ACCOUNTS of chemical research

pubs.acs.org/accounts Article

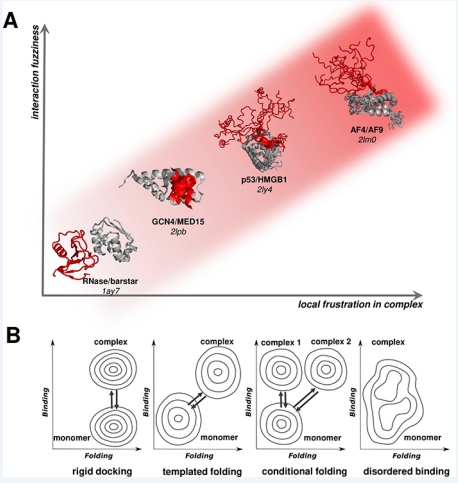
Fuzziness and Frustration in the Energy Landscape of Protein Folding, Function, and Assembly

Stefano Gianni,* María Inés Freiberger, Per Jemth, Diego U. Ferreiro, Peter G. Wolynes,* and Monika Fuxreiter*

Cite This: *Acc. Chem. Res.* 2021, 54, 1251−1259 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

CONSPECTUS: Are all protein interactions fully optimized? Do suboptimal interactions compromise specificity? What is the functional impact of frustration? Why does evolution not optimize some contacts? Proteins and their complexes are best described as ensembles of states populating an energy landscape. These ensembles vary in breadth from narrow ensembles clustered around a single average X-ray structure to broader ensembles encompassing a few different functional "taxonomic" states on to near continua of rapidly interconverting conformations, which are called "fuzzy" or even "intrinsically disordered". Here we aim to provide a comprehensive framework for confronting the structural and dynamical continuum of protein assemblies by combining the concepts of energetic frustration and interaction fuzziness. The diversity of the protein structural ensemble arises from the frustrated conflicts between the interactions that create the energy landscape. When frustration is minimal after folding, it results in a narrow ensemble, but residual frustrated interactions result in fuzzy ensembles, and this fuzziness allows a versatile repertoire of biological interactions. Here we discuss how fuzziness and frustration play off each other as proteins fold and assemble, viewing their significance from energetic, functional, and evolutionary perspectives.



We demonstrate, in particular, that the common physical origin of both concepts is related to the ruggedness of the energy landscapes, intramolecular in the case of frustration and intermolecular in the case of fuzziness. Within this framework, we show that alternative sets of suboptimal contacts may encode specificity without achieving a single structural optimum. Thus, we demonstrate that structured complexes may not be optimized, and energetic frustration is realized via different sets of contacts leading to multiplicity of specific complexes. Furthermore, we propose that these suboptimal, frustrated, or fuzzy interactions are under evolutionary selection and expand the biological repertoire by providing a multiplicity of biological activities. In accord, we show that non-native interactions in folding or interaction landscapes can cooperate to generate diverse functional states, which are essential to facilitate adaptation to different cellular conditions. Thus, we propose that not fully optimized structures may actually be beneficial for biological activities of proteins via an alternative set of suboptimal interactions. The importance of such variability has not been recognized across different areas of biology.

This account provides a modern view on folding, function, and assembly across the protein universe. The physical framework presented here is applicable to the structure and dynamics continuum of proteins and opens up new perspectives for drug design involving not fully structured, highly dynamic protein assemblies.

KEY REFERENCES

Fuxreiter, M. Fuzziness in Protein Interactions-A Historical Perspective. *J. Mol. Biol.* **2018**, *430*, 2278−2287.[1] *10-years overview of the fuzziness concept and application areas.*

Chen, M.; Chen, X.; Schafer, N. P.; Clementi, C.; Komives, E. A.; Ferreiro, D. U.; Wolynes, P. G. Surveying biomolecular frustration at atomic resolution. *Nat. Commun.* **2020**, *11*, 5944.[2] *A comprehensive overview of the frustration concept and its role in shaping protein function.*

Freiberger, M. I.; Wolynes, P. G.; Ferreiro, D. U.; Fuxreiter, M. Frustration in protein complexes leads to interaction versatility. 2020, *BioRxiv*. https://www.biorxiv.org/content/10.1101/2020.11.11.378091v1.[3] *Systematic analysis of 160 fuzzy complexes demonstrates a high degree of frustration and reveals the origin of specificity without an optimal structure.*

Jemth, P.; Karlsson, E.; Vogeli, B.; Guzovsky, B.; Andersson, E.; Hultqvist, G.; Dogan, J.; Guntert, P.; Riek, R.; Chi, C. N.

Received: December 2, 2020
Published: February 8, 2021

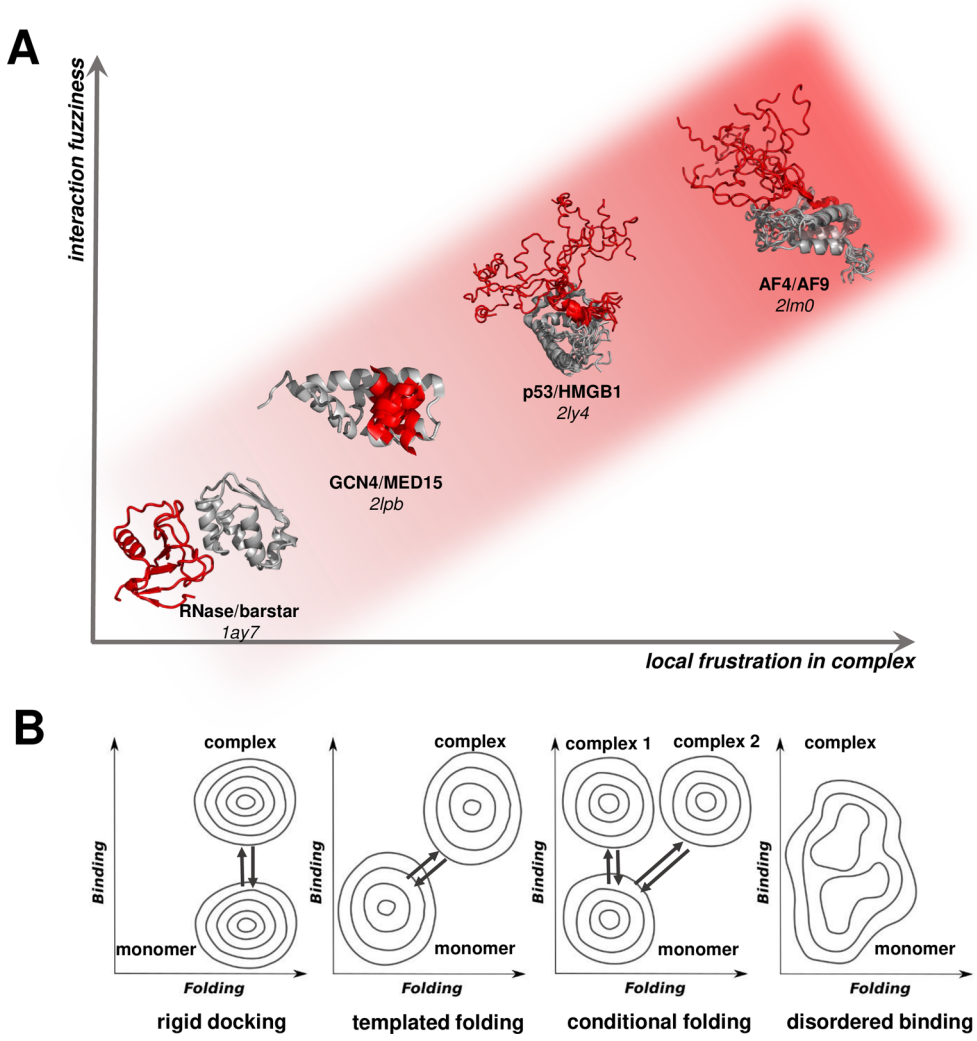


Figure 1. (A) Local frustration of the bound complex correlates to protein interaction fuzziness. Interactions between the two partners exhibit minimal frustration in rigid docking, which is coupled to moderate conformational changes upon interactions (represented by the RNase/barstar complex) and high frustration in disordered binding, when the bound partners are conformational heterogeneous in the complex (represented by the AF4/AF9 complex). Templated and conditional foldings, which are accompanied by a transition from disordered to ordered forms upon binding (represented by TADs of transcription factors GCN4 and p53), have intermediate local frustration, reflecting that the interactions of the folded elements are suboptimal. Indeed, in the case of conditional folding, the same protein region may remain disordered in complex with other proteins. (B) Folding and binding energy landscapes. The contour plots illustrate how folding frustration relates to interaction fuzziness by showing the energy landscapes in the free state (monomer) and the bound state (complex) along the folding and binding coordinates. In rigid docking, interactions take place between folded partners, resulting in a structured bound complex. In templated folding, the disordered partner(s) adopt a well-defined structure upon binding. Similar to templated folding, conditional folding may result in a structured complex upon binding with some partners (complex 2), but folding may not be induced upon assembly with other partners (complex 1). In the case of disordered binding, the energy landscapes of the monomeric and bound states overlap, as no folding takes place upon binding.

Structure and dynamics conspire in the evolution of affinity between intrinsically disordered proteins. *Sci. Adv.* **2018**, *4*, eaau4130.[4] *Biophysical analysis showing the interplay among plasticity, structure, and dynamics during evolution of an interaction involving disordered proteins.*

1. INTRODUCTION

The crystallographers' paradigm that function follows structure has played a pivotal role in protein science. Yet, the puzzle of how oxygen could get into a rigid hemoglobin arose as soon as its structure was determined. Proteins can be understood only when their motions are taken into account.[5] The molecular movements and rearrangements needed for function range from ps side-chain rotations to subunit and domain rearrangements on the ms time scale and on to the most extreme examples of protein dynamics: unfolding and/or folding of an entire protein, which can take times varying from microseconds to several seconds or minutes.[6] Even for well structured proteins characterizing their motions requires the notion of an energy landscape.[7]

Decades after the first determinations of protein structures it emerged that many proteins in their functional state, unlike hemoglobin, contain regions which continuously sample widely different conformations. The prevalence of such 'intrinsically disordered regions' created a new field within protein science and made it overwhelmingly obvious that the structure−function dogma must embrace dynamics.[8,9] Appreciating the dominant role of structural heterogeneity

https://dx.doi.org/10.1021/acs.accounts.0c00813
Acc. Chem. Res. 2021, 54, 1251−1259

manifested in these proteins has led to much new insight into the molecular mechanisms of biology. A range of functional characteristics is associated precisely with the lack of a unique structure: the structural diversity enables an increased interaction capacity with a repertoire of partners. The complexity of the landscape makes possible the complex logical operations needed in the regulation and coordination of different signaling pathways.[10] We now know that functional proteins as well as their complexes display a continuum of landscape patterns, ranging from those having nearly well-defined structures to those dominated by very fast-exchanging conformations.

Here we discuss how protein structural ensembles form along the full dynamical continuum and how these ensembles have been evolutionary selected to satisfy a variety of cellular requirements. Our discussion is informed by merging two concepts: the frustration concept, which arose in statistical physics and has been developed to quantify conformational variations in structured proteins,[11] and fuzziness, which captures functional heterogeneity seen so often in protein interactions[1] (Figure 1). Both concepts are deeply related to the ruggedness of the free energy landscape, whose quantitative analysis has formed the basis of the modern unified theory of protein folding and assembly.

2. PROTEIN FOLDING—FUNNELED ENERGY LANDSCAPES

Protein folding is the process whereby a polypeptide chain is self-guided spontaneously to its native, functional conformation. Levinthal pointed out that there is not enough time for a polypeptide to sample all the conformations to find the most stable one.[12] Originally, folding was conceived as a specific sequential mechanism characterized by the accumulation of partly folded intermediates, occurring through a stepwise folding, with local elements displaying a preformed structure progressively colliding to build the native state. This view was challenged by the discovery of globular domains capable of folding via an all-or-none two-state reaction.[13] In such cases, folding must be a highly cooperative reaction, involving concurrent formation of secondary and tertiary interactions, with no detectable intermediates.[14−16] These experimental observations could be understood using statistical mechanical energy landscape theory.[17,18] This theory resolved the Levinthal paradox by introducing the notion of a "folding funnel".

The consequences and implications of the funneled energy landscape theory have been extensively analyzed and reviewed over the years as the theory has continued to develop.[19−21] From a mechanistic perspective, one of the most important predictions arising from the funneled energy landscape theory is that the strong energetic bias toward the native conformation would correspond to a transition state ensemble for folding made up of mostly native-like structural features. This observation has been validated on different protein systems, whose transition state ensemble of folding can be described as a large set of structures all stabilized by only a fraction of the interactions in the native state.[22−24] More to the point, comparative studies on homologous proteins have revealed that these structural features arising from the protein topology are very robust, and proteins sharing the same structure but quite different sequences have been shown to fold via structurally similar transition state ensembles.[24−26] These transition state ensembles are generally well predicted using models with perfectly funneled landscapes.[19,27−29]

3. CONFORMATIONAL VARIATIONS—FRUSTRATION OF THE LANDSCAPE

The resolution of Levinthal's paradox that accounts for cooperative two-state folding is that the energy landscape of proteins is only minimally frustrated. Instead of there being strong conflicts among the interactions between residues in the amino acid sequence, evolution has selected sequences such that there is a strong energetic bias toward the native conformation.[27,30] But what is protein frustration? Colloquially, frustration is a condition that arises from the inability to fulfill several goals at once. In a physical system, there is frustration when each of the energetic interactions holding the protein together cannot be individually yet simultaneously minimized by a single conformation. The funneled energy landscape theory postulates that there is a strong energetic bias toward the native state in which these frustrated conflicts are largely absent. Natural proteins have been evolutionary sculpted by natural selection to be minimally frustrated. Many sequences can satisfy the principle of minimal frustration for the same structure, underlying the fact that proteins come in families. Statistical analysis shows that the result of random mutagenesis and selection on protein structure is quantitatively consistent with a highly funneled landscape.[31]

Proteins, however, have evolved not only to fold robustly but also to be able to "function".[32] An activity such as binding or catalytic action may be in conflict with the overall structural architecture.[33] Therefore, the native bias must be compromised locally by competing interactions in protein folding, which sometimes leads to some non-native structures in the folding process. An example is the intermediate formed in the folding of Im7, a small, fast folding protein that not only must fold but also must rapidly bind to a toxin[34−36]. Non-native contacts are also sampled in the folding transition state, as experimentally observed for frataxin.[34] The organization of the conformational space of globular proteins is nevertheless overall consistent with the principle of minimal frustration, implying the presence of converging routes eventually leading to the native state.[37] For functional reasons, however, proteins may contain a significant portion of frustration, which perturbs their folding routes. Consequently, competing routes do sometimes emerge, and non-native alternative free energy minima will be populated along the folding pathway.[34,38] In our view, these examples provide exceptions that, however, corroborate the generally funnel-like nature of protein folding landscapes.

4. TEMPLATED FOLDING—PLASTICITY OF FOLDING PATHWAYS

Globular proteins usually attain a low degree of frustration finally by having a well-defined tertiary structure. However, increasing the frustration level promotes the population of new conformational substates, which results in more heterogeneous ensembles. Under such conditions, protein function can exploit then the ability to visit many conformational substates whose population can be modulated by other partners. Structural diversity can be manifested in many different ways: having different secondary structure conformations with similar probabilities;[39,40] transiently populated conformations of flexible linkers, which control the arrangements of globular

https://dx.doi.org/10.1021/acs.accounts.0c00813
Acc. Chem. Res. 2021, 54, 1251−1259

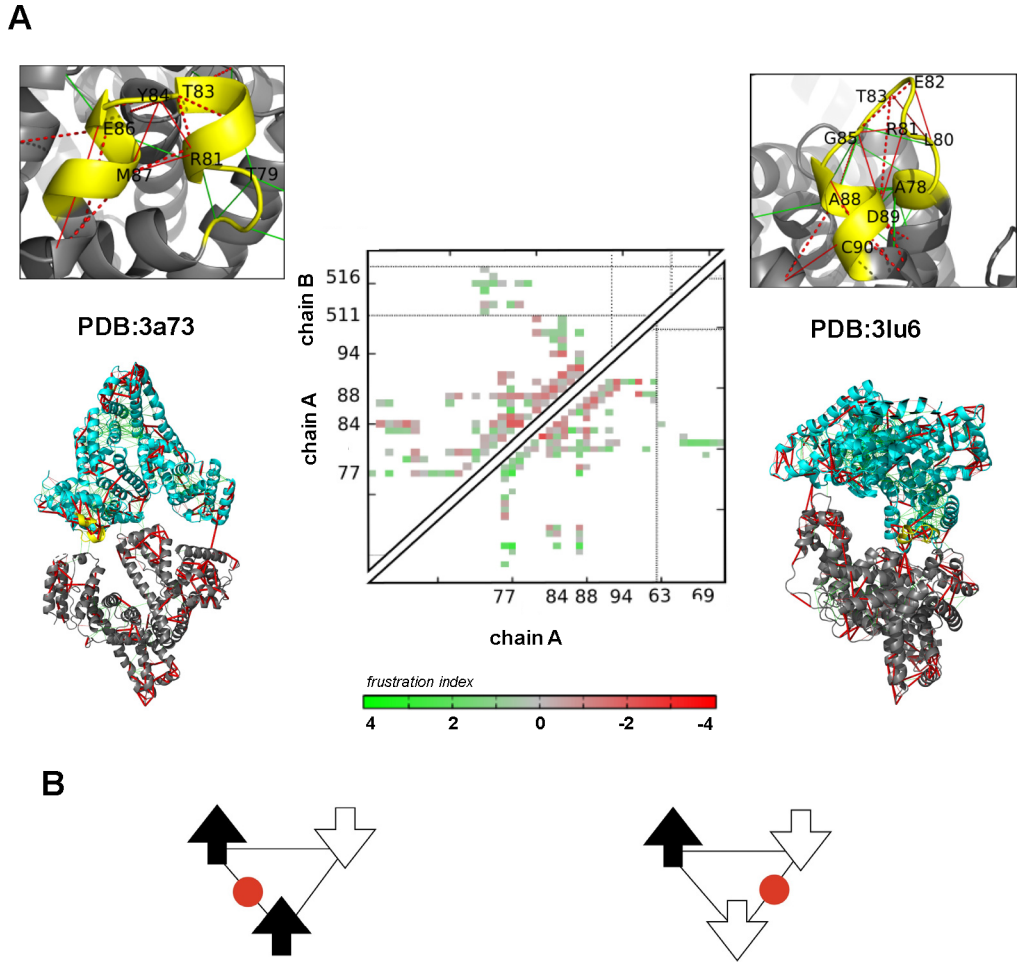


Figure 2. Specificity in a frustrated complex leads to different contact patterns with different partners. (A) Human serum albumin binds the natural ligand prostaglandin (PDB: 3a73) and a drug compound (PDB: 3lu6) in alternative binding modes via an extensive set of aromatic and electrostatic interactions. These contacts and the level of frustration can vary depending on the target and fine-tune affinities with different ligands. Local frustration values associated with the contacts are shown in Table S1. Local frustration patterns of thc are shown with minimally frustrated interactions in green and highly frustrated interactions in red. The fuzzy region (77−90 residues) is displayed as a yellow backbone. In the contact map, minimally frustrated contacts are shown in green, highly frustrated contacts are shown in red, and neutral contacts are shown in gray (as shown in the horizontal bar below the contact map). Some contacts are not displayed for better visualization. (B) Frustration in magnets also leads to alternative spin arrangements. A triangular ferromagnetic lattice is represented with the spins as arrows, and the favorable antiferromagnetic interactions are represented with lines. There is no way to arrange the spins such that they are all satisfied; in any case, an unfavorable interaction remains (red dot). This triangular lattice is thus energetically frustrated.

domains;[41] switching between folded and unfolded states;[42] as well as almost random seeming heteropolymers that interconvert between many different conformations.[43] Proteins possessing such very broad structural ensembles are often denoted as 'intrinsically disordered'. The native state of these proteins is a conformational ensemble, in contrast to a distinguished well-folded conformation.[39] Almost all natural proteins do, in fact, have some degree of frustration, sampling a continuum between order (low frustration) and disorder (high frustration).[44]

In principle, through interaction of intrinsically disordered regions with other partner proteins, proteins may decrease their local frustration. In a binding-induced folding process, highly frustrated disordered proteins adopt a well-defined conformation once bound[45,46] (Figure 1). In contrast to folding on highly funneled landscapes which often starts in clear locations,[14−16] such partner-assisted folding often may begin from several nuclei, where the physiological partner controls ("templates") the transition state of the folding reaction.[40,47−49] Thus, the transition state of templated folding is liable to change with the experimental or cellular conditions.

A typical signature of frustration in protein folding lies in the presence of transient non-native interactions along the major folding pathway. This effect has been observed in several different protein systems including the aforementioned Im7, the designed and overstabilized Top7,[50,51] alpha 3D,[52] and frataxin.[34] In these cases, large roughness of the folding energy landscape leads to multiple qualitatively different competing pathways. Different binding partners may tune the folding trajectories, so the distributions of the nucleation may differ. Consequently, although templated folding results in well-defined structures in the bound state, there is often considerable variability in the binding-induced folding process.

5. FUZZY BINDING−HETEROGENEITY IN BOUND COMPLEXES

Proteins often interact in a context-dependent manner, denoted as fuzzy binding.[53] Intrinsically disordered proteins,

https://dx.doi.org/10.1021/acs.accounts.0c00813
Acc. Chem. Res. 2021, 54, 1251−1259

in particular, may adopt folded structures in some of their complexes, while they remain disordered when interacting with other partners,[54] a phenomenon that is termed conditional folding (Figure 1). Such transitions may be also induced in the bound complex by post-translational modifications or can arise further via interactions with additional partners.[55,56] The binding can be so fuzzy so that different ordered structures are populated in the bound state, as exemplified by the p53 C-terminal region.[57] In this manner, fuzzy binding leads to a capability to interact in a versatile fashion with a defined but large set of partners. The phenomenon of fuzzy binding also suggests that templated folding can result in frustrated structures in the bound state. This may lead to conditional folding, when both ordered and disordered structures are populated in the bound state (Figure 1).[58,59]

Recently, the energy landscape theory was applied to assess the level of local frustration in protein complexes, which are formed by templated folding.[3] It has been shown that partner interactions reduce but do not always entirely eliminate local frustration in the final complexes of disordered protein regions. These results indicate that binding-induced folding often results in suboptimal interactions both at the binding interface as well as in the structured part of the protein. Protein regions which adopt ordered structures in a context-dependent manner also display frustrated patterns in their protein complexes.[3] The energy landscape theory has also illuminated how selectivity is achieved in frustrated bound complexes. Owing to the presence of multiple suboptimal conformations, binding to different partners results in distinct frustration patterns[3] (Figure 2, Table S1). Thus, the frustrated conformational energy landscape in the bound state seems to provide a functional advantage by allowing an increased interaction capacity with a limited set of partners.

The sequence properties that lead to the emergence of fuzzy binding have been recently elaborated. Interaction motifs, with distinguished composition and physicochemical features as compared to their flanking regions, will undergo disorder-to-order transition upon binding and fold in a partner induced manner.[60] In contrast, protein motifs similar to their embedding regions likely remain disordered in the bound state.[60] Within this framework, protein complexes sample a dynamical continuum from well-defined, ordered bound states to highly heterogeneous, disordered bound states, and the different modes of binding can be quantified by the conformational entropy. Frustration of the energy landscape in the bound state leads to variation between these different binding modes. Based on the local sequence biases with different possible binding sites, the pool of possible binding modes leading to fuzziness can also be quantified.[53] Frustration in the binding landscape also may stem from non-native interactions as observed for c-Myb/KIX[61−63] or ACTR/NCBD.[64,65]

All these alternative binding modes expand the functional repertoire of proteins, as the relative populations of the different conformational substates can now be shifted according to the cellular milieu.[66]

6. FUZZY INTERACTIONS−EVOLUTIONARY IMPLICATIONS

The possibility of expanding the functional repertoire through frustrated interactions suggests an interesting evolutionary mechanism. The ability of any system to respond to a selection pressure demands its constituents to be mutable and malleable. In the case of a cell, optimization of fitness is a constantly ongoing process in which conditions change over both evolutionary and ecological time scales and where surviving lineages manage to adapt via changes in expression levels of proteins as well as the protein sequences themselves. From a protein structure point of view, it is clear that both extremely well-folded as well as highly dynamic proteins are vital for life. It is not clear to what extent 'new' folded proteins arise during evolution, but it is likely that the majority of folded proteins today are descendants of ancient folded structures. On the other hand, short linear or peptide interaction motifs consisting of a few up to a dozen amino acid residues recurrently arise de novo, as demonstrated for virus-host protein−protein interactions.[67] The evolutionary history of the interaction between a longer disordered region called CID from the transcriptional coactivator ACTR and the NCBD domain of CBP/p300 has recently been investigated.[61,65] Here, a likely evolutionary scenario is that the ancestral interaction involved a weaker binding that was under positive selection and that a few key mutations in NCBD finally yielded a functional affinity. Interestingly, structural data[4] suggested that the plasticity of the disordered CID could relieve frustration and compensate for apparently detrimental point mutations.

The idea that structural pliability gives an evolutionary edge with respect to the emergence of new protein−protein interactions is appealing. Along these lines, it is likely that initial promiscuous interactions between a disordered region and a protein domain may explore alternative conformations. While many solved complexes between disordered regions and folded domains are structurally well-defined, like that for CID/NCBD, there are examples of positive selection for interactions with fuzzy complexes.[68,69] Coevolutionary analysis provides a powerful method to identify intermolecular interaction partners as well as long-range intramolecular contacts within a protein. Analysis of coevolving residues has demonstrated that the supramodular structure of the nonreceptor tyrosine kinase c-Src involves intramolecular contacts between two domains (Unique and SH3),[70] an interaction, which is structurally heterogeneous as shown by biophysical methods (Figure 3). Another example is the fuzzy interaction between S100B and ribosomal S6 kinase, which is conserved among vertebrates.[71] Thus, in a cell, disordered protein regions are constantly being exposed to various other proteins and engage in numerous promiscuous binding events, also referred to as quinary interactions.[72] Whenever such interactions are

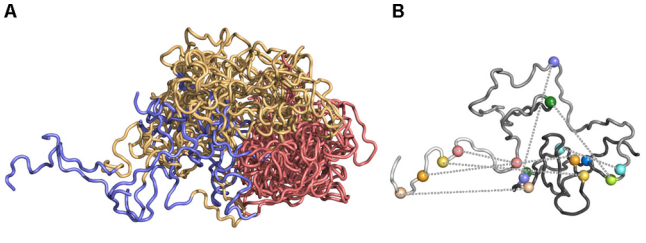


Figure 3. Fuzzy interactions with coevolutionary signals. (A) N-terminal regulatory region of c-Src, where the SH4 (teal), Unique (light orange), and SH3 (salmon) domains form a compact but highly dynamic, supramodular structure due to alternative long-range interactions. (B) Residues belonging to different domains with high coevolutionary probability on a single conformer. The three domains are shown by different shades of gray, which is gradually increasing from SH4 toward SH3. Residue pairs are displayed by identical colors.

https://dx.doi.org/10.1021/acs.accounts.0c00813
Acc. Chem. Res. 2021, 54, 1251−1259

advantageous to the organism, they might become fixed in the population by positive selection. The original interaction was likely frustrated and explored multiple conformations, but then depending on the evolutionary dice and functional restraints, further point mutations might lead the complex to evolve into a more minimally frustrated and structurally well-defined complex, like that between CID and NCBD. Alternatively, the complex may remain locally frustrated and structurally heterogeneous, fuzzy, as seen for Unique/SH3. It is clear that selection has generated a wide spectrum of local frustration levels in natural protein−protein interactions, resulting in well-defined complexes as well as sometimes structurally heterogeneous complexes.

7. A UNIFIED MODEL: FRUSTRATION VIS-À-VIS FUZZINESS

Frustration implies deviations from the perfect funnel-like energy landscape, often resulting in conformational heterogeneity of native proteins. Ruggedness of the landscape is caused by the suboptimal interactions that take place through non-native contacts, and thus we use it in thermodynamic terms. Multiple, minimalistic motifs coupled to non-native interactions may compromise folding. Ruggedness of the binding energy landscape can also arise from suboptimal, often redundant interaction motifs, for example, in low-complexity protein regions. Frustration results in alternative binding modes, which imparts plasticity to the templated folding mechanism and can enhance adaptability to a range of partners in a versatile form. Thus, ruggedness of the binding landscape leads to fuzzy binding and variations between many different interaction modes. Similar to frustration in free proteins, fuzziness of protein complexes also expands the functional repertoire to achieve fine-tuned, context-dependent regulation.[73,74]

Here we have shown that frustration and fuzziness are parallel concepts, aspects of suboptimal interactions of the energy landscape (Figure 1). This can be illustrated by the N-terminal regulatory region of c-Src, where the SH4, Unique, and SH3 domains form a compact but highly dynamic, supramodular structure due to alternative long-range interactions (Figure 3).[70] Residues mediating fuzzy interactions are well-conserved and often display coevolutionary signatures. This example illustrates the way fuzziness originating in frustrated contacts can control signaling specificity.

8. CONCLUSIONS

Several observations suggest that the contrasting requirements of physical folding and biological function may give rise to energetic frustration in natural proteins. Conformational heterogeneity promotes adaptability which itself may often be under evolutionary selection. Landscape ruggedness can be modulated by interactions with physiological partners. Suboptimal contacts with the partner often lead to fuzzy binding and give protein complexes with partner-specific, alternative binding modes. Although the concept of local frustration has been applied to folding and function of globular proteins, while fuzziness has mostly been discussed in the area of protein interactions and assemblies, these are really two sides of the same coin. Both frustration and fuzziness emerge from common physical principles of the energy landscape and can be described by a common formalism. We feel this realization can open new perspectives unifying protein folding and assembly with biological function.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.accounts.0c00813.

Table S1, frustration index for contacts displayed in Figure 2 (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Stefano Gianni − *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;* orcid.org/0000-0003-1653-1925; Email: stefano.gianni@uniroma1.it

Peter G. Wolynes − *Center for Theoretical Biological Physics, Rice University, Houston, Texas 77251-1892, United States;* orcid.org/0000-0001-7975-9287; Email: pwolynes@rice.edu

Monika Fuxreiter − *MTA-DE Laboratory of Protein Dynamics, Department of Biochemistry and Molecular Biology, University of Debrecen, H-4032 Debrecen, Hungary; Department of Biomedical Sciences, University of Padova, 35131 Padova, Italy;* orcid.org/0000-0002-4463-6727; Email: monika.fuxreiter@unipd.it

Authors

María Inés Freiberger − *Protein Physiology Lab, Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires-CONICET-IQUIBICEN, 1428 Buenos Aires, Argentina*

Per Jemth − *Department of Medical Biochemistry and Microbiology, Uppsala University, SE-75123 Uppsala, Sweden;* orcid.org/0000-0003-1516-7228

Diego U. Ferreiro − *Protein Physiology Lab, Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires-CONICET-IQUIBICEN, 1428 Buenos Aires, Argentina;* orcid.org/0000-0002-7869-4247

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.accounts.0c00813

Notes

The authors declare no competing financial interest.

Biographies

Stefano Gianni is a Professor of Molecular Biology in the Department of Biochemistry of the Sapienza University of Rome (Italy), where he is also Director of the Ph.D. Program in Biochemistry. He completed his Ph.D. in Rome, and after a post doc with Alan Fersht in Cambridge (UK), he started his independent group in 2005. His lab is known for the study of protein folding and protein−protein interactions, investigated through a combination of protein engineering and biophysical methods. Currently, he studies mechanisms of binding-induced folding of intrinsically disordered proteins.

Maria I. Freiberger is a Ph.D. student in the Department of Biological Chemistry, University of Buenos Aires (Argentina) with Dr. Diego U. Ferreiro on evolution and protein folding. She completed

https://dx.doi.org/10.1021/acs.accounts.0c00813
Acc. Chem. Res. 2021, 54, 1251−1259

her Bsc/MSc in Bioinformatics at the University of Entre Rios (Argentina) in 2017. She completed her Master's thesis analyzing local frustration in enzymes.

Per Jemth is a Professor of Protein Chemistry in the Department of Medical Biochemistry and Microbiology, Uppsala University (Sweden). He completed his Ph.D. in Uppsala, and after one postdoc in Uppsala with Ulf Lindahl and one in Cambridge (UK) with Alan Fersht, he started his independent group in 2004. His lab is known for experimental studies of protein−protein interactions with special focus on mechanism and evolution of folded as well as intrinsically disordered protein domains.

Diego U. Ferreiro is an Adjunct Professor at the Department of Biological Chemistry, Universidad de Buenos Aires, Argentina. He received his Ph.D. in Biochemistry from Buenos Aires studying protein−DNA interactions. He did his postdoc in Biophysics with Elizabeth Komives and Peter Wolynes at UCSD. His lab develops computational tools to analyze protein structures and dynamics in both physiological and evolutionary time scales.

Peter G. Wolynes is an American theoretical chemist and physicist. Since 2011 he has been a Bullard-Welch Foundation Professor of Science and Professor of Chemistry at Rice University. He is co-director of the Center for Theoretical Biological Physics at Rice University since 2019. He is widely recognized for his significant contributions to the theories of protein folding, glasses, and gene networks. Previously he was a James R. Eiszner Professor at the University of Illinois at Urbana−Champaign and the Francis H. C. Crick Chair of Physical Sciences at the University of California, San Diego.

Monika Fuxreiter is a Professor at the Department of Biomedical Sciences, University of Padova (Italy). She completed her Ph.D. in Budapest, Hungary in theoretical chemistry and protein crystallography and was a postdoc with Arieh Warshel (US), 2013 Nobel laureate in chemistry. She started her independent group in 2012 at Debrecen University (Hungary). She has made pioneering contributions to understanding the interactions of disordered proteins and coined the concept of fuzziness.

■ ACKNOWLEDGMENTS

This work was partly supported by grants from the Italian Ministero dell'Istruzione dell'Università e della Ricerca (Progetto di Interesse "Invecchiamento" to S.G.), Sapienza University of Rome (C26A155S48, B52F16003410005, and RP11715C34AEAC9B to S.G), the Associazione Italiana per la Ricerca sul Cancro (Individual Grant - IG 24551 to S.G.), the Hungarian Academy of Sciences 11015 and GINOP-2.3.2-15-2016-00044 (M.F.), and the Swedish Research Council (2020-04395 to P.J.). Also, this work was supported by the Center for Theoretical Biological Physics and sponsored by NSF grant PHY- 2019745 (to P.G.W). Additional support was also provided by the D. R. Bullard-Welch Chair at Rice University, Grant C-0016 (to P.G.W). D.U.F. is an investigator of CONICET and acknowledges grants ANPCyT - PICT2016-1467 and UBACYT 2018-20020170100540BA.

■ REFERENCES

(1) Fuxreiter, M. Fuzziness in Protein Interactions-A Historical Perspective. *J. Mol. Biol.* **2018**, *430*, 2278−2287.
(2) Chen, M.; Chen, X.; Schafer, N.; Clementi, C.; Komives, E.; Ferreiro, D.; Wolynes, P. Surveying biomolecular frustration at atomic resolution. *Nat. Commun.* **2020**, *11*, 5944.
(3) Freiberger, M. I.; Wolynes, P. G.; Ferreiro, D. U.; Fuxreiter, M. Frustration in protein complexes leads to interaction versatility. 2020, *BioRxiv*. https://www.biorxiv.org/content/10.1101/2020.11.11.378091v1 (accessed 2021-02-02).
(4) Jemth, P.; Karlsson, E.; Vogeli, B.; Guzovsky, B.; Andersson, E.; Hultqvist, G.; Dogan, J.; Guntert, P.; Riek, R.; Chi, C. N. Structure and dynamics conspire in the evolution of affinity between intrinsically disordered proteins. *Sci. Adv.* **2018**, *4*, eaau4130.
(5) Perutz, M. F. Nature of haem-haem interaction. *Nature* **1972**, *237*, 495−499.
(6) Lewandowski, J. R.; Halse, M. E.; Blackledge, M.; Emsley, L. Protein dynamics. Direct observation of hierarchical protein dynamics. *Science* **2015**, *348*, 578−581.
(7) Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G. The energy landscapes and motions of proteins. *Science* **1991**, *254*, 1598−1603.
(8) Wright, P. E.; Dyson, H. J. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **1999**, *293*, 321−331.
(9) Romero, P.; Obradovic, Z.; Kissinger, C. R.; Villafranca, J. E.; Garner, E.; Guillot, S.; Dunker, A. K. Thousands of proteins likely to have long disordered regions. *Pac. Symp. Biocomputing.* **1998**, *3*, 437−448.
(10) Bhattacharyya, R. P.; Remenyi, A.; Good, M. C.; Bashor, C. J.; Falick, A. M.; Lim, W. A. The Ste5 scaffold allosterically modulates signaling output of the yeast mating pathway. *Science* **2006**, *311*, 822−826.
(11) Ferreiro, D. U.; Komives, E. A.; Wolynes, P. G. Frustration in biomolecules. *Q. Rev. Biophys.* **2014**, *47*, 285−363.
(12) Levinthal, C. Are there pathways for protein folding? *J. Chim. Phys. Phys.-Chim. Biol.* **1968**, *65*, 44.
(13) Jackson, S. E.; Fersht, A. R. Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* **1991**, *30*, 10428−10435.
(14) Abkevich, V. I.; Gutin, A. M.; Shakhnovich, E. I. Impact of local and non-local interactions on thermodynamics and kinetics of protein folding. *J. Mol. Biol.* **1995**, *252*, 460−471.
(15) Abkevich, V. I.; Gutin, A. M.; Shakhnovich, E. I. Specific nucleus as the transition state for protein folding: evidence from the lattice model. *Biochemistry* **1994**, *33*, 10026−10036.
(16) Itzhaki, L. S.; Otzen, D. E.; Fersht, A. R. 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.* **1995**, *254*, 260−288.
(17) Bryngelson, J.; Wolynes, P. G. Spin glasses and the statistical mechanics of protein folding. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 7524−7528.
(18) Bryngelson, J.; Onuchic, J.; Socci, N.; Wolynes, P. G. Funnels, pathways, and the energy landscape of protein folding: a synthesis. *Proteins: Struct., Funct., Genet.* **1995**, *21*, 167−195.
(19) Onuchic, J. N.; Socci, N. D.; Luthey-Schulten, Z.; Wolynes, P. G. Protein folding funnels: the nature of the transition state ensemble. *Folding Des.* **1996**, *1*, 441−450.
(20) Onuchic, J. N.; Luthey-Schulten, Z.; Wolynes, P. G. Theory of protein folding: the energy landscape perspective. *Annu. Rev. Phys. Chem.* **1997**, *48*, 545−600.
(21) Socci, N. D.; Onuchic, J. N.; Wolynes, P. G. Protein folding mechanisms and the multidimensional folding funnel. *Proteins: Struct., Funct., Genet.* **1998**, *32*, 136−158.
(22) Daggett, V.; Fersht, A. R. Is there a unifying mechanism for protein folding? *Trends Biochem. Sci.* **2003**, *28*, 18−25.
(23) Sanchez, I. E.; Kiefhaber, T. Origin of unusual phi-values in protein folding: evidence against specific nucleation sites. *J. Mol. Biol.* **2003**, *334*, 1077−1085.
(24) Gianni, S.; Jemth, P. Conserved nucleation sites reinforce the significance of Phi value analysis in protein-folding studies. *IUBMB Life* **2014**, *66*, 449−452.
(25) Clarke, J.; Cota, E.; Fowler, S. B.; Hamill, S. J. Folding studies of immunoglobulin-like beta-sandwich proteins suggest that they share a common folding pathway. *Structure* **1999**, *7*, 1145−1153.

https://dx.doi.org/10.1021/acs.accounts.0c00813
Acc. Chem. Res. 2021, 54, 1251−1259

(26) Calosci, N.; Chi, C. N.; Richter, B.; Camilloni, C.; Engstrom, A.; Eklund, L.; Travaglini-Allocatelli, C.; Gianni, S.; Vendruscolo, M.; Jemth, P. Comparison of successive transition states for folding reveals alternative early folding pathways of two homologous proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 19241−19246.
(27) Oliveberg, M.; Wolynes, P. G. The experimental survey of protein-folding energy landscapes. *Q. Rev. Biophys.* **2005**, *38*, 245−288.
(28) Clementi, C.; Nymeyer, H.; Onuchic, J. N. 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.* **2000**, *298*, 937−953.
(29) Chung, H.; Piana-Agustinetti, S.; Shaw, D.; Eaton, W. A. Structural origin of slow diffusion in protein folding. *Science* **2015**, *349*, 1504−1510.
(30) Bryngelson, J. D.; Wolynes, P. G. Spin glasses and the statistical mechanics of protein folding. *Proc. Natl. Acad. Sci. U. S. A.* **1987**, *84*, 7524−7528.
(31) Morcos, F.; Schafer, N. P.; Cheng, R. R.; Onuchic, J. N.; Wolynes, P. G. Coevolutionary information, protein folding landscapes, and the thermodynamics of natural selection. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 12408−12413.
(32) Ferreiro, D. U.; Komives, E. A.; Wolynes, P. G. Frustration, function and folding. *Curr. Opin. Struct. Biol.* **2018**, *48*, 68−73.
(33) Ferreiro, D. U.; Hegler, J. A.; Komives, E. A.; Wolynes, P. G. Localizing frustration in native proteins and protein assemblies. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 19819−19824.
(34) Gianni, S.; Camilloni, C.; Giri, R.; Toto, A.; Bonetti, D.; Morrone, A.; Sormanni, P.; Brunori, M.; Vendruscolo, M. Understanding the frustration arising from the competition between function, misfolding, and aggregation in a globular protein. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 14141−14146.
(35) Sutto, L.; Latzer, J.; Hegler, J. A.; Ferreiro, D. U.; Wolynes, P. G. Consequences of localized frustration for the folding mechanism of the IM7 protein. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 19825−19830.
(36) Chen, T.; Sun Chan, H. Native contact density and nonnative hydrophobic effects in the folding of bacterial immunity proteins. *PLoS Comput. Biol.* **2015**, *11*, e1004260.
(37) Eaton, W. A.; Wolynes, P. G. Theory, simulations, and experiments show that proteins fold by multiple pathways. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, E9759−E9760.
(38) Di Silvio, E.; Brunori, M.; Gianni, S. Frustration Sculpts the Early Stages of Protein Folding. *Angew. Chem., Int. Ed.* **2015**, *54*, 10867−10869.
(39) Sormanni, P.; Piovesan, D.; Heller, G. T.; Bonomi, M.; Kukic, P.; Camilloni, C.; Fuxreiter, M.; Dosztanyi, Z.; Pappu, R. V.; Babu, M. M.; Longhi, S.; Tompa, P.; Dunker, A. K.; Uversky, V. N.; Tosatto, S. C.; Vendruscolo, M. Simultaneous quantification of protein order and disorder. *Nat. Chem. Biol.* **2017**, *13*, 339−342.
(40) Toto, A.; Camilloni, C.; Giri, R.; Brunori, M.; Vendruscolo, M.; Gianni, S. Molecular Recognition by Templated Folding of an Intrinsically Disordered Protein. *Sci. Rep.* **2016**, *6*, 21994.
(41) Chasman, D.; Cepek, K.; Sharp, P. A.; Pabo, C. O. Crystal structure of an OCA-B peptide bound to an Oct-1 POU domain/octamer DNA complex: specific recognition of a protein-DNA interface. *Genes Dev.* **1999**, *13*, 2650−2657.
(42) Bah, A.; Vernon, R. M.; Siddiqui, Z.; Krzeminski, M.; Muhandiram, R.; Zhao, C.; Sonenberg, N.; Kay, L. E.; Forman-Kay, J. D. Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature* **2015**, *519*, 106−109.
(43) Pometun, M. S.; Chekmenev, E. Y.; Wittebort, R. J. Quantitative observation of backbone disorder in native elastin. *J. Biol. Chem.* **2004**, *279*, 7982−7987.
(44) van der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R. J.; Daughdrill, G. W.; Dunker, A. K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D. T.; Kim, P. M.; Kriwacki, R. W.; Oldfield, C. J.; Pappu, R. V.; Tompa, P.; Uversky, V. N.; Wright, P. E.; Babu, M. M. Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **2014**, *114*, 6589−6631.
(45) Kussie, P. H.; Gorina, S.; Marechal, V.; Elenbaas, B.; Moreau, J.; Levine, A. J.; Pavletich, N. P. Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science* **1996**, *274*, 948−953.
(46) Mogridge, J.; Legault, P.; Li, J.; Van Oene, M. D.; Kay, L. E.; Greenblatt, J. Independent ligand-induced folding of the RNA-binding domain and two functionally distinct antitermination regions in the phage lambda N protein. *Mol. Cell* **1998**, *1*, 265−275.
(47) Bonetti, D.; Troilo, F.; Brunori, M.; Longhi, S.; Gianni, S. How Robust Is the Mechanism of Folding-Upon-Binding for an Intrinsically Disordered Protein? *Biophys. J.* **2018**, *114*, 1889−1894.
(48) Karlsson, E.; Andersson, E.; Dogan, J.; Gianni, S.; Jemth, P.; Camilloni, C. A structurally heterogeneous transition state underlies coupled binding and folding of disordered proteins. *J. Biol. Chem.* **2019**, *294*, 1230−1239.
(49) Toto, A.; Malagrino, F.; Visconti, L.; Troilo, F.; Pagano, L.; Brunori, M.; Jemth, P.; Gianni, S. Templated folding of intrinsically disordered proteins. *J. Biol. Chem.* **2020**, *295*, 6586−6593.
(50) Watters, A. L.; Deka, P.; Corrent, C.; Callender, D.; Varani, G.; Sosnick, T.; Baker, D. The highly cooperative folding of small naturally occurring proteins is likely the result of natural selection. *Cell* **2007**, *128*, 613−624.
(51) Zhang, Z.; Sun Chan, H. Competition between native topology and nonnative interactions in simple and complex folding kinetics of natural and designed proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 2920−2925.
(52) Best, R. B.; Hummer, G.; Eaton, W. A. Native contacts determine protein folding mechanisms in atomistic simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 17874−17879.
(53) Horvath, A.; Miskei, M.; Ambrus, V.; Vendruscolo, M.; Fuxreiter, M. Sequence-based prediction of protein binding mode landscapes. *PLoS Comput. Biol.* **2020**, *16*, e1007864.
(54) Miskei, M.; Antal, C.; Fuxreiter, M. FuzDB: database of fuzzy complexes, a tool to develop stochastic structure-function relationships for protein complexes and higher-order assemblies. *Nucleic Acids Res.* **2017**, *45*, D228−D235.
(55) Delaforge, E.; Kragelj, J.; Tengo, L.; Palencia, A.; Milles, S.; Bouvignies, G.; Salvi, N.; Blackledge, M.; Jensen, M. R. Deciphering the Dynamic Interaction Profile of an Intrinsically Disordered Protein by NMR Exchange Spectroscopy. *J. Am. Chem. Soc.* **2018**, *140*, 1148−1158.
(56) Tuttle, L. M.; Pacheco, D.; Warfield, L.; Luo, J.; Ranish, J.; Hahn, S.; Klevit, R. E. Gcn4-Mediator Specificity Is Mediated by a Large and Dynamic Fuzzy Protein-Protein Complex. *Cell Rep.* **2018**, *22*, 3251−3264.
(57) Oldfield, C. J.; Meng, J.; Yang, J. Y.; Yang, M. Q.; Uversky, V. N.; Dunker, A. K. Flexible nets: disorder and induced fit in the associations of p53 and 14−3-3 with their partners. *BMC Genomics* **2008**, *9*, S1.
(58) Fuxreiter, M. Classifying the Binding Modes of Disordered Proteins. *Int. J. Mol. Sci.* **2020**, *21*, 8615.
(59) Korennykh, A. V.; Egea, P. F.; Korostelev, A. A.; Finer-Moore, J.; Zhang, C.; Shokat, K. M.; Stroud, R. M.; Walter, P. The unfolded protein response signals through high-order assembly of Ire1. *Nature* **2009**, *457*, 687−693.
(60) Miskei, M.; Horvath, A.; Vendruscolo, M.; Fuxreiter, M. Sequence-Based Prediction of Fuzzy Protein Interactions. *J. Mol. Biol.* **2020**, *432*, 2289−2303.
(61) Hultqvist, G.; Aberg, E.; Camilloni, C.; Sundell, G. N.; Andersson, E.; Dogan, J.; Chi, C. N.; Vendruscolo, M.; Jemth, P. Emergence and evolution of an interaction between intrinsically disordered proteins. *eLife* **2017**, *6*, e16059.
(62) Toto, A.; Giri, R.; Brunori, M.; Gianni, S. 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.* **2014**, *23*, 962−969.

https://dx.doi.org/10.1021/acs.accounts.0c00813
Acc. Chem. Res. 2021, 54, 1251−1259

(63) Giri, R.; Morrone, A.; Toto, A.; Brunori, M.; Gianni, S. 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.* **2013**, *110*, 14942−14947.
(64) Jemth, P.; Mu, X.; Engstrom, A.; Dogan, J. A frustrated binding interface for intrinsically disordered proteins. *J. Biol. Chem.* **2014**, *289*, 5528−5533.
(65) Karlsson, E.; Paissoni, C.; Erkelens, A. M.; Tehranizadeh, Z. A.; Sorgenfrei, F. A.; Andersson, E.; Ye, W.; Camilloni, C.; Jemth, P. Mapping the transition state for a binding reaction between ancient intrinsically disordered proteins. *J. Biol. Chem.* **2020**, *295*, 17698.
(66) Pricer, R.; Gestwicki, J. E.; Mapp, A. K. From Fuzzy to Function: The New Frontier of Protein-Protein Interactions. *Acc. Chem. Res.* **2017**, *50*, 584−589.
(67) Davey, N. E.; Cyert, M. S.; Moses, A. M. Short linear motifs - ex nihilo evolution of protein regulation. *Cell Commun. Signaling* **2015**, *13*, 43.
(68) Kurzbach, D.; Schwarz, T. C.; Platzer, G.; Hofler, S.; Hinderberger, D.; Konrat, R. Compensatory adaptations of structural dynamics in an intrinsically disordered protein complex. *Angew. Chem., Int. Ed.* **2014**, *53*, 3840−3843.
(69) Schwarten, M.; Solyom, Z.; Feuerstein, S.; Aladag, A.; Hoffmann, S.; Willbold, D.; Brutscher, B. Interaction of nonstructural protein 5A of the hepatitis C virus with Src homology 3 domains using noncanonical binding sites. *Biochemistry* **2013**, *52*, 6160−6168.
(70) Arbesu, M.; Maffei, M.; Cordeiro, T. N.; Teixeira, J. M.; Perez, Y.; Bernado, P.; Roche, S.; Pons, M. The Unique Domain Forms a Fuzzy Intramolecular Complex in Src Family Kinases. *Structure* **2017**, *25*, 630−640.e4.
(71) Gogl, G.; Alexa, A.; Kiss, B.; Katona, G.; Kovacs, M.; Bodor, A.; Remenyi, A.; Nyitray, L. Structural Basis of Ribosomal S6 Kinase 1 (RSK1) Inhibition by S100B Protein: Modulation of the extracellular signal-regulated kinase (ERK) signaling cascade in a calcium-dependent way. *J. Biol. Chem.* **2016**, *291*, 11−27.
(72) Monteith, W. B.; Cohen, R. D.; Smith, A. E.; Guzman-Cisneros, E.; Pielak, G. J. Quinary structure modulates protein stability in cells. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 1739−1742.
(73) Nnamani, M. C.; Ganguly, S.; Erkenbrack, E. M.; Lynch, V. J.; Mizoue, L. S.; Tong, Y.; Darling, H. L.; Fuxreiter, M.; Meiler, J.; Wagner, G. P. A Derived Allosteric Switch Underlies the Evolution of Conditional Cooperativity between HOXA11 and FOXO1. *Cell Rep.* **2016**, *15*, 2097−2108.
(74) Miskei, M.; Gregus, A.; Sharma, R.; Duro, N.; Zsolyomi, F.; Fuxreiter, M. Fuzziness enables context dependence of protein interactions. *FEBS Lett.* **2017**, *591*, 2682−2695.

https://dx.doi.org/10.1021/acs.accounts.0c00813
Acc. Chem. Res. 2021, 54, 1251−1259