016.
- 129 Q. Duez, T. Josse, V. Lemaure, F. Chirot, C. Choi, P. Dubois, P. Dugourd, J. Cornil, P. Gerbaux, and J. De Winter, "Correlation between the Shape of the Ion Mobility Signals and the Stepwise Folding Process of Polylactide Ions," *Journal of Mass Spectrometry*, no. October 2016, pp. 133–138, 2017.
- 130 A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry*. New-York: McGraw-Hill Publishing Company, 1989.
- 131 P. Atkins and R. Friedman, *Molecular Quantum Mechanics*. Oxford: OUP Oxford, 5th ed., 2011.
- 132 M. Born and R. Oppenheimer, "Zur Quantentheorie der Molekeln," *Annalen der Physik*, vol. 389, no. 20, pp. 457–484, 1927.
- 133 P. A. Dirac, "Note on Exchange Phenomena in the Thomas Atom," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 26, no. 3, pp. 376–385, 1930.
- 134 J. C. Slater, "The Theory of Complex Spectra," *Physical Review*, vol. 34, pp. 1293–1322, nov 1929.
- 135 W. Pauli, "Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren," *Zeitschrift für Physik*, vol. 31, pp. 765–783, feb 1925.
- 136 C. C. J. Roothaan, "New Developments in Molecular Orbital Theory," *Reviews of Modern Physics*, vol. 23, pp. 69–89, apr 1951.
- 137 G. G. Hall and P. R. S. L. A, "The molecular orbital theory of chemical valency VIII. A method of calculating ionization potentials," *Proceedings of the Royal*

- Society of London. Series A. Mathematical and Physical Sciences*, vol. 205, pp. 541–552, mar 1951.
- 138 C. Møller and M. S. Plesset, “Note on an Approximation Treatment for Many-Electron Systems,” *Physical Review*, vol. 46, pp. 618–622, oct 1934.
- 139 S. L. Mayo, B. D. Olafson, and W. A. Goddard, “DREIDING: a generic force field for molecular simulations,” *The Journal of Physical Chemistry*, vol. 94, pp. 8897–8909, dec 1990.
- 140 W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, “Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids,” *Journal of the American Chemical Society*, vol. 118, pp. 11225–11236, jan 1996.
- 141 J. M. Wang, R. M. Wolf, J. W. Caldwell, P. a. Kollman, and D. a. Case, “Development and testing of a general amber force field,” *J. Comput. Chem.*, vol. 25, no. 9, pp. 1157–1174, 2004.
- 142 K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, and A. D. Mackerell, “CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields.,” *Journal of computational chemistry*, vol. 31, pp. 671–90, mar 2010.
- 143 J. Huang and A. D. MacKerell, “CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data,” *Journal of Computational Chemistry*, vol. 34, pp. 2135–2145, sep 2013.
- 144 Dassault Systèmes BIOVIA, Materials Studio, 18.0, San Diego: Dassault Systèmes, 2018.
- 145 S. Saiev, L. Bonnaud, P. Dubois, D. Beljonne, and R. Lazzaroni, “Modeling the formation and thermomechanical properties of polybenzoxazine thermosets,” *Polymer Chemistry*, vol. 8, no. 38, pp. 5988–5999, 2017.
- 146 L. Sun, L. Yang, Y.-D. Zhang, Q. Shi, R.-F. Lu, and W.-Q. Deng, “Accurate van der Waals force field for gas adsorption in porous materials,” *Journal of Computational Chemistry*, vol. 38, no. 23, pp. 1991–1999, 2017.
- 147 J. Gasteiger and M. Marsili, “A new model for calculating atomic charges in molecules,” *Tetrahedron Letters*, vol. 19, pp. 3181–3184, jan 1978.

- 148 L. Zanotto, G. Heerdt, P. C. Souza, G. Araujo, and M. S. Skaf, “High performance collision cross section calculation-HPCCS,” *Journal of Computational Chemistry*, pp. 1–7, 2018.
- 149 T. Wyttenbach, G. Von Helden, J. J. Batka, D. Carlat, and M. T. Bowers, “Effect of the long-range potential on ion mobility measurements,” *Journal of the American Society for Mass Spectrometry*, vol. 8, no. 3, pp. 275–282, 1997.
- 150 A. A. Shvartsburg and M. F. Jarrold, “An exact hard-spheres scattering model for the mobilities of polyatomic ions,” *Chemical Physics Letters*, vol. 261, no. 1-2, pp. 86–91, 1996.
- 151 M. F. Mesleh, J. M. Hunter, A. A. Shvartsburg, G. C. Schatz, and M. F. Jarrold, “Structural Information from Ion Mobility Measurements: Effects of the Long-Range Potential,” *The Journal of Physical Chemistry*, vol. 100, pp. 16082–16086, jan 1996.
- 152 J. S. Prell, *Modelling Collisional Cross Sections*. Elsevier B.V., 1 ed., 2018.
- 153 S. A. Ewing, M. T. Donor, J. W. Wilson, and J. S. Prell, “Collidoscope: An Improved Tool for Computing Collisional Cross-Sections with the Trajectory Method,” *Journal of the American Society for Mass Spectrometry*, vol. 28, no. 4, pp. 587–596, 2017.
- 154 M. Aminpour, C. Montemagno, and J. A. Tuszynski, “An overview of molecular modeling for drug discovery with specific illustrative examples of applications,” *Molecules*, vol. 24, no. 9, 2019.
- 155 V. a. Voelz, K. a. Dill, and I. Chorny, “Peptoid conformational free energy landscapes from implicit-solvent molecular simulations in AMBER,” *Biopolymers*, vol. 96, no. 5, pp. 639–650, 2011.
- 156 S. Mukherjee, G. Zhou, C. Michel, and V. A. Voelz, “Insights into Peptoid Helix Folding Cooperativity from an Improved Backbone Potential,” *Journal of Physical Chemistry B*, vol. 119, no. 50, pp. 15407–15417, 2015.
- 157 C. Decroo, E. Colson, M. Demeyer, V. Lemaur, G. Caulier, I. Eeckhaut, J. Cornil, P. Flammang, and P. Gerbaux, “Tackling saponin diversity in marine animals by mass spectrometry: data acquisition and integration,” *Analytical and Bioanalytical Chemistry*, vol. 409, pp. 3115–3126, may 2017.

- 158 W. Mästle, U. Link, W. Witschel, U. Thewalt, T. Weber, and M. Rothe, “Conformation and formation tendency of the cyclotetrapeptidecyclo(D-pro-D-pro-L-pro-L-pro): Experimental results and molecular modeling studies,” *Biopolymers*, vol. 31, no. 6, pp. 735–744, 1991.
- 159 K. B. Borisenko, H. J. Reavy, Q. Zhao, and E. W. Abel, “Adhesion of protein residues to substituted (111) diamond surfaces: An insight from density functional theory and classical molecular dynamics simulations,” *Journal of Biomedical Materials Research - Part A*, vol. 86, no. 4, pp. 1113–1121, 2008.
- 160 V. Lemaury, L. Muccioli, C. Zannoni, D. Beljonne, R. Lazzaroni, J. Cornil, and Y. Olivier, “On the Supramolecular Packing of High Electron Mobility Naphthalene Diimide Copolymers: The Perfect Registry of Asymmetric Branched Alkyl Side Chains,” *Macromolecules*, vol. 46, pp. 8171–8178, oct 2013.
- 161 B. Guillermin, V. Lemaury, B. Ernould, J. Cornil, R. Lazzaroni, J.-F. Gohy, P. Dubois, and O. Coulembier, “A one-pot two-step efficient metal-free process for the generation of PEO-b-PCL-b-PLA amphiphilic triblock copolymers,” *RSC Advances*, vol. 4, no. 20, p. 10028, 2014.
- 162 Q. Duez, S. Moins, O. Coulembier, J. De Winter, J. Cornil, and P. Gerbaux, “Assessing the Structural Heterogeneity of Isomeric Homo and Copolymers: an Approach Combining Ion Mobility Mass Spectrometry and Molecular Dynamics Simulations,” *Journal of the American Society for Mass Spectrometry*, vol. 31, pp. 2379–2388, nov 2020.
- 163 Q. Duez, H. Metwally, S. Hoyas, V. Lemaury, J. Cornil, J. De Winter, L. Konermann, and P. Gerbaux, “Effects of electrospray mechanisms and structural relaxation on polylactide ion conformations in the gas phase: Insights from ion mobility spectrometry and molecular dynamics simulations,” *Physical Chemistry Chemical Physics*, vol. 22, no. 7, pp. 4193–4204, 2020.
- 164 Dassault Systèmes BIOVIA, Materials Studio, 6.0, San Diego: Dassault Systèmes, 2011.
- 165 propane data:  
<https://webbook.nist.gov/cgi/cbook.cgi?ID=C74986&Mask=4#Thermo-Phase>, accessed 2018-06-04.
- 166 n-butane data:  
<https://webbook.nist.gov/cgi/cbook.cgi?ID=C106978&Mask=4#Thermo-Phase>, accessed 2018-06-04.

167 n-pentane data:

<https://webbook.nist.gov/cgi/cbook.cgi?ID=C109660&Mask=4#Thermo-Phase>, accessed 2018-06-04.

168 n-hexane data:

<https://webbook.nist.gov/cgi/cbook.cgi?ID=C110543&Mask=4#Thermo-Phase>, accessed 2018-06-04.

169 cyclohexane data:

<https://webbook.nist.gov/cgi/cbook.cgi?ID=C110827&Mask=4#Thermo-Phase>, accessed 2018-06-04.

170 ethyl acetate data:

<https://webbook.nist.gov/cgi/cbook.cgi?ID=C141786&Units=SI&Mask=4#Thermo-Phase>, accessed 2018-06-04.

171 H. Sun, “Compass: An ab initio force-field optimized for condensed-phase applications - Overview with details on alkane and benzene compounds,” *Journal of Physical Chemistry B*, vol. 102, no. 38, pp. 7338–7364, 1998.

172 J. R. Haler, P. Massonnet, F. Chirot, C. Kune, C. Comby-Zerbino, J. Jordens, M. Honing, Y. Mengerink, J. Far, P. Dugourd, and E. De Pauw, “Comparison of Different Ion Mobility Setups Using Poly (Ethylene Oxide) PEO Polymers: Drift Tube, TIMS, and T-Wave,” *Journal of the American Society for Mass Spectrometry*, vol. 29, no. 1, pp. 114–120, 2018.

173 L. Gu and R. A. Robinson, “Multiple proteases to localize oxidation sites,” *PLoS ONE*, vol. 10, no. 3, pp. 1–14, 2015.

174 J. R. Haler, D. Morsa, P. Lecomte, C. Jérôme, J. Far, and E. De Pauw, “Predicting Ion Mobility-Mass Spectrometry trends of polymers using the concept of apparent densities,” *Methods*, vol. 144, pp. 125–133, 2018.

175 J. Gidden, T. Wyttenbach, A. T. Jackson, J. H. Scrivens, and M. T. Bowers, “Gas-phase conformations of synthetic polymers: Poly(ethylene glycol), poly(propylene glycol), and poly(tetramethylene glycol),” *Journal of the American Chemical Society*, vol. 122, no. 19, pp. 4692–4699, 2000.

176 A. Memboeuf, K. Vékey, and G. Lendvay, “Structure and Energetics of Poly(Ethylene Glycol) Cationized by  $\text{Li}^+$ ,  $\text{Na}^+$ ,  $\text{K}^+$  and  $\text{Cs}^+$ : A First-Principles Study,” *European Journal of Mass Spectrometry*, vol. 17, pp. 33–46, feb 2011.

- 177 A. Chakrabartty and R. Baldwin, "Stability of  $\alpha$ -Helices," in *Advances in Protein Chemistry Volume 46*, vol. 46, pp. 141–176, Elsevier, 1995.
- 178 J. De Winter, V. Lemaur, R. Ballivian, F. Chiro, O. Coulembier, R. Antoine, J. Lemoine, J. Cornil, P. Dubois, P. Dugourd, and P. Gerbaux, "Size Dependence of the Folding of Multiply Charged Sodium Cationized Polylactides Revealed by Ion Mobility Mass Spectrometry and Molecular Modelling," *Chemistry - A European Journal*, vol. 17, pp. 9738–9745, aug 2011.
- 179 J. R. N. Haler, V. Lemaur, J. Far, C. Kune, P. Gerbaux, J. Cornil, and E. De Pauw, "Sodium Coordination and Protonation of Poly(ethoxy phosphate) Chains in the Gas Phase Probed by Ion Mobility-Mass Spectrometry," *Journal of the American Society for Mass Spectrometry*, vol. 31, pp. 633–641, mar 2020.
- 180 S. Tanaka and A. Nakajima, "Conformation of Poly(N-methyl glycine) Random Chains," *Polymer Journal*, vol. 1, no. 1, pp. 71–81, 1970.
- 181 S. Poyer, C. Comby-Zerbino, C. M. Choi, L. Macaleese, C. Deo, N. Bogliotti, J. Xie, J. Y. Salpin, P. Dugourd, and F. Chiro, "Conformational Dynamics in Ion Mobility Data," *Analytical Chemistry*, vol. 89, no. 7, pp. 4230–4237, 2017.
- 182 M. F. Bush, I. D. Campuzano, and C. V. Robinson, "Ion mobility mass spectrometry of peptide ions: Effects of drift gas and calibration strategies," *Analytical Chemistry*, vol. 84, no. 16, pp. 7124–7130, 2012.
- 183 B. C. Gorske and H. E. Blackwell, "Tuning peptoid secondary structure with pentafluoroaromatic functionality: a new design paradigm for the construction of discretely folded peptoid structures.," *Journal of the American Chemical Society*, vol. 128, pp. 14378–87, nov 2006.
- 184 S. A. Fowler, R. Luechapanichkul, and H. E. Blackwell, "Synthesis and Characterization of Nitroaromatic Peptoids: Fine Tuning Peptoid Secondary Structure through Monomer Position and Functionality," *The Journal of Organic Chemistry*, vol. 74, pp. 1440–1449, feb 2009.
- 185 O. Roy, G. Dumonteil, S. Faure, L. Jouffret, A. Kriznik, and C. Taillefumier, "Highly homogeneous and robust PolyProline type I helices from peptoids with non-aromatic  $\alpha$ -chiral side chains," *Journal of the American Chemical Society*, p. jacs.7b07475, 2017.
- 186 B. S. Kinnear, M. R. Hartings, and M. F. Jarrold, "Helix unfolding in unsolvated peptides," *Journal of the American Chemical Society*, vol. 123, no. 24, pp. 5660–5667, 2001.

- 187 D. T. Kaleta and M. F. Jarrold, "Peptide pinwheels," *Journal of the American Chemical Society*, vol. 124, no. 7, pp. 1154–1155, 2002.
- 188 L. Shi, A. E. Holliday, H. Shi, F. Zhu, M. A. Ewing, D. H. Russell, and D. E. Clemmer, "Characterizing intermediates along the transition from polyproline I to polyproline II using ion mobility spectrometry-mass spectrometry," *Journal of the American Chemical Society*, vol. 136, no. 36, pp. 12702–12711, 2014.
- 189 F. Haeffner, J. K. Merle, and K. K. Irikura, "N-protonated isomers as gateways to peptide ion fragmentation," *Journal of the American Society for Mass Spectrometry*, vol. 22, no. 12, pp. 2222–2231, 2011.
- 190 M. Kohtani, M. F. Jarrold, S. Wee, and R. A. J. O'Hair, "Metal Ion Interactions with Polyalanine Peptides," *The Journal of Physical Chemistry B*, vol. 108, pp. 6093–6097, may 2004.
- 191 Q. Duez, H. Metwally, and L. Konermann, "Electrospray Ionization of Polypropylene Glycol: Rayleigh-Charged Droplets, Competing Pathways, and Charge State-Dependent Conformations," *Analytical Chemistry*, vol. 90, no. 16, pp. 9912–9920, 2018.
- 192 G. von Helden, T. Wyttenbach, and M. T. Bowers, "Conformation of Macromolecules in the Gas Phase: Use of Matrix-Assisted Laser Desorption Methods in Ion Chromatography," *Science*, vol. 267, pp. 1483–1485, mar 1995.
- 193 A. T. Jackson, J. H. Scrivens, J. P. Williams, E. S. Baker, J. Gidden, and M. T. Bowers, "Microstructural and conformational studies of polyether copolymers," *International Journal of Mass Spectrometry*, vol. 238, pp. 287–297, nov 2004.
- 194 D. Van Der Spoel, P. J. Van Maaren, P. Larsson, and N. Tîmneanu, "Thermodynamics of hydrogen bonding in hydrophilic and hydrophobic media," *Journal of Physical Chemistry B*, vol. 110, no. 9, pp. 4393–4398, 2006.
- 195 S. Y. Sheu, D. Y. Yang, H. L. Selzle, and E. W. Schlag, "Energetics of hydrogen bonds in peptides," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 22, pp. 12683–12687, 2003.
- 196 D. G. Blackmond, "The Origin of Biological Homochirality," *Cold Spring Harbor Perspectives in Biology*, vol. 2, pp. a002147–a002147, may 2010.
- 197 S. D. Banik and N. Nandi, "Chirality and Protein Biosynthesis," in *Topics in current chemistry*, vol. 333, pp. 255–305, 2012.

- 198 S. Mane, "Racemic drug resolution: A comprehensive guide," *Analytical Methods*, vol. 8, no. 42, pp. 7567–7586, 2016.
- 199 J. T. Mohr, M. R. Krout, and B. M. Stoltz, "Natural products as inspiration for the development of asymmetric catalysis," *Nature*, vol. 455, no. 7211, pp. 323–332, 2008.
- 200 A. Gogoi, N. Mazumder, S. Konwer, H. Ranawat, N. T. Chen, and G. Y. Zhuo, "Enantiomeric recognition and separation by chiral nanoparticles," *Molecules*, vol. 24, no. 6, pp. 1–31, 2019.
- 201 Y. Wang and A. Chen, "Crystallization-Based Separation of Enantiomers," *Stereoselective Synthesis of Drugs and Natural Products*, pp. 1–20, 2013.
- 202 K. Robards, P. Haddad, and P. Jackson, "High-performance Liquid Chromatography—Separations," *Principles and Practice of Modern Chromatographic Methods*, pp. 305–380, 2004.
- 203 Y. Xiao, S. C. Ng, T. T. Y. Tan, and Y. Wang, "Recent development of cyclodextrin chiral stationary phases and their applications in chromatography," *Journal of Chromatography A*, vol. 1269, pp. 52–68, 2012.
- 204 T. Ikai, C. Yamamoto, M. Kamigaito, and Y. Okamoto, "Immobilized polysaccharide-based chiral stationary phases for HPLC," *Polymer Journal*, vol. 38, no. 2, pp. 91–108, 2006.
- 205 J. Haginaka, "Protein-based chiral stationary phases for high-performance liquid chromatography enantioseparations," *Journal of Chromatography A*, vol. 906, pp. 253–273, jan 2001.
- 206 K. Ohyama, K. Oyamada, N. Kishikawa, M. Wada, Y. Ohba, K. Nakashima, and N. Kuroda, "Effects of Temperature and Mobile Phase Condition on Chiral Recognition of Poly(L-phenylalanine) Chiral Stationary Phase," *Chromatographia*, vol. 74, pp. 467–470, sep 2011.
- 207 C. Zhao and N. M. Cann, "The docking of chiral epoxides on the Whelk-O1 stationary phase: A molecular dynamics study," *Journal of Chromatography A*, vol. 1149, no. 2, pp. 197–218, 2007.
- 208 C. F. Zhao and N. M. Cann, "Molecular Dynamics Study of Chiral Recognition for the Whelk-O1 Chiral Stationary Phase," *Analytical Chemistry*, vol. 80, pp. 2426–2438, apr 2008.

- 209 M. Ashtari and N. Cann, "Poly-proline-based chiral stationary phases: A molecular dynamics study of triproline, tetraproline, pentaproline and hexaproline interfaces," *Journal of Chromatography A*, vol. 1265, pp. 70–87, nov 2012.
- 210 P. Peluso, A. Dessì, R. Dallochio, V. Mamane, and S. Cossu, "Recent studies of docking and molecular dynamics simulation for liquid-phase enantioseparations," *ELECTROPHORESIS*, vol. 40, pp. 1881–1896, aug 2019.
- 211 X. Wang, C. J. Jameson, and S. Murad, "Modeling Enantiomeric Separations as an Interfacial Process Using Amylose Tris(3,5-dimethylphenyl carbamate) (ADMPC) Polymers Coated on Amorphous Silica," *Langmuir*, vol. 36, no. 5, pp. 1113–1124, 2020.
- 212 O. M. Roscioni, C. Zannoni, L. Muccioli, R. G. Della Valle, A. Pizzirusso, and M. Ricci, "Predicting the Anchoring of Liquid Crystals at a Solid Surface: 5-Cyanobiphenyl on Cristobalite and Glassy Silica Surfaces of Increasing Roughness," *Langmuir*, vol. 29, no. 28, pp. 8950–8958, 2013.
- 213 J. Du and A. N. Cormack, "Molecular dynamics simulation of the structure and hydroxylation of silica glass surfaces," *Journal of the American Ceramic Society*, vol. 88, no. 9, pp. 2532–2539, 2005.
- 214 L. T. Zhuravlev, "Concentration of Hydroxyl Groups on the Surface of Amorphous Silicas," *Langmuir*, vol. 3, no. 3, pp. 316–318, 1987.
- 215 O. M. Roscioni, L. Muccioli, A. Mityashin, J. Cornil, and C. Zannoni, "Structural Characterization of Alkylsilane and Fluoroalkylsilane Self-Assembled Monolayers on SiO<sub>2</sub> by Molecular Dynamics Simulations," *The Journal of Physical Chemistry C*, vol. 120, pp. 14652–14662, jul 2016.
- 216 R. T. Cygan, J. J. Liang, and A. G. Kalinichev, "Molecular models of hydroxide, oxyhydroxide, and clay phases and the development of a general force field," *Journal of Physical Chemistry B*, vol. 108, no. 4, pp. 1255–1266, 2004.
- 217 J. M. Rotter and H. N. Knickle, "Isobaric Vapor-Liquid Equilibrium Data for the System n-Hexane-2-Propanol," *Journal of Chemical and Engineering Data*, vol. 22, no. 3, pp. 246–248, 1977.
- 218 "CGenFF interface at paramchem.org: <https://cgenff.umaryland.edu/> (accessed Apr 15, 2019)."
- 219 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé, and K. Schulten, "Scalable molecular dynamics

- with NAMD,” *Journal of Computational Chemistry*, vol. 26, pp. 1781–1802, dec 2005.
- 220 J. C. Phillips, D. J. Hardy, J. D. C. Maia, J. E. Stone, J. V. Ribeiro, R. C. Bernardi, R. Buch, G. Fiorin, J. Hénin, W. Jiang, R. McGreevy, M. C. R. Melo, B. K. Radak, R. D. Skeel, A. Singharoy, Y. Wang, B. Roux, A. Aksimentiev, Z. Luthey-Schulten, L. V. Kalé, K. Schulten, C. Chipot, and E. Tajkhorshid, “Scalable molecular dynamics on CPU and GPU architectures with NAMD,” *The Journal of Chemical Physics*, vol. 153, p. 044130, jul 2020.
- 221 E. R. Cruz-Chu, A. Aksimentiev, and K. Schulten, “Water-silica force field for simulating nanodevices,” *Journal of Physical Chemistry B*, vol. 110, no. 43, pp. 21497–21508, 2006.
- 222 T. Darden, D. York, and L. Pedersen, “Particle mesh Ewald: An  $N \log(N)$  method for Ewald sums in large systems,” *The Journal of Chemical Physics*, vol. 98, pp. 10089–10092, jun 1993.
- 223 R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II*, vol. 31 of *Springer Series in Solid-State Sciences*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1991.
- 224 S. E. Feller, Y. Zhang, R. W. Pastor, and B. R. Brooks, “Constant pressure molecular dynamics simulation: The Langevin piston method,” *The Journal of Chemical Physics*, vol. 103, pp. 4613–4621, sep 1995.
- 225 MOZZI RL and WARREN BE, “Structure of Vitreous Silica,” *Journal of Applied Crystallography*, vol. 2, no. pt 4, pp. 164–172, 1969.
- 226 S. Izrailev, S. Stepaniants, M. Balsera, Y. Oono, and K. Schulten, “Molecular dynamics study of unbinding of the avidin-biotin complex,” *Biophysical Journal*, vol. 72, pp. 1568–1581, apr 1997.
- 227 J. M. Stubbs and J. I. Siepmann, “Aggregation in dilute solutions of 1-hexanol in n-hexane: A Monte Carlo simulation study,” *Journal of Physical Chemistry B*, vol. 106, no. 15, pp. 3968–3978, 2002.
- 228 M. Požar, B. Lovrinčević, L. Zoranić, T. Primorać, F. Sokolić, and A. Perera, “Micro-heterogeneity versus clustering in binary mixtures of ethanol with water or alkanes,” *Physical Chemistry Chemical Physics*, vol. 18, no. 34, pp. 23971–23979, 2016.

- 229 D. Van Der Spoel and H. J. Berendsen, "Molecular dynamics simulations of leu-enkephalin in water and DMSO," *Biophysical Journal*, vol. 72, no. 5, pp. 2032–2041, 1997.
- 230 S. Daly, F. Rosu, and V. Gabelica, "Mass-resolved electronic circular dichroism ion spectroscopy," *Science*, vol. 368, no. 6498, pp. 1465–1468, 2020.
- 231 O. Roy, G. Dumonteil, S. Faure, L. Jouffret, A. Kriznik, and C. Taillefumier, "Homogeneous and Robust Polyproline Type I Helices from Peptoids with Nonaromatic  $\alpha$ -Chiral Side Chains," *Journal of the American Chemical Society*, vol. 139, pp. 13533–13540, sep 2017.
- 232 G. Angelici, N. Bhattacharjee, O. Roy, S. Faure, C. Didierjean, L. Jouffret, F. Jolibois, L. Perrin, and C. Taillefumier, "Weak backbone CH  $\cdots$  O  $\cdots$  C and side chain tBu  $\cdots$  tBu London interactions help promote helix folding of achiral NtBu peptoids," *Chemical Communications*, vol. 52, no. 24, pp. 4573–4576, 2016.
- 233 H. Wu, X. Su, K. Li, H. Yu, Y. Ke, and X. Liang, "Improvement of peptoid chiral stationary phases by modifying the terminal group of selector," *Journal of Chromatography A*, vol. 1265, pp. 181–185, 2012.
- 234 F. De Riccardis, "The Challenge of Conformational Isomerism in Cyclic Peptoids," *European Journal of Organic Chemistry*, pp. 1–15, 2020.
- 235 N. Bhattacharjee, N. Bhattacharjee, L. Perrin, and F. Jolibois, "Relating circular dichroism to atomic structure by means of MD simulations and computed CD spectra with  $\alpha$ -peptoids as an example," *Physical Chemistry Chemical Physics*, vol. 22, no. 23, pp. 13192–13200, 2020.
- 236 D. Kim, Y. J. Lee, D. H. Ahn, J. W. Song, J. Seo, and H. Lee, "Peptoid-Conjugated Magnetic Field-Sensitive Exciplex System at High and Low Solvent Polarities," *Journal of Physical Chemistry Letters*, vol. 11, no. 12, pp. 4668–4677, 2020.

---

## PUBLICATIONS

---

1. S. Hoyas, V. Lemaury, Q. Duez, F. Saintmont, E. Halin, J. De Winter, P. Gerbaux, J. Cornil, "PEPDROID: Development of a Generic DREIDING-Based Force Field for the Assessment of Peptoid Secondary Structures", *Advanced Theory and Simulations*, vol. 1, 1800089, **2018**
2. E. Halin, S. Hoyas, V. Lemaury, J. De Winter, S. Laurent, J. Roithová, J. Cornil, P. Gerbaux, "Side-chain loss reactions of collisionally activated protonated peptoids: A mechanistic insight", *International Journal of Mass Spectrometry*, vol. 435, 217, **2018**
3. E. Halin, S. Hoyas, V. Lemaury, J. De Winter, S. Laurent, M. D. Connolly, R. N. Zuckermann, J. Cornil, P. Gerbaux, "Backbone Cleavages of Protonated Peptoids upon Collision-Induced Dissociation: Competitive and Consecutive B-Y and A<sub>1</sub>-Y<sub>X</sub> Reactions", *Journal of the American Society for Mass Spectrometry*, vol. 30, 2726, **2019**
4. S. Hoyas, E. Halin, V. Lemaury, J. De Winter, P. Gerbaux, J. Cornil, "Helicity of Peptoid Ions in the Gas Phase", *Biomacromolecules*, vol. 21, 903, **2020**
5. Q. Duez, H. Metwally, S. Hoyas, V. Lemaury, J. Cornil, J. De Winter, L. Konermann, P. Gerbaux, "Effects of electrospray mechanisms and structural relaxation on polylactide ion conformations in the gas phase: insights from ion mobility spectrometry and molecular dynamics simulations", *Physical Chemistry Chemical Physics*, vol. 22, 4193, **2020**
6. S. Moins, S. Hoyas, V. Lemaury, B. Orhan, K. Della Chiaie, R. Lazzaroni, D. Taton, A. P. Dove, O. Coulembier, "Stereoselective ROP of *rac*- and *meso*-Lactides Using Achiral TBD as Catalyst", *Catalysts*, vol. 10, 620, **2020**
7. B. Tassignon, P. Weber, S. Hoyas, A. Galanti, J. De Winter, J. Cornil, P. Samori, P. Gerbaux, "Synthèse et Etudes Photochimiques de Peptoides Porteurs de Groupements Azobenzènes pour le Stockage Chimique de l'Énergie Solaire", *Chimie Nouvelle*, vol. 136, 1, **2021**
8. S. Hoyas, O. M. Roscioni, C. Tonneaux, P. Gerbaux, J. Cornil, L. Muccioli, "Peptoids as a Chiral Stationary Phase for Liquid Chromatography: Insights from Molecular Dynamics Simulations", *Biomacromolecules*, vol. 22, 2573, **2021**
9. S. Hoyas, P. Weber, E. Halin, O. Coulembier, J. De Winter, J. Cornil, P. Gerbaux, "Helical peptoid ions in the gas phase: Thwarting the charge solvation effect by H-bond compensation", *Biomacromolecules*, **2021**  
10.1021/acs.biomac.1c00623
10. Q. Duez, S. Hoyas, T. Josse, J. Cornil, P. Gerbaux, J. De Winter, "Gas-phase structure of polymer ions: Tying together theoretical approaches and ion mobility spectrometry", *Mass Spectrometry Reviews*, *submitted*