Title
Protein Dynamics to Define and Refine Disordered Protein Ensembles

Journal
The Journal of Physical Chemistry B, 126(9)

ISSN
1520-6106

Authors
Naullage, Pavithra M
Haghighatlari, Mojtaba
Namini, Ashley
et al.

Publication Date
2022-03-10

DOI
10.1021/acs.jpcb.1c10925

Peer reviewed

HHS Public Access

Author manuscript

J Phys Chem B. Author manuscript; available in PMC 2023 April 23.

Published in final edited form as:

J Phys Chem B. 2022 March 10; 126(9): 1885–1894. doi:10.1021/acs.jpcb.1c10925.

Protein Dynamics to Define and Refine Disordered Protein Ensembles

Pavithra M. Naullage[1,2], **Mojtaba Haghighatlari**[1,2], **Ashley Namini**[4], **João M. C. Teixeira**[4], **Jie Li**[1,2], **Oufan Zhang**[1,2], **Claudiu C. Gradinaru**[7], **Julie D. Forman-Kay**[4,5], **Teresa Head-Gordon**[1,2,3,6,*]

[1]Pitzer Center for Theoretical Chemistry, University of California, Berkeley, CA 94720, USA.

[2]Department of Chemistry, University of California, Berkeley, California, USA

[3]Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720, USA.

[4]Molecular Medicine Program, Hospital for Sick Children, Toronto, Ontario M5G 0A4, Canada

[5]Department of Biochemistry, University of Toronto, Toronto, Ontario, M5S 1A8, Canada

[6]Department of Bioengineering, University of California, Berkeley, California, USA

[7]Department of Chemical and Physical Sciences, University of Toronto Mississauga, Mississauga, Ontario, L5L 1C6, Canada

Abstract

Intrinsically disordered proteins and unfolded proteins have fluctuating conformational ensembles that are fundamental to their biological function and impact protein folding, stability and misfolding. Despite the importance of protein dynamics and conformational sampling, time-dependent data types are not fully exploited when defining and refining disordered protein ensembles. Here we introduce a computational framework using an elastic network model and normal mode displacements to generate a dynamic disordered ensemble consistent with NMR-derived dynamics parameters, including transverse

R_2

relaxation rates and Lipari-Szabo order parameters

(

S^2

values). We illustrate our approach using the unfolded state of the drkN SH3 domain to show that the dynamical ensembles give better agreement than a static ensemble for a wide range of experimental validation data including NMR chemical shifts, J-couplings, nuclear Overhauser effects, paramagnetic relaxation enhancements, residual dipolar couplings, hydrodynamic radii, single-molecule fluorescence Förster resonance energy transfer, and small-angle X-ray scattering.

*Corresponding author, thg@berkeley.edu.

SUPPORTING INFORMATION. Different evaluations of the optimized S2-selected and R2-selected ensembles and original ensembles for ENSEMBLE, RANDOM, and MIXED pools with 8 experimental data types

INTRODUCTION

It has long been perceived that proteins need to adopt a well-defined three-dimensional structure to carry out their function, although it is increasingly clear that all proteomes encode proteins that do not adopt stable three-dimensional structure that are nevertheless important for cellular function.[1–10] Proteins that do not possess stable structural domains, known as intrinsically disordered proteins (IDPs), or contain intrinsically disordered regions (IDRs), play a central role in signaling, transcription, and regulation of cell cycle, as well as in the formation of biomolecular condensates in the cell.[11–22] In addition, unfolded states of folded proteins are increasingly appreciated as critical for understanding protein folding and stability[23–25]. Finally, both IDPs/IDRs and unfolded proteins can participate in pathological aggregation or fiber formation, with strong dependencies on the conformational equilibria in the disordered ensembles.[26–27]

Because IDPs/IDRs and unfolded states have fluctuating heterogeneous conformations under physiological conditions, it necessitates detailed investigation into the underlying conformational ensembles to enable structural correlations with function.[14, 28–29] There are a number of solution experimental techniques that can guide the construction of structural ensembles, many of which encompass key features of disordered states: fractional local structural propensities as measured by NMR chemical shifts (CSs)[28, 30–32], J-couplings (JCs)[33], and residual dipolar couplings (RDCs) [34–37]; global properties as measured through NMR hydrodynamic radii values[38] and small angle X-ray scattering (SAXS) data[39–42]; and tertiary contact information as measured through NMR paramagnetic relaxation enhancements (PREs)[43–44], nuclear Overhauser effects (NOEs)[28] and single-molecule fluorescence Förster resonance energy transfer (smFRET)[29, 45].

However, due to the fast dynamics in the disordered state, the measured NMR, SAXS and smFRET observables are highly averaged, resulting in significant challenges to bridge the connection between experimental observables and structural ensembles.[5, 45–46] Thus computational models play a critical role in constructing structural ensembles of disordered states by first creating an initial library of putative conformations, and then a selected subset is chosen based on improving the agreement with the available experimental solution data. Most computational efforts have focused on the ensemble subset selection process, such as the ENSEMBLE program[47] that uses a Monte Carlo algorithm or the ASTEROIDS approach with its evolutionary algorithm[48], to select a set of conformations for which the back-calculated data fit the available experimental data. Head-Gordon and co-workers have introduced the extended Experimental Inferential Structure Determination (X-EISD) procedure that uses a variety of experimental observables, and their known errors and variances, to determine the most probable structural ensemble from candidate structures for disordered states.[33] This Bayesian statistical method calculates log-likelihood scores of the selected ensemble corresponding to a set of experimental data and thus can be used to refine a structural ensemble of disordered states.[33] Often these approaches have relied on a somewhat arbitrary formulation of the underlying structural ensemble before optimization, such as use of the TraDES conformer generator.[49]

In addition to structural characterization, protein dynamics is fundamental in understanding function of both ordered and disordered states of proteins[50], and NMR techniques have been used to investigate dynamics of proteins through measurement of spin relaxation properties.[51] Heteronuclear spin relaxation studies typically measure internuclear bond vector dynamics (^{1}H-^{15}N) by measuring ^{15}N longitudinal R_1 and transverse R_2 relaxation rates along with the heteronuclear ^{1}H-^{15}N NOE, potentially augmented by other relaxation experiments focused on ^{15}N, ^{13}C or ^{2}H nuclei. These relaxation data can also be expressed in terms of spectral density functions and are often translated into model free Lipari-Szabo order parameters (S^2 values).[51–52] The magnitude of S^2 values can vary from 0 to 1, corresponding to completely isotropic large amplitude internal motions or complete rigidity, respectively. The model free formalism assumes that the timescales of internal motion and overall rotation are very different[53] and provide a reasonable expectation for folded states but not for disordered proteins. Thus, most NMR relaxation studies of disordered proteins use raw rates or spectral density function mapping.[52, 54–57]

While protein dynamics can potentially be of great importance for an accurate description of disordered protein conformational ensembles, there has been little effort to include dynamics data in computational tools for creating and/or refining disordered state ensembles. Molecular simulations such as molecular dynamics (MD) or Monte Carlo (MC) methods can elucidate information on protein motions.[37, 58–60] Molecular simulations such as molecular dynamics (MD) or Monte Carlo (MC) methods can elucidate information on protein motions[37, 58–60], however fixed charge force fields have been shown to be unreliable for IDPs and better polarizable force fields are costly[10]. Coarse-grained elastic network models (ENMs), on the other hand, offer simple and computationally efficient alternatives to MD or MC simulations for sampling the intrinsic motions accessible to a protein, and have been widely used to study the intrinsic dynamics of folded systems from small globular proteins to large biomolecular assemblies.[61–72] The underlying principal is that dynamic properties of globular proteins are determined mainly by the protein topology, that can be captured as a collection of harmonic springs between native contacts, such that the most important global motions are captured.[73–75] It has been shown that the normal modes generated through ENM models can accurately predict the crystallographic B-factors as well as NMR dynamics data for folded proteins.[56, 61] However, to the best of our knowledge, ENMs have not been used to explore the intrinsic dynamics of disordered proteins.

In this work, we introduce a framework for using an ENM and subsequent normal mode displacements for the creation and optimized selection of a disordered protein ensemble consistent with

R_2 rates and S^2 values for the unfolded state of the N-terminal SH3 domain of the Drosophila drk protein (drkN SH3).[57, 76] We show that dynamic ensembles created to satisfy the R_2 relaxation or S^2 order parameter restraints better agree with more commonly collected experimental solution NMR and SAXS data relative to statically generated ensembles. Furthermore, we observe that R_2 rates combined with NOEs alone can yield ensembles with the smallest RMSD with respect to all experimental data types for the unfolded state of the drkN SH3 domain. We conclude that incorporation of dynamic information in defining starting pools of conformers for disordered protein structure calculations is valuable for computational approaches to best represent the disordered protein ensemble.

METHODS

Structural ensembles of the unfolded states of the drkN SH3 domain.

To evaluate the suitability of R_2 data and S^2 values for selecting a disordered ensemble, we consider the unfolded state of the drkN SH3 domain.[57, 76] The drkN SH3 domain has been extensively studied by the Forman-Kay and Gradinaru groups, yielding a wide variety of structural and dynamic experimental data types.[33, 76–77] The domain exists in approximately 1:1 equilibrium between folded and unfolded states under typical buffer conditions[54], and the NMR data were acquired under conditions of the 50:50 mixture with slow exchange on the NMR timescale so that data for the unfolded state could be extracted separate from data for the folded state. NOE data were carefully analyzed to avoid transfer NOEs occurring due to exchange between folded and unfolded states during the mixing time.[57, 76]

We use three initial structural ensemble pools previously investigated with the X-EISD method.[33] (i) The RANDOM pool is a collection of 100,000 structures including 1000 folded structures and 999,000 increasingly unfolded structures. The RANDOM pool is not optimized with respect to the experimental data. (ii) The ENSEMBLE pool is optimized with respect to the available NMR and SAXS experimental data and contains 1700 structures.[24] (iii) The MIXED pool combines the ENSEMBLE pool structures and optimized RANDOM pool structures. There are underlying structural differences in the RANDOM, ENSEMBLE, and MIXED pools, such as the percentage of the secondary structure for each residue (local structure) and global characteristics as measured by the radius of gyration

(R_g).[33]

Anisotropic Network Model.

We use the Anisotropic Network Model (ANM) introduced by Bahar and co-workers[78], implemented in the Prody software program[79], to define the three lowest normal modes of each conformation of the original ENSEMBLE, RANDOM, and MIXED static pools. Since the frequencies of the normal modes are directly proportional to the energy required for the movement, high frequency modes describe local motions and low frequency modes represent collective conformational changes. Many studies have shown that the lowest three frequency normal modes often agree with experimentally observed functional motions for folded proteins.[68, 80] Figure 1 shows a schematic of one of the conformers within the disordered ensembles modeled as an ANM and its three lowest frequency normal modes.

In ANM, the network nodes are the positions of the α -carbons, and uniform elastic springs with force constants k connect the nodes located within a cutoff distance of R_c. The generalized form of the entire network potential of the ANM is given by:[81]

$$v = \frac{k}{2}\sum_{i<j}\left(r_{ij} - r_{ij}^{o}\right)^2 \quad (1)$$

where r_{ij} is the distance between atoms i and j and r_{ij}^{o} is the distance between the atoms in the reference structure. The normal mode analysis using the ANM requires the diagonalization of the following generalized eigenvalue problem:

$$\boldsymbol{HU} = \lambda \boldsymbol{TU} \; where \; \boldsymbol{U} = (u_1, u_2, \ldots, u_N) \quad (2)$$

where $\boldsymbol{H}$ is the Hessian matrix of the partial second derivatives of the potential energy, $\boldsymbol{T}$ the kinetic energy matrix, $\boldsymbol{U}$

are the normal modes, and
λ
is a diagonal matrix with the
λ_k
eigenvalues associated to the kth normal mode
u_k
. A detailed description of the derivation of the Hessian matrix and the normal modes can be found elsewhere.[78, 81] In this study we model the unfolded state of the drkN SH3 domain using an ANM model that has a force constant of
$k = 1$
kcal mol^{-1} and a cutoff of
$R_c = 15$ Å
.

Normal Mode Displacements.

We also further extend the ANM ensembles to generate alternate conformations along the normal mode displacement vectors. Procedurally we used the Prody software to extend normal modes calculated for the
Cα
model into an all-atom model. The extend-Model function in the Prody software[79] takes part of the normal modes for
Cα
atoms and extends it to create all other atoms in the same residue. Then we use these all-atom models to generate an ensemble of randomly sampled conformations along the three lowest frequency normal modes. Then we use the all-atom model and normal mode displacements to generate alternate conformations along the three lowest frequency normal modes. For a given normal mode,
u_i
, and their eigenvalues,
λ_i
, new conformations are sampled using the relationship in Eq. 3, where
R_0
is the active coordinate set of atoms and
$[r_1^k r_2^k \ldots r_m^k]$
are normally distributed random numbers generated for conformation
j
.

$$R_k = R_0 + s \sum_{i=1}^{m} r_i^j \lambda_i^{-0.5} u_i \quad (3)$$

We use this procedure to create the
R_2

-dynamic and S^2-dynamic pools as our putative starting ensembles which we compare to the original "static" ensemble pools.

Sub-ensemble selection using R_2 rates and S^2 values.

Within the ENSEMBLE program[47], the R_2 restraint is defined as the correlation of R_2 values with the number of contacts of the amide group.[24, 47] We use this structural definition as a dynamic proxy of the transverse R_2 relaxation rate to select sub-ensembles of structures from the original ENSEMBLE, RANDOM, and MIXED pools. We also chose to generate sub-ensembles selected using S^2 order parameters since they are conceptually easy to compare to dynamic modes that are computationally sampled. To select a sub-ensemble of structures that agrees with the S^2 values, we first compute the normalized mean square fluctuations (MSFs) of the lowest three frequency modes for each protein structure using an anisotropic network model (see below). Then, we select individual protein structures whose MSFs agree well with the S^2 order parameters.[82] The underlying hypothesis is that the higher the S^2 values for a given segment, the lower its calculated MSFs should be. To measure the agreement between MSFs and S^2, we calculate Pearson correlation coefficients for six different residue ranges as provided in Supplementary Figure S1. The R_2 rate and S^2 order parameter sub-ensemble structure pools selected via these procedures are seen to have very similar profiles (Figure 2a and 2b). Therefore, we consider both R_2 and S^2 measurements in the creation of new dynamical ensembles for disordered states.

Back-calculations and scoring of disordered ensembles using the X-EISD algorithm.

The validity of the disordered conformational ensembles refined and expanded using R_2 and S^2 criteria and the degree to which they can represent the structural information available from experimental data is measured using X-EISD.[33] The X-EISD framework uses a maximum likelihood estimator formalism to assign a log likelihood score of a simulated ensemble matching an input set of experimental data. The procedure accounts for the uncertainties in both back-calculation and experiment by optimizing over the set of "nuisance parameters" that are treated as Gaussian random variables. X-EISD can be applied to multiple data types simultaneously to generate an aggregated probabilistic score of the form as in Eq. (4).

$$\log \mathrm{p}(X, \xi \mid D, I) = \log \mathrm{p}(X \mid I) + \sum_{j=1}^{M} \log[p(d_j \mid X, \xi_j, I) p(\xi_j \mid I)] + C \tag{4}$$

Moreover, the X-EISD framework can provide a probabilistic score of an ensemble in a Markov Chain Monte Carlo (MCMC) optimization as given in eq. (5).[33] In this work, we use a simple direct maximization and perform 1000 exchange attempts to replace one conformation with another from the total pool of $N = 100$ starting structures where an exchange is accepted only when the new ensemble has a higher probabilistic X-EISD score than the previous. We perform optimization by using either a single experimental type or multiple experimental types at a given time.

$$acc(i \rightarrow j) = X - EISD_j > X - EISD_i \tag{5}$$

The back-calculations for the different experimental types are done as follows: the Karplus equation is used to back-calculate the J scalar couplings (JCs) as reported previously[33]; chemical shift (CS) calculations are performed using SHIFTX2 calculator;[83] residual dipolar couplings (RDCs) are computed using the local RDC back-calculator from the Forman-Kay group;[35] the hydrodynamic radius (R_h) is calculated using the program HYDROPRO;[84] and NOE, PRE and smFRET calculations are done using in-house codes as described previously.[33] Finally, we calculate the properties of the optimized ensemble such as the R_g distribution using the Prody implementation,[79] and secondary structure content using the implementation of the DSSP algorithm[85] within the AmberTools program cpptraj.[86]

RESULTS

While methods to develop ensembles composed of independent "static" conformational snapshots have proven very useful to the disordered protein structural community,

approaches for building meaningful ensembles consistent with all experimental data, including those that are time-dependent (e.g., NOEs, relaxation rates), require further attention. Hence the purpose of this study is to evaluate the sub-ensembles created via R_2 and S^2 selections and the degree to which they can represent the structural information available from all available experimental data for the drk SH3 domain.

Our first approach uses the correlations between R_2 rates and NH contacts or S^2 values and the low frequency modes of the ANM to create sub-ensembles of each of the original static ENSEMBLE, RANDOM, and MIXED pools; the resulting sub-ensembles are then optimized across all the experimental data types using the X-EISD optimization to create the R_2-select and S^2-select pools. Supplementary Table S1 reports the X-EISD scores as well as the root mean square deviations (RMSD) with respect to all 8 experimental data types. The R_2 and S^2 filtering step and subsequent X-EISD optimization improves chemical shifts, SAXS, and R_h values with little degradation in the JC, RDC, and smFRET experimental data types for the ENSEMBLE and MIXED pools. However, R_2 filtering significantly decreases agreement with the JC, NOE and PRE values for the RANDOM pool, which we attribute to diminishment of compact structures relative to the original unfiltered structural ensemble.

Importantly, improved ensembles are realized when we expand the R_2 and S^2 filtering selections by introducing dynamical motions of their conformations using the normal mode displacements from the ANM model, and then optimizing these dynamically expanded ensembles with all the experimental data using the X-EISD procedure. As seen in Figure 3, the resulting R_2-dynamic and the S^2

-dynamic ensembles have even better RMSD values against all independent experimental data types compared to the R_2-select and the S^2-select ensembles. This we take as our final dynamical model approach to improve experimental agreement for the unfolded state of the drkN SH3 domain.

Table 1 quantifies the overall improvement in X-EISD scores and RMSD when comparing the optimized R_2-dynamic and S^2-dynamic pools against the original and static ENSEMBLE, MIXED, and RANDOM structural ensembles. For the ENSEMBLE pool, the R_2-dynamic ensemble performs better for nearly all the experimental properties compared to the original static ENSEMBLE pool. Improvements in RMSD can be seen in local data types such as chemical shifts, J-couplings, and RDCs, long-range distance restraints such as NOEs and PREs, as well as measurements of global shape information such as R_h, with essentially equivalent performance for smFRET and SAXS. Similar improvement with respect to experimental data types can also be seen for the MIXED R_2-dynamic ensemble, although our approach is not as consistent if the underlying pool is highly non-optimal such as the RANDOM R_2-dynamic pool. It has been previously shown that the RANDOM ensemble does not have sufficient conformers to represent the drkN SH3 domain unfolded state and is outside the uncertainties for local experimental data types such as J-couplings and chemical shifts.[33] Hence the dynamical expansion can't overcome the original deficiencies of the static RANDOM ensemble.

Similar to the R_2-dynamic ensembles, the S^2-dynamic pools also have improved X-EISD scores and RMSD values with respect to the 8 data types, but quantitatively the R_2 relaxation rates seem to yield more consistent results than the S^2 order parameter. A T-test with a 95% confidence interval and a comparison of distribution of values indicate that the two dynamical selection methods are in fact statistically different from each other (Supplementary Table S1). The origin of this difference in the two

dynamical measures could arise from a number of sources: (i) the R_2 rates are distinct from S^2 values, which include R_1 rates and NOE contributions, as well as assumptions about separation of motional timescales; (ii) the approach for selection is different, comparing the correlation between R_2 rates and numbers of HN contacts vs. the correlation between S^2 values and normalized mean square fluctuations of the lowest three frequency modes using ANM; and (iii) the selection method is on a per-residue basis for R_2 rates vs. on a segmental basis for S^2 values (see Supplementary Figure S2). Summarizing, we find that using dynamic information from R_2 relaxation data, and to a lesser extent S^2 values, as a prior to select a sub-ensemble improves the agreement with experiment for the unfolded state of the drkN SH3 domain.

Lincoff et al. have shown previously that certain NMR data types, such as J-couplings or NOEs, are very valuable in refining a structural ensemble since optimization of these single data types can help improve the other experimental data types such as SAXS or PREs.[33] Figure 4 provides the single optimizations with one data type (the diagonal entries) and its influence on the RMSDs of unoptimized data types (off-diagonal entries) compared with the unoptimized scores for all experimental data types in the last row for the R_2-dynamic structural ensembles. While the single optimizations of a given experimental data type using the R_2-dynamic ensemble improve RMSD errors of all the other experimental data types with few exceptions, the most dramatic improvement is the direct optimization of NOEs that improves the RMSD significantly for JC, PRE, RDC and smFRET for the MIXED R_2-dynamic ensemble (Figure 4c). Corresponding plots are shown using the S^2-derived ensembles in Supplementary Figure S3.

It is insightful to analyze the resulting refined R_2-dynamic ensemble from the perspective of structural signatures such as the radius of

gyration (R_g) (Figure 5). We find that the change in R_g between the original and the R_2 -dynamic pools is minimal for the ENSEMBLE structural ensemble, but that is not surprising since it had already been optimized by the same experimental data types as previously reported.[33] By contrast the MIXED R_2 -dynamic ensemble has shifted to having more expanded conformers as evident in the R_g distribution. For the RANDOM and MIXED pools, the refinement based on R_2 relaxation rates decreases the collapsed peak at ~12 Å in the R_g distribution, corresponding to the folded state conformers of the drkN SH3 domain and increases the importance of more extended conformations. Similar conclusions are reached for the S_2 -dynamic ensembles as reported in the Supplementary Figure S2. This suggests that the selection and expansion based on dynamic properties enhances relevant structural properties within these disordered ensembles.

DISCUSSION AND CONCLUSIONS

In this study, we presented a computational method that uses NMR R_2 relaxation data, or S^2 order parameters in conjunction with an anisotropic network model, to define an underlying ensemble of conformations for a disordered protein that is more suitable for optimization against a range of experimental data that measure local, long range as well as global shape restraint information. Furthermore, we show that by using an anisotropic network model to generate alternate conformations along the normal modes, one can further improve the agreement of the selected ensemble with experiment.

It is understood that the S^2 value representing the internal dynamics of a protein derived from the Lipari-Szabo model free approach is not well formulated for the motions of an IDP or an unfolded protein.[53, 82] This is because the

S^2 value is based on the assumption that the correlation function that describes the dynamics of the amide bond vector can be written as a product of two exponentially decaying correlation functions: the overall tumbling motion of the protein molecule and the internal motion of the bond vector. The use of a single correlation time to characterize the overall tumbling is inappropriate for disordered proteins, in general, including the unfolded state of the drkN SH3 domain.[82] Hence the use of S^2 order parameters for disordered proteins is not rigorous due to this lack of separation of motional timescales, and we would also conclude that it was not as fully effective as the R_2 relaxation rate, even with a crude structural estimate based on structural contacts.

Our previous work in developing the X-EISD method considered the full uncertainty in the experimental and back-calculation errors to enable meaningful comparisons between ensembles and their optimization. From that work, we show that J-couplings and NOEs can be highly useful data types that can simultaneously improve other data types such as smFRET, SAXS, and chemical shifts. However, it also matters that the underlying pool have meaningful conformers, and here we emphasize that the transverse R_2 relaxation rate is another useful measurement that can aid in selection of conformers to facilitate ensemble optimization. This contributes to the important goal of defining disordered state ensembles in order to better understand the relationships between the experimental data types that are most useful for characterizing IDP and unfolded protein conformational landscapes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS.

The Berkeley and Toronto investigators thank the National Institute of Health for support under Grant 5R01GM127627-04. J.D.F. also acknowledges support from the Natural Sciences and Engineering Research Council of Canada. This research used the computational resources from the SAVIO cluster. We thank James Lincoff for assistance with the X-EISD code. We congratulate José Onuchic, an excellent scientist and community builder, in honor of his 65th birthday.

REFERENCES

1. Wright PE; Dyson HJ Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. Journal of molecular biology 1999, 293 (2), 321–331. [PubMed: 10550212]
2. Uversky VN Intrinsically disordered proteins from A to Z. The international journal of biochemistry & cell biology 2011, 43 (8), 1090–1103. [PubMed: 21501695]
3. Tompa P.Intrinsically disordered proteins: a 10-year recap. Trends in biochemical sciences 2012, 37 (12), 509–516. [PubMed: 22989858]
4. Uversky VN Unusual biophysics of intrinsically disordered proteins. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 2013, 1834 (5), 932–951. [PubMed: 23269364]

5. Ball KA; Wemmer DE; Head-Gordon T.Comparison of structure determination methods for intrinsically disordered amyloid-β peptides. The Journal of Physical Chemistry B 2014, 118 (24), 6405–6416. [PubMed: 24410358]
6. Van Der Lee R; Buljan M; Lang B; Weatheritt RJ; Daughdrill GW; Dunker AK; Fuxreiter M; Gough J; Gsponer J; Jones DT Classification of intrinsically disordered regions and proteins. Chemical reviews 2014, 114 (13), 6589–6631. [PubMed: 24773235]
7. Oldfield CJ; Dunker AK Intrinsically disordered proteins and intrinsically disordered protein regions. Annual review of biochemistry 2014, 83, 553–584.
8. Wright PE; Dyson HJ Intrinsically disordered proteins in cellular signalling and regulation. Nature reviews Molecular cell biology 2015, 16 (1), 18–29. [PubMed: 25531225]
9. Bhowmick A; Brookes DH; Yost SR; Dyson HJ; Forman-Kay JD; Gunter D; Head-Gordon M; Hura GL; Pande VS; Wemmer DE Finding our way in the dark proteome. Journal of the American Chemical Society 2016, 138 (31), 9730–9742. [PubMed: 27387657]
10. Liu M; Das AK; Lincoff J; Sasmal S; Cheng SY; Vernon RM; Forman-Kay JD; Head-Gordon T.Configurational Entropy of Folded Proteins and Its Importance for Intrinsically Disordered Proteins. International journal of molecular sciences 2021, 22 (7), 3420. [PubMed: 33810353]
11. Uversky VN; Oldfield CJ; Dunker AK Intrinsically disordered proteins in human diseases: introducing the D2 concept. Annu. Rev. Biophys 2008, 37, 215–246. [PubMed: 18573080]
12. Yoon M-K; Mitrea DM; Ou L; Kriwacki RW Cell cycle regulation by the intrinsically disordered proteins p21 and p27. Biochemical Society Transactions 2012, 40 (5), 981–988. [PubMed: 22988851]
13. Elbaum-Garfinkle S; Kim Y; Szczepaniak K; Chen CC-H; Eckmann CR; Myong S; Brangwynne CP The disordered P granule protein LAF-1 drives phase separation into droplets with tunable viscosity and dynamics. Proceedings of the National Academy of Sciences 2015, 112 (23), 7189–7194.
14. Das RK; Ruff KM; Pappu RV Relating sequence encoded information to form and function of intrinsically disordered proteins. Current opinion in structural biology 2015, 32, 102–112. [PubMed: 25863585]
15. Lin Y-H; Forman-Kay JD; Chan HS Sequence-specific polyampholyte phase separation in membraneless organelles. Physical review letters 2016, 117 (17), 178101. [PubMed: 27824447]
16. Brady JP; Farber PJ; Sekhar A; Lin Y-H; Huang R; Bah A; Nott TJ; Chan HS; Baldwin AJ; Forman-Kay JD Structural and hydrodynamic properties of an intrinsically disordered region of a germ cell-specific protein on phase separation. Proceedings of the National Academy of Sciences 2017, 114 (39), E8194–E8203.
17. Shin Y; Brangwynne CP Liquid phase condensation in cell physiology and disease. Science 2017, 357 (6357).
18. Wei M-T; Elbaum-Garfinkle S; Holehouse AS; Chen CC-H; Feric M; Arnold CB; Priestley RD; Pappu RV; Brangwynne CP Phase behaviour of disordered proteins underlying low density and high permeability of liquid organelles. Nature chemistry 2017, 9 (11), 1118–1125.
19. Lin Y-H; Forman-Kay JD; Chan HS Theories for sequence-dependent phase behaviors of biomolecular condensates. Biochemistry 2018, 57 (17), 2499–2508. [PubMed: 29509422]
20. Posey AE; Holehouse AS; Pappu RV Phase separation of intrinsically disordered proteins. Methods in enzymology 2018, 611, 1–30. [PubMed: 30471685]
21. Kim TH; Tsang B; Vernon RM; Sonenberg N; Kay LE; Forman-Kay JD Phospho-dependent phase separation of FMRP and CAPRIN1 recapitulates regulation of translation and deadenylation. Science 2019, 365 (6455), 825–829. [PubMed: 31439799]
22. Shrestha UR; Smith JC; Petridis L.Full structural ensembles of intrinsically disordered proteins from unbiased molecular dynamics simulations. Communications biology 2021, 4 (1), 1–8. [PubMed: 33398033]
23. Robic S; Guzman-Casado M; Sanchez-Ruiz JM; Marqusee S.Role of residual structure in the unfolded state of a thermophilic protein. Proceedings of the National Academy of Sciences 2003, 100 (20), 11345.
24. Marsh JA; Forman-Kay JD Ensemble modeling of protein disordered states: experimental restraint contributions and validation. Proteins 2012, 80 (2), 556–72. [PubMed: 22095648]

25. Katava M; Stirnemann G; Zanatta M; Capaccioli S; Pachetti M; Ngai KL; Sterpone F; Paciaroni A.Critical structural fluctuations of proteins upon thermal unfolding challenge the Lindemann criterion. Proceedings of the National Academy of Sciences 2017, 114 (35), 9361.
26. Chiti F; Dobson CM Protein Misfolding, Functional Amyloid, and Human Disease. Annual Review of Biochemistry 2006, 75 (1), 333–366.
27. Dyson HJ; Wright PE NMR illuminates intrinsic disorder. Current Opinion in Structural Biology 2021, 70, 44–52. [PubMed: 33951592]
28. Fisher CK; Stultz CM Constructing ensembles for intrinsically disordered proteins. Current opinion in structural biology 2011, 21 (3), 426–431. [PubMed: 21530234]
29. Gomes G-NW; Krzeminski M; Namini A; Martin EW; Mittag T; Head-Gordon T; Forman-Kay JD; Gradinaru CC Conformational ensembles of an intrinsically disordered protein consistent with NMR, SAXS, and single-molecule FRET. Journal of the American Chemical Society 2020, 142 (37), 15697–15710. [PubMed: 32840111]
30. Huang A; Stultz CM The effect of a ΔK280 mutation on the unfolded state of a microtubule-binding repeat in Tau. PLoS computational biology 2008, 4 (8), e1000155. [PubMed: 18725924]
31. Jensen MR; Markwick PR; Meier S; Griesinger C; Zweckstetter M; Grzesiek S; Bernadó P; Blackledge M.Quantitative determination of the conformational properties of partially folded and intrinsically disordered proteins using NMR dipolar couplings. Structure 2009, 17 (9), 1169–1185. [PubMed: 19748338]
32. Fisher CK; Huang A; Stultz CM Modeling intrinsically disordered proteins with bayesian statistics. Journal of the American Chemical Society 2010, 132 (42), 14919–14927. [PubMed: 20925316]
33. Lincoff J; Haghighatlari M; Krzeminski M; Teixeira JM; Gomes G-NW; Gradinaru CC; Forman-Kay JD; Head-Gordon T.Extended experimental inferential structure determination method in determining the structural ensembles of disordered protein states. Communications chemistry 2020, 3 (1), 1–12. [PubMed: 36703346]
34. Mukrasch MD; Markwick P; Biernat J; von Bergen M; Bernadó P; Griesinger C; Mandelkow E; Zweckstetter M; Blackledge M.Highly populated turn conformations in natively unfolded tau protein identified from residual dipolar couplings and molecular simulation. Journal of the American Chemical Society 2007, 129 (16), 5235–5243. [PubMed: 17385861]
35. Marsh JA; Baker JM; Tollinger M; Forman-Kay JD Calculation of residual dipolar couplings from disordered state ensembles using local alignment. Journal of the American Chemical Society 2008, 130 (25), 7804–7805. [PubMed: 18512919]
36. Nodet G; Salmon L; Ozenne V; Meier S; Jensen MR; Blackledge M.Quantitative description of backbone conformational sampling of unfolded proteins at amino acid resolution from NMR residual dipolar couplings. Journal of the American Chemical Society 2009, 131 (49), 17908–17918. [PubMed: 19908838]
37. Camilloni C; Vendruscolo M.Using Pseudocontact Shifts and Residual Dipolar Couplings as Exact NMR Restraints for the Determination of Protein Structural Ensembles. Biochemistry 2015, 54 (51), 7470–7476. [PubMed: 26624789]
38. Wilkins DK; Grimshaw SB; Receveur V; Dobson CM; Jones JA; Smith LJ Hydrodynamic Radii of Native and Denatured Proteins Measured by Pulse Field Gradient NMR Techniques. Biochemistry 1999, 38 (50), 16424–16431. [PubMed: 10600103]
39. Petsko GA; Ringe D.Fluctuations in protein structure from X-ray diffraction. Annual review of biophysics and bioengineering 1984, 13 (1), 331–371.
40. Mylonas E; Hascher A; Bernado P; Blackledge M; Mandelkow E; Svergun DI Domain conformation of tau protein studied by solution small-angle X-ray scattering. Biochemistry 2008, 47 (39), 10345–10353. [PubMed: 18771286]
41. Yang S; Blachowicz L; Makowski L; Roux B.Multidomain assembled states of Hck tyrosine kinase in solution. Proceedings of the National Academy of Sciences 2010, 107 (36), 15757–15762.
42. Różycki B; Kim YC; Hummer G.SAXS ensemble refinement of ESCRT-III CHMP3 conformational transitions. Structure 2011, 19 (1), 109–116. [PubMed: 21220121]
43. Mukrasch MD; Bibow S; Korukottu J; Jeganathan S; Biernat J; Griesinger C; Mandelkow E; Zweckstetter M.Structural polymorphism of 441-residue tau at single residue resolution. PLoS biology 2009, 7 (2), e1000034. [PubMed: 19226187]

44. Ganguly D; Chen J.Structural interpretation of paramagnetic relaxation enhancement-derived distances for disordered protein states. Journal of molecular biology 2009, 390 (3), 467–477. [PubMed: 19447112]
45. Nasir I; Onuchic PL; Labra SR; Deniz AA Single-molecule fluorescence studies of intrinsically disordered proteins and liquid phase separation. Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics 2019, 1867 (10), 980–987. [PubMed: 31054969]
46. Brookes DH; Head-Gordon T.Experimental inferential structure determination of ensembles for intrinsically disordered proteins. Journal of the American Chemical Society 2016, 138 (13), 4530–4538. [PubMed: 26967199]
47. Krzeminski M; Marsh JA; Neale C; Choy W-Y; Forman-Kay JD Characterization of disordered proteins with ENSEMBLE. Bioinformatics 2013, 29 (3), 398–399. [PubMed: 23233655]
48. Jensen MR; Salmon L. c.; Nodet G; Blackledge M. Defining conformational ensembles of intrinsically disordered and partially folded proteins directly from chemical shifts. Journal of the American Chemical Society 2010, 132 (4), 1270–1272. [PubMed: 20063887]
49. Feldman HJ; Hogue CW A fast method to sample real protein conformational space. Proteins: Struct., Func., Bioinform. 2000, 39 (2), 112–131.
50. Chen Y; Campbell SL; Dokholyan NV Deciphering protein dynamics from NMR data using explicit structure sampling and selection. Biophysical journal 2007, 93 (7), 2300–2306. [PubMed: 17557784]
51. Ortega G; Pons M; Millet O.Chapter Six - Protein Functional Dynamics in Multiple Timescales as Studied by NMR Spectroscopy. In Advances in Protein Chemistry and Structural Biology, Karabencheva-Christova T, Ed. Academic Press: 2013; Vol. 92, pp 219–251. [PubMed: 23954103]
52. Farrow NA; Zhang O; Szabo A; Torchia DA; Kay LE Spectral density function mapping using 15N relaxation data exclusively. Journal of Biomolecular NMR 1995, 6 (2), 153–162. [PubMed: 8589604]
53. Lipari G; Szabo A.Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. Journal of the American Chemical Society 1982, 104 (17), 4546–4559.
54. Farrow NA; Zhang O; Forman-Kay JD; Kay LE Comparison of the backbone dynamics of a folded and an unfolded SH3 domain existing in equilibrium in aqueous buffer. Biochemistry 1995, 34 (3), 868–878. [PubMed: 7827045]
55. Brüschweiler R.New approaches to the dynamic interpretation and prediction of NMR relaxation data from proteins. Current Opinion in Structural Biology 2003, 13 (2), 175–183. [PubMed: 12727510]
56. Temiz NA; Meirovitch E; Bahar I.Escherichia coli adenylate kinase dynamics: Comparison of elastic network model modes with mode-coupling 15N-NMR relaxation data. PROTEINS: Structure, Function, and Bioinformatics 2004, 57 (3), 468–480.
57. Marsh JA; Neale C; Jack FE; Choy W-Y; Lee AY; Crowhurst KA; Forman-Kay JD Improved structural characterizations of the drkN SH3 domain unfolded state suggest a compact ensemble with native-like and non-native structure. Journal of molecular biology 2007, 367 (5), 1494–1510. [PubMed: 17320108]
58. Robustelli P; Kohlhoff K; Cavalli A; Vendruscolo M.Using NMR Chemical Shifts as Structural Restraints in Molecular Dynamics Simulations of Proteins. Structure 2010, 18 (8), 923–933. [PubMed: 20696393]
59. Baxa MC; Haddadian EJ; Jumper JM; Freed KF; Sosnick TR Loss of conformational entropy in protein folding calculated using realistic ensembles and its implications for NMR-based calculations. Proceedings of the National Academy of Sciences 2014, 201407768.
60. Lincoff J; Sasmal S; Head-Gordon T.The combined force field-sampling problem in simulations of disordered amyloid-β peptides. The Journal of chemical physics 2019, 150 (10), 104108. [PubMed: 30876367]
61. Haliloglu T; Bahar I.Structure-based analysis of protein dynamics: Comparison of theoretical results for hen lysozyme with X-ray diffraction and NMR relaxation data. Proteins: Structure, Function, and Bioinformatics 1999, 37 (4), 654–667.

62. Chennubhotla C; Rader A; Yang L-W; Bahar I.Elastic network models for understanding biomolecular machinery: from enzymes to supramolecular assemblies. Physical biology 2005, 2 (4), S173. [PubMed: 16280623]
63. Cui Q; Bahar I.Normal mode analysis: theory and applications to biological and chemical systems. CRC press: 2005.
64. Yang L; Song G; Jernigan RL How well can we understand large-scale protein motions using normal modes of elastic network models? Biophysical journal 2007, 93 (3), 920–929. [PubMed: 17483178]
65. Eyal E; Chennubhotla C; Yang L-W; Bahar I.Anisotropic fluctuations of amino acids in protein structures: insights from X-ray crystallography and elastic network models. Bioinformatics 2007, 23 (13), i175–i184. [PubMed: 17646294]
66. Yang L-W; Eyal E; Chennubhotla C; Jee J; Gronenborn AM; Bahar I.Insights into equilibrium dynamics of proteins from comparison of NMR and X-ray data with computational predictions. Structure 2007, 15 (6), 741–749. [PubMed: 17562320]
67. Yang L-W; Chng C-P Coarse-grained models reveal functional dynamics-I. elastic network models–theories, comparisons and perspectives. Bioinformatics and Biology Insights 2008, 2, BBI. S460.
68. Peng C; Zhang L; Head-Gordon T.Instantaneous normal modes as an unforced reaction coordinate for protein conformational transitions. Biophys J 2010, 98 (10), 2356–64. [PubMed: 20483345]
69. Isin B; Tirupula KC; Oltvai ZN; Klein-Seetharaman J; Bahar I.Identification of motions in membrane proteins by elastic network models and their experimental validation. In Membrane Protein Structure and Dynamics, Springer: 2012; pp 285–317.
70. Na H; Song G.Bridging between normal mode analysis and elastic network models. Proteins: Structure, Function, and Bioinformatics 2014, 82 (9), 2157–2168.
71. Fuglebakk E; Tiwari SP; Reuter N.Comparing the intrinsic dynamics of multiple protein structures using elastic network models. Biochimica et Biophysica Acta (BBA)-General Subjects 2015, 1850 (5), 911–922. [PubMed: 25267310]
72. Zhang Y; Doruker P; Kaynak B; Zhang S; Krieger J; Li H; Bahar I.Intrinsic dynamics is evolutionarily optimized to enable allosteric behavior. Current opinion in structural biology 2020, 62, 14–21. [PubMed: 31785465]
73. Bernado P; Svergun DI Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. Molecular biosystems 2012, 8 (1), 151–167. [PubMed: 21947276]
74. Setny P; Zacharias M.Elastic network models of nucleic acids flexibility. Journal of chemical theory and computation 2013, 9 (12), 5460–5470. [PubMed: 26592282]
75. Krieger JM; Doruker P; Scott AL; Perahia D; Bahar I.Towards gaining sight of multiscale events: utilizing network models and normal modes in hybrid methods. Current opinion in structural biology 2020, 64, 34–41. [PubMed: 32622329]
76. Mazouchi A; Zhang Z; Bahram A; Gomes G-N; Lin H; Song J; Chan HS; Forman-Kay JD; Gradinaru CC Conformations of a metastable SH3 domain characterized by smFRET and an excluded-volume polymer model. Biophysical journal 2016, 110 (7), 1510–1522. [PubMed: 27074677]
77. Choy W-Y; Mulder FA; Crowhurst KA; Muhandiram D; Millett IS; Doniach S; Forman-Kay JD; Kay LE Distribution of molecular size within an unfolded state ensemble using small-angle X-ray scattering and pulse field gradient NMR techniques. Journal of molecular biology 2002, 316 (1), 101–112. [PubMed: 11829506]
78. Atilgan AR; Durell S; Jernigan RL; Demirel MC; Keskin O; Bahar I.Anisotropy of fluctuation dynamics of proteins with an elastic network model. Biophysical journal 2001, 80 (1), 505–515. [PubMed: 11159421]
79. Bakan A; Meireles LM; Bahar I.ProDy: protein dynamics inferred from theory and experiments. Bioinformatics 2011, 27 (11), 1575–1577. [PubMed: 21471012]
80. Peng C; Head-Gordon T.The dynamical mechanism of auto-inhibition of AMP-activated protein kinase. PLoS Comput Biol 2011, 7 (7), e1002082. [PubMed: 21814500]
81. Eyal E; Yang L-W; Bahar I.Anisotropic network model: systematic evaluation and a new web interface. Bioinformatics 2006, 22 (21), 2619–2627. [PubMed: 16928735]

82. Farrow NA; Zhang O; Forman-Kay JD; Kay LE Characterization of the backbone dynamics of folded and denatured states of an SH3 domain. Biochemistry 1997, 36 (9), 2390–2402. [PubMed: 9054544]

83. Han B; Liu Y; Ginzinger SW; Wishart DS SHIFTX2: significantly improved protein chemical shift prediction. Journal of biomolecular NMR 2011, 50 (1), 43. [PubMed: 21448735]

84. Ortega A; Amorós D; De La Torre JG Prediction of hydrodynamic and other solution properties of rigid proteins from atomic-and residue-level models. Biophysical journal 2011, 101 (4), 892–898. [PubMed: 21843480]

85. Kabsch W; Sander C.Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers: Original Research on Biomolecules 1983, 22 (12), 2577–2637.

86. Roe DR; Cheatham III TE PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. Journal of chemical theory and computation 2013, 9 (7), 3084–3095. [PubMed: 26583988]

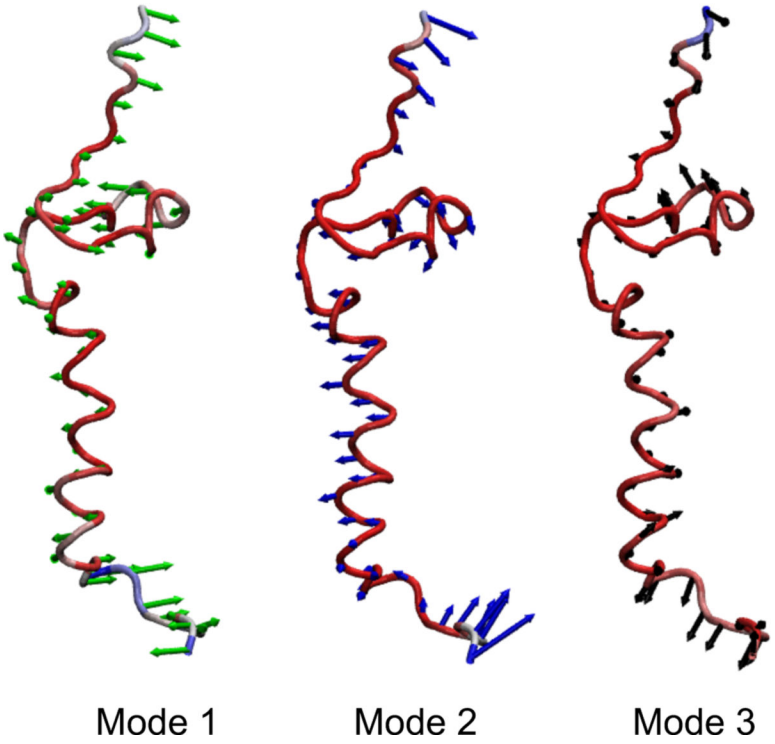

Figure 1. The dynamical description of the unfolded state of the drkN SH3 domain using an anisotropic network model.

Three lowest frequency normal modes for a single structure out of the disordered ensemble where green arrows, blue arrows, and black arrows represent the direction and amplitudes of the three lowest frequency motional modes, super-imposed on a schematic representation of the protein backbone in the unfolded state.

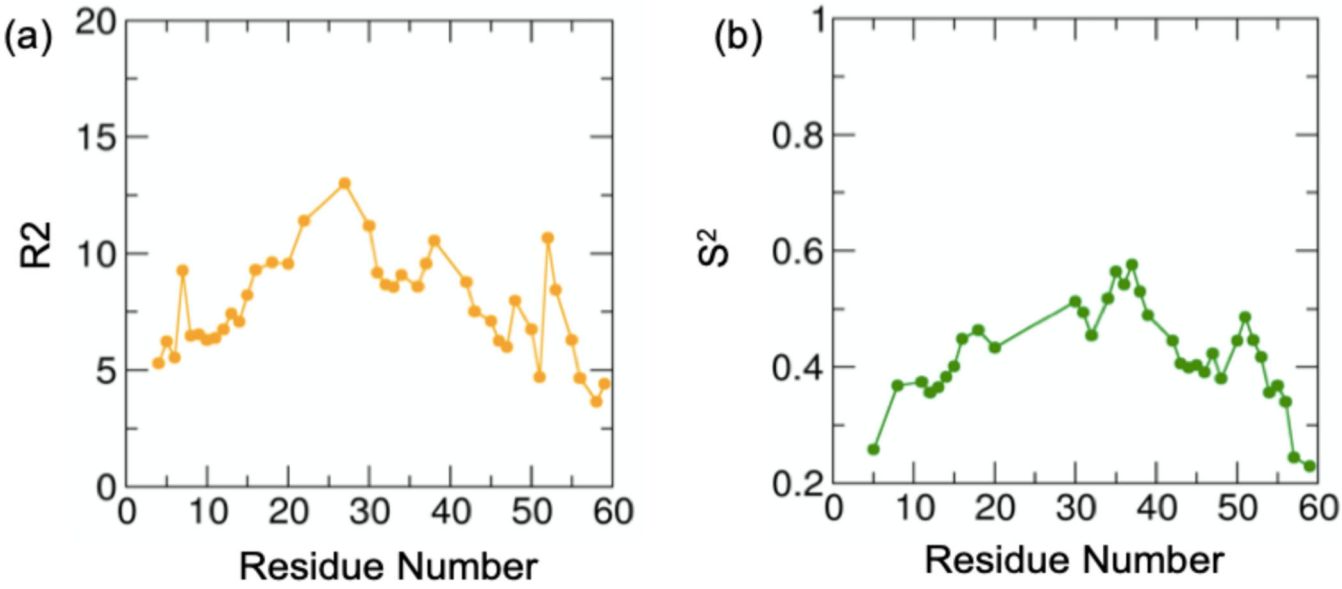

Figure 2.

R_2

and

S^2

profiles for the unfolded state of the drkN SH3 domain.

(a) ^{15}N

R_2

relaxation rates and (b)

S^2

order parameters as a function of residue number.[54, 82] Units in (a) are in 1/sec.

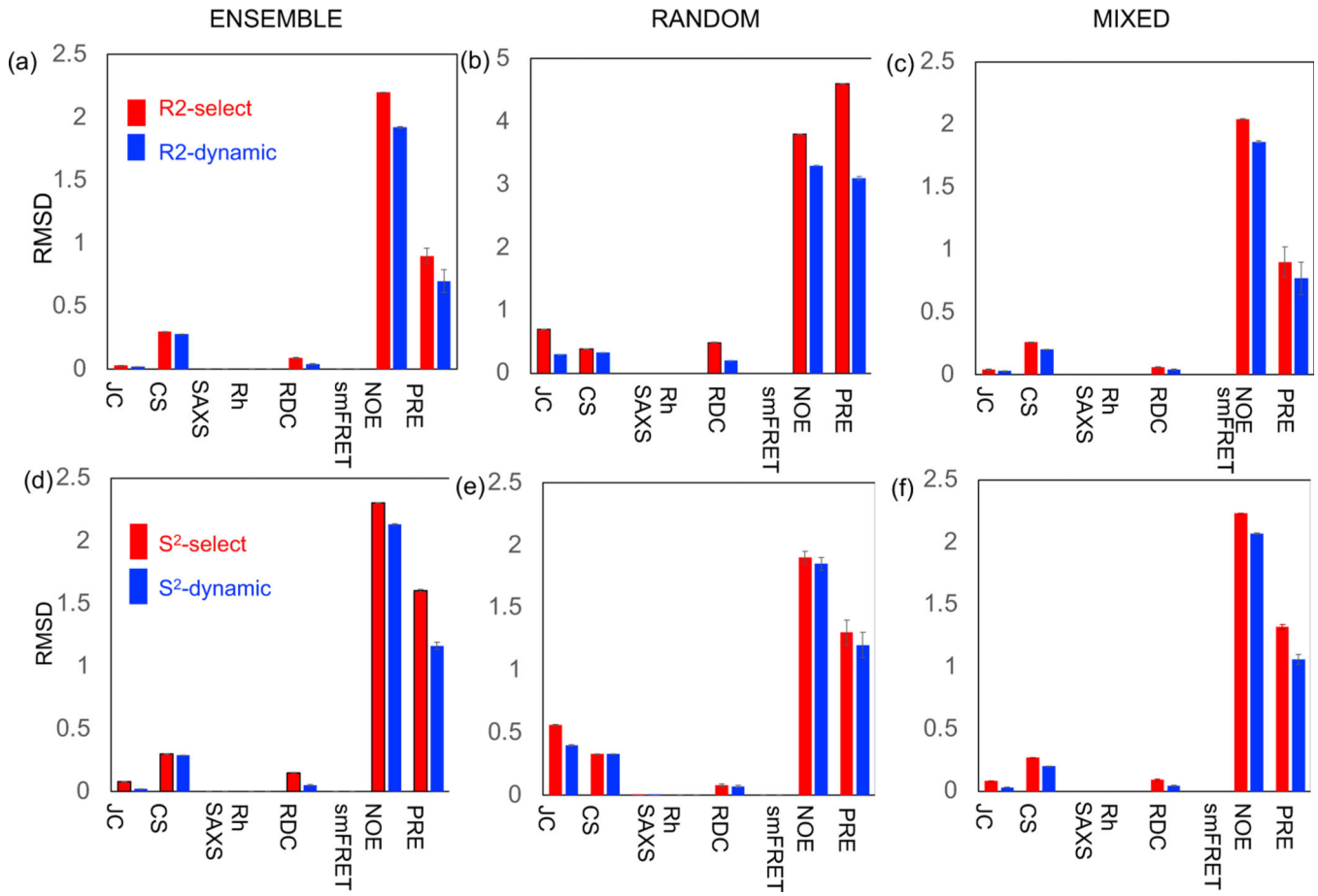

Figure 3. The change in RMSD error of the eight experimental data types (x-axis) for the R_2 and S^2 -selected ensembles (red-bar) vs the R_2 - and S^2 -dynamic ensembles (blue-bar).

The panels (a-c) show the results for R_2 and panels (d-f) show the results for S^2 for the ENSEMBLE, RANDOM, and MIXED pool optimizations using X-EISD. There are negligible errors according to the Bayesian model for SAXS, R_h and smFRET. RMSD units are different for each experimental data type and are found in Table 1.

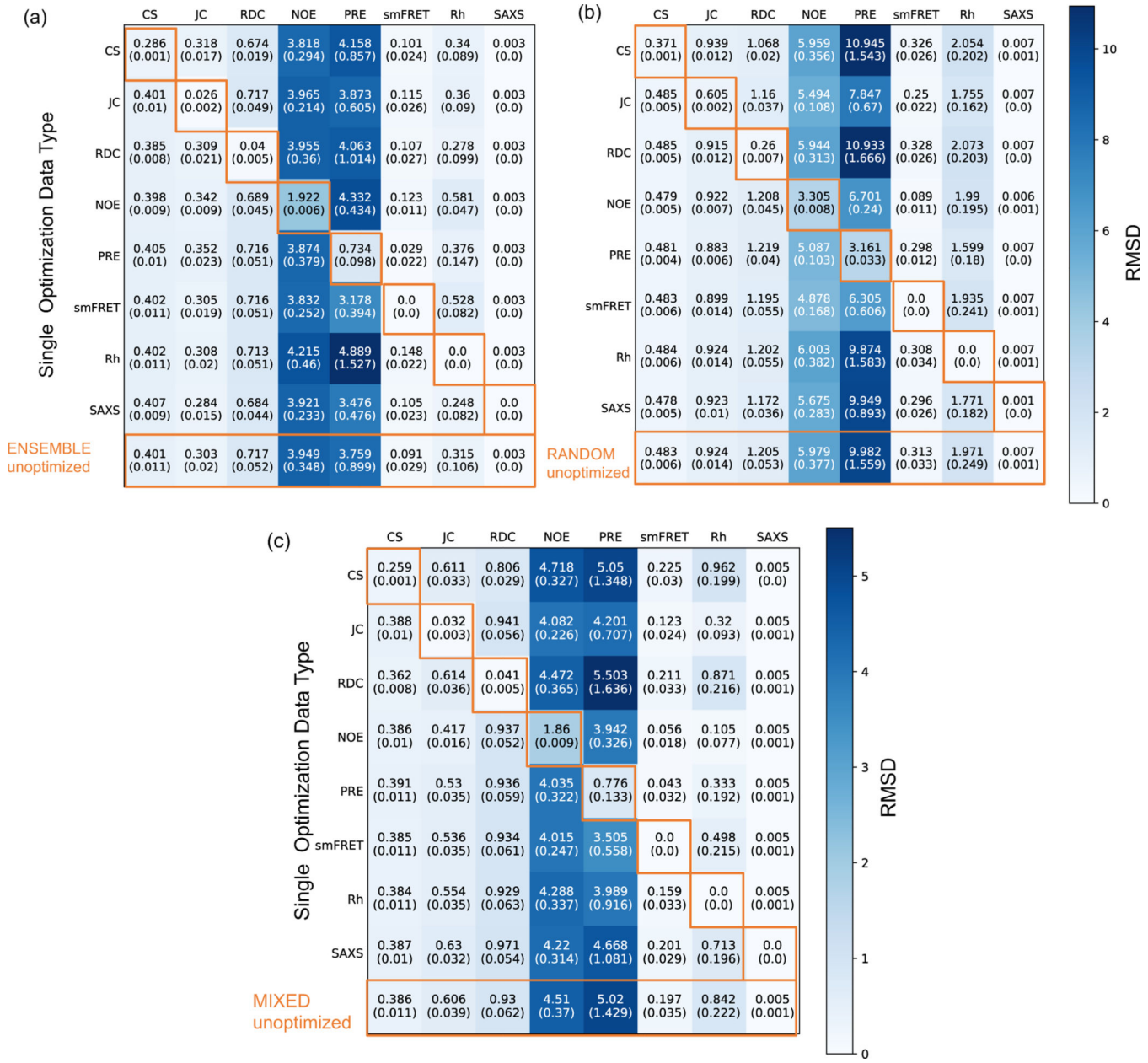

Figure 4. ***RMSD errors by optimizing the X-EISD score with a single experimental data type for*** R_2 ***-dynamic ensembles*** **derived from (a) ENSEMBLE, (b) RANDOM, and (c) MIXED pools.** The values are averages over 1000 ensembles of 100 structures each, and the numbers in parenthesis are standard deviations. The last row refers to the unoptimized R_2 -dynamic pool of structures. Units are different for each experimental data type and are found in Table 1.

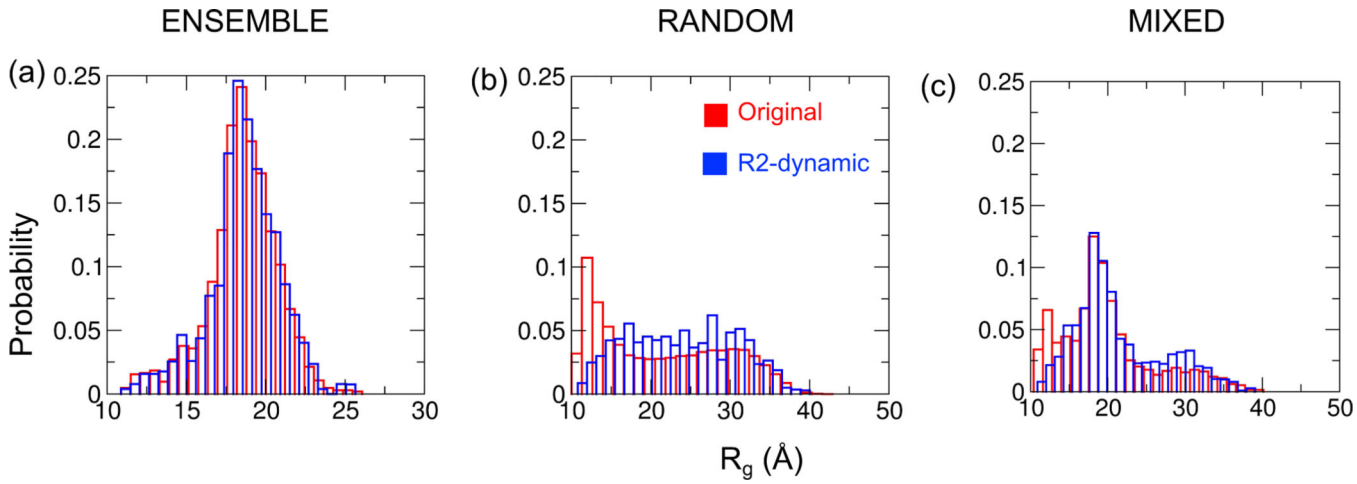

Figure 5. Radius of gyration distributions of the original (red) and R_2 -dynamic (blue) ensembles for the unfolded state of drkN SH3 domain.

Shown for the (a) ENSEMBLE, (b) RANDOM, and (c) MIXED pools.

Table 1:

Evaluation of the optimized R_2-dynamic and S^2-dynamic ensembles and original ensembles for ENSEMBLE, RANDOM, and MIXED pools with 8 experimental data types.

We report both the X-EISD score, and the root mean square deviations (RMSD) with experiments. The results shown here are optimized with all experimental data types. The experimental and back calculations errors for CSs ($\sigma_{exp} = 0.03 - 0.3$ ppm; $\sigma_q = 0.3 - 0.5$ ppm (hydrogen), 1.2–1.4 ppm (carbon)); JCs ($\sigma_{exp} = 0.5$; $\sigma_{q,A} = 0.14$, $\sigma_{q,B} = 0.03$, $\sigma_{q,C} = 0.08$); RDCs ($\sigma_{exp} = 1.0$; $\sigma_q = 0.9$); NOEs ($\sigma_{exp} = 4.0 - 5.0$Å ; $\sigma_q = 0.0001$); PREs ($\sigma_{exp} = 1.25 - 3.625$Å ; $\sigma_q = 0.0001$); smFRET <E> ($\sigma_{exp} = 0.02$; $\sigma_q = 0.007$);

R_h ($\sigma_{exp} = 0.3$; $\sigma_q = 0.8$); SAXS ($\sigma_{exp} = 0.0008 - 0.002$, $\sigma_q = 0.0006$).

Experimental data type	R2 -Dynamic Pool		S^2 -Dynamic Pool		Original Pool	
	X-EISD Score	RMSD	X-EISD Score	RMSD	X-EISD Score	RMSD
ENSEMBLE						
CS (ppm)	117.5 (0.4)	0.3 (0.005)	117.9 (0.3)	0.34 (0.004)	110.1 (0.4)	0.51 (0)
JC (Hz)	43.68 (0.41)	0.16 (0.008)	41.2 (0.4)	0.21 (0.007)	43.3 (0.5)	0.18 (0.01)
RDC (Hz)	−47.8 (0.003)	0.04 (0.005)	−50.1 (0.2)	0.54 (0.02)	−50.4 (0.2)	0.56 (0.03)
NOE (Å)	534.8 (0.4)	2.6 (0.03)	533.7 (0.4)	2.7 (0.03)	532.2 (0.5)	2.8 (0.03)
PRE (Å)	460.4 (2.1)	0.86 (0.06)	443.6 (1.0)	1.42 (0.03)	453.9 (1.7)	1.02 (0.11)
smFRET <E>	6.91 (0.05)	0.006 (0.003)	6.95 (0.03)	0.004 (0.002)	7.0 (0)	0 (0)
Rh (Å)	−0.42 (0)	0	−0.57 (0.02)	0.47 (0.03)	−0.8 (0)	0.71 (0.04)
SAXS (Intensity)	458.3 (0.2)	0.001 (0)	458.0 (0.2)	0.001 (0)	457.9 (0.2)	0.001 (0)
RANDOM						
CS (ppm)	118.3 (0.03)	0.33 (0.002)	109.59 (0.66)	0.47 (0.006)	103.6 (0.7)	0.55 (0.01)
JC (Hz)	−47.6 (1.7)	0.78 (0.006)	−34.39 (3.03)	0.73 (0.01)	−25.7 (2.4)	0.7 (0.01)
RDC (Hz)	−48.4 (0.02)	0.2 (0.007)	52.2 (0.02)	0.2 (0.007)	−55.4 (0.6)	0.98 (0.04)
NOE (Å)	506.0 (1.2)	4.3 (0.05)	526.03 (1.77)	3.20 (0.10)	528.5 (1.5)	3.06 (0.10)
PRE (Å)	261.1 (4.0)	3.4 (0.03)	429.53 (5.61)	1.69 (0.11)	450.0 (4.4)	1.24 (0.12)
smFRET <E>	3.9 (0.7)	0.05 (0.006)	6.84 (0.15)	0.009 (0.006)	6.9 (0.1)	0.01 (0)

Experimental data type	R2 -Dynamic Pool		S^2 -Dynamic Pool		Original Pool	
	X-EISD Score	RMSD	X-EISD Score	RMSD	X-EISD Score	RMSD
Rh (Å)	−0.42 (0.1)	0	−0.47 (0.1)	0	−0.4 (0)	0.14 (0.10)
SAXS (Intensity)	452.4 (0.9)	0.003 (0)	455.03 (0.81)	0.002 (0)	456.3 (0.4)	0.002 (0)
MIXED						
CS (ppm)	118.6 (0.45)	0.34 (0.006)	119.6 (0.3)	0.32 (0.005)	115.5 (0.5)	0.49 (0.01)
JC (Hz)	41.2 (0.60)	0.21 (0.01)	39.59 (0.66)	0.23 (0.01)	41.4 (0.7)	0.21 (0.01)
RDC (Hz)	−51.17 (0.35)	0.64 (0.03)	−51.16 (0.3)	0.62 (0.03)	−50.7 (0.3)	0.60 (0.03)
NOE (Å)	535.8 (0.66)	2.6 (0.04)	533.02 (0.53)	2.81 (0.04)	539.4 (0.7)	2.28 (0.07)
PRE (Å)	459.7 (2.3)	0.89 (0.06)	446.4 (1.8)	1.33 (0.03)	458.4 (4.3)	0.92 (0.11)
smFRET <E>	6.9 (0.08)	0.007 (0.004)	6.9 (0.05)	0.005 (0.004)	6.9 (0)	0.01 (0)
Rh (Å)	0.54 (0.05)	0.41 (0.09)	−0.46 (0.02)	0.21 (0.08)	−0.7 (0)	0.69 (0.05)
SAXS (Intensity)	457.5 (0.33)	0.002 (0)	457.9 (0.2)	0.001 (0)	458.0 (0.2)	0.001 (0)