Scaling Exponents and Probability Distributions of DNA End-to-End Distance

Francesco Valle,[1] Mélanie Favre,[1] Paolo De Los Rios,[2] Angelo Rosa,[3] and Giovanni Dietler[1]

[1]*Laboratory of Physics of Living Matter, IPMC, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland*
[2]*Laboratory of Statistical Biophysics, ITP, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland*
[3]*Laboratory for Computation and Visualization in Mathematics and Mechanics, IMB, Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland*

(Received 15 March 2005; published 6 October 2005)

The scaling of the average gyration radius of polymers as a function of their length can be experimentally determined from ensemble measurements, such as light scattering, and agrees with analytical estimates. Ensemble techniques, yet, do not give access to the full probability distributions. Single molecule techniques, instead, can deliver information on both average quantities and distribution functions. Here we exploit the high resolution of atomic force microscopy over long DNA molecules adsorbed on a surface to measure the average end-to-end distance as a function of the DNA length, and its full distribution function. We find that all the scaling exponents are close to the predicted 3D values ($\nu = 0.589 \pm 0.006$ and $\delta = 2.58 \pm 0.77$). These results suggest that the adsorption process is akin to a geometric projection from 3D to 2D, known to preserve the scaling properties of fractal objects of dimension $d_f < 2$.

DOI: 10.1103/PhysRevLett.95.158105 PACS numbers: 87.14.Gg, 36.20.Ey, 82.35.Gh, 87.64.Dz

The statistical properties of polymers are the focus of a strong research effort. In the limit of very long and perfectly flexible linear polymers, the situation is rather clear and the principles were laid down some time ago [1]. Self-avoiding walk (SAW) statistics describe the properties of very long polymers in two and three dimensions, and the main results can be summarized in the scaling properties of polymers, which were theoretically as well as experimentally confirmed [1]. Concerning DNA, the situation is more complex due to the elastic properties of the double helix, its polyelectrolytic properties, and its persistence length ℓ_p (see [2,3]). Experimentally, the dynamics and statics of a purely two-dimensional linear DNA chain was investigated by Maier *et al.* [4] finding results in agreement with the theoretical predictions. Moreover, Rivetti *et al.* [5] used atomic force microscopy (AFM) to investigate statistical properties of DNA yielding information about the persistence length, the kinetics, and the mode of adsorption on a substrate. Local changes in rigidity, curvature, and/or topology of a DNA molecule induced by chemical compounds or by DNA binding proteins [6–11] have also been studied.

Here we show that AFM imaging also permits us to extract information about the three-dimensional conformation of a flexible biopolymer. We have studied long DNA molecules adsorbed on a flat surface and have determined how its end-to-end distance and the associated probability distribution scale with the contour length s over 4 orders of magnitude, from $s = 1$ nm to $s = 10\,000$ nm, a range that spans values both smaller and much larger than the persistence length l_p. This is possible only with the atomic force microscope [12], which combines a very high spatial resolution with the ability to image biomolecules on large scales. Images of long DNA molecules adsorbed onto a surface thus provide information at the single molecule level.

Polymer theory [1] suggests that the average end-to-end distance $\langle R(s)\rangle$ of a polymer of contour length s should scale as $\langle R(s)\rangle \sim s^{\nu}$. The exponent ν depends on the dimension of the system, taking values $\nu = 1$, 0.75, 0.588, and $1/2$ for $d = 1, 2, 3$, and $d \geq 4$, respectively. Our experimental results show that there are two scaling regimes. At short length scales (i.e., smaller than the persistence length ℓ_p), the DNA behaves like a rigid rod. On length scales bigger than ℓ_p, a behavior is observed with exponent $\nu = 0.589 \pm 0.006$, in agreement with the numerical estimates [13–15]. Furthermore, working with images of single DNA molecules allowed us to determine the distribution of the end-to-end distances for a given contour length; this information is usually not accessible. The form of the distribution depends on the flexibility of the polymer and on the contour length considered. Although several theoretical estimates of the distributions exist [16–18], we are not aware of any direct experimental measurements of these parameters for DNA.

Linear DNA was prepared from a solution of λ-phage DNA, 48 502 base pairs (bp) long, cleaved by restriction enzymes to give a mixture of lengths from 1503 bp up to the maximum 48 502 bp. DNA molecules were prepared in a buffer solution of 10 mM Tris-HCl, pH 7.6, 1 mM EDTA with DNA concentrations ranging from 0.5 to 2 μg/ml. The Debye screening length $1/k$ is 2 nm [19]. The substrates (freshly cleaved mica) were positively charged by exposing them to 3-aminopropyltriethoxy silane vapors for two hours at room temperature in a dry atmosphere [20]. A 10 μl drop of a DNA solution was deposited onto the substrate surface for ten minutes and then rinsed with ultrapure water. The sample was finally blown dry with

 © 2005 The American Physical Society

clean nitrogen. The DNA images were recorded by means of an AFM operated in intermittent-contact mode, in order to reduce the effect of lateral forces during scanning of the surface [21]. We checked that the sample remains stable for weeks if kept in dry atmosphere. In Fig. 1 four DNA images are depicted. From similar pictures, the contour of about 60 DNA molecules was digitized using a specially designed software [22] to track the molecule backbone and to extract the coordinates of the polymer contour. The first digitization, which may lead to noisy and nonequidistant coordinates, was subsequently smoothed using the Snake algorithm [23]. Several tests were performed with molecules of known length (DNA plasmids) to check the procedure [22].

The analysis of the end-to-end distance $R(s)$ for the DNA molecules as a function of the contour length s was done by moving a window of length s along the contour of the DNA molecule for values of s going from the minimal segment value $s_{\min} = 2$ nm to the total DNA length. In this respect, what we refer to as end-to-end distance is actually an internal end-to-end distance. The small values of s have a better statistical error, since they occur more often than the maximum value. The $R(s)$ values were then averaged over all the molecules to yield the mean end-to-end $\langle R(s)\rangle$, which is plotted in Fig. 2. Since DNA is a semiflexible polymer with a persistence length ℓ_p, one expects that $R(s)$ should scale linearly for $s \ll \ell_p$ as well as according to self-avoiding walk theory for $s \gg \ell_p$, rather than with a single power law. A two-power law function has thus to be used in order to fit the data,

$$\langle R(s)\rangle \sim \left(\frac{s}{\ell}\right)^{\nu_0}\left(1+\frac{s}{\ell}\right)^{\nu_1-\nu_0}, \qquad (1)$$

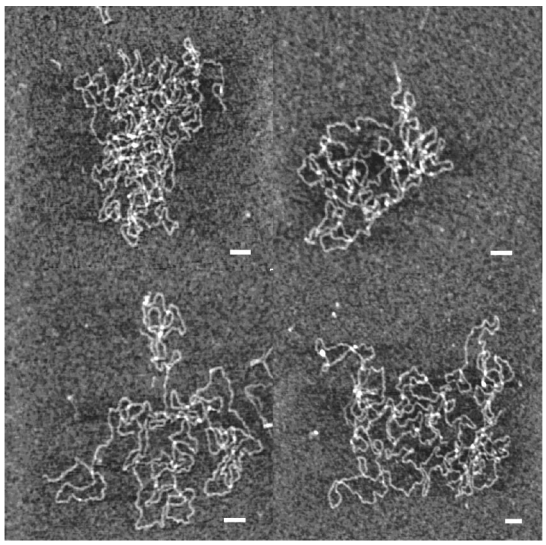

FIG. 1. (a) Intermittent-contact mode images of linear DNA molecule from an enzymatic digested lambda DNA. The original DNA has 48 502 base pairs; here we show some shorter fragments. The scale bar represents 100 nm.

where ℓ is a crossover length between the two power laws (that is likely to correspond to the persistence length ℓ_p), and ν_0 and ν_1 the scaling exponents. The fit to the data gives the following results:

$$\nu_0 = 1.030 \pm 0.017, \qquad \nu_1 = 0.589 \pm 0.006,$$
$$\ell = (44 \pm 3) \text{ nm}.$$

These results can be interpreted as follows. The value of ℓ is in good agreement with previously reported measurements of the persistence length of DNA performed by microscopy techniques [24,25]. We can therefore identify ℓ with ℓ_p, which in turn leads to a simple and intuitive interpretation of the two scaling regimes. For $s < \ell_p$ DNA behaves as a rigid rod, and the end-to-end distance scales linearly with the contour length of the polymer ($\nu_0 = 1.030$). For $s > \ell_p$ the scaling exponent ($\nu_1 = 0.589$) agrees with the numerical estimates of the exponents [13] and with experimental values for synthetic polymers by scattering methods [26,27]. Thus, the adsorbed DNA behaves as a three-dimensional polymer, which does not undergo any two-dimensional equilibration upon adsorption. This is also evident from the images of Fig. 1, where the DNA molecules show a large number of crossings. The Euclidean dimension of the surface ($d = 2$) onto which the molecule is projected is larger than the fractal dimension

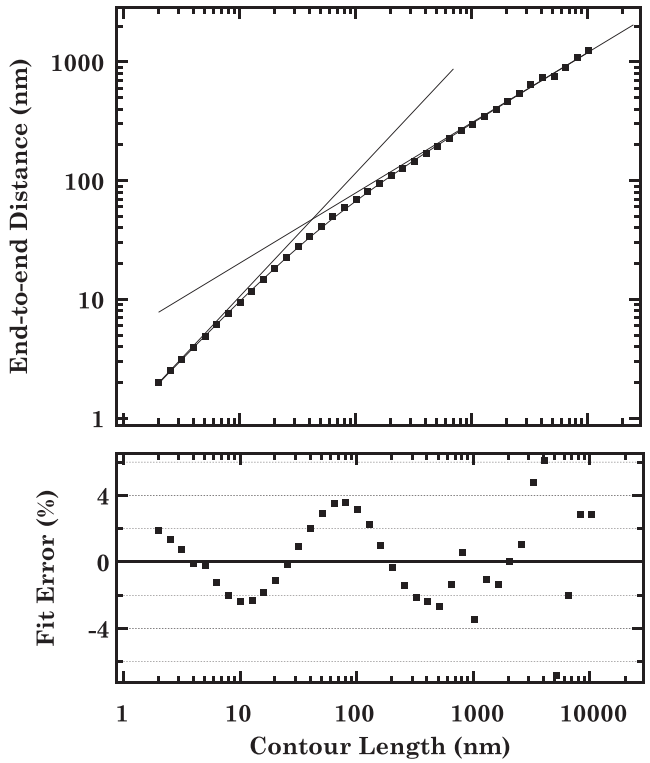

FIG. 2. Representation in double logarithmic scale of the end-to-end distance versus the contour length. Two distinct power laws are evident; the data have been fitted using Eq. (1).

($d_f = 1/\nu_1 \approx 1.7$) of the DNA, explaining the conservation of the 3D exponent upon adsorption on a surface [28]. We conclude that the λ-DNA images represent some form of a two-dimensional projection of their three-dimensional bulk conformation.

Furthermore, we can measure the probability distribution of the end-to-end distance for a determined contour length s, as it can be extracted from the microscopic conformation of each molecule. Two regimes are possible in respect to the persistence length: $s \lesssim \ell_p$, and $s > \ell_p$. For $s \lesssim \ell_p$, one expects an almost delta-function distribution, since the polymer chains do not bend over these length scales. For $s \approx \ell_p$, the distribution is still narrow but not as peaked as for the previous case, since the molecules starts to enter the semiflexible regime; an example is given in Fig. 3 for a contour length of $s_0 = 75$ nm ($\approx 2\ell_p$). This distribution was fitted with Winkler's equation [18],

$$f(s) = a \frac{s e^{-s_0/8\ell_p[1-(s/s_0)^2]}}{[1-(s/s_0)^2][2-(s/s_0)^2]^2}. \quad (2)$$

The fits shown in Fig. 3 give $\ell_p = 46.6$ nm and $s_0 = 71.4$ nm, in good agreement with the literature value for the persistence length and with the nominal total length of 75 nm. For contour lengths $s_0 \gg \ell_p$, the distribution changes dramatically and was determined for 34 different contour lengths, ranging from 200 to 4600 nm. For longer contour lengths, the distributions are difficult to determine because of the reduced number of samples available, and only the average end-to-end distance can be given as in Fig. 1. In Fig. 4 we show the distributions for $s_0 = 548$ nm ($\approx 12\ell_p$) and $s_0 = 748$ nm ($\approx 17\ell_p$). The histograms of Fig. 4 have been rescaled with $s_0^{\nu_1}$, and the way they collapse onto each other is a good *a posteriori* confirmation of the power law previously determined (Fig. 2). Starting from a SAW model [1,29], the distribution probability of the end-to-end distance as a function of the contour length s becomes

$$f(s) = a s^{d-1+\sigma} e^{-b s^{\delta}}. \quad (3)$$

The two exponents characterizing the distributions were determined from fits of the histograms for the 34 different contour lengths between 200 and 4600 nm; their averages are

$$d - 1 + \sigma = 1.33 \pm 0.22, \qquad \delta = 2.58 \pm 0.76.$$

These values have to be compared to those characterizing the corresponding two- or three-dimensional distributions of the end-to-end distances for a SAW (for $d = 2$, $\sigma = 0.44$, $\delta = 4$, and for $d = 3$, $\sigma = 0.33$, $\delta = 2.43$). The agreement between ν and δ is important because of the scaling relation $\nu = 1 - 1/\delta$ [30], which is verified within the error bar. Moreover, both ν and δ agree with their three-dimensional values. We speculate therefore that the adsorption process preserves the three-dimensional exponents. Since a SAW in 3D is a fractal object with fractal dimension $d_f = 1/\nu < 2$, its mathematical projection would preserve d_f and hence ν. We can thus, as a first approximation, think of the adsorption process as a projection from 3D to 2D. Yet, after a projection the small s behavior of $f(s)$ should be linear, which is not the case (see Fig. 3).

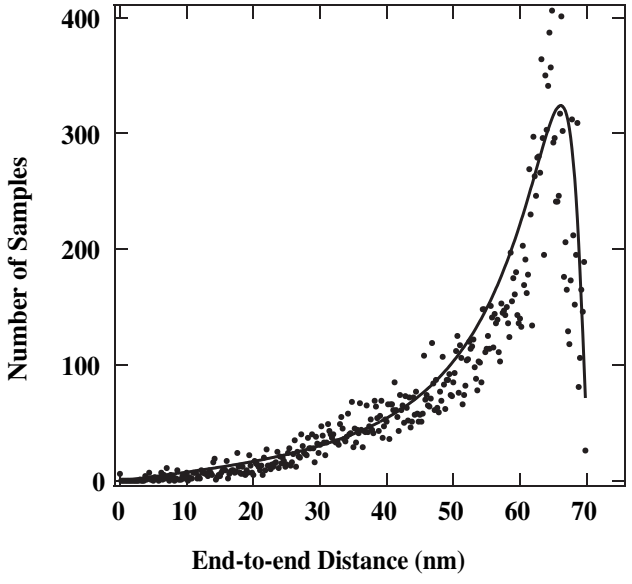

FIG. 3. Histogram representing the distribution of the end-to-end distance for a contour length $s_0 = 75$ nm. The continuous line is a fit to Eq. (2).

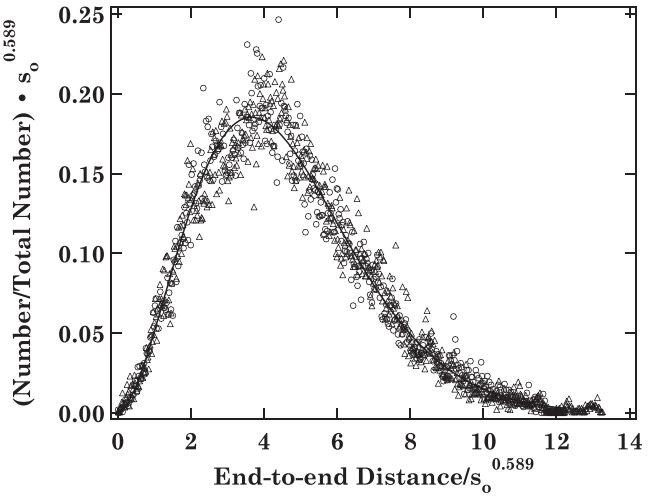

FIG. 4. Histogram representing the distribution of the end-to-end distance for two different contour length (548 nm circles, 748 nm triangles) and how they collapse onto each other once they have been rescaled using the scaling exponent $\nu_1 = 0.589$ measured in the present experiments. The solid line represents the fit of the 548 nm distribution using Eq. (3).

Our results are a confirmation that for long DNA molecules ($s_0 \sim 5$–100 persistence lengths) the end-to-end distribution matches a pure SAW distribution for large end-to-end distances. Under the conditions used in our preparation, the DNA adsorption is strong and DNA is quenched on the surface. No equilibration is taking place in two dimensions. The irreversibility of the adsorption was verified by imaging under liquid linear DNA molecules during 30 min; no changes of the conformation on short or large length scales of the molecules could be detected [31]. The problem of ''trapping'' or ''equilibration'' of DNA onto different surfaces has already been studied by Rivetti *et al.* [5], using short DNA fragments that were then analyzed with the wormlike chain (WLC) model. In their work they saw a large deviation from the expected three-dimensional WLC even for fragments of 6 kbp. In our work the 3D power law fits the experimental data up to $\approx$30 kbp or $230\ell_p$, with no *a priori* assumption about the dimensionality of the final conformation. Compared to the optical images of fluorescently marked polymers [4,32], the high resolution of AFM allows one to perform the analysis on segments as short as a few nanometers up to 10 000 nm over four decades of lengths permitting the observation of the transition from stiff to SAW polymer behavior. Short DNA molecules ($s_0 < 5$ persistence lengths) behave instead as semiflexible polymers.

Further experiments should be performed by varying the salt concentration in order to determine the contribution of the electrostatic persistence length to the total persistence length (see [3]). However, the use of high salt concentrations will also change the deposition process [5] and the adsorption might not be irreversible. A certain degree of two-dimensional equilibration might take place, influencing the persistence length [5,33]. Moreover, the theory of polyelectrolytes with stiffness should be used if salt and other parameters could be varied [2,3]. In these theories a more complex behavior is predicted: a crossover from stiff rod behavior to SAW through a region of Gaussian behavior is expected. At present our data do not allow us to make such a detailed test.

We thank A. Stasiak, J. Prost, L. Peliti, R. Metzler, and R. Winkler for suggestions, comments, and important discussions on the interpretation of the data. G. D. thanks the Swiss National Science Foundation for support through Grant No. 2100-063746.00.

[1] P. G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, 1979).
[2] K. Ghosh, Gustavo A. Carri, and M. Muthukumara, J. Chem. Phys. **115**, 4367 (2001).
[3] C. Holm, J. F. Joanny, K. Kremer, R. R. Netz, P. Reineker, C. Seidel, T. A. Vilgis, and R. G. Winkler, Adv. Polym. Sci. **166**, 67 (2004).
[4] B. Maier and J. O. Rädler, Phys. Rev. Lett. **82**, 1911 (1999).
[5] C. Rivetti, M. Guthold, and C. Bustamante, J. Mol. Biol. **264**, 919 (1996).
[6] Y. L. Lyubchenko and L. S. Shlyakhtenko, Proc. Natl. Acad. Sci. U.S.A. **94**, 496 (1997).
[7] C. Rivetti, C. Walker, and C. Bustamante, J. Mol. Biol. **280**, 41 (1998).
[8] Y. K. Jiao, D. I. Cherny, G. Heim, T. M. Jovin, and T. E. Shaffer, J. Mol. Biol. **314**, 233 (2001).
[9] J. Adamcik, V. Viglasky, F. Valle, M. Antalik, D. Podhradsky, and G. Dietler, Electrophoresis **23**, 3300 (2002).
[10] A. Scipioni, C. Anselmi, G. Zuccheri, B. Samori, and P. De Santis, Biophys. J. **83**, 2408 (2002).
[11] V. Viglasky, F. Valle, J. Adamcik, I. Joab, D. Podhradsky, and G. Dietler, Electrophoresis **24**, 1703 (2003).
[12] G. Binnig, C. F. Quate, and C. Gerber, Phys. Rev. Lett. **56**, 930 (1986).
[13] J. C. Le Guillou and J. Zinn-Justin, Phys. Rev. Lett. **39**, 95 (1977).
[14] S. Caracciolo, M. S. Causo, and A. Pelissetto, Phys. Rev. E **57**, R1215 (1998).
[15] H.-P. Hsu, W. Nadler, and P. Grassberger, Macromolecules **37**, 4658 (2004).
[16] J. Wilhelm and E. Frey, Phys. Rev. Lett. **77**, 2581 (1996).
[17] D. Thirumalai and B.-Y. Ha, cond-mat/9705200.
[18] R. G. Winkler, J. Chem. Phys. **118**, 2919 (2003).
[19] A. Vologodskii, Macromolecules **27**, 5623 (1994).
[20] Y. L. Lyubchenko, A. A. Gall, L. S. Shlyakhtenko, R. E. Harrington, B. L. Jacobs, P. I. Oden, and S. M. Lindsay, J. Biomol. Struct. Dyn. **10**, 589 (1992).
[21] R. Garcia and R. Perez, Surf. Sci. Rep. **47**, 197 (2002).
[22] J. Marek, E. Demjénová, Z. Tomori, J. Janáček, F. Valle, M. Favre, and G. Dietler, Cytometry **63A**, 87 (2005).
[23] Y. Y. Wong, P. C. Yuen, and C. S. Tong, Pattern Recognition **31**, 1669 (1998).
[24] J. Bednar, P. Furrer, V. Katritch, A. Z. Stasiak, J. Dubochet, and A. Stasiak, J. Mol. Biol. **254**, 579 (1995).
[25] T. Berge, N. S. Jenkins, R. B. Hopkirk, M. J. Waring, J. M. Edwardson, and R. M. Henderson, Nucleic Acids Res. **30**, 2980 (2002).
[26] M. Daoud, J. P. Cotton, B. Farnoux, G. Jannik, G. Sarma, H. Benoit, R. Duplessix, C. Picot, and P. G. de Gennes, Macromolecules **8**, 804 (1975).
[27] Y. Miyaki, Y. Einaga, and H. Fujita, Macromolecules **11**, 1180 (1978).
[28] B. B. Mandelbrot, *The Fractal Geometry of Nature* (W. H. Freeman and Company, Oxford, 1982).
[29] A. Y. Grosberg and A. R. Khokhlov, *Statistical Physics of Macromolecules* (American Institute of Physics, Woodbury, 1994).
[30] M. E. Fisher, J. Chem. Phys. **44**, 616 (1966).
[31] See EPAPS Document No. E-PRLTAO-95-004543 for AFM images in liquid of linear DNA at different times after deposition. This document can be reached via a direct link in the online article's HTML reference section or via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html).
[32] B. Maier and J. O. Rädler, Macromolecules **33**, 7185 (2000).
[33] M. Joanicot and B. Revet, Biopolymers **26**, 315 (1987).