Templated folding of intrinsically disordered proteins

Angelo Toto[‡], **Francesca Malagrinò**[‡], **Lorenzo Visconti**[‡], **Francesca Troilo**[‡], **Livia Pagano**[‡], **Maurizio Brunori**[‡],
Per Jemth[§1], **and Stefano Gianni**[‡2]

From the [‡]*Istituto Pasteur—Fondazione Cenci Bolognetti, Dipartimento di Scienze Biochimiche "A. Rossi Fanelli" and Istituto di Biologia e Patologia Molecolari del CNR, Sapienza Università di Roma, 00185 Rome, Italy and the* [§]*Department of Medical Biochemistry and Microbiology, Uppsala University, BMC Box 582, SE-75123 Uppsala, Sweden*

Edited by Wolfgang Peti

Much of our current knowledge of biological chemistry is founded in the structure-function relationship, whereby sequence determines structure that determines function. Thus, the discovery that a large fraction of the proteome is intrinsically disordered, while being functional, has revolutionized our understanding of proteins and raised new and interesting questions. Many intrinsically disordered proteins (IDPs) have been determined to undergo a disorder-to-order transition when recognizing their physiological partners, suggesting that their mechanisms of folding are intrinsically different from those observed in globular proteins. However, IDPs also follow some of the classic paradigms established for globular proteins, pointing to important similarities in their behavior. In this review, we compare and contrast the folding mechanisms of globular proteins with the emerging features of binding-induced folding of intrinsically disordered proteins. Specifically, whereas disorder-to-order transitions of intrinsically disordered proteins appear to follow rules of globular protein folding, such as the cooperative nature of the reaction, their folding pathways are remarkably more malleable, due to the heterogeneous nature of their folding nuclei, as probed by analysis of linear free-energy relationship plots. These insights have led to a new model for the disorder-to-order transition in IDPs termed "templated folding," whereby the binding partner dictates distinct structural transitions *en route* to product, while ensuring a cooperative folding.

The study of the mechanism whereby a polypeptide chain acquires its native conformation has represented one of the biggest challenges of biochemistry and molecular biology over the last half a century (1). In fact, because the functions of proteins are mainly dictated by their three-dimensional structure, it is not surprising that the folding process has gained the attention of different scientific communities and that the folding field has played an influential role in protein science. Moreover, medical interest arose when it was realized that many human pathologies, including Alzheimer's, Parkinson's, and Creutzfeldt–Jacob diseases, are linked to misfolding and aggregation of specific proteins (2–4).

Some of the general properties of the folding process have already been characterized in detail, and our understanding of the basic features of this reaction has grown enormously over the last few decades. The collaborative efforts of experimentalists and theoreticians have resulted in substantial breakthroughs in the characterization of folding pathways, to the point that the folding of small single-domain proteins can be successfully described at nearly atomic resolution (5–16).

The discovery that about 40% of the human proteome is essentially disordered if expressed as single proteins (17–21) challenged existing paradigms drawn from the structure-function relationship. In addition to the complexities of integrating natively disordered sequences into our concepts of cell biology, researchers were immediately keen to know more about how and when these proteins adopt structure. It has been shown that whereas some of these proteins retain a high degree of disorder in all of their physiologically relevant states, a large fraction of them undergo a disorder-to-order folding transition upon binding to specific partners, which could be other proteins or nucleic acids (22–25). From a mechanistic perspective, proteins undergoing such coupled binding and folding process represent particularly interesting systems, where supramolecular organization is coupled to bimolecular recognition.

In this review, we explore the advances made thus far in understanding IDP folding by considering the main similarities and differences emerging from the comparison of the folding of globular *versus* intrinsically disordered proteins (IDPs).[3] We will first focus on the overall features of the observed folding transitions, as mirrored by the apparent kinetics. Then we will highlight the differences and similarities of the structural features of the main transition states. These comparisons have pointed to a new model for IDP folding called "templated folding," a mechanism that appears to capture the peculiar properties of the folding-upon-binding event of IDPs and can be used as a general framework for IDP research going forward.

[3] The abbreviations used are: IDP, intrinsically disordered protein; LFER, linear free-energy relationship.

Cooperativity in globular and IDP folding

The initial view of protein folding used to depict the reaction as a stepwise formation of native-like structure (26, 27). All proteins were assumed to fold via a framework model postulating secondary structure to form before tertiary interactions were locked in place. This scenario implicitly involved the accumulation of a series of partially structured intermediates, with an increasing degree of native-like structure. The discovery that small single-domain proteins were in fact capable of folding in an all-or-none fashion therefore came as a surprise. After the first observation of this so-called "two-state folding" by Jackson and Fersht in 1991 (28), it soon became clear that the majority of proteins containing less than 100 amino acids could fold in a highly cooperative manner, despite the formation and breakage of hundreds of weak noncovalent bonds. The two-state protein folding, where only the fully native and denatured states could be experimentally detected, revolutionized our view of the folding reaction and established some of the general rules in the field (29).

A two-state reaction is characterized by the lack of low-energy intermediates, populated either at equilibrium or transiently. Consequently, the denatured and folded states are separated by a single energy barrier defining the transition state, predicting that the time evolution of any two-state folding reaction should conform to a single-exponential decay (30). However, because of the complexity of these reactions and the challenges in monitoring this transition, it is difficult to conclusively determine whether intermediates do not exist or are just elusive. Rigorous experimental tests have therefore been employed to exclude their presence. For example, the two-state assumption can be verified by comparing the consistency of the thermodynamic parameters obtained from different experimental methods, such as by using different probes (*e.g.* CD, fluorescence, calorimetry, and NMR) and comparing equilibrium with kinetic experiments (31–33).

IDPs provided an immediate challenge to these expectations: Because IDPs are by definition unstructured and dynamic in isolation (34), the existence of cooperative motions and transition states was no longer guaranteed. Instead, the encounter and binding to a physiological partner often triggered a conformational change. In these cases, a disorder-to-order transition upon recognizing the physiological partner may occur. This phenomenon has generally been referred to as "coupled folding and binding" or "induced folding." The revelation of this transitions cast doubt whether known folding rules established for globular proteins would apply for IDPs and what possibilities might exist. Does induced folding follow an induced fit or conformational selection kind of model? Would the structure of the same IDP bound to different partners be the same? Would different partners bind to the same regions of the IDP?

Importantly, it should be noticed that what "induced folding" entails might also be very diverse for different IDPs. In fact, whereas in some cases the IDP may fold to a structured conformation, some other systems may retain a considerable level of disorder even in their bound states, a phenomenon generally referred to as "fuzziness" (Fig. 1) (35). Thus, there is a range of possible scenarios from disorder-to-(complete)order to disor-

Figure 1. Induced folding of intrinsically disordered proteins. It is well-known that some IDPs undergo a disorder-to-order transition upon binding their physiological partner. Nevertheless, the level of disorder retained in the complex may vary substantially in different cases. The figure reports two cases: the complex between KIX and pKID (23) and that between GCN4 and Med15 (90).

der-to-disorder transition, but, whatever the case, binding will result in some change of structure and dynamics of the bound ground state. Our ensuing discussion on templated folding might be valid for any degree of fuzziness.
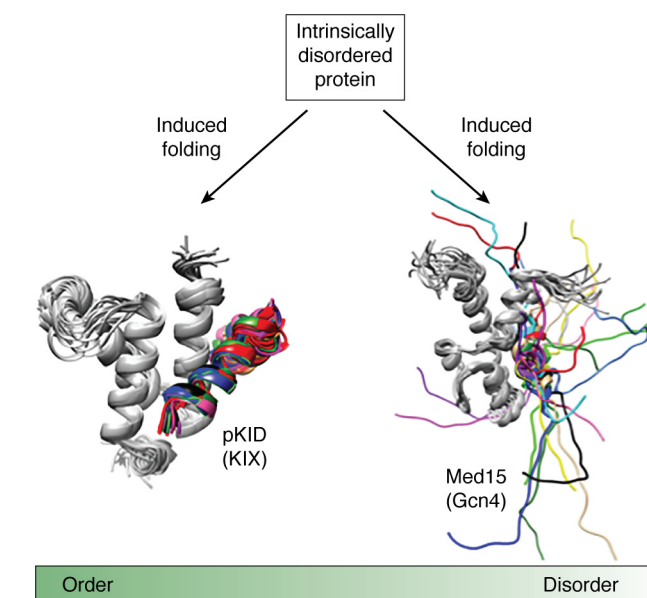
In theory, the mechanism of recognition between an IDP and its partner is expected to be a complex reaction involving, at least, the productive encounter between the two interacting units, the folding of the IDP in the bound conformation, and the locking of key stabilizing interactions. Thus, it would be expected that the mechanism of induced folding would imply the presence of multiple reaction intermediates. Nevertheless, in analogy to the folding of globular proteins, the expected theoretical complexity is contrasted by a striking simplicity of the binding kinetics.

Several IDPs were found to conform to robust two-state kinetics, displaying single-exponential time courses, as well as a linear dependence on reactant concentrations under pseudo-first-order conditions, another hallmark of two-state behavior (36–51). Note that apparent two-state does not imply that there are no intermediates, but that any intermediates are high-energy states, which rapidly convert to the equilibrium bound state or back to the free state. It is very difficult to directly assess these fast transitions between high-energy intermediates, but one of the first kinetic characterizations of an induced folding reaction by Hagen and co-workers (52) actually achieved this by using a laser-induced temperature jump spectroscopy. In this case, inducing folding with trifluoroethanol cosolvent allowed measurement of the folding and unfolding rate constant of an intrinsically disordered protease inhibitor (IA_3 from yeast), found to occur in the microsecond time range, indicating that the folding of IDPs may occur very rapidly.

Furthermore, it is important to highlight that not all of the IDPs studied to date conform to a two-state scenario. In fact,

reaction intermediates have been identified by NMR (53–55), by detecting multiexponential kinetics (56, 57) as well as by analyzing the concentration dependence of observed rate constants (56–60). Nevertheless, in most cases, these intermediates tend to be very elusive and scarcely populated, indicating that the cooperative nature of induced folding of IDPs is very similar to that of the folding of globular proteins.

The structure of the transition state: Φ value analysis

As mentioned above, examination of the transition state structure has come to be a central step in analyzing the folding pathways of globular proteins. However, whether similar considerations would be relevant for describing IDP folding was not immediately clear. The description of reaction mechanisms implies identifying all of the intermediates along the pathway and then characterizing their structure. As outlined above, in the case of protein folding, the process often takes place in a highly cooperative manner, such that most often only the fully native and denatured states may be populated. Thus, information that is accessible to the experimentalist is generally very limited as no snapshots between reactants and products can be characterized. In this context, it becomes clear how the study of the transition state is critical to pinpoint the key residues dictating folding.

Because the transition state at the top of the energetic barrier never accumulates, information about its structure must be inferred indirectly. In this context, a powerful approach, called Φ value analysis, has been conceived by Fersht and co-workers (61) and is based on the following methodology. In particular, by systematically mutating amino acid side chains while measuring the effect of each structural perturbation on the activation free-energy barrier and on the ground states, it is possible to map interaction patterns in the transition state. Quantitatively, the index of native-like structural content is measured by the Φ value, which normalizes the change in free energy of the transition state ($\Delta\Delta G^{\ddagger}$) upon mutation to that of the native state ($\Delta\Delta G$). A Φ value close to 1 is often interpreted as native-like structure in the folding transition state of that specific residue, whereas a Φ value equal to 0 suggests that the mutated residue is as unstructured in the transition state as it is in the denatured state.

The Φ value analysis represents the only experimental method available to describe the structure of a transition state, and it has been employed extensively to describe the mechanism of folding of globular proteins (62, 63) and lately the coupled binding and folding of IDPs (36–38, 40, 48, 50, 59, 60, 64–73). Interestingly, it has been generally observed that Φ values tend to be between 0 and 1, with very few cases of unusual values (*i.e.* lower than 0 or larger than 1). Hence, in analogy to what is observed for globular proteins, when folding takes place, IDPs tend to avoid misfolded conformations, and the transition state represents a distorted version of the bound state. Thus, whereas several alternative structures are populated by IDPs in their free state, the transition state of induced folding appears committed to the structure of the complex. Of additional interest, it may be noted that the magnitude of the Φ values, calculated for induced folding of different IDPs, tend to be broadly distributed along the sequence, rather than clus-

tered in regions of 0 and 1. This tendency, which has been discussed previously extensively in the case of globular proteins (6, 63), indicates that IDPs display structurally diffused nuclei in their folding transition states, where diverse residues contribute fractional formation of native-like contacts.

One of the main assumptions of the Φ value analysis lies in a negligible effect of the mutation on the structure of the denatured (or unbound) state (61, 63). In fact, when and if the structural perturbation induced by site-directed mutagenesis affects the stability of the denatured state (*e.g.* by disrupting an element of its residual structure), observed Φ values may deviate from the expected values, and unusual values may be detected (74). Because many IDPs retain elements of embryonic secondary or tertiary structure in their unbound states, it is worth considering this complication when comparing the results obtained with IDPs with those of globular proteins. As recalled above, however, none of the Φ value analyses performed to date present a significant number of unusual values, indicating that the effect of mutagenesis on the residual structure of the IDPs (studied so far) is not relevant enough to jeopardize these kinds of experiments and that such structures are relatively robust to the mutations considered. However, it will be important to conduct additional work to further support this conclusion.

Linear free-energy relationship (LFER) plots: The nucleation-condensation model versus templated folding

A general concept in chemical kinetics implies that optimal rates are obtained when reaction intermediate(s) have higher free energy than the reactants and products and are not accumulating, existing at very low concentrations (62, 75). Similarly, in the case of protein folding, optimal folding rate constants may be obtained when reaction intermediates are very unstable, so that the reaction is consistent with two-state folding (31). Under such conditions, local elements of structure display a low tendency to interact in isolation, and folding may only occur in a highly cooperative manner. This scenario, which is very recurrent in the folding of globular proteins, conforms to the so-called nucleation-condensation mechanism (31, 76, 77).

The nucleation-condensation model for globular proteins postulates the existence of a folding nucleus in globular proteins. Importantly, because such a nucleus is weak, structurally diffused, and extended, formation of the discontinuous network of interactions stabilizing the nucleus can occur only when a significant fraction of the overall fold has acquired an approximately correct overall conformation. Under such conditions, therefore, nucleation is coupled with general condensation of a native-like fold, with secondary and tertiary structure forming simultaneously and the transition state reflecting a distorted version of the native state (76, 77).

A valuable method to characterize the nucleation-condensation model is the analysis of LFER plots, also referred to as Leffler or Brønsted plots (78). This type of analysis correlates the activation free energy for a given reaction with its equilibrium free energy. LFER plots were originally introduced in physical organic chemistry to evaluate the position of the transition state along the reaction coordinate (79). In fact, by perturbing the structure and thereby the reactivity of a reactant,
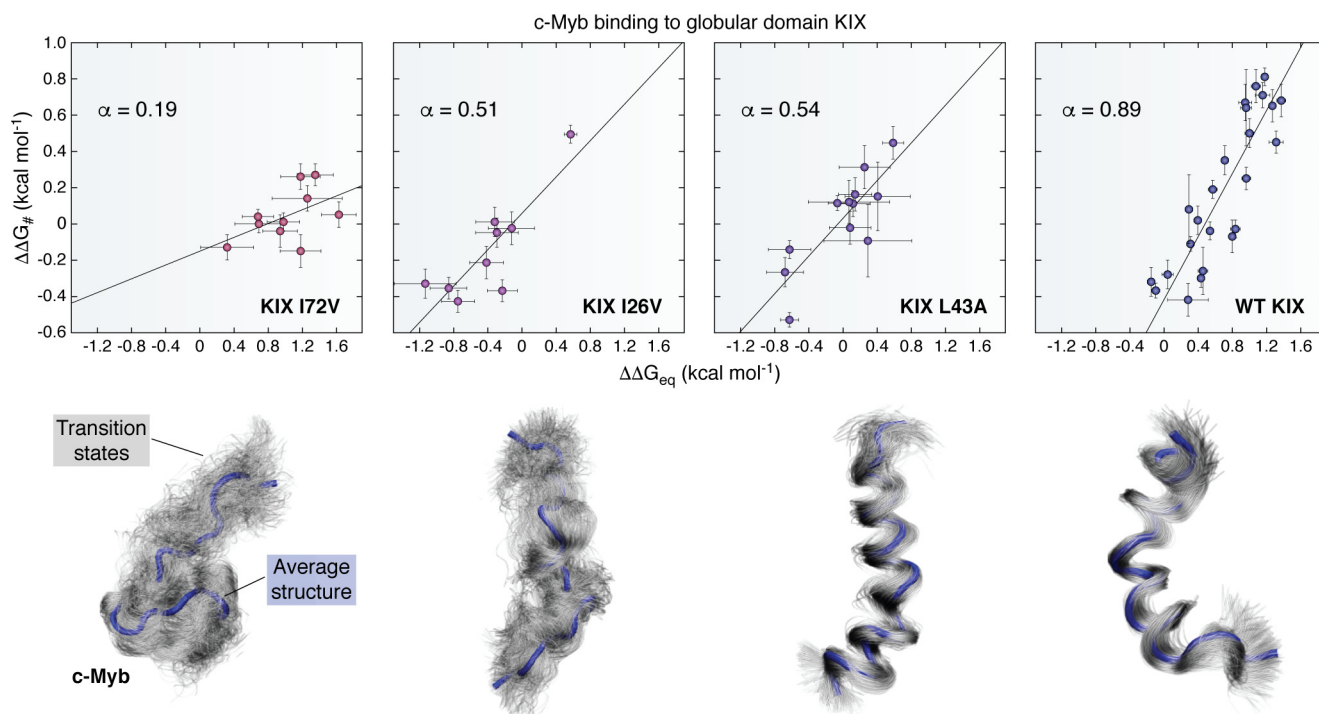
Figure 2. LFER analysis of the IDP c-Myb binding to KIX. Thermodynamic parameters were measured for the binding reaction between the intrinsically disordered protein c-Myb and the globular domain KIX, either the WT or three site-directed variants (I26V, L43A, I72V). The structural ensemble of the transition state of c-Myb (in *gray*) and its average structure (in *blue*) for the binding reaction with the different variants of KIX are shown *below* the respective LFER plot (data taken from Ref. 67). In each plot, each *point* represents to the change in activation free energy *versus* that of the bound state measured for a single site-directed mutant.

the dependence of the activation free energy on the equilibrium free energy generally results in a linear correlation, with a slope α reflecting the position of the transition state along the reaction coordinate. When applying this type of analysis to protein folding, linear LFERs may be obtained only when the transition state resembles the structure of the native state.

If two systems respond similarly to the same perturbation, induced for example by mutagenesis, it may be deduced that the systems are also similar in structure. Consequently, the linearity in the LFER plot represents strong experimental evidence that the transition state of folding reflects by-and-large a distorted version of the native state (78). Nearly all globular proteins that have been subjected to LFER analysis display a linear plot, suggesting that nucleation-condensation represents a general model for protein folding (13, 80, 81). Moreover, the slope of the LFER plot is highly similar for different proteins, with a value of α of ~0.3 (80). This finding indicates that, despite the fact that the folding nucleus is specific to each protein, the degree of native-like structure in the transition states of globular proteins is very robust.

It is of particular interest to compare the LFER analysis performed on globular proteins *versus* those of IDPs. The LFER analysis of the induced folding of different IDPs shows that this class of proteins also displays a linear correlation. Nonetheless, peculiarities have been discovered, as shown in the binding reaction of the transactivation domain of c-Myb, an IDP system, to the globular protein KIX. In this case, it was shown that the transition state contains a very high degree of native-like structure, with a value of α of 0.89. This finding indicated the

transition state of this IDP to be much more ordered than what is typically observed for the folding of globular proteins (48).

To test the robustness of the structure of the transition state of c-Myb, the LFER analysis was repeated, measuring the binding with three different site-directed mutants of KIX, displaying a different degree of hydrophobicity in their binding pocket (67). Surprisingly, whereas the linearity in the LFER plot was maintained, it was observed that mutations of KIX had a pronounced effect on the values of α calculated for the binding of c-Myb, with α progressively decreasing with decreasing hydrophobicity of the binding pocket, down to 0.19 in the case of the least hydrophobic variant (Fig. 2). This variability demonstrates a degree of structural plasticity in the transition state. Subsequently, other clear signatures of this plasticity have been observed, such as in the case of the interaction between NCBD and ACTR (82) and in the case of the induced folding of NTAIL with XD (59, 60). However, the protein BH3 appears more robust with regard to binding partners (37). Thus, at variance to what is observed for globular proteins, there is a remarkable structural malleability, or plasticity, in the transition state of induced folding, which is dictated by the binding partner.

Templated folding of IDPs

Despite the considerable interest in understanding the properties and peculiarities of IDPs, our current understanding of the mechanisms whereby binding-induced folding takes place is still relatively limited and based on the study of small protein systems. Furthermore, because the disorder-to-order transitions of IDPs are coupled to a binding reaction, it is generally

very difficult to experimentally pinpoint the characteristic features of the actual folding of the IDPs compared with those of globular proteins.

The comparison between the spontaneous folding of globular domains and the induced folding of IDPs presented in this review highlights some key differences, which can be interpreted in light of homogeneous and heterogeneous nucleation. In phase transitions, heterogeneous nucleation is a process whereby the interactions leading to the formation of the nucleus are established in contact either with the heterogeneities found in the generating phase or with a surface (83). On the other hand, homogeneous nucleation occurs through condensation of a single type of chemical compound. Thus, in binding-induced folding, the interacting partner of the IDP participates directly in the nucleation process, leading to a heterogeneous variable nucleation as opposed to the homogeneous nucleation observed in globular folding. Consequently, IDPs tend to follow a "templated folding" mechanism whereby the structure of the transition state is dictated by the nature of the interacting partner (Fig. 3). This behavior is in contrast to what would be expected from a homogeneous nucleation type mechanism, in which the disorder-to-order transitions would appear robust and imprinted in the amino acid sequence of the protein.

Whereas the folding nuclei of globular proteins are highly diffused, given that nucleation occurs homogeneously, the overall reaction appears rather robust. Thus, in spontaneous globular folding, nucleation sites are highly conserved and are maintained in different members of the same protein family, circular permutants, and truncated or circularized variants. This property is nicely illustrated by the so-called Φ-Φ plot analysis, where Φ values of corresponding residues in homologous proteins are plotted *versus* each other (13). On the other hand, the heterogeneous nucleation invoked by templated folding implies that binding-induced folding of IDPs would be more malleable, with alternative pathways and nucleation sites emerging with changing binding partners or experimental conditions (59, 60, 67, 70, 82, 84). It is likely that this behavior may also be reflected in structural malleability of the bound state upon changing experimental conditions and/or mutagenesis, a hypothesis that seems to be supported by the observed fuzziness of several IDP systems (24, 25, 35, 85, 86) as well as crystal structures of site-directed mutants (87). Such structural malleability could be further tested by monitoring the structural behavior of different variants of the same IDP by different techniques, such as NMR, SAXS, or single-molecule FRET. Furthermore, we predict that when and if the propensity to form ordered structure in an IDP is increased, the folding mechanism would most likely transition toward a classic homogeneous nucleation type scenario and would therefore appear more robust to changing conditions or binding partners. This hypothesis is readily testable by future experiments.

We note that templated folding is conceptually similar to the model of "slaving," previously introduced by Frauenfelder and collaborators (88, 89) to describe the role of solvent in controlling some of the dynamics and function of globins. In fact, by following this view, it was proposed that the conformational motions of proteins were essentially dictated by the external fluctuations of their hydration shell. Analogously, in the case of
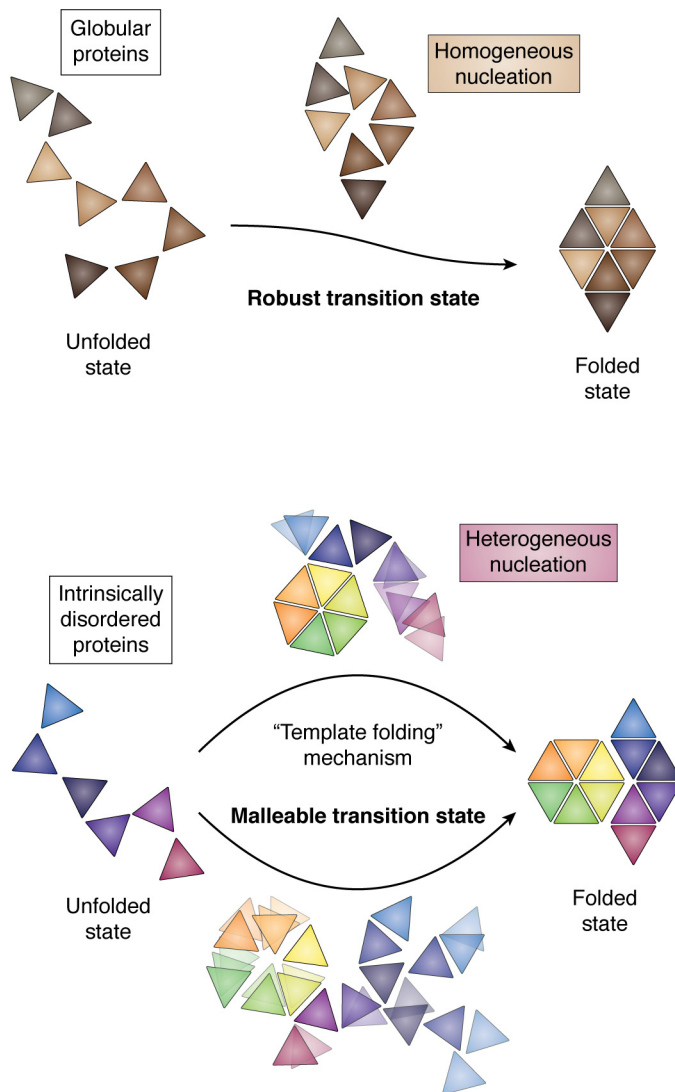

Figure 3. Homogeneous and heterogeneous nucleation in protein folding. Folding pathways of globular proteins (*top*) are characterized by a robust transition state, which forms through a homogeneous nucleation process. On the other hand, folding of IDPs may take place via heterogeneous nucleation and, therefore, may display a remarkable malleability of the transition state. The latter scenario is characteristic of the so-called "template folding" mechanism.

IDPs, the heterogeneous nature of their folding nuclei results in a remarkable malleability, which in their bound state is directly influenced by their physiological partner. We propose that templated folding represents a general mechanism whereby multiple alternative partners can recognize the same IDP and induce cooperative folding. In fact, templated folding ensures the robustness of the cooperativity and minimizes the possibility of establishing aberrant interactions with potential pathological consequences, while increasing the possibility of having an extended repertoire of different interaction partners.

References

1. Editorial (2005) So much more to know. *Science* **309,** 78–102 CrossRef Medline
2. Dobson, C. M., Knowles, T. P. J., and Vendruscolo, M. (2020) The amyloid phenomenon and its significance in biology and medicine. *Cold Spring Harb. Perspect. Biol.* **12,** a033878 Medline

3. Chiti, F., and Dobson, C. M. (2017) Protein misfolding, amyloid formation, and human disease: a summary of progress over the last decade. *Annu. Rev. Biochem.* **86,** 27–68 CrossRef Medline

4. Knowles, T. P., Vendruscolo, M., and Dobson, C. M. (2014) The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **15,** 384–396 CrossRef Medline

5. Bartlett, A. I., and Radford, S. E. (2009) An expanding arsenal of experimental methods yields an explosion of insights into protein folding mechanisms. *Nat. Struct. Mol. Biol.* **16,** 582–588 CrossRef Medline

6. Daggett, V., and Fersht, A. R. (2003) Is there a unifying mechanism for protein folding? *Trends Biochem. Sci.* **28,** 18–25 CrossRef Medline

7. Fersht, A. R. (2008) From the first protein structures to our current knowledge of protein folding: delights and scepticisms. *Nat. Rev. Mol. Cell Biol.* **9,** 650–654 CrossRef Medline

8. Nickson, A. A., Wensley, B. G., and Clarke, J. (2013) Take home lessons from studies of related proteins. *Curr. Opin. Struct. Biol.* **23,** 66–74 CrossRef Medline

9. O'Brien, E. P., Christodoulou, J., Vendruscolo, M., and Dobson, C. M. (2011) New scenarios of protein folding can occur on the ribosome. *J. Am. Chem. Soc.* **133,** 513–526 CrossRef Medline

10. Shirts, M., and Pande, V. S. (2000) COMPUTING: screen savers of the world unite!. *Science* **290,** 1903–1904 CrossRef Medline

11. Wolynes, P. G. (2005) Energy landscapes and solved protein-folding problems. *Philos. Trans. A Math. Phys. Eng. Sci.* **363,** 453–464; discussion 464–467 CrossRef Medline

12. Gianni, S., and Jemth, P. (2016) Protein folding: vexing debates on a fundamental problem. *Biophys. Chem.* **212,** 17–21 CrossRef Medline

13. Gianni, S., and Jemth, P. (2014) Conserved nucleation sites reinforce the significance of Phi value analysis in protein-folding studies. *IUBMB Life* **66,** 449–452 CrossRef Medline

14. Robustelli, P., Piana, S., and Shaw, D. E. (2018) Developing a molecular dynamics force field for both folded and disordered protein states. *Proc. Natl. Acad. Sci. U.S.A.* **115,** E4758–E4766 CrossRef Medline

15. Xiao, P., Bolton, D., Munro, R. A., Brown, L. S., and Ladizhansky, V. (2019) Solid-state NMR spectroscopy based atomistic view of a membrane protein unfolding pathway. *Nat. Commun.* **10,** 3867 CrossRef Medline

16. Luchinat, E., and Banci, L. (2018) In-cell NMR in human cells: direct protein expression allows structural studies of protein folding and maturation. *Acc. Chem. Res.* **51,** 1550–1557 CrossRef Medline

17. Dunker, A. K., Silman, I., Uversky, V. N., and Sussman, J. L. (2008) Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.* **18,** 756–764 CrossRef Medline

18. Dyson, H. J., and Wright, P. E. (2002) Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol.* **12,** 54–60 CrossRef Medline

19. Tompa, P. (2011) Unstructural biology coming of age. *Curr. Opin. Struct. Biol.* **21,** 419–425 CrossRef Medline

20. Uversky, V. N., and Dunker, A. K. (2010) Understanding protein non-folding. *Biochim. Biophys. Acta* **1804,** 1231–1264 CrossRef Medline

21. Wright, P. E., and Dyson, H. J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **293,** 321–331 CrossRef Medline

22. Demarest, S. J., Martinez-Yamout, M., Chung, J., Chen, H., Xu, W., Dyson, H. J., Evans, R. M., and Wright, P. E. (2002) Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **415,** 549–553 CrossRef Medline

23. Sugase, K., Dyson, H. J., and Wright, P. E. (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447,** 1021–1025 CrossRef Medline

24. Fuxreiter, M. (2019) Fold or not to fold upon binding: does it really matter? *Curr. Opin. Struct. Biol.* **54,** 19–25 CrossRef Medline

25. Tompa, P., and Fuxreiter, M. (2008) Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci.* **33,** 2–8 CrossRef Medline

26. Kim, P. S., and Baldwin, R. L. (1982) Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu. Rev. Biochem.* **51,** 459–489 CrossRef Medline

27. Creighton, T. E. (1988) Toward a better understanding of protein folding pathways. *Proc. Natl. Acad. Sci. U.S.A.* **85,** 5082–5086 CrossRef Medline

28. Jackson, S. E., and Fersht, A. R. (1991) Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* **30,** 10428–10435 CrossRef Medline

29. Jackson, S. E. (1998) How do small single-domain proteins fold? *Fold. Des.* **3,** R81–R91 CrossRef Medline

30. Arrhenius, S. (1889) Über die Reaktiongeschwindigkeitbei der Inversion von Rohrzucker durch Sauren. *Z. Phys. Chem.* **4,** CrossRef

31. Fersht, A. R. (1995) Optimization of rates of protein folding: the nucleation-condensation mechanism and its implications. *Proc. Natl. Acad. Sci. U.S.A.* **92,** 10869–10873 CrossRef Medline

32. Parker, M. J., Spencer, J., and Clarke, A. R. (1995) An integrated kinetic analysis of intermediates and transition states in protein folding reactions. *J. Mol. Biol.* **253,** 771–786 CrossRef Medline

33. Gianni, S., Ivarsson, Y., Jemth, P., Brunori, M., and Travaglini-Allocatelli, C. (2007) Identification and characterization of protein folding intermediates. *Biophys. Chem.* **128,** 105–113 CrossRef Medline

34. Adamski, W., Salvi, N., Maurin, D., Magnat, J., Milles, S., Jensen, M. R., Abyzov, A., Moreau, C. J., and Blackledge, M. (2019) A unified description of intrinsically disordered protein dynamics under physiological conditions using NMR spectroscopy. *J. Am. Chem. Soc.* **141,** 17817–17829 CrossRef Medline

35. Miskei, M., Horváth, A., Vendruscolo, M., and Fuxreiter, M. (2020) Sequence-based determinants and prediction of fuzzy interactions in protein complexes. *J. Mol. Biol.* S0022-2836(20)30190-X CrossRef Medline

36. Crabtree, M. D., Borcherds, W., Poosapati, A., Shammas, S. L., Daughdrill, G. W., and Clarke, J. (2017) Conserved helix-flanking prolines modulate intrinsically disordered protein:target affinity by altering the lifetime of the bound complex. *Biochemistry* **56,** 2379–2384 CrossRef Medline

37. Crabtree, M. D., Mendonça, C. A. T. F., Bubb, Q. R., and Clarke, J. (2018) Folding and binding pathways of BH3-only proteins are encoded within their intrinsically disordered sequence, not templated by partner proteins. *J. Biol. Chem.* **293,** 9718–9723 CrossRef Medline

38. Dahal, L., Kwan, T. O. C., Shammas, S. L., and Clarke, J. (2017) pKID binds to KIX via an unstructured transition state with nonnative interactions. *Biophys. J.* **113,** 2713–2722 CrossRef Medline

39. Milles, S., Mercadante, D., Aramburu, I. V., Jensen, M. R., Banterle, N., Koehler, C., Tyagi, S., Clarke, J., Shammas, S. L., Blackledge, M., Gräter, F., and Lemke, E. A. (2015) Plasticity of an ultrafast interaction between nucleoporins and nuclear transport receptors. *Cell* **163,** 734–745 CrossRef Medline

40. Rogers, J. M., Oleinikovas, V., Shammas, S. L., Wong, C. T., De Sancho, D., Baker, C. M., and Clarke, J. (2014) Interplay between partner and ligand facilitates the folding and binding of an intrinsically disordered protein. *Proc. Natl. Acad. Sci. U.S.A.* **111,** 15420–15425 CrossRef Medline

41. Rogers, J. M., Wong, C. T., and Clarke, J. (2014) Coupled folding and binding of the disordered protein PUMA does not require particular residual structure. *J. Am. Chem. Soc.* **136,** 5197–5200 CrossRef Medline

42. Shammas, S. L., Travis, A. J., and Clarke, J. (2014) Allostery within a transcription coactivator is predominantly mediated through dissociation rate constants. *Proc. Natl. Acad. Sci. U.S.A.* **111,** 12055–12060 CrossRef Medline

43. Shammas, S. L., Crabtree, M. D., Dahal, L., Wicky, B. I., and Clarke, J. (2016) Insights into coupled folding and binding mechanisms from kinetic studies. *J. Biol. Chem.* **291,** 6689–6695 CrossRef Medline

44. Åberg, E., Karlsson, O. A., Andersson, E., and Jemth, P. (2018) Binding kinetics of the intrinsically disordered p53 family transactivation domains and MDM2. *J. Phys. Chem. B* **122,** 6899–6905 CrossRef Medline

45. Dogan, J., Gianni, S., and Jemth, P. (2014) The binding mechanisms of intrinsically disordered proteins. *Phys. Chem. Chem. Phys.* **16,** 6323–6331 CrossRef Medline

46. Gianni, S., Dogan, J., and Jemth, P. (2016) Coupled binding and folding of intrinsically disordered proteins: what can we learn from kinetics? *Curr. Opin. Struct. Biol.* **36,** 18–24 CrossRef Medline

47. Gianni, S., Morrone, A., Giri, R., and Brunori, M. (2012) A folding-after-binding mechanism describes the recognition between the transactivation

domain of c-Myb and the KIX domain of the CREB-binding protein. *Biochem. Biophys. Res. Commun.* **428,** 205–209 CrossRef Medline

48. Giri, R., Morrone, A., Toto, A., Brunori, M., and Gianni, S. (2013) Structure of the transition state for the binding of c-Myb and KIX highlights an unexpected order for a disordered system. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 14942–14947 CrossRef Medline

49. Hemsath, L., Dvorsky, R., Fiegen, D., Carlier, M. F., and Ahmadian, M. R. (2005) An electrostatic steering mechanism of Cdc42 recognition by Wiskott-Aldrich syndrome proteins. *Mol. Cell.* **20,** 313–324 CrossRef Medline

50. Karlsson, O. A., Chi, C. N., Engström, A., and Jemth, P. (2012) The transition state of coupled folding and binding for a flexible β-finger. *J. Mol. Biol.* **417,** 253–261 CrossRef Medline

51. Nyqvist, I., and Dogan, J. (2019) Characterization of the dynamics and the conformational entropy in the binding between TAZ1 and CTAD-HIF-1α. *Sci. Rep.* **9,** 16557 CrossRef Medline

52. Narayanan, R., Ganesh, O. K., Edison, A. S., and Hagen, S. J. (2008) Kinetics of folding and binding of an intrinsically disordered protein: the inhibitor of yeast aspartic proteinase YPrA. *J. Am. Chem. Soc.* **130,** 11477–11485 CrossRef Medline

53. Arai, M., Sugase, K., Dyson, H. J., and Wright, P. E. (2015) Conformational propensities of intrinsically disordered proteins influence the mechanism of binding and folding. *Proc. Natl. Acad. Sci. U.S.A.* **112,** 9614–9619 CrossRef Medline

54. Desfosses, A., Milles, S., Jensen, M. R., Guseva, S., Colletier, J. P., Maurin, D., Schoehn, G., Gutsche, I., Ruigrok, R. W. H., and Blackledge, M. (2019) Assembly and cryo-EM structures of RNA-specific measles virus nucleocapsids provide mechanistic insight into paramyxoviral replication. *Proc. Natl. Acad. Sci. U.S.A.* **116,** 4256–4264 CrossRef Medline

55. Jensen, M. R., Communie, G., Ribeiro, E. A., Jr., Martinez, N., Desfosses, A., Salmon, L., Mollica, L., Gabel, F., Jamin, M., Longhi, S., Ruigrok, R. W., and Blackledge, M. (2011) Intrinsic disorder in measles virus nucleocapsids. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 9839–9844 CrossRef Medline

56. Dogan, J., Schmidt, T., Mu, X., Engström, Å., and Jemth, P. (2012) Fast association and slow transitions in the interaction between two intrinsically disordered protein domains. *J. Biol. Chem.* **287,** 34316–34324 CrossRef Medline

57. Chemes, L. B., Sánchez, I. E., and de Prat-Gay, G. (2011) Kinetic recognition of the retinoblastoma tumor suppressor by a specific protein target. *J. Mol. Biol.* **412,** 267–284 CrossRef Medline

58. Dosnon, M., Bonetti, D., Morrone, A., Erales, J., di Silvio, E., Longhi, S., and Gianni, S. (2015) Demonstration of a folding after binding mechanism in the recognition between the measles virus NTAIL and X domains. *ACS Chem. Biol.* **10,** 795–802 CrossRef Medline

59. Bonetti, D., Troilo, F., Brunori, M., Longhi, S., and Gianni, S. (2018) How robust is the mechanism of folding-upon-binding for an intrinsically disordered protein? *Biophys. J.* **114,** 1889–1894 CrossRef Medline

60. Troilo, F., Bonetti, D., Bignon, C., Longhi, S., and Gianni, S. (2019) Understanding intramolecular crosstalk in an intrinsically disordered protein. *ACS Chem. Biol.* **14,** 337–341 CrossRef Medline

61. Fersht, A. R., Matouschek, A., and Serrano, L. (1992) The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding. *J. Mol. Biol.* **224,** 771–782 CrossRef Medline

62. Fersht, A. R. (1999) *Structure and Mechanism in Protein Science*, pp. 325–365, Freeman, New York

63. Fersht, A. R., and Sato, S. (2004) Φ-Value analysis and the nature of protein-folding transition states. *Proc. Natl. Acad. Sci. U.S.A.* **101,** 7976–7981 CrossRef Medline

64. Dogan, J., Mu, X., Engström, Å., and Jemth, P. (2013) The transition state structure for coupled binding and folding of disordered protein domains. *Sci. Rep.* **3,** 2076 CrossRef Medline

65. Lindström, I., and Dogan, J. (2017) Native hydrophobic binding interactions at the transition state for association between the TAZ1 domain of CBP and the disordered TAD-STAT2 are not a requirement. *Biochemistry* **56,** 4145–4153 CrossRef Medline

66. Lindström, I., Andersson, E., and Dogan, J. (2018) The transition state structure for binding between TAZ1 of CBP and the disordered Hif-1α CAD. *Sci. Rep.* **8,** 7872 CrossRef Medline

67. Toto, A., Camilloni, C., Giri, R., Brunori, M., Vendruscolo, M., and Gianni, S. (2016) Molecular recognition by templated folding of an intrinsically disordered protein. *Sci. Rep.* **6,** 21994–22000 CrossRef Medline

68. Jemth, P., Mu, X., Engström, Å., and Dogan, J. (2014) A frustrated binding interface for intrinsically disordered proteins. *J. Biol. Chem.* **289,** 5528–5533 CrossRef Medline

69. Iešmantavičius, V., Dogan, J., Jemth, P., Teilum, K., and Kjaergaard, M. (2014) Helical propensity in an intrinsically disordered protein accelerates ligand binding. *Angew. Chem. Int. Ed. Engl.* **53,** 1548–1551 CrossRef Medline

70. Bonetti, D., Troilo, F., Toto, A., Brunori, M., Longhi, S., and Gianni, S. (2017) Analyzing the folding and binding steps of an intrinsically disordered protein by protein engineering. *Biochemistry* **56,** 3780–3786 CrossRef Medline

71. Toto, A., Bonetti, D., De Simone, A., and Gianni, S. (2017) Understanding the mechanism of binding between Gab2 and the C terminal SH3 domain from Grb2. *Oncotarget* **8,** 82344–82351 CrossRef Medline

72. Toto, A., and Gianni, S. (2016) Mutational analysis of the binding-induced folding reaction of the mixed-lineage leukemia protein to the KIX domain. *Biochemistry* **55,** 3957–3962 CrossRef Medline

73. Toto, A., Giri, R., Brunori, M., and Gianni, S. (2014) The mechanism of binding of the KIX domain to the mixed lineage leukemia protein and its allosteric role in the recognition of c-Myb. *Protein Sci.* **23,** 962–969 CrossRef Medline

74. Cho, J. H., and Raleigh, D. P. (2006) Denatured state effects and the origin of nonclassical Φ values in protein folding. *J. Am. Chem. Soc.* **128,** 16492–16493 CrossRef Medline

75. Fersht, A. R. (1974) Catalysis, binding and enzyme-substrate complementarity. *Proc. R. Soc. Lond. B Biol. Sci.* **187,** 397–407 CrossRef Medline

76. Abkevich, V. I., Gutin, A. M., and Shakhnovich, E. I. (1994) Specific nucleus as the transition state for protein folding: evidence from the lattice model. *Biochemistry* **33,** 10026–10036 CrossRef Medline

77. Itzhaki, L. S., Otzen, D. E., and Fersht, A. R. (1995) The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: evidence for a nucleation-condensation mechanism for protein folding. *J. Mol. Biol.* **254,** 260–288 CrossRef Medline

78. Fersht, A. R. (2004) Relationship of Leffler (Bronsted) α values and protein folding Φ values to position of transition-state structures on reaction coordinates. *Proc. Natl. Acad. Sci. U.S.A.* **101,** 14338–14342 CrossRef Medline

79. Leffler, J. E. (1953) Parameters for the description of transition states. *Science* **117,** 340–341 CrossRef Medline

80. Naganathan, A. N., and Muñoz, V. (2010) Insights into protein folding mechanisms from large scale analysis of mutational effects. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 8611–8616 CrossRef Medline

81. Sánchez, I. E., and Kiefhaber, T. (2003) Origin of unusual Φ-values in protein folding: evidence against specific nucleation sites. *J. Mol. Biol.* **334,** 1077–1085 CrossRef Medline

82. Karlsson, E., Andersson, E., Dogan, J., Gianni, S., Jemth, P., and Camilloni, C. (2019) A structurally heterogeneous transition state underlies coupled binding and folding of disordered proteins. *J. Biol. Chem.* **294,** 1230–1239 CrossRef Medline

83. Kelton, K. F., and Greer, A. (1988) Test of classical nucleation theory in a condensed system. *Phys. Rev. B Condens. Matter* **38,** 10089–10092 CrossRef Medline

84. Ou, L., Matthews, M., Pang, X., and Zhou, H. X. (2017) The dock-and-coalesce mechanism for the association of a WASP disordered region with the Cdc42 GTPase. *FEBS J.* **284,** 3381–3391 CrossRef Medline

85. Schuler, B., Borgia, A., Borgia, M. B., Heidarsson, P. O., Holmstrom, E. D., Nettels, D., and Sottini, A. (2020) Binding without folding: the biomolecular function of disordered polyelectrolyte complexes. *Curr. Opin. Struct. Biol.* **60,** 66–76 CrossRef Medline

86. Borgia, A., Borgia, M. B., Bugge, K., Kissling, V. M., Heidarsson, P. O., Fernandes, C. B., Sottini, A., Soranno, A., Buholzer, K. J., Nettels, D., Kragelund, B. B., Best, R. B., and Schuler, B. (2018) Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **555,** 61–66 CrossRef Medline

87. Mesrouze, Y., Bokhovchuk, F., Izaac, A., Meyerhofer, M., Zimmermann, C., Fontana, P., Schmelzle, T., Erdmann, D., Furet, P., Kallen, J., and Chène, P. (2018) Adaptation of the bound intrinsically disordered protein YAP to mutations at the YAP:TEAD interface. *Protein Sci.* **27,** 1810–1820 CrossRef Medline

88. Frauenfelder, H., Fenimore, P. W., and McMahon, B. H. (2002) Hydration, slaving and protein function. *Biophys. Chem.* **98,** 35–48 CrossRef Medline

89. Fenimore, P. W., Frauenfelder, H., McMahon, B. H., and Parak, F. G. (2002) Slaving: solvent fluctuations dominate protein dynamics and functions. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 16047–16051 CrossRef Medline

90. Tuttle, L. M., Pacheco, D., Warfield, L., Luo, J., Ranish, J., Hahn, S., and Klevit, R. E. (2018) Gcn4-mediator specificity is mediated by a large and dynamic fuzzy protein-protein complex. *Cell Rep.* **22,** 3251–3264 CrossRef Medline