Discriminating binding mechanisms of an intrinsically disordered protein via a multi-state coarse-grained model

Michael Knott[1] and Robert B. Best[1,2,a)]
[1]*Department of Chemistry, Cambridge University, Lensfield Road, Cambridge CB2 1EW, United Kingdom*
[2]*Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, USA*

(Received 14 February 2014; accepted 17 April 2014; published online 6 May 2014)

Many proteins undergo a conformational transition upon binding to their cognate binding partner, with intrinsically disordered proteins (IDPs) providing an extreme example in which a folding transition occurs. However, it is often not clear whether this occurs via an "induced fit" or "conformational selection" mechanism, or via some intermediate scenario. In the first case, transient encounters with the binding partner favour transitions to the bound structure before the two proteins dissociate, while in the second the bound structure must be selected from a subset of unbound structures which are in the correct state for binding, because transient encounters of the incorrect conformation with the binding partner are most likely to result in dissociation. A particularly interesting situation involves those intrinsically disordered proteins which can bind to different binding partners in different conformations. We have devised a multi-state coarse-grained simulation model which is able to capture the binding of IDPs in alternate conformations, and by applying it to the binding of nuclear coactivator binding domain (NCBD) to either ACTR or IRF-3 we are able to determine the binding mechanism. By all measures, the binding of NCBD to either binding partner appears to occur via an induced fit mechanism. Nonetheless, we also show how a scenario closer to conformational selection could arise by choosing an alternative non-binding structure for NCBD. © *2014 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4873710]

I. INTRODUCTION

Protein-protein interactions perform many key functions, ranging from cell adhesion to signal transduction. In many cases, the bound conformation of one, or possibly both, of the binding partners differs from the predominant conformations populated at equilibrium, creating a challenge for methods predicting bound structures and binding equilibria.[1] A related and classic problem is to discern the mechanism by which the binding takes place. Two extreme scenarios have been contemplated.[2–10] In one, commonly known as conformational selection,[5] the bound conformation must be adopted prior to the encounter with the binding partner in order for binding to occur. The other extreme viewpoint is known as induced fit, in which formation of the correct bound conformation is facilitated by the presence of the binding partner, as originally proposed by Koshland.[11] Such a mechanism has been inferred for the recognition and proof-reading of tRNA by the ribosome[12–14] and for the open/closed transition of polymerases.[15] In the case of intrinsically disordered proteins (IDPs), which often adopt specific structures upon binding to other proteins or nucleic acids,[16,17] one may ask whether the bound conformation of the IDP must be adopted prior to binding, or must be induced by the binding partner.[18–20] This question is most relevant when the bound conformation is not already the dominant population in the unbound state. Of particular interest are those IDPs that are able to interact with multiple binding partners,[21] sometimes even adopting a different folded structure in the context of a different binding partner.[20,21]

The distinction between these two binding mechanisms may seem somewhat ill-defined: on the one hand, the bound state must always exist with some finite probability even in the absence of the partner, while on the other hand, the presence of the binding partner would be expected to favour a transition to the bound conformation, if binding is favourable. However, the distinction can be made more precise by thinking about what happens during typical binding events. We sketch out the different scenarios in Fig. 1. In the absence of the partner (Fig. 1(a)), the protein flips stochastically between (in this case) two alternative conformations, only one of which has a high affinity for the binding partner. In the conformational selection scenario (Fig. 1(b)), the binding partner only binds after a transition to the bound conformation has taken place; if the protein encounters its partner while not in its bound conformation, dissociation is more likely to occur than a transition to the bound conformation. For induced fit (Fig. 1(c)), association with the binding partner accelerates transitions to the bound conformation. If the transition between bound and unbound conformations is highly cooperative, and it is hard for the non-binding conformation to make favourable contacts with the binding partner, then a conformational selection mechanism would be favoured. The opposite scenario, of rapid transitions between non-binding and binding conformations together with the ability to form binding contacts even in the unbound state, would tend to favour induced fit.[22–24]

a)Electronic mail: robertbe@helix.nih.gov

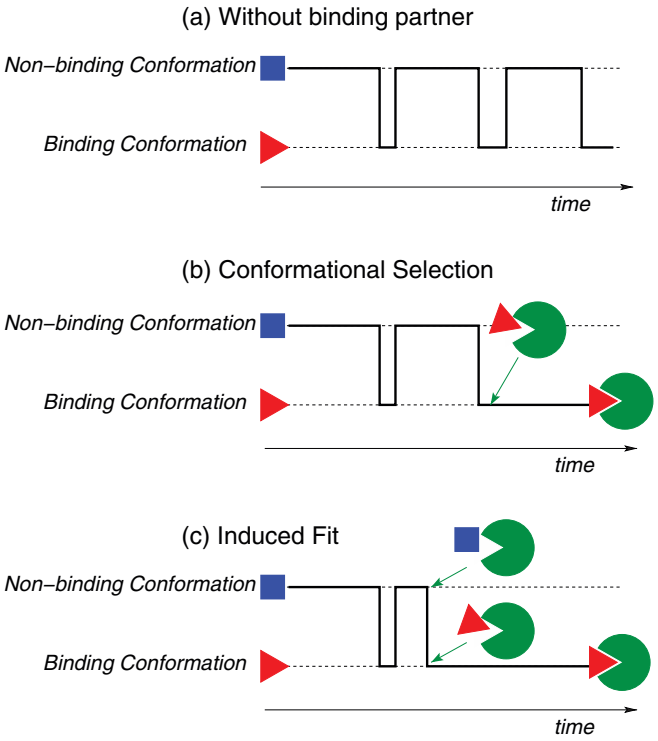

FIG. 1. Illustration of alternative binding mechanisms. (a) A schematic illustration of an intrinsically disordered protein with two alternative stable conformations within the unbound state, corresponding to states which can (red triangle) and cannot (blue square) bind to a given binding partner (green pac-man), which may or may not itself be disordered. The thick black curve represents a bistable flipping between these two conformations over time. (b) The effect of the binding partner when the binding mechanism is conformational selection (no reduction in average first passage time from non-binding to binding). (c) Situation when the binding mechanism is induced fit (reduction in first passage time for conformational change).

Experiments have yielded valuable results on many aspects of IDP structure and function. The most powerful methods are single molecule FRET experiments, which can provide information on the overall dimensions[25] and dynamics[26] of unbound IDPs, and NMR experiments which give more localised and higher resolution structural data[27,28] and also mechanistic insights.[29] Kinetic experiments can yield information on binding mechanism,[30–35] in particular when used in conjunction with a protein engineering approach. In general, however, it is challenging to resolve mechanistic scenarios from experiment alone. How can these two scenarios be distinguished in molecular simulation? One possibility is to use the binding transition state.[18,19] If this resembles the unbound state, it is clear evidence in favour of induced fit, while if the structure is almost exactly the bound one, it suggests conformational selection; intermediate scenarios are also possible. Here, we propose an alternative method, based on first passage times for conformational transitions in a multi-state coarse-grained model, as described below.

The nuclear coactivator binding domain (NCBD) of CREB-binding protein (CBP) is an intrinsically disordered protein (IDP) in the absence of any binding partner,[36] although there is, nonetheless, significant secondary structure formation.[37,38] Like many IDPs, it does adopt a structure on binding to its cognate partners, and of particular interest is

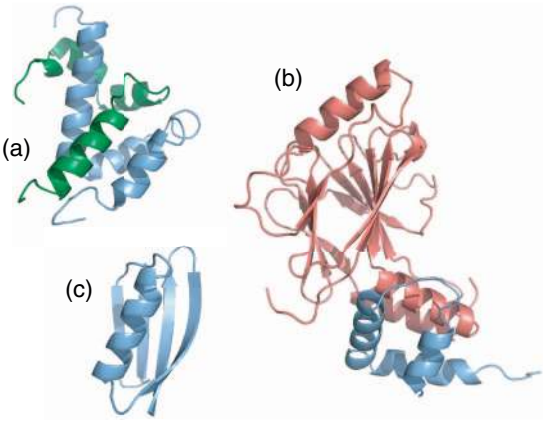

FIG. 2. Structures of IDPs and binding partners considered. (a) NCBD (blue) in complex with ACTR (green), (b) NCBD (blue) in complex with IRF3 (red), (c) protein G, used as a putative alternate conformation.

the fact that it forms two different folded conformations when bound to two different binding partners: the ACTR domain of p160 (henceforth ACTR)[36] and the interferon regulatory factor IRF-3 (henceforth IRF3),[39] for which the structures of the bound complexes are shown in Figs. 2(a) and 2(b), respectively. While other IDPs may have short recognition sequences which can adopt different structures upon binding to different partners, what is noteworthy in the case of NCBD is the size of the domain which makes the switch. NMR experiments have shown that the predominant form of NCBD in solution resembles the ACTR-bound form,[37] but that an excited state – possibly corresponding to a structure resembling the IRF3-bound form – is in exchange with the major conformer on a millisecond time scale.[40] We have previously[38] used large scale atomistic simulations to investigate structure formation in the unbound state of NCBD, and others have used a multi-scale approach to study unbound NCBD.[41] Here, we study the binding mechanism of NCBD to its partners; this complements recent experimental studies, focussed on the association of NCBD to ACTR,[33–35] as we discuss below. Although ACTR also undergoes a disorder to order transition upon binding, we have chosen to restrict our attention here to the mechanism as seen from the point of view of NCBD. Since all-atom methods would be a challenging starting point for investigation of the binding mechanism (although they may be useful at a later stage to provide more detail),[42,43] here we adopt a coarse-grained model.

Because standard structure-based models of protein folding and binding cannot capture multiple folded states, we have developed a novel multi-state structure-based model to investigate the binding mechanism of NCBD to ACTR and to IRF3. This allows the population of both bound conformers with a single energy function. Note that the unbound state is still disordered, although there are two basins of attraction within it, consistent with experiment.[40] We use this model to investigate the mechanism of binding to each binding partner. By studying the first passage times for crossing to the bound energy surface in the presence and absence of the binding partner, we are able to distinguish induced fit from conformational selection mechanisms. We find that NCBD binds to both ACTR and IRF3 by a mechanism very much like induced fit, with a

strong reduction in time taken to cross to the bound conformation in the presence of the binding partner. Using a fictitious model where the alternate folded structure of NCBD is taken to be that of protein G, we show that it is also possible to obtain a mechanism more resembling conformational selection. Analysis of binding transitions confirms the picture provided by the first passage times.

II. MODEL AND METHODS

Simulating coupled folding and binding events with atomistic simulations[44] would be a very computationally demanding task,[43] unless one had some prior insight into the mechanism, allowing for the sampling to be accelerated. We therefore adopt a coarse-graining scheme, in which the energy landscape for binding of each pair of proteins is given by a structure-based (Gō) model.[45] This is a coarse-grained model in which each residue is represented by a bead centred on the Cα carbon, and the short-range pair potential is constructed in such a way that interactions between a pair of residues can only be favourable for native contacts (i.e., when that same pair is in contact in the native state). Such a model is motivated by the hypothesis that evolution has designed sequences in such a way that competition from non-native contacts is minimised; this concept is known as the "principle of minimal frustration."[46] Gō models have been successful[47] in describing a range of phenomena, from protein folding mechanisms[48] to protein association,[49] domain swapping,[50] protein misfolding,[51] and association of IDPs with their binding partners.[18,19,52,53]

However, these models each describe only one folded (and one unfolded) free energy minimum (excepting native-like intermediates), and would therefore not be appropriate for describing the association of an IDP with two different binding partners where it binds them in very different structures. To address this shortcoming, we have constructed a double-Gō variant which combines Gō potentials derived from different structures into a hybrid system whose partition function at a defined mixing temperature T_{mix} is equal to the sum of the partition functions of the separate Gō potentials at T_{mix}.[54,55] Similar schemes for combining more than one structure-based model have been proposed based on quantum mechanical mixing[56,57] or by combining contacts from different models in a pairwise fashion.[58,59] The advantage of the present formulation is that it does not sacrifice the cooperative formation of folded structures, and allows an arbitrary number of different conformations to be included easily. Multi-Gō models of various flavours have been successfully applied to the study of problems such as conformational transitions in adenylate kinase,[56,60–62] calmodulin[55,58] and glutamine-binding protein,[57] base-flipping in B-DNA,[63] activation of src-kinase,[64] conformational exchange between different protein folds,[59,65] structural transitions in motor proteins,[66] and protein binding mechanisms.[67]

A. Double (or multi) Gō model

The double Gō method describes the system by two different energy functions E_1 and E_2. These energy functions are defined for every possible conformation of the system, but each is centred on a different conformation, $\mathbf{R}_1$ or $\mathbf{R}_2$, in which it is a minimum. The total energy E is calculated by exponential mixing:[54]

$$e^{-\beta_{\mathrm{mix}}E} = e^{-\beta_{\mathrm{mix}}E_1} + e^{-\beta_{\mathrm{mix}}E_2}, \tag{1}$$

where E_1 and E_2 are the two single-Gō energy surfaces, and $\beta_{\mathrm{mix}} = 1/k_{\mathrm{B}}T_{\mathrm{mix}}$ is an inverse mixing temperature, which is unrelated to the simulation temperature. (This can be extended to more than two energy functions simply by adding more exponential terms to the sum.) The effect of the exponential mixing is that, in a given conformation, the energy function that produces the lower energy will dominate the dynamics. In particular, if the system is near to the conformation on which one of the energy functions is centred, that function will tend to dominate. The system may be able to make transitions between the surfaces E_1 and E_2 when they are close enough.

The mixing temperature T_{mix} is the parameter that governs the extent to which the lower energy will dominate, in any given conformation. In our simulations, we used $k_{\mathrm{B}}T_{\mathrm{mix}} = 10$ kcal mol^{-1} for 1ZOQ to 1KBH in the presence of ACTR, $k_{\mathrm{B}}T_{\mathrm{mix}} = 12$ kcal mol^{-1} for 1KBH to 1ZOQ in the presence of IRF3, and $k_{\mathrm{B}}T_{\mathrm{mix}} = 16$ kcal mol^{-1} for 1PGB to 1KBH in the presence of ACTR. We can consider the relative extent to which E will change in response to a change in E_1 or E_2:

$$\frac{\partial E/\partial E_2}{\partial E/\partial E_1} = \exp\left(-\frac{\Delta E}{k_B T_{\mathrm{mix}}}\right), \tag{2}$$

where $\Delta E = E_2 - E_1$. When $k_B T_{\mathrm{mix}} \ll |\Delta E|$, the total energy is overwhelmingly governed by the lower of the contributing energies, while in the limit $k_B T_{\mathrm{mix}} \gg |\Delta E|$, both would contribute equally. In making a transition, the largest energy gap that the system can readily jump is of the order of $k_B T_{\mathrm{mix}}$.

A good choice of mixing temperature is essential to the double Gō model. It should be low enough so that when the protein is in the vicinity of one of the native structures, the energy surface derived from that structure dominates the average, i.e., when the instantaneous structure at time t, $\mathbf{R}(t)$, is near folded structure 1, the energy $E(\mathbf{R}(t))$ will be close to $E_1(\mathbf{R}(t))$ and near structure 2, close to $E_2(\mathbf{R}(t))$). Only in the more unfolded regions of configuration space will significant mixing between the two surfaces occur. In this way, the combined potential remains faithful to the original surfaces in the vicinity of their respective native states, where they are expected to be most accurate. If T_{mix} is too high, we will instead be simulating an unphysical hybrid surface in which, for example, both energy functions contribute to the average even when close to one of the native states. However, the mixing temperature must also be high enough to allow mixing when far from the native states, so that transitions between the surfaces can be observed on a reasonable time scale. The ability to find a value of T_{mix} that satisfies both requirements for a given pair of reference configurations depends on a physically reasonable choice for the individual energy functions E_1 and E_2 (see below).

B. Single Gō model

Each state of NCBD was based on a Gō type model derived from the experimental structure in complex with a binding partner: PDB entries 1KBH[36] and 1ZOQ[39] for NCBD:ACTR and NCBD:IRF3, respectively. An alternative non-binding state was also modeled on the folded structure of protein G, PDB entry 1PGB,[88] as described below. The total energy function is a simple sum over harmonic bonds and pairs of contacts. All angle and dihedral angle terms were set to zero; instead, short range contacts were allowed as native contacts (all contacts i, j with $|i-j| \geq 3$ are permitted).

Each energy function (E_1 or E_2) can be written as a sum of three terms:

$$E_{1,2} = \sum_{j=i+1} V_{ij}^{(B)}(r_{ij}) + \sum_{j>i} V_{ij}^{(R)}(r_{ij}) + \sum_{i,j\in N} V_{ij}^{(N)}(r_{ij}), \quad (3)$$

where r_{ij} is the distance between residues i and j. The purpose of the first term is to maintain the bonds between neighbouring residues

$$V_{ij}^{(B)}(r_{ij}) = \kappa^{(B)}\left(r_{ij} - r_{ij}^{(0)}\right)^2, \quad (4)$$

where $\kappa^{(B)} = 378$ kcal mol^{-1} Å^{-2} is a spring constant and $r_{ij}^{(0)}$ is the distance between residues i and j in the reference conformation ($\mathbf{R}_1$ or $\mathbf{R}_2$).

The second term $V_{ij}^{(R)}$ provides residues with repulsive cores that repel one another. Many Gō models use repulsive cores whose range depends on the native conformation. While this helps to increase the cooperativity of the model, it becomes impracticable in a double Gō model: both E_1 and E_2 need to have reasonable values for all accessible conformations, to avoid the situation where a conformation is accessible according to one energy function, but largely inaccessible according to the other because of a very high energy. This means that the repulsion should have the same range for both surfaces. This is also a more physically reasonable picture, since the closest approach of two residues should be independent of the native structure, while the attractive terms in the Gō model may be more long ranged. To achieve this, we separate the repulsive part of the Karanicolas and Brooks 12-10-6 potential[68] at the minimum, in a manner similar to the Weeks-Chandler-Anderson (WCA) approach for simple liquids:[69]

$$\frac{V_{ij}^{(R)}(r_{ij})}{\epsilon_{ij}^{(R)}} = \begin{cases} 1 + V^{\mathrm{KB}}(r_{ij}), & r_{ij} \leq \sigma_{ij}^{(R)} \\ 0, & r_{ij} \geq \sigma_{ij}^{(R)} \end{cases}, \quad (5)$$

where $\epsilon_{ij}^{(R)} = 1.0$ kcalmol^{-1} is an energetic parameter governing the strength of the repulsion, and $\sigma_{ij}^{(R)}$ is the range of the repulsion. Although the range must be the same for every energy surface, in principle it can be different for different residues. We use $\sigma_{ij}^{(R)} = 4$ Å for all residues. The Karanicolas-Brooks potential $V^{\mathrm{KB}}(r_{ij})$ is given by

$$V^{\mathrm{KB}}(r_{ij}) = 13\left(\frac{\sigma_{ij}^{(R)}}{r_{ij}}\right)^{12} - 18\left(\frac{\sigma_{ij}^{(R)}}{r_{ij}}\right)^{10} + 4\left(\frac{\sigma_{ij}^{(R)}}{r_{ij}}\right)^{6}. \quad (6)$$

In addition to the repulsion, we need an interaction between pairs of residues that are regarded as making contacts in the native conformation. We treat a pair as forming a contact if two non-hydrogen atoms, one from each residue, are found within 5 Å of one another in the native conformation. We include in the energy function E_1 the *intermolecular* interactions from structure $\mathbf{R}_2$, and vice versa. In doing so, we make the physically reasonable assumption that the IDP is not required to adopt the correct bound structure before it can form attractive pairwise interactions with the binding partner (assuming otherwise would essentially presuppose a conformational selection mechanism). The energy of interaction between pairs is

$$\frac{V_{ij}^{(N)}}{\epsilon_{ij}^{(N)}} = -\exp\left(-\frac{1}{2}\left(\frac{r_{ij} - r_{ij}^{(0)}}{s_{ij}^{(N)}}\right)^2\right). \quad (7)$$

Here, $s_{ij}^{(N)} = 0.5$ Å is a parameter governing the width of the attractive well, while $\epsilon_{ij}^{(N)}$ governs its depth. This parameter is set to 1.0 kcalmol^{-1} for all contacts between NCBD and a binding partner (ACTR or IRF3), and also for all intraprotein contacts within the binding partner. For intraprotein contacts within NCBD, it is set to 1.0 kcalmol^{-1} for the 1KBH model, but is scaled in the other models in order to keep the stability of the reference state of an isolated NCBD molecule approximately equal to that in 1KBH: to 1.183 and 1.215 kcal mol^{-1} in 1ZOQ (when binding to ACTR and IRF3, respectively) and to 0.910 kcal mol^{-1} in 1PGB.

1. Q value

The fraction of native contacts, Q, an indicator of the extent to which native structure is present in a given conformation, is defined by

$$Q = \frac{1}{N}\sum_{ij} \exp\left(-\frac{1}{2}\left(\frac{r_{ij} - r_{ij}^{(0)}}{s_{ij}^{(N)}}\right)^2\right), \quad (8)$$

where $s_{ij}^{(N)}$ is set to 0.5 Å for all contacts, and the sum is taken over the subset of N contacts (intermolecular or intramolecular or both) whose residues i and j are separated by at least five positions along the chain.

C. Binding simulations

All simulations were performed in CHARMM,[70,71] modified to include the double-Gō model. First passage binding simulations were run using Langevin dynamics with a time step of 0.01 ps and a friction coefficient of 0.1 ps^{-1}. An initial unbound ensemble was generated by running simulations with the interprotein interactions all set to be repulsive (i.e., third term in Eq. (3) is set to zero), but with all other details of the model remaining the same. Configurations from these simulations (in which the protein was on the incorrect energy surface for binding), were selected at intervals of 100 ns and used as the starting point for binding simulations. A transition between energy surfaces was considered to have occurred once the energy according to the target surface was lower than that according to the starting surface in two consecutive

snapshots separated by 0.1 ns. A total of 100 first passage time runs was conducted for each binding scenario.

III. RESULTS AND DISCUSSION

We have constructed a hybrid potential based on the bound structures of NCBD with ACTR (PDB reference: 1KBH[36]) and with IRF3 (PDB reference: 1ZOQ[39]). This allows NCBD to bind to these two binding partners in the correct native conformations, but still to populate an unbound state which is disordered. Langevin dynamics simulations in the hybrid potential allow the protein to flip between the two conformations. This can be monitored by calculating the intramolecular contact energy of NCBD, shown in Fig. 3(a), where three transitions between IRF3-bound and ACTR-bound energy surfaces are shown. Our assumption of only two stable attractors within the unbound state, based on the two bound structures, may appear an oversimplification. However, it is consistent with recent NMR dynamics experiments on unbound NCBD, which found a millisecond-time scale two-state conformational exchange between the dominant ACTR-bound conformer and an alternate, high energy state which may correspond to the IRF-3 bound form.[40] Note also, that although the unbound structures may have some similar features to the bound states on which they are modeled, their tertiary structure is still substantially disordered (as expected for the unfolded state of any Gō-like model). The experimental evidence suggests that the state similar to the ACTR-bound structure is somewhat more favoured in isolated NCBD,[37,40] with an equilibrium population of ∼90%. In our model, the state similar to the ACTR-bound structure is still the most favoured, although with slightly reduced population (∼75%). The transition time between conformations in such a model is clearly a critical determinant of the type of mechanism.[22–24] In our model, we can control the rate of transitions between the two conformations in the unbound state by varying the mixing temperature T_{mix}. We have chosen a mixing temperature in our model which approximately reproduces the time scale of transitions between unbound conformations of NCBD: in our model, the transition times are of the order ∼0.1 μs. Because we have used a low Langevin friction coefficient of 0.1 ps^{-1} (relative to ∼50 ps^{-1} for water) to accelerate dynamics, and because of the smoother energy landscape for a coarse-grained model, the simulation times are shorter than experiment by a factor of 100–1000.[19,72] Thus, the model transition time scale is roughly consistent with the experimental relaxation time of the order of ∼0.1 ms.

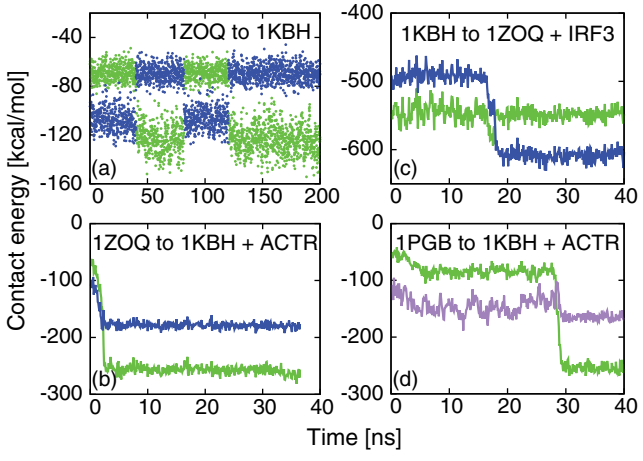

FIG. 3. Double Gō model for NCBD. (a) Energies of NCBD in the absence of binding partner, (b) in the presence of the intrinsically disordered partner ACTR and (c) in the presence of the folded binding partner IRF3. In each case the energies for the models based on the structure in complex with ACTR and with IRF3 are given in green and blue, respectively. (d) Alternative model in which the unbound state of NCBD is artificially represented by protein G (protein G energy in purple). Note that the offsets in energy between graphs are the result of differing internal energies of the binding partner (if present).

A. Transition times

Having established that the system can flip between these energy surfaces at equilibrium in the absence of a binding partner, we next conducted first-passage binding simulations, in which the system was initialised from the incorrect conformation for binding for a given partner, and simulations were run with a spherical boundary potential until a binding event occurred. The radius of the boundary sphere was 62.035 Å, which gives the same volume as a cubic box of side length 100 Å, and leads to protein concentrations of 1.66 mM. Representative examples of binding events are shown in Figs. 3(b) and 3(c) for NCBD binding to ACTR and IRF3, respectively. Due to the additional stabilising influence of the binding partner, the bound structure is much lower in energy than unbound conformations, and so unbinding transitions are not observed on the simulation time scale shown.

By running a large number of binding simulations, we can characterise the binding mechanism in each case. While there are many ways to describe binding mechanism, we have exploited the ability of our model to discriminate between alternative conformations based on their energy. We determine the first passage time for the system to cross to the bound energy surface relative to the time taken in the absence of the binding partner. Clearly, for induced fit binding, the time taken to cross to the bound surface should be reduced by the presence of the binding partner, while in an extreme conformational selection scenario, no difference in first passage time would be observed.

We show the results of this analysis in Figs. 4(a) and 4(b) for binding to ACTR and IRF3 respectively, by plotting the proportion of simulations remaining on the original energy surface at each time point. In both cases, the presence of the second protein greatly reduces the time to cross to the bound energy surface, indicating an induced-fit type of mechanism, as has been inferred for the binding of several other IDPs.[18,19,52] An induced fit mechanism is favoured by a rapidly interconverting unfolded state.[41] An induced-fit mechanism is interesting in the context of a two-state exchange in the unfolded state, which by itself would be highly suggestive of conformational selection. Our results show that an induced fit mechanism is still possible in this situation, when the binding partner is able to accelerate the transition to the binding conformation.

In order to illustrate that our model can accommodate a range of scenarios, we have constructed an alternative double-Gō model in which the alternate conformation is not one of

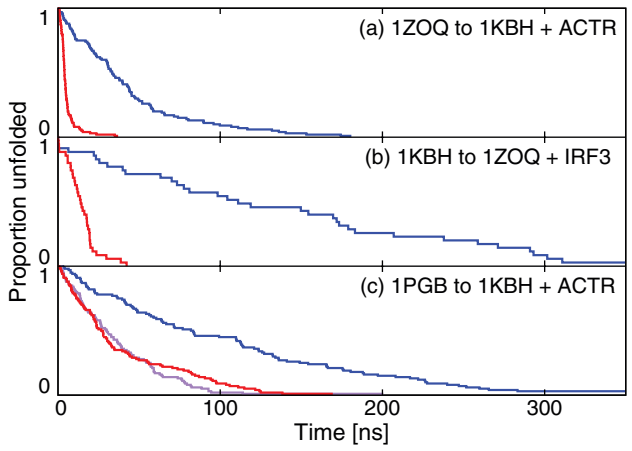

FIG. 4. Statistics of first passage times for binding. (a) NCBD binding to ACTR, (b) NCBD binding to IRF3, (c) NCBD-G binding to ACTR (here protein G rather than the IRF3-bound structure is used as the alternate conformation). Blue curves give the survival probability for the unbound conformation in the absence of a binding partner and red curves give the survival probability in the presence of the binding partner. In (c), the purple curve represents the survival probability for the NCBD-G model in which inter- and intramolecular contacts have been scaled equally in the construction of the model.

the bound states of NCBD, but rather the folded state of protein G (PDB reference: 1PGB).[73] While the two bound states of NCBD each consist of three helices with similar helical boundaries – and so may be argued to differ mainly in tertiary packing – protein G has an α/β structure bearing little resemblance to NCBD. By using this alternate conformation, we construct a hypothetical model of NCBD in which the secondary structure of the bound conformation is no longer significantly populated under unbound conditions. In this case a scenario closer to conformational selection would be anticipated, because of the larger rearrangement required to interconvert the two conformations and the slower associated time scale. In Fig. 3(d) we show a representative binding trajectory of the chimeric protein (hereafter referred to as NCBD-G), and in Fig. 4(c) we show the statistics of first passage times for NCBD-G crossing from the protein G energy surface to the ACTR-bound one, in the presence and absence of ACTR. The behaviour of this system clearly resembles conformational selection much more closely, although the ACTR does still affect the rate of energy surface crossing, indicating a small induced-fit effect.

Our analysis of first passage times provides a novel, and unequivocal method for analyzing binding mechanism, which is complementary to existing flux-based approaches. Nonetheless, we can simply relate our results to flux, if we assume that all steps are irreversible, which is valid if the reverse processes are highly improbable on the time scale of interest. In this case, the relative flux via an "induced-fit" pathway may be computed as $\phi_{\mathrm{IF}} = 1 - \tau^{+}_{\mathrm{MFPT}}/\tau^{-}_{\mathrm{MFPT}}$, where τ^{+}_{MFPT} and τ^{-}_{MFPT} are the first passage times for conformational change in NCBD in the presence and absence of the ligand, respectively. Using this approximation, we obtain relative induced fit fluxes ϕ_{IF} of 0.87 and 0.89 for binding of NCBD to ACTR and IRF3, and 0.62 for binding of NCBD-G to ACTR, confirming our earlier intepretation.

Because the relative strengths of intra- and interprotein interactions may affect the binding mechanism, we have investigated this effect in the context of the NCBD-G model. In addition to separately scaling the intramolecular contacts in order to achieve a given stability, as in the model described above, we also adopted a more conservative approach. In this alternative approach, the intermolecular contacts were scaled by the same amount as the intramolecular ones, corresponding to the assumption that inter- and intramolecular interactions should be of similar strength. Remarkably, we find that this has little effect on the binding mechanism of NCBD-G (Fig. 4(c)), with a ϕ_{IF} of 0.65 relative to 0.62 for the original model, suggesting that the mechanism is robust to small variations in scaling of intermolecular interactions.

B. Binding transition paths

Our analysis of first passage times provides quantitative evidence for the type of binding mechanism which is adopted in each case. In order to obtain more structural insight into the binding mechanism of NCBD to ACTR and IRF3, we have analysed binding transition paths in more detail. Transition paths are those segments of trajectories during which the binding and folding take place, i.e., excluding the time spent exploring unbound configurations prior to binding, or non-productive transient encounters. As is the case with protein folding, all of the mechanistic information about binding can be obtained from these transition paths.[74]

Examples of complete binding transition paths are shown in Fig. 5(a). In each case, the fractions of intermolecular contacts, and the fractions of NCBD intramolecular contacts, are shown in blue and red, respectively, each running from zero at the start of the transition to near unity at the end. Already from this single example, it is apparent that substantial formation

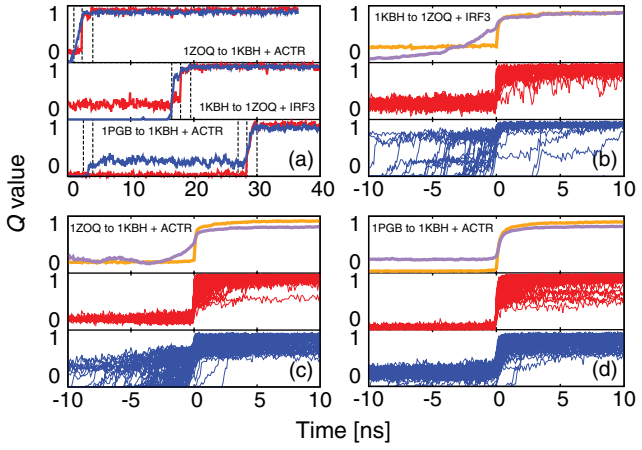

FIG. 5. Binding transition paths. (a) Examples of single binding events for each system, showing the fraction of intermolecular contacts (blue) and intramolecular contacts within NCBD (red). Broken vertical lines indicate positions of the snapshots given in Fig. 6. (b) Binding of NCBD to IRF3. Top panel: Average intermolecular Q (purple) and average intramolecular NCBD contacts (orange); middle panel: individual intramolecular Q trajectories; lower panel: individual intermolecular Q trajectories. For averaging, trajectories were aligned at the point at which the 1KBH and 1ZOQ energy surfaces cross. (c) Binding of NCBD to ACTR, legend as in (b). (d) Binding of NCBD-G to ACTR.

of intermolecular contacts can occur *before* the intramolecular structural transition takes place. We also show a transition path for the binding of the chimeric NCBD-G to ACTR in Fig. 5(a). In that case, there is clearly a partially bound intermediate which remains fully in the protein G conformation.

We have averaged the results of many transition paths similar to those in Fig. 5(a). In order to synchronise the times in each path, we define $t = 0$ to be the time at which the transition between the global E_1 and E_2 energy surfaces takes place. The results of these averages are shown in Figs. 5(b)–5(d), together with the overlaid results of all trajectories. These support the conclusion from the individual examples, that the formation of intermolecular contacts precedes any structural transition of NCBD when binding to either IRF3 or ACTR. This is reminiscent of the "fly-casting" mechanism that has been proposed to increase the rate of protein association for IDPs, in which the capture radius for binding is effectively increased by the ability of the IDP to form transient interactions with the binding partner while still unfolded.[75] An earlier simulation study using a single-Gō model, with carefully calibrated secondary structure, similarly predicted an initial formation of inter-protein contacts prior to folding of the NCBD, when binding to ACTR, consistent with our results.[53] The results we obtain with our double-Gō model are completely consistent with the conclusions of our earlier all-atom study of the isolated NCBD.[38] In that work, we found that although the unbound state was still significantly disordered, there was evidence that some features of the binding interface were already present. The presence of such an interface would help to enhance the efficiency of the induced-fit mechanism that we find here.

By contrast, the formation of bound and most intermolecular contacts occurs simultaneously for the chimeric NCBD-G when binding to ACTR. While the binding of NCBD to either target molecule clearly occurs via induced fit, the situation for the binding of NCBD-G to ACTR is somewhat less clear cut. Clearly, the final transition from the non-natively bound intermediate to the bound state occurs via something very close to ideal conformational selection. However, it is also evident that the binding must slightly alter the energy landscape of NCBD-G in such a way as to make the transition to the bound energy surface more favourable. An idealised conformational selection mechanism would require that it be more likely for the proteins to dissociate before any transition to the bound conformation could occur.[24] This would have presumably been the case if there were a greater degree of "frustration"[76] between the NCBD-like contacts to ACTR, and the contacts stabilizing protein G.

We can also follow individual transition events in order to zoom in on the structural changes occurring as the two proteins bind. In Fig. 6, we show representative snapshots drawn from specific points along the binding transition paths (indicated by broken vertical lines in Fig. 5(a)). These illustrate very nicely the features of each mechanism. In the binding of NCBD to ACTR, the two proteins initially associate in a disordered state (since ACTR is also an IDP), Fig. 6(a). At the point at which the energy surfaces cross, Fig. 6(b), substantial native-like intermolecular interactions have been formed, with the fully native-like structure being adopted only later,

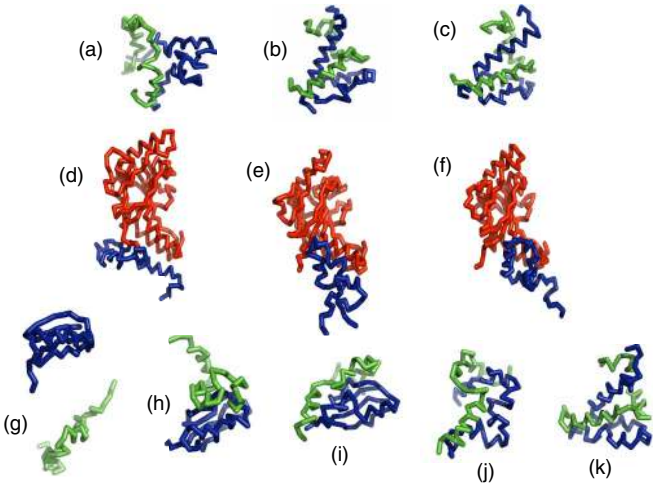

FIG. 6. Snapshots extracted from binding transitions. (a)–(c) Snapshots on the transition paths for binding of NCBD to ACTR in Fig. 5 (a); (d)–(f) analogous snapshots for binding of NCBD to IRF3; (g)-(k) snapshots for binding of NCBD-G to ACTR.

Fig. 6(c). A similar sequence of events is observed in the binding of NCBD to the folded IRF3. In this case, however, the NCBD is even more disordered at the point where the energy surfaces cross, Fig. 6(e)—the higher degree of disorder may be possible because the binding partner is already folded in this case.

Finally, a somewhat different sequence of events occurs in the binding of NCBD-G to ACTR: the NCBD initially binds to the ACTR in the folded protein G structure (Figs. 6(g) and 6(h)), and remains in this structure, unfolding only at the point at which the energy surfaces cross (Fig. 6(j)), as in a conformational selection scenario. Note, however, that the acceleration of transition rate to the NCBD conformation in the presence of ACTR indicates that the barrier for transition is lowered in the bound complex, so that this mechanism also has some features of an induced-fit pathway. The intention of this example was of course to show that there are scenarios in which our model results in prototypical conformational selection. The reason that the mechanism does not completely conform to conformational selection probably relates to the fact that the binding partner in this case is a disordered protein (ACTR): thus, even though the NCBD-G may initially be in the incorrect conformation to form favourable intermolecular contacts, the ACTR can deform around it to make these contacts, reducing the probability of encounter complex dissociation and favouring transition to the bound state.

Our finding of an induced-fit mechanism for NCBD binding to ACTR is confirmed by recent experimental evidence from a protein engineering analysis of the binding transition states, which show that, apart from weak native-like interactions, the transition state is largely disordered,[33,34] with Φ-values for binding in the range 0–0.2. We have also calculated Φ-values for binding of NCBD to ACTR based on the simulations, using a simple contact based analysis, i.e., we define the Φ value of residue i as $\Phi(i) = (q_{\ddagger}(i) - q_{\mathrm{U}}(i))/(q_{\mathrm{F}}(i) - q_{\mathrm{U}}(i))$, where $q_{\ddagger}(i)$, $q_{\mathrm{U}}(i)$ and $q_{\mathrm{F}}(i)$ are the fraction of its native contacts formed by residue i in the binding transition state,

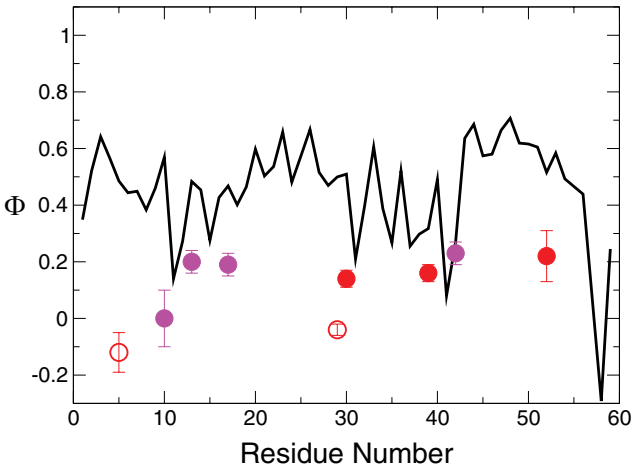

FIG. 7. Φ-values for binding of NCBD to ACTR. Black curve shows estimate from simulation, red and magenta symbols are experimental data in the absence and presence of TMAO, respectively. Negative Φ-values are given by unfilled symbols.

unfolded state and folded state, respectively. Here, we assume the binding transition state is close to where the energy surfaces cross, and use the structures closest to this point on each transition path to represent it. These simulation Φ-values are slightly higher than the experimental data (Fig. 7), but bearing in mind the approximations involved, they qualitatively agree on the partial formation of structure in the transition state.

Although NCBD-G is an example which has been somewhat artificially contrived in this work in order to illustrate a point, it does raise an intriguing possibility: could folded proteins have evolved specifically to bind to different binding partners with very different structures? In fact, it has recently been shown that it is possible to design proteins with 98% sequence identity which fold to different structures—namely, a variant of protein G (G_B98) and a three-helix bundle (G_A98).[77] While each protein has a natural ligand, the ability of the G_A98 variant to bind both ligands (human serum albumin immunoglobulin and immunoglobulin G for G_A98 and G_B98 respectively) suggests strongly that it can adopt either folded structure, depending on which ligand it is bound to.[77,78] This is consistent with the hypothesis that such bistable structures may function as evolutionary bridges between proteins carrying out different functions, supported by recent experimental and theoretical work.[79–82] In addition, several natural examples exist of proteins that populate different native states in the presence of different binding partners[83–85] and are not intrinsically disordered, indicating that this is not a property unique to IDPs. Nonetheless, it is likely to be easier for nature to design IDPs to bind in different structures, because of the larger contribution made by the binding interface to the overall stability. There is also likely to be a difference in the binding mechanism, with conformational selection likely to be more important for predominantly folded proteins which can bind in different folded states, because the barrier to transitions between these two states is likely to be larger than for IDPs. This difference in mechanism may also have functional implications: for example, induced-fit binding may be optimal for enhancing binding rates, while conformational selection may be preferred when it is necessary to damp the response to the fluctuating concentration of a ligand.

IV. CONCLUSION

We have developed a two-state coarse-grained model which is able to describe the binding of an IDP to different binding partners, even in a situation where it binds in different conformations. In the absence of a binding partner, the protein flips between two conformations in a two-state fashion, which is consistent with the conformational exchange observed in NMR relaxation-dispersion experiments. By incorporating the binding partner into the simulation, we are able to use the change in rate of flipping between these alternative conformations as a novel criterion for assessing the binding mechanism. We find that binding of NCBD to either ACTR or IRF-3 closely resembles an induced fit scenario. This is not an artefact of the model, as we show that it is possible to obtain a situation closer to conformational selection by choosing a different alternative folded structure. One limitation of our present study is the neglect of non-native contacts; while specific non-native contacts are generally not important for protein folding,[74,86] non-specific non-native interactions can help to enhance the binding rates of IDPs by increasing the chance that they remain in contact after initial encounter.[19,87] The fact that non-native contacts reduce the probability of dissociation should not alter our conclusions, however, as it would only enhance the likelihood of an induced fit mechanism. The coarse-grained models we present here may assist a more detailed investigation of the mechanism with atomistic models in two ways: (i) they may be used to guide the atomistic simulations in biased sampling, or (ii) they may be used to determine the mechanism by post-processing the trajectories using the coarse-grained energy functions to determine when the switching occurs, analogously to what has been done here.

ACKNOWLEDGMENTS

We thank Chris Baker and David de Sancho for many helpful suggestions and comments on the manuscript. M.K. and R.B.B. were supported by BBSRC Award No. BB/H006885/1. R.B.B. was supported by a Royal Society University Research Fellowship while at Cambridge and by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases of the National Institutes of Health.

[1] Y. C. Kim and G. Hummer, "Coarse-grained models for simulation of multiprotein complexes: application to ubiquitin binding," J. Mol. Biol. **375**, 1416–1433 (2008).
[2] S. Chaudhury and J. J. Gray, "Conformer selection and induced fit in flexible backbone protein-protein docking using computational and NMR ensembles," J. Mol. Biol. **381**, 1068–1087 (2008).
[3] D. D. Boehr, R. Nussinov, and P. E. Wright, "The role of dynamic conformational ensembles in biomolecular recognition," Nature Chem. Biol. **5**, 789–796 (2009).
[4] B. Ma and R. Nussinov, "Enzyme dynamics point to stepwise conformational selection in catalysis," Curr. Opin. Chem. Biol. **14**, 652–659 (2010).

[5]P. Csermely, R. Palotai, and R. Nussinov, "Induced fit, conformational selection and independent dynamic segments: an extended view of binding events," Trends. Biochem. Sci. **35**, 539–546 (2010).
[6]L. C. James, P. Roversi, and D. S. Tawfik, "Antibody multispecificity mediated by conformational diversity," Science **299**, 1362–1367 (2003).
[7]R. Jimenez, G. Salazar, K. K. Baldridge, and F. E. Romesberg, "Flexibility and molecular recognition in the immune system," Proc. Natl. Acad. Sci. U.S.A. **100**, 92–97 (2003).
[8]O. F. Lange, N.-A. Lakomek, C. Farès, G. F. Schröder, K. E. Walter, S. Becker, J. Meiler, H. Grubmuller, C. Griesinger, and B. L. de Groot, "Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution," Science **320**, 1471–1475 (2008).
[9]D. Long and R. Brüschweiler, "In silico elucidation of the recognition dynamics of ubiquitin," PLoS Comput. Biol. **7**, e1002035 (2011).
[10]D. Tobi and I. Bahar, "Structural changes involved in protein binding correlate with intrinsic motions of proteins in the unbound state," Proc. Natl. Acad. Sci. U.S.A. **102**, 18908–18913 (2005).
[11]D. E. Koshland, "Application of a theory of enzyme specificity to protein synthesis," Proc. Natl. Acad. Sci. U.S.A. **44**, 98–104 (1958).
[12]T. Pape, W. Wintermeyer, and M. W. Rodina, "Induced fit in initial selection and proofreading of aminoacyl-t-RNA on the ribosome," Eur. J. Mol. Biol. **18**, 3800–3807 (1999).
[13]J. M. Ogle, A. P. Carter, and V. Ramakrishnan, "Insights into the decoding mechanism from recent ribosome structures," Trends Biochem. Sci. **28**, 259–266 (2003).
[14]S. C. Blanchard, J. Ruben L. Gonzalez, H. D. Kim, S. Chu, and J. D. Puglisi, "tRNA dynamics on the ribosome during translation," Proc. Natl. Acad. Sci. U.S.A. **101**, 12893–12898 (2004).
[15]W. Zheng, B. R. Brooks, S. Doniach, and D. Thirumalai, "Network of dynamically important residues in the open/closed transition in polymerases is strongly conserved," Structure **13**, 565–577 (2005).
[16]R. W. Kriwacki, L. Hengst, L. Tennant, S. I. Reed, and P. E. Wright, "Structural studies of p21(Waf1/Cip1/Sdi1) in the free and Cdk2-bound state: Conformational disorder mediates binding diversity," Proc. Natl. Acad. Sci. U.S.A. **93**, 11504–11509 (1996).
[17]M. Fuxreiter, P. Tompa, I. Simon, V. N. Uversky, J. C. Hansen, and F. J. Asturias, "Malleable machines take shape in eukaryotic transcription regulation," Nat. Chem. Biol. **4**, 728–737 (2008).
[18]A. Turjanski, R. B. Best, J. S. Gutkind, and G. Hummer, "Binding-induced folding of a natively unstructured transcription factor," PLoS Comput. Biol. **4**, e1000060 (2008).
[19]D. de Sancho and R. B. Best, "Modulation of an IDP binding mechanism and rates by helix propensity and non-native interactions: association of Hif1α with CBP," Mol. Biosys. **8**, 256–267 (2012).
[20]P. E. Wright and H. J. Dyson, "Linking folding and binding," Curr. Opin. Struct. Biol. **19**, 31–38 (2009).
[21]W.-L. Hsu, C. J. Oldfield, B. Xue, J. Meng, F. Huang, P. Romero, V. N. Uversky, and A. K. Dunker, "Exploring the binding diversity of intrinsically disordered proteins involved in one-to-many binding," Protein Sci. **22**, 258–273 (2013).
[22]G. G. Hammes, Y.-C. Chand, and T. G. Oas, "Conformational selection or induced fit: a flux description of reaction mechanism," Proc. Natl. Acad. Sci. U.S.A. **106**, 13737–13741 (2009).
[23]T. Wlodarski and B. Zagrovic, "Conformational selection and induced fit mechanism underlie specificity in nonconvalent interactions with ubiquitin," Proc. Natl. Acad. Sci. U.S.A. **106**, 19346–19351 (2009).
[24]H.-X. Zhou, "From induced fit to conformational selection: a continuum of binding mechanism controlled by the timescale of conformational transitions," Biophys. J. **98**, L15–L17 (2010).
[25]H. Hofmann, A. Soranno, A. Borgia, K. Gast, D. Nettels, and B. Schuler, "Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy," Proc. Natl. Acad. Sci. U.S.A. **109**, 16155–16160 (2012).
[26]A. Soranno, B. Buchli, D. Nettels, R. R. Cheng, S. Müller-Späth, S. H. Pfeil, A. Hoffmann, E. A. Lipman, D. E. Makarov, and B. Schuler, "Quantifying internal friction in unfolded and intrinsically disordered proteins with single-molecule spectroscopy," Proc. Natl. Acad. Sci. U.S.A. **109**, 17800–17806 (2012).
[27]A. S. Maltsev, J. Ying, and A. Bax, "Impact of N-terminal acetylation of alpha-synuclein on its random coil and lipid binding properties," Biochemistry **51**, 5004–5013 (2012).
[28]M. R. Jensen, R. W. H. Ruigrok, and M. Blackledge, "Describing intrinsically disordered proteins at atomic resolution by NMR," Curr. Opin. Struct. Biol. **23**, 426–435 (2013).
[29]K. Sugase, H. J. Dyson, and P. E. Wright, "Mechanism of coupled folding and binding of an intrinsically disordered protein," Nature **447**, 1021–1027 (2007).
[30]A. Bachmann, D. Wildemann, F. Praetorius, G. Fischer, and T. Kiefhaber, "Mapping backbone and side-chain interactions in the transition state of a coupled protein folding and binding reaction," Proc. Natl. Acad. Sci. U.S.A. **108**, 3952–3957 (2011).
[31]J. M. Rogers, A. Steward, and J. Clarke, "Folding and binding of an intrinsically disordered protein: fast, but not 'diffusion limited'," J. Am. Chem. Soc. **135**, 1415–1422 (2013).
[32]S. L. Shammas, J. M. Rogers, S. A. Hill, and J. Clarke, "Slow, reversible, coupled folding and binding of the spectrin tetramerization domain," Biophys. J. **103**, 2203–2214 (2012).
[33]J. Dogan, T. Schmidt, X. Mu, A. Engström, and P. Jemth, "Fast association and slow transitions in the interaction between two intrinsically disordered protein domains," J. Biol. Chem. **287**, 34316–34324 (2012).
[34]J. Dogan, X. Mu, A. Engström, and P. Jemth, "The transition state structure for coupled binding and folding of disordered protein domains," Sci. Rep. **3**, 2076 (2013).
[35]V. Iešmantavičius, J. Dogan, P. Jemth, K. Teilum, and M. Kjaergaard, "Helical propensity in an intrinsically disordered protein accelerates ligand binding," Angew. Chem., Int. Ed. **52**, 1548–1551 (2014).
[36]S. J. Demarest, M. Martinez-Yamout, J. Chung, H. Chen, W. Xu, H. J. Dyson, R. M. Evans, and P. E. Wright, "Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators," Nature **415**, 549–553 (2002).
[37]M. Kjaergaard, K. Teilum, and F. M. Poulsen, "Conformational selection in the molten globule state of the nuclear coactivator binding domain of CBP," Proc. Natl. Acad. Sci. U.S.A. **107**, 12535–12540 (2010).
[38]M. Knott and R. B. Best, "A preformed binding interface in the unbound ensemble of an intrinsically disordered protein: evidence from molecular simulation," PLoS Comput. Biol. **8**, e1002605 (2012).
[39]B. Y. Qin, C. Liu, H. Srinath, S. S. Lam, J. J. Correia, R. Derynck, and K. Lin, "Crystal structure of IRF-3 in complex with CBP," Structure **13**, 1269–1277 (2005).
[40]M. Kjaergaard, L. Andersen, L. D. Nielsen, and K. Teilum, "A folded excited state of ligand-free nuclear coactivator binding domain (NCBD) underlies plasticity in ligand recognition," Biochemistry **52**, 1686–1693 (2013).
[41]A. N. Naganathan and M. Orozco, "The native ensemble and folding of a protein molten-globule: functional consequences of downhill folding," J. Am. Chem. Soc. **133**, 12154–12161 (2011).
[42]J. Mittal, T. H. Yoo, G. Georgiou, and T. M. Truskett, "Structural ensemble of an intrinsically disordered polypeptide," J. Phys. Chem. B **117**, 118–124 (2013).
[43]C. M. Baker and R. B. Best, "Insights into the binding of intrinsically disordered proteins from molecular dynamics simulation," WIREs Comput. Mol. Sci. **4**, 182–198 (2014).
[44]I. Staneva, Y. Huang, Z. Liu, and S. Wallin, "Binding of two intrinsically disordered peptides to a multi-specific protein: a combined Monte Carlo and molecular dynamics study," PLoS Comput. Biol. **8**, e1002682 (2012).
[45]H. Taketomi, Y. Ueda, and N. Gō, "Studies on protein folding, unfolding and fluctuations by computer simulation. I. The effects of specific amino acid sequence represented by specific inter-unit interactions," Int. J. Pept. Res. **7**, 445–459 (1975).
[46]J. D. Bryngelson and P. G. Wolynes, "Spin glasses and the statistical mechanics of protein folding," Proc. Natl. Acad. Sci. U.S.A. **84**, 7524–7528 (1987).
[47]H. S. Chan, Z. Zhang, S. Wallin, and Z. Liu, "Cooperativity, local-nonlocal coupling, and nonnative interactions: principles of protein folding from coarse-grained models," Annu. Rev. Phys. Chem. **62**, 301–326 (2011).
[48]C. Clementi, H. Nymeyer, and J. N. Onuchic, "Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? An investigation for small globular proteins," J. Mol. Biol. **298**, 937–953 (2000).
[49]Y. Levy, P. G. Wolynes, and J. N. Onuchic, "Protein topology determines binding mechanism," Proc. Natl. Acad. Sci. U.S.A. **101**, 511–516 (2004).
[50]S. Yang, S. S. Cho, Y. Levy, M. S. Cheung, H. Levine, P. G. Wolynes, and J. N. Onuchic, "Domain swapping is a consequence of minimal frustration," Proc. Natl. Acad. Sci. U.S.A. **101**, 13786–13791 (2004).
[51]M. B. Borgia, A. Borgia, R. B. Best, A. Steward, D. Nettels, B. Wunderlich, B. Schuler, and J. Clarke, "Single-molecule fluorescence reveals sequence-specific misfolding in multidomain proteins," Nature **474**, 662–665 (2011).

[52]W. Zhang, D. Ganguly, and J. Chen, "Residual structures, conformational fluctuations, and electrostatic interactions in the synergistic folding of two intrinsically disordered proteins," PLoS Comput. Biol. **8**, e1002353 (2012).
[53]D. Ganguly, W. Zhang, and J. Chen, "Synergistic folding of two intrinsically disordered proteins: searching for conformational selection," Mol. Biosyst. **8**, 198–209 (2012).
[54]R. B. Best, Y.-G. Chen, and G. Hummer, "Slow protein conformational dynamics from multiple experimental structures: the helix/sheet transition of Arc repressor," Structure **13**, 1755–1763 (2005).
[55]Y.-G. Chen and G. Hummer, "Slow conformational dynamics and unfolding of the calmodulin c-terminal domain," J. Am. Chem. Soc. **129**, 2414–2415 (2007).
[56]P. Maragakis and M. Karplus, "Large amplitude conformational change in proteins explored with a plastic network model: adenylate kinase," J. Mol. Biol. **352**, 807–822 (2005).
[57]K. Okazaki, N. Koga, S. Takada, J. N. Onuchic, and P. G. Wolynes, "Multiple-basin energy landscapes for large-amplitude conformational motions of proteins: Structure-based molecular dynamics simulations," Proc. Natl. Acad. Sci. U.S.A. **103**, 11844–11849 (2006).
[58]D. M. Zuckerman, "Simulation of an ensemble of conformational transitions in a united-residue model of calmodulin," J. Phys. Chem. B **108**, 5127–5137 (2004).
[59]A. Schug, P. C. Whitford, Y. Levy, and J. N. Onuchic, "Mutations as trapdoors to two competing native conformations of the Rop-dimer," Proc. Natl. Acad. Sci. U.S.A. **104**, 17674–17679 (2007).
[60]M. D. Daily, G. N. Phillips, and Q. Cui, "Many local motions cooperate to produce the adenylate kinase conformational transition," J. Mol. Biol. **400**, 618–631 (2010).
[61]M. D. Daily, G. N. Phillips, and Q. Cui, "Interconversion of functional motions between mesophilic and thermophilic adenylate kinases," PLoS Comput. Biol. **7**, e1002103 (2011).
[62]Y. Wang, L. Gan, E. Wang, and J. Wang, "Exploring the dynamic functional landscape of adenylate kinase modulated by substrates," J. Chem. Theory Comput. **9**, 84–95 (2013).
[63]G. de Marco and P. Várnai, "Molecular simulation of conformational transitions in biomolecules using a combination of structure-based potential and empirical valence bond theory," Phys. Chem. Chem. Phys. **11**, 10694–10700 (2009).
[64]L. Sutto, I. Mereu, and F. L. Gervasio, "A hybrid all-atom structure-based model for protein folding and large scale conformational transitions," J. Chem. Theor. Comput. **7**, 4208–4217 (2011).
[65]L. Sutto and C. Camilloni, "From A to B: a ride in the free energy surfaces of protein G domains suggests how new folds arise," J. Chem. Phys. **136**, 185101 (2012).
[66]W. Zheng, B. R. Brooks, and G. Hummer, "Protein conformational transitions explored by mixed elastic network models," Proteins **69**, 43–57 (2007).
[67]K. Okazaki and S. Takada, "Dynamic energy landscape view of coupled binding and protein conformational change: induced fit versus population-shift mechanisms," Proc. Natl. Acad. Sci. U.S.A. **105**, 11182–11187 (2008).
[68]J. Karanicolas and C. L. Brooks III, "The origins of asymmetry in the folding transition states of protein L and protein G," Prot. Sci. **11**, 2351–2361 (2002).
[69]J. D. Weeks, D. Chandler, and H. C. Andersen, "Role of repulsive forces in determining the equilibrium structure of simple liquids," J. Chem. Phys. **54**, 5237–5247 (1971).
[70]B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations," J. Comput. Chem. **4**, 187–217 (1983).
[71]B. R. Brooks, C. L. Brooks II, A. D. Mackerell, Jr., L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus, "CHARMM: The biomolecular simulation program," J. Comput. Chem. **30**, 1545–1614 (2009).
[72]R. B. Best and G. Hummer, "Diffusive model of protein folding dynamics with Kramers turnover in rate," Phys. Rev. Lett. **96**, 228104 (2006).
[73]T. Gallagher, P. Alexander, P. Bryan, and G. L. Gilliland, "Two crystal structures of the B1 immunoglobulin-binding domain of streptococcal protein G and comparison with NMR," Biochemistry **33**, 4721–4729 (1994).
[74]R. B. Best, "A "slow" protein folds quickly in the end," Proc. Natl. Acad. Sci. U.S.A. **110**, 5744–5745 (2013).
[75]B. A. Shoemaker, J. J. Portman, and P. G. Wolynes, "Speeding molecular recognition by using the folding funnel: the fly-casting mechanism," Proc. Natl. Acad. Sci. U.S.A. **97**, 8868–8873 (2000).
[76]D. U. Ferreiro, J. A. Hegler, E. A. Komives, and P. G. Wolynes, "Localizing frustration in native proteins and protein assemblies," Proc. Natl. Acad. Sci. U.S.A. **104**, 19819–19824 (2007).
[77]P. A. Alexander, Y. He, Y. Chen, J. Orban, and P. N. Bryan, "A minimal sequence code for switching protein structure and function," Proc. Natl. Acad. Sci. U.S.A. **106**, 21149–21154 (2009).
[78]Y. He, Y. Chen, P. A. Alexander, P. N. Bryan, and J. Orban, "Mutational tipping points for switching protein folds and functions," Structure **20**, 283–291 (2012).
[79]M. H. J. Cordes, R. E. Burton, N. P. Walsh, C. J. McKnight, and R. T. Sauer, "An evolutionary bridge to a new protein fold," Nat. Struct. Biol. **7**, 1129–1132 (2000).
[80]N. Tokuriki and D. S. Tawfik, "Protein dynamism and evolvability," Science **324**, 203–207 (2009).
[81]T. Sikosek, H. S. Chan, and E. Bornberg-Bauer, "Escape from adaptive conflict follows from weak functional trade-offs and mutational robustness," Proc. Natl. Acad. Sci. U.S.A. **109**, 14888–14893 (2012).
[82]T. Sikosek, E. Bornberg-Bauer, and H. S. Chan, "Evolutionary dynamics on protein bi-stability landscapes can potentially resolve adaptive conflicts," PLoS Comput. Biol. **8**, e1002659 (2012).
[83]R. L. Tuinstra, F. C. Peterson, S. K. E. S. Elgin, M. A. Kron, and B. F. Volkman, "Interconversion between two unrelated protein folds in the lymphotactin native state," Proc. Natl. Acad. Sci. U.S.A. **105**, 5057–5062 (2008).
[84]B. M. Burmann, S. H. Knauer, A. Sevostyanova, K. Schweimer, R. A. Mooney, R. Landick, I. Artsimovitch, and P. Rösch, "An α helix to β barrel domain switch transforms the transcription factor RfaH into a translation factor," Cell **150**, 291–303 (2012).
[85]E. Braselmann, J. L. Chaney, and P. L. Clark, "Folding the proteome," Trends Biochem. Sci. **38**, 337–344 (2013).
[86]B. C. Gin, J. P. Garrahan, and P. L. Geissler, "The limited role of nonnative contacts in the folding pathways of a lattice protein," J. Mol. Biol. **392**, 1303–1313 (2009).
[87]H.-X. Zhou, "Enhancement of association rates by nonspecific binding to DNA and cell membranes," Phys. Rev. Lett. **93**, 178101 (2004).
[88]A. M. Gronenborn, D. R. Filpula, N. Z. Essig, A. Achari, M. Whitlow, P. T. Wingfield, and G. M. Clore, "A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G," Science **253**, 657–661 (1991).