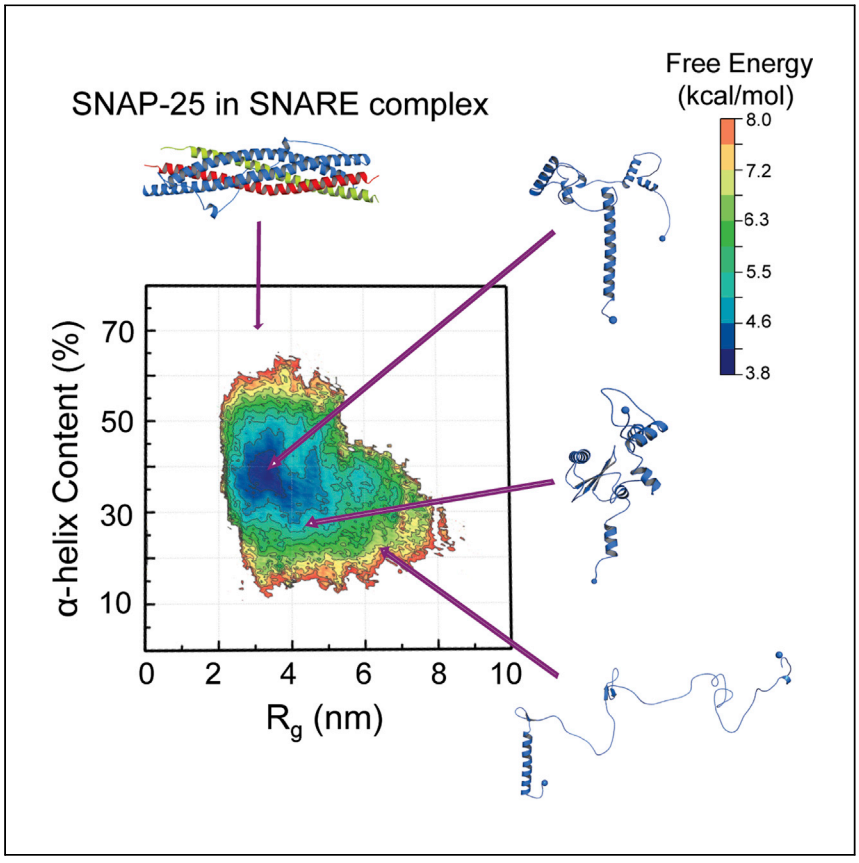
Cell Reports
Physical Science

CellPress
OPEN ACCESS

Article

Integrative structural dynamics probing of the conformational heterogeneity in synaptosomal-associated protein 25



Intrinsically disordered proteins (IDPs) play important functional roles owing to their high conformational plasticity and structural heterogeneity, which also renders the structural characterization of IDPs challenging. Saikia et al. report an integrative approach to probe IDP structural dynamics by synergistically combining simulations with label-based experiments at ensemble and single-molecule levels.

Nabanita Saikia, Inna S. Yanez-Orozco, Ruoyi Qiu, ..., Tatyana I. Smirnova, Hugo Sanabria, Feng Ding

hsanabr@clemson.edu (H.S.)
fding@clemson.edu (F.D.)

Highlights

A structural dynamics probing method integrates simulations with label-based experiments

Replica-exchange DMD simulations capture the structural dynamics of SNAP-25 *in silico*

Label-based experiments using FRET and DEER probe conformational dynamics *in vitro*

The synergy *in silico* and *in vitro* uncovers functionally important helix-coil transitions

Saikia et al., Cell Reports Physical Science *2*, 100616
November 17, 2021
https://doi.org/10.1016/j.xcrp.2021.100616

Cell Reports
Physical Science

CellPress
OPEN ACCESS

Article

Integrative structural dynamics probing of the conformational heterogeneity in synaptosomal-associated protein 25

Nabanita Saikia,[1,2] Inna S. Yanez-Orozco,[1] Ruoyi Qiu,[3] Pengyu Hao,[3] Sergey Milikisiyants,[4] Erkang Ou,[4] George L. Hamilton,[1] Keith R. Weninger,[3] Tatyana I. Smirnova,[4] Hugo Sanabria,[1,*] and Feng Ding[1,5,*]

SUMMARY

SNAP-25 (synaptosomal-associated protein of 25 kDa) is a prototypical intrinsically disordered protein (IDP) that is unstructured by itself but forms coiled-coil helices in the SNARE complex. With high conformational heterogeneity, detailed structural dynamics of unbound SNAP-25 remain elusive. Here, we report an integrative method to probe the structural dynamics of SNAP-25 by combining replica-exchange discrete molecular dynamics (rxDMD) simulations and label-based experiments at ensemble and single-molecule levels. The rxDMD simulations systematically characterize the coil-to-molten globular transition and reconstruct structural ensemble consistent with prior ensemble experiments. Label-based experiments using Förster resonance energy transfer and double electron-electron resonance further probe the conformational dynamics of SNAP-25. Agreements between simulations and experiments under both ensemble and single-molecule conditions allow us to assign specific helix-coil transitions in SNAP-25 that occur in submillisecond timescales and potentially play a vital role in forming the SNARE complex. We expect that this integrative approach may help further our understanding of IDPs.

INTRODUCTION

Intrinsically disordered proteins (IDPs) play critical roles in diverse regulatory and cellular signaling processes, and are also associated with multiple neurodegenerative diseases.[1–3] The intrinsic disorder with high conformational plasticity is essential for forming transient yet stable signaling protein complexes[4] and key transcription regulators.[5] In synapses, IDPs and proteins containing intrinsically disordered regions (IDRs) are important in neurotransmitter release,[6] the formation of junctions,[7] remodeling of the postsynaptic density,[8] and signaling in the cytoplasmic region of several membrane receptors.[9,10] The high degree of conformational heterogeneity of IDPs is due to the lack of a well-defined secondary and tertiary structure organization, the biased amino acid composition, and low sequence complexity.[11–13]

To structurally characterize IDPs or IDRs, ensemble-based nuclear magnetic resonance (NMR) with its atomic resolution and broad temporal resolution has been the gold standard methodology.[14–16] Recent advances in single-molecule experiments, however, offer new insights into the dynamic behavior of IDPs.[17] Here, we implement an alternative approach integrating replica-exchange discrete molecular dynamics (rxDMD) and label-based experiments at both ensemble and single-molecule conditions to resolve structural dynamics and heterogeneity of IDPs. DMD is a

[1]Department of Physics and Astronomy, Clemson University, Clemson, SC 29634, USA

[2]Department of Chemistry, Navajo Technical University, Chinle, AZ 86503, USA

[3]Department of Physics, North Carolina State University, Raleigh, NC 27695, USA

[4]Department of Chemistry, North Carolina State University, Raleigh, NC 27695, USA

[5]Lead contact

*Correspondence: hsanabr@clemson.edu (H.S.), fding@clemson.edu (F.D.)

https://doi.org/10.1016/j.xcrp.2021.100616

rapid and predictive molecular dynamics approach to sample protein dynamics at long timescales,[18] while replica exchange is an enhanced sampling approach for obtaining a generalized ensemble of a molecular system and exploring the free-energy landscape.[19] DMD simulations can also incorporate experimentally derived structural and dynamic information as constraints to reconstruct experimentally consistent conformational ensembles for both structured proteins and IDPs.[20,21] We use SNAP-25 (synaptosomal-associated protein of 25 kDa) as a prototypical IDP for our system of study. SNAP-25, together with the synaptic vesicle protein Synaptobrevin 2 (or VAMP 2) and plasma membrane protein Syntaxin 1a, bind together to form a coiled 4-helix bundle as part of the SNARE (soluble *N*-ethylmaleimide-sensitive factor attachment receptor) complex.[22–24] The SNARE complex is a crucial component of the eukaryotic fusion machinery at the neuronal synapses.[25] The SNARE motif of ~60–70 residues features heptad repeats via a disorder-to-order transition to form the SNARE complex.[6,26] Although the structural and functional role of the SNARE complex in eukaryotic membrane fusion machinery has been intensely studied,[27,28] probing the disordered structural heterogeneity of unbound SNAP-25 remains challenging at both experimental and computational levels.

The structural ensemble of SNAP-25 derived from unbiased rxDMD simulations shows an excellent agreement with prior ensemble experimental measurements, including radius of gyration (R_g) or hydrodynamic radii (R_H),[29] inter-dye or inter-residue distances from single-molecule Förster resonance energy transfer (smFRET), and circular dichroism (CD) spectra.[30] In addition, the incorporation of label-based experiments at the ensemble and single-molecule level on specific regions of SNAP-25 allow us to monitor the structural heterogeneity crucial in the disorder-to-order transition required for binding. The agreement between rxDMD simulations and label-based experimental methods enables us to capture a novel conformational switching behavior of SNAP-25 that could promote efficient scavenging of binding partners. Moreover, SNAP-25 retains residual and transient secondary structural elements compatible with the bound state, central to its synaptic transmission and membrane fusion function. The integration of rxDMD simulations and label-based approaches can augment standard structural characterization methods and allows us to examine the detailed complexities of IDPs and IDRs.

RESULTS AND DISCUSSION

Integrative modeling and validation of the conformational ensemble

We use an integrative modeling approach blending simulations with complementary label-based experiments to characterize the structural dynamics and heterogeneity of SNAP-25. SNAP-25 is a moderate-sized IDP with a sequence length of 206 amino acids (Figure 1A). Our workflow (Figure 1B) describes a procedure to incorporate prior structural knowledge to reconstruct the conformational ensemble and study the conformational heterogeneity of biomolecules, including IDPs. Prior knowledge includes the primary, secondary, and tertiary structural information from experimental measurements, such as R_g, R_H, and CD spectra. Starting with the extended conformation from the amino acid sequence, we prepare our protein for rxDMD simulations. Multiple simulations are performed at a wide range of temperatures, with periodic exchanges of conformations between replica according to the Metropolis criterion. Using the weighted histogram analysis method (WHAM)[31] on the simulation trajectories, we can determine various thermodynamic parameters such as specific heat (C_v) and the melting temperature (T_m) for coil-to-molten globular transition, and compute structural observables such as R_g, R_H, inter-residue distances, and estimated CD spectra, which can be compared with experiments to reconstruct the experimentally consistent structural ensemble.

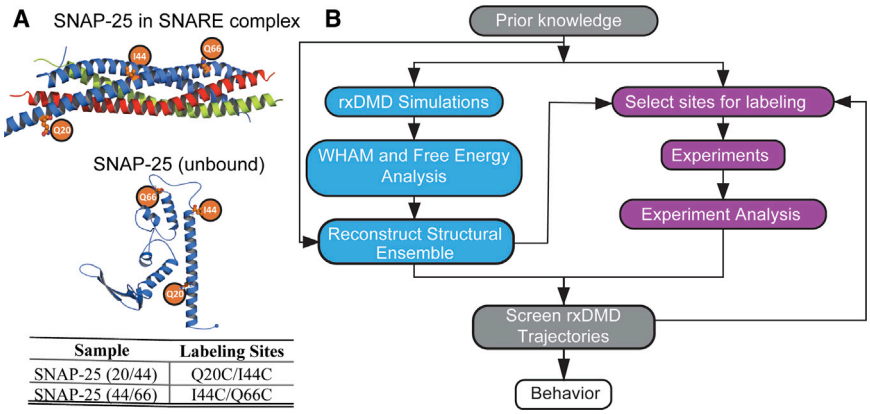


Figure 1. Flow diagram of an integrative approach to characterize the structural dynamics of IDPs

(A) SNAP-25 protein in SNARE complex and unbound with the FRET and EPR label sites SNAP-25 (20/44) and SNAP-25 (44/66). SNARE complex (PDB: 1SFC) is formed by the 3 core proteins synaptobrevin (red), syntaxin (green), and SNAP-25 (blue, 2 helices).

(B) A workflow of the integrative approach using DMD and label-based experiments to determine the structural dynamics and heterogeneity of IDPs and IDRs.

To carry out label-based experiments, we determine the preferred sites for labeling according to the dynamics and heterogeneity of a particular region of interest. The selection of labeling sites can be empirically determined or designed.[32] Alternatively, it is possible to screen the trajectories from simulations to identify ideal labeling sites. Next, we perform experiments at ensemble or single-molecule conditions. For our purposes here, we measure two pairs of selected sites that encompass the region of the disorder-to-order transition of SNAP-25 using FRET experiments in three modalities: (1) a cuvette/ensemble mode with time-resolved fluorescence, (2) freely diffusing smFRET experiments in confocal mode using multiparameter fluorescence detection (MFD), (3) immobilized smFRET under total internal reflection microscopy (TIRFM), as well as (4) double electron-electron resonance (DEER) experiments. We further screen the accessible volume (AV) using the reconstructed structural ensemble from rxDMD simulations. AV models provide a sterically allowed space of the dye attached to the protein by approximating the dye to be freely diffusing inside a uniform spatial distribution of the AV. An agreement between ensemble, single-molecule experiments, and DMD simulations allows us to map the conformational free-energy landscape and obtain local or region-wise structural heterogeneity. FRET and DEER measurements report inter-label distances; thus, multiple pairs are required for complete structural modeling, with enough sampling of the overall organization of the molecule.[32–36] Our integrative method can leverage experimental data, prove specific regions of interest that would be limited in determining structures with reasonable accuracy, and provide complementary information to NMR experiments. Similar workflows resolved the structural dynamics in enzymes,[36] the dynamic heterogeneity in multi-domain proteins,[18] and other intrinsically disordered proteins,[37,38] and has been used for determining the ensemble switching mechanism of eukaryotic thiamin riboswitch.[39] Furthermore, the integrative structural modeling using DMD and label-based experiments can solve the remaining challenges in highly dynamic systems such as IDPs, IDRs, and complex supertertiary structural dynamics.

rxDMD simulation of SNAP-25

Prior knowledge of SNAP-25 includes a known disorder-to-order conformational switching occurring at various timescales.[29] Dynamic light scattering, sedimentation

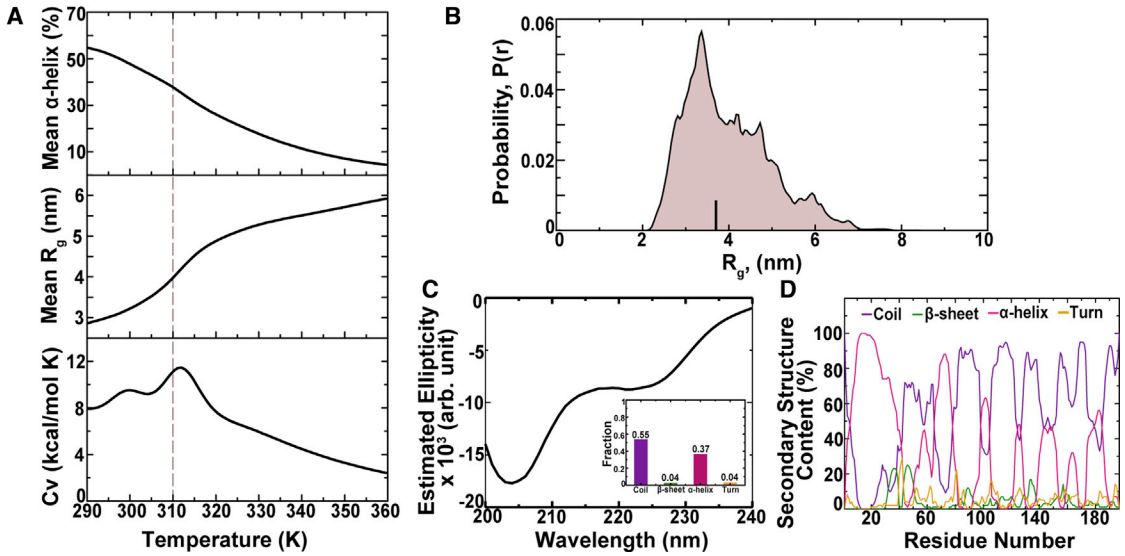


Figure 2. Ensemble analysis of SNAP-25

(A) Specific heat at constant volume (C_v) spectra, mean $\langle R_g \rangle$, and mean helicity dependence as a function of temperature. The C_v shows a coil-to-globular transition above 300 K, as shown in the dashed line.

(B) The probability distribution of R_g. The vertical line corresponds to the experimental value of 3.85 nm.[29]

(C) Estimated ellipticity in the far-UV region. The estimated CD shows 2 minima ~204 and 223 nm and a low ellipticity above 215 nm, suggesting a disordered protein structure with a small fraction of α-helix. The inset panel shows the fraction of secondary structure corresponding to α-helix, β sheet, random coil, and turn regions. A total of 37% of amino acids adopt an α-helix and 55% are involved in random coil regions. β sheet and turn regions constitute a small fraction of the residues (4%).

(D) Per residue-wise secondary structure content depicting the random coil, β sheet, α-helix, and turn regions, obtained from the conformational ensemble.

velocity analytical ultracentrifugation, and analytical size exclusion chromatography measurements showed that SNAP-25 has a mean R_H value of 3.75 ± 0.38 nm.[29] This value is consistent with the expected R_g of ~3.85 nm estimated as the root-mean-square distance for an IDP with an average persistence length of 0.6 nm and 206 amino acids in length.[29] Also, a CD spectrum study of SNAP-25 reported a low α-helical content of ~14%, suggesting a transient residual secondary structure content.[30]

Using the equilibrated trajectories in rxDMD simulations (Figures S1 and S2), we calculate the C_v, R_g, and mean α-helix content of SNAP-25 as the function of temperature. The C_v plot shows 2 peaks at ~300 and ~312 K (Figure 2A). The first peak at ~300 K corresponds to an increase in the local structural disorder, while the second peak at ~312 K denotes the midpoint of R_g increase with increasing temperature. Simulations above 312 K feature continued loss of ordered secondary and tertiary structures of the IDP (see Figure S3A). Hence, the second peak corresponds to the coil-to-molten globular transition, T_m. Based on the mean $\langle R_g \rangle$ and helicity, we chose 310 K slightly below the T_m to be the optimal temperature for studying the conformational substates and structural ensemble of SNAP-25. At 310 K in DMD simulations, the $\langle R_g \rangle$ at 4.0 nm, mean helicity, and the correspondingly estimated CD spectrum agree well with reported measurements.[29,30,40] The probability distribution of the R_g shows a prominent peak at ~3.36 nm and multiple shoulders extending from 4.0 to 6.0 nm (Figure 2B). The tailing to large R_g values in the probability distribution suggests that SNAP-25 can adopt occasionally extended conformations. Since R_g and R_H are related to each other in a non-straightforward way,[41] we also estimate R_H for the computationally derived conformational ensemble of SNAP-25 using HYDROPRO.[42] The obtained mean R_H is 3.38 ± 0.4 nm, in agreement with

experimental measurements.[29] The overall structural ensemble of SNAP-25 is disordered, with ~37% helical and ~55% coil contents. The 2 states together account for 92% of all secondary structure states. We note that the helical content in our simulations is higher than the previously reported value of ~14%.[30] However, using a single value of the mean residue ellipticity value at 222 nm, $[\Theta]_{222}$, instead of the whole spectrum to linearly extrapolate the helical content, the earlier study successfully monitored changes of the helical content upon environment changes, but likely underestimated the α-helical content of SNAP-25 since coils enriched in the IDP could offset $[\Theta]_{222}$ from helices with opposite contributions. In addition, the peptide in simulations does not include the C-terminal 9-residue coil and the His6-tag used in the prior study. However, the estimated far-UV CD spectrum based on secondary structure contents in simulations agreed well with the experimental spectrum.[30] Both computational and experimental CD spectra depict 2 minima centered ~204 and 222 nm and a low ellipticity above 215 nm, reminiscent of disordered protein structure (Figure 2C). We also compute the secondary structure probabilities for each residue (Figure 2D). Only the first 40 N-terminal residues and residues 65–80 are predominantly α-helical; other regions either adopt coil-dominant structures or undergo frequent random coil to ordered helix transitions (e.g., residues ~55, 100, 110, 145–150, 170–180).

The conformational heterogeneity in SNAP-25 ensemble is exemplified in the free-energy landscape, calculated as the 2-dimensional (2D) potential of mean force (2D-PMF) with respect to the R_g and α-helix content (Figure 3A) and as the function of R_g and coil content (Figure S3B). The energy landscape shows a minimum energy basin (denoted by α) around R_g values of 2.7–3.9 nm with helix content between 32% and 45% and the coil content between 45% and 56%. We also observe two higher energy basins, denoted as β and γ, with larger R_g values and slightly lower helical contents. We perform a clustering analysis to acquire the representative structures belonging to the identified energy basins in the energy landscape. Representative structures corresponding to the three basins display a stable α-helix in the N-terminal region around the first 40 residues and intermittent residual secondary structure regions (Figure 3A). Notably, for the energy basins β and γ, the transition to extended structures are mainly due to the intrinsically disordered regions within the protein that render high conformational flexibility. Thus, our simulations sample a heterogeneous conformational landscape with a high content of random coil and residual α-helical regions that undergo reversible order-to-disorder transitions. The contents of other secondary structure, β sheet and turn, are low.

Helix-coil transitions of the N-terminal region of SNAP-25 *in silico*

To gain insights into the tertiary interactions in the conformational ensemble of SNAP-25, we compute the residue-wise contact frequency map (Figure 3B). The contact map lacks long-range contacts, consistent with both the coil and helical content of SNAP-25. The analysis confirms the high probability of α helices (contact patterns along the diagonal) around the N-terminal residues 1–40 and residues 65–80, along with a low probability of a short β-hairpin turned around residue 40 between the 2 stable helices. The contact map for the rest of the protein also features weak helices, and local interactions resulted from frequent coil-helix transitions.

Without loss of generality, we next focus on the detailed dynamics of the N-terminal region encompassing the first 2 helices (i.e., residues 1–40 and 65–80) to characterize the structural heterogeneity of SNAP-25 due to frequent order-disorder transitions. To compare with label-based experiments that measure inter-residue distances, we propose to investigate 2 specific pairs, Q20/I44 and I44/Q66, which

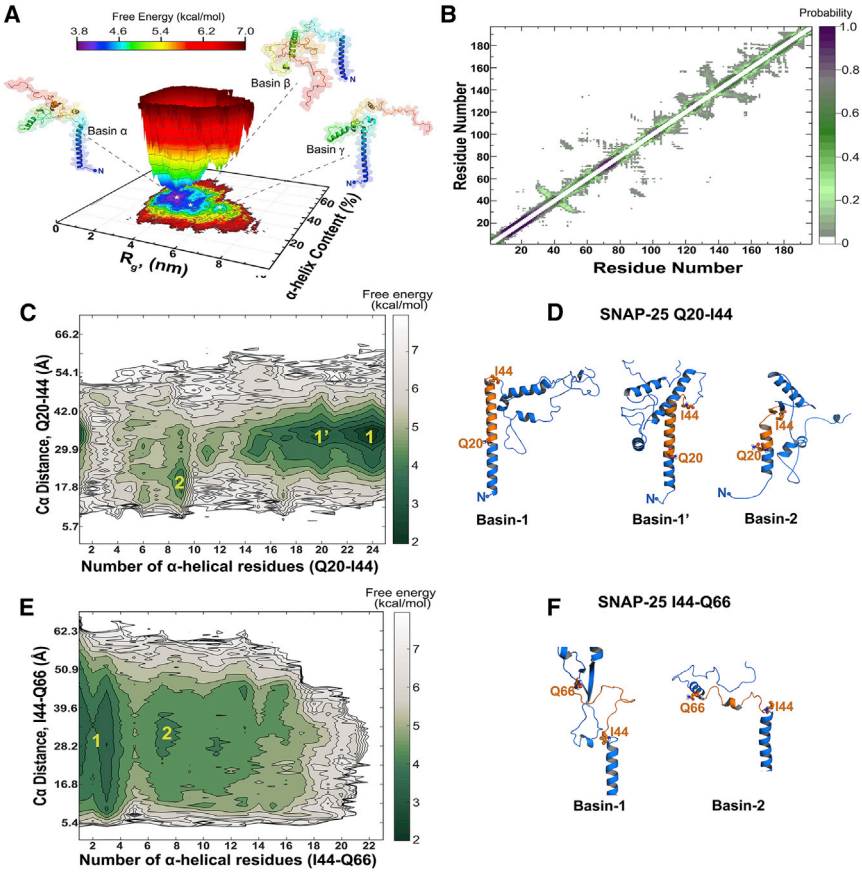


Figure 3. Free-energy landscape and helix-coil transition of the N-terminal region of SNAP-25

(A) The energy landscape as a function of R_g and α-helix content from rxDMD simulations and representative structures corresponding to 3 identified energy basins, α, β, and γ.

(B) Contact frequency map depicting tertiary contacts in SNAP-25.

(C) The 2D-PMF as a function of the Cα distance and number of α-helical residues between the residue pair Q20/I44.

(D) Central representative structures of SNAP-25 Q20/I44 (basins 1, 1′, and 2). The residues between the pairs 20/44 used for labeling the protein are highlighted in orange.

(E) The 2D-PMF between the Cα distance and number of α-helical residues between the residue pair I44/Q66. The presence of multiple shallow minima indicates many conformational substates revealing conformational heterogeneity reminiscent of IDPs and IDRs.

(F) Central representative structures of SNAP-25 I44/Q66 (basins 1 and 2). The residues between 44/66 used for labeling the protein are highlighted in orange.

probe the folding/refolding dynamics of the first 2 helices. For each pair, we compute the 2D-PMF as the function of the inter-Cα distance of the pairing residues and the number of residues in the enclosed sequence fragment adopting α-helix (Figure 3C) or random coil (Figure S4) conformations, as well as the probability density distribution of the inter-Cα distance (Figure S5) for direct comparison with distance-based FRET and DEER experiments. For the residue pair Q20/I44, the lowest energy basin has the number of α-helical residues ~24 equal to the fragment length, suggesting that the region between Q20/I44 is dominated by a complete α-helix with the corresponding inter-Cα distance 30–35 Å (basin 1, Figure 3C). Other two higher energy basins, 1′ and 2, can be observed in the PMF plot. Basin 1′ has a similar inter-Cα distance as basin 1, but with a loss of 1 α-helical turn, and basin 2 has a much smaller inter-Cα ~17.8 Å and only 9 α-helical residues. The central representative structures corresponding to 3 basins (Figure 3D) clearly demonstrate the transient

unfolding of the N-terminal α-helix in the first 40 residues. For the residue pair I44/Q66, the secondary structure of the sequence fragment is low in α-helix content and rich in random coil (Figures 3E and S4B). The computed 2D-PMF shows a minimum energy basin with a wide spread in the inter-Cα pair distance between 11 and 45 Å, consistent with a coil-rich fragment (Figure 3E). Representative structures corresponding to two identified energy basins 1 and 2 are shown in Figure 3F, which demonstrate that residues between I44 and Q66 indeed adopt coil structure along with a very low fraction of α-helix. Next, we experimentally probe the dynamics of local order-disorder transitions using two double cysteine variants for site-specific label attachment: SNAP-25 20/44 and SNAP-25 44/66.

Conformational plasticity of the N-terminal region of SNAP-25 using label-based experiments

To experimentally monitor the conformational plasticity of the N-terminal region of SNAP-25, we engineered 2 double cysteine variants at positions 20/44 and 44/66. We chose these three sites because they belong to the same helix as part of the SNARE complex (Figure 1A). The variant 20/44 monitors the helix-coil transition observed in rxDMD simulations and capture the dynamics of the residual helices, while variant 44/66 characterizes the coil region between these 2 amino acids. These 2 specific pairs, 20/44 and 44/66, can probe the folding/refolding dynamics of the first 2 helices, relevant for the SNARE complex formation. rxDMD simulations of the 2 double-cysteine mutants confirmed that the mutations needed to introduce probes did not affect the overall dynamics of SNAP-25 (Figure S6). We used the same constructs to measure the distance between labels at various modalities, including (1) a cuvette/ensemble FRET mode with time-resolved fluorescence, (2) freely diffusing smFRET experiments in confocal mode using MFD, (3) immobilized smFRET under TIRFM, and (4) DEER in frozen samples. FRET experiments in different modalities provide complementary information. smFRET in TIRFM mode monitors FRET states that are stable at the second timescale. smFRET in confocal mode captures dynamic averaging occurring faster than the observation time limited by the molecules' transient time over the confocal volume, usually on the order of milliseconds. Cuvette/ensemble time-resolved fluorescence will capture the time evolution of the FRET process in the nanosecond timescales with a high signal-to-noise ratio by recording over 10^6 photons. We refer to the identified states as limiting states.

The variant 20/44 monitors the helix-coil transition observed in DMD simulations (Figure 4). We present smFRET in TIRFM mode, where we use the donor and acceptor fluorescence signal as a function of time (Figure 4A, top) to derive a monodisperse FRET distribution that peaks at 0.8 (Figure 4A, bottom). In contrast, we required two limiting states when modeling the time-resolved fluorescence decays in cuvette experiments (Figure 4B). Our fit models the FRET-induced donor decay globally, with the fluorescence decay containing the donor-only samples as previously done.[43] Each limiting state follows a Gaussian distribution, and all of the required parameters to analyze FRET samples, including the dyes uses, the Förster radii, donor- and acceptor-only lifetime decay fits, the residual anisotropies and $\langle\kappa^2\rangle$, and the donor and acceptor distances and fractions obtained by fluorescence decays are included in Tables 1 and S1–S5. Our results suggest that the SNAP-25 within the region of 20–44 shows a transition between at least 2 conformational states resolved by FRET with a fraction of donor-only-like state. This donor-only-like fraction could correspond to the fraction of molecules containing only the donor fluorophore, molecules with the inactive acceptor, or distances beyond FRET sensitivity.

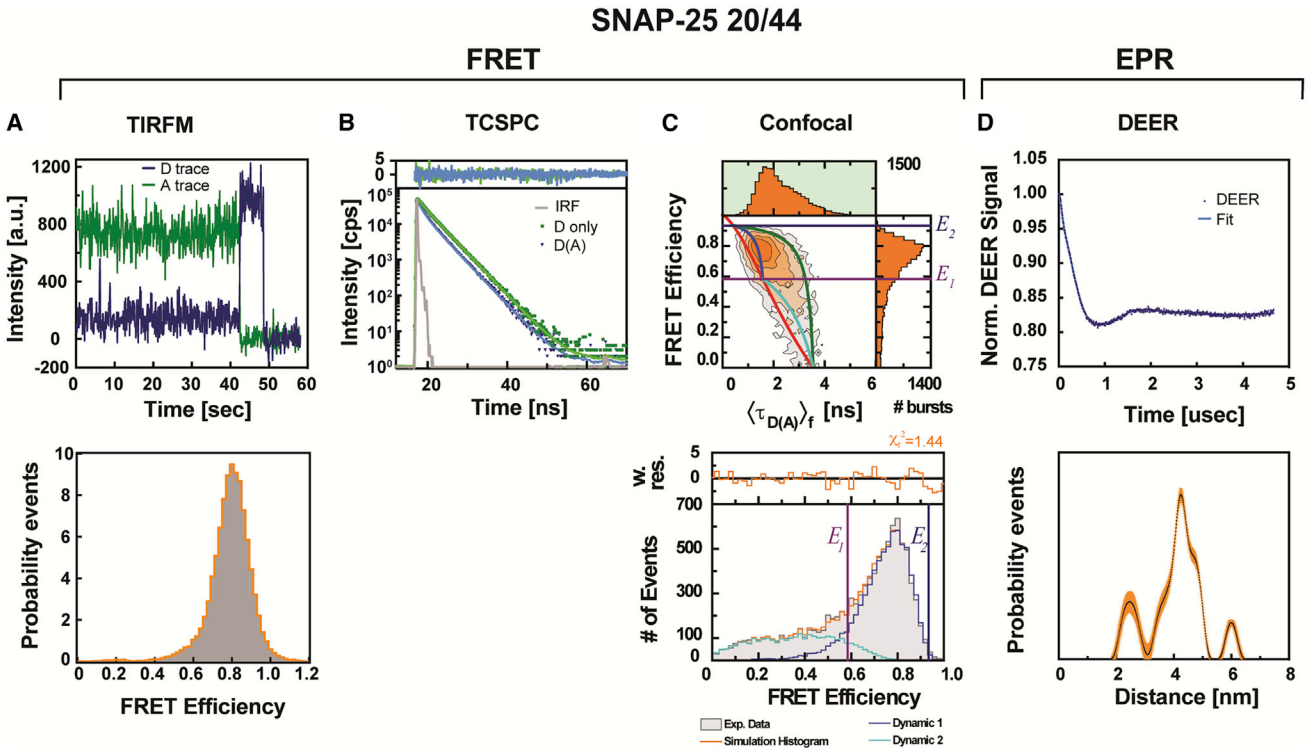


Figure 4. Label-based observations of SNAP-25 20/44

(A) Top: exemplary fluorescence traces of the donor and acceptor fluorescence signal as a function of time in immobilized FRET TIRFM experiments. Bottom:calculated FRET efficiency histogram from donor and acceptor traces. For TIRFM FRET, we used Alexa Fluor 555 and 647 maleimide as donor and acceptor fluorophores, respectively.

(B) TCSPC decays of donor-only labeled and donor in the presence of acceptor-labeled SNAP-25. Weighted residuals of the fit with a model with 2-Gaussian distributed states. Table 1 shows the values of the fit.

(C) Top: 2D MFD histogram of the FRET efficiency versus the average fluorescence lifetime per burst in confocal mode. Static (red) and dynamic (blue, green, and cyan) FRET lines as guides for the interconversion between identified states by photon distribution analysis (PDA) with a time window analysis of 5 ms. Tables S3 and S4 summarized the FRET lines. Bottom: PDA analysis with a model of 2 interconverting states and a NO FRET population (Table S9). Vertical lines correspond to the limiting states E_1 and E_2. Those states appear as horizontal lines in the MFD histogram. Correction parameters are shown in Table S10. For TCSPC and confocal FRET measurements, we used Alexa Fluor 488 and 647 maleimide as donor and acceptor fluorophores, respectively.

(D) Top: normalized DEER signal and its fit with the Tikhonov regularization to derive the inter-spin distance distribution (bottom). The orange area represents the confidence interval of the Tikhonov regularization.

Given that the FRET efficiency distribution in TIRFM mode showed a unimodal distribution (Figure 4A), we proceeded to measure FRET in a confocal setup in MFD mode (Figure 4C, top), as previously done for multiple systems.[44,45] We observe a unimodal distribution, but in this case, due to the integration time that happens in the millisecond timescale, we conclude that there is dynamic averaging between the limiting states that are identified by time-correlated single-photon counting (TCSPC) in nanoseconds. This shift occurs because the peak of the distribution shifts to the right of the static FRET line (red line in Figure 4C top and Table S6). To further corroborate this, we use photon distribution analysis (PDA) to fit the FRET efficiency distribution (Figure 4C, bottom) at multiple time windows (Δt = 0.5, 1, 2, and 5 ms; Figure S7). We identify a dynamic interconversion between the 2 limiting states, with FRET efficiency levels E_1 = 0.58 and E_2 = 0.93 and a population with NO FRET. In Table S7, we present the conversion between FRET efficiencies and the average fluorescence lifetime ($\langle\tau_{D(A)}\rangle_f$). The two FRET efficiencies agree very well with the limiting states identified by TCSPC (Table 1), and the NO FRET state is also consistent with those observations. Next, we map those FRET efficiency levels in the MFD

Table 1. Summary of experimental FRET and model distances based on the AV

Sample 20/44	x_1	$\langle R_{DA}\rangle_1$ [Å]	x_2	$\langle R_{DA}\rangle_2$ [Å]	D only	χ^2
TCSPC (2-Gaussian)	0.36	49.4 ± 2.6	0.64	33.3 ± 1.8	0.40	1.14
Screening	0.81	43.5 ± 0.1[a]	0.19	36.0 ± 1.3[a]		
Sample 44/66	x_1	$R_{DA(1)}$ [Å]			D only	χ^2
TCSPC (1-Gaussian)	1.00	39.2 ± 4.5			0.77	1.10
Screening	1.00	40.3 ± 0.5[a]				
WLC 44/66	L [Å]	κ	l_p [Å]		D only	χ^2
Experiment	79.2	0.19	15.2 ± 4.3		0.46	1.19

x_1 and x_2 are the ith population fractions and D only is the population fraction with no observed FRET.
[a]95% confidence interval from fitting Gaussian distributions to the histograms in Figure 6.

histograms (Figure 4C, top) by placing two horizontal lines corresponding to the limiting states identified by PDA. If molecules interconvert between these states, then single-molecule events should follow or fall within the dynamic FRET lines (blue, green, and cyan; Table S8), representing the limiting states' exchange processes.[46,47] The observed histogram distribution lies within these FRET lines, which means that the model is in qualitative agreement with the observation in the MFD histograms. To achieve independent validation of the FRET experiments, we use DEER to obtain the inter-spin distance distribution.

In contrast with TCSPC, the ill-poised problem uses the Tikhonov regularization algorithm to derive the inter-spin distance distribution. Figure 4D shows exemplary DEER decay, and the inter-spin distribution between homo spin labels at the same cysteine positions (20/44) as those used in FRET experiments. In the inter-spin distribution, we identify three different conformations or states. One that peaks at ~2 nm is consistent with the FRET efficiency E_2, a larger peak centered at ~4 nm is consistent with the FRET efficiency E_1, and there is also a smaller peak at ~6 nm. The latter population is consistent with the NO FRET state because the presence of the fluorescence labels will push the distance ~2 nm farther apart beyond FRET detection.

Next, we studied the variant 44/66 to characterize the coil region between these amino acids. Like the variant 20/44, TIRFM trajectories (Figure 5A, top) mostly show unimodal distributions with a mean FRET efficiency of 0.8. However, in some traces, we observe transitions showing a lower FRET efficiency of 0.3 (Figure 5A, bottom). Time-resolved fluorescence decays from TCSPC experiments (Figure 5B) were analyzed with increasing complexity levels in the model functions. The figure of merit χ^2 for a 1-Gaussian distributed state was $\chi^2_{1,Gauss}$ = 1.10. Then, we probed a 2-Gaussian model that would be consistent with the 2 states identified in TIRFM experiments. It showed a slightly worse $\chi^2_{2,Gauss}$ = 1.18 and $\langle R_{DA}\rangle^{(1)}$ = 41 and $\langle R_{DA}\rangle^{(2)}$ = 21 Å. Although the fit with the 1-Gaussian distributed state was statistically better than the 2-Gaussian distributed states, the 1-state model is not consistent with the smFRET results in MFD mode. In confocal measurements (Figure 5C), we observed a skewed distribution toward NO FRET clearly to the right of the static FRET line (red line). This line sets the limit when the host molecule—SNAP-25, in this case—is considered rigid. It is not. Thus, a single static host with a Gaussian distributed state due to the mobile fluorophores is not consistent with the MFD experiments. Next, we applied PDA and failed to fit the FRET efficiency with a 2-state model compatible with the 2-Gaussian distributed states because we could not globally fit multiple time windows with a reasonable figure of merit. With the apparent discrepancy between TCSPC and MFD, we decided to model the time-resolved fluorescence decay of the DA sample using a worm-like-chain (WLC) model (Equations 2-3 in

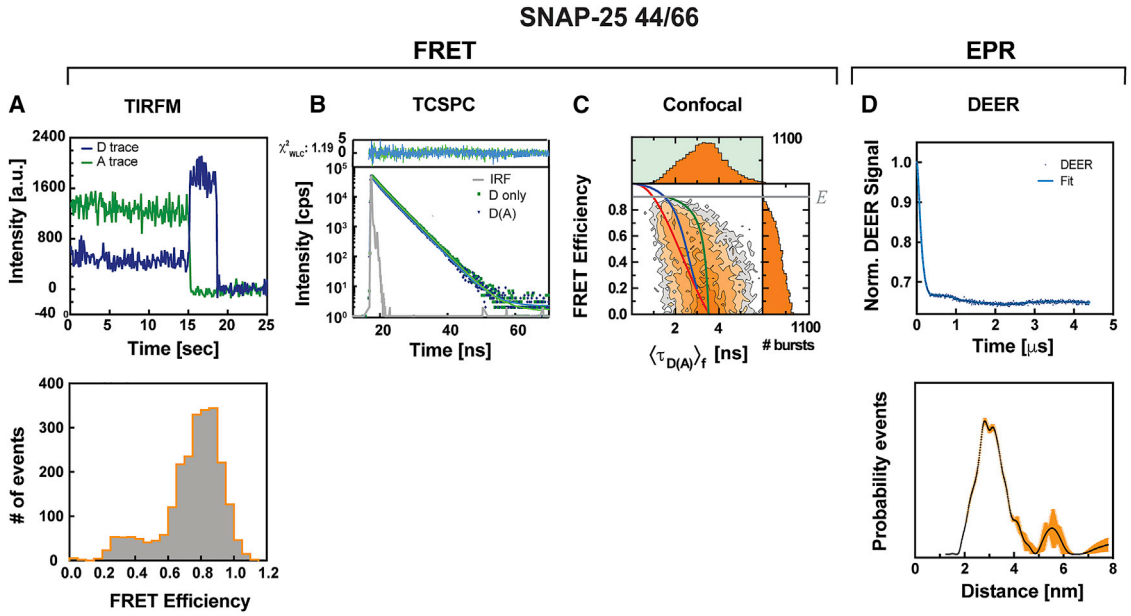


Figure 5. Label-based observations of SNAP-25 44/66

(A) Top: exemplary fluorescence traces of the donor and acceptor fluorescence signal as a function of time in immobilized FRET TIRFM experiments. Bottom: calculated FRET efficiency histogram from donor and acceptor traces. For TIRFM FRET, we used Alexa Fluor 555 and 647 maleimide as donor and acceptor fluorophores, respectively.

(B) TCSPC decays of donor-only labeled and donor in the presence of acceptor-labeled SNAP-25. Weighted residuals of the fit with a worm-like-chain (WLC) model with figure of merit χ^2_{WLC}. Table 1 shows the values of the fit.

(C) A 2D histogram of the FRET efficiency versus the average fluorescence lifetime per burst in confocal mode. Static (red) and dynamic (blue, green, and cyan) FRET lines as guides for the interconversion between identified states by PDA. Correction parameters are shown in Table S10. For TCSPC and confocal FRET measurements, we used Atto 488 and 647 maleimide as donor and acceptor fluorophores, respectively.

(D) Top: normalized DEER signal and fit with the Tikhonov regularization to derive the inter-spin distance distribution (bottom). The orange area represents the confidence interval of the Tikhonov regularization.

Supplemental information). In the WLC model, there is a single free parameter (stiffness of the chain κ) because we used the theoretical estimate from the number of amino acids and bond length between them to estimate the length of the WLC chain (L = 79 Å). The figure of merit χ^2_{WLC} = 1.19 models the decay with less free parameters than the 1- or 2-Gaussian distributed states in which each state requires 2 free parameters, the σ and mean $\langle R_{DA} \rangle$ of the distribution. Thus, we concluded that inter-dye distribution follows a WLC model. Next, we modeled a FRET line using the theoretical WLC with the expected chain length (L) (blue line in Figure 5C). The overlay of the WLC FRET line over the 2D MFD histogram captures the spread of the FRET efficiency distribution and the shift to the right of the static FRET line (red line). Thus, single-molecule events lie almost entirely in the center of the experimental distribution. Thus, we conclude that the WLC model is consistent with the TCSPC and the confocal smFRET measurement. Moreover, the green lines in the MFD histogram depict transitions between the $\langle R_{DA} \rangle_{WLC}$ and a NO FRET population, likely corresponding to acceptor photobleaching events.

Although the TCSPC and smFRET in confocal and DMD simulations agree with each other, there is an apparent disagreement between the TIRFM- and MFD-derived FRET efficiency histograms, which we resolved as follows. TIRFM FRET efficiency shows a smaller population at FRET efficiency of 0.3 (Figure 5A, bottom), which will correspond to an inter-dye distance of 58.7 Å that MFD cannot distinguish as completely separate population. However, when looking at the WLC model, the limit for the fully disordered and extended states would results in the intercepts

with the static FRET lines in MFD histogram corresponding to a FRET efficiency of 0.22. This FRET efficiency corresponds to an inter-dye distance of 60.5 Å, which is within error of the TIRFM-derived distance. If this state corresponds to an ensemble of extended disordered conformations, then molecules will sample over the ensemble of extended conformations every so often and will slowly transition back into a collapse disordered ensemble. This sampling could result in some TIRFM traces showing the extended configuration. Furthermore, the DEER decay and the corresponding inter-spin distribution (Figure 5D, top and bottom, respectively) between homo spin labels at the cysteine position (44/66) show a broad peak with a mean inter-spin distance of 33 Å, and a smaller population peaking at an inter-spin distance of 55.5 Å. This longer distance peak is consistent with both the TIRFM and confocal observation at lower FRET efficiencies (Figure S8). The fact that DEER shows distinct populations as TIRFM does resembles a phenomenon in which an apparent IDP may spontaneously switch between conformational ensembles as in previous reports.[48] Given that FRET depends on the position of the FRET labels, this behavior is likely representative of the sampling of different ensembles at different timescales. The ensembles could be representative of the different clusters with extended configurations identified in the DMD simulations.

Screening DMD simulations

To integrate the DMD simulations and the experimental observables, we used FRET-restrained positioning and screening (FPS)[34] and modeled the AV of the fluorophores (Alexa Fluor 647 and Alexa Fluor 488) and considered the last atom of the side chain of the amino acid as an approximation of the DEER label. We used derived distances from TCSPC decays for the SNAP-25 20/44 and 44/66 FRET variants due to its statistical benefit for illustration purposes. When screening, we computed the mean inter-dye distance for each screenshot sampled in the DMD simulation at 310 K. Figures 6A and 6B shows a 2D histogram correlating the AV determined inter-dye distance for the DMD snapshots against the computed α-helical content of the amino acids located in between the labels for both measured samples. When inspecting 1D $\langle R_{DA} \rangle$ histograms, we noticed that we could model them with a 2-Gaussian distribution, which we used to compare to experimentally determined distances. The 2 mean distances corresponding to 2 ensembles, $E_{1,AV}$ and $E_{2,AV}$, are within <6 Å from the experimental distances for the states E_1 and E_2, respectively (Table 1). Based on this agreement, we can assign each ensemble on the simulation to specific observables. For variant 44/66, we observed a broader monodisperse distribution due to the configurational heterogeneity of the probed region. The mean AV inter-dye distance was 39.2 Å, compared to 40.3 Å when modeling the time-resolved fluorescence decays with a Gaussian distributed state. Thus, the agreement in both variants captures a specific region of configuration and conformational dynamics.

Furthermore, by correlating the inter-dye distance and the α-helical content, we can inspect snapshots from various ensembles and visualize them. To do so, in Figures 6C and 6D, we show a representative structure of 2 ensembles (1 and 2) for each variant overlaying the corresponding AV for each dye for the 2 variants. The selected snapshots for the 20/44 depict a picture of fast helix-coli transition between low-energy barriers (~1–3 kcal/mol; Figure 3A) that is consistent with the dynamic averaging observed in faster than millisecond but slower than nanosecond timescales. From the snapshots of the 44/66 variant, we perceive a coil behavior consistent with a chain polymeric behavior on the nanosecond timescale.

It is worth mentioning that simulations do not fully capture the experimentally determined population fractions (Table 1); this is obvious in the 20/44 variant, in which the

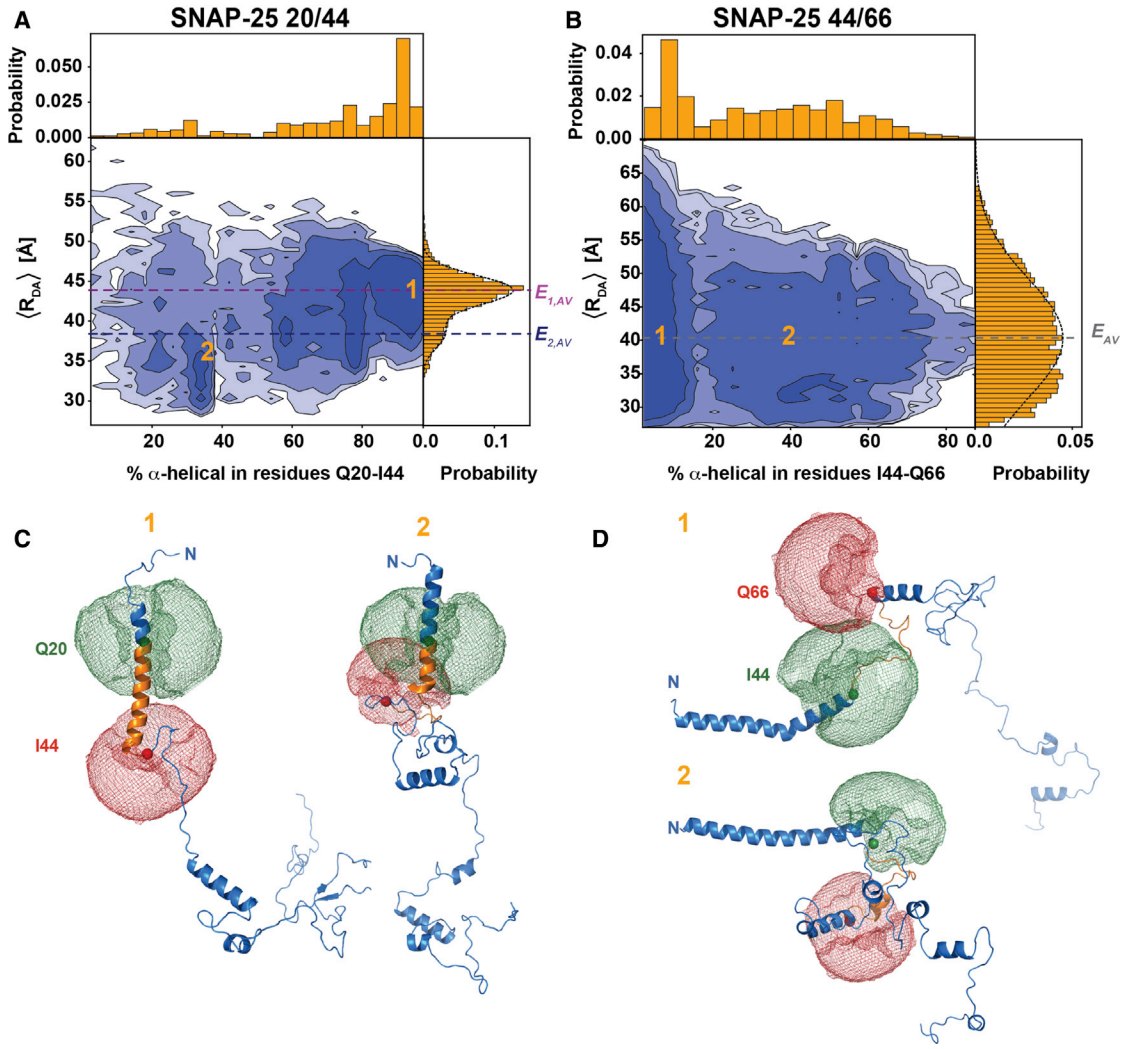


Figure 6. Integrating modeling of SNAP-25 using FRET distances and DMD simulations

(A and B) 2D histograms of the mean inter-dye distance ($\langle R_{DA} \rangle$) and the α-helical content of the amino acids within Q20 and I44 and I44 and Q66, respectively. The 1D projection histograms are shown at top and to the right.

(A) The $\langle R_{DA} \rangle$ distribution was fit with a 2-Gaussian distribution representing 2 ensembles, $E_{1,AV}$ and $E_{2,AV}$. Horizontal lines correspond to the mean inter-dye distance from each of the modeled Gaussians.

(B) A single Gaussian distribution models the inter-dye distribution with a mean E_{AV}.

(C and D) Snapshot models displaying the AV shown as clouds in green and red, representing the donor and acceptor fluorophores, respectively. Corresponding locations of the snapshots (1 and 2) at the 2D histogram are shown by the overlay numbers in (A) and (B). The variant 20/44 shows the helix-coil transition, while the 44/66 mostly shows coil behavior.

fraction for the shorter distance (33.3 Å) is 64% as determined by TCSPC, but from screening the population with a similar distance, it only shows a population fraction of 19%. This apparent discrepancy highlights the challenges of exactly matching simulations and experiments, including different environments and accuracies in both experimental measurements and force fields used in simulations. However, what is essential is identifying the basins or states from the analysis of energy landscape *in silico*, and the agreement in the highly heterogeneous configurational and conformational landscape of SNAP-25 is remarkable—within 6 Å.

The structural characterization of IDPs has become indispensable in cellular biology to help understand protein structure dynamics and the mechanisms that facilitate

their diverse cellular functions under biologically relevant conditions. In the limit of fast exchange between conformational substates (relaxation rate <10 μs), the measured spin relaxation rates report on population-weighted averages within each conformational substate on timescales up to nanoseconds.[49] Most of the time, this timescale is inadequate to quantitatively inspect the all-inclusive dynamics and accurate characterization of IDPs.[50] Incompatibility between the timescales accessible to MD simulation (order of a microsecond to thousands of microseconds) of the folding/unfolding pathways of proteins can be resolved by combining simulation with near-atomic level experiments.[51] To this end, experimental techniques (e.g., far UV CD,[52] small-angle X-ray scattering [SAXS],[53] NMR [chemical shifts, scalar ^{3}J-coupling, and residual dipolar couplings],[54,55] FRET,[56,57] electron paramagnetic resonance [EPR],[58–62] hydrodynamic radii determination methods),[63] have been useful in elucidating the structural information of disordered proteins. CD spectroscopy facilitates structural assessment of the residual secondary structure of disordered proteins under varying physiological conditions such as pH, ionic strength, solvent effects, or in presence of ligands.[52] NMR has advanced as one of the most widely used preeminent techniques in characterizing the solution structure and dynamics of IDPs in an aqueous environment and ensemble description of the conformational space.[15,64–67]

Intensity-based smFRET methods with temporal resolution ranging from 0.1 ms to >1,000 s can provide real-time information on rapidly interconverting conformational transitions and molecular interactions in IDP conformational ensembles with sensitivity to individual subpopulations.[68] The smFRET provides information on intramolecular distance by measuring the Förster transfer between donor and acceptor dyes attached to the polypeptide. Because the signal is recorded on single molecules, structural heterogeneity can be resolved that is often impossible to detect by ensemble-averaged methods.[69] Complementary to smFRET, fluorescence correlation spectroscopy (FCS), often done in a confocal system, integrates fluorescence fluctuations to provide measurements of dimensions and rapid molecular fluctuations in IDPs.[45,70] Nonetheless, experimental and spectroscopic methods are based on an ensemble average of conformations and undermine structural description of the subpopulations, thereby limiting identifying a single conformation to represent the disordered state.[71,72] Additional information is gained in FRET experiments when time-resolved fluorescence in TCSPC is monitored in addition to intensity observables,[73] thus, reaching the picosecond timescales. When TCSPC is added to single-molecule detection using intensity-based approaches, MFD is established,[74] particularly when multiple spectral windows and polarization of light are collected. EPR-like FRET is a label-based experimental approach that can model inter-spin distances and local and global motions of the labeled system.[75,76] The progressive development of pulse EPR, in particular the DEER technique, has facilitated the measurement of distances between spin-labeled sites in the 1.8- to 6.0-nm range.[77] Recent reviews have described the approaches and applications of EPR techniques to study large conformational changes in proteins and biomolecular associations.[58,76] Major differences are due to complexity in the experimental conditions, EPR experiments are mostly conducted in capillaries in ensemble conditions, and DEER measurements using nitroxide labels require low temperatures and done using frozen solutions. The systematic application of spin labeling and EPR identifies sequence-specific secondary structures, topology, and packing in the tertiary fold.[78] Also, due to concerns raised about the validity of individual spectroscopic methods for correctly quantifying the properties of unfolded proteins or IDPs, SAXS, NMR, and smFRET have only rarely been directly combined or compared.[69,79]

Dynamics simulation methods provide information on the time evolution of protein conformations and biological macromolecules, along with kinetic and thermodynamic information.[80] Computational methods such as Markov state models,[81] coarse-grained,[82] Monte Carlo simulation,[83] and temperature replica exchange[19,84,85] complement experimental techniques in elucidating realistic predictive models to circumvent the dynamic nature of interconverting ensembles and its interaction with other binding partners.[70] However, it is significant to note that conventional MD simulations hardly sample the complete free-energy landscape of proteins, as the simulated system can be trapped in local-minimum conformations. Thus, enhanced sampling methods such as replica-exchange MD[19] and metadynamics[86] methods have been developed to enhance the rugged folding landscape of proteins and IDPs.[87] Generalized ensemble methods such as rxDMD[88] coupled with the implicit solvent accelerates the sampling of the conformational dynamics of biomolecules[89] and have been shown to have high predictive power in capturing the structure and dynamics of folded protein, protein-protein interactions, and supertertiary structure of proteins *in vitro* and in live cells.[90–92]

In this study, atomistic rxDMD in combination with smFRET and DEER experiments are used to probe the structural heterogeneity and order-to-disorder transition in SNAP-25, an important prototypical neuronal IDP that couples a disorder-to-order transition to the fusion of synaptic vesicles with the plasma membrane.[6,26] Using WHAM analysis of rxDMD simulations, we identified the optimal temperature of 310 K slightly below the coil-to-molten globular transition, at which the computed R_g, R_H, secondary structure contents, and the correspondingly estimated ellipticity match previous experiments. At this temperature, the structures of SNAP 25 are rich in coil with residual helices, which undergo frequent unfolding and refolding dynamics. Our results reveal the agreement between rxDMD simulations and label-based experiments to capture a novel conformational switching behavior in SNAP-25. SNAP-25 retains residual and transient secondary structural elements compatible with the bound state. Since the protein lacks persistent tertiary interactions and the order-to-disorder transitions are limited to local helices, we monitor the dynamics of the first 2 helices, residues 1–40 and 65–80. AV screening of the structural ensemble sampled by rxDMD simulations at 310 K shows a good agreement with experiments. A 2-Gaussian inter-dye distance distribution confirms the presence of at least 2 states between Q20/I44 pairs exemplifying conformational dynamics within the N-terminal helical region. However, the broad distance distribution between I44/Q66 residue pairs confirms this region to be a mostly disordered with transient residual structure. Nevertheless, we point out that resolving these populations is extremely challenging, apparently due to the spontaneous interchange between conformational ensembles. Given that FRET measurements strongly depend on the position of FRET dye labels, the behavior is likely representative of the sampling at different timescales.

These results suggest future utility of this integrative approach, as we have made possible (1) accurate structural characterization of a moderately sized neuronal IDP in submillisecond timescales, (2) identification of the dynamic "limiting states" in the conformational exchange processes, (3) decoupling the helix-coil transition in the N-terminal α-helix region that dominates its structural properties, and (4) establishment of a workflow that successfully allows the conjunction of simulation with experiment to study the conformational changes occurring in protein at the atomic level. What makes our method reliable is the quantitatively consistent agreement for R_g and other properties of atomistic simulations to experiments. This is of the essence, as the R_g obtained with most force fields in explicit solvent

simulations are much below the experimental value,[93] which makes it essentially impossible to quantitatively compare between FRET and SAXS experiment and simulation.[94]

In summary, we systematically identify the novel molecular determinants in the conformational switching behavior of SNAP-25; while being unstructured to enable efficient scavenging of binding partners, the IDP still dynamically retains residual secondary structures compatible with a bound state, which is essential in synaptic transmission and membrane fusion. To our knowledge, this is the first demonstration of capturing a stochastic dynamic switching behavior of the native state of SNAP-25 on a much faster timescale from the combination of simulation with single-molecule experiments. The increasing convergence of timescales in experiments and simulations enable an increasingly reliable interpretation of experimental observables based on molecular simulations. Notably, this synergy in our techniques overcomes the limitation of each individual approach and augments standard structural characterization methods to allow us to examine the detailed complexities of IDPs and IDRs.

EXPERIMENTAL PROCEDURES

Resource availability

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Feng Ding (fding@clemson.edu).

Materials availability

This research did not generate any unique materials.

Data and code availability

The authors declare that the data supporting the findings in the study are available within the article and the Supplemental information. All other data are available from the lead contact upon reasonable request. The DMD simulation engine is available at http://www.moleculesinaction.com. The computably obtained conformational ensemble of SNAP-25 is publicly available from the lab website (https://dlab.clemson.edu/research/SNAP-25/). Software for analysis of single-molecule experiments, written in-house, can be downloaded from https://www.mpc.hhu.de/software.html. The fluorescence decay histograms were analyzed using ChiSurf (https://github.com/Fluorescence-Tools/chisurf).

Materials and methods

See the Supplemental experimental procedures for full details of the molecular model and simulation of SNAP-25, sample preparation, EPR spectroscopy, confocal smFRET in MFD mode, PDA, ensemble TCSPC, time-resolved fluorescence decay analysis, κ^2 and error propagation, screening molecular dynamics simulations, and comparison of EPR and FRET distances.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xcrp.2021.100616.

ACKNOWLEDGMENTS

F.D. acknowledges support from NSF CBET-1553945 and NIH R35GM119691. H.S. acknowledges support from Clemson University start-up funds, the NSF (CAREER

MCB-1749778), and the NIH (MH081923 and 1P20GM12134201). T.S. acknowledges partial support from the NSF (CHE-1508607). S.M. acknowledges the partial support of the US Department of Energy (DOE) (contract no. DE-FG02-02ER15354) for DEER experiments. EPR instrumentation at NCSU was supported by grants from the NIH (no. RR023614), the NSF (No. CHE-0840501), and North Carolina Biotechnolgy Center (no. 2009-IDG-1015).

AUTHOR CONTRIBUTIONS

F.D. and N.S. conceived and designed the DMD simulations. F.D. guided the computational modeling and simulations. N.S. performed the simulations and analyzed the data. I.S.Y.-O., G.L.H., and H.S. measured and analyzed the smFRET in confocal MFD mode. R.Q., P.H., and K.R.W. produced the protein samples and measured and analyzed the smFRET in TIRFM. S.M., E.O., and T.S. measured and analyzed the DEER experiments. N.S., I.S.Y.-O., H.S., and F.D. drafted the manuscript. K.R.W. and T.S. contributed to the editing of the manuscript. K.R.W., T.S., H.S., and F.D. advised on the overall project.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 30, 2021
Revised: August 24, 2021
Accepted: September 28, 2021
Published: October 15, 2021

REFERENCES

1. Dobson, C.M. (2003). Protein folding and misfolding. Nature *426*, 884–890.
2. Cheng, Y., LeGall, T., Oldfield, C.J., Dunker, A.K., Uversky, V.N., Gillespie, J.R., and Fink, A.L. (2006). Abundance of intrinsic disorder in protein associated with cardiovascular disease. Biochemistry *45*, 10448–10460.
3. Uversky, V.N., Davé, V., Iakoucheva, L.M., Malaney, P., Metallo, S.J., Pathak, R.R., and Joerger, A.C. (2014). Pathological unfoldomics of uncontrolled chaos: intrinsically disordered proteins and human diseases. Chem. Rev. *114*, 6844–6879.
4. Uversky, V.N. (2015). The multifaceted roles of intrinsic disorder in protein complexes. FEBS Lett. *589* (19 Pt A), 2498–2506.
5. Shammas, S.L. (2017). Mechanistic roles of protein disorder within transcription. Curr. Opin. Struct. Biol. *42*, 155–161.
6. Snead, D., and Eliezer, D. (2019). Intrinsically disordered proteins in synaptic vesicle trafficking and release. J. Biol. Chem. *294*, 3325–3342.
7. Paz, A., Zeev-Ben-Mordehai, T., Lundqvist, M., Sherman, E., Mylonas, E., Weiner, L., Haran, G., Svergun, D.I., Mulder, F.A., Sussman, J.L., and Silman, I. (2008). Biophysical characterization of the unstructured cytoplasmic domain of the human neuronal adhesion protein neuroligin 3. Biophys. J. *95*, 1928–1944.
8. Kiss-Tóth, A., Dobson, L., Péterfia, B., Ángyán, A.F., Ligeti, B., Lukács, G., and Gáspári, Z. (2019). Occurrence of ordered and disordered structural elements in postsynaptic proteins supports optimization for interaction diversity. Entropy (Basel) *21*, 761.
9. Minezaki, Y., Homma, K., and Nishikawa, K. (2007). Intrinsically disordered regions of human plasma membrane proteins preferentially occur in the cytoplasmic segment. J. Mol. Biol. *368*, 902–913.
10. Stavropoulos, I., Khaldi, N., Davey, N.E., O'Brien, K., Martin, F., and Shields, D.C. (2012). Protein disorder and short conserved motifs in disordered regions are enriched near the cytoplasmic side of single-pass transmembrane proteins. PLoS ONE *7*, e44389.
11. Wright, P.E., and Dyson, H.J. (2015). Intrinsically disordered proteins in cellular signalling and regulation. Nat. Rev. Mol. Cell Biol. *16*, 18–29.
12. Dyson, H.J., and Wright, P.E. (2005). Intrinsically unstructured proteins and their functions. Nat. Rev. Mol. Cell Biol. *6*, 197–208.
13. Bürgi, J., Xue, B., Uversky, V.N., and van der Goot, F.G. (2016). Intrinsic disorder in transmembrane proteins: roles in signaling and topology prediction. PLoS ONE *11*, e0158594.
14. Kleckner, I.R., and Foster, M.P. (2011). An introduction to NMR-based approaches for measuring protein dynamics. Biochim. Biophys. Acta *1814*, 942–968.
15. Adamski, W., Salvi, N., Maurin, D., Magnat, J., Milles, S., Jensen, M.R., Abyzov, A., Moreau, C.J., and Blackledge, M. (2019). A unified description of intrinsically disordered protein dynamics under physiological conditions using NMR spectroscopy. J. Am. Chem. Soc. *141*, 17817–17829.
16. Dyson, H.J., and Wright, P.E. (2019). Perspective: the essential role of NMR in the discovery and characterization of intrinsically disordered proteins. J. Biomol. NMR *73*, 651–659.
17. Nasir, I., Onuchic, P.L., Labra, S.R., and Deniz, A.A. (2019). Single-molecule fluorescence studies of intrinsically disordered proteins and liquid phase separation. Biochim. Biophys. Acta Proteins Proteomics *1867*, 980–987.
18. Yanez Orozco, I.S., Mindlin, F.A., Ma, J., Wang, B., Levesque, B., Spencer, M., Rezaei Adariani, S., Hamilton, G., Ding, F., Bowen, M.E., and Sanabria, H. (2018). Identifying weak interdomain interactions that stabilize the supertertiary structure of the N-terminal tandem PDZ domains of PSD-95. Nat. Commun. *9*, 3724.
19. Sugita, Y., and Okamoto, Y. (1999). Replica-exchange molecular dynamics method for protein folding. Chem. Phys. Lett. *314*, 141–151.
20. Dokholyan, N.V. (2020). Experimentally-driven protein structure modeling. J. Proteomics *220*, 103777.
21. Chen, J., Zaer, S., Drori, P., Zamel, J., Joron, K., Kalisman, N., Lerner, E., and Dokholyan, N.V. (2021). The structural heterogeneity of α-synuclein is governed by several distinct subpopulations with interconversion times slower than milliseconds. Structure *29*, 1048–1064.e6.

22. Chen, X., Tomchick, D.R., Kovrigin, E., Araç, D., Machius, M., Südhof, T.C., and Rizo, J. (2002). Three-dimensional structure of the complexin/SNARE complex. Neuron *33*, 397–409.

23. Diao, J., Liu, R., Rong, Y., Zhao, M., Zhang, J., Lai, Y., Zhou, Q., Wilz, L.M., Li, J., Vivona, S., et al. (2015). ATG14 promotes membrane tethering and fusion of autophagosomes to endolysosomes. Nature *520*, 563–566.

24. Strop, P., Kaiser, S.E., Vrljic, M., and Brunger, A.T. (2008). The structure of the yeast plasma membrane SNARE complex reveals destabilizing water-filled cavities. J. Biol. Chem. *283*, 1113–1119.

25. Ramakrishnan, N.A., Drescher, M.J., and Drescher, D.G. (2012). The SNARE complex in neuronal and sensory cells. Mol. Cell. Neurosci. *50*, 58–69.

26. Yang, H., Zhang, M., Shi, J., Zhou, Y., Wan, Z., Wang, Y., Wan, Y., Li, J., Wang, Z., and Fei, J. (2017). Brain-specific SNAP-25 deletion leads to elevated extracellular glutamate level and schizophrenia-like behavior in mice. Neural Plast. *2017*, 4526417.

27. Zhang, X., Rebane, A.A., Ma, L., Li, F., Jiao, J., Qu, H., Pincet, F., Rothman, J.E., and Zhang, Y. (2016). Stability, folding dynamics, and long-range conformational transition of the synaptic t-SNARE complex. Proc. Natl. Acad. Sci. USA *113*, E8031–E8040.

28. Ungermann, C., and Langosch, D. (2005). Functions of SNAREs in intracellular membrane fusion and lipid bilayer mixing. J. Cell Sci. *118*, 3819–3828.

29. Choi, U.B., McCann, J.J., Weninger, K.R., and Bowen, M.E. (2011). Beyond the random coil: stochastic conformational switching in intrinsically disordered proteins. Structure *19*, 566–576.

30. Fasshauer, D., Bruns, D., Shen, B., Jahn, R., and Brünger, A.T. (1997). A structural change occurs upon binding of syntaxin to SNAP-25. J. Biol. Chem. *272*, 4582–4590.

31. Kumar, S., Bouzida, D., Swendsen, R.H., Kollman, P.A., and Rosenberg, J.M. (1992). The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. J. Comput. Chem. *13*, 1011–1021.

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

33. Dimura, M., Peulen, T.O., Hanke, C.A., Prakash, A., Gohlke, H., and Seidel, C.A. (2016). Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems. Curr. Opin. Struct. Biol. *40*, 163–185.

34. Kalinin, S., Peulen, T., Sindbert, S., Rothwell, P.J., Berger, S., Restle, T., Goody, R.S., Gohlke, H., and Seidel, C.A.M. (2012). A toolkit and benchmark study for FRET-restrained high-precision structural modeling. Nat. Methods *9*, 1218–1225.

35. Boura, E., Rózycki, B., Herrick, D.Z., Chung, H.S., Vecer, J., Eaton, W.A., Cafiso, D.S., Hummer, G., and Hurley, J.H. (2011). Solution structure of the ESCRT-I complex by small-angle X-ray scattering, EPR, and FRET spectroscopy. Proc. Natl. Acad. Sci. USA *108*, 9437–9442.

36. Jeschke, G. (2012). DEER distance measurements on proteins. Annu. Rev. Phys. Chem. *63*, 419–446.

37. Medina, E., Villalobos, P., Hamilton, G.L., Komives, E.A., Sanabria, H., Ramírez-Sarmiento, C.A., and Babul, J. (2020). Intrinsically disordered regions of the DNA-binding domain of human FoxP1 facilitate domain swapping. J. Mol. Biol. *432*, 5411–5429.

38. Tsytlonok, M., Hemmen, K., Hamilton, G., Kolimi, N., Felekyan, S., Seidel, C.A.M., Tompa, P., and Sanabria, H. (2020). Specific conformational dynamics and expansion underpin a multi-step mechanism for specific binding of p27 with Cdk2/Cyclin A. J. Mol. Biol. *432*, 2998–3017.

39. Ma, J., Saikia, N., Godar, S., Hamilton, G.L., Ding, F., Alper, J., and Sanabria, H. (2021). Ensemble switching unveils a kinetic rheostat mechanism of the eukaryotic thiamine pyrophosphate riboswitch. RNA *27*, 771–790.

40. English, L.R., Tilton, E.C., Ricard, B.J., and Whitten, S.T. (2017). Intrinsic α helix propensities compact hydrodynamic radii in intrinsically disordered proteins. Proteins *85*, 296–311.

41. Lindorff-Larsen, K., Kristjansdottir, S., Teilum, K., Fieber, W., Dobson, C.M., Poulsen, F.M., and Vendruscolo, M. (2004). Determination of an ensemble of structures representing the denatured state of the bovine acyl-coenzyme a binding protein. J. Am. Chem. Soc. *126*, 3291–3299.

42. Ortega, A., Amorós, D., and García de la Torre, J. (2011). Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. Biophys. J. *101*, 892–898.

43. Sanabria, H., Rodnin, D., Hemmen, K., Peulen, T.O., Felekyan, S., Fleissner, M.R., Dimura, M., Koberling, F., Kühnemuth, R., Hubbell, W., et al. (2020). Resolving dynamics and function of transient states in single enzyme molecules. Nat. Commun. *11*, 1231.

44. Ma, J., Yanez-Orozco, I.S., Rezaei Adariani, S., Dolino, D., Jayaraman, V., and Sanabria, H. (2017). High precision FRET at single-molecule level for biomolecule structure determination. J. Vis. Exp. *123*, 55623.

45. Hamilton, G., and Sanabria, H. (2019). Multiparameter fluorescence spectroscopy of single molecules. In Spectroscopy and Dynamics of Single Molecules: Methods and Applications, C.K. Johnson, ed. (Elsevier), pp. 269–333.

46. Barth, A., Opanasyuk, O., Peulen, T.-O., Felekyan, S., Kalinin, S., Sanabria, H., and Seidel, C.A. (2021). Unraveling multi-state molecular dynamics in single-molecule FRET experiments-Part I: theory of FRET-lines. arXiv, 2107.14770v1. http://arxiv.org/abs/2107.14770.

47. Opanasyuk, O., Barth, A., Peulen, T.-O., Felekyan, S., Kalinin, S., Sanabria, H., and Seidel, C.A. (2021). Unraveling multi-state molecular dynamics in single-molecule FRET experiments–Part II: Quantitative analysis of multi-state kinetic networks. arXiv, 2107.14791v1. http://arxiv.org/abs/2107.14791.

48. Choi, U.B., Sanabria, H., Smirnova, T., Bowen, M.E., and Weninger, K.R. (2019). Spontaneous Switching among Conformational Ensembles in Intrinsically Disordered Proteins. Biomolecules *9*, 114.

49. Salvi, N., Abyzov, A., and Blackledge, M. (2019). Solvent-dependent segmental dynamics in intrinsically disordered proteins. Sci. Adv. *5*, eaax2348.

50. Otosu, T., Ishii, K., and Tahara, T. (2015). Microsecond protein dynamics observed at the single-molecule level. Nat. Commun. *6*, 7685.

51. Fersht, A.R. (2002). On the simulation of protein folding by short time scale molecular dynamics and distributed computing. Proc. Natl. Acad. Sci. USA *99*, 14122–14125.

52. Chemes, L.B., Alonso, L.G., Noval, M.G., and de Prat-Gay, G. (2012). Circular dichroism techniques for the analysis of intrinsically disordered proteins and domains. Methods Mol. Biol. *895*, 387–404.

53. Bernadó, P., and Svergun, D.I. (2012). Analysis of intrinsically disordered proteins by small-angle X-ray scattering. Methods Mol. Biol. *896*, 107–122.

54. Dyson, H.J., and Wright, P.E. (2004). Unfolded proteins and protein folding studied by NMR. Chem. Rev. *104*, 3607–3622.

55. Jensen, M.R., Zweckstetter, M., Huang, J.R., and Blackledge, M. (2014). Exploring free-energy landscapes of intrinsically disordered proteins at atomic resolution using NMR spectroscopy. Chem. Rev. *114*, 6632–6660.

56. Soranno, A., Buchli, B., Nettels, D., Cheng, R.R., Müller-Späth, S., Pfeil, S.H., Hoffmann, A., Lipman, E.A., Makarov, D.E., and Schuler, B. (2012). Quantifying internal friction in unfolded and intrinsically disordered proteins with single-molecule spectroscopy. Proc. Natl. Acad. Sci. USA *109*, 17800–17806.

57. Hofmann, H., Soranno, A., Borgia, A., Gast, K., Nettels, D., and Schuler, B. (2012). Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. Proc. Natl. Acad. Sci. USA *109*, 16155–16160.

58. Le Breton, N., Martinho, M., Mileo, E., Etienne, E., Gerbaud, G., Guigliarelli, B., and Belle, V. (2015). Exploring intrinsically disordered proteins using site-directed spin labeling electron paramagnetic resonance spectroscopy. Front. Mol. Biosci. *2*, 21.

59. Cattani, J., Subramaniam, V., and Drescher, M. (2017). Room-temperature in-cell EPR spectroscopy: alpha-Synuclein disease variants remain intrinsically disordered in the cell. Phys. Chem. Chem. Phys. *19*, 18147–18151.

60. Habchi, J., Martinho, M., Gruet, A., Guigliarelli, B., Longhi, S., and Belle, V. (2012). Monitoring structural transitions in IDPs by site-directed spin labeling EPR spectroscopy. Methods Mol. Biol. *895*, 361–386.

61. Pirman, N.L., Milshteyn, E., Galiano, L., Hewlett, J.C., and Fanucci, G.E. (2011). Characterization of the disordered-to-α-helical transition of IA_3

by SDSL-EPR spectroscopy. Protein Sci. *20*, 150–159.

62. Drescher, M. (2012). EPR in Protein Science: Intrinsically Disordered Proteins. Top. Curr. Chem. *321*, 91–119.

63. Tomasso, M.E., Tarver, M.J., Devarajan, D., and Whitten, S.T. (2016). Hydrodynamic radii of intrinsically disordered proteins determined from experimental polyproline II propensities. PLoS Comput. Biol. *12*, e1004686.

64. Felli, I.C., and Pierattelli, R. (2015). Intrinsically disordered proteins studied by NMR spectroscopy. Adv. Exp. Med. Biol. *870*, 149–185.

65. Gil, S., Hošek, T., Solyom, Z., Kümmerle, R., Brutscher, B., Pierattelli, R., and Felli, I.C. (2013). NMR spectroscopic studies of intrinsically disordered proteins at near-physiological conditions. Angew. Chem. Int. Ed. Engl. *52*, 11808–11812.

66. Gibbs, E.B., Cook, E.C., and Showalter, S.A. (2017). Application of NMR to studies of intrinsically disordered proteins. Arch. Biochem. Biophys. *628*, 57–70.

67. Brookes, D.H., and Head-Gordon, T. (2016). Experimental inferential structure determinator of ensembles for intrinsically disordered proteins. J. Am. Chem. Soc. *138*, 4530–4538.

68. LeBlanc, S.J., Kulkarni, P., and Weninger, K.R. (2018). Single molecule FRET: a powerful tool to study intrinsically disordered proteins. Biomolecules *8*, 140.

69. Aznauryan, M., Delgado, L., Soranno, A., Nettels, D., Huang, J.R., Labhardt, A.M., Grzesiek, S., and Schuler, B. (2016). Comprehensive structural and dynamical view of an unfolded protein from the combination of single-molecule FRET, NMR, and SAXS. Proc. Natl. Acad. Sci. USA *113*, E5389–E5398.

70. Brucale, M., Schuler, B., and Samorì, B. (2014). Single-molecule studies of intrinsically disordered proteins. Chem. Rev. *114*, 3281–3317.

71. Cino, E.A., Wong-ekkabut, J., Karttunen, M., and Choy, W.Y. (2011). Microsecond molecular dynamics simulations of intrinsically disordered proteins involved in the oxidative stress response. PLoS ONE *6*, e27371.

72. Navarro-Retamal, C., Bremer, A., Alzate-Morales, J., Caballero, J., Hincha, D.K., González, W., and Thalhammer, A. (2016). Molecular dynamics simulations and CD spectroscopy reveal hydration-induced unfolding of the intrinsically disordered LEA proteins COR15A and COR15B from Arabidopsis thaliana. Phys. Chem. Chem. Phys. *18*, 25806–25816.

73. Duncan, R.R., Bergmann, A., Cousin, M.A., Apps, D.K., and Shipston, M.J. (2004). Multi-dimensional time-correlated single photon counting (TCSPC) fluorescence lifetime imaging microscopy (FLIM) to detect FRET in cells. J. Microsc. *215*, 1–12.

74. Michalet, X., Colyer, R.A., Scalia, G., Ingargiola, A., Lin, R., Millaud, J.E., Weiss, S., Siegmund, O.H., Tremsin, A.S., Vallerga, J.V., et al. (2012). Development of new photon-counting detectors for single-molecule fluorescence microscopy. Philos. Trans. R. Soc. Lond. B Biol. Sci. *368*, 20120035.

75. Hubbell, W.L., Cafiso, D.S., and Altenbach, C. (2000). Identifying conformational changes with site-directed spin labeling. Nat. Struct. Biol. *7*, 735–739.

76. Hubbell, W.L., López, C.J., Altenbach, C., and Yang, Z. (2013). Technological advances in site-directed spin labeling of proteins. Curr. Opin. Struct. Biol. *23*, 725–733.

77. Pannier, M., Veit, S., Godt, A., Jeschke, G., and Spiess, H.W. (2000). Dead-time free measurement of dipole-dipole interactions between electron spins. J. Magn. Reson. *142*, 331–340.

78. McHaourab, H.S., Steed, P.R., and Kazmier, K. (2011). Toward the fourth dimension of membrane protein structure: insight into dynamics from spin-labeling EPR spectroscopy. Structure *19*, 1549–1561.

79. Boomsma, W., Ferkinghoff-Borg, J., and Lindorff-Larsen, K. (2014). Combining experiments and simulations using the maximum entropy principle. PLoS Comput. Biol. *10*, e1003406.

80. Adcock, S.A., and McCammon, J.A. (2006). Molecular dynamics: survey of methods for simulating the activity of proteins. Chem. Rev. *106*, 1589–1615.

81. Shukla, D., Hernández, C.X., Weber, J.K., and Pande, V.S. (2015). Markov state models provide insights into dynamic modulation of protein function. Acc. Chem. Res. *48*, 414–422.

82. Cragnell, C., Rieloff, E., and Skepö, M. (2018). Utilizing coarse-grained modeling and monte carlo simulations to evaluate the conformational ensemble of intrinsically disordered proteins and regions. J. Mol. Biol. *430*, 2478–2492.

83. Bhowmick, A., and Head-Gordon, T. (2015). A monte carlo method for generating side chain structural ensembles. Structure *23*, 44–55.

84. Zerze, G.H., Miller, C.M., Granata, D., and Mittal, J. (2015). Free energy surface of an intrinsically disordered protein: comparison between temperature replica exchange molecular dynamics and bias-exchange metadynamics. J. Chem. Theory Comput. *11*, 2776–2782.

85. Kulke, M., Geist, N., Möller, D., and Langel, W. (2018). Replica-based protein structure sampling methods: compromising between explicit and implicit solvents. J. Phys. Chem. B *122*, 7295–7307.

86. Piana, S., and Laio, A. (2007). A bias-exchange approach to protein folding. J. Phys. Chem. B *111*, 4553–4559.

87. Zhang, B.W., Dai, W., Gallicchio, E., He, P., Xia, J., Tan, Z., and Levy, R.M. (2016). Simulating replica exchange: Markov state models, proposal schemes, and the infinite swapping limit. J. Phys. Chem. B *120*, 8289–8301.

88. Shirvanyants, D., Ding, F., Tsao, D., Ramachandran, S., and Dokholyan, N.V. (2012). Discrete molecular dynamics: an efficient and versatile simulation method for fine protein characterization. J. Phys. Chem. B *116*, 8375–8382.

89. Anandakrishnan, R., Drozdetski, A., Walker, R.C., and Onufriev, A.V. (2015). Speed of conformational change: comparing explicit and implicit solvent molecular dynamics simulations. Biophys. J. *108*, 1153–1164.

90. Dokholyan, N.V., Buldyrev, S.V., Stanley, H.E., and Shakhnovich, E.I. (2000). Identifying the protein folding nucleus using molecular dynamics. J. Mol. Biol. *296*, 1183–1188.

91. Brodie, N.I., Popov, K.I., Petrotchenko, E.V., Dokholyan, N.V., and Borchers, C.H. (2017). Solving protein structures using short-distance cross-linking constraints as a guide for discrete molecular dynamics simulations. Sci. Adv. *3*, e1700479.

92. Basak, S., Sakia, N., Dougherty, L., Guo, Z., Wu, F., Mindlin, F., Lary, J.W., Cole, J.L., Ding, F., and Bowen, M.E. (2021). Probing interdomain linkers and protein supertertiary structure in vitro and in live cells with fluorescent protein resonance energy transfer. J. Mol. Biol. *433*, 166793.

93. Piana, S., Donchev, A.G., Robustelli, P., and Shaw, D.E. (2015). Water dispersion interactions strongly influence simulated structural properties of disordered protein states. J. Phys. Chem. B *119*, 5113–5123.

94. Best, R.B. (2020). Emerging consensus on the collapse of unfolded and intrinsically disordered proteins in water. Curr. Opin. Struct. Biol. *60*, 27–38.