MDPI

Perspective

Towards a Stochastic Paradigm: From Fuzzy Ensembles to Cellular Functions

Monika Fuxreiter *

MTA-DE Laboratory of Protein Dynamics, Department of Biochemistry and Molecular Biology,
H-4032 Debrecen, Hungary

check for
updates

Abstract: The deterministic sequence → structure → function relationship is not applicable to describe how proteins dynamically adapt to different cellular conditions. A stochastic model is required to capture functional promiscuity, redundant sequence motifs, dynamic interactions, or conformational heterogeneity, which facilitate the decision-making in regulatory processes, ranging from enzymes to membraneless cellular compartments. The fuzzy set theory offers a quantitative framework to address these problems. The fuzzy formalism allows the simultaneous involvement of proteins in multiple activities, the degree of which is given by the corresponding memberships. Adaptation is described via a fuzzy inference system, which relates heterogeneous conformational ensembles to different biological activities. Sequence redundancies (e.g., tandem motifs) can also be treated by fuzzy sets to characterize structural transitions affecting the heterogeneous interaction patterns (e.g., pathological fibrillization of stress granules). The proposed framework can provide quantitative protein models, under stochastic cellular conditions.

1. The Structure-Function Paradigm

Protein functions take place in space and time. Structure-function principles, however, relate a protein sequence to biological activity, only via the spatial coordinates of the residues [1,2]:

$$\text{SEQUENCE} \rightarrow \text{STRUCTURE} \rightarrow \text{FUNCTION}$$
$$(x,y,z) \qquad (x,y,z,t) \tag{1}$$

The three-dimensional organization of amino acids brings different chemical groups into proximity [3,4], creating specific microenvironments for biological activities. The emerging active sites, for example, can catalyze chemical reactions at significantly faster rates, than in solution [5,6]. The classical, deterministic Paradigm 1 establishes an unambiguous connection between the protein sequence and its function, via a unique structure.

2. The Ensemble View

The energy landscapes of proteins are, in reality, more complicated. Proteins fluctuate among various conformations ('macrostates') and sub-states ('microstates'), which need to be considered for their relevant functioning [7–9]. A wide spectrum of dynamical transitions [10,11]—from local movements (e.g., sidechain rotations [12]) to large-amplitude collective motions (e.g., domain repositioning [13])—generates conformational ensembles, which, however, are not trivial to link to the function. How can structure-function relationships account for protein dynamics? If protein

structure is described as an ensemble, the populations of the relevant sub-states, as well as the rate of interconversion between them, must be experimentally determined for each biological activity.

$$\text{SEQUENCE} \rightarrow \text{CONFORMATIONAL ENSEMBLE} \rightarrow \text{FUNCTION}$$

$$\{p_{cs_1}, \dots, p_{cs_N}\} \qquad\qquad (x,y,z,t) \qquad\qquad (2)$$
$$\{k_{cs_1}, \dots, k_{cs_N}\}$$

where p_{CS_i} is the probability of the given conformational sub-state CS_i, N is the number of sub-states, and $\{k_{CS_i}\}$ is the set of rates, corresponding to the conversions between $CS_i \rightarrow CS_j$, where j corresponds to all the other sub-states. Even if the number of conformational states is reduced to a few functionally relevant ones, characterizing both their thermodynamic and kinetic properties is a daunting task [14,15]. Furthermore, the deterministic relationship between the ensemble parameters and a unique function is also influenced by the environmental conditions.

3. Adaptation to Stochastic Cellular Conditions

Proteins function under rapidly changing extracellular signals and intracellular milieu, which is shaped by cellular diffusion and transport, stochastic gene expression, degradation, and other environmental fluctuations. These factors present stochastic conditions for protein evolution [16–18] leading to 'noise' in biological innovations [19], which is reflected by redundancies and ambiguities in sequences [20], structures [21], and functions [22]. On the one hand, proteins attempt to minimize functional noise. For example, higher-order structures emerge to reduce noise-to-signal ratio for low-affinity substrates [23–25]. On the other hand, ambiguities and redundancies in sequence, structure and function facilitate dynamic adaptation [26]. Proteins evolve under these two opposing constraints to optimize fitness under given cellular conditions.

4. Ambiguity and Redundancy in Sequence, Structure, and Function

The re-formatted paradigm (2), still implies that a given sequence generates a well-defined ensemble, which belongs to a specific function. The stochastic cellular conditions lead to the following observations, which violate the classical paradigm: (i) A considerable proportion of proteins exhibit multiple, simultaneous activities, often referred to as promiscuity or moonlighting [27]. (ii) Certain biological activities (i.e., signaling) are related to heterogeneous conformational ensembles, which are mixtures of different functional ensembles [28]. (iii) Some proteins exhibit a weak sequence dependence, i.e., a large degree of tolerance towards sequence modifications [29]. These observations stem from redundancies in sequence or structure, coupled to ambiguities in function. The same ensemble may perform multiple functions (*functional promiscuity*); the same sequence may be organized into multiple functional ensembles, depending on the context (*conformation and interaction heterogeneity*); and multiple sequences may encode the same conformational ensemble (*sequence redundancy*). These problems, which reflect a more complex relationship between the sequence, structure, and function of proteins, are detailed below.

5. Functional Promiscuity

Metabolic enzymes often catalyze reactions on the non-canonical substrates, some of which are also relevant physiologically [27,30,31]. Functional promiscuity may parallel organism complexity [32], or be driven by network context [33]. Promiscuous activities can serve as starting points to engineer new enzymes [34]. Tailored selection pressures may optimize latent activities to overcome the primary function by $>10^9$-fold [35]. Functional transitions are usually initiated by 'neutral drifts', with a negligible impact on the original activity [36,37]. That is, the optimization of a promiscuous function initially exploits the inherent variations in structure [38] and dynamics [39]. Functional transition of a phosphotriesterase to arylesterase [35], for example, is coupled to increasing structural divergence between the two subunits, until the two activities become comparable (Figure 1A).

In contrast, specialization for the new activity is accompanied by structural convergence (Figure 1A). Similarly, 'freezing' out unnecessary motions offers another route to optimize enzymatic efficiency [6]. Along these lines, principal modes derived from structure [40] often presage or follow the evolutionary changes [41,42].
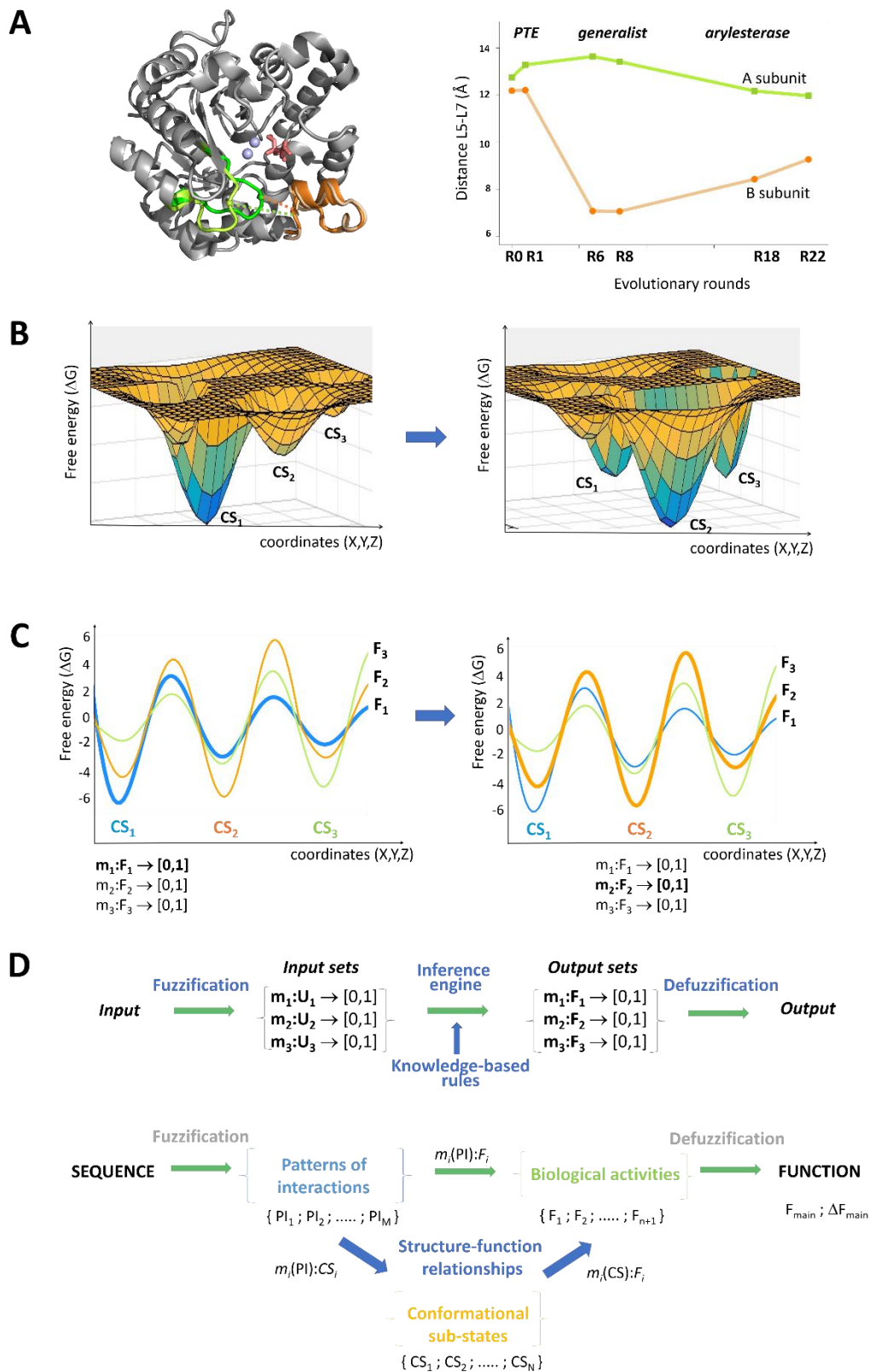
Figure 1. Towards a stochastic structure-function relationship. (**A**) Structural diversity increases with functional promiscuity. The distance between the L5 (*lime*, *green*) and L7 (*wheat*, *orange*) loops (A204 C–G273 C) deviates in the two subunits (*superimposed*) of a dimeric phosphotriesterase (PTE) enzyme (PDB code: 4xag [39]). During laboratory evolution into arylesterase, the structural difference increases as the two activities become comparable (R1 → R6), while it decreases during specialization (R8 → R22). (**B**) Free energy landscape changes upon adaptation of proteins. Functional alterations shift the relative populations of conformational sub-states, but may not impact the ruggedness of the landscape. (**C**) Conformational sub-states (CSs) contribute to multiple free landscapes. The functional noise (uncertainty of F_1, F_2, F_3) of the main activity (*bold*) can be quantified by fuzzy membership functions. (**D**) The fuzzy structure-function model. In the fuzzy inference system, the logical relationship is established between the fuzzy sets of the input and output (*top*). In proteins, fuzzification generates sets of interaction patterns amongst functional sequence motifs, which can be linked to conformational sub-states. The connection between structure and function is a knowledge-based logical rule between the set of conformational sub-states and the set of alternative functions, from which the most likely activity can be selected (*bottom*).

6. Conformational Heterogeneity

Dynamic signals perturb conformational ensembles by changing the relative populations of the different sub-states [43] (Figure 1B). The co-existence of functionally different conformations, in a broad regime, may enable the same protein to be simultaneously engaged in multiple pathways [44]. An agonist binding to a β_2-adrenergic receptor, for example, does not stabilize the active conformation of the cytoplasmic domain; it rather increases the conformational heterogeneity of the active, intermediate, and inactive states, for the complex signaling outputs [28].

Intriguing observations indicate that specific biomolecular recognition can also be achieved in heterogeneous conformational ensembles [45–47]. Although the underlying molecular forces are often puzzling [48,49], conformational ambiguities often enable context-dependent responses, via alternative interaction patterns [50,51]. Conformational heterogeneity along the binding trajectory, has recently been concluded to critically influence the structures in a complex, with different partners [52,53]. Structural ambiguities might even be a pre-requisite, for example, for efficient transcription [54] via a fuzzy 'free-for-all' mechanism [55].

Conformational heterogeneity often leads to dynamic interaction profiles, where the functional output (specificity, signal, and polymerization) is controlled by transient contacts [56,57]. Dynamic interactions may also balance between the auto-inhibited and active states [58,59] and can be significantly influenced by post-translational modifications (PTMs) [60,61]. Although the modification pattern inducing the functional response can be defined, its impact on the underlying heterogeneous conformational ensembles often remains unclear.

7. Redundant Sequence Motifs

Multiple, weakly-restrained sequence motifs are frequently distinguished in signaling pathways, via mediating protein interactions [62]. Regions linking the motifs exhibit increased conformational plasticity and reduced sensitivity to mutations or scrambling [63], leading to a phenomenon, often referred to as 'sequence independence' [64]. Tandem repeats of a few residues, for example, are often involved in the organization of higher-order structures [65], ranging from amyloids to signaling complexes and nuclear pores [66]. Motif redundancy leads to the redundancy of interaction patterns and the co-existence of different contact topologies. Although the interactions of the individual motifs are often sub-optimal, their cooperativity may result in high-affinity associations [25,67].

Both the dynamics of the motif-linking regions, and the variations in contact patterns, lead to conformational heterogeneity in higher-order assemblies [68]. The Fused in Sarcoma (Fus) protein, for example, is involved in the formation of stress granules, via a liquid–liquid phase transition, which is driven by its low-complexity (LC) domain, composed of 27 [S/G]Y[S/G] repeats. The NMR spectra of the LC domain in the droplet, is similar to that of the unbound state, witnessing

conformational heterogeneity in the assembly [69]. Single-point mutations may gradually decrease conformational heterogeneity, leading to pathological aggregation [70]. Additional studies corroborate the finding that pathological mutations initially induce minor perturbations [71], which simultaneously affect multiple conformations/interaction patterns and induce their shift towards the fibril form.

8. Generalized Structure-Function Ensembles

The experimental data summarized in the above three sections are difficult to interpret via the classical structure-function paradigm (2). We may attempt to solve these problems by treating the sequences, conformations, and functions as generalized ensembles:

$$\text{SEQUENCE } (\mu, \sigma) \rightarrow \text{CONFORMATIONAL ENSEMBLE } (\mu, \sigma) \rightarrow \text{FUNCTION } (\mu, \sigma) \qquad (3)$$

where μ is the mean, and σ is the variance of the respective distribution.

Evaluating the structure-function paradigm in the form (3), requires decoupling of all the respective activities, to analyze the underlying distributions of conformational ensembles and sequences. Careful experimental studies, along these lines [72], demonstrate that these approaches are hardly feasible. First, because the dimensionality of the problem is overwhelming, and second, the deconvolution of different functionalities may not be possible in vivo, owing to the intricate connections.

9. Fuzzy Sets Quantify Sequence and Conformation Ambiguities

I propose that the fuzzy set theory [73] offers a quantitative framework to derive stochastic structure-function relationships. In fuzzy sets $U = \{x_1, x_2, ..., x_N\}$ a membership function $m(x_i) \rightarrow [0, 1]$; $x_i \in U$ is assigned to each element, which characterizes to what extent x_i belongs to the given set. For example, the membership of a protein conformational sub-state (CS_i), in a specific functional set (F_1), can vary between 0 and 1 $(m_1(CS_i) : F_1 \rightarrow [0, 1])$, allowing the conformation to contribute to additional activities (e.g., F_2 and F_3, Figure 1C). Memberships for other possible biological functions could also be defined, using this formalism (Figure 1C). In a similar manner, memberships of sequences in given conformational ensembles, $(m_1(SEQ_i) : CS_1 \rightarrow [0, 1])$, or in given functions $(m_1(SEQ_i) : F_1 \rightarrow [0, 1])$, could also be quantified.

The structure-function paradigm could thus, be reformulated by treating the sequences and conformational ensembles as fuzzy sets:

$$\text{SEQUENCE} \rightarrow \text{CONFORMATIONAL ENSEMBLE} \rightarrow \text{FUNCTION} \qquad (4)$$

$m_i(PI) : CS_i \rightarrow [0, 1] \qquad\qquad m_i(CS) : F_i \rightarrow [0, 1]$
$m_i(PI) : F_i \rightarrow [0, 1]$

where $m_i(PI)$ is the respective membership function of a sequence, defined with respect to the conformational states (CS_i) or biological activity (F_i), as a pattern of interacting elements/motifs (PI). $m_i(CS)$ is the membership function of the conformational sub-state/ensemble (CS), in a given function.

Here sequence, structure, and function are considered as different co-existing distributions (Figure 1C), and their contributions change according to the cellular conditions. For example, in the case of a β_2-adrenergic receptor, the active, intermediate, and inactive states (represented by three ensembles) are mixed differently, depending on the signaling input. The fuzzy formalism handles combinations of activities aiming to determine the individual contributions of the different conformational ensembles.

10. The Stochastic Structure-Function Relationship

Within this framework, the structure-function relationship can be quantified by a fuzzy inference system [74,75] (Figure 1D). Parameters describing the elements of the sequence (motifs) or

conformational space (distinguished secondary structures) are used as the input, and the different biological activities serve as the output of the system. The first step is the fuzzification of the input, when the fuzzy sets and their membership functions are defined to describe the interaction patterns, and the corresponding conformational sub-states (Figure 1D). The fuzzy inputs are then combined and knowledge-based logical rules ('IF-THEN') are applied to obtain the output membership functions of the different biological activities in the system. These rules could be derived using machine-learning or neural network algorithms. Defuzzification of the output can select the most likely activity, under a given condition, while also accounting for other, promiscuous activities (Figure 1D).

The fuzzy model quantifies the functional ambiguities of the conformational sub-states:

$$\Delta F_{main} = \left(\sum_{i}^{n} \delta F_{i,main}\right)/n \tag{5}$$

where n is the number of alternative (promiscuous) activities, and F_{main} is the main function with membership function m_{max}. The contribution of function F_i, with respect to the main function, is computed from the corresponding membership functions: $\delta F_{i,\,main} = m_i / m_{max}$.

Here, the challenge is to define the membership functions. To this end, the efficiencies of the alternative activities (e.g., catalytic rates) are determined via functional assays on well-characterized conformations (e.g., crystal structures, chip-bound proteins, or those selected by conformational antibodies) or ensembles (solution techniques, NMR, FRET, and single-molecule methods). Different membership functions could be probed computationally, based on the regulatory characteristics (e.g., changing an auto-inhibited to an active state).

The fuzzy formalism (4) is particularly useful to relate the sequence sets to function. Here structural features, which could be predicted from the sequence (e.g., secondary structure elements, disordered regions, or post-translational modification sites) may serve to generate the pattern of interaction elements (PI), to define the fuzzy sets. This approach has been implemented in simulations of higher-order protein organizations [76].

11. Conclusion and Outlook

Proteins deal with uncertain information, regarding cellular conditions. The information is not only imprecise, but various components are unknown or are unpredictable, owing to the non-random fluctuations in the system. The functional characteristics of proteins need to be adjusted to this poorly defined environment. The classical models in protein science, such as the structure-function paradigm, are based on well-defined properties and cannot deal with the ambiguities related to "noise". The fuzzy set theory offers a quantitative framework to reformulate the structure-function paradigm for describing the stochastic cellular behavior of proteins (Figure 1D). This approach will provide a more holistic protein model, which can be applied to generate interaction or metabolic networks of different cell lines as well as more reliable pharmacophore models.

Conflicts of Interest: The author declared no conflicts of interest.

References

1. Perutz, M.F. Relation between structure and sequence of haemoglobin. *Nature* **1962**, *194*, 914–917. [CrossRef] [PubMed]
2. Kendrew, J.C. Myoglobin and the structure of proteins. *Science* **1963**, *139*, 1259–1266. [CrossRef] [PubMed]
3. Koshland, D.E., Jr. Properties of the active site of enzymes. *Ann. N Y Acad. Sci.* **1963**, *103*, 630–642. [CrossRef] [PubMed]

4.　Blake, C.C.; Koenig, D.F.; Mair, G.A.; North, A.C.; Phillips, D.C.; Sarma, V.R. Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Angstrom resolution. *Nature* **1965**, *206*, 757–761. [CrossRef] [PubMed]

5.　Warshel, A. Energetics of enzyme catalysis. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 5250–5254. [CrossRef] [PubMed]

6.　Blomberg, R.; Kries, H.; Pinkas, D.M.; Mittl, P.R.; Grutter, M.G.; Privett, H.K.; Mayo, S.L.; Hilvert, D. Precision is essential for efficient catalysis in an evolved Kemp eliminase. *Nature* **2013**, *503*, 418–421. [CrossRef] [PubMed]

7.　Frauenfelder, H.; Sligar, S.G.; Wolynes, P.G. The energy landscapes and motions of proteins. *Science* **1991**, *254*, 1598–1603. [CrossRef] [PubMed]

8.　Artymiuk, P.J.; Blake, C.C.; Grace, D.E.; Oatley, S.J.; Phillips, D.C.; Sternberg, M.J. Crystallographic studies of the dynamic properties of lysozyme. *Nature* **1979**, *280*, 563–568. [CrossRef] [PubMed]

9.　Frauenfelder, H.; Petsko, G.A.; Tsernoglou, D. Temperature-dependent X-ray diffraction as a probe of protein structural dynamics. *Nature* **1979**, *280*, 558–563. [CrossRef] [PubMed]

10.　Henzler-Wildman, K.; Kern, D. Dynamic personalities of proteins. *Nature* **2007**, *450*, 964–972. [CrossRef] [PubMed]

11.　Lewandowski, J.R.; Halse, M.E.; Blackledge, M.; Emsley, L. Protein dynamics. Direct observation of hierarchical protein dynamics. *Science* **2015**, *348*, 578–581. [CrossRef] [PubMed]

12.　Fontes, M.R.; Teh, T.; Kobe, B. Structural basis of recognition of monopartite and bipartite nuclear localization sequences by mammalian importin-alpha. *J. Mol. Biol.* **2000**, *297*, 1183–1194. [CrossRef] [PubMed]

13.　Diez, M.; Zimmermann, B.; Borsch, M.; Konig, M.; Schweinberger, E.; Steigmiller, S.; Reuter, R.; Felekyan, S.; Kudryavtsev, V.; Seidel, C.A.; et al. Proton-powered subunit rotation in single membrane-bound F0F1-ATP synthase. *Nat. Struct. Mol. Biol.* **2004**, *11*, 135–141. [CrossRef] [PubMed]

14.　Lindorff-Larsen, K.; Best, R.B.; Depristo, M.A.; Dobson, C.M.; Vendruscolo, M. Simultaneous determination of protein structure and dynamics. *Nature* **2005**, *433*, 128–132. [CrossRef] [PubMed]

15.　Jensen, M.R.; Zweckstetter, M.; Huang, J.R.; Blackledge, M. Exploring free-energy landscapes of intrinsically disordered proteins at atomic resolution using NMR spectroscopy. *Chem. Rev.* **2014**, *114*, 6632–6660. [CrossRef] [PubMed]

16.　Zhang, J.; Maslov, S.; Shakhnovich, E.I. Constraints imposed by non-functional protein-protein interactions on gene expression and proteome size. *Mol. Syst. Biol.* **2008**, *4*, 210. [CrossRef] [PubMed]

17.　Levy, E.D.; De, S.; Teichmann, S.A. Cellular crowding imposes global constraints on the chemistry and evolution of proteomes. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 20461–20466. [CrossRef] [PubMed]

18.　McGuffee, S.R.; Elcock, A.H. Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm. *PLoS Comput. Biol.* **2010**, *6*, e1000694.

19.　Tawfik, D.S. Messy biology and the origins of evolutionary innovations. *Nat. Chem. Biol.* **2010**, *6*, 692–696. [CrossRef] [PubMed]

20.　Butner, K.A.; Kirschner, M.W. Tau protein binds to microtubules through a flexible array of distributed weak sites. *J. Cell. Biol.* **1991**, *115*, 717–730. [CrossRef] [PubMed]

21.　Sigler, P.B. Transcriptional activation. Acid blobs and negative noodles. *Nature* **1988**, *333*, 210–212. [CrossRef] [PubMed]

22.　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. [CrossRef] [PubMed]

23.　Levy, E.D.; Boeri Erba, E.; Robinson, C.V.; Teichmann, S.A. Assembly reflects evolution of protein complexes. *Nature* **2008**, *453*, 1262–1265. [CrossRef] [PubMed]

24.　Wu, H. Higher-order assemblies in a new paradigm of signal transduction. *Cell* **2013**, *153*, 287–292. [CrossRef] [PubMed]

25.　Bienz, M. Signalosome assembly by domains undergoing dynamic head-to-tail polymerization. *Trends Biochem. Sci.* **2014**, *39*, 487–495. [CrossRef] [PubMed]

26.　Franzmann, T.M.; Jahnel, M.; Pozniakovsky, A.; Mahamid, J.; Holehouse, A.S.; Nuske, E.; Richter, D.; Baumeister, W.; Grill, S.W.; Pappu, R.V.; et al. Phase separation of a yeast prion protein promotes cellular fitness. *Science* **2018**, *359*, 5654. [CrossRef] [PubMed]

27.　Copley, S.D. Shining a light on enzyme promiscuity. *Curr. Opin. Struct. Biol.* **2017**, *47*, 167–175. [CrossRef] [PubMed]

28. Manglik, A.; Kim, T.H.; Masureel, M.; Altenbach, C.; Yang, Z.; Hilger, D.; Lerch, M.T.; Kobilka, T.S.; Thian, F.S.; Hubbell, W.L.; et al. Structural Insights into the Dynamic Process of β₂-Adrenergic Receptor Signaling. *Cell* **2015**, *161*, 1101–1111. [CrossRef] [PubMed]

29. Ross, E.D.; Baxa, U.; Wickner, R.B. Scrambled prion domains form prions and amyloid. *Mol. Cell. Biol.* **2004**, *24*, 7206–7213. [CrossRef] [PubMed]

30. Khersonsky, O.; Tawfik, D.S. Enzyme promiscuity: A mechanistic and evolutionary perspective. *Annu. Rev. Biochem.* **2010**, *79*, 471–505. [PubMed]

31. Jeffery, C.J. Moonlighting proteins: Old proteins learning new tricks. *Trends Genet.* **2003**, *19*, 415–417. [CrossRef]

32. Freilich, S.; Spriggs, R.V.; George, R.A.; Al-Lazikani, B.; Swindells, M.; Thornton, J.M. The complement of enzymatic sets in different species. *J. Mol. Biol.* **2005**, *349*, 745–763. [CrossRef] [PubMed]

33. Nam, H.; Lewis, N.E.; Lerman, J.A.; Lee, D.H.; Chang, R.L.; Kim, D.; Palsson, B.O. Network context and selection in the evolution to enzyme specificity. *Science* **2012**, *337*, 1101–1104. [CrossRef] [PubMed]

34. Arnold, F.H.; Wintrode, P.L.; Miyazaki, K.; Gershenson, A. How enzymes adapt: Lessons from directed evolution. *Trends Biochem. Sci.* **2001**, *26*, 100–106. [CrossRef]

35. Tokuriki, N.; Jackson, C.J.; Afriat-Jurnou, L.; Wyganowski, K.T.; Tang, R.; Tawfik, D.S. Diminishing returns and tradeoffs constrain the laboratory optimization of an enzyme. *Nat. Commun.* **2012**, *3*, 1257. [CrossRef] [PubMed]

36. Aharoni, A.; Gaidukov, L.; Khersonsky, O.; Mc, Q.G.S.; Roodveldt, C.; Tawfik, D.S. The 'evolvability' of promiscuous protein functions. *Nat. Genet.* **2005**, *37*, 73–76. [CrossRef] [PubMed]

37. Amitai, G.; Gupta, R.D.; Tawfik, D.S. Latent evolutionary potentials under the neutral mutational drift of an enzyme. *HFSP J.* **2007**, *1*, 67–78. [CrossRef] [PubMed]

38. James, L.C.; Roversi, P.; Tawfik, D.S. Antibody multispecificity mediated by conformational diversity. *Science* **2003**, *299*, 1362–1367. [CrossRef] [PubMed]

39. Campbell, E.; Kaltenbach, M.; Correy, G.J.; Carr, P.D.; Porebski, B.T.; Livingstone, E.K.; Afriat-Jurnou, L.; Buckle, A.M.; Weik, M.; Hollfelder, F.; et al. The role of protein dynamics in the evolution of new enzyme function. *Nat. Chem. Biol.* **2016**, *12*, 944–950. [CrossRef] [PubMed]

40. Brooks, B.; Karplus, M. Harmonic dynamics of proteins: Normal modes and fluctuations in bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 6571–6575. [CrossRef] [PubMed]

41. Marsh, J.A.; Teichmann, S.A. Parallel dynamics and evolution: Protein conformational fluctuations and assembly reflect evolutionary changes in sequence and structure. *Bioessays* **2014**, *36*, 209–218. [CrossRef] [PubMed]

42. Haliloglu, T.; Bahar, I. Adaptability of protein structures to enable functional interactions and evolutionary implications. *Curr. Opin. Struct. Biol.* **2015**, *35*, 17–23. [CrossRef] [PubMed]

43. Smock, R.G.; Gierasch, L.M. Sending signals dynamically. *Science* **2009**, *324*, 198–203. [CrossRef] [PubMed]

44. Nygaard, R.; Zou, Y.; Dror, R.O.; Mildorf, T.J.; Arlow, D.H.; Manglik, A.; Pan, A.C.; Liu, C.W.; Fung, J.J.; Bokoch, M.P.; et al. The dynamic process of beta(2)-adrenergic receptor activation. *Cell* **2013**, *152*, 532–542. [CrossRef] [PubMed]

45. Tompa, P.; Fuxreiter, M. Fuzzy complexes: Polymorphism and structural disorder in protein-protein interactions. *Trends Biochem. Sci.* **2008**, *33*, 2–8. [CrossRef] [PubMed]

46. Sharma, R.; Raduly, Z.; Miskei, M.; Fuxreiter, M. Fuzzy complexes: Specific binding without complete folding. *FEBS Lett.* **2015**, *589*, 2533–2542. [CrossRef] [PubMed]

47. Fuxreiter, M. Fuzziness in Protein Interactions-A Historical Perspective. *J. Mol. Biol.* **2018**, *430*, 2278–2287. [CrossRef] [PubMed]

48. Hadzi, S.; Mernik, A.; Podlipnik, C.; Loris, R.; Lah, J. The Thermodynamic Basis of the Fuzzy Interaction of an Intrinsically Disordered Protein. *Angew Chem. Int. Ed. Engl.* **2017**, *56*, 14494–14497. [CrossRef] [PubMed]

49. Fuxreiter, M.; Simon, I.; Bondos, S. Dynamic protein-DNA recognition: Beyond what can be seen. *Trends Biochem. Sci.* **2011**, *36*, 415–423. [CrossRef] [PubMed]

50. Buljan, M.; Chalancon, G.; Eustermann, S.; Wagner, G.P.; Fuxreiter, M.; Bateman, A.; Babu, M.M. Tissue-Specific Splicing of Disordered Segments that Embed Binding Motifs Rewires Protein Interaction Networks. *Mol. Cell* **2012**, *46*, 871–883. [CrossRef] [PubMed]

51. 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. [CrossRef] [PubMed]

52. 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. [CrossRef] [PubMed]

53. Fuxreiter, M. Fold or not to fold upon binding-does it really matter? *Curr. Opin. Struct. Biol.* **2018**, *54*, 19–25. [CrossRef] [PubMed]

54. Warfield, L.; Tuttle, L.M.; Pacheco, D.; Klevit, R.E.; Hahn, S. A sequence-specific transcription activator motif and powerful synthetic variants that bind Mediator using a fuzzy protein interface. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E3506–E3513. [CrossRef] [PubMed]

55. 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. [CrossRef] [PubMed]

56. 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. [CrossRef] [PubMed]

57. Kragelj, J.; Palencia, A.; Nanao, M.H.; Maurin, D.; Bouvignies, G.; Blackledge, M.; Jensen, M.R. Structure and dynamics of the MKK7-JNK signaling complex. *Proc. Natl. Acad. Sci. USA.* **2015**, *112*, 3409–3414. [CrossRef] [PubMed]

58. 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. [CrossRef] [PubMed]

59. Lu, A.; Magupalli, V.G.; Ruan, J.; Yin, Q.; Atianand, M.K.; Vos, M.R.; Schroder, G.F.; Fitzgerald, K.A.; Wu, H.; Egelman, E.H. Unified polymerization mechanism for the assembly of ASC-dependent inflammasomes. *Cell* **2014**, *156*, 1193–1206. [CrossRef] [PubMed]

60. Lee, G.M.; Pufall, M.A.; Meeker, C.A.; Kang, H.S.; Graves, B.J.; McIntosh, L.P. The affinity of Ets-1 for DNA is modulated by phosphorylation through transient interactions of an unstructured region. *J. Mol. Biol.* **2008**, *382*, 1014–1030. [CrossRef] [PubMed]

61. Li, J.; McQuade, T.; Siemer, A.B.; Napetschnig, J.; Moriwaki, K.; Hsiao, Y.S.; Damko, E.; Moquin, D.; Walz, T.; McDermott, A.; et al. The RIP1/RIP3 Necrosome Forms a Functional Amyloid Signaling Complex Required for Programmed Necrosis. *Cell* **2012**, *150*, 339–350. [CrossRef] [PubMed]

62. Davey, N.E.; Cyert, M.S.; Moses, A.M. Short linear motifs-ex nihilo evolution of protein regulation. *Cell Commun. Signal.* **2015**, *13*, 43. [CrossRef] [PubMed]

63. Lu, X.; Hamkalo, B.; Parseghian, M.H.; Hansen, J.C. Chromatin condensing functions of the linker histone C-terminal domain are mediated by specific amino acid composition and intrinsic protein disorder. *Biochemistry* **2009**, *48*, 164–172. [CrossRef] [PubMed]

64. Ross, E.D.; Edskes, H.K.; Terry, M.J.; Wickner, R.B. Primary sequence independence for prion formation. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 12825–12830. [CrossRef] [PubMed]

65. Nott, T.J.; Petsalaki, E.; Farber, P.; Jervis, D.; Fussner, E.; Plochowietz, A.; Craggs, T.D.; Bazett-Jones, D.P.; Pawson, T.; Forman-Kay, J.D.; et al. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol. Cell* **2015**, *57*, 936–947. [CrossRef] [PubMed]

66. Milles, S.; Mercadante, D.; Aramburu, I.V.; Jensen, M.R.; Banterle, N.; Koehler, C.; Tyagi, S.; Clarke, J.; Shammas, S.L.; Blackledge, M.; et al. Plasticity of an Ultrafast Interaction between Nucleoporins and Nuclear Transport Receptors. *Cell* **2015**, *163*, 734–745. [CrossRef] [PubMed]

67. Ciuffa, R.; Lamark, T.; Tarafder, A.K.; Guesdon, A.; Rybina, S.; Hagen, W.J.; Johansen, T.; Sachse, C. The selective autophagy receptor p62 forms a flexible filamentous helical scaffold. *Cell Rep.* **2015**, *11*, 748–758. [CrossRef] [PubMed]

68. Wu, H.; Fuxreiter, M. The Structure and Dynamics of Higher-Order Assemblies: Amyloids, Signalosomes, and Granules. *Cell* **2016**, *165*, 1055–1066. [CrossRef] [PubMed]

69. Burke, K.A.; Janke, A.M.; Rhine, C.L.; Fawzi, N.L. Residue-by-Residue View of In Vitro FUS Granules that Bind the C-Terminal Domain of RNA Polymerase II. *Mol. Cell* **2015**, *60*, 231–241. [CrossRef] [PubMed]

70. Patel, A.; Lee, H.O.; Jawerth, L.; Maharana, S.; Jahnel, M.; Hein, M.Y.; Stoynov, S.; Mahamid, J.; Saha, S.; Franzmann, T.M.; et al. A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell* **2015**, *162*, 1066–1077. [CrossRef] [PubMed]

71. Molliex, A.; Temirov, J.; Lee, J.; Coughlin, M.; Kanagaraj, A.P.; Kim, H.J.; Mittag, T.; Taylor, J.P. Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization. *Cell* **2015**, *163*, 123–133. [CrossRef] [PubMed]

72. Nagulapalli, M.; Parigi, G.; Yuan, J.; Gsponer, J.; Deraos, G.; Bamm, V.V.; Harauz, G.; Matsoukas, J.; de Planque, M.R.; Gerothanassis, I.P.; et al. Recognition pliability is coupled to structural heterogeneity: A calmodulin intrinsically disordered binding region complex. *Structure* **2012**, *20*, 522–533. [CrossRef] [PubMed]

73. Zadeh, L.A. Fuzzy sets. *Inf. Control* **1965**, *8*, 338–353. [CrossRef]

74. Zadeh, L.A. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans. Syst. Man Cybern.* **1973**, *3*, 28–44. [CrossRef]

75. Mamdani, E.H.; Assilian, S. An experiment in linguistic synthesis with a fuzzy logic controller. *Int. J. Man-Mach. Stud.* **1975**, *7*, 1–13. [CrossRef]

76. Tuu-Szabo, B.; Koczy, L.; Fuxreiter, M. Simulations of Higher-Order Protein Organizations Using a Fuzzy Framework. *Complexity* **2018**, in press.