Characterizing Single-Molecule FRET Dynamics with Probability Distribution Analysis

Yusdi Santoso,[a] Joseph P. Torella,[a, b] and Achillefs N. Kapanidis*[a]

Probability distribution analysis (PDA) is a recently developed statistical tool for predicting the shapes of single-molecule fluorescence resonance energy transfer (smFRET) histograms, which allows the identification of single or multiple static molecular species within a single histogram. We used a generalized PDA method to predict the shapes of FRET histograms for molecules interconverting dynamically between multiple states. This method is tested on a series of model systems, including both static DNA fragments and dynamic DNA hairpins. By fitting the shape of this expected distribution to experimental data, the timescale of hairpin conformational fluctuations can be recovered, in good agreement with earlier published results obtained using different techniques. This method is also applied to studying the conformational fluctuations in the unliganded Klenow fragment (KF) of *Escherichia coli* DNA polymerase I, which allows both confirmation of the consistency of a simple, two-state kinetic model with the observed smFRET distribution of unliganded KF and extraction of a millisecond fluctuation timescale, in good agreement with rates reported elsewhere. We expect this method to be useful in extracting rates from processes exhibiting dynamic FRET, and in hypothesis-testing models of conformational dynamics against experimental data.

1. Introduction

Single-molecule fluorescence methods, and in particular single-molecule fluorescence resonance energy transfer (smFRET), have provided novel insights into the structures, interactions, and dynamics of biological systems.[1–11] Since the early days of smFRET, it has been apparent that the width of a FRET distribution carries important information about the static and dynamic heterogeneity of biomolecules.[12,13] Nonetheless, most single-molecule experiments only made use of the mean values of FRET distributions,[9,14,15] partially because the mean values were easier to obtain and more straightforward to interpret. Early efforts to interpret the widths of FRET histograms were limited to establishing lower bounds for histogram width, either through computational methods[12,13,16–18] or well-characterized static controls.[19]

Pioneering works[1–3] have laid a statistical framework for interpreting not only the widths, but also the overall shapes of FRET histograms. These approaches, which we collectively call probability distribution analysis (PDA) methods, are based on the premise that shot-noise-limited FRET histograms—the widths of which are determined solely by photon statistics—can be recapitulated using the distributions of photon counts obtained directly from experimental data. Indeed, these methods were able to predict the exact shapes of FRET histograms under typical experimental conditions. These methods were subsequently extended to account for the shapes of single-molecule fluorescence anisotropy histograms[20] and to describe mixtures of multiple, static FRET species.[2] In the latter case, a particularly useful application was the ability to extract the underlying distribution of static FRET values in a nonparametric fashion, by using a maximum-likelihood estimator approach. However, no method has yet been detailed to account for the shape of smFRET histograms for species dynamically interconverting between multiple FRET states, despite the fact that such a method has previously been suggested.[1] While refs. [1,21] have applied a variant of the PDA method to study the conformational dynamics of biomolecules (DNA hairpin and LacY, respectively), the details of such methodology have not been published. Dynamic conformational transitions are associated with many biomolecular functions.[22] While smFRET experiments on diffusing molecules have been important for our understanding of such systems, analysis of these experiments has generally been limited to time traces[4,6,23] or correlation-based analyses.[4,24–27] Time trace analysis of diffusing single molecules is mainly qualitative in nature due to the low information content available in short fluorescent bursts (with duration of ≈ 1 ms for an average globular protein or a short DNA fragment). Correlation analysis, on the other hand, provides quantitative information about the timescales of conformational fluctuations; however, its interpretation is often complicated by its strong model dependence, and by the presence of optical artifacts.[28]

[a] *Y. Santoso,+ J. P. Torella,+ Dr. A. N. Kapanidis*
Department of Physics and Biological Physics Research Group
University of Oxford, Parks Road, Oxford OX1 3PU (UK)
Fax: (+ 44) 1865-272-400
E-mail: a.kapanidis1@physics.ox.ac.uk

[b] *J. P. Torella+*
Department of Systems Biology
Harvard Medical School, Boston, MA 02115 (USA)

[+] *These authors contributed equally to this work.*

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cphc.201000129.

Herein, we use an extension of the PDA method to predict the shapes of smFRET distributions for molecules dynamically interconverting between distinct FRET states, described using kinetic schemes of arbitrary complexity. Specifically, we interrogate a biologically relevant model for conformational dynamics: a simple two-state system in which a molecule stochastically interconverts between two FRET states. This method uncovers quantitative information about FRET fluctuations directly from smFRET experimental data, therefore providing a novel tool for studying the conformational dynamics of protein and nucleic acid complexes. We validated this approach by using Monte Carlo simulations of two-state systems, as well as through experiments on DNA hairpins, which served as a model system for biomolecular dynamics. We further used this method to recover the conformational-fluctuation timescale of the fingers subdomain in the DNA polymerase I of *Escherichia coli*; our results were in excellent agreement with timescales determined using different methods, thus validating the PDA-based approach.

2. Results and Discussion

2.1. Theory

FRET is a photophysical interaction (a nonradiative transfer of excited-state energy from a fluorophore to a suitable chromophore) that reports on the proximity of two sites within a biomolecule or a complex of biomolecules. Typically, each of the sites is labeled with one of two different probes, which act as either the FRET donor (D) or acceptor (A). Experimentally, the FRET efficiency is generally defined as the number of photons detected in the acceptor emission channel over the total number of detected photons over time; in studies on diffusing molecules, the number of photons involved typically refers to a single fluorescent burst (i.e. the spike in the fluorescence photons that is observed when a fluorescent molecule moves in and out of the detection volume). This apparent FRET value, or proximity ratio, is defined as $E^* = (B_A + F_A)/N$, where F_A is the number of photons (excluding background) detected in the acceptor-emission channel, B_A is the background count in the acceptor-emission channel, and N is the total number of photons detected in both donor- and acceptor-emission channels. The proximity ratio is distinct from the accurate FRET efficiency due to several factors: the presence of background and spectral crosstalk, as well as differences in the detection efficiency and quantum yield between donor and acceptor fluorophores. The background contribution arises primarily from scattering of the incident laser light off water molecules (Raman scattering).[29] Some of the crosstalk contribution arises from the fraction of donor fluorescence emitted at the acceptor-emission channel (hereafter termed "leakage"); moreover, some of the acceptor fluorophores will be directly excited by the donor-excitation laser (hereafter termed "direct excitation"). Differences in the detection efficiencies and quantum yield of the fluorophores will influence the proportion of photons detected in each channel, therefore directly affecting the value of the proximity ratio. While it is possible to correct for these effects to

obtain accurate FRET values,[30] most smFRET experiments are concerned only with relative FRET changes, the interpretation of which is not typically complicated by these artifacts. As such, the proximity ratio, which we focus on herein, is a more common measure of relative distance changes in biomolecules.[1, 12, 31]

We build the theory on the formalism laid out in ref. [3], in which the theoretical distribution of the proximity ratio, E^*, can be calculated as [Eq. (1)]:

$$P(E^*) = \sum_{\text{all } F_D, F_A, B_D, B_A \text{ yielding } E^*} P(F_D, F_A, B_D, B_A) \tag{1}$$

which is a simple sum over the probabilities of all combinations of F_D, F_A, B_D, and B_A yielding the desired E^* value. Here, F_X is the number of photons detected in the X-emission channel (not including background photons) and B_Y is the background photons arising during the molecular transit and detected in the Y-emission channel. E^* is calculated as [Eq. (2)]:

$$E^* = \frac{B_A + F_A}{B_D + B_A + F_D + F_A} = \frac{B_A + F_A}{B_D + B_A + F} = \frac{B_A + F_A}{N} \tag{2}$$

where F is the total number of fluorescence photons arising from the molecule (and not from the background), and N is the total number of photons in a burst (including background). To extend this analysis to multiple molecular states, we generalize the photon counts to include contributions from each state (i) occupied by the molecule [Eq. (3)]:

$$E^* = \frac{B_A + \sum a_i}{B_D + B_A + \sum d_i + \sum a_i} = \frac{B_A + \sum a_i}{B_D + B_A + \sum f_i} = \frac{B_A + \sum a_i}{N} \tag{3}$$

where f_i is the fluorescence photon count while the molecule is in state i, and d_i and a_i are the subsets of these photons detected in the donor and acceptor channels, respectively. We can then rewrite Equation (1) as [Eq. (4)]:

$$P(E^*) = \sum_{\text{all } (B_A + \sum a_i)/N = E^*} P(\vec{a}, \vec{d}, B_D, B_A) \tag{4}$$

where $\vec{x}$ is a vector of all photon counts x_i, for all possible molecular states i.

The background in single-molecule experiments on freely diffusing molecules is Poisson-distributed with respect to time and independent of emission from the fluorescent particle [Eq. (5)]:[1, 27]

$$P(B_\gamma | T, r_\gamma) = \frac{(r_\gamma T)^{B_\gamma} e^{-(r_\gamma T)}}{B_\gamma!} \tag{5}$$

where r_γ is the rate of background in the Y-channel and T is the burst duration. We can thus rewrite Equation (4) as

[Eq. (6)]:

$$P(E^*) = \sum_{\text{all}\,(B_A \sum a_i)/N = E^*} P(\vec{a}, \vec{d}) \cdot P(B_D|T, r_D) \cdot P(B_A|T, r_A) \cdot P(T) \quad (6)$$

where $P(T)$ is the distribution of burst durations.

Although we do not know the joint distribution $P(\vec{a}, \vec{d})$ of all a_i and d_i, we can simplify the expression by assuming the total photons emitted in each state, f_i, are known. In this case, the values of a_i each follow a binomial distribution,[1,3] with their means set by the FRET efficiency of each state, E_i [Eq. (7)]:

$$P(E^*) = \sum_{\text{all}\,(B_A \sum a_i)/N = E^*} \left[\prod_{i=1}^{m} (P(a_i|f_i, E_i) \cdot P(f_i)) \right] \times$$
$$P(B_D|T, r_D) \cdot P(B_A|T, r_A) \cdot P(T) \quad (7)$$

where m is the number of states, and [Eq. (8)]:

$$P(a_i|f_i, E_i) = \binom{f_i}{a_i} (E_i)^{a_i} (1 - E_i)^{f_i - a_i} \quad (8)$$

is the binomial distribution describing the probability of detecting a_i photons in state i, where f_i total fluorescence photons have been emitted.

The joint distribution of all f_i is not directly measurable, but each f_i can be expressed as a function of known distributions by making the simplifying assumption of uniform excitation intensity. Compared with approaches incorporating more realistic spatial brightness functions,[32] this assumption decreases the variance of photon counts assigned to each FRET state, and therefore decreases the width of the FRET histogram. However, it does not affect the means or relative occupancies of each state in the histogram (the features to which our minimization function is most sensitive), and is much less computationally expensive. As a result, this method is rapid and achieves good agreement between the expected and predicted distributions (Figure 2).

Under this assumption, the total number of fluorescence photons f_i arising from each state follows a Poisson distribution with a mean proportional to the state's brightness, b_i, and to the time spent in that state, τ_i. Estimating the distribution of F by incorporating the experimental distribution of photon counts, $P(F)$ where $F = N - B_A - B_D$,[3] and assuming we know the joint distribution of all τ_i values, the distribution of f_i can then be reduced to a simple multinomial [Eqs. (9) and (10)]:

$$P(E^*) = \sum_{\text{all}\,(B_A + \sum a_i)/N = E^*} \left\{ \prod_{i=1}^{m} \left[P(a_i|f_i, E_i) \cdot P\left(f_i \middle| F, \vec{\tau}, \vec{b}\right) \right] \right\} \times$$
$$P(\vec{\tau}|T) \cdot P(B_D|T, r_D) \cdot P(B_A|T, r_A) \cdot P(T, F) \quad (9)$$

$$P\left(f_i \middle| F, \vec{\tau}, \vec{b}\right) = (F!) \left\{ \prod_{i=1}^{m} \left[\frac{(\tau_i / \sum (b_i \tau_i))^{f_i}}{f_i!} \right] \right\} \quad (10)$$

where we replace $P(T)$ with the joint distribution $P(T,F)$, since the terms are non-independent and can be obtained jointly from the experimental data. Herein we assume that all states have equal brightness, a good assumption for the fluorophores and experimental setup used. We also note that the effects of brightness on the PDA method have been treated in depth elsewhere.[2,20]

The dwell-time distribution of times spent in each state, $P(\vec{\tau}|T)$ (see Figure 1A), can be determined based on the posited kinetic model and the known experimental distribution of burst durations, $P(T)$. We note that while the dwell-time distribution $P(\vec{\tau}|T)$ is analytically solvable in simple (i.e. two-state) cases,[18] we advocate a numerical approach, where $P(\vec{\tau}|T)$ is determined using a Monte Carlo simulation of molecular transitions given the experimental $P(T)$; such an approach allows the evaluation of systems with kinetic schemes of arbitrary complexity.

2.2. Determining the Dwell-Time Distribution

To implement a Monte Carlo determination of $P(\vec{\tau}|T)$, we need to consider the rates of interconversion between each set of states, and the equilibrium occupancy of each state. For a biological molecule with m interconverting states, where the transition from state i to state j occurs with a single-exponential rate k_{ij}, and this rate is independent of the states previously occupied by the molecule (i.e. a "memoryless" system), the kinetic fluctuations of the molecule represent a continuous-time Markov process.[33] Defining the $m \times m$ ij-th rate matrix Q as the matrix of rate constants q_{ij}, the vector containing the equilibrium frequencies of each state, π, can be calculated as [Eq. (11)]:

$$\pi Q = 0 \text{ where } -q_{ii} = \sum_{i \neq j} q_{ij}, \quad \sum \pi_i = 1 \quad (11)$$

For any given molecule, the probability that it enters the volume while in state i is therefore equal to the i-th element of π. And, since the transition to any other state j is memoryless and occurs with rate constant q_{ij}, the waiting time distribution $P(t_{ij})$ for a transition into any state j is given by a single exponential distribution [Eq. (12)]:

$$P(t_{ij}) = q_{ij} e^{-q_{ij}t} \quad (12)$$

2.3. Implementation

While it is possible to implement PDA using a precise calculation of the proximity ratio histogram by considering all possible variable combinations,[2,3] we based our implementation on the Monte Carlo method of Nir et al.[1]. We evaluated Equation (9) using a Monte Carlo approach taking into account the experimental distribution of $P(T,F)$. For simplicity, we ignored the negligible contribution of background in our measurements (typically < 4 kHz, corresponding to a signal-to-noise ratio > 10). It has been previously shown that a background of < 6 kHz for typical single-molecule FRET confocal measurements has only a marginal effect on the FRET histograms[1]
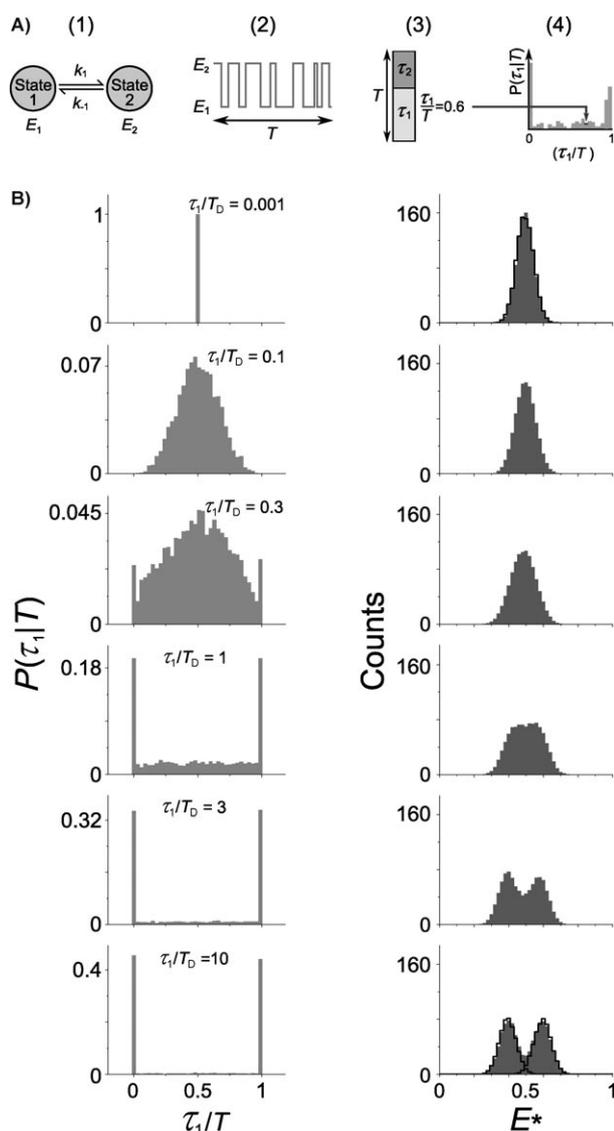
Figure 1. A) Monte Carlo simulations were used to generate distributions of dwell times, $P(\tau_i|T)$, for a two-state kinetic model. From left to right: 1) Simple two-state kinetic model, with the two states centered at E_1 and E_2, respectively; $k_1 =$ transition rate from state 1 to state 2; $k_{-1} =$ transition rate from state 2 to state 1. 2) Example time trace showing interconversions between the two states within a burst duration T. 3) τ_i is the total time spent in state i, where $\Sigma\tau_i = T$. The ratio τ_i/T gives the proportion of time within a burst spent in state i. 4) Histogram of dwell times in the first state, $P(\tau_1|T)$, was constructed by binning the τ_1/T value for each burst. B) Distributions of dwell times [$P(\tau_i|T)$, where $i = 1,2$ for a two-state system] obtained using Monte Carlo simulations (left) and their respective FRET distributions after considering the shot-noise contribution (right; dark gray histograms). $E_1 = 0.4$, $E_2 = 0.6$, $k_1 = k_{-1}$. For brevity, only the distributions of dwell times in the first state, $P(\tau_1|T)$, are plotted; the distributions of dwell time in the second state, $P(\tau_2|T)$, are identical due to the equal forward and backward rates. From top to bottom, the average dwell time in first state ($\tau_1 = 1/k_1$) was increased relative to the average diffusion time through the confocal volume (T_D). When the dwell times were much shorter than the diffusion time (fast dynamics), the two FRET states merged into one shot-noise-limited peak (e.g. $\tau_1/T_D = 0.001$; solid black line: shot-noise-limited peak with $E = 0.5$, $\sigma = 0$). On the other hand, when the dwell times were much longer (slow dynamics), the two states appeared as two well-resolved shot-noise-limited FRET peaks (e.g. $\tau_1/T_D = 10$; solid black lines: two separate shot-noise-limited peaks, $E_1 = 0.4$, $\sigma_1 = 0$ and $E_2 = 0.6$, $\sigma_2 = 0$).

(see also the Supporting Information, Figure S-2). If needed, a background contribution can be included as suggested in ref. [1].

This algorithm produces a predicted E^* histogram using the following approach:

1) Choose an oversampling factor K (typically $K = 5$) and realistic initial kinetic parameters.
2) Repeat for each collected burst:
3)   Repeat K times:
4)   Determine $\vec{\tau}$ by drawing from $P(\vec{\tau}|T)$ using Monte Carlo simulation of the kinetic model given the kinetic parameters and $P(T,F)$.
5)   Determine the number of photons f_i in each state by drawing from the multinomial distribution $P(f_i|F, \vec{\tau}, \vec{b})$.
6)   For each state, determine the photons emitted, a_i, by drawing from the binomial distribution $P(a_i|f_i, E_i)$.
7)   Add the value $\Sigma a_i/F$ to the E^* histogram.
8) Divide the final E^* histogram by K.
9) Refine the kinetic parameters using the gradient descent method and return to step (2) unless successive parameter estimations are different by less than a given tolerance (typically $< 0.1\%$ of the expected parameter value).

The E_i value in step (6) can be "broadened" by replacing the static E_i value in step (6) with a random variable drawn from a Gaussian distribution with a given mean, E_i, and standard deviation, σ_i. This broadening was necessary to account for the heterogeneity observed in experimental measurements of double-stranded DNA.[1,3] The kinetic parameters obtained at the end of step (9) represent the best approximation to the actual kinetic rates, given the proposed kinetic model.

To calculate the goodness of fit in our models, we use the reduced chi-squared statistic suggested by Antonik et al.[3] [Eq. (13)]:

$$\chi_r^2 = \frac{1}{n-m} \sum_{i=1}^{n} \frac{[\text{Freq}(X_i) - \text{Freq}_M(X_i)]^2}{\text{Freq}(X_i)} \qquad (13)$$

where m is the number of adjustable model parameters, n is the total number of nonzero bins, and Freq and Freq_M represent the frequency of data falling into bin i in either the data or prediction, respectively, and we include only those bins containing at least one data point. A perfect fit corresponds to $\chi_r^2 = 1$. To fit our data, we use minimization of χ_r^2 using a simple gradient-descent algorithm.[34] We chose this algorithm due to its implementation simplicity. We also note that the choice of minimization algorithm is not crucial for the PDA approach; moreover, better and faster algorithms can easily substitute the gradient-descent algorithm used here.

To provide a measure of the uncertainty of the fitting process, we repeated each fitting procedure five times and reported the standard deviation value of each fitted parameter. We note that a more accurate measure of the uncertainty on each parameter value can be obtained by considering the model

objective function over the space of the model parameters; such plots are generated in a later section (Figure 3).

2.4. Detectable Dynamic Timescales

To illustrate the effect of dynamics on the shape of the E^* histogram, we first considered a simple two-state kinetic system (Figure 1 A-1) in which a molecule fluctuates between FRET states E_1 and E_2 at forward and reverse rates k_1 and k_{-1}, which gives rise to fluctuations in FRET over time (Figure 1 A-2). We used a Monte Carlo method to calculate the distribution of times spent in each FRET state, $P(\tau_1, \tau_2 | T) = P(\vec{\tau} | T)$, where $\vec{\tau}$ is the joint distribution of τ_1 and τ_2 (Figure 1 A-3,4).

Figure 1 B illustrates the different $P(\vec{\tau} | T)$ achieved for different kinetic rates and dwell times, where $\tau_1 = \tau_2 = 1/k_1 = 1/k_{-1}$ (left: light gray histograms), and their effect on the resulting proximity ratio histograms, $P(E^*)$ (right: dark gray histograms). To generate the $P(\vec{\tau} | T)$ distributions, we assumed a uniform distribution of burst durations, $T = \tau_1 + \tau_2 = 1$ ms. To generate the $P(E^*)$ distributions, we assumed all bursts had the same number of photons, $F = 100$, and that the distribution was shot-noise limited [Eq. (9)]. While in our actual implementation neither burst duration nor photon count are assumed to be fixed [Eq. (9)], here we make these assumptions just for the purpose of generating Figure 1 B. From these simple examples (Figure 1 B), it is apparent that $P(\vec{\tau} | T)$ plays an important role in determining the shape of the final FRET distribution. In the limit of short dwell times, the two states fluctuate so fast that they appear as a single, shot-noise-limited state with intermediate mean E^* (top panel; compare dark gray histogram with the gray line). As the dwell time increases, this peak broadens, eventually splitting into two distinct peaks which, in the limit of very long dwell times, are themselves shot-noise-limited (bottom panel).[35] We note that the PDA method is most sensitive when $0.01 \ll (\tau_r / T_D) \ll 10$, where T_D is the average dwell time in the confocal volume (see the Supporting Information, Figure S-3). To achieve the highest sensitivity using the PDA method, one can tune the value of T_D to match the range of the conformational rates to be explored. This can be accomplished by increasing the size of the confocal spot, by using polyacrylamide gels, or by changing the viscosity of the surrounding medium.[4]

2.5. Comparison with smFRET Simulations

To validate the PDA approach, we first performed Monte Carlo simulations of diffusing single molecules (see the Experimental Section). In particular, we simulated molecules with either one or two states, and with different mean FRET values and rates of interconversion. We then used PDA to fit the simulated data using either a one- or two-state model and extracted both the kinetic parameters and a χ_r^2 value. We found excellent agreement between the original simulation parameters and the fitting results (Figure 2): the χ_r^2 values for all fits were close to 1, which indicated a good fit to the data in all cases. For single-state systems and two-state systems with equal forward and backward rates (Figure 2 A,B), the fitted parameters were well

within 2% of the original simulation parameters. For the two-state systems with differing forward and backward rates (Figure 2 C), the mean values of the fitted parameters were up to ≈20% away from the original parameters, but still within error of the fitted parameters in all cases.

Our steepest-descent minimization algorithm minimizes the sum of squared error (SSE) between the data and the PDA prediction over a significant parameter space [Eq. (14)]:

$$SSE = \sum_{i = all\ bins} [Freq(X_i) - Freq_M(X_i)]^2 \qquad (14)$$

SSE was chosen over χ_r^2 because it can be calculated using even those bins without any data points, and therefore provides a more stable minimization method in dealing with large differences between the data and the PDA prediction. A parameter minimization surface for a two-state dynamic system ($E_1 = 0.4$, $E_2 = 0.6$, and $1/k_1 = 1/k_{-1} = 1$ ms) is shown in Figure 3 A (which corresponds to the middle panel of Figure 2 B), where a minimum corresponding to best-fit rate parameters is easily identifiable. In some other cases, the SSE surface might be flat. This indicates the presence of a degeneracy problem in which many sets of parameters are equally good at minimizing the SSE (see Figure 3 B as well as the rightmost panel of Figure 2 C). Due to the stochastic nature of our PDA implementation, each run of the minimization routine will return slightly different results. The standard deviation obtained from these repeated fittings can be used as a measure of the certainty of the fitting procedure. When the SSE surface has a clear minimum, repeated PDA runs will yield a tight distribution of fitted parameters; when the SSE surface is flat, however, repeated PDA runs will stop at different points along the basin of the surface, which increases the variance in the fitted parameters. We can therefore use large standard deviations on fitted parameters as an indicator of a flat SSE surface, with the standard deviation itself providing an index of confidence in the fitted parameters.

2.6. A Static Control: Double-Stranded DNA

To further validate this approach, we analyzed two DNA samples using PDA: a short double-stranded DNA, and a dynamic DNA hairpin.

Figure 4 A shows the comparison between the PDA prediction and the experimental FRET distribution obtained for the $T_{1-Cy3B}, B_{18-ATTO647N}$ double-stranded DNA sample. The experimental FRET distribution is ≈1.8-fold wider than the prediction of the single-FRET state model (see the Supporting Information, Figure S-4, $\sigma_{experimental} = 0.061$, $\sigma_{shot\ noise} = 0.034$), consistent with refs. [1,3], which reported similar widening for double-stranded DNA samples. In ref. [1], the discrepancy between the experimental and shot-noise-only histograms was accounted for by assuming a quasi-static Gaussian distribution of distances with a standard deviation of 1.6 Å, in good agreement with the 2 Å we obtained for our sample. This analysis helped us determine the baseline broadening beyond shot noise expected for a typical double-stranded DNA molecule.
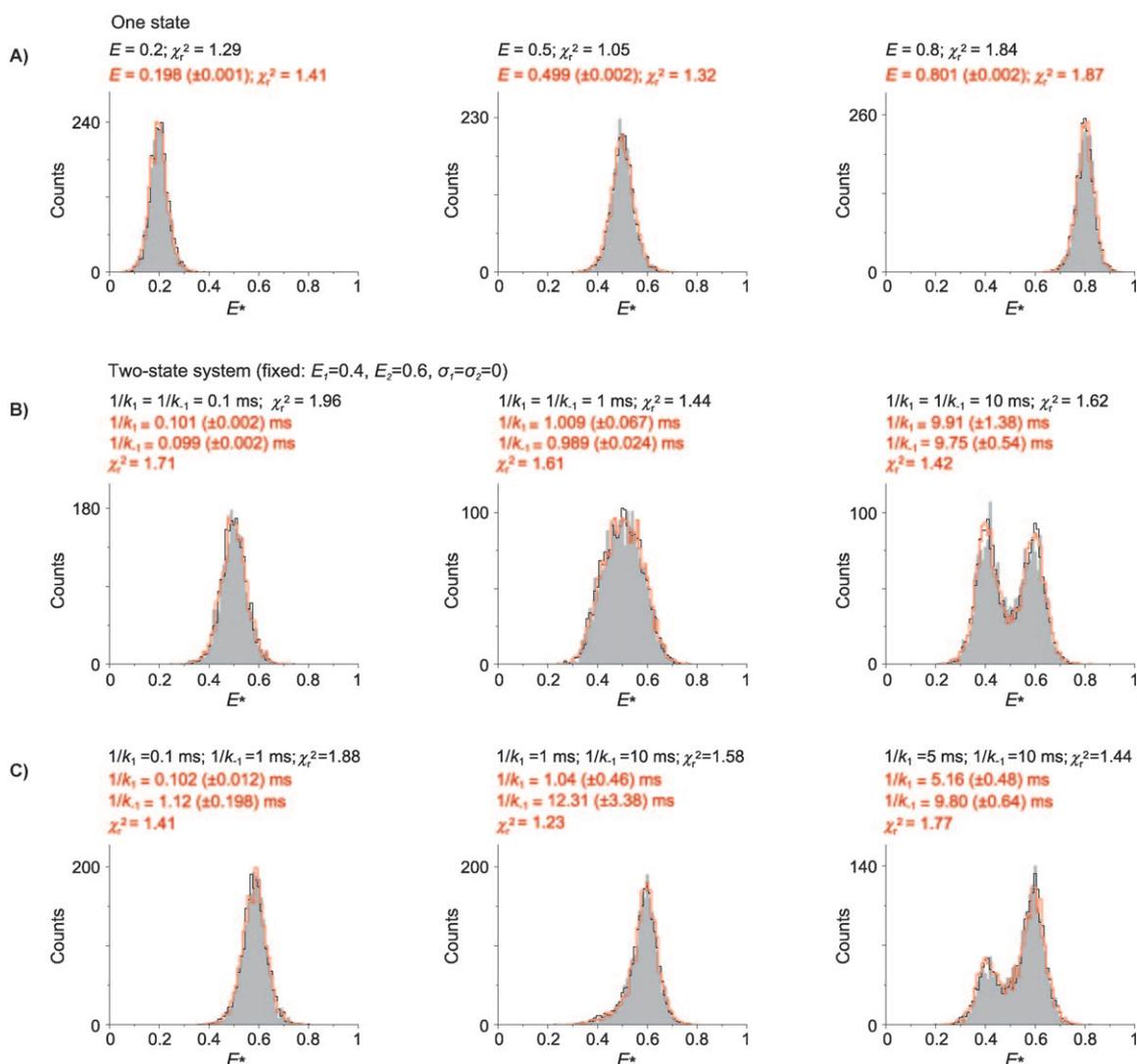
Figure 2. Comparison of Monte Carlo simulations (light gray histograms), PDA fitting results (red lines), and PDA predictions given the original kinetic parameters (black lines). The original simulation parameters and the χ_r^2 values obtained from PDA predictions made using those parameters are displayed in black above each panel; the parameters recovered through fitting and their χ_r^2 values are displayed in red; some parameters are fixed during fitting to reduce the number of free parameters. The average diffusion time of the single-molecule simulation was 1 ms. A) Simulations of a one-state system for low (left, $E = 0.2$), intermediate (middle, $E = 0.5$), and high (right, $E = 0.8$) FRET values. B,C) Simulations of a two-state system ($E_1 = 0.4$, $E_2 = 0.6$, $\sigma_1 = \sigma_2 = 0$). The rates of interconversion are as displayed on the respective panels. The simulations in (B) have equal forward and backward rates ($k_1 = k_{-1}$), while those in (C) have faster forward rates ($k_1 > k_{-1}$).

2.7. Conformational Dynamics in a DNA Hairpin

We next used PDA to analyze the conformational fluctuations in a DNA hairpin (Figure 4 B,C). DNA hairpins interconvert stochastically between a folded and an unfolded conformation (Figure 4 B). In addition to the broadening beyond shot noise caused by this interconversion, each of the two conformations may be assumed to demonstrate broadening beyond shot-noise similar to a simple, double-stranded DNA. Assuming a two-state model, with standard deviations of 2 Å in each state, the experimental FRET distribution was best fitted with folding and unfolding rates of 310($\pm$20) and 280($\pm$30) s^{-1}, which gave a total reaction time $\tau_R = 1/(k_{fold}+k_{unfold}) = 1.7(\pm 0.1)$ ms. This result is in good agreement with the 0.5–1 ms reaction time obtained in previous studies on the same hairpin using correla-

tion-based analyses.[4,5] The fluctuations are unlikely to be due to photophysical processes, since these would have been visible as an additional process in our static DNA control (see Figure 4 A), for which we used the same excitation intensities and fluorophores as in the dynamic hairpin. Moreover, triplet-state fluctuations for our acceptor (ATTO647N) are not significant (only a 6% triplet fraction with a lifetime of $\approx$ 2 µs), and therefore do not affect the timescale of the dynamics observed for the hairpin DNA.

We note that although the source of broadening in double-stranded DNAs is unknown, this method produced a good fit to the experimental hairpin data even when we increased or decreased the broadening by 20%, which suggests that it is robust to such sources of broadening (see Supporting Information, Figure S-5).
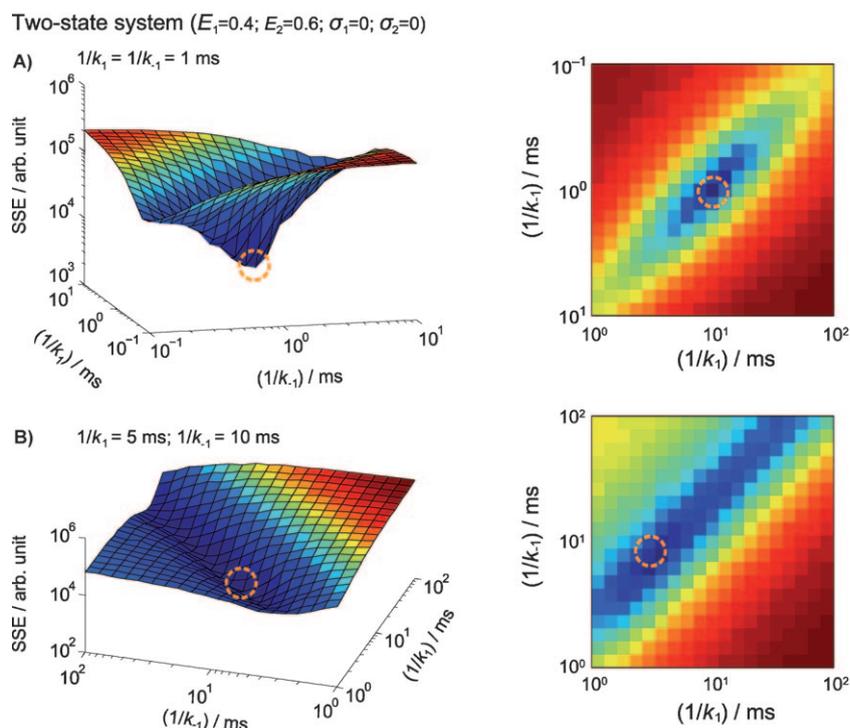
Figure 3. Left: Parameter minimization surface, where each point on the surface represents the goodness of fit (represented as sum of squared error, SSE; dark blue: low SSE, red: high SSE) given the combination of the parameter (X, Y). Here, we illustrate how parameters were optimized using simulated data from Figure 2. The combinations of (X, Y) that correspond to the original simulation parameters are marked by the orange circles. Right: 2D projection of the parameter minimization surface. A) The parameter minimization surface for a two-state system with $1/k_1 = 1/k_{-1} = 1$ ms forms a deep well around the expected rates. B) For a two-state system with unequal rates, $1/k_1 = 5$ ms and $1/k_{-1} = 10$ ms, the parameter space forms a flat and elongated basin around the expected parameter combination. Such a topographic feature gives rise to many, equally good, fitting results along the floor of the basin.

An important feature of PDA is its ability to test the consistency of the observed distribution with a particular kinetic model ("hypothesis testing"). While our fit was good ($\chi_r^2 = 1.17$), the predicted distribution also shows signs of systematic deviation from the data (Figure 4 C); for instance, the two-state model prediction contains a small peak at $E^* = 0.55$ not observed in the actual data. This could be due to more complex kinetic schemes, as identified in recent biophysical studies of DNA hairpins.[36,37] While some of these behaviors are expected to occur on timescales too short to be detected by PDA (e.g. multiple loop orientations in the open conformation), longer-lived states (e.g. mismatched closed conformations) may be addressable via PDA by using more complex models and appropriate controls.

2.8. Conformational Dynamics in the Klenow Fragment (KF) of *E. coli* DNA Polymerase I

DNA polymerase I is an important enzyme involved in DNA replication and repair.[38] One of the important characteristics of this enzyme is its remarkable fidelity, which is in part due to a series of noncovalent transitions that precede the chemical step of phosphoryl transfer and serve as kinetic checkpoints that reject inappropriate substrates early in the reaction path-

way.[39] An important conformational change is the "fingers-closing" transition, inferred from cocrystal structures, in which the addition of the correct complementary deoxyribonucleotide (dNTP) to a polymerase–DNA (Pol–DNA) binary complex results in a transition from an open to a closed conformation, thus forming a snug binding pocket around the nascent base pair (Figure 5 A).

Our previous studies on a bacterial DNA Pol I revealed millisecond-timescale dynamics in the unliganded Pol enzyme.[6] Such rapid fluctuations were not previously anticipated based on the available crystal structures.[40–43] We postulated that these rapid motions play an important role in the fast rejection of the large number of incorrect nucleotide substrates encountered by a DNA polymerase *in vivo*. The timescale and the nature of the conformational fluctuations made this system an attractive experimental system for PDA.

To study conformational dynamics in the Klenow fragment (KF) of DNA polymerase I (hereafter Pol I), we analyzed smFRET data obtained using a Pol I molecule site-specifically labeled with two fluorophores, Cy3B as the FRET donor and ATTO647N as the FRET acceptor. Characterization of the labeled enzyme has been described elsewhere.[6] The placement of the fluorophores allowed the study of the finger-closing motion through a large change in FRET (from $E^* \approx 0.5$ in the open state to $E^* \approx 0.7$ in the closed state).

The E^* histogram of unliganded Pol I exhibits a broad and flat distribution, centered on $E^* \approx 0.6$ (Figure 5 B). We have shown previously that such a distribution cannot be accounted for either by a single FRET distribution with $E^* = 0.6$, or a superposition of two shot-noise-limited FRET distributions.[6] We proposed that such broadening originated from interconversions between the open and closed states, in the form of a classical two-state kinetic model.

We used a two-state system model to fit the E^* histogram of unliganded Pol I. To reduce the number of free parameters, we fixed the mean E^* and the E^* standard deviation of the open and closed states using the PDA fitting results of the FRET histograms of the Pol–DNA binary complex (which favors the open state) and the Pol–DNA–dNTP ternary complex (which favors the closed state; see Supporting Information, Figure S-6).
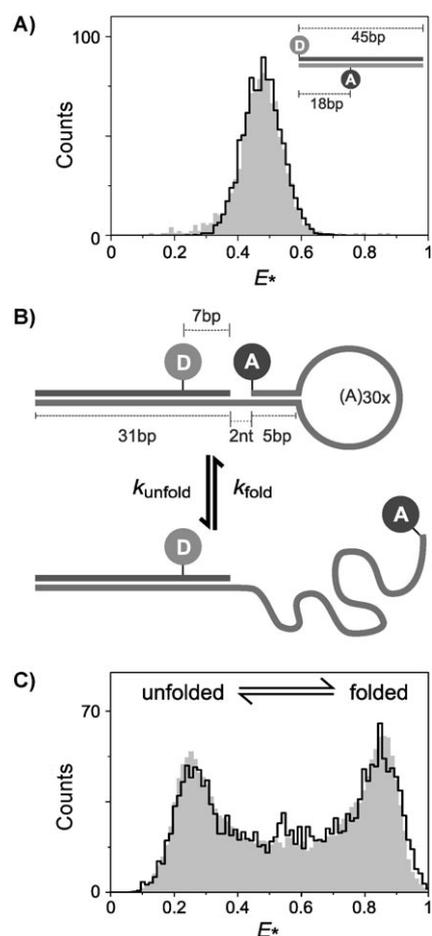
Figure 4. Comparison of experimental FRET histograms (light gray) and PDA predictions (black lines) for double-stranded DNA and DNA hairpin. A) The E^* histogram of the $T_{1-Cy3B}, B_{18-ATTO647N}$ DNA (schematic shown as inset; bp = base pair) was best fitted using a PDA prediction that incorporated a Gaussian distribution of FRET values [$E = 0.48$, $\sigma_E = 0.05(\pm 0.002)$, $\chi_r^2 = 1.22$]. The value of $\sigma_E = 0.05$ translates to a distance change of approximately 2 Å (given $R_0 = 67$ Å, determined as in ref. [48]). B) Schematic of a DNA hairpin fluctuating between the folded (high FRET) and unfolded (low FRET) states. C) The experimental E^* histogram of the hairpin in a buffer containing MgCl$_2$ (5 mM) exhibits a bimodal distribution that can be fitted to a two-state model with folding and unfolding rates of 310($\pm$21) and 278($\pm$29) s^{-1} ($E_1 = 0.25$, $E_2 = 0.86$, $\sigma_1 = \sigma_2 = 0.05$, $\chi_r^2 = 1.17$). The experimental E^* histogram was dithered prior to fitting to remove binning artifacts (see ref. [1]).

The parameter-minimization surface in Figure 5C illustrates the combinations of the opening and closing rates that achieve the best fit (i.e. minimize the SSE). The best fit is obtained for opening and closing rates of 189($\pm$23) and 191($\pm$25) s^{-1}, respectively; the minimum is clearly identifiable and localized (as in Figure 2A). These rates agree very well with the 166 s^{-1} opening and closing rates obtained using fluorescence correlation spectroscopy (FCS).[6] The excellent agreement between the actual and PDA-predicted histograms (Figure 5B) suggests that a simple, two-state kinetic model is both adequate and reasonable in accounting for the dynamics of unliganded KF.

We also used PDA to characterize the Pol–DNA binary complex. We found that the E^* distribution could not be accounted for by a simple model that assumes two static non-interconverting FRET populations (see Supporting Information, Fig-

ure S-7C). This indicates that the binary complex interconverts between the open and closed conformations, albeit very slowly (Supporting Information, Figure S-7A; closing rate = 59($\pm$5) s^{-1}, opening rate = 111($\pm$13) s^{-1}). Furthermore, the Pol–DNA–dNTP ternary complex exhibits a FRET histogram that primarily populates the closed conformation. The 5:1 ratio between the closing and opening rates obtained using PDA further confirms our previous observations (Supporting Information, Figure S-7B).[6, 44] However, the PDA method gave an opening rate of 101($\pm$13) s^{-1}, approximately threefold faster than the opening rate estimated using values from ensemble stopped-flow experiments.[6, 44] Such a difference may be due to the different labeling scheme employed previously,[44] which may in turn report slightly differently on conformational transitions of the fingers subdomain. The rate difference may also mean that a simple two-state model is not sufficient to describe the transitions in the ternary complex. The PDA-based approach should be important for testing and validating conformational models for this system, especially once higher-resolution single-molecule data become available.

2.9. Comparison with Correlation-Based Analysis

Both correlation-based analysis and PDA can provide information about the conformational transitions of biomolecules; however, since the nature of the processes measured is slightly different, the kind of information available from each method is also different. A correlation-based analysis relies on statistical analysis of fluorescence fluctuations above and beyond the background noise. To achieve an optimal signal-to-noise ratio, FCS experiments are rarely performed at the single-molecule level (i.e. at $\approx$100 pM fluorescent analyte given a femtoliter detection volume). Instead, typical FCS experiments are performed in the 1–100 nM range, thus foregoing some of the benefits of single-molecule methods. The timescale of fluorescence fluctuations is recovered by fitting a mathematical description of the process to the temporal correlation curve.

The main source of fluorescence fluctuations in freely diffusing molecules is molecular diffusion through the confocal volume. It is often difficult to determine the rates of transition that occur at a timescale similar to that of diffusion ($\tau_r \approx T_D$; where τ_r is the rate of transition); the same problems are encountered when the rates of the process of interest overlap with other processes, such as triplet-state formation. Extraction of reliable rates for such conditions is difficult and requires careful characterization of the optical setup and the interfering processes.

PDA, on the other hand, uses the distribution of photon statistics obtained from single-molecule FRET experiments. Thus, the PDA method is most sensitive to conformational transitions with rates similar to the timescale of diffusion ($\tau_r \approx T_D$) as such rates provide the largest signal, that is, the largest deviation from a simple shot-noise-limited distribution (see Supporting Information, Figure S-3). As a result, the PDA approach is less sensitive to very fast or very slow dynamics, wherein the resulting FRET distributions approach the shot-noise limit. The effects of bleaching, blinking, and triplet-state behaviors can

be important at times, but there are many methods to identify them (e.g. FCS), minimize them experimentally (e.g. by using special buffer components[45,46]) or, in the case of PDA, remove them during data analysis.[1]

Another important benefit of PDA over correlation-based analysis is the use of a Monte Carlo approach in determining the $P(\vec{\tau}|T)$ distribution. This approach allows for the incorporation of arbitrarily complex kinetic models, which cannot always be represented analytically, in contrast to the interpretation of FCS curves, which require fitting to well-defined mathematical expressions.

3. Conclusions

We have demonstrated a straightforward numerical implementation of PDA for molecules undergoing conformational dynamics. In particular, we use a generalized PDA method that allows for the incorporation of any arbitrary kinetic model of FRET dynamics. This was accomplished by incorporating the experimental joint photon count-burst duration distribution, $P(T,F)$, into the analysis, to calculate the dwell-time distribution of the molecule in each state using Monte Carlo methods.

In practice, it may not be possible to determine the true underlying kinetic model using only the PDA approach. Translating a complex molecular mechanism, with multiple states and rates, into a one-dimensional FRET distribution will inevitably lead to a degeneracy problem where more than one set of parameters adequately describe the same final FRET distribution; indeed, this problem arose previously in the work of Kalinin et al.[2] when attempting to recover underlying distributions of static states from E^* distributions. As such, PDA is best used for hypothesis testing, and rejecting, of plausible kinetic models, rather than deriving an underlying model without any prior input. The method should also prove useful for extracting parameters (rates, energies, etc.) by fitting the data to a preexisting model.

We validated this approach by comparing the PDA predictions to FRET histograms obtained from simulations, as well as single-molecule FRET experiments on double-stranded DNA and a dynamic DNA hairpin. Consistent with previous reports,[1,3] we observed a broadening of the FRET distribution beyond the shot-noise limit for double-stranded DNA. Applica-
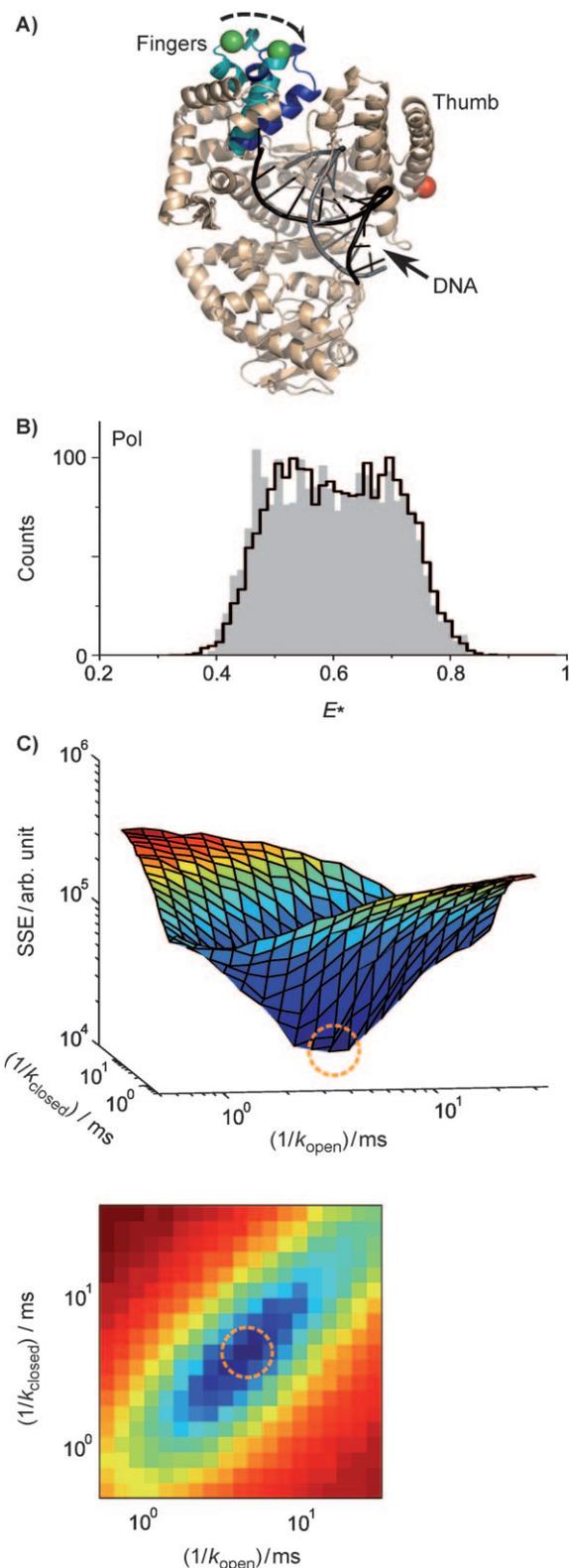
Figure 5. Use of PDA to study the conformational dynamics in unliganded DNA polymerase I. A) The crystal structures of open Pol–DNA binary complex (PDB file 1L3U) and closed Pol–DNA–dNTP ternary complex (PDB file 1LV5) are illustrated using structural data from *Bacillus stearothermophilus* DNA polymerase,[49] a close homologue of Pol I (KF). The structure shown is a superposition of the binary and ternary complexes showing the closing of the fingers subdomain upon the addition of matched dNTP to the Pol–DNA binary complex. The direction of movement of the fingers subdomain from open to closed is denoted by the dashed arrow. The α-carbon backbone of the protein is shown in beige, except for the mobile segment of the fingers subdomain, which is shown in teal in the binary complex and dark blue in the ternary complex. The DNA is shown in gray. The β carbon atoms of the two side chains used as fluorophore attachment sites are shown in space-filling representation; residue 744 (Pol residue numbers) in green (labeled with Cy3B), and residue 550 in red (labeled with ATTO647N). B) The experimental E^* histogram obtained from unliganded Pol exhibits a broad distribution that can be fitted using a two-state system (black line) with opening and closing rates of 189($\pm$23) and 191($\pm$25) s^{-1}, respectively ($\chi_r^2 = 1.94$). C) A parameter minimization surface (top) and its 2D projection (bottom) for panel (B) [dark blue: low SSE (good fit); red: high SSE (bad fit)]. The minimum SSE (best fit; shown as orange circles) is achieved when $1/k_{open} = 1/k_{closed} \approx 4$–$6$ ms.

tion of this method to the DNA hairpin uncovered millisecond folding/unfolding rates, in good agreement with rates previously reported for the same hairpin using different experimental approaches.[4,5]

Although our hairpin data were consistent with a two-state model via the chi-squared metric, we also observed a systematic deviation between our data and the model. Recent studies of DNA hairpin dynamics have identified putative intermediates in the folding pathway,[36,37] which could explain this discrepancy. However, we caution that the detection of such intermediates by PDA should involve the use of novel experimental controls, and not simply an increase in the number of free parameters. The number of fitting parameters increases nearly quadratically with the number of states [see the rate matrix Q in Eq. (11)]. As a result, a "simple" three-state model would require at least nine fitted parameters, thus exacerbating the parameter degeneracy problem. To circumvent this "parameter explosion," the number of parameters must be kept to a minimum by making reasonable simplifying assumptions and, where possible, determining the values of these parameters through independent controls.

It is instructive to compare the PDA results with results obtained using correlation-based approaches, as they provide complementary information. Once a correlation-based or other method has established that the FRET heterogeneity arises from dynamic processes, a PDA-based model can be used to extract rates, and hypothesis-test a proposed model. In contrast to FCS and other correlation approaches, the PDA-based approach is most sensitive to conformational transitions around the timescale of diffusion. Moreover, artifacts such as the presence of aggregates, which can strongly skew correlation-based analyses, do not significantly affect FRET histograms and can often be removed with data-filtering methods.

Finally, we used this method to analyze the conformational dynamics in unliganded DNA polymerase I (KF). Using a simple two-state kinetic model, we recovered interconversion rates consistent with those obtained in a previous publication.[6] Furthermore, the goodness of the fit itself (χ_r^2 near 1) suggests that this simple model is sufficient to explain the dynamics of the unliganded polymerase and to provide rate information about conformational transitions in the case of binary and ternary complexes. Regarding the ternary complex, we observed rates slightly different from those reported previously, due either to differences in labeling strategies, or to the inadequacy of a simple two-state model to explain the transitions in the ternary complex; the latter intriguing possibility may be due to a complicated conformational landscape for the ternary complex.

Our results demonstrate the utility of PDA in modeling and quantifying biomolecular dynamics. We expect that the PDA-based approach will complement other approaches, such as FCS, to enhance our understanding of the dynamics of biological systems. We note that this relatively simple analysis can also be performed retrospectively on existing single-molecule FRET data, since the required inputs—the joint distribution of photon counts and burst durations—are typically preserved in single-molecule experimental data.

Experimental Section

DNA: Amino-modified oligonucleotides (IBA, Germany) were labeled using NHS-conjugated fluorophores according to the manufacturer's instructions, and purified either on a reverse-phase C18 fast protein liquid chromatography (FPLC) column (µRPC C2/C18, GE Healthcare, UK) or via denaturing PAGE purification. Where necessary, labeled, single-stranded DNA samples were annealed in hybridization buffer [Tris (50 mM, pH 8.0), EDTA (1 mM), NaCl (500 mM)]. All DNA sequences are listed in the Supporting Information, Figure S-1.

DNA Polymerase I (KF): The expression, purification, and labeling of our doubly labeled Pol I have been described elsewhere.[6,44] We used a Pol I derivative labeled at position 550 with ATTO647N and at position 744 with Cy3B. The specificity of the labeling orientation was ≈88%. Labeled proteins were stored at −20°C in Tris–HCl (50 mM, pH 7.5), dithiothreitol (DTT; 1 mM), and glycerol [40% (v/v)]. The extent of labeling, calculated from UV spectra, was ≥70%.

Single-Molecule Experiments: Single-molecule measurements were performed on an alternating laser excitation (ALEX) microscope as described elsewhere.[4,6,15] The excitation powers were 400 and 60 µW for the 532 and 638 nm lasers, respectively. Samples were analyzed at a concentration of 50–100 pM. The double-stranded DNA sample was measured in Tris–glycine (TG) buffer (Tris (25 mM), glycine (200 mM), pH 8.0), while the hairpin sample was measured in TG buffer with additional MgCl₂ (5 mM). The Pol I sample was measured in 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)–NaOH (40 mM, pH 7.3), MgCl₂ (10 mM), DTT (1 mM), bovine serum albumin (100 µg mL^{-1}), glycerol [5% (v/v)], and mercaptoethylamine (1 mM).

Simulations of smFRET Experiments: Monte Carlo simulations of diffusing molecules in confocal microscopy have been described.[1,47] We simulated molecules diffusing through a 3D Gaussian excitation/detection volume. For each type of molecule, we defined its concentration, diffusion coefficient, a set of rates describing the interconversion between the open and closed states, and a set of fluorophore-specific parameters (stoichiometry, molecular brightness, interfluorophore distance). Simulations (written in C++) were performed using a 1 µs time step, much faster than diffusion (occurring at the ≈3 ms timescale) and laser alternation due to ALEX (10 kHz). The results were analyzed as with the experimental data.

Data Analysis: Single-molecule data (from experiments or simulations) were analyzed using custom software written in LabVIEW (National Instruments, USA). Fluorescent bursts were separated from the background by using published burst-search algorithms.[1,4] For each detected burst, we calculated a range of burst statistics, including E^* (apparent FRET), S (relative stoichiometry), total photon count due to excitation of the donor, and burst duration. The PDA algorithm was implemented in MATLAB (MathWorks, USA).

[1] E. Nir, X. Michalet, K. Hamadani, T. Laurence, D. Neuhauser, Y. Kovchegov, S. Weiss, *J. Phys. Chem. B* **2006**, *110*, 22103–22124.

[2] S. Kalinin, S. Felekyan, A. Valeri, C. Seidel, *J. Phys. Chem. B* **2008**, *112*, 8361–8374.

[3] M. Antonik, S. Felekyan, A. Gaiduk, C. Seidel, *J. Phys. Chem. B* **2006**, *110*, 6970–6978.

[4] Y. Santoso, A. Kapanidis, *Anal. Chem.* **2009**, *81*, 9561–9570.

[5] M. Wallace, L. Ying, S. Balasubramanian, D. Klenerman, *J. Phys. Chem. B* **2000**, *104*, 11551–11555.

[6] Y. Santoso, C. Joyce, O. Potapova, L. Le Reste, J. Hohlbein, J. Torella, N. Grindley, A. Kapanidis, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 715–720.

[7] P. Selvin, T. Ha, *Single Molecule Techniques: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, **2007**.

[8] W. Moerner, M. Orrit, *Science* **1999**, *283*, 1670–1676.

[9] S. Weiss, *Science* **1999**, *283*, 1676–1683.

[10] X. Xie, J. Trautman, *Annu. Rev. Phys. Chem.* **1998**, *49*, 441–480.

[11] T. Ha, *Biochemistry* **2004**, *43*, 4055–4063.

[12] M. Dahan, A. Deniz, T. Ha, D. Chemla, P. Schultz, S. Weiss, *Chem. Phys.* **1999**, *247*, 85–106.

[13] A. Deniz, M. Dahan, J. Grunwell, T. Ha, A. Faulhaber, D. Chemla, S. Weiss, P. Schultz, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 3670–3675.

[14] M. Margittai, J. Widengren, E. Schweinberger, G. Schröder, S. Felekyan, E. Haustein, M. König, D. Fasshauer, H. Grübmuller, R. Jahn, C. Seidel, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 15516–15521.

[15] A. Kapanidis, E. Margeat, S. Ho, E. Kortkhonjia, S. Weiss, R. Ebright, *Science* **2006**, *314*, 1144–1147.

[16] L. Ying, M. Wallace, S. Balasubramanian, D. Klenerman, *J. Phys. Chem. B* **2000**, *104*, 5171–5178.

[17] P. Rothwell, S. Berger, O. Kensch, S. Felekyan, M. Antonik, B. M. Wöhrl, T. Restle, R. S. Goody, C. Seidel, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1655–1660.

[18] I. Gopich, A. Szabo, *J. Chem. Phys.* **2005**, *122*, 014707.

[19] O. Coban, D. Lamb, E. Zaychikov, H. Heumann, G. Nienhaus, *Biophys. J.* **2006**, *90*, 4605–4617.

[20] S. Kalinin, S. Felekyan, M. Antonik, C. Seidel, *J. Phys. Chem. B* **2007**, *111*, 10253–10262.

[21] D. Majumdar, I. Smirnova, V. Kasho, E. Nir, X. Kong, S. Weiss, H. Kaback, *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 12640–12645.

[22] K. Henzler-Wildman, D. Kern, *Nature* **2007**, *450*, 964–972.

[23] S. Brasselet, E. Peterman, A. Miyawaki, W. Moerner, *J. Phys. Chem. B* **2000**, *104*, 3676–3682.

[24] T. Laurence, Y. Kwon, E. Yin, C. Hollars, J. Camarero, D. Barsky, *Biophys. J.* **2007**, *92*, 2184–2198.

[25] G. Li, M. Levitus, C. Bustamante, J. Widom, *Nat. Struct. Mol. Biol.* **2004**, *12*, 46–53.

[26] T. Torres, M. Levitus, *J. Phys. Chem. B* **2007**, *111*, 7392–7400.

[27] C. Eggeling, J. Fries, L. Brand, R. Günther, C. Seidel, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 1556–1561.

[28] P. Schwille, S. Kummer, A. Heikal, W. Moerner, W. Webb, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 151–156.

[29] L. Brand, C. Eggeling, C. Zander, K. Drexhage, C. Seidel, *J. Phys. Chem. A* **1997**, *101*, 4313–4321.

[30] N. Lee, A. Kapanidis, Y. Wang, X. Michalet, J. Mukhopadhyay, R. Ebright, S. Weiss, *Biophys. J.* **2005**, *88*, 2939–2953.

[31] A. Kapanidis, T. Laurence, N. Lee, E. Margeat, X. Kong, S. Weiss, *Acc. Chem. Res.* **2005**, *38*, 523–533.

[32] J. P. Torella, Y. Santoso, S. J. Holden, J. Hohlbein, A. N. Kapanidis, *unpublished results*.

[33] S. McKinney, C. Joo, T. Ha, *Biophys. J.* **2006**, *91*, 1941–1951.

[34] G. Arfken, *The Method of Steepest Descents*, 3rd ed., Academic Press, Orlando, **1985**.

[35] B. Schuler, E. Lipman, W. Eaton, *Nature* **2002**, *419*, 743–747.

[36] A. Van Orden, J. Jung, *Biopolymers* **2008**, *89*, 1–16.

[37] J. Jung, A. Van Orden, *J. Am. Chem. Soc.* **2006**, *128*, 1240–1249.

[38] C. Joyce in *Encyclopedia of Biological Chemistry*, Elsevier, San Diego, **2004**, pp. 720–725.

[39] C. Joyce, S. Benkovic, *Biochemistry* **2004**, *43*, 14317–14324.

[40] J. Kiefer, C. Mao, C. Hansen, S. Basehore, H. Hogrefe, J. Braman, L. Beese, *Structure* **1997**, *5*, 95–108.

[41] S. Korolev, M. Nayal, W. Barnes, E. Cera, G. Waksman, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 9264–9268.

[42] M. Sawaya, H. Pelletier, A. Kumar, S. Wilson, J. Kraut, *Science* **1994**, *264*, 1930–1935.

[43] M. Franklin, J. Wang, T. Steitz, *Cell* **2001**, *105*, 657–667.

[44] C. Joyce, O. Potapova, A. Delucia, X. Huang, V. Basu, N. Grindley, *Biochemistry* **2008**, *47*, 6103–6116.

[45] J. Vogelsang, R. Kasper, C. Steinhauer, B. Person, M. Heilemann, M. Sauer, P. Tinnefeld, *Angew. Chem.* **2008**, *120*, 5545–5550; *Angew. Chem. Int. Ed.* **2008**, *47*, 5465–5469.

[46] I. Rasnik, S. McKinney, T. Ha, *Nat. Methods* **2006**, *3*, 891–893.

[47] T. Wohland, R. Rigler, H. Vogel, *Biophys. J.* **2001**, *80*, 2987–2999.

[48] V. Mekler, E. Kortkhonjia, J. Mukhopadhyay, J. Knight, A. Revyakin, A. Kapanidis, W. Niu, Y. Ebright, R. Levy, R. Ebright, *Cell* **2002**, *108*, 599–614.

[49] S. Johnson, J. Taylor, L. Beese, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3895–3900.