

## Compression and Stretching of a Self-Avoiding Chain in Cylindrical Nanopores

Suckjoon Jun

*FAS Center for Systems Biology, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA*

D. Thirumalai

*Biophysics Program, IPST and Department of Chemistry and Biochemistry, University of Maryland at College Park, College Park, Maryland 20742, USA*

Bae-Yeun Ha

*Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

(Received 26 May 2008; published 22 September 2008)

Force-induced deformations of a self-avoiding chain confined inside a cylindrical cavity, with diameter  $D$ , are probed using molecular dynamics simulations, scaling analysis, and analytical calculations. We obtain and confirm a simple scaling relation  $-fD \sim R^{-9/4}$  in the strong-compression regime, while for weak deformations, we find  $fD = -A(R/R_0) + B(R/R_0)^{-2}$ , where  $A$  and  $B$  are constants,  $f$  the external force, and  $R$  the chain extension (with  $R_0$  its unperturbed value). For a strong stretch, we present a universal, analytical force-extension relation. Our results can be used to analyze the behavior of biomolecules in confinement.

DOI: 10.1103/PhysRevLett.101.138101

PACS numbers: 87.15.-v, 82.35.Lr, 82.37.Rs

Technological advances have made it possible to visualize and manipulate individual polymer molecules that are trapped in nanopores [1–4]. This enables one to test theoretical predictions on confined polymers, shedding new insights into their static and dynamic properties. Besides their technological importance, confined polymers are relevant in a number of biological processes. In the context of proteins, a newly synthesized protein exits the ribosome through a narrow cylindrical pore [5] assisted, perhaps, by a tensile force [6]. Both chaperone-assisted protein folding and protein degradation by proteasome involve encapsulation of proteins in cylindrical pores, in which they experience compression and stretching forces [7,8]. Furthermore, many bacterial species are rod-shaped and even filamentous, and their chromosomes are highly compressed inside the cell and yet well organized [9,10]. Thus, the study of cylindrically confined polymers is not only of practical importance but also is a first step towards understanding several fundamental biological processes.

Surprisingly, compared to the progress on the experimental side, and despite its fundamental and practical importance, theoretical understanding of force-induced chain deformations in a pore remains behind. In the literature, analysis of force-compression-extension measurements heavily relies on and is often limited to simple scaling arguments [1–3].

The main purpose of this Letter is to present a quantitative picture of the interplay between confinement and force-induced chain deformations (from strong compression to almost full stretching). To this end, we combine a scaling analysis, molecular dynamics (MD) simulations, and a systematic theoretical approach, and obtain force-compression-extension relations in almost the full range of

chain deformations. Our results thus provide a quantitative basis for experiments involving biomolecules (e.g., dsDNA and chromosomes) in strong confinement as in nano- and micro-channels.

Consider a self-avoiding polymer with  $N$  monomers in a cylindrical pore with diameter  $D$ , as illustrated in Fig. 1, with  $R$  being the longitudinal end-to-end distance of the chain. Owing to the simultaneous presence of self-avoidance and confinement, the chain resembles a linear string of “compression blobs” [11,12]; inside each blob, the effect of confinement is not significant. Because of the

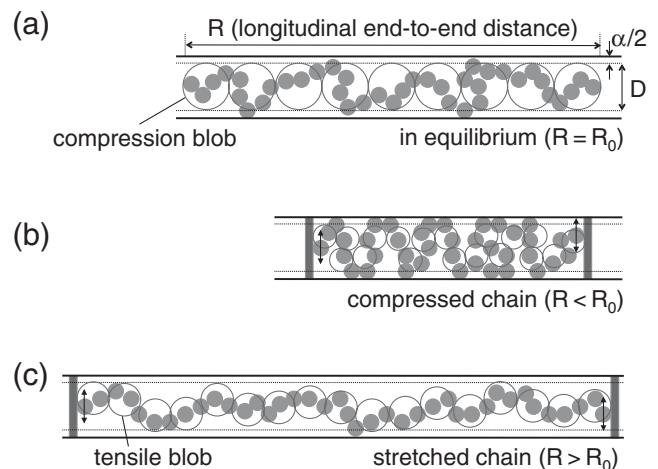


FIG. 1. Self-avoiding polymer in a nanopore (a). Simulation setting for compression (b) and stretching (c); the two chain ends are confined to, but can freely move within, the confining “piston” walls. By measuring the pressure exerted on the piston walls, we obtain force-compression-stretching relations, for a given chain size  $N$  and a pore width  $D$ .

clear separation between two length scales ( $>D$  and  $<D$ ), the conventional Flory approach may fail to capture the correct  $D$  dependence of relevant quantities (see below). The essential features of the blob-scaling approaches [11,12] can be reproduced by the following “renormalized” free energy [13],

$$\beta\mathcal{F}(R, D) = A \frac{R^2}{(N/g)D^2} + B \frac{D(N/g)^2}{R}, \quad (1)$$

where  $g$  is the number of monomers inside a compression blob of diameter  $D$ , i.e.,  $g \simeq (D/a)^{5/3}$ ,  $A$  and  $B$  are constants, and  $a$  is the monomer size. (Throughout this Letter,  $\beta = 1/k_B T$  with  $k_B T$  being the thermal energy; unless otherwise indicated,  $a$  is the unit of length.) The first term in Eq. (1) describes connectivity of blobs, while the second term, which can be rewritten as  $\sim(N/g) \times$  blob density, represents the mutual repulsion between the blobs, thus ensuring linear ordering. The second term is equivalent to correctly assigning  $k_B T$  to a binary collision of two blobs, independent of their size [11,14].

The free energy in Eq. (1) produces *not only* the expected equilibrium chain size  $R_0 \sim ND^{-2/3}$  [15], when  $\mathcal{F}$  is minimized, *but also* the correct confinement free energy  $\beta\mathcal{F}_{\text{conf}} = \beta\mathcal{F}(R = R_0, D) \sim N/g \sim ND^{-5/3}$ , which is linear in  $N$  [11,12]. The free energy cost for linear ordering is correctly counted as  $\sim k_B T$  per blob ( $D$ -independent). Also, the resulting effective Hookean spring constant of the chain,  $k_{\text{eff}} = (\partial^2 \mathcal{F} / \partial R^2)_{R_0} \sim N^{-1} D^{-1/3}$ , is consistent with the previous scaling result [1,12]. Note that, without the renormalization trick in Eq. (1), Flory theory (as employed in [3]; see also [13]) will lead to a  $D$ -independent (thus qualitatively inaccurate) spring constant  $k_{\text{eff}}$ .

An important consequence of Eq. (1) is that it predicts a universal scaling relation for (external) force-compression-extension ( $R$ ) relation,

$$D\beta f = D \frac{\partial}{\partial R} (\beta\mathcal{F}) = 2A \left( \frac{R}{R_0} \right) - B \left( \frac{R}{R_0} \right)^{-2}. \quad (2)$$

To test the validity of Eq. (2) [thus Eq. (1) as well], we have carried out molecular dynamics simulations using ESPRESSO [16] for a wide range of parameters:  $D = 4, 5, \dots, 15$  and  $N = 128, 256, 512$  (see Fig. 1). In ESPRESSO, monomers (or beads) are connected by finite extensible nonlinear elastic (FENE) (spring) bonds with a Weeks-Chandler-Andersen (WCA) potential for excluded-volume interactions (monomer-monomer as well as monomer-wall). We confined the two end beads of the chain by “piston” walls at the cylinder ends [Figs. 1(b) and 1(c)], while allowing the ends to move freely in the plane of the piston wall. By varying the wall-to-wall distance (thus  $R$ ), we obtained the longitudinal force  $f(R)$  to keep  $R$  fixed.

Figure 2(a) summarizes our simulation results, where all 24 curves have been rescaled according to Eq. (2). The data perfectly collapse onto a single master curve (the solid line) for  $R < R_0$ . Moreover, in the range of  $0.5 < R/R_0 < 1.2$ , Eq. (2) fits the rescaled data very well, thus confirm-

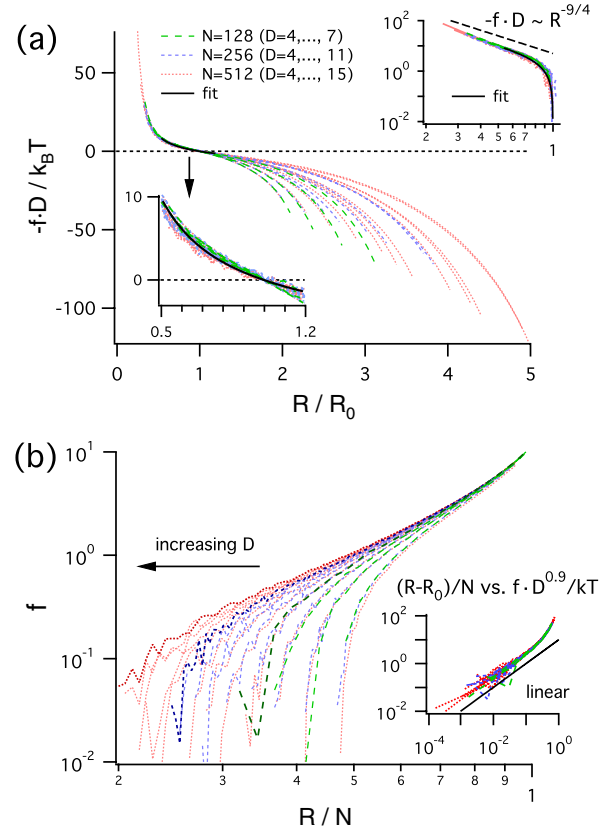


FIG. 2 (color online). [(a) and (b) have the same color scheme online] Simulation (dotted lines) and scaling (solid lines) of force-compression-stretching of a chain in a cylinder. (a)  $R/R_0$  vs  $fD/k_B T$ . All 24 simulation curves collapse onto a single universal curve for  $R < 1.2R_0$  [ $N = 128(D = 4, 5, 6, 7)$ ,  $N = 256(D = 4, 5, \dots, 11)$ ,  $N = 512(D = 4, 5, \dots, 15)$ ]. The solid line is the scaling result using Eq. (2) in the range of  $0.5 < R/R_0 < 1.2$ ; by a global fit to the simulation, we obtained  $2A = 2.84 \pm 0.03$  and  $B = 2.81 \pm 0.01$  (i.e.,  $2A \approx B$ ). The lower-left inset shows the fitted region. The upper-right inset is the log-log plot for  $R < R_0$ . (b)  $R/N$  vs  $|f|$  for force-stretching ( $R > R_0$ ). The inset shows the linear Hookean regime for weak force (see text). The sudden departure from the linear relation as  $R/N \rightarrow 1$  reflects the nonuniversal aspect of the bond potentials in simulations.

ing the scaling prediction of Eq. (1) [see the lower-left inset in Fig. 2(a)]. However, as the compression becomes strong ( $R/R_0 \ll 1$ ), the blobs break into smaller ones [see Fig. 1(b)], and the correct form of free energy is  $\beta\mathcal{F} \sim (R_g^3/V)^{1/(3\nu-1)}$ , where  $R \sim N^\nu$  is the radius of gyration of a corresponding unconfined chain (with  $\nu \approx 3/5$  the Flory exponent) and  $V \sim D^2 R$  the confining volume [14]. From this, we obtain

$$-fD \sim \left( \frac{R_0}{R} \right)^{3\nu/(3\nu-1)} \sim \left( \frac{R_0}{R} \right)^{9/4} \quad (3)$$

[see the upper-right inset in Fig. 2(a)].

Importantly, Fig. 2 shows the limitations of Eq. (2); when the chain is stretched appreciably from its equilibrium length ( $R \gtrsim 1.2R_0$ ), the force-extension curves

(FECs) do not collapse onto a single universal plot. This is not unexpected since an additional length scale, namely, the tensile blob size  $\xi$  ( $< D$ ) [11], is relevant at high  $f$ , as indicated in Fig. 1(c). This also signals the onset of a  $D$ -independent regime at large  $f$ , as the interactions between the chain and the wall are negligible [see Fig. 2(b) and 3].

In Fig. 2(b), we show FECs for  $R > R_0$  by rescaling  $R$  by  $N$ . Since  $R/N$  denotes the extension per chain segment, for the same pore diameter  $D$ , the force-extension curves tend to collapse onto one another, as expected. In a recent simulation study [17], the effective spring constant  $k_{\text{eff}}$  was shown to vary from  $k_{\text{eff}} \sim 1/ND^{0.7}$  to  $1/ND^{1.3}$  for the intermediate chain length (up to  $N = 2000$ ). This differs from the scaling result  $k_{\text{eff}} \sim 1/ND^{1/3}$  [12], which could be reached only when  $N \sim 10^4$  [17]. Our results are also in good agreement with Arnold *et al.*, and we find the linear Hookean collapse for  $k_{\text{eff}} \sim 1/ND^{0.9 \pm 0.2}$  [ $f \sim k_{\text{eff}}(R - R_0)$ ; inset of Fig. 2(b)]. The data cannot be fit using the large  $N$  result,  $k_{\text{eff}} \sim 1/ND^{1/3}$  [18].

To further elucidate the subtleties of the FECs in the presence of confinement, we have also performed analytical, self-consistent calculations following the theoretical methods described in [19,21]. Our motivation is to derive the FEC relation in the large- $N$  limit, which can be reached experimentally (if not computationally; see above). In the continuum limit, the distribution of monomers at  $\mathbf{r}(s)$  can be represented by the following Hamiltonian

$$\beta\mathcal{H} = \frac{3}{2a} \int_0^L \mathbf{r}^2(s) ds + \frac{v}{2} \int_0^L \int_0^L ds ds' \delta(\mathbf{r}(s') - \mathbf{r}(s)) - \beta f \int_0^L \dot{z}(s) ds, \quad (4)$$

where  $s$  is the contour length ( $0 \leq s \leq L = Na$ ),  $[\dots] = \partial[\dots]/\partial s$ , and  $v$  the excluded-volume parameter. Confinement effects will be taken into account through

$$\begin{aligned} La_1^{3/2} \left( \frac{1}{a} - \frac{1}{a_1} \right) &= \sqrt{\frac{6}{\pi}} \frac{1.76}{\pi} \frac{v}{\alpha_0^2 D^2} \int_0^L ds (L-s) \sqrt{s} \exp \left[ -\frac{1}{6} s a_1 (\beta f)^2 \right] \\ &= 0.4 \frac{v}{D^2} \left[ \frac{18 e^{-\beta^2 f^2 L a_1}}{L^{1/2} \beta^4 f^4 a_1^2} + \frac{\sqrt{6\pi} (-9 + L \beta^2 f^2 a_1) \text{erf}(f \sqrt{L a_1 / 6})}{\beta^5 f^5 a_1^{5/2}} \right], \end{aligned} \quad (5)$$

where  $\alpha_0 \approx 2.40$  is the smallest (first) zero of the zeroth-order Bessel function of the first kind  $J_0(x)$ , and  $\text{erf}(x)$  is the error function. When combined with  $R = \beta L a_1 f / 3$ , the solution of the SCE determines the equilibrium chain length along the pore axis.

For  $f = 0$ , the SCE leads to the equilibrium chain size  $R_0 = \sqrt{\langle (z_L - z_0)^2 \rangle}_{f=0} \sim L (\frac{va}{D^2})^{1/3}$ , which is consistent with the previous result [12,21]. For  $f > 0$ , the SCE can be solved numerically. A few comments are in order: First, the excluded-volume parameter  $v$  in the SCE can be considered as a fitting parameter (see Ref. [19]). The value of  $v$  can be adjusted to ensure the best global fit to simulations

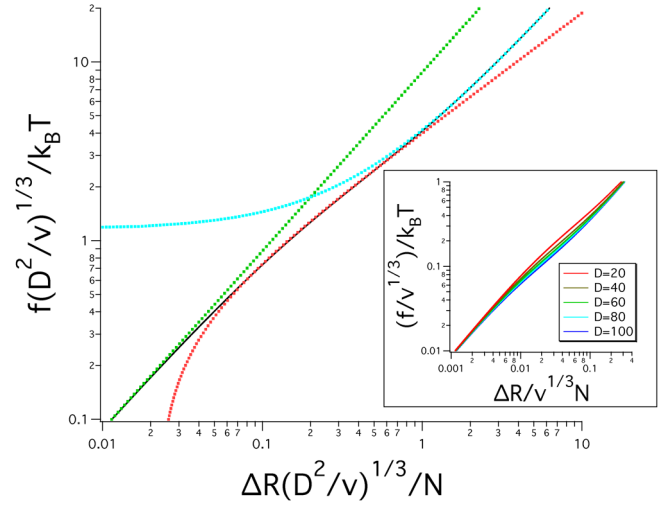


FIG. 3 (color online). Rescaled force-extension relation in the limit  $N \rightarrow \infty$ . Three distinct regimes are identified and presented by the dotted lines (green, red, and cyan), which are linear, power-law, and linear fits to the force-extension relation in their respective regimes. The power-law analysis leads to  $k_{\text{eff}} \sim 1/D^{1/3}N$ .

the boundary condition that the probability of finding any monomer on the wall is zero.

We replace  $\mathcal{H}$  by the simpler reference  $\mathcal{H}_1$  given by  $\beta\mathcal{H}_1 = \frac{3}{2a_1} \int_0^L \mathbf{r}^2(s) ds - \beta f \int_0^L \dot{z}(s) ds$  and choose an optimal value of  $a_1$ —instead of minimizing the free energy as in the Flory approach—so that  $R$  obtained using  $\mathcal{H}_1$ , i.e.,  $R \equiv \langle z_L - z_0 \rangle = \beta L a_1 f / 3$ , coincides with  $R$  calculated using  $\mathcal{H}$  to first order in  $v$  [19,21]. Additionally, we use the “ground-state dominance” approximation, which is valid in the large  $N$  limit [11]. The optimal  $a_1$  satisfies the self-consistent equation (SCE):

or experimental data. Second, our SCE does not reflect the kinetic aspects of force-extensions; chain reorientation, allowed in our SCE, is kinetically suppressed for strong confinement. (It is strictly forbidden in our simulations.) Finally, because of the ground-state-dominance approximation, the SCE becomes more accurate for large  $L$ . In the limit  $L \rightarrow \infty$ , chain reorientation becomes irrelevant in that chain alignment in the force direction is ensured for  $f \geq k_B T / R_0$  ( $\rightarrow 0$  as  $L \rightarrow \infty$ ). In this case,  $v$  dependence enters the FEC only via the combination  $D^2 v$  and, thus, can be absorbed into an effective  $D$  (see below). Our focus below is thus the large  $L$  behavior of our SCE.

In the large  $L$  limit, the SCE in Eq. (5) reduces to  $a_1^3(\frac{1}{a} - \frac{1}{a_1}) \approx 0.44 \frac{v}{D^2} \frac{1}{(\beta f)^3}$ . This equation implies that  $a_1$  is a function of  $fD^{2/3}/v^{1/3}$ :  $a_1 = a_1(fD^{2/3}/v^{1/3})$ . As a result,  $(\Delta R/N)D^{2/3}/v^{1/3} = (\beta a_1 f/3)D^{2/3}/v^{1/3}$  is a function of  $fD^{2/3}/v^{1/3}$ , where  $\Delta R = R - R_0$ . Thus, the FECs collapse onto each other in a  $fD^{2/3}/v^{1/3} - (\Delta R/N)D^{2/3}/v^{1/3}$  plot. The analytic results justify the scaling plot in Fig. 2, where the renormalized  $D_R \sim Dv^{1/2}$ . In Fig. 3, we plot  $\Delta RD^{2/3}/v^{1/3}$  as a function of  $fD^{2/3}/v^{1/3}$  in the limit  $N \rightarrow \infty$ . All the curves for different values of  $D$  collapse onto each other (as found in the simulations) for the reason described above. (Also see the inset in Fig. 3 obtained without  $D$  rescaling. This suggests that a more strongly confined chain behaves as a stronger spring).

Three different force-extension regimes are identified in Fig. 3: linear, power-law, and linear (green-, red-, and cyan-dotted lines, respectively, online). The good linear fit to the very narrow linear regime indicates a weak  $D$ -dependence of the force-extension relation [22]. For the power-law fit in the second regime, we have chosen  $RD^{2/3} \sim \text{const} + (fD^{2/3})^\alpha$  with  $\alpha \approx 1.5$ . From this,  $k_{\text{eff}}$  has been estimated to vary as  $D^{-2(1-\alpha)/3} \approx D^{-1/3}$ . The seemingly perfect fit to the force-extension in this regime justifies the scaling result  $k_{\text{eff}} \sim 1/D^{1/3}N$ . However, the force-extension in this regime is nonlinear, in contrast to what one may expect from the scaling approach [12]. Only over a narrow parameter range can the linearity be recovered. We note, however, that this power-law regime represents a much wider parameter space than the initial linear regime. The large- $f$  linear regime is where the force-extension becomes  $D$ -independent, due to almost full stretching of the chain by large force.

Our results suggest that relaxation dynamics of a chain confined in a narrow channel will show several distinct relaxation rates ( $\propto k_{\text{eff}}$ ) with different  $D$  dependence. This explains the limitations of existing scaling arguments based on a single-time scale [1–3]. The predictions based on a scaling approach are reached only when  $N$  is very large. Thus simulations or experiments on short chains need to be interpreted with caution. Our approach can be extended to incorporate other nontrivial but important effects such as chain topology (e.g., branched polymers), electrostatics, molecular crowding, and chain stiffness [23]. Most importantly, the force-compression-extension relations presented above can be used to understand the elastic response of chromosomes confined in microfluidic devices, as well as the fate of proteins in nanopores.

We thank A. Arnold and D. Frenkel for helpful discussions in the early stage of this work and J. Bechhoefer for critical reading. The work of D. T. was supported by a grant from NSF No. CHE05-14056. B. Y. H acknowledges financial supports from the NSERC (Canada), and the BK 21 Frontier Physics Research Program during his stay at the Department of Physics and Astronomy, Seoul National University, Korea.

- [1] J. O. Tegenfeldt *et al.*, Proc. Natl. Acad. Sci. U.S.A. **101**, 10979 (2004).
- [2] W. Reisner *et al.*, Phys. Rev. Lett. **94**, 196101 (2005).
- [3] C.H. Reccius, J.T. Mannion, J.D. Cross, and H.G. Craighead, Phys. Rev. Lett. **95**, 268101 (2005).
- [4] C. Dekker, Nature Nanotech. **2**, 209 (2007).
- [5] V. Ramakrishnan and P.B. Moore, Curr. Opin. Struct. Biol. **11**, 144 (2001).
- [6] D.K. Klimov, D. Newfield, and D. Thirumalai, Proc. Natl. Acad. Sci. U.S.A. **99**, 8019 (2002).
- [7] A. Martin, T.A. Baker, and R. T. Sauer, Nat. Struct. Mol. Biol. **15**, 139 (2008).
- [8] D. Thirumalai and G.H. Lorimer, Annu. Rev. Biophys. Biomol. Struct. **30**, 245 (2001).
- [9] J. Stavans and A. Oppenheim, Phys. Biol. **3**, R1 (2006).
- [10] S. Jun and B. Mulder, Proc. Natl. Acad. Sci. U.S.A. **103**, 12388 (2006); A. Arnold and S. Jun, Phys. Rev. E **76**, 031901 (2007).
- [11] P.-G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, 1979).
- [12] F. Brochard and P.-G. de Gennes, J. Chem. Phys. **67**, 52 (1977).
- [13] The original Flory approach in  $d$  dimensions leads to  $\beta\mathcal{F} \sim \frac{R^2}{Na^2} + \frac{a^d N^2}{R^d}$ . The renormalized Flory free energy in Eq. (1) can be obtained by setting  $d = 1$  and rescaling  $a \rightarrow D$  and  $N \rightarrow N/g$ . Linearity of  $\mathcal{F}$  around  $R = R_0$  in  $d = 1$  underlies the success of this rescaling scheme.
- [14] S. Jun, A. Arnold, and B.-Y. Ha, Phys. Rev. Lett. **98**, 128303 (2007).
- [15] In the experimental literature, even the basic de Gennes scaling of  $R_0 \sim ND^{-(2/3)}$  has yet to be clarified for DNA in nanochannels, where significant fluctuations of  $R_0$  (and, thus, the measurement uncertainties) are present [2].
- [16] H.-J. Limbach, A. Arnold, B.A. Mann, and C. Holm, Comput. Phys. Commun. **174**, 704 (2006).
- [17] A. Arnold, B. Borzorgui, D. Frenkel, B.-Y. Ha, and S. Jun, J. Chem. Phys. **127**, 164903 (2007).
- [18] Based on the results in Refs. [17,19] and this study, care should be exercised in analyzing data on finite-sized DNA confined to pores. The chain behavior can be sensitive to the precise choice of the pair interaction of the chain beads especially if the channel diameter approaches the bead diameter. This may be critical when coarse grained models are employed as in Ref. [20].
- [19] G. Morrison, C. Hyeon, N.M. Toan, B.-Y. Ha, and D. Thirumalai, Macromolecules **40**, 7343 (2007).
- [20] R.M. Jendrejack, *et al.*, Phys. Rev. Lett. **91**, 038102 (2003).
- [21] G. Morrison and D. Thirumalai, J. Chem. Phys. **122**, 194907 (2005).
- [22] The emergence of the  $D$ -insensitive regime (green) is intriguing. We believe that this can be attributed to the different setting for our theoretical analysis and/or partial chain backfolding. Simulations on much longer chains are needed for further clarification.
- [23] When  $D$  is comparable to but larger than the persistence length, compression blobs can overlap each other relatively easily. Chain stiffness can be implemented in our approach.