Towards improved FRET-derived biophysical models of proteins

Inaugural dissertation

for the attainment of the title of doctor
in the Faculty of Mathematics and Natural Sciences
at the Heinrich Heine University Düsseldorf

presented by

Milana Popara
from Trebinje, Bosnia and Herzegovina

Düsseldorf, June 2023

from the institute for Molecular Physical Chemistry
at the Heinrich Heine University Düsseldorf

Published by permission of the
Faculty of Mathematics and Natural Sciences at
Heinrich Heine University Düsseldorf

Supervisor: Prof. Dr. Claus A. M. Seidel
Co-supervisor: Prof. Dr. Holger Gohlke

Date of the oral examination: 04.12.2023

Declaration

I, Milana Popara, declare under oath that I have produced my thesis independently and without any undue assistance by third parties under consideration of the 'Principles for the Safeguarding of Good Scientific Practice at Heinrich Heine University Düsseldorf'.
Experimental work of this thesis was conducted at Heinrich Heine University Düsseldorf.
This dissertation, in a presented form, has not been submitted to any other institution, and no successful or unsuccessful attempts to earn a doctoral degree have been done by me.

Düsseldorf, June 2023

Abstract

Förster Resonance Energy Transfer (FRET) measurements are an ever-growing tool in structural biology that have the ability of solving highly sophisticated biological questions, which are otherwise inaccessible or challenging to be addressed by other approaches.
Through technical and methodological developments, FRET measurements have evolved to a multiparameter detecting technique, which is able to infer the structural information with Ångström accuracy. Moreover, in FRET measurements one can detect and quantify conformational dynamics with unrivaled time resolution, spanning twelve orders of magnitudes in time, from sub-ns to thousands of seconds. Despite these advantages, FRET experiments suffer from the intrinsic sparsity of information, which impedes comprehensive structural insights at a satisfying level of detail or resolution. This can be overcome through an alliance with other experimental methods and computational methods that can add insights down to the atomistic level of detail.
Driven by these challenges, efforts of this thesis were focused on two main areas. To a large extent, this thesis is aimed at answering the question if precision and accuracy of FRET measurements are at a level that is ready to take on complexity of proteins. In a blind study, we assessed the performance of FRET measurements at a single-molecule level (smFRET). Reproducibility and accuracy of smFRET experiments across instruments, analysis procedures and systems of different complexity were rigorously tested. And much beyond - we answered two questions: (1) Can FRET detect and quantify conformational dynamics on different timescales? (2) What are the minimal structural fluctuations detectable? We were able to extract maximum information from fluorescence burst traces, while pushing the limits of FRET studies and dismissing any doubts about its credibility. At the same time, this work also revealed areas where FRET measurements are still imperfect. In particular, we found that the main cause of discrepancies between lab reported values is the error in correction parameters. To alleviate this issue, in a follow-up project I established a protocol for robust determination of correction parameters.
Second major area of research within this thesis was the use of FRET in integrative manner with all-atom MD simulations, with commitment to the principle of maximum entropy. Besides establishing the workflow for FRET/MD integrative modelling, we put on a test bench robustness of posterior reconstructions, and found that while recovering atomistic models is an ill-defined problem, ensemble-integrated representations, such as inter-residue distograms and 3D density maps, can be robustly recovered. While atomistic models are a long-term dream of structural biology, we pose the question if these are indeed useful, when the target are biomolecules which exist as large quasi-continuum of states. With the established FRET/MD modelling workflow, we recovered the first-ever structural models of a steric chaperone *Pa*Lif, from the family of lipase-specific foldases, which activate the most important class of enzymes in biotechnology. Besides Lif, using FRET-based tools we decoded plasticity of U2AF2 - yet another system that exceeds the complexity of systems typically studied. All systems studied within this thesis belong to a family of multidomain proteins, which constitute the majority of proteins in both eukaryotes and prokaryotes.
The results of this thesis provide a major step ahead for quantitative FRET and FRET-based structure determination.

This thesis is based on the following articles and manuscripts:

Published articles

- Ganesh Agam[#], Christian Gebhardt[#], **Milana Popara[#]**, ..., Claus A. M. Seidel, Don C. Lamb, Thorben Cordes
  "Reliability and accuracy of single-molecule FRET studies for characterization of structural dynamics and distances in proteins"
  DOI: 10.1038/s41592-023-01807-0
  Nature Methods 20, 523-535 (**2023**)
  (Supplement A)

- Jonas Dittrich[#], **Milana Popara[#]**, Jakub Kubiak, Mykola Dimura, Bastian Schepers, Neha Verma, Birte Schmitz, Peter Dollinger, Filip Kovacic, Karl Erich Jaegerc, Claus A. M. Seidel, Thomas-Otavio Peulen, Holger Gohlke
  "Resolution of Maximum Entropy Method-derived posterior conformational ensembles of a flexible system probed by FRET and molecular dynamics simulations"
  DOI: 10.1021/acs.jctc.2c01090
  Journal of Chemical Theory and Computation 19, 8, 2389–2409 (**2023**)
  (Supplement B)

Manuscripts

- **Milana Popara**, Anders Barth, Julian Folz, Suren Felekyan, Oleg Opanasyuk, Claus A. M. Seidel
  "Increasing the accuracy of single-molecule FRET experiments- robust and unbiased estimation of correction parameters"
  (Supplement C)

[#] these authors contributed equally

Contributions

Supplement A**:**

Reliability and accuracy of single-molecule FRET studies for characterization of structural dynamics and distances in proteins

Nature Methods 20, 523-535 (**2023**)

Ganesh Agam[#], Christian Gebhardt[#], **Milana Popara[#]**, ... Claus A. M. Seidel, Don C. Lamb, Thorben Cordes

T.C. initiated the study and D.C.L. coordinated the study. G.A., C.G., **M.P.**, A.B., C.A.M.S., D.C.L. and T.C. designed research. A.B., C.A.M.S., D.C.L. and T.C. supervised the project. R.M. cloned and purified MalE variants. H.-S.K. and M.S. provided U2AF2. C.G. and G.A. performed labeling of MalE and U2AF2 variants, respectively, for shipment to participating laboratories. **M.P.**, J.F. and C.A.M.S. designed MalE mutants 4 and 5 in silico. **M.P.** and J.F. performed initial measurements on MalE mutants 4 and 5. G.A. reperformed the analysis on the provided raw data for U2AF2 and MalE-1 from eight laboratories. G.A., A.B. and **M.P.** performed ds estimation. C.G. performed FCS experiments on MalE variants, time-resolved anisotropy experiments and R_0-determination. **M.P.** performed time-resolved anisotropy analysis of single labelled MalE cysteine mutants from ensemble measurements as well of all MalE and U2AF2 dye combinations from smFRET measurements. C.G., G.A. and **M.P.** performed measurements of MalE mutants with additional dye combinations and **M.P.** and A.B. performed statistical analysis of ds and anisotropies. G.A. and A.B. performed estimation of setup-dependent parameters and PDA of U2AF2. G.A. and **M.P.** performed the filtered-FCS and A.B. performed TCSPC analysis of U2AF2. C.G., G.A. and **M.P.** performed smFRET measurements on double-stranded DNA rulers. **M.P.** performed AV and ACV modelling of dye distributions for MalE and U2AF2. G.G.M.M. performed microscale thermophoresis experiments. M.d.B. performed confocal scanning experiments for surface-immobilized MalE. **M.P.** performed and analyzed MD simulations of MalE. All authors were involved in performing comparison experiments and analyzing smFRET data. G.A. and C.G. consolidated data collection of participating laboratories. G.A. and C.G. designed Fig. 1. G.A., C.G. and **M.P.** designed Fig. 2. A.B. designed Figs. 3 and 4. A.B. and **M.P.** designed Figs. 5 and 6. G.A., C.G., **M.P.**, A.B., C.A.M.S., D.C.L. and T.C. interpreted data and wrote the manuscript in consultation with all authors.

Supplement B**:**

Resolution of Maximum Entropy Method-derived posterior conformational ensembles of a flexible system probed by FRET and molecular dynamics simulations

Journal of Chemical Theory and Computation 19, 8, 2389–2409 (**2023**)

Jonas Dittrich[#], **Milana Popara**[#], Jakub Kubiak, Mykola Dimura, Bastian Schepers, Neha Verma, Birte Schmitz, Peter Dollinger, Filip Kovacic, Karl Erich Jaegerc, Claus A. M. Seidel, Thomas-Otavio Peulen, Holger Gohlke

J.D.: investigation, formal analysis, visualization, and writing-original draft; **M.P.**: investigation, formal analysis, visualization, and writing-review and editing; J.K.: investigation, fluorescence measurements, experimental data analysis, visualization, and writing-review and editing; M.D.: software-MEM implementation; Ba.S., N.V., and Bi.S.: investigation; P.D.: sample preparation and LipA activity measurements; F.K. and K.-E.J.: project administration and funding acquisition; C.A.M.S.: conceptualization, writing-review and editing, project administration, and funding acquisition; T.-O.P.: investigation, software-ucFRET, design of analysis, formal analysis, and writing-original draft; H.G.: conceptualization, writing-review and editing, project administration, and funding acquisition.

Supplement C**:**

Increasing the accuracy of single-molecule FRET experiments- robust and unbiased estimation of correction parameters

Milana Popara, Anders Barth, Julian Folz, Suren Felekyan, Oleg Opanasyuk, Claus A. M. Seidel

M.P. performed measurements, data analysis, made the figures and wrote the original draft of the manuscript. SF performed software development. J.F., **M.P.** and C.A.M.S ordered and managed samples. J.F. performed pilot measurements for the project. All authors reviewed the manuscript. A.B. and C.A.M.S conceptualized and supervised the project.

Table of Contents

List of Figures

Chapter 1

Introduction

Grand challenge of biology. According to the UniProt database, there are over 200 million known proteins, with their number growing rapidly every year. Yet, only for a tiny fraction of all known proteins structural models could be inferred. What might appear as a paradox to that statement, is a fact that proteins are comprised of only 20 building blocks, known as amino acids. However, a range of interactions occurring between these building blocks leads to 10^{300} possible conformations of a protein, as noted by Cyrus Levinthal in 1969.[1] While in nature proteins fold spontaneously, and in a very short time, this is a tremendously large search space for prediction of protein structure using computational methods. On the other side, experimental data are without exception too sparse to provide information on biomolecular structures at atomistic level of detail. Moreover, many experimental techniques suffer from temporal and/or spatial averaging, or derive structural information from non-physiological measurement conditions. This summarizes why protein structure determination represents the grand(est) challenge of biology. However, especially in recent years, this challenge has been tackled through an alliance of multiple experimental and computational methods, resulting in so-called integrative or hybrid biomolecular models.[2-4] Moreover, through integrative approach, structural biology is nowadays able to go beyond single structures, since rising number of multi-state and time-ordered structural ensembles is being reported.[5-8]

Revolution of fluorescence spectroscopy. Part of the integrative modelling revolution are Förster Resonance Energy Transfer (FRET) measurements. More than 70 years elapsed since the pioneering work of Förster[9] on the inter-molecular non-radiative energy transfer between donor and acceptor fluorophore, which is the underlying mechanism in FRET experiment. Donor and acceptor fluorophore can be both free in the solution, but it is also possible that one or both fluorophores are attached to a molecular construct of interest, or they can even be its natural part, in which case they are known as intrinsic fluorophores. An example of intrinsic fluorophores would be aromatic amino acids, such as tryptophan. Key feature of FRET is that efficiency of energy transfer between the fluorophores is an inverse function of the sixth power of distance between the fluorophores. However, it wasn't until the discovery of principles of click chemistry[10] that FRET started to be used to explore biologically relevant processes. Click chemistry, for which the Nobel Prize in Chemistry was awarded in 2022, is an approach which allows building blocks to snap together. It enabled a new level of applications within a range

of disciplines in chemistry and related sciences. Part of this revolution was fluorescence spectroscopy, where click chemistry is used to tag biomolecules with fluorescent probes. Residue-specific labelling of biomolecules with fluorescent probes allows to obtain information about inter-residue distances within biomolecule. This meant that to work out the structure of a biomolecule, one was no longer constrained to traditional tools of structural biology, i.e. X-ray crystallography, NMR and cryo-EM.

Sensitivity and experimental design of FRET measurements. Despite long-term dominance of X-ray crystallography, and its non-questionable contributions to the science, it is hard to overlook that it is limited to providing biomolecular models in a single static pose, along with non-physiological measurement conditions (see **Chapter 4.1** for more details). Other dominant tools of structural biology, such as NMR spectroscopy and cryo-EM, have other limitations of their own, such as the size of the biomolecule that can be studied or the resolution of the obtained models (see **Chapter 4.1** for more details). While FRET measurements are not the only tool reporting on molecules directly in a solution, they allow to monitor conformational dynamics in real-time, and most importantly, one single molecule at a time. Observing single molecules, instead of ensemble averages, is the ultimate level of sensitivity, that is in FRET experiment achieved by employing a relatively simple optical design, which is centered around confocal microscope. Through its focusing power, a confocal microscope minimizes the excitation volume, and the pinhole, that lies in a plane conjugate to the focal plane of the objective lens, rejects all out-of-focus light. These two features of a confocal microscope, in a combination with small enough sample concentrations, enable us to observe single molecules at a time (for details see **Chapter 2.2**). The true potential of FRET started to be harnessed when two important features were added to aforementioned experimental design, and that is Pulsed Interleaved Excitation (PIE)[11] and Multiparameter Fluorescence Detection (MFD)[12, 13]. These two features allow selection of sub-populations and their characterization with a set of simultaneously recorded fluorescence parameters. PIE is an excitation scheme where two pulsed excitation lasers are being operated with a repetition time on nanosecond timescale, and are alternated with a delay in respect to each other, of approximately half the repetition period. Such laser operating scheme allows to excite donor and acceptor fluorophore at different points in time, and since the photon arrival time is determined relative to the excitation pulse, it is possible to disentangle donor and acceptor emitted photons despite their spectral overlap. Besides temporal classification of emitted photons, MFD schemes takes a step further, by additionally employing spectral and polarization classification of the emitted photons. This allows to simultaneously determine multiple properties of a sample, such as fluorescence

lifetime and anisotropy, as well as the quantities such as FRET efficiency and stoichiometry. Simultaneous acquisition of all these parameters minimizes ambiguity of the data interpretation, which in turn gives high fidelity to FRET measurements. Principles of PIE and MFD are discussed in more detail in **Chapter 2.3**.

Determination of absolute FRET efficiency. Essential observable in FRET experiment is efficiency of energy transfer between donor and acceptor fluorophore, that informs on their spatial separation. As the instrumentation for FRET experiments was developing, accordingly the approach for determination of FRET efficiency was changing. Two most common ways to determine absolute value of FRET efficiency are: (**a**) by measuring the fraction of FRET sensitized acceptor fluorescence out of the total fluorescence of both donor and acceptor ("*classical ratio method"*) and (**b**) by measuring the decrease of donor fluorescence lifetime in the presence of acceptor (for details see **Chapter 2.1** and **Chapter 2.4**). The first approach is intensity based, and it requires correction of raw fluorescence intensities for several experimental artifacts, namely: (*i*) background emission (*ii*) donor crosstalk into acceptor emission channel (*iii*) direct acceptor excitation by donor excitation laser (*iv*) correction for distinct fluorescence quantum yields of donor and acceptor, as well for different detection efficiencies of green and red detection channels for donor and acceptor fluorophore (*v*) correction for different excitation rates of donor and acceptor with their corresponding excitation light. However, in approach (b), which is lifetime-based, corrections (*ii*)-(*v*) are not needed. Determination of aforementioned correction parameters is a topic of Manuscript 1 (see **Chapter 6.3** and Supplement C). Most commonly, one applies both intensity and lifetime approach for a consistency check. Furthermore, a series of alternative ways to determine FRET efficiency are described and reviewed by Clegg[14], such as: (**c**) by measuring enhanced acceptor fluorescence, as the ratio of FRET sensitized acceptor fluorescence and acceptor fluorescence due to direct acceptor excitation ("$(ratio)_A$ *method"*) (**d**) by measuring decrease in fluorescence quantum yield of donor, i.e. via the ratio of donor fluorescence quantum yield in the presence and absence of the acceptor ("$(ratio)_D$ *method"*) (**e**) via anisotropy change of donor and acceptor.

FRET-based tools. In terms on instrumentation, the field has not seen further major advancements of a described PIE-MFD scheme, other than multi-color FRET experiments[15]. However, in terms of data analysis, the number of statistical tools, which allow to extract the maximum information from the noisy experimental data, is rapidly growing, with some of the popular methods being: Photon Distribution Analysis (PDA)[16, 17], filtered Fluorescence Correlation Spectroscopy (filtered FCS)[18], FRET-FCS[19], Burst Variance Analysis (BVA)[20],

FRET-2 Channel Kernel Density Estimator (FRET-2CDE)[21], and graphical tools such as FRET-lines[22, 23]. Extensive and up-to-date overview of different analysis tools and software packages is given in a community paper by Lerner *et al*.[24]

Current status of FRET measurements and scope of the thesis. In terms of system complexity, FRET historically started with artificial toy models. In the first experimental proof of Förster theory, FRET efficiency was measured in oligomeric chains of poly-L-proline, between the α-naphthyl group (donor) at the carboxyl end of polypeptide and dansyl group (acceptor) at the imino end.[25, 26] With time, rapid development of instrumentation and analysis tools enabled FRET to evolve into a method that nowadays can address systems with highest degree of complexity, i.e. multidomain proteins with dynamic super-tertiary structure[7] and intrinsically disordered proteins[8]. However, the amount of data points obtained in a FRET experiment is typically too small relative to the number of degrees of freedom in a molecule. In other words, FRET data is too sparse to deliver accurate structural models of biomolecules on its own. To alleviate this limitation, FRET teams up with other experimental methods and computation techniques such as MD simulations, which provide physical models down to atomic resolution.[5, 7] Alone or in integrative manner with other experimental techniques and computational methods, FRET is nowadays a competitive tool in the field of structural biology. It is already demonstrated that FRET can inform on small- and large-scale motions, it can measure distances in biomolecules, detect dynamics across different time regimes and unravel underlying kinetical scheme in a dynamic system.[27, 28]

The FRET community has shown that FRET measurements are reproducible, and can infer structural information with sub-nanometer precision and accuracy when using rigid double-stranded DNA molecules.[29] However, an objective evaluation of the overall performance of FRET experiments when using challenging proteins system has been lacking. Therefore, this thesis was to large extent focused on critical assessment of FRET experiments for determining distances, as well for detecting and quantifying dynamics in protein systems that display conformational dynamics across different spatial and temporal scales (see **Chapter 6.1**). As model systems we used maltose binding protein, referred to as MalE, and U2AF2, the truncated large subunit of U2 auxiliary factor of the spliceosome machinery (see **Chapter 5.2** and **Chapter 5.3**). In the second part of this thesis, we embarked on optimizing the conformational ensemble of *Pa*Lif in its apo state (see **Chapter 6.2**). *Pa*Lif belongs to a family of steric chaperones, which activate their cognate lipase (see **Chapter 5.4**). Since lipases are the most important class of enzymes in biotechnology, it is of immense importance to understand the behavior of their steric chaperones. To tackle this challenge, we joined the forces of MD

simulations and FRET measurements. For that, we established the workflow for integrative FRET/MD modelling using maximum entropy method. In a benchmark using synthetic data, we were able to find how to optimally balance these two sources of information. Moreover, we put on a test bench the robustness of optimized ensembles across representations of different level of detail- from atomistic models, over inter-residue distance histograms (distograms) to 3D density maps. Lastly, by using experimental FRET data acquired for a network of six FRET pairs, we were able to infer the first ever structural models of *Pa*Lif in its apo state.

Outline

This thesis is designed such to provide information and tools necessary for accurate smFRET measurements, and robust modelling of structural ensembles, in integrative approach using FRET measurements and structural models obtained from all-atom MD simulations. In **Chapter 2**, I describe basics of FRET and its fundamental properties. This is then followed with discussion on how the ultimate level of sensitivity, i.e. single-molecule detection, is achieved in a FRET experiment by using a confocal microscope. I dissect the reasonings behind such experimental realization, and introduce further additions to a confocal setup, i.e. principles of Pulsed Interleaved Excitation (PIE) and Multiparameter Fluorescence Detection (MFD). I finally close this chapter by showing how the structural information is extracted from measured fluorescence decays. In **Chapter 3** I outline fundamentals of MD simulations along with practical insights into methods for trajectory clustering, analysis tool frequently used to both reduce the amount of data points, but also to identify hidden trends in data. After discussing the fundamentals of FRET and MD simulations, in **Chapter 4** I provide detailed overview of current trends in integrative structural biology, as well as mathematical formalism underlying the ensemble reweighting methods. These include Bayesian inference and principle of maximum entropy. Moving forward, in **Chapter 5** I highlight the exciting biological context, structure and function of each of molecules studied within this thesis, prior to reaching **Chapter** 6, which provides overview over articles published during this thesis. Lastly, **Chapter 7** summarizes the main outcomes of this thesis, while highlighting different methodologies from the field of computer science and fluorescence, that were applied in individual or synergetic approach within this thesis.

Chapter 2

FRET: Concepts and experimental realization

2.1. Basics of FRET

Förster resonance energy transfer (FRET) is a singlet-singlet radiationless energy transfer that occurs from an excited donor fluorophore (D^*) to an acceptor fluorophore (A), as described in the pioneering work of Förster.[9] Energy transfer between excited donor and ground-state acceptor fluorophore, can be described by the following reaction scheme[30]:

$$\begin{array}{ccccc} & & \begin{array}{c} D + A \\ \uparrow k_{D,nr} \end{array} & & \begin{array}{c} D + A \\ \uparrow k_{A,nr} \end{array} \\ D + A + h\nu_{exc} & \rightarrow & \begin{array}{c} D^* + A \\ \downarrow k_{D,f} \end{array} & \xrightarrow{k_{FRET}} & \begin{array}{c} D + A^* \\ \downarrow k_{A,f} \end{array} \\ & & D + A + h\nu_D & & A + D + h\nu_A \end{array} \tag{2.1}$$

D and A represent donor and acceptor fluorophore in the ground state, while their excited state is denoted as D^* and A^*. Furthermore, $h\nu$ represents the quantum of radiation, which in the reaction scheme above appears with frequency of donor excitation, ν_{exc}, donor fluorescence, ν_D and acceptor fluorescence, ν_A.

Competitive to energy transfer (k_{FRET}) are fluorescence and other non-radiative relaxation processes of the donor ($k_{D,f}$, $k_{D,nr}$). The FRET sensitized acceptor fluorophore (A^*) can analogously depopulate its excited state via both fluorescence and non-radiative pathways ($k_{A,f}$, $k_{A,nr}$). The efficiency of energy transfer from D to A, also called FRET efficiency *E*, is given by:

$$E = \frac{k_{FRET}}{k_{FRET} + k_{D,f} + k_{D,nr}} \tag{2.2}$$

where k_{FRET}, the rate of energy transfer between donor and acceptor fluorophore, strongly depends on the spatial separation between the fluorophores, R_{DA}:

$$k_{FRET} = \frac{1}{\tau_{D(0)}} \left(\frac{R_0}{R_{DA}}\right)^6 \tag{2.3}$$

$\tau_{D(0)}$ is the average excited state lifetime of a donor in the absence of acceptor, i.e. $\tau_{D(0)} = 1/(k_{D,f} + k_{D,nr})$, and R_0 is, as described in pioneering work of Förster[9], the distance at which the rate of the energy transfer is equal to the sum of rates of all other pathways of donor excited state deactivation, i.e. $k_{FRET} = 1/\tau_{D(0)}$. Then, it becomes obvious that R_0, also known as Förster radius, is an intrinsic property of a given donor-acceptor pair, and that it represents a distance at which the energy transfer equals 0.5.

By inserting eq. 2.3 in eq. 2.2, we can redefine FRET efficiency in terms of donor-acceptor distance R_{DA} and Förster radius, R_0 (**Figure 1A**):

$$E = \frac{{R_0}^6}{{R_0}^6 + {R_{DA}}^6} \tag{2.4}$$

Furthermore, considering the definitions of donor excited state lifetime in absence and presence of acceptor, i.e. $\tau_{D(0)} = 1/(k_{D,f} + k_{D,nr})$ and $\tau_{D(A)} = 1/(k_{D,f} + k_{D,nr} + k_{FRET})$, we arrive to the formulation that underlies one of the most frequently used approaches for experimental determination of FRET efficiency:

$$E = 1 - \frac{\tau_{D(A)}}{\tau_{D(0)}} \tag{2.5}$$

The Förster radius, as a characteristic property of a given donor-acceptor pair, depends on a number of properties of the fluorophores, as well as properties of the surrounding medium:

$${R_0}^6 = \frac{9000(\ln 10)}{128\pi^5 N_A} \frac{\kappa^2 \Phi_{F,D} J(\lambda)}{n^4} \tag{2.6}$$

κ^2 is the orientation factor that describes relative orientation of transition dipole moment vectors of donor's emission and acceptor's excitation. The vector of inter-dye distance, $\vec{R}_{DA}$, and the vectors of transition dipole moments, $\vec{\mu}_D$ and $\vec{\mu}_A$, define two planes, in which transition dipole moment vectors lie. The angle between these two planes, φ, and the angles between transition dipoles and inter-dye distance vector, θ_A and θ_D, define the orientational factor according to[30]:

$$\kappa^2 = (\sin\theta_D \sin\theta_A \cos\varphi - 2\cos\theta_D \cos\theta_A)^2 \tag{2.7}$$

For graphical illustration of geometrical parameters defining orientational factor see **Figure 1C**. In a geometrical arrangement when transition dipole moment vectors are colinear, i.e. when they are in line with the inter-dye distance vector, i.e. φ, θ_D, θ_A= [0°, 0°, 0°], κ^2 has the maximum value of 4. For parallel dipoles, i.e. φ, θ_D, θ_A= [0°, 90°, 90°], $\kappa^2 = 1$.[30]

The minimum value of $\kappa^2 = 0$ can be realized for an infinite number of acceptor transition dipole orientations, all of which are perpendicular to the transition dipole moment of the donor, i.e. φ, θ_D, θ_A= [90°, 90°, 0°…360°].[30]

When the rotational motion of donor and acceptor is faster than the energy transfer time, then they sample all possible orientations relative to each other, which is known as isotropic dynamic averaging regime. In such conditions, κ^2 can be replaced with its average value $\langle\kappa^2\rangle_{iso} = \int_{\kappa^2=0}^{\kappa^2=4} \kappa^2 p(\kappa^2) d\kappa^2 = 2/3$, which is an assumption often made when calculating FRET efficiency.[30]

Additionally, the Förster radius is a function of: $\Phi_{F,D}$, fluorescence quantum yield of the donor in absence of acceptor; n, refractive index of the surrounding medium; N_A, Avogadro's number; $J(\lambda)$, the overlap integral between the normalized fluorescence emission spectrum of the donor $f_D(\lambda)$, and extinction spectrum of the acceptor $\varepsilon_A(\lambda)$ (**Figure 1B**):

$$J(\lambda) = \int_0^{\infty} f_D(\lambda)\, \varepsilon_A(\lambda) d\lambda \qquad (2.8)$$

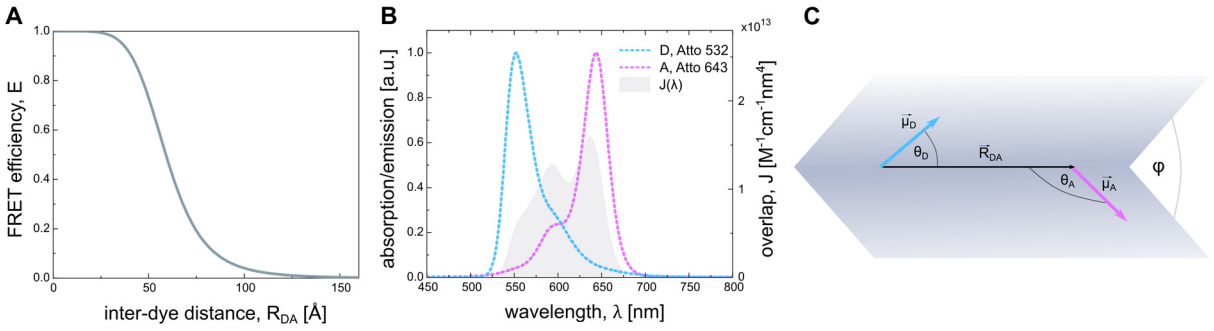


Figure 1 | Basics of FRET. ***(A)*** *Inter-dye distance dependence of FRET efficiency (eq. 2.4) for the dye pair Atto 532- Atto 643, with a Förster radius of 59 Å.* ***(B)*** *Overlap integral (grey shaded area), together with fluorescence emission spectrum of donor (blue) and absorption spectrum of acceptor dye (magenta). Overlap integral is calculated according to eq. 2.8 by assuming extinction coefficient of acceptor to be 150000 $M^{-1}cm^{-1}$ at the maximum of absorption.* ***(C)*** *Transition dipole moment vector of donor $\vec{\mu}_D$ (blue) and acceptor $\vec{\mu}_A$ (magenta), and angles θ_A, θ_D, which denote their orientation relative to the connecting distance vector $\vec{R}_{DA}$, whereas angle φ denotes the angle between the planes of dipole moment vectors. These geometrical parameters define the orientational factor κ^2 according to eq. 2.7.*

The classical definition of FRET rate constant (eq. 2.3) has a disadvantage of having both k_D and R_0 being a function of fluorescence quantum yield of donor, $\Phi_{\mathrm{F,D}}$, while in fact k_{FRET} is independent of donor quenching. Furthermore, the effect of relative dipole orientation is implicitly accounted through R_0. To circumvent these complications, reduced "spectral" Förster radius $R_{0,J}$ is defined[31]:

$$R_{0J}^6 = \frac{9000(\ln 10)}{128\pi^5 N_{\mathrm{A}}} \frac{J(\lambda)}{n^4} \tag{2.9}$$

which is less sample dependent and it allows to reformulate k_{FRET} as function of radiative rate constant of fluorescence, $k_{\mathrm{D,f}}$:

$$k_{\mathrm{FRET}} = k_{\mathrm{D,f}}\, \kappa^2 \left(\frac{R_{0\mathrm{J}}}{R_{\mathrm{DA}}}\right)^6 \tag{2.10}$$

Most importantly, such defined k_{FRET} is independent of sample-specific donor quenching.

2.2. Single-molecule resolution and confocal microscopy

The advantage of FRET over many other methods in the field of structural biology is that the measurements can be done at the single-molecule level, in real time and in solution. By monitoring single-molecules, one at a time, smFRET excels plethora of other methods, by not reporting on simple ensemble averages. Instead, smFRET provides distributions of inter-dye distances and underlying kinetic schemes, with minimal averaging that occurs on the timescale of fluorescence (few ns). Monitoring single-molecules, not only is useful for assessing heterogeneity of structures and underlying dynamics, but it also allows direct comparison with computational approaches, e.g. molecular dynamics simulations, which cannot address bulk properties, but are rather limited to characterization of individual molecules. Detection of single-molecules is the ultimate level of sensitivity and ultimate goal of structural biology, that is in an elegant way achieved in a single-molecule FRET experiment, by using a confocal microscope in combination with a Time-Correlated Single Photon Counting (TCSPC)[32] detection scheme. Why such experimental optical arrangement is a good way of detecting single-molecules, becomes obvious once we address the key experimental requirements, and that is (a) high Signal-to-Noise Ratio (SNR) and (b) high detection efficiency.[33] Bursts of fluorescence photons are interfered with unwanted background signal, that can have origins in (*i*) Rayleigh scattering by the solvent molecules at the excitation wavelength (*ii*) Raman scattering by the solvent molecules at the lower energy, i.e. higher wavelength compared to the

excitation light, and lastly (*iii*) Rayleigh and Raman scattering and fluorescence by the impurities in the solution.[33] Therefore, to achieve high SNR, a good optical arrangement should have efficient delivery of the excitation light, combined with a rather small excitation/observation volume, since that way the number of solvent molecules is minimized and thus the scattered light as well.[33] However, a small excitation volume is not sufficient to achieve ideal experimental conditions mentioned above, since unwanted signal, no matter how minimized, still goes through.

A simple confocal microscope showed to be capable to detect single-molecules while satisfying all aforementioned criteria.[33] Generally, an episcopic-fluorescence (epi-fluorescence) confocal microscope with a high Numerical Aperture (NA) objective is used, in an inverted configuration, i.e. objective pointing upwards. Epi-fluorescence is a configuration where the same optical component (dichromatic mirror) delivers the excitation light to the sample, and collects the emitted fluorescence (**Figure 2A**). Excitation light, reflected (or transmitted) on the dichromatic mirror, travels through the lens system of a microscope objective, which focuses the excitation light to a "point" that lies on a focal plane at the distance f from the objective lens. The size of the excitation/observation volume is determined by the ratio of the excitation light diameter, D, and the diameter of the back aperture of the objective, r_{BA} (**Figure 2C**).[34] When using a laser beam with smaller diameter compared to the back aperture of the objective (i.e. $D < r_{\mathrm{BA}}$), which is known as "underfilling" of the objective, excitation volume increases. Contrary to that, for overfilled back aperture of microscope objective (i.e. $D > r_{\mathrm{BA}}$), we reach diffraction limited spot ("Airy disk"), whose radius perpendicular to the light propagation is:

$$d = 1.22\frac{f\lambda}{D} \tag{2.11}$$

Airy disk represents the distance between the central maximum of the diffraction pattern to the first minimum, and it is a function of wavelength of the excitation light λ, focal length of the objective lens f, and the diameter of the excitation beam D. Focal spot size can also be expressed in other terms, such as the radius at the half maximum of the diffraction pattern or as a radius at the $1/e^2$ of the maximum intensity.

By exploiting the focusing power of the microscope objective, we minimize the excitation volume, and the amount of unwanted scattered light. However, the excitation light travels on both sides of the focal plane, i.e. it converges towards the focal points, and afterwards diverges, causing the out-of-focus fluorescence of the specimen. Here comes the main characteristic of a

confocal microscope that makes it a good optical arrangement for detection of single-molecules, which is the existence of a pinhole that lies in a plane *con*jugate to the *focal* plane of the objective lens, i.e. at the equal distance f from the objective lens as the focal plane, hence the name "confocal". The pinhole will reject all out-of-focus light and prevent its propagation towards the detection path (**Figure 2B**). That way we observe a single spot of a specimen at a time, which for sufficiently small observation volumes (few fL) and low enough sample concentrations (< 100 pM), means a single molecule at a time.

To summarize, properties of a confocal microscope that make it a suitable optical arrangement for detection of single molecules are:

- It focuses the excitation light, such that the excitation volume is minimized, and so is the unwanted background signal
- Through existence of the pinhole, that lies in a plane conjugate to the focal plane of the objective lens, it rejects any signal coming out-of-focus, allowing to study single spot of a specimen at a time, which means single molecule at a time, provided small enough concentration and small enough excitation volume.

Depending on the position of the focal plane, there are two smFRET measurement modalities. If the focal plane is inside of the solution, then freely diffusing molecules are being monitored, but if the focal plane is nearby the surface of the glass, then so-called Total Internal Reflection Fluorescence Microscopy (TIRFM)[35] is being employed, in which case the molecules are typically immobilized, and imaged on camera. Consequently, the temporal bandwidth of these two experimental designs is different. In diffusion experiment the observation time is limited by the diffusion of the molecule (typically < 10 ms), while in the TIRFM setup, the observation time extends to minutes, and is only limited by the dye photobleaching. On the other hand, TIRFM modality has a disadvantage of relatively low time resolution (miliseconds)[36], as opposed to picosecond time resolution in diffusion experiments. Higher temporal bandwidth of diffusion experiments can be achieved by increasing the observation volume, or by slowing down the molecular diffusion, for instance by encapsulating the molecule inside of liposomes. In this thesis, smFRET experiments are performed on freely diffusing molecules.

The way the specimen is seen is not solely determined by the shape of the excitation volume. The Point Spread Function (PSF), i.e. the mathematical description of the way the light travels through the entire optical arrangement is the convolution of both excitation and observation volume, where the latter is affected by each of the optical components in the detection path, e.g. objective lenses and confocal pinhole. Despite the numerous factors that contribute to its shape,

the PSF is a surprisingly simple mathematical expression. It was found that for underfilled back aperture of the objective, the shape of PSF can be approximated by a Gaussian[37, 38]:

$$\mathrm{PSF}(x,y,z) \propto \exp\left[-2\left(\frac{x^2+y^2}{w_0^2}+\frac{z^2}{z_0^2}\right)\right] \qquad (2.12)$$

with w_0 and z_0 being the waist radii at the $1/e^2$ of the maximum intensity (13.5%), perpendicular and in the direction of the light propagation (**Figure 2C**).

When using laser beam of Gaussian intensity profile, one would intuitively expect that Gaussian PSF would be achieved by overfilling the back aperture of the objective. However, Hess[37] has shown the opposite, i.e. that by underfilling of the objective, combined with the small confocal pinhole, the PSF becomes more Gaussian. While the shape of PSF is in general of no importance for intensity-based approaches where we look at the ratio of the instantaneous signals of donor and acceptor, it is essential for formal description of the molecular diffusion in a range of correlation methods, such as Fluorescence Correlation Spectroscopy (FCS).[39]

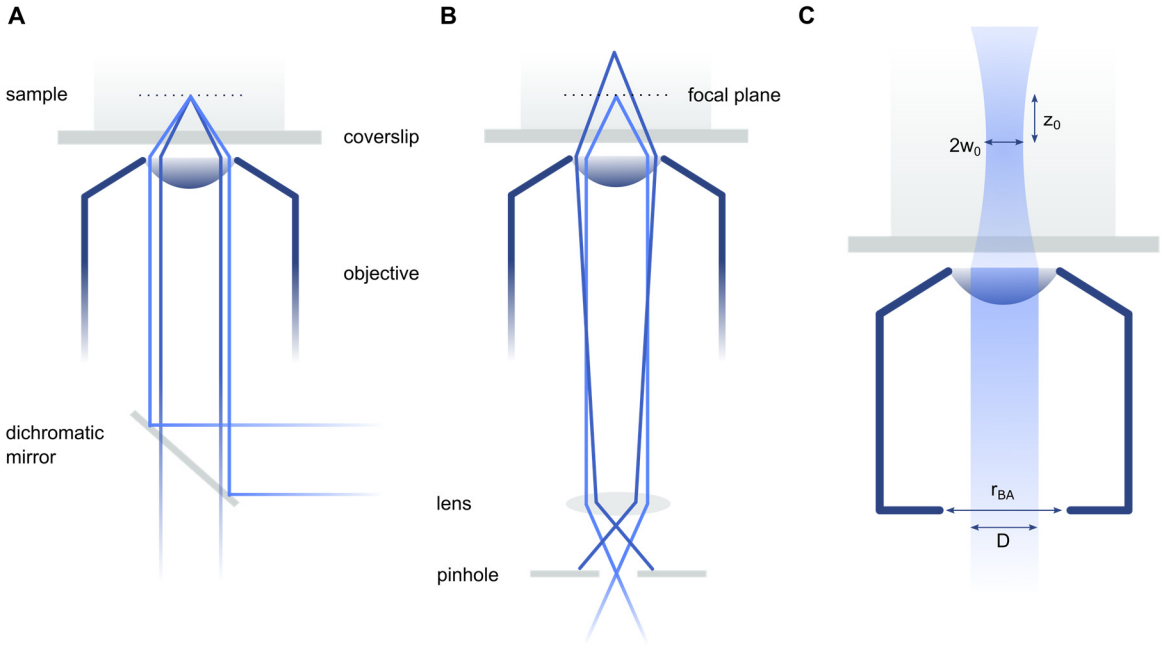


Figure 2 | Principles of confocal microscopy. ***(A)*** *Episcopic-fluorescence confocal microscope in an inverted configuration, i.e. objective pointing upwards. In epi(scopic)-fluorescence, the same optical component, i.e. dichromatic mirror, both delivers the excitation light (bright blue) to the specimen, but also collects emitted light (dark blue). Here is illustrated the case of long-pass dichromatic mirror, where excitation light is reflected, and emitted light is transmitted through the mirror.* ***(B)*** *The presence of a pinhole in a plane conjugate to the focal plane of the objective lens, allows to reject all out-of-focus light. This allows to study single spot of a sample at a time, which for sufficiently small observation volumes and small enough concentrations*

means single molecule at a time. ***(C)*** *Zoom-in to the observation volume, with waist radii,* w_0 *and* z_0*, at the* $1/e^2$ *of the maximum intensity, perpendicular and parallel to the direction of the light propagation. Size of observation volume is determined by the ratio of the laser beam diameter D, and the diameter of the back-aperture of the objective,* r_{BA}*. For overfilled objective, i.e.* $D > r_{BA}$*, we reach diffraction-limited size of volume (eq. 2.11), while for underfilled objective, i.e.* $D < r_{BA}$*, size of volume increases and its shape becomes more Gaussian-like (eq. 2.12). All illustrations are not to scale.*

As already discussed, basic confocal scheme will minimize the excitation volume and amount of unwanted scattered light, and the confocal pinhole will eliminate out-of-focus light. However, unwanted signal, even though it is minimized, still goes through and interferes with the fluorescence of a specimen. Therefore, for high SNR further additions to a basic confocal optical scheme are necessary. Rayleigh scattering at the excitation wavelength can be separated from the fluorescence via the use of dichromatic mirror, that will reflect (long-pass) or transmit (short-pass) it back to the light source. Raman scattering by the solvent molecules happens at higher wavelengths, and spectrally overlaps with fluorescence, but can be blocked from the detection path through the use of bandpass filters. This often has a cost of sub-optimal fluorescence detection, since one typically has to use narrower bandpass filters, or filters that are not around the maximum of fluorescence emission. Lastly, once the scattered light is eliminated, efficient detection of fluorescence is achieved through the use of high-performance detectors. Nowadays widely used are Avalanche Photodiodes (APDs), which are "point detectors", with typical active area of ~150 μm in diameter. APDs are designed such that they have uniform detection across the active area, high detection efficiency over a wide spectral range, low dark counts, fast response needed for time-resolved studies and excellent detection linearity.

2.3. Pulsed interleaved excitation and MFD detection scheme

When the fluorescently labelled biomolecule diffuses through the confocal volume, it emits bursts of fluorescence photons. Information contained within the burst is immense, and in order to harvest this information content, we employ spectral, polarization and temporal classification of the emitted photons. This in turn allows us to calculate a whole range of fluorescence parameters, such as fluorescence lifetime, anisotropy, FRET efficiency and stoichiometry. This is known as Multiparameter Fluorescence Detection (MFD).[13] Exemplary set of fluorescence parameters typically studied from smFRET data is given in **Figure 3**. MFD is nowadays typically employed together with Pulsed Interleaved Excitation (PIE) scheme, where two

pulsed excitation lasers with picosecond pulse width (FWHM) are interleaved on a nanosecond timescale.[11, 12]

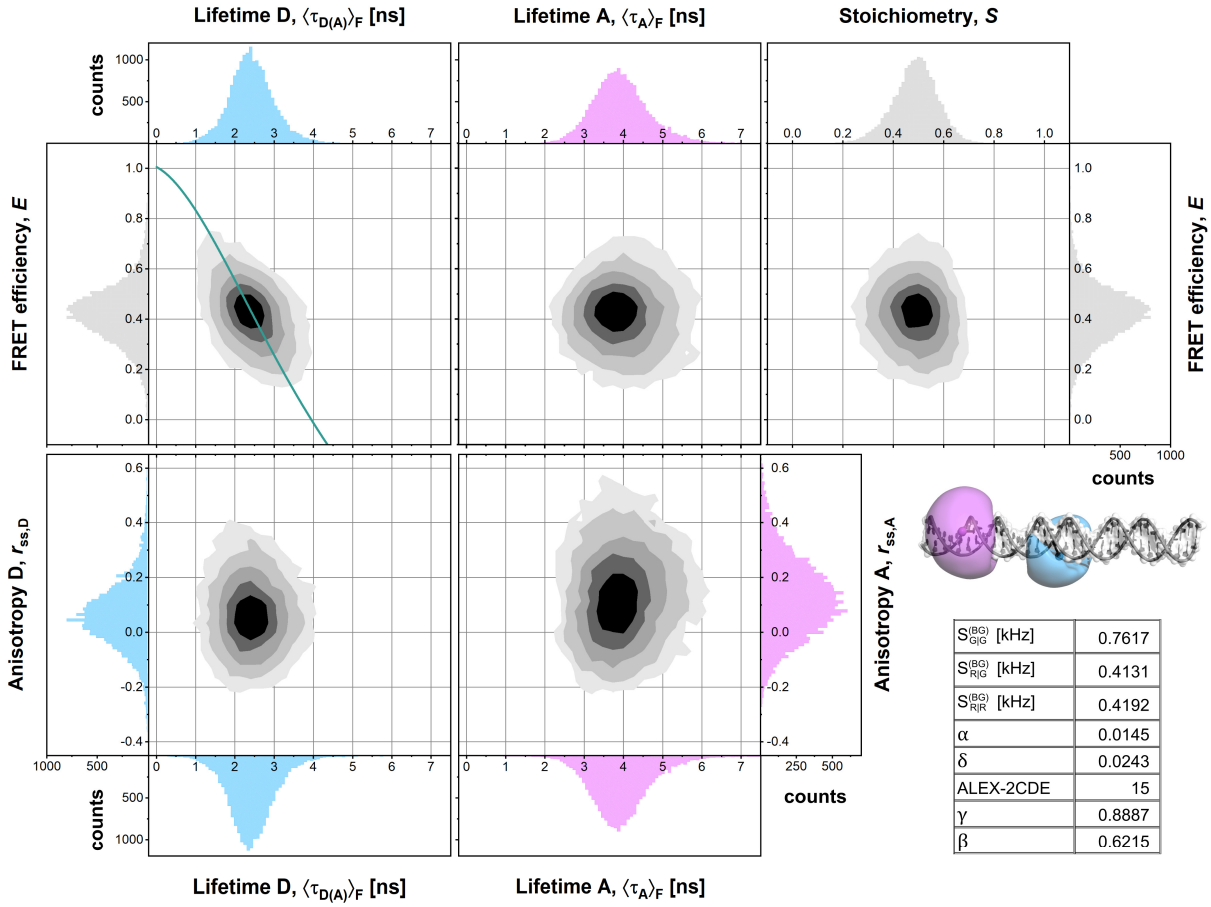


Figure 3 | PIE-MFD burst-wise fluorescence parameters *measured on single-molecule level for freely diffusing double-stranded DNA molecules labelled with Alexa Fluor 488 as donor dye and Atto647N as an acceptor dye. Visualized is only double-labelled population, obtained by applying the ALEX-2CDE filter, which removes single-labelled species and species with unstable emission of the fluorophores. Top row: (left) two-dimensional histogram of fully corrected FRET efficiency, E, versus the fluorescence intensity averaged lifetime of donor,* $\langle\tau_{D(A)}\rangle_F$*,with overlaid E-τ static FRET-line (green), which is a widely used graphical tool to detect FRET fluctuations. Displacement of the double labelled FRET population from the static FRET-line is indicative of presence of conformational dynamics, but can also be caused by dye artifacts, such as position-specific dye-protein interactions. Displayed static FRET-line was generated with following parameters:* R_0 *= 49 Å,* $\tau_{D(0)}$ *= 3.94 ns and assuming normally distributed dye positions with width of* σ_{DA} *= 6 Å. Furthermore, visualized are 2D histogram of fully corrected FRET efficiency E versus acceptor fluorescence lifetime,* $\langle\tau_A\rangle_F$*, (middle) and versus stoichiometry, S, (right). Bottom row: using polarization resolved intensities, calculated are steady-state anisotropy of donor* $r_{ss,D}$ *versus donor fluorescence lifetime (left) and steady-state anisotropy of acceptor,* $r_{ss,A}$*, versus acceptor fluorescence lifetime (middle), which can inform on size and shape of biomolecule, as well on the degree of fluorophore mobility. Steady state anisotropies are calculated accounting for polarization-dependent sensitivity of detectors (*$G_{VH/VV}$ *= 1.12 for donor,* $G_{VH/VV}$ *= 1.06 for acceptor) and for polarization mixing in the microscope objective (*l_1*= 0.0175,* l_2*= 0.0526).*

Donor and acceptor pulsed excitation lasers are operated at the same repetition frequency f_{rep}, but to ensure temporal separation of donor and acceptor excitation, lasers are delayed with respect to each other, for approximately half the repetition period, i.e. $\approx \frac{1}{2f_{\mathrm{rep}}}$. By exciting donor and acceptor at different points in time, it is possible to separate their photons, despite their spectral overlap. Such excitation scheme, combined with Time-Correlated Single Photon Counting (TCSPC)[32] detection scheme, allows for accurate FRET measurements.

Repetition period and the delay between donor and acceptor excitation laser are determined based on the fluorescence lifetime of the fluorophores. They should be sufficiently long for fluorophores to depopulate the excited state, and to avoid leakage of their emission into each other's PIE channel. To find how much at least two laser excitations have to be apart, we follow the simple equation for the fluorescence decay, i.e. $F(t) = F_0 exp(-t/\tau)$. For instance, for a fluorophore that has fluorescence lifetime of $\tau = 4.1$ ns, in order to collect 99.9% of the photons, we need two laser excitations to be at least $t = 4.1\ \mathrm{ns} \cdot \ln(1/0.001) \approx 29$ ns apart, which gives us laser repetition rate of at least $f_{\mathrm{rep}} = \frac{1}{2t} = 17$ MHz.

When talking about temporal classification of the photons, we use two terms, namely *macro-time,* t_{macro}, and *micro-time,* t_{micro}. Macro-time is simply the total number of synchronization periods from the beginning of the measurement

$$t_{\mathrm{macro}} = N_{\mathrm{sync}} \frac{1}{f_{\mathrm{rep}}} \tag{2.13}$$

while the micro-time represents the delay between the last synchronization signal and the subsequent photon detection time. Micro-time is perhaps the main experimental parameter because it encodes the fluorescence lifetime. Micro-time is determined using a Time-to-Amplitude Converter (TAC), which charges the capacitor whose final voltage is proportional to the time between the start pulse (synchronization signal) and stop pulse (photon detection).[32] However, since there are many periods when excitation is not followed by the photon detection, TAC would be triggered, but not stopped. This would require TAC reset before each next sync signal. For high-frequency measurements, performing TAC reset in each sync period is not a sustainable approach, and instead, a reverted start-stop scheme is employed, meaning that TAC is started with photon detection, and stopped by the next sync signal.[32] The output voltage of TAC is digitized using an Analog-to-Digital Converter (ADC). ADC builds a histogram, i.e. distribution of photons over the time. When ADC gets a signal from TAC, it adds an event in a respective n-th time bin, n_{ADC}. The resolution of ADC time-bins is obtained as the ratio of

maximum time covered by TAC, $T_{\mathrm{TAC}} = 1/f_{\mathrm{rep}}$, and total number of ADC bins, N_{ADC}. Then, micro-time can be determined as

$$t_{\mathrm{micro}} = n_{\mathrm{ADC}} \frac{T_{\mathrm{TAC}}}{N_{\mathrm{ADC}}} \tag{2.14}$$

which finally leads us to accurate photon-arrival time, which is obtained as a sum of t_{micro} and t_{macro}:

$$T_{\mathrm{arrival}} = N_{\mathrm{sync}} \frac{1}{f_{\mathrm{rep}}} + n_{\mathrm{ADC}} \frac{T_{\mathrm{TAC}}}{N_{\mathrm{ADC}}} \tag{2.15}$$

Obtaining photon arrival times is followed with "photon-sorting" into corresponding micro-time channels. In dual-color PIE experiment, there are four micro-time windows/channels, defined by the excitation laser X (Y|**X**) and detector Y (**Y**|X) where the photons are being registered. Donor/green micro-time windows are G|G and G|R, and similarly, acceptor/red micro-time windows are R|R and R|G. Besides temporal and spectral classification, one can further sort photons based on their polarization, which leads us to total of eight detection channels.

2.4. Burst search and burst analysis

After the photons are sorted, one can build photon traces, and perform burst search. To detect bursts of fluorescence photons, we can use the fact that background and fluorescence photons can be distinguished based on their inter-photon time. Background photons are of a lower count-rate, and subsequently they have longer inter-photon times. On the other hand, single-molecules diffusing through the confocal volume emit bursts of photons, i.e. photons of high count-rate and short inter-photon times. Therefore, if one monitors the inter-photon time trace, transit of a single-molecule through confocal volume will be detected as a drop in inter-photon time trace. Since photon counts follow Poisson distribution, their number can suddenly drop after the beginning of the burst, or before the end of a burst. This would lead to wrong identification of the burst edges. Therefore, the inter-photon time trace is usually smoothed prior to the burst search, by using a filter developed by Lee.[40, 41] Using *m* photons before and *m* photons after each of the photons *i*, a running mean $\overline{\Delta t_i}$ and variance σ_i^2 of the inter-photon times Δt_i is computed. Based on these, smoothed data $\widehat{\Delta t_i}$ is obtained:

$$\widehat{\Delta t_i} = \overline{\Delta t_i} + (\Delta t_i - \overline{\Delta t_i}) \frac{\sigma_i^2}{\sigma_i^2 + \sigma_0^2} \tag{2.16}$$

with σ_0^2 being the variance of background inter-photon times. From the equation above, it follows that for small SNR, $\sigma_i^2 \ll \sigma_0^2$ and $\widehat{\Delta t_i} \approx \overline{\Delta t_i}$. In other words, data is replaced with its local mean value. On the other hand, for high SNR, $\sigma_i^2/(\sigma_i^2 + \sigma_0^2) \approx 1$ and $\widehat{\Delta t_i} \approx \Delta t_i$. In other words, data remains unchanged. For such smoothed trace of inter-photon times, bursts are selected using an upper inter-photon time threshold, while simultaneously requiring minimum number of photons *N* in a burst. Alternatively, one can employ the "sliding-time-window" method,[42] where the local count-rate is calculated around each of the photons, within a specified time window centered around it. If the local count-rate exceeds the threshold, the central photon is then assigned to a burst. For stream of photons to be considered a burst, it is required that at least *N* successive photons exceed the local count-rate threshold.

Burst search can be performed on all photons, that are merged in a single trace, irrespective of their micro-time window origin.[43] This approach is referred to as All-Photon-Burst-Search (APBS). This is the most common burst-search approach, where all species, i.e. double labelled but also single labelled species are accounted. Dual-Channel-Burst-Search (DCBS)[42] method monitors two photon traces in parallel: (*i*) sum of all photons detected after donor laser time-on (G|G and R|G) and (*ii*) acceptor photons detected after acceptor laser time-on (R|R). For each of the two photon traces, separate burst search is performed, and only those time periods where bursts from two photon traces overlap define burst duration. That way we select only those species where both donor and acceptor are active, meaning that single-labelled species will be discarded, as well the species with unstable emission (bleached and blinking species). If the system is unknown, it is advised to use APBS to get an insight into the system and presence of artifacts.

Selected bursts can be subsequently analyzed in terms of duration, brightness, polarization, lifetime and quantities such as FRET efficiency and stoichiometry:

$$E = \frac{F_{\mathrm{R|G}}}{F_{\mathrm{R|G}} + F_{\mathrm{G|G}}} \tag{2.17}$$

$$S = \frac{F_{\mathrm{R|G}} + F_{\mathrm{G|G}}}{F_{\mathrm{R|G}} + F_{\mathrm{G|G}} + F_{\mathrm{R|R}}} \tag{2.18}$$

$F_{\mathrm{R|G}}$ denotes the fluorescence intensity of acceptor upon donor laser time-on, also known as FRET sensitized acceptor fluorescence, corrected for background, donor crosstalk into acceptor detection channel, α, and for direct excitation of acceptor with donor excitation laser, δ:

$$F_{\mathrm{R|G}} = {}^{ii}I_{\mathrm{R|G}} - \alpha\,{}^{ii}I_{\mathrm{G|G}} - \delta\,{}^{ii}I_{\mathrm{R|R}} \tag{2.19}$$

${}^{ii}I_{\mathrm{R|G}}$ denotes background corrected intensity, and in general terms, for a detection channel Y upon excitation X, it is defined as:

$${}^{ii}I_{\mathrm{Y|X}} = {}^{i}I_{\mathrm{Y|X}} - I_{\mathrm{Y|X}}^{(\mathrm{BG})} \tag{2.20}$$

with ${}^{i}I_{\mathrm{Y|X}}$ being raw, uncorrected fluorescence intensity.

Donor crosstalk into acceptor detection channel, α, is defined as ratio of detection efficiency of red and green detection path ($g_{\mathrm{R|D}}$ and $g_{\mathrm{G|D}}$, respectively) for donor dye:

$$\alpha = \frac{g_{\mathrm{R|D}}}{g_{\mathrm{G|D}}} \tag{2.21}$$

Direct excitation of acceptor with donor excitation laser, δ, is defined via acceptor excitation cross-sections for green and red-light source, $\sigma_{\mathrm{A|G}}$ and $\sigma_{\mathrm{A|R}}$, and excitation light irradiance, L_{G} and L_{R}:

$$\delta = \frac{\sigma_{\mathrm{A|G}}\,L_{\mathrm{G}}}{\sigma_{\mathrm{A|R}}\,L_{\mathrm{R}}} \tag{2.22}$$

Further definitions of crosstalk and direct excitation, and practical considerations for their determination are discussed in **Chapter 6.3** and Supplement C.

Furthermore, FRET efficiency and stoichiometry contain $F_{\mathrm{G|G}}$ term, which denotes donor fluorescence intensity after donor laser time-on:

$$F_{\mathrm{G|G}} = \gamma\,{}^{ii}I_{\mathrm{G|G}} \tag{2.23}$$

where correction factor γ compensates for different fluorescence quantum yields of donor and acceptor, and for different sensitivity of green detection channels for donor dye and red detection channels for acceptor dye:

$$\gamma = \frac{\Phi_{\mathrm{F,A}}}{\Phi_{\mathrm{F,D}}}\frac{g_{\mathrm{R|A}}}{g_{\mathrm{G|D}}} \tag{2.24}$$

The last term in definition of E and S is $F_{\mathrm{R|R}}$, fluorescence intensity of acceptor upon acceptor laser time-on:

$$F_{\mathrm{R|R}} = \frac{1}{\beta}\, {}^{ii}I_{\mathrm{R|R}} \tag{2.25}$$

where correction factor β compensates for different excitation rates of donor and acceptor with their corresponding excitation light with irradiances L_{G} and L_{R}:

$$\beta = \frac{\sigma_{\mathrm{A|R}}\, L_{\mathrm{R}}}{\sigma_{\mathrm{D|G}}\, L_{\mathrm{G}}} \tag{2.26}$$

Robust determination of the aforementioned correction parameters α, β, γ and δ is a topic of **Chapter 6.3** and Supplement C.

While FRET efficiency could be obtained also in single-color excitation schemes, stoichiometry can only be obtained through dual-color excitation scheme. For single labelled species, i.e. acceptor-only or donor-only, stoichiometry takes value of ~ 0 and ~ 1, respectively. For double labelled species it takes value around ~ 0.5. These properties of stoichiometry allow us to select sub-populations of interest, but also to study the stoichiometry of biomolecular complexes.

2.5. Inferring structural information from fluorescence decays

Structural information is typically obtained from a global analysis of fluorescence decay of donor in the absence $f_{\mathrm{D|D}}^{(\mathrm{D0})}(t)$, and in presence of acceptor $f_{\mathrm{D|D}}^{(\mathrm{DA})}(t)$, where the latter can be factorized as follows[31]:

$$f_{\mathrm{D|D}}^{(\mathrm{DA})}(t) = \varepsilon_{\mathrm{D}}(t) f_{\mathrm{D|D}}^{(\mathrm{D0})}(t) \tag{2.27}$$

$\varepsilon_{\mathrm{D}}(t)$ denotes the so-called FRET-induced donor decay, which informs on the time-scale of FRET and fraction of molecules that undergo FRET process.

The time-dependent fluorescence intensity of donor at time *t* after donor excitation, $f_{\mathrm{D|D}}(t)$, is proportional to the radiative rate constant of the fluorescence, $k_{\mathrm{D,f}}$, and the time-dependent population of the fluorescence excited state, $p_{\mathrm{D|D}}(t)$, i.e. $f_{\mathrm{D|D}}(t) = k_{\mathrm{D,f}}\, p_{\mathrm{D|D}}(t)$. For donor in the absence and presence of acceptor, the excited state depopulates in time according to:

$$p_{\mathrm{D|D}}^{(\mathrm{D0})}(t) = e^{-k_{\mathrm{D}}t} \tag{2.28}$$

$$p_{\mathrm{D|D}}^{(\mathrm{DA})}(t) = e^{-(k_{\mathrm{D}}+k_{\mathrm{FRET}})t} \tag{2.29}$$

where k_{D} is sum of rate constants of all deactivation processes of donor excited state, other than k_{FRET}. This allows us to express $f_{\mathrm{D|D}}^{(\mathrm{D0})}(t)$ and $f_{\mathrm{D|D}}^{(\mathrm{DA})}(t)$ as:

$$f_{\mathrm{D|D}}^{(\mathrm{D0})}(t) = k_{\mathrm{D,f}} e^{-k_{\mathrm{D}}t} \tag{2.30}$$

$$f_{\mathrm{D|D}}^{(\mathrm{DA})}(t) = k_{\mathrm{D,f}} e^{-(k_{\mathrm{D}}+k_{\mathrm{FRET}})t} \tag{2.31}$$

from where it follows that $\varepsilon_{\mathrm{D}}(t) = e^{-k_{\mathrm{FRET}}t}$. The essential part of fluorescence decay analysis is how $\varepsilon_{\mathrm{D}}(t)$ is modelled. In a case of a mixture of *N* FRET species, with state specific FRET rate constants $k_{\mathrm{FRET},i}$ and corresponding species fractions $x_{\mathrm{FRET},i}$, the FRET-induced donor decay can be represented as species fraction weighted sum, based on the additive property of fluorescence intensities:

$$\varepsilon_{\mathrm{D}}(t) = \sum_{i=1}^{N} x_{\mathrm{FRET},i}\, e^{-k_{\mathrm{FRET}}t} \tag{2.32}$$

For continuous distribution of FRET rate constants, $x(k_{\mathrm{FRET}})$, the sum is replaced by an integral:

$$\varepsilon_{\mathrm{D}}(t) = \int x(k_{\mathrm{FRET}}) e^{-k_{\mathrm{FRET}} t} dk_{\mathrm{FRET}} \tag{2.33}$$

Taking advantage of the Förster equation (see eq. 2.3), we can convert the time axis into inter-dye distance axis, i.e.:

$$\varepsilon_{\mathrm{D}}(R_{\mathrm{DA}}) = \int x(R_{\mathrm{DA}}) e^{\left[-\frac{t}{\tau_{\mathrm{D}(0)}}\left(\frac{R_0}{R_{\mathrm{DA}}}\right)^6\right]} dR_{\mathrm{DA}} \tag{2.34}$$

which, however, requires the analytical model for the distribution of inter-dye distances, $x(R_{\mathrm{DA}})$. For normally distributed dye positions, distribution of inter-dye distances $x(R_{\mathrm{DA}})$ is then given as the non-central χ- distribution[22], i.e. $x(R_{\mathrm{DA}}) = \chi\left(R_{\mathrm{DA}}|R_{\mathrm{mp}}, \sigma_{\mathrm{DA}}\right)$, where the non-centrality parameter is the distance between the mean dye positions, R_{mp}, and the width parameter is σ_{DA}, i.e.:

$$\chi\left(R_{\mathrm{DA}}|R_{\mathrm{mp}}, \sigma_{\mathrm{DA}}\right) = \frac{R_{\mathrm{DA}}}{R_{\mathrm{mp}}}\left[\mathcal{N}^+\left(R_{\mathrm{DA}}|R_{\mathrm{mp}}, \sigma_{\mathrm{DA}}\right) - \mathcal{N}^+\left(R_{\mathrm{DA}}|-R_{\mathrm{mp}}, \sigma_{\mathrm{DA}}\right)\right] \tag{2.35}$$

$\mathcal{N}^+\left(R_{\mathrm{DA}}|R_{\mathrm{mp}}, \sigma_{\mathrm{DA}}\right)$ represents normal distribution with a mean value R_{mp} and width σ_{DA}, taken only at the positive values of R_{DA}:

$$\mathcal{N}^+\left(R_{\mathrm{DA}}|R_{\mathrm{mp}}, \sigma_{\mathrm{DA}}\right) = \frac{1}{\sigma_{\mathrm{DA}}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{R_{\mathrm{DA}}-R_{\mathrm{mp}}}{\sigma_{\mathrm{DA}}}\right)^2} \tag{2.36}$$

The width parameter σ_{DA} is determined by the widths of spatial distributions of donor and acceptor, i.e. $\sigma_{\mathrm{DA}} = \sqrt{\sigma_{\mathrm{D}}^2 + \sigma_{\mathrm{A}}^2}$. From the study using dyes coupled to nucleic acids, and for unrestricted dye motion, it was found that σ_{D} and σ_{A} take value of ≈ 6 Å[44], which yields σ_{DA} ≈ 8.5 Å.

In the case of narrow spatial distributions of dye positions and large inter-dye distances, the χ-distribution converges to a normal distribution.[22] Using this approximation, and accounting for structural heterogeneity, i.e. a mixture of FRET species, $x(R_{\mathrm{DA}})$ is given as a discrete sum of n_{G} Gaussian distributions, of variable mean $\langle R_{\mathrm{DA},i}\rangle$, width $\sigma_{\mathrm{DA},i}$ and weight x_i:

$$x(R_{\mathrm{DA}}) = \sum_{i=1}^{n_G} x_i \mathcal{N}^{+}\left(R_{\mathrm{DA}} | \langle R_{\mathrm{DA},i} \rangle , \sigma_{\mathrm{DA},i}\right) \tag{2.37}$$

Typically, n_{G} takes value between 1-3, and is determined by comparing the scores of competing models.

However, for small inter-dye distances, and especially in the absence of any knowledge regarding the nature of inter-dye distance distribution, the best is to represent $x(R_{\mathrm{DA}})$ using χ-distribution, or in general case as a mixture of L χ-distributions of variable location R_{mp} and width σ_{DA}:

$$x(R_{\mathrm{DA}}) = \sum_{l=1}^{l=L} \chi\left(R_{\mathrm{DA}} | R_{\mathrm{mp},l}, \sigma_{\mathrm{DA}}\right) \tag{2.38}$$

where the number L of χ-distributions is determined by the bounds of R_{mp} and σ_{DA}. For R_{mp} and σ_{DA} being sampled in a range of [10 Å, 130 Å] and [3 Å, 9 Å], respectively, we obtain L = $(130 - 10)/3$ = 40. With such representation for $x(R_{\mathrm{DA}})$, no presumptions are made with respect to the nature of distribution. This is in particular suitable when the biomolecule does not occupy well defined states, but rather exists as quasi-continuum of interconverting states. Next, we optimize and sample the variable parameters and nuisance parameters, in order to find models M which are consistent with data D. By sampling over variable parameters, we obtain numerous posterior distributions $x(R_{\mathrm{DA}})$ which describe the experiment, and their probability conditioned on the experimental data D, i.e. $p(x(R_{\mathrm{DA}})|D)$. For each inter-dye distance and species fraction bin, we integrate the probabilities of all sampled distributions found in respective bin to obtain $p(x, R_{\mathrm{DA}}|D, M)$ (**Figure 4**). From there, as final representation for experimentally-derived inter-dye distance distribution, $x_{\mathrm{e}}(R_{\mathrm{DA}})$, we take $x(R_{\mathrm{DA}})$ at which we reach 0.5 cumulative sum of $p(x, R_{\mathrm{DA}}|D, M)$, and as $\Delta x_{\mathrm{e}}(R_{\mathrm{DA}})$ we take one standard deviation around $x_{\mathrm{e}}(R_{\mathrm{DA}})$.[45] This approach was used in Publication 2 (see **Chapter 6.2** and Supplement **B**) to infer structural information from fluorescence decays.

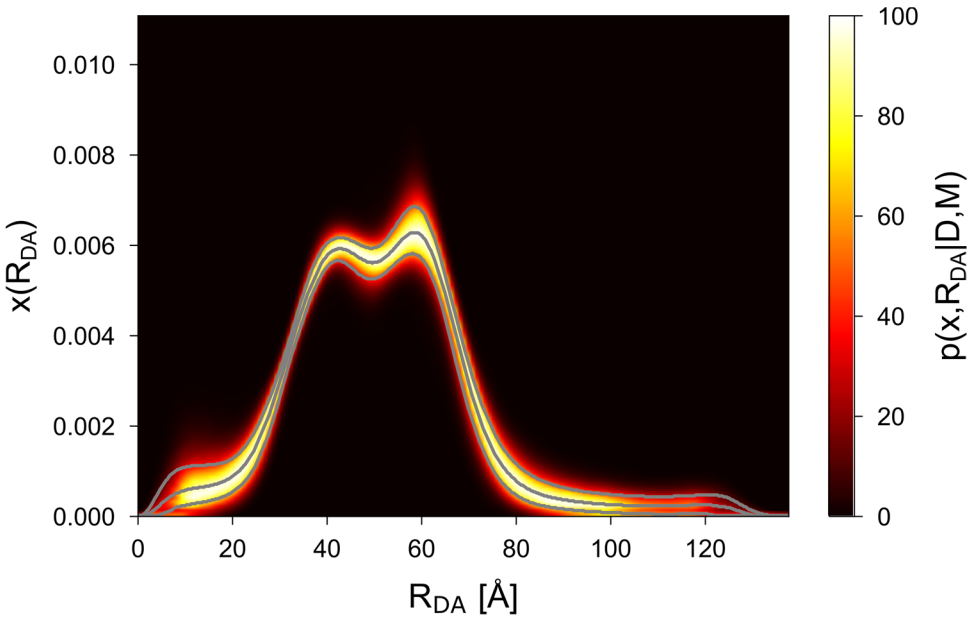


Figure 4 | Inferring distance information from fluorescence decays. *Exemplary inter-dye distance distribution obtained through Bayesian analysis of ensemble-TCSPC data, with grey lines denoting* $x_e(R_{DA})$ *and* $\Delta x_e(R_{DA})$.

Chapter 3

Computational biophysics

3.1. Basics of Molecular Dynamics Simulations

Sampling techniques of molecular conformations can be divided into two categories: stochastic (e.g. Monte Carlo) and deterministic (e.g. Molecular Dynamics).[46, 47] In Monte Carlo simulations, the configurational space is sampled by random displacement of the particles/atoms and the energy change during this displacement serves as a criterium based on which the new coordinates are being accepted or discarded. On the other hand, sampling in molecular dynamics simulations is deterministic, since in each step one solves classical Newton's equations of motions for a given force field, which gives the new velocities and positions. To simulate a system using molecular dynamics, it is necessary to (*i*) model the interactions within the system, (*ii*) choose the appropriate integration scheme based on which velocities and new positions will be propagated, and (*iii*) to choose the statistical ensemble for the simulation of the ensemble.[46] Although the natural choice for molecular dynamics is micro-canonical NVE ensemble, it is also possible to simulate the system in other ensembles as well, such as NVT or NPT. Whatsoever, depending on a system simulated, this might be a better or even a necessary choice of an ensemble, e.g. for systems with lipid bilayers.

Evolution of a system is fully determined by its potential energy $U(r^N)$, which is a function of coordinates r of *N* atoms. Within this thesis, I used AMBER functional form of the potential energy $U(r^N)$, which consists of a bonded and a non-bonded term. The bonded term in AMBER is typically described by three contributions, i.e. terms for bonds, angles and dihedrals[48]:

$$U_{\text{bonded}}(r^N) = \frac{1}{2}\sum_{\text{bonds}} k_{ij}^{r}\left(r_{ij} - r_{\text{eq}}\right)^2 + \frac{1}{2}\sum_{\text{angles}} k_{ijk}^{\theta}\left(\theta_{ijk} - \theta_{\text{eq}}\right)^2 + \frac{1}{2}\sum_{\text{dihedrals}} k_{ijkl}^{\varphi,m}\left(1 + \cos\left(m\varphi_{ijkl} - \gamma_m\right)\right) \tag{3.1}$$

The first term describes the bond stretching, and is a function of the distance r_{ij} between two bonded atoms *i* and *j*. It is typically given as the harmonic potential, with r_{eq} being the equilibrium distance at which the harmonic potential has minimum, and k_{ij}^{r} being the force constant for bond stretching. The second term in potential energy requires to know in each step

three consecutive atomic coordinates *i*, *j* and *k.* It describes the angle bending between the two vectors $\boldsymbol{r}_{ij}$ and $\boldsymbol{r}_{jk}$, which is given as $\cos\theta_{ijk} = \boldsymbol{r}_{ij} \cdot \boldsymbol{r}_{jk}$. Since angles only slightly change around their equilibrium value at room temperature, the angle potential has harmonic form as well, with characteristic force constant k_{ijk}^{θ} and equilibrium angle θ_{eq}. Lastly, the potential describing rotation around the dihedral angles φ_{ijkl}, which is a function of four coordinates *i*, *j*, *k* and *l*, typically involves expansion into periodic function of the order *m*, where *m* represents number of minima in the function. $k_{ijkl}^{\varphi,m}$ represents the height of the barrier for rotation around the dihedral angle, and γ_m is the phase factor, which corresponds to the angle value at the corresponding minimum *m*.

The non-bonded term is calculated between pairs of atoms *i* and *j* that are either in different molecules, or belong to the same molecules, but are separated by at least three bonds. In AMBER, the non-bonded term is given as sum of a pair-wise, two-body 6-12 Lennard-Jones potential and a pair-wise, two-body Coulomb potential, that describe van der Waals and electrostatic interactions, respectively:

$$U_{\mathrm{non-bonded}}(r) = \sum_i \sum_{j \neq i} \left(4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{R_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{R_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon R_{ij}} \right) \tag{3.2}$$

The Lennard-Jones 12-6 potential contains attractive part, ~ R_{ij}^{-6}, and repulsive part, ~ R_{ij}^{-12}, where R_{ij} is the distance between atoms. Furthermore, Lennard-Jones potential is a function of a diameter σ_{ij}, which is the distance between the atoms at which the potential is zero, and ε_{ij}, which the depth of the well. Both σ_{ij} and ε_{ij} are a characteristic of a pair of atoms, and are determined by Lorentz/Berthelot mixing rules.

Coulomb potential for electrostatic interactions decays as R_{ij}^{-1}, with R_{ij} being the distance between the charges q_i and q_j, in a medium with dielectric constant ε.

Equations 3.1 and 3.2 represent a very minimalistic description of the way interactions are modelled within a classical force field often used in MD simulations, originally proposed by Cornell *et al.*[48] Nevertheless, such representation of interactions is adequate for the description of most systems. For efficient calculation of long-range and short-range interaction, various computational algorithms, such as Ewald summation[49] and Verlet lists[50], were developed. Depending on the force field used, functional shape of the bonded and non-bonded potential can vary, as well the values of atom-specific parameters. Atom-specific parameters are derived either from quantum chemical calculations, or directly from experiments, or are tuned to fit

some experimental observable. Although it is in principle possible to "mix and match" parametrizations for biomolecule, ions and solvent, these usually come together in a package, and work the best in a given combination. For example, the relatively new FF99SBdisp force field[51], parametrized for both folded and unfolded proteins, should be used with its "native" water model, which is the modified TIP4P-D water model, also known as "disp-water". For a more detailed overview of force field models, I refer the reader to the comprehensive book by Leach, A. R.[47]

Once the potential energy for modelling the interactions is defined, the next step is to calculate the forces acting on each of the particles in the system:

$$f = -\frac{\partial U(r)}{\partial r} \tag{3.3}$$

and to solve Newton's equations of motion. One of the most frequently used algorithms for solving the equations of motions is the "velocity-Verlet" algorithm.[52] Main requirement of this algorithm is the use of a short time-step, which means that the forces are frequently calculated. To get to the formula that defines new atomic coordinates, we start from the Taylor expansion of the atomic coordinate around the time *t*:

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{f(t)}{m}\frac{\Delta t^2}{2!} + \mathcal{O}(\Delta t^3) \tag{3.4}$$

Estimation of new coordinates has an error that is on the order of Δt^3, with Δt being the time step in equations of motion. To calculate new atomic positions, one needs to know current coordinates, current velocities, as well the force acting on the particles in current coordinates.

Besides being necessary for the calculation of new coordinates, velocities are also needed for calculation of kinetic and total energy, as well the instantaneous temperature. To derive velocities at $t + \Delta t$ instance in time, we start from the Taylor expansion of velocity around the time *t*:

$$v(t + \Delta t) = v(t) + \dot{v}(t)\Delta t + \ddot{v}(t)\frac{\Delta t^2}{2!} + \mathcal{O}(\Delta t^3) \tag{3.5}$$

Now, to get rid of high-order derivative $\ddot{v}(t)$, we expand $\dot{v}(t + \Delta t)$ into a Taylor series:

$$\dot{v}(t + \Delta t) = \dot{v}(t) + \ddot{v}(t)\Delta t + \mathcal{O}(\Delta t^2) \tag{3.6}$$

and after multiplying both sides of the equation with $\Delta t/2$, we obtain:

$$\ddot{v}(t)\frac{\Delta t^2}{2} = \left(\dot{v}(t+\Delta t) - \dot{v}(t) + \mathcal{O}(\Delta t^3)\right)\frac{\Delta t}{2} \tag{3.7}$$

Finally, after plugging the latter expression for $\ddot{v}(t)$ back into eq. 3.5, we get:

$$v(t+\Delta t) = v(t) + \left(\dot{v}(t) + \dot{v}(t+\Delta t)\right)\frac{\Delta t}{2} + \mathcal{O}(\Delta t^3) \tag{3.8}$$

or, expressed in terms of forces:

$$v(t+\Delta t) = v(t) + \left(f(t) + f(t+\Delta t)\right)\frac{\Delta t}{2m} + \mathcal{O}(\Delta t^3) \tag{3.9}$$

To summarize, in the "velocity-Verlet" algorithm, new coordinates are calculated (eq. 3.4), which is followed by calculation of forces in the updated positions. Lastly, using forces in current and updated positions, velocities are updated (eq. 3.9). Such expressions for positions and velocity have reasonable error on the order of Δt^3. Furthermore, both positions and velocities are computed for the same instance of time.

Besides "velocity-Verlet" algorithm, numerous other integration schemes were developed, and they all in essence differ in the order of the derivative at which the Taylor expansions of $r(t+\Delta t)$ and $v(t+\Delta t)$ are truncated, and consequently in their accuracy.[46] Furthermore, they differ in the time instance at which velocities are calculated. For example, in the basic Verlet algorithm, new positions are calculated with smaller error of $\mathcal{O}(\Delta t^4)$, but velocities are calculated late in respect to the positions, i.e. at the time $\underline{t}$, $v(t)$. In "leap-frog" method[53], velocities are computed half a step ahead of the coordinates, i.e. $v\left(t+\frac{\Delta t}{2}\right)$, which is then used to calculate new positions $r(t+\Delta t)$. In principle, all of these schemes are computationally efficient. Where they differ is actually: (*a*) level of accuracy for both short and large time step (*b*) time reversibility, and (c) conservation of volume and energy.[46]

3.2. MD trajectory analysis: Clustering

For a series of purposes, it is desirable to reduce the number of structures obtained from MD simulations. For that, a wide range of clustering methods have been developed. Clustering is one of the most widespread methods of unsupervised machine learning, with implications in diverse fields. In terms of data science, clustering represents assigning unlabelled data to groups,[54] and in biophysical/statistical terms it means e.g. to choose for each of the molecular configurations to which conformational state it belongs. However, reducing the number of structures is rather a "side product", because often the essential motivation behind clustering is to find hidden structure in data.

Clustering of conformational models, obtained from e.g. MD simulations, is based on some measure of similarity between different system configurations. Measures of similarity can be structural or kinetic. In the case of kinetic measures, we group configurations that are close to one another in time. Much more used are structural measures of similarity, such as: cartesian coordinates, inter-residue distances, dihedral angles etc. Using the RMSD on cartesian coordinates is the most frequent choice of structural similarity, however, not as trivial, since it requires suitable alignment of structures. Two most popular clustering approaches are K-means and hierarchical clustering.[55, 56] K-means clustering algorithm initiates k centers (means) at random, and then it creates k clusters by assigning structures to the closest mean. After assigning structures to the k clusters, positions of means/centroids are updates. The latter two steps are repeated until convergence. K-means is a stochastic method, meaning that we do not get twice the same result, since the means are initiated at random. Further limit is that it only performs well on spherical and well-defined clusters. In terms of hyperparameters, it requires as input the number of clusters k, which is (typically) not known *a priori*. However, there are evaluation metrics one can use to determine the value of k parameter, such as the elbow method or silhouette score, where in the latter case k is chosen such that small within the cluster distance and large distance to the closest cluster is obtained.[57]

Hierarchical clustering can be done in two ways: top-down (divisive) and bottom-up (agglomerative). Within Publication 2 (see **Chapter 6.2** and Supplement B), I have used the agglomerative hierarchical clustering, i.e. bottom-up approach, where all structures are initially taken as single clusters, and then using some similarity measure they are being merged into new clusters, until one big cluster, containing all original clusters, is left. In other words, a hierarchy of clusters is being formed, which is often illustrated in a shape of a tree, also known as *dendrogram*.[55] Hierarchical clustering method requires to choose so-called "linkage method", which is a way of calculating the distance/dissimilarity between the data points in terms of some

feature. Some of the most popular linkage methods are: “single linkage”, “complete linkage”, and “average linkage”.[55, 58]

In the case of “single linkage” method[59], similarity measure between two clusters A and B is the closest distance between any pair of points of the two clusters, i.e.:

$$d(A,B) = \min_{a\in A, b\in B} d(a,b) \tag{3.10}$$

In the case of “complete linkage”[60], distance between two clusters is the maximal distance between any pair of points between the two clusters:

$$d(A,B) = \max_{a\in A, b\in B} d(a,b) \tag{3.11}$$

Lastly, in the case of “average linkage”[61], computed is the average distance between all pairs of points between two clusters

$$d(A,B) = \frac{1}{|A||B|} \sum_{a\in A, b\in B} d(a,b) \tag{3.12}$$

Complete and average linkage methods tend to create more compact clusters, compared to single linkage method. Hierarchical clustering can be performed in two ways, i.e. either we define number of clusters in advance, and then clusters are merged in the dendrogram till the defined number of clusters is met, or a full dendrogram is constructed and then a certain threshold is applied, and only those clusters that were formed at distance below the given threshold will pass. Typically, the latter approach is used.

Chapter 4

Inference methods in structural biology

4.1. Traditional techniques of structural biology

X-ray crystallography. For a long time, X-ray crystallography[62, 63] has dominated the field of structural biology. By not being limited in size of a system, it provided structural models of biomolecules ranging from few kDa to tens of MDa. This, in addition to being able to derive absolute atomic coordinates without having to employ extensive modelling procedures on raw data, largely explains the long-term dominance of X-ray crystallography. In X-ray crystallography experiment, the crucial and limiting step is the sample preparation, where the protein of interest is dissolved, and then its aggregation, nucleation and crystal formation are promoted.[64] The obtained 3D crystal is exposed to the beam of X-rays, which are according to the Bragg's law[65] diffracted upon interaction with the electrons around each of the atoms. Angle and intensity of diffraction spots, known as diffraction pattern, is being detected, based on which the structure of biomolecule can be inferred. The quality and resolution of model directly depends on the sharpness of diffraction spots, which in turn is determined by the quality of the crystal itself. Preparing a 3D crystal of sufficient size and quality, i.e. of regular/uniform structure, is far from trivial task. Many proteins defy crystallization, such as the membrane proteins of limited solubility.[64] While obtained models are of a high resolution, the ultimate limit of a technique, however, is the fact that only model of a single structure can be obtained, since measurements are performed on crystals, i.e. periodic arrays of proteins in single and static conformational state. However, in recent years efforts were made to overcome this issue through development of so-called time-resolved X-ray approaches.[66] Despite its limitations, X- ray crystallography yielded invaluable insights into biomolecular structures, and for decades it guided advances in a range of disciplines, and will continue to do so.

NMR spectroscopy. In the 1990s, NMR[67, 68] appeared as competing approach of structural biology to X-ray crystallography (**Figure 5A**), by having the advantage of measuring molecules directly in a solution. That way, NMR has capability of providing multi-state models with corresponding kinetic scheme. NMR spectra appear for atoms that have non-zero spin of the nuclei, i.e. odd number of protons and/or neutrons, upon exposure to the radio waves of the external magnetic field, whose frequency is equal to the energy difference between the nuclear spin levels. An NMR spectrum has the following characteristics: position of the lines (so-called

chemical shift), their width and multiplicity/splitting. Each of these spectral properties are affected by the local environment of the nuclei, presence of interactions, structural exchange (e.g. cis-trans exchange or between two conformational states of biomolecule) and its energy barrier.[69-71] While NMR provides high-resolution structural information on atomic/residue level, and in physiological conditions, due to the complexity of a signal and difficulty to interpret it, solution NMR is rather limited with the size of biomolecule, and typically reports on smaller size systems. For example, for structures of < 20 kDa, there are over 12 000 NMR structures deposited in PDB, but this number rapidly declines to a couple dozen structures, for molecular weights of > 40 kDa.

Cryo-Electron Microscopy (Cryo-EM). Within three decades, NMR was overtaken by rapidly growing field of cryo-electron microscopy (cryo-EM) in terms of number of deposited structures in PDB (see **Figure 5A**).[72-74] Although cryo-EM measurements are not done in solution, flash-freezing of the sample is performed using liquid nitrogen, as opposed to crystallization in X-ray crystallography. That way, instead of having all molecules in a single, static pose, cryo-EM captures molecules in multiple states. Since snapshots are static, the kinetic scheme that underlies protein function is inaccessible. However, similar to X-ray crystallography, cryo-EM is in principle adaptable to time-resolved studies, and some attempts were done in that direction.[75] In cryo-EM experiment sample is exposed to the beam of high-energy electrons, whose scattering profile is collected as stack of 2D images, that are aligned, averaged and then used to infer the 3D structure of biomolecule, or so-called 3D electron density map. Cryo-EM reconstructions are mainly of medium- to low-resolution, and most frequently maps with 3-4 Å resolution are reported, although exceptionally sharp reconstructions with features close to an atom size were reported as well.[76] For a comparison, at the resolution of ~ 5-7 Å secondary structure elements become visible, while side-chains become visible only at < 3 Å resolution.[4] Naturally, cryo-EM is biased towards large biomolecular machineries, e.g. supramolecular complexes consisted of dozens of proteins, as these can handle electron bombardment. For example, cryo-EM reported on > 6000 structures with molecular weight > 380 kDa, but on less than 100 structures with weight of < 20 kDa.

Statistics related to number of structures, molecular sizes and resolution of models for each of here presented techniques is given based on entries available in PDB by the April, 2023.

4.2. Deep-learning single-structure predictions

Christian Anfinsen postulated that, theoretically, one-dimensional amino acid sequence of a protein fully determines its 3D structure.[77] Ever since, triggered with this hypothesis, numerous efforts were invested to solve the "protein-folding problem", using various methodologies from the field of computer science. Structure prediction was initially shaped by two approaches, namely: (a) by modelling interactions between molecular constituents and which drive a protein towards its thermodynamically stable state, and (b) via bioinformatical analyses of evolutionary history of proteins.[78] Both approaches produced far inferior levels of accuracy compared to experimental approaches. To facilitate and advance the progress in predicting the structures of single proteins and domains, the Critical Assessment of Structure Prediction (CASP) challenge is held biannually (https://predictioncenter.org/), and puts on a test bench hundreds of modelling methods. Participants of CASP challenge are asked to blindly predict structures of targets divided in several categories according to their complexity, and for which experimental structure is determined but not yet publicly available. As independent model evaluation metric, CASP uses the Global Distance Test (GDT)[79]. This metric is a measure of model accuracy, and can be understood as a percentage of amino acids that are within a threshold distance from experimentally determined positions (ground truth). Since the beginning of the CASP in 1997, the best performing predictors were having median accuracy of ~ 50% in GDT units, across all targets. Exceptional results were observed in CASP14 challenge in the year 2020, when AlphaFold2[78] achieved median accuracy of 92.4% in GDT units, across all targets, which is considered competitive to experimentally derived structures.[80] In terms of physical units, AlphaFold achieved median C_α-RMSD of 0.96 Å and all-atom-RMSD of 1.5 Å for 95% residue coverage, which is in the range of the width of carbon atom itself.[78] AlphaFold2, whose exceptional performance originates from the use of novel neural networks trained on physical, geometrical and evolutionary constraints,[78] has ever since inferred structures for over 200 million proteins that are deposited with open access in AlphaFold Protein Structure Database (https://alphafold.ebi.ac.uk/) and in PDB. Not long after, several competing modelling methods reached or even outperformed the accuracy, speed and computational cost of AlphaFold2, and such an example is trRosettaX-Single.[81]

While one can say with confidence that prediction of protein structure reached the "single-structure frontier" [82], there are many remaining challenges, from multi-protein complexes, protein-ligand complexes, proteins with intrinsically disordered regions that transform to functionally relevant states only upon interaction with substrates, and the ultimate challenge of structural biology: multi-state models and the time-ordered ensembles of structures.

4.3. Integrative modelling of structural ensembles

Judging by the number of protein structures inferred by deep learning-based methods, it might appear that the future of structural biology "will not be crystallized",[72] and that computer science will dominate the structural biology. Current trends in structural biology, however, rather indicate that traditional techniques are here to stay, but in integrative manner, such that advantages of each of the techniques, including X-ray crystallography, are harvested in the best possible way to complement each other.[2-4] For example, lower-resolution cryo-EM 3D density maps can serve as information on overall topology of a complex, and its individual components can be modelled with high-resolution by e.g. X-ray crystallography or NMR spectroscopy, and docked into cryo-EM 3D density map.[83] Another typical example of complementary techniques is combination of FRET and SAXS[84-86] measurements, particularly for studies of multidomain proteins with flexible linkers or intrinsically disordered proteins.[5, 8] SAXS, similarly to cryo-EM, reports on low-resolution features, such as the overall shape of the molecule, and can be aided by FRET that reports on inter-residue distances, to distinguish between species with overall same shape but rotated domains, and which can represent two functionally different species. Many more techniques are arising in recent years, and in integrative manner with other experimental methods and/or computational approaches, they report multi-scale and/or multi-state models. These methods provide information of varying level of detail, from overall molecular shape (cryo-EM, SAXS), over interface mapping (mutagenesis[87], hydrogen-deuterium exchange (HDX)[88]) to inter-residue distance information (cross-linking mass-spectrometry (XL-MS)[89-91], FRET, EPR[92]). Methods for optimal integration of different sources of information[93] and PDB-Dev system for archiving and dissemination of integrative models is under ongoing developement.[94-96] While the number of deposited integrative models in PDB- Dev is still limited, one can already observe vast diversity, in sense of different model representations (coarse-grained, atomistic, volumetric etc.) and in terms of complexity of studied systems (**Figure 5B**). To get an insight into which experimental methods typically combine with each other, I have sourced the reports on all deposited models from PDB-Dev, extracted the type of input data that was used in each of the models, found all the unique combinations of input data types, and calculated their frequency (**Figure 5C**). This analysis revealed the complexity of the integrative modelling, with 63 unique ways of combining different data types, and cross-linking mass-spectrometry (CX-MS) being the most paired-with method. FRET typically pairs with techniques that inform on overall shape of the molecule, such as SAXS and cryo-EM.

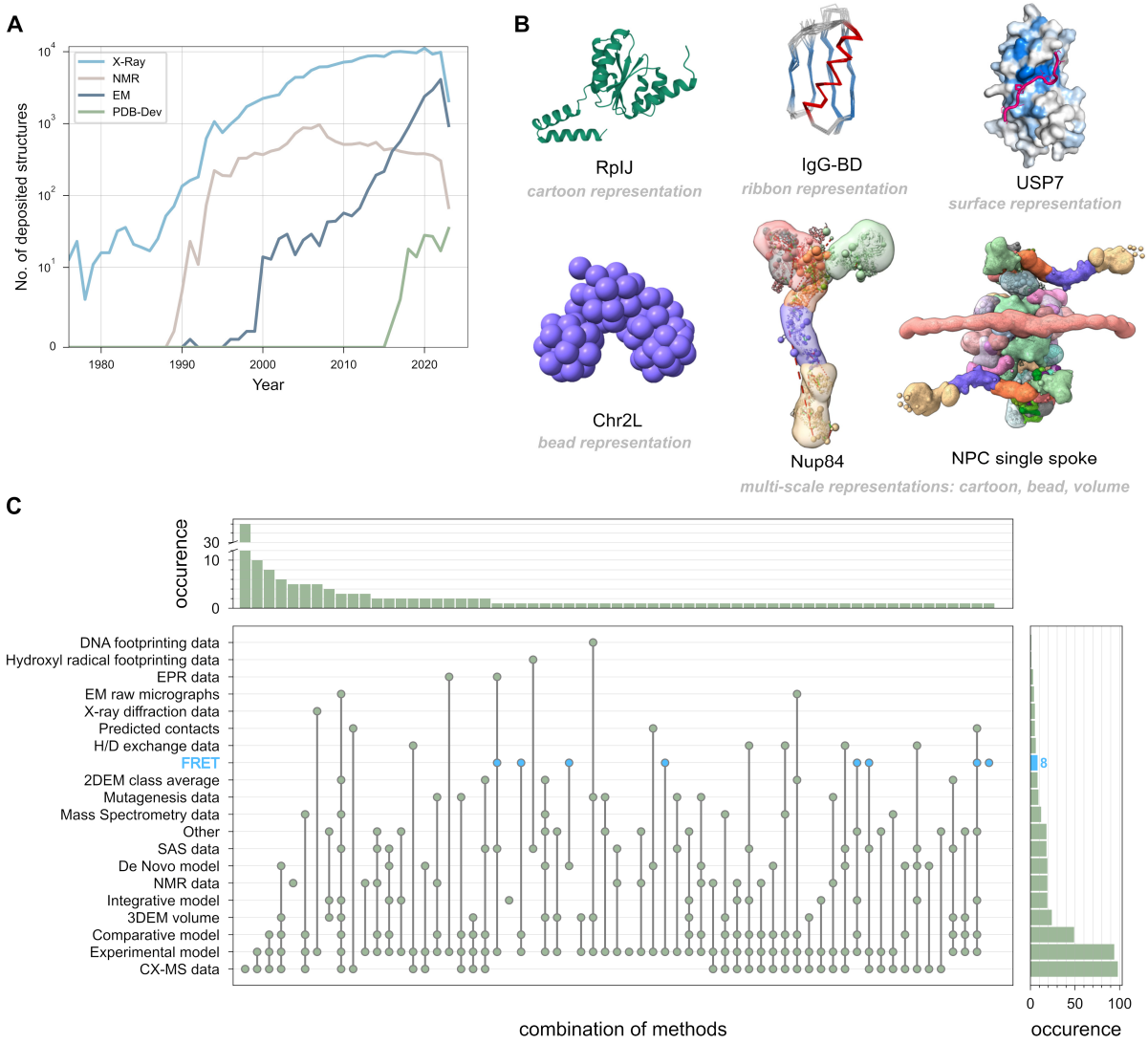


Figure 5 | Current trends in integrative modelling. ***(A)*** *Number of deposited structures per year in PDB by X-ray, NMR (solution or solid-state) and EM (electron microscopy or electron crystallography or electron tomography), compared to the number of integrative models deposited in PDB-Dev. The data reveal that while X-ray still dominates the field of structural biology, it is being challenged by rapidly growing method of cryo-EM. Moreover, through extensive collaborations, integrative characterizations of biomolecules are being carried out, and already over 100 hybrid multi-scale and/or multi-state models are disseminated to the public through PDB-Dev archiving system.* ***(B)*** *Glimpse into the different levels of complexity of models deposited in PDB-Dev: model of E. coli RpIJ (PDB-Dev entry 168), model B1 of IgG-binding domain of protein G (PDB-Dev entry 84), TRAF domain of USP7 in complex with DNA polymerase (PDB-Dev entry 70), structure of a Fly genome Chromosome 2L (PDB-Dev entry 8), followed by the structure of Nup84 sub-complex of NPC (PDB-Dev entry 1) and anatomy of the single spoke of NPC (PDB-Dev entry 10). Illustrations are taken from PDB-Dev.* ***(C)*** *Upset plot of unique combinations (63) of input data types used in integrative modelling, and their frequency, revealing that cross-linking mass-spectrometry (CX-MS) is the leading, but also the most paired-with method. Panels* ***(A)*** *and* ***(C)*** *are generated by the analysis of entries available in PDB and PDB-Dev by the April, 2023.*

We join the outcomes of different experimental and computation techniques not only because they sense different properties of a biomolecule, but also because proteins do not adopt a single structural state. Instead, they display a distribution of dynamic structures, or so-called structural ensemble. Structural ensemble is a set of conformational models with their corresponding population fractions (weights), which denote to which extent biomolecule populates corresponding state under given experimental conditions. Such an ensemble reflects true/intrinsic structural heterogeneity of the molecule, and should be distinguished from so-called "uncertainty ensembles", where multiple reported models rather represent insufficient information about the system, i.e. ambiguity of the solution. Currently, PDB-Dev counts 13 deposited multi-state models, 7 of which use single-molecule FRET measurements as input data.

Structural ensembles are extremely challenging to obtain at accurate or detailed enough level by experimental techniques or computational methods alone. Many experimental methods suffer from spatial and/or temporal averaging, in addition to being sensitive on limited spatial and/or temporal scales. Furthermore, experimental data can be sparse and provide ambiguous information. For example, FRET experiments inform on long-range features, such as tertiary, super-tertiary and quaternary structure, and not on absolute atomistic coordinates. Similar holds for techniques that inform on low-resolution features such as overall molecular shape (SAXS, cryo-EM). Likewise, computational methods suffer from inaccurate force fields and limited sampling times, which has a consequence that relevant states may never be sampled. Therefore, the most promising approach is to combine experimental data with structural models obtained from computational approaches.[97] Numerous approaches have been developed to combine these two sources of information, and they are all grouped around two principles, namely maximum entropy and maximum parsimony principle.[93, 98] In maximum entropy principle, one searches for such model which will be as close as possible to the initial one, while improving the agreement to the experimental data. The maximum entropy principle typically yields large number of conformers, which is suitable for systems that do not occupy well defined states, but exist in the solution as quasi-continuum of interconverting states. Contrary to that, in maximum parsimony approaches, one searches for a minimum number of states that can describe the experimental data. When using either of the two principles, information from different sources can be combined in two ways. In one case experimental data can be used directly as a constraint during the sampling of structural models in molecular simulations. In another case, experimental data and molecular models are generated independently from each other, and then the experimental data act *a posteriori* by refining/reweighting the ensemble from molecular

simulations (**Figure 6**).[99] In this thesis, I have used maximum entropy method to refine the conformational ensemble of lipase-specific foldase from *Pseudomonas aeruginosa* species in its apo state (for details on system see **Chapter 5.4**). As prior information in ensemble optimization I used set of conformational models obtained from all-atom MD simulations, and as experimental data I used time-resolved ensemble TCSPC measurements, acquired for a network of double-labelled cysteine variants (for details see **Chapter 6.2** and Supplement B). The principles of MEM-based ensemble optimization are described in the **Chapter 4.4**.

4.4. Ensemble reweighting methods

4.4.1. Maximum Entropy Method. Our understanding of biological systems is usually given in a shape of some model. Model, e.g. a distribution of some observable across an ensemble of structures, reflects the current state of knowledge about a system. With newly acquired data, often an update of a model/distribution has to be performed, such that it is compatible with new information available. This is a frequent challenge, given the growth of both experimental and computational methods. Bayes and Jaynes both provided very fundamental principles on how to find an updated distribution of an observable, which will be in accordance with new information. These principles are the pillars of integrative structural biology.[98]

Building upon Shannon's information theory, Jaynes established the principle on how to find the best distribution of an observable, amongst infinite number of them satisfying observed data.[100, 101] He suggested that out of all distributions, one that maintains highest uncertainty about the observable of interest, should be chosen.[102] The uncertainty, that is to be maximized, is defined as information entropy, *S*, introduced by Shannon.[101] Information entropy can be defined in terms of some discrete variable x_i as follows:

$$S(p) = -\sum_{i=1}^{N} p(x_i) log\ p(x_i) \tag{4.1}$$

One should not confuse information entropy with the thermodynamic entropy, although there is some resemblance between the two, in that both capture the randomness. There are many properties of this mathematical construct that make it the (only) good measure of uncertainty, with the fundamental ones being: (a) S vanishes as we become more and more certain about a particular outcome, i.e. $S \to 0$ when $p(x_i) \to 1$, (b) it increases the less we know about the underlying distribution, and in fact it reaches the maximum value of $\log N$ when we know

nothing about the distribution, or in other words, when the distribution is uniform, i.e. $p(x_i) = \frac{1}{N}$.

A set of conditions defining the optimal distribution according to maximum entropy principle can be summarized as follows:

$$\begin{cases} p_{\mathrm{ME}}(x) = \underset{p(x)}{arg\,max}\, S(p) \\ \sum_{i=1}^{N} p_{\mathrm{ME}}(x_i) = 1 \\ \sum_{i=1}^{N} f_k(x_i) p_{\mathrm{ME}}(x_i) = \langle f_k(x) \rangle = \langle f_k(x) \rangle^{\mathrm{exp}};\ k = 1, \dots K \end{cases} \tag{4.2}$$

According to the set of conditions given in eq. 4.2: (*i*) the optimal distribution, $p_{\mathrm{ME}}(x)$, maximizes information entropy (uncertainty), (*ii*) it sums to one, i.e. it is probability distribution, and lastly, (*iii*) the ensemble average $\langle f_k(x) \rangle$ computed for $p_{\mathrm{ME}}(x)$ is constrained by *K* experimentally derived expectation values $\langle f_k(x) \rangle^{\mathrm{exp}}$.

However, a common scenario is that there is already some existing model, called *prior* distribution, which, upon acquiring additional experimental restraints is to be converted to a *posterior* distribution, if such modification is necessary. In that case, the maximum entropy principle is rather expressed in terms of relative entropy or negative Kullback-Leibler divergence[103] $D_{\mathrm{KL}}(p_1 || p_0)$ between a prior, $p_0(x_i)$, and posterior probability distribution, $p_1(x_i)$:

$$S(p_1 || p_0) = -\sum_{i=1}^{N} p_1(x_i) ln\, \frac{p_1(x_i)}{p_0(x_i)} = -D_{\mathrm{KL}}(p_1 || p_0) \tag{4.3}$$

The Kullback-Leibler divergence is a measure of similarity between two distributions, that is minimal, i.e. it converges to zero when $p_1 = p_0$. Therefore, the probability distribution that maximizes entropy, $p_{\mathrm{ME}}(x)$, can also be defined as the distribution that is as close as possible to the prior distribution, while satisfying experimental constraints.

4.4.2. Bayesian inference. Analogous to the maximum entropy approach, the Bayesian theorem describes an update from a *prior* probability distribution $p(M|I)$ to a *posterior* probability distribution $p(M|D,I)$, as a result of acquiring data D about the system.[104] Bayes' theorem is formulated as follows:

$$p(M|D,I) \propto p(D|M,I)p(M|I) \tag{4.4}$$

I defines some prior information about the system, and can be defined for instance as $I = \{C_j, w_j^0\}_{j=1}^{j=J}$, i.e. as a set of J conformational states C_j with corresponding population fractions w_j^0, obtained from e.g. molecular dynamics simulations. D is a set of K experimental observables, $D = \{D_k\}_{k=1}^{k=K}$, and model M represents set of updated weights, $M = \{w_j\}_{j=1}^{j=J}$, such that $\sum_{j=1}^{J} w_j = 1$ and $w_j \geq 0$.

Bayes theorem is defined in terms of conditional probabilities: $p(M|D,I)$, which represents the *posterior* probability of a model M, given that both D and I are true; $p(M|I)$, which is a *prior* probability of a model given the prior information I, and lastly, $p(D|M,I)$, data likelihood conditioned on the model and prior information. The prior probability, $p(M|I)$, is expressed in terms of relative information entropy $S(w_j||w_j^0)$, or negative Kullback-Leibler divergence (eq. 4.3) between prior and posterior weights:

$$p(M|I) \propto exp\left(\theta S(w_j||w_j^0)\right) = exp\left(-\theta \sum_{j=1}^{J} w_j \, ln\frac{w_j}{w_j^0}\right) \tag{4.5}$$

where θ denotes temperature-like factor that quantifies the confidence in the prior, as will be discussed later. Data likelihood, $p(D|M,I)$, is defined by Gaussian distribution as the sum of squared deviations between observables computed for a model, $x_{\mathrm{m},k}$, and observables derived from an experiment, $x_{\mathrm{e},k}$, weighted by the experimental uncertainty, $\Delta x_{\mathrm{e},k}$:

$$p(D|M,I) \propto exp\left(-\frac{1}{2}\chi^2_{\mathrm{r,total}}\right) \tag{4.6}$$

$$\chi^2_{\mathrm{r,total}} = \frac{1}{K}\sum_{k=1}^{K} exp\left(\frac{x_{\mathrm{e},k} - x_{\mathrm{m},k}}{\Delta x_{\mathrm{e},k}}\right)^2 \tag{4.7}$$

According to Bayes' theorem, finding the optimal model, in this case optimal weights of conformational models, means to maximize posterior probability $p(M|D,I)$:

$$\underset{M}{arg\,max}\, p(M|D,I) = \underset{M}{arg\,max}\, p(D|M,I)p(M|I) \tag{4.8}$$

Now, instead of maximizing posterior probability, we maximize *log*-posterior probability, for several practical reasons. First, *log* is a monotonically increasing function, meaning that it will preserve the properties of the original function, i.e. the same ascending or descending order of probabilities. Extrema of p are the same as extrema of $\log p$. Therefore, maximizing *log*-posterior probability, is the same as maximizing probability itself. Furthermore, having the optimization problem expressed in terms of logarithm is much more convenient, since with logarithmic transformation a lot of mathematical expressions get significantly simplified. For instance, product becomes a sum ($\ln ab = \ln a + \ln b$), which is more numerically stable, and prevents underflow, since probabilities can have very small magnitudes, and their product can often go below the float precision of the computers.

However, optimizers in available statistical packages rather perform minimization of some objective function, instead of searching for its maximum. By having in mind the following equivalence:

$$\underset{M}{arg\,max}\, ln\, p(M|D,I) = \underset{M}{arg\,min}[-\ln p(M|D,I)] \tag{4.9}$$

we can redefine the goal, and say that optimal model (i.e. maximum *a posteriori* distribution, "MAP") can be found by minimizing negative *log*-posterior probability. Bayes' theorem can then be expressed as:

$$MAP = \underset{M}{arg\,min}[-\ln p(M|D,I)] = \underset{M}{arg\,min}[-\ln p(D|M,I) - \ln p(M|I)] \tag{4.10}$$

which, after inserting eq. 4.5 and eq. 4.6, transforms to:

$$MAP = \underset{M}{arg\,min}\left[1/2\,\chi^2_{\mathrm{r,total}} - \theta S\right] = \underset{M}{arg\,min}\, G \tag{4.11}$$

Solution to the optimization problem is unique, since the negative *log*-posterior is convex. Furthermore, the gradient of eq. 4.11 can be easily calculated, which makes the optimization computationally efficient, i.e. it can be performed in seconds on a current laptop, using efficient limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm.

According to the last transformation, finding the maximum *a posteriori* distribution means to minimize the energy-like function *G*, which contains two terms- one that scores the agreement of the optimized model against the experimental data ($\chi^2_{\mathrm{r,total}}$), and an entropy term, *S*, which is a negative Kullback-Leibler divergence, and which informs on the similarity between prior and posterior weights. Contribution of each term in optimization is determined by the temperature-like θ factor. When using large θ value, we express huge confidence in prior, and consequently, entropy term dominates in the optimization over the $\chi^2_{\mathrm{r,total}}$ term. In that case, in order to minimize *G*, one has to find such distribution of weights for which the entropy will be maximized. As discussed previously, entropy is maximized with minimal modification of prior information (eq. 4.3). On the other hand, for small θ, we express small confidence in prior information, and in that case minimization of *G* is dominated by $\chi^2_{\mathrm{r,total}}$ term. In other words, agreement with experimental data is prioritized, which allows large changes in the prior weights. This in turn leads to overfitting.

Therefore, it becomes obvious that prioritizing any of the two sources of information has consequences, and for reliable ensemble reweighting it is necessary to optimally balance different sources of information. Optimal balance of prior information and experimental data was one of the topics in Publication 2 (see **Chapter 6.2** and Supplement B). It was previously proposed that θ value at the corner of $\chi^2_{\mathrm{r,total}}$ versus S, which is often referred to as "L-shaped curve", should be taken for optimal balance of prior information and experimental data. [105] Through rigorous analysis using synthetic data we found that θ value at the corner of "L-shaped curve" is close to the optimal choice of θ. This confirmed that previously arbitrarily proposed corner point is indeed a meaningful choice of θ.

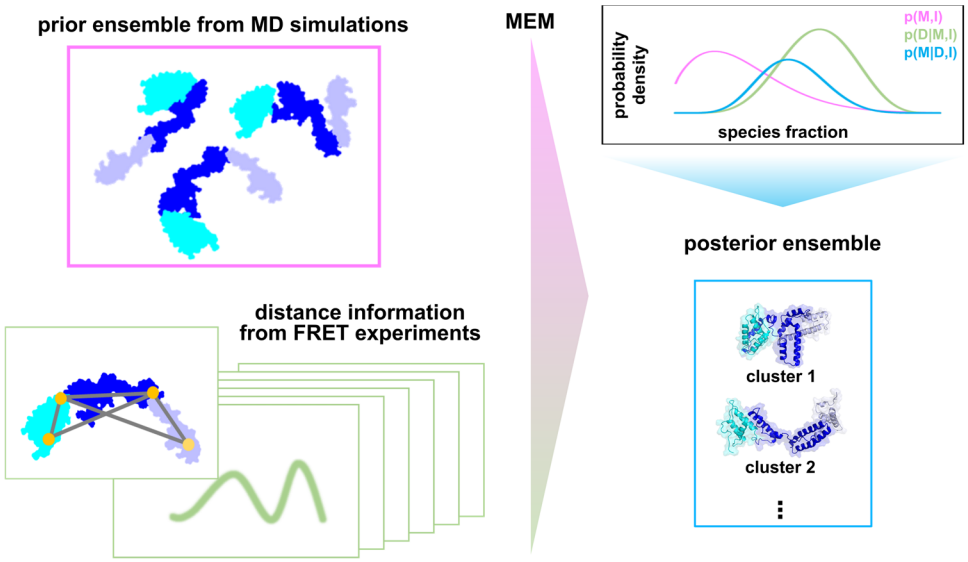


Figure 6 | Ensemble refinement. *To overcome the limitations of using experimental or computational data alone, we gather all available sources of information about the system, for instance structural models with corresponding weights obtained from all-atom MD simulations, and inter-dye distance distributions from FRET experiments, and combine them using some refinement approach, here Maximum Entropy Method (MEM). MEM is the most suitable approach for systems that do not occupy well defined states, and it provides posterior distribution, i.e. set of optimized weights, that are as close as possible to the prior weights, while simultaneously satisfying available experimental data. Figure is adapted from article by Dittrich J., Popara M., et al[45] (see Supplement B).*

4.5. Representing structural ensembles: are atomistic models needed?

Having structural biology for a long time limited to modelling biomolecules in a single, static pose, comparison of structures was typically done in terms of Root-Mean-Square Deviation (RMSD). However, developments in the field of structure determination and integrative modelling, have brought us to a stage where we talk about *ensembles* of structures, and where one-to-one comparison of individual structures is no longer a suitable approach. Moreover, it was found that RMSD is not an adequate quantitative measure of similarity, and whatsoever, it was found that it is the least representative measure of structural similarity.[106] Main reason to that is that RMSD calculation requires superimposition of structures under comparison, which is not a trivial task for structures containing loops, flexible termini or disordered regions. Therefore, two structures that are essentially identical, but differ in position of a single loop, would be classified as different structures, if judged by global RMSD value. These challenges have sparked the discussion on what are suitable ways to visualize and compare ensembles of structures.[107-110]

Within Publication 2 we proposed alternative ways to visualize and compare ensembles of conformations, namely inter-residue distance histograms (distograms) and 3D density maps (see **Chapter 6.2** and Supplement B). These model representations are in no way novel, but we were able to give reasoning why they are well suited and why they should be widely adopted. In the case of distogram representation, two-dimensional matrix of distances between all pairs of C_α atoms is computed, followed by distance histogramming across entire ensemble, by accounting for weights of individual ensemble members. To compare two distograms, we proposed Jensen-Shannon divergence[111], D_{JS}, a quantitative measure of distributions similarity, that converges to zero for identical distributions.[111] It is even more convenient to look at moments of distribution, i.e. mean, variance and skewness, and directly compute absolute deviation between these distribution features for the two ensembles in comparison. In the case of 3D density maps, which is a concept taken from EM field, we map the coordinates of ensemble members onto a grid of *nx*, *ny*, *nz* dimensions and of a specific voxel size. Voxel occupancy is computed as a sum of weights of ensemble members found in respective voxel. As a comparison measure between the density maps, we used the cross-correlation coefficient (CCC). Using synthetic benchmark, we were able to demonstrate that such ensemble representations are robust and efficient way for ensemble visualization and comparison. For details see **Chapter 6.2** and Supplement B. In the distogram representation, we integrate over ensemble members, but residue information is retained. Freely said, one can think of distograms as a "medium-resolution" representation of ensembles. In the case of 3D density maps, we go even further with reduction of the information content, and besides ensemble averaging, residue information is not conserved. Strictly speaking, one can think of 3D density maps as a "low-resolution" ensemble representation.

However, 3D density maps are informative, and there are numerous examples of industrial applications where the volumetric models were successfully used in drug design. One example would be so-called shape screening, where based on a shape of a molecule known to bind to a target, it is possible to identify new binding partners that fit into a given shape constraint.[112] Such simple approach was shown to be very successful, and well suited for early stage drug design (Schrödinger, Inc.). Sadly, there is a long-term practice of attributing negative connotation to these representations, for instance by naming them "blob" or "cartoon-like". We hope that the work presented in this thesis will contribute to removing the stigma from such model representations of ensembles.

Chapter 5

Biological context of studied systems and prior knowledge

5.1. Multi-domain proteins

All three proteins that I studied share the same architecture- they are multidomain proteins, where different modules are connected with a linker of varying level of softness. Multidomain proteins are of immense importance, as they constitute the majority of proteins in both eukaryotes and prokaryotes.[113] Linkers between the domains are found to have huge variability in amino acid propensity, secondary structure, and length.[114] As such, they shape the free-energy landscape of proteins. The function of linkers is beyond being mere connectors of the subunits.[113] They have a role in enabling domains to move from detached to closed form, and classical example for such linker-induced large-scale displacements can be found in U2AF2, the large subunit of U2 auxiliary factor of spliceosome machinery. Furthermore, when the linker acts as a hinge, it can direct correlated movements of subunits. Such an example would be clamp motion upon substrate binding ("Venus-flytrap mechanism") of D-maltodextrin binding protein (MalE). I studied U2AF2 and MalE within Publication 1 (see **Chapter 6.1** and Supplement A). As opposed to those two examples, rigid linkers, which typically exist in a shape of stable α-helices, act as spacers that keep domains at a certain end-to-end distance. This is of importance when different modules in a protein perform separate functions, or when inter-domain cavity serves for substrate docking. Example would be extended helical domain found in lipase specific foldases (Lifs), which maintains mini domains at a specific distance upon binding to lipase. Such spatial arrangement enables docking of lipase into a high contact area, which in turn results in strong binding between Lif and its cognate lipase.[115] I studied conformational landscape of Lif within Publication 2 (see **Chapter 6.2.** and Supplement B).
Linkers, their structure and intrinsic dynamics are not the topic of this thesis, but they rule the energy landscape of the studied proteins and they are the underlying factor of their structural complexity. For an interested reader, comprehensive discussion on the role of linkers in conformational dynamics and allostery can be found in a review by Papaleo *et al*.[114]

5.2. MalE

Membrane transport is one of the crucial processes in the physiology of a cell, which is mediated by a plethora of transport systems. One of these is so called Adenosine Triphosphate Binding Cassette (ABC) family of transporters.[116] ABC transporters exploit the energy coming from Adenosine Triphosphate (ATP) hydrolysis for the transport of the substrates through the membrane. Typical ABC transporter has a four-domain organization, where two domains are highly hydrophobic and are incorporated within the membrane, forming the pathway for the substrate translocation.[116] Remaining two domains are on the cytoplasm side of the membrane, and are responsible for ATP binding. These four domains represent minimal ABC construct, but additional domains can be involved, having a regulatory or substrate binding role, and can be located on either side of membrane.[116] Example of such domain would be D-maltodextrin binding protein (MalE) of *Escherichia coli*, which is periplasmic component of ABC transporter $MalFGK_2$-E, responsible for delivery of maltodextrins to membrane-associated domains of ABC. MalE follows typical structure of periplasmic proteins, i.e. it is monomeric and contains two globular rigid domains of roughly equal size, connected with a hinge which allows conformational change upon substrate binding (**Figure 7A**).[117] Hinge region simultaneously acts as a binding pocket for a substrate.

MalE is a well characterized system, and a classical example of systems used in numerous studies of ligand binding mechanism.[118] We studied MalE as part of Publication 1 (**Chapter 6.1** and Supplement A), with aim to assess the reproducibility of smFRET measurements in detecting slow, ligand induced, conformational change on a sub- second time scale. In particular, we studied reproducibility in detecting population shift between open/inactive (apo, PDB 1OMP) and closed/active (holo, PDB 1ANF) form of MalE upon maltose binding (**Figure 7A**). Holo measurements were done at the substrate concentration of 1 mM, with K_d varying between 1 and 2 μM (for details see Supplement A). Sensing of conformational change was done through smFRET measurements of stochastically labelled double-cysteine mutants at the positions K29C-S352C (MalE-1), D87C-A186C (MalE-2), A134-A186C (MalE-3), K34C-N205C (MalE-4) and T36C-N205C (MalE-5), using multiple dye pair combinations (**Figure 7A**). For further details see **Chapter 6.1** and Supplement A.

5.3. U2AF2

U2 Auxiliary Factor (U2AF) is part of spliceosome machinery, responsible for the transformation of precursor mRNA into mature mRNA. Splicing of pre-mRNA represents

removal of non-coding sequence (intron), which is followed by ligation of coding material (exons).[119] Spliceosome machinery consists of several small nuclear ribonucleoproteins (snRNP) and auxiliary factors, that in coordinating manner and through several stages bind and remove intron.[120] After the formation of mature mRNA, intron is released from the spliceosome, which is then ready for the next splicing activity.

U2AF is a heterodimer consisted of larger (65 kDa) subunit, U2AF65 (also known as U2AF2), and a smaller (35 kDa) subunit, U2AF35 (also known as U2AF1). U2AF acts in the early stage of pre-mRNA splicing, by targeting 3'- splice site of the intron. Specifically, U2AF2 subunit binds to single-stranded polypyrimidine RNA sequence (Py-tract) while U2AF1 subunit binds to 'AG' dinucleotide at the 3'- splice end of intron.

Binding of U2AF2 to Py-tract of the intron is the crucial act of U2AF in the splicing process. U2AF2 subunit has following building blocks: RS - arginine and serine rich domain, ULM-peptide motif responsible for binding of U2AF2 and U2AF1 into heterodimer[121], RNA recognition motif domains, RRM1 and RRM2, that bind to polypyrimidine sequence[122] of pre-mRNA, and lastly, C-terminal U2AF homology motif (UHM) domain, that has a role in interaction between U2AF and splicing factor 1 (SF1) of spliceosome[123]. Interaction between U2AF1 and U2AF2, is proven to have affinity-increasing effect in the recognition of weak Py-tracts.[124]

Subject of Publication 1 (**Chapter 6.1** and Supplement A) was truncated version of U2AF2, that contains only tandem RRM1,2 domains (residues 148-342), connected via flexible linker, since these are responsible for the recognition of Py-tract. For simplicity, we refer to this minimalistic construct as U2AF2. These two domains follow a typical architecture of RRMs, with two alpha helices intertwined with four antiparallel beta sheets: βαββαβ.[125] Despite having canonical topology of its building blocks, U2AF2 exists in the solution as a continuum of dynamic diverse conformations, as a result of flexible linkage between RRMs.[126]

U2AF2 has a strong affinity towards uridine rich sequences.[127] Therefore, in this work we studied the population shift between inactive (apo, PDB 2YH0) and active (holo, PDB 2YH1) form of U2AF2 upon binding to U9 RNA (**Figure 7B**), at the concentration of 5 μM, corresponding to K_d of roughly 1.2 μM. Since RRM2 has exposed binding site in the apo state, and RRM1 not, binding of shorter sequences, e.g. U4 RNA, would occur only on RRM2, without causing inter-domain rearrangement.[128] Hence, we chose longer Py-tract (U9) in order to trigger the rearrangement of RRM1 and RRM2 relative to one another. While binding of Py-tract to U2AF2 is well understood[122], characterization of dynamic repositioning of RRM1 and RRM2 relative to one another with corresponding kinetic scheme is lacking.

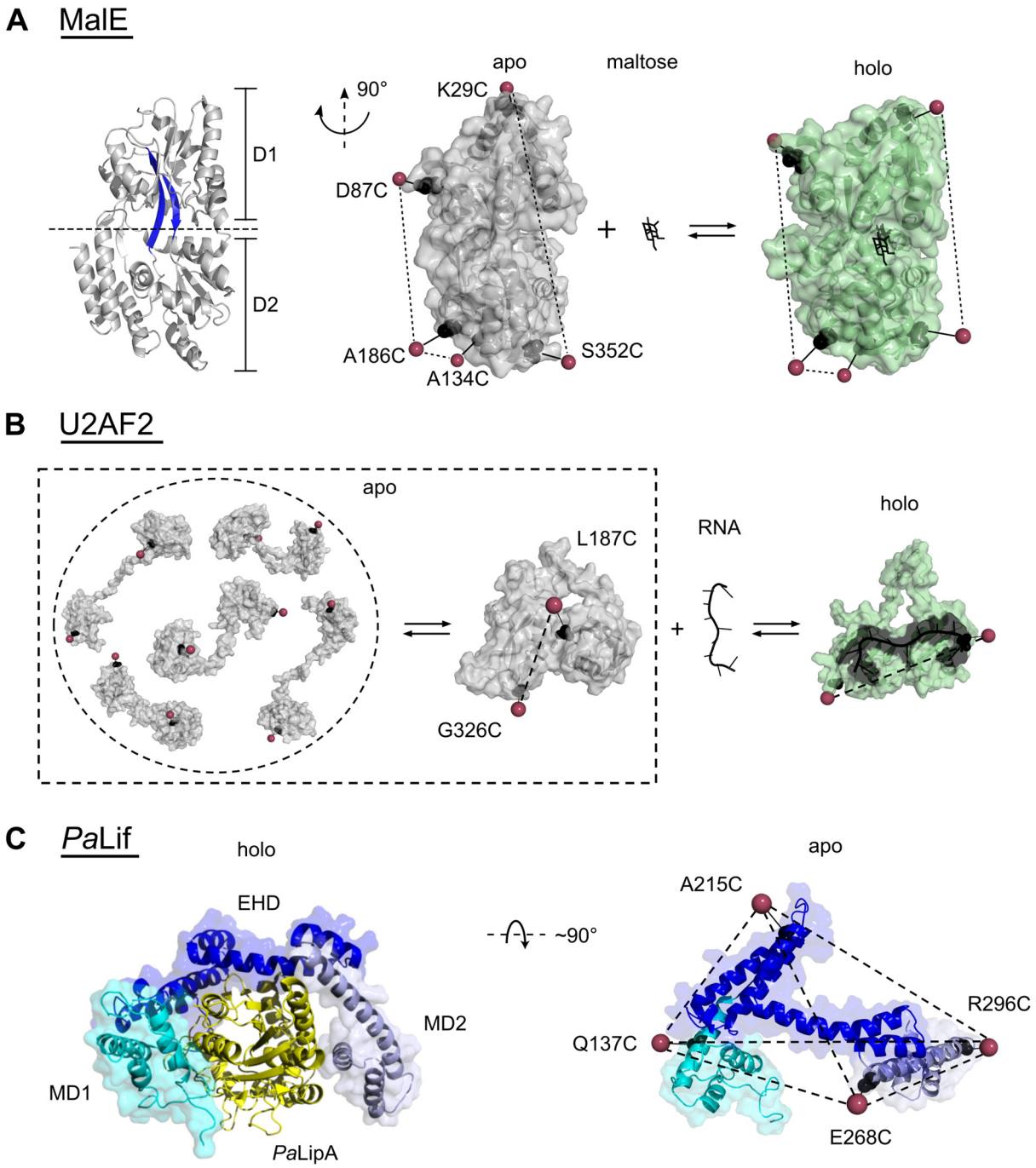


Figure 7 | Overview of studied biomolecules. ***(A)*** *Left: Crystal structure of MalE in its ligand-free apo state (PDB ID 1OMP) showing structural organization typical for periplasmic proteins, where two rigid domains D1 and D2 are linked with a hinge (highlighted in blue). Right: Conformational change between open/apo (gray, PDB ID 1OMP) and closed/holo (green, PDB ID 1ANF) form of MalE, induced with maltose binding, where hinge region acts as a binding pocket for a substrate.* ***(B)*** *Conformational ensemble of U2AF2. Apo state (gray) fluctuates between an ensemble of detached conformations[126], where for simplicity only five representative structures are displayed, and a compact conformation (PDB ID 2YHO). Binding of U9 Py-tract ligand triggers rearrangement of RRM domains relative to one another, leading to stabilization of open/holo state (green, PDB ID 2YH1).* ***(C)*** *Left: Homology model of catalytic folding domain of Lif in complex with its cognate lipase, LipA (yellow), from Pseudomonas Aeruginosa species. CFD of PaLif has headphone-like shape, and is consisted of mini domain*

1 (MD1, cyan), extended helical domain (EHD, blue) and mini domain 2 (MD2, light blue). In all three panels, cysteine mutations are depicted as black spheres and mean dye positions determined from AV simulations are shown as red spheres. Double cysteine variants, for which inter-dye distances were measured experimentally, are indicated with black dashed lines. Panels ***(A-B)*** *are adapted from Agam, G., Gebhardt, C., Popara M., et al[129] (see Supplement A) and panel* ***(C)*** *is adapted from Dittrich J., Popara M., et al[45] (see Supplement B).*

Here we employed a range of fluorescence-based methods to unravel the RRM1,2 domain arrangement as well the kinetic scheme of a dynamic exchange within a heterogeneous ensemble of U2AF2. Furthermore, U2AF2 served as more complex system to evaluate the reproducibility of smFRET measurements in detection of fast conformational dynamics on a millisecond and sub-millisecond time scale. Sensing of U2AF2 ensemble was performed using multiple dye pair combinations, stochastically labelled at the positions L187C of RRM1 and G326C of RRM2 (**Figure 7B**), which, as previously shown, do not affect structural integrity of U2AF2, nor its binding affinity towards U9 ligand.[130] For details see **Chapter 6.1** and Supplement A.

5.4. Lif

Lipase specific foldases (Lifs) are an inner-membrane bound steric chaperones, that in the periplasm of Gram-negative bacteria activate their cognate lipase.[131] Extracellular microbial lipases are the most important class of enzymes with vast number of applications in biotechnology, from detergents, over food to paper industry.[132, 133] It is therefore of great interest to understand structure and dynamics of their chaperones. The whole understanding of the role of Lif chaperones is vague, and historically they were named "foldase"[132] from the understanding that Lifs interact with unfolded lipase, and assist their folding into enzymatically active state. However, circular dichroism studies showed that secondary structure of enzymatically active lipase and lipase in absence of Lif is practically indistinguishable.[134] On the other hand, sum of CD spectra of lipase and Lif, measured separately, does not match the CD spectrum of a solution containing both Lif and lipase. On those grounds, it is concluded that it is rather Lif that undergoes structural change upon lipase binding. However, without Lif, lipase is not enzymatically active. The importance of Lif in the activity of lipase is furthermore highlighted through the fact that they are encoded together in a single operon. Their interaction is highly specific, as it was found that Lif does not activate lipase originating from different species.[135]

*Bg*Lif in complex with lipase *Bg*LipA, originating from bacterial species *Burkholderia glumae*, represent the only foldase-lipase complex for which the crystal structure has been reported (PDB 2ES4).[115] This structure reveals high contact area between Lif and its cognate lipase, which is in agreement with nM K_d value.[115, 136] Focus of this study was *Pa*Lif, responsible for the folding of its cognate lipase *Pa*LipA, originating from the *Pseudomonas aeruginosa* species (**Figure 7C**). There is 52% sequence similarity and 39% sequence identity between *P. aeruginosa* and *B. glumae* Lifs.[136] Overall, level of sequence identity between Lifs originating from different species is rather low, and only a sequence of eight amino acids is found to be conserved for all members of Lif family known by now: RXXFDY(F/C)L(S/T)A, with X being any residue.[137] This conserved sequence contains amino acids which were shown to be essential, but not sufficient for the activation of LipA.[138]

*Pa*Lif has five domains, namely: TMD- transmembrane domain that anchors Lif to the inner membrane, VLD- unstructured variable linker domain, that most likely has a role as a spacer that keeps catalytically active part of Lif far enough from inner membrane for the full accessibility to LipA, which is lastly followed by CFD- catalytic folding domain, consisted of mini domain 1 (MD1, α1-α3), extended helical domain (EHD, α4-α8) and mini domain 2 (MD2, α9-α11). Here we studied truncated *Pa*Lif, containing only catalytic folding domain (MD1-EHD-MD2, residues 66-340) (see **Figure 7C**). Conserved sequence is found to be located in the α1 helix of MD1[115], however, MD1 alone is not sufficient for activation of *Pa*LipA.[136]

So far, high-resolution structure of *Pa*Lif has not been reported. The only structural insight available is the NMR-derived structure of MD1, that is found to form a stable three-helical bundle, preceded with unstructured N-terminal tail, similar to MD1 from *Bg*Lif.[136] To unravel the conformational ensemble of apo Lif, we applied Maximum Entropy Method (MEM) on prior conformational models obtained from all-atom MD simulations. As starting structure in MD simulations we used homology model constructed based on crystal structure of *Bg*Lif:*Bg*LipA complex. We refined prior conformational ensembles using time-resolved ensemble-TCSPC data, acquired for a network of 6 FRET pairs (for details see **Chapter 6.2** and Supplement B).

Chapter 6

Overview of published articles and manuscripts

In this chapter, I give the overview of published articles and manuscripts. Presented articles demonstrate various methodologies from the field of single-molecule FRET studies, computer science and integrative structural modelling, for which I presented the basic concepts in the previous chapters. In the **Chapter 6.1** I present the results of a blind international study (Publication 1), whose aim was to assess the precision and accuracy of single-molecule FRET measurements, across different instruments, analysis procedures and when using challenging protein systems. In the **Chapter 6.2**, I give an overview of Publication 2, in which we embarked on optimizing the conformational ensemble of a flexible and disordered system. Ensemble optimization was done using maximum entropy method, by combining FRET data with conformational ensembles obtained from MD simulations. Lastly, within the Manuscript 1 I created the protocol for robust estimation of correction parameters necessary for accurate determination of FRET observables, and which I accompanied with guidelines for alignment and calibration of diffusion-based PIE-MFD confocal setup. Results of this study are summarized in **Chapter 6.3**.

6.1. Publication 1: Reliability and accuracy of single-molecule FRET studies for characterization of structural dynamics and distances in proteins

This section is based on Agam, G., Gebhardt, C., Popara M., *et al*[129] (Supplement A).

This community paper was an international blind study involving 19 laboratories world-wide, and it was a follow-up of a previous study by Hellenkamp *et al*[29], in which double-stranded DNA strands were used to evaluate the precision and accuracy of single-molecule FRET measurements. We took the next step, and using challenging protein systems we confirmed that smFRET measurements are reliable and reproducible across different instrumentations and analysis tools, even when dealing with complex systems such as proteins.

Proteins display large structural heterogeneity, they are dynamic across different spatial and temporal scales, local environments and dye mobility vary a lot within proteins, and lastly, they are more challenging to handle, as they often demonstrate substantial sensitivity to experimental conditions, such as buffer composition, storage and measurement temperature, pH etc. Despite

all presented challenges, we were able to obtain the sub-nanometer precision and accuracy as in a previous study by Hellenkamp *et al*.[29]

How consistent are FRET efficiency histograms and distances determined by different laboratories?

In the first round of blind study, as a model system we used D-maltodextrin binding protein of *Escherichia coli*, the periplasmic binding component of the ATP binding cassette, here referred to as MalE (for further information see **Chapter 5.2**). MalE has a typical architecture of the periplasmic binding proteins, i.e. it contains two roughly globular domains, or approximately equal size, connected with a hinge, which allows conformational change from opened (apo) to closed (holo) state, upon substrate binding (**Figure 8a-b**). We monitored conformational change of MalE upon binding to maltose at the concentration of 1 mM, and for that we designed three double-cysteine variants K29C-S352C, D87C-A186C and A134-A186C, to which we refer as MalE-1, MalE-2 and MalE-3. These three variants display an increase, decrease and no change in FRET efficiency upon maltose binding. Later on, additional two mutants were added, namely K34C-N205C (MalE-4) and T36C-N205C (MalE-5), which are analogous to MalE-1, but show larger FRET efficiency contrast between apo and holo state. The latter two MalE variants were studied only by a subset of laboratories. MalE variants were stochastically labelled by Alexa Fluor 546 (later referred to as Alexa546) as a donor dye and Alexa Fluor 647 (later referred to as Alexa647) as an acceptor dye. Laboratories were asked to provide fully corrected average FRET efficiencies for all three variants MalE1-3 in both apo and holo conditions. All laboratories properly observed expected FRET efficiency change, namely increase (MalE-1), decrease (MalE-2) and no change (MalE-3) in FRET efficiency upon ligand binding (**Figure 8c**). We observed excellent consistency between the $\langle E \rangle$ values reported by different laboratories, with uncertainty across all MalE variants ≤ 0.06, which in terms of inter-dye distances corresponds to the precision of ≤ 2 Å, and accuracy of ≤ 5 Å, when comparing experimentally derived distances to the structure-based derived distances. Largest uncertainty was obtained for MalE-1, and smallest for MalE-3, which interestingly also has the highest FRET efficiency of ~ 0.9.

Despite good agreement between the laboratories, we did observe, consistently across all mutants, systematic deviations of some laboratories from the lab-averaged $\langle E \rangle$ values. Therefore, we looked at the difference between apo and holo average FRET efficiency, i.e. $\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle$ (**Figure 8d**). As expected, distribution of FRET efficiency differences narrowed down two- to three-fold, which proved that observed deviations were indeed

systematic and not random. We suspected that these systematic deviations originate from the error in determination of correction parameters.

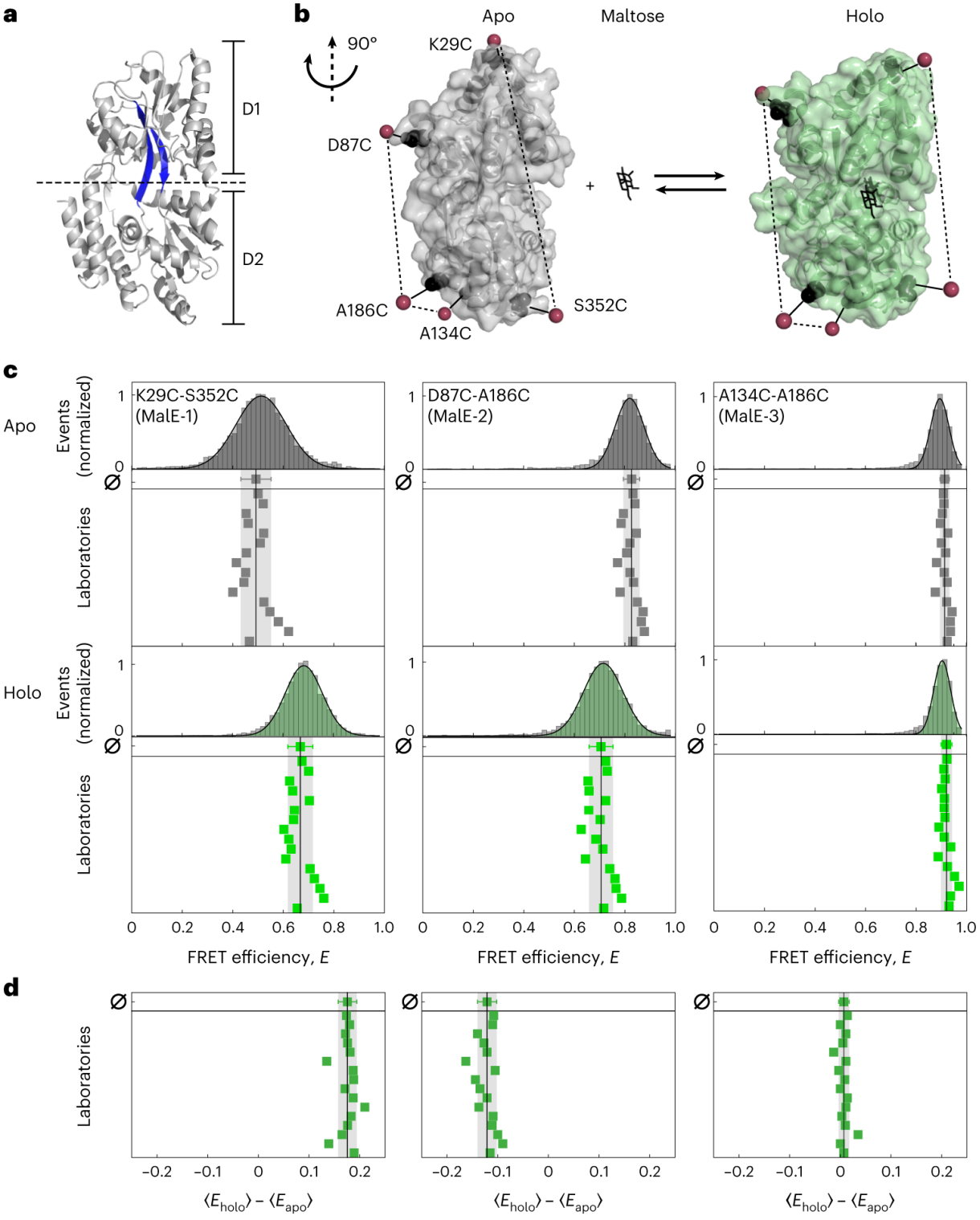


Figure 8 | Experimental design of MalE as a protein model system for smFRET studies. *(**a**) Crystal structure of MalE in its ligand-free apo state (PDB ID 1OMP) with domains D1 and D2 linked by flexible beta sheets (highlighted in blue). (**b**) The crystal structure of MalE (rotated by 90° as compared to **(a)** in the apo (gray, PDB ID 1OMP) and holo (green, PDB ID 1ANF)*

states with mutations at K29C-S352C (MalE-1), D87C-A186C (MalE-2) and A134C-A186C (MalE-3) indicated in black. Note, each mutant only contains one cysteine pair and was measured using the Alexa546–Alexa647 FRET pair. The estimated mean position of the fluorophores from AV calculations are shown as red spheres. ***(c)*** *FRET efficiency E histograms for three MalE mutants, MalE-1 (left), MalE-2 (middle) and MalE-3 (right), in the absence and presence of 1 mM maltose (bottom, green) for one exemplary dataset measured in laboratory 1. The distribution is fitted to a Gaussian distribution. The reported mean FRET efficiencies for 16 laboratories are shown below (due to experimental difficulties, the results of three laboratories were excluded). The mean FRET efficiency and the standard deviation of all 16 laboratories are given by the black line and gray area.* ***(d)*** *Individual FRET efficiency differences for each laboratory, between the apo and holo states,* $\langle E_{holo} \rangle - \langle E_{apo} \rangle$*, for MalE-1 (left), MalE-2 (middle) and MalE-3 (right). The mean FRET efficiency difference and the standard deviation of all 16 laboratories are given by the black line and gray area. Figure and figure caption are taken from Agam, G., Gebhardt, C., Popara M., et al*[129] *(Supplement A).*

Therefore, we propagated the uncertainty in FRET efficiency based on the error in correction parameters (**Figure 9A**). Propagated uncertainties due to the error in background, crosstalk, α, and direct excitation, δ, all show monotonic decrease or increase against FRET efficiency. Only the uncertainty ΔE_γ due to the error in detection efficiency parameter, γ, shows parabolic dependence:

$$\Delta E_\gamma = E(1-E)\frac{\Delta\gamma}{\gamma} \qquad (6.1.1)$$

We found that experimentally observed uncertainties in FRET efficiency all lie on a parabolic line that corresponds to relative error in γ parameter of 23% (**Figure 9B**). Therefore, error due to the γ parameter alone was sufficient to explain discrepancies between different laboratories.

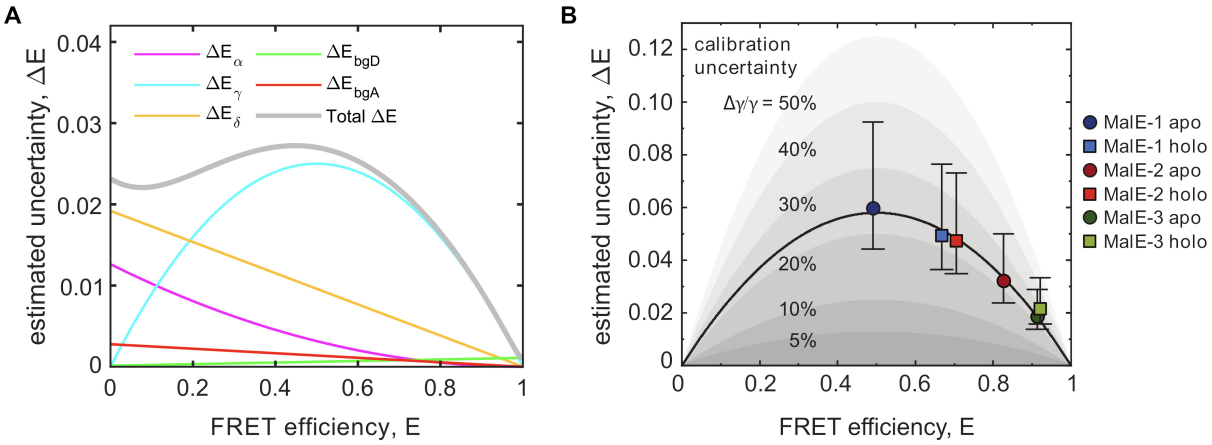


Figure 9 | Contributions of the errors in correction factors to the uncertainty in FRET efficiency. ***(A)*** *Propagated uncertainties in FRET efficiency due to donor crosstalk,* ΔE_α*, acceptor direct excitation,* ΔE_δ*, detection efficiency,* ΔE_γ*, and due to background in donor and acceptor channels,* ΔE_{bgD} *and* ΔE_{bgA}*. Total uncertainty,* ΔE*, is given as grey line.* ***(B)*** *Standard*

deviation of the reported FRET efficiencies against the average FRET efficiency for MalE1-3 variants, revealing that variability between lab-reported values arises solely due to the error in γ parameter, since all points lie on parabolic black line that corresponds to relative error in γ of 23%. Adapted from Agam, G., Gebhardt, C., Popara M., et al[129] (see Supplement A).

Second system that we studied was part of the U2 auxiliary factor (U2AF) of the spliceosome machinery. Specifically, we studied a minimalistic construct consisted of two RNA recognition motifs 1 and 2 (RRM1,2), connected with a flexible linker (see **Chapter 5.3** and **Figure 7B**). We refer to this construct as U2AF2. In the spliceosome activity U2AF2 binds to the polypyrimidine tract of the precursor mRNA, and therefore in our study we monitored RRM1,2 inter-domain rearrangement upon binding to U9 single-stranded RNA sequence, at the concentration of 5 μM. For such minimalistic construct we designed a single double-cysteine variant, L187C-G326C, for which the labelling positions are on the two RRM domains, allowing us to sense the inter-domain arrangement (**Figure 7B**). Double cysteine variant was labelled with Atto 532 as a donor dye and Atto 643 as an acceptor dye. Similarly as for MalE, we found a good agreement between average FRET efficiencies reported by different laboratories. However, especially in the case of this protein we expected that the systematic errors, due to the user bias in determination of correction parameters, would be prominent. Due to the smaller number of measurements that cover narrower range of FRET efficiencies, γ parameter could not be as robustly determined as for MalE. Therefore, to evaluate the extent of the user bias in data correction, single person reanalyzed measurements from all laboratories with optimized calibration procedure, which improved precision to a standard deviation of 0.008 from previously obtained 0.03. This again demonstrated that γ parameter is indeed the limiting factor and main cause of the observed discrepancies. Therefore, improved calibration protocols are necessary in order to further push the limits of FRET measurements.

How reliably can FRET detect dynamics across different time-scales?

In the next step we wanted to address the ability of FRET measurements to unambiguously detect and quantify dynamics across different time-scales. Burst variance analysis (BVA) and E-τ plots are frequently used tools to detect FRET dynamics. In both approaches, presence of *E*-fluctuations is visualized as a displacement of double-labelled FRET population from the static FRET-line. This displacement can be quantified using so-called apparent dynamic shift, ds, which is a minimal distance of the center of the FRET population to the static FRET-line.[22, 23] Both BVA and E-τ plots confirmed that MalE in apo and holo state does not exhibit large scale structural changes on the time-scale of < 10 ms, since in that case two distinct populations

should be visible. However, when looking at the apparent dynamic shift values, we found that in particular MalE-1 variant exceeds the reference value of a static system (dsDNA), but also the upper limit that is expected for the large-scale apo-holo transition:

$$ds_{\max} = \frac{1}{\sqrt{2}}\left(\sqrt{1 - E_{\text{apo}}} - \sqrt{1 - E_{\text{holo}}}\right)^2 \tag{6.1.2}$$

Due to that, some laboratories declared that they suspect MalE to be dynamic on sub-*ms* time-scale. In the case of U2AF2 apo, all groups unambiguously detected dynamics, since it displays prominent displacement of the double-labelled population from the static FRET-line. In the holo state, apparent dynamic shift was not so prominent compared to the static reference. Additionally, in the holo measurement around 15% of the molecules remained in the apo state at the RNA concentration used, which overlaps in the E-histogram with the holo state of the holo measurement. Thus, it was difficult to judge if U2AF2 holo was indeed dynamic or static.

What are the minimal structural fluctuations detectable?

MalE-1 variant showed not only the notably higher ds values compared to what is expected for static reference and large-scale apo-holo transition, but it also had the largest deviation of inter-dye distances from structure-based predicted values (**Figure 10d**). To understand the origins of these deviations, we studied dye behavior using fluorescence lifetime and anisotropy. We found that position S352 of MalE-1 variant displays prominent sticking for both donor and acceptor dye ($r_\infty > 0.25$), while at the position K29 of MalE-1 only donor dye largely interacts with the surface ($r_{\infty,\text{D}} > 0.3$), but acceptor has unrestrained mobility ($r_{\infty,\text{A}} \approx 0.12$) (**Figure 10b**). Accounting for position-specific sticking in the forward model helped improve the agreement between predicted and measured distances (**Figure 10a-d**). Such position-specific dye-protein interactions cause *E*-fluctuations, which manifest themselves as displacement of double-labelled population from the static FRET-line. This in turn, might be falsely interpreted as presence of conformational dynamics. Therefore, we studied additional dye combinations, for all MalE variants in both apo and holo conditions. We found that the dye pair Alexa546-Alexa647, originally measured by laboratories, has the highest combined residual anisotropy $r_{\text{c},\infty} = \sqrt{r_{\infty,\text{A}} r_{\infty,\text{D}}}$. We were able to understand that this comes due to the donor dye, because when the acceptor is changed and donor is the same, $r_{\text{c},\infty}$ remains high, but it reduces once the donor is exchanged (**Figure 10e, top**). Next, we looked into the relation between observed

apparent dynamic shift and combined residual anisotropy, and found strong correlation between the two, over all dye pairs (Pearson's r = 0.73), which even further increased when looking at individual dye-combinations (**Figure 10f**). This was indication that observed ds is caused by dye sticking interactions. To remove the data sets with prominent dye sticking, we needed to establish a robust threshold for the combined residual anisotropy, instead of using previously used rule-of-the thumb value of $r_{c,\infty}$ = 0.2.[29] To do so, we propagated uncertainty in distance, $\Delta R_{\mathrm{app}}(\kappa^2)$, based on the orientational factor κ^2 using "diffusion with traps" model[139], that takes as an input experimentally determined residual anisotropies of donor and acceptor (**Figure 10e, bottom**). We proposed a realistic threshold of 10% in distance uncertainty, and we were able to find that this corresponds to the threshold of 0.25 in terms of combined residual anisotropy, $r_{c,\infty}$. Interestingly, this is close to previously arbitrarily used threshold of 0.2. Using the established threshold, we were able to filter out the data sets with spurious dye behavior, and we performed the *P*-test analysis of remaining data sets where dye mobility is unrestrained. *P*-test revealed that, even then, ds value is significant for some of the MalE variants. This indicated that observed ds values, were perhaps caused by a combination of dye sticking and small-scale structural fluctuations. Through simple assumptions, we were able to convert experimentally observed ds values into the amplitude of structural fluctuations δR (**Figure 10g**). We found that the obtained δR exceeds the amplitude of structural fluctuations of a static system (dsDNA), by 2-3 Å for MalE-1 and its analogous variants MalE-4 and MalE-5. This was an ultimate proof that indeed small-scale structural fluctuations take place, and we were able to establish a current lower limit for the detection of structural dynamics of ≤ 5 Å.

Beyond lab comparison

In the article we went beyond lab comparison, and characterized in detail the complex structural ensemble of U2AF2 (see **Figure 7B**), using a combination of approaches, namely filtered FCS for the dynamics on microsecond time-scale, and dynamic-PDA for dynamics on millisecond time-scale. We were able to model the apo state of U2AF2 as an exchange (~ 10 ms) between the ensemble of detached conformations and a compact state. The detached ensemble is itself highly dynamic, with interconversion between the detached states happening on the order of ~ 9 μs. Then, upon ligand binding (> 100 ms), opened state is being stabilized. Lastly, we also collected all setup-related data from participating laboratories, and were able to identify the advantages and disadvantages of different experimental designs, opening a discussion regarding optimal measurement conditions.

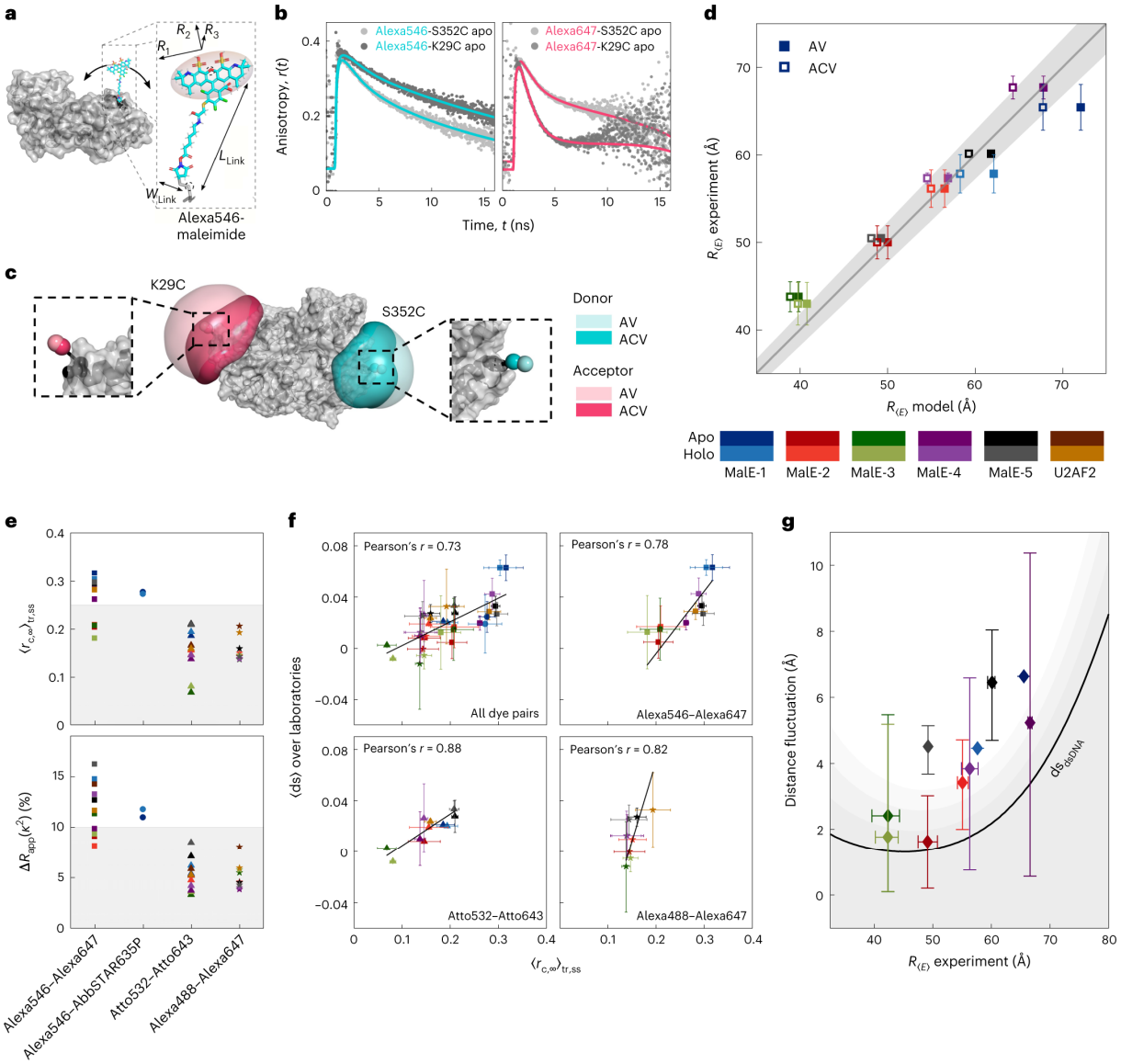


Figure 10 | Assessing the accuracy of smFRET-derived distances in MalE. ***(a–d)*** *AV calculations and model-based interdye distances.* ***(a)*** *Schematic of Alexa546 attached to MalE (PDB 1OMP) showing the parameters needed for the AV calculations using the AV3 model.* ***(b)*** *Fluorescence anisotropy decays of single-cysteine mutants for the donor (Alexa546, left) and acceptor (Alexa647, right) at the labeling positions K29C and S352C. Solid lines represent fits to a model with two or three rotational components.* ***(c)*** *AV (light color) and ACV (dark color) calculations for Alexa546 (cyan) and Alexa647 (pink) at labeling positions 352 and 29. The zoom-ins show the mean positions of the dyes based on the AV (light shade) and ACV (darker shade) models.* ***(d)*** *Comparison of the experimentally obtained FRET-averaged distance $R_{\langle E \rangle}$ with the theoretical model distances using the AV (filled squares) and ACV (empty squares) calculations. Errors represent the standard deviation in experimental distances (n = 16 laboratories for MalE mutants 1–3, n = 2 laboratories for MalE mutants 4–5, n = 7 laboratories for U2AF2). The solid line represents a 1:1 relation and the gray area indicates an uncertainty of ± 3 Å for a Förster radius of R_0 = 65 Å. MalE-4 and -5 were measured by two laboratories.* ***(e)*** *Detection of dye-specific protein interactions. Top shows the five MalE mutants and U2AF2 labelled with different dye combinations to determine the donor–acceptor-combined residual anisotropy, $\langle r_{c,\infty} \rangle_{tr,ss}$ (n = 3 laboratories). Bottom shows the distance uncertainty relating to κ^2, $\Delta R_{app}(\kappa^2)$, estimated from residual anisotropy. A maximum allowed distance uncertainty of ≤ 10% (shaded gray region) in $\Delta R_{app}(\kappa^2)$ leads to a dye-independent*

*threshold of 0.25 for $\langle r_{c,\infty} \rangle$. **(f)** Lab-averaged apparent dynamic shift $\langle ds \rangle$ versus the combined residual anisotropy $\langle r_{c,\infty} \rangle$ is shown for all measured dye pairs (top left) and individually. Error bars of the apparent ds represent the standard deviation over n = 3 laboratories. For the combined residual anisotropy, the propagated 1σ uncertainty. **(g)** The structural flexibility of MalE estimated after filtering using the distance uncertainty threshold shown in **(e)**. Error bars represent the 1σ percentiles averaged over all dye pairs (n = 1, MalE-1; n = 7, MalE-2 and MalE-3; n = 4, MalE-4 and n = 5, MalE-5). The residual distance fluctuations obtained from control measurements on dsDNA in one laboratory ($ds_{dsDNA} = 0.0026 \pm 0.0044$) are shown as a black line (gray areas represent confidence intervals of 1σ, 2σ and 3σ). Figure and figure caption are taken from Agam, G., Gebhardt, C., Popara M., et al[129] (Supplement A)*

Conclusions and significance

This article was a community-driven effort that verified smFRET measurements as mature and reliable technique for quantitative structural biology. Despite the use of various home-built instrumentations and software, we were able to demonstrate that with smFRET one can measure distances with sub-nanometer precision and accuracy, even when using challenging protein systems. Main outcomes of this community paper are:

- Across two protein systems, with different level of structural complexity, and which display dynamics across different spatial and temporal scales, we were able to measure FRET efficiencies with uncertainty of ≤ 0.06, which in terms of inter-dye distances corresponds to precision of ≤ 2 Å and accuracy of ≤ 5 Å.
- We showed that discrepancies between the lab-reported values could be fully explained by the error in γ parameter, and that optimized procedures for the determination of correction parameters are necessary to further push the limits of FRET experiments.
- We demonstrated that FRET can unambiguously detect dynamics across different timescales, and is able to distinguish true conformational dynamics from dye-artifacts.
- The article encourages to perform dye controls and/or additional variant measurements, before drawing conclusions on the presence of dynamics, and provides quantitative guidelines on how to recognize measurements with spurious dye behavior.
- We showed that in smFRET experiment one can detect small-scale fluctuations on the order of ≤ 5 Å, and that detectability of dynamics depends on myriad of factors, such as the FRET efficiency contrast between the interconverting states, kinetic averaging, photon statistics of the experimental data and systematic errors in the determination of correction parameters.

6.2. Publication 2: Resolution of Maximum Entropy Method-derived posterior conformational ensembles of a flexible system probed by FRET and molecular dynamics simulations

This section is based on the article by Dittrich J., Popara M., *et al* (Supplement B).

In this project we embarked on optimizing the conformational ensemble of lipase-specific foldase from *Pseudomonas aeruginosa* species, i.e. *Pa*Lif, using the Maximum Entropy Method (MEM). For fundamentals of ensemble optimization see **Chapter 4.4** and for biological context of *Pa*Lif see **Chapter 5.4**.

As prior information we used conformational ensembles obtained from all-atom MD simulations, and as experimental data we used inter-dye distance distributions obtained from time-resolved ensemble-TCSPC measurements. As there is no existing structural model for the *Pa*Lif, as a starting structure for MD simulations we used homology model generated using the only existing crystal structure of lipase-specific foldase, *Bg*Lif, in complex with its cognate lipase, i.e. *Bg*Lif:*Bg*LipaA.[115, 136] The obtained homology model of *Pa*Lif (see **Figure 7C**) shares the same headphone-like shape of its catalytic folding domain with the *Bg*Lif. Three domain organization can be recognized, where two mini-domains (MD1 and MD2) are connected with extended helical domain (EHD). Helical organization of EHD implies that it most likely plays a role of a rigid linker that keeps mini domains at a fixed distance, allowing the docking of LipA. Knowing from a previous NMR study that MD1 is structurally stable[136], and assuming that MD2 is predominantly stable as well, we approximated Lif as a three rigid body object, with six degrees of freedom. With such approximation, smaller number of FRET pairs was assumed to be sufficient to resolve inter-domain arrangement of Lif. We designed 6 FRET pairs, sampling both inter- and intra-domain distances (see **Figure 7C**). Fluorescence decays of each FRET pair $D_k \in \{D_k\}_{k=1}^{k=K}$, were analyzed through non-parametric Bayesian framework (eq. 2.35 and eq. 2.38) to obtain inter-dye distance distribution $x_{\mathrm{e}}(R_{\mathrm{DA}})_k$ and corresponding error $\Delta x_{\mathrm{e}}(R_{\mathrm{DA}})_k$ (for details see **Chapter 2.5**).

As prior information, we used conformational ensembles obtained from all-atom MD simulations, using distinct force fields, namely FF14SB[140], FF19SB[141] and FF99SB[142], which are tailored to ordered proteins, followed by FF14IDPS[143] (here referred to as IDPSFF) tailored to disordered proteins and FF99SBdisp[51] parameterized to describe the whole spectrum of proteins from folded to disordered. Prior to ensemble reweighting, trajectories were clustered using agglomerative hierarchical clustering method (see **Chapter 3.2**). Each obtained cluster is

represented with the respective centroid structure C_j and population fraction (weight) ω_j^0. Centroid of a cluster is a structure with smallest cumulative distance to all members in a cluster.

Ensemble optimization

To combine prior information $I = \{C_j, \omega_j^0\}_{j=1}^{j=J}$ with experimental data $D = \{x_e(R_{DA})_k, \Delta x_e(R_{DA})_k\}_{k=1}^{k=K}$, experimental observables must be first computed for a prior, using a forward model. As forward model we used AV simulations[144, 145] that provide center-to-center distance R_{mp} between the AV clouds of donor and acceptor for each of the conformational models j and for each of the k FRET variants. R_{mp} distances are then converted to inter-dye distance distributions $x_m(R_{DA})_{jk}$ using non-central χ-distribution as described previously (see eq. 2.35). Ensemble-averaged inter-dye distance distribution $x_m(R_{DA})_k$ is then computed as weighted average over all ensemble members:

$$x_m(R_{DA})_k = \sum_{j=1}^{J} \omega_j x_m(R_{DA})_{jk} \tag{6.2.1}$$

In maximum entropy method, we minimize energy-like function G (see eq. 4.11), by optimizing the weights of structural models $\{\omega_j\}_{j=1}^{j=J}$, and by varying the temperature-like parameter θ, which determines the contribution of two sources of information. Prioritizing any source of information has its consequences, and for reliable MEM reconstructions, optimal balance between experimental data and prior information is necessary. Results of MEM optimization are typically given as $\chi^2_{r,total}$ versus entropy S, which is often referred to as "L-shaped curve". Entropy S tells about the level of perturbation of prior weights (see eq. 4.3 and eq. 4.5), and $\chi^2_{r,total}$ quantifies the agreement between the distributions computed for model $x_m(R_{DA})_k$ and experimental data $x_e(R_{DA})_k$, discretized in N_b bins. Due to the pronounced sensitivity of FRET measurements in the proximity of Förster radius, i.e. in the range of $0.65R_0$-$1.5R_0$, which for a dye-pair used in this study (AlexaFluor488-AlexaFluor647, R_0 = 52 Å) equals to 30-80 Å, and less reliable performance outside of these boundaries, we performed non-uniform binning of the distance distributions, such that distances < 30 Å and > 80 Å are represented by a single bin each. The overall discrepancy between x_m and x_e per each distance bin, and for each of the k FRET measurements was calculated as follows:

$$\chi^2_{\mathrm{r,total}} = \frac{1}{K}\sum_{k=1}^{K}\frac{1}{N_b}\sum_{i=1}^{N_b}\left[\frac{x_{\mathrm{e}}(R_{\mathrm{DA}}^{(i)})_k - x_{\mathrm{m}}(R_{\mathrm{DA}}^{(i)})_k}{x_{\mathrm{e}}(R_{\mathrm{DA}}^{(i)})_k}\right]^2 \quad (6.2.2)$$

MEM benchmark using synthetic data

In the first step we wanted to assess the performance of MEM, i.e. to establish how to optimally balance different sources of information, and to assess which ensemble representations can be robustly recovered by MEM. For that we computed synthetic FRET data using our prior conformational ensembles, and that way the structural models underlying the ground truth/target *Y* are known (**Figure 11A**).

To cover different experimental scenarios, we created two target ensembles, namely *Y*(FF14SB) and *Y*(all). In the case of *Y*(FF14SB), cluster representatives and their corresponding weights obtained using FF14SB force field represent target ensemble. In the case of *Y*(all), target ensemble is a mixture of structures randomly selected from all prior ensembles, where the contribution of each FF in the target ensemble, in terms of number of structures, is proportional to the size of the prior ensemble, i.e.: 21.5% of structures in Y(all) originate from FF14SB, 17.2% from FF19SB, 6.6% from FF99SB, 16.4% from IDPSFF and 38.3% from FF99SBdisp. Uniform weights are assigned to all members in Y(all).

For target ensembles with such underlying structural models and corresponding weights, we generated synthetic fluorescence decays for a network of 50 optimally selected FRET pairs (**Figure 11B**).[145] Using previously described Bayesian framework (eq. 2.29 and eq. 2.32), synthetic fluorescence decays were converted to inter-dye distance distributions, which are then used in MEM optimization (**Figure 11C**). When using Y(FF14SB), as expected, the prior *I*(FF14SB) has initially the lowest $\chi^2_{\mathrm{r,total}}$, and $\chi^2_{\mathrm{r,total}} = 1$ is obtained with the least perturbation of prior weights (**Figure 11D**). However, on the case of other priors, we have learnt that to compare and rank the priors, one should consider not only the initial agreement with experimental data, but the entire evolution of L-shaped curve, i.e. the position of its corner (elbow), as well the reduction of $\chi^2_{\mathrm{r,total}}$ at the corner point.

How to optimally balance different sources of information?

When solely looking at $\chi^2_{\mathrm{r,total}}$, θ value which yields $S \approx -3$, here denoted as $\theta_{S\approx-3}$, appears as senseful choice of θ, since majority of the FFs have virtually converged L-curves at that point. However, we found that prior ensembles are heavily perturbed at $\theta_{S\approx-3}$, with change in

individual weights up to 9 orders of magnitude, and that few structures dominate the ensemble. Such observations are indicative of overfitting, and that such selected θ value does not provide optimal balance between different sources of information, despite good agreement with experimental data.

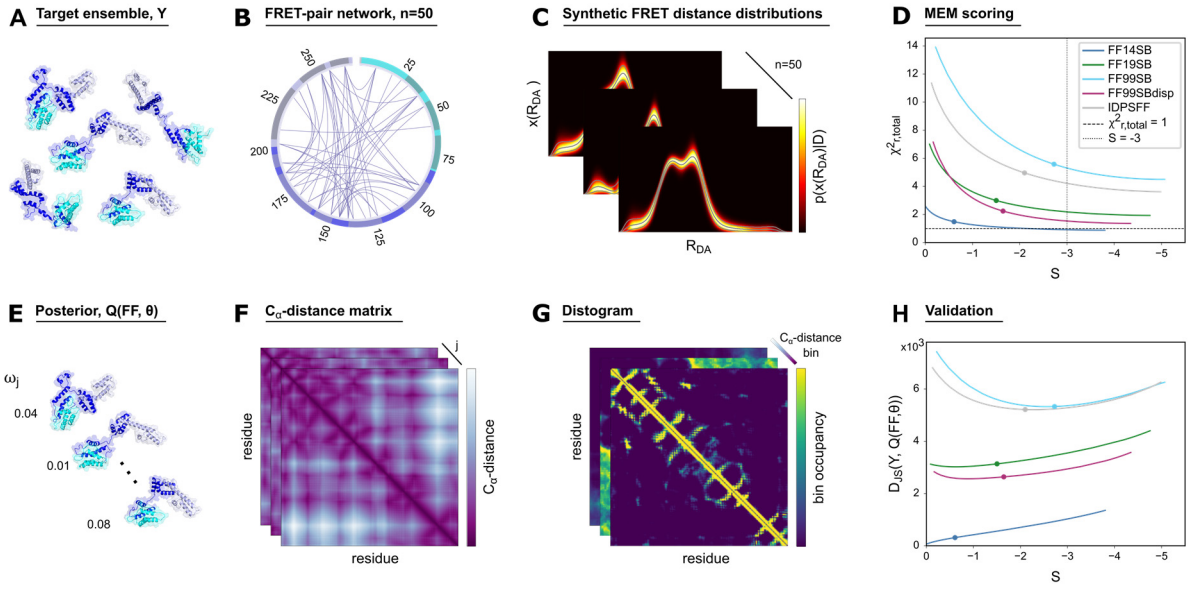


Figure 11 | MEM reweighting and validation of posterior ensembles using synthetic FRET data. *For a given target ensemble, here Y(FF14SB) (**A**) and a network of the 50 most informative FRET pairs (**B**), we simulate fluorescence decay curves and convert them by sampling probability distributions over the population of distances (synthetic distance distribution, $x_e(R_{DA})$) (**C**). The total reduced discrepancy, $\chi^2_{r,total}$, between the model distance distributions, $x_m(R_{DA})$, of the posterior conformational ensemble recovered by MEM and $x_e(R_{DA})$ for all FRET network members versus the entropy, S (**D**). The circles mark θ_c points that balance the synthetic data D and the prior I based on the L-curve criterion. For a given posterior ensemble with structures and weights (**E**), we compute residue pairwise Cα-distance matrices (**F**) and corresponding population-weighted distograms (**G**). Each point in a distogram corresponds to a distance occupancy computed for Cα atoms of structures of an ensemble. To compare two ensembles, we compute the Jensen-Shannon divergence (D_{JS}) between the posterior distogram and the distogram computed for Y, here Y(FF14SB) (**H**). Figure and figure caption are taken from the article by Dittrich J., Popara M., et al (for details see Supplement B)*

In synthetic experiment, we know the structural models underlying the ground truth, meaning that we can validate MEM results beyond looking at mere agreement with experimental data in terms of $\chi^2_{\mathrm{r,total}}$. Therefore, we computed pair-wise inter-residue C_α-distance matrices for each of the structural models in the ensembles, which were converted, using the cluster weights of ensemble members, to inter-residue distance histograms (distograms) (**Figure 11E-G**). To quantify the similarity between the distograms computed for posterior and target, we calculated

Jensen-Shannon divergence[111] between two distograms, $D_{\mathrm{JS}}(Y, Q(\mathrm{FF}, \theta))$. We found that in the case of complete prior, i.e. when optimizing I(FF14SB) against Y(FF14SB), D_{JS} continuously increases. For other FFs, D_{JS} curve has a minimum in the proximity of a corner point of the L-shaped curve (**Figure 11H**). Evolution of D_{JS} curves confirmed that MEM reconstructions are not reliable at $\theta_{S\approx-3}$, and that instead, more senseful choice of θ is at the corner point of L-shaped curve, here referred to as θ_{c}, when we are much closer to the minimum of D_{JS} curve.

Which model representations can be robustly recovered by MEM?

Atomistic models. It has already been shown that when using temporal and/or ensemble averaged experimental data, resolving ensemble members at the atomistic level is an ill-defined problem, meaning that MEM will allow multiple solutions dependent on prior, which can all fit the same averaged data.[98, 104, 146-148] In our case, we did not use a simple ensemble averaged observable, but a *distribution* of distances from minimally averaged, time-resolved FRET measurements, where the averaging occurs on a nanosecond time-scale, i.e. during the fluorescence lifetime of the donor. Nevertheless, FRET informs on long-range features, and not on absolute atomic positions. Therefore, it is not expected that unique solution can be obtained for different priors, when individual structures at atomistic resolution and their weights are considered. We computed C_α-RMSD matrices between the top 50 most populated clusters in posterior and target ensemble, and we found that even the best fitting force field, i.e. Q(FF14SB, θ_{c}), is on average ≈ 20 Å away from the target, which is the level of dissimilarity that one would obtain when comparing two randomly chosen proteins. Furthermore, remarkable dissimilarity was also obtained when cross-comparing posterior ensembles- we found no structural similarity amongst most populated clusters, and no correlation in weights for similar structures between two posterior ensembles. For details see Supplement B.

Inter-residue distograms. We asked ourselves, if there is instead a model representation of reduced level of detail compared to atomistic models, that can be robustly recovered by MEM. As a first step in resolution reduction, we computed inter-residue distance histograms (distograms), as a "medium-resolution" representation of an ensemble (for details see **Chapter 4.5** and Supplement B). This representation integrates over the ensemble members, but retains the residue information. We then reduce the information complexity even further, by looking at the moments of the inter-residue distograms, i.e. mean $\bar{R}$ and standard deviation σ_R. To quantify the similarity between two distograms, we computed squared relative differences of $\bar{R}$

and σ_R ($\delta^2(\bar{R})$ and $\delta^2(\sigma_R)$). For *Y*(FF14SB) and *I*(FF14SB), which is the case of complete prior, we get an excellent agreement between the distograms in terms of both the mean and standard deviation (**Figure 12A**). For *Y*(FF14SB) and incomplete prior, e.g. *I*(FF99SBdisp), initial deviations are larger, but nevertheless MEM improves $\delta^2(\bar{R})$ 4-fold, while σ_R on average remains unchanged (**Figure 12B**).

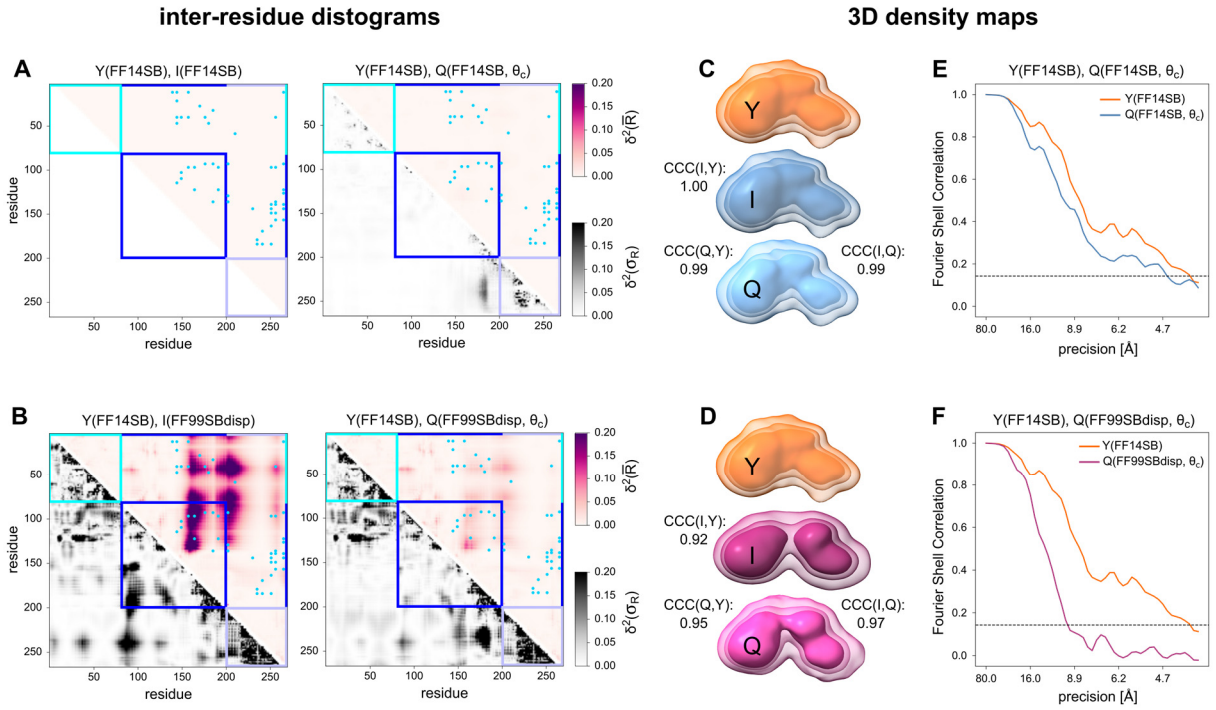


Figure 12 | Robustness of distogram and 3D density map ensemble representations. Line 1*: case Y = I, a single FF is the target Y(FF14SB) and the identical FF was chosen as prior I.* ***Line 2****: case Y ≠ I, a single FF is the target Y(FF14SB) and another FF was chosen as prior I(F99SBdisp)* ***(A, B)*** *Residue-wise squared relative deviation of prior (left) and posterior (right) of the distogram mean distance $\delta^2(\bar{R})$ (upper-right triangles) and standard deviation $\delta^2(\sigma_R)$ (lower-left triangles) to Y. Cyan, dark blue, and light blue squares and bars mark the residue range of the MD1, EHD, and MD2, respectively. Blue dots indicate FRET pairs.* ***(C, D)*** *Density maps of the target ensemble Y (orange), the prior I (FF14SB - blue, FF99SBdisp - dark pink), and the posterior Q (FF14SB - light blue, FF99SBdisp - light pink), outline the extent of the experimental ensembles at 50%, 68% and 90% of the density-weighted volume. The cross-correlation coefficients (CCCs) displayed to the left of the densities quantify similarities of prior and posterior entire density maps to the target ensemble, Y, and the ones on the right similarities of prior and posterior density maps.* ***(E, F)*** *Fourier Shell Correlations (FSCs) of the two half-density maps provide precision estimates for target Y(orange) and posterior Q (FF14SB - blue, FF99SBdisp - pink) density maps. The horizontal dashed line at FSC = 0.143 serves for a precision estimate of a density map. Figure and figure caption are adapted from the article by Dittrich J., Popara M., et al (for details see Supplement B)*

3D density maps. As further reduced ensemble representation, we used 3D density maps, a concept adopted from cryo-EM field. In such representation, we have further information loss, by averaging over both the ensemble members and residues (for details see **Chapter 4.5** and Supplement B). As a similarity measure between two density maps we used cross-correlation coefficient (CCC). We found that CCCs to the target ensemble are larger for the posterior than for the prior, which proves that MEM reliably recovers ensemble-integrated features (**Figure 12C-D**). Furthermore, we estimated the precision of the posterior density maps, by means of Fourier shell correlation (FSC) curves[149]. Ensembles were split into two sets of structures of the same size, by using interleaved ensemble member numbers. Subsequently, FSC curves were computed between two half-maps computed for two sets of structures. Precision of 3D density map was found to be in range of 5-10 Å, and is determined at the point where FSC curve crosses the threshold of 0.143[150] (**Figure 12E-F**).

MEM refinement using experimental data

When using the experimental data, we found that for all priors we needed a large entropy value in order to obtain good agreement with experimental data (**Figure 13A**). Even at the corner point of L-shaped curves, perturbation of weights spans across several orders of magnitude. This is indication that none of the priors is suitable to describe the data. However, we wanted to assess if the posterior ensembles nevertheless converged to unique solution in terms of model representations with reduced level of detail, such as inter-residue distograms or 3D density maps. For the two best fitting FFs, namely FF14SB and IDPSFF, with $\chi^2_{\mathrm{r,total}}$ = 2.011 and $\chi^2_{\mathrm{r,total}}$ = 2.069 at their respective corner points, we found that they better agree in distogram representation before MEM, than after MEM refinement (**Figure 13B)**. Particularly high differences in terms of means of distograms were observed in EHD region as well for distances between MD1 and EHD. This is likely a consequence of much smaller number of experimental FRET restraints (6) compared to the synthetic benchmark (50), and a consequence of EHD internal conformations not being sufficiently probed.

Likewise, similarity between the 3D density maps of FF14SB and IDPSFF is not improved for posteriors at θ_c (CCC=0.88) compared to the priors (CCC=0.88). However, analyzing the CCC values of all pairwise combinations of FFs, we could identify that our five distinct priors converged to three classes of posterior density maps (**Figure 13C**). First class, consisted of FF14SB, FF19SB and FF99SBdisp, retains a headphone-like shape of the homology model.

Second and third class are formed by FF99SB and IDPSFF. FF99SB density map is more compact, while the one of IDPSFF is extended. However, the solutions of two best fitting FFs are notably distinct, and with the data available we could not judge which one, or a mixture of both describes a true behavior of Lif in solution. To solve this ambiguity of the solution, additional experimental data is necessary, such as SAXS scattering profiles. We computed theoretical scattering profiles for the posterior ensembles, and showed that their differences are large enough to score differently against experimental SAXS data (for details see Supplement B).

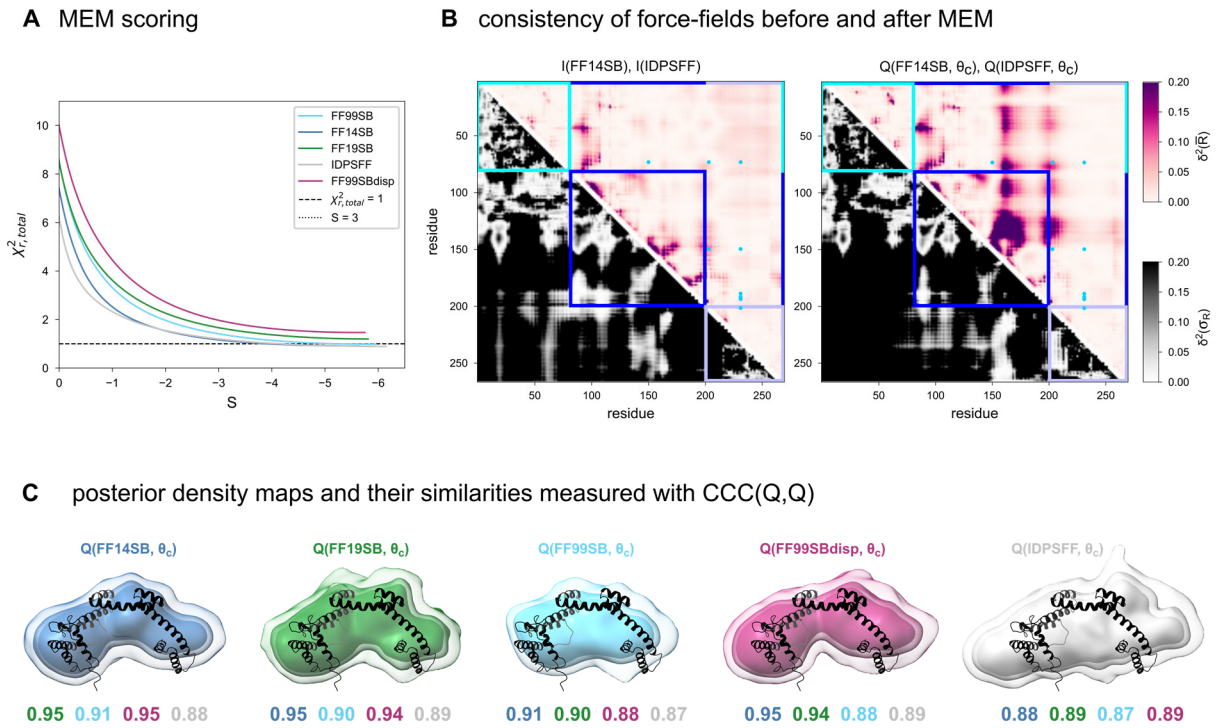


Figure 13 | Experimental inter-residue distograms and posterior density maps of Lif. (A) *The total reduced discrepancy $\chi^2_{r,total}$ between experimental, $x_e(R_{DA})$, and model distance distributions, $x_m(R_{DA})$, versus entropy S.* ***(B)*** *Residue-wise squared relative deviation of distogram mean distance $\bar{R}$ and standard deviation σ_R between FF14SB and IDPSFF prior (left) and posterior (right) ensembles.* ***(C)*** *Posterior density maps of Q(FF, θ_c) contoured at 50, 68, and 90% of the density-weighted volume aligned to all atoms of the homology model are displayed as cartoon. Numbers below the density maps represent Cross Correlation Coefficient (CCC) values between the posterior density map of a respective force field to all the others. CCC(Q,Q) values are color-coded according to the force field coloring scheme. Figure and figure caption are adapted from the article by Dittrich J., Popara M., et al (for details see Supplement B)*

Conclusions and significance

Lipase specific foldases (Lifs) are responsible for the activation of their cognate lipases in the periplasm of Gram-negative bacteria. Lipases are the most important class of enzymes in biotechnology, and it is therefore essential to understand the behavior of their chaperons. In this article we provided first-ever insights into the structural ensemble of *Pa*Lif in apo state.
Main outcomes of this study are:

- We established the workflow for integrative modelling using time-resolved FRET data and structural ensembles obtained from all-atom MD simulations.
- We found that in the absence of posterior uncertainty estimates, varying priors is an approach to estimate the robustness of MEM results.
- Through the benchmark using synthetic data, we found that optimal balance between different sources of information is obtained using θ value at the corner of L-shaped curve.
- To rank and compare the priors, one has to consider not only the initial $\chi^2_{r,total}$, but the entire evolution of the L-shaped curve, i.e. the position of the corner point, as well the reduction of $\chi^2_{r,total}$ at the corner point.
- Resolving the individual ensemble members at the atomistic resolution is an ill-defined problem, meaning that MEM will yield multiple solutions dependent on the prior, that will all fit the same experimental data.
- Depending on the agreement of the prior and experimental data, MEM can reliably recover only ensemble representations of reduced level of detail, such as inter-residue distance histograms (distograms) and 3D density maps.
- Using experimental fluorescence decays acquired for a network of 6 FRET pairs, we did not obtain consistency of posterior ensembles when using medium-resolution distogram representation. However, using 3D density map representation we obtained consistent result for three out of five priors.
- 3D density maps of two best fitting solutions have distinct shape, i.e. headphone-like and extended, and at present we cannot judge which one describes apo Lif in solution.

6.3. Manuscript 1: Increasing the accuracy of single-molecule FRET experiments – robust and unbiased estimation of correction parameters

This section is based on the manuscript in preparation by Popara M. *et al* (Supplement C).

In this work I established robust and unbiased workflow (see **Figure 14**) for estimation of correction parameters, that are necessary for the accurate determination of quantitative FRET observables, such as FRET efficiency and stoichiometry (see **Chapter 2.4**). Motivation for this project came from the results of a blind study (see **Chapter 6.1** and Supplement A**)**, where we showed that the observed deviations between the lab-reported FRET efficiencies can be fully explained by the systematic errors in determination of correction parameters. In particular, deviations could be explained by the error in detection efficiency parameter alone. That is why the particular focus of this work was devoted to determination of detection efficiency parameter. In a single-molecule FRET experiment, donor and acceptor emitted photons are distorted due to the background signal, spectral overlap of the fluorophores, and the non-ideal optical configuration of the setup that allows photons to be misguided to wrong detection channel. To get the accurate FRET quantities, such as FRET efficiency and stoichiometry, a set of correction parameters needs to be applied to raw, distorted, signal. First, fluorescence photons emitted by a single molecule in a focus are always interfered with unwanted background photons, that can originate from sample impurities, Rayleigh scattering at the excitation wavelength, or dark counts of detector. Estimation and correction for background emission is the first step towards accurate FRET observables. Afterwards, four additional correction parameters are applied to the data, namely:

(*i*) **crosstalk $\boldsymbol{\alpha}$**, which corrects FRET sensitized acceptor signal for the leakage of donor photons into acceptor detection channel (eq. 2.19 and eq. 2.21). Absence of this correction falsely increases FRET sensitized acceptor signal, which leads to overestimation of FRET efficiency. As shown already (see **Figure 9A**), the uncertainty in *E* due to donor crosstalk is largest for the low FRET efficiencies, but is significant also for the range of intermediate FRET efficiencies where most of the measurements are being performed. Crosstalk α is determined from the donor-only sub-population, by performing the robust linear regression between the corresponding background-corrected photon counts:

$$N_{\mathrm{R|G}} - S_{\mathrm{R|G}}^{(\mathrm{BG})} T = \alpha (N_{\mathrm{G|G}} - S_{\mathrm{G|G}}^{(\mathrm{BG})} T) \tag{6.3.1}$$

$N_{\mathrm{Y|X}}$ and $S_{\mathrm{Y|X}}^{(\mathrm{BG})}$ denote raw photon counts and background count-rate in a detection channel Y after X excitation, and T represents burst duration. In the protocol, provided are instructions on how to select donor-only population in order to assure the least amount of user bias. This is followed with the guidelines concerning the regression fit, and with recommendations on how to validate the obtained α value.

(*ii*) **direct excitation $\boldsymbol{\delta}$**, which corrects FRET sensitized acceptor signal for the direct acceptor excitation with the donor excitation laser (eq. 2.19 and eq. 2.22). Similarly as for crosstalk, absence of this correction leads to overestimation of FRET efficiency, and it has increasing effect towards low FRET efficiencies (**Figure 9A**). Direct excitation correction parameter is determined from acceptor-only sub-population, by performing robust regression between the corresponding background-corrected photon counts:

$$N_{\mathrm{R|G}} - S_{\mathrm{R|G}}^{(\mathrm{BG})}T = \delta(N_{\mathrm{R|R}} - S_{\mathrm{R|R}}^{(\mathrm{BG})}T) \tag{6.3.2}$$

In the protocol, provided are instructions for selecting acceptor-only sub-population, and recommendations for cross-validation of obtained δ value.

(*iii*) **detection efficiency $\boldsymbol{\gamma}$**, which represents correction of donor signal after donor laser time-on (eq. 2.23 and eq. 2.24) for distinct fluorescence quantum yields of donor and acceptor, as well for different detection sensitivities of green and red detection channels for donor and acceptor fluorophore.

(*iv*) **correction parameter $\boldsymbol{\beta}$**, which is applied to acceptor signal upon acceptor laser time-on (eq. 2.25 and eq. 2.26), and which compensates for different excitation rates of donor and acceptor with their corresponding excitation light.

In the three methods proposed, γ and β correction parameters are determined jointly, using double-labelled donor-acceptor (DA) sub-population. Therefore, the first step is to ensure robust selection of DA species. For that I proposed selection method based on ALEX-2CDE filter[21], that in the least invasive way removes single-labelled species (DO, AO), as well the species with unstable emission of the fluorophores. Three methods are proposed, namely: (M- 1) non-linear regression of burst-wise values of ${}^{iii}S_{\mathrm{app}}$ versus ${}^{iii}E_{\mathrm{app}}$; (M-2) linear regression between $1/\langle {}^{iii}S_{\mathrm{app}}\rangle$ and $\langle {}^{iii}E_{\mathrm{app}}\rangle$ of the population means; (M-3) linear regression on photon counts.

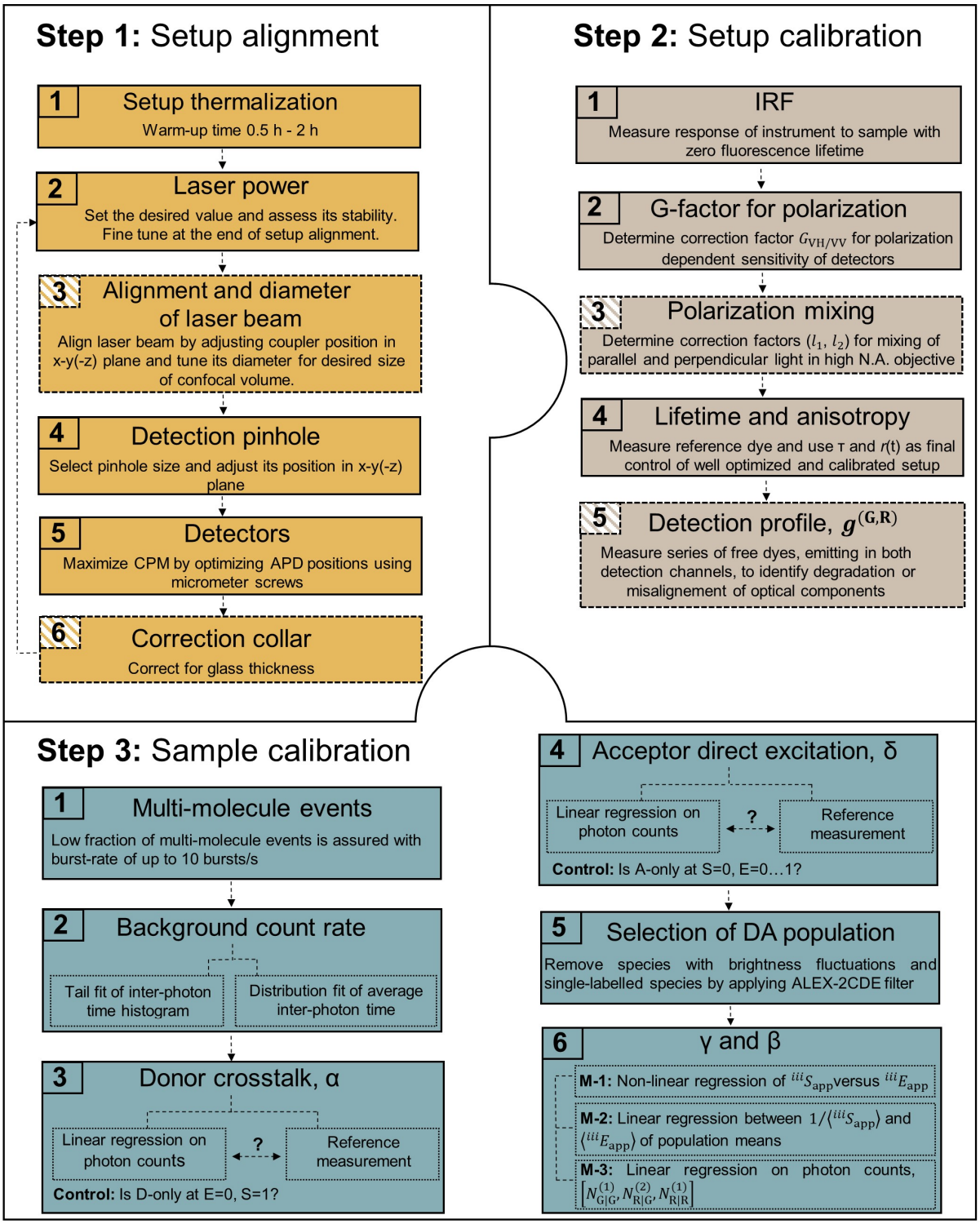


Figure 14 | Workflow diagram for MFD-PIE confocal setup alignment and setup and sample calibration. *Steps with dotted frame are typically not needed in day-to-day alignement/calibration.*

In method M-1, one performs non-linear regression of burst wise values of ${}^{iii}S_{\mathrm{app}}$ versus ${}^{iii}E_{\mathrm{app}}$, which are background, α and δ corrected. Such procedure requires either multiple DA

species with different FRET efficiencies, or a single sample that displays sufficient broadening of FRET efficiency distribution. The underlying equation for this approach is:

$$^{iii}S_{\mathrm{app}} = (1 + \gamma\beta + (1 - \gamma)\beta\,{}^{iii}E_{\mathrm{app}})^{-1} \tag{6.3.3}$$

This method is based on the homogeneous approximation, i.e. it requires uniform fluorophore properties across all samples used in regression.

In method M-2 one performs linear regression between $1/\langle {}^{iii}S_{\mathrm{app}}\rangle$ and $\langle {}^{iii}E_{\mathrm{app}}\rangle$, where $\langle {}^{iii}S_{\mathrm{app}}\rangle$ and $\langle {}^{iii}E_{\mathrm{app}}\rangle$ are the means of the ${}^{iii}S_{\mathrm{app}}$ and ${}^{iii}E_{\mathrm{app}}$ histograms of selected DA population.[151] γ and β are then obtained from the intercept and slope of the linear regression:

$$1/\langle {}^{iii}S_{\mathrm{app}}\rangle = 1 + \gamma\beta + (1 - \gamma)\beta\langle {}^{iii}E_{\mathrm{app}}\rangle \tag{6.3.4}$$

This method requires minimum two DA populations for the regression, with uniform fluorophore properties. In the article by Agam, G., Gebhardt, C., Popara M., *et al*[129] (Supplement A), we evaluated how much is the uncertainty of γ parameter affected by the FRET efficiency contrast between the populations used in γ determination (**Figure 15**). In the simplest case of two samples, propagated uncertainty Δγ, based on the errors in the population mean estimates $\langle {}^{iii}S_{\mathrm{app},j}\rangle$ and $\langle {}^{iii}E_{\mathrm{app},j}\rangle$, is given by:

$$\Delta\gamma = \sqrt{\sum_{j=1,2} \frac{\partial\gamma}{\partial\langle {}^{iii}S_{\mathrm{app},j}\rangle}\Delta\langle {}^{iii}S_{app,j}\rangle + \frac{\partial\gamma}{\partial\langle {}^{iii}E_{\mathrm{app},j}\rangle}\Delta\langle {}^{iii}E_{\mathrm{app},j}\rangle} \tag{6.3.5}$$

where the partial derivatives are calculated according to the chain rule:

$$\frac{\partial\gamma}{\partial u} = \frac{\partial\gamma}{\partial\Sigma}\cdot\frac{\partial\Sigma}{\partial u} + \frac{\partial\gamma}{\partial\Omega}\cdot\frac{\partial\Omega}{\partial u} \tag{6.3.6}$$

with $u \in \{\langle {}^{iii}S_{\mathrm{app},1}\rangle, \langle {}^{iii}E_{\mathrm{app},1}\rangle, \langle {}^{iii}S_{\mathrm{app},2}\rangle, \langle {}^{iii}E_{\mathrm{app},2}\rangle\}$, and Ω and Σ being the intercept and slope of the regression.

We found that all partial derivatives $\frac{\partial\Sigma}{\partial u}$ and $\frac{\partial\Omega}{\partial u}$ are an inverse function of apparent FRET efficiency difference:

$$\frac{\partial\gamma}{\partial\Sigma}=-\frac{\gamma}{\beta};\qquad \frac{\partial\gamma}{\partial\Omega}=\frac{1-\gamma}{\beta};$$

$$\frac{\partial\Sigma}{\partial\langle^{iii}S_{\mathrm{app},1}\rangle}=\frac{1}{\langle^{iii}E_{\mathrm{app},2}\rangle-\langle^{iii}E_{\mathrm{app},1}\rangle}\cdot\frac{1}{\langle^{iii}S_{\mathrm{app},1}\rangle^2};\qquad \frac{\partial\Sigma}{\partial\langle^{iii}S_{\mathrm{app},2}\rangle}=\frac{-1}{\langle^{iii}E_{\mathrm{app},2}\rangle-\langle^{iii}E_{\mathrm{app},1}\rangle}\cdot\frac{1}{\langle^{iii}S_{\mathrm{app},2}\rangle^2};$$

$$\frac{\partial\Sigma}{\partial\langle^{iii}E_{\mathrm{app},1}\rangle}=-\frac{\Sigma}{\langle^{iii}E_{\mathrm{app},2}\rangle-\langle^{iii}E_{\mathrm{app},1}\rangle};\qquad \frac{\partial\Sigma}{\partial\langle^{iii}E_{\mathrm{app},2}\rangle}=\frac{\Sigma}{\langle^{iii}E_{\mathrm{app},2}\rangle-\langle^{iii}E_{\mathrm{app},1}\rangle};\tag{6.3.7}$$

$$\frac{\partial\Omega}{\partial\langle^{iii}S_{\mathrm{app},1}\rangle}=\frac{-\langle^{iii}E_{\mathrm{app},2}\rangle}{\langle^{iii}E_{\mathrm{app},2}\rangle-\langle^{iii}E_{\mathrm{app},1}\rangle}\cdot\frac{1}{\langle^{iii}S_{\mathrm{app},1}\rangle^2};\qquad \frac{\partial\Omega}{\partial\langle^{iii}S_{\mathrm{app},2}\rangle}=\frac{\langle^{iii}E_{\mathrm{app},1}\rangle}{\langle^{iii}E_{\mathrm{app},2}\rangle-\langle^{iii}E_{\mathrm{app},1}\rangle}\cdot\frac{1}{\langle^{iii}S_{\mathrm{app},2}\rangle^2};$$

$$\frac{\partial\Omega}{\partial\langle^{iii}E_{\mathrm{app},1}\rangle}=\frac{-\langle^{iii}E_{\mathrm{app},2}\rangle}{\langle^{iii}E_{\mathrm{app},2}\rangle-\langle^{iii}E_{\mathrm{app},1}\rangle}\Sigma;\qquad \frac{\partial\Omega}{\partial\langle^{iii}E_{\mathrm{app},2}\rangle}=\frac{\langle^{iii}E_{\mathrm{app},1}\rangle}{\langle^{iii}E_{\mathrm{app},2}\rangle-\langle^{iii}E_{\mathrm{app},1}\rangle}\Sigma.$$

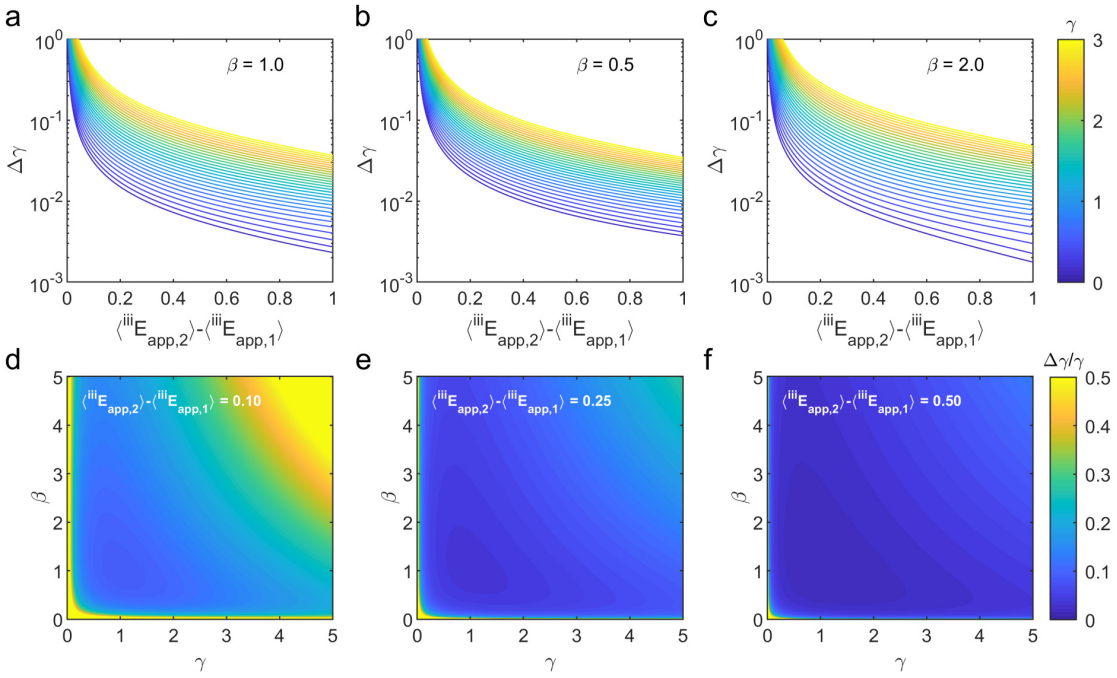


Figure 15 | The propagated error of the γ-factor based on the uncertainty of the population mean estimate in the apparent FRET efficiency histogram. *For the calculations shown here, the uncertainties of the population mean estimates in the $\langle^{iii}E_{app}\rangle$ vs. $\langle^{iii}S_{app}\rangle$ plot are given by $\Delta\langle^{iii}E_{app}\rangle=\Delta\langle^{iii}S_{app}\rangle=0.05/\sqrt{1000}$. The position of the two populations is chosen to be symmetric with respect to $\langle^{iii}E_{app}\rangle=0.5$, i.e., $\langle^{iii}E_{app,j}\rangle=0.5\pm\left(\Delta\langle^{iii}E_{app,2}\rangle-\Delta\langle^{iii}E_{app,1}\rangle\right)/2$. **(a-c)** The absolute uncertainty $\Delta\gamma$ is evaluated as a function of the difference of the apparent FRET efficiencies of the two populations, $\langle^{iii}E_{app,2}\rangle-\langle^{iii}E_{app,1}\rangle$, at the indicated values of the γ factor (see color bar) and β values of 1 **(a)**, 0.5 **(b)**, and 2 **(c)**. **(d-f)** The relative uncertainty $\Delta\gamma/\gamma$ is evaluated as a function of the absolute values of γ and β, and is plotted for difference between the apparent FRET efficiencies of 0.1 **(g)**, 0.25 **(h)**, and 0.5 **(i)**. Figure and figure caption are adapted from Agam, G., Gebhardt, C., Popara M., et al[129] (Supplement A).*

We observed that the total absolute uncertainty of detection efficiency, Δγ, is indeed the highest for the small FRET efficiency contrast between the DA populations used in regression, and it drops as the contrast increases. However, we also found that the absolute uncertainty in γ

parameter depends on the absolute values of both γ and β parameters (**Figure 15a-c**). To further explore this dependence, we computed the relative error Δγ/γ for a range of γ and β values, at the fixed differences $\langle {}^{iii}E_{\mathrm{app},2}\rangle - \langle {}^{iii}E_{\mathrm{app},1}\rangle$ of 0.1, 0.25 and 0.5 (**Figure 15d-f**). We found that relative uncertainty of γ is the lowest for intermediate values of γ and β around the value of 1. Interestingly, Δγ/γ remains constant in large range of γ and β values, from ~0.1 to ~3, where most experiments are performed. In such range of γ and β values, neither does the FRET efficiency contrast plays notable role. Only at extreme values of γ and β, do they have large effect on γ uncertainty. These findings were experimentally confirmed in a previous study. [152]

As opposed to M-1 and M-2 methods, which rely on ratiometric quantities ${}^{iii}E_{\mathrm{app}}$ and ${}^{iii}S_{\mathrm{app}}$, method M-3 takes advantage of photon counts directly. This approach is based on the following relation:

$$N_{\mathrm{R|R}}^{(1)} = \gamma\beta N_{\mathrm{G|G}}^{(1)} + \beta N_{\mathrm{R|G}}^{(2)} \tag{6.3.8}$$

which is derived under the assumption of D:A=1:1 labelling stoichiometry.[153] Essentially, in this approach one performs a plane fit in 3D space, which is defined by the coordinates $[N_{\mathrm{G|G}}^{(1)}, N_{\mathrm{R|G}}^{(2)}, N_{\mathrm{R|R}}^{(1)}]$. While on one hand the pitfall of this approach is that it requires labelling stoichiometry of a FRET construct to be known and equal to 0.5, on the other hand this approach has advantage of not being sensitive to the asymmetry or presence of shoulders in the histogram, as is the case for methods M-1 and M-2.

For the benchmark of designed protocol for determination of correction parameters, we used a set of double-stranded DNA molecules, with high-, medium- and low-FRET efficiencies, covering distances of ~ 50-70 Å. Considering dsDNA molecule to be static and provided photophysical stability of the tethered fluorophores and absence of dye-protein sticking interactions, no displacement of the DA population from the static FRET-line should be observed, when all levels of correction are applied. Therefore, one way to assess the accuracy of determined correction parameters is by quantifying the displacement of DA sub-populations from the static FRET-line. As an additional way of evaluating parameter accuracy, we quantified the agreement between experimentally derived FRET efficiencies and values predicted using structural model of double-stranded DNA and AV forward model. No notable displacement from the static FRET-line was observed, and moreover, we found agreement between experimental and predicted FRET-averaged inter-dye distances, with RMSD across all dye-pairs and samples on the order of ~3 Å, which is close to the accuracy previously reported

for dsDNA.[29] Furthermore, these two evaluation procedures did not reveal discrepancies between different methods of γ and β determination. However, a blind test, involving seven laboratory members, revealed superiority of methods M-1 and M-3, compared to method M-2 where we observed larger discrepancy between user-reported γ and β values. This can be explained with the fact that method M-2 in the first step requires determination of population means in $\langle {}^{iii}E_{\mathrm{app}} \rangle$ - $\langle {}^{iii}S_{\mathrm{app}} \rangle$ histogram. Inconsistencies in fit model and assignment of the main population are underlying cause of the observed discrepancies between the user reported values.

Conclusions and significance

In a recent FRET community paper[129] (see Supplement A), it was identified that the error in determination of correction parameters is the main cause of the discrepancies between FRET efficiencies determined by different laboratories. This suggested the necessity for optimized protocol for the determination of correction parameters, that will minimize the extent of user bias, and further push the limits of FRET measurements in terms of precision and accuracy. Driven by these insights, this project was initiated, and building upon existing practices in the field, we designed minimally user-biased protocol. The main outcomes of this study are:

- As a first step towards accurate FRET measurements, we provided elaborate instructions on how to align and optimize confocal setup for MFD-PIE diffusion-based FRET experiments.
- Coming to sample-related correction parameters, for each of the steps we defined unambiguous steps on how to select a sub-population (DO, AO and DA), necessary for determination of the respective correction parameter.
- Presented are theoretical foundations underlying determination of corrections parameters, followed by discussions on their application versatility and performance.
- All correction parameters can be obtained from the measurement of the sample of interest. No separate measurements, which could introduce the error, are needed.
- Robustness of proposed protocol is evaluated in a blind study comprising seven lab members.
- Accuracy of fitted correction parameters is confirmed through two independent evaluation procedures.
- If nature of the samples allows, users should apply multiple methods for determination of γ and β parameters and check for their consistency, rather than exclusively using single approach.

Chapter 7

Summary and Perspectives

In this thesis, I showed the potential of smFRET measurements for dissecting the inner workings of proteins, from unraveling the structural features down to Ångström resolution, to resolving structural dynamics across multiple time-scales, with unparalleled time resolution. The challenge of observing conformational dynamics lies in a fact that dynamics occurs in unsynchronized manner in a molecular ensemble. However, in FRET experiment we are able to observe single-molecules at a time, circumventing the need for synchronization of the entire ensemble. Moreover, such ultimate level of sensitivity in smFRET experiment, where the ensemble averaging is circumvented, and temporal averaging occurs on the timescale of fluorescence lifetime (few ns), is achieved through relatively simple optical design, centered around a confocal microscope.

Although the studies demonstrating potentials of FRET are numerous, assessment of robustness of FRET measurements was done only for simple toy systems, such as double-stranded DNA molecules, dsDNA, that essentially behave as rigid rods.[29] Therefore, as one of the first aims of this thesis, overall performance of FRET measurements was critically assessed in a truly blind study on substantially more complex systems.

Using two protein systems, i.e. maltose-binding protein, MalE, and large subunit of U2 auxiliary factor, U2AF2, we verified that FRET measurements are reproducible and accurate across instrumentations, analysis tools and dynamic protein systems of different structural complexity and with varying dye environments. We found that all laboratories were able to observe expected conformational change upon ligand binding, with standard deviation across reported mean FRET efficiencies up to 0.06, which in terms of inter-dye distances corresponds to precision of ≤ 2 Å and accuracy ≤ 5Å. Deviations between the laboratories could be fully explained by the error and user bias in determination of correction parameters, and in particular by the error in detection efficiency parameter. Driven with these findings, in a follow-up project I established a robust protocol on how to determine accurate values of correction parameters, with minimum number of steps required and least amount of user bias (Supplement C). Furthermore, we went beyond evaluating precision and accuracy of FRET measurements. In a rigorous approach we showed how the presence of position-specific dye-protein sticking interactions can lead to false-positive identification of conformational dynamics, highlighting the importance of dye controls. Through elaborate analysis of fluorescence anisotropy, we were

able to establish a robust threshold for exclusion of measurements with dye artifacts. Using this exclusion rule, we were left with artifact-free measurements, for which we were ultimately able to detect small-scale structural fluctuations on the order of 5 Å for MalE, a current lower limit for detection of structural fluctuations. Lastly, on the example of U2AF2, we showed how conformational dynamics can be studied through the combination of multiple tools, such as PDA and filtered FCS.

The second main research area of this thesis was aimed at facilitating translation of distance information encoded in fluorescence decays into accurate structural models (Supplement B). We created a set of guidelines on how to harvest the information content from FRET measurements and combine it with MD simulations, in order to achieve detailed and robust ensembles of biomolecular structures. Our workflow is centered around the principle of maximum entropy, a method of ensemble reweighting that is the most suitable approach for systems with large number of states. As a target system we used lipase-specific foldase from *Pseudomonas aeruginosa* species in its apo state.

In the synthetic benchmark, we investigated how to properly balance prior information and experimental data, as well how to compare and rank different priors. Furthermore, we addressed the robustness of three different types of model representations for ensembles: atomistic models, inter-residue distance histograms (distograms) and 3D density maps. We found that even for the complete prior, at atomistic resolution MEM yields ambiguous solutions heavily dependent on the prior. However, the fact that individual structures with atomistic resolution could not be achieved, should not be taken as shortcoming, but as inherent characteristic of a method when using averaged experimental data. To this also contributes if the experimental method reports on long-range features, such as tertiary, super-tertiary and quaternary structure, and not directly on absolute atomic positions. To achieve posterior distribution with distogram level of resolution, good coverage of the conformational space should be assured, meaning that prior ensemble has to be part of the true ensemble underlying experimental data. Even in that case, means of distograms are better recovered than the width of distograms. In the case of 3D density maps, we found that MEM robustly recovers such ensemble representation, for both complete and incomplete priors. Lastly, using experimental fluorescence decays of limited set of FRET pairs (6), we could obtain consistent 3D density maps for 3 out of the 5 priors. However, two best fitting solutions are different, and at present, without additional information available, we cannot judge which, or if the combination of both solutions describes apo *Pa*Lif.

Overall, our results indicate that maximum entropy principle is a promising route for ensemble reweighting. However, for systems with large number of degrees of freedom, good coverage of conformational space in the prior, and as large as possible amount of experimental data points, are a prerequisite in order to achieve robust posterior reconstructions with high level of structural detail. Furthermore, due to the absence of posterior uncertainty estimates, varying priors is crucial in order to get an insight in how robust posterior reconstructions are.

We believe that these results represent major step forward for integrative modelling of biomolecular ensembles, and that we opened an important discussion on how the ensembles should be visualized, compared and disseminated. We posed an important question if atomistic models are indeed necessary or even useful when ensembles are large. Contrary to that, results of this thesis put a positive light on volumetric representations, as these can be robustly obtained irrespective of the completeness of the prior information. Moreover, even though volumetric representations are of lower resolution, they still have high information content, and have a proven track record of successful applications in pharmaceutical industry.

In the future it would be interesting to see how MEM performs when additional experimental methods, informing on orthogonal features, are included, either as a restraint or as a *post-hoc* validation. It would be also of interest to see how much the local-model accuracy would improve if fluorescence anisotropy measurements, which sense local environment and internal motions[154], are included.

To conclude, results of this thesis, which comprised different methodologies from the field of computer science and fluorescence spectroscopy, demonstrated that FRET measurements are an outstanding tool of structural biology, but they also revealed areas where FRET measurements are still imperfect, and challenges yet to be overcome.

Bibliography

1. Levinthal, C. In *How to fold graciously*, Mossbauer Spectroscopy in Biological Systems: Proceedings of a Meeting held at Allerton House, Monticello, Illinois, Debrunner P., T. J. C. M., Munck E. , Ed. University of Illinois Press, Urbana: Monticello, Illinois, 1969; pp 22-26.

2. Ward, A. B.; Sali, A.; Wilson, I. A., Biochemistry. Integrative structural biology. *Science* **2013,** *339* (6122), 913-5.

3. Rout, M. P.; Sali, A., Principles for Integrative Structural Biology Studies. *Cell* **2019,** *177* (6), 1384-1403.

4. Seffernick, J. T.; Lindert, S., Hybrid methods for combined experimental and computational determination of protein structure. *J Chem Phys* **2020,** *153* (24), 240901.

5. Peulen, T. O.; Hengstenberg, C. S.; Biehl, R.; Dimura, M.; Lorenz, C.; Valeri, A.; Folz, J.; Hanke, C. A.; Ince, S.; Vopel, T.; Farago, B.; Gohlke, H.; Klare, J. P.; Stadler, A. M.; Seidel, C. A. M.; Herrmann, C., Integrative dynamic structural biology unveils conformers essential for the oligomerization of a large GTPase. *Elife* **2023,** *12*.

6. Molnar, K. S.; Bonomi, M.; Pellarin, R.; Clinthorne, G. D.; Gonzalez, G.; Goldberg, S. D.; Goulian, M.; Sali, A.; DeGrado, W. F., Cys-scanning disulfide crosslinking and bayesian modeling probe the transmembrane signaling mechanism of the histidine kinase, PhoQ. *Structure* **2014,** *22* (9), 1239-1251.

7. Hamilton, G. L.; Saikia, N.; Basak, S.; Welcome, F. S.; Wu, F.; Kubiak, J.; Zhang, C.; Hao, Y.; Seidel, C. A. M.; Ding, F.; Sanabria, H.; Bowen, M. E., Fuzzy supertertiary interactions within PSD-95 enable ligand binding. *Elife* **2022,** *11*.

8. Naudi-Fabra, S.; Tengo, M.; Jensen, M. R.; Blackledge, M.; Milles, S., Quantitative Description of Intrinsically Disordered Proteins Using Single-Molecule FRET, NMR, and SAXS. *J Am Chem Soc* **2021,** *143* (48), 20109-20121.

9. Forster, T., Zwischenmolekularer Energieubergang in Farbstofflosungen. *Angew Chem-Ger Edit* **1948,** *60* (6), 163-164.

10. Kolb, H. C.; Finn, M. G.; Sharpless, K. B., Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angew Chem Int Ed Engl* **2001,** *40* (11), 2004-2021.

11. Muller, B. K.; Zaychikov, E.; Brauchle, C.; Lamb, D. C., Pulsed interleaved excitation. *Biophys J* **2005,** *89* (5), 3508-3522.

12. Kudryavtsev, V.; Sikor, M.; Kalinin, S.; Mokranjac, D.; Seidel, C. A. M.; Lamb, D. C., Combining MFD and PIE for Accurate Single-Pair Forster Resonance Energy Transfer Measurements. *Chemphyschem* **2012,** *13* (4), 1060-1078.

13. Widengren, J.; Kudryavtsev, V.; Antonik, M.; Berger, S.; Gerken, M.; Seidel, C. A., Single-molecule detection and identification of multiple species by multiparameter fluorescence detection. *Anal Chem* **2006,** *78* (6), 2039-50.

14. Clegg, R. M., Fluorescence Resonance Energy-Transfer and Nucleic-Acids. *Method Enzymol* **1992,** *211*, 353-388.

15. Hohng, S.; Joo, C.; Ha, T., Single-molecule three-color FRET. *Biophys J* **2004,** *87* (2), 1328-1337.

16. Gopich, I. V.; Szabo, A., Single-molecule FRET with diffusion and conformational dynamics. *J Phys Chem B* **2007,** *111* (44), 12925-32.

17. Kalinin, S.; Valeri, A.; Antonik, M.; Felekyan, S.; Seidel, C. A., Detection of structural dynamics by FRET: a photon distribution and fluorescence lifetime analysis of systems with multiple states. *J Phys Chem B* **2010,** *114* (23), 7983-95.

18. Felekyan, S.; Kalinin, S.; Sanabria, H.; Valeri, A.; Seidel, C. A., Filtered FCS: species auto- and cross-correlation functions highlight binding and dynamics in biomolecules. *Chemphyschem* **2012,** *13* (4), 1036-53.

19. Torres, T.; Levitus, M., Measuring conformational dynamics: a new FCS-FRET approach. *J Phys Chem B* **2007,** *111* (25), 7392-400.

20. Torella, J. P.; Holden, S. J.; Santoso, Y.; Hohlbein, J.; Kapanidis, A. N., Identifying molecular dynamics in single-molecule FRET experiments with burst variance analysis. *Biophys J* **2011,** *100* (6), 1568-77.

21. Tomov, T. E.; Tsukanov, R.; Masoud, R.; Liber, M.; Plavner, N.; Nir, E., Disentangling subpopulations in single-molecule FRET and ALEX experiments with photon distribution analysis. *Biophys J* **2012,** *102* (5), 1163-73.

22. Barth, A.; Opanasyuk, O.; Peulen, T. O.; Felekyan, S.; Kalinin, S.; Sanabria, H.; Seidel, C. A. M., Unraveling multi-state molecular dynamics in single-molecule FRET experiments. I. Theory of FRET-lines. *J Chem Phys* **2022,** *156* (14).

23. Opanasyuk, O.; Barth, A.; Peulen, T. O.; Felekyan, S.; Kalinin, S.; Sanabria, H.; Seidel, C. A. M., Unraveling multi-state molecular dynamics in single-molecule FRET experiments. II. Quantitative analysis of multi-state kinetic networks. *J Chem Phys* **2022,** *157* (3).

24. Lerner, E.; Barth, A.; Hendrix, J.; Ambrose, B.; Birkedal, V.; Blanchard, S. C.; Borner, R.; Sung Chung, H.; Cordes, T.; Craggs, T. D.; Deniz, A. A.; Diao, J.; Fei, J.; Gonzalez, R. L.; Gopich, I. V.; Ha, T.; Hanke, C. A.; Haran, G.; Hatzakis, N. S.; Hohng, S.; Hong, S. C.; Hugel, T.; Ingargiola, A.; Joo, C.; Kapanidis, A. N.; Kim, H. D.; Laurence, T.; Lee, N. K.; Lee, T. H.; Lemke, E. A.; Margeat, E.; Michaelis, J.; Michalet, X.; Myong, S.; Nettels, D.; Peulen, T. O.; Ploetz, E.; Razvag, Y.; Robb, N. C.; Schuler, B.; Soleimaninejad, H.; Tang, C.; Vafabakhsh, R.; Lamb, D. C.; Seidel, C. A.; Weiss, S., FRET-based dynamic structural biology: Challenges, perspectives and an appeal for open-science practices. *Elife* **2021,** *10*.

25. Schuler, B.; Lipman, E. A.; Steinbach, P. J.; Kumke, M.; Eaton, W. A., Polyproline and the "spectroscopic ruler" revisited with single-molecule fluorescence. *Proc Natl Acad Sci U S A* **2005,** *102* (8), 2754-9.

26. Stryer, L.; Haugland, R. P., Energy Transfer - a Spectroscopic Ruler. *P Natl Acad Sci USA* **1967,** *58* (2), 719-&.

27. Kilic, S.; Felekyan, S.; Doroshenko, O.; Boichenko, I.; Dimura, M.; Vardanyan, H.; Bryan, L. C.; Arya, G.; Seidel, C. A. M.; Fierz, B., Single-molecule FRET reveals multiscale chromatin dynamics modulated by HP1alpha. *Nat Commun* **2018,** *9* (1), 235.

28. Sanabria, H.; Rodnin, D.; Hemmen, K.; Peulen, T. O.; Felekyan, S.; Fleissner, M. R.; Dimura, M.; Koberling, F.; Kuhnemuth, R.; Hubbell, W.; Gohlke, H.; Seidel, C. A. M., Resolving dynamics and function of transient states in single enzyme molecules. *Nat Commun* **2020,** *11* (1), 1231.

29. Hellenkamp, B.; Schmid, S.; Doroshenko, O.; Opanasyuk, O.; Kuhnemuth, R.; Rezaei Adariani, S.; Ambrose, B.; Aznauryan, M.; Barth, A.; Birkedal, V.; Bowen, M. E.; Chen, H.; Cordes, T.; Eilert, T.; Fijen, C.; Gebhardt, C.; Gotz, M.; Gouridis, G.; Gratton, E.; Ha, T.; Hao, P.; Hanke, C. A.; Hartmann, A.; Hendrix, J.; Hildebrandt, L. L.; Hirschfeld, V.; Hohlbein, J.; Hua, B.; Hubner, C. G.; Kallis, E.; Kapanidis, A. N.; Kim, J. Y.; Krainer, G.; Lamb, D. C.; Lee, N. K.; Lemke, E. A.; Levesque, B.; Levitus, M.; McCann, J. J.; Naredi-Rainer, N.; Nettels, D.; Ngo, T.; Qiu, R.; Robb, N. C.; Rocker, C.; Sanabria, H.; Schlierf, M.; Schroder, T.; Schuler, B.; Seidel, H.; Streit, L.; Thurn, J.; Tinnefeld, P.; Tyagi, S.; Vandenberk, N.; Vera, A. M.; Weninger, K. R.; Wunsch, B.; Yanez-Orozco, I. S.; Michaelis, J.; Seidel, C. A. M.; Craggs, T. D.; Hugel, T., Precision and accuracy of single-molecule FRET measurements-a multi-laboratory benchmark study. *Nat Methods* **2018,** *15* (9), 669-676.

30. Meer, B. W. v. d.; Coker, G.; Chen, S. Y. S., *Resonance energy transfer : theory and data*. VCH: New York, 1994; p x, 177 p.

31. Peulen, T. O.; Opanasyuk, O.; Seidel, C. A. M., Combining Graphical and Analytical Methods with Molecular Simulations To Analyze Time-Resolved FRET Measurements of Labeled Macromolecules Accurately. *J Phys Chem B* **2017,** *121* (35), 8211-8241.

32. Becker, W., *Advanced time-correlated single photon counting techniques*. Springer: Berlin ; New York, 2005; p xix, 401 p.

33. Gell, C.; Brockwell, D.; Smith, A., *Handbook of single molecule fluorescence spectroscopy*. Oxford University Press: Oxford ; New York, 2006; p xvi, 262 p.

34. Enderlein, J.; Gregor, I.; Patra, D.; Dertinger, T.; Kaupp, U. B., Performance of fluorescence correlation spectroscopy for measuring diffusion and concentration. *Chemphyschem* **2005,** *6* (11), 2324-36.

35. Hildebrandt, L. L.; Preus, S.; Birkedal, V., Quantitative single molecule FRET efficiencies using TIRF microscopy. *Faraday Discuss* **2015,** *184*, 131-42.

36. Holden, S. J.; Uphoff, S.; Hohlbein, J.; Yadin, D.; Le Reste, L.; Britton, O. J.; Kapanidis, A. N., Defining the limits of single-molecule FRET resolution in TIRF microscopy. *Biophys J* **2010,** *99* (9), 3102-11.

37. Hess, S. T.; Webb, W. W., Focal volume optics and experimental artifacts in confocal fluorescence correlation spectroscopy. *Biophys J* **2002,** *83* (4), 2300-2317.

38. Rigler, R.; Mets, U.; Widengren, J.; Kask, P., Fluorescence Correlation Spectroscopy with High Count Rate and Low-Background - Analysis of Translational Diffusion. *Eur Biophys J Biophy* **1993,** *22* (3), 169-175.

39. Aragon, S. R.; Pecora, R., Fluorescence Correlation Spectroscopy as a Probe of Molecular-Dynamics. *J Chem Phys* **1976,** *64* (4), 1791-1803.

40. Lee, J. S., Digital image enhancement and noise filtering by use of local statistics. *IEEE Trans Pattern Anal Mach Intell* **1980,** *2* (2), 165-8.

41. Enderlein, J.; Robbins, D. L.; Ambrose, W. P.; Goodwin, P. M.; Keller, R. A., Statistics of single-molecule detection. *J Phys Chem B* **1997,** *101* (18), 3626-3632.

42. Nir, E.; Michalet, X.; Hamadani, K. M.; Laurence, T. A.; Neuhauser, D.; Kovchegov, Y.; Weiss, S., Shot-noise limited single-molecule FRET histograms: comparison between theory and experiments. *J Phys Chem B* **2006,** *110* (44), 22103-24.

43. Eggeling, C.; Berger, S.; Brand, L.; Fries, J. R.; Schaffer, J.; Volkmer, A.; Seidel, C. A., Data registration and selective single-molecule analysis using multi-parameter fluorescence detection. *J Biotechnol* **2001,** *86* (3), 163-80.

44. Sindbert, S.; Kalinin, S.; Nguyen, H.; Kienzler, A.; Clima, L.; Bannwarth, W.; Appel, B.; Muller, S.; Seidel, C. A., Accurate distance determination of nucleic acids via Forster resonance energy transfer: implications of dye linker length and rigidity. *J Am Chem Soc* **2011,** *133* (8), 2463-80.

45. Dittrich, J.; Popara, M.; Kubiak, J.; Dimura, M.; Schepers, B.; Verma, N.; Schmitz, B.; Dollinger, P.; Kovacic, F.; Jaeger, K. E.; Seidel, C. A. M.; Peulen, T. O.; Gohlke, H., Resolution of Maximum Entropy Method-Derived Posterior Conformational Ensembles of a Flexible System Probed by FRET and Molecular Dynamics Simulations. *J Chem Theory Comput* **2023,** *19* (8), 2389-2409.

46. Frenkel, D.; Smit, B., *Understanding molecular simulation : from algorithms to applications*. 2nd ed.; Academic Press: San Diego, 2002; p xxii, 638 p.

47. Leach, A. R., *Molecular modelling : principles and applications*. 2nd ed.; Prentice Hall: Harlow, England ; New York, 2001; p xxiv, 744 p., 16 p. of plates.

48. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A., A 2nd Generation Force-Field for the Simulation of Proteins, Nucleic-Acids, and Organic-Molecules. *Journal of the American Chemical Society* **1995,** *117* (19), 5179-5197.

49. Ewald, P. P., Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Ann. Phys.* **1921,** *64*, 253–287.

50. Verlet, L., Computer Experiments on Classical Fluids .I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys Rev* **1967,** *159* (1), 98-+.

51. Robustelli, P.; Piana, S.; Shaw, D. E., Developing a molecular dynamics force field for both folded and disordered protein states. *Proc Natl Acad Sci U S A* **2018,** *115* (21), E4758-E4766.

52. Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R., A Computer-Simulation Method for the Calculation of Equilibrium-Constants for the Formation of Physical Clusters of Molecules - Application to Small Water Clusters. *J Chem Phys* **1982,** *76* (1), 637-649.

53. Hockney, R. W.; Eastwood, J. W., *Computer simulation using particles*. McGraw-Hill International Book Co.: New York, 1981; p xix, 540 p.

54. Kaufman, L.; Rousseeuw, P. J., *Finding groups in data : an introduction to cluster analysis*. Wiley: New York, 1990; p xiv, 342 p.

55. Gordon, A. D., A Review of Hierarchical-Classification. *J Roy Stat Soc a Sta* **1987,** *150*, 119-137.

56. MacQueen, J. In *Some methods for classification and analysis of multivariate observations*, Le Cam, L., Neyman, J., Ed. University of California Press, Berkeley: pp 281-297.

57. Rousseeuw, P. J., Silhouettes - a Graphical Aid to the Interpretation and Validation of Cluster-Analysis. *J Comput Appl Math* **1987,** *20*, 53-65.

58. Johnson, S. C., Hierarchical Clustering Schemes. *Psychometrika* **1967,** *32* (3), 241-254.

59. Sneath, P. H. A., The Application of Computers to Taxonomy. *J Gen Microbiol* **1957,** *17* (1), 201-226.

60. Mcquitty, L. L., Hierarchical Linkage Analysis for the Isolation of Types. *Educ Psychol Meas* **1960,** *20* (1), 55-67.

61. Mcquitty, L. L., Expansion of Similarity Analysis by Reciprocal Pairs for Discrete and Continuous Data. *Educ Psychol Meas* **1967,** *27* (2), 253-&.

62. Brooks-Bartlett, J. C.; Garman, E. F., The Nobel Science: One Hundred Years of Crystallography. *Interdiscipl Sci Rev* **2015,** *40* (3), 244-264.

63. Ilari, A.; Savino, C., Protein structure determination by x-ray crystallography. *Methods Mol Biol* **2008,** *452*, 63-87.

64. McPherson, A., Current approaches to macromolecular crystallization. *Eur J Biochem* **1990,** *189* (1), 1-23.

65. Bragg, W. L. In *The diffraction of short electromagnetic waves by a crystal*, Proceedings of the Cambridge Philosophical Society, 1913a.

66. Hekstra, D. R., Emerging Time-Resolved X-Ray Diffraction Approaches for Protein Dynamics. *Annu Rev Biophys* **2023,** *52*, 255-274.

67. Wüthrich, K., *NMR of proteins and nucleic acids*. Wiley: New York, 1986; p xv, 292.

68. Wurz, J. M.; Kazemi, S.; Schmidt, E.; Bagaria, A.; Guntert, P., NMR-based automated protein structure determination. *Arch Biochem Biophys* **2017,** *628*, 24-32.

69. Berjanskii, M. V.; Wishart, D. S., Unraveling the meaning of chemical shifts in protein NMR. *Bba-Proteins Proteom* **2017,** *1865* (11), 1564-1576.

70. Mielke, S. P.; Krishnan, V. V., Characterization of protein secondary structure from NMR chemical shifts. *Prog Nucl Magn Reson Spectrosc* **2009,** *54* (3-4), 141-165.

71. Saito, H.; Ando, I.; Ramamoorthy, A., Chemical shift tensor - The heart of NMR: Insights into biological aspects of proteins. *Prog Nucl Mag Res Sp* **2010,** *57* (2), 181-228.

72. Callaway, E., The revolution will not be crystallized: a new method sweeps through structural biology. *Nature* **2015,** *525* (7568), 172-4.

73. Callaway, E., The Protein-Imaging Technique Taking over Structural Biology. *Nature* **2020,** *578* (7794), 201-201.

74. Nwanochie, E.; Uversky, V. N., Structure Determination by Single-Particle Cryo-Electron Microscopy: Only the Sky (and Intrinsic Disorder) is the Limit. *Int J Mol Sci* **2019,** *20* (17).

75. Maeots, M. E.; Enchev, R. I., Structural dynamics: review of time-resolved cryo-EM. *Acta Crystallogr D Struct Biol* **2022,** *78* (Pt 8), 927-935.

76. Peplow, M., Cryo-Electron Microscopy Reaches Resolution Milestone. *ACS Cent Sci* **2020,** *6* (8), 1274-1277.

77. Anfinsen, C. B., Principles that govern the folding of protein chains. *Science* **1973,** *181* (4096), 223-30.

78. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Zidek, A.; Potapenko, A.; Bridgland, A.; Meyer, C.; Kohl, S. A. A.; Ballard, A. J.; Cowie, A.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Adler, J.; Back, T.; Petersen, S.; Reiman, D.; Clancy, E.; Zielinski, M.; Steinegger, M.; Pacholska, M.; Berghammer, T.; Bodenstein, S.; Silver, D.; Vinyals, O.; Senior, A. W.; Kavukcuoglu, K.; Kohli, P.; Hassabis, D., Highly accurate protein structure prediction with AlphaFold. *Nature* **2021,** *596* (7873), 583-589.

79. Zemla, A.; Venclovas; Moult, J.; Fidelis, K., Processing and evaluation of predictions in CASP4. *Proteins* **2001,** *Suppl 5*, 13-21.

80. Kryshtafovych, A.; Schwede, T.; Topf, M.; Fidelis, K.; Moult, J., Critical assessment of methods of protein structure prediction (CASP)-Round XIV. *Proteins* **2021,** *89* (12), 1607-1617.

81. Wang, W. K.; Peng, Z. L.; Yang, J. Y., Single-sequence protein structure prediction using supervised transformer protein language models. *Nat Comput Sci* **2022,** *2* (12), 804-+.

82. Lane, T. J., Protein structure prediction has reached the single-structure frontier. *Nat Methods* **2023,** *20* (2), 170-173.

83. Wang, H. W.; Wang, J. W., How cryo-electron microscopy and X-ray crystallography complement each other. *Protein Sci* **2017,** *26* (1), 32-39.

84. Jeffries, C. M.; Ilavsky, J.; Martel, A.; Hinrichs, S.; Meyer, A.; Pedersen, J. S.; Sokolova, A. V.; Svergun, D. I., Small-angle X-ray and neutron scattering. *Nat Rev Method Prime* **2021,** *1* (1).

85. Brosey, C. A.; Tainer, J. A., Evolving SAXS versatility: solution X-ray scattering for macromolecular architecture, functional landscapes, and integrative structural biology. *Curr Opin Struct Biol* **2019,** *58*, 197-213.

86. Mertens, H. D.; Svergun, D. I., Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *J Struct Biol* **2010,** *172* (1), 128-41.

87. Shortle, D.; DiMaio, D.; Nathans, D., Directed mutagenesis. *Annu Rev Genet* **1981,** *15*, 265-94.

88. Englander, S. W.; Kallenbach, N. R., Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q Rev Biophys* **1983,** *16* (4), 521-655.

89. Holding, A. N., XL-MS: Protein cross-linking coupled with mass spectrometry. *Methods* **2015,** *89*, 54-63.

90. Leitner, A.; Faini, M.; Stengel, F.; Aebersold, R., Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines. *Trends Biochem Sci* **2016,** *41* (1), 20-32.

91. Sinz, A.; Arlt, C.; Chorev, D.; Sharon, M., Chemical cross-linking and native mass spectrometry: A fruitful combination for structural biology. *Protein Sci* **2015,** *24* (8), 1193-209.

92. Jeschke, G., DEER Distance Measurements on Proteins. *Annu Rev Phys Chem* **2012,** *63*, 419-446.

93. Orioli, S.; Larsen, A. H.; Bottaro, S.; Lindorff-Larsen, K., How to learn from inconsistencies: Integrating molecular simulations with experimental data. *Prog Mol Biol Transl Sci* **2020,** *170*, 123-176.

94. Berman, H. M.; Adams, P. D.; Bonvin, A. A.; Burley, S. K.; Carragher, B.; Chiu, W.; DiMaio, F.; Ferrin, T. E.; Gabanyi, M. J.; Goddard, T. D.; Griffin, P. R.; Haas, J.; Hanke, C. A.; Hoch, J. C.; Hummer, G.; Kurisu, G.; Lawson, C. L.; Leitner, A.; Markley, J. L.; Meiler, J.; Montelione, G. T.; Phillips, G. N., Jr.; Prisner, T.; Rappsilber, J.; Schriemer, D. C.; Schwede, T.; Seidel, C. A. M.; Strutzenberg, T. S.; Svergun, D. I.; Tajkhorshid, E.; Trewhella, J.; Vallat, B.; Velankar, S.; Vuister, G. W.; Webb, B.; Westbrook, J. D.; White, K. L.; Sali, A., Federating Structural Models and Data: Outcomes from A Workshop on Archiving Integrative Structures. *Structure* **2019,** *27* (12), 1745-1759.

95. Sali, A.; Berman, H. M.; Schwede, T.; Trewhella, J.; Kleywegt, G.; Burley, S. K.; Markley, J.; Nakamura, H.; Adams, P.; Bonvin, A. M.; Chiu, W.; Peraro, M. D.; Di Maio, F.; Ferrin, T. E.; Grunewald, K.; Gutmanas, A.; Henderson, R.; Hummer, G.; Iwasaki, K.; Johnson, G.; Lawson, C. L.; Meiler, J.; Marti-Renom, M. A.; Montelione, G. T.; Nilges, M.; Nussinov, R.; Patwardhan, A.; Rappsilber, J.; Read, R. J.; Saibil, H.; Schroder, G. F.; Schwieters, C. D.; Seidel, C. A.; Svergun, D.; Topf, M.; Ulrich, E. L.; Velankar, S.; Westbrook, J. D., Outcome of the First wwPDB Hybrid/Integrative Methods Task Force Workshop. *Structure* **2015,** *23* (7), 1156-67.

96. Vallat, B.; Webb, B.; Fayazi, M.; Voinea, S.; Tangmunarunkit, H.; Ganesan, S. J.; Lawson, C. L.; Westbrook, J. D.; Kesselman, C.; Sali, A.; Berman, H. M., New system for archiving integrative structures. *Acta Crystallogr D Struct Biol* **2021,** *77* (Pt 12), 1486-1496.

97. Bottaro, S.; Lindorff-Larsen, K., Biophysical experiments and biomolecular simulations: A perfect match? *Science* **2018,** *361* (6400), 355-360.

98. Bonomi, M.; Heller, G. T.; Camilloni, C.; Vendruscolo, M., Principles of protein structural ensemble determination. *Curr Opin Struct Biol* **2017,** *42*, 106-116.

99. Gama Lima Costa, R.; Fushman, D., Reweighting methods for elucidation of conformation ensembles of proteins. *Curr Opin Struct Biol* **2022,** *77*, 102470.

100. Jaynes, E. T., Information Theory and Statistical Mechanics. *Phys Rev* **1957,** *106* (4), 620-630.

101. Shannon, C. E., A Mathematical Theory of Communication. *Bell Syst Tech J* **1948,** *27* (4), 623-656.

102. Jaynes, E. T., On the Rationale of Maximum-Entropy Methods. *P Ieee* **1982,** *70* (9), 939-952.

103. Kullback, S.; Leibler, R. A., On Information and Sufficiency. *Ann Math Stat* **1951,** *22* (1), 79-86.

104. Rieping, W.; Habeck, M.; Nilges, M., Inferential structure determination. *Science* **2005,** *309* (5732), 303-6.

105. Kofinger, J.; Stelzl, L. S.; Reuter, K.; Allande, C.; Reichel, K.; Hummer, G., Efficient Ensemble Refinement by Reweighting. *J Chem Theory Comput* **2019,** *15* (5), 3390-3401.

106. Kufareva, I.; Abagyan, R., Methods of protein structure comparison. *Methods Mol Biol* **2012,** *857*, 231-57.

107. Lindorff-Larsen, K.; Ferkinghoff-Borg, J., Similarity measures for protein ensembles. *PLoS One* **2009,** *4* (1), e4203.

108. Tiberti, M.; Papaleo, E.; Bengtsen, T.; Boomsma, W.; Lindorff-Larsen, K., ENCORE: Software for Quantitative Ensemble Comparison. *PLoS Comput Biol* **2015,** *11* (10), e1004415.

109. Yang, S.; Salmon, L.; Al-Hashimi, H. M., Measuring similarity between dynamic ensembles of biomolecules. *Nat Methods* **2014,** *11* (5), 552-4.

110. Viswanath, S.; Sali, A., Optimizing model representation for integrative structure determination of macromolecular assemblies. *Proc Natl Acad Sci U S A* **2019,** *116* (2), 540-545.

111. Lin, J. H., Divergence Measures Based on the Shannon Entropy. *Ieee T Inform Theory* **1991,** *37* (1), 145-151.

112. Sastry, G. M.; Dixon, S. L.; Sherman, W., Rapid shape-based ligand alignment and virtual screening method based on atom/feature-pair similarities and volume overlap scoring. *J Chem Inf Model* **2011,** *51* (10), 2455-66.

113. Ma, B.; Tsai, C. J.; Haliloglu, T.; Nussinov, R., Dynamic allostery: linkers are not merely flexible. *Structure* **2011,** *19* (7), 907-17.

114. Papaleo, E.; Saladino, G.; Lambrughi, M.; Lindorff-Larsen, K.; Gervasio, F. L.; Nussinov, R., The Role of Protein Loops and Linkers in Conformational Dynamics and Allostery. *Chem Rev* **2016,** *116* (11), 6391-423.

115. Pauwels, K.; Lustig, A.; Wyns, L.; Tommassen, J.; Savvides, S. N.; Van Gelder, P., Structure of a membrane-based steric chaperone in complex with its lipase substrate. *Nat Struct Mol Biol* **2006,** *13* (4), 374-5.

116. Higgins, C. F., ABC transporters: from microorganisms to man. *Annu Rev Cell Biol* **1992,** *8*, 67-113.

117. Berntsson, R. P.; Smits, S. H.; Schmitt, L.; Slotboom, D. J.; Poolman, B., A structural classification of substrate-binding proteins. *FEBS Lett* **2010,** *584* (12), 2606-17.

118. Kim, E.; Lee, S.; Jeon, A.; Choi, J. M.; Lee, H. S.; Hohng, S.; Kim, H. S., A single-molecule dissection of ligand binding to a protein with intrinsic dynamics. *Nat Chem Biol* **2013,** *9* (5), 313-8.

119. Berget, S. M.; Moore, C.; Sharp, P. A., Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci U S A* **1977,** *74* (8), 3171-5.

120. Matera, A. G.; Wang, Z., A day in the life of the spliceosome. *Nat Rev Mol Cell Biol* **2014,** *15* (2), 108-21.

121. Kielkopf, C. L.; Rodionova, N. A.; Green, M. R.; Burley, S. K., A novel peptide recognition mode revealed by the X-ray structure of a core U2AF35/U2AF65 heterodimer. *Cell* **2001,** *106* (5), 595-605.

122. Sickmier, E. A.; Frato, K. E.; Shen, H.; Paranawithana, S. R.; Green, M. R.; Kielkopf, C. L., Structural basis for polypyrimidine tract recognition by the essential pre-mRNA splicing factor U2AF65. *Mol Cell* **2006,** *23* (1), 49-59.

123. Berglund, J. A.; Abovich, N.; Rosbash, M., A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes Dev* **1998,** *12* (6), 858-67.

124. Voith von Voithenberg, L.; Sanchez-Rico, C.; Kang, H. S.; Madl, T.; Zanier, K.; Barth, A.; Warner, L. R.; Sattler, M.; Lamb, D. C., Recognition of the 3' splice site RNA by the U2AF heterodimer involves a dynamic population shift. *Proc Natl Acad Sci U S A* **2016,** *113* (46), E7169-E7175.

125. Ito, T.; Muto, Y.; Green, M. R.; Yokoyama, S., Solution structures of the first and second RNA-binding domains of human U2 small nuclear ribonucleoprotein particle auxiliary factor (U2AF(65)). *EMBO J* **1999,** *18* (16), 4523-34.

126. Huang, J. R.; Warner, L. R.; Sanchez, C.; Gabel, F.; Madl, T.; Mackereth, C. D.; Sattler, M.; Blackledge, M., Transient electrostatic interactions dominate the conformational equilibrium sampled by multidomain splicing factor U2AF65: a combined NMR and SAXS study. *J Am Chem Soc* **2014,** *136* (19), 7068-76.

127. Singh, R.; Valcarcel, J.; Green, M. R., Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. *Science* **1995,** *268* (5214), 1173-6.

128. Mackereth, C. D.; Madl, T.; Bonnal, S.; Simon, B.; Zanier, K.; Gasch, A.; Rybin, V.; Valcarcel, J.; Sattler, M., Multi-domain conformational selection underlies pre-mRNA splicing regulation by U2AF. *Nature* **2011,** *475* (7356), 408-U174.

129. Agam, G.; Gebhardt, C.; Popara, M.; Machtel, R.; Folz, J.; Ambrose, B.; Chamachi, N.; Chung, S. Y.; Craggs, T. D.; de Boer, M.; Grohmann, D.; Ha, T.; Hartmann, A.; Hendrix, J.; Hirschfeld, V.; Hubner, C. G.; Hugel, T.; Kammerer, D.; Kang, H. S.; Kapanidis, A. N.; Krainer, G.; Kramm, K.; Lemke, E. A.; Lerner, E.; Margeat, E.; Martens, K.; Michaelis, J.; Mitra, J.; Moya Munoz, G. G.; Quast, R. B.; Robb, N. C.; Sattler, M.; Schlierf, M.; Schneider, J.; Schroder, T.; Sefer, A.; Tan, P. S.; Thurn, J.; Tinnefeld, P.; van Noort, J.; Weiss, S.; Wendler, N.; Zijlstra, N.; Barth, A.; Seidel, C. A. M.; Lamb, D. C.; Cordes, T., Reliability and accuracy of single-molecule FRET studies for characterization of structural dynamics and distances in proteins. *Nat Methods* **2023,** *20* (4), 523-535.

130. Sanchez-Rico, C.; Voith von Voithenberg, L.; Warner, L.; Lamb, D. C.; Sattler, M., Effects of Fluorophore Attachment on Protein Conformation and Dynamics Studied by spFRET and NMR Spectroscopy. *Chemistry* **2017,** *23* (57), 14267-14277.

131. Pauwels, K.; Van Molle, I.; Tommassen, J.; Van Gelder, P., Chaperoning Anfinsen: the steric foldases. *Mol Microbiol* **2007,** *64* (4), 917-22.

132. Jaeger, K. E.; Ransac, S.; Dijkstra, B. W.; Colson, C.; van Heuvel, M.; Misset, O., Bacterial lipases. *FEMS Microbiol Rev* **1994,** *15* (1), 29-63.

133. Jaeger, K. E.; Reetz, M. T., Microbial lipases form versatile tools for biotechnology. *Trends Biotechnol* **1998,** *16* (9), 396-403.

134. El Khattabi, M.; Van Gelder, P.; Bitter, W.; Tommassen, J., Role of the lipase-specific foldase of Burkholderia glumae as a steric chaperone. *J Biol Chem* **2000,** *275* (35), 26885-91.

135. El Khattabi, M.; Ockhuijsen, C.; Bitter, W.; Jaeger, K. E.; Tommassen, J., Specificity of the lipase-specific foldases of gram-negative bacteria and the role of the membrane anchor. *Mol Gen Genet* **1999,** *261* (4-5), 770-6.

136. Viegas, A.; Dollinger, P.; Verma, N.; Kubiak, J.; Viennet, T.; Seidel, C. A. M.; Gohlke, H.; Etzkorn, M.; Kovacic, F.; Jaeger, K. E., Structural and dynamic insights revealing how lipase binding domain MD1 of Pseudomonas aeruginosa foldase affects lipase activation. *Sci Rep* **2020,** *10* (1), 3578.

137. Rosenau, F.; Tommassen, J.; Jaeger, K. E., Lipase-specific foldases. *Chembiochem* **2004,** *5* (2), 152-61.

138. Shibata, H.; Kato, H.; Oda, J., Random mutagenesis on the Pseudomonas lipase activator protein, LipB: exploring amino acid residues required for its function. *Protein Eng* **1998,** *11* (6), 467-72.

139. Kalinin, S., Fulle, S., Hanke, C.A., Peulen, T. O., Sindbert, S., Felekyan, S., Kühnemuth, R., Gohlke, H., Seidel, C. A. M., Diffusion with traps: experiment, simulation, and theory to describe the dynamics of flexibly linked fluorophores in biomolecular FRET. *In preparation*.

140. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J Chem Theory Comput* **2015,** *11* (8), 3696-713.

141. Tian, C.; Kasavajhala, K.; Belfon, K. A. A.; Raguette, L.; Huang, H.; Migues, A. N.; Bickel, J.; Wang, Y.; Pincay, J.; Wu, Q.; Simmerling, C., ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J Chem Theory Comput* **2020,** *16* (1), 528-552.

142. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C., Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **2006,** *65* (3), 712-25.

143. Song, D.; Luo, R.; Chen, H. F., The IDP-Specific Force Field ff14IDPSFF Improves the Conformer Sampling of Intrinsically Disordered Proteins. *J Chem Inf Model* **2017,** *57* (5), 1166-1178.

144. Kalinin, S.; Peulen, T.; Sindbert, S.; Rothwell, P. J.; Berger, S.; Restle, T.; Goody, R. S.; Gohlke, H.; Seidel, C. A., A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nat Methods* **2012,** *9* (12), 1218-25.

145. Dimura, M.; Peulen, T. O.; Sanabria, H.; Rodnin, D.; Hemmen, K.; Hanke, C. A.; Seidel, C. A. M.; Gohlke, H., Automated and optimally FRET-assisted structural modeling. *Nat Commun* **2020,** *11* (1), 5394.

146. Ravera, E.; Sgheri, L.; Parigi, G.; Luchinat, C., A critical assessment of methods to recover information from averaged data. *Phys Chem Chem Phys* **2016,** *18* (8), 5686-701.

147. Gomes, G. W.; Namini, A.; Gradinaru, C. C., Integrative Conformational Ensembles of Sic1 Using Different Initial Pools and Optimization Methods. *Front Mol Biosci* **2022,** *9*, 910956.

148. Thomasen, F. E.; Lindorff-Larsen, K., Conformational ensembles of intrinsically disordered proteins and flexible multidomain proteins. *Biochem Soc Trans* **2022,** *50* (1), 541-554.

149. Harauz, G.; Vanheel, M., Exact Filters for General Geometry 3-Dimensional Reconstruction. *Optik* **1986,** *73* (4), 146-156.

150. Rosenthal, P. B.; Henderson, R., Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J Mol Biol* **2003,** *333* (4), 721-45.

151. Lee, N. K.; Kapanidis, A. N.; Wang, Y.; Michalet, X.; Mukhopadhyay, J.; Ebright, R. H.; Weiss, S., Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. *Biophys J* **2005,** *88* (4), 2939-53.

152. Vandenberk, N.; Barth, A.; Borrenberghs, D.; Hofkens, J.; Hendrix, J., Evaluation of Blue and Far-Red Dye Pairs in Single-Molecule Forster Resonance Energy Transfer Experiments. *J Phys Chem B* **2018,** *122* (15), 4249-4266.

153. Coullomb, A.; Bidan, C. M.; Qian, C.; Wehnekamp, F.; Oddou, C.; Albiges-Rizo, C.; Lamb, D. C.; Dupont, A., QuanTI-FRET: a framework for quantitative FRET measurements in living cells. *Sci Rep* **2020,** *10* (1), 6504.

154. Mockel, C.; Kubiak, J.; Schillinger, O.; Kuhnemuth, R.; Della Corte, D.; Schroder, G. F.; Willbold, D.; Strodel, B.; Seidel, C. A. M.; Neudecker, P., Integrated NMR, Fluorescence, and Molecular Dynamics Benchmark Study of Protein Mechanics and Hydrodynamics. *J Phys Chem B* **2019,** *123* (7), 1453-1480.

List of Abbreviations

ABC	adenosine triphosphate binding cassette
ADC	analog-to-digital converter
APD	avalanche photodiode
APBS	all-photon burst search
ATP	adenosine triphosphate
ALEX-2CDE	ALEX-2-channel kernel-based density estimator function
BVA	burst variance analysis
CD	circular dichroism
CCC	cross correlation coefficient
CASP	Critical Assessment of Structure Prediction
cryo-EM	cryogenic electron microscopy
DCBS	dual-channel burst search
EM	electron microscopy
EPR	electron paramagnetic resonance
FCS	fluorescence correlation spectroscopy
FRET	Förster resonance energy transfer
FSC	Fourier shell correlation
FWHM	full width at half maximum
GDT	global distance test
HDX	hydrogen-deuterium exchange
Lif	lipase-specific foldase
MD	molecular dynamics
MalE	maltose-binding protein
MAP	maximum *a posteriori*
MEM	maximum entropy method
MFD	multiparameter fluorescence detection
NA	numerical aperture
NMR	nuclear magnetic resonance
PDA	photon distribution analysis
PDB	Protein Data Bank
PIE	pulsed interleaved excitation

PSF	point spread function
RMSD	root-mean-square deviation
RRM	RNA recognition motif
SAXS	small angle X-ray scattering
smFRET	single-molecule FRET
SNR	signal-to-noise ratio
TAC	time-to-amplitude converter
TCSPC	time-correlated single photon counting
TIRFM	total internal reflection fluorescence microscopy
U2AF	U2 auxiliary factor
XL-MS	cross-linking mass-spectrometry

Acknowledgements

First and foremost, I would like to express deepest gratitude to my parents, who all the way supported my scientific curiosity and enabled me access to the best educational facilities, as well for supporting my international endeavors.

Next, I would like to thank my supervisor Prof. Dr. Claus A. M. Seidel, for giving me the opportunity to conduct doctoral research under his mentoring, as well for always ensuring comfortable working environment, with all financial and infrastructure resources needed. I am also very grateful for introducing me to relevant and productive collaborations.

I am also grateful to my mentor Prof. Dr. Holger Gohlke, for all the valuable insights into the world of computational biophysics, as well for very successful collaboration. I am also thankful for the access to computer cluster and any support I needed in that respect.

Furthermore, I thank to Suren Felekyan for all the support with experiments, setups and data analysis, as well to Christian Hanke, for assistance with MD simulations, for proof-reading of this thesis, and for help with any technical and scientific challenge I encountered. I am so grateful for constant support of both of you, even when you were not part of the project.

From the FRET community paper, I am very thankful to Prof. Don C. Lamb, for making sure that we are always on the right path, and for being so kind, patient and charismatic. Anders Barth, I cannot emphasize how much I learnt from you, and how much you inspired me through our joint work. Ganesh Agam, I thank you for always being responsive and in positive mood, no matter how many times you had to send us samples. From the integrative modelling project, I would like to acknowledge Mykola Dimura for educating me on the concepts from statistics and computational modelling, Thomas Peulen for all the programming support and for giving me examples of efficient and creative coding, Jonas Dittrich for being effective and reliable collaborator, and last but not least Jakub Kubiak for being welcoming and friendly towards me since the day one of my PhD, for many discussions we had and for all the support beyond science.

Special gratitude also goes to Bianca Reschke for proof-reading of this thesis, Veronika Mendorf, Bekir Bulat and Matthias Fröscher for their administrative and technical support, Martin Schramm for being so kind to fix my bike several times, my office mate Aiswaria Prakash for always patiently listening to me, as well to all the other current and former members of the lab.

Last but not least, I want to thank Dragana Sretenovic and Julian Folz for the friendship we developed, and for many good memories we had during the years.

Supplement A. Reliability and accuracy of single-molecule FRET studies for characterization of structural dynamics and distances in proteins

nature methods

Analysis

https://doi.org/10.1038/s41592-023-01807-0

Reliability and accuracy of single-molecule FRET studies for characterization of structural dynamics and distances in proteins

Received: 2 August 2022

Accepted: 31 January 2023

Published online: 27 March 2023

Check for updates

A list of authors and their affiliations appears at the end of the paper

Single-molecule Förster-resonance energy transfer (smFRET) experiments allow the study of biomolecular structure and dynamics in vitro and in vivo. We performed an international blind study involving 19 laboratories to assess the uncertainty of FRET experiments for proteins with respect to the measured FRET efficiency histograms, determination of distances, and the detection and quantification of structural dynamics. Using two protein systems with distinct conformational changes and dynamics, we obtained an uncertainty of the FRET efficiency ≤0.06, corresponding to an interdye distance precision of ≤2 Å and accuracy of ≤5 Å. We further discuss the limits for detecting fluctuations in this distance range and how to identify dye perturbations. Our work demonstrates the ability of smFRET experiments to simultaneously measure distances and avoid the averaging of conformational dynamics for realistic protein systems, highlighting its importance in the expanding toolbox of integrative structural biology.

Förster-resonance energy transfer (FRET) studies have become a widely used approach to complement classical structural biology techniques[1–4]. They provide information on the structure and conformational heterogeneity of biomolecules over a distance range of 30 to 120 Å and, when performed on single molecules, contribute additional information regarding conformational dynamics on the timescales of nanoseconds to seconds[1,2,5–10]. They also allow for quantitative assessment of structural dynamics and heterogeneity of conformational ensembles. This information is not easily accessible by X-ray crystallography, cryogenic-electron microscopy or cross-linking mass-spectrometry, which provide structural information of solution structures but lack temporal information. FRET can also be used to resolve (parts of) structures in an integrative manner (refs. 11–17) and has the unique ability to provide correlated information on structure and dynamics[1,2].

Hellenkamp et al. presented a quantitative multilaboratory smFRET blind study assessing the validity of using smFRET for structural measurements. This study used static double-stranded DNA (dsDNA) that demonstrated a high reproducibility between the different laboratories with an uncertainty of ≤6 Å for the FRET-derived distances[18]. These results strongly supported the idea that standardized smFRET measurements in combination with standardized data analysis routines are a useful addition to the integrative modeling of static biomolecular structures[12,19,20].

Here, we assessed whether the established procedures translate to more flexible biomacromolecules such as proteins that undergo conformational changes. Compared to dsDNA, proteins are more challenging systems, because the local environments and flexibility of the tethered dyes can vary considerably. Site-specific dye labeling of proteins usually requires the introduction of point mutations (for example, cysteines or nonnatural amino acids), which can affect its structure and function[1]. Moreover, proteins require careful handling and storage due to sample instability and aggregation, and are sensitive to experimental conditions, buffer composition, pH, temperature, surface interactions and so on. In a blind comparison study involving

e-mail: a.barth@tudelft.nl; cseidel@hhu.de; d.lamb@lmu.de; cordes@bio.lmu.de

19 laboratories using diffusion-based confocal smFRET, we investigated the maltose-binding protein (MalE) and the U2 Auxiliary Factor 2 (U2AF2), which display conformational dynamics on different time and length scales. We addressed two key questions: (1) how consistently can smFRET efficiency histograms (and the derived distances) be determined by different laboratories for protein samples prepared with stochastic fluorophore labeling? (2) How reliably can smFRET measurements detect structural dynamics in these proteins and what are the minimal structural fluctuations detectible?

Our study confirmed the reproducibility of accurate FRET efficiency histograms and the ability of smFRET to detect and quantify conformational dynamics on the submillisecond timescale. We demonstrate reproducible FRET efficiency values with uncertainties ≤0.06 corresponding to a distance precision of ≤2 Å and an accuracy ≤5 Å in MalE. Moreover, we compare the variability of setup-dependent parameters and identified the main sources of calibration uncertainty. To push the detection limits for structural dynamics, we refined established experimental and data analysis procedures and studied distinct dye pairs to identify and eliminate dye-specific effects. With this, we could detect distance fluctuations on the order of 5 Å in the FRET-sensitive range. Our work demonstrates that smFRET is able to characterize challenging and realistic protein systems with conformational dynamics on timescales from nanoseconds to seconds, highlighting its importance in the expanding toolbox of integrative structural biology[19–21].

Results

We chose two protein systems with conformational dynamics on different timescales. Our first target was the MalE protein of *Escherichia coli*, the periplasmic component of the ATP binding cassette transporter $MalFGK_2$-E (refs. 22–24). MalE exhibits a typical periplasmic-binding protein fold[25,26] composed of two rigid domains connected by a flexible two-segment ß-stranded hinge (Fig. 1a). This structure enables an allosterically driven motion from an open to closed state upon maltose binding on the subsecond timescale (Supplementary Fig. 1). As a second system, we chose the large subunit of U2AF2 from the pre-messenger RNA (mRNA) splicing machinery[27]. Its two RNA recognition motif domains (RRM1,2) are connected by a long flexible linker and bind single-stranded Py-tract RNA[28]. For U2AF2, the two domains fluctuate between an ensemble of detached conformations and a compact conformation in the apo state[29], whereas ligand binding stabilizes an open conformation (Fig. 2a)[30].

SmFRET experiments were blindly performed by 19 laboratories for MalE and by seven laboratories for U2AF2 using different implementations of diffusion-based confocal spectroscopy with alternating excitation, that is, microsecond-ALEX (alternating laser excitation mode)[31] for intensity-based analysis and nsALEX[32] or pulsed-interleaved excitation (PIE)[33] for intensity- and lifetime-based analyses (Supplementary Fig. 2). To avoid additional complexity and to restrict any preknowledge regarding the samples, the proteins were labeled and checked for functionality before being delivered to the participants. Information regarding the identity of the proteins and ligands, labeling positions, labeling efficiency, and expected FRET efficiencies and changes were not provided. The laboratories were informed about which fluorophores were coupled. We adapted a data analysis routine similar to ref. 18 to determine setup-independent accurate FRET efficiency E values from the photon counts detected in the donor (D) and acceptor (A) detection channels during a single-molecule event. The procedure is described in the Methods and includes subtraction of background signals from all channels and the determination of four correction factors: (α) for spectral crosstalk of D fluorescence into the A channel, (β) for normalization of direct D and A excitation fluxes, (γ) for differences in D and A quantum yields and detection efficiencies and (δ) for the ratio of indirect and direct A excitation (Supplementary Fig. 3 and Supplementary Tables 1 and 2)[34]. The use of ALEX or PIE (Supplementary Note 1) was crucial for corrections of the photon counts to reflect the actual D and A signal and exclusion of single-molecule events from incompletely labeled molecules or ones showing photo- blinking and bleaching[8,9,18,34,35].

MalE

We prepared three double-cysteine variants of MalE with interresidue distances that cover a large part of the dynamic range of FRET (Fig. 1b, Methods and Supplementary Fig. 4). The variants were designed to show a decrease (MalE-1, K29C-S352C), an increase (MalE-2, D87C-A186C) or an unaltered interdye distance (MalE-3, A134C-A186C) upon maltose binding. All variants of MalE were stochastically labeled in one of the laboratories at the given positions with the donor Alexa Fluor 546 (Alexa546) and acceptor Alexa Fluor 647 (Alexa647). Before shipment, we confirmed the functionality of the labeled protein by ligand titrations using smFRET and microscale thermophoresis, and verified that maltose did not affect the dye properties (Supplementary Figs. 5 and 6). To allow a comparison, participants were asked to provide mean FRET efficiencies using Gaussian fits for apo and holo FRET efficiency histograms (Fig. 1c) and to determine a global γ value for all measurement conditions (Supplementary Note 2 and Supplementary Fig. 3). For this workflow, participants used custom or publicly available software packages.

FRET efficiency histograms for representative experiments on MalE in the apo (no ligand) and the holo state (1 mM maltose) are shown in Fig. 1c with mean values reported by 16 laboratories. They show very good agreement and reproducibility. It was not possible to extract accurate FRET efficiency values from three laboratories due to, for example, missing or suboptimal laser lines (Supplementary Table 1 and Supplementary Note 2). All laboratories observed the expected changes for MalE-1, MalE-2 and no shift for MalE-3. This indicates that the samples did not degrade during shipment on dry ice and storage in the laboratories at 4 °C. MalE-1 showed an average FRET efficiency of 0.49 ± 0.06 in the apo state that increased to 0.67 ± 0.05 in the holo state. MalE-2 showed the expected decrease in FRET efficiency from 0.83 ± 0.03 to 0.71 ± 0.05 in the apo and holo states, respectively (Fig. 1c). MalE-3, with both labels on one lobe, showed no significant change in FRET efficiency (E_{apo} = 0.91 ± 0.02, E_{holo} = 0.92 ± 0.02).

The standard deviation of the determined mean FRET efficiency over all laboratories was less than ±0.06, similar to the precision found for dsDNA previously[18] (Extended Data Table 1 and Supplementary Table 3). We observe the highest standard deviation for MalE-1 and the lowest values of ±0.02 for MalE-3, which also has the highest FRET efficiency. We observed systematic deviations of the reported FRET efficiency values for the apo and holo states from the mean value. Hence, we analyze the individual FRET efficiency differences, $\langle E_{holo}\rangle - \langle E_{apo}\rangle$, between the apo and holo states for the different laboratories (Fig. 1d). The distributions indeed narrow for all samples by approximately twofold because systematic deviations cancel out ($\sigma_{\langle E_{holo}\rangle-\langle E_{apo}\rangle}$ for MalE-1 ±0.02, MalE-2 ±0.02 and MalE-3 ±0.01: Fig. 1d, Extended Data Table 1 and Supplementary Table 3).

U2AF2

For the second protein, U2AF2, we chose the published double-cysteine variant L187C-G326C of the minimal RRM1,2 construct, where we previously verified that protein function is not affected by labeling (Fig. 2a)[36,37]. The construct was labeled stochastically on the two RRM domains with the dye pair Atto532–Atto643. A subset of seven groups measured the sample. To investigate the consistency of the obtained FRET efficiency histograms, we plotted the smFRET histograms from individual laboratories (Fig. 2b,c, row 1) as well as the average distribution illustrated by the mean and standard deviation (row 2). All groups found a single broad distribution (Fig. 2b, row 1, apo) with an average E = 0.74 ± 0.03 (row 2). In the presence of 5 µM ligand (U9 RNA, K_d of roughly 1.3 µM), a second narrower peak at lower E appears

https://doi.org/10.1038/s41592-023-01807-0

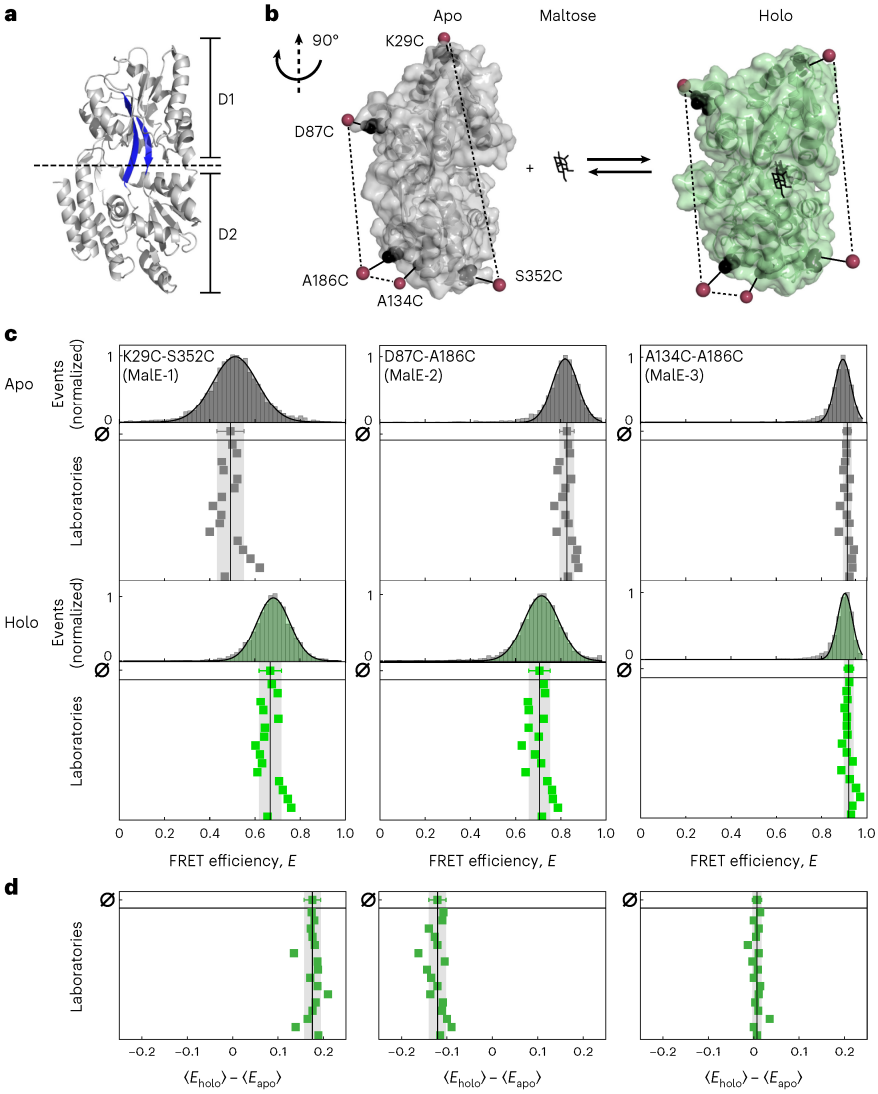


Fig. 1 | Experimental design of MalE as a protein model system for smFRET studies. a, Crystal structure of MalE in its ligand-free apo state (PDB ID 1OMP) with domains D1 and D2 linked by flexible beta sheets (highlighted in blue). **b**, The crystal structure of MalE (rotated by 90° as compared to **a** in the apo (gray, PDB ID 1OMP) and holo (green, PDB ID 1ANF) states with mutations at K29C-S352C (MalE-1), D87C-A186C (MalE-2) and A134C-A186C (MalE-3) indicated in black. Note, each mutant only contains one cysteine pair and was measured using the Alexa546–Alexa647 FRET pair. The estimated mean position of the fluorophores from AV calculations are shown as red spheres. **c**, FRET efficiency *E* histograms for three MalE mutants, MalE-1 (left), MalE-2 (middle) and MalE-3 (right), in the absence and presence of 1 mM maltose (bottom, green) for one exemplary dataset measured in laboratory 1. The distribution is fitted to a Gaussian distribution. The reported mean FRET efficiencies for 16 laboratories are shown below (due to experimental difficulties, the results of three laboratories were excluded; Supplementary Table 1). The mean FRET efficiency and the standard deviation of all 16 laboratories are given by the black line and gray area. **d**, Individual FRET efficiency differences for each laboratory, between the apo and holo states, $\langle E_{holo}\rangle - \langle E_{apo}\rangle$, for MalE-1 (left), MalE-2 (middle) and MalE-3 (right). The mean FRET efficiency difference and the standard deviation of all 16 laboratories are given by the black line and gray area.

(Fig. 2c, row 1) with an average $E = 0.46 \pm 0.04$ (row 2) as expected for the open conformation of the holo state[30,36]. Notably, a fraction of around 15% of ligand-free protein remains in the sample at the RNA concentration used (Supplementary Fig. 7).

For the apo state, we obtained a similar standard deviation of ±0.03 as found for MalE, however, a clear outlier was apparent (Supplementary Table 4). To test whether user bias affected the reported results, a single person reanalyzed the datasets. This person developed an optimal procedure for determining the correction factors for this challenging sample (Supplementary Note 3) and improved the agreement to a standard deviation of ±0.008 with no change in mean *E* (Fig. 2d,e and Supplementary Table 4). The reanalysis revealed the detection correction factor γ to be the main cause of the deviations between the measurements. As a single population of the apo state did not allow for a robust determination of the γ factor[34,35], it was best to estimate the γ factor from a global analysis of the apo and holo measurements. This was possible since the quantum yield of the fluorophores remained unchanged upon RNA binding (Supplementary Table 5). We also reanalyzed data from the same seven laboratories for MalE-1 apo and obtained nearly identical mean FRET efficiencies

https://doi.org/10.1038/s41592-023-01807-0

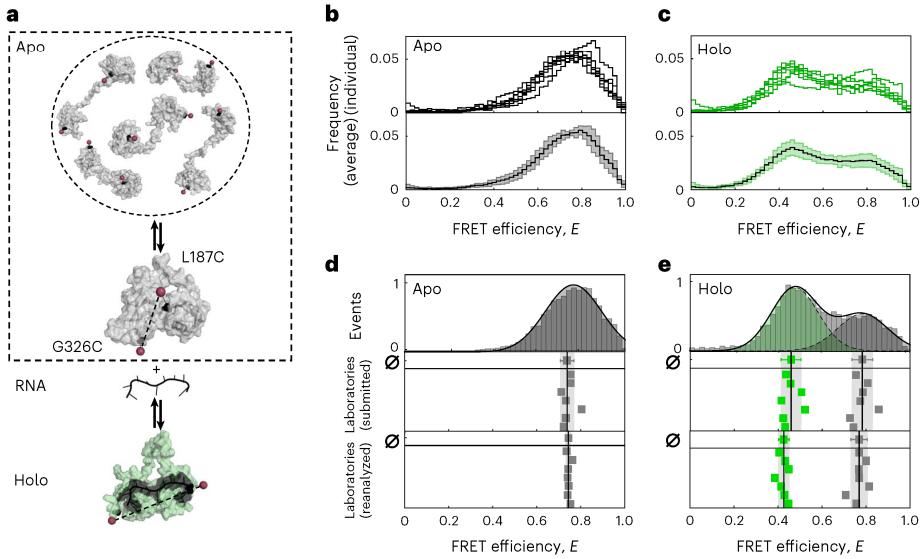


Fig. 2 | The experimental system of U2AF2 (RRM1, 2) and a comparison of FRET efficiency histograms from seven different laboratories. a, Schematic of the dynamics of U2AF2. The apo state (in gray, top) undergoes fast exchange between an ensemble of detached structures of which five representative structures are displayed. A slower exchange occurs between the dynamic detached ensemble and a compact conformation (PDB ID 2YHO) shown below. The holo state (in green, PDB ID 2YH1) bound to a U9 RNA ligand (in dark gray) assumes a well-defined, open conformation. Positions of cysteine mutations introduced for labeling (L187 in RRM1 and G326 in RRM2) are depicted as black spheres with the mean dye position determined by AV calculations indicated by red spheres. **b**,**c**, SmFRET efficiency histograms reported by the seven participating laboratories for apo (**b**) and holo (**c**) measurements of U2AF2. The top shows the individual FRET efficiency histograms and the bottom shows the average FRET efficiency histogram (solid line) with standard deviation (light area). **d**, SmFRET efficiency *E* histograms of U2AF2 in the apo state. The top shows a representative 1D FRET efficiency histogram with a Gaussian fit (laboratory 1). The middle shows the reported mean FRET efficiencies reported by seven laboratories. The mean value from all datasets is 0.739 ± 0.029, shown above with the corresponding standard deviation in gray. The bottom shows the extracted mean FRET values after reanalysis of the collected data. After reanalysis, the agreement improved to 0.742 ± 0.008. **e**, SmFRET efficiency histogram comparisons of U2AF2 in the holo state. 5 µM of U9 RNA was used to obtain the holo state. The top shows a representative 1D FRET efficiency histogram of laboratory 1 fitted to two Gaussian distributions to determine the FRET efficiencies of the different subpopulations, yielding mean FRET efficiencies of 0.44 for RNA-bound and 0.76 for the RNA-free conformation. The middle shows the mean FRET efficiencies reported by the seven laboratories. The mean values from all seven of the datasets were 0.45 ± 0.04 for the RNA-bound conformation (in green) and 0.78 ± 0.04 for the RNA-free conformation (in gray). The bottom shows the reanalysis of the holo measurements yielding values of 0.42 ± 0.02 and 0.77 ± 0.03 for RNA-bound and RNA-free fractions, respectively.

and standard deviations (0.49 ± 0.05 versus 0.47 ± 0.06, Supplementary Fig. 8). This indicates that user bias was less pronounced when a global, well-defined analysis procedure for determining γ was provided over several samples covering a substantial fraction of the FRET range (Supplementary Note 2).

For the holo state of U2AF2, good agreement between laboratories was obtained for the peak positions with a standard deviation of ±0.03 and ±0.02 for the high- and low-FRET peaks, respectively. A minimal improvement resulted from the reanalysis (Supplementary Table 4). In contrast to the agreement in FRET efficiency, we observed variations in the relative amplitudes of the two populations: 0.58 + 0.08 for the holo state and 0.42 ± 0.08 for the apo population (Fig. 2c and Supplementary Table 4). We attribute this to potentially reduced protein activity, degradation of the RNA ligand and sensitivity of conformational dynamics to the experimental conditions, for example, temperature, ligand concentration, buffer composition, salt concentration or the presence of stabilizers such as bovine serum albumin (BSA) (Supplementary Fig. 7).

Setup-dependent parameters and correction factors

The quality of smFRET experiments is determined by the statistics of the measurement and the performance of the setup to maximize photon collection and thereby minimize shot noise. To this end, we quantified the number of bursts, average photon count rate, burst duration and the number of photons in the D and A channels for the MalE measurements from eight laboratories (Fig. 3a and Supplementary Fig. 9). On average, participants collected 6,000 bursts (minimum 500, maximum 21,000) of molecules carrying both fluorophores. The required number of bursts for a smFRET analysis depends on the goal of the experiment. To determine the average FRET efficiency from a single population, as performed for MalE, roughly 1,000 bursts of double-labeled molecules may be sufficient. For advanced analysis methods such as time-correlated single photon counting (TCSPC) for lifetime analysis, burst-wise fluorescence correlation spectroscopy (FCS) or a photon distribution analysis (PDA) that are applied to subensembles, higher burst numbers of >5,000 are desired. Typical count rates per single-molecule event were found to be 60 + 20 kHz, with an average burst of 90 ± 40 photons and 1.7 ± 0.9 ms duration (Fig. 3a and Supplementary Fig. 9). The average count rate and burst duration depend on the size of the confocal volume, where smaller sizes result in higher count rates but shorter burst durations. We observe a negative correlation between burst duration and average count rate (Fig. 3b, Pearson's $r = -0.58$ and Supplementary Fig. 10). The large spread of the burst duration arises from the fact that some participants applied a diffraction-limited observation volume, while others underfilled the objective lens to create a larger confocal volume with a diameter of roughly 1 µm (assuming that the detection pinhole corresponds to the excitation volume). We also observed a small positive correlation between detected photon numbers and burst duration (Fig. 3c, Pearson's $r = 0.54$ and Supplementary Fig. 10). This suggests that larger volumes, in combination with high irradiances,

https://doi.org/10.1038/s41592-023-01807-0

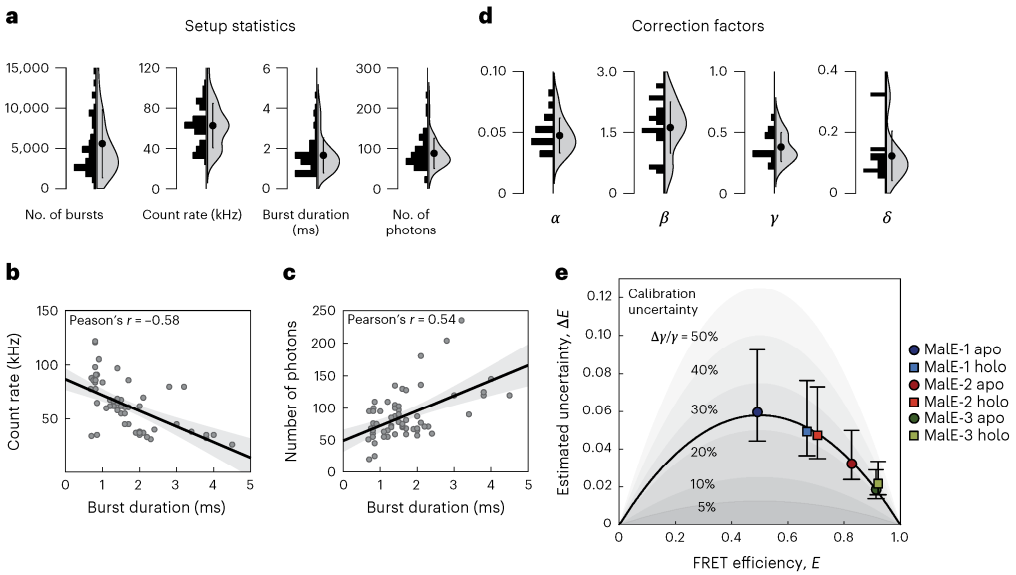


Fig. 3 | Setup-dependent parameters and calibration uncertainty. a, The distribution of the parameters quantifying the statistics of the measurements and the performance of the setups used for both MalE and U2AF2 measurements are shown as histograms and violin plots for the measurements from eight laboratories. The circle and whiskers in the violin plot indicate the mean and standard deviation ($n = 64$, averaged over eight samples measured in the eight different laboratories). Sample-dependent distributions of the shown parameters are given in Supplementary Fig. 9. **b**,**c**, Pairwise plots of the average count rate (**b**) and the number of photons (**c**) against the burst duration. The same datasets are plotted as used for **a**. While the count rate decreases slightly for longer burst durations, a positive correlation is observed for the acquired number of photons per burst and the burst duration, indicating that larger observation volumes result in a higher accumulated signal per molecule. Correlations between all parameters are shown in Supplementary Fig. 10. Error bands indicate the 95% confidence intervals of the regression. **d**, The distributions of the four correction factors for the calculation of accurate FRET efficiencies for all the MalE measurements are shown as histograms and violin plots for the measurements from all laboratories. The circle and whiskers in the violin plot indicate the mean and standard deviation ($n = 64$, averaged over eight samples measured in the eight different laboratories). **e**, A plot of the standard deviation of the reported FRET efficiencies from 16 laboratories (as a measure of the experimental uncertainty) against the average FRET efficiency for the MalE mutants 1–3 reveals that lower uncertainties are observed for higher FRET efficiencies. The black line represents a fit of the estimated uncertainties under the assumption that the variations arise solely due to uncertainty in the γ factor (equation (1)). The inferred relative uncertainty of the γ factor is around 23%. Shaded areas indicate relative uncertainties of 5–50%. Error bars indicate 95% confidence intervals around the average value.

yield the highest number of photons per burst[38]. Smaller volumes generally allow for higher burst collection rates with higher count rates and thus shorter interphoton times, enabling fast transitions on the sub-µs timescale to be resolved[39,40]. Longer burst durations offer the benefit that slower dynamics can be studied.

For an accurate analysis, the correction factors for donor spectral crosstalk (α), excitation flux (β), detection efficiency and quantum yields (γ) and direct acceptor excitation (δ) must be determined (see ref. 18, Supplementary Table 5). We plot the distribution of the correction factors used to determine accurate FRET efficiencies for MalE in Fig. 3d from 16 laboratories (Supplementary Table 1). Besides fluorophore properties, these also depend on setup-specific parameters including dichroic mirrors, emission filters, detectors, excitation wavelengths and laser power. Nonetheless, we observed a well-defined distribution for α of 0.05 ± 0.01, which is determined by the emission filters and detectors in both detection channels. A larger spread was observed for β values of 1.6 ± 0.6 and δ of 0.12 ± 0.08. These depend on the ratio of the excitation powers, where most participants used about half the laser power for direct acceptor excitation (45 ± 27 µW) in comparison to the donor excitation (78 ± 58 µW), resulting in similar count rates after donor and acceptor excitation. The agreement between the reported FRET efficiency values clearly shows that the diverse experimental settings are compensated by the correction procedure applied here.

For γ, which is the most difficult factor to determine, we observed an average of 0.4 ± 0.1 (Supplementary Fig. 3b). It depends on the acceptor-to-donor ratio of the detection efficiencies, g, and the effective fluorescence quantum yields, ϕ_F, as $\gamma = g_A\phi_{F,A}/g_D\phi_{F,D}$ (ref. 18). Similar to crosstalk, γ strongly depends on the emission filters and the type of detectors used. Due to $\phi_{F,A}$ of roughly 0.32 (acceptor) and $\phi_{F,D}$ of roughly 0.72 (donor), all laboratories reported γ factors below 1. Despite the large spread in the reported values, we observed very good agreement for the reported FRET efficiencies in our blind study. Our analysis identified γ as the key factor limiting the consistency, which is supported by the following arguments: (1) in Fig. 1d, the spread of $\langle E_{holo}\rangle - \langle E_{apo}\rangle$ is smaller (for example, 0.06 to 0.02 for MalE-1) than for absolute E values in Fig. 1c, suggesting that errors in E are systematic rather than random. (2) The observed spread in reported FRET efficiencies depends on the absolute FRET efficiency measured for MalE (Fig. 1c,d). (3) We also calculated the uncertainty in the FRET efficiency calculation using error propagation for crosstalk, direct excitation and background correction in the donor and acceptor channels. The reported uncertainty can be attributed mainly to the γ factor (Fig. 3e and Supplementary Note 4) with the error of the γ factor, $\Delta\gamma$, that propagates into an uncertainty in the reported FRET efficiencies, ΔE:

$$\Delta E = E(1-E)\frac{\Delta\gamma}{\gamma} \quad (1)$$

Notably, the observed experimental ΔE is well described by equation (1) (black line in Fig. 3e), yielding a relative uncertainty of

$\Delta\gamma/\gamma = 23\%$ corresponding to $\Delta\gamma \cong 0.07$. The improved agreement between measurements on reanalysis for U2AF2 (Fig. 2d) suggests that the accuracy of the analysis could be improved by standardized procedures for the determination of all correction factors, which differ depending on the number of populations in the measurement and whether the FRET efficiency peak is dynamically averaged (Supplementary Note 2).

Detection and quantification of conformational dynamics

Fluorescence trajectories of immobilized molecules provide access to kinetics on the millisecond to second timescales via a dwell-time analysis (Supplementary Fig. 1)[41–43]. For freely diffusing molecules, millisecond dynamics can be studied in the same fashion when molecules diffuse slowly[44,45]. The detection and quantification of submillisecond conformational dynamics in quickly diffusing molecules (with the maximum timescale limited by the burst duration) is possible via FRET–FCS[44,46,47], filtered-FCS[48,49], burst-variance analysis (BVA)[50], FRET–two-channel kernel-based density distribution estimator[51], dynamic PDA[52], FRET efficiency E versus fluorescence-weighted average donor lifetime $\langle\tau_{D(A)}\rangle_F$ analysis (E–τ plots)[52,53], nanosecond-FCS[54], recurrence analysis of single particles[55], photon-by-photon maximum likelihood approaches[40,56–59] and Monte Carlo diffusion-enhanced photon inference (MC-DEPI)[60]. To assess how consistently dynamics can be detected in smFRET measurements, we asked the participants to evaluate whether the proteins were static or dynamic on the (sub-) millisecond timescale and which method they used to come to this conclusion (Supplementary Table 6).

BVA and E–τ plots are frequently used techniques to visualize FRET dynamics by comparing the measured data to theoretical expectations. BVA detects dynamics by estimating the standard deviation of the FRET efficiency over individual bursts, using a predefined photon window (typically ≥100 µs depending on the molecular brightness). Due to FRET dynamics, the standard deviation of the FRET signal within a burst (red line in Fig. 4a) can be higher than expected from shot noise (black semicircle in Fig. 4a), which becomes visible as a deviation or apparent dynamic shift, ds[50]. In the E–τ plots, the observed FRET efficiency determined via intensity (Fig. 4b) is a species-weighted average and, in the presence of dynamics, the position along the y axis depends on the fraction of time spent in the respective states. The fluorescence lifetime of the donor (Fig. 4b, $\langle\tau_{D(A)}\rangle_F$, x axis) is a photon-weighted average, because only a single lifetime is determined. Hence, it is weighted toward the lifetime of low-FRET states as they emit more donor photons[52,53], shifting the data to the right of the 'static' FRET line. E–τ plots can detect dynamics on the nanosecond to millisecond timescale. Here, we have included an additional correction that considers distance fluctuations of the flexible dye linkers (6 Å) resulting in a slightly curved 'static' FRET line[52,61]. To quantify dynamics between two distinct states, a theoretical 'dynamic' FRET line (red, Fig. 4b) is overlaid. Again, ds is defined as the deviation of the observed data from the theoretical static line (Fig. 4b and Supplementary Note 6). It is important to mention that FRET dynamics, and the related ds, are not always of conformational origin.

MalE exhibits slow ligand-driven dynamics on the subsecond timescale between high- and low-FRET states (Supplementary Fig. 1)[62]. Here, we investigated whether the apo and/or holo states undergo dynamics faster than the timescale of diffusion. Both techniques reveal that MalE exhibits no large FRET-fluctuations on the ms timescale (Fig. 4c,d and Supplementary Fig. 11). Almost all groups confirmed this assessment and only three groups concluded that MalE is dynamic without further justification (Supplementary Table 6). To investigate the presence of potential dynamics in more detail, we determined the ds for a subset of the data (eight laboratories for BVA, Fig. 4e and five for E–τ, Fig. 4f, Supplementary Note 5, and Supplementary Table 7). As a static control, we determined the ds of the dsDNA rulers used in ref. 18 (mean ± 1 s.d. as determined from laboratories 1 and 2) shown in gray in Fig. 4e,f (Supplementary Table 8). The ds did not exceed that of the static reference when using BVA for all MalE mutants. From the E–τ plots, however, ds was higher than for dsDNA, especially for MalE-1. This sample clearly exceeds what is predicted for a static system or even what is predicted for dynamics between the apo and holo states (Fig. 4f, red lines and Supplementary Note 6). Hence, some laboratories categorized MalE as dynamic. The cause of this ds, which must originate from FRET dynamics that are faster than around 100 µs, will be discussed in detail below.

In contrast to MalE, all groups found U2AF2 to be dynamic as was expected for two domains connected by a flexible linker (Fig. 4c–f and Supplementary Table 6). The ligand-free apo state shows pronounced deviations from the behavior for static molecules both in the BVA and E–τ plots, while the RNA-bound holo state shows a notable ds for BVA but not for the E–τ analysis (Fig. 4c–f). It was challenging to assess whether the holo state is truly static or dynamic since it contained a measurable fraction of apo protein, which overlaps with the holo population. Hence, U2AF2 is a challenging test case, yet, dynamics were unambiguously detected in all laboratories demonstrating the reliability of smFRET for investigating dynamic systems.

Accuracy of FRET-derived distances and structural modeling

Accurate FRET efficiencies need to be converted into distances for comparison with structures or to use them as distance constraints in integrative FRET-assisted structural modeling[1,5,7,15,63,64]. SmFRET experiments yield FRET efficiencies as a result of dynamically, nonlinearly averaged distances due to the flexible fluorophore linkers. To assess the accuracy of our measurements, we applied the accessible volumes (AV) approach[5,6,64,65], which uses a coarse-grained dye model to estimate the FRET efficiency averaged model distance $R^{model}_{\langle E\rangle}$ between the two dyes. For this, all possible positions of the fluorophores are averaged, taking into account linker conformations and steric hindrances (Fig. 5a–c, Methods and ref. 6). For AV calculations, we assume fast rotational and slow positional averaging with respect to the fluorescence lifetime. Prediction of measured distances via FRET values based on the flexibility and attachment points of a fluorophore is an area of active research and alternative methods are being developed, for example, rotamer libraries[66] or molecular dynamics simulations[67,68].

The average experimental FRET efficiencies from the individual smFRET histograms $\langle E\rangle$ for MalE (Fig. 2) were used to determine $R_{\langle E\rangle}$ for each laboratory (Extended Data Table 1 and Supplementary Table 3) using the Förster equation (equation (2)):

$$R_{\langle E\rangle} = R_0\left(\frac{1}{\langle E\rangle} - 1\right)^{\frac{1}{6}} \quad (2)$$

The Förster radius of Alexa546–Alexa647 on MalE was determined to be $R_0 = 65 \pm 3$ Å (Supplementary Note 7). Figure 5d displays the correlation between the experimental observable $R_{\langle E\rangle}$ and predicted $R^{model}_{\langle E\rangle}$ using apo and holo structures exhibiting an uncertainty of 3–5 Å over all variants. In agreement with the predictions by Peulen et al.[69], this accuracy is achieved despite stochastic protein labeling, which could result in different charge environments and AVs of the fluorophores depending on the labeling positions. This is evident by the varying dye behavior at different locations (Fig. 5b). During the study, three laboratories studied additional MalE variants (MalE-4, K34C-N205C and MalE-5, T36C-N205C) with a larger FRET efficiency contrast between the apo and holo states, complementing the results of the other variants (Extended Data Table 1).

Figure 5d reveals the largest deviation between experimental and predicted distances for MalE-1, which also had the highest ds values (Fig. 4f and Supplementary Fig. 11). Therefore, we investigated the role of dye–protein interactions using single-cysteine variants of MalE by measuring the fluorescence lifetimes, and time-resolved and steady-state anisotropies (Supplementary Note 8, Supplementary

https://doi.org/10.1038/s41592-023-01807-0

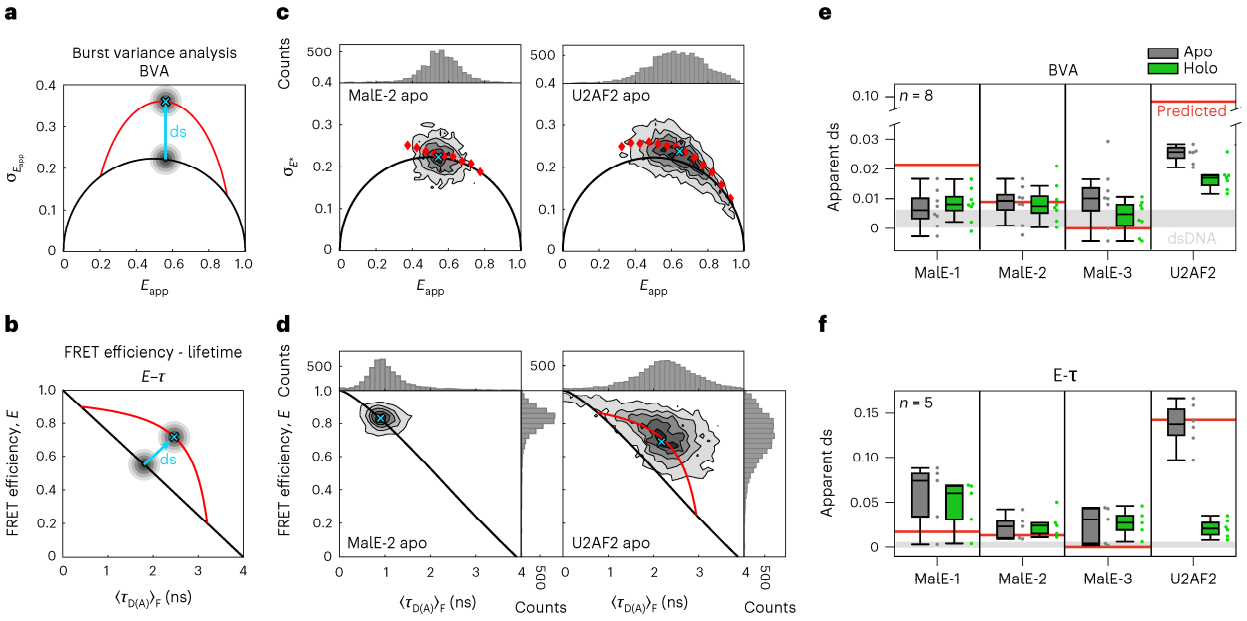


Fig. 4 | Detection and characterization of conformational dynamics on the submillisecond timescale in MalE and U2AF2. a,**b**, Schematic representations of BVA (**a**) and E–τ (**b**) plots. The ds is defined as the excess standard deviation compared to the static line (shown in black). Dynamic FRET lines are indicated in red. **c**, BVA of MalE-2 labeled with Alexa546–Alexa647 without maltose (apo, left) and U2AF2 labeled with Atto532–Atto643 without RNA (apo, right). Here, the BVA is based on a photon binning of five photons. Red diamonds indicate the average standard deviation of all bursts within a FRET efficiency range of 0.05. The mean positions of the populations (cyan crosses) were determined by fitting a two-dimensional Gaussian distribution to the data (Supplementary Note 5). **d**, The plots of the FRET efficiency E versus intensity-weighted average donor lifetime $\langle\tau_{D(A)}\rangle_F$ of the same measurement as in **c**. The donor-only population was excluded from the plot. For MalE-2, the population falls on the static FRET line, while a clear ds is observed for U2AF2. The endpoints of the dynamic FRET line for U2AF2 were determined from a subensemble analysis of the fluorescence decay. **e**,**f**, The apparent ds of the peak of the population was determined graphically from BVA (eight laboratories for MalE and seven laboratories for U2AF2, respectively) (**e**) and E–τ (five laboratories) (**f**) plots (Methods). For U2AF2 in the holo state, the ds was assessed only for the low-FRET RNA-bound population. Boxes indicate the median and 25/75% quartiles of the data. Whiskers extend to the lowest or highest data point within 1.5 times the interquartile range. The gray area indicates the ds obtained for the dsDNA used in a previous study[18] based on measurements performed in laboratory 1 for BVA (ds_{DNA} = 0.0033 ± 0.0033) and laboratory 2 for the E–τ plot (ds_{DNA} = 0.0026 ± 0.0044). The horizontal red lines indicate the expected ds for a potential conformational exchange between the apo and holo states. We computed the expected change in FRET efficiency using their structural models in the PDB (Supplementary Note 6 and Supplementary Table 9).

Tables 5, 11 and 12 and Fig. 5b). Labeling at residue 352 promotes dye sticking to the protein surface indicated by multiexponential fluorescence lifetimes and a high residual anisotropy, r_∞, for both fluorophores ($r_\infty > 0.25$). Labeling at residue 29 only shows sticking for the donor ($r_{\infty,D} > 0.30$, $r_{\infty,A}$ roughly 0.12). At other positions (for example, residue 186), free rotation is possible for both dyes (Supplementary Tables 5 and 11). These position-specific interactions can cause the observed deviations between the experiment and structural model (Fig. 5d and Extended Data Table 1) and high ds values for MalE-1 (Fig. 4f). By using the accessible contact volume (ACV) approach[63], which accounts for dye–protein interactions, the root-mean-average deviation between the structural model and experimental values decreased from 3 Å for AV to 2 Å (Fig. 5c). For protein labeling on opposite sides, dye–protein interactions in the ACV model result in reduced model distances and improved accuracy for all outliers (Fig. 5d and Extended Data Table 1).

It was suggested to use the combined residual anisotropy of D and A ($r_{c,\infty} = \sqrt{r_{\infty,D}\, r_{\infty,A}}$) for filtering out dye-related artifacts in FRET-assisted structural modeling with an empirical threshold of $r_{c,\infty} < 0.2$ (refs. 13,70). To further investigate dye-specific sticking, three laboratories studied MalE mutants with the additional dye pairs Alexa546–AbbSTAR635P, Atto532–Atto643 and Alexa Fluor 488 (Alexa488)–Alexa647 and determined the residual anisotropies and distance uncertainties based on the orientation factor κ^2 (Fig. 5e, top, Supplementary Tables 13 and 14 and Supplementary Notes 8 and 9). The dye pair Alexa546–Alexa647 showed the highest combined anisotropies (Supplementary Fig. 12a and Supplementary Table 13), which is attributed to the donor Alexa546 as the combined anisotropy also remains high for Alexa546–AbbSTAR635P but is reduced for Alexa488–Alexa647. To derive a robust and well-defined threshold for recognizing measurements with dye artifacts, we determined the uncertainty in the FRET-derived distances, $\Delta R_{app}(\kappa^2)$, that originates from the uncertainty of the orientation factor κ^2. Previous approaches estimated the uncertainty in κ^2 from the residual anisotropy in terms of rotational restrictions (wobbling-in-a-cone model)[70–73]. Here, we used a 'diffusion with traps' model, which assumes two dye populations (free and trapped) and relates the residual anisotropies to the fraction of dyes interacting with the surface of the biomolecule (Supplementary Note 9). Based on the estimated distance uncertainty, we propose a threshold of $\Delta R_{app}(\kappa^2) < 10\%$ to identify measurements with dye-related artifacts (Fig. 5e, bottom). This threshold corresponds to a combined residual anisotropy of 0.25, similar to the previously suggested empirical threshold value of around 0.2 (refs. 13,70).

Next, we investigated whether dye sticking could cause the ds in the E–τ plot for MalE-1 with Alexa546–Alexa647 (Fig. 4f). To be observable in the E–τ plot, the exchange between the free and trapped dye species must occur faster than the diffusion time of roughly 1 ms, otherwise the two species would be observable as individual peaks. We observed a correlation between the laboratory-averaged $\langle ds \rangle$ and $\langle r_{c,\infty} \rangle$

https://doi.org/10.1038/s41592-023-01807-0

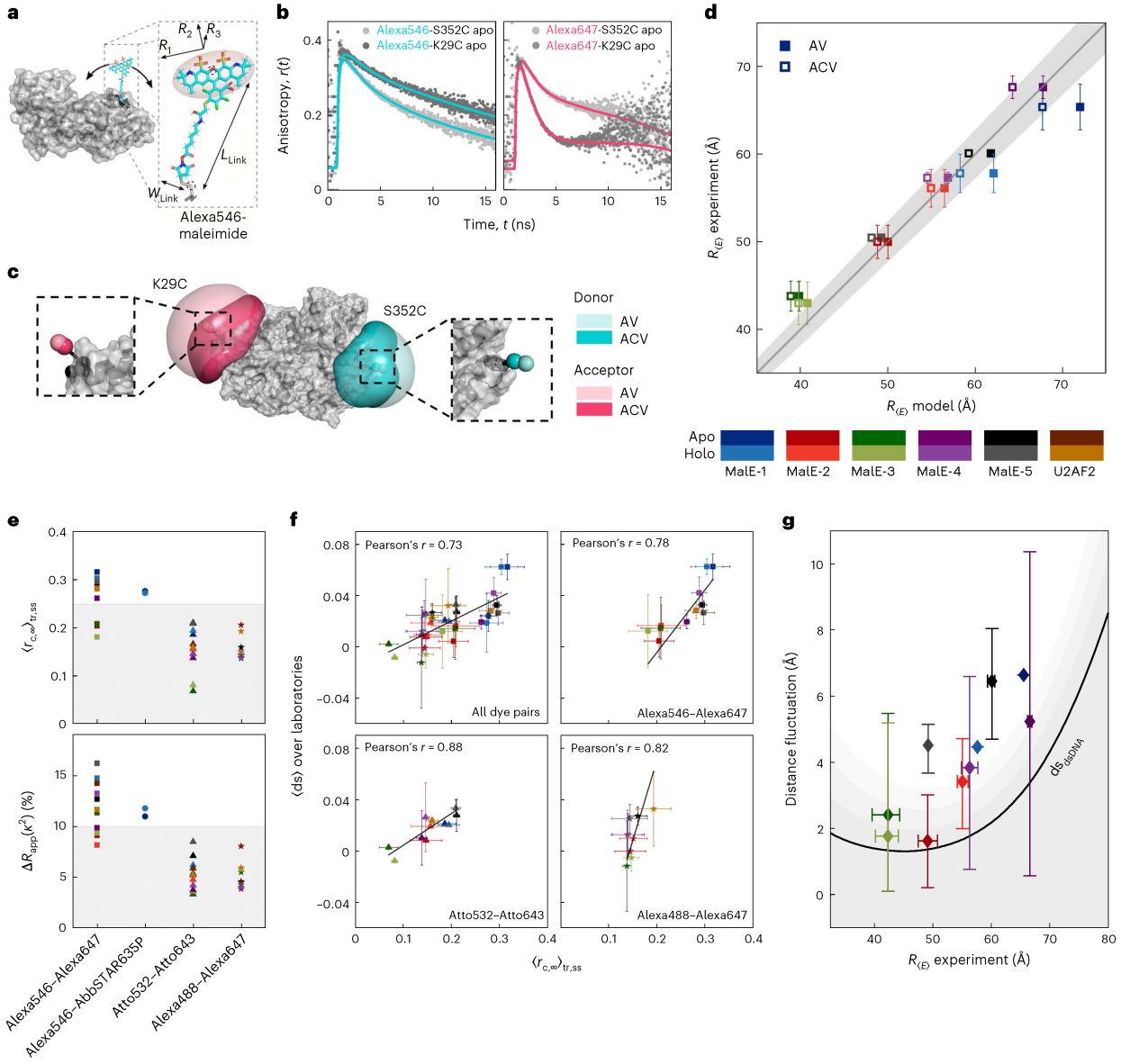


Fig. 5 | Assessing the accuracy of smFRET-derived distances in MalE. a–**d**, AV calculations and model-based interdye distances. **a**, Schematic of Alexa546 attached to MalE (PDB 1OMP) showing the parameters needed for the AV calculations using the AV3 model[6] (Supplementary Table 10). **b**, Fluorescence anisotropy decays of single-cysteine mutants for the donor (Alexa546, left) and acceptor (Alexa647, right) at the labeling positions K29C and S352C. Solid lines represent fits to a model with two or three rotational components (Supplementary Tables 11 and 12 and Supplementary Note 8). **c**, AV (light color) and ACV (dark color) calculations for Alexa546 (cyan) and Alexa647 (pink) at labeling positions 352 and 29. The zoom-ins show the mean positions of the dyes based on the AV (light shade) and ACV (darker shade) models. **d**, Comparison of the experimentally obtained FRET-averaged distance $R_{\langle E \rangle}$ with the theoretical model distances using the AV (filled squares) and ACV (empty squares) calculations. Errors represent the standard deviation in experimental distances ($n = 16$ laboratories for MalE mutants 1–3, $n = 2$ laboratories for MalE mutants 4–5, $n = 7$ laboratories for U2AF). The solid line represents a 1:1 relation and the gray area indicates an uncertainty of ±3 Å for a Förster radius of $R_0 = 65$ Å. MalE-4 and -5 were measured by two laboratories. **e**, Detection of dye-specific protein interactions. Top shows the five MalE mutants and U2AF2 labeled with different dye combinations to determine the donor–acceptor-combined residual anisotropy, $\langle r_{c,\infty} \rangle_{tr,ss}$ ($n = 3$ laboratories). Bottom shows the distance uncertainty relating to κ^2, ΔR_{app} (κ^2), estimated (Supplementary Note 8). A maximum allowed distance uncertainty of ≤10% (shaded gray region) in ΔR_{app} (κ^2) leads to a dye-independent threshold of 0.25 for $\langle r_{c,\infty} \rangle$. **f**, The apparent dynamic shift $\langle ds \rangle$ versus the combined residual anisotropy $\langle r_{c,\infty} \rangle$ is shown for all measured dye pairs (top left) and individually. Error bars of the apparent ds represent the standard deviation over $n = 3$ laboratories. For the combined residual anisotropy, the propagated 1σ uncertainty (Supplementary Note 8). **g**, The structural flexibility of MalE estimated after filtering using the distance uncertainty threshold shown in **e** (Supplementary Note 12). Error bars represent the 1σ percentiles averaged over all dye pairs ($n = 1$, MalE-1; $n = 7$, MalE-2 and MalE-3; $n = 4$, MalE-4 and $n = 5$, MalE-5). The residual distance fluctuations obtained from control measurements on dsDNA in one laboratory ($ds_{dsDNA} = 0.0026 \pm 0.0044$) are shown as a black line (gray areas represent confidence intervals of 1σ, 2σ and 3σ).

https://doi.org/10.1038/s41592-023-01807-0

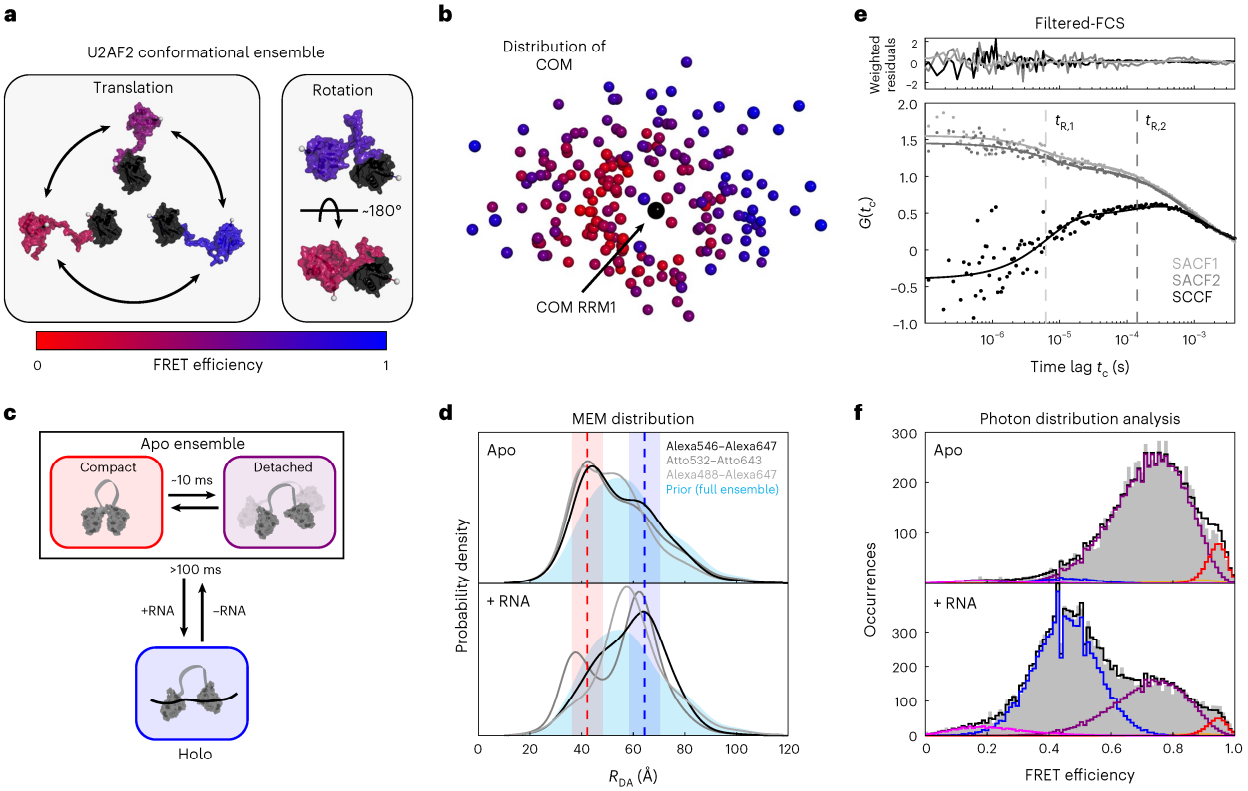


Fig. 6 | Structural characterization of U2AF2. a, Structural flexibility of U2AF2 is given by translational (left) and rotational (right) movement of the two domains. Representative structures are taken from the ensemble determined using NMR and SAXS measurements[29]. **b**, Degeneracy of structural states in FRET measurements. The position of the two domains of U2AF2 is illustrated by the COM of the C_α atoms in RRM2 (residues 260–329, colored) with respect to RRM1 (residues 150–227, black) for the 200 structures of the conformational ensemble[29]. The COM of RRM2 is color-coded according to the FRET efficiency determined using AV3 calculations. **c**, A schematic of the kinetic model used for the global dynamic PDA of U2AF2 (Supplementary Note 17). **d**, Distance distributions obtained from a donor fluorescence decay analysis by a model-free MEM approach (Supplementary Note 15). The distance distribution from the NMR–SAXS ensemble[29] (light blue) was used as the prior distribution. The expected interdye distances for the compact apo and open holo states are shown as red and blue dashed lines (PDB 2YH0 and 2YH1). Shaded areas indicate the distance broadening due to the flexible dye linkers of 6 Å. The distribution in the donor–acceptor distance R_{DA} for different dye pairs is shown. **e**, Filtered-FCS reveals conformational dynamics in the U2AF2 apo ensemble on two timescales, $t_{R,1} = 9 \pm 3$ and $t_{R,2} = 300 \pm 90$ µs, average and standard deviation ($n = 3$, results from laboratory 1 are shown). The two species were defined at the lower and upper edge of the FRET efficiency histogram shown in Fig. 2b, top panel (see Methods and Supplementary Note 16 for details). The species autocorrelation functions (SACFs) and one of the two species cross-correlation functions (SCCFs) are shown. The weighted residuals are shown above. **f**, The PDA analysis was conducted globally over both apo (top) and holo (bottom) measurements using time windows of 0.5, 1.0, 1.5 and 2.0 ms (the 1.0 ms time window histograms are shown). A relaxation time of roughly 10 ms for the dynamics between the detached ensemble and compact apo state with a small amplitude was determined (orange curve) (Supplementary Fig. 16 and Supplementary Note 17).

over all dye pairs (Pearson's r – 0.73), with a stronger correlation when each dye pair is investigated individually (Fig. 5f). As conformational dynamics should be label independent, dye sticking is likely responsible for the observed ds values. The x intercept of the linear fit is between 0.1 and 0.2, suggesting a dye-dependent anisotropy threshold needs to be considered. When applying the criteria $\langle r_{c,\infty} \rangle < 0.25$ to MalE-1 (Supplementary Fig. 12b), only the dye pair Atto532–Atto643 should be used for distance determination, which also showed a markedly reduced ds (Supplementary Fig. 12c). Lifetime analysis of MalE-1 donor-only molecules showed donor quenching only at position 352, which confirms that labeling at this position is problematic (Supplementary Fig. 12c, Supplementary Note 10 and Supplementary Table 5).

Using the above criteria of $\langle r_{c,\infty} \rangle < 0.25$ to minimize the influence of dye artifacts on the ds, we hypothesized that the remaining ds could be indicative of low-amplitude, fast conformational fluctuations. A P test analysis between the ds for dsDNA and protein samples ($P < 0.05$) indicated that the ds is still significant for various protein variants after filtering out dye artifacts (Supplementary Note 11, Supplementary Table 8). To estimate the conformational fluctuations necessary to generate the observed ds (Fig. 4f and Supplementary Table 8), we assume that dynamics occur between two nearby states with interdye distances of $R_{\langle E \rangle} \pm \delta R$ where δR is the amplitude of the fluctuation[61] (Fig. 5g, Supplementary Note 12 and Supplementary Table 8). This inferred fluctuation provides an upper bound for the conformational flexibility because factors such as calibration errors, dye blinking or photoisomerization could contribute to the observed ds. We consider the ds obtained from dsDNA as the lower limit (black line in Fig. 5g, $ds_{DNA} = 0.0026 \pm 0.0044$: Supplementary Note 12), which defines the current detection limit for dynamics in smFRET experiments. The MalE variants 1, 4 and 5 exceed the ds for dsDNA by 2–3 Å (Fig. 5g, Supplementary Fig. 13 and Supplementary Table 8). Consistent with the smFRET results, all-atom molecular dynamics simulations of MalE using the ff14SB force field[74] (Supplementary Note 13) suggest thermally induced conformational fluctuations with a standard

deviation up to roughly 3 Å at the labeled residues in MalE-1, MalE-4 and MalE-5. This is larger than the typical fluctuations of about 1 Å (ref. 75) and leads to a broadening of the interresidue distance distributions for these FRET pairs. We conclude that the observed ds in the experiments can be explained by a combination of measurement uncertainty and small-scale structural fluctuations. Note that such small-scale fluctuations can be amplified in FRET experiments when the dye linker acts as a lever arm for appropriate labeling positions. A detailed discussion of the theoretical limits for detecting dynamics in smFRET experiments using BVA or the E–τ is given in Supplementary Note 14.

Quantitative analysis of U2AF2

The structural characterization of U2AF2 is more complex than for MalE and a simple distance comparison is not possible. Nonetheless, we asked what information smFRET measurements could provide for such a dynamic system. We first surveyed the structural information available on apo U2AF2 from nuclear magnetic resonance (NMR) and small-angle X-ray scattering (SAXS) data[29]. The highly flexible linker allows for a heterogeneous ensemble of U2AF2 conformations (Fig. 6a). To assess how this translates into a smFRET distribution, we quantified the FRET efficiency using AV calculations for all 200 conformers from the NMR–SAXS-derived ensemble of apo U2AF2 (ref. 29). Notably, conformations with similar center-of-mass (COM) distances between the domains showed different FRET efficiencies (Fig. 6a,b), because domain rotations result in distinct interdye distances for identical COM (Fig. 6a, right). Due to this degeneracy, a single-distance probe is insufficient to capture the full structural complexity.

The observed ds in the apo state suggests the presence of conformational dynamics (Fig. 4d–f). To decipher the underlying kinetics and their temporal hierarchy, we applied three analyses. First, we investigated the interdye distance distribution of the apo and holo states from the donor lifetime using a model-free maximum entropy method (MEM) (Fig. 6c,d and Supplementary Note 15)[76]. As a prior, we used the NMR–SAXS structural ensemble. This analysis yielded consistent results for all three dye pairs studied for U2AF2. The MEM analysis revealed peaks in the probability density at the expected distances for the compact apo conformation and RNA-bound holo structure (Fig. 6d, dashed lines). We note that the fluorescence lifetime analysis resolves states on the nanosecond timescale and is therefore less sensitive to dynamic averaging.

Second, to assess the dynamics on the microsecond timescale, three groups performed filtered-FCS and found at least two relaxation times (9 ± 3 and 300 ± 90 µs; Fig. 6e, Supplementary Table 15 and Supplementary Note 16), which were independent of the dyes used (Supplementary Fig. 15). We assign the fast process to dynamics within the detached domains and the slower process to interconversion between compact conformations within the conformational ensemble.

Last, we investigated dynamics on the millisecond timescale using a dynamic PDA. A global analysis of the apo and holo measurements was performed using the kinetic model shown in Fig. 6c (Supplementary Note 17 and Supplementary Table 16). The apo state was treated as a two-state system with slow dynamics between a detached ensemble and a well-defined, compact apo conformation. The rapid dynamics within the detached ensemble is empirically described using a broad, static distribution. For the holo measurement, we account for the residual population of apo molecules. Exchange between the holo and apo states is irrelevant as the binding and dissociation of RNA occurs on timescales of more than 100 ms (ref. 36). This model incorporates all information and is sufficient to describe the smFRET efficiency histograms. The dynamic PDA analysis returned a relaxation time of roughly 10 ms for the dynamics between the detached ensemble and compact apo state (Fig. 6f, orange curve, Supplementary Fig. 16 and Supplementary Table 16). We also determined an interdye distance of $R_{\langle E\rangle} = 61$ Å in the RNA-bound holo state, which is in good agreement with 63 Å from the RNA-bound conformation (Protein Data Bank (PDB) 2YH1).

Discussion

We show that smFRET can provide accurate distances of conformational states and reliable information on conformational dynamics in proteins. Since all experiments were performed using established techniques and analyzed with freely available software[5,6,34,35,77–79], such information is accessible to any group with similar expertise. Despite the challenges of protein samples, we achieved a similar precision in FRET efficiencies as reported for dsDNA[18] (between ±0.02 and ±0.06) (Extended Data Table 1). The reproducibility in excluding large-scale conformational dynamics for MalE on a timescale <10 ms while detecting large-scale submillisecond dynamics in U2AF2 shows that the community can deal with dynamic protein systems. In addition, we could consistently establish the timescales and hierarchy of the exchange dynamics in such a complex protein system as U2AF2. The study of complex dynamics is improvable by probing additional distances[5,13,17,80–83].

The high level of agreement is notable given the diversity of the setups (Fig. 3 and Supplementary Fig. 2) and the number of possible pitfalls. A large contribution to the spread in the reported mean FRET efficiencies was caused by systematic errors in the data analysis. This is supported by a comparison of the FRET efficiency changes ($\langle E_{holo}\rangle - \langle E_{apo}\rangle$) instead of absolute FRET efficiency values (Fig. 1d), which reduced the spread of roughly threefold. Having a single person reanalyze the data led to a similar decrease in the uncertainty of the FRET efficiency for the apo state of U2AF2 (Fig. 2d). Determination of γ was most crucial and the optimal approach depends on the details of the studied system (Supplementary Note 2). In the intensity-based approach of Lee et al.[34], multiple samples with uniform fluorophore properties are required or individual corrections need to be made. When using the approach of Kudryavtsev et al.[35] via E–τ calibration, the system needs to be static and a single population suffices. A protocol with unambiguous instructions for the calibration steps and minimized number of user-dependent steps would enhance the accuracy of FRET measurements.

From accurate FRET efficiencies, we obtained reproducible interdye distances with a precision of 3 Å and an accuracy of 5 Å against structural models of MalE (Extended Data Table 1). This is similar to what was determined for dsDNA samples. This is a very positive outcome, given that dsDNA features a consistent, homogenous chemical environment for each labeling position, in contrast to the variable dye environment experienced in proteins. The distance determination could be improved by including the interaction of the fluorophores with the protein surface using ACV calculations (Fig. 5d and Methods)[63]. Furthermore, we give experimental support (Fig. 5f) for only using dyes with a combined residual anisotropy of $r_{c,\infty} < 0.25$, as suggested previously[13,70]. Proteins often exist within a family of conformations as we observed for U2AF2 (Fig. 6d). Determining how to best deal with distance distributions for conformational ensembles is one of the challenges for structural biology.

Investigating different dye pairs allowed us to reduce dye artifacts, leading to more accurate FRET efficiencies and reliable detection of the dynamics. Hence, we investigated the detection limits for ds and studied its relation to conformational dynamics with a subset of laboratories. Besides conformational motions, dynamic FRET shifts can occur in different directions and have several origins including structural instabilities[37] or photophysics (as shown in Fig. 5f)[44]. Thus, it is advisable to verify the key findings in smFRET measurements with at least two dye pairs and/or with different residue combinations in the protein. Once the non-FRET-dynamic contributions are minimized, we still observed significant residual ds for MalE. Consistent with molecular dynamics simulations (Supplementary Note 13), we interpret these shifts as small-scale conformational dynamics and established a current lower limit for the detection of structural changes via smFRET on the order of ≤5 Å. In summary, the consensus of smFRET experiments on two protein systems exhibiting dynamic behavior on different spatiotemporal scales obtained blindly from 19 laboratories offers strong

support for its use as a robust, versatile and quantitative tool for the coming age of dynamic structural biology. In this context, it will be crucial to integrate the correlated structural and dynamic information provided by smFRET[1] with structural information provided by other experimental techniques as well as artificial intelligence-based protein structural prediction[83]. Considering that protein structure prediction has reached the single-structure frontier[84], the information from smFRET experiments could leverage the power of artificial intelligence to resolve more complex multi-state and ensemble structural models[83]. Vice versa, the power of artificial intelligence and deep learning can be used to increase the throughput for the design and analysis of smFRET experiments[85–87].

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41592-023-01807-0.

References

1. Lerner, E. et al. FRET-based dynamic structural biology: challenges, perspectives and an appeal for open-science practices. *eLife* **10**, e60416 (2021).
2. Lerner, E. et al. Toward dynamic structural biology: two decades of single-molecule Förster resonance energy transfer. *Science* **359**, eaan1133 (2018).
3. Algar, W. R., Hildebrandt, N., Vogel, S. S. & Medintz, I. L. FRET as a biomolecular research tool—understanding its potential while avoiding pitfalls. *Nat. Methods* **16**, 815–829 (2019).
4. Hildebrandt, N. in *FRET—Förster Resonance Energy Transfer* (eds Medintz, I. & Hildebrandt, N.) 105–163 (Wiley, 2013).
5. Muschielok, A. et al. A nano-positioning system for macromolecular structural analysis. *Nat. Methods* **5**, 965–971 (2008).
6. Kalinin, S. et al. A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nat. Methods* **9**, 1218–1225 (2012).
7. Craggs, T. D. & Kapanidis, A. N. Six steps closer to FRET-driven structural biology. *Nat. Methods* **9**, 1157–1159 (2012).
8. Voith von Voithenberg, L. & Lamb, D. C. Single pair Förster resonance energy transfer: a versatile tool to investigate protein conformational dynamics. *BioEssays* **40**, 1700078 (2018).
9. Hohlbein, J., Craggs, T. D. & Cordes, T. Alternating-laser excitation: single-molecule FRET and beyond. *Chem. Soc. Rev.* **43**, 1156–1171 (2014).
10. Krainer, G., Hartmann, A. & Schlierf, M. FarFRET: extending the range in single-molecule FRET experiments beyond 10 nm. *Nano Lett.* **15**, 5826–5829 (2015).
11. Muschielok, A. & Michaelis, J. Application of the nano-positioning system to the analysis of fluorescence resonance energy transfer networks. *J. Phys. Chem. B.* **115**, 11927–11937 (2011).
12. Sali, A. et al. Outcome of the first wwPDB Hybrid/Integrative Methods Task Force Workshop. *Structure* **23**, 1156–1167 (2015).
13. Hellenkamp, B., Wortmann, P., Kandzia, F., Zacharias, M. & Hugel, T. Multidomain structure and correlated dynamics determined by self-consistent FRET networks. *Nat. Methods* **14**, 176–182 (2017).
14. Choi, U. B. et al. Single-molecule FRET-derived model of the synaptotagmin 1-SNARE fusion complex. *Nat. Struct. Mol. Biol.* **17**, 318–324 (2010).
15. Dimura, M. et al. Automated and optimally FRET-assisted structural modeling. *Nat. Commun.* **11**, 5394 (2020).
16. Lerner, E., Ingargiola, A. & Weiss, S. Characterizing highly dynamic conformational states: the transcription bubble in RNAP-promoter open complex as an example. *J. Chem. Phys.* **148**, 123315 (2018).
17. Craggs, T. D. et al. Substrate conformational dynamics facilitate structure-specific recognition of gapped DNA by DNA polymerase. *Nucleic Acids Res.* **47**, 10788–10800 (2019).
18. Hellenkamp, B. et al. Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. *Nat. Methods* **15**, 669–676 (2018).
19. Rout, M. P. & Sali, A. Principles for integrative structural biology studies. *Cell* **177**, 1384–1403 (2019).
20. Sali, A. From integrative structural biology to cell biology. *J. Biol. Chem.* **296**, 100743 (2021).
21. Burley, S. K. et al. PDB-Dev: a prototype system for depositing integrative/hybrid structural models. *Structure* **25**, 1317–1318 (2017).
22. Davidson, A. L., Dassa, E., Orelle, C. & Chen, J. Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol. Mol. Biol. Rev.* **72**, 317–364 (2008).
23. Mächtel, R., Narducci, A., Griffith, D. A., Cordes, T. & Orelle, C. An integrated transport mechanism of the maltose ABC importer. *Res. Microbiol.* **170**, 321–337 (2019).
24. Malik, A. Protein fusion tags for efficient expression and purification of recombinant proteins in the periplasmic space of *E. coli*. *3 Biotech* **6**, 44 (2016).
25. Berntsson, R. P. A., Smits, S. H. J., Schmitt, L., Slotboom, D. J. & Poolman, B. A structural classification of substrate-binding proteins. *FEBS Lett.* **584**, 2606–2617 (2010).
26. Fukami-Kobayashi, K., Tateno, Y. & Nishikawa, K. Domain dislocation: a change of core structure in periplasmic binding proteins in their evolutionary history. *J. Mol. Biol.* **286**, 279–290 (1999).
27. Banerjee, H., Rahn, A., Davis, W. & Singh, R. Sex lethal and U2 small nuclear ribonucleoprotein auxiliary factor (U2AF65) recognize polypyrimidine tracts using multiple modes of binding. *RNA* **9**, 88–99 (2003).
28. Sickmier, E. A. et al. Structural basis for polypyrimidine tract recognition by the essential pre-mRNA splicing factor U2AF65. *Mol. Cell* **23**, 49–59 (2006).
29. Huang, J. R. et al. Transient electrostatic interactions dominate the conformational equilibrium sampled by multidomain splicing factor U2AF65: a combined NMR and SAXS study. *J. Am. Chem. Soc.* **136**, 7068–7076 (2014).
30. MacKereth, C. D. et al. Multi-domain conformational selection underlies pre-mRNA splicing regulation by U2AF. *Nature* **475**, 408–413 (2011).
31. Kapanidis, A. N. et al. Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules. *Proc. Natl Acad. Sci. USA* **101**, 8936–8941 (2004).
32. Kapanidis, A. N. et al. Alternating-laser excitation of single molecules. *Acc. Chem. Res.* **38**, 523–533 (2005).
33. Müller, B. K., Zaychikov, E., Bräuchle, C. & Lamb, D. C. Pulsed interleaved excitation. *Biophys. J.* **89**, 3508–3522 (2005).
34. Lee, N. K. et al. Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. *Biophys. J.* **88**, 2939–2953 (2005).
35. Kudryavtsev, V. et al. Combining MFD and PIE for accurate single-pair Förster resonance energy transfer measurements. *Chem. Phys. Chem.* **13**, 1060–1078 (2012).
36. Von Voithenberg, L. V. et al. Recognition of the 3′ splice site RNA by the U2AF heterodimer involves a dynamic population shift. *Proc. Natl Acad. Sci. USA* **113**, E7169–E7175 (2016).
37. Sánchez-Rico, C., Voith von Voithenberg, L., Warner, L., Lamb, D. C. & Sattler, M. Effects of fluorophore attachment on protein conformation and dynamics studied by spFRET and NMR spectroscopy. *Chemistry* **23**, 14267–14277 (2017).

38. Eggeling, C., Widengren, J., Rigler, R. & Seidel, C. A. M. Photobleaching of fluorescent dyes under conditions used for single-molecule detection: evidence of two-step photolysis. *Anal. Chem.* **70**, 2651–2659 (1998).
39. Chung, H. S., McHale, K., Louis, J. M. & Eaton, W. A. Single-molecule fluorescence experiments determine protein folding transition path times. *Science* **335**, 981–984 (2012).
40. Ramanathan, R. & Muñoz, V. A method for extracting the free energy surface and conformational dynamics of fast-folding proteins from single molecule photon trajectories. *J. Phys. Chem. B.* **119**, 7944–7956 (2015).
41. McKinney, S. A., Joo, C. & Ha, T. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys. J.* **91**, 1941–1951 (2006).
42. 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).
43. 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).
44. Margittai, M. et al. Single-molecule fluorescence resonance energy transfer reveals a dynamic equilibrium between closed and open conformations of syntaxin 1. *Proc. Natl Acad. Sci. USA* **100**, 15516–15521 (2003).
45. Diez, M. et al. Proton-powered subunit rotation in single membrane-bound F 0F1-ATP synthase. *Nat. Struct. Mol. Biol.* **11**, 135–141 (2004).
46. Torres, T. & Levitus, M. Measuring conformational dynamics: a new FCS-FRET approach. *J. Phys. Chem. B.* **111**, 7392–7400 (2007).
47. Felekyan, S., Sanabria, H., Kalinin, S., Kühnemuth, R. & Seidel, C. A. M. Analyzing Förster resonance energy transfer with fluctuation algorithms. *Methods Enzymol.* **519**, 39–85 (2013).
48. 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. *Chem. Phys. Chem.* **13**, 1036–1053 (2012).
49. Olofsson, L. et al. Fine tuning of sub-millisecond conformational dynamics controls metabotropic glutamate receptors agonist efficacy. *Nat. Commun.* **5**, 5206 (2014).
50. Torella, J. P., Holden, S. J., Santoso, Y., Hohlbein, J. & Kapanidis, A. N. Identifying molecular dynamics in single-molecule FRET experiments with burst variance analysis. *Biophys. J.* **100**, 1568–1577 (2011).
51. Tomov, T. E. et al. Disentangling subpopulations in single-molecule FRET and ALEX experiments with photon distribution analysis. *Biophys. J.* **102**, 1163–1173 (2012).
52. Kalinin, S., Valeri, A., Antonik, M., Felekyan, S. & Seidel, C. A. M. Detection of structural dynamics by FRET: a photon distribution and fluorescence lifetime analysis of systems with multiple states. *J. Phys. Chem. B.* **114**, 7983–7995 (2010).
53. Gopich, I. V. & Szabo, A. Theory of the energy transfer efficiency and fluorescence lifetime distribution in single-molecule FRET. *Proc. Natl Acad. Sci. USA* **109**, 7747–7752 (2012).
54. Nettels, D., Gopich, I. V., Hoffmann, A. & Schuler, B. Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proc. Natl Acad. Sci. USA* **104**, 2655–2660 (2007).
55. Hoffmann, A. et al. Quantifying heterogeneity and conformational dynamics from single molecule FRET of diffusing molecules: recurrence analysis of single particles (RASP). *Phys. Chem. Chem. Phys.* **13**, 1857–1871 (2011).
56. Gopich, I. V. & Szabo, A. Decoding the pattern of photon colors in single-molecule FRET. *J. Phys. Chem. B.* **113**, 10965–10973 (2009).
57. Chung, H. S. & Gopich, I. V. Fast single-molecule FRET spectroscopy: theory and experiment. *Phys. Chem. Chem. Phys.* **16**, 18644–18657 (2014).
58. 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).
59. Harris, P. D. et al. Multi-parameter photon-by-photon hidden Markov modeling. *Nat. Commun.* **13**, 1000 (2022).
60. Ingargiola, A., Weiss, S. & Lerner, E. Monte Carlo diffusion-enhanced photon inference: distance distributions and conformational dynamics in single-molecule FRET. *J. Phys. Chem. B.* **122**, 11598–11615 (2018).
61. Barth, A. et al. Unraveling multi-state molecular dynamics in single-molecule FRET experiments. I. Theory of FRET-lines. *J. Chem. Phys.* **156**, 141501 (2022).
62. De Boer, M. et al. Conformational and dynamic plasticity in substrate-binding proteins underlies selective transport in ABC importers. *eLife* **8**, e44652 (2019).
63. Dimura, M. et al. Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems. *Curr. Opin. Struct. Biol.* **40**, 163–185 (2016).
64. Sindbert, S. et al. Accurate distance determination of nucleic acids via Förster resonance energy transfer: implications of dye Linker length and rigidity. *J. Am. Chem. Soc.* **133**, 2463–2480 (2011).
65. Steffen, F. D., Sigel, R. K. O. & Börner, R. An atomistic view on carbocyanine photophysics in the realm of RNA. *Phys. Chem. Chem. Phys.* **18**, 29045–29055 (2016).
66. Klose, D. et al. Resolving distance variations by single-molecule FRET and EPR spectroscopy using rotamer libraries. *Biophys. J.* **120**, 4842–4858 (2021).
67. Reinartz, I. et al. Simulation of FRET dyes allows quantitative comparison against experimental data. *J. Chem. Phys.* **148**, 123321 (2018).
68. Hoefling, M. et al. Structural heterogeneity and quantitative FRET efficiency distributions of polyprolines through a hybrid atomistic simulation and monte carlo approach. *PLoS ONE* **6**, 19791 (2011).
69. Peulen, T. O., Opanasyuk, O. & Seidel, C. A. M. Combining graphical and analytical methods with molecular simulations to analyze time-resolved FRET measurements of labeled macromolecules accurately. *J. Phys. Chem. B.* **121**, 8211–8241 (2017).
70. Dale, R. E., Eisinger, J. & Blumberg, W. E. The orientational freedom of molecular probes. The orientation factor in intramolecular energy transfer. *Biophys. J.* **26**, 161–193 (1979).
71. Dale, R. E. & Eisinger, J. Intramolecular distances determined by energy transfer. Dependence on orientational freedom of donor and acceptor. *Biopolymers* **13**, 1573–1605 (1974).
72. Ivanov, V., Li, M. & Mizuuchi, K. Impact of emission anisotropy on fluorescence spectroscopy and FRET distance measurements. *Biophys. J.* **97**, 922–929 (2009).
73. Eilert, T., Kallis, E., Nagy, J., Röcker, C. & Michaelis, J. Complete kinetic theory of FRET. *J. Phys. Chem. B* **122**, 11677–11694 (2018).
74. Maier, J. A. et al. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
75. Zaccai, G. How soft is a protein? A protein dynamics force constant measured by neutron scattering. *Science* **288**, 1604–1607 (2000).
76. Vinogradov, S. A. & Wilson, D. F. Recursive maximum entropy algorithm and its application to the luminescence lifetime distribution recovery. *Appl. Spectrosc.* **54**, 849–855 (2000).
77. Ingargiola, A., Lerner, E., Chung, S. Y., Weiss, S. & Michalet, X. FRETBursts: an open source toolkit for analysis of freely-diffusing Single-molecule FRET. *PLoS ONE* **11**, 39198 (2016).

78. Schrimpf, W., Barth, A., Hendrix, J. & Lamb, D. C. PAM: a framework for integrated analysis of imaging, single-molecule, and ensemble fluorescence data. *Biophys. J.* **114**, 1518–1528 (2018).
79. Ambrose, B. et al. The smfBox is an open-source platform for single-molecule FRET. *Nat. Commun.* **11**, 5641 (2020).
80. Knight, J. L., Mekler, V., Mukhopadhyay, J., Ebright, R. H. & Levy, R. M. Distance-restrained docking of rifampicin and rifamycin SV to RNA polymerase using systematic FRET measurements: developing benchmarks of model quality and reliability. *Biophys. J.* **88**, 925–938 (2005).
81. Kapanidis, A. N. et al. Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science* **314**, 1144–1147 (2006).
82. Sanabria, H. et al. Resolving dynamics and function of transient states in single enzyme molecules. *Nat. Commun.* **11**, 1231 (2020).
83. Berman, H. M. et al. Federating structural models and data: outcomes from a workshop on archiving integrative. *Structures. Structure* **27**, 1745–1759 (2019).
84. Lane, T. J. Protein structure prediction has reached the single-structure frontier. *Nat. Methods* **20**, 170–173 (2023).
85. 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).
86. Thomsen, J. et al. DeepFRET, a software for rapid and automated single-molecule FRET data classification using deep learning. *eLife* **9**, e60404 (2020).
87. Wanninger, S. et al. Deep-learning assisted, single-molecule imaging analysis (Deep-LASI) of multi-color DNA Origami structures. Preprint at *bioRxiv* https://doi.org/10.1101/2023.01.31.526220 (2023).

Ganesh Agam[1,33], Christian Gebhardt[2,33], Milana Popara[3,33], Rebecca Mächtel[2], Julian Folz[3], Benjamin Ambrose[4], Neharika Chamachi[5], Sang Yoon Chung[6], Timothy D. Craggs[4], Marijn de Boer[7], Dina Grohmann[8], Taekjip Ha[9], Andreas Hartmann[5], Jelle Hendrix[10,11], Verena Hirschfeld[12], Christian G. Hübner[12], Thorsten Hugel[13,14], Dominik Kammerer[15,16], Hyun-Seo Kang[17], Achillefs N. Kapanidis[15,16], Georg Krainer[5,18], Kevin Kramm[8], Edward A. Lemke[19,20,21], Eitan Lerner[22], Emmanuel Margeat[23], Kirsten Martens[24], Jens Michaelis[25], Jaba Mitra[9,26], Gabriel G. Moya Muñoz[2], Robert B. Quast[23], Nicole C. Robb[15,16,30], Michael Sattler[17,27], Michael Schlierf[5,28], Jonathan Schneider[2], Tim Schröder[1], Anna Sefer[25], Piau Siong Tan[19,20], Johann Thurn[13,31], Philip Tinnefeld[1], John van Noort[24], Shimon Weiss[6,29], Nicolas Wendler[2], Niels Zijlstra[2], Anders Barth[3,32] ✉, Claus A. M. Seidel[3] ✉, Don C. Lamb[1] ✉ & Thorben Cordes[2] ✉

[1]Department of Chemistry, Ludwig-Maximilians University München, München, Germany. [2]Physical and Synthetic Biology, Faculty of Biology, Ludwig-Maximilians University München, Planegg-Martinsried, Germany. [3]Molecular Physical Chemistry, Heinrich-Heine University Düsseldorf, Düsseldorf, Germany. [4]Department of Chemistry, University of Sheffield, Sheffield, UK. [5]B CUBE – Center for Molecular Bioengineering, Technische Universität Dresden, Dresden, Germany. [6]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA, USA. [7]Molecular Microscopy Research Group, Zernike Institute for Advanced Materials, University of Groningen, AG Groningen, the Netherlands. [8]Department of Biochemistry, Genetics and Microbiology, Institute of Microbiology, Single-Molecule Biochemistry Laboratory, University of Regensburg, Regensburg, Germany. [9]Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine and Howard Hughes Medical Institute, Baltimore, MD, USA. [10]Dynamic Bioimaging Laboratory, Advanced Optical Microscopy Center and Biomedical Research Institute, Hasselt University, Agoralaan C (BIOMED), Hasselt, Belgium. [11]Department of Chemistry, KU Leuven, Leuven, Belgium. [12]Institute of Physics, University of Lübeck, Lübeck, Germany. [13]Institute of Physical Chemistry, University of Freiburg, Freiburg, Germany. [14]Signalling Research Centers BIOSS and CIBSS, University of Freiburg, Freiburg, Germany. [15]Department of Physics, Clarendon Laboratory, University of Oxford, Oxford, UK. [16]Kavli Institute of Nanoscience Discovery, University of Oxford, Oxford, UK. [17]Bayerisches NMR Zentrum, Department of Bioscience, School of Natural Sciences, Technical University of München, Garching, Germany. [18]Yusuf Hamied Department of Chemistry, University of Cambridge, Cambridge, UK. [19]Biocenter, Johannes Gutenberg University Mainz, Mainz, Germany. [20]Institute of Molecular Biology, Mainz, Germany. [21]Structural and Computational Biology Unit, European Molecular Biology Laboratory, Heidelberg, Germany. [22]Department of Biological Chemistry, The Alexander Silberman Institute of Life Sciences, and The Center for Nanoscience and Nanotechnology, Faculty of Mathematics and Science, The Edmond J. Safra Campus, The Hebrew University of Jerusalem, Jerusalem, Israel. [23]Centre de Biologie Structurale (CBS), University of Montpellier, CNRS, INSERM, Montpellier, France. [24]Biological and Soft Matter Physics, Huygens–Kamerlingh Onnes Laboratory, Leiden University, Leiden, the Netherlands. [25]Institute for Biophysics, Ulm University, Ulm, Germany. [26]Materials Science and Engineering, University of Illinois Urbana-Champaign, Urbana, IL, USA. [27]Institute of Structural Biology, Molecular Targets and Therapeutics Center, Helmholtz Center Munich, Munich, Germany. [28]Cluster of Excellence Physics of Life, Technische Universität Dresden, Dresden, Germany. [29]California NanoSystems Institute, University of California, Los Angeles, CA, USA. [30]Present address: Warwick Medical School, The University of Warwick, Coventry, UK. [31]Present address: Institute of Technical Physics, German Aerospace Center (DLR), Stuttgart, Germany. [32]Present address: Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Delft, the Netherlands. [33]These authors contributed equally: Ganesh Agam, Christian Gebhardt, Milana Popara. ✉e-mail: a.barth@tudelft.nl; cseidel@hhu.de; d.lamb@lmu.de; cordes@bio.lmu.de

Methods

Sample preparation of proteins

Double-cysteine mutants of MalE were prepared and labeled using established protocols[62]. Human RRM1,2 L187C-G326C mutant (U2AF2-148-342) was obtained and purified as described in Mackereth et al.[30].

Fluorescence labeling of proteins

All fluorophores were purchased as maleimide derivatives from commercial suppliers as listed in Supplementary Table 19. MalE was stochastically labeled as described previously[88] with fluorophores as indicated in the text with a combined labeling efficiency higher than 70% resulting in a donor–acceptor pairing of at least 20%. Protein stability and functionality (ligand binding) was verified by affinity measurements using microscale thermophoresis[89]. All preparations, that is, MalE-wildtype, unlabeled cysteine mutants and fluorophore-labeled variants, showed an affinity for maltose between roughly 1 and 2 µM (Supplementary Fig. 5) consistent with previously published $K_d$values for wildtype MalE[90,91]. The stability and labeling of the sample were verified by FCS (Supplementary Fig. 18), which excluded the presence of larger aggregates in the samples and confirms that MalE is functional.

U2AF2 was stochastically labeled as described previously in Voith von Voithenberg et al.[36]. The combined labeling efficiencies for the labeling reactions were 20 and 14% for the Alexa546–Alexa647 and Atto532–Atto643 pairs, respectively. For Alexa488–Alexa647, the combined labeling efficiency was found to be 10%. The functionality of the labeled U2AF protein was checked with affinity measurements for U9 RNA, which was found to be 1.2 µM (ref. 30), consistent with the previous reports[36] (Supplementary Fig. 7d).

Sample handling

Both protein systems required special handling due to sample instability or aggregate formation, which are both problematic for long-term storage and shipping. The labeled MalE proteins were stored in 50 mM Tris-HCl pH 7.4, 50 mM KCl with 1 mg ml^{-1} BSA at 4 °C for less than 7 d. U2AF2 was stored in 20 mM potassium phosphate buffer pH 6.5, 50 mM NaCl and kept in the fridge until used. Both samples were loaded in low-binding Eppendorf tubes (Eppendorf Germany, catalog no. 0030108094) and shipped on ice in a cooling box with overnight shipping to avoid unnecessary freezing and thawing. MalE stock solutions were on the order of 10 to 100 nM concentration and the sent stock solution of U2AF2 was 5–10 µM concentration. Dilution buffers for apo and holo measurements were provided. SmFRET experiments were carried out by diluting the labeled proteins to concentrations of roughly 50 pM in 50 mM Tris-HCl pH 7.4, 50 mM KCl supplemented with the ligand maltose at 1 mM concentration. Labeled U2AF2 protein was measured at roughly 40–100 pM in 20 mM potassium phosphate buffer pH 6.5, 50 mM NaCl. Purchased U9 RNA (Biomers.net GmbH and IBA Solutions for Life Sciences) was dissolved in RNA-free water and added directly to the solution at a final concentration of 5 µM for the holo measurements. Both proteins were studied on coverslips typically passivated with 1 mg ml^{-1} BSA in buffer before adding the sample. The measurements were performed without any photostabilizer to keep the measurements as simple as possible to avoid any further source for discrepancies between the groups, for example, degradation of photostabilizer or use of different photostabilizer concentrations.

SmFRET data acquisition and analysis

Data acquisition and correction procedures were performed for confocal measurements as described by Hellenkamp et al.[18]. The samples were measured using ALEX or PIE on a confocal microscope as sketched in Supplementary Fig. 2. A description of the experimental procedures of all laboratories is given in Supplementary Note 18.

Briefly, the three recorded intensity time traces for each single-molecule event are:

$$\text{donor emission after donor excitation}: {}^{\mathrm{i}}I_{\mathrm{Dem|Dex}},$$

$$\text{acceptor emission after donor excitation (FRET signal)}: {}^{\mathrm{i}}I_{\mathrm{Aem|Dex}},$$

$$\text{and acceptor emission after acceptor excitation}: {}^{\mathrm{i}}I_{\mathrm{Aem|Dex}}.$$

The apparent (raw) FRET efficiency is computed as:

$$E_{\mathrm{app}} = \frac{{}^{\mathrm{i}}I_{\mathrm{Aem|Dex}}}{{}^{\mathrm{i}}I_{\mathrm{Dem|Dex}} + {}^{\mathrm{i}}I_{\mathrm{Aem|Dex}}}, \tag{3}$$

Recorded intensities were corrected for background contributions as:

$${}^{\mathrm{ii}}I_{\mathrm{Dem|Dex}} = {}^{\mathrm{i}}I_{\mathrm{Dem|Dex}} - I^{(\mathrm{BG})}_{\mathrm{Dem|Dex}}, \tag{4}$$

$${}^{\mathrm{ii}}I_{\mathrm{Aem|Dex}} = {}^{\mathrm{i}}I_{\mathrm{Aem|Dex}} - I^{(\mathrm{BG})}_{\mathrm{Aem|Dex}}, \tag{5}$$

$${}^{\mathrm{ii}}I_{\mathrm{Aem|Aex}} = {}^{\mathrm{i}}I_{\mathrm{Aem|Aex}} - I^{(\mathrm{BG})}_{\mathrm{Aem|Aex}}, \tag{6}$$

where $I^{(\mathrm{BG})}_{\mathrm{Dem|Dex}}$, $I^{(\mathrm{BG})}_{\mathrm{Aem|Dex}}$ and $I^{(\mathrm{BG})}_{\mathrm{Aem|Aex}}$ are the respective background signals. Correction factors for spectral crosstalk, α and direct excitation, δ, were determined from the donor- and acceptor-only populations[34]. The corrected acceptor fluorescence after donor excitation, $F_{\mathrm{A|D}}$, is computed as:

$$F_{\mathrm{A|D}} = {}^{\mathrm{ii}}I_{\mathrm{Aem|Dex}} - \alpha\,{}^{\mathrm{ii}}I_{\mathrm{Dem|Dex}} - \delta\,{}^{\mathrm{ii}}I_{\mathrm{Aem|Aex}} \tag{7}$$

The γ and β factors, correcting for differences in the detection yield and excitation fluxes of the donor and acceptor dyes, were estimated using a global correction procedure following the approach of Lee et al. (Supplementary Fig. 3)[34]. Alternatively, when pulsed excitation was used and the sample is known to be static, the γ factor can be determined by fitting the measured population to the static FRET line[35,92]. This allows a robust determination of the γ factor when only a single species is present but requires a static sample and the appropriate static FRET line (Supplementary Note 2).

The accurate FRET efficiency E and stoichiometry S values were then calculated as:

$$E = \frac{F_{\mathrm{A|D}}}{\gamma\,{}^{\mathrm{ii}}I_{\mathrm{Dem|Dex}} + F_{\mathrm{A|D}}}, \tag{8}$$

$$S = \frac{\gamma\,{}^{\mathrm{ii}}I_{\mathrm{Dem|Dex}} + F_{\mathrm{A|D}}}{\gamma\,{}^{\mathrm{ii}}I_{\mathrm{Dem|Dex}} + F_{\mathrm{A|D}} + {}^{\mathrm{ii}}I_{\mathrm{Aem|Aex}}/\beta}. \tag{9}$$

Conversion of accurate FRET efficiencies into distances were done using equation (2) with Förster radii determined as described in Supplementary Note 7.

Detection of protein dynamics

In this work, we used the following two approaches to detect conformational dynamics:

BVA. In BVA, the presence of dynamics is determined by looking for excess variance in the FRET efficiency data beyond the shot-noise limit. The standard deviation ($\sigma_{E_{\mathrm{app}}}$) of the apparent FRET efficiency (E_{app}) is calculated using a fixed photon window of $n = 5$ over the time period of the individual bursts given by:

$$\sigma_{E_{app}} = \sqrt{\frac{E_{app}(1-E_{app})}{n}}, \quad (10)$$

The shot-noise limited standard deviation of the apparent FRET efficiency is generally described by a semicircle[50] (Fig. 4a and Supplementary Fig. 11a–d). In the presence of dynamics, the standard deviation for the FRET efficiency within a burst becomes higher than that expected from shot noise. Photophysical effects such as photobleaching and blinking also give rise to the higher standard deviation beyond the shot-noise limit. Typically, BVA is sensitive to fluctuations in the FRET signal of ≥100 µs, but this depends on the brightness of the burst and the photon window used.

FRET efficiency versus fluorescence-weighted average donor lifetime analysis (*E*–τ plots). Two-dimensional histograms of the FRET efficiency E and donor fluorescence lifetime $\langle\tau_{D(A)}\rangle_F$ (Fig. 4b and Supplementary Fig. 11e–h) were created for single-molecule measurements using multiparameter fluorescence detection (MFD) in combination with PIE[35], described below. Static FRET lines were calculated using the following equation:

$$E = 1 - \frac{\tau_{D(A)}}{\tau_{D(0)}} \quad (11)$$

and further modified for linker dynamics[61]. Deviations of FRET populations from the static FRET line can indicate FRET dynamics, which can be due to conformational fluctuations or photophysical dynamics. In addition, a time-resolved FRET analysis of TCSPC data can accurately resolve the distance heterogeneities by revealing multiple components in the decay curve and recovering their specific species fractions and FRET rate constants[69]. Dynamics are thus detected from the presence of multiple components in the subensemble decay of a single FRET population. In addition, dynamics that are slower than the fluorescence lifetime (roughly 5 ns) are not averaged in the FRET lifetime analysis leading to the detection of the full conformational distribution.

MFD with PIE

MFD, introduced by Eggeling et al.[93], combines spectral and polarized detection with picosecond pulsed lasers and TCSPC, allowing the simultaneous detection of intensity, lifetime, anisotropy and spectral range of the fluorescence signal of single molecules. nsALEX or PIE additionally provides the acceptor lifetime information[35]. Due to the availability of the lifetime information when using pulsed excitation, this approach is well suited for using *E*–τ-based analyses.

Dye simulations (AV and ACV)

The AV approach uses a simple coarse-grained dye model[64] defined by five parameters: the width and length of the linker, and three radii that define the fluorophore volume (Fig. 5a and Supplementary Table 10). Using these parameters, AV simulations for both fluorophores were calculated by considering the linker flexibility and steric hindrances of the labeled molecule (Fig. 5a). In the ACV model[63], the position of the dyes is biased toward the protein surface, resulting in a reduction of the interdye distance for the given labeling positions. To do this, the residual anisotropy was used to estimate the fraction of sticking dyes. In the computation of the FRET-averaged model distances, the occupancy of a thin surface layer (roughly 3 Å) was then increased such that its fraction matches the amount of interacting dye detected in the experiment (Fig. 5b and Supplementary Table 10).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data for all figures, all supplementary figures, the raw data for MalE measurements from all laboratories (with the exception of one mutant from one laboratory) and the raw data for all U2AF2 measurements have been uploaded to Zenodo (https://doi.org/10.5281/zenodo.7472900). PDB IDs used are 1OMP, 1ANF, 2YHO and 2YH1. Source data are provided with this paper.

Code availability

The software used for data analysis are available from the respective laboratories: laboratory no. 1, PAM (PIE Analysis with MATLAB) software package (ref. 79 in the main text); laboratory no. 2, Home-written LabView-based software (ref. 15 in Supplementary Information); Laboratory no. 3, FRETBursts toolkit (ref. 14 in Supplementary Information); laboratory no. 4, PAM (PIE Analysis with MATLAB) software package; laboratory no. 5, PAM (PIE Analysis with MATLAB) software package; laboratory no. 6, PAM (PIE Analysis with MATLAB) software package; laboratory no. 7, PAM (PIE Analysis with MATLAB) software package; laboratory no. 8, PAM (PIE Analysis with MATLAB) software package v.2.0; laboratory no. 9, PAM (PIE Analysis with MATLAB) software package; laboratory no. 11, data were analyzed with the burst analysis toolbox (BAT, V2018 and V2019) and filtered and visualized with T3ee (V2018, V2019) (ref. 84 in Supplementary Information); laboratory no. 12: PAM (PIE Analysis with MATLAB) software package; laboratory no. 13, IgorPro 8 (Wavemetrics); laboratory no. 14, PAM (PIE Analysis with MATLAB) software package; laboratory no. 15, Software Package for Multiparameter Fluorescence Spectroscopy, Full Correlation and Multiparameter Fluorescence Imaging developed in C.A.M. Seidel's laboratory (http://www.mpc.uni-duesseldorf.de/seidel/); laboratory no. 16, ALEX-suite software package (ref. 88 of Supplementary Information); laboratory no. 17, FRETBursts analysis software (ref. 14 of Supplementary Information) and laboratory no. 18, FRETBursts toolkit (ref. 14 of Supplementary Information).

References

88. Gouridis, G. et al. Conformational dynamics in substrate-binding domains influences transport in the ABC importer GlnPQ. *Nat. Struct. Mol. Biol.* **22**, 57–64 (2015).
89. Jerabek-Willemsen, M. et al. MicroScale thermophoresis: interaction analysis and beyond. *J. Mol. Struct.* **1077**, 101–113 (2014).
90. Hall, J. A., Gehring, K. & Nikaido, H. Two modes of ligand binding in maltose-binding protein of *Escherichia coli*: correlation with the structure of ligands and the structure of binding protein. *J. Biol. Chem.* **272**, 17605–17609 (1997).
91. Kim, E. et al. A single-molecule dissection of ligand binding to a protein with intrinsic dynamics. *Nat. Chem. Biol.* **9**, 313–318 (2013).
92. Sisamakis, E., Valeri, A., Kalinin, S., Rothwell, P. J. & Seidel, C. A. M. Accurate single-molecule FRET studies using multiparameter fluorescence detection. *Methods Enzymol.* **475**, 455–514 (2010).
93. Eggeling, C. et al. Data registration and selective single-molecule analysis using multi-parameter fluorescence detection. *J. Biotechnol.* **86**, 163–180 (2001).

Acknowledgements

Work in the laboratory of T.C. was financed by a European Research Council (ERC) Starting grant (no. ERC-StG 638536—SM-IMPORT), German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) within grant nos. GRK2062 (project C03) and SFB863 (project A13) and an Alexander von Humboldt postdoctoral fellowship (to N.Z.). T.C., P.T. and D.C.L. acknowledge the support of the Center for integrated protein science Munich and the Center for NanoScience. D.C.L. acknowledges the support of the Nanosystems Initiative Munich and LMUinnovative program BioImaging

Network. We also acknowledge support via the SFB1035 (DFG, Sonderforschungsbereich 1035 project no. 201302640, project no. A11 to D.C.L. and project no. B03 to M. Sattler). D.C.L. and P.T. acknowledge support by the Federal Ministry of Education and Research (BMBF) and the Free State of Bavaria under the Excellence Strategy of the Federal Government and the Länder through the ONE MUNICH Project Munich Multiscale Biofabrication. C.A.M.S. acknowledges the support by the ERC (grant no. 671208 (hybridFRET)) and by the DFG (grant nos. SE 1195/17-1 and CRC 1208 (project no. A08)). A.B. acknowledges funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 101029907. Research in the contributing authors' laboratories was financed by the following sources: P.T. acknowledges the support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project ID 201269156, SFB 1032 (A13) and project ID 267681426. T.D.C. was supported by the Biotechnology and Biological Sciences Research Council (BBSRC) (grant no. BB/T008032/1) and Engineering and Physical Sciences Research Council (EPSRC) (grant no. EP/V034804/1), B.A. was supported by an EPSRC Prize Fellowship. BMBF grant nos. 03Z2EN11 and 03Z22E511 as well as DFG grant no. SCHL 1896/4-1 (to M. Schlierf). SFB960 project A7 (to D.G.) US National Institutes of Health grant no. GM122569 (to T. Ha). J.H. acknowledges the Research Foundation Flanders (FWO) (project nos. G0B4915, G0B9922N and G0H3716N) and is indebted to Johan Hofkens at KU Leuven for the used smFRET infrastructure. ERC grant agreement no. 681891 (Prosint) and DFG under Germany's Excellence Strategy (CIBSS EXC-2189 project ID 390939984) and the SFB1381 program (project ID 403222702) (to T. Hugel). Royal Society Dorothy Hodgkin Research Fellowship DKR00620 and a Research Grant for Research Fellows no. RGF\R1\180054 (to N.C.R.), by the Wellcome Trust (grant no. 110164/Z/15/Z to A.N.K.). The Israel Science Foundation (grant nos. 556/22 to E.L., 3565/20 to E.L., within the KillCorona – Curbing Coronavirus Research Program), the National Institutes of Health (grant no. R01 GM130942 to S.W. and to E.L. as a subaward), by the Milner Fund (to E.L.) and by the Hebrew University of Jerusalem (start-up funds to E.L.). Agence Nationale pour la Recherche (grant nos. ANR 18-CE11-0004-02, ANR-19-CE44-0009-02, ANR-21-CE11-0034-01, ANR-21-CE11-0026-03 and ANR-10-INBS-04, 'Investments for the future' to E.M.). E.A.L. acknowledges funding by the ERC ADG MultiOrganelleDesign and the SFB1551 (project ID 464588647).

Author contributions

T.C. initiated the study and D.C.L. coordinated the study. G.A., C.G., A.B., C.A.M.S., D.C.L. and T.C. designed research. A.B., C.A.M.S., D.C.L. and T.C. supervised the project. R.M. cloned and purified MalE variants. H.-S.K. and M.S. provided U2AF2. C.G. and G.A. performed labeling of MalE and U2AF2 variants, respectively, for shipment to participating laboratories. M.P., J.F. and C.A.M.S. designed MalE mutants 4 and 5 in silico. M.P. and J.F. performed initial measurements on MalE mutants 4 and 5. G.A. reperformed the analysis on the provided raw data for U2AF2 and MalE-1 from eight laboratories. G.A., A.B. and M.P. performed ds estimation. C.G. performed FCS experiments on MalE variants, time-resolved anisotropy experiments and R_0-determination. M.P. performed time-resolved anisotropy analysis of single labeled MalE cysteine mutants from ensemble measurements as well of all MalE and U2AF2 dye combinations from smFRET measurements. C.G., G.A. and M.P. performed measurements of MalE mutants with additional dye combinations and M.P. and A.B. performed statistical analysis of ds and anisotropies. G.A. and A.B. performed estimation of setup-dependent parameters and PDA of U2AF2. G.A. performed the filtered-FCS and A.B. performed TCSPC analysis of U2AF2. C.G., G.A. and M.P. performed smFRET measurements on dsDNA rulers. M.P. performed AV and ACV modeling of dye distributions for MalE and U2AF2. G.G.M.M. performed microscale thermophoresis experiments. M.d.B. performed confocal scanning experiments for surface-immobilized MalE. All authors were involved in performing comparison experiments and analyzing smFRET data. G.A. and C.G. consolidated data collection of participating laboratories. G.A. and C.G. designed Fig. 1. G.A., C.G. and M.P. designed Fig. 2. A.B. designed Figs. 3 and 4. A.B. and M.P. designed Figs. 5 and 6. G.A., C.G., M.P., A.B., C.A.M.S., D.C.L. and T.C. interpreted data and wrote the manuscript in consultation with all authors.

Competing interests

T.D.C. and A.N.K. are founders of different companies selling single-molecule fluorescence microscopes (Exciting Instruments, Oxford Nanoimager). The other authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41592-023-01807-0.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41592-023-01807-0.

Correspondence and requests for materials should be addressed to Anders Barth, Claus A. M. Seidel, Don C. Lamb or Thorben Cordes.

Peer review information *Nature Methods* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editor: Rita Strack, in collaboration with the *Nature Methods* team.Peer reviewer reports are available.

https://doi.org/10.1038/s41592-023-01807-0

Extended Data Table. 1 | Average of mean FRET efficiency and standard deviation for MalE and U2AF2 samples reported by the participating laboratories

		Experimental Values			Modeled Distances	
Sample	**Condition**	$\mu_{\langle E\rangle} \pm \sigma_{\langle E\rangle}$	$\mu_{\langle E_{\mathrm{holo}}\rangle-\langle E_{\mathrm{apo}}\rangle} \pm \sigma_{\langle E_{\mathrm{holo}}\rangle-\langle E_{\mathrm{apo}}\rangle}$	$\mu_{R_{\langle E\rangle}} \pm \sigma_{R_{\langle E\rangle}}$ [Å]	$R^{\mathrm{AV}}_{\langle E\rangle}$ [Å]	$R^{\mathrm{ACV}}_{\langle E\rangle}$ [Å]
MalE-1	apo	0.49±0.06	0.177±0.019	65.4±2.6	72.0	67.7
	holo	0.67±0.05		57.8±2.1	62.1	58.3
MalE-2	apo	0.83±0.03	-0.121±0.019	50.0±1.9	50.1	48.8
	holo	0.71±0.05		56.1±2.1	56.5	55.0
MalE-3	apo	0.913±0.019	0.007±0.010	43.8±1.7	39.9	38.9
	holo	0.920±0.021		43.0±2.4	40.8	39.8
MalE-4*	apo	0.442±0.025		67.6±1.2	67.8	64.3
	holo	0.678±0.017		57.4±0.7	56.9	54.6
MalE-5*	apo	0.613±0.003		60.22±0.15	61.8	59.3
	holo	0.821±0.001		50.43±0.08	49.3	48.2
U2AF2**	apo	0.74±0.03		49.6±1.3		
	holo***	0.46±0.04		60.8±1.7		

The calculated average $\mu_{\langle E\rangle}$ and standard deviation $\sigma_{\langle E\rangle}$ of the mean FRET efficiency values provided by the participating labs are given for all three studied mutants of MalE labeled with Alexa546 and Alexa647 under both apo and holo conditions (see Supplementary Table 3). The calculated mean and standard deviation of the difference in the reported mean FRET efficiency between the apo and holo ($\langle E_{\mathrm{holo}}\rangle - \langle E_{\mathrm{apo}}\rangle$) for the three MalE mutants are given by $\mu_{\langle E_{\mathrm{holo}}\rangle-\langle E_{\mathrm{apo}}\rangle}$ and $\sigma_{\langle E_{\mathrm{holo}}\rangle-\langle E_{\mathrm{apo}}\rangle}$ respectively (see Supplementary Table 3). The calculated average $\mu_{R_{\langle E\rangle}}$ and standard deviation $\sigma_{R_{\langle E\rangle}}$ of the mean distances were derived according to Eq. 2. The modeled distances $R^{\mathrm{AV}}_{\langle E\rangle}$ and $R^{\mathrm{ACV}}_{\langle E\rangle}$ are derived using accessible volume (AV) and accessible contact volume (ACV) calculations respectively, as described in the Methods. We also give the average and standard deviation for the FRET values determined for U2AF2 labeled with Atto532-Atto643 under both apo and holo conditions (Supplementary Table 4). *Only studied by two laboratories. **Due to the fast-structural dynamics in the sample, only 7 labs studied this mutant and distances were not determined. *** Only the holo state under holo condition was considered.

nature methods

Analysis https://doi.org/10.1038/s41592-023-01807-0

Reliability and accuracy of single-molecule FRET studies for characterization of structural dynamics and distances in proteins

In the format provided by the authors and unedited

Supplementary Notes

Supplementary Note 1: Comparison between µsALEX and PIE.

With the lifetime information, it is possible to directly access distances and distance distributions via analysis of the donor fluorescence decay[1,2] and detect dynamics on the nano- to millisecond timescale in the E-τ plot, as we have shown for the apo state of U2AF2. The advantage of µsALEX is that it is less expensive to implement than nsALEX/PIE and that the continuous wave excitation exerts less photophysical stress on the fluorophores (e.g., bleaching or blinking) compared to the high peak irradiance for pulsed excitation[3]. NsALEX/PIE have the advantage that they provide the fluorescence lifetime information of the donor and acceptor fluorophore[4]. Via the lifetime information, it is also possible to detect changes of the donor and acceptor quantum yields due to protein or ligand interactions with the dye. In addition, the faster alternation timescale of nsALEX/PIE enables higher time-resolution for fluorescence correlation spectroscopy and reduces the spread in the stoichiometry distribution due to diffusion of molecules through the confocal volume during the alternation period. Combining pulsed excitation with polarized excitation and detection[4,5] further adds the fluorescence anisotropy information on the single-molecule or sub-ensemble level, allowing one to monitor size changes of the biomolecule (e.g., through the binding of interaction partners) and detect changes of the local environment of the fluorophore such as dye sticking (as seen for MalE-1) directly from the single-molecule dataset.

The spread in the results in Fig. 3d indicates that the direct probing of the acceptor in ALEX/PIE experiments was applied differently by the participating labs. Part of the participants either used high laser power for acceptor excitation to obtain reliable information on the labeling stoichiometry, acceptor lifetime and anisotropy by acquiring a maximum number of photons. Others kept the acceptor excitation power as low as possible to minimize acceptor saturation and photobleaching[6]. Regardless of which approach was used, accurate FRET efficiencies can be determined from the experimentally determined correction factors. Acceptor photobleaching results in a higher amount of donor-only molecules. In addition, acceptor saturation leads to dark states that still quench the donor via FRET and appears below the static FRET line. Acceptor blinking leads to a mixture of donor-only signal, which, in turn, results in false-positive dynamics.

Supplementary Note 2: Discussion of γ-factor estimation.

Obtaining a reliable γ-factor is very crucial for smFRET data analysis and proper determination of accurate FRET values. While correction for background, spectral crosstalk and direct excitation can be performed reliably, the detection efficiency correction factor γ is much more difficult to determine and the best approach depends on the sample at hand and the measurement modality (see Supplementary Note 1, Supplementary Note 18). Popular approaches for determining the γ-factor are 1) requiring that multiple populations on the E-S plot have the same stoichiometry value[7] or 2) ensuring that the FRET efficiency determined from intensity and that from the fluorescence lifetime fall on the static FRET line, provided the sample is static[4]. Alternative methods include fitting a plane to the photon intensities, direct determination of the detection efficiencies using theoretical computation and a series of calibration dyes (Popara et al, in preparation), using calibration standards such as the DNA FRET pairs used in the Hellenkamp et al study or using a mixture of free fluorophores with identical absorption at the donor excitation wavelength[8,9]. The latter two approaches allow determination of the detection efficiency correction factor independent of the sample of interest and allow a good control of the alignment of the system when the same standards are continually used. However, it is very often the case that the quantum yield of the fluorophores changes upon labeling and may also depend upon the local environment. Furthermore, fluorophores may undergo spectral shifts upon labeling. Both effects would lead to an inappropriate γ-factor when they are not corrected for.

An incorrect γ-factor introduces systematic errors in the FRET efficiency and the derived distances, especially for intermediate FRET efficiencies. The difficulties arising from the γ-factor can be illustrated with data from one lab used sub-optimal filter combinations for the dye pair Alexa Fluor 546 (Alexa546)/Alexa Fluor 647 (Alexa647). This resulted in an inefficient detection of the red fluorescent signal and a γ-factor of 0.09 (Supplementary Table 1; lab #18). Consequently, the determined accurate FRET efficiencies were unreliable, i.e., deviating largely from the expected values, and were hence excluded for the calculation of the averages in Fig. 1. For another laboratory, the correction factors could not be determined due to lack of a functional red laser (Supplementary Table 1; lab #19). Hence, accurate FRET values could not be calculated. In a third case, the individual populations could not be clearly separated in the data due to photoblinking and/or photobleaching that may be potentially caused by the high laser powers used (Supplementary Table 1; lab #17). An incorrect γ-factor can additionally distort the dynamic shift in the E-τ plot resulting either in unphysically negative or artificially positive apparent dynamic shifts. The absence of negative values for the apparent dynamic shifts in Fig. 4f indicates that the five participating groups estimated the γ-factor well.

In this study, we asked all groups to use a global γ-factor for the data analysis of MalE, meaning that one common γ-factor was determined for all six data sets (Supplementary Figure 3). A global γ-factor works well when all samples have the same photophysical behavior. However, when the labeling position and/or conformation of the biomolecule induces a change of the fluorescence quantum yield of the fluorophore by local quenching or enhancement, a global γ-factor is no longer strictly appropriate. Here, we observed variations of the fluorescence

lifetimes of the dyes on the order of 5-10% (Supplementary Table 11). This indicates no large variations of the fluorescence quantum yields between the different samples and justifies the global γ correction. On the other hand, a local γ correction can only be done when multiple conformations with different FRET efficiencies are present in a single measurement, or by using the lifetime information when the sample is known to be static. The γ-factors reported in this study for the MalE system with dye pair Alexa546-Alexa647 were generally low in the range from 0.2 to 0.6, with an average of 0.39±0.12 (Supplementary Table 1). Interestingly, the spread in FRET efficiency between the different groups can be described by an apparent uncertainty in the γ-factor of ~23% (Fig. 3e) corresponding to an absolute error of ±0.07. This suggests that the uncertainty in the determination of the γ-factor contribute significantly to the uncertainty in the overall FRET efficiency (see Supplementary Note 4).

Error propagation in the population estimate on the γ-factor

For the popular approach[7] in determining the γ and β correction factors, we investigated how the difference in FRET efficiency between two populations with apparent FRET efficiencies $E_{PR,i}$ and the uncertainty in estimating the individual populations propagates into an uncertainty in the detection efficiency correction factor. According to Lee et al.[7], the γ- and β-correction factors are estimated from the slope Σ and the offset Ω of a plot of $1/S_{app}$ vs E_{PR} as:

$$\gamma = \frac{\Omega - 1}{\Omega + \Sigma - 1}; \qquad \beta = \Omega + \Sigma - 1. \tag{1}$$

We assume that the population centers in the E_{PR} vs. S_{app} plot are estimated with respective uncertainties ΔE_{PR} and ΔS_{app}. From the given apparent FRET efficiencies $E_{PR,i}$, we compute the apparent stoichiometries $S_{app,i}$ for given values of the γ and β correction factors as (Lee et al. 2005):

$$S_{app,i} = \left(1 + \gamma\beta + \beta(1-\gamma)E_{PR,i}\right)^{-1}. \tag{2}$$

The parameters of the linear regression Σ and Ω are given by:

$$\Sigma = \frac{\frac{1}{S_{app,2}} - \frac{1}{S_{app,1}}}{E_{PR,2} - E_{PR,1}}; \qquad \Omega = \frac{1}{S_{app,1}} - \Sigma\, E_{PR,1} = \frac{1}{S_{app,2}} - \Sigma\, E_{PR,2}. \tag{3}$$

The propagated error $\Delta\gamma$ of the population estimates $E_{PR,i}$ and $S_{app,i}$ on the γ-factor is given by:

$$\Delta\gamma = \sqrt{\sum_{i=1,2}\left(\frac{\partial\gamma}{\partial S_{app,i}}\Delta S_{app,i}\right)^2 + \left(\frac{\partial\gamma}{\partial E_{PR,i}}\Delta E_{PR,i}\right)^2}, \tag{4}$$

where the partial derivates can be obtained using the chain rule from the partial derivatives of Σ and Ω:

$$\frac{\partial\gamma}{\partial u} = \frac{\partial\gamma}{\partial\Sigma}\cdot\frac{\partial\Sigma}{\partial u} + \frac{\partial\gamma}{\partial\Omega}\cdot\frac{\partial\Omega}{\partial u}, \tag{5}$$

where $u \in \{S_{app,1}, E_{PR,1}, S_{app,2}, E_{PR,2}\}$. The required partial derivatives are given by:

$$\frac{\partial\gamma}{\partial\Sigma} = -\frac{\gamma}{\beta}; \qquad \frac{\partial\gamma}{\partial\Omega} = \frac{1-\gamma}{\beta};$$

$$\frac{\partial\Sigma}{\partial S_{app,1}} = \frac{1}{E_{PR,2}-E_{PR,1}}\cdot\frac{1}{S_{app,1}^2}; \qquad \frac{\partial\Sigma}{\partial S_{app,2}} = \frac{1}{E_{PR,2}-E_{PR,1}}\left(-\frac{1}{S_{app,2}^2}\right);$$

$$\frac{\partial\Sigma}{\partial E_{PR,1}} = -\frac{\Sigma}{E_{PR,2}-E_{PR,1}}; \qquad \frac{\partial\Sigma}{\partial E_{PR,2}} = \frac{\Sigma}{E_{PR,2}-E_{PR,1}}; \tag{6}$$

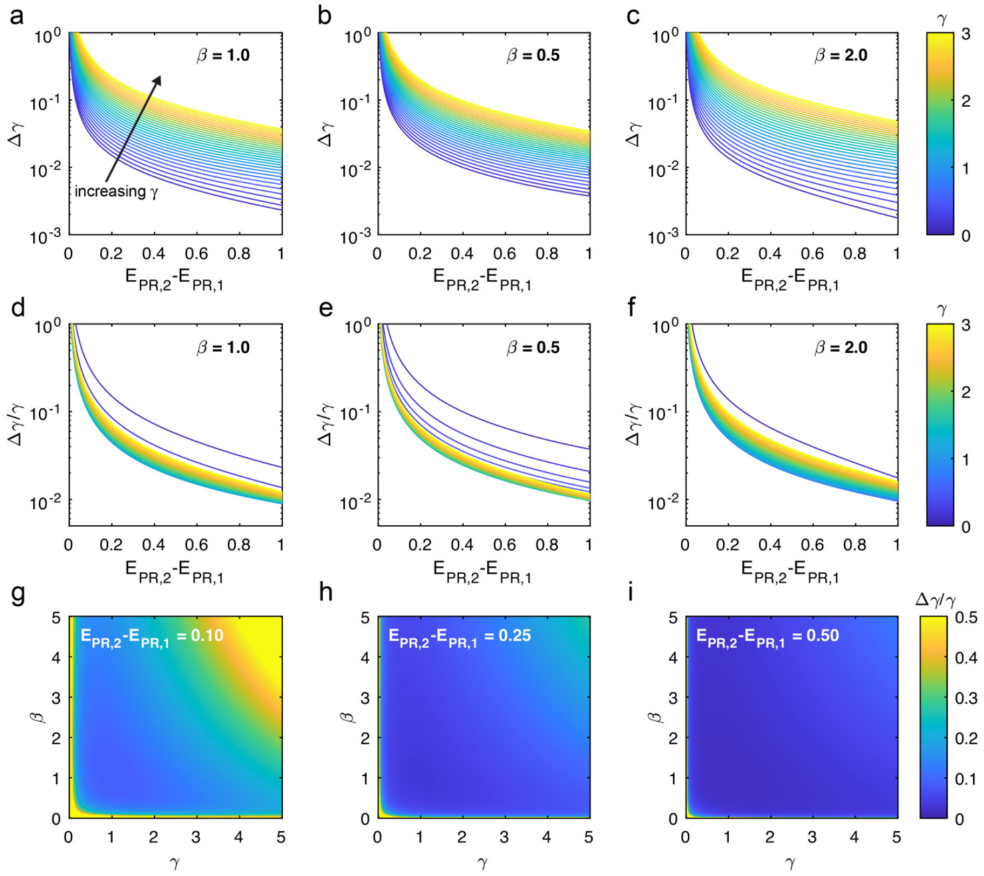
$$\frac{\partial\Omega}{\partial S_{app,1}} = \frac{-E_{PR,2}}{E_{PR,2}-E_{PR,1}}\cdot\frac{1}{S_{app,1}^2}; \qquad \frac{\partial\Omega}{\partial S_{app,2}} = \frac{E_{PR,1}}{E_{PR,2}-E_{PR,1}}\cdot\frac{1}{S_{app,2}^2};$$

$$\frac{\partial\Omega}{\partial E_{PR,1}} = \frac{-E_{PR,2}}{E_{PR,2}-E_{PR,1}}\Sigma; \qquad \frac{\partial\Omega}{\partial E_{PR,2}} = \frac{E_{PR,1}}{E_{PR,2}-E_{PR,1}}\Sigma.$$

As expected intuitively, all contributions to the uncertainty of the γ-factor scale with the inverse of the difference in the apparent FRET efficiencies. This behavior is confirmed in a plot of the absolute uncertainty $\Delta\gamma$ against the contrast of the apparent FRET efficiencies, $E_{PR,2}-E_{PR,1}$, evaluated at different absolute values of γ (Supplementary Figure SN2.1a). For the computations, the positions of the two populations were chosen to be symmetric with respect to $E_{PR}=0.5$, i.e., $E_{PR,i}=0.5\pm\left(E_{PR,2}-E_{PR,1}\right)/2$, and we assumed that 1000 events are collected for each population with a width $\sigma_{E_{PR}}=\sigma_{S_{app}}=0.05$, resulting in uncertainties of $\Delta E_{PR}=\Delta S_{app}=0.05/\sqrt{1000}=0.0016$ as given by the standard of the mean.

As expected, the uncertainty $\Delta\gamma$ scales with the absolute value of the γ-factor. Interestingly, we also observe a dependence of $\Delta\gamma$ on the value of the β correction factor (Supplementary Figure SN2.1a-c). When computing the relative error $\Delta\gamma/\gamma$ (Supplementary Figure SN2.1d-f), most curves at intermediate values of γ collapse onto a single line, suggesting that the relative uncertainty is mostly independent of the absolute value of the correction factors γ and β. To explore the dependence of the relative error $\Delta\gamma/\gamma$ on the absolute values of γ and β, we estimate $\Delta\gamma/\gamma$ at fixed values of $E_{PR,2}-E_{PR,1}$ of 0.1, 0.25, and 0.5 (Supplementary Figure SN2.1g-i). As expected, the uncertainty is highest for lowest contrast between the populations. In all cases, $\Delta\gamma/\gamma$ is minimal at intermediate values of γ and β around 1 and remains approximately constant over a large interval of γ and β values from 0.1 to 3. Notably, most experimental values for γ and β fall in this range (compare Fig. 3d).

We conclude that, with respect to the linear regression, the contrast between the apparent FRET efficiencies, $E_{PR,2}-E_{PR,1}$, is limiting for the uncertainty of the γ-factor, which scales approximately as $1/\left(E_{PR,2}-E_{PR,1}\right)$. Importantly, the absolute values of the γ and β correction factors have no large influence on the uncertainty of the γ-factor over a wide range of values that covers most experimental setups. Only at extreme values below 0.1 or above 3 are the absolute values of the γ and β correction factors limiting for the accuracy of the γ-factor estimation. Experimentally, this was investigated by Vandenberk et al[10]. No strong correlation between the determined γ-factor and ΔE_{PR} were observed. This suggests that it is not the contrast between apparent FRET efficiencies that limits the accuracy of the γ-factor estimation.



Supplementary Figure SN2.1: The propagated error of the γ-factor based on the uncertainty of the population estimate of the apparent FRET efficiency, E_{PR}. For the calculations shown here, the uncertainties of the position estimates of the populations in the E_{PR} vs. S_{app} plot are given by $\Delta E_{PR} = \Delta S_{app} = 0.05/\sqrt{1000}$. The position of the two populations is chosen to be symmetric with respect to $E_{PR} = 0.5$, i.e., $E_{PR,i} = 0.5 \pm (E_{PR,2} - E_{PR,1})/2$. **a-c)** The absolute uncertainty $\Delta\gamma$ is evaluated as a function of the difference of the apparent FRET efficiencies of the two states, $E_{PR,2} - E_{PR,1}$, at the indicated absolute values of the γ-factor (see color bar) and β values of 1 (a), 0.5 (b), and 2 (c). **d-f)** The same data shown in a-c is normalized with respect to the absolute value of γ to compute the relative uncertainty $\Delta\gamma/\gamma$. **g-i)** The relative uncertainty $\Delta\gamma/\gamma$ is evaluated as a function of the absolute values of γ and β, and are plotted for contrast between the apparent FRET efficiencies of 0.1 (g), 0.25 (h), and 0.5 (i).

Supplementary Note 3: Data re-evaluation procedure for U2AF2.

The apo population for U2AF2 is a single, dynamic population. Hence, *E*-*S* plots or *E*-τ plots are not suitable for determination of the *γ*-factor. In addition, there was an additional subpopulation with an unusually high stoichiometry. Therefore, we had to find another approach, discussed below, to accurately estimate the *γ*-factor for this system. In general, the best protocol for *γ*-factor determination will depend on the properties of the sample being measured. Reanalysis of the collected raw data of U2AF2 from different labs was performed using the PAM software[11] (Supplementary Table 17). First, a burst search was performed using an all photon burst search with a threshold of 50-100 photons per sliding time window of 500 µs depending on the dataset. For one set of measurements, a lower threshold of 20 photons per 500 µs time window was necessary. To remove blinking and bleaching events, an ALEX-2CDE filter with a lower limit of 5 and an upper limit of 25 was used depending on the data set. Values may differ depending on the excitation intensities and sample concentrations used for the measurements[12]. After burst selection, background subtraction and correction for crosstalk and direct excitation were performed as discussed in the data analysis section. Briefly, for background subtraction, the background signal was obtained from the buffer measurement and subtracted it from the burst signal. Crosstalk and direct excitation corrections were performed by calculating the signal in the acceptor channel for donor only and acceptor only species respectively, after donor excitation and subtracting them from the burst signal. To determine the detection correction factor, we used the approach of Lee *et al.*[7] (i.e. fitting a line to $1/S_{PR}$ vs E_{PR}) as the sample is dynamic and a lifetime approach is not possible. The apo configuration shows a single, dynamically averaged population so that we had to combine data from both the holo and apo measurements. We verified that there was no significant change in quantum yield of the donor and acceptor fluorophores in the absence and presence of ligand by measuring the fluorescence lifetime of the donor-only species and the acceptor lifetime with direct excitation. However, we did observe an additional subpopulation with a higher stoichiometry value (Supplementary Figure 17a). The acceptor is slightly quenched in this population, but not enough to explain the stoichiometry shift. As the origins of this population are unclear and simple explanations are insufficient to describe the observed properties, this population was not incorporated in the calculation of the *γ*-value. The average values of S_{PR} and E_{PR} for the three populations (apo, holo – low FRET, holo – high FRET) were determined from the peak values of a 2D-Gaussian fit in the ES-histograms for the respective populations. From these peak values, a straight line was fit to the three data points of $1/S_{PR}$ vs E_{PR}. Reanalysis of the data by a single person led to a further improvement of the consistency between laboratories (Fig. 2d-e). Part of the discrepancy came from the fact that the individual labs used a global *γ* approach but did not compensate for the presence of the second population. Further reasons for the discrepancies in the measured data from different laboratories arise from the fact that the dynamics and RNA binding are temperature dependent (and temperature was not specified) and that the RNA concentration used in the holo experiments was insufficient to saturate protein binding leading to a mixture of apo and holo proteins in the "holo" measurements.

Supplementary Note 4: Error propagation of the uncertainty of the γ-factor on the FRET efficiency *E*.

To support the hypothesis that the spread of the reported FRET efficiency E values for the different MalE mutants is caused by inaccuracies of the experimental calibration, we derive below how the uncertainty of the detection efficiency correction factor γ propagates into the uncertainty of the measured FRET efficiency. We also investigated the propagation of error for all other correction factors (donor leakage, acceptor direct excitation, and background in the donor and acceptor channels) to show that only the propagated uncertainty of the γ-factor follows the parabolic trend of the observed uncertainty in Fig. 3e. Lastly, we discuss the relative contributions of the uncertainties of the different correction factors.

I. Error propagation of the γ-factor on *E*

To support the hypothesis that the spread of the reported FRET efficiency E values for the different MalE mutants is caused by inaccuracies of the experimental calibration, we derive how the uncertainty of the detection efficiency correction factor γ propagates into the uncertainty of the measured FRET efficiency. Using the nomenclature introduced by Hellenkamp et al.[13] , the apparent FRET efficiency before γ-factor correction ${}^{\mathrm{iii}}E_{\mathrm{app}}$ is given by:

$$ {}^{\mathrm{iii}}E_{\mathrm{app}} = \frac{F_{A|D}}{{}^{\mathrm{ii}}I_{\mathrm{D_{em}|D_{ex}}} + F_{A|D}}, \tag{4.1} $$

where ${}^{\mathrm{ii}}I_{\mathrm{D_{em}|D_{ex}}}$ is the background-corrected donor intensity after donor excitation, and $F_{A|D}$ is the cross-talk and direct excitation corrected acceptor fluorescence after donor excitation, given by:

$$ F_{A|D} = {}^{\mathrm{ii}}I_{\mathrm{A_{em}|D_{ex}}} - \alpha\,{}^{\mathrm{ii}}I_{\mathrm{D_{em}|D_{ex}}} - \delta\,{}^{\mathrm{ii}}I_{\mathrm{A_{em}|A_{ex}}}. \tag{4.2} $$

Here, ${}^{ii}I_{\mathrm{A_{em}|D_{ex}}}$ is the background-corrected acceptor intensity after donor excitation, ${}^{\mathrm{ii}}I_{\mathrm{A_{em}|A_{ex}}}$ is the background-corrected acceptor intensity after acceptor excitation, and α and δ are the correction factor for donor crosstalk and acceptor direct excitation respectively.

The fully corrected FRET efficiency E is then given by:

$$ E = \frac{F_{A|D}}{\gamma\,{}^{\mathrm{ii}}I_{\mathrm{D_{em}|D_{ex}}} + F_{A|D}} = \left(\gamma\left(\frac{1}{{}^{\mathrm{iii}}E_{\mathrm{app}}} - 1\right) + 1\right)^{-1} \tag{4.3} $$

Using standard error propagation, the uncertainty of the FRET efficiency E due to the uncertainty of the γ-factor is given by:

$$ \Delta E = \left|\frac{\partial E}{\partial \gamma}\right| \Delta\gamma, \tag{4.4} $$

where $\Delta\gamma$ is the uncertainty in γ, and $|\partial E/\partial\gamma|$ is the partial derivative of E with respect to γ. The partial derivative is given by:

$$\left|\frac{\partial E}{\partial \gamma}\right| = \frac{E(1-E)}{\gamma}$$

Thus, the propagated uncertainty is given by:

$$\Delta E = E(1-E)\frac{\Delta\gamma}{\gamma} \tag{4.5}$$

This equation was fit to the experimental data in Fig. 3e of the main text, yielding an estimated uncertainty of the γ-factor calibration of $\Delta\gamma/\gamma = 0.23$.

To estimate the absolute (rather than the relative) error in γ from the data, we calculated $\Delta\gamma$ directly from the reported FRET efficiency E and γ-factor for each lab by rearranging Eq. 4.5:

$$\Delta\gamma = \gamma\frac{\Delta E}{E(1-E)} \tag{4.6}$$

From the obtained values for $\Delta\gamma$ for the different labs, we obtain an average value of $\Delta\gamma = 0.071 \pm 0.051$ (mean ± standard deviation).

II. Error propagation of the remaining correction factors on *E*

Hellenkamp et al.[13] have provided explicit expressions for the propagation of all relevant error sources into the derived FRET-averaged distance $R_{DA}^{\langle E\rangle}$. Analogous expressions have been provided by Peulen et al.[1] Here, we compare the uncertainty of the FRET efficiency ΔE, which relates to the uncertainty of the distance as:

$$\Delta E = \left|\frac{\partial E}{\partial R_{DA}^{\langle E\rangle}}\right| \Delta R_{DA}^{\langle E\rangle}, \tag{4.7}$$

where $\Delta R_{DA}^{\langle E\rangle}$ is the distance uncertainty and $\left|\frac{\partial E}{\partial R_{DA}^{\langle E\rangle}}\right|$ is the absolute value of the partial derivative with:

$$E = \left(1 + \left(\frac{R_{DA}^{\langle E\rangle}}{R_0}\right)^6\right)^{-1} \tag{4.8}$$

and:

$$\left|\frac{\partial E}{\partial R_{DA}^{\langle E\rangle}}\right| = \frac{6E(1-E)}{R_{DA}^{\langle E\rangle}}. \tag{4.9}$$

For the propagated uncertainty of the correction factors for donor leakage (α) and acceptor direct excitation (δ) on E, we obtain according to Hellenkamp et al.[13]:

$$\Delta E_\alpha = \frac{(1-E)^2}{\gamma}\Delta\alpha, \tag{4.10}$$

$$\Delta E_\delta = (1-E)\beta\Delta\delta, \tag{4.11}$$

where β is the correction factor for the different excitation flux of the donor and acceptor fluorophores. For the contributions of constant background signal in the donor or acceptor channel (bgD and bgA, respectively), we obtain according to Hellenkamp et al.[13]:

$$\Delta E_{\mathrm{bgD}} = [\gamma E + \alpha(1-E)]\frac{\Delta I_D^{\mathrm{BG}}}{\langle F\rangle}, \tag{4.12}$$

$$\Delta E_{\mathrm{bgA}} = (1-E)\frac{\Delta I_A^{\mathrm{BG}}}{\langle F\rangle}, \tag{4.13}$$

where ΔI_D^{BG} and ΔI_A^{BG} are the uncertainties of the estimated background signal in the donor and acceptor detection channels and $\langle F\rangle$ is the average sum of the corrected donor and acceptor fluorescence collected during a single-molecule event.

The total uncertainty of the FRET efficiency is then given by:

$$\Delta E = \sqrt{(\Delta E_\alpha)^2 + \left(\Delta E_\gamma\right)^2 + (\Delta E_\delta)^2 + \left(\Delta E_{\mathrm{bgD}}\right)^2 + \left(\Delta E_{\mathrm{bgA}}\right)^2} \tag{4.14}$$

III. Identification of the uncertainty of the γ-factor as the dominant error source

Our assumption that the uncertainty of the γ-factor is the dominant source of calibration error is based on the following experimental, methodical, and theoretical arguments.

a) Inaccuracies of other correction factors propagate into the determined γ-factor

The corrections applied here and in Hellenkamp et al.[13] proceed through a step-wise workflow of background subtraction and correction for donor crosstalk and acceptor direct excitation, after which the correction factors γ and β are estimated from the apparent FRET efficiency and stoichiometry (see Online Methods). Any inaccuracies of the background count rates and correction factors α and δ must thus propagate to the correction factor γ. Hence, we argue that the γ-factor effectively consolidates the uncertainties of all correction factors and serves as an adequate measure for the total calibration uncertainty.

b) The γ-factor adds the largest uncertainty to E over the probed distance range

To compare the relative contributions of the different correction factors to the total uncertainty ΔE, we assess the expected propagated errors according to Eq. 4.5 and 4.10-14. For the absolute values of the correction factors α, β, γ and δ, we use the average values of the participating labs reported for the MalE system (Fig. 3d and Supplementary Table 1). For the relative uncertainties, we follow the estimates given in Hellenkamp et al.[13] of $\Delta\alpha/\alpha = 0.1$, $\Delta\gamma/\gamma = 0.1$ and $\Delta\delta/\delta = 0.1$. This results in the following values for the correction factors and their uncertainties: $\alpha = 0.048 \pm 0.005$, $\beta = 1.6$ (no uncertainty needed for propagation), $\gamma = 0.38 \pm 0.04$, and $\delta = 0.12 \pm 0.01$.

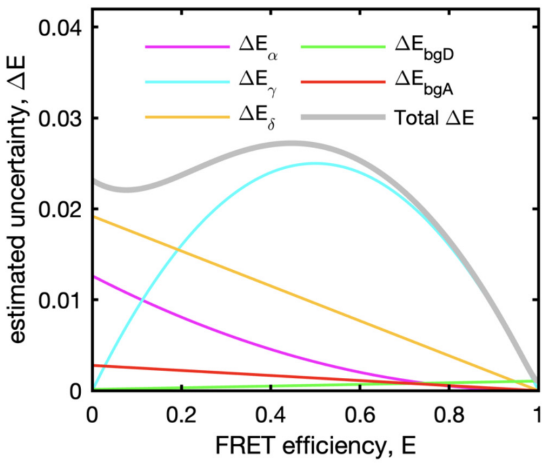
The uncertainty of the background signals ΔI_D^{BG} and ΔI_A^{BG} has previously been estimated to be of the order of ~1 photon per burst with a mean number of fluorescence photons per burst of $\langle F\rangle = 50$.[13] Here, we argue that the error due to background signal was previously overestimated. We found that an average of 90 ± 40 photons were detected by the participants over a typical burst duration of 1.7 ± 0.9 ms. Background count rates can generally be

quantified from buffer measurements or even extracted from burst measurements[4,14,15] with a high accuracy and reproducibility of $\Delta I^{BG}/I^{BG} \leq 0.1$. Given typical background count rates of 1-2 kHz, we expect an average of 2-3 background photons per burst with an absolute uncertainty of 0.2-0.3. Therefore, we assume $\Delta I_D^{BG} = \Delta I_A^{BG} = 0.25$ and $\langle F \rangle = 90$, leading to a lower value for the factor $\Delta I^{BG}/\langle F \rangle$ in Eq. 4.12-13 of 0.003 compared to the value used in Hellenkamp et al.[13] of $\Delta I^{BG}/\langle F \rangle = 0.02$. When the background is not constant but fluctuates, the uncertainty in the background will increase.

Based on the estimated correction factor values and uncertainties, we plot the propagated error of the different calibration factors in Supplementary Fig. SN4.1. Over the range of mean FRET efficiency values probed in this study for the MalE system (i.e. between 0.49-0.92), clearly the uncertainty caused by the γ-factor provides the largest contribution to the total calibration uncertainty. At FRET efficiency values below 0.25, the contribution of the correction factors for direct excitation (δ) and crosstalk (α) become significant. Note that we assumed a lower contribution of the background signal to the uncertainty compared to Hellenkamp et al.[13] as discussed above.

c) The observed trend of ΔE can only be described by an uncertainty in γ

The dependence of the observed uncertainty of the FRET efficiency ΔE on the FRET efficiency, E, clearly follows a parabolic trend (Fig. 3e). Based on the error propagation of the different correction factors, only the γ-factor follows a parabolic shape of the form $\Delta E_\gamma \propto E(1-E)$ (Eq. 4.5), while the remaining propagated uncertainties show either a monotonic decreasing or increasing behavior (Eq. 4.10-13). This supports the notion that the uncertainty of E is dominated by a calibration uncertainty of γ.



Supplementary Figure SN4.1. Error propagation of calibration uncertainties on the FRET efficiency, ΔE. The contributions of the uncertainty in different correction factors to the total uncertainty are given as colored lines (ΔE_α, crosstalk; ΔE_δ, direct excitation; ΔE_γ, γ-factor; and ΔE_{bgD} and ΔE_{bgA}, background in donor and acceptor channels). The total uncertainty, ΔE, is given as a gray line, evaluated according to Eq. 4.14. The following values and uncertainties were used for the correction factors: $\alpha = 0.048 \pm 0.005$, $\beta = 1.6$, $\gamma = 0.38 \pm 0.04$, $\delta = 0.12 \pm 0.01$, and $\Delta I^{BG}/\langle F \rangle = 0.003$.

Supplementary Note 5: Estimation of the experimental dynamic shift.

In BVA, the standard deviation of the apparent FRET efficiency, $\sigma_{E_{\mathrm{app}}}$, is plotted against the apparent FRET efficiency, E_{app} (Fig. 4a). For the E-τ plot, usually the accurate FRET efficiency E is plotted against the intensity-averaged donor fluorescence lifetime $\langle\tau_{D(A)}\rangle_F$ (Fig. 4b). For the estimation of the dynamic shift, $\langle\tau_{D(A)}\rangle_F$ is normalized with respect to the lifetime of the donor in the absence of the acceptor, $\tau_{D(0)}$, to constrain both axes to the interval of [0,1].

To determine the peak position of the single-molecule population, we fit the distribution of single molecule events in the E-$\langle\tau_{D(A)}\rangle_F/\tau_{D(0)}$ or BVA histograms using a superposition of M two-dimensional Gaussian distributions. Each Gaussian distribution is described by an amplitude A_i, a vector of the central coordinates $\boldsymbol{\mu_i}$, and a covariance matrix $\boldsymbol{\Sigma_i}$:

$$f(\boldsymbol{x}) = \sum_{i=1}^{M} A_i \cdot (2\pi)^{-1}|\boldsymbol{\Sigma_i}|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu_i})\boldsymbol{\Sigma_i}^{-1}(\boldsymbol{x}-\boldsymbol{\mu_i})^{'}\right), \tag{5.1}$$

where $\boldsymbol{\mu_i}$ is a row vector of length 2, and $\boldsymbol{\Sigma}$ is a *2*-by-*2* symmetric matrix whose diagonal elements and non-diagonal elements are the variances and covariances, respectively. Depending on whether the analysis is applied to the E-τ or BVA plot, the coordinate row vector $\boldsymbol{x}$ is given by $\boldsymbol{x} = \left(\langle\tau_{D(A)}\rangle_F/\tau_{D(0)}, E\right)$ or $\boldsymbol{x} = \left(E_{\mathrm{app}}, \sigma_{E_{\mathrm{app}}}\right)$, respectively. The model parameters are determined using a maximum likelihood estimation. The required number of populations needed to adequately describe the data is determined visually. When more than one bivariate normal distributions ($M > 1$) is needed to describe the data, the position of the component with the highest amplitude is taken. After the position of the major population is determined, we estimate the dynamic shift as described below.

For BVA, the dynamic shift is defined as the distance of the population center along the vertical (σ_{BVA}) axis from the static FRET-line, given by:

$$\sigma_{E_{\mathrm{app}}} = \sqrt{\frac{E_{\mathrm{app}}\left(1-E_{\mathrm{app}}\right)}{n}}, \tag{5.2}$$

where n is the number of photons per window used to estimate $\sigma_{E_{\mathrm{app}}}$ (here, $n = 5$). To provide consistency with the procedure proposed by Torella et al.[16], we compute the dynamic shift for the main population based on the average value of $\sigma_{E_{\mathrm{app}}}$ over a FRET efficiency interval given by its estimated mean and width, $[\mu_{E_{\mathrm{app}}} - w_{E_{\mathrm{app}}}, \mu_{E_{\mathrm{app}}} + w_{E_{\mathrm{app}}}]$, where $\mu_{E_{\mathrm{app}}}$ and $w_{E_{\mathrm{app}}}$ are the center position and width of the main population in the $\left(E_{\mathrm{app}}, \sigma_{E_{\mathrm{app}}}\right)$ plot.

For the E-τ plot, the dynamic shift is defined as the minimum distance from the peak of the population to the static FRET line, given by:

$$E = 1 - \frac{\langle\tau_{D(A)}\rangle_F}{\tau_{D(0)}}. \tag{5.3}$$

Generally, a linker correction for the static FRET-line is performed[17]. As a result, the static FRET-line is slightly curved. The dynamic shift is thus not necessarily the orthogonal distance to the static FRET line and was determined numerically. Lastly, the dynamic shift is assigned a positive or negative sign depending on whether the population lies above (positive) or below (negative) the static FRET line.

Supplementary Note 6: Prediction of the expected dynamic shifts.

Analytical expressions were derived to predict the maximum expected dynamic shift for a dynamic exchange between two conformational states with FRET efficiencies $E^{(1)}$ and $E^{(2)}$ in the E-τ and BVA plots.

I. The dynamic shift in the E-τ plot

For the E-τ plot, i.e. a plot of the intensity averaged FRET efficiency E against the normalized intensity averaged donor fluorescence lifetime $\langle\tau_{D(A)}\rangle_F/\tau_{D(0)}$, the maximum dynamic shift is given[18] :

$$ds_{(E-\tau)} = \frac{1}{\sqrt{2}}\left(\sqrt{1-E^{(1)}} - \sqrt{1-E^{(2)}}\right)^2. \tag{6.1}$$

This equation is derived by considering the maximum deviation between the ideal static FRET-line, given by:

$$E^{(\mathrm{stat})}_{(E-\tau)} = 1 - \frac{\langle\tau_{D(A)}\rangle_F}{\tau_{D(0)}}, \tag{6.2}$$

where $\langle\tau_{D(A)}\rangle_F$ and $\tau_{D(0)}$ are the intensity-weighted average donor fluorescence lifetimes in the presence and absence of the acceptor, respectively. Binary dynamic exchange between two limiting conformational states with FRET efficiencies $E^{(1)}$ and $E^{(2)}$ and the corresponding donor fluorescence lifetimes $\tau^{(1)}_{D(A)}$ and $\tau^{(2)}_{D(A)}$ is described by the dynamic FRET-line for exchange between two states[17]:

$$E^{(\mathrm{dyn})}_{(E-\tau)} = 1 - \frac{\tau^{(1)}_{D(A)}\tau^{(2)}_{D(A)}}{\tau_{D(0)}\left(\tau^{(1)}_{D(A)} + \tau^{(2)}_{D(A)} - \langle\tau_{D(A)}\rangle_F\right)}. \tag{6.3}$$

Using the relations:

$$\tau^{(i)}_{D(A)} = \tau_{D(0)}\left(1-E^{(i)}\right);\ \ i = 1,2 \tag{6.4}$$

we can then write Eq. 6.3 as a function of the FRET efficiencies:

$$E^{(\mathrm{dyn})}_{(E-\tau)} = 1 - \frac{\left(1-E^{(1)}\right)\left(1-E^{(2)}\right)}{\left(2 - E^{(1)} - E^{(2)} - \frac{\langle\tau_{D(A)}\rangle_F}{\tau_{D(0)}}\right)} \tag{6.5}$$

The difference between the static and dynamic FRET-lines along the FRET efficiency axis, Δ_E, as a function of $\langle\tau_{D(A)}\rangle_F$ is then given by:

$$\Delta_E = E^{(\mathrm{dyn})}_{(E-\tau)} - E^{(\mathrm{stat})}_{(E-\tau)} \tag{6.6}$$

$$\Delta_E(\langle\tau_{D(A)}\rangle_F) = \frac{\langle\tau_{D(A)}\rangle_F}{\tau_{D(0)}} - \frac{\left(1-E^{(1)}\right)\left(1-E^{(2)}\right)}{\left(2-E^{(1)}-E^{(2)}-\frac{\langle\tau_{D(A)}\rangle_F}{\tau_{D(0)}}\right)} \tag{6.7}$$

The function $\Delta_E(\langle\tau_{D(A)}\rangle_F)$ is unimodal and the maximum depends on the FRET efficiencies of the limiting states, $E^{(1)}$ and $E^{(2)}$:

$$\Delta_{E,max} = \left(\sqrt{1-E^{(1)}} - \sqrt{1-E^{(2)}}\right)^2. \tag{6.8}$$

We define the *dynamic shift*, $ds_{(E-\tau)}$, as the maximum deviation of the dynamic FRET-line measured orthogonal to the static FRET-line (see Fig. 4b), which introduces the factor of $1/\sqrt{2}$:

$$ds_{(E-\tau)} \stackrel{\text{def}}{=} \frac{\Delta_{E,max}}{\sqrt{2}} = \frac{1}{\sqrt{2}}\left(\sqrt{1-E^{(1)}} - \sqrt{1-E^{(2)}}\right)^2. \tag{6.9}$$

II. The dynamic shift in BVA

For the BVA plot, the maximum dynamic shift is defined as the maximum distance of the dynamic FRET-line from the static FRET-line along the y-axis, i.e., the estimated standard deviation of the apparent FRET efficiency, $\sigma_{E_{\text{app}}}$ (see Fig. 4a). The static FRET-line in BVA, describing the shot noise variance, is given by[16] :

$$\sigma_{E_{\text{app}}}^{(\text{stat})} = \sqrt{\frac{E_{\text{app}}(1-E_{\text{app}})}{n}}, \tag{6.10}$$

where n is the photon averaging window used to estimate the standard deviation. For the dynamic FRET-line in BVA describing the exchange between two limiting conformational states with apparent FRET efficiencies $E_{\text{app}}^{(1)}$ and $E_{\text{app}}^{(2)}$, we need to consider the excess variance caused by the conformational exchange. The contributions of shot noise, described by Eqn. 6.10, and the conformational exchange are not additive. The variance of the signal does not depend on the time dependence of the fluctuations. Thus, it is not important to know at what time points transitions between the limiting states occurred. It is only important what fraction of time the molecule spent in each individual state. For a two-state system, the expected variance in the presence of shot-noise and conformational dynamics is given by:

$$\begin{aligned}\text{Var}^{(\text{dyn})}\left(E_{\text{app}}\right) &= f_1\left[\frac{E_{\text{app}}^{(1)}\left(1-E_{\text{app}}^{(1)}\right)}{n} + {E_{\text{app}}^{(1)}}^2\right] \\ &+ f_2\left[\frac{E_{\text{app}}^{(2)}\left(1-E_{\text{app}}^{(2)}\right)}{n} + {E_{\text{app}}^{(2)}}^2\right] - \left(f_1 E_{\text{app}}^{(1)} + f_2 E_{\text{app}}^{(2)}\right)^2,\end{aligned} \tag{6.11}$$

where f_1 and f_2 are the fraction of time spent in the respective state with $f_1 + f_2 = 1$, and the apparent FRET efficiency is given by $E_{\mathrm{app}} = f_1 E_{\mathrm{app}}^{(1)} + f_2 E_{\mathrm{app}}^{(2)}$. The standard deviation of the apparent FRET efficiency for a dynamic exchange is then given by:

$$\sigma_{E_{\mathrm{app}}}^{(\mathrm{dyn})} = \sqrt{\mathrm{Var}^{(\mathrm{dyn})}\left(E_{\mathrm{app}}\right)}. \tag{6.12}$$

The dynamic FRET-line in BVA is then obtained by varying the fraction of time spent in the limiting states, $f_1 \in [0,1]$.

To derive Eq. 6.11, we start by expressing the variance as the difference between the expected value of the squared apparent FRET efficiency $\mathrm{E}\left[E_{\mathrm{app}}^2\right]$ and the square of the expected value $\mathrm{E}\left[E_{\mathrm{app}}\right]^2$, that is:

$$\mathrm{Var}\left(E_{\mathrm{app}}\right) = \mathrm{E}\left[E_{\mathrm{app}}^2\right] - \mathrm{E}\left[E_{\mathrm{app}}\right]^2. \tag{6.13}$$

Here, $\mathrm{E}[X^m]$ is the expected value of a continuous random variable X^m with $m \in \{1,2\}$, defined by:

$$\mathrm{E}[X^m] = \int X^m P(X)\, dX. \tag{6.14}$$

In BVA, the number of photons per sampling window, n, is constant. The probability to observe an apparent FRET efficiency, E_{app} , is thus equal to the probability of observing $n_A = nE_{\mathrm{app}}$ acceptor photons among the n detected photons, which is given by a binomial distribution:

$$P\left(E_{\mathrm{app}}\right) = P\left(n_A = nE_{\mathrm{app}} \middle| E_{\mathrm{app}}^{(i)}, n\right) = \binom{n}{n_A} {E_{\mathrm{app}}^{(i)}}^{n_A} \left(1 - E_{\mathrm{app}}^{(i)}\right)^{n-n_A}, \tag{6.15}$$

where n_A is the number of acceptor photons and $E_{\mathrm{app}}^{(i)}$ is the ideal apparent FRET efficiency of the molecule in state *i*.

We assume that the conformational dynamics are slow compared to the sampling frequency of the standard deviation, which is defined by the average time needed to detect n photons. Then, in the presence of conformational dynamics between two states with apparent FRET efficiencies $E_{\mathrm{app}}^{(1)}$ and $E_{\mathrm{app}}^{(2)}$, the probability to observe a given average apparent FRET efficiency E_{app} in a sampling window of n photons is given by the weighted average of the binomial distributions for the two limiting states:

$$P\left(E_{\mathrm{app}}\right) = f_1 P\left(n_A = n\, E_{\mathrm{app}} \middle| E_{\mathrm{app}}^{(1)}, n\right) + f_2 P\left(n_A = n\, E_{\mathrm{app}} \middle| E_{\mathrm{app}}^{(2)}, n\right), \tag{6.16}$$

where f_1 and f_2 are the probabilities that the molecule is in state 1 or 2, respectively (with $f_1 + f_2 = 1$). This follows because the molecule is found exclusively in one of the two limiting states during each sampling window of n photons.

We can now calculate the expected value of the apparent FRET efficiency E_{app} as:

$$\mathrm{E}\left[E_{\mathrm{app}}\right] = f_1 E_{\mathrm{app}}^{(1)} + f_2 E_{\mathrm{app}}^{(2)}. \tag{6.17}$$

The expected value of the squared apparent FRET efficiency, $\mathrm{E}\left[E_{\mathrm{app}}^2\right]$, is given by:

$$\begin{aligned}\mathrm{E}\left[E_{\mathrm{app}}^2\right] &= \int E_{\mathrm{app}}^2 P\left(E_{\mathrm{app}}\right) dE_{\mathrm{app}} \\ &= f_1 \int E_{\mathrm{app}}^2 P\left(n_A = n\,E_{\mathrm{app}} \middle| E_{\mathrm{app}}^{(1)}, n\right) dE_{\mathrm{app}} \\ &+ f_2 \int E_{\mathrm{app}}^2 P\left(n_A = n\,E_{\mathrm{app}} \middle| E_{\mathrm{app}}^{(2)}, n\right) dE_{\mathrm{app}}\end{aligned} \tag{7.18}$$

The integrals represent the expected value of the square, $\mathrm{E}[X^2]$, i.e., the second moments, for a binomial random variable. Using the definition of the variance as given in Eq. 6.13, $\mathrm{E}[X^2]$ can be calculated as:

$$\mathrm{E}[X^2] = \mathrm{Var}(X) + \mathrm{E}[X]^2 = np(1-p) + n^2p^2, \tag{6.19}$$

where n and p are the attempt number (i.e. number of photons in the averaging window, n) and success probability of the binomial process ($p = E_{\mathrm{app}}^{(i)}$), and X is the number of successes (here, $X = n_A = nE_{\mathrm{app}}$). The apparent FRET efficiency is given by $E_{\mathrm{app}} = n_A/n$. Thus, for state i, the expected value of the squared apparent FRET efficiency, $\mathrm{E}^{(i)}\left[E_{\mathrm{app}}^2\right]$, is given by:

$$\mathrm{E}^{(i)}\left[E_{\mathrm{app}}^2\right] = \frac{E_{\mathrm{app}}^{(i)}\left(1 - E_{\mathrm{app}}^{(i)}\right)}{n} + E_{\mathrm{app}}^{(i)}. \tag{6.20}$$

Finally, the variance of E_{app} is given by:

$$\begin{aligned}\mathrm{Var}^{(\mathrm{dyn})}\left(E_{\mathrm{app}}\right) &= \mathrm{E}\left[E_{\mathrm{app}}^2\right] - \mathrm{E}\left[E_{\mathrm{app}}\right]^2 \\ &= f_1 \left[\frac{E_{\mathrm{app}}^{(1)}\left(1 - E_{\mathrm{app}}^{(1)}\right)}{n} + {E_{\mathrm{app}}^{(1)}}^2\right] \\ &+ f_2 \left[\frac{E_{\mathrm{app}}^{(2)}\left(1 - E_{\mathrm{app}}^{(2)}\right)}{n} + {E_{\mathrm{app}}^{(2)}}^2\right] - \left(f_1 E_{\mathrm{app}}^{(1)} + f_2 E_{\mathrm{app}}^{(2)}\right)^2,\end{aligned} \tag{6.21}$$

where the fractions of the limiting states are defined by the observed average apparent FRET efficiency $E_{\mathrm{app}} = f_1 E_{\mathrm{app}}^{(1)} + f_2 E_{\mathrm{app}}^{(2)}$. This is Eqn. 6.11. The fractions are given by:

$$f_1 = \frac{E_{\mathrm{app}} - E_{\mathrm{app}}^{(2)}}{E_{\mathrm{app}}^{(1)} - E_{\mathrm{app}}^{(2)}}; \quad f_2 = 1 - f_1. \tag{6.22}$$

The dynamic shift in BVA is then defined as the maximum difference between the static and dynamic FRET-lines:

$$ds_{(BVA)} = \max_{E_{\mathrm{app}}} \left[\sigma_{E_{\mathrm{app}}}^{(\mathrm{dyn})}\left(E_{\mathrm{app}}\right) - \sigma_{E_{\mathrm{app}}}^{(\mathrm{stat})}\left(E_{\mathrm{app}}\right)\right]. \tag{6.23}$$

Since no analytical expression was found for Eq. 6.23, it was evaluated numerically.

Supplementary Note 7: R_0 determination for the Alexa546-Alexa647 dye pair for MalE.

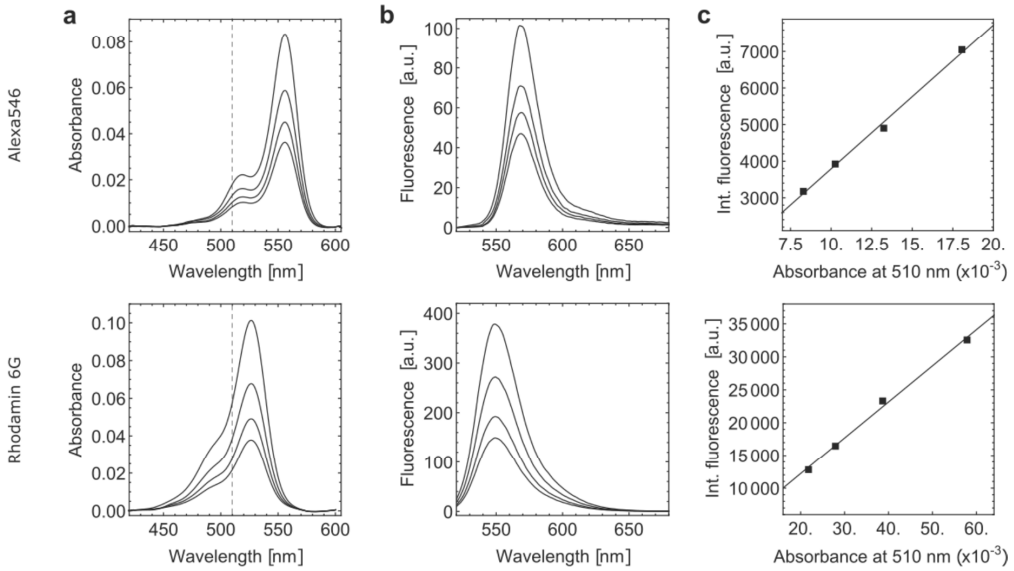
The Förster radius, R_0, is given by[19]

$$R_0{}^6 = \frac{9\,ln\,(10)}{128\,\pi^5\,N_A}\frac{\kappa^2}{n_{im}^4}\,Q_D\,\frac{\int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\,\lambda^4\,d\lambda}{\int_0^\infty F_D(\lambda)\,d\lambda}. \qquad (7.1)$$

where κ^2 is the orientation factor, Q_D the quantum yield of the donor, $F_D(\lambda)$ is the fluorescence emission spectrum of the donor, $\varepsilon_A(\lambda)$ is the absorption spectrum of the acceptor (scaled with the appropriate absorption coefficient), N_A is Avogadro's number and n_{im} is the index of refraction in the intervening medium. For an appropriate R_0 determination, it is necessary to determine the fluorescence quantum yield of the donor labeled to the molecule of interest as well as the overlap integral.

7.1 Fluorescence quantum yield $\Phi_{F,D}$

The fluorescence quantum yield of the donor dye Alexa546 covalently bound to the protein was determined using Rhodamine 6G as a reference.



Supplementary Figure SN7.1. Exemplarily fluorescence quantum yield determination for MalE-1. (a) The absorption spectrum of MalE-1 labeled with Alexa546 (top) and Rhodamine 6G (bottom) at 0.5, 0.625, 0.75 and 1 µM concentration. **(b)** The emission spectrum of MalE-1 labeled with Alexa546 (top) and Rhodamine 6G (bottom) at 0.5, 0.625, 0.75 and 1 µM concentration excited at 510 nm. **(c)** The integrated fluorescence of MalE-1 labeled with Alexa546 (top) and Rhodamine 6G (bottom) plotted against absorbance at 510 nm at 0.5, 0.625, 0.75 and 1 µM concentration and linear fit to the data points (line). Assuming a fluorescence quantum yield of 91% for Rhodamine 6G, the quantum yield for MalE-1 labeled with Alexa546 was determined using the ratio of the slopes.

The absorption (Supplementary Figure SN7.1a) and emission spectra (Supplementary Figure SN7.1b) at 510 nm excitation wavelength were measured for four different dye concentrations

in the range of 0.5-1.0 µM for the sub-stoichiometrically labeled donor only samples and compared to Rhodamine 6G in water. A linear fit to the integrated fluorescence intensity $\int_0^\infty F_D(\lambda)\, d\lambda$ for 510 nm excitation as a function of concentration provides two slopes: m_{Al546} and m_{R6G} (Supplementary Figure SN7.1c).

The fluorescence quantum yield of the donor Alexa546 is calculated from the slopes of the linear fit as

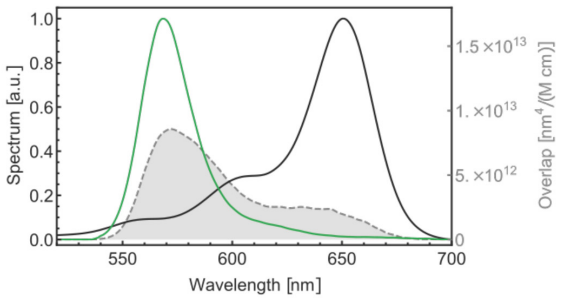
$$\boldsymbol{\Phi_{F,D}} = \frac{m_{Al546}}{m_{R6G}} \boldsymbol{\Phi_{F,R6G}} \tag{7.2}$$

where $\boldsymbol{\Phi_{F,R6G}} = 91 \pm 2\%$ is taken from literature[20]. The quantum efficiency of Alexa546 bound to the protein was found to be $72 \pm 4\%$. This is similar to the quantum yield of Alexa546 alone in PBS given in the molecular probes handbook of 0.79[21].

7.2 Overlap integral J

The overlap integral was calculated using the emission spectrum $F_D(\lambda)$ of the donor only sample and a normalized absorption spectrum of the acceptor $\bar{\varepsilon}_A$ scaled to the literature extinction coefficient $\varepsilon_A(\lambda) = \varepsilon_{A_{max}}\, \bar{\varepsilon}_A$ using:

$$J = \frac{\int_0^\infty F_D(\lambda)\varepsilon_A(\lambda)\, \lambda^4\, d\lambda}{\int_0^\infty F_D(\lambda)\, d\lambda}. \tag{7.3}$$



Supplementary Figure SN7.2. The overlap integral of Alexa546-Alexa647. The normalized emission spectrum of Alexa546 (green) and absorption spectrum of Alexa647 (black) plotted along with the overlap integral shown in gray for the MalE-1 mutant.

The overlap integral is illustrated in Supplementary Figure SN7.2 (gray area) resulting from the donor emission (green) and acceptor absorption (black) spectra.

7.3 Förster radius calculation

For calculating R_0, we have used the following values:

Orientation factor κ^2: 2/3
Averaged refractive index n_{im} (ref. [22]): 1.4
Extinction coefficient at maximum $\varepsilon_{A_{max}}$ (ref. [22]): 265,000 $OD/(M\, cm)$
Fluorescence quantum yield $\boldsymbol{\Phi_{F,D}}$: $72 \pm 4\%$
Overlap integral J: $7.0 \pm 0.1 \times 10^{15}\ nm^4/(M\, cm)$

(assuming that the given extinction coefficient is precise) and determined to be $R_0 = 65 \pm 3$ Å considering uncertainties in κ^2 of 10% and in n_{im} of 5%. This differs significantly from the manufacture's published R_0 of 74 Å[23].

The absorption and emission spectra were measured using singly labeled donor and acceptor mutants for MalE-1, MalE-2 and MalE-3. For other dye combinations used in this study, the following Förster radii were used: R_0 (Alexa546-Abberior STAR635P) = 62.5 Å, R_0 (Atto532-Atto643) = 59.0 Å and R_0 (Alexa488-Alexa647) = 52.0 Å.

Supplementary Note 8: Estimation of the residual anisotropies.

To improve the robustness of the analysis, the residual anisotropies were determined via two approaches, from the steady-state intensities ("*ss*") and from the time-resolved data ("*tr*"). The values reported in Fig. 5e-f of the main text correspond to an average value of the two estimates.

a) Residual anisotropies from steady-state intensities

Using single-molecule multiparameter fluorescence detection (MFD)[5], steady-state anisotropies, r_{ss}, in each detected single-molecule event were calculated from the background-corrected photon counts registered in the parallel and perpendicular detection channels as described in detail in reference [24]. Next, we plotted the obtained r_{ss} versus the intensity averaged fluorescence lifetime $\langle\tau\rangle_F$ to analyze their relation by the Perrin equation[25,26]. Here, we assume that the anisotropy decay can be described by two depolarization components[27]: (i) the global rotation of the protein, ρ_{global}, of approx. 20 ns, and (ii) a second component of ~ 0.5 ns describing the local dye motions, ρ_{linker}, with amplitudes $x_{\mathrm{global}} = 1 - x_{\mathrm{linker}}$ and x_{linker}, respectively:

$$r_{ss} = r_0\left(\frac{x_{\mathrm{linker}}}{1+\frac{\langle\tau\rangle_F}{\rho_{\mathrm{linker}}}} + \frac{1-x_{\mathrm{linker}}}{1+\frac{\langle\tau\rangle_F}{\rho_{\mathrm{global}}}}\right). \tag{8.1}$$

The residual anisotropy is defined as the amplitude of the global rotational motion as:

$$r_{\infty,ss} = r_0(1 - x_{\mathrm{linker}}), \tag{8.2}$$

where x_{linker} is given by:

$$x_{\mathrm{linker}} = \left(\frac{r_{ss}}{r_0} - \frac{1}{1+\frac{\langle\tau\rangle_F}{\rho_{\mathrm{global}}}}\right) \Big/ \left(\frac{1}{1+\frac{\langle\tau\rangle_F}{\rho_{\mathrm{linker}}}} - \frac{1}{1+\frac{\langle\tau\rangle_F}{\rho_{\mathrm{global}}}}\right), \tag{8.3}$$

and the fundamental anisotropy r_0 is assumed to be 0.38 for all dyes used here.

b) Residual anisotropies from the time-resolved anisotropy analysis

For the single-cysteine variants, the time-resolved anisotropy decay, $r(t)$, was obtained from ensemble TCSPC experiments. For some variants where single-cysteine mutants were not available, $r(t)$ was obtained from the single-molecule MFD experiments using a sub-ensemble selection of double-labeled molecules by applying an adequate value of the ALEX-2CDE filter. When needed, the bursts are additionally filtered using the stoichiometry value. From the selected double-labeled molecules, the fluorescence decays were obtained for the donor emission channel after donor excitation ($D_{ex}|D_{em}$) and the acceptor emission channel after acceptor excitation ($A_{ex}|A_{em}$). Note that, using this selection, the donor fluorescence lifetime is shortened by FRET to the acceptor. The anisotropy information was obtained from a global fit of the fluorescence decays $F_{VV}(t)$, $F_{VH}(t)$ and $F_{VM}(t)$, where VV and VH denote vertically (V)

and horizontally (H) polarized emitted light upon vertical excitation (V), respectively and VM represents the fluorescence lifetime decay measured at the magic angle (~54.7°) of the emission polarizer after vertical excitation. Here, the VM decay was not measured but rather constructed by mixing the VV and VH decays and with the experimentally determined g-factor and correcting for polarization mixing in the beam path of the high NA objective using the factors l_1 and l_2[24,28] as follows:

$$F_{VM}(t) = (1 - 3l_2)F_{VV}(t) + g_{VH/VV}(2 - 3l_1)F_{VH}(t) \tag{8.4}$$

Since correction for polarization mixing is only needed when using a high NA objective, they were not needed for the analysis of ensemble TCSPC measurements of single-cysteine mutants (Supplementary Table 11) and the l_1 and l_2 correction factors were set to zero. For the anisotropy analysis of sub-ensemble data (Supplementary Table 5) acquired using confocal single-molecule setups, values of l_1= 0.0308 and l_2 = 0.0368 were used for samples labeled with Alexa Fluor 546 or Atto532 as the donor dye, and l_1= 0.0175 and l_2 = 0.0526 for samples labeled with Alexa Fluor 488 as the donor dye. These values were determined with the following procedure using the enhanced Yellow Fluorescent Protein (eYFP). eYFP has a slow rotational correlation time of ρ = 16 ns, a monoexponential lifetime decay with τ = 3.2 ns and a fundamental anisotropy of r_0 = 0.39. To determine the l_1 and l_2 correction factors, we performed a global fit of polarization resolved signals of eYFP with fixed ρ and r_0. When τ is previously determined from a measurement at the magic angle, it can be fixed as well in the fit, reducing the number of free fit parameters.[21] Including the magic angle decay $F_{VM}(t)$ in a global analysis significantly improves the stability of the fit and increases the robustness of the fluorescence lifetime estimation.

The experimental decay curves $F(t)$ are described based on the idealized model decays $f(t)$. The relation between time-resolved anisotropy and polarization resolved fluorescence decays $f(t)$ is given as follows[28] :

$$f_{VV}(t) = \frac{1}{3} f_{VM}(t)\big(1 + (2 - 3l_1)r(t)\big); \tag{8.5}$$

$$f_{VH}(t) = \frac{1}{3} f_{VM}(t)\big(1 - (1 - 3l_2)r(t)\big), \tag{8.6}$$

The magic angle decay, $f_{VM}(t)$, is modelled by a multi-exponential function:

$$f_{VM}(t) = F_0 \sum_i x_i \exp\left(-\frac{t}{\tau_i}\right), \tag{8.7}$$

where τ_i denotes the fluorescence lifetimes and x_i are the fraction of molecules with the corresponding lifetime, and F_0 is a scaling factor that normalizes the integral of the model decay to the integrated experimental signal (i.e., the total measured counts). Similarly, the time-resolved anisotropy is modelled as:

$$r(t) = r_0\left(x_{\text{dye}}\exp\left(-\frac{t}{\rho_{\text{dye}}}\right) + x_{\text{linker}}\exp\left(-\frac{t}{\rho_{\text{linker}}}\right) + x_{\text{global}}\exp\left(-\frac{t}{\rho_{\text{global}}}\right)\right), \tag{8.8}$$

where ρ_j is the depolarization time of species j with anisotropy amplitude $b_j = r_0 x_j$. As above, we consider the global rotation of the molecule on timescales longer than 10 ns ($\rho_{\text{global}} >$ 10 ns) and the local mobility of the tethered label including the linker on the nanosecond scale, which is accounted for using two components, one for the linker movement (ρ_{linker} ~ 1-5 ns) and the other for fast rotation of the dye (ρ_{dye} ~ 0.3-0.5 ns). Due to the slow, global rotation, this equation becomes equivalent to the cone-in-cone model[29]. The contribution of a specific process to the overall depolarization of the fluorescence signal is determined by the species fraction x_j, with $\sum_j x_j = 1$. The fundamental anisotropy is taken to be $r_0 = 0.38$ for all dyes. The contribution of global rotation of the macromolecule to the total depolarization of emitted light represents the residual anisotropy, $r_{\infty,tr}$, given by $r_{\infty,tr} = r_0 x_{\text{global}}$.

The model decays are then convolved with the corresponding polarized component of the instrument response function (IRF):

$$F_{VV}(t) = f_{VV}(t) \otimes IRF_{VV}(t) + sc_{VV} IRF_{VV}(t) + B_{VV}, \tag{8.9}$$

$$F_{VH}(t) = g_{VH/VV}\, f_{VH}(t) \otimes IRF_{VH}(t) + sc_{VH} IRF_{VH}(t) + B_{VH}, \tag{8.10}$$

$$F_{VM}(t) = f_{VM}(t) \otimes IRF_{VM}(t) + sc_{VM} IRF_{VM}(t) + B_{VM}, \tag{8.11}$$

Here, B_{VV}, B_{VH} and B_{VM} account for constant background signal, the factors sc_{VV}, sc_{VH} and sc_{VM} describe the contribution of scattered laser light, and the factor $g_{VH/VV}$ corrects for the different detection efficiencies in the two detection channels VV and VH and $\otimes$ denotes a circular convolution. Here we apply a circular convolution because the full microtime range is used for the analysis, resulting in a periodic signal. Note that the idealize curves, $f(t)$, have been normalized to the total measured number of counts in Eqs. 8.5 and 8.6 via Eq. 8.7. Prior to the convolution, the IRF is corrected for uncorrelated background signal, e.g., due to the detector dark counts, and a time shift of the IRF is applied to correct for shifts that arise due to count rate differences between the IRF and fluorescence decay measurements. Similar to the $F_{VM}(t)$ decay, its corresponding $IRF_{VM}(t)$ was constructed by mixing $IRF_{VV}(t)$ and $IRF_{VH}(t)$ according to Eq. 8.4.

The quality of the fit was judged using the reduced chi-squared, χ_r^2. Analysis of the time-resolved anisotropy was done using the software package ChiSurf, a GUI-based suite composed of a collection of Python scripts for various types of fluorescence data analysis, available at https://github.com/Fluorescence-Tools/chisurf.

c) Computation of the average residual and combined anisotropies from steady-state and time-resolved measurements

For the steady-state anisotropy, assumptions regarding the rotational times are needed, which improves the robustness of this calculation. Time-resolved anisotropy requires no assumptions, but results in higher uncertainties. Hence, to increase the overall robustness of the analysis, we take the average of the residual anisotropies from the two approaches to compute the average residual anisotropy, $\langle r_\infty \rangle_{tr,ss}$:

$$\langle r_\infty \rangle_{tr,ss} = \frac{r_{\infty,tr} + r_{\infty,ss}}{2} \tag{8.12}$$

The uncertainty $\Delta\langle r_\infty \rangle_{tr,ss}$ was obtained as the standard deviation of the two estimates, corresponding to:

$$\Delta\langle r_\infty \rangle_{tr,ss} = \frac{|r_{\infty,tr} - r_{\infty,ss}|}{\sqrt{2}} \tag{8.13}$$

The combined residual anisotropy was then computed as a geometric average:

$$\langle r_{c,\infty} \rangle_{tr,ss} = \sqrt{\langle r_{\infty,D} \rangle_{tr,ss} \cdot \langle r_{\infty,A} \rangle_{tr,ss}}, \tag{8.14}$$

and the corresponding error was obtained using standard error propagation as:

$$\Delta\big(\langle r_{c,\infty} \rangle_{tr,ss}\big) = \frac{1}{2}\sqrt{\frac{\langle r_{\infty,D} \rangle_{tr,ss}}{\langle r_{\infty,A} \rangle_{tr,ss}}(\Delta\langle r_{\infty,A} \rangle_{tr,ss})^2 + \frac{\langle r_{\infty,A} \rangle_{tr,ss}}{\langle r_{\infty,D} \rangle_{tr,ss}}(\Delta\langle r_{\infty,D} \rangle_{tr,ss})^2}. \tag{8.15}$$

Supplementary Note 9: Calculation of distance uncertainties according to the "diffusion with traps" model.

Anisotropy measurements of fluorophores tethered to the surface of biomolecules indicate that the dye motion is commonly not isotropic despite the use of long, flexible linkers. Estimating a precise value of the orientational factor κ^2 is generally impossible as it would require knowledge of the mutual orientations of the transition dipole moments of the donor and acceptor molecules and their orientation with respect to the inter-dye distance vector. However, one can restrict the possible κ^2 values to a range that is compatible with the experiments with the help of the experimental anisotropy data. Previous approaches have estimated the uncertainty in κ^2 from the residual anisotropy in terms of rotational restrictions[30,31]. Here, we adopt a different approach termed the *diffusion with traps* (DWT) model[32] that provides an uncertainty in the measured donor-acceptor separation due to the unknown orientation factor, κ^2, based on the experimentally obtained fractions of trapped dye species and FRET efficiencies. The fraction of trapped dye species is obtained from the fluorescence anisotropy decays of the donor and acceptor fluorophores as a ratio of the residual, r_∞, to the fundamental anisotropy, r_0, which is often referred to as the second rank order parameter, $S_{A/D}^{(2)}$:

$$S_{A/D}^{(2)} = \frac{r_{\infty,A/D}}{r_{0,A/D}}, \tag{9.1}$$

where the subscript denotes that the order parameter is computed either for the donor (D) or acceptor (A) fluorophore.

Inspired by observations from MD simulations, the DWT model assumes that the fluorophores are either completely free, i.e., undergoing isotropic reorientation, or remain immobile for an extended period of time, meaning that they are trapped. It is also assumed that the exchange between these two species is slower than the donor fluorescence lifetime (i.e., static on the timescale of FRET) but faster than the diffusion time through the confocal volume in smFRET experiments of ~1-5 ms. This approximation is supported from MD data where the residence times of these states are found to be on the order of 10-100 ns or longer.

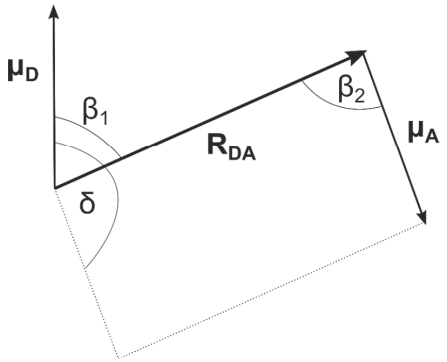
The DWT model assumes a mixture of trapped and free dye species and considers all four mobility scenarios for the donor and acceptor when computing the FRET efficiency, that is:

(I) $D_{\text{free}} - A_{\text{free}}$
(II) $D_{\text{free}} - A_{\text{trapped}}$
(III) $D_{\text{trapped}} - A_{\text{free}}$ and
(IV) $D_{\text{trapped}} - A_{\text{trapped}}$.

To calculate the range of possible κ^2 values, we generate random orientations of the dipole moment vectors, $\boldsymbol{\mu_D}$ and $\boldsymbol{\mu_A}$. For a given orientation of dipole vectors, the orientation factor κ^2 is given by[32]:

$$\kappa^2 = \frac{2}{3} + \frac{2}{3}S_D^{(2)}S_{\beta_1}^{(2)} + \frac{2}{3}S_A^{(2)}S_{\beta_2}^{(2)} + \frac{2}{3}S_D^{(2)}S_A^{(2)}\left[S_\delta^{(2)} + 6S_{\beta_1}^{(2)}S_{\beta_2}^{(2)} + 1 + 2S_{\beta_1}^{(2)} + 2S_{\beta_2}^{(2)} - 9\cos\beta_1\cos\beta_2\cos\delta\right] \quad (9.2)$$

where the angles β_1 and β_2 define the dipole orientations with respect to the donor-acceptor distance vector $\mathbf{R_{DA}}$ for D and A, respectively, and δ corresponds to the angle between the D and A transition dipole vectors, as illustrated below:



The second-rank order parameters for the angles are defined as follows:

$$S_\delta^{(2)} = \frac{1}{2}(3cos^2\delta - 1) \quad (9.3)$$

$$S_{\beta_1}^{(2)} = \frac{1}{2}(3cos^2\beta_1 - 1) \quad (9.4)$$

$$S_{\beta_2}^{(2)} = \frac{1}{2}(3cos^2\beta_2 - 1) \quad (9.5)$$

In the first step, generated are N random vectors, which represent random orientations of transition dipole vectors of donor and acceptor dyes. From those vectors, the angles β_1, β_2 and δ are calculated. Since we do not use the FRET sensitized acceptor anisotropy decay, δ is allowed to vary between 0° and 90°. However, it is possible to limit the range of possible δ - and κ^2- values by using the residual FRET sensitized acceptor anisotropy, $r_{\infty,A(D)}$.

In the next step, the FRET efficiency is computed as a function of κ^2 by considering all possible scenarios (I - IV) weighted by their respective probability of occurrence based on the estimated fractions of free/trapped dyes from the order parameters:

$$E(\kappa^2) = \left(1 - S_D^{(2)}\right)\left(1 - S_A^{(2)}\right)E + S_D^{(2)}S_A^{(2)}E\left(\kappa^2(D_{\text{trapped}} - A_{\text{trapped}})\right) + S_D^{(2)}\left(1 - S_A^{(2)}\right)E\left(\kappa^2(D_{\text{trapped}} - A_{\text{free}})\right) + \left(1 - S_D^{(2)}\right)S_A^{(2)}E\left(\kappa^2(D_{\text{free}} - A_{\text{trapped}})\right) \quad (9.6)$$

where $E = E(\kappa^2(D_{\text{free}} - A_{\text{free}})$ is the experimentally measured FRET efficiency, and:

$$E\left(\kappa^2(D_{\text{trapped/free}} - A_{\text{trapped/free}})\right) = \frac{1}{1 + \frac{2/3}{\kappa^2(D_{\text{trapped/free}} - A_{\text{trapped/free}})}\left(\frac{1}{E} - 1\right)}, \tag{9.7}$$

where the factor $2/3$ represents the isotropically averaged κ^2 value. The value of $\kappa^2(D_{\text{trapped/free}} - A_{\text{trapped/free}})$ is calculated according to Eq. 9.2 using the following assignment of the order parameters:

$$\begin{aligned} &\kappa^2\left(D_{\text{trapped}} - A_{\text{trapped}}\right)\text{: } S_D^{(2)} = S_A^{(2)} = 1 \\ &\kappa^2\left(D_{\text{trapped}} - A_{\text{free}}\right)\text{: } S_D^{(2)} = 1, \quad S_A^{(2)} = 0 \\ &\kappa^2\left(D_{\text{free}} - A_{\text{trapped}}\right)\text{: } S_D^{(2)} = 0, \quad S_A^{(2)} = 1 \end{aligned} \tag{9.8}$$

The dynamically averaged $\langle\kappa^2\rangle$ that corresponds to the obtained FRET efficiency $E(\kappa^2)$ is then obtained as:

$$\langle\kappa^2\rangle = \frac{2}{3}\frac{\frac{1}{E} - 1}{\frac{1}{E(\kappa^2)} - 1}, \tag{9.9}$$

where again the factor $2/3$ represents the isotropic average of κ^2.

This procedure is repeated for 10,000 randomly generated orientations of the dye vectors for a pre-defined distance, $R_{DA}(2/3) = R_0(1/E - 1)^{\frac{1}{6}}$, where E is the measured FRET efficiency and R_0 is the Förster radius assuming a κ^2 of $2/3^{\text{rds}}$. The obtained dynamically-averaged $\langle\kappa^2\rangle$ values are then converted into a normalized distance R_{app} according to the following equation:

$$R_{\text{app}} = \frac{R_{DA}(\langle\kappa^2\rangle)}{R_{DA}(2/3)} = \left(\frac{3}{2}\langle\kappa^2\rangle\right)^{-1/6} \tag{9.10}$$

The relative standard error in distance due to κ^2, $\Delta R_{\text{app}}(\kappa^2)$, is then defined as the standard deviation of the obtained R_{app} distribution normalized to its mean (Fig. 5e). Here, we assume all orientations of the fluorophore in the trapped states are possible. A more precise determination would require a detailed knowledge of the trapped states and the orientation distribution function of the dyes in the various traps.

Supplementary Note 10: Donor quenching estimation at different labeling positions on the MalE protein.

To explain the observed quenching at particular locations in MalE, we used coarse-grained Brownian dynamics simulations as described previously[1]. First, we identified the positions of the amino acids that tend to quench. These are typically MET, TYR, TRP and HIS. Next, we assigned a quenching rate constant to each of them. We found that the experimental data is explained best with a quenching rate constant of k_q= 2 ns^{-1}. After the position and quenching rate of the problematic amino acids are identified, we determine all sterically allowed positions of the donor dye using Accessible Volume simulations[27]. Within the accessible volume, we simulate diffusion of a dye using Brownian dynamics. As it is known that dyes diffuse slower in the vicinity of biomolecule's surface due to the non-specific sticking interactions, a heterogeneous diffusion model was applied. In this model, the diffusion coefficient of D = 10 Å^2 ns^{-1} is decreased by a factor of 10 in the vicinity of the surface. For each dye position within the AV, we measured the distance between the dye and all C_β atoms of the protein. Whenever the minimum distance was below the threshold of $R_{surface}$ = 8 Å, the slow diffusion coefficient (D = 1 Å^2 ns^{-1}) was used. Next, for each dye position, we estimated the fluorescence properties by measuring the distance of the fluorophore to the amino acids that can quench. The quenching was approximated by a step function such that whenever the dye was found within a distance of R_{rad} = 8.5 Å, quenching occurs at the given rate of k_q. Hence, during quenching, the emission rate of the donor is given by $k_D = \tau_0^{-1} + k_q$ and, for distances > R_{rad}, the donor is unquenched $k_D = \tau_0^{-1}$. Simulations were performed using the open-source GUI version of *QuEST – Quenching Estimator* software (GitHub page: https://github.com/Fluorescence-Tools/quest). Parameters used in the simulations are summarized in the Supplementary Table SN10.1 and our estimated fluorescence quantum yields for the donor at different labeling positions of MalE protein are provided in Supplementary Table SN10.2.

Supplementary Table SN10.1. Parameters used for the coarse-grained MD simulations of Alexa546 quenching at the different labeling positions of MalE in the apo and holo states.

Dye species	**τ_0 / ns**	**Dye properties/ Å**		**D_{free} / Å² ns⁻¹**	**$R_{surface}$ / Å**	**$D_{surface}$ / Å² ns⁻¹**	**R_{rad} / Å**	**Quencher**	**k_q /ns⁻¹**	**BD simulations**	
Alexa Fluor 546	4.1	L_{length}	20.5	10	8	1	8.5	MET	2	Simulation time / µs	10
								TYR		Time step / ps	2
		L_{width}	4.5					TRP		Grid size / Å	0.5
		R_{dye}	3.5					HIS			

Supplementary Table SN10.2. Simulated fluorescence quantum yield $\phi_{F,D}$ and fraction of dye states that collided with quenching amino acids for Alexa546 dye on the MalE protein. Using the coarse-grained Brownian dynamics simulations (see used simulation parameters in Supplementary Table SN10.1), we estimated quenching of the donor dye at the labeling positions of MalE. The positions that show a prominent quenching in BD simulations are S352C (apo/holo), K34C(apo/holo) and T36C (apo/holo), as given by frequent collisions with quenching amino acids. Using species averaged fluorescence lifetime $\langle\tau\rangle_x$ obtained by fitting ensemble time-resolved lifetime measurements of single labelled cysteine mutants (Supplementary Table 11), one can determine experimental fluorescence quantum yields, according to $\phi_{F,D} = \phi_{F,D,ref}\frac{\langle\tau_{D(0)}\rangle_x}{\tau_{D(0)}}$, with $\phi_{F,D,ref}$ = 0.72 (see Supplementary Note 7) and $\tau_{D(0)}$ = 4.1 ns. As predicted by BD simulations, experiments confirm that at position S352C dye is prone to more frequent collisions with quenchers, resulting in lower fluorescence quantum yield compared to other labelling positions.

Position	**State**	**Fraction of collisions with quenchers / %**	**Simulated $\phi_{F,D}$**	**Experimental $\phi_{F,D}$***
K29C	apo	10.84	0.64	0.71
	holo	14.45	0.60	0.71
S352C	apo	54.77	0.50	0.66
	holo	42.73	0.52	0.67
D87C	apo	23.16	0.57	0.68
	holo	28.65	0.53	0.69
A186C	apo	20.95	0.62	0.69
	holo	20.94	0.61	0.69
A134C	apo	7.92	0.64	0.69
	holo	5.96	0.65	0.70
K34C	apo	38.62	0.58	-
	holo	32.88	0.58	-
T36C	apo	49.94	0.50	-
	holo	48.02	0.53	-
N205C	apo	28.94	0.52	-
	holo	25.08	0.55	-

*Ensemble lifetime measurements of the single-labeled cysteine mutants K34C, T36C and N205C are not available.

Supplementary Note 11: Determination of statistical significance of the dynamic shifts of MalE.

To assess the statistical significance of the excess dynamic shift obtained for the measured MalE mutants, we compared the distribution of the measured dynamic shift between the different labs to the dynamic shifts observed for dsDNA. We computed the *p*-value for each of them using a chi-square score. With the *p*-value, we test whether distributions of ds values differ between DNA samples (reference, *r*) and the measured MalE and U2AF2 samples (*s*). Our null hypothesis is that measured MalE and U2AF2 samples appear as static as DNA. The *p*-value is the probability that the data is observed given the null hypothesis and can be used to decide whether the null hypothesis is valid or should be rejected given. Small *p*-values, particularly those below 0.05 (corresponding to a 2σ confidence interval), leads to the rejection of the null hypothesis.

Assuming that the reference DNA samples and samples under the tested hypothesis (MalE and U2AF2 variants) both have normally distributed *ds* values with mean $\langle ds \rangle_r$, $\langle ds \rangle_s$ and standard error of the mean SEM_r, SEM_s respectively, the chi-square test is then performed as follows:

$$\chi^2 = \left(\frac{\langle ds \rangle_r - \langle ds \rangle_{tot}}{SEM_r} \right)^2 + \left(\frac{\langle ds \rangle_s - \langle ds \rangle_{tot}}{SEM_s} \right)^2 \tag{11.1}$$

with $\langle ds \rangle_{tot}$ being:

$$\langle ds \rangle_{tot} = \frac{\frac{\langle ds \rangle_r}{SEM_r{}^2} + \frac{\langle ds \rangle_s}{SEM_s{}^2}}{\frac{1}{SEM_r{}^2} + \frac{1}{SEM_s{}^2}} \tag{11.2}$$

Using the probability density function of χ^2 distribution:

$$f(\chi^2 | N_{dof}) = \frac{1}{2^{\frac{N_{dof}}{2}} \Gamma\left(\frac{N_{dof}}{2}\right)} (\chi^2)^{\frac{N_{dof}}{2}-1} e^{-\frac{\chi^2}{2}} \tag{12.3}$$

where N_{dof} is number of degrees of freedom (N_{dof} =$N_{measurements} - N_{fit.params}$; here, $N_{dof} = 1$), one can determine the *p*-value, also referred to as *significance* or *certainty* that a reference distribution (DNA) does not match the sampled one (MalE, U2AF2). The p-value quantifies how unlikely it is to obtain a χ^2 value that is larger than the one observed between reference and measured sample. Commonly, the *p*-value is defined as the area under the right tail of the $f(\chi^2 | N_{dof})$ function:

$$p = \int_{\chi^2}^{+\infty} f(\chi^2 | N_{dof}) d\,\chi^2 \tag{11.4}$$

Using the described methodology in Supplementary Note 5, three labs estimated the apparent dynamic shift, *ds*, for different dye combinations of MalE and U2AF2 mutants (Supplementary

Table 8). Additionally, one lab determined *ds* values for the DNA rulers, that would serve as a reference value of what would be the observed *ds* value for static systems. Furthermore, using *ds* values reported from all three labs, we calculated *p*-values for all measured variants (Supplementary Table SN11.1). *p*-values below 0.05 are obtained for the apo state of MalE-1, MalE-5 and U2AF2, and for the holo state of MalE-1, MalE-2, MalE-4, MalE-5 and U2AF2. For those samples, the distribution of *ds* values cannot be explained with the one of reference DNA sample and the samples appear to be more dynamic compared to dsDNA.

Furthermore, *p*-values were computed using the residual dynamic shift values after filtering out dyes with pronounced sticking interactions. Significant shrinking in average *ds* value was obtained mainly for MalE-1, where the *p*-value could not be estimated due to only one point being left after dye filtering. An increase in the *p*-value can be observed for a few mutants, meaning that, after removal of sticking artifacts, they do not appear significantly different compared to a static DNA sample. This effect is particularly pronounced for the MalE-4 apo/holo sample.

Supplementary Table SN11.1. *p*-values for MalE and U2AF2 samples before and after filtering dye-pairs with pronounced sticking interactions. One sided *p*-values were computed according to Eq. 11.4.

	***p*-value all dyes**		***p*-value dyes ($\Delta R_{app}(\kappa^2) \leq 10\%$)**	
	apo	**holo**	**apo**	**holo**
MalE-1	0.0004	0.0024	n.a*	n.a*
MalE-2	0.7841	0.0376	0.7841	0.0376
MalE-3	0.9290	0.8973	0.5959	0.8973
MalE-4	0.0679	0.0051	0.4009	0.2024
MalE-5	1.20e-6	1.38e-6	0.0002	1.08e-5
U2AF2	1.11e-52	0.0009	6.91e-36	0.0260

*p-value could not be estimated due to only one point being left after filtering dyes with pronounced sticking interactions

Supplementary Note 12: Estimation of conformational flexibility from the residual dynamic shift for MalE.

The dynamic shift of a population in the E-τ plot is defined as the minimum distance to the reference static FRET-line given by:

$$E = 1 - \frac{\langle \tau_{D(A)} \rangle_F}{\tau_{D(0)}}, \tag{12.1}$$

where $\langle \tau_{D(A)} \rangle_F$ is the intensity-weighted average donor fluorescence lifetime and $\tau_{D(0)}$ is the donor-only lifetime (see Supplementary Note 5). We use the normalized donor fluorescence lifetime $\langle \tau_{D(A)} \rangle_F / \tau_{D(0)}$ to quantify the dynamic shift as this ensures that both axes in the E-τ plot range from 0 to 1. We have previously derived an expression for the maximum dynamic shift of a two-state system from the static FRET-line in a plot of E against $\langle \tau_{D(A)} \rangle_F / \tau_{D(0)}$ as a function of the limiting FRET efficiencies of the two states, E_1 and E_2 (Eq. 6.1 and ref. [18]):

$$\mathrm{ds} = \frac{1}{\sqrt{2}} \left(\sqrt{1 - E_1} - \sqrt{1 - E_2} \right)^2 \tag{12.2}$$

Here, we use Eq. 12.2 to estimate the conformational flexibility from the measured dynamic shift under the following assumptions.

1) The dynamics are fast compared to the burst duration (< 100 µs) so that complete averaging occurs.

2) The two states are equally populated, that is the equilibrium constant of the dynamics, $K = k_{12}/k_{21}$, is equal to 1.

3) The dynamics are symmetric around the FRET-averaged mean interdye distance $R_{\langle E \rangle}$, i.e. the two states have interdye distances of $R_{\langle E \rangle} \pm \delta R$.

Under these assumptions, the dynamic population will be at a defined position off the static line that falls between the two limiting states. The dynamic shift then only depends on the amplitude of the distance fluctuation, δR:

$$\mathrm{ds}(\delta R) = \frac{1}{\sqrt{2}} \left(\sqrt{1 + \left(\frac{R_0}{R_{\langle E \rangle} - \delta R} \right)^6} - \sqrt{+ \left(\frac{R_0}{R_{\langle E \rangle} + \delta R} \right)^6} \right)^2 . \tag{12.3}$$

We numerically solve this equation to translate the measured dynamic shift of a given dye pair into an apparent distance fluctuation given the Förster radius R_0 and FRET-averaged interdye distance $R_{\langle E \rangle}$. The estimated distance fluctuations δR are given in Supplementary Table 8.

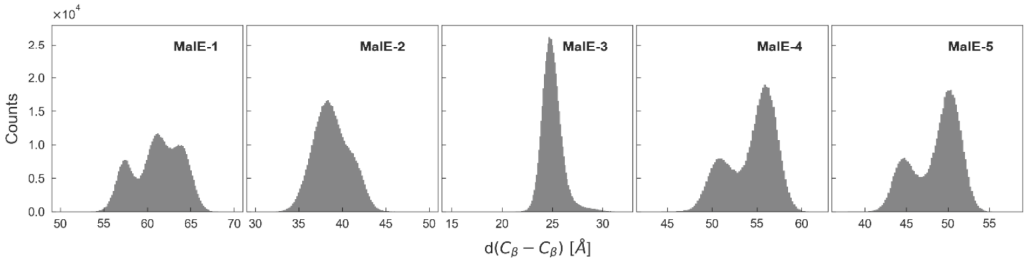
Supplementary Note 13: Molecular dynamics (MD) simulations of MalE.

Setup of topology and coordinate files. All-atom MD simulations were performed for MalE using the Amber18 suite[33] of programs with the FF14SB protein force-field[34] and the TIP4P-Ew water model[35]. For the initial structure, the crystal structure from PDB ID 1OMP was used. Using the LEaP program, the initial structure was solvated using TIP4P-Ew water molecules in an octahedral box such that the distance between the edge of water box and closest solute atom was at least 11Å. K^+/Cl^- ions were added to neutralize the system and then additional ions were added to reach a concentration of 50 mM for both ions. To allow for a larger time step of 4 fs, Hydrogen Mass Repartitioning (HMR)[36] was applied using ParmEd in the Amber software suite.

Minimization and Thermalization. After the topology and coordinate files were created and solvated, energy minimization of the system was performed while the positional restraints were applied to the solute atoms. Minimization was performed in two phases: first with high and then with low force constants for the positional restraints. In the first phase, the positional restraint was applied using a harmonic potential with a force constant of 25 kcal mol^{-1} $Å^{-2}$ and minimization was done through a total of 15,000 steps (5,000 steps of steepest descent minimization followed by 10,000 steps of conjugate gradient minimization). In the second phase, the force constant was reduced to 5 kcal mol^{-1} $Å^{-2}$ and minimization was done in the same manner as the first phase with 5,000 steps of steepest descent minimization followed by 10,000 steps of conjugate gradient minimization. After minimization, the system was heated from 100-300 K over 100 ps using NVT-MD simulations with the force constant of 5 kcal mol^{-1} $Å^{-2}$. The temperature change was applied in steps as follows. In the first 80 ps, the temperature was gradually raised from 100 K to the target 300 K. In the last 20 ps, the system was kept at the target temperature. To adjust the solvent density after increasing the temperature, heating was followed with a 300 ps of MD simulations using the NPT ensemble and the same force constant. Lastly, the force constant of the harmonic restraints was gradually reduced from 5 kcal mol^{-1} $Å^{-2}$ to 1 kcal mol^{-1} $Å^{-2}$ over 5×20 ps of NVT-MD simulations and then unrestrained NVT-MD production runs were performed.

Production runs. In total, 5 independent production replicas with a length of 2 µs each were performed. Conformations were saved every 20 ps. The SHAKE algorithm[37] was used in the production runs to constrain the bond lengths involving hydrogen atoms and long-range electrostatic interaction were treated using the Particle Mesh Ewald method[38] with a non-bonded cutoff of 8 Å.

Trajectories were analyzed using the MDTraj[39] library. The distances between the C_β atoms of the labeling positions for the five MalE mutants were calculated (Supplementary Figure SN13.1). For mutants MalE-1, MalE-4 and MalE-5, larger C_β-C_β distance fluctuations were observed.



Supplementary Figure SN13.1. Distributions of C_β-C_β distances from MD simulations of MalE proteins in the apo state. For each of the MalE mutants in the absence of fluorescent labels, the distance distribution between C_β atoms of the labeling positions is shown from a total of five production runs of 2µs length each. Consistent with smFRET experiments, a higher backbone-to-backbone distance fluctuations are detected for mutants MalE-1, MalE-4 and MalE-5. The obtained mean and standard deviations are $\langle d(C_\beta - C_\beta)\rangle_{MalE-1} =$ (61.2 ± 2.7) Å, $\langle d(C_\beta - C_\beta)\rangle_{MalE-2} =$ (38.8 ± 2.1) Å, $\langle d(C_\beta - C_\beta)\rangle_{MalE-3} =$ (25.0 ± 1.0) Å, $\langle d(C_\beta - C_\beta)\rangle_{MalE-4} =$ (54.4 ± 2.7) Å and $\langle d(C_\beta - C_\beta)\rangle_{MalE-5} =$ (48.6 ± 2.8) Å.

Supplementary Note 14: Theoretical limits for detecting dynamics.

The theoretical description of the dynamic shifts in the E-τ and BVA plots as given in Supplementary Note 5 allows us to assess the detection limit for conformational dynamics with respect to the uncertainty of the experiment. To define the detectability of a given shift from the static FRET-line in an experiment, it is necessary to quantify the experimental uncertainties associated with the FRET indicators E and $\langle\tau_{D(A)}\rangle_F$ for the E-τ plot, and E_{app} and $\sigma_{E_{\mathrm{app}}}^{(\mathrm{stat})}$ for the BVA plot. The experimental uncertainty contains contributions from the photon counting statistics of the experiment and systematic errors due to the calibrations or timescales of the dynamic exchange.

We first consider only the statistical error based on the standard error of the mean (σ_{SEM}) of a population, given by:

$$\sigma_{SEM} = \frac{\sigma_{SN}}{\sqrt{N_b}}, \tag{14.1}$$

where σ_{SN} is the shot-noise broadened width of a population and N_b is the number of single molecule bursts in the population. A given dynamic shift is considered detectable if it is larger than the SEM of a population on the static FRET-line.

I. Detection limit for dynamics in the E-τ plot

Shot-noise broadening in the E-τ plot

To simplify the expressions for the shot-noise broadening of a population in the E-τ plot, we assume that all detected single-molecule events have the same number of photons, N_p. For the FRET efficiency, the population width is then given by[40]:

$$\sigma_{SN}^{E} = \sqrt{\frac{E(1-E)}{N_p}}. \tag{14.2}$$

For the donor fluorescence lifetime $\tau_{D(A)}$, the variance of the estimator in the absence of background and for an instrument response with zero width (δ-function) is (ref [41,42]):

$$\mathrm{Var}\left(\tau_{D(A)}\right) = \frac{1}{N_p}{\tau_{D(A)}}^2\frac{k^2}{r^2}(1-e^{-r})\left(\frac{e^{\frac{r}{k}}(1-e^{-r})}{\left(e^{\frac{r}{k}}-1\right)^2} - \frac{k^2}{e^r-1}\right)^{-1}, \tag{14.3}$$

where k is the number of TCSPC detection channels and $r = T/\tau_{D(A)}$ is the ratio of the detection time window T and the fluorescence lifetime $\tau_{D(A)}$. The complex expression for the variance arises because only a part of the decay is seen in the finite time window T and due to the discretization of the time axis into k bins. In the limit of a large time window T and a high number of bins k, the expression simplifies to:

$$\lim_{r\to\infty,\ k\to\infty} \mathrm{Var}\left(\tau_{D(A)}\right) = \frac{1}{N_p}{\tau_{D(A)}}^2 \tag{14.4}$$

The same result is obtained by considering that the sum of independent exponential random variables δt_i (i.e. delay times) follows the Erlang distribution with variance:

$$\mathrm{Var}\left(\sum_{i=1}^{N_p} \delta t_i\right) = N_p {\tau_{D(A)}}^2 \tag{14.5}$$

from which the variance of the lifetime estimate is obtained as:

$$\mathrm{Var}\left(\frac{1}{N_p}\sum_{i=1}^{N_p} \delta t_i\right) = \frac{1}{N_p^2}\mathrm{Var}\left(\sum_{i=1}^{N_p} \delta t_i\right) = \frac{1}{N_p}{\tau_{D(A)}}^2 \tag{14.6}$$

Thus, for the normalized donor lifetime, we obtain the shot-noise limited width of the burst-wise distribution as:

$$\sigma_{SN}^{\tau} = \frac{\tau_{D(A)}}{\sqrt{N_p}\ \tau_{D(0)}} \tag{14.7}$$

To obtain the width of the distribution along the vector of the dynamic shift, i.e., orthogonal to the static FRET-line, the two contributions are combined geometrically:

$$\sigma_{SN}^{(E-\tau)} = \frac{1}{\sqrt{2}}\sqrt{{\sigma_{SN}^{E}}^2 + {\sigma_{SN}^{\tau}}^2}, \tag{14.8}$$

The standard error of the mean of the population, determined for example by Gaussian fitting, depends on the number of single molecule events in the population N_b, and is given by:

$$\sigma_{SEM}^{(E-\tau)} = \frac{\sigma_{SN}^{(E-\tau)}}{\sqrt{N_b}} \tag{14.9}$$

Effect of calibration error in the γ-factor

An incorrect calibration of the detection efficiency correction factor γ can lead to the detection of false-positive apparent dynamic shifts by shifting the population away from the static FRET line. However, only an underestimation of the γ-factor, leading to an overestimation of the FRET efficiency E, will shift the population above the static FRET-line, whereas an overestimation of γ will result in an unphysical shift below the static FRET-line. The propagated uncertainty of the γ-factor on the FRET efficiency is derived in Supplementary Note 4 (Eq. 4.5) and is given by:

$$\sigma_{E,\gamma} = E(1-E)\frac{\Delta\gamma}{\gamma}. \tag{14.10}$$

The calibration error, $\sigma_{E,\gamma}$, is combined with the shot-noise related uncertainty $\sigma_{SN}^{(E-\tau)}$ as:

$$\sigma_{SN,\gamma}^{(E-\tau)} = \sqrt{\sigma_{SN}^{(E-\tau)^2} + \sigma_{E,\gamma}^2}, \tag{14.11}$$

leading to an estimation of the combined standard error of the mean of:

$$\sigma_{SEM}^{(E-\tau)} = \frac{\sigma_{SN,\gamma}^{(E-\tau)}}{\sqrt{N_b}}. \tag{14.12}$$

II. Detection limit for dynamics in the BVA plot

Shot-noise broadening in the BVA plot

For BVA, we only have to consider the broadening along the y-axis, i.e., the uncertainty in the variance estimate. Assuming the apparent FRET efficiency, E_{app}, follows a Gaussian distribution with width parameter $\sigma_{E_{\mathrm{app}}}$, the estimated standard deviation in BVA, $\tilde{\sigma}_{E_{\mathrm{app}}}$, follows a scaled chi-squared distribution:

$$(M-1)\frac{\tilde{\sigma}_{E_{\mathrm{app}}}^2}{\sigma_{E_{\mathrm{app}}}^2} \sim \chi_{M-1}^2. \tag{14.13}$$

Here, $\sigma_{E_{\mathrm{app}}} = \sqrt{\frac{E_{\mathrm{app}}(1-E_{\mathrm{app}})}{n}}$ and M is the number of samples for the standard deviation estimate given by $M = N_p/n$, where N_p is the number of photons and n is the photon averaging window used for BVA. The upper 1σ confidence interval was used for the standard deviation estimate $\tilde{\sigma}_{E_{\mathrm{app}}}$ and is given by:

$$P\left((M-1)\frac{\tilde{\sigma}_{E_{\mathrm{app}}}^2}{\sigma_{E_{\mathrm{app}}}^2} \leq (M-1)\frac{\tilde{\sigma}_{E_{\mathrm{app}},\mathrm{UL}}^2}{\sigma_{E_{\mathrm{app}}}^2}\right) = T_M(1), \tag{14.14}$$

where $\tilde{\sigma}_{E_{\mathrm{app}},\mathrm{UL}}^2$ is the 1σ upper limit and $T_M(x)$ is the cumulative Student's t-distribution with M degrees of freedom. Assuming an average of N_p = 100 photons per burst and a photon averaging window of n = 5, then M = 20 and $T_{20}(1) \approx 0.84$. The upper limit $\tilde{\sigma}_{E_{\mathrm{app}},\mathrm{UL}}$ is given by:

$$\tilde{\sigma}_{E_{\mathrm{app}},\mathrm{UL}} = \sqrt{\frac{E_{\mathrm{app}}(1-E_{\mathrm{app}})}{n}}\sqrt{\frac{\chi_{\mathrm{inv},M-1}^2(T_M(1))}{M-1}}, \tag{14.15}$$

where $\chi_{\mathrm{inv},M-1}^2(x)$ is the inverse cumulative chi-squared distribution with $M-1$ degrees of freedom.

The shot-noise broadening around the true value on the static FRET-line is then given by:

$$\sigma_{SN}^{(BVA)} = \tilde{\sigma}_{E_{\mathrm{app}},\mathrm{UL}} - \sqrt{\frac{E_{\mathrm{app}}(1-E_{\mathrm{app}})}{N}}, \tag{14.16}$$

from which the standard error of the mean is obtained by dividing by the square root of the number of bursts (Eqn. 14.9).

Kinetic averaging in BVA

As BVA relies on the sampling of the FRET efficiency distribution from the photon time trace, it is subject to kinetic averaging during the sampling time, reducing the detectable variance. Thus, the dynamic shift is decreased for dynamics that occurs on the order of the sampling time or faster.

The sampling time, T, depends on the observed signal count rate, S, and the photon number, n, used to sample the FRET efficiency. The delay time between subsequent photon detection events is exponentially distributed. The time that it takes to detect n photons is then described by the Erlang distribution with a shape parameter given by n and the rate S. The average sampling time $\langle T \rangle$ is then given by:

$$\langle T \rangle = n/S\,. \tag{14.17}$$

For typical experiments, S = 100 kHz and n = 5, yielding a sampling time of $\langle T \rangle$ = 50 µs.

For BVA, it is assumed that the molecule is either in state 1 or 2 (with apparent FRET efficiencies $E_{\mathrm{app}}^{(1)}$ or $E_{\mathrm{app}}^{(2)}$) at each sampling point of the FRET efficiency. Here, we consider a two-state system with rates k_{12} and k_{21}:

$$E_{\mathrm{app}}^{(1)} \underset{k_{21}}{\overset{k_{12}}{\rightleftarrows}} E_{\mathrm{app}}^{(2)} \tag{14.18}$$

If the dynamics are fast, there is a certain probability that the molecule interconverts during the sampling time T, which reduces the observed variance. The average time for the molecule to convert from any state is given by the kinetic relaxation time τ_r:

$$\tau_r = \frac{1}{k_{12}+k_{21}} \tag{14.19}$$

The number of interconversions k during the sampling time T is Poisson distributed with the average value $\mu = T/\tau_r$:

$$P(k) = \frac{(T/\tau_r)^k}{k!} e^{-T/\tau_r}. \tag{14.20}$$

The probability to convert at least once during a time interval T is then given by:

$$P_{\mathrm{convert}} = P(k \geq 1) = 1 - P(k=0) = 1 - e^{-\frac{T}{\tau_r}}. \tag{14.21}$$

For simplification, we assume that, when an interconversion occurred, the sampling window converges to the average apparent FRET efficiency $\langle E_{\mathrm{app}} \rangle$. Then, any sampling of the apparent

FRET efficiency in which an interconversion occurred will have zero contribution to the variance and the dynamic contribution to the variance is simply reduced by the probability that no interconversion event occurred:

$$\mathrm{Var}'_{\mathrm{dyn}} = (1 - P_{\mathrm{convert}})\mathrm{Var}_{\mathrm{dyn}} = e^{-\frac{T}{\tau_r}}\mathrm{Var}_{\mathrm{dyn}}. \qquad (14.22)$$

By mixing the shot-noise and conformational contributions to the variance of the FRET efficiency, the dynamic FRET-line for BVA in the presence of fast conformational dynamics becomes:

$$\begin{aligned}\mathrm{Var}\left(E_{\mathrm{app}}, P_{\mathrm{convert}}\right) &= (1 - P_{\mathrm{convert}})\left(f_1\left[\frac{E_{\mathrm{app}}^{(1)}\left(1 - E_{\mathrm{app}}^{(1)}\right)}{n} + {E_{\mathrm{app}}^{(1)}}^2\right]\right. \\ &\left. + (1 - f_1)\left[\frac{E_{\mathrm{app}}^{(2)}\left(1 - E_{\mathrm{app}}^{(2)}\right)}{n} + {E_{\mathrm{app}}^{(2)}}^2\right]\right) + P_{\mathrm{convert}}\frac{\langle E_{\mathrm{app}}\rangle\left(1 - \langle E_{\mathrm{app}}\rangle\right)}{n} \\ &- \langle E_{\mathrm{app}}\rangle^2\end{aligned}$$

where f_1 and $(1 - f_1)$ are the fraction of time the molecule spends in state 1 and 2, respectively.

III. Discussion of detection limits

Using the described formalism, we tested whether hypothetical conformational dynamics between the apo and holo states of the studied systems, MalE-1 to 5 and U2AF, could be detected in a given situation. The FRET efficiencies of the apo and holo states were obtained from AV simulations on the available crystal structures as described in the online methods (Supplementary Table 10).

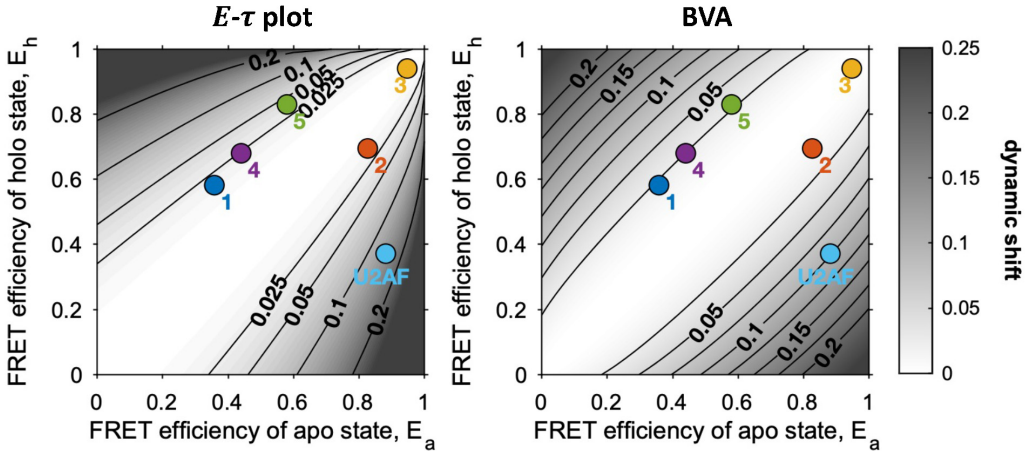
We first visualized the expected dynamic shift for the E-τ and BVA plots as a function of the FRET efficiencies of the limiting states (Supplementary Figure SN14.1). Along the diagonal, dynamic shifts are low (<0.025) because the FRET efficiency contrast is small whereas off-diagonal combinations of FRET efficiencies result in large dynamic shifts. While the BVA plot is symmetric between low and high FRET efficiencies, the E-τ plot is better at resolving a dynamic exchange at high FRET efficiencies of the limiting states compared to the low FRET efficiency region. This asymmetry arises because the uncertainty in the fluorescence lifetime is largest for long fluorescence lifetimes and thus for low FRET efficiencies.

The experimentally studied MalE mutants all showed dynamic shifts below 0.03 while, for U2AF, a dynamic shift of > 0.1 is expected for both the E-τ and BVA plots. To put these expected shifts into perspective with respect to the experimental uncertainty, we plot the dynamic shift normalized to the expected standard error of the mean for different numbers of detected photons per burst and number of bursts in Supplementary Figure SN14.2 for the E-τ plot and Supplementary Figure SN14.3 for the BVA plot. Dynamics are detectable when the ratio of the dynamic shift to the experimental uncertainty, ds/σ_{SEM}, exceeds one. Clearly, the

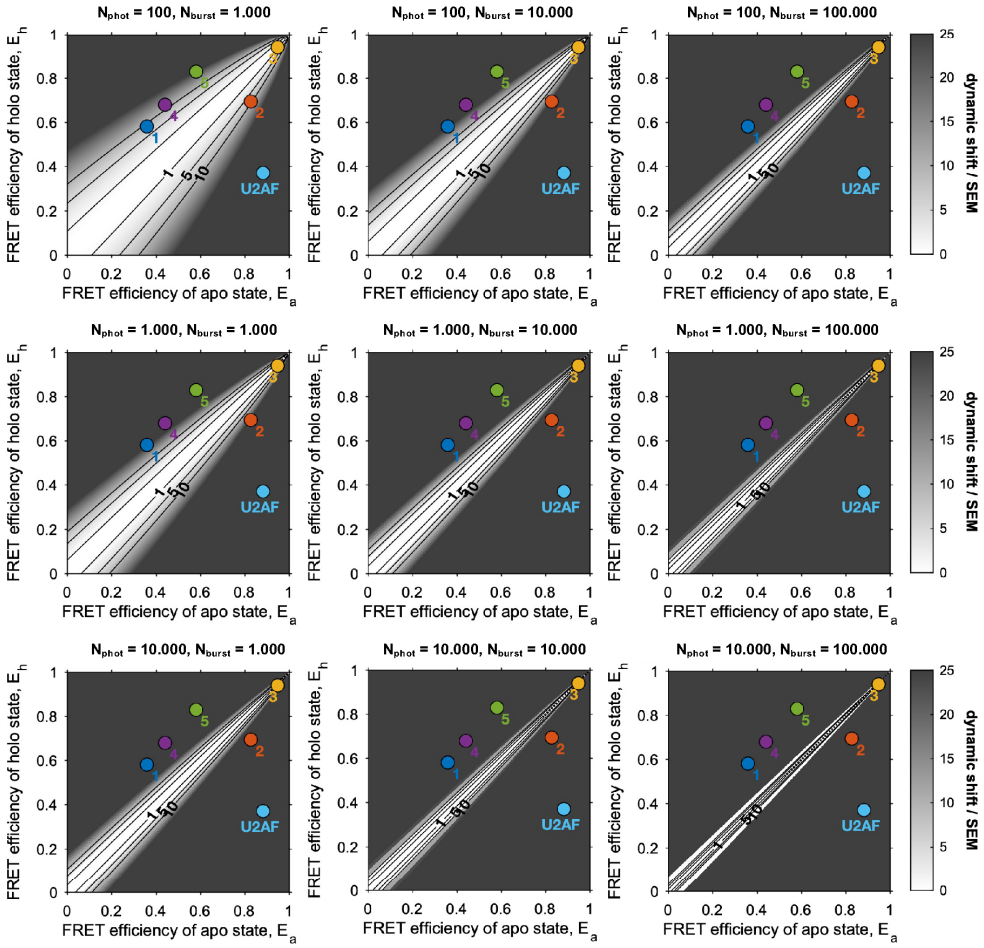
sensitivity to detect a dynamic exchange increases for both the E-τ and BVA plot with increasing number of photons and bursts. For typical experimental values (100 photons per burst and 1000 bursts), the dynamic exchange is predicted to be detectable as $ds/\sigma_{SEM} > 10$ for all experimental systems.

The sensitivity of the E-τ plot crucially depends on the accuracy of the correction factors used to compute accurate FRET efficiencies, in particular the detection efficiency correction factor γ. With higher uncertainty, $\Delta\gamma/\gamma$, the sensitivity of the E-τ plot decreases (Supplementary Figure SN14.4). At a relative uncertainty of $\Delta\gamma/\gamma = 0.1$, the sensitivity reduces to an extent that potential dynamics between the apo and holo states of the different MalE mutants would become undetectable. On the other hand, the large-scale dynamics for U2AF would remain detectable even at a high calibration uncertainty of $\Delta\gamma/\gamma = 0.3$. Note, the uncertainty discussed here is the uncertainty of the detection correction factor in a single lab and not the distribution of $\Delta\gamma/\gamma$ values calculated between labs in Fig. 3c.

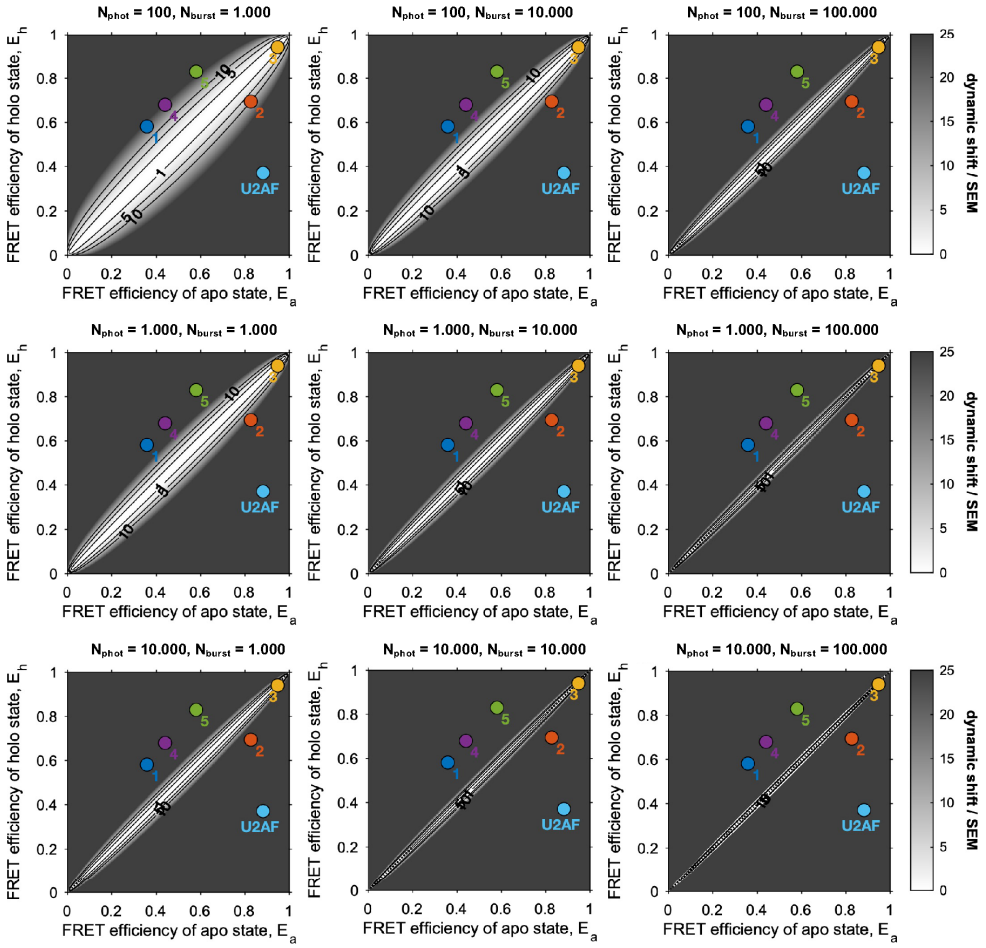
The BVA plot loses sensitivity when the timescale of dynamics approaches the sampling window used for the estimation of the variance of the FRET efficiency distribution (Supplementary Figure SN14.5). For slow conformational dynamics with a relaxation time of $\tau_r = 10$ ms, the potential exchange between the apo and holo states is detectable with $ds/\sigma_{SEM} > 10$. As the relaxation time approaches the sampling window (here, $T \approx 50$ µs), the sensitivity is reduced significantly. At $\tau_r = 20$ µs, the exchange for most MalE mutants is on the border of the detection limit while, for $\tau_r = 10$ µs, even potential dynamics of U2AF would become undetectable in BVA.



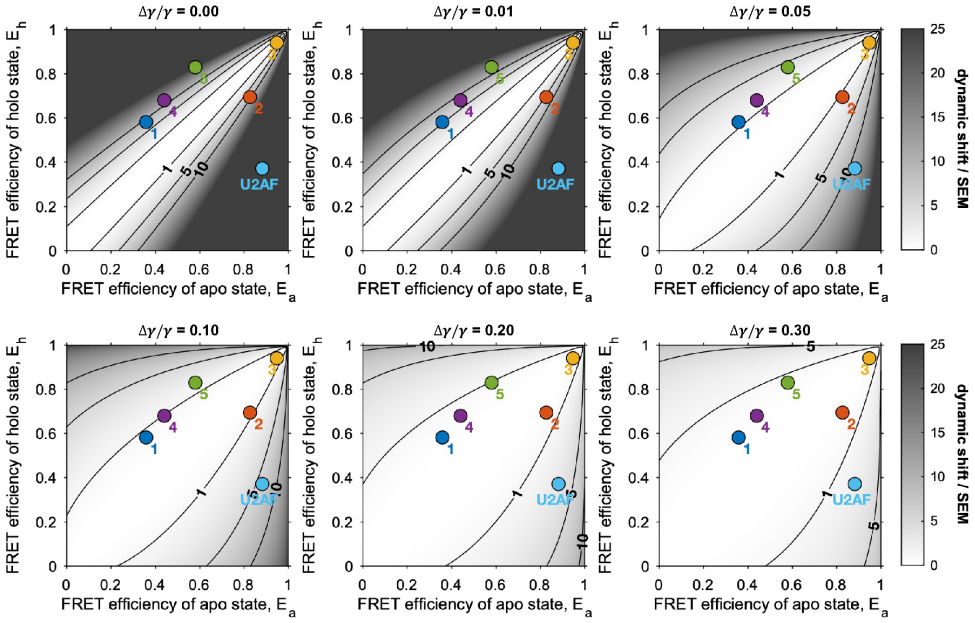
Supplementary Figure SN14.1: The dynamic shift in the E-τ (left) and the BVA (right) plots as a function of the FRET efficiencies of the limiting states (apo and holo) undergoing hypothetical dynamic exchange. The positions of the studied experimental systems are shown as colored markers (dark blue: MalE-1, red: MalE-2, yellow: MalE-3, purple: MalE-4, green: MalE-5, light blue: U2AF). The theoretical FRET efficiencies of the apo and holo states for the experimental systems were estimated from the PDB structures using AV simulations (see Online methods, Supplementary Tables 9 and 10).



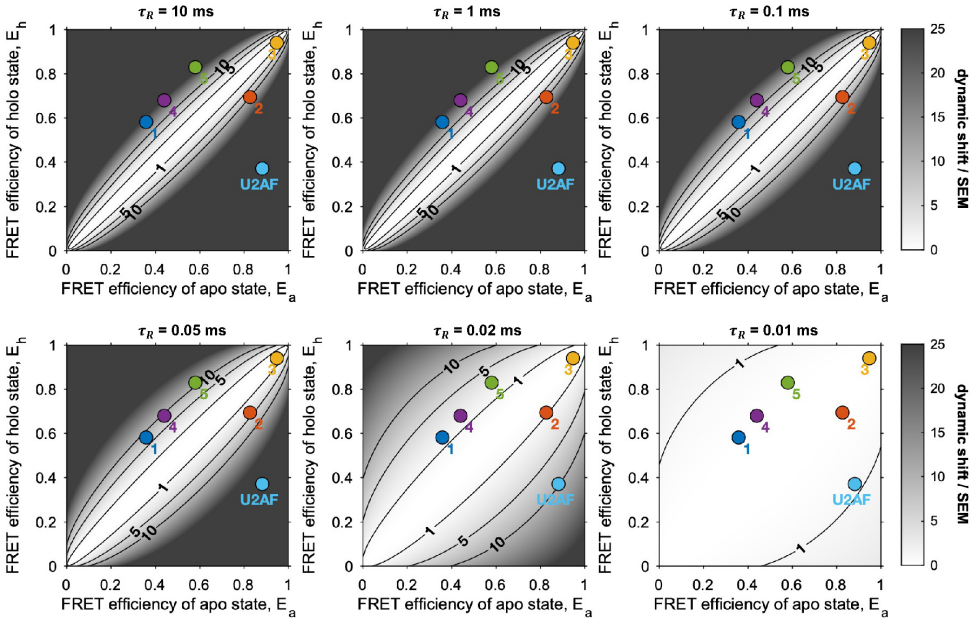
Supplementary Figure SN14.2: Detectability of conformational dynamics in the *E*-τ plot as a function of the photon-counting statistics. The detectability is defined as the ratio of the dynamic shift over the theoretical measurement uncertainty given by the standard error of the mean of the dynamic population. Dynamics are undetectable for ratios below one. The positions of the studied experimental systems are shown as colored markers (dark blue: MalE-1, red: MalE-2, yellow: MalE-3, purple: MalE-4, green: MalE-5, light blue: U2AF). The theoretical FRET efficiencies of the apo and holo states for the experimental systems were estimated from the PDB structures using AV simulations (see Online methods, Supplementary Tables 9 and 10).



Supplementary Figure SN14.3: Detectability of conformational dynamics in the BVA plot as a function of the photon-counting statistics. The detectability is defined as the ratio of the dynamic shift over the theoretical measurement uncertainty given by the standard error of the mean of the dynamic population. Dynamics are undetectable for ratios below one. The positions of the studied experimental systems are shown as colored markers (dark blue: MalE-1, red: MalE-2, yellow: MalE-3, purple: MalE-4, green: MalE-5, light blue: U2AF). The theoretical FRET efficiencies of the apo and holo states for the experimental systems were estimated from the PDB structures using AV simulations (see Online methods, Supplementary Tables 9 and 10).



Supplementary Figure SN14.4: Detectability of conformational dynamics in the E-τ plot as a function of the calibration error of the detection efficiency correction factor γ. The relative uncertainty of the γ-factor is given by $\Delta\gamma/\gamma$. The detectability is defined as the ratio of the dynamic shift over the theoretical measurement uncertainty given by the standard error of the mean of the dynamic population. Dynamics are undetectable for ratios below one. The positions of the studied experimental systems are shown as colored markers (dark blue: MalE-1, red: MalE-2, yellow: MalE-3, purple: MalE-4, green: MalE-5, light blue: U2AF). It is assumed that the dynamic population contains 1000 bursts of 100 photons. The theoretical FRET efficiencies of the apo and holo states for the experimental systems were estimated from the PDB structures using AV simulations (see Online methods, Supplementary Tables 9 and 10).



Supplementary Figure SN14.5: Detectability of conformational dynamics in the BVA plot as a function of the time scale of the dynamic exchange, quantified by the relaxation rate τ_r. The detectability is defined as the ratio of the dynamic shift over the theoretical measurement uncertainty given by the standard error of the mean of the dynamic population. Dynamics are undetectable for ratios below one. The positions of the studied experimental systems are shown as colored markers (dark blue: MalE-1, red: MalE-2, yellow: MalE-3, purple: MalE-4, green: MalE-5, light blue: U2AF). It is assumed that the dynamic population contains 1000 bursts of 100 photons and a count rate of 100 kHz, and that an averaging window of 5 photons is used for the BVA analysis. The theoretical FRET efficiencies of the apo and holo states for the experimental systems were estimated from the PDB structures using AV simulations (see Online methods, Supplementary Tables 9 and 10).

Supplementary Note 15: Model-free analysis of fluorescence decays for U2AF2.

Sub-ensemble fluorescence decays. Sub-ensemble donor fluorescence decays were generated from single molecule experiments by selecting the double-labeled population using the ALEX-2CDE filter[12] with an upper limit of 15. An additional stoichiometry cut was applied ($0.4 \leq S \leq 0.55$) to remove the dye-related artifact at high FRET efficiency that showed a higher stoichiometry of ~0.6 (Supplementary Figure 17). For the donor-only reference decays, a stoichiometry threshold of $S \geq 0.98$ was used.

Fitting procedure. The ideal fluorescence decays $f(t)$ were convoluted with the instrument response function and corrected for background contributions to the TCSPC pattern due to, for example, the contribution of scattered laser light, autofluorescence of the buffer, dark counts and uncorrelated background signal, to obtain the model decay $F(t)$:

$$F(t) = F_0 \cdot f(t) \otimes IRF(t - t_{\text{shift}}) + sc\, BG(t), \tag{15.1}$$

where $\otimes$ here denotes a linear convolution, F_0 is the initial amplitude of the fluorescence decay, $IRF(t)$ is the instrument response function (which is shifted by the time t_{shift}), $BG(t)$ is the normalized background/scatter pattern obtained from a buffer measurement and sc is the background/scatter amplitude, which is estimated from the cumulative duration of all analyzed bursts T_{bursts} and the background count rate c_{BG} as $sc = c_{BG} T_{\text{bursts}}$. Note that, contrary to Supplementary Note 8, we apply the linear convolution operation here for the analysis of smFRET data recorded using PIE because the analysis is performed only a selected time interval of the microtime histogram (i.e., a PIE channel) to restrict the analysis to the fluorescence decay of the donor. Hence, the signal here is not periodic as it is in the ensemble TCSPC experiments.

Fits were optimized using the reduced chi-square defined by:

$$\chi^2_{\text{red}} = \frac{1}{N_{\text{data}} - N_{\text{param}}} \sum \frac{\left(F_{exp}^{(k)} - F_{model}^{(k)}\right)^2}{w_k^2}, \tag{15.2}$$

where $F_{exp}^{(k)}$ and $F_{model}^{(k)}$ are the intensities of the measured and model decay in the TCSPC bin k, w_k is the respective weight given by $w_k = \sqrt{F_{exp}^{(k)}}$ based on the Poisson statistics of the detected signal, and N_{data} and N_{param} are the number of bins in the TCSPC histogram and the number of independent fit parameters, respectively.

Pre-fitting to estimate the background parameters. In a first step, we performed a global fit of the donor-only fluorescence decay $F_{DO}(t)$ and the FRET-induced donor decay $F_{DA}(t)$ using a two-component Gaussian distribution for the interdye distance. This allows us to estimate the fluorescence lifetimes and respective amplitudes of the donor-only sample as well as the parameters t_{shift} and F_0 of the FRET-induced donor decay. The donor-only decay is described using two lifetime components:

$$f_{DO}(t) = \sum_{i=1}^{2} x_{D(0)}^{(i)} \exp\left(-t/\tau_{D(0)}^{(i)}\right), \tag{15.3}$$

where $\tau_{D(0)}^{(i)}$ and $x_{D(0)}^{(i)}$ are the lifetime and fraction of donor-only species i. The presence of the acceptor acts as an additional process that depopulates the donor excited state at a rate of:

$$k_{RET}(R_{DA}) = \frac{1}{\tau_{D(0)}}\left(\frac{R_0}{R_{DA}}\right)^6, \tag{15.4}$$

where R_0 is the Förster radius and R_{DA} is the donor-acceptor separation distance. Note that $\tau_{D(0)}$ refers here to the lifetime of the species with the respective quantum yield that is used for the calculation of R_0, which can be different from the $\tau_{D(0)}^{(i)}$ obtained for the donor-only decay. For a given distance distribution, $p(R_{DA})$, the FRET-induced donor decay is then given by:

$$f_{DA}(t) = \left(1 - x_{\mathrm{DOnly}}\right) f_{DO}(t)\left(\int p(R_{DA}) \exp[-k_{RET}(R_{DA})\, t]\, dR_{DA}\right) + x_{\mathrm{DOnly}} f_{DO}(t), \tag{15.5}$$

where x_{DOnly} is the contribution of a donor-only signal due to acceptor photoblinking or photobleaching.

The two-component Gaussian distance distribution is given by:

$$p_{2G}(R_{DA}) = \sum_{i=1}^{2} x_{DA}^{(i)}\left(\sqrt{2\pi}\sigma_{DA,i}\right)^{-1} \exp\left[-\frac{\left(R_{DA} - \langle R_{DA}^{(i)}\rangle\right)^2}{2\sigma_{DA,i}^2}\right], \tag{15.6}$$

where $x_{DA}^{(i)}$ is the amplitude, $\langle R_{DA}^{(i)}\rangle$ the average interdye distance and $\sigma_{DA,i}$ the width of component i. The model is globally optimized with respect to the amplitude and lifetimes of the donor-only components. All parameters, except for the distance distribution $p(R_{DA})$, were fixed for the maximum entropy method model-free analysis discussed below.

Model-free analysis. The maximum entropy method (MEM) is an approach to extract the most unbiased distribution of a given parameter that provides a satisfactory fit to the experimental data[43–45]. Instead of minimizing the reduced chi-square, χ^2_{red}, the following functional is maximized:

$$\Theta = vS - \chi^2_{\mathrm{red}}, \tag{15.7}$$

where v is a constant scaling factor and S is the entropy functional of the parameter distribution. The entropy, S, of a discrete probability distribution p_i is defined by:

$$S = -\sum_i p_i \log\frac{p_i}{m_i}, \tag{15.8}$$

where p_i is the distribution of the parameter of interest and m_i describes the prior knowledge of the parameter distribution. We applied the MEM analysis to extract the distribution of interdye distances R_{DA}, $p(R_{DA})$. The measured FRET-induced donor fluorescence decay $f_{DA}^{exp}(t)$ is described as a superposition of exponential functions given by:

$$f_{DA}^{exp}(\mathrm{t}) = \left(1 - x_{\mathrm{DOnly}}\right) f_{DO}(t)\left(\sum_j p\left(R_{DA}^{(j)}\right) \exp\left[-k_{RET}\left(R_{DA}^{(j)}\right)t\right]\right) + x_{\mathrm{DOnly}} f_{DO}(t) \tag{15.9}$$

where the summation is performed over a range of interdye distances $R_{DA}^{(j)}$ from 10 to 150 Å using a step size of 0.7 Å. Maximization of Θ is performed as described in Vinogradov and Wilson[46] over a wide range of values for the regularization parameter v. The choice of the regularization parameter v was done by visual inspection of the L-curve plot of the negative entropy, -S, against the reduced chi-squared χ^2_{red}. The resulting values for χ^2_{red} and the regularization parameter v are given in Supplementary Table SN15.1. All analyses were performed using the *TauFit* module of the PAM software package[11].

Prior distribution. The prior distribution is based on the full apo ensemble derived in Huang et al.[47] For each of the 200 structures in the ensemble, the average interdye distance $\langle R_{DA}\rangle$ was determined using AV simulations[27,48]. To account for additional broadening due to the flexible dye linker, a kernel density estimate using a Gaussian kernel with a fixed width of 6 Å was performed to obtain the prior distribution as shown in Fig. 6d of the main text (light blue curve). We also performed a kernel density estimation of the interdye distance distribution without explicitly accounting for linker broadening using a Gaussian kernel by the *ksdensity* function of MATLAB, which computes the theoretically optimal bandwidth for normally distributed data[49]. This procedure returned a similar bandwidth of 5.7 Å. The different histograms and the kernel density estimate of the full apo ensemble are compared in Supplementary Figure SN15.1.

Deconvolution of the probability distribution obtained by the MEM analysis. The obtained distribution of the donor-acceptor distance R_{DA} from the MEM analysis is broadened due to the flexible dye linkers. The magnitude of this additional broadening has previously been characterized to be on the order of ~6 Å[27]. To approximate the underlying distribution of mean donor-acceptor distances $\langle R_{DA}\rangle$, we performed a deconvolution of the R_{DA} distribution obtained by MEM.

The deconvolution is performed using a Gaussian kernel with a width of σ_{DA} = 6 Å, defined on a distance grid of $\langle R_{DA}\rangle \in [10\,\text{Å}, 150\,\text{Å}]$ with a resolution of 0.5 Å. The kernel matrix $\boldsymbol{Q}$ is defined as:

$$\boldsymbol{Q} = \begin{pmatrix} g_Q\left(R_{DA}^{(1)}; \langle R_{DA}\rangle^{(1)}\right) & \cdots & g_Q\left(R_{DA}^{(1)}; \langle R_{DA}\rangle^{(M)}\right) \\ \vdots & \ddots & \vdots \\ g_Q\left(R_{DA}^{(N)}; \langle R_{DA}\rangle^{(1)}\right) & \cdots & g_Q\left(R_{DA}^{(N)}; \langle R_{DA}\rangle^{(M)}\right) \end{pmatrix}, \tag{15.10}$$

where the indices N and M represent the number of sampling points of the discrete distributions for R_{DA} and $\langle R_{DA}\rangle$, respectively, and the kernel functions are given by:

$$g_Q\left(R_{DA}^{(n)}; \langle R_{DA}\rangle^{(m)}\right) = \left(\sqrt{2\pi}\sigma_{DA}\right)^{-1} \exp\left[-\frac{\left(R_{DA}^{(n)} - \langle R_{DA}\rangle^{(m)}\right)^2}{2\sigma_{DA}^2}\right].$$

The discretized probability distribution of mean donor-acceptor distances $\langle R_{DA}\rangle$, $\boldsymbol{p}_{\langle \boldsymbol{R_{DA}}\rangle}$, is expressed as a row vector:

$$\boldsymbol{p}_{\langle \boldsymbol{R_{DA}}\rangle} = \begin{pmatrix} p_{\langle R_{DA}\rangle}^{(1)} \\ p_{\langle R_{DA}\rangle}^{(2)} \\ \vdots \end{pmatrix}, \tag{15.11}$$

and the discrete distribution of R_{DA}, $\boldsymbol{p}_{R_{DA}}$, is obtained as:

$$\boldsymbol{p}_{R_{DA}} = \boldsymbol{Q} \cdot \boldsymbol{p}_{\langle R_{DA} \rangle}. \tag{15.12}$$

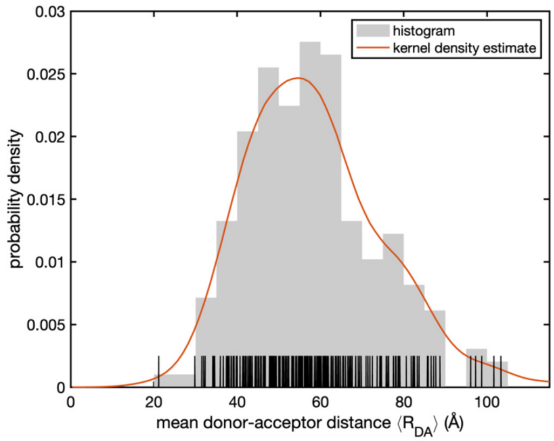
To obtain an estimate of $\boldsymbol{p}_{\langle R_{DA} \rangle}$, we minimized the absolute value of the difference between the measured distribution $\boldsymbol{p}_{R_{DA}}^{(\mathrm{exp})}$ and the distribution obtained by Eq. 15.12 above:

$$\min \left\| \boldsymbol{p}_{R_{DA}}^{(\mathrm{exp})} - \boldsymbol{Q} \cdot \boldsymbol{p}_{\langle R_{DA} \rangle} \right\|, \tag{15.13}$$

under the constraint that all elements of $\boldsymbol{p}_{\langle R_{DA} \rangle}$ must be positive. In addition, we used the kernel density estimate of the distribution of $\langle R_{DA} \rangle$ as obtained from the full apo ensemble as a starting point for $\boldsymbol{p}_{\langle R_{DA} \rangle}$. The optimization is performed using the *fmincon* function of MATLAB. The deconvoluted $\boldsymbol{p}_{\langle R_{DA} \rangle}$ distributions are shown in Supplementary Figure 14.

Supplementary Table SN15.1: Reduced chi-squared χ^2_{red} and regularization parameter v for the model-free analysis of fluorescence decays of U2AF2 measured in Lab #2.

	Dye Pair					
	Atto532-Atto643		Alexa546-Alexa647		Alexa488-Alexa647	
Sample	$\chi^2_{\mathrm{red.}}$	v	$\chi^2_{\mathrm{red.}}$	v	$\chi^2_{\mathrm{red.}}$	v
U2AF2 apo	1.30	5.09	1.28	3.20	1.29	2.21
U2AF2 holo	1.03	1.05	1.24	1.05	1.13	1.05



Supplementary Figure SN15.1: Probability distribution of the mean donor-acceptor separation obtained for the full apo-ensemble of U2AF2 as reported in Huang et al.[47] For each of the 200 structures in the ensemble, the mean donor-acceptor distance $\langle R_{DA} \rangle$ for the dye pair Atto532-Atto643 was computed from AV simulations. The vertical black lines indicate the individual values of $\langle R_{DA} \rangle$ and the gray bars shows the distance histogram computed with a bin width of 5 Å. The kernel density estimate (red line) was computed using a Gaussian kernel with a bandwidth of 6 Å, providing a smooth estimate of the probability density that is in good agreement with the histogram. The kernel density estimate was used as the prior distribution for the maximum entropy analysis of the fluorescence decays (Fig. 6d).

Supplementary Note 16: Filtered-FCS Analysis of the U2AF2 kinetics.

Filtered-FCS (fFCS) calculates the correlation functions for distinct species using statistical weighting based on the TCSPC patterns for the different species[50]. For the fFCS of the apo state of U2AF2 shown in Fig. 6e in the main text, the correlation functions were calculated as described in Barth et al.[51] using the PAM software package[11]. Briefly, two species with low and high FRET efficiency (LF, HF) were selected based on the FRET efficiency thresholds of $E \leq 0.6$ and $E \geq 0.9$, respectively. Filter functions for the fFCS analysis were calculated based on the concatenated TCSPC patterns of the donor and FRET-sensitized acceptor fluorescence decays. As a third species, a background pattern from a buffer measurement was included to account for the contribution of scattered laser light and constant background signal. "Purified" fFCS correlation functions were computed only for the detected and filtered double-labeled bursts. Signal occurring within 5 ms before or after the burst was included to obtain a better estimate of the diffusional part of the correlation function. To avoid detector afterpulsing in the species autocorrelation functions, the signal detected in the parallel and perpendicular channels of the MFD setup was filtered separately and cross-correlated. The two species autocorrelation functions (SACF) and two species cross-correlation functions (SCCF) were globally analyzed using a model with one diffusion term and two kinetic terms:

$$\begin{aligned} G(t_c) &= G_{\text{diff}}(t_c) G_{\text{kin}}(t_c) \\ G_{\text{diff}}(t_c) &= \frac{1}{\langle N \rangle}\left(1+\frac{t_c}{t_{\text{diff}}}\right)^{-1}\left(1+\frac{t_c}{p\,t_{\text{diff}}}\right)^{-\frac{1}{2}} \\ G_{\text{kin}}^{SACF}(t_c) &= 1 + A_1 e^{-\frac{t_c}{t_{R,1}}} + A_2 e^{-\frac{t_c}{t_{R,2}}} \\ G_{\text{kin}}^{SCCF}(t_c) &= 1 - A_1 e^{-\frac{t_c}{t_{R,1}}} - A_2 e^{-\frac{t_c}{t_{R,2}}} \end{aligned} \qquad (16.1)$$

Here, $G_{\text{diff}}(t_c)$ and $G_{\text{kin}}(t_c)$ are the diffusion and kinetic part of the correlation function $G(t_c)$ at lag time t_c. For the diffusion component, $\langle N \rangle$ is the average number of particles in the observation volume, t_{diff} is the diffusion time and the structural factor p is given by the ratio of the axial to the lateral width of the observation volume (typically, p = 5-10). For the kinetic component, $t_{R,1}$ and $t_{R,2}$ are the kinetic relaxation times with amplitudes A_1 and A_2 respectively. For the analysis, the kinetic relaxation times and diffusion time are optimized globally over all four curves. The amplitude of the diffusional part (given by $\langle N \rangle^{-1}$) and the kinetic terms (A_1 and A_2) were optimized individually for each of the four curves. This is necessary as the amplitude information in fFCS is less reliable compared to the time evolution of the curves due to the effect of imperfect filters[50].

Supplementary Note 17: Dynamics of U2AF2.

From previous work, we were aware that U2AF2 is a dynamic protein system whose conformational distribution can be shifted by the addition of an RNA ligand[52]. From the BVA and E-τ plots (Fig.4), it is evident that the apo protein is dynamic on the sub-millisecond timescale. The E-τ plot already contains information regarding the timescale of the dynamics of the sample with respect to the duration of the burst. A single transition is already sufficient to cause a shift for that burst from the static FRET line in the E-τ plot. Hence, dynamics that are up to an order of magnitude slower than the diffusion time can still be detected. In this case, individual "static" populations should be observable on the static FRET line with a smearing between the states coming from the few bursts where transitions occurred during the burst. For faster dynamics, several transitions occur during a burst and the individual "static" populations disappear with only a single, dynamically averaged state observable on the dynamic FRET line. The position of the population along the dynamic FRET line depends on the equilibrium between the different states[18]. For U2AF2, the dynamics are fast enough that we only observed an average FRET state and the individual populations are not visible.

The dynamics of U2AF2 is further complicated by the fact that, as already known from NMR and SAXS data, there is not a single closed or a single open conformation, but a family of conformations[47]. To get an idea of the distribution of FRET efficiencies in the sample, we performed a model-free analysis of the donor fluorescence lifetime using the maximum entropy method[43–45] to infer the underlying distribution of interdye distances (Supplementary Note 15). Our analysis revealed a higher population of compact states compared to the full structural ensemble of Huang et al[47] (Fig. 6d in the main text and Supplementary Fig. 14). We note that smFRET and NMR provide complementary information as the two techniques are sensitive to distinct distance ranges of 1-3 nm for paramagnetic relaxation enhancement (PRE) measurements and 4-12 nm for smFRET[53]. The structural ensemble was selected to satisfy the experimental averages of the measured PREs while the information from SAXS constrains the overall structural extension of the ensemble. The slight deviation of the smFRET information from the published structural ensemble indicates that a further refinement of the structural ensemble could be performed. We speculate that an alternative ensemble could be found that provides better agreement with the smFRET data while still satisfying the NMR/SAXS constraints[53].

After establishing the distribution of FRET efficiencies observable within the sample, we next determined the timescale of the dynamics. To get a model-independent estimation of the timescale of kinetics faster than the burst duration, we utilized an FCS approach. One possibility to visualize the dynamics is by performing a cross-correlation analysis between the donor and acceptor signals (FRET-FCS)[54,55]. We choose to use filtered-FCS (fFCS)[50,56] (Supplementary Note 16) as it increases the contrast of the FCS signal in comparison to FRET-FCS by weighing the correlation signals depending on the difference in the lifetime, color and anisotropy of fluorescence fluctuations between the bursts showing the lowest and highest FRET values. For the Atto532-Atto643 labeled protein, we observed two relaxation times with values of ~ 10 µs and ~ 200 µs (Fig. 6e, Supplementary Table 15). We assign the faster kinetics

to dynamics of the detached domains and the slower rate to interconversion between compact conformations within the conformational ensemble.

A second approach for extracting the dynamic timescales from a sample is to use the dynamic photon distribution analysis (PDA) method[17]. Dynamic PDA allows one to delineate the conformational heterogeneity over sub-millisecond transitions from the width of the FRET histograms beyond shot-noise by analyzing the raw photon counts[57,5817]. The FRET efficiency histograms are calculated using different time-windows for the binning and then fit using a global model (Fig. 6f in the main text and Supplementary Figure 16). There is little change in the FRET efficiency histograms when varying the integration time between 0.5 ms, 1.0 ms, 1.5 ms and 2.0 ms. From a global fit to the different histograms for the apo and holo conditions, the states as well as the conversion rates between them (provided they are between ca 100 µs and the timescale of diffusion) can be extracted. To simplify the complexity of the dynamic PDA model, we used a broad Gaussian function to empirically describe the fast, dynamic ensemble observed with fFCS and used dynamic PDA to look for slower kinetics. We observed a slow interconversion on the timescale of 10-100 ms between the very-high-FRET state at $E \sim 0.95$ (consistent with the closed NMR structure, PDB 2YH0[59]) and the main, dynamically averaged population at $E \sim 0.85$. The high FRET population (E = 0.85) is clearly distinct from a second peak in the E-τ plot of E ~ 0.95 and has a very short donor lifetime. The very-high-FRET signal is also reminiscent of what is observed for dye-dye interactions[60]. If this is true, the rates given by dynamic PDA should change when a different dye-pair is used whereas the protein dynamics should remain unchanged. Indeed, when evaluating measurements using the Alexa546 – Alexa647 dye pair, the fast dynamics remains unchanged (90 – 200 µs) whereas the amplitude of this very-high-FRET peak and the rate of the slow dynamics changes slightly (Supplementary Figure 19 and Supplementary Table 18). This demonstrates the importance of performing smFRET measurements with at least two different dye-pairs[60]. Here, the high stoichiometry state at very-high FRET efficiency was excluded and the observed dynamics suggest there is a transient stabilization of the closed structure.

Due to the complexity of the dynamics, we did not ask the various groups to provide a detailed analysis of the kinetics. We did ask laboratories, when possible, to provide an estimate of the timescale of the dynamics, when present. Five labs could contribute to a dynamic quantification of U2AF2 (Supplementary Table 15). These all correctly reported a quasi two-state behavior of the U2AF2 system. A comparison of kinetic rates based on further evaluation with dynamic PDA and fFCS for U2AF2 from different labs showed good consistency. Specifically, the calculated relaxation times for the apo state evaluated with fFCS were consistent with an approximately two-fold variation across labs (~200, ~320 and ~370 µs from three different labs). A dynamic PDA analysis for the holo-state kinetic rate estimations were not fully consistent most likely due to the variation seen in the FRET histograms (perhaps due to differences in temperature). Three among the five groups provided the kinetic rates for the holo -state with only ~20% variation of the reported relaxation times ($\tau_R = 1/(k_{12}+k_{21})$) of 1.25 ms, 1.42 and 1.6 ms (see Supplementary Table 15).

Challenges with the U2AF2 sample and analysis:

From the detailed analysis described above, it is clear that U2AF2 is not a simple two-state dynamic system. Apo U2AF2 exhibits dynamics on the microsecond timescale between a family of open and of closed conformations. Both apo and holo measurements exhibit a slow dynamic exchange between a very-high-FRET populations (at $E \sim 0.95$) and the other FRET states.

Beyond the complicated dynamics, other factors also impacted the analysis of the dynamic state. First of all, the measured RNA ligand concentration was not high enough to saturate binding to U2AF2. From the measured binding affinity of $K_d \approx 1.2$ µM (Supplementary Figure 7d), the ligand concentration of 5 µM used in these measurements lead to approximately 85% of the sample having a ligand and 15% remaining in the apo conformation. This corresponds well to the measured ratio of the two populations. An alternative explanation is that the protein exchanges between a quasi-static state and a dynamic conformation where one of the RRM domains transiently releases the RNA strand. However, there is no evidence that the RNA-bound state of U2AF2 interconverts with the apo-like conformation on the timescale of a burst (~ 10 ms). This is consistent with what is expected assuming a K_d of 1.2 µM and a measured k_{on} of 0.7 μM^{-1} s^{-1}, leading to a k_{off} of ~ 1s. This timescale is too slow to be detected with the solution-based measurements. Hence, we attributed the dynamics detected by BVA and the E-τ plots in the holo measurements to the significant presence of proteins not having an RNA ligand bound (Fig.5). Also, the timescale of the kinetics for the dynamic population was similar to that of the apo state (~ 100 - 200 µs). The second FRET peak with a FRET efficiency of $E \sim 0.44$ observed in the presence of RNA is close to the static line suggesting that the RNA bound conformation is static.

A second challenge in analyzing the U2AF2 data comes from the fraction of molecules with a higher than normal stoichiometry (Supplementary Figure 17). The origin of this population is not known. It could be due to dye-dye interactions where a change in fluorescence intensity yields a different stoichiometry value. Two populations with similar FRET efficiencies but different stoichiometries indicates inconsistencies in the dataset and most likely require individual γ factors for a proper analysis. At any rate, both populations should not be included in determining a single detection correction factor.

In summary, apo U2AF2 appears to undergo fluctuations between families of open and closed states with dynamics on the timescale of 100-200 µs. Dynamics on the millisecond timescale are also observed due to the formation of stable compact states. Fluctuations between an open and a closed conformation is a fair approximation of the dynamics, but an ensemble of open and closed conformations are needed to fully describe the measured data as shown in Fig. 6c-f.

Supplementary Note 18: Overview of set-ups and analysis software used across all labs.

Here, we provide detailed information regarding the different instrumentation used by the labs involved in the study. Important parameters are summarized in Supplementary Table SN18.1.

Lab#	Mode	Laser power measured at	Diffusion time τ_D (ms)	Donor P_D (µW)	Donor excitation wavelength (nm)	Filter used for donor detection (wavelength/ bandwidth)	[a]Background corrected Donor only Brightness and S/N[b] ratio of apo MalE-1 by Burst FCS (kHz)	Acceptor (µW)	Acceptor excitation wavelength (nm)	Filter used for acceptor detection (wavelength/ bandwidth)	[c]Background corrected Brightness and S/N ratio of direct acceptor excitation by Burst FCS (kHz)	[d]Background corrected Brightness and S/N ratio of direct acceptor excitation over a full trace (kHz)	Irradiance estimate ($P_D/\pi w_r^2 \sim P_D/4\pi D\tau_D$) ($W/m^{2}*10^{-3}$)
1	PIE	at the sample	1.202	45.1 94.6	532 485	582/75 525/50	290 / 201	23.3 22.3	640	700/75 670/30	153 / 163	99.2 / 105.5	3.11
2	PIE	at the sample	1.443	80 60	532 485	595/50 530/43	161 / 221	10	640	730/140 720/150	49.8 / 383	16.5 / 127	4.59
3	ALEX	at the sample	0.298	60	532	582/75	271 / 240	25	640	647 LP	407 / 947	222 / 517	16.66
4	PIE	before the objective	0.151	55	532	582/75	402 / 1436	50	640	647 LP + 750 SP	283 / 1415	154 / 770	30.15
5	ALEX	before the objective	0.485	60	532	582/75	435 / 659	25	640	700/75	243 / 451	133 / 246	10.24
6	PIE	-	1.182	-	530	582/75	170 / 32.6	-	640	700/75	159 / 90.2	125 / 71.0	-
7	ALEX	at the sample	0.419	100	515	571/72	208 / 274	230	638	679/40	404 / 1496	107 / 397	19.75
8	PIE	before the objective	0.257	60	531	582/75	183 / 201	40	640	700/75	108.5 / 213	113 / 222	19.32
9	PIE	at the laser exit port through the polarizer	1.586	35	560	612/69	144 / 86	100	660	700/75	57.9 / 170	23.9 / 70.3	1.83
10	ALEX	before the objective	2.349	250	532	585/70	40.1 / 16.0	60	640	650 LP	45.9 / 58.9	4.7 / 6.0	8.81
11	PIE	before the objective	0.541	110	530	582/75	307 / 236	90	640	700/75	174 / 218	108 / 135	16.82
12	PIE	at the sample	0.223	60	532	582/64	473 / 875	30	640	690/70	250 / 926	142 / 524	22.27
13	*POE	at the sample	0.239	40	532	582/75	51 / -	5	635	685/70	-	10 / -	13.85
14	PIE	at the sample	0.689	84	532	582/75	426 / 265	42	635	705/100	676 / 650	139 / 133	10.09
15	PIE	before the objective	0.336	50	532	585/65	395 / 238	25	635	700/75	198 / 267	111 / 150	12.32
16	ALEX	before the beamsplitter	0.743	100	532	580/60	81.7 / 102.1	35	638	665 LP	119 / 198	39.9 / 66.5	11.14

Supplementary Table SN18.1: Summary of the parameters regarding the setups used by the labs involved referring to the FRET subpopulations of the Alexa546-Alexa647 labeled apo MalE-1 samples. For the diffusion coefficient, we used the value of 96.2 $\mu m^2/s$.

[a]The donor-only brightness was determined by selecting donor-only molecules from the smFRET measurements and performing a burst-FCS analysis[54,61,62] on the selected burst with a 20 ms time window added to each side of the burst. From the amplitude of the autocorrelation function (ACF), the average number of molecules can be determined $\left(G(0) = \frac{\gamma_c}{\langle N \rangle}\right)$ where γ_c corrects for the geometry of the sample ($\gamma_c = 2^{-1.5}$ for a 3D Gaussian PSF). The molecular brightness at the center of the probe volume is then determined using the average intensity of the photons detected during the selected bursts divided by average number of molecules in the volume ($\varepsilon = \langle F \rangle / \langle N \rangle$). The molecular brightness was then corrected for background. This was done using the equation, $G_m(0) = \left(\frac{\langle F \rangle}{\langle F \rangle + \langle B \rangle}\right)^2 G(0)$, where $G_m(0)$ is the measured autocorrelation amplitude, $\langle F \rangle$ is the average fluorescence intensity coming from the labeled molecules, $\langle B \rangle$ is the background and $G(0)$ is the autocorrelation amplitude of the species of interest. The background was estimated from the distribution of inter-photon times determined from the actual measurement. Assuming that background does not contribute to the shape of the ACF, we correct the amplitude by dividing the ACF by the fractional intensity of the signal squared[5,63].

[b]S/N ratio: The signal/noise ratio was calculated as $\frac{\langle F \rangle}{\langle B \rangle}$ for the respective channel.

[c]We estimate the molecular brightness of the acceptor molecules using two approaches. Similar to the donor-only calculation, we performed burst FCS analysis on selected bursts as mentioned above. The amplitude of the ACF was corrected for background.

[d]We also calculated the molecular brightness of the acceptor using a full FCS analysis on all detected acceptor photons after acceptor excitation during the entire measurement. No burst selection was performed. The amplitude of the ACF was corrected for background. Note that the brightness obtained with burst FCS analysis is overestimated compared to the full FCS analysis as only molecules going near the center of the confocal volume are included in the burst analysis.

*POE: Pulsed Overlaid Excitation[64], similar to Periodic Acceptor Excitation[65], which uses cw donor excitation and bursts or pulses of acceptor excitation.

Lab#1

All sample solutions were measured in Lab-Tek I chamber slides at a concentration of ~50-100 pM. Single-molecule FRET experiments with MFD-PIE were performed on a homebuilt confocal setup as described previously[4].

For samples labeled with Alexa546-Alexa647 and Atto532-Atto643, our two-color green-red setup was used: Fluorescent donor molecules were excited by an amplified, frequency-doubled diode laser (PicoTA 530, PicoQuant, Berlin) at 532 nm and acceptor molecules were excited by a pulsed diode laser (LDH-D-C 640, PicoQuant) at 640 nm. Green laser was operated at a power of 45.1 µW and red laser at 23.3 µW, both measured at the sample, operated at a repetition rate of 26.66 MHz and synchronized with a delay of 18 ns. The laser light was guided into the epi-illuminated confocal microscope base (Nikon Eclipse TE300) and focused by a 60X water immersion objective (Plan Apo IR 60x/1.27, Nikon). For excitation, only parallel and perpendicular polarized light was selected using a Glan-Thompson polarizer (GTHM Polarizer, Thorlabs) before focusing into the objective. The emitted fluorescence was collected through the objective and spatially filtered using a pinhole with 75 µm diameter. The fluorescence emission was first separated for polarization using a polarizing beam-splitter (PBS3, Thorlabs) before being spectrally separated into donor and acceptor channels by a dichroic mirror (640DCXR; AHF Analysentechnik). Fluorescence emission was filtered (donor: Brightline HQ582/75, acceptor: Brightline HQ700/75, AHF Analysentechnik) and focused on avalanche photodiodes (SPCM-AQR, Perkin-Elmer) for detection. The detector outputs were recorded using four TCSPC cards (SPC154; Becker and Hickl)

For samples labeled with Alexa488-Alexa647, our three-color setup was used as described previously[66]. For this dye-pair combination, we switched off the 560 nm laser and associated detectors. In general, the three-color set-up has three pulsed lasers with a ~20 ns delay between the laser pulses (pulse frequency of 16.7 MHz) (PicoQuant, Germany; LDH-D- C-485, LDH-D-TA-560, LDH-D-C-640). The lasers were synchronized using a laser driver (PicoQuant, Germany; Sepia II). A 60x water immersion objective with 1.27 N.A. (Nikon, Germany; Plan Apo IR 60x 1.27 WI) was used for focusing the lasers into the sample. The measured laser powers at the sample were ~94.6 µW for blue, ~69.3 µW for green and ~22.3 µW for red laser. Fluorescence was collected by the same objective, separated from the laser excitations using a polychroic mirror (AHF Analysentechnik; zt405/488/633, Germany) and confocal geometry was achieved using a 50 µm pinhole. Light passing through the pinhole was further separated by a polarizing beam splitter (Thorlabs, Germany). Separation of blue and red wavelengths were performed with a dichroic mirror (AHF Analysentechnik; 640DCXR). Emission filters (AHF Analysentechnik; ET525/50, ET670/30) were placed right before the APD detectors (LaserComponents, 2x COUNT-100B for blue detection; Perkin Elmer, 2x SPCM-AQR14 for red detection). Photons were recorded and synchronized to the lasers pulses using a TCSPC module (PicoQuant; HydraHarp400).

Data analysis was performed using the PAM (PIE Analysis with Matlab) software package as described elsewhere[11]. Single-molecule events were identified using a sliding time window analysis with a threshold of 50 photons, a time window of 500 µs and a minimum photon number of 10. To remove photo-blinking and -bleaching events, the ALEX-2CDE filter was applied using an upper threshold of 12.

Lab#2

All sample solutions were measured in NUNC chambers (Lab-Tek, Thermo Scientific) with 500 µL sample volume and a pM concentration. Single-molecule FRET experiments with PIE were performed on a homebuilt confocal setup as described previously[26].

Setup #1 for samples labeled with Alexa Fluor 546 – Alexa Fluor 647, Alexa Fluor 546 – AbberiorSTAR635P and Atto532 – Atto643: fluorescent donor molecules are excited by a pulsed white light laser source (SuperK EXTREME, NKT Photonics) at 532 nm, using a modulator (SuperK Varia, NKT Photonics), or with a pulsed diode laser LDH-P-FA 530B. Donor excitation laser is operated with repetition rate of 25 MHz, and with power of 80 µW measured at the sample. Acceptor molecules are excited by a pulsed diode laser (LDH-D-C 640), operated at 25 MHz and 8 µW power measured at the sample. Laser light is guided into the epi-illuminated confocal microscope (Olympus IX71, Hamburg, Germany) by dichroic beamsplitter F68-532_zt532/640NIRrpo (AHF, Germany) and focused on a sample by a water immersion objective (UPlanSApo 60x/1.2w, Olympus Hamburg, Germany). Emitted fluorescence is collected through the objective and spatially filtered using a pinhole with 100 µm diameter and further split into parallel and perpendicular components using polarizing beam splitter cube (VISHT11, Gsänger). Light is then spectrally split into "green" and "red" spectral windows by a dichroic mirror (T640lpxr, AHF, Germany). Fluorescence emission was filtered (donor: 47-595/50 ET, acceptor: HQ 730/140, AHF, Germany) prior to detection using avalanche photodiodes (SPCM-AQRH 14, Excelitas). The detector outputs were recorded by a TCSPC module (HydraHarp 400, PicoQuant). To account for polarization mixing in the microscope objective, eYFP (enhanced Yellow Fluorescent Protein) was used as a calibration sample using the procedure described in Supplementary Note 8. The polarization correction factors were determined to be: l_1= 0.0308 and l_2 = 0.0368.

Setup #2 for samples labeled with Alexa Fluor 488 – Alexa Fluor 647: donor molecules were excited by a pulsed diode laser (LDH-D-C 485, PicoQuant) at 485 nm. Acceptor molecules were excited with pulsed diode laser (LDH-D-C 640, PicoQuant) at 635 nm. Lasers were operated with the repetition frequency of 32 MHz. Laser powers were measured at the sample and were 60 µW for donor excitation laser and 10 µW for acceptor excitation laser. Laser light is guided into the epi-illuminated confocal microscope (Olympus IX71, Hamburg, Germany) by dichroic beamsplitter FF500/646-Di01 (Semrock, USA), and focused on the sample by a water immersion objective (UPlanSApo 60x/1.2 NA, Olympus Hamburg, Germany). The emitted fluorescence is collected through the objective and focused on a 100 µm pinhole. Using a polarizing beam splitter cube, emitted light is divided into its parallel and perpendicular components. This is then followed by light being split into two spectral windows, "green" and "red", using longpass filter Q595, and then again using 50/50 beam splitters resulting in a total of eight detection channels. Additionally, bandpass filters are placed in front of the detectors (FF01-530/43-25; AHF, Tübingen, Germany for donor molecules and HQ 720/150 nm; AHF, Tübingen, Germany for acceptor molecules). Detection is performed using eight avalanche photodiodes (4 green channels: τ-SPAD (PicoQuant, Germany) and 4 red channels: SPCM-AQR-14 (Perkin Elmer). The detector outputs were recorded by a TCSPC module (HydraHarp 400, PicoQuant). As for setup #1, the correction factors for polarization mixing in the microscope objective were determined to be l_1= 0.0175 and l_2 = 0.0526.

For both setups data analysis was performed using home-written LabView-based software. Burst search was performed as described[15], using APBS (All Photon Burst Search) method and inter-photon times as threshold.

Lab#3

Sample solutions were measured with 100 µl drop on coverslip with concentration of around 50 pM. Single-molecule FRET experiments with ALEX were performed on a homebuilt confocal microscope as described previously[67]. The fluorescent donor molecules are excited by a diode laser OBIS 532-100-LS (Coherent, USA) at 532 nm operated at 60 µW at the sample in alternation mode. The fluorescent acceptor molecules are excited by a diode laser OBIS 640-100-LX (Coherent, USA) at 640 nm operated at 25 µW at the sample in alternation mode (100 µs alternation period). The lasers are combined by an aspheric fiber port (PAF2S-11A) and coupled into a polarization maintaining single-mode fiber P3-

488PM- FC-2 (Thorlabs, USA). The laser light is guided into the epi-illuminated confocal microscope (Olympus IX71, Hamburg, Germany) by dual-edge beamsplitter ZT532/640rpc (Chroma/AHF) focused by a water immersion objective (UPlanSApo 60x/1.2w, Olympus Hamburg, Germany). The emitted fluorescence is collected through the objective and spatially filtered using a pinhole with 50 µm diameter and spectrally split into donor and acceptor channel by a single-edge dichroic mirror H643 LPXR (AHF). Fluorescence emission was filtered (donor: BrightLine HC 582/75 (Semrock/AHF), acceptor: Longpass 647 LP Edge Basic (Semroch/AHF), focused on avalanche photodiodes (SPCM-AQRH-64, Excelitas). The detector outputs were recorded by a NI-Card PCI-6602 (National Instruments, USA).

Data analysis was performed using home written software package as described[67]. Single-molecule events were identified using a All-Photon-Burst-Search algorithm with a threshold of 15, a time window of 500 µs and a minimum total photon number of 150.

Lab#4

All sample solutions are measured in home-built 60 µL chambers with ~100-pM concentration. PIE-FRET experiments are carried out on a home build confocal microscope. The fluorescent donor molecules are excited by a pulsed diode laser (LDH-P- FA-530B, PicoQuant), at 532 nm operated at 20 MHz, with an excitation power of 55 µW at the sample in PIE experiment. The fluorescent acceptor molecules are excited by a pulsed diode laser (LDH-D-C-640, PicoQuant), at 639 nm operated at 20 MHz, with an excitation power of 50 µW at the sample in PIE experiment. The laserpulses are altered on the nanosecond timescale by a multichannel picosecond diode laser driver (PDL 828 "Sepia II", PicoQuant GmbH) with an oscillator module (SOM 828, PicoQuant GmbH). The lasers were coupled into a single mode fiber (P3-488PM-FC, Thorlabs GmbH) to obtain a Gaussian beam profile. Circular polarized light is obtained by a linear polarizer (LPVISE100-A, Thorlabs GmbH) and a quarter-wave plate (AQWP05M- 600, Thorlabs GmbH). The laser light is guided into the epi-illuminated confocal microscope (Olympus IX71, Hamburg, Germany) by dual-edge beam splitter (z532/633, AHF analysentechnik AG) focused by an oil immersion objective (UPLSAPO100XO, NA 1.40, Olympus Hamburg, Germany). The emitted fluorescence is collected through the objective and spatially filtered using a pinhole with 50 µm diameter and spectrally split into donor and acceptor channel by a single-edge dichroic mirror (640DCXR, AHF Analysentechnik AG, Germany). Fluorescence emission is filtered (donor: Brightline HC582/75 (AHF Analysentechnik AG), RazorEdge LP 532 (Laser 2000 GmbH), acceptor: (Shortpass 750, AHF Analysentechnik AG; RazorEdge LP 647, Laser 2000 GmbH), focused on avalanche photodiodes (SPCM-AQRH-14-TR, Excelitas Technoligies GmbH & Co. KG). The detector outputs were recorded by a TCSPC module (HydraHarp 400, PicoQuant). The setup is controlled by a commercial software package (SymPhoTime64, Picoquant GmbH).

Data analysis is performed using PAM software package as described[11] . Single-molecule events are identified using a two channel APBS-algorithm with a threshold of 5 photons per time window, a time window of 500 µs and a minimum photon number of 20. To remove photo-blinking and -bleaching events, the ALEX-2CDE filter was applied using an upper threshold of 10^{12}.

Lab#5

All sample solutions were measured in 200 µl PBS buffer with a labelled protein concentration of 25-100 pM. In short: Single-molecule FRET experiments with ALEX were performed on a homebuilt confocal microscope as described previously[67–69] .The fluorescent donor molecules are excited by a

spectrally filtered laser beam of a pulsed supercontinuum source (SuperK Extreme, NKT Photonics) with an acousto-optical tunable filter (AOTFnc-VIS, EQ Photonics), at 532 nm and 640 nm. The laser light is guided into a single-mode fiber (PM-S405-XP, Thorlabs) and the collimated beam (Focusing collimator MB06, Q-Optics/Linos) was coupled into an oil-immersion objective (60×, NA 1.35, UPLSAPO 60XO, Olympus) by using a dichroic beam splitter (zt532/642rpc, AHF Analysentechnik) mounted on an inverse microscope body (IX71, Olympus). The emitted fluorescence is collected through the objective and spatially filtered using a pinhole with 50 µm diameter and spectrally split into donor and acceptor channel by a single-edge dichroic mirror (640DCXR, AHF Analysentechnik). Fluorescence emission was filtered (donor: Brightline HC582/75, acceptor ET700/75; AHF Analysentechnik), focused on avalanche photodiodes (Tau-SPAD, PicoQuant). The detector outputs were recorded by a TCSPC module (HydraHarp 400, PicoQuant).

Data analysis was performed as described[67–69]. Single-molecule events were identified using an All Photon Burst Search algorithm with a threshold of 15, a time window of 500 µs and a minimum photon number of 200.

Lab#6

Single molecule measurements were carried out on a home-built confocal microscope[70] .Pulsed green and red laser light (532nm, LDH-P-FA-530 and 640nm, LDH-D-C-640, respectively, PicoQuant) was polarised, overlaid and focused on the sample by a 60x water immersion objective (CFI Plan Apo VC 60XC/1.2 WI, Nikon). Excitation light was separated from the emitted light by a dichroic mirror (F53-534 Dual Line beamsplitter z 532/633, AHF). The emitted light was then guided through a further dichroic mirror (F33-647 beam splitter 640 DCXR, AHF) to separate donor and acceptor fluorescence. After spectral separation, pinholes with a diameter of 150 micrometer refined the detection volume to about 8fL. Finally, the two photon streams were separated by polarizing beam splitters into their parallel and perpendicular parts and recorded by single-photon detectors (two SPCMAQR-14, PerkinElmer and two PDMseries APDs, Micro Photon Devices). Time-correlated single photon counting with picosecond resolution and data collection was performed by a HydraHarp400 (PicoQuant) and the Symphotime 32 software (PicoQuant). The G-factors for correcting the difference in sensitivity of the two polarization channels were 1.2.

The sample was measured within a BSA passivated well. Data analysis was performed using the PAM software package. Bursts were identified first using APBS with at minimum of 50 photons per burst, a 500 µs time window and at least 20 photons per time window to determine correction factors α and δ. Then a DCBS was performed with similar parameters and at least 12 photons in the green and 10 in the red channel. The other two correction factors were determined from the bursts identified in the second burst search.

Lab#7

All sample solutions were measured a drop on a coverslip, sealed in an airtight chamber, with concentration ~50 pM. The general scheme of the setup is described[71] . In short, Single-molecule FRET experiments with ALEX were performed on a homebuilt confocal setup, the smfBox. The fluorescent donor molecules are excited by a continuous wave laser at 515 nm operated at 100 µW at the sample. The fluorescent acceptor molecules are excited by a continuous wave laser, at 638 nm operated at 230 µW at the sample. The laser light is guided into the custom-built microscope body and focused by an oil immersion objective (UPLSAPO 60× NA = 1.35, Olympus Hamburg, Germany). The emitted fluorescence is collected through the objective and spatially filtered using a pinhole with 20 µm

diameter and spectrally split into donor and acceptor channels by a single-edge dichroic mirror (NC395323 - T640lpxr, Chroma, USA). Fluorescence emission was filtered (donor: FF01-571/72-25, acceptor: FF01-679/41-25, Semrock, USA), focused on avalanche photodiodes (SPCM-AQRH-14, Excelitas). The detector outputs were recorded by a national instruments card (PCIe-6353), with acquisition controlled using custom software available on our github (see smfBox reference above). A full description of the procedure is available[72].

Data analysis was performed using the PAM software package as described[11] .Single-molecule events were identified using a dual channel burst search algorithm with a threshold of 5 photons, a time window of 500 μs and a minimum photon number of 50.

Lab#8

All sample solutions were measured as a drop on a passivated with BSA (1mg/ml) coverslip with protein concentration from 20 to 50pM. SmFRET measurements were carried out on a custom built confocal setup using time-correlated single photon counting (TCSPC), which combines pulsed interleaved excitation with multi-parameter fluorescence detection (PIE-MFD) [4]. In brief, the fluorescent donor molecules are excited by a pulsed diode laser (LDH-P-FA 530L, PicoQuant) at 531nm, operated at 20MHz with 60μW power. The fluorescent acceptor molecules are excited by a pulsed diode laser (LDH-D-C 640, PicoQuant) at 640nm, operated at 20MHz with 40μW power. The laser light is coupled into a fiber collimator (60FC-4-RGBV11-47, Schäfter + Kirchhoff GmbH, Hamburg). After the collimator, the laser light is reflected by a dichroic mirror (HC quadband laser beamsplitter R405/488/532/635, AHF Analysentechnik AG, Tübingen) and then focused into the sample by a 1.2 NA water immersion objective (CFI Plan Apochromat VC 60x, Nikon GmbH, Düsseldorf). The laser driver, Sepia II (PicoQuant), operates the lasers such that they are pulsed (20 MHz) and shifted (by ~25ns) with respect to each other. The emitted fluorescence is collected by the objective and spatially filtered using a pinhole with 75μm diameter and spectrally split into two beams: parallel and perpendicular with respect to the excitation light by a polarizing beam splitter cube (PBS201, Thorlabs, Munich). After the PBS, the two beams are then split into donor and acceptor channel according to wavelength (HC BS 649, AHF Analysentechnik), resulting in 2 beams per polarization (green and red). Fluorescence emission was filtered (donor: 582/75 Brightline HC, acceptor: 700/75 ET bandpass, AHF Analysentechnik) and focused on a single-photon avalanche diode (τ-SPAD-100, PicoQuant) by a lens of 100 mm focal length. The detector outputs were recorded by a TCSPC module HydraHarp 400 (PicoQuant) which is synchronized with the laser driver. The photon arrival times were recorded with 16ps resolution for microtime and synchronization period of 50ns for macrotime.

Data analysis was performed using (PAM–PIE analysis with MATLAB v2.0 (develop branch up to commit 320364c4) software package as described [11]. Single-molecule events were identified using APBS as burst search method with a threshold of 100 photons per burst, a time window of 500 μs and a minimum photon number of 5. To remove photo-blinking and -bleaching events, the ALEX-2CDE filter was applied using an upper threshold of 10 [12].

Lab#9

The smFRET experiments were performed on a custom-built confocal detection-based microscope. The general scheme of the setup was previously described[73] All sample solutions were measured in 8-well chamber slides with a final volume of 200 μl at sample concentration of 50 pM. A pulsed laser diode (LDH-TA-560, Picoquant, Germany), pulsed at 40 MHz, 35 μW, is used for donor excitation; A pulsed laser diode (LDH-D-C-660, Picoquant, Germany), pulsed at 40 MHz, 100 μW, is used for acceptor excitation. We used pulsed interleaved excitation (PIE) scheme to alternatingly excite donor and

acceptor fluorophores to retrieve the stoichiometry (S) information[74] .The lasers were cleaned up with excitation filters (Brightline FF01-572/15 and Brightline FF01-650/13, Semrock), passed through the polarizer (GL-10A, Thorlabs) before entering the objective lens (60×, water immersion, NA=1.27, Nikon) through the central dichroic mirror (ZT405/488/561/660/905rpc-UF3, Chroma). The fluorescence emission signal was collected through the same objective lens and spatially filtered using a pinhole with 100 µm diameter and split into parallel and perpendicular polarization axis. Fluorescence emission was then filtered (donor: Brightline HC612/69, acceptor: ET700/75, Chroma) after separated into donor and acceptor channels via the dichroic mirror (FF650-DI01, Semrock). Photon signals were detected on photon counting detectors (τ-SPAD, Picoquant, Berlin) and recorded by a TCSPC module (Hydraharp400, Picoquant, Berlin). Acquired data were subject to multi-parameter fluorescence analysis and processed burst-wise for fluorescence intensities and lifetime with a threshold of 30 photons per burst[5,26]. All acquired data was performed by a custom written Igor-Program (Wavemetrics)[73] and the burst variance analysis (BVA) was analyzed by using algorithm from PAM-PIE with MATLAB[11] .

Lab#10

All sample solutions were measured as drops on a coverslip with a concentration of 50 pM. A custom-built confocal microscope was used for single-molecule FRET experiments as previously described[75–79] and the setup was modified to allow alternating-laser excitation of donor and acceptor fluorophores[80,81]. For this purpose, the fibre-coupled outputs of a 532-nm laser (operated at 250 µW) and a 640-nm laser (operated at 60 µW) were alternated with a modulation frequency of 20 kHz. Both beams were spatially filtered and coupled into an oil-immersion objective (60x 1.35 NA, UPLSAPO 60XO, Olympus). The same objective was used to collect the resulting fluorescence; the emission was separated from excitation light by a dichroic mirror, focused onto a 200-µm pinhole, and subsequently split spectrally on two avalanche photodiodes (SPCM-AQR-14, PerkinElmer, UK) detecting the donor and acceptor fluorescence with two distinct spectral filters (green: 585DF70; red: 650LP). Custom-written LabVIEW software was used to register and evaluate the detected signal. Fluorescence photons were assigned to either donor or acceptor-based excitation with respect to their photon arrival time. Two characteristic ratios, the fluorophore stoichiometry S and apparent FRET efficiency E*, were calculated for each fluorescent burst, yielding a two-dimensional histogram 6-7. One-dimensional E* distributions for donor-acceptor species were obtained using a $0.4 < S < 0.8$ threshold. These E* distributions were fitted with Gaussian functions, yielding the mean E* value for each distribution.

Lab#11

All sample solutions were measured in custom made glass chambers at a concentration of 15 pM. The general scheme of the setup is described in Krainer et al. 2018[82]. In short, single-molecule FRET experiments were performed on a custom-built confocal microscope as described previously by Hartmann et al. 2015[83]. The fluorescent donor and acceptor labelled molecules were excited in PIE mode with 25 MHz repetition rate by a green pulsed diode laser (LDH-P-FA 530L, PicoQuant, Berlin, Germany), at 530 nm wavelength, and a red pulsed diode laser (LDH-D-C 640, PicoQuant), at 640 nm wavelength with laser powers of 110 µW and 90 µW before objective, respectively. The laser light is guided into the inverted confocal microscope (Nikon Eclipse Ti - Nikon, Tokyo, Japan) by a dual-edge beam splitter zt532/640rpc (Chroma, Bellows Falls, VT, USA) and focused by a water immersion objective (FI Plan Apo WI 60x (NA 1.2), Nikon). The emitted fluorescence is collected through the same objective and spatially filtered using a pinhole with 50 µm diameter. In order to detect the fluorescence anisotropy, the spatially filtered light is separated according to its polarization by a

polarizing beam splitter (PBS201, Thorlabs, Newton, NJ, USA) and guided in the parallel and perpendicular detection channel. In each detection channel the light is spectrally split into donor and acceptor emission by a single-edge dichroic mirror (FF650-Di01, Semrock, New York, NY, USA). The polarization separated and spectrally split fluorescence emission was bandpass filtered (donor: FF01-582/75, Semrock, acceptor: ET700/75M, Chroma) and focused on avalanche photodiodes (SPCM-AQRH-14, Excelitas, Waltham, MS, USA). The detector outputs were recorded by a TCSPC module (HydraHarp 400, PicoQuant).

Data analysis was performed using a custom written software package as described in Hartmann et al. 2017[84]. Single-molecule events were identified from the acquired photon stream as fluorescence bursts with a maximum inter-photon time of 50 µs containing a minimum total number of 40 photons for protein 1 and 100 photons for protein 2 after background correction and a Lee filter with window size four. To remove photo-blinking and bleaching events, the ALEX-2CDE filter as described by Tomov et al. 2012[12] was applied using an upper threshold of 8.

Lab#12

All sample solutions were measured in Cellview chamber slides (Greiner BioOne, Frickenhausen, Germany) with a dilution of 1 to 600 from the received stocks. Single-molecule FRET experiments with PIE were performed on a commercial MicroTime 200 confocal microscope (PicoQuant, Berlin, Germany) as described [85]. The fluorescent donor molecules are excited by a pulsed diode laser (LDH-P-FA 530B, PicoQuant), at 532 nm operated at 20 MHz, 60 µW at the sample in the PIE excitation mode . The fluorescent acceptor molecules are excited by a pulsed diode laser (LDH-D-C-640, PicoQuant, Berlin, Germany; clean up filter: zet636/20x, Chroma, Bellow Falls, VT, USA), at 640 nm operated at 20 MHz, 30 µW at the sample in the PIE excitation mode. The laser light is guided into the inverted IX73 confocal microscope (Olympus, Hamburg, Germany) by a dual-edge beam splitter ZT532/640rpc-UF3 (Chroma, Bellow Falls, VT, USA) and then focused by a water immersion objective (UPlanSApo 60x/1.2w, Olympus, Hamburg, Germany). The emitted fluorescence is collected through the objective and spatially filtered using a pinhole with 50 µm diameter and spectrally split into donor and acceptor channel by a single-edge dichroic mirror (T635lpxr; Chroma, Bellow Falls, VT, USA). Fluorescence emission was filtered (donor: ff01-582/64; Semrock, Rochester, NY, USA; acceptor: H690/70; Chroma, Bellow Falls, VT, USA) and focused on avalanche photodiodes (SPCM-AQRH-14-TR, Excelitas Technologies, Waltham, MS, USA). The detector outputs were recorded by a TCSPC module (HydraHarp 400, PicoQuant, Berlin, Germany). The setup was controlled with the SymPhoTime64 software package (PicoQuant, Berlin, Germany).

Data analysis was performed using the PAM software package as described [11]. Single-molecule events were identified using an APBS and DCBS with a threshold of 100 photons per burst, a sliding time window of 500 µs and a minimum photon number of 10.

Lab#13

All sample solutions were measured as a drop on a coverslip (Roth, Karlsruhe, heated to 500 °C for 2 h), covered with a humidity chamber with a dilution of 1 to 300 (sample 1) and 1 to 200 (sample 2 and 3) of the delivered stock solution. The general scheme of the setup is described[86]. In short: Single-molecule FRET experiments with POE (Pulsed Overlayed Excitation) were performed on a homebuilt confocal microscope as described[86] .The fluorescent donor molecules are excited by a cw DPSS "Crysta Laser" (GCL-005-L, Laser2000, Wessling) at 532 nm with 40 µW at the sample. The fluorescent acceptor molecules are excited by a pulsed diode laser (LDH-P-C 635, PicoQuant GmbH), at 635 nm operated at 10 MHz, 5 µW at the sample in the POE experiment. The laser light is guided through a

single-mode fibre (SMC-460, Schäfter&Kirchoff), directed by dual-band beamsplitter (Z532/633, Chroma, Bellows Falls, USA) and then focused by a water immersion objective (CFI PlanApochromat 60x WI, Nikon, Japan). The emitted fluorescence is collected through the objective and spatially filtered using a pinhole with 50 µm diameter and spectrally split into donor and acceptor channel by a single-edge dichroic mirror (BS640DCXR, Chroma, Bellows Falls, USA). Fluorescence emission was filtered (donor: FF01_582_75, Semrock, USA) acceptor: HQ685_70, Chroma, Bellows Falls, USA) and focused on avalanche photodiodes (SPCM-AQR 14, Perkin Elmer, Fremont, USA). The detector outputs were recorded by a TCSPC module (TimeHarp200, PicoQuant GmbH, Berlin, Germany).

Data analysis was performed using IgorPro 8 (Wavemetrics, Portland OR, USA). Single-molecule events were identified using a bin-selection algorithm with a threshold of 40 photons in the sum of donor and acceptor channel upon donor excitation and a threshold of 10 photons in the acceptor channel upon acceptor excitation.

Lab#14

Our multi-parameter fluorescence detection setup equipped with pulsed interleaved excitation is conceptually identical to the confocal microscope described previously[4]. For excitation, a pulsed supercontinuum laser was used with wavelength selector at 532±5 nm (Solea, Picoquant, Berlin, Germany), and a spectrally filtered (Chroma z635/10x, Picoquant) 635-nm laser diode (LDH-P-C-635B, Picoquant). Both lasers were alternated at 26.67 MHz (PDL 828 Sepia2, Picoquant), delayed ~18-ns with respect to each other and combined via a dichroic mirror (Chroma T560lpxr, F48-559, AHF) in a single-mode optical fiber (coupler: 60FC-4-RGBV11-47, fiber: PMC-400Si-2.6-NA012-3-APC-150-P, Schäfter und Kirchhoff GmbH, Hamburg, Germany). After collimation (60FC-L-4-RGBV11-47, SuK GmbH), the linear polarization was cleaned up (Codixx VIS-600-BC-W01, F22-601, AHF) and the light was reflected on a 3-mm thick excitation polychroic mirror (Laser Beamsplitter zt532/640rpc, F58-PQ09, AHF) upward and into the back port of the microscope (IX70, Olympus Belgium NV, Berchem, Belgium) via two mirrors and upward to the sample (3-mm thick Full Reflective Ag Mirror, F21-005, AHF, mounted in a TIRF Filter Cube for BX2/IX2, F91-960, AHF) to the objective (UPLSAPO-60XW, Olympus). Sample emission was focused through a 75-µm pinhole (P75S, Thorlabs, Munich, Germany) via an achromatic lens (AC254-200-A-ML, Thorlabs), collimated again (AC254-50-A-ML, Thorlabs) and spectrally split (Laser Beamsplitter H 643 LPXR superflat, F48-643, AHF). The yellow range was filtered (582/75 BrightLine HC, F37-582, AHF) and polarization was split (PBS251, Thorlabs). The red range was also filtered (Chroma ET705/100m, AHF) and polarization was split (PBS252, Thorlabs). Photons were detected on four avalanche photodiodes (yellow photons: Laser Components COUNT Blue, red photons: Perkin-Elmer or EG&G SPCM-AQR12/14) all of which were connected to a time-correlated single photon counting (TCSPC) device (SPC-630, Becker & Hickl GmbH, Berlin, Germany) over a router (HRT-82, Becker & Hickl) and power supply (DSN 102, Picoquant). Signals were stored in 12-bit first-in- first-out (FIFO) files. All analyses of experimental data were performed in the software package PAM[11].

Lab#15

All sample solutions were measured in Corning 384 well non-binding plates at concentrations ranging between 20-60 pM. The general scheme of the setup is described[87].In short: Single-molecule FRET experiments with PIE - MFD were performed on a homebuilt confocal microscope as described previously[87].Excitation was performed with a pulsed SC450-4-20Mhz supercontinuum source (Fianium, Southampton, UK). It runs at 20MHz, and has a power density >2mW/nm over the 450-800nm range, with an average pulse width of 100-150ps. The collimated, unpolarized output of the source is divided by a 50:50 beamsplitter cube (BS016, Thorlabs, NJ, USA), thus generating two beams.

The beams are spectrally filtered using excitation bandpass filters at wavelength 532/10 (prompt beam) and 635/10 (delayed beam) to excite donor and acceptor molecules, respectively. The delayed beam has a path length of ~ 8m relative to the prompt beam, generating a ~24 ns delay in the pulse. The two beam paths are then recombined using a 50:50 beamsplitter cube (BS016, Thorlabs, NJ, USA) and focused using a 10x objective into a single-mode fiber (SMF) (P1-460A-FC, Thorlabs, NJ, USA), by which the beams become spatially overlapped and filtered. The output of the fiber is collimated using a 10x microscope objective lens (04OAS010; CVI Melles Griot, Albuquerque, NM, USA), polarized by a polarizing beamsplitter cube (PBS; PBS201, Thorlabs, NJ, USA) and coupled into an inverted microscope (Eclipse Ti, Nikon, France). Power was controlled using ½ (prompt: WPH05M-532 and delayed WPH05M-633, Thorlabs, NJ, USA) and ¼ waveplates (prompt: WPQ05M-532 and delayed WPQ05M-633, Thorlabs, NJ, USA) placed in the prompt and delayed beams before recombination, resulting in 50 µW for the prompt and 25 µW for the delayed beam at the entrance into the microscope. The light is reflected by a dichroic mirror that matches the excitation/emission wavelengths (FF545/650-Di01, Semrock, Rochester, NY, USA) and coupled into a Nikon 100x, NA1.4 objective. Emitted photons are then collected by the objective and focused by the tube lens on a pinhole of 150µm. The emission photon stream is collimated and divided using a polarizing beamsplitter cube (PBS; PBS201, Thorlabs, NJ, USA). In each created polarization channel, the photons are spectrally separated using dichroic mirrors (BS 649, Semrock, Rochester, NY, USA) and filtered using high quality emission filters (parallel: ET BP 585/65 and ET BP 700/75 and perpendicular: HQ 590/75 M and HQ 660 LP, Chroma, Bellows Falls, VT, USA). Single photons are detected using Single Photon Avalanche Diodes. We use two MPD-1CTC (MPD, Bolzano, Italy) for the donor wavelength channels and two SPCM AQR-14 (Perkin Elmer, Fremont, CA, USA) for the acceptor wavelength channels. The output of the detectors is coupled into a TCSPC counting board (SPC-150, Becker&Hickl, Berlin, Germany), through a HRT41 router (B&H), using appropriate pulse inverters and attenuators. The sync signal that triggers the TCSPC board is provided by picking a small fraction of the light from the prompt path (reflected by a coverslip), and focusing it on an avalanche diode (APM-400, B&H).

Data analysis was performed using the Software Package for Multiparameter Fluorescence Spectroscopy, Full Correlation and Multiparameter Fluorescence Imaging developed in C.A.M. Seidel's lab (http:// www.mpc.uni-duesseldorf.de/seidel/). A single-molecule event was defined as a burst containing of at least 40 photons with a maximum allowed interphoton time of 0.3 ms and a Lee-filter of 20. Photobleaching events were identified base on |TGX-TRR|<1 ms as described [4].

Lab#16

All sample solutions were measured in chamber with concentration 100 pM. The general scheme of the setup is similar to Supplementary Figure 2. In short: Single-molecule FRET experiments with ALEX were performed on a homebuilt confocal microscope as described previously. The light from 532nm (Compass 215M-50, Coherent, USA) and 638nm (25mW Red flame, Coherent, USA) cw-laser sources is alternately directed to an IX71 inverted microscope (Olympus, Japan) every 25 µs, reflected on a beamsplitter (Z488-533-633RPC, Chroma, USA) and focused through a water-immersion objective (UPlanApo 60x/1.2w, Olympus, Japan) to excite fluorescent molecules. The light intensities were 100 µW for 532nm and 35 µW for 638nm, measured before the beamsplitter. The emitted fluorescence is collected through the objective, spatially filtered using a 100 µm pinhole, and then spectrally split into two photon streams by a dichroic mirror (635 DCXR, Chroma, USA). Individual photon streams were filtered (for donor: HQ580/60m, for acceptor: HQ665lp, Chroma, USA) and detected by single photon-avalanche photodiodes (SPCM-AQR-14, PerkinElmer, USA). The detector outputs (photon arrival times) were recorded by a counter/timer device module (PCI-6602, National Instruments, USA).

Data analysis was performed using the ALEX-suite software package as described[88].Single-molecule events were identified using an all-photon-burst-search (APBS) and a dual-channel-burst-search

(DCBS) algorithm with a threshold of 150, a time window of 50 μs and a minimum photon number of 50.

Lab#17

All sample solutions were measured in microscopy coverslip wells (μ-Slide 18 Well, Ibidi, GmbH) as was previously described in the analysis of smFRET within-burst dynamics of a DNA hairpin[89]. Sample excitation in all nsALEX/PIE measurements was achieved using pulsed picosecond fiber laser (λ = 532 nm, pulse width of 100 ps FWHM, operating at 20 MHz repetition rate and 100 μW measured at the back aperture of the objective lens; FL-532-PICO, CNI, China), and pulsed picosecond diode laser (λ = 642 nm, pulse width of 100 ps FWHM, operating at 20 MHz repetition rate and 60 μW measured at the back aperture of the objective lens; QuixX® 642-140 PS, Omicron, GmbH), delayed by 25 ns, for donor and acceptor excitation, respectively. Excitation path: (i) polarization maintaining optical fiber, (ii) beam shaping through quarter waveplate and a linear polarizer, (iii) dichroic beamsplitter with high reflectivity at 532 and 640 nm (ZT532/640rpc, Chroma, USA), (iv) high numerical aperture (NA) super apochromatic objective lens (60X, NA = 1.2, water immersion, Olympus, Japan). Emission path: (i) fluorescence was collected through the same objective lens, (ii) focused with an achromatic lens (f = 100 mm) onto a 100 μm diameter pinhole (variable pinhole, motorized, tunable from 20 μm to 1 mm), (iii) re-collimated with an achromatic lens (f = 100 mm), (iv) donor and acceptor fluorescence were split into two separate detection channels using a dichroic mirror with cutoff wavelength at λ = 652 nm (FF652-Di01-25x36, Semrock Rochester NY, USA), (v) donor and acceptor fluorescence were further filtered using 585/40 nm (FF01-585/40-25, Semrock Rochester NY, USA) and 698/70 nm (FF01-698/70-25, Semrock Rochester NY, USA) bandpass filters, respectively, and (vi) donor and acceptor fluorescence photons were detected using two hybrid photomultipliers (Model R10467U-40, Hamamatsu, Japan). Signal acquisition path: detectors were routed through a 4-to-1 router to a time-correlated single photon counting (TCSPC) module (SPC-150, Becker & Hickl, GmbH) as its START signal (the STOP signal is routed from the laser controller). Data acquisition was performed using the VistaVision software (version 4.2.095, 64-bit, ISS™, USA) in the time-tagged time-resolved (TTTR) file format, and then transformed into the photon HDF5 file format[90] for input in the FRET Bursts[14] analysis software. Data analysis: (i) the data was split into different photon streams, based on the detector they were originating from (donor or acceptor detectors) and on their photon nanotimes, whether it belonged to the donor or acceptor excitation time windows in nsALEX/PIE. The photon streams further used were $D_{ex.}D_{em.}$, $D_{ex.}A_{em.}$ and $A_{ex.}A_{em.}$ (ii) The background for each photon stream was assessed per each 30 s of acquisition. (iii) Photon bursts were identified as time periods were the instantaneous photon count rate of a sliding window of m=20 consecutive photons were at least F=6 times higher than the background rate. Both ACBS and DCBS were tested[91], where ACBS was used for attaining all data including donor- and acceptor-only bursts, and DCBS was used for attaining FRET-active bursts. (iv) After ACBS, bursts were further selected if they included at least 30 photons in all photon streams. After DCBS, bursts were further selected if they included at least 30 photons in the donor excitation photon streams, $D_{ex.}D_{em.}$, $D_{ex.}A_{em.}$, and at least 30 photons in the acceptor excitation stream, $A_{ex.}A_{em.}$. (v) Using the selected bursts after ACBS, donor-only and acceptor-only sub-populations were further selected, and then used for calculating the α and δ. (vi) Using the selected bursts after DCBS, the FRET-active bursts were further used for estimating the β and γ correction factors. (vii) burst-wise parameters, E and S values, as well as mean donor/acceptor nanotime for E-τ maps and E standard deviation for BVA were calculated per each burst and accumulated in 1D and 2D histograms. Sum of gaussian fitting was used for extracting the mean values of sub-populations of burst-wise parameters. All burst-wise parameters were corrected for background and the α, β, γ and δ correction factors.

Lab#18

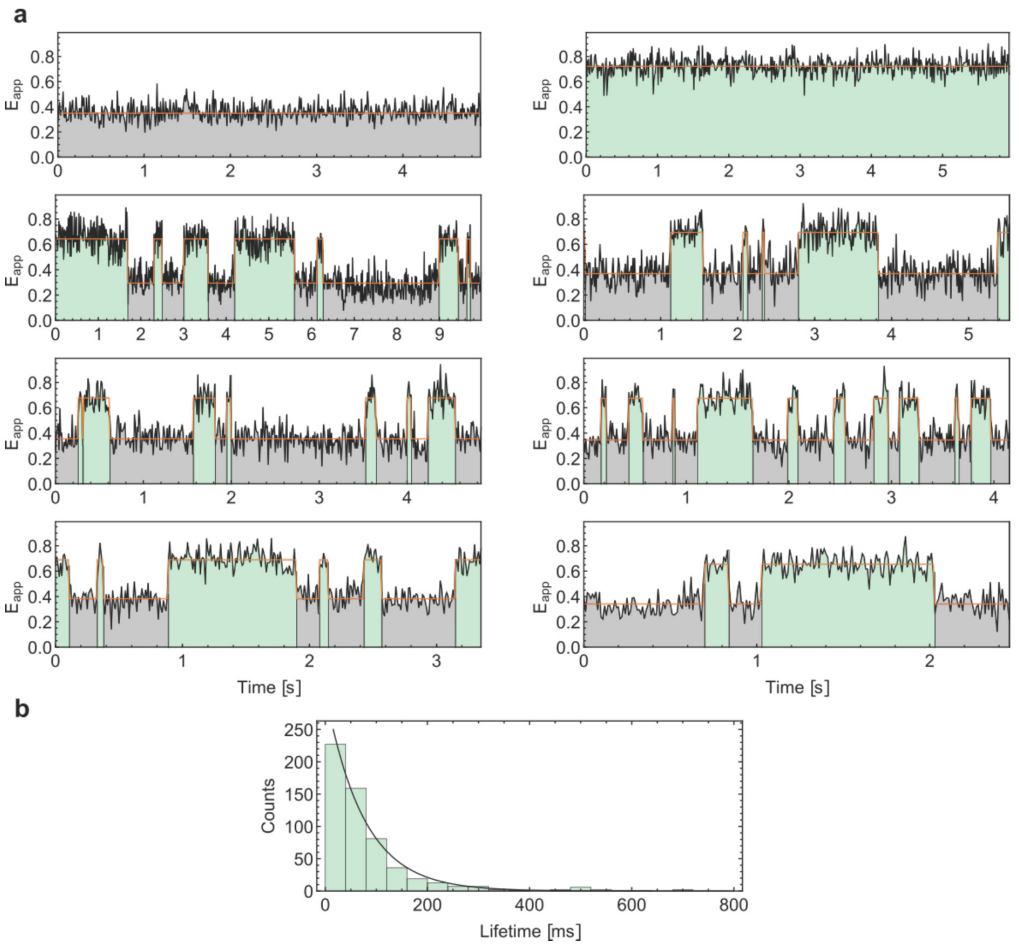
All sample solutions were measured as a 50ul drop on a coverslip covered with a humidity chamber with a dilution of 1 to 1000 (sample 1) and 1 to 400 (sample 2 and 3) of the delivered stock solution. Measurements were performed on a home build confocal microscope, as previously described[92] . Using an ICHROME MLE-SFG laser module (Toptica, Germany) as excitation source, alternating 25 µs excitation pulses of 514 nm (80 µW) and 632 nm (80 µW). The excitation beam was directed via a fiber coupler and a dichroic mirror (z514/640rpc, Chroma, USA) through the water-immersion objective (60x, NA 1.2, Olympus, Japan), and focused 50 µm above the glass-sample interface. Fluorescence was spatially filtered with a 50 µm pinhole in the image plane and split by a second dichroic mirror (640dcxr, Chroma, USA). The fluorescence signals were further filtered (hq570/100nm and hq700/75nm, Chroma, USA) for green and red detection and focused on the active area of two single photon avalanche photodiodes (SPADs, SPCM AQR-14, Perkin Elmer, USA). The photodiodes were read out with a TimeHarp 200 photon counting board (Picoquant, Germany), and the arrival times were stored in t3r (time-tagged to timeresolved) files. Each dataset contained ~3000 bursts and was analyzed using the FRETBursts toolkit[14] .A single-molecule event was defined as a burst containing at least 40 photons with a maximum interphoton time of 100 µs.

Lab#19

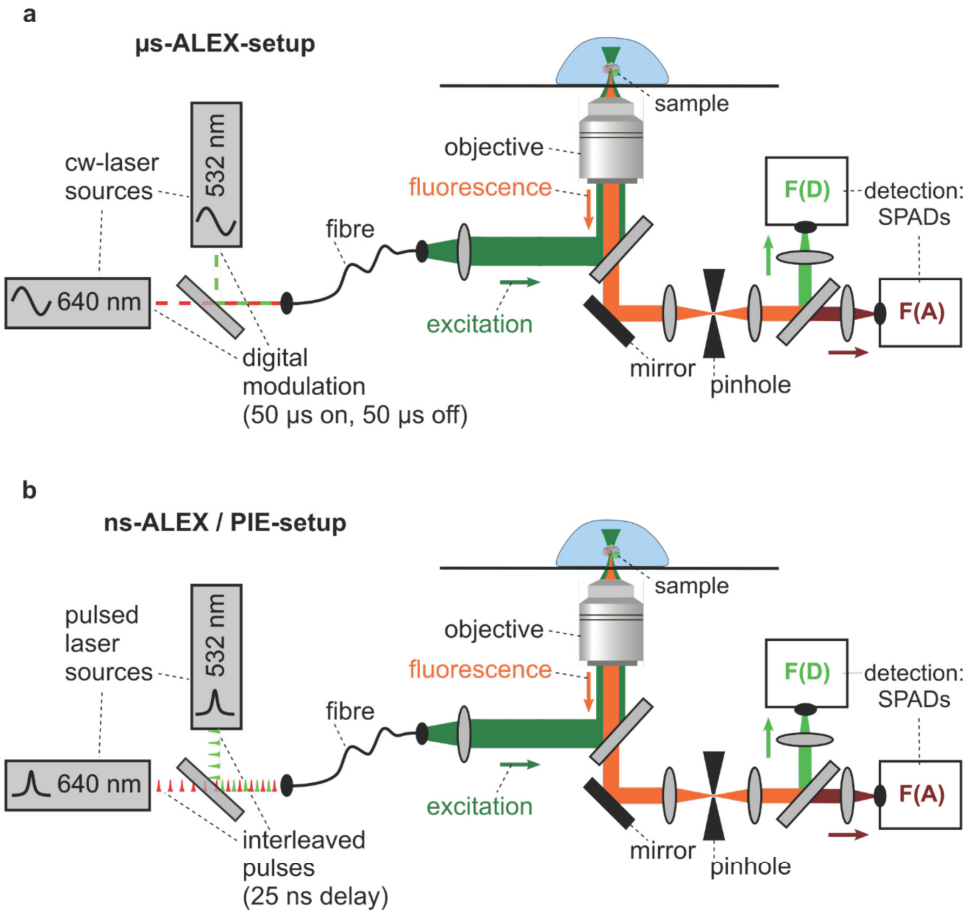
Single-molecule FRET experiments were performed on a homebuilt confocal microscope as described previously[93]. Briefly, a confocal excitation laser (532 nm, World StarTech) was focused on the sample through the side port of the microscopy (Olympus) with a 100X oil immersion objective (Olympus). The excitation laser intensity was maintained at 50 µW for all experimental conditions. The fluorescence emission was collected through the objective, spatially filtered through a 100 µM pinhole, spectrally filtered through a band pass filter (HQ580/60 m, Chroma), split into donor and acceptor channels by a dichroic mirror (HQ680/60 m, Chroma) and subsequently detected by two avalanche photodiodes (Excelitas Technologies).

All sample solutions were measured in BSA (Bovine serum albumin, 1 mg/mL) passivated chambers with sample concentrations around 50 pM. Baseline fluorescence signal was estimated from imaging chamber containing imaging buffer only. Single molecule events were identified after subtraction of baseline signal, as bursts with minimum 40 counts over an exposure time of 1 ms.

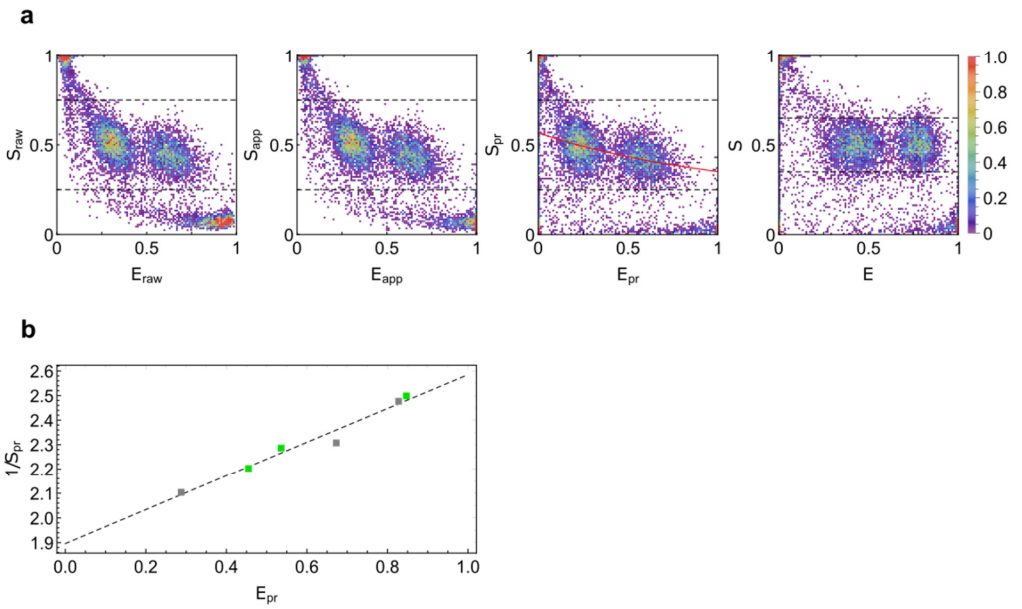
Supplementary Figures



Supplementary Figure 1: Ligand-induced slow conformational dynamics of MalE switching between the apo- and holo states. (a) Time traces of immobilized MalE 36C/352C molecules labeled with Alexa555-Alexa647 measured in buffer containing no ligand (top row, left, grey), 1 mM maltose (top row, right, green) and 1 µM maltose , close to the K_d (rows 2-4)[94]. The samples were measured in a scanning confocal microscope as described in reference [68]. FRET states and lifetimes are extracted from a fitted two-state Hidden-Markov-Model as described in reference [95]. The traces show ligand-induced interconversion of states on the >10 millisecond timescale. **(b)** A dwell time histogram of the duration of the holo state at 1 µM maltose with an exponential fit (solid line) shows a mean dwell time of 75 ms.



Supplementary Figure 2: Schematics of the experimental setups. (a) A schematic of a confocal microscope setup used for the acquisition of diffusion-based smFRET data using alternating-laser excitation (ALEX). Continuous wave laser sources (here with wavelengths of 532 nm and 640 nm) excite the sample alternatively for periods of 50 µs through a microscope objective. F(D) and F(A) indicate the donor and acceptor detection channels, respectively. **(b)** A schematic of a confocal microscope setup used for the acquisition of diffusion-based smFRET data using nsALEX / pulsed interleaved excitation (PIE). Pulsed laser sources (here with wavelengths of 532 nm and 640 nm) alternatingly excite the sample with picosecond pulses delayed by ~25 ns.

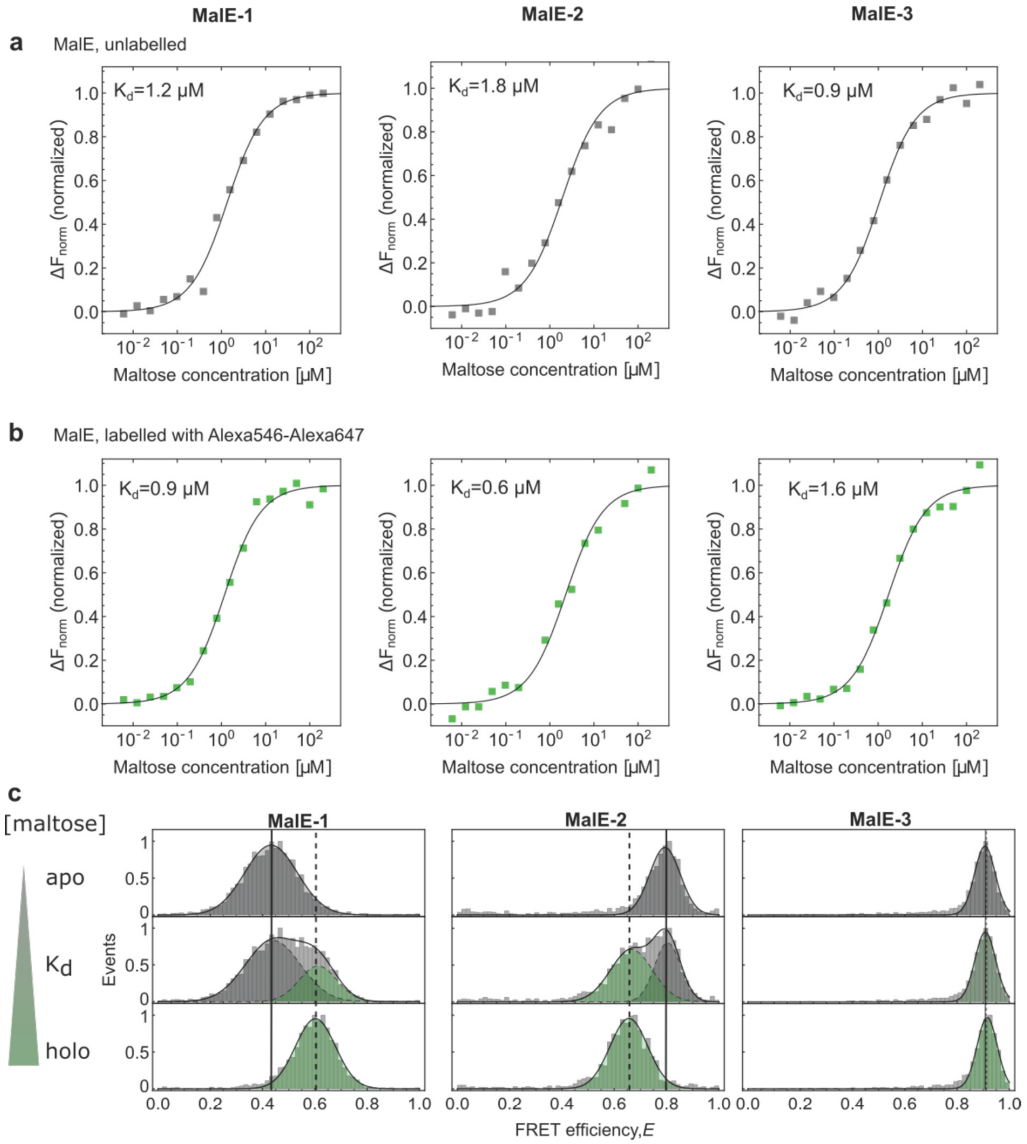


Supplementary Figure 3: Correction procedure when using a global γ. (a) *ES*-diagrams of two selected and merged data sets (Alexa546-Alexa647 labeled MalE-1 apo and MalE-2 apo) showing the results from the different correction steps; from left to right: raw data, background corrected apparent FRET efficiency, crosstalk and direct excitation corrected proximity ratio, E_{pr}, with fitted γ curve (red), and the γ corrected FRET efficiency versus stoichiometry plots. **(b)** Proximity ratio, E_{pr} and stoichiometry S_{pr} of all MalE mutants in the apo (gray dots) and holo states (green dots) with the linear fit (dashed line) used for a global γ correction. For more details for global γ estimation, please see Supplementary Note 2.

Supplementary Figure 4: Primer design and sequences for creation of MalE variants

The sequence of MalE (blue, DNA sequence in black) and the primers (in red) used for generating the mutants are given below. The procedures have been published previously[68].

```
   1                                                                GTCGGTAAG
   1  K  I  E  E  G  K  L  V  I  W  I  N  G  D  K  G  Y  N  G  L  A  E  V  G  K
   1 AAAATCGAAGAAGGTAAACTGGTAATCTGGATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAG

  76 AAATTCGAGWRMGATACCGG 3' Lys29Cys
  26  K  F  E  K  D  T  G  I  K  V  T  V  E  H  P  D  K  L  E  E  K  F  P  Q  V
  76 AAATTCGAGAAAGATACCGGAATTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCCACAGGTT

  51  A  A  T  G  D  G  P  D  I  I  F  W  A  H  D  R  F  G  G  Y  A  Q  S  G  L
 151 GCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTACGCTCAATCTGGCCTG

 226               CCGGACAAAGCGTTCCAGKRCAAGCTGTATCCG 3' Asp87Cys
  76  L  A  E  I  T  P  D  K  A  F  Q  D  K  L  Y  P  F  T  W  D  A  V  R  Y  N
 226 TTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAGCTGTATCCGTTTACCTGGGATGCCGTACGTTACAAC

 101  G  K  L  I  A  Y  P  I  A  V  E  A  L  S  L  I  Y  N  K  D  L  L  P  N  P
 301 GGCAAGCTGATTGCTTACCCGATCGCTGTTGAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCG

 376            GAAGAGATCCCGKSSCTGGATAAAGAAC 3' Ala134Cys
 126  P  K  T  W  E  E  I  P  A  L  D  K  E  L  K  A  K  G  K  S  A  L  M  F  N
 376 CCAAAAACCTGGGAAGAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC

 151  L  Q  E  P  Y  F  T  W  P  L  I  A  A  D  G  G  Y  A  F  K  Y  E  N  G  K
 451 CTGCAAGAACCGTACTTCACCTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAAACGGCAAG

 526                    GTGGATAACKSYGGCGCGAAAGCG 3' Ala186Cys
 176  Y  D  I  K  D  V  G  V  D  N  A  G  A  K  A  G  L  T  F  L  V  D  L  I  K
 526 TACGACATTAAAGACGTGGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCTGGTTGACCTGATTAAA

 201  N  K  H  M  N  A  D  T  D  Y  S  I  A  E  A  A  F  N  K  G  E  T  A  M  T
 601 AACAAACACATGAATGCAGACACCGATTACTCCATCGCAGAAGCTGCCTTTAATAAAGGCGAAACAGCGATGACC

 226  I  N  G  P  W  A  W  S  N  I  D  T  S  K  V  N  Y  G  V  T  V  L  P  T  F
 676 ATCAACGGCCCGTGGGCATGGTCCAACATCGACACCAGCAAAGTGAATTATGGTGTAACGGTACTGCCGACCTTC

 251  K  G  Q  P  S  K  P  F  V  G  V  L  S  A  G  I  N  A  A  S  P  N  K  E  L
 751 AAGGGTCAACCATCCAAACCGTTCGTTGGCGTGCTGAGCGCAGGTATTAACGCCGCCAGTCCGAACAAAGAGCTG

 276  A  K  E  F  L  E  N  Y  L  L  T  D  E  G  L  E  A  V  N  K  D  K  P  L  G
 826 GCGAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATAAAGACAAACCGCTGGGT

 301  A  V  A  L  K  S  Y  E  E  E  L  A  K  D  P  R  I  A  A  T  M  E  N  A  Q
 901 GCCGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATCCACGTATTGCCGCCACCATGGAAAACGCCCAG

 976                                                              GATCAACGCC
 326  K  G  E  I  M  P  N  I  P  Q  M  S  A  F  W  Y  A  V  R  T  A  V  I  N  A
 976 AAAGGTGAAATCATGCCGAACATCCCGCAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCC

1051 GCCWGCGGTCGTCAG 3' Ser352Cys
 351  A  S  G  R  Q  T  V  D  E  A  L  K  D  A  Q  T
1051 GCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACT
```

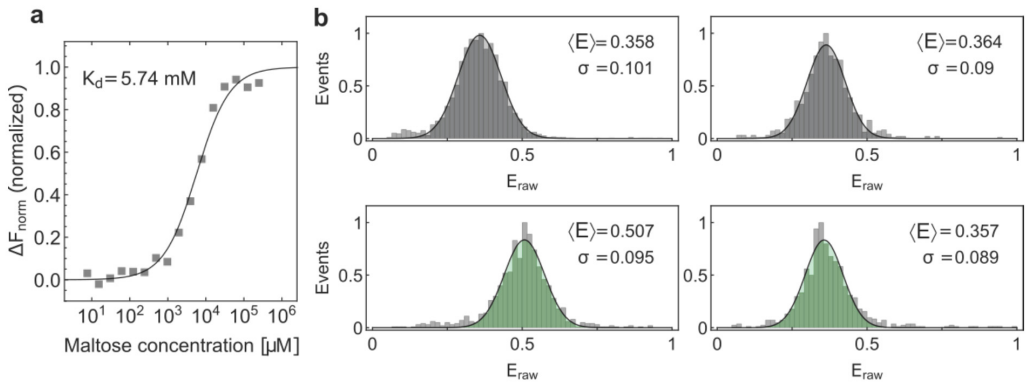


Supplementary Figure 5: Binding affinity measurements of maltose to MalE using microscale thermophoresis. (a) The binding affinities of maltose to MalE were measured using microscale thermophoresis (Monolith NT.LabelFree, Nanotemper) where the ratio of fluorescence before and after heating $\Delta F_{norm} = F_{cold}/F_{hot}$ was recorded at different maltose concentrations[96]. Data points show ΔF_{norm} normalized to the minimal and maximal fluorescence intensities for the unlabeled mutants MalE-1 (left), MalE-2 (middle), and MalE-3 (right). The curves were fitted with a standard model for receptor-ligand kinetics

$$\Delta F_{norm} = \frac{K_d + c_P + c_{malt} - \sqrt{(K_d + c_P + c_{malt})^2 - 4 c_P c_{malt}}}{2 c_P},$$

where K_d is the dissociation constant, c_P the protein concentration (0.25 µM in the experiment), and c_{malt} the maltose concentration. The fits to the binding model (solid line) yield K_d-values of 1.2 µM (left), 1.8 µM (middle), and 0.9 µM (right), respectively. This variation in the K_d values is within the typical reproducibility of the measurement determined to be K_d =1.2 ± 0.5 µM from measurements on19 different mutants. **(b)** The binding affinities of maltose to fluorescently-labeled MalE (Alexa547 and Alexa647) mutants MalE-1 (left), MalE-2 (middle), and MalE-3 (right) were measured using microscale thermophoresis. These experiments yielded K_d-values of 0.9 µM (left), 0.6 µM (middle), and 1.6 µM (right), respectively. **(c)** FRET efficiency E histogram for the MalE mutant MalE-1 (left), the mutant MalE-2 (middle), and the mutant MalE-3 (right) in the presence of

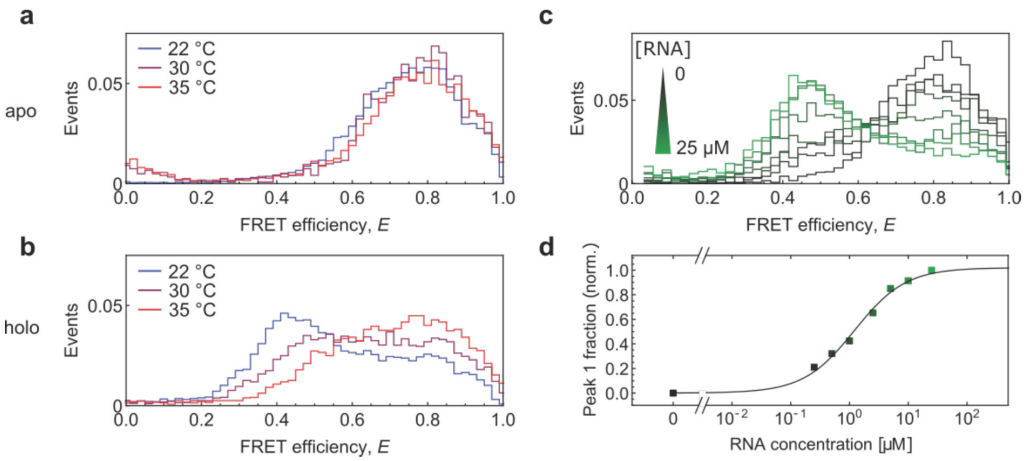
0 (top), 1 μM (middle), and 1 mM maltose (bottom). The peak position of the apo (solid line) and holo (dashed line) populations, determined for this experiment, are shown. The data is from Lab #3.



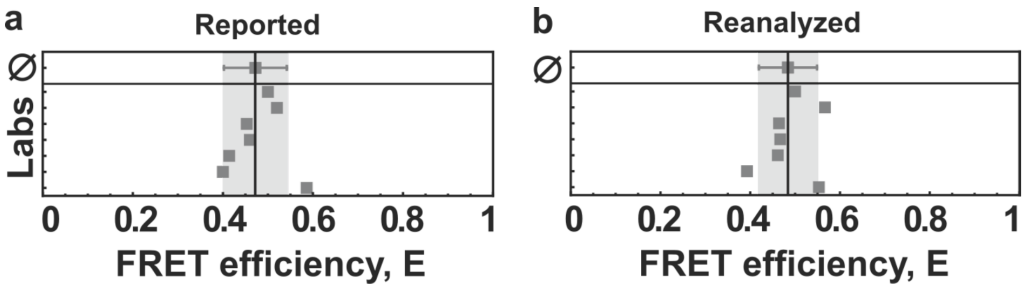
Supplementary Figure 6: Millimolar maltose concentration does not influence fluorophore properties. To investigate the potential influence of maltose on the photophysical properties of the used fluorophores, measurements were performed using the D65A closing deficient mutant of MalE-1. **(a)** The binding affinity of maltose to the MalE-1 mutant D65A were measured using microscale thermophoresis. The ratio of fluorescence before and after heating $\Delta F_{norm} = {F_{cold}}/{F_{hot}}$ was recorded at different maltose concentrations[96]. Data points show ΔF_{norm} normalized to minimal and maximal fluorescence intensities for the unlabeled mutants. The curves were fitted with a standard model for receptor-ligand kinetics:

$$\Delta F_{norm} = \frac{K_d + c_P + c_{malt} - \sqrt{(K_d + c_P + c_{malt})^2 - 4 c_P\, c_{malt}}}{2\, c_P},$$

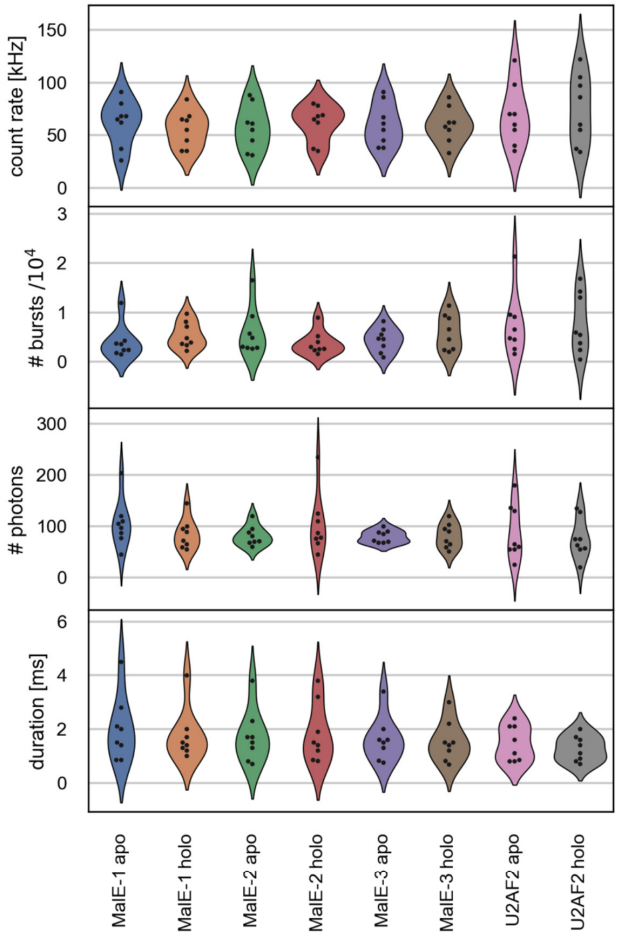
where K_d is the dissociation constant, c_P the protein concentration (0.25 µM in the experiment) and c_{malt} the maltose concentration. A K_d of 5.7 mM was measured. **(b)** Raw FRET efficiency E_{raw} histograms for the mutant MalE-1 labeled with Alexa546-Alexa647 without maltose (top, left) and with added 1 mM maltose (bottom, left) compared to same measurement with the MalE-1 closing deficient mutant D65A (right). No influence of the maltose is visible in the FRET efficiency histogram for the closing deficient mutant.



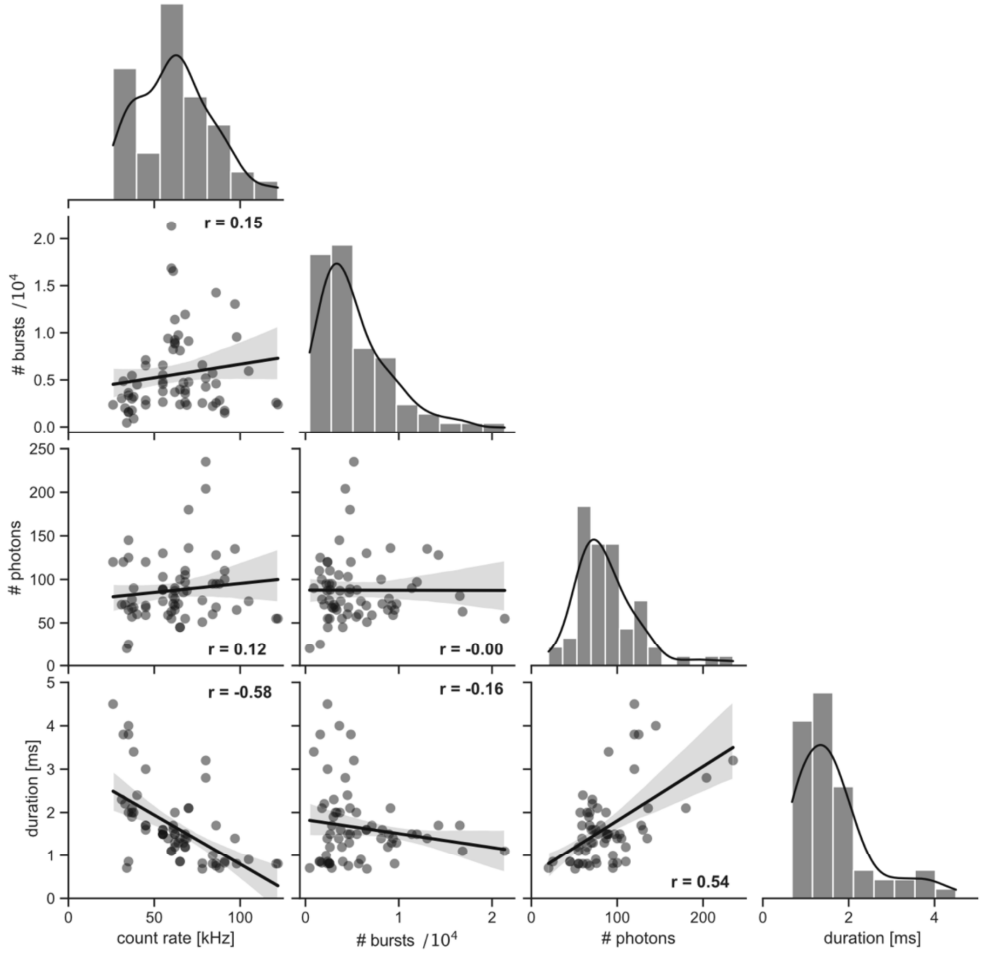
Supplementary Figure 7: Temperature and concentration dependence of RNA binding to U2AF2. **(a, b)** SmFRET histograms of **(a)** apo and **(b)** holo measurements at 22°C (blue), 30°C (purple) and 35°C (red). **(c)** SmFRET histograms for U9 RNA titration measurements with U2AF2 (low to high RNA concentrations are shown in a color gradient from black to light green (0, 0.25, 0.5, 1.0, 2.5, 5, 10 and 25 µM). **(d)** The area under peak 1 of the FRET histograms (0.1-0.6 FRET efficiency) from panel c is plotted versus the U9 RNA concentration to estimate the K_d. For normalization, the area for the apo measurement was set to zero and for the holo measurement at 25 µM was set to 1. The affinity of U9 RNA binding to U2AF62 was estimated to be ~1.2 µM using the standard model for receptor-ligand kinetics as described in Supplementary Figure 5.



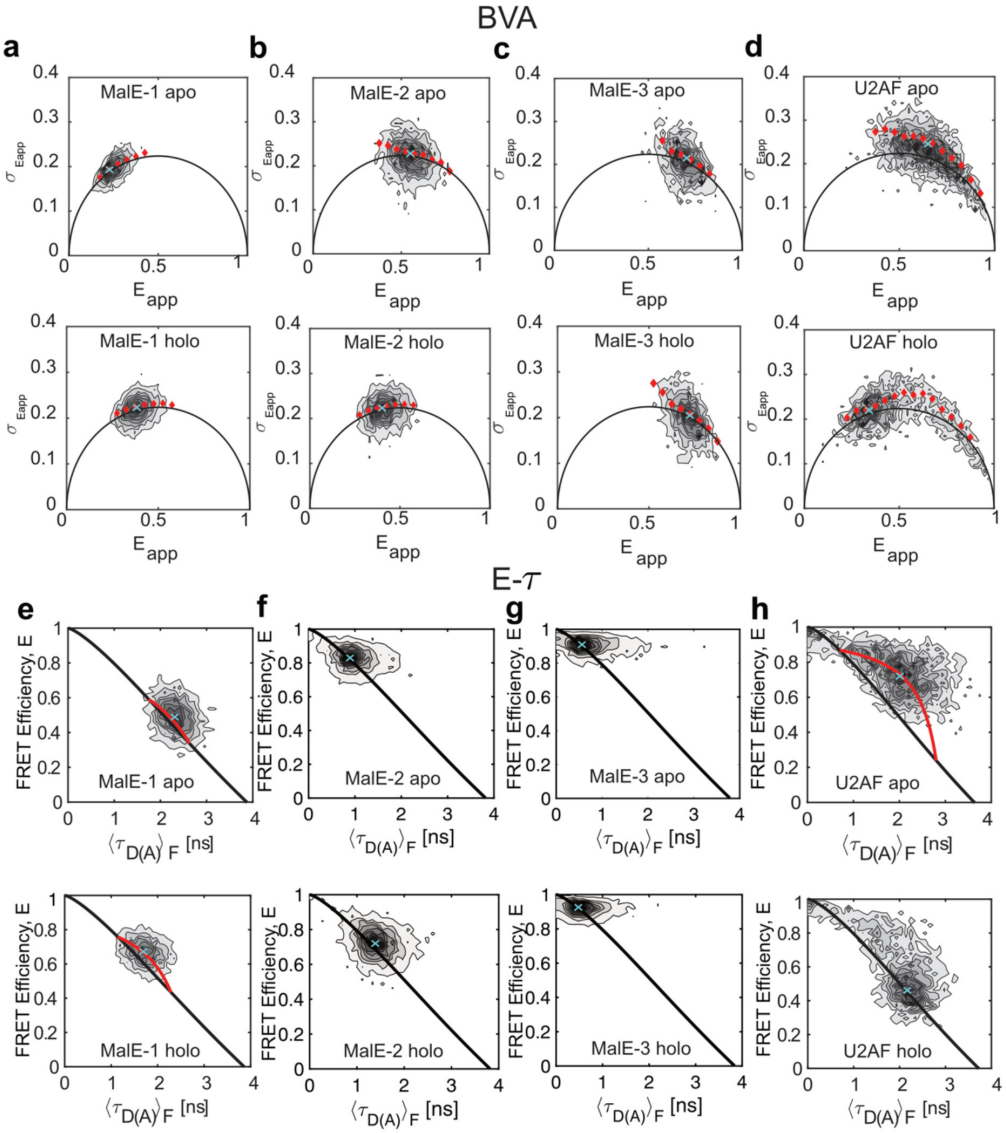
Supplementary Figure 8: Comparison of the FRET efficiency reported for different labs and after reanalysis for the MalE-1 mutant under apo conditions. **(a, b)** The determined average FRET efficiency values (**a**) reported by the different laboratories and (**b**) after reanalysis are shown as squares for measurements of MalE-1 in the apo states for 7 laboratories. The mean value (upper data point) from all data sets with the corresponding standard deviation is shown in grey. For the details of the reanalysis procedure, please refer to the Supplementary Note 3. The mean FRET Efficiency and standard deviation was 0.474 and 0.063 for the reported values and 0.487 and 0.054 after reanalysis respectively. For reanalysis purposes, we used data from the same 7 labs which measured the dynamics of U2AF.



Supplementary Figure 9: Sample-dependent setup statistics. Sample-dependent distributions of the setup-dependent parameters are given as violin plots for (*top*) the average count rate, (*second from the top*) number of detected bursts, (*third from the top*) average number of photons per burst and (*bottom*) average burst duration. The procedure to obtain the values for the above parameters from the measurement data collected from 8 labs for both MalE and U2AF samples is as follows: The collected raw data from different labs was analyzed using the PAM software[1111]. Briefly, first, a burst search was performed using an all photon burst search with a threshold of 50-100 photons per sliding time window of 500 µs depending on the dataset. For one set of measurements, a lower threshold of 20 photons per 500 µs time window was necessary. After burst selection, background subtraction and correction for crosstalk and direct excitation were performed as discussed in the data analysis section. To remove blinking and bleaching events as well as for selecting out the double-labeled molecules, an ALEX-2CDE filter[12] with a lower limit of 5 and an upper limit of 25 was used depending on the data set. The ALEX-2CDE filter values differed depending on the excitation intensities and sample concentrations used for the measurements. The values for all the parameters are a median of the values obtained for the double-labeled molecules for each measurement. These values were made available with PAM software.

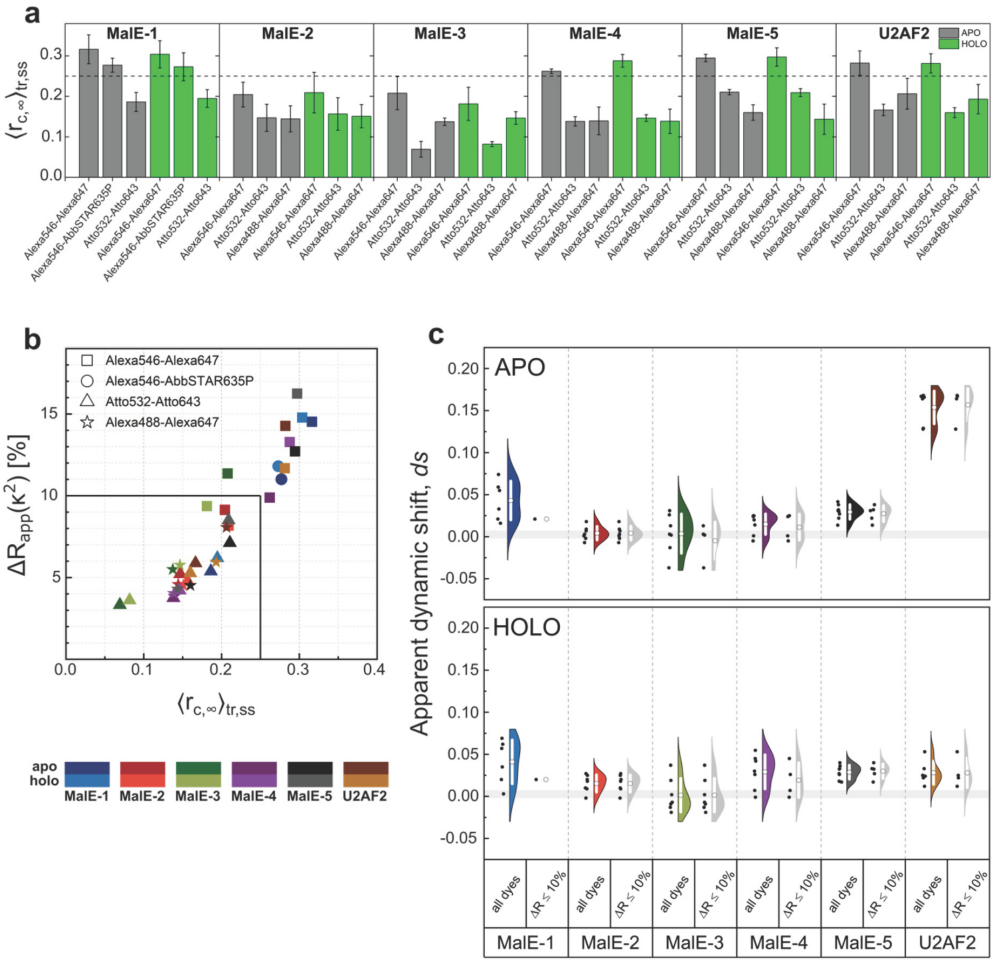


Supplementary Figure 10: Correlations between setup-dependent parameters. Pairwise plots of the number of detected bursts, average number of photons per burst, average burst duration and average burstwise count rate shown in Supplementary Figure 9. *r* is the Pearson's correlation coefficient. Error bands indicate the 95% confidence intervals of the regression. The one-dimensional projections show the distribution of the parameters as histograms (gray bars) and kernel density estimates (black lines).

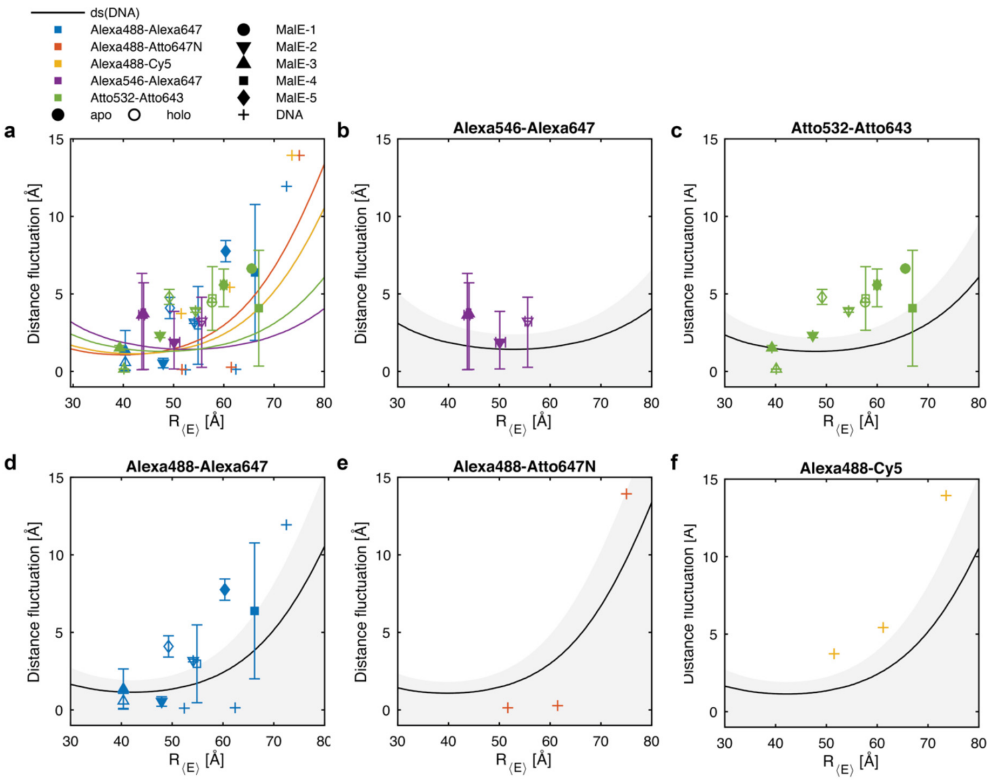


Supplementary Figure 11: Overview of conformational dynamics and determination of the dynamic shift on the sub-millisecond timescale in MalE labeled with Alexa546-Alexa647 and U2AF2 labeled with Atto532-Atto643. (a-d) Burst variance analysis (BVA) of MalE-1 (a), MalE-2 (b), MalE-3 (c), and U2AF (d) under both apo (upper panel) and holo (lower panel) conditions. In BVA, the standard deviation $\sigma_{E_{app}}$ of the apparent FRET efficiency E_{app} is compared to the shot-noise limit given by $\sigma_{E_{app}} = \sqrt{E_{app}(1-E_{app})/n}$ (black line, here n = 5). Single-molecule events with conformational dynamics show an increased variance and follow the dynamic line (red diamonds). Red diamonds indicate the average standard deviation of all bursts within a FRET efficiency range of 0.05. The mean positions of the populations (cyan crosses) were determined by fitting a two-dimensional Gaussian distribution to the data (see Supplementary Note 5). The dynamic shift, *ds*, is defined as the excess standard deviation compared to the static line. A clear deviation from the static line is observed for the apo state of U2AF2. For U2AF2 under holo conditions, the *ds* was determined for the major holo state population. Please note the leftover minor apo state population in holo condition has a similar deviation as for the apo condition. **(e-h)** Plots of the FRET efficiency E versus intensity-weighted average donor lifetime $\langle\tau_{D(A)}\rangle_F$ (E-τ) for MalE-1 (e), MalE-2 (f), MalE-3 (g), and U2AF (h) under both apo (upper panel) and holo (lower panel)

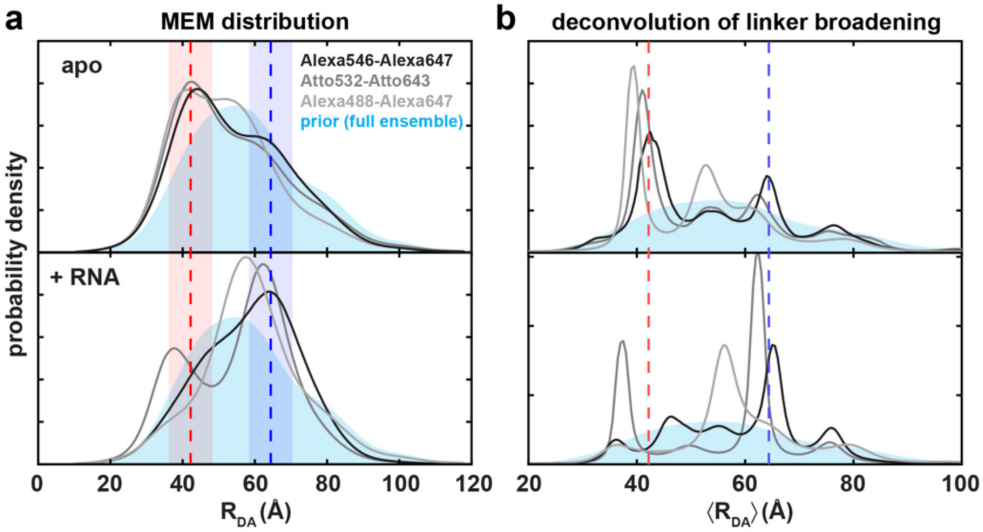
conditions. In the E-τ plot, the intensity-based FRET efficiency E is plotted against the intensity-weighted average donor fluorescence lifetime, $\langle\tau_{D(A)}\rangle_F$. The static FRET-line is given by the Förster relation as $E = 1 - \frac{\langle\tau_{D(A)}\rangle_F}{\tau_{D(0)}}$ (black). The static lines are slightly curved as they account for the flexibility of the dye linkers. Molecules undergoing dynamics are shifted from the static line and follow a dynamic FRET-line (red). The mean positions of the populations (cyan crosses) were determined by fitting a two-dimensional Gaussian distribution to the data (see Supplementary Note 5). For a given population, the dynamic shift is defined as the displacement of the population orthogonal to the static FRET-line. A clear dynamic shift is observed for U2AF2. The slight dynamic shift observed in MalE-1 is due to the high anisotropy of attached dyes as identified later in the study (see Fig. 5e,f and Supplementary Figure 12). The end points of the dynamic FRET-line for MalE-1 and U2AF2 were determined from a sub-ensemble analysis of the fluorescence decay. For U2AF2 under holo conditions, the *ds* was determined for the major holo state population. Please note that the leftover minor apo state population in the holo measurement has a similar deviation as for the apo condition.



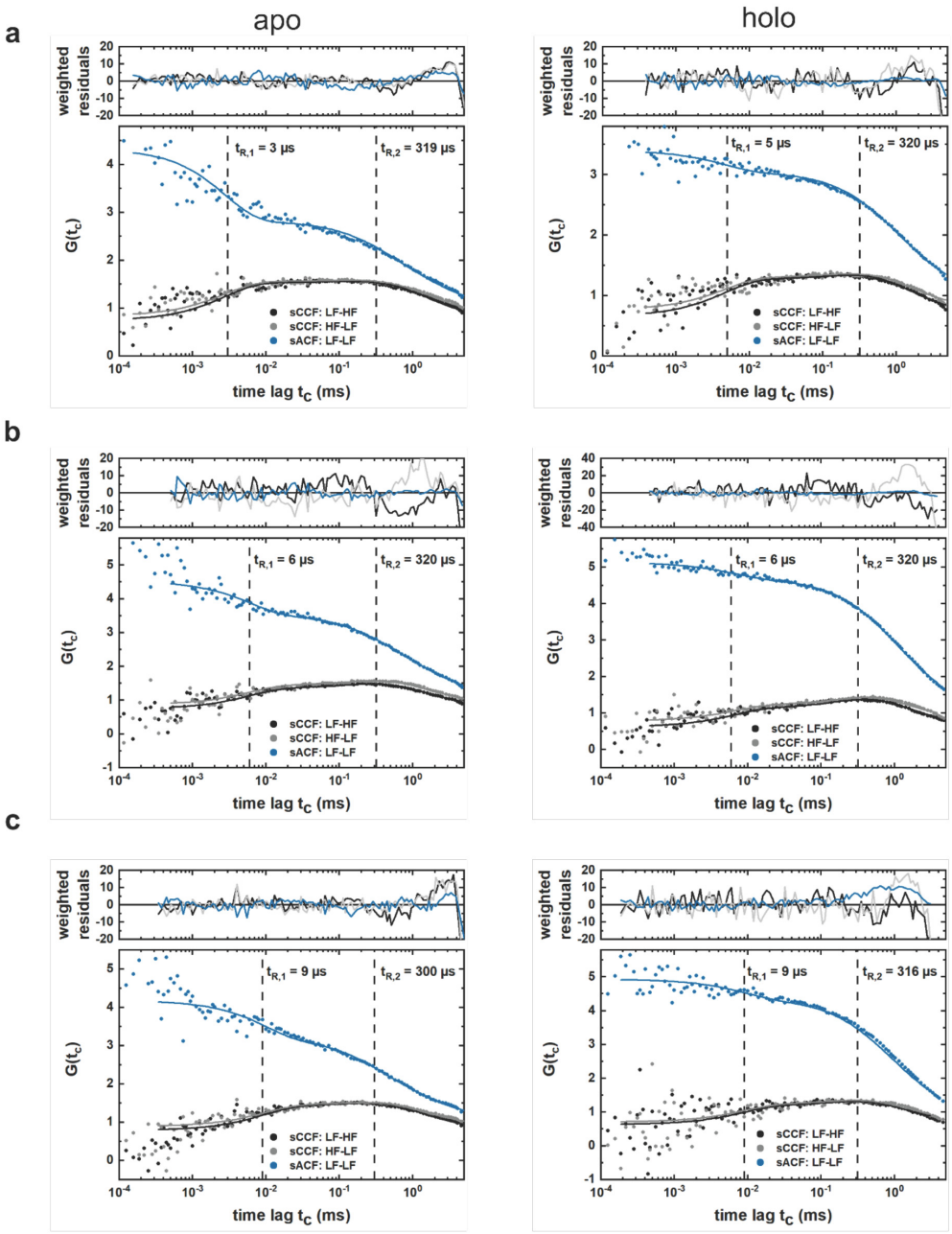
Supplementary Figure 12: Dye-specific sticking interactions in proteins and residual apparent dynamic shift upon removal of sticky dyes. (a) Combined residual anisotropies $\langle r_{c,\infty}\rangle_{tr,ss}$ for the different dye combinations of MalE and U2AF2 variants (see Supplementary Note 8 and Supplementary Table 13). Error bars represent the propagated 1σ uncertainty of the donor and acceptor residual anisotropies on the reported combined residual anisotropy as described in Supplementary Note 8, section c, equation 8.12-8.15. **(b)** The distance uncertainty due to the orientation factor κ^2, $\Delta R_{\mathrm{app}}(\kappa^2)$ (see Supplementary Note 9 and Supplementary Table 14), and its correlation with the combined residual anisotropy $\langle r_{c,\infty}\rangle_{tr,ss}$. A threshold of 10% in terms of distance uncertainty (solid black line) corresponds to $\langle r_{c,\infty}\rangle_{tr,ss}$= 0.25, which was used to filter-out sticking-prone dyes. **(c)** Violin plots of the dynamic shift values reported by three labs (see Supplementary Table 8) including all measured dye combinations and after filtering of the dyes to remove those with pronounced sticking interactions using a threshold of $\Delta R_{\mathrm{app}}(\kappa^2) = 10\%$ corresponding to $\langle r_{c,\infty}\rangle_{tr,ss}$ = 0.25. In the violins, the white circle and whisker denote the mean ± 1σ. The shaded grey area gives the upper and lower limit of apparent dynamic shift obtained from DNA samples with different FRET efficiencies and with different dye combinations (Supplementary Table 8). Before filtering, the number of measurements over the three different labs and three different dye pairs are n=5 for MalE-1, n=7 for MalE variants 2-4, n=8 for MalE-5, and n=6 for U2AF (Supplementary Table 8). After filtering, the remaining number of measurements are n=1 for MalE-1, n=7 for MalE-2 and MalE-3, n=4 for MalE-4, n=5 for MalE-5, and n=4 for U2AF.



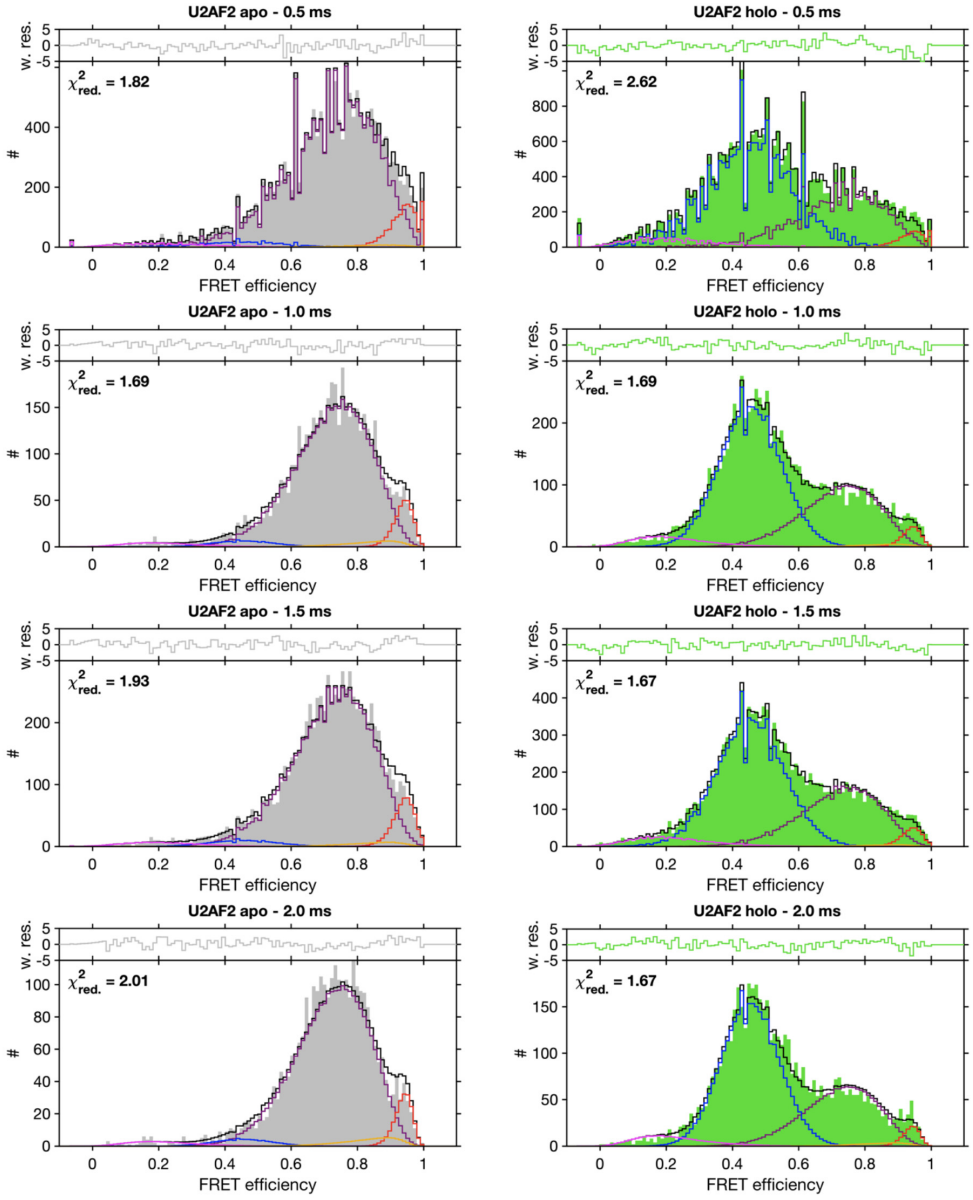
Supplementary Figure 13: Estimated conformational flexibility of MalE based on the residual dynamic shift for different donor-acceptor pairs. (a) The estimated distance fluctuations are plotted against the FRET-averaged interdye distance for the give MalE mutants in the apo and holo state labeled with the dye pairs Alexa546-Alexa647, Atto532-Atto643 and Alexa488-Alexa647. Additional control measurements on dsDNA are shown as crosses for the dye pairs Alexa488-Alexa647, Alexa488-Atto647N and Alexa488-Cy5. The lines indicate the apparent distance fluctuation obtained for the dsDNA control measurements, calculated based on a dynamic shift of $ds_{DNA} = 0.0026 \pm 0.0044$ measured in one lab. Gray areas indicate the 1σ confidence interval for the apparent distance fluctuations obtained for the dsDNA control measurements. All values are given in Supplementary Table 8. Error bars represent the 1σ percentiles of the distance fluctuation estimated for the remaining measurements after filtering for the individual dye pairs (for Alexa546-Alexa647, n=3 for MalE variants 2 and 3; for Atto532-Atto643, n=1 for MalE-1, n=2 for MalE variants 2-4, and n=3 for MalE-5; for Alexa488-Alexa647, n=2 for MalE variants 2-5). **b-f)** Individual plots of the data shown in panel a for the different dye combinations.



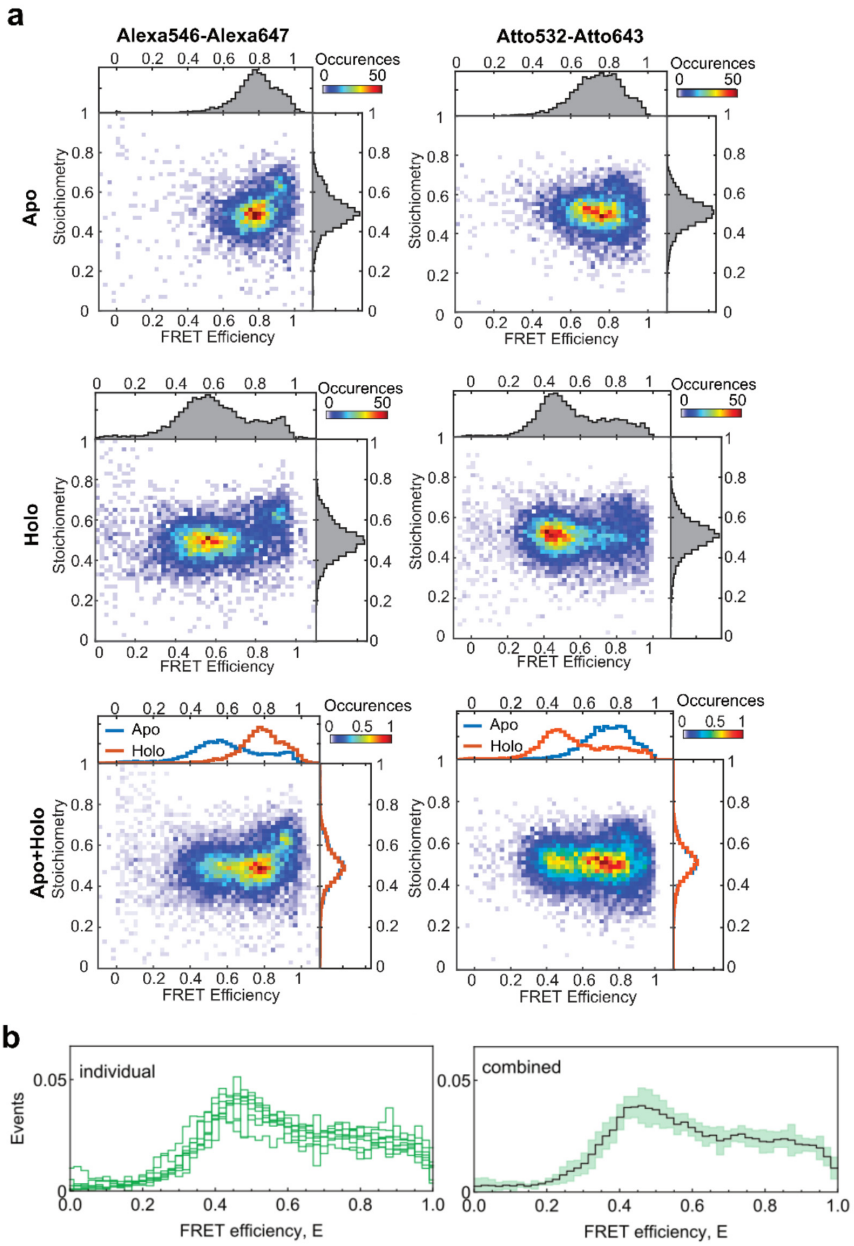
Supplementary Figure 14: Comparison of the distance distributions obtained using different donor-acceptor dye pairs. Comparison of the distance distributions obtained using different donor-acceptor dye pairs. **(a)** The donor-acceptor distance R_{DA} distribution determined from the donor fluorescence decay using a model-free MEM approach for different dye-pairs where the distance distribution of the NMR/SAXS full ensemble of structures[47] (shown in light blue) was used as the prior distribution. The expected interdye distances for the resolved structure of the compact apo and open holo states are shown as red and blue dashed lines (PDB: 2YH0, 2YH1) with the shaded areas indicating the distance broadening due to the flexible dye linkers of 6 Å. This panel is identical to Fig. 6d. **(b)** The distribution of the mean donor-acceptor distance R_{DA} obtained by the MEM in panel a deconvolved as described in Supplementary Note 15 to remove the broadening of the linker. Well-defined peaks at distances similar to the compact apo (red dashed line) and open-holo (blue dashed line) conformations become clearly visible.



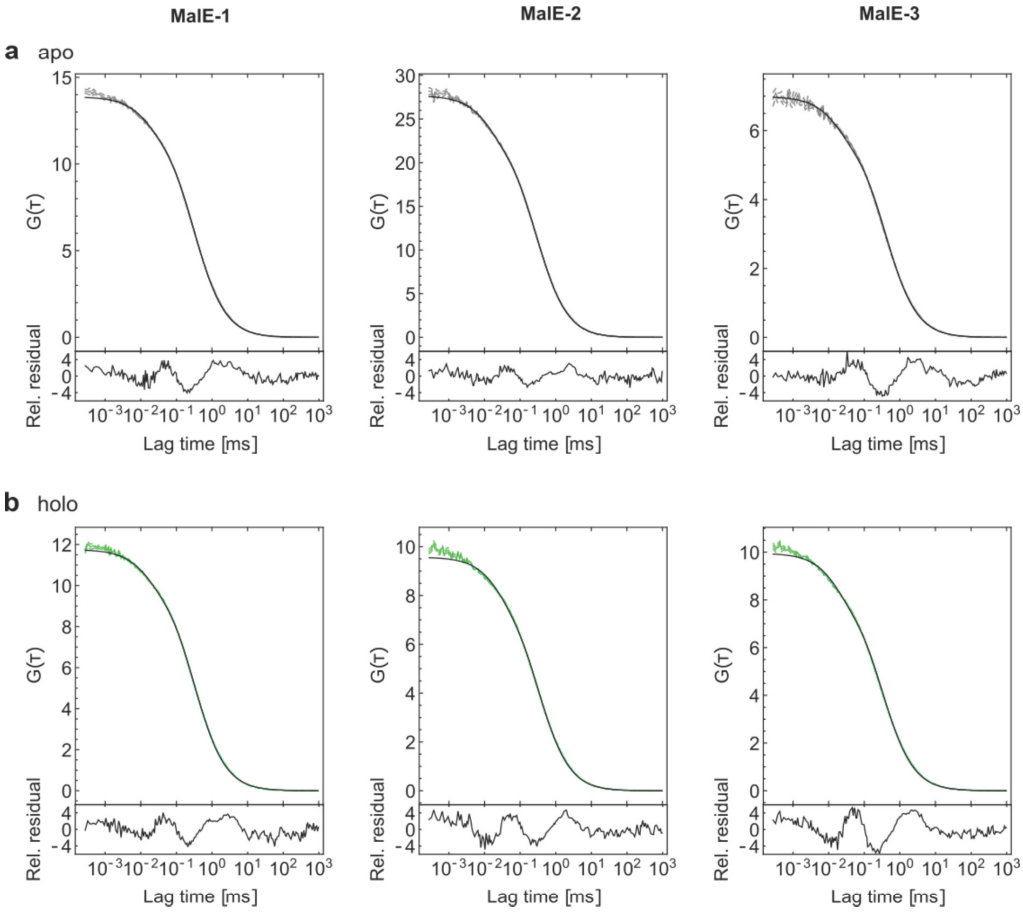
Supplementary Figure 15: Filtered fluorescence correlation spectroscopy curves for U2AF2. fFCS curves are shown for three different dye combinations, namely (a) Alexa546-Alexa647 (b)Atto532-Atto643 (c) Alexa488-Alexa647. In the analysis, a global fit of the two species autocorrelation functions, sACF, and two species cross-correlation functions, sCCF was performed. For simplicity, only one sACF function is shown. The fit model consisted of a diffusion term and two kinetic terms with corresponding relaxation times, $t_{R,1}$ and $t_{R,2}$ (Supplementary Note 16). The obtained relaxation times (denoted as vertical lines) are consistent between different dye combinations as well as between apo and holo states.



Supplementary Figure 16: A global dynamic photon distribution analysis (PDA) of apo and holo measurements of U2AF2 labeled with the Atto532-Atto643 dye pair. Different binning times (from top to bottom: 0.5 ms, 1 ms, 1.5 ms and 2.0 ms) for *left*: apo measurements (in grey) and *right*: holo measurement (in green) are shown. The analysis was performed globally over the apo and holo measurements using integration windows of 0.5, 1.0, 1.5 and 2.0 ms and shared distances for the compact apo conformation (red), the detached apo ensemble (purple) and the holo state (blue). The dynamic interconversion between the compact apo state and apo conformational ensemble is shown in yellow. An additional low-FRET population (magenta) had to be included, which most likely originates from photobleaching. All distances and kinetic rates for the apo ensemble are globally optimized while the amplitudes of the apo and holo populations were kept constant within the integration time windows for the apo and holo states. See Supplementary Note 17 for details.



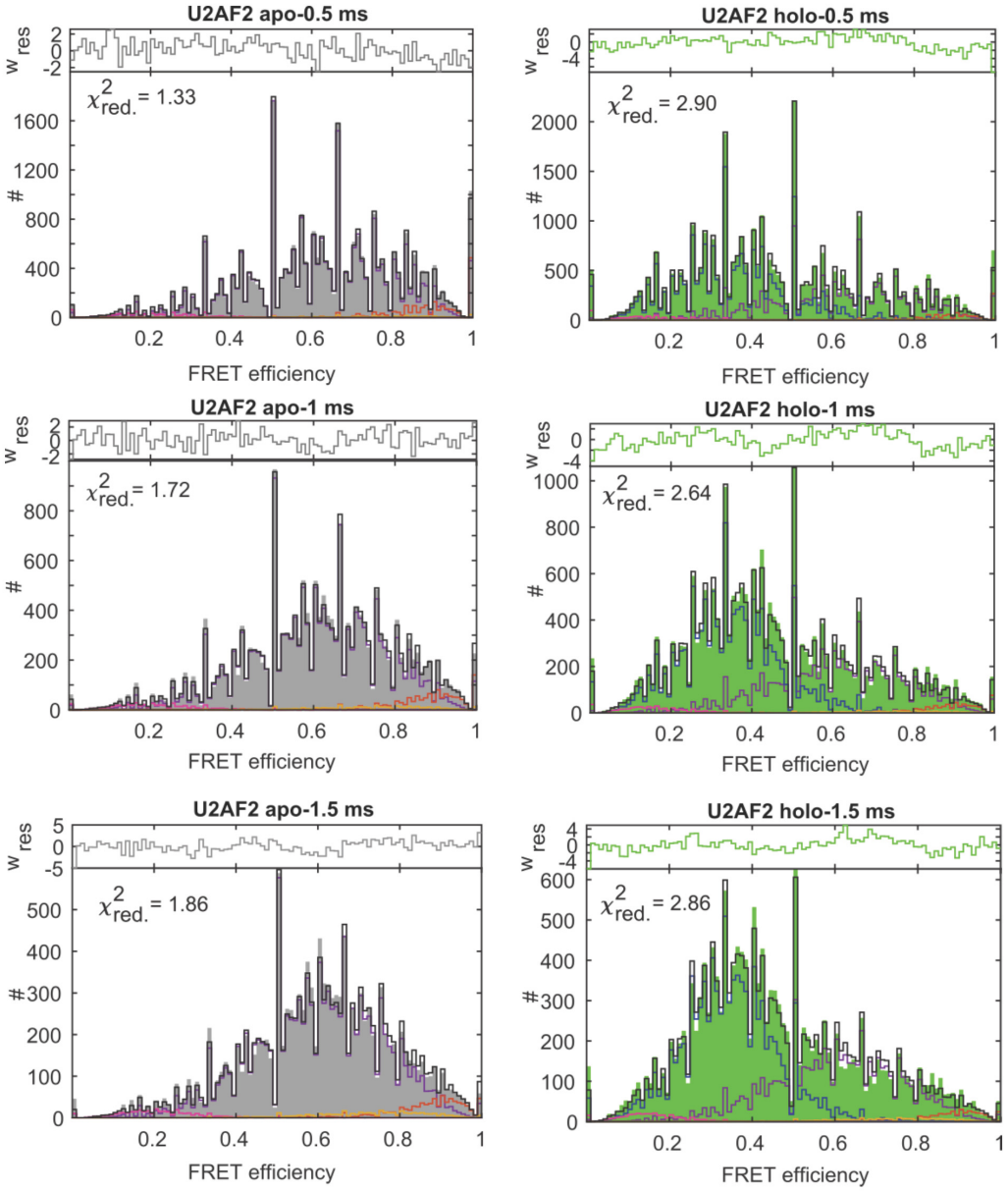
Supplementary Figure 17: The very-high-FRET population in U2AF2. (a) FRET efficiency versus stoichiometry plots are shown for two different dye combinations (Alexa546-Alexa647, left panel and Atto532-Atto643, right panel). These are plotted for apo (upper graph), holo (middle graph) and combined both apo and holo (lower graphs). A very-high-FRET populations with a different stoichiometry is visible in all measurements but with different amplitudes. **(b)** The FRET efficiency histograms (left) from the individual laboratories and (right) the combined histogram showing the mean (solid line) and a standard deviation (pale) after avoiding the subpopulation with slight acceptor quenching by using only bursts with a stoichiometry between 0.2-0.5 to build the smFRET histograms measured by the different labs.



Supplementary Figure 18: FCS experiments to rule out the presence of large aggregates. Fluorescence Correlation Spectroscopy (FCS) measurements for the mutants MalE-1 (left), MalE-2 (middle), and MalE-3 (right) in **(a)** the apo state and **(b)** in the holo state are plotted. The two orthogonally oriented polarizations in the donor detection channel were correlated to remove the detector afterpulsing at short lag times (green line). All correlation curves were fitted with a standard model including a triplet fraction (black line):

$$G(t) = \frac{\gamma}{\langle N \rangle}\left(1 + \frac{t}{\tau_{\text{diff}}}\right)^{-1}\left(1 + \frac{t}{\tau_{\text{diff}}p^2}\right)^{-\frac{1}{2}}\left(1 + \frac{T}{1-T}e^{-\frac{t}{\tau_{\text{trip}}}}\right),$$

where T is the triplet fraction, τ_{trip} the triplet lifetime, p is the structural or elongation factor of the confocal volume given as the ratio of the axial and lateral dimensions, τ_{diff} is the diffusion time, and $\gamma = 2^{-3/2}$ is the geometric factor[97]. The confocal instrument was calibrated using free Alexa546 dye with a published[98] diffusion coefficient of 341 $\mu m^2 s^{-1}$ for which a diffusion time of 95±10 µs was found. The overall diffusion time for the six measurements was 325±40 µs. None of the correlation curves show any indication of the presence of protein aggregates.



Supplementary Figure 19: A global dynamic photon distribution analysis (PDA) of the apo and holo measurements of U2AF2 labeled with the Alexa546-Alexa647 dye-pair. Left: apo measurement (in grey), right: holo measurement (in green). The analysis was performed globally over the apo and holo measurements for integration windows of (*top*) 0.5 ms, (*middle*) 1.0 ms, and (*bottom*) 1.5 ms using shared distances for the compact apo conformation (red), the detached apo ensemble (purple) and the holo state (blue). The dynamic interconversion between the compact apo state and the apo conformational ensemble is shown in yellow. All distances and kinetic rates for the apo ensemble are globally optimized while the amplitudes of the apo and holo populations were kept constant within the integration time windows for the apo and holo states. For details, see Supplementary Note 17.

Supplementary Tables

Supplementary Table 1: FRET efficiency correction factors reported by the participating labs for MalE constructs labeled with Alexa546-Alexa647. Overview of all correction factors for MalE and the resulting change in the FRET efficiency upon application of the correction factors is shown as an example for the holo state of the MalE-2 mutant. The correction factors: α is spectral crosstalk of donor fluorescence into the acceptor channel, ß is the normalization of direct donor and acceptor excitation fluxes, γ is the differences in donor and acceptor quantum yields and detection efficiencies and δ is the ratio of indirect and direct A excitation. $\langle E_{app} \rangle$ is the uncorrected FRET efficiency and $\langle E \rangle$ is the corrected FRET efficiency.

Lab#	$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$	$\boldsymbol{\gamma}$	$\boldsymbol{\delta}$	$\langle E_{app} \rangle$	$\langle E \rangle$
1	0.03	2.38	0.23	0.08	0.43	0.72
2	0.05	0.5	0.34	0.32	0.54	0.73
3	0.05	1.70	0.51	0.07	0.56	0.66
4	0.04	2.65	0.31	0.06	0.42	0.66
5	0.04	1.56	0.34	0.14	0.54	0.72
6	0.08	1.59	0.54	0.12	0.57	0.66
7	0.07	0.64	0.63	0.12	0.59	0.70
8	0.04	1.53	0.48	0.11	0.47	0.63
9	0.06	1.84	0.47	0.10	0.57	0.69
10	0.03	1.88	0.31	0.11	0.46	0.72
11	0.03	2.37	0.32	0.07	0.43	0.64
12	0.04	1.99	0.25	0.07	0.47	0.74
13	0.06	0.60	0.34	0.32	0.58	0.76
14	0.04	2.00	0.23	0.09	0.41	0.77
15	0.05	1.42	0.33	0.05	0.57	0.79
16	0.05	1.31	0.46	0.14	0.60	0.72
17 *	0.06	1.26	0.55	0.06		
18 **	0.01	4.86	0.09	0.10	0.18	0.59
19 ***	0.18				0.59	

*Due to measurement problems with a large photobleaching and/or blinking contribution, data could not be fitted with a Gaussian distribution (the FRET population was small with most bursts containing photobleaching events). Data were not considered in the evaluation of the mean and standard deviation.

**The measurements were performed in a regime with a $\gamma < 0.1$, where the error in γ is significantly increased. Data were not considered in the evaluation of the mean and standard deviation.

***Due to measurement problems, the data could not be corrected for direct excitation (δ) and γ. These data were not considered in the evaluation of mean and standard deviation.

Supplementary Table 2: FRET efficiency correction factors reported by the 7 labs for apo and holo measurements on U2AF2 labeled with Atto532-Atto643. α is the spectral crosstalk of donor fluorescence into the acceptor channel, ß is the normalization of direct donor and acceptor excitation fluxes, γ is the differences in donor and acceptor quantum yields and detection efficiencies and δ is the ratio of indirect and direct acceptor excitation.

Lab#	α	β	γ	δ
1	0.02	0.78	0.59	0.06
2	0.06	-	1.1	0.23
3	0.04	0.64	0.80	0.05
4	0.03	1.05	0.83	0.09
8	0.03	0.91	0.73	0.05
11	0.05	-	0.64	0.02
14	0.05	-	0.64	0.09

Supplementary Table 3: Reported mean FRET Efficiency and standard deviation for MalE samples by the participating laboratories: The mean FRET efficiency values $\langle E \rangle$ and distribution widths σ_E provided by the participating labs for all the three studied mutants of MalE labeled with Alexa546 and Alexa647 under both apo and holo conditions are listed. The difference in the reported mean FRET efficiency for individual laboratories between the apo and holo state was calculated as $\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle$. The calculated mean and standard deviation of the reported FRET efficiencies over all labs are given by $\mu_{\langle E \rangle}$ and $\sigma_{\langle E \rangle}$ and for the FRET efficiency difference, $\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle$, by $\boldsymbol{\mu}_{\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle}$ and $\sigma_{\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle}$, respectively.

	MalE-1					MalE-2					MalE-3					MalE-4			MalE-5		
Lab#	Apo		Holo		$\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle$	Apo		Holo		$\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle$	Apo		Holo		$\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle$	Apo	Holo	$\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle$	Apo	Holo	$\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle$
	$\langle E \rangle$	σ_E	$\langle E \rangle$	σ_E		$\langle E \rangle$	σ_E	$\langle E \rangle$	σ_E		$\langle E \rangle$	σ_E	$\langle E \rangle$	σ_E		$\langle E \rangle$	$\langle E \rangle$		$\langle E \rangle$	$\langle E \rangle$	
1	0.500	0.128	0.674	0.098	0.174	0.832	0.061	0.724	0.089	-0.108	0.908	0.039	0.922	0.035	0.014	0.424	0.666	0.242	0.615	0.822	0.207
2	0.520		0.700		0.180	0.840		0.730		-0.11	0.910		0.910		0	0.46	0.69	0.23	0.61	0.82	0.21
3	0.453	0.096	0.641	0.077	0.188	0.794	0.058	0.654	0.072	-0.14	0.904	0.040	0.915	0.039	0.011						
4	0.461	0.169	0.637	0.119	0.176	0.785	0.096	0.658	0.126	-0.127	0.896	0.066	0.901	0.064	0.005						
5	0.522		0.703		0.181	0.845		0.724		-0.121	0.925		0.912		-0.013						
6	0.509	0.129	0.644	0.109	0.135	0.820	0.073	0.657	0.090	-0.163	0.900	0.054	0.911	0.060	0.011						
7	0.454	0.119	0.641	0.086	0.187	0.807	0.058	0.702	0.068	-0.105	0.917	0.046	0.914	0.044	-0.003						
8	0.414	0.130	0.602	0.105	0.188	0.771	0.057	0.627	0.101	-0.144	0.881	0.044	0.890	0.041	0.009						
9	0.451	0.147	0.622	0.118	0.171	0.820	0.069	0.685	0.095	-0.135	0.911	0.054	0.911	0.053	0						
10	0.444	0.180	0.631	0.144	0.187	0.834	0.081	0.713	0.102	-0.121	0.923	0.062	0.937	0.073	0.014						
11	0.4	0.112	0.610	0.087	0.210	0.781	0.063	0.644	0.075	-0.137	0.876	0.054	0.887	0.047	0.011						
12	0.523	0.118	0.706	0.081	0.183	0.849	0.052	0.74	0.073	-0.109	0.921	0.039	0.924	0.037	0.003						
13	0.547	0.151	0.723	0.105	0.176	0.872	0.065	0.761	0.093	-0.111	0.942	0.051	0.952	0.043	0.010						
14	0.580		0.745		0.165	0.865		0.765		-0.1	0.935		0.970		0.035						
15	0.621	0.114	0.76	0.072	0.139	0.877	0.047	0.787	0.066	-0.09	0.936	0.035	0.936	0.034	0						
16	0.466	0.100	0.655	0.082	0.189	0.831	0.050	0.716	0.067	-0.115	0923	0.038	0.93	0.038	0.007						
$\boldsymbol{\mu}_{\langle E \rangle}$ or $\boldsymbol{\mu}_{\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle}$	0.492		0.667		0.177	0.826		0.705		-0.121	0.913		0.920		0.007	0.442	0.678	0.236	0.613	0.821	0.209
$\boldsymbol{\sigma}_{\langle E \rangle}$ or $\boldsymbol{\sigma}_{\langle E_{\mathrm{holo}} \rangle - \langle E_{\mathrm{apo}} \rangle}$	0.060		0.049		0.019	0.032		0.047		0.019	0.019		0.021		0.010	0.025	0.017	0.008	0.004	0.001	0.002

Supplementary Table 4: Reported and reanalyzed mean FRET efficiency and standard deviations for the U2AF2 experiments: The U2AF2 protein was labeled with Atto532-Atto643 and labs provided the number of events for FRET efficiency over 51 bins between 0 and 1. To obtain the mean FRET efficiency $\langle E \rangle$ and distribution width σ_E values for U2AF2, the reported events were fitted with one and two Gaussian distribution functions for apo and holo conditions respectively. For reanalysis, the original datasets from the individual labs were obtained and reanalyzed according to the procedure detailed in the Supplementary Note 3. U2AF2 under holo conditions has a major fraction of holo state and a minor fraction of apo state due to incomplete saturation of the RNA ligand. Only 7 labs participated in the U2AF2 study because of its complexity. The calculated mean and standard deviation of the reported and reanalyzed FRET efficiencies over all labs for each sample are given by $\mu_{\langle E \rangle}$ and $\sigma_{\langle E \rangle}$ respectively.

Lab #	Reported						Reanalyzed					
	Apo		+RNA				Apo		+RNA			
			Holo state		Apo state				Holo state		Apo state	
	$\langle E \rangle$	σ_E	$\langle E \rangle$	σ_E	$\langle E \rangle$	σ_E	$\langle E \rangle$	σ_E	$\langle E \rangle$	σ_E	$\langle E \rangle$	σ_E
1	0.748	0.192	0.436	0.150	0.756	0.182	0.736	0.200	0.405	0.158	0.769	0.243
2	0.749	0.202	0.455	0.168	0.811	0.174	0.757	0.207	0.432	0.157	0.800	0.208
3	0.708	0.228	0.506	0.206	0.785	0.142	0.738	0.262	0.449	0.163	0.767	0.229
4	0.728	0.205	0.414	0.197	0.773	0.152	0.736	0.211	0.385	0.158	0.768	0.179
8	0.795	0.169	0.521	0.206	0.857	0.117	0.742	0.217	0.414	0.192	0.814	0.216
11	0.729	0.204	0.423	0.133	0.727	0.231	0.736	0.196	0.429	0.112	0.707	0.269
14	0.718	0.199	0.430	0.164	0.743	0.195	0.749	0.214	0.448	0.131	0.755	0.241
$\mu_{\langle E \rangle}$	0.74		0.46		0.78		0.742		0.423		0.77	
$\sigma_{\langle E \rangle}$	0.03		0.04		0.04		0.008		0.023		0.03	

Supplementary Table 5: Global fit of the polarization-resolved and magic-angle fluorescence decays from sub-ensemble data of MalE and U2AF2 samples. The different rotational correlation times, ρ_j, with corresponding amplitudes, b_j, for the $A_{ex}|A_{em}$ channels (acceptor dye) and $D_{ex}|D_{em}$ channels (donor dye) on a sub-ensemble DA population describe different depolarization processes. The fluorescence lifetimes, τ_i, and fraction of molecules with a given lifetime, x_i, were obtained from the magic-angle decay function, which was fitted globally with the polarization-resolved fluorescence decays, $f_{VV}(t)$ and $f_{VH}(t)$, as described in Supplementary Note 8 Eqn. 8.4 - 8.11. The fit quality was judged by the reduced Chi-squared value, χ_r^2. Fitted residual anisotropies for different dye combinations are used in the calculation of relative distance uncertainties (Fig. 5e), and it is shown that they are correlated to the observed dynamic shift (Fig. 5f).

	$A_{ex}\|A_{em}$ channels										
Sample	**Dye combination**	**Condition**	ρ_1 [ns]	b_1= $r_{\infty,tr}$	ρ_2 [ns]	b_2	τ_1 [ns]	x_1	τ_2 [ns]	x_2	χ_r^2
MalE-1	Alexa546-Alexa647	apo	>50	0.331	0.84	0.049	1.69	0.50	0.93	0.50	0.96
	Alexa546-AbberiorSTAR635P	apo	20	0.209	0.16	0.171	4.00	1.00			0.98
	Atto532-Atto643	apo	20	0.181	0.27	0.199	3.95	1.00			0.99
	Alexa546-Alexa647	holo	20	0.314	0.10	0.066	1.66	0.57	0.95	0.43	1.39
	Alexa546-Abberior STAR635P	holo	20	0.182	0.12	0.198	4.00	1.00			0.98
	Atto532-Atto643	holo	20	0.212	0.32	0.168	3.80	1.00			1.15
MalE-2	Alexa546-Alexa647	apo	20	0.212	0.24	0.168	1.39	0.61	0.75	0.39	1.15
	Atto532-Atto643	apo	20	0.160	0.33	0.220	3.61	1.00			1.08
	Alexa488-Alexa647	apo	>50	0.150	0.57	0.230	1.30	1.00			0.94
	Alexa546-Alexa647	holo	20	0.228	0.25	0.152	1.48	0.48	0.91	0.52	1.10
	Atto532-Atto643	holo	20	0.186	0.45	0.194	3.62	1.00			1.00
	Alexa488-Alexa647	holo	20	0.150	0.49	0.230	1.72	0.21	1.22	0.79	1.11
MalE-3	Alexa546-Alexa647	apo	20	0.215	0.27	0.165	1.37	0.58	0.75	0.42	1.09
	Atto532-Atto643	apo	20	0.149	0.31	0.231	3.78	1.00			1.1
	Alexa488-Alexa647	apo	20	0.192	0.62	0.188	1.37	1.00			1.08
	Alexa546-Alexa647	holo	20	0.201	0.28	0.179	1.45	0.49	0.82	0.51	1.16
	Atto532-Atto643	holo	20	0.147	0.28	0.233	3.67	1.00			1.12
	Alexa488-Alexa647	holo	20	0.185	0.44	0.195	1.33	1.00			1.16
MalE-4	Alexa546-Alexa647	apo	>50	0.205	0.59	0.175	1.30	0.68	0.60	0.32	0.83
	Atto532-Atto643	apo	20	0.144	0.42	0.236	3.84	1.00			1.09
	Alexa488-Alexa647	apo	20	0.113	0.29	0.267	1.50	0.58	0.88	0.42	1.16
	Alexa546-Alexa647	holo	20	0.217	0.32	0.163	1.67	0.27	0.91	0.73	0.9
	Atto532-Atto643	holo	20	0.140	0.29	0.240	3.53	0.84	5.08	0.16	1.07
	Alexa488-Alexa647	holo	20	0.121	0.30	0.259	1.84	0.23	1.12	0.77	1.16
MalE-5	Alexa546-Alexa647	apo	>50	0.275	0.68	0.105	1.65	0.24	1.00	0.76	1.20
	Atto532-Atto643	apo	20	0.198	0.24	0.182	3.70	1.00			1.04
	Alexa488-Alexa647	apo	20	0.160	0.24	0.220	1.54	0.43	1.06	0.57	1.16
	Alexa546-Alexa647	holo	>50	0.296	1.20	0.084	1.87	0.16	1.00	0.84	1.08
	Atto532-Atto643	holo	20	0.201	0.41	0.179	3.40	1.00			1.16
	Alexa488-Alexa647	holo	20	0.108	0.34	0.272	1.68	0.38	1.00	0.62	1.11
U2AF2	Alexa546-Alexa647	apo	21	0.292	0.10	0.088	1.70	0.82	0.87	0.18	0.98
	Atto532-Atto643	apo	4	0.128	0.03	0.252	4.27	1.00			0.99
	Alexa488-Alexa647	apo	18	0.208	0.14	0.172	1.49	0.92	3.40	0.08	1.03
	Alexa546-Alexa647	holo	24	0.230	0.10	0.150	1.70	0.72	1.00	0.28	1.05
	Atto532-Atto643	holo	3.4	0.117	0.06	0.263	4.10	1.00			1.00
	Alexa488-Alexa647	holo	18	0.204	0.20	0.176	1.41	0.92	2.90	0.08	1.05

$D_{ex}\|D_{em}$ channels															
Sample	**Dye pair**	**Condition**	ρ_1 [ns]	b_1= $r_{\infty,tr}$	ρ_2 [ns]	b_2	ρ_3 [ns]	b_3	τ_1 [ns]	x_1	τ_2 [ns]	x_2	τ_3 [ns]	x_3	χ_r^2
MalE-1	Alexa546-Alexa647	apo	80	0.340	0.63	0.040			2.73	0.52	1.12	0.48			1.16
	Alexa546-Abberior STAR635P	apo	>50	0.327	1.60	0.050			2.87	0.49	1.18	0.51			1.03
	Atto532-Atto643	apo	20	0.229	0.21	0.151			2.22	0.79	3.62	0.21			1.24
	Alexa546-Alexa647	holo	46	0.329	0.16	0.050			1.69	0.48	0.77	0.41	3.09	0.11	1.19
	Alexa546-Abberior STAR635P	holo	>50	0.341	1.61	0.039			2.38	0.38	0.78	0.62			1.19
	Atto532-Atto643	holo	20	0.221	0.28	0.159			2.25	0.64	0.91	0.36			1.19
MalE-2	Alexa546-Alexa647	apo	>50	0.252	4.96	0.128			0.95	0.42	0.30	0.53	3.06	0.05	1.2
	Atto532-Atto643	apo	20	0.198	0.26	0.182			0.58	0.67	1.70	0.33			1.1
	Alexa488-Alexa647	apo	20	0.136	0.19	0.244			0.95	0.54	2.30	0.46			1.1
	Alexa546-Alexa647	holo	20	0.255	0.37	0.125			1.30	0.51	0.38	0.44	3.26	0.05	1.01
	Atto532-Atto643	holo	20	0.207	0.15	0.173			2.02	0.50	0.77	0.50			1.10
	Alexa488-Alexa647	holo	20	0.133	0.23	0.247			0.70	0.27	2.5	0.73			1.14
MalE-3	Alexa546-Alexa647	apo	20	0.255	0.26	0.125			0.33	0.81	1.04	0.18	6.30	0.01	1.03
	Atto532-Atto643	apo	20	0.048	0.14	0.332			0.45	0.87	2.02	0.13			1.00
	Alexa488-Alexa647	apo	20	0.087	0.10	0.293			0.66	0.69	2.10	0.31			1.14
	Alexa546-Alexa647	holo	20	0.187	0.25	0.193			0.33	0.78	1.02	0.19	3.20	0.03	1.13
	Atto532-Atto643	holo	20	0.044	0.15	0.336			0.45	0.87	2.24	0.13			1.09
	Alexa488-Alexa647	holo	20	0.116	0.15	0.264			0.47	0.64	1.80	0.36			0.96
MalE-4	Alexa546-Alexa647	apo	>50	0.344	0.57	0.036			2.60	0.69	0.80	0.31			1.04
	Atto532-Atto643	apo	20	0.126	1.40	0.05	0.10	0.202	2.82	0.8	1.17	0.20			1.13
	Alexa488-Alexa647	apo	20	0.13	0.15	0.250			3.30	0.87	1.53	0.13			1.18
	Alexa546-Alexa647	holo	>50	0.348	1.35	0.032			2.26	0.38	0.80	0.62			1.00
	Atto532-Atto643	holo	20	0.166	0.10	0.214			1.64	0.57	0.40	0.23	3.00	0.2	1.12
	Alexa488-Alexa647	holo	20	0.116	0.13	0.264			2.80	0.68	1.09	0.32			1.02
MalE-5	Alexa546-Alexa647	apo	>50	0.335	0.81	0.04			2.35	0.49	0.80	0.50			1.07
	Atto532-Atto643	apo	20	0.238	0.15	0.142			2.36	0.58	0.60	0.42			1.11
	Alexa488-Alexa647	apo	20	0.143	0.13	0.236			3.00	0.85	1.11	0.14			1.14
	Alexa546-Alexa647	holo	>50	0.338	0.29	0.042			1.65	0.34	0.40	0.66			1.1
	Atto532-Atto643	holo	20	0.233	0.15	0.147			1.70	0.42	0.50	0.58			1.16
	Alexa488-Alexa647	holo	20	0.118	0.10	0.262			2.37	0.51	0.74	0.49			1.07
U2AF2	Alexa546-Alexa647	apo	22	0.276	0.10	0.104			2.50	0.39	0.60	0.61			1.17
	Atto532-Atto643	apo	18	0.253	0.21	0.127			2.59	0.41	0.53	0.59			1.27
	Alexa488-Alexa647	apo	18	0.200	0.19	0.180			2.95	0.44	0.67	0.56			1.16
	Alexa546-Alexa647	holo	25	0.294	0.10	0.086			2.49	0.43	0.77	0.57			1.14
	Atto532-Atto643	holo	18	0.25	0.17	0.130			2.48	0.57	0.65	0.43			1.29
	Alexa488-Alexa647	holo	18	0.18	0.17	0.200			2.84	0.67	0.73	0.33			1.09

Supplementary Table 6: The statement of the individual laboratories regarding the dynamics of MalE and U2AF2. MalE-1: 29C/352C, MalE-2: 87C/186C; MalE-3: 134C/186C. "-" = no statement, "n/a"=not applicable due to experimental limitations (instrumentation & established evaluation routines).

Lab#	Method	Sub-ms dynamics in MalE-1/2/3	Sub-ms dynamics in U2AF2 (apo / holo)
1	BVA + E-Tau	no / no / no	yes / yes
2	E-Tau (fFCS, PDA)	yes / yes / yes	yes / yes
3	BVA + E-Tau	no / no / no	yes / yes
4	BVA + E-Tau	no / no / no	yes / yes
5	n/a	n/a	n/a
6	E-Tau (fFCS)	no / no / -	yes / yes
7	BVA	no / - / yes	n/a
8	BVA	no / no / no	yes / yes
9	BVA + E-Tau	no / no / no	n/a
10	BVA	no / no / no	n/a
11	BVA + E-Tau	no / no / no	yes / yes
12	BVA	no / no / no	n/a
13	fFCS	no / no / yes	n/a
14	BVA + E-Tau	no / no / no	n/a
15	n/a	n/a	n/a
16	BVA	no / no / no	n/a
17	n/a	n/a	n/a
18	n/a	n/a	n/a
19	n/a	n/a	n/a

Supplementary Table 7: The apparent dynamic shifts determined for both MalE and U2AF2 samples for the data collected from 8 labs. MalE and U2AF2 samples were labeled with Alexa546-Alexa647 and Atto532-Atto643 dye pairs respectively. The apparent dynamic shifts for the data collected in different labs are shown below for both BVA and E-τ plots. The dynamic shift, *ds* of the peak of the population was determined graphically as explained in Supplementary Note 5. For U2AF2 in the holo state, the dynamic shift was assessed only for the low-FRET RNA-bound population. Note the negative dynamic shifts which occur due to dye artifacts.

Lab#			1	2	3	4	7	8	11	14
BVA	MalE-1	Apo	0.0049	0.0007	0.0128	-0.0027	0.0167	0.0075	0.0042	0.0093
		Holo	0.0083	0.0022	0.0135	-0.001	0.0166	0.0073	0.0079	0.0098
	MalE-2	Apo	0.0083	0.0009	0.014	0.0023	0.0167	0.0103	0.0080	0.0105
		Holo	0.0070	0.0025	0.0143	0.0004	0.021	0.008	0.0062	0.0097
	MalE-3	Apo	0.0127	-0.0002	0.0081	-0.0043	0.0166	0.01	0.0107	0.0293
		Holo	0.0068	-0.0036	0.0087		-0.0043	0.0024	0.0077	0.0028
	MalE-4	Apo	0.0022		-0.0031					
		Holo	0.004		0.0070					
	MalE-5	Apo	0.0016		0.0031					
		Holo	0.003		0.0009					
	U2AF2	Apo	0.0283	0.0206	0.0217	0.026		0.0261	0.0363	0.0256
		Holo	0.0131	0.018	0.0259	0.0117		0.0178	0.0160	0.0171
E-τ	MalE-1	Apo	0.074	0.059	0.055	0.003		0.034		0.088
		Holo	0.069	0.057	0.062	0.004		0.031		0.068
	MalE-2	Apo	0.003	-0.007	0.018	0.01		0.029		0.042
		Holo	0.027	-0.002	0.025	0.015		0.011		0.05
	MalE-3	Apo	0.031	-0.013	0.026	0.004		0.044		0.043
		Holo	0.019	-0.019	0.037	0.035		0.027		0.046
	MalE-4	Apo	0.025	0.014	0.02					
		Holo	0.055	0.03	0.042					
	MalE-5	Apo	0.042	0.027	0.03					
		Holo	0.036	0.018	0.026					
	U2AF2	Apo	0.167	0.134	0.097	0.122		0.141		0.159
		Holo	0.022	0.008	0.019	0.035		0.012		0.029

Supplementary Table 8: The apparent dynamic shift values for different dye combinations of MalE and U2AF2 FRET variants as determined by three labs. The estimated distance fluctuation δR and FRET-averaged distance $R_{\langle E \rangle}$ is given for the measurements that passed the filtering procedure based on the estimated distance uncertainty from the residual anisotropies. Furthermore, the *ds* value for DNA rulers for each construct (LF: low-FRET, MF: medium-FRET, HF: high-FRET) and dye labels are provided and well as the average value over all measurement with the given error being the standard error of the mean, SEM.

			Lab#1			*Lab#2*			*Lab#3*		
Sample	**Dye combination**	**state**	***ds***	$R_{\langle E \rangle}$ [Å]	δR [Å]	***ds***	$R_{\langle E \rangle}$ [Å]	δR [Å]	***ds***	$R_{\langle E \rangle}$ [Å]	δR [Å]
MalE-1	Alexa546-Alexa647	apo	0.074			0.059			0.055		
MalE-1	Alexa546-Abb. STAR635P	apo	0.016			0.033					
MalE-1	Atto532-Atto643	apo	0.021	65.5	6.6						
MalE-1	Alexa546-Alexa647	holo	0.069			0.057			0.062		
MalE-1	Alexa546-Abb. STAR635P	holo	0.003			0.035					
MalE-1	Atto532-Atto643	holo	0.020	57.6	4.5						
MalE-2	Alexa546-Alexa647	apo	0.003	49.8	1.6	-0.007	49.3	0.2	0.018	51.2	3.9
MalE-2	Atto532-Atto643	apo	0.007	47.3	2.2	0.009	47.3	2.5			
MalE-2	Alexa488-Alexa647	apo	0.001	47.9	0.8	-0.002	47.9	0.2			
MalE-2	Alexa546-Alexa647	holo	0.027	55.1	4.8	-0.002	55.3	0.0	0.025	56.3	4.7
MalE-2	Atto532-Atto643	holo	0.020	54.4	4.0	0.018	54.4	3.8			
MalE-2	Alexa488-Alexa647	holo	0.010	54.1	3.3	0.009	54.1	3.1			
MalE-3	Alexa546-Alexa647	apo	0.031	44.2	5.7	-0.013	44.4	0.0	0.026	43.6	5.3
MalE-3	Atto532-Atto643	apo	0.003	39.3	1.7	0.002	39.3	1.4			
MalE-3	Alexa488-Alexa647	apo	- 0.037	40.4	0.0	0.013	40.4	2.6			
MalE-3	Alexa546-Alexa647	holo	0.019	44.2	4.5	-0.019	43.1	0.0	0.037	43.8	6.3
MalE-3	Atto532-Atto643	holo	-0.009	40.1	0.0	-0.007	40.1	0.0			
MalE-3	Alexa488-Alexa647	holo	-0.013	40.4	0.0	0.002	40.4	1.1			
MalE-4	Alexa546-Alexa647	apo	0.025			0.014			0.02		
MalE-4	Atto532-Atto643	apo	-0.005	66.9	0.0				0.025	66.9	7.8
MalE-4	Alexa488-Alexa647	apo	0.001	66.2	2.0	0.024	66.2	10.8			
MalE-4	Alexa546-Alexa647	holo	0.055			0.03			0.042		
MalE-4	Atto532-Atto643	holo	0.007	57.7	2.7				0.045	57.7	6.8
MalE-4	Alexa488-Alexa647	holo	-0.001	54.9	0.0	0.026	54.9	5.5			
MalE-5	Alexa546-Alexa647	apo	0.042			0.027			0.03		
MalE-5	Atto532-Atto643	apo	0.014	60.6	4.2	0.038	59.4	6.6	0.031	59.4	6.0
MalE-5	Alexa488-Alexa647	apo	0.032	60.4	8.4	0.022	60.4	7.1			
MalE-5	Alexa546-Alexa647	holo	0.036			0.018			0.026		
MalE-5	Atto532-Atto643	holo	0.032	49.1	4.7	0.027	49.2	4.3	0.04	49.2	5.3
MalE-5	Alexa488-Alexa647	holo	0.017	49.3	3.4	0.033	49.3	4.8			
U2AF	Alexa546-Alexa647	apo	0.166			0.129					
U2AF	Atto532-Atto643	apo	0.167			0.164					
U2AF	Alexa488-Alexa647	apo	0.168			0.128					
U2AF	Alexa546-Alexa647	holo	0.033			0.024					
U2AF	Atto532-Atto643	holo	0.022			0.025					
U2AF	Alexa488-Alexa647	holo	0.053			0.012					
DNA	Alexa488-Atto647N	LF				0.007	75.0	13.9			
DNA	Alexa488-Cy5	LF				0.015	73.6	13.9			
DNA	Alexa488-Alexa647	LF				0.012	72.5	11.9			
DNA	Alexa488-Atto647N	MF				-0.008	61.5	0.3			
DNA	Alexa488-Cy5	MF				0.011	61.2	5.4			
DNA	Alexa488-Alexa647	MF				-0.003	62.4	0.0			
DNA	Alexa488-Atto647N	HF				-0.012	51.7	0.0			
DNA	Alexa488-Cy5	HF				0.017	51.5	3.7			
DNA	Alexa488-Alexa647	HF				-0.016	52.4	0.0			
DNA	average over different dye combinations	-				0.0026 ± 0.0044					

Supplementary Table 9: The expected FRET efficiencies and dynamic shifts for the different experimental systems based on structural models from the PDB for the dye pair Alexa546-Alexa647. FRET efficiencies were predicted by AV simulations using the parameters given in Supplementary Table 10 based on the following PDB IDs: MalE apo - 1OMP, MalE holo - 1ANF, U2AF2 apo - 2YHO, U2AF2 holo - 2YH1. The expected dynamic shifts were calculated as described in Supplementary Note 6 using Eq. 6.1 for the E-τ plot and Eq. 6.23 for BVA.

	FRET efficiency, *E*		**Expected dynamic shift, ds**	
Sample	**apo**	**holo**	***E*-τ plot**	**BVA**
MalE-1	0.358	0.582	0.0169	0.0215
MalE-2	0.827	0.695	0.0131	0.0089
MalE-3	0.947	0.941	0.0002	0.0000
MalE-4	0.443	0.680	0.0231	0.0240
MalE-5	0.578	0.826	0.0382	0.0282
U2AF2	0.882	0.372	0.1426	0.0984

Supplementary Table 10: Parameters used for the AV and ACV calculations. The fraction of trapped dye is computed from the fundamental and residual anisotropy as $x_{\mathrm{trapped}} = r_{\infty,tr}/r_0$ (see Supplementary Table 5 and Supplementary Note 8, Eqn. 8.8).

	Dye species	
	Alexa546	**Alexa647**
Dye parameters		
L_{length} / Å	20.5	21.0
L_{width} / Å	4.5	4.5
$R_{\mathrm{dye},1}$ / Å	5.0	11.0
$R_{\mathrm{dye},2}$ / Å	4.5	4.7
$R_{\mathrm{dye},3}$ / Å	1.5	1.5
AV parameters		
Grid step / Å	0.9	
Allowed sphere radius / Å	1.0	
ACV parameters		
Contact volume thickness / Å	3.0	
Mutant	**Fraction of trapped dye (apo/holo)**	
29	0.87/0.83	0.33/0.36
352	0.72/0.73	0.68/0.74
87	0.54/0.62	0.44/0.44
186	0.38/0.42	0.22/0.22
134	0.90/0.90	0.45/0.43
34-205*	0.91/0.92	0.54/0.57
36-205*	0.88/0.89	0.72/0.78

*Anisotropy measurements of single mutants were not performed. Hence, trapped fractions were obtained from smFRET measurements and, as such, are an average of the two labeling positions due to the stochastic labeling of the protein.

Supplementary Table 11: Global fit of polarization-resolved fluorescence decays from ensemble measurements of single mutants of MalE. Rotational correlation times, ρ_j, with corresponding amplitudes, b_j, for Alexa647 and Alexa546 dyes from ensemble TCSPC measurements of single cysteine MalE mutants are reported (Supplementary Note 8, Eqn. 8.8). The rotational correlation times and fluorescence lifetimes as well their corresponding amplitudes were obtained from a global fit of the polarization-resolved fluorescence decays, $f_{VV}(t)$ and $f_{VH}(t)$, as described in Supplementary Note 8. The fit quality was judged by χ_r^2. The fit results indicate strong sticking interactions for positions K29C, A134C and S352C for the Alexa546 dye while, for the Alexa647 fluorophore, pronounced sticking interactions were found only at position S352C. From the fluorescence lifetime analysis, it can be seen that positions D87C and S352C are prone to quenching for the donor dye. Fitted polarization-resolved decays are combined into anisotropy decay and displayed in Fig. 5b for two representative mutation sites, namely S352C and K29C. Furthermore, obtained residual anisotropies are used in computation of Accessible Contact Volumes (Fig. 5c), which improved the agreement between modelled and measured distances (Fig. 5d).

Dye	Mutant	Condition	ρ_1 [ns]	b_1= $r_{\infty,tr}$	ρ_2 [ns]	b_2	ρ_3 [ns]	b_3	τ_1 [ns]	x_1	τ_2 [ns]	x_2	χ_r^2
Alexa546	K29C	apo	26	0.332	1.193	0.048			4.03	1.00			1.01
	D87C	apo	20	0.206	3.500	0.087	0.572	0.087	3.85	1.00			1.03
	A134C	apo	20	0.342			0.138	0.038	3.93	1.00			0.99
	A186C	apo	20	0.146	3.187	0.089	0.580	0.145	3.94	1.00			1.07
	S352C	apo	20	0.272	2.442	0.053	0.594	0.055	3.98	0.90	1.8	0.10	1.05
	K29C	holo	20	0.316			0.413	0.064	4.05	1.00			0.90
	D87C	holo	20	0.236	1.198	0.144			3.91	1.00			0.95
	A134C	holo	20	0.343			0.477	0.037	3.99	1.00			0.94
	A186C	holo	20	0.161			0.973	0.218	3.93	1.00			0.94
	S352C	holo	20	0.276	1.100	0.104			4.02	0.90	1.94	0.10	0.92
Alexa647	K29C	apo	>50	0.125			0.639	0.255	1.878	0.14	1.165	0.86	1.19
	D87C	apo	20	0.169			0.748	0.210	1.729	0.18	1.174	0.82	1.07
	A134C	apo	20	0.170			0.539	0.210	1.833	0.18	1.205	0.82	1.15
	A186C	apo	20	0.082	1.108	0.189	0.428	0.109	1.557	0.17	1.145	0.83	1.05
	S352C	apo	30	0.258			0.695	0.122	1.733	0.42	1.203	0.58	1.10
	K29C	holo	20	0.138			0.578	0.241	1.885	0.13	1.169	0.87	1.22
	D87C	holo	20	0.167			0.691	0.213	1.745	0.18	1.174	0.82	1.11
	A134C	holo	20	0.164			0.536	0.216	1.817	0.18	1.202	0.82	1.17
	A186C	holo	20	0.084			0.756	0.296	1.660	0.09	1.172	0.91	1.03
	S352C	holo	20	0.283			0.498	0.097	1.757	0.42	1.209	0.58	1.18

Supplementary Table 12: Steady-state and residual time-resolved anisotropy values of single mutant MalE samples labelled with Alexa546 and Alexa647, respectively. In addition to time-resolved anisotropies (Supplementary Table 11), for comparison we also give the steady-state anisotropy of the same fluorophores as free dyes and coupled to double-stranded DNA. Residual anisotropies obtained from time-resolved analysis (Supplementary Note 8, Eqn. 8.4 - 8.11) of donor and acceptor dye are used in computation of Accessible Contact Volumes (ACVs) (Fig. 5c), which improved the agreement between modelled and measured distances (Fig. 5d).

	$D_{ex}\|D_{em}$ channels (donor dye)		$A_{ex}\|A_{em}$ channels (acceptor dye)	
Sample	**Steady-state anisotropy, r_{ss}**	**Residual anisotropy, $r_{\infty,tr}$**	**Steady-state anisotropy, r_{ss}**	**Residual anisotropy, $r_{\infty,tr}$**
Free dye				
Alexa546	0.035±0.003	0.01±0.02	-	
Alexa647	-		0.120±0.007	0.02±0.02
DNA-Standards				
8 base-pairs	0.114±0.003		0.184±0.012	
33 base-pairs	0.134±0.002		0.159±0.011	
Donor-only strand	0.134±0.003		-	
Acceptor-only strand	-		0.172±0.010	
Protein single mutants				
K29C, apo	0.285±0.017	0.332	0.198±0.015	0.125
K29C, holo	0.280±0.017	0.316	0.199±0.018	0.138
D87C, apo	0.231±0.012	0.206	0.217±0.016	0.169
D87C, holo	0.225±0.005	0.236	0.229±0.017	0.167
A134C, apo	0.290±0.016	0.342	0.215±0.019	0.170
A134C, holo	0.281±0.007	0.343	0.216±0.003	0.164
A186C, apo	0.176±0.018	0.146	0.186±0.014	0.082
A186C, holo	0.161±0.010	0.161	0.186±0.016	0.084
S352C, apo	0.247±0.007	0.272	0.272±0.015	0.258
S352C holo	0.243±0.010	0.276	0.263±0.002	0.283

Supplementary Table 13: Combined residual anisotropies of additional dye combinations for double-labeled MalE and U2AF2 samples. Residual anisotropies of different donor and acceptor pairs used for labeling of MalE and U2AF2 variants. Residual anisotropies are computed using two approaches: from time-resolved analysis of the polarization-resolved fluorescence decays and from steady-state anisotropy measurements using a two-component Perrin equation (Supplementary Note 8, Eqn. 8.1 - 8.3). The results from the two approaches were averaged. Furthermore, using the residual anisotropies of donor and acceptor fluorophores, a combined residual anisotropy for a given FRET pair is obtained (Supplementary Note 8, Eqn. 8.12 - 8.15).

Sample	Dye combination	state	$\langle r_{\infty,D} \rangle_{tr,ss}$	$\langle r_{\infty,A} \rangle_{tr,ss}$	$\langle r_{c,\infty} \rangle_{tr,ss}$
MalE-1	Alexa546-Alexa647	apo	0.350 ± 0.014	0.286 ± 0.064	0.316 ± 0.036
MalE-1	Alexa546-Abb. STAR635P	apo	0.336 ± 0.013	0.228 ± 0.027	0.277 ± 0.017
MalE-1	Atto532-Atto643	apo	0.194 ± 0.049	0.179 ± 0.004	0.186 ± 0.023
MalE-1	Alexa546-Alexa647	holo	0.339 ± 0.014	0.272 ± 0.059	0.304 ± 0.034
MalE-1	Alexa546-Abb. STAR635P	holo	0.336 ± 0.006	0.222 ± 0.056	0.273 ± 0.035
MalE-1	Atto532-Atto643	holo	0.200 ± 0.029	0.189 ± 0.033	0.195 ± 0.022
MalE-2	Alexa546-Alexa647	apo	0.237 ± 0.021	0.176 ± 0.050	0.204 ± 0.031
MalE-2	Atto532-Atto643	apo	0.154 ± 0.063	0.141 ± 0.027	0.147 ± 0.033
MalE-2	Alexa488-Alexa647	apo	0.110 ± 0.037	0.190 ± 0.056	0.144 ± 0.032
MalE-2	Alexa546-Alexa647	holo	0.257 ± 0.002	0.170 ± 0.082	0.209 ± 0.050
MalE-2	Atto532-Atto643	holo	0.161 ± 0.065	0.152 ± 0.048	0.156 ± 0.040
MalE-2	Alexa488-Alexa647	holo	0.117 ± 0.023	0.195 ± 0.063	0.151 ± 0.029
MalE-3	Alexa546-Alexa647	apo	0.257 ± 0.003	0.168 ± 0.066	0.208 ± 0.041
MalE-3	Atto532-Atto643	apo	0.035 ± 0.019	0.139 ± 0.014	0.069 ± 0.019
MalE-3	Alexa488-Alexa647	apo	0.090 ± 0.004	0.211 ± 0.026	0.137 ± 0.009
MalE-3	Alexa546-Alexa647	holo	0.212 ± 0.036	0.155 ± 0.065	0.181 ± 0.041
MalE-3	Atto532-Atto643	holo	0.048 ± 0.006	0.139 ± 0.012	0.082 ± 0.006
MalE-3	Alexa488-Alexa647	holo	0.104 ± 0.017	0.206 ± 0.030	0.147 ± 0.016
MalE-4	Alexa546-Alexa647	apo	0.345 ± 0.002	0.199 ± 0.009	0.262 ± 0.006
MalE-4	Atto532-Atto643	apo	0.140 ± 0.020	0.136 ± 0.011	0.138 ± 0.011
MalE-4	Alexa488-Alexa647	apo	0.116 ± 0.020	0.167 ± 0.077	0.139 ± 0.034
MalE-4	Alexa546-Alexa647	holo	0.352 ± 0.006	0.235 ± 0.025	0.288 ± 0.016
MalE-4	Atto532-Atto643	holo	0.165 ± 0.001	0.130 ± 0.015	0.146 ± 0.008
MalE-4	Alexa488-Alexa647	holo	0.110 ± 0.008	0.174 ± 0.075	0.138 ± 0.030
MalE-5	Alexa546-Alexa647	apo	0.326 ± 0.012	0.266 ± 0.013	0.295 ± 0.009
MalE-5	Atto532-Atto643	apo	0.232 ± 0.009	0.191 ± 0.009	0.211 ± 0.006
MalE-5	Alexa488-Alexa647	apo	0.135 ± 0.012	0.190 ± 0.042	0.160 ± 0.019
MalE-5	Alexa546-Alexa647	holo	0.330 ± 0.011	0.268 ± 0.040	0.297 ± 0.023
MalE-5	Atto532-Atto643	holo	0.232 ± 0.001	0.188 ± 0.018	0.209 ± 0.010
MalE-5	Alexa488-Alexa647	holo	0.120 ± 0.003	0.171 ± 0.089	0.143 ± 0.037
U2AF	Alexa546-Alexa647	apo	0.305 ± 0.041	0.261 ± 0.044	0.282 ± 0.030
U2AF	Atto532-Atto643	apo	0.229 ± 0.034	0.121 ± 0.010	0.166 ± 0.014
U2AF	Alexa488-Alexa647	apo	0.168 ± 0.045	0.253 ± 0.064	0.207 ± 0.038
U2AF	Alexa546-Alexa647	holo	0.332 ± 0.054	0.238 ± 0.012	0.281 ± 0.024
U2AF	Atto532-Atto643	holo	0.241 ± 0.013	0.106 ± 0.016	0.160 ± 0.012
U2AF	Alexa488-Alexa647	holo	0.150 ± 0.042	0.248 ± 0.062	0.193 ± 0.036

Supplementary Table 14: Computed distance uncertainties for MalE and U2AF2 samples with different dye combination using a "Diffusion with traps" (DWT) model and a "Wobbling in cone" (WIC) model. Uncertainties in the calculated distance due to uncertainities in the orientation factor, κ^2, are computed using the residual anisotropies of donor and acceptor fluorophores for a given FRET pair. For filtering out dye combinatins with specific sticking interactions, a threshold of 10% from the DWT model in the distance uncertainty was used. Details on uncertainty calculations using the DWT model[32] can be found in Supplementary Note 9. For details of the WIC model, see ref [27].

Sample	Dye combination	state	$\Delta R_{app}(\kappa^2)$ [%] DWT model	$\Delta R_{app}(\kappa^2)$ [%] WIC model
MalE-1	Alexa546-Alexa647	apo	14.5	16.1
MalE-1	Alexa546-Abb. STAR635P	apo	11.0	13.3
MalE-1	Atto532-Atto643	apo	5.4	8.7
MalE-1	Alexa546-Alexa647	holo	14.8	14.9
MalE-1	Alexa546-Abb. STAR635P	holo	11.8	13.1
MalE-1	Atto532-Atto643	holo	6.2	9.0
MalE-2	Alexa546-Alexa647	apo	9.1	9.6
MalE-2	Atto532-Atto643	apo	5.2	7.2
MalE-2	Alexa488-Alexa647	apo	4.6	7.4
MalE-2	Alexa546-Alexa647	holo	8.2	9.6
MalE-2	Atto532-Atto643	holo	4.8	7.6
MalE-2	Alexa488-Alexa647	holo	4.5	7.5
MalE-3	Alexa546-Alexa647	apo	11.4	9.6
MalE-3	Atto532-Atto643	apo	3.3	5.2
MalE-3	Alexa488-Alexa647	apo	5.5	7.3
MalE-3	Alexa546-Alexa647	holo	9.4	8.5
MalE-3	Atto532-Atto643	holo	3.6	5.3
MalE-3	Alexa488-Alexa647	holo	5.8	7.4
MalE-4	Alexa546-Alexa647	apo	9.9	12.5
MalE-4	Atto532-Atto643	apo	3.7	7.0
MalE-4	Alexa488-Alexa647	apo	3.9	7.2
MalE-4	Alexa546-Alexa647	holo	13.3	14.1
MalE-4	Atto532-Atto643	holo	4.2	7.4
MalE-4	Alexa488-Alexa647	holo	4.0	7.0
MalE-5	Alexa546-Alexa647	apo	12.7	14.3
MalE-5	Atto532-Atto643	apo	7.1	9.7
MalE-5	Alexa488-Alexa647	apo	4.5	7.8
MalE-5	Alexa546-Alexa647	holo	16.2	14.4
MalE-5	Atto532-Atto643	holo	8.5	9.6
MalE-5	Alexa488-Alexa647	holo	4.3	7.2
U2AF	Alexa546-Alexa647	apo	14.3	13.4
U2AF	Atto532-Atto643	apo	5.9	8.2
U2AF	Alexa488-Alexa647	apo	8.1	9.6
U2AF	Alexa546-Alexa647	holo	11.7	13.5
U2AF	Atto532-Atto643	holo	5.3	8.1
U2AF	Alexa488-Alexa647	holo	5.9	9.1

Supplementary Table 15: Analysis of the U2AF2 dynamics. Results collected from 5 participating labs on the dynamics of U2AF2 labelled with the Atto532-Atto643 dye-pair. Timescales and rates were derived by dynamic PDA, filtered-FCS, and FRET-FCS as reported by the various groups. R_1 and R_2 are the distances corresponding to the compact and open states respectively for extracting the interconversion rates by dynamic PDA. Interconversion rates from closed to open and open to closed are denoted as k_{12} and k_{21} with their respective relaxation time τ_R. See Supplementary Note 17 for more details on the dynamics of U2AF2. The two extracted relaxation times from FRET-FCS and fFCS are denoted as t_{R1} and t_{R2}. See Supplementary Note 16 for more details on the filtered-FCS analysis. Average values of the rates and relaxation times, are given in the last row. * Only Lab#2 performed the FRET-FCS analysis and is not considered for average values for filtered-FCS.

Lab#	Method			Apo					Holo				
		R_1 (Å)	R_2 (Å)	t_{R1} (µs)	t_{R2} (µs)	k_{12} (ms^{-1})	k_{21} (ms^{-1})	τ_R (ms)	t_{R1} (µs)	t_{R2} (µs)	k_{12} (ms^{-1})	k_{21} (ms^{-1})	τ_R (ms)
1	Filtered-FCS			200	10								
	Dynamic PDA	38	59			0.43	0.07	2			0.52	0.14	1.42
2	Filtered-FCS			320	6				320	6			
	FRET-FCS*			32	2				321				
	Dynamic PDA	37	60			2.3	0.39	0.37			2.05	1.29	0.29
8	Dynamic PDA	38	56								0.62	0.18	1.25
11	Dynamic PDA	38	56								0.63	0.83	0.68
14	Filtered-FCS			370	12								
	Dynamic PDA	37	60			2.3	4.6	0.14	617	15	0.47	0.15	1.6
Average				296	9.3	1.67	1.68	0.83			0.85	0.51	1.05

Supplementary Table 16: The global dynamic photon distribution analysis (PDA) of apo and holo U2AF2 labeled with the Atto532-Atto643 dye pair. Parameters of the global dynamic PDA model are given, which is composed of a two-state dynamic system (apo ensemble) and two static states (holo and low-FRET) as described in the main text and Supplementary Note 17. The fit was performed globally over the apo and holo measurements using time windows of 0.5, 1.0, 1.5 and 2.0 ms, respectively. See Supplementary Figure 16 for an overview of the fits. The global reduced chi-square of the fit was 1.69. The analysis was performed using correction factors of: γ = 0.59 for the detection correction factor, a direct excitation correction of δ = 0.024, a crosstalk value of α = 0.02 and a Förster radius of 59 Å. Background count rates were 0.82 kHz and 0.28 kHz in the donor and FRET detection channels respectively. Fitting was performed using the *PDAFit* module of the PAM software package[11] .Errors were approximated from the covariance matrix given from the fit routine.

			Sample	
Population		**Parameter**	**apo**	**holo**
apo ensemble	compact state	$R_{compact}$ [Å]	37.4 ± 0.1	
		$\sigma_{compact}$ [Å]	2.1 ± 0.4	
	detached ensemble	$R_{detacted}$ [Å]	50.0 ± 0.1	
		$\sigma_{detached}$ [Å]	4.5 ± 0.2	
	kinetic rates	$k_{c->d}$ [ms^{-1}]	0.14 ± 0.11	
		$k_{d->c}$ [ms^{-1}]	0.013 ± 0.011	
	amplitude	A_{apo}	0.95 ± 0.02	0.35 ± 0.05
holo	open conformation	R_{open}[Å]	60.8 ± 0.1	
		σ_{open} [Å]	2.9 ± 0.2	
	amplitude	A_{holo}	0.03 ± 0.01	0.60 ± 0.03
low-FRET	apparent distance	R_{LF}[Å]	73.9 ± 4	
		σ_{LF} [Å]	7.2 ± 1.4	
	amplitude	A_{LF}	0.02 ± 0.01	0.05 ± 0.03

Supplementary Table 17: Correction factors obtained after reanalysis of the U2AF2 datasets from 7 different labs. Raw datasets were collected from 8 different labs for U2AF2 labeled with Atto532-Atto643. The reanalysis procedure is described in the Supplementary Note 3.

Lab#	***α***	***δ***	***γ***	***β***
1	0.02	0.06	0.73	0.72
2	0.04	0.20	1.26	0.22
3	0.03	0.06	0.98	0.77
4 *	0.03	0.08	0.98	1.15
8	0.03	0.05	1.15	0.80
11	0.03	0.07	0.65	1.21
14	0.03	0.05	0.92	1.00

* In these measurements, we observed a large contribution of the subpopulation with quenched acceptor lifetime (holo-high FRET) for double-labeled molecules. Hence, one population from apo and only one population from holo (holo-low FRET) conformation was used to calculate the γ globally.

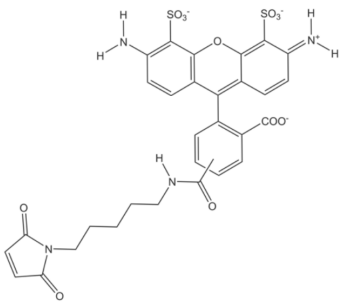
Supplementary Table 18: Dynamic photon distribution analysis (PDA) of apo and holo U2AF2 labeled with Alexa546-Alexa647 dye-pair. Parameters of the global PDA model are given, which is composed of a two-state dynamic system (apo ensemble) and two static states (holo and low-FRET) as described in the Supplementary Note 17. The fit was performed globally over the apo and holo measurements using time windows of 0.5, 1.0, and 1.5 ms respectively. See Supplementary Figure 19 for an overview of the fits. The analysis was performed using a detection correction factor of $\gamma = 0.32$, a direct excitation correction of $\delta = 0.10$, a crosstalk value of $\alpha = 0.036$ and a Förster radius of 65 Å. Background count rates were 0.918 kHz and 0.281 kHz in the donor and FRET detection channels respectively. Fitting was performed using the *PDAFit* module of the PAM software package[11].

			Sample	
Population		**Parameter**	**apo**	**holo**
apo ensemble	compact state	$R_{compact}$ [Å]	35.4	
		$\sigma_{compact}$ [Å]	3.5	
	detached ensemble	$R_{detacted}$ [Å]	50.3	
		$\sigma_{detached}$ [Å]	5.0	
	kinetic rates	$k_{c->d}$ [ms^{-1}]	0.16	
		$k_{d->c}$ [ms^{-1}]	0.014	
	amplitude	A_{apo}	0.97	0.36
holo	open conformation	R_{open}[Å]	-	62.0
		σ_{open} [Å]	-	2.9
	amplitude	A_{holo}	-	0.59
low-FRET	apparent distance	R_{LF}[Å]	72.5	
		σ_{LF} [Å]	5.0	
	amplitude	A_{LF}	0.03	0.05

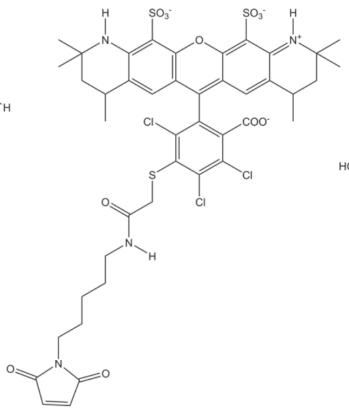
Supplementary Table 19: Detailed information on dye maleimides. Dye-maleimide conjugates were purchased from the indicated companies. Structural formulas of dyes are illustrated below, with the exception of Atto643 where the chemical structure of the dye is unknown. For dyes Atto532 and Abberior STAR635P, linkage length to the maleimide functional group is not indicated by the producer. Those dyes are therefore illustrated without the maleimide modification.

Fluorophore	Company	Catalog Number
Alexa488 C5 Maleimide	ThermoFisher Scientific	A10254
Alexa546 C5 Maleimide	ThermoFisher Scientific	A10258
Alexa647 C2 Maleimide	ThermoFisher Scientific	A20347
Atto532 Maleimide	ATTO-TEC	AD 532-45
Atto643 Maleimide	ATTO-TEC	AD 643
Abberior STAR635P Maleimide	Abberior	ST635P

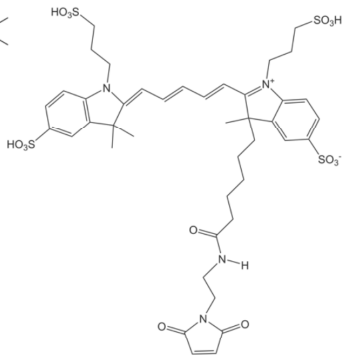
Alexa 488 C5 Maleimide



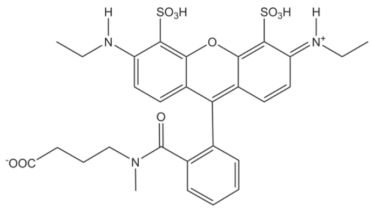
Alexa 546 C5 Maleimide



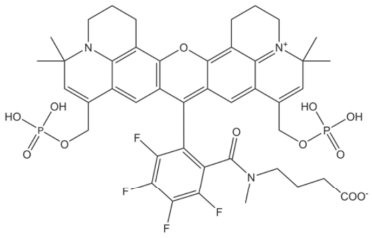
Alexa 647 C2 Maleimide



Atto 532



Abberior STAR635P



Supplementary References:

1. Peulen, T. O., Opanasyuk, O. & Seidel, C. A. M. Combining Graphical and Analytical Methods with Molecular Simulations to Analyze Time-Resolved FRET Measurements of Labeled Macromolecules Accurately. *J. Phys. Chem. B* **121**, 8211–8241 (2017).
2. Haas, E., Ephraim-Katchalski-Katzir & Steinberg, I. Z. Effect of the Orientation of Donor and Acceptor on the Probability of Energy Transfer Involving Electronic Transitions of Mixed Polarization. *Biochemistry* **17**, 5064–5070 (1978).
3. Kong, X., Nir, E., Hamadani, K. & Weiss, S. Photobleaching pathways in single-molecule FRET experiments. *J. Am. Chem. Soc.* **129**, 4643–4654 (2007).
4. Kudryavtsev, V. *et al.* Combining MFD and PIE for accurate single-pair Förster resonance energy transfer measurements. *ChemPhysChem* **13**, 1060–1078 (2012).
5. Eggeling, C. *et al.* Data registration and selective single-molecule analysis using multi-parameter fluorescence detection. *J. Biotechnol.* **86**, 163–180 (2001).
6. Lerner, E. *et al.* FRET-based dynamic structural biology: Challenges, perspectives and an appeal for open-science practices. *eLife* **10**, e60416 (2021).
7. Lee, N. K. *et al.* Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. *Biophys. J.* **88**, 2939–2953 (2005).
8. Metskas, L. A. & Rhoades, E. Single-molecule FRET of intrinsically disordered proteins. *Annu. Rev. Phys. Chem.* **71**, 391–414 (2020).
9. Ferreon, A. C. M., Gambin, Y., Lemke, E. A. & Deniz, A. A. Interplay of α-synuclein binding and conformational switching probed by single-molecule fluorescence. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 5645–5650 (2009).
10. Vandenberk, N., Barth, A., Borrenberghs, D., Hofkens, J. & Hendrix, J. Evaluation of Blue and Far-Red Dye Pairs in Single-Molecule FRET Experiments - SUP. *J. Phys. Chem. B* **32**, acs.jpcb.8b00108 (2018).
11. Schrimpf, W., Barth, A., Hendrix, J. & Lamb, D. C. PAM: A Framework for Integrated Analysis of Imaging, Single-Molecule, and Ensemble Fluorescence Data. *Biophys. J.* **114**, 1518–1528 (2018).
12. Tomov, T. E. *et al.* Disentangling subpopulations in single-molecule FRET and ALEX experiments with photon distribution analysis. *Biophys. J.* **102**, 1163–1173 (2012).
13. Hellenkamp, B. *et al.* Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. *Nat. Methods* **15**, 669–676 (2018).
14. Ingargiola, A., Lerner, E., Chung, S. Y., Weiss, S. & Michalet, X. FRETBursts: An open source toolkit for analysis of freely-diffusing Single-molecule FRET. *PLoS One* **11**, 39198 (2016).
15. Fries, J. R., Brand, L., Eggeling, C., Köllner, M. & Seidel, C. A. M. Quantitative identification of different single molecules by selective time-resolved confocal fluorescence spectroscopy. *J. Phys. Chem. A* **102**, 6601–6613 (1998).

16. Torella, J. P., Holden, S. J., Santoso, Y., Hohlbein, J. & Kapanidis, A. N. Identifying molecular dynamics in single-molecule fret experiments with burst variance analysis. *Biophys. J.* **100**, 1568–1577 (2011).

17. Kalinin, S., Valeri, A., Antonik, M., Felekyan, S. & Seidel, C. A. M. Detection of Structural Dynamics by FRET: A Photon Distribution and Fluorescence Lifetime Analysis of Systems with Multiple States. *J. Phys. Chem. B* **114**, 7983–7995 (2010).

18. Barth, A. *et al.* Unraveling multi-state molecular dynamics in single-molecule FRET experiments. I. Theory of FRET-lines. *J. Chem. Phys.* **156**, 141501 (2022).

19. Förster, T. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann. Phys.* **437**, 55–75 (1948).

20. Rurack, K. Fluorescence Quantum Yields: Methods of Determination and Standards. in *Standardization and Quality Assurance in Fluorescence Measurements I* 101–145 (2008). doi:10.1007/4243_2008_019

21. Wiederschain, G. Y. The Molecular Probes handbook. A guide to fluorescent probes and labeling technologies. *Biochem.* **76**, 1276–1276 (2011).

22. Voith von Voithenberg, L. & Lamb, D. C. Single Pair Förster Resonance Energy Transfer: A Versatile Tool To Investigate Protein Conformational Dynamics. *BioEssays* **40**, 1–14 (2018).

23. Spectra, E. Alexa Fluor ® Dyes — Across the Spectrum. 1–3 (2014).

24. Schaffer, J. *et al.* Identification of Single Molecules in Aqueous Solution by Time-Resolved Fluorescence Anisotropy. *J. Phys. Chem. A* **103**, 331–336 (1999).

25. Perrin, F. Polarisation de la lumière de fluorescence. Vie moyenne des molécules dans l'etat excité. *J. Phys. le Radium* **7**, 390–401 (1926).

26. Sisamakis, E., Valeri, A., Kalinin, S., Rothwell, P. J. & Seidel, C. A. M. Accurate Single-Molecule FRET Studies Using Multiparameter Fluorescence Detection. *Methods Enzymol.* **475**, 455–514 (2010).

27. Sindbert, S. *et al.* Accurate distance determination of nucleic acids via Förster resonance energy transfer: Implications of dye Linker length and rigidity. *J. Am. Chem. Soc.* **133**, 2463–2480 (2011).

28. Koshioka, M., Sasaki, K. & Masuhara, H. Time-Dependent Fluorescence Depolarization Analysis in Three-Dimensional Microspectroscopy. *Appl. Spectrosc.* **49**, 224–228 (1995).

29. Schröder, G. F., Alexiev, U. & Grubmüller, H. Simulation of fluorescence anisotropy experiments: Probing protein dynamics. *Biophys. J.* **89**, 3757–3770 (2005).

30. Dale, R. E. & Eisinger, J. Intramolecular distances determined by energy transfer. Dependence on orientational freedom of donor and acceptor. *Biopolymers* **13**, 1573–1605 (1974).

31. Ivanov, V., Li, M. & Mizuuchi, K. Impact of emission anisotropy on fluorescence spectroscopy and FRET distance measurements. *Biophys. J.* **97**, 922–929 (2009).

32. Kalinin, S., Fulle, S., Hanke, C.A., Peulen, T. O., Sindbert, S., Felekyan, S., Kühnemuth, R., Gohlke, H., Seidel, C. A. M. Diffusion with traps: experiment,

simulation, and theory to describe the dynamics of flexibly linked fluorophores in biomolecular FRET. *Prep.*

33. Case, D. A. *et al. AMBER 2018, University of California. University of California, San Francisco* (2018).

34. Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).

35. Horn, H. W. *et al.* Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *J. Chem. Phys.* **120**, 9665–9678 (2004).

36. Hopkins, C. W., Le Grand, S., Walker, R. C. & Roitberg, A. E. Long-time-step molecular dynamics through hydrogen mass repartitioning. *J. Chem. Theory Comput.* **11**, 1864–1874 (2015).

37. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).

38. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).

39. McGibbon, R. T. *et al.* MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. *Biophys. J.* **109**, 1528–1532 (2015).

40. Gopich, I. V. & Szabo, A. Single-molecule FRET with diffusion and conformational dynamics. *J. Phys. Chem. B* **111**, 12925–12932 (2007).

41. Zander, C. *et al.* Detection and characterization of single molecules in aqueous solution. *Appl. Phys. B Lasers Opt.* **63**, 517–523 (1996).

42. Hall, P. & Selinger, B. *Better estimates of exponential decay parameters. Journal of Physical Chemistry* **85**, (1981).

43. Livesey, A. K. & Skilling, J. Maximum entropy theory. *Acta Crystallogr. Sect. A* **41**, 113–122 (1985).

44. Brochon, J. C. Maximum entropy method of data analysis in time-resolved spectroscopy. *Methods Enzymol.* **240**, 262–311 (1994).

45. Skilling, J. & Bryan, R. K. Maximum entropy image reconstruction: general algorithm. *Mon. Not. R. Astron. Soc.* **211**, 111–124 (1984).

46. Vinogradov, S. A. & Wilson, D. F. Recursive maximum entropy algorithm and its application to the luminescence lifetime distribution recovery. *Appl. Spectrosc.* **54**, 849–855 (2000).

47. Huang, J. R. *et al.* Transient electrostatic interactions dominate the conformational equilibrium sampled by multidomain splicing factor U2AF65: A combined NMR and SAXS study. *J. Am. Chem. Soc.* **136**, 7068–7076 (2014).

48. Kalinin, S. *et al.* A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nat. Methods* **9**, 1218–1225 (2012).

49. Bowman, A. W. & Azzalini, A. Computational aspects of nonparametric smoothing with illustrations from the sm library. *Comput. Stat. Data Anal.* **42**, 545–560 (2003).

50. 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).

51. Barth, A. *et al.* Dynamic interactions of type I cohesin modules fine-tune the structure of the cellulosome of Clostridium thermocellum. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E11274–E11283 (2018).

52. Von Voithenberg, L. V. *et al.* Recognition of the 3′ splice site RNA by the U2AF heterodimer involves a dynamic population shift. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E7169–E7175 (2016).

53. Naudi-Fabra, S., Tengo, M., Jensen, R., Blackledge, M. & Milles, S. Quantitative Description of Intrinsically Disordered Proteins Using Single-Molecule FRET, NMR, and SAXS. *J. Am. Chem. Soc* **143**, 50 (2021).

54. Margittai, M. *et al.* Single-molecule fluorescence resonance energy transfer reveals a dynamic equilibrium between closed and open conformations of syntaxin 1. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 15516–15521 (2003).

55. Torres, T. & Levitus, M. Measuring conformational dynamics: A new FCS-FRET approach. *J. Phys. Chem. B* **111**, 7392–7400 (2007).

56. Felekyan, S., Sanabria, H., Kalinin, S., Kühnemuth, R. & Seidel, C. A. M. Analyzing Förster resonance energy transfer with fluctuation algorithms. *Methods Enzymol.* **519**, 39–85 (2013).

57. Antonik, M., Felekyan, S., Gaiduk, A. & Seidel, C. A. M. Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis. *J. Phys. Chem. B* **110**, 6970–6978 (2006).

58. Kalinin, S., Felekyan, S., Antonik, M. & Seidel, C. A. M. Probability distribution analysis of single-molecule fluorescence anisotropy and resonance energy transfer. *J. Phys. Chem. B* **111**, 10253–10262 (2007).

59. MacKereth, C. D. *et al.* Multi-domain conformational selection underlies pre-mRNA splicing regulation by U2AF. *Nature* **475**, 408–413 (2011).

60. Sánchez-Rico, C., Voith von Voithenberg, L., Warner, L., Lamb, D. C. & Sattler, M. Effects of Fluorophore Attachment on Protein Conformation and Dynamics Studied by spFRET and NMR Spectroscopy. *Chem. - A Eur. J.* **23**, 14267–14277 (2017).

61. Neubauer, H. *et al.* Orientational and dynamical heterogeneity of rhodamine 6G terminally attached to a DNA helix revealed by NMR and single-molecule fluorescence spectroscopy. *J. Am. Chem. Soc.* **129**, 12746–12755 (2007).

62. Laurence, T. A. *et al.* Correlation spectroscopy of minor fluorescent species: Signal purification and distribution analysis. *Biophys. J.* **92**, 2184–2198 (2007).

63. Ivanchenko, S. & Lamb, D. C. Fluorescence Correlation Spectroscopy: Principles and Developments. in *Supramolecular Structure and Function 10* 1–30 (2011). doi:10.1007/978-94-007-0893-8_1

64. Henzler-Wildman, K. A. *et al.* Intrinsic motions along an enzymatic reaction trajectory. *Nature* **450**, 838–844 (2007).

65. Doose, S., Heilemann, M., Michalet, X., Weiss, S. & Kapanidis, A. N. Periodic acceptor excitation spectroscopy of single molecules. *Eur. Biophys. J.* **36**, 669–674 (2007).

66. Barth, A., Voith Von Voithenberg, L. & Lamb, D. C. Quantitative Single-Molecule Three-Color Förster Resonance Energy Transfer by Photon Distribution Analysis. *J. Phys. Chem. B* **123**, 6901–6916 (2019).

67. Gouridis, G. *et al.* Conformational dynamics in substrate-binding domains influences transport in the ABC importer GlnPQ. *Nat. Struct. Mol. Biol.* **22**, 57–64 (2015).

68. De Boer, M. *et al.* Conformational and dynamic plasticity in substrate-binding proteins underlies selective transport in ABC importers. *eLife* **8**, e44652 (2019).

69. de Boer, M., Gouridis, G., Muthahari, Y. A. & Cordes, T. Single-Molecule Observation of Ligand Binding and Conformational Changes in FeuA. *Biophys. J.* **117**, 1642–1654 (2019).

70. Wolf, S. *et al.* Hierarchical dynamics in allostery following ATP hydrolysis monitored by single molecule FRET measurements and MD simulations. *Chem. Sci.* **12**, 3350–3359 (2021).

71. Ambrose, B. *et al.* The smfBox is an open-source platform for single-molecule FRET. *Nat. Commun.* **11**, 5641 (2020).

72. Abdelhamid, M. A. S., Rhind-Tutt, A. V., Ambrose, B. & Craggs, T. D. Making Precise and Accurate Single-Molecule FRET Measurements using the Open-Source smfBox. *J. Vis. Exp.* e62378 (2021). doi:10.3791/62378

73. Tan, P. S. & Lemke, E. A. Probing Differential Binding Mechanisms of Phenylalanine-Glycine-Rich Nucleoporins by Single-Molecule FRET. in *Methods in Enzymology* **611**, 327–346 (2018).

74. Müller, B. K., Zaychikov, E., Bräuchle, C. & Lamb, D. C. Pulsed interleaved excitation. *Biophys. J.* **89**, 3508–3522 (2005).

75. Robb, N. C. *et al.* The transcription bubble of the RNA polymerase-promoter open complex exhibits conformational heterogeneity and millisecond-scale dynamics: Implications for transcription start-site selection. *J. Mol. Biol.* **425**, 875–885 (2013).

76. Santoso, Y. *et al.* Conformational transitions in DNA polymerase I revealed by single-molecule FRET. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 715–720 (2010).

77. Santoso, Y., Torella, J. P. & Kapanidis, A. N. Characterizing single-molecule FRET dynamics with probability distribution analysis. *ChemPhysChem* **11**, 2209–2219 (2010).

78. Robb, N. C. *et al.* Single-molecule FRET reveals the pre-initiation and initiation conformations of influenza virus promoter RNA. *Nucleic Acids Res.* **44**, 10304–10315 (2016).

79. Tomescu, A. I., Robb, N. C., Hengrung, N., Fodor, E. & Kapanidis, A. N. Single-molecule FRET reveals a corkscrew RNA structure for the polymerase-bound influenza virus promoter. *Proc. Natl. Acad. Sci.* **111**, E3335-3342 (2014).

80. Kapanidis, A. N. *et al.* Fluorescence-aided molecule sorting: Analysis of structure and

interactions by alternating-laser excitation of single molecules. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 8936–8941 (2004).

81. Kapanidis, A. N. *et al.* Alternating-laser excitation of single molecules. *Acc. Chem. Res.* **38**, 523–533 (2005).

82. Krainer, G. *et al.* Ultrafast Protein Folding in Membrane-Mimetic Environments. *J. Mol. Biol.* **430**, 554–564 (2018).

83. Hartmann, A., Krainer, G., Keller, S. & Schlierf, M. Quantification of Millisecond Protein-Folding Dynamics in Membrane-Mimetic Environments by Single-Molecule Förster Resonance Energy Transfer Spectroscopy. *Anal. Chem.* **87**, 11224–11232 (2015).

84. Hartmann, A. Observing Biomolecular Dynamics from Nanoseconds to Hours with Single-Molecule Fluorescence Spectroscopy. 139 (2017).

85. Kramm, K. *et al.* DNA origami-based single-molecule force spectroscopy elucidates RNA Polymerase III pre-initiation complex stability. *Nat. Commun.* **11**, 2828 (2020).

86. Kahra, D. *et al.* Conformational plasticity and dynamics in the generic protein folding catalyst SlyD unraveled by single-molecule FRET. *J. Mol. Biol.* **411**, 781–790 (2011).

87. Olofsson, L. & Margeat, E. Pulsed interleaved excitation fluorescence spectroscopy with a supercontinuum source. *Opt. Express* **21**, 3370 (2013).

88. Fuertes, G. *et al.* Decoupling of size and shape fluctuations in heteropolymeric sequences reconciles discrepancies in SAXS vs. FRET measurements. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E6342–E6351 (2017).

89. Harris, P. D. *et al.* Multi-parameter photon-by-photon hidden Markov modeling. *Nat. Commun.* **13**, 1000 (2022).

90. Ingargiola, A., Laurence, T., Boutelle, R., Weiss, S. & Michalet, X. Photon-HDF5: An Open File Format for Timestamp-Based Single-Molecule Fluorescence Experiments. *Biophys. J.* **110**, 26–33 (2016).

91. Nir, E. *et al.* Shot-noise limited single-molecule FRET histograms: Comparison between theory and experiments. *J. Phys. Chem. B* **110**, 22103–22124 (2006).

92. Buning, R., Kropff, W., Martens, K. & van Noort, J. spFRET reveals changes in nucleosome breathing by neighboring nucleosomes. *J. Phys. Condens. Matter* **27**, 064103 (2015).

93. Zhou, R., Schlierf, M. & Ha, T. Force-Fluorescence Spectroscopy at the Single-Molecule Level. *Methods Enzymol.* **475**, 405–426 (2010).

94. Nikaido, H. Maltose transport system of Escherichia coli: An ABC-type transporter. *FEBS Lett.* **346**, 55–58 (1994).

95. McKinney, S. A., Joo, C. & Ha, T. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys. J.* **91**, 1941–1951 (2006).

96. Jerabek-Willemsen, M. *et al.* MicroScale Thermophoresis: Interaction analysis and beyond. *J. Mol. Struct.* **1077**, 101–113 (2014).

97. Widengren, J. & Schwïlle, P. Characterization of photoinduced isomerization and

back-isomerization of the cyanine dye cy5 by fluorescence correlation spectroscopy. *J. Phys. Chem. A* **104**, 6416–6428 (2000).

98. Petrášek, Z. & Schwille, P. Precise measurement of diffusion coefficients using scanning fluorescence correlation spectroscopy. *Biophys. J.* **94**, 1437–1448 (2008).

Supplement B. Resolution of Maximum Entropy Method-derived posterior conformational ensembles of a flexible system probed by FRET and molecular dynamics simulations

JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

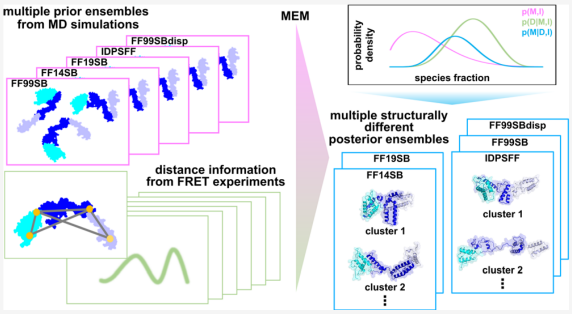
Resolution of Maximum Entropy Method-Derived Posterior Conformational Ensembles of a Flexible System Probed by FRET and Molecular Dynamics Simulations

Jonas Dittrich,[▽] Milana Popara,[▽] Jakub Kubiak, Mykola Dimura, Bastian Schepers, Neha Verma, Birte Schmitz, Peter Dollinger, Filip Kovacic, Karl-Erich Jaeger,* Claus A. M. Seidel,* Thomas-Otavio Peulen,* and Holger Gohlke*

Cite This: *J. Chem. Theory Comput.* 2023, 19, 2389−2409 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Maximum entropy methods (MEMs) determine posterior distributions by combining experimental data with prior information. MEMs are frequently used to reconstruct conformational ensembles of molecular systems for experimental information and initial molecular ensembles. We performed time-resolved Förster resonance energy transfer (FRET) experiments to probe the interdye distance distributions of the lipase-specific foldase Lif in the *apo* state, which likely has highly flexible, disordered, and/or ordered structural elements. Distance distributions estimated from ensembles of molecular dynamics (MD) simulations serve as prior information, and FRET experiments, analyzed within a Bayesian framework to recover distance distributions, are used for optimization. We tested priors obtained by MD with different force fields (FFs) tailored to ordered (FF99SB, FF14SB, and FF19SB) and disordered proteins (IDPSFF and FF99SBdisp). We obtained five substantially different posterior ensembles. As in our FRET experiments the noise is characterized by photon counting statistics, for a validated dye model, MEM can quantify consistencies between experiment and prior or posterior ensembles. However, posterior populations of conformations are uncorrelated to structural similarities for individual structures selected from different prior ensembles. Therefore, we assessed MEM simulating varying priors in synthetic experiments with known target ensembles. We found that (i) the prior and experimental information must be carefully balanced for optimal posterior ensembles to minimize perturbations of populations by overfitting and (ii) only ensemble-integrated quantities like inter-residue distance distributions or density maps can be reliably obtained but not ensembles of atomistic structures. This is because MEM optimizes ensembles but not individual structures. This result for a highly flexible system suggests that structurally varying priors calculated from varying prior ensembles, e.g., generated with different FFs, may serve as an *ad hoc* estimate for MEM reconstruction robustness.



1. INTRODUCTION

To understand protein function, proteins are best described by energy landscapes, molecular ensembles, and kinetic pathways,[1] which challenges structural biology.[2] Classical experimental techniques such as X-ray crystallography, cryo-EM, or NMR can resolve fascinating snapshots of single atomistic structures. Many proteins require flexibility to function.[1,3] If the number of conformers is small and they are long-lived, the classical techniques can still resolve conformers at high resolution by varying experimental conditions.[4−7] However, multidomain proteins with unstructured segments and intrinsically disordered proteins (IDPs) are particularly challenging for structural biology due to a large number of possible conformations.[8−12] Structural ensembles were proposed for molecular systems of varying sizes and complexities, starting from small folded proteins that can be studied accurately by NMR techniques, such as ubiquitin, for which structures and an associated dynamic model, a so-called kinetic ensemble, were reported.[13,14] Multidomain proteins with flexible linkers, such as U2AF65,[15] calmodulin,[16] or TIA-1,[17] are more challenging, as the size and associated number of degrees of freedom are substantially larger. Proteins with disorder regions, such as ESCRT complexes,[18] are even more complex, and the ultimate challenge in terms of the number of degrees of freedom is the IDPs.[19,20]

Received: November 1, 2022
Published: April 6, 2023

For large ensembles, no technique alone can resolve molecular ensembles in atomistic detail, as most biophysical experiments report on observables that are averages over ensembles of conformations. To overcome the limitations of individual techniques, integrative molecular modeling combines experimental information such as NMR,[15] small angle X-ray scattering (SAXS),[21] electron paramagnetic resonance (EPR),[22] Förster resonance energy transfer (FRET),[23−27] and cross-linking[28] with computational information (e.g., from molecular simulations using force fields (FFs)).[25−27,29]

Bayes theorem provides a formal framework to integrate simulated ensemble and experimental data. Combined with the maximum entropy principle, it is a suitable refinement approach for flexible systems, where no assumption about the number of states is made, as opposed to maximum-parsimony methods, where the minimum number of structures that describe experimental data is selected.[8,30] Bayes theorem expresses information in the form of probabilities, p. The theorem combines information on a model, M, experimental data, D, and prior information, I (eq 1)[31]

$$p(M|D, I) = \frac{p(D|M, I)p(M|I)}{p(D|I)} \quad (1)$$

where $p(M|D, I)$ is the posterior model density, e.g., the probability density of a conformational ensemble conditioned on D and I, $p(D|M, I)$ is the data likelihood, $p(M|I)$ is the prior model density, and $p(D|I) = Z$ is the model evidence or marginal likelihood (in analogy to statistical mechanics, $p(D|I)$ is sometimes called partition function, Z). Conformational ensembles M can be determined by restrained simulations,[32−36] where D is incorporated into the structural modeling framework.[27,37] Alternatively, $p(M|I)$ is sampled independently of $p(D|M, I)$, and $p(M|D, I)$ is computed using the independently sampled prior and likelihood. The latter approach is easy to implement, as the sampling is independent of evaluating the likelihood function. However, few sampled models might satisfy D and I either due to insufficient sampling or inaccuracies of I used for sampling (FF parameters).

Maximum entropy methods (MEMs) are statistical approaches based on Bayes' theorem that combine I, D, and M.[38] MEMs determine the maximum *a posteriori* (MAP) distribution, Q, for $p(D|M, I)$ and an entropy prior, $p(M|I)$. When MEMs are applied to recover ensembles of states for D, a set of predetermined states with corresponding population fractions is the prior information I, and the population of the states are parameters of M.[39−41] States are most commonly defined by structural similarities, and MEMs are used to optimize populations of conformers. When optimizing M, the strict convexity of the entropy prior asserts that even for underdetermined systems unique regular solutions are found.[42] However, $p(M|D, I)$ depends on I, the state distribution of the initially simulated ensemble, and D, the experimental data. Hence, both I and D impact the ensembles recovered by MEMs.

Here, we probe the robustness of MEM, i.e., the precision and accuracy, for recovering the structural diversity of the ensemble and resolving individual structures in the ensemble with respect to I, considering experimental data and its corresponding precision. For that, we determine posterior ensembles for the folding domain (FOD) of *Pseudomonas aeruginosa* lipase-specific foldase (Lif)[43,44] (Figure 1). Lif is a steric chaperone with a three-domain organization common

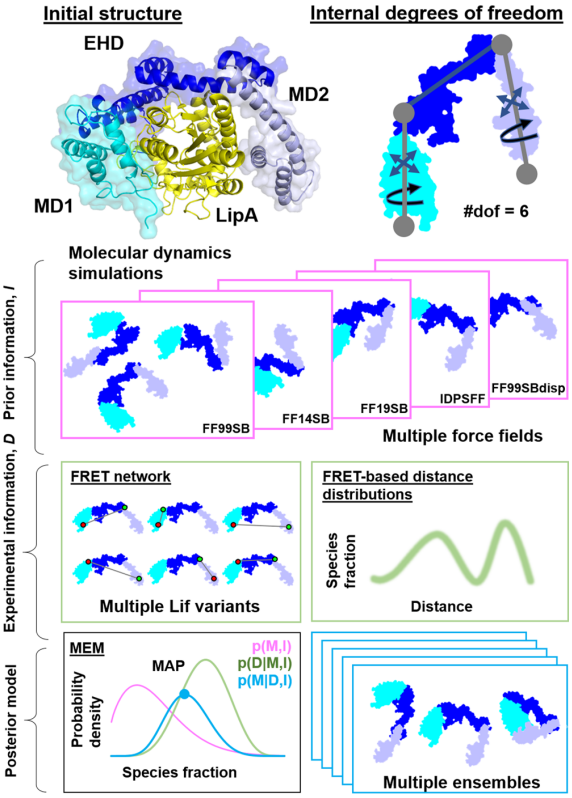


Figure 1. Maximum entropy-based ensemble refinement of lipase-specific foldase (Lif). Top left: the quaternary structure of the *P. aeruginosa* Lif:LipA complex homology model is shown, with the foldase Lif minidomains MD1 (residues 65−146, cyan), the extended helical domain EHD (residues 147−265, blue), and MD2 (residues 266−340, light blue). Lif forms a unique headphone-like fold around the lipase LipA (yellow). Top right: simplified Lif model. Three domains are assumed to act as three rigid bodies connected by two hinges (between EHD and MDs, allowing three degrees of freedom each for the MDs), yielding a model with six internal degrees of freedom. The all-atom structural information is fed into molecular dynamics simulations with five different FFs. The obtained conformational ensembles serve as a basis to calculate the interdye distance distribution in MD, i.e., the prior information I (magenta frames). FRET data acquired for the distance network provides information on interdye distance distributions D (green frames). The prior information I is combined with the FRET data D using the maximum entropy method (MEM, black frame). MEM determines the maximum a posteriori (MAP) distribution for the species fractions to produce posterior structure weights Q that represent molecular ensembles (blue frames).

with other steric chaperones.[43,45,46] Its folding domain is anchored to the inner bacterial membrane by a conserved transmembrane domain and an alanine/proline-rich flexible linker.[47] Lif activates the prefolded lipase LipA with a mechanism independent of ATP[48] by binding with the folding domain (FOD),[47] which can be divided into three subdomains: minidomains 1 and 2 (MD1, MD2), connected by the extended helical domain (EHD).[43,49] As such, Lif follows the typical architecture of multidomain proteins, where different domains of a protein are connected by linkers with different amounts of softness. The α-helical arrangement of the

https://doi.org/10.1021/acs.jctc.2c01090
J. Chem. Theory Comput. 2023, 19, 2389−2409

EHD suggests it to be a "rigid linker" that keeps MD1 and MD2 as a molecular spacer at a certain distance.

We, therefore, geometrically approximated Lif as a 3-body object for which a small set of distance measurements should potentially resolve the domain arrangement (Figure 1) while preserving biological activity (Table S1). We used time-resolved FRET experiments and molecular dynamics (MD) simulations to acquire structural data on Lif. FRET is the distance-dependent energy migration from an excited donor fluorophore to an acceptor fluorophore;[50] time-resolved FRET experiments inform on conformational distributions with minimal conformational averaging, as snapshots of the conformers are obtained on the nanosecond time scale. Our MEM scoring uses the known single photon counting statistics (Poisson). Making use of defined data and data noise, our FRET-based MEM has a well-defined minimum that is convenient and robust for validating the consistency of the FF-dependent prior and posterior ensembles to experiments.

For integrative FRET modeling by MEM, we established the following workflow: we developed a Bayesian tool (ucFRET) that transforms time-resolved FRET data into interdye distance distributions as opposed to simple ensemble averages. For resolving conformational ensembles, we acquired data on a set of FRET experiments with a network of FRET pairs. These experiments served as experimental data, D. As input information, we used distinct conformational ensembles obtained from MD simulations with different force fields optimized for folded and unfolded proteins. To compute experimental observables as prior I for combining experimental data and computational models, we employed a forward model based on accessible volume (AV) simulations.[25,37] We optimized our prior ensembles against the FRET data and found that our priors describe the FRET data to a different extent. By varying the contribution of experimental and prior information, we searched for optimally balanced solutions (L-curve analysis).

Next, we assessed the performance of our MEM approach. We tested the effect of prioritizing experiments over the prior information to evaluate how uniquely conformations are resolved within a large ensemble at atomistic resolution and whether ensembles from distinct force fields yield similar conformers. This objective is a great challenge in general and especially for FRET because FRET mainly informs on long-range features.[37] To optimize our analysis and assess the robustness of MEM ensemble reconstructions, we performed synthetic experiments on defined conformational ensembles (target ensemble Y). We found the optimal balance for prior and experimental information that minimized the discrepancy to the target ensemble to be close to the corner of the L-curve. Using ensemble representations of varying details (atomistic representations, inter-residue distograms, and three-dimensional (3D) densities), we assessed the reliability of MEM-derived ensembles and consequences for the achieved resolution of posterior models. We found that distograms and 3D density maps are robustly recovered representations that are more appropriate to capture and identify relevant features of large and diverse posterior ensembles than individual structures. Finally, for our experimental FRET data, we find that three out of five posterior ensembles have a headphone shape as identified by the density map representation. According to the L-curve criterion, ensembles with a headphone or expanded shape are both feasible for *apo*Lif in solution.

Our results stress the need for robust uncertainty estimates of posterior ensembles. In the absence of uncertainty estimates, the distinct obtained posterior ensembles highlight the importance of using structurally varying priors to probe the robustness of MEM reconstructions.

2. MATERIALS AND METHODS

2.1. Maximum Entropy Method. MEMs used for integrative modeling are based on MaxEnt[51] methods developed for image restoration.[52] MEMs maximize the posterior model density according to eq 2

$$\arg\max_M p(M|D, I) = \arg\max_M p(D|M, I)p(M|I) \tag{2}$$

where $\arg\max_M p(M|D, I) = Q$ is the MAP distribution, i.e., the posterior model density that is maximal for M, D, and I. A set of J states defines I. In structural integrative modeling problems, a state corresponds to a conformation C_j with a prior weight $w_j^{(0)}$. Hence, $I = \{w_j^{(0)}, C_j\}_{j=1}^{j=J}$. The set of weights defines the model, $M = \{w_j\}_{j=1}^{j=J}$. Weights belong to the $J-1$ simplex, i.e., $\sum_j w_j = 1$ and $w_j \geq 0$ for $j \in \{1, \cdots, J\}$. In MEM, $p(M|I)$ is an entropy prior (eq 3)

$$p(M|I) = \prod_{j=1}^{J} \exp(-\theta \cdot w_j \ln(w_j/w_j^{(0)})) = \exp(\theta S) \tag{3}$$

where $S = -\sum_{j=1}^{J} w_j \ln(w_j/w_j^{(0)})$ and θ is a temperature-like factor that controls the weight of the prior. $p(D|M, I) = \exp(-H)$ is the data likelihood, where H depicts the agreement of the model with D. Computing the data likelihood requires a forward model and a noise model. The forward model predicts D for M, assuming no experimental noise. The noise model quantifies the distribution of the differences between D and the forward model,[53] which originates from random experimental error, systematic experimental error, and error in the forward model; noise models generally assume that systematic and forward model errors are 0.

The forward model, the data, the noise, and the prior information affect the outcome of the modeling. In FRET experiments, the forward model requires a model of the dyes. Such dye models can compute for a given conformation the distribution of FRET rate constants. Dye models are essential parts of FRET forward models and need to be critically evaluated on a case-by-case basis. A full atomistic treatment requires separate simulations for each FRET pair as an exhaustive sampling of the dye's conformational space is challenging. We use a coarse-grained implicit dye model that was shown to be accurate. We further assume slow translational diffusion and fast rotational diffusion on the time scale of fluorescence (<4 ns). The dye model output (FRET rate constants) can be converted into time-resolved fluorescence intensities. This conversion is computationally expensive, as it requires the convolution with the instrument response function (IRF) among other experimental nuisances (e.g., background, IRF corrections, incomplete FRET labeling). Due to these nuisances, direct scoring against fluorescence decays is computationally costly. Thus, instead of approaching the modeling problem as a single-sided problem, where fluorescence decays are computed for MD simulations, we approach the modeling from two sides. First, we transform by Bayesian modeling time-resolved fluorescence intensities (and the associated data noise) into posterior model densities. We marginalize nuisance to obtain probability

https://doi.org/10.1021/acs.jctc.2c01090
J. Chem. Theory Comput. 2023, 19, 2389−2409

densities over distance distributions. Next, for scoring of ensembles, we compute, using implicit dye models as a forward model, ensemble distance distributions. In our workflow, both the output of the Bayesian fluorescence modeling and the output of the structural forward model are distributions that abstract molecular models and experimental data, respectively. This intermediate representation allows for the computationally efficient scoring of ensembles.

In MEM, the posterior model density is defined by eq 4

$$p(M|D, I) = Z^{-1} \exp(-H + \theta S) = Z^{-1} \exp(-G) \quad (4)$$

where $G = H - \theta S$ is a free-energy-like function. For normal distributed data, $H(D, M, I) = 1/2 \cdot \chi^2$, where χ^2 is the sum of squared deviations between the forward model and the data weighted by the inverse of the noise. The maximum of $p(M|D, I)$ is obtained by solving a constrained optimization problem that maximizes $\ln(p(M|D, I))$. The MAP distribution is uniquely defined and can be determined without an expensive computation of Z^{-1} by minimizing $G = H - \theta S$ using computationally inexpensive gradient descent methods.

Here, $\{C_j\}$ and $\{w_j^{(0)}\}$ are sets of representative conformations and corresponding cluster weights from clustered MD trajectories. The initial weight $w_j^{(0)}$ of the cluster j corresponds to the fraction of that cluster within the ensemble of MD trajectories for a respective FF. The MD analysis is described in detail in a separate section.

D is a set of K FRET experiments, $D = \{D_k\}_{k=1}^{k=K}$. Each D_k measures a donor−acceptor (DA) distance, R_{DA}, distribution, $x_e(R_{DA})_k$, with the corresponding uncertainty, $\Delta x_e(R_{DA})_k$. In a forward model, for D_k, a weighted average model R_{DA} distribution, $x_m(R_{DA})_k$, is computed in two steps: first, for each D_k and entropy value, $x_m(R_{DA})_{jk}$ is computed from the accessible volume (AV) center-to-center distance,[54−56] R_{MP}, and a noncentral χ-distribution (eq 5)

$$x_m(R_{DA})_{jk} = (\sigma_{DA}\sqrt{2\pi})^{-1} \frac{R_{DA}}{R_{MP}} \left(e^{-\left(\frac{R_{DA}-R_{MP,jk}}{2\sigma_{DA}}\right)^2} - e^{-\left(\frac{R_{DA}+R_{MP,jk}}{2\sigma_{DA}}\right)^2} \right) \quad (5)$$

where σ_{DA} represents the overall broadness of $x_m(R_{DA})_{jk}$. We used σ_{DA} = 8.5 Å, as this value is consistent in reference experiments[54] and MD simulations.[57] Second, $x_m(R_{DA})_k$ is computed over all C_j using w_j as the weighting factor (eq 6)

$$x_m(R_{DA})_k = \sum_{j=1}^{J} w_j x_m(R_{DA})_{jk} \quad (6)$$

To compute χ^2, we discretize x_m and x_e into N_b bins of 1.9 Å, considering that FRET experiments are most sensitive around the Förster radius R_0. For our experiments, the reliable range is $0.65R_0$−$1.5R_0$[58] For the FRET pair used here (AlexaFluor488, AlexaFluor647), R_0 = 52 Å; thus, the reliable range is 30−80 Å. Accordingly, distances <30 and >80 Å are represented by a single bin each. To allow for nonuniform binning, bin counts are normalized by respective bin widths. We quantify the discrepancy between M and D_k by eq 7

$$\chi^2_{r,k} = \frac{1}{N_b}\chi^2_k = \frac{1}{N_b} \sum_{i=1}^{N_b} \left[\frac{x_e(R_{DA}^{(i)})_k - x_m(R_{DA}^{(i)})_k}{\Delta x_e(R_{DA}^{(i)})} \right]^2 \quad (7)$$

where N_b is the total number of bins and $x_e(R_{DA}^{(i)})_k$, $x_m(R_{DA}^{(i)})_k$, and $\Delta x_e(R_{DA}^{(i)})$ are the experimental species fraction, the model species fraction, and the experimental uncertainty for the $R_{DA}^{(i)}$ bin, respectively. The overall discrepancy for all $D_k \in D$ is $\chi^2 = \sum_k \chi^2_k$. To obtain a metric less dependent on $|D|$ = K, we normalize $\chi^2_{r,k}$ by the number of measurements (eq 8)

$$\chi^2_{r,total} = \frac{1}{K} \sum_{k=1}^{K} \chi^2_{r,k} \quad (8)$$

MEMs minimize $G = 1/2\chi^2_{r,total} - \theta S$ by varying $\{w_j\}$.[59] We minimized G by our in-house open-source software MaxEnt-pRda (https://github.com/Fluorescence-Tools/MaxEnt-pRda). MaxEnt-pRda makes use of the Jacobian and relies on optimization algorithms provided by the scipy.optimize module. In line with others,[60] the L-BFGS-B algorithm gave the best combination of convergence speed and reliability. For the system investigated here, MaxEnt-pRda with L-BFGS-B converges in about a hundred iterations independent of the initial weights, i.e., in seconds on a current laptop.

Note that S is the negative Kullback−Leibler divergence between $w_j^{(0)}$ and the varied weight w_j. When $w_j = w_j^{(0)}$, S = 0. The more negative S, the less similar a posterior ensemble is to the original ensemble. The temperature-like factor θ is a separately defined hyperparameter that controls how strongly deviations between $\{w_j\}$ and $\{w_j^{(0)}\}$ are penalized. Large θ leads to small deviations of $\{w_j\}$ from $\{w_j^{(0)}\}$.[59] Small θ can cause large changes in the $\{w_j\}$ and best agreement with D but can lead to overfitting.[61] We scan θ to obtain a broad range of $\chi^2_{r,total}$−S pairs.

2.2. FRET Network. To acquire information on intramolecular distances by FRET experiments, we generated double cysteine variants and labeled the cysteines by AlexaFluor488-C_5-maleimide and AlexaFluor647-C_2-maleimide (Thermo Fisher Scientific) as donor and acceptor fluorophores, respectively. The set of double cysteine variants, referred to by pairs of substituted amino acid numbers, defines a network: 137−215, 137−268, 137−296, 215−268, 215−296, 255−296, 258−296, 259−296, 260−296, and 268−296. We performed ensemble time-correlated single photon counting (eTCSPC) experiments. Moreover, we performed FRET-FCS analysis[62] on data acquired on a multiparameter fluorescence detection (MFD) setup.[63,64] FCS determines a set of auto- and cross-correlation functions based on spectral channels sensitive to donor and acceptor fluorescence. The analysis of the correlation curves of FRET-labeled samples recovers relaxation times that inform us of the molecular dynamics.[62] Details on the eTCSPC and FCS analysis are provided in Supplemental Materials and Methods. The fit results are compiled in Tables S2 and S3.

2.3. Nonparametric Bayesian Analysis of eTCSPC Data (ucFRET). To maximize the precision and accuracy of recovered distance information, which serves as an input for MEM ensemble refinement and account for data noise (photon noise) ambiguities, we processed the raw eTCSPC data, D, in a Bayesian framework using the ucFRET software in four steps (i−iv). We processed the data for the FRET pairs individually and repeated the steps i−iv for all distance measurements.

In step (i), we gather all data D_l that informs us of distances between labeling sites. Here, D_l is the fluorescence decay of the FRET pair in the absence, $f_{D|D}^{(D0)}$, and presence of FRET, $f_{D|D}^{(DA)}$.

https://doi.org/10.1021/acs.jctc.2c01090
J. Chem. Theory Comput. 2023, 19, 2389−2409

In step (ii), we define a representation suitable for D_k. The key determinant of the representation is the DA distance distribution, $x(R_{DA})_k$. We use mixture models, e.g., of Gaussian distributions, to parameterize $x(R_{DA})_k$. As a forward model to compute model fluorescence decays, we use a model for small organic fluorophores homogeneously quenched by FRET.[58] Briefly, the model factorizes $f_{DID}^{(DA)}$ into the FRET-induced decay, $\epsilon_{D,k}(t) = \int x(R_{DA})_k \cdot e^{-k_{RET}(R_{DA})t}\, dx$, and $f_{DID}^{(D0)}$. We use two representation types for $x(R_{DA})_k$: (a) mixtures of Gaussians of variable location, height, and width and (b) mixtures of multiple noncentral χ-distributions, χ_m (eq 5), of variable location, $R_{MP,m}$, and width, σ_{DA}; $x(R_{DA})_k = \sum_{l=1}^{L} \chi_m(R_{DA}|R_{MP,l}, \sigma_{DA})$. In representation (a), the number of Gaussians, n_G, is a hyperparameter determined by comparing scores of competing models. Typically, n_G is in the range of 1−3. In representation (b), we set the number of χ distributions, L, by the bounds of $R_{MP,m}$ and σ_{DA}. We sample σ_{DA} and $R_{MP,m}$ in the range of [3.0 Å, 9.0 Å] and [10 Å, 130 Å], respectively. Thus, we use $L = (130-10)/3 = 40$ as the number of χ-distributions. Representation (b) minimizes presumptions on the number of states and is better suited for systems with an undefined or large number of states. For scoring of ensembles, we used posteriors of the χ-distribution mixture models. Dyes that are normally distributed have a χ-distributed interdye distribution.

In step (iii), we optimize and sample over variable parameters and nuisance parameters of the representation to find models consistent with D_l using a scoring function that ranks models. The exact scoring function (negative log-likelihood function) considers that fluorescence decays are counting histograms with the Poisson noise as described previously.[65] We optimize representation parameters by minimizing the discrepancy between a computed model, M_k, and D_k. By minimizing the discrepancy, we obtain an M_k that best describes D_k. Moreover, we sample over variable parameters to obtain parameter probability distributions and parameter uncertainties.[27,58]

In step (iv), we analyze and interpret the optimization and sampling results. For representation (a), we obtain n_G posterior distributions of population fractions, distribution widths, and interdye distances. The population fractions inform us of the relative abundance of molecules in a state, the distribution widths of conformational heterogeneity, and the interdye distances of structural features. For representation (b), we obtain a posterior distribution over a set of separation distances, $\{R_{MP,lk}\}_{l=1}^{l=L}$, and a width σ_{DA} that relate to $p(x(R_{DA})|D_k)$, the probability of $x(R_{DA})$ conditioned on D_k. For MEM scoring, we convert $p(x(R_{DA})|D_k)$ to $p(x, R_{DA}|D_k, M_k)$ (eq 9)

$$p(x, R_{DA}|D_k, M_k) = \iint \delta(R_{DA} - R'_{DA}) p(x_m(R_{DA})|D_k)\, dR'_{DA}\, dp(x_m(R_{DA})) \quad (9)$$

and use the $p(x, R_{DA}|D_k, M_k)$ mean and standard deviation as $x_e(R_{DA})$ and $\Delta x_e(R_{DA})$ estimate, respectively, for MEM ensemble reweighting (eq 7). Assuming a fixed experimentally determined Förster radius, we sample from the posterior of the parameters using ensemble slice samplers.[66] The software implementation of the analysis framework is publicly accessible (https://hub.docker.com/r/tpeulen/ucfret).

2.4. Molecular Dynamics Simulations. All-atom MD simulations were performed for Lif using five FFs, namely, FF99SB,[67] FF14SB,[68] FF19SB,[69] FF14IDPS[70] (here referred to as IDPSFF), and FF99SBdisp.[71] The simulations using FF99SB were performed with the Amber11 software suite[72] with the TIP3P water model,[73] and simulations using FF14SB or FF19SB were performed with Amber20.1[72] and the OPC water model.[74] Simulations using IDPSFF were performed with Amber20.1 and the recommended TIP3P water model,[70] and simulations using FF99SBdisp were performed with Amber19.4[72] and a modified version of the TIP4P-D water model, which is native to FF99SBdisp[75] (Table S4). Besides the diverse protein FFs and water models, we used different procedures for the minimization and thermalization of the systems to support generating independent structural ensembles. Details on the minimization and thermalization procedures are provided in the Supplemental Material and Methods. The SHAKE algorithm[76,77] was used in the production runs to constrain the bond lengths of hydrogen atoms, and the particle mesh Ewald method[78] was used to treat long-range electrostatic interactions with a direct-space, nonbonded cutoff of 8 Å. The integration step for the MD simulations was 2 fs. Production runs of 10 independent replicas with a length of 1 μs each were performed for all FFs. Conformations were saved at 20 ps intervals (settings collected in Table S5).

Clustering of conformations and geometrical analyses were performed with pytraj and CPPTRAJ[79] of Amber20.1.[72] To reduce the number of conformers obtained from the five sets of MD simulations, we employed hierarchical clustering using the best-fit coordinate RMSD of all atoms. We employed a maximum distance of 4 Å between the clusters as cluster criterion (average linkage) for each set of simulations. The sets of trajectories were sieved prior to clustering, i.e., the clustering was performed on every 25th frame of the trajectory set, which is equal to a time step of 0.5 ns, and the remaining frames were added to the respective clusters after the clustering. For all consecutive calculations, such as the determination of interdye distances and MEM reweighting, the pooled cluster representatives (which are the conformations closest to the cluster average) were used as $\{C_j\}$. With these settings, we obtained a total of 15,521 clusters of Lif structures. Initial cluster weights $w_j^{(0)}$ were calculated as the size of cluster j divided by the number of structures within the set of simulations. To estimate the conformational space explored by Lif, we performed a principal component analysis (PCA)[80,81] of the C_α atoms for all 15,521 cluster representatives. MEM cluster weight optimization was performed for each FF individually.

Dye positions and, therefore, donor and acceptor dye distances were estimated using the free and open-source software Olga[37] (available at https://github.com/Fluorescence-Tools/Olga).

2.5. Computation of Synthetic FRET Data. To assess the accuracy of MEM, we computed synthetic FRET data for a set of FRET pairs. In the computations, the fluorophores in a FRET pair were attached to Lif by flexible linkers with a length of 20 Å and had a Förster radius, R_0, of 52 Å. To pick the donor and acceptor attachment sites, we use a method based on an automated pair selection planning approach that selects an optimal set of FRET pairs by maximizing the information from FRET experiments as implemented in the Olga[37] software. An optimal set of FRET pairs maximizes the precision of a structural model for a given ensemble that satisfies the corresponding distance restraints. When selecting

https://doi.org/10.1021/acs.jctc.2c01090
J. Chem. Theory Comput. 2023, 19, 2389−2409

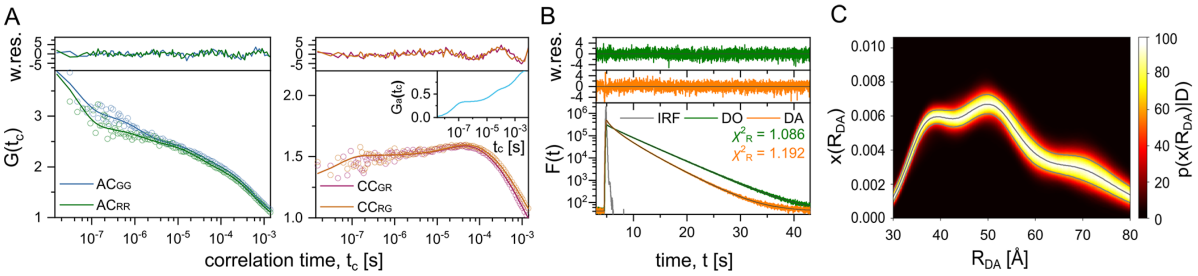


Figure 2. Fluorescence spectroscopy of Lif variant 215−296. (A) Fluorescence correlation spectroscopy (FCS). Global formal fit of the auto- (left) and cross- (right) correlation curves of donor and acceptor channels (SI eqns S1−S5). The inset in the right plot shows the donor−acceptor anticorrelation curve (SI eqn S3), indicating the presence of fast and slow Lif dynamics on the time scale reaching hundreds of microseconds. The results are compiled in Table S2. (B) Exemplary fluorescence decays measured by eTCSPC for the FRET sample (DA) and the donor-only reference sample (DO). The DO sample (green) is fitted with a multi-exponential fit with three fluorescence lifetimes τ_i and species fractions given in parentheses: 4.19 ns (0.781), 2.20 ns (0.131), 0.36 ns (0.088). The best donor−acceptor distance distribution found in Bayesian sampling for the DA sample is displayed as a fit to the orange line. The weighted residuals for the models are plotted on top. (C) Bayesian analysis of fluorescence decays. Probability of interdye distance distribution, $p(x(R_{DA})|D_k)$, conditioned on fluorescence decays of the donor in the presence and absence of the acceptor molecules. Gray lines indicate mean and 3-sigma limits of experimental model density, which are used as $x_e(R_{DA})_k$ and $\Delta x_e(R_{DA})_k$ estimates for MEM ensemble reweighting. All R_{DA} values below 30 Å and all above 80 Å are considered one additional data point each in reweighting. See Figure S4 for all other investigated Lif variants.

FRET pairs, only noncharged and polar amino acids spaced, on average, five residues apart were considered attachment sites to make a choice more experimentally plausible.

The automated FRET pair selection was run on the MD ensembles computed for individual FFs. For each of the five MD simulations for each, we obtained 30 FRET pairs. Among the 5 ×30 pairs, there are 120 unique FRET pairs that were used in a second step to select 50 FRET pairs that are the most informative for all FFs combined (see Supplemental Materials and Methods).

We computed fluorescence decays for an ensemble Y as previously described.[58] Briefly, for each conformer and FRET pair in a target ensemble, Y, we simulated the positional distribution of the donor and acceptor dyes around their attachment sites using the AV approach.[54] For the corresponding interdye distance distributions, we computed FRET rate constants assuming that the dyes diffuse slowly on the time scale of fluorescence. Using the homogeneous approximation that assumes uncorrelated quenching of the donor by other processes and quenching by FRET, we computed fluorescence decays of the donor in the presence of FRET.[58] The fluorescence decay of Y is the weighted average of fluorescence decays of all structures. Experimental data is perturbed by nuisances. Thus, we convolved fluorescence decays of Y with an experimental instrument response function, added nuisances such as scattered light and fluorescence of incompletely labeled molecules to the fluorescence decays, and simulated shot-noise by random draws from Poissonian distributions.[82]

2.6. Ensemble Comparison. Various representations and metrics were proposed to evaluate the similarity between conformational ensembles and determine the quality of resolved ensembles.[83−86] To assess the consistency between posterior ensembles of different FFs after FRET-MEM refinement, we used the Jensen−Shannon divergence D_{JS}[87] of residue pairwise C_α-distance histograms (distograms) and 3D density maps computed for the ensembles; the 3D density maps provide a low-resolution model representation. 3D density maps were computed by mapping coordinates of all atoms of ensemble members onto a grid of uniform dimensions (nx, ny, nz) = (100, 100, 100) and isotropic voxel size of 2 Å. Voxel occupancy of a grid corresponds to a sum of ensemble member weights found in respective voxels. Computed 3D density maps were stored in the MRC2014 file format,[88] a standard for cryo-EM data, using the *mrcfile* python library.[89] For the precision estimate of density maps, we computed the gold-standard Fourier shell correlation (FSC) curves[90] between two independent reconstructions that are free of spurious correlations.[91,92] For that, structures of a given ensemble were split into two sets of the same size. Here, we used interleaved frames, i.e., odd and even frame numbers. For each set of frames and corresponding weights, we computed 3D density half-maps. Subsequently, FSC curves between Fourier transforms of half-maps were computed using the *EMDA* python module.[93] Finally, the precision of a full 3D density map was assigned to a point where the FSC curve crosses the threshold of 0.143.[91] All density maps were displayed using the 3D rendering program UCSF Chimera,[94] where the cross-correlation coefficient (CCC) between a pair of density maps was computed as a similarity measure between ensembles.

3. RESULTS

3.1. Prior Structural Knowledge. The *P. aeruginosa* Lif structure is unknown. However, Lif from *P. aeruginosa* and the *Burkholderia glumae* lipase-specific foldase (*Bg*Lif) have a 39% sequence identity and 52% sequence similarity. The *Bg*Lif FOD embraces its cognate lipase LipA (*Bg*LipA) to form a complex with a headphone-like structure. A homology model of *P. aeruginosa* Lif/LipA based on the *Bg*Lif/*Bg*LipA shows the same tertiary structure, with the FOD consisting of 11 α-helices connected by loops (Figure 1).[43] CD spectroscopy indicates that *Bg*Lif and *Bg*LipA undergo structural changes upon interaction.[46] Knowing that MD1 of the FOD is structurally stable[43] and assuming that MD2 and EHD are predominantly stable with the linkers in between acting as hinges in the absence of LipA, Lif can be geometrically approximated by a 3-body object with two hinge regions, having six internal degrees of freedom (Figure 1). Thus, a combination of fluorescence spectroscopy techniques and MD simulations seems well suited to characterize the conformational space of Lif.

https://doi.org/10.1021/acs.jctc.2c01090
J. Chem. Theory Comput. 2023, 19, 2389−2409

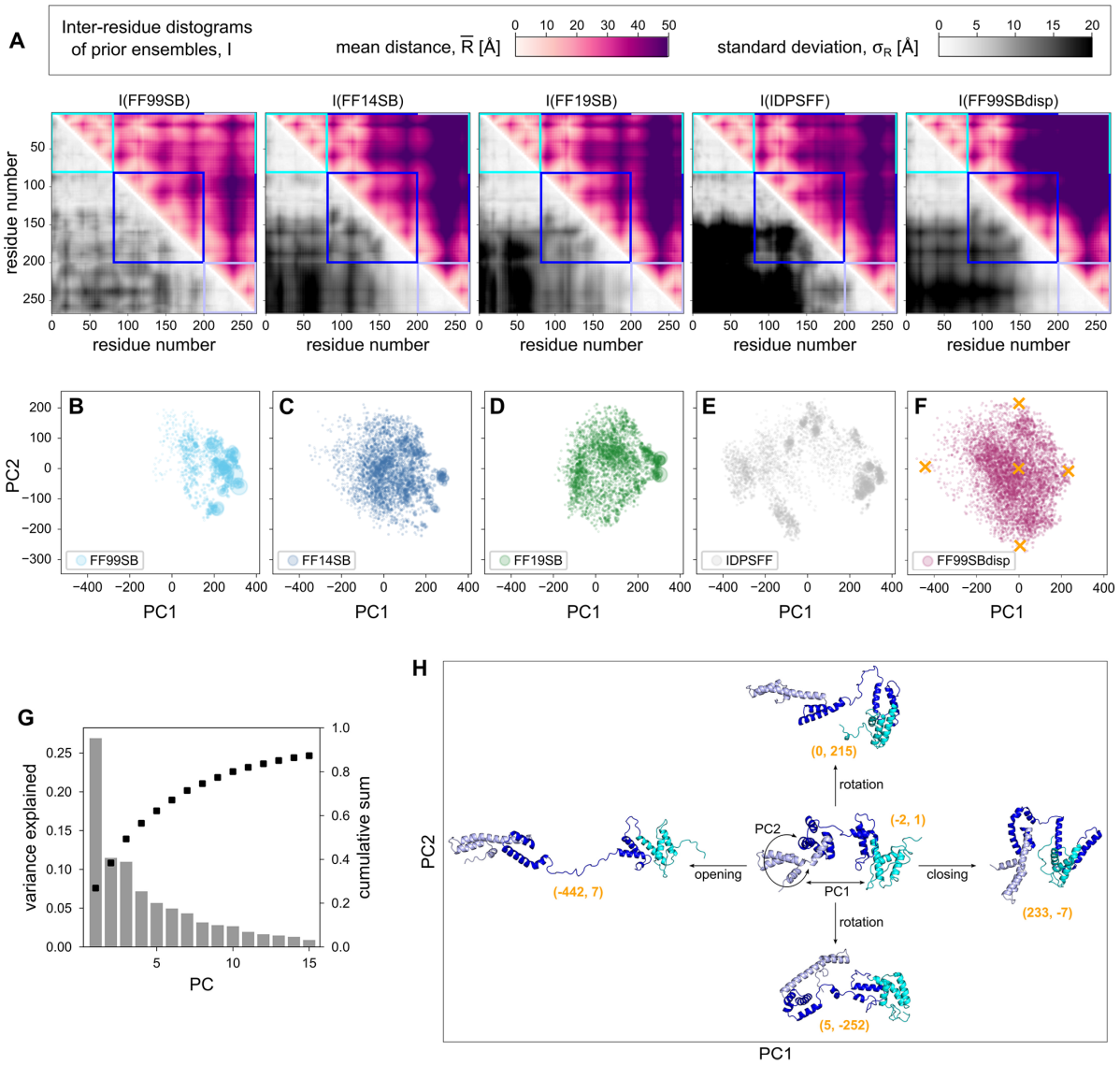


Figure 3. Distogram and principal component analysis of prior conformational ensembles. (A) Inter-residue distograms of prior Lif ensembles I obtained by MD simulations using five FFs. This representation displays the diversity of an ensemble by integrating over ensemble members while retaining the residue (sequence) information. We compute the pairwise mean, $\overline{R}$, and the standard deviation, σ_R. (B–F) Projection of C_α atom positions of cluster representatives for a given FF onto the first two principal components (PC1 and PC2) obtained from a PCA. The PCA was performed on the pooled cluster representatives of all FFs. The marker size for each representative structure scales linearly with the cluster weight. (G) Variance explained by the first 15 PCs (bars) and the cumulative explained variance (squares). (H) Representative Lif structures for different PC1/PC2 values (given in parentheses below each structure) taken from MD simulations with FF99SBdisp (orange crosses in panel F). PC1 describes the opening/closing of MD1 and MD2; PC2 describes the rotation of MD2. Lif domains are colored as in Figure 1.

3.2. Fluorescence Information. To probe the structure and dynamics of Lif, we generated a set of six double cysteine variants that define a FRET network (Figure 1). Cysteine mutations and fluorescent dyes might affect the conformations adopted by Lif. However, activity and binding assays indicate that the mutations and the fluorophores only weakly affect the ability of Lif to bind and activate LipA (Figure S1 and Table S1). We performed FRET-FCS and eTCSPC experiments on the six FRET-labeled double cysteine Lif variants. The FRET-FCS experiment provides information on time scales of exchange between conformational states; the eTCSPC experiment on a FRET sample yields information on a donor–acceptor distance distribution, $x_e(R_{DA})_k$. Our FCS data indicates that unbound Lif exhibits complex structural dynamics, covering time scales from nanoseconds to hundreds of microseconds (Figure 2A). Lif variants show various amplitudes of FRET anticorrelation, which depend on the contrast of FRET efficiencies during dynamics[62] (all curves are shown in Figure S2).

https://doi.org/10.1021/acs.jctc.2c01090
J. Chem. Theory Comput. 2023, 19, 2389–2409

We performed eTCSPC experiments on FRET-labeled double cysteine Lif variants to obtain information on experimental interdye distance distributions $x_e(R_{DA})_k$. First, we applied the Gaussian model (eq S8) to get an insight into the inter-dye distance distribution. As an example, the fluorescence decays of Lif variant 215−296 (donor-only and donor-acceptor sample) display broad $x(R_{DA})$: $\langle R_{DA}\rangle$ = 46 Å, σ_{DA} = 16 Å. Similar features were observed for the other labeled variants (decays in Figure S3 together with fit results in Table S3). The broad $x(R_{DA})$ of *apo*Lif is a direct indication of conformational heterogeneity.[23,95,96] Since there is no indication that unbound Lif adopts discrete states, we used a Bayesian approach (ucFRET) that applies χ distributions with variable location and width (eq 5) in a mixture model to represent $x(R_{DA})$, and use $x(R_DA)$ posterior model densities (eq 9) (Figures 2B and S4) as information for later steps (Figure S5).

Our Bayesian analysis converts fluorescence decays into interdye distance distributions. We estimate the conversion precision by the posterior distribution (Figure 2C). For scoring, we use marginal distributions that do not consider correlations and underestimate the sampling precision. In addition to the precision (data noise), the model and calibration accuracy contribute to the overall uncertainty. The experiments (and simulations) were performed on dyes flexibly tethered to their host. The fluorescence of such dyes can be approximated by models where the translational dye diffusion is slow compared to the time scale of fluorescence (<4 ns).[25,54] The distance distribution between dyes normally distributed around their attachment site is χ-distributed.[97] Thus, $x(R_{DA})$ is represented by a χ distribution mixture model. Time-resolved anisotropy experiments on such dyes showed that the rotational dye diffusion is fast compared to the time scale of fluorescence. Thus, FRET rate constants can be converted to distances using a single Förster radius. Slower rotating dyes require more complex conversion schemes.[98]

3.3. Prior Conformational Ensembles from MD Simulations. Although the FOD of Lif bound to LipA is likely structured,[43] the unbound FOD might have partially disordered regions. Thus, we performed MD simulations of the unbound FOD with five FFs optimized for folded or disordered proteins: we used standard Amber FFs optimized for folded proteins, the classic FF99SB and the more recent FF14SB as well as FF19SB, and, in addition, IDPSFF (a modified FF14SB) and FF99SBdisp that supposedly better capture features of intrinsically disordered/partially disordered proteins. We used FF99SB[67] together with the TIP3P water model in a previous study.[43] FF14SB is an updated FF99SB with improved accuracy of protein side chain and backbone parameters.[68] FF19SB is a further improvement of FF14SB, which tackles an underestimated helix stability in FF14SB.[69] IDPSFF is based on FF14SB, with the backbone dihedral terms of all 20 natural amino acids corrected for IDPs, enhancing the unfolding of collapsed proteins.[70] FF99SBdisp was reported to be capable of describing folded proteins and disordered proteins equally well. Together with its native water model,[71] FF99SBdisp applies modest changes to the backbone torsion and the strength of a backbone O−H Lennard-Jones pair potential to describe both ordered and disordered proteins accurately. For each FF, we performed 10 replicas of MD simulations of 1 μs length each. The replicas for each FF were pooled, and the conformations were clustered. The downstream analysis was performed on cluster representatives and cluster populations, i.e., cluster weights (see the Material and Methods section: clustering and geometry analyses). For FF99SB, FF14SB, FF19SB, IDPSFF, and FF99Sbdisp, we obtained 910, 3640, 2852, 2561, and 5558 clusters, respectively, for an all-atom RMSD cluster threshold of 4 Å. Surprisingly, IDPSFF, although tailored for more flexible proteins, samples more similar conformations than FF14SB and FF19SB, as indicated by the similar number of clusters. Here, we note that the cluster representatives, although capturing the conformational information within a cluster, do not necessarily conserve FRET properties in all cases due to variances in the estimation of the dye position and, therefore, interdye distances. The potential error introduced by performing the downstream analysis on cluster representatives, however, is likely to be similar for all tested FFs, as clustering was performed uniformly. To display the influence of FF on the sequence-dependent structural features of the five prior Lif ensembles, i.e., the ensembles obtained from unbiased simulations, we generate pairwise C_α inter-residue distance histograms (distograms) and show the first (mean, $\overline{R}$) and second moment (standard deviation, σ_R) (Figure 3A) revealing that MD1 (cyan box) and MD2 (light blue box) are structurally more stable in all FFs (small σ_R), while the EHD domain (blue box) is structurally variable. This has consequences for the heterogeneity of the conformational Lif ensembles, that gradually increases from FF99SB, FF19SB, FF14SB, FF99SBdisp to IDPSFF. To conclude, we obtained five different prior ensembles that cover a wide conformational space that can be tested by MEM.

Next, we further characterize the spatial domain configurations in the prior ensembles by principal component analysis to prepare the ground for a subsequent comparison of the prior I with the experimentally derived FRET data (Figure S4) by calculating interdye distance distributions for the prior ensembles (Figure S5).

We mapped the sampled conformational spaces for each FF onto the two main principal components (PCs) obtained from a PCA of the C_α atom positions of all cluster representatives (Figure 3B−F). The first two and three principal components (PCs) account for 38 and 49% of the variance, respectively (Figure 3G). The first PC (PC1) captures an opening and closing motion of Lif (Figure 3H). PC2 and PC3 capture a rotation of MD2 and MD1, respectively. We use PC1 and PC2 to highlight differences among the FFs (see Figure S6 for a projection of the first three PCs). Projection of the FF99SB-derived cluster representatives onto PC1 and PC2 covers the narrowest range of values and favors a compact conformation (Figure 3B). Both FF14SB and FF19SB sample similar conformational spaces (Figure 3A,C,D). IDPSFF samples open conformations more frequently than the other FFs; however, clusters with the highest weights are associated with compact conformations (Figure 3A,E). FF99SBdisp, by contrast, samples the PC1−PC2 space most uniformly of all FFs (Figure 3A,F).

As the stability of secondary structure elements often influences the overall rigidity and flexibility of the protein, we compare the FF's tendency to conserve structural elements such as α-helices. In all simulations, MD1 and MD2 were folded in more than 70% of the frames on average (Figure S7). Lif flexibility was dominated by the flexibility of the EHD composed of helices H4−H8. For these helices, varying secondary structure propensities dependent on the used FF were found. In MD simulations with FF14SB, FF19SB, and

https://doi.org/10.1021/acs.jctc.2c01090
J. Chem. Theory Comput. 2023, 19, 2389−2409

FF99Sdisp, H4−H8 are predominantly helical (>70%). FF99SBdisp seemingly promotes protein flexibility (Figure 3F) while simultaneously preserving the protein's secondary structure (Figure S7). In IDPSFF and FF99SB, the Lif secondary structure is less stable (33 and 59%, respectively; Figure S7). For IDPSFF, which preserves the secondary structure of the EHD the least, this leads to broad coverage of the space spanned by the first two PCs (Figure 3E), yielding more open and closed conformations and fewer intermediate states. By contrast, for FF99SB, the partial collapse of the secondary structure of H4−H8 seemingly promotes the compaction of the protein, paralleled by the lowest number of clusters and the least coverage of the space spanned by the first two PCs (Figure 3B).

Next, we compared interdye distributions obtained from experimental data $x_e(R_{DA})_k$ with $x_m(R_{DA})_k$ based on the forward modeling of ensembles from different FFs (Figure S5). Prior ensembles describe the experimental data poorly. Overall discrepancies ($\chi^2_{r,total}$) are between 13.6 and 20.8 (Table S6). Hence, the donor−acceptor distance distributions from the simulations and experiments in general disagree. For single distances, a better agreement (minimum discrepancy, χ^2) is found for FF99SB simulations and the residue pairs 258−296 (within MD2, $\chi^2_{r,258-296}$ = 2.80) and 259−296 (within MD2, $\chi^2_{r,259-296}$ = 0.84). For other pairs, however, FF99SB simulations underestimate the distances, indicating that Lif collapsed in these simulations without consecutive unfolding or opening events (Figure S8). For FF14SB and FF19SB simulations, we find closed and open conformations. The conformational ensemble from FF19SB simulations agrees slightly better with experiments than FF14SB (Figures S9 and S10). This is particularly prominent for distances 137−215 (within MD1), 137−296 (MD1-MD2), and 137−268 (MD1-EHD). IDPSFF produced ensembles that disagree the least with the experimental data. These ensembles describe the distances 258−296 ($\chi^2_{r,258-296}$ = 2.11) similarly well as FF99SB and 137−296 slightly better ($\chi^2_{r,137-296}$ = 36.81) (Figure S11). FF99SBdisp simulations produced ensembles that agree the least with experiments. Although the FF99SBdisp ensemble describes the distances 137−215 ($\chi^2_{r,137-215}$ = 2.06) and 137−296 ($\chi^2_{r,137-296}$ = 1.15) well, it disagrees with the remaining set of distances (Figure S12). An overview of all χ^2_r for all distances and FFs is provided in Table S7. Hence, although some distances agree better with the experiments, there is no general trend that distances within a minidomain or between them stand out.

In conclusion, the chosen FFs produce structurally different ensembles, covering different regions of the combined conformational space, and show variance in secondary structure conservation. When comparing the prior I to experimental data D, we find poor agreement. Here, we note that due to the flexibility of the protein, the sampling of conformations is potentially poor for all tested FFs. Nevertheless, this approach pictures a valid real-life scenario in which a potentially poorly described prior needs to be refined using additional information.

3.4. MEM Cluster Reweighting with Experimental FRET Data. To overcome the disagreements between MD simulations and experimental data, we used the experimental data in MEM refinement. Our FRET-FCS experiments revealed the presence of fast and slow Lif dynamics on the time scale reaching hundreds of microseconds (Figure 2A). Since the latter cannot be sufficiently sampled, we combined experimental distance information D, with simulated distance information based on conformational ensembles from the MD simulations, the prior information I, consisting of the structure representatives of clusters and corresponding cluster weights of the MD trajectories for the different FFs. We used MEM to find conformational ensembles describing both the experimental data and the MD simulations (Figure 1) by optimizing cluster weights according to eqs 2−4 (Figure 4A). As above, we quantify the discrepancy with D using $\chi^2_{r,total}$ (eq 8). $\chi^2_{r,total} \approx 1$ is of particular interest as ensembles with $\chi^2_{r,total} > 1$ disagree with experimental data (under the assumptions of the used forward model). To quantify differences of prior and posterior weights of conformations, we use the entropy, S (eq 3). For $S \approx -3$, posterior ensembles from priors of FFs best describing D achieve $\chi^2_{r,total} \approx 1$ (Figure 4B). At $S \approx -3$, the observed $\chi^2_{r,total}$ values are virtually converged for all FFs when compared to the $\chi^2_{r,total}$ values obtained at $S < -3$ (Figure 4B). Thus, we use $S \approx -3$ as a reference point for comparing posterior ensembles for different I.

In Figure 4A, we exemplarily depict initial and MEM reweighted distance distributions for the FRET pair 215−296 and FF14SB (for remaining FRET pairs and other FFs, see Figure S5). Unbiased, the ensemble has a discrepancy (eq 7) between the model and the data of $\chi^2_{(215-296)}$ (FF14SB) = 35.21. For $S = -3$, we find $\chi^2_{(215-296)}$ (*FF14SB*) = 0.90. The dependency of $\chi^2_{r,total}$ on S is depicted in Figure 4B for all tested FFs. For $S = -3$, $\chi^2_{r,total}$ (*FF99SB*) = 1.41, $\chi^2_{r,total}$ (*FF14SB*) = 1.16, $\chi^2_{r,total}$ (*FF19SB*) = 1.66, $\chi^2_{r,total}$ (*IDPSFF*) = 1.20, and $\chi^2_{r,total}$ (*FF99SBdisp*) = 1.97. To assess the significance of $\chi^2_{r,total}$ differences, the effective number of observations and the model complexity need to be considered. In a conservative significance estimate, we treat the 10 experimental distance measurements as independent and assume that only the mean, mode, skewness, and kurtosis of a distance distribution are the observed features. Moreover, we assume that Lif has six effective degrees of freedom. Thus, the total degree of freedom is 10 ×4−6 = 34, and 0.06 is an upper estimate for the p-value (F-value = 1.97/1.16 = 1.70) of the comparison of FF14SB and FF99SBdisp MEM ensembles. Consequently, the FF14SB MEM ensemble describes experiments better than the MEM FF99SBdisp. The MEM reweighting of all ensembles for all measured distances is shown in Figures S8−S12.

Overall, posterior ensembles from FF14SB and IDPSFF agree best with experimental results, yielding the lowest $\chi^2_{r,total}$ for the given entropy threshold. This result is surprising as IDPSFF was modified from FF14SB to handle intrinsically disordered proteins by enhancing the protein flexibility, which often preserves the secondary structure to a lower degree.[70] Surprisingly, FF99SBdisp, designed for both ordered and disordered proteins and which samples most uniformly the PC1−PC2 space, disagreed the most with experiments before and after MEM reweighting.

3.5. Posterior Analysis. MEM favored a few clusters. After reweighting, the 50 clusters with the largest weights account for more than 60% of the overall weights; the top 500 clusters account for more than 95% of the overall weights (Figure 4C). For FF14SB, before applying MEM, 677 and 2790 clusters represent 60 and 95% of the conformational ensemble, respectively.

We next compared reweighted ensembles among different FFs and discussed the results for the FFs FF14SB, FF19SB, and IDPSFF because before reweighting, FF14SB and FF19SB are the most similar (Figure 3), and after reweighting, FF14SB

https://doi.org/10.1021/acs.jctc.2c01090
J. Chem. Theory Comput. 2023, 19, 2389−2409

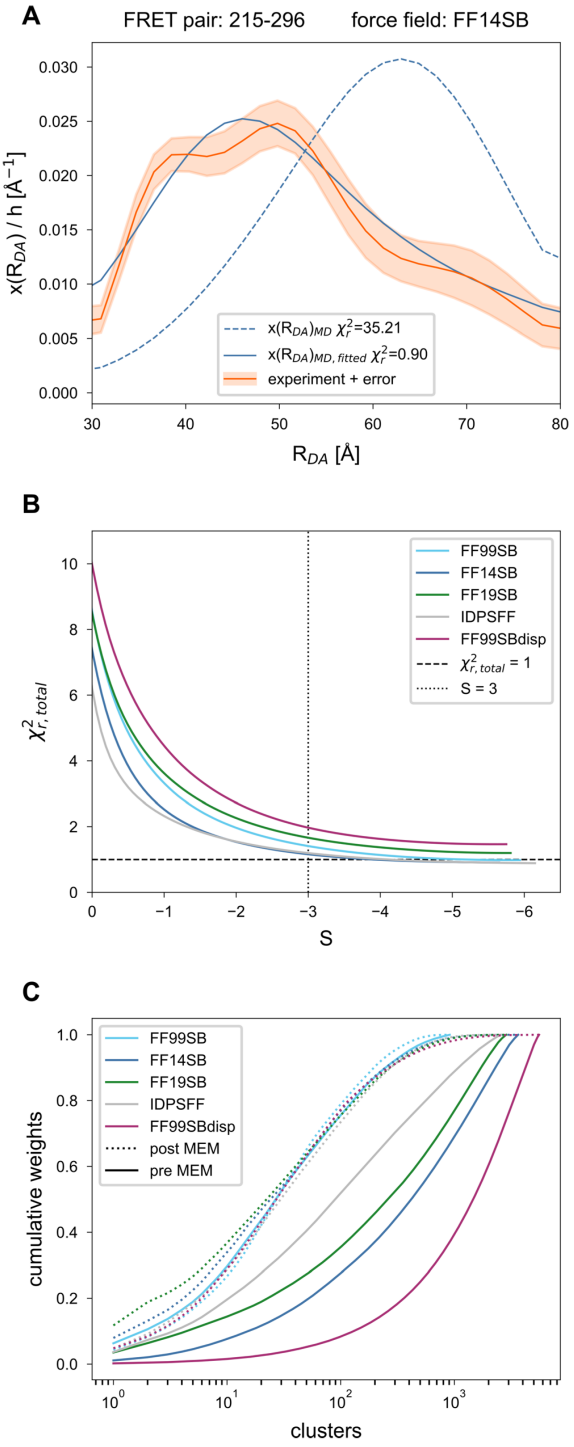


Figure 4. MEM reweighting of cluster weights increases the agreement with experiments while simultaneously reducing the number of relevant conformations. (A) R_{DA} distribution normalized by bin widths (h) for FRET pair 215−296 obtained from 10 × 1 μs of MD simulations using I(FF14SB) (dashed blue line) and post MEM reweighting Q (solid blue line) in comparison to the experimental distribution of R_{DA} (orange line) and its error (orange-shaded area, using a confidence threshold of 0.68) obtained from ucFRET. (B) The total reduced discrepancy $\chi^2_{r,total}$ between experimental, $x_e(R_{DA})$, and model distance distributions, $x_m(R_{DA})$, versus entropy S. (C) Cumulative distribution of prior (solid lines) and posterior (dotted lines) cluster weights for five different FFs (indicated by color).

and IDPSFF ensembles described the data best and second best. The results for all remaining pairwise FF comparisons are depicted in Figures S13−S21. We compare reweighted ensembles based on two hypotheses.

First, we hypothesize that structures ranked similarly are structurally similar. To test our hypothesis, we computed all pairwise cluster representatives in terms of the C_α-RMSD and displayed these similarities for the top 50 clusters in matrices (Figure 5A,C). By definition of the cluster criterion (see above), the average C_α-RMSD of structures within a cluster differ at most by 4 Å from their representative. For FF14SB and IDPSFF, the top 50 clusters with similar weights generally have different structures (Figure 5A). The two most populated clusters (with weights of 7.9 and 5.2% (FF14SB) as well as 4.7 and 4.3% (IDPSFF)) are remarkably dissimilar: we find C_α-RMSD differences in the range of 18.1−29.3 Å (Figure 5B). For FF14SB and FF19SB, which produced comparable prior ensembles, the two most populated clusters (with weights of 7.9 and 5.2% (FF14SB) as well as 11.8 and 7.1% (FF19SB)) are remarkably dissimilar, too (Figure 5D): we find C_α-RMSD differences in the range of 11.4−20.8 Å. This result is surprising, as the conformational ensembles sampled by FF14SB and FF19SB were more similar before reweighting (C_α-RMSD differences in the range of 10.2−14.4 Å). Moreover, a visual comparison of the structures reveals different low-resolution properties such as the interdomain contacts, shapes, or radius of gyration. To conclude, we find structurally different posterior ensembles for FFs that produce posterior ensembles well describing the data (FF14SB and IDPSFF) and FFs that produce similar prior ensembles (FF14SB and FF19SB).

Furthermore, we hypothesize that clusters of high structural similarity between different posteriors have similar weights. To test our hypothesis, we compare weights of structurally similar clusters. We determined 50 cluster pairs iteratively: we selected the most similar cluster representatives between two FFs based on C_α-RMSD, discarding cluster representatives already assigned to a structure pair. The data (Figure 5E,F) reveals no correlation in the posterior weights for similar structures selected from different prior I, and this finding is independent of the magnitude of the structural similarity.

3.6. Assessment of MEM with Synthetic FRET Data.

To assess which model resolution is possible from MEM reconstruction independent of the prior I, we computed synthetic FRET data. We consider the ensemble on which synthetic FRET data is computed the known target ensemble, Y. We then use Y to assess the accuracy and precision of the reconstructed posterior ensembles. We computed synthetic FRET data for two target ensembles. First, MD ensembles obtained for FF14SB depict the target ensemble, i.e., Y(FF14SB) (Figure 6A). Second, Y is a mixture of 488 structures randomly selected from a combined ensemble from MD simulations of all studied FFs, i.e., Y(all); the contribution in terms of the number of structures of each FF in target Y corresponds to the sizes of prior ensembles (FF14SB = 21.5%,

https://doi.org/10.1021/acs.jctc.2c01090
J. Chem. Theory Comput. 2023, 19, 2389−2409

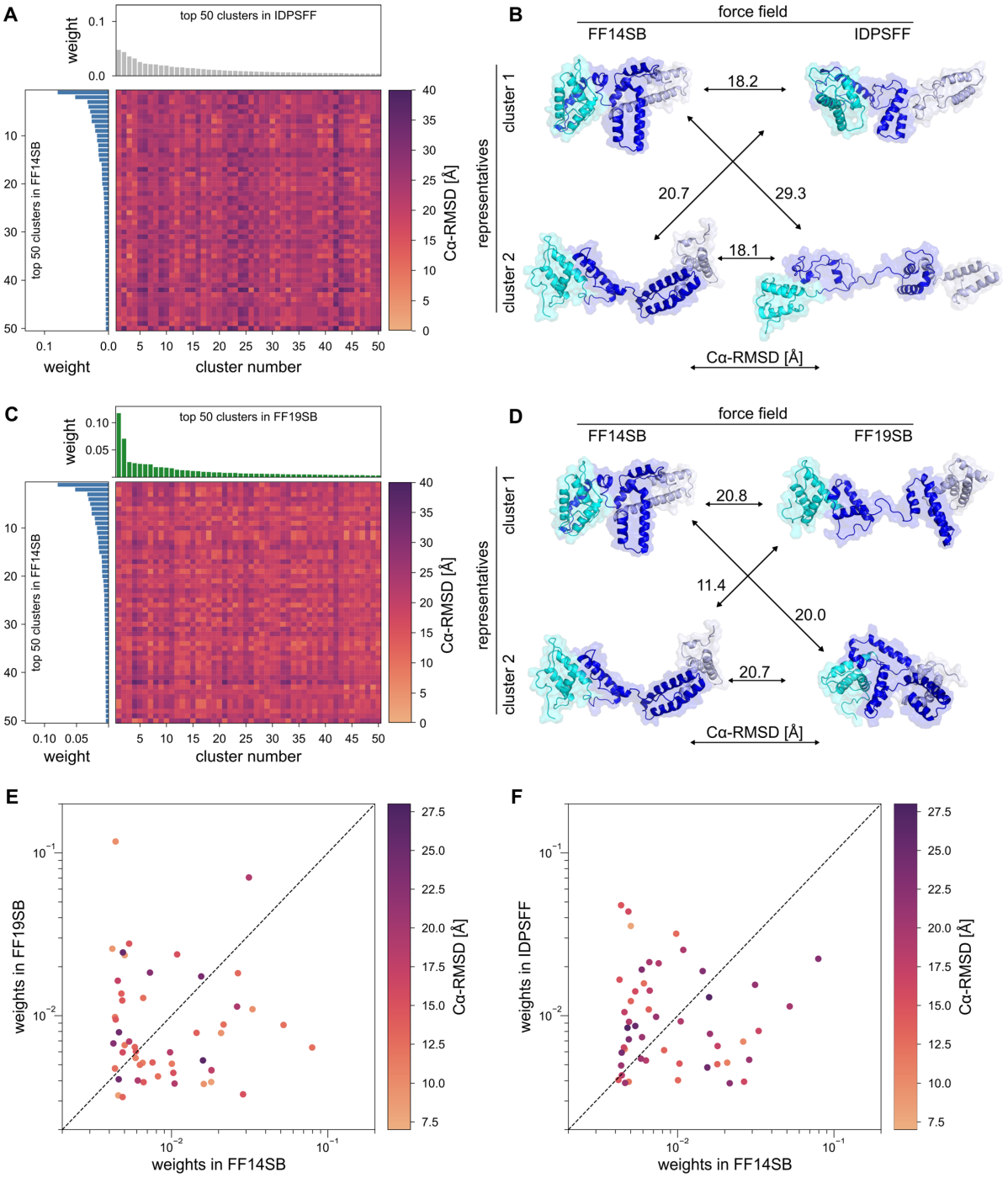


Figure 5. Comparison of posterior ensembles of FF14SB to FF19SB and IDPSFF after reweighting at an entropy threshold of $S = -3$. (A, C) C_α-RMSD matrix for the top 50 cluster representatives reveals structural dissimilarity between cluster representatives despite a reasonable agreement with experimental data in both cases ($\chi^2_{r,total}$ (*FF14SB*) = 1.16, $\chi^2_{r,total}$ (*FF19SB*) = 1.66, and $\chi^2_{r,total}$ (*IDPSFF*) = 1.20). Bar graphs depict the corresponding posterior cluster weight for the 50 most populated clusters. (B, D) Comparison of representatives of the two most populated clusters obtained from MEM. The pairwise structural similarity is given as C_α RMSD. RMSD values in all cases exceed typical values for thermal fluctuations by 1 Å.[99] (E, F) Comparison of posterior weights of the pairwise most similar structures from the top 50 cluster representatives reveals no correlation in the posterior weights.

https://doi.org/10.1021/acs.jctc.2c01090
J. Chem. Theory Comput. 2023, 19, 2389−2409

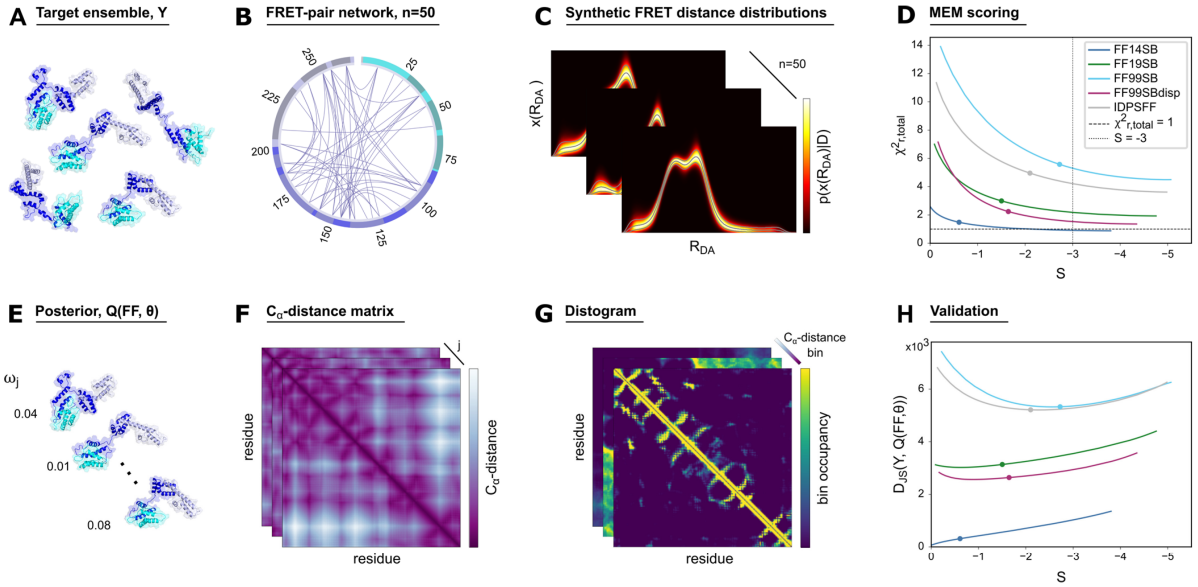


Figure 6. MEM reconstruction based on synthetic FRET data. For a given target ensemble, here *Y*(FF14SB) (A) and a network of the 50 most informative FRET pairs (B), we simulate fluorescence decay curves and convert them by sampling probability distributions over the population of distances (synthetic distance distribution, $x_e(R_{DA})$) (C). The total reduced discrepancy, $\chi^2_{r,total}$, between the model distance distributions, $x_m(R_{DA})$, of the posterior conformational ensemble recovered by MEM and $x_e(R_{DA})$ for all FRET network members versus the entropy, *S* (D). The circles mark θ_c points that balance the synthetic data *D* and the prior *I* based on the L-curve criterion. For a given posterior ensemble with structures and weights (E), we compute residue pairwise C_α-distance matrices (F) and corresponding population-weighted distograms (G). Each point in a distogram corresponds to a distance occupancy computed for C_α atoms of structures of an ensemble. To compare two ensembles, we compute the Jensen−Shannon divergence (D_{JS}) between the posterior distogram and the distogram computed for *Y*, here *Y*(FF14SB) (H).

FF19SB = 17.2%, FF99SB = 6.6%, IDPSFF = 16.4%, FF99SBdisp = 38.3%) used in the target ensemble. Uniform weights are assigned to all members of *Y*(all). We then selected the 50 most informative FRET pairs[37] (Figure 6B). For each FRET pair, we simulated fluorescence intensity decays for donor dyes in the presence of FRET and applied the ucFRET method to recover FRET distances (Figure 6C), which were, in turn, used to optimize the prior ensembles using MEM.

3.7. MEM Cluster Reweighting with Synthetic FRET Data. Predictably, we find that the agreement between synthetic FRET data and the posterior MEM ensembles quantified by $\chi^2_{r,total}$ improved with decreasing *S* (Figure 6D for *Y*(FF14SB), Figure S22 for *Y*(all), and Table S8). We quantify differences between conformational ensembles using the Jensen−Shannon divergence ($D_{JS}(Y, Q)$)[87] of the inter-residue distograms computed for the posterior ensembles, *Q* (Figure 6G). When refining *I*(FF14SB) for *Y*(FF14SB), $\chi^2_{r,total}$ is initially the lowest, as expected, and reaches $\chi^2_{r,total} = 1$ with the least perturbation compared to any other *I* (Figure 6D). Initial discrepancies between the synthetic data and the posterior ensemble for small perturbations are likely due to numerical inaccuracies, experimental noise, and systematic errors due to different forward models for structures (i.e., AV calculations) and the model used to recover experimental $x_e(R_{DA})$ (systematic errors associated with the forward model describing the fluorescence decays).[19] For *Y*(FF14SB) and *I*(FF14SB), D_{JS} continuously increases. With decreasing *S*, deviations between *I* and *Q*(FF, θ) are less penalized, and the synthetic FRET data is better described, which may be due to overfitting (Figures 6D and S22). In line with others,[41] for large θ, the initial ensemble remains unperturbed, and for small θ, a few structures dominate the reweighted ensemble (Figure S23). For the other *I* except FF99SBdisp and FF19SB, $D_{JS}(Y, Q(FF, \theta))$ has a minimum in the proximity of the θ parameter obtained from the L-shaped curve analysis of the $\chi^2_{r,total}$−*S* plot as described by Köfinger et al.[41] The proximity of the D_{JS} minimum justifies the previously used rule-of-thumb of using the "corner" of the $\chi^2_{r,total}$−*S* curve[39,41,59] as a sensible choice for θ, termed θ_c; the corner is the point of maximal positive curvature.[100] In line with previous studies,[20,101] we find that when a large magnitude of *S* is needed to describe the experimental data, this indicates that the prior is not suited to represent the experiment. Additionally, we suggest that to compare and rank priors, one has to consider not only the initial $\chi^2_{r,total}$ but the entire evolution of the L-shaped curve, including the position of the "corner" and the reduction of the $\chi^2_{r,total}$ value at θ_c. As to FF99SBdisp and FF19SB, the L-shaped curves reveal that both priors initially fit poorly to *Y*(FF14SB) and require similarly strong perturbations to reach the "corner". Still, the FF99SBdisp ensemble shows a larger reduction of χ^2_r compared to FF19SB (Figure 6D).

For *I*(FF14SB), which matches *Y*(FF14SB), we found that computing the posterior using $\theta = \theta_c$ is close to the optimal reweighting as judged by the JS divergence and leads to a weaker perturbation of cumulative weights, compared to the stronger perturbation at $\theta_{S=-3}$ (Figure 7A). In C_α-RMSD matrices that compare pairwise *Y* to *Q*(FF14SB, θ), for θ_c under these ideal conditions, we find clusters with smaller C_α RMSD than for $\theta_{S=-3}$ (Figure 7B,C). The change of individual weights is considerable, however, ranging over 3 orders of magnitude and more even at optimally chosen $\theta = \theta_c$ (Figure S23). The C_α-RMSD matrices show that cluster populations

https://doi.org/10.1021/acs.jctc.2c01090
J. Chem. Theory Comput. 2023, 19, 2389−2409

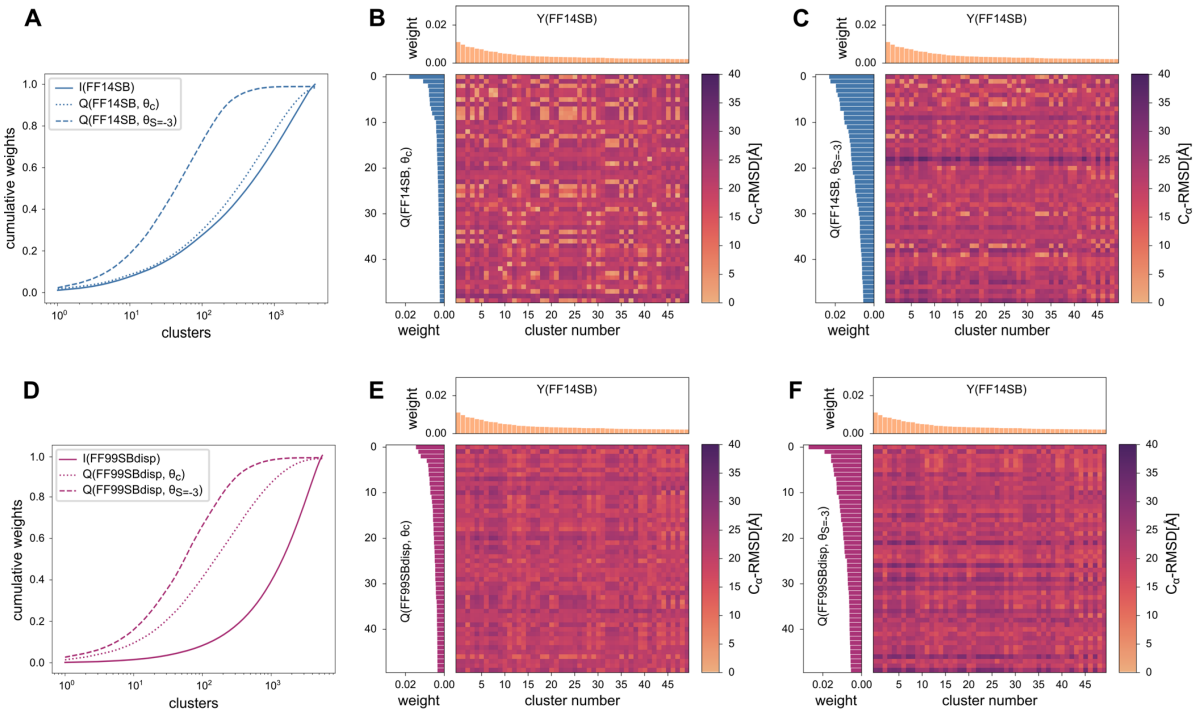


Figure 7. Cumulative weight distribution of prior I and posterior ensembles Q and pairwise Cα-RMSD matrices of the target ensemble, Y(FF14SB), and the posterior ensemble, Q, computed for θ_c and $\theta_{S=-3}$. (A) Cumulative weight for prior I(FF14SB) and posteriors Q(FF14SB, θ_c) and Q(FF14SB, $\theta_{S=-3}$). (B) Pairwise Cα-RMSD matrices of Y(FF14SB) and Q(FF14SB, θ_c) and (C) Y(FF14SB) and Q(FF14SB, $\theta_{S=-3}$). (D) Cumulative weights for I(FF99SBdisp) and Q(FF99SBdisp, θ_c) and Q(FF99SBdisp, $\theta_{S=-3}$). (E) Pairwise Cα-RMSD matrices for Y(FF14SB) and Q(FF99SBdisp, θ_c) and (F) for Y(FF14SB) and Q(FF99SBdisp, $\theta_{S=-3}$). In all panels, clusters are sorted by descending weights.

are uncorrelated to structural similarities and, therefore, indicate that individual atomistic structures of the MEM posterior have no relevance. This is expected as MEM does not optimize weights in a unique way given ensemble-average data as an experimental input.[8,9] In the case of FF99SBdisp, which is the second-best FF based on D_{JS} (Figure 6H) and provides a real-life case ($I \neq Y$), the weights are perturbed more substantially (Figures 7D and S23), and structures of Q(FF99SBdisp, θ_c) deviate by ∼20 Å on average from Y (Figure 7E,F), despite using a network of 50 optimally chosen FRET pairs for the generation of D. Note that for proteins with chain lengths between 100 and 200 residues (Lif considered here has 270 residues), average C_α-RMSD values against ∼2500 alternative folds taken from the PDB range from ∼15 to 20 Å.[102] The average deviation between Y(FF14SB) and Q(FF99SBdisp, θ_c) found here is thus comparable to the structural dissimilarity of two randomly chosen proteins.

3.8. Robustness of Ensemble Features Resolved by MEM. In line with others,[8,9,103] we found that FRET-based MEM is not robust at providing unique solutions when individual (atomistic) structures and corresponding weights are considered. Thus, we test if there are well-resolvable ensemble features beyond the radius of gyration, a low-information ensemble feature widely used to characterize flexible molecules.[20] We compute inter-residue distograms (Figure 6F−G) as a representation for the ensemble diversity. This representation integrates over ensemble members but retains the residue (sequence) information. We compute the pairwise mean, $\overline{R}$, and the standard deviation, σ_R, and the squared relative differences of $\overline{R}$ and σ_R ($\delta^2(\overline{R})$ and $\delta^2(\sigma_R)$ between two distograms) (Figure 8A−C). Comparing the target ensemble Y(FF14SB) with the posterior ensemble Q(FF14SB, θ_c), we find an excellent agreement between $\overline{R}$ (upper triangle) and σ_R (lower triangle) of the Y and Q distograms (Figure 8A). As expected, based on the PCA (Figure 3), the initial deviations between Y(FF14SB) and I(FF99SBdisp) are larger (Figure 8B) compared to I(FFSB14). Nevertheless, we find that MEM improves $\delta^2(\overline{R})$ 4-fold, corresponding to the improvement of the average absolute difference of $\overline{R}$, $\Delta(\overline{R})$, from 5.3 to 2.4 Å, while σ_R on average remains unchanged. For the case of Y(all) and Q(FF14SB, θ_c), $\delta^2(\overline{R})$ improves 10-fold ($\Delta(\overline{R})$ from 5.0 to 1.0 Å), while $\delta^2(\sigma_R)$ improves slightly, from 0.048 to 0.034 (Figure 8C). We find that mean distances, $\overline{R}$, are optimized overall better than σ_R. Moreover, the largest differences in $\overline{R}$ occur within EHD (dark blue frame) and between any pair of subunits, highlighting differences in the local model accuracy.

The average $\delta^2(\overline{R})$ within EHD improves 4-fold for Y(FF14SB) and Q(FF19SBdisp, θ_c) and 12-fold for Y(all) and Q(FF14SB, θ_c), corresponding to improvements of $\Delta(\overline{R})$ from 4.8 to 2.4 Å and 4.0 to 0.9 Å.

As an alternative, we characterized the structural diversity of ensembles by computing density maps of the target ensemble Y, the prior I, and the posterior Q ensembles (Figure 8D−F). Prior to the computation of density maps, target, Y, prior, I, and posterior ensemble, Q, were aligned to all atoms of the first conformer in the corresponding ensemble, which is the most populated cluster in the prior in the case of I and Q. In this

https://doi.org/10.1021/acs.jctc.2c01090
J. Chem. Theory Comput. 2023, 19, 2389−2409

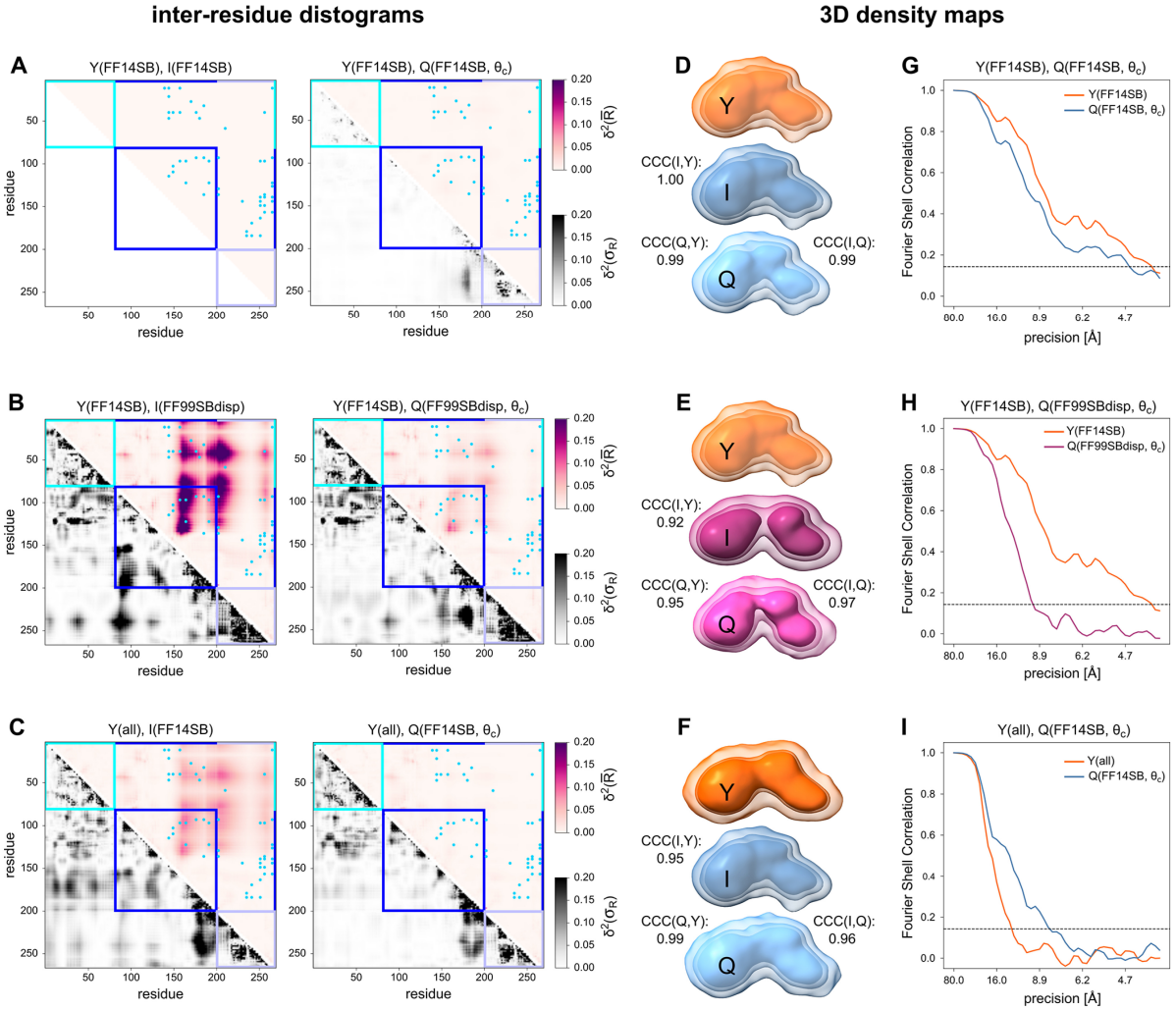


Figure 8. Deviation of prior I and posterior ensemble Q from the target ensemble, Y, for three synthetic experiments (A−C), posterior ensemble density maps (D−F), and ensemble precision estimates (G−I). First row: case $Y = I$, a single FF is the target Y(FF14SB) and the identical FF was chosen as prior I. Second row: case $Y \neq I$, a single FF is the target Y(FF14SB) and another FF was chosen as prior I(F99SBdisp) Third row: case $Y \neq I$, a mixed ensemble from subensembles of all five FFs Y(all) and only the ensemble of one FF was chosen as prior I(FF14SB). Y(all) consists of FF14SB (fraction of 21.5% of Y), FF19SB (17.2%), FF99SB (6.6%), IDPSFF(16.4%), and FF99SBdisp (38.3%), where the fractions in terms of the number of structures of each FF in Y correspond to the sizes of prior ensembles. (A−C) Residue-wise squared relative deviation of prior (left) and posterior (right) of the distogram mean distance $\delta^2(\bar{R})$ (upper-right triangles) and standard deviation $\delta^2(\sigma_R)$ (lower-left triangles) to Y. Cyan, dark blue, and light blue squares and bars mark the residue range of the MD1, EHD, and MD2, respectively. Blue dots indicate FRET pairs. (D−F) Density maps of the target ensemble (orange), Y, the prior (FF14SB, blue; FF99SBdisp, dark pink), I, and the posterior (FF14SB, light blue; FF99SBdisp, light pink), Q, outline the extent of the experimental ensembles at 50%, 68%, and 90% of the density-weighted volume. The cross-correlation coefficients (CCCs) displayed to the left of the densities quantify similarities of prior and posterior entire density maps to the target ensemble, Y, the ones on the right similarities of prior and posterior density maps. (G−I) Fourier shell correlations (FSCs) of the two half-density maps provide precision estimates for target Y (orange) and posterior Q (FF14SB, blue; FF99SBdisp, pink) density maps. The horizontal dashed line at FSC = 0.143 marks a precision and resolution estimate for the gold standard used in electron microscopy.

model representation, we reduce the data further by averaging ensembles over clusters and residues. Representing the ensembles as density maps allows us to use the cross-correlation coefficient (CCC) of two maps as a similarity metric for two ensembles. Overall, the CCCs to the target ensemble are larger for the posterior than for the prior (Figure 8D−F; CCCs displayed to the *left* of the densities), indicating that MEM refines ensemble-integrated features. CCCs between prior and posterior are of the same magnitude, indicating the retained structural similarity between the two despite MEM refinement. We estimated the ensemble precision of the density maps by computing Fourier shell correlation (FSC) curves of two independent half-maps (Figure 8G−I). Using an FSC threshold of 0.143, the average posterior ensemble precision is in the range of 5−10 Å. We use FSCs and CCCs to characterize ensembles and not individual structures. An ensemble with an FSC precision of 5 Å can be composed of structures that have individually a lower

https://doi.org/10.1021/acs.jctc.2c01090
J. Chem. Theory Comput. 2023, 19, 2389−2409

Table 1. Summary of Ensemble Representations That Can Be Robustly Recovered by MEM at $\theta = \theta_c$, Depending on the Agreement between Prior and Experimental Data

completeness of prior	atomistic models[b]	distogram representation[a] mean	distogram representation[a] std	3D density map[c]
prior is complete θ_c is large, i.e., small perturbation of prior weights is needed to describe the data.	−	+	+	+
prior is incomplete θ_c is small, i.e., large perturbation of prior weights is necessary, and few structures dominate the ensemble	−	+	−	+

[a]Averaging over ensemble members. [b]No averaging. [c]Structural diversity: averaging over ensemble members and residues.

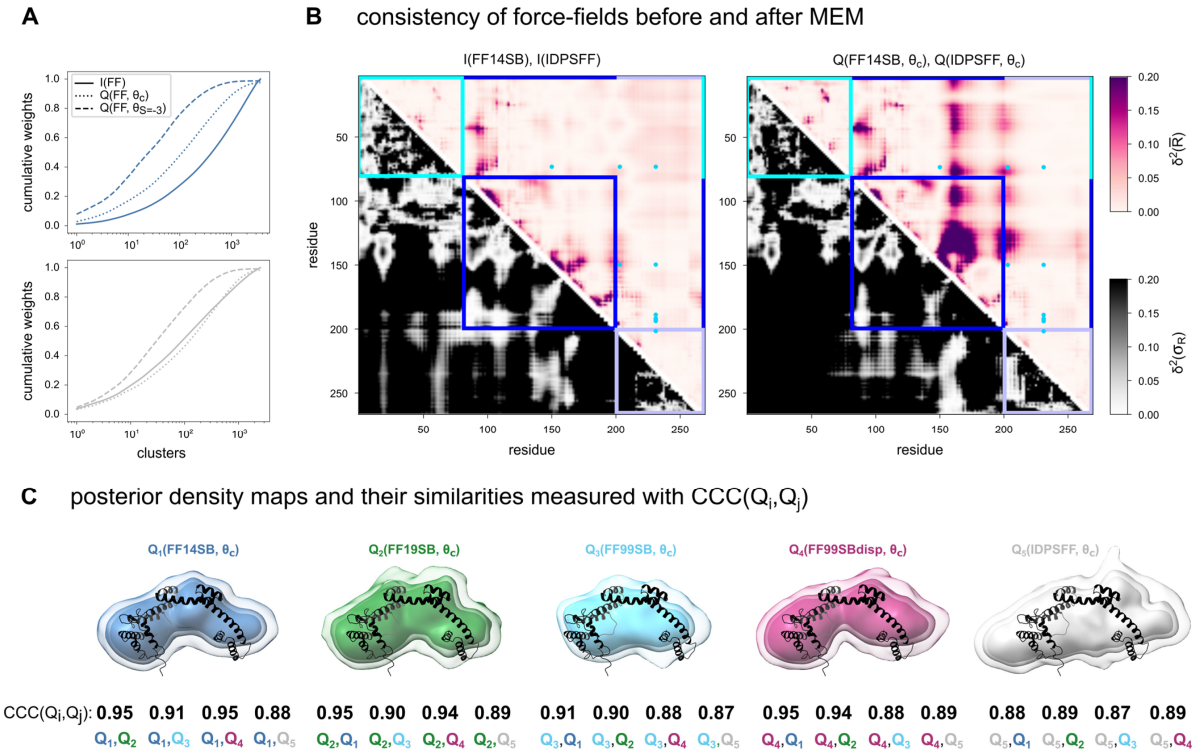


Figure 9. Experimental inter-residue distograms and posterior density maps of Lif. (A) Cumulative distributions of prior and posterior weights for FF14SB (top) and IDPSFF (bottom) ensembles reveal that smaller perturbations of prior weights occur at θ_c compared to $\theta_{S=-3}$. (B) Residue-wise squared relative deviation of distogram mean distance $\overline{R}$ and standard deviation σ_R between FF14SB and IDPSFF prior (left) and posterior (right) ensembles. (C) Posterior density maps of Q(FF, θ_c) contoured at 50, 68, and 90% of the density-weighted volume aligned to all atoms of the homology model are displayed as cartoons. Numbers below the density maps represent cross-correlation coefficient (CCC) values between the posterior density map of a respective force field to all of the others. CCC(*Q*, *Q*) values are color-coded according to the force field coloring scheme.

precision. Thus, FSC ensemble precision estimates should not be compared to FSC resolution estimates of individual structures, e.g., used in cryo-electron microscopy. Moreover, note that sampling reflects the coverage of the clustered conformational ensemble. This is reflected by *Y*'s sampling precision of ≈4 Å (orange curves in Figure 8G,H.) As a logical consequence, reweighting ensembles by MEM affects precision, i.e., the difference in the precision between the two FSC curves can be attributed to MEM. In the case of the mixed ensemble *Y*(all), MEM even improved the precision.

To summarize, based on the benchmark using synthetic data, we have refined our approach to optimally balance different sources of information using the L-shaped curve criterion. Furthermore, we show that to rank the priors and judge their agreement with experimental data, one should not solely rely on χ_r^2 but instead analyze the entire evolution of the L-shaped curve. We demonstrate that prioritizing one source of information over the other, i.e., relying solely on χ_r^2, leads to overfitting and increased disagreement between optimized ensembles and the ground truth (Figure 6H). In line with others, we show that, even with a network of 50 FRET pairs and in an ideal case scenario when $Y = I$, achieving atomistic resolution of the posterior ensembles is neither realistic nor feasible (Figure 7B). However, we found that representations of the ensemble diversity on a reduced level of detail (distograms and density maps) can be robustly recovered depending on the agreement between the prior and experimental data and only when optimally balancing the two sources of information by using the corner point of the L-shaped curve. An overview of which ensemble representation can be robustly recovered by MEM depending on the agreement of prior and experiments is provided in Table 1.

https://doi.org/10.1021/acs.jctc.2c01090
J. Chem. Theory Comput. 2023, 19, 2389−2409

3.9. Reevaluation of the Experimental Data. Analyzing synthetic data for known target ensembles led us to refine our MEM approach. Therefore, we reevaluated our experimental data using low-resolution model representations of Lif obtained at the "corner" point of the L-shaped curves (Table S9). As for the synthetic data, for large θ, the prior remains unperturbed, and for small θ, a few structures tend to dominate the reweighted ensemble, which also holds at θ_c (Figure S24). This feature is prominent for FF14SB (Figure 9A, top) and also for IDPSFF, which, at θ_c, has an almost unaltered weight distribution compared to the prior, although the individual weights are perturbed as for FF14SB (Figure 9A, bottom; Figure S24). Since posterior ensembles could not be resolved in atomistic detail (Figure 5), we compare the experimentally determined posterior ensembles at an ensemble-integrated level using distograms.

Interestingly, among the FF14SB and IDPSFF ensembles that fit the experimental data best, the prior ensembles agree better in the distogram comparison than the posterior ensembles at θ_c (Figure 9B). This is in line with PCA analyses, which reveal that the conformational space of the two prior ensembles is different (Figure 3) but that this difference is exacerbated for the posteriors (Figures S25−S26). Note that the number of experimental restraints (effective 6) is much smaller than those from the synthetic data (50) and that internal conformations of EHD are not probed. In accordance, FF14SB and IDPSFF posterior distograms disagree the most for residues 150−180 of the EHD (Figure 9B). In IDPSFF, this region is unstructured in the first 20 residues, whereas it is mainly helical in FF19SB (Figure S7).

Furthermore, the similarity between density maps of FF14SB and IDPSFF ensembles is likewise not improved at θ_c compared to the priors ($CCC_{I(FF14SB,IDPSFF)}$ = 0.88, $CCC_{Q(FF14SB,IDPSFF,\theta_c)}$ = 0.88). Averaged over all FFs, we observe a slight improvement in CCCs ($\langle CCC\rangle_{I(FFs)}$ = 0.88, $\langle CCC\rangle_{Q(FFs,\theta_c)}$ = 0.91).

Based on CCCs ≥ 0.94 for all pairwise combinations of FFs (Figure 9C (bottom line); Table S10), we classified Lif density maps into three categories (Figure 9C). The first category comprises FF14SB, FF19SB, and FF99SBdisp and resembles a headphone shape; the second and third categories are, respectively, formed by FF99SB and IDPSFF, which differ equally from other FFs as from each other. The FF99SB density map represents more compact Lif, whereas the IDPSFF density map is more extended.

Overall, with our experimental data, MEM does not improve the agreement between ensemble features of different priors even at the medium-resolution level of distograms. At the low-resolution level of density maps, we obtained consistent posterior solutions for three of the five FFs. Still, the solutions with the lowest $\chi^2_{r,total}$ (FF14SB and IDPSFF) are distinct, and, at present, we cannot judge which one of the posterior ensembles, or a mixture of both, describes Lif in solution. To evaluate if further experimental data, such as SAXS measurements, could alleviate this ambiguity, we computed SAXS profiles (for details, see Supplemental Materials and Methods) of the posterior ensembles at $\theta = \theta_c$ (Figure S27). Observed differences in the predicted SAXS profiles, particularly noticeable in the Kratky plot, suggest that posterior ensembles possibly score differently against experimental SAXS data.

4. DISCUSSION

In our study, we found that in the absence of error estimation procedures, varying priors is a prerequisite for assessing the robustness of MEM results. The optimal balance between prior information and experimental data is achieved using the L-curve criterion. In view of the many degrees of freedom of the molecular system, the convergence of MEM results can only be achieved when using a distogram or 3D density map representation, depending on the consistency of priors and experimental data.

MEMs optimize posterior conformer populations by balancing information from different sources. Thus, the success of MEM ensemble refinement depends on the (i) sampling convergence and completeness of the prior, (ii) completeness of the experimental information (data sparsity), and (iii) availability of posterior uncertainty estimates. First, issues of sampling convergence and completeness of the prior ensembles can be addressed by more sampling. For *apo*Lif, we did not cover the conformational space sufficiently in 10 unbiased MD simulations with times of 1 μs each per FF, as indicated by the structural differences of the prior ensembles revealed by PCA and exchange time scales reaching hundreds of microseconds determined experimentally. More sophisticated sampling schemes, such as replica exchange simulations,[104,105] modeling transition networks,[106,107] adaptive sampling of Markov state models,[108] or machine learning-based approaches,[109] are needed. Second, the six initially chosen informative labeling site variants were insufficient to cover local conformational heterogeneity as the conformational space of the EHD was severely underestimated. To overcome data sparsity, an extension of the FRET network is necessary. However, this is related to a considerable experimental effort requiring iterative structural and biochemical analysis cycles to identify new labeling sites suitable for FRET experiments.[37] Third, another fundamental restriction is that MEMs in their current implementation lack uncertainty estimates of the posterior. MEMs yield MAP estimates given an entropy prior and the data likelihood, but these are point estimates without information on the variance. Even if MEM reconstructions ascertain that regular solutions are found for underdetermined systems, it is currently not possible to judge the robustness of the solution. Therefore, future ensemble refinement methods will need to consider varying priors, e.g., using methods such as *k*-fold cross-validation, Bayesian model validation, or subsampling of posterior models for estimating posterior uncertainties.[110] Overall, the presented study depicts likely challenges one might face when investigating a flexible system in an integrated approach using MEM and highlights potential pitfalls.

Previously, MEM refinement was applied to IDPs,[19,20] ordered systems,[13,59] and artificial model systems such as Ala_5 pentapeptide.[41] IDPs preclude a structural analysis in atomistic detail, as they are very flexible and do not adopt defined structures. Ordered systems likely can be well studied by the classical FFs, and the effective degrees of freedom of the systems that need to be covered experimentally are small. Model systems can be exhaustively sampled. Hence, in these cases, the challenges mentioned above in MEM refinement were potentially less of an issue.

Considering Lif with its ordered structural parts combined with highly flexible regions, we aimed at generating conformational ensembles for the *apo* state by combining structural

https://doi.org/10.1021/acs.jctc.2c01090
J. Chem. Theory Comput. 2023, 19, 2389−2409

ensembles from MD simulations with FRET data using MEM. As it is not clear to what extent Lif is folded, we chose classical FFs (FF99SB, FF14SB, and FF19SB), an FF tailored to IDPs (IDPSFF), and an FF designed for ordered and disordered proteins (FF99SBdisp). This allowed assessing the recovered posterior ensemble dependency on priors and identifying inconsistencies in posteriors.

In a benchmark based on synthetic FRET data, we find that the optimal conformational ensembles are found when prior and experimental information are balanced using the L-curve criterion. Still, we were unable to obtain atomistic-level information because the ensemble and the associated conformational space are too large. MEM performed best at recovering ensemble-integrated properties (both mean and std of distograms) when the prior is complete, and only small perturbations of the ensemble are required. For cases where the prior is incomplete, e.g., due to an inaccurate FF or insufficient sampling, the recovery of these properties is less reliable, i.e., only the mean of distograms but not std and ensemble shapes are recovered. In our case study, we find that at an entropy lower than our estimated corner of the L-shaped curve ($S = -3$), also medium-resolution features (means of distograms) diverged more from the target ensemble despite lower χ^2_r, indicating overfitting. Hence, we recommend not relying solely on the agreement with experimental data (χ^2_r) but instead to balance sources of information with the L-corner criterion. Thus, scoring by MEM is convenient and robust for validating the overall consistency of the FF-dependent prior and posterior ensembles to experiments because the minimum of the quality parameter, χ^2_r, is well-known and defined and overfitting can be controlled by the L-corner criterium.

Similarly, when using experimental data, the posterior ensembles differ structurally, also for FFs that produce posterior ensembles that best describe the data (FF14SB and IDPSFF) or FFs that produce similar prior ensembles (FF14SB and FF19SB). Likewise, no correlation between the posterior cluster weights was found for similar structures selected from different prior ensembles, independent of the magnitude of the structural similarity. No consistency between posteriors was found for ensemble-integrated features such as distograms either. Only when considering posterior density maps as low-resolution information, MEM yields three distinct outcomes for the five priors. Three posteriors suggest that despite the tendency of Lif to adopt compact and extended conformations, "headphone"-like extended conformations dominate the conformational landscape, reflecting that Lif consists of ordered structural parts combined with highly flexible regions.

In line with others,[9,61,111] we conclude that using cluster-resolved, atomistic-level representations for MEM with ensemble-averaged experimental observables is an ill-defined problem. Accordingly, for this, we obtained multiple different ensembles that depend on their prior. Going to medium- and low-resolution model representations (means of distograms and density maps), we obtain partially consistent results based on different priors, making the problem well defined. Furthermore, our results caution against the interpretation of MEM reconstructions for such heterogeneous systems as studied here when single-point estimates of conformational ensembles are interpreted without additional uncertainty estimates. Using structurally varying priors, e.g., generated with different FFs, is an *ad hoc* approach to estimating the robustness of MEM reconstructions.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.2c01090.

It contains Supplementary Materials and Methods on the generation of Lif variants, FRET spectroscopy and analysis, MD simulations, and the generation of informative FRET pairs; 10 Supplemental Tables and 26 Supplemental Figures with additional data; and Supplemental References (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Karl-Erich Jaeger − *Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany; Institute of Bio- and Geosciences (IBG-1: Biotechnology), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany*; Phone: (+49) 2461 613716; Email: k.-e.jaeger@fz-juelich.de

Claus A. M. Seidel − *Institute for Molecular Physical Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany*; orcid.org/0000-0002-5171-149X; Phone: (+49) 211 81 15881; Email: cseidel@hhu.de

Thomas-Otavio Peulen − *Department of Bioengineering and Therapeutic Sciences and Department of Pharmaceutical Chemistry and Quantitative Biosciences Institute (QBI), University of California, San Francisco, San Francisco 94143 California, United States of America*; Phone: (+49) 211 81 15881; Email: thomas.otavio.peulen@gmail.com

Holger Gohlke − *Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany; John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC), and Institute of Bio- and Geosciences (IBG-4: Bioinformatics), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany*; orcid.org/0000-0001-8613-1447; Phone: (+49) 211 81 13662; Email: gohlke@uni-duesseldorf.de; Fax: (+49) 211 81 13847

Authors

Jonas Dittrich − *Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany*; orcid.org/0000-0003-2377-2268

Milana Popara − *Institute for Molecular Physical Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany*; orcid.org/0000-0003-2626-6096

Jakub Kubiak − *Institute for Molecular Physical Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany*

Mykola Dimura − *Institute for Molecular Physical Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany*

Bastian Schepers − *Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany*

Neha Verma − *Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany*

Birte Schmitz − *Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany*

Peter Dollinger − *Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany*

Filip Kovacic − *Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.2c01090

Author Contributions

[▽]J.D. and M.P. contributed equally.

Author Contributions

J.D.: investigation, formal analysis, visualization, and writing—original draft; M.P.: investigation, formal analysis, visualization, and writing—review and editing; J.K.: investigation, fluorescence measurements, experimental data analysis, visualization, and writing—review and editing; M.D.: software-MEM implementation; Ba.S., N.V., and Bi.S.: investigation; P.D.: sample preparation and LipA activity measurements; F.K. and K.-E.J.: project administration and funding acquisition; C.A.M.S.: conceptualization, writing—review and editing, project administration, and funding acquisition; T.-O.P.: investigation, software-ucFRET, design of analysis, formal analysis, and writing—original draft; H.G.: conceptualization, writing—review and editing, project administration, and funding acquisition.

Funding

This study was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through funding no. GO 1367/1-1 and INST 208/704-1 FUGG to H.G., JA 448/8-1 to K.-E.J., and SE 1195/16-1 to C.A.M.S., and in part by CRC 1208 (project number 267205415, project A02 to F.K. and K.-E.J., A03 to H.G., and A08 to C.A.M.S.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

H.G. gratefully acknowledges the computational support provided by the "Center for Information and Media Technology" (ZIM) at the Heinrich Heine University Düsseldorf and the computing time provided by the John von Neumann Institute for Computing (NIC) on the supercomputer JUWELS at Jülich Supercomputing Centre (JSC) (user IDs: HKF7, VSK33, HDD16). The authors are grateful to Florian Bleffert for help with Lif purification and labeling.

■ REFERENCES

(1) Henzler-Wildman, K.; Kern, D. Dynamic personalities of proteins. *Nature* **2007**, *450*, 964−972.
(2) Kern, D. From structure to mechanism: skiing the energy landscape. *Nat. Methods* **2021**, *18*, 435−436.
(3) Smock, R. G.; Gierasch, L. M. Sending Signals Dynamically. *Science* **2009**, *324*, 198−203.
(4) Dashti, A.; Mashayekhi, G.; Shekhar, M.; Ben Hail, D.; Salah, S.; Schwander, P.; des Georges, A.; Singharoy, A.; Frank, J.; Ourmazd, A. Retrieving functional pathways of biomolecules from single-particle snapshots. *Nat. Commun.* **2020**, *11*, 4734.
(5) Neudecker, P.; Robustelli, P.; Cavalli, A.; Walsh, P.; Lundström, P.; Zarrine Afsar, A.; Sharpe, S.; Vendruscolo, M.; Kay, L. E. Structure of an Intermediate State in Protein Folding and Aggregation. *Science* **2012**, *336*, 362−366.
(6) Kato, H.; van Ingen, H.; Zhou, B.-R.; Feng, H.; Bustin, M.; Kay, L. E.; Bai, Y. Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 12283−12288.
(7) Bertram, K.; Agafonov, D. E.; Liu, W.-T.; Dybkov, O.; Will, C. L.; Hartmuth, K.; Urlaub, H.; Kastner, B.; Stark, H.; Lührmann, R. Cryo-EM structure of a human spliceosome activated for step 2 of splicing. *Nature* **2017**, *542*, 318−323.
(8) Thomasen, F. E.; Lindorff-Larsen, K. Conformational ensembles of intrinsically disordered proteins and flexible multidomain proteins. *Biochem. Soc. Trans.* **2022**, *50*, 541−554.
(9) Bonomi, M.; Heller, G. T.; Camilloni, C.; Vendruscolo, M. Principles of protein structural ensemble determination. *Curr. Opin. Struct. Biol.* **2017**, *42*, 106−116.
(10) Kruschel, D.; Zagrovic, B. Conformational averaging in structural biology: issues, challenges and computational solutions. *Mol. Biosyst.* **2009**, *5*, 1606−1616.
(11) Borgia, A.; Borgia, M. B.; Bugge, K.; Kissling, V. M.; Heidarsson, P. O.; Fernandes, C. B.; Sottini, A.; Soranno, A.; Buholzer, K. J.; Nettels, D.; Kragelund, B. B.; Best, R. B.; Schuler, B. Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **2018**, *555*, 61−66.
(12) Tompa, P. On the supertertiary structure of proteins. *Nat. Chem. Biol.* **2012**, *8*, 597−600.
(13) Lindorff-Larsen, K.; Best, R. B.; DePristo, M. A.; Dobson, C. M.; Vendruscolo, M. Simultaneous determination of protein structure and dynamics. *Nature* **2005**, *433*, 128−132.
(14) Smith, C. A.; Mazur, A.; Rout, A. K.; Becker, S.; Lee, D.; de Groot, B. L.; Griesinger, C. Enhancing NMR derived ensembles with kinetics on multiple timescales. *J. Biomol. NMR* **2020**, *74*, 27−43.
(15) Huang, J.-r.; Warner, L. R.; Sanchez, C.; Gabel, F.; Madl, T.; Mackereth, C. D.; Sattler, M.; Blackledge, M. Transient Electrostatic Interactions Dominate the Conformational Equilibrium Sampled by Multidomain Splicing Factor U2AF65: A Combined NMR and SAXS Study. *J. Am. Chem. Soc.* **2014**, *136*, 7068−7076.
(16) Bertini, I.; Gupta, Y. K.; Luchinat, C.; Parigi, G.; Peana, M.; Sgheri, L.; Yuan, J. Paramagnetism-Based NMR Restraints Provide Maximum Allowed Probabilities for the Different Conformations of Partially Independent Protein Domains. *J. Am. Chem. Soc.* **2007**, *129*, 12786−12794.
(17) Larsen, A. H.; Wang, Y.; Bottaro, S.; Grudinin, S.; Arleth, L.; Lindorff-Larsen, K. Combining molecular dynamics simulations with small-angle X-ray and neutron scattering data to study multi-domain proteins in solution. *PLoS Comput. Biol.* **2020**, *16*, e1007870.
(18) Boura, E.; Różycki, B.; Herrick, D. Z.; Chung, H. S.; Vecer, J.; Eaton, W. A.; Cafiso, D. S.; Hummer, G.; Hurley, J. H. Solution structure of the ESCRT-I complex by small-angle X-ray scattering, EPR, and FRET spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 9437−9442.
(19) Crehuet, R.; Buigues, P. J.; Salvatella, X.; Lindorff-Larsen, K. Bayesian-Maximum-Entropy Reweighting of IDP Ensembles Based on NMR Chemical Shifts. *Entropy* **2019**, *21*, 898.
(20) Hermann, M. R.; Hub, J. S. SAXS-Restrained Ensemble Simulations of Intrinsically Disordered Proteins with Commitment to the Principle of Maximum Entropy. *J. Chem. Theory Comput.* **2019**, *15*, 5103−5115.
(21) Kikhney, A. G.; Svergun, D. I. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett.* **2015**, *589*, 2570−2577.
(22) Weickert, S.; Cattani, J.; Drescher, M. Intrinsically Disordered Proteins (IDPs) Studied by EPR and in-cell EPR. In *Electron Paramagnetic Resonance*; The Royal Society of Chemistry, 2019; Vol. *26*, pp 1−37.
(23) Haas, E.; Wilchek, M.; Katchalski-Katzir, E.; Steinberg, I. Z. Distribution of end-to-end distances of oligopeptides in solution as estimated by energy transfer. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 1807−1811.
(24) Lerner, E.; Barth, A.; Hendrix, J.; Ambrose, B.; Birkedal, V.; Blanchard, S. C.; Börner, R.; Sung Chung, H.; Cordes, T.; Craggs, T.

D.; Deniz, A. A.; Diao, J.; Fei, J.; Gonzalez, R. L.; Gopich, I. V.; Ha, T.; Hanke, C. A.; Haran, G.; Hatzakis, N. S.; Holing, S.; Hong, S.-C.; Hugel, T.; Ingargiola, A.; Joo, C.; Kapanidis, A. N.; Kim, H. D.; Laurence, T.; Lee, N. K.; Lee, T.-H.; Lemke, E. A.; Margeat, E.; Michaelis, J.; Michalet, X.; Myong, S.; Nettels, D.; Peulen, T.-O.; Ploetz, E.; Razvag, Y.; Robb, N. C.; Schuler, B.; Soleimaninejad, H.; Tang, C.; Vafabakhsh, R.; Lamb, D. C.; Seidel, C. A. M.; Weiss, S. FRET-based dynamic structural biology: Challenges, perspectives and an appeal for open-science practices. *eLife* **2021**, *10*, e60416.

(25) Kalinin, S.; Peulen, T.; Sindbert, S.; Rothwell, P. J.; Berger, S.; Restle, T.; Goody, R. S.; Gohlke, H.; Seidel, C. A. A toolkit and benchmark study for FRET-restrained high-precision structural modeling. *Nat. Methods* **2012**, *9*, 1218−25.

(26) Dimura, M.; Peulen, T. O.; Hanke, C. A.; Prakash, A.; Gohlke, H.; Seidel, C. A. M. Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems. *Curr. Opin. Struct. Biol.* **2016**, *40*, 163−185.

(27) Sanabria, H.; Rodnin, D.; Hemmen, K.; Peulen, T.-O.; Felekyan, S.; Fleissner, M. R.; Dimura, M.; Koberling, F.; Kühnemuth, R.; Hubbell, W.; Gohlke, H.; Seidel, C. A. M. Resolving dynamics and function of transient states in single enzyme molecules. *Nat. Commun.* **2020**, *11*, 1231.

(28) Molnar, K. S.; Bonomi, M.; Pellarin, R.; Clinthorne, G. D.; Gonzalez, G.; Goldberg, S. D.; Goulian, M.; Sali, A.; DeGrado, W. F. Cys-Scanning Disulfide Crosslinking and Bayesian Modeling Probe the Transmembrane Signaling Mechanism of the Histidine Kinase, PhoQ. *Structure* **2014**, *22*, 1239−1251.

(29) Furtmann, F.; Porta, N.; Hoang, D. T.; Reiners, J.; Schumacher, J.; Gottstein, J.; Gohlke, H.; Smits, S. H. J. Characterization of the nucleotide-binding domain NsrF from the BceAB-type ABC-transporter NsrFP from the human pathogen Streptococcus agalactiae. *Sci. Rep.* **2020**, *10*, 15208.

(30) Ravera, E.; Sgheri, L.; Parigi, G.; Luchinat, C. A critical assessment of methods to recover information from averaged data. *Phys. Chem. Chem. Phys.* **2016**, *18*, 5686−5701.

(31) Jaynes, E. T. The Relation of Bayesian and Maximum Entropy Methods. In *Maximum-Entropy and Bayesian Methods in Science and Engineering*; Springer, 1988; pp 25−29.

(32) Robustelli, P.; Kohlhoff, K.; Cavalli, A.; Vendruscolo, M. Using NMR Chemical Shifts as Structural Restraints in Molecular Dynamics Simulations of Proteins. *Structure* **2010**, *18*, 923−933.

(33) Sinelnikova, A.; Spoel, D. v. d. NMR refinement and peptide folding using the GROMACS software. *J. Biomol. NMR* **2021**, *75*, 143−149.

(34) Pitera, J. W.; Chodera, J. D. On the Use of Experimental Observations to Bias Simulated Ensembles. *J. Chem. Theory Comput.* **2012**, *8*, 3445−3451.

(35) Marinelli, F.; Faraldo-Gómez, J. D. Ensemble-Biased Metadynamics: A Molecular Simulation Method to Sample Experimental Distributions. *Biophys. J.* **2015**, *108*, 2779−2782.

(36) White, A. D.; Dama, J. F.; Voth, G. A. Designing Free Energy Surfaces That Match Experimental Data with Metadynamics. *J. Chem. Theory Comput.* **2015**, *11*, 2451−2460.

(37) Dimura, M.; Peulen, T.-O.; Sanabria, H.; Rodnin, D.; Hemmen, K.; Hanke, C. A.; Seidel, C. A. M.; Gohlke, H. Automated and optimally FRET-assisted structural modeling. *Nat. Commun.* **2020**, *11*, 5394.

(38) Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379−423.

(39) Bottaro, S.; Bengtsen, T.; Lindorff-Larsen, K. Integrating Molecular Simulation and Experimental Data: A Bayesian/Maximum Entropy Reweighting Approach. In *Structural Bioinformatics*; Springer, 2020; pp 219−240.

(40) Boomsma, W.; Ferkinghoff-Borg, J.; Lindorff-Larsen, K. Combining Experiments and Simulations Using the Maximum Entropy Principle. *PLoS Comput. Biol.* **2014**, *10*, e1003406.

(41) Köfinger, J.; Stelzl, L. S.; Reuter, K.; Allande, C.; Reichel, K.; Hummer, G. Efficient Ensemble Refinement by Reweighting. *J. Chem. Theory Comput.* **2019**, *15*, 3390−3401.

(42) De Martino, A.; De Martino, D. An introduction to the maximum entropy approach and its application to inference problems in biology. *Heliyon* **2018**, *4*, e00596.

(43) Viegas, A.; Dollinger, P.; Verma, N.; Kubiak, J.; Viennet, T.; Seidel, C. A. M.; Gohlke, H.; Etzkorn, M.; Kovacic, F.; Jaeger, K.-E. Structural and dynamic insights revealing how lipase binding domain MD1 of *Pseudomonas aeruginosa* foldase affects lipase activation. *Sci. Rep.* **2020**, *10*, 3578.

(44) Rosenau, F.; Tommassen, J.; Jaeger, K.-E. Lipase-Specific Foldases. *ChemBioChem* **2004**, *5*, 152−161.

(45) Pauwels, K.; Van Molle, I.; Tommassen, J.; Van Gelder, P. Chaperoning Anfinsen: the steric foldases. *Mol. Microbiol.* **2007**, *64*, 917−922.

(46) El Khattabi, M.; Van Gelder, P.; Bitter, W.; Tommassen, J. Role of the Lipase-specific Foldase of *Burkholderia glumae* as a Steric Chaperone. *J. Biol. Chem.* **2000**, *275*, 26885−26891.

(47) Frenken, L. G. J.; de Groot, A.; Tommassen, J.; Verrips, C. T. Role of the lipB gene product in the folding of the secreted lipase of *Pseudomonas glumae*. *Mol. Microbiol.* **1993**, *9*, 591−599.

(48) Ellis, R. J. Steric chaperones. *Trends Biochem. Sci* **1998**, *23*, 43−45.

(49) Pauwels, K.; Lustig, A.; Wyns, L.; Tommassen, J.; Savvides, S. N.; Van Gelder, P. Structure of a membrane based steric chaperone in complex with its lipase substrate. *Nat. Struct. Mol. Biol.* **2006**, *13*, 374−375.

(50) Förster, T. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann. Phys.* **1948**, *437*, 55−75.

(51) Jaynes, E. T. Information Theory and Statistical Mechanics. *Phys. Rev.* **1957**, *106*, 620−630.

(52) Skilling, J. The Axioms of Maximum Entropy. In *Maximum-Entropy and Bayesian Methods in Science and Engineering*; Springer, 1988; pp 173−187.

(53) Schneidman-Duhovny, D.; Pellarin, R.; Sali, A. Uncertainty in integrative structural modeling. *Curr. Opin. Struct. Biol.* **2014**, *28*, 96−104.

(54) Sindbert, S.; Kalinin, S.; Nguyen, H.; Kienzler, A.; Clima, L.; Bannwarth, W.; Appel, B.; Müller, S.; Seidel, C. A. M. Accurate Distance Determination of Nucleic Acids via Förster Resonance Energy Transfer: Implications of Dye Linker Length and Rigidity. *J. Am. Chem. Soc.* **2011**, *133*, 2463−2480.

(55) Klose, D.; Holla, A.; Gmeiner, C.; Nettels, D.; Ritsch, I.; Bross, N.; Yulikov, M.; Allain, F. H. T.; Schuler, B.; Jeschke, G. Resolving distance variations by single-molecule FRET and EPR spectroscopy using rotamer libraries. *Biophys. J.* **2021**, *120*, 4842−4858.

(56) Hellenkamp, B.; Schmid, S.; Doroshenko, O.; Opanasyuk, O.; Kühnemuth, R.; Rezaei Adariani, S.; Ambrose, B.; Aznauryan, M.; Barth, A.; Birkedal, V.; Bowen, M. E.; Chen, H.; Cordes, T.; Eilert, T.; Fijen, C.; Gebhardt, C.; Götz, M.; Gouridis, G.; Gratton, E.; Ha, T.; Hao, P.; Hanke, C. A.; Hartmann, A.; Hendrix, J.; Hildebrandt, L. L.; Hirschfeld, V.; Hohlbein, J.; Hua, B.; Hübner, C. G.; Kallis, E.; Kapanidis, A. N.; Kim, J.-Y.; Krainer, G.; Lamb, D. C.; Lee, N. K.; Lemke, E. A.; Levesque, B.; Levitus, M.; McCann, J. J.; Naredi-Rainer, N.; Nettels, D.; Ngo, T.; Qiu, R.; Robb, N. C.; Röcker, C.; Sanabria, H.; Schlierf, M.; Schröder, T.; Schuler, B.; Seidel, H.; Streit, L.; Thurn, J.; Tinnefeld, P.; Tyagi, S.; Vandenberk, N.; Vera, A. M.; Weninger, K. R.; Wünsch, B.; Yanez-Orozco, I. S.; Michaelis, J.; Seidel, C. A. M.; Craggs, T. D.; Hugel, T. Precision and accuracy of single-molecule FRET measurements—a multi-laboratory benchmark study. *Nat. Methods* **2018**, *15*, 669−676.

(57) Gansen, A.; Valeri, A.; Hauger, F.; Felekyan, S.; Kalinin, S.; Tóth, K.; Langowski, J.; Seidel, C. A. M. Nucleosome disassembly intermediates characterized by single-molecule FRET. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 15308−15313.

(58) Peulen, T.-O.; Opanasyuk, O.; Seidel, C. A. M. Combining Graphical and Analytical Methods with Molecular Simulations To Analyze Time-Resolved FRET Measurements of Labeled Macromolecules Accurately. *J. Phys. Chem. B* **2017**, *121*, 8211−8241.

https://doi.org/10.1021/acs.jctc.2c01090
J. Chem. Theory Comput. 2023, 19, 2389−2409

(59) Różycki, B.; Kim, Y. C.; Hummer, G. SAXS Ensemble Refinement of ESCRT-III CHMP3 Conformational Transitions. *Structure* **2011**, *19*, 109−116.
(60) Orioli, S.; Larsen, A. H.; Bottaro, S.; Lindorff-Larsen, K. How to learn from inconsistencies: Integrating molecular simulations with experimental data. In *Progress in Molecular Biology and Translational Science*; Strodel, B.; Barz, B., Eds.; Academic Press, 2020; Chapter 3, Vol. *170*, pp 123−176.
(61) Hummer, G.; Köfinger, J. Bayesian ensemble refinement by replica simulations and reweighting. *J. Chem. Phys.* **2015**, *143*, 243150.
(62) Felekyan, S.; Sanabria, H.; Kalinin, S.; Kuhnemuth, R.; Seidel, C. A. M. Analyzing Forster Resonance Energy Transfer with Fluctuation Algorithms. *Fluorescence Fluctuation Spectroscopy (Ffs), Pt B* **2013**, *519*, 39−85.
(63) Sisamakis, E.; Valeri, A.; Kalinin, S.; Rothwell, P. J.; Seidel, C. A. M.Accurate Single-Molecule Fret Studies Using Multiparameter Fluorescence Detection. In *Methods in Enzymology*; Academic Press, 2010; Vol. *475*, pp 455−514.
(64) Kudryavtsev, V.; Sikor, M.; Kalinin, S.; Mokranjac, D.; Seidel, C. A. M.; Lamb, D. C. Combining MFD and PIE for Accurate Single-Pair Forster Resonance Energy Transfer Measurements. *ChemPhysChem* **2012**, *13*, 1060−1078.
(65) Maus, M.; Cotlet, M.; Hofkens, J.; Gensch, T.; De Schryver, F. C.; Schaffer, J.; Seidel, C. A. M. An Experimental Comparison of the Maximum Likelihood Estimation and Nonlinear Least-Squares Fluorescence Lifetime Analysis of Single Molecules. *Anal. Chem.* **2001**, *73*, 2078−2086.
(66) Karamanis, M.; Beutler, F. Ensemble slice sampling. *Stat. Comput.* **2021**, *31*, 61.
(67) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.* **2006**, *65*, 712−725.
(68) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696−3713.
(69) Tian, C.; Kasavajhala, K.; Belfon, K. A. A.; Raguette, L.; Huang, H.; Migues, A. N.; Bickel, J.; Wang, Y.; Pincay, J.; Wu, Q.; Simmerling, C. ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J. Chem. Theory Comput.* **2020**, *16*, 528−552.
(70) Song, D.; Luo, R.; Chen, H.-F. The IDP-Specific Force Field ff14IDPSFF Improves the Conformer Sampling of Intrinsically Disordered Proteins. *J. Chem. Inf. Model.* **2017**, *57*, 1166−1178.
(71) Robustelli, P.; Piana, S.; Shaw, D. E. Developing a molecular dynamics force field for both folded and disordered protein states. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, E4758−E4766.
(72) Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E., III; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Greene, D.; Harris, R.; Homeyer, N.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. *AMBER*; University of California: San Francisco, 2021.
(73) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(74) Izadi, S.; Anandakrishnan, R.; Onufriev, A. V. Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* **2014**, *5*, 3863−3871.
(75) Gopal, S. M.; Wingbermühle, S.; Schnatwinkel, J.; Juber, S.; Herrmann, C.; Schäfer, L. V. Conformational Preferences of an Intrinsically Disordered Protein Domain: A Case Study for Modern Force Fields. *J. Phys. Chem. B* **2021**, *125*, 24−35.
(76) Hammonds, K. D.; Ryckaert, J.-P. On the convergence of the SHAKE algorithm. *Comput. Phys. Commun.* **1991**, *62*, 336−351.
(77) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(78) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(79) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084−3095.
(80) Galindo-Murillo, R.; Roe, D. R.; Cheatham, T. E. On the absence of intrahelical DNA dynamics on the μs to ms timescale. *Nat. Commun.* **2014**, *5*, 5152.
(81) Galindo-Murillo, R.; Roe, D. R.; Cheatham, T. E. Convergence and reproducibility in molecular dynamics simulations of the DNA duplex d(GCACGAACGAACGAACGC). *Biochim. Biophys. Acta, Gen. Subj.* **2015**, *1850*, 1041−1058.
(82) Becker, W. *The bh TCSPC Handbook*, 9th ed.; Becker & Hickl GmbH, 2021.
(83) Yang, S.; Salmon, L.; Al-Hashimi, H. M. Measuring similarity between dynamic ensembles of biomolecules. *Nat. Methods* **2014**, *11*, 552−554.
(84) Tiberti, M.; Papaleo, E.; Bengtsen, T.; Boomsma, W.; Lindorff-Larsen, K. ENCORE: Software for Quantitative Ensemble Comparison. *PLoS Comput. Biol.* **2015**, *11*, e1004415.
(85) Lindorff-Larsen, K.; Ferkinghoff-Borg, J. Similarity Measures for Protein Ensembles. *PLoS One* **2009**, *4*, e4203.
(86) McClendon, C. L.; Hua, L.; Barreiro, G.; Jacobson, M. P. Comparing Conformational Ensembles Using the Kullback−Leibler Divergence Expansion. *J. Chem. Theory Comput.* **2012**, *8*, 2115−2126.
(87) Lin, J. Divergence measures based on the Shannon entropy. *IEEE Trans. Inf. Theory* **1991**, *37*, 145−151.
(88) Cheng, A.; Henderson, R.; Mastronarde, D.; Ludtke, S. J.; Schoenmakers, R. H. M.; Short, J.; Marabini, R.; Dallakyan, S.; Agard, D.; Winn, M. MRC2014: Extensions to the MRC format header for electron cryo-microscopy and tomography. *J. Struct. Biol.* **2015**, *192*, 146−150.
(89) Burnley, T.; Palmer, C. M.; Winn, M. Recent developments in the CCP-EM software suite. *Acta Crystallogr., Sect. D: Struct. Biol.* **2017**, *73*, 469−477.
(90) Henderson, R.; Sali, A.; Baker, M. L.; Carragher, B.; Devkota, B.; Downing, K. H.; Egelman, E. H.; Feng, Z.; Frank, J.; Grigorieff, N.; Jiang, W.; Ludtke, S. J.; Medalia, O.; Penczek, P. A.; Rosenthal, P. B.; Rossmann, M. G.; Schmid, M. F.; Schröder, G. F.; Steven, A. C.; Stokes, D. L.; Westbrook, J. D.; Wriggers, W.; Yang, H.; Young, J.; Berman, H. M.; Chiu, W.; Kleywegt, G. J.; Lawson, C. L. Outcome of the First Electron Microscopy Validation Task Force Meeting. *Structure* **2012**, *20*, 205−214.
(91) Rosenthal, P. B.; Henderson, R. Optimal Determination of Particle Orientation, Absolute Hand, and Contrast Loss in Single-particle Electron Cryomicroscopy. *J. Mol. Biol.* **2003**, *333*, 721−745.
(92) Scheres, S. H. W.; Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* **2012**, *9*, 853−854.
(93) Warshamanage, R.; Yamashita, K.; Murshudov, G. N. EMDA: A Python package for Electron Microscopy Data Analysis. *J. Struct. Biol.* **2022**, *214*, 107826.
(94) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605−1612.
(95) Lakowicz, J. R.; Gryczynski, I.; Wiczk, W.; Laczko, G.; Prendergast, F. C.; Johnson, M. L. Conformational distributions of melittin in water / methanol mixtures from frequency-domain measurements of nonradiative energy transfer. *Biophys. Chem.* **1990**, *36*, 99−115.
(96) Lakowicz, J. R.; Gryczynski, I.; Laczko, G.; Wiczk, W.; Johnson, M. L. Distribution of distances between the tryptophan and the N-

terminal residue of melittin in its complex with calmodulin, troponin C, and phospholipids. *Protein Sci.* **1994**, *3*, 628−637.
(97) Barth, A.; Opanasyuk, O.; Peulen, T.-O.; Felekyan, S.; Kalinin, S.; Sanabria, H.; Seidel, C. A. M. Unraveling multi-state molecular dynamics in single-molecule FRET experiments. I. Theory of FRET-lines. *J. Chem. Phys.* **2022**, *156*, 141501.
(98) Hummer, G.; Szabo, A. Dynamics of the Orientational Factor in Fluorescence Resonance Energy Transfer. *J. Phys. Chem. B* **2017**, *121*, 3331−3339.
(99) Zaccai, G. How Soft Is a Protein? A Protein Dynamics Force Constant Measured by Neutron Scattering. *Science* **2000**, *288*, 1604−1607.
(100) Hansen, P. C.; O'Leary, D. P. The Use of the L-Curve in the Regularization of Discrete Ill-Posed Problems. *SIAM J. Sci. Comput.* **1993**, *14*, 1487−1503.
(101) Rangan, R.; Bonomi, M.; Heller, G. T.; Cesari, A.; Bussi, G.; Vendruscolo, M. Determination of Structural Ensembles of Proteins: Restraining vs Reweighting. *J. Chem. Theory Comput.* **2018**, *14*, 6632−6641.
(102) Reva, B. A.; Finkelstein, A. V.; Skolnick, J. What is the probability of a chance prediction of a protein structure with an rmsd of 6 å? *Fold Des.* **1998**, *3*, 141−147.
(103) Gomes, G.-N. W.; Namini, A.; Gradinaru, C. C. Integrative Conformational Ensembles of Sic1 Using Different Initial Pools and Optimization Methods. *Front. Mol. Biosci.* **2022**, *9*, 910956.
(104) Hansmann, U. H. E. Parallel tempering algorithm for conformational studies of biological molecules. *Chem. Phys. Lett.* **1997**, *281*, 140−150.
(105) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141−151.
(106) Noé, F.; Fischer, S. Transition networks for modeling the kinetics of conformational change in macromolecules. *Curr. Opin. Struct. Biol.* **2008**, *18*, 154−162.
(107) Noé, F.; Krachtus, D.; Smith, J. C.; Fischer, S. Transition Networks for the Comprehensive Characterization of Complex Conformational Change in Proteins. *J. Chem. Theory Comput.* **2006**, *2*, 840−857.
(108) Bowman, G. R.; Ensign, D. L.; Pande, V. S. Enhanced Modeling via Network Theory: Adaptive Sampling of Markov State Models. *J. Chem. Theory Comput.* **2010**, *6*, 787−794.
(109) Noé, F.; De Fabritiis, G.; Clementi, C. Machine learning for protein folding and dynamics. *Curr. Opin. Struct. Biol.* **2020**, *60*, 77−84.
(110) Lele, S. R. How Should We Quantify Uncertainty in Statistical Inference? *Front. Ecol. Evol.* **2020**, *8*, 35.
(111) Rieping, W.; Habeck, M.; Nilges, M. Inferential Structure Determination. *Science* **2005**, *309*, 303−306.

SUPPORTING INFORMATION

Resolution of Maximum Entropy Method-derived posterior conformational ensembles of a flexible system probed by FRET and molecular dynamics simulations

Jonas Dittrich[a,†], Milana Popara[b,†], Jakub Kubiak[b], Mykola Dimura[b], Bastian Schepers[a], Neha Verma[a], Birte Schmitz[a], Peter Dollinger[c], Filip Kovacic[c], Karl-Erich Jaeger[c,d,*], Claus A.M. Seidel[b,*], Thomas-Otavio Peulen[e,*], Holger Gohlke[a,f,*]

[†]These authors contributed equally.

[a]Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

[b]Institute for Molecular Physical Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany
[c]Institute of Molecular Enzyme Technology Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany
[d]Institute of Bio- and Geosciences (IBG-1: Biotechnology), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany
[e]Department of Bioengineering and Therapeutic Sciences and Department of Pharmaceutical Chemistry and Quantitative Biosciences Institute (QBI), University of California, San Francisco, San Francisco 94143, CA, United States of America
[f]John von Neumann Institute for Computing (NIC), Jülich Supercomputing Centre (JSC), and Institute of Bio- and Geosciences (IBG-4: Bioinformatics), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

Table of Contents

Supplemental Materials and Methods

Generation of Lif variants. Lif variants with Cys substitutions and missing residues 1-64 were created by standard "QuikChange" (Qiagen) or SLIC[1] PCR methods using Phusion DNA polymerase (Thermo Fisher Scientific), complementary mutagenic oligonucleotide pairs, and the pEHTHis19 plasmid as a template in whole plasmid amplification according to Viegas et al..[2] Lif variants carrying an N-terminal His10-tag were expressed in *E. coli* BL21(DE3) and purified by immobilized metal affinity chromatography using Ni-NTA resins (Qiagen) according to a protocol of Hausmann et al..[3] Lif variants were labeled with malemide-functionalized dyes, Alexa Fluor 488 C5 maleimide (donor) and Alexa Fluor 647 C2 maleimide (acceptor) (Thermo Fisher Scientific), in two steps: i) 100 nmol of purified Lif in labeling buffer (50 mM Tris-HCl pH 8.0 at 4 °C), 0.8 molar parts of Alexa488 dissolved in dimethyl sulfoxide (DMSO), and 5 mM Tris(2-carboxyethyl) phosphine (pH 8) were agitated slowly for 3 h at 25 °C. Single donor-labeled Lif was purified at 10 °C via anion exchange chromatography using a ResourceQ column (1 ml capacity) (GE Healthcare) using 50 mM Bis-Tris pH 6.3 buffer and a linear NaCl gradient 0-500 mM; ii) labeling with Alexa647 dye and purification of double-labeled Lif variants was performed following the donor labeling procedure using a four-fold molar excess of the acceptor dye.

Fluorescence correlation spectroscopy. Single-molecule multiparameter fluorescence detection (MFD) measurements were performed as described previously.[4] Measurements were performed at room temperature (21°C ± 1°C) in 50 mM Tris-HCl, pH 8.0. The laser power in the sample was L_G = 60 µW. MFD data were processed as described by Sisamakis *et al.*[5] FCS was performed according to Felekyan *et al.*[6] Burst IDs of FRET-labelled molecules were used for autocorrelation of "green channels and red channels, G_{AC}, and cross-correlation, G_{CC}, of green channels with red channels. Four correlations curves were fitted with global correlation times denoting FRET dynamics using the following model functions:

$$G_{CC}(t_c) = \mathrm{b}_0 + \frac{1}{N} G_{dif}(t_c) \times G_{ph}(t_c) \times G_a(t_c) \qquad \text{eq. S1}$$

with:

$$G_{dif}(t_c) = \left(1 + \frac{t_c}{t_{dif}}\right)^{-1} \left(1 + \left(\frac{\omega_0}{z_0}\right)^2 \times \frac{t_c}{t_{dif}}\right)^{-1/2} \qquad \text{eq. S2}$$

$$G_a(t_c) = \left(1 - a \times \sum_{i=1}^{3} a_i \, exp\left(\frac{-t_c}{t_{a,i}}\right)\right) \quad \text{eq. S3}$$

$$G_{ph}(t_c) = \left(1 - a_{ph} \, exp\left(\frac{-t_c}{t_{ph}}\right)\right) \quad \text{eq. S4}$$

$$G_{AC}(t_c) = \mathrm{b}_0 + \left(1 + \frac{t_c}{t_{dif}}\right)^{-1} \left(1 + \left(\frac{\omega_0}{z_0}\right)^2 \times \frac{t_c}{t_{dif}}\right)^{-1/2} \left(1 - \sum_{i=1}^{4} \left(b_i + b_i exp\left(\frac{-t_c}{t_{b,i}}\right)\right)\right) \quad \text{eq. S5}$$

The observation volume is approximated by a 3D Gaussian with $1/e^2$ radii in the lateral (ω_0) and the axial direction (z_0). t_{dif} is the diffusion time, a_{ph} and t_{ph} is the amplitude and time of the anti-bunching term related to acceptor photo-bleaching, respectively, a, a_i and $t_{a,i}$ are the amplitudes and relaxation times of the anti-correlation terms, and, b_i and $t_{b,i}$ are the amplitudes and times of the bunching term.

Generation of the donor fluorescence intensity decays. Ensemble time-correlated single photon-counting (TCSPC) measurements of high-precision fluorescence intensity decays were performed on a FluoTime300 High Performance Fluorescence Lifetime Spectrometer equipped with HydraHarp 400 Multichannel Picosecond Event Timer & TCSPC Module and Hybrid Photomultiplier Detector Assembly (Picoquant GmbH, Germany). The excitation source was a SuperK EXTREME supercontinuum white light laser (NKT Photonics, Denmark), set to 488 nm and 19.51 MHz pulsing frequency, passed through BP488/10x (Chroma Technology, USA) interference filter for excitation beam cleaning. The emission wavelength was set to 525 nm and the long-pass LP495nm filter was used to reduce the scattered light contribution. All measurements were performed using a magic-angle polarizer configuration (54.7°) and with the temperature control set to 20.0°C ± 0.1°C in 50 mM Tris-HCl pH 8.0. Time-resolved fluorescence intensity decays were collected using 8 ps bins up to 100000 photons in the most populated bin.

Analysis of TCSPC decays. In time-resolved fluorescence decay experiments, the fluorescence intensity decay of a donor fluorophore in the presence of FRET, $f_{D|D}^{(DA)}(t)$, and in the absence of FRET, $f_{D|D}^{(D0)}(t)$, inform on the distance between the donor and the acceptor, R_{DA}. $f_{D|D}^{(D0)}(t)$ was modeled as a sum of exponentials with species fraction x_i and lifetime τ_i (eq. 6 and eq. 7):

$$f_{D|D}^{(D0)}(t) = \sum_{i=1}^{N} x_{D,i} e^{-t/\tau_{D,i}} \qquad \text{eq. S6}$$

and $f_{D|D}^{(DA)}(t)$ was used to model donor-acceptor distance distribution $x(R_{DA})$:

$$f_{D|D}^{(DA)}(t) = f_{D|D}^{(D0)}(t) \int_{R_{DA}} x(R_{DA}) \exp\left(-t k_0 (R_0/R_{DA})^6\right) dR_{DA} + x_0 f_{D|D}^{(D0)}(t) \qquad \text{eq. S7}$$

R_0 is the Förster radius (here R_0 = 52 Å for donor quantum yield $\Phi_{F,D}$ = 0.8 and radiative rate of unquenched donor dye k_0 = 0.25 ns^{-1}), x_0 is the fraction of the molecules that do not carry an acceptor dye.[7] The distance distribution $x(R_{DA})$ was modeled assuming a normal distribution[8] characterized by parameters mean donor-acceptor distance $\langle R_{DA} \rangle$ and width σ_{DA} (eq. S8):

$$x(R_{DA}) = \sum_{i=1}^{N} \frac{1}{\sigma_{DA}\sqrt{\pi/2}} \exp\left(-2\left(\frac{R_{DA} - \langle R_{DA} \rangle}{\sigma_{DA}}\right)^2\right) \qquad \text{eq. S8}$$

or alternatively using the ucFRET approach.

Molecular dynamics simulations. The structure of Lif used for MD simulations was generated by homology modeling, using Lif from *B. glumae* (PDB ID: 2ES4) as a template (see Viegas et al.[2] for a detailed description of the homology modeling). All-atom molecular dynamics (MD) simulations were performed for the five investigated force fields FF99SB[9], FF14SB[10], FF19SB[11], FF14IDPSFF[12] (here referred to as IDPSFF), and FF99SBdisp[13]. For the simulations using FF99SB, FF14SB, FF19SB, and IDPSFF, the initial Lif structure was solvated in a truncated octahedral box of the corresponding water model (Table S4) using the LEaP program from the Amber20.1 software suite[14]; a minimum water shell thickness of 11.0 Å around the solute was used. For the IDPSFF, the obtained topology file was modified using the grid-based energy correction maps method.[12] The initial structure for FF99SBdisp was solvated in a cubic box using the FF99SB-disp water model and a minimum water shell thickness of 11.0 Å around the solute using GROMACS 2019.3.[15] The obtained topology and coordinate files were converted into the amber format using the "gromber" command in ParmEd[16]. Na^+ counterions were added to neutralize the charges. An overview of the force field combinations is given in Table S4.

For FF99SB, the starting structure was minimized by applying 50 steps of steepest descent minimization followed by 450 steps of conjugate gradient minimization. During the minimization, the solute atoms were restrained by applying decreasing harmonic potentials, with a force constant of 25 kcal mol^{-1} $Å^{-2}$ initially, reduced to 5 kcal mol^{-1} $Å^{-2}$ in the last round. Afterward, the systems were heated from 100 K to 300 K in 50 ps of NVT MD simulations with a force constant of 5 kcal mol^{-1} $Å^{-2}$. Afterward, MD simulations of 250 ps length were performed using the NPT ensemble to adjust the density with the same force constant. At last, the force constant of the harmonic restraints was constantly reduced to zero over 100 ps of NVT MD simulations.

For FF19SB and FF14SB, the starting structures were minimized by applying 50 steps of steepest descent minimization followed by 200 steps of conjugate gradient minimization. During the minimization, the solute atoms were restrained by applying decreasing harmonic potentials, with a force constant of 25 kcal mol^{-1} $Å^{-2}$ initially, reduced to 5 kcal mol^{-1} $Å^{-2}$ in the last round. Afterward, the systems were heated from 100 K to 300 K in 50 ps of NVT MD simulations with a force constant of 5 kcal mol^{-1} $Å^{-2}$. Subsequently, MD simulations of 150 ps length were performed using the NPT ensemble to adjust the density with the same force constant. At last, the force constant of the harmonic restraints was constantly reduced to zero over 5*10 ps of NVT MD simulations, after which 50 ps of unrestrained MD simulations followed.

For IDPSFF, the starting structures were minimized by applying 5000 steps of steepest descent minimization followed by 5000 steps of conjugate gradient minimization. During the minimization, the solute atoms were restrained by applying decreasing harmonic potentials, with a force constant of 25 kcal mol^{-1} $Å^{-2}$ initially, reduced to 5 kcal mol^{-1} $Å^{-2}$ in the second round, and unrestrained in the last round. The systems were heated from 0 K to 100 K in 20 ps of NVT MD simulations with a force constant of 5 kcal mol^{-1} $Å^{-2}$. Afterward, the systems were heated from 100 K to 300 K in 100 ps of NPT MD simulations to also adjust the density with the same force constant, followed by another 100 ps at 300 K. At last, the force constant of the harmonic restraints was constantly reduced to zero over 80 ps of NVT MD simulations, followed by additional 200 ps of unrestrained NVT MD simulations.

To ensure the independence of the replicas using FF99SB, FF14SB, FF19SB, and IDPSFF, production runs were carried out at temperatures of 300.0 K + T, where T was incremented by 0.1 K for each consecutive run.

For FF99SBdisp, the starting structures were minimized by applying 5000 steps of steepest descent minimization followed by 10000 steps of conjugate gradient minimization. During the minimization, the solute atoms were restrained by applying decreasing harmonic potentials, with a force constant of 25 kcal mol^{-1} $Å^{-2}$ initially, reduced to 5 kcal mol^{-1} $Å^{-2}$ in the last round. The systems were heated from 100 K to 300 K in 100 ps of NVT MD simulations with a force constant of 5 kcal mol^{-1} $Å^{-2}$. The initial temperature during the second heating step was incremented by 0.1 K for each consecutive run to ensure the independence of the consecutive production runs. Afterward, MD simulations of 300 ps length were performed using the NPT ensemble to adjust the density with the same force constant. At last, the force constant of the harmonic restraints was constantly reduced to zero over 100 ps of NVT MD simulations, followed by 20 ps of unrestrained NVT MD simulations. Production runs were carried out at a temperature of 300.0 K. An overview of the performed steps during system minimization and thermalization is given in Table S5.

Generation of the informative FRET pairs. The procedure of selecting the most informative FRET pairs was described by Dimura et al.[17] Here, we combined conformational ensembles generated with all five investigated force fields and applied the procedure for a selection of non-charged and non-hydrophobic amino-acids (Ser, Thr, Cys, Asn, Gln, Ala, Met, His) that were spaced by, on average, five residues to make the labeling experimentally plausible. The selected amino acids (numbers: 74, 80, 84, 90, 95, 103, 110, 112, 118, 122, 128, 134, 138, 157, 161, 167, 174, 179, 184, 187, 191, 194, 201, 204, 209, 214, 222, 229, 235, 241, 245, 250, 256, 262,

267, 271, 275, 282, 295, 300, 306, 311, 316, 321, 330) provide 990 possible pairs, of which the 50 most informative FRET pairs were selected for the synthetic experiment (275_122, 316_161, 245_110, 235_90, 267_157, 222_103, 209_316, 187_256, 174_214, 222_330, 74_209, 204_321, 250_300, 229_161, 191_330, 241_157, 229_316, 74_330, 201_250, 209_103, 179_209, 235_306, 256_161, 204_330, 214_321, 194_267, 209_295, 250_316, 321_103, 222_161, 167_214, 74_222, 330_157, 209_84, 245_300, 74_204, 201_321, 214_103, 187_229, 214_316, 250_95, 321_161, 222_321, 187_330, 250_157, 209_110, 250_306, 204_316, 201_267, 241_103, 209_330).

Prediction of SAXS profiles for posterior ensembles. The *saxs* module[18] from the Integrative Modeling Platform[19] (IMP) was used to predict SAXS profiles of posterior conformational ensembles of *apo* Lif. The computation of the SAXS intensity profile, $I(q)$, as a function of the momentum transfer q, is based on the Debye formula[20] (eq. 9):

$$I(q) = \sum_{i=1}^{N}\sum_{j=1}^{N} f_i(q)\, f_j(q) \frac{sin(qd_{ij})}{qd_{ij}} \qquad \text{eq. S9}$$

with $f_i(q)$ being the atomic form factor and d_{ij} the distance between any pair of atoms. Momentum transfer q is a function of the wavelength of incident X-rays, λ, and scattering angle 2θ relative to the incident X-ray beam (eq. 10):

$$q = \frac{4\pi sin(\theta)}{\lambda} \qquad \text{eq. S10}$$

Scattering profiles of ensemble members were predicted in q range from 0.0-0.6 $Å^{-1}$, with a step of dq = 0.0012. Ensemble averaged scattering intensity of posterior ensembles is obtained as a weighted linear combination of SAXS profiles of ensemble members, by taking the weights at the corner point of the L-shaped curve, i.e. at $\theta = \theta_c$. The shape of the scattering profile $I(q)$ versus q informs about the flexibility of a macromolecule, such that for folded proteins it has a characteristic shape with rises and dips, while for unfolded proteins the scattering profile is flat (featureless). For multi-domain proteins connected with flexible linkers, the scattering profile is in between these two scenarios, i.e., it has concealed (smoothed) features.[21] Transformation of scattering profile, $I(q) \cdot q^2$ versus q, known as Kratky plot, informs about the compactness of the macromolecule, and takes on a parabolic to hyperbolic shape going from folded to unfolded proteins.

Supplemental Tables

Table S1. Equilibrium dissociation and activation constants (K_D and K_{act}) for in vitro reaction between pre-active LipA and Lif.

Lif variant	K_D ± SEM [nM]	K_{act} ± SEM [nM]
137-215	21 ± 7	63 ± 24
137-268	16 ± 5	73 ± 9
137-296	25 ± 4	40 ± 6
215-268	24 ± 6	43 ± 4
215-296	25 ± 7	85 ± 9
268-296	29 ± 6	24 ± 4
wt	29 ± 9*	19 ± 2

*The dissociation constant for the wild type (wt) Lif could not be determined with the FRET assay. Hence, we use the value obtained from Viegas et al.[2].

Table S2. Parameters of FRET-FCS global fit (eq. S1-5) of apoLif 215-296 variant.

	CC_{GR}	CC_{RG}		AC_{GG}	AC_{RR}
χ^2	4.18	3.82	χ^2	1.90	1.87
b_0	0.517	0.631	b_0	0.994	0.804
N	0.586	0.529	N	0.309	0.289
t_{dif}, ms	0.835	0.835	t_{dif}, ms	0.835	0.835
ω_0/z_0	5.060	5.060	ω_0/z_0	5.060	5.060
a_1	0.332	0.332	b_1	0.351	0.423
$t_{a,1}$, ms	0.00004	0.00004	$t_{b,1}$, ms	0.00004	0.00004
a_2	0.252	0.252	b_2	0.111	0.121
$t_{a,2}$, ms	0.027	0.027	$t_{b,2}$, ms	0.027	0.027
a_3	0.416	0.416	b_3	0.362	0.211
$t_{a,3}$, ms	1.364	1.364	$t_{b,3}$, ms	1.364	1.364
a_{ph}	0.175	0.312	b_4	0.166	0.079
t_{ph}, ms	40.6	40.6	$t_{b,4}$, ms	0.0013	0.0013
a	0.442	0.487			

Table S3. Model-based FRET analysis of the fluorescence decays shown in Figure S3**.** Fit parameters of fluorescence intensity donor decay (eq. S6) and FRET-induced donor decay (eTCSPC) with donor-acceptor distance distribution modeled by generalized Gaussian distribution (eq. S7 and eq. S8) of fluorescently labeled apoLif variants.

	137-215	**137-268**	**137-296**	**215-268**	**215-296**	**268-296**
Donor decay parameters $f_{D|D}^{(D0)}(t)$, eq. 6						
χ^2_r	1.054	1.035	1.019	1.126	1.086	1.060
τ_1, ns	4.22	4.23	4.22	4.18	4.19	4.13
τ_2, ns	2.19	2.28	2.29	2.11	2.20	1.94
τ_3, ns	0.347	0.42	0.370	0.365	0.364	0.377
x_1	0.819	0.749	0.825	0.695	0.781	0.654
x_2	0.108	0.158	0.108	0.174	0.131	0.204
x_3	0.074	0.094	0.067	0.130	0.088	0.142
$\langle\tau\rangle_x$, ns	3.71	3.57	3.75	3.32	3.59	3.15
$x(R_{DA})$ model parameters, eq. 8						
χ^2_r	1.123	1.105	1.057	1.206	1.081	1.090
x_0	0.000	0.055	0.345	0.005	0.023	0.015
$\langle R_{DA}\rangle$, Å	63.9	57.9	69.9	20.9	46.0	43.8
σ_{DA}, Å	10.3	19.5	17.5	21.5	16.4	16.5
κ	0.404	1.000	0.312	-0.006	-0.087	0.539
$\langle\tau\rangle_x$, ns	2.21	1.74	2.87	0.43	1.48	0.71

Table S4. Force field combinations used in MD simulations.

Protein FF	Water FF	Ion FF
FF14SB	OPC	ionsjc-tip4pew*
FF19SB	OPC	ionsjc-tip4pew*
FF99SB	TIP3P	ionsjc-tip3p*
FF99SBdisp	TIP4P-D*	Ions-charm22*
IDPSFF	TIP3P*	ionsjc-tip3p*

*Recommended for the respective protein/water FF.

Table S5. Overview of minimization and thermalization steps.

Force field step	FF99SB	IDPSFF	FF14SB / FF19SB	FF99SBdisp
SD 25[a]	50 steps	5000 steps	50 steps	5000 steps
CG 25[a]	450 steps	5000 steps	200 steps	10000 steps
SD 5[b]	50 steps	5000 steps	50 steps	5000 steps
CG 5[b]	450 steps	5000 steps	200 steps	10000 steps
SD 0[c]	-	5000 steps	-	-
CG 0[c]	-	5000 steps	-	-
H 0-100 K[d]	-	NVT 20 ps	-	-
H 100-300 K[d]	NVT 50 ps	NPT 100 ps	NVT 50 ps	NVT 100 ps
NPT	250 ps	100 ps	150 ps	300 ps
RR[e]	100 ps	80 ps	50 ps	100 ps
NVT	-	200 ps	50 ps	20 ps

[a] Steepest descent (SD) / conjugate gradient (CG) minimization with restraints of 25 kcal mol^{-1} $Å^{-2}$.
[b] Steepest descent (SD) / conjugate gradient (CG) minimization with restraints of 5 kcal mol^{-1} $Å^{-2}$.
[c] Steepest descent (SD) / conjugate gradient (CG) minimization with no restraints.
[d] Heating.
[e] Restraint reduction from 5 kcal mol^{-1} $Å^{-2}$ to 0 kcal mol^{-1} $Å^{-2}$.

Table S6. Overall discrepancy of simulation data with experiments prior and post MEM cluster reweighting at the entropy threshold S ≈ -3.

FF	$\chi^2_{r,total}$ prior MEM	$\chi^2_{r,total}$ post MEM
FF14SB	18.4	1.16
FF19SB	17.0	1.66
FF99SB	15.0	1.41
FF99SBdisp	20.8	1.97
IDPSFF	13.6	1.20

Table S7. χ^2_r for all individual measured distances for all five force fields prior to MEM reweighting.

	MD1	**MD1-EHD**	**MD1-MD2**	**EHD**	**EHD-MD2**	**MD2**	**MD2**	**MD2**	**MD2**	**H9**
	137-215	**137-268**	**137-296**	**215-268**	**215-296**	**255-296**	**258-296**	**259-296**	**260-296**	**268-296**
FF14SB	16.42	4.08	7.22	18.9	35.21	19.97	26.94	7.68	7.17	40.07
FF19SB	4.29	4.26	6.79	15.51	39.15	18.11	26.29	8.22	3.12	45.74
FF99SB	16.36	5.09	52.07	8.55	14.24	6.7	2.8	0.84	2.18	41.29
FF99SB disp	2.06	21.56	1.15	16.29	34.73	27.89	27.58	10.54	9.23	57.00
IDPSFF	30.68	2.26	36.81	5.65	20.84	1.71	2.11	0.79	1.99	33.37

Table S8. L-shaped curve parameters for the synthetic experiments.

***Y*(FF14SB)**

	θ_c				$S\sim -3$			
FF	S	$\chi^2_{r,total}$	θ	$D_{JS}(Y,Q)$	S	$\chi^2_{r,total}$	θ	$D_{JS}(Y,Q)$
FF14SB	-0.609	1.483	0.40	313	-2.954	0.915	0.035	1011
FF19SB	-1.502	2.994	0.50	3135	-3.018	2.175	0.150	3556
FF99SB	-2.725	5.578	0.65	5330	-2.958	5.342	0.550	5354
FF99SBdisp	-1.646	2.241	0.50	2638	-2.867	1.565	0.150	2899
IDPSFF	-2.101	4.958	0.55	5214	-3.031	4.195	0.300	5287

Y(all)

	θ_c				$S\sim -3$			
FF	S	$\chi^2_{r,total}$	θ	$D_{JS}(Y,Q)$	S	$\chi^2_{r,total}$	θ	$D_{JS}(Y,Q)$
FF14SB	-1.295	1.797	0.50	1547	-2.997	1.112	0.100	1970
FF19SB	-1.472	2.679	0.50	2437	-2.861	2.003	0.150	2736
FF99SB	-2.891	7.108	0.75	6426	-2.994	7.000	0.700	6444
FF99SBdisp	-0.886	1.559	0.45	1209	-3.014	0.914	0.050	1640
IDPSFF	-1.842	2.825	0.50	2837	-2.970	2.105	0.200	2838

Table S9. Corner point parameters of L-shaped curves using experimental FRET data.

FF	S	$\chi^2_{r,total}$	θ_c
FF14SB	-1.382	2.011	0.60
FF19SB	-1.718	2.534	0.60
FF99SB	-1.628	2.321	0.60
FF99SBdisp	-1.935	2.796	0.70
IDPSFF	-1.237	2.069	0.50

Table S10. Cross-correlation coefficients (CCCs) of pairwise combinations of density maps obtained with different force fields.

prior

	FF14SB	**FF19SB**	**FF99SB**	**FF99SBdisp**	**IDPSFF**
FF14SB	1	0.961	0.863	0.933	0.879
FF19SB	0.961	1	0.886	0.895	0.895
FF99SB	0.863	0.886	1	0.781	0.906
FF99SBdisp	0.933	0.895	0.781	1	0.825
IDPSFF	0.879	0.895	0.906	0.825	1

θ_c

	FF14SB	**FF19SB**	**FF99SB**	**FF99SBdisp**	**IDPSFF**
FF14SB	1	0.950	0.907	0.948	0.883
FF19SB	0.950	1	0.901	0.939	0.892
FF99SB	0.907	0.901	1	0.877	0.869
FF99SBdisp	0.948	0.939	0.877	1	0.891
IDPSFF	0.883	0.892	0.869	0.891	1

Supplemental Figures

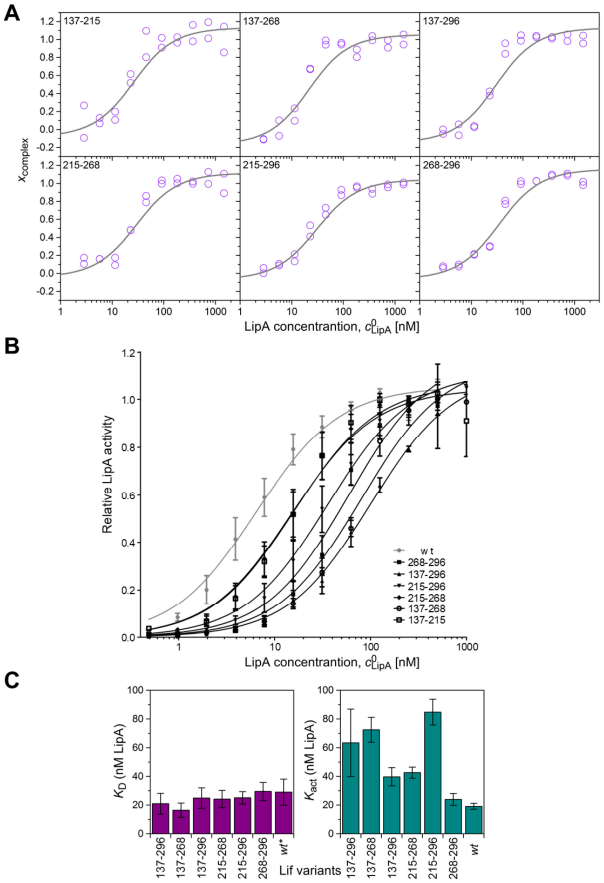


Figure S1. **Activity and affinity of fluorescently labeled Lif variants to LipA.** (A) Equilibrium complex formation. Labeled Lif variants (10 nM) were titrated with pre-active LipA (from 0.5 to 1,000 nM), incubated over night at 4 °C in TB buffer followed by fluorescence measurement. Proximity ratio P = SR/(SG+SR) was converted into the fraction of the complex $x_{complex} = (P - P_0) / (P_{end} - P_0)$, where P_0 and P_{end} are proximity ratios of free Lif and Lif with maximum concentration of LipA. Binding model with the equilibrium dissociation constant K_D was applied to two sets of LipA dilution series. (B) Kinetics of LipA activation. Labeled Lif variants and the wild-type (wt) (2.5 nM) were titrated with pre-active LipA (from 0.5 to 1,000 nM), incubated over night at 4 °C in TB buffer followed by lipase activity assay. The activation constants (K_{act}) were calculated by one site binding hyperbola fit. Results are mean ± SEM of three independent experiments, with each set in triplicates. (C) Equilibrium dissociation constant (K_D); uncertainty is given as standard error of the mean from the fit. Wild-type variant (wt*) was taken from the fluorescence anisotropy binding assay.[2] The activation constant (K_{act}) is referred to as the effective concentration of LipA activated with labeled Lif variants. The results are mean ± SEM of three independent experiments, with each set in triplicates. *K* values are given in **Table S1**.

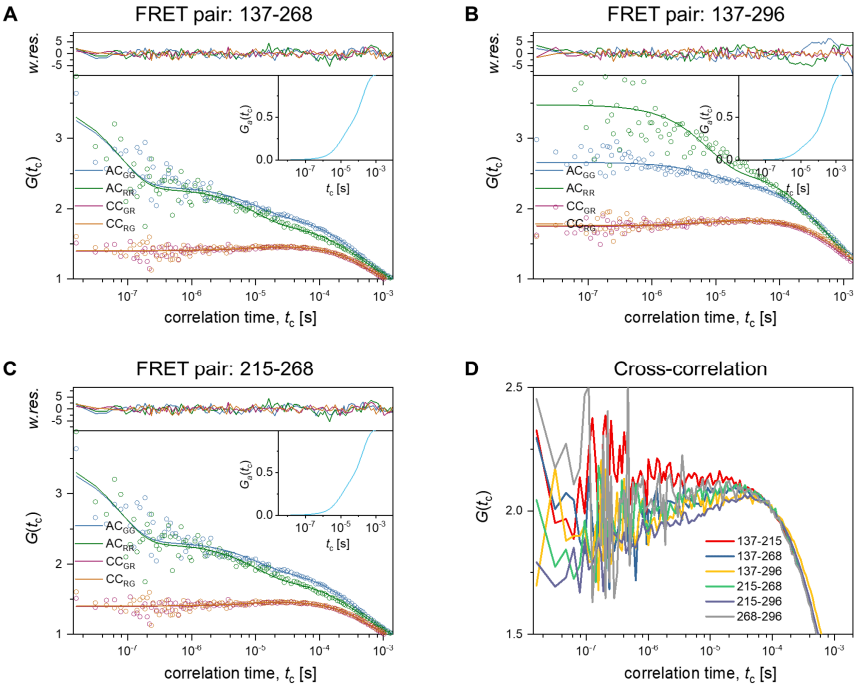


Figure S2. **Dynamics of apoLif probed by dual-color FCS.** (A-C) Global formal fit of the auto- and cross-correlation curves of donor and acceptor channels (eqs. S1-5). The insert shows the donor-acceptor anti-correlation curve (eq. S3) indicating the presence of fast and slow Lif dynamics on the timescale reaching hundreds of microseconds. Fit results are given in **Table S2**. (D) Comparison of anti-correlation amplitudes of donor-acceptor cross-correlation of six FRET variants. Anti-correlation amplitude depends on FRET contrast (range of dynamics).[6]

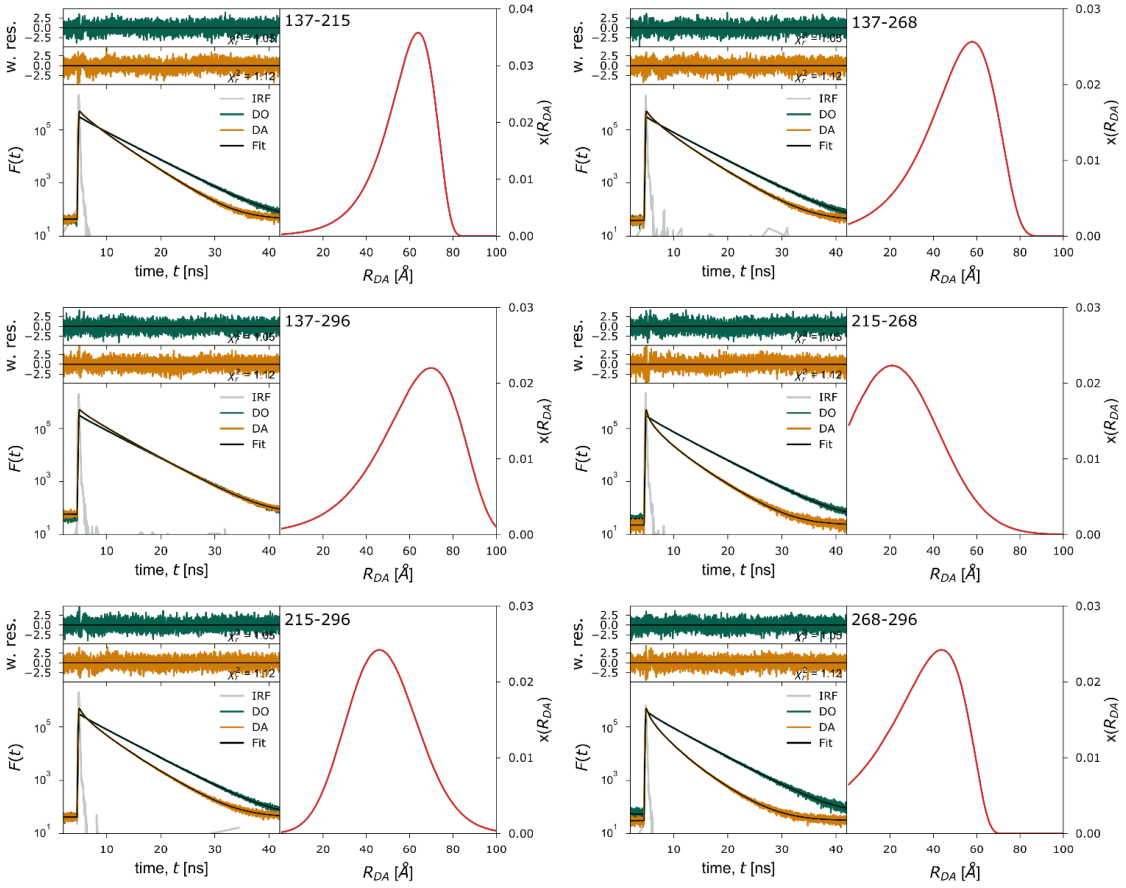


Figure S3. **TCSPC analysis of donor decay in the presence of acceptor for apoLif.** Fluorescence decays of donor-only reference sample, DO, were modeled with 3-exponential lifetime decay, and fluorescence decays of the donor in the presence of the acceptor, DA, were modeled with single Gaussian distribution for Lif (eqs. S6-8). Donor-acceptor distance distributions, $x(R_{DA})_k$, report consistently for all variants a broad distribution for *apo* Lif. Fit results are given in **Table S3**.

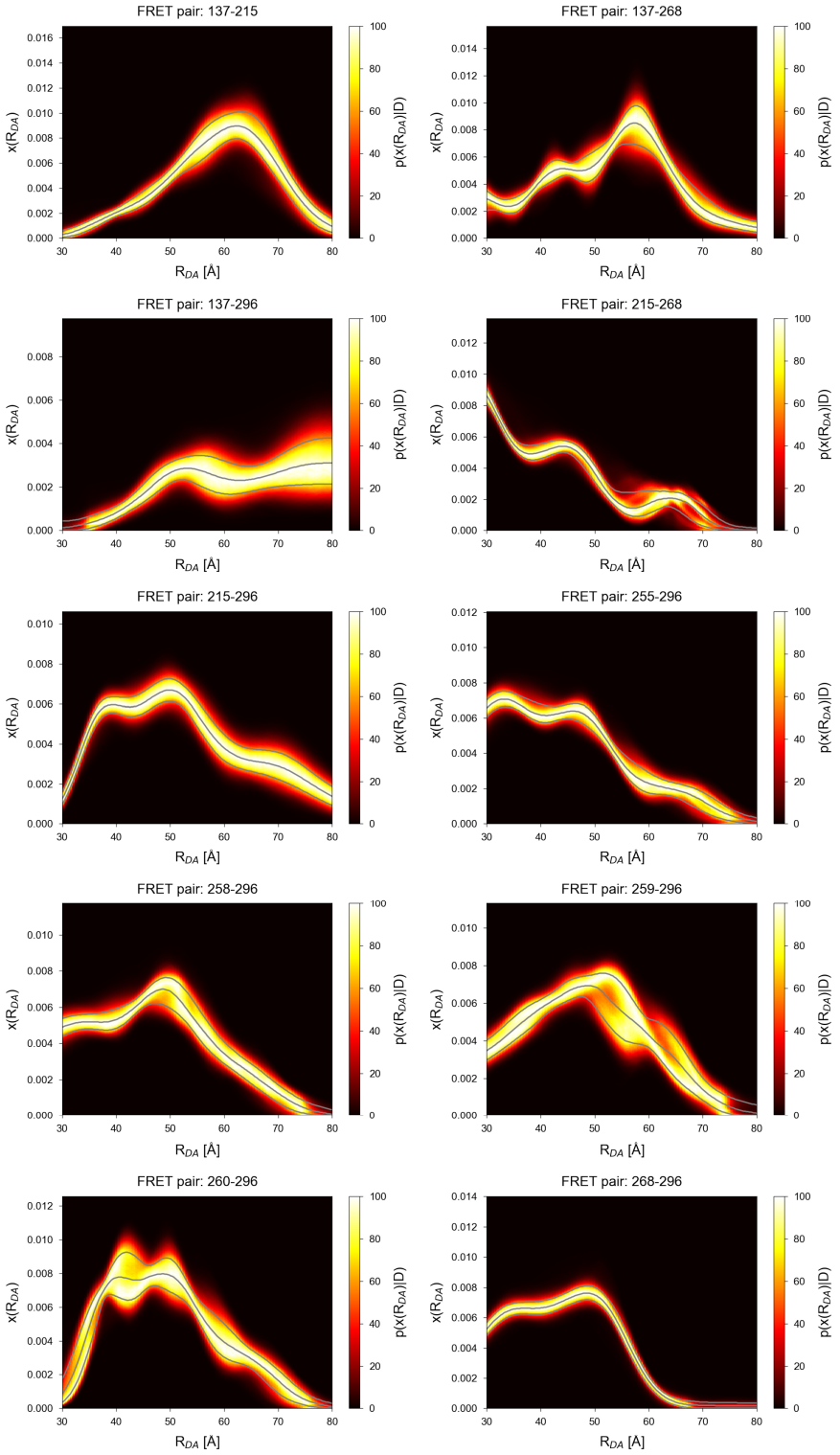


Figure S4. **R_{DA} distributions $x(R_{DA})$ for the apo Lif labeling variants obtained by ucFRET.**

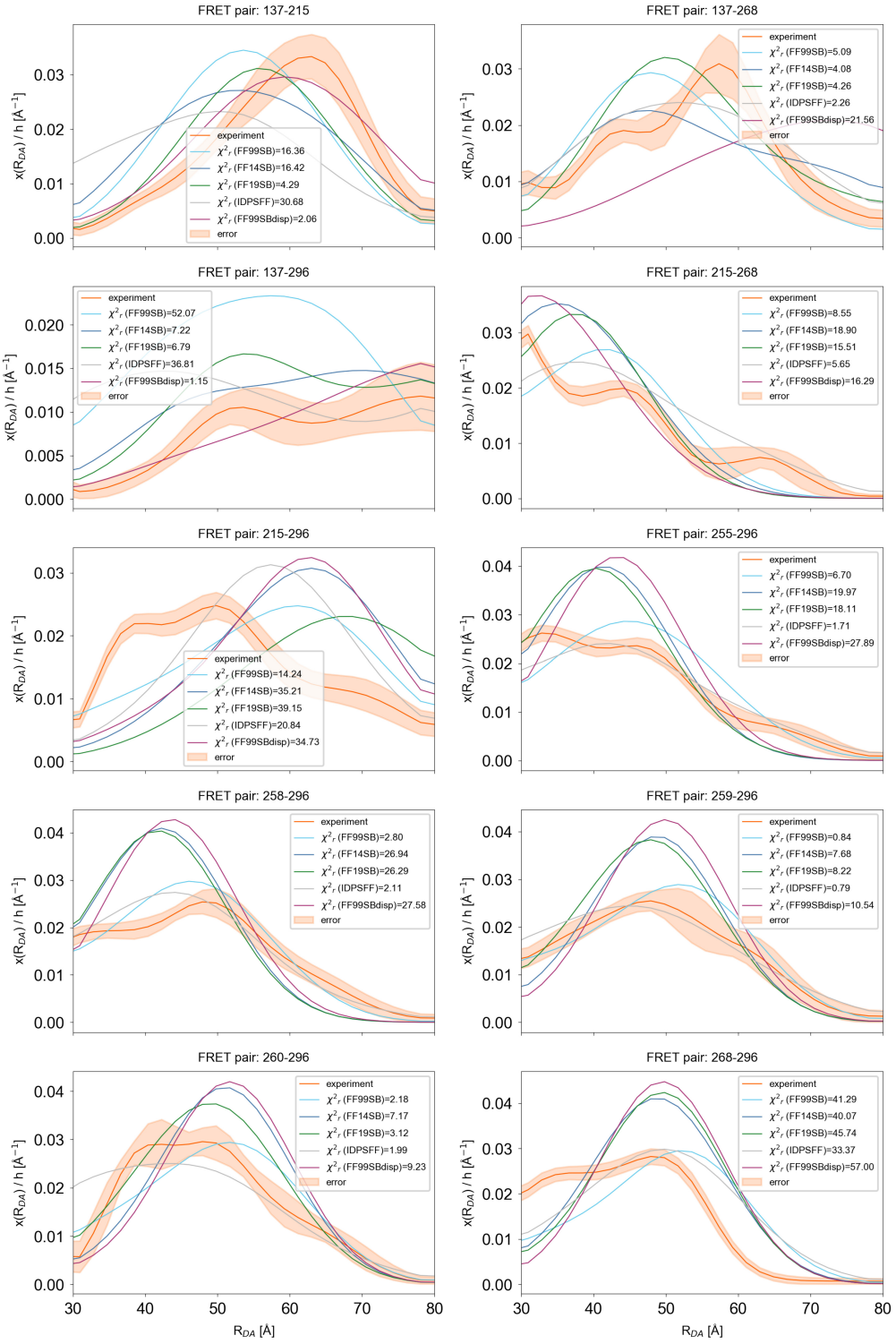


Figure S5. **Prior x(R$_{DA}$) distributions observed in MD for all FRET pairs.** Different FFs are color coded as indicated by the legend.

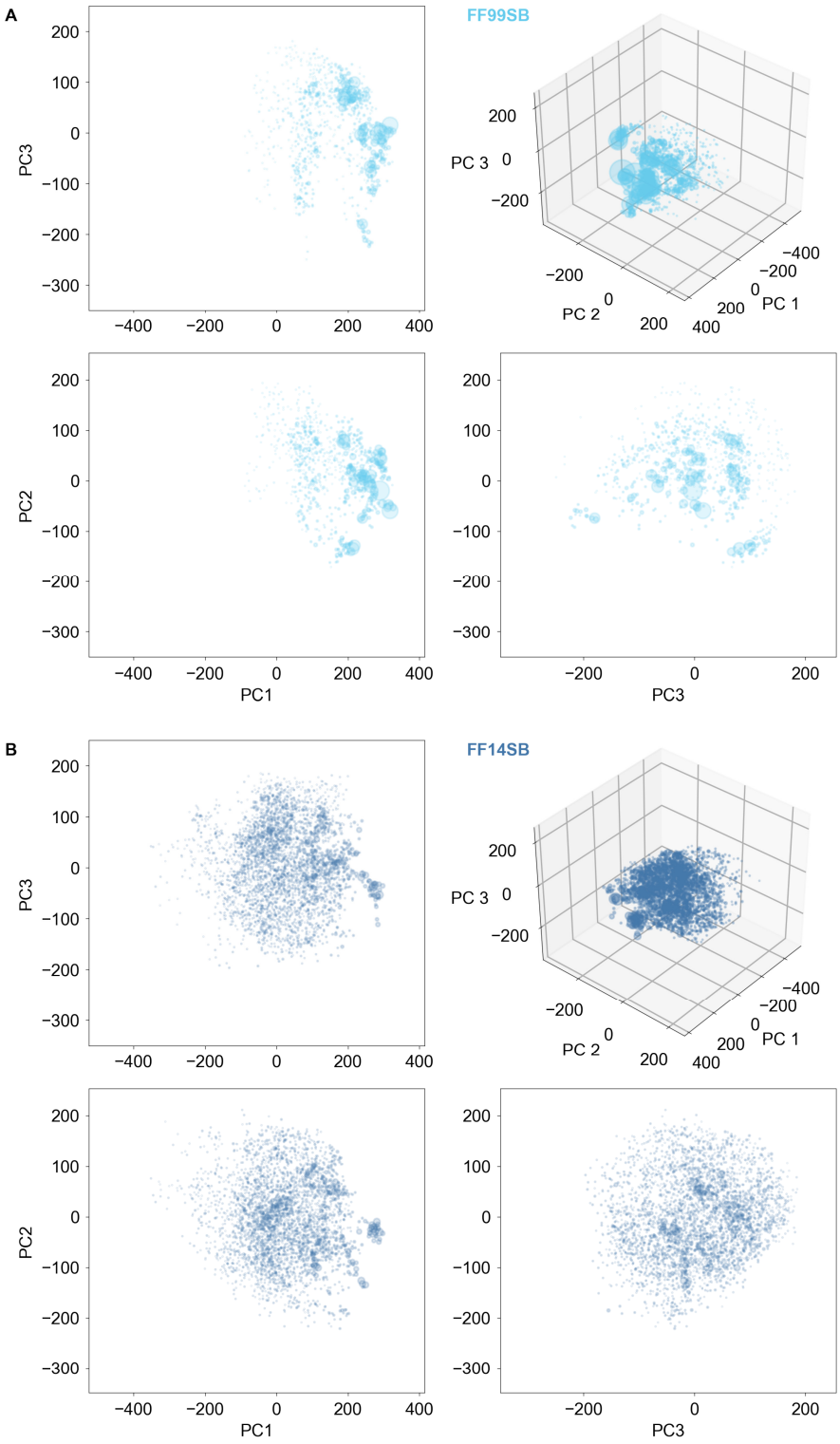


Figure S6: **continued**

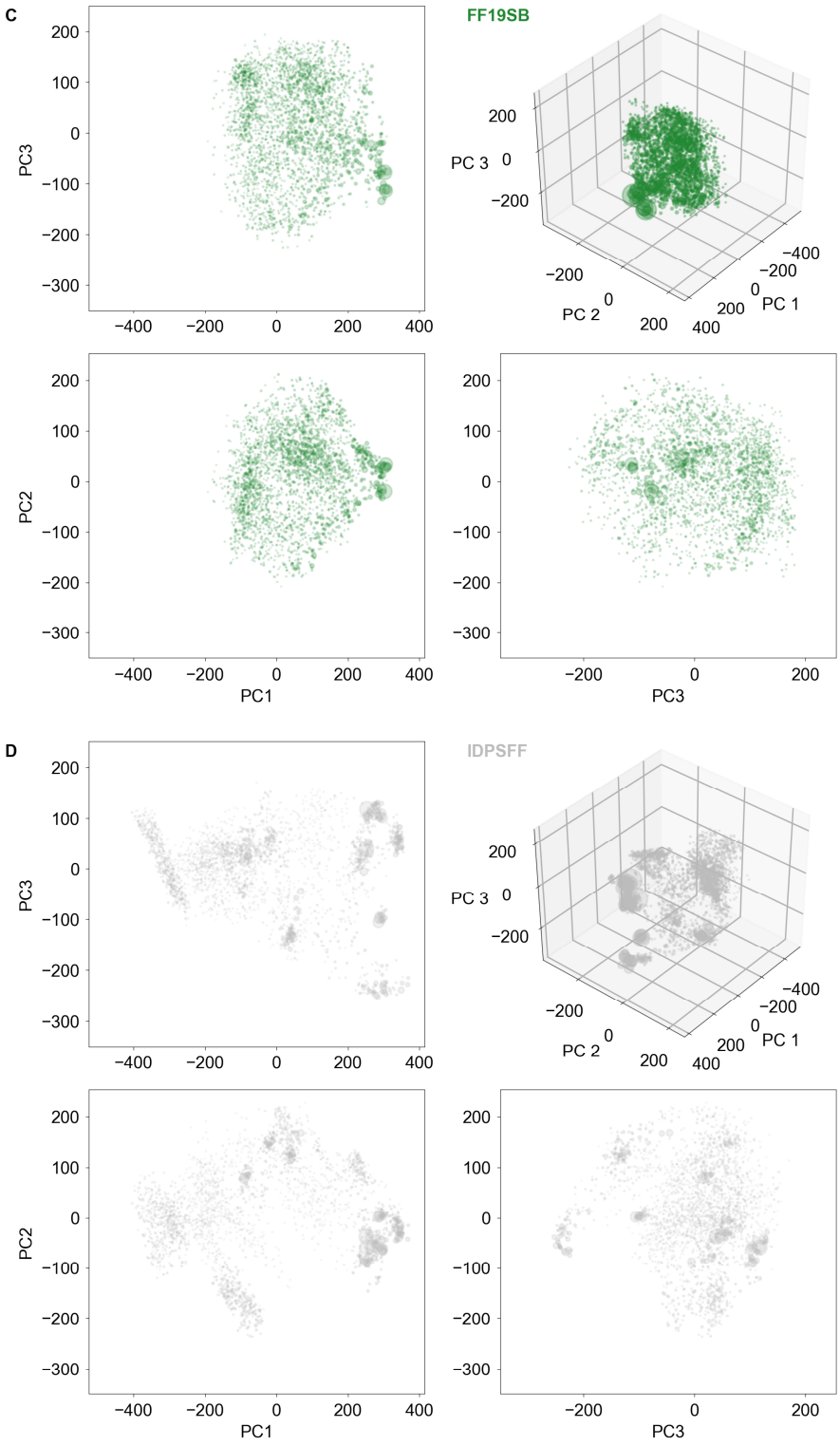


Figure S6: **continued**

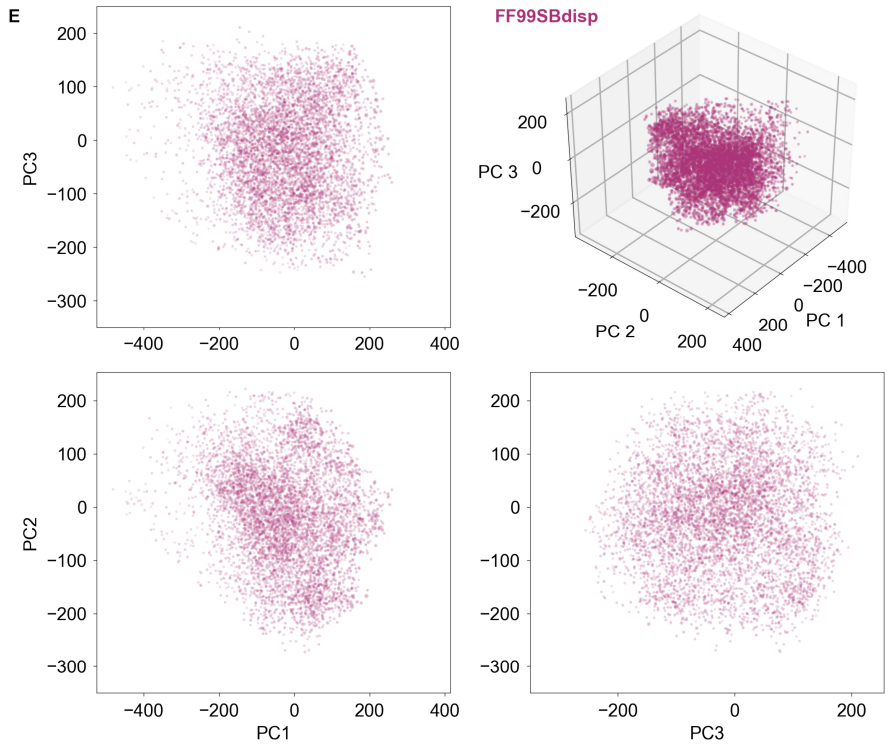


Figure S6. **Principal component analysis of prior conformational ensembles. A-E** Projection of C_α atom positions of cluster representatives for a given FF onto the first three principal components (PC1, PC2, and PC3) obtained from a PCA. The PCA was performed on the pooled cluster representatives of all FFs. The marker size for each representative structure scales linearly with the cluster weight.

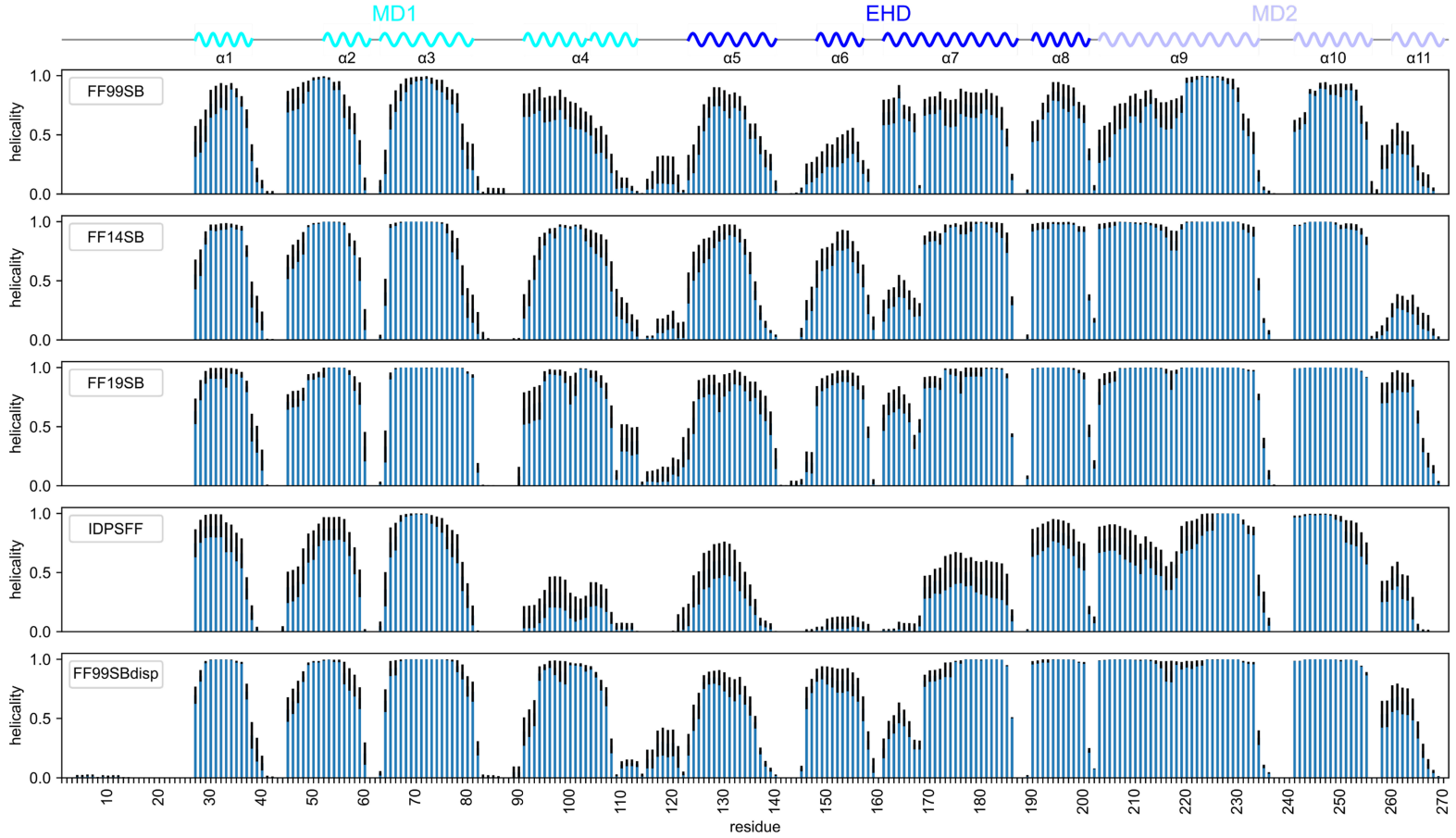


Figure S7. **Helicality of Lif throughout 10x1 μs MD simulation.** Blue bars denote the mean value per residue, black bars indicate the standard deviation over the ten replicas.

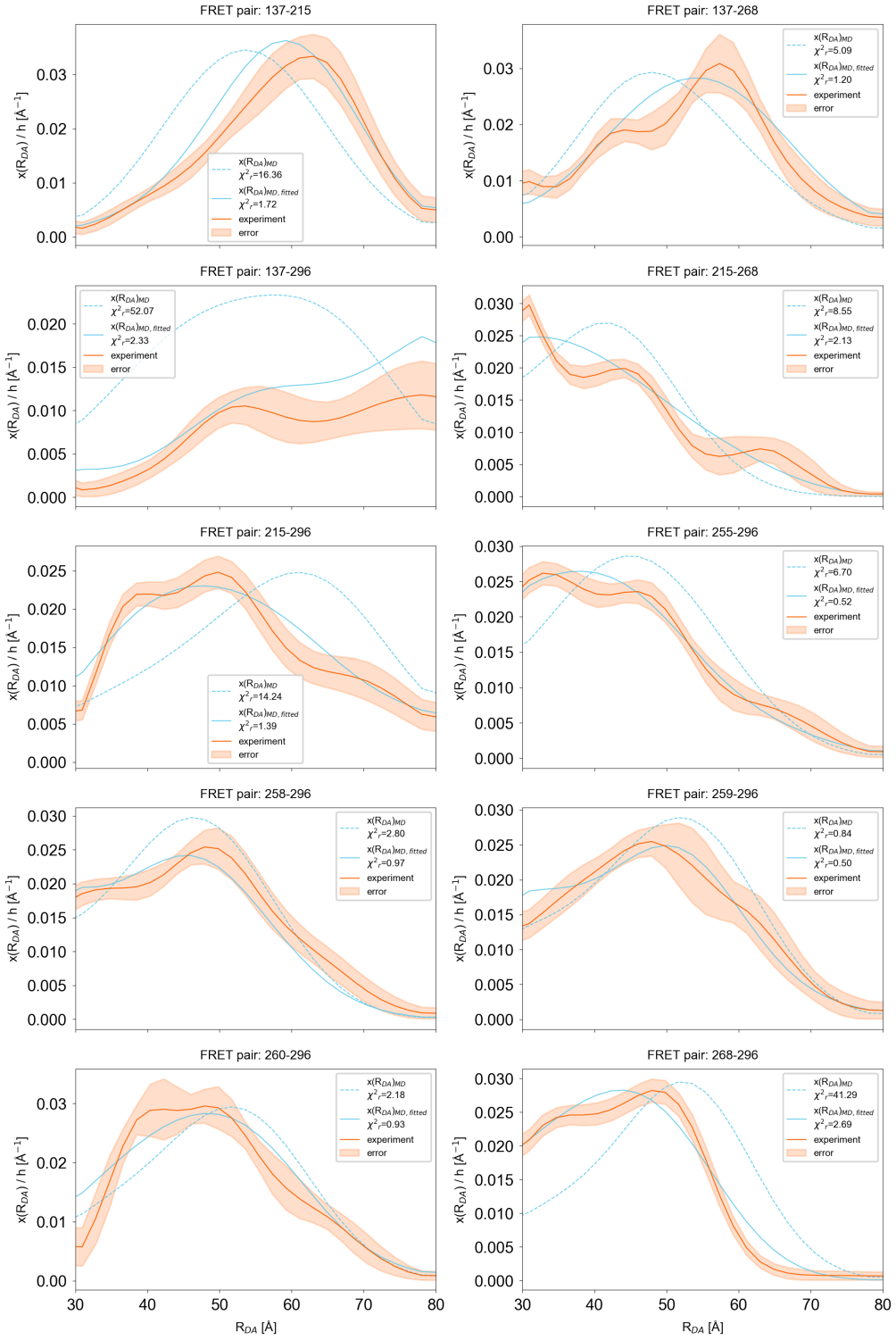


Figure S8. **MEM cluster reweighting for the FF99SB MD ensemble.**

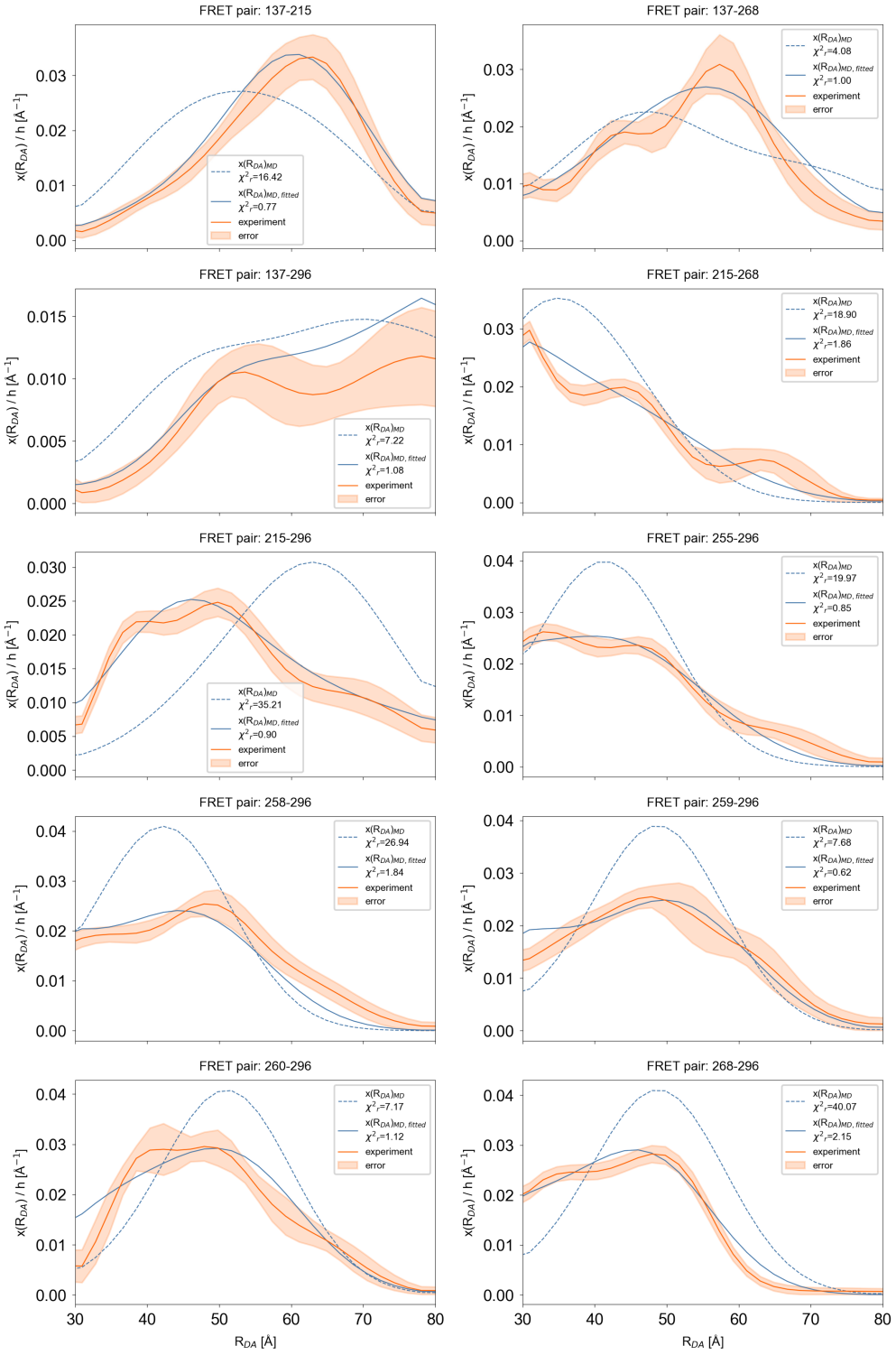


Figure S9. **MEM cluster reweighting for the FF14SB MD ensemble.**

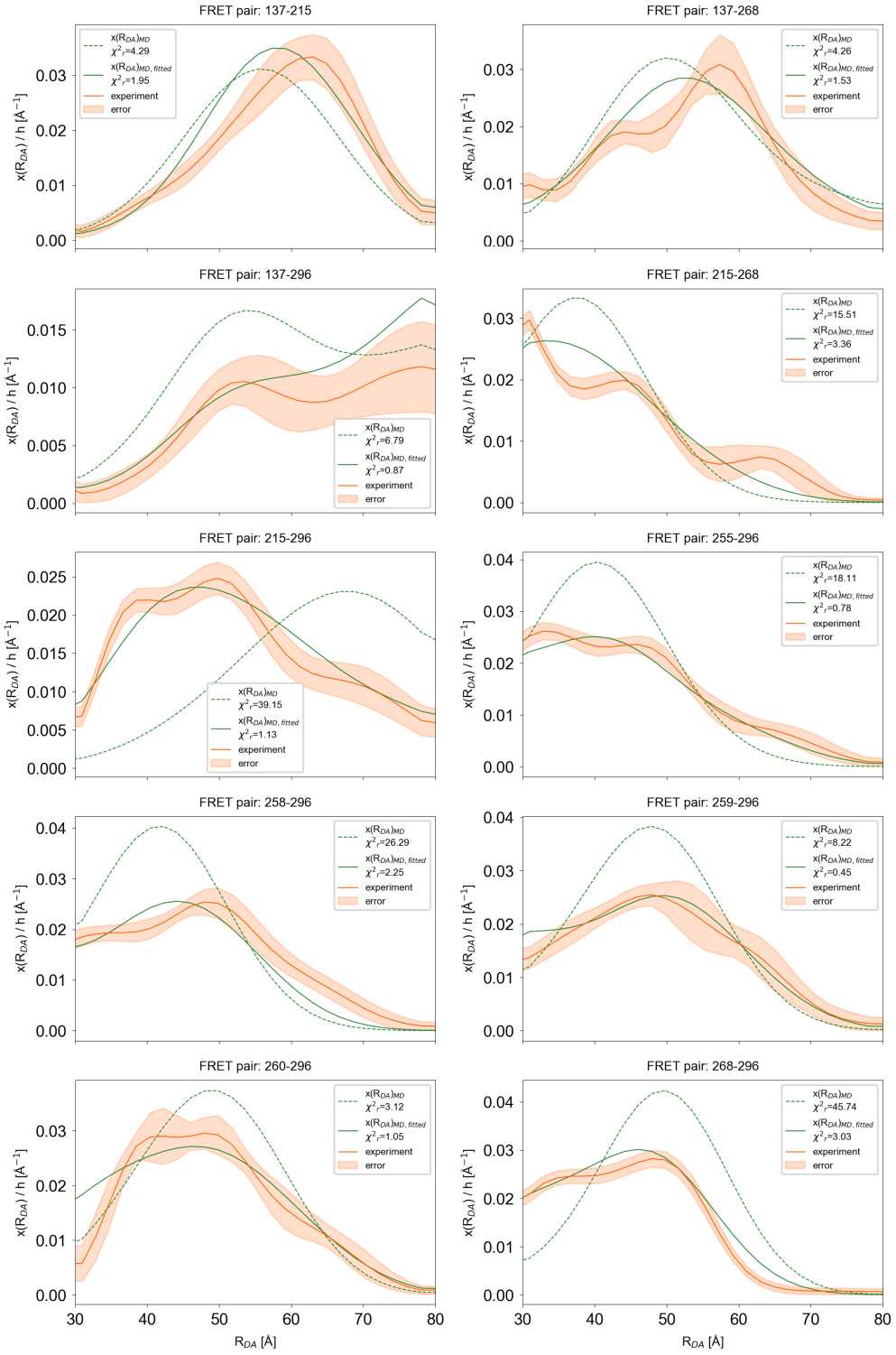


Figure S10. **MEM cluster reweighting for the FF19SB MD ensemble.**

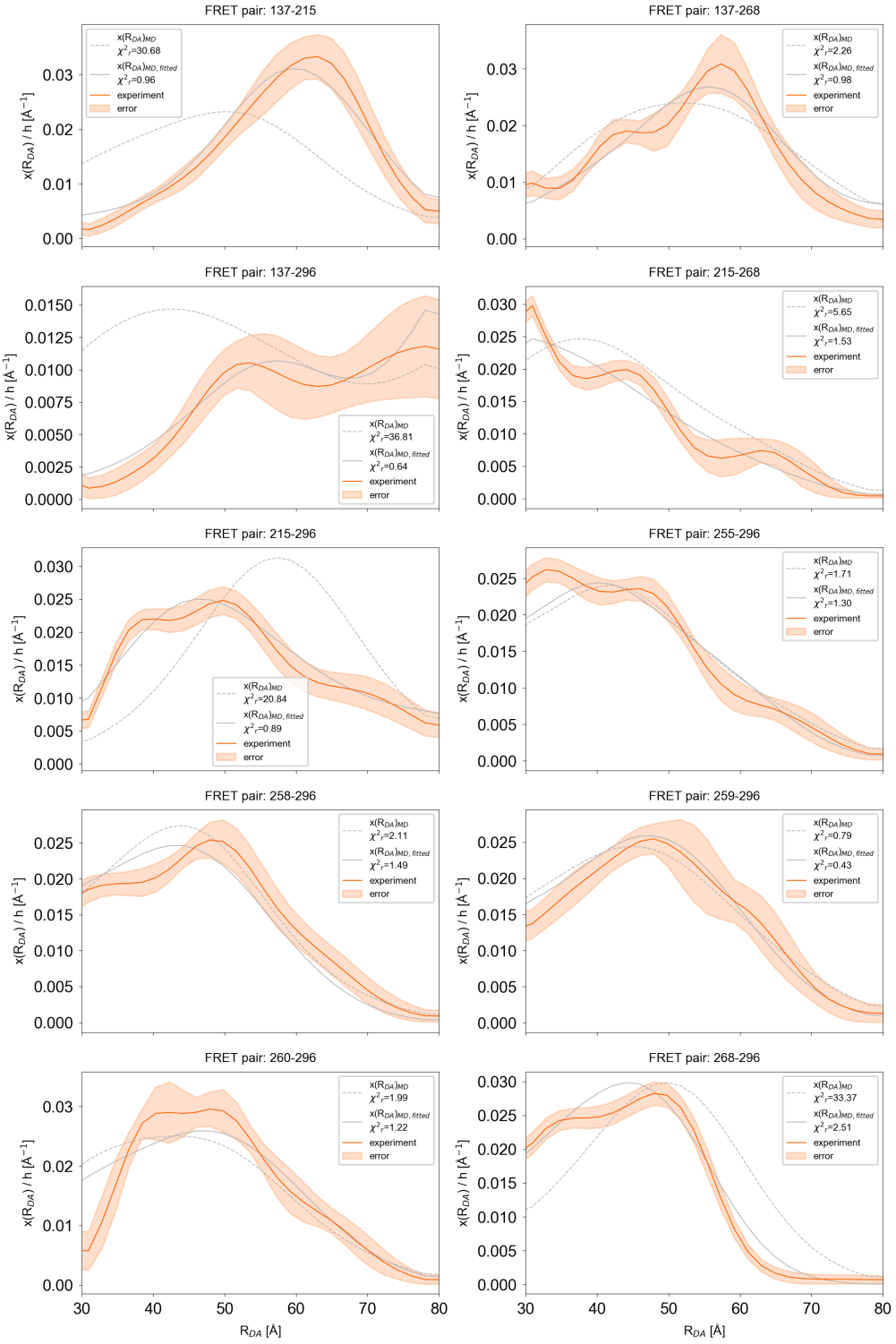


Figure S11. **MEM cluster reweighting for the IDPSFF MD ensemble.**

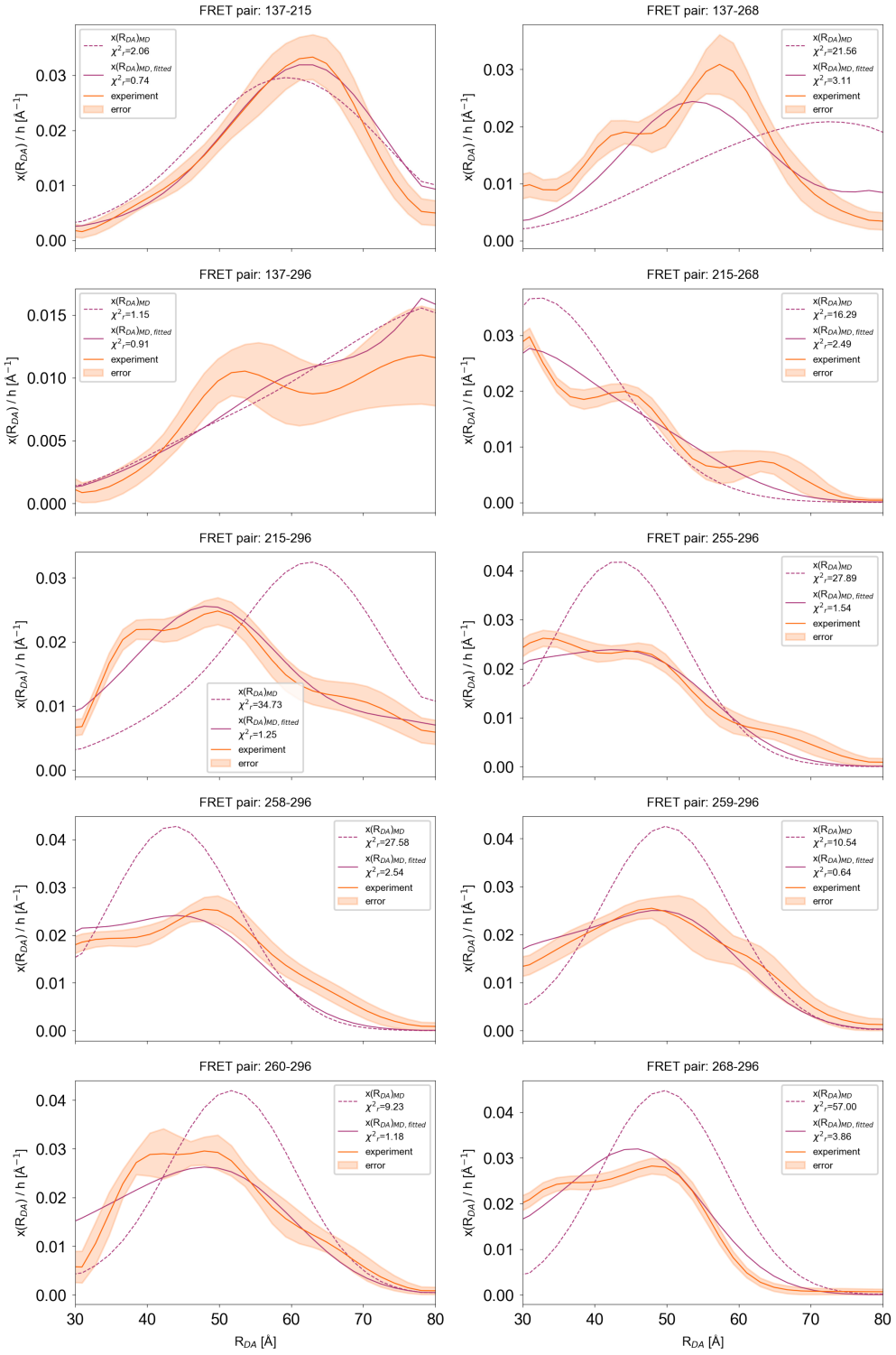


Figure S12. **MEM cluster reweighting for the FF99SBdisp MD ensemble.**

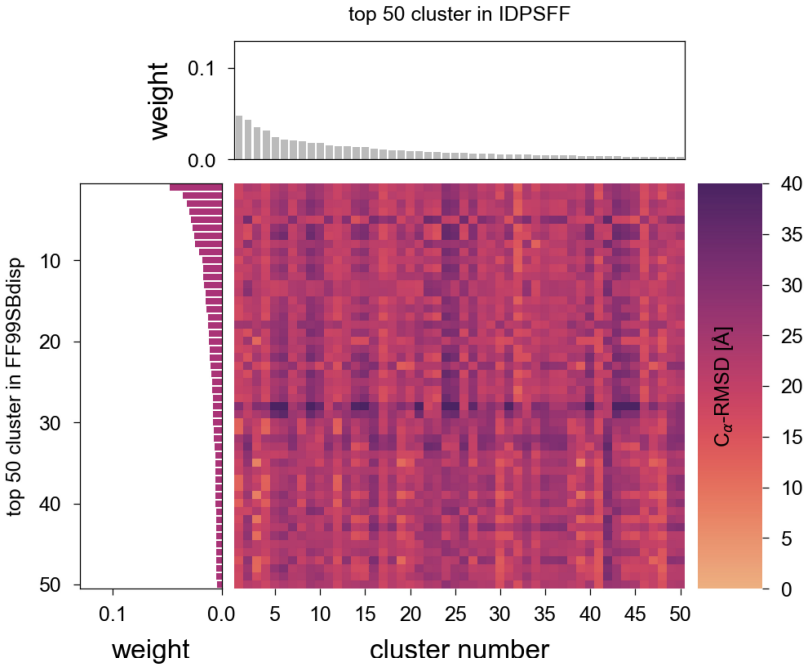


Figure S13. **C_α-RMSD matrix for the top 50 cluster representatives of IDPSFF and FF99SBdisp after reweighting using an entropy threshold of S = -3.**

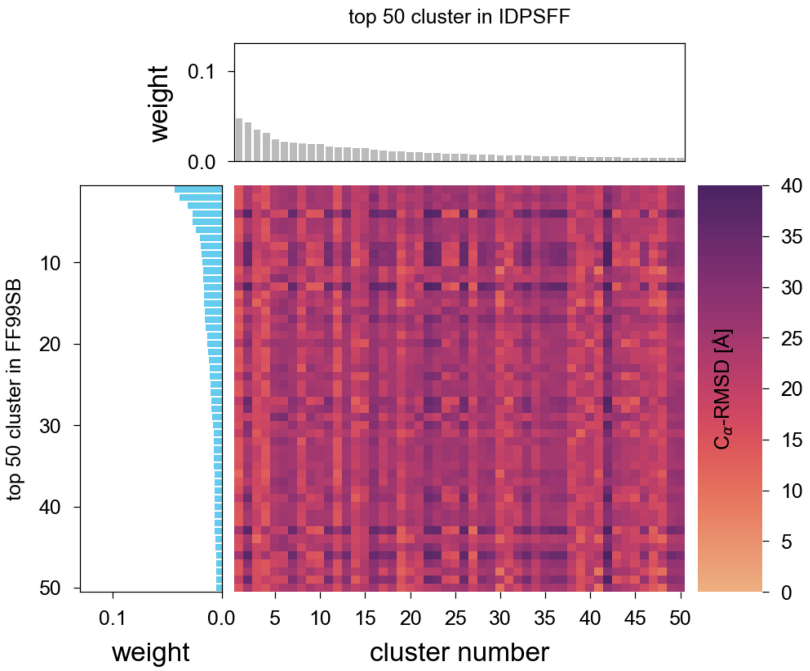


Figure S14. **Cα-RMSD matrix for the top 50 cluster representatives of IDPSFF and FF99SB after reweighting using an entropy threshold of S = -3.**

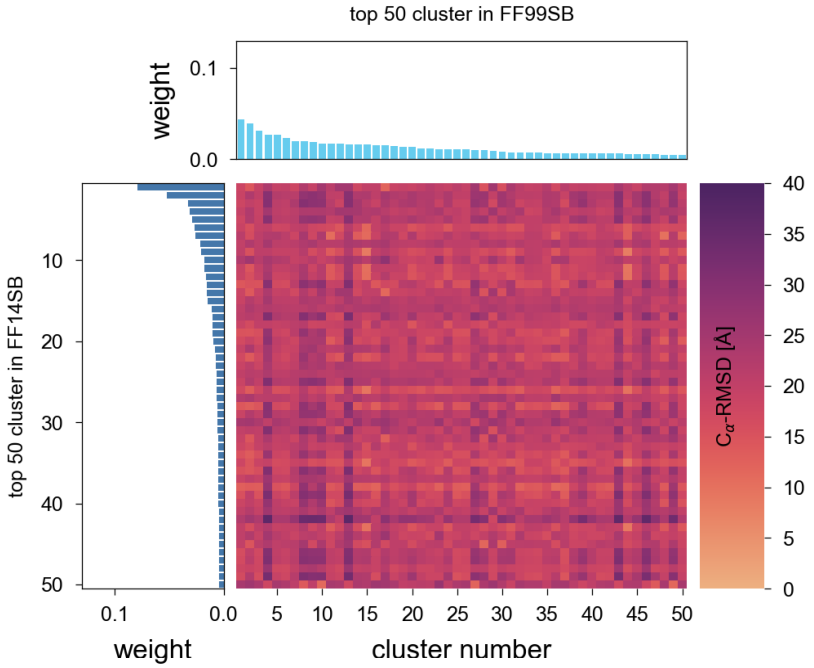


Figure S15. **Cα-RMSD matrix for the top 50 cluster representatives of FF99SB and FF14SB after reweighting using an entropy threshold of S = -3.**

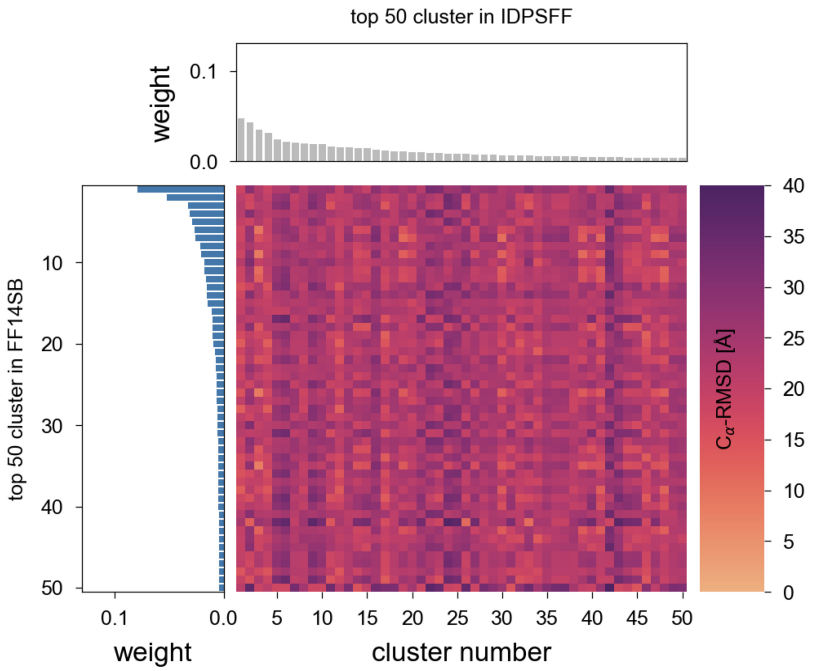


Figure S16. **Cα-RMSD matrix for the top 50 cluster representatives of IDPSFF and FF14SB after reweighting using an entropy threshold of S = -3.**

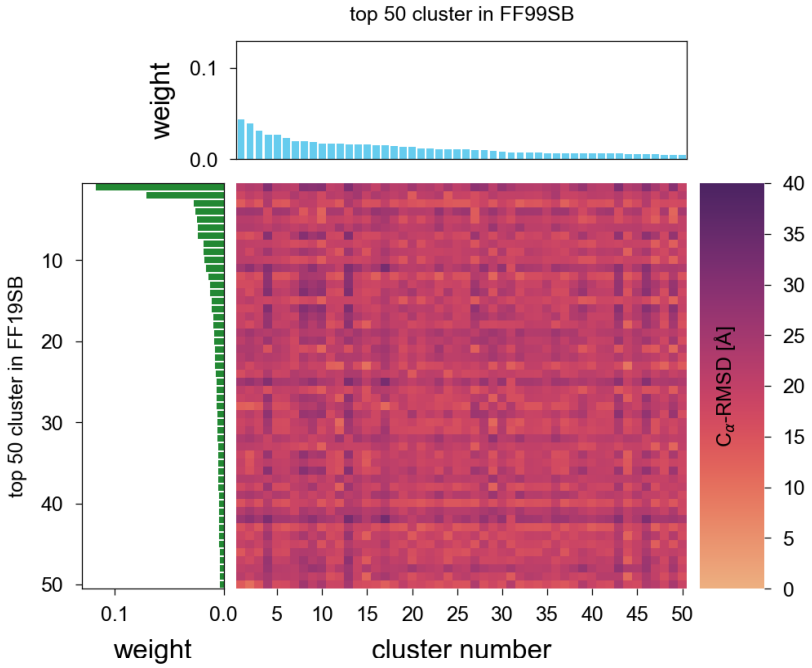


Figure S17. **Cα-RMSD matrix for the top 50 cluster representatives of FF99SB and FF19SB after reweighting using an entropy threshold of S = -3.**

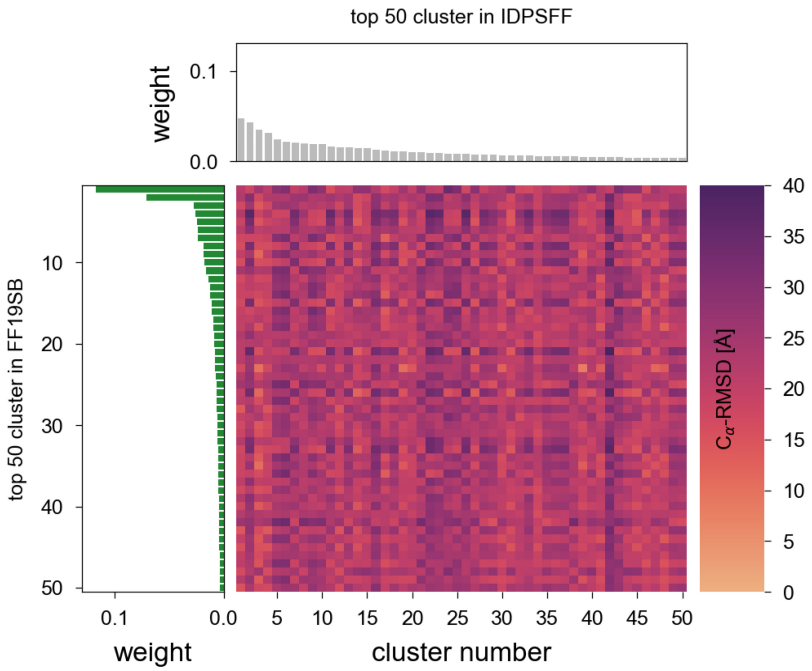


Figure S18. **Cα-RMSD matrix for the top 50 cluster representatives of IDPSFF and FF19SB after reweighting using an entropy threshold of S = -3.**

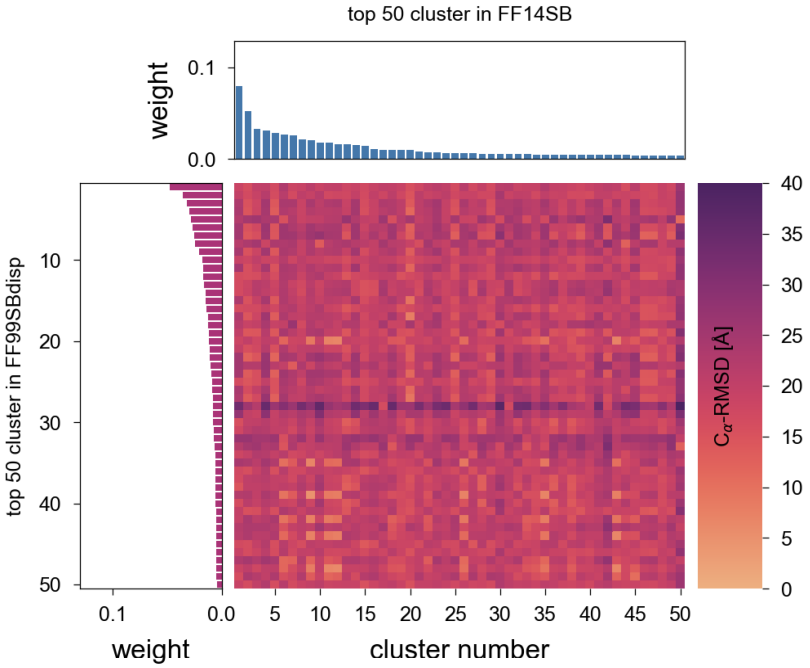


Figure S19. **Cα-RMSD matrix for the top 50 cluster representatives of FF14SB and FF99SBdisp after reweighting using an entropy threshold of S = -3.**

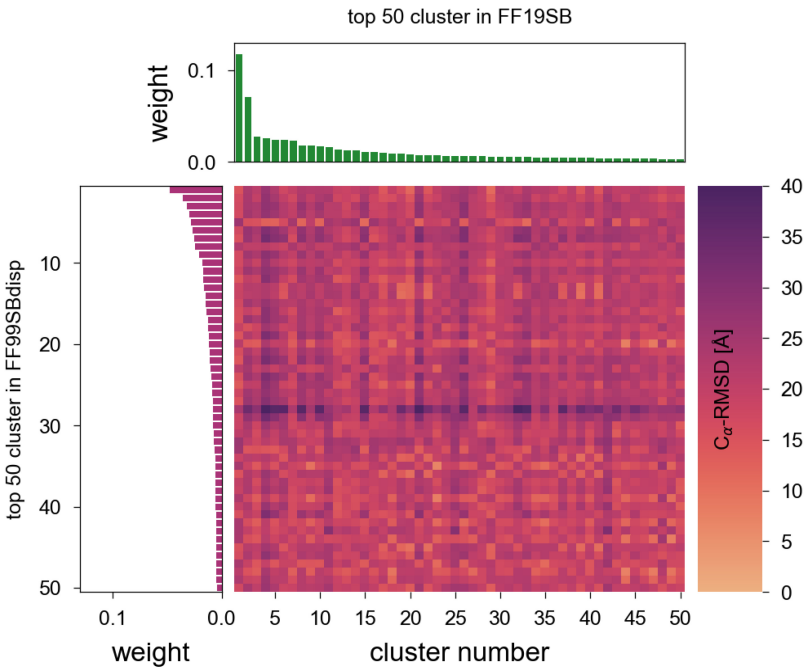


Figure S20. **Cα-RMSD matrix for the top 50 cluster representatives of FF19SB and FF99SBdisp after reweighting using an entropy threshold of S = -3.**

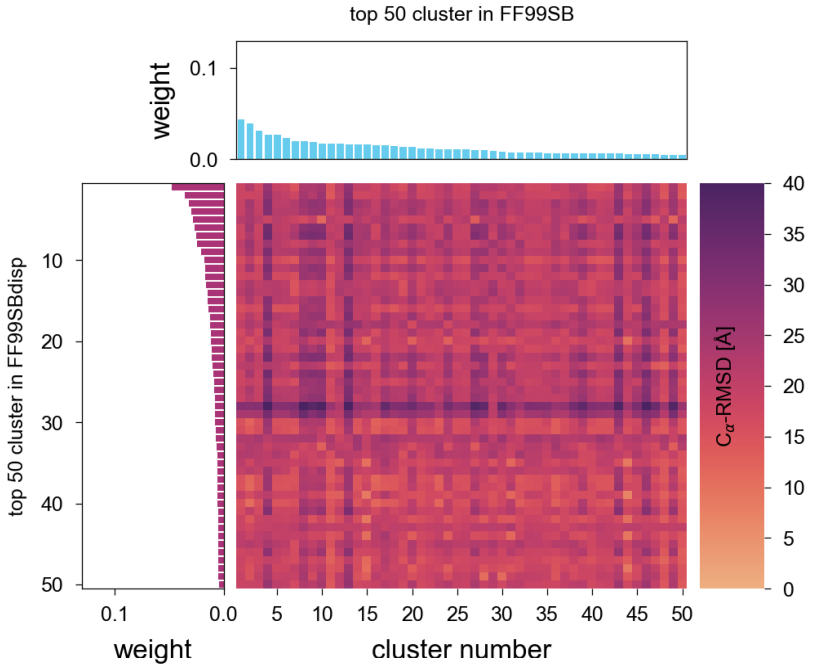


Figure S21. **Cα-RMSD matrix for the top 50 cluster representatives of FF99SB and FF99SBdisp after reweighting using an entropy threshold of S = -3.**

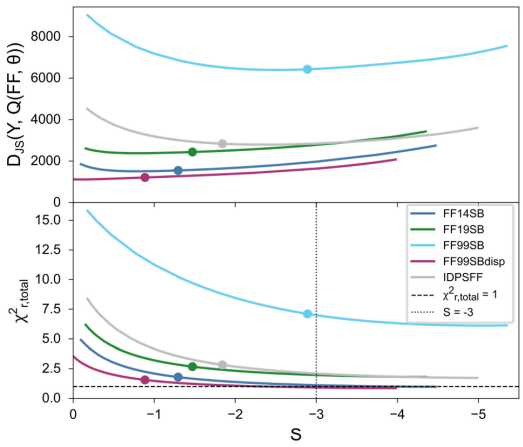


Figure S22. **MEM scoring and validation of reweighted ensembles using synthetic data when Y(all).** Discrepancy, χ_r^2, between the model distance distributions of the posterior conformational ensembles recovered by MEM and experimental distance distributions versus the entropy, S (bottom). To validate reweighted ensembles, the Jensen-Shannon divergence is computed between posterior distograms and ground truth, Y (top). Circles mark θ_c points that balance the experimental and the prior information based on the L-shaped curve criterion.

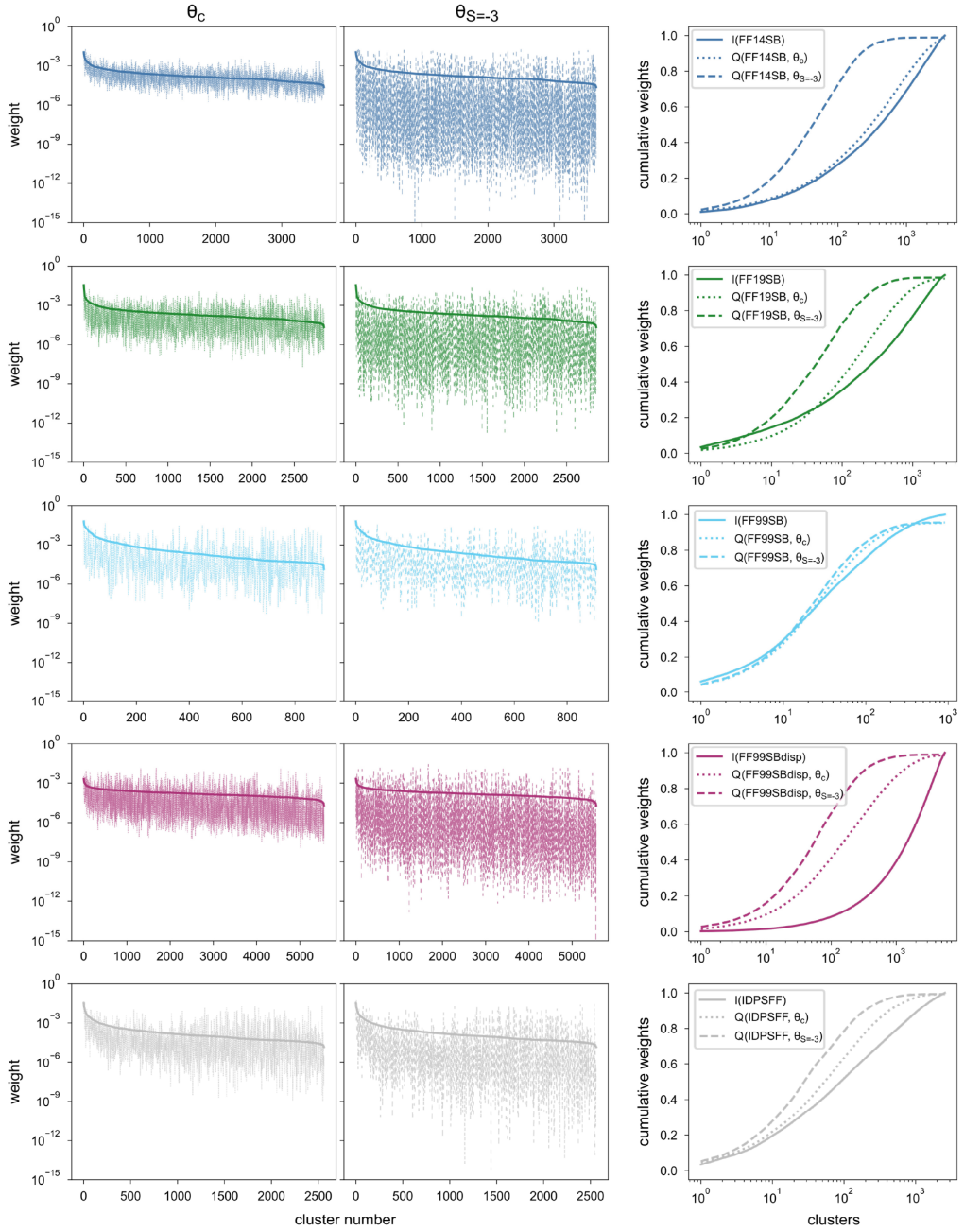


Figure S23. **Individual and cumulative weights of prior and posterior ensembles (at θ_C and $\theta_{S=-3}$) from a synthetic experiment when the target ensemble is Y(FF14SB).**

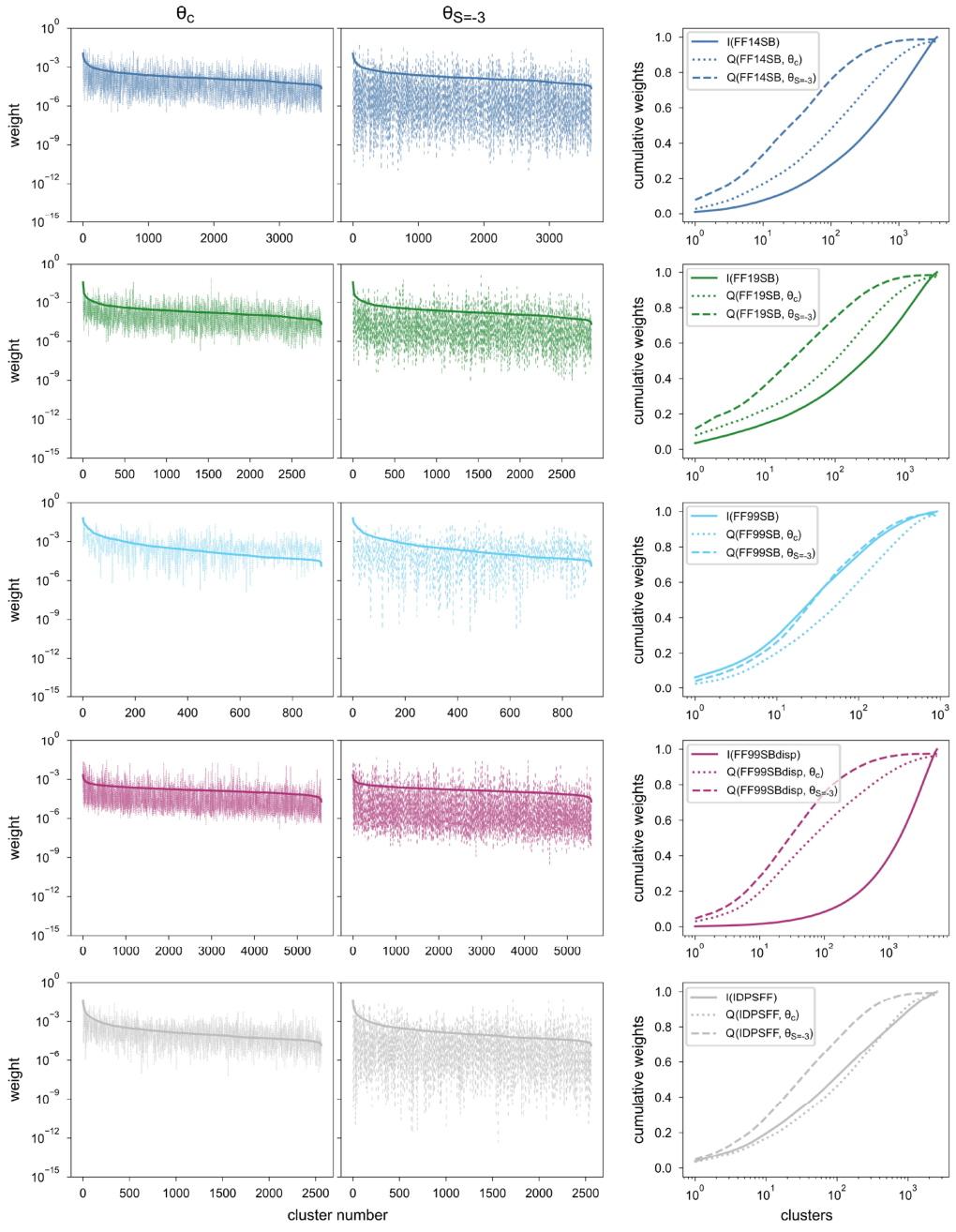


Figure S24. **Individual and cumulative weights of prior and posterior ensembles (at θ_C and $\theta_{S=-3}$) for experimental data.**

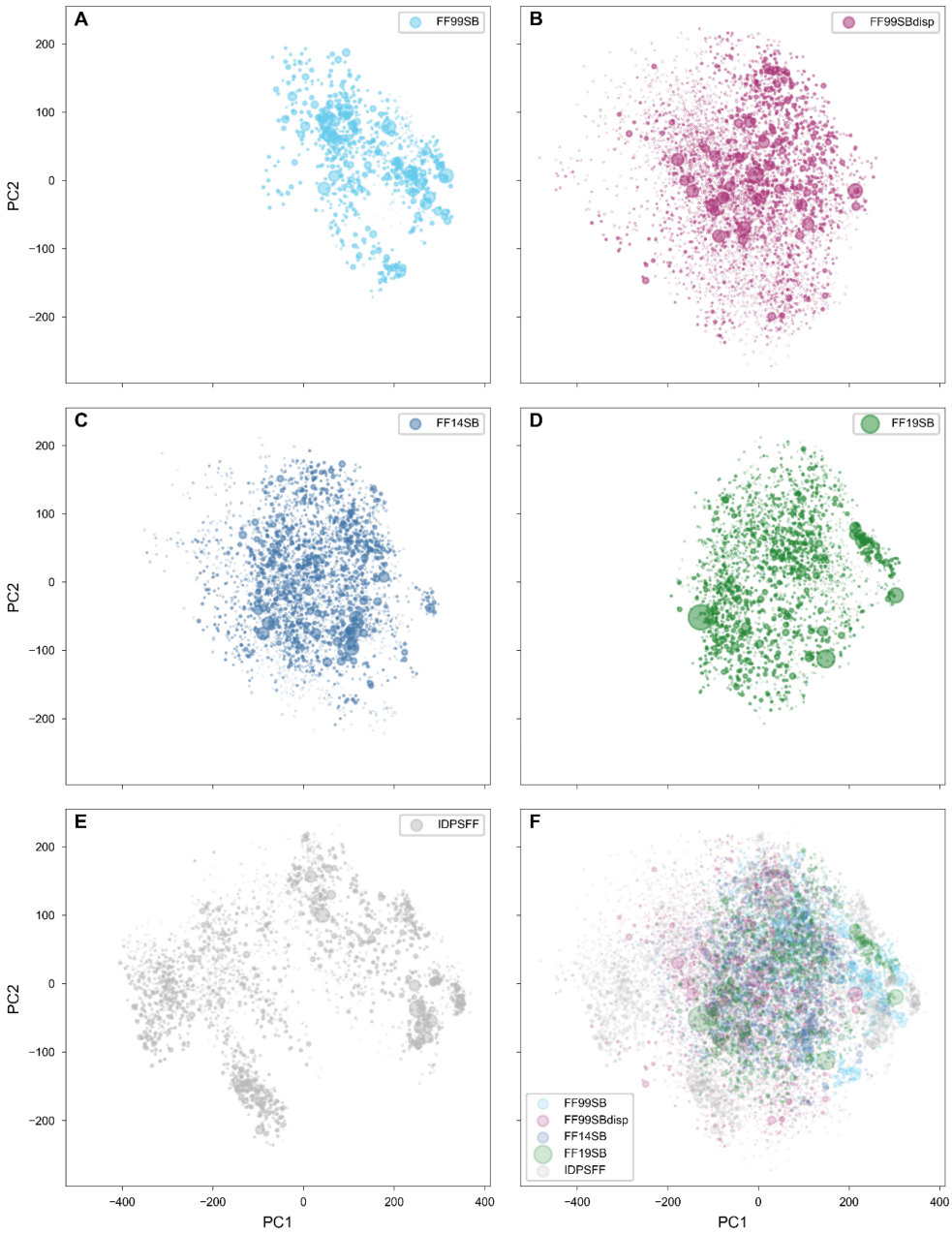


Figure S25. **PCA of optimized ensembles at at θ_C using experimental data.** The marker size for each representative structure scales linearly with the cluster weight.

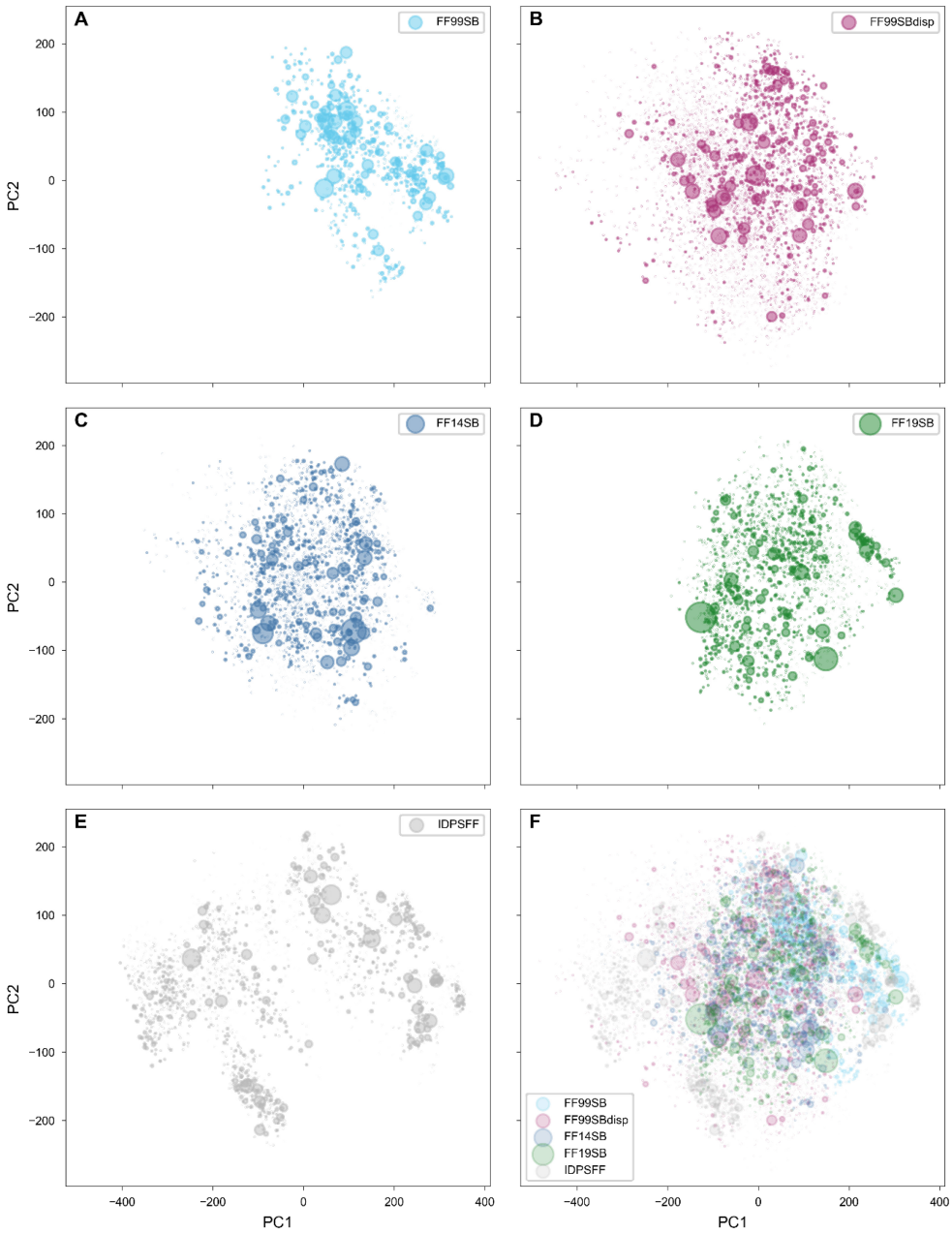


Figure S26. **PCA of optimized ensembles at $\theta_{S=-3}$ using experimental data.** The marker size for each representative structure scales linearly with the cluster weight.

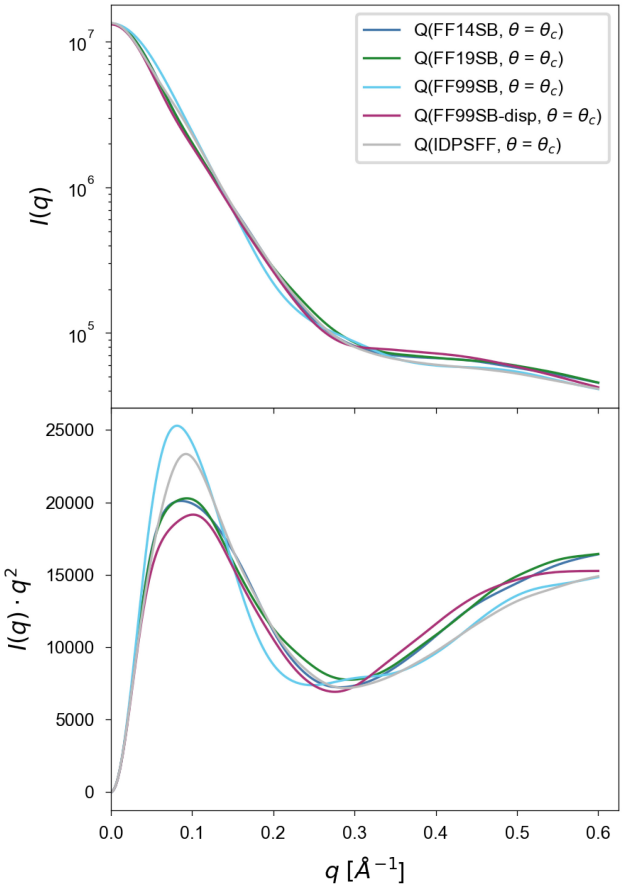


Figure S27. **SAXS profiles of posterior ensembles. Top:** Scattering profiles of posterior ensembles reveal smoothed features, characteristic of multidomain proteins with flexible linkers. **Bottom:** A commonly used transformation of the scattering profile is displayed, known as Kratky plot, which gives additional information about the compactness (folding) of the macromolecule.

Supplemental References

1. Li, M. Z.; Elledge, S. J., Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nat. Methods* **2007,** *4* (3), 251-256.
2. Viegas, A.; Dollinger, P.; Verma, N.; Kubiak, J.; Viennet, T.; Seidel, C. A. M.; Gohlke, H.; Etzkorn, M.; Kovacic, F.; Jaeger, K.-E., Structural and dynamic insights revealing how lipase binding domain MD1 of *Pseudomonas aeruginosa* foldase affects lipase activation. *Sci. Rep.* **2020,** *10* (1), 3578.
3. Hausmann, S.; Wilhelm, S.; Jaeger, K.-E.; Rosenau, F., Mutations towards enantioselectivity adversely affect secretion of Pseudomonas aeruginosa lipase. *FEMS Microbiol. Lett.* **2008,** *282* (1), 65-72.
4. Tsytlonok, M.; Sanabria, H.; Wang, Y.; Felekyan, S.; Hemmen, K.; Phillips, A. H.; Yun, M.-K.; Waddell, M. B.; Park, C.-G.; Vaithiyalingam, S.; Iconaru, L.; White, S. W.; Tompa, P.; Seidel, C. A. M.; Kriwacki, R., Dynamic anticipation by Cdk2/Cyclin A-bound p27 mediates signal integration in cell cycle regulation. *Nat. Commun.* **2019,** *10* (1), 1676.
5. Sisamakis, E.; Valeri, A.; Kalinin, S.; Rothwell, P. J.; Seidel, C. A. M., Chapter 18 - Accurate Single-Molecule FRET Studies Using Multiparameter Fluorescence Detection. In *Methods Enzymol.*, Walter, N. G., Ed. Academic Press: 2010; Vol. 475, pp 455-514.
6. 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 Enzymology*, Tetin, S. Y., Ed. Academic Press: 2013; Vol. 519, pp 39-85.
7. Peulen, T.-O.; Opanasyuk, O.; Seidel, C. A. M., Combining Graphical and Analytical Methods with Molecular Simulations To Analyze Time-Resolved FRET Measurements of Labeled Macromolecules Accurately. *J. Phys. Chem. B* **2017,** *121* (35), 8211-8241.
8. Sanabria, H.; Rodnin, D.; Hemmen, K.; Peulen, T.-O.; Felekyan, S.; Fleissner, M. R.; Dimura, M.; Koberling, F.; Kühnemuth, R.; Hubbell, W.; Gohlke, H.; Seidel, C. A. M., Resolving dynamics and function of transient states in single enzyme molecules. *Nat. Commun.* **2020,** *11* (1), 1231.
9. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C., Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics* **2006,** *65* (3), 712-725.
10. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C., ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **2015,** *11* (8), 3696-3713.
11. Tian, C.; Kasavajhala, K.; Belfon, K. A. A.; Raguette, L.; Huang, H.; Migues, A. N.; Bickel, J.; Wang, Y.; Pincay, J.; Wu, Q.; Simmerling, C., ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J. Chem. Theory Comput.* **2020,** *16* (1), 528-552.
12. Song, D.; Luo, R.; Chen, H.-F., The IDP-Specific Force Field ff14IDPSFF Improves the Conformer Sampling of Intrinsically Disordered Proteins. *J. Chem. Inf. Model.* **2017,** *57* (5), 1166-1178.
13. Robustelli, P.; Piana, S.; Shaw, D. E., Developing a molecular dynamics force field for both folded and disordered protein states. *PNAS* **2018,** *115* (21), E4758-E4766.
14. Case, D. A.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; Cheatham III, T. E.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Ghoreishi, D.; Gilson, M. K.; Gohlke, H.; Goetz, A.

W.; Greene, D.; Harris, R.; Homeyer, N.; Izadi, S.; Kovalenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Mermelstein, D. J.; Merz, K. M.; Miao, Y.; Monard, G.; Nguyen, C.; Nguyen, H.; Omelyan, I.; Onufriev, A.; Pan, F.; Qi, R.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Smith, J.; Salomon-Ferrer, R.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiao, L.; York , D. M.; Kollman, P. A. *AMBER*, University of California: San Francisco, 2021.

15. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C., GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005,** *26* (16), 1701-1718.
16. Shirts, M. R.; Klein, C.; Swails, J. M.; Yin, J.; Gilson, M. K.; Mobley, D. L.; Case, D. A.; Zhong, E. D., Lessons learned from comparing molecular dynamics engines on the SAMPL5 dataset. *J. Comput. Aided Mol. Des.* **2017,** *31* (1), 147-161.
17. Dimura, M.; Peulen, T.-O.; Sanabria, H.; Rodnin, D.; Hemmen, K.; Hanke, C. A.; Seidel, C. A. M.; Gohlke, H., Automated and optimally FRET-assisted structural modeling. *Nat. Commun.* **2020,** *11* (1), 5394.
18. Schneidman-Duhovny, D.; Hammel, M.; Tainer, John A.; Sali, A., Accurate SAXS Profile Computation and its Assessment by Contrast Variation Experiments. *Biophys. J.* **2013,** *105* (4), 962-974.
19. Russel, D.; Lasker, K.; Webb, B.; Velázquez-Muriel, J.; Tjioe, E.; Schneidman-Duhovny, D.; Peterson, B.; Sali, A., Putting the Pieces Together: Integrative Modeling Platform Software for Structure Determination of Macromolecular Assemblies. *PLoS Biol.* **2012,** *10* (1), e1001244.
20. Debye, P., Zerstreuung von Röntgenstrahlen. *Annalen der Physik* **1915,** *351* (6), 809-823.
21. Brosey, C. A.; Tainer, J. A., Evolving SAXS versatility: solution X-ray scattering for macromolecular architecture, functional landscapes, and integrative structural biology. *Curr. Opin. Struct. Biol.* **2019,** *58*, 197-213.

Supplement C. Increasing the accuracy of single-molecule FRET experiments-robust and unbiased estimation of correction parameters

Increasing the accuracy of single-molecule FRET experiments - robust and unbiased estimation of correction parameters

Milana Popara[1], Anders Barth[1,a], Julian Folz[1], Suren Felekyan[1], Oleg Opansyuk[1], Claus A. M. Seidel[1,*]

[1]Institut for Molecular Physical Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

[a]Present address: Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Delft, Netherlands

[*]Author to whom correspondence should be addressed:
Claus A. M. Seidel, Phone: (+49) 211 81 15881, E-Mail: cseidel@hhu.de

Keywords: PIE, smFRET, correction parameters, calibration

Abstract

Single-molecule FRET experiments are a prime example of ever-growing scientific method, whose impact can be credited to the ability to provide site-specific structural information, with minimal spatial and temporal averaging, and temporal resolution spanning across twelve decades in time (sub-*ns* to thousands of *s*). Despite wide-spread use of smFRET and new measurement modalities arising, few long-standing challenges of this scientific method are yet to be overcome. In a typical single-molecule FRET experiment, donor and acceptor raw signal are distorted due to spectral overlap, and using a set of correction parameters they can be converted into accurate quantities, such as FRET efficiency and stoichiometry. After the background subtraction, further four levels of correction to raw data are needed, namely α, β, γ and δ, which are known as correction parameters. Major advancement in determination of correction parameters happened with introduction of Pulsed Interleaved Excitation (PIE) scheme, but ever since, several correction procedures were proposed, with no consensus in the community on the optimal approach. Moreover, neither of the approaches gives clear guidelines on the selection of subpopulations necessary for determination of each of four correction parameters. And while it has been recognized that the user bias in correction procedures is significant, so far there were no attempts to alleviate this issue. Driven with this challenge, and building upon the previous contributions in the community, we designed protocol for data correction, with the smallest number of selection steps to minimize the extent of user bias. For each of the correction steps provided are unambiguous instructions for

population selection and filtering, and proposed are complementary measurements for parameter cross-validation, although all necessary information can be obtained from the measurement of the sample. Robustness of proposed protocol is assessed in a blind challenge. Furthermore, we complemented data correction workflow with guidelines for setup alignment and calibration.

1. Introduction

Vast number of experimental methods with different spatial and temporal resolution are being employed to inform on structure and dynamics of biomacromolecules and their complexes. Method of Förster Resonance Energy Transfer (FRET) stands out, as it can inform on both the small- and large-scale conformational motions, conformational states and corresponding kinetical schemes[1-3]. Alone or in integrative approach with data from other sources, it was demonstrated that FRET measurements are able to tackle all levels of biomolecular complexity, from folded proteins[4], over multi-domain proteins[5] to IDPs[6, 7]. Main observable in FRET experiment is energy transfer from donor to acceptor fluorophore, i.e. FRET efficiency E, which informs about their spatial separation. There are different experimental approaches to determine FRET efficiency, such as: (**a**) by measuring the fraction of FRET sensitized acceptor fluorescence out of the total fluorescence of both donor and acceptor (*classical ratio method*) (**b**) by measuring enhanced acceptor fluorescence, as the ratio of FRET sensitized acceptor fluorescence and acceptor fluorescence due to direct acceptor excitation ($(ratio)_A$ *method*) (**c**) by measuring decrease in fluorescence quantum yield of donor, i.e. via the ratio of donor fluorescence quantum yield in the presence and absence of the acceptor ($(ratio)_D$ *method*) (**d**) via decrease of donor fluorescence lifetime in the presence of acceptor (**e**) via anisotropy change of donor and acceptor. Clegg[8] wrote a comprehensive overview on how each of the methods should be handled, and in which scenarios, considering their strengths and weaknesses. Out of the approaches listed here, *classical ratio method* is most commonly employed. In this work, we focus on a confocal measurement modality where freely diffusing molecules in solution are studied, and at single-molecule concentration (smFRET). Major advancement in smFRET studies has been made with introduction of Pulsed Interleaved Excitation (PIE), where donor and acceptor excitation light are being interleaved on a nanosecond timescale. By exciting donor and acceptor fluorophore at different points in time, it is possible to separate donor and acceptor emitted photons, despite the difficulty of their spectral overlap. For details on how a PIE microscope works, reader is referred to pioneering

work of Müller *et al*[9] and a review by Hendrix *et al*.[10] Step forward was made when PIE was combined with Multiparameter Fluorescence Detection (MFD).[11] MFD measurement modality allows simultaneous detection of several fluorescence parameters, which in turn minimizes data interpretation ambiguity. Moreover, due to both temporal and spectral separation of detected photons PIE-MFD measurement modality allows for double-labelled donor-acceptor species (DA) to be separated from single-labelled donor-only (D-only, D0) and acceptor-only species (A-only, A0), and furthermore using those sub-populations it is possible to determine four levels of correction to measured fluorescence signals: (*i*) $\boldsymbol{\alpha}$ − correction of FRET sensitized acceptor signal for donor crosstalk (leakage) into acceptor emission channel (*ii*) $\boldsymbol{\delta}$ − correction of FRET sensitized acceptor signal for direct acceptor excitation with donor excitation laser (*iii*) $\boldsymbol{\gamma}$ − correction of donor signal upon donor excitation for distinct fluorescence quantum yields of donor and acceptor, as well for different detection efficiencies of green or red detection path for donor or acceptor fluorophore (*iv*) $\boldsymbol{\beta}$ − correction of acceptor signal upon acceptor excitation for different excitation rate of acceptor and donor with their corresponding excitation light. These four correction parameters, in addition to background correction, allow to convert semiquantitative observables, i.e. apparent FRET efficiency and apparent stoichiometry, into their fully-corrected (accurate) equivalents, making it possible for FRET to deliver structural information with interdye distances accuracy of ≤ 5 Å.[12, 13]

Recent international blind study, where 19 laboratories were asked to report accurate FRET efficiencies on two proteins undergoing conformational change, has thought us a lot about the objective performance of FRET experiments.[13] While on one hand this study was encouraging in terms of observed precision, accuracy and sensitivity of FRET experiments, it also revealed limitations, yet to be overcome. The study showed that the spread of reported accurate FRET efficiencies was mainly caused by the error in data calibration, where predominant error contributor was found to be detection efficiency correction parameter, γ. These findings indicated that the routines for data correction needed further improvement, and they triggered a vivid debate in the community on proper procedures for data correction. However, despite these observations, up to date there is no consensus in the community on the optimal correction procedure. A generalized protocol is lacking and in-house procedures for estimation of correction parameters and filtering of data are often not enclosed in the publications. Apart from affecting the accuracy of smFRET measurements, error in data correction is limiting even further our possibilities. With emerging use of graphical tools, such as FRET-lines[14, 15], one could go beyond resolving accurate FRET efficiencies, and obtain more information on the dynamic systems, such as to identify limiting states and their connectivity map. However, for

reliable interpretation of a system using graphical tools, it is of crucial importance to properly determine correction parameters.

Historically, there are two most commonly referred correction procedures, but of rather limited application, considering how detection efficiency parameter, γ, is determined. In one case, γ is obtained from the linear regression of $1/\langle {}^{iii}S_{\mathrm{app}}\rangle$ versus $\langle {}^{iii}E_{\mathrm{app}}\rangle$, where $\langle {}^{iii}S_{\mathrm{app}}\rangle$ and $\langle {}^{iii}E_{\mathrm{app}}\rangle$ represent population means of double-labelled sub-populations.[16, 17] Clearly, the disadvantage is the necessity to have either multiple samples in calibration (i.e. minimum two), or that a single sample shows distance-related E broadening. Second, this approach requires uniform acceptor properties for all samples ("homogeneous approximation"), i.e. this method represents global γ correction. Second historical reference for the smFRET data calibration is given in Kudriavtstev *et al.*[11] Two approaches are proposed there, with the first one requiring minimum two samples to be used in the calibration, and where γ is tuned till the donor-acceptor sub-populations get the same stoichiometry distribution, since stoichiometry should be independent of the FRET efficiency. In the second approach, calibration can be done on just one sample, however, γ is tuned till the donor-acceptor sub-population(s) fall on the static $E-\tau$ FRET-line. This method requires that the nature of the system is known, i.e. system must be static, which limits the application versatility of the method. Additionally, requirement is the use of static FRET-line. Defining accurate static FRET-line as a reference for calibration is far from trivial if donor has multi-exponential lifetime decay. Good example is widely used Atto 550 dye, which has three distinct lifetimes originating from different isomeric states. Furthermore, typically used standard deviation of ~6 Å [18, 19] for linker-induced broadening of inter-dye distance distributions is determined in benchmark studies involving rigid DNA molecules, and hence may not be accurate for more complex environments such as those in proteins. To avoid the use of static FRET-line, possible workaround here is to decompose the detection ratio, γ, on two factors, with one being the ratio of donor and acceptor effective quantum yields $\Phi_{\mathrm{F,A}}/\Phi_{\mathrm{F,D}}$, and the second factor being the ratio of instrument detection efficiency for acceptor and donor fluorophore, $g_{\mathrm{R|A}}/g_{\mathrm{G|D}}$. This way nature of the system does not have to be known in advance. However, pitfall is that both terms require additional measurements (spectra of dyes and optical components) and reference values (quantum yields), which can result in significant error. Furthermore, neither of the aforementioned approaches give clear guidelines on selecting donor-only (D-only, D0), acceptor-only (A-only, A0) and donor-acceptor (DA) sub-populations, which is the first step in determination of correction parameters.

Driven with all aforementioned challenges and debates in the community, in this work we evaluated performance of multiple methods for determination of correction parameters, and building upon the previous contributions in the community we designed a minimally user-biased protocol (**Figure 1**). Our protocol comprises instructions on how to select sub-populations necessary for determination of correction parameters (e.g. D-only, A-only and double-labelled donor-acceptor population, DA) as well the instructions for their determination. We further propose complementary measurements for cross-validation of obtained correction parameters, although all necessary information can be obtained from the measurement of the sample. Furthermore, we complemented data correction workflow with guidelines for setup alignment and calibration. Robustness of protocol for determination of correction parameters is tested in a blind challenge, and furthermore, their accuracy is assessed through the agreement between derived and modelled inter-dye distances and by quantifying the displacement of DA sub-populations from the static $E - \tau$ FRET-line. Careful implementation of the here proposed workflow should enhance the accuracy of FRET measurements and allow use of graphical tools for the analysis of complex biomolecular systems.

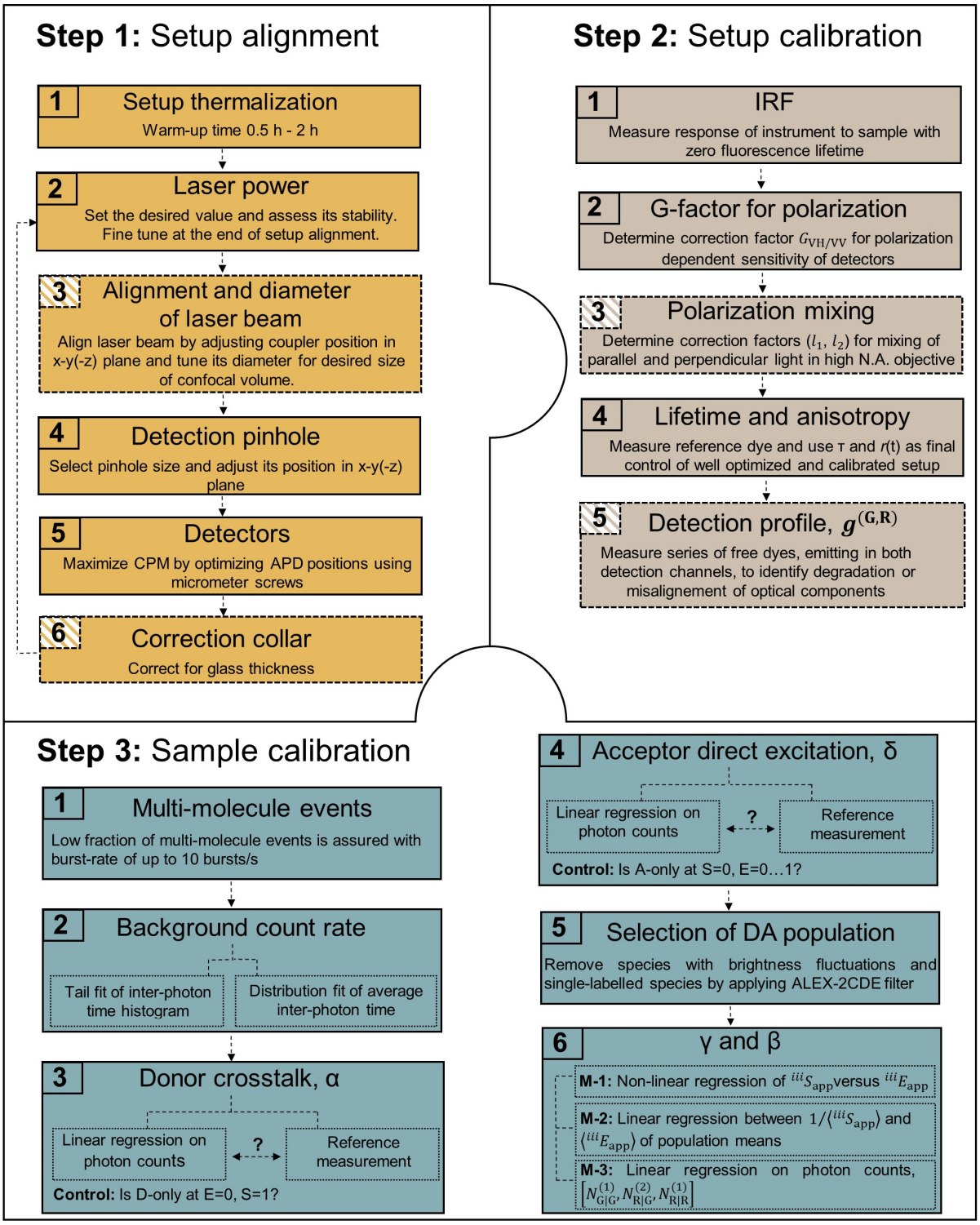


Figure 1. Workflow diagram for MFD-PIE confocal setup alignment and setup/sample calibration. Steps with dotted frame are typically not needed in day-to-day alignment/calibration.

2. Theory

Nomenclature, definitions and E, S-correction procedure. In a dual-color PIE excitation scheme, there are four micro-time data channels, that are defined by the excitation laser X (Y|**X)** and detector Y where the photons are being registered (**Y**|X). Each of two detectors (green and red) has two micro-time windows, defined with excitation light. Green micro-time detection windows are: G|G and G|R. Similarly, red micro-time detection windows are: R|G and R|R. When PIE is combined with MFD, additional polarization separation of photons is being performed in each of the four micro-time windows, resulting in total of eight independent detection channels. Fluorescence intensity in green detection channel after red excitation can be ignored, i.e. ${}^{i}I_{\mathrm{G|R}} \approx 0$, since the red laser normally does not excite the donor fluorophore. Fluorescence intensities detected in other three micro-time channels, namely: donor fluorescence intensity after green laser time-on, ${}^{i}I_{\mathrm{G|G}}$, fluorescence intensity of acceptor upon green laser time-on, ${}^{i}I_{\mathrm{R|G}}$, and fluorescence intensity of acceptor upon red laser excitation, ${}^{i}I_{\mathrm{R|R}}$, constitute main parameters for determination of apparent FRET efficiency and stoichiometry:

$$ {}^{i}E_{\mathrm{app}} = \frac{{}^{i}I_{\mathrm{R|G}}}{{}^{i}I_{\mathrm{R|G}} + {}^{i}I_{\mathrm{G|G}}} \tag{1} $$

$$ {}^{i}S_{\mathrm{app}} = \frac{{}^{i}I_{\mathrm{R|G}} + {}^{i}I_{\mathrm{G|G}}}{{}^{i}I_{\mathrm{R|G}} + {}^{i}I_{\mathrm{G|G}} + {}^{i}I_{\mathrm{R|R}}} \tag{2} $$

As the first step toward accurate FRET efficiency, background intensities $I_{\mathrm{Y|X}}^{(\mathrm{BG})}$, that originate from scattered light, detector dark counts and sample impurities, are subtracted from raw fluorescence intensities in each of the micro-time detection channels:

$$ {}^{ii}I_{\mathrm{Y|X}} = {}^{i}I_{\mathrm{Y|X}} - I_{\mathrm{Y|X}}^{(\mathrm{BG})} \tag{3} $$

and background-corrected FRET observables ${}^{ii}E_{\mathrm{app}}$ and ${}^{ii}S_{\mathrm{app}}$ are calculated.

Background corrected fluorescence intensity of acceptor upon green laser time-on, ${}^{ii}I_{\mathrm{R|G}}$, also known as FRET sensitized acceptor emission, is in the next step corrected for donor crosstalk (leakage) into acceptor emission channel after green laser time-on, given as correction parameter α, and for direct acceptor excitation with donor excitation laser, given as correction parameter δ:

$$F_{\mathrm{R|G}} = {}^{ii}I_{\mathrm{R|G}} - \alpha\,{}^{ii}I_{\mathrm{G|G}} - \delta\,{}^{ii}I_{\mathrm{R|R}} \tag{4}$$

Stoichiometry and FRET efficiency values can now be updated, using $F_{\mathrm{R|G}}$, ${}^{ii}I_{\mathrm{G|G}}$ and ${}^{ii}I_{\mathrm{R|R}}$:

$${}^{iii}E_{\mathrm{app}} = \frac{F_{\mathrm{R|G}}}{F_{\mathrm{R|G}} + {}^{ii}I_{\mathrm{G|G}}} \tag{5}$$

$${}^{iii}S_{\mathrm{app}} = \frac{F_{\mathrm{R|G}} + {}^{ii}I_{\mathrm{G|G}}}{F_{\mathrm{R|G}} + {}^{ii}I_{\mathrm{G|G}} + {}^{ii}I_{\mathrm{R|R}}} \tag{6}$$

Donor crosstalk is defined as the ratio of detection efficiency of red detection path for a donor dye, $g_{\mathrm{R|D}}$, and detection efficiency of greed detection path for donor dye, $g_{\mathrm{G|D}}$:

$$\alpha = \frac{g_{\mathrm{R|D}}}{g_{\mathrm{G|D}}} \tag{7}$$

while direct acceptor excitation with donor excitation laser is defined via acceptor excitation cross-sections for green and red-light source, $\sigma_{\mathrm{A|G}}$ and $\sigma_{\mathrm{A|R}}$, and corresponding excitation light irradiances, L_{G} and L_{R}:

$$\delta = \frac{\sigma_{\mathrm{A|G}}\,L_{\mathrm{G}}}{\sigma_{\mathrm{A|R}}\,L_{\mathrm{R}}} \tag{8}$$

Furthermore, to account for different detection sensitivity of green detection channels for donor dye, $g_{\mathrm{G|D}}$, and red detection channels for acceptor dye, $g_{\mathrm{R|A}}$, and for different fluorescence quantum yields of donor, $\Phi_{\mathrm{F,D}}$, and acceptor, $\Phi_{\mathrm{F,A}}$, donor fluorescence intensity after donor laser time-on is multiplied with γ correction factor:

$$\gamma = \frac{\Phi_{\mathrm{F,A}}}{\Phi_{\mathrm{F,D}}} \cdot \frac{g_{\mathrm{R|A}}}{g_{\mathrm{G|D}}} \tag{9}$$

$$F_{\mathrm{G|G}} = \gamma\,{}^{ii}I_{\mathrm{G|G}} \tag{10}$$

Lastly, to account for different excitation rates of donor and acceptor with their corresponding excitation light source, fluorescence intensity of acceptor fluorescence upon red excitation is divided with β correction parameter:

$$\beta = \frac{\sigma_{\mathrm{A|R}}\,L_{\mathrm{R}}}{\sigma_{\mathrm{D|G}}\,L_{\mathrm{G}}} \tag{11}$$

$$F_{\mathrm{R|R}} = \frac{1}{\beta}\,{}^{ii}I_{\mathrm{R|R}} \tag{12}$$

with $\sigma_{\mathrm{A|R}}$ being the acceptor excitation cross-section for the red light, and $\sigma_{\mathrm{D|G}}$ donor excitation cross-section for the green light.

Finally, using fully corrected fluorescence intensities, as defined in eqs. (4), (10) and (12), accurate FRET efficiency and stoichiometry can be obtained:

$$E = \frac{F_{\mathrm{R|G}}}{F_{\mathrm{R|G}} + F_{\mathrm{G|G}}} \tag{13}$$

$$S = \frac{F_{\mathrm{R|G}} + F_{\mathrm{G|G}}}{F_{\mathrm{R|G}} + F_{\mathrm{G|G}} + F_{\mathrm{R|R}}} \tag{14}$$

Detailed overview over symbols used in this work, as well their definitions are comprised in **Table 1.**

Table 1. Used symbols and their definitions.

Variable	description	eqs
Abbreviations		
a) detection channel names		
G|G	intensity in green (donor) channel upon green (donor) excitation	
R|G	intensity in red (acceptor) channel upon green (donor) excitation	
R|R	intensity in red (acceptor) channel upon red (acceptor) excitation	
b) species names		
A-only or A0	acceptor-only species	
D-only or D0	donor-only species	
DA	double-labelled donor-acceptor species	
c) other indices used		
D or A	donor or acceptor	
BG	background	
i	uncorrected (raw) quantity	
ii	background corrected quantity	
iii	background, crosstalk and direct excitation corrected quantity	
PIE-MFD quantities		
a) correction parameters		
$\alpha = \frac{g_{\mathrm{R\|D}}}{g_{\mathrm{G\|D}}}$	leakage (crosstalk) of donor fluorescence into the acceptor detection channel	4, 7, 29, 30, 31
$\delta = \frac{\sigma_{\mathrm{A\|G}}\, L_{\mathrm{G}}}{\sigma_{\mathrm{D\|R}}\, L_{\mathrm{R}}}$	direct acceptor excitation by the donor excitation laser	4, 8, 32, 33, 34
$\gamma = \frac{\Phi_{\mathrm{F,A}}}{\Phi_{\mathrm{F,D}}} \cdot \frac{g_{\mathrm{R\|A}}}{g_{\mathrm{G\|D}}}$	correction for distinct fluorescence quantum yields of donor and acceptor, and different detection efficiencies of green and red detection channels for donor and acceptor fluorophore	9, 10, 35, 38, 41
$\beta = \frac{\sigma_{\mathrm{A\|R}}\, L_{\mathrm{R}}}{\sigma_{\mathrm{D\|G}}\, L_{\mathrm{G}}}$	correction for different excitation rates of acceptor and donor with their corresponding excitation light	11, 12, 35, 37, 41
b) Photon counts, signal count-rates and fluorescence intensities		
$S_{\mathrm{Y\|X}}$	uncorrected (raw) count rate in detection channel Y after X excitation (X and Y can be G or R)	

$S_{\mathrm{Y\|X}}^{(\mathrm{BG})}$	background count rate	
$N_{\mathrm{Y\|X}}$	uncorrected (raw) photon counts	
$N_{\mathrm{Y\|X}}^{(1)}$	background corrected photon counts	
$N_{\mathrm{Y\|X}}^{(2)}$	background, crosstalk, α, and direct excitation, δ, corrected photon counts	
${}^{i}I_{\mathrm{Y\|X}}$	uncorrected (raw) fluorescence intensity	
${}^{ii}I_{\mathrm{Y\|X}} = {}^{i}I_{\mathrm{Y\|X}} - I_{\mathrm{Y\|X}}^{(\mathrm{BG})}$	background corrected fluorescence intensity	3
$F_{\mathrm{R\|G}} = {}^{ii}I_{\mathrm{R\|G}} - \alpha\,{}^{ii}I_{\mathrm{G\|G}} - \delta\,{}^{ii}I_{\mathrm{R\|R}}$	background, crosstalk and direct excitation corrected FRET-sensitized acceptor fluorescence intensity	4
$F_{\mathrm{G\|G}} = \gamma\,{}^{ii}I_{\mathrm{G\|G}}$	detection efficiency corrected fluorescence intensity of donor upon donor laser time-on	10
$F_{\mathrm{R\|R}} = \frac{1}{\beta}\,{}^{ii}I_{\mathrm{R\|R}}$	fluorescence intensity of acceptor upon acceptor laser time-on, corrected for different excitation rates of D and A	12
c) Burst-wise quantities		
T	burst duration	
${}^{i}E_{\mathrm{app}} = \frac{{}^{i}I_{\mathrm{R\|G}}}{{}^{i}I_{\mathrm{R\|G}} + {}^{i}I_{\mathrm{G\|G}}}$ ${}^{i}S_{\mathrm{app}} = \frac{{}^{i}I_{\mathrm{R\|G}} + {}^{i}I_{\mathrm{G\|G}}}{{}^{i}I_{\mathrm{R\|G}} + {}^{i}I_{\mathrm{G\|G}} + {}^{i}I_{\mathrm{R\|R}}}$	fully uncorrected apparent FRET efficiency and stoichiometry, calculated from raw photon counts	1, 2
${}^{ii}E_{\mathrm{app}} = \frac{{}^{ii}I_{\mathrm{R\|G}}}{{}^{ii}I_{\mathrm{R\|G}} + {}^{ii}I_{\mathrm{G\|G}}}$ ${}^{ii}S_{\mathrm{app}} = \frac{{}^{ii}I_{\mathrm{R\|G}} + {}^{ii}I_{\mathrm{G\|G}}}{{}^{ii}I_{\mathrm{R\|G}} + {}^{ii}I_{\mathrm{G\|G}} + {}^{i}I_{\mathrm{R\|R}}}$	background corrected apparent FRET efficiency and stoichiometry	
${}^{iii}E_{\mathrm{app}} = \frac{F_{\mathrm{R\|G}}}{F_{\mathrm{R\|G}} + {}^{ii}I_{\mathrm{G\|G}}}$ ${}^{iii}S_{\mathrm{app}} = \frac{F_{\mathrm{R\|G}} + {}^{ii}I_{\mathrm{G\|G}}}{F_{\mathrm{R\|G}} + {}^{ii}I_{\mathrm{G\|G}} + {}^{ii}I_{\mathrm{R\|R}}}$	background, crosstalk, α, and direct excitation, δ, corrected apparent FRET efficiency and stoichiometry	5, 6
$E = \frac{F_{\mathrm{R\|G}}}{F_{\mathrm{R\|G}} + F_{\mathrm{G\|G}}}$ $S = \frac{F_{\mathrm{R\|G}} + F_{\mathrm{G\|G}}}{F_{\mathrm{R\|G}} + F_{\mathrm{G\|G}} + F_{\mathrm{R\|R}}}$	fully corrected (accurate) FRET efficiency and stoichiometry, after correction for: background, donor crosstalk, α, acceptor direct excitation, δ, detection ratio, γ, and ratio of excitation flux, β	13, 14
d) other variables used		

$G_{\mathrm{VH/VV}}$	correction factor compensating for polarization dependent sensitivity of the detectors	18, 20
l_1, l_2	correction factors for polarization mixing in high-NA microscope objectives	
$\sigma_{\mathrm{A\|G}}$, $\sigma_{\mathrm{A\|R}}$	excitation cross-sections of acceptor for green and red light source	8
$\sigma_{\mathrm{D\|G}}$	excitation cross-section of donor for green light source	11
L_{G}, L_{R}	irradiance of green and red excitation light	8, 11
Δt	inter-photon time	
Δt_{av}	average inter-photon time	
$g_{\mathrm{G\|D}}$, $g_{\mathrm{R\|A}}$	detection efficiency of a green (G) or red (R) detection path for donor, D, or acceptor, A, fluorophore	9
$\Phi_{\mathrm{F,D}}$, $\Phi_{\mathrm{F,A}}$	fluorescence quantum yield of donor and acceptor	9
ALEX-2CDE	ALEX two-channel kernel-based density estimator function	

3. Protocol

This protocol starts with steps for setup alignment and calibration (**step 1**-**2**), followed by the steps for determination of sample-related correction parameters (**step 3**). Diagram comprising all the recommended steps for setup alignment and calibration, as well for data correction is given in **Figure 1**Error! Reference source not found..

3.1. Confocal PIE-MFD setup optimization and calibration

3.1.1. Step 1: Setup alignment

Setup alignment instructions are tailored to the typical architecture of the PIE-MFD confocal workstation as previously described.[12] For detailed description of optical design of our setup see **Supplemental Note 1**.

Caution: Users should familiarize themselves beforehand with laser safety, to avoid injuries due to the exposure to open laser radiation during the setup alignment.

Step 1.1: Setup thermalization. Prior to any setup alignment or measurements, one should allow for thermal stabilization of the setup. Ideally, the microscope is placed in a temperature-controlled room. The entire setup should be thermally stabilized for at least 30 minutes up to two hours before usage, to avoid drift of the optical components during the experiment. Importantly, the laser source should be allowed to stabilize to avoid slow fluctuations of the incident laser power.

Step 1.2: Laser power. As the next step, one should set the desired laser power. After the setup is optimized, laser power is finely tuned again prior to measurements. Commonly, the aim is to have as much signal as possible, however, one has to make a compromise such that dye photobleaching and photoblinking are avoided. However, if high laser powers are required, one could reduce dye photo-damage by using chemical photo-stabilizers, such as Trolox. In any case, high laser powers that produce count rates of more than 200 kHz should be avoided, as detector saturation occurs in this range for APD detectors due to detector dead times of typically ~50-100 ns. After setting up the laser power, its stability should be assured. Occasionally, stability of the laser power should be done over extended periods of the time. In day-to-day work, multiple power checks during the day are commonly sufficient.

Step 1.3a: Laser beam alignment. To align the laser beam arising from a fiber, we recommend mounting a crosshair on the microscope, and adjust the laser output coupler in the lateral (*x*-*y*) plane using micrometer screws, until equal illumination of all four crosshair arms is observed in image on the ceiling. Any adjustment of the coupler position in the axial (*z*) direction requires re-alignment in lateral plane. Furthermore, any adjustment of the excitation path requires a careful realignment of the detection path. Generally, this step is not needed in day-to-day alignment if the laser output coupler is stably mounted.

Step 1.3b: Size of confocal volume. The size of the observation volume can be controlled by tuning the diameter of the laser beam at the back aperture of the objective. Shape of this dependence has been given in the work of Enderlein *et al*. [20] By underfilling back aperture of the objective (i.e. radius of back aperture is larger than the radius of laser beam), observation volume increases. Then, observation volume, also called point spread function (PSF), can be approximated by three-dimensional Gaussian function.[21, 22] Although for intensity-based approaches shape of PSF is of no importance, this plays a large role for formal description of molecular diffusion for a range of correlation spectroscopy methods, as it will be shown later. Furthermore, bigger observation volume means longer burst duration, which allows slow processes to be studied. However, for underfilled objective, size of observation volume is more sensitive to the changes in laser diameter, causing such experiments harder to be reproduced.[20] On the other hand, smaller observation volume generally provides higher count rates at equal laser power due to higher irradiances, yielding a higher signal to noise ratio (SNR) and allows to study fast processes.[13] However, smaller observation volumes are also more non-Gaussian.[21] Therefore, choice of observation volume size has to be a trade-off between a high SNR and validity of a Gaussian function assumption. There are several ways to determine the size of confocal volume, and here we focus on approaches that rely on FCS. One way to determine the size of observation volume is via the fit of FCS curve of a free dye with known translational diffusion coefficient (e.g. Rhodamine 110, D_{trans} = 4.4·10^{-6} cm^2/s at T= 22.5 °C[23]). Given that dye diffusion is the only source of the fluorescence intensity fluctuation, and that the observation volume can be approximated by a three-dimensional Gaussian function, the following fit model can be applied to the data[24]

$$G(t_{\mathrm{c}}) = 1 + \frac{1}{N} G_{\mathrm{trans}}(t_{\mathrm{c}}) = 1 + \frac{1}{N} \cdot \left[\frac{1}{1 + \frac{t_{\mathrm{c}}}{t_{\mathrm{trans}}}} \cdot \frac{1}{\sqrt{1 + \frac{t_{\mathrm{c}}}{\left(\frac{z_0}{\omega_0}\right)^2 t_{\mathrm{trans}}}}} \right] \qquad (15)$$

with the fit parameters being as follows: t_c- correlation lag time, N - average number of molecules in the detection volume obtained as $1/G\ (t_c \to 0)$, t_{trans} – translational diffusion time of a molecule, z_0 and ω_0 - axial and lateral $1/e^2$ radii of the detection volume. From the relationship between known translational diffusion coefficient, D_{trans}, and fitted translational diffusion time, t_{trans}, one can estimate the lateral radius $\omega_0 = \sqrt{4D_{trans}t_{trans}}$.

Next, axial radius z_0 is obtained from the fitted ratio $\frac{z_0}{\omega_0}$ and previously calculated value of ω_0.

Finally, size of observation volume is obtained according to the:

$$V = \pi^{3/2} z_0 \omega_0^2 \tag{16}$$

With the approach described above we found in our measurements size of observation volume to be ~6 fL, corresponding to lateral and axial radii of ω_0~ 0.6 μm and z_0~ 3.7 μm, and diffusion time of Rhodamine 110 of t_{trans} ~ 0.170 ms.

Another approach to estimate the size of detection volume is by measuring FCS curves for a concentration series of a free dye. By plotting the estimated number of molecules in detection volume, *N*, versus the sample concentration, *c*, size of observation volume can be determined from the slope of approximately linear dependence $N = VN_A c$.

At lower sample concentrations, dependence deviates from a linear one, due to the prominent contribution of uncorrelated background signal, which lowers the correlation amplitude. Although this approach does not require any assumption to be made in regard to the shape of PSF, a wide range of sample concentrations has to be measured and tedious correction for background contribution at lower concentrations has to be performed.

Step 1.4: Detection pinhole. Next, pinhole size should be selected. Optimal pinhole size should correspond to one Airy Unit (1 AU), which represents the diameter of the Airy disk. Airy disk is the inner circle of so called "Airy pattern", which is obtained as a consequence of light diffraction at the objective aperture. Airy disk contains approximately 84% of diffracted light, and remaining light is distributed across concentric circles. The more the pinhole exceeds the diameter of Airy disk, more out-of-focus light is transmitted. Optimal size of pinhole, equal to 1 AU, is a function of objective magnification, M, numerical aperture, NA, and wavelength of transmitted light:

$$d = \frac{1.22 \cdot \lambda}{NA} \cdot M \tag{17}$$

However, due to aberrations and underfilling of the back aperture of the objective, diameter of Airy disk is usually underestimated.

Alternatively, one can vary pinhole size till the maximum dye brightness is obtained, or specifically maximum *cpm* (counts per molecule (per time). Dye brightness, or *cpm*, is simply determined by dividing the average count rate with the number of molecules in the detection volume. For this, one constructs FCS curve from a nM measurement of a known dye (e.g. Rhodamine 110). Besides obtaining maximum *cpm*, approximation of 3D-Gaussian shaped detection volume should be satisfied, as measured by the fit quality of FCS curve. For our setup with 60x/1.2 NA objective and dual edge dichroic beamsplitter FF500/646-Di01, we found rather a large pinhole of 100 µm to be suitable. Besides selecting pinhole size, its position should be optimized. Pinhole should be positioned at the focal plane of the tube lens. Using micrometer screws, optimization of its position in lateral (x-y) plane is commonly sufficient in day-to-day setup alignment.

Step 1.5: Detectors alignment. For alignment of each of detectors one can measure nM water solution of dye emitting in a given detection channel. For instance, Rhodamine 110 for green channels, and Rhodamine 101 for red channels. APD positions are then optimized using micrometer screws, such that maximum *cpm* is obtained.

Step 1.6: Objective correction collar. To correct for the thickness of the cover glass, most water-immersion objectives are equipped with a correction collar, that allows the adjustment of the central lens group position to coincide with the glass thickness. Objective correction collar position should correspond to the manufacturer information on the glass thickness. For instance, if one uses #1.5 cover glass, correction collar should be set to the position 0.17 mm. However, thickness of a cover glass usually varies few percent compared to the value reported by manufacturer. Enderlein et al[20] have shown that deviation in cover-slide thickness mainly affects axial radius (i.e. elongation of focal volume) and that absolute deviation of 10 µm in glass thickness can cause an increase in focal volume by approximately a factor of 2. Therefore, to avoid deformation of focal volume due to variation in cover glass thickness, it is recommended to optimize the position of correction collar before the experiments such that the *cpm* is maximized.

3.1.2. Step 2: Setup calibration

Step 2.1: Instrument Response Function. Fluorescence intensity decay of a sample of interest is analyzed in convolution with Instrument Response Function (IRF), which represents response of the instrument to zero lifetime sample. MilliQ water normally satisfies the purity requirement for this purpose. Depending on the age of the filter, MilliQ water can be contaminated by fluorescent background (seen as an exponential tail in the TCSPC histogram), in which case pure water can also be purchased from suppliers. It is also possible to measure the IRF using fluorescent dyes with a short lifetime (< 50 ps), such as azo dye Allura Red AC (CAS: 25956-17-6) with fluorescence lifetime of ~10 ps, or by measuring any dye in a highly concentrated potassium-iodide solution which lowers the lifetime of dye by collisional quenching [25]. This has the benefit that the IRF is measured using a dye with identical spectral signature as in the sample of interest. However, small but finite lifetimes of the reference sample can affect the extracted lifetimes, especially in the sub-nanosecond regime.

Step 2.2: G-factor for polarization. For each of the detection channels one should determine G-factor for polarization, $G_{\mathrm{VH/VV}}$, a parameter that quantifies polarization dependent sensitivity of the detectors, defined as the ratio of horizontally and vertically polarized light. For each of the channels, nM measurement of a dye emitting in given channel should be performed. For macro-times t_i of at least five times the dye rotational correlation time ρ, any observed polarization must be due to $G_{\mathrm{VH/VV}} \neq 1$. Thus, to obtain the $G_{\mathrm{VH/VV}}$, polarization resolved fluorescence decays, F_{VH} and F_{VV}, are integrated over $t_i \gg \rho$ TCSPC channels[26]:

$$G_{\mathrm{VH/VV}} = \frac{\sum_{t_i \gg \rho}^{t_{\max}} F_{\mathrm{VH}}(t_i)}{\sum_{t_i \gg \rho}^{t_{\max}} F_{\mathrm{VV}}(t_i)} \tag{18}$$

Another approach to determine $G_{\mathrm{VH/VV}}$ factor is to fit the uncorrected time-resolved anisotropy decay of a free dye, with a mono-exponential decay:

$$r(t) = (r_0 - r_\infty) \exp\left(-\frac{t}{\rho}\right) + r_\infty \tag{19}$$

If F_{VH} and F_{VV} were to be the same, anisotropy would be zero. However, if the anisotropy decay appears with the offset, it is due to the polarization sensitivity of the detectors, and $G_{\mathrm{VH/VV}}$ factor can then be directly determined from the offset of anisotropy decay [11]:

$$G_{\mathrm{VH/VV}} = \frac{1 - r_\infty}{1 + 2r_\infty} \tag{20}$$

Step 2.3: Correction for polarization mixing. Mixing of the polarization, of both incident and emitted light, happens due to the light refraction on the high NA objective lenses. To correct for this effect in the calculation of anisotropy, introduced are correction factors l_1 and l_2. [27] For their determination used is a sample with slow rotational correlation time, and of simple and known characteristics. Good example is mutant of GFP protein, known as eYFP (enhanced Yellow Fluorescence Protein) with mono-exponential lifetime decay, with fluorescence lifetime of τ = 3.2 ns, fundamental anisotropy of r_0 = 0.39 and rotational correlation time of ρ=16 ns [28]. nM solution of eYFP in the TRIS buffer (pH=8.5) is prepared and polarization resolved signals are recorded of both the sample and buffer. In the next step one performs a global fit of parallel and perpendicular decays with fixed ρ = 16 ns, and l_1 and l_2 being the fit parameters. In addition, if fluorescence lifetime is previously fitted from a magic angle decay, it can be as well fixed in a global fit, to reduce the number of free fit parameters. Normally, l_1 and l_2 parameters are determined once for a given objective and size of observation volume, and not in day-to-day calibration.

Step 2.4: Lifetime and anisotropy of a known dye. Perform a pM water solution measurement of a known dye with single exponential decay (e.g. Rhodamine 110). Use scatter-corrected steady-state anisotropy and fitted lifetime as final control of well performed setup optimization and calibration.

Step 2.5: Regularized optimization of instrument detection efficiency profile, $g^{(\mathrm{G,R})}$. Detection efficiency of a setup is determined by the shape functions $g^{\mathrm{G}}(\lambda)$ and $g^{\mathrm{R}}(\lambda)$, which represent spectral response of green and red detection path. Shape function $g^{\mathrm{G}}(\lambda)$ and $g^{\mathrm{R}}(\lambda)$ are a product of spectral response of each of the setup components: microscope objective, dichroic beamsplitter, longpass filter, bandpass filters and detectors (see **Figure 2a**). Spectral response of filters is usually known with more certainty, as these can be easily determined in the house using conventional absorption spectrometer. On the other hand, up-to-date spectral response of detectors is harder to obtain, and variation between modules or a change in their performance due to the detector ageing is often an issue. Furthermore, spectral response of some components, such as dichroic beamsplitter, is affected by setup misalignment, due to its angle-dependency.

Therefore, as the final control of well-performing and aligned setup, we propose optimization approach that provides a $g^{(\mathrm{G,R})}$ – profile using a series of dyes measured free in solution, emitting in both detection channels, G and R. From the optimized $g^{(\mathrm{G,R})}$ profile, one can identify if and which detection channel under-performs and at which spectral part. Full mathematical formalism underlying optimization of $g^{(\mathrm{G,R})}$ profile is given in **Supplemental Note 3**. In a nutshell, for each of the j dyes determined is the ratio of background-corrected measured intensities in green and red detection channels:

$$\xi_j = \frac{f_j^{(\mathrm{G})}}{f_j^{(\mathrm{R})}} \tag{21}$$

where observed intensity $f_j^{(\Lambda)}$ of a given detection window $\Lambda = \{\mathrm{G}, \mathrm{R}\}$ is the result of the product of instrument spectral response $g^{(\Lambda)}(\lambda)$ and fluorescence emission spectrum of a dye, $p_j(\lambda)$:

$$f_j^{(\Lambda)} = f_{0,j} \int_{\Lambda} g^{(\Lambda)}(\lambda) p_j(\lambda) d\lambda \tag{22}$$

with $f_{0,j}$ being total radiative intensity of a dye.

Ideally, ratios predicted using fluorescence emission spectra and $g^{(\Lambda)}$ profile, $\xi_{j,\mathrm{predicted}}$, should be equal to measured ratios $\xi_{j,\mathrm{measured}}$. As both spectra and emission intensities are measured with some accuracy, we perform minimization of difference between predicted and measured ratios ξ_j, by optimizing the instrument spectral response, $g^{(\Lambda)}$. Arising ambiguity of the optimized solution can be circumvented by requiring certain properties of the solution, such that only positive values of $g^{(\mathrm{G,R})}(\lambda)$ are allowed, and that no sharp changes are allowed, besides at its edges. Smoothness regularized optimization of $g^{(\mathrm{G,R})}(\lambda)$ is formulated as

$$\begin{gathered} min\left(g^{(\mathrm{G,R})\mathrm{T}}\left(\mu\, S^{(\mathrm{G,R})} + P^{(\mathrm{G,R})T} P^{(\mathrm{G,R})}\right) g^{(\mathrm{G,R})}\right) \\ subject\ to \quad I g^{(\mathrm{G,R})} \geq 0 \end{gathered} \tag{23}$$

where μ is regularization term, penalizing sharp changes of the solution other than at its edges. $S^{(\mathrm{G,R})}$ is a "smoothness" block matrix that is based on the central-difference differentiation operator of first order and $P^{(G,R)}$ is 1×2 block matrix $P^{(\mathrm{G,R})} = \left(P^{(\mathrm{G})}, -\Xi\ P^{(\mathrm{R})}\right)$ with j rows.

$P^{(\Lambda)}$ contains in each of the j rows normalized dye fluorescence emission spectrum with sum in rows being equal to one, and Ξ is diagonal matrix with measured $\xi_{j,\mathrm{measured}}$ ratios along the diagonal.

In **Supplemental Table S3** given is a set of 13 dyes, that are a good example of dyes emitting in both detection channels. Their corresponsind fluorescence emission spectra are given in **Supplemental Figure S1**. Dyes with very low ratios, i.e. $\xi_{j,\mathrm{measured}} \sim 0$, are not recommended, as those will not contribute to the solution. In **Figure 2** illustrated is the approach of regularized optimization of instrument detection profile $g^{(\mathrm{G,R})}$. We have found the approach to be more successful if input are not fluorescence emission spectra $p_j(\lambda)$, but rather fluorescence emission spectra multiplied with transmission spectra of all optical components (objective, dichroic beamsplitter, longpass filter, bandpass filters and detectors), i.e. $g^{(\Lambda)}(\lambda)p_j(\lambda)$. In that case, optimized $g^{(\mathrm{G,R})}$ solution is not the detection profile, but the correction to it. In **Supplemental Table S4** we compile measured, predicted and optimized intensity ratios ξ_j. We show that approach of regularized optimization of $g^{(\mathrm{G,R})}$ profile on average improves agreement with measured intensity ratios for a given set of dyes (**Supplemental Figure S2**). Obtained $g^{(\mathrm{G,R})}$ profile in our case is given as top stack in **Figure 2c.** Chosen $g^{(\mathrm{G,R})}$ solution corresponds to μ = 0.9 value, selected such that it satisfies two criteria: (*i*) has the smallest sum over all dyes of squared deviations of the eq. S13 from zero (see **Supplemental Note 3**), and (*ii*) is smooth within each of Λ windows. Larger value of correction in the right edge of G window, obtained in the case of our setup, we attribute to underestimated detection profile of detectors, as these are associated with highest uncertainty. We identify this approach as very practical way of evaluating setup performance, i.e. identifying potential degradation of optical components, setup misalignment and sub-optimal probe detection.

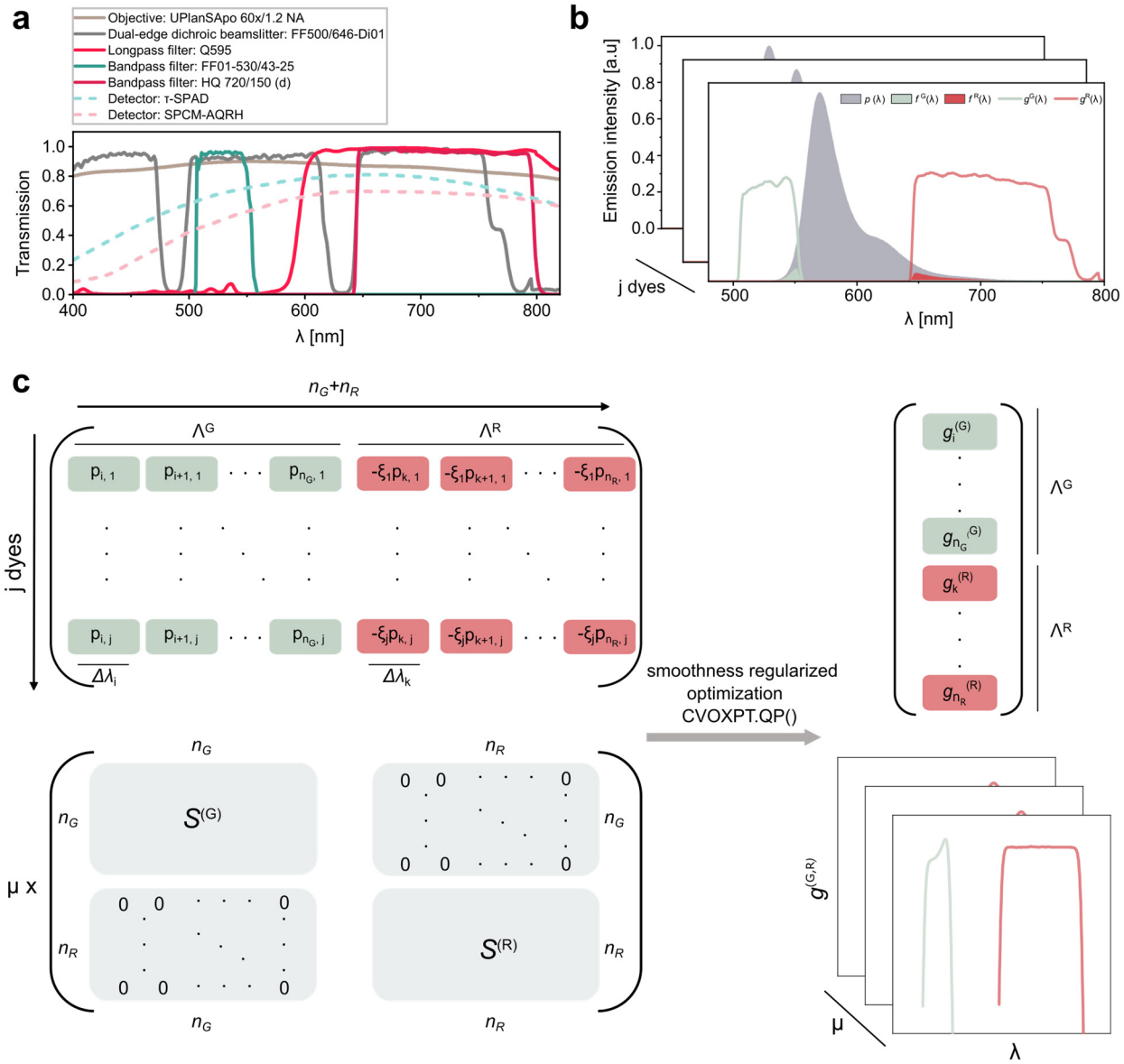


Figure 2. Smoothness regularized optimization of instrument detection efficiency profile $g^{(G,R)}$ using a set of dyes measured free in solution. (a) Transmission spectra of optical components in our MFD confocal setup **(b)** Fluorescence emission spectra of a set of dyes measured free in solution, $p(\lambda)$, as well the observed emission intensities, $f(\lambda)$, after multiplication with instrument detection profile, $g^{(G,R)}$. **(c)** Illustration of input and output data for the determination of instrument detection profile, $g^{(G,R)}$. Optimized $g^{(G,R)}$ solution satisfies measured intensity ratios ξ_j of a set of j dyes, is positive and regularized via parameter μ, which penalizes its sharp changes within a given spectral window $\Lambda = \{G, R\}$ except at its edges. For that, "smoothness" operator $S^{(\Lambda)}$ is applied to the detection profile, and solution with minimum derivative is taken.

3.2. Step 3: Sample calibration

For the evaluation of the protocol we used a set of double-stranded DNA molecules, with varying base-pair separation between donor and acceptor. In total, we designed three dsDNA molecules, with high- (HF), medium- (MF) and low-FRET efficiency (LF), covering range of FRET efficiencies from ~0.1 to ~0.5. As robustness of smFRET measurements depends on the fluorophore choice,[13] we critically assessed our protocol on samples with different dye-pair combinations. As donor we used bright and photostable blue emitting rhodamine dye Alexa Fluor 488 (here referred to as AF488)[29], and as acceptor we used far-red dyes, i.e. carbopyronine dye Atto647N, and cyanine dyes Alexa Fluor 647 (here referred to as AF647) and Cy5. Features of Atto647N are long fluorescence lifetime ($\tau(H_20) = 3.71$ ns), big fluorescence quantum yield ($\Phi_F(H_20) = 0.64$) and increased hydrophobicity.[30] On the other hand, cyanine dyes AF647 and Cy5 have shorter fluorescence lifetime ($\tau(H_20) = 1.02$ ns, $\tau(H_20) = 0.99$ ns), lower fluorescence quantum yield ($\Phi_F(H_20) = 0.34$, $\Phi_F(H_20) = 0.29$), and are associated with cis-trans isomerization induced blinking.[30-32] For extensive overview of photophysical features of used dyes reader is referred to work of Vandenberk *et al*.[33] Not only do these dyes display diverse photophysical features, but they are also widely used in smFRET studies. In all three samples (HF, MF and LF), position of acceptor is fixed, which allows homogeneous approximation and global γ correction. For further details on samples, including sequence, labelling positions and labelling chemistry, see **Supplemental Note 2** and **Supplemental Table S1.**
Procedure for determination of correction parameters is illustrated on the example of dsDNA molecules labelled with AF488-Atto647N, and the full overview of all experiment- and data processing-specific metadata of all dye combinations is collected in **Supplemental Table S2**.

Step 3.1: Presence of multi-molecule events

First step in data correction procedure is to ensure that samples are measured at sufficiently low concentration, such that each burst of photons is generated by a single molecule passing through confocal volume.[34] As previously described[35, 36], probability to find n molecules in the detection volume follows a Poisson distribution:

$$P_n(N_{\mathrm{av}}) = \frac{{N_{\mathrm{av}}}^n}{n!} e^{-N_{\mathrm{av}}} \tag{24}$$

and is function of the sample concentration with the average number of molecules in detection volume, N_{av}. If more than one molecule diffuses through the detection volume, and are

separated less than the mean transit time, they will produce a single unresolved fluorescence burst. The multi-molecule probability can then be calculated as:

$$\begin{aligned} P_{\text{multi-molecule}} = P_{n>1}(N_{\text{av}}) &= 1 - P_{n=0}(N_{\text{av}}) - P_{n=1}(N_{\text{av}}) \\ &= 1 - e^{-N_{\text{av}}}(N_{\text{av}} + 1) \end{aligned} \tag{25}$$

To obtain the fraction of multi-molecule events in the histogram, multi-molecule probability $P_{n>1}(N_{\text{av}})$ has to be normalized with the burst probability, $P_{n>0}(N_{\text{av}})$:

$$F_{\text{multi-molecule}} = \frac{P_{n>1}(N_{\text{av}})}{P_{n>0}(N_{\text{av}})} = 1 - \frac{N_{\text{av}} e^{-N_{\text{av}}}}{1 - e^{-N_{\text{av}}}} \tag{26}$$

For burst rate of 1 burst/s (including all species, i.e. D-only, A-only, DA and BG), and confocal volume with 1 ms burst duration, $N_{\text{av}} = 0.001$, which yields fraction of multi-molecule events of 0.05%. Similarly, for burst rate of 10 bursts/s, $N_{\text{av}} = 0.01$ and multi-molecule fraction is 0.5%. On the other hand, significant value of 5% is reached at the burst rate of 100 bursts/s. Therefore, we recommend measurements to be performed with burst-rate of up to 10 bursts/s.

Step 3.2: Determination of buffer background count rate

Single-molecules in a focus emit fluorescence photons with high count rate, i.e. short inter-photon times. However, in any real system single-molecule photons are interfered with unwanted background signal. Background photons can have various origins, such as sample impurities, scattering at the excitation wavelength and dark counts of detectors. Photons of such origin are with low count rate, and subsequently with longer inter-photon times. Therefore, all recorded photons can be divided in two categories according to inter-photon times. Proper estimation of pure background count rate is essential for all further calibration steps, and is based on the aforementioned fact that we can distinguish background and single-molecule photons based on inter-photon times, Δt. We give an overview over two ways to determine the background count-rate, both of which perform well and provide consistent results. Two approached are illustrated in **Figure 3.** Lastly, these two approaches are more favorable compared to simple determination of mean background count-rate, where discrimination of pure background signal from potential fluorescence impurities is not possible.

a) Tail fit of inter-photon time histogram [37]

Straight tail of semi-log inter-photon time distribution can be fitted with single exponential function (**Figure 3a**), which indicates that these photons obey to Poisson distribution, and represent background:

$$P(\Delta t) = N_0 \exp(-S^{(\mathrm{BG})} \Delta t) \tag{27}$$

At the short inter-photon times, data deviates from exponential function, which is due to high count rate processes (i.e. fluorescence bursts of single-molecules).

In this approach one defines a threshold after which the distribution can be explained with mono-exponential function. Using non-linear least-squares curve fitting, or Maximum Likelihood Estimator (MLE), one finally obtains background count rate, $S^{(\mathrm{BG})}$, as the inverse of exponential time constant.

b) Distribution fit of the average inter-photon time Δt_{av} [36]

To better discriminate background from fluorescence signal, one can smooth the trace of the inter-photon times, by calculating average inter-photon time Δt_{av} over n subsequent inter-photon times, $\Delta t_{\mathrm{av}}(n) = 1/n \sum_{i=1}^{n} \Delta t_i$. Distribution of average inter-photon times is fitted with gamma distribution (**Figure 3b**):

$$P(\Delta t_{\mathrm{av}}, n) = \frac{(nS^{(\mathrm{BG})})^n}{(n-1)!} \Delta t^{n-1} \exp(-n\Delta t S^{(\mathrm{BG})}) \tag{28}$$

It is recommended to use 5-10 subsequent photons for the estimation of the average inter-photon time. Higher photon numbers lead to data over-smoothing, and wrong estimation of the background count rate.

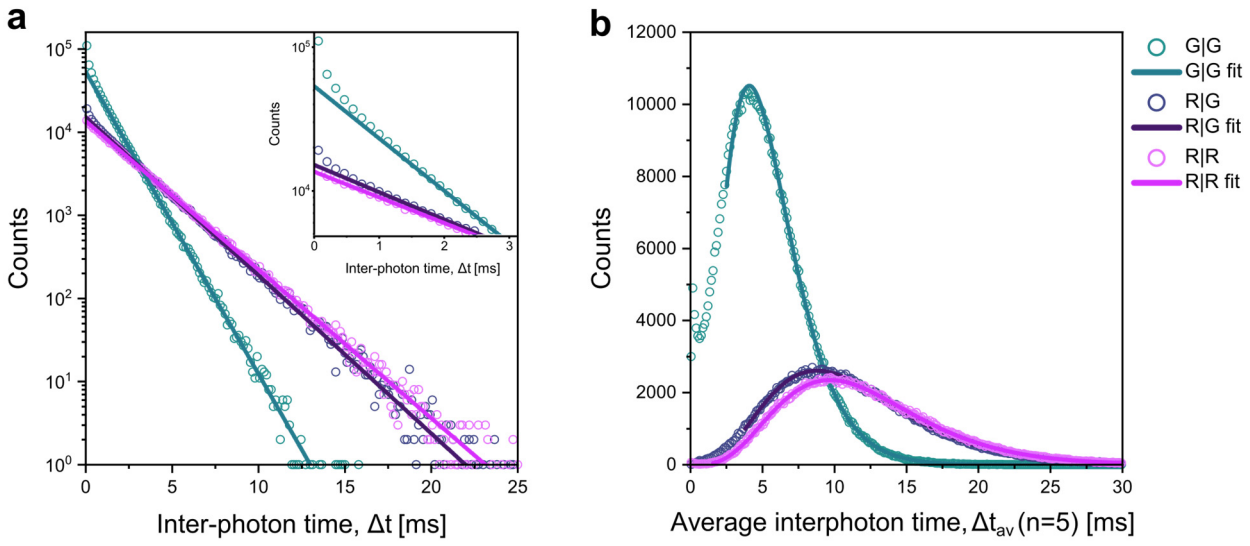


Figure 3. Step 3.2: Background estimation based on inter-photon times. (a) Tail fit of inter-photon times using single-exponential function for each micro-time detection window (G|G, R|G, R|R) according to eq. (27) **(b)** Fit of average inter-photon time distribution using gamma distribution, Γ, for each micro-time detection window (see eq. (28)).

Step 3.3: Donor crosstalk

After the background determination for each of the detection channels, next step needed for accurate *E* and *S* determination is correction for the leakage (crosstalk) of donor emission into acceptor detection channel. The absence of donor crosstalk correction falsely increases the intensity in acceptor emission channel, $F_{R|G}$, and subsequently larger FRET efficiency is measured than expected. As shown previously [12, 17, 38], the error of measured *E* due to donor crosstalk increases towards low- *E* species. They show that while error is negligible for high- *E* species, in the range of intermediate FRET efficiencies, where most measurements are performed, error is already significant. Finally, for low- *E* species, *E* error goes up to 50%, implying that careful crosstalk evaluation is essential for accurate *E*- determination.

For robust crosstalk determination, recommended is to perform a cut of D-only (D0) species from a 2D histogram of fully uncorrected apparent FRET efficiency, ${}^{i}E_{\mathrm{app}}$, versus fully uncorrected apparent stoichiometry, ${}^{i}S_{\mathrm{app}}$, with ${}^{i}S_{\mathrm{app}} > 0.95$, while along ${}^{i}E_{\mathrm{app}}$ axis selection should be made in amount ~ ±0.015, symmetrically around the center of the population (**Figure 4a**). Subsequently, for such selected D-only sub-population, historically it was proposed to histogram the background corrected ${}^{ii}E_{\mathrm{app}}^{(\mathrm{D0})}$ values and fit with a single or a mixture of Gaussian functions to determine the mean value. [11] Next, using the fitted mean value of the distribution, crosstalk can be calculated as follows:

$$\alpha = \frac{\langle {}^{\mathrm{ii}}E_{\mathrm{app}}^{(\mathrm{D0})} \rangle}{1 - \langle {}^{\mathrm{ii}}E_{\mathrm{app}}^{(\mathrm{D0})} \rangle} \tag{29}$$

However, since such approach is user-biased when identifying the mean of the D-only sub-population, here we advocate to instead perform linear bisquare robust regression of background corrected photon counts (**Figure 4b**) following the relation:

$$N_{\mathrm{R|G}} - S_{\mathrm{R|G}}^{(\mathrm{BG})} T = \alpha (N_{\mathrm{G|G}} - S_{\mathrm{G|G}}^{(\mathrm{BG})} T) \tag{30}$$

We highly advocate for the use of robust regression[39], as it is less prone to the effect of outliers compared to ordinary least squares. Robust regression optimizes the weights of each data point through iteratively reweighted least-squares method, and is less sensitive to extreme outliers in data. Here we chose bisquare function for weight calculation, and tune constant of K = 4.685.

Validation of fitted α value can be performed by monitoring if D-only sub-population is positioned at ${}^{ii*}E_{\mathrm{app}}^{(\mathrm{D0})} = 0$ and ${}^{ii*}S_{\mathrm{app}}^{(\mathrm{D0})} = 1$ (**Figure 4c**). Here, asterix in superscript denotes that besides for background, apparent FRET efficiency and stoichiometry are also corrected for donor crosstalk.

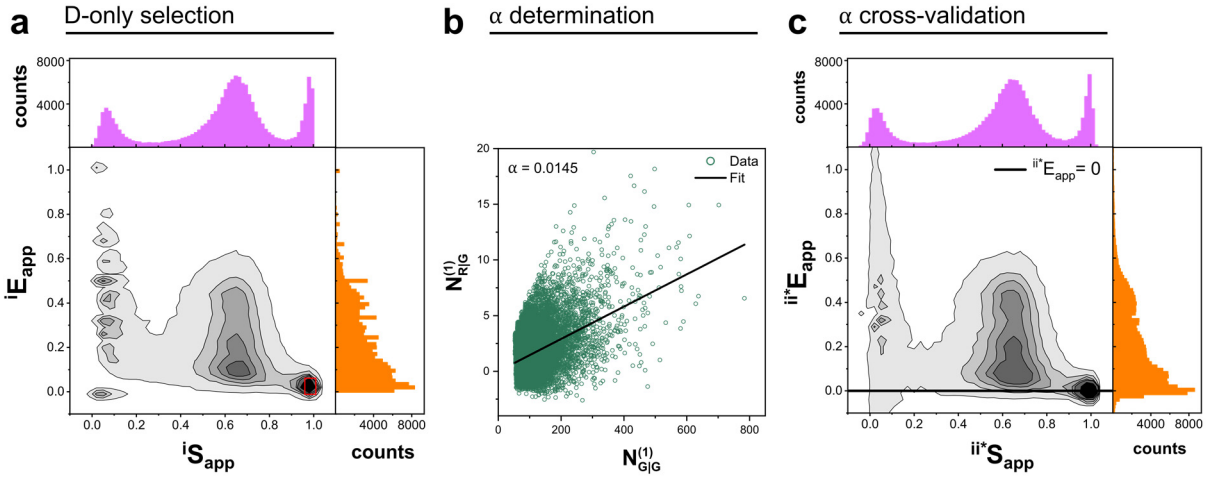


Figure 4. Step 3.3: Determination of donor crosstalk into acceptor emission channel. (a) Selection of donor-only sub-population in ${}^{i}S_{\mathrm{app}} - {}^{i}E_{\mathrm{app}}$ 2D histogram (red rectangle). Only stoichiometry values of ${}^{i}S_{\mathrm{app}} > 0.95$ should be selected, while on ${}^{i}E_{\mathrm{app}}$ axis the selection is made ±0.015 symmetrically around the center od D-only population. **(b)** Robust linear regression on background corrected photon counts $N_{\mathrm{R|G}}^{(1)} = f(N_{\mathrm{G|G}}^{(1)})$, where the proportionality constant corresponds to donor crosstalk, α (see eq. (30)). **(c)** Cross-validation of the obtained crosstalk value, by monitoring the position of D-only sub-population after BG and α correction. Ideally, D-only species should be at ${}^{ii*}E_{\mathrm{app}}^{(\mathrm{D0})} = 0$ and ${}^{ii*}S_{\mathrm{app}}^{(\mathrm{D0})} = 1$.

Control measurement. Furthermore, possible is to cross-validate obtained value from a separate nM measurement of D-only sample. Measurement of a free dye is not recommended due to the spectral shift occurring after coupling dye to biomolecule. From the background corrected mean count rates in detection channels G|G and R|G, one can calculate crosstalk as follows:

$$\alpha = \frac{S_{\mathrm{R|G}} - S_{\mathrm{R|G}}^{(\mathrm{BG})}}{S_{\mathrm{G|G}} - S_{\mathrm{G|G}}^{(\mathrm{BG})}} \tag{31}$$

Step 3.4: Direct excitation

Besides for donor crosstalk, fluorescence intensity of FRET sensitized acceptor emission, $F_{R|G}$, needs to be corrected for direct acceptor excitation by the donor excitation laser. The absence of this correction, leads to overestimation of measured FRET efficiency. [12, 17, 38] demonstrated, that as in the case of donor crosstalk, direct excitation has increasing effect towards low-*E* species. That being so, accurate distance measurements are only possible with meticulous determination of both the crosstalk and direct excitation.

As for crosstalk, determination of direct excitation is as well possible from the single-molecule measurement of the system of interest, using acceptor-only (A-only, A0) sub-population, selected using a 2D histogram of fully uncorrected apparent FRET efficiency and stoichiometry, ${}^{i}S_{\mathrm{app}} - {}^{i}E_{\mathrm{app}}$. ${}^{i}S_{\mathrm{app}}$ threshold should be set symmetrically around the center of A-only sub-population, while for ${}^{i}E_{\mathrm{app}}$ we recommend selecting higher FRET efficiencies (${}^{i}E_{\mathrm{app}} \geq 0.5$), given that these are associated with direct excitation of the acceptor (**Figure 5a**). Lower ${}^{i}E_{\mathrm{app}}$ values are hard to disentangle from species that have unstable emission, and which manifest as tailing between A-only and DA species. Following the selection of acceptor-only sub-population, similar as for donor crosstalk, historically it was recommended to histogram background corrected apparent stoichiometry values, ${}^{ii}S_{\mathrm{app}}^{(\mathrm{A0})}$ and fit that distribution with a Gaussian fit model. From obtained mean value, one calculates direct acceptor excitation as follows [11]:

$$\delta = \frac{\langle {}^{ii}S_{\mathrm{app}}^{(\mathrm{A0})} \rangle}{1 - \langle {}^{ii}S_{\mathrm{app}}^{(\mathrm{A0})} \rangle} \tag{32}$$

However, to avoid user bias associated with this approach, we suggest to instead perform robust linear regression between background corrected photon counts $N_{\mathrm{R|R}}^{(1)}$ and $N_{\mathrm{R|G}}^{(1)}$ (**Figure 5b**):

$$N_{\mathrm{R|G}} - S_{\mathrm{R|G}}^{(\mathrm{BG})} T = \delta (N_{\mathrm{R|R}} - S_{\mathrm{R|R}}^{(\mathrm{BG})} T) \tag{33}$$

Such determined direct excitation parameter, locates the acceptor-only sub-population at the ${}^{iii}S_{\mathrm{app}}^{(\mathrm{A0})} = 0$ and ${}^{iii}E_{\mathrm{app}}^{(\mathrm{A0})} = 0 \ldots 1$ (**Figure 5c**).

Control measurement. To avoid acceptor photobleaching, one commonly performs sm-PIE measurements with lower power of acceptor excitation laser, which on the other hand leaves us with smaller number of bursts in A-only population. This makes direct excitation determination

less robust compared to crosstalk determination. For that reason, in such circumstances, we speak for additional controls of fitted direct excitation parameter. Highly robust control would be to perform nM measurement of A-only sample. Use of free dye measurement should be avoided, since free dye can have a spectral shift compared to dye coupled to a biomolecule. From the background corrected mean count-rates in detection channels R|G and R|R, one can calculate direct excitation as follows:

$$\delta = \frac{S_{\mathrm{R|G}} - S_{\mathrm{R|G}}^{(\mathrm{BG})}}{S_{\mathrm{R|R}} - S_{\mathrm{R|R}}^{(\mathrm{BG})}} \tag{34}$$

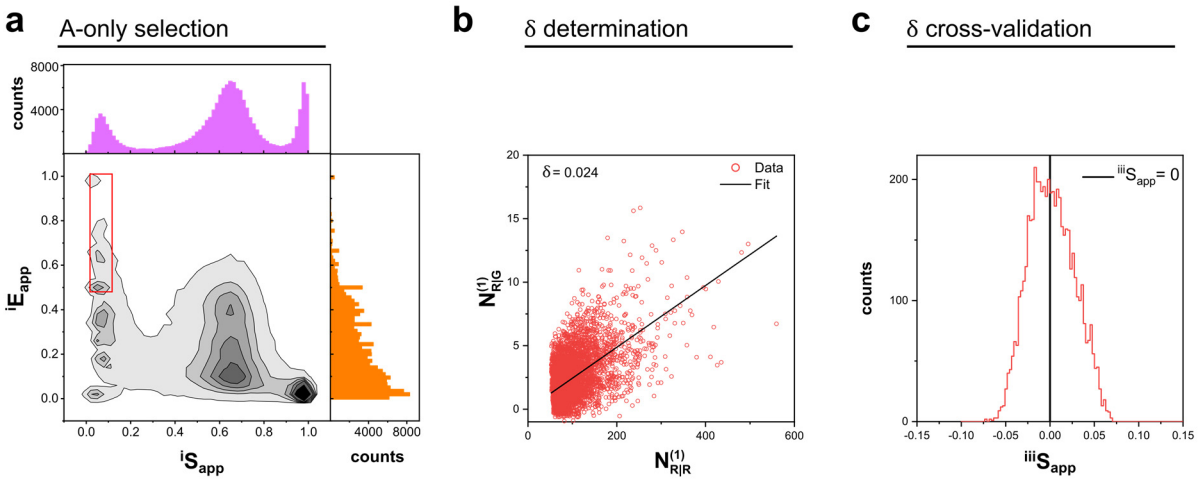


Figure 5. Step 3.4: Determination of acceptor direct excitation with donor excitation laser. (a) Selection of acceptor-only sub-population in ${}^{i}S_{\mathrm{app}} - {}^{i}E_{\mathrm{app}}$ 2D histogram (red rectangle). Recommended is to select ${}^{i}E_{\mathrm{app}} \geq 0.5$ to avoid bleached and blinking acceptor species that smear between DA and A-only sub-populations. ${}^{i}S_{\mathrm{app}}$ selection is made symmetrically around the center of the population. **(b)** Robust linear regression on background corrected photon counts $N_{\mathrm{R|G}}^{(1)} = f(N_{\mathrm{R|R}}^{(1)})$, where the proportionality constant corresponds to direct acceptor excitation, δ (see eq. (33)). **(c)** Cross-validation of the obtained δ value, by monitoring the position of A-only sub-population in ${}^{iii}S_{\mathrm{app}}^{(\mathrm{A0})}$ 1D histogram. Ideally, A-only species should be at ${}^{iii}S_{\mathrm{app}}^{(\mathrm{A0})} = 0$.

Step 3.5: Selection of DA sub-population via "ALEX-2CDE" filter

Detection efficiency γ and excitation ratio β are determined using DA species. Since photobleaching and photoblinking result in brightness fluctuations of the donor and/or acceptor fluorophore, it is essential to filter out these events, such that only those DA molecules that have stable emission of both fluorophores are used for γ and β determination. Selecting DA sub-population via stoichiometry cut is not the proper way to deal with this, because bleaching and blinking events manifest themselves as strong tailing between double-labelled, DA, and A-only and/or D-only sub-populations, causing the DA species of interest to be concealed.

It was found that bursts containing photobleached acceptor species can be removed by analyzing macroscopic photon arrival times.[11, 40] With stable acceptor emission, burst-averaged macroscopic photon arrival time in acceptor detection channel after donor excitation, $T_{R|G}$, is equal to burst-averaged macroscopic photon arrival time in donor detection channel after donor excitation, $T_{G|G}$, i.e. $T_{G|G} - T_{R|G} \approx 0$. However, in the presence of acceptor photobleaching within a burst, $T_{R|G}$ gets lower than $T_{G|G}$. This is the key feature for filtering out bursts with unstable acceptor emission based on average macro arrival times of photons. However, in this approach, it is very difficult to distinguish photobleaching from slow conformational dynamics. Therefore, by filtering bursts according to $T_{G|G} - T_{R|G}$ difference, we potentially eliminate very valuable bursts containing information on slow conformational dynamics. To overcome this drawback, it was proposed to instead monitor the difference between burst-averaged macroscopic arrival time of photons detected in acceptor detection channel after acceptor excitation, $T_{R|R}$, which are independent of FRET, and burst-averaged macroscopic photon arrival time of all photons after donor excitation, $T_{X|G}$. However, since background photons can be wrongly assigned to a burst at its edges, this approach suffers from using a simple average over macro arrival times of photons as it does not consider local count rate around each of the photons.

Convenient way to remove DA species with brightness fluctuations, as well single-labelled species, while maintaining enough bursts in the DA sub-population, is to apply **A**LEX **2**-**C**hannel kernel-based **D**ensity **E**stimator function, "ALEX-2CDE" (**Figure 6**), developed by Tomov et al[41]. The essence of this approach is in estimation of local density of photons around each of detected photons. Local density of photons is obtained using kernel density estimate of exponential kind, which is expressed in terms of photon arrival times. For density estimation, considered are only photon arrival times of those photons that arrive within $5 \times tw$ before and $5 \times tw$ after each central photon. Time-window, *tw*, used for estimation of local photon density

should be smaller than burst duration, but longer than inter-photon time. Considering those factors and also in terms of CPU time, we found the optimal choice of time-window to be 100 µs. For the fixed brightness of donor and acceptor, ALEX-2CDE filter converges to the value of 0. On the other hand, for fluctuating brightness of donor and/or acceptor, ALEX-2CDE > 0. After application of the ALEX-2CDE filter, no further ${}^{iii}E_{\mathrm{app}}$ or ${}^{iii}S_{\mathrm{app}}$ cuts should be applied to double labelled, DA, sub-population.

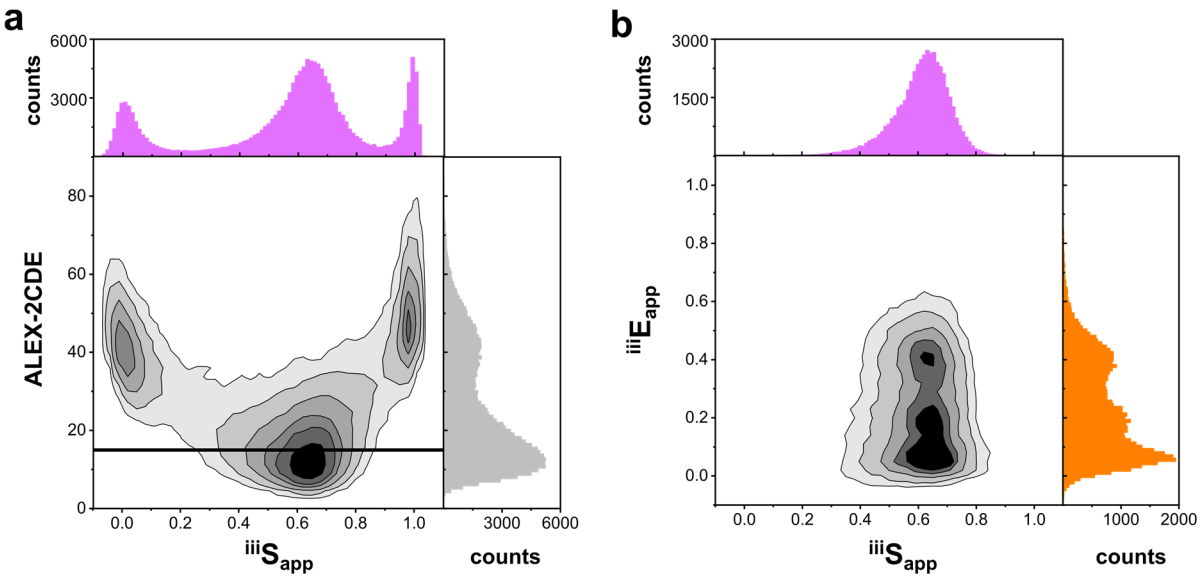


Figure 6. Step 3.5: Selection of DA population using ALEX-2CDE filter. (a) Selection of double-labelled population used for γ and β determination is made in 2D histogram of ALEX-2CDE filter values against ${}^{iii}S_{\mathrm{app}}$. Selection is made such to remove single labelled species, A-only and D-only, as well species with non-stable emission that smear between A-only/D-only and DA species. Here ALEX-2CDE ≤ 15 (horizontal black line). **(b)** Visualization of obtained DA population after applying ALEX-2CDE filter.

Step 3.6: Determination of γ and β correction factors

a) (**M-1**) *Non-linear regression of burst-wise* ${}^{iii}S_{app}$ *versus* ${}^{iii}E_{app}$

In this method for γ and β determination, the underlying relation between ${}^{iii}S_{\mathrm{app}}$ and ${}^{iii}E_{\mathrm{app}}$ is as follows:

$$ {}^{iii}S_{\mathrm{app}} = (1 + \gamma\beta + (1-\gamma)\beta\, {}^{iii}E_{\mathrm{app}})^{-1} \tag{35}$$

Due to the large spread of $1/{}^{iii}S_{\mathrm{app}}$ burst-wise values, instead of fitting the inverse of ${}^{iii}S_{\mathrm{app}}$ versus ${}^{iii}E_{\mathrm{app}}$, we perform direct non-linear regression of burst-wise values of ${}^{iii}S_{\mathrm{app}}$ versus ${}^{iii}E_{\mathrm{app}}$. Such fit is performed on DA populations of multiple samples with different FRET efficiencies (**Figure 7a**), or on a single sample that shows sufficient broadening of ${}^{iii}E_{\mathrm{app}}$. However, if acceptor properties are not uniform across the samples, then the homogeneous approximation does not apply. In that case, regression is performed either on the subset of DA populations, or local γ correction is applied. For the entire goodness-of-fit landscape see **Supplemental Note 4**, where two aforementioned scenarios are examined: global and local γ correction. Since the benchmark samples are DNA rulers with labelling stoichiometry D:A=1:1, and which are considered to be static, we do cross-validation of obtained γ and β parameters by monitoring if after the correction all DA populations lie on the S = 0.5 line and static E-τ FRET-line (**Figure 7b-c**).

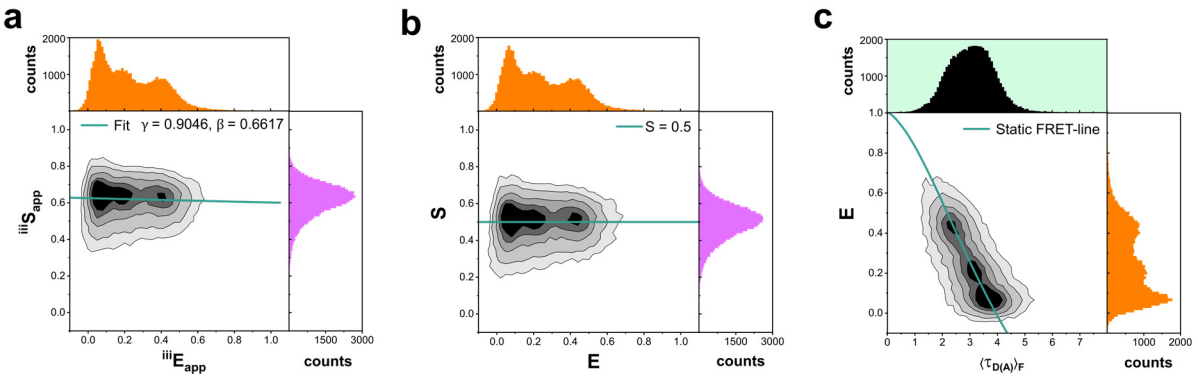


Figure 7. Step 3.6: γ and β determination via non-linear regression of burst-wise values of ${}^{iii}S_{app}$ **versus** ${}^{iii}E_{app}$ **(M-1). (a)** Non-linear regression on burst-wise values of ${}^{iii}S_{\mathrm{app}}$ versus ${}^{iii}E_{\mathrm{app}}$. **(b)** Control of obtained γ and β parameters, by monitoring whether the DA populations are centered at S = 0.5. **(c)** Control of obtained γ and β parameters, by monitoring the position of DA populations relative to the static FRET-line. For stiff molecule such as DNA, no shift from static FRET-line is expected, provided photophysical stability of fluorophores and the absence of position-specific dye-protein sticking interactions. For the derivation of static FRET-line and used parameters, see **Supplemental Note 5** and **Supplemental Table S5**.

b) (**M-2**) *Linear regression between* $1/\langle {}^{iii}S_{app}\rangle$ *and* $\langle {}^{iii}E_{app}\rangle$ *of population means*

As an alternative to performing non-linear regression of ${}^{iii}S_{\mathrm{app}}$ versus ${}^{iii}E_{\mathrm{app}}$, one could directly perform linear regression between $1/\langle {}^{iii}S_{\mathrm{app}}\rangle$ and $\langle {}^{iii}E_{\mathrm{app}}\rangle$, where $\langle {}^{iii}S_{\mathrm{app}}\rangle$ and $\langle {}^{iii}E_{\mathrm{app}}\rangle$ are the means of the DA sub-populations [17] (**Figure 8a**)

$$1/\langle {}^{iii}S_{\mathrm{app}}\rangle = a + b\langle {}^{iii}E_{\mathrm{app}}\rangle \tag{36}$$

Correction factors γ and β are then obtained from the intercept, *a*, and the slope, *b*, of the linear fit (**Figure 8b**)

$$\beta = a + b - 1 \tag{37}$$

$$\gamma = (a - 1)/(a + b - 1) \tag{38}$$

If the population means do not follow straight line, this is indication that homogeneous approximation does not hold, and if possible, regression should be applied on a subset of species (minimum two). Otherwise, method is not applicable. While on one hand this method has advantage of not being sensitive to shoulders in the distributions, as is the case when fitting burst-wise values of ${}^{iii}S_{\mathrm{app}}$ versus ${}^{iii}E_{\mathrm{app}}$, there is a user bias while performing Gaussian fit to the subpopulations, and when assigning the main population. Furthermore, it was previously shown that the uncertainty in such determined γ parameter is the function of inverse difference between the population mean estimates, but this only makes notable effect at extreme values of γ and β, i.e. < 0.1 and >3.[13] For intermediate values of γ and β, where most measurement are performed, FRET efficiency contrast between species does not play a major role, as it was experimentally observed in earlier study.[33]

Analogously to previous method, cross-validation of obtained γ and β parameters is performed by monitoring if after the correction all DA populations lie on the S = 0.5 line and static E-τ FRET-line (**Figure 8c-d**).

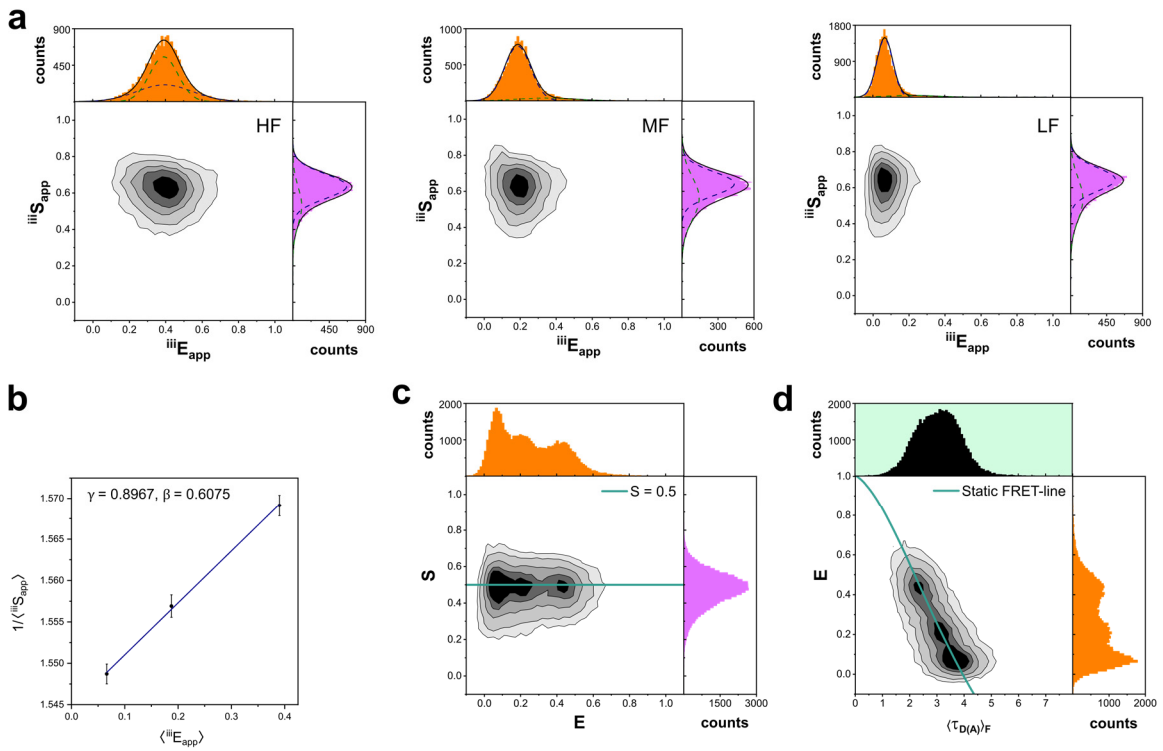


Figure 8. Step 3.6: γ and β determination via linear regression on population means (M-2). (a) 1D $^{iii}S_{app}$ and $^{iii}E_{app}$ histograms are fitted using the superposition of *N* normal distributions. For *N>1*, Gaussian with maximum amplitude is taken as the main population **(b)** Fit of $1/\langle^{iii}S_{app}\rangle$ versus $\langle^{iii}E_{app}\rangle$ of DA population means. γ and β are determined using slope and intercept of linear fit using eqs. (36)-(38). **(c)** and **(d)** Control of obtained γ and β parameters, by monitoring whether the DA populations are centered at S=0.5 and if they lie on static FRET-line. For the derivation of static FRET-line and used parameters, see **Supplemental Note 5** and **Supplemental Table S5**.

c) (**M-3**) *Linear regression on photon counts,* $N_{R|R}^{(1)} = f\left(N_{G|G}^{(1)},\ N_{R|G}^{(2)}\right)$

Instead of using ratiometric quantities S_{app} and E_{app} one can directly use underlying photon counts, as described previously.[42] The underlying equation of this method is obtained from the definition of stoichiometry:

$$S = \frac{\gamma N_{\mathrm{G|G}}^{(1)} + N_{\mathrm{R|G}}^{(2)}}{\gamma N_{\mathrm{G|G}}^{(1)} + N_{\mathrm{R|G}}^{(2)} + \frac{1}{\beta} N_{\mathrm{R|R}}^{(1)}} \tag{39}$$

by rewriting it as follows

$$\frac{S}{1-S} N_{\mathrm{R|R}}^{(1)} = \gamma\beta N_{\mathrm{G|G}}^{(1)} + \beta N_{\mathrm{R|G}}^{(2)} \tag{40}$$

It can be recognized that this represents equation of a plane in 3D space, defined by $[N_{\mathrm{G|G}}^{(1)}, N_{\mathrm{R|G}}^{(2)}, N_{\mathrm{R|R}}^{(1)}]$. However, in the next step, assumption about the stoichiometry of the sample has to be made. For D:A =1:1, this equation reduces to

$$N_{\mathrm{R|R}}^{(1)} = \gamma\beta N_{\mathrm{G|G}}^{(1)} + \beta N_{\mathrm{R|G}}^{(2)} \tag{41}$$

where multiplication factors γ and β are determined using plane fit (**Figure 9a**).
Although eq. (41) can be mathematically reformulated as $N_{\mathrm{G|G}}^{(1)} = f(N_{\mathrm{R|G}}^{(2)}, N_{\mathrm{R|R}}^{(1)})$ or $N_{\mathrm{R|G}}^{(2)} = f(N_{\mathrm{G|G}}^{(1)}, N_{\mathrm{R|R}}^{(1)})$, that way we add another dependency in addition to the one we want to extract, since $N_{\mathrm{G|G}}^{(1)}$ and $N_{\mathrm{R|G}}^{(2)}$ are correlated due to FRET. Therefore, for determination of γ and β parameters we recommend to use eq. (41), with $N_{\mathrm{R|R}}^{(1)}$ on *z*-axis.

Potential drawback of this method is that it requires to know the donor:acceptor labelling stoichiometry of the sample in advance. However, for the constructs used for intramolecular FRET experiments, this information is in general known. Otherwise, separate construct with known stoichiometry has to be measured, with the same FRET pair, and with the same setup conditions. Cross-validation of obtained γ and β parameters is given in **Figure 9b-c.**

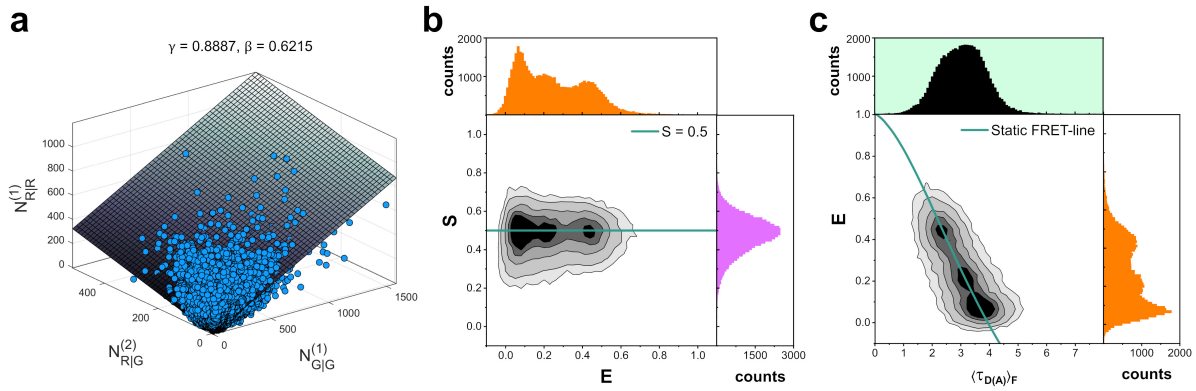


Figure 9. Step 3.6: γ and β determination via linear regression on photon counts (M-3). (a) Plane fit in 3D space of photon counts $[N_{G|G}^{(1)}, N_{R|G}^{(2)}, N_{R|R}^{(1)}]$ according to (41). **(b)** and **(c)** Control of obtained γ and β parameters, by monitoring whether the DA populations are centered at S=0.5 and if they lie on static FRET-line. For the derivation of static FRET-line and used parameters, see **Supplemental Note 5** and **Supplemental Table S5**.

Assessment of the extent of user bias in calibration workflow

To assess the user bias when determining correction parameters, we performed a blind challenge amongst seven lab members. Participants were provided with a set of DNA measurements for a global calibration (i.e. HF-, MF- and LF-sample), and calibration protocol, presented in this work. Distribution of reported values for calibration parameters is shown in **Figure 10**. Even though that somewhat bigger spread of ALEX-2CDE threshold values were obtained, in all cases threshold was properly selected, and DA sub-populations were cleaned-up from single-labelled species and species with brightness fluctuations. However, due to different ALEX-2CDE thresholds applied, we observe inconsistencies in determination of γ and β parameters using the method **M-2** of *Linear regression between* $1/\langle {}^{iii}S_{\mathrm{app}}\rangle$ *and* $\langle {}^{iii}E_{\mathrm{app}}\rangle$ *of population means* ((36)). ALEX-2CDE filter directly affects the mean of the DA population, and on top of that users were inconsistent in the decision of how many Gaussians should be fitted to the DA sub-population, as well in the assignment of the main population. Especially troublesome is fitting of stoichiometry distribution, where asymmetry of the distribution is often observed. Other than ALEX-2CDE filter value, and γ and β calibration method **M-2**, in other cases we found rather consistent results.

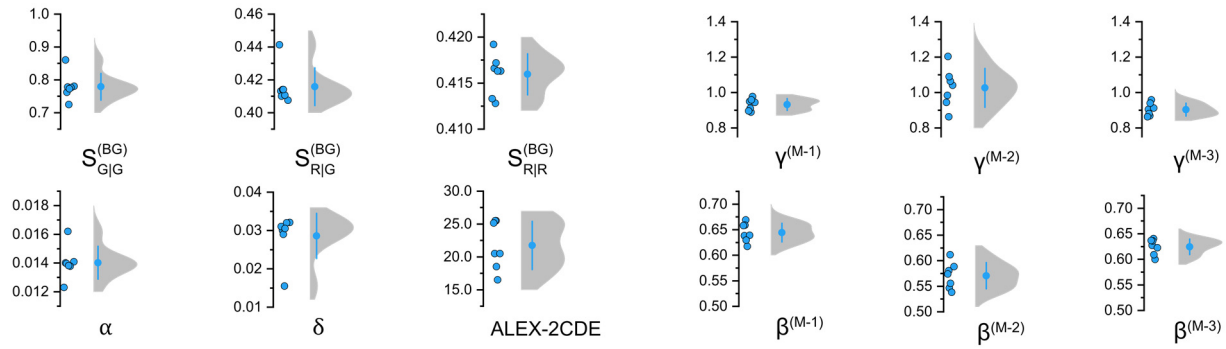


Figure 10. Results of a blind calibration challenge. Reported values of correction parameters by lab members (N=7). Blue circle in the distribution plot denotes mean of the distribution, while whiskers span a range of one standard deviation. Background count rates are reported in kHz. Superscript for γ and β denotes different calibration method: **M-1:** Non-linear regression of burst-wise ${}^{iii}S_{\mathrm{app}}$ versus ${}^{iii}E_{\mathrm{app}}$ values; **M-2:** Linear regression between $1/\langle {}^{iii}S_{\mathrm{app}}\rangle$ and $\langle {}^{iii}E_{\mathrm{app}}\rangle$ of population means; **M-3:** linear regression on photon counts, $N_{\mathrm{R|R}}^{(1)} = f\left(N_{\mathrm{G|G}}^{(1)}, N_{\mathrm{R|G}}^{(2)}\right)$. Results of blind calibration challenge revealed that method of linear regression between $1/\langle {}^{iii}S_{\mathrm{app}}\rangle$ and $\langle {}^{iii}E_{\mathrm{app}}\rangle$ of population means (M-2) should be applied with caution, since we observed high user bias when determining the mean of the DA populations.

Evaluation of calibration parameters accuracy

a) Static FRET-line

When all levels of correction are performed (background, α, δ, γ and β) no displacement of the DA population from the static FRET-line should be observed, considering dsDNA molecule to be static and provided photophysical stability of the fluorophores and absence of dye sticking. For detailed description and derivation of equation underlying static FRET-line see **Supplemental Note 5**. In this work, we use static FRET-line to evaluate the accuracy of different calibration methods. We determined apparent dynamic shift of the DA population from the static FRET-line, ds_{app}, for each of the calibration methods, as described previously[13] and in **Supplemental Note 6**. In **Figure 11** illustrated is determination of ds_{app} values for a set of DNA rulers labelled with AF488-Atto647N using method M-1. In **Figure 12a** provided are ds_{app} box plots for all three dye combinations and three methods for γ and β determination. Determined ds_{app} values comply with the static nature of the measured system.

Lastly, it is worth to mention that such evaluation of calibration methods accuracy carries assumption of σ_{DA} = 6 Å linker-induced broadening of inter-dye distance distribution. Ideally, width of R_{DA} distribution when computing static FRET-line should be determined individually for each system under the study, by performing Gaussian fit of ensemble-TCSPC decays, with variance of the $p(R_{\mathrm{DA}})$ distribution being the fit parameter (see **Supplemental Note 5**). However, such approach typically overestimates broadening of the inter-dye distance distribution, due to the donor-only contribution. Alternative would be to use sub-ensemble TCSPC decays of DA sub-population to restrain the fitting.

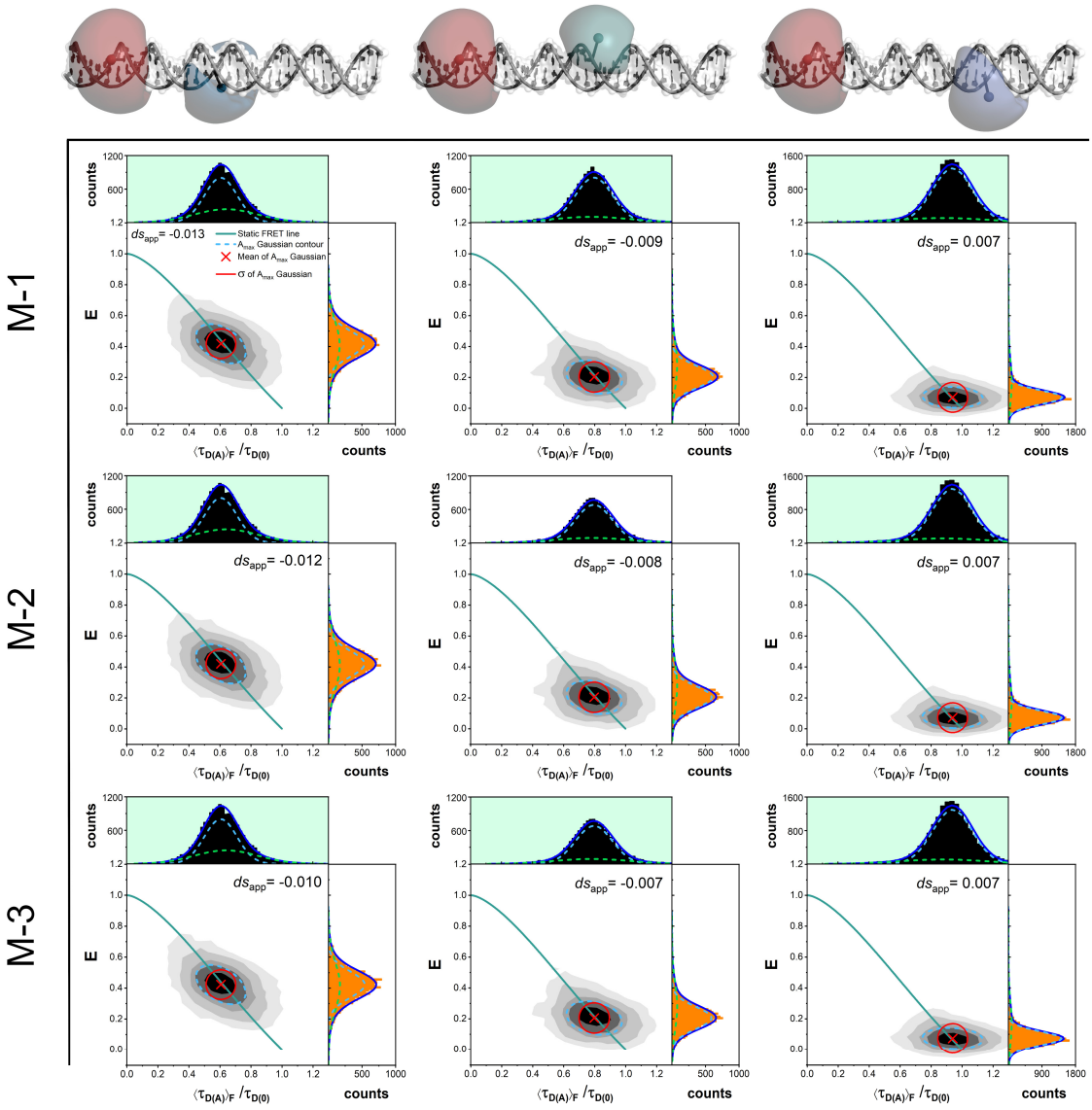


Figure 11. Apparent dynamic shift for different methods of γ and β determination. Displacement of DA sub-populations from static FRET-line was determined for HF (left column), MF (middle column) and LF sample (right column) labelled with AF488-Atto 647N, and for different methods of γ and β determination, namely: M-1: Non-linear regression of burst-wise ${}^{iii}S_{\mathrm{app}}$ versus ${}^{iii}E_{\mathrm{app}}$ values; M-2: Linear regression between $1/\langle {}^{iii}S_{\mathrm{app}}\rangle$ and $\langle {}^{iii}E_{\mathrm{app}}\rangle$ of population means; M-3: linear regression on photon counts, $N_{\mathrm{R|R}}^{(1)} = f\left(N_{\mathrm{G|G}}^{(1)},\ N_{\mathrm{R|G}}^{(2)}\right)$. For the derivation of static FRET-line and used parameters, see **Supplemental Note 5** and **Supplemental Table S5**.

b) Structural models

In order to assess how well the experimentally determined FRET observables agree with structural model of dsDNA used in this study, we simulated Accessible Volume (AV)[19] of the fluorophores using Olga software.[43] See **Supplemental Note 7** for detailed description of AV simulations and used parameters. In **Figure 12b** plotted are experimental versus modelled E-averaged inter-dye distances, $R_{\langle E \rangle}$. Consistent distances are obtained using independent measurements with three different dye pairs, and using different calibration methods. RMSD = 3.31 Å between experimental and modelled $R_{\langle E \rangle}$ distances agrees with what was reported earlier for dsDNA molecules (RMSD = 3.34 Å for static DNA model and RMSD = 2.74 Å for dynamic DNA model[12]).

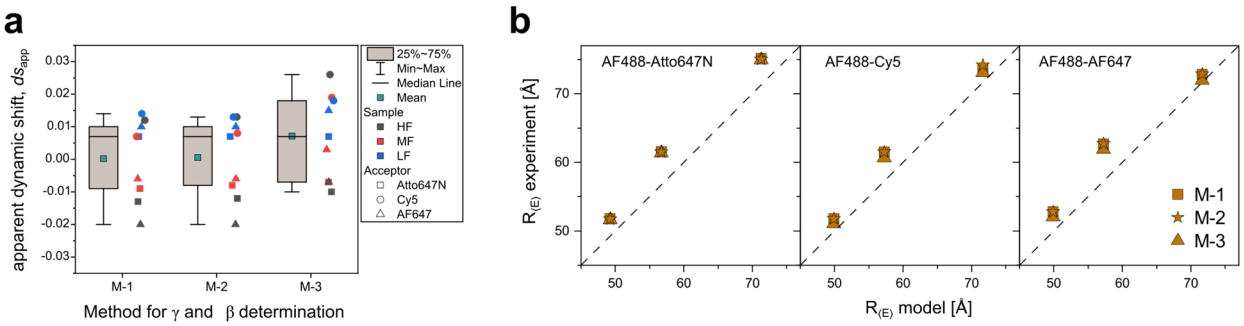


Figure 12. Assessing the accuracy of calibration parameters. (a) Box plots of dynamic shift values determined for HF, MF and LF samples with three different dye combinations and for different calibration methods. Box length denotes interquartile range, horizontal line inside a box is median, while blue square is the mean of distribution. Whiskers spread from minimum to maximum observed value. **(b)** Agreement between measured and modelled FRET-efficiency averaged inter-dye distances, showing that the same distances are recovered irrespective of dye pair or calibration method used.

5. Conclusions

In this study we presented robust and unbiased workflow for estimation of calibration parameters, allowing for determination of accurate FRET efficiency and stoichiometry. As estimation of each of correction parameters is preceded with selection of D-only, A-only or DA sub-population, and neither of the existing calibration methods gives clear guidelines on the robust selection of those, we refined existing approaches, such that for each of the steps in calibration, proposed are clear, unambiguous instructions. As a benchmark system we used a set of dsDNA rulers with varying FRET efficiencies, such that position of acceptor dye is fixed, and the one of donor is varied. This allowed global γ correction, since donor dye has uniform properties even in distinct environments. We primarily focused on determination of γ and β parameters, and for that we put on a test three commonly used methods: (M-1) Non-linear regression of burst-wise values of ${}^{iii}S_{\mathrm{app}}$ versus ${}^{iii}E_{\mathrm{app}}$, (M-2) Linear regression between $1/\langle {}^{iii}S_{\mathrm{app}}\rangle$ and $\langle {}^{iii}E_{\mathrm{app}}\rangle$ of the population means, (M-3) Linear regression on photon counts. We presented their theoretical foundations, application versatility and assessed their performance in the blind calibration challenge.

When applicable, users should evaluate different methods and check for their consistency, rather than to exclusively use one of the methods. Method of non-linear regression of burst-wise values of ${}^{iii}S_{\mathrm{app}}$ versus ${}^{iii}E_{\mathrm{app}}$ (M-1) and method of linear regression of $1/\langle {}^{iii}S_{\mathrm{app}}\rangle$ versus $\langle {}^{iii}E_{\mathrm{app}}\rangle$ (M-2) depend on ratiometric quantities of unknown statistics, and are sensitive to shoulders in the distribution. Furthermore, through the blind calibration challenge, we identified that method M-2, is particularly prone to user bias. We found that this method is the most sensitive on the way DA sub-population is selected, since ALEX-2CDE filter directly shifts the mean of DA species in ${}^{iii}S_{\mathrm{app}}$ - ${}^{iii}E_{\mathrm{app}}$ histogram. Furthermore, users were inconsistent with the fit model and at assigning the main population. On contrary, methods M-1 and M-3 give more narrow distribution of γ and β values.

After evaluating the extent of the user bias, in the next step we assessed the accuracy of the calibration methods: (*i*) by using static FRET-line as a reference and by calculating the displacement of DA sub-populations from the static FRET-line via apparent dynamic shift, ds_{app} (*ii*) by comparing derived and modelled FRET efficiency averaged inter-dye distances. For these two criteria, we did not observe any significant distinction between the calibration methods, i.e. they all give consistent inter-dye distances irrespective of the dye-pair used, and they comply with the static nature of the system, since obtained ds_{app} values are in all cases negligibly small.

However, even though that the method of linear regression on photon counts (M-3) seems to tick all boxes, its pitfall is that it requires stoichiometry of a FRET construct to be known. If stoichiometry is to be determined from the experiment, then calibration has to be performed on independent FRET construct of known stoichiometry, measured with the same fluorophore pair and under the same setup conditions.

Finally, we devoted particular attention to setup alignment and calibration, with elaborate instructions for handling entire setup configuration. As ultimate control of optimal setup performance, we proposed framework for regularized optimization of instrument detection profile, using a set of dyes emitting in both spectral windows, which are measured free in solution. We identify this approach as practical way of evaluating setup performance, i.e. for identifying potential degradation of optical components or setup misalignment.

Acknowledgements

Olga Doroshenko, Oleg Opanasyuk, Jakub Kubiak, Alexander Larbig, Bianca Reschke and Noah Salama are thanked for their participation in blind calibration challenge.

Associated content

Supporting information. This material is available free of charge via the internet at xxx.

Author information

Corresponding author

Corresponding Author: Claus A. M. Seidel

Address: Institute for Molecular Physical Chemistry, Heinrich Heine University Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany

Phone: (+49) 211 81 15881; Fax: (+49) 211 81 11612803

E-mail: cseidel@hhu.de

Author Contributions

M.P. performed measurements, data analysis, made the figures and wrote the original draft of the manuscript. S.F. performed software development. J.F., M.P. and C.A.M.S ordered and managed samples. J.F. performed pilot measurements for the project. All authors reviewed the manuscript. A.B. and C.A.M.S conceptualized and supervised the project

ORCID

Milana Popara: 0000-0003-2626-6096

Anders Barth: 0000-0003-3671-3072

Julian Folz: 0000-0001-8232-8787

Suren Felekyan: 0000-0002-0554-4520

Claus A. M. Seidel: 0000-0002-5171-149X

Notes

The authors declare no competing financial interest.

References

1. Gansen, A.; Felekyan, S.; Kuhnemuth, R.; Lehmann, K.; Toth, K.; Seidel, C. A. M.; Langowski, J., High precision FRET studies reveal reversible transitions in nucleosomes between microseconds and minutes. *Nat Commun* **2018,** *9*.

2. Sanabria, H.; Rodnin, D.; Hemmen, K.; Peulen, T. O.; Felekyan, S.; Fleissner, M. R.; Dimura, M.; Koberling, F.; Kuhnemuth, R.; Hubbell, W.; Gohlke, H.; Seidel, C. A. M., Resolving dynamics and function of transient states in single enzyme molecules. *Nat Commun* **2020,** *11* (1).

3. Kilic, S.; Felekyan, S.; Doroshenko, O.; Boichenko, I.; Dimura, M.; Vardanyan, H.; Bryan, L. C.; Arya, G.; Seidel, C. A. M.; Fierz, B., Single-molecule FRET reveals multiscale chromatin dynamics modulated by HP1 alpha. *Nat Commun* **2018,** *9*.

4. Margittai, M.; Widengren, J.; Schweinberger, E.; Schroder, G. F.; Felekyan, S.; Haustein, E.; Konig, M.; Fasshauer, D.; Grubmuller, H.; Jahn, R.; Seidel, C. A. M., Single-molecule fluorescence resonance energy transfer reveals a dynamic equilibrium between closed and open conformations of syntaxin 1. *P Natl Acad Sci USA* **2003,** *100* (26), 15516-15521.

5. Hamilton, G. L.; Saikia, N.; Basak, S.; Welcome, F. S.; Wu, F.; Kubiak, J.; Zhang, C. C.; Hao, Y.; Seidel, C. A.; Ding, F.; Sanabria, H.; Bowen, M. E., Fuzzy supertertiary interactions within PSD-95 enable ligand binding. *Elife* **2022,** *11*.

6. Gomes, G. N. W.; Krzeminski, M.; Namini, A.; Martin, E. W.; Mittag, T.; Head-Gordon, T.; Forman-Kay, J. D.; Gradinaru, C. C., Conformational Ensembles of an Intrinsically Disordered Protein Consistent with NMR, SAXS, and Single-Molecule FRET. *J Am Chem Soc* **2020,** *142* (37), 15697-15710.

7. Naudi-Fabra, S.; Tengo, M.; Jensen, M. R.; Blackledge, M.; Milles, S., Quantitative Description of Intrinsically Disordered Proteins Using Single-Molecule FRET, NMR, and SAXS. *J Am Chem Soc* **2021,** *143* (48), 20109-20121.

8. Clegg, R. M., Fluorescence Resonance Energy-Transfer and Nucleic-Acids. *Methods in Enzymology* **1992,** *211*, 353-388.

9. Muller, B. K.; Zaychikov, E.; Brauchle, C.; Lamb, D. C., Pulsed interleaved excitation. *Biophys J* **2005,** *89* (5), 3508-3522.

10. Hendrix, J.; Lamb, D. C., Pulsed Interleaved Excitation: Principles and Applications. *Fluorescence Fluctuation Spectroscopy (Ffs), Pt A* **2013,** *518*, 205-243.

11. Kudryavtsev, V.; Sikor, M.; Kalinin, S.; Mokranjac, D.; Seidel, C. A. M.; Lamb, D. C., Combining MFD and PIE for Accurate Single-Pair Forster Resonance Energy Transfer Measurements. *Chemphyschem* **2012,** *13* (4), 1060-1078.

12. Hellenkamp, B.; Schmid, S.; Doroshenko, O.; Opanasyuk, O.; Kuhnemuth, R.; Adariani, S. R.; Ambrose, B.; Aznauryan, M.; Barth, A.; Birkedal, V.; Bowen, M. E.; Chen, H. T.; Cordes, T.; Eilert, T.; Fijen, C.; Gebhardt, C.; Gotz, M.; Gouridis, G.; Gratton, E.; Ha, T.; Hao, P. Y.; Hanke, C. A.; Hartmann, A.; Hendrix, J.; Hildebrandt, L. L.; Hirschfeld, V.; Hohlbein, J.; Hua, B. Y.; Hubner, C. G.; Kallis, E.; Kapanidis, A. N.; Kim, J. Y.; Krainer, G.; Lamb, D. C.; Lee, N. K.; Lemke, E. A.; Levesque, B.; Levitus, M.; McCann,

J. J.; Naredi-Rainer, N.; Nettels, D.; Ngo, T.; Qiu, R. Y.; Robb, N. C.; Rocker, C.; Sanabria, H.; Schlierf, M.; Schroder, T.; Schuler, B.; Seidel, H.; Streit, L.; Thurn, J.; Tinnefeld, P.; Tyagi, S.; Vandenberk, N.; Vera, A. M.; Weninger, K. R.; Wunsch, B.; Yanez-Orozco, I. S.; Michaelis, J.; Seidel, C. A. M.; Craggs, T. D.; Hugel, T., Precision and accuracy of single-molecule FRET measurements-a multi-laboratory benchmark study (vol 15, pg 984, 2018). *Nat Methods* **2018,** *15* (11), 984-984.

13. Agam, G.; Gebhardt, C.; Popara, M.; Machtel, R.; Folz, J.; Ambrose, B.; Chamachi, N.; Chung, S. Y.; Craggs, T. D.; de Boer, M.; Grohmann, D.; Ha, T.; Hartmann, A.; Hendrix, J.; Hirschfeld, V.; Hubner, C. G.; Hugel, T.; Kammerer, D.; Kang, H. S.; Kapanidis, A. N.; Krainer, G.; Kramm, K.; Lemke, E. A.; Lerner, E.; Margeat, E.; Martens, K.; Michaelis, J.; Mitra, J.; Moya Munoz, G. G.; Quast, R. B.; Robb, N. C.; Sattler, M.; Schlierf, M.; Schneider, J.; Schroder, T.; Sefer, A.; Tan, P. S.; Thurn, J.; Tinnefeld, P.; van Noort, J.; Weiss, S.; Wendler, N.; Zijlstra, N.; Barth, A.; Seidel, C. A. M.; Lamb, D. C.; Cordes, T., Reliability and accuracy of single-molecule FRET studies for characterization of structural dynamics and distances in proteins. *Nat Methods* **2023,** *20* (4), 523-535.

14. Barth, A.; Opanasyuk, O.; Peulen, T. O.; Felekyan, S.; Kalinin, S.; Sanabria, H.; Seidel, C. A. M., Unraveling multi-state molecular dynamics in single-molecule FRET experiments. I. Theory of FRET-lines. *J Chem Phys* **2022,** *156* (14).

15. Opanasyuk, O.; Barth, A.; Peulen, T. O.; Felekyan, S.; Kalinin, S.; Sanabria, H.; Seidel, C. A. M., Unraveling multi-state molecular dynamics in single-molecule FRET experiments. II. Quantitative analysis of multi-state kinetic networks. *J Chem Phys* **2022,** *157* (3), 031501.

16. Kapanidis, A. N.; Lee, N. K.; Laurence, T. A.; Doose, S.; Margeat, E.; Weiss, S., Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *P Natl Acad Sci USA* **2004,** *101* (24), 8936-8941.

17. Lee, N. K.; Kapanidis, A. N.; Wang, Y.; Michalet, X.; Mukhopadhyay, J.; Ebright, R. H.; Weiss, S., Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation. *Biophys J* **2005,** *88* (4), 2939-2953.

18. Kalinin, S.; Valeri, A.; Antonik, M.; Felekyan, S.; Seidel, C. A. M., Detection of Structural Dynamics by FRET: A Photon Distribution and Fluorescence Lifetime Analysis of Systems with Multiple States. *J Phys Chem B* **2010,** *114* (23), 7983-7995.

19. Sindbert, S.; Kalinin, S.; Hien, N.; Kienzler, A.; Clima, L.; Bannwarth, W.; Appel, B.; Muller, S.; Seidel, C. A. M., Accurate Distance Determination of Nucleic Acids via Forster Resonance Energy Transfer: Implications of Dye Linker Length and Rigidity. *J Am Chem Soc* **2011,** *133* (8), 2463-2480.

20. Enderlein, J.; Gregor, I.; Patra, D.; Dertinger, T.; Kaupp, U. B., Performance of fluorescence correlation spectroscopy for measuring diffusion and concentration. *Chemphyschem* **2005,** *6* (11), 2324-2336.

21. Hess, S. T.; Webb, W. W., Focal volume optics and experimental artifacts in confocal fluorescence correlation spectroscopy. *Biophys J* **2002,** *83* (4), 2300-2317.

22. Rigler, R.; Mets, U.; Widengren, J.; Kask, P., Fluorescence Correlation Spectroscopy with High Count Rate and Low-Background - Analysis of Translational Diffusion. *Eur Biophys J Biophy* **1993,** *22* (3), 169-175.

23. Gendron, P. O.; Avaltroni, F.; Wilkinson, K. J., Diffusion Coefficients of Several Rhodamine Derivatives as Determined by Pulsed Field Gradient-Nuclear Magnetic Resonance and Fluorescence Correlation Spectroscopy. *J Fluoresc* **2008,** *18* (6), 1093-1101.

24. Aragon, S. R.; Pecora, R., Fluorescence Correlation Spectroscopy as a Probe of Molecular-Dynamics. *J Chem Phys* **1976,** *64* (4), 1791-1803.

25. Liu, M. W.; Jia, M. H.; Pan, H. F.; Li, L.; Chang, M. F.; Ren, H.; Argoul, F.; Zhang, S. J.; Xu, J. H., Instrument Response Standard in Time-Resolved Fluorescence Spectroscopy at Visible Wavelength: Quenched Fluorescein Sodium. *Appl Spectrosc* **2014,** *68* (5), 577-583.

26. Sisamakis, E.; Valeri, A.; Kalinin, S.; Rothwell, P. J.; Seidel, C. A. M., Accurate Single-Molecule Fret Studies Using Multiparameter Fluorescence Detection. *Method Enzymol* **2010,** *475*, 455-514.

27. Koshioka, M.; Sasaki, K.; Masuhara, H., Time-Dependent Fluorescence Depolarization Analysis in 3-Dimensional Microspectroscopy. *Appl Spectrosc* **1995,** *49* (2), 224-228.

28. Schaffer, J.; Volkmer, A.; Eggeling, C.; Subramaniam, V.; Striker, G.; Seidel, C. A. M., Identification of single molecules in aqueous solution by time-resolved fluorescence anisotropy. *J Phys Chem A* **1999,** *103* (3), 331-336.

29. Panchuk-Voloshina, N.; Haugland, R. P.; Bishop-Stewart, J.; Bhalgat, M. K.; Millard, P. J.; Mao, F.; Leung, W. Y.; Haugland, R. P., Alexa dyes, a series of new fluorescent dyes that yield exceptionally bright, photostable conjugates. *J Histochem Cytochem* **1999,** *47* (9), 1179-88.

30. Maillard, J.; Klehs, K.; Rumble, C.; Vauthey, E.; Heilemann, M.; Furstenberg, A., Universal quenching of common fluorescent probes by water and alcohols. *Chem Sci* **2021,** *12* (4), 1352-1362.

31. Kohn, F.; Hofkens, J.; Gronheid, R.; Van der Auweraer, M.; De Schryver, F. C., Parameters influencing the on- and off-times in the fluorescence intensity traces of single cyanine dye molecules. *J Phys Chem A* **2002,** *106* (19), 4808-4814.

32. Widengren, J.; Schwille, P., Characterization of photoinduced isomerization and back-isomerization of the cyanine dye Cy5 by fluorescence correlation spectroscopy. *J Phys Chem A* **2000,** *104* (27), 6416-6428.

33. Vandenberk, N.; Barth, A.; Borrenberghs, D.; Hofkens, J.; Hendrix, J., Evaluation of Blue and Far-Red Dye Pairs in Single-Molecule Forster Resonance Energy Transfer Experiments. *J Phys Chem B* **2018,** *122* (15), 4249-4266.

34. Gopich, I. V., Concentration effects in "Single-Molecule" spectroscopy. *J Phys Chem B* **2008,** *112* (19), 6214-6220.

35. Edman, L.; Mets, U.; Rigler, R., Conformational transitions monitored for single molecules in solution. *P Natl Acad Sci USA* **1996,** *93* (13), 6710-6715.

36. Fries, J. R.; Brand, L.; Eggeling, C.; Kollner, M.; Seidel, C. A. M., Quantitative identification of different single molecules by selective time-resolved confocal fluorescence spectroscopy. *J Phys Chem A* **1998,** *102* (33), 6601-6613.

37. Ingargiola, A.; Lerner, E.; Chung, S.; Weiss, S.; Michalet, X., FRETBursts: An Open Source Toolkit for Analysis of Freely-Diffusing Single-Molecule FRET. *Plos One* **2016,** *11* (8).

38. Peulen, T. O.; Opanasyuk, O.; Seidel, C. A. M., Combining Graphical and Analytical Methods with Molecular Simulations To Analyze Time-Resolved FRET Measurements of Labeled Macromolecules Accurately. *J Phys Chem B* **2017,** *121* (35), 8211-8241.

39. DuMouchel, W. H., and F. L. O'Brien In *Integrating a Robust Option into a Multiple Regression Computing Environment*, Computer Science and Statistics: Proceedings of the 21st Symposium on the Interface, Alexandria, VA, 1989; American Statistical Association: Alexandria, VA, 1989.

40. Eggeling, C.; Widengren, J.; Brand, L.; Schaffer, J.; Felekyan, S.; Seidel, C. A. M., Analysis of photobleaching in single-molecule multicolor excitation and forster resonance energy transfer measurement. *J Phys Chem A* **2006,** *110* (9), 2979-2995.

41. Tomov, T. E.; Tsukanov, R.; Masoud, R.; Liber, M.; Plavner, N.; Nir, E., Disentangling Subpopulations in Single-Molecule FRET and ALEX Experiments with Photon Distribution Analysis. *Biophys J* **2012,** *102* (5), 1163-1173.

42. Coullomb, A.; Bidan, C. M.; Qian, C.; Wehnekamp, F.; Oddou, C.; Albiges-Rizo, C.; Lamb, D. C.; Dupont, A., QuanTI-FRET: a framework for quantitative FRET measurements in living cells. *Sci Rep-Uk* **2020,** *10* (1).

43. Dimura, M.; Peulen, T. O.; Sanabria, H.; Rodnin, D.; Hemmen, K.; Hanke, C. A.; Seidel, C. A. M.; Gohlke, H., Automated and optimally FRET-assisted structural modeling. *Nat Commun* **2020,** *11* (1).

SUPPORTING INFORMATION

Increasing the accuracy of single-molecule FRET experiments - robust and unbiased estimation of correction parameters

Milana Popara[1], Anders Barth[1,a], Julian Folz[1], Suren Felekyan[1], Oleg Opansyuk[1], Claus A. M. Seidel[1,*]

[1]Institut für Physikalische Chemie, Lehrstuhl für Molekulare Physikalische Chemie, Heinrich-Heine-Universität, Düsseldorf, Germany
[a]Present address: Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Delft, Netherlands

[*]Author to whom correspondence should be addressed:

Claus A. M. Seidel, Phone: (+49) 211 81 15881, E-Mail: cseidel@hhu.de

Keywords: PIE, smFRET, correction parameters, calibration

Table of Contents

Supplemental Notes

Supplemental Note 1: Experimental setup and burst analysis

All sample solutions were measured in NUNC chambers (Lab-Tek, Thermo Scientific) with 500 µL sample volume at a concentration of ~10 pM. Single-molecule MFD measurements were performed at a room temperature on a homebuilt, dual-color and dual-polarization confocal setup, as described previously.[1] Donor molecules (AlexaFluor488) were excited by a pulsed diode laser (LDH-D-C 485, PicoQuant) at 485 nm, operated with repetition frequency of 32 MHz (16 MHz for samples with Atto647N as acceptor dye due to its longer fluorescence lifetime). Acceptor molecules (AlexaFluor647/Cy5/Atto647N) were excited with pulsed diode laser (LDH-D-C 640, PicoQuant) at 635 nm. Laser powers were measured at the sample and were 60 µW for donor excitation laser and 10 µW for acceptor excitation laser (20 µW for samples with Atto647N due to the smaller extinction coefficient compared to other acceptor dyes). Laser pulses were delayed with respect to each other by ~15 ns (~31 ns for samples with Atto647N as acceptor dye). Laser light is guided into the epi-illuminated confocal microscope (Olympus IX71, Hamburg, Germany) by a dichroic beamsplitter FF500/646-Di01 (Semrock, USA), and focused on the sample by a water immersion objective (UPlanSApo 60x/1.2 NA, Olympus Hamburg, Germany). The emitted fluorescence is collected through the objective and focused on a 100 µm pinhole. Using a polarizing beam splitter cube, emitted light is divided into its parallel and perpendicular components, followed by splitting into two spectral windows, "green" and "red", using long pass beamsplitter Q595 LPXR (AHF, Germany), and then again using non-polarizing 50/50 beam splitters resulting in a total of eight detection channels. Additionally, bandpass filters are placed in front of the detectors (FF01-530/43-25; AHF, Tübingen, Germany for donor molecules and HQ 720/150 nm; AHF, Tübingen, Germany for acceptor molecules). Detection is performed using eight avalanche photodiodes (4 green detectors (τ-SPAD, PicoQuant, Germany) and 4 red detectors (SPCM-AQR-14, Perkin Elmer). The detector outputs were recorded by a TCSPC module (HydraHarp 400, PicoQuant). Burst search was performed using home-written LabView-based software based on the inter-photon times. Before the burst selection, the trace of inter-photon times is smoothed using Lee-filter[2] with sliding window of the size 5 (m=2) and standard deviation of the background inter-photon times of σ_0=10 [3]. Bursts were selected using inter-photon time threshold of 100 µs. Minimum required number of photons in a burst is 60. Such settings correspond to approximately 10 kHz threshold in count-rate.

Supplemental Note 2: Benchmark samples

For the evaluation of correction procedures, we used a set of DNA rulers with varying base-pair separation between donor and acceptor dye, yielding a range of FRET efficiencies. The base-pair separation of the donor and acceptor labelling sites is as follows: 19 bp for low-FRET (LF), 14 bp for medium-FRET (MF) and 10 bp for high-FRET (HF) sample. These constructs cover distances of ~50-70 Å. The set of DNA rulers was measured with following three dye combinations, with labelling modification and linker length specified in bracket: AlexaFluor488 (TFP ester/C6 linker) – Atto647N (NHS ester/C2 linker); AlexaFluor488 (TFP ester/C6 linker) – Cy5 (NHS ester/C2 linker); AlexaFluor488 (TFP ester/C6 linker) – AlexaFluor647 (NHS ester/C2 linker). AlexaFluor488 and AlexaFluor647 are abbreviated as AF488 and AF647 throughout this work. Our choice of dyes was driven by many reports where these dyes were used and evaluated as suitable dye choice for burst-wise single-molecule experiments [4]. Measurements of DNA rulers were performed in buffer containing 5 mM $MgCl_2$, 150 mM NaCl, and 50 mM TRIS at pH 7.9. Labelling sites and DNA sequences are given in **Table S1.**

Supplemental Note 3: Regularized optimization of instrument spectral response.

Observed signal of j-th dye in detection window $\Lambda = \{G, R\}$ is defined as follows:

$$f_{\mathrm{j}}^{(\Lambda)} = f_{0,\mathrm{j}} \int_{\Lambda} g^{(\Lambda)}(\lambda) p_{\mathrm{j}}(\lambda) \mathrm{d}\lambda \qquad \int p_{\mathrm{j}}(\lambda)\, \mathrm{d}\lambda = 1 \tag{S1}$$

where $f_{0,j}$ is the total radiative intensity of the j-th dye, $p_j(\lambda)$ is the unit area normalized fluorescence emission spectrum of the j-th dye and $g^{(\Lambda)}(\lambda)$ is the instrument spectral response of a given spectral window $\Lambda = \{G, R\}$. This formulation can be discretized and given in a form of a matrix

$$f_j^{(\Lambda)} = F_{0,j} \sum_{i=1}^{n_\Lambda} g_i^{(\Lambda)}\, p_{ij} \tag{S2}$$

$$f^{(\Lambda)} = F_0\, P\, g^{(\Lambda)} \tag{S3}$$

where F_0 is a diagonal matrix with diagonal elements being radiative intensities $f_{0,j}$.
$g^{(\Lambda)}$ is a column vector, whose elements are integrals over the i^{th} λ – bin

$$g_i^{(\Lambda)} = \int_{\Delta\lambda_i} g^{(\Lambda)}(\lambda) d\lambda \tag{S4}$$

and P is a matrix with sum in rows being equal to one, and whose elements are analogously integrals over the i^{th} λ – bin

$$p_{ij} = \int_{\Delta\lambda_i} p_j(\lambda) A d\lambda \tag{S5}$$

For two detection windows, e.g. green (G) and red (R), we can write the system of matrix equations:

$$f^{(G)} = F_0\, P^{(G)}\, g^{(G)} \tag{S6}$$

$$f^{(R)} = F_0\, P^{(R)}\, g^{(R)} \tag{S7}$$

Quantities $f_{0,j}$ are unknown, but can be easily removed by dividing the observed signals $f^{(G)}$ and $f^{(R)}$. Such ratio of observed signal of j-th dye we denote as ξ_j:

$$\xi_j = \frac{f_j^{(G)}}{f_j^{(R)}} \tag{S8}$$

This can be rewritten as

$$f_j^{(G)} = \xi_j f_j^{(R)} \tag{S9}$$

$$f^{(G)} = \Xi\, f^{(R)} \tag{S10}$$

where Ξ is the diagonal matrix of the $f_j^{(G)}/f_j^{(R)}$ ratios. Using the definitions of $f^{(G)}$ and $f^{(R)}$ from above (see eq. (S6) and eq. (S7)), we can rewrite previous equation as follows:

$$F_0\, P^{(G)}\, g^{(G)} = \Xi\, F_0\, P^{(R)}\, g^{(R)} \tag{S11}$$

After multiplying each side of the equation with inverted matrix F_0^{-1} one obtains

$$P^{(G)}\, g^{(G)} = \Xi\ P^{(R)}\, g^{(R)} \tag{S12}$$

One can rewrite this equation as

$$P^{(G,R)} g^{(G,R)} = 0 \tag{S13}$$

$$P^{(G,R)} = \left(P^{(G)},\ 0^{(R)}\right) - \left(0^{(G)},\ \Xi\ P^{(R)}\right) = \left(P^{(G)}, -\Xi\ P^{(R)}\right) \tag{S14}$$

$$g^{(G,R)} = \begin{pmatrix} g^{(G)} \\ g^{(R)} \end{pmatrix} \tag{S15}$$

where $P^{(G,R)}$ is row block matrix with size $(n_G + n_R) \times j$ and $g^{(G,R)}$ is column block vector with size $n_G + n_R$.

As elements of matrix $P^{(G,R)}$ are measured with some accuracy, the equation $P^{(G,R)} g^{(G,R)} = 0$ is never fulfilled in practice. Therefore, instead of looking for exact solution, we are looking for such $g^{(G,R)}$ that minimizes the norm of the vector $P^{(G,R)} g^{(G,R)}$ (optimization problem). The arising ambiguity of the solution can be dealt with by requiring certain properties of the vector

$g^{(G,R)}$, such that only positive values are allowed. In that case the optimization problem can be defined as follows:

$$\begin{aligned} &min\left(\left\|P^{(G,R)}g^{(G,R)}\right\|^2\right) \\ &subject\ to\ \ Ig^{(G,R)} \geq 0 \end{aligned} \tag{S16}$$

Smoothness regularized optimization

One can impose further restrictions on the $g^{(G,R)}(\lambda)$ solution. For instance, we can require that the shape of $g^{(G,R)}(\lambda)$ does not have sharp changes, i.e. is smooth. This is achieved by searching the minimum norm of the derivative vector for each wavelength window $\Lambda = \{G, R\}$:

$$\min\left(\left\|\partial_\lambda g^{(\Lambda)}(\lambda)\right\|^2\right), \Lambda = \{G, R\} \tag{S17}$$

where the derivative $\partial_\lambda g^{(\Lambda)}(\lambda)$ is defined as

$$\partial_\lambda g^{(\Lambda)}(\lambda) = \frac{dg^{(\Lambda)}(\lambda)}{d\lambda} \tag{S18}$$

or in a case of finite-difference approximation

$$\partial g_i^{(\Lambda)} = \frac{\Delta g_i^{(\Lambda)}}{\Delta\lambda_i} \tag{S19}$$

where $\Delta g_i^{(\Lambda)}$ can be defined as forward difference

$$\Delta g_i^{(\Lambda)+} = g_{i+1}^{(\Lambda)} - g_i^{(\Lambda)} \tag{S20}$$

as a backward difference

$$\Delta g_i^{(\Lambda)-} = g_i^{(\Lambda)} - g_{i-1}^{(\Lambda)} \tag{S21}$$

or as a central difference

$$\Delta g_i^{(\Lambda)0} = \frac{1}{2}(g_{i+1}^{(\Lambda)} - g_{i-1}^{(\Lambda)}) \tag{S22}$$

Therefore, the vector of $g^{(\Lambda)}$differences can be expressed as

$$\Delta g^{(\Lambda)} = D\, g^{(\Lambda)} \tag{S23}$$

$$D = \frac{1}{2}\begin{bmatrix} 0 & 0 & 0 & & & \\ -1 & 0 & 1 & \cdots & 0 & \\ 0 & -1 & 0 & & & \\ & \vdots & & \ddots & \vdots & \\ & & & 0 & 1 & 0 \\ & 0 & & \cdots & -1 & 0 & 1 \\ & & & & 0 & 0 & 0 \end{bmatrix} \tag{S24}$$

where *D* is first-order, central-difference differentiation operator with 1's on upper diagonal, -1's on lower diagonal and 0's on main diagonal. First and last row elements are treated specially, i.e. set to zero. $Dg^{(\Lambda)}$ represents a measure of the variation or a measure of "smoothness" of $g^{(\Lambda)}$.

The squared norm of vector $\partial g^{(\Lambda)}$ is:

$$\left\|\partial g^{(\Lambda)}\right\|^2 = \partial g^{(\Lambda)\mathrm{T}} \partial g^{(\Lambda)} = g^{(\Lambda)\mathrm{T}}\, S\, g^{(\Lambda)} \tag{S25}$$

where *S* is a "smoothness" matrix/operator

$$S = \frac{1}{\Delta\lambda^2} D^{\mathrm{T}} D \tag{S26}$$

A sharp change of derivative on the border of two spectral windows, G and R, is allowed. Therefore, we build the final smoothing operator as block matrix:

$$S^{(G,R)} = \begin{pmatrix} \frac{1}{\Delta\lambda^{(G)2}} D^{(\mathrm{G})\mathrm{T}} D^{(G)} & 0 \\ 0 & \frac{1}{\Delta\lambda^{(R)2}} D^{(\mathrm{R})\mathrm{T}} D^{(R)} \end{pmatrix} \tag{S27}$$

Matrices $D^{(G)}$ and $D^{(R)}$ are of sizes $(n_G \times n_G)$ and $(n_R \times n_R)$ respectively. Finally, the smoothness regularized optimization problem can be formulated as:

$$min\left(g^{(G,R)\mathrm{T}}\left(\mu\, S^{(G,R)} + P^{(G,R)T}P^{(G,R)}\right)g^{(G,R)}\right) \quad \text{(S28)}$$
$$subject\ to\ \ Ig^{(G,R)} \geq 0$$

where $\boldsymbol{\mu}$ is regularization term, penalizing sharp changes of the $g^{(G,R)}$ solution.

In our workflow, we multiply the spectra of dyes, $P^{(G)}$and $P^{(R)}$, with the spectra of dichroic, filters and detector, and the optimized $g^{(\mathrm{G,R})}$ solution in that case is not the detection profile, but the correction to it. Smoothness regularized optimization of the $g^{(G,R)}$ solution was performed using in-house Python script, based on quadratic programming module of CVXOPT library.

Supplemental Note 4: Goodness-of-fit landscape for methods M-1 and M-3

Here we show γ and β goodness-of-fit landscape, where as a goodness of fit statistics we used sum of weighted squared residuals, here denoted as *sse ("sum of squared errors")*. Furthermore, we examined how is fit landscape affected by limited data availability, i.e. when only one sample is available for the calibration. We considered two scenarios: global calibration of three populations ("global γ"), and calibration of a single population ("local γ"). In both cases, we evaluate following two methods: *"Non-linear regression of burst-wise* $^{iii}S_{app}$ *versus* $^{iii}E_{app}$*"* (**M-1**), and *"Linear regression on photon counts"* (**M-3**). In the first step generated is a search space of all possible combinations of γ and β values for which we would like to evaluate goodness of a fit. In the next step, for each of the γ and β combinations, computed is a curve (M-1) or a plane (M-3), and estimated is the goodness of a given fit model, as a sum of a weighted squared residuals, *sse*. As weights, used are robust weights, calculated according to bisquare weight function.[5] Finally, we normalize *sse* with the minimum observed sse_{min}, to ensure equal scaling of the *z*-axis for different calibration methods. Contour lines are displayed in range from $1 \times sse/sse_{min}$ to $1.1 \times sse/sse_{min}$. It can be seen in **Figure SN1** that in both scenarios, quite broad range of γ and β values satisfy the data. Lastly, we demonstrate that when using a single population in a fit ambiguity of the optimal solution visibly increases, as well the discrepancy between methods M-1 and M-3.

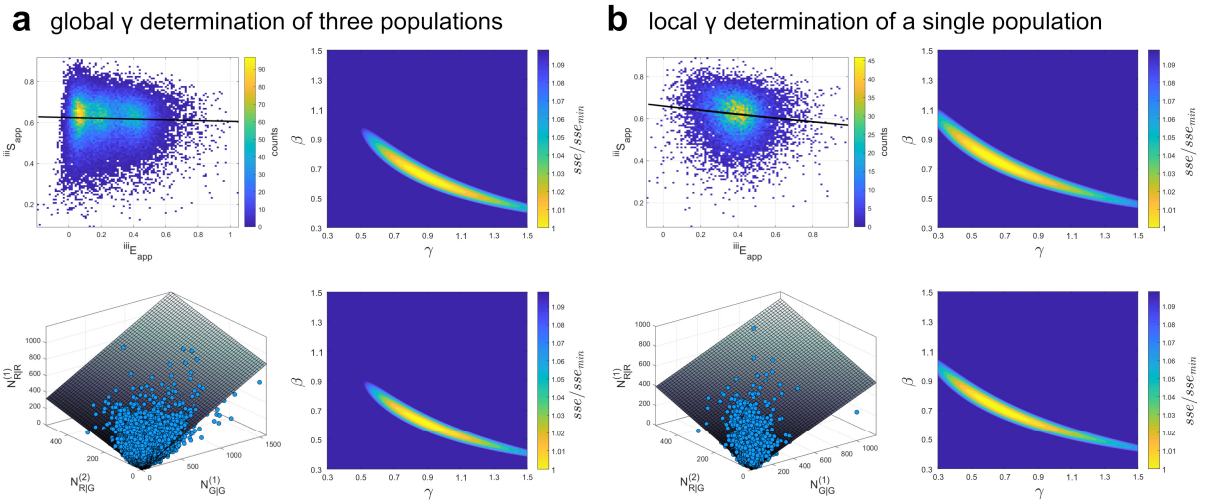


Figure SN4.1. γ and β goodness-of-fit landscape using methods M-1 and M-3. (a) global calibration of three samples (HF, MF and LF) and (b) local calibration of a single sample (HF) for a dye combination AF488-Atto647N.

Supplemental Note 5: Static FRET-line.

If one assumes a fixed distance between donor and acceptor, then only a single FRET rate is expected from single donor lifetime, $\tau_{\mathrm{D(A)}}$. With that being the case, static FRET-line can be described using two following simple equations:

$$E = 1 - \frac{\tau_{\mathrm{D(A)}}}{\tau_{\mathrm{D(0)}}} \tag{S29}$$

$$\frac{F_{\mathrm{G|G}}}{F_{\mathrm{R|G}}} = \frac{\phi_{\mathrm{F,D}}}{\phi_{\mathrm{F,A}}} \cdot \left(\frac{\tau_{\mathrm{D(0)}}}{\tau_{\mathrm{D(A)}}} - 1 \right)^{-1} \tag{S30}$$

In such approximation, to obtain static FRET-line, FRET rate, k_{FRET}, is varied, giving us a relation between all possible E or $\frac{F_{\mathrm{G|G}}}{F_{R|G}}$ values and all possible $\tau_{\mathrm{D(A)}}$ values. Unfortunately, this simple scenario does not explain experimental data. However, even for the structurally static systems, DA sub-population is not lying on a static FRET-line defined as in (S29) and (S30). This displacement largely is caused by the approximation of dyes being fixed in space, while in experiment dyes are coupled to the biomolecules via ~20 Å long chemical linkers. This means that there is a spatial distribution of donor and acceptor positions, leading to the distribution of lifetimes, $p(k_{\mathrm{FRET}})$, which then translates into distribution of interdye distances, $p(R_{\mathrm{DA}})$ [6]. For dye positions being normally distributed around the mean dye position, $p(R_{\mathrm{DA}})$ is given as non-central χ-distribution, with non-centrality parameter being distance between mean dye positions, R_{mp}, and width parameter σ_{DA}. However, for long interdye distances and small variance, χ-distribution converges to the normal distribution. As in experiments only the distances > 35Å are measured, and the linker mobility induced distribution broadening is in the range of 5-10 Å, we can then express $p(R_{\mathrm{DA}})$ as follows:

$$p(R_{\mathrm{DA}}) = \frac{1}{\sigma_{\mathrm{DA}}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{R_{\mathrm{DA}} - \langle R_{\mathrm{DA}} \rangle}{\sigma_{\mathrm{DA}}}\right)^2} \tag{S31}$$

To correct static FRET-line for dye linker mobility, we vary $\langle R_{\mathrm{DA}} \rangle$, and for each of them $p(R_{\mathrm{DA}})$ is calculated, using a fixed value of 6Å for standard deviation, as this is found to be in agreement with series of benchmark experiments.[7] Using computed probabilities, $p(R_{\mathrm{DA}})$, we can further calculate species-weighted average fluorescence lifetime, $\langle \tau_{\mathrm{D(A)}} \rangle_x = \sum p(R_{\mathrm{DA}}) \tau_{\mathrm{D(A)}}$. Static FRET-line corrected for linker mobility, can be then expressed as follows:

$$E = 1 - \frac{\sum p(R_{\mathrm{DA}})\tau_{\mathrm{D(A)}}}{\tau_{\mathrm{D(0)}}} \tag{S32}$$

$$\frac{F_{\mathrm{G|G}}}{F_{\mathrm{R|G}}} = \frac{\phi_{\mathrm{F,D}}}{\phi_{\mathrm{F,A}}} \cdot \left(\frac{\tau_{\mathrm{D(0)}}}{\sum p(R_{\mathrm{DA}})\tau_{\mathrm{D(A)}}} - 1 \right)^{-1} \tag{S33}$$

For different values of σ_{DA}, shape of static FRET-line changes for the short lifetimes, i.e. high FRET efficiencies. Finally, it is important to note that all the static FRET-line derivations above assumed single donor-only lifetime, $\tau_{\mathrm{D(0)}}$. This condition is often not met, and if donor possesses multiple lifetimes due to, for instance, multiple isomeric states, then such approximated static FRET-line will not properly describe the system.

Parameters used to generate static FRET-lines for different dye combinations are given in **Table S5.**

Supplemental Note 6: Apparent dynamic shift determination

To assess the accuracy of procedures for determination of correction parameters, determined is displacement of DA sub-populations from a static FRET-line. We refer to this displacement to as apparent dynamic shift, ds_{app}. In the first step, ensured is equal scaling in the range 0-1 for both axes in the $E - \langle \tau_{\mathrm{D(A)}} \rangle_{\mathrm{F}}$ plot, by normalizing *x*-axis with $\tau_{\mathrm{D(0)}}$. In the next step, performed is a fit of the DA population using a superposition of *N* bivariate normal distributions. In a case that a mixture of Gaussians is needed to describe the population, mean of the Gaussian with maximum amplitude is taken. Next, computed are distances of the mean point to all points in the static FRET-line. Finally, the shortest distance is taken as a value for ds_{app}. Due to the curvature of the static FRET-line that accounts for linker broadening, the shortest distance is not necessarily the orthogonal one. Lastly, dynamic shift is assigned a positive or negative sign if the population is above or below the static FRET-line.

Supplemental Note 7: AV simulations

AV represents sterically allowed space of the dye coupled to biomolecule. In the AV method, fluorophore moiety is modelled as a sphere with radius R_{dye} or as an ellipsoid with radii $R_{dye\ (1)}$, $R_{dye\ (2)}$, $R_{dye\ (3)}$. Model with three radii is generally recommended for sterically complex dye environments. Chemical linker that couples dye to the biomolecule is modelled as a cylinder with length L_{length} and width L_{width}. The AV algorithm searches for the dye positions from attachment point and within the linker length, and this way a 3D grid is built around the attachment point, with the resolution of 0.9 Å. However, only those positions that do not cause steric clashes with the biomolecule are taken. To allow that linker actually can move around its attachment point, there is certain tolerance for linker clashes, such that all positions within $0.5 \times L_{width}$ are allowed.

All sterically allowed points in the AV cloud are equally weighted when calculating FRET observables. Such approach provided accurate inter-dye distances when host molecule for dyes are nucleic acids.[7, 8]

Here we used R_{dye} = 3.5 Å, L_{width} = 4.5 Å, L_{length} = 21 Å for AF488 and L_{length} = 22.5 Å for acceptor dyes.

AV simulations were performed using open-source software Olga[9] available at github.com/Fluorescence-Tools/Olga.

Supplemental Tables

Table S1. Sequence of dsDNA molecules and labelling sites. Labelling sites for the donor and acceptor are denoted in green and red. In the sample name, abbreviations LF, MF and HF indicate low-, medium- and high-FRET efficiency.

Name	Labelling position	Sequence and labelling sites
LF	T 28	5'- GAG CTG AAA GTG TCG AGT TTG TTT GAG TGT TTG TCT GG-3'
	T 31	3'- CTC GAC TTT CAC AGC TCA AAC AAA CTC ACA AAC AGA CC-5'
MF	T 23	5'- GAG CTG AAA GTG TCG AGT TTG TTT GAG TGT TTG TCT GG-3'
	T 31	3'- CTC GAC TTT CAC AGC TCA AAC AAA CTC ACA AAC AGA CC-5'
HF	T 19	5'- GAG CTG AAA GTG TCG AGT TTG TTT GAG TGT TTG TCT GG-3'
	T 31	3'- CTC GAC TTT CAC AGC TCA AAC AAA CTC ACA AAC AGA CC-5'

Table S2. Documentation of sample properties, experimental design and parameters used in data processing and analysis

A. Sample specification	
Samples	Double-stranded DNA rulers: • LF: sample with low-FRET efficiency; 19 bp separation • MF: sample with medium-FRET efficiency; 14 bp separation • HF: sample with high-FRET efficiency; 10 bp separation For sequence information see main text **Table S1.**
Dyes and labelling modification/linkage length	Donor dye: • Alexa Fluor 488 (TFP ester/C6 linker) Acceptor dye: • Atto647N (NHS ester/C2 linker) • Alexa Fluor 647 (NHS ester/C2 linker) • Cy5 (NHS ester/C2 linker)
Labelling positions	LF sample: T28 (D-strand); T31 (A-strand) MF sample: T23 (D-strand); T31 (A-strand) HF sample: T19 (D-strand); T31 (A-strand)
Buffer	5 mM $MgCl_2$, 150 mM NaCl, and 50 mM TRIS at pH 7.9
Additives	n.a.
B. Measurement specification	
Measurement modality	confocal smFRET measurements of freely diffusing molecules using Pulsed Interleaved excitation (PIE)
Laser	Donor excitation laser: LDH-D-C 485 (PicoQuant, Germany) Acceptor excitation laser: LDH-D-C 640 (PicoQuant, Germany)
Laser power at the sample (measured on a dry objective)	Donor excitation laser: 60 µW Acceptor excitation laser: 10µW (20 µW for samples with Atto647N)
Laser repetition rate	32 MHz (16 MHz for samples with Atto647N)
Laser repetition time	31.25 ns (62.5 ns for samples with Atto647N)
Delay between the lasers	~ 15 ns (~ 31 ns for samples with Atto647N)
Time resolution	1 ps (2 ps for samples with Atto647N)
Confocal microscope	Olympus IX71 (Hamburg, Germany)
Dichroic beamsplitter	FF500/646-Di01 (Semrock, USA)
Objective	water immersion objective UPlanSApo 60x/1.2 NA (Olympus Hamburg, Germany)
Observation volume size	~6 fL
Pinhole size	100 µm
Fluorescence beamsplitter	longpass beamsplitter Q595 LP DCXR (AHF, Germany),
Bandpass filters	donor detection channels: FF01-530/43-25 (AHF, Tübingen, Germany) acceptor detection channels: and HQ 720/150 nm (AHF, Tübingen, Germany)

Detectors	4 donor detection channels: τ-SPAD (PicoQuant, Germany): • #0, #4 – perpendicular signal detectors • #1, #5 – parallel signal detectors 4 acceptor detection channels: SPCM-AQR-14 (Perkin Elmer): • #2, #6 – perpendicular signal detectors • #3, #7 – parallel signal detectors
Data acquisition electronics	TCSPC and Time Tagging Electronics HydraHarp 400 (PicoQuant, Germany)
Raw data format	HydraHarp 400 time-tagged time-resolved (TTTR) file (*.ht3)
Sample chamber	8 well NUNC chambers with #1.5 borosilicate glass bottom (Lab-Tek, Thermo Scientific); sample volume: 0.5 mL
Temperature	20°C
C. Data processing/ parameters and software	
Software	MFD Spectroscopy and Imaging Software package for confocal fluorescence spectroscopy and imaging experiments using Multiparameter Fluorescence Detection (MFD) data includs various analysis tools (FCS, fFCS, PDA, seTCSPC, trace analysis, 2D simulation of MFD diagrams and Multiparameter Fluorescence Image Spectroscopy (MFIS)).
Burst search method	Inter-photon time- based burst selection with Lee filter smoothing of inter-photon time trace
Selection method	All photons
Lee filter parameters	m=2; σ_0=10
Inter-photon time threshold	100 μs
Minimum number of photons in a burst	60
Japanese correction factors for polarization mixing	l_1 = 0.0175, l_2= 0.0526
Time window for computation of ALEX-2CDE filter	100 μs
dsDNA-Alexa Fluor 488- Atto647N	
$G_{\mathrm{VH/VV}}$	donor detection channels: 1.12 acceptor detection channels: 1.06
Background count-rate	$S_{\mathrm{G\|G}}^{(\mathrm{BG})}$ = 0.7617 kcps $S_{\mathrm{R\|G}}^{(\mathrm{BG})}$ = 0.4131 kcps $S_{\mathrm{R\|R}}^{(\mathrm{BG})}$ = 0.4192 kcps
${}^{i}S_{\mathrm{app}} - {}^{i}E_{\mathrm{app}}$ selection of D0 population	-0.017 < ${}^{i}E_{\mathrm{app}}$ < 0.074 0.955 < ${}^{i}S_{\mathrm{app}}$< 1.002
Donor crosstalk, α	0.0145
${}^{i}S_{\mathrm{app}} - {}^{i}E_{\mathrm{app}}$ selection of A0 population	${}^{i}E_{\mathrm{app}}$ > 0.478 0.017 < ${}^{i}S_{\mathrm{app}}$< 0.117

Direct excitation of acceptor, δ	0.0243
ALEX-2CDE threshold for selection of DA population	15.5
β and γ parameters	• Non-linear regression of burst-wise ${}^{iii}S_{\mathrm{app}}$ versus ${}^{iii}E_{\mathrm{app}}$ values (M-1) γ = 0.9046; β =0.6617 • Linear regression between $1/\langle {}^{iii}S_{\mathrm{app}}\rangle$ and $\langle {}^{iii}E_{\mathrm{app}}\rangle$ of the population means (M-2) γ = 0.8967; β = 0.6075 • Linear regression on photon counts (M-3) γ = 0.8887; β = 0.6215
dsDNA-Alexa Fluor 488- Alexa Fluor 647	
$G_{\mathrm{VH/VV}}$	donor detection channels: 1.25 acceptor detection channels: 1.11
Background countrate	$S_{\mathrm{G\|G}}^{(\mathrm{BG})}$ = 0.7901 kcps $S_{\mathrm{R\|G}}^{(\mathrm{BG})}$ = 0.4866 kcps $S_{\mathrm{R\|R}}^{(\mathrm{BG})}$ = 0.3519 kcps
${}^{i}S_{\mathrm{app}} - {}^{i}E_{\mathrm{app}}$ selection of D0 population	-0.002 < ${}^{i}E_{\mathrm{app}}$< 0.058 0.968 < ${}^{i}S_{\mathrm{app}}$ < 1.012
Donor crosstalk, α	0.0118
${}^{i}S_{\mathrm{app}} - {}^{i}E_{\mathrm{app}}$ selection of A0 population	${}^{i}E_{\mathrm{app}}$ > 0.478 0.028 < ${}^{i}S_{\mathrm{app}}$ < 0.139
Direct excitation of acceptor, δ	0.0321
ALEX-2CDE threshold for selection of DA population	20.5
β and γ parameters	• Non-linear regression of burst-wise ${}^{iii}S_{\mathrm{app}}$ versus ${}^{iii}E_{\mathrm{app}}$ values (M-1) γ = 0.5449; β = 0.5849 • Linear regression between $1/\langle {}^{iii}S_{\mathrm{app}}\rangle$ and $\langle {}^{iii}E_{\mathrm{app}}\rangle$ of the population means (M-2) γ = 0.5454; β = 0.5476 • Linear regression on photon counts (M-3) γ = 0.5066; β = 0.5631
dsDNA-Alexa Fluor 488- Cy5	
$G_{\mathrm{VH/VV}}$	donor detection channels: 1.14 acceptor detection channels: 1.14
Background count-rate	$S_{\mathrm{G\|G}}^{(\mathrm{BG})}$ = 0.7095 kcps $S_{\mathrm{R\|G}}^{(\mathrm{BG})}$ = 0.4274 kcps $S_{\mathrm{R\|R}}^{(\mathrm{BG})}$ = 0.3658 kcps
${}^{i}S_{\mathrm{app}} - {}^{i}E_{\mathrm{app}}$ selection of D0 population	-0.001 < ${}^{i}E_{\mathrm{app}}$ < 0.058 0.966 < ${}^{i}S_{\mathrm{app}}$ < 1
Donor crosstalk, α	0.0134

${}^{i}S_{\mathrm{app}} - {}^{i}E_{\mathrm{app}}$ selection of A0 population	${}^{i}E_{\mathrm{app}} > 0.478$ $0.028 < {}^{i}S_{\mathrm{app}} < 0.139$
Direct excitation of acceptor, δ	0.0384
ALEX-2CDE threshold for selection of DA population	18.5
β and γ parameters	• Non-linear regression of burst-wise ${}^{iii}S_{\mathrm{app}}$ versus ${}^{iii}E_{\mathrm{app}}$ values (M-1) γ = 0.5045; β = 0.5107 • Linear regression between $1/\langle {}^{iii}S_{\mathrm{app}} \rangle$ and $\langle {}^{iii}E_{\mathrm{app}} \rangle$ of the population means (M-2) γ = 0.5034; β = 0.4757 • Linear regression on photon counts (M-3) γ = 0.4644; β = 0.4814

Table S3. List of dyes used for regularized optimization of instrument detection profile. Dyes are selected such that they emit in both detection channels. All measurements are performed in Milli-Q water (18 MΩ·cm), with exception of Rhodamine B, where triethylamine (Fluka) was added in traces to the measurement solution to ensure formation of Rhodamine B- zwitterion. Background signal was recorded for each sample separately. This is especially crucial for reproducibility of smaller intensity ratios.

Dye	Producer	Stock solution solvent
Alexa Fluor 488 free acid, 5 – isomer/ experimental sample	Molecular Probes	water
Atto 488 – COOH	ATTO-TEC	water
Rhodamine 110 – chloride	Sigma Aldrich	ethanol (Merck Supelco)
Oregon Green 514 carboxylic acid, mixed isomers	Molecular Probes	N, N – DMF (Merck Supelco)
Rhodamine 123	Sigma Aldrich	ethanol (Merck Supelco)
Rhodamine 19 – perchlorate	Fluka	ethanol (Merck Supelco)
Atto 532 – COOH	ATTO-TEC	water
Rhodamine 6G	Lambda Physik	water
Atto 542 – COOH	ATTO-TEC	water
Rhodamine B	Fluka	water
Alexa Fluor 546 free acid (isomer mix?/experimental sample)	Molecular Probes	water
5 – carboxy tetramethylrhodamine (5-TAMRA)	Biosearch	methanol (Acros Organics)
Atto 550 – COOH	ATTO-TEC	DMSO (Fisher Scientific)

Table S4. Measured, predicted and optimized ratios of green to red fluorescence intensity.

Dye	$\xi_{j,\text{measured}}$	$\xi_{j,\text{predicted}}$	$\xi_{j,\text{optimized}}$	$\frac{\xi_{j,\text{predicted}}}{\xi_{j,\text{measured}}}$	$\frac{\xi_{j,\text{optimized}}}{\xi_{j,\text{measured}}}$
Alexa Fluor 488 free acid, 5 – isomer/ experimental sample	93.2809	93.1178	91.6162	0.9983	0.9822
Atto 488 – COOH	69.7432	72.2228	73.3419	1.0356	1.0516
Rhodamine 110 – chloride	60.5077	61.1656	61.7754	1.0109	1.0210
Oregon Green 514 carboxylic acid, mixed isomers	48.8029	45.7916	45.2581	0.9383	0.9274
Rhodamine 123	40.8533	43.4413	44.7400	1.0633	1.0951
Rhodamine 19 – perchlorate	16.2032	15.4369	16.8769	0.9527	1.0416
Atto 532 – COOH	9.4590	8.7309	9.5257	0.9230	1.0070
Rhodamine 6G	8.5998	7.7937	8.5223	0.9063	0.9910
Atto 542 – COOH	2.8493	2.5236	2.7775	0.8857	0.9748
Rhodamine B	0.5932	0.5742	0.5929	0.9680	0.9996
Alexa Fluor 546 free acid (isomer mix?/experimental sample)	0.4879	0.4268	0.4667	0.8747	0.9565
5-carboxy tetramethylrhodamine (5-TAMRA)	0.3532	0.3015	0.3283	0.8537	0.9297
Atto 550 – COOH	0.2755	0.2560	0.2806	0.9291	1.0182

Table S5. Parameters of static FRET-line for different dye combinations used in this study.

Dye pair	$\tau_{D(0)}$ [ns]	σ_{DA} [Å]	R_0 [Å]
AlexaFluor488-Atto647N	3.94	6	49
AlexaFluor488-AlexaFluor647	3.94	6	52
AlexaFluor488-Cy5	3.94	6	52

Supplemental Figures

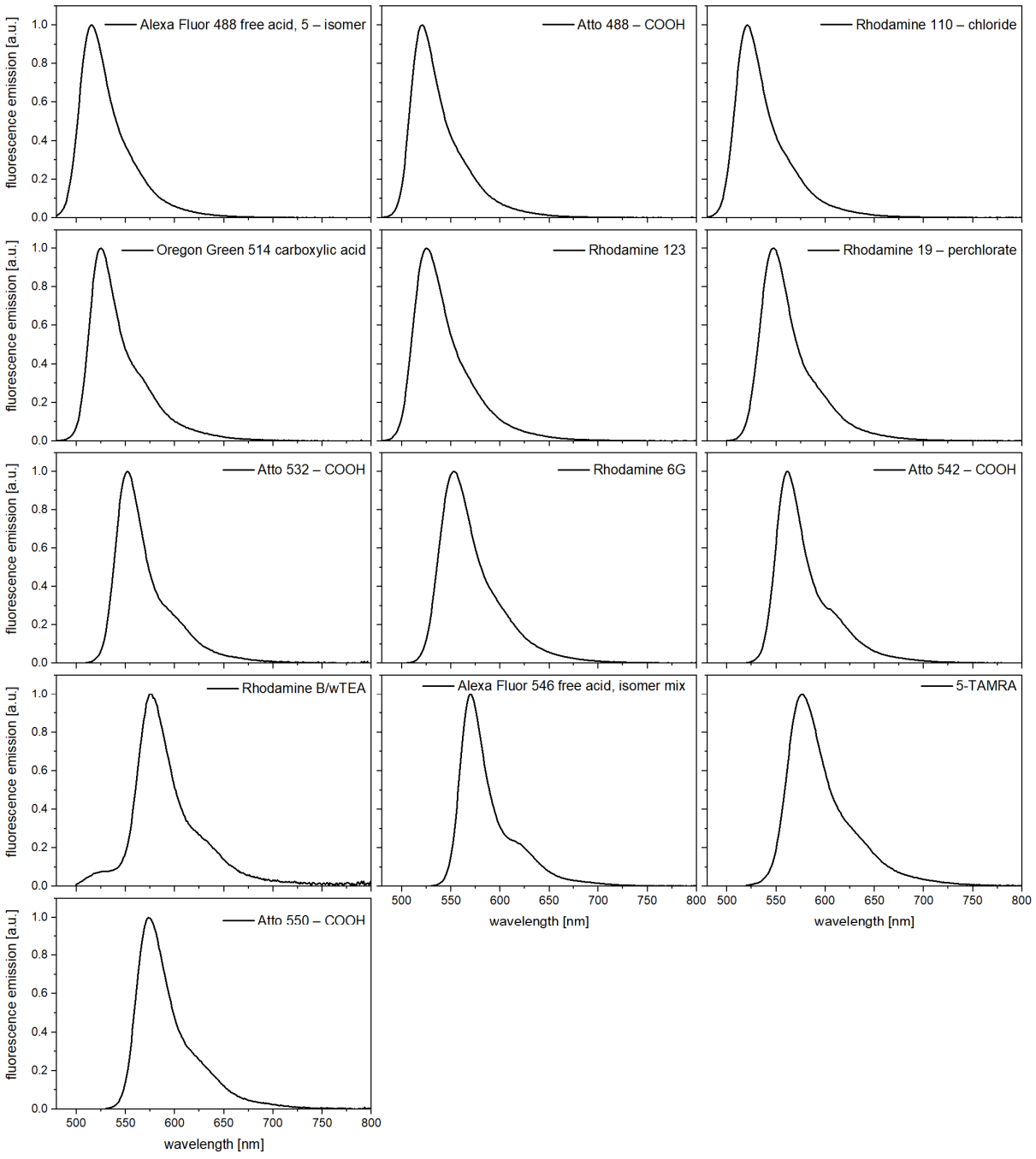


Figure S1. Spectra of free dyes used for regularized optimization of instrument spectral response.

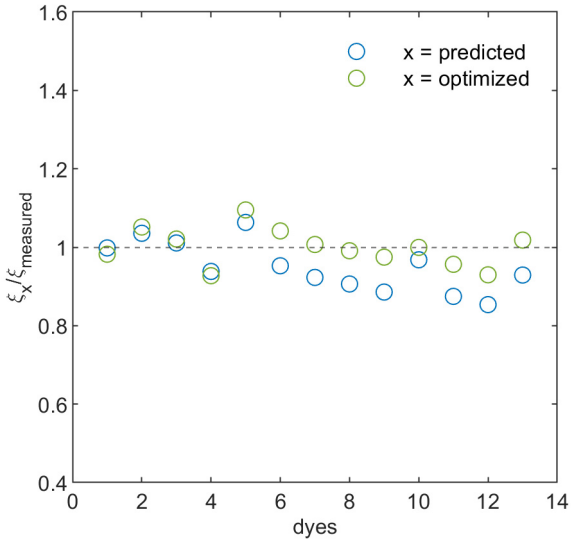


Figure S2. Discrepancy between measured and predicted/optimized ratios of green to red fluorescence intensity.

Supplemental References

1. Sisamakis, E.; Valeri, A.; Kalinin, S.; Rothwell, P. J.; Seidel, C. A. M., Accurate Single-Molecule Fret Studies Using Multiparameter Fluorescence Detection. *Method Enzymol* **2010,** *475*, 455-514.

2. Lee, J. S., Digital image enhancement and noise filtering by use of local statistics. *IEEE Trans Pattern Anal Mach Intell* **1980,** *2* (2), 165-8.

3. Enderlein, J.; Robbins, D. L.; Ambrose, W. P.; Goodwin, P. M.; Keller, R. A., Statistics of single-molecule detection. *J Phys Chem B* **1997,** *101* (18), 3626-3632.

4. Vandenberk, N.; Barth, A.; Borrenberghs, D.; Hofkens, J.; Hendrix, J., Evaluation of Blue and Far-Red Dye Pairs in Single-Molecule Forster Resonance Energy Transfer Experiments. *J Phys Chem B* **2018,** *122* (15), 4249-4266.

5. DuMouchel, W. H., and F. L. O'Brien In *Integrating a Robust Option into a Multiple Regression Computing Environment*, Computer Science and Statistics: Proceedings of the 21st Symposium on the Interface, Alexandria, VA, 1989; American Statistical Association: Alexandria, VA, 1989.

6. Haas, E.; Wilchek, M.; Katchalskikatzir, E.; Steinberg, I. Z., Distribution of End-to-End Distances of Oligopeptides in Solution as Estimated by Energy-Transfer. *P Natl Acad Sci USA* **1975,** *72* (5), 1807-1811.

7. Sindbert, S.; Kalinin, S.; Hien, N.; Kienzler, A.; Clima, L.; Bannwarth, W.; Appel, B.; Muller, S.; Seidel, C. A. M., Accurate Distance Determination of Nucleic Acids via Forster Resonance Energy Transfer: Implications of Dye Linker Length and Rigidity. *J Am Chem Soc* **2011,** *133* (8), 2463-2480.

8. Hellenkamp, B.; Schmid, S.; Doroshenko, O.; Opanasyuk, O.; Kuhnemuth, R.; Adariani, S. R.; Ambrose, B.; Aznauryan, M.; Barth, A.; Birkedal, V.; Bowen, M. E.; Chen, H. T.; Cordes, T.; Eilert, T.; Fijen, C.; Gebhardt, C.; Gotz, M.; Gouridis, G.; Gratton, E.; Ha, T.; Hao, P. Y.; Hanke, C. A.; Hartmann, A.; Hendrix, J.; Hildebrandt, L. L.; Hirschfeld, V.; Hohlbein, J.; Hua, B. Y.; Hubner, C. G.; Kallis, E.; Kapanidis, A. N.; Kim, J. Y.; Krainer, G.; Lamb, D. C.; Lee, N. K.; Lemke, E. A.; Levesque, B.; Levitus, M.; McCann, J. J.; Naredi-Rainer, N.; Nettels, D.; Ngo, T.; Qiu, R. Y.; Robb, N. C.; Rocker, C.; Sanabria, H.; Schlierf, M.; Schroder, T.; Schuler, B.; Seidel, H.; Streit, L.; Thurn, J.; Tinnefeld, P.; Tyagi, S.; Vandenberk, N.; Vera, A. M.; Weninger, K. R.; Wunsch, B.; Yanez-Orozco, I. S.; Michaelis, J.; Seidel, C. A. M.; Craggs, T. D.; Hugel, T., Precision and accuracy of single-molecule FRET measurements-a multi-laboratory benchmark study (vol 15, pg 984, 2018). *Nat Methods* **2018,** *15* (11), 984-984.

9. Dimura, M.; Peulen, T. O.; Sanabria, H.; Rodnin, D.; Hemmen, K.; Hanke, C. A.; Seidel, C. A. M.; Gohlke, H., Automated and optimally FRET-assisted structural modeling. *Nat Commun* **2020,** *11* (1).