Boulder Notes on: Colloidal particle motion as a diagnostic of DNA conformational transitions

Philip C. Nelson

draft June 28, 2007

Abstract

Tethered particle motion is an experimental technique to monitor conformational changes in single molecules of DNA in real time. This article reviews some recent work on theoretical problems inherent in the interpretation of TPM experiments, both in equilibrium and dynamical aspects.

1 Introduction

Molecular biophysics attempts to explain life processes by understanding the behavior of individual macromolecules. Broadly speaking, past work has relied on (a) traditional light microscopy, which can discern the actions of individual objects, in physiological conditions and in real time, but is limited to scales bigger than half a micrometer, (b) electron microscopy and x-ray diffraction, which resolve fractions of a nanometer but give little information about dynamics, and (c) electrophysiology, which gives detailed dynamic information but with an indirect readout, applicable mainly to one class of molecular device. The recent rise of single molecule biophysics has complemented these approaches, in some ways combining their strengths. Working at the single molecule level allows us to see differences in behavior between different molecules, as well as making it possible to extract detailed kinetic information.

Tethered particle motion (TPM) is an experimental technique to monitor conformational changes in single molecules of DNA in real time. In this technique, a large colloidal particle (typically half a micrometer in diameter) is attached to the end of a DNA molecule of interest (the "tether"); the

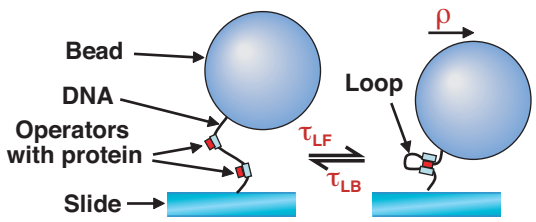

Figure 1: A DNA molecule flexibly links a bead to a surface. The motion of the bead's center is observed and tracked, for example as described in Refs. [1, 2]. In each video frame, the position vector, usually projected to the xy plane, is found. After drift subtraction, the mean of this position vector defines the anchoring point; the vector $\mathbf{r}$ discussed in this article is always understood to be the drift-subtracted, projected position, measured relative to this anchoring point. The length of this vector is called ρ. In this cartoon, the conformational change of interest is loop formation: Regulatory proteins, for example dimers of lambda repressor cI, bind to specific "operator" sites on the DNA. A loop forms when two (or more) such dimers subsequently bind to each other. A closely related system involves the lactose repressor protein, LacR (also called LacI): LacR exists in solution as a tetramer, which has two binding sites for DNA. Looping occurs when such a tetramer binds first to one, then to the other of two distant operator sites on the DNA. In each case, recent experiments have sought to determine the values of the time constants $\tau_{\rm LF}$ and $\tau_{\rm LB}$ for loop formation and breakdown. f:cartoon

other end is fixed to the microscope coverslip, and the resulting motion is passively observed. This review will outline the basics of the method, then turn to more recent applications to the kinetics of DNA looping.

2 TPM technique {s:tt}

The TPM technique uses the motion of a large, optically resolvable object (a bead) to report on changes in a nanoscale object of interest (a DNA molecule). The change of interest could include the progress of a processive motor that walks along DNA, as well as loop or kink formation on the DNA as proteins bind to it (Fig. 1). Experimental aspects of implementing the technique, including the attachment of the DNA of interest to the mobile bead at one end and the immobile surface at the other, are discussed in the original articles, including [3, 4, 5, 6, 7, 2, 8, 9, 10, 11].

Once the DNA/bead constructs are created, typically they are attached to the surface with a density small enough to prevent mutual interference, yet large enough that ten or more beads are

simultaneously visible in a microscope frame. The tethered beads are then observed, for example using differential interference contrast microscopy. The observations are typically at video rate, either gathering light throughout the entire frame time (typically 33 msec), or (with an electronic shutter) for some fraction of the frame (for example, 1 msec [2]). Automated data acquisition software can then track one or more beads in real time, characterizing each bead's motion; again details are available in the articles cited above. Some implementations do not track individual trajectories, instead observing the blurred average image of each bead [5, 10]; this review will focus on particle-tracking implementations, e.g. [6, 2].

Real-time tracking allows the experimenter to accept or reject individual beads prior to the start of datataking. For example, beads stuck to the surface will display little or no Brownian motion and can be rejected. During datataking each bead's image is fitted to a standard form to infer its position projected to the microscope focal plane; some experiments also report on the third component of bead position [7]. Each bead wanders about an average position, its anchoring point; Fig. 2 shows a typical time series for the bead's distance from its anchoring point.

Typically at least several acceptable beads will survive these initial cuts in a single microscope frame; the experimenter can select those for tracking, then observe their motion for an hour or more (Fig. 2). This simultaneous acquisition confers some advantages over other single molecule methods: First, it allows a degree of parallel processing, increasing experimental throughput. Second, as described later the availability of multiple time series, in time registry, assists in the removal of instrumental vibration and drift. The very long observation times available open the possibility of observing very slow kinetics. In contrast, other single molecule measurements such as Förster resonance energy transfer are limited in duration to the lifetime of a single flurophore.

After datataking, data must generally be be examined for failures of the particle-tracking software: Occasional individual outlier points can be deleted from the time series or replaced by the mean of flanking points, whereas time series where the tracking software loses a bead altogether must be discarded or truncated. Next, some correction must be applied for the effects of instrumental drift. Slow drift can be partially removed by low-pass filtering the bead position time series. Most experiments take the additional step of running a sliding window over the data, typically of 4 s duration, and within each window computing the variance of bead position, yielding a measure

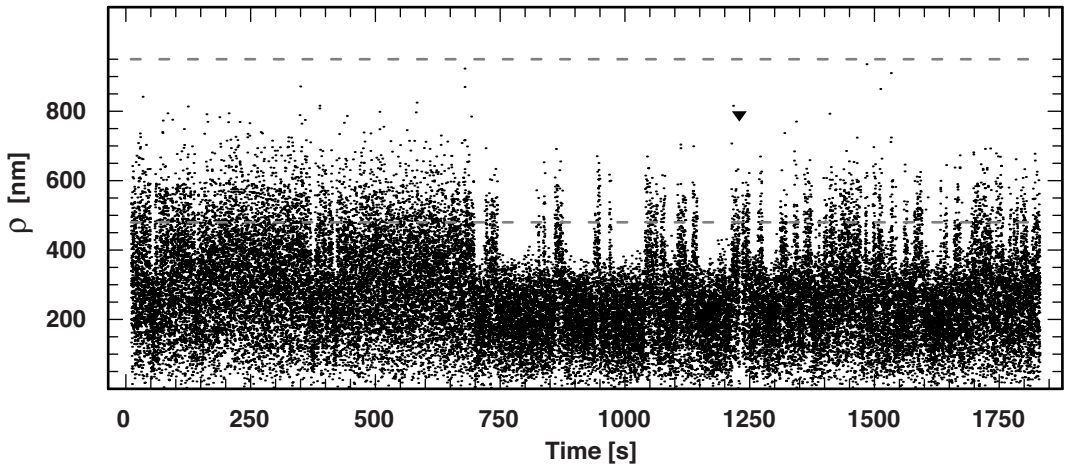

Figure 2: Typical time series of bead positions. DNA constructs of total length 3477 basepairs (bp) were attached at one end to a glass coverslip and at the other to a 480 nm diameter bead. The vertical axis gives the distance of the bead center from its attachment point, as reported by particle-tracking software, after drift subtraction. The trace shown passed the tests discussed in Ref. [2], for example the ones that eliminate doubly-tethered beads. The DNA construct contained two sets of three operator sites. The two sets of operators were separated by 2317 bp. The system contained cI repressor protein dimers at concentration 200 nM; repressor proteins bind to the operator sites on the DNA, and to each other, looping the DNA as in Fig. 1. A sharp transition can be seen from a regime of no loop formation to one of dynamical loop formation at $\sim$650 s. The dashed lines represent the two values of $\rho_{\max}$ corresponding respectively to the looped and unlooped states; control data in these two states was observed never to exceed these values. A brief sticking event, indicated by the inverted triangle, was excised from the data prior to analysis. (Experimental data from [12].) f:FIGTIMERHO

of the amplitude of the tethered motion. This "variance-filtering" method has the advantage of removing instrumental drift on time scales slower than the window width, but also degrades the ability to see true dynamics in the bead on time scales faster than the window.

A better approach might be to include fixed (or "stuck") beads in the sample, track them, and subtract their position traces from those of the mobile beads; in practice, however, it is generally not convenient to prepare such samples. An intermediate method is to track several beads simultaneously, and to use their common mean position as a proxy for the absent fiducial fixed bead position. This "mean-bead" subtraction method can remove most drift motion, while preserving information about truly random Brownian motion in the bead traces [2, 12]. (The residual, unremoved instrumental motion can readily be characterized by preparing samples consisting entirely

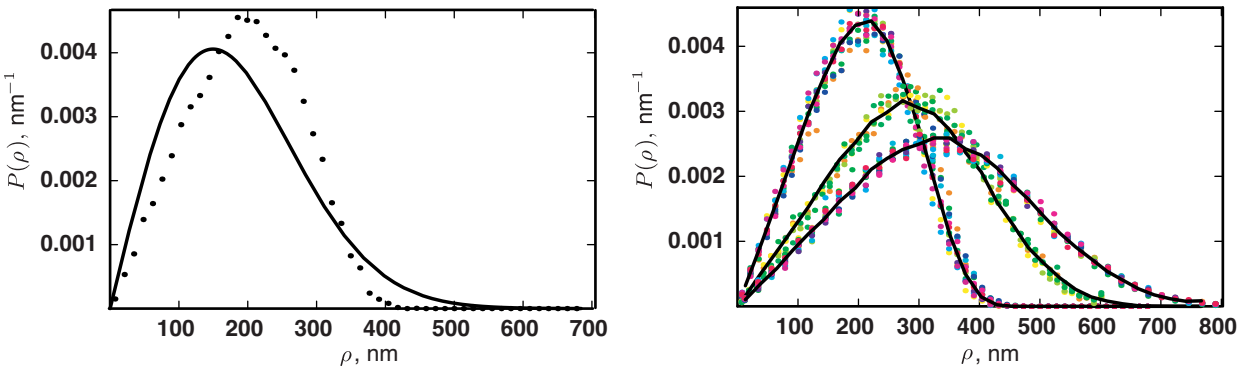

Figure 3: *Left:* A typical observed probability distribution function $P(\rho)$ for a 1096 bp-long tether (dots) is distinct from that of a two-dimensional Gaussian with the same mean-square deviation (curve). That is, the solid curve is the function $(2\rho/\sigma^2)\exp(-(\rho/\sigma)^2)$, where $\sigma = 211\,\text{nm}$. *Right:* Probability distributions (histograms) $P(\rho)$ for tethers of length $L = 957$, 2211, and 3477 bp (left to right). Dots: For each length, distributions from several tethers are represented in different colors. Curves: Predictions from theoretical model described in Sect. 3.1, with $\xi = 43\,\text{nm}$. The bead radius R_{bead} was known to be 240 nm from the manufacturer's specifications. (Data from [2].) f:distrs

of stuck beads and applying the same analysis to them.)

Unremoved drift can make the distribution of bead positions appear stretched in the direction of drift. Therefore, a final step of bead selection involves assessing the circular symmetry of the observed bead positions *after* drift subtraction. Beads whose distributions have principal axes differing by more than 5% may be attached to the surface by multiple tethers, or be otherwise faulty; such traces should be rejected. The surviving beads have equilibrium position distributions that depend only on distance from the anchoring point; Fig. 3 shows some examples.

3 Theory of tethered Brownian motion

3.1 Equilibrium fluctuations {s:ef}

The shapes of the distributions in Fig. 3 are not Gaussian, and the question arises whether we can predict them from first principles. Beyond such basic polymer science questions, an *a priori* knowledge of, say, the mean-square bead excursion for simple tethers is useful for interpreting experiments where the tether conformation is not known in advance. For example, tether length may be changing in time, in a way we would like to measure, as a processive enzyme walks

along the DNA. Or, effective tether length may change stochastically as proteins bind to the DNA, kinking or looping it. A third motivation for such a theoretical study is that by comparing theory to experiment, we gain confidence both that the experiment is working as desired and that the underlying polymer theory, which we may wish to apply to other problems, is adequate.

Although the end-end distribution of a semiflexible polymer such as DNA is a classical problem in polymer physics, the present problem differs from that one in several respects. For example, the DNA is not isolated, but instead is confined between two surfaces. DNA attached to a single planar surface experiences an effective entropic stretching force due to the steric exclusion from half of space; a similar effective repulsion exists between the DNA and the large bead. Far more important than these effects, however, is the steric exclusion of the bead from the wall. Ref. [13] argued that the effect of this exclusion would be to create an entropic stretching force on the DNA of the general form $F_{\rm eff} = k_{\rm B}T/z$ for $z < 2R_{\rm bead}$, where z is the height of the DNA endpoint from the wall, $R_{\rm bead}$ is the radius of the bead, and $k_{\rm B}T$ is the thermal energy at temperature T. (For $z > 2R_{\rm bead}$ the entropic force is zero.) Even if we do not measure z directly, this stretching force affects the distributions of projected distance, reducing its spread.

Additional subtleties of the problem include the fact that the polymer itself has two additional length scales in addition to the bead radius, namely its persistence length ξ and finite total length L, and the fact that we do not observe the polymer endpoint, but rather the center of the attached bead. Some of these effects have been studied in an analytical formalism for the low-force case [14], but for *zero* applied stretching force the steric constraints, not fully treatable in that formalism, become important. For this reason Refs. [13, 2] developed a Monte Carlo calculation method (a similar method was independently used for a study of DNA cyclization by Czapla et al. [15]).

The Monte Carlo code generated chains realized as strings of rotation matrices, all close to the identity matrix. Each such matrix represents the orientation of one polymer chain segment relative to its predecessor and was drawn from a Gaussian distribution, as expected for bend angles subject to a harmonic restoring torque [16, 17]. The width of the Gaussian bend distribution was chosen to reproduce the persistence length of DNA, which was either supposed to be known from other experiments in similar solvent conditions, or taken as a fitting parameter. The biotin and digoxigenin linkages attaching the DNA to bead and wall were treated as freely flexible pivots,

and so the orientation of the first chain segment, and that of the bead relative to the last segment, were taken uniformly distributed in the half-spaces allowed by the respective surfaces. Once a set of relative orientations was selected, the corresponding absolute orientations were found by successive matrix multiplication. The spatial location of each chain segment was then found by following the 3-axis of each orientation triad, and the chain and bead were then checked for steric clashes. Chain/bead configurations surviving this check were then entered into histograms of projected 2D distance between bead center and the anchoring point. The corresponding distributions for bead position are seen in Fig. 3 to fit the experimental observations quite well after adjusting a single parameter, the chain persistence length ξ; the fit value accords with published values obtained by force–extension measurement.

We can see the trends in the data more clearly if we reduce the distributions to their root-mean-square excursion, a quantity often used in experiments to characterize tethered particle motion. To facilitate fitting the theory to the data, the RMS excursion radius was computed for one fixed value of ξ and a grid of different values of $R_{\rm bead}$ and L in the range of interest, then summarized by an interpolating function. Dimensional analysis shows that $\sqrt{\langle\rho^2\rangle}/\xi$ can be written as a dimensionless function of L/ξ and $R_{\rm bead}/\xi$, so this calculation suffices to find the RMS excursion at any $\xi,\ L,$ and $R_{\rm bead}$ [2]. The resulting interpolating function can then be readily fit to experimental data.

The above procedure works well for TPM data taken with a fast shutter speed [2]. For longer exposure times, however, the theory must be corrected to account for the blurring of the image due to Brownian motion during each exposure [11]. Fig. 4 shows a global fit of the corrected theory to excursion data for three different bead radii. The discrepancy between theory and experiment may reflect unremoved instrumental drift, for example high-frequency motion that the filtering algorithm could not distinguish from true Brownian motion. Alternatively, the effective bead size may be slightly different from the nominal value, or effectively different due to surface irregularity. Finally, hydrodynamic effects slowing the bead's motion when close to the surface may set an equilibration time scale longer than the observation time.

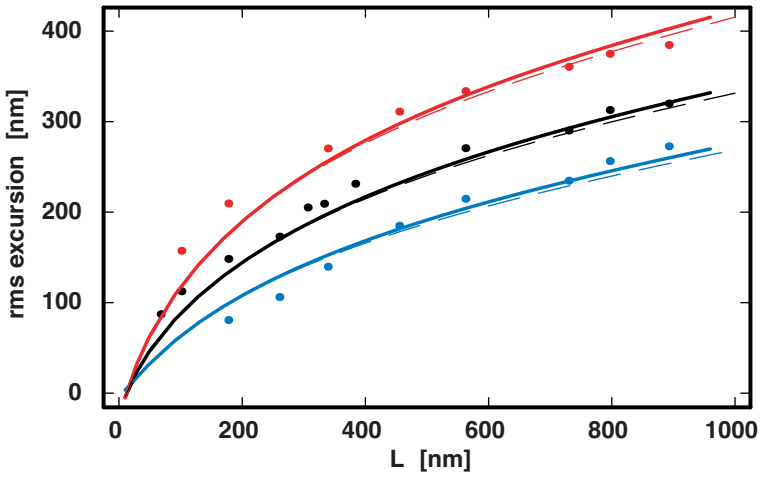

Figure 4: *Dots:* Experimental values for RMS motion of bead center for the three different bead sizes: Top to bottom, $R_{\mathrm{bead}} = 485,\ 245$, and $100\,\mathrm{nm}$. Each dot represents the average of approximately 20–200 different observed beads with the given tether length. *Curves:* Theoretically predicted RMS motion, corrected for the blurring effect of our long shutter time. The solid curves assume $\xi = 50\,\mathrm{nm}$; the dashed curve assumes $\xi = 47\,\mathrm{nm}$. There are *no other fit parameters;* the theoretical model uses values for bead diameter given by the manufacturer's specification. (Data and theory from [11].) `f:calibrplot`

3.2 Dynamics of tethered particle motion

The previous subsection discussed equilibrium bead fluctuations, but for the study of dynamic phenomena we also wish to understand the *dynamics* of tethered motion. A simple model captures qualitatively the expected character of the motion [18]. We imagine dividing time into very small intervals. At each time step, the change in bead position relative to the previous step is a combination of a random, diffusive step (Brownian motion), plus a deterministic drift step determined by the particle's current position in an effective potential well. The potential well comes from the entropic elasticity of the tether. The friction constant for the deterministic component of the motion is related to the diffusion constant for the Brownian component by the Einstein relation. If we roughly model the potential as that of a Gaussian chain, then the problem becomes the fluctuations of an overdamped harmonic oscillator, which can be solved analytically for the RMS excursion and the temporal autocorrelation function [19].

Indeed, experimental data for the equilibrium excursion distribution and the 2D autocorrelation function conform to the qualitative forms expected from the above description (namely the Gaus-

sian distribution and exponential falloff in autocorrelation respectively) [18]. However, for more detailed comparison of theory to experiment we would need to account for many modifications to the simple picture. For example, the friction and diffusion constants are much bigger than the values from Stokes's law, due to wall-proximity effects. Nor is the motion in the third (height) direction decoupled from projected-plane motion, as assumed in the simple model above. Nor is a DNA tether's entropic elasticity well represented by the Gaussian-chain formula. For all these reasons, prediction of TPM dynamics from first principles is a daunting task. But we can at least be fairly confident that whatever this motion, on video time scales it will have a Markov property: Each step will be drawn from a distribution depending only on the position at the start of that step (not on previous steps). Refs. [12, 20] extracted the single-step probability distributions from pairs of adjacent video frames, found a convenient parametric representation for these empirical functions, and then showed that they could be used to generate simulated TPM data with the same equilibrium and dynamics over long times as the real data. Knowing the dynamics of tethers that are not undergoing discrete conformational changes sets the stage for investigating tethers that *are* changing, for example by loop formation.

4 DNA looping

4.1 Background

All cells use genetic regulatory mechanisms that involve, among other things, the formation of loops in DNA similar to the one shown in Fig. 1. Particularly well studied cases involve bacteria, for example the lactose metabolism and lambda phage genetic switches [21, 22, 23, 24]. Reconstituting DNA looping behavior *in vitro*, and studying its dependence on physical parameters of the system, is thus an important step in clarifying the mechanism of gene regulation. A particular puzzling aspect of regulation by looping is that even very short loops, corresponding to tightly bent DNA, seem to form readily. In some cases, DNA-bending proteins appear to assist in loop formation; in others, specially flexible or pre-bent DNA sequence may help too. But a number of experiments appear to point to a surprising rate of loop formation even in the absence of such assistance.

4.2 TPM measurements of looping

TPM experiments are beginning to shed new light on looping [5, 1, 8, 25, 9, 12, 10, 20, 26, 27]. These experiments appear to be exempt from some of the complexities in corresponding DNA cyclization experiments. Moreover, by watching repeated loop formation and breakdown in real time, they can yield not only equilibrium constants but also corresponding rates, allowing us to quantify separately the kinetic barriers to loop formation (such as DNA bending energy) and breakdown (controlled by the unbinding kinetics of regulatory proteins and their operator sequences in the DNA) [9]. Finally, the TPM technique allows us to monitor a single molecule continuously as we change its surroundings, for example before and after the addition of the regulatory proteins [12].

Looping events are sometimes clearly discernible by eye in the microscopy, and in any case can be seen as sharp transitions in the time series after drift subtraction (Fig. 2). But a glance at the figure makes it clear that the interesting looping transitions are partially obscured by the bead's Brownian motion. Sect. 6 below will outline some new statistical methods to draw conclusions about looping from such data.

5 Theoretical models of DNA looping

5.1 Equilibrium calculation of looping J factor

Before discussing looping kinetics, however, we again begin with equilibrium. Classical works pointed out that the equilibrium constant for looping, or for cyclization (ring formation), can usefully be decomposed into *(i)* a binding constant for free operator DNA fragments to bind to the repressor protein, and *(ii)* a "J factor" contribution, describing the effect of the DNA chain joining the two operators [28, 29, 30, 31]. The J factor can roughly be regarded as the concentration of one operator in the vicinity of the other. Its overall dependence on the length of the intervening DNA between the operators can be qualitatively understood as reflecting two competing phenomena. First, a shorter tether confines the second operator into a smaller region about the first one, increasing the effective concentration. But if the required loop is too short, then forming it will entail a large bending elastic energy cost, depressing the probability by a Boltzmann factor. For these reasons, the cyclization J factor exhibits a peak at DNA length about 300 basepairs [32].

Later work extended Shimada and Yamakawa's calculation in many ways, using a variety of mathematical techniques [32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 15]. Recent work has attempted to extract the looping J factor from *in vivo* experimental data [44, 45], but any such determination contains uncertainties due to the complex world inside a living cell.

To clarify the situation, other experiments cited earlier have sought to quantify looping *in vitro*. Confronting these experiments with theory has required new calculations to incorporate the specific features of TPM. For example, shortening the DNA construct increases the entropic force exerted by bead–wall avoidance, discouraging looping [13]. To see what a measurement of TPM looping equilibrium tells us, Ref. [27] calculated the expected local concentration of one operator near the other, based on a particular mathematical model of DNA elasticity. Those authors chose a harmonic-elasticity model (a generalization of the traditional wormlike chain model), to see if it could adequately explain their experimental results, or if, on the contrary, some non-harmonic model (for example the one proposed in [46, 47]) might be indicated.

As in Sect. 3.1 above, the nonlocal steric constraints dictated a Monte Carlo approach. The authors of Ref. [27] modified the Gaussian sampling method outlined in Sect. 3.1 to generate many simulated DNA chains, apply steric constraints [13], and report what fraction of accepted chain/bead configurations had the two operator sites separated by 7 nm (the distance between operator centers in the LacR–DNA complex, as seen in Protein Data Bank structure 1LBG). The standard elastic model as an isotropic rod is inadequate for the description of DNA loops only a few helical repeats in length (see for instance [15]), so the authors modified the elasticity model to account for bend anisotropy and bend–roll coupling. The chain generation accounted for bead–wall, bead–chain, and wall–chain avoidance, but not chain–chain, which is much less important for the short tethers used in TPM experiments.

The result of the simulation was that the looping J factor for short loops was less than 0.015 times as great for the constructs with interoperator spacing around 100 bp as for those with spacing around 300 bp; this ratio was about a hundred times smaller than the experimentally determined ratio of 1.7 at optimum helical phasing [27]. The authors concluded that the hypotheses of harmonic elasticity, plus a rigid V-shaped protein coupler, could not explain their results. One possible explanation, for which other support has been growing, is the alternate hypothesis of DNA elastic

breakdown at high curvature [46, 47].

6 Determination of kinetics and event detection by hidden Markov modeling {s:dk}

The experimental determination of the fraction of time spent in the looped state seems to require that we know the looping state of the DNA tether at every moment of time. Indeed if that information were known, then we could also find the kinetic rate constants mentioned earlier, by tabulating the dwell times for loops to form and break down and then fitting them to exponential (or multiexponential) distributions. Unfortunately, data such as those in Fig. 2 do not admit an obvious point-by-point determination of looping state, in part because the probability distributions of bead positions in the different looped states have large overlap (Fig. 3).

One popular approach to this problem has been to run a sliding window across the data, typically of 4 s width, and within the window to apply the variance filter mentioned in Sect. 2. The filtered signal is local measure of the amplitude of the tethered Brownian motion; loop formation and breakdown events are defined as time points where it crosses a predetermined threshold [9], and then dwell time histograms are generated. There are several difficulties with this window/threshold approach, however [12, 20], including a strong dependence of the answers on the choice of window size. One approach to (one part of) this problem is to correct for missed transitions [9]. An alternative method, generally useful for obtaining kinetic and equilibrium constants from noisy data, is hidden Markov modeling (HMM) [48], which uses no filter at all.

The key point in HMM is the realization that it is not necessary to identify unambiguously the states at each time point. Instead, a kinetic model is assumed for the underlying "hidden" Markov process of interest (here the loop/unloop transitions). The model contains a small number of unknown parameters (here, the rate constants $\vec{\alpha} = (\tau_{\rm LF}, \tau_{\rm LB})$ for loop formation and breakdown). The model also assumes that given the underlying hidden state q, the probability distribution of the observed quantity (here the bead position $\mathbf{r}$) is known. That is, we choose a transition matrix $T(q_{n+1}|q_n;\vec{\alpha})$ and measure the conditional probabilities $P(\mathbf{r}|q)$, for example by studying DNA tethers that are forced to be permanently looped or unlooped.

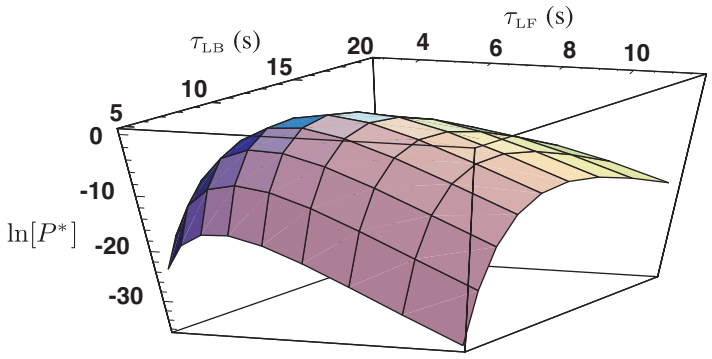

Figure 5: Evaluation of $\log[P^*_{\rm tot}(\mathcal{O})]$ on a logarithmically-spaced grid of $\tau_{\rm LF}, \tau_{\rm LB}$ lifetimes corresponding to data from Fig. 2 (from [20]). f:FIGLOGLIK

If we knew the exact states of the hidden variable at all times, $\{q_i, i = 1, \ldots, N\}$, then we could find the likelihood, or the probability that a particular time series of the observed quantity, $\{\mathbf{r}_i, i = 1, \ldots, N\}$, would arise. This likelihood contains contributions from both the transition matrix and the conditional probabilities. Since we *don't* know the underlying hidden variable, we marginalize it by summing the likelihood over *all* possible time series $\{q_i\}$. Substituting an actual experimentally observed time series $\{\mathbf{r}_i\}$ gives its likelihood to have been generated from the underlying model, as a function of the model's parameters, the unknown rates $\vec{\alpha}$. If the available time series is very long, we expect that this likelihood $P^*(\vec{\alpha})$ will be very sharply peaked about the true values of the rate constants $\vec{\alpha}$. That is, our best guess for the system's rate constants is the set of "maximum likelihood" values.

Hidden Markov analysis has recently been brought to bear on photon-counting single molecule techniques such as FRET [49, 50]. To adapt it to TPM experiments requires a slight generalization, however [12, 20], because the observed bead position, being a form of Brownian motion, has some memory. Indeed, both the observed motion and also the underlying hidden state have transition matrices describing the Markov character of their time evolution. Nevertheless, we can measure experimentally the transition matrix for the observed variable $\mathbf{r}$, using control experiments, and so reduce our problem to maximizing a likelihood P^* that depends only on two unknown rate constants for the hidden state transitions. The resulting function generally displays a single smooth peak (Fig. 5), and so can be maximized numerically by standard methods.

Acknowledgements

Nily Dan.

John Beausang, Rob Phillips, Lin Han, Hernan Garcia, Laura Finzi, David Dunlap, Chiara Zurla, Yale Goldman, Kevin Towles, . . .

This work was partially supported by NSF Grants DMR-0404674 and DGE-0221664, the Penn Nano/Bio Interface Center (NSF DMR04-25780), and by the Human Frontier Science Programme.

References

[1] C. Zurla, A. Franzini, G. Galli, D. D. Dunlap, D. E. A. Lewis, S. Adhya, and L. Finzi. Novel tethered particle motion analysis of cI protein-mediated DNA looping in the regulation of bacteriophage lambda. *Journal of Physics-Condensed Matter*, 18(14):S225–S234, 2006.

[2] P. C. Nelson, C. Zurla, D. Brogioli, J. F. Beausang, L. Finzi, and D. Dunlap. Tethered particle motion as a diagnostic of DNA tether length. *Journal of Physical Chemistry B*, 110(34):17260–17267, 2006.

[3] D. A. Schafer, J. Gelles, M. P. Sheetz, and R. Landick. Transcription by single molecules of RNA–polymerase observed by light-microscopy. *Nature*, 352(6334):444–448, 1991. ●●.

[4] H. Yin, R. Landick, and J. Gelles. Tethered particle motion method for studying transcript elongation by a single RNA-polymerase molecule. *Biophysical Journal*, 67(6):2468–2478, 1994.

[5] L. Finzi and J. Gelles. Measurement of lactose repressor-mediated loop formation and breakdown in single DNA-molecules. *Science*, 267(5196):378–380, 1995. ●●.

[6] N. Pouget, C. Dennis, C. Turlan, M. Grigoriev, M. Chandler, and L. Salomé. Single-particle tracking for DNA tether length monitoring. *Nucl. Acids Res.*, 32:e73, 2004.

[7] S. Blumberg, A. Gajraj, M. W. Pennington, and J. C. Meiners. Three-dimensional characterization of tethered microspheres by total internal reflection fluorescence microscopy. *Biophysical Journal*, 89(2):1272–1281, 2005.

[8] B. van den Broek, F. Vanzi, D. Normanno, F. S. Pavone, and G. J. L. Wuite. Real-time observation of DNA looping dynamics of type IIE restriction enzymes NaeI and NarI. *Nucleic Acids Research*, 34(1):167–174, 2006. ●●.

[9] F. Vanzi, C. Broggio, L. Sacconi, and F. S. Pavone. Lac repressor hinge flexibility and dna looping: Single molecule kinetics by tethered particle motion. *Nucleic Acids Research*, 34(12):3409–3420, 2006.

[10] O. K. Wong, M. Guthold, D. A. Erie, and J. Gelles. Interconvertible lactose repressor–DNA looped complexes revealed by single-molecule experiments. unpublished, 2007.

[11] L. Han et al. [[calibration]]. In preparation (to appear on `arXiv.org`)., 2007.

[12] J. F. Beausang, C. Zurla, C. Manzo, D. Dunlap, Laura Finzi, and P. C. Nelson. DNA looping kinetics analyzed using diffusive hidden Markov model. *Biophys. J.*, 92:L64–L66, 2007. ●.

[13] D. E. Segall, P. C. Nelson, and R. Phillips. Volume-exclusion effects in tethered-particle experiments: Bead size matters. *Physical Review Letters*, 96(8):088306–(1–4), 2006. ●.

[14] Y. Seol, J. Li, P. C. Nelson, T. T. Perkins, and M. D. Betterton. Elasticity of short DNA molecules: a combined experimental and theoretical study. Unpublished (to appear on `arXiv.org`).

[15] L. Czapla, D. Swigon, and W. K. Olson. Sequence-dependent effects in the cyclization of short DNA. *Journal of Chemical Theory and Computation*, 2(3):685–695, 2006. ●.

[16] A. V. Vologodskii. *Topology and physics of circular DNA*. CRC Press, Boca Raton, 1992.

[17] Philip Nelson. *Biological Physics: Energy, Information, Life*. W. H. Freeman and Co., New York, updated first edition, 2008.

[18] J. F. Beausang, C. Zurla, L. Finzi, L. Sullivan, and P.C. Nelson. Elementary simulation of tethered Brownian motion. *Am. J. Phys.*, 75:520–523, 2007.

[19] F. Reif. *Fundamentals of statistical and thermal physics*. McGraw-Hill, New York, 1965. Sect. 15.10.

[20] J F Beausang and P C Nelson. Diffusive hidden Markov model characterization of DNA looping dynamics in tethered particle experiments. Submitted; available on arXiv.org., 2007.

[21] K. Rippe, P. H. von Hippel, and J. Langowski. Action at a distance: DNA-looping and initiation of transcription. *TIBS*, 20:500–506, 1995.

[22] KS Matthews and JC Nichols. Lactose repressor protein: Functional properties and structure. *Prog. Nucleic Acid Res. Mol. Biol.*, 58:127–164, 1998.

[23] B. Revet, B. von Wilcken-Bergmann, H. Bessert, A. Barker, and B. Müller-Hill. Four dimers of lambda repressor bound to two suitably spaced pairs of lambda operators form octamers and DNA loops over large distances. *Current Biology*, 9(3):151–154, 1999.

[24] L. Saiz and J. M. G. Vilar. DNA looping: the consequences and its control. *Current Opinion in Structural Biology*, 16(3):344–350, 2006.

[25] D. Normanno, F. Vanzi, and F. S. Pavone. Unpublished., 2007.

[26] L. Han et al. [[sequence]]. In preparation (to appear on arXiv.org)., 2007.

[27] L. Han et al. [[short dna; concentration]]. In preparation (to appear on arXiv.org)., 2007.

[28] H. Jacobson and W. H. Stockmayer. Intramolecular reaction in polycondensations. I. The theory of linear systems. *J. Chem. Phys.*, 18:1600–1606, 1950.

[29] D. Y. Yoon and P. J. Flory. Moments and distribution functions for polymer chains of finite length. II. Polymethylene chains. *J. Chem. Phys.*, 61:5366–5380, 1974.

[30] U. W. Suter, M. Mutter, and P. J. Flory. Macrocyclization equilibria. 2. Poly(dimethylsiloxane). *J. Am. Chem. Soc.*, 98:5740–5745, 1976.

[31] N. L. Marky and W. K. Olson. Loop formation in polynucleotide chains. I. Theory of hairpin loop closure. *Biopolymers*, 21:2329–2344, 1982.

[32] J. Shimada and H. Yamakawa. Ring-closure probabilities for twisted wormlike chains – applications to DNA. *Macromolecules*, 17:689–698, 1984.

[33] P. J. Hagerman. Analysis of the ring-closure probabilities of isotropic wormlike chains: Application to duplex DNA. *Biopolymers*, 24:1881–1897, 1985.

[34] S. D. Levene and D. M. Crothers. Ring closure probabilities for DNA fragments by Monte Carlo simulation. *J. Mol. Biol.*, 189:61–72, 1986.

[35] A. A. Podtelezhnikov, C. D. Mao, N. C. Seeman, and A. Vologodskii. Multimerization-cyclization of DNA fragments as a method of conformational analysis. *Biophys. J.*, 79:2692–2704, 2000.

[36] Y. L. Zhang and D. M. Crothers. Statistical mechanics of sequence-dependent circular DNA and its application for DNA cyclization. *Biophysical Journal*, 84(1):136–153, 2003.

[37] A. J. Spakowitz and Z. G. Wang. Exact results for a semiflexible polymer chain in an aligning field. *Macromolecules*, 37(15):5814–5823, 2004.

[38] A. J. Spakowitz and Z. G. Wang. End-to-end distance vector distribution with fixed end orientations for the wormlike chain model. *Physical Review E*, 72(4), 2005. Part 1.

[39] Jie Yan, Ryo Kawamura, and John F. Marko. Statistics of loop formation along double helix DNAs. *Phys. Rev. E*, 71:061905–(1–17), 2005.

[40] N. Douarche and S. Cocco. Protein-mediated DNA loops: Effects of protein bridge size and kinks. *Physical Review E*, 72(6):061902–(1–10), 2005.

[41] Y. Zhang, A. E. McEwen, D. M. Crothers, and S. D. Levene. Statistical-mechanical theory of DNA looping. *Biophysical Journal*, 90(6):1903–1912, 2006.

[42] P. A. Wiggins and P. C. Nelson. Generalized theory of semiflexible polymers. *Phys. Rev. E*, 73:031906–(1–13), 2006.

[43] A. J. Spakowitz. Wormlike chain statistics with twist and fixed ends. *Europhys. Lett.*, 73:684–690, 2006.

[44] Lacramioara Bintu, Nicolas E Buchler, Hernan G Garcia, Ulrich Gerland, Terence Hwa, Jané Kondev, and Rob Phillips. Transcriptional regulation by the numbers: models. *Current Opinion in Genetics and Development*, 15:116–124, 2005.

[45] L. Saiz, J. M. Rubi, and J. M. G. Vilar. Inferring the in vivo looping properties of DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 102(49):17642–17645, 2005.

[46] J. Yan and J. F. Marko. Localized single-stranded bubble mechanism for cyclization of short double helix DNA. *Phys. Rev. Lett.*, 93:108108–(1–4), 2004. •.

[47] P. A. Wiggins, P. C. Nelson, and R. Phillips. Exact theory of kinkable elastic polymers. *Phys. Rev. E*, 71:021909–(1–19), 2005. •.

[48] L. R. Rabiner. A tutorial on hidden Markov-models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989. ••.

[49] M. Andrec, R. M. Levy, and D. S. Talaga. Direct determination of kinetic rates from single-molecule photon arrival trajectories using hidden Markov models. *Journal of Physical Chemistry A*, 107(38):7454–7464, 2003. •.

[50] S. A. McKinney, C. Joo, and T. Ha. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophysical Journal*, 91(5):1941–1951, 2006. •.