bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Inferring kinetic rate constants from single-molecule FRET trajectories – a blind benchmark of kinetic analysis tools

Markus Götz[1,*], Anders Barth[2,3], Søren S.-R. Bohr[4], Richard Börner[5,6], Jixin Chen[7], Thorben Cordes[8], Dorothy A. Erie[10], Christian Gebhardt[8], Mélodie C.A.S. Hadzic[5], George L. Hamilton[11], Nikos S. Hatzakis[4], Thorsten Hugel[12,13], Lydia Kisley[14,15], Don C. Lamb[9], Carlos de Lannoy[16], Chelsea Mahn[17], Dushani Dunukara[14], Dick de Ridder[16], Hugo Sanabria[11], Julia Schimpf[12,18], Claus A.M. Seidel[2], Roland K.O. Sigel[5], Magnus Berg Sletfjerding[4], Johannes Thomsen[4], Leonie Vollmar[12,18], Simon Wanninger[9], Keith R. Weninger[17], Pengning Xu[17], Sonja Schmid[19,*]

1) Centre de Biologie Structurale, CNRS UMR 5048, INSERM U1054, Univ Montpellier, 60 rue de Navacelles, 34090 Montpellier, France. Current address: PicoQuant GmbH, Rudower Chaussee 29, 12489 Berlin, Germany

2) Institut fu□r Physikalische Chemie, Lehrstuhl fu□r Molekulare Physikalische Chemie, Heinrich-Heine-Universita□t, Universitätsstr. 1, 40225 Du□sseldorf, Germany

3) present address: Department of Bionanoscience, Kavli Institute of Nanoscience Delft, Delft University of Technology, Van der Maasweg 9, 2629 HZ Delft, The Netherlands

4) Department of Chemistry & Nano-science Center, University of Copenhagen, 2100 Copenhagen, Denmark & Novo Nordisk Foundation Centre for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, 2100 Copenhagen, Denmark

5) Department of Chemistry, University of Zurich, 8057 Zurich, Switzerland

6) present address: Laserinstitut Hochschule Mittweida, University of Applied Sciences Mittweida, 09648 Mittweida, Germany

7) Department of Chemistry and Biochemistry, Ohio University, Athens, Ohio, United States

8) Physical and Synthetic Biology, Faculty of Biology, Ludwig-Maximilians-Universität München, Großhadernerstr. 2-4, 82152 Planegg-Martinsried, Germany

9) Department of Chemistry and Center for Nano Science (CeNS), Ludwig Maximilians-Universität München, Butenandtstraße 5-13, 81377 München, Germany

10) Department of Chemistry, University of North Carolina, Chapel Hill, NC 27599, United States; and Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC 27599, United States

11) Department of Physics and Astronomy, Clemson University, Clemson, South Carolina, 29634, United States

12) Institute of Physical Chemistry, University of Freiburg, Freiburg, Germany

13) Signalling Research Centers BIOSS and CIBSS, University of Freiburg, Freiburg, Germany

14) Department of Physics, Case Western Reserve University, Cleveland, Ohio, United States

15) Department of Chemistry, Case Western Reserve University, Cleveland, Ohio, United States

16) Bioinformatics Group, Wageningen University, Droevendaalsesteeg 1, 6708PB Wageningen, The Netherlands

17) Department of Physics, North Carolina State University, Raleigh, NC, 27695, United States

18) Spemann Graduate School of Biology and Medicine (SGBM), University of Freiburg

19) NanoDynamicsLab, Laboratory of Biophysics, Wageningen University, Stippeneng 4, 6708WE Wageningen, The Netherlands

* Correspondence should be addressed to: Markus Götz goetz@picoquant.com and Sonja Schmid schmid@nanodynlab.org

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ABSTRACT

Single-molecule FRET (smFRET) is a versatile technique to study the dynamics and function of biomolecules since it makes nanoscale movements detectable as fluorescence signals. The powerful ability to infer quantitative kinetic information from smFRET data is, however, complicated by experimental limitations. Diverse analysis tools have been developed to overcome these hurdles but a systematic comparison is lacking. Here, we report the results of a blind benchmark study assessing eleven analysis tools used to infer kinetic rate constants from smFRET trajectories. We tested them against simulated and experimental data containing the most prominent difficulties encountered in analyzing smFRET experiments: different noise levels, varied model complexity, non-equilibrium dynamics, and kinetic heterogeneity. Our results highlight the current strengths and limitations in inferring kinetic information from smFRET trajectories. In addition, we formulate concrete recommendations and identify key targets for future developments, aimed to advance our understanding of biomolecular dynamics through quantitative experiment-derived models.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

How does biomolecular function arise from structural dynamics? This largely unsolved question is central for the understanding of life at the molecular scale. However, the transitions between various conformational states have remained challenging to detect, quantify, and interpret. Over the past two decades, single-molecule Förster resonance energy transfer (smFRET) detection has emerged as a powerful technique to study the dynamics of single biomolecules under physiological conditions using fluorescence as a readout[1]. A unique aspect of smFRET is its ability to link space and time, i.e., to connect structural with kinetic information under both equilibrium and non-equilibrium conditions, which is often unachievable using ensemble methods. By measuring the distance-dependent energy transfer from a donor to an acceptor fluorophore, distances in the range of 4 to 12 nm can be measured with sub-nanometer precision and accuracy[2]. Various experimental implementations exist that allow one to measure smFRET on diverse timescales from picoseconds to hours. All of this makes smFRET an ideal tool in the growing field of dynamic structural biology[3].

To study conformational dynamics of one single molecule for an extended time (seconds to minutes), dye-labeled biomolecules are most commonly immobilized on passivated glass slides and imaged using camera-based brightfield detection, or confocal detection using avalanche photodiodes[2]. The resulting fluorescence time traces have a typical time resolution of about 10 ms - 100 ms, and the observation time per single molecule is limited by photobleaching, leading to an average bandwidth of less than three orders of magnitude in time[4–6]. Conformational transitions of the biomolecule change the inter-dye distance leading to discrete steps in the fluorescence signal and the FRET efficiency (**Fig. 1**). This desired time-resolved distance information is convoluted with largely Gaussian noise in the experiment (from autofluorescence background, detector noise, laser fluctuations, etc.). Moreover, noise and photobleaching are intrinsically coupled: increasing the laser power for a better signal-to-noise ratio causes faster photobleaching, which reduces the temporal bandwidth of the experiment. As a result, signal interpretation in terms of biomolecular states and specific transitions between them is not trivial.

A multitude of analytical approaches have been developed to infer the number of functional states and quantify kinetic rate constants from noisy experimental data. Frequently, hidden Markov models (HMMs)[7] are used to infer an idealized state sequence from which dwell-time distributions are compiled, which are then fit (with exponentials) to obtain kinetic rate constants[8,9]. Alternatively, the transition matrix that is part of every HMM can directly be converted to kinetic rate constants. The HMM formalism is based on a discrete memoryless Markov process that infers a set of parameters (probabilities of states, transitions, and observations) to describe the observed sequence of FRET efficiencies. Many extensions of the HMM formalism have been developed[10–15] including Bayesian approaches[16–19], and very fast kinetics (low energy barrier crossings) can be inferred from single-photon arrival times[54,20,55].

Often, multiple input models are compared based on a scarcity criterion to avoid bias in the selection of the optimal model size (i.e., the number of states and rate constants), and hence the number of free parameters[8,21–23]. Other analysis approaches, such as correlation analysis[24–28] and discretization

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

methods based on cluster analysis[29–32], may treat the raw data in a model-free way while the extraction of individual rate constants (rather than residence times only) still relies on an initial guess of a model. The growing number of analytical methods renders it increasingly difficult to decide on the optimal tool for a specific application and to judge whether the described improvements justify the time cost of implementation. Hence, it was identified during a round table discussion of the smFRET community (Fluorescence subgroup, Biophysical Society Meeting 2019, Baltimore, US) that a critical assessment of the available tools is needed.

Here, we present the results of the first comparative multi-laboratory study that provides a systematic evaluation of eleven analysis tools (summarized in **Box 1**) using simulated as well as experimental data of varied complexity. Three of the analysis tools were utilized under different conditions, leading to a comparison of 14 different analyses. While clearly not all existing analysis tools could be covered (new tools are released continuously and not all authors decided to participate), this blind study (illustrated in **Fig. 1**) allows us for the first time to directly assess the performance of the different analysis approaches for the inference of kinetic information from single-molecule FRET trajectories and to identify their strengths and weaknesses. Specifically, we assess the accuracy of the inferred kinetic model (i.e., the kinetic rate constants and their connectivity) plus the associated uncertainties, and this for kinetic models of varied size, from the simple case of a two-state system (**Figure 2**) to the more complex case of a non-equilibrium three-state system (**Figure 3**), and finally to degenerate multi-state systems (**Figures 4, 5**). All analyses were performed by the expert labs of each tool to ensure optimal implementation (see Methods for details).

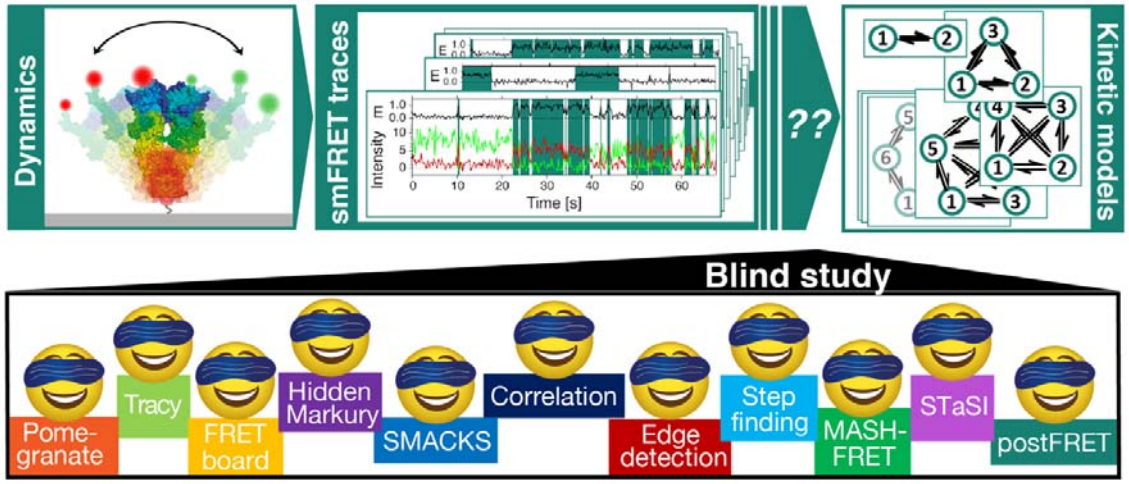

Figure 1 | This blind study reports on the performance of diverse analysis tools to describe single-molecule dynamics with quantitative kinetic rate constants. Biomolecular dynamics of proteins and nucleic acids can be detected by smFRET and other single-molecule techniques. Extracting testable kinetic rate models from the experimental time traces is complicated by experimental shortcomings. Multiple labs joined forces to directly test the performance of diverse analytical approaches to infer kinetic rate constants in a blind study.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Box 1: Overview of the kinetic analysis approaches and software tools considered herein, grouped based on similarity. All tools are detailed in the Supplementary Methods.

#	Tool name	Key-words	Description	Model selection	Uncer-tainty
1	Pome-granate	HMM	The python package Pomegranate is used for efficient and iterative modeling, fitting and evaluation of state numbers using the BIC. Dwell time analysis is subsequently performed after defining all transitions using a multivariate Gaussian fitting scheme and unbinned maximum likelihood fitting. (CI: confidence interval)	BIC	95 % CI
2	Tracy	HMM	Global HMM analysis was performed while setting the FRET efficiency and sigma as parameters to be learned. The state transitions and the state dwell times are selected by the user in a transition density plot and fitted with an exponential to obtain the rate constants.	Manual inspection	95% CI
3	FRETboard[33]	Semi-supervision, remotely served	A semi-supervised classification tool served remotely through a browser window. Users supervise the training of a classification model of choice, by manually correcting classification of example traces until the quality of automated classification is satisfactory.	Path probability	95% CI
4	Hidden-Markury	1D/2D-HMM	Hidden-Markury is a trace analysis software based on an interactive Jupyter notebook script, supporting global 1D FRET efficiency traces or 2D donor & acceptor photon streams, optionally treating degenerate states, forbidden transitions, fixed model parameters.	BIC	Sub-sampling
5	SMACKS[13]	1-3D-HMM	Semi-ensemble HMM is used to extract one kinetic model from many smFRET fluorescence traces without prior discretization in two steps: (1) per trace HMM optimization (2) global per dataset optimization of the kinetic model, with pre-trained intensity parameters.	BIC	95% CI
6	SMACKS (SS)	1-3D-HMM	Test for user bias in semi-supervised inference: independent second analysis using SMACKS by S. Schmid.	BIC	95% CI
7	Correlation	Discretized correlation	An unbiased, model-independent approach to obtain quantitative relaxation times from the negative amplitude of the cross-correlation function[24,25]. To enable a quantitative analysis of multi-state systems, a filtered correlation analysis[34] is performed based on the state sequence obtained with a step-finding algorithm[35].	BIC	95%
8	Edge finding (CK)	CK filter	The Chung Kennedy non-linear filter is applied to the time records of donor, acceptor and/or FRET efficiency to identify state transition points as sudden increases in the standard deviation of points in forward/backward predictor windows. Transition edges are confirmed by a two-sample student's t-test on the forward/backward windows.	Manual inspection	Not assessed
9	Edge finding (k-means)	k-means clusters	All data points in either the donor and acceptor or the FRET efficiency time traces are assigned to distinct clusters. The mean value of each cluster is calculated and the points are reassigned to clusters to iteratively minimize the differences between the point values. Transition edges are identified as cluster assignment changes.	Manual inspection	Not assessed
10	Step finding	Line fitting	The entire dataset is iteratively fit with an increasing number of line segments. The addition of line segments is accepted if the overall fit quality is improved significantly. Rate constants are derived from dwell time analysis of line segments, which are assigned to a FRET state based on their mean FRET efficiencies. (SEM: standard error of the mean.)	Manual inspection	SEM / 68% CI
11	STaSI	Student's t-test	Detects step transition using the Student's t-test. Segments are grouped into states by hierarchical clustering. The optimum number of states is established using a minimum description length equation that sums the goodness of fit measured using the L1 norm to consider the sparseness of the states and transitions. (MDL: minimum description length.)	MDL	95% CI
12	MASH-FRET (bootstrap)[36]	STaSI, vbFRET, bootstrap	A MATLAB-based GUI for the simulation and analysis of smFRET videos and fluorescence time traces[52]. Initial FRET states are obtained using STaSI and a BIC selection on 2D-Gaussian mixtures that model the global transition density plot. Refined FRET states, transition rate constants and uncertainties are then obtained using vbFRET and single exponential fit on bootstrapped dwell time histograms.	BIC	95%
13	MASH-FRET (prob.)[53]	STaSI vbFRET DPH HMM	The degeneracies of FRET states are estimated from ensemble dwell time histograms by performing a BIC selection on phase-type distributions. The fix-sized transition rate matrix is finally optimized using the Baum-Welch algorithm on hard-assigned FRET state trajectories.	BIC	95%

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

14	postFRET	Monte Carlo	Simple thresholding is used for an initial assessment of the rate constants. A computationally-intensive Monte Carlo simulation is then used to find simulated trajectories that contain the same rate and error pattern as the experimental ones to guess a possible truth. Compare the two and adjust the guess for the next iteration. Noisy data is binned for the initial thresholding. (LAD: least absolute deviations.)	LAD	68% CI

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

RESULTS

The archetypal 2-state system

We first consider the simplest case of a kinetic 2-state system, which could represent alternation between two conformations of a biomolecule in dynamic equilibrium, or transient biomolecular interactions. The kinetics of this system are described by two rate constants (**Figure 2a**). In a blind study, we analysed simulated and experimental smFRET data using the diverse set of analysis tools summarized in **Box 1** and detailed in the Supplementary Methods. Simulated test data (described in Supplementary Note 1) has the advantage that the underlying ground truth (GT, i.e., the simulation input) is known, which facilitates the evaluation of the inferred results, while, for experimental data, the GT is naturally not known. **Figure 2b** depicts an example of the simulated traces. We note that it closely resembles the experimental trace in **Figure 2e.** Based on a dataset of such simulated traces (n=75), all laboratories inferred FRET efficiencies (**Fig. 2c**) and rate constants (**Fig. 2d**), which agree very well: the FRET efficiencies deviate by less than 17 % from the GT, and the inferred rate constants deviate with a maximum of 12 % from the GT (5 % average deviation), with a slight systematic underestimation in most cases, i.e., the determined rate constants were slower. *Pomegranate*, *FRETboard*, and *Step finding* infer the most accurate rate constants under the tested conditions (**Table** 1). The equilibrium constants $K=k_{21}/k_{12}$ vary generally less since systematic deviations balance each other in this case (Supplementary Fig. 1a). In contrast, the reported uncertainty measures vary greatly, independent of the analysis type (0.4 % to 21 % relative to the inferred rate constant). For comparison, we estimated the minimal uncertainty given the finite size of the dataset, by quantifying the standard deviation of the rate constants obtained from one million simulated samples (Supplementary Note 2). This standard deviation is ≥3 % of the rate constants for the provided dataset (gray and light gray bars in Fig. 2b shown for 1σ and 2σ, respectively). Thus, most analysis tools reported reasonable uncertainty estimates, while some tools reported uncertainties that are smaller than this lower limit (*Tracy*, *Correlation*, *STaSI*) or provided no uncertainty measures (*Edge finding*). *FRETboard* reported consistently very large uncertainties, which was solved in their latest software version 0.0.3 (Ref[33]). *Step finding* initially found erroneously large uncertainties that have been corrected in the latest software version. We note that various methods are currently in use for estimating uncertainties which complicates the direct comparison.

Next, we consider experimental data (see Methods), which naturally contains all typical noise sources and experimental artefacts (**Fig. 2e-g**). As there is no GT for experimental data, we assessed the consistency of the inferred FRET efficiencies and rate constants using the coefficient of variation (CV, i.e., the standard deviation divided by the mean). We found excellent agreement for all inferred FRET efficiencies (CV ≤ 2 %). The rate constants vary by 12 % and 16 % (CV for k_{12} and k_{21}, respectively), consistent with the variation found for simulated data (**Fig. 2d**). Again, no correlation of the rate constants with respect to the analysis approach is evident, but the tendency of a given tool for large or small uncertainties is conserved (**Fig. 2d and 2g**), with *FRETboard* and *Step finding* reporting the largest uncertainties, and *STaSI*, *MASH-FRET* (prob.), *postFRET*, and *Correlation* the smallest uncertainties. In most cases, the equilibrium constants (Supplementary Fig. 1b) agree well with each

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

other and with the equilibrium populations of the FRET histogram, while some results are inconsistent with the latter (*Hidden-Markury*, *Correlation*, *STaSI*, and *postFRET*).

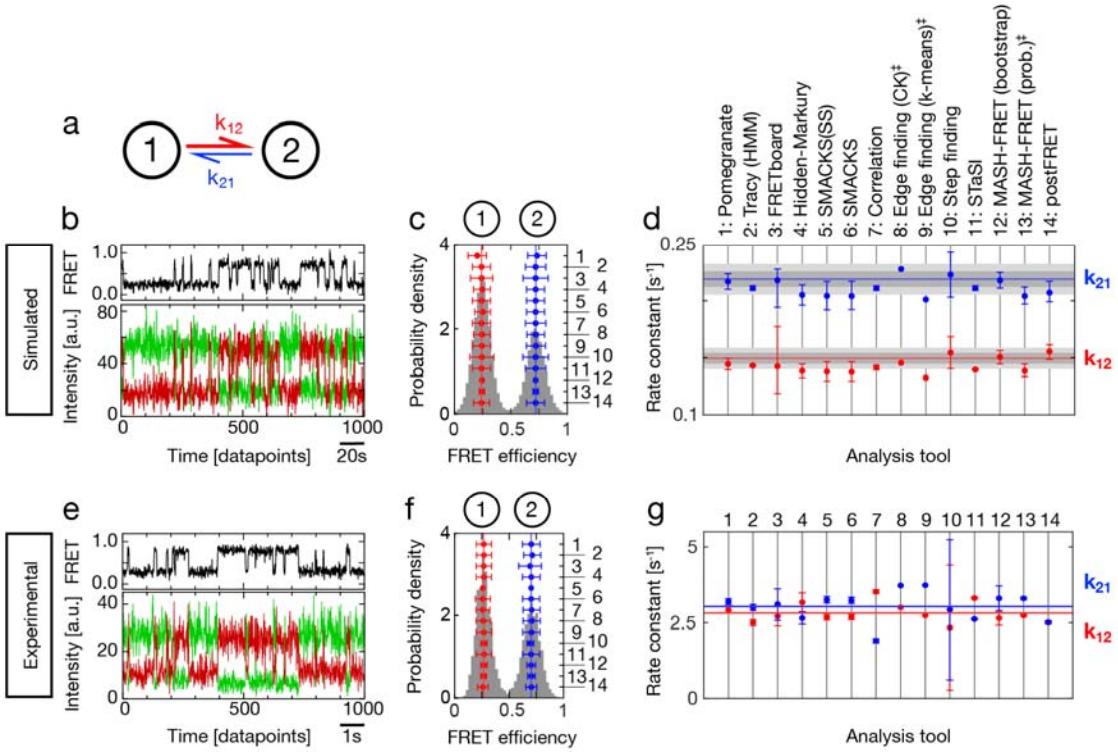

Figure 2 | Quantification of simulated and experimental kinetics between two states. a Illustration of the kinetic model with two states (circles) connected by forward and backward rate constants: k_{12} and k_{21}. **b** A simulated FRET trace showing the donor and acceptor fluorescence intensity (green, red) and the FRET efficiency (FRET, black), representative for the dataset used in c, d: n(traces)=75, n(datapoints)=59,486, sampling rate=5 Hz. **c** FRET efficiency histogram (gray) with assigned states on top and inferred FRET efficiencies in red and blue. Numbers on the right axis refer to the analysis tools specified in d. Vertical lines indicate the mean over all tools. **d** Rate constants and uncertainties inferred from the dataset in c by different labs using the respective analysis tools. The ground truth (GT) is indicated by horizontal red and blue lines, the intrinsic uncertainty of the dataset (see text) is represented by dark gray (1σ) and light gray (2σ) intervals. **e** An experimental time trace with colors as in b, representative for the dataset used in f, g with n(traces)=19, n(datapoints)=226,100, using 10 ms time bins resulting in 100 Hz sampling, kindly provided by B. Schuler. **f** FRET histogram with color code and axis labels as in c. No uncertainties were submitted for tool #5. **g** Inferred rate constants from the experimental dataset in f. Color code as in d. Horizontal red and blue lines indicate the mean of the inferred rate constants. Supplementary Fig. 2 shows the experimental data and analysis with ten times higher time resolution. ‡ denotes results that were submitted after the GT was known. The model size was restricted to two states. *FRETboard* and *Step finding* found erroneously large uncertainty intervals, which has been corrected in their latest software versions.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 1: Summary of the test conditions for the individual datasets, including the prior information on ground truth (GT) and number (N) states, as well as three data characteristics: kinetic heterogeneity, photo-physics, and signal-to-noise ratio (SNR). [a] Exceptions are labelled with a dagger in all figures. [b] See simulation parameters in Supplementary Table N1. [c] The SNR was obtained from the FRET efficiency histogram using Gaussian fits and SNR = $|\mu_1 - \mu_2| / \sqrt{{\sigma_1}^2 + {\sigma_2}^2}$.

	GT known? [a]	N states predefined?	Kinetic heterogeneity	Photo-physics [b]	SNR [c]
Figure 2 (sim.)	No	Yes, 2.	No	Clean	4
Figure 2 (exp.)	No	Yes, 2.	Not observed	Mainly clean	4
Figure 3 (sim.)	No	No	No	Intensity variation	3
Figure 4 (sim.)	No	No	Yes	Intensity variation & blinking	4
Figure 5 (exp.)	No	No	Yes observed	Intensity variation & blinking	3

One important factor in dynamic smFRET data is the signal-to-noise ratio (SNR), which depends on the acquired signal per data point and can be controlled by the integration time (also known as exposure time). We explicitly tested the effect of a ten-fold shorter integration time. On the one hand, this offers better sampling of fast kinetics due to the increased time resolution (1 kHz instead of 0.1 kHz sampling), but, on the other hand, it results in a lower signal-to-noise ratio which is more challenging for state identification. In addition, at 1 kHz sampling, the data shows single-photon discretization and non-Gaussian noise (Supplementary Fig. 2a,b), thus deviating from the basic assumptions underlying most of the considered analysis tools. Indeed, the overall agreement of the rate constants at this lower SNR was reduced: CV= 33 % and 45 % for k_{12} and k_{21}, respectively (Supplementary Fig. 2c), indicating that the benefit of the increased time resolution is minor in this case. Nevertheless, the equilibrium constants agree very well again (CV= 2 %, when excluding the two clear outliers in Supplementary Fig. 2d) due to the cancelation of systematic shifts for both rate constants (Supplementary Fig. 2e). Comparing the rate constants inferred at 1 kHz and 0.1 kHz sampling, *pomegranate*, *Tracy*, *Correlation*, *MASH-FRET*, and *Step finding* reported similar values (Supplementary Fig. 2e), while *STaSI* inferred slower rate constants for faster sampling. Conversely, *FRETboard*, and *SMACKS* inferred faster rate constants for faster sampling, either due to fitting noise or due to short events that are missed at lower time resolution. The latter is less plausible, given that the inferred rate constants are 20-fold smaller than the 0.1 kHz sampling rate. Thus, a comparison between 0.1 kHz and 1 kHz sampling can serve to estimate the robustness of the analysis tools towards non-Gaussian noise. Taken together, fundamentally different analysis approaches inferred consistent rate constants and FRET efficiencies from a simple, two-state system both for simulated data and experimental data with varied SNR.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Directional sequences in a non-equilibrium steady-state system

Many biomolecular systems involve more than just two functionally relevant states, leading to more intricate kinetic models with more rate constants and, hence, more degrees of freedom. Such systems with three or more states can show a conceptually unique thermodynamic phenomenon: the non-equilibrium steady-state, in which a biomolecule, such as a motor protein or a molecular machine such as F_0F_1-ATP synthase, is driven by continuous external energy input, e.g. in the form of a chemical gradient[37], light[38,39], or ATP. As a result, conformational states may appear in a preferred sequence order, causing a non-zero net flow, e.g. for the 3-state system depicted in **Fig. 3a**:

$$\Delta G_{1\to2\to3\to1} = -k_B T * ln\left(\frac{k_{21}\cdot k_{32}\cdot k_{13}}{k_{12}\cdot k_{23}\cdot k_{31}}\right) \neq 0 \quad (1)$$

The unique ability to directly observe the non-equilibrium steady-state is a prime example of the merits of single-molecule studies. Hence, we investigated it explicitly, using smFRET data simulated with a kinetic 3-state model and a non-zero counter-clockwise flow: $\Delta G_{1\to2\to3\to1} < 0$ (**Fig. 3a,b**). As an additional challenge, this data contained fluorescence intensity variation between individual dye molecules, as observed in experimental data due to varied local dye environment and orientation, inhomogeneities in excitation intensity and polarisation, and also variations in detection efficiency[40].

All analysis tools found the three clearly separated FRET efficiency populations (**Fig. 3c**), while the inferred rate constants varied more than for the 2-state systems above (**Fig. 3d**). Most tools systematically underestimated k_{13} and k_{31} and overestimated all other rate constants. This may be attributed to the inevitable effect of time discretization and related intensity averaging: when a transition between the high- and low-FRET states happens during a time bin, time-weighted averaging (camera blurring) of the FRET efficiencies occurs, leading, in some cases, to mid-FRET observations that are indistinguishable from those caused by a *bona fide* biomolecular conformation. While, at the single datapoint level this discretization artefact cannot be prevented, the inference accuracy may be improved by treating discretization-induced averaging explicitly in the analysis[31,41]; or using pulsed illumination to reduce blurring[42,43]. Overall, *postFRET* and *Tracy* inferred the most accurate rate constants with average GT deviations of 9 % and 14 %, respectively. As shown in **Fig. 3e**, qualitatively, the net flow was correctly resolved (most accurately by *postFRET*, *Tracy*, and *FRETboard*), while quantitatively it was mostly underestimated, which we attribute to the aforementioned systematic misallocation of transitions between states 1 and 3. For this simulated dataset, the theoretical lower limit of the uncertainty (as introduced above for the 2-state system) is smaller because the dataset is larger. About half of the tools reported uncertainties that are in line with this lower limit (grey intervals), while the other half reported none or too small uncertainties. Altogether, the rate constants of the non-equilibrium 3-state system with intensity variation were less accurate than those of the 2-state system, and also the uncertainty estimation was challenging in this case. Nevertheless, the steady-state flow was qualitatively well resolved by most tools.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

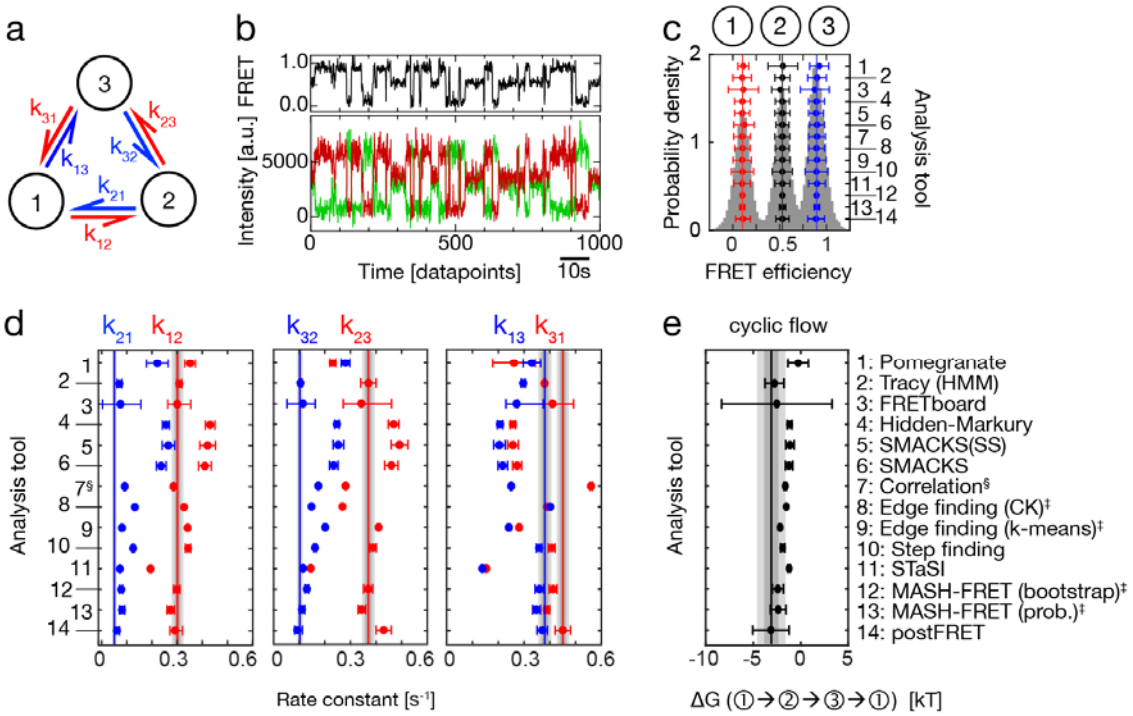

Figure 3 | Quantitative analysis of a non-equilibrium steady-state system. a An illustration of the simulated three-state model with a counter-clockwise net flow. States (circles) are connected by forward and reverse rate constants as specified. **b** A simulated smFRET trace with donor and acceptor fluorescence intensity (green, red) and FRET efficiency (FRET, black), representative for the dataset used in c, d, e: n(traces)=150, n(datapoints)=82,594, sampling rate=10 Hz. **c** SmFRET histogram overlaid with the inferred FRET efficiencies (right axis, numbers as in e) and assigned states on top. **d** Inferred rate constants are shown in red and blue as specified. Vertical lines indicate the GT. The intrinsic uncertainty of the dataset is represented by dark gray (1σ) and light gray (2σ) intervals. Analysis tools are numbered as in e. **e** The inferred cyclic flow in the counter-clockwise direction determined by calculating ΔG from Equation (1) and compared with the GT value (solid vertical line). The uncertainty intervals (dark and light gray) are plotted as in d. Additional simulations to validate the dataset are shown in Supplementary Fig. 3. ‡ denotes results that were submitted after the GT was known. Edge finding did not report uncertainties. § denotes that the misassignment of start and end states was corrected after the GT was known.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

States with overlapping FRET efficiencies

Many biological systems show multi-exponential dwell-time distributions with long and short dwell times for the same apparent FRET state[44–46,6]. This can, for example, arise when the one-dimensional reaction coordinate spanned by the FRET pair is not sufficient to uniquely identify structural states in 3D space. Such kinetic heterogeneity is difficult to interpret because transitions between states with identical or overlapping FRET efficiencies cannot be directly observed in the recorded time traces, while they can often be inferred kinetically. To investigate this case, we simulated kinetic heterogeneity based on a four-state model (**Fig. 4a**) where states 1 and 2 have the same low-FRET efficiencies, and states 3 and 4 have the same high-FRET efficiencies. Again, the fluorescence traces included intensity variations between FRET pairs as observed in the experiment (introduced in the previous section), and also donor and acceptor blinking was included, as an additional imperfection of the data. **Fig. 4b** shows example traces from the simulation and **Fig. 4c** shows the FRET efficiency histogram with two peaks. Without *a priori* knowledge of the model size, most tools identified the correct number of two apparent FRET states, while *FRETboard* used three FRET states to describe the data. *Edge finding* was not developed to deal with such kinetic heterogeneity, and *Pomegranate*, *Correlation, STaSI* and *MASH-FRET (bootstrap)* reported FRET efficiencies but no kinetic models. In the following, we use cumulative dwell-time distributions derived from each inferred model (**Fig. 4d**, detailed in Supplementary Note 3) to compare models with the correct number of FRET states but differences in the kinetic model, such as the connectivity of states or the number of hidden states (rate constants of all inferred models are reported in the Supplementary Table 1, and in the Supplementary Datafiles). Out of the seven independently inferred kinetic models, the two models without kinetic heterogeneity (by *Step finding* and *postFRET*) show the largest deviations from the GT, as these models cannot reproduce the multi-exponential nature of the dwell-time distribution. On the other hand, the four models inferred by the HMM-based *Hidden-Markury, SMACKS, SMACKS(SS),* as well as *MASH-FRET (prob.)* show good agreement with the GT and overlay the GT in the low- and high-FRET case (compare **Fig. 4d**). A quantitative comparison of these four models and their uncertainties is provided in Supplementary Figure 4. While model selection remains a main challenge in inferring kinetic information from smFRET trajectories, it is encouraging that several analysis tools can deduce kinetic models that closely reproduce the GT even under difficult conditions involving kinetic heterogeneity.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

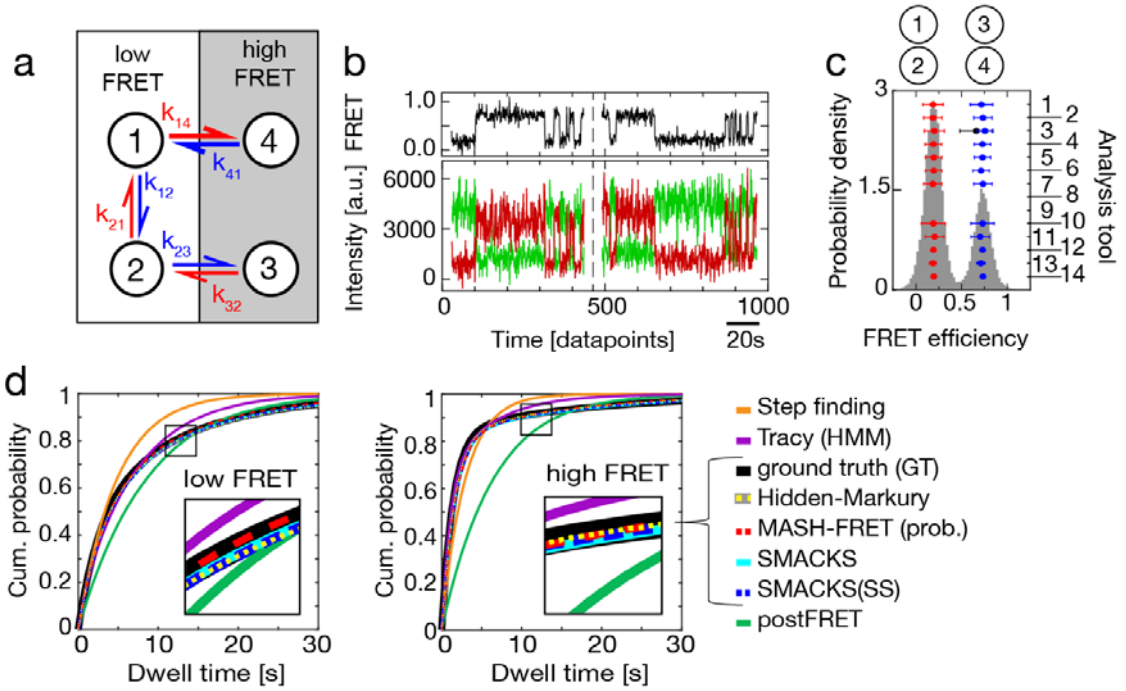

Figure 4 | Resolving kinetic heterogeneity: states with indistinguishable FRET efficiencies but different kinetics. a An illustration of the simulated GT model with states (circles) connected by forward and reverse rate constants. States 1 and 4 as well as states 2 and 3 have indistinguishable FRET efficiencies, causing kinetic heterogeneity. **b** Two simulated FRET traces offset in time with donor and acceptor fluorescence intensity (green, red) and FRET efficiency (FRET, black) are shown, representative for the dataset used in c, d: n(traces)=250, n(datapoints)=56,794, sampling rate=5 Hz. **c** FRET histogram with inferred FRET efficiencies overlaid (right axis: legend as in Box 1 and in all Figures). **d** Comparison of cumulative dwell time distributions derived from the kinetic models with two FRET states (detailed in Supplementary Note 3). The GT histogram is shown as a bold black line. Insets show zoomed-in views of the data indicated by the squares. The rate constants of the four models that most closely reproduce the GT are quantitatively compared in the Supplementary Figure 4. All inferred values of all models are reported in the Supplementary Tables 1 and in the Supplementary Datafiles. ‡ denotes results that were submitted after the GT was known. No results were reported by *Edge finding*. Participants were informed that kinetic heterogeneity may be involved, but not in which configuration.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Full complexity of a black-box experiment

Encouraged by the previous results, we tested all tools vis-à-vis the full experimental complexity to see if they perform similarly as in the simulated case (**Fig. 4**). Three experimental datasets of the same biological system (protein binding to a fluorescently labelled DNA, see Methods), under different experimental conditions and thus different kinetic behaviour, served as a test case. However, the analysts had no prior information on the molecular system causing the dynamics. This means that all the effects discussed so far could potentially be present in these experimental datasets: multiple FRET states, diverse noise sources, fluorophore blinking, directional steady-state flow and kinetic heterogeneity. In addition, the fluorescence intensity variation between single molecules was particularly high in these datasets (see **Figure 5a,d,g**), which complicated the inference of the number of states and rates involved (subsequently referred to as model selection). Under these complex conditions, the inferred number of FRET states (**Fig. 5b,e,h**) varied more than in the simulated case (discussed in **Fig. 4**). Most tools found two FRET states (**Fig. 5b,e,h,** some of them including kinetic heterogeneity), but also three, four, or more different FRET states were reported (Supplementary Figure 5), and the kinetic rate constants varied accordingly. Given the inherent lack of GT information in experimental data, we cannot quantitatively assess the accuracy in this comparison. To balance this fact, we qualitatively compare the inferred results for all three datasets. The 6-7 models with two FRET states are compared in (**Fig. 5c,f,i)**. Other models with three, four, or more FRET states are compared in Supplementary Figures 5-7. (All inferred rate constants are given in Supplementary Tables 2-4 and Supplementary Datafiles). Again, we use cumulative dwell-time distributions (cf. last section) derived from each inferred model (**Fig. 5c,f,i**) to compare models with the same number of FRET states but possibly different state connectivity. The distributions are thus single- or double-exponential depending on the reported kinetic model. The five tools that inferred two FRET states and qualitatively similar kinetic models under all three conditions despite different analysis approaches, are the HMM-based *Hidden-Markury* and *SMACKS*, as well as *Step finding*, *postFRET and MASH-FRET (prob.)*. While *postFRET* consistently inferred slower rate constants, the qualitative agreement among the other five tools is surprisingly good (CV ≤ 25 % for the average residence time) despite the complexity of the input data, the missing prior knowledge about the system, and the different analysis approaches used.

Altogether, we conclude that model selection and state allocation are currently the key challenges in the analysis of kinetic data. In this study, we focused only on the analysis of fluorescence intensity and FRET efficiency data. The addition of complementary information from simulations or experiments (e.g., static molecular structures and other observables, such as fluorescence lifetimes, anisotropy, and more) may help to elucidate complicated or otherwise underdetermined systems[28,47,48].

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

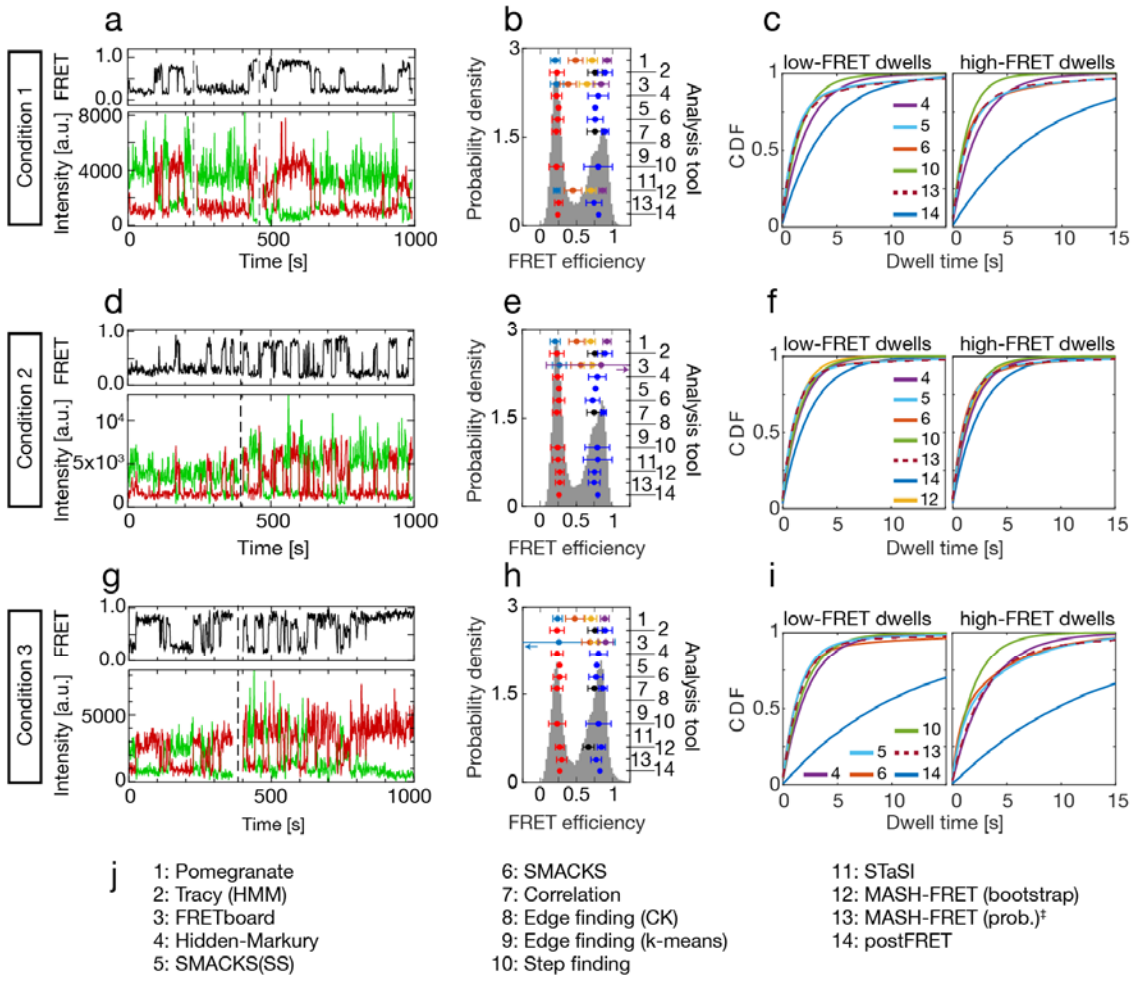

Figure 5 | Increased experimental complexity. Results inferred from three experimental datasets where naturally no GT exists. **a,d,g** Experimental traces, offset in time and separated by dashed vertical lines, with donor and acceptor fluorescence intensity (green, red) and FRET efficiency (FRET, black), representative for the datasets used in (b,c), (e,f), (h,i), respectively, with n(traces): 134, 163, 118; and n(datapoints): 36,604, 37,067, 43,512; sampling rate=33 Hz. All three datasets were kindly provided by M. Schlierf. **b,e,h** FRET efficiency histograms and FRET efficiencies inferred by the analysis tools numbered as in j. For clarity, only the smallest reported model is shown for each analysis tool, up to a maximum of four FRET states. All inferred FRET efficiencies are shown in Supplementary Figure 5, and all inferred results are provided in the Supplementary Tables 2-4 and in the Supplementary Datafiles. Purple arrow in (e): the error bar extends to 1.61. Teal arrow in (h): the error bar extends to -0.53. **c,f,i** Cumulative distribution functions (CDF) of the dwell-times simulated using the inferred kinetic models with two FRET states, obtained with the tools numbered as in j. **j** Legend with all analysis tools. No results were reported by *Edge finding*. ‡ denotes results that were submitted after all other results were known.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

DISCUSSION

In this blind study, we compared eleven kinetic analysis tools for the inference of quantitative kinetic rate constants based on single-molecule FRET trajectories. We explicitly considered the major (kinetic) challenges that the single-molecule experimentalists are typically confronted with: determining the best model to describe the data, especially with multiple FRET states, a varying signal-to-noise ratio, directional non-equilibrium steady-state flow, and kinetic heterogeneity (i.e., states with indistinguishable FRET efficiency but distinct kinetics). We assessed the inferred FRET efficiencies, rate constants, and the reported uncertainties, based on three simulated datasets and four experimental datasets from two biological systems measured using two different setups in different laboratories. The simulated data allowed us to directly assess the accuracy of the inferred rate constants using the known ground truth model and to judge the plausibility of the reported uncertainty measures, while the experimental data shows the relevance and validity of this study.

We found that the number of states was correctly inferred by all tools, as long as their FRET efficiencies were clearly separated (**Figs. 2 and 3**). In the presence of kinetic heterogeneity with overlapping FRET states, model selection was more challenging (**Fig. 4**). In this case, three tools successfully inferred models that accurately reproduce the dwell-time distribution of the GT despite overlapping FRET states (*Hidden-Markury*, *MASH-FRET*, *SMACKS*). In general, the accuracy of the rate constants inferred by all tools decreased with increasing model size and complexity, where time discretization artefacts and inter-trace intensity variation become increasingly challenging. The equilibrium constants and steady-state flow were more accurately inferred than individual rate constants due to the cancellation of systematic errors (Supplementary Figs. 1 and 2d,e, **Fig. 3**). Caution is advised with the uncertainties of rate constants since different uncertainty measures are reported by different approaches. Even for small models (**Figs. 2 and 3**), we found that some uncertainty estimates were smaller than the uncertainties caused by the finite dataset size, while interestingly, more plausible uncertainties were reported for the more complex model in **Fig. 4** (Supplementary Fig. 4). In general, the comparison of uncertainties is complicated by the fact that no common standard exists and the mathematical interpretation of the reported uncertainty intervals differs from tool to tool.

When comparing various analysis frameworks, model-free approaches are generally considered advantageous for an unbiased data analysis. However, HMM-based tools (that compare several input models based on scarcity criteria) were found to be more robust towards data heterogeneity (**Figs. 4 and 5**, Supplementary Fig. 2). Nevertheless, we did not observe a clear overall clustering of the inferred rate constants with the underlying analysis framework, likely due to differences in the data handling beyond the used algorithms (e.g. supervised, semi-supervised, or unsupervised inference). The total analysis durations (processing and computation) ranged from a few minutes to several hours depending on the analysis tool and the model size, with *StaSI* and *StepFinding* ranking among the fastest, and *SMACKS* among the slower tools. In the course of this study, multiple conceptual oversights could be found and solved in a number of tools, which is a direct constructive result of this

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

collaborative comparison study that led to the general recommendations stated in **Box 2**. Additionally, a simple shareable smFRET data format was introduced (Supplementary Note 4) and utilized by all twelve labs working in diverse software environments. We anticipate that this data format will facilitate future collaborations and significantly lower the barrier for an experimentalist to adopt a newly developed analysis tool if it supports the accepted format.

Box 2: General recommendations for users and developers of kinetic inference tools.

In the course of this study, several difficulties with the analysis of kinetic data have become apparent. Out of this experience, we have compiled a list of recommendations for those developing and using kinetic analysis tools.

(i) As a general consistency test, the inferred kinetic model (connectivity and rate constants) can be simulated and the output of the simulation compared to the original input data. For example, the simulator used herein is publicly available as a simple and powerful (Matlab) tool to test whether the proposed model can generate data analogous to the original input, e.g. regarding FRET histogram, smFRET traces, etc.

(ii) Potential biases in the analysis (e.g. regarding model size, state occupation, etc.) can be revealed by subjecting the re-simulated data (with known ground truth) to the same analysis approach as the experimental data.

(iii) Where possible, kinetic models with a specific number and connectivity of states are preferred over mean residence times, since the latter leave the individual transition rate constants undetermined for more than 2 states.

(iv) Benchmarking new analysis tools using datasets of varied complexity – including models with more than 2 states – can reveal systematic errors, e.g. regarding the weighting of multiple rate constants that depopulate a given state, an issue encountered in this study.

(v) Benchmarking new software with established test data helps the potential users to judge the added benefits of newly introduced analysis tools. The diverse datasets used herein are publicly available and can serve to assess a tool's performance under varied experimental conditions.

(vi) Supporting broadly accepted file formats for newly developed analysis tools facilitates fast dissemination in the field. We offer the simple format described in Supplementary Note 4, which proved to be very useful for this study.

Looking ahead, a particularly promising outlook is the possibility to characterize individual states with individual noise patterns more accurately, using machine learning. Recently, deep learning approaches have been developed for the unbiased selection of single molecule traces for further kinetic analysis[49,50]. Similar approaches could be envisioned for a model-free kinetic analysis, which bears the potential to improve model selection significantly[51,18]. Demonstrating such new tools using public training datasets and supporting the simple file format introduced in this study, will accelerate the dissemination of the newest theoretical developments within the community of single-molecule experimentalists.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In conclusion, this first blind study on kinetic inference from smFRET data further validates the use of smFRET in deciphering biomolecular rates. It unequivocally reveals the current strengths and weaknesses of the various analysis approaches when tested against frequently encountered phenomena in smFRET experiments, and provides a reference standard for the continuous developments in this active field. We anticipate that this study will serve the community as a guide for data interpretation, spark future developments in kinetic inference, and therefore help to advance our understanding of biomolecular dynamics leading to function.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

METHODS

Procedure of this benchmark study

The need for a comparison of analysis tools for smFRET trajectories has grown with the increasing number of smFRET users and published tools. This was addressed at the Biophysical Society Meeting 2019 (Baltimore, US) by initiating a kinetic software challenge, short kinSoftChallenge. In line with more efforts to assess, promote, and potentially standardize experimental and analytical smFRET procedures (Refs[2,3] and Gebhardt *et al.*, in preparation), the kinSoftChallenge represents an important step aimed to improve the reliability and accuracy of kinetic inference from smFRET trajectories. In a first round of the study (July 2019 to November 2019), the participants received three simulated datasets (shown in **Figs. 2, 3,** and **4**). In the second round (December 2019 to February 2020), the participants analyzed the experimental dataset shown in **Fig. 5**. Experimental data with high and low SNR was compared in a third round (November to December 2020, shown in **Fig. 2**, and Supplementary Fig. 2). The individual test conditions are described in the text and summarized in **Table 1**. All challenge rounds were conducted as blind studies, i.e., the participants did not have ground truth information during data analysis (exceptions are labeled with a dagger in all Figures).

The simulations are detailed in the Supplementary Note 1. The experimental data shown in **Figure 2** and Supplementary Figures 1,2,3 was kindly provided by Benjamin Schuler. It shows the interaction between the nuclear-coactivator binding domain of CBP/p300 (NCBD) and the intrinsically disordered activation domain of the steroid receptor coactivator 3 (ACTR), measured using confocal single-photon detection[5]. The experimental data shown in **Figure 5** and Supplementary Figure 5 was kindly provided by Michael Schlierf. It shows binding of single-strand binding proteins (SSB) to a fluorescently labelled DNA hairpin, measured in prism-type total-internal reflection fluorescence (TIRF) mode using camera-based detection (EMCCD)[4]. Supplementary Figures and Notes on the simulations, validations, minimal uncertainty estimation, and the file format used herein are provided as **Supplementary Information,** along with detailed descriptions of all analysis tools and additional data tables. All inferred results as well as simulation scripts and parameters are provided as **Supplementary Datafiles**.

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Acknowledgements

We thank Benjamin Schuler and Michael Schlierf for providing experimental smFRET data. We thank the 2019 chair of the Biological fluorescence subgroup of the Biophysical Society (BPS), Paul Wiseman, and the co-chairs of the FRET in Biophysics Discussion Forum, Claus Seidel and Hugo Sanabria, for providing a platform to initiate the kinSoftChallenge. MG was funded by the Deutsche Forschungsgemeinschaft (German Research Foundation) – Project no. 431471305. RKOS thanks the SNF (200020_165868 and 200020_192153) and UZH for financial support. RB thanks the University of Applied Sciences Mittweida for financial support. LV, JS, TH were supported by the Collaborative Research Centre SFB1381 funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 403222702 – SFB 1381. SS was supported by the Postdoc.Mobility fellowship no. P400PB_180889 by the Swiss National Science Foundation. JC thanks US National Human Genome Research Institute (NHGRI) Project-ID 1R15HG009972. LK and DDM thank the Case Western Reserve University College of Arts and Sciences for support. DAE supported by NIH grant R35 GM127151. KRW supported by NIH grants R01 GM132263 and R01 GM118508. CDL and DDR were supported by grant 16SMPS05 from the Institutes Organization of the Dutch Research Council (NWO-I, formerly FOM). NSH acknowledges support from Carlsberg foundation Distinguished associate professor program (CF16-0797) Vellux foundation center of excellence BIONEC (grant no 18333) and the NovoNordisk foundation (NNF14CC00001 and NNF16OC0021948), GH and HS acknowledge support by NIH 1P20GM130451 and 2R01MH0 81923-11A1, and NSF 1749778. AB acknowledges support by the European Research Council through the Advanced Grant 2014 hybridFRET (number 671208).

Author contributions

All authors analysed data and/or discussed the results and contributed to writing the manuscript. MG and SS compiled all results, prepared figures, and wrote the article draft together with AB. MG performed and analysed simulations. SS designed and initiated the study, and organized the collaboration in consultation with all authors.

Additional information

Supplementary figures, notes, and methods are provided in the Supplementary Information file. All inferred results and simulation parameters are provided in the Supplementary Datafiles.

Competing interests

The authors declare no competing interests.

Code and data availability

Simulation code and all smFRET datasets are available at www.kinsoftchallenge.com and https://doi.org/10.5281/zenodo.5701310 .

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

1. Lerner, E. *et al.* Toward dynamic structural biology: Two decades of single-molecule Förster resonance energy transfer. *Science* **359**, (2018).
2. Hellenkamp, B. *et al.* Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. *Nat. Methods* **15**, 669–676 (2018).
3. Lerner, E. *et al.* FRET-based dynamic structural biology: Challenges, perspectives and an appeal for open-science practices. *eLife* **10**, e60416 (2021).
4. Grieb, M. S. *et al.* Dynamic stepwise opening of integron attC DNA hairpins by SSB prevents toxicity and ensures functionality. *Nucleic Acids Res.* **45**, 10555–10563 (2017).
5. Zosel, F., Soranno, A., Buholzer, K. J., Nettels, D. & Schuler, B. Depletion interactions modulate the binding between disordered proteins in crowded environments. *Proc. Natl. Acad. Sci.* **117**, 13480–13489 (2020).
6. Schmid, S. & Hugel, T. Controlling protein function by fine-tuning conformational flexibility. *eLife* **9**, e57180 (2020).
7. Rabiner, L. R. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE* **77**, 257–286 (1989).
8. McKinney, S. A., Joo, C. & Ha, T. Analysis of Single-Molecule FRET Trajectories Using Hidden Markov Modeling. *Biophys. J.* **91**, 1941–1951 (2006).
9. Blanco, M. & Walter, N. G. Chapter 9 - Analysis of Complex Single-Molecule FRET Time Trajectories. in *Methods in Enzymology* (ed. Walter, N. G.) vol. 472 153–178 (Academic Press, 2010).
10. König, S. L. B. *et al.* BOBA FRET: Bootstrap-Based Analysis of Single-Molecule FRET Data. *PLOS ONE* **8**, e84157 (2013).
11. Keller, B. G., Kobitski, A., Jäschke, A., Nienhaus, G. U. & Noé, F. Complex RNA Folding Kinetics Revealed by Single-Molecule FRET and Hidden Markov Models. *J. Am. Chem. Soc.* **136**, 4534–4543 (2014).
12. Pirchi, M. *et al.* Photon-by-Photon Hidden Markov Model Analysis for Microsecond Single-Molecule FRET Kinetics. *J. Phys. Chem. B* **120**, 13065–13075 (2016).
13. Schmid, S., Götz, M. & Hugel, T. Single-Molecule Analysis beyond Dwell Times: Demonstration and Assessment in and out of Equilibrium. *Biophys. J.* **111**, 1375–1384 (2016).
14. Zarrabi, N., Schluesche, P., Meisterernst, M., Börsch, M. & Lamb, D. C. Analyzing the Dynamics of Single TBP-DNA-NC2 Complexes Using Hidden Markov Models. *Biophys. J.* **0**, (2018).
15. Hon, J. & Gonzalez, R. L. Bayesian-Estimated Hierarchical HMMs Enable Robust Analysis of Single-Molecule Kinetic Heterogeneity. *Biophys. J.* (2019) doi:10.1016/j.bpj.2019.02.031.
16. Bronson, J. E., Fei, J., Hofman, J. M., Gonzalez, R. L. & Wiggins, C. H. Learning Rates and States from Biophysical Time Series: A Bayesian Approach to Model Selection and Single-Molecule FRET Data. *Biophys. J.* **97**, 3196–3205 (2009).
17. van de Meent, J.-W., Bronson, J. E., Wiggins, C. H. & Gonzalez, R. L. Empirical Bayes Methods Enable Advanced Population-Level Analyses of Single-Molecule FRET Experiments. *Biophys. J.* **106**, 1327–1337 (2014).
18. Sgouralis, I. *et al.* A Bayesian Nonparametric Approach to Single Molecule Förster Resonance Energy Transfer. *J. Phys. Chem. B* **123**, 675–688 (2019).
19. Kinz-Thompson, C. D., Ray, K. K. & Gonzalez, R. L. Bayesian Inference: The Comprehensive Approach to Analyzing Single-Molecule Experiments. *Annu. Rev. Biophys.* **50**, 191–208 (2021).
20. Chung, H. S. *et al.* Extracting Rate Coefficients from Single-Molecule Photon Trajectories and FRET Efficiency Histograms for a Fast-Folding Protein. *J. Phys. Chem. A* **115**, 3642–3656 (2011).
21. Liu, Y., Park, J., Dahmen, K. A., Chemla, Y. R. & Ha, T. A Comparative Study of Multivariate and Univariate Hidden Markov Modelings in Time-Binned Single-Molecule FRET Data Analysis. *J. Phys. Chem. B* **114**, 5386–5403 (2010).
22. Greenfeld, M., Pavlichin, D. S., Mabuchi, H. & Herschlag, D. Single Molecule Analysis Research Tool (SMART): An Integrated Approach for Analyzing Single Molecule Data. *PLOS ONE* **7**, e30024 (2012).
23. Kelly, D., Dillingham, M., Hudson, A. & Wiesner, K. A New Method for Inferring Hidden Markov Models from Noisy Time Sequences. *PLOS ONE* **7**, e29703 (2012).
24. Gurunathan, K. & Levitus, M. FRET Fluctuation Spectroscopy of Diffusing Biopolymers: Contributions of Conformational Dynamics and Translational Diffusion. *J. Phys. Chem. B* **114**, 980–986 (2010).
25. Felekyan, S., Sanabria, H., Kalinin, S., Kühnemuth, R. & Seidel, C. A. M. Chapter Two - Analyzing Förster Resonance Energy Transfer with Fluctuation Algorithms. in *Methods in*

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Enzymology (ed. Tetin, S. Y.) vol. 519 39–85 (Academic Press, 2013).
26. Schuler, B. Perspective: Chain dynamics of unfolded and intrinsically disordered proteins from nanosecond fluorescence correlation spectroscopy combined with single-molecule FRET. *J. Chem. Phys.* **149**, 010901 (2018).
27. Chung, H. S., Cellmer, T., Louis, J. M. & Eaton, W. A. Measuring ultrafast protein folding rates from photon-by-photon analysis of single molecule fluorescence trajectories. *Chem. Phys.* **422**, 229–237 (2013).
28. Kilic, S. *et al.* Single-molecule FRET reveals multiscale chromatin dynamics modulated by HP1α. *Nat. Commun.* **9**, 235 (2018).
29. Watkins, L. P. & Yang, H. Detection of Intensity Change Points in Time-Resolved Single-Molecule Measurements. *J. Phys. Chem. B* **109**, 617–628 (2005).
30. Shuang, B. *et al.* Fast Step Transition and State Identification (STaSI) for Discrete Single-Molecule Data Analysis. *J. Phys. Chem. Lett.* **5**, 3157–3161 (2014).
31. Chen, J., Pyle, J. R., Sy Piecco, K. W., Kolomeisky, A. B. & Landes, C. F. A Two-Step Method for smFRET Data Analysis. *J. Phys. Chem. B* **120**, 7128–7132 (2016).
32. Gauer, J. W. *et al.* Chapter Ten - Single-Molecule FRET to Measure Conformational Dynamics of DNA Mismatch Repair Proteins. in *Methods in Enzymology* (eds. Spies, M. & Chemla, Y. R.) vol. 581 285–315 (Academic Press, 2016).
33. de Lannoy, C. V., Filius, M., Kim, S. H., Joo, C. & de Ridder, D. FRETboard: Semisupervised classification of FRET traces. *Biophys. J.* **120**, 3253–3260 (2021).
34. Felekyan, S., Kalinin, S., Sanabria, H., Valeri, A. & Seidel, C. A. M. Filtered FCS: Species Auto- and Cross-Correlation Functions Highlight Binding and Dynamics in Biomolecules. *ChemPhysChem* **13**, 1036–1053 (2012).
35. Aggarwal, T., Materassi, D., Davison, R., Hays, T. & Salapaka, M. Detection of Steps in Single Molecule Data. *Cell. Mol. Bioeng.* **5**, 14–31 (2012).
36. Hadzic, M. C. A. S., Börner, R., König, S. L. B., Kowerko, D. & Sigel, R. K. O. Reliable State Identification and State Transition Detection in Fluorescence Intensity-Based Single-Molecule Förster Resonance Energy-Transfer Data. *J. Phys. Chem. B* **122**, 6134–6147 (2018).
37. Noji, H., Yasuda, R., Yoshida, M. & Kinosita, K. Direct observation of the rotation of F1-ATPase. *Nature* **386**, 299–302 (1997).
38. Croce, R. & van Amerongen, H. Light harvesting in oxygenic photosynthesis: Structural biology meets spectroscopy. *Science* **369**, eaay2058 (2020).
39. Kühlbrandt, W. Bacteriorhodopsin — the movie. *Nature* **406**, 569–570 (2000).
40. Edel, J. B., Eid, J. S. & Meller, A. Accurate Single Molecule FRET Efficiency Determination for Surface Immobilized DNA Using Maximum Likelihood Calculated Lifetimes. *J. Phys. Chem. B* **111**, 2986–2990 (2007).
41. Kilic, Z., Sgouralis, I. & Pressé, S. Generalizing HMMs to Continuous Time for Fast Kinetics: Hidden Markov Jump Processes. *Biophys. J.* **120**, 409–423 (2021).
42. Farooq, S. & Hohlbein, J. Camera-based single-molecule FRET detection with improved time resolution. *Phys. Chem. Chem. Phys.* **17**, 27862–27872 (2015).
43. Nicholson, D. A. & Nesbitt, D. J. Pushing Camera-Based Single-Molecule Kinetic Measurements to the Frame Acquisition Limit with Stroboscopic smFRET. *J. Phys. Chem. B* **125**, 6080–6089 (2021).
44. Frauenfelder, H., Sligar, S. G. & Wolynes, P. G. The energy landscapes and motions of proteins. *Science* **254**, 1598–1603 (1991).
45. Sunney Xie, X. Single-molecule approach to dispersed kinetics and dynamic disorder: Probing conformational fluctuation and enzymatic dynamics. *J. Chem. Phys.* **117**, 11024–11032 (2002).
46. Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450**, 964–972 (2007).
47. Feng, X. A., Poyton, M. F. & Ha, T. Multicolor single-molecule FRET for DNA and RNA processes. *Curr. Opin. Struct. Biol.* **70**, 26–33 (2021).
48. Schuler, B. *et al.* Binding without folding – the biomolecular function of disordered polyelectrolyte complexes. *Curr. Opin. Struct. Biol.* **60**, 66–76 (2020).
49. Thomsen, J. *et al.* DeepFRET, a software for rapid and automated single-molecule FRET data classification using deep learning. *eLife* **9**, e60404 (2020).
50. Li, J., Zhang, L., Johnson-Buck, A. & Walter, N. G. Automatic classification and segmentation of single-molecule fluorescence time traces with deep learning. *Nat. Commun.* **11**, 5833 (2020).
51. Sgouralis, I. & Pressé, S. An Introduction to Infinite HMMs for Single-Molecule Data Analysis. *Biophys. J.* **112**, 2021–2029 (2017).
52. Börner, R. et al. Simulations of camera-based single-molecule fluorescence experiments. *PLOS ONE* **13**, e0195277. (2018)

bioRxiv preprint doi: https://doi.org/10.1101/2021.11.23.469671; this version posted March 21, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

53 Hadzic, MCAS., Sigel, RKO., & Börner, R. Single-molecule kinetic studies of nucleic acids by Förster resonance energy transfer. *Methods in Molecular Biology* (in press, 2021).
54 Gopich, I. V. & Szabo, A. Decoding the Pattern of Photon Colors in Single-Molecule FRET. *J. Phys. Chem. B* **113**, 10965–10973 (2009).
55 Harris, P. D. *et al.* Multi-parameter photon-by-photon hidden Markov modeling. *Nat Commun* **13**, 1000 (2022).