

Review

Not peer-reviewed version

---

# Intrinsic Disorder and Other Malleable Arsenals of Evolved Protein Multi- functionality

---

Asifa Aftab , Anirneya Basu , [Seema Nath](#) , [Sankar Basu](#) \*

Posted Date: 18 March 2024

doi: 10.20944/preprints202402.1029.v4

Keywords: Protein multi-functionality; Intrinsically Disordered Proteins; Gene duplication; Domain Shuffling; Protein moonlighting



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Review

# Intrinsic Disorder and Other Malleable Arsenals of Evolved Protein Multi-Functionality

Asifa Aftab <sup>1,ξ</sup>, Anirneya Basu <sup>2,ξ</sup>, Seema Nath <sup>3</sup> and Sankar Basu <sup>2,\*</sup>

<sup>1</sup> Department of Zoology, Asutosh College, (affiliated with University of Calcutta), Kolkata - 700026, India

<sup>2</sup> Department of Microbiology, Asutosh College (affiliated with University of Calcutta), Kolkata - 700026, India

<sup>3</sup> Department of Biochemistry and Structural Biology, University of Texas Health Science Center at San Antonio, Texas - 78229, USA

\* Correspondence: author (ORCID: 0000-0003-1393-1982, Email: sankarchandra.basu@asutoshcollege.in)

ξ These authors contributed equally to the work.

**Abstract:** The evolution of human societal dynamics over the course of decades and centuries has led to an ever-increasingly complex digital modern life, attributed by the need for multi-tasking, inclusive of that with media. Is there an empirical mapping between this macroscopic (digital) evolution (in terms of multi-tasking) and the microscopic evolution of the functional molecular units within cells (i.e., proteins) and their manifold cross-talks? Multi-tasking requires a certain degree of intrinsic or adapted flexibility which is as true at the bio-molecular level. Evolutionary diversification of structure-function relationship in proteins emphasizes the functional importance of intrinsically disordered entities (IDPs/IDRs) which are highly dynamic biological soft matters. Multi-functionality is favorably supported by their fluid-like shapes and other new malleable members of the protein family. The new wealth of knowledge coming from recent research on protein intrinsic disorder, fold-switching, moonlighting, hub proteins etc. has given us new insights into protein structure – functional relationship. This has led to a paradigm shift in protein science, clearly diverging from the traditional 'one structure – one function' model applicable to enzyme classes to a more complex understanding of protein functionality. Various evolutionary toys and tools to fit the cumulative multi-functional demands in higher organisms are being explored. This commentary covers the different evolutionary arsenals to achieve the growing multi-functionality and argues in favor of protein intrinsic disorder as probably the sharpest weapon of all.

**Keywords:** protein multi-functionality; intrinsically disordered Proteins; gene duplication; domain shuffling; protein moonlighting

---

## Introduction

*'My own prejudices are exactly opposite of to functionalists': "If you want to understand function, study structure."*

- Francis Crick

One of the hallmark of modern digital life is the need to multitask. The individuality of modern times and the rapid progress of technology have led to an increase in this need, especially with media (Popławska et al., 2021). On average, American youths spend 7.5 hours a day with media, with 29% of that time spent by simultaneous web-serving on different sites (Uncapher et al., 2017). An interesting analogy perhaps can not be left unnoticed at the cellular level with progressive knowledge about protein multi-functionality (Espinosa-Cantú et al., 2020), often associated with health and disease. With research over more than the past two decades, we now have a wealth of new knowledge on protein intrinsic disorder (Uversky, 2013), fold-switching (Bryan and Orban, 2010), domain shuffling (Kawashima et al., 2009), moonlighting (Espinosa-Cantú et al., 2015), hub proteins (Cino et al., 2013) etc. supporting protein multi-functionality in various ways. This knowledge expands our

understanding far beyond the over-simplistic ‘one structure – one function’ model applicable to enzyme classes (Alberts et al., 2002), and, rather, leads us to a more sophisticated comprehension of protein functionality, both for enzymes and non-covalent protein binding (Mannige, 2014). The evolution of multi-functional proteins and multi-enzyme complexes (Alberts et al., 2002; Jeffery, 2003) seems inevitable to fit the increasing (energy) demand in eukaryotes and higher organisms that are continuously under micro-evolutionary selection pressure. This commentary examines the different evolutionary strategies employed to achieve the growing protein multi-functionality and argues in favor of protein intrinsic disorder as probably the most powerful and demanded weapon.

### **Protein Multi-Functionality from an Evolutionary Perspective**

Proteins can serve multiple functions due to evolutionary pressures that lead to adaptations in their structure and function. The fundamental evolutionary basis of protein multi-functionality is rooted to an activity – stability trade-off between conservation of the core protein function (and corresponding key residue positions in the structure) and variation at other mutable peripheral positions allowing the scope for novel functions to evolve (Sikosek and Chan, 2014; Tokuriki et al., 2008). This trade-off can be attained by various evolutionary mechanisms. One mechanism is gene duplication – where redundant copies evolve novel functions (Espinosa-Cantú et al., 2015). Evolutionary changes like mutations or domain shuffling (Anonymous, n.d.) can also increase a protein's functional repertoire, enabling it to perform multiple roles crucial for an organism's survival. Compared to prokaryotes, eukaryotic proteins have more intrinsic disorder with more than 30% of all eukaryotic proteins consisting of long disordered regions comprising of  $\geq 50$  consecutive amino acid residues (Dunker et al., 2001). In comparative genomics studies, it has been observed that bacteriophages and their host bacteria co-evolve on the chromosomal level (Brüssow et al., 2004). Viruses can alter their host cells through short linear motifs in their intrinsically disordered regions (IDRs), which lead to functional binding promiscuity (Davey et al., 2011). Understanding the evolutionary pathways of multi-functional proteins is crucial to unraveling their complexities and biological significance. This commentary presents a comprehensive comparison of different tools for evolved protein multi-functionality and highlights intrinsic disorder as a major means of achieving it.

### **Biophysical Basis of Multi-Functionality in IDPs**

Intrinsically disordered proteins (IDPs) lack a stable fold and any recognizable domains therein. Unlike well-folded globular proteins with a deep energy-well (global minima), they map to rugged energy landscapes with many equivalent local minima. This imparts in them structural plasticity (flexibility) for which they are represented as conformational ensembles rather than a single structure. This in turn, helps them inherit their characteristic binding promiscuity upon encountering different patterns to aid multi-functionality (Fersht, 2009) without showing any overall preference in chaperone binding *in vivo* (Hegyi and Tompa, 2008). Their binding promiscuity stems from a phenomenon called ‘coupled folding and binding’ (Sugase et al., 2007), where they undergo transition from disorder to different structured states upon binding to different partners. The root cause of this phenomenon lies in their non-disjoint flexible backbone trajectory that enables them to accommodate different combinations of side chain rotamers, consistent with different befitting surfaces of ordered protein partners (Dunker et al., 2001). Protein folding and binding (in globular proteins) are analogous processes and can be bridged by concepts like that of complementarity (in geometry and electrostatics) (Basu et al., 2012). In case of IDPs, they keep on hovering across their rugged energy landscape in search for suitable intra and/or intermolecular interactions to stabilize them (Levy et al., 2005; Tsai et al., 1999). Eukaryotic IDPs stay disordered under normal physiological conditions and only fold into ordered structures when they come in contact with their ‘cellular targets’ (Dyson and Wright, 2005, 2002; Uversky, 2002; Wright and Dyson, 1999). It has been theorized that disordered proteins bind weakly and non-specifically to the target and aligns structurally to a befitting surface (and becomes structured) as it approaches the cognate binding site(s) (Shoemaker et al., 2000). In many cases, especially in signal transduction pathways, the bindings are essentially

transient (meta-stable). IDPs can also escape protein degradation by undergoing a co-translational folding mechanism involving the ribosomal surface and molecular chaperons (Simister et al., 2011; Waudby et al., 2019). These adaptabilities enable IDPs to engage in numerous cellular processes, contributing to their multi-functionality despite lacking a defined structure. Importantly, IDRs of proteins also serve as promising (fuzzy) drug-targets (Kamagata et al., 2019, p. 53; Saurabh et al., 2023), wherein a whole new approach for drug-development has lately been initiated, accounting for an acceptable representation of their conformational ensemble (Saurabh et al., 2023) as the receptor surface (in contrast to well demarcated drug-binding pockets of the folded proteins), thereby increasing the interacting cross-section for the ligands (drugs). The promiscuous binding nature of IDRs are also capitalized to make it potential drug-target, e.g., in case of castration resistant prostate cancer, the disordered N-terminal domain of androgen receptor are being targeted to overcome existing drug-resistance (Yi et al., 2023). Formation of binding competent transient structures (conformational clusters) induced by molecular crowding in the close vicinity of IDPs/IDRs is another unique and idiosyncratic mechanisms to exhibit binding promiscuity and multi-functionality – as demonstrated in intrinsically disordered proteins of the Gab family (Gab1) during signal transduction (Gruber et al., 2022). It is presumed that these contribute to the spatial organization of complex components. These observations can show us the path which we can take to decipher and understand the mechanism of the assembly of very large and distinct signal transduction protein complexes (viz., ‘signalosomes’ that are stimulus specific) in response to certain stimuli in a short period of time (Simister et al., 2011).

### Weaponry of Evolved Protein Multi-Functionality

Evolved protein multi-functionality harnesses several armaments (molecular evolutionary strategies) in its armoury (**Figure 1**).

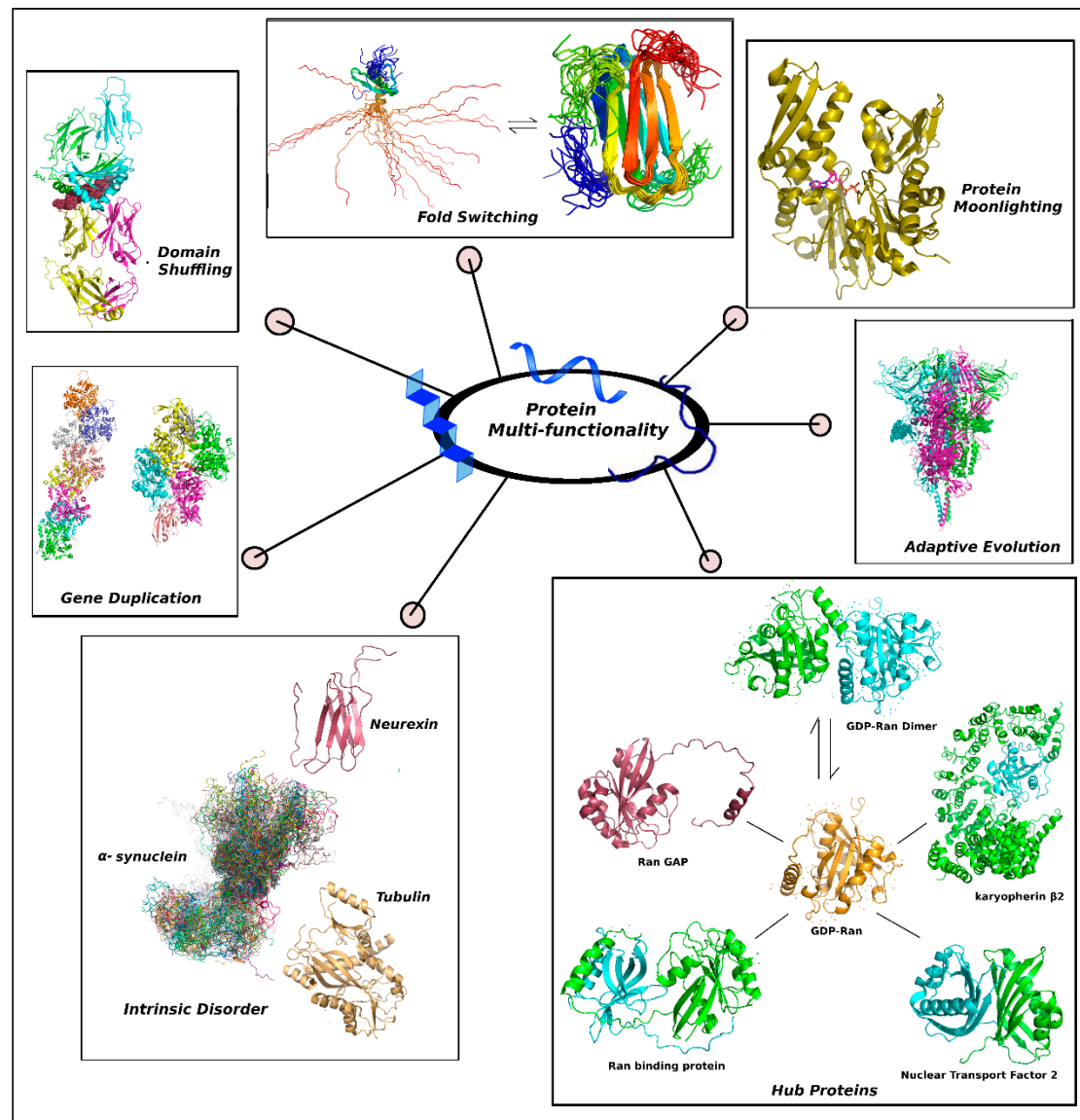
Following is a comparative discussion of these evolutionary tools.

**i. Gene Duplication & Functional Divergence:** Gene duplications create redundant gene copies, allowing one copy to retain the original function while the paralog (often varying at their oligomeric states (Mallik et al., 2022)) accumulates mutations at a higher rate and is often fixed in the population by acquiring an adaptive function according to the classical model of divergence by neo-functionalization (Anonymous, n.d.). To that end, accelerated evolution in retained paralogs (e.g., Rck1/Rck2, Ptc2/Ptc3, Sim1/Sun4, Ktr5/Ktr6 (Hughes and Friedman, 2003) paralogs) have been observed through evolving post-translational regulation mechanisms (diversified short linear motif like sequences) (Nguyen Ba et al., 2014). At the other end, if we consider the model of sub-functionalization, after gene duplication and divergence, the biological functions of ancestor get partitioned between two paralogs. Sub-functionalization may be of two types: qualitative and quantitative. Qualitative subfunctionalization of the molecular functions that trade-off between each other in the ancestral gene. Each paralog may then evolve towards the optimization of the retained function. Alternatively, quantitative subfunctionalization occurs when neutral evolution results in complementary loss-of-function mutations between the paralogs. In this model, both duplicates become indispensable as they together provide the ancestral functional requirements (Fewell and Woolford, 1999; He and Zhang, 2005; Lynch and Force, 2000).

**ii. Domain Shuffling:** Reorganization of protein domains can create multi-functional proteins by combining existing functional units in new ways. It may come through horizontal gene transfer (e.g., from eukaryotes to prokaryotes) or by in-del mutations of genes, post duplication. One common way in which domain shuffling leads to novel functions is by the shuffling of exons (exon shuffling, analogous to alternative splicing at the m-RNA level) followed by in-del mutations. Usually this is established by a mapping of exons and domains (e.g., a single exon coding for a single complete domain) (Anonymous, n.d.). Insertion of a ‘nested’ domain may also interrupt the linear sequence of a structural domain. Such insertions often map to disordered loops in the parent structure. For example, this has been found in phospholipase C $\gamma$  wherein an insert of ~300 residues (comprising of one SH3 and two SH2 domains) separates one of its two Pleckstrin Homology (PH) domains (Fadri et al., 2005). Certain domains (e.g., the Xlink domain) of the protein aggrecan (Kawashima et al., 2009),



the most abundant noncollagenous protein in cartilage, is also said to have been created by domain shuffling in ancestral vertebrates. Multi-domain proteins (Vishwanath et al., 2018) are also useful tool to express protein multi-functionality alongside with providing folding benefits and structural stability.



**Figure 1. The composite figure portrays different evolutionary tools to express protein multi-functionality.** The represented repertoire includes (i) Gene Duplication: Actins (tested structurally in yeasts, camelidae, insects) separated by two insertions - deletions (in-del) mutations (Q211→D231, D299→T365): *ARP4* (PDB ID: 5NBM) where filament formation is inhibited and its heteromerized paralog *ACT1* (6BNO) (Mallik et al., 2022), (ii) Domain Shuffling: human Aggrecan core peptide (highlighted as dots) presented by class II major histocompatibility complexes (7RDV) (Kawashima et al., 2009), (iii) Fold Switching: human Lymphotoxin (1J8I ↔ 2JP1) (Bryan and Orban, 2010), (iv) Protein Moonlighting: Yeast Heat shock protein Hsp70 bound with ADP (3QFU) (Jeffery, 2018), (v) Adaptive Evolution: the Spike protein with highly mutable FLCSp<sub>Spike</sub> (highlighted as dots) from SARS-CoV-2 (6XR8), (vi) Hub Proteins: Canine GDP-Ran (monomer: 1QG4 ↔ dimer: 1BYU) with its interactome consisting of Ran binding protein (RBP): 1RRP, Ran GAP: 1K5D, karyopherin β2: 1QBK, Nuclear Transport Factor 2 (NTF-2): 5BXQ etc. (Higurashi et al., 2008) and (iii) Intrinsic Disorder: human alpha synuclein (a conformational ensemble picked up from its MD-simulated trajectory (Bandyopadhyay and Basu, 2020)) with its two cognate globular binding partners: Tubulin (4YRL) and β-neurexin 1 (3MW2).

**iii. Protein Moonlighting:** In contrast to gene-fusion, alternative splicing or functional peptides resulting from multiple proteolysis, protein moonlighting (Jeffery, 2018, 2003) refers to multi-functionality evolved in proteins (especially enzymes) without requiring any change in their primary sequence, typically expressed via alternative sites to that of the primary active site – which often maps to a pocket for catalysis (Jeffery, 1999; Piatigorsky, 2009). In these proteins both classic and non-classic type protein functions co-exist wherein the former refers to enzymatic activities (i.e., involving covalent bond breaking and making) while the later refers to protein – protein interactions (via alternative part of the protein's surface). However these alternative sites are different to that of allosteric regulations often found with enzymes like phosphofructokinase, hemoglobins etc. Heat shock proteins (HSPs) are classic examples of protein moonlighting.

**iv. Fold-switching Proteins:** Fold-switching proteins (Bryan and Orban, 2010), a newly emerging class of proteins undergo a distinct switching of their folds by remodelling their secondary structures upon change in environmental (physiological) conditions, for example, a change in pH (Baruah and Biswas, 2015). Upon fold-switching, they respond to cellular stimuli enabling them to perform important alternative regulatory (e.g., transcriptional regulation) functions of the cell (demonstrated in proteins like RfaH, KaiB etc.) (Bernhardt and Hansmann, 2018; Kim and Porter, 2021). Another dramatic example is Lymphotactin which undergoes conformational switching (Bryan and Orban, 2010).

**v. Adaptive Evolution:** Environmental pressures, such as genetic drift, natural selection etc. drive proteins to adapt, acquiring new functions that enhance an organism's survival fitness. Adaptive mutations are largely amino-acid substitutions that occur at the protein surface with a high degree of solvent accessibility for these exposed residues that make them most prone to mutations. Population genomics studies in model systems (*Drosophila* & *Arabidopsis*) surveyed a multitude of genomic, structural and functional descriptors and revealed that (i) the rate of adaptive substitutions are different for different functional classes (with the fastest rates of protein adaptation observed in proteins involved in translation, degradation and signalling) while (ii) intermolecular interactions (e.g., host-pathogen coevolution) is a major determinant for adaptive evolution (Moutinho et al., 2019). Multifunctional viral proteins are classic examples of adaptive evolution (Hasiów-Jaroszewska et al., 2014). The recent case is of course the Spike protein of the Coronavirus rapidly undergoing mutations (particularly at the solvent exposed disordered loop containing the crucial Furin like cleavage site or FLCSSpike (Balaram, 2021; Roy et al., 2022)) from SARS-CoV-2 → omicron (Araf et al., 2022), deltacron (Maulud et al., 2022) etc. Significant patterns of co-occurrence of adaptive events have also been identified in the RNA binding domains with functional overlapping of the HC-Pro of the potyvirus (established by covariation analyses) (Hasiów-Jaroszewska et al., 2014).

**vi. Intrinsically Disordered Proteins:** IDPs are biological soft matters (Bandyopadhyay and Basu, 2020) that are highly dynamic and biologically active (Uversky, 2016a). Unlike globular proteins, they do not have enough hydrophobic residues to trigger a hydrophobic collapse. Instead, they have high amounts of polar and charged residues (Basu and Biswas, 2018; Már et al., 2023; Sun et al., 2013; Uversky, 2016a) which contribute to less sequence complexity in the absence of folding (Már et al., 2023; Uversky, 2016a). This results in partial temporal order by hydrogen bonding, water mediated contacts (indirect readouts) (Reid et al., 2023) and formation of transient interchangeable salt-bridges (Basu and Biswas, 2018). IDPs do not have a characteristic deep well in their energy landscapes like globular proteins, which means they do not conform to a lone stable 3D structure under physiological conditions. They have an affinity to undergo transition from disorder to order and back to disorder (Basu and Biswas, 2018; Sun et al., 2013). This makes them highly flexible and adaptable. Partially disordered proteins have intrinsically disordered regions that may be present in varying degrees. These regions often map to hybrid proteins that have both ordered and disordered regions (Sun et al., 2013; Uversky, 2016a). A classic example of this is p53 (Xue et al., 2013).

Recurrent salt bridges (especially, those with short-range contact orders) impart local temporal structural rigidity in IDPs. However, it is crucial to maintain a balance between the number of salt bridges that allow flexibility and prevent complete rigidity, as seen in globular proteins (Basu and Biswas, 2018). Studies (Bandyopadhyay and Basu, 2020; Basu and Biswas, 2018; Roy et al., 2022) have

demonstrated that salt bridges in IDPs are typically not stable (or, persistent) and tend to dissolve and reform frequently with various interchangeable counter-ionic partners. This phenomenon is referred to as 'transient salt bridge dynamics'. This is necessary to accommodate the high occurrence of oppositely charged residues and to allow for sampling of different conformations, leading to an ensemble. These conformations are not random but revolve around a finite number of structurally degenerate conformational clusters (Bandyopadhyay and Basu, 2020). Phase transitions among these clusters are often triggered by switching of transient salt bridges, demonstrating critical behavior similar to a sand-pile model. The presence of these transient or flitting salt bridges may stabilize the IDP in a conformationally dependent manner, locked by befitting surfaces of its globular counterparts (Bandyopadhyay and Basu, 2020). This is especially relevant in the case of cell signalling, such as suppressors of cytokine signalling (SOCs) (Bandyopadhyay and Basu, 2020), where IDRs in eukaryotic transcription factors (Már et al., 2023) are evolving with high sequence heterogeneity and demonstrated dynamic multi-functionality through their binding promiscuity.

IDPs, lacking a fixed structure or folding code, exist as highly dynamic '*dancing protein clouds*' (Uversky, 2016a) that can adopt different shapes depending on their local environment. When IDPs interact with ordered proteins, their binding contributes to at least partial folding, which depends on the binding partner. Different binding partners can induce different folds (Uversky, 2016a; Wright and Dyson, 1999), making them highly adaptable. Additionally, IDPs exhibit fractals and heterogeneity, meaning that they neither converge to a steady state nor diverge to infinity, but rather stay within a defined and chaotic region.

**vii. Hub Proteins:** Hub proteins (Higurashi et al., 2008) are well-structured proteins with a (hub-like) high degree in a protein – protein interaction (PPI) network. They can interact with multiple partners, even those associated with very different protein networks, leading to diverse biological processes. What sets hub proteins apart is their tendency to interact with disordered partners (IDPs/IDRs) (Cino et al., 2013), allowing them to interact with structurally diverse partners (Sun et al., 2013). In simpler terms, hub proteins can fold into different ordered conformations when they bind to different Molecular Recognition Features (MoRFs) of their preferentially disordered binding partners. This feature makes them highly adaptable and valuable in a wide range of biological processes. One example of Hub proteins is the canine GDP-Ran, the active monomeric form of which interacts with RBP, Ran GAP, karyopherin  $\beta$ 2, NTF-2 etc. apart from forming a biological dimer (Higurashi et al., 2008).

### Necessity of IDPs as SOS (Ad Hoc) Tools for Multi-Functionality in Higher Organisms

The oversimplified 'one gene - one enzyme' hypothesis (Beadle and Tatum, 1941) has long been outdated with an evolving definition of 'gene' (Portin and Wilkins, 2017), and, perhaps even more so, with the growing knowledge of IDPs in recent times. In a human cell, there are approximately  $10^4$  genes that code for proteins, while the number of different proteins is around  $10^6$ . This suggests that there must be mechanisms that cause originally coded proteins to change biophysically, allowing them to acquire multiple functions (Sun et al., 2013; Uversky, 2016b). One such mechanism is the diverse functionality exhibited by IDPs (Clark et al., 2015; Sickmeier et al., 2007), which are more common in eukaryotes than in prokaryotes. The acquired multi-functionality in IDPs is supported by their fluid like flexibility, conformational dynamics (Bandyopadhyay and Basu, 2020; Basu and Biswas, 2018) and binding promiscuity (Morris et al., n.d.). Eukaryotic proteins are mostly intrinsically disordered, with around 23-28% (Basile et al., 2019; Zamora-Briseño et al., 2021) showing this characteristic, while about 70% of signalling proteins seem to possess intrinsic disorder. IDPs are highly flexible and can undergo conformational changes to fit the surfaces of their binding partners. They exhibit high binding plasticity and a low affinity – high specificity trade-off as their conformational flexibility enables them to become almost tailor made for its partner(s) in each event of binding (Sun et al., 2013). IDPs and IDRs thus complement the repertoire of ordered proteins (i.e., globular, fibrous, membrane proteins) by providing multi-functionality while the ordered proteins with one distinct native global minima structure carry out their routine (primarily, housekeeping) functions (Uversky, 2016a). In particular, IDPs are essential for cell signalling pathways, as they allow

for high specificity, transitory (switch-like) and reversible (Wright and Dyson, 2015) interactions (formation of meta-stable 'fuzzy complexes' upon binding to suitable partners) that are not possible with ordered proteins. Ordered and disordered proteins work together to bring about cellular functions efficiently, with the presence of long IDRs being more common in eukaryotes than in Archaea and Bacteria (Wright and Dyson, 2015). The complexity of an organism is directly proportional to the demand for IDRs, as higher organisms often require more cellular signalling that relies on these proteins.

### **p53: Example of a Unique Idiosyncratic Multi-Functional Hybrid Protein with Functionally Crucial IDRs**

Hybrid proteins contain structured regions that are connected by disordered loops (i.e., IDRs). IDRs are directly correlated with sequence diversity, making them robust for their regulatory functions. A prime example of this is p53, a protein found in both vertebrates and invertebrates, which has a unique structure-functional mapping. Its primary function is to suppress tumors by regulating cell cycle and control. However, it also has many other related non-enzymatic biological functions, such as PPI and DNA-binding. It can form different biologically active multimers like homo-tetramers and isoform-based hetero-tetramers. Additionally, it undergoes alternative splicing and has many preferentially localized pre- and post-translational modifications that lead to various isoforms known as 'proteoforms'. These combinations, along with the presence of multiple disorder-based protein binding sites, allow p53 to adopt meta-stable states upon interacting with many binding partners in a switch-like transient manner, characteristic of signal transducers and eukaryotic transcription factors (Már et al., 2023; Uversky, 2016b). This is possible due to the flexibility and sequence diversity offered by its IDRs. While acting as a tumor suppressor, it binds to DNA via its highly-conserved, well-structured DNA-binding domains. The flanking and interconnecting IDRs often promote these bindings to different partners transiently (Xue et al., 2013). These IDRs situated amidst structured domains in hybrid proteins have high amino acid substitution rates, leading to high sequence heterogeneity. The resultant expressed structural heterogeneity can be categorized into foldons, inducible foldons, semi-foldons, non-foldons, and unfoldons (Uversky, 2016b, 2013), underscoring their promiscuous binding capacity and their significance in PPI networks and signaling pathways. With over 1000 binding partners, p53's intrinsic disorder is essential for its functionality. This intricate interplay between protein variation, intrinsic disorder, and functionality underscores the complexity of the biological machinery, with implications for understanding disease pathogenesis and the regulation of cellular processes.

### **Conclusion**

The long-established model of protein structure-function relationship by 'induced fit' and 'lock-n-key' model was questioned by the rising evidence of IDRs of proteins which are often seen as functionally complementary to that of structurally ordered proteins. The IDR is coined as an umbrella term to indicate either the full protein or any part of it which may adopt energetically favorable complex and heterogeneous spatiotemporal structures under various suitable conditions. Genetic drift and positive selection contribute to the rapid evolution and diversification of IDP/IDRs which could be functionally classified as chaperones, recognizers, protein assemblers, scavengers, signal transducers, effectors etc. IDPs are highly populated with structure-breaking residues like Gly and Pro and other polar residues and lack hydrophobic and aromatic residues which are the core components of a properly folded globular protein. Different structural mosaic patterns found in IDPs render them multi-functionality and promiscuous binding nature. These two aspects are two sides of a coin presenting advantages and disadvantages simultaneously. The energy efficiency attributed to IDP/IDRs in the cellular milieu stems from its multifunctional nature enabling them to transiently bind to a vast spectrum of interacting partners. But this apparent energy-efficient strategy to perform various physiological functions often gets backfired by 'high specificity – low affinity' binding, attributed to neurodegenerative diseases, cardiovascular diseases, and cancer. The promiscuous binding of IDP/IDR comes into play to alleviate the pathological conditions where the protein-protein



interactions are being targeted as potential therapeutic measures in drug development. The rapid advancements in bioinformatics tools are crucial for deciphering the complicated world of IDP/IDRs and utilizing it to our advantages.

**Acknowledgement:** We thank Dr. Ritobrita Chakraborty, Post Doctoral Researcher, University of Pennsylvania, Philadelphia, USA for crucial discussions on the topic during the work.

**Funding:** The project was self-funded.

**Competing Interests:** *The authors have no relevant financial or non-financial interests to disclose.*

**Authors Contributions:** S.B. conceptualized the problem. A.A. and A.B. wrote the first draft together. Everybody participated in the literature survey. S.B. made the figure and edited the manuscript rigorously with crucial inputs from S.N.. S.N. wrote the conclusion. All authors read and approved the final manuscript.

**Data Availability:** Not applicable

## References

- Alberts B, Johnson A, Lewis J, et al. Protein Function. In: Molecular Biology of the Cell. 4th Edition Garland Science; 2002.
- Anonymous. Domain Shuffling - an Overview | ScienceDirect Topics. n.d. Available from: <https://www.sciencedirect.com/topics/neuroscience/domain-shuffling> [Last accessed: 2/1/2024a].
- Anonymous. Evolution by Gene Duplication | SpringerLink. n.d. Available from: <https://link.springer.com/book/10.1007/978-3-642-86659-3> [Last accessed: 1/23/2024b].
- Araf Y, Akter F, Tang Y-D, et al. Omicron Variant of SARS-CoV-2: Genomics, Transmissibility, and Responses to Current COVID-19 Vaccines. *J Med Virol* 2022;94(5):1825–1832; doi: 10.1002/jmv.27588.
- Balaram P. The Murky Origins of the Coronavirus SARS-CoV-2, the Causative Agent of the COVID-19 Pandemic. *CURRENT SCIENCE* 2021;120(11):4.
- Bandyopadhyay A and Basu S. Criticality in the Conformational Phase Transition among Self-Similar Groups in Intrinsically Disordered Proteins: Probed by Salt-Bridge Dynamics. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 2020;1868(10):140474; doi: 10.1016/j.bbapap.2020.140474.
- Baruah A and Biswas P. Designing pH Induced Fold Switch in Proteins. *J Chem Phys* 2015;142(18):185102; doi: 10.1063/1.4920938.
- Basile W, Salvatore M, Bassot C, et al. Why Do Eukaryotic Proteins Contain More Intrinsically Disordered Regions? *PLoS Comput Biol* 2019;15(7):e1007186; doi: 10.1371/journal.pcbi.1007186.
- Basu S, Bhattacharyya D and Banerjee R. Self-Complementarity within Proteins: Bridging the Gap between Binding and Folding. *Biophys J* 2012;102(11):2605–2614; doi: 10.1016/j.bpj.2012.04.029.
- Basu S and Biswas P. Salt-Bridge Dynamics in Intrinsically Disordered Proteins: A Trade-off between Electrostatic Interactions and Structural Flexibility. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 2018;1866(5):624–641; doi: 10.1016/j.bbapap.2018.03.002.
- Beadle GW and Tatum EL. Genetic Control of Biochemical Reactions in *Neurospora*\*. *Proceedings of the National Academy of Sciences* 1941;27(11):499–506; doi: 10.1073/pnas.27.11.499.
- Bernhardt NA and Hansmann UHE. Multifunnel Landscape of the Fold-Switching Protein RfaH-CTD. *J Phys Chem B* 2018;122(5):1600–1607; doi: 10.1021/acs.jpcb.7b11352.
- Brüssow H, Canchaya C and Hardt W-D. Phages and the Evolution of Bacterial Pathogens: From Genomic Rearrangements to Lysogenic Conversion. *Microbiol Mol Biol Rev* 2004;68(3):560–602; doi: 10.1128/MMBR.68.3.560-602.2004.
- Bryan PN and Orban J. Proteins That Switch Folds. *Curr Opin Struct Biol* 2010;20(4):482–488; doi: 10.1016/j.sbi.2010.06.002.
- Cino EA, Killoran RC, Karttunen M, et al. Binding of Disordered Proteins to a Protein Hub. *Sci Rep* 2013;3(1):2305; doi: 10.1038/srep02305.
- Clark SA, Jespersen N, Woodward C, et al. Multivalent IDP Assemblies: Unique Properties of LC8-Associated, IDP Duplex Scaffolds. *FEBS Letters* 2015;589(19, Part A):2543–2551; doi: 10.1016/j.febslet.2015.07.032.
- Davey NE, Travé G and Gibson TJ. How Viruses Hijack Cell Regulation. *Trends Biochem Sci* 2011;36(3):159–169; doi: 10.1016/j.tibs.2010.10.002.
- Dunker AK, Lawson JD, Brown CJ, et al. Intrinsically Disordered Protein. *Journal of Molecular Graphics and Modelling* 2001;19(1):26–59; doi: 10.1016/S1093-3263(00)00138-8.

- Dyson HJ and Wright PE. Coupling of Folding and Binding for Unstructured Proteins. *Curr Opin Struct Biol* 2002;12(1):54–60; doi: 10.1016/s0959-440x(02)00289-0.
- Dyson HJ and Wright PE. Intrinsically Unstructured Proteins and Their Functions. *Nature Reviews Molecular Cell Biology* 2005;6(3):197–208; doi: 10.1038/nrm1589.
- Espinosa-Cantú A, Ascencio D, Barona-Gómez F, et al. Gene Duplication and the Evolution of Moonlighting Proteins. *Frontiers in Genetics* 2015;6.
- Espinosa-Cantú A, Cruz-Bonilla E, Noda-Garcia L, et al. Multiple Forms of Multifunctional Proteins in Health and Disease. *Front Cell Dev Biol* 2020;8:451; doi: 10.3389/fcell.2020.00451.
- Fadri M, Daquinag A, Wang S, et al. The Pleckstrin Homology Domain Proteins Slm1 and Slm2 Are Required for Actin Cytoskeleton Organization in Yeast and Bind Phosphatidylinositol-4,5-Bisphosphate and TORC2. *Mol Biol Cell* 2005;16(4):1883–1900; doi: 10.1091/mbc.E04-07-0564.
- Fersht PT Alan. Structure and Function of Intrinsically Disordered Proteins. Chapman and Hall/CRC: New York; 2009.; doi: 10.1201/9781420078930.
- Fewell SW and Woolford JL. Ribosomal Protein S14 of *Saccharomyces Cerevisiae* Regulates Its Expression by Binding to RPS14B Pre-mRNA and to 18S rRNA. *Mol Cell Biol* 1999;19(1):826–834; doi: 10.1128/MCB.19.1.826.
- Gruber T, Lewitzky M, Machner L, et al. Macromolecular Crowding Induces a Binding Competent Transient Structure in Intrinsically Disordered Gab1. *Journal of Molecular Biology* 2022;434(5):167407; doi: 10.1016/j.jmb.2021.167407.
- Hasiów-Jaroszewska B, Fares MA and Elena SF. Molecular Evolution of Viral Multifunctional Proteins: The Case of Potyvirus HC-Pro. *J Mol Evol* 2014;78(1):75–86; doi: 10.1007/s00239-013-9601-0.
- He X and Zhang J. Rapid Subfunctionalization Accompanied by Prolonged and Substantial Neofunctionalization in Duplicate Gene Evolution. *Genetics* 2005;169(2):1157–1164; doi: 10.1534/genetics.104.037051.
- Hegyi H and Tompa P. Intrinsically Disordered Proteins Display No Preference for Chaperone Binding In Vivo. *PLOS Computational Biology* 2008;4(3):e1000017; doi: 10.1371/journal.pcbi.1000017.
- Higurashi M, Ishida T and Kinoshita K. Identification of Transient Hub Proteins and the Possible Structural Basis for Their Multiple Interactions. *Protein Sci* 2008;17(1):72–78; doi: 10.1110/ps.073196308.
- Hughes AL and Friedman R. Parallel Evolution by Gene Duplication in the Genomes of Two Unicellular Fungi. *Genome Res* 2003;13(5):794–799; doi: 10.1101/gr.714603.
- Jeffery CJ. Moonlighting Proteins. *Trends Biochem Sci* 1999;24(1):8–11; doi: 10.1016/s0968-0004(98)01335-8.
- Jeffery CJ. Multifunctional Proteins: Examples of Gene Sharing. *Ann Med* 2003;35(1):28–35; doi: 10.1080/07853890310004101.
- Jeffery CJ. Protein Moonlighting: What Is It, and Why Is It Important? *Philos Trans R Soc Lond B Biol Sci* 2018;373(1738):20160523; doi: 10.1098/rstb.2016.0523.
- Kamagata K, Mano E, Itoh Y, et al. Rational Design Using Sequence Information Only Produces a Peptide That Binds to the Intrinsically Disordered Region of P53. *Sci Rep* 2019;9(1):8584; doi: 10.1038/s41598-019-44688-0.
- Kawashima T, Kawashima S, Tanaka C, et al. Domain Shuffling and the Evolution of Vertebrates. *Genome Res* 2009;19(8):1393–1403; doi: 10.1101/gr.087072.108.
- Kim AK and Porter LL. Functional and Regulatory Roles of Fold-Switching Proteins. *Structure* 2021;29(1):6–14; doi: 10.1016/j.str.2020.10.006.
- Levy Y, Cho SS, Onuchic JN, et al. A Survey of Flexible Protein Binding Mechanisms and Their Transition States Using Native Topology Based Energy Landscapes. *Journal of Molecular Biology* 2005;346(4):1121–1145; doi: 10.1016/j.jmb.2004.12.021.
- Lynch M and Force A. The Probability of Duplicate Gene Preservation by Subfunctionalization. *Genetics* 2000;154(1):459–473; doi: 10.1093/genetics/154.1.459.
- Mallik S, Tawfik DS and Levy ED. How Gene Duplication Diversifies the Landscape of Protein Oligomeric State and Function. *Curr Opin Genet Dev* 2022;76:None; doi: 10.1016/j.gde.2022.101966.
- Mannige RV. Dynamic New World: Refining Our View of Protein Structure, Function and Evolution. *Proteomes* 2014;2(1):128–153; doi: 10.3390/proteomes2010128.
- Már M, Nitsenko K and Heidarsson PO. Multifunctional Intrinsically Disordered Regions in Transcription Factors. *Chemistry – A European Journal* 2023;29(21):e202203369; doi: 10.1002/chem.202203369.

- Maulud SQ, Hasan DA, Ali RK, et al. Deltacron: Apprehending a New Phase of the COVID-19 Pandemic. *Int J Surg* 2022;102:106654; doi: 10.1016/j.ijssu.2022.106654.
- Morris OM, Torpey JH and Isaacson RL. Intrinsically Disordered Proteins: Modes of Binding with Emphasis on Disordered Domains. *Open Biol* n.d.;11(10):210222; doi: 10.1098/rsob.210222.
- Moutinho AF, Trancoso FF and Dutheil JY. The Impact of Protein Architecture on Adaptive Evolution. *Mol Biol Evol* 2019;36(9):2013–2028; doi: 10.1093/molbev/msz134.
- Nguyen Ba AN, Strome B, Hua JJ, et al. Detecting Functional Divergence after Gene Duplication through Evolutionary Changes in Posttranslational Regulatory Sequences. *PLoS Comput Biol* 2014;10(12):e1003977; doi: 10.1371/journal.pcbi.1003977.
- Piatigorsky J. Gene Sharing and Evolution: The Diversity of Protein Functions. In: *Gene Sharing and Evolution* Harvard University Press; 2009; doi: 10.4159/9780674042124.
- Popławska A, Szumowska E and Kuś J. Why Do We Need Media Multitasking? A Self-Regulatory Perspective. *Front Psychol* 2021;12:624649; doi: 10.3389/fpsyg.2021.624649.
- Portin P and Wilkins A. The Evolving Definition of the Term “Gene.” *Genetics* 2017;205(4):1353–1364; doi: 10.1534/genetics.116.196956.
- Reid KM, Poudel H and Leitner DM. Dynamics of Hydrogen Bonds between Water and Intrinsically Disordered and Structured Regions of Proteins. *J Phys Chem B* 2023;127(37):7839–7847; doi: 10.1021/acs.jpbc.3c03102.
- Roy S, Ghosh P, Bandyopadhyay A, et al. Capturing a Crucial ‘Disorder-to-Order Transition’ at the Heart of the Coronavirus Molecular Pathology—Triggered by Highly Persistent, Interchangeable Salt-Bridges. *Vaccines* 2022;10(2):301; doi: 10.3390/vaccines10020301.
- Saurabh S, Nadendla K, Purohit SS, et al. Fuzzy Drug Targets: Disordered Proteins in the Drug-Discovery Realm. *ACS Omega* 2023;8(11):9729–9747; doi: 10.1021/acsomega.2c07708.
- Shoemaker BA, Portman JJ and Wolynes PG. Speeding Molecular Recognition by Using the Folding Funnel: The Fly-Casting Mechanism. *Proc Natl Acad Sci U S A* 2000;97(16):8868–8873; doi: 10.1073/pnas.160259697.
- Sickmeier M, Hamilton JA, LeGall T, et al. DisProt: The Database of Disordered Proteins. *Nucleic Acids Res* 2007;35(Database issue):D786–D793; doi: 10.1093/nar/gkl893.
- Sikosek T and Chan HS. Biophysics of Protein Evolution and Evolutionary Protein Biophysics. *J R Soc Interface* 2014;11(100):20140419; doi: 10.1098/rsif.2014.0419.
- Simister PC, Schaper F, O'Reilly N, et al. Self-Organization and Regulation of Intrinsically Disordered Proteins with Folded N-Termini. *PLOS Biology* 2011;9(2):e1000591; doi: 10.1371/journal.pbio.1000591.
- Sugase K, Dyson HJ and Wright PE. Mechanism of Coupled Folding and Binding of an Intrinsically Disordered Protein. *Nature* 2007;447(7147):1021–1025; doi: 10.1038/nature05858.
- Sun X, Rikkerink EHA, Jones WT, et al. Multifarious Roles of Intrinsic Disorder in Proteins Illustrate Its Broad Impact on Plant Biology. *The Plant Cell* 2013;25(1):38–55.
- Tokuriki N, Stricher F, Serrano L, et al. How Protein Stability and New Functions Trade Off. *PLOS Comput Biol* 2008;4(2):e1000002; doi: 10.1371/journal.pcbi.1000002.
- Tsai CJ, Kumar S, Ma B, et al. Folding Funnels, Binding Funnels, and Protein Function. *Protein Sci* 1999;8(6):1181–1190; doi: 10.1110/ps.8.6.1181.
- Uncapher MR, Lin L, Rosen LD, et al. Media Multitasking and Cognitive, Psychological, Neural, and Learning Differences. *Pediatrics* 2017;140(Suppl 2):S62–S66; doi: 10.1542/peds.2016-1758D.
- Uversky VN. Natively Unfolded Proteins: A Point Where Biology Waits for Physics. *Protein Sci* 2002;11(4):739–756; doi: 10.1110/ps.4210102.
- Uversky VN. Unusual Biophysics of Intrinsically Disordered Proteins. *Biochim Biophys Acta* 2013;1834(5):932–951; doi: 10.1016/j.bbapap.2012.12.008.
- Uversky VN. Dancing Protein Clouds: The Strange Biology and Chaotic Physics of Intrinsically Disordered Proteins. *J Biol Chem* 2016a;291(13):6681–6688; doi: 10.1074/jbc.R115.685859.
- Uversky VN. P53 Proteoforms and Intrinsic Disorder: An Illustration of the Protein Structure–Function Continuum Concept. *Int J Mol Sci* 2016b;17(11):1874; doi: 10.3390/ijms17111874.
- Vishwanath S, de Brevern AG and Srinivasan N. Same but Not Alike: Structure, Flexibility and Energetics of Domains in Multi-Domain Proteins Are Influenced by the Presence of Other Domains. *PLoS Comput Biol* 2018;14(2):e1006008; doi: 10.1371/journal.pcbi.1006008.
- Waudby CA, Dobson CM and Christodoulou J. Nature and Regulation of Protein Folding on the Ribosome. *Trends in Biochemical Sciences* 2019;44(11):914–926; doi: 10.1016/j.tibs.2019.06.008.

- Wright PE and Dyson HJ. Intrinsically Unstructured Proteins: Re-Assessing the Protein Structure-Function Paradigm. *J Mol Biol* 1999;293(2):321–331; doi: 10.1006/jmbi.1999.3110.
- Wright PE and Dyson HJ. Intrinsically Disordered Proteins in Cellular Signaling and Regulation. *Nat Rev Mol Cell Biol* 2015;16(1):18–29; doi: 10.1038/nrm3920.
- Xue B, Brown CJ, Dunker AK, et al. Intrinsically Disordered Regions of P53 Family Are Highly Diversified in Evolution. *Biochim Biophys Acta* 2013;1834(4):725–738; doi: 10.1016/j.bbapap.2013.01.012.
- Yi Q, Liu W, Seo JH, et al. Discovery of a Small-Molecule Inhibitor Targeting the Androgen Receptor N-Terminal Domain for Castration-Resistant Prostate Cancer. *Molecular Cancer Therapeutics* 2023;22(5):570–582; doi: 10.1158/1535-7163.MCT-22-0237.
- Zamora-Briseño JA, Pereira-Santana A, Reyes-Hernández SJ, et al. Towards an Understanding of the Role of Intrinsic Protein Disorder on Plant Adaptation to Environmental Challenges. *Cell Stress and Chaperones* 2021;26(1):141–150; doi: 10.1007/s12192-020-01162-5.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.