

Review

Molecular modeling studies of peptoid polymers

Laura J. Weiser and Erik E. Santiso *

Department of Chemical and Biomolecular Engineering, NC State University, Campus Box 7905, Raleigh, NC 27695-7905, United States

* **Correspondence:** Email: eesantis@ncsu.edu; Tel: +1-919-515-2520.

Abstract: Peptoids, or poly-N-substituted glycines, are synthetic polymers composed of a protein backbone with side chains attached to the nitrogen atoms rather than the α -carbons. Peptoids are biomimetic and protease resistant and have been explored for a variety of applications including pharmaceuticals and coatings. They are also foldamer-type materials that can adopt diverse structures based on the sequences of their side chains. Design of new peptoid sequences may lead to the creation of many interesting materials. Given the large number of possible peptoid side chains, computer models predicting peptoid structure-side chain relationships are desirable. In this paper, we provide a survey of computational efforts to understand and predict peptoid structures. We describe simulations at several levels of theory, along with their assumptions and results. We also discuss some challenges for future peptoid computational research.

Keywords: peptoids; foldamer; molecular simulation; biomimetic

Abbreviations

AMBER	Assisted Model Building with Energy Refinement
B3LYP	Becke, 3-parameter, Lee-Yang-Parr hybrid functional
CHARMM	Chemistry at Harvard Macromolecular Mechanics
BCC	Bond-Centered Charges
CD	Circular Dichroism
GAFF	General AMBER Force Field
GBSA	Generalized Born Surface Area implicit solvent model
HF	Hartree-Fock
IEFPCM	Integral Equation Formalism for the Polarizable Continuum Model
MF-CG-TOID	Molecular Foundry Coarse-Grained Peptoid model

MFTOID	Molecular Foundry Peptoid Force Field
MP2	Møller-Plesset perturbation theory
NMR	Nuclear Magnetic Resonance spectroscopy
PCM	Polarizable Continuum Model
PPI Helix	Polyproline type I Helix
PPII Helix	Polyproline type II Helix
REMD	Replica Exchange Molecular Dynamics

1. Introduction

Peptoids are synthetic, protein-like polymers composed of poly-glycine backbones with side chains attached to their backbone nitrogen atoms [1,2] (See Figure 1). Peptoids are protease resistant and biocompatible and can be designed to mimic peptides [3], diffuse through membranes [4], and bind biological targets [5]. They have potential applications as drug molecules [6], antimicrobials [7], lung surfactants [8], and catalysts [9]. Peptoids can be synthesized with specific sequences in a step-by-step approach. Their synthesis, applications, and experimental characterizations have been extensively reviewed elsewhere [1,2].

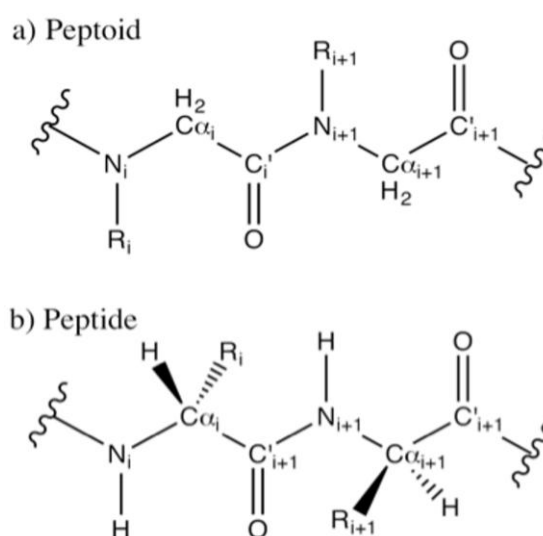


Figure 1. A Peptoid (a) and a peptide (b) Peptoids and peptides share the same backbone, but peptoids have side chains attached to their nitrogen atoms. This removes backbone chirality and backbone hydrogen bonding capability. The repeat unit of the peptoid backbone resembles the amino acid glycine, with a side chain replacing the nitrogen-bound hydrogen atom. Due to this, peptoids are also called poly-N-substituted glycines.

Peptoids are a kind of foldamer [10], i.e., linear polymers that fold into diverse 3D structures depending on the sequence of their attached side chains—just as amino acids combine to yield a diverse library of proteins. Single peptoid chains have been shown to adopt ordered structures including polyproline type I helices (PPI) [11], polyproline type II helices (PPII) [12], a threaded loop [13], flat “ribbons” [14] and “ Σ -strands” [15], which are shown in Figure 2. Groups of peptoid

chains can further self-assemble into nanostructures including microspheres [16], superhelices [17], and thin, bilayered nanosheets [18]. The chief focuses of this review are simulation studies that enhance our understanding of peptoids as foldamers, i.e., those studies that describe the relation between side chains and backbone conformations.

As researchers attempt to develop peptoids into useful materials, a key challenge is identifying side chains, and combinations of side chains, that drive backbone folding in a predictable way. This challenge is closely related to the well-known protein folding problem [19], whereby researchers aspire to predict folded protein structures from sequences of amino acids. Peptoid researchers have mainly approached this problem by identifying side chains expected to exert local structural preferences through sterics [11,12], electrostatics [20], hydrogen bonding [21], or $n-\pi^*$ bonding [22]. These candidates are then synthesized into peptoid polymers and characterized using experimental methods such as X-ray crystallography [23], circular dichroism [24], and NMR [25].

Molecular modeling has played a prominent role in understanding the folding behavior of peptoids. Modeling studies have helped characterize and even predict some of the structures described above. Given the large number of possible peptoid side chains (potentially synthesized from over 300 commercially available primary amines [2]), and thus the astronomical number of possible combinations of side chains, molecular simulations will continue to play a crucial role in peptoid development, and will benefit from the extensive body of work devoted to the protein folding problem.

Computational approaches to the protein folding problem consist of methods, models, and algorithms that read in amino acid sequences and return 3-D structure predictions. These predictions can be based on physical, force field models with parameters derived from theory or tuned to experimental data [26,27], and/or they can be heuristically scored according to similarity with existing sequences [28,29,30]. We focus our discussion on the former approach, as it is most transferrable to peptoids: little experimental data exists for most side chains and theoretical data is easier to obtain.

Protein folding can be studied at several scales of complexity. Quantum mechanical modeling can determine small molecule, monomer-level energy preferences, at high resolution but also high computational costs. All-atom methods model molecules as combinations of atoms that interact via potential energy functions (“force fields”), whose parameters describe atomic-scale interactions such as covalent bonds, hydrogen bonds, van der Waals forces, electrostatics, and solvent effects. These force fields may be tuned to both *ab initio* and experimental data. They are utilized with computer algorithms that intensively search for favorable folded structures. Finally, coarse-grained models blend multiple atoms into a single sub unit or “bead” and define force field parameters that describe interactions of these beads. Coarse-grained studies have reduced resolution compared to all-atom studies but are also less computationally intensive. Because of this, they are more practical to study larger simulations involving multiple chains, such as aggregation. More details on the different levels of molecular models can be found in [31] and [32].

Peptoid modeling can be approached using a similar framework: using quantum methods to characterize monomer subunits, and building these into force fields that can describe longer chains, or systems of chains. Protein force fields may sometimes be used for peptoid simulation because the backbones of peptides and peptoids are composed from the same atoms. However, some key differences between peptoids and peptides can limit this transferability. Force fields for peptoids should be carefully verified and parameterized when appropriate.

Three key features of peptoid backbones can grant them local folding behavior significantly different from that of proteins: (1) peptoid backbones do not form hydrogen bonds, (2) their backbones are achiral, and (3) the peptoid sp^2 amide bond can isomerize and exist as both a *cis* or a *trans* conformer. In proteins, hydrogen bonding between nearby backbone amide and carboxyl groups stabilizes highly ordered, localized secondary structures, including α -helices and β -sheets. Because the peptoid side chain replaces the amide hydrogen, peptoids do not form the same secondary structures. Furthermore, in proteins, each C_α atom (with the exception of glycine) is a chiral center with (L) chirality, granting the rotational directionality to the backbone seen in the handedness of α -helices and the twists of β -sheets. Peptoid backbones resemble glycine, the only achiral amino acid. Therefore, unless chirality is added via a chiral side chain, peptoids do not favor a specific twist. Moreover, it is possible to build peptoids with side chains of alternating handedness, leading to unique secondary structures [15]. Finally, in proteins, the sp^2 amide bond is almost always in a *trans* configuration, where the carboxyl oxygen and amide hydrogen rest on opposite sides of the peptide bond, as the *cis* conformation leads to a steric clash between side chains. The peptoid amide bond, in contrast, can exist as a *cis* or a *trans* conformer, generally determined by the size of the side chain residue.

The differences mentioned above must be addressed when developing peptoid models. In the case of chirality and hydrogen bonding differences, this can be achieved by careful adaption of protein models. Protein force field parameters can be tuned to favor secondary structures not present in peptoids [33], so protein force fields transferred to peptoids should be verified and tuned against peptoid experimental data. Addressing the amide bond isomerization can be more challenging. The effect of a side chain on a peptoid secondary structure can differ depending on the isomerization state of the peptide bond. Due to this issue, all *ab initio* characterizations of side chain energetics should be examined in both the *cis* and the *trans* states. Ideally, parameters in all-atom and coarse-grained force fields should be tuned to accurately represent both configurations.

Properly sampling peptoid *cis/trans* isomerization is an unsolved challenge. Isomerization itself is a rare event that can have an activation barrier of 15 kcal/mol [34] and characteristic times slower than a microsecond [35]. The physical mechanism of isomerization involves breaking the peptoid double bond and rotating through a partial- sp^3 transition structure [36]. Force field parameters enforce this high energy barrier, and in molecular dynamics, a simulation method where Newton's equations of motion are integrated in time to explore the conformational space [31], the high-energy peptide bond isomerization is unlikely to occur at room temperature. Therefore, to make comparisons between *cis* and *trans* structures, this isomerization must be studied through enhanced sampling techniques that force the system across the barrier. These include umbrella sampling [37] and replica exchange molecular dynamics [38].

In this paper, we survey the literature on modeling studies of peptoids in light of the issues discussed above. We have organized this review as follows: In section 2.1, we describe insights into monomer-level backbone energetics found from *ab initio* studies. These include studies that describe differences between peptoid and peptide structures, and studies characterizing the effects of specific side chains on the peptoid backbone. In section 2.2, we discuss the predictive power of peptoid modeling using combined all-atom/semi-empirical and combined all-atom empirical techniques. Section 2.3 describes studies using all-atom simulations of peptoid chains. These studies have two goals: to characterize the transferability of protein force field models to peptoid systems and to identify peptoids that form local secondary structures. In section 2.4, we review a coarse-grained

model for peptoids. Section 2.5 describes how peptoid residues were added to the protein structure prediction program ROSETTA. Finally, we summarize the current state of peptoid modeling, and suggest future directions for computational studies of peptoids.

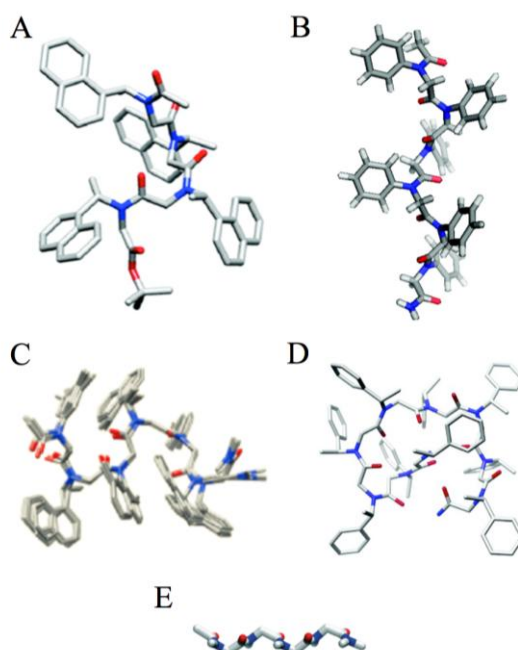


Figure 2. Known Peptoid Secondary Structures. (A) The PPI helix (shown here as (S)-N-(1-naphthylethyl)glycine) [11,39], (B) the N-(phenyl)glycine PPII helix [12], (C) the peptoid ribbon [14], (D) the threaded loop [13], and (E) the Σ -strand [15]. A, C, and D are shown without their backbone hydrogen atoms, and E is shown without its side chains. The secondary structures are characterized by their patterns of ϕ and ω dihedrals defined in Figure 3. The PPI and PPII helices have repeating ($\phi = (-)75^\circ$, $\omega \sim 0^\circ$) and ($\phi = (-)75^\circ$, $\omega = 180^\circ$) dihedrals, respectively. The threaded loop (D) is created by 9-mer NSPE peptoids and is stabilized by a hydrogen bonding interaction between side chains and their NH_2 groups. The ribbon (C) has alternating $\omega = 0^\circ$ and $\omega = 180^\circ$ dihedrals and the Σ -strand (E) has $\omega = 180^\circ$ and alternating $\phi = -90^\circ$, $\psi = 120^\circ$ and $\phi = -120^\circ$, $\psi = 90^\circ$ dihedrals. Structure A Reprinted with permission from Stringer et al. [39]. Structure B reprinted with permission from Shah et al. [12]. Structure C reprinted with permission from Crapster et al. [14]. Structure D Reprinted (adapted) with permission from Huang et al. [13]. Structure E reproduced with permission from Mannige et al. [15]. Copyright (A: 2011, B: 2008, D: 2006) American Chemical Society.

2. Modeling Studies of Peptoids

2.1. *Ab Initio* Studies of Peptoids

Ab initio simulations can be used to calculate highly accurate energy minimized structures for small molecules and have been used to characterize monomer-level structural preferences for peptoids. An advantage to this approach is that *ab initio* calculations are based wholly on quantum

theory and do not rely on system-specific force field parameters or tuning to experimental data. A disadvantage is that *ab initio* simulations are computationally expensive for all but the smallest molecules. Therefore, *ab initio* peptoid modeling has mostly been used to characterize short-range energy preferences of peptoid monomers and to refine structure predictions from all-atom simulations (see section 2.2). This section focuses on *ab initio* characterization of peptoid monomers, but also describes the *ab initio* characterization of the binding structure of a peptoid trimer.

Ab initio peptoid simulations frequently describe peptoid backbone energetics by calculating Ramachandran plots. Widely used to visualize protein secondary structure preferences, Ramachandran plots describe peptoid conformations in terms of the backbone dihedrals ϕ and ψ [40]. These dihedrals, a peptoid Ramachandran plot, and the locations of key peptoid structures are shown in Figure 3. Peptoid Ramachandran plots can look different from those of proteins due to structural differences between peptoids and peptides and due to plotting conventions. Because of their achiral backbones, peptoid Ramachandran plots can have energy minima with both positive and negative ϕ . Protein ϕ dihedrals are almost always in the -180° – 0° range because almost all amino-acids are (L) chiral. Additionally, peptoid Ramachandran plots should be plotted in pairs, with separate ϕ – ψ plots for the *cis* and *trans* conformers. A 3rd backbone dihedral, ω , describes the *cis/trans* isomerization. To compare *cis* ($\omega = 0^\circ$) and *trans* ($\omega = 180^\circ$) probabilities in the same figure, plots of ω versus ϕ or ω versus ψ , may be more appropriate (See Figure 4) [41]. By convention, peptoid Ramachandran plots may also be plotted on 0° – 360° axes rather than -180° – 180° axis so that the α_D minimum is centered.

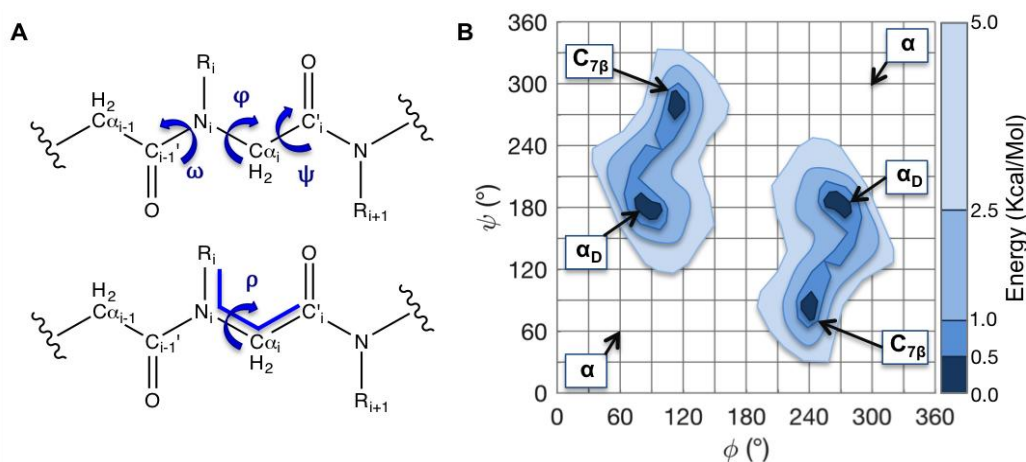


Figure 3. The peptoid backbone dihedrals and a peptoid Ramachandran plot. (A) Peptoid backbone conformations can be described by the dihedrals ϕ , ψ , and ω . The dihedral ω describes the isomerization of the planar peptoid bond. In the *cis* ($\omega = 0^\circ$) configuration, the side chain and the carbonyl oxygen are on the same side of the bond. In the *trans* ($\omega = 180^\circ$) configuration (shown) the side chain and the oxygen lie on opposite sides of the amide bond. The ρ dihedral, closely related to ϕ , is an additional parameter used to parameterize peptoid force fields, and is also shown. Ramachandran plots (B) are used to show backbone (ϕ – ψ) conformational preferences. Ramachandran plots are 2-D histograms that graph either the probability or the conformational energy of each ϕ/ψ pair of dihedrals. The figure shown corresponds to the *trans* isomer. This figure is based on data from Butterfoss [41].

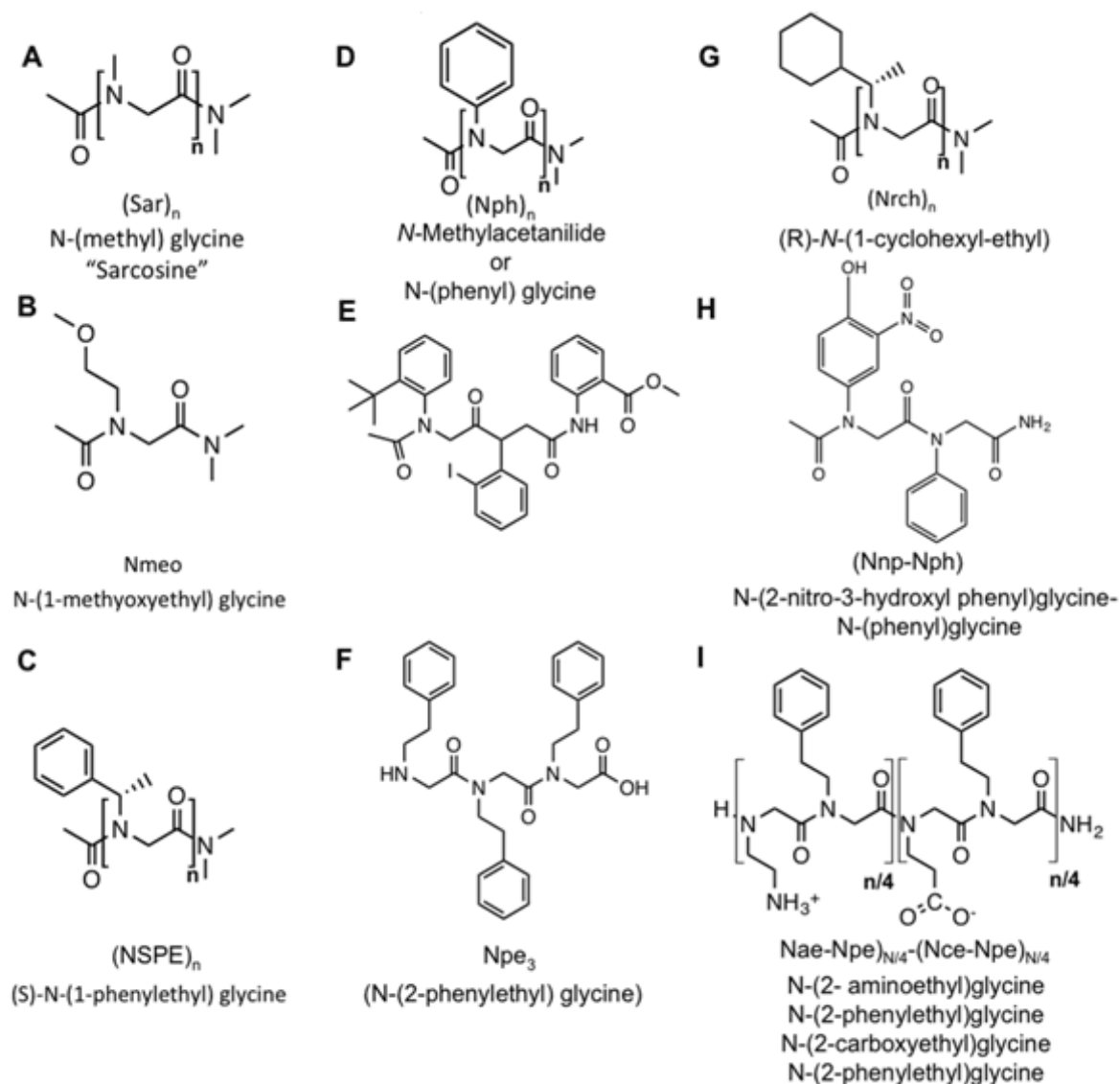


Figure 4. A partial summary of the peptoids described in this work. (A) Sarcosine (B) N-(1-methoxyethyl)glycine, (C) (S)-N-(1-phenylethyl)glycine, (D) N-Methylacetanilide or N-(phenyl)glycine, (E) The aryl trimer studied by Butterfoss et al., (F) (N-(2-phenylethyl)glycine)₃, (G) (R)-N-(1-cyclohexyl-ethyl), (H) N-(2-nitro-3-hydroxyl phenyl)glycine-N-(phenyl)glycine, (I) (N-(2-aminoethyl)glycine-N-(2-phenylethyl)glycine)_{N/4}-(N-(2-carboxyethyl)glycine-N-(2-phenylethyl)glycine).

Möhle and Hoffman [42] performed the first *ab initio* studies of a peptoid in 1996, comparing monomer conformations of the peptoid sarcosine (N-methylamide) to the small peptides glycine and alanine. Using HF-6-31G* and MP2-6-31G* calculations, they identified and named two vacuum-phase minimum energy conformations that are unique to peptoids, $C_{7\beta}$ ($\varphi \sim 130^\circ$, $\psi \sim 78^\circ$) and α_D ($\varphi \sim 75^\circ$, $\psi \sim 170^\circ$), and observed a local minimum at the α -helical (α) dihedrals ($\varphi \sim 55^\circ$, $\psi \sim 55^\circ$ and $\varphi \sim 305^\circ$, $\psi \sim 305^\circ$; $\omega \sim 180^\circ$) or ($\varphi \sim 55^\circ$, $\psi \sim 55^\circ$ and $\varphi \sim 305^\circ$, $\psi \sim 305^\circ$; $\omega \sim 0^\circ$). All three minima exist for both the *cis* and *trans* sarcosine conformers. Möhle and Hoffman found that $C_{7\beta}$ and α_D are the lowest energy conformations for *trans* sarcosine, and *cis* α_D is the global minimum in the

vacuum phase. Using a polarized continuum model [43,44] to represent solvation effects, they showed that the solvation stabilizes the α_D minimum in both *cis* conformations and that the *cis* and *trans* configurations are equally favorable in polar solvents.

Butterfoss et al. expanded on Möhle and Hoffman's work by exploring how longer backbones and more complex side chains affect the features of the peptoid energy surface [41]. Using the B3LYP//HF-6-31G* level of theory, they created Ramachandran plots for a sarcosine dimer and for monomers with N-(methoxyethyl) (Nmeo) and (S)-N-(1-phenylethyl) (NSPE) side chains. Chemical structures of these side chains are shown in Figure 4B and 4C. They overlaid these plots with experimental dihedral measurements from X-ray crystallographic and NMR studies of a variety of peptoid oligomers. To ensure that the Ramachandran energies were calculated with sufficient accuracy, Butterfoss et al. compared minimum energy surfaces calculated at the MP2-6-31G**//HF-6-31G* and B3LYP/6-31G**//HF/6-31G* levels of theory. They found that these scans were very similar, particularly around the energy minima, suggesting that B3LYP/6-31G* calculations are sufficient for calculating peptoid structure energies.

Butterfoss et al.'s results showed that peptoid side chains exert strong control over the peptoid backbone dihedrals. As in Möhle and Hoffman's work, Butterfoss et al.'s sarcosine Ramachandran plot featured $C_{7\beta}$ and α_D minima. These same minima were present for the NSPE and NMeo monomers, however, with the larger side chains, the α_D minimum was increasingly favored over $C_{7\beta}$. The chiral N side chain imparts a twisting structural preference on the peptoid backbone. This causes the Ramachandran plot to become asymmetric, as shown in Figure 5. In accordance with Möhle and Hoffman's solvated simulations, the experimental dihedrals summarized in Butterfoss et al. were largely localized to the α_D minima, confirming that it is the most common solvated conformation. In fact, the *trans* $C_{7\beta}$ minimum was only observed experimentally as part of a strained turn in a cyclic tetramer.

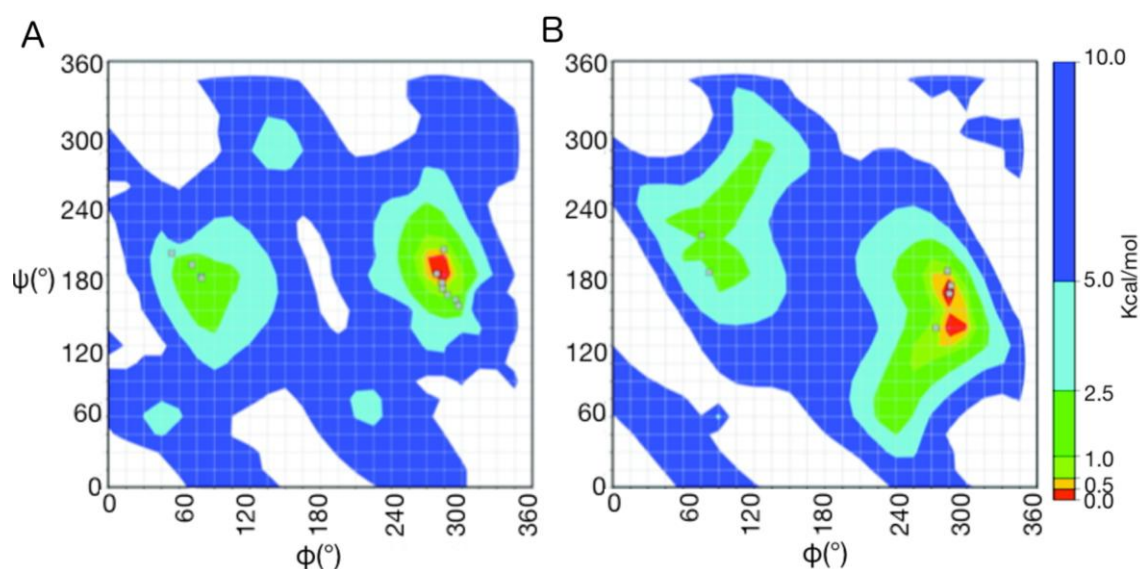


Figure 5. *Ab initio* Ramachandran plot for *cis* (A) and *trans* (B). Both plots represent data at the HF-6-31G* level of theory. The “o” and “+” markers represent cyclic and linear experimental data. Adapted with permission from Butterfoss et al. [41].

Both Butterfoss and Mühle found that the peptoid nitrogen is slightly pyramidal, causing the ω dihedral to deviate from 0° or 180° in these *cis* and *trans* configurations. Butterfoss and coworkers observed this aplanarity in both their experimental and computational results. This flexibility can be seen in traditional Ramachandran plots in the form of wider, flatter contours, or it can be plotted explicitly in ϕ - ω and ψ - ω plots as shown in Figure 6.

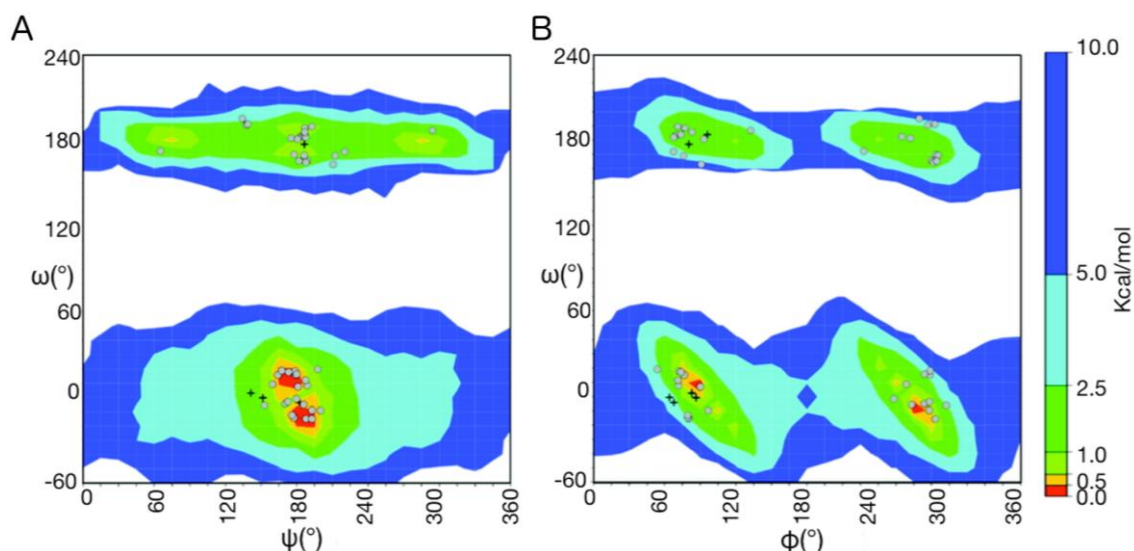


Figure 6. ψ - ω and ϕ - ω *ab initio* Ramachandran plots. ϕ - ψ Ramachandran plots are unable to describe isomerization of the ω dihedral. Butterfoss generated these (A) ψ - ω and (B) ϕ - ω plots as an alternative (*ab initio* HF-6-31G* data). ω dihedrals which deviate from 0° or 180° describe pyramidal nitrogen atoms. The “o” and “+” markers represent cyclic and linear experimental data. Adapted with permission from Butterfoss et al. [41], copyright (2009) American Chemical Society.

While most *ab initio* peptoid simulations have been used to characterize backbone energies, Parker et al. [45] used *ab initio* simulations for a different purpose: to understand the binding structures of a uranyl chelating peptoid. Parker et al. sought to identify peptoids that could selectively bind and extract uranyl (UO_2^{2-}) ions from seawater. Using a combinatorial screening approach, they identified peptoids with three consecutive carboxyl side chains as strong uranyl binders. They then used B3LYP/6-31G(d,g) calculations to find energy minimized structures, and the integral equation formalism for the polarizable continuum model (IEFPCM) to calculate solvation free energies [46]. Parker et al. identified several possible uranyl binding conformations for a carboxyl peptoid trimer. Two of them are shown in Figure 7. Carboxyl groups and water molecules can bind to uranyl around the linear ion’s equatorial plane. Parker et al. identified and minimized binding configurations where the trimer binds uranyl through 1, 2, or 3 carboxyl groups. They suggested that the binding free energy of these complexes results from a tradeoff between binding enthalpy of each group and lost chain entropy due to conformational restriction.

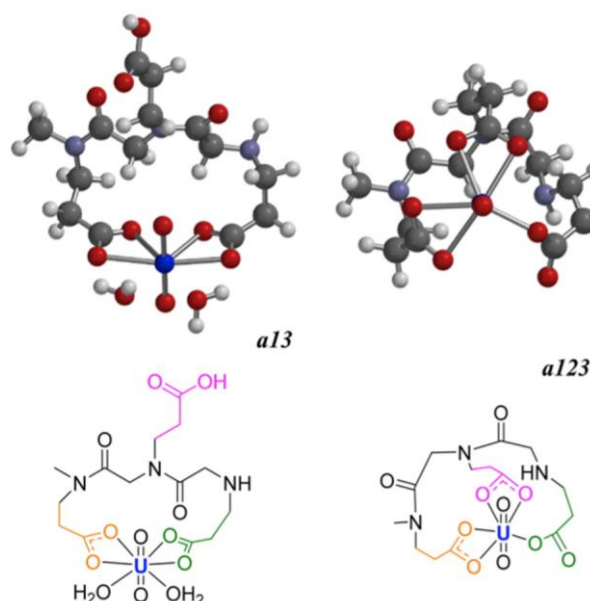


Figure 7. Two possible binding configurations for a uranyl chelating peptoid. The peptoid can bind the uranyl ion with one, two, or three carboxyl groups. The binding enthalpy released on additional carboxyl binding is counteracted by the entropy loss in the restricted backbone chain. Adapted with permission from Parker et al. [45]. Copyright (2016) American Chemical Society.

2.2. Peptoid Structure Prediction with Combined Methods

Attempts to study longer peptoid chains with available models pose a problem: *ab initio* methods are computationally expensive for all but the smallest molecules, while existing all-atom force fields may not adequately model peptoid energetics. A productive solution has been to conduct simulation studies combining different levels of theory. In this approach, existing all-atom force fields are used to explore many possible conformations and identify potential minima, then semi-empirical or *ab initio* methods are used to more precisely characterize these minimum energy structures. This kind of approach has played a role in the prediction and discovery of novel peptoid secondary structures, including the (S)-N-(1-phenylethyl)glycine (NSPE)-induced polyproline I helix (PPI) [11] and the N-(phenyl)-glycine-induced polyproline II helix (PPII) [12], and in the NMR characterization of the peptoid ribbon [14].

Armand et al. [11] predicted the structure of the NSPE PPI helix in 1997, using the AMBER all-atom force field [47], and the AMSOL [48] semi-empirical force field with the AM1 parameter set. Previous experimental studies of peptoids with repeating NSPE side chains had resulted in Circular Dichroism (CD) spectra consistent with helical secondary structures [24]. With this knowledge, Armand et al. performed conformational searches of NSPE octamers by varying each of the ϕ and ψ dihedrals over a range of 0–360°. They calculated conformational energies using the torsional and nonbonded parameters of the AMBER force field, holding the molecule's angles and bond lengths fixed. To refine their structure characterization, they used AMSOL to minimize NSPE dimers.

Armand et al. showed that (NSPE)₈ favored a PPI structure with *cis* ($\omega = 0^\circ$) peptoid bonds and a negative ($-120^\circ < \phi < -75^\circ$) twist, and identified the steric interactions that stabilize this

conformation. In particular, they suggested that steric clashes between the bulky side chain and the peptoid backbone enforce a preference for the *cis* isomer, and that the chiral center orients to minimize steric clashes between its bulky benzene and methyl groups and the neighboring carbonyl oxygen. Through these interactions, the chiral side chain stabilizes a negative ϕ for NSPE but positive ϕ for its isomer (R)-N-(1-phenylethyl)glycine (NRPE).

In a later paper, Shah et al. used a combination of all-atom and semi-empirical studies to show that peptoids with simple benzene side chains (Figure 4D) form PPI helices (Figure 2) [12]. Experimental studies show that phenyl side chains promote the *trans* isomer of peptoid monomers [49]. Shah et al. were interested in other possible structural preferences imparted by this side chain. They generated an *ab initio* Ramachandran plot at the B3LYP/6-311+G(2d,p)//HF/6-31G* level of theory for an N-methylacetanilide dimer in the *trans* configuration. This scan identified minima in the ϕ - ψ energy surface at ($\phi = 60^\circ$, $\psi = -150^\circ$) ($+\phi$) and ($\phi = -60^\circ$, $\psi = 150^\circ$) ($-\phi$). To understand if these local energy preferences resulted in long-range order, Shah studied N-methylacetanilide hexamers whose units had different combinations of these minima. The structures of these hexamers were then optimized using the Sigma simulation package [50] with the cedar all-atom potential [51]. These calculations showed that steric clashes prohibited sequences with mixtures of residues in the $+\phi$ and $-\phi$ conformations. The minimum energy structures were helices with repeating $-\phi$ or $+\phi$ turns.

In a similar study, Butterfoss et al. attempted blind structure prediction for three peptoid oligomers [52]. Each oligomer, (E, F, and C in Figure 4) contained side chains that stabilized its crystals via different forces: (E) a trimer with three bulky, ortho-substituted benzene side chains that stabilize its structure via side chain steric effects, (F) (Npe)₃, a trimer with three N-2-phenylethyl side chains whose crystals are stabilized by π - π interactions between benzene rings in neighboring molecules and by end group effects, (C) a cyclic (NSPE)₉ molecule that forms needle like crystals whose side chains encage a single ethanol molecule. Prior to simulations, newly solved crystal structures for these chains were hidden from the modelers. Butterfoss et al. ran all-atom simulations, using parameters from the generalized amber force field (GAFF) [53] and represented solvent effects using the Generalized Born implicit solvent model [54]. To overcome *cis/trans* barriers and allow sampling of all ω conformations, they used replica exchange molecular dynamics (REMD) [38]. For the trimers, quantum structure calculations were performed, independent of the REMD, at the HF/3-21G* level with energies calculated at the M052X/6-311G** level. For the nonamer characterization, REMD and low-level quantum mechanics calculations were performed sequentially, using quantum mechanics to optimize the minimum energy structures identified by REMD.

Butterfoss et al. showed that the success of computational structure predictions with GAFF depends heavily on the particular forces that stabilize a peptoid structure. Both REMD and QM successfully predicted the crystal structure of the aryl trimer but identified a highly degenerate set of low energy conformations for (Npe)₃. This is likely because the crystal structure of (E) is stabilized by its own side chains, while that of (Npe)₃ is stabilized by interchain interactions. The cyclic (NSPE)₉ simulations predicted the correct *cis/trans* patterning of the backbone ω dihedrals but failed to predict trends in the ϕ dihedrals and the orientations of the side chain dihedrals. Butterfoss re-ran a DFT calculation of NSPE including an explicit ethanol molecule, and showed that this predicted the correct structure. This suggests that, in this case, and other cases where complexation between solvent molecules and the peptoid side chains may be possible, the solvent must be explicitly included in the structure calculation.

2.3. All-atom Molecular Dynamics Studies of Peptoids

Quantum studies can precisely describe the properties of very small peptoids, but all-atom molecular dynamics (MD) studies are necessary to characterize longer peptoid chains. The quantum studies described previously were able to use monomer level energy preferences to predict structures for peptoids composed of the same repeated side chain. The next step in rational peptoid design is to predict how sequences composed of differing side chains combine to lead to folded structures or nanomaterials. All-atom simulations can build upon quantum work by making such predictions.

All-atom force fields extend the range of size and length scales accessible to simulation by using effective interaction potentials between the atoms in the system. The parameters in these effective potentials can be fitted to experimental data and also to reproduce the results of quantum calculations, thus bridging the gap between the electronic and the atom-wise description of the system. The literature on force field development is extensive and methods vary depending on the force field, for an introduction to the subject the reader is referred to force field specific references [53,55,56,57].

The quality of all-atom simulation results is dictated by the quality of the all-atom force field used. Most peptoid MD simulations have been performed with either protein-specific force fields such as those of CHARMM [55] or AMBER [56], or generalized force fields such as OPLS [57] or GAFF [53]. A priori, both of these classes of force fields should not be expected to represent peptoid structures accurately, as protein force fields may result in an unrealistic preference for protein-specific secondary structures, such as α -helices, and general force fields may need extra tuning to correctly predict peptoid-specific structures. More recently, newer, peptoid-specific force fields have been developed to address these concerns.

Möhle and Hoffman ran the first all-atom simulations of sarcosine monomers, comparing these MD calculations against their *ab initio* results described in section 2.1 [58]. Using the CHARMM22 force field, they simulated sarcosine molecules both in vacuum and explicitly solvated with TIP3P water [59], and used holonomic constraints [60] to calculate relative free energies for the sarcosine minimum energy structures. In these simulations, CHARMM22 replicated the *ab initio trans* vacuum phase minima α_D and $C_{7\beta}$ but overstabilized *cis* α . In spite of this, both *cis* and *trans* simulations in water replicated the *ab initio* results. The solvated simulations showed that both *trans* and *cis* conformers preferred the α_D minimum when solvated in explicit water.

Voelz et al. sought to determine whether the GAFF all-atom force field [53] could simulate the dihedral preferences of small peptoids. Their work targeted the *ab initio* and experimental data published by Butterfoss et al. [41] and described in section 2.1. They parameterized GAFF for peptoids using Antechamber automatic typing [61] and AM1-BCC charge fitting [62]. To ensure that both *cis* and *trans* configurations of the amide bond were explored, they used REMD to overcome high energy barriers [38]. GAFF's ability to recreate the *ab initio* peptoid dihedrals was compared to other AMBER force fields and to the OPLS force field in implicit and explicit solvent. GAFF simulations using the generalized Born surface area (GBSA) solvent model [54] predicted peptoid dihedrals better than other force fields and better than explicitly solvated simulations. GAFF recreated most of the *ab initio* dihedrals well but failed to reproduce the $-\phi$ preference observed for NSPE. GAFF was able to successfully reproduce the *cis/trans* patterning of the peptoid oligomers (Sar)₈, a (R)-N-(1-cyclohexyl-ethyl)glycine pentamer (Nrch)₅, and an N-(2-nitro-3-hydroxyl phenyl)glycine-N-(phenyl)glycine (Nnp-Nph) dimer (Figure 4A, G, and H).

Following up on Voelz's study, Mukherjee et al. developed a tuned force field, called GAFF- ϕ , to improve GAFF's sampling of the NSPE peptoid helix [63]. To create GAFF- ϕ , Mukherjee et al. added a single additional ϕ dihedral parameter to GAFF. This was sufficient to fix the ϕ distributions for NSPE residues. REMD simulations of chains of 3 to 15 NSPE residues showed that, in addition to improving the ϕ sampling, the ratio of *cis* and *trans* ω orientations was closer to experiment for GAFF- ϕ than for unfitted GAFF. This suggests that peptoid ϕ and ω energies are correlated.

Using GAFF- ϕ , Mukherjee et al. characterized the cooperative folding of the NSPE helix. To do this, they adapted the Lifson-Roig helix-coil theory [64] to describe peptoid helix formation. Briefly, the Lifson-Roig theory for proteins categorizes a chain's residues, based on their dihedrals, as either α -helical helical or non-helical coiled. A statistical weight for nucleation or elongation is then assigned to each helical residue. Which type of weight is assigned depends on whether one or both of the residue's neighbors are also helical. The combination of residue weights can then be used to calculate a chain's thermodynamic properties including the enthalpy of helix formation. Mukherjee sought to study the formation of PPI helices by NSPE peptoids, and defined helical conformations using the PPI dihedrals ($|\omega| < 90^\circ$ with the $-\phi$ conformation defined in section 2.2). They then fit LR nucleation and elongation weights to their GAFF- ϕ simulation data and used these weights to calculate the enthalpies and entropies of helix formation. The results showed that helix formation of 1-NSPE is entropically favorable, likely driven by steric interactions between the bulky side chains.

More recently, Mirijanian et al. [65] developed the MFTOID (short for Molecular Foundry Peptoid) force field, a CHARMM-based all-atom force field parameterized specifically for peptoids. CHARMM [55] is an additive force field, whose energy function includes contributions from charges, van der Waals forces, and bonded terms including bonds, angles, dihedrals, and improper dihedrals. Mirijanian transferred most MFTOID parameters from the CHARMM22 force field by assigning peptoid atoms parameters from similar atoms in the peptide backbone. The Lennard-Jones, bond, and angle parameters for the peptoid nitrogen were transferred from the secondary peptide backbone nitrogen. Mirijanian et al. fit the charges of the amide nitrogen and carboxyl carbon, as well as the Lennard-Jones parameters of the carboxyl carbon. They also refit the ω dihedral potential to allow *cis/trans* isomerization of the amide bond, and the ρ dihedral potential, which is closely related to ϕ (see Figure 3A). All other bonded parameters, including the ϕ and ψ dihedral potential terms were left unchanged.

To validate MFTOID, Mirijanian et al. carried out two different types of simulations. First, they used umbrella sampling [37] to create solvated and unsolvated free energy Ramachandran maps of sarcosine dipeptoids. These maps showed qualitative agreement with previous studies of sarcosine: the vacuum phase minima included α_D , $C_{7\beta}$ and α -helix, and the *trans* α_D configuration was stable in solvent. Unfortunately MFTOID over-stabilizes the *cis* α -helical conformation, the same effect observed by Mühle et al. [42] in their vacuum phase, CHARMM22 simulations of sarcosine. In MFTOID, the α conformation is also a global minimum in solvent. Mirijanian et al. also ran molecular dynamics simulations of N-2-phenylethyl trimer crystals. These simulations used experimental crystal structures composed of 256 trimers as starting configurations. They showed that after a 10 ns simulation at 300 K, MFTOID reproduces the starting crystal structure with a Root Mean Square Deviation (RMSD) of approximately 1.0 Å.

MFTOID parameters were tuned exclusively to sarcosine, and the transferability of the parameters to other peptoids is a potential issue. Prior work suggests that transferability should be high: Butterfoss's quantum results show that side chains bound to the backbone through a methylene

group have similar dihedral minima to sarcosine. Additionally, Voelz's work with GAFF showed that many sterically dominated peptoid structures may not require precisely-tuned parameters to reproduce the correct minimum energy structures. However, peptoids with certain side chains, for example N-alkoxy-glycines [66] whose side chains attach to the backbone through a nitrogen-oxygen bond, exhibit different backbone behavior than sarcosine, and will likely need to be reparameterized.

Mannige et al. [15] used MFTOID in combination with experimental data to explore the structure of bilayered peptoid microsheets. The work focused on microsheets previously described by Nam et al. [67] composed of diblock peptoid oligomers of alternating charged and hydrophobic side chains of the form $(\text{Nae-Npe})_{n/4}-(\text{Nce-Npe})_{n/4}$ (Figure 4I). In water, these peptoids form large sheets, two chains thick, with the hydrophobic phenyl-ethyl side chains sequestered inside. MFTOID quantitatively and qualitatively reproduced the features of these sheets. It reproduced X-ray and NMR data describing sheet thickness and strand separation to an RMSD within 1 Å.

In the same work, Mannige et al. recognized a novel peptoid secondary structure defined by residues with an alternating pattern of dihedrals. They dubbed this structure a “ Σ -strand” and suggested that it was the first of a class of possible peptoid secondary structures described by the alternating conformations of two residues rather than a single repeating dihedral motif. A Ramachandran scheme for this type of secondary structure is shown in Figure 8.

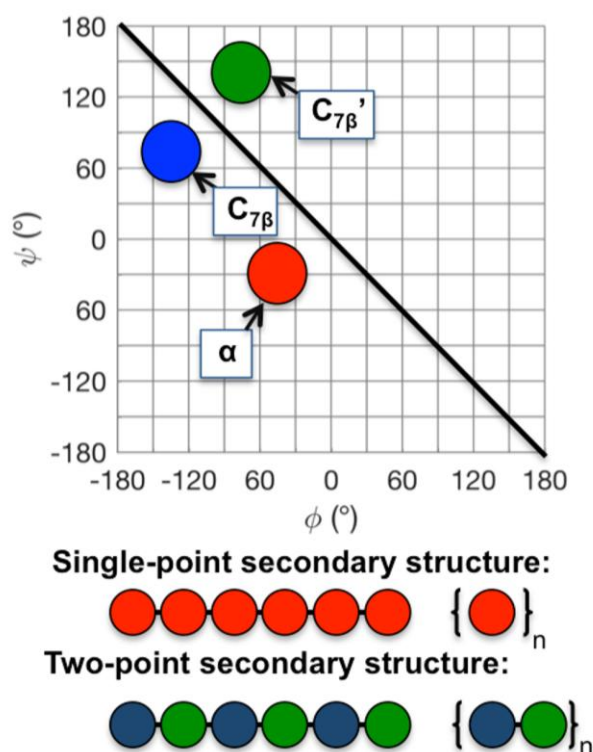


Figure 8. The Σ -strand alternating motif. Mannige et al. [15] described a secondary structure of alternating residues with alternating rotational states. Single-point secondary structures, like α -helices, contain a single repeating dihedral conformation. Two-point or Σ -strand motifs have alternating dihedral pairs, in this example they alternate between the $C_{7\beta}$ and $C_{7\beta}'$ conformations. This figure is based on a scheme by Mannige [15].

In subsequent work, Mannige, Kundu, and Whitlam derived a new order parameter called the “Ramachandran Number” (R) to describe alternating structures such as Σ -strands [68]. Using a thoughtfully chosen gridding of the Ramachandran space, they derived an expression for R that combines information about the ϕ and ψ dihedrals and the radius of r_g of the structure. The number is obtained by defining a grid of values of ϕ and ψ in the Ramachandran plot, which is rotated, shifted and rescaled so that the important secondary structures are mapped onto unique integer numbers. The Ramachandran Number can be plotted as a histogram “barcode” or as a per residue graph allowing visualization of patterned peptoid structures including Σ -strands. This is shown in Figure 9.

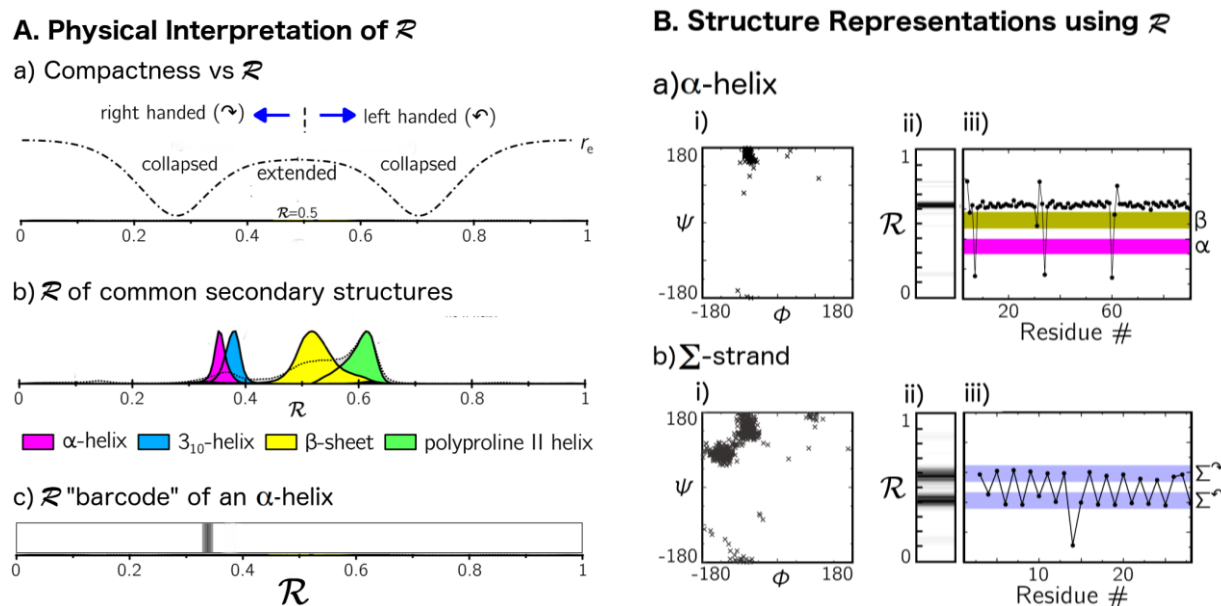


Figure 9. The Ramachandran Number. The Ramachandran Number R is a parameter derived by Mannige et al. which combines a residue’s ϕ and ψ dihedrals and chain compactness into a single number. (A) (a) shows the physical information encoded in R . $R > 0.5$ describes a left handed twist, $R < 0.5$ describes a right handed twist. $R = 0$, $R = 0.5$, and $R = 1.0$ describe fully extended backbones, while intermediate values are more compact. (b) The distribution of R values for common peptide secondary structures. (c) A Ramachandran number histogram or “barcode” representation of the α -helix, (the distribution in (b) shown as a heat map). (B) shows different structure visualizations for an α -helix (a) and for the alternating Σ -strand (b). (i) A Ramachandran plot, (ii) a barcode of the Ramachandran number, (iii) An R versus residue # plot, which, unlike the histograms, can show the alternating pattern of dihedrals in the Σ -strand. This figure is reproduced from figures in [68].

2.4. Coarse-grained Peptoid Simulations

All-atom simulations are useful for studying the conformational energies of individual peptoid chains, however, they are too computationally demanding for studying slow processes involving the interactions of many chains. All-atom simulations are best for simulations up to a few microseconds, and up to a few million atoms. Simulations of self-assembling peptoid nanomaterials, such as peptoid

sheets, require force fields usable across millisecond time scales or longer for many hundreds of peptoid chains. One way this type of simulation can be achieved is by using coarse-grained models that incorporate many peptoid atoms into a single bead.

Coarse-grained models extend the range of length and time scales that can be simulated beyond the limits of all-atom models by defining larger entities (“beads”) that encompass groups of atoms, as the basic building blocks of the system. The interactions between these effective building blocks are then fitted to reproduce the results of all-atom simulations of smaller cognate systems. Some of the methods used to develop these models include Boltzmann Inversion [69], which builds effective potentials by fitting to the all-atom pair distribution functions, Force Matching [70], which fits the effective potential to the net forces acting between the building blocks in the all-atom simulations, and Relative Entropy Minimization [71], which seeks to minimize the information loss in replacing the atomistic description with the coarse grained one. A detailed description of these approaches is beyond the scope of this paper, for a recent exhaustive review the reader is referred to [32]. It is also possible to build coarse-grained models by directly fitting to experimental data in what is called “top-down” coarse-graining, but such approaches have not, to our knowledge, been used for peptoids.

Haxton et al. [72] developed the first coarse-grained model for peptoids and named it MF-CG-TOID (short for Molecular Foundry-Coarse-Grained-Peptoid). They developed this model specifically to study $(\text{Nca-Npe})_{n/4}$ - $(\text{Nae-Npe})_{n/4}$ self assembly into bilayered sheets. The $(\text{Nca-Npe})_{n/4}$ - $(\text{Nae-Nce})_{n/4}$ polymers are described in the review of Mannige’s work above. Haxton et al. deliberately avoided a coarse-graining scheme based on beads with radially-symmetric Lennard-Jones potentials and their bonded interactions, arguing that this way of partitioning the model’s potential energy is artificial and non-ideal. Instead, each MF-CG-TOID bead was represented by a location (\vec{r}) and an orientation parameter (\hat{x}). Figure 10 contains a scheme showing \vec{r} and \hat{x} parameter for a pair of peptoid side chains.

Haxton et al. parameterized 4 coarse-grained beads: one bead representing a backbone unit, and one for each of the 3 side chains in the $(\text{Nca-Npe})_{n/4}$ - $(\text{Nae-Npe})_{n/4}$ system. Most of the force field interaction potentials are functions of these parameters: backbone-backbone and backbone-side chain bonded parameters, and nonbonded interactions between like-side chains. Nonbonded interactions between unlike-beads and between backbone-beads were represented by hard-sphere potentials, which depend only on a size parameter. Parameterization involved fitting 115 parameters through direct Boltzmann Inversion of MFTOID and published all-atom data, with 4 parameters fit to the target system (phenylethyl side chain size, interaction strength and bonded interaction with the backbone, and the backbone bead size). Haxton et al. performed Monte Carlo simulations using their force field to model structures in different stages of $(\text{Nae-Npe})_{n/4}$ - $(\text{Nce-Npe})_{n/4}$ bilayered sheet formation [73]. MF-CG-TOID qualitatively represented the globular structure of individual peptoids in water, and their unfolding when exposed to an air surface. Using the model, they were able to match X-Ray spectroscopy data, such as chain spacing for single layer sheets at the air-water interface, and free floating peptoid bilayers. MF-CG-TOID has been used for an additional study of $(\text{Nae-Npe})_{n/4}$ - $(\text{Nce-Npe})_{n/4}$ sheet formation [74].

MF-CG-TOID reproduces $(\text{Nca-Npe})_{n/4}$ - $(\text{Nae-Npe})_{n/4}$ sheets with excellent precision, but using it to simulate other peptoid side chains may be challenging, particularly due to the number of parameters that must be fit for each new side chain. Its direct Boltzmann inversion scheme requires an accurate all-atom peptoid model. Therefore, adding novel peptoids to MF-CG-TOID would

potentially require two parameterization steps: one adding the new side chain to MFTOID, and a second parameterization using the Boltzmann inversion procedure.

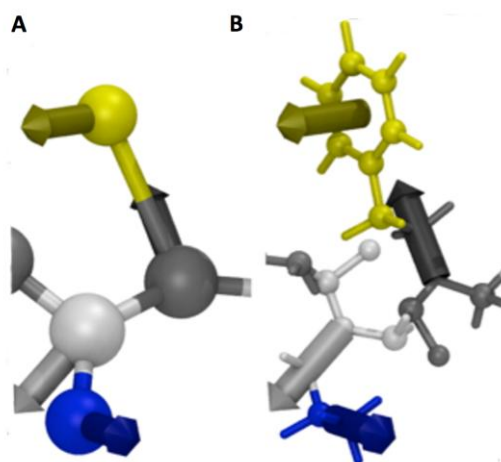


Figure 10. MF-CG-TOID coarse-grained peptoid model. (A) The MF-CG-TOID coarse-grained model for an (Nae)–(Npe) unit. Each peptoid monomer is represented by two coarse-grained sites. Each site is described by two degrees of freedom: a position \vec{r} represented in the diagram by spheres and a director \hat{x} represented by the arrows. All bonded and nonbonded interaction parameters in MF-CG-TOID are functions of both parameters ($X(\vec{r}_i, \hat{x}_i)$). (B) The all atom structure upon which this model was based. Figure reproduced with permission from Haxton et al. [72]. Copyright (2014) American Chemical Society.

2.5. Peptoid Simulations in Rosetta

For researchers looking to study peptoids with existing protein structure prediction software, Drew et al. added peptoids to the Rosetta framework [75].

Rosetta [76] is a molecular modeling and protein design utility with broad functionality, including algorithms for *de novo* protein structure prediction. Rosetta’s folding algorithm differs from the all-atom and coarse-grained approaches described above, which use physics-based models and long simulations (or enhanced-sampling methods) to identify minimum energy configurations. Instead, Rosetta identifies thousands of possible folded structures through a conformational search algorithm based on Metropolis Monte Carlo, first sampling large backbone (ϕ – ψ) motions and then refining the local side chain structure based on a pre-calculated “rotamer library” of allowable side chain (χ) dihedral conformations. These structures are then minimized using physics-based molecular mechanics models. Candidate structures are those with the lowest minimized energies.

To incorporate peptoids into ROSETTA, functions for use in the Monte Carlo search and energy parameters for the molecular mechanics minimization are needed. Drew defined a peptoid backbone unit to have the same backbone atoms as a peptide subunit, determined molecular mechanics parameters for over 50 side chains, and modified Rosetta’s conformational search algorithms to allow sampling of backbone *cis/trans* isomerization and nitrogen flexibility. Peptoid molecular mechanics parameters were taken from CHARMM. A CHARMM proline nitrogen atom

was used to represent the peptoid backbone nitrogen, and dihedral energy parameters were tuned to quantum data. To enable conformational search of both *cis* and *trans* isomers, Drew et al. modified the rotamer creation function to sample both isomers. They then added two types of Monte Carlo moves to the search algorithm: one sampling *cis/trans* flips and one sampling small backbone nitrogen pyramidalization by making small adjustments to the ω dihedral.

In subsequent work, Renfrew et al. [77] developed a peptoid rotamer library for use with Drew's framework and validated Rosetta-derived peptoid structure predictions against known experimental structures. Rosetta's peptide rotamer libraries include sterically allowed configurations for the side chain χ_1 and χ_2 dihedrals (shown in Figure 11). For peptides, these libraries are defined with respect to (ϕ, ψ) conformations and each (χ_1, χ_2) rotamer is weighted with a probability derived from data from the Protein Data Bank. Renfrew showed that peptoid rotamer libraries should be defined with respect to (ϕ, ψ, ω) combinations, to account for peptoid nitrogen pyramidalization. They also compensated for the lack of peptoid experimental data by deriving rotamer weights based on the Rosetta molecular mechanics force field and quantum mechanical calculations. Renfrew et al. validated their model against experimental data for the common peptoid side chains Nph, Nspe, Nmeo, and N-(S)-1-Naphthylethyl (Ns1e), and showed that the rotamer library was able to represent the preferred rotamers of the peptoid within a RMSD of 0.5 Å.

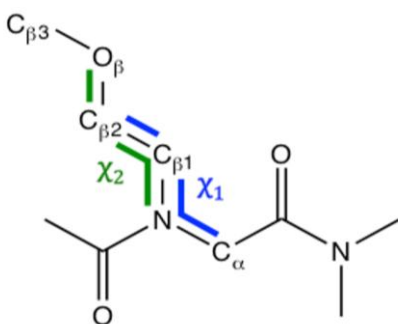


Figure 11. The rotamer dihedrals. Rosetta rotamer libraries describe the side chain conformation in terms of the dihedrals χ_1 and χ_2 . Renfrew developed rotamer libraries for 54 peptoid side chains.

The peptoid residues included in Rosetta can be used to predict peptoid structures or structures of polymers combining peptoid and peptide residues.

3. Conclusions and Future Directions

In recent years, computational studies of peptoids have multiplied and progress has been made towards predicting peptoid conformations based on the sequence of side chains attached to their backbones. Approaches to predict folded peptoid structures depend on the nature of the side chains and the length of the chain under study. From quantum chemistry, we can obtain fundamental understanding of the peptoid backbone and the influence of individual side chains on its conformation, but longer chains cannot be studied due to computational constraints. All-atom models can partially overcome these limitations and predict the folding behavior of several short peptoid chains, however they are limited by the quality of the all-atom force field used. Existing protein and

general force fields have been shown to adequately predict the backbone dihedrals of several solved crystal structures, and new peptoid-specific force fields have been developed in the last few years. However, peptoid force fields are only parameterized for a limited set of side chains. Finally, coarse-grained models have shown promise in the study of $(\text{Nae-Npe})_{n/4}$ – $(\text{Nce-Npe})_{n/4}$ bilayer sheet formation.

While peptoid modeling studies have made much progress in the last 20 years, there is still a need to develop and improve peptoid models. This will require expanding the number of side chains parameterized for all-atom and coarse-grained peptoid force fields and, at the same time, expanding the number of experimentally solved structures available to tune and verify these force fields. Additionally, it will require developing simulation techniques that encourage peptoids to explore both the *cis* and the *trans* peptide bond conformations. This may involve use of enhanced sampling techniques such as REMD, umbrella sampling, and metadynamics [78]. Nevertheless, in recent years there has been a big improvement in the simulation models available for peptoids, and we expect to see much better tools to predict peptoid structure and properties in the near future.

Conflict of Interest

All authors declare no conflict of interest in this paper.

References

1. Sun J, Zuckermann RN (2013) Peptoid Polymers: A Highly Designable Bioinspired Material. *ACS Nano* 7: 4715–4732.
2. Seo J, Lee BC, Zuckermann RN (2011) Peptoids: Synthesis, Characterization, and Nanostructures. *Compr Biomater* 2: 53–76.
3. Chongsiriwatana NP, Patch JA, Czyzewski AM, et al. (2008) Peptoids that mimic the structure, function, and mechanism of helical antimicrobial peptides. *Proc Natl Acad Sci USA* 105: 2794–2799.
4. Vollrath SBL, Fürniss D, Schepers U, et al. (2013) Amphiphilic peptoid transporters—synthesis and evaluation. *Org Biomol Chem* 11: 8197–8201.
5. Li N, Zhu F, Gao F, et al. (2010) Blockade of CD28 by a synthetic peptoid inhibits T-cell proliferation and attenuates graft-versus-host disease. *Cell Mol Immunol* 7: 133–142.
6. Dohm MT, Kapoor R, Barron AE (2011) Peptoids: Bio-Inspired Polymers as Potential Pharmaceuticals. *Curr Pharm Design* 17: 2732–2747.
7. Statz AR, Park JP, Chongsiriwatana NP, et al. (2008) Surface-immobilised antimicrobial peptoids. *Biofouling* 24: 439–448.
8. Seuryneck SL, Patch JA, Barron AE (2005) Simple, helical peptoid analogs of lung surfactant protein B. *Chem Biol* 12: 77–88.
9. Maayan G, Ward MD, Kirshenbaum K (2009) Folded biomimetic oligomers for enantioselective catalysis. *Proc Natl Acad Sci USA* 106: 13679–13684.
10. Gellman SH (1998) Foldamers: A Manifesto. *Accounts Chem Res* 31: 173–180.
11. Armand P, Kirshenbaum K, Falicov A, et al. (1997) Chiral N-substituted glycines can form stable helical conformations. *Fold Design* 2: 369–375.

12. Shah NH, Butterfoss GL, Nguyen K (2008) Oligo(N-aryl glycines): A New Twist on Structured Peptoids. *J Am Chem Soc* 130: 16622–16632.
13. Huang K, Wu CW, Sanborn TJ, et al. (2006) A threaded loop conformation adopted by a family of peptoid nonamers. *J Am Chem Soc* 128: 1733–1738.
14. Crapster JA, Guzei IA, Blackwell HE (2013) A Peptoid Ribbon Secondary Structure. *Angew Chem Int Ed* 52: 5079–5084.
15. Mannige RV, Haxton TK, Proulx C, et al. (2015) Peptoid nanosheets exhibit a new secondary-structure motif. *Nature* 526: 415–420.
16. Hebert ML, Shah DS, Blake P, et al. (2013) Tunable peptoid microspheres: effects of side chain chemistry and sequence. *Org Biomol Chem* 11: 4459–4464.
17. Murnen HK, Rosales AM, Jaworski JN, et al. (2010) Hierarchical Self-Assembly of a Biomimetic Diblock Copolypeptoid into Homochiral Superhelices. *J Am Chem Soc* 132: 16112–16119.
18. Sanii B, Kudirka R, Cho A, et al. (2011) Shaken, Not Stirred: Collapsing a Peptoid Monolayer To Produce Free-Floating, Stable Nanosheets. *J Am Chem Soc* 133: 20808–20815.
19. Dill KA, MacCallum JL (2012) The Protein-Folding Problem, 50 Years On. *Science* 338: 1042–1046.
20. Gorske BC, Blackwell HE (2006) Tuning peptoid secondary structure with pentafluoroaromatic functionality: A new design paradigm for the construction of discretely folded peptoid structures. *J Am Chem Soc* 128: 14378–14387.
21. Stringer JR, Crapster JA, Guzei IA, et al. (2010) Construction of Peptoids with All Trans-Amide Backbones and Peptoid Reverse Turns via the Tactical Incorporation of N-Aryl Side Chains Capable of Hydrogen Bonding. *J Org Chem* 75: 6068–6078.
22. Gorske BC, Nelson RC, Bowden ZS, et al. (2013) “Bridged” $n \rightarrow \pi^*$ Interactions Can Stabilize Peptoid Helices. *J Org Chem* 78: 11172–11183.
23. Wu CW, Kirshenbaum K, Sanborn TJ, et al. (2003) Structural and spectroscopic studies of peptoid oligomers with alpha-chiral aliphatic side chains. *J Am Chem Soc* 125: 13525–13530.
24. Kirshenbaum K, Barron AE, Goldsmith RA, et al. (1998) Sequence-specific polypeptoids: A diverse family of heteropolymers with stable secondary structure. *Proc Natl Acad Sci USA* 95: 4303–4308.
25. Armand P, Kirshenbaum K, Goldsmith RA, et al. (1998) NMR determination of the major solution conformation of a peptoid pentamer with chiral side chains. *Proc Natl Acad Sci USA* 95: 4309–4314.
26. Dill KA (1990) Dominant forces in protein folding. *Biochemistry* 29: 31.
27. Dill KA, Ozkan SB, Shell MS, et al. (2008) The protein folding problem. *Annu Rev Biophys* 37: 289–316.
28. Sali A, Blundell TL (1993) Comparative protein modeling by satisfaction of spatial restraints. *J Mol Biol* 234: 779–815.
29. Shen MY, Sali A (2006) Statistical potential for assessment and prediction of protein structures. *Protein Sci* 15: 2507–2524.
30. Rohl CA, Strauss CEM, Misura KMS, et al. (2004) Protein structure prediction using Rosetta. *Method Enzymol* 383: 66–93.
31. Leach AR (2001) *Molecular modelling: principles and applications*, England: Pearson/Prentice Hall.

32. Shell MS (2016) Coarse-Graining with the Relative Entropy, In: Rice SA, Dinner AR, *Advances in Chemical Physics*, Malden: Wiley-Blackwell, 395–441.
33. Mackerell AD, Feig M, Brooks CL (2004) Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J Comput Chem* 25: 1400–1415.
34. Feigel M (1983) Rotation barriers of amides in the gas phase. *J Phys Chem* 87: 3054–3058.
35. Sui Q, Borchardt D, Rabenstein DL (2007) Kinetics and equilibria of cis/trans isomerization of backbone amide bonds in peptoids. *J Am Chem Soc* 129: 12042–12048.
36. Duffy EM, Severance DL, Jorgensen WL (1992) Solvent effects on the barrier to isomerization for a tertiary amide from ab initio and Monte Carlo calculations. *J Am Chem Soc* 114: 7535–7542.
37. Torrie GM, Valleau JP (1977) Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *J Comput Phys* 23: 187–199.
38. Sugita Y, Okamoto Y (1999) Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett* 314: 141–151.
39. Stringer JR, Crapster JA, Guzei IA, et al. (2011) Extraordinarily Robust Polyproline Type I Peptoid Helices Generated via the Incorporation of alpha-Chiral Aromatic N-1-Naphthylethyl Side Chains. *J Am Chem Soc* 133: 15559–15567.
40. Ramachandran GN, Ramakrishnan C, Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. *J Mol Biol* 7: 95–99.
41. Butterfoss GL, Renfrew PD, Kuhlman B, et al. (2009) A Preliminary Survey of the Peptoid Folding Landscape. *J Am Chem Soc* 131: 16798–16807.
42. Mohle K, Hofmann HJ (1996) Peptides and peptoids—A systematic structure comparison. *J Mol Model* 2: 307–311.
43. Miertuš S, Scrocco E, Tomasi J (1981) Electrostatic interaction of a solute with a continuum. A direct utilization of ab initio molecular potentials for the prevision of solvent effects. *Chem Phys* 55: 117–129.
44. Pascual-Ahuir JL, Silla E, Tomasi J, et al. (1987) Electrostatic interaction of a solute with a continuum. Improved description of the cavity and of the surface cavity bound charge distribution. *J Comput Chem* 8: 778–787.
45. Parker BF, Knight AS, Vukovic S, et al. (2016) A Peptoid-Based Combinatorial and Computational Approach to Developing Ligands for Uranyl Sequestration from Seawater. *Ind Eng Chem Res* 55: 4187–4194.
46. Cancès E, Mennucci B, Tomasi J (1997) A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics. *J Chem Phys* 107: 3032–3041.
47. Cornell W, Cieplek P, Bayly CI, et al. (1995) A Second Generation Force-Field for the Simulation of Proteins, Nucleic-Acids, and Organic-Molecules. *J Am Chem Soc* 117: 5179–5197.
48. Hawkins GD, Cramer CJ, Truhlar DG (1998) Universal Quantum Mechanical Model for Solvation Free Energies Based on Gas-Phase Geometries. *J Phys Chem B* 102: 3257–3271.
49. Bradley EK, Kerr JM, Richter LS, et al. (1997) NMR Structural Characterization of Oligo-N-Substituted Glycine Lead Compounds from a Combinatorial Library. *Mol Divers* 3: 1–15.

50. Mann G, Yun RH, Nyland L, et al. (2002) The Sigma MD Program and a Generic Interface Applicable to Multi-Functional Programs with Complex, Hierarchical Command Structure, In: Schlick T, Gan HH, *Computational Methods for Macromolecules: Challenges and Applications*, Springer, Berlin, Heidelberg, 129–145.
51. Hermans J, Berendsen HJC, Van Gunsteren WF, et al. (1984) A consistent empirical potential for water–protein interactions. *Biopolymers* 23: 1513–1518.
52. Butterfoss GL, Yoo B, Jaworski JN (2012) De novo structure prediction and experimental characterization of folded peptoid oligomers. *Proc Natl Acad Sci USA* 109: 14320–14325.
53. Wang JM, Wolf RM, Caldwell JW, et al. (2004) Development and testing of a general amber force field. *J Comput Chem* 25: 1157–1174.
54. Onufriev A, Bashford D, Case DA (2004) Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins* 55: 383–394.
55. MacKerell AD, Bashford D, Bellott M, et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102: 3586–3616.
56. Case DA, Cheatham TE, Darden T, et al. (2005) The Amber biomolecular simulation programs. *J Comput Chem* 26: 1668–1688.
57. Jorgensen WL, Maxwell DS, TiradoRives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc* 118: 11225–11236.
58. Moehle K, Hofmann HJ (1996) Peptides and peptoids—A quantum chemical structure comparison. *Biopolymers* 38: 781–790.
59. Jorgensen W, Chandrasekhar J, Madura J, et al. (1983) Comparison of Simple Potential Functions for Simulating Liquid Water. *J Chem Phys* 79: 926–935.
60. Tobias DJ, Brooks CL (1988) Molecular dynamics with internal coordinate constraints. *J Chem Phys* 89: 5115–5127.
61. Wang J, Wang W, Kollman PA, et al. (2006) Automatic atom type and bond type perception in molecular mechanical calculations. *J Mol Graph Model* 25: 247–260.
62. Jakalian A, Jack DB, Bayly CI (2002) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J Comput Chem* 23: 1623–1641.
63. Mukherjee S, Zhou G, Michel C, et al. (2015) Insights into Peptoid Helix Folding Cooperativity from an Improved Backbone Potential. *J Phys Chem B* 119: 15407–15417.
64. Lifson S, Roig A (1961) On the Theory of Helix—Coil Transition in Polypeptides. *J Chem Phys* 34: 1963–1974.
65. Mirijanian DT, Mannige RV, Zuckermann RN, et al. (2014) Development and use of an atomistic CHARMM-based forcefield for peptoid simulation. *J Comput Chem* 35: 360–370.
66. Jordan PA, Bishwajit P, Butterfoss GL, et al. (2011) Oligo(N-alkoxy glycines): trans substantiating peptoid conformations. *J Pept Sci* 96: 617–626.
67. Nam KT, Shelby SA, Cho PH, et al. (2010) Free-floating ultrathin two-dimensional crystals from sequence-specific peptoid polymers. *Nat Mater* 9: 454–460.
68. Mannige RV, Kundu J, Whitlam S (2016) The Ramachandran Number: An Order Parameter for Protein Geometry. *PLoS One* 11: e0160023.
69. Reith D, Putz M, Muller-Plathe F (2003) Deriving effective mesoscale potentials from atomistic simulations. *J Comput Chem* 24: 1624–1636.

70. Izvekov S, Voth GA (2005) Multiscale coarse graining of liquid-state systems. *J Chem Phys* 123: 134105.
71. Shell MS (2008) The relative entropy is fundamental to multiscale and inverse thermodynamic problems. *J Chem Phys* 129: 144108.
72. Haxton TK, Mannige RV, Zuckermann RN, et al. (2015) Modeling Sequence-Specific Polymers Using Anisotropic Coarse-Grained Sites Allows Quantitative Comparison with Experiment. *J Chem Theory Comput* 11: 303–315.
73. Sanii B, Haxton TK, Olivier GK, et al. (2014) Structure-Determining Step in the Hierarchical Assembly of Peptoid Nanosheets. *ACS Nano* 8: 11674–11684.
74. Haxton TK, Zuckermann RN, Whitlam S (2016) Implicit-Solvent Coarse-Grained Simulation with a Fluctuating Interface Reveals a Molecular Mechanism for Peptoid Monolayer Buckling. *J Chem Theory Comput* 12: 345–352.
75. Drew K, Renfrew PD, Butterfoss GL (2013) Adding Diverse Noncanonical Backbones to Rosetta: Enabling Peptidomimetic Design. *PLoS One* 8: e67051.
76. Kaufmann KW, Lemmon GH, DeLuca SL, et al. (2010) Practically Useful: What the Rosetta Protein Modeling Suite Can Do for You. *Biochemistry* 49: 2987–2998.
77. Renfrew PD, Craven TW, Butterfoss GL, et al. (2013) A Rotamer Library to Enable Modeling and Design of Peptoid Foldamers. *J Am Chem Soc* 136: 8772–8782.
78. Laio A, Gervasio FL (2008) Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Rep Prog Phys* 71: 126601.



AIMS Press

© 2017 Erik E. Santiso, et al., licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)