Macromolecules

Article

pubs.acs.org/Macromolecules

Impact of Hydrophobic Sequence Patterning on the Coil-to-Globule Transition of Protein-like Polymers

Hannah K. Murnen,[†] Alexei R. Khokhlov,[‡,§] Pavel G. Khalatur,[§] Rachel A. Segalman,[†,∥] and Ronald N. Zuckermann*[,⊥,∥]

[†]Department of Chemical and Biomolecular Engineering, University of California, Berkeley, Berkeley, California 94720, United States

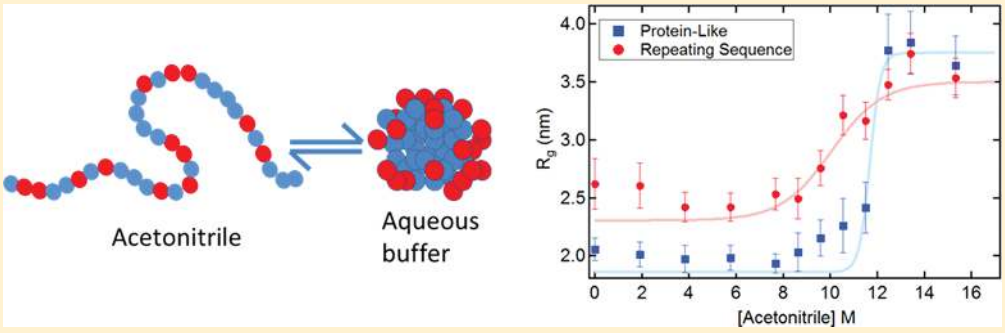
[‡]Department of Physics, Moscow State University, Moscow, Russia

[§]Department of Advanced Energy Related Nanomaterials, Ulm University, Ulm D-89069, Germany

[∥]Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

[⊥]The Molecular Foundry, Lawrence Berkeley National laboratory, Berkeley, California 94720, United States

Supporting Information



ABSTRACT: 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.

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. Dill and others have postulated that hydrophobicity serves as one of the most important driving forces for protein assembly.[1−3] To this end, they have developed a computational framework known as the HP model in which only two types of monomers, hydrophobic (H) and polar (P), are considered. With this dramatically reduced set of interactions, polymer sequences can be computationally designed to fold into defined structures.[3] A key experimental realization of this theory has been work by 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 residues.[4] Overwhelmingly, the mutated proteins still formed 4-helix bundles, demonstrating that the particular side chain was not as

Received: April 5, 2012
Revised: May 28, 2012
Published: June 8, 2012

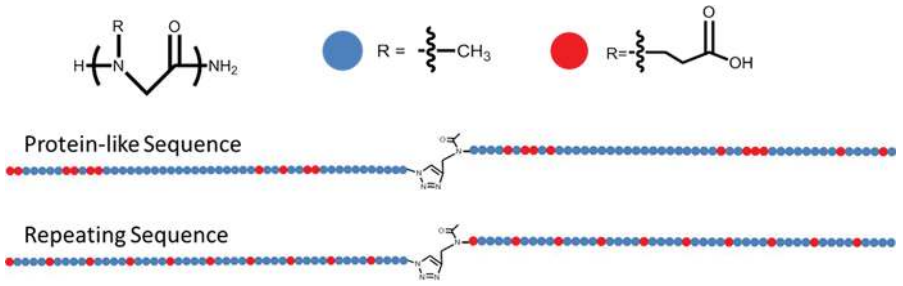


Figure 1. Protein-like and repeating sequence polypeptoid 100mers. The polypeptoids were synthesized by clicking two HPLC-purified 50mers together. Each monomer is represented by a circle where the red circles are the hydrophilic and polar *N*-(2-carboxyethyl)glycine (P) monomer while the blue circles are hydrophobic *N*-methylglycine (H) monomers. The protein-like sequence contains block sections of each type of monomer, while the repeating sequence 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.

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 folding.[5−9] 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 protein,[10] and hydrophobic interactions have been shown to have a large impact on this transition.[11,12] In fact, theoretical works have shown that for many polypeptide sequences, the likelihood of the chain forming a compact globule is relatively high and that the easiest way to disrupt compact globule formation is to exchange polar and nonpolar residues.[13,14] These works demonstrate that the position of hydrophobic and polar residues is very important for structure formation and folding. 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 monomers,[15] the degree of hydrophobicity of the hydrophobic monomers,[16] the overall length of the chain,[17] the importance of large contiguous sections of H or P monomers,[18] and even the ability of ionic interactions to impact the collapse process.[19] 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.[11,12,20] They also predicted that the coil-to-globule transitions 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 efforts[21−24] 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 co-workers elegantly demonstrated the chemical labeling of surface exposed monomers in a collapsed polystyrene globule, generating protein-like sequences.[25] However, assessment of exact chain sequence information was not possible.

Given the difficulties mentioned above, polypeptoids or poly(N-substituted glycine)s are an ideal model system for understanding sequence effects on the hydrophobic collapse of polymers 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 backbones[26] and self-assemble into protein-like structures in aqueous solution,[27−32] 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 lack chirality and the ability to form intrachain hydrogen bonds (due to the absence of any backbone hydrogen bond donors), allowing the study of hydrophobic forces in isolation. This makes polypeptoids 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 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. Growing interest in intrinsically disordered proteins including those with repeating[33] and polar sequences has provided both experimental[34,35] and theoretical[36] results showing that they also undergo coil-to-globule transitions. 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.

EXPERIMENTAL METHODS

System and Sequence Design. The protein-like sequence generation was carried out according to methods previously published.[11,12,20] The length and specific side chains were designed specifically for this synthetic system. The H and P monomer structures and their relative fractions were chosen to provide enough hydrophobicity to form a hydrophobic core and enough hydrophilicity to maintain solubility in aqueous solution. 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. This process is detailed in Figure 2. As a control sequence, a repeating sequence of the same monomer composition (80% H and 20% P) was chosen so as to minimize the length of any continuous region of H or P monomers. Thus, the repeating sequence consists of a simple pentamer repeat of (PHHHH), limiting the longest stretch of hydrophobic monomers to four residues.

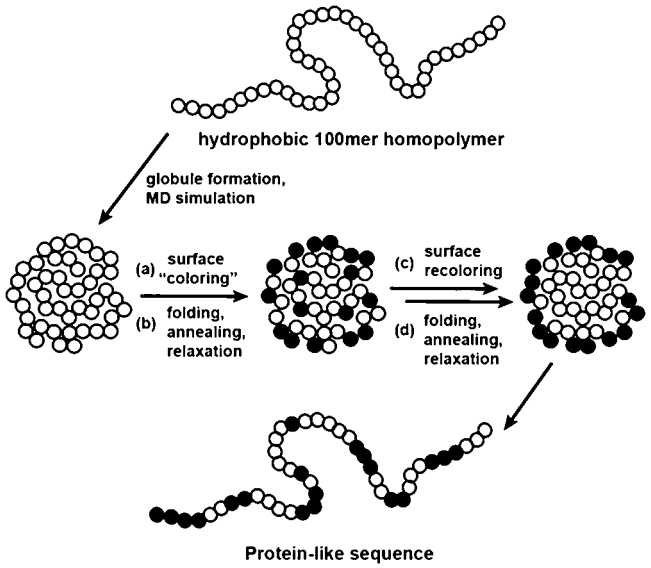


Figure 2. Protein-like polypeptoid sequences were generated by an iterative process involving successive rounds of globule formation using all-atom molecular dynamics and the addition/redistribution of polar residues on the globule surface. Open circles represent the hydrophobic monomer with a methyl side chain; black circles represent the polar monomer with a carboxylate side chain.

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.[37,38] All primary amine monomers, solvents, and reagents were purchased from commercial sources and used without further purification. The β-alanine-O-tBu·HCl was purchased from ChemImpex and used after freebasing by extraction from dichloromethane (DCM) and basic water. The resulting compound was confirmed by ^{1}H NMR. The β-alanine submonomer was dissolved in *N*-methylpyrrolidinone at a concentration of 1.5 M, while the methylamine was used directly as purchased as a 40% w/v solution in water. All displacement times were 60 min for the first 15 monomer addition cycles, 90 min for the next 15 and 120 min for the remainder. All polypeptoids were acetylated on the resin and purified using reverse phase HPLC as previously described.[37] The mass and purity were confirmed using reverse phase analytical HPLC and MALDI and representative traces and spectra are shown in Table 1 and Figure 3. 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 propargylamine 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 1) was performed by reacting the propargylated compound **A** (23 mM) with 2 equiv of azide-modified compound **B** in 200 μL of water. Copper(II) sulfate (5 equiv) and ascorbic acid (6 equiv) were added, and the solution was mixed at 70 °C for 24 h. The excess copper was removed by stirring over basic alumina for 2 h. 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 3.

Table 1. HPLC and Mass Spectrometry of 50mer and 100mer Peptoids

compound	length	observed mass/expected mass	% purity[a]
protein-like_a	50	4290.3/4288.4	98
protein-like_b	50	4232.0/4232.4	97
repeating_a	50	4289.6/4288.4	97
repeating_b	50	4235.9/4232.4	96
protein-like	100	8512/8517.5	97
repeating	100	8514/8517.5	99

[a]After HPLC purification.

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 ∼30 s to ensure dissolution of the molecules and were then annealed at 70 °C for 4 h.

Two-State Model. The equation below was fit to the R_g measurements with $R_{g,coil}$ and $R_{g,globule}$ calculated by averaging the three points at either end of the curve. The percentage of acetontrile in the solution is

$$E = \frac{R_{g,coil} + R_{g,globule}\exp(\Delta G_{CG} - m[acn])RT}{1 + \exp(\Delta G_{CG} - m[acn])RT}$$

represented by [acn]. R is the ideal gas constant (kcal/(mol K)$^{-1}$) and T is temperature in kelvin. $\Delta G_{CG}(H_2O)$ 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

dx.doi.org/10.1021/ma300707t | *Macromolecules* 2012, 45, 5229−5236

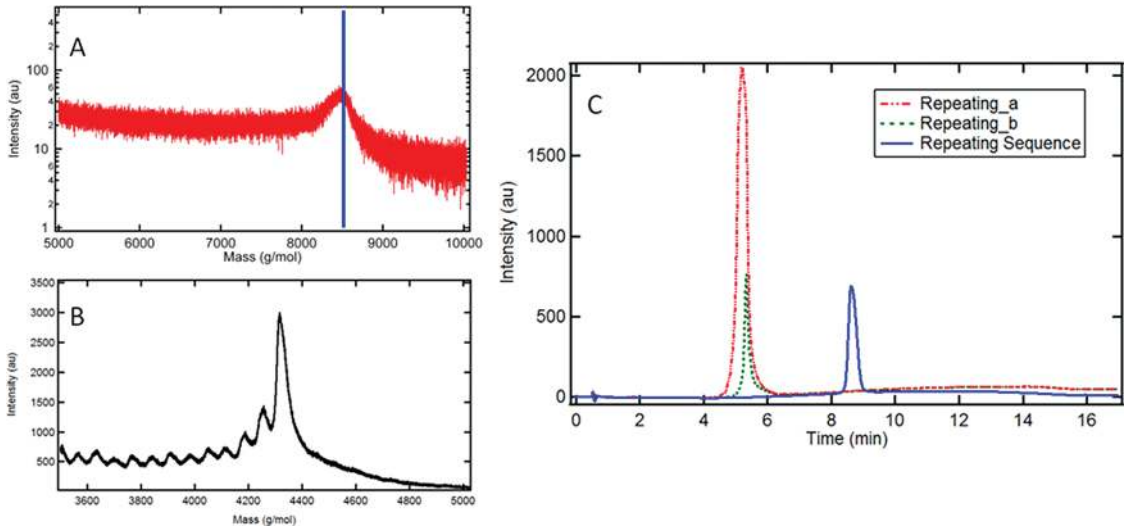
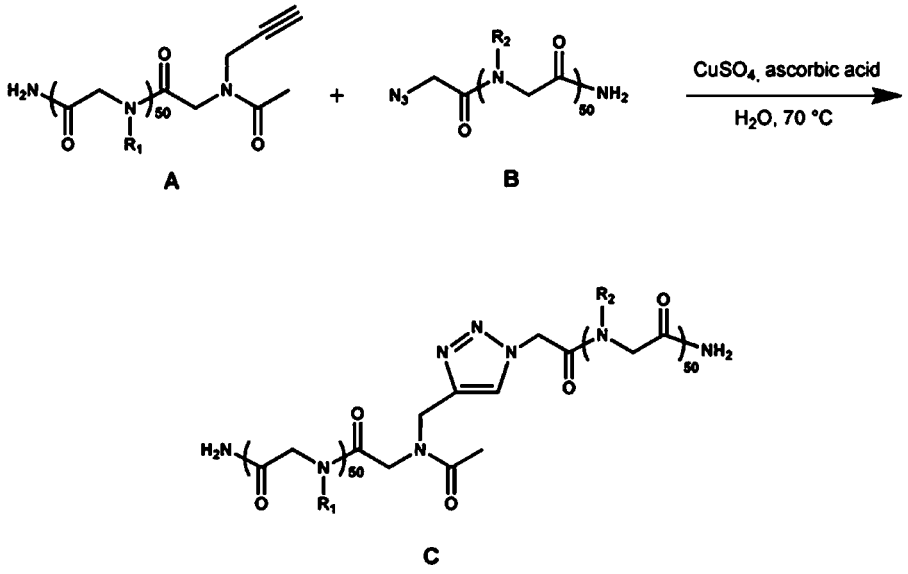


Figure 3. Representative analytical traces and MALDI spectra. MALDI spectrum of peptoid 100mer (a) shows a molecular weight of 8517 while the spectrum of a propargylated 50mer, repeating_a (b), shows a molecular weight of 4289. (c) Analytical reverse-phase HPLC traces show both 50mers eluted just after 5 min in a 0−50% acetonitrile gradient over 17 min, while the 100mer product eluted at close to 9 min.

Scheme 1. Azide−Alkyne Cycloaddition Reaction Was Employed To Ligate Two 50mers To Create a 100mer of Defined Sequence



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 a 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 λ = 1.240 Å 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 postion chamber intensity using Nika version 1.18. At SSRL, the beamline was configured with an X-ray wavelength of λ = 1.488 Å 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 were plotted on a Guinier plot and fitted to obtain a radius of gyration.

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 1). The simulation work leading to the generation of the protein-like sequence is detailed in the Experimental Methods as well as the synthetic details. The H monomer, an *N*-methylglycine (or sarcosine), was chosen to be

dx.doi.org/10.1021/ma300707t | *Macromolecules* 2012, 45, 5229−5236

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. Experiments were performed at pH 8 in order to maintain the 2-carboxyethyl side chains in the predominantly negative state, making them highly polar and precluding any hydrogen bonding between them and the backbone. 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 created to date. The peptoid polymers were dissolved at a concentration of 1 mg/mL (120 μM) in 25 mM 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 the protein-like sequence in buffer was determined to be 2.1 ± 0.09 nm, while the R_g of the repeating sequence was 2.6 ± 0.21 nm in buffer. 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.[39] 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 radius of gyrations can be converted into densities using the volume of a polymer and the volume of the globule and water for the resulting makeup of the globule (see Supporting Information for this calculation). Using this conversion, the density of the protein-like sequence is 1.070 g/cm^3 vs 1.033 g/cm^3 for the repeating sequence. The higher density of the protein-like sequence as compared to the repeating sequence indicates that it is more tightly folded in its globule state even though the two polymers are of identical molecular weight and chemical composition. The longer blocks of hydrophobic and hydrophilic residues in the protein-like sequence allow the molecule to collapse more fully while the repeating sequence remains in a loosely collapsed state, resulting in a lower density.

Equilibrium denaturant titration with increasing concentrations of a less polar 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), 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 states with R_g's of about 3.5 nm. The repeating sequence begins to show an increased R_g around 7 M acetonitrile after which its R_g slowly increases, demonstrating

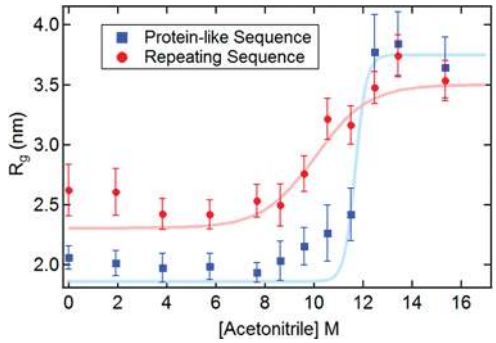


Figure 4. Equilibrium denaturant titration of peptoid globules with acetonitrile. The radius of gyration as determined by SAXS was determined as a function of acetonitrile concentration for the protein-like and repeating sequences. The polymers were dissolved at 1 mg/mL (120 μM) in 25 mM Tris, pH 8.0. The blue squares are for the protein-like sequence, while the red circles are for the repeating sequence and the lines were fit using a two-state model.

a gradual unfolding transition. In contrast, the protein-like sequence stays folded until about 10 M 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 can be approximated by fitting the R_g data using a two-state model as described by Santoro and Bolen[40] and as previously utilized for folded polypeptoids.[28] Using the method detailed in the Supporting 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 2). The larger $\Delta G_{CG}(H_2O)$ for the protein-like sequence (42.5 kcal/mol vs 7.7 kcal/mol) can be attributed to a lower free energy for the globule state of the protein-like sequence as compared to the repeating sequence. The unfolding energy is therefore significantly larger ($\Delta\Delta G_{CG}(H_2O)$ is 34.8 kcal/mol). The value of m for the protein-like sequence is also significantly greater than that for compound **2** (3.6 kcal mol^{-1} M^{-1} vs 0.7 kcal mol^{-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 site.[41] 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 transition,[42]

Table 2. Calculated Values from a Two-State Model Fitted to the R_g Values Obtained from Equilibrium Acetonitrile Titration of Both Peptoids[a]

compound	$\Delta G_{CG}(H_2O)$ (kcal/mol)	m (kcal/(mol M^{-1}))
protein-like	42.5 ± 1.0	3.63 ± 0.08
repeating	7.7 ± 0.5	0.73 ± 0.05

[a]Higher values of both $\Delta G_{CG}(H_2O)$ and m for the protein-like sequence indicate increased driving force and cooperativity for the collapse transition.

dx.doi.org/10.1021/ma300707t | *Macromolecules* 2012, 45, 5229−5236

indicating that perhaps the coil−globule transition has an inherently higher energy difference that that of protein-folding. Regardless, the difference in $\Delta G_{CG}(H_2O)$ for the two compounds 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 the protein-like sequence is 3.4 nm while the R_h of the repeating sequence is 3.3 nm (Figure 5).

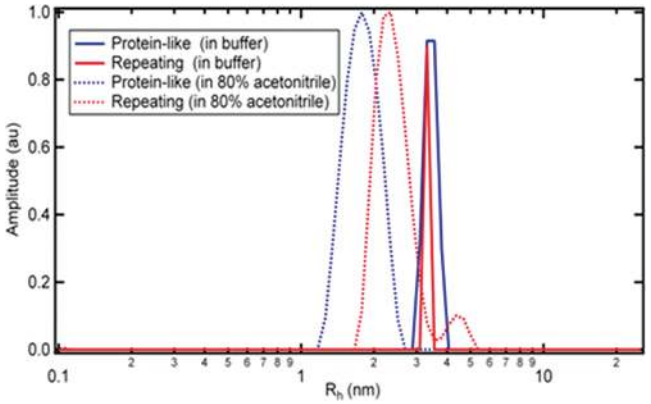


Figure 5. Hydrodynamic radius as measured using dynamic light scattering. The blue lines are for protein-like while the red lines are the repeating sequence. The R_h's that are measured are then used to calculate ρ, the ratio of R_g/R_h.

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 polymer.[43] This value is theoretically predicted for both a tightly compacted spherical globule (0.77) and for an expanded coil (1.5−2).[44] The R_h was measured in aqueous conditions as well as in 80% acetonitrile solution and ρ was subsequently calculated (Table 3).

Table 3. Characteristic Ratio, ρ, for Each Compound in Buffer and in 80% Acetonitrile

compound	ρ (in buffer)	ρ (in 80% acetonitrile)
protein-like sequence	0.62	1.97
repeating sequence	0.79	1.60

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 nonspherical 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.

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 environment,[45] leading to its use as a probe for hydrophobic cores in molten globule or folded proteins.[46] 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 (120 μM) as well as a control solution of a homopolypeptoid 36mer containing only relatively polar *N*-2-methoxyethyl side chains. In Figure 6, the emission peaks for each of these polymers is

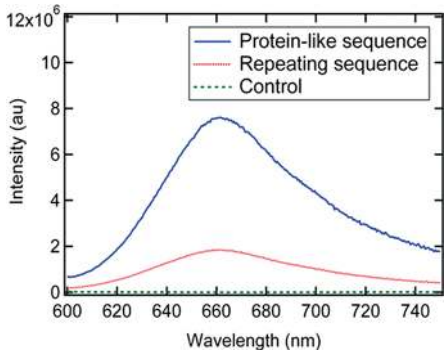


Figure 6. Nile red fluorescence showing an increased emission peak from the protein-like sequence. All solutions were excited at 586 nm, and the emission peaks are shown above. The repeating sequence clearly has a decreased emission peak when compared to the protein-like compound, indicating a looser globule formation. The control sequence is a homopolymer of 36 *N*-(2-methoxyethyl)glycine residues, and as expected it shows no fluorescence.

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, the protein-like sequence has a strong emission peak, indicating that the hydrophobic core of the collapsed globule is the most isolated from the hydrophilic environment. The repeating sequence shows weaker fluorescence (~20% of the protein-like molecule) due to the fact that it is forming a globule in solution as evidenced by the value of ρ, indicating globule formation (Table 3). There is essentially no shift of the peak maximum (protein-like sequence, 663 nm; repeating sequence, 662 nm), suggesting that the signal is still coming from Nile red molecules located within a hydrophobic region in the repeating case, but the region is providing less intense signal than that from the protein-like molecule. This supports the conclusion from the previous scattering (SAXS) measurements that a globule is formed from the repeating sequence. However, this globule appears to be less well folded, leading to a lower intensity of emission as compared to the protein-like molecule. 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.

■ CONCLUSION

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

dx.doi.org/10.1021/ma300707t | *Macromolecules* 2012, 45, 5229−5236

allowed the direct study of the impact of hydrophobic sequence patterning in complete isolation from 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 was 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 vehicles.[47] 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.

■ ASSOCIATED CONTENT

Supporting Information

Guinier plots used to calculate the radius of gyration for each solution and the globule density calculations. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Dr. 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 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.

■ 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) Lau, K. F.; Dill, K. A. Theory for protein mutability and biogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87* (2), 638−642.

(14) Chan, H. S.; Dill, K. A. Sequence-space soup of proteins and copolymers. *J. Chem. Phys.* **1991**, *95* (5), 3775−3787.

(15) Ashbaugh, H. S. Tuning the Globular Assembly of Hydrophobic/Hydrophilic Heteropolymer Sequences. *J. Phys. Chem. B* **2009**, *113* (43), 14043−14046.

(16) 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.

(17) 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.

(18) 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.

(19) 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.

(20) 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, 2006; Vol. *195*, pp 1−100.

(21) 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−+.

(22) 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.

(23) 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.
(24) 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.
(25) 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.
(26) 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.
(27) 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.
(28) 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.
(29) 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.
(30) 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.
(31) 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.
(32) 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.
(33) Jorda, J.; Xue, B.; Uversky, V. N.; Kajava, A. V. Protein tandem repeats - the more perfect, the less structured. *FEBS J.* **2010**, *277* (12), 2673−2682.
(34) Crick, S. L.; Jayaraman, M.; Frieden, C.; Wetzel, R.; Pappu, R. V. Fluorescence correlation spectroscopy shows that monomeric polyglutamine molecules form collapsed structures in aqueous solutions. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (45), 16764−16769.
(35) Mukhopadhyay, S.; Krishnan, R.; Lemke, E. A.; Lindquist, S.; Deniz, A. A. A natively unfolded yeast prion monomer adopts an ensemble of collapsed and rapidly fluctuating structures. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (8), 2649−2654.
(36) Vitalis, A.; Lyle, N.; Pappu, R. V. Thermodynamics of beta-Sheet Formation in Polyglutamine. *Biophys. J.* **2009**, *97* (1), 303−311.
(37) 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.
(38) Figliozzi, G. M.; Goldsmith, R.; Ng, S. C.; Banville, S. C.; Zuckermann, R. N. Synthesis of N-substituted glycine peptoid libraries. *Comb. Chem.* **1996**, *267*, 437−447.
(39) Abrego, J. R. B.; Craievich, A. F.; Mascarenhas, Y. P.; Laure, C. J. SAXS study of Crotapotin at low pH. *Biophys. J.* **1993**, *64* (2), 560−564.
(40) Santoro, M. M.; Bolen, D. W. Unfolding free-energy changes determined by the linear extrapolation method 0.1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* **1988**, *27* (21), 8063−8068.
(41) 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.
(42) Regan, L.; Degrado, W. F. Characterization of a helical protein designed from 1st principles. *Science* **1988**, *241* (4868), 976−978.
(43) Tanford, C. *Physical Chemistry of Macromolecules*; John Wiley and Sons: New York, 1961.
(44) Burchard, W. Static and dynamic light-scattering from branched polymers and bio-polymers. *Adv. Polym. Sci.* **1983**, *48*, 1−124.
(45) Sackett, D. L.; Wolff, J. Nile red as a polarity-sensitive fluorescent probe of hydrophobic protein surfaces. *Anal. Biochem.* **1987**, *167* (2), 228−234.
(46) Hawe, A.; Sutter, M.; Jiskoot, W. Extrinsic fluorescent dyes as tools for protein characterization. *Pharm. Res.* **2008**, *25* (7), 1487−1499.
(47) Davis, M. E.; Chen, Z.; Shin, D. M. Nanoparticle therapeutics: an emerging treatment modality for cancer. *Nat. Rev. Drug Discovery* **2008**, *7* (9), 771−782.