Extension of DNA in a Nanochannel is a Rod-to-Coil Transition

Douglas R. Tree,[1] Yanwei Wang,[2] and Kevin D. Dorfman[1,*]

[1]*Department of Chemical Engineering and Materials Science, University of Minnesota — Twin Cities, 421 Washington Ave. SE, Minneapolis, MN 55455*

[2]*Department of Polymer Science and Engineering, College of Chemistry, Chemical Engineering and Materials Science, Soochow University, 199 Ren-ai Road, Suzhou, 215123, P.R. China*

(Dated: March 26, 2013)

Abstract

DNA confinement in nanochannels is emerging as an important tool for genomics and an excellent platform for testing the theories of confined wormlike polymers. Using cutting-edge, large scale Monte Carlo simulations of asymptotically long wormlike chains, we show that, in analogy to the rod-to-coil transition for free wormlike polymers, there exists a universal, Gauss-de Gennes regime that connects the classic Odijk and de Gennes regimes of channel-confined chains. For DNA in a nanochannel, this Gauss-de Gennes regime spans practically the entire experimentally relevant range of channel sizes, including the nanochannels used in an incipient genome mapping technology.

PACS numbers: 87.15.ak, 87.15.hj, 87.14.gk

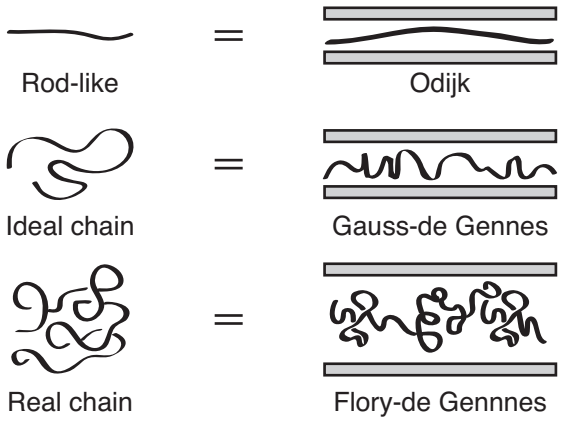

FIG. 1. Illustration of the analogy between free solution and confined configurations of a wormlike chain. The classical theories renormalize the chain into a series of subchains, where these subchains are either rod-like (Odijk) or excluded volume blobs (de Gennes). We demonstrate here the existence of a universal Gauss-de Gennes regime in confinement that connects the (rod-like) Odijk and (excluded volume) de Gennes regime. For clarity, we refer to the classic de Gennes regime as the "Flory-de Gennes" regime to highlight its excluded volume nature.

When a wormlike polymer such as DNA is confined in a long channel whose width is smaller than the polymer's free solution radius of gyration, steric interactions with the walls cause the polymer to extend along the channel axis. The classical theories describing this phenomenon, sketched in Fig. 1, were described by Odijk [1] and de Gennes [2], respectively, over 30 years ago. However, these theories are only valid in the impractical cases of very strong ($D \ll l_p$) or very weak ($D \gg l_p$) confinement, respectively, where D is the channel size and l_p is the persistence length of the chain. In this Letter, we establish that the relevant intermediate regime for DNA extension in a nanochannel is a universal de Gennes-like regime with ideal blobs. We arrived at this conclusion by recognizing the connection with the rod-to-coil transition for free wormlike polymers [3–5] illustrated in Fig. 1. In addition to describing the experimentally relevant phenomena for DNA, this connection leads to a complete description of the universal regimes for all long channel-confined wormlike chains that we validated using large-scale Pruned-Enriched Rosenbluth Method (PERM) simulations [6, 7].

To understand the analogy in Fig. 1, let us first recall the regimes of chain conformations in free solution. Three regimes characterize the normalized end-to-end distance of a wormlike

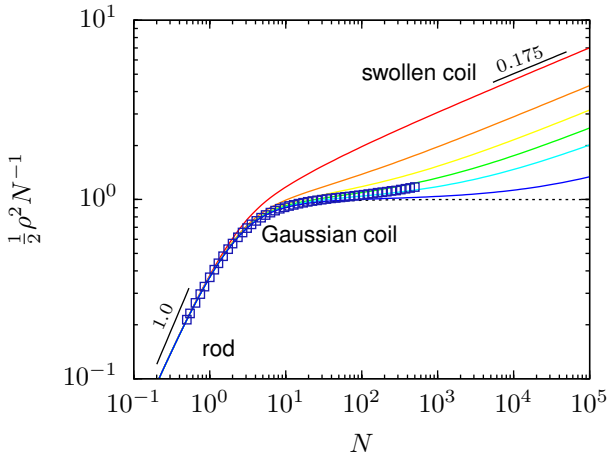

FIG. 2. (color online) The normalized mean-square end-to-end distance of a wormlike chain in free solution as a function of normalized chain length from Eq. 1 ($\epsilon = 0$, - -), renormalization group theory [4] with (descending from top to bottom in the figure) $\epsilon = 2$, 0.5, 0.2, 0.1 (DNA), 0.05, 0.01, and PERM simulations for $\epsilon = 0.05$ ($\square$, $L_{\max}/w = 2 \times 10^4$). The most flexible chain corresponds to an upper bound $\epsilon = 2$, where the Kuhn length of the chain equals its width [12].

chain, $\rho \equiv \langle R^2 \rangle^{1/2}/l_p$. This distance depends on two dimensionless numbers: $N \equiv L/l_p$, the number of persistence lengths in a chain of length L, and an anisotropy parameter $\epsilon \equiv w/l_p$, which measures the relative strength of the excluded volume interactions (quantified by the effective chain width w) to the bending energy. In the limit of negligible excluded volume interactions [8], $z \equiv \epsilon N^{1/2} \ll 1$, the Benoit-Doty equation for a continuous wormlike chain gives [5]

$$\rho^2/2 = N - 1 + \exp(-N) \tag{1}$$

This model predicts a stiff chain with $\rho \sim N$ for $N \lesssim 1$ and Gaussian statistics with $\rho \sim N^{1/2}$ for $1 \lesssim N \lesssim \epsilon^{-2}$. For a sufficiently long chain $N \gtrsim \epsilon^{-2}$, excluded volume interactions are important and $\rho \sim \epsilon^{2\nu-1} N^{\nu}$, where $\nu = 0.5877$ is the modern value of the Flory parameter [9]. As evident in Fig. 2, the scaling in the Gaussian regime is not exactly that for an ideal chain due to finite excluded volume effects. Moreover, weakly anisotropic chains — such as DNA, which is only a moderately stiff biopolymer ($\epsilon \approx 0.1$) [10] — have a very narrow Gaussian regime. However, in the limit $\epsilon \to 0$, the Gaussian regime spans an infinite amount of chain length and is thus a universal regime. Accordingly, many biopolymers [11] are stiff enough to exhibit broad Gaussian regimes.

We will demonstrate, via simulations, that three similar regimes characterize the confinement free energy of an asymptotically long wormlike chain confined in channel. Following Odijk [1] and de Gennes [2, 13], the chain is renormalized into N/g units containing g persistence lengths per unit. This in turn implies that the chain properties (confinement free energy and extension) are extensive, as has been shown many times [1, 2, 13, 14] for an infinite chain in a quasi-1D geometry.

The Odijk regime [1] in Fig. 1 corresponds to rod-like behavior over the length scale D. For channel sizes $\delta \equiv D/l_p \lesssim 1$ [1], the stiff chain projects a distance of $\lambda \sim (D^2 l_p)^{1/3}$ before deflecting off of the walls. This makes the number of persistence lengths in the correlation volume

$$g = \lambda/l_p \sim \delta^{2/3} \tag{2}$$

Assuming an energy of $k_B T$ per independent segment [1] gives the dimensionless confinement free energy,

$$\mathcal{F} \equiv \frac{\Delta F_c}{N k_B T} \sim 1/g \sim \delta^{-2/3} \tag{3}$$

The extension is given by the projection of the deflection segment length onto the channel axis $X = (N/g)\lambda \cos\theta$ [1] which simplifies to

$$\langle X/L \rangle = 1 - \alpha\, \delta^{2/3} \tag{4}$$

where the prefactor $\alpha = 0.18274$ for a square nanochannel is given by high resolution simulations [15]. Analogous to the rod-like behavior in free solution, the thermodynamics of the Odijk regime is independent of the width of the chain.

Continuing with the analogy, the de Gennes regime [2] in Fig. 1 corresponds to real chain statistics, which leads us to call it the "Flory-de Gennes" regime. Here, as was the case for real chains in free solution, we need to account for the finite chain width. To do so, we use the concept of a "blob" to denote a section of the chain with g persistence lengths that has a correlation length equal to the channel size D. Recalling that the Flory radius for a chain in a good solvent is $R_F/l_p \approx \epsilon^{2\nu-1} N^{\nu}$ [8], the blobs have the size

$$\delta \approx \epsilon^{2\nu-1} g^{\nu} \tag{5}$$

With the assumption that the free energy scales as $k_B T$ per blob [2], we have $\mathcal{F} \sim 1/g$, or

$$\mathcal{F} \sim \left(\delta \epsilon^{1-2\nu}\right)^{-1/\nu} \tag{6}$$

Following the same reasoning, the extension $\langle X \rangle$ is also extensive in the number of blobs, $\langle X \rangle \cong (N/g)D$. Substituting Eq. (5) in the latter gives the scaling

$$\langle X/L \rangle \cong \delta^{1-1/\nu} \epsilon^{2-1/\nu} \tag{7}$$

Since the Flory-de Gennes regime corresponds to the onset of excluded volume interactions [16], we would expect this regime to start when the excluded volume parameter for a blob reaches

$$z_{\text{blob}} \equiv \epsilon g^{1/2} \approx 1 \tag{8}$$

We thus find that $g \gtrsim \epsilon^{-2}$ corresponds to the Flory-de Gennes regime. Recall that the excluded volume scaling in free solution begins when $N \gtrsim \epsilon^{-2}$. We thus infer that g in confinement is the analogue of N in free solution. Additionally we note that by combining Eqs. (5) and (8), we can find the boundary of the Flory-de Gennes regime limit in terms of the channel size, $\delta \gtrsim \epsilon^{-1}$, which proves more useful since the channel size is an experimental observable.

For intermediate channel sizes $1 \lesssim \delta \lesssim \epsilon^{-1}$, the g persistence lengths inside D^3 exhibit approximately Gaussian statistics. The derivation of the confinement free energy follows that for the Flory-de Gennes regime with $\nu = 1/2$, leading

$$\mathcal{F} \sim \delta^{-2} \tag{9}$$

Since this regime consists of blobs with Gaussian statistics, we refer to it as the "Gauss-de Gennes" regime. This free energy scaling is the same as that of a channel-confined phantom chain originally derived by Cassassa [2, 14]. As is the case in free solution, the scaling of $\mathcal{F}$ for chains with a finite value of ϵ will not be exactly equal to Eq. (9). This arises from the weakness (rather than absence) of excluded volume at the persistence length scale.

In the Gauss-de Gennes regime, the intra-polymer correlations are screened at the channel wall [13, 17], which gives $g \sim \delta^2$ persistence lengths per correlation length, D. Since the extended chain consists of N/g such correlation lengths, the corresponding fractional extension is

$$\langle X \rangle / L \sim \delta^{-1} \tag{10}$$

The latter scaling has been observed in a number of previous simulations (see [10, 18–20] and supporting Fig. S6), but its origin and universal nature (or lack thereof) have been

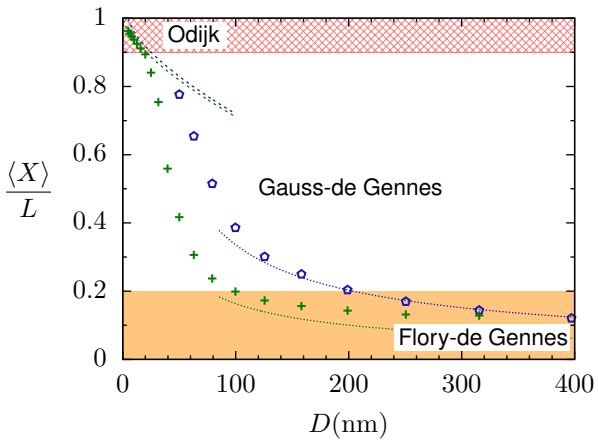

FIG. 3. (color online) Comparison of the fractional extension of the chain predicted by Odijk [1] and de Gennes [2] and simulations of an asymptotically long DNA chain ($\pentagon$, $l_p = 50$ nm, $w = 5$ nm, $\epsilon = 0.1$) using the Pruned-Enriched Rosenbluth Method (PERM). The extent of the Gauss-de Gennes regime increases for more filamentous chains (+, $l_p = 50$ nm, $w = 0.5$ nm, $\epsilon = 10^{-2}$). The shading corresponds to the regimes for $\epsilon = 0.1$.

elusive until now because DNA is not an especially stiff biopolymer. Given this fact one may be tempted to dismiss the regime as unimportant, but consider the case of DNA in a high ionic strength buffer ($\epsilon = 0.1$) which is highlighted in Fig. 3. Although the Gauss-de Gennes regime spans less than a decade in dimensionless channel size, these sizes encompass the typical channels used in experiments [17, 21–24]. Morover, the Flory-de Gennes regime corresponds to $\lesssim 20\%$ extension and the Odijk regime corresponds to $\gtrsim 90\%$ extension, leaving the Gauss-de Gennes regime to span a significant portion of the practically relevant range of fractional extensions for genomic mapping. However, this regime is not in principle limited to a small range of channel sizes. For stiff enough chains, the range of applicable channel sizes $1 \lesssim \delta \lesssim \epsilon^{-1}$ will span many decades, showing the existence of a universal regime.

We have tested this scaling theory in square channels using Pruned-Enriched Rosenbluth Method (PERM) simulations of asymptotically long chains that are long enough to suppress any end effects. PERM is a biased chain-growth Monte Carlo algorithm originally introduced for lattice chains by Grassberger [6]. In the algorithm, "tours" of chains are grown and the Rosenbluth weight of the chain is controlled by selective pruning and enrichment, thus

overcoming the attrition problem for the Rosenbluth-Rosenbluth chain growth algorithm [25] for long self-avoiding chains. Choosing efficient parameters for executing the original PERM algorithm is somewhat of an art, and we have followed a parameterless version by Prellberg [7] that simplified the calculation considerably. Our optimized implementation of PERM (see supporting information) allowed us to sample long chain lengths (typically 2×10^4 touching beads of size w) while spanning four decades in the dimensionless channel size δ and three decades in the anisotropy ϵ. For DNA with $w = 5$ nm, our data typically correspond to contour lengths of 100 μm, a full order of magnitude longer than traditional Markov chain Monte Carlo techniques [10, 17–20, 33]. This combination of asymptotically long chains, a thorough exploration of the (δ,ϵ) phase space (see supporting information), and the large range of confinement free energies allowed us to draw meaningful conclusions about universality. In addition to providing the chain conformational properties shown in Fig. 2 and Fig. 3, PERM can provide thermodynamic properties like the confinement free energy.

To clearly see the analogy with the rod-to-coil transition in free solution, we also need the equivalent of Eq. (1) for the confinement free energy of an ideal wormlike chain ($\epsilon = 0$). To a good approximation, the confinement free energy of a chain in a channel is equal to twice the confinement free energy of a chain confined to a slit [14, 26]

$$\mathcal{F} = (2/3)\pi^2\delta^{-2} \tag{11}$$

Additionally, extensive computational work on strongly confined wormlike chains has yielded an accurate prefactor to the Odijk expression for square channels [15]

$$\mathcal{F} = 2.2072\,\delta^{-2/3} \tag{12}$$

Following Chen and Sullivan [27] we propose an interpolation formula of the form

$$\mathcal{F} = \frac{(2/3)\pi^2\delta^{-2}}{(5.147\delta^{-2} + 3.343\delta^{-1} + 1)^{2/3}} \tag{13}$$

Taking the limit $\delta \to \infty$ yields Eq. (11) and $\delta \to 0$ matches Eq. (12). The remaining constant for the δ^{-1} term is used to fit the shape of the crossover region obtained from PERM simulations in the absence of excluded volume (see supporting information).

The similarity between Fig. 2 and Fig. 4 confirms the analogy between bulk and confinement, and the plateau in Fig. 4 validates the presence of a Gauss-de Gennes regime

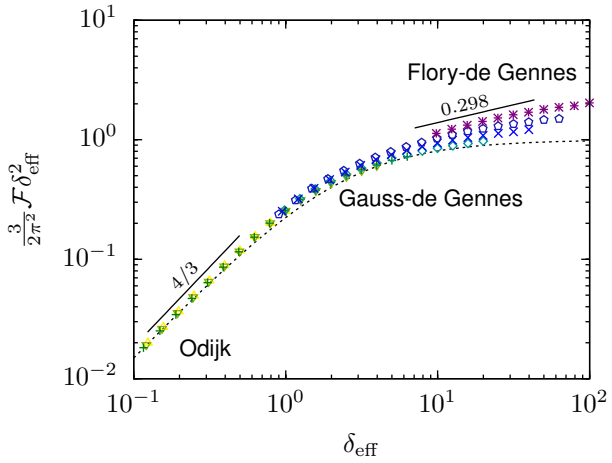

FIG. 4. (color online) The normalized free energy of confinement as a function of normalized channel width, $\delta_{\rm eff} = (D - w)/l_p$ from Eq. 13 ($\epsilon = 0$, - -) and PERM simulations for $\epsilon = 2 \times 10^{-3}$ ($\triangle$), 5×10^{-3} ($\bigtriangledown$), 0.01 ($+$), 0.02 ($\diamond$), 0.05 ($\times$), 0.1 ($\pentagon$, DNA), and 0.2 ($\ast$).

in confinement that connects the Odijk and Flory-de Gennes regimes. Compared to free solution, the Gauss-de Gennes regime in confinement is less prominent than the Gaussian regime in free solution because (i) the upper bound is lower in confinement (ϵ^{-1} versus ϵ^{-2}) and (ii) it is challenging to simulate extremely long chains with small ϵ at very high spatial resolution. Nevertheless, Fig. 4 clearly demonstrates the three regimes, including the scaling exponent predicted by Eq. (9).

The close agreement here between the scaling theory and simulations has parallels with DNA confined in a sphere [28], but calls into question previous theories for the thermodynamics of a channel-confined chain between the Odijk and the Flory-de Gennes regimes. Most treatments apply Flory theory for a confined chain [16, 29–31] notwithstanding the fact that the accuracy of Flory theory predictions in free solution relies on a serendipitous cancellation of errors that are not *a priori* applicable in confinement. For example, the scaling $\mathcal{F} \sim \delta^{-4/3}$ predicted by Flory theory [22] for the "extended de Gennes" regime [10, 16, 29, 31] is not evident in our simulations. Other theories have attempted to incorporate backfolding of the chain to explain the transition between the Odijk and Flory-de Gennes regime [17, 19, 20, 29, 32]. The analogy between free solution and confinement makes the role of backfolding clear — it is simply the transition from rod-like to ideal statistics in the correlation volume.

Our results provide not only a complete description of the universal regimes of any long,

channel-confined wormlike chain, but also have practical implications for genomic mapping in nanochannels [23, 24]. Our simulations predict that the Odijk regime is valid for an effective channel size $\delta_{\text{eff}} \equiv (D-w)/l_p \leq 0.3$ and the Flory-de Gennes regime begins at $\delta_{\text{eff}} \geq (2\epsilon)^{-1}$ (see supporting information). For DNA in a nanochannel, the Odijk extension [1, 15] applies for channels smaller than 20 nm, whereas the Flory-de Gennes extension [2] only starts to apply for channels larger than around 200 nm, where stretching is insubstantial. Since almost all experiments [22] and the commercial nanochannel technology [23] operate between these limiting cases, it is unsurprising that the experimental data are not described by the Odijk or de Gennes theories. Additionally, the Gauss-de Gennes regime certainly has implications for dynamics, which have recently been shown to be very sensitive to the anisotropy ϵ over similar ranges of extension [33]. Future device design, as well as fundamental work, will need to account for the nature of the rod-to-coil transition of the subchains comprising nanoconfined polymers.

We thank Drew Gustafson and Igal Szleifer (Northwestern University) for useful discussions, and Frank S. Bates and Timothy P. Lodge for comments on earlier versions of this manuscript. This work was supported by the NIH (R01-HG005216) and was carried out in part using computing resources at the University of Minnesota Supercomputing Institute.

* dorfman@umn.edu

[1] T. Odijk, Macromolecules **16**, 1340 (1983).

[2] M. Daoud and P.-G. de Gennes, J. Phys. France **38**, 85 (1977).

[3] J. Moon and H. Nakanishi, Phys. Rev. A **44**, 6427 (1991); J. W. Halley, D. Atkatz, and H. Nakanishi, J. Phys. A **23**, 3297 (1990); H.-P. Hsu, W. Paul, and K. Binder, Europhys. Lett. **92**, 28003 (2010).

[4] Z. Y. Chen and J. Noolandi, J. Chem. Phys. **96**, 1540 (1992).

[5] H. Benoit and P. Doty, J. Phys. Chem. **57**, 958 (1953).

[6] P. Grassberger, Phys. Rev. E **56**, 3682 (1997).

[7] T. Prellberg and J. Krawczyk, Phys. Rev. Lett. **92**, 120602 (2004).

[8] M. Rubinstein and R. Colby, *Polymer Physics* (Oxford University Press, 2003).

[9] B. Li, N. Madras, and A. D. Sokal, J. Stat. Phys. **80**, 661 (1995).

[10] Y. Wang, D. R. Tree, and K. D. Dorfman, Macromolecules **44**, 6594 (2011).

[11] Z. Dogic, J. Zhang, A. W. C. Lau, H. Aranda-Espinoza, P. Dalhaimer, D. E. Discher, P. A. Janmey, R. D. Kamien, T. C. Lubensky, and A. G. Yodh, Phys. Rev. Lett. **92**, 125503 (2004).

[12] A. Y. Grosberg and A. R. Khokhlov, *Statistical Physics of Macromolecules* (American Institute of Physics, 1994).

[13] P.-G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, NY, 1979).

[14] E. F. Casassa, Macromolecules **9**, 182 (1976).

[15] T. W. Burkhardt, Y. Yang, and G. Gompper, Phys. Rev. E **82**, 041801 (2010).

[16] F. Brochard-Wyart, T. Tanaka, N. Borghi, and P. G. de Gennes, Langmuir **21**, 4144 (2005).

[17] E. Werner, F. Persson, F. Westerlund, J. O. Tegenfeldt, and B. Mehlig, Phys. Rev. E **86**, 041802 (2012).

[18] P. Cifra, Z. Benková, and T. Bleha, J. Phys. Chem. B **113**, 1843 (2009); P. Cifra, J. Chem. Phys. **131**, 224903 (2009).

[19] L. Dai, S. Y. Ng, P. S. Doyle, and J. R. C. van der Maarel, ACS Macro Letters **1**, 1046 (2012).

[20] P. Cifra, J. Chem. Phys. **136**, 024902 (2012).

[21] J. O. Tegenfeldt, C. Prinz, H. Cao, S. Y. Chou, W. W. Reisner, R. Riehn, Y. M. Wang, E. D. Cox, J. C. Sturm, P. Silberzan, and R. H. Austin, Proc. Natl. Acad. Sci. USA **101**, 10979 (2004); W. Reisner, K. J. Morton, R. Riehn, Y. M. Wang, Z. Yu, M. Rosen, J. C. Sturm, S. Y. Chou, E. Frey, and R. H. Austin, Phys. Rev. Lett. **94**, 196101 (2005).

[22] W. Reisner, J. N. Pedersen, and R. H. Austin, Rep. Prog. Phys. **75**, 106601 (2012).

[23] E. T. Lam, A. Hastie, C. Lin, D. Ehrlich, S. K. Das, M. D. Austin, P. Deshpande, H. Cao, N. Nagarajan, M. Xiao, and P.-Y. Kwok, Nat. Biotech. **30**, 771 (2012).

[24] K. D. Dorfman, S. B. King, D. W. Olson, J. D. P. Thomas, and D. R. Tree, Chem. Rev. (in press, doi:10.1021/cr3002142).

[25] M. N. Rosenbluth and A. W. Rosenbluth, J. Chem. Phys. **23**, 356 (1955).

[26] Y. Wang, G. H. Peters, F. Y. Hansen, and O. Hassager, J. Chem. Phys. **128**, 124904 (2008).

[27] J. Z. Y. Chen and D. E. Sullivan, Macromolecules **39**, 7769 (2006).

[28] T. Sakaue, Macromolecules **40**, 5206 (2007).

[29] T. Odijk, Phys. Rev. E **77**, 060901(R) (2008).

[30] S. Jun, D. Thirumalai, and B.-Y. Ha, Phys. Rev. Lett. **101**, 138101 (2008).

[31] F. Brochard-Wyart and E. Raphael, Macromolecules **23**, 2276 (1990).

[32] T. Su, S. K. Das, M. Xiao, and P. K. Purohit, PLoS One **6**, e16890 (2011).

[33] D. R. Tree, Y. Wang, and K. D. Dorfman, Phys. Rev. Lett. **108**, 228105 (2012).

Supplemental Information for "Extension of DNA in a Nanochannel is a Rod-to-Coil Transition"

Douglas R. Tree,[1] Yanwei Wang,[2] and Kevin D. Dorfman[1,*]

[1] *Department of Chemical Engineering and Materials Science, University of Minnesota — Twin Cities, 421 Washington Ave. SE, Minneapolis, MN 55455*

[2] *Jiangsu Key Laboratory of Advanced Functional Polymer Design and Application, Department of Polymer Science and Engineering, College of Chemistry, Chemical Engineering and Materials Science, Soochow University, 199 Ren-ai Road, Suzhou, 215123, P.R. China*

(Dated: March 26, 2013)

* dorfman@umn.edu

I. SIMULATION METHOD

In our implementation we have chosen an *off-lattice*, touching-bead model [1] in which we implement a discrete wormlike chain potential

$$\beta U_{\text{bend}} = \kappa \sum_{j=1}^{N_b-2} (1 - \mathbf{u}_{j+1} \cdot \mathbf{u}_j) \tag{S1}$$

where $\beta = (k_B T)^{-1}$ is the inverse absolute temperature and $\mathbf{u}$ is the bond vector connecting bead j to bead $j+1$. The parameters in this equation are the number of beads (N_b), the bending potential constant (κ), and the bead size (w). In this model, choosing w is especially consequential because it (1) fixes the bond length, (2) sets the excluded volume to the value: $v \cong {l_p}^2 w$ (which is athermal because of hard beads and walls) and (3) imposes the length scale below which self-interactions are disallowed.

Simulations are run in either free solution or in an infinintely long square channel of size D. When confined, the channel walls are hard and are defined by the potential

$$\beta U_{\text{wall}} = \begin{cases} \infty & \max(|r_x|, |r_y|) \geq (D-w)/2 \\ 0 & \text{otherwise} \end{cases} \tag{S2}$$

Here $|r_x|$ and $|r_y|$ are the absolute values of the chain position in the x and y directions respectively, D is the channel width and the z-axis is left open.

The bending potential is responsible for setting the value of the persistence length, which is defined as

$$\langle \mathbf{u}_{j+1} \cdot \mathbf{u}_j \rangle = \exp(-jw/l_p) \tag{S3}$$

Fortuitously, the probability density function for a given bond angle defined as $\mathbf{u}_{j+1} \cdot \mathbf{u}_j = \cos\theta_j$ can be solved analytically [2] and is given by

$$P(\theta_j) = \frac{\kappa \exp\left[-\kappa(1-\cos\theta_j)\right]}{1-\exp(-2\kappa)}. \tag{S4}$$

For illustrative purposes Fig. S1 shows several plots of Eq. (S4) for various values of κ.

The closed form solution in Eq. (S4) is beneficial for two reasons. First, it allowed us to code PERM more efficiently, as the simulation requires generating many bond vectors according to this distribution. Second, it allowed us to obtain the persistence length for given values of κ and w. This can be done by calculating Flory's characteristic ratio [3, 4],

$$C_\infty = \frac{b}{w} = \frac{1 + \langle \cos\theta_j \rangle}{1 - \langle \cos\theta_j \rangle} \tag{S5}$$

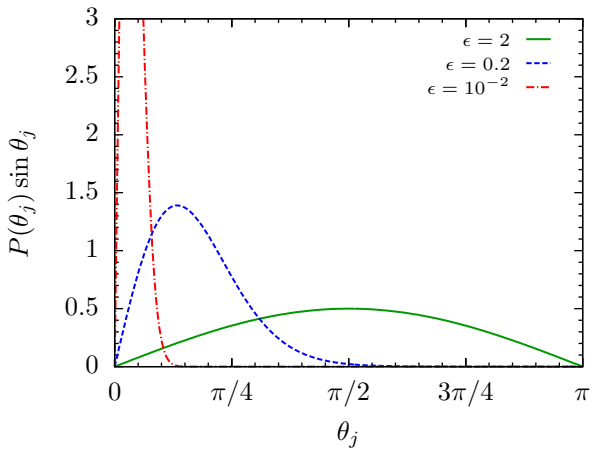

FIG. S1. Various curves of Eq. (S4) for three values of κ: 10^{-5} (solid green), 5.499 (blue dashed), 100.5 (red dash-dot). These values illustrate the bond angle probability density function when chains are very stiff ($\kappa = 100.5$) or very flexible ($\kappa = 10^{-5}$, practically a freely-joined chain).

which gives the Kuhn length $b = 2l_p$. Evaluating this expression leads to [2]

$$l_p = \left(\frac{w}{2}\right) \frac{\kappa - 1 + \kappa \coth \kappa}{\kappa + 1 - \kappa \coth \kappa} \tag{S6}$$

which simplifies to the well-known expression for the persistence length,

$$\frac{l_p}{w} = \kappa - 1/2 \tag{S7}$$

when $\kappa \gg 1$.

In addition to using an analytical distribution to generate the bond vectors, we further accelerated the simulations by taking advantage of the fact that the confinement free energy is extensive in N for sufficiently long chains. PERM relies on an estimate of the partition function in order to bias the sampling. In the "blind" version of the PERM algorithm [5], the estimate is progressively built at the beginning of each simulation run, a process which dominates the simulation time. To speed up the convergence of the estimate, we first ran a blind simulation for short chains and linearly extrapolated the partition estimate to get a good initial guess for the partition function for the longer chains. This guess is only used to set the biasing in PERM, and does not affect the final partition function estimate. We were able to run "non-blind" simulations [5] for the longest chains, a technique which reduced the simulation time considerably.

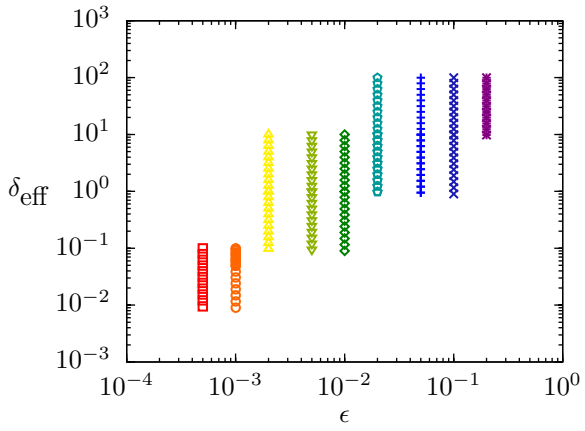

FIG. S2. Range of values for the effective channel size available to the chain, $\delta_{\rm eff} \equiv (D-w)/l_p$, and the chain anisotropy, $\epsilon \equiv w/l_p$, used in the PERM simulations. The phase space explored here is orders of magnitude larger than previous studies. The umbrella-like overlap between different values of ϵ allows us to produce the universal free energy curve in Fig. 4 of the main text.

In our implementation, we employed a master/slave parallel algorithm *without* Markovian anticipation [6] on a DELL linux cluster. (Markovian anticipation is not trivial to implement with an off-lattice model.) The hardcore excluded volume interaction calculations took advantage of neighbor lists, and data analysis was done on the fly since recording each tour's configuration is prohibitively expensive. Most of the PERM data (unless indicated otherwise) was taken for chains with 2×10^4 beads in five batches of 10^4 tours for error estimation; in all cases the error of the data shown is smaller than the symbol size. Our simulations spanned the wide range of channel sizes and chain anisotropies shown in Fig. S2, corresponding to four decades in the dimensionless channel size $\delta \equiv D/l_p$ and three decades in the anisotropy $\epsilon \equiv w/l_p$. In contrast, most previous work, focusing on DNA in nanochannels [7–13], spans less than two decades in channel size (say, 10 nm to 500 nm) and uses a single value of the anistropy ϵ corresponding to DNA or, at most, a change in ϵ by a factor of around 5. Moreover, these previous simulations typically use around 10^3 beads to represent the DNA.

The free energy of a given chain was obtained from PERM's estimate of the canonical partition function [14] which has an unconfined, ideal wormlike chain standard state ($\epsilon = 0$).

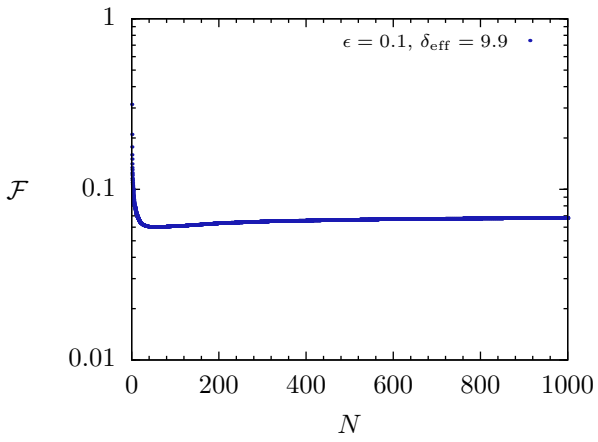

FIG. S3. Example of the convergence of the confinement free energy, $\mathcal{F} \equiv \beta \Delta F_c N^{-1}$ as a function of the size of the chain, $N \equiv L/l_p$. The data appearing in Fig. 4 of the main text correspond to the plateau region for a given simulation for a channel size δ and chain anisotropy ϵ.

Thus, the confinement free energy

$$\Delta F_c \equiv F_{\rm confined} - F_{\rm bulk} \tag{S8}$$

requires simulations of both confined and free solution chains for chains with non-zero excluded volume. In order to assure that the free energy calculations were accurate, we verified that all of the free energy calculations shown in the main text did in fact reach the asymptotic limit where $\mathcal{F} \sim N^0$ where $N \equiv L/l_p$ is the number of persistence lengths of the chain and $\mathcal{F} \equiv \beta \Delta F_c N^{-1}$ is the dimensionless free energy. Figure S3 shows an example set of data ($\delta_{\rm eff} = 9.9$ and $\epsilon = 0.1$) which corresponds to DNA in a high ionic strength buffer in a 500 nm channel. As we can see, the simulation spans 1000 persistence lengths, corresponding to a contour length of 50 μm, or about 150 kbp (slightly smaller than T4 DNA). The plateau for $\mathcal{F}$ in Fig. S3 is the long-chain asymptotic value for this particular combination of $\delta_{\rm eff}$ and ϵ appearing in the main text. We constructed similar plots for every combination of $\delta_{\rm eff}$ and ϵ and included only those simulations which reached a plateau value.

II. CONFINEMENT FREE ENERGY OF AN IDEAL SEMIFLEXIBLE CHAIN

Following Chen and Sullivan [15] we propose an interpolation formula of the form

$$\mathcal{F} = \frac{\frac{2}{3}\pi^2 \delta^{-2}}{(C_2 \delta^{-2} + C_1 \delta^{-1} + 1)^{2/3}} \tag{S9}$$

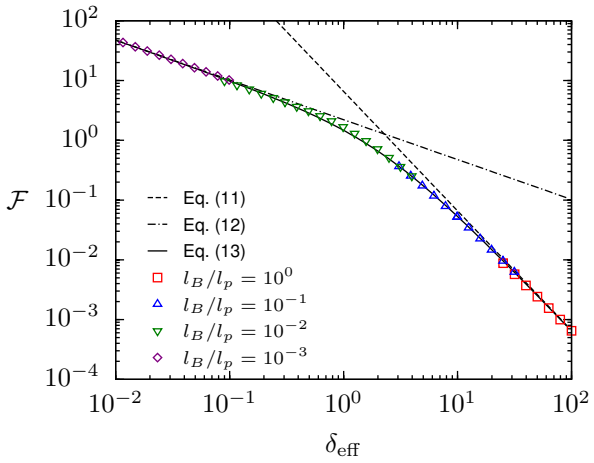

FIG. S4. Free energy of a confined semiflexible chain without excluded volume. Simulation data is taken for chains of different persistence length where the bond length is represented as l_B. For each simulated chain $\epsilon = 0$, except for $l_B/l_p = 10^{-3}$, which has a small but negligible $\epsilon = 10^{-3}$. The data are used to validate Eq. (S9) and calculate C_1 by a least-squares regression.

The choice $C_2 = 5.147$ matches previous calculations [16] for the Odijk regime. The remaining constant is used to fit the shape of the crossover region from PERM simulations in the absence of any excluded volume. A least squares fit yields $C_1 = 3.343$ and the interpolation shown in Fig. S4.

III. NUMERICAL PREFACTORS FOR THE ODIJK AND FLORY-DE GENNES EXTENSION

In several instances in the main text, PERM results for the extensions are compared to Odijk and de Gennes theories with exact prefactors. The Odijk regime curve corresponds to

$$\langle X/L \rangle = 1 - 0.18274 \left(\frac{D-w}{l_p} \right)^{2/3} \tag{S10}$$

which is the prediction of the Odijk theory [17] using the prefactor computed by Burkhardt et al. [16] for a square nanochannel. The quantity $D - w$ is the effective width of the nanochannel available to the chain.

For the Flory-de Gennes regime, the extension is

$$\langle X/L \rangle = (1.033 \pm 0.005) \left(\frac{D-w}{l_p} \right)^{1-1/\nu} \left(\frac{w}{l_p} \right)^{2-1/\nu} \tag{S11}$$

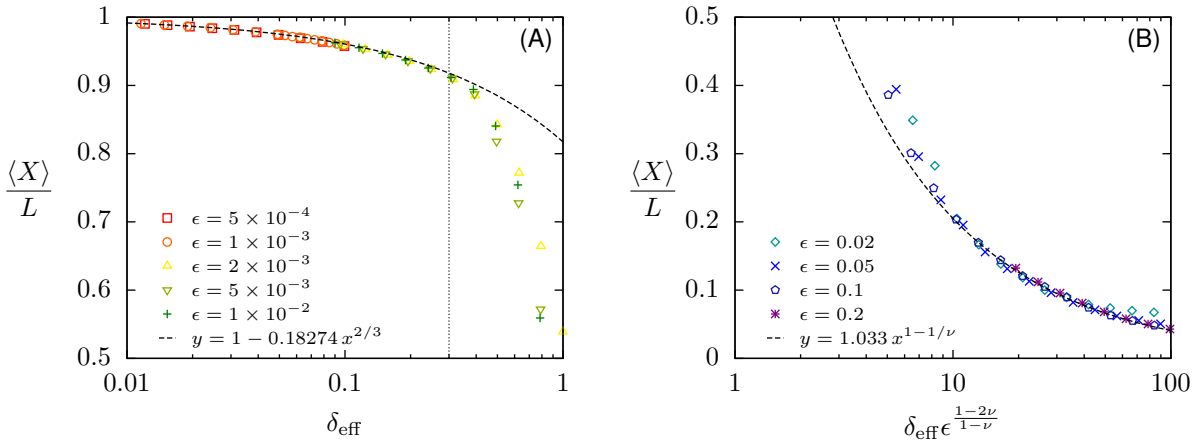

FIG. S5. Fractional extension of a semiflexible chain confined in a square nanochannel collapsed to the (A) Odijk regime and (B) Flory-de Gennes regime. (A) The data collapses before $\delta_{\text{eff}} = 0.3$ as indicated by the dotted vertical line. (B) Note here that the dependent variable is set to $x = \delta_{\text{eff}} \epsilon^{\frac{1-2\nu}{1-\nu}}$ and that $x^{1-1/\nu}$ yields the right hand side of Eq. (S11). This was done to make the ordinate linear in the channel size.

where the prefactor is obtained by a fit to the collapsed PERM data shown in Fig. S5. The prefactor of almost unity is a satisfying test of the de Gennes theory for a finite width chain, derived above.

In Fig. 3 of the main text, the data shown corresponds to $\epsilon = 10^{-1}$ and $\epsilon = 10^{-2}$ for a range of values of $\delta \equiv D/l_p$. These dimensionless data were converted using a persistence length $l_p = 50$ nm and two different values of the width, $w = 5$ nm and $w = 0.5$ nm. The corresponding channel sizes D follow from the definition of δ.

We also stated that the Odijk regime is valid for an effective channel size $\delta_{\text{eff}} \equiv (D - w)/l_p \leq 0.3$ and the Flory-de Gennes regime begins at $\delta_{\text{eff}} \geq (2\epsilon)^{-1}$. The approximate numerical values 0.3 and 1/2 were obtained by inspection of the data in Fig. S5A and S6A in the region where the data appear to collapse onto the universal curves. Additionally, one can again note that these regimes show the lack of collapse to the Odijk theory for a fractional extension below 90%. A similar general statement is not possible for the de Gennes theory, since this value depends on ϵ, as is evident in Fig. S6A.

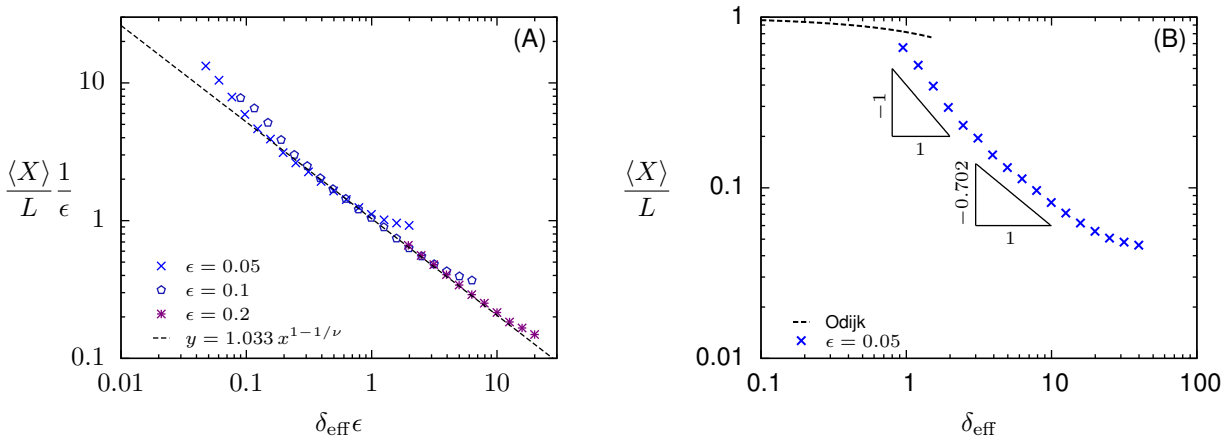

FIG. S6. (A) Log-log plot of the fractional extension from Fig. S5B, which shows the collapse to the Flory-de Gennes regime around $\delta_{\rm eff}\epsilon \approx 1/2$. This corresponds to about 200 nm for DNA ($\epsilon = 0.1$) (B) Average fractional extension versus dimensionless channel size for DNA ($\epsilon = 0.1$) obtained from PERM simulations. Both the Flory-de Gennes scaling and the Gauss-de Gennes scaling are indicated.

IV. SCALING FOR THE EXTENSION IN THE GAUSS-DE GENNES REGIME

Figure 3 of the main text shows the extension data on a linear plot in dimensional units, which is the easiest way to make a connection to the experimental data. Figure S6B shows the same data in a log-log plot using the effective dimensionless channel size, $\delta_{\rm eff}$, for the abscissa. These data correspond to asymptotically long chains; for 2×10^4 beads of size 5 nm, these data are for chains 100 μm long (approximately 300 kilobase pairs). Our extension results complement previous work [7–11] that used chains that are several microns long.

[1] Y. Wang, W. F. Reinhart, D. R. Tree, and K. D. Dorfman, Biomicrofluidics **6**, 014101 (2012).

[2] J. A. Schellman, Biopolymers **13**, 217 (1974).

[3] M. Rubinstein and R. Colby, *Polymer Physics* (Oxford University Press, 2003).

[4] H. Jian, A. V. Vologodskii, and T. Schlick, J. Comput. Phys. **136**, 168 (1997).

[5] T. Prellberg and J. Krawczyk, Phys. Rev. Lett. **92**, 120602 (2004).

[6] H. Frauenkron, M. S. Causo, and P. Grassberger, Phys. Rev. E **59**, R16 (1999).

[7] Y. Wang, D. R. Tree, and K. D. Dorfman, Macromolecules **44**, 6594 (2011).

[8] L. Dai, S. Y. Ng, P. S. Doyle, and J. R. C. van der Maarel, ACS Macro Letters **1**, 1046 (2012).

[9] P. Cifra, Z. Benková, and T. Bleha, J. Phys. Chem. B **113**, 1843 (2009).

[10] P. Cifra, J. Chem. Phys. **131**, 224903 (2009).

[11] P. Cifra, J. Chem. Phys. **136**, 024902 (2012).

[12] E. Werner, F. Persson, F. Westerlund, J. O. Tegenfeldt, and B. Mehlig, Phys. Rev. E **86**, 041802 (2012).

[13] D. R. Tree, Y. Wang, and K. D. Dorfman, Phys. Rev. Lett. **108**, 228105 (2012).

[14] P. Grassberger, Phys. Rev. E **56**, 3682 (1997).

[15] J. Z. Y. Chen and D. E. Sullivan, Macromolecules **39**, 7769 (2006).

[16] T. W. Burkhardt, Y. Yang, and G. Gompper, Phys. Rev. E **82**, 041801 (2010).

[17] T. Odijk, Macromolecules **16**, 1340 (1983).