Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems

Mykola Dimura[1,2,3], Thomas O Peulen[1,3], Christian A Hanke[1], Aiswaria Prakash[1], Holger Gohlke[2] and Claus AM Seidel[1]

Förster Resonance Energy Transfer (FRET) combined with single-molecule spectroscopy probes macromolecular structure and dynamics and identifies coexisting conformational states. We review recent methodological developments in integrative structural modeling by satisfying spatial restraints on networks of FRET pairs (hybrid-FRET). We discuss procedures to incorporate prior structural knowledge and to obtain optimal distance networks. Finally, a workflow for hybrid-FRET is presented that automates integrative structural modeling and experiment planning to put hybrid-FRET on rails. To test this workflow, we simulate realistic single-molecule experiments and resolve three protein conformers, exchanging at 30 μs and 10 ms, with accuracies of 1–3 Å RMSD versus the target structure. Incorporation of data from other spectroscopies and imaging is also discussed.

Addresses
[1] Chair for Molecular Physical Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany
[2] Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, 40225 Düsseldorf, Germany

Corresponding author: Seidel, Claus AM (cseidel@hhu.de)
[3] Contributed equally.

Current Opinion in Structural Biology 2016, **40**:163–185

This review comes from a themed issue on **Biophysical and molecular biological methods**

Edited by **Petra Fromme** and **Andrej Sali**

For a complete overview see the Issue and the Editorial

http://dx.doi.org/10.1016/j.sbi.2016.11.012

Introduction

The measurement of Förster resonance energy transfer (FRET) [1••,2••,3] from a donor (D) to an acceptor (A) fluorophore has become a popular biophysical method that can yield unique insights into the structure and the structural exchange dynamics of labeled biomolecular systems. FRET has applications in two major research areas. The first application makes use of the high time resolution and single-molecule capabilities to study the kinetics and to detect intermediates of exchanging systems with a limited number of DA pairs [4]. The second application utilizes FRET to study structures of biomolecules in solution using a larger number of DA pairs, to obtain detailed structural insights on biomolecules in solution [5]. The first approach is successful in correlating structural dynamics and function by directly mapping the timescales of exchange and the pathways between conformational states with biomolecules on the single-molecule (sm) level [4,6]. Various equilibrium processes were studied such as folding of proteins [7,8] or nucleic acids [9–11], assembly and disassembly of complexes [12–14], and enzymatic cycles [15,16]. A versatile set of single-molecule based measurements and analysis techniques allows one to cover a wide range of time scales ranging from picoseconds [17], tens of nanoseconds [18,19], microseconds [13,20•,21•,22,23•,24•] to milliseconds and seconds [25•,26,27••]. Moreover, perturbation techniques such as temperature modulation [28] and microfluidic mixers for non-equilibrium experiments [29•] widen the use of smFRET experiments even further.

In the second approach, FRET is applied as a spectroscopic ruler for quantitative structural studies by measuring a larger set of single-pair DA pairs, which is necessary for interrogating the most important degrees of freedom of the system (i.e., the data sparseness must be sufficiently low [30•]). In the absence of prior structural knowledge, such a set should ideally be a connected network. Pioneering work on the ensemble and single-molecule level determined quantitative FRET-based structural models of essentially static systems such as polyprolines [5,31], various nucleic acids [32–34,35•], and large biomolecular complexes [36,37]. This approach benefits from single-molecule methods, which increase the effective resolution, to discriminate among coexisting states in solution. Combining methodologies of both major application areas harbors the potential to resolve structural models of dynamically exchanging coexisting states.

FRET measurements can only report inter-dye distances, and the number of FRET pairs is limited. Integrated methods can leverage experimental data that by itself would be insufficient to determine structures with satisfactory accuracy [38,39]. Therefore, integrative structural modeling relying on molecular simulations as well as FRET data is essential for achieving FRET-based

atomistic structural models. In the following article, we refer to this hybrid approach as hybrid-FRET modeling. Due to the long-range FRET information (typically 20–100 Å), the hybrid-FRET modeling approach is especially suitable to characterize the tertiary, super-tertiary [40], and quaternary structures of biomolecular systems. Benchmark studies for FRET-restrained structural hybrid modeling of biomolecules and their complexes versus crystal structures have demonstrated the accuracy of the integrative approach [41•,42••]. A dramatic improvement in the precision and accuracy (root mean square deviation (RMSD) of 0.5 Å) is achieved by explicitly considering spatial distributions of dye positions, which greatly reduces uncertainties due to flexible dye linkers [42••].

There are many reasons for further improving FRET-based structural modeling. In view of the huge dynamic time range from picoseconds to hours that can be covered by fluorescence, smFRET experiments and hybrid-FRET modeling are an especially promising combination to determine structure and dynamics simultaneously. Many other biophysical techniques have shown that a dynamic view on protein structures can give many important insights because the dynamic properties control their functions, as shown for various research areas such as signaling [43], enzymatic catalysis [44], (mis)folding [45,46], and modulation of allostery [47]. Several structured [48–50] and partially unstructured [51,52,53•] proteins were characterized by the FRET approach. Even dynamic and large biomolecular systems were studied such as ligand-induced conformational changes of the membrane-bound SNARE complex [54] and the conformational dynamics of single HIV-1 envelope trimers on the surface of native virions [55].

The dynamic smFRET experiments combined with integrative modeling should be especially useful to detect and directly characterize the conformational heterogeneity of biomolecules in the presence of thermally excited conformational states with a lifetime in the microsecond time range, which complements recently improved NMR spectroscopy methods [56]. Moreover, there are many reasons (e.g., limited solubility, unsuitable molecular size, limited stability, inadequate purity, or large conformational heterogeneity) why the structures of biological macromolecules often prove intractable to mainstream structural biology methods (i.e., crystallography, NMR spectroscopy, and cryo-electron microscopy) [38]. Exploiting the time resolution, sensitivity, and selectivity of fluorescence, integrative structural modeling, combining FRET experiments with computer simulations and comparative protein structure modeling [57] as prior information is developing into a promising complementary method for structural biology of dynamic biomolecules.

This perspective article focuses on advances in FRET-based structural modeling and the application to dynamic biomolecular systems. To this end, we first review FRET techniques that can yield information for structural modeling of biomolecules. We then summarize the current state of the art techniques for modeling the FRET dye label behavior, which is important for FRET-based structural modeling. Subsequently, we introduce a typical workflow for hybrid-FRET modeling, which is then applied in a test case. We demonstrate an application of fluorescence methods to a three-state model system to show the potential of FRET-based structural modeling and to assess the currently achievable precision and accuracy. For this test case, we introduce an automated procedure for determining the most informative single-pair FRET labeling sites and simulate experimental data under realistic conditions, which are then used for structural modeling. In the outlook, we discuss the potential directions for future improvements with respect to experimental techniques and computational modeling to improve the precision and accuracy of hybrid models. Finally, we discuss applications of hybrid-FRET methods in imaging to realize an integrated molecular fluorescence microscope combining optical and computational microscopy (i.e., coarse-grained and all-atom molecular (dynamics) simulations) [58] to describe suitably labeled biomolecular systems by atomistic structural models.

FRET-based structural modeling

Experimental techniques

Considering the conformational species (i), FRET between the tethered fluorophores D and A with the DA-distance $R^{(i)}_{DA}$ is characterized by the yield of energy transfer from D to A, usually called transfer efficiency $E^{(i)}$ or FRET efficiency (Eq. (1a)). This yield is the ratio of the rate constant $k^{(i)}_{RET}$ relative to all other processes depopulating the excited donor with the total rate constant $k_{D,0}$. Eq. (1b), derived by Theodor Förster [1••,2••], allows one to compute $R^{(i)}_{DA}$ in units of the coupling constant R_0 (Förster radius). R_0 is specific for the used dye pair. It depends on the refractive index of the medium n, the spectral overlap integral of the D fluorescence and the A absorption spectrum J, the factor κ^2 for the relative orientation of the D and A dipole, and the D fluorescence quantum yield in absence of an acceptor, $\Phi_{F,D(0)}$ [3].

$$E^{(i)} = \underbrace{\frac{k^{(i)}_{RET}}{k_{D,0}+k^{(i)}_{RET}}}_{(a)} = \underbrace{\frac{1}{1+\left(R^{(i)}_{DA}/R_0\right)^6}}_{(b)} = \underbrace{\frac{F^{(i)}_A}{\gamma F^{(i)}_D + F^{(i)}_A}}_{(c)} = \underbrace{1-\frac{\tau^{(i)}_{D(A)}}{\tau_{D(0)}}}_{(d)} \quad (1)$$

The FRET-efficiency E can be determined by fluorescence intensities of the sensitized acceptor emission F_A and the quenched donor fluorescence F_D (Eq. (1c)). The

accuracy of E relies critically on several factors: (1) the purity of the samples, that is, the Donor-only reference for a correct calibration and the degree of labeling of the FRET sample; and (2) the properties of the labels and the setup. The setup and the samples should be carefully characterized for computing the experimental calibration factor $\gamma = (x_{b,A}\, g_A \Phi_{F,A})/(g_D \Phi_{F,D(0)})$ given by the spectrum-dependent detection efficiencies of the D and A fluorescence (g_D and g_A), and the fluorescence quantum yields $\Phi_{F,D(0)}$ and $\Phi_{F,A}$ of D and A respectively. Often $\Phi_{F,D(0)}$ and $\Phi_{F,A}$ are assumed to be identical for all conformational species (for details and a more complex analysis, see [59••,60••,61•]). In single-molecule experiments, the excitation irradiances are usually high. Thus, saturation effects of the fluorophores due to formation of long lived dark-states must be considered in the fluorescence quantum yield. For example, the cyanine-based acceptor dyes such as Cy5 and Alexa647, undergo photo-induced cis-trans isomerization [62], such that the fraction of bright trans species $x_{b,A}$ drops significantly below one. $x_{b,A}$ can be obtained by analyzing the acceptor autocorrelation function of the FRET experiment [63,64•].

Time-resolved measurements of donor fluorescence decays $f(t)$ by, for example, time-correlated single photon counting (TCSPC) are useful for precise FRET measurements, because they offer three main advantages. (1) The FRET efficiency can be determined without instrumental calibrations via the excited state lifetimes τ of the D-only reference sample $\tau_{D(0)} = 1/k_{D,0}$ and the DA sample $\tau^{(i)}_{D(A)} = (1/k_{D,0} + 1/k^{(i)}_{RET})$ (Eq. (1d)). These correspond to the slopes of the decay curves (see Box 1, panel 1). (2) The joint decay analysis of the D-only reference $f_{D(0)}(t)$ and the FRET sample $f_{D(A)}(t)$ allows relating the FRET rate constant $k^{(i)}_{RET}$ directly to $R^{(i)}_{DA}$ (Eq. (2), for details see [65]). (3) The curvature of $f_{D(A)}(t)$ contains information on the heterogeneity of the FRET sample such that a distance distribution $x(k^{(i)}_{RET})$ with species fractions $x^{(i)}$ can directly be resolved (Eq. (2)).

$$
\begin{aligned}
&f_{D(0)}(t) = e^{-(k_{D,0})\cdot t},\\
&f_{D(A)}(t) = f_{D(0)}(t)\cdot \int_0^\infty x(k^{(i)}_{RET})\cdot e^{-k^{(i)}_{RET}\cdot t}\, dk^{(i)}_{RET} \text{ with}\\
&k^{(i)}_{RET} = k_{D,0}(\Phi_{F,D(0)})\cdot \left(\frac{R_0(n,\kappa^2,\Phi_{F,D(0)})}{R^{(i)}_{DA}}\right)^6
\end{aligned}
\tag{2}
$$

Depending on the complexity of the sample, FRET experiments can be performed on the single-molecule, sub-ensemble (selectively averaged single-molecule events), and ensemble levels, where each technique has its advantages. Intensity-based ensemble FRET measurements are relatively easy to perform; however, these experiments yield only average observables due to ensemble averaging over the mixture of the molecules in the sample (Eq. (3)), and the sample must be well-characterized (purity, degree of labeling, homogeneity, fluorescence quantum yields of D and A) [59••]. Ideally, for such a mixture of N species with fractions $x^{(i)}$ and transfer efficiencies $E^{(i)}$, the average transfer efficiency is given by:

$$
E = \sum_{i=1}^{N} x^{(i)} E^{(i)} \tag{3}
$$

Single-molecule measurements are currently the most common type of FRET experiments, because they can resolve distributions of FRET observables and provide kinetic information over 15 orders of magnitude in time. The main advantage of sm fluorescence spectroscopy is the ability to resolve static (multiple distinguishable static states) and dynamic (interconverting states) heterogeneities. There are two main measurement setup types for smFRET which differ significantly in their time resolution: (1) total internal reflection fluorescence (TIRF) microscopy [66•] with emCCD camera detection and (2) confocal detection with fast single-photon counting point detectors [67]. TIRF is widely used for immobilized molecules. The experimental time resolution is usually limited by the frame rate (~60 Hz) of the camera. Recently Farooq and Hohlbein [68] presented a generalized excitation scheme (sALEX) that combines the concepts of alternating laser excitation (ALEX) [69,70] and stroboscopic illumination [71,72] to improve the time resolution at least 10 fold. Confocal setups can be used for studying immobilized and freely diffusing molecules. Felekyan *et al.* [73] presented a dead-time free configuration with four detectors and two data acquisition boards, which enables time-resolved measurements over 15 orders of magnitude in time (Box 1). The immobilization of single molecules on a coated glass surface can be either directly accomplished via tags [66•] or indirectly via encapsulation in a liposome, which is immobilized afterwards [74].

Sm experiments with confocal detection of freely diffusing molecules are characterized by reproducible and stable detection efficiencies. In combination with pulsed excitation or pulsed interleaved excitation (PIE) [64•,75•] time-resolved fluorescence spectroscopy with multi-parameter fluorescence detection (MFD) [60••,76,77] can be conveniently realized for all FRET relevant fluorescence observables in a single measurement with the same sample. Most importantly, time-resolved anisotropy decay curves are obtained to estimate the uncertainty of the FRET orientation factor κ^2 [61•]. Additionally, a series of combined 2D fluorescence parameter histograms can be generated, which allows determining all necessary FRET calibration parameters [60••,64•]. Finally, MFD increases the species selectivity, which is essential for

Box 1 A combination of fluorescence spectroscopy methodologies records dynamics over 15 orders of magnitude in time and captures a large variety of biologically relevant processes such as conformational transitions or folding/unfolding events [22,29•,60••,150].

The fluorescence lifetime of typical fluorophores decays in the pico-second to nanosecond range and sets the lower limit to interrogate conformational dynamics. The upper limit is set by the time a single molecule can be observed. Immobilized molecules can be studied for seconds while fluorescence bursts due to freely diffusing molecules are observed only for milliseconds in confocal setups.

Experimental information can be obtained by (1) time-correlated single photon counting (TCSPC) histograms, (2) fluorescence correlation spectroscopy (filtered FCS), (3) burst-wise single-molecule analysis, and (4) the analysis of FRET efficiency traces. The influence of the exchange dynamics on the observed data is outlined for a system with two FRET states, Low-FRET (LF) and High-FRET (HF), which are either in fast (sub milliseconds, lower row) or in slow (seconds, upper row) exchange.

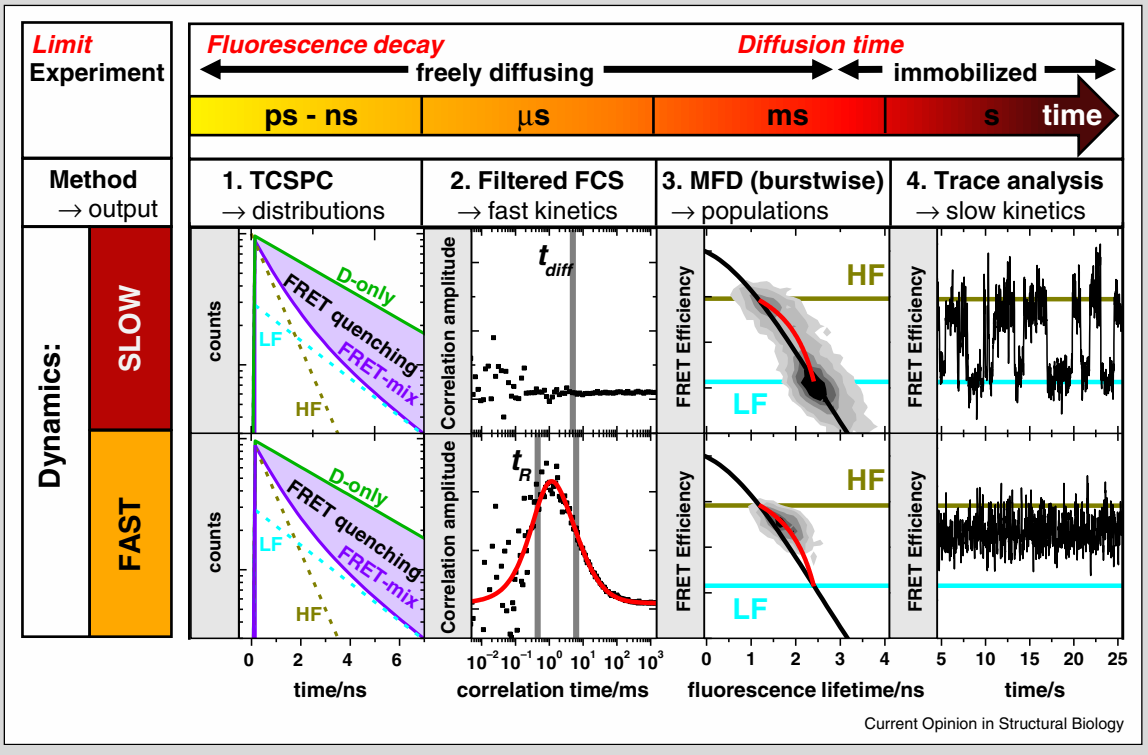
(1) In the analysis, fluorescence decays of FRET rate constants of HF and LF are averaged over the fluorescence lifetime of the dye (nanoseconds). Therefore, a joint analysis of a FRET sample and a D-only sample (shown in green) provides a static picture of a molecular ensemble and decomposes into HF and LF components shown as dashed dark yellow and cyan lines, respectively.

(2) In FCS, fluorescent photons are correlated to provide kinetic information in the submicro-second to milliseconds regime. Filtered FCS (fFCS) additionally utilizes the photon distribution with respect to an excitation pulse [23•] to highlight the kinetic exchange between the HF and LF species by the species cross-correlation functions (SCC, black dots). If the kinetic exchange is significantly slower than the observation time (t_{diff} – diffusion time), no correlation is observed. Otherwise, a pronounced anti-correlation in the SCC is visible, which is analyzed (red curve) to yield the relaxation time (t_R) of the HF/LF exchange.

(3) In burst-wise analysis, fluorescence is integrated over milliseconds to provide the FRET efficiency and fluorescence lifetime of freely diffusing single molecules. A 2D-histogram allows interrogating the kinetics between states and reveals dynamics by changes in the shapes. Theoretical static (black) and dynamic FRET lines (red) serve as references for the location of the FRET populations in the 2D parameter histograms. A shift of the observed population towards longer lifetimes with respect to the static FRET line indicates conformational dynamics.

(4) By recording fluorescence intensities of immobilized or liposomal encapsulated single molecules over a longer time with a lower time resolution, processes significantly slower than the typical burst duration (>ms) are interrogated. FRET efficiency time traces of single molecules reveal stochastic jumps between HF and LF if the integration time of each frame is shorter than the dwell time of the state (upper plot); these jumps are averaged if the dwell time of the states are short (lower plot).

Box 1



Fluorescence timescales to study molecular kinetics using FRET.

sub-ensemble TCSPC of sub-populations, that is, selective averaging by integrating all single-molecule events of a resolved FRET species to reduce the data noise dramatically (for details, see Box 1).

To perform quantitative FRET measurements, all procedures for determining the calibration factor γ in Eq. (1) (TIRF [66•,78,79•,80], confocal [60••,64•]) rely on well characterized standards, which are usually DNA rulers

labeled with a FRET pair using the dyes of interest. A worldwide consortium of research groups on FRET is currently working to establish a set of common recommendations for measurement procedures, data analysis, and joint reference molecules to maximize the accuracy of the FRET observables measured in different laboratories.

Description of the labels

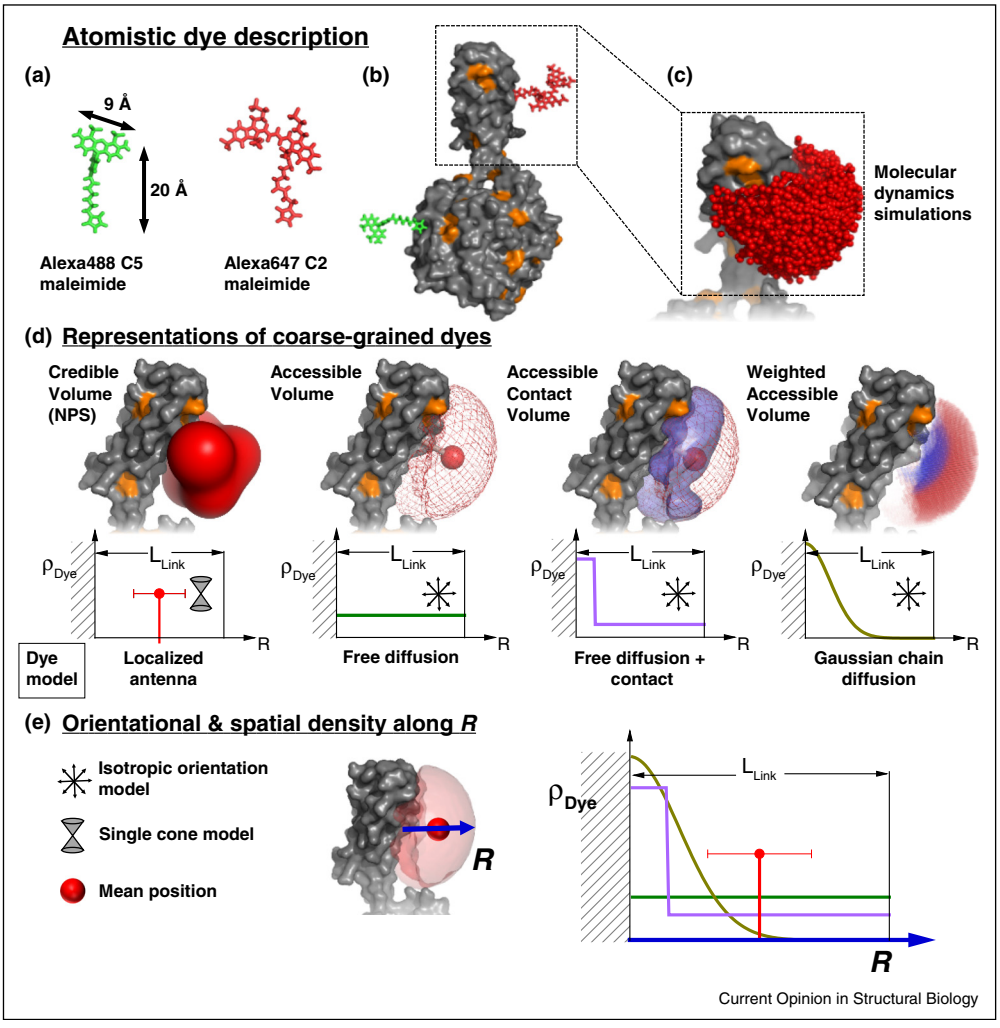
Both EPR spectroscopy with the double electron–electron resonance (DEER) method [81] and FRET spectroscopy are frequently used to obtain structural information on proteins by selectively coupling labels via flexible linkers (i.e., nitroxide spin-labels or donor and acceptor dyes, respectively). Thus, the labels can explore a large conformational space, such that a distribution of inter-label distances rather than a single inter-label distance is observed experimentally. In structural analysis, both techniques face the same problem of describing the localization of their respective labels and connecting the measured distances with structural information on the protein. This is a typical inverse problem, where both EPR [81–84] and FRET [42••,85••,86••] spectroscopies have developed similar solution strategies of assuming a structural model to compute a specific distance distribution between the labels and comparing the simulated and experimental values. Analogous to super-resolution microscopy, where the knowledge of the optical point spread function allows for localizing single emitters far beyond the optical resolution limit, a detailed knowledge of the label distribution for a given biomolecular structure is required to maximize the structural resolution of the respective technique. The so far used dye models are reviewed in this section, and their implementation in various toolkits for integrative FRET-based structural modeling is described in Section 'Procedures for hybrid-FRET modeling'.

All-atom molecular dynamics (MD) simulations [33,87–93] seem to be the most appropriate solution to describe the label localization in integrative modeling (Figure 1a–c). However, while the relevant time-scale of fluorescence and biomolecular dynamics covers picoseconds to hours (Box 1), current MD simulations reach only the micro- to millisecond timescale. Furthermore, MD simulations are time-consuming and limited in their capability to sample the conformational space, in particular since the dye distribution changes with conformational changes of the biomolecule. MD simulations [87,89,94] as well as experimental fluorescence [61•,95,96•] and EPR [97] studies find long-lived conformational states, where the labels are trapped on the biomolecular surface. Trapping of cyanine dyes is known to change their fluorescence quantum yields [96•] such that this process is utilized by others for probing protein-nucleic acid interactions by protein-induced fluorescence enhancement (PIFE) [98,99•]. To avoid quenching and mobility reduction by the surface effects, most commercially available dyes used for FRET experiments have long flexible linkers (~20 Å), and the fluorophore carries sulfonic acid groups for improved solubility. The large flexibility of the linker also ensures a high orientational mobility of the dyes such that the assumption of the isotropic average for the dipole orientation factor (i.e., $\kappa^2 = 2/3$) is feasible, which is an important factor determining the accuracy of distance calculations by FRET measurements. In conclusion, the atomistic treatment of dye labels, together with the biomolecular system, by MD simulations is helpful for analyzing protein label interactions. However, it is currently not feasible to use MD simulations as a rapid routine method to predict the conformational sampling of the labels, which is needed for finding the structural model within a large ensemble that is most consistent with FRET data, as shown below.

Computationally faster algorithms use a coarse-grained representation, e.g. the label is approximated by a flexibly linked sphere ([83,85••], see Figure 1a–c), to compute the sterically allowed accessible volume (AV), which reproduced MD simulations well [61•]. The use of a geometric search algorithm reduces the computational time for the AV of a single FRET-label to less than 50 milliseconds on a conventional desktop computer [42••]. The computed AV represents uniform population densities ρ_{Dye} of the spatial dye localizations (Figure 1d). The mesh represents the surface of the dye positions most distant from the attachment point and includes all sterically allowed dye positions. To compute fluorescence observables the spatial population densities $\rho_{Dye}(x,y,z)$ of the dye localizations is needed.

Two convenient assumptions for limiting cases were initially used for the models describing the spatial and orientational dye density in various hybrid-FRET modeling approaches: (1) The original dye model of the Nano Positioning System (NPS) assumed a label adopting a single position with a defined conformation of the linker and the fluorophore wobbling in a cone (*Localized antenna/Single cone model*) [85••]. This means that all fluorophores of an ensemble of molecules are located at the same position (defined by a delta function δ) with respect to the macromolecule: $\rho_{Dye}(x,y,z) = \delta(x_{Dye},y_{Dye},z_{Dye})$. Since the position $(x_{Dye},y_{Dye},z_{Dye})$ of the dye is initially unknown, multiple measurements are used to localize it within a credible volume for a given confidence level (Figure 1d, left). (2) As an approximation of a freely diffusing fast rotating dye all orientations and all positions within the AV are equally populated; $\rho_{Dye}(x,y,z)$ is constant and the fluorophore is isotropically oriented within the AV (*Free diffusion/Iso model*) [42••,61•]. The effect of different dye models on the spatial population densities $\rho_{Dye}(x,y,z)$ along the vector R, pointing from the linker attachment site to the mean dye position in the AV, is visualized in the lower panel of Figure 1d.

Figure 1



Description of fluorescent labels in simulation. Treatment of dye labels in FRET-based modeling. **(a)** The donor and acceptor dyes, for example, Alexa488 C5 maleimide (green) and Alexa647 C2 maleimide (red), respectively, are typically attached to the biomolecule via long (~20 Å) flexible linkers. **(b)** Dye labels attached to the protein Atlastin-1 (grey surface) to give an impression of the sizes. Quenching amino acids on the protein surface are highlighted in orange. **(c)** Molecular dynamics simulations provide the spatial distribution of dye molecules. **(d)** Representation of coarse-grained dye labels. On the top different dye representations are shown. On the bottom the spatial density ρ_{Dye} along a vector R starting at the attachment point in the direction of the dye mean position is shown for the corresponding dye models. The original Nano Positioning System (NPS) [85••] assumes the accessible volume as *prior* information (uncertainty), which is reduced by a set of FRET measurements with fixed satellites resulting in an uncertainty distribution (red) of a putatively *fixed* localized dye (antenna). The orientation of the dyes follows a diffusion in a cone model highlighted by a pictogram. Following R, the dye is located at a specific position (vertical red line) with an uncertainty. Accessible volume (AV, mesh) models provide the sterically allowed space of the dye molecule attached to the protein as calculated by the FPS program [42••]. Here, the linkers of Alexa488 and Alexa647 are approximated as flexible tubes with width of L_{width} = 4.5 Å and length of L_{link} = 20.5 Å and L_{link} = 21.0 Å, respectively. Three radii were used to describe the dyes (5.0 Å, 4.5 Å, 1.5 Å) and (11.0 Å, 4.7 Å, 1.5 Å) for Alexa488 and Alexa647, respectively. The large sphere indicates the mean dye position. For a dye freely diffusing inside of its AV a uniform spatial distribution is assumed. The accessible and contact volume (ACV, [102]) provides a similar description as the AV, but defines an area close to the surface as contact volume (violet). In this figure the density ρ_{Dye} in the contact volume is weighted six times stronger and defined as part of the AV that is closer than 3 Å from the macromolecular surface. For the corresponding model, where a dye freely diffuses within the AV and its diffusion is hindered close to the surface, the spatial density ρ_{Dye} along R is approximated by a step function: ρ_{Dye} ($R < 3$ Å) = $6 \cdot \rho_{Dye}$ ($R \geq 3$ Å). The weighted accessible volume is a modification of AV where ρ_{Dye} along R is approximated by a Gaussian chain-inspired empirical weighting function [101]. To illustrate the effect of the proposed weighting of ρ_{Dye} the weighted AV is colored from blue (high density) to red (low density). **(e)** Comparison of the normalized spatial population densities ρ_{Dye} of the above dye models.

More recently, other dye models were suggested [100•,101,102]. To consider small fractions of fluorescent fluorophores trapped on the biomolecular surface [102], the classical AV description was extended (using the *Free diffusion/Iso model*) by defining an additional interacting surface layer (the contact volume of defined thickness is highlighted in violet in Figure 1d) with a preferential surface residence of the dye, which may be approximated by a step function. This surface preference is specific for the dye and the labeling site. This combination yields the *Free diffusion + Contact/Iso model* for an accessible and contact volume (ACV), which predicts the compaction of the spatial dye distribution due to trapped dye fractions without significant increase of the computational cost. Höfig *et al.* [101] introduced a weighted AV description, where the densities ρ_{Dye} are enriched by empirical weighting with a Gaussian function leading to a higher local dye surface preference but only close to the attachment point (*Gaussian-chain (GC) diffusion/Iso model*) (Figure 1d,e).

The comparison of the normalized densities ρ_{Dye} of the four dye models in Figure 1e shows that the mean dye positions of an ACV or a weighted AV are closer to the protein interface as compared to the AV with a uniform density. This dye model specific shift of the mean positions becomes important for the predicted interdye distances if the other dye of the FRET pair is attached on the opposite side of the protein. To avoid an arbitrary choice of dye models, it is important to define rational criteria based on experimental evidence through observable fluorescence parameters. For example, time-resolved fluorescence data map the translational and orientational dye diffusion and the resulting DA distance distributions by fluorescence quenching, anisotropy and FRET.

So far we have discussed four dye models that have been developed for synthetic organic fluorescent labels, which are usually not too large (Figure 1a,b) so that their steric requirements are small. However, if the more bulky fluorescent proteins (FPs), such as GFP and mCherry, are used as labels in fusion proteins, we have to account for the significant steric hindrance of the FPs and their linkers. For this case, self-avoiding walk (SAW) weighting of the linker configuration (*SAW-Diffusion/Iso model)* was applied in our group. The scaled weighting factors were determined by computationally tractable MD simulations for sampling configurations of an explicit linker/label combination alone, followed by calculations of conformational free energies to weight each linker/label configuration in the presence of the biomolecule [103]. We applied this dye model in live cell hybrid-FRET studies for analyzing the architecture of biomolecular complexes labeled with FPs ([103,104•], further details see outlook). Notably, this *SAW-Diffusion/Iso model* has similar effects on the conformational space in the AV as the rotamer library model that is widely used for the analysis of DEER experiments in EPR spectroscopy [81]. Unlike the weighted AV description, the steric effects in the SAW-weighting lead to a strong reduction of ρ_{Dye} in proximity to the attachment site so that ρ_{Dye} has a maximum far away from the surface.

Error estimation

For quantitative FRET studies, a number of control tasks should be performed for an appropriate analysis and interpretation of the experimental observables [60••]: (1) analysis of the fluorescence properties of D or A by analyzing multiple fluorescence parameters with respect to quenching and trapping [37,60••], which could affect R_0 [105] and the calibration of the intensity-based experiments; (2) interpretation of discrete distributions of FRET efficiency E and anisotropy in confocal single-molecule experiments by photon distribution analysis, which allows separating structural heterogeneities from stochastic variations [106•,107]; (3) validation of the chosen dye model by analyzing (i) the linker-induced distance distributions of single FRET states by TCSPC or sub-ensemble TCSPC, (Eq. (2)), and (ii) the time-resolved anisotropy to scrutinize the dye mobility [61•] for verifying the assumption of $\kappa^2 = 2/3$; (4) verification by sub-ensemble TCSPC and FCS (Box 1) that a population with a single FRET efficiency is indeed a single species and that it is not narrowed by a dynamic exchange (i.e. that it is actually a mixture of fast inter-converting states); and (5) verification by functional assays (e.g. binding constants of ligands, catalytic activity, structural stability sensed by thermal or chemical denaturation) that the label does not disturb the biomolecule.

The analysis of FRET measurements provides the inter-dye distances R_{DA} as sparse input data. To take advantage of this information, rigorous error estimation is necessary for integrative modeling such that all uncertainties that have been considered correctly propagate to the final structural model [30•]. According to the error propagation rules, the overall uncertainties of the experimental inputs (the inter-dye distances R_{DA}) for the search of structural model contain three major contributions.

(1) In single photon counting, the statistical relative error is inversely proportional to the square root of the registered photons that translate to the experimental distance uncertainty, $\Delta R^2_{\mathrm{DA}}(E)$, via the experimentally determined FRET efficiencies E or FRET rate constants k_{RET}. Systematic error in $\Delta R^2_{\mathrm{DA}}(E)$ is introduced primarily by two factors: (i) Preprocessing of experimental data by approximate fluorescence models (e.g. Eq. (2)) and (ii) Inaccurate instrumental calibrations and reference samples. Furthermore, if dynamic multi-component systems are studied, $\Delta R^2_{\mathrm{DA}}(E)$ usually also contains the uncertainty due

to data ambiguity and incoherence, resulting in FRET observables being correlated [30•]. To obtain a distance set for a specific FRET species (i.e. conformer) from analyzing a large set of studied FRET pairs, species fractions and specific kinetic properties [27••], which are global (joint) for the whole data set, are the best choices to assign FRET observables to the corresponding specific FRET species. The reason is that a global analysis of multiple independent datasets by a joint model minimizes the ambiguity by making use of correlations between otherwise independent model parameters. A global analysis procedure will be demonstrated below using a sample test case.

(2) The uncertainty $\Delta R^2_{\mathrm{DA}}(\kappa^2)$ of using the correct orientation factor κ^2 directly influences R_0, so that the accuracy of the computed distances, R_{DA}, is affected. The error range of κ^2 can be estimated by anisotropy measurements [61•].

(3) The uncertainty $\Delta R^2_{\mathrm{DA}}(\vec{R}_D, \vec{R}_A)$ of using an appropriate dye model describing the spatial positions $\vec{R}_D, \vec{R}_A$ of D and A, respectively, influences the accuracy of predicting appropriate distances from structural models.

In the end, all the errors above propagate to structural models obtained via sampling of the conformational space and define the accuracy of the structural models with respect to the true structure and the precision of the model ensemble.

Procedures for hybrid-FRET modeling

Due to the sparse information provided by experimental techniques such as FRET or EPR measurements, given the large number of degrees of freedom in atomistic models of biomolecules, a complete determination of biomolecular models with atomic resolution using only these experimental data is impossible. Fortunately, computational methods can provide complementary information, such as detailed structural features of biomolecules. At the same time, many computational methods also face challenges such as sufficient sampling and correct identification of near-native structures in a structural ensemble. The combination of computational methods with experimental input within a hybrid approach can facilitate generation and verification of the detailed biomolecular structures [108•], because experimental FRET data can provide external information on the architecture of proteins and biomolecular complexes to guide the modeling procedure.

Although many FRET-based structural studies were performed since Förster's publication in 1948 [1••], the obtained structural models were shown primarily as cartoons. In recent years, a number of integrative computational modeling methods [33,51,85••,86••,89,109] were developed to derive and present structural models in a more quantitative manner. The high accuracy and performance of hybrid-FRET modeling was demonstrated in two benchmark studies with static structures [41•,42••]. In the last decade, many applications for small and large complex biomolecular systems were published, albeit the models remained on a rather descriptive level. Considering flexibly linked dyes, currently three software toolkits, Nano Positioning System (NPS) [85••], Crystallography & NMR System (CNS) [88,110] and FRET-restrained positioning and screening (FPS) [42••], are publicly available; they implement different methodologies, use partially different assumptions and dye models. Therefore, common standards for FRET analysis, generally available joint toolkits with well-defined workflows and widely accepted procedures should be established to facilitate quantitative hybrid-FRET modeling.

The key to hybrid modeling lies in the fine interplay between the computer simulations and the experiments to accomplish the most effective synergies between the strengths of both sides. On the experimental side, accuracy is achieved by appropriate consideration of spatial dye distributions, while precision estimation stems from rigorous error analysis. Moreover, the establishment of a statistically appropriate quantitative scoring function for judging agreement between the structural models and FRET observables is far from trivial. Three major routes currently exist to find a structural model satisfying FRET observables best. First, in the screening approach used by FPS, computational methods are used to initially generate a conformational ensemble that is subsequently quantitatively evaluated in terms of its agreement with the experimental FRET data. Second, in the multibody docking approach, used by CNS and FPS, the integration of FRET data into the computational modeling is achieved by addition of experiment-based distances with corresponding uncertainties as restraints that define a harmonic potential for the fluorophore center positions treated as pseudoatoms rigidly connected to the biomolecule [42••,86••]. The defined ''FRET-forces'' induce the docking process of the labeled domains and molecules. Many iterations (typically $> 10\,000$) with distinct starting conditions (i.e. different randomly assigned orientations of domains and molecules, different relative conformations of flexible domains) are usually performed for each set of FRET distance restraints to cover the configurational space. As a third option, the Markov Chain Monte Carlo (MCMC) [111••] approach, among others, allows circumventing the difficulties with geometric restraints. Examples for the use of MCMC guiding in conjunction with data from FRET experiments include the generation of an open conformation of Syntaxin 1 [48] and the determination of structures of large biomolecular complexes, such as a RNAP complex [36].

Appropriate dye models, in combination with quantitative FRET studies, are crucial for the accuracy of

integrative modeling. Previously [37,85••], studies utilizing the toolkit for the Nano Positioning System (NPS) have assumed the *Localized antenna/Single cone model* and have used Bayesian parameter estimation for three dimensional dye localization to extract structural information from a network of FRET measurements. The AV has been used as a *prior*, which is refined by experiments to resolve potential locations of the dyes as credible dye volume (*posterior*). The inclusion of more complex dye models into NPS has been introduced recently [100•,112]. In its current stage of development, NPS tests the consistency of a given structural model with DA distance sets by comparing the overlap of these credible dye volumes (*posterior*, depicted in Figure 1d) with the AVs (*prior*) of the considered structural model. Initially, CNS and FPS both used the *Free Diffusion/Iso model* to consider the distribution of flexibly linked dyes. In a benchmark study with FPS, docking a DNA primer-template to HIV-1 reverse transcriptase, Kalinin et al. have shown that it is crucial for the accuracy of the docked complex to explicitly consider the spatial dye distribution [42••]. CNS uses fluorophore center positions as pseudoatoms rigidly connected to the biomolecule. The position of the pseudoatom is taken as the average position of the fluorophore relative to the molecule as obtained from a molecular dynamics simulation [86••,88]. Owing to the explicit treatment of the dye distance distributions, FPS can serve as an appropriate interface for this experimental input. Thus additional dye models have been included in FPS meanwhile for improving the accuracy of hybrid-FRET modeling for specific combinations of systems and labels: (1) *Free diffusion/Iso model* for organic fluorophores tethered to nucleic acids [42••,61•], (2) *Free diffusion + Contact/Iso model* for organic fluorophores tethered to proteins (this work), and (3) *SAW-Diffusion/Iso model* for fusion proteins [103,104•]. Finally, Preus *et al.* [35•] developed the toolkit FRET-matrix for modeling FRET between probes possessing limited diffusional and rotational freedom. This toolkit is especially useful for localized fluorophores in nucleic acids as a replacement for one of the natural bases [113] giving the advantage of reporting from internal sites of interest.

Currently, structural models from hybrid-FRET modeling cannot be deposited in the Worldwide Protein data Bank (wwPDB) [114,115]. At present, the authors have two alternatives to archive the coordinates of suggested structural models: (1) provide them as supplementary information with the publication or (2) deposit them in the Model Archive (MA, www.modelarchive.org), which assigns a unique stable accession code (DOI) to each deposited model. The Model Archive is being developed following a community recommendation during a workshop on applications of protein models in biomedical research [57] as part of the Protein Model Portal (http://www.proteinmodelportal.org; [116]). Advantageously, it offers the opportunity to apply consistent assessment and validation criteria to the complete set of structural models available for proteins. Moreover, it allows providing in depth information about the simulations performed and the parameters and constraints used. Therefore we decided to deposit all data of the presented test case study in the Model Archive with the DOI: 10.5452/ma-a2hbq.

Finally, it is important to note that the wwPDB is aware of the need for a large and general repository for structural models of biomolecular systems that have been obtained by integrative modeling using varied types of experimental data and theoretical information [38]. Establishing community-wide accepted standards for measuring, analyzing, and describing FRET data is an additional step necessary for organized data deposition.

Hybrid-FRET modeling on rails: a case study

Which information can be obtained by a hybrid-FRET study? The aim of hybrid-FRET modeling is to describe macromolecules that potentially adopt multiple dynamically exchanging conformations in thermal equilibrium by three-dimensional structural models derived from prior structural knowledge and multiple quantitative FRET measurements [42••,88]. To provide an informative answer, our hybrid-FRET modeling procedure is organized such that a spectrum of possible solutions (structural arrangements) is collected from all structural background information (e.g. X-ray structures, homology/comparative models and/or at least educated guesses) whose suitability is under question. We use this initial information (*prior*) in step I of the workflow considering two perspectives. From the first perspective, experiments are designed to *confirm* our *prior* structural knowledge. If the experiments prove to be inconsistent with this *prior*, we consider the design also from the second perspective that the experimental information can be used to generate a three-dimensional structural model as *posterior* hypothesis. To test our *prior* (i.e. the structural background information), we seek DA pairs with the largest power in proving our null hypothesis that the initial structural information is incorrect (step I of the workflow). We identify such DA pairs by exploring a range of possible conformational motions of the system, given the *prior* knowledge. At this stage, unbiased and extensive coverage of the structural space is more important than the accuracy of structures or the density of the generated ensemble. To extract the most useful information from the system of interest, we established an iterative workflow for hybrid-FRET modeling consisting of five steps:

Step I: Defining specific questions about the biomolecular system of interest and initializing modeling by collecting *prior* knowledge to generate an initial structural ensemble and to determine the most informative DA pairs.

Step II: Performing calibrated FRET measurements and quantitative analysis.

Step III: Computational modeling to improve the sampling density of the initial ensemble guided by the obtained experimental FRET data. Identifying the best structures by screening and computing the precision of the obtained ensemble.

Step IV: Judging the agreement between the modeled structure and experimental data to decide whether the steps I to III should be repeated.

Step V: Assessing the quality of the obtained conformational ensemble with respect to the FRET data (precision and accuracy) and to the structural modeling (judgement by short-range stereochemical criteria).

Choosing informative DA pairs — an experimental design problem. The choice of DA pairs for FRET experiments is an essential initial step. It represents a classical experimental design problem [117] which influences the obtained structural model in terms of the precision and accuracy. More explicitly, the question is, which DA pairs should be measured such that the obtained information is maximized with respect to its power to discriminate structures of the initial ensemble. Two cases need to be distinguished for the selection of DA pairs. If structural pre-knowledge is unavailable or not fully considered, the localization of the fluorophores themselves is of interest. Thus, connected networks of DA pairs are needed and the dyes are localized by trilateration approaches [37,100•].

In the second case, if structural pre-knowledge is available, usually not all DA pairs are equally informative. Therefore it is advantageous to select a small set of most informative DA pairs because in real experiments, acquisition of reliable data is challenging due to the involvement of multiple steps, that is, mutagenesis, protein expression, purification, biochemical activity assays and measurements.

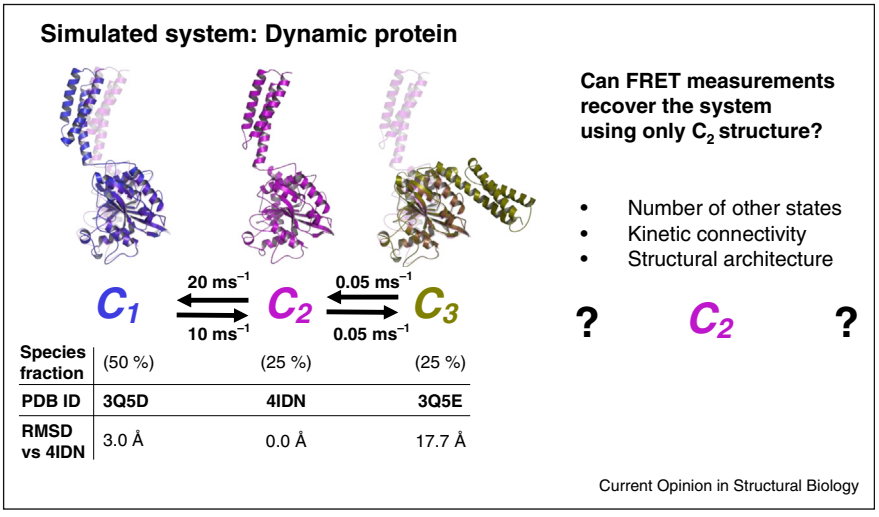
Here, we define distance sets to be the most informative if they lead to the highest expected precision of a structural model. To achieve high precision with a minimum set of DA pairs $\{DA^{(i)}\}$ for a well-planned experiment, we formalize the design of $\{DA^{(i)}\}$ to automate generation of FRET-based structural models. Thus, we establish a quality parameter, $\langle \Delta_{set} \rangle_{ref}$, for a given $\{DA^{(i)}\}$ by defining its expected precision (uncertainty). First, each conformer is used as a reference to compute the reference-specific precision Δ_{set} for $\{DA^{(i)}\}$. Here the precision corresponds to the weighted average C_α atom RMSD over all conformers with respect to the chosen reference conformer. The weight of the conformer is determined by its confidence in being a worse model compared to the reference model as judged by squared, error weighted deviations. Second, the average over the reference-specific precisions yields $\langle \Delta_{set} \rangle_{ref}$. To find an informative distance set $\{DA^{(i)}\}$, we apply a greedy backward elimination feature selection algorithm [118] to minimize $\langle \Delta_{set} \rangle_{ref}$. In this algorithm, DA pairs that lead to the smallest decline of $\langle \Delta_{set} \rangle_{ref}$ (i.e. smallest increase in RMSD) were iteratively excluded from the set. As the set of possible DA pairs is limited to several thousand pairs, the computational burden of this algorithm is low enough to complete within ~24 h on a desktop PC.

For the first iteration of the hybrid-FRET modeling workflow (Figure 3), ten distances were selected (Figure 4a). In real experiments, suggested DA pairs might be discarded to minimize errors (see Section 'Error estimation'). Structural aspects relevant for biochemical activity (binding pockets, stabilizing salt bridges, and so on) and fluorescence properties (quenching amino acids) should be considered in the selection procedure to minimize the effects of the labels on the biomolecular function and to optimize their fluorescence properties.

Generating hybrid-FRET models for a heterogeneous system in the presence of fast dynamics for the test case Atlastin-1. To study the accuracy and precision of hybrid-FRET modeling, we designed a test case and simulated a protein with three exchanging conformers. The used structures of the conformers are structurally well-characterized by X-ray crystallography. We simulated typical data traces of single-molecule FRET experiments with multi-parameter fluorescence detection for typical conditions [60••,119]. In this way, the known X-ray structures can serve as unquestioned references for the assessment of the structural models obtained by hybrid-FRET modeling using our FPS toolkit.

We chose to study the cytoplasmic part (residues 1–446) of the large GTPase human Atlastin-1, which belongs to the dynamin superfamily. This part consists of two flexibly connected domains, the N-terminal GTPase domain (G domain) and the C-terminal middle domain, for which three distinct conformations in the functional cycle were found by X-ray crystallography. For the simulations of smFRET experiments, we chose the monomeric subunits of two crystallographic dimers (PDB-IDs 3Q5D [120] and 4IDN [121]) and a tetramer (PDB-ID 3Q5E [120]), depicted in Figure 2, named here C_1, C_2, and C_3. To the best of our knowledge, nothing is known about conformational dynamics of Atlastin-1 in solution. Thus, in a Gedankenexperiment, we assumed the existence of a putative dynamic exchange equilibrium with the linear reaction scheme $C_1 — C_2 — C_3$ (Figure 2). The conformer C_3 differs significantly from the conformers C_1 and C_2. In contrast, the structures of conformers C_1 and C_2 are similar to each other (RMSD 3.0 Å). Therefore, we simulated a fast exchange between C_1 and C_2 (kinetic

Figure 2



Presentation of the test case Atlastin. Overview of the benchmark test system. For the benchmark test, we used three crystal structures of Atlastin-1, which are termed C_1, C_2, and C_3 (PDB-ID: 3Q5D (chain A) [120], 4IDN (chain A) [121], and 3Q5E (chain A) [120], respectively). We simulate the kinetic exchange between C_1 and C_2 to be fast, and the exchange between C_2 and C_3 to be slow. Our task was to recover the number of other states, their kinetic connectivity, and their structures, given only the structure (PDB-ID: 4IDN) and information from smFRET spectroscopy.

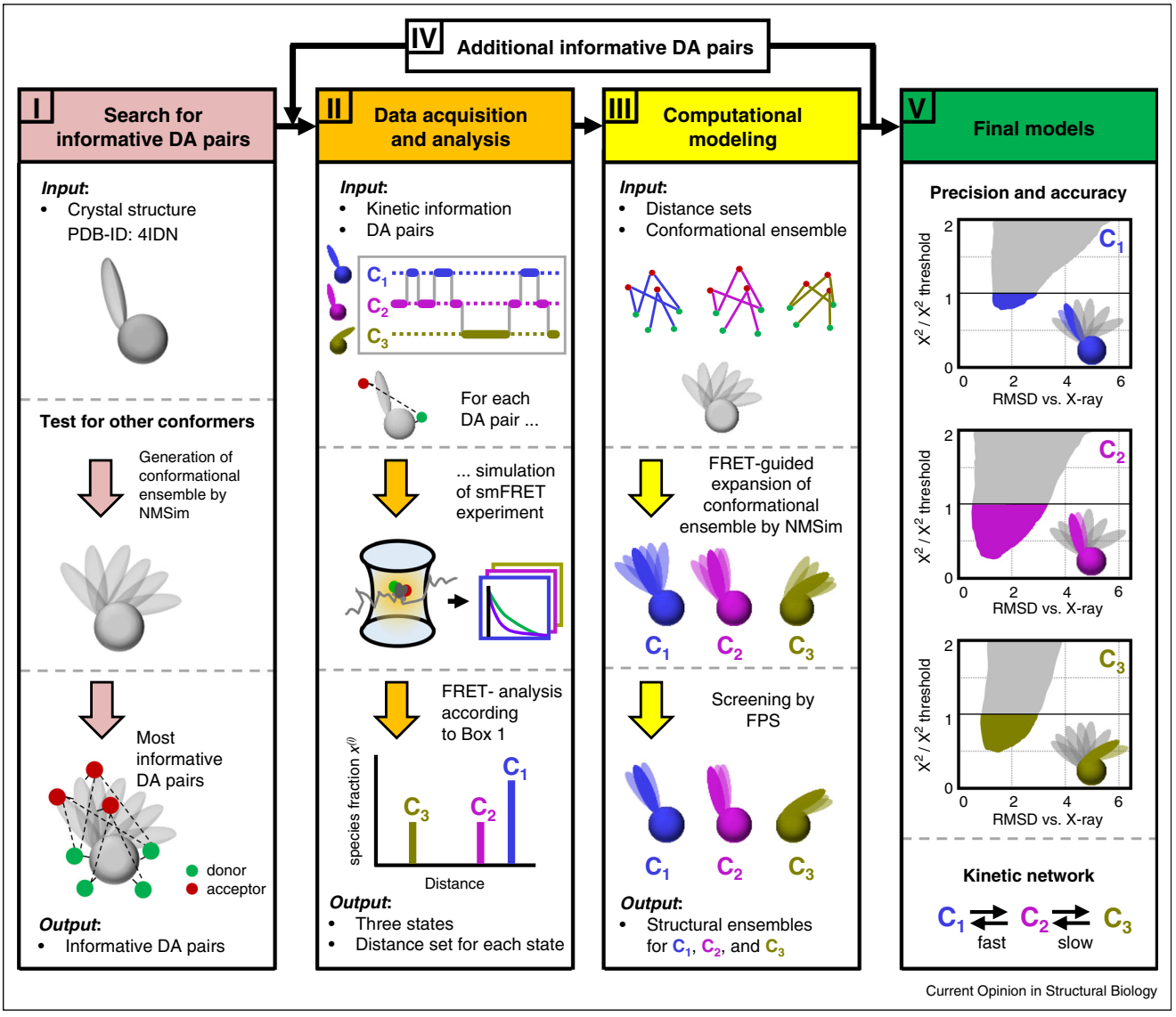
relaxation time $t_{R,1-2} = 30$ μs) and a slow exchange between C_2 and C_3 ($t_{R,2-3} = 10$ ms). In the test case, we utilized the crystal structure (PDB-ID: 4IDN) as a *prior* for structural simulations. Note, however, that all experimental FRET observables were simulated from the three ''true'' crystallographic structures to test whether they could be recovered by hybrid-FRET modeling. Thus, we aimed at finding suitable structural models for C_1, C_2, and C_3 and recovering the kinetic exchange equilibrium by treating the simulated photon traces identical to single-molecule FRET experiments using a data generator (for details see [119]) that generated photon streams of typical single-molecule experiments by Brownian dynamics (BD) simulations.

In real experiments, fluorescence decays are complex due to DA distance distributions, brightness distributions due to the confocal excitation profile and experimental nuisances such as instrumental response function and detector dark counts. All these aspects are reproduced by the simulations of freely diffusing molecules, to generate realistic photon traces. Technical details are given in detail in [119] and are outlined below. As in real single photon counting experiments, the Poissonian statistics determines the experimental noise and thus the statistical errors of the subsequent analysis. The simulations correspond to ~3 measurement hours with current instrumentation in the Seidel group. In this test case, we want to study whether the simulated data of typical experimental quality allows us to recover accurate hybrid-FRET models and explore experimental limits on their precision.

FRET on rails: Step I. We started our hybrid-FRET modeling (Figure 3) by testing the null hypothesis that our *prior* will *not* be described by experiments. To design an optimal set of DA pairs for this test, we use the toolkit NMSim to generate an ensemble unbiased by experimental FRET data with 400 000 conformers (RMSD from PDB-ID: 4IDN up to 26 Å) as alternative candidates to the *prior*. NMSim is a normal mode-based geometric simulation approach for multiscale modeling of protein conformational changes that incorporates information about preferred directions of protein motions into a geometric simulation algorithm [122,123]. The obtained ensemble was clustered and used to select an optimal set of DA pairs according to three criteria. At first, all residues that are positioned on the protein surface and are located far enough from quenching amino acid residues (Trp, Tyr, His, Met) [124] were selected. Secondly, from all pairwise combinations of these labeling positions, those that result in average inter-dye distances >30 Å were selected. Thirdly, we determined the most informative distances from a matrix of inter-dye distances as described above.

FRET on rails: Steps II + III. Next, we have to test the null hypothesis using this initial distance set. In real situations, we have to prepare a set of samples and perform a series of measurements of freely diffusing molecules. However, in this test study, we replace the experimental data acquisition by simulations for a set of 10 hypothetical FRET samples with the mixtures of appropriately labeled Atlastin-1 conformers.

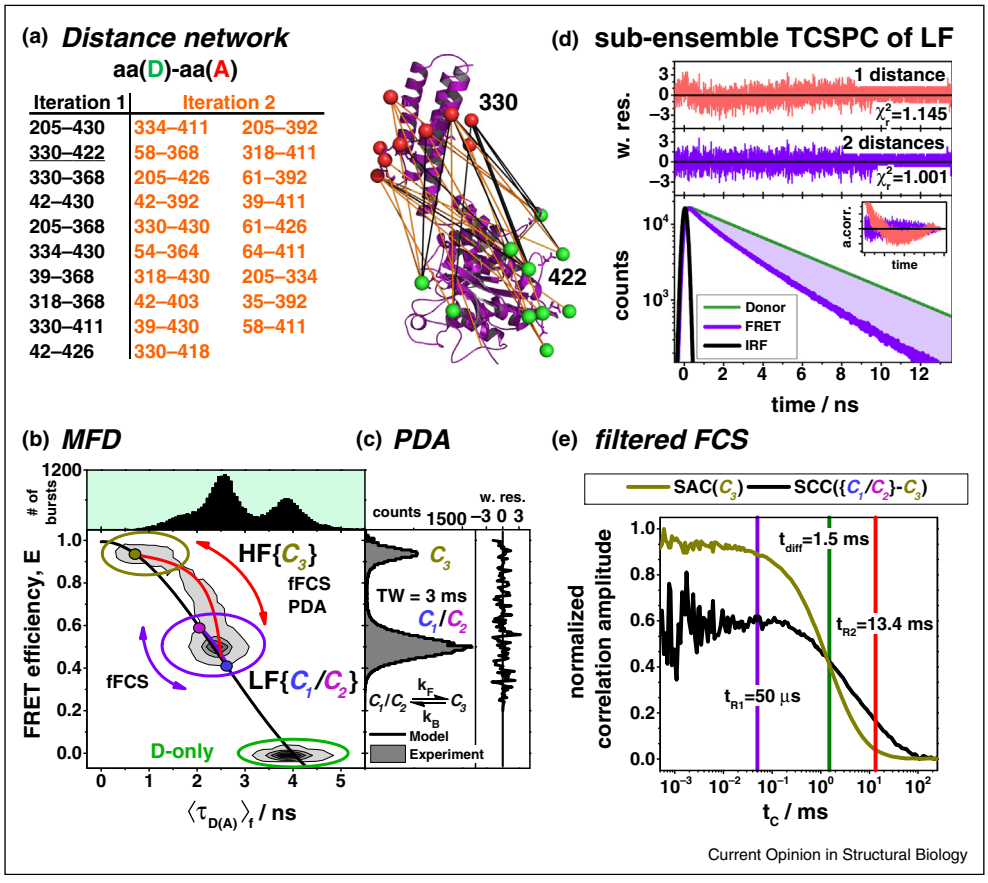
Figure 3



Complete workflow of FRET-restrained structural modeling. Workflow for the benchmark test using three distinct crystal structures of Atlastin-1. **Step I:** Starting from the crystal structure (PDB-ID: 4IDN), we generated a conformational ensemble by NMSim, which was subsequently clustered. From this, we derived the most informative donor–acceptor (DA) pairs, which can distinguish between these clusters. These DA pairs serve as one of the inputs for the next step. **Step II:** The previously determined DA pairs, crystal structures, and the kinetic scheme (Figure 2) were used to simulate smFRET experiments. In the simulations, the molecules randomly exchange their conformation, diffuse freely, and emit fluorescence if they diffuse through the confocal detection volume. The simulated photon trajectories were analyzed according to Box 1 to yield for each DA pair a conformer-specific distance and a species fraction $x^{(i)}$. This results in a distance set for each of the three conformers, which is used in the next step. **Step III:** Structural modeling was based on the distance sets determined in step II and the conformational ensembles generated in step I. Using fitted distance sets, we applied FRET-guided NMSim to expand the conformational ensemble of each conformer. Subsequently, we screened the resulting conformational ensemble with the FPS software to identify conformers that are in agreement with the simulated experiment. **Step IV:** A second round repeating steps II and III, which considers additional DA pairs, improved the discrimination power and increased the resolution. **Step V:** The final models were chosen by selecting those that are within the 68% confidence level.

In *in silico* MFD-experiments, we simulate: (1) diffusion of individual molecules; (2) dynamic exchange of the states; (3) the confocal excitation profile and observation volume; (4) the photon emission as Poissonian process, which follows the conformer-specific fluorescence decay $f_{D(A)}(t)$ (Eq. (2)) described by the excited state depopulation with $k_{D,0}$ and FRET-induced quenching with a distributed $k^{(i)}_{RET}$ due to flexibly linked dyes; and (5) experimental nuisances such as the instrumental response function (IRF), background fluorescence and experimental calibration

Figure 4



Analysis of simulated FRET data. Fluorescence analysis workflow. **(a)** The distance network used to simulate the fluorescence experiments was constructed in two iterations. After a first analysis round (iteration 1, black), the network was extended by additional distances (iteration 2, orange). The analysis workflow is shown for the highlighted DA pair. **(b)** Confocal sm experiments with pulsed excitation (73.56 MHz) and time-resolved multi-parameter fluorescence detection (i.e. spectral and polarization resolved detection) [60••] were simulated for every DA pair using a data generator as described in detail in [119], that generated photon streams of typical single-molecule experiments by Brownian dynamics (BD) simulations. The fluorescence decay of D and A in absence of FRET was modeled by single exponential decays using a fluorescence lifetime of D of $\tau_{D(0)}$ = 4 ns and A of τ_A = 1 ns, respectively. Their time-resolved anisotropies were modeled using mean rotational correlation times of 1 ns. The fluorescence signal was modeled to yield γ = 1 as calibration parameter. The DA distance distributions were simulated by ACV simulations, which resulted in an average DA distance half-width of 10.4 Å. The optical detection profile of the setup was simulated by a 3D-Gaussian, which decayed at a radius of ω_0 = 0.5 µm (*x, y* direction) and z_0 = 2.25 µm (*z* direction) to a value of $1/e^2$. To assure single-molecule conditions we simulate experiments with a mean number of 0.004 molecules in the focus which diffuse with a diffusion time t_{diff} = 1.5 ms as defined by FCS. First, single molecule bursts of the interconverting conformers (C_1, C_2, and C_3) were classified in 2D frequency histograms by their FRET-efficiency and intensity-weighted average donor fluorescence lifetime $\langle\tau_{D(A)}\rangle_f$. The number of molecules (bursts) is scaled from light grey (lowest) to black (highest). The 1D histogram of $\langle\tau_{D(A)}\rangle_f$ is given as a projection on the top. In the 2D-histogram, three peaks a high FRET (HF, yellow), low FRET (LF, violet) and a peak corresponding to FRET-inactive molecules (D-only, green) are visible. These peaks can be described by static and dynamic FRET lines. The static FRET line (black) describes the theoretical relationships between the FRET efficiency and $\langle\tau_{D(A)}\rangle_f$ for all molecules in the absence of protein dynamics. Deviation from the static FRET line towards longer fluorescence lifetimes indicates conformational dynamics. Dynamic FRET lines are defined by the limiting states of the dynamic processes. The FRET-efficiencies and $\langle\tau_{D(A)}\rangle_f$ of the conformers C_1, C_2, and C_3 are shown as yellow, magenta, and blue filled circles, respectively. Dynamic FRET lines of the C_1–C_2 transition and a C_3 – C_1/C_2 transition are shown in violet and red, respectively. C_1 and C_2 are in fast exchange (violet arrow) while C_3 exchanges slowly with C_2 (red arrow). Therefore, only an average of C_1 and C_2 LF{C_1/C_2} is resolved. **(c)** Dynamic PDA (time-window, TW = 3 ms) characterizes the slow C_3 – C_1/C_2 exchange by the analysis of FRET-efficiency histograms (data, grey) by a kinetic two-state model (fit, black line) and recovers fluorescence-averaged distances of C_3 (33.6 Å) and C_1/C_2 (52.5 Å). Weighted residuals are shown to the right. **(d)** Sub-ensemble donor fluorescence decay analysis of the LF population resolves C_1 and C_2 as individual components. The donor fluorescence decay in the absence of FRET serves as reference. The instrument response function (IRF) is shown as black line. The magenta region shows the number of photons of the donor quenched due to FRET. On top of the fluorescence decay, the weighted residuals (w.res.) of a one-component (1 distance, red) and a two-component model (2 distances, violet) are shown. As visualized by the auto-correlation of the weighted residuals, the one-component model is significantly worse and is therefore discarded (p value > 0.99) **(e)** Filtered fluorescence correlation spectroscopy (fFCS) computes the species-specific cross-correlation (SCC) between HF{C_3} and LF{C_1/C_2}. The species auto correlation (SAC) of C_3 recovers the diffusion time t_{diff}. The presence of two anticorrelation terms in the SCC indicates three kinetic states. The analysis of the SCC recovers characteristic relaxation times t_{R1} and t_{R2} of the C_1, C_2, and C_3 kinetics.

factors. Finally, to generate brightness Q and fluorescence decays $f_{D(A)}(t)$ for each state in the simulated single-molecule experiment, we used the crystal structures as input. This involved two steps. First, DA distance distributions $x(R_{DA})$ of the selected FRET pair were calculated using two inputs, the crystal structures of C_1, C_2, and C_3, respectively, and the *Free diffusion+Contact/Iso model* defining the corresponding accessible contact volumes. Second, the obtained respective $x(R_{DA})$ for C_1, C_2, and C_3 were used to calculate the FRET observables (fluorescence intensity decays $f(t)$ and transfer efficiencies E) according to Eqs. (1)–(3) (the assumed fluorescence parameters of the dyes are compiled in the caption of Figure 4). The transfer efficiencies E were used to calculate the brightness Q as described previously [60••]. All simulation parameters are compiled in the caption of Figure 4b.

The analysis of the fluorescence signals of the sm experiments simulated in step II of the workflow follows the procedures described in Box 1. First, the bright fluorescence bursts due to single-molecule events with durations of a few milliseconds were identified and selected in the time trace by separating them from the low background of 1–2 kHz. Second, for each single-molecule burst, all fluorescence parameters were determined as described in [60••]. To check for the presence of exchange dynamics, the FRET indicators, FRET efficiency E and the donor fluorescence lifetime $\tau_{D(A)}$ were plotted in two dimensional frequency MFD histograms (Figure 4b) to analyze the location of the bursts using the static (black) and dynamic (red) FRET lines [60••].

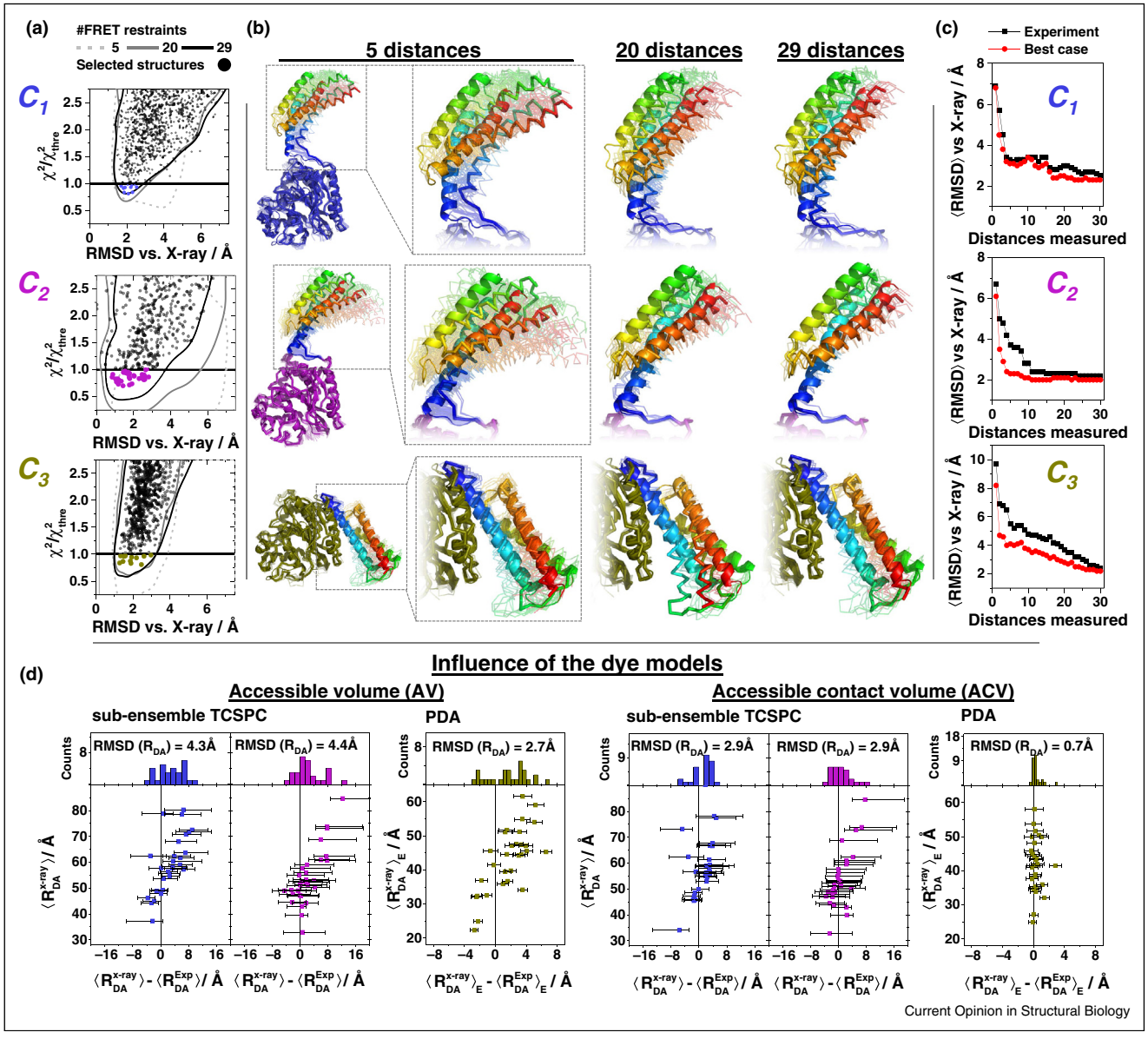
The number of conformers and their species fractions were identified as follows. A time window-based analysis by dynamic photon distribution analysis (dynPDA) [24•] resolved the FRET averaged distances $\langle R_{DA}\rangle_E$ of a small population (~25%) with high FRET efficiencies, referred to as HF population, which is in slow exchange with a second population with a lower FRET efficiency (~75%), referred to as LF population (Figure 4c). To test whether these populations are homogeneous, we performed a sub-ensemble TCSPC analysis of both populations. While the fluorescence decay of the HF population could be fitted with a single distance distribution (i.e. a single FRET species) (Eq. (2)), two distance distributions (i.e. two FRET species) were needed to describe the decay of the LF population (Figure 4d). The necessity of two FRET species in the LF population was judged by comparing the goodness of the fits for one and two distance distributions using weighted residuals (w.res.), the autocorrelation function (a.corr.) of the residuals, and χ_r^2. Overall, sub-ensemble TCSPC and PDA analysis allowed us to resolve three FRET species. PDA identifies two dynamically exchanging populations. One of these populations is resolved into two distinct populations by sub-ensemble TCSPC (seTCSPC). We formally assign the HF population to the conformer (C_3) and the LF population to a dynamic mixture of C_1 and C_2, which is separated from C_3 by a large kinetic barrier with a relaxation time of ~10 ms.

To resolve the heterogeneity of the averaged $\{C_1/C_2\}$ population, the fluorescence decays of the respective sub-ensembles for all 10 simulated single-pair FRET experiments were analyzed by a joint model, which assumed global species fractions of the $\{C_1/C_2\}$ population and Gaussian-distributed distances. Only this joint analysis is capable of recovering the experimental global species fractions $x_1 = 50\%$ and $x_2 = 25\%$, which agree well with input values used in the simulation (Figure 2). The presence of exchange kinetics is independently detected by filtered fluorescence correlation spectroscopy (fFCS) [23•,119] (Figure 4e) by computing the species cross-correlation function (SCC($\{C_1/C_2\} - C_3$)) that analyzes the exchange between the fast interchanging populations $\{C_1/C_2\}$ and C_3. The SCC exhibits two relaxation times, which is additional evidence for the existence of three conformational states. The fast relaxation time $\tau_{R1} = 50$ μs exactly corresponds to the one used in the simulations to describe the exchange between C_1 and C_2. Even though the lifetimes of the conformers C_1 and C_2 are short (100 and 50 μs, respectively), MFD, se-TCSPC and fFCS analysis unequivocally identified three FRET species together with their species fractions, respective distances and corresponding exchange kinetics. The errors of these distances $\Delta R^2_{\mathrm{DA}}(E)$ were determined by propagating the experimental noise to the model parameters. As this analysis revealed three states, we accept the null hypothesis and reject the initial assumption that the crystal structure describes the experimental sample.

Therefore, in step III of the workflow, we explore the information contained in the experiments by screening the entire ensemble using FPS [42••]. This screen selects an ensemble of iteration 1 for the structural models of C_1, C_2 and C_3 as *posterior* hypothesis.

FRET on rails: Step IV. To improve the precision of our *posterior* model, we performed the iteration step IV by repeating the DA pair selection procedure now using the reduced ensemble of iteration 1. Based on this analysis, we chose 19 additional informative DA pairs (Figure 4a). For these pairs, we again simulated smFRET experiments following step II as described above. After the analysis, our distance set for each conformer contained 29 distances. We then used these expanded distance sets of iterations 1 and 2 to model the respective structures applying FRET-guided NMSim simulations and subsequent screening by FPS (step III). We monitored the improvement of accuracy and precision for each conformer (Figure 5c). Since the RMSD value from the X-ray structure levels off at 25–29 distances for the best-case prediction and the recovered experimental accuracy for

Figure 5



Assessment of the hybrid-FRET structural models. Outcome of hybrid-FRET structural modeling of the three Atlastin-1 conformers. **(a)** The sum of squared deviations weighted by the estimated experimental error χ^2 relative to a threshold value $\chi^2_{threshold}$ estimated for the confidence level of 68% with respect to the best structure is plotted against C_α root-mean-square deviations (RMSD) from the corresponding crystal structure for 5 (dashed contour), 20 (grey contour), and 29 (black contour) measured distances. Black dots represent conformers corresponding to the set of 29 distances. Contours are drawn using a kernel density estimate. Colored dots represent selected conformations for the set of 29 distances. RMSD versus X-ray is calculated excluding flexible loops using the residues 35–99, 122–147, 157–189, 209–237, 257–277, 292–332, and 349–437. **(b)** Overlay of the crystal structures (cartoon representation), selected ensembles (transparent ribbon, 68% confidence), and best (χ^2_{min} structures (solid ribbon) for the sets of 5, 20, and 29 distances. **(c)** Improvement of the corresponding accuracy with respect to the number of distances measured. Accuracy is calculated as χ^2-weighted average C_α RMSD from the corresponding crystal structure. Black line represents the improvement using experimentally measured distances, the red line represents the best-case scenario where all the distances measured agree perfectly with the ones predicted for the crystal structure. **(d)** Comparison for the deviations between measured distances and distances predicted from the crystal structure by AV and ACV dye models.

all conformers, further measurements would not improve the accuracy of the obtained structures for Atlastin-1. However, in the case of remaining ambiguity for the solutions, additional iterations through steps IV and II (determining DA pairs, simulating and analyzing smFRET experiments), and step III (structural modeling and subsequent screening) could improve the models even further.

FRET on rails: Step V (Figure 3). We screened all generated structures using FPS, which yielded FRET-selected ensembles of iteration 2 for each conformer C_1, C_2, and C_3, using the confidence level threshold of 68%.

Accuracy and precision of FRET-derived structural models for the test case Atlastin-1. The accuracy and precision of the hybrid-FRET structural models obtained by the iterative workflow (Figure 3) is summarized in Figure 5a. We judge the accuracy of the method by plotting the conformer specific *f*-value ($\chi^2/\chi^2_{threshold}$ ratio) versus the C_α RMSD of the modeled structures from the corresponding target crystal structure. The $\chi^2_{threshold}$ value is estimated for the confidence level of 68%. For all states, the accuracy of the selected structures at a confidence level of 68% using 29 distances ranges between 1 and 3 Å. As we are interested in large-scale rearrangements of the super-tertiary structures, we exclude minor rearrangements of the sub-domains for the estimation of our accuracy (Figure 5a). To minimize the effect of non-uniform sampling, we use cluster representatives with a RMSD threshold of 1 Å (complete linkage) [125]. Cluster representatives below a confidence level of 68% for 29 distances are shown as colored dots. An overlay of the selected cluster representatives and the best model structure (highlighted by solid ribbons) visualizes the precision, which is given by the structural diversity within the selection (Figure 5b). To emphasize the differences, we aligned the representatives to the rigid G domain of the conformers. Within our precision (confidence level 68%) we can distinguish all three conformers, even C_1 and C_2 (i.e. the conformational ensembles do not overlap).

To capture the general dependence of the accuracy on the number of measured distances, we calculate an average RMSD that is weighted by the probability of agreement with the experimental distances as judged by χ^2 of the corresponding structural model. In Figure 5c, this average RMSD is used as a measure of accuracy, and shown as a function of the number of measured distances sorted by their information content. Interestingly, this RMSD decays fast with the number of measured distances and levels off in the studied system at an RMSD value of 2 Å. The precision and accuracy of the structural model is expected to depend on three major factors: (1) the noise (uncertainty) of the experimental data, (2) its sparseness and (3) outliers. Indeed, by increasing the number of distances, the effect of noise and data sparseness decrease for all conformers as highlighted in Figure 5c. The precision of the FRET- ensemble (4 to 5 Å) selected by 5 distances improves to 2 Å when 29 distances are measured. The influence of experimental noise and sparseness is seen by comparing the RMSD dependence on the number of measured distances in Figure 5c for the best case with accurate distances (black) and the real case with experimentally determined DA distances (red). Notably, already small optimally informative set of DA pairs may provide a high accuracy (Figure 5c, red), because in the structural models obtained by NMSim (or any other computer simulation procedure) C_α atoms are interdependent. Therefore, only 29 FRET measurements with realistic errors (Figure 5c, black) localize the C_α positions to an RMSD of ~2 Å; the accuracy in our test does not improve significantly even if more than 15 measurements are performed. Notably, even for a large optimal experimental dataset, a residual RMSD error will be observed, as intrinsic degrees of backbone freedom within the structural model (relatively small fluctuations of C_α atoms) are not sensed by FRET. More FRET measurements would not alleviate this problem; that is, this test study reached the maximum possible precision. Finally, the third effect considering the possibility that the measurements might contain a low number of outliers, may be minimized by increasing the number of measurements, reducing the risk of systematic errors.

In crystallography, short range (up to 3 Å) stereo-chemical criteria, that is, deviations of atomic distances, angles, and dihedrals from their mean values, are used as quality criteria for obtained structures. Using such criteria (MolProbity [126]), the FRET-selected models generated by NMSim generally score better than corresponding X-ray structures. Therefore, we assess the overall quality of the structural model by experimental quality measures. As such, deviations between experimentally determined distances ($R_{DA}{}^{Exp}$) and corresponding model distances ($R_{DA}{}^{Model}$) are utilized as a quality indicator. Knowing the target structures in our test case, we used this indicator to assess the accuracy of different dye models (Figure 5d). We simulated our experimental information using the ACV dye model (Figure 3, step II) and, accordingly, the ACV model clearly outperformed the alternative AV model with an RMSD(R_{DA}) that is lower by ~1.5–2 Å when comparing $R_{DA}{}^{Exp}$ with $R_{DA}{}^{Model} = R_{DA}{}^{X\text{-}ray}$ (Figure 5d). Thus, parameters of coarse grained dye models such as the preferential surface residence of the dye (ACV model) may be refined experimentally by calibration studies.

Outlook

Hybrid/integrative modeling: Combining FRET spectroscopy with other fluorescence and biophysical methods. To interrogate different molecular length scales, to address distinct sample properties (e.g. backbone, side chain, shape) and to cover different time scales, hybrid modeling can utilize the complementary information from a multitude of experimental techniques such as small-angle neutron scattering (SANS), small-angle X-ray scattering (SAXS), NMR-spectroscopy, EPR-spectroscopy and FRET-spectroscopy in one study [53•,109,127–129]. Integrative modeling is a thriving direction in methodological development directed towards

a multi-dimensional structural and dynamical description of complex biomacromolecular systems [38]. However, the error assessment of the data can often be treated only semi-quantitatively, because it is difficult, perhaps impossible, to determine specific errors of each method and their relative weights in a joint analysis.

In single photon counting it is straightforward to estimate error of FRET and other fluorescence measurements due to the Poisson statistics for photon noise. Besides FRET, other fluorescence methods are usually convenient and can also provide restraints for structural modeling as described in the following examples. Photo-induced electron-transfer (PET) probes the close proximity of the fluorophore to certain electron-rich quenching amino acid residues [130,131] and hence senses conformational dynamics [132]. Fluorescence polarization senses dye mobility that can be influenced by local and global rotation [133] and by complex formation where the label can be trapped [134]. The fluorescence spectra of polarity-sensitive fluorophores, such as intrinsic fluorescent amino acid tryptophan, sense their exposure to water [135]. FRET between the fluorophores and PET with electron rich amino acid residues are often competing processes in quenching the donor fluorescence, so that this scenario was either judged as valuable information [136] or as experimental nuisance [17]. Both methods can be combined to obtain distance (PET: short range, FCS: long range) [137] as well as kinetic information [138]. Information similar to that from PET can be obtained from protein-induced fluorescence enhancement (PIFE) experiments with cyanine dyes [98,99•]. Finally, SAXS and polarization resolved FCS measurements [139,140] contain information on the anisotropic rotational diffusion of macromolecules, which reports on the size and the global shape of the macromolecules.

Maximizing data utilization by structure-based forward modeling. In forward modeling [111••], the chosen model parameters are verified by a direct comparison to the experimental data. In fluorescence experiments with single photon counting detection, forward modeling of the registered fluorescence decays and FRET efficiency histograms has a long tradition (e.g. convolution of instrument response function instead of deconvolution), because in this way the Poisson statistics of the counted photons is preserved and the experimental noise (shot-noise) is exactly known. However, since structural modeling is typically performed using distance restraints (in backward modeling; i.e. the fit is decoupled from structural model), experimental data is preprocessed by applying fluorescence models (for instance Eqs. (1)–(3)). This can lead to overestimation of the errors, if the model parameters (distances) are correlated. However, to obtain independent errors for the distance restraints, the multi-dimensional uncertainties of the distances are projected to one dimension (i.e. marginalized). In this process, the valuable information on correlations is lost. Similar to peak assignment problems in NMR spectroscopy [111••], correlations can be preserved by forward modeling which utilizes a structural model as a global model for the experimental data, maximizing the recovered information and minimizing ambiguities. Dye models must be applied to directly link the structural model to the fluorescence decay. For this purpose AV, ACV or more complex spatial dye distributions can be used. In this forward fitting approach the structural model is varied for optimal agreement with experiment.

Hybrid-FRET studies in live cells. The FRET technique offers unique advantages for characterizing large biomolecules with high specificity and sensitivity in living cells or *in vitro*. Time-resolved FRET studies can deal with heterogeneous samples and make use of the single-molecule advantage so that a quantitative FRET analysis can be combined with normal [141,142] and super-resolution [143] imaging to gain insight into the biological function of biomolecules in their cellular context (e.g. localizing complexes inside of cells [141], quantifying binding to interaction partners [104•], and validating structural models [103,144,145]). For example, we could show by detailed hybrid-FRET studies of murine immune defense Guanylate binding proteins with GFP and mCherry fusions that they undergo reversible structural transitions between monomeric, dimeric and oligomeric states as revealed by a quantitative analysis of homo- and hetero-FRET [104•]. Notably, the degree of oligomerization is specific to the localization in live cells. Detailed hybrid-FRET studies allowed us to characterize the structures of the dimer and oligomer.

Comparative studies of cell-like environment and dilute solutions were used to extrapolate excluded volume effects from *in vitro* experiments to live-cells [146•], and to identify biomolecular stabilization mechanism caused by macromolecular crowding [147]. Microinjection allows one to control the concentration of biomolecules labeled by photostable flurophores and therefore enables the study of crowding effects in live cells [28]. In this way, single-molecule conditions can be achieved by tracking microinjected *in vitro* purified FRET-labeled proteins to study complex formation and conformational changes of individual proteins [148]. This approach, combined with fast confocal detection, was used to probe protein dynamics from millisecond down to the nanosecond regime [149].

In conclusion, the presented hybrid-FRET methods actually allow realizing an integrated molecular fluorescence microscope combining optical and computational microscopy [58] at a huge spatial and temporal range to display suitably labeled biomolecular systems at unprecedented resolution by atomistic structural models.

Further information

For the presented test case study on Atlastin-1, all structural models presented in Figure 5, additional detailed information on the applied procedures and the simulated sm FRET data are deposited in the Model Archive with the DOI: 10.5452/ma-a2hbq.

Conflict of interest

Nothing declared.

Acknowledgments

We thank Stanislav Kalinin, Jens Michaelis and Oleg Opanasyuk for fruitful discussions on integrative FRET-modeling. We are grateful to Suren Felekyan for supporting the fluorescence simulations and Katherina Hemmen for her support in preparing figures. This work was supported by the European Research Council through the Advanced Grant 2014 hybridFRET (671208). The authors acknowledge the stimulating networking support by the COST Action CM1306 – Understanding Movement and Mechanism in Molecular Machines.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. •• Förster T: **Zwischenmolekulare Energiewanderung und Fluoreszenz**. *Ann Phys* 1948, **437**:55-75.
In this seminal paper Förster derived the correct distance and orientation dependence of the FRET-rate constant for the first time. Additionally, energy migration within regular crystallographic lattices (the second part of this paper) is studied, which is still of interest for crystal like structures for instance, formed by membrane associated proteins such as GPCRs.

2. •• Förster T: **Experimentelle und theoretische Untersuchung des zwischenmolekularen Übergangs von Elektronenanregungsenergie**. *Z Naturforsch A* 1949, **4**:321-327.
In this second seminal paper Förster derived experimental observables starting from simple schemes of rate constants. Equations for fluorescence decays of donors randomly surrounded by acceptor fluorophores are presented. In this sense this fundamental paper is the basis for understanding most FRET-experiments with flexibly coupled dyes.

3. van der Meer BW, Cooker G, Chen SY: *Resonance Energy Transfer: Theory and Data*. VCH Publishers; 1994.

4. Weiss S: **Fluorescence spectroscopy of single biomolecules**. *Science* 1999, **283**:1676-1683.

5. Stryer L: **Fluorescence energy transfer as a spectroscopic ruler**. *Annu Rev Biochem* 1978, **47**:819-846.

6. Ha T, Enderle T, Ogletree DF, Chemla DS, Selvin PR, Weiss S: **Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor**. *Proc Natl Acad Sci U S A* 1996, **93**:6264-6268.

7. Pirchi M, Ziv G, Riven I, Cohen SS, Zohar N, Barak Y, Haran G: **Single-molecule fluorescence spectroscopy maps the folding landscape of a large protein**. *Nat Commun* 2011, **2**:493-499.

8. Chung HS, Eaton WA: **Single-molecule fluorescence probes dynamics of barrier crossing**. *Nature* 2013, **502**:685-688.

9. Zhuang XW, Kim H, Pereira MJB, Babcock HP, Walter NG, Chu S: **Correlating structural dynamics and function in single ribozyme molecules**. *Science* 2002, **296**:1473-1476.

10. Kim HD, Nienhaus GU, Ha T, Orr JW, Williamson JR, Chu S: **Mg^{2+}-dependent conformational change of RNA studied by fluorescence correlation and FRET on immobilized single molecules**. *Proc Natl Acad Sci U S A* 2002, **99**:4284-4289.

11. Al-Hashimi HM, Walter NG: **RNA dynamics: it is about time**. *Curr Opin Struct Biol* 2008, **18**:321-329.

12. Roy R, Kozlov AG, Lohman TM, Ha T: **SSB protein diffusion on single-stranded DNA stimulates RecA filament formation**. *Nature* 2009, **461**:1092-1097.

13. Olofsson L, Felekyan S, Doumazane E, Scholler P, Fabre L, Zwier JM, Rondard P, Seidel CAM, Pin JP, Margeat E: **Fine tuning of sub-millisecond conformational dynamics controls metabotropic glutamate receptors agonist efficacy**. *Nat Commun* 2014, **5**:e5206.

14. Robb NC, Te Velthuis AJ, Wieneke R, Tampe R, Cordes T, Fodor E, Kapanidis AN: **Single-molecule FRET reveals the pre-initiation and initiation conformations of influenza virus promoter RNA**. *Nucleic Acids Res* 2016, **44**.

15. Diez M, Zimmermann B, Börsch M, König M, Schweinberger E, Steigmiller S, Reuter R, Felekyan S, Kudryavtsev V, Seidel CAM, Gräber P: **Proton-powered subunit rotation in single membrane-bound F_0F_1-ATP synthase**. *Nat Struct Mol Biol* 2004, **11**:135-141.

16. Lu HP: **Sizing up single-molecule enzymatic conformational dynamics**. *Chem Soc Rev* 2014, **43**:1118-1143.

17. Winkler JR: **FRETting over the spectroscopic ruler**. *Science* 2013, **339**:1530-1531.

18. Nettels D, Gopich IV, Hoffmann A, Schuler B: **Ultrafast dynamics of protein collapse from single-molecule photon statistics**. *Proc Natl Acad Sci U S A* 2007, **104**:2655-2660.

19. Soranno A, Buchli B, Nettels D, Cheng RR, Muller-Spath S, Pfeil SH, Hoffmann A, Lipman EA, Makarov DE, Schuler B: **Quantifying internal friction in unfolded and intrinsically disordered proteins with single-molecule spectroscopy**. *Proc Natl Acad Sci U S A* 2012, **109**:17800-17806.

20. • Gopich I, Szabo A: **Theory of photon statistics in single-molecule Förster resonance energy transfer**. *J Chem Phys* 2005, **122**:014707.
Gopich and Szabo describe FRET-efficiency histograms (FEHs) of single-molecule FRET experiments by the photon statistics. As they consider the effect of conformational dynamics on FEH, they set the basis for quantitative time-window analysis of FEHs of dynamically exchanging species.

21. • Chung HS, Gopich IV: **Fast single-molecule FRET spectroscopy: theory and experiment**. *Phys Chem Chem Phys* 2014, **16**:18644-18657.
In this perspective, the authors demonstrate the improvement of the dynamic range and accuracy of sm fluorescence spectroscopy at a given photon count rate by considering each and every photon and introducing a maximum likelihood (ML) method. By analyzing photon trajectories with recorded photon colors and inter-photon times of smFRET studies (protein folding), their ML method recovers parameters of a kinetic exchange model.

22. Chung HS, McHale K, Louis JM, Eaton WA: **Single-molecule fluorescence experiments determine protein folding transition path times**. *Science* 2012, **335**:981-984.

23. • Felekyan S, Sanabria H, Kalinin S, Kühnemuth R, Seidel CAM: **Analyzing Förster Resonance Energy Transfer with fluctuation algorithms**. *Methods Enzymol* 2013, **519**:39-85.
This review describes the theory and application of fluorescence correlation spectroscopy (FCS) with different analysis procedures (classical FRET-FCS and FRET-specific species correlation by filtered FCS) to analyze FRET fluctuations in order to distinguish between different species. Filtered FCS is especially useful for systems with complex dynamics.

24. • Kalinin S, Valeri A, Antonik M, Felekyan S, Seidel CAM: **Detection of structural dynamics by FRET: a photon distribution and fluorescence lifetime analysis of systems with multiple states**. *J Phys Chem B* 2010, **114**:7983-7995.
The authors present a method to analyze FRET efficiency histograms, taking into account shot noise and background contributions. The method can be applied to resolve the dynamically exchanging states and their kinetics. Furthermore, the authors present analysis tools to identify conformational dynamics by comparing the intensity-based and time-resolved FRET indicators F_D, F_A and τ_D (Eq. (1)) by calculating FRET-lines for the static and dynamic case, respectively (see Box 1). The provided theory is valuable to investigate dynamic systems.

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

• The authors present a robust and unbiased algorithm to recover kinetic schemes from time-binned sm FRET trajectories. A number of different states and state-to-state transition probabilities can be recovered, together with the corresponding FRET efficiencies for transitions slower than the bin width. This is an effective analysis method for widely applied TIRF measurements.

26. König SLB, Hadzic M, Fiorini E, Börner R, Kowerko D, Blanckenhorn WU, Sigel RKO: **BOBA FRET: Bootstrap-based analysis of single-molecule FRET data**. *PLoS ONE* 2013, **8**:e84157.

27. Keller BG, Kobitski A, Jaschke A, Nienhaus GU, Noe F: **Complex RNA folding kinetics revealed by single-molecule FRET and Hidden Markov Models**. *J Am Chem Soc* 2014, **136**:4534-4543.

•• A rigorous hidden Markov Model (HMM) analysis procedure is shown to recover states based on kinetics rather than just FRET efficiencies. For Diels–Alderase ribozyme, a seven state model with fastest interconversion timescale of 3 ms is recovered. The approach enables reconstruction of complex kinetic schemes at high temporal resolution.

28. Gao M, Gnutt D, Orban A, Appel B, Righetti F, Winter R, Narberhaus F, Muller S, Ebbinghaus S: **RNA hairpin folding in the crowded cell**. *Angew Chem Int Ed* 2016, **55**:3224-3228.

29. Schuler B, Hofmann H: **Single-molecule spectroscopy of protein folding dynamics-expanding scope and timescales**. *Curr Opin Struct Biol* 2013, **23**:36-47.

• The authors review a broad spectrum of sm fluorescence methods based on sm FRET and photo-induced electron transfer (PET) that have been developed for probing protein structure and dynamics, especially in structurally heterogeneous systems. They give an excellent overview on the applications of the kinetic experiments by discussing many processes coupled with protein folding such as misfolding, binding, transition path times and interactions with chaperones.

30. Schneidman-Duhovny D, Pellarin R, Sali A: **Uncertainty in integrative structural modeling**. *Curr Opin Struct Biol* 2014, **28**:96-104.

• This review defines the need for standards and tools to assess the input data and resulting models in integrative structural modeling. To tackle problems in integrative structural modeling they define standards for a common terminology. Furthermore, the authors point out possible sources of model uncertainty.

31. Stryer L, Haugland RP: **Energy transfer: a spectroscopic ruler**. *Proc Natl Acad Sci U S A* 1967, **58**:719-726.

32. Murchie AI, Clegg RM, von Kitzing E, Duckett DR, Diekmann S, Lilley DMJ: **Fluorescence energy transfer shows that the four-way DNA junction is a right-handed cross of antiparallel molecules**. *Nature* 1989, **341**:763-766.

33. Wozniak AK, Schröder GF, Grubmüller H, Seidel CAM, Oesterhelt F: **Single-molecule FRET measures bends and kinks in DNA**. *Proc Natl Acad Sci U S A* 2008, **105**:18337-18342.

34. Preus S, Wilhelmsson LM: **Advances in quantitative FRET-based methods for studying nucleic acids**. *ChemBioChem* 2012, **13**:1990-2001.

35. Preus S, Kilsa K, Miannay FA, Albinsson B, Wilhelmsson LM: **FRETmatrix: a general methodology for the simulation and analysis of FRET in nucleic acids**. *Nucleic Acids Res* 2013, **41**:e18.

• The presented methodology addresses the problem of deriving structures using fluorophores with fixed positions and fixed dipoles using information from fluorescence decays combined with fits directed by the structural model. The resulting toolkit FRETmatrix (http://www.fluortools.com/software/fretmatrix) is applied to study nucleic acid structures in 3D.

36. Mekler V, Kortkhonjia E, Mukhopadhyay J, Knight J, Revyakin A, Kapanidis AN, Niu W, Ebright YW, Levy R, Ebright RH: **Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex**. *Cell* 2002, **108**:599-614.

37. Nagy J, Grohmann D, Cheung ACM, Schulz S, Smollett K, Werner F, Michaelis J: **Complete architecture of the archaeal RNA polymerase open complex from single-molecule FRET and NPS**. *Nat Commun* 2015, **6**:e6161.

38. Sali A, Berman HM, Schwede T, Trewhella J, Kleywegt G, Burley SK, Markley J, Nakamura H, Adams P, Bonvin A, Chiu W *et al.*: **Outcome of the first wwPDB Hybrid/Integrative methods task force workshop**. *Structure* 2015, **23**:1156-1167.

39. Schneider M, Belson A, Rappsilber J, Brock O: **Blind testing of cross-linking/mass spectrometry hybrid methods in CASP11**. *Proteins* 2016, **84**:152-163.

40. Tompa P: **On the supertertiary structure of proteins**. *Nat Chem Biol* 2012, **8**:597-600.

41. Knight JL, Mekler V, Mukhopadhyay J, Ebright RH, Levy RM: **Distance-restrained docking of rifampicin and rifamycin SV to RNA polymerase using systematic FRET measurements: developing benchmarks of model quality and reliability**. *Biophys J* 2005, **88**:925-938.

• This benchmark study demonstrates positioning accuracies between 7 and 10 Å for small molecules relative to the crystallographically defined binding sites without explicitly considering dye distributions.

42. Kalinin S, Peulen T, Sindbert S, Rothwell PJ, Berger S, Restle T, Goody RS, Gohlke H, Seidel CAM: **A toolkit and benchmark study for FRET-restrained high-precision structural modeling**. *Nat Meth* 2012, **9**:1218-1227.

•• The authors present a comprehensive toolkit (FRET Positioning and Screening, FPS, http://www.mpc.hhu.de/software/fps.html) for quantitative FRET-restrained modeling for applications in structural biology, which takes the spatial distributions of flexible linked dyes into account. In a benchmark study (docking a DNA primer-template to HIV-1 reverse transcriptase) the RMSD of the structural model to the known X-ray structure was only 0.5 Å. The FPS toolkit has two main functionalities: (1) Planning FRET experiments by calculating fluorescence observables for the given structural models; (2) FRET-based screening and multibody docking determine quantitative structural models.

43. Smock RG, Gierasch LM: **Sending signals dynamically**. *Science* 2009, **324**:198-203.

44. Henzler-Wildman K, Kern D: **Dynamic personalities of proteins**. *Nature* 2007, **450**:964-972.

45. Neudecker P, Robustelli P, Cavalli A, Walsh P, Lundstrom P, Zarrine-Afsar A, Sharpe S, Vendruscolo M, Kay LE: **Structure of an intermediate state in protein folding and aggregation**. *Science* 2012, **336**:362-366.

46. Banerjee PR, Deniz AA: **Shedding light on protein folding landscapes by single-molecule fluorescence**. *Chem Soc Rev* 2014, **43**:1172-1188.

47. Ferreon ACM, Ferreon JC, Wright PE, Deniz AA: **Modulation of allostery by protein intrinsic disorder**. *Nature* 2013, **498**:390-394.

48. Margittai M, Widengren J, Schweinberger E, Schröder GF, Felekyan S, Haustein E, König M, Fasshauer D, Grubmüller H, Jahn R, Seidel CAM: **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* 2003, **100**:15516-15521.

49. McCann JJ, Zheng LQ, Chiantia S, Bowen ME: **Domain orientation in the N-terminal PDZ tandem from PSD-95 is maintained in the full-length protein**. *Structure* 2011, **19**:810-820.

50. McCann JJ, Zheng L, Rohrbeck D, Felekyan S, Kühnemuth R, Sutton RB, Seidel CAM, Bowen ME: **Supertertiary structure of the synaptic MAGuK scaffold proteins is conserved**. *Proc Natl Acad Sci U S A* 2012, **109**:15775-15780.

51. Merchant KA, Best RB, Louis JM, Gopich IV, Eaton WA: **Characterizing the unfolded states of proteins using single-molecule FRET spectroscopy and molecular simulations**. *Proc Natl Acad Sci U S A* 2007, **104**:1528-1533.

52. Kellner R, Hofmann H, Barducci A, Wunderlich B, Nettels D, Schuler B: **Single-molecule spectroscopy reveals chaperone-mediated expansion of substrate protein**. *Proc Natl Acad Sci U S A* 2014, **111**:13355-13360.

53. Aznauryan M, Delgado L, Soranno A, Nettels D, Huang JR, Labhardt AM, Grzesiek S, Schuler B: **Comprehensive structural and dynamical view of an unfolded protein from the**

combination of single-molecule FRET, NMR, and SAXS. *Proc Natl Acad Sci U S A* 2016, **113**:E5389-E5398.
In this integrative structural study, the authors combine three of the most powerful biophysical methods (NMR spectroscopy, small angle X-ray scattering and smFRET measurements) available to obtain a comprehensive view on the structure and dynamics of the unfolded protein, ubiquitin, that would not be available from any of the individual methods.

54. Choi UB, Zhao ML, Zhang YX, Lai Y, Brunger AT: **Complexin induces a conformational change at the membrane-proximal C-terminal end of the SNARE complex**. *Elife* 2016, **5**:e16886.

55. Munro JB, Gorman J, Ma XC, Zhou Z, Arthos J, Burton DR, Koff WC, Courter JR, Smith AB, Kwong PD, Blanchard SC *et al.*: **Conformational dynamics of single HIV-1 envelope trimers on the surface of native virions**. *Science* 2014, **346**:759-763.

56. Kay LE: **New views of functionally dynamic proteins by solution NMR spectroscopy**. *J Mol Biol* 2016, **428**:323-331.

57. Schwede T, Sali A, Honig B, Levitt M, Berman HM, Jones D, Brenner SE, Burley SK, Das R, Dokholyan NV, Dunbrack RL Jr *et al.*: **Outcome of a workshop on applications of protein models in biomedical research**. *Structure* 2009, **17**:151-159.

58. Lee EH, Hsin J, Sotomayor M, Comellas G, Schulten K: **Discovery through the computational microscope**. *Structure* 2009, **17**:1295-1306.

59. •• Clegg RM: **Fluorescence resonance energy transfer and nucleic acids**. *Methods Enzymol* 1992, **211**:353-388.
This review provides an introduction to FRET and presents elegant and simple methods to determine absolute FRET efficiencies by steady state fluorometers. This allows one to obtain structural insights on the ensemble level using standard instrumentation.

60. •• Sisamakis E, Valeri A, Kalinin S, Rothwell PJ, Seidel CAM: **Accurate single-molecule FRET studies using multiparameter fluorescence detection**. *Methods Enzymol* 2010, **475**:455-514.
This review gives an introduction to single-molecule FRET experiments and analyses. It focuses on confocal detection of freely diffusing molecules by multi-parameter fluorescence detection. The authors describe qualitative and *quantitative* analysis methods for single-molecule FRET measurements and their prerequisites. Quantitative analysis of FRET measurements is the basis for FRET-based structural modeling.

61. • Sindbert S, Kalinin S, Hien N, Kienzler A, Clima L, Bannwarth W, Appel B, Muller S, Seidel CAM: **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.
This study relates dye linker length flexibilities to experimental FRET observables by measuring the fluorescence lifetimes and anisotropies of donor (Alexa 488) and acceptor (Cy5) dyes, respectively, tethered to dsRNA and dsDNA. For dyes with short linkers and unknown local environments, positional-related and orientational-related uncertainties outweigh better defined dye positions. For FRET-based structure determination of very high accuracy using dyes with long linkers, a procedure (accessible volume simulations) to model the spatial dye distributions is introduced.

62. Widengren J, Schweinberger E, Berger S, Seidel CAM: **Two new concepts to measure fluorescence resonance energy transfer via fluorescence correlation spectroscopy: theory and experimental realizations**. *J Phys Chem A* 2001, **105**:6851-6866.

63. Rothwell PJ, Berger S, Kensch O, Felekyan S, Antonik M, Wöhrl BM, Restle T, Goody RS, Seidel CAM: **Multi-parameter single-molecule fluorescence spectroscopy reveals heterogeneity of HIV-1 Reverse Transcriptase:primer/template complexes**. *Proc Natl Acad Sci U S A* 2003, **100**:1655-1660.

64. • Kudryavtsev V, Sikor M, Kalinin S, Mokranjac D, Seidel CAM, Lamb DC: **Combining MFD and PIE for accurate single-pair Förster Resonance Energy Transfer measurements**. *ChemPhysChem* 2012, **13**:1060-1078.
A combination of multiparameter fluorescence detection (MFD), pulsed interleaved excitation (PIE) and smFRET experiments on freely diffusing molecules is presented. MFD monitors a variety of fluorescence parameters simultaneously while PIE directly excites the acceptor to probe its presence and photoactivity. A single MFD-PIE measurement provides all necessary calibration factors for accurate analysis of smFRET histograms and excludes artifacts due to acceptor blinking or photobleaching. The performance of MFD-PIE measurements is demonstrated studying the interdomain separation in DnaK, the major bacterial heat shock protein 70 (Hsp70).

65. Vöpel T, Hengstenberg CS, Peulen TO, Ajaj Y, Seidel CAM, Herrmann C, Klare JP: **Triphosphate induced dimerization of human guanylate binding protein 1 involves association of the C-terminal helices: a joint double electron–electron resonance and FRET study**. *Biochemistry* 2014, **53**:4590-4600.

66. • Roy R, Hohng S, Ha T: **A practical guide to single-molecule FRET**. *Nat Meth* 2008, **5**:507-516.
The majority of the single-molecule FRET community performs total internal reflection (TIRF) experiments. This guide gives an introduction to FRET, TIRF, and discusses potential artifacts relevant for quantitative FRET-based structural modeling.

67. Rigler R, Mets Ü, Widengren J, Kask P: **Fluorescence correlation spectroscopy with high count rate and low background: analysis of translational diffusion**. *Eur Biophys J* 1993, **22**:169-175.

68. Farooq S, Hohlbein J: **Camera-based single-molecule FRET detection with improved time resolution**. *Phys Chem Chem Phys* 2015, **17**:27862-27872.

69. Lee NK, Kapanidis AN, Wang Y, Michalet X, Mukhopadhyay J, Ebright RH, Weiss S: **Accurate FRET measurements within single diffusing biomolecules using alternating-laser excitation**. *Biophys J* 2005, **88**:2939-2953.

70. Hohlbein J, Craggs TD, Cordes T: **Alternating-laser excitation: single-molecule FRET and beyond**. *Chem Soc Rev* 2014, **43**:1156-1171.

71. Elf J, Li GW, Xie XS: **Probing transcription factor dynamics at the single-molecule level in a living cell**. *Science* 2007, **316**:1191-1194.

72. Flors C, Hotta J, Uji IH, Dedecker P, Ando R, Mizuno H, Miyawaki A, Hofkens J: **A stroboscopic approach for fast photoactivation-localization microscopy with Dronpa mutants**. *J Am Chem Soc* 2007, **129**:13970-13977.

73. Felekyan S, Kühnemuth R, Kudryavtsev V, Sandhagen C, Becker W, Seidel CAM: **Full correlation from picoseconds to seconds by time-resolved and time-correlated single photon detection**. *Rev Sci Instrum* 2005, **76**:083104.

74. Boukobza E, Sonnenfeld A, Haran G: **Immobilization in surface-tethered lipid vesicles as a new tool for single biomolecule spectroscopy**. *J Phys Chem B* 2001, **105**:12165-12170.

75. • Müller BK, Zaychikov E, Bräuchle C, Lamb DC: **Pulsed interleaved excitation**. *Biophys J* 2005, **89**:3508-3522.
The authors present the method and analysis procedures of pulsed interleaved excitation (PIE) for the first time. PIE measurements significantly extend the capabilities of multiple-color fluorescence imaging, fluorescence cross-correlation spectroscopy (FCCS), and smFRET measurements.

76. Kühnemuth R, Seidel CAM: **Principles of single molecule multiparameter fluorescence spectroscopy**. *Single Mol* 2001, **2**:251-254.

77. Widengren J, Kudryavtsev V, Antonik M, Berger S, Gerken M, Seidel CAM: **Single-molecule detection and identification of multiple species by multiparameter fluorescence detection**. *Anal Chem* 2006, **78**:2039-2050.

78. McCann JJ, Choi UB, Zheng L, Weninger K, Bowen ME: **Optimizing methods to recover absolute FRET efficiency from immobilized single molecules**. *Biophys J* 2010, **99**:961-970.

79. • Hildebrandt LL, Preus S, Birkedal V: **Quantitative single molecule FRET efficiencies using TIRF microscopy**. *Faraday Discussions* 2015, **184**:131-142.
The authors introduced new software (http://www.isms.au.dk) to determine quantitative FRET efficiencies directly from individual fluorescence time traces of surface immobilized DNA molecules without the need for external calibrants.

80. Preus S, Hildebrandt LL, Birkedal V: **Optimal background estimators in single-molecule FRET microscopy**. *Biophys J* 2016, **111**:1278-1286.

81. Jeschke G: **DEER distance measurements on proteins**. In *Annual Review of Physical Chemistry*, Vol. 63. Edited by Johnson MA, Martinez TJ. 2012:419-446.

82. Columbus L, Hubbell WL: **A new spin on protein dynamics**. *Trends Biochem Sci* 2002, **27**:288-295.

83. Cai Q, Kusnetzow AK, Hideg K, Price EA, Haworth IS, Qin PZ: **Nanometer distance measurements in RNA using site-directed spin labeling**. *Biophys J* 2007, **93**:2110-2117.

84. Hagelueken G, Abdullin D, Schiemann O: **mtsslSuite: Probing biomolecular conformation by spin-labeling studies**. In *Electron Paramagnetic Resonance Investigations of Biological Systems by Using Spin Labels, Spin Probes, and Intrinsic Metal Ions, Pt A. 563.* Edited by Qin PZ, Warncke K. 2015:595-622.

85. •• Muschielok A, Andrecka J, Jawhari A, Bruckner F, Cramer P, Michaelis J: **A nano-positioning system for macromolecular structural analysis**. *Nat Meth* 2008, **5**:965-971.
The nano positioning system (NPS) methodology (https://www.uni-ulm.de/nawi/nawi-biophys/forschung/forschung-michaelis/method-development/software/nps/) determines the spatially distributed uncertainties from a set of FRET experiments, assuming the dyes adopt fixed positions in space. A Bayesian approach to address such a problem is presented which localizes dyes with fixed locations.

86. •• Brunger AT, Strop P, Vrljic M, Chu S, Weninger KR: **Three-dimensional molecular modeling with single molecule FRET**. *J Struct Biol* 2011, **173**:497-505.
Technical details on implementing FRET-restraints into modelling techniques originally developed for NMR to FRET are given. The presented implementation is publicly available in the established software CNS (Crystallography and NMR system; http://cns-online.org/v1.3/).

87. Klose D, Klare JP, Grohmann D, Kay CWM, Werner F, Steinhoff HJ: **Simulation vs. reality: a comparison of in silico distance predictions with DEER and FRET measurements**. *PLoS ONE* 2012, **7**:e39492.

88. Choi UB, Strop P, Vrljic M, Chu S, Brunger AT, Weninger KR: **Single-molecule FRET-derived model of the synaptotagmin 1-SNARE fusion complex**. *Nat Struct Mol Biol* 2010, **17**:318-324.

89. Hoefling M, Lima N, Haenni D, Seidel CAM, Schuler B, Grubmüller H: **Structural heterogeneity and quantitative FRET efficiency distributions of polyprolines through a hybrid atomistic simulation and Monte Carlo approach**. *PLoS ONE* 2011, **6**:e19791.

90. Graen T, Hoefling M, Grubmuller H: **AMBER-DYES: characterization of charge fluctuations and force field parameterization of fluorescent dyes for molecular dynamics simulations**. *J Chem Theory Comput* 2014, **10**:5505-5512.

91. Best RB, Hofmann H, Nettels D, Schuler B: **Quantitative interpretation of FRET experiments via molecular simulation: force field and validation**. *Biophys J* 2015, **108**:2721-2731.

92. Milas P, Gamari BD, Parrot L, Krueger BP, Rahmanseresht S, Moore J, Goldner LS: **Indocyanine dyes approach free rotation at the 3′ terminus of A-RNA: a comparison with the 5′ terminus and consequences for Fluorescence Resonance Energy Transfer**. *J Phys Chem B* 2013, **117**:8649-8658.

93. Liao J-M, Wang Y-T, Chen C-L: **Computer simulation to investigate the FRET application in DNA hybridization systems**. *Phys Chem Chem Phys* 2011, **13**:10364-10371.

94. Walczewska-Szewc K, Corry B: **Accounting for dye diffusion and orientation when relating FRET measurements to distances: three simple computational methods**. *Phys Chem Chem Phys* 2014, **16**:12317-12326.

95. Kalinin S, Sisamakis E, Magennis SW, Felekyan S, Seidel CAM: **On the origin of broadening of single-molecule FRET efficiency distributions beyond shot noise limits**. *J Phys Chem B* 2010, **114**:6197-6206.

96. • Levitus M, Ranjit S: **Cyanine dyes in biophysical research: the photophysics of polymethine fluorescent dyes in biomolecular environments**. *Q Rev Biophys* 2010:1-29.
This review focuses on the photophysical and photochemical properties of the most frequently used cyanine dyes in single-molecule FRET studies. Non-fluorescent dark states and methods to enhance dye fluorescence are discussed. Effects of dark states in the context of single-molecule FRET and localization studies are described.

97. Abdullin D, Hagelueken G, Schiemann O: **Determination of nitroxide spin label conformations via dPELDOR and X-ray crystallography**. *Phys Chem Chem Phys* 2016, **18**:10428-10437.

98. Hwang H, Myong S: **Protein induced fluorescence enhancement (PIFE) for probing protein-nucleic acid interactions**. *Chem Soc Rev* 2014, **43**:1221-1229.

99. • Lerner E, Ploetz E, Hohlbein J, Cordes T, Weiss S: **A quantitative theoretical framework for protein-Induced fluorescence enhancement-Förster-Type Resonance Energy Transfer (PIFE-FRET)**. *J Phys Chem B* 2016, **120**:6401-6410.
Protein-induced fluorescence enhancement (PIFE) serves as a complementary molecular ruler at molecular distances inaccessible to other spectroscopic rulers such as FRET or photoinduced electron transfer (PET). The authors provide experimental evidence that PIFE is proportional to the contact volume of the dye's accessible volume with protein interface.

100. • Beckers M, Drechsler F, Eilert T, Nagy J, Michaelis J: **Quantitative structural information from single-molecule FRET**. *Faraday Discuss* 2015, **184**:117-129.
The classic NPS methodology is extended by assumptions (e.g. dynamic dipole averaging), which drastically increases the precision in locating the position of the dyes by multiple FRET measurements. This study demonstrates that the dye description is essential to achieve a high precision to localize dyes.

101. Höfig H, Gabba M, Poblete S, Kempe D, Fitter J: **Inter-dye distance distributions studied by a combination of single-molecule FRET-filtered lifetime measurements and a weighted accessible volume (wAV) algorithm**. *Molecules* 2014, **19**:19269-19291.

102. Steffen FD, Sigel RKO, Börner R: **An atomistic view on carbocyanine photophysics in the realm of RNA**. *PCCP* 2016, **18** Advance Article.

103. Greife A, Felekyan S, Ma QJ, Gertzen CGW, Spomer L, Dimura M, Peulen TO, Wöhler C, Häussinger D, Gohlke H, Keitel V *et al.*: **Structural assemblies of the di- and oligomeric G-protein coupled receptor TGR5 in live cells: an MFIS-FRET and integrative modelling study**. *Sci Rep* 2016. Advance Article.

104. • Kravets E, Degrandi D, Ma QJ, Peulen TO, Klumpers V, Felekyan S, Kühnemuth R, Weidtkamp-Peters S, Seidel CAM, Pfeffer K: **Guanylate binding proteins directly attack *Toxoplasma gondii* via supramolecular complexes**. *eLife* 2016, **5**:e11479.
The authors use multi-parameter fluorescence image spectroscopy in live cells to validate structural models and determine equilibrium constants to study the role of immuno-proteins in intracellular pathogen defense. This study shows the application of forward modeling using live-cell data.

105. Iqbal A, Arslan S, Okumus B, Wilson TJ, Giraud G, Norman DG, Ha T, Lilley DMJ: **Orientation dependence in fluorescent energy transfer between Cy3 and Cy5 terminally attached to double-stranded nuclelic acids**. *Proc Natl Acad Sci U S A* 2008, **105**:11176-11181.

106. • Antonik M, Felekyan S, Gaiduk A, Seidel CAM: **Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions via photon distribution analysis**. *J Phys Chem B* 2006, **110**:6970-6978.
The authors established the probability distribution analysis (PDA) method for the quantitative analysis of FRET efficiency (*E*) histograms to determine with high precision (<1% of the Förster radius) the originating E value of a shot-noise-limited distribution. Any deviations can be assigned to static and/or dynamic heterogeneities of the molecular system.

107. Kalinin S, Felekyan S, Antonik M, Seidel CAM: **Probability distribution analysis of single-molecule fluorescence anisotropy and resonance energy transfer**. *J Phys Chem B* 2007, **111**:10253-10262.

108. • Raval A, Piana S, Eastwood MP, Shaw DE: **Assessment of the utility of contact-based restraints in accelerating the prediction of protein structure using molecular dynamics simulations**. *Protein Sci* 2016, **25**:19-29.
An excellent demonstration of how sparse restraints similar to those provided by FRET, enable MD-based protein structure determination.

109. Boura E, Rozycki B, Chung HS, Herrick DZ, Canagarajah B, Cafiso DS, Eaton WA, Hummer G, Hurley JH: **Solution structure of the ESCRT-I and -II supercomplex: Implications for membrane budding and scission**. *Structure* 2012, **20**:874-886.

110. Brunger AT: **Version 1.2 of the crystallography and NMR system**. *Nat Protocols* 2007, **2**:2728-2733.

111. •• Rieping W, Habeck M, Nilges M: **Inferential structure determination**. *Science* 2005, **309**:303-306.
The authors describe a forward modeling Bayesian approach to improve the structural quality of a model by taking into account indirect structural data, such as NMR data. Similar to NMR, ambiguous experimental data poses a challenge to structural modeling by FRET

112. Muschielok A, Michaelis J: **Application of the nano-positioning system to the analysis of fluorescence resonance energy transfer networks**. *J Phys Chem B* 2011, **115**:11927-11937.

113. Dumat B, Larsen AF, Wilhelmsson LM: **Studying Z-DNA and B- to Z-DNA transitions using a cytosine analogue FRET-pair**. *Nucleic Acids Res* 2016, **44**:e101.

114. Berman H, Henrick K, Nakamura H: **Announcing the worldwide Protein Data Bank**. *Nat Struct Biol* 2003, **10** 980-980.

115. Berman HM, Burley SK, Kleywegt GJ, Markley JL, Nakamura H, Velankar S: **The archiving and dissemination of biological structure data**. *Curr Opin Struct Biol* 2016, **40**:17-22.

116. Haas J, Roth S, Arnold K, Kiefer F, Schmidt T, Bordoli L, Schwede T: **The Protein Model Portal — a comprehensive resource for protein structure and model information**. *Database* 2013:bat031.

117. NIST/SEMATECH: *e-Handbook of Statistical Methods*. 2016:. http://www.itl.nist.gov/div898/handbook/.

118. Guyon I, Elisseeff A: **An introduction to variable and feature selection**. *J Mach Learn Res* 2003, **3**:1157-1182.

119. Felekyan S, Kalinin S, Sanabria H, Valeri A, Seidel CAM: **Filtered FCS: species auto- and cross-correlation functions highlight binding and dynamics in biomolecules**. *ChemPhysChem* 2012, **13**:1036-1053.

120. Byrnes LJ, Sondermann H: **Structural basis for the nucleotide-dependent dimerization of the large G protein atlastin-1/SPG3A**. *Proc Natl Acad Sci U S A* 2011, **108**:2216-2221.

121. Byrnes LJ, Singh A, Szeto K, Benvin NM, O'Donnell JP, Zipfel WR, Sondermann H: **Structural basis for conformational switching and GTP loading of the large G protein atlastin**. *EMBO J* 2013, **32**:369-384.

122. Ahmed A, Gohlke H: **Multiscale modeling of macromolecular conformational changes combining concepts from rigidity and elastic network theory**. *Proteins* 2006, **63**:1038-1051.

123. Ahmed A, Rippmann F, Barnickel G, Gohlke H: **A normal mode-based geometric simulation approach for exploring biologically relevant conformational transitions in proteins**. *J Chem Inf Model* 2011, **51**:1604-1622.

124. Chen HM, Ahsan SS, Santiago-Berrios MB, Abruna HD, Webb WW: **Mechanisms of quenching of Alexa fluorophores by natural amino acids**. *J Am Chem Soc* 2010, **132**:7244-7245.

125. Mullner D: **fastcluster: fast hierarchical, agglomerative clustering routines for R and Python**. *J Stat Softw* 2013, **53**:1-18.

126. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC: **MolProbity: all-atom structure validation for macromolecular crystallography**. *Acta Crystallogr Sect D: Biol Crystallogr* 2010, **66**:12-21.

127. Delaforge E, Milles S, Bouvignies G, Bouvier D, Boivin S, Salvi N, Maurin D, Martel A, Round A, Lemke EA, Jensen MR *et al.*: **Large-scale conformational dynamics control H5N1 influenza polymerase PB2 binding to Importin alpha**. *J Am Chem Soc* 2015, **137**:15122-15134.

128. Choudhury AR, Sikorska E, van den Boom J, Bayer P, Popenda L, Szutkowski K, Jurga S, Bonomi M, Sali A, Zhukov I, Passamonti S *et al.*: **Structural model of the bilitranslocase transmembrane domain supported by NMR and FRET data**. *PLoS ONE* 2015:10.

129. Boura E, Rozycki B, Herrick DZ, Chung HS, Vecer J, Eaton WA, Cafiso DS, Hummer G, Hurley JH: **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.

130. Doose S, Neuweiler H, Sauer M: **Fluorescence quenching by photoinduced electron transfer: a reporter for conformational dynamics of macromolecules**. *ChemPhysChem* 2009, **10**:1389-1398.

131. Noe F, Doose S, Daidone I, Lollmann M, Sauer M, Chodera JD, Smith JC: **Dynamical fingerprints for probing individual relaxation processes in biomolecular dynamics with simulations and kinetic experiments**. *Proc Natl Acad Sci U S A* 2011, **108**:4822-4827.

132. Schulze A, Beliu G, Helmerich DA, Schubert J, Pearl LH, Prodromou C, Neuweiler H: **Cooperation of local motions in the Hsp90 molecular chaperone ATPase mechanism**. *Nat Chem Biol* 2016, **12**:628-635.

133. Schröder GF, Alexiev U, Grubmüller H: **Simulation of fluorescence anisotropy experiments: Probing protein dynamics**. *Biophys J* 2005, **89**:3757-3770.

134. Cristovao M, Sisamakis E, Hingorani MM, Marx AD, Jung CP, Rothwell PJ, Seidel CAM, Friedhoff P: **Single-molecule multiparameter fluorescence spectroscopy reveals directional MutS binding to mismatched bases in DNA**. *Nucleic Acids Res* 2012, **40**:5448-5464.

135. Lakowicz JR: *Principles of Fluorescence Spectroscopy*. Springer Science & Business Media; 2013.

136. Yang H, Luo GB, Karnchanaphanurach P, Louie TM, Rech I, Cova S, Xun LY, Xie XS: **Protein conformational dynamics probed by single-molecule electron transfer**. *Science* 2003, **302**:262-266.

137. Mansoor SE, DeWitt MA, Farrens DL: **Distance mapping in proteins using fluorescence spectroscopy: the tryptophan-induced quenching (TrIQ) method**. *Biochemistry* 2010, **49**:9722-9731.

138. Haenni D, Zosel F, Reymond L, Nettels D, Schuler B: **Intramolecular distances and dynamics from the combined photon statistics of single-molecule FRET and Photoinduced Electron Transfer**. *J Phys Chem B* 2013, **117**:13015-13028.

139. Kask P, Piksarv P, Pooga M, Mets Ü, Lippmaa E: **Separation of the rotational contribution in fluorescence correlation experiments**. *Biophys J* 1989, **55**:213-220.

140. Pieper CM, Enderlein J: **Fluorescence correlation spectroscopy as a tool for measuring the rotational diffusion of macromolecules**. *Chem Phys Lett* 2011, **516**:1-11.

141. Stahl Y, Grabowski S, Bleckmann A, Kuhnemuth R, Weidtkamp-Peters S, Pinto KG, Kirschner GK, Schmid JB, Wink RH, Hulsewede A, Felekyan S *et al.*: **Moderation of Arabidopsis root stemness by CLAVATA1 and ARABIDOPSIS CRINKLY4 receptor kinase complexes**. *Curr Biol* 2013, **23**:362-371.

142. Somssich M, Ma QJ, Weidtkamp-Peters S, Stahl Y, Felekyan S, Bleckmann A, Seidel CAM, Simon R: **Real-time dynamics of peptide ligand-dependent receptor complex formation in planta**. *Sci Signal* 2015:8.

143. Blom H, Widengren J: **STED microscopy — towards broadened use and scope of applications**. *Curr Opin Chem Biol* 2014, **20**:127-133.

144. Bonomi M, Pellarin R, Kim SJ, Russel D, Sundin BA, Riffle M, Jaschob D, Ramsden R, Davis TN, Muller EGD, Sali A: **Determining protein complex structures based on a Bayesian model of in vivo Förster Resonance Energy Transfer (FRET) data**. *Mol Cell Proteomics* 2014, **13**:2812-2823.

145. Vogel SS, van der Meer BW, Blank PS: **Estimating the distance separating fluorescent protein FRET pairs**. *Methods* 2014, **66**:131-138.

146. • Gnutt D, Gao M, Brylski O, Heyden M, Ebbinghaus S: **Excluded-volume effects in living cells**. *Angew Chem Int Ed* 2015, **54**:2548-2551.

The effect of macromolecular crowding inside living cells is studied by a FRET based sensor. This study is of relevance to relate *in vitro* FRET-studies to live-cell measurements.

147. Senske M, Tork L, Born B, Havenith M, Herrmann C, Ebbinghaus S: **Protein stabilization by macromolecular crowding through enthalpy rather than entropy**. *J Am Chem Soc* 2014, **136**:9036-9041.

148. Sakon JJ, Weninger KR: **Detecting the conformation of individual proteins in live cells**. *Nat Meth* 2010, **7**:203-205.

149. König I, Zarrine-Afsar A, Aznauryan M, Soranno A, Wunderlich B, Dingfelder F, Stuber JC, Pluckthun A, Nettels D, Schuler B: **Single-molecule spectroscopy of protein conformational dynamics in live eukaryotic cells**. *Nat Meth* 2015, **12**:773-779.

150. Ratzke C, Mickler M, Hellenkamp B, Buchner J, Hugel T: **Dynamics of heat shock protein 90 C-terminal dimerization is an important part of its conformational cycle**. *Proc Natl Acad Sci U S A* 2010, **107**:16101-16106.