Backfolding of wormlike chains confined in nanochannels

Abhiram Muralidhar,[†] Douglas R. Tree,[‡,¶] and Kevin D. Dorfman[*,†]

Department of Chemical Engineering and Materials Science, University of Minnesota – Twin Cities, 421 Washington Ave. SE, Minneapolis, MN 55455, USA, and Materials Research Laboratory, University of California, Santa Barbara, CA 93106, USA

E-mail: dorfman@umn.edu

[*]To whom correspondence should be addressed

[†]Department of Chemical Engineering and Materials Science, University of Minnesota – Twin Cities, 421 Washington Ave. SE, Minneapolis, MN 55455, USA

[‡]Materials Research Laboratory, University of California, Santa Barbara, CA 93106, USA

[¶]Previous address: Department of Chemical Engineering and Materials Science, University of Minnesota – Twin Cities, 421 Washington Ave. SE, Minneapolis, MN 55455, USA

Abstract

Using Pruned-Enriched Rosenbluth Method (PERM) simulations of a discrete wormlike chain model, we provide compelling evidence in support of Odijk's prediction of two distinct Odijk regimes for a long wormlike chain confined in a nanochannel. In both cases, the chain of persistence length $l_{\rm p}$ is renormalized into a series of deflection segments of characteristic length $D^{2/3}l_{\rm p}^{1/3}$, where D is the channel size. In the first (classic) Odijk regime, these deflection segments are linearly ordered. In the second Odijk regime, thin, long wormlike chains can backfold at a length scale quantified by the global persistence length. We have measured this quantity by simulations and modified Odijk's global persistence length theory to account for thermal fluctuations. The global persistence length, which is defined to be independent of the effect of excluded volume, provides the requisite closure to Odijk's scaling theory for the second regime and thus allows us to resolve much of the confusion surrounding the so-called "transition" regime for DNA confined in a nanochannel. We show that Odijk's theory for the backfolded regime correctly describes both the average chain extension and the variance about this extension for wormlike chains in channel sizes between the classic Odijk regime and the de Gennes blob regimes, with our data spanning several decades in terms of Odijk's scaling parameter ξ. Although the backfolded Odijk regime occupies a very narrow range of $D/l_{\rm p}$, it is indeed a regime when viewed in terms of ξ and grows in size with increasing monomer anisotropy.

1 Introduction

The problem of a wormlike chain confined in a channel has attracted significant attention since seminal experiments on DNA in nanochannels revealed a discrepancy between theory and experiment.[1] Subsequent research, driven primarily by simulation, has explained much of the discrepancy. For a wormlike chain of persistence length l_p and width w confined in a square channel of size D, there are now simulation data that support the existence of at least three regimes. The chain statistics in very tight channels, where $D \ll l_\mathrm{p}$, correspond to the classic Odijk regime[2] of linearly ordered deflection segments. The prefactors for the chain statistics in the classic Odijk regime are known[3] and in quantitative agreement with simulations.[4] When the confinement is relatively weak, confined chains are in the de Gennes regime and arrange themselves into a one-dimensional array of compression blobs.[5] While the original de Gennes model of isometric compression blobs only extends down to channels of size $D \approx l_\mathrm{p}^2/w$, there is now convincing evidence[6–8] for a so-called "extended de Gennes" regime[9–11] consisting of anisometric blobs that describes the chain statistics over the range $l_\mathrm{p} \lesssim D \lesssim l_\mathrm{p}^2/w$.

The key outstanding question for channel-confined wormlike chains, which we address here, concerns the chain statistics for channel sizes between $D/l_\mathrm{p} \ll 1$ and $D/l_\mathrm{p} \approx 1$. In the context of DNA in a nanochannel,[11,12] this range of channel sizes is often referred to as the "transition regime" despite the fact that simulations of DNA models in channel confinement[4,7,11,13] indicate that the transition from the extended de Gennes regime to the classic Odijk regime spans less than a decade in D/l_p. Indeed, since DNA is generally taken as the model polymer for studying channel-confined chains at the single molecule level, it is not even clear whether there exists a universal regime spanning channel sizes from $D/l_\mathrm{p} \ll 1$ to $D/l_\mathrm{p} \approx 1$ or just some gradual transition from the classic Odijk regime to the extended de Gennes regime. In this sense, the moniker "transition regime" splits the difference; these channel sizes either correspond to a transition or a regime, but probably not both. While the range of channel sizes spanning the gap between the classic Odijk regime and the extended de

Gennes regime may be narrow, these channel sizes are highly relevant for practical purposes, in particular in genomics,[12,14,15] since they encompass fractional extensions from about 20% to 85% for DNA.[4]

There exist three different explanations for the transition regime: (i) an ideal blob (Gauss-de Gennes) model,[4,16,17] which is inspired by simulation data[4,11,13] for DNA indicating that the mean extension in the transition regime scales like $\langle X \rangle \sim D^{-1}$, albeit over a narrow range of channel sizes; (ii) a cooperative backfolding model,[18] which was inferred from parameterized simulation data for chains with the relatively small ratios of $l_{\rm p}/w$ that characterize DNA; and (iii) a theory by Odijk[10] where the deflection segments can form hairpins with a characteristic length scale g, which Odijk calls the global persistence length.[19] Explanations (i-iii) are conveniently divided into two categories, based upon their fundamental explanation of the transition regime. Both the ideal blob and Odijk explanations posit a universal regime, in the sense that the behavior should persist over many decades in some relevant scaling parameter. By contrast, the backfolding model supposes that no such universal regime exists, and aims to describe a transition between the classic Odijk and extended de Gennes regimes.

Of the arguments supporting a universal regime, Odijk's theory[10] seems the most persuasive, since it merges Flory theory arguments that have been successful in explaining the blob regimes[12] with the undisputed presence of deflection segments in very tight confinement. Moreover, Odijk's theory suggests that the relevant scaling parameter for backfolded chains is not $D/l_{\rm p}$, and thus permits the possibility of a universal regime spanning many decades in another scaling parameter, ξ, for sufficiently large ratios of $l_{\rm p}/w$, even though it only may span a narrow range in $D/l_{\rm p}$. Unfortunately, the closure of the Odijk scaling theory[10] requires a model for the global persistence length that correctly accounts for the free energy costs for backfolding in the absence of excluded volume. Odijk[19] provided a mechanical approximation for the global persistence length, which indicates that an enormous contour length is required to observe backfolding of deflection segments.[10,19] However, Odijk was rather critical of his mechanical model of the global persistence length,[19] since it failed

to account for fluctuation effects.

In addition to its aforementioned positive qualities, Odijk's global persistence length theory is particularly appealing in the light of recent work on strongly confined wormlike chains in slits.[20] In agreement with a Flory theory by Odijk,[10] there is now numerical evidence that strongly confined chains perform a two-dimensional walk of deflection segments.[20] This walk gives rise to a number of different "sub-regimes," depending on the chain contour length, persistence length and excluded volume strength. Similar to the way in which blob regimes in channels parallel those in slits, we expect to find multiple regimes of strongly confined wormlike chains in channels. Indeed, Odijk's theory for channel confinement[10,19] posits that a wormlike chain performs a one-dimensional walk of deflection segments with regimes that depend the chain contour length and excluded volume strength.

In the present contribution, we compute the global persistence length via off-lattice, Pruned-Enriched Rosenbluth Method (PERM) simulations[21,22] of a discrete wormlike chain model,[23] which we have used recently[4,20,24,25] to study very long wormlike chains in confinement. We find that Odijk's mechanical theory[19] overestimates the global persistence length by several orders of magnitude over the range of channel sizes where chains of contour length $L \gg g$ are still accessible by simulation. By correcting Odijk's theory for thermal fluctuations,[19] we provide the requisite closure to test Odijk's proposition of a regime of backfolded deflection segments.[10] We show that this regime indeed exists, as the scaling theory not only collapses the data for both the average chain extension and the variance about the mean extension for channel sizes $D \lesssim l_{\mathrm{p}}$, but provides a collapse of the data over more than a decade in terms of Odijk's scaling parameter ξ (see Eq. 15). These results suggest that Odijk's theory,[10,19] as modified herein, is the correct description of the so-called "transition" regime for DNA in a nanochannel. Moreover, moving beyond the specific case of DNA, we confirm that the monomer anisotropy l_{p}/w plays a key role in demarcating the transition to the classic Odijk regime of linearly ordered deflection segments, with the width of the backfolded regime increasing with the monomer anisotropy. As a result, the backfolded Odijk

regime is very narrow for almost all practical situations, most notably for DNA.

2 Theoretical Background

2.1 Scaling theory

Odijk[10] classified channel confinement for $D \lesssim l_p$ into two regimes, based on whether or not the chains should be able to backfold. The regime with no backfolding of deflection segments is the classic Odijk regime,[2] which has been studied and confirmed in various theoretical[11,18,26,27] and experimental studies.[28–31] The second regime, characterized by backfolded deflection segments, was cast by Odijk into a modified Flory theory where the polymer is treated as a renormalized "chain of deflection segments" confined in a channel.

To see clearly the connection between Odijk's theory for a confined chain and standard Flory theory, let us first recall the well established Flory theory results for a real chain in free solution.[32] The chain is described as a series of segments of width w with a typical bending length scale l_p. The persistence length l_p is defined independent of excluded volume effects, arising from the intrinsic stiffness of the polymer.[33,34] The excluded volume strength can be quantified by the z parameter,

$$z = \frac{N_{lp}^2 v'_{ex}}{R^3} = \frac{w}{l_p}\left(\frac{L}{l_p}\right)^{1/2} \tag{1}$$

which measures the total excluded volume of a "gas" of $N_{lp} = L/l_p$ monomers uniformly distributed in a volume R^3 when the coil obeys ideal chain statistics, $R^2 = Ll_p$, with an excluded volume per contact of

$$v'_{ex} = l_p^2 w \tag{2}$$

The chain exhibits swollen coil behavior if $z > 1$ and Gaussian coil statistics if $z < 1$.

In free solution, the Flory free energy is the sum the entropic and excluded volume

energetic contributions,

$$\frac{F}{k_{\mathrm{B}}T} = \frac{R^2}{Ll_{\mathrm{p}}} + \frac{N_{\mathrm{lp}}^2 v'_{\mathrm{ex}}}{R^3} \tag{3}$$

where R^2 is the square of the end-to-end distance of the chain, k_{B} is the Boltzmann constant and T is the absolute temperature. Recall that the first term corresponds to the entropic elasticity of an ideal chain and the second term is the z parameter evaluated at some value of R. As noted by de Gennes, the first term is inaccurate because of the assumption of ideality, while the second term is erroneous as it does not consider correlations between monomer positions.[35] However, cancellation of errors in the Flory free energy gives rise to remarkably good scaling estimates of properties such as the average end-to-end distance.[32,35]

Odijk's theory[10] for a backfolded wormlike chain confined in a channel of size D between $D \ll l_p$ and $D \approx l_{\mathrm{p}}$ is built on an analogy with the Flory theory in free solution (Eq. 3). However, instead of considering a chain of monomers of size w and persistence length l_{p}, Odijk considers a chain of deflection segments, where the renormalized monomers have a length $\lambda = D^{2/3} l_{\mathrm{p}}^{1/3}$ and a width w. The excluded volume per contact is

$$v_{\mathrm{ex}} \approx \lambda^2 w \left(D/l_{\mathrm{p}}\right)^{1/3} \tag{4}$$

where the additional factor $\sin\delta = \left(D/l_{\mathrm{p}}\right)^{1/3}$ accounts for the orientation of deflection segments, which are aligned at an average angle, δ, between each other.[10] The characteristic length scale for bending in this renormalized chain of deflection segments, which corresponds to the formation of a hairpin, is the global persistence length g. Like the persistence length in free solution, the global persistence length in confinement is defined independent of excluded volume interactions.

For a confined chain of deflection segments, the analogy to the z parameter in Eq. 1 is[10]

$$Z = \frac{N_\lambda^2 v_{\mathrm{ex}}}{(Lg)^{1/2} D^2} \tag{5}$$

where $N_\lambda \equiv L/\lambda$ is the number of deflection segments in a chain of length L and $R^2 = Lg$ is the size of an ideal chain of such deflection segments with global persistence length g. Accordingly, the denominator of Eq. 5 is the total volume available to such an ideal chain in a channel of size D. In the development of his theory, Odijk[10] also defined a second parameter related to the excluded volume interactions for a single hairpin of length g,

$$\xi = \frac{n_\lambda^2 v_{\rm ex}}{gD^2} = \frac{gw}{D^{5/3} l_{\rm p}^{1/3}} \tag{6}$$

where $n_\lambda \simeq g/\lambda$ is the number of deflection segments in a single hairpin and gD^2 is the volume occupied by that hairpin.

Excluded volume thus plays a role at two different length scales in the Odijk theory. On one hand, the ability to backfold is controlled by ξ; if $\xi > 1$ then the volume gD^2 available to a hairpin is too small to contain the excluded volume $n_\lambda^2 v_{\rm ex}$ caused by the turn. On the other hand, if the chain can backfold, the swelling of the chain of deflection segments by excluded volume is controlled by Z; if $Z > 1$, then excluded volume amongst all of the deflection segments plays an important role. Odijk's theory[10] thus consists of three possible cases: (i) $\xi > 1$, where the chains cannot backfold, (ii) $\xi < 1$ and $Z > 1$, where the chains backfold and experience excluded volume interactions, and (iii) $\xi < 1$ and $Z < 1$, where the chains backfold but do not experience significant excluded volume interactions. Let us examine these three cases in turn.

If $\xi > 1$, then the excluded volume caused by the backfolded deflection segments in a hairpin exceeds the volume available in the channel containing those deflection segments. Backfolding is thus suppressed and the deflection segments are linearly ordered.[2,10] When excluded volume interactions are strong enough to prevent backfolding, the mean extension of the confined chain is close to its contour length L,[2,3]

$$\langle X \rangle = L\left[1 - \alpha\left(\frac{D}{l_{\rm p}}\right)^{2/3}\right] \tag{7}$$

and the variance $\delta X^2 \equiv \langle (X - \langle X \rangle)^2 \rangle$ about this average extension is

$$\delta X^2 = \beta \frac{D^2}{l_\mathrm{p}} L. \tag{8}$$

Burkhardt et al.[3] computed the prefactors for square channels to be $\alpha = 0.18274$ and $\beta = 0.00956$. In the above expressions, we have used the mean span

$$\langle X \rangle \equiv \langle x_\mathrm{max} - x_\mathrm{min} \rangle \tag{9}$$

as the measure of the extension of the chain, where x_max and x_min refer to the maximum and minimal axial positions respectively.[11] For strongly stretched chains, or for asymptotically long chains in channel-confinement, the mean span and mean end-to-end distance are indistinguishable.[24]

If $\xi < 1$, then the deflection segments are able to backfold due to the weakness of excluded volume of a single hairpin interacting with itself. In this sense, a single hairpin is pseudo-ideal. For long chains, there are also excluded volume interactions between hairpins. The overall strength of the hairpin-hairpin excluded volume interaction is determined by Z, with the value $Z \approx 1$ demarcating the boundary between ideal and real behavior for the chain as a whole. For the case $\xi < 1$, Odijk's modification of the Flory theory in Eq. 3 leads to[10]

$$\frac{F}{k_\mathrm{B}T} \cong \frac{L}{\lambda} + \frac{X^2}{Lg} + \frac{N_\lambda^2 v_\mathrm{ex}}{XD^2}. \tag{10}$$

The first term in Eq. 10, representing the confinement of the deflection segments,[2] does not appear in the original theory[10] because it does not contribute to the calculation of the average chain extension $\langle X \rangle$, even though it is the dominant term in the calculation of F. We include it here since (i) we know from previous simulation work[4] that the confinement free energy is extensive in chain length and exhibits Odijk scaling $F \sim D^{-2/3}$ well past $D/l_\mathrm{p} \ll 1$ and (ii) the neglect of such terms in the Flory theories for the confinement free energy in blob

regimes[9] led to some confusion about the existence of the extended de Gennes regime.[4,7]

Let us consider the two cases in turn. For the case $\xi < 1$ and $Z > 1$, minimizing the free energy in Eq. 10 with respect to X leads to the average extension[10]

$$\langle X \rangle \sim L\xi^{1/3}. \tag{11}$$

and the second derivative of the free energy produces the variance

$$\delta X^2 \sim gL, \tag{12}$$

For the case where $\xi < 1$ and $Z < 1$, the extension of the chain of deflection segments will exhibit ideal scaling[19]

$$\langle X \rangle \sim \sqrt{gL}, \tag{13}$$

with a persistence length g. Accordingly, the variance for $\xi < 1$ and $Z < 1$ is the same as Eq. 12. The crossover contour length, L_*, between Eq. 11 and Eq. 13, is given by $Z = 1$, namely[10]

$$L_* \cong g\xi^{-2/3}. \tag{14}$$

Note that these are the limiting cases for $\xi < 1$, and we would expect to see smooth transitions between real and ideal scaling for the extension near $Z = 1$.

For the backfolded Odijk regime ($\xi < 1$), in addition to L_*, there is also a second crossover length $L_{**} = g$ where the chains are so short that they cannot even form a global persistence length. Interestingly, for $L < L_{**}$, so long as these chains have a contour length of at least $L > \lambda$, they exhibit the same extension and variance as the classic Odijk regime given by Eq. 7 and Eq. 8 respectively, independent of the value of ξ. The linear ordering of such short chains is driven by an insufficient number of segments to form a global hairpin, rather than excluded volume interactions suppressing backfolding. In addition, there should exist other sub-regimes for even shorter chains $L < \lambda$ corresponding to the wall-induced orientation of a

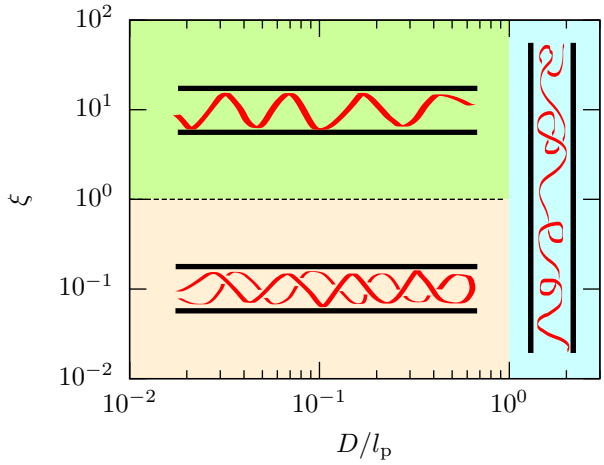

Figure 1: Phase diagram for confined wormlike chains based on the scaling theory of Odijk.[10] Channel sizes $D/l_\mathrm{p} \lesssim 1$ correspond to Odijk regimes, where $\xi \gtrsim 1$ is the classic Odijk regime[2] and $\xi \lesssim 1$ is the backfolded Odijk regime. For wider channels with $D \gtrsim l_\mathrm{p}$, there is a gradual transition to blob regimes. This schematic is simplified to emphasize the importance of both D/l_p and ξ; note that ξ depends on D/l_p via Eq. 15.

rod, analogous to the case of a slit.[20]

Figure 1 summarizes our discussion of Odijk's scaling theory[10] so far. This figure emphasizes that the chain statistics depend on two factors: (i) the channel size relative to the persistence length, D/l_p, and (ii) the relative strength of excluded volume to the channel volume in a hairpin, ξ. Note that while both parameters are required to classify the various regimes, ξ and D/l_p are not independent variables. A rearrangement of Eq. 6 yields

$$\xi \simeq \left(\frac{g}{l_\mathrm{p}}\right)\left(\frac{l_\mathrm{p}}{D}\right)^{5/3}\left(\frac{w}{l_\mathrm{p}}\right), \tag{15}$$

which shows how ξ is related to other relevant dimensionless length scales. In particular, ξ has both an explicit dependence of $(D/l_\mathrm{p})^{-5/3}$ and an implicit dependence on D/l_p through the global persistence length. As the channel size increases and the chain exits the Odijk regimes, we expect that $g \to l_\mathrm{p}$.[19] In other words, the global persistence length limits to the persistence length of the chain in the absence of confinement, which further emphasizes that both types of persistence lengths are defined independent of excluded volume interactions.

2.2 Odijk theory for the Global Persistence Length

The challenge in applying Odijk's theory[10] is that it requires a model for the global persistence length g to close the system of equations. Odijk developed a theory for the statistical mechanics of hairpins for ideal wormlike chains confined in square channels, an outcome of which is an expression for g.[19] By determining the optimum chord length of the hairpin that minimizes the bending penalty for hairpin formation, and estimating the free energy of confinement of these hairpins, Odijk arrived at an expression for the global persistence length,[19]

$$g = \alpha \bar{r} \exp\left(\frac{\bar{F}}{k_{\mathrm{B}}T}\right). \tag{16}$$

The first term in Eq. 16 is the constant $\alpha = 3.3082$, which Odijk obtained by numerical integration from his model. The quantity $\bar{r}(D/l_{\mathrm{p}})$ is the average length of a hairpin chord,

$$\frac{\bar{r}}{l_{\mathrm{p}}} = \frac{1}{6}\left\{\left[E_m^2 + 6\sqrt{2}E_m\left(\frac{D}{l_{\mathrm{p}}}\right)\right]^{1/2} - E_m\right\}, \tag{17}$$

where a separate numerical integration produced $E_m = 1.5071$. Odijk's derivation corresponds to the mechanical limit, $\bar{F} = \bar{F}_{\mathrm{mc}}$, where $\bar{F}_{\mathrm{mc}}(D/l_p)$ has the form

$$\frac{\bar{F}_{\mathrm{mc}}}{k_{\mathrm{B}}T} = E_m\left(\frac{l_{\mathrm{p}}}{\bar{r}}\right) - 3\ln\left(\frac{D - \bar{r}\sqrt{2}}{D}\right) - \ln\left(\frac{8}{3\pi}\right) \tag{18}$$

and accounts for bending energy of the hairpin in addition to its translational and orientational entropy.

Odijk's mechanical approximation for the global persistence length has no adjustable parameters, which makes it a remarkable result. However, at the conclusion of his paper,[19] Odijk provides an enumerated list of shortcomings in this model. Thus, in order to test Odijk's scaling theory, we need to first validate (and, if necessary, modify) Odijk's mechanical model for the global persistence length.

3 Simulation Methodology

From a simulation standpoint, testing Odijk's theory is reasonably difficult. The challenge becomes clear when we consider Eq. 16 in the limit of a very small channel,

$$\frac{g}{l_\mathrm{p}} \simeq c_1 \left(\frac{l_\mathrm{p}}{D}\right)^2 \exp\left[c_2\left(\frac{l_\mathrm{p}}{D}\right)\right] \quad \text{for } D/l_\mathrm{p} \to 0 \tag{19}$$

where c_1 and c_2 are $O(1)$ constants. Thus, in the limit of narrow channels, the contour length between hairpins diverges exponentially and hairpin formation becomes a rare event. From a practical standpoint, this exponential divergence makes the observation of hairpins in simulations increasingly difficult with decreasing channel size, necessitating simulation of extraordinarily long chains. Coarse-graining of the model to access these length scales, however, is limited by the ability of a coarse-grained model to resolve the sub-persistence length scale information characterizing deflection segments, which is critical in this extreme-confinement regime.[7]

To address this challenge, we exploit the capability of the Pruned-Enriched Rosenbluth Method (PERM)[21,22,36] to simulate long wormlike chains in narrow square channels with a sufficiently fine-grained model, namely the discrete wormlike chain model.[23] Our model consists of a discretized chain of beads connected by rods of length a. For real chains, we use $a = w$, and in the case of ideal chains, a is chosen such that $2l_\mathrm{p} \geq a$. Particularly for simulations in the Odijk regime, we opt for $D/a \geq 5$ in order to resolve deflection segments.[24] We have described our simulation approach in some detail elsewhere[4,24,36] and provide a brief description of the methodology in the Supporting Information.

4 Results

4.1 Measuring the Global Persistence Length

The global persistence length g is defined independent of any excluded volume, as it plays the one-dimensional analog to the three-dimensional persistence length l_p in free solution, and the effect of excluded volume only enters the scaling theory through the third term in Eq. 10. We thus only need to work for the moment with ideal wormlike chains ($w = 0$). In the context of the phase diagram in Figure 1, ideal wormlike chains are either linearly ordered (if $L < g$) or backfolded ($L > g$) since $\xi = 0$ for $w = 0$ by Eq. 6.

To measure g as a function of the confinement, D/l_p, we take advantage of qualitative similarities between the properties of a wormlike chain in channel confinement and the properties of single wormlike chains in the nematic phase of lyotropic liquid crystalline solutions, which has been studied extensively.[10,19,37–39] Just as wormlike chains align parallel to the director axis in the nematic phase,[40] wormlike chains in confinement align parallel to the axis of the channel. There is no polarity associated with the channel axis, i.e. the unit vectors $\hat{\mathbf{e}}_x$ and $-\hat{\mathbf{e}}_x$ are equivalent for a channel whose axis is parallel to the x axis, much the same as the director vector. Furthermore, wormlike chains in the nematic phase form hairpins that are separated by a length scale resembling the global persistence length.[37,39] Accordingly, it is convenient to define an orientational order parameter for the channel confined chains as the second Legendre polynomial of $\cos\theta$,[27,39,41]

$$m \equiv \langle P_2(\cos\theta)\rangle = \frac{1}{2}\left(3\left\langle\cos^2\theta\right\rangle - 1\right), \tag{20}$$

where θ is the angle between the channel axis and a tangent to the polymer molecule. The angle brackets in Eq. 20 denote both an ensemble average over multiple chains and an average over the whole contour for long chains.

In the long-chain limit, the confined chain behaves like a one-dimensional ideal wormlike

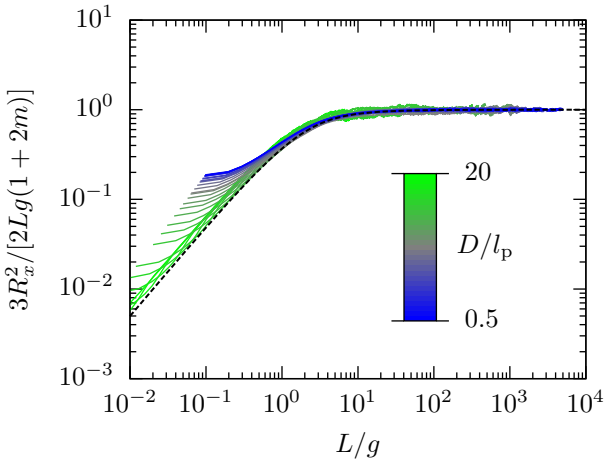

Figure 2: RMS end-to-end distance versus the contour length rescaled using the global persistence length for ideal wormlike chains with 10 beads per persistence length (l_p/a = 10). The colored curves correspond to simulation data for chains confined in 29 channels logarithmically spaced between D/l_p = 0.5 to D/l_p = 20. The black dashed line shows the result from Eq. 21 for $m = 0$, which is the expected behavior in free solution.

chain aligned along the x-axis with an effective persistence length of g, albeit with small-scale excursions perpendicular to the channel.[10,19,37,38] From the analogy between a wormlike chain in the nematic phase and a wormlike chain in channel confinement, the projection of the mean square end-to-end distance on the x-axis in the limit $L \gg g$ can be written as[39,42]

$$R_x^2 = \frac{1}{3}(1+2m)\left[2gL - 2g^2\left(1-\exp\left(-L/g\right)\right)\right]. \tag{21}$$

The prefactor $(1+2m)/3 = \langle\cos^2\theta\rangle$ in Eq. 21, though not considered in Refs. 10 and 19, accounts for the alignment of the chain along the channel axis. For a completely aligned chain ($m = 1$), Eq. 21 reduces to the mean square end-to-end distance of a 1D wormlike chain with persistence length g. On the contrary, for an isotropic chain in free solution ($m = 0$), Eq. 21 gives the expression for the projection of the end-to-end distance for a 3D wormlike chain with persistence length g. It should be noted that the scaling of R_x and X must be the same in the long-chain limit; we use R_x^2 here as it has the relatively simple relation to g embodied in Eq. 21.

Having first established that our simulations give the expected result for the orientational

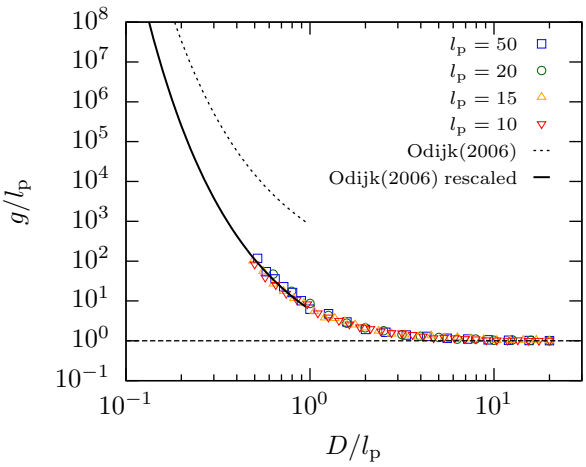

Figure 3: Collapse of global persistence length data for different values of l_p and D. The dotted line shows the result from Eq. 16. The solid line corresponds to Eq. 23. An extrapolation of Eq. 23 is also shown for $D/l_p < 0.5$, where simulation data are absent. All values of l_p are in units of a.

order parameter (see Supporting Information), we calculated g by fitting our PERM simulation data for various values of l_p and D to Eq. 21. Figure 2 shows the rescaled value of R_x^2 plotted as a function of L/g, for ideal wormlike chains confined in channels of sizes ranging from $D/l_p = 0.5$ to $D/l_p = 20$. In accordance with Eq. 21, the data for all chains collapse onto a single curve when $L/g \gtrsim 10$, since this collapse is enforced in our fits (see Supporting Information). We also observe that the end-to-end distance in the unconfined direction approaches the curve for free solution ($m = 0$) as one increases the channel size, confirming that the size of ideal wormlike chains parallel to the axis of the channel is unaffected by confinement for $D \gg l_p$.[24]

As we obtained the values for g for each value of D/l_p to produce Figure 2, we are now in a position to test Odijk's theory[19] for the global persistence length. Figure 3 demonstrates that the global persistence length is indeed a universal quantity, as all our data fall on a single curve as a function of D/l_p, irrespective of the value of l_p. It is worth pointing out that g is almost 10 times l_p even for channels as big as $D = l_p$. These results are consistent with previous observations of the increase in the apparent persistence length of wormlike chains in channel and tube confinement.[43,44] Moreover, for $D \gg l_p$, the global persistence

length approaches the native persistence length of the molecule, suggesting that the length scale for backfolding approaches the persistence length in the de Gennes regime.[19]

Our simulation method for computing g has some advantages compared to Odijk's theoretical approach[19] since we can reliably determine g in the wide-channel limit and thus obtain the right limiting value of $g = l_p$. However, our method is much more limited in the small-channel limit. We cannot directly verify the behavior of g/l_p as $D/l_p \to 0$, since our calculation of g relies on simulations of chains of contour length of the order of several global persistence lengths. Because the value of g exponentially increases with decreasing D/l_p (Eq. 19), obtaining an estimate of g for $D/l_p < 0.5$ becomes prohibitively expensive. This exponential growth poses an intrinsic problem for simulations of the global persistence length. For example, while other coarse-grained models such as the freely-jointed rod model[6,7] have been very successful at modeling the extended de Gennes regime, they cannot be used to overcome this issue either, since these models lack the sub-persistence length information critical in resolving the deflection segments.

Figure 3 also compares our simulation data with Eq. 16 for $D < l_p$, as this is roughly the range in which Odijk's theory[19] should be valid. Although our simulation data do not agree quantitatively with his theory, we find a similar qualitative trend — in both Odijk's theory and our simulations, g/l_p rapidly increases as D/l_p decreases to values much lesser than unity. However, it appears from Figure 3 that Odijk's theory overestimates g by about two orders of magnitude in the range of our simulation data.

Odijk derived the expression in Eq. 16 in the mechanical limit, neglecting any fluctuations of the hairpin. In addition, the approximations used by Odijk en route to Eq. 16 consider only the leading order terms in the perturbations to the shape of the hairpin about the lowest energy state, resulting in systematic errors in the estimation of g. Odijk recognized these potential shortcomings in his derivation[10] and proposed that the free energy in the global

persistence length of Eq. 16 should be of the form

$$\bar{F} = \bar{F}_{\mathrm{mc}} + H(D/l_{\mathrm{p}}) \tag{22}$$

where the second term corrects for the approximations in the mechanical limit.[19]

Although there is no obvious way to compute $H(D/l_{\mathrm{p}})$ theoretically,[19] we can easily estimate this quantity from the simulation results in Figure 3. For each channel size D/l_{p}, we use Eq. 22 as the free energy in Eq. 16 and solve for the correction term H. We find that H is a weak function of D/l_{p} for the channel sizes $D \leq l_{\mathrm{p}}$ where we could measure the global persistence length, and it is almost constant in the range of our simulations (see Supporting Information). The best fit constant to the data is $H/k_{\mathrm{B}}T = -4.91$, leading to the global persistence length

$$g = \alpha \bar{r} \exp\left(\frac{\bar{F}_{\mathrm{mc}}}{k_{\mathrm{B}}T} - 4.91\right). \tag{23}$$

As we see in Figure 3, this corrected form of the global persistence length for $0.5 < D/l_{\mathrm{p}} < 1$ is in good quantitative agreement with our simulation data, as would be expected from the relative insensitivity of H to the channel size.

From a practical standpoint, we can treat Eq. 23 as a useful functional form, valid over the range of D/l_{p} where we have obtained our data. In other words, Eq. 23 provides the closure to Odijk's Flory theory and thus allows us to evaluate the predictions of that theory independent of the theory for the global persistence length. However, we need to be cautious in extrapolating Eq. 23 for channels where $D/l_{\mathrm{p}} < 0.5$. Both Odijk's model in Eq. 16 and our modified version in Eq. 23 agree with the asymptotic result in Eq. 19, albeit with different prefactors. However, it seems reasonable that $H \to 0$ as $D/l_{\mathrm{p}} \to 0$, since extreme confinement would suppress the thermal fluctuations. This logic implies that H is indeed a function of D/l_{p} over some range of channel sizes. Thus, we would expect that the actual global persistence length for $D < l_{\mathrm{p}}/2$ would lie somewhere between the extrapolation of Eq. 23 in Figure 3 and the mechanical model.

Nonetheless, the quantitative difference between the value of g from our simulations and Odijk's mechanical theory seems to clarify some apparent discrepancies in the literature. Odijk used Eqs. 16-18 and an equation similar to Eq. 21 to calculate the mean end-to-end distance of "ideal" wormlike chains with the same persistence length of DNA and showed that resultant values are in agreement with experimental data from Reisner et al.[1] This agreement now appears like a fortuitous cancellation of errors, as Odijk's approach for estimating the extension of DNA did not account for excluded volume, while making use of extraordinarily high values for g. Furthermore, Wang et al.[11] observed that the values for the extension of DNA in their simulations were much lower than what one would predict from the mechanical theory. Our results for g, which are two orders of magnitude less than Odijk's prediction, are consistent with this observation.

4.2 Test of Scaling Theory

Now that we have computed the global persistence length, we are in a position to evaluate Odijk's scaling theory for the backfolding regime. Because we are addressing the problem numerically, we can only evaluate his theory for cases where we can obtain g by the procedure outlined above. Fortunately, this does not limit our analysis of real chains ($w \neq 0$), since the computationally accessible contour lengths of real chains are a subset of those we can achieve with ideal chains. This is due to the fact that ideal chain simulations are cheaper than real chain simulations, since we do not need to make expensive excluded volume calculations. As a result, if we cannot simulate an ideal chain to sufficient length to measure g for a given channel size, then we also cannot simulate a real chain to sufficient length $L > L_*$ to reach the long chain limit, since $L_* \gg g$.

In what follows, we only present real chain data where we reach the long chain limit $\langle X \rangle \sim L$. Furthermore for the variance, we restrict ourselves to the subset of the latter data where we also reach the long-chain limit $\delta X^2 \sim L$ (see Supporting Information). The value of g for all real chains confined in channels of size D has been computed for equivalent ideal

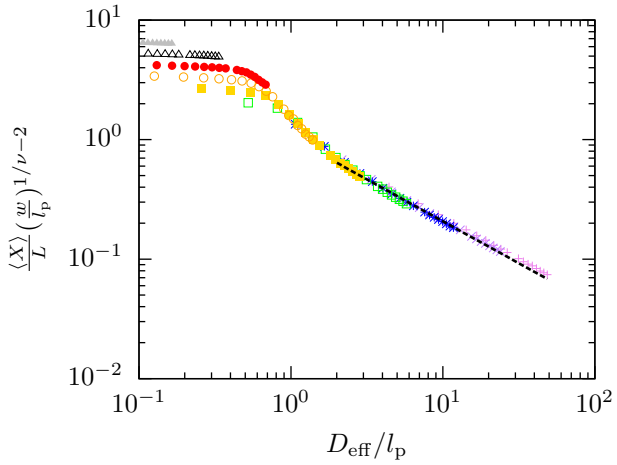

Figure 4: Collapse of extension data using the scaling for the blob regimes. The dashed line indicates the scaling $\langle X\rangle/L = 1.04278\,(D_{\rm eff}/l_p)^{-0.701}\,(w/l_p)^{0.298}$. The l_p/w values used here are 2.04 (violet +), 4.15 (purple ×), 8.45 (blue ∗), 17.20 (green □), 35.04 (gold ■), 71.35 (orange ○), 145.3 (red •), 295.5 (black △) and 602.7 (gray ▲).

chains confined in channels with size $D_{\rm eff}$, except for cases where Eq. 23 has been used as an approximation. Note that we do not need to test the confinement free energy, since we showed that $F \sim D^{-2/3}$ for these channel sizes in a previous publication.[4] This scaling of the confinement free energy corresponds to the confinement of ideal wormlike chains, which is accounted for by the L/λ term in Eq. 10. The magnitude of the excluded volume term is small compared to this term and thus does not contribute at the leading order to the scaling of the free energy.[20]

We begin our test of the scaling theory with the extension. We first need to identify the blob-to-deflection segment transition in the phase diagram of Figure 1. To do this, we repeat the analysis by Dai et al.[7] for our results with touching beads ($a = w$) and plot our extension data rescaled in the same way as Dai et al.[7] in Figure 4. For real chains, the channel size $D_{\rm eff} = D - w$ represents the cross-sectional width available to the chain. For the extended de Gennes regime,[7,9,11] we would expect extension to have the scaling $\langle X\rangle \sim D_{\rm eff}^{(\nu-1)/\nu}$, where $\nu = 0.587597(7)$ is the Flory exponent.[45] We observe that although the data seem to collapse all the way down to channel sizes $D \approx 0.8 l_p$, there is a clear change in slope around $D_{\rm eff} \approx 2 l_p$, indicating that the chains are losing their blob character and are are no longer in the extended

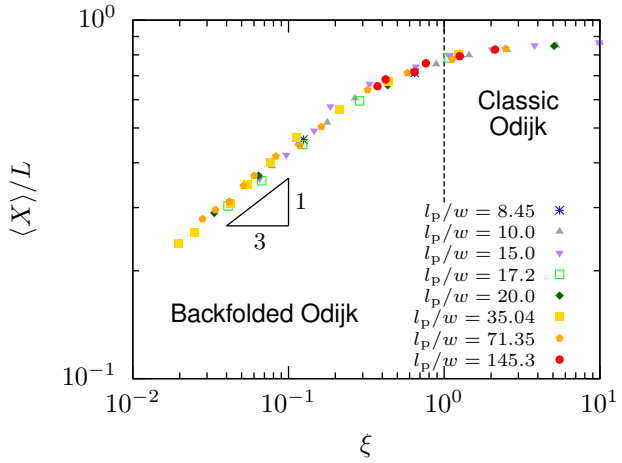

Figure 5: Plot of the fractional extension versus the parameter ξ for real wormlike chains with different monomer anisotropies. The vertical dashed line denotes the boundary between the two regimes according to scaling theory, $\xi = 1$. We calculated ξ from Eq. 6 by using g data obtained from simulations of ideal wormlike chains for the same $D_{\text{eff}}/l_{\text{p}}$ values as those of the corresponding real chains. Only data from channel sizes with $D_{\text{eff}} < 2l_{\text{p}}$ are shown. The backfolded Odijk regime encompasses fractional extensions from 20% to 80%, which is the typical range of the "transition" regime. A fit to the data for $\xi < 0.2$ reveals a relation given by $\langle X\rangle/L = 0.92\xi^{1/3}$.

de Gennes regime for smaller channels. This value, $D_{\text{eff}} \approx 2l_{\text{p}}$, for the transition out of the extended de Gennes regime is in agreement with the freely-jointed rod model of Dai et al.[7] These results show that the data are insensitive to the detailed model of the chain, confirming the universal nature of the de Gennes blob regime.

Having confirmed that there is a transition to the blob regime at $D/l_{\text{p}} \approx 2$, we now turn our attention to examining Odijk's scaling theory for backfolded chains. Figure 5 shows that the Odijk backfolding model in Eq. 11 with computed global persistence length values from ideal wormlike chain simulations collapses all of the data for $\xi \ll 1$ and $D_{\text{eff}} < 2l_{\text{p}}$, independent of the particular values of l_{p}/w and D/l_{p}. As we would expect, the data do not exhibit the scaling $\langle X\rangle/L \sim \xi^{1/3}$ for $D_{\text{eff}} > 2l_{\text{p}}$, even when $\xi \ll 1$, since these channels correspond to the blob regimes (see Supporting Information). While the range of $D_{\text{eff}}/l_{\text{p}}$ for the backfolded Odijk regime in Figure 4 may be small, this regime spans many decades in ξ, thereby identifying ξ as the proper scaling variable. The lower bound on the value of ξ for the backfolded Odijk regime for a given l_{p}/w ratio can be estimated by using a

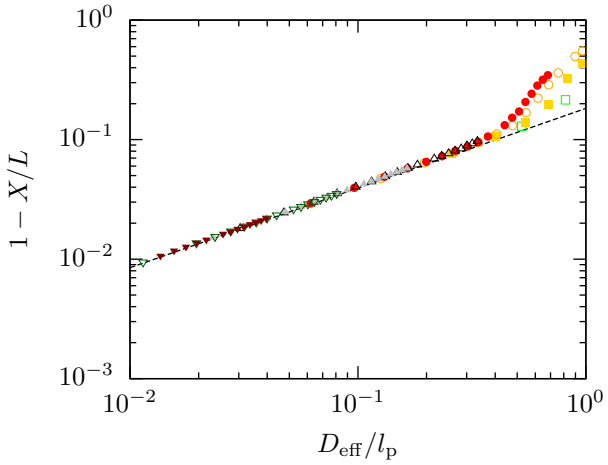

Figure 6: Collapse of the fractional extension onto the classic Odijk regime in extreme confinement. The $l_{\rm p}/w$ values are 17.2 (green $\square$), 35.03 (gold $\blacksquare$), 71.35 (orange $\circ$), 145.32 (red $\bullet$), 295.96 (black $\triangle$) 602.74 (gray $\blacktriangle$), 1227.54 (dark-green $\triangledown$), 2500 (brown $\blacktriangledown$). The dashed line is Eq. 7.

channel size $D \approx 2l_{\rm p}$ as the boundary. At this channel size, $g \approx 2l_{\rm p}$ (from Figure 3), implying that the lower bound is $\xi \approx 0.63(w/l_{\rm p})$ from Eq. 15. Note that the range of fractional extensions corresponding to the backfolded Odijk regime in Fig. 5 is between 0.2 and 0.8, which is reminiscent of DNA extensions in the transition regime.[11,46] This rapid increase in the extension over a small range of channel sizes is due to the brisk increase in the scaling parameter ξ with decrease in D, as the extension is related to ξ by Eq. 11. The rise in ξ is in turn related to the exponential upsurge of g, since ξ is proportional to g (Eq. 6).

The deviation from the scaling $\langle X \rangle \sim \xi^{1/3}$ takes place in the range $\xi \approx 1$, as predicted by Odijk's scaling theory. The accuracy of his theory is remarkable in this respect, since the statistics used to estimate the excluded volume (in particular the angle δ) are only valid for $D \ll l_{\rm p}$ but still provide a reasonable estimate out to $D \approx l_{\rm p}$. For $\xi > 1$, the chain approaches full extension, as predicted by the Odijk scaling theory. Indeed, as we see in Figure 6, the prefactors from Burkhardt et al.[3] lead to quantitative agreement with all of our simulation data once the chains crossover to the classic Odijk regime.

It is natural to wonder why the calculations by Gompper and coworkers[3,47] work so well for real chains[4,11] even though they take advantage of an ideal chain model. Indeed, the

case is similar for analyses of the chain extension using numerical solutions of the modified diffusion equation for the propagator.[41] In these ideal chain models, the extension in the Odijk regime was evaluated either by computing the rate of change of transverse displacement along the length of the chain[3,47] or the average angle between a deflection segment and the channel axis.[41] Both calculations inherently assume that there is no backfolding. Although the results for extension from simulations of Gompper and coworkers[3,47] and the numerical estimates of Chen[41] are in agreement with one another, it is now clear that these results are valid only for short chains in the backfolded Odijk regime ($L \ll g$) and real chains in the classic Odijk regime. Because ideal wormlike chains in the extreme confinement limit ($D \lesssim l_p$) are always in the backfolded Odijk regime ($\xi = 0$ identically), these results[3,41,47] for extension of ideal chains do not apply in the long-chain limit. Note that similar behavior is also observed with the Odijk regime for wormlike chains confined in slits.[20]

The results for the Odijk backfolding theory are even more impressive for the chain variance in Figure 7. For relatively large channels ($D_{eff}/l_p \approx 3$ to $D_{eff}/l_p \approx 10$), the variance is independent of channel size, as observed previously[6] in studies of the extended de Gennes regime. Our results also confirm the expected scaling in the de Gennes regime, $\delta X^2 \sim D_{eff}^{1/3}$,[12] for the largest channels. Once the chain passes into the backfolded Odijk regime, the fluctuations in the extension increase rapidly; note that the ordinate in Figure 7(a) spans nine orders of magnitude. As real chains cross over from the extended de Gennes regime to the backfolded Odijk regime, their rescaled variance $\delta X^2/Ll_p$ follows the $g/4l_p$ curve, as shown in Figure 7(b). This implies that

$$\delta X^2 \approx 0.25Lg \tag{24}$$

in accordance with Eq. 12. The collapse of the data onto Odijk's theory is even more apparent in Fig. 8, where we use the scaling variable ξ rather than the channel size D_{eff}/l_p. Similar to what we saw for the extension in Fig. 5, the data collapse onto the Odijk theory for $\xi < 0.1$,

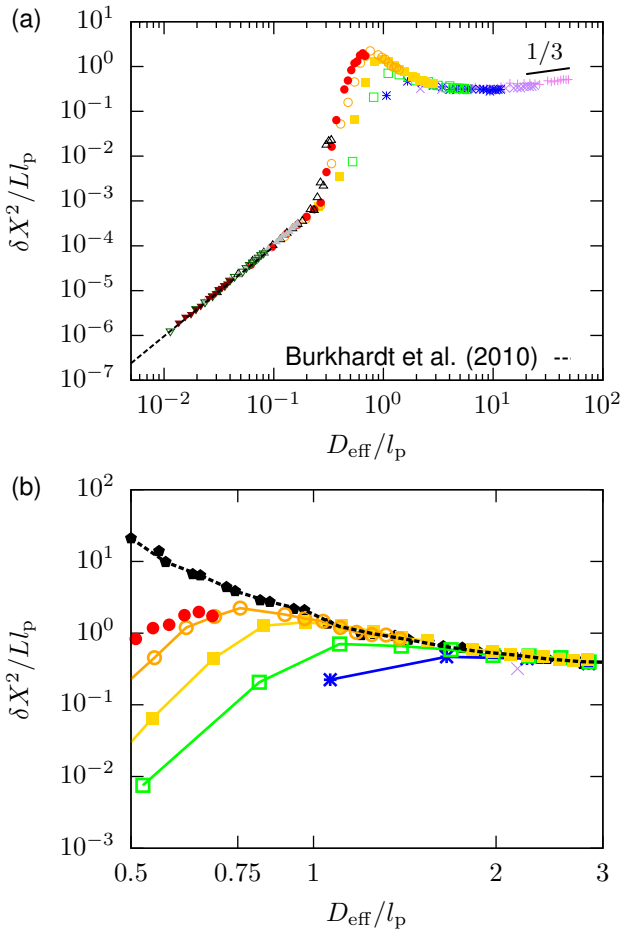

Figure 7: (a) Normalized variance in the chain extension as a function of channel size. The l_p/w values 2.04 (violet +), 4.15 (purple ×), 8.45 (blue ∗), 17.20 (green □), 35.04 (gold ■), 71.35 (orange ∘), 145.3 (red •), 295.5 (black △) and 602.7 (gray ▲), 1227.54 (dark-green ▽), 2500 (brown ▼). The dashed line is Eq. 8 for the classic Odijk regime. The solid line shows the slope $\delta X^2 \sim D_{\text{eff}}^{1/3}$ expected in the de Gennes regime. (b) The region in part (a) where the normalized variance appears to rise up with decreasing channel size. The solid black pentagons are simulation data for $g/4l_p$ for ideal chains of l_p = 15 and l_p = 10. The black dashed line, which joins the simulation data for $g/4l_p$, and the colored solid lines for the variance data are meant as a guide to the eye.

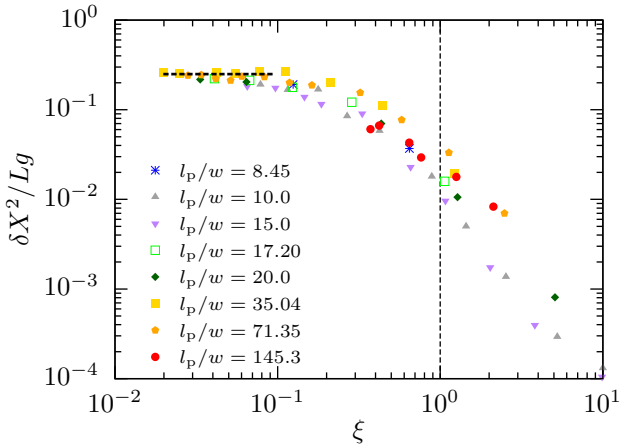

Figure 8: Collapse of the variance in the chain extension to Eq. 12 for $\xi < 0.1$ and $D_{\text{eff}} < 2l_p$. The horizontal black dashed line denotes $\delta X^2/Lg = 0.25$ from Eq. 24 and the vertical dashed line shows the boundary between the backfolded and classic Odijk regimes according to scaling theory ($\xi = 1$).

given that $D_{\text{eff}} < 2l_p$. The collapse of the data for the variance in Fig. 8 looks sparser than that for the extension in Fig. 5 for two reasons. First, it is difficult to reach the asymptotic limit for the chain variance (see Supporting Information). Second, the scaling of the variance in the backfolded Odijk regime given by Eq. 12 holds only for $\xi < 0.1$, while the scaling of the extension in Eq. 11 is valid up to $\xi \approx 0.4$. While this result is indeed a scaling law in terms of g, the global persistence length itself is not a power law in channel size D, so there is no simple power law of the form $\delta X^2 \sim D^x$. Note that the variance data in Figure 7(b) fall off the $g/4l_p$ curve at different values of D_{eff}/l_p depending on the monomer anisotropy, l_p/w, suggesting that the width of the backfolded regime is a function of l_p/w. This dependence of the width of the backfolded Odijk regime on the monomer anistropy will be addressed in detail in Section 5.1.

Eventually, all real chains need to reach a channel size where $\xi > 1$. Since excluded volume terminates the backfolded Odijk regime, chains with higher monomer anisotropy remain in the backfolded Odijk regime down to smaller channel widths. The global persistence length increases almost exponentially as the channel size decreases, so the maximum in the variance of the extension also increases rapidly with monomer anisotropy. This peak in the variance

in chain extension, taken together with the freely draining hydrodynamics for these channel sizes,[46] also explain the peak in the relaxation time for DNA in nanochannels around 100 nm.[1,48] Similar to the chain extension, there is a broad transition for the δX^2 out of the backfolded Odijk regime and into the classic Odijk regime. As we see in Figure 7, all real chains eventually collapse onto Eq. 8 for sufficiently small channels. As was the case with the chain extension, the prefactor computed by Burkhardt et al.[3] is in remarkable agreement with our simulation data.

The peak in the variance of the chain extension around $D \approx l_\mathrm{p}$ might lead one to think that it represents a discontinuity in the second derivative of the confinement free energy. This effect, coupled with the smoothness of the confinement free energy[4] and the mean extension over the same range of channel sizes, may lead one to speculate further that the blob-to-deflection segment transition is a second-order phase transition. This supposition is supported by the difficulty in obtaining variance data in the long-chain limit (see Supporting Information). Indeed, the confinement free energy reaches the long chain limit very quickly[4] compared to the variance in extension. However, the Odijk scaling theory makes it clear that no such phase transition exists at $D \approx l_\mathrm{p}$. The variance in the extension increases exponentially for small channel sizes through the scaling of Eqs. 12 and 19. However, the variance in the extension only diverges for ideal chains ($w = 0$), and this divergence only takes place in the limit of vanishing channel size. As a result, our current understanding of the Odijk theory and our simulation data suggest that there is no phase transition between blobs and deflection segments.

5 Discussion

5.1 Possible Experimental Tests of the Odijk Theory

From a practical standpoint, is it possible to verify the backfolded Odijk regime experimentally? To help answer this question, we computed the channel size corresponding to $\xi = 1$

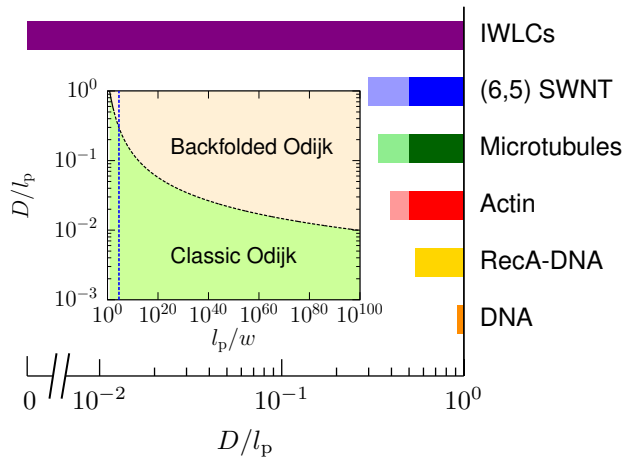

Figure 9: The range of the backfolded Odijk regime for DNA in a high ionic strength buffer ($l_p/w = 10$),[4] RecA-DNA complex ($l_p/w = 230$ assuming $w = 5\,\text{nm}$),[49] actin ($l_p/w = 2375$),[30,31] microtubules ($l_p/w = 9600$)[31,50] and (6,5) single-walled carbon nanotubes (SWNT, $l_p/w = 34210$).[51] For ideal wormlike chains, $l_p/w = \infty$ and the backfolded Odijk regime extends all the way down to $D/l_p = 0$. Darker shade indicates the region in which we have simulation data for g, while lighter shade signifies the region wherein Eq. 23 has been extrapolated to obtain the value of g. The inset shows the range of the Odijk regimes as a function of the value of monomer aspect ratio, l_p/w. The vertical blue line corresponds to the (6,5) SWNT, the wormlike macromolecule with the highest l_p/w value considered here. We set the upper limit of the backfolded Odijk regime to the scaling theory value of $D/l_p = 1$ for simplicity, although this limit is closer to $D/l_p \approx 2$.

for a number of experimentally relevant wormlike chains, from the fairly flexible example of double-stranded DNA to carbon nanotubes. In most cases, we expect the backfolded Odijk regime to extend below the channel sizes where we obtained data for g. In these instances, we have extrapolated our result in Eq. 23 to smaller channel sizes, with the caveat that this extrapolation must fail at some point.

As we can see in Figure 9, the backfolded Odijk regime occupies a very narrow range in channel sizes due to the exponential increase in the global persistence length as the channel size decreases, even for extremely stiff molecules such as carbon nanotubes. The inset of Figure 9 highlights the narrowness of the backfolded Odijk regime; in order for the backfolded Odijk regime to span two decades in channel size, we would need to have a monomer anisotropy $l_p/w = O(10^{100})$. This is clearly well outside the range of any of the typical materials used to study channel confinement. Likewise, while DNA ($l_p/w = 10$) is

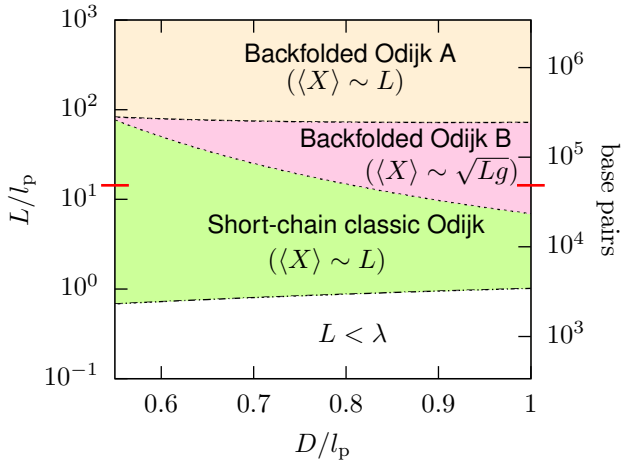

Figure 10: Scaling of the end-to-end distance as a function of contour length of RecA-DNA (l_p/w = 210).[49] The red lines on the y-axes correspond to λ-DNA, which is most commonly used in experiments.

considered a stiff molecule in many polymer physics situations, the l_p/w ratio of DNA in this context is so low that the backfolded Odijk regime is almost non-existent. The compression of the backfolded Odijk regime thus poses a significant problem for experiments if we use the channel size as the control parameter; we can explore a wide range of ξ values over a narrow range of channel sizes, but doing so in experiments requires an exquisite control over the channel size that is unlikely to be achieved[1] for the persistence length of DNA. We can open up a wider range of channel sizes by increasing the persistence length, provided that wormlike chains with these larger persistence lengths are available with contour lengths such that $L \gg g$.

While it seems unlikely that the backfolded Odijk regime will be verified by experiments that involve changing the channel size, it may be possible to test Odijk's theory experimentally instead by fixing the channel size and changing the molecular weight of the confined chain. For example, consider the phase diagram for RecA-DNA in Figure 10. Relatively short RecA-coated DNA has recently been used to study confined chains in the Odijk regime.[49] To construct this phase diagram, we assumed that the effective width of RecA-coated DNA is the same as naked DNA in the absence of a better estimate. For chains with $L < g$, we would expect to observe the classic Odijk regime ($\langle X\rangle \sim L$) because there are not a sufficient

number of deflection segments to produce a global hairpin. This appears to be the case in experiments so far using RecA-DNA[49] (and actin filaments[29]) based on the microscopy images. However, if we increase the contour length, such stiff molecules would eventually exhibit a regime with backfolding and weak excluded volume interactions (i.e., $g < L < L_*$), where $\langle X \rangle \sim L^{1/2}$. A further increase in contour length would lead to an onset of excluded volume interactions and a return to the scaling $\langle X \rangle \sim L$ for the backfolded Odijk regime. Counterintuitively, this test of the Odijk regime is best done in relatively wide channels ($D \approx l_p$) rather than very small channels; as we see in Figure 10, the range of molecular weights with $\langle X \rangle \sim L^{1/2}$ widens substantially as the channel size increases. Indeed, with the estimates we used to construct Figure 10, it may be possible to observe the first transition using RecA-coated DNA at reasonable molecular weights. It seems unlikely that we would reach the transition back to $\langle X \rangle \sim L$ at L_*, since this would require RecA-coated DNA of several hundred thousand base pairs. The upside of such an experiment is that, in order to observe a reasonable range of molecular weights where $g < L < L_*$, the experiment needs to be performed in a channel close to the persistence length. Thus, the fabrication of the experimental system should be relatively easy, with the biochemistry being the limiting step.

5.2 Implications for Genomic Mapping

While the exponential increase in the global persistence length as the channel size decreases makes fundamental studies of the different Odijk regimes difficult, it is critical to recent successes using 45 nm channels for genomic mapping of DNA.[15,52] In nanochannel mapping,[53,54] DNA are labeled with sequence specific probes and stretched by injecting them into a channel. The physical distances between neighboring probes are measured using fluorescence microscopy, and these data can be converted to genomic distances from the fractional extension of the chain. When implemented in a massively parallel way,[15] this method can be used to obtain information on structural variations in genomes that are difficult to assess by other techniques.

Clearly, the key to successful genomic mapping is to obtain linearly ordered probes; if the probes are disordered due to backfolding of the chain, this may be interpreted as a genomic reorganization rather than just a physical reorganization inside the channel. Based on Odijk's theory, there are two strategies to obtain linearly ordered deflection segments. In the first strategy, we can work with a chain in the Odijk regime $D \lesssim l_{\rm p}$ so long as the chain length satisfies $L \ll g$, even if $\xi < 1$. This approach will lead to ordered deflection segments and, in principle, can be implemented in circa 100 nm channels. Unfortunately, working with such small fragments of DNA makes genome assembly more difficult and leads to a large variance in chain extension.[55] The second strategy involves working with very long molecules, hundreds of kilobases in size, with the largest value of ξ achievable in experiments. It is obvious that forming a global hairpin is a problem, since it leads to scrambling of the genomic information. However, since global hairpin formation is a stochastic process, it is possible to get linearly ordered genomic data from a single molecule that did not happen to form a global hairpin. While this appears to be a possible strategy, we know from Odjik's theory that there will still be large fluctuations in the distance between barcodes *in the absence* of a global hairpin. It is the second derivative of the Flory free energy in Eq. 10 that drives the fluctuations, which is affected by the statistical probability of forming a hairpin but does not require that such a hairpin actually form. This theoretical insight is supported by the evolution of the genome mapping technology, where large fluctuations[55] motivated the switch from 100 nm channels[54] to 45 nm channels[15] in the current commercial technology. Once the system is firmly entrenched in the classic Odijk regime, not only are the deflection segments linearly ordered but the variance in the extension also drops precipitiously. As can be seen from Figure 7(a) a small decrease in channel size going from the backfolded Odijk regime to the classic Odijk regime can result in a decrease in variance by 3 to 4 orders of magnitude.

Since DNA is a polyelectrolyte, its persistence length and effective width can be manipulated by changing the ionic strength of the buffer. Thus, in an experimental system, the

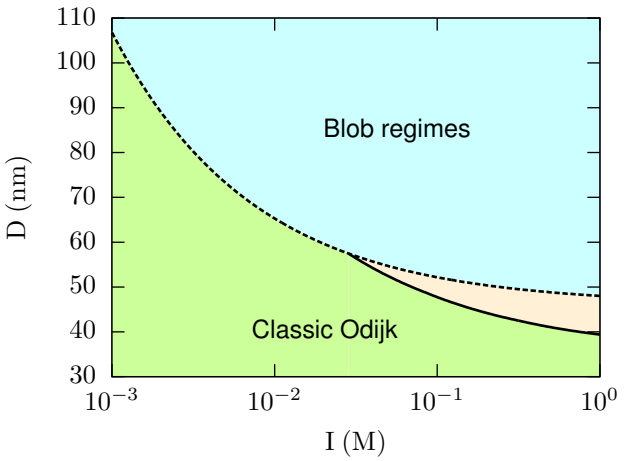

Figure 11: Phase diagram for DNA confinement in a nanochannel as a function of channel size and the ionic strength of the buffer. The onset of the blob regimes is denoted as $D = l_p$ and the transition between Odijk regimes is at $\xi = 1$ using Eq. 23. The beige-shaded region is the backfolded Odijk regime.

extent of confinement can be controlled by either changing the channel size or changing the ionic strength of the system.[28,53,56] Figure 11 provides a phase diagram for DNA in channel confinement as a function of ionic strength and channel size. To construct this diagram, we used Dobrynin's theory[57] for the persistence length of DNA[58] and Stigter's theory[59] for the effective width. Our results indicate that double-stranded DNA rapidly reaches the classic Odijk regime and thus should not exhibit many hairpins in 45 nm channels at 100 mM salt, which is similar to the point where the genomic mapping technology operates. Hairpin formation could still occur due to kinetically trapped, frozen hairpin states as the DNA enters the nanochannel.[60] Our work certainly did not account for such cases, as our results are restricted to equilibrium thermodynamics.

5.3 Implications for Theories of the Transition Regime

Our results strongly support Odijk's theory of two deflection segment regimes, one with linearly ordered deflection segments and a second with backfolding. In this sense, they clarify the confusion in the literature surrounding the so-called "transition regime" for DNA in a nanochannel. Only by considering a wide range of monomer anisotropies, up to approxi-

mately fifteen times stiffer than DNA, using many beads per persistence length, were we able to observe the generality of Odijk's scaling theory in Figures 5 and 7. We thus recognize many of the prior generalizations related to the transition regime are extrapolations of artifacts arising from the modest monomer anisotropy of DNA and the limited chain lengths used in previous work.

With respect to the ideal blob (Gauss-de Gennes) model,[4,16,61] there is indeed an apparent exponent $\langle X \rangle \sim D^{-1}$ at the start of the backfolded Odijk regime in Figure 4 for monomer anisotropies similar to DNA[4,7,11,13] using many different polymer models. Moreover, one of the striking results from Dai et al.[7] is the collapse of the data in the transition regime using the ordinate and abscissa of Figure 4, which we initially thought might support a regime with the scaling

$$\frac{\langle X \rangle}{L} \sim \left(\frac{D}{l_\mathrm{p}}\right)^{x} \left(\frac{l_\mathrm{p}}{w}\right)^{1/\nu - 2}. \tag{25}$$

However, in order for Eq. 25 to be congruent with Odijk's theory with a scaling exponent $x = -1$, the global persistence length would need to follow the power law $g/l_\mathrm{p} \sim (l_\mathrm{p}/D)^{4/3}$. This is clearly not the case from the simulation data in Figure 3, even for channel sizes $D \approx l_\mathrm{p}$.

It is thus natural to wonder why the data collapse in both the blob regimes and the backfolded Odijk regime using the approach in Figure 4. First, note that the exponent in the ordinate of Figure 4 was chosen to ensure collapse in the blob regimes.[7] If we use the the classical value of the Flory exponent, $\nu = 3/5$, then $1/\nu - 2 = -1/3$. Second, the mean extension in the backfolded Odijk regime (Eq. 11) can be rewritten as

$$\frac{\langle X \rangle}{L} \left(\frac{w}{l_\mathrm{p}}\right)^{-1/3} \simeq \left(\frac{g}{l_\mathrm{p}}\right)^{1/3} \left(\frac{l_\mathrm{p}}{D}\right)^{5/9} = f\left(\frac{D}{l_\mathrm{p}}\right), \tag{26}$$

where the left hand hand side is the ordinate of Figure 4 and f is some function of D/l_p. Thus, the collapse of the data for different l_p/w values in Figure 4 in the range $D_\mathrm{eff} \approx l_\mathrm{p}$ to $D_\mathrm{eff} \approx 2l_\mathrm{p}$ is a serendipitous consequence of the scaling, $\langle X \rangle \sim \xi^{1/3}$, in the backfolded Odijk

regime rather than evidence of a scaling law for $\langle X \rangle$ in terms of D/l_p.

Unlike the ideal-blob model, the cooperative backfolding model of Dai et al.[18] is not directly contradicted by the results presented herein. Furthermore, Odijk's theory clarifies the region in phase space where the cooperativity model is potentially applicable. In the cooperativity model, the extension of the chain proximate to the classic Odijk regime is described by conformations consisting of regions of S-loops, and regions of linearly ordered deflection segments.[18] This premise is consistent with the physics of the transition at $\xi \approx 1$ in Odijk's theory.[10] For $\xi \ll 1$, the chain can have many hairpins present simultaneously at some channel location x. Indeed, this is the basis for the Z parameter in Eq. 5. As ξ increases, the ability to backfold decreases. As a result, the number of hairpins present at a cross-sectional slice, x, should decrease, with an S-loop being the limiting case when $\xi \approx 1$.

In the context of elucidating the universal behavior of a confined wormlike chain, modeling the transitions between regimes is generally not a priority. However, in the context of DNA in a nanochannel, Fig. 9 shows that the backfolded Odijk regime is almost entirely suppressed and Fig. 11 shows that the genomic mapping technologies lie near $\xi = 1$. Thus, there is a practical impetus to develop a model for the transition to the classic Odijk regime for DNA. Intriguingly, Fig. 5 suggests that ξ may still be the relevant parameter for describing the extension even as ξ approaches unity. While the extension data in Fig. 5 no longer follow the scaling law $\langle X/L \rangle \sim \xi^{1/3}$ as $\xi \to 1$, the data still collapse when plotted as a function of ξ well past $\xi = 1$. While the Flory theory in Eq. 10 is not the correct approach to describing the transition at $\xi = 1$, the physics embodied in ξ (and, more importantly, the global persistence length g) suggests the possibility of a more fundamental, predictive model of the transition to the classic Odijk regime at $\xi \approx 1$.

5.4 Comparison with Slit Confinement

It is worthwhile to briefly compare Odijk's theory[10] for the extreme confinement limits, $D \lesssim l_p$ to recent results for confinement in slits[20] of size $H \lesssim l_p$. Similar to the case in

channels, slit confinement also leads to different types of behavior in strong confinement depending on the chain length, channel size and excluded volume interactions. One of the advantages in describing slit confinement is that we know that the global persistence length in a slit is bounded between $g = 2l_p$, i.e. the persistence length of a wormlike chain in two dimensions, and $g = l_p$.

The situation for short chains ($\lambda \ll L \ll g$) is identical for slits and channels; the walls orient the stiff chain, leading to an extension of the form of Eq. 7. The key difference is the upper bound in molecular weight; since the global persistence length grows rapidly in a channel as its cross-section decreases, the short-chain behavior persists up to much higher molecular weights in a channel than it does in a slit.

As the chain length increases beyond the global persistence length, both slit-confined and channel-confined chains undergo ideal random walks in the unconfined dimension(s) provided that the chain is thin enough so that excluded volume interactions are weak ($\xi < 1$). The random walk in the channel is one-dimensional, while the random walk in the slit is two-dimensional, but the scaling of the size of the chain with molecular weight in both cases is $L^{1/2}$. In both slits and channels, the ideality of the random walk arises because the chains can bend along the unconfined dimension(s) but they are too short to exhibit substantial excluded volume interactions. Naturally, in both slits and channels, this scaling persists to high molecular weights as $w \to 0$, since the chain can bend many times without experiencing substantial excluded volume interactions.

The key distinction between slits and channels takes place for long chains. In the slit case, there exists a single long chain regime.[20] Here, chains with contour length $L \gg Hl_p/w$ cross-over to a self-avoiding random walk in two dimensions, where the RMS end-to-end distance scales like $R \sim L^{3/4}$. In channels, Odijk's theory[10] also predicts that chains with $L > L_*$ and $\xi < 1$ have the scaling for a self-avoiding random walk in one-dimension, namely $\langle X \rangle \sim L$. Likewise, chains with $\xi > 1$ always have this extensive scaling, independent of L, as long as $L > \lambda$. However, the dependence of $\langle X \rangle$ on the parameters w, l_p, and D depends

on whether excluded volume interactions are strong ($\xi > 1$) or weak ($\xi < 1$).

6 Concluding Remarks

In the present contribution, we have used large scale PERM simulations of an ideal wormlike chain model to correct Odijk's theory for the global persistence length[19] to account for fluctuation effects and other approximations down to channel sizes $D/l_p = 0.5$. This result provided the closure for Odijk's scaling theory for confined wormlike chains.[10] The theory correctly predicts the confinement free energy,[4] average chain extension, and the variance in the extension, as well as the point at which excluded volume interactions suppress backfolding. We have tested his theory from $l_p/w = 8.45$ to $l_p/w = 145$ using a simulation model that resolves highly anisotropic chains with many beads per persistence length.

The concordance between the results presented here and Odijk's theory implies that prior models of the so-called transition regime,[4,16,18] while in agreement with simulation data for a limited range of monomer anisotropies, are not the correct description for the full range of parameters. In fact, the sharp change in extension over a narrow range of channel sizes in the transition regime is a result of the rapid increase of the global persistence length, as one enters the backfolded Odijk regime from the extended de Gennes regime. Consequently, given the confusion disseminated by the term "transition regime", we propose abandoning it in favor of the label "backfolded Odijk" regime for the universal regime when $\xi \ll 1$ and reserve the term "transition" for crossovers between universal regimes.

Our results, when combined with recent work on the extended de Gennes regime,[7,8] validate Odijk's description of channel-confined wormlike chains[10] in Figure 12. While the classic Odijk regime[2] and the de Gennes regime[5] have been accepted dogma for several decades, they represent limiting cases that are difficult to access experimentally and thus not a good description of typical experimental data.[1,25,62,63] Although the effect of excluded volume was not considered in the original theory of the classic Odijk regime,[2] our results

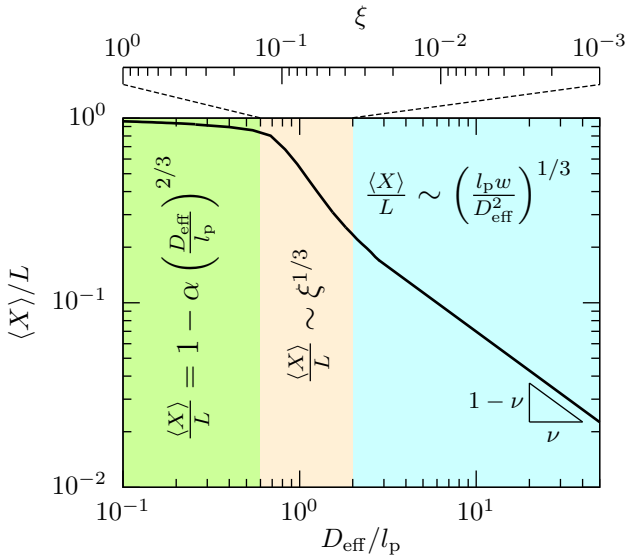

Figure 12: Summary of the regimes of confinement for a wormlike chain in a nanochannel and the corresponding fractional extension. The blue-shaded region corresponds to the de Gennes regime and the extended de Gennes regime, as they exhibit identical scaling of the extension. Note that the lower bound in ξ for the backfolded Odijk regime is a function of $w/l_{\rm p}$ and is approximately given by $0.63(w/l_{\rm p})$. However, we used the lower bound value of 10^{-3} only to illustrate that this regime can span many decades in ξ.

indicate that excluded volume is indeed necessary to suppress backfolding of long chains in this regime. Most experimental systems either lie in the extended de Gennes regime or, for a small range of channel sizes (but a large range of $\langle X\rangle$ and ξ), the backfolded Odijk regime. Moreover, since the transitions between confinement regimes are not sharp and the molecular weights required to reach the long-chain limit are large, a substantial portion of the experimental data do not lie firmly in any well defined regime.

While we feel that the Odijk theory[10] is now validated, this is not the first time that the problem of channel-confined wormlike chains has been declared nearly "solved." Naturally, the Odijk theory still requires an experimental test, and this may not be an easy task. Aside from the challenges we highlighted already concerning the narrowness of the backfolded Odijk regime and the need to use very high molecular weight chains, most experimental systems use rectangular channels due to fabrication limitations or the desire to increase experimental throughput via tapered nanochannels.[62] We recently found that such channels can exhibit

mixing of regimes, where the two channel dimensions lie in different regimes and thus affect the chain statistics differently.[25] We expect that the behavior of g in this case will be a non-trivial function of the aspect ratio of the channel, and the narrowness of the backfolded Odijk regime will make it difficult to isolate this regime. Thus, while simulation and theory now seem to be in accord, substantial challenges remain in the experimental area.

Acknowledgement

We thank Liang Dai and Patrick Doyle (MIT) for sharing their simulation data[18] and for useful discussions. We also thank Yanwei Wang (Soochow University) for early discussions on the Odijk theory. This work was supported by the National Science Foundation (CBET-1262286). In addition, D.R.T. thanks the University of Minnesota for support through a Doctoral Dissertation Fellowship. Computational resources were provided in part by the University of Minnesota Supercomputing Institute.

Supporting Information Available

Brief description of simulation methodology, results for order parameter as a function of channel size, calculation of global persistence length, comparision between global persistence length obtained by our method with Odijk's mechanical theory, contour length dependence on $\langle X \rangle$ and δX^2 and all extension data (including the blob regimes) plotted as a function of ξ. This material is available free of charge via the Internet at `http://pubs.acs.org/`.

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) Odijk, T. *Macromolecules* **1983**, *16*, 1340–1344.

(3) Burkhardt, T. W.; Yang, Y.; Gompper, G. *Phys. Rev. E* **2010**, *82*, 041801.

(4) Tree, D. R.; Wang, Y.; Dorfman, K. D. *Phys. Rev. Lett.* **2013**, *110*, 208103.

(5) Daoud, M.; de Gennes, P. G. *J. Phys.-Paris* **1977**, *38*, 85–93.

(6) Dai, L.; Doyle, P. S. *Macromolecules* **2013**, *46*, 6336–6344.

(7) Dai, L.; van der Maarel, J.; Doyle, P. S. *Macromolecules* **2014**, *47*, 2445–2450.

(8) Werner, E.; Mehlig, B. *arXiv:1404.6643 [cond-mat.soft]* **2014**,

(9) Brochard-Wyart, F.; Tanaka, T.; Borghi, N.; de Gennes, P. G. *Langmuir* **2005**, *21*, 4144–4148.

(10) Odijk, T. *Phys. Rev. E* **2008**, *77*, 060901.

(11) Wang, Y.; Tree, D. R.; Dorfman, K. D. *Macromolecules* **2011**, *44*, 6594–6604.

(12) Reisner, W.; Pedersen, J. N.; Austin, R. H. *Rep. Prog. Phys.* **2012**, *75*, 106601.

(13) Cifra, P.; Benková, Z.; Bleha, T. *J. Phys. Chem. B* **2009**, *113*, 1843–1851.

(14) Dorfman, K. D.; King, S. B.; Olson, D. W.; Thomas, J. D. P.; Tree, D. R. *Chem. Rev.* **2013**, *113*, 2584–2667.

(15) Lam, E. T.; Hastie, A.; Lin, C.; Ehrlich, D.; Das, S. K.; Austin, M. D.; Deshpande, P.; Cao, H.; Nagarajan, N.; Xiao, M.; Kwok, P.-Y. *Nat. Biotechnol.* **2012**, *30*, 771–776.

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

(17) Raphael, E.; Pincus, P. *J. Phys. II* **1992**, *2*, 1341–1344.

(18) Dai, L.; Ng, S. Y.; Doyle, P. S.; van der Maarel, J. R. C. *ACS Macro Lett.* **2012**, *1*, 1046–1050.

(19) Odijk, T. *J. Chem. Phys.* **2006**, *125*, 204904.

(20) Tree, D. R.; Reinhart, W. F.; Dorfman, K. D. *Macromolecules* **2014**, *47*, 3672–3684.

(21) Grassberger, P. *Phys. Rev. E* **1997**, *56*, 3682–3693.

(22) Prellberg, T.; Krawczyk, J. *Phys. Rev. Lett.* **2004**, *92*, 120602.

(23) Wang, J.; Gao, H. *J. Chem. Phys.* **2005**, *123*, 084906.

(24) Muralidhar, A.; Tree, D. R.; Wang, Y.; Dorfman, K. D. *J. Chem. Phys.* **2014**, *140*, 084905.

(25) Gupta, D.; Sheats, J.; Muralidhar, A.; Miller, J. J.; Huang, D. E.; Mahshid, S.; Dorfman, K. D.; Reisner, W. *J. Chem. Phys.* **2014**, *140*, 214901.

(26) Cifra, P. *J. Chem. Phys.* **2012**, *136*, 024902.

(27) Wagner, F.; Lattanzi, G.; Frey, E. *Phys. Rev. E* **2007**, *75*, 050902.

(28) Kim, Y.; Kim, K. S.; Kounovsky, K. L.; Chang, R.; Jung, G. Y.; dePablo, J. J.; Jo, K.; Schwartz, D. C. *Lab Chip* **2011**, *11*, 1721–1729.

(29) Nöding, B.; Köster, S. *Phys. Rev. Lett.* **2012**, *108*, 088101.

(30) Köster, S.; Stark, H.; Pfohl, T.; Kierfeld, J. *Biophys. Rev. Lett.* **2007**, *2*, 155–166.

(31) Köster, S.; Steinhauser, D.; Pfohl, T. *J. Phys.: Condens. Matter* **2005**, *17*, S4091.

(32) Rubinstein, M.; Colby, R. H. *Polymer Physics*; Oxford University Press: Oxford, 2003.

(33) Hsu, H.-P.; Paul, W.; Binder, K. *Europhys. Lett.* **2010**, *92*, 28003.

(34) Hsu, H.; Paul, W.; Binder, K. *Macromolecules* **2010**, *43*, 3094–3102.

(35) de Gennes, P. G. *Scaling Concepts in Polymer Physics*; Cornell University Press: Ithaca and London, 1979.

(36) Tree, D. R.; Muralidhar, A.; Doyle, P. S.; Dorfman, K. D. *Macromolecules* **2013**, *46*, 8369–8382.

(37) Vroege, G. J.; Odijk, T. *Macromolecules* **1988**, *21*, 2848–2858.

(38) Odijk, T. *J. Chem. Phys.* **1996**, *105*, 1270–1286.

(39) Spakowitz, A. J.; Wang, Z.-G. *J. Chem. Phys.* **2003**, *119*, 13113–13128.

(40) de Gennes, P. G. In *Polymer Liquid Crystals*; A. Ciferri, R. B. M., W. R. Krigbaum, Ed.; Academic Press, 1982; pp 115–131.

(41) Chen, J. Z. Y. *Macromolecules* **2013**, *46*, 9873–9844.

(42) Tkachenko, A.; Rabin, Y. *Macromolecules* **1995**, *28*, 8646–8656.

(43) Cifra, P.; Benková, Z.; Bleha, T. *J. Phys. Chem. B* **2008**, *112*, 1367–1375.

(44) Cifra, P.; Benková, Z.; Bleha, T. *Phys. Chem. Chem. Phys.* **2010**, *12*, 8934–8942.

(45) Clisby, N. *Phys. Rev. Lett.* **2010**, *104*, 055702.

(46) Tree, D. R.; Wang, Y.; Dorfman, K. D. *Phys. Rev. Lett.* **2012**, *108*, 228105.

(47) Yang, Y.; Burkhardt, T. W.; Gompper, G. *Phys. Rev. E* **2007**, *76*, 011804.

(48) Tree, D. R.; Wang, Y.; Dorfman, K. D. *Biomicrofluidics* **2013**, *7*, 054118.

(49) Frykholm, K.; Alizadehheidari, M.; Fritzsche, J.; Wigenius, J.; Modesti, M.; Persson, F.; Westerlund, F. *Small* **2014**, *10*, 884–887.

(50) van den Heuvel, M. G. L.; Bolhuis, S.; Dekker, C. *Nano Lett.* **2007**, *7*, 3138–3144.

(51) Fakhri, N.; MacKintosh, F. C.; Lounis, B.; Cognet, L.; Pasquali, M. *Science* **2010**, *330*, 1804–1807.

(52) Hastie, A. R.; Dong, L.; Smith, A.; Finkelstein, F.; Lam, E. T.; Huo, N.; Cao, H.; Kwok, P.-Y.; Deal, K. R.; Dvorak, J.; Luo, M.-C.; Gu, Y.; Xiao, M. *PLoS One* **2013**, *8*, e55864.

(53) Jo, K.; Dhingra, D.; Odijk, T.; De Pablo, J.; Graham, M.; Runnheim, R.; Forrest, D.; Schwartz, D. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 2673–2678.

(54) Das, S. K.; Austin, M. D.; Akana, M. C.; Deshpande, P.; Cao, H.; Xiao, M. *Nucleic Acids Res.* **2010**, *38*, e177.

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

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

(57) Dobrynin, A. *Macromolecules* **2006**, *39*, 9519–9527.

(58) Hsieh, C.-C.; Balducci, A.; Doyle, P. S. *Nano Lett.* **2008**, *8*, 1683–1688.

(59) Stigter, D. *Biopolymers* **1977**, *16*, 1435–1448.

(60) Levy, S. L.; Mannion, J. T.; Cheng, J.; Reccius, C. H.; Craighead, H. G. *Nano Lett.* **2008**, *8*, 3839–3844.

(61) Cifra, P.; Bleha, T. *Soft Matter* **2012**, *8*, 9022–9028.

(62) Persson, F.; Utko, P.; Reisner, W.; Larsen, N. B.; Kristensen, A. *Nano Lett.* **2009**, *9*, 1382–1385.

(63) Werner, E.; Westerlund, F.; Tegenfeldt, J. O.; Mehlig, B. *Macromolecules* **2013**, *46*, 6644–6650.

Graphical TOC Entry