ACS Macro Letters

Letter

pubs.acs.org/macroletters

Conformation Model of Back-Folding and Looping of a Single DNA Molecule Confined Inside a Nanochannel

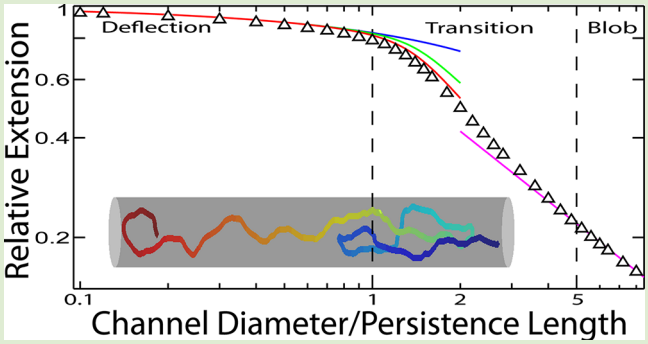
Liang Dai,[†,∥] Siow Yee Ng,[‡,∥] Patrick S. Doyle,[†,§] and Johan R. C. van der Maarel*[,‡,†]

[†]BioSystems and Micromechanics, Singapore−MIT Alliance for Research and Technology, Singapore

[‡]Department of Physics, National University of Singapore, Singapore 117542

[§]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, United States

ABSTRACT: Currently, no theory is available to describe the conformation of DNA confined in a channel when the nanochannel diameter is around the persistence length. Back-folded hairpins in the undulating wormlike chain conformation result in the formation of loops, which reduces the stretch of the molecule in the longitudinal direction of the channel. A cooperativity model is used to quantify the frequency and size of the loop domains. The predictions agree with results from the Monte Carlo simulation.



Advances in nanofabrication have made it possible to fabricate quasi one-dimensional devices with a cross-sectional diameter on the order of tens to hundreds of nanometers. These nanochannels serve as a platform for studying, among others, single DNA molecules.[1−7] Furthermore, confinement in a nanospace results in significant modification of certain important biophysical phenomena, such as the knotting probability of circular DNA and the effect of macromolecular crowding on the conformation and folding of DNA.[8−11] The conformation and dynamics of DNA in confinement have also been investigated with computer simulations.[12−15] All of these results are invaluable in, for example, the development of a better genome analysis platform and our understanding of biological processes such as DNA packaging in viruses or segregation of DNA in bacteria.[16,17]

The conformation of a wormlike chain in a nanochannel is determined by the persistence length P, width w, unconfined radius of gyration R_g, and the cross-sectional diameter of the channel D. We have investigated the relative extension $L_\parallel/L$, that is, the stretch of the molecule along the direction of the channel divided by its contour length, of DNA with Monte Carlo simulation (details will be presented below). A typical result is shown in Figure 1. Due to the interplay of the various length scales, regimes are established along the curve of $L_\parallel/L$ as a function of D/P. Odijk's deflection and Daoud and de Gennes's blob regimes are the two extreme regimes originally proposed.[18,19] In the deflection regime (i.e., $D < P$), the chain is undulating since it is deflected by the walls. As a result of the undulation with deflection length $\lambda \sim D^{2/3}P^{1/3}$, $L_\parallel/L$ is reduced with respect to its fully stretched value of unity according to

$$L_\parallel^{\mathrm{d}}/L = 1 - c(D/P)^{2/3} \quad (1)$$

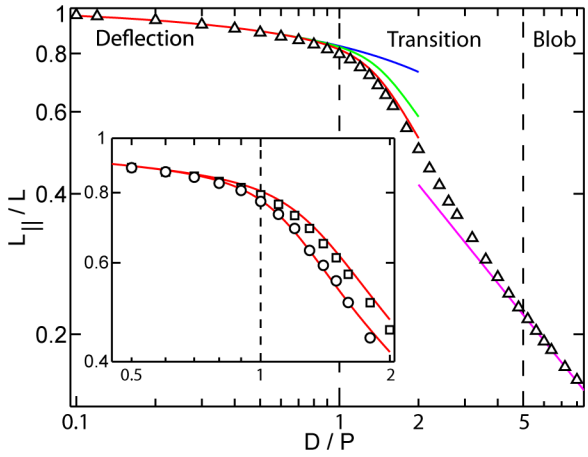


Figure 1. Monte Carlo results of the relative extension versus channel diameter divided by persistence length (50 nm). The chain width w = 10 nm and contour length L = 8 μm. The curves represent deflection (blue), deflection with S-loops (green), deflection with S- and C-loops (red), and blob (magenta) theories. The inset shows the results for w = 5 (○) and 7.5 (□) nm.

with $c = 0.1701$ for a circular cross-section.[20] In the range of $P^2/w < D < R_g$, the chain is in the blob regime.[21] In this regime, the chain statistics is described as a linearly packed array of subcoils (blobs). The relative extension is given by the scaling law

$$L_\parallel^{\mathrm{b}}/L \propto (D/P)^{-0.701} \quad (2)$$

Received: June 24, 2012
Accepted: July 31, 2012
Published: August 2, 2012

In the transition regime, the situation is less clear. Several attempts have been reported to bridge the gap. For the chain to crossover from the blob regime, the blob size reduces such that the volume interaction energy per blob becomes less than thermal energy kT. This results in modified or the same scaling of $L_{\parallel}/L$ with D, depending on the presumed chain statistics within the blob.[5,13] On the other hand, for the crossover from the deflection regime, it has been proposed that the chain performs a one-dimensional random walk through the formation of back-folded hairpin conformations.[21,22]

Here, we seek the conformation of the chain in the range $1 < D/P < 2$. In this range, the confined chain is neither in the full deflection conformation nor a series of blobs representation. However, it cannot resist thermal fluctuation that leads to hairpins along the chain. Indeed, folded structures of length 150−250 nm exist for DNA confined in channels with a cross-sectional diameter of ∼100 nm.[7] As shown by the snapshot of an equilibrated Monte Carlo conformation in Figure 2, there

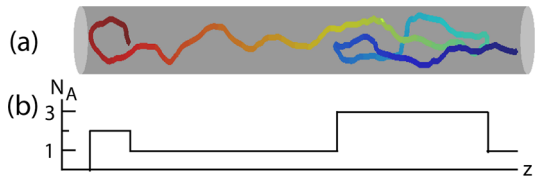


Figure 2. (a) Snapshot of equilibrated Monte Carlo conformation of DNA (w = 5 nm, P = 50 nm, L = 8 μm) inside a D = 50 nm channel. Only a section of the chain is shown. (b) Number of segments profile along the channel.

are two types of back-folded chain configurations. The C-loop results from the formation of a single hairpin and occurs at the end of the chain. An S-loop comes from a pair of hairpins somewhere in the middle. Both types of loops affect $L_{\parallel}/L$, but the effect of the C-loops becomes vanishingly small for very long chains. Accordingly, we primarily focus on S-loop formation, but eventually, we will also include C-loops in our prediction for the stretch. Note that for $1 < D/P < 2$, more than three parallel chain segments inside the channel is improbable, and a one-dimensional random walk is not established.

The reduction in $L_{\parallel}/L$ by S-loop formation is determined by the average contour length L_s stored inside an S-loop and the number of S-loops f_s per unit contour length. In the derivation of $L_{\parallel}/L$, we make two assumptions. (i) We assume that the extension of each chain segment in an S-loop is not affected by the presence of the other two segments. (ii) The contour length of a hairpin chord is assumed to be $\pi D/2$. The relative extension then reads

$$L_{\parallel}^{ds}/L = L_{\parallel}^{d}/L(1 - 2f_s L_s/3 - \pi D f_s/3) \tag{3}$$

where the final term takes into account the fact that the hairpin chords do not contribute to the overall stretch. Expressions for L_s and f_s are derived from a Bragg−Zimm type cooperativity model.[23] Loop formation can be viewed as a dynamic process under constant thermal fluctuation. A pair of hairpins is initially created followed by the growth of an S-loop domain. Nucleation is hence determined by the free energy cost of hairpin formation, which is predominantly bending energy. Domain growth is mainly controlled by excluded volume interaction among parallel chain segments inside the loop.

To implement the cooperativity model, the worm-like chain must be discretized into a sequence of units. Each unit can either be in a deflection or S-loop state. We define the basic unit as a chain segment with a contour length of πD, as being the length stored in two hairpin chords. The smallest S-loop consists hence of a single unit. For a chain with contour length L, the total number of units $N = L/(\pi D)$. The free energy of the chain with a certain sequence of units can be written as

$$F_{conf} = N_s F_s + 2N_d F_u \tag{4}$$

with N_s and N_d being the number of units in the S-loop state and the number of S-loops, respectively. The excess in free energy of a unit in the S-loop state with respect to the deflection state is denoted by F_s (the deflection state has been chosen to be the reference state). Note that each S-loop is sided by two hairpin junctions. The free energy of nucleation of an S-loop is hence $2F_u$, with F_u being the required free energy to create a hairpin.

The cooperativity model can be solved by using the transfer matrix method.[24,25] In the ground state dominance, the thermally averaged free energy of the chain reads

$$F/kT = -N \ln\left[\frac{1 + s + ((1 - s)^2 + 4su)^{1/2}}{2}\right] \tag{5}$$

with the nucleation and propagation parameters $u = \exp(-2F_u/kT)$ and $s = \exp(-F_s/kT)$, respectively. The ensemble-averaged $\langle N_s\rangle$ and $\langle N_d\rangle$ can be derived by differentiation of eq 5 with respect to F_s and $2F_u$, respectively. The average number of S-loops per unit contour length $f_s = \langle N_d\rangle/L$ then takes the form

$$f_s = \frac{u}{2\pi D(1 - u)}\left[\frac{1 + s}{((1 - s)^2 + 4su)^{1/2}} - 1\right] \tag{6}$$

and for the average contour length stored in an S-loop $L_s = \pi D\langle N_s\rangle/\langle N_d\rangle$, we obtain

$$L_s = \pi D\left[1 + \frac{1}{2u}(s - 1 + ((1 - s)^2 + 4su)^{1/2})\right] \tag{7}$$

The cooperativity model requires two input parameters u and s, related to F_u and F_s, respectively. A first principles derivation of these free energies is a difficult task. We expect that F_u is dominated by the elastic bending energy associated with the formation of a hairpin, but there is also an entropic contribution. We propose the simple form

$$F_u/kT = c_1\pi P/D - c_2 \tag{8}$$

where c_1 is a prefactor pertaining to the bending energy $\pi P/D$ and c_2 includes the entropic contribution. The free energy for the growth of the S-loop, F_s is mainly determined by excluded volume interaction. One may assume that the segments inside an S-loop are partially aligned rather than isotropically oriented. For a unit of length πD, the excluded volume is given by $E_x \approx 1/2\ (\pi D)^2 w(D/P)^{1/3}$.[21] The corresponding free energy (per unit volume) takes the form

$$F_s/kT \approx \frac{E_x}{\pi^2 D^3/12} = c_3\frac{6w}{D^{2/3}P^{1/3}} \tag{9}$$

where c_3 is another prefactor of order unity. The feasibility of the cooperativity model for the description of S-loop formation and in particular the expressions for F_u and F_s will be gauged with Monte Carlo simulation.

In the Monte Carlo protocol, the DNA chain is modeled as a string of $(N + 1)$ beads, which are connected by N inextensible bonds of length l_b.[15] The simulation model consists of bending

dx.doi.org/10.1021/mz300323a | ACS Macro Lett. 2012, 1, 1046−1050

energy between adjacent bonds and two interaction terms. The bending rigidity is set to reproduce a persistence length of 50 nm. The interactions are hard-sphere repulsion between DNA beads and hard-wall repulsion between the beads and the wall. If the center of a bead is beyond the channel wall, the potential becomes infinitely large, and the configuration is rejected. The effective channel diameter is hence the real diameter minus the diameter of the bead. For most simulations, the contour length of the DNA chain was fixed at 8 μm. The diameter of the bead was set equal to the bond length l_b, which is equivalent to the chain width w. Accordingly, for chains with a fixed contour length of 8 μm, the number of beads is 1601, 1068, or 801 for w = 5, 7.5, or 10 nm, respectively. In each Monte Carlo cycle, we carried out one crankshaft and one reptation move. The simulation started from a random conformation and usually reached equilibrium after 10^7 cycles. In the production run, we generated 10^{10} cycles and recorded the conformation every other 10^5 cycles. For each conformation, we calculated the extension as the maximum span of the DNA molecule along the channel axis. Finally, the extension was averaged over the ensemble of 10^5 conformations. To track the S-loops, we first generated the segment density profile along the channel for each conformation. This density profile is subsequently examined for domains with a contour length exceeding πD and containing three parallel chain sections (see Figure 2). The number of S-loops and the number of units in the S-loop state are then collected. This procedure is carried out for each conformation in the ensemble, so that we obtain the averages $\langle N_d \rangle$ and $\langle N_s \rangle$. From the latter averages we can calculate, among other quantities, the average number of S-loops per unit contour length $f_s = \langle N_d \rangle / L$ and the average contour length stored in an S-loop $L_s = \pi D \langle N_s \rangle / \langle N_d \rangle$.

The average number of S-loops per unit contour length and the average contour length stored in an S-loop are shown in Figure 3 as a function of channel diameter divided by persistence length. The statistical fluctuation in the results for w = 5 nm is somewhat larger as compared to the ones for w = 7.5 and 10 nm due to the larger number of beads and concomitant longer relaxation time. Note that the cooperativity model is only relevant in the range $1 < D/P < 2$; for larger channel diameters loops can no longer be discerned, and the chain starts to coil. With increasing channel diameter and/or smaller chain width, both f_s and L_s increase monotonously. These effects are most conveniently discussed in terms of F_u and F_s, respectively. The latter free energies can be obtained through u and s and the values of f_s and L_s from the simulation. For this purpose, we have solved eqs 6 and 7 for u and s in terms of f_s and L_s. The analytical solution takes the form

$$u = \frac{\pi^2 D^2 f_s}{(\pi D - L_s)(f_s(L_s + \pi D) - 1)} \tag{10}$$

and

$$s = \frac{(f_s L_s - 1)(L_s - \pi D)}{L_s(f_s(L_s + \pi D) - 1)} \tag{11}$$

The results of F_u and F_s pertaining to the values of f_s and L_s in Figure 3 are shown in Figure 4. As a result of the decrease in F_u and F_s with increasing D and/or decreasing w, the occurrence as well as the average length of the S-loops increases.

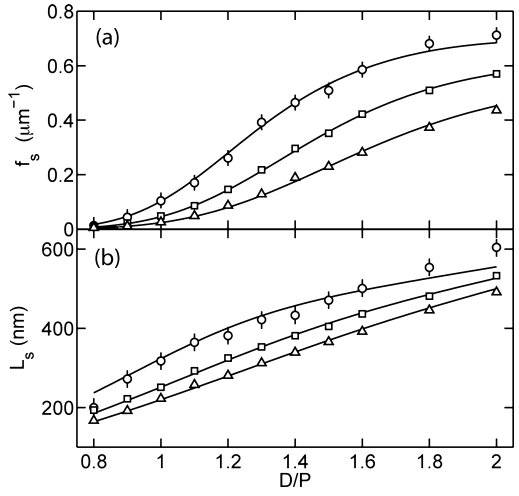


Figure 3. (a) Average number of S-loops per unit contour length vs channel diameter divided by persistence length. The symbols are results obtained from simulations with different chain widths: △, w = 10.0 nm; □, w = 7.5 nm; ○, w = 5.0 nm. For all simulations P = 50 nm and L = 8 μm. The curves represent the cooperativity model. (b) As in panel a, but for the average contour length stored inside an S-loop.

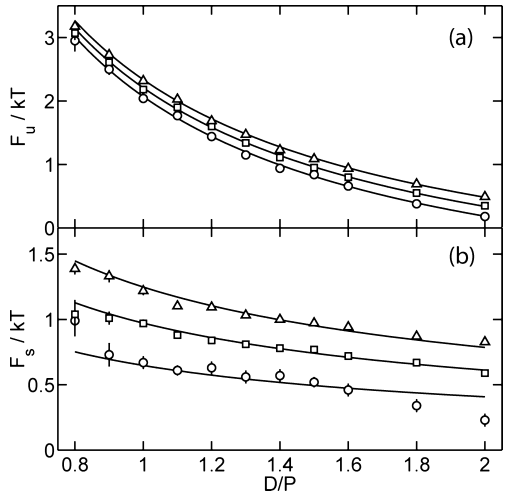


Figure 4. (a) Free energy cost of hairpin formation vs channel diameter divided by persistence length. The symbols are as in Figure 3. The curves represent the prediction based on bending energy and a constant entropy. (b) As in panel a, but for the free energy pertaining to the growth of the S-loop. The curves represent the prediction based on excluded volume.

The free energy of hairpin formation is mainly determined by elastic bending energy, but there is also an entropic contribution. We have fitted eq 8 to F_u obtained from the simulation (see Figure 4a). The values of the fitted constants c_1 and c_2 are collected in Table 1. The prefactor pertaining to the bending energy c_1 does not depend on the width of the chain w beyond the accuracy of the fit. Furthermore, its average value

Table 1. Values of c_1, c_2, and c_3 Resulting from the Fits of Equations 8 and 9 to the Simulation Data in Figure 4

w (nm)	c_1	c_2	c_3
5	1.20 ± 0.03	1.70 ± 0.07	1.08 ± 0.05
7.5	1.19 ± 0.01	1.53 ± 0.03	1.08 ± 0.01
10	1.18 ± 0.02	1.37 ± 0.04	1.04 ± 0.01

dx.doi.org/10.1021/mz300323a | ACS Macro Lett. 2012, 1, 1046−1050

exceeds unity by about 20%, because of hairpin tightening by depletion from the wall.[22] The entropy term c_2 significantly increases with decreasing w, related to reduced volume interaction of the hairpin chord. Given the good agreement between eq 8 and the results for F_u, we have refrained from a more elaborate analysis.[21,22]

The free energy of growth of the S-loop is expected to be dominated by excluded volume interaction among the partially aligned segments within the loop. Accordingly, we have fitted eq 9 to F_s obtained from the simulation. As can be seen in Figure 4b, there is equally good agreement between the prediction for F_s and the simulation results. The adjusted parameter c_3 is constant within the accuracy of the fit, irrespective the value of w (Table 1). This confirms that F_s is proportional to the diameter of the chain, in agreement with an excluded volume mechanism for the growth of the S-loop domains.

The predictions of the cooperativity model, together with eqs 8 and 9 for the free energies of nucleation and propagation, can also directly be compared to the values of f_s and L_s obtained from the simulation. As can be seen in Figure 3, the predicted values based on the values of the parameters in Table 1 are in equally good agreement. Note however that there is a one-to-one correspondence between the data in Figures 3 and 4, so this agreement does not bear more significance.

We now compare the prediction of the cooperativity model for the relative extension in eq 3 with the data obtained from the simulations in Figure 1. Undulation theory without the inclusion of loops in eq 1 describes the extension well for $D/P < 1$. For larger channel diameters, the stretch drops and eventually follows the prediction of the blob model of eq 2 in the range $D/P > P/w$. Inclusion of S-loops in the undulation theory gives a significantly better agreement for $1 < D/P < 2$, but the predicted extensions are still somewhat larger than the simulation results. Note that our model considers an infinitely long chain, but in simulations as well as experiments a DNA molecule has a finite length. The stretch is further reduced due to C-loop formation at both ends. For each end, we assume a Boltzmann distribution in the number of back-folded units of length $\pi D/2$. With the relevant nucleation and propagation parameters $u' = \exp(-F_u/kT)$ and $s' = \exp(-F_s/3kT)$, the relative extension of a C-loop takes the form

$$L_{\parallel}^{dc}/L_c = L_{\parallel}^{d}/L(s'/(1+s')) \tag{12}$$

with $L_{\parallel}^{d}/L$ from eq 1 and its averaged contour length

$$L_c = \pi D u'(1+s')/(2(1-s')(1+u'-s')) \tag{13}$$

Note that the formation of a C-loop is energetically more favorable than an S-loop, due to the reduced free energies for nucleation (a single hairpin) and propagation (two rather than three parallel strands). The complete relative extension follows from the weighed average of the middle and end sections of the chain

$$L_{\parallel}^{dsc}/L = (1-2L_c/L)L_{\parallel}^{ds}/L + (2L_c/L)L_{\parallel}^{dc}/L_c \tag{14}$$

with $L_{\parallel}^{ds}/L$ and $L_{\parallel}^{dc}/L_c$ given by eqs 3 and 12, respectively. Once S and C-loops are included and irrespective of the chain width, fairly good agreement is observed in the relevant range $1 \leq D/P \leq 2$ (see Figure 1). With decreasing contour length, the relative weight of the C-loops in the expression for the stretch increases. Accordingly, the relative extension decreases for shorter chains. This has been verified with additional simulations of chains with contour lengths of 2 and 4 μm. It should be noted however that this applies only to the transition regime. For wider channels with $D > 2P$, the relative extension increases due to progressively weaker confinement.[13]

We have applied a cooperativity model to describe the stretch of a wormlike chain confined in a nanochannel with a diameter of the order of the chain's persistence length. The formation of back-folded hairpins can be considered as the initial stage in the transition from the deflection to the locally coiled, blob regime.[7] Although the range of applicability of the cooperativity model is limited, it is highly relevant from an experimental point of view with wormlike chain persistence lengths and channel cross sections typically in the range of tens to a few hundreds of nanometers. For larger channel diameters $D > 2P$, the cooperativity model no longer provides a satisfactory description of the stretch. Fortunately, as can be gauged from Figure 1, in this range the extension data are already close to the prediction of the blob model. Nevertheless, more work is needed to further bridge the gap.

AUTHOR INFORMATION

Corresponding Author

*E-mail: johanmaarel@gmail.com.

Author Contributions

$^{\parallel}$Contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work is supported by the Singapore-MIT Alliance for Research and Technology, National Science Foundation grant CBET-0852235, and MOE Academic Research Fund MOE2009-T2-2-005. The authors thank CSE in National University of Singapore for providing the computational resources.

REFERENCES

(1) Reisner, W.; Morton, K. J.; Riehn, R.; Wang, Y. M.; Yu, Z.; Rosen, M.; Sturm, J. C.; Chou, S. Y.; Frey, E.; Austin, R. H. *Phys. Rev. Lett.* **2005**, *94*, 196101.

(2) Mannion, J. T.; Reccius, C. H.; Cross, J. D.; Craighead, H. G. *Biophys. J.* **2006**, *90*, 4538.

(3) Reisner, W.; Beech, J. P.; Larsen, N. B.; Flyvbjerg, H.; Kristensen, A.; Tegenfeldt, J. O. *Phys. Rev. Lett.* **2007**, *99*, 058302.

(4) Jo, K.; Dhingra, D. M.; Odijk, T.; de Pablo, J. J.; Graham, M. D.; Runnheim, R.; Forrest, D.; Schwartz, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2673.

(5) Zhang, C.; Zhang, F.; van Kan, J. A.; van der Maarel, J. R. C. *J. Chem. Phys.* **2008**, *128*, 225109.

(6) Persson, F.; Utko, P.; Reisner, W.; Larsen, N. B.; Kristensen, A. *Nano Lett.* **2009**, *9*, 1382.

(7) Su, T.; Das, S. K.; Xiao, M.; Purohit, P. K. *PLoS One* **2011**, *6*, e16890.

(8) Dai, L.; van der Maarel, J. R. C.; Doyle, P. S. *ACS Macro Lett.* **2012**, *1*, 732.

(9) Zhang, C.; Shao, P. G.; van Kan, J. A.; van der Maarel, J. R. C. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 16651.

(10) Jones, J. J.; van der Maarel, J. R. C.; Doyle, P. S. *Nano Lett.* **2011**, *11*, 5047.

(11) Zhang, C.; Gong, Z.; Guttula, D.; Malar, P. P.; van Kan, J. A.; Doyle, P. S.; van der Maarel, J. R. C. *J. Phys. Chem. B* **2012**, *116*, 3031.

(12) Cifra, P.; Benkova, Z.; Bleha, T. *J. Phys. Chem. B* **2008**, *112*, 1367.

(13) Wang, Y.; Tree, D. R.; Dorfman, K. D. *Macromolecules* **2011**, *44*, 6594.

(14) Cifra, P. *J. Chem. Phys.* **2012**, *136*, 024902.
(15) Dai, L.; Jones, J. J.; van der Maarel, J. R. C.; Doyle, P. S. *Soft Matter* **2012**, *8*, 2972.
(16) Williams, M. C. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 11125.
(17) Jun, S.; Mulder, B. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 12388.
(18) Odijk, T. *Macromolecules* **1983**, *16*, 1340.
(19) Daoud, M.; de Gennes, P. G. *J. Phys. (Paris)* **1977**, *38*, 85.
(20) Yang, Y. Z.; Burkhardt, T. W.; Gompper, G. *Phys. Rev. E* **2007**, *76*, 011804.
(21) Odijk, T. *Phys. Rev. E* **2008**, *77*, 060901(R).
(22) Odijk, T. *J. Chem. Phys.* **2006**, *125*, 204904.
(23) Zimm, B. H.; Bragg, J. K. *J. Chem. Phys.* **1959**, *31*, 526.
(24) Kramers, H. A.; Wannier, G. H. *Phys. Rev.* **1941**, *60*, 252.
(25) Grosberg, A. Y.; Khokhlov, A. R. *Statistical Physics of Macromolecules*; AIP Press: New York, 1994.