# Revisiting the Helical Cooperativity of Synthetic Polypeptides in Solution

Yuan Ren,[†] Ryan Baumgartner,[§] Hailin Fu,[†] Paul van der Schoot,[⊥] Jianjun Cheng,*[,§,∥] and Yao Lin*[,†,‡]

[†]Department of Chemistry, [‡]Polymer Program, Institute of Material Science, University of Connecticut, Storrs, Connecticut 06269, United States

[§]Department of Chemistry, [∥]Department of Materials Science and Engineering, University of Illinois at Urbana–Champaign, Urbana, Illinois 61801, United States;

[⊥]Department of Applied Physics, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands

**S** Supporting Information

**ABSTRACT:** Using synthetic polypeptides as a model system, the theories of helix–coil transition were developed into one of the most beautiful and fruitful subjects in macromolecular science. The classic models proposed by Schellman and Zimm–Bragg more than 50 years ago, differ in the assumption on whether the configuration of multiple helical sequences separated by random coil sections is allowed in a longer polypeptide chain. Zimm also calculated the critical chain lengths that facilitate such interrupted helices in different solvent conditions. The experimental validation of Zimm's prediction, however, was not carefully examined at that time. Herein, we synthesize a series of homopolypeptide samples with different lengths, to systematically examine their helix–coil transition and folding cooperativity in solution. We find that for longer chains, polypeptides do exist as interrupted helices with scattered coil sections even in helicogenic solvent conditions, as predicted in the Zimm–Bragg model. The critical chain lengths that facilitate such interrupted helices, however, are substantially smaller than Zimm's original estimation. The inaccuracy is in part due to an approximation that Zimm made in simplifying the calculation. But more importantly, we find there exist intramolecular interactions between different structural segments in the longer polypeptides, which are not considered in the classic helix–coil theories. As such, even the Zimm–Bragg model in its exact form cannot fully describe the transition behavior and folding cooperativity of longer polypeptides. The results suggest that long "all-helix" chains may be much less prevalent in solution than previously imagined, and a revised theory is required to accurately account for the helix–coil transition of the longer chains with potential "non-local" intramolecular interactions.
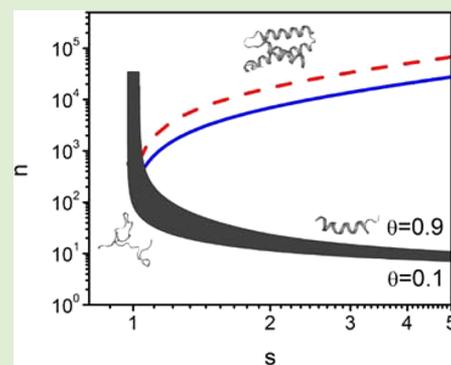
## INTRODUCTION

Numerous polypeptide chains, synthesized chemically or biologically, can reversibly transform between $\alpha$-helical and randomly coiled forms in solution, upon changing solution parameters such as temperature, solvent composition, or pH.[1–14] The sharpness or cooperativity of this helix–coil transition depends on the length, sequence, and interactions of the polypeptide chains. The statistical mechanics describing the helix–coil transition of polypeptides were developed by Schellman,[15] Gibbs,[16] Zimm,[17,18] Lifson,[19] and Nagai,[20] mostly in the 1950s and 1960s. There have been many applications of the theory in both biological and synthetic macromolecules in the last half century, as well as in the context of supramolecular polymers.[21–23] The theory of helix–coil transition is thus regarded as one of the most elegant and thoroughly examined areas of polymer physical chemistry. Interestingly, the early development of helix–coil theory relied largely on synthetic homopolypeptides as the model system (e.g., poly-γ-benzyl-L-glutamate or PBLG, or poly-L-lysine).[1–3] One of the earliest theories was Schellman's model (also called the "zipper" model), which assumes that each amino acid residue of a

polypeptide chain exists in either a helical or a coil state, with all helical units occurring contiguously in a single region (like a zipper).[15] This assumption simplifies the statistical thermodynamic treatment of the helix–coil transition and provides a closed form of solutions. While the theory accurately described the behavior of short polypeptide chains, it seemed to overestimate the folding cooperativity for longer polypeptide chains (e.g., PBLG). In longer chains, the entropic advantages of having more than one helical sequence within a single chain may well compensate the energetic cost of nucleating a new helix. Hence, Zimm and Bragg considered the possibility of forming two or more helical sequences separated by coil sections (interrupted helix) in a longer polypeptide chain and utilized matrix theory to treat the partition function for a more general analysis. In their classic paper, they compared their theoretical transition curves with the experimental data obtained for three PBLG samples with an average degree of

**Scheme 1. Different Conformations of Chains Considered in the Theories of Helix−Coil Transition[a]**



(A)   (B)   (C)   (D)   (E)
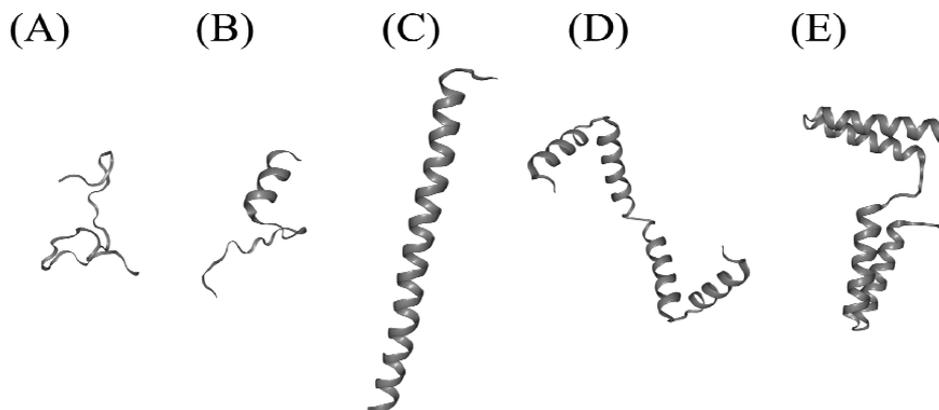
[a]From left to right: (A) all-coil, (B) partial-helix, (C) all-helix, (D) an interrupted helix with scattered coil sections, (E) an interrupted helix with sections that interact.

polymerization (DP) of 1500, 46, and 26 and found reasonably good agreements between them.[17]

A relevant question for an experimentalist is, for a given system and under a specific environmental condition, how do we know whether the polypeptide should be described as being composed of single helices or interrupted helices? This question is particularly relevant to increasing use of long, rod-like helical polypeptides or other foldable polymers in various contemporary applications such as protein mimics, drug delivery, molecular electronics, and self-organized materials.[24−36] Theoretically, one does expect the formation of breaks between helical stretches when the length of the chain exceeds the correlation length, which is set by the level of cooperativity of the helix−coil transition.

In both Schellman's and Zimm−Bragg's models, the helix−coil transition can be described as a function of three key parameters: $n$, number of residues or DP of the polymer; $s$, a propagation equilibrium constant that is equal to the probability of helical continuation relative to the probability of termination; and $\sigma$, a factor much less than unity and represents the barrier in starting a helical section. This last quantity describes the cooperativity of the helix−coil transition: the smaller its value, the sharper the helix−coil transition. In Schellman's model,[15] as $s$ begins to increase by changing either the temperature or the solvent composition, all-coil (or random) chains are converted into a mixture of all-coil chains and chains containing one continuous helical section, and eventually into all-helix chains with the occasional disorder at the ends (Scheme 1A−C). Zimm's model[17] predicts the same behavior for the low molecular weight of polypeptides, however, for sufficiently high molecular weight (e.g., DP > 1000 for $\sigma = 10^{-4}$ according to ref 17), all-coil chains will first enter a transition region in which the chains have alternating random and helical sections, and eventually into helices with scattered random sections and terminal disorder (interrupted helices, e.g., see Scheme 1D). For very high molecular weight polypeptides, all-helix chains are not possible, no matter what solution parameters are (e.g., even when $s$ is considerably larger than unity). Fundamentally, this reflects the absence of a phase transition in quasi one-dimensional systems characterized by short-ranged interactions, even in the thermodynamic limit (e.g., for infinite chains).

As the polymer chain length ($n$) increases under conditions that favor the formation of helices ($s > 1$), however, the predicted breakdown of single helices into interrupted helices has not been carefully examined. There may be at least two reasons behind this. First, early interest in this area was focused primarily on establishing a basic folding mechanism in proteins; polypeptides with more than a few hundred residues were not as relevant as the shorter chains. Second, high molecular weight polypeptides are mainly synthesized by ring-opening polymerization (ROP) of $\alpha$-amino acid-$N$-carboxyanhydride (NCA) monomers, which at that time, were plagued by undesired side reactions that give rise to products with a broad molecular weight distribution.[37,38] This difficulty, however, has been overcome by the development of using high vacuum techniques and modern initiators that largely eliminate the side reactions for a "living" polymerization process.[39−45] Herein, we utilized the controlled ROP methods[43,44] to synthesize a series of PBLG based polymers which spanned both large and small DPs and examined the helix−coil transition in solution thoroughly using nuclear magnetic resonance spectroscopy (NMR). We confirm that, as predicted by Zimm−Bragg model, for longer chains in solution, polypeptides do exist as interrupted helices with scattered coil sections even at high values of $s$; and in that region, Schellman's model cannot accurately account for the helix−coil transitions in the experiments. However, the transition of single helices into interrupted helices with increasing $n$ occurs much earlier than that originally described in Zimm's phase diagram. The discrepancy is due to two reasons. First, an approximation utilized by Zimm in the calculation[17] led to overestimation of the critical $n$ that allows for the formation of interrupted helix. This can be easily remedied by removing the approximation and recalculating the phase diagram. Second, we find that intramolecular interactions may exist from different segments of the longer chains; even the units involved in contacts are separated along the chains. As these types of "non-local" interaction were not considered in the classic helix−coil theories, even Zimm−Bragg model in its exact form cannot fully describe the transition behavior of longer PBLGs.

## ■ EXPERIMENTAL SECTION

**General.** All reagents and solvents were purchased from Sigma-Aldrich and used as received unless otherwise specified. $\gamma$-Benzyl-$L$-glutamic acid was purchased from Chem-Impex. All polymerizations were mixed in an MBraun glovebox under argon. Dichloromethane (DCM) and 1,4-dioxane were prepared by refluxing over $CaH_2$ for 24 h distilling, and purging with nitrogen gas for 15 min. Dry hexanes and THF were prepared by passing nitrogen purged solvents through

activated alumina columns. Anhydrous chloroform was purchased from Sigma-Aldrich and used as received. Anhydrous triethylamine was purchased form Sigma-Aldrich and used as received for polymerizations. All dry solvents were stored over 4 Å sieves in the glovebox. All vials used to handle trimethylsilyl (TMS) protected amines were silanzed by allowing vials to sit over vapor of chloro-trimethylsilane for 4 h in a desiccator under static vacuum. Vials were rinsed with deionized water, dried at 100 °C, and stored in the glovebox. γ-benzyl-L-glutamate-N-carboxyanhydride (BLG-NCA) and hexylamine-TMS were synthesized according to literature procedures.[43,44]

**Instrumentation.** NMR spectra were recorded on a Bruker DRX 500 MHz spectrometer for polymer characterization. Chemical shifts are referenced to residual protons in the deuterated NMR solvents. MestReNova was used to analyze all spectra. Gel permeation chromatography (GPC) was performed on a system equipped with a Model 1200 isocratic pump (Agilent Technology) in series with a 717 Autosampler (Waters) and size exclusion columns ($10^2$ Å, $10^3$ Å, $10^4$ Å, $10^5$ Å, $10^6$ Å Phenogel columns, 5 $\mu$m, 300 × 7.8 mm, Phenomenex), which were maintained at a temperature of 60 °C. A DAWN HELEOS (Wyatt Technology) multiangle laser light scattering (MALLS) operating at a wavelength of 658 nm and an Optilab rEX refractive index detector (Wyatt Technology) operating at a wavelength of 658 nm were used as detectors. The mobile phase consisted of N,N-dimethylformamide (DMF) containing 0.1 M LiBr at a flow rate of 1 mL min$^{-1}$. Samples were filtered through a 0.45 $\mu$m PTFE filter before analysis. Absolute molecular weights of poly(γ-L-benzyl-glutamate) (PBLG) samples were determined using ASTRA 6.1.1.17 software (Wyatt Technology) and calculated using a d$n$/d$c$ value of 0.0930.

**Synthesis of Homo PBLGs.** (Table 1, Entries 1−7) Linear PBLGs were synthesized either via hexylamine-N-TMS mediated NCA

**Table 1. Synthesis of PBLGs**

| entry | sample | $M_w$ (kDa) | $M_n$ (kDa) | PDI |
|---|---|---|---|---|
| 1 | PBLG$_{42}$ | 10.7 | 9.1 | 1.18 |
| 2 | PBLG$_{90}$ | 20.5 | 19.7 | 1.04 |
| 3 | PBLG$_{147}$ | 33.7 | 32.1 | 1.05 |
| 4 | PBLG$_{232}$ | 36.1 | 50.8 | 1.17 |
| 5 | PBLG$_{308}$ | 50.3 | 67.4 | 1.13 |
| 6 | PBLG$_{484}$ | 131 | 106 | 1.24 |
| 7 | PBLG$_{1228}$ | 312 | 269 | 1.16 |

polymerization (Procedure A) in DCM or by triethylamine initiated polymerization in 1,4-dioxane and chloroform (Procedure B). The $M_n$, $M_w$, and PDI ($M_w/M_n$) were determined by GPC and summarized in Table 1. $^1$H NMR (TFA-$d$, 500 MHz): δ 7.39−7.25 (ArH−), 5.08 (C$_6$H$_5$CH$_2$−), 4.57 (CH−NH−, αH−), 2.90 (−CH2−,γH−), 2.44, 2.11 (−CH2−, βH−).

*Procedure A.* In a glovebox, BLG-NCA (60.0 mg, 0.23 mmol) was dissolved in DCM (0.46 mL) to yield a 0.5 M solution. Then, hexylamine-N-TMS was added as a solution in DCM at a designed [M]:[I] ratio = 20, 50, 100, 165, or 230 and the reactions were stirred for 24 h. After completion of the reactions, the polymer solutions were then precipitated into hexanes and dried *in vacuo* to yield PBLG as a white solid (yields typically >80%). [M]$_0$:[I]$_0$ ratios = 20, 50, 100, 165, and 230 yielded polymers of $M_n$ = 9.1, 19.7, 32.1, 50.8, and 67.4 kDa, respectively.

*Procedure B.* In a glovebox, BLG-NCA (50.0 mg, 0.19 mmol) was dissolved in 1 mL of either dry 1,4-dioxane or chloroform (CHCl$_3$), yielding a 0.2 M solution. Thereafter, triethylamine (20 $\mu$L, 0.14 mmol) was added and the reactions were stirred for 24 h. The then viscous polymer solutions were precipitated into diethyl ether and dried *in vacuo* at room temperature to yield the final polymer PBLG in 22% (9 mg, 1,4-dioxane, DP = 1228, $M_n$ = 269 kDa) or 46% (19 mg, CHCl$_3$, DP = 484, $M_n$ = 106 kDa) yield as pure white solids.

**Trifluoroacetic Acid (TFA) Induced Helix−Coil transitions of PBLGs.** Solvent induced helix−coil transition studies were carried out

on a Bruker DRX 500 MHz spectrometer, using premeasured compositions of TFA-$d$ and CDCl$_3$ as the solvent at 300 K. PBLG has been extensively studied as a model for the helix−coil transition in polypeptides, and the assignments for the helix−coil transition in both the helix and coil form have been well established and were followed in this study.[46−49] PBLG exists as a helix in chloroform and transits into coil structures with increasing amount of TFA added into the solution. In these studies, at least 1% TFA-$d$ has been added into the solution to prevent the aggregation of PBLG chains. $^1$H NMR was used to identify different secondary structures of PBLGs based on the chemical shifts of α-CH.

**Temperature-Induced Helix−Coil Transitions of PBLGs.** Temperature-induced helix−coil transition studies were carried out on a Bruker DRX 500 MHz spectrometer, using predetermined compositions of TFA-$d$ and CDCl$_3$ as the solvent. The effect of temperature on the helix−coil transition has also been extensively studied.[3] PBLG tends to exist as a helix at higher temperature and converts into coil structures when temperature decreased. Since the helix−coil transition of PBLG is affected by both temperature and TFA concentration, different amounts of TFA-$d$ were added into the solution so that the temperature range of helix−coil transition was accessible under the instrumental conditions. $^1$H NMR was used to identify different secondary structures of PBLGs based on the chemical shifts of α-CH. NMR calibration tube containing methanol or ethylene glycol was used to measure the exact temperature of sample chamber by measuring the chemical shift separation between the OH resonances and CH$_n$ resonance.[50−52] Before each NMR acquisition, the samples were allowed sufficient time to reach equilibrium at the set temperature.

**NOESY Study of Homo PBLGs.** Polymers for NOESY experiments were dissolved in 99:1 (v:v) CDCl3:TFA-$d$ and sealed in NMR tubes to prevent solvent evaporation. 2D NOESY experiments were performed on a Bruker DRX-500 MHz spectrometer with the $(\pi/2)$-t1-$(\pi/2)$-$\tau_m$-$(\pi/2)$-t2 pulse sequence. Spectra were acquired with t$_1$ at 0.15 ms. The $\pi/2$ pulse width was 8.4 $\mu$s, $\tau_m$ was 100 ms, and the delay between acquisitions was 2 s.

## ■ THEORETICAL BASIS

We first compare the different predictions from the Schellman and Zimm−Bragg models on how the average conformational property varies with chain length. For both models, the average fractional helicity θ can be expressed as

$$\theta = \left(\frac{1}{n}\right)\left[\frac{\partial(\ln q)}{\partial(\ln s)}\right] \tag{1}$$

with $q$ the partition function of the chain, which is model dependent.

In Schellman's model,[15] the partition function is $q = 1 + \sum_{k=1}^{n}(n - k + 1)\sigma s^k$ for a chain of $n$ units, thus $\theta$ can be expressed as

$$\theta = \frac{\sigma s}{(s-1)^3}\left[\frac{ns^{n+2} - (n+2)s^{n+1} + (n+2)s - n}{n(1 + (\sigma s/(s-1)^2)(s^{n+1} + n - (n+1)s))}\right] \tag{2}$$

For small $\sigma$, it predicts a relatively sharp coil-to-helix transition when $s$ increases, and the sharpness of this transition (or apparent cooperativity, S) can be defined as the slope of a $\theta$ versus $s$ plot at midpoint of transition.[7] We can evaluate the dependence of S on chain length, by taking a simple derivative of $\theta$ with respect to $s$, and evaluating the maximum value of the derivative. The black line in Figure 1 shows that the sharpness of the helix−coil transition increases almost linearly with $n$, based on Schellman's model. The model predicts infinite cooperativity for infinitely long chains. This implies that

Schellman's model implicitly presumes long-range interactions between the monomer units along the backbone of the chain.

In the Zimm−Bragg model,[17] the partition function is defined as

$$q = [\lambda_1^{n+1}(1 - \lambda_2) - \lambda_2^{n+1}(1 - \lambda_1)]/(\lambda_1 - \lambda_2) \quad (3)$$

where $\lambda_1$ and $\lambda_2$ are the two eigenvalues of the matrix operator as

$$\lambda_{1,2} = \{(1 + s) \pm [(1 - s)^2 + 4\sigma s]^{1/2}\}/2 \quad (4)$$

By combining eqs 1, 3 and 4, the apparent cooperativity based on the Zimm−Bragg model can be evaluated in the same way. As shown from the red line in Figure 1, it only agrees with
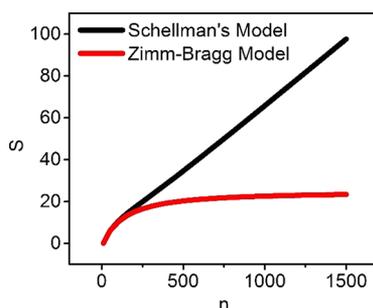


**Figure 1.** Plots of the apparent cooperativity versus degree of polymerization based on Schellman's model (in black line) and Zimm−Bragg model (in red line), with $\sigma = 10^{-4}$.

Schellman's model for small $n$, and becomes saturated when $n$ gets larger (e.g., $\sim 250$ for $\sigma = 10^{-4}$). This behavior arises from the model's consideration of the entropically favorable formation of two or more helical sequences separated by coil sections (interrupted helix) in a longer polypeptide chain. As Schellman's model does not consider the configuration of multiple helical sequences in the partition function, the gap between the two model-based predictions can be regarded as a measure of the increasing statistical weight of chain configurations with two or more helical sections in solution. As shown in Figure 1, for helical polymers with $\sigma$ of $10^{-4}$, the entropic effect for $n$ larger than $\sim 300$ cannot be ignored.

Zimm and Bragg[17] had previously predicted that a critical $n$ exists for the transition of single helices into interrupted-helices at specific $s$, by calculating the average number of helical regions $\langle j \rangle$ from the partition function:

$$\langle j \rangle = \partial(\ln q)/\partial(\ln \sigma) \quad (5)$$

Their result, however, suggested that interrupted helices are insignificant at helicogenic condition for polymers with $n < 1000$ and $\sigma = 10^{-4}$ (see Figure 5 in ref 17 or the red dash line in Figure 2). The quick comparison in Figure 1 suggests that Zimm and Bragg might have considerably overestimated the critical $n$ required for transitioning from single helices to interrupted-helices at $s > 1$ in their diagram. We find that the inaccuracy originates from the approximation used in their calculation of the average number of helical region $\langle j \rangle$, in which Zimm took $q \cong \lambda_1^{n+1}(1 - \lambda_2)/(\lambda_1 - \lambda_2)$ to simplify the calculation. However, this approximation is only valid for large $n$. To correct this, we replot the transition at various $n$ and $s$ values in the $n$−$s$ diagram (the blue solid line in Figure 2) without taking the approximation in calculating the partition function of Zimm−Bragg. The line, which runs diagonally through the helical region, defines the conditions under which half the
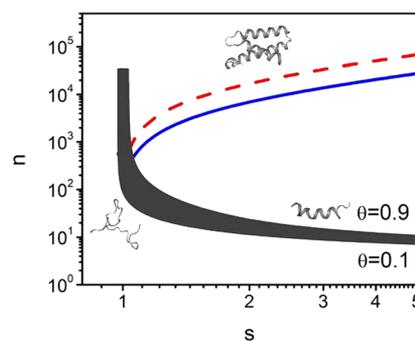


**Figure 2.** $n$−$s$ plane for $\sigma = 10^{-4}$. The boundary of the transition is chosen to be $\theta = 0.1$ and $\theta = 0.9$. The diagonal line indicates the conditions where half of the chains in the solution contain a single unbroken helix; the other half contain two helical sections connected with a coil segment. The solid line represents the calculation without taking the approximation in calculating the partition function of Zimm−Bragg, while the dash line represents the calculation with the approximation originally taken by Zimm and Bragg.

chains contain only one helical section, while the other half contains two (i.e., $< j \geq 1.69$, to be consistent with Zimm's original figure). In comparison with the original prediction of Zimm, this line shifts considerably to lower values of $n$. This result implies that in many more situations than originally projected, longer helical polypeptides or other foldable polymers in solution may not exist as single helices with only occasional disorder at the ends. Rather, they may tend to form multiple helical sections with scattered coil sections, and this flexibility could even allow for some intrachain interactions from different sections.

To picture the breakdown of the single helices into interrupted helices with increasing $n$ in helicogenic conditions, we utilize an intuitive "toy" model to estimate the critical $n$ ($n^*$), above which two helical sections are statistically more populated than single helices in solution. Consider a polymeric chain consisting of $n$ units that has undergone the coil to helix transition. Conditions are such that the expected value of the number consecutive helical repeat units is equal to $k \leq n$, making the fractional helicity equal to $\theta = k/n$. Here, $s\sigma$ is essentially the Boltzmann factor of the free energy difference between the helical and nonhelical state of a link, and $\ln \sigma < 0$ is the energy gain of two neighboring helical links that make contact, in units of thermal energy. The helical segment of $n$ units is free to start at any point along the chain, $i = 1, 2, ..., n - k + 1$. The partition function of this configuration reads:

$$q_1(k) = (n - k + 1)s^k\sigma \quad (6)$$

where the first term accounts for the degeneracy of the configuration. The probability distribution $P(k)$ of $k$ consecutive helical links is $P_1(k) = q_1(k)/q_1$ with $q_1 \equiv 1 + \sum_{k=1}^{n} q_1(k)$. Hence, the expected value of the number of helical segments can be calculated from $\langle k \rangle = \partial(\ln q_1)/\partial(\ln s)$. The fraction of helical bonds is then $\theta \equiv \langle k \rangle/n = (1/n)[\partial(\ln q_1)/\partial(\ln s)]$.

We then ask the question, for what values of $n$ (where $n \geq k - 1$) does a chain consisting of $k$ contiguous helical repeat units break into two helical subsections that have at least one nonhelical link separating them? The partition function of this configuration reads

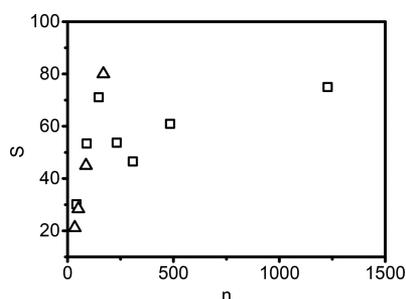$$q_2(k) = \frac{1}{2}(k - 1)(n - k)(n - k + 1)s^k\sigma^2 \quad (7)$$

**Figure 3.** TFA-induced helix-to-coil transition of PBLGs at room temperature (the data from the newly synthesized samples are in open squares, the old data from ref 49 are shown in open triangles for comparison). Linear correlation of measured apparent cooperativity (S) with DP was found for short PBLG chains ($n < 250$), while deviation was observed for longer PBLG chains.

The chain transitions from one helical stretch to two helical stretches provided $q_2(k) \geq q_1(k)$, or when

$$\frac{1}{2}(k-1)(n^* - k)\sigma \geq 1 \tag{8}$$

Inserting the expected value $\langle k \rangle = n\theta$, we find

$$n^* \geq \sqrt{\frac{2}{\sigma(1-\theta)\left(\theta - \frac{1}{n}\right)}} \tag{9}$$

Of course, $\theta$ depends on $n$ as well (e.g., in the simplest case, by the eq 2 based on Schellman's model), so the problem should be solved self-consistently for given values of the energetic parameters $s$ and $s\sigma$ to obtain $n^*$. This can be easily accomplished graphically by combining eq 9 and eq 2. The $n^*$

determined from this method (Figure S1) predicts a similar diagonal line as the blue line shown in Figure 2.

In the next section, we discuss the experimental results we obtained from examining the average conformational property and helical parameters of PBLGs with different chain lengths in solution, and how they compare with the theoretical predictions.

## ■ RESULTS AND DISCUSSION

Using NMR analysis,[46−48] we can experimentally measure the average fractional helicity of PBLGs as a function of temperature at a given solvent composition, or as a function of added solvent composition at a specific temperature. The sharpness of the transition obtained from the experiments on the PBLGs with different DPs can be used to qualitatively compare to the theoretical predictions in Figure 1, with the caveat that $s$, the propagation equilibrium constant, is a function of both solvent composition and the temperature. Previously, we have evaluated the trifluoroacetic acid (TFA) induced helix−coil transition of four PBLG samples (DP = 34, 51, 87, 169 with a narrow molecular weight distribution) in solution, and found that the sharpness of the transition to approximately follow a linear dependence on their chain length (shown in open triangles in Figure 3).[49] To examine the transition behaviors in a broader range of $n$, we now synthesized seven additional samples with (DP = 42, 90, 147, 232, 308, 484, 1228) via controlled NCA polymerization, adapting the methods we previously reported.[43,44] We refer to the Experimental Section and Supporting Information for the details, and Table 1 and Figure S2 for the characterization of the polymers. We then carried out both the temperature and TFA-induced helix-to-coil
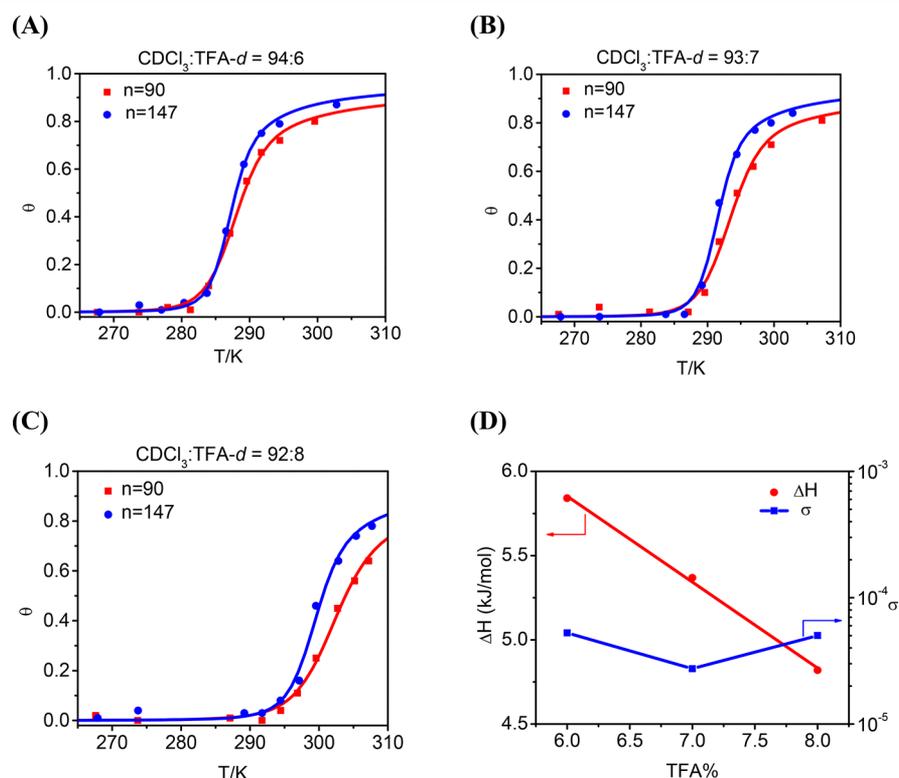


**Figure 4.** Model-based analysis of temperature-induced helix−coil transition of two PBLG samples ($n = 90$ and $147$) at specific solvent conditions. (A) CDCl$_3$:TFA-$d$ = 94:6, (B) CDCl$_3$:TFA-$d$ = 93:7, (C) CDCl$_3$:TFA-$d$ = 92:8, and (D) the dependence of the helix parameters obtained from panels A−D on the solvent composition.
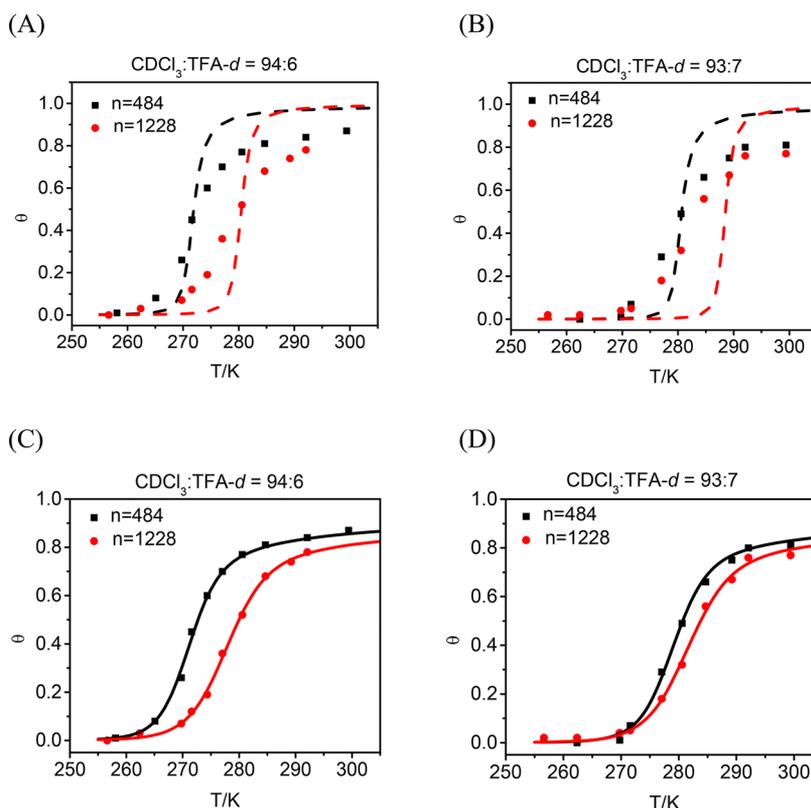
**Figure 5.** Model-based analysis of temperature-induced helix−coil transition of two PBLG samples ($n$ = 484 and 1228) at specific solvent conditions. (A) CDCl$_3$:TFA-$d$ = 94:6, the dash lines were based on Zimm−Bragg model with $\sigma$ = 5.3 × 10$^{-5}$ and $\Delta H_g$ = 5.8 kJ/mol, as obtained in Figure 4D; (B) CDCl$_3$:TFA-$d$ = 93:7, the dash lines were based on Zimm−Bragg model with $\sigma$ = 2.7 × 10$^{-5}$ and $\Delta H_g$ = 5.4 kJ/mol, as obtained in Figure 4D; (C) same as (A), but the value $n$ used in the model was instead set to be 62 for PBLG$_{484}$ and 49 for PBLG$_{1228}$ to get a good curve fit; (D) same as panel B, but $n$ used in the model was instead set to be 62 for PBLG$_{484}$ and 49 for PBLG$_{1228}$.

transition studies on these samples using NMR (please see Figure S3−S33 in the Supporting Information).

Figure 3 summarizes the measured sharpness of transition from the PBLG samples versus their chain length, as obtained from the TFA-induced transition studies at 300 K. While the sharpness of the transition of short chains ($n$ < 200 for the PBLGs in the specific solution conditions) increases almost linearly with $n$, in agreement with both Schellman's and Zimm's prediction, Zimm's model is necessary to describe the saturation of the apparent cooperativity for longer chains. Apparently, considering the chain configuration of multiple helical sections is necessary even for PBLGs of a couple of hundred repeat units. Long PBLG chains may well remain as interrupted helices even under strongly helicogenic conditions. To further elucidate the dependence of the helix−coil transition on the chain length, we selected two samples of relatively short chains (DP = 90 and 147), and two others with the highest molecular weight (DP = 484 and 1228), to perform the model-based analysis on the full set of data of thermally induced helix−coil transitions.

Figure 4A shows the fraction of helical residues as a function of temperature, for two samples of low molecular weight ($n$ = 90 and 147) in a mixture of chloroform-$d$ and TFA-$d$ (94:6). Following Zimm's approach[18] to relate the parameter $s$ and $\sigma$ to experimental variables, we also assume that the equilibrium constant $s$ depends on the temperature but $\sigma$ does not. Additionally, the relation of $s$ to temperature should follow the Van't Hoff equation as

$$\ln s = \left( \frac{\Delta H_g}{R} \right) \left[ \frac{T - T_m}{TT_m} \right] \tag{10}$$

where $\Delta H_g$ is the enthalpy change for addition of one helical unit onto a preexisting helical section, and $T_m$ is the temperature at the midpoint of the transition. By combining eq 1, 3 and 4, we found that the optimized fit to the transitions in Figure 4A can be obtained by assigning $\sigma$ = 5.3 × 10$^{-5}$, and $\Delta H_g$ = 5.8 kJ/mol. The same approach has been utilized to compare experimental data points and the theoretical transition curves for two PBLG samples at the other solvent conditions (CDCl$_3$:TFA-$d$ = 93:7 in Figure 4B and 92:8 in Figure 4C). The result demonstrates that the Zimm's model is able to describe the data from these experiments in excellent agreement. Figure 4D shows the dependence of $\Delta H_g$ and $\sigma$ on the amount of added TFA-$d$ in the solvent. Previously, Zimm guessed that $\sigma$ might be also independent of solvent, as this nucleation constant should depend only on the polymers (e.g., the difficulty in making the first hydrogen bond by immobilizing six dihedral rotation angles).[17] Our result seems to agree with Zimm's proposal, although more solvent conditions should be tested to draw a firm conclusion on this point. And with a $\sigma$ on the order of 10$^{-4}$, chain configuration with two or more helical sections should be statistically more populated than single helices even in helicogenic conditions for $n$ > ∼400, as predicted by the corrected $n$−$s$ diagram of Figure 2.

The thermally induced helix−coil transition for the two highest molecular weight PBLGs synthesized ($n$ = 484 and 1228), however, are not well described by the Zimm−Bragg
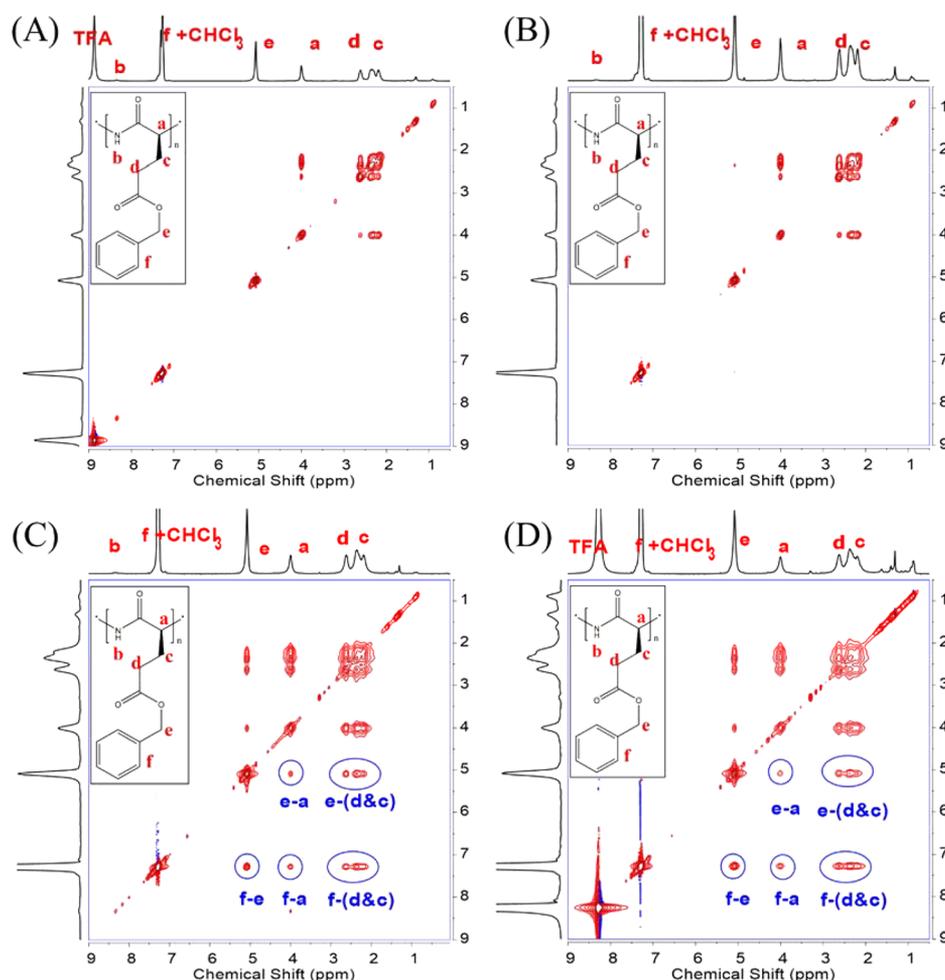
**Figure 6.** $^1H$ NMR 2D NOESY spectra of (A) PBLG$_{90}$, (B) PBLG$_{147}$, (C) PBLG$_{484}$, and (D) PBLG$_{1228}$ obtained by the $(\pi/2)$-t1-$(\pi/2)$-$\tau_m$-$(\pi/2)$-t2 pulse sequence.

model. The comparison of the experimental results and the theoretical curves using the values of $\sigma$ and $\Delta H_g$ in Figure 4D is shown in Figure 5A,B. The helix−coil transitions are much broader than predicted by Zimm−Bragg model (i.e., a much weaker folding cooperativity). An acceptable fit to the experimental data can only be approached by applying artificially small $n$ (e.g., ∼ 62 for PBLG$_{484}$ and ∼49 for PBLG$_{1228}$) in the Zimm−Bragg model, as shown in Figure 5C,D. This means that these two polymers with high molecular weights have a folding cooperativity like a chain of merely 50−60 repeats. This type of behavior is typically found only in situations where the continuity of the H-bond networks along the polypeptide chain is disrupted by nonlocal interactions, e.g., from neighboring grafted chains in comb-like macromolecules.[49] We further performed the model-based analysis on the PBLG$_{232}$ and PBLG$_{308}$ (Figure S36) and found a similar behavior. The maximal contiguous helix seems to be around 200 repeats for the specific system. The result indicates that completely unperturbed helical chains are not always possible for these high molecular weight PBLGs in solution (even when $s$ is considerably larger than unity at higher temperature). Certain interactions can interrupt the folding continuity of these longer chains, causing a smaller folding cooperativity than that predicted by Zimm−Bragg model. The overall structure of these interrupted helices may thus possess considerable flexibility arising from the coil sections in between the helical sections. The structural sections

in the same chain may interact with each other either in the form of bundles between two helical sections, or as entanglements or by hydrogen bonding between two coil sections. These types of nonlocal interactions were not taken into account in Zimm's model or other classic theories of helix−coil transitions, but was recently considered by Ghosh and Dill for helix bundle proteins.[53] In order to probe whether these nonlocal interactions do exist in PBLGs with high molecular weight, we carried out Nuclear Overhauser Enhancement Spectroscopy (NOESY) experiments on the samples.

Figure 6 compares $^1H$ NMR 2D NOESY spectra of four PBLG samples in 99:1 CDCl$_3$:TFA-$d$ at 25 °C (a strongly helicogenic condition), obtained with a mixing time of 100 ms. Cross peaks connecting several of the resonances were observed for the aromatic and benzylic methylene protons in PBLG$_{484}$ and PBLG$_{1228}$, showing that these protons are in close contact with $\alpha$-CH, $\beta$-CH, and $\gamma$-CH, resulting in an exchange of magnetization (within ∼0.5 nm). Similar intramolecular interactions was also found in PBLG$_{232}$ and PBLG$_{308}$ (Figure S34−35), albeit weaker than that of PBLG$_{484}$ and PBLG$_{1228}$. By contrast, no NOE cross peaks were observed for the aromatic or benzylic methylene protons in PBLG$_{90}$ and PBLG$_{147}$, or any other PBLG samples[49] we measured previously with $n$ < 200. We note here, the intermolecular association of PBLGs is very unlikely in the examined solution conditions with the added TFA. Doty previously studied the association of PBLGs caused

by various solvents and concluded that the mostly likely association is of end-to-end type,[54] which does not interrupt the helical continuity of the chains. Presumably, the flexible coil region in-between the helical sections provide the opportunity for separated helical stretches to interact and gain additional stability (e.g., solvophobically). As such, the long chain PBLGs (e.g., $n > 300$) in solution may indeed exist as the configurations shown in Scheme 1E, in which the individual segments of the chain may interact. These nonlocal interactions may also explain the difficulty in the use of more-ideal Zimm−Bragg model to analyze the helix−coil transition curves for high molecular weight samples in Figure 5. A statistical mechanics treatment similar to the recent Ghosh-Dill model[53] may allow for incorporating the chain configurations with nonlocal interactions into the partition function to describe the folding behaviors of long helical polypeptide, and is an interesting subject of future studies.

## CONCLUSIONS

In conclusion, we examined the helix−coil transition of a series of synthetic polypeptides with different lengths in solution and found that for longer chains, it becomes very difficult for the polymer chain to remain as a single contiguous helix, even under helicogenic conditions. The critical chain lengths above which the entropic effect causes this transition from all-helix to interrupted helix conformation are substantially smaller than estimated by Zimm in his original diagram. For these interrupted helices, there may exist nonlocal interactions between the different sections of polypeptides so that their helix−coil transitions deviate from the prediction based on the classic models. These additional conformations need to be considered in the partition function in the statistical mechanics of helix−coil transition of longer polypeptide chains. The results also suggest that long contiguous helical chains may be much less prevalent in solution than previously expected. Indeed, the interrupted helices were previously discovered in the melt state of different polypeptide materials, and the dynamic nature of these broken helical segments have been thoroughly investigated in a series of papers from Floudas and co-workers.[34,55−57] It will be interesting to compare the helical persistence of polypeptides or the correlation length of these helices measured in the solution and melt states, to obtain a unified understanding on this topic.

## ASSOCIATED CONTENT

### ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biomac.7b00534.

> Material synthesis, experimental procedures, and characterization data (PDF)

## AUTHOR INFORMATION

### Corresponding Authors
*E-mail: yao.lin@uconn.edu (Y.L.).
*E-mail: jianjunc@illinois.edu (J.C.).

### ORCID Ⓓ
Jianjun Cheng: 0000-0003-2561-9291
Yao Lin: 0000-0001-5227-2663

### Notes
The authors declare no competing financial interest.

## REFERENCES

(1) Blout, E. R.; Lenormant, H. Reversible Configurational Changes in Poly-L-Lysine Hydrochloride Induced by Water. *Nature* **1957**, *179* (4567), 960−963.

(2) Doty, P.; Wada, A.; Yang, J.; Blout, E. R. Polypeptides. VIII. Molecular configurations of poly-L-glutamic acid in water-dioxane solution. *J. Polym. Sci.* **1957**, *23* (104), 851−861.

(3) Doty, P.; Yang, J. T. Polypeptides. VII. Poly-γ-benzyl-L-glutamate: The helix-coil transition in solution. *J. Am. Chem. Soc.* **1956**, *78* (2), 498−500.

(4) Davidson, B.; Fasman, G. D. The conformational transitions of uncharged poly-L-lysine. α helix-random coil-β structure. *Biochemistry* **1967**, *6* (6), 1616−1629.

(5) Bovey, F. A. *Polymer Conformation and Configuration*; Academic Press: New York, 1969.

(6) Poland, D.; A, S. H. *Theory of Helix-Coil Transitions in Biopolymers*; Academic Press: New York, 1970.

(7) Cantor, C. R.; Schimmel, P. R. *Biophysical Chemistry, Part III*; W.H. Freeman: San Francisco, 1980.

(8) Pauling, L.; Corey, R. B.; Branson, H. R. The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Natl. Acad. Sci. U. S. A.* **1951**, *37* (4), 205−211.

(9) Chakrabartty, A.; Baldwin, R. L. Stability of α-Helices. *Adv. Protein Chem.* **1995**, *46*, 141−176.

(10) Presta, L. G.; Rose, G. D. Helix Signals in Proteins. *Science* **1988**, *240* (4859), 1632−1641.

(11) Gans, P. J.; Lyu, P. C.; Manning, M. C.; Woody, R. W.; Kallenbach, N. R. The helix−coil transition in heterogeneous peptides with specific side-chain interactions: Theory and comparison with CD spectral data. *Biopolymers* **1991**, *31* (13), 1605−1614.

(12) Huang, C. Y.; Klemke, J. W.; Getahun, Z.; DeGrado, W. F.; Gai, F. Temperature-dependent helix-coil transition of an alanine based peptide. *J. Am. Chem. Soc.* **2001**, *123* (38), 9235−9238.

(13) Dill, K. A. Polymer principles and protein folding. *Protein Sci.* **1999**, *8* (6), 1166−1180.

(14) DeGrado, W. F.; Wasserman, Z. R.; Lear, J. D. Protein design, a minimalist approach. *Science* **1989**, *243* (4891), 622−628.

(15) Schellman, J. A. The factors affecting the stability of hydrogen-bonded polypeptide structures in solution. *J. Phys. Chem.* **1958**, *62* (12), 1485−1494.

(16) Gibbs, J. H.; DiMarzio, E. A. Statistical mechanics of helix-coil transitions in biological macromolecules. *J. Chem. Phys.* **1959**, *30* (1), 271−282.

(17) Zimm, B. H.; Bragg, J. K. Theory of the phase transition between helix and random coil in polypeptide chains. *J. Chem. Phys.* **1959**, *31* (2), 526−535.

(18) Zimm, B. H.; Doty, P.; Iso, K. DETERMINATION OF THE PARAMETERS FOR HELIX FORMATION IN POLY-γ-BENZYL-L-GLUTAMATE. *Proc. Natl. Acad. Sci. U. S. A.* **1959**, *45* (11), 1601−1607.

(19) Lifson, S.; Roig, A. On the theory of helix-coil transition in polypeptides. *J. Chem. Phys.* **1961**, *34* (6), 1963−1974.

(20) Nagai, K. Configurational Changes of Polypeptide Molecules in the Helix-Coil Transition Region, I. *J. Phys. Soc. Jpn.* **1960**, *15* (3), 407−416.

(21) Teramoto, A.; Fujita, H., Conformation-dependent properties of synthetic polypeptides in the helix-coil transition region. In *Macroconformation of Polymers*; Springer Berlin Heidelberg: Berlin, Heidelberg, 1975; pp 65−149.

(22) Qian, H.; Schellman, J. A. Helix-coil theories: A comparative study for finite length polypeptides. *J. Phys. Chem.* **1992**, *96* (10), 3987−3994.

(23) Gestel, J. V.; Schoot, P. V. D.; Michels, M. A. J., Growth and Chirality amplification in Helical Supramolecular Polymers. In *Molecular Gels: Materials with Self-Assembled Fibrillar Networks*; Weiss, R. G.; Terech, P., Eds.; Springer Netherlands: Dordrecht, 2006; pp 79−97.

(24) Gower, L. A.; Tirrell, D. A. Calcium carbonate films and helices grown in solutions of poly(aspartate). *J. Cryst. Growth* **1998**, *191* (1−2), 153−160.

(25) van Hest, J. C. M.; Tirrell, D. A. Protein-based materials, toward a new level of structural control. *Chem. Commun.* **2001**, *19*, 1897−1904.

(26) Deming, T. J. Synthetic polypeptides for biomedical applications. *Prog. Polym. Sci.* **2007**, *32* (8−9), 858−875.

(27) Nowak, A. P.; Breedveld, V.; Pakstis, L.; Ozbas, B.; Pine, D. J.; Deming, T. J. Rapidly recovering hydrogel scaffolds from self-assembling diblock copolypeptide amphiphiles. *Nature* **2002**, *417* (6887), 424−428.

(28) Regan, L.; DeGrado, W. F. Characterization of a helical protein designed from first principles. *Science* **1988**, *241* (4868), 976−978.

(29) Bryson, J. W.; Betz, S. F.; Lu, H. S.; Suich, D. J.; Zhou, H. X.; O'Neil, K. T.; DeGrado, W. F. Protein design: A hierarchic approach. *Science* **1995**, *270* (5238), 935−941.

(30) Dong, H.; Shu, J. Y.; Dube, N.; Ma, Y.; Tirrell, M. V.; Downing, K. H.; Xu, T. 3-Helix micelles stabilized by polymer springs. *J. Am. Chem. Soc.* **2012**, *134* (28), 11807−11814.

(31) Shu, J. Y.; Panganiban, B.; Xu, T. Peptide-polymer conjugates: From fundamental science to application. *Annu. Rev. Phys. Chem.* **2013**, *64*, 631−657.

(32) Yin, L.; Tang, H.; Kim, K. H.; Zheng, N.; Song, Z.; Gabrielson, N. P.; Lu, H.; Cheng, J. Light-responsive helical polypeptides capable of reducing toxicity and unpacking DNA: Toward nonviral gene delivery. *Angew. Chem., Int. Ed.* **2013**, *52* (35), 9182−9186.

(33) Jaworek, T.; Neher, D.; Wegner, G.; Wieringa, R. H.; Schouten, A. J. Electromechanical properties of an ultrathin layer of directionally aligned helical polypeptides. *Science* **1998**, *279* (5347), 57−60.

(34) Floudas, G.; Spiess, H. W. Self-assembly and dynamics of polypeptides. *Macromol. Rapid Commun.* **2009**, *30* (4−5), 278−298.

(35) MacPhee, C. E.; Woolfson, D. N. Engineered and designed peptide-based fibrous biomaterials. *Curr. Opin. Solid State Mater. Sci.* **2004**, *8* (2), 141−149.

(36) Wang, J.; Xia, H.; Zhang, Y.; Lu, H.; Kamat, R.; Dobrynin, A. V.; Cheng, J.; Lin, Y. Nucleation-controlled polymerization of nanoparticles into supramolecular structures. *J. Am. Chem. Soc.* **2013**, *135* (31), 11417−11420.

(37) Bamford, C. H.; Elliott, A.; Hanby, W. E. *Synthetic Polypeptides*; Academic Press: New York, 1956.

(38) Kricheldorf, H. R. *α-Aminoacid-N-Carboxy-Anhydrides and Related Heterocycles*; Springer-Verlag Berlin Heidelberg: Berlin, Heidelberg, New York, London, Paris, Tokyo, 1987.

(39) Hadjichristidis, N.; Iatrou, H.; Pitsikalis, M.; Sakellariou, G. Synthesis of well-defined polypeptide-based materials via the ring-opening polymerization of α-amino acid N-carboxyanhydrides. *Chem. Rev.* **2009**, *109* (11), 5528−5578.

(40) Aliferis, T.; Iatrou, H.; Hadjichristidis, N. Living polypeptides. *Biomacromolecules* **2004**, *5* (5), 1653−1656.

(41) Deming, T. J. Facile synthesis of block copolypeptides of defined architecture. *Nature* **1997**, *390* (6658), 386−389.

(42) Deming, T. J. Polypeptide materials: New synthetic methods and applications. *Adv. Mater.* **1997**, *9* (4), 299−311.

(43) Lu, H.; Cheng, J. Hexamethyldisilazane-Mediated Controlled Polymerization of α-Amino Acid N-Carboxyanhydrides. *J. Am. Chem. Soc.* **2007**, *129* (46), 14114−14115.

(44) Lu, H.; Wang, J.; Lin, Y.; Cheng, J. One-pot synthesis of brush-like polymers via integrated ring-opening metathesis polymerization and polymerization of amino acid N-carboxyanhydrides. *J. Am. Chem. Soc.* **2009**, *131* (38), 13582−13583.

(45) Baumgartner, R.; Fu, H.; Song, Z.; Lin, Y.; Cheng, J. Cooperative polymerization of α-helices induced by macromolecular architecture. *Nat. Chem.* **2017**, *9*, 614−622.

(46) Bradbury, E. M.; Carpenter, B. G.; Crane-Robinson, C.; Rattle, H. W. E. Application of high resolution nuclear magnetic resonance to conformational analyses of polypeptides in solution. *Nature* **1968**, *220* (5162), 69−72.

(47) Bradbury, E. M.; Crane-Robinson, C.; Goldman, H.; Rattle, H. W. E. Proton magnetic resonance and the helix-coil transition. *Nature* **1968**, *217* (5131), 812−816.

(48) Markley, J. L.; Meadows, D. H.; Jardetzky, O. Nuclear magnetic resonance studies of helix-coil transitions in polyamino acids. *J. Mol. Biol.* **1967**, *27* (1), 25−40.

(49) Wang, J.; Lu, H.; Ren, Y.; Zhang, Y.; Morton, M.; Cheng, J.; Lin, Y. Interrupted helical structure of grafted polypeptides in brush-like macromolecules. *Macromolecules* **2011**, *44* (22), 8699−8708.

(50) Ammann, C.; Meier, P.; Merbach, A. A simple multinuclear NMR thermometer. *J. Magn. Reson.* **1982**, *46* (2), 319−321.

(51) Kaplan, M. L.; Bovey, F. A.; Cheng, H. N. Simplified method of calibrating thermometric nuclear magnetic resonance standards. *Anal. Chem.* **1975**, *47* (9), 1703−1705.

(52) Spees, W. M.; Song, S. K.; Garbow, J. R.; Neil, J. J.; Ackerman, J. J. H. Use of ethylene glycol to evaluate gradient performance in gradient-intensive diffusion MR sequences. *Magn. Reson. Med.* **2012**, *68* (1), 319−324.

(53) Ghosh, K.; Dill, K. A. Theory for protein folding cooperativity: Helix bundles. *J. Am. Chem. Soc.* **2009**, *131* (6), 2306−2312.

(54) Doty, P.; Bradbury, J. H.; Holtzer, A. M. Polypeptides. IV. The Molecular Weight, Configuration and Association of Poly-γ-benzyl-L-glutamate in Various Solvents. *J. Am. Chem. Soc.* **1956**, *78* (5), 947−954.

(55) Papadopoulos, P.; Floudas, G.; Schnell, I.; Klok, H. A.; Aliferis, T.; Iatrou, H.; Hadjichristidis, N. "Glass transition" in peptides: Temperature and pressure effects. *J. Chem. Phys.* **2005**, *122*, 224906.

(56) Gitsas, A.; Floudas, G.; Mondeshki, M.; Butt, H. J.; Spiess, H. W.; Iatrou, H.; Hadjichristidis, N. Effect of chain topology on the self-organization and dynamics of block copolypeptides: From diblock copolymers to stars. *Biomacromolecules* **2008**, *9* (7), 1959−1966.

(57) Papadopoulos, P.; Floudas, G.; Klok, H. A.; Schnell, I.; Pakula, T. Self-assembly and dynamics of poly(γ-benzyl-L-glutamate) peptides. *Biomacromolecules* **2004**, *5* (1), 81−91.