Transition between two regimes describing internal fluctuation of DNA in a nanochannel

Tianxiang Su[1], Somes K. Das[2], Ming Xiao[2], Prashant K. Purohit[1,*]

1. Department of Mechanical Engineering and Applied Mechanics, University of Pennsylvania, Philadelphia, PA 19104, USA.
2. Bionanomatrix Inc, 3701 Market Street, 4th floor, Philadelphia, PA, 19104, USA.

∗ E-mail: purohit@seas.upenn.edu

Abstract

We measure the thermal fluctuation of the internal segments of a piece of DNA confined in a nanochannel about 50−100nm wide. This local thermodynamic property is key to accurate measurement of distances in genomic analysis. For DNA in ∼100nm channels, we observe a critical length scale ∼10μm for the mean extension of internal segments, below which the de Gennes' theory describes the fluctuations with no fitting parameters, and above which the fluctuation data falls into Odijk's deflection theory regime. By analyzing the probability distributions of the extensions of the internal segments, we infer that folded structures of length 150−250nm, separated by ∼10μm exist in the confined DNA during the transition between the two regimes. For ∼50nm channels we find that the fluctuation is significantly reduced since the Odijk regime appears earlier. This is critical for genomic analysis. We further propose a more detailed theory based on small fluctuations and incorporating the effects of confinement to explicitly calculate the statistical properties of the internal fluctuations. Our theory is applicable to polymers with heterogeneous mechanical properties confined in non-uniform channels. We show that existing theories for the end-to-end extension/fluctuation of polymers can be used to study the internal fluctuations only when the contour length of the polymer is many times larger than its persistence length. Finally, our results suggest that introducing nicks in the DNA will not change its fluctuation behavior when the nick density is below 1 nick per kbp DNA.

Introduction

Stretching DNA in nanochannels has emerged as an important technique for separating DNA, performing genome mapping, and also studying repressor-DNA interactions, *etc* [1–3]. On the other hand, DNA confined in nanochannels also serves as a simplified model for studying single polymer behavior in concentrated polymeric solutions and melts [4, 5]. For these reasons, mechanical behaviors of DNA inside nanochannels have attracted a long-standing interest. The two most well-known scaling theories in this field are those described by de Gennes [5] and by Odijk [6]. de Gennes' blob theory, which was later generalized by Schaefer and Pincus [7], assumes that the channel width D is much greater than the persistence length ξ_p of the polymer. It models the moderately confined DNA as a chain of spherical blobs inside a cylindrical channel and gives the following expression for the end-to-end extension $\langle x\rangle$ of the polymer [5, 7, 8]:

$$\frac{\langle x\rangle}{L} = A\left(\frac{w\,\xi_p}{D^2}\right)^{1/3}, \tag{1}$$

where L, w are the contour length and effective molecule width of the DNA respectively. The prefactor A is found to be close to unity [9]. Odijk's theory, on the other hand, works for DNA under strong confinement in which $D << \xi_p$. In this regime, the polymer is deflected back and forth by the channel

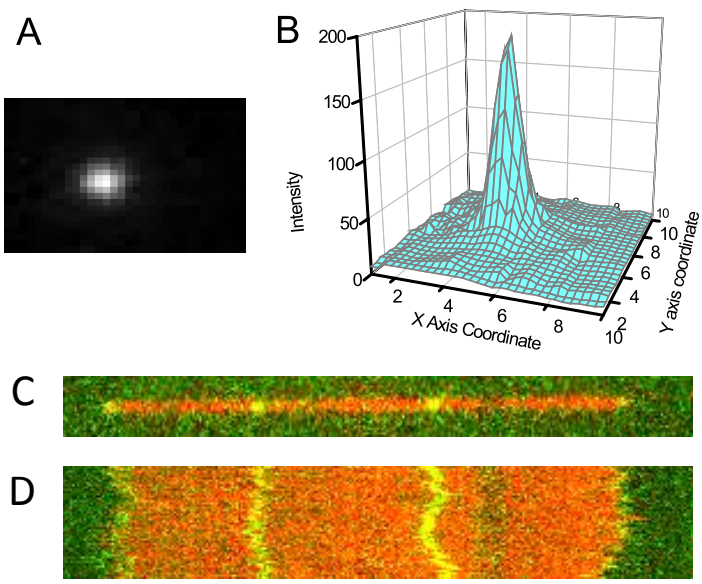

Figure 1. Measurement of the fluctuations of the internal segments of confined DNA. (A) Image of a dye label (Alexa-546) on a DNA backbone (backbone not shown) with 80ms exposure time. (B) 2D surface plot of the raw image (intensity of the dye vs. the X Y coordinates). (C) Image of one T4 DNA fragment ($\sim$ 36 microns) with backbone (red) and internal labels (green). (D) Time series (8 seconds) of the DNA showing the fluctuations of backbone and internal labels. In (D), the red trace is the backbone and the green traces are the trajectories of internal dye labels.

walls and the end-to-end extension is predicted to be [6]:

$$\frac{\langle x \rangle}{L} \approx 1 - \alpha_\circ \left(\frac{D}{\xi_p}\right)^{2/3}, \tag{2}$$

where $\alpha_\circ = 0.17$ is a constant whose value was determined recently by simulations [10]. Aside from the scaling theories, Wang and Gao [11] showed that the end-to-end extension of a strongly confined polymer in the Odijk regime can be derived analytically by modeling the confinement effect as a quadratic potential $U = 1/2\,\Xi|\vec{r}_\perp|^2$. Here Ξ is the stiffness of the effective quadratic potential, which depends on the channel width D, and $\vec{r}_\perp$ is the transverse displacement of the polymer from the axis of the nanochannel. Wang and Gao considered a confined chain under end-to-end applied force F and obtained an expression for the total extension $\langle x \rangle$ as a function of Ξ and F. We set $F = 0$pN, substitute the relation between Ξ and D (see Supporting Information) into their expression, and find:

$$\frac{\langle x \rangle}{L} = 1 - \frac{1}{5}\left(\frac{D}{\xi_p}\right)^{2/3}, \tag{3}$$

which is the same as Eq.2, confirming the scaling theory of Odijk, and at the same time validating the use of quadratic confinement potentials in the strongly confined regime.

Both de Gennes' and Odijk's theories have been tested by experiments as well as simulations over the years [10, 12–16]. However, most of the studies so far have focused on the properties of the entire DNA, for example, the *end-to-end* extension $\langle x \rangle$, the corresponding *end-to-end* fluctuation σ_x, and also the relaxation time τ of the *entire* DNA *etc.* Local properties of a confined polymer, on the other hand, like

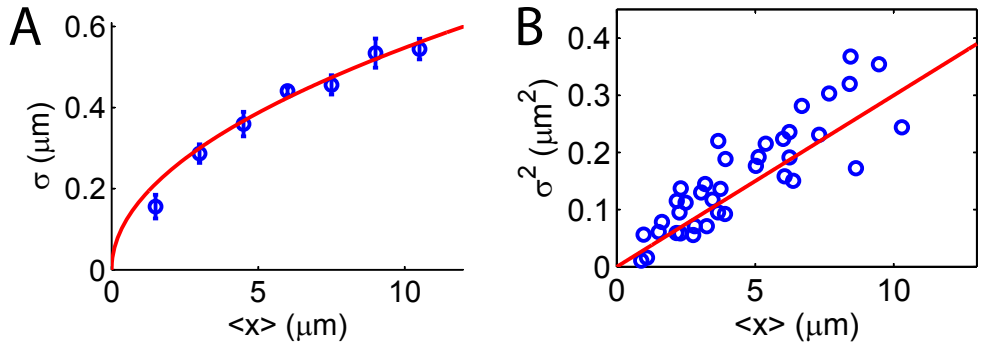

Figure 2. Internal fluctuation of λ DNA confined in a 80nm×130nm channel. (A) The measured rms fluctuation σ versus mean extension $\langle x\rangle$ for the internal segments of the DNA agrees very well with de Genne's theory with no fitting parameters (red curve, Eq.4). (B) A linear $\sigma^2 - \langle x\rangle$ profile confirms the 0.5 power law of $\sigma \sim \langle x\rangle^{1/2}$ of the de Gennes' theory. Note, however, that here we have maximum $\langle x\rangle \lesssim 10\mu$m. As shown in a subsequent figure (Fig.4) and in the text, for longer polymer with a maximum $\langle x\rangle \gtrsim 10\mu$m, the data deviates significantly from de Gennes' theory and even the 0.5 power law is lost.

the extension and fluctuation of its internal segments, are rarely investigated. In fact, local conformation and alignment of the confined DNA have been probed only recently [17,18]. It is also not well understood whether the existing theories developed for an entire piece of DNA can be applied locally for its internal segments. These are important issues because, if one considers the case of genome mapping, it is the local fluctuation of the internal segments that determines the resolution of the mapping.

In this paper, we measure the longitudinal *internal* fluctuation of a piece of DNA confined in rectangular channels about 50−100nm wide. We show that neither de Gennes' blob theory nor Odijk's deflection theory can completely describe the measured internal fluctuation versus mean extension profile. A critical length scale of $\sim 10\mu$m for the mean extension is observed, below which the internal DNA segments are more 'blob'-like, and above which Odijk's deflection theory works better. From the histograms of extension of the internal segments, we further infer that there exist folded structures of length 150−250nm separated by $\sim 10\mu$m along the backbone of the DNA during the transition between the two regimes. To justify the use of existing theories for studying the internal fluctuation, we focus on the Odijk regime and propose a method to explicitly calculate the internal fluctuation of a strongly confined DNA. We model the confinement effects by quadratic potentials and show that one can use the existing theories for end-to-end extension/fluctuation to describe the internal segments of the DNA when the contour length of the polymer is many times larger than its persistence length. Our model, which views the confined DNA as a discrete wormlike chain, can describe the fluctuations of heterogeneous polymers confined in non-uniform channels. It is also capable of capturing effects, like the influence of nicking sites on the DNA fluctuation profiles, which we will discuss at the end of the paper.

Results and Discussion

To visualize the internal segments, dye-labeled (Alexa-546) nucleotides are introduced into the backbones of the nicked λ DNA (48.5kbp, $L \approx 16.5\mu$m), $T4$ DNA (166kbp, $L \approx 56.4\mu$m) and bacterial artificial chromosome (BAC) human DNA clones (MCF7 BAC clone 9I10, fragmented, full length $\sim$ 180kbp, $L \approx 61.2\mu$m) (Fig.1) [19]. The DNA molecules are then driven by electric field into the nanochannels. With the Alexa-546 labels excited by light, extension of each internal segment is recorded frame-by-frame. Average extension $\langle x\rangle$ and the root mean square (rms) fluctuation $\sigma = \sqrt{\langle x^2\rangle - \langle x\rangle^2}$ for each internal

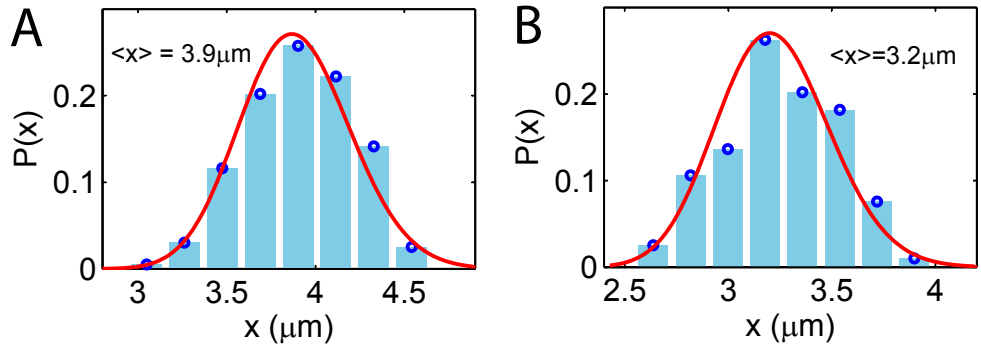

Figure 3. Probability distributions $P(x)$ for 2 internal segments of λ DNA inside a 80nm×130nm channel. The experimental data is fitted to Eq.7 (red). The fitting value C (Eq.7), when plugged back to Eq.6-2, recovers de Gennes's formula Eq.4.

segmenet are calculated and plotted in the $\sigma - \langle x \rangle$ profile.

In Fig.2, we first show the result for λ DNA confined in a 80nm×130nm channel. The maximum $\langle x \rangle$, which is roughly the mean extension of the entire DNA, is about 10μm, in agreement with the measurements of Tegenfeldt *et al* [12]. The internal fluctuation σ increases with $\langle x \rangle$ with a 0.5 power law. This 0.5 power law and even the magnitude of the fluctuation can be well captured by de Gennes' theory (discussed below) with no fitting parameters.

The longitudinal fluctuation of the confined DNA in de Gennes' theory can be evaluated using the effective stiffness k_{eff} of the polymer: $\sigma^2 = k_B T / k_{\text{eff}} \cong (4/15) L (w \xi_p D)^{1/3}$ [12, 20]. Using this expression and Eq.1 to eliminate L, we get the relation between σ and $\langle x \rangle$:

$$\sigma \cong \sqrt{\frac{4D}{15}} \cdot \sqrt{\langle x \rangle}. \tag{4}$$

Therefore, de Gennes' theory predicts a 0.5 power law for the $\sigma - \langle x \rangle$ profile. It is interesting to note that the prefactor in Eq.4 depends only on the channel width D, but not on the effective molecule width w, nor on the persistence length ξ_p. This implies that the $\sigma - \langle x \rangle$ profile is independent of the ionic strength of the experimental buffer. To compare the theory with the measured internal fluctuation, we plot Eq.4 together with the experimental data in Fig.2. Surprisingly, the data matches with the theory very well without any fitting parameters. Both the 0.5 power law and the magnitude of the fluctuation are correctly predicted by Eq.4.

de Gennes' theory also gives the distribution of the extension $P(x)$, which we can compare to our measurement. We consider the recently proposed "renormalized" Flory-type free energy $\mathcal{F}$ for a confined polymer [21] and its corresponding prediction of the longitudinal fluctuation:

$$\beta \mathcal{F} = A \frac{x^2}{(N/g)\, D^2} + B \frac{D (N/g)^2}{x}, \qquad \sigma^2 = \left(\frac{\partial^2 (\beta \mathcal{F})}{\partial x^2} \right)^{-1}, \tag{5}$$

where $\beta = 1/k_B T$, A, B are two constants, N, g are the total number of monomers and the number of monomers inside a blob respectively [21]. Both of the relations can be rewritten in terms of $\langle x \rangle$ (which is the solution of $\partial \mathcal{F} / \partial x = 0$) as:

$$\beta \mathcal{F} = C \left(\frac{x^2}{2D\langle x \rangle} + \frac{\langle x \rangle^2}{Dx} \right), \qquad \sigma = \sqrt{\frac{D}{3C}}\ \sqrt{\langle x \rangle}, \tag{6}$$

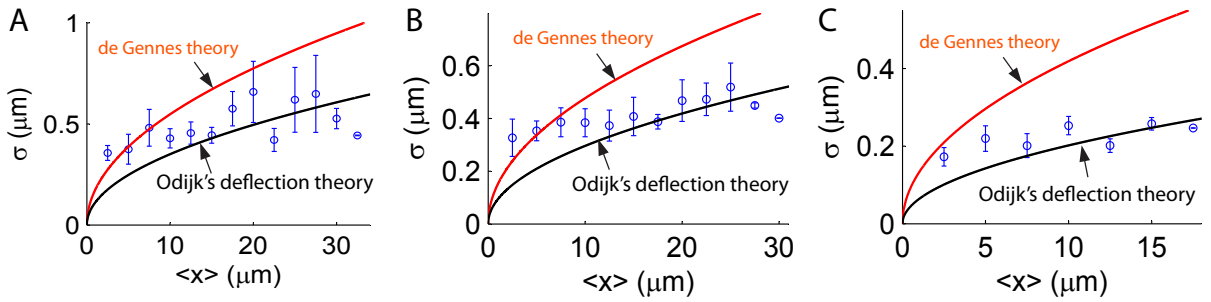

Figure 4. Fluctuation of the internal segments of (A) T4 DNA in 80nm×130nm, (B) T4 DNA in 60nm×100nm and (C) λ DNA in 50nm×70nm channels. For all cases, the maximum mean extension $\langle x\rangle > 10\mu$m. For (A) and (B), the data $\langle x\rangle \lesssim 10\mu$m agrees with de Gennes's theory (red, no fitting parameters). Deviation from de Gennes' theory begins at a critical $\langle x\rangle \sim 10\mu$m, above which the data falls into the black curve predicted by the deflection theories of Odijk [6], Wang and Gao [11]. For tighter channels (C), the transition occurs earlier with most data falling in the deflection regime.

with $C = (2A)^{2/3}B^{1/3}$ being a constant. The probability distribution $P(x)$ is therefore:

$$P(x) = P_0\ \exp(-\beta\mathcal{F}) = P_0 \exp\left[-C\left(\frac{x^2}{2D\langle x\rangle} + \frac{\langle x\rangle^2}{Dx}\right)\right]. \tag{7}$$

Here P_0 is a constant determined by the normalization condition. In our experiments, we record the extension x of each internal segment frame-by-frame and then calculate the distribution $P(x)$ for *each* segment. Fig.3 shows the measured $P(x)$ for two internal segments and their fitting results to Eq.7 (red). The result again implies that, for λ DNA confined in a 80nm×130nm channel, the behavior of the internal segments can be well captured by de Gennes' theory. Moreover, by fitting the distribution $P(x)$ to Eq.7, we obtain the constant C, which, when plugged back into Eq.6-2, yields: $\sigma = 0.58\sqrt{D}\sqrt{\langle x\rangle} \approx \sqrt{4D/15}\sqrt{\langle x\rangle}$ (here $D = \sqrt{80\times 130} = 102$nm). Therefore, starting from the "renormalized" Flory-type free energy Eq.5, we recover Eq.4 with the same prefactor. This indicates that the prefactor in Eq.4 is quite accurate although it is derived from a scaling theory. It also explains why Eq.4 matches with the measured $\sigma - \langle x\rangle$ profile without any fitting parameters (Fig.2). It is important to note that, for λ DNA confined in a 80nm×130nm channel, the maximum $\langle x\rangle$ is less than $\sim 10\mu$m (Fig.2). We shall show next that for longer DNA whose maximum $\langle x\rangle$ is greater than $\sim 10\mu$m, the measurement no longer agrees with de Gennes' theory. In particular, the 0.5 power law in the $\sigma - \langle x\rangle$ profile is lost.

Fig.4A shows the $\sigma - \langle x\rangle$ profile for the internal segments of T4 DNA in a 80nm×130nm channel. The maximum $\langle x\rangle$, which is roughly the mean extension of the entire DNA, is about 30μm, in agreement with the simulation result of Jung *et al* [14]. Fitting of $\sigma \sim \langle x\rangle^\gamma$ to the experimental data yields $\gamma = 0.19$, which is very different from the prediction of de Gennes' theory (Eq.4). Similar results are found for DNA in channels of different sizes: $\gamma = 0.15$ for T4 DNA confined in 60nm×100nm channels (Fig.4B) and $\gamma = 0.11$ for λ DNA in 50nm×70nm channels (Fig.4C). In all these cases the maximum $\langle x\rangle$ is greater than 10μm. We note, however, that in Fig.4, the experimental data for segments with $\langle x\rangle \lesssim 10\mu$m still matches with de Gennes' theory (except for the 50×70nm channel case, which we will explain later). It is the data with $\langle x\rangle \gtrsim 10\mu$m that deviates significantly from de Gennes' prediction. In fact, if we plot the fluctuation results for short segments with $\langle x\rangle \lesssim 10\mu$m for λ and T4 DNA together, the two profiles are almost identical, satisfying de Gennes' theory (see Supporting Information Fig. S1).

To rule out the possibility that the observed difference between λ DNA and T4 DNA stems from sequence variations, we perform the same experiments on the bacterial artificial chromosome (BAC) human DNA clones (MCF7 BAC clone 9I10), which also has maximum $\langle x\rangle \gtrsim 10\mu$m. As shown in Fig. 5, the results for the BAC DNA are almost identical to those for the T4 DNA. In particular, for

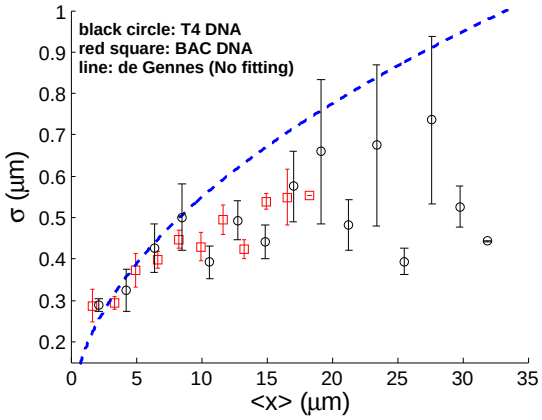

Figure 5. Internal fluctuation σ versus mean extension $\langle x \rangle$ for BAC (red squares) and T4 DNA (black circles) in a 80nm×130nm channel. This figure shows that DNAs from two different sources give almost identical results, which suggests that agreement with de Gennes theory for short internal segments, and deviation from de Gennes' theory for long internal segments, are both sequence independent.

small $\langle x \rangle < 10\mu$m, both match with de Gennes' prediction without any fitting parameters, while for $\langle x \rangle > 10\mu$m, both identically deviate from de Gennes' prediction. This suggests that the deviation from de Gennes' theory for long internal segments truly stems from segment size, not from sequence variations.

To better understand the deviation from de Gennes' prediction, we further look into the local structures of the confined DNA. Odijk showed recently that even in a 135nm channel, DNA can fold back on itself, giving rise to a global persist ence length much larger than 50nm, the intrinsic persistence length of the DNA [18, 22]. Because of this, Odjik argued that the transition from Odijk's regime to de Gennes' regime could be delayed with the increase of the channel size [18]. To check whether such local folded structures exist in the DNA in our experiments, we measure the extension distribution $P(x)$ for each single internal segment (see "Materials and Methods" for details). We find that for most internal segments whose mean extension is longer than 10μm, the distribution $P(x)$ shows two or more peaks (Fig.6B-C). From this observation, we infer that there indeed exist some folded structures in those internal segments – one peak in the distribution corresponds to the folded configuration, and the second peak corresponds to the extended configuration (Fig.6). The existence of folded structures can be also inferred from the typical extension x versus time plot as shown in Fig. 6D, where the steps in x correspond to different states of the internal segments. Furthermore, we find that in the distribution $P(x)$, the measured distances between any two peaks are always integral multiples of $400-500$nm, indicating that the difference in extension of a single folded structure and its extended form is about 500nm, ten times the persistence length of the DNA. This further implies that each branch of the folded structure is about $150-250$nm, if we assume each folded structure has two (loop) or three (hairpin) branches (Fig.6). Also, by checking the location of the internal segments that show multiple-peak distributions, we find that the folded structures are separated by $\sim 10\mu$m, which roughly agrees with the value of $\langle x \rangle$ above which de Gennes' theory fails to match with the experimental data (Fig.4). In the following we show that for $\langle x \rangle \gtrsim 10\mu$m the fluctuation data is better described by Odijk's deflection theory.

To exactly (rather than in a scaling sense) evaluate the fluctuation of DNA in the Odijk deflection regime, we extend the theory recently developed by Wang and Gao [11]. This theory represents the DNA as a strongly confined wormlike chain (fluctuating elastic rod) subjected to an additional end-to-end force F and produces the relation between the mean extension $\langle x \rangle$ and Ξ, the stiffness of the effective

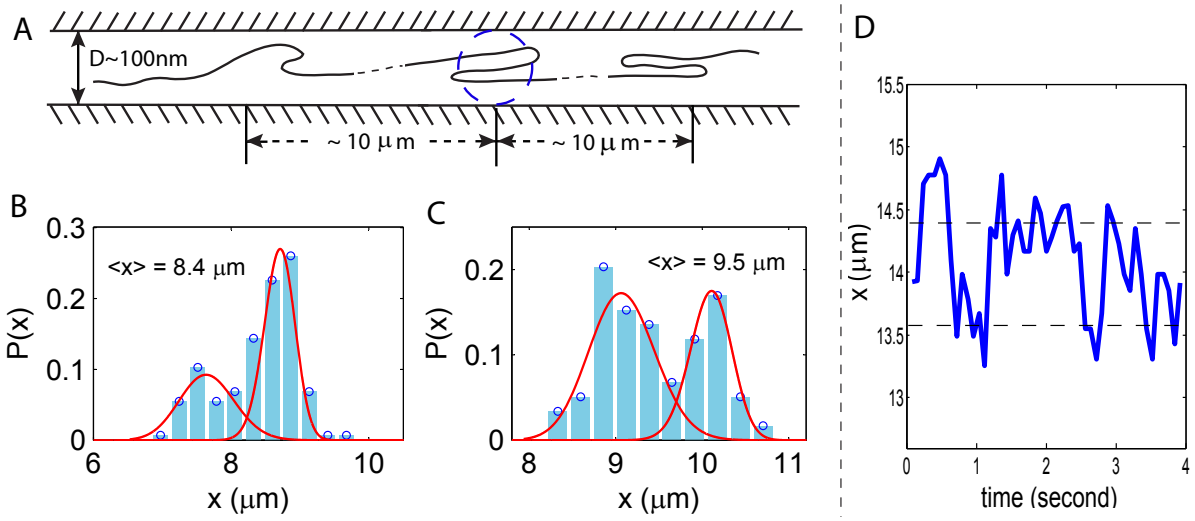

Figure 6. (A) Folded structures in the backbone of confined DNA. Each branch of the structure is about $150 - 250$nm, about the width of the channel size. The structures are separated by a distance $\sim 10\mu$m. (B, C) Distribution of extension $P(x)$ for 2 internal segments that contain the folded structures. In disagreement with de Gennes' prediction, the distributions show 2 peaks, from which we infer the existence of the folded structures. However, the structures are not stable as the two peaks in the distributions are comparable in height. The red curves fitted to the left peaks on the histogram are from de Gennes' theory (Eq.7) and the ones superimposed on the right peaks are from the deflection theory (Eq.10). (D) Extension x versus time for a single internal segment that shows two peaks in the distribution $P(x)$. The extension of this particular internal segment seems to fluctuate around two values shown by the dashed lines. This gives rise to the two peaks seen in the probability distribution.

confinement potential (which is a function of the channel width D):

$$\langle x\rangle = L - \frac{k_B T L}{2\sqrt{\kappa}} \frac{1}{\sqrt{F + 2\sqrt{\Xi(D)\kappa}}}, \tag{8}$$

where again, $k_B T$ is the thermal energy, κ is the bending modulus of the polymer, and in a rectangular channel the stiffness of the confinement potential can be expressed as $\Xi = 4c^4 \left[k_B T/(\kappa^{1/4} D^2)\right]^{4/3}$, with c being a constant. Using Eq.8, we calculate the effective stiffness of the DNA as $k_{\text{eff}} = (\partial\langle x\rangle/\partial F)^{-1}$, and then evaluate the fluctuation as $\sigma^2 = k_B T/k_{\text{eff}}$:

$$\sigma = \frac{D}{2\sqrt{8\xi_p c^3}} \cdot \left[1 - \frac{1}{4c}\left(\frac{D}{\xi_p}\right)^{2/3}\right]^{-1/2} \sqrt{\langle x\rangle}. \tag{9}$$

Leaving c as a free parameter, we fit Eq.9 to the experimental data with $\langle x\rangle > 10\mu$m in Fig.4A-C (black curves) and obtain $c = 1.03$, 0.94 and 0.99 respectively. For the BAC DNA confined in 80nm×130nm channels shown in Fig. 5, we obtain $c = 0.9$ from a similar fit. The fact that all the four sets of experimental data for different channel widths yield the same $c \approx 1$ makes sense because c is expected to be a universal constant independent of D. Moreover, the constant c comes from the expression for the free energy of confined chains in the Odijk regime and it has been estimated by Burkhardt to be $c = 1.1$ [23], which is very close to our fitting results. This strongly suggests that in the large mean extension regime $\langle x\rangle > 10\mu$m, the DNA segments are better described by the deflection theory.

Furthermore, from Fig.4A to C, we observe that the length of the error bars decreases with the decrease of the channel size. The reason for this may be that for moderately confined DNA, the local folded structures can form and unravel with comparable rates, as indicated by the similar height of the two peaks in the distribution in Fig.6B-C. Therefore, the behaviors of the confined polymer is a competition between de Gennes' type and Odijk type regimes and the error bar is large. As the channel size becomes smaller, Odijk's theory begins to dominate, resulting in smaller error bars.

By integrating the force-extension relation Eq.8, we obtain the free energy expression $G(x)$ in the Odijk (or Wang and Gao) deflection regime (see Supporting Information), which further leads to the distribution for the extension $P(x)$:

$$P(x) = P_0 \exp\left(Bx - \frac{A}{L - x}\right), \tag{10}$$

where $A = L^2/4\xi_p$, $B = 4c^2\xi_p^{1/3}/D^{4/3}$ and P_0 is the normalization factor. We fit this expression to the right peaks in Fig.6B-C and find that reasonable parameters ($L \approx 15\mu$m, $\xi_p \approx 50$nm) give excellent matches with the measured probability distributions in experiments. In fact, we can use this free energy expression to understand the transition from a different point of view. We note that the internal segments are expected to stay in the regime with lower free energy, and that regime transition occurs when the free energies in the two regimes are equal. By comparing the free energies in the two regimes, we draw a phase diagram on the $L - D$ plane in Fig. 7. The result shows that as D decreases, the transition length L decreases. Theoretically, the phase diagram involves an undetermined constant, which we fit such that transition occurs in the range $L \approx 8 - 12\mu$m when $D = 100$nm. Then the result shows that at $D = 60$nm, the transition length is $3 - 5\mu$m, which roughly agrees with our experimental result for λ DNA in a 50nm×70nm channel (Fig. 4C). The phase diagram shows that transition from de Gennes' to Odijk's regime can occur when D decreases with L fixed, or when L increases with D fixed.

We also measure the end-to-end extension for DNA with different lengths (longer than 10 microns) in a 60nm×100nm channel and the result agrees with Odijk's theory (Fig. S3).

In the above analysis, we have applied the theories (de Gennes, Odijk, Wang and Gao) for the *end-to-end* extension/fluctuation to evaluate the *internal, or local* extension/fluctuation of a confined DNA.

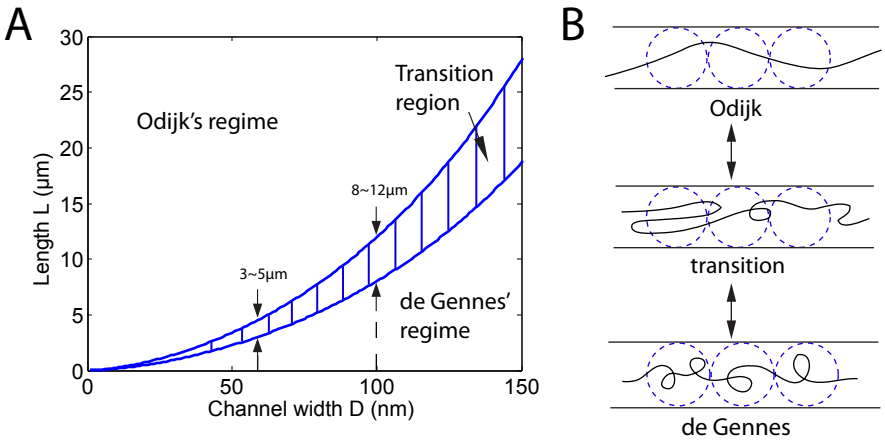

Figure 7. (A) Phase diagram showing two regimes on the $L - D$ plane, assuming $\xi_p = 50$nm for DNA. Transition from de Gennes' to Odijk's regime can occur when D decreases with L fixed, or when L increases with D fixed. (B) DNA with local folded structures as an intermediate state between de Gennes's and Odijk's regimes. In experiments, we observe heterogeneity in the intensity profile of YOYO-1 dye along the backbone of a confined DNA, which suggests the existence of the local folded structures (see Supporting Information Fig. S2).

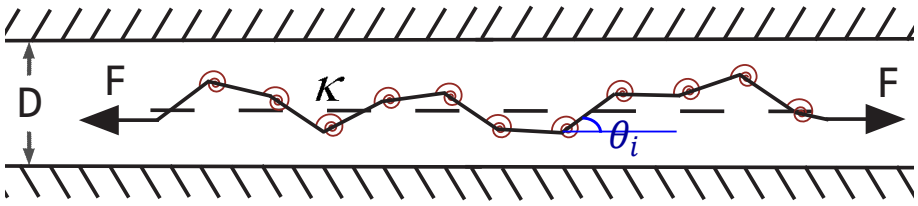

Figure 8. Discrete wormlike chain model for confined DNA in a nanochannel. The confined wormlike chain, subjected to and end-to-end applied force in general, has bending energy represented by a spring of stiffness κ at each node.

The assumption behind this is that when the internal segments are much longer than the persistence length of the DNA, the behavior of the segments is not very different from that of the entire DNA (with the same length) because the boundary conditions do not play a significant role [24–26]. To verify such an assumption, we explicitly calculate the internal fluctuation in Odijk's regime by extending a theory we developed earlier [26], and then compare our results to the theories developed for an entire piece of DNA.

Following the procedure in ref. [26], we model the polymer as a confined discrete $N-$segment wormlike chain, or fluctuating elastic rod (Fig.8). The Hamiltonian consists of 3 terms (Eq.11): (1) bending energy, (2) confinement energy, and (3) potential energy of an end-to-end applied force as shown in Fig.8.

$$\mathcal{H} = \int_0^L \frac{\kappa(s)}{2} \left| \frac{d\hat{t}}{ds} \right|^2 ds + \int_0^L \frac{\Xi}{2} y^2 ds - F\Delta x \tag{11}$$

$$= \frac{1}{2}\vec{\theta}^T \cdot \mathbf{K}\vec{\theta} - FL. \tag{12}$$

In the bending energy term, $\kappa(s)$ is the bending modulus of the DNA and it can vary along the arc length s so that the polymer is not necessarily homogeneous in mechanical properties. $\hat{t}$ is the tangent

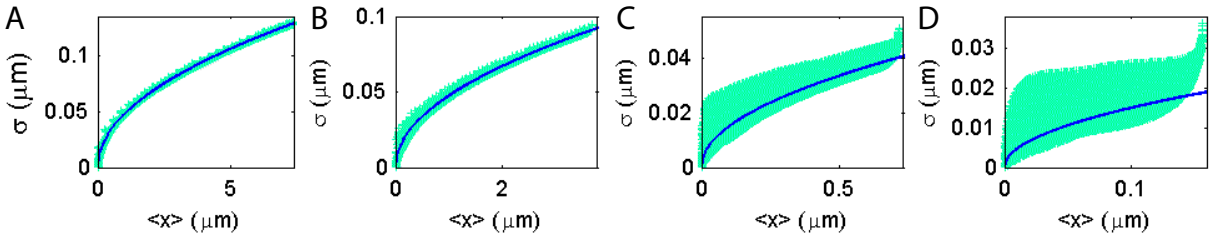

Figure 9. Fluctuation versus mean extension of internal segments of the strongly confined DNA in 60nm channels (Eq.13 and Eq.14). The contour lengths of the DNA are (A) $L = 10\mu$m, (B) 5μm, (C) 1μm and (D) 250nm. For a long DNA (A and B), data from internal segments of various locations of the chain collapse on the a curve with 0.5 power law (light green). The result agrees with Eq.9 (blue), which is derived for the end-to-end fluctuation of a confined DNA. For short DNA however (C and D), no power law is found as data from various locations of the chain do not collapse onto a single curve (light green). Therefore, formulae derived for the end-to-end fluctuation of the confined DNA, such as Eq.9 (blue), cannot be used for internal fluctuation. The boundary effect is so significant that the rms fluctuation σ not only depends on $\langle x \rangle$, but also on the location of the internal segments.

vector along the polymer. For the confinement potential term, we follow Wang and Gao's approach [11] and use an effective quadratic energy characterized by the coefficient Ξ, with y being the transverse displacement. In general, Ξ can be a function of the arc length s in case the confinement is not uniform. Also, for 3D chains in rectangular channels, Ξ can be different in the two transverse directions. For the potential energy term, we consider the chain subjected to an end-to-end force F, which can be set to zero if no force is applied. $\Delta x = x(L) - x(0)$ is the end-to-end extension of the chain. Up to a second order approximation, the Hamiltonian can be written in matrix form as shown in Eq.12, with θ_i being the discretized tangent angles and $\mathbf{K}$ being the $N \times N$ stiffness matrix of the chain [26].

It has been shown that when there are no constraints on twist (as is the case here), thermodynamic properties of a 3D chain can be easily generated from those of two 2D chains [26]. Therefore, for simplicity, here we describe the theory for 2D chains and plot the results for the corresponding 3D chains.

To get the internal fluctuation, we first need to calculate (1) the partition function, and (2) the angle fluctuation $\langle \theta_i \theta_j \rangle$. These are evaluated in the "Materials and Methods" section. Finally, for any internal segment between node i and node j of the discrete chain, the mean extension $\langle x_{ij} \rangle$ and the corresponding rms fluctuation can be explicitly calculated as:

$$\frac{\langle x_{ij} \rangle}{l} = (j - i) - \frac{\langle \theta_{i+1}^2 \rangle + \cdots + \langle \theta_j^2 \rangle}{2}, \tag{13}$$

$$\frac{\sigma_{ij}^2}{l^2} = \frac{\langle (x_j - x_i)^2 \rangle - \langle x_j - x_i \rangle^2}{l^2} = \frac{1}{2} \sum_{m=i+1}^{j} \sum_{n=i+1}^{j} \langle \theta_m \theta_n \rangle^2, \tag{14}$$

where l is the segment length of the discrete chain. In Fig.9, we consider DNA in 60nm×60nm channels and plot σ_{ij} versus $\langle x_{ij} \rangle$ for all the pairs of internal nodes (i, j) and see if the profiles match with the theories developed for the entire piece of DNA. Fig.9(A) shows the result for a chain with contour length $L = 10\mu$m, which is much larger than its persistence length $\xi_p = 50$nm. The internal fluctuation profile agrees exactly with Eq.9, which is derived for the end-to-end fluctuations. In particular, all the data collapses into a single curve with 0.5 power law. As the contour length of the polymer decreases, however, (Fig.9B-D), the internal fluctuation profile begins to scatter around the curve for the end-to-end fluctuation. This implies that, for short chains, the magnitude of internal fluctuation can be different, even if two internal segments have the same mean extension. The magnitude of the fluctuation depends

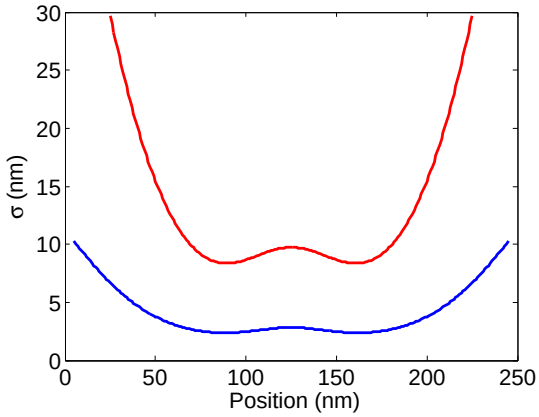

Figure 10. Fluctuation as a function as the position of an internal segment for a short chain. The contour length of the entire chain is short (250nm), so that the fluctuation not only depends on the length of the internal segment, but also on its position. Here we plot the fluctuation versus position for internal segments with the same size: 50nm (red) and 10nm (blue). For the internal segments close to the boundaries, the fluctuation is larger because they have more freedom compared to the segments inside the chain.

strongly on where the internal segment is located. In fact, we show in Fig. 10 that the internal segments located at the two boundaries have larger fluctuation because they have more freedom to fluctuate compared to the segments inside the chain. The strong boundary effects on short chains (such as, DNA with contour length 0.6-7μm) have been discussed by several groups recently [24–26]. Our results suggest that the accuracy of DNA sizing depends on the DNA contour length. For a short DNA with contour length $L < 1\mu$m confined in a 60nm×60nm channel, the uncertainty of the measurement will be high. For the experimental results we discussed earlier, the λ DNA, T4 DNA and BAC DNA all have contour lengths of tens of microns, for which boundary effects can be neglected. Therefore, it is safe to use the formulae for end-to-end extension/fluctuation to estimate the internal properties of the confined DNA in our experiments.

To measure the internal fluctuation, we have introduced nicks into the DNA so that internal sites along the DNA can be labeled. Since the theory discussed above allows for arbitrary bending modulus $\kappa(s)$ as a function of the arc length s, we can model the effect of nicking by setting $\kappa = 0$ on some nodes of the discrete chain and see whether the nicks have significant effects on the behavior of the DNA. For simplicity, we assume here that the nicks are equally spaced along the chain. Fig.11 shows that the fluctuation profile does not significantly deviate from the homogeneous chain with uniform κ when there are less than 50 nicks along a 18μm chain (∼50kbp DNA in a 60nm×60nm channel). In our experiments, the fluorescent tagging is introduced at the nicking endonuclease recognition sequence sites, which have much lower density than 1 nick/kbp in λ, T4 and BAC DNA. Therefore, the nicks will not significantly affect the DNA internal fluctuation.

To summarize, in this paper, we have investigated the thermal fluctuations of the internal segments of a piece of confined DNA in a nanochannel. The channel size is on the order of the persistence length of the DNA and we have compared the fluctuation data to several theories in literature. We have found that for channel widths on the order of 100nm there exists a critical length scale ∼10μm for the mean extension of an internal segment below which the de Gennes' theory describes the internal fluctuations and above which the data agree better with Odijk's deflection theory. For long DNAs confined in nanochannels we have inferred that there are folded structures whose branches are about 3 times the persistence length of DNA which are separated by segments with mean extension ∼10μm. We surmise that these folded

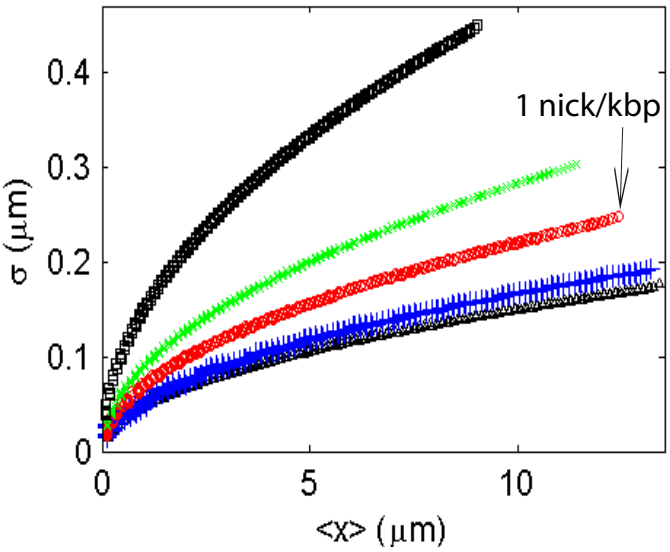

Figure 11. Fluctuation of a 18μm long chain with persistence length $\xi_p = 50$nm confined in a 60nm×60nm channel. From bottom to top: (1) △: no nicks; (2) +: 10 nick in 18μm; (3) ◯: 50 nicks in 18μm; (4) ×: 100 nicks in 18μm; (5) □: 200 nicks in 18μm. This figure shows that when the density of nicks is lower than 50 nicks per 18μm, or 1 nick per kbp of DNA, the fluctuation profile is almost the same as that for a chain without nicks.

structures are indicative of a transition from the Odijk regime, in which the DNA is relatively straight, to the deGennes regime, in which the DNA is more blob-like. We have also presented a more detailed theory based on small fluctuations and incorporating the effects of confinement. We have shown that one can use the existing theories for end-to-end extension/fluctuations to study the statistical properties of internal segments only when the contour length of the chain is much larger than the persistence length of the molecule so that boundary effects play no role. Our calculations suggest that introducing nicks into the DNA can change its fluctuation behavior if the density of nicks is greater than about 1 nick per kbp DNA.

Materials and Methods

Sequence specific labeling and DNA staining

In a 20μl reaction native, duplex DNA samples 50ng/μl (λ, T4 DNA and also MCF7 BAC clone 9I10) are incubated with 0.5U of Nb.BbvCI (0.5U/μl) (NEB, Ipswich, MA) in 1× NEB buffer 2 (NEB) for 1 hr at 37°C and 20 min at 65°C. The nicked DNA samples (12.5ng/μl) are then incubated for 30 min at 50°C in 1×NEB thermopol buffer with DNA polymerase Vent (exo-) (NEB) at 0.5U/μl in presence of a mixture of 75nM dAGC and 75nM Alexa-546 labeled dUTP. Then, the DNA (4ng/μl) samples are stained with intercalating dye YOYO-1 iodide at 1 dye molecule per 10 base pairs of DNA (Invitrogen Inc, Carlsbad, CA) in presence of 0.4M DTT (Promega Inc, Madison, WI).

Loading DNA into nanochannels

Fabrication of silicon based nanochannel chips has been described elsewhere [27,28]. The DNA sample is diluted by 2 times using the flow buffer consisting of 1×TBE, 3.6% Tween, and 10% Polyvinylpyrrolidone

(PVP). Ultrapure distilled water is used for making solutions (Invitrogen Corp., Ultrapure water). The DNA molecules are driven by electric field ($3-5$V) at the port of entrance of the chip and allowed to populate there for $2-3$ minutes [29]. Under higher voltage ($\sim$10V), the populated molecules are moved to the locos and then through the micro pillar structure of the chip to convert from a compact globular conformation to an open relaxed one. At the 300nm channel area the molecules adopt a more relaxed linear form with some heterogeneity on the backbone. With one end entering the nanochannel under the electric field, the DNA molecules elongate to a linear conformation with almost homogeneous backbone. Most of the structural heterogeneity progressively disappears as it interacted with the nanochannels, adopting fully confined equilibrium conformation after the field is off (relaxation time $10-15$s). A buffer consisting 0.5×TBE, 1.8% Tween 20, 5% PVP has been used to flow the DNA molecules resulting in a stretch of 65%.

Microscopy and image processing

The epi-fluorescence imaging is done in Olympus microscope (Model IX-71, Olympus America Inc, Melville, NY) using a 100×SAPO objective (Olympus SApo 100X/1.4 oil). YOYO-1, the DNA backbone staining dye ($\sim$491nm absorption, $\sim$509nm emission) is excited using 488nm laser (BCD1, Blue DDD Laser Systems, CVI Melles Griot, Rochester, NY) whereas Alexa-546 ($\sim$550nm absorption, $\sim$570nm emission) is excited using 543nm green laser (Voltex Inc, Colorado Springs, CO). The same filter cube consisting triple band dichroic and dual band pass emission filters (Z488/532/633rpc, z488/543m respectively) (Custom made, Chroma Technology Corp. Rockingham, VT) is used for detection of YOYO-1 and Alexa-546 emission by alternative laser excitation (using external laser shutters, Thorlabs, Newton, NJ). The emission signal is magnified 1.6× and detected by a back-illuminated, thermoelectric cooled charge coupled device (EMCCD) detector (iXon) (Andor, Ireland). About 200 sequential images of the labeled DNAs confined in nanochannels are recorded at $60-80$ms exposure time in blue-green alternative laser excitation.

Recording and calculations

The intensity profile $I(x,y)$ of each Alexa-546 label is fitted by a 2D Gaussian function to determine the position of the label (x_c, y_c) in the channel (Fig.1B). The position of each internal label is followed frame-by-frame at a time interval of about 160ms. The probability distribution, the mean value and the corresponding standard deviation of the distance between each pair of internal labels are calculated.

Partition function and angle fluctuation

The partition function for a confined DNA, whose Hamiltonian is expressed in Eq.12, is: $Z = \int \exp\left(-\mathcal{H}/k_BT\right) d\vec{\theta} = \exp\left(FL/k_BT\right) \sqrt{(2\pi k_BT)^N / \det \mathbf{K}}$, where N is the number of segments in the discrete chain. The angle fluctuation or correlation is the Boltzmann weighted average of $(\theta_i\theta_j)$ over all the configurations [26, 30]:

$$\langle \theta_i \theta_j \rangle = \frac{1}{Z} \int \theta_i \cdot \theta_j \, \exp\left(-\frac{\mathcal{H}}{k_BT}\right) d\vec{\theta} = k_BT \left(\mathbf{K}^{-1}\right)_{ij}. \tag{15}$$

Using Eq.15, we can explicitly calculate the mean extension and fluctuation of the internal segments (Eq.13-14).

Acknowledgments

References

1. Riehn R, Lu M, Wang YM, Lim SF, Cox EC, et al. (2005) Restriction mapping in nanofluidic devices. *P Natl Acad Sci USA* 102:10012-10016.

2. Douville N, Huh D, Takayama S (2008) DNA linearization through confinement in nanofluidic channels. *Anal Bioanal Chem* 391:2395-2409.

3. Wang YM, Tegenfeldt JO, Reisner W, Riehn R, Guan XJ, et al. (2005) Single-molecule studies of repressor-DNA interactions show long-range interactions. *P Natl Acad Sci USA* 102:9796-9801.

4. Doi M, Edwards SF (1986) The Theory of Polymer Dynamics. (Clarendon Press, Oxford).

5. de Gennes PG (1979) Scaling Concepts in Polymer Physics. (Cornell University Press, Ithaca, NY).

6. Odijk T (1983) On the statistics and dynamics of confined or entangled stiff polymers *Macromolecules* 16:1340-1344.

7. Schaefer DW, Joanny JF, Pincus P. (1980) Dynamics of semiflexible polymers in solution. *Macromolecules* 13:1280-1289.

8. Daoud M, Degennes PG (1977) Statistics of macromolecular solutions trapped in small pores. *J Phys France* 38:85-93

9. Reisner W, Beech JP, Larsen NB, Flyvbjerg H, Kristensen A, et al. (2007) Nanoconfinement-enhanced conformational response of single DNA molecules to changes in ionic environment. *Phys Rev Lett* 99:058302.

10. Yang Y, Burkhardt TW, Gompper G (2007) Free energy and extension of a semiflexible polymer in cylindrical confining geometries. *Phys Rev E* 76:011804.

11. Wang J, Gao H (2007) Stretching a stiff polymer in a tube. *J Mater Sci* 42:8838-8843.

12. Tegenfeldt JO, Prinz C, Cao H, Chou S, Reisner WW, et al. (2004) The dynamics of genomic-length DNA molecules in 100-nm channels. *P Natl Acad Sci USA* 101:10979-10983.

13. Reisner W, Morton KJ, Riehn R, Wang YM, Yu Z, et al. (2005) Statics and dynamics of single DNA molecules confined in nanochannels. *Phys Rev Lett* 94:196101.

14. Jung Y, Jun S, Ha B (2009) Self-avoiding polymer trapped inside a cylindrical pore: flory free energy and unexpected dynamics. *Phys Rev E* 79:061912.

15. Bonthuis DJ, Meyer C, Stein D, Dekker C (2008) Conformation and dynamics of DNA confined in slitlike nanofluidic channels. *Phys Rev Lett* 101:108303.

16. Cifra P, Benkova Z, Bleha T (2008) Effect of confinement on properties of stiff biological macromolecules. *Faraday Discuss* 139:377-392.

17. Persson F, Westerlund F, Tegenfeldt JO, Kristensen A (2009) Local conformation of confined DNA studied using emission polarization anisotropy. *Small* 5:190-193.

18. Odijk T (2006) DNA confined in nanochannels: hairpin tightening by entropic depletion. *J Chem Phys* 125:204904.

19. Xiao M, Phong A, Ha C, Chan TF, Cai D, et al. (2007) Rapid DNA mapping by fluorescent single molecule detection. *Nucleic Acids Res* 35:e16.

20. Persson F, Tegenfeldt JO (2010) DNA in nanochannels-directly visualizing genomic information. *Chem Soc Rev* 39:985-999.

21. Jun S, Thirumalai D, Ha B (2008) Compression and stretching of a self-avoiding chain in cylindrical nanopores. *Phys Rev Lett* 101:138101.

22. Odijk T (2008) Scaling theory of DNA confined in nanochannels and nanoslits. *Phys Rev E* 77:060901.

23. Burkhardt TW (1997) Free energy of a semiflexible polymer in a tube and statistics of a randomly-accelerated particle. *J Phys A-Math Gen* 30:L167-L172.

24. Seol Y, Li J, Nelson PC, Perkins TT, Betterton MD (2007) Elasticity of short DNA molecules: theory and experiment for contour lengths of 0.6-7 μm. *Biophys J* 93:4360-4373.

25. Purohit PK, Arsenault ME, Goldman Y, Bau HH (2008) The mechanics of short rod-like molecules in tension. *Int J Nonlin Mech* 43:1056-1063.

26. Su T, Purohit PK (2010) Thermomechanics of a heterogeneous fluctuating chain. *J Mech Phys Solids* 58:164-186.

27. Cao H, Yu ZN, Wang J, Tegenfeldt JO, Austin RH, et al. (2002) Fabrication of 10 nm enclosed nanofluidic channels. *Appl Phys Lett* 81:174-176.

28. Cao H, Tegenfeldt JO, Austin RH, Chou SY (2002) Gradient nanostructures for interfacing microfluidics and nanofluidics. *Appl Phys Lett* 81:3058-3060.

29. Das SK, Austin MD, Akana MC, Deshpande P, Cao H, et al. (2010) Single molecule linear analysis of DNA in nano-channel labeled with sequence specific fluorescent probes. *Nucleic Acids Res* 38:e177.

30. Zhang Y, Crothers DM (2003) Statistical mechanics of sequence-dependent circular DNA and its application for DNA cyclization. *Biophys J* 84:136-153.

Supporting Information: Transition between two regimes describing internal fluctuation of DNA in a nanochannel

Tianxiang Su[1], Somes K. Das[2], Ming Xiao[2], Prashant K. Purohit[1,*]

1. Department of Mechanical Engineering and Applied Mechanics, University of Pennsylvania, Philadelphia, PA 19104, USA.
2. Bionanomatrix Inc, 3701 Market Street, 4th floor, Philadelphia, PA, 19104, USA.

∗ E-mail: purohit@seas.upenn.edu

1 End-to-end extension $\langle x \rangle$ versus channel width D relation for Wang and Gao's theory

For a strongly confined DNA under force F, the total extension $\langle x \rangle$ as a function as F and Ξ (stiffness of the effective confinement potential) is found to be [1]:

$$\langle x \rangle = L - \frac{k_B T L}{2\sqrt{\kappa}} \frac{1}{\sqrt{F + 2\sqrt{\Xi(D)\kappa}}}, \tag{1}$$

where κ is the bending modulus of the polymer, Ξ relates to the channel width D in the following way:

$$\Xi = \frac{c^4}{4} \left(\frac{k_B T}{\kappa^{1/4} D^2} \right)^{4/3}, \tag{2}$$

and $c = 2.5$ is a constant for a cylindrical channel [1]. Setting $F = 0$ and plugging Eq.2 into Eq.1, we obtain:

$$\langle x \rangle = L \left[1 - \frac{1}{5} \left(\frac{D}{\xi_p} \right)^{2/3} \right]. \tag{3}$$

2 Fluctuation for short internal segments

The fluctuation for short internal segments is expected to be in the de Gennes' moderately confined regime. In Fig. S1, we plot the internal fluctuation profiles for short segments with $\langle x \rangle < 10\mu$m for 4 different sets of DNA: (1) λ DNA, (2) T4 DNA, (3) fragmented T4 DNA and (4) BAC DNA. Note that here we not only have short DNA, like λ DNA, but also long DNA like T4 and BAC DNA, but we discuss only the short internal segments on them in this section. The results for all the 4 sets of DNA are almost identical, and they all match with de Gennes' theory with NO fitting parameters. This result suggests that for all internal segments with $\langle x \rangle \leq 10\mu$m, irrespective of the sequence and length, de Gennes' theory works.

3 Heterogeneity on the backbone of DNA

In experiments, we observe heterogeneity in the intensity profile of the YOYO-1 dye along the backbone of the DNA (Fig.S2), which can be evidence for the formation of local folded structures.

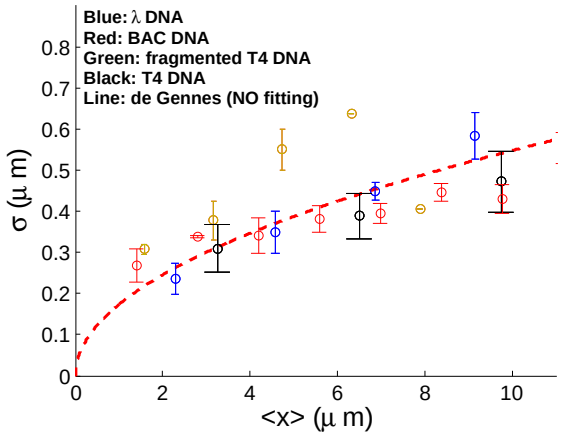

Figure S1: σ **versus** $\langle x \rangle$ **profile for the** $\langle x \rangle \leq 10\mu$**m region.** Fluctuation of short internal DNA segments from different sources matches with de Gennes' theory with NO fitting parameters.

For the left figure of Fig.S2, in the intensity profiles corresponding to 0s and 1.6s, localized peaks are clearly apparent (Fig.S2). These localized peaks could represent the deGennes' blobs or local folded structures. On the other hand, regions of uniform intensity could correspond to DNA in the Odijk regime. Heterogeneity in the DNA backbone fluorescence intensity is also shown in the right figure of Fig.S2. This figure shows two internal labels coming together, which is evidence for formation of local folded structure. Although these images gives us a visual picture of DNA confined to a nanochannel, we believe that the two-peak probability distributions shown in the main text provide much stronger evidence of the transition between the deGennes' and Odijk regimes than the fluorescence intensity profiles.

4 Total extension versus L relation

As another evidence that the deflection theory works for segments with $\langle x \rangle \gtrsim 10\mu$m, we measure the end-to-end extension for DNA with different lengths (but with mean end-to-end extension greater than 10 microns) in a 60nm×100nm channel and plot the result against the contour length (Fig.S3). A linear relation is found with a fitting result of $\langle x \rangle = 0.5L$. This is consistent with the deflection theories (formulae shown below), which, with numerical values plugged in, gives $\langle x \rangle \approx 0.7L$.

$$\frac{\langle x \rangle}{L} \approx 1 - \alpha_\circ \left(\frac{D}{\xi_p}\right)^{2/3}, \quad \alpha_\circ = 0.17, \quad \text{(Odijk, [2])} \tag{4}$$

$$\frac{\langle x \rangle}{L} = 1 - \frac{1}{5}\left(\frac{D}{\xi_p}\right)^{2/3}, \quad \text{(Wang and Gao, [1])} \tag{5}$$

5 Distribution of extension in the deflection regime

The force-extension relation for a 3D confined chain in Odijk's regime is given in Eq.1 [1].

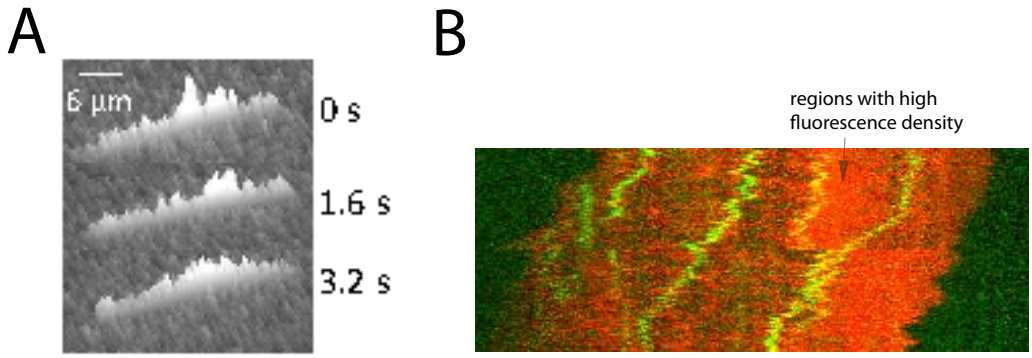

Figure S2: (A) The backbone intensity images of a confined DNA fragment ($\sim$34μm) stained with YOYO-1 iodide in a 80nm$\times$130nm channel. The images are recorded at time interval of 1.6s. From the heterogeneity of the intensity profile, we infer that there exist some local structures on the backbone. (B) Images of the time series (8 seconds) of a T4 DNA fragment ($\sim$32μm). The backbone of the DNA is shown in red and the internal dyes are shown in green. The region with high fluorescence density is the area with local folded structures. The green traces are the trajectories of internal dye labels in the time series. This image shows two internal dyes coming together, which is evidence of formation of local folded structures.

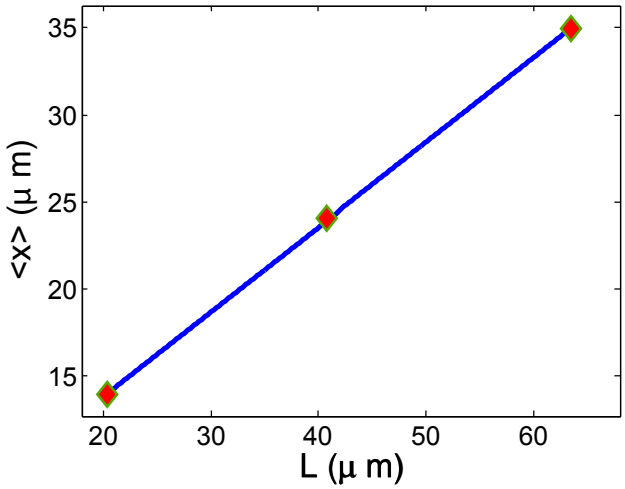

Figure S3: **Mean end-to-end extension $\langle x \rangle$ versus contour length L of confined DNA in a 60nm$\times$100nm channel.** The fitting result is $x = 0.5L$, which is consistent with the prediction of the Odijk deflection theory: $x = 0.7L$.

Since $dG = -xdF$, we can integrate Eq.1 to obtain the free energy $G = G(F,T)$:

$$\begin{aligned} G &= -\int xdF & (6)\\ &= -\int \left[L - \frac{k_B T L}{2\sqrt{\kappa}} \frac{1}{\sqrt{F + 2\sqrt{\Xi\kappa}}} \right] dF & (7)\\ &= -\frac{(k_B T L)^2}{4\kappa} \frac{2x - L}{(L-x)^2} + \text{const} & (8) \end{aligned}$$

This is the free energy in a fixed force ensemble, i.e $G(F,T) = E - TS - Fx$. We need the free energy in a fixed extension ensemble. Therefore:

$$\begin{aligned} G(x,T) &= G(F,T) + Fx & (9)\\ &= \frac{(k_B T L)^2}{4\kappa} \frac{1}{L-x} - 2\sqrt{\Xi K}x + \text{const} & (10) \end{aligned}$$

Denote $\rho = x/L$, then the free energy is:

$$\frac{G(\rho,T)}{k_B T} = \frac{A}{1-\rho} - B\rho + \text{const}, \tag{11}$$

where

$$A = \frac{L}{4\xi_p}, \quad B = \frac{4c^2 \xi_p^{1/3} L}{D^{4/3}}, \quad c = 1.1 \tag{12}$$

Therefore, the probability distribution is:

$$P(\rho) = P_0 \exp\left[B\rho - \frac{A}{1-\rho} \right] \tag{13}$$

with A, B given in Eq. 12 and P_0 being the normalization constant.

References

[1] Wang J, Gao H (2007) Stretching a stiff polymer in a tube. *J Mater Sci* 42:8838-8843.

[2] Odijk T (1983) On the statistics and dynamics of confined or entangled stiff polymers. *Macromolecules* 16:1340-1344.