PORTLAND PRESS

Review Article

Fuzzy protein theory for disordered proteins

Monika Fuxreiter[1,2]

[1]Laboratory of Protein Dynamics, University of Debrecen, Hungary; [2]Department of Biomedical Sciences, University of Padova, Italy

Correspondence: Monika Fuxreiter (monika.fuxreiter@unipd.it)

OPEN ACCESS

Why proteins are fuzzy? Constant adaptation to the cellular environment requires a wide range of changes in protein structure and interactions. Conformational ensembles of disordered proteins in particular exhibit large shifts to activate or inhibit alternative pathways. Fuzziness is critical for liquid–liquid phase separation and conversion of biomolecular condensates into fibrils. Interpretation of these phenomena presents a challenge for the classical structure-function paradigm. Here I discuss a multi-valued formalism, based on fuzzy logic, which can be applied to describe complex cellular behavior of proteins.

Introduction

Traditionally, we consider a statement either true or false (0 or 1, Boolean logic). In our everyday life, however, we use more complex descriptions. Somebody is not simply tall or short, but tall to a certain degree (e.g. relatively tall, tall for the age, tallest in the team) usually defined on a given scale. These statements conform to a many-valued or fuzzy logic [1], as we assign a degree to the truth value. This means that more values can be true, but to different degrees (e.g. somebody can be tall and short, depending on the person of comparison). Similarly, the organization of a protein chain is along the continuum between an ordered state dominated by one conformation, and a disordered state generated by multiple conformations [2]. Different protein regions may have different degrees of preference for ordered or disordered states [3].

Most proteins are capable to perform multiple activities under physiological conditions [4–9]. In the stochastic cellular environment, the different protein functions are realized to different extents depending on the intra- or extracellular signals, localization or milieu. The classical one gene → one structure → one function paradigm, which is based on binary logic, is challenged to interpret this phenomenon. The simultaneous occurrence of multiple activities, however, can be described based on fuzzy logic. In artificially intelligent devices a fuzzy logic-based control system is used to adapt to non-precisely defined or unforeseen conditions, such as the cellular world for proteins. Here I propose a structure-function framework based on multi-valued, fuzzy logic and illustrate this approach using interactions of disordered proteins [10].

A binding mode continuum between order and disorder

The wide range of changes in structure and dynamics upon interactions of disordered proteins result in a continuum of bound states between order and disorder [11,12]. Based on the change in conformational entropy upon binding we can classify different 'binding modes', which represent different degrees of conformational heterogeneity of the bound state ensemble. Ordered binding modes are achieved via disorder-to-order transitions resulting in well-defined contact patterns between specific residues [13,14]. Disordered binding modes involve many binding configurations, which are resulted by disorder-to-disorder transitions [15–17]. In these cases, interactions in the bound state ensemble are mediated by redundant motifs.

Ordered and disordered binding modes have different sequence codes [18]. Binding motifs mediating ordered interactions exhibit a distinct composition as compared with their flanking regions. This is exemplified by short hydrophobic motifs within disordered regions [19]. In contrast, the interacting

elements mediating disordered binding modes have a similar composition to their flanking regions (e.g. tandem motifs or low-complexity sequences) [20,21]. Based on these local compositional biases, the probabilities of transitions towards ordered or disordered bound states (p_{DO} and p_{DD}) can be predicted from the sequence. The method has been published [18], therefore only the key points and equations are summarized below. Predictions were validated using ∼2000 protein complexes, by comparing the unbound and bound states of the same protein regions [18]. Using this approach, ∼40% of disordered proteins in yeast were predicted to fold upon partner interactions, in good agreement with proteomic results *in vivo* [22].

Recognizing that the bound state of a protein can be to some extent ordered and to some extent disordered and simultaneous quantification of these two properties conforms to fuzzy logic.

Fuzzy binding of disordered proteins

Fuzzy binding means that the interaction behavior of a protein region changes according to the cellular context. In particular, binding to different partners or changes in cellular conditions (localization, posttranslational modification) may result in different bound state ensembles with different degrees of conformational heterogeneity, even if the same region is involved in the interaction. For example, binding to E-cadherin induces the ordering of the 134–161 residues in β-catenin (PDB:1i7w; [23]), while this region remains disordered in the complex with ICAT β-catenin interacting protein (PDB: 1m1e [24]). Ordered binding modes can be induced by posttranslational modifications, exemplified by the phosphorylation of the N-terminal autoinhibitory region of glycogen-synthase kinase 3 in the insulin pathway (GSK-3, PDB: 4nm3 [25]), which region remains disordered when bound to the Wnt receptor LRP6 motif in the Wnt pathway (PDB: 4nm5 [25]). Phosphorylation in the bound complex may also induce a conversion between disordered and ordered binding modes, which may give rise to oligomerisation or higher-order assembly [26,27].

Increasing experimental evidence shows that context-dependent variation of binding modes can activate different cellular pathways. Thus, fuzzy binding introduces uncertainty into predicting function, as the same sequence encodes different bound state ensembles, with different biological outcomes. Indeed, fuzzy binding is exploited to regulate signaling specificity [7,28–30].

Fuzzy binding is represented by binding mode landscapes

Context-dependent interaction behavior of proteins requires novel computational approaches to predict fuzzy binding from sequence.

The degree of 'fuzziness' can be defined by binding mode entropy (S_{bind}) [31], which quantifies the change in bound-state conformational heterogeneity with different partners and conditions. This is derived from the distributions of binding entropies $[\pi_{DD}(R_i)]_N$ with different interaction partners and plausible binding sites, as illustrated in Figure 1A for four interaction sites of the p53 tumor suppressor.

The x-axis shows the $\pi_{DD}(R_i)$ values, which quantify the probability of disorder-to-disorder transitions ranging from 0 (complete ordering upon binding) to 1 (complete disordering upon binding) (they are computed as detailed in ref [18]). The distribution corresponds to distinct binding events under different conditions (partner, cellular milieu, posttranslational modification), which may result in different binding entropies. The y-axis shows the $\pi_{DD}(R_i)$ frequencies in 10 bins of the [0,1] scale:

$$f[\pi_{DD}(R_i)] = \frac{n_R[\pi_{DD}(R_i)]}{N}$$

where n_R is the number of binding regions within a given bin.

Fuzzy binding can be described by binding mode landscapes, which are derived from the distributions of binding entropy associated with distinct binding events [31]. The x-axis quantifies the conformational heterogeneity in the bound state, reflecting the most likely mode of binding. The y-axis quantifies the binding mode entropy (S_{bind}), reflecting changes in conformational heterogeneity with different partners. The methods to derive these quantities have been published [31], here only the key points are reviewed.

The x-axis shows the most likely mode of binding (p_{DD}), which is computed from the $[\pi_{DD}(R_i)]_N$ distributions displayed in Figure 1A:
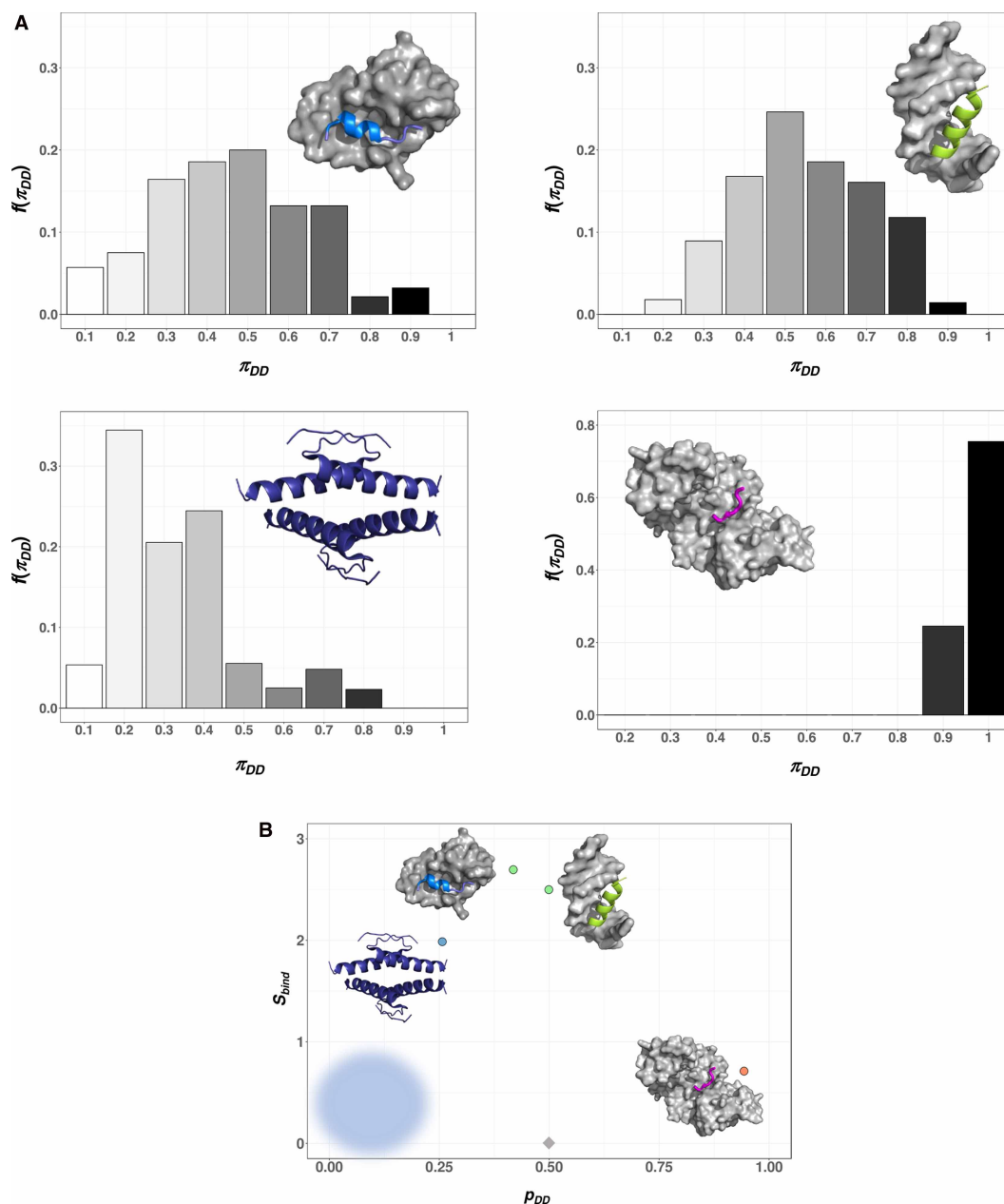
$$p_{DD}(A_i) = \text{median}[\pi_{DD}(R_i)]_N$$

Figure 1. Representation of the binding mode landscape concept.   Part 1 of 2

(**A**) The interaction behavior of four different regions of the p53 tumor suppressor: top left is the mdm2 binding region (residues 19–26, PDB:1ycr, *blue*); top right is the DNA recognition helix (residues 278–285, PDB: 2ady, *lime*); bottom left is the oligomerisation domain (residues 325–345, PDB: 1c26, *dark blue*); bottom right is the C-terminal peptide bound to sirtuin (residues 378–386, PDB: 4zzj, *magenta*). The partners are shown by grey surfaces. The interaction sites, which bind to mdm2 and DNA (top panels) have broad $[\pi_{DD}(R_i)]_N$ distributions, indicating a wide range of conformational heterogeneity in the bound state. In contrast, the p53 C-terminal peptide exhibits a narrow $[\pi_{DD}(R_i)]_N$ distribution, indicating that it mostly remains disordered (bottom right), when bound to different partners. The oligomerisation domain (bottom left) shows high frequency ordered binding modes, but can also visit more disordered bound configurations. (**B**) Representation of the four p53 regions on a binding mode landscape. The C-terminal peptide of p53 (*magenta*), which is predicted to remain conformationally heterogeneous in its complexes ($p_{DD} = 0.9$), has the lowest variation in binding modes ($S_{bind} = 0.7$). In contrast, the mdm2 binding helix (lime) may exhibit both ordered and disordered binding configurations ($p_{DD} = 0.4$), indicating a wide range of conformational heterogeneity in the bound state ensemble with different partners ($S_{bind} = 2.7$). The DNA recognition helix (lime) may also exhibit high binding mode entropy indicating context-dependence ($S_{bind} = 2.5$). The oligomerisation domain (blue) has

Figure 1. Representation of the binding mode landscape concept.

a considerably larger probability for ordering upon binding (p_{DD} = 0.3, S_{bind} = 2.0). The blue area on the bottom left of the figure represents disorder-to-order binding of disordered regions by conformational selection and induced fit mechanisms. These are weakly context-dependent binding modes and have low S_{bind} values. The grey diamond (p_{DD} = 0.5, S_{bind} = 0) represents a hypothetical 'lock-and-key' mechanism for disordered regions, when the ensemble does not change upon binding (no change in conformational entropy) and is not influenced by the context. This scenario however, is not realized in Nature.

The y-axis is defined by the Shannon-entropy computed from the frequencies of different $\pi_{DD}(R_i)$ values:

$$S_{bind}(A_i) = \sum f[\pi_{DD}(R_i)] \log_2 f[\pi_{DD}(R_i)]$$

where the summation goes over the number of bins used to compute the frequencies of the $\pi_{DD}(R_i)$ values.

The binding mode landscape concept is illustrated in Figure 1B for the p53 interaction sites. The C-terminal region of p53 (residues 378–386) for example, has a low S_{bind} value, as it tends to bind via short motifs to sirtuins or Cdk2/cyclin-A with a large degree of conformational heterogeneity [32,33]. In contrast, the N-terminal transactivation region of p53, which serves as a network hub, has a high S_{bind} value, reflecting selective interactions with Mdm2 (residues 19–26) [34], HMGB [35] or CBP/p300 [36] and many other partners via both ordered and disordered binding modes. The recognition helix (residues 278–285) has a similarly high S_{bind} value, reflecting conditional folding upon DNA interactions. The tetramerisation domain (residues 325–345) of p53 is biased towards ordered binding modes [37] and in accord has a lower S_{bind} value (Figure 1B).

Taken together, the binding mode entropy is derived from different, hypothetical binding events and is computed over bound-states ensembles with different partners. In contrast, the binding entropy is related to the conformational heterogeneity of the bound-state ensemble in a single binding event. Representing the binding entropy and binding mode entropy on the binding mode landscape characterizes fuzzy binding.

Fuzzy binding in biomolecular condensates: from liquids to fibrils

Biomolecular condensates formed by liquid–liquid phase separation open new frontiers in our understanding of cellular processes [38,39]. The molecular driving forces, which organize these assemblies however, have not been fully elucidated, owing to the complex stoichiometry and the wide variety of sequence patterns in these membraneless organelles [40,41]. In general, weak interactions amongst redundant binding motifs is considered as the organizing principle of these higher-order assemblies [42].

Within the framework of fuzzy binding, biomolecular condensates exhibit disordered binding modes with many different bound configurations mediated by low-complexity motifs [43]. Such conformational heterogeneity (generated by disorder-to-disorder transitions) is required for driving droplet formation. Thus, protein regions capable of forming disordered binding modes can spontaneously phase separate and serve as scaffolds for droplets [44]. Other components of condensates, termed as clients, require a partner-induced change in their binding modes. These proteins exhibit fuzzy binding and shift from ordered to disordered bound states in a context-dependent manner [27].

Aging or familial mutations can induce the conversion of liquid droplets to fibrils, which are implicated in neurodegenerative diseases [45,46]. Phase separation may facilitate aggregation [47,48], but their relationship has not been fully elucidated. Fibrillization of liquid droplets takes place via a gradual change in binding modes from disordered to ordered states [45,46]. Some mutations, such as glycine mutations in Fus reduce fluidity and thus shift the system towards more ordered binding modes [49]. Other mutations affect interactions with other biomolecules, in particular with RNA, which induce changes from disordered to ordered bound states. According to fuzzy logic, disordered and ordered binding modes co-exist, and their gradual shift drives the system towards the amyloid state.

Why fuzzy?

The concept of fuzziness reflects a lack of precision, which we aim to avoid in scientific studies. On the other hand, the fuzzy set theory provides a quantitative approach to handle uncertainty. Cellular behavior of proteins is stochastic, with a high level of noise and is driven by many, often conflicting signals. Proteins have evolved to perform multiple activities [4–6,50] for adapting to the cellular context of a network of proteins [51–53]. It

is beyond the binary logic of the classical structure-function paradigm to describe how proteins can simultaneously carry out multiple functions.

Let's consider the problem of biomolecular recognition by dynamic conformational ensembles [54]. The classical lock-and-key mechanism, when the binding does not influence the conformational ensemble would not operate here. Instead, the bound state of dynamic conformational ensembles can be achieved via conformational selection or induced fit mechanisms depending on the protein concentration [55]. Both scenarios are coupled to disorder-to-order transitions and result in a well-defined bound state, which weakly depends on the context (low p_{DD} and S_{bind}, Figure 1B). Such a well-defined bound-state can be described by a deterministic one structure → one function model, which conforms to binary logic. In contrast, the fuzzy framework can describe dynamic conformational ensembles with multiple activities by establishing one structure → multiple functions relationships [56]. In addition, the fuzzy model can explain considerable conformational shifts in the ensemble with a minor impact on the original function, leading to multiple structures → one function relationships [57,58]. These functional redundancies are in accord with the messiness of evolutionary innovations [59].

Outlook: describing complex cellular behavior

The cellular world is very complex, where proteins need to perceive their environments to maximize their chance of successfully achieving their biological roles. To solve similar adaptation problems, artificially intelligent devices often use a fuzzy control system, which exploits multiple, simultaneous activities.

Structural studies demonstrate the conformational heterogeneity in cellular systems. Description of these ensembles [60,61] under a variety of conditions is a key to establish structure-function relationships using a multi-valued (fuzzy) formalism. Another essential component is to characterize the activities performed by the different ensembles [62] and to what extent these activities are modulated by shifting the populations of the different conformational sub-states. As in most cases this information is not available, this can be a bottleneck for developing many-valued structure-function models.

The interaction partners of disordered proteins, for example, are often not known, in particular under cellular conditions. However, as also illustrated in this paper, context-dependent interaction behavior can be predicted from the sequence. Binding mode landscapes describe fuzzy binding using the (i) the degree of conformational heterogeneity of the bound state ensembles ('binding mode'), and (ii) the binding mode entropy. Entropy is usually a measure of uncertainty. Binding mode entropy is a measure of binding mode uncertainty. We need to accept that the protein sequence also encodes multiple interaction behaviors. Recognizing such uncertainty will contribute to understanding cellular activities of proteins. In particular, how the same sequence contributes to different signaling pathways, even with opposite biological outcomes.

Application of the fuzzy structure-function model thus opens perspectives to identify context-dependent regulatory motifs in protein networks. This approach can also be used to describe the regulated assembly/disassembly of membraneless organelles [44,63]. Fuzzy structure-function relationships can be integrated into a fuzzy inference system [56], which can model cellular networks with different biological outcomes, similarly to the control system of artificially intelligent devices. Overall, we do not need to abandon the classical structure-function paradigm to understand cellular behavior, just expand it to a multi-valued formalism.

Perspectives

- Predicting complex interaction behavior of disordered proteins

- Modeling cellular networks including context-dependence of individual proteins

- Deriving sequence codes for biomolecular condensates based on fuzzy binding

- Modeling the conversion of droplets to amyloids, and identifying pathological mutations

- Predicting cellular consequences of changing interaction modes

- Development a general structure-function paradigm for proteins

Competing Interests

The author declares that there are no competing interests associated with this manuscript.

Abbreviations

GSK-3, glycogen-synthase kinase 3.

References

1　Zadeh, L.A. (1965) Fuzzy sets. *Inform. Control.* **8**, 338–353 https://doi.org/10.1016/S0019-9958(65)90241-X

2　Sormanni, P., Piovesan, D., Heller, G.T., Bonomi, M., Kukic, P., Camilloni, C. et al. (2017) Simultaneous quantification of protein order and disorder. *Nat. Chem. Biol.* **13**, 339–342 https://doi.org/10.1038/nchembio.2331

3　van der Lee, R., Buljan, M., Lang, B., Weatheritt, R.J., Daughdrill, G.W., Dunker, A.K. et al. (2014) Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **114**, 6589–6631 https://doi.org/10.1021/cr400525m

4　Nobeli, I., Ponstingl, H., Krissinel, E.B. and Thornton, J.M. (2003) A structure-based anatomy of the *E. coli* metabolome. *J. Mol. Biol.* **334**, 697–719 https://doi.org/10.1016/j.jmb.2003.10.008

5　Nam, H., Lewis, N.E., Lerman, J.A., Lee, D.H., Chang, R.L., Kim, D. et al. (2012) Network context and selection in the evolution to enzyme specificity. *Science* **337**, 1101–1104 https://doi.org/10.1126/science.1216861

6　Korcsmaros, T., Farkas, I.J., Szalay, M.S., Rovo, P., Fazekas, D., Spiro, Z. et al. (2010) Uniformly curated signaling pathways reveal tissue-specific cross-talks and support drug target discovery. *Bioinformatics* **26**, 2042–2050 https://doi.org/10.1093/bioinformatics/btq310

7　Manglik, A., Kim, T.H., Masureel, M., Altenbach, C., Yang, Z., Hilger, D. et al. (2015) Structural insights into the dynamic process of beta2-adrenergic receptor signaling. *Cell* **161**, 1101–1111 https://doi.org/10.1016/j.cell.2015.04.043

8　Bah, A., Vernon, R.M., Siddiqui, Z., Krzeminski, M., Muhandiram, R., Zhao, C. et al. (2015) Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature* **519**, 106–109 https://doi.org/10.1038/nature13999

9　Ciuffa, R., Lamark, T., Tarafder, A.K., Guesdon, A., Rybina, S., Hagen, W.J. et al. (2015) The selective autophagy receptor p62 forms a flexible filamentous helical scaffold. *Cell Rep.* **11**, 748–758 https://doi.org/10.1016/j.celrep.2015.03.062

10　Fuxreiter, M. (2018) Fuzziness in protein interactions-a historical perspective. *J. Mol. Biol.* **430**, 2278–2287 https://doi.org/10.1016/j.jmb.2018.02.015

11　Miskei, M., Antal, C. and Fuxreiter, M. (2017) FuzDB: database of fuzzy complexes, a tool to develop stochastic structure-function relationships for protein complexes and higher-order assemblies. *Nucleic Acids Res.* **45**, D228–DD35 https://doi.org/10.1093/nar/gkw1019

12　Fuxreiter, M. (2018) Fold or not to fold upon binding: does it really matter? *Curr. Opin. Struct. Biol.* **54**, 19–25 https://doi.org/10.1016/j.sbi.2018.09.008

13　Borcherds, W., Theillet, F.X., Katzer, A., Finzel, A., Mishall, K.M., Powell, A.T. et al. (2014) Disorder and residual helicity alter p53-Mdm2 binding affinity and signaling in cells. *Nat. Chem. Biol.* **10**, 1000–1002 https://doi.org/10.1038/nchembio.1668

14　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 https://doi.org/10.1074/jbc.R115.692715

15　Kurzbach, D., Schwarz, T.C., Platzer, G., Hofler, S., Hinderberger, D. and Konrat, R. (2014) Compensatory adaptations of structural dynamics in an intrinsically disordered protein complex. *Angew. Chem. Int. Ed. Engl.* **53**, 3840–3843 https://doi.org/10.1002/anie.201308389

16　Borgia, A., Borgia, M.B., Bugge, K., Kissling, V.M., Heidarsson, P.O., Fernandes, C.B. et al. (2018) Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **555**, 61–66 https://doi.org/10.1038/nature25762

17　Lee, S., Kim, J., Han, S. and Park, C.J. (2020) Recognition and unfolding of c-MYC and telomeric G-quadruplex DNAs by the recQ C-terminal domain of human bloom syndrome helicase. *ACS Omega* **5**, 14513–14522 https://doi.org/10.1021/acsomega.0c01176

18　Miskei, M., Horvath, A., Vendruscolo, M. and Fuxreiter, M. (2020) Sequence-based prediction of fuzzy protein interactions. *J. Mol. Biol.* **432**, 2289–2303 https://doi.org/10.1016/j.jmb.2020.02.017

19　Warfield, L., Tuttle, L.M., Pacheco, D., Klevit, R.E. and Hahn, S. (2014) A sequence-specific transcription activator motif and powerful synthetic variants that bind mediator using a fuzzy protein interface. *Proc. Natl Acad. Sci. U.S.A.* **111**, E3506–E3513 https://doi.org/10.1073/pnas.1412088111

20　Helander, S., Montecchio, M., Pilstal, R., Su, Y., Kuruvilla, J., Elven, M. et al. (2015) Pre-anchoring of Pin1 to unphosphorylated c-Myc in a fuzzy complex regulates c-Myc activity. *Structure* **23**, 2267–2279 https://doi.org/10.1016/j.str.2015.10.010

21　Hadzi, S., Mernik, A., Podlipnik, C., Loris, R. and Lah, J. (2017) The thermodynamic basis of the fuzzy interaction of an intrinsically disordered protein. *Angew. Chem. Int. Ed. Engl.* **56**, 14494–7 https://doi.org/10.1002/anie.201707853

22　Leuenberger, P., Ganscha, S., Kahraman, A., Cappelletti, V., Boersema, P.J., von Mering, C. et al. (2017) Cell-wide analysis of protein thermal unfolding reveals determinants of thermostability. *Science* **355**, eaai7825 https://doi.org/10.1126/science.aai7825

23　Huber, A.H. and Weis, W.I. (2001) The structure of the beta-catenin/E-cadherin complex and the molecular basis of diverse ligand recognition by beta-catenin. *Cell* **105**, 391–402 https://doi.org/10.1016/S0092-8674(01)00330-0

24　Daniels, D.L. and Weis, W.I. (2002) ICAT inhibits beta-catenin binding to Tcf/Lef-family transcription factors and the general coactivator p300 using independent structural modules. *Mol. Cell* **10**, 573–584 https://doi.org/10.1016/S1097-2765(02)00631-7

25　Stamos, J.L., Chu, M.L., Enos, M.D., Shah, N. and Weis, W.I. (2014) Structural basis of GSK-3 inhibition by N-terminal phosphorylation and by the Wnt receptor LRP6. *eLife* **3**, e01998 https://doi.org/10.7554/eLife.01998

26　Korennykh, A.V., Egea, P.F., Korostelev, A.A., Finer-Moore, J., Zhang, C., Shokat, K.M. et al. (2009) The unfolded protein response signals through high-order assembly of Ire1. *Nature* **457**, 687–693 https://doi.org/10.1038/nature07661

27　Wippich, F., Bodenmiller, B., Trajkovska, M.G., Wanka, S., Aebersold, R. and Pelkmans, L. (2013) Dual specificity kinase DYRK3 couples stress granule condensation/dissolution to mTORC1 signaling. *Cell* **152**, 791–805 https://doi.org/10.1016/j.cell.2013.01.033

28 Delaforge, E., Kragelj, J., Tengo, L., Palencia, A., Milles, S., Bouvignies, G. et al. (2018) Deciphering the dynamic interaction profile of an intrinsically disordered protein by NMR exchange spectroscopy. *J. Am. Chem. Soc.* **140**, 1148–1158 https://doi.org/10.1021/jacs.7b12407

29 Miskei, M., Gregus, A., Sharma, R., Duro, N., Zsolyomi, F. and Fuxreiter, M. (2017) Fuzziness enables context dependence of protein interactions. *FEBS Lett.* **591**, 2682–2695 https://doi.org/10.1002/1873-3468.12762

30 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 https://doi.org/10.1074/jbc.RA118.005854

31 Horvath, A., Miskei, M., Ambrus, V., Vendruscolo, M. and Fuxreiter, M. (2020) Sequence-based prediction of protein binding mode landscapes. *PLoS Comput. Biol.* **16**, e1007864 https://doi.org/10.1371/journal.pcbi.1007864

32 Dai, H., Case, A.W., Riera, T.V., Considine, T., Lee, J.E., Hamuro, Y. et al. (2015) Crystallographic structure of a small molecule SIRT1 activator-enzyme complex. *Nat. Commun.* **6**, 7645 https://doi.org/10.1038/ncomms8645

33 Lowe, E.D., Tews, I., Cheng, K.Y., Brown, N.R., Gul, S., Noble, M.E. et al. (2002) Specificity determinants of recruitment peptides bound to phospho-CDK2/cyclin A. *Biochemistry* **41**, 15625–15634 https://doi.org/10.1021/bi0268910

34 Kussie, P.H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A.J. et al. (1996) Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science* **274**, 948–953 https://doi.org/10.1126/science.274.5289.948

35 Rowell, J.P., Simpson, K.L., Stott, K., Watson, M. and Thomas, J.O. (2012) HMGB1-facilitated p53 DNA binding occurs via HMG-Box/p53 transactivation domain interaction, regulated by the acidic tail. *Structure* **20**, 2014–2024 https://doi.org/10.1016/j.str.2012.09.004

36 Lee, C.W., Ferreon, J.C., Ferreon, A.C., Arai, M. and Wright, P.E. (2010) Graded enhancement of p53 binding to CREB-binding protein (CBP) by multisite phosphorylation. *Proc. Natl Acad. Sci. U.S.A.* **107**, 19290–5 https://doi.org/10.1073/pnas.1013078107

37 Kitayner, M., Rozenberg, H., Kessler, N., Rabinovich, D., Shaulov, L., Haran, T.E. et al. (2006) Structural basis of DNA recognition by p53 tetramers. *Mol. Cell* **22**, 741–753 https://doi.org/10.1016/j.molcel.2006.05.015

38 Banani, S.F., Lee, H.O., Hyman, A.A. and Rosen, M.K. (2017) Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell. Biol.* **18**, 285–298 https://doi.org/10.1038/nrm.2017.7

39 Boeynaems, S., Alberti, S., Fawzi, N.L., Mittag, T., Polymenidou, M., Rousseau, F. et al. (2018) Protein phase separation: a New phase in cell biology. *Trends Cell Biol.* **28**, 420–435 https://doi.org/10.1016/j.tcb.2018.02.004

40 Wang, J., Choi, J.M., Holehouse, A.S., Lee, H.O., Zhang, X., Jahnel, M. et al. (2018) A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell* **174**, 688–99 e16 https://doi.org/10.1016/j.cell.2018.06.006

41 Vernon, R.M., Chong, P.A., Tsang, B., Kim, T.H., Bah, A., Farber, P. et al. (2018) Pi-Pi contacts are an overlooked protein feature relevant to phase separation. *eLife* **7**, e31486 https://doi.org/10.7554/eLife.31486

42 Li, P., Banjade, S., Cheng, H.C., Kim, S., Chen, B., Guo, L. et al. (2012) Phase transitions in the assembly of multivalent signalling proteins. *Nature* **483**, 336–340 https://doi.org/10.1038/nature10879

43 Burke, K.A., Janke, A.M., Rhine, C.L. and Fawzi, N.L. (2015) Residue-by-residue view of in vitro FUS granules that bind the C-Terminal domain of RNA polymerase II. *Mol. Cell* **60**, 231–241 https://doi.org/10.1016/j.molcel.2015.09.006

44 Hardenberg, M., Horvath, A., Ambrus, V., Fuxreiter, M. and Vendruscolo, M. (2020) Widespread occurrence of the droplet state of proteins in the human proteome. *bioRxiv* https://doi.org/10.1101/2020.10.21.348532

45 Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A.P., Kim, H.J. et al. (2015) Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123–133 https://doi.org/10.1016/j.cell.2015.09.015

46 Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y. et al. (2015) A liquid-to-solid phase transition of the ALS protein FUS accelerated by disease mutation. *Cell* **162**, 1066–1077 https://doi.org/10.1016/j.cell.2015.07.047

47 Wegmann, S., Eftekharzadeh, B., Tepper, K., Zoltowska, K.M., Bennett, R.E., Dujardin, S. et al. (2018) Tau protein liquid-liquid phase separation can initiate tau aggregation. *EMBO J.* **37**, e98049 https://doi.org/10.15252/embj.201798049

48 Ray, S., Singh, N., Kumar, R., Patel, K., Pandey, S., Datta, D. et al. (2020) alpha-Synuclein aggregation nucleates through liquid-liquid phase separation. *Nat. Chem.* **12**, 705–716 https://doi.org/10.1038/s41557-020-0465-9

49 Niaki, A.G., Sarkar, J., Cai, X., Rhine, K., Vidaurre, V., Guy, B. et al. (2020) Loss of dynamic RNA interaction and aberrant phase separation induced by two distinct types of ALS/FTD-linked FUS mutations. *Mol. Cell* **77**, 82–94 e4 https://doi.org/10.1016/j.molcel.2019.09.022

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

51 Bensimon, A., Heck, A.J. and Aebersold, R. (2012) Mass spectrometry-based proteomics and network biology. *Annu. Rev. Biochem.* **81**, 379–405 https://doi.org/10.1146/annurev-biochem-072909-100424

52 Li, A., Cornelius, S.P., Liu, Y.Y., Wang, L. and Barabasi, A.L. (2017) The fundamental advantages of temporal networks. *Science* **358**, 1042–1046 https://doi.org/10.1126/science.aai7488

53 Vidal, M., Cusick, M.E. and Barabasi, A.L. (2011) Interactome networks and human disease. *Cell* **144**, 986–998 https://doi.org/10.1016/j.cell.2011.02.016

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

55 Hammes, G.G., Chang, Y.C. and Oas, T.G. (2009) Conformational selection or induced fit: a flux description of reaction mechanism. *Proc. Natl Acad. Sci. U.S.A.* **106**, 13737–13741 https://doi.org/10.1073/pnas.0907195106

56 Fuxreiter, M. (2018) Towards a stochastic paradigm: from fuzzy ensembles to cellular functions. *Molecules* **23**, 3008 https://doi.org/10.3390/molecules23113008

57 Hong, N.S., Petrovic, D., Lee, R., Gryn'ova, G., Purg, M., Saunders, J. et al. (2018) The evolution of multiple active site configurations in a designed enzyme. *Nat. Commun.* **9**, 3900 https://doi.org/10.1038/s41467-018-06305-y

58 Piazza, I., Kochanowski, K., Cappelletti, V., Fuhrer, T., Noor, E., Sauer, U. et al. (2018) A Map of protein-metabolite interactions reveals principles of chemical communication. *Cell* **172**, 358–72 e23 https://doi.org/10.1016/j.cell.2017.12.006

59 Tawfik, D.S. (2010) Messy biology and the origins of evolutionary innovations. *Nat. Chem. Biol.* **6**, 692–696 https://doi.org/10.1038/nchembio.441

60 Baldwin, A.J. and Kay, L.E. (2009) NMR spectroscopy brings invisible protein states into focus. *Nat. Chem. Biol.* **5**, 808–814 https://doi.org/10.1038/nchembio.238

61   Bonomi, M. and Vendruscolo, M. (2019) Determination of protein structural ensembles using cryo-electron microscopy. *Curr. Opin. Struct. Biol.* **56**, 37–45 https://doi.org/10.1016/j.sbi.2018.10.006

62   Nagulapalli, M., Parigi, G., Yuan, J., Gsponer, J., Deraos, G., Bamm, V.V. et al. (2012) Recognition pliability is coupled to structural heterogeneity: a calmodulin intrinsically disordered binding region complex. *Structure* **20**, 522–533 https://doi.org/10.1016/j.str.2012.01.021

63   Tuu-Szabo, B., Koczy, L. and Fuxreiter, M. (2018) Simulations of higher-Order protein organizations using a fuzzy framework. *Complexity* **2018**, 6360846 https://doi.org/10.1155/2018/6360846