Structure of DNA Coils in Dilute and Semidilute Solutions

Manish Nepal,[1] Alon Yaniv,[1] Eyal Shafran,[1] and Oleg Krichevsky[1,2,*]

[1]*Physics Department, Ben-Gurion University, Beer-Sheva 84105, Israel*

[2]*Ilse Kats Center for Nanoscience, Ben-Gurion University, Beer-Sheva 84105, Israel*

(Received 13 March 2012; revised manuscript received 10 October 2012; published 29 January 2013)

We apply scanning fluorescence correlation spectroscopy to study the structure of individual DNA coils in dilute and semidilute solutions. In dilute solutions, over two decades in length, from 0.6 to 46 μm, DNA behave as ideal chains, in agreement with theoretical predictions and in disagreement with prior experiments. In semidilute solutions, up to very high densities, the structures of individual DNA coils are independent of concentration, unlike flexible coils that shrink with increasing density. Our experimental findings are consistent with the marginal solution theory of semiflexible polymers.

DOI: 10.1103/PhysRevLett.110.058102 PACS numbers: 87.14.gk, 87.15.N−

In solutions of flexible polymers we distinguish between good, θ—and bad solvents based on the second virial coefficient v of interactions between polymer segments in the solvent, and at least two concentration regimes: below and above coil overlap concentration, or, respectively, dilute and semidilute solutions [1]. Over a very narrow range of parameters around θ temperature, Van der Waals attraction compensates excluded volume repulsion between the segments, so that $v = 0$. Thus, in dilute-θ solutions polymers behave as ideal chains with a random walk statistics: in particular, their gyration radius R_g scales with polymer length L as $R_g \sim L^{1/2}$. In semidilute-θ solutions, the solution correlation length ξ dependence on monomer concentration c is determined by three-body collisions between segments of interpenetrating chains leading to $\xi \propto c^{-1}$, while the structure of individual coils is unaffected by the weak interactions with other chains [2]. In *good* solvents, repulsion dominates ($v > 0$) resulting in the swelling of polymer coils in dilute-good solutions so that $R_g \sim L^{0.59}$ [1,3]. In semidilute-good solutions, the interactions between interpenetrating chains lead to $\xi \propto c^{-3/4}$ and to the screening of interactions between same chain segments, so that the individual polymers adopt the ideal chain statistics and their coils contract. In bad solvents, the attraction between polymer segments is dominant ($v < 0$) bringing about chain collapse.

Semiflexible polymers bend only weakly in space: their persistence length l_p is much larger than the monomer size a (while smaller than L). This means that there is a lot of ''free'' volume within their coils, and the probability of collisions between the same chain monomers is low. In a Flory-type calculation [3], the Schaefer *et al.* theory for semiflexible polymers [4] predicts a wide range of dilute-θ conditions: such chains behave as ideal up to the contour length of $L_\tau \sim 8l_p^3/a^2 \gg a$. For semidilute solutions, the theory [4] predicts a totally new, marginal, regime intermediate between θ and good solvent types in which the correlation length should obey Edwards' scaling [5] with concentration $\xi \sim c^{-1/2}$, different from other regimes. Yet as in θ conditions, despite this collective screening effect, the structure of each individual chain is predicted to remain largely unaffected by the presence of other chains [4,6]. The attempts to test this theory using synthetic polymers have been inconclusive [4,7–10].

A particularly striking discrepancy between theory and experiments relates to the structure of DNA coils in dilute solutions. DNA has been the most prominent model of semiflexible polymers (and polymers in general) over the past several decades. Its persistence length of $l_p \approx 50$ nm is much larger than its double-helix diameter of 2 nm. In physiological conditions (pH6–8, 100–200 mM of monovalent salts) electrolyte ions screen the electrostatic interactions between negatively charged DNA segments beyond Debye-Hückel distance of ~1 nm. The double-helix and the Debye-Hückel layer widths add up to an effective DNA diameter of $d \approx 4$ nm. Theoretical estimates then show that in physiological conditions, DNA up to $L_\tau \approx 8l_p^3/d^2 \approx 60$ μm in length should have negligible excluded-volume interactions and behave as an ideal chain [11].

Surprisingly though, this prediction has never been confirmed experimentally. In fact, a variety of experimental results actually contradict it: the diffusion coefficient of DNA coils of lengths $2.6 < L < 130$ μm [12], the segment distribution of DNA of lengths $15 < L < 60$ μm tethered to a substrate [13], AFM measurements of DNA end-to-end distance distribution for $0.5 < L < 15$ μm [14]—all reveal the scaling $R_g \propto L^{0.57-0.59}$ indicating strong excluded volume interactions. While none of these studies present a direct measurement of coil structure in solution, even static light scattering (SLS) experiments that do measure the structure factor of solutions, reveal a significant 30–40% expansion of 14 μm DNA chains [15,16] far exceeding theoretical estimations (~ 4%) [17,18]. A known issue in SLS studies is the very poor scattering by DNA, that requires extrapolation procedures from higher concentrations to measure DNA structure in dilute solutions [16]. SLS is also sensitive to scattering by dust particles, that are not

trivial to remove because of the closeness of their sizes to those of DNA molecules [15].

The disagreement between experiment and theory on DNA coil structure questions the most basic calculation in polymer physics—Flory estimation of excluded volume interactions within a polymer chain [3]—as applied to semiflexible chains. Here we present the first measurements of the structure of DNA coils showing that over almost two decades in length, from ~0.6 to ~46 μm, DNA behaves as a nearly ideal polymer in accordance with theoretical predictions: their structure factor $S(q)$ is described by the Debye expression for ideal chains [19] and their $R_g(L)$ dependence follows that of an ideal wormlike chain. We use a new approach we introduced recently that is based on scanning fluorescence correlation spectroscopy (SFCS) of fluorescently labeled DNA molecules [20]. With the help of this technique we showed previously [20] that the screening length of semidilute marginal solutions of DNA indeed obeys Edwards' scaling as predicted by Schaefer *et al.* [4]. Here, we couple our method with specific fluorescence labeling to also test another prediction of the theory: the invariance of the structure of individual coils in marginal semidilute solutions, a question that could not be addressed previously in experiments.

In our approach, the emission I_{em} arriving from a diffraction limited confocal volume is monitored as the sample is moved at a constant speed v through the beam. If the speed is sufficiently high so that there is no significant DNA internal dynamics while the molecule is passed through the beam, then the temporal correlation function of emission fluctuations $G(t) = \langle \delta I_{\mathrm{em}}(0)\delta I_{\mathrm{em}}(t)\rangle / \langle I_{\mathrm{em}}^2 \rangle$ reflects monomer-monomer density spatial correlation function $g(r) = \langle \delta c(0)\delta c(r)\rangle$ where $\delta c(r) = c(r) - \langle c \rangle$ with $r = vt$. We showed [20] that the measured $G(t)$ is a convolution of $g(r)$ and of the function characterizing setup excitation-detection profile, so that it can be expressed in Fourier space as:

$$G(\vec{r}) \propto \int d\vec{q} |I(\vec{q})|^2 S(q) e^{i\vec{q}\,\vec{r}}, \tag{1}$$

where $I(\vec{q})$ is the optical transfer function of the setup.

The measurement setup and experimental conditions were similar to those used in Ref. [20]. Briefly, DNA molecules of varying lengths from 0.6 to 46 μm were dissolved in phosphate buffered saline and placed onto the piezodriven stage of a home-built confocal SFCS setup. The samples were scanned at ~4 mm/s speed in a pattern of imperfect circles of 60 μm diameter whose center moves along a 40 μm diameter circle. Such a complex pattern was designed to minimize the chances of repeated passage of the beam through the same point in the sample. Although the formalism of our method is strictly correct only for scanning along the line with constant velocity, it can be applied in our setup since the characteristic curvature radius of sampling trajectory is much larger than the typical sizes of DNA coils (up to ~1 μm in R_g). The scanning speed was held constant to within 1–2% (standard deviation). We verified that the scanning speed is high enough for the sample to be considered "frozen" by obtaining similar results with scan velocities as low as 1 mm/s. The effective point spread function of the setup $I(\vec{r})$ was assumed to be Gaussian and was calibrated as in [20]. The characteristic dimensions of the confocal volume are $w_{xy} \approx 0.25$ μm and $w_z \approx 1.4$ μm in lateral and axial directions, respectively.

In dilute solutions, the majority of the measurements were carried out on DNA stained nonspecifically with ethidium bromide (EtBr) dye. At EtBr concentrations in the 5 to 20 μM range in phosphate buffered saline, dye intercalation extends DNA contour length by a factor of ~1.4 but does not affect its persistence length [21,22]. We use 10 μM EtBr concentrations at which the dye binds DNA in ~1:5 dye/base pair ratio. Short DNA fragments of different lengths ($L < 48\,502$ bp) were obtained by applying commercially available restriction enzymes to λ-DNA (48 502 bp), pUC18 (2686 bp), and pBR322 (4361 bp). The longest DNA probed in this study were λ-DNA and its dimers. In order to obtain the dimers, λ-DNA were annealed to open their overhangs, ligated to each other and separated with the help of field inversion gel electrophoresis [23]. λ-DNA and its dimers were annealed at 65 °C for 10 min before each measurement in order to keep the chains linear. In part of the experiments, λ-DNA were hybridized to their complementary oligonucleotides in order to further restrict polymer ability to form a closed circle: the results were similar to those on annealed molecules.

In another approach to sample labeling, we stained DNA molecules uniformly along their contour with covalently bound carboxyrhodamine 6G (Rh6G) dye. This was achieved by synthesizing DNA fragments of different lengths through Polymerase Chain Reaction (PCR) with part of the native nucleotide bases (dTTP) in solution substituted by an analog (amino-allyl-dUTP, AA-dUTP) having a chemically reactive site (amino group) [24]. Following PCR, succinimidyl ester moiety of Rh6G is reacted onto the amino group of AA-dUTP. Practically, the modified base is introduced in ~1:200 ratio to native bases.

The DNA concentrations for all samples were kept well in the dilute regime: less than 4 μg/mL for λ-DNA and shorter molecules and less ~1 μg/mL for λ-DNA dimers, while the overlap concentration for λ-DNA is ~10–15 μg/mL [20].

The examples of measured SFCS correlation functions for DNA in dilute solutions are shown vs $r^2 = (vt)^2$ in Fig. 1. As expected, spatial correlations extend to larger distances for longer chains. All the correlation functions can be fit well with Eq. (1) where the Debye structure factor for ideal polymers is used [19]:

$$S(q) = \frac{2}{(qR_g)^4}(q^2R_g^2 + e^{-q^2R_g^2} - 1) \tag{2}$$

with R_g as a fit parameter (Fig. 1).

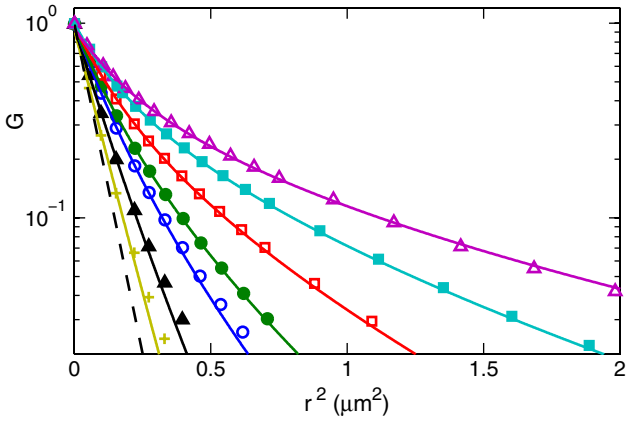

FIG. 1 (color online). Examples of normalized SFCS correlation functions for DNA in dilute solutions. DNA lengths: 0.64 (crosses), 2.1 (full triangles), 4.5 (open circles), 8.0 (full circles), 11.5 (open squares), 23.1 (full squares), and 46.2 μm (open triangles). Solid lines represent fits to the experimental data with Eq. (1) assuming ideal structure of the chains [Eq. (2)]. The dashed line represents the calibration of the excitation-emission profile of the optical setup corresponding to the SFCS correlation function of independent point sources.

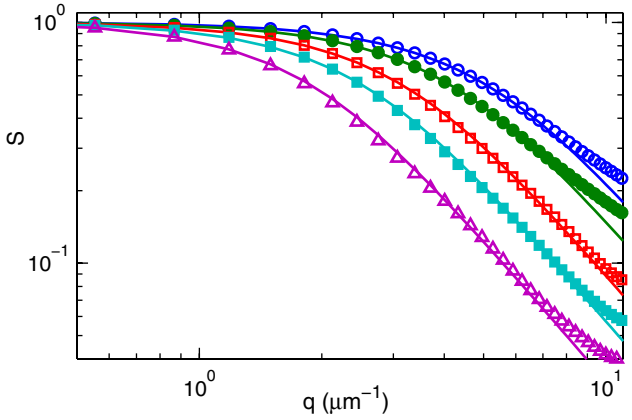

FIG. 2 (color online). Symbols represent measured structure factors $S(q)$ for dilute DNA solutions. DNA lengths: 4.5 (open circles), 8.0 (full circles), 11.5 (open squares), 23.1 (full squares), and 46.2 μm (open triangles). Lines are fits of $S(q)$ with Debye expression (2).

For DNA longer than 4 μm, the structure factors can be extracted by Fourier transforming Eq. (1) as in Ref. [20] in a sufficiently large dynamic range to be fit directly with Eq. (2) (for shorter DNA the measured correlation functions are close to that of uncorrelated point sources and the extracted $S(q)$ are reliable in a dynamic range much smaller than a decade). The fits are good for $q < 9\ \mu\text{m}^{-1}$ (Fig. 2) for all DNA lengths $L > 4\ \mu$m tried. The upward trend in the extracted $S(q)$ for $q > 9\ \mu\text{m}^{-1}$ appears to be an artifact of the deviations of the optical field from Gaussianity (as measured through direct mapping of $I(\vec{r})$ by scanning 100 nm gold beads). The overall quality of the fits in Figs. 1 and 2 points to the negligible excluded volume effects in agreement with theoretical predictions.

The gyration radii R_g obtained from the fits are plotted vs DNA contour length L in Fig. 3. The data from both sets of the samples, those stained with intercalating dye and with covalently bound labels, exhibit excellent agreement. The $R_g(L)$ dependence follows closely a $R_g \propto L^{0.52\pm0.02}$ power law which indicates the absence of any major coil expansion. Alternatively, we can describe $R_g(L)$ dependence with the expression for an ideal wormlike chain [25],

$$R_g^2(L) = \frac{Ll_p}{3} - l_p^2 + \frac{2l_p^3}{L} - \frac{2l_p^4}{L^2}\left[1 - \exp\left(-\frac{L}{l_p}\right)\right]. \quad (3)$$

Indeed, the above expression fits the data very well (Fig. 3) giving an estimate for a DNA persistence length of 51 ± 1 nm, consistent with the known value of ~50 nm.

Furthermore, the SFCS approach in combination with specific fluorescent labeling allows us to address a diverse set of questions, similar to the application of neutron scattering to flexible polymer systems where chains can be selectively deuterated [26–28]. Here, we probe the structure of individual DNA coils in semidilute solutions (in contrast, our previous study focused on the overall structure of such solutions [20]). A small fraction of covalently labeled DNA chains was mixed with a large amount of unlabeled chains (apart from a single point in the dilute regime where the concentrations of both types of chains are small). The labeled DNA has a length of 4.8 μm (14.2 kbp) and its concentration is kept low at ~1–2 μg/mL, so that labeled chains do not interact directly. Unlabeled DNA is of 16.5 μm (λ-DNA) and its concentration is varied from the

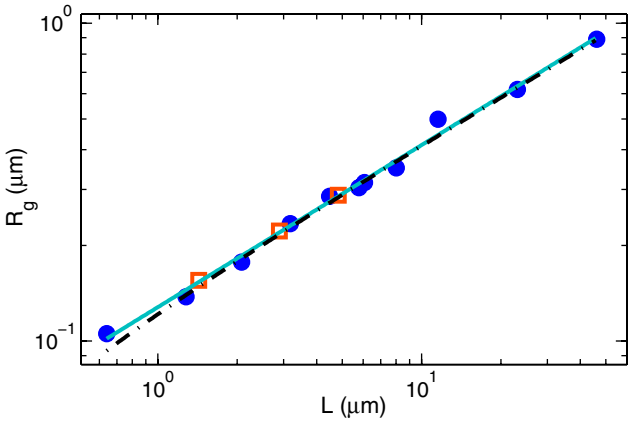

FIG. 3 (color online). The dependence of R_g on DNA contour length. Data were obtained on DNA stained with EtBr (circles) and with covalently bound Rh6G dye (squares). The presented values of R_g were extracted from the direct fits to SFCS correlation functions as in Fig. 1. The fits to $S(q)$ (Fig. 2) give R_g values lower by ~5% (not shown). Error bars are close to symbol size. Solid line: The best power law fit $R_g \propto L^{0.52\pm0.02}$. Dashed line: Fit with Eq. (3) for wormlike chain giving $l_p = 51 \pm 1$ nm.

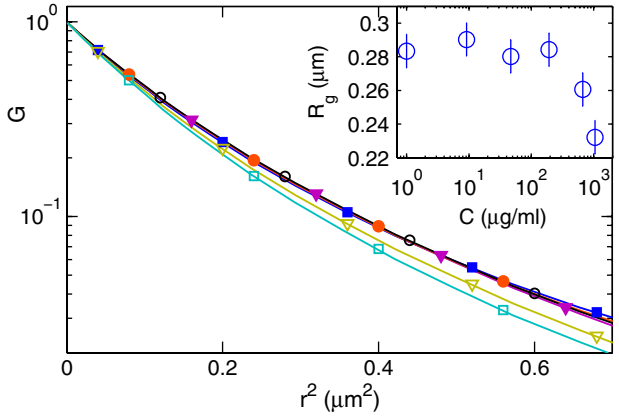

FIG. 4 (color online). Normalized SFCS correlation functions of individual covalently labeled 4.8 μm DNA chains mixed with unlabeled 16.5 μm chains at different densities. The lines marked with different symbols correspond to total DNA concentrations of: 1 (full triangles), 9 (full squares), 47 (full circles), 190 (open circles), 670 (open triangles), and 1040 μg/mL (open squares). Symbols are placed sparsely (not at every data point) to avoid clutter.

dilute regime $\sim$1 μg/mL through the onset of the semidilute regime at $\sim$10 μg/mL up to 1040 μg/mL [29].

The measured correlation functions for different concentrations of unlabeled DNA are presented in Fig. 4. Since unlabeled DNA is "invisible," the correlation functions reflect the structure of labeled chains. Up to the concentration $\sim$600 μg/mL there are no changes in the chain conformation. We stress that this happens deep in the semidilute regime. In this concentration range, the polymer solution screening length changes about tenfold from $\sim$1 μm to 100 nm [20]. The results point to an amazing feature of these solutions: large changes in the collective behavior (screening of spatial correlations) are brought about by the tiny rearrangements in the structure of individual coils. This contrasts with the behavior of flexible chains that significantly change their conformations and contract in semidilute solutions. Yet our results obtained on DNA correspond to the theoretical expectations for marginal solutions [4,30].

At high concentrations the SFCS correlation functions indicate a weak contraction of the DNA coil with increasing DNA density: $\sim$18% contraction over 5.5 fold increase in DNA concentration from 190 to 1040 μg/mL. We note that in these conditions, the solution is at the crossover to yet another new, concentrated regime in which the screening length is comparable to the DNA persistence length. Theory predicts no change in ξ and R_g with concentration in this regime [20]. So, the observed contraction of DNA coils at the highest probed concentrations contradicts theoretical predictions. Our measurements of ξ cover only a part of this range up to $\sim$400 μg/mL and within the whole range are consistent with the scaling predicted for marginal solutions [20]. Understanding the discrepancy between theory and experiment at the crossover to concentrated regime will require further measurements.

To conclude, we implemented a new approach to measure the structure of individual DNA coils in dilute and semidilute solutions. Up to 46 μm contour length, DNA coils do not exhibit significant excluded volume effects: their structure factor can be well described by the Debye expression for ideal chains and their $R_g(L)$ dependence follows closely that of an ideal wormlike chain. We present the first data on the structure of individual DNA coils in semidilute solutions. Up to the concentrated regime where solution screening length becomes comparable to DNA persistence length, the structure of DNA coils is unaffected by the presence of other coils. Overall, our data are consistent with the theoretical predictions for marginal solutions [4].

This work was supported by Israel Science Foundation Grant No. 984/09 and German-Israeli Foundation Grant No. 972-146.14/2007. M. N. was supported by EC FP7 Marie Curie program (Grant No. 215148).

*okrichev@bgu.ac.il

[1] P.-G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca and London, 1993).
[2] M. Daoud and G. Jannink, J. Phys. (Paris) **37**, 973 (1976).
[3] P. J. Flory, J. Chem. Phys. **17**, 303 (1949).
[4] D. W. Schaefer, J. F. Joanny, and P. Pincus, Macromolecules **13**, 1280 (1980).
[5] S. F. Edwards, Proc. Phys. Soc. **88**, 265 (1966).
[6] R. Verma, J. C. Crocker, T. C. Lubensky, and A. G. Yodh, Macromolecules **33**, 177 (2000).
[7] W. Brown and T. Nicolai, Colloid Polym. Sci. **268**, 977 (1990).
[8] P. Wiltzius, H. R. Haller, D. S. Cannell, and D. W. Schaefer, Phys. Rev. Lett. **51**, 1183 (1983).
[9] A. Bennett, P. Daivis, R. Shanks, and R. Knott, Polymer **45**, 8531 (2004).
[10] T. Uematsu, C. Svanberg, M. Nydén, and P. Jacobsson, Phys. Rev. E **68**, 051803 (2003).
[11] J. F. Marko and E. D. Siggia, Phys. Rev. E **52**, 2912 (1995).
[12] R. M. Robertson, S. Laib, and D. E. Smith, Proc. Natl. Acad. Sci. U.S.A. **103**, 7310 (2006).
[13] R. Lehner, J. Koota, G. Maret, and T. Gisler, Phys. Rev. Lett. **96**, 107801 (2006).
[14] F. Valle, M. Favre, P. De Los Rios, A. Rosa, and G. Dietler, Phys. Rev. Lett. **95**, 158105 (2005).
[15] J. Harpst and J. Dawson, Biophys. J. **55**, 1237 (1989).
[16] J. A. Harpst, Biophys. Chem. **11**, 295 (1980).
[17] T. Odijk, Biopolymers **18**, 3111 (1979).
[18] H. Yamakawa, *Modern Theory of Polymer Solutions* (Harper & Row, New York, 1971).
[19] P. Debye, J. Phys. Colloid Chem. **51**, 18 (1947).
[20] E. Shafran, A. Yaniv, and O. Krichevsky, Phys. Rev. Lett. **104**, 128101 (2010).
[21] S. B. Smith, Y. Cui, and C. Bustamante, Science **271**, 795 (1996).

[22] M. C. F. M. S. Rocha and O. N. Mesquita, J. Chem. Phys. **127**, 105 108 (2007).

[23] The conditions for field inversion electrophoresis: $0.2x$ TAE buffer, agarose gel concentration of 0.75%, switching frequency 0.25 Hz, 3 s forward and 1 s reverse pulse, 8 V/cm, 12 h total run time.

[24] A single forward primer CCGTTCTTCTTCGTCATAAC and different reverse primers GCA CTC TTT CTC GTA GGT ACT, CGCTTTATTACCATCCTCAG, and CACG CAGGGGAAATATCTTT were used to obtain DNA of 4.2, 8.5, and 14.2 kbp, respectively. In the reaction mixture (50 μL) AA-dUTP to dTTP ratio was 1:1, the concentrations of dNTPs, Mg, and primers were 0.4 mM, 5 mM, and 5 μM respectively, and the amounts of template and of polymerase (Vent exo-, New England Biolabs, Ipswich, MA) were 5 ng and 0.5 μL respectively. Template denaturation at 98 °C for 10 s was followed by 30 s annealing at 3 °C above the lower annealing temperature recommended by the primer manufacturer (Midland Certified) for a given primer pair. Extension temperature was set at 68 °C and extension time was kept at 45 s per 1 kbp for the first 15 cycles and increased to 2 min/1 kbp for the next 20 cycles. Following PCR, DNA was purified from salts, unused primers, and dNTPs (DNA purification kit, GE Healthcare), reacted overnight with Rh6G dye in 0.1 M $NaHCO_3$ buffer and then separated from nonspecific PCR products and unreacted dye by gel electrophoresis. In the synthesized sample the typical dye to base pair ratio was $\sim$1:200 as measured by the amplitude of FCS correlation functions.

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

[26] M. Daoud, J. P. Cotton, B. Farnoux, G. Jannink, G. Sarma, H. Benoit, R. Duplessix, C. Picot, and P.-G. deGennes, Macromolecules **8**, 804 (1975).

[27] B. Farnoux, F. Boué, J. P. Cotton, M. Daoud, G. Jannink, M. Nierlich, and P.-G. deGennes, J. Phys. (Paris) **39**, 77 (1978).

[28] L. H. Sperling, Polym. Eng. Sci. **24**, 1 (1984).

[29] The "background" 16 μm long DNA has $c^* \approx 10\,\mu$g/mL. The overlap concentration for labeled chains of 4.8μm is about twice larger.

[30] R. Verma, J. C. Crocker, T. C. Lubensky, and A. G. Yodh, Phys. Rev. Lett. **81**, 4004 (1998).