Biophysical Journal, Volume 98

Supporting Material

Denoising Single-Molecule FRET Trajectories with Wavelets and Bayesian Inference

J. Nick Taylor, Dmitrii E. Makarov, and Christy F. Landes

Biophysical Journal, Volume 98

Supporting Material

Denoising Single-Molecule FRET Trajectories with Wavelets and Bayesian Inference

J. Nick Taylor, Dmitrii E. Makarov, and Christy F. Landes

Part S1: Regions in smFRET Time Trajectories

Recall that two-channel smFRET experiments result in two data vectors containing detected numbers of acceptor and donor photons in discrete time steps. We term these numbers of detected photons as N_A for acceptor and N_D for donor. Also recall that these vectors contain three distinct regions, as illustrated in Fig. S1. The utilization of photobleaching events on both fluorophores involved in these measurements allows for the collection of local background signals for each single molecule. Poissonian shot-noise in each of the background signals arises, producing a Poisson distribution around each channel's mean number of photons. We define the mean and variance of these distributions of on detection channel α as b_α ($\alpha = A, D$).

Imperfect separation of donor and acceptor photons by a dichroic mirror leads to the observation of donor-emitted photons on the acceptor detector. This is referred to as crosstalk, and we produce the crosstalk region by observing a molecule after photobleaching of the acceptor but before that of the donor. Due to the absence of acceptor fluorophore emission in this region, photons detected above background levels on the acceptor channel are assumed to be emitted by the donor fluorophore. The number of crosstalk photons is thus estimated as $p_x = N_A - b_A$. Because the crosstalk photons are emitted by the donor but counted as acceptor photons, the actual number p_D of donor-emitted photons in the crosstalk region is different from the number of observed photons N_D, and is estimated as

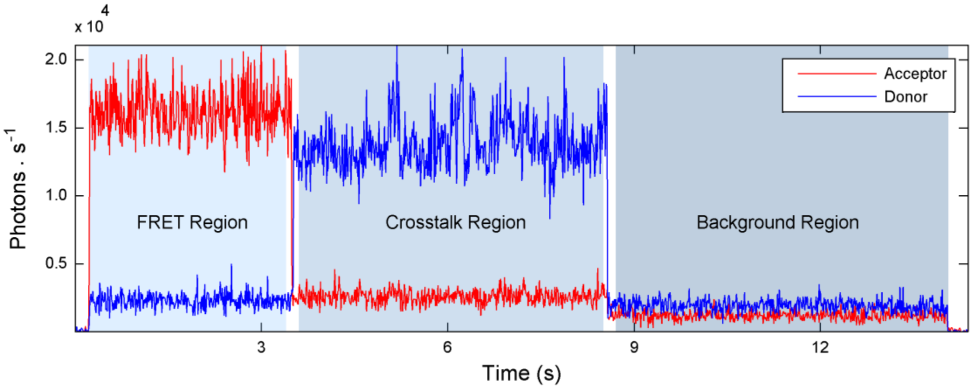

Figure S1. Regions in an smFRET time trajectory. The region to the far left shows the FRET region, the center region shows the crosstalk region, and to the right is the background region.

$$p_D \equiv N_D - b_D + p_x. \qquad (S1)$$

The ratio of the number of crosstalk photons p_x to the number of photons emitted by the donor fluorophore p_D is calculated at every time step in the region, and the mean of these values is taken to be the crosstalk parameter

$$x \equiv \langle p_x / p_D \rangle. \tag{S2}$$

To estimate the efficiency of energy transfer in the FRET region of an smFRET trajectory, again we must correct for the difference between the numbers N_D and N_A of detected photons, respectively, on the donor and the acceptor channels, and the numbers n_D and n_A of photons actually emitted by the respective fluorophores. The latter can be estimated as

$$n_D \equiv N_D - b_D + n_x \tag{S3a}$$

$$n_A \equiv N_A - b_A - n_x, \tag{S3b}$$

where N_α and b_α (α = A or D) are defined previously, and n_x is the number of crosstalk photons calculated at each time step using the relation

$$n_x = xn_D = \frac{x}{1-x}(N_D - b_D). \tag{S4}$$

The energy transfer efficiency, E, is calculated at each time step in the FRET region using the standard relationship

$$E = \frac{n_A}{n_A + n_D}, \tag{S5}$$

In the context of the present discussion, we refer to the efficiency E as “observed efficiency”. We determined the correction factor commonly known as γ (1, 2) to be within error of unity for our apparatus using the fluorophores Cy3 and Cy5.

Part S2: Photoblink Removal in smFRET Time Trajectories

Here we describe the Bayesian photoblink detection algorithm in detail. Recall that we need the conditional probability distributions of the acceptor photon count N_A given two alternatives, the “blink” hypothesis (B), and the “no blink” hypothesis (NB). A Poisson process describes photon emission in each case, but their properties differ. Specifically, in the absence of a blink the distribution of N_A can be approximated by

$$P(N_A|NB) = \frac{1}{(2\pi\mu_{NB})^{\frac{1}{2}}} \exp\left[-\frac{(N_A - \mu_{NB})^2}{2\mu_{NB}}\right], \tag{S6}$$

where $\mu_{NB} = \langle N_A \rangle$ is the calculated mean of detected acceptor photons for the trajectory under consideration. In writing Eq. S6, we have replaced the Poisson distribution by a Gaussian one whose mean and variance are both equal to μ_{NB}. This approximation is valid assuming $\mu_{NB} >> 1$, which is always the case for typical time steps used (~10 ms).

Similarly, the conditional probability distribution for N_A given the “blink” hypothesis can be written as a normal distribution. To calculate its width and its mean we note that the mean number of detected acceptor counts arising from such a situation will be the same that arises in the crosstalk region. We express this mean as

$$\mu_B = \langle b_A \rangle + x\langle n_D \rangle, \tag{S7}$$

where $\langle b_A \rangle$ is the mean of the acceptor background signal, x is the crosstalk value as calculated from Eq. S2, and $\langle n_D \rangle$ is the mean of donor fluorophore-emitted photons over the FRET region as calculated from Eq. S3a. Again, we note that the width of this distribution is induced by shot-noise, and thus assume the distribution's variance to be equal to its mean, μ_B. These assignments result in the following expression for the prior probability distribution of N_A given the "blink" hypothesis:

$$P(N_A|B) = \frac{1}{(2\pi\mu_B)^{\frac{1}{2}}} \exp\left[-\frac{(N_A - \mu_B)^2}{2\mu_B}\right]. \tag{S8}$$

The posterior probability of the blink hypothesis B given the observation N_A is now given by Bayes' theorem (3):

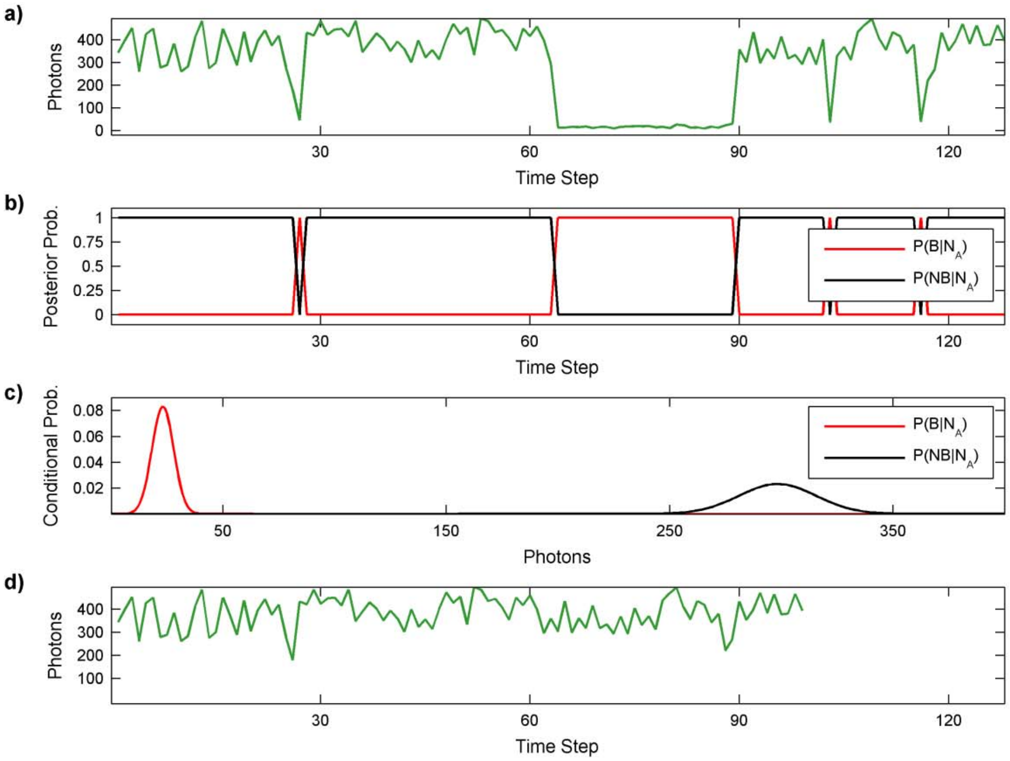

Figure S2. Photoblinks in an smFRET time trajectory. a) An acceptor photon count trajectory that contains photoblinks on long and short time-scales. b) The calculated posterior probability of the occurrence of a photoblink (4) and the absence of a photoblink (black) at each time step. c) The prior probability distributions of a photoblink (black, left) and the absence of a photoblink (red, right). d) The same acceptor photon trajectory with photoblinks removed.

$$\mathrm{P(B|N_A)} = \frac{\mathrm{P(N_A|B)P(B)}}{\mathrm{P(N_A|B)P(B) + P(N_A|NB)P(NB)}}, \qquad \text{(S9)}$$

where P(B) is the prior probability of hypothesis B and P(NB) is the prior probability of the no blink hypothesis NB.

To evaluate Eq. S9 we need the probabilities P(B) and P(NB). These probabilities represent the probability of the hypothesis H (= B or NB) being true *prior* to taking the observation N_A into consideration. We do not assume to know these values in advance, but instead obtain self-consistency through an iterative process. Initial guesses of P(B) = 0.001 and P(NB) = 0.999 are input, each posterior probability is calculated at each time step, and those having Bayes factors – i.e., the ratio of $P(B|N_A)$ to $P(NB|N_A)$ – greater than 2 are labeled as blinks. The fraction of time steps labeled as photoblinks is defined as the blink fraction, f_B. This value is compared with the value of P(B) at each iteration. Unless a deviation between these values of less than 5% is obtained, the substitutions $\mathrm{P(B)} = f_B$, and $\mathrm{P(NB)} = 1 - f_B$ are made, and the process repeats until this condition is met.

Fig. S2 illustrates the process of the identification and removal of photoblinks from an acceptor fluorophore-emitted time trajectory. Fig. S2a depicts a trajectory that contains both long and short time-scale photoblinks. Fig. S2b shows the posterior probabilities evaluated at each time step from the probability distribution of the blink hypothesis B. The conditional probability distributions $P(N_A|B)$ and $P(N_A|NB)$ are shown in Fig. S2c. Lastly, the acceptor fluorophore-emitted photon counts after removal of time steps that contain photoblinks are shown in Fig. 1d. The time steps that are removed Bayes factors greater than 2, which is chosen because the number of data points removed becomes, to a large extent, constant for Bayes factors larger than 2. Typical Bayes factors obtained for time steps containing photoblinks are on the order of 10^{10}.

Part S3: Details of the Denoising Algorithm

Recall that, in the context of smFRET time trajectories, the trajectory consists of two data vectors containing detected numbers of photons in discrete time steps. Here we consider only the acceptor photon trajectory $\mathbf{N}_A$ (= $N_A(0)$, $N_A(\Delta t)$,…), in discrete time steps Δt. We recall that this vector is written in the form

$$\mathbf{N}_\mathrm{A} = \mathbf{S}_\mathrm{A} + \sigma\mathbf{Z}, \qquad \text{(S10)}$$

where, at each time step Δt, **Z** is a Gaussian white noise component, and each element of **Z** is independently and identically distributed on a normal distribution with mean 0 and variance 1, σ is a known noise level, and $\mathbf{S}_A$ is the “true” signal that we wish to recover. Similarly to the smoothing method described in the main text, we accomplish the recovery of the true signal $\mathbf{S}_A$ in three steps: (1) transform the observed data $\mathbf{N}_A$ into the wavelet domain, (2) suppress the presumed noise component of the signal, and (3) invert the wavelet transform to obtain the denoised signal. A key assumption of most denoising schemes is that the noise is additive – that the strength of the noise is independent of the signal. In the case of smFRET experiments, the shot-noise strength is dependent on the brightness of the fluorophores. Here however, we approximate the noise strength with its average value.

The first step in this procedure is accomplished by the multiresolution approximation of Mallat (5). It is a pyramidal algorithm that consists of multiple decomposition levels, each of which reduces the resolution of the signal by a factor of 2. Transformation of the signal to the first decomposition level produces two components, one containing information about the low frequency part of the Fourier spectrum, and the other containing information about the high frequency part of the Fourier spectrum.

The bases of each orthogonal complement are built by dilating and translating a unique function. In the case of the low frequency complement, this function is known as the scaling function, and in the case of the high frequency complement, it is known as the wavelet function. The digital filter transforming the signal to the low frequency basis is a low pass filter, $\mathbf{W}_{\mathrm{lo}}$, and the counterpart filter transforming the signal to the high frequency complement is a high pass filter, $\mathbf{W}_{\mathrm{hi}}$. The filters used in the Haar wavelet transform (6) are described by the vectors

$$\mathbf{W}_{\mathrm{lo}} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, \dots, 0\right) \tag{S12a}$$

$$\mathbf{W}_{\mathrm{hi}} = \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, \dots, 0\right). \tag{S12b}$$

By applying the low pass filter, $\mathbf{W}_{\mathrm{lo}}$ to the time series $\mathbf{N}_{\mathrm{A}}$, one obtains the approximation $\mathbf{A}_1$ containing the signal at a reduced resolution

$$\mathbf{A}_1 = (\mathbf{W}_{\mathrm{lo}} * \mathbf{N}_{\mathrm{A}})_{\downarrow 2}. \tag{S13a}$$

Here, the symbol "*" denotes a convolution, and ↓2 represents the downsampling operation, where alternating elements of the convolution's output are removed. The high frequency, or detail component $\mathbf{D}_1$ is similarly produced by the convolution of the signal with the high pass filter

$$\mathbf{D}_1 = (\mathbf{W}_{\mathrm{hi}} * \mathbf{N}_{\mathrm{A}})_{\downarrow 2}. \tag{S13b}$$

The approximation and detail signals at subsequent decomposition levels are obtained by substituting the approximation from the previous level for the signal $\mathbf{N}_{\mathrm{A}}$ in Eqs. S13a and S13b.

The second step in the denoising procedure is the suppression of the noise component. As in the example above, the noise in a discrete-time photon signal is in the high frequency part of the Fourier spectrum, and as such, is contained within the high frequency component of the wavelet decomposition, the detail signal. The simplest way to reduce the noise is to set any detail component that exceeds a certain tolerance value to zero. An improved procedure utilizes "soft thresholding" (7), where noise suppression is achieved by shrinking any detail component Y of a decomposition according to:

$$\mathrm{Y} \rightarrow \mathrm{sgn}(\mathrm{Y})\max(0, |\mathrm{Y}| - \tau) \equiv \mathrm{T}(\mathrm{Y}, \tau). \tag{S14}$$

Here, τ is the universal threshold value given by Donoho (7)

$$\tau = \sigma[2 \log \mathrm{n}]^{\frac{1}{2}}. \tag{S15}$$

where, n is the number of data points in the original time series (i.e., the dimensionality of $\mathbf{N}_A$). Given that this method is designed to remove shot-noise, we approximate the noise parameter σ as the fluctuation about the mean intensity induced by shot-noise. For the acceptor photon trajectory $\mathbf{N}_A$ with mean intensity μ_A, we have $\sigma = [\mu_A]^{1/2}$. Likewise, for a donor photon trajectory with mean intensity μ_D, we have $\sigma = [\mu_D]^{1/2}$. Inserting these values in Eq. 15, we generate thresholds τ_A and τ_D, respectively, for each of the acceptor and donor photon trajectories

$$\tau_A = [2\mu_A \log n]^{\frac{1}{2}} \tag{S16a}$$

$$\tau_D = [2\mu_D \log n]^{\frac{1}{2}}. \tag{S16b}$$

Returning to the detail signal obtained in Eq. 13b, we now apply, element-wise with threshold calculated from Eq. 16a, the thresholding operation of Eq. 14, and obtain thresholded detail signal $\mathbf{D}_1{}^T$

$$\mathbf{D}_1^T = T(\mathbf{D}_1, \tau_A). \tag{S17}$$

Thresholded details at subsequent decomposition levels are obtained by applying the thresholding operator with the same threshold as in Eq. 16a.

The denoised signal $\mathbf{S}_A$ is obtained by inverting the decomposition procedure described above. Firstly, we define the low pass and high pass reconstruction filters, $\mathbf{W}_{lo}{}^{-1}$ and $\mathbf{W}_{hi}{}^{-1}$ respectively, as the reverse of their decomposition counterparts (6)

$$\mathbf{W}_{lo}{}^{-1} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0, \ldots, 0\right) \tag{S18a}$$

$$\mathbf{W}_{hi}{}^{-1} = \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0, \ldots, 0\right). \tag{S18b}$$

Next, we dyadically upsample the highest level approximation and thresholded detail components by inserting zeros between each element of each vector. After upsampling, the approximation vector is convolved with $\mathbf{W}_{lo}{}^{-1}$, and the detail vector with $\mathbf{W}_{hi}{}^{-1}$. The sum of the output of each convolution is then the reconstructed approximation of the next resolution level. This reconstructed approximation is then upsampled, convolved, and combined with the next level's thresholded details. This procedure is summarized by Eq. S19 for the reconstruction of the denoised signal $\mathbf{S}_A$ from a level 1 approximation and thresholded details, where the superscript ↑2 represents the upsampling operation

$$\mathbf{S}_A = \left(\mathbf{W}_{lo}{}^{-1} * (\mathbf{A}_1)^{\uparrow 2}\right) + \left(\mathbf{W}_{hi}{}^{-1} * (\mathbf{D}_1^T)^{\uparrow 2}\right). \tag{S19}$$

Evaluation of the mean-squared error of simulated signals decomposed to various resolution levels revealed that this quantity is largely minimized for signals reconstructed from the third resolution level. As such, all decompositions were processed to the third level. Due to the nature of the decomposition, the number of elements in the time series must be a power of 2. Each signal was extended with a vector containing m elements, where the value of each element is the mean intensity of the signal, and where m is chosen such that $n + m = 2^j$, with j being an integer. The dimensionality of the signal, n, remained unchanged in Eqs. S16. This extension was discarded after reconstruction of the denoised signal. Each of the acceptor and donor photon signals in a particular smFRET trajectory were decomposed separately using the digital filters in Eqs. S12a and S12b. Thresholds for each signal were generated using Eqs. S16a and S16b, and the detail signals of each decomposition were thresholded via the soft thresholding operator in Eq. S14. Each of the acceptor and donor signals were then reconstructed as in Eq. S19.

Fig. S3 illustrates the denoising of an acceptor photon trajectory. The fluorophore-emitted acceptor photon signal is shown in Fig. S3a. Convolution of the signal vector with the low pass filter results in the first level approximation shown in Fig. S3b after downsampling. The second and third level approximations are generated by convolution of the previous level's approximation with the low pass filter. The detail coefficient vectors shown in Fig. S3c are similarly produced by convolution of the signal and approximation vectors with the high pass filter. The universal threshold as calculated from Eq. S16a is shown as dashed lines in Fig. S3c. Detail coefficients whose magnitudes are smaller than the threshold are set to zero, and the threshold is subtracted from detail coefficients having magnitudes larger than the threshold with retention of these coefficients' signs.

Fig. S3d shows the original acceptor signal in red overlaid with its denoised representation in black. The denoised signal was reconstructed by upsampling the third level

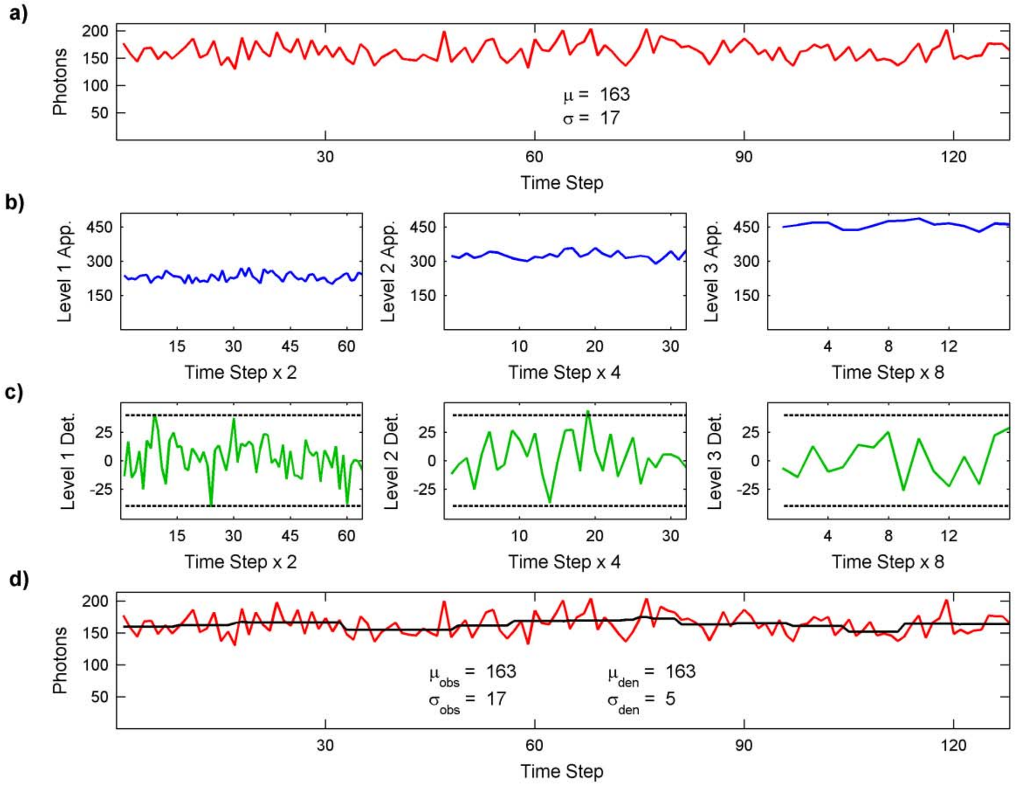

Figure S3. Denoising an acceptor photon trajectory. a) The original trajectory. b) The first, second, and third level approximation coefficients shown left to right. c) The first, second, and third level detail coefficients shown left to right, along with the detail threshold is shown as dotted lines. d) The original signal (4) overlaid with the denoised signal (black). The standard deviation of the original signal is reduced by a factor of 3, illustrating that the denoising process has been a success.

approximation and thresholded details, convolution of the upsampled coefficients with their respective low or high pass reconstruction filters, and addition of the two upsampled vectors, producing the second level reconstruction. This reconstruction vector was then combined with thresholded second level details in the same manner, and the resulting vector combined with the thresholded first level details to produce the reconstructed signal shown in black in Fig. S3d. It is seen that, while the mean of the denoised signal remains the same as its noisy counterpart, comparison of the standard deviations of each trajectory shows the noise level to be reduced by approximately 300% for this particular photon trajectory.

Part S4: Photoblink Removal from Simulated Data

To compare the effects of the Bayesian photoblink detection algorithm to a more traditional method of photoblink removal, simulated trajectories were constructed using kinetic Monte Carlo simulation. Three states were included in this simulation, the observable state with central efficiency of 0.15 and average lifetime of ~ 1 s, an acceptor photoblinking state, and a donor photoblinking state. The lifetimes used for the photoblinking states were allowed to vary, producing photoblinks with lifetimes ranging from 10 ms (1 time step) to 5 seconds. By inspection of the sample trajectory shown in Fig. S4a, one can clearly see that, in the case of these trajectories, manual photoblink removal would most likely have a negative effect on the data. Human bias would arise due to the slight difference between what is actual data and what is a photoblink.

Photoblinks were filtered from the simulated trajectories by 2 different methods: (1) the Bayesian method presented here, and (2) a simpler method that removed time steps with acceptor intensities below a predefined threshold τ_B, given by

$$\tau_B = \mu_B + 2(\mu_B)^{\frac{1}{2}}. \qquad (S20)$$

where, μ_B is as defined in Eq. S7.

Fig. S4 shows the results of this comparison. A sample trajectory containing unfiltered acceptor (4) and donor (blue) counts is shown (Fig. S4a) to have photoblinks of both the acceptor and donor varieties on both long and short time scales, as well as the presence of the observable state with a central efficiency of 0.15. Figs. S4b and S4c show the same trajectory after photoblink filtering by the thresholding method and by the Bayesian method, respectively. It is clear in these trajectories that a larger fraction (~ 15% larger) of the original data is retained by the Bayesian operation, but it remains unclear if the remaining data is of the observable variety or of the photoblink variety.

Fig. S4d shows a distribution of the unfiltered data’s calculated efficiencies. While we can see the presence of the observable state at a its central efficiency of 0.15, we also see that the distribution is marred by the noise arising from the presence of photoblinks. Fig. S4e shows the corresponding distribution after photoblink removal using the thresholding method, and Fig. S4f shows the distribution after photoblink removal using the Bayesian method.

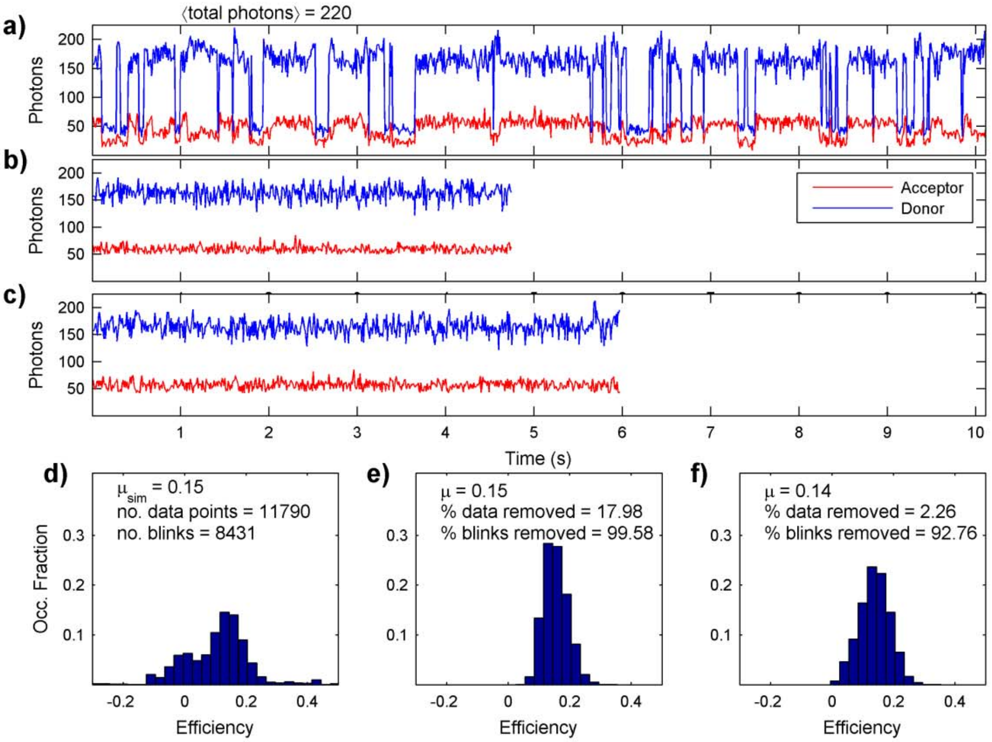

Figure S4. Comparison of photoblink removal methods. a) A sample trajectory from the simulated data. b) The sample trajectory after removal of photoblinks by the thresholding method. c) The sample trajectory after removal of photoblinks by the Bayesian method. d) The distribution of efficiencies produced by the unfiltered, simulated data. e) The efficiency distribution after photoblink removal using the thresholding method. f) The efficiency distribution after photoblink removal using the Bayesian method.

It is clear from the statistics shown in Fig. S4e that the thresholding method successfully removes photoblinks from the trajectories (with > 99% success), but it is also clear that it removes a large portion of the actual data as well (~ 18%). This is a less than desirable quality, in that, if nearly 20% of the observable data is removed, then nearly 20% of the information acquired about this system is lost. In comparison, the Bayesian method also successfully removes photoblinks (with ~ 93% success), and removes a much smaller fraction of the observable data (~ 2.3%). While a portion of the original photoblinks remain, their contribution to the overall outcome is small, as is shown by the distribution in Fig. S4f. Perhaps more importantly, the human element of photoblink detection, whether it be in the manual selection of photoblink regions, or in the selection of an intensity threshold, has been completely eliminated.

Part S5: The Performance of Bayesian Photoblink Detection in Relation to K_{eq}

Given the substantial variation of equilibrium constants in various biological interactions, the performance of the Bayesian photoblink filter was tested on simulated, two-state equilibria with equilibrium constants ranging over 4 orders of magnitude. The smFRET trajectories for each equilibrium were simulated using kinetic Monte Carlo simulations with the mean number of total photons at each time step being 220. Each simulated equilibrium contained over 100,000 data points prior to photoblink removal, and consisted of two efficiency states, one with a central efficiency of 0.2 (state 0.2), and the other with central efficiency of 0.8 (state 0.8). The forward transition is taken to be the transition from state 0.8 to state 0.2, such that a $K_{eq} < 1$ corresponds to state 0.8 being favored in the equilibrium. The average photoblinking lifetime was ~ 0.4 s in all cases, and photoblinks were removed as described above.

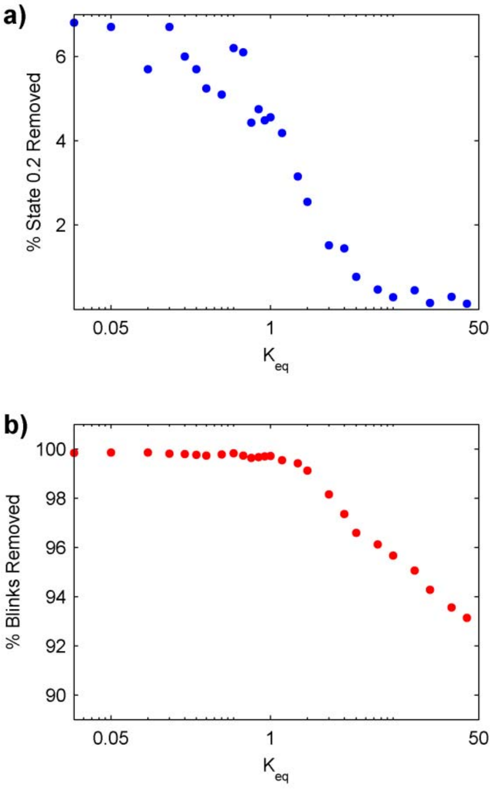

Figure S5. The performance of the Bayesian photoblink filter vs. K_{eq}. a) The percentage of state 0.2 data points removed vs. K_{eq}. b) The percentage of photoblinks removed vs. K_{eq}.

Fig. S5 shows the results of these simulations. In Fig. S5a, the percentage of state 0.2 data points that were removed in each simulation are plotted logarithmically versus K_{eq}. It is shown by Fig. S5a that the percentage of state 0.2 removed never exceeds 6.8%, and that, as the equilibrium shifts toward this state, the percentage of this state's data points that are removed trends toward zero. Fig. S5b shows the percentage of photoblinks removed versus K_{eq}, again plotted logarithmically. As shown in Fig. S5b, the percentage of photoblinks removed by the Bayesian photoblink filter only begins to fall after state 0.2 becomes favored in the equilibrium. This is simply a result of the mean acceptor intensity falling as the mean efficiency of the trajectory falls, and of the prior probability of a photoblink being less than that of the absence of a photoblink. We therefore conclude that, regardless of which state is favored in the equilibrium, as well as the extent to which a that state is favored, the Bayesian photoblink filter performs capably under all tested circumstances.

Part S6: Denoising a Simulated System with Well-Defined States and Dynamics

As a means to quantify the effects of the wavelet denoising algorithm to smFRET trajectories, a two state equilibrium having an equilibrium constant of 0.33 was simulated using kinetic Monte Carlo methods. The states were assigned mean efficiencies of 0.2 and 0.8, respectively, with the state having mean efficiency of 0.8 being favored in the equilibrium. Acceptor and donor photon trajectories that include shot-noise were constructed from the simulated trajectories. Each of the simulated trajectories was denoised by the wavelet denoising algorithm as well as the hidden-Markov model (HaMMy) described by McKinney, *et al* (8).

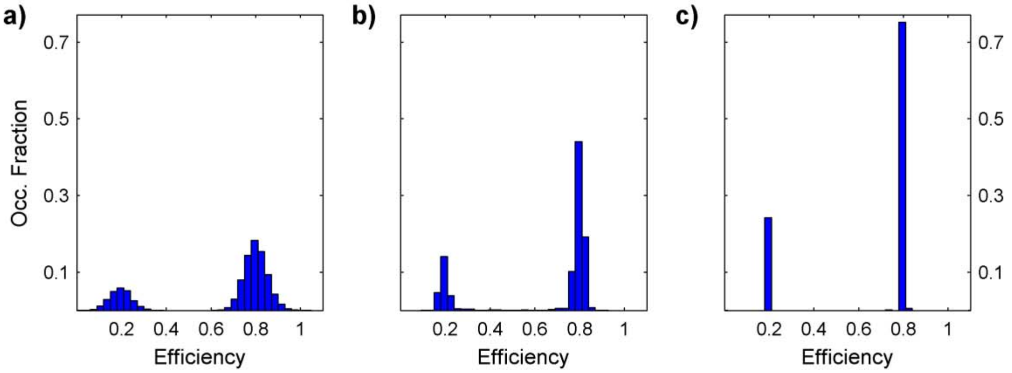

Figure S6. Efficiency distributions of a simulated two state system. a) The shot-noise induced efficiency distribution of a simulated two state system having states with mean efficiencies of 0.2 and 0.8. b) The efficiency distribution of the system after denoising. c) The efficiency distribution as generated by the hidden-Markov model HaMMy (8).

Fig. S6a shows the efficiency distribution generated by the simulated photon trajectories. Each state's standard deviation was calculated as measurement error and found to be approximately 0.06 efficiency units for both states at a mean of 220 total photons per time step. The standard deviation of each state's distribution was also found by least-squares regression to normal distributions, and these values show good agreement as the standard deviations of both states were again found to be approximately 0.06 efficiency units.

Figure S6b shows the efficiency distribution produced after wavelet denoising. The standard deviation of each state's efficiency distribution was again estimated by least-squares regression to normal distributions. These values were again found to be approximately equal at 0.02 efficiency units, translating to a reduction in the width of each state's efficiency distribution of 300%. It is shown in Figure S6c that HaMMy identifies each state nicely by collapsing the broad efficiency distribution of each of the states nearly to a single efficiency. Although the denoising algorithm does not reduce the width of each distribution as radically as HaMMy, the wavelet denoising algorithm is not, as is HaMMy, constructed to do so. The comparison does show, however, that the denoising algorithm narrows each state's efficiency distribution by 300%, thereby proving its viability in the case of a system having well-defined states and kinetics while offering the advantage of making no *a priori* assumptions regarding the state's thermodynamic or kinetic properties.

Part S7: Denoising a Single smFRET Trajectory

Here we examine a single smFRET trajectory obtained in the aV aptamer studies described previously (9). Fig. S7a depicts the fluorophore-emitted photon signals in 10 ms time steps, with the acceptor signal shown in red and the donor signal shown in blue. The trajectory includes photoblink anomalies of both the acceptor and donor varieties on both short and long time-scales. Additionally, intensity variation in of each of the signals is observed, and arises both from energy transfer efficiency fluctuations and shot-noise. Fig. S7b shows the photoblink-filtered and wavelet-denoised versions of the same acceptor and donor trajectories. Time steps identified as photoblinks were removed from the acceptor (4) and donor (cyan) trajectories shown, and each is overlaid with its wavelet-denoised complement (denoised acceptor in black and denoised donor in blue). Fig. S7c shows the energy transfer efficiency calculated from each complementary pair of signals in Fig. S7b using Eq. S5. The observed efficiency trajectory was calculated at each time step using the photoblink-filtered pair of acceptor and donor signals shown in Fig. S7b, and this trajectory is overlaid in black by the denoised efficiency calculated at each time step using the wavelet-denoised acceptor and donor signals shown in Fig. S7b. Fig. S7d shows a histogram of the observed efficiencies in Fig. S7c, and Fig. S7e shows that of the denoised efficiencies shown in Fig. S7c.

The method presented to detect photoblinks, as shown by Fig. S7b, proves to be quite discriminatory in its application to experimental data. Not only is the method able to distinguish both acceptor and donor photoblinks, but making a decision at every time step allows for the distinction of both long and short time-scale photoblinks as well. Additionally, given that the expected minimum energy transfer efficiency of a valid state of the system is larger than approximately 20%, the method shows the ability to distinguish a photoblink from a steep drop in signal intensity that arises from a conformational change.

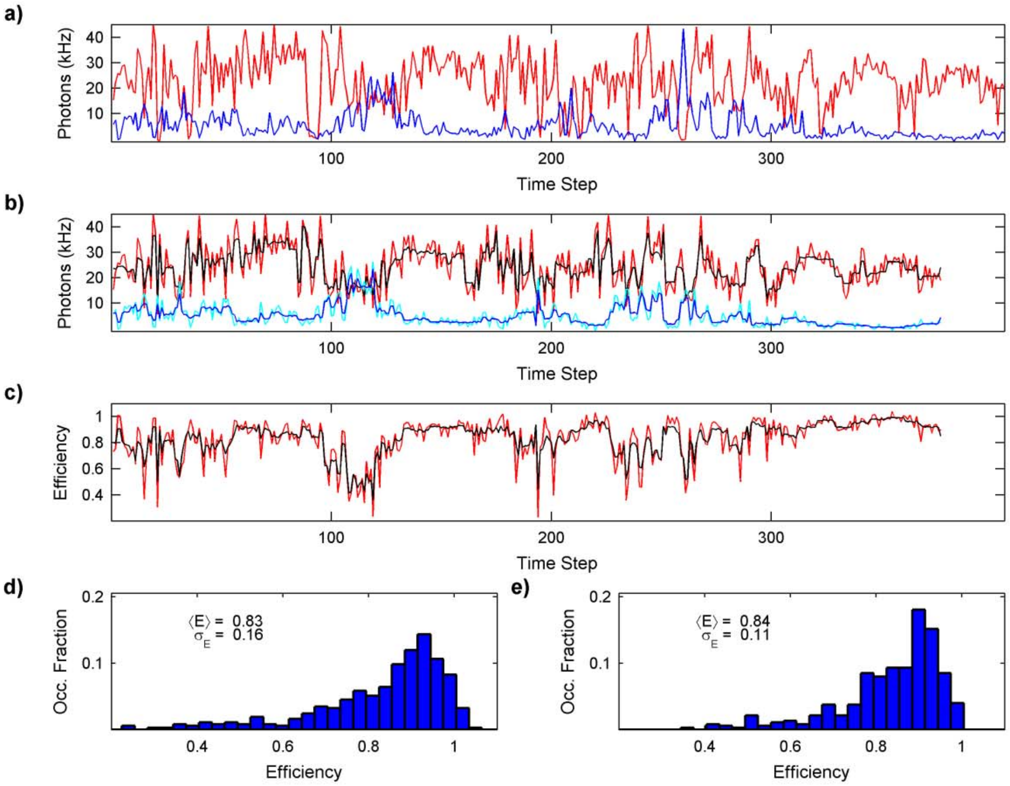

Figure S7. Denoising an experimental smFRET trajectory. a) The original fluorophore-emitted acceptor (4) and donor (blue) photon trajectories. b) The original acceptor photons (4) are overlaid with their denoised counterparts (black), and the original donor photons (cyan) are overlaid with their denoised counterparts (blue). c) The smFRET efficiency calculated from the original acceptor and donor photon counts (4) is overlaid by that calculated from the denoised acceptor and donor photon counts (black). d) The efficiency histogram generated by the noisy data in b). e) The efficiency histogram generated by the denoised data shown in b).

Also, as shown by Fig. S7b, the wavelet-denoising algorithm proves to be quite effective in its application to experimental data. While fluctuations in the signals' intensities due to experimental considerations such as apparatus limitations and fluorophore orientations are unavoidable, it is shown in the denoised signals that small fluctuations arising from quantifiable sources are virtually eliminated. Of equal, or perhaps greater, importance, it is also shown that large intensity fluctuations that are induced by conformational changes in the system are allowed to remain.

The significance of these aspects is clear, and is illustrated by the efficiency trajectories in Fig. S7c. The observed trajectory of calculated efficiencies (4) fluctuates wildly from time step to time step due to small, insignificant changes in the signal intensity on one or both of the detection channels. The denoised complement (black) to this trajectory, however, does not exhibit such excessive fluctuation. We still observe major fluctuations in efficiency, and these are virtually unchanged from the major fluctuations we see in the observed efficiency trajectory. Anomalously large and fast efficiency fluctuations have been reduced, and we observe a smoother trajectory as well as more accurate representation of our physical system.

Figs. S7d and S7e further validate this point by showing that, while the shape of each efficiency distribution is approximately the same, and their mean efficiencies are virtually unchanged, the denoised standard deviation is reduced by approximately 30%. This validates that occurrences of anomalously large or small efficiencies have been reduced. On the whole, the application of the photoblink detection and wavelet-denoising algorithms is shown by Figure S7 to improve the quality of this experimental smFRET trajectory.

Part S8: HIV-1 TAR DNA: Denoising a Single State Experimental System

Here we apply the photoblink detection and wavelet-denoising algorithms to a collection of smFRET trajectories. The trajectories are collected from studies, described previously (10), on the transactivation response (TAR) region of HIV-1 viral DNA. The collection contains nearly 16,000 data points, and was chosen to represent a predominantly single state system. The secondary structure of the TAR hairpin is shown in Fig. S7a, and the depicted conformation is expected to yield smFRET efficiencies approaching unity.

Application of the photoblink detection algorithm to this collection of trajectories yields, after computation of energy transfer efficiency at each of the individual data points, the global efficiency histogram that is shown in Fig. S8b. As expected, the system does produce smFRET efficiencies approaching unity, and the normal distribution about the mean energy transfer efficiency of 0.96 signifies that we do observe a predominantly single state in this collection of trajectories. The width of this distribution is small, again signifying a single state, and we observe this value to be 0.09 efficiency units.

Calculating and compiling each data point's energy transfer efficiency from wavelet-denoised complements produces the global efficiency histogram shown in Fig. S8c. While we do not observe a shift in the mean smFRET efficiency, we do observe a sizeable reduction in the magnitude of the distribution's standard deviation. The width of this distribution, while small to begin with, is reduced by approximately 45% by the denoising algorithm. The significance is clear – in the characterization of a single state system, eliminating the artifacts of shot-noise and photoblinks results in a more precise representation of the state's structure and energetics.

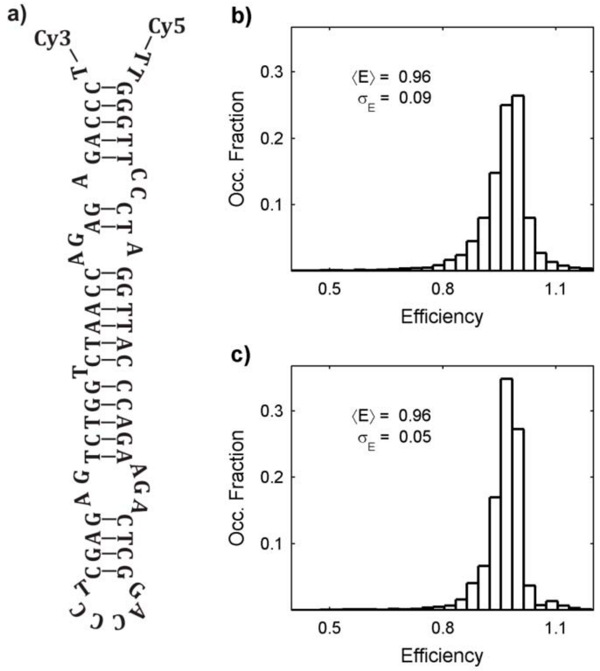

Figure S8. Single state TAR DNA. a) The secondary structure of TAR. B) The efficiency distribution of observed and blink-filtered data acquired from experiments involving TAR in 2 mM Mg^{2+} buffer solution. c) The corresponding denoised data.

Part S9: Acceptor Photobleaching: Denoising a 2-State Experimental System

As a model two state system, irreversible acceptor photobleaching from a high efficiency state is chosen. This collection of trajectories was also chosen from studies on the TAR region of HIV-1 viral DNA described previously (10). This system is treated as a purely two state system: energy transfer is either "on", as shown to the left in Fig. S9a, or "off", as shown in the center of Figure S8a. As the crosstalk region of this collection of trajectories is being considered as part of the FRET region, the crosstalk value x is fixed to a characteristic value of 0.11. Additionally, given that one of the states in our model has an efficiency that is expected to approach zero, the previously discussed caveat arises. We circumvent this caveat by simply marking the time step at which the acceptor photobleach occurs. The photoblink detection method is applied to time steps prior to this time step as previously described, and for time steps after the photobleach we substitute the donor photon signal for the acceptor and proceed in the same manner.

Fig. S9b shows a global efficiency histogram of approximately 18,000 blink-filtered data points in this collection of smFRET trajectories. The "on" state is represented to the right, and shows a mean efficiency of 0.98 with a standard deviation of 0.17. The "off" state is represented to the left, and this distribution shows a mean efficiency of 0.01 with equal width of 0.17. To avoid bias, instead of selecting the limits of each state manually, each mean and standard deviation is produced by a least-squares regression to the sum of two normal distributions.

Figure S9c shows the wavelet-denoised complements of Figure S9b. Again we obtain the characteristics of each state's efficiency distribution by a least-squares regression, and again we find dramatic improvement. While each state's mean value has remained unchanged, their distributions have narrowed significantly. In the case of the "off" state, we see that the distribution has narrowed by approximately 20%, and in that of the "on" state, we see a larger narrowing of just under 25%. As such, we conclude that the wavelet-denoising algorithm is capable of removing quantifiable noise components of each state's efficiency distribution, resulting in more precise description of each state.

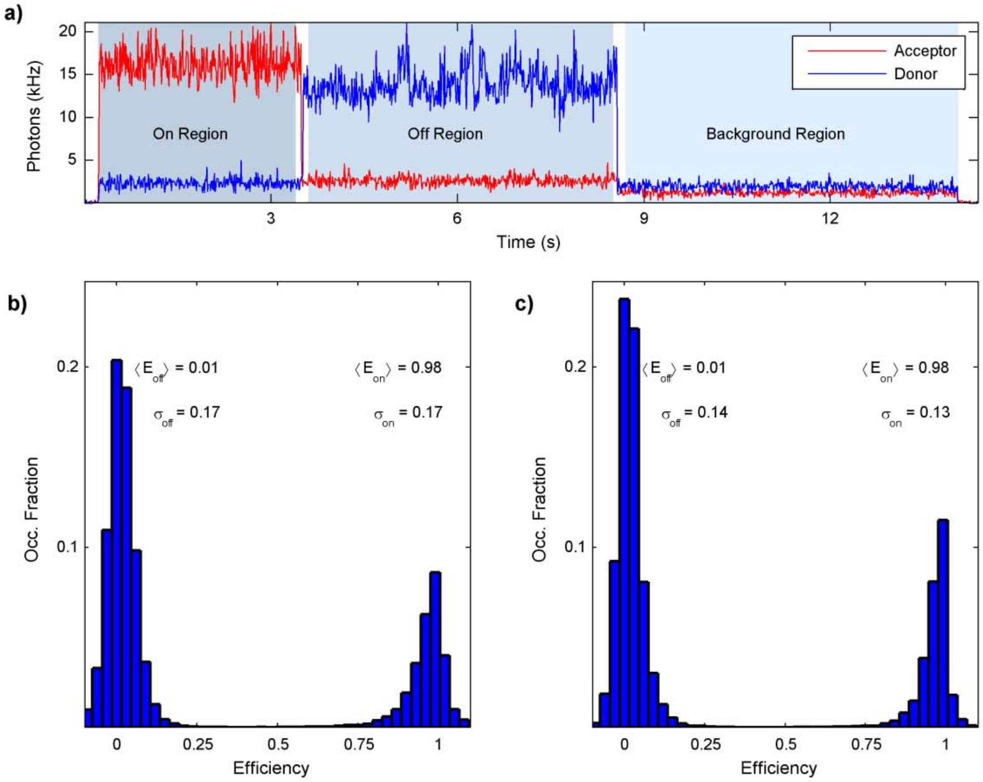

Figure S9. Irreversible acceptor photobleaching as a purely two state system. a) A model trajectory where the "on" state is represented by the region before the acceptor photobleach, and the "off" state is represented by the region after the acceptor photobleach. b) The efficiency distribution compiled from blink-filtered trajectories. c) The corresponding denoised efficiency distribution.

REFERENCES

1. Nir, E., X. Michalet, K. M. Hamadani, T. A. Laurence, D. Neuhauser, Y. Kovchegov, and S. Weiss. 2006. Shot-noise limited single-molecule FRET histograms: Comparison between theory and experiments. *J. Phys. Chem. B.* 110:22103-22124.
2. Sabanayagam, C. R., J. S. Eid, and A. Meller. 2005. Using fluorescence resonance energy transfer to measure distances along individual DNA molecules: Corrections due to nonideal transfer. *J. Chem. Phys.* 122:061103-061105.
3. Winkler, R. L. 1972. An introduction to bayesian inference and decision. Holt, Rinehart and Winston, New York.
4. Haar, A. 1910. Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen.* 69:331-371.
5. Mallat, S. G. 1989. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Patt. An. Mach. Int.* 11:674-693.
6. Daubechies, I. 1992. *10 Lectures on Wavelets.* Society for Industrial and Applied Mathematics, Philadelphia, PA.

7. Donoho, D. L. 1995. De-noising by soft-thresholding. *IEEE Trans. Inf. Theory*. 41:613-627.

8. McKinney, S. A., C. Joo, and T. Ha. 2006. Analysis of single-molecule FRET trajectories using hidden-Markov modeling. *Biophys. J.* 91:1941-1951.

9. Taylor, J. N., Q. Darugar, K. Kourentzi, R. C. Willson, and C. F. Landes. 2008. Dynamics of an anti-VEGF DNA aptamer: A single-molecule study. *Biochem. Biophys. Res. Comm.* 373:213-218.

10. Darugar, Q., H. Kim, R. J. Gorelick, and C. Landes. 2008. Human t-cell lymphotropic virus type 1 nucleocapsid protein-induced structural changes in transactivation response DNA hairpin measured by single-molecule fluorescence resonance energy transfer. *J. Virology*. 82:12164-12171.