Materials, Methods, and Derivations:

Supporting Online Material to accompany
A. E. Cohen and W. E. Moerner, "Dynamics of Single DNA Molecules in Equilibrium"

Data analysis

Below we give algorithms for analyzing and interpreting DNA conformational fluctuations. Throughout, we index quantities both by their time of measurement, t, and their measurement number, k.

Expression of the dynamics in the PC basis

Let $\boldsymbol{U}_p$ be the p^{th} eigenvector, and represent each image $S_i(t)$ as a vector, $\mathbf{S}(t)$, of length 1024. Then the time-dependent amplitude in $\boldsymbol{U}_p$ is given by the dot product, $a_p(t) = \delta\boldsymbol{S}(t)\cdot\boldsymbol{U}_p$. By construction $\langle a_p \rangle = 0$, so each a_p describes deviations from the average conformation. The eigenvalues are given by $\lambda_p = \langle a_p^2 \rangle$; i.e. each eigenvalue is proportional to the fraction of the variance of the entire data set that falls along its corresponding eigenvector.

The overall sign of each eigenvector is arbitrary: if $\boldsymbol{U}_\text{p}$ is an eigenvector, so too is $-\boldsymbol{U}_\text{p}$. This sign ambiguity implies that the time-dependent amplitudes, $a_\text{p}(t)$ have an arbitrary sign. Thus in the time-dependent covariance matrix $\tilde{\boldsymbol{\rho}}(\tau)$, one may arbitrarily switch the sign of the p^{th} row (for all τ), provided one also switches the sign of the p^{th} column (leaving the diagonal element positive).

Calculation of the linear and nonlinear dynamics

In both the Rouse and Zimm models, one expects the vector of amplitudes $\mathbf{a}(k)$ in the principal components to evolve linearly subject to a transition matrix $\mathbf{M}$ and white noise $\boldsymbol{\xi}$:

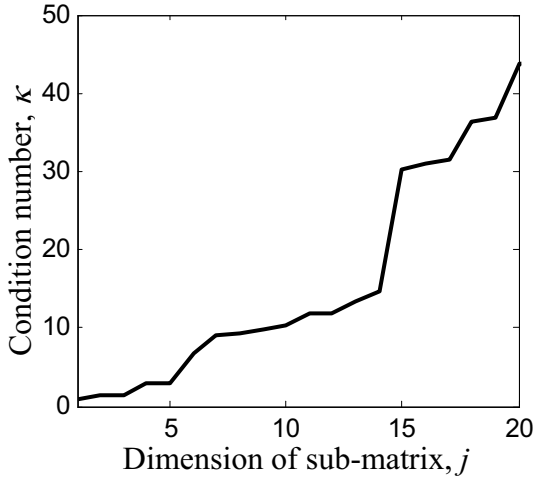
$$\mathbf{a}(k+1) = \mathbf{M}\mathbf{a}(k) + \boldsymbol{\xi}(k). \tag{0.1}$$

The challenge is to extract a best-fit $\mathbf{M}$ from the record of $\mathbf{a}(k)$, and then to determine whether Eq. (0.1) adequately describes the dynamics. Multiplying Eq. (0.1) on the right by $\mathbf{a}^\text{T}(h)$ (with $h < k$) and taking a time average yields

$$\tilde{\boldsymbol{\rho}}(k+1-h) = \mathbf{M}\tilde{\boldsymbol{\rho}}(k-h), \tag{0.2}$$

i.e. the covariance matrix of the vector $\mathbf{a}$ evolves deterministically under $\mathbf{M}$, and so in principle $\mathbf{M}$ can be extracted from any pair of samples of $\tilde{\boldsymbol{\rho}}$. In practice, $\tilde{\boldsymbol{\rho}}(0)$ is contaminated by measurement noise, so we calculated $\mathbf{M}$ from $\mathbf{M} = \tilde{\boldsymbol{\rho}}(2)\tilde{\boldsymbol{\rho}}^{-1}(1)$. To verify the stability of the matrix inversion, we calculated the condition number of the first $j\times j$ sub-matrices of $\tilde{\boldsymbol{\rho}}(1)$, for j between 1 and 25. Fig. S3 shows that the matrix has small condition number for $j < 15$. The Brownian contributions to $\mathbf{a}$ are obtained from $\boldsymbol{\xi}(k) = \mathbf{a}(k+1) - \mathbf{M}\mathbf{a}(k)$.

Fig. S3 Condition number of the first $j \times j$ submatrices of the covariance matrix at lag = 1. For $j < 15$, the sub-matrix is well-conditioned, validating the procedure used for calculating the transition matrix, **M**.



To check the validity of the estimate of ξ we first calculated the linear, time-dependent correlation

$$\frac{\left\langle \xi_p(t+\tau) a_q(t) \right\rangle}{\mathrm{var}(\xi_p)^{1/2}\, \mathrm{var}(a_q)^{1/2}}. \tag{0.3}$$

This second-order correlation differs from Eq. 9 in the Text because $\xi_p(t)$ is not squared in (0.3). Only the diagonal terms of (0.3) are nonzero, as shown in Figure S4a. The unusual time-dependence of the diagonal elements can be understood from a scalar analogue of Eq. (0.1) with noisy observations. The governing equations are:

$$a(k+1) = Ma(k) + \xi(k) \tag{0.4}$$

$$b(k) = a(k) + \chi(k), \tag{0.5}$$

where b is the observed variable and χ is independent identically distributed Gaussian measurement noise. After subtracting off the linear dynamics, the residuals are:

$$\begin{aligned} q(k) &\equiv b(k+1) - Mb(k) \\ &= \xi(k) + \chi(k+1) - M\chi(k) \end{aligned}. \tag{0.6}$$

Apart from a constant factor, our estimate of the diagonal elements of (0.3) in this scalar analogue is given by $\left\langle q(k+h)b(k) \right\rangle$. Several special cases need to be considered to calculate this quantity.

a) $h = 0$

$$\begin{aligned} \left\langle q(k)b(k) \right\rangle &= \left\langle \left[\xi(k) + \chi(k+1) - M\chi(k) \right] \left[a(k) + \chi(k) \right] \right\rangle \\ &= -M \left\langle \chi^2 \right\rangle \end{aligned} \tag{0.7}$$

b) $h = -1$

$$\begin{aligned} \left\langle q(k-1)b(k) \right\rangle &= \left\langle \left[\xi(k-1) + \chi(k) - M\chi(k-1) \right] \left[Ma(k-1) + \xi(k-1) + \chi(k) \right] \right\rangle \\ &= \left\langle \chi^2 \right\rangle + \left\langle \xi^2 \right\rangle \end{aligned} \tag{0.8}$$

c) $h \le -2$

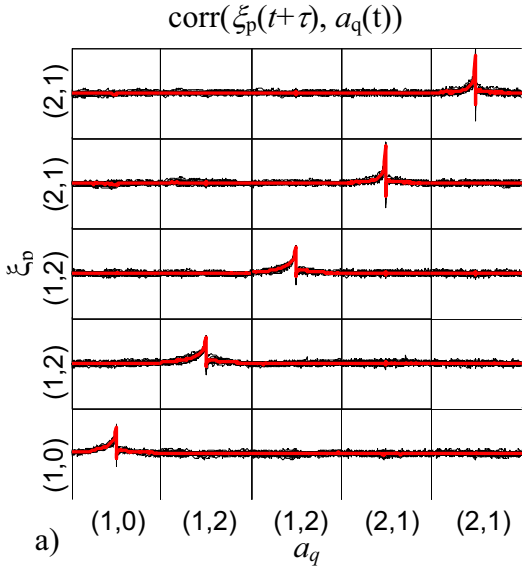
$$\left\langle q(k+h)b(k) \right\rangle = M^{h-1} \left\langle \xi^2 \right\rangle \tag{0.9}$$

d) $h \ge 1$

$$\left\langle q(k+h)b(k) \right\rangle = 0. \tag{0.10}$$

The four regimes embodied in Eqs. (0.7)-(0.10) are clearly seen in Fig. S4a. The correlation grows exponentially at negative times, jumps positive at $h = -1$, jumps negative at $h = 0$, and is zero for positive times.

Importantly, there is no indication from Fig. S4a that anything unusual is going on in the off-diagonal elements. Only when ξ_p is squared do the nonlinear couplings appear. The overall sign of each column of $\tilde{\boldsymbol{\rho}}^{(3)}$ is arbitrary because the sign of each component of **a** is arbitrary.



corr(ξ_p^2, a_q)

(2, 1)	0.19	0.004	0.021	0.018	-0.004
(2, 1)	0.19	-0.006	0.020	0.014	-0.009
(1, 2)	0.065	0.010	-0.021	-0.011	-0.043
(1, 2)	0.049	-0.007	0.003	-0.016	-0.039
(1, 0)	0.14	-0.002	0.027	-0.007	-0.000
ξ_p^2 / a_q	**(1, 0)**	**(1, 2)**	**(1, 2)**	**(2, 1)**	**(2, 1)**

b)

Fig. S4 Statistical properties of the residuals after fitting to the linear model of Eq. (0.1). a) Cross-correlation of the Brownian displacements and the measured mode amplitudes in the first 5 eigenstates (Eq. (0.3)). The black lines show the calculation for individual molecules of DNA and the red lines are the average. The unusual time-dependence of the diagonal elements is explained by a model of a linear autoregressive process with measurement noise. Each box has a time axis of $\tau = (-450, 450$ ms), and a vertical axis of (-0.6, 0.6). b) Table of numerical values at $\tau = 0$ of the third-order cross-correlation in the text, $\tilde{\boldsymbol{\rho}}^{(3)}(\tau)$. Statistical errors on all elements are approximately ±0.006. These are the peak-heights of the plot in Fig. 2c.