Principal-components analysis of shape fluctuations of single DNA molecules

Adam E. Cohen* and W. E. Moerner

Department of Chemistry, Stanford University, Stanford, CA 94305

Edited by Robert J. Silbey, Massachusetts Institute of Technology, Cambridge, MA, and approved March 15, 2007 (received for review October 25, 2006)

Thermal fluctuations agitate molecules in solution over a broad range of times and distances. By passively watching the shape fluctuations of a thermally driven biomolecule, one can infer properties of the underlying interactions that determine the motion. We applied this concept to single molecules of fluorescently labeled λ-DNA, a key model system for polymer physics. In contrast to most other single-molecule DNA experiments, we examined the unstretched, equilibrium state of DNA by using an anti-Brownian electrokinetic trap to confine the center of mass of the DNA without perturbing its internal dynamics. We analyze the long-wavelength conformational normal modes, calculate their spring constants, and measure linear and nonlinear couplings between modes. The modes show strong signs of nonlinear hydrodynamics, a feature of the underlying equations of polymer dynamics that has not previously been reported and is neglected in the widely used Rouse and Zimm approximations.

DNA dynamics | single molecule | hydrodynamic interactions | polymer physics

The simplest model of a linear polymer is a chain of beads joined by springs. In the Rouse model (1), each bead is a Brownian diffuser with the same drag and diffusion coefficients it would have in the absence of other beads. This model neglects the fact that each bead is subject to the time-varying flow fields produced by the other diffusing beads. This hydrodynamic interaction (HI) renders the underlying dynamics nonlinear. The Zimm model (2) includes hydrodynamics but restores linearity through a mean-field approximation: each bead is made to interact with the average conformation of its neighbors. Subsequent work has applied sophisticated mathematical techniques to calculate corrections resulting from fluctuating hydrodynamics (3–5), but the overall significance of internal HIs to polymer dynamics remains unresolved (6).

The HI is the dominant long-range force for biomolecules in aqueous buffers: it couples motion of one part of a molecule to motion of possibly remote parts of the same molecule. Thus, understanding HIs is crucial to understanding the rates of molecular events that involve large-scale conformational change such as folding of proteins and RNA, packaging of DNA, motion of molecular motors, and motion of DNA-binding proteins.

Long before polymers were studied at the single-molecule level, many clever experiments applied light scattering and neutron scattering as indirect probes of polymer dynamics (7, 8). However, these experiments were (*i*) limited to probing only the lowest one or two internal relaxations and (*ii*) only yielded second-order ensemble-averaged correlation functions without measuring the entire distribution of underlying states. This second property of scattering techniques makes them insensitive to deviations from the linearized Zimm theory.

Higgins and Benoit (9) and Quake *et al*. (10) used laser tweezers to study the dynamics of partially extended DNA in solution yet failed to find deviations from the Zimm theory. This negative result is not surprising, because extending a molecule weakens the internal HIs, rendering the nonlinearity harder to detect. Subsequent experiments have probed DNA under a wide range of twists and extensions (for reviews see refs. 11 and 12). In contrast, the equilibrium (i.e., unstretched) dynamics of single DNA molecules have received relatively little attention (13–15) because of the challenge of following a molecule as it diffuses away from the field of view.

Experimental Method

Our experimental method consists of (*i*) acquiring a large number of images showing the free-solution shape fluctuations of single fluorescently labeled molecules of DNA [see supporting information (SI) Movie 1]; (*ii*) identifying the conformational normal modes (analogous to the wavefunctions of an electron or the normal modes of a drum); and (*iii*) determining whether the dynamics in these modes can be fit to a linear model, as required by the Zimm theory. The observed deviations from Zimm theory verify that HI effects are present in DNA and that HI may play an important role in the dynamics of more complex biomolecular systems.

To study a single molecule in equilibrium, one would like to eliminate the motion of the center of mass without affecting internal motions. Active feedback provides a means to accomplish this elimination. One tracks the Brownian motion of a molecule and then imposes a body force that counteracts this motion. A variety of schemes have been proposed (16–18) and implemented (19–22), differing in the method of tracking and the source of the restoring force. Here we apply our anti-Brownian electrokinetic (ABEL) trap that uses video tracking and electrokinetic feedback and is capable of trapping objects as small as individual proteins in solution (23). The hardware and software have been described in detail (19, 20, 24, 25). In brief, the molecule to be trapped is confined to a thin fluid layer [$\approx$1 μm thick, slightly larger than the radius of gyration R_g for λ-DNA ($\approx$700 nm)] in a glass microfluidic cell. An automated video tracking system follows the Brownian displacements (via fluorescence microscopy) and, for every frame, applies feedback voltages to the cell to induce an electrokinetic drift that approximately cancels the Brownian motion. Motions at frequencies higher than the update rate are not affected.

Molecules of double-stranded λ-DNA fluorescently labeled with YOYO-1 (Molecular Probes) were held in the ABEL trap, and a two-dimensional projection of their conformational motions was recorded with video microscopy at a time resolution of 4.5 ms per frame. Twenty-one separate molecules were trapped, each for between 9 and 18 s (yielding between 2,000 and 4,000 images per molecule; total data set: 58,421 frames). Several frames showing the shape fluctuations that are the focus of this article are shown in Fig.

Author contributions: A.E.C. and W.E.M. designed research; A.E.C. performed research; A.E.C. contributed new reagents/analytic tools; A.E.C. analyzed data; and A.E.C. and W.E.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Abbreviations: HI, hydrodynamic interaction; ABEL, anti-Brownian electrokinetic; PC, principal component; PCA, principal-components analysis.

*To whom correspondence should be sent at the present address: Department of Chemistry, 12 Oxford Street, Harvard University, Cambridge, MA 02138. E-mail: acohen@post.harvard.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0610396104/DC1.

© 2007 by The National Academy of Sciences of the USA

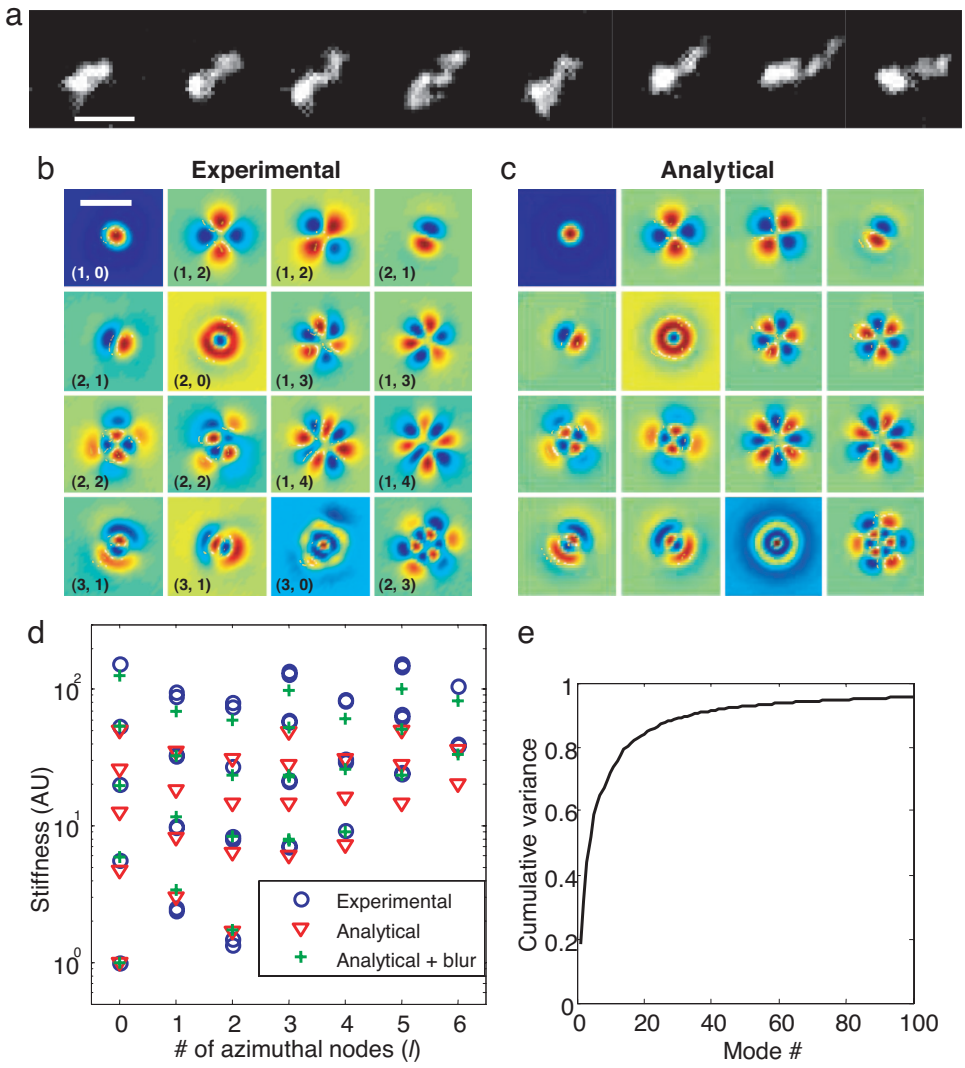

Fig. 1. Shape fluctuations of λ-DNA. (*a*) Series of images showing equilibrium conformational fluctuations of fluorescently labeled λ-DNA. Here every eighth image from the full movie provided in SI Movie 1 is shown. (Scale bar, 2 μm.) (*b*) Experimentally determined first 16 PCs of conformational fluctuations, ordered by their associated eigenvalues. Symmetry forbids the existence of (1, 1) modes (analogous to the 2p hydrogen wavefunctions), because amplitude in these modes leads to a displacement of the center of mass. The color scale maps the most positive and most negative excursions to red and blue colors, respectively, and the color corresponding to zero at the edges of the PCs varies from panel to panel. (Scale bar, 2 μm.) (*c*) PCs predicted by a random-walk model. Because of the rotational symmetry, experimental and theoretical eigenfunctions are often aligned along different axes. (*d*) Stiffness of 45 of the low-energy modes. All modes with $l \neq 0$ are 2-fold degenerate. A random-walk model, taking into account finite imaging resolution, yields a similar spectrum of eigenvalues. AU, arbitrary units. (*e*) Fraction of the total variance of the data accounted for by the first p modes. The first 34 modes account for 90% of the variance.

1*a*, with the full video given in SI Movie 1 along with details of data analysis and calculations. The molecules were labeled with a uniform density of fluorophores, so the fluorescence intensity at each point in an image was proportional to the density of DNA averaged over the point-spread function of the microscope at a corresponding point in the sample.

An important question is whether the trapping feedback fields affect the conformation or dynamics of the DNA. To test for such interactions experimentally, we calculated the correlation between the applied voltage and the measured shape of the molecule as follows. Let $I(\mathbf{r}, t)$ be the intensity distribution (after removing center-of-mass motion). Let $\langle I(\mathbf{r}, t)\rangle$ be the time-average distribution and $\delta I(\mathbf{r}, t) \equiv I(\mathbf{r}, t) - \langle I(\mathbf{r}, t)\rangle$ be the instantaneous deviation from this average. The feedback voltages are V_x and V_y. We found that $\langle V_x(t_2)\delta I(\mathbf{r}, t_1)\rangle \approx 0$ and $\langle V_y(t_2)\delta I(\mathbf{r}, t_1)\rangle \approx 0$ to within the experimental uncertainty for all t_2, t_1, and $\mathbf{r}$. Thus, coupling between the feedback voltage and the conformation is small enough to be neglected here. This finding is consistent with the theoretical argument that electrokinetic force and drag act uniformly along the DNA backbone (26) and the experimental observation that the free-solution mobility of DNA is independent of contour length or conformation (27). We cannot rule out the possibility that higher-order effects modify the conformation to a small extent, but our current measurements are not sensitive to such perturbations. Finite element simulations show that the electric field is homogeneous to within <0.1% over the size of the DNA molecule (25).

Another possible source of bias in the data is the finite z thickness of the trapping region, which is only slightly larger than the radius of gyration of λ-DNA (1 μm vs. 700 nm). This confinement is expected to have little effect on the distribution of states; excluded volume interactions are very weak in λ-DNA, so a modest confinement of the random walk in one dimension does not significantly affect the random walks in the perpendicular dimensions. This assumption is confirmed by the analysis below, which shows

that the DNA shape is well described by a pure random walk. Of greater concern is the impact of the walls of the trap on the dynamics. The "no-slip" condition on the walls leads to a screening of HI on a length scale of ≈ 1 μm. Thus, the HIs described below will be slightly stronger in a truly three-dimensional polymer. The diffusion coefficient of the center of mass of λ-DNA in the ABEL trap is 24% lower than in bulk (28). Internal HIs propagate over shorter distances than does the disturbance generated by diffusion of the center of mass. Thus, the perturbation to internal HI effects is expected to be $<24\%$.

Principal-Components Analysis

We now analyze the shapes of λ-DNA, first considering only the distribution of shapes and then the dynamics (i.e., how one shape becomes another). How can one describe the variability among the observed shapes without selecting an arbitrary descriptive statistic as the quantity of interest? Principal-components analysis (PCA) provides a systematic expansion procedure for coarse-graining over atomic degrees of freedom while preserving the large-distance dynamics that are relevant for many functions. Furthermore, PCA is unbiased in the sense that it uses the data to determine the characteristic motions without requiring the experimenter to specify a model of the underlying process.

The process of PCA consists of (*i*) obtaining the covariance matrix for some randomly fluctuating N-dimensional quantity and (*ii*) calculating the eigenvectors of this covariance matrix. The eigenvectors with the largest eigenvalues are called the principal components (PCs). The PCs are an efficient basis in that if the fluctuations of the system are Gaussian, then projection of the dynamics onto the first m eigenvectors (ranked by descending eigenvalue) accounts for a larger fraction of the total variance than would a projection onto any other basis with m elements.

There are two approaches to PCA on a material system, which, in analogy to fluid dynamics, we call the Lagrangian and Eulerian approaches. In Lagrangian fluid dynamics, one follows the trajectories of distinct fluid elements (imagine tracer particles) as a function of time. In Eulerian fluid dynamics, one calculates fluid properties (e.g., pressure, velocity, shear) as a function of position in some fixed reference frame. The Eulerian approach leads to the well known Navier–Stokes equations and is generally preferred because of its simplicity. The Eulerian approach takes advantage of a symmetry of the fluid (i.e., that all elements of the fluid are identical), so there is no need to follow a particular fluid element (29).

A similar situation prevails when one wishes to perform PCA on the fluctuations of a material body. There are two ways one can construct the covariance matrix. One can follow the trajectory $x_i(t)$ of mass element i and calculate covariance matrices such as $C(i, j, \tau) = \langle x_i(t)x_j(t + \tau)\rangle$. This approach is necessary when considering a multicomponent system such as a protein, in which two elements i and j may not be interchangeable (e.g., they have different chemical properties). We call this the Lagrangian approach. For a homogeneous system, however, one can adopt a reduced description, considering only the density $\rho(x)$. Then, one can calculate covariance matrices such as $C(x_1, x_2, \tau) = \langle \rho(x_1, t)\rho(x_2, t + \tau)\rangle$. We call this the Eulerian approach.

Starting with the pioneering work of Karplus and Kushick (30), PCA has been widely applied to analyze molecular dynamics trajectories, but always in the Lagrangian perspective. Such an approach is productive, because the secondary structure of proteins is fairly rigid, so many motions involve entire domains. For a review of PCA of proteins, see refs. 31 and 32. The standard Rouse and Zimm models of polymer dynamics are also developed in the Lagrangian language (following every mass element). Indeed, performing PCA on a random walk in the Lagrangian perspective yields precisely the Rouse polymer modes.

For determining many dynamical properties of polymers, only the density distribution is needed. The present experiments are not sensitive to the underlying sequence of base pairs, so we approximate the DNA as a homogeneous polymer. Furthermore, we do not know which piece of the polymer contributes to each piece of the image. Thus, it is most appropriate to perform PCA in the Eulerian perspective.

We performed PCA (33) on 58,421 video images of DNA to identify spatially separated parts of the DNA that fluctuate in synchrony (see *SI Appendix*). Fig. 1*b* shows the first 16 PCs for λ-DNA. Starting at the upper left, the dominant PC is a radial breathing mode, followed by a pair of degenerate modes in which the molecule stretches along one axis and contracts along an orthogonal axis, followed by more complex deformations. Each PC is indexed by (n, l), where n is the number of radial nodes, and l is the number of azimuthal nodes. PCA implies a picture of the molecule as a gel-like solid with a spectrum of long-wavelength collective motions.

Each PC has associated with it an eigenvalue, λ_p, that is equal to the fraction of the variance of the entire data set that falls along the PC. The equipartition theorem implies that $\frac{1}{2}k_p\lambda_p = \frac{1}{2}k_BT$, where k_B is the Boltzmann constant, and T is the temperature, from which one can extract the stiffness, k_p, of mode p. The persistence length is much less than our optical resolution, so the stiffness is entirely caused by the entropic cost of deforming the molecule. The stiffnesses of the first 45 PCs (shown in Fig. 1*d*) follow an unexplained semiregular pattern. The PCs and the spectrum of eigenvalues appear qualitatively similar to atomic wavefunctions and energy levels familiar from quantum mechanics, but the details of the shape and the underlying equations are completely different.

Use of the PCs achieves a large reduction in the amount of data required to describe the conformational fluctuations compared with the raw images. Although each image contains 1,024 pixels, 90% of the variance in the data set is contained in the first 34 PCs (Fig. 1*e*). The remaining variance is mostly caused by measurement noise. By working in the PC basis, we suggest that numerical simulations of polymer dynamics could be rendered more efficient. Rather than simulating the trajectory of each mass element, one could simulate the dynamics of the PCs. This approach would only be beneficial for studying near-equilibrium fluctuations; otherwise the PCs cease to be an efficient basis.

PCs of a Random Walk

Here we develop a semianalytical description of the observed PCs. As a minimal model we assume that each image recorded by the camera shows a pure two-dimensional random walk. In this section we derive the PCs of a d-dimensional random walk with a fixed center of mass. Although these results are presented in the context of polymer physics, a coarse-grained description of the shapes of random walks may prove useful in other disciplines as well.

Consider a one-dimensional random walk of N Gaussian steps joining $n + 1$ mass elements, each step of variance a^2. We will work in the limit of large N and small a, keeping $Na^2 = 1$, and keeping the center of mass of the walk fixed at the origin. We think of the polymer as a density distribution $\rho(x)$. In the Eulerian perspective, the covariance matrix depends on fluctuations about the mean density distribution, $\langle\rho(x)\rangle$, so we start by discussing this distribution.

The mean distribution of density about the center of mass of a random walk is not Gaussian (34). Each segment obeys a Gaussian density distribution about the center of mass, but the width of this distribution varies along the chain: the ends wander further than the middle. The total density distribution is the sum of many Gaussian distributions of distinct widths and, thus, is not a Gaussian distribution. Yamakawa (34) showed that the total density distribution is

$$\langle \rho(x) \rangle = \sum_{\varepsilon=0}^{N} \sqrt{\frac{3}{2\pi F(\varepsilon)}} \exp\left[-\frac{3x^2}{2F(\varepsilon)} \right] \quad [1]$$

with

$$F(\varepsilon) \equiv 3\frac{\varepsilon^2}{N^2} - 3\frac{\varepsilon}{N} + 1. \quad [2]$$

This function looks qualitatively like a Gaussian distribution near $x = 0$ but has fat tails relative to a Gaussian distribution. For a multidimensional Gaussian random walk, the displacements along orthogonal axes are statistically independent, so the total probability density is the product of the one-dimensional probability densities along each of the axes. We previously showed that Eq. **1** agrees well with the observed density distribution of λ-DNA, whereas a Gaussian distribution does not (28).

Now we calculate the density–density covariance of an ensemble of one-dimensional random walks in the Eulerian perspective. The covariance matrix is

$$C(x_1, x_2) = \langle \rho(x_1)\rho(x_2) \rangle - \langle \rho(x_1) \rangle \langle \rho(x_2) \rangle \quad [3]$$

The second term on the right-hand side of Eq. **3** is obtained from Eq. **1**. The challenging task is to calculate the first term, $\langle \rho(x_1)\rho(x_2) \rangle$.

If we assume that each piece of the random walk contributes a point-like density, the density at position x_1 is

$$\rho(x_1) = \sum_{\alpha=0}^{N} \delta(x_1 - x_\alpha),$$

where α is an index of the mass elements. The product of the densities at two positions is

$$\rho(x_1)\rho(x_2) = \sum_{\alpha,\beta=0}^{N} \delta(x_1 - x_\alpha)\delta(x_2 - x_\beta).$$

We take the Fourier transforms of the above equation with respect to x_1 and x_2 to obtain

$$\rho(k_1)\rho(k_2) = \sum_{\alpha,\beta=0}^{N} \exp[i(k_1 x_\alpha + k_2 x_\beta)].$$

The position of mass element α is then expanded in Rouse modes as

$$x_\alpha = \sum_{\nu=1}^{\infty} c_\nu \cos\left(\frac{\pi\nu\alpha}{N} \right)$$

and similarly for x_β. Leaving out the $\nu = 0$ term guarantees that the center of mass remains fixed at $x = 0$ (i.e., $\Sigma_\alpha x_\alpha = 0$ for all $\{c_\nu\}$). Thus, we have

$$\rho(k_1)\rho(k_2) = \sum_{\alpha,\beta=0}^{N} \exp\left[i \sum_{\nu=1}^{\infty} c_\nu \left(k_1 \cos\left(\frac{\pi\nu\alpha}{N} \right) + k_2 \cos\left(\frac{\pi\nu\beta}{N} \right) \right) \right].$$

The coefficients c_ν are Gaussian distributed and statistically independent, with $\langle c_\nu c_\mu \rangle = 1/(\pi\nu)^2 \delta_{\nu\mu}$. After taking the average over all conformations (i.e., all $\{c_\nu\}$), we obtain

$$\langle \rho(k_1)\rho(k_2) \rangle = \sum_{\alpha,\beta=0}^{N} \exp\left[-\sum_{\nu=1}^{\infty} \frac{1}{(\pi\nu)^2} \left(k_1 \cos\left(\frac{\pi\nu\alpha}{N} \right) + k_2 \cos\left(\frac{\pi\nu\beta}{N} \right) \right)^2 \right]. \quad [4]$$

The next challenge is to evaluate the sum

$$S = \sum_{\nu=1}^{\infty} \frac{1}{(\pi\nu)^2} \left(k_1 \cos\left(\frac{\pi\nu\alpha}{N} \right) + k_2 \cos\left(\frac{\pi\nu\beta}{N} \right) \right)^2. \quad [5]$$

A similar sum appears in the theory of scattering from polymer solutions (35), but in that case the sum depends on only one k variable. Eq. **4** becomes the formula for the static structure factor when $k_2 = -k_1$. This difference arises because scattering experiments always probe spatially averaged quantities. Direct imaging experiments, on the other hand, allow us to compare fluctuations at pairs of distinct points, x_1 and x_2, or alternatively at distinct k vectors, k_1 and k_2.

When $k_1 = -k_2$, the sum in Eq. **5** evaluates to $S = k_1^2|\alpha - \beta|/2$ (35). In the more general case where k_1 and k_2 vary independently, the sum evaluates to

$$S = \frac{1}{12}[2k_1^2 F(\alpha) + 2k_2^2 F(\beta) + k_1 k_2 G(\alpha, \beta)], \quad [6]$$

where $F(\cdot)$ is as defined in Eq. **2**, and

$$G(\alpha, \beta) \equiv 3\left(\frac{\alpha}{N} + \frac{\beta}{N} - 1 \right)^2 + 3\left(\left| \frac{\alpha}{N} - \frac{\beta}{N} \right| - 1 \right)^2 - 2.$$

Substituting Eq. **6** into Eq. **4** and taking the inverse Fourier transform yields

$$\langle \rho(x_1)\rho(x_2) \rangle = \sum_{\alpha,\beta=0}^{N} \frac{6}{\pi\sqrt{16F(\alpha)F(\beta) - G(\alpha,\beta)^2}} \cdot \exp\left[\frac{-24F(\beta)x_1^2 - 24F(\alpha)x_2^2 + 12G(\alpha,\beta)x_1 x_2}{16F(\alpha)F(\beta) - G(\alpha,\beta)^2} \right]. \quad [7]$$

Eq. **7** is the fundamental result of this section. It gives the joint probability of two pieces of the random walk being at positions x_1 and x_2, with the center of mass at the origin. Unfortunately, the sum over α and β must be evaluated numerically. In d dimensions, the fluctuations along orthogonal axes are statistically independent, so the total probability density is the product of the one-dimensional probability densities along each of the axes.

We evaluated the two-dimensional version of Eq. **7**, where $\mathbf{r}_1$ and $\mathbf{r}_2$ each were selected from a 32×32 grid, to obtain the density–density covariance matrix for two-dimensional random walks. This covariance matrix was diagonalized numerically, and Fig. 1 *c–e* shows a comparison between the experimental and theoretical PCs and eigenvalues. To improve the correspondence between the theoretical and experimental eigenvalues we added to the model the effect of finite imaging resolution. The analytical PCs were convolved with a Gaussian distribution, and the analytical eigenvalues were then multiplied by the ensuing decrease in mean-square amplitude of the corresponding eigenvectors. This procedure led to good correspondence between the theoretical and experimental eigenvalues. At present it is not clear whether the remaining differences arise from noise in the data or from a physical process

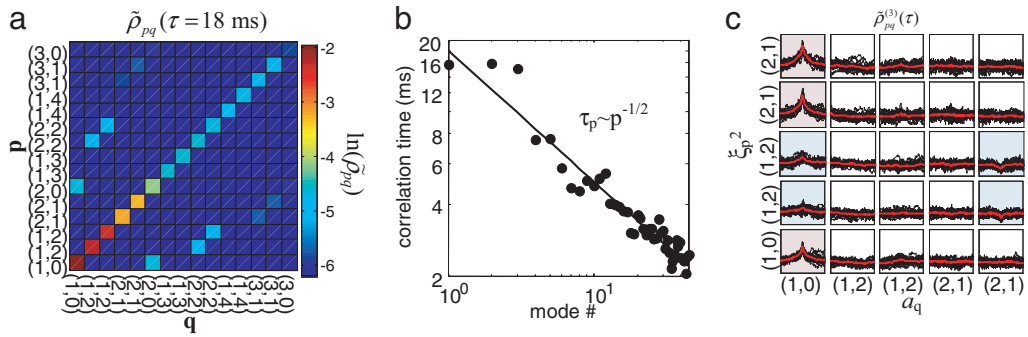

Fig. 2. Linear and nonlinear dynamics of the PCs. (*a*) Covariance matrix of amplitudes in the first 15 PCs at $\tau = 18$ ms. The off-diagonal elements indicate that the PCs are not eigenstates of the time-evolution operator. The only significant transitions conserve *l* (i.e., are vertical on Fig. 1*d*) and change *n* by ± 1. (*b*) Power-law scaling of the relaxation times in the first 45 PCs. The relaxation time was extracted from the short-time autocorrelation of the mode amplitudes. (*c*) Nonlinear couplings in the first 5 PCs. Each correlation function $\tilde{\rho}^{(3)}_{pq}(\tau)$ in the 5 × 5 array probes the effect of amplitude in mode a_q on the magnitude of the thermal fluctuations ξ_p^2. Strong nonlinear interactions are shaded pink, and weak ones are in blue. Each black line is the calculation for a single molecule of λ-DNA, and the red lines are the ensemble average. Each box has a time axis of $\tau = (-450, 450$ ms) and a vertical axis $\tilde{\rho}^{(3)}_{pq} = (-0.1, 0.3)$. A table of values of $\tilde{\rho}^{(3)}_{pq}(0)$ is provided in *SI Appendix*.

in the DNA (such as excluded volume) that has been left out of the model.

Dynamics of the PCs

The PCA is only sensitive to shape deformations that occur within the same video frame, so it provides no information on the dynamics (i.e., how a fluctuation at one time impacts the fluctuations at a later time). If the video images of the DNA were randomly reordered, the results of PCA would not change. To obtain a more detailed picture of the dynamics, we decomposed each image into the basis of PCs and examined the time dependence of the mode amplitudes, $a_p(t)$.

The amplitude in each PC ebbs and flows stochastically as thermal fluctuations and viscous damping add and remove energy. The most general second-order quantity characterizing these dynamics is the time-dependent covariance matrix, $\tilde{\rho}_{pq}(\tau) = \langle a_p(t + \tau)a_q(t)\rangle$, where the indices p and q may be truncated (16) at a small value (e.g., 15 in Fig. 2*a*). If the PCs were also the eigenstates of time evolution, then $\tilde{\rho}_{pq}(\tau)$ would be diagonal for all τ. However, Fig. 2*a* shows that off-diagonal terms arise. The only significant off-diagonal elements connect mode $p = (n \pm 1, l)$ to mode $q = (n, l)$, suggesting that conservation of azimuthal mode number is a selection rule for DNA conformational transitions. That is, only vertical transitions are allowed on the plot in Fig. 1*d*. This selection rule for DNA conformational transitions implies that the interaction that causes transitions has radial symmetry. We are not aware of a detailed explanation for this phenomenon, but it is physically plausible that the molecule, once extended with a particular number of azimuthal nodes, must contract back to a more compact object before it can extend out into a conformation with a different number of azimuthal nodes.

The autocorrelations of the PC amplitudes [i.e., the diagonal elements $\tilde{\rho}_{pp}(\tau)$] show nonexponential decay in time, a phenomenon due partially to the mode-mixing shown in Fig. 2*a* and partially to the nonlinear interaction discussed below. Nonetheless, a characteristic relaxation time can be associated with each PC by examining its slope near zero time lag by using the formula

$$\frac{1}{\tau_p} = \frac{1}{\delta t}\frac{\tilde{\rho}_{pp}(1) - \tilde{\rho}_{pp}(2)}{\tilde{\rho}_{pp}(1)}.$$

The values of $\tilde{\rho}_{pp}$ at lags 1 and 2 were used here (rather than at lags 0 and 1) to avoid contaminations from shot noise and other δ-correlated noise sources. These relaxation times are shown in Fig. 2*b* and are well fit by a power-law $\tau_p \sim p^{\alpha}$, with $\alpha = -0.55 \pm 0.05$ (95% confidence interval).

Why do the relaxation times of the PC modes scale like $\tau_p \sim p^{-1/2}$ while the relaxation times of the Zimm modes scale like $p^{-3/2}$? In the Zimm model, the modes are sinusoids, like the vibrations of a violin string, with a characteristic wavevector $k_p \sim p$. Zimm showed that a fluctuation with a wavevector of k has a relaxation time $\tau \sim k^{-3/2}$. In the PC model the fundamental motions are those of a three-dimensional spherically symmetric elastic continuum. In k space, the number of modes with a wavevector less than k in magnitude scales as $p \sim k^3$, and so the pth mode will have a wavevector $k_p \sim p^{1/3}$. The relation $\tau \sim k^{-3/2}$ still applies, whence $\tau_p \sim p^{-1/2}$. Thus, the essential difference between the Zimm and PC models is that the Zimm model deals with the vibrations of a one-dimensional string, whereas the PC model deals with the vibrations of a three-dimensional elastic continuum.

Now we quantify nonlinear interactions between the modes. The approach is to fit a linear model to the dynamics and then to look for higher-order correlations among the residuals. If the time evolution of the system were linear and Markovian (as required by both the Rouse and Zimm models), then the vector of amplitudes in each eigenstate would evolve according to a Langevin equation,

$$\mathbf{a}(k + 1) = \mathbf{M}\mathbf{a}(k) + \boldsymbol{\xi}(k), \qquad [8]$$

where $\mathbf{M}$ is a transition matrix, $\boldsymbol{\xi}$ is a vector of Gaussian white noise describing the effect of thermal fluctuations, and k is the frame index ($t = k\delta t$) (see *SI Appendix*). We considered only the first 15 eigenvectors and estimated the elements of $\mathbf{M}$ by a least-squares fit to the data. The time series of the residuals $\boldsymbol{\xi}(k)$ was examined for signs of nonlinear dynamics. For example, a conformation-dependent internal friction would lead to a nonzero value of the third-order correlation function:

$$\tilde{\rho}^{(3)}_{pq}(\tau) = \frac{\langle \xi_p^2(t + \tau)a_q(t)\rangle}{\text{var}(\xi_p)\text{var}(a_q)^{1/2}}. \qquad [9]$$

This correlation function characterizes the effect of amplitude in mode q affecting the friction (and hence the thermal fluctuations $\boldsymbol{\xi}$) in mode p. Fig. 2*c* shows that for several p and q, $\tilde{\rho}^{(3)}_{pq}(\tau) \neq 0$, indicating that the *ansatz* of Eq. **8** is only approximately true. Excitation in the (1, 0) mode leads to the largest nonlinear effects, strongly affecting the dynamics in the mode (1, 0) itself as well as the two (2, 1) modes and weakly affecting the dynamics in the two (1, 2) modes. Details of the calculation are given in *SI Appendix*. These nonzero correlations show that the intrinsic

nonlinear couplings in the DNA-shape dynamics are directly observable in our measurements. These effects are not predicted by the linearized Zimm model and cannot be detected by traditional scattering techniques.

We have shown that active trapping of single molecules of DNA in the ABEL trap provides rich dynamical information about the shapes and fluctuations of the molecules. The results presented here serve as a benchmark against which to test analytical theories and numerical simulations of polymer dynamics. The selection rules for conformational transitions (Fig. 2*a*) and the pattern of nonlinear coupling between modes (Fig. 2*c*) both await theoretical explanation.

In our analysis we found strong signs of dynamic internal interactions (forces depending on the relative motion of chain elements) but not of static internal interactions (forces depending on the relative position of chain elements). We justified neglect of static interactions on the basis of the good agreement between the PCA eigenstates of the data and those of a pure random walk. The appearance of features that resemble condensed globules, or pearls, in the movies is, we believe, an artifact of the finite resolution of the imaging; all features smaller than a diffraction-limited cutoff take on a globular appearance. The possibility remains that PCA is particularly insensitive to the effects of static interactions and that a different statistical description would be preferable. Such static interactions could arise from solvent-induced attraction, electrostatic repulsion, or excluded volume. However, the present experiment was performed under "good-solvent" conditions, for which the DNA is expected to adopt a random coil conformation.

It will be interesting to see how the descriptive statistics presented here change when a trapped molecule of DNA is subjected to physical and chemical perturbations, such as changes in pH, temperature, ionic strength, or the addition of proteins that interact with DNA. In particular, it would be interesting to add condensing agents to study the dynamics of the coil-to-globule transition (36), which is a topic of much interest for polyelectrolytes (37). One may also use more complex labeling schemes, such as labeling the ends of the polymer with a different color from the center. We hope that our experiments on DNA will spur efforts toward a better understanding of the role of HIs in biomolecular processes.

Materials and Methods

DNA Preparation. Double-stranded λ phage DNA (Molecular Probes) was dissolved in a buffer of 10 mM Tris·HCl, 10 mM NaCl, and 1 mM EDTA (pH 8.0). The fluorescent dye YOYO-1 (Molecular Probes) was added at a concentration of 1 dye/10 bp of DNA, and the mixture was incubated at room temperature in the dark for 30 min. An oxygen-scavenger system of glucose (4.5 mg/ml), glucose oxidase (0.43 mg/ml), catalase (72 μg/ml), and 2-mercaptoethanol (5 μl/ml) was added to the solution to reduce photobleaching. An antiadsorption polymer (POP-6, without denaturant, Applied Biosystems) was added at a concentration of 10% to prevent DNA sticking to the walls of the cell. The molecules were excited at 488 nm, and fluorescence was collected through a 495-nm-long pass filter. Under the experimental conditions, λ-DNA has a persistence length of $l_p \approx 60$ nm and a contour length of $L \approx 20$ μm.

Image Acquisition and Preprocessing. The ABEL trap cell was placed on an inverted optical microscope (TE300, Nikon), and epifluorescence images were acquired by using an oil-immersion objective with a numerical aperture of 1.3. The video images were formatted for data analysis as follows. Each frame was 32×32 pixels, with a pixel width corresponding to 118 nm in the sample plane. The small image size was chosen to allow a fast frame rate on the electron-multiplying CCD camera (Cascade 512B, Roper Scientific, Trenton, NJ). A background image (acquired under identical conditions to the data except with no DNA in the field of view) was subtracted from each frame. In a small fraction of the frames (≈5%), a second DNA molecule was seen floating through the field of view. In these frames, the pixels affected by the second molecule were manually set to the background level. Images were shifted to remove residual center-of-mass fluctuations uncompensated by the ABEL trap by using a bicubic interpolation to localize the center of mass to less than the pixel size. The total intensity of each frame was normalized to account for the slow rate of photobleaching of the YOYO-1 during the trapping period. For the present analysis, the data from all 21 molecules was aggregated except for the third-order correlation functions in Fig. 2*c*.

The $1{,}024 \times 1{,}024$ equal-time covariance matrix, $C(x_i, x_j)$, was numerically diagonalized in Matlab (Mathworks, Natick, MA). The eigenvectors with the largest eigenvalues were converted to 32×32 images. Eigenvectors 11–14 are nearly degenerate and arrived mixed together because of the presence of statistical noise in the covariance matrix. The space of four-dimensional rotations among these eigenvectors was manually searched to find linear combinations that had manifest symmetry. These linear combinations were taken to be the "true" eigenvectors.

We thank John Brauman, Joel E. Cohen, and Kit Werley for a careful reading of this manuscript and Willy Wiyatno of Applied Biosystems for samples of an antiadsorption polymer. This work was supported in part by National Science Foundation Grant CHE-0554681 and National Institutes of Health Grant 1R21-RR023149. A.E.C. acknowledges the support of a Hertz Fellowship.

1. Rouse PE (1953) *J Chem Phys* 21:1272–1280.
2. Zimm BH (1956) *J Chem Phys* 24:269–278.
3. Bixon M, Zwanzig R (1978) *J Chem Phys* 68:1890–1895.
4. Wang SQ, Freed KF (1986) *J Chem Phys* 85:6210–6224.
5. Wang SQ, Freed KF (1987) *J Chem Phys* 86:3021–3031.
6. Fujita H (1988) *Macromolecules* 21:179–185.
7. Berne BJ, Pecora R (2000) *Dynamic Light Scattering: With Applications to Chemistry, Biology, and Physics* (Dover, Mineola, NY).
8. Higgins JS, Benoit HC (1994) *Polymers and Neutron Scattering* (Oxford Univ Press, Oxford).
9. Quake SR, Babcock HP, Chu S (1997) *Nature* 388:151–154.
10. Hatfield JW, Quake SR (1999) *Phys Rev Lett* 82:3548–3551.
11. Strick TR, Dessinges, MN, Charvin G, Dekker NH, Allemand JF, Bensimon D, Croquette V (2003) *Rep Prog Phys* 66:1–45.
12. Shaqfeh ESG (2005) *J Non-Newtonian Fluid Mech* 130:1–28.
13. Shusterman R, Alon S, Gavrinyov T, Krichevsky O (2004) *Phys Rev Lett* 92:048303.
14. Lumma D, Keller S, Vilgis T, Radler JO (2003) *Phys Rev Lett* 90:218301.
15. Haber C, Ruiz SA, Wirtz D (2000) *Proc Natl Acad Sci USA* 97:10792–10795.
16. Enderlein J (2000) *Appl Phys B* 71:773–777.
17. Andersson SB (2005) *Appl Phys B* 80:809–816.
18. Berglund AJ, Mabuchi H (2004) *Appl Phys B* 78:653–659.
19. Cohen AE (2005) *Phys Rev Lett* 94:118102.
20. Cohen AE, Moerner WE (2005) *Appl Phys Lett* 86:093109.
21. Berglund AJ, Mabuchi H (2005) *Opt Express* 13:8069–8082.
22. Armani M, Chaudhary S, Probst R, Shapiro B (2006) *J Microelectromech Syst* 15:945–956.
23. Cohen AE, Moerner WE (2006) *Proc Natl Acad Sci USA* 103:4362–4365.
24. Cohen AE, Moerner WE (2005) *Proc SPIE Int Soc Opt Eng* 5699:296–305.
25. Cohen AE, Moerner WE (2005) *Proc SPIE Int Soc Opt Eng* 5930:191–198.
26. Viovy JL (2000) *Rev Mod Phys* 72:813–872.
27. Olivera BM, Baine P, Davidson N (1964) *Biopolymers* 2:245–257.
28. Cohen AE, Moerner WE (2007) *Phys Rev Lett* 98:116001.
29. Salmon R (1988) *Annu Rev Fluid Mech* 20:225–256.
30. Karplus M, Kushick JN (1981) *Macromolecules* 14:325–332.
31. Hayward S, Go N (1995) *Annu Rev Phys Chem* 46:223–250.
32. Balsera MA, Wriggers W, Oono Y, Schulten K (1996) *J Phys Chem* 100:2567–2572.
33. Jackson JE (2003) *A User's Guide to Principal Components* (Wiley, Hoboken, NJ).
34. Yamakawa H (1971) *Modern Theory of Polymer Solutions* (Harper & Row, New York).
35. Doi M, Edwards SF (1988) *The Theory of Polymer Dynamics* (Oxford Univ Press, Oxford).
36. Ueda M, Yoshikawa K (1996) *Phys Rev Lett* 77:2133–2136
37. Dobrynin AV, Rubinstein M, Obukhov SP (1996) *Macromolecules* 29:2974–2979.