Macromolecules

pubs.acs.org/Macromolecules Article

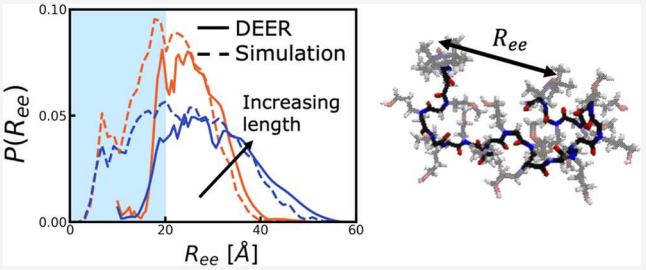
Quantifying Polypeptoid Conformational Landscapes through Integrated Experiment and Simulation

Sally Jiao, Audra DeStefano, Jacob I. Monroe, Mikayla Barry, Nicholas Sherck, Thomas Casey, Rachel A. Segalman, Songi Han,* and M. Scott Shell*

Cite This: *Macromolecules* 2021, 54, 5011−5021 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: We combine experiment and simulation to develop a powerful, validated approach for characterizing the conformational landscapes of disordered polypeptoids. Polypeptoids have become an important class of polymers, capable of precisely defined sequences while remaining gram-synthesizable—properties that have driven a rapidly expanding set of applications, including antifoulants, therapeutics, sensing, and directed self-assembly. The characterization of polypeptoid structure provides critical molecular insight into sequence−structure−function relationships. Structurally disordered polypeptoids require new approaches to interrogate their wide range of conformations in solution. Here, we measure full end-to-end distance distributions, instead of configurational averages, using double electron−electron resonance (DEER) spectroscopy and enhanced sampling molecular modeling. We demonstrate excellent agreement between the experiments and simulations for a set of model hydrophilic polypeptoids. Moreover, we illustrate the utility of this combined experiment−simulation approach in probing structure−function relationships by characterizing the basic polymer physics of this polypeptoid series, demonstrating that the polypeptoids probed here exhibit excluded volume behavior.



1. INTRODUCTION

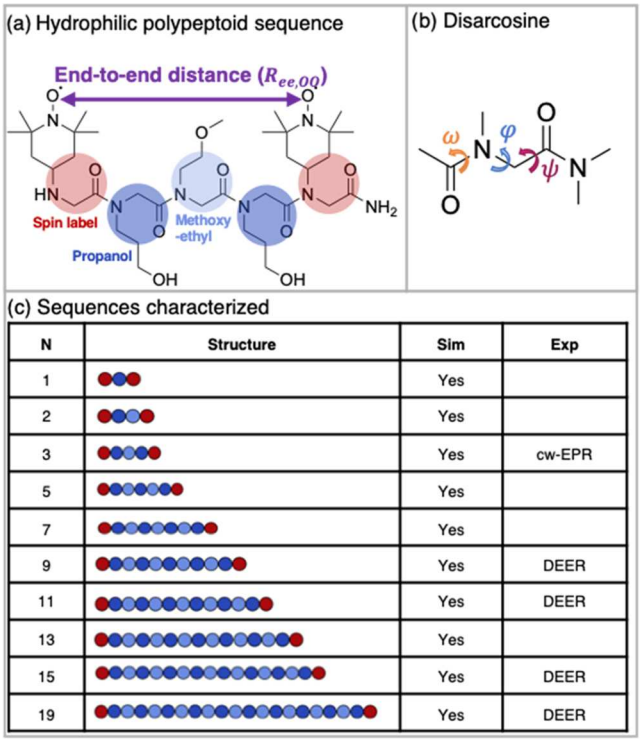
Polypeptoids, or N-substituted glycines, possess both the robustness of synthetic polymers and the tunability of polypeptides and other biomolecules.[1] They can now be routinely synthesized at gram scale, sequence-specifically up to 50 monomers in length in a process that eliminates the need for protection/deprotection procedures common in polypeptide synthesis, allowing access to a wide variety of chemistries.[1] As such, they have been extensively explored for a broad range of applications, including as antifoulants,[2] antimicrobials,[3−5] other therapeutics,[6] delivery agents,[7,8] sensing applications[9] such as stimuli-responsive materials,[10] binder materials for batteries,[11] crystallization modulators,[12] and programmable self-assembly.[13−15] Many of these applications leverage the sequence-specificity of polypeptoids that has been shown to significantly modulate secondary[16−20] and self-assembled structures[21,22] and therefore function. Thus, advancement of polypeptoid-containing materials and the fundamental design rules that they follow requires elucidation of sequence−structure−function relationships.

Many polypeptoids used in these applications are structurally disordered and adopt a wide range of conformations. Sequence has been shown to tune functional properties even in these structurally disordered cases. For example, variations in the location and number of hydrophobic or charged monomers can modify antifouling and fouling release,[23,24] loading and release in drug delivery,[25] gene transfection efficiency,[26] and the degree of melting point depression,[27] among other properties. Further evidence suggests that shifts in the conformational distributions of these disordered polypeptoids, modulated by the polypeptoid sequence, contribute to changes in function.[27−29] However, the difficulty of probing the structure of disordered polypeptoids both experimentally and computationally has hindered the development of fundamental design rules to guide the discovery of novel and efficacious polypeptoids. Experiments typically report on dominant or average structures, obscuring variations within the ensemble of conformations. On the other hand, modeling provides direct access to the conformational landscape through insight into atomic-scale structures and structural correlations. However, polypeptoid simulations encounter major sampling challenges due to the long time scales associated with conformational transitions, which has also hampered the validation of polypeptoid simulation models.

Received: March 11, 2021
Revised: April 29, 2021
Published: May 25, 2021

https://doi.org/10.1021/acs.macromol.1c00550
Macromolecules 2021, 54, 5011−5021

In this work, we pursue a novel combined experimental–computational approach of measuring and comparing structural ensembles, as opposed to dominant or average structures, to elucidate a fundamental picture of polypeptoid conformational landscapes. Specifically, we measure the probability distribution of the distance between the ends (end-to-end distance) of a hydrophilic polypeptoid of varying length (Figure 1a,c) using double electron–electron resonance



N	Structure	Sim	Exp
1		Yes	
2		Yes	
3		Yes	cw-EPR
5		Yes	
7		Yes	
9		Yes	DEER
11		Yes	DEER
13		Yes	
15		Yes	DEER
19		Yes	DEER

Figure 1. Hydrophilic polypeptoids used here to cross-validate the experimental and computational workflows are synthesized, characterized, and simulated. (a) Polypeptoids comprise an alternating sequence of hydrophilic monomers terminated on both ends by spin labels for electron paramagnetic resonance, which measures the distance between the oxygen radicals, $R_{ee,OO}$, shortened to R_{ee} in the text. (b) Disarcosine is simulated to validate the expanded ensemble sampling. The arrows show rotation around the ω, φ, and ψ dihedrals in the polypeptoid backbone. (c) Polypeptoids of varying lengths are probed both through DEER and cw-EPR experiments and through MD simulation. Note that N is the number of hydrophilic monomers, not including the terminal spin labels.

(DEER) spectroscopy and leverage recently developed inversion techniques to accurately extract these broad distance distributions. We measure the same distributions computationally using an expanded ensemble simulation technique that directly speeds up the sampling of the slow backbone bond rotations to resolve the full equilibrium conformational ensemble. The hydrophilic polypeptoid we choose to study (Figure 1a) comprises an alternating sequence of methoxyethyl and propanol side chains. The former have been previously studied as promising antifoulants[2] and as model polar side chains in self-assembling systems[22] and the latter enhance the polypeptoid's solubility in water.

DEER is an electron paramagnetic resonance technique that provides unique access to conformational ensemble information by measuring the full distance distribution from a composite time-domain signal that reports on distances between pairs of spin labels.[30] It is especially advantageous for comparison with simulations because its probes are relatively small,[31] minimally perturb the conformational ensemble,[31] and can be directly simulated.[32] DEER has traditionally been used to probe distances in structured biomacromolecules[30] or average distances in unstructured systems. In recent years, its ability to detail the full probability distribution over accessible distance ranges has garnered significant interest. The application of DEER to characterize disordered biological and synthetic polymeric systems through this distance distribution is emerging,[33−39] and methodological developments are continually improving the accurate resolution of the shape as well as the width of broad distributions.[40−44] Recent work on aqueous solutions of polyethylene oxide, a canonical disordered, synthetic polymer, demonstrates consistency between the broad distance distributions determined by DEER and simulation,[39] suggesting that DEER is well positioned to characterize the conformational landscape of disordered macromolecules. Single-molecule Förster resonance energy transfer (FRET) also probes conformational distributions,[45] but DEER requires fewer assumptions[46,47] and captures far more molecules, resulting in more accurate distance distributions.

The accessible distance range for DEER is generally considered to be between 20 and 80 Å.[30,48] Supplementing DEER at shorter distances, we use the analysis of continuous wave electron paramagnetic resonance (cw-EPR) line shapes where the dipolar interaction between two spin labels with distances between 8 and 25 Å leads to broadening of the spectrum that can be related to the distances between them by spectral simulation. Unlike DEER, cw-EPR cannot fully resolve multifeatured distributions but is a complementary probe of the existence and approximate width and character of the distribution in the short-distance region.

Simulations of polypeptoids present unique challenges because the polypeptoid amide bond isomerizes on long time scales (seconds[49]) relative to achievable simulation lengths due to a large (∼17 kcal/mol[50]) barrier to isomerization.[51] Polypeptoid forcefields have a correspondingly large barrier in the ω amide-bond dihedral potential, which hinders simulation studies of appreciable chain lengths since the number of isomerization states expands combinatorically with the degree of polymerization. To accurately probe conformational space, the sampling technique must accelerate cis–trans isomerizations for each amide bond. Here, we present and validate a technique to sample amide-bond dihedral isomerization in long polypeptoids by simulating in an expanded ensemble of states in which the amide-bond dihedral potentials are scaled down, after which properties are rigorously reweighted back to the original state. This approach, which has previously been used to enhance conformational sampling in small molecules[52] but has not yet been applied to polypeptoids, addresses many of the challenges of select earlier simulation strategies that do not extend well to longer chains in explicit solvent.[53−55] More recently, the metadynamics technique has provided an efficient and targeted solution to specifically enhance sampling of the slow degrees of freedom.[56−58] However, the successful tuning of key algorithm convergence parameters can be difficult and a significant amount of simulation time is spent exploring regions of phase space inconsequential to the equilibrium ensemble. The approach presented here similarly targets specific degrees of freedom, the amide-bond isomerizations, and demonstrates

https://doi.org/10.1021/acs.macromol.1c00550
Macromolecules 2021, 54, 5011−5021

excellent convergence properties while remaining extensible to longer chains.

For the model polypeptoid studied here, the combined simulation and DEER workflows show excellent agreement, both in the shape of the end-to-end distance distributions and in the scaling of the mean end-to-end distance with increasing polypeptoid length that furthermore suggests that these hydrophilic polypeptoids obey excluded volume behavior. The distance distribution extracted using cw-EPR for a shorter polypeptoid also agrees with simulation, offering an important supplemental experimental probe of short polypeptoids.

2. METHODS

2.1. Polypeptoid Synthesis. Polypeptoid sequences varying the number of alternating hydrophilic monomers (Figure 1a), N = 3, 9, 11, 15, and 19, and containing spin labels on both ends are synthesized using an automated Prelude peptide synthesizer following previously established methods.[59] Full details on polypeptoid synthesis are included in Section S1.

2.2. Experimentally Probing Conformational Landscapes with 4-Pulse DEER. DEER spectroscopy probes conformational distributions, rather than just averaged quantities, because the time-domain signal is a composite representing the product of individual dipole−dipole interactions of the molecular ensemble that can be decomposed into its constituents. Figure 2a−c provides an overview of the workflow by which DEER spectroscopy generates an end-to-end distance distribution. DEER involves the application of microwave pulses to doubly spin-labeled macromolecules in a static magnetic field according to the sequence illustrated in Figure 2a. The microwave pulses with the (probe) frequency ω_A excite a subset of electron spins (called "ω_A spins"), while the pulse with the (pump) frequency ω_B excites a separate subset (called "ω_B spins"). The ω_A sequence leads to a refocused electron spin echo (ESE) at time $(2\tau_1 + \tau_2)$ with an amplitude that depends primarily on the bandwidths of the pulses and the relaxation times of the excited spins. With the application of the pump pulse at ω_B that results in the excitation of ω_B spins, any of the ω_A spins that are dipolar-coupled to ω_B spins will experience changes in their local magnetic fields, communicated through their dipolar couplings. Since the relaxation mechanisms that govern the contributions of ω_A spins to the refocused ESE depend on the local magnetic fields at the ω_A spins, this will have the effect of attenuating the ESE relative to the ω_A sequence in the absence of the pump pulse at ω_B. As the ω_B pulse is moved in time between the second and third ω_A pulses, the ESE amplitude is modulated at the frequencies corresponding to the strengths of the dipolar couplings between ω_A and ω_B spins. In this way, a time-domain signal is generated that consists of frequencies that can be related to dipole−dipole distances. In the time domain, the DEER signal consists of a form factor, $F(t)$, that contains the dipole−dipole interaction information of interest and a background decay. The background decay is owed primarily to intermolecular interactions and is removed using the division of the overall signal by (typically) a stretched-exponential function.[40] We collect the DEER time-domain signal using a Bruker QT-II resonator in a pulsed Q-band Bruker E580 Elexsys spectrometer with a 300 W TWT amplifier (Applied Systems Engineering, Model 177 Ka), as described in Section S2. Raw time-domain data are shown in Figure S7, and background-corrected time-domain data are shown in Figure 6.

Most commonly, the conversion of $F(t)$ to a distance distribution, $P(r)$, is achieved using Tikhonov regularization, where the mathematical inversion of $F(t)$ occurs through a modified least-squares regression to find a solution for $P(r)$.[40] This method, however, can struggle to find a unique solution, especially when $F(t)$ represents a broad and/or multimodal $P(r)$. Method development for deriving reliable $P(r)$ from $F(t)$ is an active area of research, and several recent publications propose various solutions, e.g., neural network processing,[60] DeerLab,[40] and singular value decomposition.[43,44] Of particular interest are the techniques using singular value

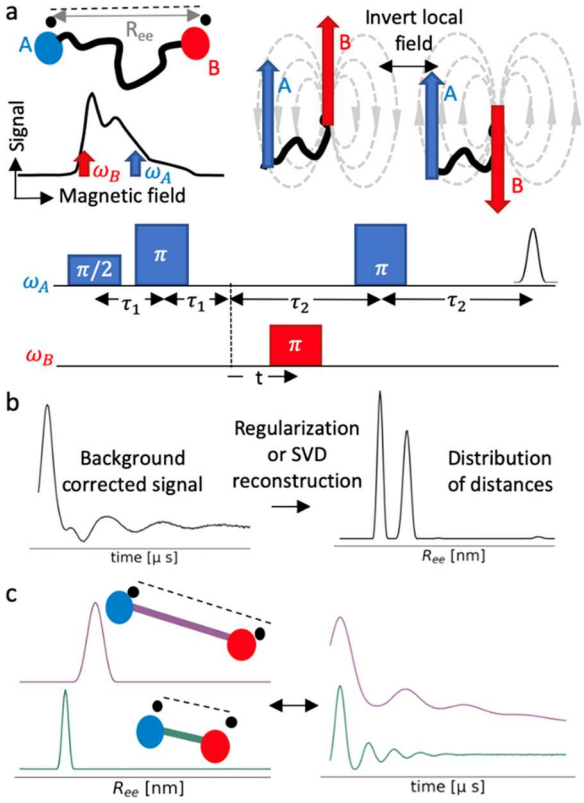


Figure 2. DEER enables the resolution of distributions of distances. (a) A pump pulse at frequency ω_B inverts the population of spins at that frequency (represented here by spin B), consequently inverting the local field of spin B on spins at the observer frequency, ω_A (spin A, in this case). Interaction between the ω_A and ω_B spin populations changes the frequency of spin A, observed in the echo intensity measured at ω_A as a function of the time corresponding to the pump pulse position. In the pulse schematic shown, π pulses fully invert spins, while $\pi/2$ pulses tip the spins only 90°. The times between observer frequency pulses are denoted as τ_1 and τ_2. (b) Inversion of this time-domain signal yields a distribution of distances in the form of a $P(R_{ee})$. (c) Distance between the spin labels determines the frequency of the oscillations in the time-domain data with shorter distances resulting in faster oscillations. The observed time-domain signal is a product of the frequencies corresponding to each distance in the ensemble.

decomposition (SVD) that have been developed by Srivastava and Freed (designated as SF-SVD). These methods involve factorization of the time-domain signal into discrete distance regions, such that $P(r)$ is constructed from a linear combination of individual "singular-valued" components.[43,44] This approach has been shown to be robust for analyzing broad and multimodal distributions[61,62] and hence is the primary approach used in this paper. Each distance distribution is obtained by denoising the data using WavPDS[63] and then fitting the denoised data with the software SVDReconstruction.[43] Both software packages are available via the National Biomedical Research Center for AdvanCed ESR Technology hosted by the Cornell Center for Advanced Computing. We present $P(R_{ee})$ found with SF-SVD in the main text and present those found via Tikhonov regularization using the LongDistances software package[64] in Section S2.

The lower limit of the DEER technique is approximately 20 Å,[30,48] as discussed in Section S9.[30,48,65,66] For this reason, DEER-derived distributions are presented with the region below 20 Å in shaded blue.

2.3. Measuring Distances below the DEER-Resolvable Region with cw-EPR. To complement DEER in the short-distance regime, cw-EPR spectroscopy probes distances between 8 and 25 Å.[67] We acquire the cw-EPR spectra on a Bruker EMXplus X-band EPR spectrometer equipped with a Bruker ER4119HS-W1 resonator, as

https://doi.org/10.1021/acs.macromol.1c00550
Macromolecules 2021, 54, 5011−5021

described in Section S4. This technique extracts end-to-end distances from the effects of dipolar coupling between nearby unpaired electrons. When the extent of EPR line broadening by the dipole−dipole interactions exceeds the intrinsic cw-EPR line width of the spin label under the given experimental condition, its effect results in spectral broadening that can be related to dipolar distances through spectral simulation.[68] Line broadening is observed for doubly labeled polypeptoids where dipolar coupling originates from distances typically below 18−25 Å.[48,67] Quantitative analyses leading to distance distributions are performed using ShortDistances spectral simulation software.[69,70] As Figure 3 illustrates, deconvolution analysis

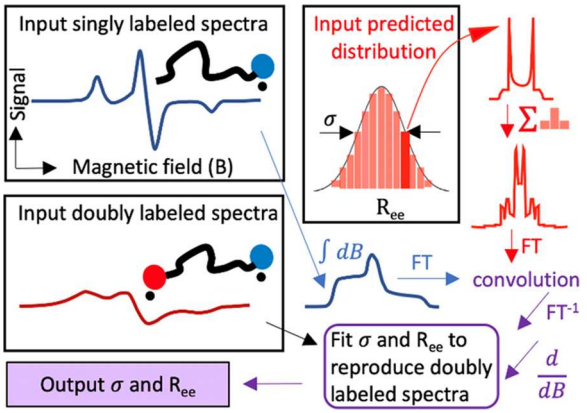


Figure 3. cw-EPR probes distances less than 25 Å. Short distances inaccessible with DEER can be determined from dipolar broadening accessed with cw-EPR. Deconvolution analysis iteratively fits a Gaussian approximation of the distance distribution, $P(R_{ee})$, parameterized by average distance, $\langle R_{ee}\rangle$, and width, σ, to the EPR spectrum of a doubly labeled sample by convoluting the EPR spectrum of a singly labeled sample with the dipolar broadening from the distance distribution.

starts with generating Gaussian distance distributions parameterized by an average distance, $\langle R_{ee}\rangle$, and width, σ (full width at half-maximum). The distributions are convoluted with EPR spectra of singly labeled samples (representing spectra unaffected by dipolar broadening) and compared to the experimental EPR spectra for doubly labeled samples. The parameters, $\langle R_{ee}\rangle$ and σ, are iteratively varied until the parameters best reproduce the broadened EPR spectrum and then are used to calculate the $P(R_{ee})$. We determine the 95% confidence intervals for the resulting distance distribution from the covariance matrix of the fit parameters (Section S4).

2.4. Polypeptoid Forcefield Parameterization and Simulation. In this study, we primarily utilize the MFTOID[53] forcefield to simulate polypeptoids but also compare results with other peptoid models: GAFF-φ,[54] a modified AMBER99SB-ILDN forcefield,[56] and a modified CGenFF forcefield.[57] Section S8 gives details of the forcefield parameterization. With any given forcefield, we create and geometry-minimize initial polypeptoid structures in the fully extended conformation using Avogadro 1.2.0.[71] We place each polypeptoid in a box of length at least 20 Å longer than its end-to-end distance when fully extended (Table S2) and solvate the boxes using GROMACS 2016.1.[72,73] We perform simulations using the OpenMM simulation engine.[74] Electrostatic interactions are computed using particle-mesh Ewald. Lennard-Jones nonbonded interactions are cut off at 10 Å. Hydrogen bonds are constrained with SHAKE,[75] and water is kept rigid with SETTLE.[76] We simulate Langevin dynamics with an integration time step of 2 fs, a temperature of 300 K, and a friction coefficient of 0.01 ps^{-1}. The pressure is set to 1 bar using a Monte Carlo (MC) barostat with 500 fs between MC barostat moves.

2.5. Expanded Ensemble Sampling. To sample across slow polypeptoid isomerization time scales, we develop a new sampling approach based on an expanded ensemble,[77,78] allowing the system to visit states in which the dihedral potentials for the ω dihedral (Figure 1b) are "softened", i.e., the force constant is scaled by a factor, λ_{ee}, that varies between 0 and 1. When simulating with the modified CGenFF forcefield, we multiply all backbone dihedral potentials (φ and ψ in Figure 1b, in addition to ω) by λ_{ee} since isomerization of the φ dihedrals is also slow (Section S12). To simulate in the expanded ensemble, we perform Monte Carlo (MC) moves to change the λ_{ee}-state every 20 ps. For each MC move, we use metropolized independence sampling[79] to compute the proposal and acceptance probabilities for each λ_{ee}-state. These probabilities are biased by weights for each state that are modified in the equilibration period to achieve a flat sampling of states and then fixed in the production period. Based on the probabilities, a new λ_{ee}-state is proposed and accepted or rejected. For each system, there are a discrete set of λ_{ee}-states, distributed so that the self-transition probabilities are nearly flat. We reweight the properties computed during the expanded ensemble production period back to the original, unsoftened state using MBAR.[80,81]

2.6. Comparison of Polypeptoid Forcefields. Simulation forcefield models for polypeptoids have seen far less development and refinement than their peptide counterparts. Several polypeptoid forcefields have been developed starting from general or protein forcefields and reparameterizing to fit either experimental or quantum mechanical measurements of short polypeptoids. Previous validation of forcefields has shown that they accurately reproduce the configurations of very short[82] or highly structured polypeptoids.[54,83] Many studies, for instance, compare simulations with crystal structures,[53,55,82,84] so that only sequences that show a single dominant conformation can be evaluated. However, to date, there has been little work testing whether these polypeptoid forcefields accurately reproduce the broader and multifeatured conformational distributions of longer and disordered systems.

Figure 4 shows the end-to-end distance distribution of the hydrophilic polypeptoid of length N = 11 measured by DEER (top)

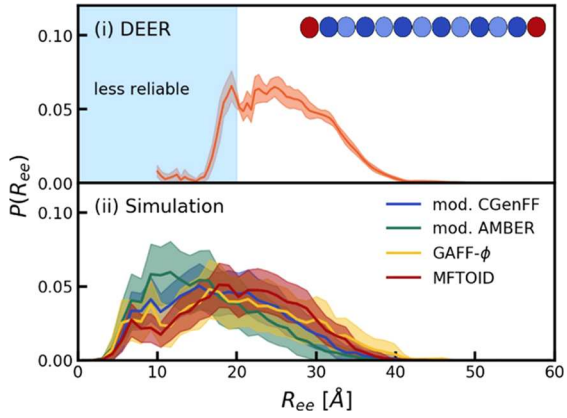


Figure 4. Experiments validate models for disordered polypeptoids. For the N = 11 polypeptoid, (top) the end-to-end distribution measured using DEER and fit with SF-SVD. The shaded region around the distribution gives 95% confidence intervals, computed as described in Section S2. The shaded blue box denotes the region below the accessible region for DEER. (Bottom) End-to-end distributions computed from simulations using various polypeptoid forcefields. Shaded regions show a 68% confidence interval computed using bootstrapping (Section S13).

and computed from simulations using various forcefields (bottom). In the accessible region, the MFTOID and GAFF-φ forcefields show better agreement with the experiment than the modified AMBER and modified CGenFF forcefields. Moving beyond visual assessment, to quantify the agreement between the experimentally (P_{DEER}) and computationally (P_{sim}) derived end-to-end distance distributions, we compute the Jensen−Shannon distance (JS).[85] Smaller distances reflect greater overlap between two probability distributions and thus better quantitative agreement between the experiment and simulation

https://doi.org/10.1021/acs.macromol.1c00550
Macromolecules 2021, 54, 5011−5021

$$\mathrm{JS}(P_{\mathrm{DEER}}, P_{\mathrm{Sim}}) = \sqrt{\frac{1}{2}\left[S_{\mathrm{rel}}\left(P_{\mathrm{DEER}}, \frac{P_{\mathrm{sim}} + P_{\mathrm{DEER}}}{2}\right) + S_{\mathrm{rel}}\left(P_{\mathrm{sim}}, \frac{P_{\mathrm{sim}} + P_{\mathrm{DEER}}}{2}\right)\right]} \quad (1)$$

Here, S_{rel} is the relative entropy

$$S_{\mathrm{rel}}(P_1, P_2) = \int P_1(R_{\mathrm{ee}}) \ln \frac{P_1(R_{\mathrm{ee}})}{P_2(R_{\mathrm{ee}})}\, \mathrm{d}R_{\mathrm{ee}} \quad (2)$$

We limit computations to the DEER-accessible region (>20 Å) for both distributions. Distances between the experimental distribution and simulated distributions using the MFTOID, GAFF-φ, modified AMBER, and modified CGenFF forcefields are 0.16 ± 0.06, 0.19 ± 0.07, 0.30 ± 0.09, and 0.22 ± 0.07, respectively (ranges give a 95% confidence interval). An alternative metric gives similar trends (Table S22). The modified AMBER and modified CGenFF forcefields therefore show the weakest agreement with the experimental distribution, while MFTOID and GAFF-φ seem comparable in accuracy. The latter, however, enforces an asymmetry in the φ dihedral angle distribution that is not physical for the achiral polypeptoids studied here, and thus the MFTOID forcefield is the model we adopt for the remainder of this work.

We attribute these differences to the polypeptoid forcefield parameters themselves and not the different water models used, as Figure S31 shows that the $P(R_{\mathrm{ee}})$ for N = 11 with the MFTOID forcefield is consistent for the mTIP3P and SPC/E water models. Below the DEER-accessible region (<20 Å), the forcefields also show differences in the height of the low-distance peak, suggesting that the strength of spin-label aggregation varies among the forcefields. These differences could be due to the polypeptoid forcefield or the spin-label forcefield, but the latter is less likely because forcefields implemented with the same spin-label parameters (MFTOID and modified CGenFF) still give different $P(R_{\mathrm{ee}})$'s in this region. Furthermore, the backbone end-to-end distance distributions for the labeled and unlabeled N = 11 polypeptoids, using the MFTOID forcefield, are consistent (Figure S34), suggesting that the presence of the spin label does not significantly perturb the polypeptoid backbone conformational distribution. Thus, small deviations in the spin-label parameters, at least for the MFTOID forcefield, should not have a significant effect on the overall computed conformational landscape of the polypeptoid, especially for longer lengths.

3. RESULTS AND DISCUSSION

3.1. Expanded Ensemble Simulation Accurately and Efficiently Samples Isomerization States. The expanded ensemble simulation method reproduces the free-energy difference between the cis and trans states computed from umbrella sampling (Figure S23) for disarcosine (Figure 1b). The 12 kcal/mol isomerization barrier also agrees with the value reported by Zhao et al., who used parallel-bias metadynamics to sample across the isomerization barrier.[58] The uncertainty in the free-energy landscape computed using the expanded ensemble is larger at the barriers because the expanded ensemble only circumvents the high free-energy barrier for isomerization through modulation of the dihedral potential, so that the sampling of the high free-energy regions may still be poor. On the other hand, approaches such as umbrella sampling and metadynamics directly bias the dihedrals, encouraging flat sampling of the space of dihedral angles. However, the equilibrium conformational landscape will be dominated by configurations within the cis or trans basins, not at the high free-energy barriers. The expanded ensemble technique, then, allows crossing of high free-energy barriers without spending unnecessary time thoroughly sampling these regions.

As previously reported for small molecules,[52] the expanded ensemble significantly enhances the efficiency of the simulation. Scaling the dihedral potential by a factor λ_{ee} provides an order-of-magnitude increase in the isomerization rate, comparing the rate of transitions between cis and trans states when the ω dihedral potential is completely removed (λ_{ee} = 0) to when it is unsoftened (λ_{ee} = 1) (Figure S20); extremely high temperatures are required to achieve the same isomerization rates (Figure S21), rendering temperature replica exchange impractical. The ability to access states at or close to λ_{ee} = 0 then makes the correlation time for the ω dihedrals tractable (Figure S22), even accounting for the additional computational cost of simulating the system in λ_{ee}-states that are far from the unbiased state (Figure S28). As a result, equilibration and production runs on the order of microseconds are sufficient to sample the end-to-end distance distribution of these polypeptoids (Section S11). This means that achieving an unbiased sampling of conformational space, assuming that all other degrees of freedom are fully sampled over, is now computationally tractable. The following results from MD simulation all rely on this sampling approach.

3.2. End-to-End Distance Distributions from Experiments and Simulation. The computational and experimental workflows presented here enable detailed characterization of the conformational distribution of polypeptoids with increasing chain lengths. We synthesize the polypeptoid with alternating hydrophilic methoxyethyl (NME) and propanol groups (Figure 1a), varying the number of hydrophilic monomers: N = 9, 11, 15, and 19. We attach spin labels to both ends of all polypeptoids for DEER measurements. We carry out 4-pulse DEER experiments and obtain the end-to-end distance distributions, $P(R_{\mathrm{ee}})$, through SF-SVD, as described in Section 2. Unless otherwise noted, we use "end-to-end distance" to refer to the distance between the oxygen atoms on the spin labels. We simulate the same polypeptoids (including spin labels) as well as other lengths (N = 1, 2, 3, 5, 7, 13) using the expanded ensemble sampling approach and the MFTOID forcefield (see Section 2.6 for justification).

Figure 5 shows the end-to-end distance distributions for lengths N = 9, 11, 15, and 19. In the accessible region, simulations and DEER show excellent agreement, which offers strong cross-validation of both workflows. The locations of the peaks and the shapes of the long-distance tails of the distributions are especially well reproduced. Both methods consistently show not only the expected shift in the mean end-to-end distance but also a broadening of the distribution with length, characteristic of disordered polymers. Below the DEER-accessible region (<20 Å), the simulated distributions show a smaller peak at around 10 Å that decreases in height with increasing polypeptoid length. This peak suggests that the spin labels on both ends have a slight tendency to aggregate with each other (Figure S39), as was previously observed in simulations of end-labeled polyethylene oxide.[39] Recalculating the $P(R_{\mathrm{ee}})$ by removing the contribution from these short-distance conformations, which are not captured by DEER, further improves the agreement between DEER and simulation (Figures S40 and S41).

In the DEER-accessible region (>20 Å), Figure 5 demonstrates a few small inconsistencies between DEER and simulation. To determine whether these inconsistencies are reproduced in the time domain or stem from instabilities in the fitted distance distribution due to the ill-posed nature of the inversion, we back-calculate the DEER time-domain signal,

https://doi.org/10.1021/acs.macromol.1c00550
Macromolecules 2021, 54, 5011−5021

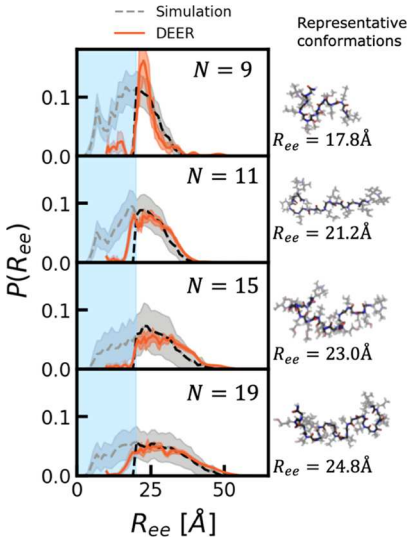


Figure 5. End-to-end distance distributions measured with DEER and computed from simulation are in excellent agreement. The gray lines are the simulated distributions (using the MFTOID forcefield), and the gray-shaded regions give a 68% confidence interval computed from bootstrapping (Section S13). The orange lines and shaded regions are the fitted DEER distributions and uncertainty computed, respectively, as described in the main text and Section S2. All distributions are normalized such that the integral above 20 Å is 1. The light blue-shaded box denotes the region below the accessible region for DEER. The snapshots to the right of the figures are representative conformations chosen to be close to the mean end-to-end distances and are visualized in VMD.[86]

$F(t)$, from the MD-simulated and SF-SVD fitted distance distributions, performing the forward convolution that the fitting procedures attempt to invert (Section S9).

Figure 6 shows the time-domain signals computed through convolution (Section S9) using the distributions in Figure 5 ("back-calculated MD" and "back-calculated SF-SVD" signals), along with the background-corrected experimental time-domain data from DEER ("raw" signal). The shifts in the back-calculated MD signal with increasing polypeptoid lengths

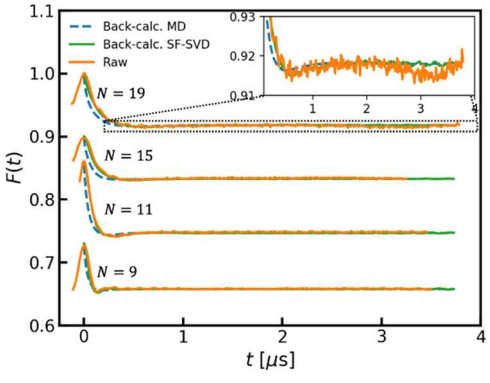


Figure 6. Raw DEER signals are consistent with signals back-calculated from simulated and fit end-to-end distance distributions. Orange lines show the raw DEER signals, while the blue and green lines show the signals back-calculated from the MD-simulated $P(R_{ee})$ (simulated using the MFTOID forcefield) and the SF-SVD-fit $P(R_{ee})$, respectively. For all signals, $V(0) = 1$, but the plots are shifted vertically (Table S21) for each polypeptoid length for clarity. (Inset) Low-frequency oscillations in the raw signal for the $N = 19$ polypeptoid.

are qualitatively consistent with that of the raw signal. In particular, the slower decay and broadening of the minimum at larger lengths are well reproduced. For all lengths, the back-calculated MD signal shows a steeper decay near $t = 0$. This feature reflects the presence of shorter end-to-end distances in the MD-simulated $P(R_{ee})$ that are not present in the DEER-derived $P(R_{ee})$. As discussed earlier, these short distances are inaccessible to DEER. Recomputing the back-calculated MD signal using only the portion of the MD-simulated $P(R_{ee})$ above the lower accessible limit of DEER (20 Å) leads to a better agreement in the short-time decay (Figure S19). For $N = 15$ and 19, the raw signal displays a low-frequency oscillation (Figure 6, inset) that arises from longer distances in the DEER-derived $P(R_{ee})$ and that is not reproduced in the back-calculated MD signal shown in Figure 6. Thus, the differences between the simulation and DEER end-to-end distance distributions at longer lengths in Figure 5 are real and not artifacts of the inversion. This discrepancy at longer lengths could be due to inaccuracies in the forcefield that lead to oversampling of shorter distances. For intrinsically disordered polypeptides, it was previously reported that peptide forcefields bias conformational distributions toward more compact structures than is realistic.[87,88]

The full distance distributions in Figure 5 enhance our understanding of shifts in the conformational landscape with additional monomers. Figure S35 compares simulated distance distributions to those for a theoretical excluded volume polymer, suggesting that the polypeptoids follow excluded volume scaling, characteristic of polymers in a good solvent.[89] To confirm, Figure 7 shows that the mean distances computed from both the simulated and experimental distance distributions follow excluded volume scaling. The mean end-to-end distances along the backbone, between terminal nitrogens instead of the spin-label oxygen radicals (see Figure 7, inset), computed from simulations and fit to $\langle R^2_{ee,NN}\rangle^{1/2} = bN^{\alpha}$, give a scaling exponent and prefactor of $\alpha = 0.61 \pm 0.02$ and $b = 4.0$

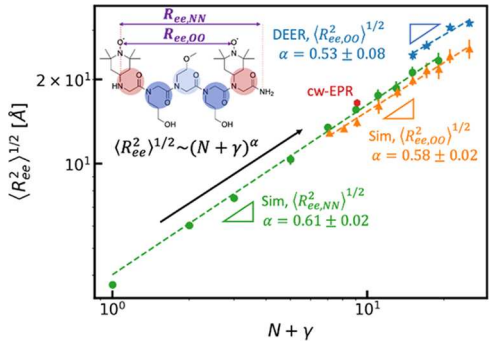


Figure 7. Scaling of the mean end-to-end distance shows reasonable agreement between experiments and simulations and suggests excluded volume scaling. The mean end-to-end distance computed from the DEER distributions (blue stars) and from simulation (orange triangles give the mean distance between the oxygen radicals, and green circles give the mean distance along the backbone) using the MFTOID forcefield. The red hexagon shows the mean end-to-end distance from cw-EPR. Dashed lines show fits to a power law, shifted by an effective number of additional monomers γ for the spin-label end-to-end distances, as described in the text. The labels report the fitted exponents. (Inset) Spin-label end-to-end distance, $R_{ee,OO}$, is measured by DEER and cw-EPR and is computed from simulations, while the backbone end-to-end distance, $R_{ee,NN}$, is only accessible in simulations.

https://doi.org/10.1021/acs.macromol.1c00550
Macromolecules 2021, 54, 5011−5021

± 0.2, respectively. The value of the exponent is consistent with excluded volume scaling (α = 0.588) and reasonably consistent with Zhao et al. who reported a scaling exponent of 0.66 ± 0.01 for the radius of gyration of polysarcosine,[58] another relatively hydrophilic polypeptoid, computed from molecular dynamics simulations. As far as we know, no experimental works have previously reported scaling exponents for polypeptoids.

However, the mean distance that is computed from the simulated and DEER distributions in Figure 5 is not the backbone distance but the spin-label end-to-end distance, which incurs an additional contribution (relative to the backbone) from the spin labels themselves. To account for this effect, we fit mean spin-label distances to a shifted power law, $\langle R^2_{ee,OO}\rangle^{1/2} = b(N+\gamma)^{\alpha}$, where b = 4.0 is the prefactor fit from the backbone end-to-end distances and γ accounts for the effective increase in the number of monomers due to the addition of the spin labels and the distance between the backbone and the oxygen radical on the spin label. Previous work has used a similar offset parameter to fit mean end-to-end distances from FRET.[90,91] The fitted exponent for the simulated mean spin-label distances is α = 0.58 ± 0.02, corroborating excluded volume scaling. The fitted effective increase in the number of monomers is γ = 6.0, suggesting that each spin label effectively behaves as three additional monomers. We then fit the DEER-derived mean spin-label distances using this shift factor, i.e., to $\langle R^2_{ee,OO}\rangle^{1/2} = b(N+6.0)^{\alpha}$. The resulting exponent is α = 0.53 ± 0.08 and again agrees with the excluded volume scaling within uncertainty. Thus, the scaling exponents computed from simulations and DEER are consistent. However, the mean end-to-end distances from simulations are systematically smaller, consistent with the previous work comparing mean distances from DEER and simulation for poly(ethylene oxide).[39] This is partly due to the challenges in resolving the region below 20 Å for DEER. Computing the mean of the portion above the lower accessible limit for DEER (20 Å) for both the DEER-resolved and simulated distributions reduces the discrepancy (Figure S41).

We next use cw-EPR to probe the distance distribution of the N = 3 polypeptoid, whose end-to-end distance distribution largely lies below 20 Å and thus below the accessible region of DEER. Figure 8 shows that the Gaussian approximation of the distance distribution from the cw-EPR experiment agrees with the longer-distance peak in the simulated distribution. The simulated distribution also contains a shorter-distance peak at around 8 Å that is also seen in Figure 5 for longer polypeptoids and here similarly indicates spin-label aggregation. The mean end-to-end distance from cw-EPR line-shape analysis is 16.5 ± 0.1 Å. Fitting to the power law, $\langle R^2_{ee,OO}\rangle^{1/2} = b(N+6.0)^{\alpha}$, using both the DEER and cw-EPR mean distances yields a scaling exponent of α = 0.63 ± 0.06 and hence remains consistent with the excluded volume scaling seen in Figure 7. Cw-EPR cannot resolve distances below ∼8 Å, so a fraction of the low-distance population predicted from the simulations would not contribute to the cw-EPR-resolved distribution. However, cw-EPR clearly extends our insight into the experimental conformational distributions at distances below the DEER-accessible regime.

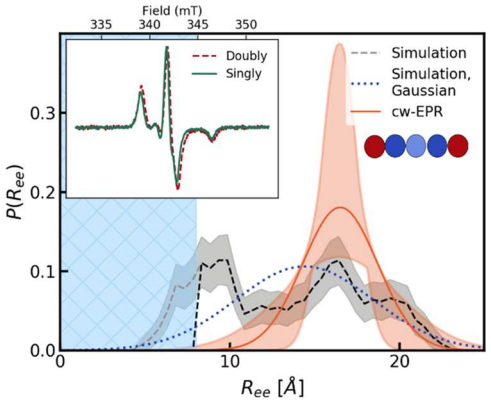


Figure 8. Continuous wave electron paramagnetic resonance allows access to the low-distance region and agrees with simulations. For the N = 3 polypeptoid, the Gaussian approximation of the end-to-end distance distribution from cw-EPR (orange) agrees with the simulated distribution (using the MFTOID forcefield) in the accessible region for cw-EPR (black dashed line gives the full distribution; the blue line is a Gaussian with the same mean and standard deviation as the simulated distribution in the cw-EPR-accessible region). Shaded gray regions show a 68% confidence interval computed using bootstrapping (Section S13). The shaded orange region encompasses all Gaussian distributions with mean and standard deviation within the 95% confidence intervals obtained from the covariance matrix of $\langle R_{ee}\rangle$ and σ (Section S4). All distributions are normalized such that the integral above 8 Å is 1. The shaded and hatched blue box denotes the region below the accessible region for cw-EPR. (Inset) Raw cw-EPR spectra for the doubly labeled 3mer (dashed red) and singly labeled 3mer (solid green) normalized by the maximum peak height.

4. CONCLUSIONS

Using a powerful experimental technique, DEER, and a state-of-the-art enhanced sampling approach to overcome long isomerization time scales in polypeptoid simulations, we develop a combined experimental−simulation approach to quantify the conformational landscapes of disordered hydrophilic polypeptoids through the measurement of end-to-end distance distributions. The large widths of the distributions indicate that these hydrophilic polypeptoids are indeed disordered and would thus be difficult to structurally characterize solely by examination of dominant conformations, instead requiring the resolution of full distributions. Validation of the expanded ensemble sampling technique and the MFTOID forcefield allows us to computationally probe the conformational distribution of polypeptoids of appreciable lengths (here, tested up to N = 19, while we estimate up to N ≈ 30 is reasonable for a single GPU). We demonstrate excellent agreement between distance distributions resolved through simulations and dipolar EPR spectroscopy for varying lengths of a model hydrophilic polypeptoid. These capabilities offer fundamental insight into the excluded volume polymer scaling behavior of this particular polypeptoid system.

These results establish highly transferable experimental and computational workflows to characterize disordered polypeptoid conformations with emerging structural characteristics beyond excluded volume properties. Though polypeptoid sequences have already been shown to modulate structure and function, the workflows developed here will enable more controlled investigations of sequence−structure−property relationships in the highly tunable, sequence-programmable polypeptoid platform. For instance, the determination of changes in the conformational ensemble allows us to isolate the effects of backbone conformation and side-chain chemistries on functional properties. This is especially important because more significant variations in sequence, such as incorporating an irregular sequence of hydrophobic

https://doi.org/10.1021/acs.macromol.1c00550
Macromolecules 2021, 54, 5011−5021

side chains, or movement of the polypeptoid from the solution to a surface will lead to significant perturbations in structure that can complicate the sequence−property relationship and cannot be fully captured by existing polymer physics models. Beyond intrachain distances, the ability to resolve intermolecular distance distributions, if spin labels are placed on separate molecules, can provide insight into self-assembled polypeptoid structures or even the mechanism of self-assembly. For these more complex systems, the ability to evaluate full distance distributions both experimentally and computationally is critical. The versatility of the experimental technique, as the spin labels can be placed in other locations along the polypeptoids, provides an exceptional degree of insight into broad and multifeatured conformational landscapes. These capabilities will further contribute to the design of novel materials that utilize disordered polypeptoids.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.macromol.1c00550.

Details regarding polypeptoid synthesis; DEER and cw-EPR characterizations; simulation sizes and lengths; forcefield parameterizations; back-calculation of time-domain signals; computational sampling, convergence, and uncertainty estimates; measures of distances between distributions; impact of the water model; polypeptoid amide-bond dihedral preferences; computational comparison of labeled and unlabeled polypeptoids; comparison to theoretical end-to-end distance distributions; scaling of the mean end-to-end distance from LongDistances and from the modified CGenFF forcefields; scaling of the radius of gyration from the MFTOID and modified CGenFF forcefields; removal of the short-distance spin-label-aggregated states from the simulated distance distributions; dependence of the predicted simulated distribution on the cw-EPR bounds; visualization of the conformational distribution; and conformational trends with increasing polypeptoid lengths (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Songi Han − *Department of Chemical Engineering, University of California, Santa Barbara, California 93106, United States; Department of Chemistry and Biochemistry, University of California, Santa Barbara, California 93106, United States;* orcid.org/0000-0001-6489-6246; Email: songi@chem.ucsb.edu

M. Scott Shell − *Department of Chemical Engineering, University of California, Santa Barbara, California 93106, United States;* orcid.org/0000-0002-0439-1534; Email: shell@ucsb.edu

Authors

Sally Jiao − *Department of Chemical Engineering, University of California, Santa Barbara, California 93106, United States;* orcid.org/0000-0003-1729-0414

Audra DeStefano − *Department of Chemical Engineering, University of California, Santa Barbara, California 93106, United States*

Jacob I. Monroe − *Department of Chemical Engineering, University of California, Santa Barbara, California 93106, United States*

Mikayla Barry − *Department of Materials, University of California, Santa Barbara, California 93106, United States*

Nicholas Sherck − *Department of Chemical Engineering, University of California, Santa Barbara, California 93106, United States;* orcid.org/0000-0001-5503-1494

Thomas Casey − *Department of Chemistry and Biochemistry, University of California, Santa Barbara, California 93106, United States*

Rachel A. Segalman − *Department of Chemical Engineering, University of California, Santa Barbara, California 93106, United States; Department of Materials, University of California, Santa Barbara, California 93106, United States;* orcid.org/0000-0002-4292-5103

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.macromol.1c00550

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the Center for Materials for Water and Energy Systems (M-WET), an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award #DE-SC0019272. Use was made of computational facilities purchased with funds from the National Science Foundation (OAC-1925717) and administered by the Center for Scientific Computing (CSC). The CSC is supported by the California NanoSystems Institute and the Materials Research Science and Engineering Center (MRSEC; NSF DMR 1720256) at UC Santa Barbara. S.J., J.I.M., and M.B. acknowledge support by the National Science Foundation Graduate Research Fellowship (DGE 1650114, DGE 1144085). M.B. also recognizes support from the Office of Naval Research (ONR) under Award #N00014-17-1-2047. A.D. acknowledges support by the National Defense Science & Engineering Graduate (NDSEG) Fellowship Program. The authors gratefully acknowledge access to the WavPDS and SVDReconstruction software through the National Biomedical Center for Advanced ESR Technology (ACERT). ACERT is hosted by the Cornell Center for Advanced Computing and supported by the National Institute of General Medical Sciences. Conversations with M. Srivastava regarding the use of WavPDS and SVDReconstruction were particularly helpful. The authors acknowledge the use of MRL Shared Experimental Facilities, which are supported by the MRSEC Program of the NSF under Award No. DMR 1720256, a member of the NSF-funded Materials Research Facilities Network (www.mrfn.org). The authors also acknowledge helpful discussions regarding DEER acquisition and analysis with T. Keller and K. Tsay and conversations with R. Zuckermann regarding polypeptoid synthesis. The authors thank A. Prakash and J. Pfaendtner for providing modified AMBER99SB-ILDN forcefield files for peptoids and helpful discussions concerning their implementation and R. Jain and E. Santiso for helpful discussions concerning the implementation of the modified CGenFF forcefield.

REFERENCES

(1) Knight, A. S.; Zhou, E. Y.; Francis, M. B.; Zuckermann, R. N. Sequence Programmable Peptoid Polymers for Diverse Materials Applications. *Adv. Mater.* **2015**, *27*, 5665−5691.

(2) Saxena, V.; Merrilees, M. G. L.; Lau, K. H. A. Antifouling Peptoid Biointerfaces. In *Biointerface Engineering: Prospects in Medical Diagnostics and Drug Delivery*, Chandra, P.; Pandey, L. M., Eds.; Springer: Singapore, 2020; pp 55−73.

(3) Green, R. M.; Bicker, K. L. Evaluation of Peptoid Mimics of Short, Lipophilic Peptide Antimicrobials. *Int. J. Antimicrob. Agents* **2020**, *56*, No. 106048.

(4) Khara, J. S.; Mojsoska, B.; Mukherjee, D.; Langford, P. R.; Robertson, B. D.; Jenssen, H.; Ee, P. L. R.; Newton, S. M. Ultra-Short Antimicrobial Peptoids Show Propensity for Membrane Activity Against Multi-Drug Resistant Mycobacterium Tuberculosis. *Front. Microbiol.* **2020**, *11*, No. 417.

(5) Kim, J. H.; Kim, S. C.; Kline, M. A.; Grzincic, E. M.; Tresca, B. W.; Cardiel, J.; Karbaschi, M.; Dehigaspitiya, D. C.; Chen, Y.; Udumula, V.; Jian, T.; Murray, D. J.; Yun, L.; Connolly, M. D.; Liu, J.; Ren, G.; Chen, C.-L.; Kirshenbaum, K.; Abate, A. R.; Zuckermann, R. N. Discovery of Stable and Selective Antibody Mimetics from Combinatorial Libraries of Polyvalent, Loop-Functionalized Peptoid Nanosheets. *ACS Nano* **2020**, *14*, 185−195.

(6) Mackwitz, M. K. W.; Hesping, E.; Eribez, K.; Schöler, A.; Antonova-Koch, Y.; Held, J.; Winzeler, E. A.; Andrews, K. T.; Hansen, F. K. Investigation of the in Vitro and in Vivo Efficacy of Peptoid-Based HDAC Inhibitors with Dual-Stage Antiplasmodial Activity. *Eur. J. Med. Chem.* **2021**, *211*, No. 113065.

(7) Asfaw, K. G.; Liu, Q.; Xu, X.; Manz, C.; Purper, S.; Eghbalian, R.; Münch, S. W.; Wehl, I.; Bräse, S.; Eiche, E.; Hause, B.; Bogeski, I.; Schepers, U.; Riemann, M.; Nick, P. A Mitochondria-Targeted Coenzyme Q Peptoid Induces Superoxide Dismutase and Alleviates Salinity Stress in Plant Cells. *Sci. Rep.* **2020**, *10*, No. 11563.

(8) Mukudai, S.; Kraja, I.; Bing, R.; Nalband, D. M.; Tatikola, M.; Hiwatashi, N.; Kirshenbaum, K.; Branski, R. C. Implementing Efficient Peptoid-Mediated Delivery of RNA-Based Therapeutics to the Vocal Folds. *Laryngoscope Investigative Otolaryngology* **2019**, *4*, 640−644.

(9) Wang, M.; Song, Y.; Mu, P.; Cai, X.; Lin, Y.; Chen, C.-L. Peptoid-Based Programmable 2D Nanomaterial Sensor for Selective and Sensitive Detection of H_2S in Live Cells. *ACS Appl. Bio Mater.* **2020**, *3*, 6039−6048.

(10) Li, Y.; Tom, J. C.; Biehl, P.; Ling, J.; Schacher, F. H. Block Polypeptoids: Synthesis, Characterization, and Response Toward Irradiation with UV Light and Temperature. *Macromolecules* **2020**, *53*, 5218−5226.

(11) Zhang, Q.; Zhang, C.; Luo, W.; Cui, L.; Wang, Y.; Jian, T.; Li, X.; Yan, Q.; Liu, H.; Ouyang, C.; Chen, Y.; Chen, C.; Zhang, J. Sequence-Defined Peptoids with -OH and -COOH Groups As Binders to Reduce Cracks of Si Nanoparticles of Lithium-Ion Batteries. *Adv. Sci.* **2020**, *7*, No. 2000749.

(12) Chen, C.-L.; Qi, J.; Zuckermann, R. N.; DeYoreo, J. J. Engineered Biomimetic Polymers as Tunable Agents for Controlling $CaCO_3$ Mineralization. *J. Am. Chem. Soc.* **2011**, *133*, 5214−5217.

(13) Zhang, S.; Chen, J.; Liu, J.; Pyles, H.; Baker, D.; Chen, C.; De Yoreo, J. J. Engineering Biomolecular Self-Assembly at Solid−Liquid Interfaces. *Adv. Mater.* **2020**, 1905784. DOI: 10.1002/adma.201905784.

(14) Xuan, S.; Zuckermann, R. N. Engineering the Atomic Structure of Sequence-Defined Peptoid Polymers and Their Assemblies. *Polymer* **2020**, *202*, No. 122691.

(15) Cai, B.; Li, Z.; Chen, C.-L. Programming Amphiphilic Peptoid Oligomers for Hierarchical Assembly and Inorganic Crystallization. *Acc. Chem. Res.* **2021**, *54*, 81−91.

(16) Shah, N. H.; Butterfoss, G. L.; Nguyen, K.; Yoo, B.; Bonneau, R.; Rabenstein, D. L.; Kirshenbaum, K. Oligo(*N*- Aryl Glycines): A New Twist on Structured Peptoids. *J. Am. Chem. Soc.* **2008**, *130*, 16622−16632.

(17) 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*, 4303−4308.

(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*, 4309−4314.

(19) 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*, No. 3673.

(20) Wijaya, A. W.; Nguyen, A. I.; Roe, L. T.; Butterfoss, G. L.; Spencer, R. K.; Li, N. K.; Zuckermann, R. N. Cooperative Intramolecular Hydrogen Bonding Strongly Enforces *Cis* -Peptoid Folding. *J. Am. Chem. Soc.* **2019**, *141*, 19436−19447.

(21) 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*, 586−595.

(22) Patterson, A. L.; Danielsen, S. P. O.; Yu, B.; Davidson, E. C.; Fredrickson, G. H.; Segalman, R. A. Sequence Effects on Block Copolymer Self-Assembly through Tuning Chain Conformation and Segregation Strength Utilizing Sequence-Defined Polypeptoids. *Macromolecules* **2019**, *52*, 1277−1286.

(23) van Zoelen, W.; Buss, H. G.; Ellebracht, N. C.; Lynd, N. A.; Fischer, D. A.; Finlay, J.; Hill, S.; Callow, M. E.; Callow, J. A.; Kramer, E. J.; Zuckermann, R. N.; Segalman, R. A. Sequence of Hydrophobic and Hydrophilic Residues in Amphiphilic Polymer Coatings Affects Surface Structure and Marine Antifouling/Fouling Release Properties. *ACS Macro Lett.* **2014**, *3*, 364−368.

(24) Lau, K. H. A.; Sileika, T. S.; Park, S. H.; Sousa, A. M. L.; Burch, P.; Szleifer, I.; Messersmith, P. B. Molecular Design of Antifouling Polymer Brushes Using Sequence-Specific Peptoids. *Adv. Mater. Interfaces* **2015**, *2*, No. 1400225.

(25) Maron, E.; Kochovski, Z.; Zuckermann, R. N.; Börner, H. G. Peptide-Assisted Design of Peptoid Sequences: One Small Step in Structure and Distinct Leaps in Functions. *ACS Macro Lett.* **2020**, *9*, 233−237.

(26) Zhu, L.; Simpson, J. M.; Xu, X.; He, H.; Zhang, D.; Yin, L. Cationic Polypeptoids with Optimized Molecular Characteristics toward Efficient Nonviral Gene Delivery. *ACS Appl. Mater. Interfaces* **2017**, *9*, 23476−23486.

(27) Huang, M. L.; Ehre, D.; Jiang, Q.; Hu, C.; Kirshenbaum, K.; Ward, M. D. Biomimetic Peptoid Oligomers as Dual-Action Antifreeze Agents. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 19922−19927.

(28) van Zoelen, W.; Zuckermann, R. N.; Segalman, R. A. Tunable Surface Properties from Sequence-Specific Polypeptoid−Polystyrene Block Copolymer Thin Films. *Macromolecules* **2012**, *45*, 7072−7082.

(29) Leng, C.; Buss, H. G.; Segalman, R. A.; Chen, Z. Surface Structure and Hydration of Sequence-Specific Amphiphilic Polypeptoids for Antifouling/Fouling Release Applications. *Langmuir* **2015**, *31*, 9306−9311.

(30) Jeschke, G. DEER Distance Measurements on Proteins. *Annu. Rev. Phys. Chem.* **2012**, *63*, 419−446.

(31) Chiang, Y.-W.; Borbat, P. P.; Freed, J. H. The Determination of Pair Distance Distributions by Pulsed ESR Using Tikhonov Regularization. *J. Magn. Reson.* **2005**, *172*, 279−295.

(32) Sezer, D.; Freed, J. H.; Roux, B. Parametrization, Molecular Dynamics Simulation, and Calculation of Electron Spin Resonance Spectra of a Nitroxide Spin Label on a Polyalanine α-Helix. *J. Phys. Chem. B* **2008**, *112*, 5755−5767.

(33) Kaminker, R.; Kaminker, I.; Gutekunst, W. R.; Luo, Y.; Lee, S.; Niu, J.; Han, S.; Hawker, C. J. Tuning Conformation and Properties of Peptidomimetic Backbones through Dual N/C_α-Substitution. *Chem. Commun.* **2018**, *54*, 5237−5240.

https://doi.org/10.1021/acs.macromol.1c00550
Macromolecules 2021, 54, 5011−5021

(34) Morimoto, J.; Fukuda, Y.; Kuroda, D.; Watanabe, T.; Yoshida, F.; Asada, M.; Nakamura, T.; Senoo, A.; Nagatoishi, S.; Tsumoto, K.; Sando, S. A Peptoid with Extended Shape in Water. *J. Am. Chem. Soc.* **2019**, *141*, 14612−14623.
(35) Kotler, S. A.; Tugarinov, V.; Schmidt, T.; Ceccon, A.; Libich, D. S.; Ghirlando, R.; Schwieters, C. D.; Clore, G. M. Probing Initial Transient Oligomerization Events Facilitating Huntingtin Fibril Nucleation at Atomic Resolution by Relaxation-Based NMR. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 3562−3571.
(36) Schmidt, T.; Schwieters, C. D.; Clore, G. M. Spatial Domain Organization in the HIV-1 Reverse Transcriptase P66 Homodimer Precursor Probed by Double Electron-Electron Resonance EPR. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 17809−17816.
(37) Schmidt, T.; Jeon, J.; Okuno, Y.; Chiliveri, S. C.; Clore, G. M. Submillisecond Freezing Permits Cryoprotectant-Free EPR Double Electron−Electron Resonance Spectroscopy. *ChemPhysChem* **2020**, *21*, 1224−1229.
(38) Schmidt, T.; Louis, J. M.; Clore, G. M. Probing the Interaction between HIV-1 Protease and the Homodimeric P66/P66' Reverse Transcriptase Precursor by Double Electron-Electron Resonance EPR Spectroscopy. *ChemBioChem* **2020**, *21*, 3051−3055.
(39) Sherck, N.; Webber, T.; Brown, D. R.; Keller, T.; Barry, M.; DeStefano, A.; Jiao, S.; Segalman, R. A.; Fredrickson, G. H.; Shell, M. S.; Han, S. End-to-End Distance Probability Distributions of Dilute Poly(Ethylene Oxide) in Aqueous Solution. *J. Am. Chem. Soc.* **2020**, *142*, 19631−19641.
(40) Fábregas Ibáñez, L.; Jeschke, G.; Stoll, S. DeerLab: A Comprehensive Software Package for Analyzing Dipolar Electron Paramagnetic Resonance Spectroscopy Data. *Magn. Reson.* **2020**, *1*, 209−224.
(41) Edwards, T. H.; Stoll, S. A Bayesian Approach to Quantifying Uncertainty from Experimental Noise in DEER Spectroscopy. *J. Magn. Reson.* **2016**, *270*, 87−97.
(42) Edwards, T. H.; Stoll, S. Optimal Tikhonov Regularization for DEER Spectroscopy. *J. Magn. Reson.* **2018**, *288*, 58−68.
(43) Srivastava, M.; Freed, J. H. Singular Value Decomposition Method to Determine Distance Distributions in Pulsed Dipolar Electron Spin Resonance. *J. Phys. Chem. Lett.* **2017**, *8*, 5648−5655.
(44) Srivastava, M.; Freed, J. H. Singular Value Decomposition Method To Determine Distance Distributions in Pulsed Dipolar Electron Spin Resonance: II. Estimating Uncertainty. *J. Phys. Chem. A* **2019**, *123*, 359−370.
(45) de Souza, E. S.; Hirata, I. Y.; Juliano, L.; Ito, A. S. End-to-End Distance Distribution in Bradykinin Observed by Förster Resonance Energy Transfer. *Biochim. Biophys. Acta, Gen. Subj.* **2000**, *1474*, 251−261.
(46) 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*, 18964−18969.
(47) Riback, J. A.; Bowman, M. A.; Zmyslowski, A. M.; Plaxco, K. W.; Clark, P. L.; Sosnick, T. R. Commonly Used FRET Fluorophores Promote Collapse of an Otherwise Disordered Protein. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 8889−8894.
(48) Banham, J. E.; Baker, C. M.; Ceola, S.; Day, I. J.; Grant, G. H.; Groenen, E. J. J.; Rodgers, C. T.; Jeschke, G.; Timmel, C. R. Distance Measurements in the Borderline Region of Applicability of CW EPR and DEER: A Model Study on a Homologous Series of Spin-Labelled Peptides. *J. Magn. Reson.* **2008**, *191*, 202−218.
(49) 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*, 12042−12048.
(50) Gasparro, F. P.; Kolodny, N. H. NMR Determination of the Rotational Barrier in N,N-Dimethylacetamide. *J. Chem. Educ.* **1977**, *54*, 258.
(51) Weiser, L.J.; Santiso, E.E. Molecular Modeling Studies of Peptoid Polymers. *AIMS Mater. Sci.* **2017**, *4*, 1029−1051.
(52) Paluch, A. S.; Mobley, D. L.; Maginn, E. J. Small Molecule Solvation Free Energy: Enhanced Conformational Sampling Using Expanded Ensemble Molecular Dynamics Simulation. *J. Chem. Theory Comput.* **2011**, *7*, 2910−2918.
(53) Mirijanian, D. T.; Mannige, R. V.; Zuckermann, R. N.; Whitelam, S. Development and Use of an Atomistic CHARMM-Based Forcefield for Peptoid Simulation. *J. Comput. Chem.* **2014**, *35*, 360−370.
(54) Mukherjee, S.; Zhou, G.; Michel, C.; Voelz, V. A. Insights into Peptoid Helix Folding Cooperativity from an Improved Backbone Potential. *J. Phys. Chem. B* **2015**, *119*, 15407−15417.
(55) Butterfoss, G. L.; Yoo, B.; Jaworski, J. N.; Chorny, I.; Dill, K. A.; Zuckermann, R. N.; Bonneau, R.; Kirshenbaum, K.; Voelz, V. A. De Novo Structure Prediction and Experimental Characterization of Folded Peptoid Oligomers. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 14320−14325.
(56) Prakash, A.; Baer, M. D.; Mundy, C. J.; Pfaendtner, J. Peptoid Backbone Flexibilility Dictates Its Interaction with Water and Surfaces: A Molecular Dynamics Investigation. *Biomacromolecules* **2018**, *19*, 1006−1015.
(57) Weiser, L. J.; Santiso, E. E. A CGenFF-based Force Field for Simulations of Peptoids with Both *Cis* and *Trans* Peptide Bonds. *J. Comput. Chem.* **2019**, *40*, 1946−1956.
(58) Zhao, M.; Sampath, J.; Alamdari, S.; Shen, G.; Chen, C.-L.; Mundy, C. J.; Pfaendtner, J.; Ferguson, A. L. A MARTINI-Compatible Coarse-Grained Model for the Mesoscale Simulation of Peptoids. *J. Phys. Chem. B* **2020**, *124*, 7745−7764.
(59) Figliozzi, G. M.; Goldsmith, R.; Ng, S. C.; Banville, S. C.; Zuckermann, R. N. Synthesis of N-Substituted Glycine Peptoid Libraries. In *Methods Enzymology*; Elsevier: 1996; Vol. *267*, pp 437−447.
(60) Worswick, S. G.; Spencer, J. A.; Jeschke, G.; Kuprov, I. Deep Neural Network Processing of DEER Data. *Sci. Adv.* **2018**, *4*, No. eaat5218.
(61) Merz, G. E.; Borbat, P. P.; Muok, A. R.; Srivastava, M.; Bunck, D. N.; Freed, J. H.; Crane, B. R. Site-Specific Incorporation of a Cu^{2+} Spin Label into Proteins for Measuring Distances by Pulsed Dipolar Electron Spin Resonance Spectroscopy. *J. Phys. Chem. B* **2018**, *122*, 9443−9451.
(62) Dikiy, I.; Edupuganti, U. R.; Abzalimov, R. R.; Borbat, P. P.; Srivastava, M.; Freed, J. H.; Gardner, K. H. Insights into Histidine Kinase Activation Mechanisms from the Monomeric Blue Light Sensor EL346. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 4963−4972.
(63) Srivastava, M.; Georgieva, E. R.; Freed, J. H. A New Wavelet Denoising Method for Experimental Time-Domain Signals: Pulsed Dipolar Electron Spin Resonance. *J. Phys. Chem. A* **2017**, *121*, 2452−2465.
(64) Altenbach, C. *LongDistances*, 2019. http://www.biochemistry.ucla.edu/Faculty/Hubbell/.
(65) Schweiger, A.; Jeschke, G. *Principles of Pulse Electron Paramagnetic Resonance*; Oxford University Press: New York, 2001.
(66) Doll, A.; Qi, M.; Godt, A.; Jeschke, G. CIDME: Short Distances Measured with Long Chirp Pulses. *J. Magn. Reson.* **2016**, *273*, 73−82.
(67) Rabenstein, M. D.; Shin, Y. K. Determination of the Distance between Two Spin Labels Attached to a Macromolecule. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 8239−8243.
(68) Steinhoff, H. J.; Radzwill, N.; Thevis, W.; Lenz, V.; Brandenburg, D.; Antson, A.; Dodson, G.; Wollmer, A. Determination of Interspin Distances between Spin Labels Attached to Insulin: Comparison of Electron Paramagnetic Resonance Data with the X-Ray Structure. *Biophys. J.* **1997**, *73*, 3287−3298.
(69) Altenbach, C. *ShortDistances*, 2014. http://www.biochemistry.ucla.edu/Faculty/Hubbell/.
(70) Altenbach, C.; Oh, K.-J.; Trabanino, R. J.; Hideg, K.; Hubbell, W. L. Estimation of Inter-Residue Distances in Spin Labeled Proteins at Physiological Temperatures: Experimental Strategies and Practical Limitations. *Biochemistry* **2001**, *40*, 15471−15482.
(71) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An Advanced Semantic Chemical Editor, Visualization, and Analysis Platform. *J. Cheminform.* **2012**, *4*, No. 17.

https://doi.org/10.1021/acs.macromol.1c00550
Macromolecules 2021, 54, 5011−5021

(72) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1*−*2*, 19−25.
(73) Bondi, A. Van Der Waals Volumes and Radii. *J. Phys. Chem. A* **1964**, *68*, 441−451.
(74) Eastman, P.; Swails, J.; Chodera, J. D.; McGibbon, R. T.; Zhao, Y.; Beauchamp, K. A.; Wang, L.-P.; Simmonett, A. C.; Harrigan, M. P.; Stern, C. D.; Wiewiora, R. P.; Brooks, B. R.; Pande, V. S. OpenMM 7: Rapid Development of High Performance Algorithms for Molecular Dynamics. *PLoS Comput. Biol.* **2017**, *13*, No. e1005659.
(75) Andersen, H. C. Rattle: A "Velocity" Version of the Shake Algorithm for Molecular Dynamics Calculations. *J. Comput. Phys.* **1983**, *52*, 24−34.
(76) Miyamoto, S.; Kollman, P. A. Settle: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992**, *13*, 952−962.
(77) Lyubartsev, A. P.; Martsinovski, A. A.; Shevkunov, S. V.; Vorontsov-Velyaminov, P. N. New Approach to Monte Carlo Calculation of the Free Energy: Method of Expanded Ensembles. *J. Chem. Phys.* **1992**, *96*, 1776−1783.
(78) Escobedo, F. A.; Martínez-Veracoechea, F. J. Optimized Expanded Ensembles for Simulations Involving Molecular Insertions and Deletions. I. Closed Systems. *J. Chem. Phys.* **2007**, *127*, No. 174103.
(79) Chodera, J. D.; Shirts, M. R. Replica Exchange and Expanded Ensemble Simulations as Gibbs Sampling: Simple Improvements for Enhanced Mixing. *J. Chem. Phys.* **2011**, *135*, No. 194110.
(80) Shirts, M. R.; Chodera, J. D. Statistically Optimal Analysis of Samples from Multiple Equilibrium States. *J. Chem. Phys.* **2008**, *129*, No. 124105.
(81) Chodera, J. D.; Swope, W. C.; Pitera, J. W.; Seok, C.; Dill, K. A. Use of the Weighted Histogram Analysis Method for the Analysis of Simulated and Parallel Tempering Simulations. *J. Chem. Theory Comput.* **2007**, *3*, 26−41.
(82) Voelz, V. A.; Dill, K. A.; Chorny, I. Peptoid Conformational Free Energy Landscapes from Implicit-Solvent Molecular Simulations in AMBER. *Biopolymers* **2011**, *96*, 639−650.
(83) Mannige, R. V.; Haxton, T. K.; Proulx, C.; Robertson, E. J.; Battigelli, A.; Butterfoss, G. L.; Zuckermann, R. N.; Whitelam, S. Peptoid Nanosheets Exhibit a New Secondary-Structure Motif. *Nature* **2015**, *526*, 415−420.
(84) Brandt, W.; Herberg, T.; Wessjohann, L. Systematic Conformational Investigations of Peptoids and Peptoid-Peptide Chimeras. *Biopolymers* **2011**, *96*, 651−668.
(85) Lin, J. Divergence Measures Based on the Shannon Entropy. *IEEE Trans. Inform. Theory* **1991**, *37*, 145−151.
(86) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.
(87) Piana, S.; Donchev, A. G.; Robustelli, P.; Shaw, D. E. Water Dispersion Interactions Strongly Influence Simulated Structural Properties of Disordered Protein States. *J. Phys. Chem. B* **2015**, *119*, 5113−5123.
(88) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2017**, *14*, 71−73.
(89) Rubinstein, M.; Colby, R.H. *Polymer Physics*; Oxford University Press: Oxford, 2003.
(90) McCarney, E. R.; Werner, J. H.; Bernstein, S. L.; Ruczinski, I.; Makarov, D. E.; Goodwin, P. M.; Plaxco, K. W. Site-Specific Dimensions Across a Highly Denatured Protein; A Single Molecule Study. *J. Mol. Biol.* **2005**, *352*, 672−682.
(91) Fuertes, G.; Banterle, N.; Ruff, K. M.; Chowdhury, A.; Mercadante, D.; Koehler, C.; Kachala, M.; Estrada Girona, G.; Milles, S.; Mishra, A.; Onck, P. R.; Gräter, F.; Esteban-Martín, S.; Pappu, R. V.; Svergun, D. I.; Lemke, E. A. Decoupling of Size and Shape Fluctuations in Heteropolymeric Sequences Reconciles Discrepancies in SAXS vs. FRET Measurements. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, E6342−E6351.

https://doi.org/10.1021/acs.macromol.1c00550
Macromolecules 2021, 54, 5011−5021