

UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Solution Self-Assembly of Sequence Specific Biomimetic Polymers

Permalink

<https://escholarship.org/uc/item/8fk384nq>

Author

Murnen, Hannah

Publication Date

2012

Peer reviewed|Thesis/dissertation

Solution Self-Assembly of Sequence Specific Biomimetic Polymers

by

Hannah Murnen

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemical Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Rachel A. Segalman, Chair

Professor Nitash P. Balsara

Professor Sanjay Kumar

Spring 2012

Solution Self-Assembly of Sequence Specific Biomimetic Polymers

© 2012

By Hannah Murnen

Abstract

Solution Self-Assembly of Sequence Specific Biomimetic Polymers

By

Hannah Murnen

Doctor of Philosophy in Chemical Engineering

University of California, Berkeley

Professor Rachel A. Segalman, Chair

Biological molecules, such as polypeptides, form the basis for most of life's functions. These simple linear polymers made up of only 20 amino acids can fold to form complex and intricate structures that span several length scales. The structures that arise from these molecules stem from their high level of molecular control. Each molecule is an exact sequence with complete monodispersity. From this angstrom level of control, micron scale structures can be assembled. This process is one of the most studied in the history of science, but due to the diversity of amino acids and potential interactions, it is still nearly impossible for scientists to look at a peptide sequence and predict a folded structure. Likewise, it is very difficult to choose a desired structure for a particular function and reverse engineer the linear sequence that will provide that structure. Therefore, there is the need for simplified systems to understand the interactions involved in protein folding and to begin to build the knowledge necessary for engineering similar types of functional structures.

This thesis uses sequence specific biomimetic polymers, namely polypeptoids or poly N-substituted glycines, to fundamentally probe the chain conformation and assembly properties of sequence specific polymers. Polypeptoids or N-substituted glycines, are sequence specific biomimetic chains that have the same backbone as polypeptides. However, rather than the side chain being attached through the alpha carbon, it is attached to the backbone nitrogen. This chemical alteration has several implications for the system including the elimination of backbone hydrogen bonding and chirality. This chemical alteration yields a more flexible and tunable chain where intramolecular interactions can be modulated by the introduction of different side chains. In addition, the synthesis of polypeptoids is a simple two step submonomer addition that uses a primary amine as the submonomer. This results in a very high yield synthesis with virtually limitless possibilities for side chains due to the commercial availability of a wide variety of primary amines.

Using this modular system, my thesis focuses on understanding the solution self assembly of a sequence specific biomimetic polymer. Much of the work has focused on understanding the single chain conformation and collapse of polypeptoids in order to apply this information to larger self assembly systems. The persistence lengths of several polypeptoids have been measured including those containing secondary structure or ionic groups in order to understand the effect of these factors on the chain conformation. Additionally, the collapse or folding of a single polypeptoid chain into a globule structure is discussed. The impact of monomer sequence

on this collapse was investigated and shown to have an important effect both on the coil to globule transition as well as the resulting globule structure. Finally, a hierarchical super helix formed through the assembly of an amphiphilic diblock copolypeptoid is discussed. Using chemical modifications coupled with x-ray scattering, the super structure was shown to include ordering stemming from the angstrom level packing of molecules all the way up to the micron scale diameter of the helix.

Table of Contents

Table of Contents	i
List of Figures	iii
List of Tables	v
Acknowledgements	vi
Chapter 1. Introduction	1
1.1 Complexities of Macromolecular Self-Assembly	1
1.1.1 Single chain folding and conformations	1
1.1.2 Mesoscale self-assembly	5
1.2 Bioinspired Polymers and Self-Assembly	8
1.2.1 Synthetic polypeptides	8
1.2.2 Peptidomimetic materials	9
1.2.3 Polypeptoids	9
1.3 Motivation and Thesis Outline	12
1.4. References	13
Chapter 2. Determination of the persistence length of helical and non-helical polypeptoids in solution	20
2.1. Introduction	20
2.2. Experimental Section	21
2.3. Results and Discussion	24
2.3.1. Circular Dichroism	24
2.3.2. Small Angle Neutron Scattering (SANS)	26
2.4. Conclusions	33
2.5. Appendix	34
2.5. Acknowledgements	41
2.6. References	41
Chapter 3: Persistence length of polyelectrolytes with precisely located charges	45
3.1. Introduction	45
3.2. Experimental methods	47
3.3 Results and Discussions	50
3.4 Conclusions	59
3.5. Acknowledgements	59
3.6. References	59
Chapter 4. Experimental Validation of the HP Model for Globule Formation	62

4.1. Introduction	62
4.2. Experimental Section	64
4.3. Results and Discussion.....	70
4.4. Conclusions	78
4.4. Appendix	79
4.5. Acknowledgements	83
4.6. References	83
Chapter 5. Hierarchical Self-assembly of a Biomimetic Diblock Copolypeptoid into Homochiral Superhelices	86
5.1. Introduction	86
5.2. Experimental Section	87
5.3. Results and Discussion.....	91
5.4. Conclusions	102
5.5. Appendix	103
5.6. Acknowledgements	109
5.7. References	109
Chapter 6. Conclusions and Future Outlook.....	113

List of Figures

Figure 1.1. Illustration of the various parameters that define a polymer conformation.	3
Figure 1.2. Coil to globule transition of a polymer.....	3
Figure 1.3. Israelachvili packing factors.....	7
Figure 1.4. Chemical structures of several peptidomimetics.....	11
Figure 1.5. Two-step submonomer synthesis of polypeptoids.....	11
Figure 2.1. Chemical structures of the helical and nonhelical compound.....	22
Figure 2.2. CD spectra for different chain lengths of a helical polypeptoid as well as a nonhelical polypeptoid.....	25
Figure 2.3. Heating (a) and solvent (b) do not significantly affect the CD spectra of 2c.....	27
Figure 2.4. Small angle neutron scattering (SANS) shows the relatively short persistence length of both polypeptoids.....	28
Figure 2.5. The persistence length, as determined using the worm-like chain model, of polypeptoids ranging in length from 18-48 monomers plotted versus the number of monomers.....	32
Figure 2.A.1. Guinier plots for all compounds.....	36
Figure 2.A.2. Fitted persistence lengths using the wormlike chain equation over a series of polypeptoid chain lengths.....	37
Figure 2.A.3. Compound 3 contains a racemic mixture of α -chiral side chains and is proposed to be non-helical.....	38
Figure 2.A.4. Circular dichroism at varying temperatures for each helical polypeptoid.....	40
Figure 3.1. The titration curves of $p(\text{Nce})_{36}\text{ac}$ and $p(\text{NceNme})_{18}\text{ac}$	53
Figure 3.2. Kratky plots for each polymer solution.....	55
Figure 3.3. The persistence lengths for $p(\text{Nce})_{36}\text{ac}$ and $p(\text{NceNme})_{18}\text{ac}$	56
Figure 3.4. Electrostatic contribution to the chain persistence length as a function of the reduced Debye screening length.....	58
Figure 4.1. Protein-like and repeating sequence polypeptoid 100mers.....	65
Figure 4.2. Representative analytical traces and MALDI spectra.....	68
Figure 4.3. Equilibrium denaturant titration of peptoid globules with acetonitrile.....	72
Figure 4.4. Hydrodynamic radius as measured using dynamic light scattering.....	74
Figure 4.5. Nile red fluorescence showing an increased emission peak from the protein-like sequence.....	77
Figure 5.1. The sheet structures formed from the diblock copolypeptoid $p\text{Npe}_{15}\text{Nce}_{15}$	92
Figure 5.2. A model of the proposed self assembly process.....	93
Figure 5.3. Helix formation of $p\text{Npe}_{15}\text{Nce}_{15}$	94
Figure 5.4. SAXS pattern of giant super helices.....	95

Figure 5.5. X-ray scattering on helical samples of <i>pNpe</i> ₁₅ <i>Nce</i> ₁₅ , <i>pNbn</i> ₁₅ <i>Nce</i> ₁₅ and <i>pNpp</i> ₁₅ <i>Nce</i> ₁₅	97
Figure 5.6. Crystal structure of a model cyclic dipeptoid 1,4-bis-(2-phenethyl)-piperazine-2,5-dione.....	99
Figure 5.7. Sequences with varying locations of charges.....	101
Figure 5.A.1. Self assembly quantification.....	103
Figure 5.A.2. Titration curve.	104
Figure 5.A.3. Solid SAXS matches liquid SAXS.....	105
Figure 5.A.4. Crystallization of <i>pNpp</i> ₁₅	106
Figure 5.A.5. The image gallery of helices.....	107
Figure 5.A.6. Circular dichroism spectra of a 300 μM solution of helices.	108

List of Tables

Table 2.1. All the compounds used in this study	23
Table 2.2. Fitted parameters for Compounds 1 and 2 with $n = 6$	30
Table 2.A.1. The contour length and the persistence length as fit using the semiflexible cylinder model.....	39
Table 3.1. The polymers used, their purity, observed mass, and chemical structures	48
Table 3.2. All solutions used in this study.	49
Table 4.1. HPLC and mass spectrometry of 50mer and 100mer peptoids.	67
Table 4.2. Calculated values from a two-state model.....	73
Table 4.3. The characteristic ratio, ρ , for each compound in buffer and in 80% acetonitrile.....	75
Table 5.1. Chemical structure of the peptoid monomers used and their abbreviations.	89
Table 5.2. All of the polymers synthesized and used in this article.....	90

Acknowledgements

First and foremost I need to thank my advisors. Rachel has provided enormous support throughout my entire PhD. She has allowed me to find my own way through graduate school while still guiding me with strong scientific principles. She has even allowed me to indulge in my own learning opportunities through internships and other extracurricular activities for which I am truly grateful. Ron Zuckermann has been a perfect balance of rigorous scientific advice and reminding me not to take myself too seriously. His levity and honesty has been much appreciated.

Secondly I need to thank my current and former labmates. Adrienne Rosales and I have spent more time in lab together than is probably a good idea and we are still great friends so I consider that a success! Along with Adrienne, Megan Hoarfrost and Bryan McCulloch have helped make the Segalman group what it is today and I have enjoyed working (and playing) with them over the last 4 and half years. Miguel Modestino and Victor Ho have been here nearly as long as I have and I will miss the wonderful discussions we have had while in lab together. Barbara Eckerdt, Boris Russ, Shannon Yee, Nelson Coates, Wendy Van Zoelen, Hilda Buss, Eddie Buehler, Cynthia Chen, William Chang, Jonathan King, Yanika Schneider, Yuefei Tao, Brad Olsen, Justin Virgili, Jibin Sun, Saar Kirmayer, Bryan Boudouris, Jon Malen, Joe Feser, Young rae Hong, Kevin See, Kasper Moth-Poulsen, Robert Wang, Joe Chen and Peter Doak have all provided assistance and entertainment along the way and I am thankful for all their help.

Finally, my family and friends have provided a ton of support. Both of my sisters are amazing people and I am so lucky to have such great relationships with them. Along with my sisters, my mom and dad have been extremely supportive. John has provided wonderful distractions from grown-up life whenever I see him. I am excited to be moving to the east coast where I can be closer to all of them! All of my friends, near and far have been wonderful at listening to me when I need to vent and celebrating when things are going well. Finally, Jon is everything I could ever hope for in a partner and more.

Chapter 1. Introduction

Biological molecules, such as polypeptides or DNA, form the basis for most of life's functions. In the case of polypeptides, these very simple polymers made up of a linear combination of only 20 amino acids can fold to form complex and intricate structures that span several length scales. These sophisticated structures stem from the high level of molecular control present in the polypeptide system. Each molecule is an exact sequence with complete monodispersity. From this angstrom level of control, micron scale structures can be assembled. Due to the diversity of amino acids and potential interactions, it is still a hard task for scientists to take a peptide sequence and predict a folded structure. Likewise, it is very difficult to choose a desired structure for a particular function and reverse engineer the linear sequence that will provide that structure. Therefore, there is the need for simplified systems to understand the interactions involved in protein folding and to build the knowledge necessary for engineering similar types of functional structures.

In this work, simplified biomimetic polymers, poly N-substituted glycines, or polypeptoids are used to probe the solution assembly of sequence specific polymers. Starting from the very smallest length scale, the persistence length and single chain conformations of various polypeptoids are probed. This understanding is used to develop a hierarchical micron-sized super helix comprised of a simple amphiphilic polypeptoid diblock.

1.1 Complexities of Macromolecular Self-Assembly

The self-assembly of macromolecules can be incredibly complex due to the many different forces at play. In addition to atomistic level interactions such as ionic or hydrophobic forces, the situation is complicated by the fact that atoms are connected together in long chains leading to entropic considerations as well. In the case of biological macromolecules, the incorporation of many different monomers adds more interactions that can drive self-assembly. Due to this complexity, predicting the assembly of these biological molecules is still an unsolved problem. In general, understanding the fundamental behavior of macromolecules at small length scales can lead to further understanding about how large scale assemblies occur and how scientists can control that assembly.

1.1.1 Single chain folding and conformations

The behavior of a chain in solution depends on many variables such as the length of the polymer and the interactions of the chain components, both with themselves and with the solvent. It is not straightforward to describe the conformation of a chain in 3-dimensional space and thus, several measurable parameters (illustrated in Figure 1.1) are needed. The end to end distance R , is simply the distance between two ends of a polymer while the radius of gyration, R_g , another measure of chain size, is an average of the distance from the center of mass of a polymer to each segment of the chain (Figure 1.1a). The persistence length, L_p , is a measure of how stiff a polymer is. The larger the L_p is, the more the polymer behaves like a stiff rod, rather than a cooked piece of spaghetti (Figure 1.1b). The R_h or hydrodynamic radius is determined by measuring the diffusion of the polymer through a solvent with viscosity, η , and then using Stokes-Einstein relationship (below) to calculate a hydrodynamic radius. The R_h is an indirect measurement of size and due to the fact that the Stokes Einstein relationship assumes a sphere, it is not directly related to a physical dimension of a polymer.

$$D = \frac{kT}{6\pi\eta R_h}$$

All of these measures help to express how the polymer is arranged in a particular solution and are useful in understanding how several chains would interact with each other.

Much of the work in understanding polymer conformation has started from basic synthetic polymers where the arrangements in solution tend to be simple because there are only a few types of monomers. In comparison to biological molecules, synthetic polymers are relatively well understood in terms of chain conformation and self-assembly properties in solution. Numerous studies have considered the simplest case where the polymer is dissolved in a theta solvent where the interactions between the monomers and solvent are essentially equal^{1,2}. In this regime, the polymer has a random walk behavior such that each segment of the polymer is randomly located in 3 dimensional space next to the previous segment. This is known as a Gaussian coil conformation. In this regime, the persistence length is low and the polymer has predictable scaling relationships between the chain dimensions and the molecular length of the polymer. For example, the end to end distance of a polymer, R , can be shown both theoretically³ and experimentally⁴ to scale as $N^{1/2}$ where N is the number of monomers in the chain. These scaling relationships can change for polymers who adopt a rigid conformation, such as those containing conjugation in their backbone or those residing in a secondary structure such as an alpha helix. In this case, the L_p is much larger and R scales linearly with N . Understanding these scaling relationships allows the prediction of polymer configurations based on the molecular structure of the polymer.

The single chain conformation of a polymer can also change as conditions change. For example, when a polymer is in a poor solvent where the interactions between the solvent and the polymer are unfavorable, the polymer might collapse into a globule (Figure 1.2). It is possible to change the interaction of the polymer with the solvent, and therefore the polymer conformation, by changing factors such as the temperature, pH, or even the solvent composition. Poly(*N*-isopropyl acrylamide), or PNIPAM, holds particular interest in the study of these transitions, because in contrast to most polymers, it has a lower critical solution temperature such that if the temperature is increased, the solubility of the polymer in water decreases.⁵ The transition from coil to globule can be probed through the measurement of R_g and R_h at various temperatures.⁶ These measurements show the coil to globule transition occurring between 30°C and 35°C and there have been numerous studies probing the effect of comonomers,^{7,8} surfactants⁹ or different solvents¹⁰ on this transition. One can imagine that a polymer designed to collapse or unfold upon the introduction of a trigger would be ideal for applications such as drug delivery and chemical sensors.

Polymers with higher levels of sequence complexity such as biologically occurring polypeptides can have correspondingly higher levels of functionalities. The ability to understand and predict their self-assembly could therefore hold potential for designing synthetic mimics that can perform some of these same functions. However, understanding the single chain folding of sequence specific biopolymers such as polypeptides is inherently more difficult than synthetic homopolymers.

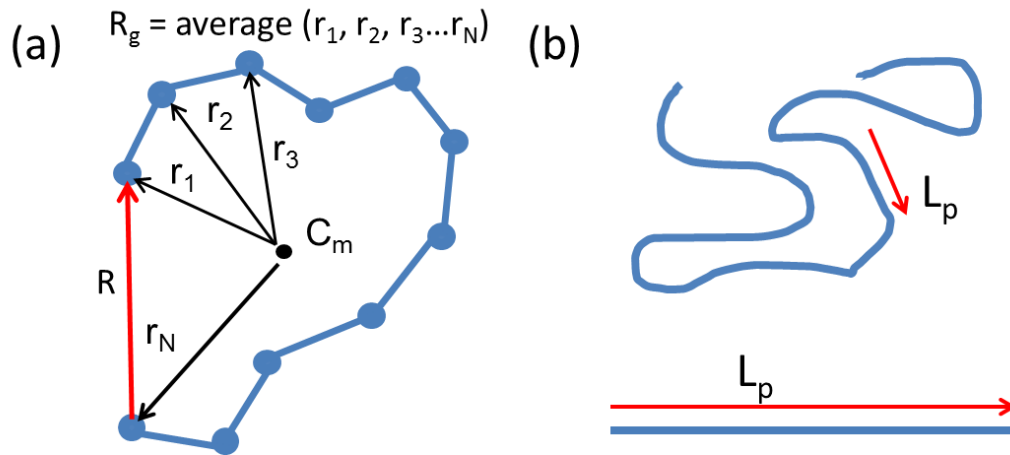


Figure 1.1. Illustration of the various parameters that define a polymer conformation.
 In (a) the radius of gyration (R_g) and the end to end to distance R , are demonstrated for a polymer chain. In (b), the persistence length is shown for both a flexible polymer with a relatively short L_p and a rod-like polymer with a very large L_p .

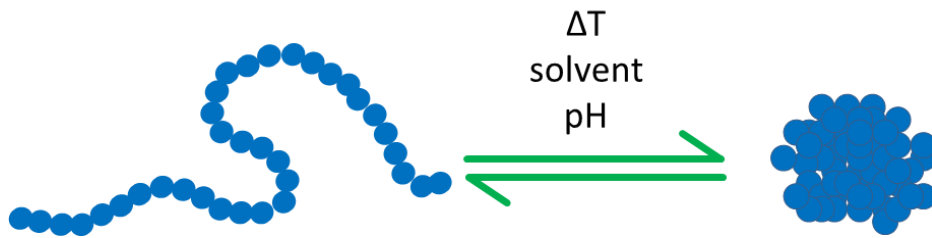


Figure 1.2. Coil to globule transition of a polymer
 A change in temperature, solvent or pH may trigger the collapse of the polymer by changing the interaction between the solvent and the polymer. The transition is usually reversible such that the polymer can unfold back into its coil state.

While the state of the art method for characterization of folded polypeptide structures is x-ray crystallography to obtain complete crystal structures, it is possible to gain a lot of information from simpler measurements of polymer conformation. For example, the persistence length of a polypeptide changes dramatically when it is in an unstructured conformation¹¹⁻¹³ (<1 nm) versus when it is adopting a secondary structure such as an alpha helix¹⁴ (~200 nm). The R_g can also be used to probe the arrangement of a polypeptide. Semisotnov and coworkers looked at several different proteins and showed that in their native state, their R_g 's were significantly smaller than in the coil state due to the compact nature of the folded structure.¹⁵ Pollack and coworkers took this even further and tracked the folding of a protein in real-time by continually measuring the R_g using small angle x-ray scattering during the initiation of the folding process.¹⁶ While these types of measurements are far less detailed than the crystal structure of a folded protein, they provide insight into the fundamental configuration of the polypeptide.

Beyond the characterization of existing folded polypeptide structures, there has been a lot of work in both de novo protein design and ab initio protein structure prediction. Although these are difficult topics, much progress has been made in both areas. Starting in the late 1980s and early 1990s, progress was made on the ability to predict a protein structure from the polypeptide sequence primarily through the identification of sequence portions that were similar to solved protein structures.¹⁷⁻¹⁹ This type of method resulted in predictions that were accurate at best about 70% of the time.²⁰ Many of these predictions focused on the packing of the core, usually a hydrophobic region at the center of a folded protein under the hypothesis that if the core was folded correctly, the structure would extend to the outer residues. These methods resulted in the prediction of folded helical bundle proteins and demonstrated the importance of hydrophobic forces in driving folding.²¹⁻²³

As computing power improved, molecular dynamics models such as CHARMM²⁴ were developed to calculate the energy of a given protein structure and help to determine the most stable structure for a sequence. Recently David Baker's group at University of Washington has developed a similar program entitled Rosetta to compute the energy of interactions between polypeptide pieces that is used to find the lowest energy conformation for a protein sequence.^{25, 26} The game Fold-it uses Rosetta and allows people to manipulate polypeptide sequences to find the lowest energy conformation.²⁷ It has been shown that people are often better than a computer at finding the lowest energy structure for a polypeptide and the goal of the game is to find the successful strategies used by human protein folders and teach them to computer protein folding programs. Other groups have employed similar computational methods,^{28, 29} but there are several limitations to such methods. The first is that while these methods can generally provide fairly good agreement with known structures (similar energy values for folded structures), they still have some inaccuracies in the exact placement of various amino acids.³⁰ In addition, they require large amounts of computer power and can be very time intensive due to the large number of potential configurations. Therefore, there is still the need for improved understanding of polypeptide folding systems in order to provide better folding rules for the computer algorithms.

The previous paragraphs discussed taking a polypeptide sequence and finding what structure it will fold into. It would also be useful to be able to perform the opposite procedure and predict a polypeptide sequence that will give a desired structure. In this area, Steve Mayo and coworkers have been one of the most successful groups. Their approach uses computational methods that incorporate both theory and experimental testing.^{31, 32} In one example, they were able to predict a sequence that folds into a zinc finger protein from a combinatorial library of 1.9×10^{27} possible

amino acid sequences.³³ Several groups have developed similar techniques for design,³⁴⁻³⁶ but in general these types of calculations have comparable limitations to those of the previous paragraphs in that they require large amounts of computer power and still have limited precision. In order to decrease the computing power necessary to make these predictions, it is necessary to develop better rules for a computer to use as it performs the algorithms. One way to go about developing better rules is to improve our understanding of the self-assembly process. In any given polypeptide chain there may be many different monomers all contributing hydrophobic, electrostatic and steric interactions as well as the potential for backbone hydrogen bonding and chirality influences. Simplifying the number of interactions both by decreasing the types of monomers as well as simplifying the chemistry can provide opportunities for increased understanding. Further discussion (section 1.2) will probe this possibility through the use of simplified peptidomimetic polymer systems.

It is often desirable to design larger assemblies that incorporate multiple macromolecules. These types of supramolecular assemblies can provide functions ranging from artificial catalysis to filtration among many others. The next section will discuss the rules for designing the assembly of mesoscale structures from both synthetic and biological macromolecules.

1.1.2 Mesoscale self-assembly

The patterning of polymers on both the nanoscale and mesoscale can be desirable for a wide range of applications from catalysis to membrane development. Reaching micron and larger dimensions requires assembly of multiple polymer chains where multiple monomer types can provide interactions to drive the assembly of larger structures. The simplest way to increase monomer types is to build a diblock copolymer where two types of monomers are connected in a block formation. In addition to interactions with the solvent, each of the monomers also has an interaction with the other monomer. This small increase in sequence complexity leads to much larger multi-chain assemblies in contrast to the single chain globules seen from homopolymers.

Amphiphilic diblocks are one of the most common systems designed to build supramolecular assemblies. In these systems, one of the blocks is hydrophilic while the other is hydrophobic. When the molecule is dissolved in aqueous solution, the hydrophobic block wants to isolate itself from the solution while the hydrophilic block wants to remain well dissolved. However, the two blocks are covalently connected forcing them to balance their enthalpic drive to segregate and the entropic penalty for stretching away from each other. Generally, these types of sequences can assemble into several different structures in solution including micelles, vesicles, bilayers, etc. The particular structure that is formed is related to many factors including the relative selectivity of the solvent for either block, the length of the polymer, and the volume fraction of each block.

As a demonstration of these principles, Eisenberg and coworkers have performed numerous experiments to understand the solution assembly of polystyrene-*b*-poly acrylic acid.³⁷⁻⁴¹ In this system, the polystyrene is hydrophobic while the poly acrylic acid is hydrophilic. The relative block lengths and the solvent content (a mixture of water, tetrahydrofuran and dioxane) control whether the polymers assemble into vesicles, spherical micelles, or rod-like micelles. In order to generalize this phenomenon, Israelachvili developed a method to use the critical packing factor, which is related to the diblock copolymer shape, to predict what type of structure the molecule will assemble into.⁴² The critical packing factor, p , takes into account the volume of the hydrophobic block, v_h , the packing area of the polar block, a_p , and the length of the hydrophobic block, l_h . This is most commonly used for surfactants where the polar block can be thought of as

a head group and in this case the packing area is simply the cross sectional area of the head group.

$$p = \frac{v_h}{a_p l_h}$$

This concept is illustrated in Figure 1.3 where the potential self-assembly structures are correlated with their respective packing factors. Using these types of theoretical predictions, it is clear that understanding the single chain configuration of a polymer is imperative in being able to design mesoscale structures using that polymer.

As seen in Figure 1.3, most of the structures assembled from block copolymers contain chains that are nonspecifically packed into tubes, spheres or bilayers. For some applications such as templating or tissue scaffolding, it may be necessary to obtain more hierarchical structures where order spans from the angstrom-scale packing of atoms up to the micron scale. For this, it is necessary to turn to more complex sequences that contain many monomers and therefore more interactions that can drive ordering of a much more specific nature. For example, numerous studies have focused on the naturally occurring polypeptide known to form amyloid fibrils. In these structures, the peptide forms anti-parallel β -sheet structures which then stack on top of each other to form fibrils.⁴³⁻⁴⁵ These fibrils show scattering peaks that correspond to spacings of 4.75 Å and 4.3 Å and yet the dimensions of the fibrils are on the micron scale. The monodispersity and sequence specificity of the molecule allows it to pack at many levels, leading to the hierarchical assembly across several length scales. In addition, because the sequence of the molecule is controllable, scientists can make known alterations to the sequence and see how those alterations change the overall structure.⁴⁵⁻⁴⁷ This tactic provides a good understanding of how each interaction contributes to the larger structure and how the molecule is arranged within the structure.

In addition to amyloid fibers, proteins have been shown to assemble into nanotubules,⁴⁸ fibrils,^{49, 50} and sheets⁵¹ that can span microns in size. As in the case of the amyloid fibrils, these structures are generally very hierarchical in nature and the exact monomer sequence has a large effect on the overall structure. In addition, because of the sequence specificity of polypeptides, it is possible to introduce functional moieties into the supramolecular structures at desired locations, leading to interesting functions. For example, protein nanotubes have been designed to contain a particular ligand binding site on their exterior allowing delivery of the tube to a desired location, perhaps for the purpose of drug delivery or chemical sensing.⁵² The ability to introduce these functional groups is a direct result of the molecular control available in the polypeptide system. Coupled with the hierarchical assembly, this makes polypeptide assembly a powerful tool. This is in contrast to synthetic systems where the polydispersity and non-specific interactions lead to systems with poor molecular control and therefore low levels of order.

However, polypeptide systems are inherently complex. Thus, while it is possible to take biologically occurring systems and tweak them for our needs, it is difficult to de novo design a sequence to assemble into a desired structure. Therefore scientists have looked to simplified systems to develop a set of assembly rules to allow the targeted design of sequences for particular structures and functions. The next section will discuss some of these bioinspired systems and what has been learned about their self assembly.

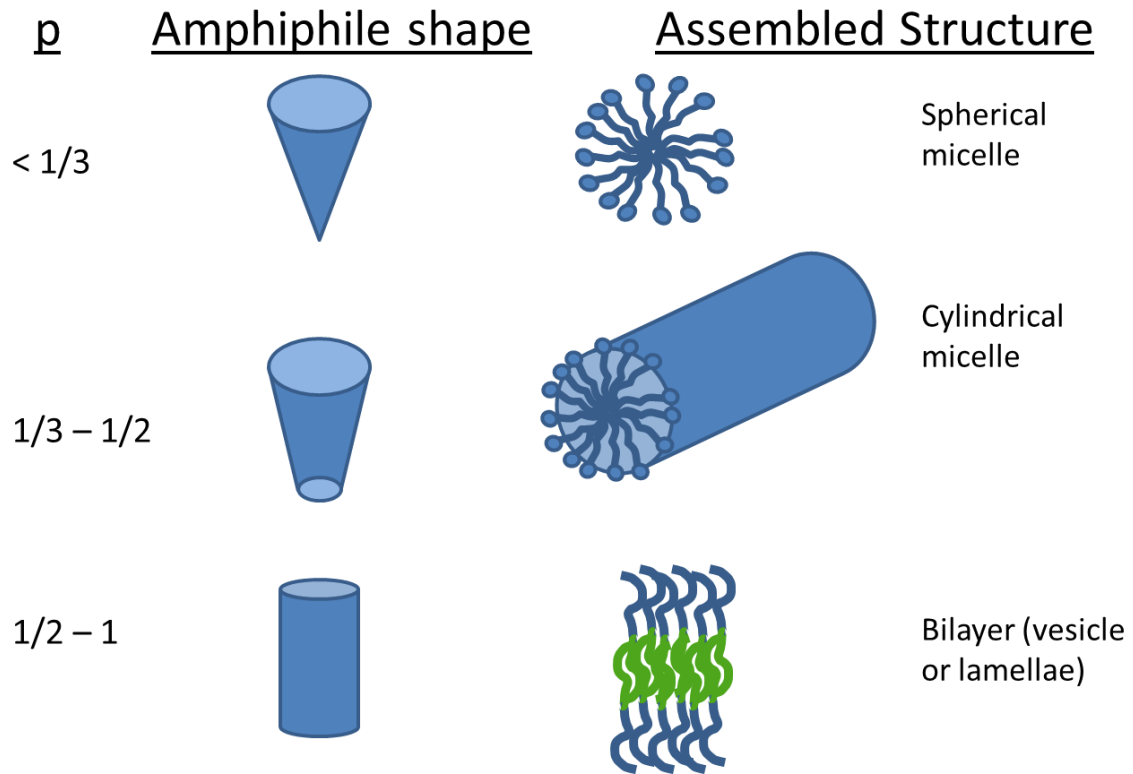


Figure 1.3. Israelachvili packing factors

The packing factor predicts what structure an amphiphile will assemble into based on its molecular shape. This figure is modified from reference 20.

1.2 Bioinspired Polymers and Self-Assembly

As discussed in the previous section, interesting mesoscale assembly can occur in the case of both sequence specific biological macromolecules as well as simpler synthetic polymers that have far less complicated sequences and interactions. However, the complexity of the mesoscale assembly is often tied to the complexity of the underlying polymer. Controlling and harnessing that complexity to build desired mesoscale structures is often difficult due to the many and sometimes competing interactions at play. Therefore, researchers have turned to systems that simplify various portions of biological polypeptides while still retaining some of the attributes of these biological polymers. These systems can either come in the form of a new type of chemistry such as beta-peptides or peptoids to name a few; or they can maintain the same chemistry, polypeptides, but use far simpler sequences such as block copolymers of polypeptides. Both of these types of bioinspired systems will be discussed here.

1.2.1 Synthetic polypeptides

Synthetic polypeptides often have simpler sequences leading to more predictable assembly properties than their natural analogs. One of the ways to create synthetic polypeptides is to take fragments of polypeptides that are known to direct assembly. This technique allows the development of materials that utilize the known assembly properties of a particular polypeptide piece while attempting to add new functionalities. For example, a beta-sheet forming fragment of 159 residues has been shown to assemble into a fibril structure which is then used to template an ordered array of gold nanoparticles.⁵³ Surprisingly, the structure directing polypeptide piece can also be quite short. Thermoresponsive hydrogels can be synthesized through the use of polymers containing elastin-mimetic peptide fragments that are only 15 residues long.⁵⁴ While fragments are a relatively easy way to harness some of the functionalities available to polypeptides, they do not always provide new information about why the molecules assemble as they do. A large area of research uses synthetic peptides with simplified sequences to probe the root causes of assembly.

The simplest class of synthetic polypeptides is homopolypeptides. Gallot and coworkers performed several studies in the 70s and 80s looking at the assembly of diblock copolymers containing polypeptide blocks⁵⁵⁻⁵⁸ showing that predominantly the block copolymers assembled into lamella structures in concentrated aqueous solutions. Building off of this previous work, Tim Deming's group further developed the N-carboxyanhydride synthesis of these materials to allow the synthesis of non-sequence-specific low polydispersity homopolypeptides with large molecular weights and their corresponding block copolymers.^{59, 60} They studied these block copolypeptides in a number of configurations, showing the production of stable vesicles^{61, 62} and hydrogels^{63, 64} from their materials. Most recently, block copolypeptides were designed to create stable water-oil-water double emulsions⁶⁵ demonstrating the variety of available mesoscale structures from polypeptides. Other groups have also utilized block copolypeptides to design tailored supramolecular structures. Poly sarcosine-b-(leucine-aminobutyric acid) was shown to assemble into nanotubes of varying lengths and sizes where the Leu-AIB block forms an alpha helix, forcing the molecule into the assembled nanotube.⁶⁶ The ability of the peptide portion to change secondary structure due to a trigger such as a pH or temperature change allows the resulting mesoscale structure to be environmentally responsive. Thermally responsive micelles have been developed from a triblock copolypeptide where an increase in temperature caused the middle block to become alpha helical, inducing self assembly into micelles.⁶⁷ In another example, gelation was induced for a polypeptide diblock upon the beta-sheet formation of one of

the blocks.⁶⁸ These examples serve to show how the secondary structure of the peptide block can be used to direct larger mesoscale assembly.

While it is clear that polypeptides offer many advantages in designing supramolecular assembly systems, it is still difficult to isolate interactions due to the chemically complex nature of polypeptides. The potential for backbone hydrogen bonding and chirality are constantly present, and can impact any self assembly process. The idealized HP model⁶⁹ where amino acids are simply hydrophobic or polar is difficult to experimentally realize due to these complicating possible interactions. In addition, the submonomers needed for polypeptide synthesis; either amino acids in the case of solid phase synthesis or in the case of Deming's N-carboxyanhydride method, the corresponding NCA's; can be somewhat limiting in terms of the available side chains and if one is expressing polypeptides in bacteria, one is limited to the natural amino acids. In attempting to isolate and investigate specific interactions, it would be helpful to have a system where the main interactions are a result of the side chains (allowing various interactions to be introduced modularly) and where there are a lot of easily accessible and cheap side chains. Various peptidomimetic materials that attempt to address some of these issues while still retaining some of the benefits of polypeptides have been synthesized and assembled and the next section will discuss these systems.

1.2.2 Peptidomimetic materials

Peptidomimetic materials are those that have very similar chemistries to polypeptides but have some key chemical alteration. They are generally sequence specific and have a similar chemical makeup to that of polypeptides. They are designed to preserve some of the attributes of polypeptides while allowing access to new or different properties. For example, beta-peptides have the same backbone as polypeptides, but have an extra carbon inserted into the backbone such that their side chains are bound to the beta carbon, rather than the alpha carbon. They were originally designed as a possible route to creating antibiotics that would be less susceptible to proteases and antibiotic resistance due to their non-natural chemical makeup. They have been shown to assemble into secondary structures such as alpha-helices⁷⁰ and pleated beta-sheets.⁷¹

While beta-peptides are just one example, there are many types of peptidomimetic materials. These include alpha-aminoxy acids,^{72, 73} gamma-peptides,⁷⁴⁻⁷⁶ and azapeptides⁷⁷ to name just a few. However, the synthetic chemistry for many of these types of materials can be complicated, resulting in both added expense and limited length of the resulting polymers. Additionally, many of them alter the backbone of the polypeptide, thus making it harder to translate knowledge from the synthetic system directly to the case of polypeptides. Poly N-substituted glycines or polypeptoids have distinct advantages in both of these areas.

1.2.3 Polypeptoids

Poly N-substituted glycines, or polypeptoids are a peptidomimetic motif that holds promise for a new class of self-assembly materials. Polypeptoids have the same backbone as polypeptides, but the side chain is attached to the nitrogen rather than the alpha carbon. This subtle chemical change eliminates backbone hydrogen bonding and chirality and also has implications for the synthesis of the molecule. Importantly, peptoids can retain some biological activity as well as the ability to assemble into secondary structures. They also have the same backbone and therefore dimensions as polypeptides making them an ideal model material to use for studying the assembly of sequence specific materials in solution. The scalable synthesis of sequence specific polypeptoids was developed in the early 1990s. The chemistry involves a 2 step submonomer

synthesis using primary amines as the side chain vehicle.⁷⁸ In this process a functionalized cross-linked polystyrene bead (rink amide resin) is deprotected and the resulting amine is then bromoacylated using bromoacetic acid and diisopropylcarbodiimide (Figure 1.5).

The bromine is displaced by a primary amine containing the desired side chain. These two steps can be repeated with each cycle adding another monomer. Each step is robust with coupling efficiencies greater than 99% for most monomers allowing for the synthesis of relatively long (>50 monomers) sequence specific chains before purification becomes onerous. This process can be performed using a synthetic robot in order to minimize experimental effort. It is even possible to use commercial robotic peptide synthesizers with only small modifications. After the polymer is finished synthesizing, a solution of trifluoroacetic acid is used to cleave the polymer from the polystyrene bead. The polypeptoid is then purified using reverse phase HPLC. The resulting structure can be confirmed using MALDI mass spectrometry.

As mentioned earlier, peptoids retain the peptide backbone, but do not have backbone hydrogen bonding or chirality. Interestingly, they are still capable of regular sub-structures as demonstrated through the creation of alpha helices stabilized by sterics rather than hydrogen bonding.⁷⁹⁻⁸¹ The structures are referred to as secondary structures even though they lack hydrogen bonds due to their repetitive and regular arrangement of atoms. These helices are stable to proteases allowing them to maintain biological activity under conditions that would normally denature a polypeptide.^{82, 83} By inserting different side chains, polypeptoids have been induced to form other secondary structures including a threaded loop⁸⁴⁻⁸⁶ and a β -sheet like structure.⁸⁷ In addition to secondary structures, peptoids can have functions similar to those of biological molecules. For example, Lee et al. have shown the capability of polypeptoids to mimic polypeptides by binding zinc within a multi-helical domain.⁸⁸ Polypeptoids also exhibit antimicrobial behavior in several different systems^{82, 89, 90}. These properties serve to demonstrate the utility of polypeptoids in mimicking not just the structure of biological polypeptides, but also the function.

Despite the studies described above regarding biological function of polypeptoids, their use for materials applications has been relatively limited. Only recently have researchers begun to examine designed peptoid sequences as building blocks for assembled materials. For example, Zuckermann et al. showed that two complementary binary sequences could form sheets of several nanometers thick⁹¹ and even that a single sequence containing 3 types of monomers could form the same sheet structures through ionic and hydrophobic interactions.^{92, 93} The Zhang group has demonstrated the synthesis of non-sequence-specific cyclic polypeptoids for solid state assembly applications.^{94, 95} In conclusion, the potential for polypeptoids as new materials and as a model system for developing self-assembly rules for sequence-specific polymers is large.

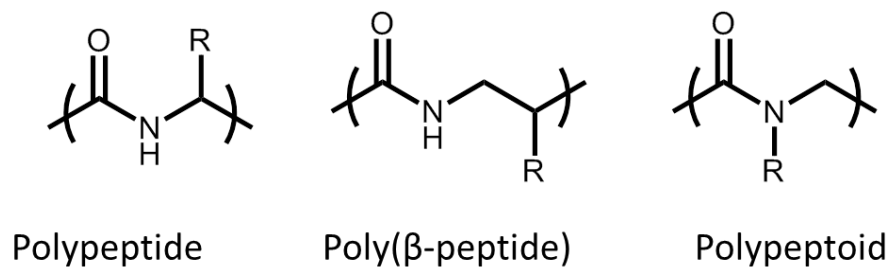


Figure 1.4. Chemical structures of several peptidomimetics

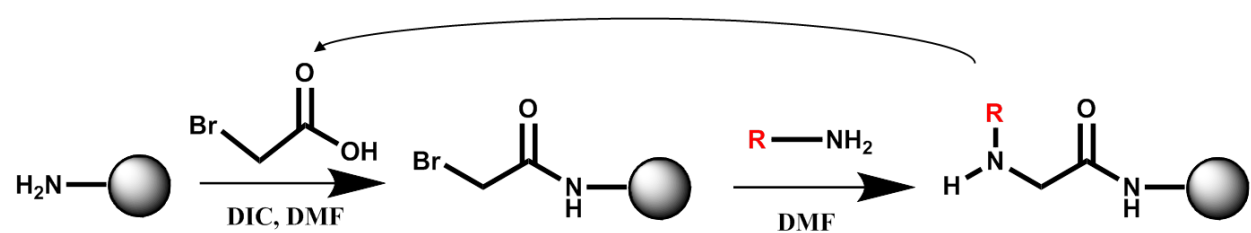


Figure 1.5. Two-step submonomer synthesis of polypeptoids

1.3 Motivation and Thesis Outline

Mimicking biological self-assembly holds promise for allowing the development of materials with complex functions that can be designed from the bottom up. In order to realize this potential, the forces that drive this type of self-assembly from sequence specific polymers must be understood far better than they currently are. Polypeptides themselves are inherently complex due to the many interactions that stem from the amino acids as well as the intrinsic backbone hydrogen bonding and chirality. A simplified model system such as poly N-substituted glycines or polypeptoids, will allow the isolation of various forces through side chain introductions.

This thesis will explore the self-assembly of polypeptoids across several length scales. Chapter 2 looks at the single chain conformation of a polypeptoid and how the persistence length can be controlled through the introduction of secondary structure. Chapter 3 explores this further and investigates the effect of charged side chains on the persistence length and chain conformation. In Chapter 4 the assembly of a single molecule into a collapsed globule is discussed. Finally, in Chapter 5 a novel supramolecular chiral structure assembled from a diblock copolypeptoid is presented along with extensive characterization through the use of chemical analogs and small angle x-ray scattering.

1.4. References

1. Dealy, J. M., Larson, R.G., *Structure and Rheology of Molten Polymers: From Structure to Flow Behavior and Back Again*. Hanser Verlag: 2006.
2. Young, R. J., Lovell, P.A., *Introduction Polymer*. CRC press: Boca Raton, 1981.
3. Flory, P. J., *Principles of Polymer Chemistry*. Cornell University Press: London, UK, 1953.
4. Kunst, E. D., Size, viscosity and precipitation of polymer molecules in solution. *Recl. Trav. Chim. Pays-Bas-J. Roy. Neth. Chem. Soc.* **1950**, 69 (1), 125-140.
5. Zhang, G. Z.; Wu, C., Folding and formation of mesoglobules in dilute copolymer solutions. In *Conformation-Dependent Design of Sequences in Copolymers I*, Khokhlov, A. R., Ed. Springer-Verlag Berlin: Berlin, 2006; Vol. 195, pp 101-176.
6. Wang, X. H.; Qiu, X. P.; Wu, C., Comparison of the coil-to-globule and the globule-to-coil transitions of a single poly(N-isopropylacrylamide) homopolymer chain in water. *Macromolecules* **1998**, 31 (9), 2972-2976.
7. Turner, K.; Zhu, P. W.; Napper, D. H., Coil-to-globule transitions of interfacial copolymers in better than theta-solvents. *Colloid Polym. Sci.* **1996**, 274 (7), 622-627.
8. Siu, M.; Zhang, G. Z.; Wu, C., Effect of comonomer distribution on the coil-to-globule transition of a single AB copolymer chain in dilute solution. *Macromolecules* **2002**, 35 (7), 2723-2727.
9. Zhu, P. W.; Napper, D. H., Effects of anionic surfactant on the coil-to-globule transition of interfacial poly(N-isopropylacrylamide). *Langmuir* **1996**, 12 (25), 5992-5998.
10. Rao, J. Y.; Xu, J.; Luo, S. Z.; Liu, S. Y., Cononsolvency-induced micellization of pyrene end-labeled diblock copolymers of N-isopropylacrylamide and oligo(ethylene glycol) methyl ether methacrylate. *Langmuir* **2007**, 23 (23), 11857-11865.
11. Rief, M.; Gautel, M.; Oesterhelt, F.; Fernandez, J. M.; Gaub, H. E., Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science* **1997**, 276 (5315), 1109-1112.
12. Ohashi, T.; Hale, C. A.; de Boer, P. A. J.; Erickson, H. P., Structural evidence that the P/Q domain of ZipA is an unstructured, flexible tether between the membrane and the C-terminal FtsZ-binding domain. *J. Bacteriol.* **2002**, 184 (15), 4313-4315.
13. Oberhauser, A. F.; Marszalek, P. E.; Erickson, H. P.; Fernandez, J. M., The molecular elasticity of the extracellular matrix protein tenascin. *Nature* **1998**, 393 (6681), 181-185.
14. Temyanko, E.; Russo, P. S.; Ricks, H., Study of rodlike homopolypeptides by gel permeation chromatography with light scattering detection: Validity of universal calibration and stiffness assessment. *Macromolecules* **2001**, 34 (3), 582-586.
15. Semisotnov, G. V.; Kihara, H.; Kotova, N. V.; Kimura, K.; Amemiya, Y.; Wakabayashi, K.; Serdyuk, I. N.; Timchenko, A. A.; Chiba, K.; Nikaido, K.; Ikura, T.; Kuwajima, K., Protein globularization during folding. A study by synchrotron small-angle X-ray scattering. *J. Mol. Biol.* **1996**, 262 (4), 559-574.
16. Pollack, L.; Tate, M. W.; Darnton, N. C.; Knight, J. B.; Gruner, S. M.; Eaton, W. A.; Austin, R. H., Compactness of the denatured state of a fast-folding protein measured by

submillisecond small-angle x-ray scattering. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (18), 10115-10117.

17. Taylor, W. R.; Orengo, C. A., A holistic approach to protein-structure alignment. *Protein Eng.* **1989**, *2* (7), 505-519.

18. Sander, C.; Schneider, R., Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins* **1991**, *9* (1), 56-68.

19. Vriend, G.; Sander, C., Detection of common 3-dimensional substructures in proteins. *Proteins-Structure Function and Bioinformatics* **1991**, *11* (1), 52-58.

20. Rost, B.; Sander, C., Prediction of protein secondary structure at better than 70-percent accuracy. *J. Mol. Biol.* **1993**, *232* (2), 584-599.

21. Harbury, P. B.; Zhang, T.; Kim, P. S.; Alber, T., A switch between 2-stranded, 3-stranded and 4-stranded coiled coils in GCN4 leucine-zipper mutants. *Science* **1993**, *262* (5138), 1401-1407.

22. Munson, M.; O'Brien, R.; Sturtevant, J. M.; Regan, L., Redesigning the hydrophobic core of a 4-helix-bundle protein. *Protein Sci.* **1994**, *3* (11), 2015-2022.

23. Munson, M.; Balasubramanian, S.; Fleming, K. G.; Nagi, A. D.; O'Brien, R.; Sturtevant, J. M.; Regan, L., What makes a protein a protein? Hydrophobic core designs that specify stability and structural properties. *Protein Sci.* **1996**, *5* (8), 1584-1593.

24. Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M., CHARMM - A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **1983**, *4* (2), 187-217.

25. Rohl, C. A.; Strauss, C. E. M.; Misura, K. M. S.; Baker, D., Protein structure prediction using Rosetta. *Numerical Computer Methods, Pt D* **2004**, *383*, 66-+.

26. Das, R.; Baker, D., Macromolecular modeling with Rosetta. In *Annual Review of Biochemistry*, Annual Reviews: Palo Alto, 2008; Vol. 77, pp 363-382.

27. Khatib, F.; Cooper, S.; Tyka, M. D.; Xu, K. F.; Makedon, I.; Popovic, Z.; Baker, D.; Foldit, P., Algorithm discovery by protein folding game players. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (47), 18949-18953.

28. Butterfoss, G. L.; Kuhlman, B., Computer-based design of novel protein structures. In *Annual Review of Biophysics and Biomolecular Structure*, Annual Reviews: Palo Alto, 2006; Vol. 35, pp 49-65.

29. Chikenji, G.; Fujitsuka, Y.; Takada, S., Shaping up the protein folding funnel by local interaction: Lesson from a structure prediction study. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (9), 3141-3146.

30. Potapov, V.; Cohen, M.; Schreiber, G., Assessing computational methods for predicting protein stability upon mutation: good on average but not in the details. *Protein Eng. Des. Sel.* **2009**, *22* (9), 553-560.

31. Dahiyat, B. I.; Mayo, S. L., Protein design automation. *Protein Sci.* **1996**, *5* (5), 895-903.

32. Dahiyat, B. I.; Sarisky, C. A.; Mayo, S. L., De novo protein design: Towards fully automated sequence selection. *J. Mol. Biol.* **1997**, *273* (4), 789-796.

33. Dahiyat, B. I.; Mayo, S. L., De novo protein design: Fully automated sequence selection. *Science* **1997**, *278* (5335), 82-87.
34. Kuhlman, B.; Dantas, G.; Ireton, G. C.; Varani, G.; Stoddard, B. L.; Baker, D., Design of a novel globular protein fold with atomic-level accuracy. *Science* **2003**, *302* (5649), 1364-1368.
35. Harbury, P. B.; Plecs, J. J.; Tidor, B.; Alber, T.; Kim, P. S., High-resolution protein design with backbone freedom. *Science* **1998**, *282* (5393), 1462-1467.
36. RamirezAlvarado, M.; Blanco, F. J.; Serrano, L., De novo design and structural analysis of a model beta-hairpin peptide system. *Nat. Struct. Biol.* **1996**, *3* (7), 604-612.
37. Zhang, L. F.; Eisenberg, A., Structures of "crew-cut" aggregates of polystyrene-b-poly(acrylic acid) diblock copolymers. *Macromol. Symp.* **1997**, *113*, 221-232.
38. Zhang, L. F.; Eisenberg, A., Multiple morphologies of crew-cut aggregates of polystyrene-b-poly(acrylic acid) block-copolymers *Science* **1995**, *268* (5218), 1728-1731.
39. Luo, L. B.; Eisenberg, A., Thermodynamic size control of block copolymer vesicles in solution. *Langmuir* **2001**, *17* (22), 6804-6811.
40. Zhang, L. F.; Barlow, R. J.; Eisenberg, A., Scaling relations and coronal dimensions in aqueous block polyelectrolyte micelles. *Macromolecules* **1995**, *28* (18), 6055-6066.
41. Zhang, L. F.; Eisenberg, A., Multiple morphologies and characteristics of "crew-cut" micelle-like aggregates of polystyrene-b-poly(acrylic acid) diblock copolymers in aqueous solutions. *J. Am. Chem. Soc.* **1996**, *118* (13), 3168-3181.
42. Israelachvili, J. N.; Mitchell, D. J.; Ninham, B. W., Theory of self-assembly of hydrocarbon amphiphiles into micelles and bilayers. *Journal of the Chemical Society-Faraday Transactions II* **1976**, *72*, 1525-1568.
43. Eanes, E. D.; Glenner, G. G., X-ray diffraction studies on amyloid filaments *J. Histochem. Cytochem.* **1968**, *16* (11), 673-&.
44. Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T.; Madsen, A. O.; Riek, C.; Eisenberg, D., Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature* **2007**, *447* (7143), 453-457.
45. Balbach, J. J.; Ishii, Y.; Antzutkin, O. N.; Leapman, R. D.; Rizzo, N. W.; Dyda, F.; Reed, J.; Tycko, R., Amyloid fibril formation by A beta(16-22), a seven-residue fragment of the Alzheimer's beta-amyloid peptide, and structural characterization by solid state NMR. *Biochemistry* **2000**, *39* (45), 13748-13759.
46. Burdick, D.; Soreghan, B.; Kwon, M.; Kosmoski, J.; Knauer, M.; Henschen, A.; Yates, J.; Cotman, C.; Glabe, C., Assembly and aggregation properties of synthetic Alzheimers A4/Beta amyloid peptide analogs. *J. Biol. Chem.* **1992**, *267* (1), 546-554.
47. Westermark, P.; Engstrom, U.; Johnson, K. H.; Westermark, G. T.; Betsholtz, C., Islet amyloid polypeptide - pinpointing amino-acid-residues linked to amyloid fibril formation. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87* (13), 5036-5040.

48. Ballister, E. R.; Lai, A. H.; Zuckermann, R. N.; Cheng, Y.; Mougous, J. D., In vitro self-assembly from a simple protein of tailorable nanotubes building block. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (10), 3733-3738.
49. Lamm, M. S.; Rajagopal, K.; Schneider, J. P.; Pochan, D. J., Laminated morphology of nontwisting beta-sheet fibrils constructed via peptide self-assembly. *J. Am. Chem. Soc.* **2005**, *127* (47), 16692-16700.
50. Castelletto, V.; Hamley, I. W.; Hule, R. A.; Pochan, D., Helical-Ribbon Formation by a beta-Amino Acid Modified Amyloid beta-Peptide Fragment. *Angew. Chem.-Int. Edit.* **2009**, *48* (13), 2317-2320.
51. Zhang, S. G.; Marini, D. M.; Hwang, W.; Santoso, S., Design of nanostructured biological materials through self-assembly of peptides and proteins. *Curr. Opin. Chem. Biol.* **2002**, *6* (6), 865-871.
52. Woods, R. D.; Takahashi, N.; Aslam, A.; Pleass, R. J.; Aizawa, S. I.; Sockett, R. E., Bifunctional nanotube scaffolds for diverse ligands are purified simply from Escherichia coli strains coexpressing two functionalized flagellar genes. *Nano Lett.* **2007**, *7* (6), 1809-1816.
53. Sharma, N.; Top, A.; Kiick, K. L.; Pochan, D. J., One-Dimensional Gold Nanoparticle Arrays by Electrostatically Directed Organization Using Polypeptide Self-Assembly. *Angew. Chem.-Int. Edit.* **2009**, *48* (38), 7078-7082.
54. Wright, E. R.; Conticello, V. P., Self-assembly of block copolymers derived from elastin-mimetic polypeptide sequences. *Adv. Drug Deliv. Rev.* **2002**, *54* (8), 1057-1073.
55. Douy, A.; Gallot, B., Block copolymers with a polyvinyl and a polypeptide block - Factors governing the folding of the polypeptide chains. *Polymer* **1982**, *23* (7), 1039-1044.
56. Douy, A.; Gallot, B., Structure of AB block copolymers with a polypeptide block - Effect of chain conformation. *Polym. Eng. Sci.* **1977**, *17* (8), 523-526.
57. Billot, J. P.; Douy, A.; Gallot, B., Synthesis and structural study of block copolymers with a hydrophobic polyvinyl block and a hydrophilic polypeptide block - Copolymers polystyrene-poly(L-lysine) and polybutadiene-poly(L-lysine). *Makromolekulare Chemie-Macromolecular Chemistry and Physics* **1976**, *177* (6), 1889-1893.
58. Douy, A.; Gallot, B., Amphipathic block copolymers with 2 Polypeptide blocks - Synthesis and structural study of poly(N-epsilon-trifluoroacetyl-L-L-Lysine) polysarcosine copolymers *Polymer* **1987**, *28* (1), 147-154.
59. Deming, T. J., Polypeptide materials: New synthetic methods and applications. *Adv. Mater.* **1997**, *9* (4), 299-&.
60. Deming, T. J., Facile synthesis of block copolypeptides of defined architecture. *Nature* **1997**, *390* (6658), 386-389.
61. Bellomo, E. G.; Wyrsta, M. D.; Pakstis, L.; Pochan, D. J.; Deming, T. J., Stimuli-responsive polypeptide vesicles by conformation-specific assembly. *Nat. Mater.* **2004**, *3* (4), 244-248.
62. Holowka, E. P.; Pochan, D. J.; Deming, T. J., Charged polypeptide vesicles with controllable diameter. *J. Am. Chem. Soc.* **2005**, *127* (35), 12423-12428.

63. Nowak, A. P.; Breedveld, V.; Pakstis, L.; Ozbas, B.; Pine, D. J.; Pochan, D.; Deming, T. J., Rapidly recovering hydrogel scaffolds from self-assembling diblock copolypeptide amphiphiles. *Nature* **2002**, *417* (6887), 424-428.
64. Li, Z. B.; Deming, T. J., Tunable hydrogel morphology via self-assembly of amphiphilic pentablock copolypeptides. *Soft Matter* **2010**, *6* (11), 2546-2551.
65. Hanson, J. A.; Chang, C. B.; Graves, S. M.; Li, Z. B.; Mason, T. G.; Deming, T. J., Nanoscale double emulsions stabilized by single-component block copolypeptides. *Nature* **2008**, *455* (7209), 85-U54.
66. Ueda, M.; Makino, A.; Imai, T.; Sugiyama, J.; Kimura, S., Rational design of peptide nanotubes for varying diameters and lengths. *J. Pept. Sci.* **2011**, *17* (2), 94-99.
67. Kim, W.; Thevenot, J.; Ibarboure, E.; Lecommandoux, S.; Chaikof, E. L., Self-Assembly of Thermally Responsive Amphiphilic Diblock Copolypeptides into Spherical Micellar Nanoparticles. *Angew. Chem.-Int. Edit.* **2010**, *49* (25), 4257-4260.
68. Gibson, M. I.; Cameron, N. R., Organogelation of sheet-helix diblock copolypeptides. *Angew. Chem.-Int. Edit.* **2008**, *47* (28), 5160-5162.
69. Lau, K. F.; Dill, K. A., A lattice statistical-mechanics model of the conformational and sequence-spaces of proteins. *Macromolecules* **1989**, *22* (10), 3986-3997.
70. Seebach, D.; Matthews, J. L., beta-peptides: a surprise at every turn. *Chem. Commun.* **1997**, (21), 2015-2022.
71. Seebach, D.; Abele, S.; Gademann, K.; Jaun, B., Pleated sheets and turns of beta-peptides with proteinogenic side chains. *Angew. Chem.-Int. Edit.* **1999**, *38* (11), 1595-1597.
72. Yang, D.; Qu, J.; Li, B.; Ng, F. F.; Wang, X. C.; Cheung, K. K.; Wang, D. P.; Wu, Y. D., Novel turns and helices in peptides of chiral alpha-aminoxy acids. *J. Am. Chem. Soc.* **1999**, *121* (3), 589-590.
73. Li, X.; Yang, D., Peptides of aminoxy acids as foldamers. *Chem. Commun.* **2006**, (32), 3367-3379.
74. Seebach, D.; Beck, A. K.; Bierbaum, D. J., The world of beta- and gamma-peptides comprised of homologated proteinogenic amino acids and other components. *Chem. Biodivers.* **2004**, *1* (8), 1111-1239.
75. Seebach, D.; Hook, D. F.; Glattli, A., Helices and other secondary structures of beta- and gamma-peptides. *Biopolymers* **2006**, *84* (1), 23-37.
76. Bouillere, F.; Thetiot-Laurent, S.; Kouklovsky, C.; Alezra, V., Foldamers containing gamma-amino acid residues or their analogues: structural features and applications. *Amino Acids* **2011**, *41* (3), 687-707.
77. Gante, J., Azapeptides. *Synthesis* **1989**, (6), 405-413.
78. Figliozzi, G. M.; Goldsmith, R.; Ng, S. C.; Banville, S. C.; Zuckermann, R. N., Synthesis of N-substituted glycine peptoid libraries. *Methods Enzymol.* **1996**, *267*, 437-447.
79. Armand, P.; Kirshenbaum, K.; Goldsmith, R. A.; Farr-Jones, S.; Barron, A. E.; Truong, K. T. V.; Dill, K. A.; Mierke, D. F.; Cohen, F. E.; Zuckermann, R. N.; Bradley, E. K., NMR

- determination of the major solution conformation of a peptoid pentamer with chiral side chains. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (8), 4309-4314.
80. Wu, C. W.; Sanborn, T. J.; Huang, K.; Zuckermann, R. N.; Barron, A. E., Peptoid oligomers with alpha-chiral, aromatic side chains: Sequence requirements for the formation of stable peptoid helices. *J. Am. Chem. Soc.* **2001**, *123* (28), 6778-6784.
81. Wu, C. W.; Sanborn, T. J.; Zuckermann, R. N.; Barron, A. E., Peptoid oligomers with alpha-chiral, aromatic side chains: Effects of chain length on secondary structure. *J. Am. Chem. Soc.* **2001**, *123* (13), 2958-2963.
82. Statz, A. R.; Park, J. P.; Chongsiriwatana, N. P.; Barron, A. E.; Messersmith, P. B., Surface-immobilised antimicrobial peptoids. *Biofouling* **2008**, *24* (6), 439-448.
83. Chongsiriwatana, N. P.; Patch, J. A.; Czyzewski, A. M.; Dohm, M. T.; Ivankin, A.; Gidalevitz, D.; Zuckermann, R. N.; Barron, A. E., Peptoids that mimic the structure, function, and mechanism of helical antimicrobial peptides. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (8), 2794-2799.
84. Huang, K.; Wu, C. W.; Sanborn, T. J.; Patch, J. A.; Kirshenbaum, K.; Zuckermann, R. N.; Barron, A. E.; Radhakrishnan, I., A threaded loop conformation adopted by a family of peptoid nonamers. *J. Am. Chem. Soc.* **2006**, *128* (5), 1733-1738.
85. Yoo, B.; Kirshenbaum, K., Peptoid architectures: elaboration, actuation, and application. *Curr. Opin. Chem. Biol.* **2008**, *12* (6), 714-721.
86. Shin, S. B. Y.; Yoo, B.; Todaro, L. J.; Kirshenbaum, K., Cyclic peptoids. *J. Am. Chem. Soc.* **2007**, *129* (11), 3218-3225.
87. Crapster, J. A.; Stringer, J. R.; Guzei, I. A.; Blackwell, H. E., Design and Conformational Analysis of Peptoids Containing N-Hydroxy Amides Reveals a Unique Sheet-Like Secondary Structure. *Biopolymers* **2011**, *96* (5), 604-616.
88. Lee, B. C.; Chu, T. K.; Dill, K. A.; Zuckermann, R. N., Biomimetic nanostructures: Creating a high-affinity zinc-binding site in a folded nonbiological polymer. *J. Am. Chem. Soc.* **2008**, *130* (27), 8847-8855.
89. Patch, J. A.; Barron, A. E., Mimicry of bioactive peptides via non-natural, sequence-specific peptidomimetic oligomers. *Curr. Opin. Chem. Biol.* **2002**, *6* (6), 872-877.
90. Fowler, S. A.; Blackwell, H. E., Structure-function relationships in peptoids: Recent advances toward deciphering the structural requirements for biological function. *Org. Biomol. Chem.* **2009**, *7* (8), 1508-1524.
91. Nam, K. T.; Shelby, S. A.; Choi, P. H.; Marciel, A. B.; Chen, R.; Tan, L.; Chu, T. K.; Mesch, R. A.; Lee, B. C.; Connolly, M. D.; Kisielowski, C.; Zuckermann, R. N., Free-floating ultrathin two-dimensional crystals from sequence-specific peptoid polymers. *Nat. Mater.* **2010**, *9* (5), 454-460.
92. Sanii, B.; Kudirka, R.; Cho, A.; Venkateswaran, N.; Olivier, G. K.; Olson, A. M.; Tran, H.; Harada, R. M.; Tan, L.; Zuckermann, R. N., Shaken, Not Stirred: Collapsing a Peptoid Monolayer To Produce Free-Floating, Stable Nanosheets. *J. Am. Chem. Soc.* **2011**, *133* (51), 20808-20815.

93. Kudirka, R.; Tran, H.; Sanii, B.; Nam, K. T.; Choi, P. H.; Venkateswaran, N.; Chen, R.; Whitelam, S.; Zuckermann, R. N., Folding of a Single-Chain, Information-Rich Polypeptoid Sequence into a Highly Ordered Nanosheet. *Biopolymers* **2011**, *96* (5), 586-595.
94. Guo, L.; Zhang, D. H., Cyclic Poly(alpha-peptoid)s and Their Block Copolymers from N-Heterocyclic Carbene-Mediated Ring-Opening Polymerizations of N-Substituted N-Carboxyanhydrides. *J. Am. Chem. Soc.* **2009**, *131* (50), 18072-+.
95. Guo, L.; Li, J. H.; Brown, Z.; Ghale, K.; Zhang, D. H., Synthesis and Characterization of Cyclic and Linear Helical Poly(alpha-peptoid)s by N-Heterocyclic Carbene-Mediated Ring-Opening Polymerizations of N-Substituted N-Carboxyanhydrides. *Biopolymers* **2011**, *96* (5), 596-603.

Chapter 2. Determination of the persistence length of helical and non-helical polypeptoids in solution

Reproduced with permission from Adrienne Rosales, Steven Kline, Ronald Zuckermann, and Rachel Segalman Soft Matter (2012) **8**, 3673-3680, *Reproduced by permission of The Royal Society of Chemistry* (<http://pubs.rsc.org/en/content/articlelanding/2012/sm/c2sm07092h>)

2.1. Introduction

The shape of a polymer chain is a reflection of its intramolecular interactions and can directly influence a large number of characteristics, including mechanical properties and self-assembly behavior. These intramolecular interactions include sterics, hydrogen bonding, hydrophobic interactions, and aromatic interactions, among others. The chain conformation in turn affects how the polymer can interact with the other chains around it, which influences both the mechanical properties of the polymer and its self-assembly into various structures. Classical polymers self-assemble via a balance of enthalpic interactions and entropic chain stretching penalties, which can both be complicated by the conformation of the polymer chain. In polymers with more complicated chemical interactions, such as polyelectrolytes and conjugated polymers, chains are often significantly stiffer (have a longer persistence length) than classical materials. The polymer backbone itself plays a large role in intramolecular interactions leading to chain shape, and polymers with highly conjugated or sterically hindered backbones, such as polyphenylene vinylene, have longer persistence lengths in solution (6-40 nm)¹ than those with backbones composed of more aliphatic linkages, such as polystyrene (1 nm).² Side chains can also play a significant role in the rigidity of a polymer chain. For example, a subtle difference of one carbon in a polysilylene side chain can increase the persistence length from 6.2 nm up to 85 nm.³ Charged side chains introduce another level of complexity due to the ionic interactions between groups that can lead to chain stiffening.⁴

Secondary structure can also have a large impact on the persistence length of a chain. Helical secondary structure in particular correlates with increasing persistence lengths in polymers. Helical polyisocyanates⁵ have a persistence length of 40 nm to 50 nm, and α -helical poly(γ -benzyl-L-glutamate)⁶ (PBLG) has a persistence length up to approximately 200 nm at high molecular weights. Several design methods exist for the formation of a polymer helix; in particular, designing side chain interactions is an interesting route to controlling chain shape. Side chain interactions have a large influence on the formation of secondary structures in biological polymers and thus directly influence the persistence length of these polymers as well. For example, the addition of long stretches of prolines in a polypeptide induces the formation of a helix. Using FRET experiments, the persistence length of a polyproline type II helix was estimated to range from 9 nm to 13 nm⁷ using a chain that was 20 monomers long in the *all-trans* form.

Here, we evaluate the effect of subtle changes in side chain size and chirality on the persistence length of N-substituted glycine polymers, also known as polypeptoids. Although polypeptoids are a relatively new material, they have recently attracted much attention in polymeric studies,⁸⁻¹¹ meaning there is a need to quantify their properties. Quantifying polymer properties, such as flexibility and persistence length, is important for modeling these systems and further understanding their self-assembly.¹²

Polypeptoids are sequence-specific, biomimetic polymers that have been shown to form stable secondary structures in solution depending upon the types of side chains incorporated into the polymer^{9, 13-15}. In particular, polypeptoids with bulky α -chiral side chains form a polyproline type I-like helix in solution. Unlike the α -helices of polypeptides, the polypeptoid helices are stabilized by steric interactions, rather than hydrogen bonding. Both theoretical¹⁶ and experimental¹⁷⁻²⁰ studies have shown that the preferred conformation of the peptoid helix is entirely composed of *cis*-amide bonds with a periodicity of three residues per turn and a pitch of approximately 6 Å. These studies showed that a polypeptoid with α -chiral aromatic side chains prefers the all *cis* conformation to a *trans* conformation in the ratio of 2:1. In addition, the handedness of the helix is determined by the handedness of the α -chiral side chains, as the peptoid backbone is devoid of chirality. The presence of a helical fold would lead one to expect a stiffening of the polymer chain. However, our small angle neutron scattering (SANS) measurements have shown that helical polypeptoids have persistence lengths much smaller than expected. In fact, they are nearly as flexible as polypeptoids without any secondary structure.

2.2. Experimental Section

Materials

Compounds **1** and **2** (Figure 2.1) were synthesized using a step-wise solid-phase submonomer synthesis method²¹ on a custom robotic synthesizer or a commercial Aapptec Apex 396 robotic synthesizer. All polypeptoids were acetylated on the resin, cleaved from the resin using 95% v/v trifluoroacetic acid in water, and purified using reverse-phase HPLC, as previously described.^{8, 10} The synthesis was confirmed by electrospray mass spectrometry on an Agilent 1100 series LC/MSD trap system (Agilent Technologies, Santa Clara, CA) and by matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) mass spectrometry on a 4800 series MALDI-TOF (Applied Biosystems, Carlsbad, CA) with a laser power of 5000. MALDI samples were prepared using a 1:1 ratio of polypeptoid in acetonitrile (1 mg/mL) and 1,8,9-anthracenetriol (10 mg/mL in tetrahydrofuran). The monomer sequences for the polypeptoids studied here are denoted in Figure 2.1, with *Nme* = *N*-(2-methoxyethyl)glycine, *Npe* = *N*-(2-phenylethyl)glycine, and *Nrpe* = (*R*)-*N*-(1-phenylethyl)glycine. Several polymers were made with *n* (the number of repeat units as designated in Figure 2.1) varying from 3 to 8. The majority of this publication will discuss the polymers where *n* is equal to 6, forming a polypeptoid with 36 total monomers.

Circular Dichroism (CD)

CD measurements were performed on a J-185 CD spectrometer (Jasco Inc., Easton, MD). Stock solutions of the polypeptoids were made in tared vials using 5 mg/mL of peptoid powder in acetonitrile. The stock solutions were then diluted to a concentration of 0.08 mg/mL before acquiring CD spectra. CD spectra were acquired using a quartz cell (Hellma USA, Plainview, NY) with a path length of 1 mm. A scan rate of 50 nm/min was used, and 3 measurements were averaged for each compound.

Small Angle Neutron Scattering (SANS)

SANS studies were conducted at the NG3 SANS line at the National Institute of Standards and Technology (NIST) Center for Neutron Research in Gaithersburg, Maryland and at the CG-3

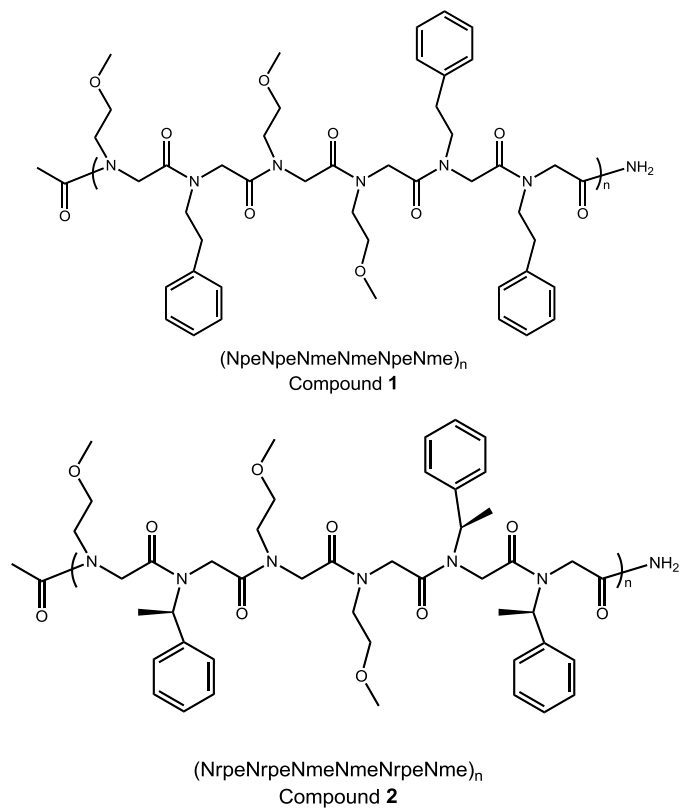
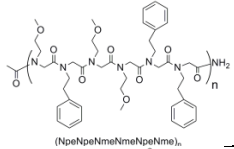
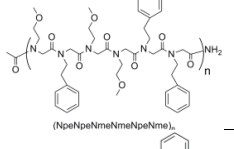
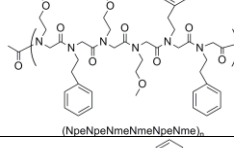
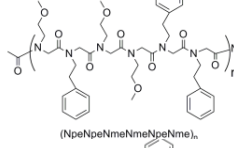
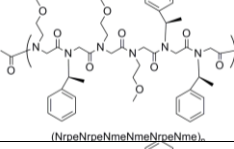
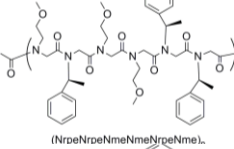
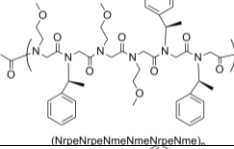
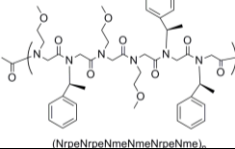


Figure 2.1. Chemical structures of the helical and nonhelical compound.

1 contains achiral aromatic side chains, while 2, a helix-forming polypeptoid, contains alpha-chiral side chains. The value of n ranges from 3 to 8 for chains of varying lengths.

Table 2.1. All the compounds used in this study

Compound	Repeat Unit	n (number of repeat units)	N (total number of monomers)	M_{obs}/M_{theo}	Structure
Compound 1a		3	18	2548.5 / 2546	Non-helical
Compound 1b		4	24	3379.4 / 3375	Non-helical
Compound 1c		6	36	5033.2 / 5033	Non-helical
Compound 1d		8	48	6701.6 / 6691.1	Non-helical
Compound 2a		3	18	2549.2 / 2546	Helical
Compound 2b		4	24	3377.5 / 3375	Helical
Compound 2c		6	36	5028.2 / 5033	Helical
Compound 2d		8	48	6604.4 / 6691.1	Helical

Bio-SANS line at the High Flux Isotope Reactor at Oak Ridge National Laboratory (ORNL) in Oak Ridge, Tennessee. Samples were prepared at a concentration of 10 mg/mL in deuterated acetonitrile to enhance the contrast between the polypeptoids and the solvent. Quartz banjo cells (Hellma USA, Plainview, NY) with a path length of 1 mm and 2 mm were used at NIST and ORNL, respectively, in a temperature controlled multiple position sample holder. A neutron wavelength of 6 Å was used at both beamlines, and data were collected at two different instrument configurations (1.3 m and 4 m at NIST, and 1.7 m and 14.5 m at ORNL). The data were reduced using the NCNR SANS reduction macros²² and the Spice SANS reduction program in Igor Pro.

2.3. Results and Discussion

2.3.1. Circular Dichroism

Compounds **1** and **2** were designed to be non-structured and helical, respectively, through the incorporation of aromatic side chains with tunable chirality. Compound **2** contains 50% α -chiral aromatic side chains while compound **1** contains 50% achiral aromatic side chains instead. Previous literature has shown that **2** forms a polyproline type I-like helix in solution with all *cis* amide bonds¹⁴ as described above. However, **1** was designed to be minimally structured by removing the α -chiral substituent that provides the steric influence for secondary structure formation. Indeed, circular dichroism (Figure 2.2) in acetonitrile shows that there is helix formation of **2** as demonstrated by the characteristic peaks at 192 nm, 202 nm, and 218 nm. This helix formation is constant across several polymers of varying molecular weights (**2a** through **2d**) with little deviation in the per-residue molar ellipticity. Previously, it was shown that as the chain length of a polypeptoid containing 100% N_{rpe} residues increases, the per-residue molar ellipticity also increases until a chain length of 13 residues is reached. After 13 residues, the ellipticity remains approximately constant for longer chain lengths, suggesting that the overall fraction of helical isomers is stabilized.²⁰ A similar trend was also observed for peptoid helices consisting of bulky *N*-1-naphthylethyl side chains, except in that case, the per-residue molar ellipticity reached a maximum after only 5 monomer units.¹⁷

It would be helpful to gain some quantitative insight about the helical content of these molecules using CD, as is often done for proteins. However, CD is a quantitative technique for proteins because there are a vast number of known protein structures, allowing the development of algorithms that can compute a reliable estimate of the fraction of α -helices, β -sheets, and random coils by comparing new CD data to that of the known structures²³. There are few X-ray solved structures of polypeptoids, meaning that it is not possible to reliably calculate the fraction of helicity for polypeptoids from CD. Because CD does not yield information about the population of conformers for polypeptoids, polypeptoid secondary structure has previously been established using a combination of 2D NMR, X-ray crystallography, and molecular modeling studies. These studies first confirmed the presence of a helical conformation in a very short polypeptoid (5 monomers in length) containing bulky chiral, aromatic residues ((*S*)-*N*-(1-phenylethyl)glycine, N_{spe}).¹⁶ However, the 2D NMR studies for N_{spe}₅ also show the presence of other isomers and conformations in the amount of approximately 40% in methanol solution.^{17, 18} Thus, the α -helix-like CD signature observed for (N_{spe})₅ and other peptoid helices is from an ensemble of closely related conformations in rapid equilibrium with one another.

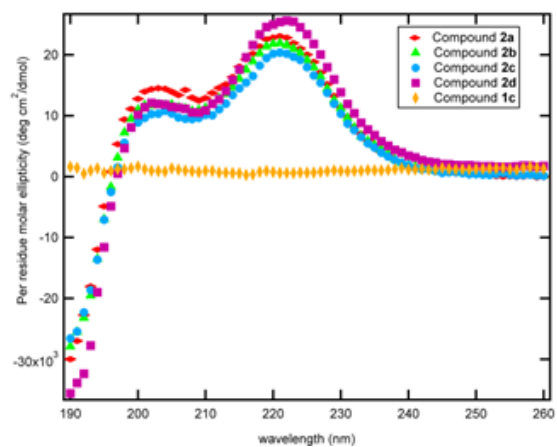


Figure 2.2. CD spectra for different chain lengths of a helical polypeptoid as well as a nonhelical polypeptoid

The presence of these other polypeptoid conformers is most likely due to the *cis/trans* isomerization of the backbone amide bonds, which may enable the polypeptoid backbone to sample many conformations. To probe whether the fraction of helices can be controlled, the effect of temperature and solvent on the per-residue molar ellipticity was examined. As **2c** is heated from 20°C to 70°C in acetonitrile (2.3a), there is no change in the spectrum shape and only a slight decrease in the peak intensity at 218 nm from 19,200 degcm²dmol⁻¹ to 15,300 degcm²dmol⁻¹. This result is consistent with the observation by Wu et. al that Nrpe₆, Nrpe₁₂, and Nrpe₁₈ all retain their helical CD signature at increased temperature, suggesting that the peptoid helices are stable to thermal unfolding because they are sterically constrained rather than hydrogen bond-stabilized.²⁰ This is the case for all of the helical polypeptoids investigated here, as shown in the Supporting Information. In addition, changing the solvent from acetonitrile to methanol has a minimal effect on the spectrum shape and intensity at increased temperature (2.3b). This result is also consistent with previous studies: Armand et. al previously observed a peptoid helix in methanol solution using 2D NMR. These experiments indicate that the peptoid helices are stable to both temperature and solvent.

All of the achiral polypeptoids studied here (**1a** to **1d**) show no net ellipticity because they do not contain chiral residues. Hence, little information about their structure can be gained using CD. Small angle neutron scattering (SANS) is therefore used to probe the difference in chain statistics for these two series of molecules.

2.3.2. Small Angle Neutron Scattering (SANS)

Because **2** forms a helix in solution, it was anticipated that its chain would be stiffer than the corresponding analog **1**; however, the SANS experiments detailed here show that the difference in chain stiffness is not as large as expected. Plotting the scattering intensity, *I*, versus the scattering angle, *q*, (Figure 2.3a) for the two 36-mer compounds yielded several insights about the polypeptoid chains. First, the two sequences both have typical scattering patterns for a single chain in dilute solution. It is expected that the intensity should scale with *q* as -2 for a random coil conformation and scale as -1 for a rod-like conformation. For a single chain, one should see the change in scaling behavior provided the appropriate *q*-range. For both polypeptoids, there was an exponential decrease over the *q*-range from 0.07 Å⁻¹ to 0.22 Å⁻¹ with a scaling of approximately -1.5. Around *q* ≈ 0.22 Å⁻¹, the intensity scales as 0.6 - 0.8. The deviation from -2 scaling indicates that the polypeptoids are not forming Gaussian coils in solution, while the deviation from -1 scaling is most likely a result of noise in the data at high *q*.

To see the change in scaling better, a Kratky plot was used (Figure 2.4b), which plots *Iq*² vs. *q*. This has the effect of making the intensity data in the Gaussian regime tend toward a horizontal asymptote. The *q*-value at which the data deviate from this asymptote and begin to increase linearly (with an intercept coinciding with the origin) is inversely related to the persistence length. Because the intensity does not quite scale as -2 with *q*, it is difficult to pinpoint the exact transition. However, a good approximation was made by selecting the point at which *Iq*² deviates from a straight line passing through the origin (the red line in Figure 2.4b). A straight line passing through the origin on a Kratky plot corresponds to the scattering function for a rod; thus, the departure from this behavior indicates scattering from a molecule above its persistence length.

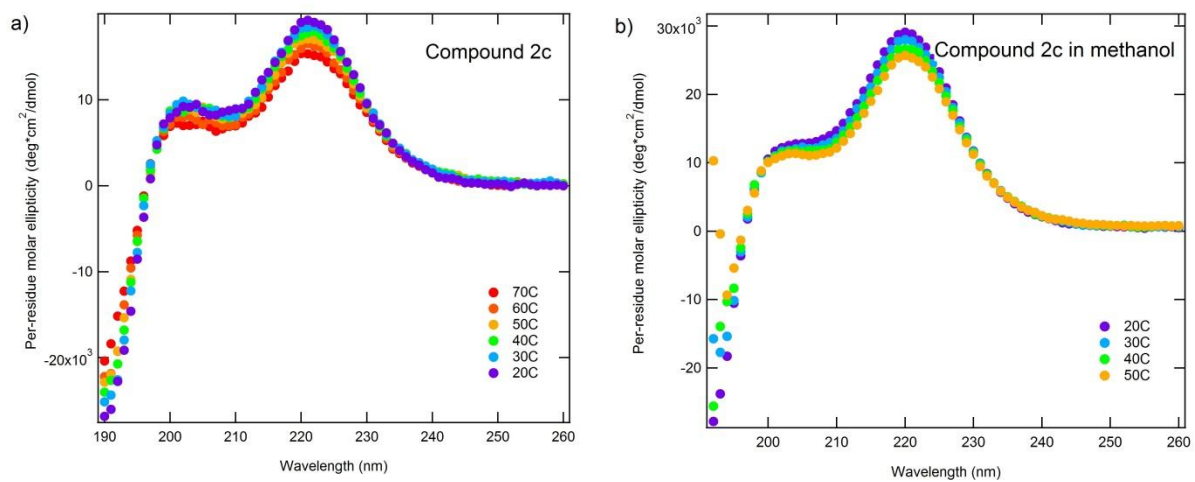


Figure 2.3. Heating (a) and solvent (b) do not significantly affect the CD spectra of 2c.

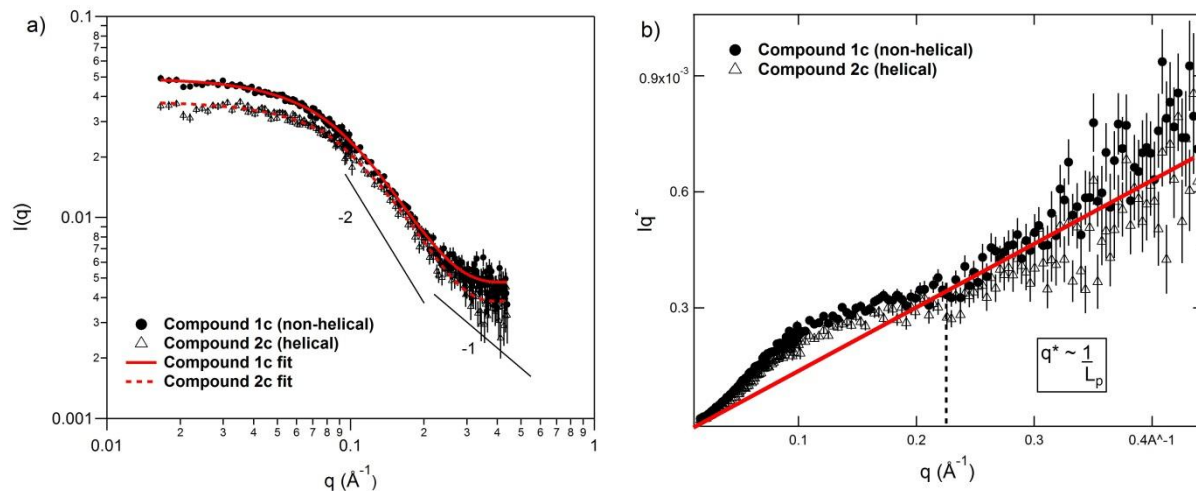


Figure 2.4. Small angle neutron scattering (SANS) shows the relatively short persistence length of both polypeptoids.

The model fits are shown in (a), demonstrating good fits for the data. In addition, lines with scalings of -1 (rigid rod) and -2 (Gaussian coil) have been added as references. The intensity presented here is absolute intensity in units of cm^{-1} . The Kratky plot in (b) emphasizes the change in scaling behavior. Incoherent background has been subtracted from the data before the fits.

The approximate q -value for this transition is marked by the dashed line in Figure 2.4b, and it is clear that it is quite similar for both sequences. To determine the persistence length, the equation $L_p = \frac{k}{q^*}$ was used, where k is a proportionality constant and q^* is the q -value for the transition in the Kratky plot. The value of the constant k given in literature is $6/\pi$ by Kratky²⁴ and Koyama²⁵, or 2.87, according to Burchard and Kajiwara²⁶. The proportionality constants available indicate that the persistence length for both **1** and **2** is on the order of 0.8 nm to 1.3 nm. Thus, both **1** and **2** are more flexible than other helical polymers, including helical polypeptides. In fact, these polypeptoids have a flexibility very similar to that of polystyrene² (~ 1 nm).

Semi-flexible cylinder model

Because the determination of the inflection point on the Kratky plot can be somewhat subjective, a more precise determination of the persistence length can be obtained by modeling and comparison to experimental data. The NIST NCNR analysis macro has been used to model these molecules as semi-flexible cylinders with excluded volume,²⁷ according to the Kratky and Porod model of a wormlike chain. In this model, the cylinder of the chain is assumed to be composed of a series of connected locally stiff chain segments. The length of these segments is called the Kuhn length and is calculated by holding the scattering length densities and the contour length constant and fitting a Kuhn length and a radius to the semi-flexible cylinder. The equations for this model are described in the Supporting Information. Previously, 2D solution NMR was used to estimate a pitch of approximately 0.6 nm for a similar peptoid helix¹⁸. Based on this value for the pitch and the number of turns expected in **2c** (12, as there are 3 residues per turn), it is expected that the peptoid helix will have a contour length of approximately 7.2 nm. This value was used as the contour length for **2c**. For compound **1c**, the contour length was held at 13.0 nm, which corresponds to the distance along the peptoid backbone if all of the amide bonds are in the *trans* configuration. Table 1 shows the results of the fit. Interestingly, the persistence lengths are found to be quite similar: approximately 0.5 nm for **1c** and 1.0 nm for **2c**. Compound **2c** has a longer persistence length than **1c**, but it clearly is not stiff relative to other polymers containing secondary structure, such as PBLG molecules.⁶ The persistence lengths calculated from this semiflexible cylinder model match relatively well with the range estimated from the Kratky plot. Additionally, if the contour lengths are allowed to be fit by the semiflexible cylinder model, the same trend holds; compound **2c** has a shorter contour length and a larger persistence length than **1c**. These results are presented in the Appendix (Table 2.A.1).

The model fit indicates that the helical conformation of **2c** most likely imparts some rigidity to the polymer. However, **2c** still has a relatively short persistence length when compared to other helical polymers. Although the circular dichroism implies that the helical conformation is favored, the short persistence length indicates that the polymer can sample many different conformations and only a portion of the polymer chains adopt a full helical conformation at any given point in time. In addition, the helix observed here may simply be quite flexible. Previous intrinsic viscosity measurements of polymeric helical (*S*)-*N*-(1-phenylethyl)glycines (from 4 – 40 kg/mol) in DMF were consistent with random coil behavior such that Guo et al. concluded the persistence length of these chains must be less than 9 nm.¹¹ Flexible helical polymers have also previously been seen, such as in the case of chiral poly(2-oxazoline)s²⁸ where the polymer showed CD signal, indicating helix formation, but scattering indicated a random coil chain conformation.

Table 2.2. Fitted parameters for Compounds 1 and 2 with n = 6.

Compound 2 has a shorter contour length and a longer persistence length than 1, indicating its helical conformation imparts some stiffness to the chain.

Polypeptoid	Contour Length (nm)	Persistence Length* (nm)	Persistence Length** (nm)	Radius (nm)*
1c	13.0	0.51 ± 0.04	0.56	0.93 ± 0.2
2c	7.2	1.05 ± 0.08	1.12	0.99 ± 0.3

*As fitted by the flexible cylinder model

**As fitted by the wormlike chain model

Polyproline helices have also been observed to be flexible; in this case, the proline group creates a tertiary amide similar to those in the polypeptoid backbone. For tertiary amides, the energy barrier to rotation about the C-N bond is much lower than for the secondary amides that dominate proteins. In addition, the *cis/trans* configurations are much closer in energy for tertiary amides,²⁹⁻³¹ and thus the isomerization can occur much more readily. Previously, the activation energy for the *cis/trans* isomerization in dimer peptoids was measured to be on the order of 17 – 20 kcal/mol,³² which is similar to the energies measured for prolyl peptide bonds.^{29, 33, 34} It is well known that prolyl peptide bonds are expected to have much higher *cis:trans* ratios (1:3) compared to planar peptide bonds (~1:1000)^{35, 36} and that the polyproline I helix, which consists of all *cis* bonds, has a much shorter persistence length than that of a traditional peptide α -helix such as PBLG.⁷ Because the peptoid helices studied here were also shown to have a relatively high *cis:trans* ratio (~2-3:1),¹⁶⁻¹⁸ it should not be entirely unexpected that they are quite flexible. Recently, the introduction of much bulkier side groups (*N*-1-naphthylethyl) has been shown to raise the *cis:trans* ratio to >19:1,¹⁷ suggesting that even larger substituents can increase the energy barrier of backbone rotation and therefore increase chain stiffness.

Wormlike Chain Model

Further information about the chain conformation and the persistence length can be obtained by evaluating the radius of gyration over a series of chain lengths and fitting the wormlike chain formula. To obtain the radius of gyration (R_g), a line was fit to the scattering data in a Guinier plot ($\ln I(q)$ vs. q^2). For compounds **1c** and **2c**, the R_g 's differ slightly, yielding a value of 14.7 ± 0.1 Å and 13.3 ± 0.1 Å, respectively. This small decrease in R_g is most likely due to the more compact packing of the helical compound that stems from its secondary structure.

The wormlike chain formula³⁷ relates R_g to L_p :

$$R_g^2 = \frac{LL_p}{3} - L_p^2 + \frac{2L_p^3}{L} \left[1 - \frac{L_p}{L} \left(1 - e^{-L/L_p} \right) \right]$$

where L is the contour length as calculated using the geometry of the molecule and the previous 2D NMR studies. Using this value and the measured R_g , it is possible to solve the equation for the persistence length. Table 1 shows the values for the persistence lengths as determined by this method. This analysis supports the conclusions drawn from the semiflexible cylinder model. The persistence length for **2c** is slightly longer than that of **1c** (1.12 nm vs. 0.56 nm), but they are both short on an absolute scale.

The persistence length as calculated by the wormlike chain model is plotted against the number of monomers for chains of 18, 24, 36, and 48 residues. The first conclusion from this plot is that **2** consistently has a higher persistence length than **1**. Additionally, the shorter molecules have higher fitted persistence lengths, especially in the case of the helical molecule. The helical chain of 18 monomers could not fit the wormlike chain model with any reasonable value of the persistence length, indicating that the 18-mers are too short to be treated using this analysis. This is probably not due to actual differences in the number of residues present in the helical conformation as CD has shown that all of the different length polymers have very similar per-residue molar ellipticities. It is more likely that the wormlike chain model is not valid for short chains where the persistence length is not sufficiently shorter than the contour length. However,

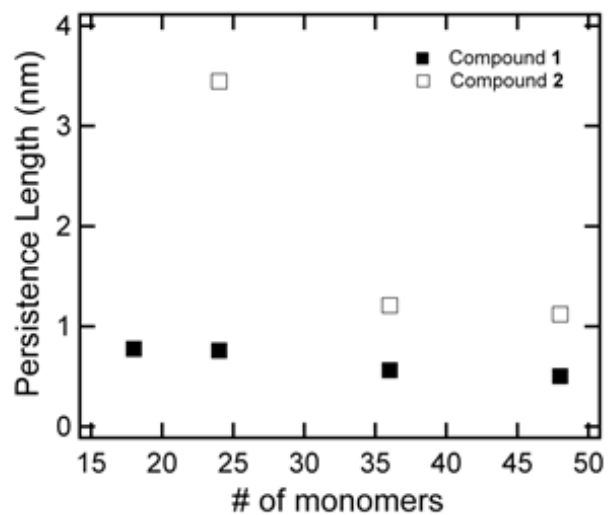


Figure 2.5. The persistence length, as determined using the worm-like chain model, of polypeptoids ranging in length from 18-48 monomers plotted versus the number of monomers.
The shorter chains have significantly higher persistence lengths, indicating that the molecules are too short to be treating using this method.

the chains of 36 and 48 monomers have reached an asymptote in their persistence length, suggesting that these polymers are of sufficient length to treat using the wormlike chain model. It is possible to fit the wormlike chain simultaneously to the polymers of different lengths and obtain a persistence length for each type of polymer. This analysis is presented in the Appendix (Figure 2.A.1) and yields very similar values for the persistence lengths.

2.4. Conclusions

In conclusion, the SANS study presented here yields two main insights into the chain conformation of polypeptoids. First, the polypeptoids studied here are inherently flexible in solution with persistence lengths ranging from 0.5 nm to 1.0 nm. Second, both the semiflexible cylinder model and the wormlike chain model indicate that inducing helicity by introducing bulky α -chiral side chains into a 36-mer polypeptoid results in a small increase in the persistence length or rigidity of the molecule. However, the fitted persistence length is still quite short in comparison to other helical polymers, suggesting that the polymer retains considerable conformational freedom. In agreement with previous 2D solution NMR studies and intrinsic viscosity measurements, the SANS data presented here indicates that helical polypeptoids with α -chiral, bulky phenyl side chains prefer an all *cis*-amide bond configuration in solution but can readily isomerize to sample other conformations as well.

2.5. Appendix

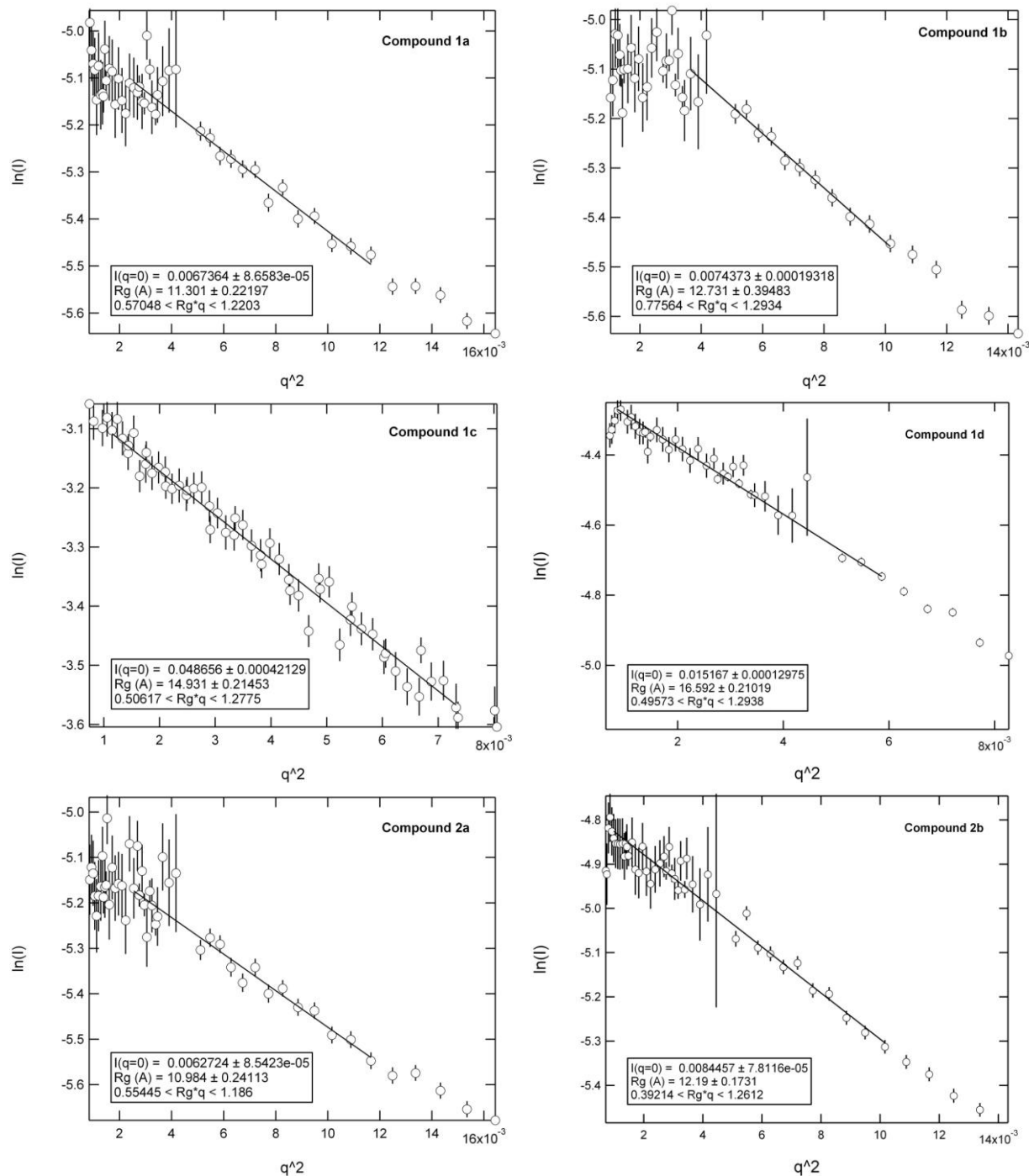
2.5.1. Semi-flexible Cylinder Model Equations

The following equations are used to fit the scattering data and calculate a persistence length.

$$\begin{aligned} & \text{---} \quad \text{---} \quad \text{---} \quad \text{---} \\ & \text{---} \quad \text{---} \quad \text{---} \end{aligned}$$

2.5.2. Guinier Plots

The following plots were used to deduce the R_g 's of Compounds **1** and **2**.



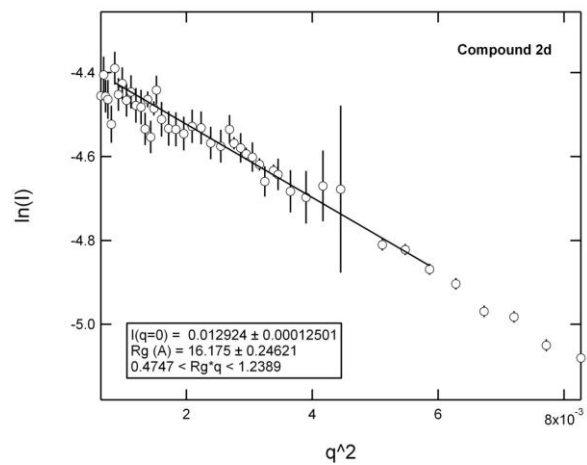
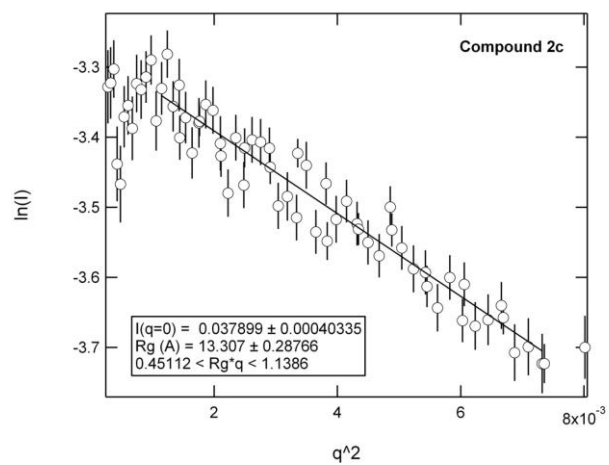


Figure 2.A.1. Guinier plots for all compounds

2.5.3. Wormlike Chain Analysis for Different Polypeptoid Chain Lengths

The radius of gyration (R_g) for polypeptoids 18, 24, 36, and 48 monomers long were obtained using a Guinier analysis on the SANS scattering curves for each polypeptoid (above). This molecular weight series was measured for both the non-helical (**1**) and the helical (**2**) compound. For compound **1**, the contour length was calculated as product of the number of monomers and the length of each monomer, and the wormlike chain equation was fit to the data by changing the persistence length, L_p . For compound **2**, the contour length was calculated using the helical pitch (0.6 nm) previously measured by 2D solution NMR experiments in acetonitrile. The fitted L_p was 0.6 nm for **1** and 1.1 nm for **2**, which agrees well with the values obtained from the semiflexible cylinder model fit.

Despite the good agreement, the wormlike chain equation does not quite fit the data, especially at lower polypeptoid chain lengths for both **1** and **2**. As stated in the manuscript, this error is due to the 18 and 24-mers simply being too short to be modeled by this equation.

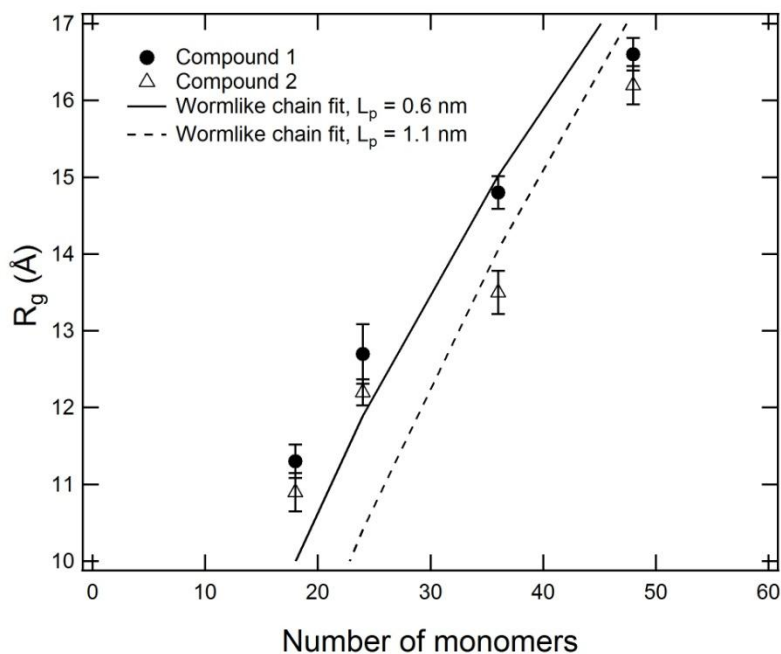


Figure 2.A.2. Fitted persistence lengths using the wormlike chain equation over a series of polypeptoid chain lengths.

The fitted L_p of **2** is longer than that of **1**, indicating it is indeed stiffer.

2.5.4. Persistence Length of a Polypeptoid with a Racemic Mixture of Monomers

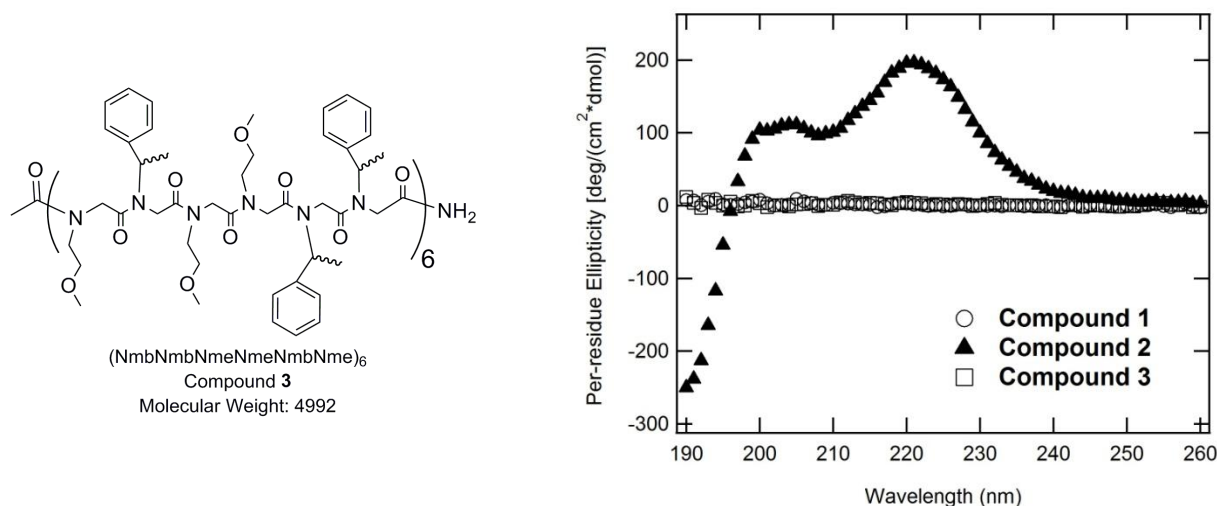


Figure 2.A.3. Compound 3 contains a racemic mixture of α -chiral side chains and is proposed to be non-helical.

The circular dichroism shows no net ellipticity

As another point of comparison, a 36-mer polypeptoid was synthesized using the same monomers as **2** but with a racemic mixture instead of an enantiomerically pure mixture of the α -methylbenzyl side chain. The structure for this compound (compound **3**) is given in 2.A.3. Circular dichroism data is provided in Figure 2.A.3b, showing that **3** is truly racemic and that there is no net ellipticity due to an equal mixture of *R* and *S* enantiomers. The helical structure of **2** results from the steric hindrance of the bulky side chains with the same chirality; thus, the racemic nature of **3** should disrupt any helical structure.

SANS measurements probed the structure of **3** in the same way as for **1** and **2**. A Guinier analysis of the scattering curve yields an R_g of 14.5 ± 0.1 Å, which is closer to the value of the R_g for **1** (14.7 Å) than **2** (13.3 Å). In addition, using the wormlike chain formula and the semiflexible cylinder model to calculate a persistence length for **3** yields a value of 0.62 nm and 0.39 nm, respectively. These results indicate that **3** is quite flexible, which is similar to **1** and **2**.

2.5.5. Semiflexible Cylinder Model of the 36-mer Polypeptoids Using a Fitted Contour Length

The semiflexible cylinder model was allowed to fit the contour length, persistence length, and cylinder radius simultaneously. Using this model, the non-helical sequence, **1**, has a much longer contour length than the chiral sequence, **2**. The secondary structure of **2c** is thought to cause the contour length to be much shorter as the polymer is adopting a helix conformation with the majority of the amide bonds in the *cis* geometry. This fitted value is less than the value estimated by NMR (7.2 nm) by about 2 nm. The fitted contour length for **1c** is less than the calculated fully extended (all-*trans*) chain length (13 nm), which is reasonable given that the flexible polypeptoid backbone chain likely exists with both *cis* and *trans* amide bonds at any given moment. Furthermore, a contour length of 8.5 ± 0.6 nm is obtained for compound **3**, suggesting that **3** is not as compact as **2c**. The fitted persistence length is also smaller for **3** (0.9 ± 0.02 nm) than **2c** (1.97 ± 0.17 nm), supporting the idea that the racemic nature of **3** results in a more flexible molecule. The results of this model fit are consistent with the trends identified from the semiflexible cylinder fit with a fixed contour length and the wormlike chain equation fit, but there is a larger error in the persistence length.

Polypeptoid	Contour Length (nm)*	Persistence Length* (nm)	Radius (nm)*
1	10.58 ± 4.8	0.66 ± 0.03	0.93 ± 0.2
2	4.93 ± 2.8	1.97 ± 0.17	0.99 ± 0.3
3	8.5 ± 0.6	0.9 ± 0.02	0.96 ± 0.2

Table 2.A.1. The contour length and the persistence length as fit using the semiflexible cylinder model.

Consistent with other analyses explored in this paper, the helical compound has a slightly higher persistence length than the non-helical version.

2.5.6. Circular Dichroism Measurements

Extensive CD measurements have been taken at various temperatures and solvents. The most relevant graphs have been included in the manuscript. The remaining data is summarized here.

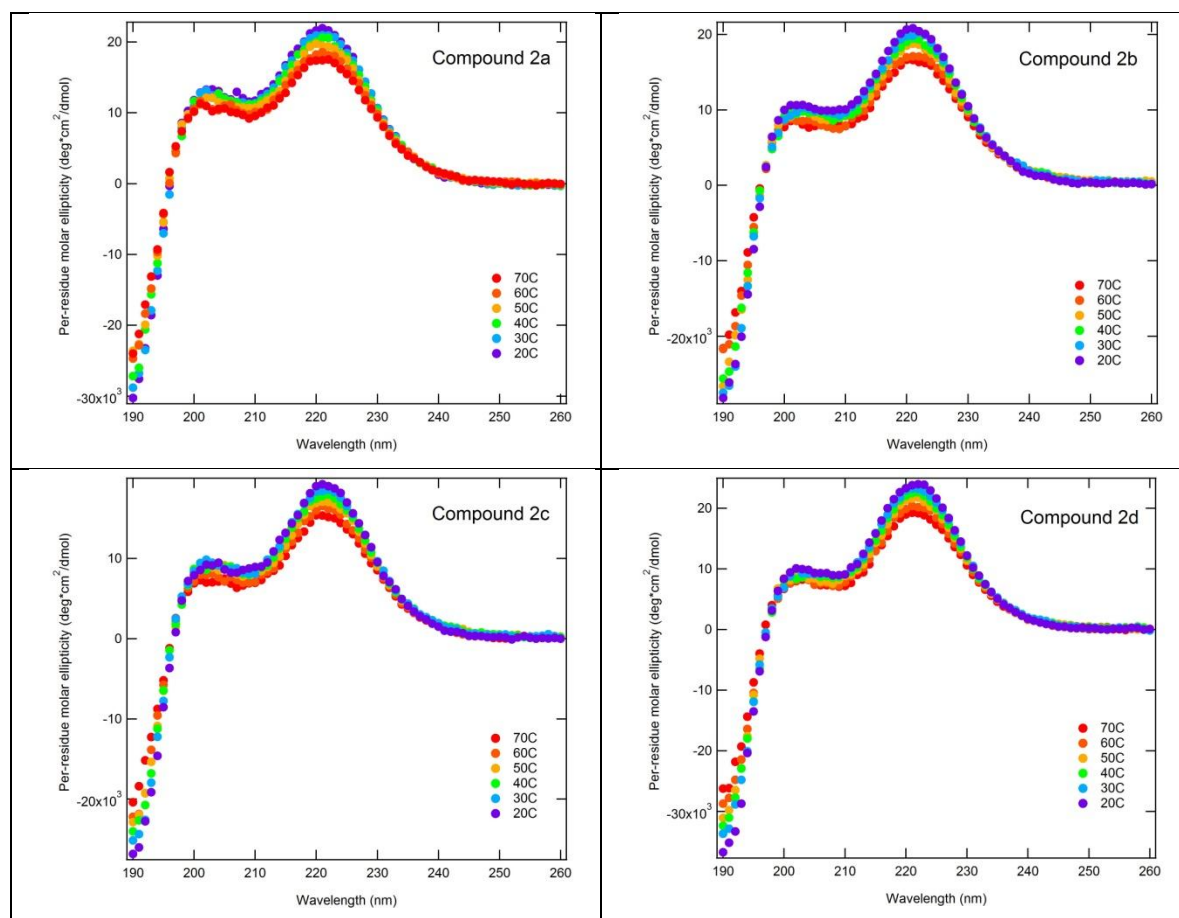


Figure 2.A.4. Circular dichroism at varying temperatures for each helical polypeptoid.

2.5. Acknowledgements

We gratefully acknowledge funding from the Office of Naval Research via a Presidential Early Career Award in Science and Engineering. A.M.R. and H.K.M also gratefully acknowledge the National Science Foundation and the Department of Defense for graduate fellowships (respectively). Polypeptoid synthesis and associated chemical characterization were performed at the Molecular Foundry, a Lawrence Berkeley National Laboratory user facility supported by the Office of Science, Office of Basic Energy Sciences, U.S. Department of Energy, under Contract DE-AC02-05CH11231. The neutron scattering in this work is based on activities at the NIST Center for Neutron Research, which is supported in part by the National Science Foundation under Agreement No. DMR-0454672. Certain trade names and company products are identified to adequately specify the experimental procedure. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products are necessarily best for the purpose. A portion of this research was also performed at Oak Ridge National Laboratory. The authors thank Dr. Volker S. Urban at Oak Ridge National Laboratory for assistance on SANS data collection. The SANS studies at Oak Ridge National Laboratory's Center for Structural Molecular Biology were supported by the Office of Biological and Environmental Research, using facilities supported by the DOE, managed by UT-Battelle, LLC, under Contract No. DE-AC05-00OR22725.

2.6. References

1. Gettinger, C. L.; Heeger, A. J.; Drake, J. M.; Pine, D. J., A photoluminescence study of poly(phenylene vinylene) derivatives - The effect of intrinsic persistence length. *J. Chem. Phys.* **1994**, *101* (2), 1673-1678.
2. Brulet, A.; Boue, F.; Cotton, J. P., About the experimental determination of the persistence length of wormlike chains of polystyrene. *Journal De Physique Li* **1996**, *6* (6), 885-891.
3. Terao, K.; Terao, Y.; Teramoto, A.; Nakamura, N.; Terakawa, I.; Sato, T., Stiffness of polysilylenes depending remarkably on a subtle difference in chiral side chain structure: Poly{n-hexyl-[(S)-2-methylbutyl]silylene} and poly{n-hexyl-[(S)-3-methylpentyl]silylene}. *Macromolecules* **2001**, *34* (8), 2682-2685.
4. Muller-Spath, S.; Soranno, A.; Hirschfeld, V.; Hofmann, H.; Ruegger, S.; Reymond, L.; Nettels, D.; Schuler, B., Charge interactions can dominate the dimensions of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (33), 14609-14614.
5. Gu, H.; Nakamura, Y.; Sato, T.; Teramoto, A.; Green, M. M.; Andreola, C., Global conformations of chiral polyisocyanates in dilute solution. *Polymer* **1999**, *40* (4), 849-856.
6. Temyanko, E.; Russo, P. S.; Ricks, H., Study of rodlike homopolypeptides by gel permeation chromatography with light scattering detection: Validity of universal calibration and stiffness assessment. *Macromolecules* **2001**, *34* (3), 582-586.

7. Best, R. B.; Merchant, K. A.; Gopich, I. V.; Schuler, B.; Bax, A.; Eaton, W. A., Effect of flexibility and cis residues in single-molecule FRET studies of polyproline. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (48), 18964-18969.
8. Murnen, H. K.; Rosales, A. M.; Jaworski, J. N.; Segalman, R. A.; Zuckermann, R. N., Hierarchical Self-Assembly of a Biomimetic Diblock Copolypeptoid into Homochiral Superhelices. *J. Am. Chem. Soc.* **2010**, *132* (45), 16112-16119.
9. Nam, K. T.; Shelby, S. A.; Choi, P. H.; Marciel, A. B.; Chen, R.; Tan, L.; Chu, T. K.; Mesch, R. A.; Lee, B. C.; Connolly, M. D.; Kisielowski, C.; Zuckermann, R. N., Free-floating ultrathin two-dimensional crystals from sequence-specific peptoid polymers. *Nat. Mater.* **2010**, *9* (5), 454-460.
10. Rosales, A. M.; Murnen, H. K.; Zuckermann, R. N.; Segalman, R. A., Control of Crystallization and Melting Behavior in Sequence Specific Polypeptoids. *Macromolecules* **2010**, *43* (13), 5627-5636.
11. Guo, L.; Li, J. H.; Brown, Z.; Ghale, K.; Zhang, D. H., Synthesis and Characterization of Cyclic and Linear Helical Poly(alpha-peptoid)s by N-Heterocyclic Carbene-Mediated Ring-Opening Polymerizations of N-Substituted N-Carboxyanhydrides. *Biopolymers* **2011**, *96* (5), 596-603.
12. "Thermodynamics of Polymer Blends", N. P. B., H.B. Eitouni, Physical Properties of Polymers Handbook, Second Edition, Chapter 19, p. 339-356, J.E. Mark, editor, Springer, New York, 2007.
13. Fowler, S. A.; Blackwell, H. E., Structure-function relationships in peptoids: Recent advances toward deciphering the structural requirements for biological function. *Organic & Biomolecular Chemistry* **2009**, *7* (8), 1508-1524.
14. Kirshenbaum, K.; Barron, A. E.; Goldsmith, R. A.; Armand, P.; Bradley, E. K.; Truong, K. T. V.; Dill, K. A.; Cohen, F. E.; Zuckermann, R. N., Sequence-specific polypeptoids: A diverse family of heteropolymers with stable secondary structure. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (8), 4303-4308.
15. Yoo, B.; Kirshenbaum, K., Peptoid architectures: elaboration, actuation, and application. *Current Opinion in Chemical Biology* **2008**, *12* (6), 714-721.
16. Armand, P.; Kirshenbaum, K.; Falicov, A.; Dunbrack, R. L.; Dill, K. A.; Zuckermann, R. N.; Cohen, F. E., Chiral N-substituted glycines can form stable helical conformations. *Folding & Design* **1997**, *2* (6), 369-375.
17. Stringer, J. R.; Crapster, J. A.; Guzei, I. A.; Blackwell, H. E., Extraordinarily Robust Polyproline Type I Peptoid Helices Generated via the Incorporation of $\hat{\pm}$ -Chiral Aromatic N-1-Naphthylethyl Side Chains. *Journal of the American Chemical Society* **2011**.
18. Armand, P.; Kirshenbaum, K.; Goldsmith, R. A.; Farr-Jones, S.; Barron, A. E.; Truong, K. T. V.; Dill, K. A.; Mierke, D. F.; Cohen, F. E.; Zuckermann, R. N.; Bradley, E. K., NMR determination of the major solution conformation of a peptoid pentamer with chiral side chains. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (8), 4309-4314.

19. Wu, C. W.; Sanborn, T. J.; Huang, K.; Zuckermann, R. N.; Barron, A. E., Peptoid oligomers with alpha-chiral, aromatic side chains: Sequence requirements for the formation of stable peptoid helices. *J. Am. Chem. Soc.* **2001**, *123* (28), 6778-6784.
20. Wu, C. W.; Sanborn, T. J.; Zuckermann, R. N.; Barron, A. E., Peptoid oligomers with alpha-chiral, aromatic side chains: Effects of chain length on secondary structure. *J. Am. Chem. Soc.* **2001**, *123* (13), 2958-2963.
21. Figliozzi, G. M.; Goldsmith, R.; Ng, S. C.; Banville, S. C.; Zuckermann, R. N., Synthesis of N-substituted glycine peptoid libraries. *Combinatorial Chemistry* **1996**, *267*, 437-447.
22. Kline, S. R., Reduction and analysis of SANS and USANS data using IGOR Pro. *J. Appl. Crystallogr.* **2006**, *39*, 895-900.
23. Kelly, S. M.; Jess, T. J.; Price, N. C., How to study proteins by circular dichroism. *BBA-Proteins Proteomics* **2005**, *1751* (2), 119-139.
24. Kratky, O., Das studium geloster fadenmolekule mittels der rontgenkleinwinkelmethode. *Kolloid-Zeitschrift and Zeitschrift Fur Polymere* **1962**, *182* (1-2), 7-&.
25. Koyama, R., Light-scattering of stiff chain polymers. *Journal of the Physical Society of Japan* **1973**, *34* (4), 1029-1038.
26. Burchard, W.; Kajiwara, K., Statistics of stiff chain molecules .1. Particle scattering factor. *Proceedings of the Royal Society of London Series a-Mathematical and Physical Sciences* **1970**, *316* (1525), 185-&.
27. Pedersen, J. S.; Schurtenberger, P., Scattering functions of semiflexible polymers with and without excluded volume effects. *Macromolecules* **1996**, *29* (23), 7602-7612.
28. Bloksma, M. M.; Rogers, S.; Schubert, U. S.; Hoogenboom, R., Secondary structure formation of main-chain chiral poly(2-oxazoline)s in solution. *Soft Matter* **2010**, *6* (5), 994-1003.
29. Wedemeyer, W. J.; Welker, E.; Scheraga, H. A., Proline cis-trans isomerization and protein folding. *Biochemistry* **2002**, *41* (50), 14637-14644.
30. Zimmerman, S. S.; Scheraga, H. A., Stability of cis, trans, and nonplanar peptide groups. *Macromolecules* **1976**, *9* (3), 408-416.
31. Maignet, B.; Perahia, D.; Pullman, B., Molecular orbital calculations on conformation of polypeptides and proteins .4. Conformation of prolyl and hydroxyprolyl residues. *J. Theor. Biol.* **1970**, *29* (2), 275-&.
32. Sui, Q.; Borchardt, D.; Rabenstein, D. L., Kinetics and equilibria of cis/trans isomerization of backbone amide bonds in peptoids. *J. Am. Chem. Soc.* **2007**, *129* (39), 12042-12048.
33. Fischer, S.; Dunbrack, R. L.; Karplus, M., Cis-trans imide isomerization of the proline dipeptide. *J. Am. Chem. Soc.* **1994**, *116* (26), 11931-11937.
34. Steinberg, I. Z.; Harrington, W. F.; Berger, A.; Sela, M.; Katchalski, E., The Configurational Changes of Poly-L-proline in Solution. *J. Am. Chem. Soc.* **1960**, *82* (20), 5263-5279.
35. Weiss, M. S.; Jabs, A.; Hilgenfeld, R., Peptide bonds revisited. *Nat. Struct. Biol.* **1998**, *5* (8), 676-676.

36. Chakrabarti, P.; Pal, D., The interrelationships of side-chain and main-chain conformations in proteins. *Prog. Biophys. Mol. Biol.* **2001**, 76 (1-2), 1-102.
37. Benoit, H.; Doty, P., Light Scattering from Non-Gaussian Chains. *The Journal of Physical Chemistry* **1953**, 57 (9), 958-963.

Chapter 3: Persistence length of polyelectrolytes with precisely located charges

Reproduced with the permission of Adrienne M. Rosales, Andrey V. Dobrynin, Ronald Zuckermann, Rachel Segalman, 2012 in preparation

The conformation of biological polyelectrolytes in salt solutions is closely related to their ability to fold into desired structures. In particular, their persistence length has a large impact on how the chain can arrange itself. In this work, poly N-substituted glycines (polypeptoids), have been designed to include specifically located charges such that the relationship between their persistence lengths, the salt concentration, and the spacing of the charges along the backbone can be elucidated. It is shown that at low salt concentrations, polypeptoids with charge groups located closer to each other along the polymer backbone are stiffer than those with the charge groups spaced further apart. At high salt concentrations, the persistence length decreases for both macromolecules as the electrostatic repulsions are screened. The data has been analyzed using a semi-flexible model and it is shown that the persistence length has a quadratic dependence on the Debye screening length for these materials. It is also shown that the bending rigidity is higher for a molecule with alternating chargeable groups than that of a molecule containing all chargeable groups. This difference can be attributed to the longer side chains forcing the macromolecule to adopt a stiffer conformation.

3.1. Introduction

The conformation of biopolymers such as DNA and proteins is highly influenced by the presence of charges along their backbone. Interactions between charged groups play a dominant role in polyelectrolyte conformation due to electrostatic induced stiffening and elongation of the polyelectrolytes.^{1,2} However, at physiological conditions, solutions of biopolymers contain various salts. These salts screen the long-range electrostatic repulsion between ionized groups, resulting in more flexible polyelectrolytes.³ Past a certain point, the presence of salt leads to counterion condensation,⁴ which reduces the charge density of the polyelectrolyte and can also affect the chain conformation. Understanding the relationship between polyelectrolyte conformation, charged group spacing, and salt concentration is a complex problem that has been addressed in numerous computational, theoretical and experimental works^{3,5-7}. This study examines the relationship between a measure of the chain conformation, the persistence length, and salt concentration in polypeptoids, a polyelectrolyte model system that contains precisely placed charged groups.

Much of the debate on developing models for polyelectrolyte systems has centered on the flexibility of the underlying chain and how it impacts the scaling of the persistence length with regards to salt concentration.⁸ On the one hand, theory originating with Kuhn, Kunzle, and Katchalsky⁹ shows that persistence length scales with the inverse of the square root of the salt concentration (i.e., a linear dependence on the Debye screening length); this calculation corresponds to a flexible or semi-flexible chain with electrostatic corrections. On the other hand, theory based on the work of Odijk,^{1,10-12} Skolnick, and Fixman² (OSF) shows the persistence length scales with the inverse of the salt concentration (i.e., a quadratic dependence on the Debye screening length); OSF theories assume a locally stiff chain for which electrostatic interactions can modify the bending properties. In both cases, theory has been developed over the past several decades, and neither fully accounts for experimental observations.

Experimental studies of biologically relevant polyelectrolytes in water have mostly focused on DNA. The effect of salt concentration on the persistence length of DNA has been studied using a variety of techniques, ranging from light scattering^{13,14} to force-extension¹⁵ experiments. However, the measurement of the persistence length can be convoluted by these techniques. For example, at high and moderate salt concentrations, a persistence length obtained by light scattering measurements for sufficiently long macromolecules can include an excluded volume contribution. In addition, force-extension experiments can underestimate values of the persistence length due to enthalpic effects associated with stretching the molecules at forces above 5 pN. Because of these issues, the relationship between salt concentration and DNA persistence length is difficult to quantify, and experimental observations often contradict each other. Baumann et. al¹⁵ measured the electrostatic persistence length of DNA using force-measuring laser tweezers, and they found quantitative agreement with OSF predictions. However, Makita and coworkers¹⁶ measured the persistence length of DNA using fluorescence microscopy and found that it exhibited a linear dependence on the Debye screening length. As for other biopolymers, light scattering has been used to determine a scaling relationship for the persistence length of hyaluronan,¹⁷ a polysaccharide, and salt, but there are very few studies for polypeptides. Given the difficulties in obtaining a definitive result for the verification of scaling laws, it is useful to investigate a simplified system in which factors such as chain length and intermolecular interactions can be tuned directly.

Biological polyelectrolytes have a sequence specific arrangement of the charged groups in their backbone. The sequence of charged groups is essential to the structure and function of the overall molecule. However, polypeptides and other natural biopolymers also have many other interactions that impact their persistence length. For instance, polypeptides are inherently chiral and form intermolecular hydrogen bonds. They can also have strong hydrophobic interactions that drive a chain to collapse. Therefore, to directly investigate the effect of charged group sequence on chain conformation, a greatly simplified biomimetic system is ideal. Poly N-substituted glycines, or polypeptoids, are a sequence specific, biomimetic system with the same backbone as polypeptides, but the side chain is attached to the nitrogen rather than the backbone α -carbon. This N-substitution eliminates chirality and hydrogen bonding in the backbone, allowing the interactions in the system to be tuned by the introduction of different side chains. Furthermore, these materials are synthesized using a solid state submonomer process that leads to sequence specific, monodisperse chains. Thus, polypeptoids are a convenient system in which to introduce ionizable side groups and vary the exact spacing between those groups.

In addition to serving as a model biomimetic polymer, polypeptoids are an interesting self-assembling material in their own right. When polypeptoids contain charged groups, they can assemble in solution into structures such as a supramolecular helix¹⁸ and nanometer-thick sheets.¹⁹⁻²¹ To model and understand these assemblies, it is helpful to have a basic understanding of polypeptoid behavior with regard to ionic interactions. In this work, polypeptoids with two different spacings of carboxyl groups have been synthesized, and their persistence lengths have been measured using small angle neutron scattering (SANS). The results show that at low salt concentrations, a polypeptoid in which every monomer contains an ionizable group has a higher persistence length than a polypeptoid in which ionizable and neutral monomers alternate. At high salt concentrations, this difference is far smaller, and both polypeptoids show a decrease in persistence length. The data has been analyzed using a semi-flexible model and it is shown that the persistence length has a quadratic dependence on the Debye screening length. It is also shown that the bending rigidity is higher for the molecule with alternating charged groups which

can be attributed to the longer side chains forcing the macromolecule to adopt a stiffer conformation.

3.2. Experimental methods

Synthesis

Polypeptoids were synthesized on a commercial robotic synthesizer Aapptec Apex 396 on 100 mg of Rink amide polystyrene resin (0.6 mmol/g, Novabiochem, San Diego, CA) using the procedure previously detailed.^{18, 19} All primary amine monomers, solvents, and reagents were purchased from commercial sources and used without further purification. The β -alanine was purchased from ChemImpex and used after freebasing from the β -alanine O-tBu ester hydrochloride submonomer by extraction from dichloromethane (DCM) and basic water. The resulting compound was confirmed by ¹H NMR. All primary amine submonomers were 1.5M in N-methyl pyrrolidinone and all displacement times were 60 minutes for the first 15 additions, 90 minutes for the next 15 and 120 minutes for the remainder. All polypeptoids were acetylated on the resin and purified using reverse phase HPLC as previously described.¹⁹ The mass and purity were confirmed using reverse phase analytical HPLC and MALDI. All polymers synthesized are listed in Table 3.1 along with their purities and molecular weights.

Small angle neutron scattering (SANS)

SANS studies were conducted at the NG3 SANS line at the National Institute of Standards and Technology (NIST) Center for Neutron Research in Gaithersburg, Maryland and at the CG-3 Bio-SANS line at the High Flux Isotope Reactor at Oak Ridge National Laboratory (ORNL) in Oak Ridge, Tennessee. Samples were prepared at a concentration of 10 mg/mL in deuterated water to enhance the contrast between the polypeptoids and the solvent. Quartz banjo cells (Hellma USA, Plainview, NY) with a path length of 1 mm and 2 mm were used at NIST and ORNL, respectively, in a temperature controlled multiple position sample holder. A neutron wavelength of 6 Å was used at both beamlines, and data were collected at two different instrument configurations (1.3 m and 4 m at NIST, and 1.7 m and 14.5 m at ORNL). The data were reduced using the NCNR SANS reduction macros²⁰ and the Spice SANS reduction program in Igor Pro.

Titration

The charge state of each molecule was probed through titrations with sodium hydroxide, NaOH. Solutions of 0.5M NaOH were added 2 μ L at a time to solutions of $p(\text{Nce})_{36}\text{ac}$ (2.1 mM, 10 mg/ml) and $p(\text{NceNme})_{18}\text{ac}$ (2.2 mM, 10 mg/ml) while the solutions were stirring. The pH was continuously read and after each addition was allowed to equilibrate until the pH value stabilized. These pH values were plotted against an x-axis of equivalents of NaOH per COOH group.

Solutions

Solutions (Table 3.2) were made by dissolving the desired polymer at a concentration of 10 mg/ml (2.1mM for $p(\text{Nce})_{36}\text{ac}$ and 2.2mM for $p(\text{NceNme})_{18}\text{ac}$) in water. Sodium hydroxide was added by dripping in the correct volume of 0.5M NaOH. The correct volume was determined by calculating the molar concentration of COOH groups and then providing the correct molar equivalent of NaOH for that number of COOH groups. The solutions were allowed to equilibrate for 1-2 days.

Table 3.1. The polymers used, their purity, observed mass, and chemical structures

Polymer	Abbreviation	Purity	Obs/Actual MW	Structure
Poly(N-carboxyethyl glycine-N-methoxyethyl glycine) ₁₈ acetyl	$p(\text{NceNme})_{18}\text{ac}$	97.6%	4487.1/4488.2	
Poly(N-carboxyethyl glycine) ₃₆ acetyl	$p(\text{Nce})_{36}\text{ac}$	95.4%	4737.0/4737.6	

Table 3.2. All solutions used in this study.

Sample	Polymer	NaOH equivalent	[NaOH] M	Overall fraction of monomers that are charged, α	Measured L_p (nm)	Measured R_g (nm)
1	<i>p</i> (Nce) ₃₆ ac	1	0.04	0.48	0.71	0.74
2	<i>p</i> (Nce) ₃₆ ac	3	0.15	1	0.91	1.05
3	<i>p</i> (Nce) ₃₆ ac	10	0.68	1	0.48	1.21
4	<i>p</i> (Nce) ₃₆ ac	30	2.19	1	0.29	1.56
5	<i>p</i> (NceNme) ₁₈ ac	1	0.03	0.18	0.56	0.86
6	<i>p</i> (NceNme) ₁₈ ac	3	0.08	0.50	0.65	1.29
7	<i>p</i> (NceNme) ₁₈ ac	10	0.36	0.50	0.56	1.68
8	<i>p</i> (NceNme) ₁₈ ac	30	1.15	0.50	0.43	1.52

3.3 Results and Discussions

Two polymers were designed to probe the effect of charge location on the chain conformations in aqueous solution. The use of polypeptoids allowed the specification of the exact location for each monomer as well as exact monodispersity for each chain. It was anticipated that this would eliminate ambiguities present in studies of synthetic polyelectrolytes due to polydispersity and uncertainty of monomer location. In addition, the simplified nature of the polypeptoid system in comparison to other biopolymers allows the isolation of the charged groups without the impact of backbone hydrogen bonding or chirality. The first polymer, $p(\text{NceNme})_{18}\text{ac}$ (Table 1) was designed to contain alternating ionizable monomers, N-carboxyethyl glycine, and neutral monomers, N-methoxyethyl glycine. The second polymer, $p(\text{Nce})_{36}\text{ac}$, contained all ionizable groups. When $p(\text{Nce})_{36}\text{ac}$ is completely charged, the negative charges are adjacent to each other. The location of these charges was expected to have an impact on the chain behavior in solution because the charges that are directly next to each other will have a stronger electrostatic repulsion than the charges that are spaced further apart.

The carboxylic acid side chain is pH dependent, and a titration with sodium hydroxide (Figure 1) shows the relative charge states of each polymer. Interestingly, it is easier to deprotonate $p(\text{Nce})_{36}\text{ac}$ than $p(\text{NceNme})_{18}\text{ac}$. It was expected that the shorter spacing between charges in $p(\text{Nce})_{36}\text{ac}$ would make it more difficult to remove protons because of the proximity and therefore repulsion of the resulting negative charges. However, the opposite trend was seen. Less equivalents of NaOH were required to reach the equivalence point and eventually full deprotonation for $p(\text{Nce})_{36}\text{ac}$. This may be due to the fact that when the ionizable groups are directly next to each other they can more easily share sodium ions, allowing the charges to be more evenly distributed. In the case of the alternating polymer where the charges are fixed at least 0.74 nm apart from each other along the backbone, they may be more isolated meaning that each deprotonated carboxylic acid has its own corresponding sodium atom, requiring more sodium hydroxide molecules to fully deprotonate the molecule.

Solutions were made with each polymer and various concentrations of NaOH (Table 2). The first solution was made with 1 equivalent of NaOH, where the polymers were only partially charged. From the titration curve equivalence point, it is calculated that the acid groups on $p(\text{Nce})_{36}\text{ac}$ for this solution (sample 1) are 48% charged while those on the $p(\text{NceNme})_{18}\text{ac}$ (sample 5) are 35% charged. Because only 50% of the monomers in $p(\text{NceNme})_{18}\text{ac}$ are ionizable, the overall fraction of monomers that are charged, α , is actually half of that, or 18%. The subsequent solutions were all made at greater concentrations of sodium hydroxide (3x, 10x, 30x) where all of the carboxylic acids are negatively charged. In these solutions, adding more sodium hydroxide simply provides more salt ions to screen the electrostatic repulsion.

Table 2: All solutions used in this study.

Sample	Polymer	NaOH equivalent	[NaOH] M	Overall fraction of monomers that are charged, α	Measured L_p (nm)	Measured R_g (nm)
1	<i>p</i> (Nce) ₃₆ ac	1	0.04	0.48	0.71	0.74
2	<i>p</i> (Nce) ₃₆ ac	3	0.15	1	0.91	1.05
3	<i>p</i> (Nce) ₃₆ ac	10	0.68	1	0.48	1.21
4	<i>p</i> (Nce) ₃₆ ac	30	2.19	1	0.29	1.56
5	<i>p</i> (NceNme) _{18a} _c	1	0.03	0.18	0.56	0.86
6	<i>p</i> (NceNme) _{18a} _c	3	0.08	0.50	0.65	1.29
7	<i>p</i> (NceNme) _{18a} _c	10	0.36	0.50	0.56	1.68
8	<i>p</i> (NceNme) _{18a} _c	30	1.15	0.50	0.43	1.52

In order to obtain information about the conformation of these molecules at the various salt concentrations, small angle neutron scattering (SANS) experiments were performed. As a first assessment of the polymer conformation, the SANS data was plotted on a Kratky plot, which graphs Q^2I versus Q (Figure 2). The shape of this plot provides information about the polymer and its arrangement in solution. Figure 2a shows the Kratky plot for $p(\text{Nce})_{36}\text{ac}$ at various salt concentrations. The data for both the 1x and 3x solutions increase monotonically, indicating that the molecule is dispersed in solution. However, for both the 10x and 30x solutions, there is a distinct difference in the shape of the data. The data now shows a hump with a peak around $0.1\text{--}0.15 \text{ \AA}^{-1}$. This hump indicates that the macromolecule has a more coiled configuration in solution. This is likely due to the fact that when the amount of sodium in the solution is relatively low (in the 1x and 3x case), the ionized groups cause the molecule to remain extended. However, when more salt is added (in the case of 10x and 30x solutions), the molecule can also form complexes between the positive sodium and the negative carboxylic acids. This is known as the counterion condensation effect.⁴ It pulls the molecule into a more coiled state. The same trend is seen for the $p(\text{NceNme})_{18}\text{ac}$. The decrease in chain size due to counterion condensation has been seen experimentally^{25, 26} and in computer simulations.^{27, 28} While R_g measurements would be helpful in confirming this analysis, low signal to noise in the low q range (Guinier region) of the scattering data makes it difficult to obtain reliable R_g values. In addition, the R_g measurement may be convoluted by any complexation of the polypeptoid with salt.

In order to gain further understanding of the effect of electrostatic interactions on chain configuration, we used a semi-flexible chain model²⁹ to fit the SANS data and to obtain a value of persistence length, L_p , for each salt concentration. These data are summarized in Table 2. In agreement with previous measurements of polypeptoid persistence lengths³⁰, these molecules are quite flexible with persistence lengths ranging from 0.3 nm up to approximately 1 nm. Dependence of the persistence length on salt concentration is shown in Figure 3. The chain persistence length first increases with increasing salt concentration and then begins to decrease with addition of further salt. The initial increase is due to the fact that the NaOH is behaving as a titrant increasing the fraction of ionized groups along the polymer backbone. These ionized groups stiffen the polymer chain, resulting in an increase in persistence length. As more sodium hydroxide is added, the persistence length decreases, which is expected because the added salt ions screen electrostatic interactions between ionized groups along the polymer backbone making the macromolecules more flexible. This is in line with the qualitative results from the Kratky plots where it was seen that at higher salt concentrations the molecules adopted more coiled conformations (and therefore are less stiff). There is also a difference between the two polymers, particularly at low NaOH concentrations. The $p(\text{Nce})_{36}\text{ac}$ has a larger L_p at these concentrations which is due to the close proximity of the charged groups along the backbone and the larger overall amount of charge as compared to $p(\text{NceNme})_{18}\text{ac}$. These two factors both serve to increase the persistence length for $p(\text{Nce})_{36}\text{ac}$, demonstrating that the specific location of the charges has the ability to tune the polymer conformation in solution. As discussed in the introduction, the most common treatment for the theoretical scaling of persistence length with salt for polyelectrolytes is the work of Odijk^{1, 10-12} and Skolnick and Fixman² who showed that for locally stiff polyelectrolytes, the persistence length should scale directly with the inverse of the salt concentration. This relationship holds well for low salt concentrations and for relatively stiff polyelectrolytes. However, at high salt concentrations and for more flexible polyelectrolytes, there is significant deviation from the OSF model. Several experimental³¹ and theoretical^{32, 33}

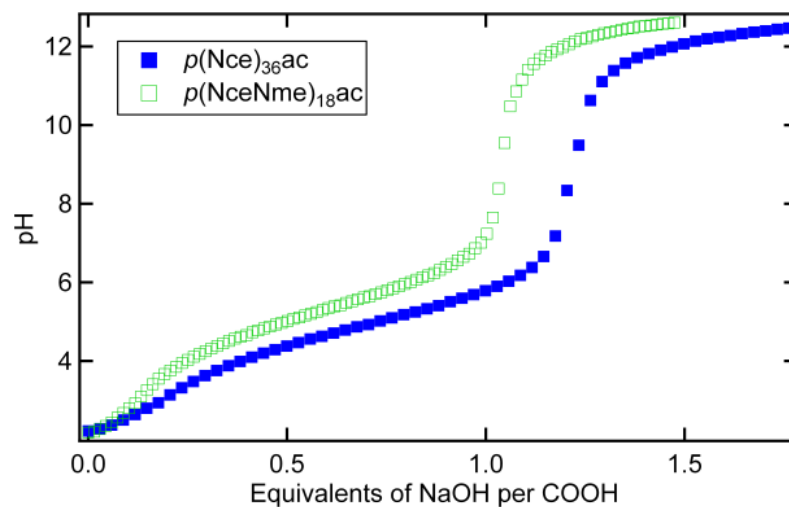


Figure 3.1. The titration curves of $p(\text{Nce})_{36}\text{ac}$ and $p(\text{NceNme})_{18}\text{ac}$

studies have shown that the persistence length of flexible polyelectrolytes is actually inversely proportional to the square root of the salt concentration, or chain persistence length scales linearly with the Debye screening length. These results hold especially well for polyelectrolytes that do not assume locally stiff conformations. Given the very low persistence length values measured here, it's likely that the polyelectrolytes described in this paper do not adopt locally stiff conformations. Therefore, it is possible to probe the scaling of the persistence length with salt by plotting an equation of the form

$$L_p = f + g[\text{NaOH}]^{-0.5}$$

with constants f and g being fitting parameters. Using a linear error minimization, they can be calculated to have the values of $f=-1.6$ and $g=2.1$ respectively. The negative value of parameter f is unphysical since it would correspond to a negative bare persistence length. This indicates that the simple scaling analysis is not applicable to the data (most likely due to the short length of the chains) and it is necessary to delve more deeply into the data analysis.

Below we will account for the finite chain length and map our polymers into a discrete chain model with internal bending rigidity³⁴. In the framework of this model,³⁴ chain properties are described by two independent parameters: effective chain bond length b and chain bending constant K . For effective bond length, we will use a projection length of a bond in zig-zag conformation which gives a value of $b=0.37$ nm. In the case of charged polymers, a chain bending constant K has contributions from the bare chain bending rigidity K_0 and one coming from electrostatic interactions between ionized groups along the polymer backbone.³⁵

$$K \approx K^0 + \frac{l_B \alpha^2}{12b} \sum_{m=1}^N \left(1 - \frac{m}{N}\right) \exp(-\kappa b m) (1 + \kappa b m) m \quad (2)$$

where l_B is the Bjerrum length and α is the fraction of ionized groups along the polymer backbone. Note that in the case of infinitely long chains, $N \gg 1$, eq 1 can be reduced to the classical OSF expression for chain persistence lengths.^{1, 2, 10} Persistence length of a chain with bending rigidity K and bond length b is equal to

$$L_p = \left(\frac{b}{2}\right) \frac{1 + \coth K - K^{-1}}{1 - \coth K + K^{-1}} \approx \begin{cases} bK, & \text{for } K \gg 1 \\ b/2, & \text{for } K \ll 1 \end{cases} \quad (3)$$

This equation was used to fit the measured persistence length as a function of NaOH concentration as shown in Figure 4 by considering K^0 in eq 1 as fitting parameter. For both peptoids we used the ionization degree α obtained from the titration curves (see Table 2). The fitting results are summarized in Table 3. As one can see from the Table, the agreement between experimental and calculated values is very good. The biggest difference is observed for values of the persistence length at highest salt concentration. The value of the bare bending constant K^0 is larger for the $p(\text{NceNme})_{18\text{ac}}$ peptoid which could be due to the longer side chain attached to the main backbone causing some bending rigidity to occur. In Figure 4, we plot the electrostatic contribution to the chain bending constant. All of the data collapses onto a universal line confirming quadratic dependence of the chain persistence length on the Debye screening length.

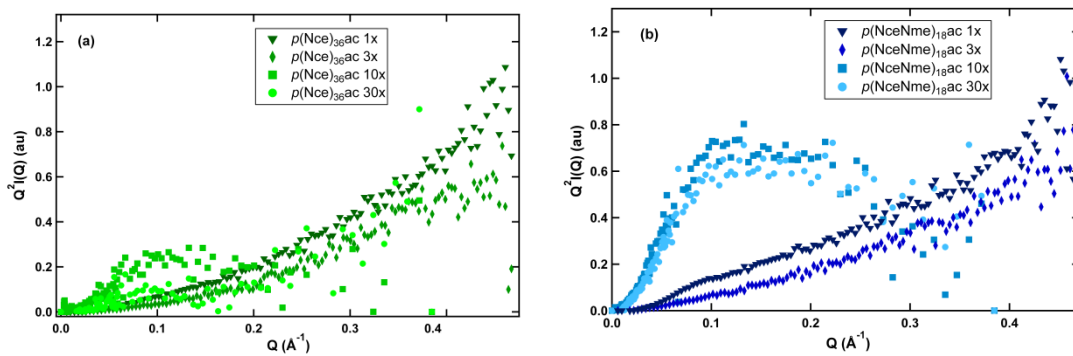


Figure 3.2. Kratky plots for each polymer solution.

(a) includes the traces for $p(\text{Nce})_{36}\text{ac}$ while (b) includes the traces for $p(\text{NceNme})_{18}\text{ac}$. For both polymers the solutions with 10x and 30x NaOH equivalents show peaks while the solutions with 1x and 3x NaOH equivalents monotonically increase.

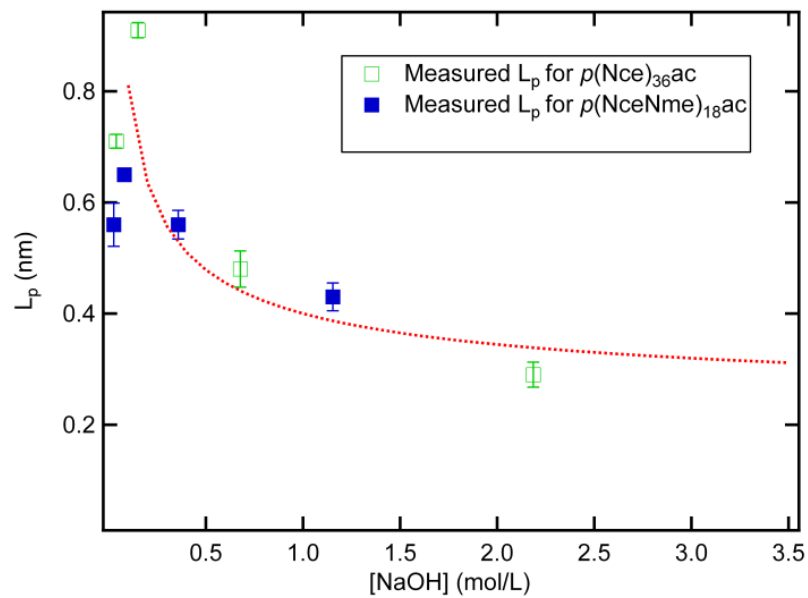


Figure 3.3. The persistence lengths for $p(\text{Nce})_{36}\text{ac}$ and $p(\text{NceNme})_{18}\text{ac}$.
 The red dashed line is the best fit line using the expression $L_p = f + g[\text{NaOH}]^{-0.5}$

Table 3.3. Best fit values for bare bending constant K^0 and chain persistence length L_p using Equations 2 and 3.

Sample	Polymer	K^0	L_p (fit)	L_p (nm)
1	$p(\text{Nce})_{36}$	0.975	0.71	0.71
2	$p(\text{Nce})_{36}$	0.975	0.87	0.91
3	$p(\text{Nce})_{36}$	0.975	0.45	0.48
4	$p(\text{Nce})_{36}$	0.975	0.38	0.29
5	$p(\text{NceNme})_{18}$	1.43	0.52	0.56
6	$p(\text{NceNme})_{18}$	1.43	0.69	0.65
7	$p(\text{NceNme})_{18}$	1.43	0.51	0.56
8	$p(\text{NceNme})_{18}$	1.43	0.47	0.43

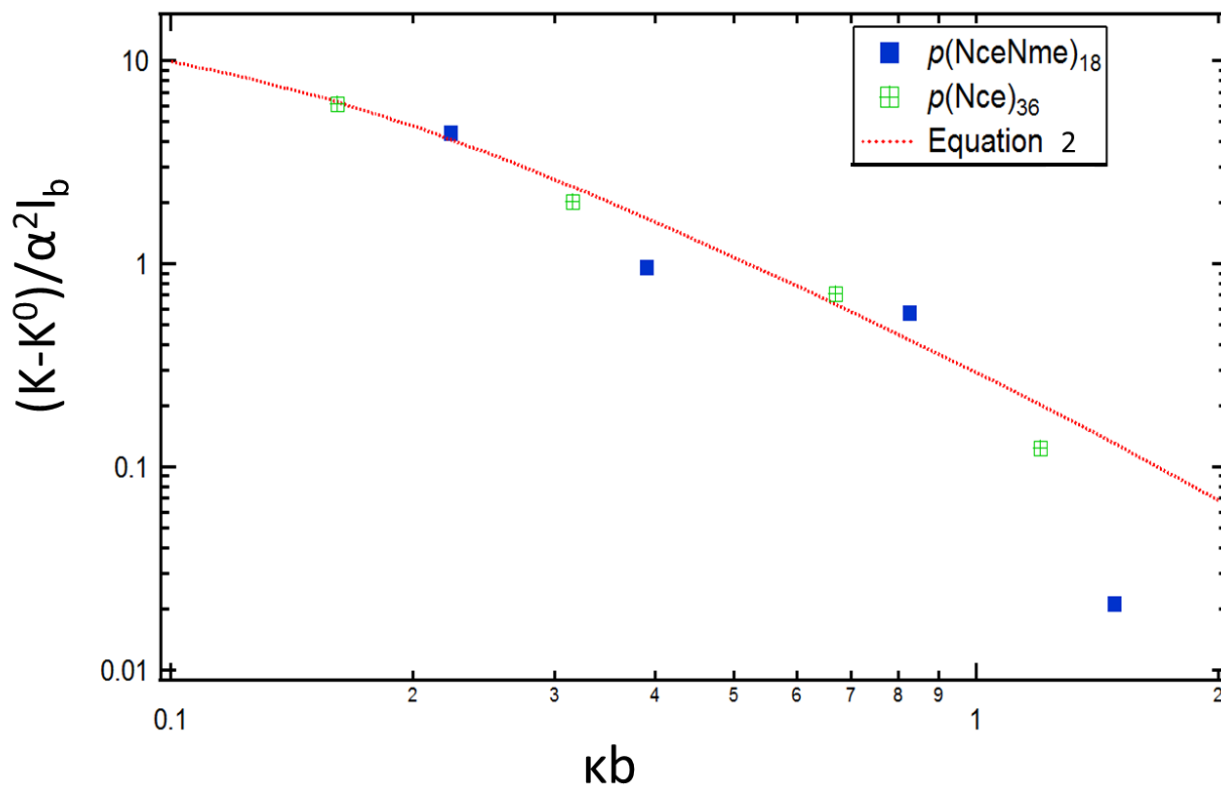


Figure 3.4. Electrostatic contribution to the chain persistence length as a function of the reduced Debye screening length.

Experimental data for $p(\text{NceNme})_{18}$ are shown as open circles and those for $p(\text{Nce})_{36}$ shown as filled circles. The solid line corresponds to equation 2.

3.4 Conclusions

In conclusion, a model system has been used to study the effect of charge placement on the single chain conformation of a polyelectrolyte in aqueous solution. Polypeptoids containing ionizable groups on every monomer or containing ionizable groups on every other monomer were synthesized and their single chain conformation analyzed using small angle neutron scattering. It was shown that at low salt concentrations, the polypeptoid with closer charge placements had a higher persistence length due to the increased level of ionic repulsion between groups. At higher salt concentrations, both polymers showed a decrease in persistence length that scales as the inverse square root of the positive ion concentration, matching previous theoretical and experimental work. These results will aid in the understanding of biopolymer behavior, and will also help in designing and understanding polypeptoid mesoscale assemblies.

3.5. Acknowledgements

We gratefully acknowledge funding from the Office of Naval Research via a Presidential Early Career Award in Science and Engineering. A.M.R. and H.K.M also gratefully acknowledge the National Science Foundation and the Department of Defense for graduate fellowships (respectively). Polypeptoid synthesis and associated chemical characterization were performed at the Molecular Foundry, a Lawrence Berkeley National Laboratory user facility supported by the Office of Science, Office of Basic Energy Sciences, U.S. Department of Energy, under Contract DE-AC02-05CH11231. A portion of the neutron scattering in this work is based on activities at the NIST Center for Neutron Research, which is supported in part by the National Science Foundation under Agreement No. DMR-0454672. Certain trade names and company products are identified to adequately specify the experimental procedure. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products are necessarily best for the purpose. The authors thank Dr. Steven Kline for assistance on SANS data collection. A portion of this research was also performed at Oak Ridge National Laboratory. The authors thank Dr. Volker S. Urban at Oak Ridge National Laboratory for assistance on SANS data collection. The SANS studies at Oak Ridge National Laboratory's Center for Structural Molecular Biology were supported by the Office of Biological and Environmental Research, using facilities supported by the DOE, managed by UT-Battelle, LLC, under Contract No. DE-AC05-00OR22725.

3.6. References

1. Odijk, T., Polyelectrolytes near rod limit. *J. Polym. Sci. Pt. B-Polym. Phys.* **1977**, *15* (3), 477-483.
2. Skolnick, J.; Fixman, M., Electrostatic persistence length of a wormlike polyelectrolyte. *Macromolecules* **1977**, *10* (5), 944-948.
3. Dobrynin, A. V.; Colby, R. H.; Rubinstein, M., Scaling theory of polyelectrolyte solutions. *Macromolecules* **1995**, *28* (6), 1859-1871.
4. Manning, G. S., Limiting laws and counterion condensation in polyelectrolyte solutions .I. colligative properties. *J. Chem. Phys.* **1969**, *51* (3), 924-&.
5. Prabhu, V. M., Counterion structure and dynamics in polyelectrolyte solutions. *Curr. Opin. Colloid Interface Sci.* **2005**, *10* (1-2), 2-8.

6. Spiteri, M. N.; Boue, F.; Lapp, A.; Cotton, J. P., Persistence length for a PSSNa polyion in semidilute solution as a function of the ionic strength. *Phys. Rev. Lett.* **1996**, *77* (26), 5218-5220.
7. Ullner, M.; Jonsson, B.; Peterson, C.; Sommelius, O.; Soderberg, B., The electrostatic persistence length calculated from Monte Carlo, variational and perturbation methods. *J. Chem. Phys.* **1997**, *107* (4), 1279-1287.
8. Dobrynin, A. V.; Rubinstein, M., Theory of polyelectrolytes in solutions and at surfaces. *Progress in Polymer Science* **2005**, *30* (11), 1049-1118.
9. Katchalsky, A.; Kunzle, O.; Kuhn, W., Behavior of polyvalent polymeric ions in solution. *Journal of Polymer Science* **1950**, *5* (3), 283-300.
10. Odijk, T., Possible scaling relations for semidilute polyelectrolyte solutions. *Macromolecules* **1979**, *12* (4), 688-693.
11. Odijk, T.; Houwaart, A. C., Theory of excluded-volume effect of a polyelectrolyte in a 1-1 electrolyte solution. *J. Polym. Sci. Pt. B-Polym. Phys.* **1978**, *16* (4), 627-639.
12. Odijk, T.; Mandel, M., Influence of chain-flexibility on colligative properties of polyelectrolyte solutions. *Physica A* **1978**, *93* (1-2), 298-306.
13. Sobel, E. S.; Harpst, J. A., Effects of Na⁺ on the persistence length and excluded volume of T7 bacteriophage DNA. *Biopolymers* **1991**, *31* (13), 1559-1564.
14. Wang, L. X.; Bloomfield, V. A., Small-angle x-ray scattering of semidilute rodlike DNA solutions-polyelectrolyte behavior. *Macromolecules* **1991**, *24* (21), 5791-5795.
15. Baumann, C. G.; Smith, S. B.; Bloomfield, V. A.; Bustamante, C., Ionic effects on the elasticity of single DNA molecules. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94* (12), 6185-6190.
16. Makita, N.; Ullner, M.; Yoshikawa, K., Conformational change of giant DNA with added salt as revealed by single molecular observation. *Macromolecules* **2006**, *39* (18), 6200-6206.
17. Buhler, E.; BouÅ©, F. o., Chain Persistence Length and Structure in Hyaluronan Solutions:â€ Ionic Strength Dependence for a Model Semirigid Polyelectrolyte. *Macromolecules* **2004**, *37* (4), 1600-1610.
18. Murnen, H. K.; Rosales, A. M.; Jaworsk, J. N.; Segalman, R. A.; Zuckermann, R. N., Hierarchical Self-Assembly of a Biomimetic Diblock Copolypeptoid into Homochiral Superhelices. *J. Am. Chem. Soc.* **2010**, *132* (45), 16112-16119.
19. Sanii, B.; Kudirka, R.; Cho, A.; Venkateswaran, N.; Olivier, G. K.; Olson, A. M.; Tran, H.; Harada, R. M.; Tan, L.; Zuckermann, R. N., Shaken, Not Stirred: Collapsing a Peptoid Monolayer To Produce Free-Floating, Stable Nanosheets. *J. Am. Chem. Soc.* **2011**, *133* (51), 20808-20815.
20. Kudirka, R.; Tran, H.; Sanii, B.; Nam, K. T.; Choi, P. H.; Venkateswaran, N.; Chen, R.; Whitlam, S.; Zuckermann, R. N., Folding of a Single-Chain, Information-Rich Polypeptoid Sequence into a Highly Ordered Nanosheet. *Biopolymers* **2011**, *96* (5), 586-595.
21. Nam, K. T.; Shelby, S. A.; Choi, P. H.; Marciel, A. B.; Chen, R.; Tan, L.; Chu, T. K.; Mesch, R. A.; Lee, B. C.; Connolly, M. D.; Kisielowski, C.; Zuckermann, R. N., Free-floating

ultrathin two-dimensional crystals from sequence-specific peptoid polymers. *Nat. Mater.* **2010**, *9* (5), 454-460.

22. Figliozzi, G. M.; Goldsmith, R.; Ng, S. C.; Banville, S. C.; Zuckermann, R. N., Synthesis of N-substituted glycine peptoid libraries. *Methods Enzymol.* **1996**, *267*, 437-447.

23. Rosales, A. M.; Murnen, H. K.; Zuckermann, R. N.; Segalman, R. A., Control of Crystallization and Melting Behavior in Sequence Specific Polypeptoids. *Macromolecules* **2010**, *43* (13), 5627-5636.

24. Kline, S. R., Reduction and analysis of SANS and USANS data using IGOR Pro. *J. Appl. Crystallogr.* **2006**, *39*, 895-900.

25. Klooster, N. T. M.; Vandertouw, F.; Mandel, M., Solvent effects in polyelectrolyte solutions .1. Potentiometric and viscometric titration of poly(acrylic acid) in methanol and counterion specificity. *Macromolecules* **1984**, *17* (10), 2070-2078.

26. Starodoubtsev, S. G.; Khokhlov, A. R.; Sokolov, E. L.; Chu, B., Evidence for polyelectrolyte ionomer behavior in the collapse of polycationic gels. *Macromolecules* **1995**, *28* (11), 3930-3936.

27. Khokhlov, A. R.; Kramarenko, E. Y., Weakly charged polyelectrolytes: Collapse induced by extra ionization. *Macromolecules* **1996**, *29* (2), 681-685.

28. Winkler, R. G.; Gold, M.; Reineker, P., Collapse of polyelectrolyte macromolecules by counterion condensation and ion pair formation: A molecular dynamics simulation study. *Phys. Rev. Lett.* **1998**, *80* (17), 3731-3734.

29. Pedersen, J. S.; Schurtenberger, P., Scattering functions of semiflexible polymers with and without excluded volume effects. *Macromolecules* **1996**, *29* (23), 7602-7612.

30. Rosales, A. M.; Murnen, H. K.; Kline, S. R.; Zuckermann, R. N.; Segalman, R. A., Determination of the persistence length of helical and non-helical polypeptoids in solution. *Soft Matter* **2012**, *8* (13), 3673-3680.

31. Ghosh, S.; Xiao, L.; Reed, C. E.; Reed, W. F., Apparent persistence lengths and diffusion behavior of high molecular weight hyaluronate. *Biopolymers* **1990**, *30* (11-12), 1101-1112.

32. Tricot, M., Comparison of experimental and theoretical persistence length of some polyelectrolytes at various ionic strengths. *Macromolecules* **1984**, *17* (9), 1698-1704.

33. Barrat, J. L.; Joanny, J. F., Persistence length of polyelectrolyte chains. *Europhys. Lett.* **1993**, *24* (5), 333-338.

34. Dobrynin, A. V.; Carrillo, J. M. Y.; Rubinstein, M., Chains Are More Flexible Under Tension. *Macromolecules* **2010**, *43* (21), 9181-9190.

35. Dobrynin, A. V.; Carrillo, J. M. Y., Swelling of biological and semiflexible polyelectrolytes. *J. Phys.-Condes. Matter* **2009**, *21* (42), 11.

Chapter 4. Experimental Validation of the HP Model for Globule Formation

Reproduced with permission from Alexei R. Khokhlov, Pavel G. Khalatur, Rachel A. Segalman, Ronald N. Zuckermann, 2012 in preparation

Understanding the driving forces for the collapse of a polymer chain from a random coil to a globule would be invaluable in enabling scientists to predict the folding of polypeptide sequences into defined tertiary structures. The HP model considers hydrophobic collapse to be the major driving force for protein-folding. However, due to the inherent presence of chirality and hydrogen bonding in polypeptides, it has been difficult to experimentally test the ability of hydrophobic forces to independently drive structural transitions. In this work, we use polypeptoids, which lack backbone hydrogen bonding and chirality, to probe the exclusive effect of hydrophobicity on the coil to globule collapse. Two sequences containing the same composition of only hydrophobic “H” *N*-methylglycine and polar “P” *N*-(2-carboxyethyl)glycine monomers are shown to have very different globule collapse behaviors due only to the difference in their monomer sequence. As compared to a repeating sequence with an even distribution of H and P monomers, a designed protein-like sequence collapses into a more compact globule in aqueous solution as evidenced by small angle x-ray scattering, dynamic light scattering and probing with environmentally-sensitive fluorophores. The free energy change for the coil to globule transition was determined by equilibrium denaturant titration with acetonitrile. Using a two-state model, the protein-like sequence is shown to have a much greater driving force for globule formation, as well as a higher *m*-value, indicating increased cooperativity for the collapse transition. This difference in globule collapse behavior validates the ability of the HP model to describe structural transitions based solely on hydrophobic forces.

4.1. Introduction

Protein folding is an inherently complex process involving a multiplicity of forces and interactions. Predicting a tertiary structure from a polypeptide sequence has presented a long-standing challenge to the scientific community. It is a fairly well established principle that hydrophobicity serves as one of the most important driving forces for protein assembly. Capitalizing on this idea, Ken Dill and others have developed a computational framework known as the HP model in which only two types of monomers, hydrophobic (H) and polar (P) are considered¹⁻³. With this dramatically reduced set of interactions, polymer sequences can be computationally designed to fold into defined structures³. A key experimental realization of this theory has been work by Michael Hecht’s group to analyze the HP sequence patterns in combinatorial variants of existing proteins. For example, a sequence known to form a 4-helix bundle was randomized while still maintaining the same pattern of H and P residues⁴. Overwhelmingly, the mutated proteins still formed 4-helix bundles, demonstrating that the particular side chain was not as important as the hydrophobic or polar character of the amino acid.

Given that hydrophobicity has been shown to have such a strong effect on protein folding, efforts have focused on ways to design polypeptide sequences that include targeted hydrophobic and polar regions to induce folding⁵⁻⁹. However, it is difficult to isolate the hydrophobic interactions, making *de novo* protein design a complex process in which the effect of different forces can be convoluted. Scientists have turned to simpler transitions such as globule formation

to try and understand the forces that influence molecular assembly. Globule formation has been described as one of the first steps along the path to a folded protein¹⁰, and hydrophobic interactions have been shown to have a large impact on this transition^{11, 12}. Many groups have used theoretical means to postulate how a monomer sequence in the HP model affects globule formation, including the effect of the relative fraction of hydrophobic monomers to polar monomers¹³, the degree of hydrophobicity of the hydrophobic monomers¹⁴, the overall length of the chain¹⁵, the importance of large contiguous sections of H or P monomers¹⁶, and even the ability of ionic interactions to impact the collapse process¹⁷. One of the most extensive efforts has been the theoretical work performed by Khokhlov and Khalatur (KK) demonstrating that copolymers with blocky (or protein-like) distributions of monomers form more stable globules than corresponding random sequences where the monomers are more evenly distributed throughout the chain^{11, 12, 18}. They also predicted that the coil to globule transition for the protein-like molecules were sharper and occurred at a higher temperature than those for the random sequences. Both results showed a clear difference in collapse behavior based solely on the sequence of the molecule.

Experimental efforts in this area have lagged behind computational efforts¹⁹⁻²² due to the synthetic challenges inherent in generating precise sequence-specific polymers in quantity. The use of polypeptides would seem obvious, but due to their inherent chirality and backbone hydrogen bonding, it is difficult to directly study the impact of hydrophobic interactions. Alternatively, in most synthetic polymer systems where it might be possible to create a system with isolated HP interactions, it is nearly impossible to obtain the level of sequence control necessary to probe the effect of H and P monomer sequence on the coil to globule transition. Genzer and coworkers elegantly demonstrated the chemical labeling of surface exposed monomers in a collapsed polystyrene globule, generating protein-like sequences²³. However, assessment of exact chain sequence information was not possible.

Given the difficulties mentioned above, polypeptoids or poly N-substituted glycines are an ideal model system for understanding sequence effects on the hydrophobic collapse of polymer chains. Their stepwise submonomer synthesis is efficient, with 99% or greater conversion for most monomer additions and provides sequence control, allowing for the creation of sequence-specific chains in the 50 monomer range with excellent precision. In addition to their synthetic tractability, polypeptoids are known to possess flexible backbones²⁴, and self-assemble into protein-like structures in aqueous solution²⁵⁻³⁰ making them ideal candidates for studying fundamental self-assembly in the form of coil to globule transitions. The peptoid backbone is nearly identical to that of polypeptides, making chain measurements biologically relevant. Finally, and perhaps most importantly, polypeptoids also have no backbone hydrogen bonding or chirality, allowing the study of hydrophobic forces in isolation, making them an excellent model system for experimentally validating the HP model.

In this work, we designed and synthesized two polypeptoid 100mer sequences containing only the H and P monomers in order to probe their coil to globule transitions (Figure 4.1). One of the sequences was designed using the KK method to be a “protein-like” sequence that contained longer stretches of both the hydrophobic and polar monomers. The other “repeating” sequence contained the shortest possible stretches of both monomers. Small angle x-ray scattering (SAXS), dynamic light scattering (DLS), and environmentally-sensitive fluorescence probe measurements were used to show that the protein-like sequence collapses into a tighter globule in aqueous solution than that formed by the repeating sequence. In addition, the transition of the protein-like molecule from the globule state to the coil state was shown to exhibit a significantly

higher unfolding free energy and that the transition is more cooperative than that of the repeating sequence molecule, indicating the importance of sequence on the folding behavior of a biomimetic molecule.

4.2. Experimental Section

Sequence design

The sequence generation was carried out according to methods previously published^{11, 12, 18}. The length and specific side chains were designed specifically for this synthetic system. A sequence design procedure aimed at obtaining the protein-like polypeptide chain containing 20% aspartic acid residues denoted as “P” and 80% alanine residues denoted as “H” was carried out. The resulting polypeptide chain was then translated into a polypeptoid chain by using an *N*-methyl side chain for the H residue and an *N*-2-carboxyethyl side chain as the P residue. First, a homopolymer globule from 100 H units was generated, using Accelrys Discovery Studio 2.5 and all-atom molecular dynamics with the AMBER96 force field. To construct a target protein-like H/P sequence, we perform globule surface "coloring". By "coloring" we mean the change of a given monomer type: monomer units in the center of the globule are assigned to be H-type units, while monomer units belonging to globular surface are assigned to be P-type units. The "coloring" is applied to the units mostly exposed to water at the surface of the globule. After the formation of initial H-P sequence, we allow the macromolecule to undergo a coil-to-globule transition and then we "recolor" a refolded globule. Each globular structure is obtained during 1 ns simulated annealing run (we start from 1000 K and then cool the system to 300 K). After that the system is again relaxed during another 1 ns run. In this way, we obtain a heteropolymer chain with a new primary sequence and all steps described above are performed again for the chain with a new sequence. After several attempts, we reach the regime when practically all of P units remain robustly located at the globular surface even after refolding, while the globular core is composed mostly of H units.

Synthesis and conjugation

Polypeptoids were synthesized on a commercial robotic synthesizer Aapptec Apex 396 on 100 mg of Rink amide polystyrene resin (0.6 mmol/g, Novabiochem, San Diego, CA) using the procedure previously detailed^{31, 32}. All primary amine monomers, solvents, and reagents were purchased from commercial sources and used without further purification. The β -alanineOtbu-HCl was purchased from ChemImpex and used after freebasing by extraction from dichloromethane (DCM) and basic water. The resulting compound was confirmed by ¹H NMR. All primary amine submonomers were dissolved in *N*-methyl pyrrolidinone at a concentration of 1.5 M and all displacement times were 60 minutes for the first 15 monomer additions cycles 90 minutes for the next 15 and 120 minutes for the remainder. All polypeptoids were acetylated on the resin and purified using reverse phase HPLC as previously described³¹. The mass and purity were confirmed using reverse phase analytical HPLC and MALDI and representative traces and spectra are shown below. Upon the synthesis of the 50mer components, the two molecules were then clicked together using an alkyne-azide reaction. An alkyne group was added to the N-terminus of the 50mer polypeptoid still attached to the polystyrene resin by the addition of a propargyl amine submonomer in a 51st monomer addition cycle. For the other component, an azide group was added by first bromacylating the polymer while it was still attached to the polystyrene resin and then substituting for the bromine using sodium azide. The click reaction (scheme 4.1) was performed by reacting the propargylated compound **A**, (23 mM)

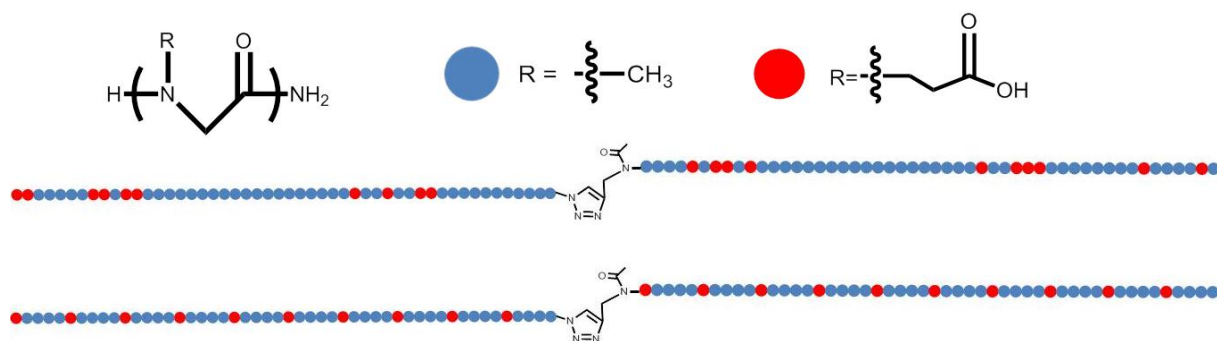
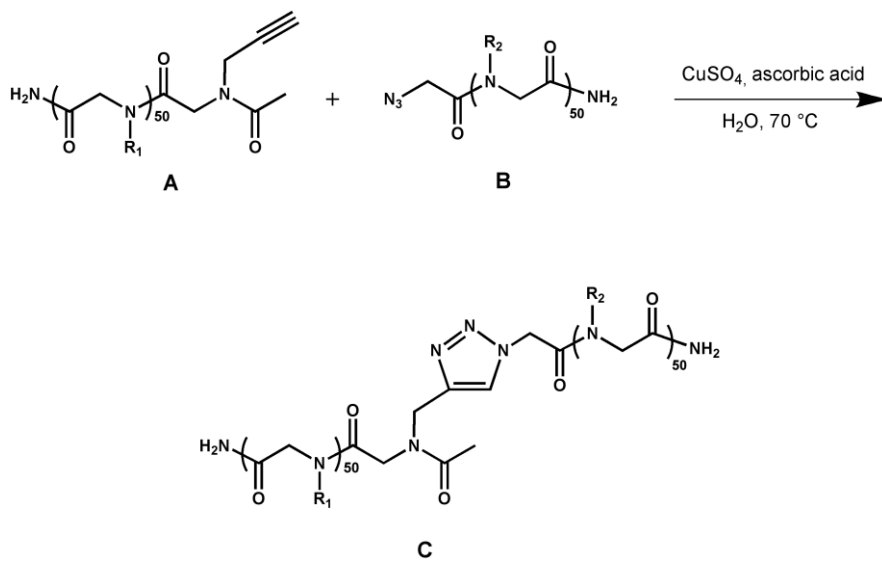


Figure 4.1. Protein-like and repeating sequence polypeptoid 100mers.

The polypeptoids were synthesized by clicking two HPLC-purified 50mers together. Each monomer is represented by a circle where the red circles are the hydrophilic and polar *N*-(2-carboxyethyl)glycine (P) monomer while the blue circles are hydrophobic *N*-methylglycine (H) monomers. Compound 1 is composed of a protein-like sequence of monomers with block sections of each type of monomer, while compound 2 has an even distribution of monomers. Both molecules have an identical composition of exactly 80 hydrophobic monomers and 20 hydrophilic monomers and a molecular weight of 8517 g/mol.



Scheme 4.1. The click reaction was employed to ligate two 50-mers to create a 100mer of defined sequence.

Table 4.1. HPLC and mass spectrometry of 50mer and 100mer peptoids.

Compound	Length	Observed mass/Expected mass	% Purity*
1a	50	4290.3/4288.4	98%
1b	50	4232.0/4232.4	97%
2a	50	4289.6/4288.4	97%
2b	50	4235.9/4232.4	96%
1	100	8512/8517.5	97%
2	100	8514/8517.5	99%

*after HPLC purification.

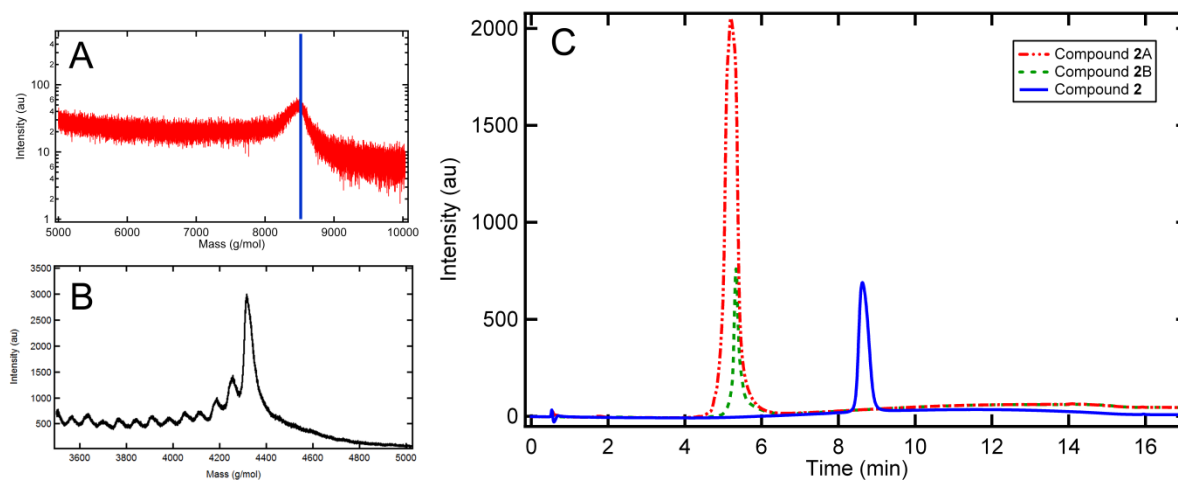


Figure 4.2. Representative analytical traces and MALDI spectra.

MALDI spectra of peptoid 100mer (a) shows a molecular weight of 8517 while the spectra of a propargylated 50mer, compound **2a** (b), shows a molecular weight of 4289. The analytical traces show both 50mers are eluted after 5 minutes in a 0-50% gradient over 17 minutes while the 100mer eluted at close to 9 minutes.

with 2 equivalents of azide-modified compound **B**, (46 mM) in 200 μL of water. Copper(II) sulfate (5x, 115 mM) and ascorbic acid (6x, 138 mM) were added and the solution was mixed at 70°C for 24 hours. The excess copper was removed by stirring over basic alumina for 2 hours. The resulting compound was purified by a 0-30% acetonitrile gradient on reverse phase HPLC. Representative analytical HPLC and MALDI analysis of the 50mer and 100mer peptoids are shown in Figure 4.2.

Self-Assembly Solutions

In order to probe aqueous assembly, the molecules were dissolved at 1 mg/mL (0.12 mM) in 25 mM Tris HCl pH 8 buffer. The solutions were sonicated for approximately 30 seconds to ensure dissolution of the molecules and were then annealed at 70°C for 4 hours.

Two-State Model

The equation below was fit to the R_g measurements with $R_{g, \text{coil}}$ and $R_{g, \text{globule}}$ calculated by averaging the 3 points at either end of the curve. The percentage of acetonitrile in the solution is represented by [acn]. R is the ideal gas constant (kcal/mol K^{-1}) and T is temperature in Kelvin. $\Delta G_{\text{CG}}(\text{H}_2\text{O})$ is the free energy change for the coil to globule transition and m is the denaturant

dependence of free energy. These two values were calculated by a least-squares fit of the above equation to the measured data.

Fluorescence probe measurements

Nile red (9-diethylamino-5-benzo[α]phenoxazinone), was added at 2 μM concentration to the self assembly solutions. The Nile red was purchased from Sigma Aldrich and used as is without further purification. The solution was placed into a quartz cuvette inside a Jobin Yvon FluoroMax Fluorometer and excited at 586 nm. The resulting emission spectrum was then collected at a scan rate of 1 nm/s.

Dynamic Light Scattering

The hydrodynamic radius of the molecules was characterized using dynamic light scattering on a Viscotek DLS model #802. The solution was loaded into a quartz cuvette and correlation functions were obtained for a period of 4 s. A diffusion coefficient was found through the correlation functions and Stokes Einstein equation used to find a hydrodynamic radius assuming a spherical particle.

Small Angle X-Ray Scattering

Small angle x-ray scattering (SAXS) was performed at Beamline 1-4 at Stanford Synchrotron Radiation Laboratory and Beamline 7.3.3 at the Advanced Light Source at Lawrence Berkeley National Laboratory. At the ALS, the beamline was configured with an X-ray wavelength of $\lambda = 1.240 \text{ \AA}$ and focused to a spot size of 50 μm by 300 μm . Two-dimensional scattering patterns were collected on an ADSC CCD detector with an active area of 188 mm by 188 mm. The isotropic scattering patterns were radially averaged, and the scattering intensity was corrected with the post-ion chamber intensity using Nika version 1.18. At SSRL, the beamline was

configured with an X-ray wavelength of $\lambda = 1.488 \text{ \AA}$ and focused to a 0.5 mm diameter spot. A single quadrant of a two-dimensional scattering pattern was collected on a CCD detector with an active area of 25.4 by 25.4 mm. The two-dimensional profiles were radially averaged and corrected for detector null signal, dark current, and empty cell scattering. In both cases, the data was plotted on a Guinier plot and fitted to obtain a radius of gyration.

4.3. Results and Discussion

In this study, two 100mer polypeptoids of exact sequence and absolute monodispersity were synthesized, one with a protein-like sequence as designed using KK theoretical methods^{11,12}, and another with an even distribution of “H” and “P” monomers (Figure 4.1). The simulation work leading to the generation of the protein-like sequence for compound **1** is detailed in the Experimental Methods as well as the synthetic details. The H monomer, an *N*-methylglycine (or sarcosine), was chosen to be sufficiently hydrophobic to drive globule formation, but not too hydrophobic such that the polymer would become insoluble in aqueous solution. The P monomer was designed to be polar and negatively charged and thus an *N*-(2-carboxyethyl)glycine side chain, similar to glutamic acid, was chosen. Probing protein folding on a scale relevant to biological molecules requires chains of at least 100 residues. However, stepwise synthesis of a 100mer is difficult even with the high coupling efficiencies of submonomer peptoid synthesis. Therefore segment condensation was used to link two HPLC purified solid phase synthesized 50mers to yield a peptoid 100mer, the longest sequence-specific polypeptoid ever created. The peptoid polymers were dissolved at a concentration of 1 mg/mL (120 μM) in 25mM Tris HCl, pH 8.0. It was expected that in aqueous solution, the molecules would collapse into a globule, and upon the introduction of a nonpolar solvent, such as acetonitrile, they would unfold their hydrophobic cores and the transition from globule to random coil could be analyzed.

The first objective in characterizing these compounds was to determine their radii of gyration, as this provides much insight into the globule conformation. Using small angle x-ray scattering, the R_g of compound **1**, the protein-like sequence, in buffer was determined to be $2.1 \pm 0.09 \text{ nm}$ while the R_g of compound **2**, the repeating sequence, was $2.6 \pm 0.21 \text{ nm}$ in buffer (Figure 4.3). These values are slightly larger than the R_g for a protein of an analogous size. For example, crotapotin, a protein with a molecular weight of 9.0 kDa has a measured R_g of 1.4 nm ³³. This is likely due to the tighter packing of a folded protein structure, and is not unexpected since the polyanionic nature of these polypeptoids will result in increased electrostatic repulsion in smaller globule sizes. Native polypeptides typically have a mix of negative and positive charges yielding denser structures due to electrostatic interactions. The smaller R_g of compound **1** as compared to the R_g of compound **2**, even though the two polymers are of identical molecular weight and chemical composition, indicates that it is more tightly folded in its globule state. The longer blocks of hydrophobic and hydrophilic residues in compound **1** allow the molecule to collapse more fully while compound **2** remains in a loosely collapsed state resulting in a larger R_g .

Equilibrium denaturant titration with increasing concentrations of a nonpolar solvent such as acetonitrile was expected to cause the molecules to unfold from a globule to their coil state. Therefore, solutions at varying percentages of acetonitrile were made and the R_g for each solution determined using SAXS. Plotting the R_g of each compound against the acetonitrile concentration in solution (Figure 4.3), the unfolding of each polymer is readily apparent. Both molecules remain collapsed at low acetonitrile concentrations with nearly constant R_g 's and, upon the addition of acetonitrile, unfold to nearly identical R_g 's of about 3.5 nm. Compound **2** begins to show an increased R_g around 7M acetonitrile after which its R_g slowly increases,

demonstrating a gradual unfolding transition. In contrast, the protein-like compound **1** stays folded until about 10M acetonitrile, and the subsequent transition is much sharper, indicating that the globule is more stable and the unfolding transition is highly cooperative.

The difference in the two unfolding transition curves as seen in Figure 4.3 can be approximated by fitting the R_g data using a two-state model as described by Santoro and Bolen³⁴ and as previously utilized for folded polypeptoids²⁶. Using the method detailed in supplemental information, $\Delta G_{CG}(H_2O)$, the free energy of the coil to globule transition, and m , the denaturant dependence of that folding free energy $\Delta G_{CG}(H_2O)$, were calculated (Table 4.2). The larger $\Delta G_{CG}(H_2O)$ for compound **1** (42.5 kcal/mol vs. 7.7 kcal/mol) can be attributed to a lower free energy for the globule state of compound **1** as compared to compound **2**. The unfolding energy is therefore significantly larger ($\Delta\Delta G_{CG}(H_2O)$ is 34.8 kcal/mol). The value of m for compound **1** is also significantly greater than that for compound **2** (3.6 kcal mole⁻¹ M⁻¹ vs. 0.7 kcal mole⁻¹ M⁻¹), which can be attributed to the difference in the coil to globule transition behavior of the protein-like molecule. Larger m values are associated with more cooperativity in folding as well as increased hydrophobic burial within a folded site³⁵. The protein-like molecule therefore has increased hydrophobic burial due to the ability of the chain to fold and accommodate the hydrophobic residues at the center of its collapsed globule. The values of $\Delta G_{CG}(H_2O)$ seen here are significantly larger than those usually seen for protein-folding (which are usually < 10 kcal/mol), although previous researchers have shown a value of 22 kcal/mol for a molten globule transition³⁶ indicating that perhaps the coil-globule transition has an inherently higher energy difference than that of protein-folding. Regardless, the difference in $\Delta G_{CG}(H_2O)$ for compounds **1** and **2** matches well with the behavior previously predicted by the modeling work of KK^{11,12}, where thermal unfolding was used to compare globule stability. These simulations predicted a lower energy for a collapsed protein-like molecule as well as an increased unfolding cooperativity, qualitatively matching the measurements made here.

In conjunction with the radius of gyration, the hydrodynamic radius, R_h , as measured by dynamic light scattering, can be used to probe the shape of the peptoid polymer chains. In aqueous buffer solution, the R_h of compound **1** is 3.4 nm while the R_h of compound **2** is 3.3 nm (Figure 4.4). While these values are very similar for the two molecules, the characteristic ratio of R_g/R_h , defined as ρ , can reveal information about the shape of the polymer³⁷. This value is theoretically predicted for both a tightly compacted spherical globule (0.77) and for an expanded coil (1.5-2)³⁸. The R_h was measured in aqueous conditions as well as in 80% acetonitrile solution and ρ was subsequently calculated (Table 4.3). Interestingly, for both molecules the R_g increased with the introduction of acetonitrile while the R_h decreased slightly. The increase in R_g is expected, as the mean distance between the N and C termini increases as they transition to the coil state. The decrease in R_h is most likely due to the non-spherical nature of the coils, while the calculation of R_h uses the Stokes-Einstein equation and therefore assumes a sphere. There is also an increase in the breadth of the peak for the R_h in the case of acetonitrile solutions, which is likely due to the coil state being more polydisperse than the globule state. Importantly, the ratio ρ indicates that for both compounds, a globule is formed in aqueous solution with ratios close to 0.77 and a coil conformation is adopted in the acetonitrile solution with ratios above 1.5.

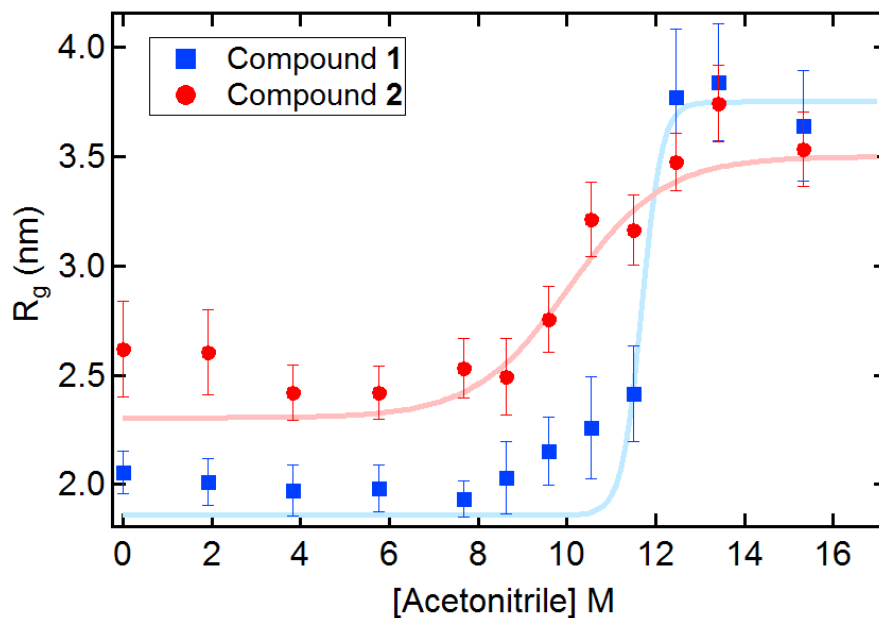


Figure 4.3. Equilibrium denaturant titration of peptoid globules with acetonitrile.

The radius of gyration as determined by SAXS was determined as a function of acetonitrile concentration for the protein-like compound **1** and repeating sequence compound **2**. The polymers were dissolved at 1 mg/mL (120 μ M) in 25 mM Tris, pH 8.0. The blue squares are for the compound **1** while the red circles are for compound **2**. The lines were fit using a two state model.

Table 4.2. Calculated values from a two-state model

The model was fit to the R_g values obtained from equilibrium acetonitrile titration of peptoids **1** and **2**. Higher values of both $\Delta G_{CG}(\text{H}_2\text{O})$ and m for compound 1 indicate increased driving force and cooperativity for the collapse transition.

Compound	$\Delta G_{CG}(\text{H}_2\text{O})$ (kcal/mol)	m (kcal/mol M ⁻¹)
1	42.5 ± 1.0	3.63 ± 0.08
2	7.7 ± 0.5	0.73 ± 0.05

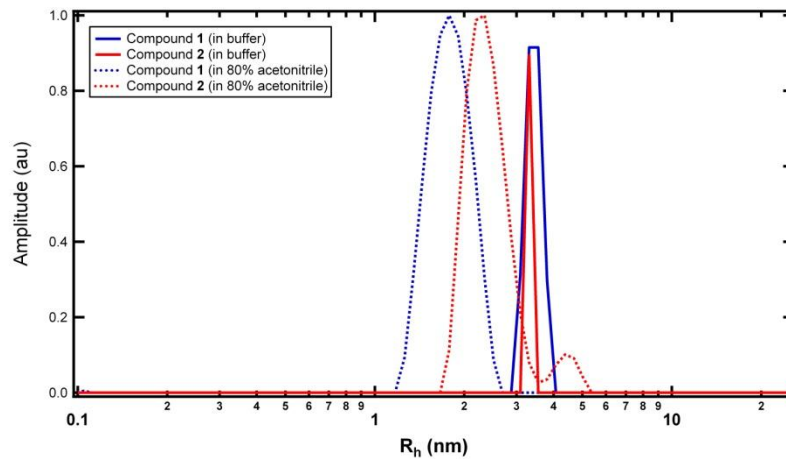


Figure 4.4. Hydrodynamic radius as measured using dynamic light scattering. The blue lines are for compound 1 while the red lines are compound 2. The R_h 's that are measured are then used to calculate ρ , the ratio of R_g/R_h .

Table 4.3. The characteristic ratio, ρ , for each compound in buffer and in 80% acetonitrile.

Compound	ρ (in buffer)	ρ (in 80% acetonitrile)
1	0.62	1.97
2	0.79	1.60

Further information about the nature of the globules formed can be gained through the use of environmentally-sensitive fluorescent probes. The fluorescence behavior of dyes such as Nile Red or ANS is strongly dependent upon the polarity of its surrounding environment³⁹, leading to its use as a probe for hydrophobic cores in molten globule or folded proteins⁴⁰. The fluorescence increases when the dye is located in a hydrophobic region, allowing the study of the arrangement of H and P residues within a collapsed structure. In this study, Nile red was added at 2 μM to buffer solutions of both compounds **1** and **2** (120 μM) as well as a control solution of a homopolypeptoid 36mer containing only relatively polar *N*-2-methoxyethyl side chains. In Figure 3.5, the emission peaks for each of these polymers is shown. As expected, the control polymer containing only methoxyethyl side chains shows no emission peaks due to its well solvated conformation in aqueous solution. In contrast, compound **1** has a strong emission peak, indicating that the hydrophobic core of the collapsed globule is the most isolated from the hydrophilic environment. Compound **2** shows weaker fluorescence (~20% of compound **1**) due to the fact that it is forming a globule in solution as evidenced by the value of ρ indicating globule formation (Table 4.3). There is essentially no shift of the peak maximum (compound **1**, 663 nm; compound **2**, 662 nm) suggesting that the signal is still coming from Nile red molecules located within a hydrophobic region in the compound **2** case, but the region is providing less intense signal than that from compound **1**. This supports the conclusion from the previous scattering (SAXS) measurements that a globule is formed from compound **2**. However, this globule appears to be less well folded, leading to a lower intensity of emission as compared to compound **1**. The larger blocks of H and P monomers along the protein-like chain allow it to fold into a more intact globule, while the repeating sequence cannot adopt a conformation in which the hydrophobic residues can be isolated and thus shows a lower emission peak from the Nile red.

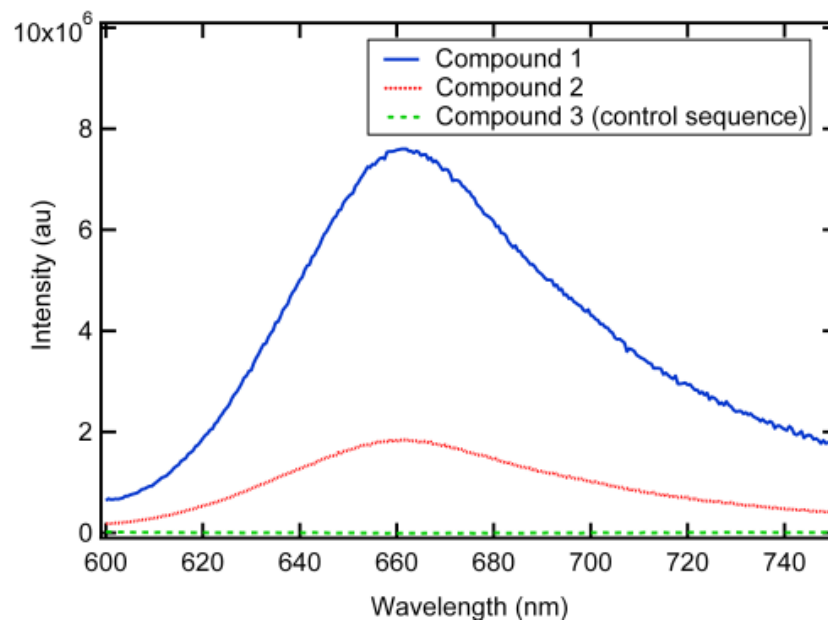


Figure 4.5. Nile red fluorescence showing an increased emission peak from the protein-like sequence. All solutions were excited at 586 nm and the emission peaks are shown above. Compound 2 clearly has a decreased emission peak when compared to compound 1, indicating a looser globule formation. Compound 3 is a control homopolymer of 36 *N*-(2-methoxyethyl)glycine residues and as expected it shows no fluorescence.

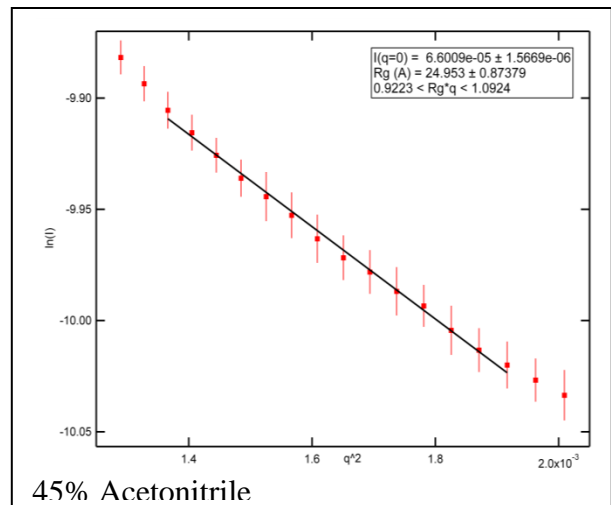
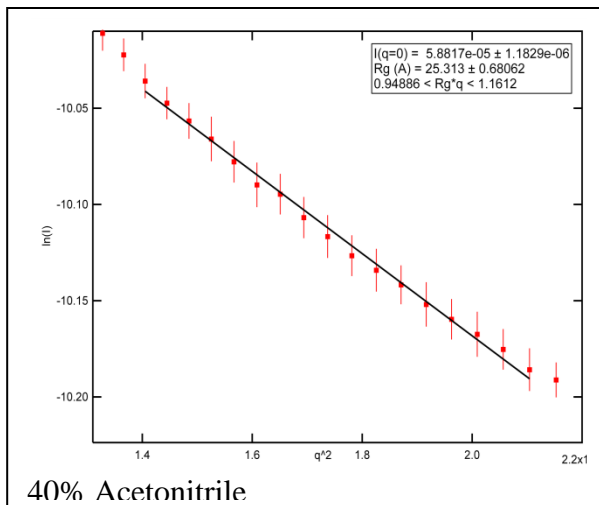
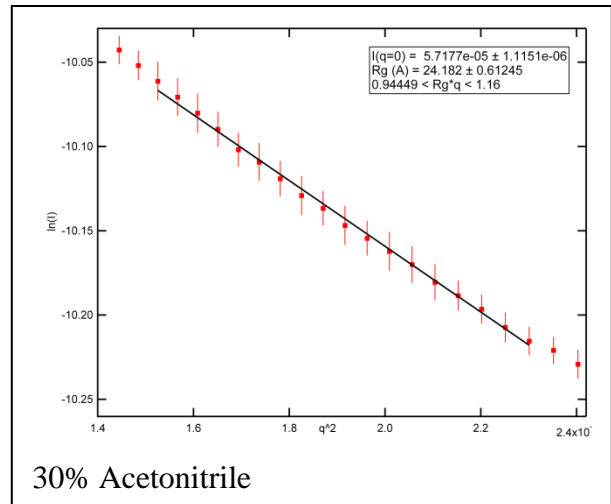
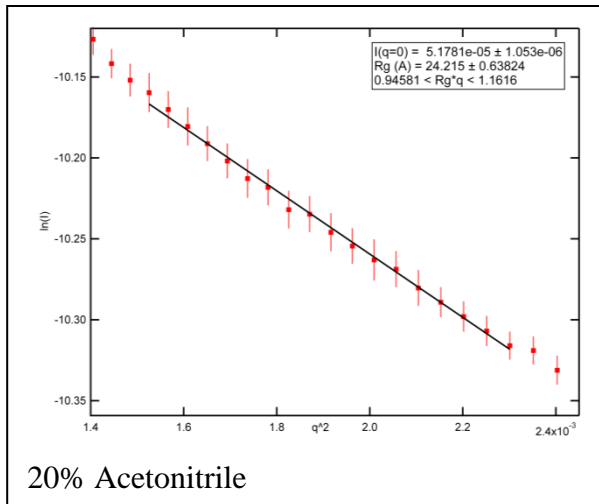
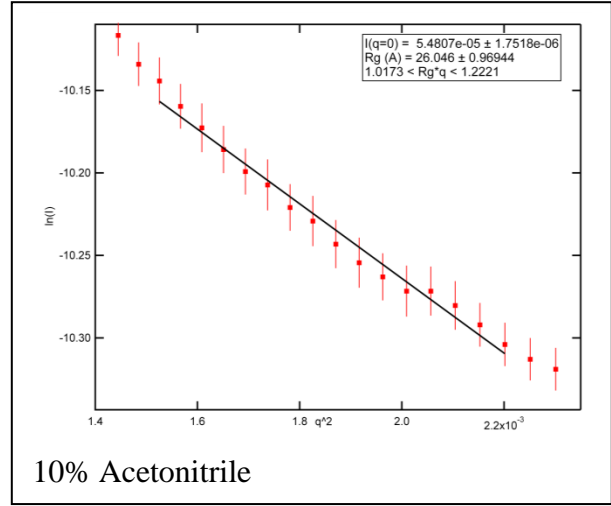
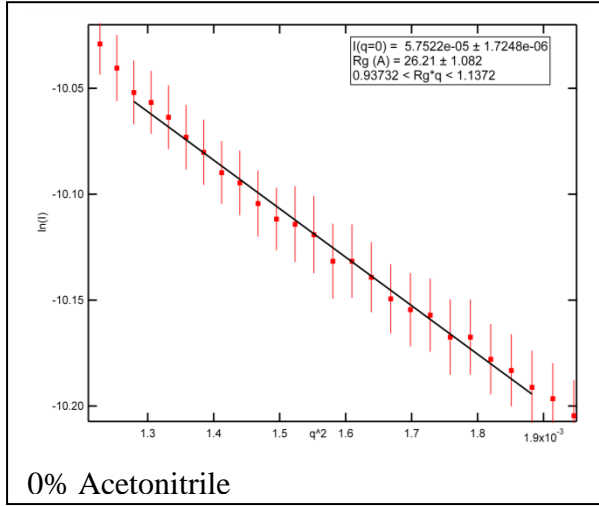
4.4. Conclusions

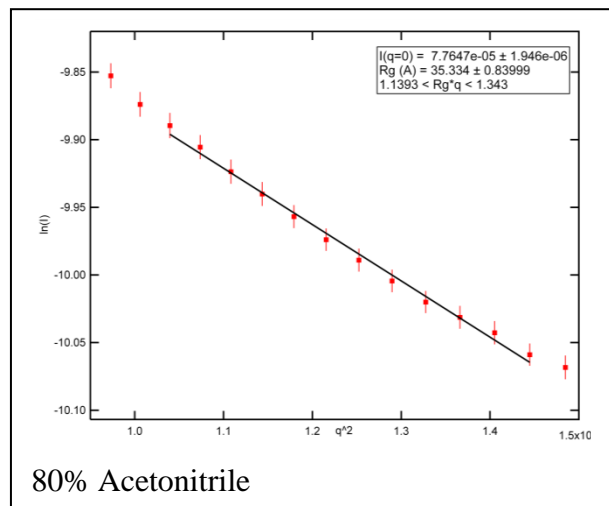
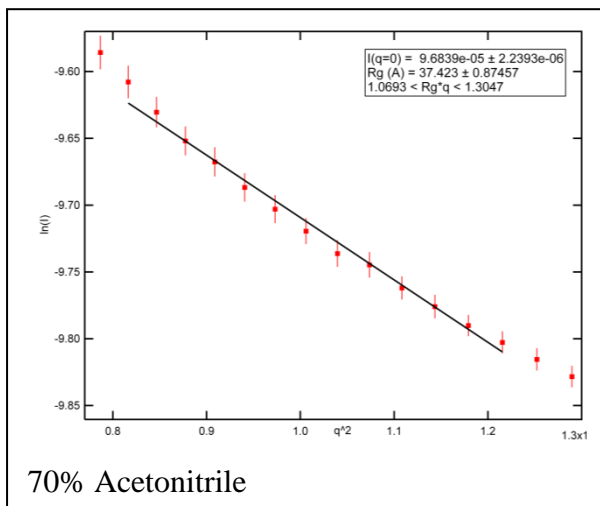
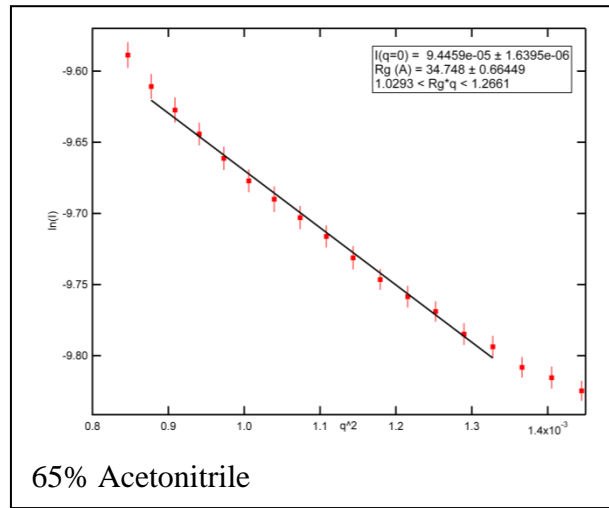
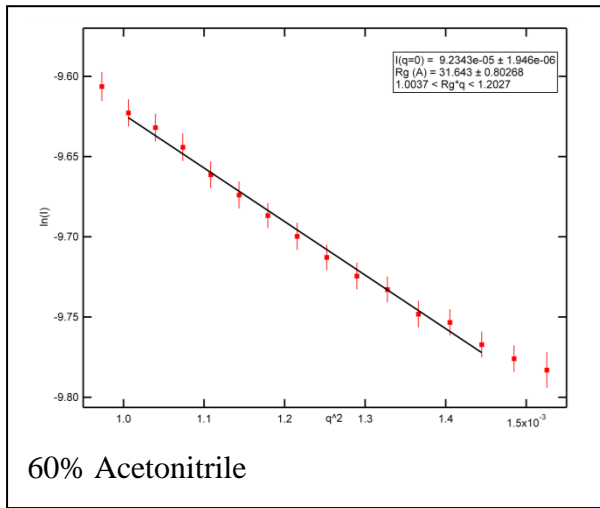
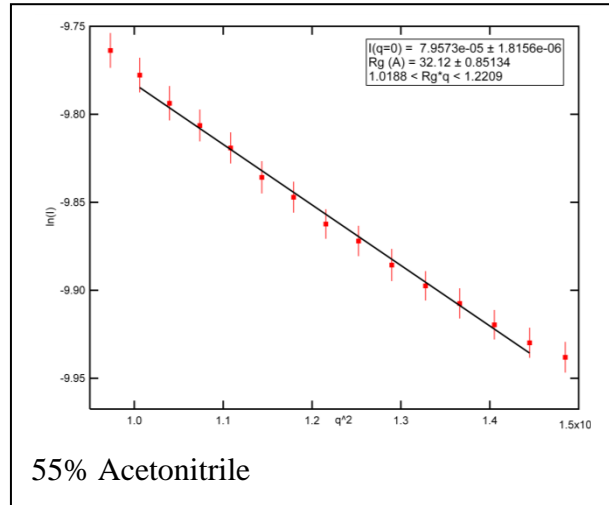
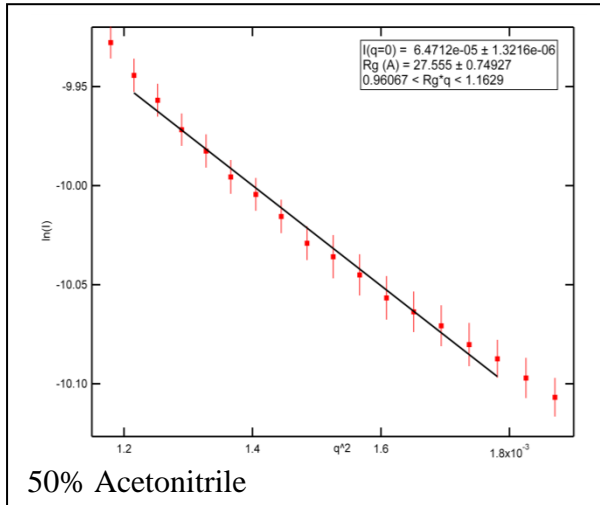
In conclusion, a protein-like polymer sequence was generated using theoretical computational methods based on an HP model. The model polypeptoid system utilized in this work allowed the direct study of the impact of hydrophobic sequence patterning in complete isolation from any other interfering factors such as hydrogen bonding or chirality. Taken together, the SAXS, DLS and fluorescence data demonstrate a fundamental difference between the coil to globule transition behavior of the protein-like polymer sequence which has longer blocks of H and P monomers, and a regular, repeating sequence containing the exact same monomer composition, and has the shortest possible stretches of H and P monomers. Size measurements, through light and x-ray scattering, as well as fluorescence measurements show a distinct difference in the globule formed by the two polypeptoids in aqueous solution. Additionally, the unfolding of the two molecules due to the addition of a hydrophobic solvent, acetonitrile, was shown to be markedly different. A significantly larger ΔG for the transition from coil to globule as well as a higher m value were measured for the protein-like polypeptoid, indicating increased cooperativity and buried hydrophobic residues for the protein-like sequence as compared to a control repeating sequence.

This provides one of the first experimental results supporting previous theoretical work demonstrating the impact of hydrophobic sequence patterning on coil to globule collapse. Furthermore, this work provides some basic rules to enable the design of folded, single-chain biomimetic nanostructures. Using this approach, it should be possible to create water-soluble polymer micelles with defined interior and exterior residues, with controlled diameters and controlled densities, which could enable new class of drug delivery vehicles⁴¹. More generally, the ability to fold non-natural polymers into defined structures could allow a new generation of robust protein-mimetic materials capable of specific molecular recognition and catalysis.

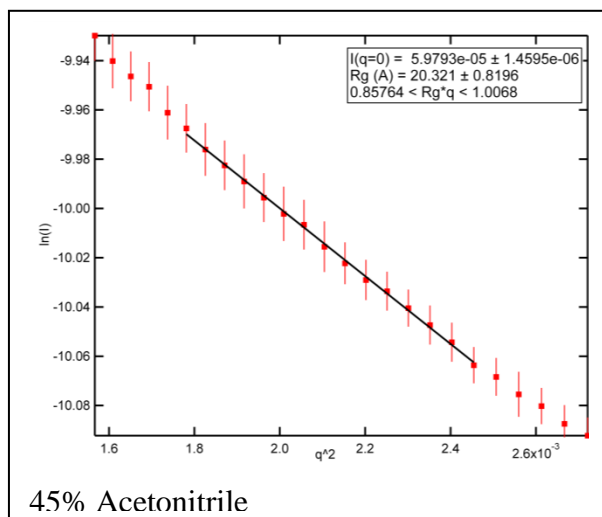
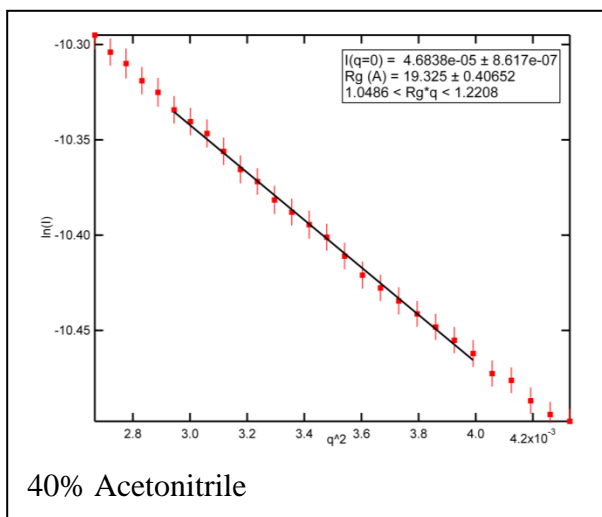
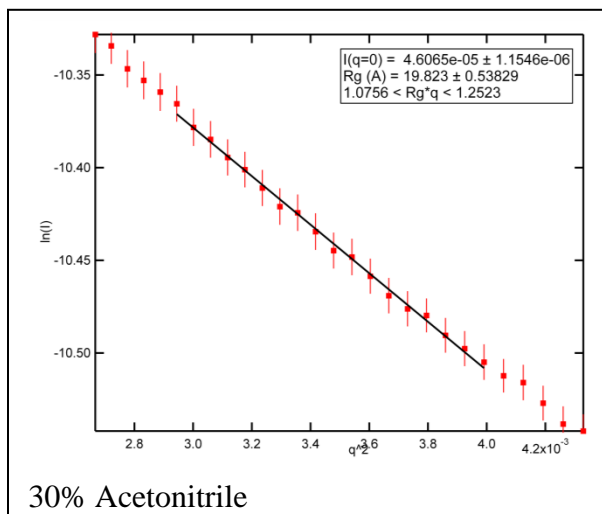
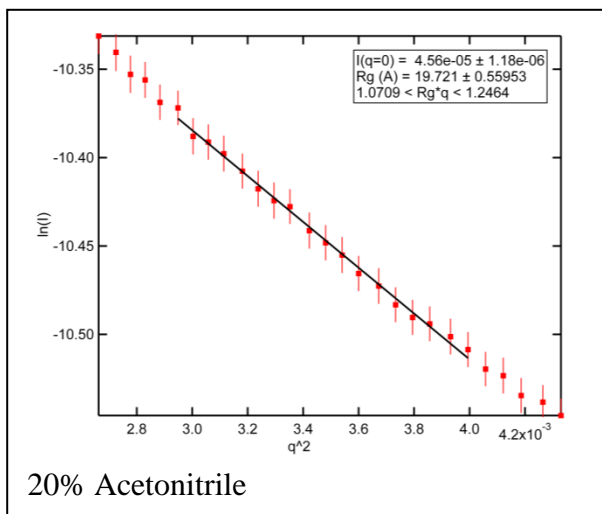
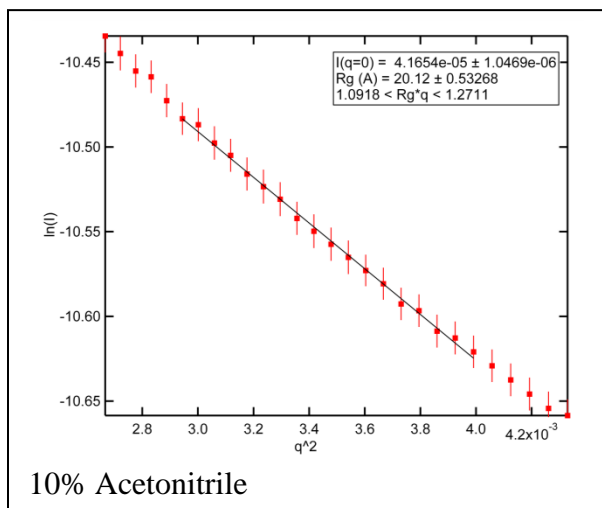
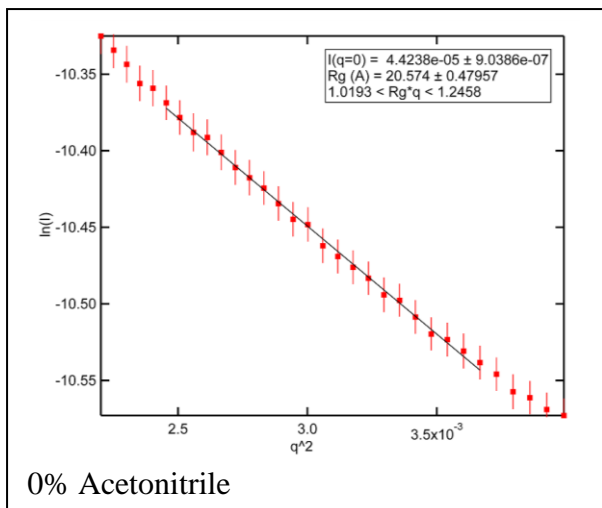
4.4. Appendix

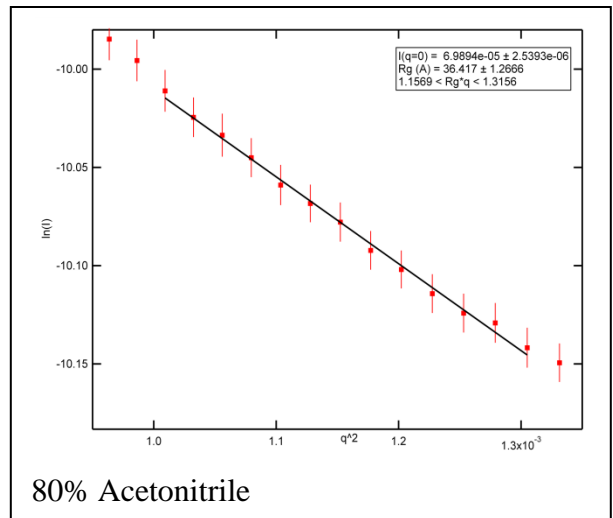
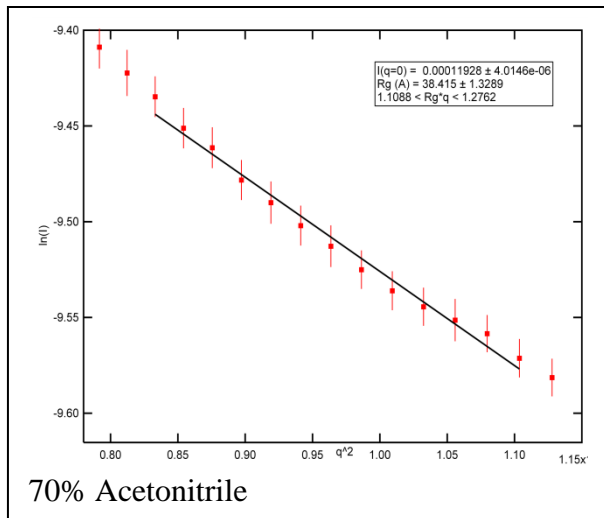
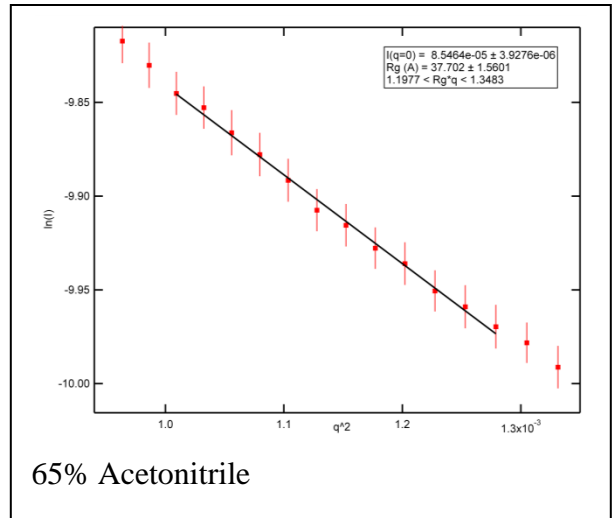
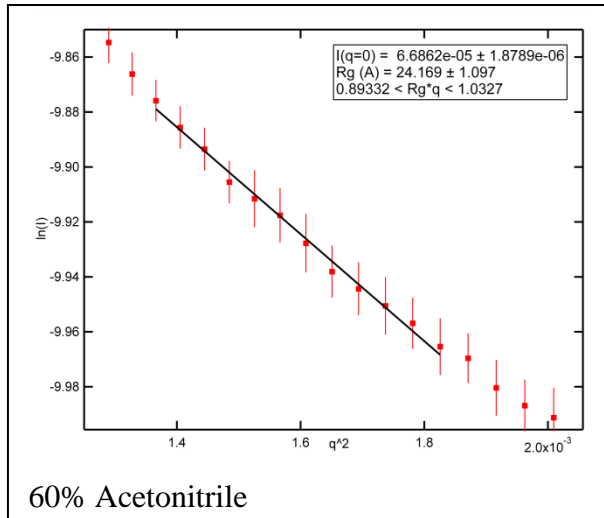
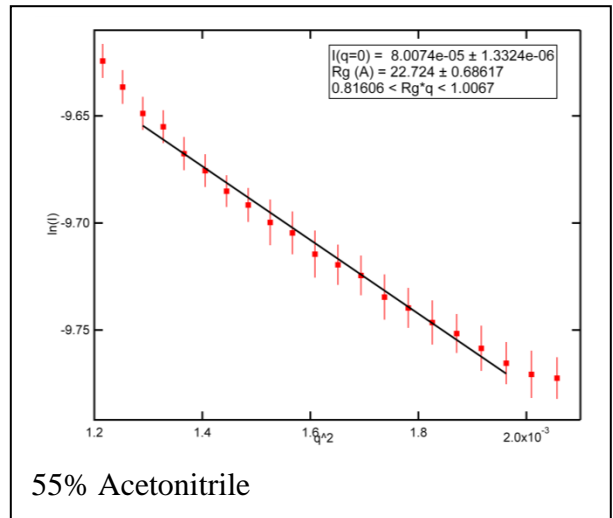
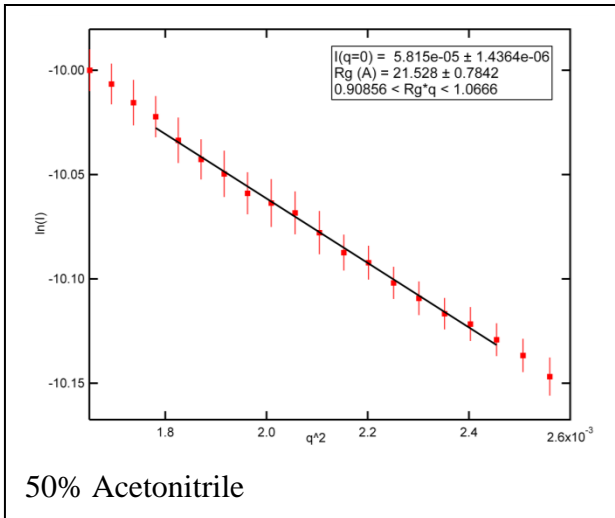
Repeating Sequence Guinier Fits





Protein-like Guinier Plots





4.5. Acknowledgements

The authors would like to thank Byoung-Chul Lee for helpful discussions. We gratefully acknowledge funding from the Office of Naval Research via a Presidential Early Career Award in Science and Engineering. H.K.M also gratefully acknowledges the Department of Defense for a graduate fellowship. Polypeptoid synthesis and associated chemical characterization were performed at the Molecular Foundry, a Lawrence Berkeley National Laboratory user facility supported by the Office of Science, Office of Basic Energy Sciences, U.S. Department of Energy, under Contract DE-AC02-05CH11231. Parts of the X-ray scattering studies were carried out at the Advanced Light Source (ALS) on Beamline 7-3-3. The Advanced Light Source is supported by the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Additional X-ray scattering studies were carried out at the Stanford Synchrotron Radiation Laboratory (SSRL), a national user facility operated by Stanford University on behalf of the U.S. Department of Energy on Beamline 1-4.

4.6. References

1. Dill, K. A.; Bromberg, S.; Yue, K. Z.; Fiebig, K. M.; Yee, D. P.; Thomas, P. D.; Chan, H. S., Principles of protein-folding - A perspective from simple exact models. *Protein Sci.* **1995**, *4* (4), 561-602.
2. Zarrine-Afsar, A.; Wallin, S.; Neculai, A. M.; Neudecker, P.; Howell, P. L.; Davidson, A. R.; Chan, H. S., Theoretical and experimental demonstration of the importance of specific nonnative interactions in protein folding. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (29), 9999-10004.
3. Sun, S. J.; Brem, R.; Chan, H. S.; Dill, K. A., Designing amino acid sequences to fold with good hydrophobic cores. *Protein Eng.* **1995**, *8* (12), 1205-1213.
4. Hecht, M. H.; Das, A.; Go, A.; Bradley, L. H.; Wei, Y. N., De novo proteins from designed combinatorial libraries. *Protein Sci.* **2004**, *13* (7), 1711-1723.
5. Kortemme, T.; Ramirez-Alvarado, M.; Serrano, L., Design of a 20-amino acid, three-stranded beta-sheet protein. *Science* **1998**, *281* (5374), 253-256.
6. Venkatraman, J.; Shankaramma, S. C.; Balaram, P., Design of folded peptides. *Chem. Rev.* **2001**, *101* (10), 3131-3152.
7. Desjarlais, J. R.; Handel, T. M., De-novo design of the hydrophobic cores of proteins. *Protein Sci.* **1995**, *4* (10), 2006-2018.
8. Baker, D., A surprising simplicity to protein folding. *Nature* **2000**, *405* (6782), 39-42.
9. Plaxco, K. W.; Riddle, D. S.; Grantcharova, V.; Baker, D., Simplified proteins: minimalist solutions to the 'protein folding problem'. *Curr. Opin. Struct. Biol.* **1998**, *8* (1), 80-85.
10. Arai, M.; Kuwajima, K., Role of the molten globule state in protein folding. *Adv. Protein Chem.* **2000**, *53*, 209-282.
11. Khokhlov, A. R.; Khalatur, P. G., Protein-like copolymers: Computer simulation. *Physica A* **1998**, *249* (1-4), 253-261.

12. Khokhlov, A. R.; Khalatur, P. G., Conformation-dependent sequence design (engineering) of AB copolymers. *Phys. Rev. Lett.* **1999**, *82* (17), 3456-3459.
13. Ashbaugh, H. S., Tuning the Globular Assembly of Hydrophobic/Hydrophilic Heteropolymer Sequences. *J. Phys. Chem. B* **2009**, *113* (43), 14043-14046.
14. Siu, M.; Zhang, G. Z.; Wu, C., Effect of comonomer distribution on the coil-to-globule transition of a single AB copolymer chain in dilute solution. *Macromolecules* **2002**, *35* (7), 2723-2727.
15. Jennings, D. E.; Kuznetsov, Y. A.; Timoshenko, E. G.; Dawson, K. A., A lattice model Monte Carlo study of coil-to-globule and other conformational transitions of polymer, amphiphile, and solvent. *J. Chem. Phys.* **2000**, *112* (17), 7711-7722.
16. Dasmahapatra, A. K.; Nanavati, H.; Kumaraswamy, G., Pathway to copolymer collapse in dilute solution: Uniform versus random distribution of comonomers. *J. Chem. Phys.* **2007**, *127* (23), 7.
17. Jamadagni, S. N.; Bosoy, C.; Garde, S., Designing Heteropolymers To Fold into Unique Structures via Water-Mediated Interactions. *J. Phys. Chem. B* **2010**, *114* (42), 13282-13288.
18. Khalatur, P. G.; Khokhlov, A. R., Computer-aided conformation-dependent design of copolymer sequences. In *Conformation-Dependent Design of Sequences in Copolymers I*, Khokhlov, A. R., Ed. Springer-Verlag Berlin: Berlin, 2006; Vol. 195, pp 1-100.
19. Semler, J. J.; Jhon, Y. K.; Tonelli, A.; Beevers, M.; Krishnamoorti, R.; Genzer, J., Facile method of controlling monomer sequence distributions in random copolymers. *Adv. Mater.* **2007**, *19* (19), 2877-+.
20. Burova, T. V.; Grinberg, N. V.; Grinberg, V. Y.; Ang, Y. T.; Zhang, G. Z.; Khokhlov, A. R., Order-disorder conformational transitions of N-isopropylacrylamide-sodium styrene sulfonate copolymers in aqueous solutions. *Macromolecules* **2008**, *41* (16), 5981-5984.
21. Lozinsky, V. I.; Simenel, I. A.; Kulakova, V. K.; Kurskaya, E. A.; Babushkina, T. A.; Klimova, T. P.; Burova, T. V.; Dubovik, A. S.; Grinberg, V. Y.; Galaev, I. Y.; Mattiasson, B.; Khokhlov, A. R., Synthesis and studies of N-vinylcaprolactam/N-vinylimidazole copolymers that exhibit the "proteinlike" behavior in aqueous media. *Macromolecules* **2003**, *36* (19), 7308-7323.
22. Jhon, Y. K.; Semler, J. J.; Genzer, J.; Beevers, M.; Gus'kova, O. A.; Khalatur, P. G.; Khokhlov, A. R., Effect of Comonomer Sequence Distribution on the Adsorption of Random Copolymers onto Impenetrable Flat Surfaces. *Macromolecules* **2009**, *42* (7), 2843-2853.
23. Han, J. W.; Jeon, B. H.; Ryu, C. Y.; Semler, J. J.; Jhon, Y. K.; Genzer, J., Discriminating Among Co-monomer Sequence Distributions in Random Copolymers Using Interaction Chromatography. *Macromol. Rapid Commun.* **2009**, *30* (18), 1543-1548.
24. Rosales, A. M., Murnen, H.K., Kline, S.R., Zuckermann, R.N., Segalman, R.A., Determination of the persistence length of helical and non-helical polypeptoids in solution. *Soft Matter* **2012**, *8*, 3673-3680.
25. Lee, B. C.; Chu, T. K.; Dill, K. A.; Zuckermann, R. N., Biomimetic nanostructures: Creating a high-affinity zinc-binding site in a folded nonbiological polymer. *J. Am. Chem. Soc.* **2008**, *130* (27), 8847-8855.

26. Lee, B. C.; Zuckermann, R. N.; Dill, K. A., Folding a nonbiological polymer into a compact multihelical structure. *J. Am. Chem. Soc.* **2005**, *127* (31), 10999-11009.
27. Murnen, H. K.; Rosales, A. M.; Jaworski, J. N.; Segalman, R. A.; Zuckermann, R. N., Hierarchical Self-Assembly of a Biomimetic Diblock Copolypeptoid into Homochiral Superhelices. *J. Am. Chem. Soc.* **2010**, *132* (45), 16112-16119.
28. Sanii, B.; Kudirka, R.; Cho, A.; Venkateswaran, N.; Olivier, G. K.; Olson, A. M.; Tran, H.; Harada, R. M.; Tan, L.; Zuckermann, R. N., Shaken, Not Stirred: Collapsing a Peptoid Monolayer To Produce Free-Floating, Stable Nanosheets. *J. Am. Chem. Soc.* **2011**, *133* (51), 20808-20815.
29. Kudirka, R.; Tran, H.; Sanii, B.; Nam, K. T.; Choi, P. H.; Venkateswaran, N.; Chen, R.; Whitelam, S.; Zuckermann, R. N., Folding of a Single-Chain, Information-Rich Polypeptoid Sequence into a Highly Ordered Nanosheet. *Biopolymers* **2011**, *96* (5), 586-595.
30. Nam, K. T.; Shelby, S. A.; Choi, P. H.; Marciel, A. B.; Chen, R.; Tan, L.; Chu, T. K.; Mesch, R. A.; Lee, B. C.; Connolly, M. D.; Kisielowski, C.; Zuckermann, R. N., Free-floating ultrathin two-dimensional crystals from sequence-specific peptoid polymers. *Nat. Mater.* **2010**, *9* (5), 454-460.
31. Rosales, A. M.; Murnen, H. K.; Zuckermann, R. N.; Segalman, R. A., Control of Crystallization and Melting Behavior in Sequence Specific Polypeptoids. *Macromolecules* **2010**, *43* (13), 5627-5636.
32. Figliozzi, G. M.; Goldsmith, R.; Ng, S. C.; Banville, S. C.; Zuckermann, R. N., Synthesis of N-substituted glycine peptoid libraries. *Combinatorial Chemistry* **1996**, *267*, 437-447.
33. Abrego, J. R. B.; Craievich, A. F.; Mascarenhas, Y. P.; Laure, C. J., SAXS study of crotopotin at low pH. *Biophys. J.* **1993**, *64* (2), 560-564.
34. Santoro, M. M.; Bolen, D. W., Unfolding free-energy changes determined by the linear extrapolation method .1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* **1988**, *27* (21), 8063-8068.
35. Myers, J. K.; Pace, C. N.; Scholtz, J. M., Denaturant M-values and heat-capacity changes - Relation to changes in accessible surface-areas of protein unfolding. *Protein Sci.* **1995**, *4* (10), 2138-2148.
36. Regan, L.; Degrado, W. F., Characterization of a helical protein designed from 1st principles. *Science* **1988**, *241* (4868), 976-978.
37. Tanford, C., *Physical Chemistry of Macromolecules*. John Wiley and Sons: New York, 1961.
38. Burchard, W., 1983; Vol. 48.
39. Sackett, D. L.; Wolff, J., Nile red as a polarity-sensitive fluorescent-probe of hydrophobic protein surfacesI. *Anal. Biochem.* **1987**, *167* (2), 228-234.
40. Hawe, A.; Sutter, M.; Jiskoot, W., Extrinsic fluorescent dyes as tools for protein characterization. *Pharm. Res.* **2008**, *25* (7), 1487-1499.
41. Davis, M. E.; Chen, Z.; Shin, D. M., Nanoparticle therapeutics: an emerging treatment modality for cancer. *Nat. Rev. Drug Discov.* **2008**, *7* (9), 771-782.

Chapter 5. Hierarchical Self-assembly of a Biomimetic Diblock Copolypeptoid into Homochiral Superhelices

Reproduced with permission from Adrienne M. Rosales, Jonathan N. Jaworski, Rachel A. Segalman, Ronald N. Zuckermann, 2010, Journal of American Chemical Society, 132, 45, 16112-16119. © American Chemical Society, 2010.

5.1. Introduction

Hierarchical self assembly is a hallmark of biological materials. Systems ranging from nacre¹ to collagen fibrils² have been heralded for advantageous properties, such as their mechanical strength, that stem from unique layered structures. The precise order of these materials on the micron and millimeter scales arises from atomically-defined interactions at the nanometer and even subnanometer level. Understanding the relationship between these interactions has great implications for the design of new materials with controllable order across many length scales³.

Although examples of hierarchical polypeptide structures abound in nature⁴⁻⁶, the *de novo* design of such systems is still a major challenge.⁷ The molecular complexity of polypeptide interactions makes it difficult to engineer their self-assembly. Hydrophobic and ionic forces are joined by backbone chirality and hydrogen bonding, making it challenging to isolate or understand the effect of any parameter in particular. Thus, most efforts in the *de novo* design of folded and self-assembling peptides have focused on relatively short chain lengths⁸. The utility of engineered peptide structures in the design of structured biomaterials has been proven by the diversity of achievable structures including flat tapes, tubes and spheres⁹, as well as by the insights gained into mechanisms ranging from amyloid fibril formation¹⁰ to antimicrobial peptide cytotoxicity¹¹. *De novo* peptide systems are thus attractive for specific biotechnological applications³, but simpler biomimetic polymer systems may allow the development of straightforward design rules for the engineering of self assembled materials. Therefore, a tunable and synthetically robust system that can mimic the atomic level ordering in biological systems while allowing system engineering is desired for both materials applications and fundamental investigations of biomacromolecular self assembly.

Polymer scientists have developed comparatively simpler model systems that allow controlled engineering of self assembled structures in aqueous solution. Charged amphiphilic block copolymers have emerged as a particularly interesting class of nanoscale building blocks due to their ability to build hierarchical levels of structure by drawing upon the interplay between the ionic and hydrophobic interactions^{12, 13}. For example, the most well studied system in this category, poly(styrene)-*b*-poly(acrylic acid), has been shown to self-assemble into a variety of structures in solution including hierarchical compound micelles, spheres, rods and vesicles. The identity of the self assembled structure depends on the solvent and the relative mole fractions of the chargeable block¹⁴. Additionally, the charges on many ionizable polymers, including polyacrylic acid, are amphoteric leading to pH dependent supramolecular structures¹⁵⁻¹⁷. However, while these structures can have internal ordering qualitatively similar to biological structures such as amyloid¹⁸ or collagen fibers¹⁹, the inherent polydispersity in main chain length and lack of sequence specificity pose a limit for the achievable order on the atomic level³. The lack of ability to introduce functional monomers (chiral, charged, hydrophobic, etc.) at specific locations means that only complete changes to entire block chemistries are possible.

Here we use peptoid polymers to explore the aqueous self-assembly of amphiphilic diblock copolymers. Polypeptoid chemistry is attractive as it combines the sequence specificity of biological systems with the more controllable intra/intermolecular interactions, robustness and ease of synthesis of traditional synthetic polymers. Polypeptoids are a class of biomimetic sequence specific polymers synthesized via a solid-supported submonomer method²⁰. The backbone is identical to that of a polypeptide, but the side chain is attached to the nitrogen rather than the alpha carbon. This difference eliminates hydrogen bonding in the backbone and also eliminates the main chain chirality, allowing the control of desired interactions through the introduction of specific side chains. Additionally, the use of a primary amine as the submonomer opens up a wealth of chemical functionalities for the side chains, making it possible to control and fine-tune the intra- and intermolecular forces simply by changing individual side chains²¹. Simple polypeptoid sequence patterns have been shown to form very well ordered nanoscale materials²². These characteristics make polypeptoids an ideal system for understanding macromolecular self assembly and building robust materials with atomic level ordering.

The lack of backbone chirality in polypeptoids allows the introduction of chirality through side chain insertion only when desired. This type of chiral control has been used in synthetic systems to influence the handedness of the resulting self assembled super structures, particularly in helices. Nolte *et al.*, were able to derive chirality in a self assembled super helix from the handedness of the alpha-helix of a polyisocyanopeptide building block²³. In addition to supramolecular chirality arising from molecular chirality, there are also cases where supramolecular helical chirality results from achiral molecules such as small molecules²⁴⁻²⁶ and liquid crystals²⁷⁻²⁹. In these cases, the chiral self assembled structures are a racemic mixture or are influenced via mechanical forces such as stirring²⁶. The sequence specificity of the polypeptoids and the lack of backbone chirality provide a framework for investigating the effects of molecular chirality on supramolecular chirality.

We present here a hierarchically self-assembled super helix structure with uniform pitch (429 ± 105 nm) and diameter (624 ± 69 nm) arising from a partially charged amphiphilic diblock copolypeptoid. The super helices are remarkably homochiral despite the achiral nature of all components. The adaptable chemistry of the system has been used to make systematic changes to the polymer in the form of closely related analogs, and from those changes an in depth understanding of the internal structure and the role of charge location and density in the self-assembly has been obtained. The origin of the chirality has also been investigated and further work will focus on this issue.

5.2. Experimental Section

Synthesis

Polypeptoids were synthesized on a custom robotic synthesizer or a commercial Aapptec Apex 396 robotic synthesizer on 100 mg of Rink amide polystyrene resin (0.6 mmol/g, Novabiochem, San Diego). All primary amine submonomers, solvents, and reagents described here were purchased from commercial sources and used without further purification. *N*-(2-carboxyethyl)glycine was made from the beta-alanine *O*-*t*Bu ester hydrochloride submonomer. The submonomer was freebased by extraction from dichloromethane and basic water. A similar procedure was used to freebase the beta-alaninamide hydrochloride monomer. In this case the freebase solvent was ethyl acetate/methanol 3:2 (v/v). The polypeptoid synthesis procedure was a modified version of methods previously described³⁰ using the primary amines shown in Table 1

in 1.5 M concentration. Sixty minute displacement times were used for the first 15 residues and 90 minutes for the remaining residues. All other synthesis conditions were identical to those previously reported. Peptoid chains were cleaved from the resin by addition of 4.0 mL of a 95% trifluoroacetic acid (TFA), 5% water solution for 60 minutes, which was then evaporated off under a stream of nitrogen gas. This treatment also served to remove the tert-butyl protecting groups from the carboxyethyl groups. Following cleavage, peptoids were dissolved in 4.0 mL of 1:1 (v/v) acetonitrile/water and lyophilized twice to a fluffy white powder.

Each polypeptoid was characterized by analytical reverse-phase HPLC using a C4 column (Vydac 214TP, 5 μ m, 4.6 x 150 mm) on a Varian ProStar system (Palo Alto, CA). The column was maintained at 60°C while a 30 minute linear gradient of 5-95% solvent B in solvent A was used (solvent A = 0.1% TFA in water, solvent B = 0.1% TFA in acetonitrile). All peptoids were purified by reverse-phase prep HPLC on a Varian ProStar system equipped with a Varian Model 345 UV-Vis Dual Wavelength detector (214 and 260 nm) and a C4 column (Vydac HPLC Protein C4 column, 10-15 μ m, 22 x 250 mm) using a linear gradient of 50-100% solvent B in solvent A over 40 minutes at a flow rate of 10 ml/min (solvent A = 0.1% TFA in water, solvent B = 0.1% TFA in acetonitrile). All of the sidechains used are listed in Table 1 along with their abbreviations. Each polymer is named by using the abbreviations of the appropriate sidechains along with a subscript indicating the number of repeats of a given monomer in a row. All of the polymers synthesized are shown in Table 2 along with their purities and observed molecular weight. The purity was determined using the analytical reverse-phase HPLC detailed above and the molecular weight was determined using an Applied Biosystems MALDI TOF/TOF Analyzer 4800 with a 1:1 (v/v) mixture of peptoid (2 mg/mL in 1:1 acetonitrile:water) and 1, 8, 9,-dianthracenetriol dissolved in tetrahydrofuran at 10 mg/mL.

Self assembly solutions

The amphiphilic molecules were dissolved in water at a concentration between 1-10 mg/mL using sodium hydroxide (NaOH) to adjust the pH to the desired value. Once the NaOH was added the solutions were allowed to sit at room temperature undisturbed. In order to image the self-assembled structures using AFM or SEM, a drop of solution was placed on an oxygen plasma cleaned silicon wafer. After waiting 10 minutes, the excess liquid was wicked away and the substrate washed once with water. To image using TEM, the same technique was used but a carbon-coated copper 200 mesh grid was used instead of the silicon wafer.

X-ray diffraction

The x-ray diffraction experiments were performed at beamline 8.3.1 and beamline 7.3.3 at the Advanced Light Source at Lawrence Berkeley National Laboratory. Samples were prepared by evaporating the solvent from the solutions using a Genevac. Data presented in the supplemental information shows that the solid state patterns had identical peak locations to those found in solution scattering patterns (Figure 5.A.3). However, signal to noise was greatly increased and acquisition time decreased by using solid samples. X-rays of 11.11 keV were focused onto the sample and a two dimensional CCD array was used to collect the scattered x-rays after transmission through the sample. The signal was then radially integrated to obtain a 1D plot of intensity versus scattering angle.

Table 5.1. Chemical structure of the peptoid monomers used and their abbreviations.

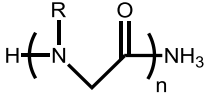
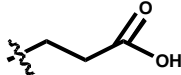
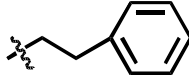
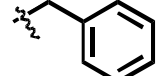
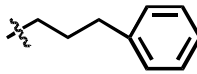
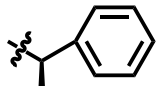
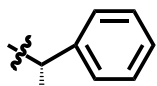
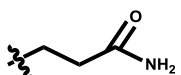

Side chain (R=)	Abbreviation
	
	Nce = <i>N</i> -(2-carboxyethyl)glycine
	Npe = <i>N</i> -(2-phenylethyl)glycine
	Nbn = <i>N</i> -(benzyl)glycine
	Npp = <i>N</i> -(3-phenyl-1-propyl)glycine
	Nrpe = <i>N</i> -((R)-(+)-1-phenylethyl)glycine
	Nspe = <i>N</i> -((S)-(-)-1-phenylethyl)glycine
	Nbm = <i>N</i> -(2-carboxamidoethyl)glycine
	Nme = <i>N</i> -(2-methoxyethyl)glycine

Table 5.2. All of the polymers synthesized and used in this article.

The monomers and their abbreviations are shown in Table 5.1. The subscript in the name indicates the repeat units of that monomer. The observed molecular weights are from MALDI-TOF and the purity was obtained on an analytical HPLC after prep HPLC purification. ND=not determined.

Name	Molecular Weight	Observed Molecular Weight	Purity
<i>pNpe</i> ₁₅ <i>Nce</i> ₁₅	4371.5	4372.1	99%
<i>pNpe</i> ₂₀ <i>Nce</i> ₂₀	5823.3	5825.2	99%
<i>pNrpe</i> ₁₅ <i>Nce</i> ₁₅	4371.5	4365.8	98%
<i>pNspe</i> ₁₅ <i>Nce</i> ₁₅	4371.5	4377.3	99%
<i>pNce</i> ₁₅ <i>Npe</i> ₁₅	4371.5	4371.0	98%
<i>pNpe</i> ₁₄ <i>Nrpe</i> ₁ <i>Nce</i> ₁₅	4371.5	4371.0	97%
<i>pNpe</i> ₁₄ <i>Nspe</i> ₁ <i>Nce</i> ₁₅	4371.5	4372.0	99%
<i>pNpp</i> ₁₅ <i>Nce</i> ₁₅	4582.1	4581.8	96%
<i>pNbn</i> ₁₅ <i>Nce</i> ₁₅	4161.3	4164.0	ND

5.3. Results and Discussion

A model diblock polypeptoid system analogous to poly(styrene-*b*-acrylic acid), one of the most well studied chargeable amphiphilic block copolymers was chosen.^{14, 31} The diblock copolypeptoid, *N*-(2-phenethyl)glycine-*b*-*N*-(2-carboxyethyl)glycine, *pNpe*₁₅*Nce*₁₅ (Table 5.2), has one hydrophobic block containing 2-phenethyl sidechains and one hydrophilic block containing 2-carboxyethyl sidechains. The carboxyl groups can be deprotonated to become negatively charged to differing degrees depending on pH. In addition to similarities to poly(styrene-*b*-poly acrylic acid), the phenethyl and carboxyethyl sidechains are identical to those known to form β -amyloid fibrils when found on a polypeptide chain.³² In the case of polypeptides, it has been shown that a sequence as short as just a dimer of phenylalanine can form nanotubes demonstrating the importance of aromatic interactions to the self-assembly process.^{33, 34}

After automated synthesis and HPLC purification, *pNpe*₁₅*Nce*₁₅ was dissolved in water at a concentration of 0.1 mM and the pH adjusted to 6.8 using 0.5 equivalents of NaOH per carboxyl side chain. Sheet like structures formed within 24 hours and were imaged using scanning electron microscopy (SEM, Figure 5.1a) and atomic force microscopy (AFM, Figure 5.1b). The sheets range from several hundred nanometers up to many microns in length and width, and their edges appear quite straight. AFM analysis of 10 sheets showed the sheet thickness to be very uniform at 7.8 ± 0.53 nm. Given that the fully extended length of a single peptoid chain is approximately 11 nm, it is thought that the sheets consist of interdigitated bilayers formed due to the amphiphilic nature of the molecules. In this scenario, as illustrated in Figure 5.2, the hydrophobic portion of the molecule is embedded in the interior of the sheet in order to minimize contact with water while the charged hydrophilic block faces outward, exposed to the aqueous solution. Superhelical structures, seen in Figure 5.3, appear after 4-7 days in solution.

Over this transitional period, no intermediate structures between the sheets and helices were observed. However, the coexistence of sheets and helices in the same sample has been observed, indicating that any intermediate structure must be relatively short lived. The wide time range of self-assembly indicates the pathway is kinetically controlled and it is possible there are several self-assembly pathways. The superhelical structures are abundant and stable after formation, lasting many months in solution. Filtration and HPLC analysis showed that over 83% of the peptoid mass is present in the self-assembled helical structures (Appendix, Figure 5.A.1). Analysis of one hundred SEM images of the helices shows self-assemblies with surprisingly uniform helix diameters of 624 ± 69 nm and lengths ranging from 2-40 μ m. These structures are significantly larger than the traditional types of block copolymer micellular structures even though the polymer is of a relatively low molecular weight (4371 g/mol). In addition, the helices are regular with a pitch of 429 ± 105 nm. Perhaps most remarkably, while the base polymer is achiral, the giant super helices are homochiral, with all helices having left-handed symmetry. An image gallery of 15 helices is presented in the appendix (Figure 5.A.5).

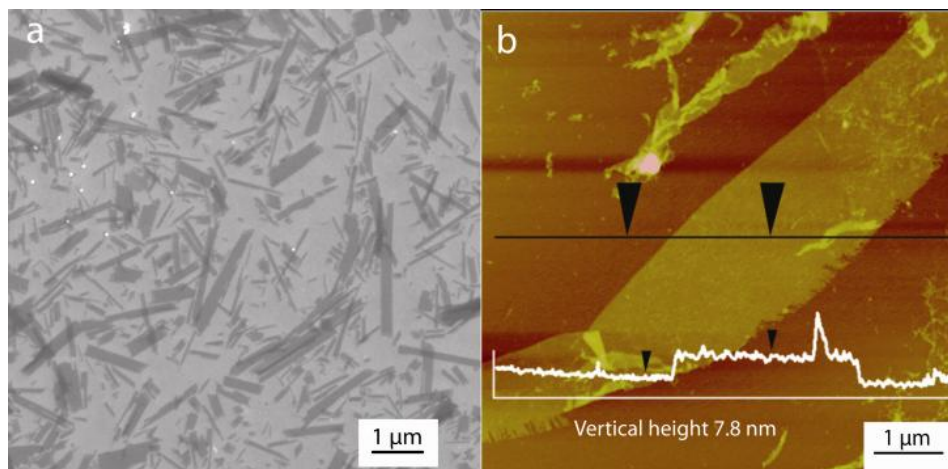


Figure 5.1. The sheet structures formed from the diblock copolypeptoid $pNpe_{15}Nce_{15}$
Assembly occurs after 24 hours in an aqueous solution at pH=6.8. Scanning electron microscopy (a) and atomic force microscopy (b) were used to observe these structures. AFM analysis determined the thickness of the sheets to be 7.8 ± 0.53 nm.

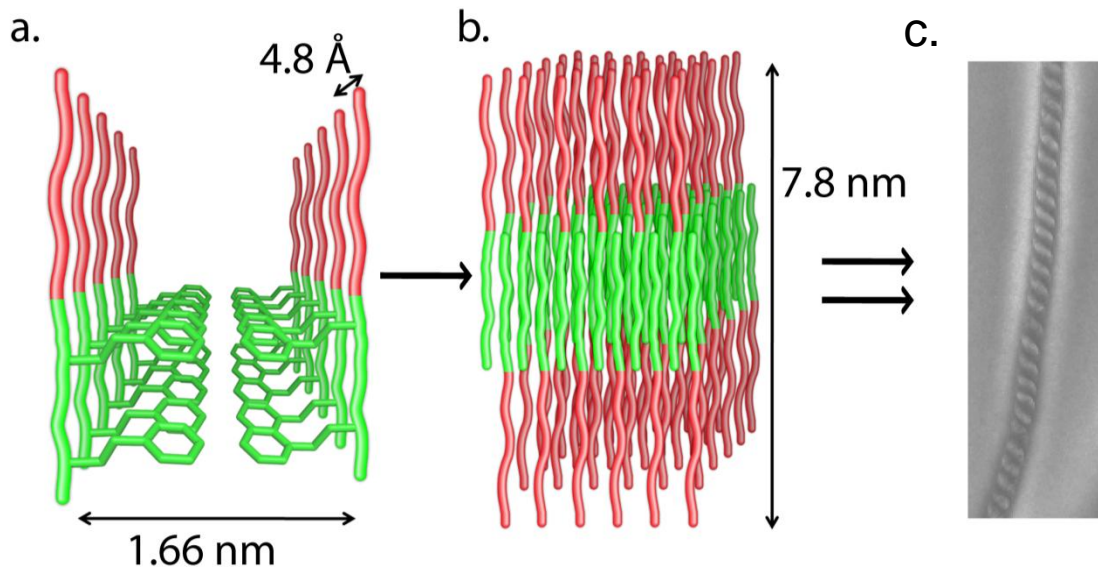


Figure 5.2. A model of the proposed self assembly process.

The green represents the hydrophobic portion of the chain while the red represents the hydrophilic block. The chains initially crystallize with the aromatic groups facing each other (a). This spacing (1.66nm) along with the distance between two chains laterally (4.8 Å) are verified in x-ray scattering. The chains further arrange into 2-dimensional sheets (b) with a height of 7.8 nm as verified by AFM and x-ray scattering. The sheets are layered within the helices as evidenced by lamellar x-ray scattering. The exact mechanism for the assembly of super helices from the sheets is difficult to ascertain due to the lack of observed intermediate structures.

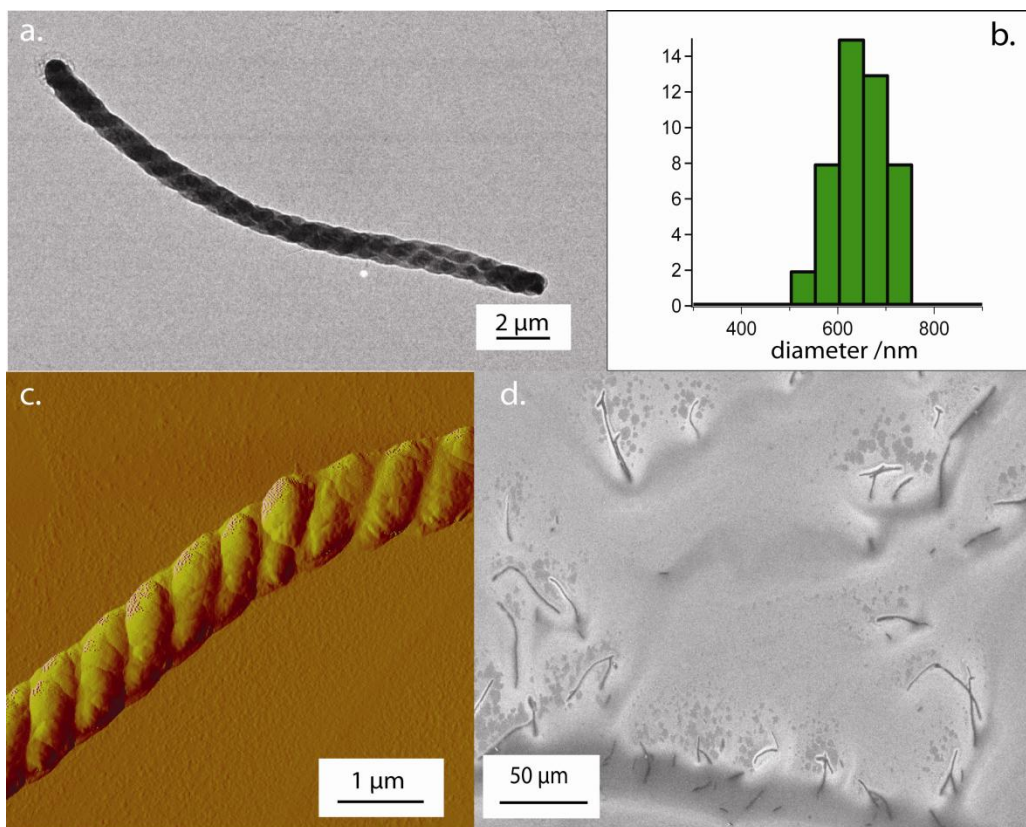


Figure 5.3. Helix formation of $pNpe_{15}Nce_{15}$

Helix formation from $pNpe_{15}Nce_{15}$ occurs after 3-7 days in aqueous solution at a pH of 6.8. The helices are 624 ± 69 nm in diameter (the histogram is shown in b) and range from 2-40 μm in length. They can be seen in TEM (a), AFM (c) or SEM (d). A zoomed out image (d) shows the abundance of the structures within one sample.

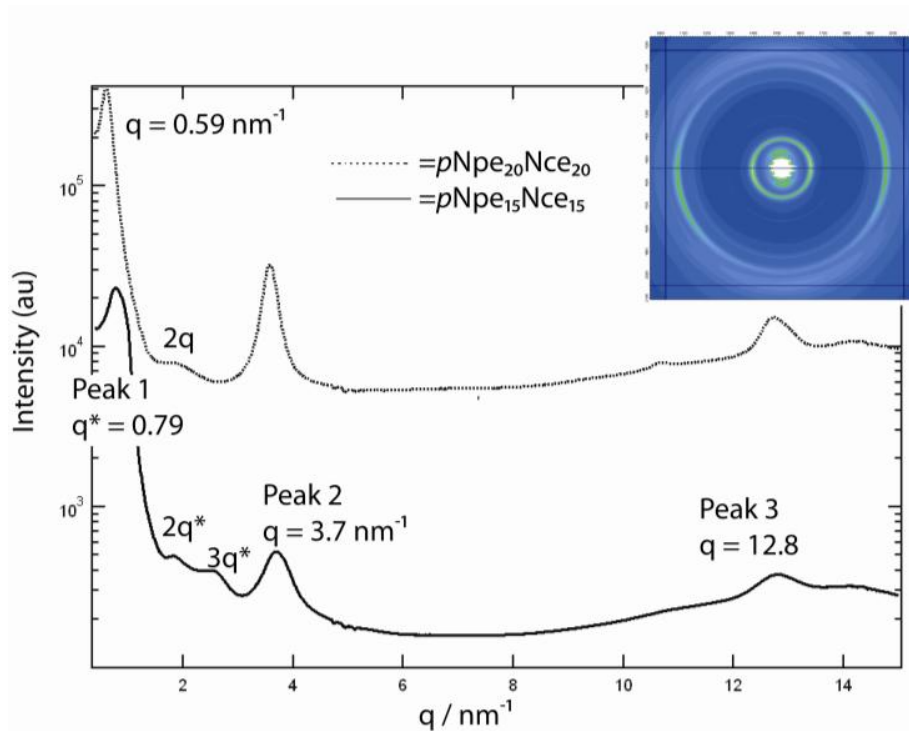


Figure 5.4. SAXS pattern of giant super helices.

Synchrotron x-ray scattering was performed on an evaporated helix sample to investigate the internal ordering. The dotted line here represents $pNpe_{15}Nce_{15}$ while the solid line is $pNpe_{20}Nce_{20}$. The peaks marked q^* , $2q^*$, and $3q^*$ indicate a lamellar stacking with a d spacing of 7.8 nm, very similar to the thickness of the sheets. The peaks marked 2 and 3 are crystalline peaks at d spacings of 1.66 nm and 4.8 Å respectively and are hypothesized to be intrachain packing as modeled in Figure 5.1

Internal Structure

The regularity of the super helix structure as seen in microscopy is indicative of internal ordering on the nanometer scale. Synchrotron x-ray scattering was used to probe this internal ordering. As shown through X-ray diffraction, repeat integer peaks starting at $q^* = 0.79 \text{ nm}^{-1}$ (labeled peak 1 on the solid curve in Figure 5.4) indicate an internal lamellar spacing within the helices with a d-spacing of 7.8 nm. This lamellar spacing closely matches the measured height of the sheets by AFM (Figure 5.1b). From this similarity it is hypothesized that these sheets are stacked within the helices creating the lamellar peaks in x-ray scattering. However, due to the lack of observation of intermediate structures, it is difficult to understand how the sheets transition from single sheets into helical stacks. In addition, there are also peaks at $q = 3.7 \text{ nm}^{-1}$ and 12.8 nm^{-1} which correspond to d-spacings of 1.66 nm and 4.8 Å (labeled peaks 2 and 3 respectively in Figure 5.4). These peaks are attributed to crystalline packing between chains (Figure 5.2a). The 1.66 nm dimension corresponds to the distance between two chains packed inside the supramolecular helix (sidechain crystallinity) with the 2-phenethyl groups are facing each in other in what is most likely an edge to face orientation as seen in Figure 5.2a. The 4.8 Å dimension corresponds to the neighboring intrachain distance as shown in Figure 1.2b. Partial alignment achieved through centrifugal evaporation results in an anisotropic 2D scattering pattern (Figure 5.4, inset). Importantly, the crystalline peaks are perpendicular to the lamellar peaks (the arc at the lowest q) as would be expected from the model in Figure 5.2. Due to the presence of multiple grains and the twisting of the helix, both the crystalline peaks are seen in the meridional direction even though in any individual sheet they are perpendicular to each other. This scattering pattern is similar to that observed in amyloid fibrils³⁵⁻³⁷ and helices formed from amyloid beta peptide fragments³⁸ where the meridional peaks are cited as evidence of a cross- β structure. It should be noted that the above data is exclusively of samples made by evaporating helices in solution to form a powder for diffraction studies. However, X-ray scattering from helices suspended in solution produced identical results and is presented in the supplemental information (Figure 5.A.4).

To further confirm the assignment of the lamellar and intramolecular crystalline peaks, specific chemical modifications were made to both the main chain and the sidechain lengths of the polymer. The resulting changes in the x-ray scattering peaks were used to verify the origin of the peaks. Firstly, the overall length of the polymer was increased from 15 monomers of each block to 20 monomers of each block forming $p\text{Npe}_{20}\text{Nce}_{20}$. This had the effect of decreasing the q -value for the primary peak, q^* , from 0.79 nm^{-1} to 0.59 nm^{-1} , demonstrating an increase in the lamellar spacing by 2.64 nm (Figure 5.4, dotted line).

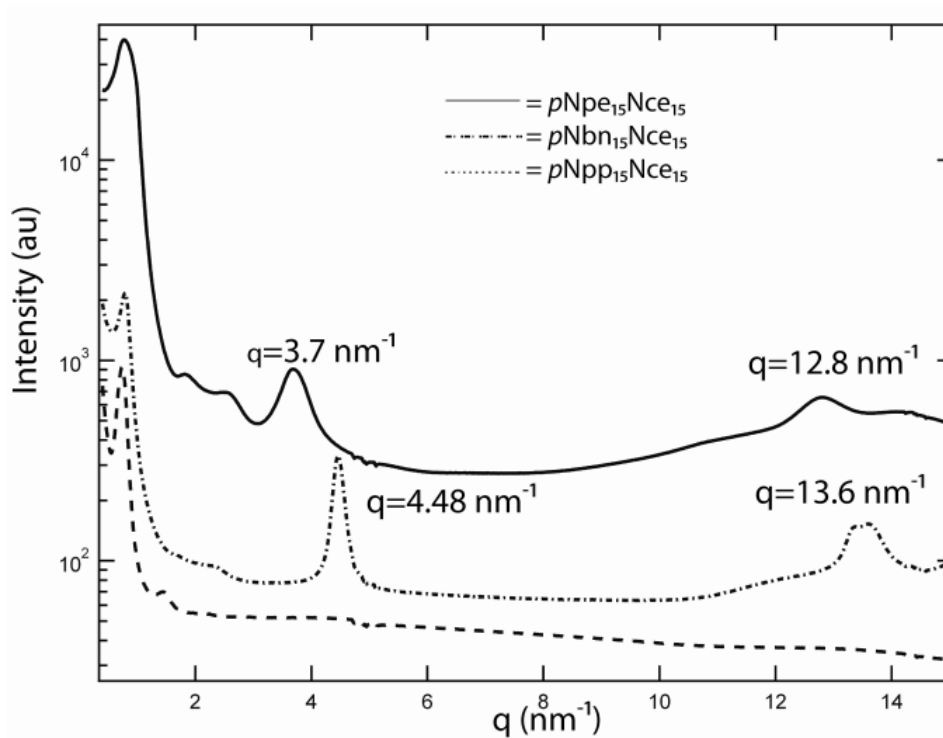


Figure 5.5. X-ray scattering on helical samples of $pNpe_{15}Nce_{15}$, $pNbn_{15}Nce_{15}$ and $pNpp_{15}Nce_{15}$.

Helices composed of $pNpe_{15}Nce_{15}$ (solid), $pNbn_{15}Nce_{15}$ (perforated) and $pNpp_{15}Nce_{15}$ (dashed) show several length scales of ordering in SAXS. The lamellar peaks at $d=7.8$ nm remain in each sample. However, in the $pNbn_{15}Nce_{15}$ sample, the crystalline peaks have shifted to higher q indicating a smaller spacing. The shift in the peak originally at 3.7 nm⁻¹ has now shifted to 4.48 nm⁻¹ which corresponds to two C-C bonds. The peak originally at 12.8 nm⁻¹ has shifted to 13.6 nm⁻¹. It is not clear where the size of this shift originates.

This was corroborated by AFM analysis showing the thickness of the sheets formed by the $pNpe_{20}Nce_{20}$ to be 9.8 nm. The scattering and AFM confirmed that the lamellar q^* , $2q^*$, and $3q^*$ peaks do stem from polymer chains extended in the lengthwise direction and also that these lamellar peaks are linked to the sheet thickness. As predicted by the model in Figure 1.2 the location of the crystalline peaks between sidechains (peaks 2 and 3 in Figure 5.4) was not altered by this chemical modification since they stem only from the packing of the side chains, which have not been changed in this case.

Additional chemical modifications were made to investigate the higher q peaks attributed to intrachain packing. The length of the 2-phenylethyl side chain was shortened by 1 methylene unit to create *N*-(benzyl)glycine-*b*-*N*-(2-carboxyethyl)glycine ($pNbn_{15}Nce_{15}$, Table 5.2). This molecule also forms a super helix when self assembled in aqueous solution. The peak which originally corresponded to a d-spacing of 1.66 nm shifts to reflect a d-spacing of 1.37 nm, resulting in a difference of 2.9 Å. This is reasonable given a C-C bond length of 1.54 Å and a crystalline arrangement with the phenyl groups facing each other such that a one carbon change in the side chain linkage actually results in a distance decrease of two carbon-carbon bonds (Figure 5.2). The peak originally corresponding to 4.8 Å spacing shifts to a d-spacing of 4.5 Å. This peak, as shown in Figure 1.2a does not depend directly on side chain length so it is likely that with a smaller side chain, the backbones can simply pack slightly closer together. The length of the phenyl side chain was also increased by 1 methylene unit to form *N*-(3-phenylpropyl)glycine-*b*-*N*-(2-carboxyethyl)glycine ($pNpp_{15}Nce_{15}$, Table 5.2), which also formed superhelices. However, in this case the side chain crystalline peaks disappear indicating that crystallization is not present within the helices. The longer side chains are more flexible and therefore more difficult to crystallize. This is corroborated by differential scanning calorimetry data (supplemental information, Figure 5.A.5) indicating that the *N*-(3-phenylpropyl)glycine homopolymer does not crystallize in the solid state whereas the 2-phenethyl homopolymer does²¹. Crystallization of the hydrophobic block is therefore not an essential factor in the overall formation of the super helices.

Simply using chemical modifications can only give indirect evidence of the atomic structure within the helix. To gain further insight into the details of the atomic order in the hydrophobic block, a model compound was synthesized and crystallized. Because symmetric cycloalkanes and *N*-methylated cyclic dipeptides have been shown to readily crystallize^{39,40}, we prepared a cyclic dipeptoid, 1,4-bis-(2-phenethyl)-piperazine-2,5-dione, that displays the same *N*-2-phenethyl sidechain groups present in the diblock. It is hypothesized that cyclized dipeptoids could serve as a model system to gain insight into the packing of these sidechains within a larger structure. A single crystal of a cyclic *N*-2-phenethyl dipeptoid, as shown in Figure 5.6, was synthesized and its atomic structure solved by x-ray crystallography as a comparison to $pNpe_{15}Nce_{15}$. The packing arrangement of the cyclic *N*-2-phenethyl dipeptoid is useful for understanding the hydrophobic domain in the $pNpe_{15}Nce_{15}$ super helices. The phenethyl groups are aligned with one another in a staggered edge-to-face conformation to form a plane (Figure 5.6). Lamellar stacks of these planes result in aromatic faces pointing directly toward each other. The spacings observed from the x-ray scattering of the $pNpe_{15}Nce_{15}$ helices (1.66 nm and 4.8 Å, respectively, Figure 5.4) match the dimensions shown in the diketopiperazine crystal structure corresponding to the sidechain and main chain packing distances (1.64 nm and 4.7 Å, respectively, Figure 5.7a). Similarly, the spacings observed from the x-ray scattering of the

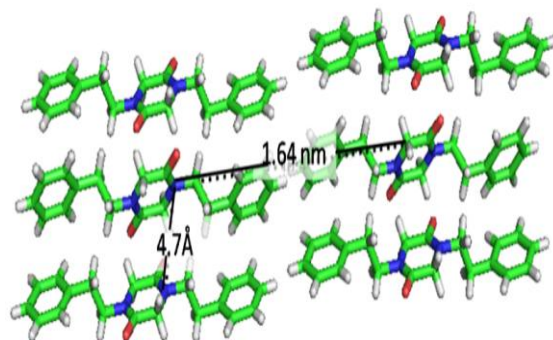


Figure 5.6. Crystal structure of a model cyclic dipeptoid 1,4-bis-(2-phenethyl)-piperazine-2,5-dione
The crystal structure shows the packing geometry of the 2-phenylethyl groups. Green represents carbon atoms, blue represents nitrogen, red represents oxygen, and white represents hydrogen. The dimensions shown match those seen in x-ray scattering of a super helix.

*p*Nbn₁₅Nce₁₅ helices (1.37 nm and 4.5 Å, respectively, Figure 5.5) match the side chain packing and the backbone (central ring) spacing dimensions shown in the previously reported⁴¹ crystal structure of a 1,4-dibenzyl-piperazine-2,5-dione (1.34 nm and 4.5 Å, respectively). This is consistent with the shortening of the sidechain by one methylene unit. The agreement of these structures with the x-ray scattering data for both the *N*-2-phenethyl and *N*-benzyl structures strongly supports the model proposed for the chain conformation within the lamellar stacks of the super helix (Figure 5.3).

Ionic interactions

The interplay of ionic interactions and hydrogen bonding of the carboxyethyl groups is clearly central to the self-assembly of the superhelices involving many discrete molecules. These interactions can be exactly varied via sidechain chemistry, solution interactions and sequence control in the synthesis of the polypeptoids. Charge density and distribution were varied using the pH of the solution as well as side chain substitutions. Helix formation occurred in the buffered region of the molecule where between 1/2 and 2/3 of the carboxylates are charged (titration curve is shown in the supplemental information, Figure 5.A.3). It is important to note that if the self-assembly is carried out at a pH greater than 9.5, helices do not form. In fact, the sheet structures persist due to the high level of deprotonation of the carboxylic acids and the resulting electrostatic repulsion. At a pH less than 5.5, the molecule is not soluble and no organized self assembly occurs. Therefore, some intermediate level of charge was necessary for super helix self assembly. Further study of the dependence of charge location on the formation of super helices was performed by substituting a nonchargeable monomer of similar hydrophilicity, *N*-(methoxyethyl)glycine at particular locations along the hydrophilic block, as shown in Figure 5.7. Given a fixed number of charges, the location of the charges does not affect helix formation. To investigate whether hydrogen bonding alone is sufficient to cause super helix formation, a hydrogen bonding, but non ionic, side chain, *N*-(3-aminopropionamide)glycine was used for the hydrophilic block (*p*Npe₁₅Nbm₁₅, Table 5.2). This sequence does not self-assemble into organized structures demonstrating that ionic interactions are indeed necessary for helix self-assembly to occur.

Chirality

Chirality is a ubiquitous presence in biological molecules resulting from the inherent chirality of the molecular building blocks (amino acids and nucleotides). In this study, the component polymers that make up the giant super helices are completely achiral, making the resulting supramolecular homochirality an unusual and fascinating result. There are no chiral materials present in the self-assembly solutions and yet, without exception all of the helices observed (hundreds) were left handed (see the Appendix for an image gallery, Figure 5.A.5). In previous examples of giant helices from block copolymers, the chirality has come from the molecular chirality of the polymer which is not present in this case. Additionally any chiral superstructures resulting from achiral building blocks were a racemic mixture or influenced in a mechanical manner such as stirring. While the authors of this paper know of no other examples of homochiral superstructures resulting from achiral building blocks without external stimulation, supramolecular chirality can be affected by very subtle influences as evidenced by the significant body of research on the “soldiers and sergeants theory”. It has been shown that the introduction of a small number of chiral groups can cause an otherwise achiral entity to behave in a homochiral manor. One example of this is benzene tricarboxamide molecules which self

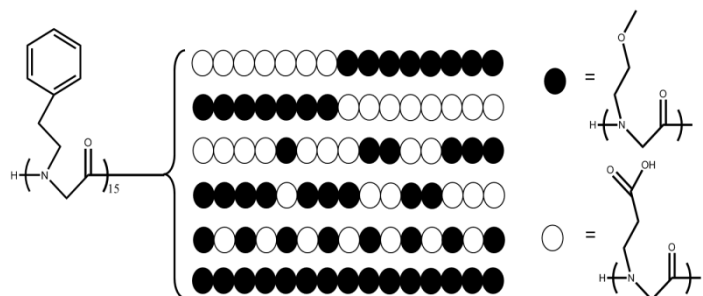


Figure 5.7. Sequences with varying locations of charges.

A set of closely related sequences was designed to pin the chargeable groups at specific locations. The hydrophobic portion of the molecule was held constant while the hydrophilic block was altered. The solid circles represent 2-methoxyethyl sidechains which have a similar hydrophilicity to the carboxyethyl sidechains (open circles) but cannot be charged.

assemble into chiral columnar stacks⁴². A very small excess of a chiral sidechain on one of the benzene groups can tip the self assembly such that a homochiral columnar stack results⁴³⁻⁴⁶.

Keeping the soldiers and sergeants principle in mind several analogs of the parent molecule were synthesized. Chiral groups were introduced at particular locations in an attempt to influence the supramolecular homochirality. It was thought that perhaps a chiral contaminant had been unknowingly introduced to the system and by intentionally inserting a "sergeant" the chirality could be controlled and thus understood. Initially, a single chiral side chain was inserted at the interface of the two blocks. N-((R)-(+)-1-Phenylethyl)glycine (Nrpe) and N-((S)-(-)-1-Phenylethyl)glycine (Nspe) were chosen for their similarity to Npe. The molecules created, (*p*Npe₁₄Nrpe₁Nce₁₅ and *p*Npe₁₄Nspe₁Nce₁₅, Table 2 both formed left handed super helices. Hypothesizing that perhaps one chiral group was not powerful enough to affect the overall structure, the entire Npe block was replaced with either Nspe (*p*Nspe₁₅Nce₁₅) or Nrpe (*p*Nrpe₁₅Nce₁₅) (see Table 5.1). Again, both of the resulting molecules formed identical left handed helix structures. In addition the molecule was synthesized in the reverse order, such that the N-terminus and C-terminus were reversed (*p*Nce₁₅Npe₁₅, Table 5.2). The reasoning here was that perhaps the end groups were providing a source of asymmetry. However, left handed helix structures again assembled. Finally, rather than using an achiral counterion (NaOH) to adjust pH, α -methyl benzyl amine was used, again resulting in left handed helices. The homochirality has proved remarkably robust as none of these approaches changed the overall super helix chirality. It has been hypothesized that surface effects could account for the homochirality. However, self-assembly has occurred in different container types indicating that surface probably does not play a large role in the self-assembly. As expected, circular dichroism (Appendix, Figure 5.A.7) has shown no optical rotation of light demonstrating that the chirality is not on molecular length scale. The shape of the molecule is another potential source of asymmetry and neutron scattering is being pursued as a future experiment to test this possibility. The last possibility is the presence of unequal surface stresses on the lamellae as they form. This has previously been shown to cause preferential bending of the lamellae in one direction of the other⁴⁷⁻⁴⁹. It is not clear in this case what would cause unequal surface stresses although they cannot be ruled out as a potential chirality inducer.

5.4. Conclusions

In conclusion, a remarkable homochiral biomimetic structure has been discovered resulting from the self-assembly of an amphiphilic partially charged diblock copolypeptoid. The hierarchical internal ordering of the assemblies has been characterized in detail using x-ray scattering coupled with precise chemical modifications. The crystal structure of a small model molecule supports the model of self assembly. While the origin of the homochirality of these structures remains a mystery, it is clear that the interplay of hydrophobic and electrostatic forces is crucial to the formation of such a complex structure.

5.5. Appendix

Quantification of self assembly, the titration curve for $pNpe_{15}Nce_{15}$, a comparison of x-ray scattering in the solid state and in solution, differential scanning calorimetry data highlighting the crystallization differences between $pNpe_{15}$ and $pNpp_{15}$ in the solid state, an image gallery display the left-handedness of all super helices, and circular dichroism showing no optical rotation of light

5.5.1. Quantification of self-assembly

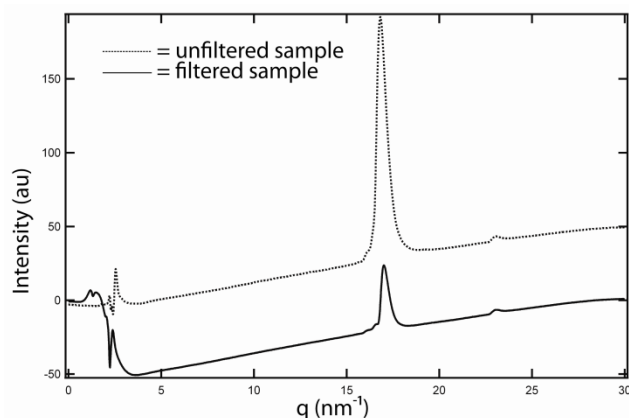


Figure 5.A.1. Self assembly quantification.

The HPLC traces of a one week old self assembly solution before filtering and after filtering. After filtering, the area of the peak in HPLC has shrunk by 83% indicating that 83% of the peptoid material was participating in a self assembled structure larger than 100K which would be true of any sheets or helices.

To assess the amount of material taking part in self-assembled structures, a 100K Pall centrifugal filter was used. After 1 week a self-assembly solution was divided in half with one half going through the Pall filter and one half remaining unfiltered. The two samples were then put through the analytical HPLC. As Figure 5.A.1 shows, after filtering a large amount of material was removed from the sample indicating that it is part of aggregates larger than 100K.

5.5.2. Titration

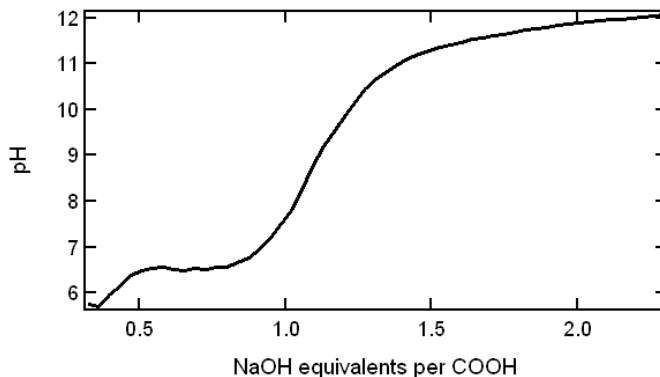


Figure 5.A.2. Titration curve.

The titration curve above was produced by dripping sodium hydroxide into a solution of $pNpe_{15}Nce_{15}$ while stirring continuously

Sodium hydroxide was added in a drip wise fashion while the solution was being stirred using a magnetic stir bar and plate. The pH was continually read using a Fisher Scientific AB15 pH meter. The titration curve is shown in Figure 5.A.2 with the x-axis converted to molar equivalents of OH⁻ groups per carboxylic acid group. A plateau appears at pH 6.5 when 0.5 equivalents of NaOH have been added. This is followed by a sharp increase in pH starting around 1.0 equivalents. This increase indicates that 100% of the carboxylic acid groups are deprotonated and any additional NaOH simply raises the pH of the solution. The self-assembly solutions are mixed at pH =6.5 using 0.5 equivalents of sodium hydroxide.

5.5.3. X-ray Scattering

X-ray scattering was performed on both solid and liquid samples. The solid samples were obtained by taking a self-assembly solution and evaporating the liquid. This solid sample was then placed in the x-ray beam. For a liquid sample, the self-assembly solution was concentrated 10x by evaporating the liquid and then drawn up into a capillary sample tube for shooting. The two methods for scattering gave qualitatively the same results, but the liquid sample scattering was much weaker and it was harder to identify peaks. Therefore all data in the main body of this report is on solid samples.

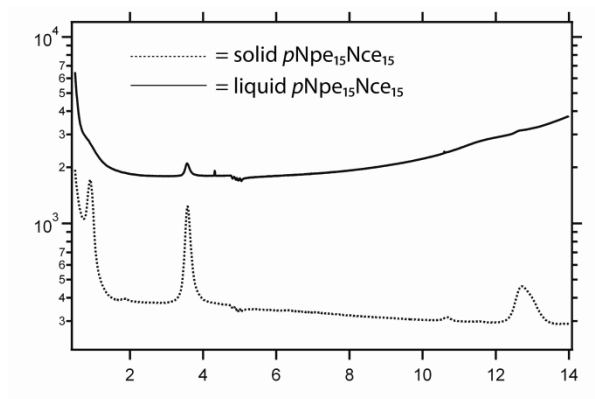


Figure 5.A.3. Solid SAXS matches liquid SAXS.

X-ray scattering on a liquid (dotted line) and solid (solid line) sample. The peaks qualitatively are the same although the liquid sample is far weaker.

5.5.4. Crystallization of $pNpp_{15}$

The self assembled super helix resulting from a diblock of $pNpp_{15}Nce_{15}$ was found not to have any internal crystal structure. The x-ray scattering trace for this structure is shown in the main body of this publication. It is hypothesized that the longer side chain length of the phenylpropyl group prevents the molecule from crystallizing. Differential Scanning Calorimetry data supports this hypothesis. The traces for the homopolymers of phenylpropyl, $pNpp_{15}$ and phenethyl, $pNpe_{15}$ show very different melting behavior (Figure 5.A.4). $pNpe_{15}$ has a clear melting and crystallization transition while $pNpp_{15}$ shows no melting transition before degradation around 250°C. This lack of crystallization in the solid state seems to be repeated in the solution state assembly as evidenced by the lack of intrachain crystallization peaks within the x-ray scattering patterns for the super helix structure. Interestingly, this lack of crystallization does not seem to inhibit helix formation.

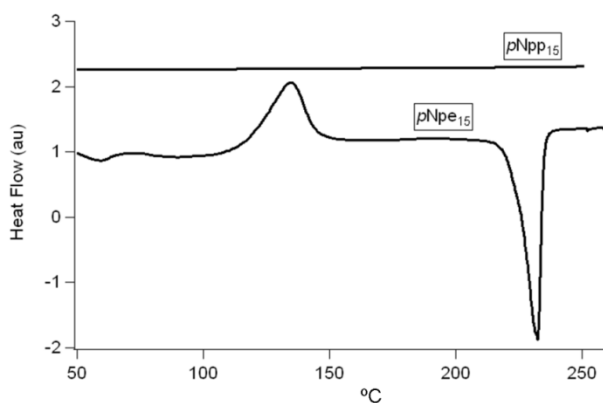


Figure 5.A.4. Crystallization of $pNpp_{15}$.

A DSC of both $pNpe_{15}$ and $pNpp_{15}$ demonstrating the lack of crystallization in the case of $pNpp_{15}$

5.5.6. Image Gallery

The image gallery below shows 15 helices aligned in a vertical direction such that it is possible to see their handedness. All of the helices are left handed, meaning that the pitch is tilted up and to the left in the images.

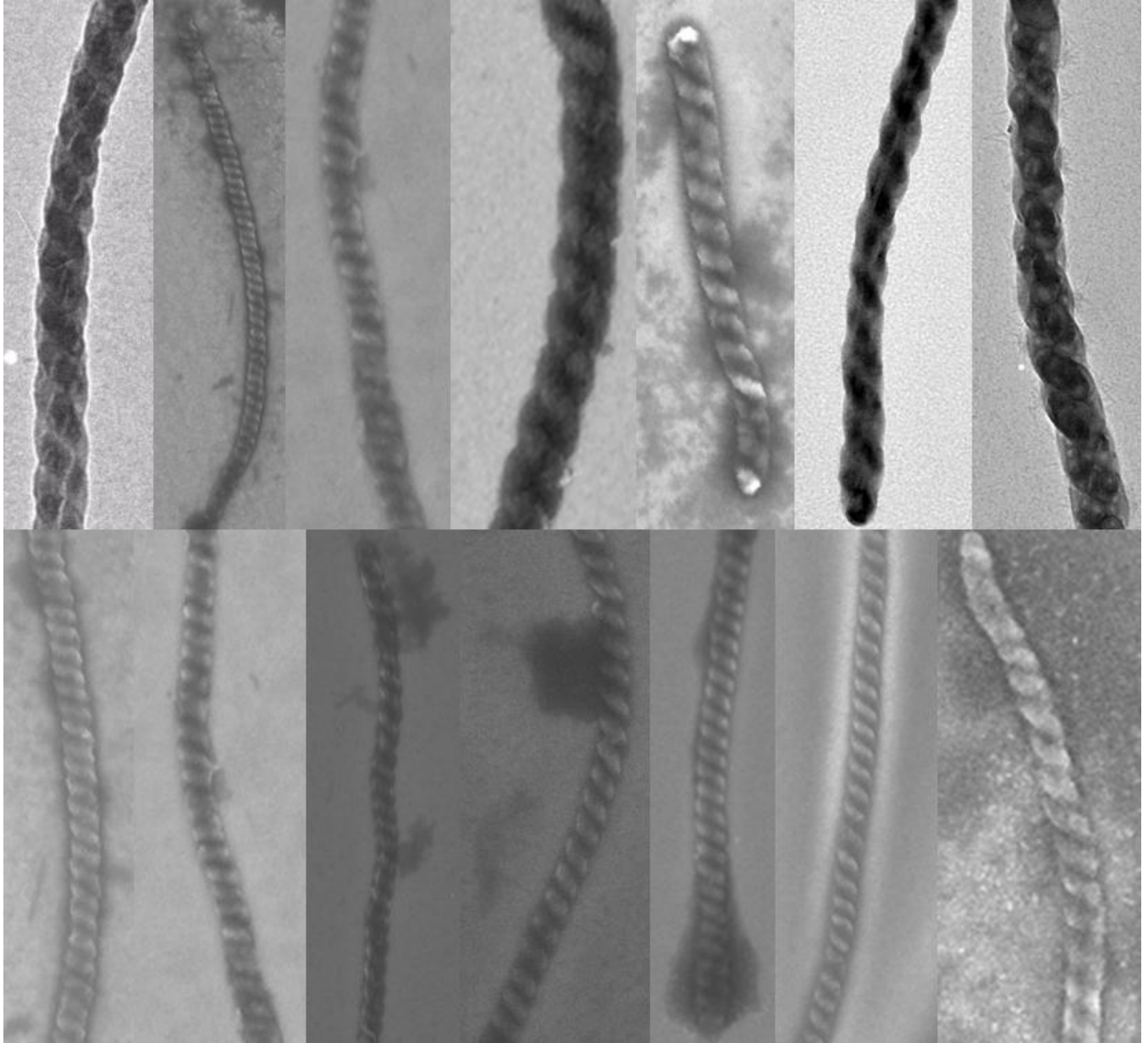


Figure 5.A.5. The image gallery of helices

5.5.7. Circular Dichroism

Circular dichroism was performed on a 300 μM solution to confirm the lack of chirality on a molecular level. As expected no rotation of light was seen, indicating a lack of molecular chirality.

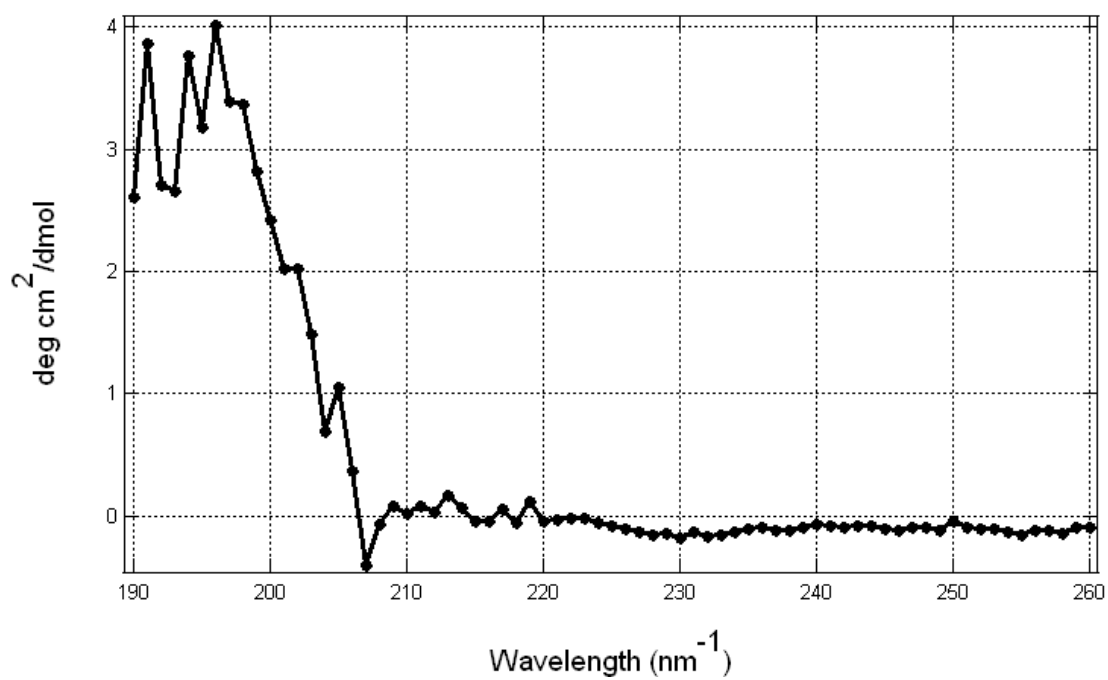


Figure 5.A.6. Circular dichroism spectra of a 300 μM solution of helices.
There is no molecular chirality as evidenced by the absence of optical rotation of light

5.6. Acknowledgements

This work was supported by the Office of Naval Research in the form of a Presidential Early Career Award in Science and Engineering (PECASE) for RAS. Polypeptoid synthesis and associated chemical characterization were performed at the Molecular Foundry, and XRD experiments were performed at the Advanced Light Source (ALS). Both are Lawrence Berkeley National Laboratory user facilities supported by the Office of Science, Office of Basic Energy Sciences, U.S. Department of Energy, under Contract No. DE-AC02-05CH11231. The authors thank Dr. James Holton and George Meigs for experimental assistance. H.K. Murnen acknowledges the Department of Defense for a NDSEG fellowship and A.M. Rosales acknowledges the National Science Foundation for a graduate fellowship. J.N. Jaworski acknowledges the Defense Threat Reduction Agency for financial support.

5.7. References

1. Fritz, M.; Morse, D. E., The formation of highly organized biogenic polymer/ceramic composite materials: the high-performance microaluminate of molluscan nacre. *Current Opinion in Colloid & Interface Science* 1998, 3 (1), 55-62.
2. Ottani, V.; Martini, D.; Franchi, M.; Ruggeri, A.; Raspanti, M., Hierarchical structures in fibrillar collagens. *Micron* 2002, 33 (7-8), 587-596.
3. Rajagopal, K.; Schneider, J. P., Self-assembling peptides and proteins for nanotechnological applications. *Curr. Opin. Struct. Biol.* 2004, 14 (4), 480-486.
4. Wong, G. C. L.; Tang, J. X.; Lin, A.; Li, Y. L.; Janmey, P. A.; Safinya, C. R., Hierarchical self-assembly of F-actin and cationic lipid complexes: Stacked three-layer tubule networks. *Science* 2000, 288 (5473), 2035-+.
5. Raviv, U.; Needleman, D. J.; Li, Y. L.; Miller, H. P.; Wilson, L.; Safinya, C. R., Cationic liposome-microtubule complexes: Pathways to the formation of two-state lipid-protein nanotubes with open or closed ends. *Proc. Natl. Acad. Sci. U. S. A.* 2005, 102 (32), 11167-11172.
6. Audette, G. F.; van Schalk, E. J.; Hazes, B.; Irvin, R. T., DNA-binding protein nanotubes: Learning from nature's nanotech examples. *Nano Lett.* 2004, 4 (10), 1897-1902.
7. Schueler-Furman, O.; Wang, C.; Bradley, P.; Misura, K.; Baker, D., Progress on modeling of protein structures and interactions. *Science* 2005, 310 (5748), 638-642.
8. Venkatraman, J.; Shankaramma, S. C.; Balaram, P., Design of folded peptides. *Chem. Rev.* 2001, 101 (10), 3131-3152.
9. Gazit, E., Self-assembled peptide nanostructures: the design of molecular building blocks and their technological utilization. *Chem. Soc. Rev.* 2007, 36 (8), 1263-1269.
10. de la Paz, M. L.; Goldie, K.; Zurdo, J.; Lacroix, E.; Dobson, C. M.; Hoenger, A.; Serrano, L., De novo designed peptide-based amyloid fibrils. *Proc. Natl. Acad. Sci. U. S. A.* 2002, 99 (25), 16052-16057.
11. Liu, D. H.; DeGrado, W. F., De novo design, synthesis, and characterization of antimicrobial beta-peptides. *J. Am. Chem. Soc.* 2001, 123 (31), 7553-7559.

12. Hales, K.; Pochan, D. J., Using polyelectrolyte block copolymers to tune nanostructure assembly. *Current Opinion in Colloid & Interface Science* 2006, *11* (6), 330-336.
13. Stuart, M. A. C.; Hofs, B.; Voets, I. K.; de Keizer, A., Assembly of polyelectrolyte-containing block copolymers in aqueous media. *Current Opinion in Colloid & Interface Science* 2005, *10* (1-2), 30-36.
14. Zhang, L. F.; Eisenberg, A., Multiple Morphologies of Crew-Cut Aggregates of Polystyrene-B-Poly(Acrylic Acid) Block-Copolymers. *Science* 1995, *268* (5218), 1728-1731.
15. Li, G. Y.; Song, S.; Guo, L.; Ma, S. M., Self-assembly of thermo- and pH-Responsive poly(acrylic acid)-b-poly(N-isopropylacrylamide) micelles for drug delivery. *J. Polym. Sci. Pol. Chem.* 2008, *46* (15), 5028-5035.
16. Xue, Y. N.; Huang, Z. Z.; Zhang, J. T.; Liu, M.; Zhang, M.; Huang, S. W.; Zhuo, R. X., Synthesis and self-assembly of amphiphilic poly(acrylic acid-b-DL-lactide) to form micelles for pH-responsive drug delivery. *Polymer* 2009, *50* (15), 3706-3713.
17. Gil, E. S.; Hudson, S. M., Stimuli-responsive polymers and their bioconjugates. *Prog. Polym. Sci.* 2004, *29* (12), 1173-1222.
18. Tycko, R., Progress towards a molecular-level structural understanding of amyloid fibrils. *Curr. Opin. Struct. Biol.* 2004, *14* (1), 96-103.
19. Orgel, J.; Irving, T. C.; Miller, A.; Wess, T. J., Microfibrillar structure of type I collagen in situ. *Proc. Natl. Acad. Sci. U. S. A.* 2006, *103* (24), 9001-9005.
20. Figliozzi, G. M.; Goldsmith, R.; Ng, S. C.; Banville, S. C.; Zuckermann, R. N., Synthesis of N-substituted glycine peptoid libraries. In *Combinatorial Chemistry*, Academic Press Inc: San Diego, 1996; Vol. 267, pp 437-447.
21. Rosales, A. M.; Murnen, H. K.; Zuckermann, R. N.; Segalman, R. A., Control of Crystallization and Melting Behavior in Sequence Specific Polypeptoids. *Macromolecule*, 2010.
22. Nam, K. T.; Shelby, S. A.; Choi, P. H.; Marciel, A. B.; Chen, R.; Tan, L.; Chu, T. K.; Mesch, R. A.; Lee, B.-C.; Connolly, M. D.; Kisielowski, C.; Zuckermann, R. N., Free-floating ultrathin two-dimensional crystals from sequence-specific peptoid polymers. *Nat Mater* 2010, *9* (5), 454-460.
23. Cornelissen, J.; Fischer, M.; Sommerdijk, N.; Nolte, R. J. M., Helical superstructures from charged poly(styrene)-poly(isocyanodipeptide) block copolymers. *Science* 1998, *280* (5368), 1427-1430.
24. Jeong, K. U.; Jin, S.; Ge, J. J.; Knapp, B. S.; Graham, M. J.; Ruan, J. J.; Guo, M. M.; Xiong, H. M.; Harris, F. W.; Cheng, S. Z. D., Phase structures and self-assembled helical suprastructures via hydrogen bonding in a series of achiral 4-biphenyl carboxylic acid compounds. *Chem. Mat.* 2005, *17* (11), 2852-2865.
25. Yang, W. S.; Chai, X. D.; Chi, L. F.; Liu, X. D.; Cao, Y. W.; Lu, R.; Jiang, Y. S.; Tang, X. Y.; Fuchs, H.; Li, T. J., From achiral molecular components to chiral supermolecules and supercoil self-assembly. *Chem.-Eur. J.* 1999, *5* (4), 1144-1149.

26. Ribo, J. M.; Crusats, J.; Sagues, F.; Claret, J.; Rubires, R., Chiral sign induction by vortices during the formation of mesophases in stirred solutions. *Science* 2001, 292 (5524), 2063-2066.
27. Hough, L. E.; Jung, H. T.; Kruerke, D.; Heberling, M. S.; Nakata, M.; Jones, C. D.; Chen, D.; Link, D. R.; Zasadzinski, J.; Heppke, G.; Rabe, J. P.; Stocker, W.; Korblova, E.; Walba, D. M.; Glaser, M. A.; Clark, N. A., Helical Nanofilament Phases. *Science* 2009, 325 (5939), 456-460.
28. Lin, S. C.; Lin, T. F.; Ho, R. M.; Chang, C. Y.; Hsu, C. S., Hierarchical Superstructures with Helical Sense in Self-Assembled Achiral Banana-Shaped Liquid Crystalline Molecules. *Adv. Funct. Mater.* 2008, 18 (21), 3386-3394.
29. Jeong, K. U.; Yang, D. K.; Graham, M. J.; Tu, Y. F.; Kuo, S. W.; Knapp, B. S.; Harris, F. W.; Cheng, S. Z. D., Construction of chiral propeller architectures from achiral molecules. *Adv. Mater.* 2006, 18 (24), 3229-+.
30. Figliozzi, G. M.; Goldsmith, R.; Ng, S. C.; Banville, S. C.; Zuckermann, R. N., Synthesis of N-substituted glycine peptoid libraries. *Combinatorial Chemistry* 1996, 267, 437-447.
31. Zhang, L. F.; Eisenberg, A., Multiple morphologies and characteristics of "crew-cut" micelle-like aggregates of polystyrene-b-poly(acrylic acid) diblock copolymers in aqueous solutions. *J. Am. Chem. Soc.* 1996, 118 (13), 3168-3181.
32. Pashuck, E. T.; Stupp, S. I., Direct Observation of Morphological Transformation from Twisted Ribbons into Helical Ribbons. *J. Am. Chem. Soc.* 2010.
33. Reches, M.; Gazit, E., Casting metal nanowires within discrete self-assembled peptide nanotubes. *Science* 2003, 300 (5619), 625-627.
34. Smith, A. M.; Williams, R. J.; Tang, C.; Coppo, P.; Collins, R. F.; Turner, M. L.; Saiani, A.; Ulijn, R. V., Fmoc-Diphenylalanine self assembles to a hydrogel via a novel architecture based on pi-pi interlocked beta-sheets. *Adv. Mater.* 2008, 20 (1), 37-+.
35. Blake, C.; Serpell, L., Synchrotron X-ray studies suggest that the core of the transthyretin amyloid fibril is a continuous beta-sheet helix. *Structure* 1996, 4 (8), 989-998.
36. Sunde, M.; Serpell, L. C.; Bartlam, M.; Fraser, P. E.; Pepys, M. B.; Blake, C. C. F., Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol.* 1997, 273 (3), 729-739.
37. Serpell, L. C., Alzheimer's amyloid fibrils: structure and assembly. *Biochim. Biophys. Acta-Mol. Basis Dis.* 2000, 1502 (1), 16-30.
38. Castelletto, V.; Hamley, I. W.; Hule, R. A.; Pochan, D., Helical-Ribbon Formation by a beta-Amino Acid Modified Amyloid beta-Peptide Fragment. *Angew. Chem.-Int. Edit.* 2009, 48 (13), 2317-2320.
39. Palacin, S.; Chin, D. N.; Simanek, E. E.; MacDonald, J. C.; Whitesides, G. M.; McBride, M. T.; Palmore, G. T. R., Hydrogen-bonded tapes based on symmetrically substituted diketopiperazines: A robust structural motif for the engineering of molecular solids. *J. Am. Chem. Soc.* 1997, 119 (49), 11807-11816.

40. Benedetti, E.; Marsh, R. E.; Goodman, M., Conformational studies on peptides - X-ray structure determinations of 6 N-methylated cyclic dipeptides derived from alanine, valine, and phenylalanine. *J. Am. Chem. Soc.* 1976, 98 (21), 6676-6684.
41. Nunez, L.; Brown, J. D.; Donnelly, A. M.; Whitlock, C. R.; Dobson, A. J., 1,4-Dibenzylpiperazine-2,5-dione. *Acta Crystallogr. Sect. E.-Struct Rep. Online* 2004, 60, O2076-O2078.
42. Green, M. M.; Reidy, M. P.; Johnson, R. J.; Darling, G.; O'leary, D. J.; Willson, G., Macromolecular Stereochemistry - the out-of-Proportion Influence of Optically-Active Co-Monomers on the Conformational Characteristics of Polyisocyanates - the Sergeants and Soldiers Experiment. *J. Am. Chem. Soc.* 1989, 111 (16), 6452-6454.
43. Wilson, A. J.; van Gestel, J.; Sijbesma, R. P.; Meijer, E. W., Amplification of chirality in benzene tricarboxamide helical supramolecular polymers. *Chemical Communications* 2006, (42), 4404-4406.
44. Otani, T.; Araoka, F.; Ishikawa, K.; Takezoe, H., Enhanced Optical Activity by Achiral Rod-Like Molecules Nanosegregated in the B-4 Structure of Achiral Bent-Core Molecules. *J. Am. Chem. Soc.* 2009, 131 (34), 12368-12372.
45. Brunsveld, L.; Schenning, A.; Broeren, M. A. C.; Janssen, H. M.; Vekemans, J.; Meijer, E. W., Chiral amplification in columns of self-assembled N,N',N''-tris((S)-3,7-dimethyloctyl)benzene-1,3,5-tricarboxamide in dilute solution. *Chemistry Letters* 2000, (3), 292-293.
46. Lightfoot, M. P.; Mair, F. S.; Pritchard, R. G.; Warren, J. E., New supramolecular packing motifs: pi-stacked rods encased in triply-helical hydrogen bonded amide strands. *Chemical Communications* 1999, (19), 1945-1946.
47. Keith, H. D.; Padden, F. J., Twisting orientation and the role of transient states in polymer crystallization. *Polymer* 1984, 25 (1), 28-42.
48. Lotz, B.; Cheng, S. Z. D., A critical assessment of unbalanced surface stresses as the mechanical origin of twisting and scrolling of polymer crystals. *Polymer* 2005, 46 (3), 577-610.
49. Ye, H.-M.; Wang, J.-S.; Tang, S.; Xu, J.; Feng, X.-Q.; Guo, B.-H.; Xie, X.-M.; Zhou, J.-J.; Li, L.; Wu, Q.; Chen, G.-Q., Surface Stress Effects on the Bending Direction and Twisting Chirality of Lamellar Crystals of Chiral Polymer. *Macromolecules* 2010.

Chapter 6. Conclusions and Future Outlook

Self-assembly of biomimetic macromolecules holds promise in the area of hierarchical nanostructures for applications such as templating inorganic materials, synthetic catalysis, or scaffolds for cell growth. The many different monomers present in these types of systems as well as the complex interactions between those monomers leads to correspondingly complex self-assembly behavior. Understanding and predicting this assembly requires a thorough understanding of the underlying polymer physics behavior of such macromolecules.

In this work a simplified biomimetic polymer system, specifically polypeptoids, has been studied in order to further develop the understanding of sequence specific self assembly. This work has investigated the persistence length of polypeptoids and how changes both in secondary structure (Chapter 2) and ionic groups (Chapter 3) can affect the persistence length. In Chapter 4, the transition of a polypeptoid chain containing only two types of monomers (hydrophobic and polar) from coil to globule was investigated and it was shown that the sequence of the monomers has a large affect on how the collapse behavior of the chain. Finally, in Chapter 5, an amphiphilic diblock copolypeptoid was shown to assemble into a super helix with ordering on many length scales ranging from angstroms up to the micron scale.

While this work provides a solid understanding of chain conformation for these biomimetic molecules there is much to do in order to fully understand self assembly of sequence specific molecules. For example, the driving forces for assembly into large super structures such as the super helix presented in Chapter 5 are still not completely understood. Further work on understanding the intermediate structures that make up the super structure may provide insight into the driving force for the assembly. In addition, while the superhelix is not immediately applicable for any function, the ability to design hierarchical materials could be very advantageous for mechanically robust materials (nacre, bone mimicking materials, etc) and it may be possible to use some of the same strategies shown here to develop these more functional materials.

This work has also provided a platform from which we can progress towards artificial folded structures. Here, the relatively simple coil to globule transitions were demonstrated, but further work on specific interactions will hopefully lead to more unique, well folded structures. For example, rather than using many non-specific and weakly hydrophobic groups to drive collapse, one could envision using fewer, stronger, and more specific hydrophobes such as phenyl groups to create targeted interactions for the molecule. This tactic might allow the repeatable folding of unique structures rather than the featureless globules seen here. Using known secondary structure motifs for polypeptoids such as alpha helices and beta-sheets will hopefully allow the development of tertiary structures that incorporate multiple functionalities. This will open up new frontiers in developing functional biomimetic materials as well as provide new insight into protein folding processes that can be used to aid in the prediction and development of polypeptides for particular functions.