

Extension of nanoconfined DNA: Quantitative comparison between experiment and theoryV. Iarko,¹ E. Werner,¹ L. K. Nyberg,² V. Müller,² J. Fritzsche,³ T. Ambjörnsson,⁴ J. P. Beech,⁵ J. O. Tegenfeldt,^{5,6} K. Mehlig,⁷ F. Westerlund,² and B. Mehlig¹¹*Department of Physics, University of Gothenburg, 412 96 Göteborg, Sweden*²*Department of Biology and Biological Engineering, Chalmers University of Technology, 412 96 Göteborg, Sweden*³*Department of Applied Physics, Chalmers University of Technology, 412 96 Göteborg, Sweden*⁴*Department of Astronomy and Theoretical Physics, Lund University, 22 100 Lund, Sweden*⁵*Department of Physics, Division of Solid State Physics, Lund University, 22 100 Lund, Sweden*⁶*NanoLund, Lund University, 22 100 Lund, Sweden*⁷*Department of Public Health and Community Medicine, University of Gothenburg, 413 46 Göteborg, Sweden*

(Received 5 June 2015; revised manuscript received 2 November 2015; published 2 December 2015)

The extension of DNA confined to nanochannels has been studied intensively and in detail. However, quantitative comparisons between experiments and model calculations are difficult because most theoretical predictions involve undetermined prefactors, and because the model parameters (contour length, Kuhn length, effective width) are difficult to compute reliably, leading to substantial uncertainties. Here we use a recent asymptotically exact theory for the DNA extension in the “extended de Gennes regime” that allows us to compare experimental results with theory. For this purpose, we performed experiments measuring the mean DNA extension and its standard deviation while varying the channel geometry, dye intercalation ratio, and ionic strength of the buffer. The experimental results agree very well with theory at high ionic strengths, indicating that the model parameters are reliable. At low ionic strengths, the agreement is less good. We discuss possible reasons. In principle, our approach allows us to measure the Kuhn length and the effective width of a single DNA molecule and more generally of semiflexible polymers in solution.

DOI: [10.1103/PhysRevE.92.062701](https://doi.org/10.1103/PhysRevE.92.062701)

PACS number(s): 87.14.gk, 36.20.Ey, 87.15.-v

I. INTRODUCTION

Nanoconfinement has recently been studied intensively [1–6] as a means of stretching DNA molecules in order to study local properties (e.g., a DNA sequence [7,8]). A fundamental question is how the physical properties of the DNA and the solution affect the extent to which the molecule is stretched by confinement. Experimentally this question has been investigated in detail, varying the confinement, the length of the DNA molecule, and the properties of the solution (see Ref. [1] for a review).

It is commonly assumed that DNA can be modeled as a semiflexible polymer with hard-core repulsive interactions [9–15]. Measurements [16–19] and theoretical considerations [9,16] indicate that a wormlike chain model may be a good approximation. A recent study [3] compares experimental results for the extension of confined λ -DNA to results of computer simulations of a self-avoiding discrete wormlike chain model, indicating that it may describe the experimental results well.

Yet quantitative comparison between experiments and theoretical model calculations has remained difficult for at least two reasons. First, the model parameters (contour length, Kuhn length ℓ_K , and effective width w_{eff}) of the semiflexible polymer are difficult to determine reliably. Second, the model is hard to analyze theoretically. Therefore, very few exact predictions exist for experimentally measurable properties. However, a recent asymptotically exact theory [20,21] for the extension of a confined self-avoiding semiflexible polymer in the so-called “extended de Gennes regime” [11,13,22] overcomes the second difficulty: it makes precise predictions for the prefactors and exponents defining scaling laws as a function of the physical parameters; see Eq. (3) below. This

opens the possibility to experimentally determine ℓ_K and w_{eff} by measurements of confined DNA. In this article, we report on experimental results mapping out how the extension of confined DNA in the extended de Gennes regime depends on the channel geometry, the ionic strength of the solution, and the dye-loading of the molecule.

At high ionic strengths, we find very good agreement between experiment and theory using approximations for ℓ_K and w_{eff} that are commonly employed [23–27], and taking into account how the contour length depends on the dye loading. The comparison between experiment and theory is so precise that it enables us to detect subtle alignment effects [28] at the border of the extended de Gennes regime. At low ionic strengths, the agreement is not as good. This may indicate that theoretical estimates of ℓ_K and w_{eff} must be improved. We expect it is possible experimentally to determine precisely ℓ_K and w_{eff} by extending the approach described in this article.

II. EXTENDED de GENNES REGIME

Consider a semiflexible polymer of contour length L , Kuhn length ℓ_K [9], and excluded volume v per Kuhn length segment. The excluded volume is often written in terms of an effective width w_{eff} , defined by the relation

$$v \equiv (\pi/2)\ell_K^2 w_{\text{eff}}. \quad (1)$$

This expression for the excluded volume is based on Onsager’s result [29] for the excluded volume of a cylinder of length ℓ_K and diameter w_{eff} , which in the limit $\ell_K \gg w_{\text{eff}}$ reduces to the expression above. We phrase our results in terms of the effective width w_{eff} . This is customary, but we stress that, strictly speaking, it is the excluded volume that determines the statistics of the polymer, and that even for DNA models

with only hard-core repulsion between rodlike segments, the effective width does not equal the actual width of the rods, except in the limit $\ell_K \gg w_{\text{eff}}$. For the commonly used touching-bead model [12], w_{eff} does not equal the bead diameter even in this limit, and it is not known how the excluded volume depends on the model parameters (bead size and bending stiffness).

The polymer is confined to a channel with cross section $D_W \times D_H$. The polymer exhibits different confinement regimes distinguished by different laws for the polymer extension in the channel direction [21]. The extended de Gennes regime is defined by the conditions

$$\ell_K \ll D_H \ll \ell_K^2/w_{\text{eff}} \quad \text{and} \quad D_W^2 \ll D_H \ell_K^2/w_{\text{eff}}, \quad (2)$$

where we assume that $D_W \geq D_H$. For a square channel, the corresponding conditions were previously derived and discussed in Refs. [10,11,13,20,22]. In regime (2), exact expressions for the mean μ and the variance σ^2 of the extension in the channel direction are known [20,21], provided that the contour length is long enough:

$$\mu/L = 0.9338(84)[\ell_K w_{\text{eff}}/(D_W D_H)]^{1/3}, \quad (3a)$$

$$\sigma/(L\ell_K)^{1/2} = 0.364(17). \quad (3b)$$

The errors quoted for the coefficients reflect strict bounds [20] derived from the exact results of Ref. [30]. Provided L is known, Eq. (3) allow us to infer ℓ_K and w_{eff} from measurements of nanoconfined DNA molecules.

III. CALCULATION OF PARAMETERS

We now discuss how the parameters L , ℓ_K , and w_{eff} are commonly estimated for DNA, and what the main uncertainties are. We first discuss bare DNA before considering the effect of staining with fluorescent dye (YOYO-1).

The contour length of bare DNA is 0.34 nm per base pair [31]. DNA is commonly modeled as a wormlike chain, for which the Kuhn length is twice the persistence length, $\ell_K = 2\ell_P$ [9]. The persistence length has been measured by a number of different techniques [16,17,19,32] yielding $\ell_P \approx 45\text{--}50$ nm at high ionic strength ($I_s \approx 100$ mM) [33], and increasing at lower I_s . However, there is no consensus regarding how rapidly the persistence length increases with decreasing ionic strength [33,34]. Three different theoretical expressions for the dependence of the persistence length upon the ionic strength are in common use. The classical theory of Odijk [23] and Skolnick and Fixman [24] (OSF) predicts that the persistence length is given as a sum of two terms: a ‘‘bare’’ persistence length, which is independent of I_s , plus an electrostatic persistence length, proportional to I_s^{-1} . Estimating the effective line charge density of the DNA backbone by Manning’s condensation theory [35] and assuming that the bare persistence length is 50 nm yields [32]

$$\ell_P = 50 \text{ nm} + 0.0324 I_s^{-1} \text{ M nm}. \quad (4)$$

By contrast, Dobrynin and Rubinstein [36] propose that the electrostatic persistence length should be proportional to $I_s^{-1/2}$. Dobrynin [25] suggests the empirical formula

$$\ell_P = 46.1 \text{ nm} + 1.92 I_s^{-1/2} \text{ M}^{1/2} \text{ nm}. \quad (5)$$

TABLE I. Numerical values for the ionic strength I_s (calculated in Appendix C), the Kuhn length ℓ_K according to three different theories (see text), and the effective width w_{eff} according to Stigter’s theory [27,37].

Buffer	$0.05 \times \text{TBE}$	$0.5 \times \text{TBE}$	$2 \times \text{TBE}$	$5 \times \text{TBE}$
I_s (mM)	3.81	24.9	78.4	178.0
ℓ_K (nm) (OSF)	117	103	101	100
ℓ_K (nm) (Dobrynin)	154	117	106	101
ℓ_K (nm) (Manning)	140	124	112	103
w_{eff} (nm) (Stigter)	26	10	6.2	4.6

Finally, Manning [26] derives an expression for the ionic strength dependence of the persistence length by comparing a DNA molecule to a hypothetical molecule that is identical to real DNA except that its phosphate groups are neutral rather than negatively charged. The resulting Eq. (25) of Ref. [26] exhibits a more complicated dependence on the ionic strength than Eqs. (4) and (5). The resulting values of ℓ_K from the three competing theories are given in Table I and plotted in Fig. 1.

The standard theory for the effective width of DNA was derived by Stigter [37]. He computed the excluded volume between two long, strongly charged cylinders in NaCl solution, and he applied this calculation to DNA to obtain an estimate of w_{eff} . Linear interpolation on a doubly logarithmic scale of the effective widths given in Table I of Ref. [27] yields the values tabulated in Table I. There are many sources of uncertainty when applying this theory to our system. Stigter’s calculation for w_{eff} assumes that the Kuhn length segments can be approximated by infinitely long cylinders with an intrinsic width of 1.2 nm, and that the effective line charge of DNA is given by $0.73e^-$ per phosphate group. The approximation of infinite cylinders is problematic when the Kuhn length and the effective width are of the same order, i.e., for low ionic strengths. According to Stigter, the value 1.2 nm has an uncertainty of about 20%, leading to an uncertainty for the effective width of 5–10 % [37], with a larger effect at large ionic strengths. The effective line-charge estimate is based on measurements in NaCl solutions, which do not generalize to other ions [38].

The DNA contour length is expected to increase in proportion to the amount of dye bound. We assume that dye

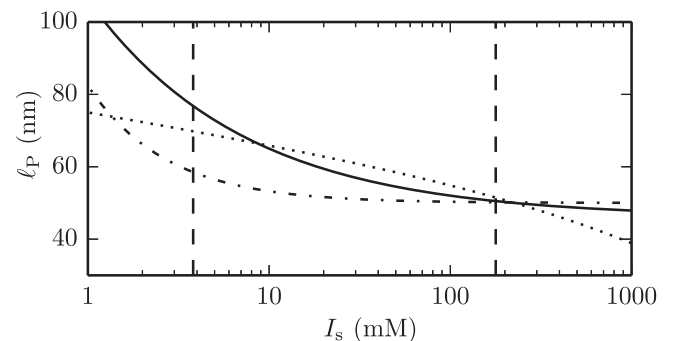


FIG. 1. Comparison of different theories for the dependence of the persistence length of DNA upon the ionic strength. Dash-dotted line: OSF theory, Eq. (4). Solid line: Dobrynin’s formula, Eq. (5). Dotted line: Manning’s theory, Eq. (25) of Ref. [26]. Dashed lines indicate the range of ionic strengths explored in this paper.

intercalation extends the bare contour length of the DNA molecule by 0.44 nm per dye molecule [3]. Estimates of this number range from about 0.4 to 0.5 nm, with large uncertainties in the individual estimates [39–41]; the uncertainty is at least 10%. At a dye loading of 1 mol per 10 base pairs, this corresponds to an uncertainty in the contour length of $\approx 2\%$. There is no consensus regarding how intercalating dye molecules affect the parameters ℓ_K and w_{eff} . Reference [42] finds that the Kuhn length decreases with increasing dye load, whereas Refs. [40,41] find no dependence. Since the dye molecules are positively charged, the effective width might decrease with dye load, but the magnitude of this effect is not known. Lacking a better estimate, it is commonly assumed that YOYO-binding does not affect these parameters [3,43,44]. In summary, while the contour length of DNA is known to a rather high accuracy, there is substantial uncertainty regarding the parameters ℓ_K and w_{eff} .

IV. EXPERIMENTAL METHOD

The experimental data are obtained by measuring the extension of single DNA molecules in nanochannels under different conditions [Figs. 2(a)–2(c)]. We use linear λ - and T4GT7-DNA (T4-DNA for short) with definite contour lengths of 48 502 and 165 647 base pairs, respectively [this ensures that L is large enough in our experiments; it exceeds $(D_W^2 D_H^2 \ell_K / w_{\text{eff}}^2)^{1/3}$ by an order of magnitude [21]]. The molecules are stained with YOYO dye and suspended inside a channel in a TBE (Tris-borate-EDTA) buffer.

The first experiment is a reanalysis of data presented in Ref. [45]. In this experiment [Fig. 2(a)], λ -DNA is inserted into a nanochannel of height $D_H = 150$ nm and width $D_W = 108$ nm. We discuss the uncertainty in the channel dimensions

in Appendix A. DNA extensions are measured at different buffer conditions ($0.05\times$, $0.5\times$, $2\times$, and $5\times$ TBE) and at different dye loads. To estimate the dye load of a molecule, we assume that it is proportional to the luminosity, and that the largest observed luminosity corresponds to full intercalation (one dye molecule per four base pairs). In this way, we obtain an estimate for the amount of dye bound to the molecule by linear interpolation.

In experiment 2 [Fig. 2(b)], T4-DNA is inserted into a nanofunnel, with fixed height $D_H = 120$ nm and gradually changing width from $D_W = 92$ to 815 nm over a length of 500 μm . These experiments are performed at two different buffer concentrations ($0.05\times$ and $2\times$ TBE).

In experiment 3 [Fig. 2(c)], T4-DNA is inserted into a channel with $D_W = 302$ nm, $D_H = 300$ nm. The buffer concentration is varied ($0.05\times$, $0.5\times$, $2\times$, and $5\times$ TBE). In experiments 2 and 3, the average dye load at $0.05\times$ TBE is approximately 1 dye molecule per 10 base pairs. Assuming that the dye load is proportional to luminosity, we estimate the dye load in experiment 2 at $2\times$ TBE to 1 dye molecule per 45 base pairs, and 1 per 12, 16, and 28 base pairs at $0.5\times$, $2\times$, and $5\times$ TBE in experiment 3.

For each molecule, 200 frames are recorded. Figure 2(d) shows an example of a fluorescence-intensity trace (“kymograph”) obtained in this way. Each row in the kymograph shows the fluorescence intensity in a given frame averaged over the channel cross section. Bright regions correspond to high intensity indicating where the DNA molecule is located. The extension of the molecule along the channel is simply the width of the bright region.

For a given set of parameters, we estimate the mean and standard deviation of the extension by a linear mixed model that takes into account the fact that the measured extensions are correlated in time. Details concerning the experimental method and the data analysis are given in Appendixes A and B.

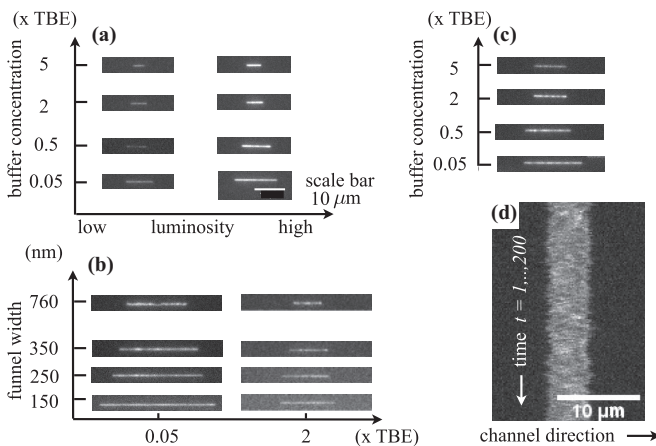


FIG. 2. (a) Experiment 1. λ -DNA in a 150 nm \times 108 nm channel, different buffer concentrations, and different luminosities corresponding to different dye loadings. Shown are representative video frames [the scale bar applies to panels (b) and (c) as well]. (b) Experiment 2. T4-DNA in a nanofunnel in $0.05\times$ and $2\times$ TBE solution, varying funnel width D_W at constant $D_H = 120$ nm. (c) Experiment 3. T4-DNA in a 302 nm \times 300 nm channel, with different buffer concentrations. (d) Time trace of the fluorescence intensity for λ -DNA in a 108 nm \times 150 nm channel in $5\times$ TBE solution, center-of-mass motion subtracted.

V. RESULTS

We first compare our experimental measurements to Eq. (3) using Dobrynin’s theory for ℓ_K and Stigter’s theory for w_{eff} . The results are shown in Fig. 3. We plot two theoretical curves. The solid curve uses the actual channel size $D_H \times D_W$. The dashed curve compensates for the repulsive interaction with the negatively charged walls [1] by using an “effective channel size” $(D_H - \delta) \times (D_W - \delta)$. We take $\delta = w_{\text{eff}}$, but it is not known how accurate this estimate is. In fact, not even the sign of the correction to Eq. (3a) is obvious, as one effect of violating the condition $D_H \gg \ell_K$ is that the polymer segments are more evenly distributed throughout the channel [46]. This leads to a smaller extension than predicted by Eq. (3a), i.e., in the opposite direction to the effect of the repulsive interaction between the DNA and the channel walls. Since the standard deviation is independent of channel size in the extended de Gennes regime, the compensation does not affect this comparison.

The results of experiment 1 are shown in panels (a) and (b). At low relative luminosity (small dye-to-basepair ratio), the average extension is well described by Eq. (3a). For the standard deviation there are larger differences between experiment and Eq. (3b). Possible reasons are discussed below.

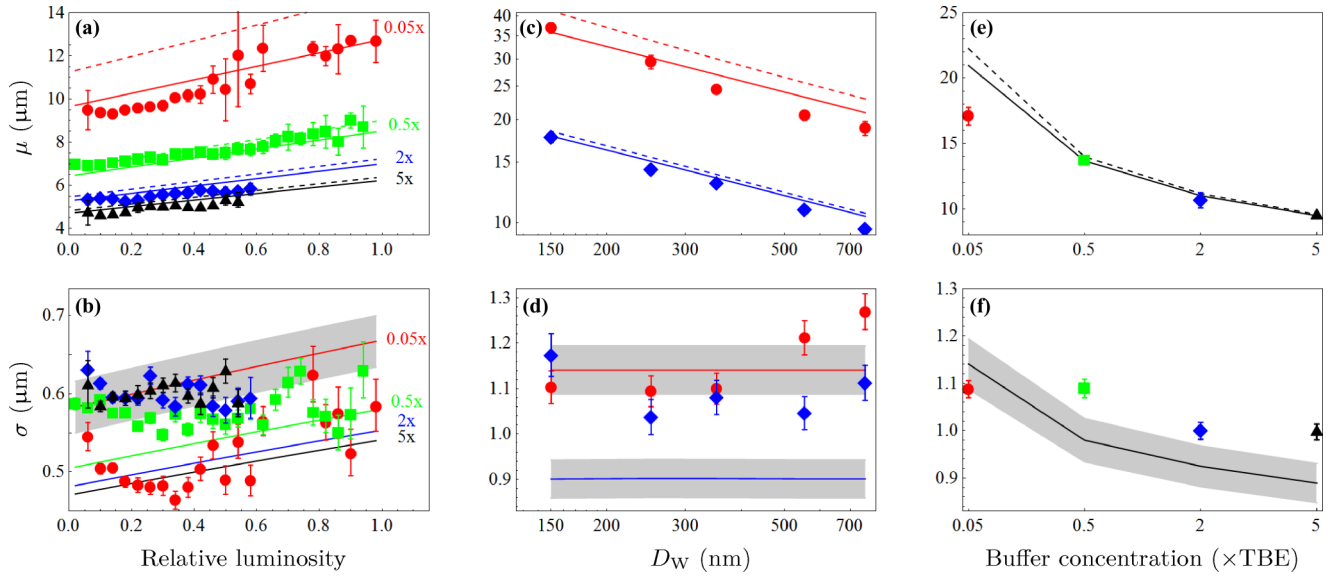


FIG. 3. (Color online) Experimental results for $0.05 \times$ TBE (red \circ), $0.5 \times$ TBE (green \square), $2 \times$ TBE (blue \diamond), and $5 \times$ TBE (black \triangle). (a) and (b) Experiment 1. Mean and standard deviation of the extension of λ -DNA in a narrow nanochannel, as a function of relative luminosity. Theory [Eq. (3)], solid lines. The rigorous bounds on the prefactor in Eq. (3b) are indicated as a shaded region for $0.05 \times$ TBE; they are of the same order for the other cases. The corresponding uncertainty for the extension is much smaller and not shown. The dashed line shows theory corrected for wall repulsion (see the text). (c) and (d) Experiment 2. Same, but for T4-DNA in a nanofunnel with varying width D_W . Note that panel (c) is a log-log plot. (e) and (f) Experiment 3. Same, but for T4-DNA in a wider square nanochannel, as a function of buffer concentration ($x \times$ TBE). Error bars correspond to 95% confidence intervals from the statistical analysis; the measurement uncertainty is not taken into account.

We turn now to the effect of increasing the dye-to-basepair ratio of the DNA. The theoretical lines are calculated under the assumption that each dye molecule increases the contour length by 0.44 nm but leaves the Kuhn length and the effective width unchanged. This yields estimates of the mean and standard deviation that overestimate the observables at high ionic strengths and high dye loads. A simple explanation would be that the persistence length decreases slightly with increasing dye load, in agreement with Ref. [42] though not with Refs. [40,41]. Note that since experiments 2 and 3 were performed at low dye-to-basepair ratios, such a decrease would not significantly influence the interpretations of these experiments.

The results of experiment 2 are shown in panels (c) and (d). Again, the experimental results are in qualitative agreement with the theoretical predictions. We see that the average extension agrees well with the theoretical prediction. However, the model predictions underestimate the standard deviation for the larger ionic strength and for the largest channel at the lower ionic strength.

It is important to note that for experiment 2 we do not expect perfect agreement with Eq. (3), since the condition $D_H \gg \ell_K$ is not satisfied, or only weakly satisfied. However, as long as $D_W \gg \ell_K$, the violation of the condition for D_H only affects the prefactors but not the power of D_W in Eq. (3). This follows from the fact that a mapping to a one-dimensional model is possible also when $D_H \approx \ell_K$ [21]. In accordance with this prediction, the data points satisfying $D_W \gg \ell_K$ in panel (c) obey the scaling $\mu \propto D_W^{-1/3}$ of [Eq. (3a)]. Similarly the data points at $2 \times$ TBE that satisfy $D_W \gg \ell_K$ show a variance that is approximately independent of D_W , in agreement with Eq. (3b),

Fig. 3(d). For the two rightmost data points at $0.05 \times$ TBE, the condition $D_W^2 \ll D_H \ell_K^2 / w_{\text{eff}}$ is violated. At this point, the variance is expected to start increasing as D_W increases further [21], in perfect agreement with what is observed.

For square channels simulations [11,15,22] show that the mean extension increases more rapidly with decreasing $D = D_H = D_W$ than Eq. (3a) predicts, when $D \approx \ell_K$. The reason is that there is a tendency for the DNA molecule to align with the wall, and that the presence of the walls makes it more difficult for the molecule to change direction in the channel, forming a “hairpin” [15,28]. This can explain why the average extension appears to increase slightly faster with decreasing width than Eq. (3a) predicts, for the leftmost points in panel (c). Such a trend has also been observed in previous measurements in rectangular channels [28,47].

Now consider the standard deviation. The alignment and correlation effects mentioned above cause σ to be overestimated [28]. But when μ approaches the maximal extension L , fluctuations are suppressed [15]. These two effects could explain why, in experiments 1 and 2, σ is larger than predicted by theory at high ionic strengths but smaller for low ionic strengths and small channel sizes. It must also be noted that the standard deviation is difficult to estimate precisely, as it is not very much larger than the pixel size in the image (159 nm), and it may depend on the assumptions entering into the statistical analysis (see Appendix B). An additional source of uncertainty specific to experiment 2 is that D_W changes over the span of the molecule. For the most extended condition ($\mu = 37 \mu\text{m}$), the channel width at either end of the molecule differs by approximately 25 nm from the stated width, measured at the center of the molecule.

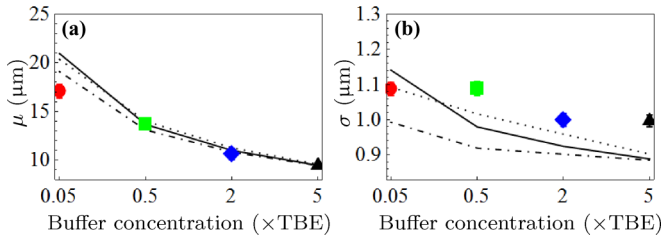


FIG. 4. (Color online) Comparison of different theories for the persistence length. Symbols: same experimental measurements as in Figs. 3(e) and 3(f). Lines: predictions for the extension statistics [Eq. (3)], using different theories for ℓ_K . Dash-dotted line: OSF theory, Eq. (4). Solid line: Dobrynin's expression, Eq. (5). Dotted line: Manning's theory, Eq. (25) of Ref. [26].

The results of experiment 3 are shown in panels (e) and (f). Here Eq. (2) is well satisfied. Simulations indicate [15,22] that the alignment effects discussed above have little influence on μ , in square channels with $D \approx 3\ell_K$. Equally sensitive simulation results for σ have not been published, but simulations of the alignment effect [15,28] indicate that Eq. (3b) underestimates σ by approximately 10%. We find that for the three largest ionic strengths, measurements are in excellent agreement with theory. The mean extension [panel (f)] agrees very well with the theoretical prediction of Eq. (3a), and Eq. (3b) underestimates the standard deviation [panel (f)] by about 10%, just as the measurements of alignment effects would suggest.

Intriguingly, the relation between measurements and predictions is different at $0.05 \times \text{TBE}$ than at high ionic strengths. Both mean and standard deviation are smaller than expected. This is particularly surprising considering that alignment effects should be even stronger at low ionic strength (where ℓ_K is larger). The discrepancy might indicate that the standard model does not describe the physical parameters well at such low ionic strengths, possibly because the high relative concentration of BME significantly changes the buffer conditions, lowering the pH from ≈ 8.5 to ≈ 7.5 . Yet we note that it is hard to ensure uniform dye coverage under these conditions [45], which makes the experimental measurements more uncertain.

Finally, in Fig. 4 we again compare the results of experiment 3 against Eq. (3), this time using not only Dobrynin's formula (5) for the Kuhn length, but also the OSF expression (4) and Manning's prediction [Eq. (25) of Ref. [26]]. Figure 4 shows that at high ionic strengths, the mean extension is well described by any of the three theories. At low ionic strengths, none of the theories accurately describes the measured extension, although the disagreement is smaller when using OSF theory instead of Dobrynin's formula in Eq. (3a). On the other hand, employing OSF theory in Eq. (3b) yields a prediction that differs somewhat more from experimental measurements of the standard deviation at higher ionic strengths. We conclude that further studies are necessary to resolve how the Kuhn length of DNA depends on the ionic strength.

VI. CONCLUSIONS

We have compared measurements of the extension statistics of confined DNA to asymptotically exact predictions. At high

ionic strengths, we find excellent agreement between experiments and theoretical predictions for the average extension, indicating that standard theories for the physical parameters L , ℓ_K , and w_{eff} of DNA are accurate under these conditions.

The results of this study also raise new questions. First, measurements of the average extension at low ionic strengths deviate from the theoretical predictions. This might indicate that common estimates for w_{eff} and ℓ_K of stained DNA fail under these conditions. Second, our measurements of the standard deviation of the extension differ from the predicted values by approximately 10%. We believe that this discrepancy is caused by alignment effects [15,28] at the border of the extended de Gennes regime. A quantitative understanding of these effects would enable one to use the methods of this study to precisely determine how ℓ_K and w_{eff} depend on the ionic strength of the solution. Careful measurements over longer times may even make it possible to determine ℓ_K and w_{eff} for single molecules. One strength of the method is that the simultaneous measurement of both the mean and the standard deviation of the extension allows one to determine both ℓ_K and w_{eff} from a single measurement series. Alternatively, if ℓ_K is already known to high precision, then a comparison between measurements of the average extension and Eq. (3a) allows for a precise determination of w_{eff} , a parameter that is otherwise difficult to measure.

Finally, by comparing measurements at many different channel sizes to theoretical predictions and computer simulations, it should be possible to investigate also DNA-wall interactions, a question about which little is known, and one that is hard to describe theoretically.

Note added. Recently, a paper [48] appeared online that also compares experimental measurements of channel-confined DNA to the theoretical predictions of Ref. [20].

ACKNOWLEDGMENTS

This work was made possible through support from the Swedish Research Council (Vetenskapsrådet), the Göran Gustafsson Foundation for Research in Natural Sciences and Medicine (B.M.), Chalmers Area of Advance in Nanoscience and Nanotechnology (F.W.), and NanoLund (J.P.B. and J.O.T.). We thank Charleston Noble and Erik Lagerstedt for helpful discussions.

APPENDIX A: EXPERIMENTAL PROCEDURE

1. Channel manufacture

The channels were manufactured as described in Ref. [49]. The nanochannels used in experiment 1 have a depth of $D_H = 150$ nm. The channel cross section is not perfectly rectangular, but rather trapezoidal, with a width at the top (bottom) of the channel of 130 nm (87 nm). We assume that these channels can be approximated by rectangular channels with a width $D_W = 108$ nm, which is the arithmetic mean of the measurements at the top and bottom of the channel. The funnels used in experiment 2 have a depth of approximately $D_H = 120$ nm, and D_W increases from 92 nm (top: 111 nm, bottom: 73 nm) at the narrow end to 815 nm (top: 830 nm, bottom: 800 nm) at the wide end—over a length of 500 μm . Finally, the channels used in experiment 3 have a depth $D_H = 300$ nm and a width

$D_w = 302$ nm (top: 330 nm, bottom: 275 nm). The dimensions are measured before the channel is closed by the lid. Since the bonding process changes the depth of the channel, our values for the depth of the channels have a relatively large uncertainty of about 10 nm.

2. Buffer preparation

The buffers were obtained by diluting $10 \times$ TBE tablets from Medicago to the desired concentration. $1 \times$ TBE contains 0.089 M Tris, 0.089 M borate, and 0.0020 M EDTA. Right before the experiments, 3% BME (3 μ L of BME to 97 μ L of sample solution) was added to prevent photoniccking of the DNA. In addition to this, the micro- and nanochannels in the chip were flushed with the right buffer concentration, containing 3% BME, to keep the same environment in the chip as in the sample solution.

3. Dye intercalation

Since intercalation of YOYO-dye molecules extends the contour length of the DNA, it is important to estimate the dye load as accurately as possible. In experiment 1, the dye load was not equilibrated between molecules, resulting in heterogeneous staining [45]. In experiments 2 and 3, the aim was instead to achieve homogeneous staining. Toward that end, we first mixed the sample in high ionic strength ($5 \times$ TBE), let it rest for at least 20 min, and then diluted to the desired ionic strength [45].

4. DNA insertion

The nanofluidic chips consist of four loading wells connected two and two by microchannels. The microchannels are in turn connected by many nanochannels. A sample solution was inserted into one of the loading wells. Then the DNA was carried to the nanochannels by pressure-driven flow (N_2 gas for T4-DNA experiments and air for λ -DNA experiments). The DNA was forced into the nanochannels by applying pressure over two wells that are connected by a microchannel. Once inserted in the channel, the pressure was switched off. Before imaging, the DNA was allowed to relax for about 30 s to reach its equilibrium extension.

5. Video recording

Videos were recorded using a Photometrics EvolveTM EMCCD camera. The image pixel size is 159.2 nm. For all measurements, the exposure time was 100 ms per photograph, but the delay between frames differed between experiments. For experiment 1, the delay between frames was 84 ms, yielding a frame rate of 184 ms/frame. For experiments 2 and 3 with T4-DNA, the correlation time is significantly longer, so to minimize problems with photobleaching and nicking, the frame rate was decreased. For experiment 2, the frame rate was 0.5 s/frame for the measurements at $0.05 \times$ TBE, and 2 s/frame for the ones at $2 \times$ TBE. For experiment 3, the frame rate was 1 s/frame for all measurements.

6. Number of molecules

In experiment 1, 2388 molecules were analyzed in total. Since the relative intensities are distributed unevenly, so are the number of molecules in each bin. This leads to very different error estimates for different bins in the estimation of the mean and standard deviation. We excluded bins with two molecules or fewer from the analysis. Another three molecules were excluded because the extension could not be successfully extracted from all frames. In experiment 2, nine to ten molecules were analyzed per data point under $0.05 \times$ TBE, and seven molecules under $2 \times$ TBE. In experiment 3, 30–35 molecules were analyzed per measured data point. One outlier at $0.5 \times$ TBE was removed, as its extension changed abruptly halfway through the data series.

APPENDIX B: DATA ANALYSIS

1. Extraction of DNA extension from kymographs

Consider a raw kymograph of the form provided in the main text [Fig. 1(d)]. The dark regions in this image correspond to background, while the bright regions correspond to fluorescence from DNA. The fluorescence intensities measured for both regions are subject to noise. This makes it difficult to reliably identify the end points of the bright regions in an automated fashion. In previous studies, this was achieved by fitting the difference of two sigmoidal functions [1] to each time frame. This method was used for analyzing experiment 1 [45]. In experiments 2 and 3, we used a computationally faster method that yields very similar results for the mean, and it agrees to within about 5% for the standard deviation of the extension. The new method for detecting end positions relies on the assumption that the fluorescence intensities from DNA and from the background assume consistent and easily differentiated values, and it proceeds as follows. First, each frame in the raw kymograph is smoothed using a moving average (we use an averaging window that is five pixels wide). Second, each frame is segmented into binary low- and high-intensity regions using Otsu's method [50,51]. Third, we even out gaps between high-intensity regions that are equal to or shorter than five pixels. Fourth, the largest connected high-intensity component is found, and its edges are identified. The distance between these edges is the extension of the DNA molecule.

2. Statistical analysis

The experimental data consist of repeated measurements of the extension of individual DNA molecules at given values of parameters and are modeled using a random-coefficient model, a variation of a linear-regression model adapted to correlated data [52]. Let y_{ij} denote the extension of molecule i ($i = 1, \dots, n$) at time t_j ($j = 1, \dots, 200$). The extension is modeled as a linear function of time as $y_{ij} = \mu + \beta t_j + u_i + e_{ij}$, where μ and β denote the overall mean value and slope. To account for the within-molecule correlation of data points, the error term is split into two uncorrelated components, a molecule-specific correction to the mean, u_i , and an error term e_{ij} that measures the deviation of a data point from the molecule-specific regression line, $\mu + \beta t_j + u_i$. Both u_i and e_{ij} are assumed to be independently normally distributed with

mean zero and variance r^2 and σ^2 , respectively. The model describes the mean extension of molecules at a particular set of external parameters, and the linear time dependence allows to include small unmeasured changes of external conditions, for instance drift into wider or narrower channel regions, or change of buffer concentration. The random intercept adjusts for the initial conditions at the start of the measurement of each time series. A random intercept model is adequate as each set of molecules can be viewed as a random sample representative for all molecules at this particular set of external parameters. On average, the covariance of two arbitrary data points within molecules is equal to the variance of intercepts between molecules. The ratio of variances, $r^2/(r^2 + \sigma^2)$, is both a measure for the proportion of variation that is explained by differences between molecules, and the correlation of data points within a typical time series. By isolating the molecule-specific random variation due to unknown fluctuations of external parameters, we are able to obtain an improved value for the residual variance of the extension (σ^2) that can be compared with the theory. We observed only very weak time trends ($|\beta| \ll \mu/200$), indicating that photodamage does not constitute a significant problem in our experiments. Nevertheless we evaluated the mean value μ at $t = 100$, i.e., we replaced t_j with $(t_j - 100)$ in the above model. Results are given in terms of point estimates together with a 95% confidence interval (95% CI). The statistical analyses were performed by using SAS (version 9.4; SAS Institute, Cary, NC).

APPENDIX C: CALCULATION OF IONIC STRENGTH

The parameters w and ℓ_K of the DNA molecule depend on the ionic strength of the surrounding solution. The ionic strength is defined as

$$I_s = \frac{1}{2} \sum_i c_i z_i^2. \quad (\text{C1})$$

Here, the sum runs over all ions in the solution, c_i is the concentration of ion species i , and z_i is its valence. To compute the ionic strength, we must calculate the equilibrium concentrations c_i of all ions. Since there is some confusion in the literature (discussed below) about the ionic strength of TBE buffer, we document our calculation in some detail.

We follow the method outlined in Sec. 12-2 of [53]. Denote by $C[X]$ the total amount of all species of a substance X , in neutral and ionic form. At $N \times$ TBE, the solution contains $C[\text{T}] = N \times 0.089$ M Tris, $C[\text{Bo}] = N \times 0.089$ M borate, $C[\text{E}] = N \times 0.0020$ M EDTA, and $C[\beta] = 0.429$ M BME. The equilibrium constants that determine the chemical equilibrium are Tris: $\text{p}K_b = 5.94$; borate: $\text{p}K_a = 9.24$, EDTA: $\text{p}K_a = \{1.99, 2.67, 6.16, 10.26\}$; BME: $\text{p}K_a = 9.6$. At the resulting pH value of approximately 8.5, most of the EDTA is triply ionized, and we can safely ignore the minute concentrations of neutral and singly ionized EDTA. We denote the molar concentration of a species by $[X]$, and its activity coefficient by γ_X . The activity coefficients are assumed to be given by the Davies equation, as stated in Ref. [53]. (The empirical prefactors of this equation differ between sources. For our ionic strength calculations, the difference between

different formulations makes a difference of 2% at most.) The system of equations that must be solved is

$$\frac{[\text{TH}^+]\gamma_{\text{TH}^+}[\text{OH}^-]\gamma_{\text{OH}^-}}{[\text{T}]\gamma_{\text{T}}} = 10^{-5.94}, \quad (\text{C2})$$

$$\frac{[\text{Bo}^-]\gamma_{\text{Bo}^-}[\text{H}^+]\gamma_{\text{H}^+}}{[\text{HBo}]\gamma_{\text{HBo}}} = 10^{-9.24}, \quad (\text{C3})$$

$$\frac{[\text{HE}^{3-}]\gamma_{\text{HE}^{3-}}[\text{H}^+]\gamma_{\text{H}^+}}{[\text{H}_2\text{E}^{2-}]\gamma_{\text{H}_2\text{E}^{2-}}} = 10^{-6.16}, \quad (\text{C4})$$

$$\frac{[\text{E}^{4-}]\gamma_{\text{E}^{4-}}[\text{H}^+]\gamma_{\text{H}^+}}{[\text{HE}^{3-}]\gamma_{\text{HE}^{3-}}} = 10^{-10.26}, \quad (\text{C5})$$

$$\frac{[\beta^-]\gamma_{\beta^-}[\text{H}^+]\gamma_{\text{H}^+}}{[\text{H}\beta]\gamma_{\text{H}\beta}} = 10^{-9.6}, \quad (\text{C6})$$

$$[\text{H}^+]\gamma_{\text{H}^+}[\text{OH}^-]\gamma_{\text{OH}^-} = 10^{-14.0}, \quad (\text{C7})$$

$$[\text{T}] + [\text{TH}^+] = C[\text{T}], \quad (\text{C8})$$

$$[\text{HBo}] + [\text{Bo}^-] = C[\text{Bo}], \quad (\text{C9})$$

$$[\text{H}_2\text{E}^{2-}] + [\text{HE}^{3-}] + [\text{E}^{4-}] = C[\text{E}], \quad (\text{C10})$$

$$[\text{H}\beta] + [\beta^-] = C[\beta], \quad (\text{C11})$$

$$[\text{TH}^+] + [\text{H}^+] = [\text{OH}^-] + [\text{Bo}^-] + 2[\text{H}_2\text{E}^{2-}] + 3[\text{HE}^{3-}] + 4[\text{E}^{4-}] + [\beta^-]. \quad (\text{C12})$$

Equations (C2)–(C7) are the equilibrium conditions from the law of mass action, Eqs. (C8)–(C11) ensure that the total concentration of a substance is conserved, and Eq. (C12) ensures charge neutrality. The activity coefficients γ_X depend on the ionic strength I_s according to the Davies equation [53],

$$\log_{10} \gamma_X = -0.51 z_X^2 \left(\frac{\sqrt{I_s}}{1 + \sqrt{I_s}} - 0.3 I_s \right). \quad (\text{C13})$$

Here, z_X is the valence of species X , and I_s is measured in units of M. Finally, the ionic strength itself is given according to Eq. (C1) by

$$I_s = 1/2([\text{TH}^+] + [\text{H}^+] + [\text{OH}^-] + [\text{Bo}^-] + 4[\text{H}_2\text{E}^{2-}] + 9[\text{HE}^{3-}] + 16[\text{E}^{4-}] + [\beta^-]). \quad (\text{C14})$$

We solve the system of equations (C2)–(C14) iteratively. Starting from the initial guess $\gamma_X = 1$ for all species, Eqs. (C2)–(C12) are solved with the MATHEMATICA routine FINDINSTANCE. Plugging the resulting concentrations into Eqs. (C13) and (C14) yields new values for γ_X , which are then used in Eqs. (C2)–(C12). This process was repeated until the ionic strength converged.

We tested our calculations by comparing with Table 1 of Ref. [44]. We reproduce the reported ionic strengths for 0% BME and 0.5% BME to within 1%.

- [1] W. Reisner, J. N. Pedersen, and R. H. Austin, *Rep. Prog. Phys.* **75**, 106601 (2012).
- [2] J. J. Jones, J. R. C. van der Maarel, and P. S. Doyle, *Phys. Rev. Lett.* **110**, 068101 (2013).
- [3] D. Gupta, J. Sheats, A. Muralidhar, J. J. Miller, D. E. Huang, S. Mahshid, K. D. Dorfman, and W. Reisner, *J. Chem. Phys.* **140**, 214901 (2014).
- [4] A. Khorshid, P. Zimny, D. Tétreault-La Roche, G. Massarelli, T. Sakaue, and W. Reisner, *Phys. Rev. Lett.* **113**, 268104 (2014).
- [5] M. Alizadehheidari, E. Werner, C. Noble, M. Reiter-Schad, L. K. Nyberg, J. Fritzsche, B. Mehlig, J. O. Tegenfeldt, T. Ambjörnsson, F. Persson, and F. Westerlund, *Macromolecules* **48**, 871 (2015).
- [6] W. F. Reinhart, J. G. Reifengerger, D. Gupta, A. Muralidhar, J. Sheats, H. Cao, and K. D. Dorfman, *J. Chem. Phys.* **142**, 064902 (2015).
- [7] E. Lam, A. Hastie, C. Lin, D. Ehrlich, S. K. Das, M. D. Austin, P. Deshpande, H. Cao, N. Nagarajan, M. Xiao, and P.-Y. Kwok, *Nat. Biotechnol.* **30**, 771 (2012).
- [8] A. N. Nilsson, G. Emilsson, L. K. Nyberg, C. Noble, L. S. Stadler, J. Fritzsche, E. R. B. Moore, J. O. Tegenfeldt, T. Ambjörnsson, and F. Westerlund, *Nucl. Acids Res.* **42**, e118 (2014).
- [9] A. Y. Grosberg and A. R. Khokhlov, *Statistical Physics of Macromolecules* (AIP Press, New York, 1994).
- [10] T. Odijk, *Phys. Rev. E* **77**, 060901 (2008).
- [11] Y. Wang, D. R. Tree, and K. D. Dorfman, *Macromolecules* **44**, 6594 (2011).
- [12] D. R. Tree, Y. Wang, and K. D. Dorfman, *Phys. Rev. Lett.* **108**, 228105 (2012).
- [13] L. Dai and P. S. Doyle, *Macromolecules* **46**, 6336 (2013).
- [14] A. Muralidhar, D. R. Tree, Y. Wang, and K. D. Dorfman, *J. Chem. Phys.* **140**, 084905 (2014).
- [15] A. Muralidhar, D. R. Tree, and K. D. Dorfman, *Macromolecules* **47**, 8446 (2014).
- [16] P. J. Hagerman, *Annu. Rev. Biophys. Biophys. Chem.* **17**, 265 (1988).
- [17] S. B. Smith, L. Finzi, and C. Bustamante, *Science* **258**, 1122 (1992).
- [18] M. D. Wang, H. Yin, R. Landick, J. Gelles, and S. M. Block, *Biophys. J.* **72**, 1335 (1997).
- [19] C. Bouchiat, M. D. Wang, J.-F. Allemand, T. Strick, S. M. Block, and V. Croquette, *Biophys. J.* **76**, 409 (1999).
- [20] E. Werner and B. Mehlig, *Phys. Rev. E* **90**, 062602 (2014).
- [21] E. Werner and B. Mehlig, *Phys. Rev. E* **91**, 050601(R) (2015).
- [22] L. Dai, J. van der Maarel, and P. S. Doyle, *Macromolecules* **47**, 2445 (2014).
- [23] T. J. Odijk, *Sci. B Polym. Phys.* **15**, 477 (1977).
- [24] J. Skolnick and M. Fixman, *Macromolecules* **10**, 944 (1977).
- [25] A. V. Dobrynin, *Macromolecules* **39**, 9519 (2006).
- [26] G. S. Manning, *Biophys. J.* **91**, 3607 (2006).
- [27] D. Stigter, *Cell Biophys.* **11**, 139 (1987).
- [28] E. Werner, F. Persson, F. Westerlund, J. O. Tegenfeldt, and B. Mehlig, *Phys. Rev. E* **86**, 041802 (2012).
- [29] L. Onsager, *Ann. N.Y. Acad. Sci.* **51**, 627 (1949).
- [30] R. van der Hofstad, F. den Hollander, and W. König, *Probab. Theor. Relat. Fields* **125**, 483 (2003).
- [31] R. R. Sinden, *DNA Structure and Function* (Elsevier, Amsterdam, 2012).
- [32] C. G. Baumann, S. B. Smith, V. A. Bloomfield, and C. Bustamante, *Proc. Natl. Acad. Sci. USA* **94**, 6185 (1997).
- [33] A. Brunet, C. Tardin, L. Salomé, P. Rousseau, N. Destainville, and M. Manghi, *Macromolecules* **48**, 3641 (2015).
- [34] A. Savel'yev, *Phys. Chem. Chem. Phys.* **14**, 2250 (2012).
- [35] G. S. Manning, *J. Chem. Phys.* **51**, 924 (1969).
- [36] A. Dobrynin and M. Rubinstein, *Prog. Polym. Sci.* **30**, 1049 (2005).
- [37] D. Stigter, *Biopolymers* **16**, 1435 (1977).
- [38] J. A. Schellman and D. Stigter, *Biopolymers* **16**, 1415 (1977).
- [39] F. Johansen and J. P. Jacobsen, *J. Biomol. Struct. Dyn.* **16**, 205 (1998).
- [40] K. Günther, M. Mertig, and R. Seidel, *Nucl. Acids Res.* **38**, 6526 (2010).
- [41] B. Kundukad, J. Yan, and P. S. Doyle, *Soft Matter* **10**, 9721 (2014).
- [42] C. U. Murade, V. Subramaniam, C. Otto, and M. L. Bennink, *Nucl. Acids Res.* **38**, 3423 (2010).
- [43] W. Reisner, K. J. Morton, R. Riehn, Y. M. Wang, Z. Yu, M. Rosen, J. C. Sturm, S. Y. Chou, E. Frey, and R. H. Austin, *Phys. Rev. Lett.* **94**, 196101 (2005).
- [44] C.-C. Hsieh, A. Balducci, and P. S. Doyle, *Nano Lett.* **8**, 1683 (2008).
- [45] L. Nyberg, F. Persson, B. Åkerman, and F. Westerlund, *Nucl. Acids Res.* **41**, e184 (2013).
- [46] J. Z. Y. Chen, D. E. Sullivan, and X. Yuan, *Europhys. Lett.* **72**, 89 (2005).
- [47] F. Persson, P. Utko, W. Reisner, N. B. Larsen, and A. Kristensen, *Nano Lett.* **9**, 1382 (2009).
- [48] D. Gupta, J. J. Miller, A. Muralidhar, S. Mahshid, W. Reisner, and K. D. Dorfman, *ACS Macro Lett.* **4**, 759 (2015).
- [49] F. Persson and J. O. Tegenfeldt, *Chem. Soc. Rev.* **39**, 985 (2010).
- [50] N. Otsu, *IEEE Trans. Syst., Man Cybern.* **9**, 62 (1979).
- [51] P.-S. Liao, T.-S. Chen, and P.-C. Chung, *J. Inf. Sci. Eng.* **17**, 713 (2001).
- [52] T. Snijders and R. Bosker, *Multilevel Analysis* (Sage, London, 1999).
- [53] D. C. Harris, *Quantitative Chemical Analysis*, 8th ed. (Freeman, New York, 2010).