

Review

Intrinsically Disordered Proteins in a Physics-Based World

Timothy H. Click [†], Debabani Ganguly [†] and Jianhan Chen ^{*}

Department of Biochemistry, Kansas State University, Manhattan, KS 66506, USA

[†] These authors contributed equally to this work.

^{*} Author to whom correspondence should be addressed; E-Mail: jianhanc@ksu.edu;
Tel.: +1-785-532-2518; Fax: +1-785-532-7278.

Received: 19 October 2010; in revised form: 17 December 2010 / Accepted: 17 December 2010 /
Published: 21 December 2010

Abstract: Intrinsically disordered proteins (IDPs) are a newly recognized class of functional proteins that rely on a lack of stable structure for function. They are highly prevalent in biology, play fundamental roles, and are extensively involved in human diseases. For signaling and regulation, IDPs often fold into stable structures upon binding to specific targets. The mechanisms of these coupled binding and folding processes are of significant importance because they underlie the organization of regulatory networks that dictate various aspects of cellular decision-making. This review first discusses the challenge in detailed experimental characterization of these heterogeneous and dynamics proteins and the unique and exciting opportunity for physics-based modeling to make crucial contributions, and then summarizes key lessons from recent *de novo* simulations of the structure and interactions of several regulatory IDPs.

Keywords: conformational selection; generalized Born; implicit solvent; induced folding; molecular dynamics; p21; p27; p53; pKID; replica exchange

1. Introduction

The prevalence and fundamental roles of intrinsically disordered proteins (IDPs) in biology were not recognized until the late 1990s. Observations accumulated that many functional proteins, particularly those involved in transcription and translation regulation, appeared to be intrinsically unstructured [1]. This prompted Wright and Dyson to publish their seminal review in 1999 [2], calling

for a reassessment of the protein structure-function paradigm. Since then, study of IDPs has rapidly evolved into a field of “growing prominence” and “one of the most exciting undertakings of structural biology” as predicted [3,4]. It is now established that IDPs are: 1. highly abundant in biology, with about one-third of eukaryotic proteins predicted to be IDPs [5,6]; 2. frequently involved in crucial areas such as signaling and regulation [7–9]; and 3. extensively associated with human diseases including cancer, diabetes, neurodegenerative diseases and amyloidoses [10,11]. Many IDPs undergo folding transitions upon binding to specific targets [12], even though some IDPs remain unstructured in the bound states [13,14]. The mechanisms of these coupled binding-folding interactions are of significant importance because they underlie the organization of important regulatory networks that inform various aspects of cellular decision-making, cellular fate, and cellular signaling. Diseases and disorders can result from IDP mis-folding as well as mis-signaling and mis-regulation [11]. While the prevalence of IDPs as functional proteins was not fully recognized until the 1990s as stated above, one reviewer pointed out several pieces of work that were early hargingers. To illustrate this, three examples have been chosen here from a much larger collection. As early as 1953, optical rotation suggested that the milk protein casein is similar to globular proteins unfolded by 5 M guanidinium hydrochloride [15]. By the mid-1960s, optical rotation was used to develop a conformational classification of proteins, which included proteins that were characterized as disordered [16]. By the early 1970s, studies on fibrinogen strongly indicated that this protein contained large regions that lack structure [17], and more recently, this early history was reviewed in a very interesting fashion [18].

For signaling and regulation, the disordered nature of IDPs is believed to offer several unique advantages [2,3], including high specificity/low affinity binding, inducibility by posttranslational modifications, and structural plasticity for binding multiple targets. The last property appears to be a hallmark of IDPs that allows one-to-many and many-to-one signaling [19,20]. Intrinsic thermo-instability can also offer a robust mechanism for allosteric coupling [21]. However, the molecular mechanism of the coupled binding and folding interaction of IDPs has remained largely elusive except in a very small number of cases [22–24]. In particular, while it is recognized that residual structures often persist in unbound IDPs [7], their functional implications are under much debate [25]. Clearly, the extent of residual structure modulates the entropic cost of folding, and thus affects the binding affinity. It has been further proposed that preformed structural elements might serve as initial contact points and facilitate IDP folding on the substrate surface [26,27]. On the other hand, increasing the amount of local structures in the unbound state actually reduces the binding rate for several IDPs [28,29], emphasizing the importance of intrinsic flexibility in facile recognition of IDPs.

Important progresses have been made over the last ten years or so in prediction, identification and general characterization of IDPs [30–34]. In particular, sequence analysis can be quite reliably applied to predict disordered protein segments with averaged sensitivity and specificity scores greater than 0.8 [35]. The Database of Protein Disorder (DisProt, version 5.3, 09/21/2010) currently contains 1,284 experimentally verified disordered regions within 594 proteins, and over forty atomistic structures of IDP complexes are now available [36]. One of the key challenges in IDP research at present is to understand mechanistically how intrinsic flexibility is exploited together with nascent structures and other biophysical signatures of IDPs for biological functions such as efficient and versatile binding. Such understanding is necessary for deciphering the elaborate physiological control of IDP function and how such control might fail in human diseases. This review will discuss the key challenges in

experimental characterization of unbound IDPs, how physics-based *de novo* simulations might help to provide important missing details of IDP structure and interaction, and important lessons that have been derived from a limited number of recent *de novo* simulations of regulatory IDPs. We note that a significant amount of computational work exists on studies of disordered proteins involved in protein aggregation and amyloidogenesis [37]. This review focuses on signaling and regulatory IDPs that often undergo coupled binding and folding during function. Studies of amyloidogenic proteins will only be discussed in the context of common challenges and opportunities for *de novo* simulations.

2. Challenges in Detailed Experimental Characterization of IDPs

Structural information of the bound state alone is not sufficient to establish biologically relevant regulatory mechanisms. For IDPs, the nature of the unbound state holds important clues to function. Due to the heterogeneous and dynamical nature, detailed characterization of unbound IDPs has proven to be a principal challenge. Consequently, a lack of understanding of the nature of residual structures in unbound IDPs hinders further clarification of their functional roles. Among a wide range of biophysical techniques available for characterizing disordered protein states [34], biomolecular NMR is probably the most comprehensive [38]. Many observables can be measured for multiple sites throughout the protein to infer (transient) organizations at the secondary and tertiary levels, including: chemical shift, coupling constant, nuclear Overhauser effect (NOE), residual dipolar coupling (RDC), paramagnetic resonance enhancement (PRE), and spin relaxation. Chemical shifts, coupling constants and NOEs can be used to determine secondary structure propensities semi-quantitatively at the residue level [39–42]. PRE and RDC are powerful tools for uncovering the existence of transient tertiary organizations [43–45]. They have been applied to derive important insights on many disordered protein states (see recent reviews [7,38,46,47]). However, a quantitative, structural interpretation of these experimental observables measured on unbound IDPs is generally not feasible. A key issue is that only ensemble-averaged properties can be measured in general and they must be represented as averaged quantities of a heterogeneous structure ensemble. Coupled with typical scarcity of data, the structural calculation is severely underdetermined and a unique structure ensemble cannot be determined solely from the experimental restraints. This critical limitation has not been fully appreciated in the literature. Certain intrinsic properties of RDC and PRE further complicate the structural interpretation. For example, PRE is extremely sensitive to the electron-nucleus distance with r^{-6} dependence [48]. While such sensitivity allows detection of weakly populated transient contacts, it also renders PRE largely insensitive to the majority of accessible conformations. Using theoretical PRE data sets derived from simulated disordered protein states, we recently demonstrated that ensemble structural calculation protocols with PRE distance restraints [49,50] could generate misleading ensembles that reflect little on the true underlying protein state [51]. Direct incorporation of RDC in disordered ensemble calculation has not been feasible because of the intricate relationship between observed data and individual ensemble structure, even though strategies exist for globular proteins [52–54]. Thus, RDC has been mainly used for validation of various (coil) models of unfolded states [55,56] or as filters to select conformations from pre-generated structural pools [57,58]. Importantly, this general difficulty in structural interpretation arises mainly from the fundamental limitations of ensemble-averaged properties measured on disordered protein states rather than

technical ones. There is simply not sufficient constraint for independent identification of representative ensembles from experimental restraints alone. A few exceptions may include cases where additional information on the accessible conformational substates is available or can be assumed [59,60], or where unusually large amount of experimental data were measured [61].

3. Opportunities and Challenges for *de Novo* Simulations

The substantial challenge in detailed experimental characterization of IDPs presents a unique and exciting opportunity for molecular modeling to make critical contributions. In particular, atomistic simulations using physics-based empirical molecular mechanics force fields [62,63] arguably provides the ultimate level of detail necessary for understanding disordered protein states. The experimental data can be used as independent validation of simulated ensembles instead of as restraints. Such a strategy avoids the under-determined structural calculation problem that tends to generate a false sense of excellent agreement with experiment. At the same time, the dynamic and unstructured nature of IDPs also presents substantial new challenges, pushing the limit on both force field accuracy and conformational sampling capability. Traditional explicit inclusion of water molecules arguably provides the most realistic description of solvent, but also significantly increases the system size (~10-fold). This can lead to prohibitive computational cost if one wants to sufficiently sample the broad manifold of functionally relevant states of IDPs. Moreover, mainly optimized for folded native states, current explicit solvent protein force fields are known to suffer from systematic biases in describing peptide conformational equilibria, such as tendency to over-stabilize helices [64,65] and peptide-peptide interactions [66]. These existing force field limitations need to be carefully considered in simulations of IDPs. Thanks to development of powerful computational hardware and advanced sampling techniques, important advances are being made recently in optimization of explicit solvent protein force fields [67–69]. Nonetheless, correcting the systematic biases in a consistent and transferable fashion has been difficult, and the computational cost is one of the main obstacles.

Alternatively, implicit solvent has emerged as a powerful approach for atomistic simulation of protein conformational equilibria that provides a balance between description of the essential physics and computational feasibility [70]. The principle and practice of implicit solvent is well documented [70,71]. The basic idea is to capture the mean influence of water on the solute via direct estimation of the solvation free energy. As such, only the solute is represented atomistically, and the system size is reduced ~10-fold. Important advances have been made recently to greatly improve the efficiency and achievable accuracy of implicit solvent, particularly with the generalized Born (GB) approximation. For example, protein simulations in the GBSW implicit solvent is ~30 times faster than equivalent ones in explicit water [72]. Importantly, a substantial gain in efficiency allows careful optimization of implicit solvent protein force fields to suppress certain systematic biases that have plagued explicit solvent ones [62]. The key is to capture the delicate balance of competing solvation and intramolecular interactions on the peptide and protein level. We recently re-balanced the GBSW protein force field based on pair-wise interactions of side chain analogs and conformational properties of model peptides [73]. The optimized force field does not only recapitulate the experimental structures *and* stabilities of several helical peptides and a series of β -hairpins with a wide range of stability, but also folds hairpin trpzip2 and mini-protein Trp-cage [73]. The same force field has also been successfully applied to pKa

prediction [74], pH-dependent protein folding [75–77], structure refinement [78], and recently to simulation of regulatory IDPs [79,80]. Similar optimization efforts have also led to substantial improvement in other GB models [81,82]. An ABSINTH implicit solvent has also been developed and optimized specifically for IDPs [83].

It is important to recognize that inherent and methodological drawbacks do exist in implicit solvent [70]. They need to be carefully considered in the interpretation of implicit solvent simulations. This is particularly important considering that validation based on ensemble-averaged properties is not conclusive because of the same reasons detailed above for difficulty in structural calculations. Implicit solvent will not properly describe short-range effects where the detailed interplay of a few non-bulk-like water molecules is important. With a lack of solvent granularity, continuum models do not capture all the fine structures in potentials of mean force (PMFs) of interactions [73,84], and the conformational diffusion kinetics is altered. The temperature dependence cannot be described accurately in general either. Besides these intrinsic limitations, implicit solvent might be further limited by the specific methodology for calculating the solvation free energy as well as the physical parameters of the solvation model and underlying protein force fields. In particular, the surface-area (SA)-based treatment of nonpolar solvation appears to be a key methodological limitation in the current GB implicit solvent [85,86]. Nevertheless, a substantial reduction in the computational cost without compromising the essential physics is an important advantage of implicit solvent. By taking care in interpretation and validation of simulations, one can expect that *de novo* simulations to provide reliable details on the structure and interaction of IDPs.

Even with the dramatic reduction in system size using implicit solvent, it is challenging to sufficiently sample biologically accessible, functionally relevant conformational space of IDPs. Conventional constant temperature molecular dynamics (MD) is generally insufficient for achieving convergence in simulated structural ensembles of IDPs. The difficulty arises not only because of the large and complex conformational space of proteins, but also due to significant energy barriers that might separate different conformational subspaces. It is necessary to exploit various advanced techniques to enhance sampling [87,88]. One particularly simple yet effective technique is replica exchange (REX) [89]. The basic idea is to simulate multiple independent replicas of the system at different temperatures, typically distributed exponentially between the temperature of interest and a maximum temperature. Periodically, replicas attempt to exchange simulation temperatures according to a Metropolis criterion that preserves the detailed balance and ensures proper canonical ensembles at all temperatures. The resulting random walk in the temperature space helps the system to avoid being kinetically trapped in states of local energy minima. The required number of replicas increases with the system size N (as a function of its squared root [89]). With ~ 10 -fold smaller system size, implicit solvent is thus particularly suitable for REX. Recent theoretical considerations [90–92] and actual simulations of small peptides [93–96] generally confirm that REX can enhance protein conformational sampling as long as the activation enthalpies (of conformational transitions) are positive. In particular, it is important to specify a maximum REX temperature slightly above where the folding rate maximizes [91,92]. The key to obtain well-converged structural ensembles of IDPs is efficient sampling of transitions between conformational substates. However, the nature of these substates and the energy barriers of their inter-conversion are not known. Therefore, the efficacy of REX for IDPs is not very obvious, nor is the optimal choice of key REX-MD parameters such as the number of

replicas, range and distribution of simulation temperatures and exchange attempt frequency. On the other hand, several recent experimental studies suggest that the average roughness of protein energy surface likely exceeds 5 RT [97–99], which can lead to a strong temperature dependence of conformation diffusion [100]. The roughness consideration argues strongly for the general ability of REX to enhance sampling of protein conformations including those of IDPs.

Another particularly attractive approach to overcome the sampling bottleneck is to combine large numbers of equilibrium and/or generalized ensemble simulations (e.g., on distributed computing platforms [101]) using network methods like the Markov State Models (MSMs) [102–104]. Even though it is yet to be applied to IDPs, this strategy has provided unprecedented detail on energy landscapes of several proteins under both stable and unstable conditions [105–107]. It is important to emphasize that sampling efficiency also strongly depends on the force field quality. In particular, modern protein force fields, with explicit or implicit solvent, tend to overestimate the strength of protein-protein interactions [66]. Consequently, non-specific collapsed protein states are often over-stabilized, which severely hinders conformational sampling. When the goal is not to generate proper thermodynamic ensembles at physiological temperatures, temperature-induced unfolding/unbinding simulations can be used to infer on the mechanism of coupled binding and folding of IDPs given the complex structures. This strategy has been quite successful in studies of protein folding [108,109], and has already been applied to understand IDP interactions [110,111]. An interesting advantage of high-temperature simulations is that they are less sensitive to imperfections in the force fields, besides reduced computational cost. A key concern is that the transition states or the most probable transition paths might depend on the temperature. However, IDP complexes tend to be less stable than globular ones, and the temperature required for unbinding simulations will likely be moderate.

4. Key Lessons from Recent *de Novo* Simulations of Regulatory IDPs

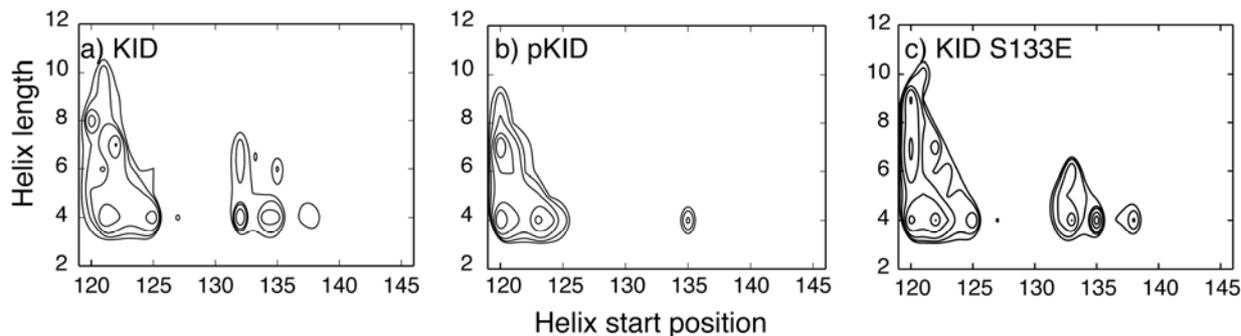
4.1. Unbound IDPs: Nascent Structures and Dependence on Post-Translational Modifications

As discussed above, the nature of the unbound state of an IDP holds important clues to how it might interact with specific targets. Abundant experimental evidence exists to suggest that various levels of residual structures persist in unbound IDPs. Such examples include several domains of transcription factor CREB [7], activation and regulatory domains of tumor suppressor p53 [112–114], the steroid receptor coactivator ACTR [115], the synuclein family [116,117], and cyclin-dependent kinase (Cdk) regulators p21 and p27 [118,119]. A key goal of *de novo* simulations of IDPs has thus been to provide molecular detail of such residual structures in unbound IDPs. In the following, we summarize key results from several recent *de novo* simulations of free regulatory IDPs where the goal is to understand coupled binding and folding. We note that physics-based simulations have also provided important insights into the intrinsic conformational properties of IDPs and how they might depend on certain physiochemical properties [120,121]. In particular, IDPs are known to have low sequence complexity with enriched charges. Several recent studies have shown that the structure of IDPs is correlated with the charge content [121,122]. It has been further suggested charges on IDPs can directly affect function, in both protein-DNA [123] and protein-protein interactions [80].

The kinase-inducible domain (KID) of CREB is one of the most extensively studied IDPs. In cell, CREB regulates transcription in response to cAMP signaling partially by binding to the coactivator

CBP (CREB binding protein) [124]. This recognition requires phosphorylation of Ser133 of KID (pKID) and involves direct interaction with the KIX domain of CBP [125,126]. When in complex with KIX, pKID adopts a helix-linker-helix structure that involves residues 119 to 146 [127]. However, free KID and pKID lack stable tertiary structures in solution. NMR chemical shift analysis has estimated that helix α A (residues 120–129) is about 50–60% folded and helix α B (residues 134–144) is only 10–15% formed [128]. Furthermore, secondary chemical shift analysis does not reveal significant differences in mean residue helicities upon phosphorylation [128]. Accordingly, the principal role of pSer133 has been attributed to mediating intermolecular interactions. Slot *et al.* have investigated the intrinsic structural properties of the unbound KID and pKID using explicit solvent MD simulations [129]. It was argued that phosphorylation could reduce the loop flexibility connecting two helices in the bound structure and induce a transient structural element that resembles native-like conformation. However, these simulations were conducted by using a simulated annealing protocol in explicit solvent. The total simulation length is only 2.0 nanoseconds (ns) with the last 1.0 ns as the production stage at 300 K. This is too short to sufficiently sample relevant conformations in the disordered state, as reflected in the narrow backbone ϕ/ψ distributions reported [129]. Persistence of native-like loop conformation in pKID could be an artifact of limited sampling initiated from the folded structure. This concern substantially weakens the reliability of the proposed effects of phosphorylation. Well-converged structural ensembles of KID and pKID were later calculated using 200-ns REX-MD simulations in the optimized GBSW force field [73]. The simulated ensembles are mainly validated by comparing the mean residue helicities with the NMR results. Further structural analyses show that both KID and pKID are compact and mainly occupy a small number of helical substates (see Figure 1). Interestingly, even though in agreement with NMR pSer133 only leads to marginal helicity changes on the ensemble level, the underlying conformational substates differ significantly. In particular, pSer133 appears to restrict the accessible conformational space of the loop connecting two helical segments and thus can reduce the entropic cost of KID folding upon binding to KIX. This entropic contribution supplements the salt-bridge interactions between KID pSer133 and KIX Lys662 and Tyr658. Replacing pSer133 with a Glu residue fails to induce similar structural changes (see Figure 1c). This explains why KID S133E does not bind to KIX nearly as strong as pKID [130], even though Glu can interact with similar strength with Lys or Tyr [131]. However, in contrast to Slot *et al.*'s simulations, pS133 does not appear to induce native-like loop conformations. The expanded role of phosphorylation in regulating the KID/KIX recognition was not recognized in the NMR studies due to lack of experimentally detectable conformational changes on the ensemble level. This highlights the importance of combining simulation and experiment for detailed characterization of unbound IDPs. The ability of well-converged *de novo* simulations to “predict” important features of the disordered states of a 28-residue IDP is nontrivial and represents an encouraging progress towards accurate simulation of IDP conformational equilibria.

Figure 1. 2D probability distributions of helix starting positions and lengths in the unbound structure ensembles of (a) KID, (b) pKID and (c) KID S133E at 302 K calculated from REX-MD folding simulations in GBSW implicit solvent [79]. Contours are drawn at 0.002, 0.005, 0.01, 0.02, 0.04, and 0.06 levels.



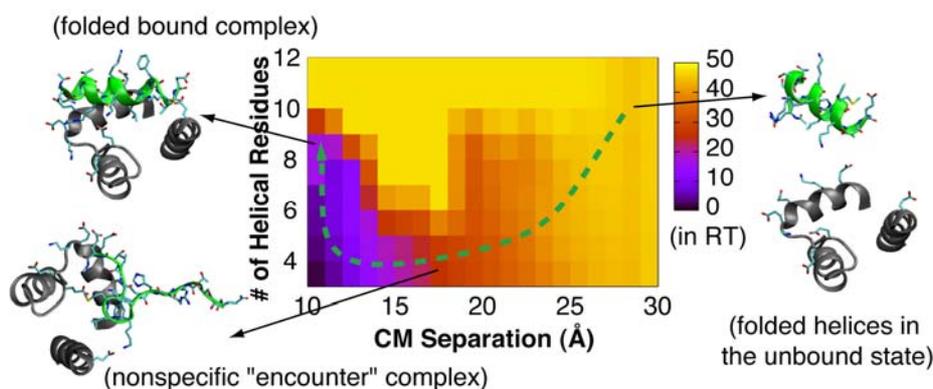
REX-MD simulations in implicit solvent have also been used to compute the structure ensemble of unbound p53 extreme C-terminus [80]. The results demonstrate that the free p53 peptide samples several distinct conformations, which, importantly, correspond to those experimentally observed when bound to different partners [19]. Presence of folded-like nascent structures has also been suggested in integrated computational and experimental studies of several other regulatory IDPs. Kriwacki and coworkers combined NMR data with a 100-ns explicit solvent simulation and suggested that sub-domains within free p27 kinase inhibitory domain exhibit partially folded nascent structures that resemble the bound conformation observed in the ternary complex of p27/Cdk2/cyclin A [118]. Yoon *et al.* characterized the residual structural elements of the C-terminal segment of p21, also using a combination of NMR and MD simulations, and identified helical and extended conformations that resemble those in complex with either calmodulin or proliferating cell nuclear antigen (PCNA) [119]. However, the role of nascent structures within these free IDPs in binding is not obvious, even though one might be tempted to speculate that such preformed structures might facilitate binding by providing initial binding sites (i.e., conformational selection-like mechanisms) [26,27].

4.2. Mechanism of Coupled Binding and Folding: Intrinsic Flexibility vs. Preformed Structures

Compared to studies of unbound IDPs, there are even fewer *de novo* computational analyses of the molecular mechanisms of coupled binding and folding of IDPs. This can be primarily attributed to much greater computational cost. Rigorous thermodynamic and/or kinetic characterization of coupled binding and folding of IDPs is generally not feasible. In a pioneering study, Verkhivker *et al.* investigated the mechanism of p27 binding to Cdk2/cyclin A using high-temperature unfolding and unbinding simulations with a simplified all-atom protein energy function with physics-motivated terms [132]. The calculated transition state ensemble (TSE) suggests an induced folding-like mechanism, where molecular recognition is initiated by formation of nonspecific encounter complexes with unstructured p27. This finding supports the “fly-casting” mechanism proposed by Wolynes and coworkers [133,134], which argues that nonspecific binding of unstructured conformations confers a kinetic advantage for binding. Another key lesson obtained from this study is that intermolecular interactions appear to dedicate the folding mechanism of p27 and overwhelm local folding preferences

in unbound p27. This conclusion might prove to be applicable to other regulatory IDPs. In particular, structural plasticity for adopting distinctive folded states is believed to be a hallmark of regulatory IDPs. Therefore, the binding partner ultimately determines the IDP topology in the bound state, *i.e.*, intermolecular interactions do overwhelm local folding preferences of IDPs prior to binding. Indeed, induced-folding was also observed in a recent high-temperature simulation of pKID/KIX [135].

Figure 2. 2D free energy surface of the folding and binding of p53 to S100B($\beta\beta$). Folding is described by the number of helical residues, and the binding by the center-of-mass (CM) separation [80]. Representative structures are shown for regions that correspond to the unbound, nonspecific contact, and bound states. Charged residues on p53 and near its binding site on S100B($\beta\beta$) are shown in sticks



The only attempt to rigorously calculate the binding and folding free energy surfaces of IDPs reported so far is on the p53/S100B($\beta\beta$) interaction [80]. A protocol that combines REX and umbrella sampling was used, and several approximations were necessary to further reduce the computational cost even with implicit solvent. In contrast to the speculation of conformational selection based on the presence of substantial helical content in the unbound state, the calculated energy surfaces reveal that the p53 extreme C-terminus initially binds to S100B($\beta\beta$) in an unfolded conformation and then quickly folds - an example of induced folding (e.g., see the green path in Figure 2). The residual structure in the unbound state thus mainly modulates the binding affinity, while the intrinsic flexibility is critical for the binding rate. This study illustrates that the presence of folded-like conformations in unbound IDPs is not sufficient evidence to establish conformational selection. It is important to directly examine the thermodynamics and/or kinetics of coupled binding and folding for mechanistic characterization. This study also reveals important remaining limitations in the current implicit solvent force field, which does not only substantially over-estimate the interaction strength but also fails to predict the experimental structure as the global free energy minimum.

5. Conclusions

IDPs are an important class of functional proteins, with high abundance and fundamental roles in biology and broad association with human diseases. Important progresses have been made in the prediction, identification and high-level characterization of disordered protein segments. Nevertheless, much needs to be learned about the nature of the unbound state and the molecular mechanism of coupled binding and folding. Detailed experimental characterization of IDPs is challenging due to the

heterogeneous and dynamic nature. IDP thus represents an exciting opportunity for physics-based modeling to make critical contributions. At the same time, simulation of IDP presents substantial new challenges that push the limit on both protein force field accuracy and conformational sampling capability. New modeling strategies need to be explored to address these challenges. Implicit solvent, coupled with advanced sampling techniques, provides a particularly suitable means for *de novo* characterization of IDP structure and interaction with a necessary balance between accuracy and efficiency. Only a small number of *de novo* simulations have been reported on regulatory IDPs so far, reflecting the existence of substantial challenges in the current simulation methodologies. Nonetheless, important insights have been obtained on the nascent structures and how they might contribute to molecular recognition of IDPs. With the sampling limitation alleviated by ever-increasing computing power and sophisticated algorithms that exploit abundant computing resources, the critical bottleneck for physics-based simulations of IDPs is in the achievable level accuracy of current protein force fields. This underpins intensive ongoing efforts in improving the ability of explicit and implicit solvent protein force fields to describe protein conformational equilibria. It is important to diligently validate *de novo* simulations using available experimental data and proper (positive and negative) controls. We anticipate an integrated computational and experimental strategy to address specific questions on the structure and interaction of biologically important IDPs, and are excited by the unique role that physics-based simulations will be expected (and required) to play.

Acknowledgements

We thank the Terry C. Johnson Center for Basic Cancer Research for an Innovative Research Award, and this work was supported by an NSF CAREER Award (MCB 0952514). This work is contribution 10-174-J from the Kansas Agricultural Experiment Station.

References

1. Kriwacki, R.W.; Hengst, L.; Tennant, L.; Reed, S.I.; Wright, P.E. Structural studies of p21(Waf1/Cip1/Sdi1) in the free and Cdk2-bound state: Conformational disorder mediates binding diversity. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 11504–11509.
2. Wright, P.E.; Dyson, H.J. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **1999**, *293*, 321–331.
3. Tompa, P. Intrinsically unstructured proteins. *Trends Biochem. Sci.* **2002**, *27*, 527–533.
4. Fink, A.L. Natively unfolded proteins. *Curr. Opin. Struct. Biol.* **2005**, *15*, 35–41.
5. Dunker, A.K.; Brown, C.J.; Lawson, J.D.; Iakoucheva, L.M.; Obradovic, Z. Intrinsic disorder and protein function. *Biochemistry* **2002**, *41*, 6573–6582.
6. Dunker, A.K.; Lawson, J.D.; Brown, C.J.; Williams, R.M.; Romero, P.; Oh, J.S.; Oldfield, C.J.; Campen, A.M.; Ratliff, C.R.; Hipps, K.W.; Ausio, J.; Nissen, M.S.; Reeves, R.; Kang, C.H.; Kissinger, C.R.; Bailey, R.W.; Griswold, M.D.; Chiu, M.; Garner, E.C.; Obradovic, Z. Intrinsically disordered protein. *J. Mol. Graph. Model* **2001**, *19*, 26–59.
7. Dyson, H.J.; Wright, P.E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197–208.

8. Uversky, V.N.; Oldfield, C.J.; Dunker, A.K. Showing your ID: Intrinsic disorder as an ID for recognition, regulation and cell signaling. *J. Mol. Recognit.* **2005**, *18*, 343–384.
9. Smock, R.G.; Gierasch, L.M. Sending signals dynamically. *Science* **2009**, *324*, 198–203.
10. Iakoucheva, L.M.; Brown, C.J.; Lawson, J.D.; Obradovic, Z.; Dunker, A.K. Intrinsic disorder in cell-signaling and cancer-associated proteins. *J. Mol. Biol.* **2002**, *323*, 573–584.
11. Uversky, V.N.; Oldfield, C.J.; Dunker, A.K. Intrinsically disordered proteins in human diseases: Introducing the D-2 concept. *Annu. Rev. Biophys.* **2008**, *37*, 215–246.
12. Wright, P.E.; Dyson, H.J. Linking folding and binding. *Curr. Opin. Struct. Biol.* **2009**, *19*, 31–38.
13. Mittag, T.; Orlicky, S.; Choy, W.Y.; Tang, X.J.; Lin, H.; Sicheri, F.; Kay, L.E.; Tyers, M.; Forman-Kay, J.D. Dynamic equilibrium engagement of a polyvalent ligand with a single-site receptor. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 17772–17777.
14. Mittag, T.; Marsh, J.; Grishaev, A.; Orlicky, S.; Lin, H.; Sicheri, F.; Tyers, M.; Forman-Kay, J.D. Structure/function implications in a dynamic complex of the intrinsically disordered Sic1 with the Cdc4 subunit of an SCF ubiquitin ligase. *Structure* **2010**, *18*, 494–506.
15. Meekin, T.J. Milk Proteins. *J. Food Protect.* **1952**, *15*, 57–63.
16. Jirgensons, B. Classification of proteins according to conformation. *Macromol. Chem. Phys.* **1966**, *91*, 74–86.
17. Doolittle, R.F. Structural aspects of the fibrinogen to fibrin conversion. *Adv. Protein Chem.* **1973**, *27*, 1–109.
18. Doolittle, R.F.; Kollman, J.M. Natively unfolded regions of the vertebrate fibrinogen molecule. *Proteins* **2006**, *63*, 391–397.
19. Oldfield, C.J.; Meng, J.; Yang, J.Y.; Yang, M.Q.; Uversky, V.N.; Dunker, A.K. Flexible nets: disorder and induced fit in the associations of p53 and 14-3-3 with their partners. *BMC Genomics* **2008**, *9*, S1.
20. Dunker, A.K.; Cortese, M.S.; Romero, P.; Iakoucheva, L.M.; Uversky, V.N. Flexible nets - The roles of intrinsic disorder in protein interaction networks. *FEBS J.* **2005**, *272*, 5129–5148.
21. Hilser, V.J.; Thompson, E.B. Intrinsic disorder as a mechanism to optimize allosteric coupling in proteins. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 8311–8315.
22. Sugase, K.; Dyson, H.J.; Wright, P.E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **2007**, *447*, 1021–1025.
23. Narayanan, R.; Ganesh, O.K.; Edison, A.S.; Hagen, S.J. Kinetics of folding and binding of an intrinsically disordered protein: The inhibitor of yeast aspartic proteinase YPrA. *J. Am. Chem. Soc.* **2008**, *130*, 11477–11485.
24. Lu, Q.; Lu, H.P.; Wang, J. Exploring the mechanism of flexible biomolecular recognition with single molecule dynamics. *Phys. Rev. Lett.* **2007**, *98*, 128105.
25. Keskin, O.; Gursoy, A.; Ma, B.; Nussinov, R. Principles of protein-protein interactions: What are the preferred ways for proteins to interact? *Chem. Rev.* **2008**, *108*, 1225–1244.
26. Receveur-Brechot, V.; Bourhis, J.M.; Uversky, V.N.; Canard, B.; Longhi, S. Assessing protein disorder and induced folding. *Proteins* **2006**, *62*, 24–45.
27. Fuxreiter, M.; Simon, I.; Friedrich, P.; Tompa, P. Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *J. Mol. Biol.* **2004**, *338*, 1015–1026.

28. Zor, T.; Mayr, B.M.; Dyson, H.J.; Montminy, M.R.; Wright, P.E. Roles of phosphorylation and helix propensity in the binding of the KIX domain of CREB-binding protein by constitutive (c-Myb) and inducible (CREB) activators. *J. Biol. Chem.* **2002**, *277*, 42241–42248.
29. Bienkiewicz, E.A.; Adkins, J.N.; Lumb, K.J. Functional consequences of preorganized helical structure in the intrinsically disordered cell-cycle inhibitor p27(Kip1). *Biochemistry* **2002**, *41*, 752–759.
30. Dunker, A.K.; Silman, I.; Uversky, V.N.; Sussman, J.L. Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.* **2008**, *18*, 756–764.
31. Bourhis, J.M.; Canard, B.; Longhi, S. Predicting protein disorder and induced folding: From theoretical principles to practical applications. *Curr. Protein Pept. Sci.* **2007**, *8*, 135–149.
32. Bracken, C.; Iakoucheva, L.M.; Rorner, P.R.; Dunker, A.K. Combining prediction, computation and experiment for the characterization of protein disorder. *Curr. Opin. Struct. Biol.* **2004**, *14*, 570–576.
33. Bartlett, A.I.; Radford, S.E. An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nat. Struct. Mol. Biol.* **2009**, *16*, 582–588.
34. Eliezer, D. Biophysical characterization of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **2009**, *19*, 23–30.
35. Noivirt-Brik, O.; Prilusky, J.; Sussman, J.L. Assessment of disorder predictions in CASP8. *Proteins* **2009**, *77*, S210–S216.
36. Meszaros, B.; Tompa, P.; Simon, I.; Dosztanyi, Z. Molecular principles of the interactions of disordered proteins. *J. Mol. Biol.* **2007**, *372*, 549–561.
37. Keskin, O.; Gursoy, A.; Ma, B.; Nussinov, R. Towards drugs targeting multiple proteins in a systems biology approach. *Curr. Top Med. Chem.* **2007**, *7*, 943–951.
38. Dyson, H.J.; Wright, P.E. Unfolded proteins and protein folding studied by NMR. *Chem. Rev.* **2004**, *104*, 3607–3622.
39. Wishart, D.S.; Sykes, B.D. The C-13 Chemical-Shift Index—a simple method for the identification of protein secondary structure using C-13 chemical-shift data. *J. Biomol. NMR* **1994**, *4*, 171–180.
40. Wang, A.C.; Bax, A. Determination of the backbone dihedral angles phi in human ubiquitin from reparametrized empirical Karplus equations. *J. Am. Chem. Soc.* **1996**, *118*, 2483–2494.
41. Cornilescu, G.; Delaglio, F.; Bax, A. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* **1999**, *13*, 289–302.
42. Marsh, J.A.; Singh, V.K.; Jia, Z.C.; Forman-Kay, J.D. Sensitivity of secondary structure propensities to sequence differences between alpha- and gamma-synuclein: Implications for fibrillation. *Protein Sci.* **2006**, *15*, 2795–2804.
43. Gillespie, J.R.; Shortle, D. Characterization of long-range structure in the denatured state of staphylococcal nuclease. 1. Paramagnetic relaxation enhancement by nitroxide spin labels. *J. Mol. Biol.* **1997**, *268*, 158–169.
44. Gillespie, J.R.; Shortle, D. Characterization of long-range structure in the denatured state of staphylococcal nuclease. 2. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. *J. Mol. Biol.* **1997**, *268*, 170–184.

45. Shortle, D.; Ackerman, M.S. Persistence of native-like topology in a denatured protein in 8 M urea. *Science* **2001**, *293*, 487–489.
46. Mittag, T.; Forman-Kay, J.D. Atomic-level characterization of disordered protein ensembles. *Curr. Opin. Struct. Biol.* **2007**, *17*, 3–14.
47. Vendruscolo, M. Determination of conformationally heterogeneous states of proteins. *Curr. Opin. Struct. Biol.* **2007**, *17*, 15–20.
48. Bloembergen, N.; Morgan, L.O. Proton relaxation times in paramagnetic solutions effects of electron spin relaxation. *J. Chem. Phys.* **1961**, *34*, 842–850.
49. Lindorff-Larsen, K.; Kristjansdottir, S.; Teilum, K.; Fieber, W.; Dobson, C.M.; Poulsen, F.M.; Vendruscolo, M. Determination of an ensemble of structures representing the denatured state of the bovine acyl-coenzyme A binding protein. *J. Am. Chem. Soc.* **2004**, *126*, 3291–3299.
50. Vise, P.; Baral, B.; Stancik, A.; Lowry, D.F.; Daughdrill, G.W. Identifying long-range structure in the intrinsically unstructured transactivation domain of p53. *Proteins* **2007**, *67*, 526–530.
51. Ganguly, D.; Chen, J. Structural interpretation of paramagnetic relaxation enhancement-derived distances for disordered protein states. *J. Mol. Biol.* **2009**, *390*, 467–477.
52. Delaglio, F.; Kontaxis, G.; Bax, A. Protein structure determination using molecular fragment replacement and NMR dipolar couplings. *J. Am. Chem. Soc.* **2000**, *122*, 2142–2143.
53. Hus, J.C.; Marion, D.; Blackledge, M. Determination of protein backbone structure using only residual dipolar couplings. *J. Am. Chem. Soc.* **2001**, *123*, 1541–1542.
54. Rathinavelan, T.; Im, W. A novel strategy to determine protein structures using exclusively residual dipolar coupling. *J. Comput. Chem.* **2008**, *29*, 1640–1649.
55. Jha, A.K.; Colubri, A.; Freed, K.F.; Sosnick, T.R. Statistical coil model of the unfolded state: Resolving the reconciliation problem. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13099–13104.
56. Bernado, P.; Blanchard, L.; Timmins, P.; Marion, D.; Ruigrok, R.W.H.; Blackledge, M. A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 17002–17007.
57. Marsh, J.A.; Baker, J.M.R.; Tollinger, M.; Forman-Kay, J.D. Calculation of residual dipolar couplings from disordered state ensembles using local alignment. *J. Am. Chem. Soc.* **2008**, *130*, 7804–7805.
58. Jensen, M.R.; Salmon, L.; Nodet, G.; Blackledge, M. Defining conformational ensembles of intrinsically disordered and partially folded proteins directly from chemical shifts. *J. Am. Chem. Soc.* **2010**, *132*, 1270–1272.
59. Felitsky, D.J.; Lietzow, M.A.; Dyson, H.J.; Wright, P.E. Modeling transient collapsed states of an unfolded protein to provide insights into early folding events. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 6278–6283.
60. Jensen, M.R.; Houben, K.; Lescop, E.; Blanchard, L.; Ruigrok, R.W.; Blackledge, M. Quantitative conformational analysis of partially folded proteins from residual dipolar couplings: application to the molecular recognition element of Sendai virus nucleoprotein. *J. Am. Chem. Soc.* **2008**, *130*, 55–61.
61. Marsh, J.A.; Neale, C.; Jack, F.E.; Choy, W.Y.; Lee, A.Y.; Crowhurst, K.A.; Forman-Kay, J.D. Improved structural characterizations of the drkN SH3 domain unfolded state suggest a compact ensemble with native-like and non-native structure. *J. Mol. Biol.* **2007**, *367*, 1494–1510.

62. Mackerell, A.D. Empirical force fields for biological macromolecules: Overview and issues. *J. Comput. Chem.* **2004**, *25*, 1584–1604.
63. Ponder, J.W.; Case, D.A. Force fields for protein simulations. *Adv. Protein Chem.* **2003**, *66*, 27–85.
64. Yoda, T.; Sugita, Y.; Okamoto, Y. Secondary-structure preferences of force fields for proteins evaluated by generalized-ensemble simulations. *Chem. Phys.* **2004**, *307*, 269–283.
65. Best, R.B.; Buchete, N.V.; Hummer, G. Are current molecular dynamics force fields too helical? *Biophys. J.* **2008**, *95*, L07–L09.
66. Kang, M.; Smith, P.E. A Kirkwood-Buff derived force field for amides. *J. Comput. Chem.* **2006**, *27*, 1477–1485.
67. Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J.L.; Dror, R.O.; Shaw, D.E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* **2010**, *78*, 1950–1958.
68. Best, R.B.; Mittal, J. Balance between alpha and beta structures in ab initio protein folding. *J. Phys. Chem. B* **2010**, *114*, 8790–8798.
69. Mittal, J.; Best, R.B. Tackling force-field bias in protein folding simulations: folding of Villin HP35 and Pin WW domains in explicit water. *Biophys. J.* **2010**, *99*, L26–L28.
70. Chen, J.H.; Brooks, C.L.; Khandogin, J. Recent advances in implicit solvent based methods for biomolecular simulations. *Curr. Opin. Struct. Biol.* **2008**, *18*, 140–148.
71. Roux, B.; Simonson, T. Implicit solvent models. *Biophys. Chem.* **1999**, *78*, 1–20.
72. Im, W.P.; Lee, M.S.; Brooks, C.L. Generalized born model with a simple smoothing function. *J. Comput. Chem.* **2003**, *24*, 1691–1702.
73. Chen, J.H.; Im, W.P.; Brooks, C.L. Balancing solvation and intramolecular interactions: Toward a consistent generalized born force field. *J. Am. Chem. Soc.* **2006**, *128*, 3728–3736.
74. Khandogin, J.; Brooks, C.L. Toward the accurate first-principles prediction of ionization equilibria in proteins. *Biochemistry* **2006**, *45*, 9363–9373.
75. Khandogin, J.; Chen, J.H.; Brooks, C.L. Exploring atomistic details of pH-dependent peptide folding. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 18546–18550.
76. Khandogin, J.; Brooks, C.L. Linking folding with aggregation in Alzheimer's beta-amyloid peptides. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 16880–16885.
77. Khandogin, J.; Raleigh, D.P.; Brooks, C.L. Folding intermediate in the villin headpiece domain arises from disruption of a N-terminal hydrogen-bonded network. *J. Am. Chem. Soc.* **2007**, *129*, 3056–3057.
78. Chen, J.H.; Brooks, C.L. Can molecular dynamics simulations provide high-resolution refinement of protein structure? *Proteins* **2007**, *67*, 922–930.
79. Ganguly, D.; Chen, J. Atomistic details of the disordered states of KID and pKID. implications in coupled binding and folding. *J. Am. Chem. Soc.* **2009**, *131*, 5214–5223.
80. Chen, J.H. Intrinsically disordered p53 extreme C-terminus binds to S100B(beta-beta) through "fly-casting". *J. Am. Chem. Soc.* **2009**, *131*, 2088–2089.
81. Okur, A.; Strockbine, B.; Hornak, V.; Simmerling, C. Using PC clusters to evaluate the transferability of molecular mechanics force fields for proteins. *J. Comput. Chem.* **2003**, *24*, 21–31.

82. Jang, S.; Kim, E.; Pak, Y. Direct folding simulation of alpha-helices and beta-hairpins based on a single all-atom force field with an implicit solvation model. *Proteins* **2007**, *66*, 53–60.
83. Vitalis, A.; Pappu, R.V. ABSINTH: A New Continuum Solvation Model for Simulations of Polypeptides in Aqueous Solutions. *J. Comput. Chem.* **2009**, *30*, 673–699.
84. Masunov, A.; Lazaridis, T. Potentials of mean force between ionizable amino acid side chains in water. *J. Am. Chem. Soc.* **2003**, *125*, 1722–1730.
85. Chen, J.H.; Brooks, C.L. Critical importance of length-scale dependence in implicit modeling of hydrophobic interactions. *J. Am. Chem. Soc.* **2007**, *129*, 2444–2445.
86. Chen, J.H.; Brooks, C.L. Implicit modeling of nonpolar solvation for simulating protein folding and conformational transitions. *Phys. Chem. Chem. Phys.* **2008**, *10*, 471–481.
87. Scheraga, H.A.; Khalili, M.; Liwo, A. Protein-folding dynamics: Overview of molecular simulation techniques. *Annu. Rev. Phys. Chem.* **2007**, *58*, 57–83.
88. Lei, H.; Duan, Y. Improved sampling methods for molecular simulation. *Curr. Opin. Struct. Biol.* **2007**, *17*, 187–191.
89. Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141–151.
90. Zuckerman, D.M.; Lyman, E. A second look at canonical sampling of biomolecules using replica exchange simulation. *J. Chem. Theor. Comput.* **2006**, *2*, 1200–1202.
91. Zheng, W.; Andrec, M.; Gallicchio, E.; Levy, R.M. Simulating replica exchange simulations of protein folding with a kinetic network model. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 15340–15345.
92. Nymeyer, H. How efficient is replica exchange molecular dynamics? An analytic approach. *J. Chem. Theor. Comput.* **2008**, *4*, 626–636.
93. Zhang, W.; Wu, C.; Duan, Y. Convergence of replica exchange molecular dynamics. *J. Chem. Phys.* **2005**, *123*, 154105:1–154105:9.
94. Periolo, X.; Mark, A.E. Convergence and sampling efficiency in replica exchange simulations of peptide folding in explicit solvent. *J. Chem. Phys.* **2007**, *126*, 014903:1–014903:11.
95. Sindhikara, D.; Meng, Y.L.; Roitberg, A.E. Exchange frequency in replica exchange molecular dynamics. *J. Chem. Phys.* **2008**, *128*, 024103:1–024103:10.
96. Rao, F.; Caflisch, A. Replica exchange molecular dynamics simulations of reversible folding. *J. Chem. Phys.* **2003**, *119*, 4035–4042.
97. Rico, F.; Moy, V.T. Energy landscape roughness of the streptavidin-biotin interaction. *J. Mol. Recognit.* **2007**, *20*, 495–501.
98. Janovjak, H.; Sapra, K.T.; Kedrov, A.; Muller, D.J. From valleys to ridges: Exploring the dynamic energy landscape of single membrane proteins. *ChemPhysChem* **2008**, *9*, 954–966.
99. Kapon, R.; Nevo, R.; Reich, Z. Protein energy landscape roughness. *Biochem. Soc. Trans.* **2008**, *36*, 1404–1408.
100. Zwanzig, R. Diffusion in a rough potential. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 2029–2030.
101. Shirts, M.; Pande, V.S. Computing-Screen savers of the world unite! *Science* **2000**, *290*, 1903–1904.
102. Bowman, G.R.; Beauchamp, K.A.; Boxer, G.; Pande, V.S. Progress and challenges in the automated construction of Markov state models for full protein systems. *J. Chem. Phys.* **2009**, *131*, 124101:1–124101:11.

103. Bowman, G.R.; Huang, X.H.; Pande, V.S. Using generalized ensemble simulations and Markov state models to identify conformational states. *Methods* **2009**, *49*, 197–201.
104. Bowman, G.R.; Ensign, D.L.; Pande, V.S. Enhanced modeling via network theory: Adaptive sampling of markov state models. *J. Chem. Theor. Comput.* **2010**, *6*, 787–794.
105. Bowman, G.R.; Pande, V.S. Protein folded states are kinetic hubs. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 10890–10895.
106. Voelz, V.A.; Bowman, G.R.; Beauchamp, K.; Pande, V.S. Molecular Simulation of ab Initio Protein Folding for a Millisecond Folder NTL9(1-39). *J. Am. Chem. Soc.* **2010**, *132*, 1526–1528.
107. Voelz, V.A.; Singh, V.R.; Wedemeyer, W.J.; Lapidus, L.J.; Pande, V.S. Unfolded-state dynamics and structure of protein L characterized by simulation and experiment. *J. Am. Chem. Soc.* **2010**, *132*, 4702–4709.
108. Schaeffer, R.D.; Fersht, A.; Daggett, V. Combining experiment and simulation in protein folding: closing the gap for small model systems. *Curr. Opin. Struct. Biol.* **2008**, *18*, 4–9.
109. Daggett, V. Protein folding-simulation. *Chem. Rev.* **2006**, *106*, 1898–1916.
110. Chen, H.F.; Luo, R. Binding induced folding in p53-MDM2 complex. *J. Am. Chem. Soc.* **2007**, *129*, 2930–2937.
111. Verkhivker, G.M.; Bouzida, D.; Gehlhaar, D.K.; Rejto, P.A.; Freer, S.T.; Rose, P.W. Simulating disorder-order transitions in molecular recognition of unstructured proteins: Where folding meets binding. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5148–5153.
112. Joerger, A.C.; Fersht, A.R. Structure-function-rescue: the diverse nature of common p53 cancer mutants. *Oncogene* **2007**, *26*, 2226–2242.
113. Lane, D.P.; Hupp, T.R. Drug discovery and p53. *Drug. Discov. Today* **2003**, *8*, 347–355.
114. Wells, M.; Tidow, H.; Rutherford, T.J.; Markwick, P.; Jensen, M.R.; Mylonas, E.; Svergun, D.I.; Blackledge, M.; Fersht, A.R. Structure of tumor suppressor p53 and its intrinsically disordered N-terminal transactivation domain. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 5762–5767.
115. Demarest, S.J.; Martinez-Yamout, M.; Chung, J.; Chen, H.W.; Xu, W.; Dyson, H.J.; Evans, R.M.; Wright, P.E. Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **2002**, *415*, 549–553.
116. Bertocini, C.W.; Rasia, R.M.; Lamberto, G.R.; Binolfi, A.; Zweckstetter, M.; Griesinger, C.; Fernandez, C.O. Structural characterization of the intrinsically unfolded protein beta-synuclein, a natural negative regulator of alpha-synuclein aggregation. *J. Mol. Biol.* **2007**, *372*, 708–722.
117. Sung, Y.H.; Eliezer, D. Residual structure, backbone dynamics, and interactions within the synuclein family. *J. Mol. Biol.* **2007**, *372*, 689–707.
118. Sivakolundu, S.G.; Bashford, D.; Kriwacki, R.W. Disordered p27(Kip1) exhibits intrinsic structure resembling the Cdk2/cyclin A-bound conformation. *J. Mol. Biol.* **2005**, *353*, 1118–1128.
119. Yoon, M.K.; Venkatachalam, V.; Huang, A.; Choi, B.S.; Stultz, C.M.; Chou, J.J. Residual structure within the disordered C-terminal segment of p21(Waf1/Cip1/Sdi1) and its implications for molecular recognition. *Protein Sci.* **2009**, *18*, 337–347.
120. Tran, H.T.; Mao, A.; Pappu, R.V. Role of backbone-solvent interactions in determining conformational equilibria of intrinsically disordered proteins. *J. Am. Chem. Soc.* **2008**, *130*, 7380–7392.

121. Mao, A.H.; Crick, S.L.; Vitalis, A.; Chicoine, C.L.; Pappu, R.V. Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 8183–8188.
122. Marsh, J.A.; Forman-Kay, J.D. Sequence determinants of compaction in intrinsically disordered proteins. *Biophys. J.* **2010**, *98*, 2383–2390.
123. Vuzman, D.; Azia, A.; Levy, Y. Searching DNA via a "Monkey Bar" mechanism: The Significance of disordered tails. *J. Mol. Biol.* **2010**, *396*, 674–684.
124. Sugase, K.; Dyson, H.J.; Wright, P.E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **2007**, *447*, 1021–1025.
125. Parker, D.; Ferreri, K.; Nakajima, T.; Morte, V.J.L.; Evans, R.; Koerber, S.C.; Hoeger, C.; Montminy, M.R. Phosphorylation of CREB at Ser-133 induces complex formation with CREB-binding protein via a direct mechanism. *Mol. Cell. Biol.* **1996**, *16*, 694–703.
126. Chrivia, J.C.; Kwok, R.P.S.; Lamb, N.; Hagiwara, M.; Montminy, M.R.; Goodman, R.H. Phosphorylated creb binds specifically to the nuclear-protein Cbp. *Nature* **1993**, *365*, 855–859.
127. Radhakrishnan, I.; PerezAlvarado, G.C.; Parker, D.; Dyson, H.J.; Montminy, M.R.; Wright, P.E. Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: A model for activator: Coactivator interactions. *Cell* **1997**, *91*, 741–752.
128. Radhakrishnan, I.; Perez-Alvarado, G.C.; Dyson, H.J.; Wright, P.E. Conformational preferences in the Ser(133)-phosphorylated and non-phosphorylated forms of the kinase inducible transactivation domain of CREB. *FEBS Lett.* **1998**, *430*, 317–322.
129. Solt, I.; Simon, I.; Tompa, P.; Fuxreiter, M. Phosphorylation induced preformed structural element in KID contributes to recognition by KIX. *FEBS J.* **2005**, *272*, 391–392.
130. Parker, D.; Jhala, U.S.; Radhakrishnan, I.; Yaffe, M.B.; Reyes, C.; Shulman, A.I.; Cantley, L.C.; Wright, P.E.; Montminy, M. Analysis of an activator : coactivator complex reveals an essential role for secondary structure in transcriptional activation. *Mol. Cell* **1998**, *2*, 353–359.
131. Mandell, D.J.; Chorny, I.; Groban, E.S.; Wong, S.E.; Levine, E.; Rapp, C.S.; Jacobson, M.P. Strengths of hydrogen bonds involving phosphorylated amino acid side chains. *J. Am. Chem. Soc.* **2007**, *129*, 820–827.
132. Verkhivker, G.M.; Bouzida, D.; Gehlhaar, D.K.; Rejto, P.A.; Freer, S.T.; Rose, P.W. Simulating disorder-order transitions in molecular recognition of unstructured proteins: where folding meets binding. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5148–5153.
133. Shoemaker, B.A.; Portman, J.J.; Wolynes, P.G. Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 8868–8873.
134. Trizac, E.; Levy, Y.; Wolynes, P.G. Capillarity theory for the fly-casting mechanism. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 2746–2750.
135. Chen, H.F. Molecular dynamics simulation of phosphorylated KID post-translational modification. *PLoS ONE* **2009**, *4*, e6516.