Supplemental Information (SI)

3D single-molecule tracking enables direct hybridization kinetics measurement in solution

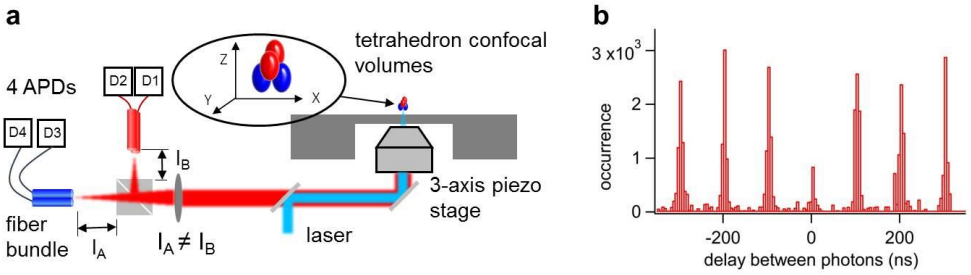
Cong Liu, Judy M. Obliosca, Yen-Liang Liu, Yu-An Chen, Ning Jiang and Hsin-Chih Yeh*

Department of Biomedical Engineering, University of Texas at Austin, Austin, Texas, 78712, USA

Content

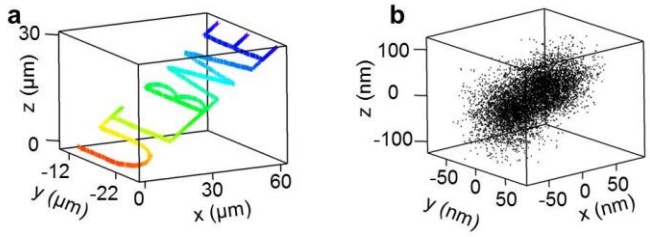
Supplemental Note S1: 3D single-molecule tracking microscope

The 3D tracking microscope is built around an Olympus IX-71 microscope. Pulsed laser (average power 100 μW, repetition rate 10 MHz) from PicoQuant LDH-P-C-640B is reflected by the dichroic mirror (Semrock FF650-Di01-25x36), and then focused by a 60× NA=1.2, water immersion objective (Olympus, UPLSAPO 60XW). The laser beam size is controlled by a Keplerian beam expander for slightly underfilling the objective. The fluorescence is collected by the same objective, and filtered by ET700/75m (Chroma). Based on the number of photons collected in two optical fiber bundles (Polymicro, 50-μm core diameters, 55-μm center-to-center spacing) in every 5 ms, a xyz piezo stage (PI, P-733K130) with 30×30×30 μm travel range is used to reposition the fluorescent molecule to the center of excitation focus. The detailed description of the 3D tracking setup can be found in our previous publications[1,2].



(**a**) Schematics of the 3D confocal tracking microscope (**b**) Photon-pair correlation histogram obtained from 190 5'-Atto647N-ssDNA tracks. The suppressed central peak (antibunching[3]) demonstrates that the trajectories are of single ssDNA molecules.

To characterize the tracking accuracy of our setup, an immobilized fluorescent bead (F8783, Ø=20 nm) is programmed to move in a pattern that reads "UTBME" (see figure below), with a speed of 5 μm/s. The same piezo stage that is used to move the bead in that pattern is also used for tracking in a time-multiplexed way. By comparing the predefined trajectory of the fluorescent bead with its trajectory inferred from the tracking mechanism, the RMS tracking accuracy has been determined to be 25 nm in xy, and 81 nm in z.



(**a**) 3D trajectory of the fluorescent bead inferred from the tracking mechanism (**b**) the scatter plot of tracking error in x, y and z.

Supplemental Note S2: Fluorescence lifetime measurement

The fluorescence lifetime of tracked molecules is measured in TTTR (time-tagged time-resolved) mode with the TCSPC system PicoHarp 300 (in combination with PHR 800, synchronized with the 3D tracking system), which allows for simultaneous measurement of 4 channel signals. The photon arrival times are registered with 128 ps resolution, and post-processed by custom-written MATLAB scripts to build fluorescence decay histograms for each channel individually, with an integration time 15 ms.

To combine the 4 histograms into a single one for lifetime fitting, time delays among the four channels caused by cable length differences and misalignment have to be properly compensated by shifting the histograms along the time axis until their peaks overlap. The amount of shifts (in units of 128 ps), however, is determined by a separate experiment, in which the time delays among the 4 channels can be calibrated with a much higher resolution (4 ps) and better signal-to-noise ratio (~15 million fluorescence photons). In that experiment, the fluorescence decay histograms (one for each channel) of 50 nM ATTO633 labeled ssDNA are obtained with the same TCSPC setup (integration time 600 seconds, resolution 4ps, laser power 0.25 μW), and the relative time delays T_i (i=1-4) of peaks in the four histograms are measured. For all the following single-molecule experiments, a fixed amount of shift $-T_i$ is then applied to the i-th channel histogram.

Supplemental Note S3: Fluorescence lifetime fitting

We use maximum likelihood estimation (MLE)[4] to fit the fluorescence lifetime. It has been shown that 100 detected photons are sufficient to determine a single exponential decay by MLE[5].

In time-correlated single-photon counting, assume S signal counts are accumulated in k bins, where n_i is the number of detected photon in i-*th* bin. f is the number of fitted parameters. The fluorescence lifetime pattern C is generated by the convolution of the instrument response function (IRF) with a single exponential decay $e^{-t/\tau}$ (Equation (1)). IRF is obtained from laser light scattered by ludox (20 wt % suspension in water).

$$c_i(\tau, s) = \mathrm{IRF} \otimes (e^{-t/\tau}) \tag{1}$$

Taking a fraction γ^2 of constant background into consideration, Equation (2) gives m_i, the probability of that a photon will fall into channel i for a certain fluorescence lifetime pattern.

$$m_i(\tau, s, \gamma) = \frac{\gamma^2}{k} + (1-\gamma^2)\frac{c_i(\tau, s)}{\sum c_i(\tau, s)} \tag{2}$$

Normalization of $2I^*$ by the degree of freedom ($k-1-f$) leads to reduced $2I_r^*$, which is 1 for an optimal fit.

$$2I_r^* = \frac{2}{k-1-f}\sum_{i=1}^{k} n_i \ln\left(\frac{n_i}{S m_i(\tau, s, \gamma)}\right) \tag{3}$$

To determine the background fraction γ^2 and the shift s of the IRF with respect to the fluorescence decay, we add up fluorescence decays of all single-molecules tracked. This integration analysis will ensure good signal-to-noise ratio. s and γ were first allowed to vary freely with τ. The values of s and γ giving a minimal $2I_r^*$ were kept constant in the further analysis of individual single-molecule fluorescence decays, *i.e.* the only remaining adjustable variable is the parameter of interest τ. It doesn't matter which IRF of the four channels you are using for lifetime fit, because the results are almost identical.

Supplemental Note S4: Signal-to-noise ratio (SNR) characterizing

The four-channel-summed photon count rate averaged from ~3000 single-molecule trajectories is 7.97 kHz, and the photon count rate without the presence of reporter strands is 3.37 kHz. This background noise includes both the detector dark count and the emission from 1 μM Iowa Black® FQ. SNR is therefore determined to be (7.97-3.37)/3.37=1.4. The same filters, laser power, and glycerol (70 wt %) and Tris-HCl buffer (20 mM, pH 8.0) concentration are used in the measurements.

Supplemental Note S5: ebFRET algorithm benchmark (simulation)

The ebFRET package [6] developed by Wiggins lab is benchmarked by semi-ideal lifetime traces generated *in silico*.

$$\tau(t_i) = \begin{cases} \epsilon \sim N(\tau_H, \sigma_H^2), & if\ s(t_i) = \text{'hybridized'} \\ \epsilon \sim N(\tau_L, \sigma_L^2), & if\ s(t_i) = \text{'melted'} \end{cases}$$

where τ_H (2.4 ns) and τ_M (3.6 ns) are constants determined from our experimental data when the quencher concentration is 600 nM, $s(t_i)$ is the state of molecule being tracked at time index t_i. σ_H^2 and σ_L^2 are variance of MLE fluorescence lifetime estimation. We estimate σ_H^2 and σ_L^2 based on a set of empirical parameters, assuming a background-free mono-exponential fluorescence decay model[7]:

$$var_N(\tau, T, k) = \frac{1}{N}\tau^2 \frac{k^2}{r^2}[1 - \exp(-r)]\left(\frac{\exp\left(\frac{r}{k}\right)[1 - \exp(-r)]}{\left[\exp\left(\frac{r}{k}\right) - 1\right]^2} - \frac{k^2}{\exp(r) - 1}\right)^{-1}$$

Where N is the number photons used for fitting, k is the number of channels, T is the measurement window, τ is the lifetime and $r = \frac{T}{\tau}$. In our experiments the resolution is 128 ps, k=110, T=14.08 ns, and N=135 on average for a 15 ms integration window. Using the above equation, $\sqrt{var(\tau)}/\tau$ is determined to be 11%. The number 11% sets the limit of sensitivity that can ultimately be attained.

In our simulations, the true transition rates k'_{on} and k'_{off} are defined to be 5 s^{-1} and 10 s^{-1} respectively. The track durations follow a geometric distribution:

$$p(x = k\Delta t) = (1 - p)^{k-1}p$$

Where $1 - p$ is the probability of successful tracking in one time step, assuming the single-molecule tracking experiment is a sequence of Bernoulli trails. p is determined to be 0.13 from out experimental data.

Given k'_{on}, k_{off}, and $\Delta t = t_{i+1} - t_i$ =15 ms, $s(t_i)$ as a Markov chain is generated by Matlab package pmtk3.

Supplemental Note S6: Convert transition matrix to annealing-melting rates

A two-state hidden Markov chain is used to model single-molecule tracks. In this model, the melted state (ssDNA) can be denoted as state 1, whereas the hybridized state (dsDNA) can be denoted as state 2. The state transition matrix p_{ij} ($1 \leq i, j \leq 2$) describes the probability of the DNA molecule transitioning from state i to state j in one step.

Since ebFRET only reports the estimated transition matrix, instead of the annealing-melting rates of our interest, proper conversion is needed. Based on the relation shown in Equation (4)-(6) [8], annealing (k_{on}) and metling rate (k_{off}) can be derived from Equation (7)-(8).

$$p_{12} = \frac{k_{\mathrm{on}}}{k_{\mathrm{on}} + k_{\mathrm{off}}}[1 - e^{-(k_{\mathrm{on}} + k_{\mathrm{off}})\Delta t}] \tag{4}$$

$$p_{21} = \frac{k_{\mathrm{off}}}{k_{\mathrm{on}} + k_{\mathrm{off}}}[1 - e^{-(k_{\mathrm{on}} + k_{\mathrm{off}})\Delta t}] \tag{5}$$

$$p_{11} = 1 - p_{12}, p_{22} = 1 - p_{21} \tag{6}$$

$$k_{on} = -\frac{p_{12}\ln(1 - p_{21} - p_{12})}{(p_{12} + p_{21})\Delta t} \tag{7}$$

$$k_{off} = -\frac{p_{21}\ln(1 - p_{21} - p_{12})}{(p_{12} + p_{21})\Delta t} \tag{8}$$

The variance of estimated k_{on} can be calculated via Equation (9)

$$\mathrm{Var}(k_{on}) = \left(\frac{\partial k_{on}}{\partial p_{12}}\right)^2 Var(p_{12}) + \left(\frac{\partial k_{on}}{\partial p_{21}}\right)^2 Var(p_{21}) + 2\left(\frac{\partial k_{on}}{\partial p_{12}}\right)\left(\frac{\partial k_{on}}{\partial p_{21}}\right) Cov(p_{12}, p_{21}) \tag{9}$$

where

$$\frac{\partial k_{on}}{\partial p_{12}} = -\frac{p_{12}(p_{12} + p_{21}) + p_{21}(p_{12} + p_{21} - 1)\ln(1 - p_{12} - p_{21})}{\Delta t(p_{12} + p_{21} - 1)(p_{12} + p_{21})^2}$$

$$\frac{\partial k_{on}}{\partial p_{21}} = \frac{p_{12}((p_{12} + p_{21} - 1)\ln(1 - p_{12} - p_{21}) - p_{12} - p_{21})}{\Delta t(p_{12} + p_{21} - 1)(p_{12} + p_{21})^2}$$

$$\mathrm{Cov}(p_{12}, p_{21}) = 0$$

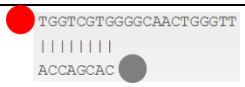
The variance of estimated k_{off} can be calculated by replacing k_{on} with k_{off} in Equation (9), and

$$\frac{\partial k_{off}}{\partial p_{12}} = \frac{p_{21}((p_{12} + p_{21} - 1)\ln(1 - p_{12} - p_{21}) - p_{12} - p_{21})}{\Delta t(p_{12} + p_{21} - 1)(p_{12} + p_{21})^2}$$

$$\frac{\partial k_{off}}{\partial p_{21}} = -\frac{p_{21}(p_{12}+p_{21}) + p_{12}(p_{12}+p_{21}-1)\ln(1-p_{12}-p_{21})}{\Delta t(p_{12}+p_{21}-1)(p_{12}+p_{21})^2}$$

Supplemental Table S1: Quenching efficiency (ATTO633 w/ Iowa Black® FQ) characterization

To optimize the donor-quencher spacing for transient binding observation, we measured the fluorescence lifetime of a set of stable DNA duplex (end-labeled by ATTO633 and Iowa Black® FQ) on an ensemble level with our confocal TCSPC system. The resolution is 4ps, and integration time is 600 seconds. The fluorescence lifetime as a function of donor-quencher spacing is shown below. Two different types of quencher-labeled ssDNA were prepared. The first type is longer (by 2, 4 or 6 nt) than the dye-labeled ssDNA, and the formed dsDNA have dye and quencher at the same side. The other type is shorter than the dye-labeled ssDNA, and has locked nucleic acids (LNA) embedded to increase the binding affinity. The formed dsDNA have dye and quencher at opposite sides.

Lifetime (ns)	donor-quencher spacing	DNA duplex schematic	DNA duplex ID
0.98	2 nt		ATTO633_FQ_2nt
1.18	4 nt		ATTO633_FQ_4nt
1.38	6 nt		ATTO633_FQ_6nt
2.01	8 nt		ATTO633_FQ_8nt
2.96	10 nt		ATTO633_FQ_10nt
3.85	14 nt		ATTO633_FQ_14nt
4.16	No quencher		/

The DNA duplexes were hybridized in-house in 100 mM pH 8.0 Tris-HCl buffer. The sequences of those duplexes can be found in Supplemental Table S2, and the ssDNA used for hybridization are detailed in Supplemental Table S3.

Supplemental Table S2: List of DNA duplex

ID	Strand 1	Strand 2	ΔG (kcal/mole)	Base pairs	# of LNA
ATTO633_FQ_2nt	ATTO633_21nt	3'FQ_23nt	-45.41	21	0
ATTO633_FQ_4nt	ATTO633_21nt	3'FQ_25nt	-45.41	21	0
ATTO633_FQ_6nt	ATTO633_21nt	3'FQ_27nt	-45.41	21	0
ATTO633_FQ_8nt	ATTO633_21nt	FQ_5LNA_8nt	/	8	5
ATTO633_FQ_10nt	ATTO633_21nt	FQ_5LNA_10nt	/	10	5
ATTO633_FQ_14nt	ATTO633_21nt	FQ_3LNA_14nt	/	14	3

Supplemental Table S3: List of ssDNA

ID	Sequence
ATTO633_21nt	/5ATTO633N/TGG TCG TGG GGC AAC TGG GTT
3'FQ_23nt	AAC CCA GTT GCC CCA CGA CCA TT/3AIABkFQ/
3'FQ_25nt	AAC CCA GTT GCC CCA CGA CCA TTT T/3AIBkFQ/
3'FQ_27nt	AAC CCA GTT GCC CCA CGA CCA TTT TTT/3AIBFQ/
FQ_5LNA_8nt	/5IABkFQ/**C**A**C** **GAC** CA
FQ_5LNA_10nt	/5IABkFQ/**CC**C A**C**G **AC**C A
FQ_3LNA_14nt	/5IABkFQ/**T**TG CC**C** CAC GA**C** CA

Bold: Locked Nucleic Acid (LNA)

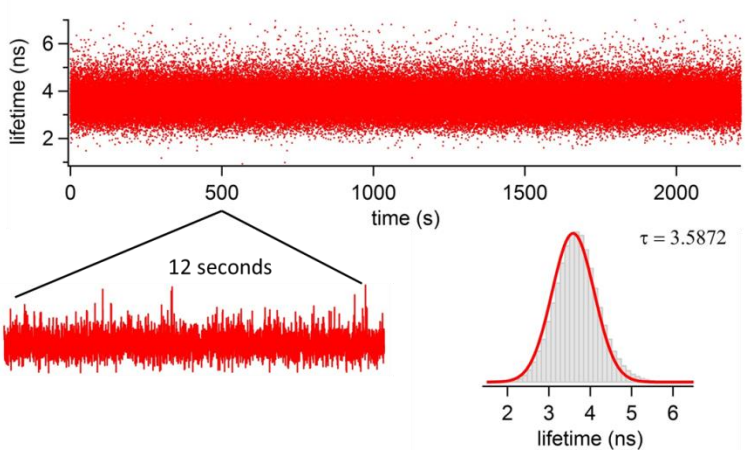
■ : Quencher

■ : Fluorophore

Supplemental Table S4: Summary of measured DNA hybridization kinetics by different labs

Lab	reporter sequence	k_{on} ($\times 10^6 M^{-1}s^{-1}$)	$k_{off}(s^{-1})$	detection environment
Harris[9]	ATGGGATATA (10 nt)	1.64	4.3×10^{-2}	surface
Moerner[10]	TCATACTAA (9 nt)	0.254	3.26	2D solution
Tinnefeld[11]	TAGATGTAT (9 nt)	2.3	1.6	DNA origami (surface)
Nesbitt[12]	GGGTTGGT (8 nt)	3.5	0.72	Surface
Ha[13]	ACAAGTCCT (9 nt)	1.1	0.1	surface
This work	TGGGCGGG (8nt)	5.13	9.55	3D solution

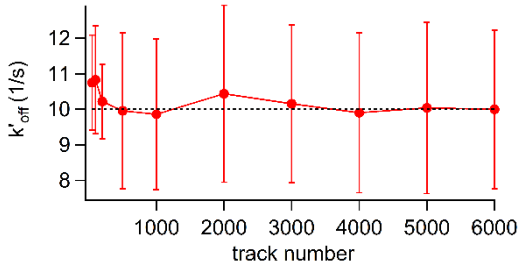
Supplemental Figure S1: Stitched lifetime trace from 6,780 ATTO633-ssDNA molecules (total track duration 2,247 seconds) does not show fluorescence lifetime switching



To show that ATTO633 itself doesn't exhibit any fluorescence lifetime switching, we performed single-molecule tracking of 50 pM reporter strand-1 (/5'ATTO633/TGGGCGGG) in 70 wt % glycerol and 20 mM pH 8.0 Tris-HCl buffer.

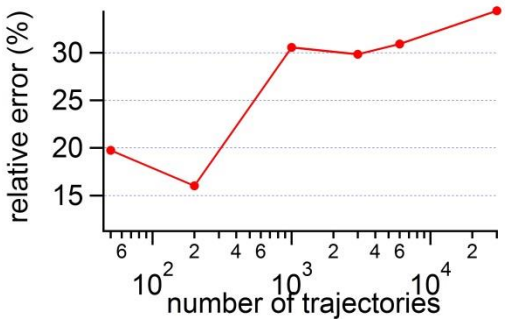
The measured fluorescence lifetime (~3.6 ns) is smaller than the lifetime (~4.16 ns) of ATTO633 in 20 mM Tris-HCl buffer because the addition of glycerol increases the refractive index of the solution. This is in agreement with the Strickler Berg equation relating the fluorophore's radiative rate (k_r) and its absorption and emission spectra[14,15].

Supplemental Figure S2: Estimated k_{off} as a function of number of lifetime traces used for ebFRET analysis



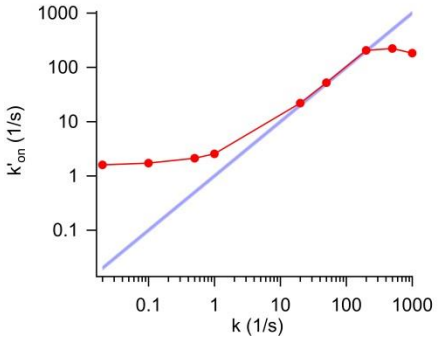
The error bars denote standard deviation. The black dotted line represents the true k_{off}. The relative error of k_{off} estimation is smaller than 4.3% as long as more than 500 lifetime traces are available. In ebFRET GUI, "Restarts" is set to be 2, "Precision" is set to be 1E-6, as recommended by the user manual.

Supplemental Figure S3: vbFRET benchmarking – relative error of estimated k'_{on} as a function of number of lifetime traces



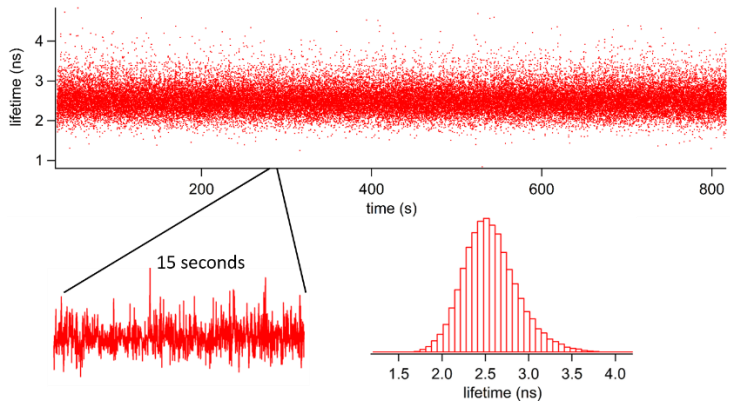
The vbFRET algorithm[16] is benchmarked following a similar approach employed in the last section. k'_{on} and k_{off} are made equal (5 s^{-1}) in this simulation. In vbFRET GUI, "Fitting attempts per trace" is set to be 10. The relative error generally increases with the number of lifetime traces, suggesting vbFRET cannot extract the annealing/melting rate reliably, possibly due to the short tracking durations. The error is always larger than 15%, which is much worse than ebFRET.

Supplemental Figure S4: vbFRET benchmarking – estimated k'_{on} (red) as a function of true k



k'_{on} and k_{off} are made equal in this simulation. The purple shaded error bar represents $\pm$10% relative error of k'_{on} estimation. 3,000 lifetime traces are used for vbFRET analysis. The relative error is smaller than 10% when k is within 20-200 s^{-1}, which is an order of magnitude smaller than the region in which ebFRET can estimate k'_{on} accurately.

Supplemental Figure S5: Stitched lifetime trace from 1,638 dsDNA molecules (total duration 808 seconds) doesn't show quencher induced donor blinking



Dark quenchers are expected to exhibit less complex photophysics due to their substantially reduced excited state lifetimes. However, Holzmeister *et al.* has observed frequent blinking of Cy5 induced by dark quenchers BBQ650 and BHQ-2 in two widely used oxygen scavenger systems (GOC system: glucose, glucose oxidase and catalase, PCD/PCA system: Trolox, Protocatechuate 3,4-dioxygenase (PCD), 3,4-dihydroxybenzoic acid (PCA))[17].

With the presence of donor blinking, the observed dwell time in neither hybridized nor melted state follows the single-rate exponential distribution, so it's incorrect to model the lifetime trace with hidden Markov process. To examine whether ATTO633 blinks when Iowa Black® FQ is the quencher, we performed single-molecule tracking of 50 pM DNA duplex ATTO633_FQ_8nt in 70 wt % dextran, and 25mM Tris-HCl buffer. ATTO633_FQ_8nt is hybridized in-house from two 5'-modified ssDNA: ATTO633_21nt and FQ_5LNA_8nt (see Supplemental Table S2-S3).

We observed no blinking of ATTO633 when Iowa Black® FQ is the dark quencher molecule. This is in accordance with previously published results where the combination of ATTO532 and BBQ650 shows no blinking [17].

References

(1) C. Liu, E. P. Perillo, Q. Zhuang, K. T. Huynh, A. K. Dunn, H.-C. Yeh, "3D single-molecule tracking using one- and two-photon excitation microscopy," *Proceedings of SPIE*, **8950**, 89501C, 2014.

(2) C. Liu, Q. Zhuang, H. C. Yeh, "Three Dimensional Single-Molecule Tracking with Confocal-Feedback Microscope," *9th IEEE International Conference on Nano/Micro Engineered and Molecular Systems (NEMS)*, 481-484, 2014.

(3) B. Lounis, W. E. Moerner, "Single photons on demand from a single molecule at room temperature," *Nature*, **407**, 491-493, 2000.

(4) L. Brand, C. Eggeling, C. Zander, K. Drexhage, C. Seidel, "Single-molecule identification of Coumarin-120 by time-resolved fluorescence detection: Comparison of one-and two-photon excitation in solution," *The Journal of Physical Chemistry A*, **101**, 4313-4321, 1997.

(5) C. Zander, L. Brand, C. Eggeling, K.-H. Drexhage, C. A. Seidel In *BiOS*; International Society for Optics and Photonics: 1997, p 552-558.

(6) J.-W. van de Meent, J. E. Bronson, C. H. Wiggins, R. L. Gonzalez, "Empirical Bayes methods enable advanced population-level analyses of single-molecule FRET experiments," *Biophysical Journal*, **106**, 1327-1337, 2014.

(7) M. Köllner, J. Wolfrum, "How many photons are necessary for fluorescence-lifetime measurements?," *Chemical Physics Letters*, **200**, 199-204, 1992.

(8) R. Das, C. W. Cairo, D. Coombs, "A hidden Markov model for single particle tracks quantifies dynamic interactions between LFA-1 and the actin cytoskeleton," *PLoS Computational Biology*, **5**, e1000556, 2009.

(9) E. M. Peterson, M. W. Manhart, J. M. Harris, "Single-Molecule Fluorescence Imaging of Interfacial DNA Hybridization Kinetics at Selective Capture Surfaces," *Analytical Chemistry*, **88**, 1345-1354, 2016.

(10) Q. Wang, W. Moerner, "Single-molecule motions enable direct visualization of biomolecular interactions in solution," *Nature Methods*, **11**, 2014.

(11) R. Jungmann, C. Steinhauer, M. Scheible, A. Kuzyk, P. Tinnefeld, F. C. Simmel, "Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on DNA origami," *Nano Letters*, **10**, 4756-4761, 2010.

(12) N. F. Dupuis, E. D. Holmstrom, D. J. Nesbitt, "Single-molecule kinetics reveal cation-promoted DNA duplex formation through ordering of single-stranded helices," *Biophysical Journal*, **105**, 756-766, 2013.

(13) I. I. Cisse, H. Kim, T. Ha, "A rule of seven in Watson-Crick base-pairing of mismatched sequences," *Nature Structural & Molecular Biology*, **19**, 623-627, 2012.

(14) K. Suhling, J. Siegel, D. Phillips, P. M. French, S. Lévêque-Fort, S. E. Webb, D. M. Davis, "Imaging the environment of green fluorescent protein," *Biophysical Journal*, **83**, 3589-3595, 2002.

(15) C. Tregidgo, J. A. Levitt, K. Suhling, "Effect of refractive index on the fluorescence lifetime of green fluorescent protein," *Journal of Biomedical Optics*, **13**, 031218-031218-031218, 2008.

(16) J. E. Bronson, J. Fei, J. M. Hofman, R. L. Gonzalez, C. H. Wiggins, "Learning rates and states from biophysical time series: a Bayesian approach to model selection and single-molecule FRET data," *Biophysical Journal*, **97**, 3196-3205, 2009.

(17) P. Holzmeister, B. Wünsch, A. Gietl, P. Tinnefeld, "Single-molecule photophysics of dark quenchers as non-fluorescent FRET acceptors," *Photochemical & Photobiological Sciences*, **13**, 853-858, 2014.