Quantitative Description of Intrinsically Disordered Proteins Using Single-Molecule FRET, NMR, and SAXS

Samuel Naudi-Fabra, Maud Tengo, Malene Ringkjøbing Jensen, Martin Blackledge, Sigrid Milles

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
pubs.acs.org/JACS
Article

Quantitative Description of Intrinsically Disordered Proteins Using Single-Molecule FRET, NMR, and SAXS

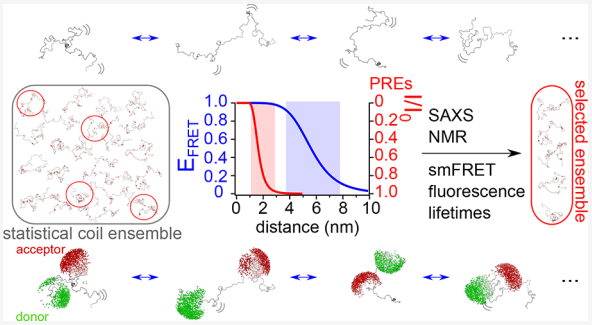
Samuel Naudi-Fabra, Maud Tengo, Malene Ringkjøbing Jensen, Martin Blackledge, and Sigrid Milles*

Cite This: J. Am. Chem. Soc. 2021, 143, 20109–20121

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Studying the conformational landscape of intrinsically disordered and partially folded proteins is challenging and only accessible to a few solution state techniques, such as nuclear magnetic resonance (NMR), small-angle scattering techniques, and single-molecule Förster resonance energy transfer (smFRET). While each of the techniques is sensitive to different properties of the disordered chain, such as local structural propensities, overall dimension, or intermediate- and long-range contacts, conformational ensembles describing intrinsically disordered proteins (IDPs) accurately should ideally respect all of these properties. Here we develop an integrated approach using a large set of FRET efficiencies and fluorescence lifetimes, NMR chemical shifts, and paramagnetic relaxation enhancements (PREs), as well as small-angle X-ray scattering (SAXS) to derive quantitative conformational ensembles in agreement with all parameters. Our approach is tested using simulated data (five sets of PREs and 15 FRET efficiencies) and validated experimentally on the example of the disordered domain of measles virus phosphoprotein, providing new insights into the conformational landscape of this viral protein that comprises transient structural elements and is more compact than an unfolded chain throughout its length. Rigorous cross-validation using FRET efficiencies, fluorescence lifetimes, and SAXS demonstrates the predictive nature of the calculated conformational ensembles and underlines the potential of this strategy in integrative dynamic structural biology.



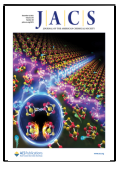
INTRODUCTION

Intrinsically disordered proteins (IDPs) play important roles in many biological systems and exert their tasks thanks to their ability to sample conformational ensembles that can have different degrees of compactness and that often comprise transiently folded regions functioning as interaction sites.[1,2] Although IDPs are known to be devoid of stable secondary and tertiary structures, primary structure determines their function and modulates the conformations sampled on a rapid time scale: small motifs can locally enrich the IDP in hydrophobic amino acids, and clusters of charged residues may lead to self-repulsion, thus affecting the properties of the chain.[3−5]

Single-molecule Förster resonance energy transfer (smFRET) has demonstrated to be a very powerful tool to access the dimension of the unfolded chain through the measurement of energy transfer between site-specifically attached donor and acceptor fluorophores as a function of their distance.[6,7] The technique is compatible with very large IDPs,[8] covering distances that range from 2 to 10 nm approximately, and structural information can be obtained in the presence of transiently folded or folded domains,[9] in complex environments, and even within the living cell.[10,11] Obtaining quantitative structural insight has, however, remained challenging in particular as the distance between the fluorophores, rather than between their attachment points in the protein backbone, is determined experimentally, and the chemical composition of the dyes and their linkers therefore has to be taken into account in structural modeling. For folded proteins, recent advances have overcome this problem by generating structural models explicitly considering the attached fluorophores mainly through calculation of the volumes that the fluorophores can occupy when attached to a specific site in the protein (accessible volumes, AVs).[12−15]

Determination of distances for IDPs suffers from the additional challenge that the measured FRET efficiency (E_{FRET}) describes an ensemble of distances rather than an individual distance, which has frequently been taken into account by assuming the sampling of a Gaussian chain (or other polymer-) distribution between the fluorophores.[16] These distributions can be expressed as a function of the number of amino acids between the attachment points of the fluorophores, and in order to consider the contribution of the fluorophores and their linkers to the measured distance, they are usually assumed to contribute a number of additional

Received: June 17, 2021
Published: November 24, 2021



https://doi.org/10.1021/jacs.1c06264
J. Am. Chem. Soc. 2021, 143, 20109−20121

residues. Although this approach has led to distance distributions in agreement with conformational ensembles derived from other experimental techniques (nuclear magnetic resonance, NMR, and small-angle X-ray scattering, SAXS),[17] the number of amino acids that has to be added to consider the dyes and their linkers is not unambiguous.[17−19] This has consequences when distances within IDPs are measured by different techniques. Radii of gyration (R_G) measured using SAXS and those inferred from end-to-end distances (R_E) using smFRET have apparently disagreed for a long time.[20−22] A number of approaches have been presented to resolve this controversy, employing improved analysis procedures and explicit ensembles, generated using Bayesian statistics or maximum entropy approaches, in agreement with smFRET and SAXS.[19,23−25] In this context, fluorophores have been attached *in silico* to describe measured E_{FRET} of individual distances (one distance per protein).[19,26] While these approaches are promising, the study of IDPs demands a systematic analysis integrating distance information between different regions of the protein, its global extension, but also local structural information to accommodate heterogeneity in compaction, as well as population of transiently structured elements.

Here, we propose an approach for the systematic integration of various solution state structural data of IDPs based on the implementation of FRET efficiencies into the algorithm ASTEROIDS that derives representative structural ensembles of IDPs from NMR and SAXS data describing both local conformational propensities and long-range distance information.[27,28] Our approach is based on the selection of smaller ensembles from a large statistical coil ensemble (calculated using flexible-meccano[29] and of an extension approximately equal to a fully unfolded protein[30]) solely using experimental data, and the fluorophores are explicitly taken into account through the per-conformer calculation of AVs. This strategy does not require a conversion between different distance measures (e.g., R_G and R_E), nor does it require an approximation of the dyes/linker length in the context of a polymer model and therefore allows describing IDPs of varying degrees of compactness along their sequence, theoretically even including entirely folded domains. We first selected and cross-validated conformational ensembles using a large set of *in silico* PRE (paramagnetic relaxation enhancement) and FRET data. Finally, we validate our approach with respect to experimental FRET efficiencies, SAXS data, as well as NMR chemical shifts and PREs, obtaining new insights into the conformational landscape of an intrinsically disordered region of the measles virus phosphoprotein. Notably, in addition to a number of FRET efficiencies and SAXS data, we also use experimental fluorescence lifetimes of the FRET-labeled protein to cross-validate our conformational ensemble, demonstrating correct sampling of the ensemble itself as well as the dye AVs. We demonstrate complementarity between different parameters (particularly FRET and PREs) and the importance of using distance information across the IDP sequence to generate meaningful conformational ensembles. The presented approach now allows addressing dynamic integrated structural biology quantitatively and in a predictive manner.

RESULTS

FRET Distance Networks in Conformational Ensembles. In order to determine conformational ensembles based on experimental smFRET data, we build on an approach that has been developed and frequently used for calculating conformational ensembles based on diverse NMR parameters and SAXS.[1,31−34] A large ensemble of conformers (e.g., 10 000) is calculated based on a statistical distribution of Φ and Ψ angles of the protein backbone using the software flexible-meccano.[29] From this large ensemble, smaller subensembles that describe the experimental data are selected using the genetic algorithm ASTEROIDS.[35]

Distance measurements through FRET rely on the attachment of a donor and an acceptor fluorophore to specific sites within the protein chain. Our goal being to describe the experimental FRET efficiencies directly, the fluorescent dyes thus have to be accounted for in the conformational ensemble. We calculated accessible volumes for the fluorophores Alexa488 and Alexa594 attached to cysteines via maleimide chemistry and comprising a C_5 linker connecting the Cys side chain and the fluorophore, as previously described.[12,36] As a first step, we calculated a conformational ensemble of a 110 amino acid long model protein containing two cysteines as dye attachment points, and we calculated AVs for every conformer in the ensemble. Both sampling of the AV and sampling of the different conformers were assumed to be on a time scale significantly slower than the fluorescence lifetime, as suggested by AV sampling based on molecular dynamics simulations of fluorescently labeled DNA.[37] We first used the cysteine side chains as attachment points. In order to allow labeling positions that are not native cysteines and that are experimentally generated through point mutations, we estimated the average distance between the C_β atom and the SH and elongated the linker length in the simulation accordingly (see Methods). The distance distributions calculated on a 100 conformer ensemble with attachment points at either the SH or C_β with their respective parametrization of the linker can be considered equal (SI Figure 1).

For the selection of meaningful ensembles, AVs have to be calculated and FRET efficiencies determined for all conformers in the large flexible-meccano ensemble before selection using ASTEROIDS. Since AV calculation is time-consuming, the iterative sampling of positions in the AV was optimized to 500 iterations (Figure 1) and the pairwise distance calculation coarsened (see Materials and Methods).

Benchmarking an Ensemble Selection Using FRET against *in Silico* Data. After optimizing AV calculations for multiconformational ensembles, we investigated whether FRET efficiencies (E_{FRET}) could be used in the context of the ensemble selection algorithm ASTEROIDS. For this, we used an IDP sequence of 155 amino acids in length, for which we calculated an ensemble comprising a long-range contact between amino acid segments 15−25 and 90−100 and for which we generated 15 *in silico* E_{FRET} (SI Figure 2) using AV calculations as described above. In order to obtain distances that adequately reflect the long-range behavior of the ensemble, we selected labeling positions covering different regions of the protein and care was taken to cover both short and long amino acid distances between the attachment points of the labels so as to address FRET efficiencies throughout the sensitive regime of FRET (around 2−10 nm).

From a large statistical-coil ensemble calculated using flexible-meccano, we then selected smaller subensembles of 200 conformers in size using ASTEROIDS based on six of the 15 *in silico* FRET efficiencies (Figure 2). When the remaining

https://doi.org/10.1021/jacs.1c06264
J. Am. Chem. Soc. 2021, 143, 20109−20121

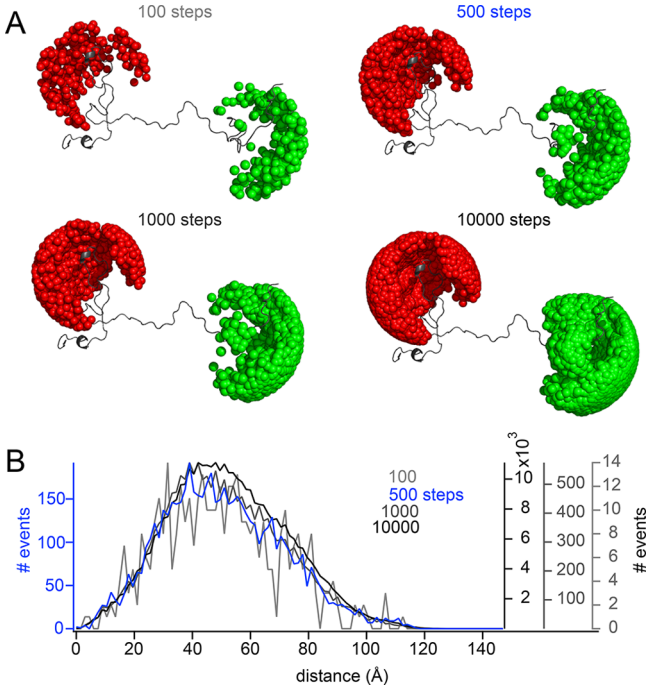


Figure 1. Influence of step size on conformational ensembles. (A) Examples of conformations of a model protein with accessible volumes (AVs) of Alexa488 (green) and Alexa594 (red), calculated using 100, 500, 1000, or 10 000 iterations (steps) for position determination. (B) Distance histogram over accessible volumes calculated over a 100 conformer ensemble using 100 (light gray), 500 (blue), 1000 (dark gray), and 10 000 (black) dye positions sampled iteratively.

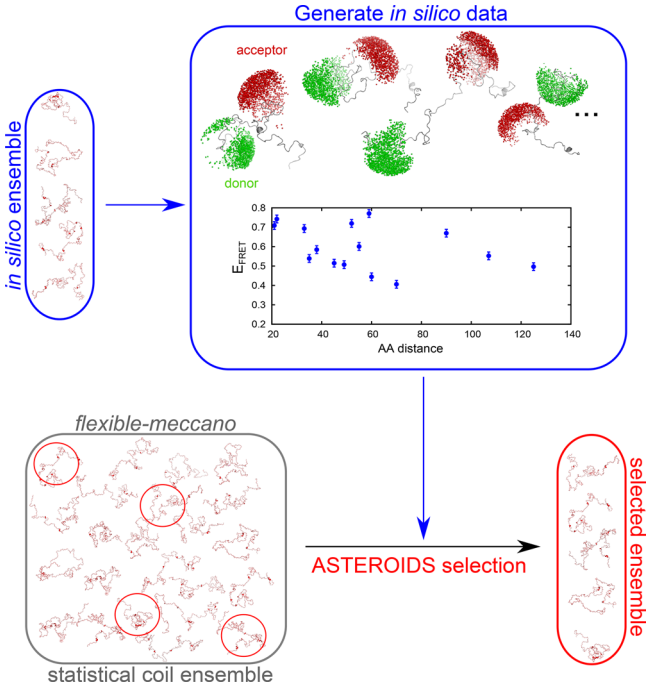


Figure 2. Scheme of incorporation of FRET distances into ASTEROIDS based on simulated data. An *in silico* ensemble of conformations is generated, for which accessible volumes occupied by Alexa488 and Alexa594 are computed and FRET data are calculated (blue frame). FRET efficiencies are used as an input for ASTEROIDS selection (red frame) from a pool of statistical coil conformers (calculated from flexible-meccano, gray frame).

nine FRET efficiencies, that were not used in the selection, were back-calculated from the selected ASTEROIDS ensemble, the *in silico* FRET efficiencies of the input ensemble comprising a long-range contact were predicted with high accuracy (SI Figure 3A).

The FRET efficiencies used in this selection were chosen to represent varying distances across the sequence of the protein, and sufficient sampling of the different regions of the protein is indeed crucial for reproducing the long-range characteristics of the ensemble with confidence. If only three FRET efficiencies were used in the selection, even when distributed along the sequence, the remaining FRET efficiencies not used in the selection were only poorly predicted by the ASTEROIDS ensemble and the long-range distances of the simulated target ensemble much less well captured (SI Figure 4).

PREs and FRET Distances Provide Complementary Long-Range Distance Information. Through paramagnetic relaxation enhancements, NMR also offers a probe for longer range distances that can reach up to around 2.5 nm.[38] For this, a paramagnetic probe (usually a spin radical) is attached to a site-specifically engineered cysteine within the protein chain, and its effects on spin relaxation of the different $^{1}H_N$ nuclei within the protein backbone are measured and depend on the inverse sixth order of the respective distance from the spin radical. PREs thus have a distance dependence similar to FRET with, however, different sensitive regimes (Figure 3A–C). Indeed, the distance windows at which FRET and PREs are sensitive, respectively, are entirely complementary, and only both techniques together are expected to provide insights into both intermediate (around 1–3 nm) and long-range (around 4–8 nm) distance ranges.

While the contribution of paramagnetic relaxation can be directly determined through the measurement of spin relaxation, we do not have access to the FRET rate (k_{ET}) itself, which can only be measured indirectly through the fluorescence lifetime of the donor in the absence (τ_D) and presence ($\tau_{D(FRET)}$) of the acceptor,

$$\tau_{D(FRET)} = \frac{\tau_D}{1 + k_{ET}\tau_D} \quad (1)$$

or through the FRET efficiency,

$$E_{FRET} = k_{ET}\tau_{D(FRET)} = \frac{k_{ET}\tau_D}{1 + k_{ET}\tau_D} \quad (2)$$

leading to a dampened dependence between the measurement parameter (τ_D or E_{FRET}) and the donor–acceptor distance for short distances. A similar dependence can be obtained if peak intensity ratios (I/I_0) of the para- and diamagnetic PRE sample are considered, allowing a visual inspection of the complementary distance ranges (Figure 3B). We calculated PREs[39] using five different attachment sites for a spin radical in our long-range ensemble and used these *in silico* PREs to select smaller subensembles of 200 conformers using ASTEROIDS. Although all PREs are captured very well in these ensembles, they fail to reproduce the expected FRET efficiencies (SI Figure 3C,D). This observation remains true also if fast (faster than the fluorescence lifetime) sampling of the AVs was assumed (SI Figure 5). The selection based on six FRET efficiencies described above, on the other hand, also fails to reproduce the expected PREs, thus illustrating the expected

https://doi.org/10.1021/jacs.1c06264
J. Am. Chem. Soc. 2021, 143, 20109–20121

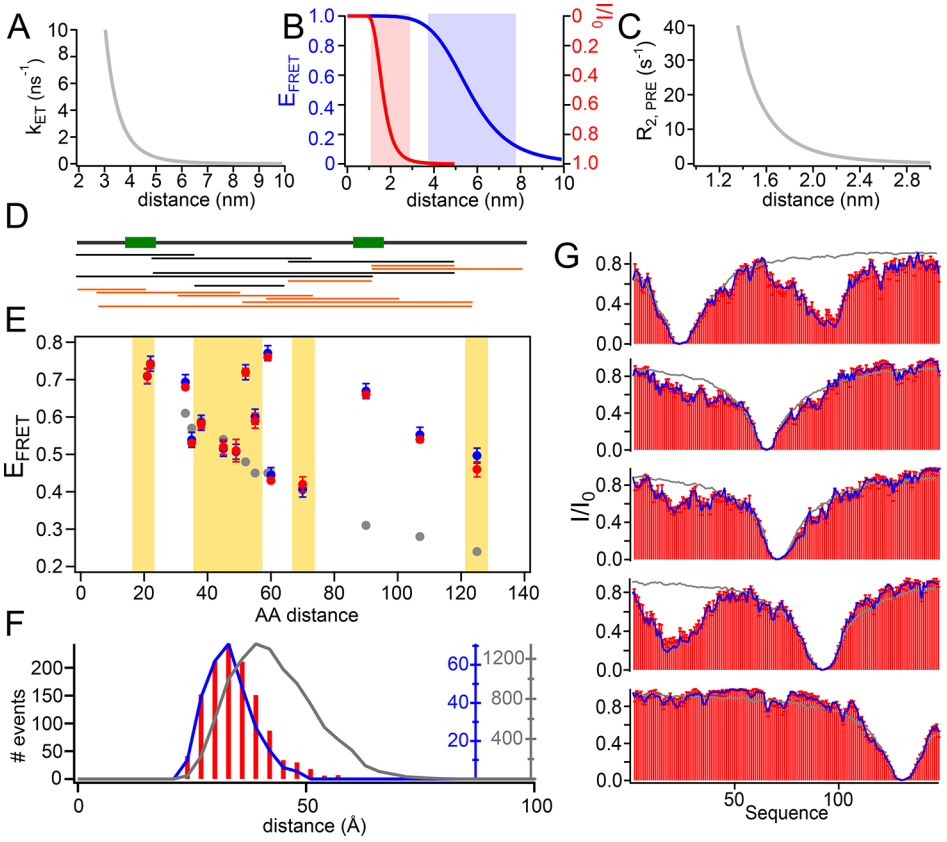


Figure 3. ASTEROIDS selection based on PREs and FRET efficiencies. (A) Dependence of the FRET rate (k_{ET}) on distance with a Förster radius of 56 Å and a fluorescence lifetime of the donor (τ_D) of 4 ns. (B) Dependence of the FRET efficiency (E_{FRET}) on distance for a k_{ET} as displayed in A (blue curve, blue *y* axis). The red curve with the red *y* axis shows the dependence of peak intensity ratios (I/I_0) for a paramagnetic as compared to a diamagnetic sample at a PRE rate described in C. Red and blue shading illustrate the distance ranges to which FRET and PREs are sensitive. (C) PRE rate ($R_{2,PRE}$) dependence on the distance between the proton and electron spins for τ_c = 5 ns, with $\tau_c = \tau_r\tau_s/(\tau_r + \tau_s)$, τ_r being the rotational correlation time of the protein and τ_s the effective electron relaxation time (see ref 38). Note that reorientation dynamics of the spin label was not taken into account for this illustration, but was considered in the ensemble calculations. (D) Schematic of a 155 residue long IDP comprising a long-range interaction between regions indicated by green boxes. Below: Distances for which E_{FRET} has been calculated after *in silico* addition of the fluorescent dyes. Black protein constructs have been used in the ASTEROIDS selection; orange protein constructs have been used for cross validation (corresponding E_{FRET} above yellow background in E). See also SI Figure 2 for a more detailed scheme. (E) FRET efficiencies plotted against the amino acid distance between the labels of a flexible-meccano ensemble (gray), the simulated ensemble with a long-range contact (blue), and an ensemble selected based on six FRET efficiencies and five PRE labeling positions (red). Only E_{FRET} on a white background have been used in the selection. Cross-validated distances are on a yellow background. Error bars on the blue points indicate the error in FRET efficiency that was allowed in the selection (0.02). Error bars on the red points refer to the standard deviation of E_{FRET} calculated from six independent selections. (F) Histogram of average pairwise C_α–C_α distances of the flexible-meccano ensemble (gray), the simulated ensemble (blue), and the ensemble selected based on PREs and six different FRET distances (red bars). (G) PREs of a flexible-meccano ensemble (gray lines), of the simulated ensemble with a long-range contact (blue lines), and of the selected ensemble (red bars). All simulated PREs (in blue) were used in the selection. Red error bars are standard deviations over six independent selections.

complementary distance ranges to which PREs and FRET are sensitive (SI Figure 3B).

An ensemble that has been selected based on five *in silico* PRE labeling sites and six FRET efficiencies, however, leads to an excellent reproduction of all *in silico* PREs and E_{FRET} (Figure 3), and this ensemble also reliably reproduces the expected average pairwise as well as specific C_α–C_α distance distributions (Figure 3F, SI Figure 6) that can be calculated directly from the selected ensemble without additional approximation concerning fluorescent dyes and their linkers (or PRE labels).

Indeed FRET efficiencies and PREs are both necessary to correctly describe a conformational ensemble that populates various intermediate- and long-range distances. Including only FRET or only PREs into a selection can only be expected to reproduce the respective other parameter for a very narrow distance window and depending on the properties of the pool of conformers from which ensembles are selected. We demonstrate this on the example of a new set of *in silico* data, in which we allowed the long-range contact to reach up to 50 rather than 20 Å, to which FRET efficiencies, but not PREs are sensitive. In this case, selection based on six FRET efficiencies leads to agreement with the *in silico* PREs, which are not noticeably different from a flexible-meccano-derived statistical coil (SI Figure 7).

Analysis of Ensemble Sizes. Ensemble selections based on *in silico* data back-calculated from a known target ensemble also allowed us to test the number of conformers required to represent the data and sufficient to reliably reproduce the statistics of the target ensemble. We have thus performed selections of 10, 20, 50, 100, 200, and 400 conformers per ensemble and calculated average absolute deviations from the

https://doi.org/10.1021/jacs.1c06264
J. Am. Chem. Soc. 2021, 143, 20109−20121

in silico data. This analysis indicates that reproduction of the data improves as the ensemble size increases (Figure 4A and B), reaching excellent agreement with the *in silico* data starting from around 200 conformers per ensemble. Reproduction of the $C_\alpha-C_\alpha$ distance distributions between the labeling sites is comparatively poor at low numbers of conformers, and only starts improving once an ensemble size of approximately 100 conformers is reached. Reproduction further improves with increasing numbers of conformers (Figure 4C and SI Figure 6). We thus conclude that, overall, an ensemble size of 200 conformers, as proposed earlier for ensembles selected based on PREs and residual dipolar couplings,[39] is a good size to reconcile reproducibility, statistics, and computation speed.

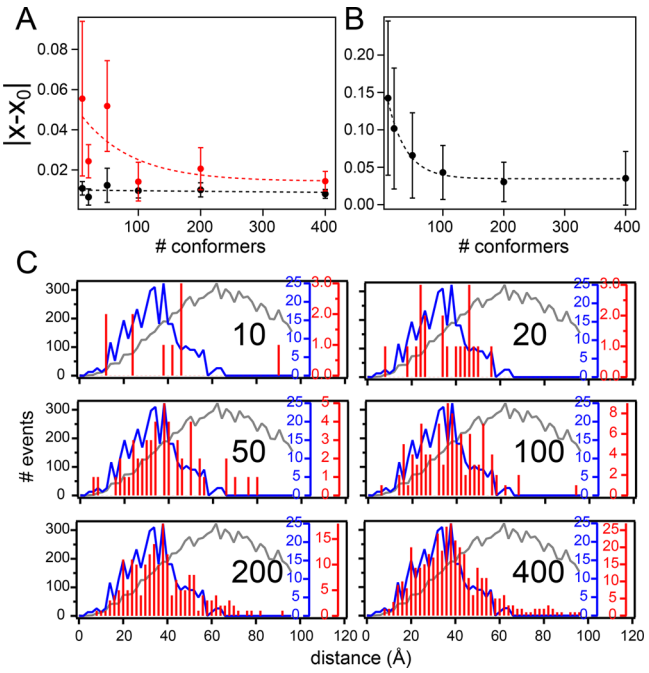


Figure 4. Varying the size of the selected ensemble. (A and B) Averaged absolute deviations of the FRET efficiency (A) or PRE (B) as calculated from the selected ensemble (x) from the respective values of the target *in silico* ensemble (x_0). Error bars show the corresponding standard deviations. Red points illustrate data not used in the selection. Ensemble sizes were 10, 20, 50, 100, 200, or 400 conformers. Dashed lines represent exponential fits representing the trend of the data. (C) $C_\alpha-C_\alpha$ distances between the *in silico* labeling sites 2 and 92 for different ensemble sizes. In red are the distances calculated from one selection based on six FRET efficiencies and using five PRE labeling sites. The expected $C_\alpha-C_\alpha$ distances are shown in blue; the distances obtained from a flexible-meccano statistical coil ensemble in gray. Black numbers inside the graphs indicate the numbers of conformers used in the selected ensembles.

Description of Experimental FRET, PREs, and Chemical Shift Data. While our comprehensive *in silico* data set demonstrates how to accurately describe long-range distances within intrinsically disordered proteins, we aimed to test the validity of this approach on experimental data. For this, we used a 110 residue long protein from the disordered N-terminus of the measles virus phosphoprotein (P_{1-100}). This protein has been extensively characterized by NMR spectroscopy[1,40] and harbors two transient α-helices, as can be inferred from backbone chemical shifts (Figure 5D). We acquired nine FRET efficiencies, PREs from five different labeling sites (Figure 5A–C), a full set of backbone chemical shifts[1] sensitive to local structural propensities, and SAXS reporting on the distribution of R_G, i.e., the overall dimension of the protein. FRET efficiencies, obtained from random labeling of two engineered cysteines with Alexa488 and Alexa594 using maleimide chemistry, were recorded on a custom-built single-molecule fluorescence spectrometer. The corrected (see Methods for details) FRET histograms were fit with double-Gaussians describing populations at $E_{FRET} = 0$ (donor only population) and at $E_{FRET} > 0$, which was extracted for ensemble selection or cross-validation (Figure 5A, SI Figure 8 and SI Table 1). Comparison of the experimentally obtained FRET efficiencies with efficiencies expected from a flexible-meccano statistical coil ensemble suggests that P_{1-100} samples a conformational ensemble that is slightly more compact than a random coil.

Conformational ensembles comprising 200 conformers (see SI Figure 9 for an assessment of ensemble sizes) were selected using ASTEROIDS based on all PREs, chemical shifts (N, H_N, C_O, C_α, C_β), and six of the nine experimental E_{FRET}. FRET efficiencies were included in the ensemble selection as described above, and the selected ensemble reliably reproduced the data used in the selection (Figure 5D, E, and F) as well as the four FRET efficiencies that have not been used in the selection (Figure 5F). A SAXS curve that was acquired from P_{1-100} and not used in the selection was also well described by the ASTEROIDS ensemble selected based on PREs, chemical shifts, and FRET efficiencies, suggesting that the ensemble also captured the overall dimension of the protein (Figure 5G). Analysis of the experimental SAXS curve as well as the SAXS curve back-calculated from the ensemble using extended Guinier analysis[41] yielded comparable R_G values, which were also in agreement with the average R_G calculated directly from the selected conformational ensemble (SI Figure 10). The scaling exponent calculated from the selected ensembles, indicative of solvent quality, was determined to be 0.52, in agreement with θ-solvent conditions (SI Figure 10A) under which excluded volume interactions cancel out.[25,42]

Our experimental data combined with ASTEROIDS selections based on only FRET or only PREs show that long-range and intermediate- range distances of the conformational ensemble are only correctly sampled when combining both sets of data (SI Figure 11). This is in agreement with the theoretical complementarity of FRET and PREs regarding their sensitive distance ranges (Figure 3B), as shown on the example of an *in silico* data set (SI Figure 3). It is interesting to note that integration of PREs into the selection also improves the reproduction of two of the experimental FRET efficiencies, indicating that the FRET efficiencies alone might not sufficiently cover all relevant protein regions in the case of P_{1-100}.

As, for this experimental data set, it is *a priori* not known on what time scale the fluorescent dyes sample the accessible volume, we additionally considered the other extreme case of AV sampling significantly faster than the fluorescence lifetime. FRET efficiencies of all conformers in the pool from which ensembles were selected were thus calculated under this assumption, and an ASTEROIDS selection was performed on the basis of six FRET efficiencies, five sets of PREs, and chemical shifts. This ensemble reproduces the FRET efficiencies not used in the selection less well than when slow (slower than the fluorescence lifetime) AV sampling was assumed (compare SI Figure 5B to Figure 5F). We thus

https://doi.org/10.1021/jacs.1c06264
J. Am. Chem. Soc. 2021, 143, 20109−20121

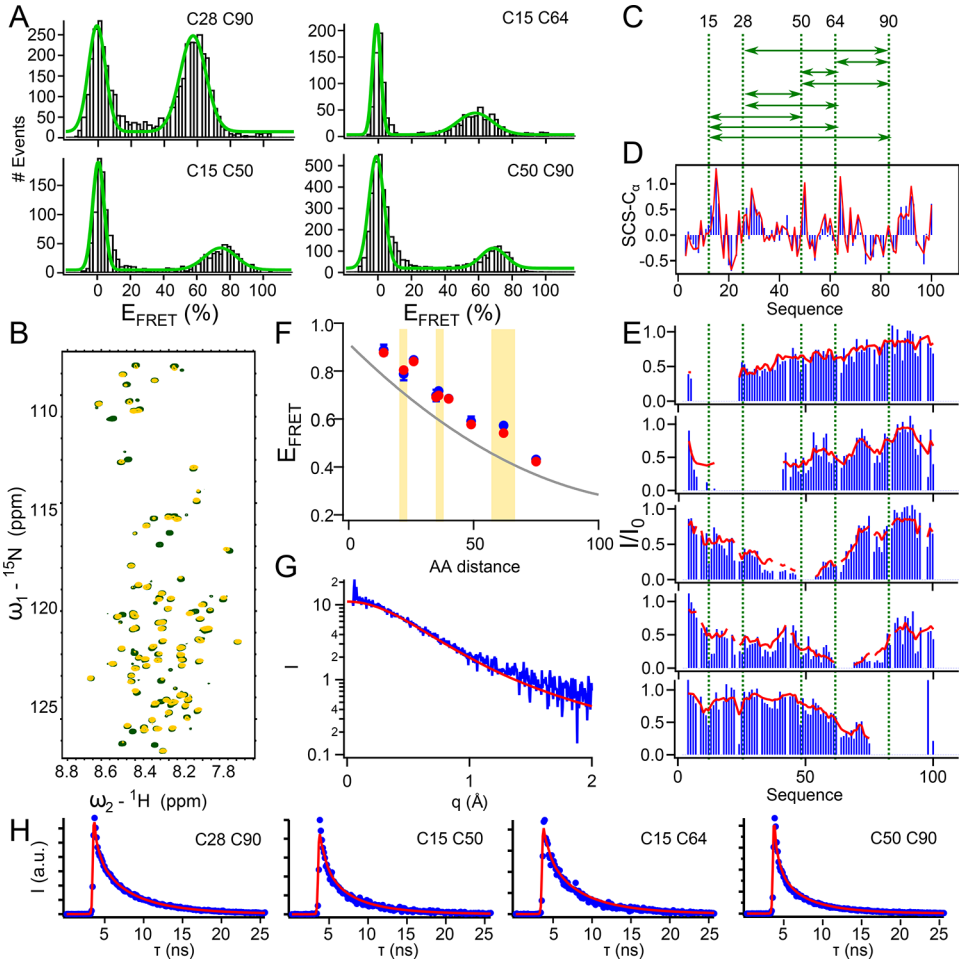


Figure 5. Description of experimental FRET, PREs, and SAXS by a common multiconformational model. (A) Experimental FRET histograms of P_{1-100} (black bars) with double Gaussian fit (green) from which E_{FRET} of the nonzero population was extracted. (B) $^{1}H-^{15}N$ heteronuclear single quantum coherence (HSQC) spectrum of P_{1-100} C64 unlabeled (green) and labeled with MTSL (yellow). (C) Visualization of FRET distances for which data have been acquired. (D) C_α secondary chemical shifts of P_{1-100} calculated based on experimental chemical shifts (blue) and based on chemical shifts calculated from an ensemble selected based on five PRE labeling positions, six FRET efficiencies and chemical shifts (red). (E) Experimental (blue) PREs and PREs calculated from the selected ensemble (red). All PREs were used in the selection. PRE labeling sites are indicated by green dashed lines (note that the same cysteines have been used for PRE and FRET labeling). Intensity ratios between the PRE labeled (I) and unlabeled (I_0) peaks are shown. (F) FRET efficiencies (E_{FRET}) of P_{1-100} plotted against the amino acid distance between the fluorophores. The gray line indicates values expected from a flexible-meccano statistical coil (polynomial fit of *in silico* data presented in Figure 3). Experimental data are shown in blue with error bars resulting from standard deviations calculated from independent measurements. Red points indicate E_{FRET} calculated from the ASTEROIDS selection. Data points plotted in front of a yellow background were not used in the selection. (G) Experimental SAXS curve (blue) and SAXS curve back-calculated from the ASTEROIDS ensemble (red). SAXS data were not used in the selection. (H) Cumulated fluorescence lifetime histograms calculated from the FRET population of the single molecule data (corresponding to FRET mutants shown in (A)). Blue points are experimental data, and red curves are decays back-calculated from the selected ensemble, comprising a scattering contribution and scaled to best fit the experimental data.

conclude that "slow" AV sampling is appropriate for the P_{1-100} experimental FRET data. We note, however, that more rapid diffusion of fluorescent dyes has been observed for other experimental systems.[43,44]

As an additional cross-validation of both AV sampling and calibrations employed for the experimental smFRET experiments, we labeled one sample of P_{1-100} (C28−C64) with a different dye pair (Alexa488/Alexa647) and determined its FRET efficiency (SI Figure 12). In parallel, we simulated the Alexa488/Alexa647 dye pair onto the ensemble selected based on smFRET (Alexa488/Alexa594), PREs, and chemical shifts. The difference between experimental E_{FRET} (0.52) and E_{FRET} expected from the selected ensemble (0.56) is below the common error determined by a recent multilaboratory study.[45]

While, in all ASTEROIDS selections, an error of 0.02 for E_{FRET} was allowed in agreement with the measurement error over several independent measurements, a larger allowed error might be considered appropriate[45] as the measured quantum yields, Förster distance R_0, or determination of spectral crosstalk is also error prone. ASTEROIDS selections based on six FRET efficiencies, five sets of PREs, and chemical shifts allowing an error of 0.06, however, are in very good agreement with those selected allowing an error of 0.02 in the case of P_{1-100} (SI Figure 13).

Reproduction of Experimental Fluorescence Lifetimes by Conformational Ensembles. In addition to intensity-based FRET efficiencies, calculated as a function of the number of emitted photons (cross-talk and background

https://doi.org/10.1021/jacs.1c06264
J. Am. Chem. Soc. 2021, 143, 20109−20121

corrected; see Methods) of the donor (I_D) and the acceptor (I_A),

$$\langle E_{FRET} \rangle = \frac{I_A}{I_A + \gamma I_D} \quad (3)$$

and corrected for differences in quantum yield and detection efficiency in the green and red channel (γ), fluorescence lifetimes provide a complementary measure for distance distributions of a conformational ensemble.[46,47] While, for a static donor−acceptor distance, E_{FRET} can be calculated from fluorescence lifetimes of the donor in the absence (τ_D) and presence of the acceptor ($\tau_{D(FRET)}$; see also eq 2), this is not the case for distances with dynamics longer than the fluorescence lifetime and shorter than the interphoton time (usually on the order of tens of microseconds):[47,48]

$$\langle E_{FRET} \rangle \neq 1 - \frac{\langle \tau_{D(FRET)} \rangle}{\tau_D} \quad (4)$$

Indeed, taking into account fluorescence lifetimes in the conformational ensemble of an IDP is complex, as every conformer in the ensemble contributes a single-exponential decay to the time-resolved fluorescence intensity of a time-correlated single photon counting (TCSPC) experiment, and the resulting multiexponential intensity decay is then experimentally convolved with the instrument response function (IRF) of the smFRET setup.[17]

In order to test whether the distance distributions of our conformational selection are in agreement with our experimental fluorescence lifetimes, we first extracted the fluorescence intensity decays of the FRET population from our single-molecule data (SI Figure 14A). The IRF was measured independently under the same experimental conditions, described with a double Gaussian function, and convoluted with the multiexponential decays expected for our conformational ensemble. The resulting decay curves described the experimental intensity decays remarkably well (Figure 5H, SI Figure 14B), indicating that our conformational ensemble correctly reproduces another set of independent long-range data that was not used in the ASTEROIDS selection process, thus confirming the validity of the selected ensemble as well as the time scales applied for motional sampling of both dyes and proteins within the ensemble.

DISCUSSION

A molecular description of the conformational landscape sampled by IDPs and proteins containing intrinsically disordered regions (IDRs) is of paramount interest, as IDPs and IDRs are enriched in several essential biological processes, such as signaling,[49,50] cellular transport processes,[51,52] and gene regulation,[53,54] and their misregulation is often also linked to disease.[55] Although multiconformational models have been conceived using mainly NMR and small angle scattering data,[29,39,56−58] and in some individual cases single-molecule FRET efficiencies,[19,26] those approaches fall short in integrating specific long-range and short-range information in a predictive manner.

We now demonstrate a tool-set to integrate the three most powerful techniques for the analysis of IDPs: NMR, SAXS, and single-molecule FRET. We show the integration of several FRET efficiencies into ensemble selections, and we reproduce them with confidence. We perform the selection using the experimentally obtained FRET efficiencies rather than their inferred distances and reproduce the corresponding fluorescence lifetimes.

Modeling of the fluorophores in terms of accessible volumes[12] on top of the pool of conformers from which the ensembles are selected is key to allowing an integration of parameters from techniques that have different experimental requirements: the attachment of fluorophores or spin radicals for single-molecule FRET and PREs, or no labeling/stable isotope labeling for SAXS/NMR. This approach assumes that the conformational ensemble remains quasi-identical in the presence and absence of the different labels (FRET/PRE) and that the parametrization of the AVs accurately reproduces the volumes sampled by the fluorophores. Successful cross-validation of a number of FRET efficiencies (including one with a different dye pair) not used in the selection and a SAXS curve suggest that these assumptions are indeed correct. The selection of explicit ensembles combined with the *in silico* attachment of labels also allows for its use if complex distance distributions are sampled that include transiently folded protein regions or even entire folded domains.[1,49,59] Distance distributions within the protein backbone can be directly calculated from the selected ensemble. While we employ the genetic algorithm ASTEROIDS[27] to select conformational ensembles in agreement with the experimental data, our developments concerning the integration of fluorophore AVs into conformational ensembles as well as insights into sampling of (sufficient) FRET distances along the protein sequence can also be used with other ensemble selection approaches.[19,23,26]

Importantly, we show that we can reproduce not only the FRET efficiencies that were used for the ensemble selection and cross-validate additional FRET efficiencies but also their corresponding fluorescence lifetimes. As fluorescence lifetimes of a FRET sample also depend on the distance distribution between the two attached fluorophores[47] and have thus frequently been used in the analysis of folded as well as intrinsically disordered proteins,[17,33,48,60−62] these results are particularly remarkable testifying to the predictive nature of our ensembles by reproducing an independent data set.

We show that PREs and FRET efficiencies provide complementary intermediate- and long-range information on the conformational ensemble, and it is worth noting that the ensembles selected on the basis of chemical shifts, PREs, and FRET efficiencies also reproduce an independently measured SAXS curve. This shows that these fundamentally different experimental techniques effectively agree with each other, therefore also supporting recent advances resolving[19,23] the long-lasting controversy concerning compaction of IDPs measured by smFRET and SAXS.[19−24]

Apart from contributing distance ranges much longer than those accessible by PREs, including smFRET into the calculation of conformational ensembles of IDPs or proteins comprising intrinsically disordered regions has far-reaching consequences regarding the applicability of ensemble calculation: Since smFRET is not limited by the size of the protein, nor any dynamic time scale sampled by the protein, FRET efficiencies can also be measured under conditions where NMR line broadening leads to factual disappearance of the signal.[33,63] Furthermore, the low protein concentrations used in an smFRET experiment (in the picomolar range) also allow accessing aggregation-prone proteins[54,64] or performing experiments within the cell under physiological conditions.[10,11] Using FRET efficiencies for the calculation of conformational ensembles thus allows addressing the conformational land-

https://doi.org/10.1021/jacs.1c06264
J. Am. Chem. Soc. 2021, 143, 20109−20121

scape of IDPs under conditions that are not accessible by any other technique.

CONCLUSION

With the integrated use of NMR, SAXS, and single-molecule FRET to calculate multiconformational models that satisfy all data, we now demonstrate how different experimental techniques can synergize to reliably describe IDPs, and we demonstrated this on the example of the measles virus phosphoprotein. With the increasing awareness of the importance of IDPs, in particular also in liquid−liquid phase separation,[65,66] we expect this tool-set to make an important impact in integrative multiconformational modeling of dynamic systems.

MATERIALS AND METHODS

Accessible Volume Calculations. AV calculations were based on procedures described previously.[12,36] Briefly, positions that the dyes are expected to sample were calculated considering a linker length, as well as three radii (R_1, R_2, R_3). Pairwise distances between the positions sampled by the donor and the acceptor fluorophore were then calculated with a coarsening step size of 200 with respect to the position list. Distance histograms as well as average FRET efficiencies were compared over an ensemble of 200 conformers using a step size of 10, 50, and 200.

For the calculation of FRET efficiencies on large conformational ensembles, the calculation speed of the AV had to be optimized: Positions describing the accessible volumes were sampled in an iterative way. A total of 100, 500, 1000, and 10 000 iteration steps were tested for reproducibility of distance histograms and average FRET efficiencies over a 100 conformer ensemble. A total of 500 iterations led to sufficiently accurate distance histograms that reproduce FRET efficiencies reliably.

In order to avoid "mutating" amino acids into cysteines *in silico*, AVs were calculated from the CB atom of the respective amino acid. The linker length in the simulations was optimized to take the distance between CB and SH of a cysteine into account. The estimate (L = 22.83 Å) was based on geometrical considerations, and an ensemble by which the AV was calculated from the CB as attachment point has been verified to reproduce the distance histograms and FRET efficiencies calculated from the same ensemble, but with SH as attachment point (L = 21 Å).

The following dye parametrization was used in the calculation:

Table 1

	Alexa488	Alexa594	Alexa647
L	22.83	22.83	22.83
R_1	6.8	7.6	11
R_2	3.9	4.1	4.7
R_3	1.5	2.2	1.5

Scripts provided by Walczewska-Szewc et al.[36] have been adapted to contain the changes above. Attachment points were read from the respective PDB files of the conformational ensemble in an automated way using in-house software and were then used for the calculation of AVs and distance histograms. Parametrization for Alexa647 was adapted from Peter et al.[67] PDB files containing full side chains were used for AV calculation. Conformer-wise FRET files were then generated as an input for ASTEROIDS[27,28] selection, containing the different FRET distances used in the selection.

Incorporation of FRET Efficiencies into Multiconformational Models. AVs were calculated per conformer as described. Pairwise distances between the sampled volumes of the two fluorophores are calculated and converted into FRET efficiencies according to

$$\varepsilon(r) = \frac{1}{1 + \left(\frac{r}{R_0}\right)^6} \tag{5}$$

with the Förster distance R_0 and the distance r between the sampled points in the AV. The average FRET efficiency of one conformer is then calculated as the average of ε over all pairwise distances n:

$$\langle E_{\text{conf}} \rangle = \left(\sum_{i=1}^{n} \varepsilon(r_i) \right) \frac{1}{n} \tag{6}$$

in accordance with a sampling of the AV on a time scale significantly longer than the fluorescence lifetime. The average FRET efficiency of the ensemble comprising all conformers (which is then compared to the measured E_{FRET}) can be described as

$$\langle E_{\text{ens}} \rangle = \int_0^\infty \varepsilon(r)\, P(r)\, \mathrm{d}r \tag{7}$$

with $\varepsilon(r)$ as described in eq 5 and $P(r)$ describing the distance distribution containing all pairwise distances of the AVs for every conformer. Computationally, for an ensemble of m conformers, $\langle E_{\text{ens}} \rangle$ can be calculated as

$$\langle E_{\text{ens}} \rangle = \left(\sum_{i=1}^{m} \langle E^i_{\text{conf}} \rangle \right) \frac{1}{m} \tag{8}$$

with $\langle E^i_{\text{conf}} \rangle$ describing the average FRET efficiency of the ith member of the ensemble as described in eq 6.

For considerations assuming a sampling of the AV that is much faster than the fluorescence lifetime, pairwise positions of the fluorophores were first determined and their sixth power was calculated and then averaged per conformer.[68,69] The FRET efficiency was calculated from these averaged distances on a conformer-by-conformer basis:

$$E^{\text{fast}}_{\text{conf}} = \frac{1}{1 + \frac{\langle r^6 \rangle}{R_0^6}} \tag{9}$$

R_0 used in the calculations was determined experimentally. The quantum yield of P_{1-100} labeled with Alexa488 was determined by the comparative method[70] described previously with fluorescein (in 0.1 M NaOH, Φ = 0.95, n = 1.334)[71] and Rhodamine 6G (in ethanol, Φ = 0.94, n = 1.361)[72] as quantum yield standards. The overlap integral $J(\lambda)$ was determined from P_{1-100} samples labeled with Alexa488 and Alexa594.[46] Rapid orientation averaging of the fluorescent dyes was assumed, leading to the common assumption of $\kappa^2 = {}^2/_3$. Fluorescence anisotropies measured on the different P_{1-100} samples suggested that this assumption was valid (SI Table 2). R_0 was then calculated according to[46]

$$R_0 = \left[\frac{9 \ln(10)}{128 \pi^5 N_A} \frac{J(\lambda) \kappa^2 \phi_D}{n^4} \right]^{1/6} \tag{10}$$

with the Avogadro number N_A, the overlap integral $J(\lambda)$, the orientation factor κ^2, the quantum yield of the donor in the absence of acceptor Φ_D, and the refractive index n. n = 1.3 was used for P_{1-100} in its measurement buffer. An R_0 of 56 Å was obtained for the dye pair Alexa488/Alexa594 in 50 mM Na-phosphate pH 6, 150 mM NaCl, and 2 mM dithiothreitol (DTT). The same R_0 was used to compute the *in silico* data set.

Generation of *in Silico* Ensemble. Flexible-meccano[29] was used to generate a large conformational ensemble (10 000 conformers) of a 155 amino acid long IDP. The centers of mass of all C_α atoms from residues 15−25 as well as residues 90−100 were calculated, and all conformers with a distance of less than 20 Å between these centers of mass were selected. AVs of Alexa488 and Alexa594 were computed as described above for 15 *in silico* "labeling positions" (SI Figure 2), and FRET efficiencies for this ensemble comprising a long-range contact were calculated as described above and used as an input for

https://doi.org/10.1021/jacs.1c06264
J. Am. Chem. Soc. 2021, 143, 20109−20121

ASTEROIDS or for cross-validation of ensembles selected using ASTEROIDS. Expected PREs for this ensemble were calculated as described elsewhere (labeling sites were residues 23, 65, 70, 92, and 130).[39] τ_c and τ_e were set to 5 and 0.5 ns, respectively. $\tau_C = \tau_r\tau_s/(\tau_r + \tau_s)$ describes the rotational correlation time of the protein (τ_r) and the effective electron relaxation time (τ_s), and $\tau_e = 1/(\tau_i^{-1} + \tau_r^{-1} + \tau_s^{-1})$ depends on the effective correlation time of the spin label (τ_i) according to a model-free expression of the spectral density function.[38,39] ^{1}H R_2 relaxation was assumed to be 18 s^{-1} throughout the protein.

Selection of Conformational Ensembles. Ensembles of 200 conformers were selected from a large statistical coil ensemble (10 000 conformers), generated through flexible-meccano,[29] using the genetic algorithm ASTEROIDS.[27] Selection based on PREs and chemical shifts was performed as described previously.[39] Selection based on FRET efficiencies allowed an error of 0.02 and was weighed 50% as compared to an NMR experiment (e.g., all PREs arising from one spin labeling site).

For P_{1-100}, ensembles were first selected based on only chemical shifts during four iterations of flexible-meccano/ASTEROIDS. A large conformational ensemble (10 000 conformers) was then calculated based on the resulting Φ/Ψ angles, from which subensembles were selected based on FRET, PREs, and chemical shifts. FRET efficiencies not used in the selection were back-calculated as described above. SAXS curves were back-calculated using CRYSOL.[73] Chemical shifts were calculated using SPARTA.[74]

Back-Calculation of Fluorescence Lifetimes. Distance distributions between the donor and acceptor fluorophores were calculated from the selected conformational ensembles, and the corresponding fluorescence lifetime decays were calculated as[23]

$$I(t) = \mathrm{IRF} \otimes \int P(r) \mathrm{e}^{t/\tau_{D(FRET)}(r)}\, \mathrm{d}r \tag{11}$$

with the instrument response function (IRF) experimentally determined and described by a double Gaussian function, and the fluorescence lifetime of the donor in the presence of the acceptor $\tau_{D(FRET)}$ calculated for every distance r according to

$$\tau_{D(FRET)}(r) = \tau_D\left(1 - \frac{1}{1 + \left(\frac{r}{R_0}\right)^6}\right) \tag{12}$$

The fluorescence lifetime of the donor in the absence of the acceptor (τ_D) and the Förster distance (R_0) were experimentally determined. A scattering contribution was added to the fluorescence lifetime decays, and both decay and scattering were scaled independently to best fit the experimental data.

Protein Production. P_{1-100} tagged with 8 His was expressed and purified as described earlier.[1,40] Briefly, a pET41c(+) plasmid containing P_{1-100} was transformed into Rosetta (λDE3)/pRARE (Novagen), and cultures were grown at 37 °C in lysogeny broth (LB) medium until an optical density (OD) of >0.6. Expression was induced with 1 mM isopropyl-β-D-thiogalactopyranoside and continued at 20 °C overnight. Cells were lysed by sonication in 20 mM Tris pH 8/150 mM NaCl and purified using standard Ni purification. The protein was eluted from the Ni resin by adding 400 mM imidazole to the lysis buffer. The protein was then further purified on a Superdex 75 column (GE Healthcare) in 50 mM Na-phosphate, pH 6, 150 mM NaCl, and 2 mM DTT. Expression of protein labeled with ^{15}N followed the same procedure, except that the protein was expressed in M9 minimal medium. All experiments were conducted in 50 mM Na-phosphate pH 6, 150 mM NaCl, and 2 mM DTT. DTT was not contained in buffers used for PRE experiments.

Protein Labeling with Fluorophores or Spin Radical Labels. P_{1-100} was randomly labeled with Alexa488 and Alexa594 essentially as described previously.[8,75] Briefly, 20 mM DTT was added to the protein sample and incubated overnight at 4 °C. The protein was then dialyzed into degassed 50 mM Na-phosphate pH 7 and 150 mM NaCl buffer until all DTT was washed out. Alexa488 and Alexa594 were added simultaneously at an excess of approximately 5× compared to protein. Labeling was allowed to proceed 30 min at room temperature, followed by 4 °C overnight. The labeled protein was then separated from excess dye by size exclusion chromatography on an Enrich SEC70 (Biorad) column using 50 mM Na-phosphate buffer (pH 6), 150 mM NaCl, and 2 mM DTT.

Labeling of ^{15}N P_{1-100} single cysteine mutants for PREs was achieved using *S*-(1-oxyl-2,2,5,5-tetramethyl-2,5-dihydro-1*H*-pyrrol-3-yl)methylmethanesulfonothioate (MTSL) and followed essentially the same procedure as for fluorescence labeling. The final buffer used for size exclusion chromatography, however, did not contain DTT.

Experimental NMR Data. All NMR experiments were performed at a temperature of 19 °C. The assignment of P_{1-100} was obtained previously[1] and used as an input for ASTEROIDS[39] selection as well as for calculation of secondary chemical shifts and secondary structure propensities[76] (SSPs).

For calculation of PREs, HSQC spectra of the different Cys mutants of ^{15}N P_{1-100} were measured in the presence and absence of the MTSL label. Spectra were processed with NMRPipe,[77] peak intensities were extracted from the respective spectra, and the ratio between MTSL labeled and unlabeled peak intensities was determined and used as an input for ASTEROIDS.

Experimental Single-Molecule FRET data. Single-molecule fluorescence spectroscopy was measured on a custom setup built around an Olympus IX73 microscope equipped with a 60× water immersion objective (NA 1.2). A pulsed laser diode (40 MHz, LDH 485, Picoquant, Berlin, Germany) was fed through a λ/4 plate and focused onto the sample to excite freely diffusing P_{1-100} molecules with circularly polarized light. Fluorescence emission was spatially filtered through a pinhole with a 100 μm diameter, separated into green (Alexa488) and orange (Alexa594) fluorescence, and focused onto two PMA hybrid detectors (Picoquant). Photons were counted using a Hydraharp (Picoquant). smFRET experiments were performed at room temperature.

FRET histograms were calculated using custom code written in Python. Lists of photon arrival times were first extracted using a code written in C, adapted from a demo-code provided by Picoquant.[78] Photon streams were then binned with a 1 ms bin width and subjected to a Lee filter before bust integration and thresholding.[79] A threshold of at least 50 photons was used. Fluorescence intensities were corrected for background contribution, spectral crosstalk, differences in quantum yield (determined as described previously[70]), and differences in detection efficiencies between the green and the orange channel.

Microtimes were extracted for bursts corresponding to the FRET peak and to the 0-FRET peak separately, and population averaged lifetime histograms were built. The instrument response function was measured on buffer under the same conditions as the single-molecule experiments, and lifetimes of the donor were extracted through fitting the lifetime histograms of the 0-FRET population with a single-exponential function convolved with the IRF.

Corrections Employed in the smFRET Experiments. Buffer background was measured using the same conditions as for the single molecule experiments, and bin-wise background contributions were determined for the donor and acceptor channel and subtracted from the bin-wise photon counts in the single-molecule FRET experiments.

Differences in detection efficiencies and quantum yields were included in the correction factor γ (see eq 3:

$$\gamma = \gamma_{\mathrm{instrument}} \frac{\Phi_{Ac}}{\Phi_{Do}} \tag{13}$$

with $\gamma_{\mathrm{instrument}} = \frac{\eta_{Ac}}{\eta_{Do}}$ being the difference in detection efficiency between acceptor (η_{Ac}) and donor (η_{Do}) signal of the instrument determined as described in Ferreon et al., 2009.[80] Briefly, fluorescence of free donor and acceptor dyes in the measurement buffer was measured on an ensemble fluorescence spectrometer (PTI Quantamaster) and on the single-molecule fluorescence setup at the same excitation wavelength. Ensemble fluorescence spectra were corrected for detection differences at different wavelengths, and the total signal

https://doi.org/10.1021/jacs.1c06264
J. Am. Chem. Soc. 2021, 143, 20109−20121

was extrapolated to the full emission spectra. Plots displaying the integrated ensemble fluorescence versus fluorescence recorded on the single-molecule setup were fitted with a line for donor and acceptor fluorescence independently. The ratio of the slopes (m_{Ac}/m_{Do}) was determined to be $\gamma_{instrument}$ and is 0.81 for the Alexa488/Alexa594 dye pair and 0.83 for the Alexa488/Alexa647 dye pair in 50 mM Na-phosphate pH 6, 150 mM NaCl, and 2 mM DTT. The spectral properties of fluorescently labeled P_{1-100} were equal to those of the free dyes in the same buffer. $\gamma_{instrument}$ was corrected on a daily basis based on a short measurement of Rhodamine 6G,[46] which emits into the donor and acceptor channel of the smFRET setup.

Fluorescence quantum yields of the donor (Φ_{Do}) and acceptor (Φ_{Ac}) were determined from singly labeled P_{1-100} proteins in the measurement buffer using the comparative method described by Williams et al. as described above.[70] Rhodamine 101 (in ethanol, Φ = 1.0, n = 1.36)[81] was used as a quantum yield standard for Alexa594-labeled proteins (see section Incorporation of FRET Efficiencies into Multiconformational Models for standards used for Alexa488-labeled proteins). For Alexa647-labeled proteins (SI Figure 12), cresyl violet (in ethanol rather than methanol, Φ = 0.54, n = 1.33)[82] was added as an additional quantum yield standard. A refractive index of n = 1.3 was used for all P_{1-100} samples. The quantum yields determined for the different P_{1-100} single cysteine constructs were very similar, and their average quantum yields were thus used both for γ correction and for the calculation of the Förster distance (R_0).

Leakage was determined from measurements undertaken in the context of γ correction by calculating the ratio of donor fluorescence arriving in the acceptor versus the donor channel of smFRET setup. These values were corrected on a daily basis using the Rhodamine 6G calibration measurement and validated by ensuring that the donor-only peak of the single-molecule FRET histograms was situated at a FRET efficiency of 0.

To estimate the contribution of direct excitation, an IDP sample labeled with Alexa488 and Alexa594 separated by 164 amino acids was prepared, which is not expected to yield $E_{FRET} > 0$.[46] While we cannot entirely exclude that this is indeed not the case, the contribution of direct excitation was tentatively attributed to be 0.2 photon per 1 ms under this assumption. Since application of this correction yields differences in E_{FRET} of only around 0.01 to 0.03, we decided not to apply this correction. This remains true if the ratio of extinction coefficients between the donor and acceptor at the excitation wavelength is used to correct for direct excitation.[83,84] In order to test the validity of this approximation, the DNA sample "4-mid", labeled with Atto488/Atto594 used in Hellenkamp et al., 2018,[45] has been measured and corrected using the same procedure (SI Figure 12B, γ was determined independently, and quantum yields as well as R_0 were used as described by atto-tec[85]), leading to E_{FRET} = 0.39 compared to 0.41 ± 0.04 as reported by Hellenkamp et al.[45]

Experimental SAXS Data. SAXS experiments were measured for five different concentrations of P_{1-100} from 0.5 to 2 mg/mL at 20 °C on BM29 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. Scattering was measured at a wavelength of 0.992 Å, and samples were exposed during 10 frames. Frames not impacted by radiation damage were averaged. Buffer scattering curves were subtracted from the scattering curves of P_{1-100}.

Theoretical Comparison between FRET and PRE Rates. Figure 3A–C were generated considering a static measured distance for both FRET (k_{ET}) and PRE rates ($R_{2,PRE}$).

$$k_{ET} = \left(\frac{R_0}{r}\right)^6 \frac{1}{\tau_D} \tag{14}$$

was calculated with a Förster distance (R_0) of 56 Å and a fluorescence lifetime of the donor (τ_D) of 4 ns. E_{FRET} was calculated from k_{ET} as described in eq 2. r is the distance between donor and acceptor fluorophores.[46]

The PRE transverse relaxation rate ($R_{2,PRE}$) was calculated according to

$$R_{2,PRE} = \frac{1}{15}\left(\frac{\mu_0}{4\pi}\right)^2 \gamma_H^2 g_e^2 \mu_B^2 s_e(s_e + 1)[4J(0) + 3J(\omega_H)] \tag{15}$$

with the electron g-factor g_e, the gyromagnetic ratio of the observed proton γ_H, the electron spin s_e, the Bohr magneton μ_B, and the proton frequency ω_H.

A spectral density function of

$$J(\omega) = r^{-6}\frac{\tau_c}{1 + (\omega\tau_c)^2} \tag{16}$$

was used with $\tau_c = \tau_r\tau_s/(\tau_r + \tau_s)$, τ_r being the rotational correlation time of the protein and τ_s the effective electron relaxation time. τ_c was set to 5 ns for Figure 3C. r is the distance between the ^1H_N nuclei and the PRE label.[38,39]

Note that for the calculation of PREs in the context of a conformational ensemble of an IDP, a model-free expression of the spectral density function was used, describing the internal motion of the IDP as well as the motion of the spin label:

$$J(\omega) = r^{-6}\left[\frac{S^2\tau_c}{1 + (\omega\tau_c)^2} + \frac{(1 - S^2)\tau_e}{1 + (\omega\tau_e)^2}\right] \tag{17}$$

The order parameter S^2 denotes the motion of the dipolar interaction vector, τ_c is as described above, and $\tau_e = 1/(\tau_i^{-1} + \tau_r^{-1} + \tau_s^{-1})$ additionally depends on the effective correlation time of the spin label (τ_i).[39]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.1c06264.

Additional figures showing a comparison between the *in silico* attachment points of accessible volumes, an overview of the protein constructs used for the *in silico* data set, ASTEROIDS selection using only *in silico* FRET or PREs, different numbers of FRET efficiencies, different numbers of conformers, and different allowed errors, comparison of different dye averaging regimes, FRET histograms of P_{1-100}, description of fluorescence lifetimes by an ASTEROIDS ensemble of P_{1-100}, validation samples (DNA and P_{1-100} labeled with a different dye pair), fits of P_{1-100} SAXS data, as well as two tables summarizing the experimental fluorescence data and FRET efficiencies obtained from selections (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Sigrid Milles − *Institut de Biologie Structurale, Université Grenoble Alpes-CEA-CNRS, 38044 Grenoble, France;* orcid.org/0000-0001-9362-9606; Email: sigrid.milles@ibs.fr

Authors

Samuel Naudi-Fabra − *Institut de Biologie Structurale, Université Grenoble Alpes-CEA-CNRS, 38044 Grenoble, France*

Maud Tengo − *Institut de Biologie Structurale, Université Grenoble Alpes-CEA-CNRS, 38044 Grenoble, France*

Malene Ringkjøbing Jensen − *Institut de Biologie Structurale, Université Grenoble Alpes-CEA-CNRS, 38044 Grenoble, France;* orcid.org/0000-0003-0419-2196

Martin Blackledge − *Institut de Biologie Structurale, Université Grenoble Alpes-CEA-CNRS, 38044 Grenoble, France;* orcid.org/0000-0003-0935-721X

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.1c06264

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This project has received funding from the European Research Council (ERC) Starting Grant MultiMotif to S.M. under the European Union's Horizon 2020 research and innovation program (grant agreement no. 802209). We also acknowledge funding from the French National Research Agency (ANR) through an ANR T-ERC MultiMotif (ANR-17-ERC3-0004) and through the framework of the "Investissements d'Avenir" program (ANR-15-IDEX-02), DisRegulate, as well as from the Commissariat à l'Énergie Atomique et aux Énergies Alternatives (CEA) though DRF impulsions (dynamicFRET) to S.M. This work used the platforms of the Grenoble Instruct-ERIC center (ISBG ; UAR 3518 CNRS-CEA-UGA-EMBL) within the Grenoble Partnership for Structural Biology (PSB), supported by FRISBI (ANR-10-INBS-0005-02) and GRAL, financed within the University Grenoble Alpes graduate school (Ecoles Universitaires de Recherche) CBH-EUR-GS (ANR-17-EURE-0003). We thank Caroline Mas for access to the Biophysical platform. The SAXS experiments were performed on beamline BM29 at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. We are grateful to local contact Petra Pernod for providing assistance in using beamline BM29. The Institut de Biologie Structurale acknowledges integration into the Interdisciplinary Research Institute of Grenoble.

■ REFERENCES

(1) Milles, S.; Jensen, M. R.; Lazert, C.; Guseva, S.; Ivashchenko, S.; Communie, G.; Maurin, D.; Gerlier, D.; Ruigrok, R. W. H.; Blackledge, M. An Ultraweak Interaction in the Intrinsically Disordered Replication Machinery Is Essential for Measles Virus Function. *Sci. Adv.* **2018**, *4* (8), eaat7778.

(2) Schneider, R.; Maurin, D.; Communie, G.; Kragelj, J.; Hansen, D. F.; Ruigrok, R. W. H.; Jensen, M. R.; Blackledge, M. Visualizing the Molecular Recognition Trajectory of an Intrinsically Disordered Protein Using Multinuclear Relaxation Dispersion NMR. *J. Am. Chem. Soc.* **2015**, *137* (3), 1220−1229.

(3) Bugge, K.; Brakti, I.; Fernandes, C. B.; Dreier, J. E.; Lundsgaard, J. E.; Olsen, J. G.; Skriver, K.; Kragelund, B. B. Interactions by Disorder − A Matter of Context. *Frontiers in Molecular Biosciences* **2020**, *7*, 110.

(4) Müller-Späth, S.; Soranno, A.; Hirschfeld, V.; Hofmann, H.; Rüegger, S.; Reymond, L.; Nettels, D.; Schuler, B. From the Cover: Charge Interactions Can Dominate the Dimensions of Intrinsically Disordered Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (33), 14609−14614.

(5) Mao, A. H.; Crick, S. L.; Vitalis, A.; Chicoine, C. L.; Pappu, R. V. Net Charge per Residue Modulates Conformational Ensembles of Intrinsically Disordered Proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (18), 8183−8188.

(6) Metskas, L. A.; Rhoades, E. Single-Molecule FRET of Intrinsically Disordered Proteins. *Annu. Rev. Phys. Chem.* **2020**, *71* (1), 391−414.

(7) Schuler, B.; Eaton, W. A. Protein Folding Studied by Single-Molecule FRET. *Curr. Opin. Struct. Biol.* **2008**, *18* (1), 16−26.

(8) Milles, S.; Lemke, E. A. Single Molecule Study of the Intrinsically Disordered FG-Repeat Nucleoporin 153. *Biophys. J.* **2011**, *101* (7), 1710−1719.

(9) Borgia, A.; Borgia, M. B.; Bugge, K.; Kissling, V. M.; Heidarsson, P. O.; Fernandes, C. B.; Sottini, A.; Soranno, A.; Buholzer, K. J.; Nettels, D.; Kragelund, B. B.; Best, R. B.; Schuler, B. Extreme Disorder in an Ultrahigh-Affinity Protein Complex. *Nature* **2018**, *555* (7694), 61−66.

(10) König, I.; Zarrine-Afsar, A.; Aznauryan, M.; Soranno, A.; Wunderlich, B.; Dingfelder, F.; Stüber, J. C.; Plückthun, A.; Nettels, D.; Schuler, B. Single-Molecule Spectroscopy of Protein Conformational Dynamics in Live Eukaryotic Cells. *Nat. Methods* **2015**, *12* (8), 773−779.

(11) Sakon, J. J.; Weninger, K. R. Detecting the Conformation of Individual Proteins in Live Cells. *Nat. Methods* **2010**, *7* (3), 203−205.

(12) Kalinin, S.; Peulen, T.; Sindbert, S.; Rothwell, P. J.; Berger, S.; Restle, T.; Goody, R. S.; Gohlke, H.; Seidel, C. A. M. A Toolkit and Benchmark Study for FRET-Restrained High-Precision Structural Modeling. *Nat. Methods* **2012**, *9* (12), 1218−1225.

(13) Hellenkamp, B.; Wortmann, P.; Kandzia, F.; Zacharias, M.; Hugel, T. Multidomain Structure and Correlated Dynamics Determined by Self-Consistent FRET Networks. *Nat. Methods* **2017**, *14* (2), 174−180.

(14) Muschielok, A.; Andrecka, J.; Jawhari, A.; Brückner, F.; Cramer, P.; Michaelis, J. A Nano-Positioning System for Macromolecular Structural Analysis. *Nat. Methods* **2008**, *5* (11), 965−971.

(15) Dimura, M.; Peulen, T.-O.; Sanabria, H.; Rodnin, D.; Hemmen, K.; Hanke, C. A.; Seidel, C. A. M.; Gohlke, H. Automated and Optimally FRET-Assisted Structural Modeling. *Nat. Commun.* **2020**, *11* (1), 5394.

(16) Schuler, B.; Soranno, A.; Hofmann, H.; Nettels, D. Single-Molecule FRET Spectroscopy and the Polymer Physics of Unfolded and Intrinsically Disordered Proteins. *Annu. Rev. Biophys.* **2016**, *45*, 207−231.

(17) Aznauryan, M.; Delgado, L.; Soranno, A.; Nettels, D.; Huang, J.-R.; Labhardt, A. M.; 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* (37), E5389−5398.

(18) McCarney, E. R.; Werner, J. H.; Bernstein, S. L.; Ruczinski, I.; Makarov, D. E.; Goodwin, P. M.; Plaxco, K. W. Site-Specific Dimensions across a Highly Denatured Protein; a Single Molecule Study. *J. Mol. Biol.* **2005**, *352* (3), 672−682.

(19) Fuertes, G.; Banterle, N.; Ruff, K. M.; Chowdhury, A.; Mercadante, D.; Koehler, C.; Kachala, M.; Estrada Girona, G.; Milles, S.; Mishra, A.; Onck, P. R.; Gräter, F.; Esteban-Martín, S.; Pappu, R. V.; Svergun, D. I.; Lemke, E. A. Decoupling of Size and Shape Fluctuations in Heteropolymeric Sequences Reconciles Discrepancies in SAXS vs. FRET Measurements. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114* (31), E6342−E6351.

(20) Yoo, T. Y.; Meisburger, S. P.; Hinshaw, J.; Pollack, L.; Haran, G.; Sosnick, T. R.; Plaxco, K. Small-Angle X-Ray Scattering and Single-Molecule FRET Spectroscopy Produce Highly Divergent Views of the Low-Denaturant Unfolded State. *J. Mol. Biol.* **2012**, *418* (3−4), 226−236.

(21) Watkins, H. M.; Simon, A. J.; Sosnick, T. R.; Lipman, E. A.; Hjelm, R. P.; Plaxco, K. W. Random Coil Negative Control Reproduces the Discrepancy between Scattering and FRET Measurements of Denatured Protein Dimensions. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (21), 6631−6636.

(22) Sherman, E.; Haran, G. Coil-Globule Transition in the Denatured State of a Small Protein. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (31), 11539−11543.

(23) Borgia, A.; Zheng, W.; Buholzer, K.; Borgia, M. B.; Schüler, A.; Hofmann, H.; Soranno, A.; Nettels, D.; Gast, K.; Grishaev, A.; Best, R. B.; Schuler, B. Consistent View of Polypeptide Chain Expansion in Chemical Denaturants from Multiple Experimental Methods. *J. Am. Chem. Soc.* **2016**, *138* (36), 11714−11726.

(24) Riback, J. A.; Bowman, M. A.; Zmyslowski, A. M.; Knoverek, C. R.; Jumper, J. M.; Hinshaw, J. R.; Kaye, E. B.; Freed, K. F.; Clark, P. L.; Sosnick, T. R. Innovative Scattering Analysis Shows That Hydrophobic Disordered Proteins Are Expanded in Water. *Science* **2017**, *358* (6360), 238−241.

(25) Best, R. B. Computational and Theoretical Advances in Studies of Intrinsically Disordered Proteins. *Curr. Opin. Struct. Biol.* **2017**, *42*, 147−154.
(26) Gomes, G.-N. W.; Krzeminski, M.; Namini, A.; Martin, E. W.; Mittag, T.; Head-Gordon, T.; Forman-Kay, J. D.; Gradinaru, C. C. Conformational Ensembles of an Intrinsically Disordered Protein Consistent with NMR, SAXS, and Single-Molecule FRET. *J. Am. Chem. Soc.* **2020**, *142* (37), 15697−15710.
(27) Nodet, G.; Salmon, L.; Ozenne, V.; Meier, S.; Jensen, M. R.; Blackledge, M. Quantitative Description of Backbone Conformational Sampling of Unfolded Proteins at Amino Acid Resolution from NMR Residual Dipolar Couplings. *J. Am. Chem. Soc.* **2009**, *131* (49), 17908−17918.
(28) Jensen, M. R.; Salmon, L.; Nodet, G.; Blackledge, M. Defining Conformational Ensembles of Intrinsically Disordered and Partially Folded Proteins Directly from Chemical Shifts. *J. Am. Chem. Soc.* **2010**, *132* (4), 1270−1272.
(29) Ozenne, V.; Bauer, F.; Salmon, L.; Huang, J.-R.; Jensen, M. R.; Segard, S.; Bernadó, P.; Charavay, C.; Blackledge, M. Flexible-Meccano: A Tool for the Generation of Explicit Ensemble Descriptions of Intrinsically Disordered Proteins and Their Associated Experimental Observables. *Bioinformatics* **2012**, *28* (11), 1463−1470.
(30) Bernadó, P.; Blackledge, M. A Self-Consistent Description of the Conformational Behavior of Chemically Denatured Proteins from NMR and Small Angle Scattering. *Biophys. J.* **2009**, *97* (10), 2839−2845.
(31) Schwalbe, M.; Ozenne, V.; Bibow, S.; Jaremko, M.; Jaremko, L.; Gajda, M.; Jensen, M. R.; Biernat, J.; Becker, S.; Mandelkow, E.; Zweckstetter, M.; Blackledge, M. Predictive Atomic Resolution Descriptions of Intrinsically Disordered HTau40 and α-Synuclein in Solution from NMR and Small Angle Scattering. *Structure* **2014**, *22* (2), 238−249.
(32) Whitley, M. J.; Xi, Z.; Bartko, J. C.; Jensen, M. R.; Blackledge, M.; Gronenborn, A. M. A Combined NMR and SAXS Analysis of the Partially Folded Cataract-Associated V75D ΓD-Crystallin. *Biophys. J.* **2017**, *112* (6), 1135−1146.
(33) Delaforge, E.; Milles, S.; Bouvignies, G.; Bouvier, D.; Boivin, S.; Salvi, N.; Maurin, D.; Martel, A.; Round, A.; Lemke, E. A.; Ringkjøbing Jensen, M.; Hart, D. J.; Blackledge, M. Large-Scale Conformational Dynamics Control H5N1 Influenza Polymerase PB2 Binding to Importin α. *J. Am. Chem. Soc.* **2015**, *137* (48), 15122−15134.
(34) Jensen, M. R.; Yabukarski, F.; Communie, G.; Condamine, E.; Mas, C.; Volchkova, V.; Tarbouriech, N.; Bourhis, J.-M.; Volchkov, V.; Blackledge, M.; Jamin, M. Structural Description of the Nipah Virus Phosphoprotein and Its Interaction with STAT1. *Biophys. J.* **2020**, *118* (10), 2470−2488.
(35) Jensen, M. R.; Zweckstetter, M.; Huang, J.; Blackledge, M. Exploring Free-Energy Landscapes of Intrinsically Disordered Proteins at Atomic Resolution Using NMR Spectroscopy. *Chem. Rev.* **2014**, *114* (13), 6632−6660.
(36) 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* (24), 12317−12326.
(37) Sindbert, S.; Kalinin, S.; Nguyen, H.; Kienzler, A.; Clima, L.; Bannwarth, W.; Appel, B.; Müller, S.; Seidel, C. A. M. Accurate Distance Determination of Nucleic Acids via Förster Resonance Energy Transfer: Implications of Dye Linker Length and Rigidity. *J. Am. Chem. Soc.* **2011**, *133* (8), 2463−2480.
(38) Clore, G. M.; Iwahara, J. Theory, Practice, and Applications of Paramagnetic Relaxation Enhancement for the Characterization of Transient Low-Population States of Biological Macromolecules and Their Complexes. *Chem. Rev.* **2009**, *109* (9), 4108−4139.
(39) Salmon, L.; Nodet, G.; Ozenne, V.; Yin, G.; Jensen, M. R.; Zweckstetter, M.; Blackledge, M. NMR Characterization of Long-Range Order in Intrinsically Disordered Proteins. *J. Am. Chem. Soc.* **2010**, *132* (24), 8407−8418.
(40) Milles, S.; Jensen, M. R.; Communie, G.; Maurin, D.; Schoehn, G.; Ruigrok, R. W. H.; Blackledge, M. Self-Assembly of Measles Virus Nucleocapsid-like Particles: Kinetics and RNA Sequence Dependence. *Angew. Chem., Int. Ed.* **2016**, *55* (32), 9356−9360.
(41) Zheng, W.; Best, R. B. An Extended Guinier Analysis for Intrinsically Disordered Proteins. *J. Mol. Biol.* **2018**, *430* (16), 2540−2553.
(42) Hofmann, H.; Soranno, A.; Borgia, A.; Gast, K.; Nettels, D.; Schuler, B. Polymer Scaling Laws of Unfolded and Intrinsically Disordered Proteins Quantified with Single-Molecule Spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (40), 16155−16160.
(43) Peulen, T.-O.; Opanasyuk, O.; Seidel, C. A. M. Combining Graphical and Analytical Methods with Molecular Simulations To Analyze Time-Resolved FRET Measurements of Labeled Macromolecules Accurately. *J. Phys. Chem. B* **2017**, *121* (35), 8211−8241.
(44) Dingfelder, F.; Benke, S.; Nettels, D.; Schuler, B. Mapping an Equilibrium Folding Intermediate of the Cytolytic Pore Toxin ClyA with Single-Molecule FRET. *J. Phys. Chem. B* **2018**, *122* (49), 11251−11261.
(45) Hellenkamp, B.; Schmid, S.; Doroshenko, O.; Opanasyuk, O.; Kühnemuth, R.; Rezaei Adariani, S.; Ambrose, B.; Aznauryan, M.; Barth, A.; Birkedal, V.; Bowen, M. E.; Chen, H.; Cordes, T.; Eilert, T.; Fijen, C.; Gebhardt, C.; Götz, M.; Gouridis, G.; Gratton, E.; Ha, T.; Hao, P.; Hanke, C. A.; Hartmann, A.; Hendrix, J.; Hildebrandt, L. L.; Hirschfeld, V.; Hohlbein, J.; Hua, B.; Hübner, C. G.; Kallis, E.; Kapanidis, A. N.; Kim, J.-Y.; Krainer, G.; Lamb, D. C.; Lee, N. K.; Lemke, E. A.; Levesque, B.; Levitus, M.; McCann, J. J.; Naredi-Rainer, N.; Nettels, D.; Ngo, T.; Qiu, R.; Robb, N. C.; Röcker, C.; Sanabria, H.; Schlierf, M.; Schröder, T.; Schuler, B.; Seidel, H.; Streit, L.; Thurn, J.; Tinnefeld, P.; Tyagi, S.; Vandenberk, N.; Vera, A. M.; Weninger, K. R.; Wünsch, B.; Yanez-Orozco, I. S.; Michaelis, J.; Seidel, C. A. M.; Craggs, T. D.; Hugel, T. Precision and Accuracy of Single-Molecule FRET Measurements-a Multi-Laboratory Benchmark Study. *Nat. Methods* **2018**, *15* (9), 669−676.
(46) Sisamakis, E.; Valeri, A.; Kalinin, S.; Rothwell, P. J.; Seidel, C. A. M. Accurate Single-Molecule FRET Studies Using Multiparameter Fluorescence Detection. *Methods Enzymol.* **2010**, *475*, 455−514.
(47) Gopich, I. V.; Szabo, A. Theory of the Energy Transfer Efficiency and Fluorescence Lifetime Distribution in Single-Molecule FRET. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (20), 7747−7752.
(48) Kalinin, S.; Valeri, A.; Antonik, M.; Felekyan, S.; Seidel, C. A. M. Detection of Structural Dynamics by FRET: A Photon Distribution and Fluorescence Lifetime Analysis of Systems with Multiple States. *J. Phys. Chem. B* **2010**, *114* (23), 7983−7995.
(49) Delaforge, E.; Kragelj, J.; Tengo, L.; Palencia, A.; Milles, S.; Bouvignies, G.; Salvi, N.; Blackledge, M.; Jensen, M. R. Deciphering the Dynamic Interaction Profile of an Intrinsically Disordered Protein by NMR Exchange Spectroscopy. *J. Am. Chem. Soc.* **2018**, *140* (3), 1148−1158.
(50) Sun, X.; Dyson, H. J.; Wright, P. E. A Phosphorylation-Dependent Switch in the Disordered P53 Transactivation Domain Regulates DNA Binding. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118* (1), e2021456118.
(51) Milles, S.; Mercadante, D.; Aramburu, I. V.; Jensen, M. R.; Banterle, N.; Koehler, C.; Tyagi, S.; Clarke, J.; Shammas, S. L.; Blackledge, M.; Gräter, F.; Lemke, E. A. Plasticity of an Ultrafast Interaction between Nucleoporins and Nuclear Transport Receptors. *Cell* **2015**, *163* (3), 734−745.
(52) Busch, D. J.; Houser, J. R.; Hayden, C. C.; Sherman, M. B.; Lafer, E. M.; Stachowiak, J. C. Intrinsically Disordered Proteins Drive Membrane Curvature. *Nat. Commun.* **2015**, *6*, 7875.
(53) Berlow, R. B.; Dyson, H. J.; Wright, P. E. Hypersensitive Termination of the Hypoxic Response by a Disordered Protein Switch. *Nature* **2017**, *543* (7645), 447−451.
(54) Ferreon, A. C. M.; Ferreon, J. C.; Wright, P. E.; Deniz, A. A. Modulation of Allostery by Protein Intrinsic Disorder. *Nature* **2013**, *498* (7454), 390−394.

https://doi.org/10.1021/jacs.1c06264
J. Am. Chem. Soc. 2021, 143, 20109−20121

(55) Uversky, V. N.; Oldfield, C. J.; Dunker, A. K. Intrinsically Disordered Proteins in Human Diseases: Introducing the D2 Concept. *Annu. Rev. Biophys.* **2008**, *37*, 215−246.
(56) Krzeminski, M.; Marsh, J. A.; Neale, C.; Choy, W.-Y.; Forman-Kay, J. D. Characterization of Disordered Proteins with ENSEMBLE. *Bioinformatics* **2013**, *29* (3), 398−399.
(57) Kashtanov, S.; Borcherds, W.; Wu, H.; Daughdrill, G. W.; Ytreberg, F. M. Using Chemical Shifts to Assess Transient Secondary Structure and Generate Ensemble Structures of Intrinsically Disordered Proteins. In *Intrinsically Disordered Protein Analysis: Vol. 1, Methods and Experimental Tools*; Uversky, V. N., Dunker, A. K., Eds.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2012; pp 139−152 DOI: 10.1007/978-1-61779-927-3_11.
(58) Bernadó, P.; Svergun, D. I. Analysis of Intrinsically Disordered Proteins by Small-Angle X-Ray Scattering. *Methods Mol. Biol.* **2012**, *896*, 107−122.
(59) Jensen, M. R.; Communie, G.; Ribeiro, E. A.; Martinez, N.; Desfosses, A.; Salmon, L.; Mollica, L.; Gabel, F.; Jamin, M.; Longhi, S.; Ruigrok, R. W. H.; Blackledge, M. Intrinsic Disorder in Measles Virus Nucleocapsids. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (24), 9839−9844.
(60) Holmstrom, E. D.; Nettels, D.; Schuler, B. Conformational Plasticity of Hepatitis C Virus Core Protein Enables RNA-Induced Formation of Nucleocapsid-like Particles. *J. Mol. Biol.* **2018**, *430* (16), 2453−2467.
(61) Soranno, A.; Buchli, B.; Nettels, D.; Cheng, R. R.; Müller-Späth, S.; Pfeil, S. H.; Hoffmann, A.; Lipman, E. A.; Makarov, D. E.; Schuler, B. Quantifying Internal Friction in Unfolded and Intrinsically Disordered Proteins with Single-Molecule Spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109* (44), 17800−17806.
(62) Sanabria, H.; Rodnin, D.; Hemmen, K.; Peulen, T.-O.; Felekyan, S.; Fleissner, M. R.; Dimura, M.; Koberling, F.; Kühnemuth, R.; Hubbell, W.; Gohlke, H.; Seidel, C. A. M. Resolving Dynamics and Function of Transient States in Single Enzyme Molecules. *Nat. Commun.* **2020**, *11*, 1231.
(63) Tsytlonok, M.; Sanabria, H.; Wang, Y.; Felekyan, S.; Hemmen, K.; Phillips, A. H.; Yun, M.-K.; Waddell, M. B.; Park, C.-G.; Vaithiyalingam, S.; Iconaru, L.; White, S. W.; Tompa, P.; Seidel, C. A. M.; Kriwacki, R. Dynamic Anticipation by Cdk2/Cyclin A-Bound P27 Mediates Signal Integration in Cell Cycle Regulation. *Nat. Commun.* **2019**, *10* (1), 1676.
(64) McKibben, K. M.; Rhoades, E. Independent Tubulin Binding and Polymerization by the Proline-Rich Region of Tau Is Regulated by Tau's N-Terminal Domain. *J. Biol. Chem.* **2019**, *294* (50), 19381−19394.
(65) Mitrea, D. M.; Kriwacki, R. W. Phase Separation in Biology; Functional Organization of a Higher Order. *Cell Commun. Signaling* **2016**, *14*, 1.
(66) Nasir, I.; Onuchic, P. L.; Labra, S. R.; Deniz, A. A. Single-Molecule Fluorescence Studies of Intrinsically Disordered Proteins and Liquid Phase Separation. *Biochim. Biophys. Acta, Proteins Proteomics* **2019**, *1867* (10), 980−987.
(67) Peter, M. F.; Gebhardt, C.; Mächtel, R.; Glaenzer, J.; Thomas, G. H.; Cordes, T.; Hagelueken, G. Cross-Validation of Distance Measurements in Proteins by PELDOR/DEER and Single-Molecule FRET. 2020, BioRxiv, DOI: 10.1101/2020.11.23.394080 (accessed Aug 10, 2021).
(68) Schuler, B.; Lipman, E. A.; Steinbach, P. J.; Kumke, M.; Eaton, W. A. Polyproline and the "Spectroscopic Ruler" Revisited with Single-Molecule Fluorescence. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (8), 2754−2759.
(69) Woźniak, A. K.; Schröder, G. F.; Grubmüller, H.; Seidel, C. A. M.; Oesterhelt, F. Single-Molecule FRET Measures Bends and Kinks in DNA. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (47), 18337−18342.
(70) Williams, A. T. R.; Winfield, S. A.; Miller, J. N. Relative Fluorescence Quantum Yields Using a Computer-Controlled Luminescence Spectrometer. *Analyst* **1983**, *108* (1290), 1067−1071.
(71) Brannon, J. H.; Magde, D. Absolute Quantum Yield Determination by Thermal Blooming. Fluorescein. *J. Phys. Chem.* **1978**, *82* (6), 705−709.
(72) Fischer, M.; Georges, J. Fluorescence Quantum Yield of Rhodamine 6G in Ethanol as a Function of Concentration Using Thermal Lens Spectrometry. *Chem. Phys. Lett.* **1996**, *260* (1), 115−118.
(73) Svergun, D.; Barberato, C.; Koch, M. H. J. CRYSOL − a Program to Evaluate X-Ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Crystallogr.* **1995**, *28* (6), 768−773.
(74) Shen, Y.; Bax, A. Protein Backbone Chemical Shifts Predicted from Searching a Database for Torsion Angle and Sequence Homology. *J. Biomol. NMR* **2007**, *38* (4), 289−302.
(75) Milles, S.; Tyagi, S.; Banterle, N.; Koehler, C.; VanDelinder, V.; Plass, T.; Neal, A. P.; Lemke, E. A. Click Strategies for Single-Molecule Protein Fluorescence. *J. Am. Chem. Soc.* **2012**, *134* (11), 5187−5195.
(76) Marsh, J. A.; Singh, V. K.; Jia, Z.; Forman-Kay, J. D. Sensitivity of Secondary Structure Propensities to Sequence Differences between α- and γ-Synuclein: Implications for Fibrillation. *Protein Sci.* **2006**, *15* (12), 2795−2804.
(77) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes. *J. Biomol. NMR* **1995**, *6* (3), 277−293.
(78) PicoQuant/PicoQuant-Time-Tagged-File-Format-Demos. https://github.com/PicoQuant/PicoQuant-Time-Tagged-File-Format-Demos (accessed Apr 16, 2021).
(79) Enderlein, J.; Robbins, D. L.; Ambrose, W. P.; Goodwin, P. M.; Keller, R. A. The Statistics of Single Molecule Detection: An Overview. *Bioimaging* **1997**, *5* (3), 88−98.
(80) Ferreon, A. C. M.; Gambin, Y.; Lemke, E. A.; Deniz, A. A. Interplay of Alpha-Synuclein Binding and Conformational Switching Probed by Single-Molecule Fluorescence. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (14), 5645−5650.
(81) Karstens, T.; Kobs, K. Rhodamine B and Rhodamine 101 as Reference Substances for Fluorescence Quantum Yield Measurements. *J. Phys. Chem.* **1980**, *84* (14), 1871−1872.
(82) Magde, D..; Brannon, J. H.; Cremers, T. L.; Olmsted, J.. Absolute Luminescence Yield of Cresyl Violet. A Standard for the Red. *J. Phys. Chem.* **1979**, *83* (6), 696−699.
(83) Nettels, D.; Haenni, D.; Maillot, S.; Gueye, M.; Barth, A.; Hirschfeld, V.; Hübner, C.; Léonard, J.; Schuler, B. Excited-State Annihilation Reduces Power Dependence of Single-Molecule FRET Experiments. *Phys. Chem. Chem. Phys.* **2015**, *17* (48), 32304−32315.
(84) Lee, N. K.; Kapanidis, A. N.; Wang, Y.; Michalet, X.; Mukhopadhyay, J.; Ebright, R. H.; Weiss, S. Accurate FRET Measurements within Single Diffusing Biomolecules Using Alternating-Laser Excitation. *Biophys. J.* **2005**, *88* (4), 2939−2953.
(85) ATTO-TEC GmbH - ATTO-TEC GmbH https://www.atto-tec.com/?language=de (accessed Aug 13, 2021).

https://doi.org/10.1021/jacs.1c06264
J. Am. Chem. Soc. 2021, 143, 20109−20121