CrossMark

Investigating the Role of Large-Scale Domain Dynamics in Protein-Protein Interactions

Elise Delaforge [1], Sigrid Milles [1], Jie-rong Huang [1], Denis Bouvier [1],
Malene Ringkjøbing Jensen [1], Michael Sattler [2,3], Darren J. Hart [1] and Martin Blackledge [1]*

[1] Institut de Biologie Structurale, CEA, Centre National de la Recherche Scientifique, University Grenoble Alpes, Grenoble, France, [2] Institute of Structural Biology, Helmholtz Zentrum München, Neuherberg, Germany, [3] Center for Integrated Protein Science Munich at Biomolecular NMR, Technische Universität München, Garching, Germany

Intrinsically disordered linkers provide multi-domain proteins with degrees of conformational freedom that are often essential for function. These highly dynamic assemblies represent a significant fraction of all proteomes, and deciphering the physical basis of their interactions represents a considerable challenge. Here we describe the difficulties associated with mapping the large-scale domain dynamics and describe two recent examples where solution state methods, in particular NMR spectroscopy, are used to investigate conformational exchange on very different timescales.

Keywords: multi-domain proteins, free-energy landscape, conformational dynamics, nuclear magnetic resonance (NMR), small angle scattering, chemical exchange saturation transfer (CEST), single molecule Förster resonance energy transfer (FRET)

Over the last four decades, X-ray crystallography, NMR and increasingly electron microscopy have provided unique insight into the nature of functionally-essential interactions between a vast array of biologically active molecules. This remarkable success has often overlooked the importance of the molecular motions that are required for function, with structural biologists tending to focus their attention on highly stable, high affinity biomolecular interactions. Proteins are however intrinsically dynamic, they can exhibit conformational modes of vastly differing amplitudes, from local bond fluctuations to folding/unfolding transitions, on timescales varying from femtoseconds to days. Interactions between proteins are also often required to be weak, for example in signaling pathways where efficiency and reversibility of information transfer are of key importance. The resulting focus on interactions that are strong enough to allow structure determination in terms of a single set of three-dimensional coordinates thus provides a distorted perspective on the interactome.

In order to fully understand the molecular basis of interactions between physiological partners it is necessary to map the modulation of the free energy landscapes of the interacting molecules throughout the interaction trajectory. This aim is the focus of considerable interest, requiring experimental, theoretical and analytical development. For this reason solution state methods have been increasingly used to actively investigate the nature of transient interactions between biomolecules (Vaynberg and Qin, 2006; Sugase et al., 2007; Bashir et al., 2010; Salmon et al., 2011).

Such considerations are particularly relevant for a highly abundant class of proteins whose physical characteristics are defined by their dynamic nature. The unexpected discovery that a high fraction of eukaryotic, but also prokaryotic and viral genomes code for proteins domains whose

functional state is natively unfolded (Dyson and Wright, 2005; Uversky and Dunker, 2010), has imposed a new perspective on our understanding of molecular recognition. In contrast to folded proteins, the primary sequence of intrinsically disordered proteins (IDPs) does not code for a single, energetically stable fold, but occupies a flatter free-energy surface, spanning a continuum of different conformations. The dynamic nature of IDPs allows access to functional modes that are otherwise inaccessible to folded proteins. For example, the dynamic intrinsically disordered domains that fill the nuclear pore (nucleoporins), exhibiting fast association and dissociation rates with nuclear transport receptors, thereby facilitating highly specific but rapid transport of cargoes into the nucleus via a continuum of ultra-weak interactions (Milles et al., 2015). Interactions between IDPs often show kinetics that are more complex than simple two-state mechanisms, remaining dynamic within the complex (Kragelj et al., 2015), exhibiting "fuzzy" interactions that involve for example local conformational funneling into partner-specific bound conformations (Schneider et al., 2015), or involving transient, non-specific interactions to enhance affinity (Fink, 2005; Dunker et al., 2008; Tompa and Fuxreiter, 2008; Wright and Dyson, 2009; Van Roey et al., 2012; Forman-Kay and Mittag, 2013; Kosol et al., 2013).

The specific subject of this mini-review concerns the large family of proteins that comprise both folded and intrinsically disordered domains. Multi-domain proteins are ubiquitous in all studied interactomes (Vogel et al., 2004). They are characterized by large-amplitude motions that play important roles in almost every aspect of biomolecular function (Bahar et al., 2007; Smock and Gierasch, 2009; Tzeng and Kalodimos, 2011; Mackereth and Sattler, 2012). Such large-scale modes are often mediated by intrinsically disordered linkers that define the conformational freedom available to the different domains (Shamoo et al., 1995; Ma et al., 2011). Solution-state approaches are essential to probe the ensemble of conformational states sampled by such complex macromolecules. In particular, nuclear magnetic resonance (NMR), small-angle X-ray scattering (SAXS) and single molecule Förster resonance energy transfer (FRET) can be used to investigate the nature of the ensembles of interchanging conformers present in solution (Baber et al., 2001; Margittai et al., 2003; Bernadó et al., 2004, 2007; Henzler-Wildman et al., 2007; Ryabov and Fushman, 2007; Boehr et al., 2009; Clore and Iwahara, 2009; Bashir et al., 2010; Bernadó and Blackledge, 2010; Bertini et al., 2010; Rambo and Tainer, 2010; Boura et al., 2011; Mackereth et al., 2011; Rózycki et al., 2011; Camilloni et al., 2012; Mackereth and Sattler, 2012; Rezaei-Ghaleh et al., 2013; Russo et al., 2013; Hennig et al., 2014).
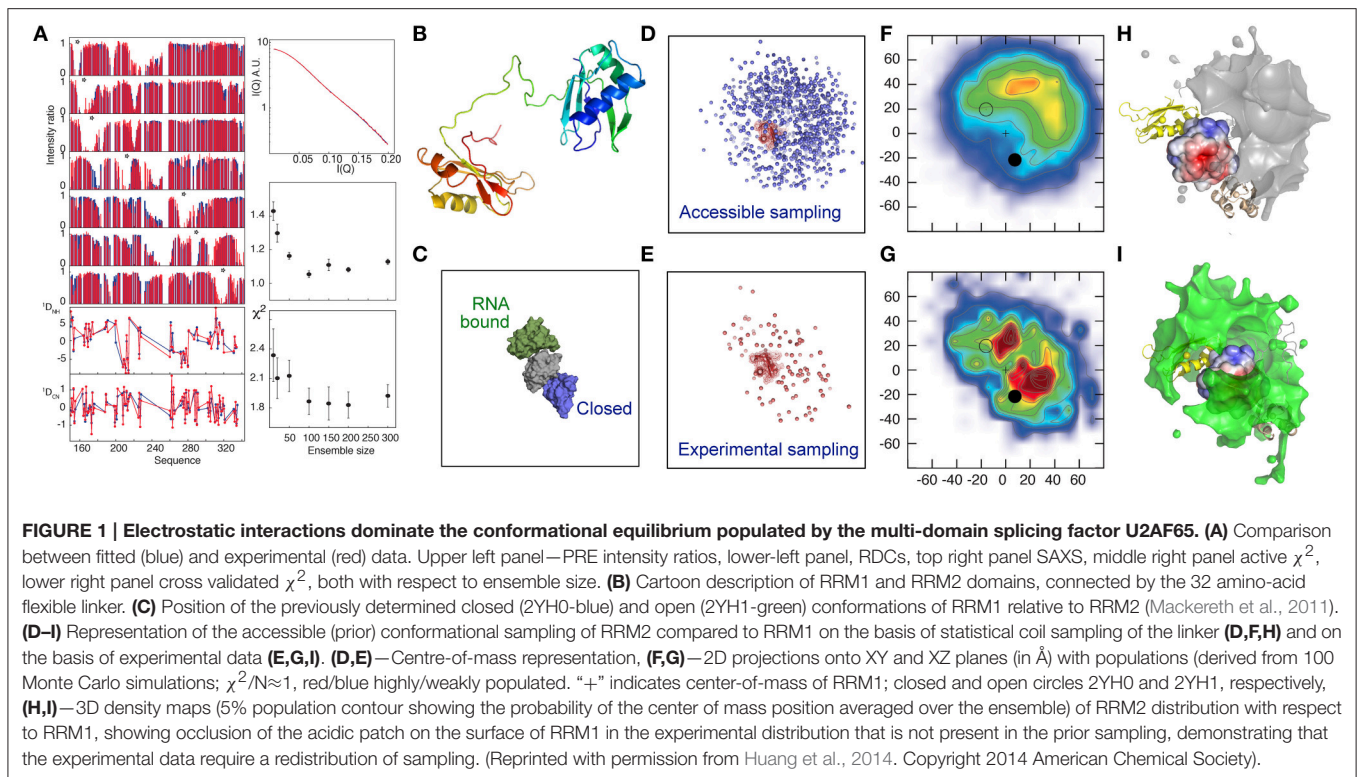
The highly dynamic nature of multi-domain proteins necessitates the development of analytical approaches for the interpretation of the available experimental data in terms of representative ensembles. Although the description of the available conformational space already represents a demanding task, due to the risk of over-fitting, a second, equally formidable requirement is the estimation of the populations of different sub-states and, if possible, their rates of interconversion. The time-scales of this dynamic exchange, and the conformational dynamics that occur within different sub-states also define the

details of the analytical approaches that can be applied, rendering the task yet more daunting.

Here, we will discuss the application of NMR, SAXS, and single molecule FRET to the study of the conformational sampling of two highly flexible multi-domain proteins, the human U2AF65 protein that plays an essential role in the spliceosome assembly (Banerjee et al., 2003; Wahl et al., 2009; Mackereth et al., 2011), and the 627-NLS domain of the PB2 segment of influenza polymerase (Tarendeau et al., 2008; Delaforge et al., 2015). These two cases illustrate very different exchange regimes, impacting the appropriate interpretation of the experimental data.

NMR spin relaxation is potentially a very powerful means to describe both local and global dynamics of macromolecules in solution. Nevertheless, quantitative analysis of the conformational space sampled by two domains relative to each other is highly challenging (Baber et al., 2001; Wong et al., 2009; Ryabov et al., 2012; Xia et al., 2013). More often, paramagnetic relaxation enhancements (PREs) are used to detect weakly populated close contacts between domains, for example transient encounter contacts in protein-protein and protein-nucleic acid complexes (Iwahara et al., 2004; Tang et al., 2006; Volkov et al., 2006). The level of information can be enhanced considerably if the paramagnetic center exhibits anisotropic magnetic susceptibility, in which case pseudo-contact shifts are induced that are sensitive to the distribution of distances and orientations of vectors connecting the observed spins and the electron spin, relative to the susceptibility tensor. Under such conditions, or when dissolved in a dilute liquid crystalline solution, residual dipolar couplings (RDCs) can also be measured to provide information about the distribution of orientations of the structured domains relative to each other (Tolman and Ruan, 2006; Salmon and Blackledge, 2015). In addition, SAXS reports on the pairwise distribution functions of all conformations averaged over the ensemble (Bernadó et al., 2007).

Phenomenologically, large levels of conformational disorder are often manifest by the inability to interpret the experimental data in terms of a single conformation, so that it becomes necessary to invoke the presence of multiple conformations present simultaneously in solution. Substantial efforts targeting ensemble descriptions of flexible multi-domain proteins have been directed toward using NMR and SAXS data to account for this conformational heterogeneity. As mentioned above, care must be exercised in order to avoid over-fitting and to introduce some estimate of uncertainty of populations and representative conformers in the ensemble descriptions. Examples of recent analytical approaches include an estimation of the maximum occurrence of each possible protein conformation on the basis of experimental NMR or SAXS data (Bertini et al., 2007, 2010; Ravera et al., 2014), or weighted-ensemble selection from molecular dynamics (MD) simulation or available crystal structures (Yang et al., 2010; Francis et al., 2011; Russo et al., 2013). Conformational space can also be sampled using replica exchange MD (Sgourakis et al., 2007; Wu et al., 2009; Terakawa and Takada, 2011; Knott and Best, 2012; Narayanan et al., 2012; Zhang et al., 2012; Mittal et al., 2013; Wang et al., 2013) using the experimental data as constraints to guide ensemble distributions

FIGURE 1 | Electrostatic interactions dominate the conformational equilibrium populated by the multi-domain splicing factor U2AF65. (A) Comparison between fitted (blue) and experimental (red) data. Upper left panel—PRE intensity ratios, lower-left panel, RDCs, top right panel SAXS, middle right panel active χ^2, lower right panel cross validated χ^2, both with respect to ensemble size. **(B)** Cartoon description of RRM1 and RRM2 domains, connected by the 32 amino-acid flexible linker. **(C)** Position of the previously determined closed (2YH0-blue) and open (2YH1-green) conformations of RRM1 relative to RRM2 (Mackereth et al., 2011). **(D–I)** Representation of the accessible (prior) conformational sampling of RRM2 compared to RRM1 on the basis of statistical coil sampling of the linker **(D,F,H)** and on the basis of experimental data **(E,G,I)**. **(D,E)**—Centre-of-mass representation, **(F,G)**—2D projections onto XY and XZ planes (in Å) with populations (derived from 100 Monte Carlo simulations; $\chi^2/N\approx1$, red/blue highly/weakly populated. "+" indicates center-of-mass of RRM1; closed and open circles 2YH0 and 2YH1, respectively, **(H,I)**—3D density maps (5% population contour showing the probability of the center of mass position averaged over the ensemble) of RRM2 distribution with respect to RRM1, showing occlusion of the acidic patch on the surface of RRM1 in the experimental distribution that is not present in the prior sampling, demonstrating that the experimental data require a redistribution of sampling. (Reprinted with permission from Huang et al., 2014. Copyright 2014 American Chemical Society).

that reproduce the experimental data on average (Im et al., 2012; Cavalli et al., 2013; Roux and Weare, 2013). Alternatively rigid-body modeling can be used to develop representative ensembles of conformers (Deshmukh et al., 2013).

Conformational sampling can be achieved with high efficiency using restraint-free Monte-Carlo approaches, exploiting for example statistical coil models to describe the backbone dihedral angle distributions of the inter-domain linker. The nature of the intrinsic conformational equilibrium can then be examined by generating representative ensembles of conformers, using an adaptation of the ASTEROIDS approach for ensemble representations of intrinsically disordered systems (Nodet et al., 2009; Salmon et al., 2010; Ozenne et al., 2012; Guerry et al., 2013). Comparison with the experimental data is then used to identify sub-ensembles of conformers that, in combination, represent the Boltzmann distribution in solution. Ensemble descriptions of highly disordered systems are faced with two problems; identification of representative conformational states, and determination of their relative populations. ASTEROIDS uses a genetic algorithm to identify combinations of conformational states that when considered together reproduce the experimental data within estimated experimental uncertainty. Different conformers (i) have populations given by $p_i = 1/n$ where n is the number of conformers in the ensemble. Populations are not optimized, so that if a given state requires a higher weight to fulfill experimental data, additional conformers with similar characteristics will be present. The optimal number of structures necessary to reproduce the complexity of the experimental data can be estimated by cross-validation of independent

experimental data that are not used in the analysis (Salmon et al., 2010; Guerry et al., 2013; Huang et al., 2014). The single-step selection of ASTEROIDS avoids additional optimization of the weights of specific conformers, and is therefore compatible with robust statistical analysis allowing estimation of the confidence levels of both conformation and population.

This approach was adapted to map the free-energy landscape of the RNA Recognition Motifs RRM1 and RRM2 domains of the U2AF65 protein. The two RRMs are connected by a flexible linker and adopt multiple domain arrangements as indicated by a combined analysis of PRE, RDC and SAXS data (Huang et al., 2014). The analysis reveals a heterogeneous ensemble of states dominated by highly populated, but very different relative positions of RRM1 and RRM2, despite the fact that these conformations are not strongly represented in the unrestrained ensemble of states. These conformations were found to resemble previously proposed "closed" and "open" states (Madl et al., 2010; Simon et al., 2010; Mackereth et al., 2011; Hennig et al., 2015), where the latter corresponds to the RNA-bound form of the protein (**Figure 1**), supporting the role of conformational selection from the free-state ensemble as a driving force of this interaction (Mackereth et al., 2011; Mackereth and Sattler, 2012). "Open" and "closed" conformations were found to lie within a continuous density of states, indicating that transition could occur between these states without invoking large conformational jumps. Inspection of the nature of the interacting surfaces dominating the ensemble suggested that the "closed" state was stabilized by electrostatic interactions. This prediction was supported by an observed weakening of

transient contacts in this interface, as detected by a reduction of distinctive PREs with increasing ionic concentration, providing an independent support for the nature of the solution state ensemble that is derived uniquely from experimental data.

In the latter example the experimental data are assumed to report on an ensemble of conformations that are in rapid exchange with respect to the difference in NMR chemical shifts, RDCs and PREs. In effect this assumes that all conformations exchange on timescales faster than or equal to tens of microseconds. Interpretation of the experimental data calculated for each independent sub-state can then be averaged to assess the ability of the ensemble to reproduce the experimental data. This exchange regime will not always be respected, for example in the case of strong interactions between domains, where high activation energies of dissociation may result in much slower exchange between states. Such an exchange regime was recently studied in the 627-NLS multi-domain component of Influenza A polymerase.

The viral RNA polymerase complex is made up of three separate proteins, PA (acidic protein), and PB1 and PB2 (basic proteins 1 and 2) that are imported into the host nucleus following translation to assemble into new polymerase heterotrimers that further catalyze viral RNA replication and transcription. PB2 enters the nucleus by binding to Importin α (Impα) via its C-terminal domain, termed the 221 amino-acid 627-NLS domain (Tarendeau et al., 2008; Kuzuhara et al., 2009) containing a nuclear localization signal (NLS) peptide. The 627-NLS domain is of particular interest, as a large proportion of mutations characterizing adaptation of avian viruses to human hosts are located on its surface (Tarendeau et al., 2008; Mehle and Doudna, 2009). Crystal structures of 627-NLS have been determined, in isolation from both avian and human forms of Influenza (Tarendeau et al., 2008), and more recently in the context of the entire polymerase complex (Pflug et al., 2014), in all cases exhibiting effectively identical compact conformations. This conformation was, however, incompatible with binding of the NLS domain to Impα, suggesting that large-scale conformational rearrangement of either 627-NLS or Impα would be necessary for successful interaction (Boivin and Hart, 2011).

A recent NMR study (Delaforge et al., 2015) of 627-NLS unambiguously revealed the presence of a more complex conformational equilibrium in solution, with approximately twice as many peaks present in the ^{1}H-^{15}N TROSY spectrum as expected, and one set of peaks exhibiting the same chemical shifts as the 627 and NLS domains in isolation. This spectrum
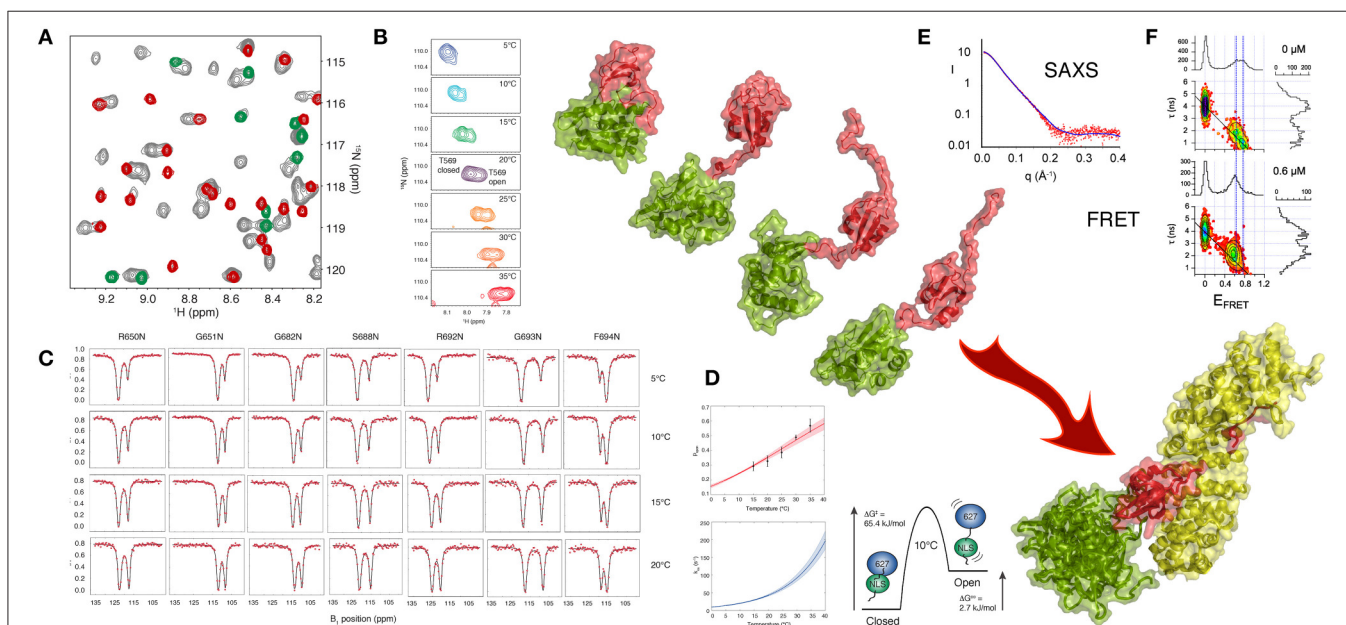
FIGURE 2 | Large Scale Conformational Dynamics Control Influenza Polymerase PB2 627-NLS domain Binding to Importin α. (A) The first evidence that the 627-NLS domain of influenza polymerase PB2 subunit samples two conformations in solution is provided by the ^{15}N-^{1}H correlation spectrum that exhibits two sets of peaks in solution (gray), one set corresponds to resonance positions of either the free 627 (green) or the free NLS (red) domains, suggesting that the two-domain protein exchanges between open and closed forms in solution. **(B)** This exchange is temperature dependent, with increasing population of the open form at higher temperatures as shown in this example showing two peaks reporting on the environment sensed by T569 in the two (open and closed) forms of the protein. The majority of peaks show such behavior, with a large range of ^{15}N and ^{1}H shifts between the two forms. **(C)** Chemical exchange saturation transfer (CEST) of seven resolved peaks, showing "dips" reporting on the open and closed forms of the protein. **(D)** CEST profiles provide information about population, exchange rates and structure, and simultaneous analysis of all spectra in **(C)** allows for the determination of enthalpic and entropic contributions to the equilibrium thermodynamics. **(E)** SAXS of the complex suggests localization of the 627 domain in the vicinity of the C terminus of Importin α. **(F)** smFRET also shows that 627-NLS exhibits domain dynamics in the free form, and demonstrated that only the open form remains in the bound form equilibrium (concentration of titrated Importin is shown in the panels). Large-scale domain dynamics are therefore essential for binding to Importin α. (Reprinted with permission from Delaforge et al., 2015. Copyright 2015 American Chemical Society).

is characteristic of slow exchange (in this case $k_{ex} = 30\,s^{-1}$ at 15°C) between a closed form of the protein (the crystalline conformation) and a previously uncharacterized open form where the two domains retain their secondary and tertiary structure, but dislocate, and evolve independently of each other (**Figure 2**). One set of peaks therefore report on the ensemble of conformations that are in fast exchange, while the other set of peaks report on the closed state that is stabilized by a tripartite salt bridge implicating highly conserved basic and acidic amino acids in the interface between the 627 and NLS domains, and in the 13-amino acid inter-domain linker. The linker becomes flexible upon dislocation, and, as in the case of the U2AF65 RRM1-RRM2 domains, provides the degrees of freedom required for the large-scale domain dynamics characterizing the open form.

Single molecule FRET can also be used to probe the distance distribution and inter-domain dynamics of the two folded domains of 627-NLS. Two populations reporting on the open and closed forms are again observed. The estimated timescale of the different motional modes, characterized by fast interconversion within the open state ($\tau_{ex} < 50\,\mu s$) and slow interconversion between the open and the closed states ($\tau_{ex} > 20\,ms$) also confirms the similarity of the closed state to the crystal conformation, showing negligible rapid motion, and the presence of significant fast domain dynamics in the open state, in good agreement with the NMR analysis.

Exchange between open and closed conformations of 627-NLS was characterized using NMR lineshape analysis and Chemical Exchange Saturation Transfer (CEST), revealing a strong temperature dependence, ranging from 10 to $100\,s^{-1}$ over the range 5–30°C, with the population of the open form increasing significantly from predominantly closed at 5°C ($p = 0.2$) to approximately equal populations at 30°C. Small angle X-ray and neutron scattering data were also measured over the entire temperature range, resulting in very good agreement between expected population-weighted scattering curves from open and closed ensembles and the experimentally determined populations from NMR exchange analysis. Simultaneous analysis of CEST data from multiple different sites throughout the protein as a function of temperature, using an Eyring relationship, provides unique insight into the thermodynamics of the temperature-dependent equilibrium. The activation energy of opening and closing is dominated by enthalpic contributions, while the open state exhibits both entropic and enthalpic contributions compared to the closed state, suggesting additional non-specific non-bonding interactions between the surfaces, as in the case of RRM1/RRM2. Characterization of the temperature-dependence of the equilibrium is of functional interest because viral replication involves adaptation from the warmer bird intestine to the cooler human respiratory system (Massin et al., 2001). Comparison of the temperature dependence of reconfiguration dynamics of 627-NLS from human and bird-adapted proteins is currently underway in our laboratories.

The physiological interaction between 627-NLS and Impα was also investigated using solution methods. The population of the closed form of the protein disappears upon interaction, as shown using single molecule FRET, with the open form maintaining the characteristics of fast distance fluctuations between the attached dyes and therefore suggesting fast inter-domain dynamics in the bound state. Although SAXS measurements of the complex also indicate flexibility of 627 with respect to Impα when bound to NLS, they also reveal that the 627 domain is localized primarily in the vicinity of the C terminus of Impα, possibly stabilized by interactions with the 34 amino acid long intrinsically disordered C terminal tail of Impα.

These observations strongly suggest that while the closed form is known to be necessary for function within the polymerase-RNA complex (Pflug et al., 2014), the open form of 627-NLS is required for interaction with Impα. Interestingly, mutation of R650A in the 627 domain removes the possibility of forming the salt bridge, resulting in suppression of the closed form in solution. This same mutation was independently shown to abrogate polymerase activity in the nucleus, while still allowing for nuclear import (Kirui et al., 2014), substantiating our model whereby the open-form mediates nuclear import. Further structural investigations of different forms of the RNA polymerase complex suggest that an open 627-NLS may also play additional roles in remodeling the polymerase structure during the viral cycle (Hengrung et al., 2015; Thierry et al., 2016).

The two studies described above demonstrate the functional importance of large-amplitude dynamics mediated by the intrinsically disordered linker peptides in multi-domain assemblies. In both cases the analysis suggests a role for conformational selection from an intrinsic pre-existing equilibrium in the interaction with physiological partners. Both studies show how solution-state approaches can be used to understand the role of complex dynamic conformational equilibria in biomolecular function and, more specifically, that these mechanisms could not be understood without a detailed description of the ensemble of states sampled in solution.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Baber, J. L., Szabo, A., and Tjandra, N. (2001). Analysis of slow interdomain motion of macromolecules using NMR relaxation data. *J. Am. Chem. Soc.* 123, 3953–3959. doi: 10.1021/ja0041876

Bahar, I., Chennubhotla, C., and Tobi, D. (2007). Intrinsic dynamics of enzymes in the unbound state and, relation to allosteric regulation. *Curr. Opin. Struct. Biol.* 17, 633–640. doi: 10.1016/j.sbi.2007.09.011

Banerjee, H., Rahn, A., Davis, W., and Singh, R. (2003). Sex lethal and U2 small nuclear ribonucleoprotein auxiliary factor (U2AF65) recognize polypyrimidine tracts using multiple modes of binding. *RNA* 9, 88–99. doi: 10.1261/rna.2131603

Bashir, Q., Volkov, A. N., Ullmann, G. M., and Ubbink, M. (2010). Visualization of the encounter ensemble of the transient electron transfer complex of cytochrome c and cytochrome c peroxidase. *J. Am. Chem. Soc.* 132, 241–247. doi: 10.1021/ja9064574

Bernadó, P., Fernandes, M. X., Jacobs, D. M., Fiebig, K., García de la Torre, J., and Pons, M. (2004). Interpretation of NMR relaxation properties of Pin1, a two-domain protein, based on Brownian dynamic simulations. *J. Biomol. NMR* 29, 21–35. doi: 10.1023/B:JNMR.0000019499.60777.6e

Bernadó, P., and Blackledge, M. (2010). Structural biology: proteins in dynamic equilibrium. *Nature* 468, 1046–1048. doi: 10.1038/4681046a

Bernadó, P., Mylonas, E., Petoukhov, M., Blackledge, M., and Svergun, D. (2007). Structural characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem. Soc.* 129, 5656–5664. doi: 10.1021/ja069124n

Bertini, I., Giachetti, A., Luchinat, C., Parigi, G., Petoukhov, M. V., Pierattelli, R., et al. (2010). Conformational space of flexible biological macromolecules from average data. *J. Am. Chem. Soc.* 132, 13553–13558. doi: 10.1021/ja1063923

Bertini, I., Gupta, Y. K., Luchinat, C., Parigi, G., Peana, M., Sgheri, L., et al. (2007). Paramagnetism-based NMR restraints provide maximum allowed probabilities for the different conformations of partially independent protein domains. *J. Am. Chem. Soc.* 129, 12786–12794. doi: 10.1021/ja0726613

Boehr, D. D., Nussinov, R., and Wright, P. E. (2009). The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* 5, 789–796. doi: 10.1038/nchembio.232

Boivin, S., and Hart, D. J. (2011). Interaction of the influenza A virus polymerase PB2 C-terminal region with importin α isoforms provides insights into host adaptation and polymerase assembly. *J. Biol. Chem.* 286, 10439–10448. doi: 10.1074/jbc.M110.182964

Boura, E., Rózycki, B., Herrick, D. Z., Chung, H. S., Vecer, J., Eaton, W. A., et al. (2011). Solution structure of the ESCRT-I complex by small-angle X-ray scattering, EPR, and FRET spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* 108, 9437–9442. doi: 10.1073/pnas.1101763108

Camilloni, C., Robustelli, P., De Simone, A., Cavalli, A., and Vendruscolo, M. (2012). Characterization of the Conformational Equilibrium between the two major substates of RNase a using NMR chemical shifts. *J. Am. Chem. Soc.* 134, 3968–3971. doi: 10.1021/ja210951z

Cavalli, A., Camilloni, C., and Vendruscolo, M. (2013). Molecular dynamics simulations with replica-averaged structural restraints generate structural ensembles according to the maximum entropy principle. *J. Chem. Phys.* 138:094112. doi: 10.1063/1.4793625

Clore, G. M., and Iwahara, J. (2009). Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chem. Rev.* 109, 4108–4139. doi: 10.1021/cr900033p

Delaforge, E., Milles, S., Bouvignies, G., Bouvier, D., Boivin, S., Salvi, N., et al. (2015). Large-scale conformational dynamics control H5N1 influenza polymerase PB2 binding to importinα. *J. Am. Chem. Soc.* 137, 15122–15134. doi: 10.1021/jacs.5b07765

Deshmukh, L., Schwieters, C. D., Grishaev, A., Ghirlando, R., Baber, J. L., and Clore, G. M. (2013). Structure and dynamics of full-length HIV-1 capsid protein in solution. *J. Am. Chem. Soc.* 135, 16133–16147. doi: 10.1021/ja406246z

Dunker, A. K., Silman, I., Uversky, V. N., and Sussman, J. L. (2008). Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.* 18, 756–764. doi: 10.1016/j.sbi.2008.10.002

Dyson, H. J., and Wright, P. E. (2005). Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* 6, 197–208. doi: 10.1038/nrm1589

Fink, A. L. (2005). Natively unfolded proteins. *Curr. Opin. Struct. Biol.* 15, 35–41. doi: 10.1016/j.sbi.2005.01.002

Forman-Kay, J. D., and Mittag, T. (2013). From sequence and forces to structure, function, and evolution of intrinsically disordered proteins. *Structure* 21, 1492–1499. doi: 10.1016/j.str.2013.08.001

Francis, D. M., Rozycki, B., Koveal, D., Hummer, G., Page, R., and Peti, W. (2011). Structural basis of p38α regulation by hematopoietic tyrosine phosphatase. *Nat. Chem. Biol.* 7, 916–924. doi: 10.1038/nchembio.707

Guerry, P., Salmon, L., Mollica, L., Ortega Roldan, J.-L., Markwick, P., van Nuland, N. A. J., et al. (2013). Mapping the population of protein conformational energy sub-States from NMR dipolar couplings. *Angew. Chem. Int. Ed Engl.* 52, 3181–3185. doi: 10.1002/anie.201209669

Hengrung, N., El Omari, K., Serna Martin, I., Vreede, F. T., Cusack, S., Rambo, R. P., et al. (2015). Crystal structure of the RNA-dependent RNA polymerase from influenza C virus. *Nature* 527, 114–117. doi: 10.1038/nature15525

Hennig, J., Militti, C., Popowicz, G. M., Wang, I., Sonntag, M., Geerlof, A., et al. (2014). Structural basis for the assembly of the Sxl-Unr translation regulatory complex. *Nature* 515, 287–290. doi: 10.1038/nature13693

Hennig, J., Warner, L. R., Simon, B., Geerlof, A., Mackereth, C. D., and Sattler, M. (2015). Structural analysis of protein-RNA complexes in solution using NMR paramagnetic relaxation enhancements. *Methods Enzymol.* 558, 333–362. doi: 10.1016/bs.mie.2015.02.006

Henzler-Wildman, K., Lei, M., Thai, V., Kerns, S., Karplus, M., and Kern, D. (2007). A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* 450, 913. doi: 10.1038/nature06407

Huang, J.-R., Warner, L. R., Sanchez, C., Gabel, F., Madl, T., Mackereth, C. D., et al. (2014). Transient electrostatic interactions dominate the conformational equilibrium sampled by multidomain splicing factor U2AF65: a combined NMR and SAXS study. *J. Am. Chem. Soc.* 136, 7068–7076. doi: 10.1021/ja502030n

Im, W., Jo, S., and Kim, T. (2012). An ensemble dynamics approach to decipher solid-state NMR observables of membrane proteins. *Biochim. Biophys. Acta* 1818, 252–262. doi: 10.1016/j.bbamem.2011.07.048

Iwahara, J., Schwieters, C. D., and Clore, G. M. (2004). Characterization of nonspecific protein-DNA interactions by 1H paramagnetic relaxation enhancement. *J. Am. Chem. Soc.* 126, 12800–12808. doi: 10.1021/ja046246b

Kirui, J., Bucci, M. D., Poole, D. S., and Mehle, A. (2014). Conserved features of the PB2 627 domain impact influenza virus polymerase function and replication. *J. Virol.* 88, 5977–5986. doi: 10.1128/JVI.00508-14

Knott, M., and Best, R. B. (2012). A preformed binding interface in the unbound ensemble of an intrinsically disordered protein: evidence from molecular simulations. *PLoS Comput. Biol.* 8:e1002605. doi: 10.1371/journal.pcbi.1002605

Kosol, S., Contreras-Martos, S., Cedeño, C., and Tompa, P. (2013). Structural characterization of intrinsically disordered proteins by NMR spectroscopy. *Mol. Basel Switz.* 18, 10802–10828. doi: 10.3390/molecules180910802

Kragelj, J., Palencia, A., Nanao, M. H., Maurin, D., Bouvignies, G., Blackledge, M., et al. (2015). Structure and dynamics of the MKK7-JNK signaling complex. *Proc. Natl. Acad. Sci. U.S.A.* 112, 3409–3414. doi: 10.1073/pnas.1419528112

Kuzuhara, T., Kise, D., Yoshida, H., Horita, T., Murazaki, Y., Nishimura, A., et al. (2009). Structural basis of the influenza A virus RNA polymerase PB2 RNA-binding domain containing the pathogenicity-determinant lysine 627 residue. *J. Biol. Chem.* 284, 6855–6860. doi: 10.1074/jbc.C800224200

Ma, B., Tsai, C.-J., Haliloglu, T., and Nussinov, R. (2011). Dynamic allostery: linkers are not merely flexible. *Structure* 19, 907–917. doi: 10.1016/j.str.2011.06.002

Mackereth, C. D., Madl, T., Bonnal, S., Simon, B., Zanier, K., Gasch, A., et al. (2011). Multi-domain conformational selection underlies pre-mRNA splicing regulation by U2AF. *Nature* 475, 408–U174. doi: 10.1038/nature10171

Mackereth, C. D., and Sattler, M. (2012). Dynamics in multi-domain protein recognition of RNA. *Curr. Opin. Struct. Biol.* 22, 287–296. doi: 10.1016/j.sbi.2012.03.013

Madl, T., Felli, I. C., Bertini, I., and Sattler, M. (2010). Structural analysis of protein interfaces from 13C direct-detected paramagnetic relaxation enhancements. *J. Am. Chem. Soc.* 132, 7285–7287. doi: 10.1021/ja1014508

Margittai, M., Widengren, J., Schweinberger, E., Schröder, G. F., Felekyan, S., Haustein, E., et al. (2003). Single-molecule fluorescence resonance

energy transfer reveals a dynamic equilibrium between closed and open conformations of syntaxin 1. *Proc. Natl. Acad. Sci. U.S.A.* 100, 15516–15521. doi: 10.1073/pnas.2331232100

Massin, P., van der Werf, S., and Naffakh, N. (2001). Residue 627 of PB2 is a determinant of cold sensitivity in RNA replication of avian influenza viruses. *J. Virol.* 75, 5398–5404. doi: 10.1128/JVI.75.11.5398-5404.2001

Mehle, A., and Doudna, J. A. (2009). Adaptive strategies of the influenza virus polymerase for replication in humans. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21312–21316. doi: 10.1073/pnas.0911915106

Milles, S., Mercadante, D., Aramburu, I. V., Jensen, M. R., Banterle, N., Koehler, C., et al. (2015). Plasticity of an ultrafast interaction between nucleoporins and nuclear transport receptors. *Cell* 163, 734–745. doi: 10.1016/j.cell.2015.09.047

Mittal, J., Yoo, T. H., Georgiou, G., and Truskett, T. M. (2013). Structural ensemble of an intrinsically disordered polypeptide. *J. Phys. Chem. B.* 117, 118–124. doi: 10.1021/jp308984e

Narayanan, C., Weinstock, D. S., Wu, K.-P., Baum, J., and Levy, R. M. (2012). Investigation of the polymeric properties of α-synuclein and comparison with NMR experiments: a replica exchange molecular dynamics study. *J. Chem. Theory Comput.* 8, 3929–3942. doi: 10.1021/ct300241t

Nodet, G., Salmon, L., Ozenne, V., Meier, S., Jensen, M. R., and Blackledge, M. (2009). Quantitative description of backbone conformational sampling of unfolded proteins at amino acid resolution from NMR residual dipolar couplings. *J. Am. Chem. Soc.* 131, 17908–17918. doi: 10.1021/ja9069024

Ozenne, V., Schneider, R., Yao, M., Huang, J.-R., Salmon, L., Zweckstetter, M., et al. (2012). Mapping the potential energy landscape of intrinsically disordered proteins at amino Acid resolution. *J. Am. Chem. Soc.* 134, 15138–15148. doi: 10.1021/ja306905s

Pflug, A., Guilligay, D., Reich, S., and Cusack, S. (2014). Structure of influenza A polymerase bound to the viral RNA promoter. *Nature* 516, 355–360. doi: 10.1038/nature14008

Rambo, R. P., and Tainer, J. A. (2010). Bridging the solution divide: comprehensive structural analyses of dynamic RNA, DNA, and protein assemblies by small-angle X-ray scattering. *Curr. Opin. Struct. Biol.* 20, 128–137. doi: 10.1016/j.sbi.2009.12.015

Ravera, E., Salmon, L., Fragai, M., Parigi, G., Al-Hashimi, H., and Luchinat, C. (2014). Insights into domain-domain motions in proteins and RNA from solution NMR. *Acc. Chem. Res.* 47, 3118–3126. doi: 10.1021/ar5002318

Rezaei-Ghaleh, N., Klama, F., Munari, F., and Zweckstetter, M. (2013). Predicting the rotational tumbling of dynamic multidomain proteins and supramolecular complexes. *Angew. Chem. Int. Ed Engl.* 52, 11410–11414. doi: 10.1002/anie.201305094

Roux, B., and Weare, J. (2013). On the statistical equivalence of restrained-ensemble simulations with the maximum entropy method. *J. Chem. Phys.* 138:084107. doi: 10.1063/1.4792208

Rózycki, B., Kim, Y. C., and Hummer, G. (2011). SAXS ensemble refinement of ESCRT-III CHMP3 conformational transitions. *Structure* 19, 109–116. doi: 10.1016/j.str.2010.10.006

Russo, L., Maestre-Martinez, M., Wolff, S., Becker, S., and Griesinger, C. (2013). Interdomain dynamics explored by paramagnetic NMR. *J. Am. Chem. Soc.* 135, 17111–17120. doi: 10.1021/ja408143f

Ryabov, Y., Clore, G. M., and Schwieters, C. D. (2012). Coupling between internal dynamics and rotational diffusion in the presence of exchange between discrete molecular conformations. *J. Chem. Phys.* 136, 34108. doi: 10.1063/1.3675602

Ryabov, Y. E., and Fushman, D. (2007). A model of interdomain mobility in a multidomain protein. *J. Am. Chem. Soc.* 129, 3315–3327. doi: 10.1021/ja067667r

Salmon, L., and Blackledge, M. (2015). Investigating protein conformational energy landscapes and atomic resolution dynamics from NMR dipolar couplings: a review. *Rep. Prog. Phys.* 78:126601. doi: 10.1088/0034-4885/78/12/126601

Salmon, L., Nodet, G., Ozenne, V., Yin, G., Jensen, M., Zweckstetter, M., et al. (2010). NMR characterization of long-range order in intrinsically disordered proteins. *J. Am. Chem. Soc.* 132, 8407–8418. doi: 10.1021/ja101645g

Salmon, L., Ortega Roldan, J.-L., Lescop, E., Licinio, A., van Nuland, N., Jensen, M. R., et al. (2011). Structure, dynamics, and kinetics of weak protein-protein complexes from NMR spin relaxation measurements of titrated solutions. *Angew. Chem. Int Ed Engl.* 50, 3755–3759. doi: 10.1002/anie.201100310

Schneider, R., Maurin, D., Communie, G., Kragelj, J., Hansen, D. F., Ruigrok, R. W. H., et al. (2015). Visualizing the molecular recognition trajectory of an intrinsically disordered protein using multinuclear relaxation dispersion NMR. *J. Am. Chem. Soc.* 137, 1220–1229. doi: 10.1021/ja511066q

Sgourakis, N. G., Yan, Y., McCallum, S. A., Wang, C., and Garcia, A. E. (2007). The Alzheimer's peptides Aβ40 and 42 adopt distinct conformations in water: a combined MD/NMR study. *J. Mol. Biol.* 368, 1448–1457. doi: 10.1016/j.jmb.2007.02.093

Shamoo, Y., Abdul-Manan, N., and Williams, K. R. (1995). Multiple RNA binding domains (RBDs) just don't add up. *Nucleic Acids Res.* 23, 725–728. doi: 10.1093/nar/23.5.725

Simon, B., Madl, T., Mackereth, C. D., Nilges, M., and Sattler, M. (2010). An efficient protocol for NMR-spectroscopy-based structure determination of protein complexes in solution. *Angew. Chem. Int. Ed.* 49, 1967–1970. doi: 10.1002/anie.200906147

Smock, R. G., and Gierasch, L. M. (2009). Sending signals dynamically. *Science* 324, 198–203. doi: 10.1126/science.1169377

Sugase, K., Dyson, H. J., and Wright, P. E. (2007). Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* 447, 1021. doi: 10.1038/nature05858

Tang, C., Iwahara, J., and Clore, G. M. (2006). Visualization of transient encounter complexes in protein-protein association. *Nature* 444, 383–386. doi: 10.1038/nature05201

Tarendeau, F., Crepin, T., Guilligay, D., Ruigrok, R. W. H., Cusack, S., and Hart, D. J. (2008). Host determinant residue lysine 627 lies on the surface of a discrete, folded domain of influenza virus polymerase PB2 subunit. *PLoS Pathog.* 4:e1000136. doi: 10.1371/journal.ppat.1000136

Terakawa, T., and Takada, S. (2011). Multiscale ensemble modeling of intrinsically disordered proteins: p53 N-terminal domain. *Biophys. J.* 101, 1450–1458. doi: 10.1016/j.bpj.2011.08.003

Thierry, E., Guilligay, D., Kosinski, J., Bock, T., Gaudon, S., Round, A., et al. (2016). Influenza polymerase can adopt an alternative configuration involving a radical repacking of PB2 domains. *Mol. Cell* 61, 125–137. doi: 10.1016/j.molcel.2015.11.016

Tolman, J., and Ruan, K. (2006). NMR residual dipolar couplings as probes of biomolecular dynamics. *Chem. Rev.* 106, 1720–1736. doi: 10.1021/cr040429z

Tompa, P., and Fuxreiter, M. (2008). Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci.* 33, 2–8. doi: 10.1016/j.tibs.2007.10.003

Tzeng, S.-R., and Kalodimos, C. G. (2011). Protein dynamics and allostery: an NMR view. *Curr. Opin. Struct. Biol.* 21, 62–67. doi: 10.1016/j.sbi.2010.10.007

Uversky, V. N., and Dunker, A. K. (2010). Understanding protein non-folding. *Biochim. Biophys. Acta* 1804, 1231–1264. doi: 10.1016/j.bbapap.2010.01.017

Van Roey, K., Gibson, T. J., and Davey, N. E. (2012). Motif switches: decision-making in cell regulation. *Curr. Opin. Struct. Biol.* 22, 378–385. doi: 10.1016/j.sbi.2012.03.004

Vaynberg, J., and Qin, J. (2006). Weak protein-protein interactions as probed by NMR spectroscopy. *Trends Biotechnol.* 24, 22–27. doi: 10.1016/j.tibtech.2005.09.006

Vogel, C., Bashton, M., Kerrison, N. D., Chothia, C., and Teichmann, S. A. (2004). Structure, function and evolution of multidomain proteins. *Curr. Opin. Struct. Biol.* 14, 208–216. doi: 10.1016/j.sbi.2004.03.011

Volkov, A. N., Worrall, J. A. R., Holtzmann, E., and Ubbink, M. (2006). Solution structure and dynamics of the complex between cytochrome c and cytochrome c peroxidase determined by paramagnetic NMR. *Proc. Natl. Acad. Sci. U.S.A.* 103, 18945–18950. doi: 10.1073/pnas.0603551103

Wahl, M. C., Will, C. L., and Lührmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. *Cell* 136, 701–718. doi: 10.1016/j.cell.2009.02.009

Wang, Y., Chu, X., Longhi, S., Roche, P., Han, W., Wang, E., et al. (2013). Multiscaled exploration of coupled folding and binding of an intrinsically disordered molecular recognition element in measles virus nucleoprotein. *Proc. Natl. Acad. Sci. U.S.A.* 110, E3743–E3752. doi: 10.1073/pnas.1308381110

Wong, V., Case, D. A., and Szabo, A. (2009). Influence of the coupling of interdomain and overall motions on NMR relaxation. *Proc. Natl. Acad. Sci. U.S.A.* 106, 11016–11021. doi: 10.1073/pnas.0809994106

Wright, P. E., and Dyson, H. J. (2009). Linking folding and binding. *Curr. Opin. Struct. Biol.* 19, 31–38. doi: 10.1016/j.sbi.2008.12.003

Wu, K.-P., Weinstock, D. S., Narayanan, C., Levy, R. M., and Baum, J. (2009). Structural reorganization of α-synuclein at low pH observed by NMR and REMD simulations. *J. Mol. Biol.* 391, 784–796. doi: 10.1016/j.jmb.2009.06.063

Xia, J., Deng, N., and Levy, R. M. (2013). NMR relaxation in proteins with fast internal motions and slow conformational exchange: model-free framework and Markov state simulations. *J. Phys. Chem. B.* 117, 6625–6634. doi: 10.1021/jp400797y

Yang, S., Blachowicz, L., Makowski, L., and Roux, B. (2010). Multidomain assembled states of Hck tyrosine kinase in solution. *Proc. Natl. Acad. Sci. U.S.A.* 107, 15757–15762. doi: 10.1073/pnas.1004569107

Zhang, W., Ganguly, D., and Chen, J. (2012). Residual structures, conformational fluctuations, and electrostatic interactions in the synergistic folding of two intrinsically disordered proteins. *PLoS Comput. Biol.* 8:e1002353. doi: 10.1371/journal.pcbi.1002353