*Review*

# In the Beginning: Let Hydration Be Coded in Proteins for Manifestation and Modulation by Salts and Adenosine Triphosphate

**Jianxing Song** ORCID

Department of Biological Sciences, Faculty of Science, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260, Singapore; dbssjx@nus.edu.sg or songjianxing1000@gmail.com; Tel.: +65-65161013; Fax: +65-67792486

**Abstract:** Water exists in the beginning and hydrates all matter. Life emerged in water, requiring three essential components in compartmentalized spaces: (1) universal energy sources driving biochemical reactions and processes, (2) molecules that store, encode, and transmit information, and (3) functional players carrying out biological activities and structural organization. Phosphorus has been selected to create adenosine triphosphate (ATP) as the universal energy currency, nucleic acids for genetic information storage and transmission, and phospholipids for cellular compartmentalization. Meanwhile, proteins composed of 20 $\alpha$-amino acids have evolved into extremely diverse three-dimensional forms, including folded domains, intrinsically disordered regions (IDRs), and membrane-bound forms, to fulfill functional and structural roles. This review examines several unique findings: (1) insoluble proteins, including membrane proteins, can become solubilized in unsalted water, while folded cytosolic proteins can acquire membrane-inserting capacity; (2) Hofmeister salts affect protein stability by targeting hydration; (3) ATP biphasically modulates liquid–liquid phase separation (LLPS) of IDRs; (4) ATP antagonizes crowding-induced protein destabilization; and (5) ATP and triphosphates have the highest efficiency in inducing protein folding. These findings imply the following: (1) hydration might be encoded in protein sequences, central to manifestation and modulation of protein structures, dynamics, and functionalities; (2) phosphate anions have a unique capacity in enhancing $\mu$s-ms protein dynamics, likely through ionic state exchanges in the hydration shell, underpinning ATP, polyphosphate, and nucleic acids as molecular chaperones for protein folding; and (3) ATP, by linking triphosphate with adenosine, has acquired the capacity to spacetime-specifically release energy and modulate protein hydration, thus possessing myriad energy-dependent and -independent functions. In light of the success of AlphaFolds in accurately predicting protein structures by neural networks that store information as distributed patterns across nodes, a fundamental question arises: Could cellular networks also handle information similarly but with more intricate coding, diverse topological architectures, and spacetime-specific ATP energy supply in membrane-compartmentalized aqueous environments?

**Keywords:** water; ATP; protein hydration; protein folding; intrinsically disordered regions (IDRs); protein phase separation; cellular networks; artificial intelligence (AI)

## 1. Introduction

The book of Genesis in the Bible begins with the verse: "Now the earth was formless and empty, darkness was over the surface of the deep, and the Spirit of God was hovering over the waters" (Genesis 1:2). Water ($H_2O$) is thought to have formed in the early universe when hydrogen, the simplest and most abundant element, combined with oxygen produced in the interiors of stars, creating water molecules in interstellar space. Although water is abundant in the cosmos, not all planets exist under the right conditions to retain or accumulate it. A planet's ability to host water depends on factors such as its formation

process, proximity to its star, size, and geological history [1]. Covering two-thirds of our planet, water is a unique and ubiquitous substance in our world and has by all accounts mysterious properties that set it aside from other molecules [2]. At least to our experience, life is invariably linked to the presence of water, as the earliest life forms likely originated in aqueous environments [2–4]. Often referred to as the 'matrix of life', water is not merely a passive backdrop but plays active and crucial roles in molecular biology [4].

Hydration is a fundamental and universal phenomenon that is common to all classes of compounds regardless of organic and inorganic chemicals [5–9]. It refers to the interaction between water molecules and other substances, which can occur through various mechanisms such as hydrogen bonding, electrostatic interactions, and van der Waals forces. The extent and nature of hydration depend on the chemical properties of the substances involved, such as polarity, charge, and molecular structure. Inorganic chemicals, particularly salts, ions, and minerals, often exhibit hydration through the formation of hydration shells around ions [5,6]. Organic molecules can also undergo hydration, particularly those with functional groups capable of hydrogen bonding or those that are polar. Hydration can significantly influence the solubility, stability, and reactivity of organic molecules in aqueous environments [7–9]. It is well-established that water mediates or even governs numerous vital biological processes such as protein folding, enzyme catalysis, membrane self-assembly, and substrate recognition through the hydration of a wide range of biomacromolecules and small molecules, including salts.

Proteins are the most essential functional players in almost all biological processes within living cells and even viruses [10]. Their multifaceted roles highlight their indispensability for life, encompassing structural support, intricate biochemical regulation, and defense mechanisms. Key functions include the following: (1) Structural Support: collagen, the most abundant protein in the human body, constitutes ~30% of total protein content and forms the extracellular matrix in connective tissues such as skin, tendons, and bones. (2) Enzymatic Activity: enzymes accelerate biochemical reactions by lowering activation energy, enabling reactions to occur millions of times faster. (3) Transport and Storage: proteins facilitate the movement of substances across membranes or within the body. For example, hemoglobin in red blood cells binds oxygen, with ~98.5% of blood oxygen transported via hemoglobin. (4) Immune Defense: proteins are vital to immunity. Antibodies (immunoglobulins), produced by B-cells, exist in billions of unique forms adapted to specific antigens. Complement proteins (~30 types) enhance pathogen destruction via opsonization, lysis, or inflammation. (5) Signaling: proteins function as hormones or receptors, enabling intercellular communication. Insulin, a 51-amino-acid protein hormone, regulates blood glucose levels. Receptors like G-protein-coupled receptors (over 800 types in humans) detect signals and initiate intracellular responses. (6) Movement: motor proteins, such as myosin and actin, drive cellular and organismal movement. In muscles, myosin heads interact with actin filaments, generating contraction forces at a rate of 5–6 ATP molecules per second. (7) Regulation: proteins modulate gene expression and metabolic pathways. For example, transcription factor p53, a tumor suppressor, regulates genes in response to stress, influencing apoptosis, the cell cycle, and DNA repair. (8) Energy Source: proteins yield ~4 kcal/g of energy when metabolized, particularly during fasting or intense exercise. (9) Buffers: proteins like hemoglobin help maintain pH homeostasis, binding hydrogen ions to stabilize blood pH around 7.4 [10].

As illustrated in Figure 1A, proteins are composed of 20 α-amino acids, each with distinct physicochemical properties. Based on their amino acid sequences, proteins encoded by higher eukaryotic genomes can be categorized into two main types [11]: those with random, high-complexity sequences (I of Figure 1A) and those with non-random, low-complexity sequences (II of Figure 1A). Interestingly, in cell-like buffers containing high salt concentrations, only a subset of proteins within the first category can fold into well-defined, soluble structures [12–15]. Conversely, many proteins in the second category remain fully functional despite lacking well-defined structures, and are thus referred to as intrinsically disordered proteins (IDPs) [15–17]. Additionally, approximately 25–30% of proteins are

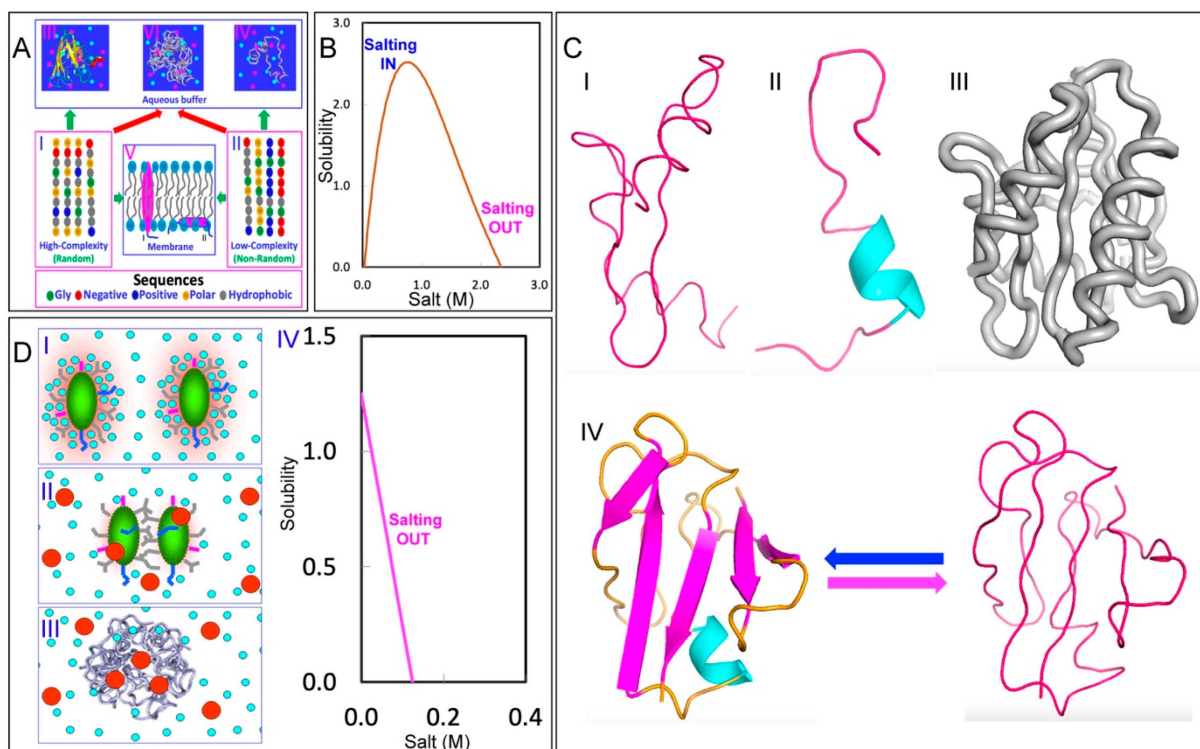associated with phospholipids [18], earning them the designation of membrane proteins (MPs) (Figure 1A).



**Figure 1.** The effect of salts on protein solubility beyond "salting-in/salting-out". (**A**) Sequence–structure relationship of proteins. Based on the sequences as represented by five types of amino acids, proteins can be classified into high complexity or random (I), and low-complexity or non-random (II) sequences. A portion of proteins of the high-complexity sequence can fold into uniquely folded structures soluble with high concentrations of salts (III), while many of proteins of the low-complexity sequence remain intrinsically disordered (IV). Furthermore, ~30% proteins of both high- and low-complexity sequences can fold in the membrane environments (V). Interestingly, a large number of proteins of both high- and low complexity sequences appear to be insoluble in vivo with high concentrations of salts (VI). (**B**) The classic curve of protein solubility versus salt concentration for a well-folded protein which consists of "salting-in" and "salting-out" phases. (**C**) Solution conformations of insoluble proteins solubilized in unsalted water. Four groups of conformations have been observed so far: (I) highly disordered state without any stable secondary and tertiary structures; (II) partially folded state with some secondary but without tertiary structures; (III) molten globule-like state with secondary structures as well as dynamic tertiary packing; and (IV) coexisting folded and unfolded states exchanging in equilibrium. (**D**) A model to rationalize how salt ions affect the solubility of an insoluble protein. (I) An unrefoldable and insoluble protein in unsalted water with the solution pH several units away from its pI. Small cyan spheres stand for water molecules and green ellipsoids for protein molecules with a large number of hydrophobic side chains exposed. (II) Protein molecules in the presence of a small number of salt ions (larger red spheres). In addition to imposing non-specific electrostatic screening, the presence of salts also provides specific anion binding to protein residues, thus altering the surface electrostatic potential. Furthermore, salts also change the hydration structure as well as protein dynamics, as represented by the broken lines of green ellipsoids. (III) The complex interplay of these salt effects results in aggregation of the protein. (IV) The unique curve of protein solubility versus salt concentration for an insoluble protein. Figure 1A was adapted from Figure 1a in Ref. [19].

Water has been proposed to "slave" proteins, implying that the dynamic behavior of proteins is intricately coupled with the dynamics of their surrounding water molecules,

specifically the hydration shell [20–24]. Protein hydration involves the interaction of water molecules with the protein's surface. These water molecules form hydrogen bonds with polar and charged groups, engage in hydrophobic interactions, and significantly influence protein folding, stability, and dynamics. Through hydrogen bonding, hydrophobic interactions, electrostatic screening, and dynamic exchange, water plays a crucial role in controlling protein behavior, facilitating folding, conformational changes, and allosteric regulation. This intricate interplay underscores the essential role of water in the life and functionality of proteins [25–29].

Nevertheless, despite exhaustive studies, it still remains extremely challenging to understand protein hydration using both experimental and computational methods due to the complexity and dynamic nature of the interactions between water molecules and protein surfaces. These complexities include the following: (1) the dynamic and heterogeneous nature of hydration, which involves a rapidly changing network of water molecules that can form and break hydrogen bonds in picoseconds. In this regard, techniques such as nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography provide only static or averaged views, making it difficult to capture the full dynamic picture [20]. On the other hand, simulating protein hydration requires capturing fast and localized interactions between water and protein, necessitating a high temporal and spatial resolution. Molecular dynamics (MD) simulations must balance accuracy with computational feasibility [30]. (2) Complex interactions and multiplicity of states. Therefore, different techniques often provide complementary but not complete information. For instance, X-ray crystallography can pinpoint positions of water molecules tightly bound to proteins in crystal structures, but misses the dynamically interacting water molecules in solution [20]. For simulations, its accurate modeling requires considering many-body interactions, the polarizability of water molecules, and long-range electrostatic effects, which are computationally demanding [20]. (3) Influence of environment and conditions. As protein hydration is sensitive to environmental conditions such as temperature, pH, and ionic strength, reproducing physiological conditions accurately in experiments can be challenging [30,31]. Simulating the exact conditions found in vivo, including the presence of cosolutes and crowding effects, adds layers of complexity to computational models [32]. (4) Scale and resolution limitations. Techniques like neutron scattering and infrared spectroscopy provide insights into hydration but often lack the resolution to distinguish between water molecules bound to specific protein sites [31,33]. MD simulations require balancing fine-grained details (e.g., quantum effects in hydrogen bonding) with the need to simulate large systems over biologically relevant timescales [30].

This review examines several unique findings on protein structures, membrane-interactions, dynamics, and liquid–liquid phase separation (LLPS) and aggregation, which appear to be fundamentally related to protein hydration and its modulation by salt ions and ATP: (1) insoluble proteins, including membrane proteins, can be solubilized in unsalted water, while folded cytosolic proteins can be "unlocked" to release aggregation-prone and membrane-inserting regions; (2) kosmotropic and chaotropic salts affect protein stability and dynamics critically by altering hydration; (3) ATP biphasically modulates LLPS of Arg-/Lys-containing intrinsically disordered regions (IDRs); (4) ATP antagonizes crowding-induced protein destabilization likely by mediating protein hydration; (5) ATP and triphosphates have the highest efficiency in inducing protein folding. The analysis suggests the following: (1) hydration appears to be coded in protein sequences, being central to manifestation and modulation of protein structures, dynamics, and functionalities; (2) phosphate anions have a unique capacity in enhancing μs-ms dynamics of proteins, likely through ionic state exchanges in the hydration shell, underpinning ATP, polyphosphate, and nucleic acids as molecular chaperones for protein folding; and (3) ATP, by linking triphosphate with adenosine, has acquired the capacity to spacetime-specifically release energy and modulate protein hydration, thus possessing myriad energy-dependent and -independent functions.

Finally, three key challenges are discussed in this review. First, molecular mechanisms by which hydration is coded into protein sequences and how protein hydration is modulated remain to be elucidated. Second, in light of the central role of the phosphate-containing molecules in cellular processes, understanding how phosphate ions affect their structures, dynamics, and interactions is essential. In particular, investigations are needed to address whether phosphate anions may govern the dynamics of nucleic acids and phospholipids, which may thus possess currently unknown functions. Third, when considering the success of AlphaFold, a neural-network-based AI system that accurately predicts protein structures by encoding information as distributed patterns across its network nodes, a fundamental question arises: Could cellular networks, which are fundamental to life, also manage information in a similar distributed manner? In particular, cellular networks are far more complex, involving intricate coding mechanisms and diverse topological architectures, and are driven by spacetime-specific ATP energy supplies in membrane-compartmentalized aqueous environments. In this context, could quantum principles be at play, thus leading to the emergence of phenomena such as consciousness.

## 2. Unusual Salt Effect on Protein Solubility Beyond "Salting-In" and "Salting-Out"

Protein solubility is essential not only for researchers and industries that work with proteins in solution, such as structural biologists and the pharmaceutical sector, but also because it is intimately linked to protein aggregation and amyloidogenesis, which are universal hallmarks of an increasing spectrum of human diseases and aging [34–41]. These diseases include neurodegenerative disorders like Alzheimer's disease (AD), Parkinson's disease (PD), Huntington's disease (HD), spinocerebellar ataxias (SCA), amyotrophic lateral sclerosis (ALS), and frontotemporal dementia (FTD) [37–39], and even extend to the aging process [40], affecting organisms as diverse as humans and *E. coli* cells [41]. Thermodynamically, solubility is defined as the concentration of protein in a saturated solution that is in equilibrium with its solid phase, whether crystalline or amorphous, under specific conditions. Protein solubility is influenced by both extrinsic and intrinsic factors. Extrinsic factors include pH, ionic strength, temperature, and the presence of solvent additives, which can sometimes be adjusted to improve solubility. However, intrinsic factors, primarily the structures and physicochemical properties on the protein's surface, also play a significant role in determining solubility [42–45].

Among the earliest recognized extrinsic factors affecting protein solubility are salts, which are ionic compounds composed of positively charged cations and negatively charged anions, held together by electrostatic forces. When salts dissolve in water, they dissociate into their constituent ions, separating into positively charged cations and negatively charged anions. Salts are quite prevalent and important among the chemicals on Earth due to their various roles in natural processes, industry, and daily life. Many salts naturally exist as minerals, and seawater is especially abundant in salts, with sodium chloride (NaCl) being the most common. Seawater contains approximately 3.5% dissolved salts by weight [46]. In many natural environments, inorganic salts are more prevalent than organic ones. Oceans, soils, and rocks are predominantly composed of inorganic salts. Although organic salts are less common, they are vital in biological systems, functioning as electrolytes and aiding in osmoregulation within the body [47,48].

In 1988, Franz Hofmeister published a landmark study on the influence of salts on protein solubility, laying the groundwork for the concepts of the "salting-in" and "salting-out" effects, as well as the Hofmeister series [49]. In this pioneering study, Hofmeister observed a universal phenomenon: the solubility of proteins in aqueous solutions could be dramatically influenced by salts in a dual manner. At low concentrations (usually <300–500 mM), salts enhance protein solubility (salting-in effect), but reduce it at higher concentrations (salting-out effect), as illustrated in Figure 1B. Hofmeister's research established the first phenomenological theory for the effects of salts on protein solubility, leading to a dogmatic concept in biochemistry that guides research and practical applications in protein chemistry [50–54].

Intriguingly, a notable subset of proteins appears to be completely insoluble under various buffer conditions, as evidenced by structural genomics projects that cloned and expressed proteins on a proteome-wide scale, but found that nearly half of them were insoluble [55]. This issue is further reflected in eukaryotic cells, which struggle with protein aggregation even under normal, unstressed conditions. Remarkably, studies on human cell lines have estimated that approximately 30% of newly synthesized proteins aggregate and are swiftly degraded by proteasomes [56–61]. These proteins likely fail to achieve their native structures due to errors in translation or in post-translational modifications critical for proper folding [56–59]. However, the intrinsic factors underlying this complete insolubility remained unknown previously.

In 2005, I was inspired to consider that these completely insoluble proteins could actually be solubilized in unsalted water [61]. This idea was later confirmed by the study on 11 proteins that were previously deemed unrefoldable and insoluble in various buffers. Marvelously, all of them were successfully solubilized in unsalted water, allowing high-resolution NMR investigations [62]. The upper limit of this discovery was further explored by examining a 25-residue integral membrane peptide from the influenza M2 channel, one of the most hydrophobic protein sequences known in nature. To the astonishment, this peptide could also be solubilized in unsalted water, forming a highly helical conformation even in the absence of any lipid molecules [63]. Since this discovery, "completely insoluble" mutants and proteins have been characterized, which are implicated in human diseases such as cancers and ALS [63–73]. Later, other research groups have also shown that proteins that were previously considered to be insoluble could, in fact, be solubilized in unsalted water [74–76]. Notably, all proteins in the total cellular extract of human cells, including the 30% that are membrane proteins, have been shown to be highly soluble in pure water when the mixture is free of nucleic acids [76]. These results together logically suggest that despite existing in the background, salts play a decisive role in controlling protein aggregation [63].

This discovery allowed us to characterize the high-resolution conformations of these 'insoluble proteins' by NMR spectroscopy, which led to the classification of them into four distinct groups (Figure 1C): Group 1 includes proteins that are highly disordered and lack both secondary and tertiary structures [62,64,69–71,73,74]; Group 2 consists of proteins that have some helical secondary structures but no tertiary packing [62,65,75]; Group 3 comprises proteins that possess secondary structures to some degrees along with loose tertiary packing, resembling molten globule states [62,63,77]; and Group 4 involves proteins that coexist between the folded and unfolded states, undergoing dynamic exchanges on a millisecond timescale [67,68,72]. Remarkably, all these insoluble proteins have been shown to lack tight tertiary packing.

The results thus revealed a previously unrecognized regime in protein behaviors, where the classic "salting-in/salting-out" principle appears to only apply to well-folded proteins. In contrast, unrefoldable and insoluble proteins, including the most hydrophobic integral membrane protein fragments, can only be solubilized in water with minimal salt concentrations. In 2008, I proposed a model (Figure 1D) to rationalize why these unrefoldable and insoluble proteins remain soluble in unsalted water but aggregate upon the introduction of even small amounts of salt ions [61–63]. Briefly, the lack of a tight tertiary structure of these proteins exposes a considerable number of hydrophobic side chains to the surrounding water. When dissolved in water with the minimal salt ions and a pH differing from their isoelectric point (pI), these proteins acquire a significant number of net charges, thus leading to strong electrostatic repulsion between molecules and a substantial hydration shell, which form an energy barrier to inhibit intermolecular interactions. As a result, protein aggregation is largely suppressed in the absence of salts (I in Figure 1D). However, the addition of even small amounts of salt ions can shield these repulsive electrostatic forces and partially disrupt the hydration shell (II in Figure 1D), thus allowing hydrophobic interactions to predominate, which result in pronounced protein aggregation (III in Figure 1D), eventually manifesting as the salting-out phase (IV in Figure 1D).

This model suggests that DRiPs [56–59], which are "intrinsically insoluble" in vitro, may also inevitably aggregate in vivo due to their inability to fold into stable tertiary structures and the high ion concentrations in cellular environments. Unlike "misfolded proteins", the aggregation of these "intrinsically insoluble proteins" cannot be prevented by chaperone systems. Therefore, despite the metabolic cost, cells likely degrade these proteins immediately after synthesis to minimize potential damage. If not effectively degraded, their aggregation could trigger various human diseases, including cancers and neurodegenerative disorders, and even contribute to aging.

Protein hydration appears to play a crucial role in solubilization of these insoluble proteins in unsalted water, as well as aggregation or amyloid formation upon the introduction of salts. In unsalted water, the disordered structures of these proteins expose their polypeptide chains extensively, leading to significant hydration. The interaction with a large number of water molecules creates a high energy barrier that prevents the transition to aggregation [78–82]. However, when salt ions are introduced, they neutralize the repulsive forces and disrupt the hydration shell, driving the proteins to form aggregates with a reduced water content [82–84]. In this process, the release of ordered water molecules from their association with the proteins into the bulk solvent increases the entropy of the water, which in turn entropically promotes aggregation.

This discovery offers a potential solution to the longstanding 'chicken-and-egg paradox' regarding the origin of integral membrane proteins. Specifically, it has been puzzling how these highly hydrophobic, water-insoluble proteins could have reached the early membranes, as "even if occasionally synthesized, they would likely remain stuck in the ribosome" [85]. However, the model suggests a resolution: emerging evidence indicates that proteins and primitive membranes containing integral membrane proteins may have emerged in unsalted, slightly acidic prebiotic oceans [85,86]. Remarkably, these conditions closely resemble the aqueous conditions used to solubilize insoluble or membrane proteins, namely largely unsalted and mildly acidic [87–96]. Consequently, as illustrated in I of Figure 2A, even the most hydrophobic integral membrane peptides would not be trapped in the early protein-synthesizing machinery. Instead, they would be soluble in this prebiotic, unsalted medium. Furthermore, the concentrations of most proteins in primitive oceans were likely very low, allowing them to diffuse freely and encounter primeval lipid molecules. However, the increase in salt concentrations would drive the spontaneous assembly of lipid molecules and hydrophobic peptides, forming primitive membranes associated with different membrane proteins including integral membrane proteins (II of Figure 2A).

This discovery may also offer insights into another longstanding mystery related to protein diversification. It has been suggested that the space of realized protein folds represents only a fraction, about one-tenth of the possible protein fold space [97], indicating that much of the sequence space remains unexplored in Earth's life forms. This limited exploration might be linked to ocean salinity. The primitive machinery responsible for protein generation is believed to have existed before the development of membrane-enclosed cells [98]. In unsalted oceans, proteins with the highly randomized sequences could be synthesized in the primitive machinery, all of which are soluble and could freely diffuse. Because in unsalted water, protein–protein interactions are greatly suppressed due to strong electrostatic repulsion between individual protein molecules, a slight increase in ocean salinity might have reduced these repulsive forces, thus facilitating the formation of various protein-based complexes. However, once membrane-enclosed cells emerged and ocean salinity increased, highly hydrophobic proteins would have been prone to aggregation and could potentially damage membranes. This may have driven the evolution of mechanisms that restricted the random sampling of protein sequence space. In this context, modern proteins, regardless of being well-folded, intrinsically disordered, or membrane-associated, likely evolved from a limited pool of primordial proteins that were randomly generated in unsalted oceans. This hypothesis aligns with the proposal that modern proteins appear designed so that their intrinsic repulsive interactions in pure water are sufficient to

counteract attractive forces, thereby preventing severe precipitation or aggregation [62,63]. Additionally, it is plausible that the 20 natural L-, α-amino acids were selected because their polymerized forms, proteins, were soluble in the primitive unsalted oceans [63].
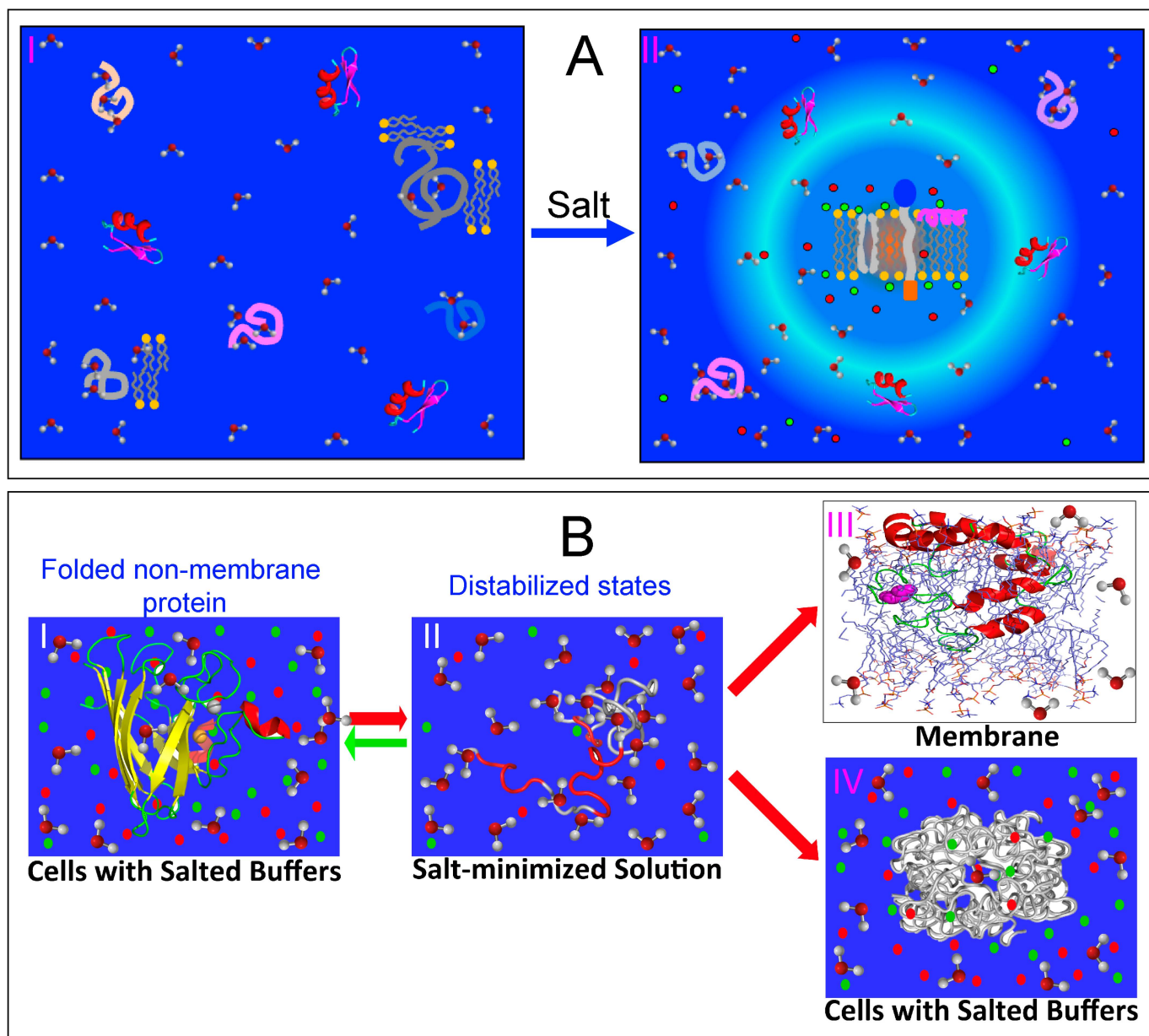


**Figure 2.** Emergence of membrane proteins and transformation of a folded cytosolic protein into a membrane-interacting protein. (**A**) (I) In the primitive water body with minimized salt concentrations, all proteins regardless of sequences are soluble. (II) The increase in the salt concentrations drives highly hydrophobic proteins to form membrane proteins. (**B**) Via disease-associated factors, a folded cytosolic protein (I) can be disrupted into a highly disordered form, which is only soluble in salt-minimized water (II). As this highly disordered form universally contains hydrophobic and/or amphiphilic patches, it acquires the novel capacity to interact with membranes energetically, driven by folding into non-native helix/loop structures in membranes (III), or becomes aggregated in the presence of high salt concentrations (IV).

## 3. Transformation of the Folded Cytosolic Proteins into Membrane-Interacting Proteins

In cells, where aqueous solutions and membrane systems coexist, proteins are classified as membrane or non-membrane proteins. However, it remains unclear whether

non-membrane proteins can transform into membrane proteins. Interestingly, on the other hand, many intrinsically disordered and partially soluble proteins associated with human diseases have been shown to contain regions that can insert into membranes by forming amphiphilic or hydrophobic helices. They include prion protein involved in spongiform transmissible encephalopathies [99], amyloid beta-peptides in Alzheimer's disease [100], tau tangles in Alzheimer's disease [101], α-synuclein in Parkinson's disease [102,103], huntingtin in Huntington's disease [104], and islet amyloid polypeptide in type II diabetes [105]. Previously, a hydrophobic region within the prion-like domain of TDP-43 was identified as capable of inserting into membranes, transitioning from a partially folded helical conformation to a well-defined Ω–loop–helix motif in vitro [69]. Recently, this region has been shown indeed to be associated with mitochondria [106]. Intriguingly, on the other hand, many well-folded non-membrane proteins have been found to convert into unfolded and insoluble forms by genetic mutations or environmental factors, many of which suddenly become cytotoxic and associated with neurodegenerative diseases and aging. However, the mechanisms underlying their gain of cytotoxicity remained poorly understood.

Previous studies on the unrefoldable and insoluble forms of ALS-related proteins, including P56S-MSP, L126Z-SOD1, nascent SOD1, and C71G-Profilin1 [107], revealed that these insoluble forms are either unfolded or coexist with their unfolded states. Most unexpectedly, these unfolded states have been decoded to gain a novel ability to interact with membranes, driven by the formation of helices or loops over amphiphilic or hydrophobic regions. These regions, which universally exist in all folded proteins, are typically locked in their folded native states. However, upon destabilization or disruption of their folded structures by genetic mutations and environmental damage, these regions become exposed completely or dynamically, thus becoming able to interact with membranes. As a consequence, the mutated or damaged proteins gain cytotoxicity by interacting with membranes, with a mechanism that has no fundamental difference from those underlying the membrane-interacting capacities of amyloid beta-peptides, α-synuclein, and even the membrane toxin melittin. These findings suggest that most, if not all, proteins contain segments capable of folding into distinct structures depending on whether they are in aqueous or membrane environments. Abnormal membrane interactions may initiate disease and aging processes, and when coupled with protein aggregation, can lead to severe proteotoxicity. This may result in the formation of inclusions composed of damaged membranous organelles and protein aggregates, further exacerbating cellular damage.

Indeed, ALS-causing SOD1 mutants were extensively shown to associate tightly with the mitochondrial membrane, behaving similarly to integral membrane proteins, with this association being resistant to high ionic strength and high-pH conditions [108–111]. Notably, the abnormal insertion of SOD1 mutants into the ER membrane has been identified as sufficient to induce ER stress, an event that initiates a cascade of cell-specific damage in ALS pathogenesis [110]. A recent study revealed that Lewy bodies are, in fact, a dense mixture of membranous components along with α-synuclein aggregates. The formation of these structures is believed to result from α-synuclein's dual abilities to aggregate and interact with membranes [112]. This suggests that the combined abilities of aggregation and non-specific membrane interaction may be a fundamental mechanism by which aggregation-prone proteins exert high proteotoxicity, leading to cell death by disrupting membranous organelles, as seen in Lewy bodies. In this context, the less specific a protein's membrane interaction, the greater its potential membrane toxicity. This observation explains why aggregation-prone proteins are generally toxic, contributing not only to neurodegenerative diseases but also to the aging process, even in organisms as simple as *E. coli* [41,111–113]. Nevertheless, the most prominent human diseases caused by protein aggregation are neurodegenerative disorders and cardiac dysfunction, likely because neurons and cardiomyocytes are rarely replaced [114]. This suggests that the ability of protein sequences to transform their structure based on the environment may be a general mechanism underlying proteotoxicity.

Therefore, the capacity for membrane interaction appears to be inherently encoded in all well-folded, non-membrane proteins (I of Figure 2B), although it is typically locked within their native folded structures. When these proteins become destabilized or their structures are disrupted due to genetic factors, such as disease-associated genetic variations, including mutations, truncations, or splicing alterations; as well as due to pathological, environmental, or aging-related factors, such as post-translational modifications (PTMs), disulfide bond reduction, cofactor loss, or oxidative modifications/fragmentation, their conformations become disordered to varying degrees. These disordered states are soluble only in salt-free water (II of Figure 2B). Remarkably, they can undergo distinct transitions depending on their environments. For example, they may transform into membrane-associated proteins (II of Figure 2B) upon encountering membranes or aggregate within cells in the presence of saline buffers (IV of Figure 2B). These transitions are tightly regulated by the intricate interplay between proteins, water, salt ions, and phospholipid molecules. Importantly, in all cases, the transition is associated with the release of bound water molecules into the bulk solvent, a process that can entropically drive the transition.

## 4. Novel Insights into the Hofmeister Effects on Protein Structures, Stability, and Dynamics

Franz Hofmeister's pioneering work not only introduced the 'salting-in/salting-out' principle but also led to the development of the Hofmeister series, which categorizes ions based on their influence on protein stability in aqueous solutions [49,54,115–118]. As depicted in Figure 3A, ions that promote protein stability are referred to as kosmotropes, while those that decrease stability are known as chaotropes. Strong kosmotropes include $Na_2SO_4$ and $Na_2HPO_4$, whereas GdmCl and NaSCN are strong chaotropes. NaCl, positioned in the middle of the series, is considered neutral. Remarkably, Hofmeister effects are observed across numerous fields, including medicine, biology, chemistry, and industrial science. Despite their widespread relevance, the precise microscopic mechanisms underlying the Hofmeister series remain poorly understood. It is generally thought that these effects result from intricate and specific interactions between ions and proteins, as well as between ions and the water molecules surrounding proteins, which influence protein hydration. However, the extreme complexity of systems involving ions, counterions, solvents, and co-solutes, each playing distinct roles, presents a significant challenge in uncovering the detailed microscopic mechanisms involved. Indeed, the morphological polymorphism of aggregation and amyloid fibrillation appears to further depend on other factors such as cation valency and ionic strength [119,120].

NMR spectroscopy is a powerful tool for characterizing protein conformation, stability, binding, and dynamics [121–126], providing valuable insights into the mechanisms by which Hofmeister series ions influence proteins. However, two key challenges arise: (1) At low concentrations, hydrophobic side chains in most proteins remain buried and inaccessible to ions until global unfolding occurs. (2) Upon global unfolding, however, significant changes in protein NMR resonances can result from both conformational changes and ion–protein interactions, making it difficult to distinguish between these two effects. To address these challenges, a folded 39-residue WW4 domain, lacking any cysteine residues (Figure 3B), was recently chosen to evaluate the effects of five salts—neutral NaCl, kosmotropic $Na_2SO_4$ and $Na_2HPO_4$, as well as chaotropic GdmCl and NaSCN—on its conformation, thermal stability, binding, and backbone dynamics at low salt concentrations (≤200 mM) using CD and NMR spectroscopy [124,125].
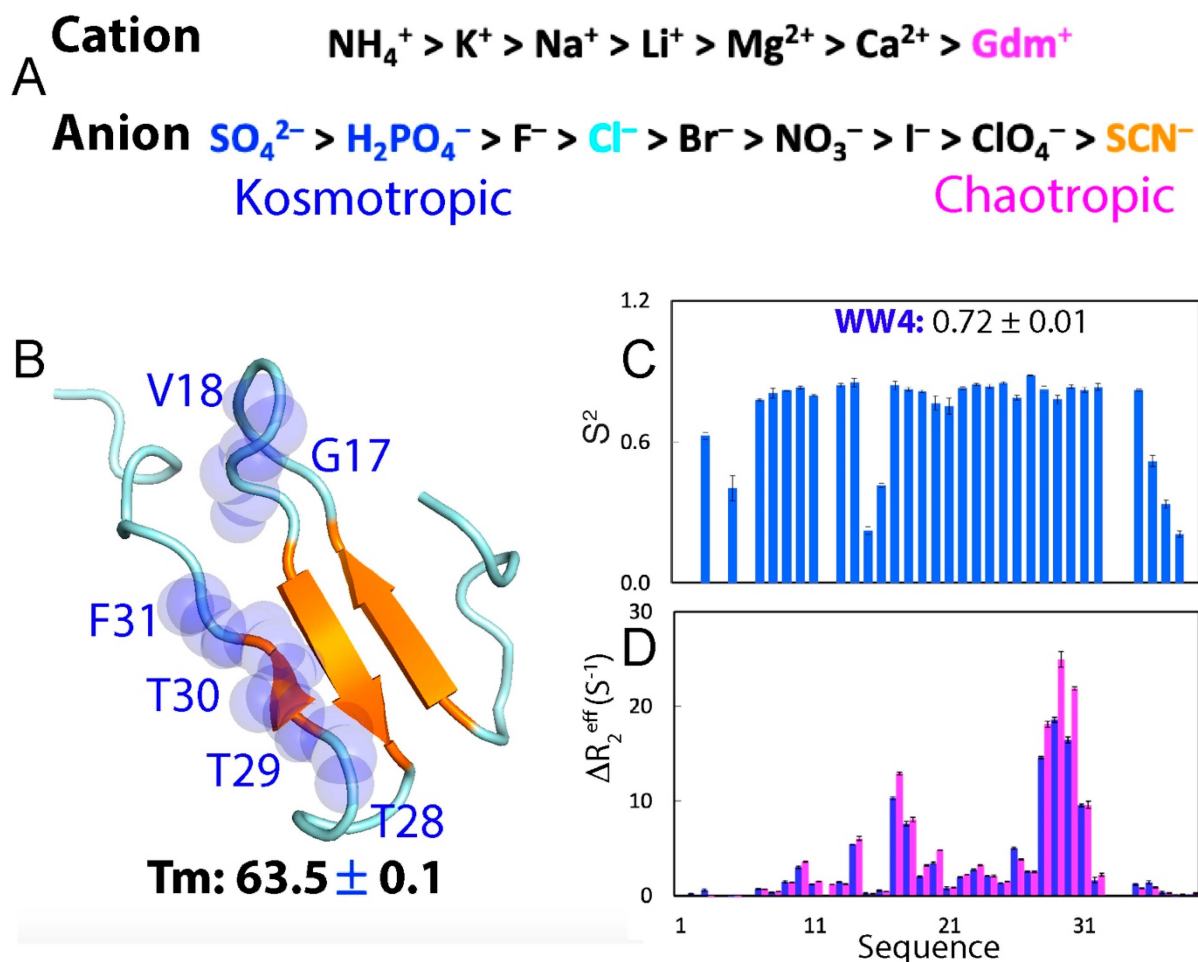
**Figure 3. The Hofmeister effects on the WW4 domain.** (**A**) The Hofmeister series of common cations and anions. (**B**) NMR structure of the WW4 domain with the melting temperature (Tm) labeled, in which 6 residues with significant μs-ms dynamics are displayed in spheres. (**C**) Squared generalized order parameters ($S^2$) of WW4. (**D**) Differences in effective transverse relaxation rates ($\Delta R^2_{eff}$) at 80 and 960 Hz, for WW4 domain at 500 MHz (blue) and 800 MHz (purple). Figure 3B was adapted from Figure 7 in Ref. [125].

The WW4 domain from the ubiquitin ligase WWP1 adopts the classic WW fold [126] and offers several significant advantages for studying ion–protein interactions: (1) Unlike larger folded proteins with buried hydrophobic side chains, WW4 consists of a flat, three-stranded β-sheet, making its hydrophilic and hydrophobic residues largely accessible even in the native state (Figure 1B). This accessibility allows WW4 to interact with salt ions even at low concentrations, before global unfolding occurs. As a result, any binding events can be detected by NMR spectroscopy, which is sensitive enough to resolve residue-specific interactions and perturbations over a wide range of affinities, including dissociation constants (Kd) in the millimolar range [120–128]. (2) Despite its small size, WW4 exhibits remarkably high thermodynamic stability, undergoing reversible thermal unfolding with a melting temperature (Tm) of 63.5 °C. This Tm is higher than that of many larger folded proteins, such as the 87-residue RRM domain of FUS, which has a Tm of ~52 °C [127], and the 140-residue human profilin 1 (hPFN1), with a Tm of ~56 °C [128]. (3) Regarding ps-ns backbone dynamics, most non-Proline residues in WW4 show order parameters ($S^2$) greater than 0.7, with an average value of 0.72, indicating that WW4 is well-folded (Figure 3C). The exceptions are the terminal residues Asn1-Leu5 and Asn36-Ser39, as well as loop residues Arg15-Glu16. Additionally, six residues (Gly17-Val18 and Thr28-Thr30-Phe31) display significant μs-ms backbone dynamics (Figure 3D), undergoing exchanges

with a minor conformation that constitutes less than 10% of the population. This minor conformation differs from the major state by less than 0.5 ppm in the backbone $^{15}N$ chemical shift, suggesting that the two exchanging conformations are highly similar [124].

The effects of five salts on the conformation, thermal stability, binding, and backbone dynamics of the WW4 domain were quantified using CD and NMR spectroscopy, as summarized in Figure 4A: (1) Up to a concentration of 200 mM, none of the five salts caused detectable changes to the tertiary structure of WW4. However, they had differential effects on its thermal stability. GdmCl and NaSCN destabilized WW4, reducing the melting temperature (Tm) by ~9.0 °C and ~3.2 °C, respectively [124]. In contrast, $Na_2SO_4$ and $Na_2HPO_4$ stabilized WW4, increasing the Tm by ~5 °C, while NaCl had no effect [125]. These results perfectly align with the classic Hofmeister series ranking (Figure 3A). (2) At the microscopic level, the binding profiles of the five salts to WW4 varied greatly. The cations sodium ($Na^+$) and guanidinium ($Gdm^+$) showed no detectable binding to the four aliphatic hydrophobic side chains, with only minor binding to amide protons. Among the five anions, only $SCN^-$ bound to aliphatic hydrophobic side chains. The four anions exhibited distinct binding profiles to amide protons. For instance, $Cl^-$ weakly bound to the Arg27 amide proton, while $SCN^-$, $SO_4^{2-}$, and $HPO_4^{2-}$ bound to different sets of residues with varying affinities: $SCN^-$ bound to Arg27, Thr28, Thr29, and Thr30; $SO_4^{2-}$ bound to Arg27 and Phe31; and $HPO_4^{2-}$ bound to Trp9, His24, and Asn36. Notably, only $SCN^-$ and $SO_4^{2-}$ showed saturable binding profiles, with $SO_4^{2-}$ having a higher affinity than NaSCN. (3) In terms of $^{15}N$ backbone dynamics, all five salts had no significant effect on ps-ns time scale dynamics up to 200 mM. However, their effects on the μs-ms time scale were highly distinct. NaCl and $Na_2SO_4$ had no notable impact, GdmCl reduced μs-ms dynamics, and NaSCN and $Na_2HPO_4$ greatly enhanced them. Remarkably, even at 20 mM, $Na_2HPO_4$ significantly increased μs-ms dynamics. Given that NaCl had no effect, the reduction in μs-ms dynamics by GdmCl is likely attributable to the guanidinium cation. Notably, the chloride anion was inert in terms of all these effects. Overall, the effects of the five salts on WW4 are consistent with previous observations in the field: although the ranking of salt effects on thermodynamic stability is consistent with the Hofmeister series across proteins, the underlying microscopic mechanisms vary significantly [49–54,115–119].

These findings prompt us to ask a fundamental question: What underlying mechanisms drive this unique phenomenon for the salts' effects on protein stability? Protein thermodynamic stability is determined by a complex interplay of several factors. Among these, the hydrophobic effect, electrostatic interactions (such as hydrogen bonds and salt bridges), and the formation of hydration shells through interactions with water play critical roles. Each protein's stability emerges from a distinct combination of these factors, creating a finely tuned equilibrium. Ultimately, this stability reflects a balance between favorable enthalpic contributions, arising from covalent bonds and non-covalent interactions, and unfavorable entropic contributions, which are largely due to the reduction in conformational entropy that occurs upon folding.

The interaction of salt ions with proteins is primarily influenced by two key factors: charge density and degree of hydration [6,50,115–119]. As shown in Figure 4A, the $Gdm^+$ cation and $SCN^-$ anion have large ionic volumes but low charge numbers, resulting in low charge densities. Additionally, they exhibit weak hydration, with only 4–6 water molecules surrounding $Gdm^+$ and 2–4 water molecules surrounding $SCN^-$. Consequently, both ions are thought to disrupt the hydration structure of proteins. In contrast, the $SO_4^{2-}$ and $PO_4^{3-}$ anions have both large ionic volumes and high charge numbers, leading to high charge densities. These anions also demonstrate significant hydration, with 12 water molecules surrounding $SO_4^{2-}$ and 16–20 water molecules around $PO_4^{3-}$. Therefore, both ions are considered to stabilize the hydration structure of proteins. The $Cl^-$ anion, meanwhile, has a medium ionic volume and low charge number, resulting in medium charge density, and exhibits a moderate hydration degree with 6–7 water molecules surrounding it.
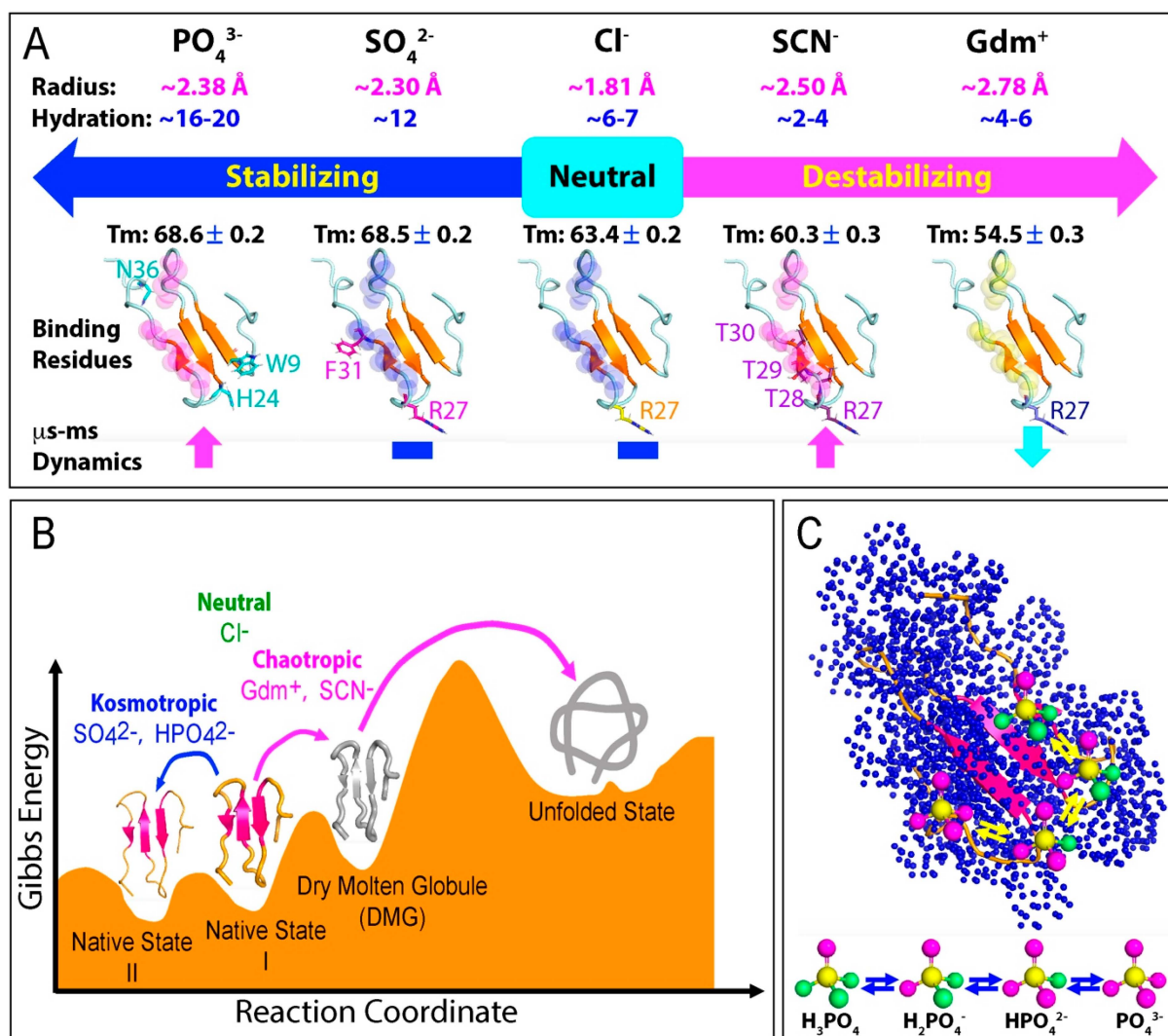
**Figure 4. Proposed mechanisms for the Hofmeister effects.** (**A**) The effects of the four anions and the Gdm+ cation on the conformation, thermodynamic stability, binding residues, and backbone dynamics of the WW4 domain. The ions are ranked based on the Hofmeister series, with the ionic radius and water molecule numbers indicated. The thermodynamic stability is reported by the melting temperature (Tm) values of WW4 in the absence and in the presence of, respectively, the five salts at 200 mM. In the structure of WW4, the spheres are utilized to indicate six residues: Gly17, Val18, Arg27, Thr28, Thr29, and Thr30, with significant μs-ms backbone dynamics ($\Delta R^2_{eff} > 4$ Hz at the 500 MHz field), while the sticks are used to show the residues with significant binding as defined by the chemical shift difference (CSD) in the presence of, respectively, the five salts at 200 mM. The changes in μs-ms backbone dynamics upon adding, respectively, the five salts at 200 mM are also indicated. (**B**) A diagram illustrating the effects of five salts on the folding and stability of the WW4 domain. (**C**) The proposed mechanism for the phosphate anion to provoke μs-ms dynamics of the WW4 domain by exchanges of the ionic states on a μs-ms time scale. Figure 4A was adapted from Figure 7 in Ref. [125].

In this context, the effects of the five salts on the WW4 domain can be explained as follows (Figure 4A,B): NaCl is inert and has no impact on stability, while both the Gdm$^+$ cation and Cl$^-$ anion exhibit only minor interactions with WW4. The pronounced destabilizing effect of GdmCl can thus be attributed to the Gdm$^+$ cation, which significantly disrupts the hydration structure of WW4, leading to reduced μs-ms backbone dynamics. Similarly, NaSCN destabilizes WW4 primarily through the SCN$^-$ anion, which breaks down the hydration structure. Additionally, the SCN$^-$ anion interacts extensively with hydrophobic

side chains, disrupting critical hydrophobic interactions essential for protein thermodynamic stability. This disruption is anticipated to increase μs-ms backbone dynamics, akin to the elevated dynamics observed by NMR in a 37-residue small protein upon pH- and temperature-induced disruption of its hydrophobic core [129–131]. In contrast, both $SO_4^{2-}$ and $PO_4^{3-}$ anions exhibit high charge densities and hydration degrees, suggesting that $Na_2SO_4$ and $Na_2HPO_4$ significantly enhance the thermal stability of WW4 by stabilizing its hydration structure. Furthermore, these two anions can interact extensively with WW4 residues but not hydrophobic side chains. Notably, $Na_2SO_4$ did not significantly affect the μs-ms backbone dynamics of WW4, whereas $Na_2HPO_4$ notably increased these dynamics.

In summary, three key factors appear to govern how salts influence protein thermodynamic stability and dynamics: (1) The intrinsic ability of salts to disrupt or stabilize the hydration structure of a protein serves as the primary factor. (2) The capacity of salt ions to bind with protein residues is an additional factor that may vary depending on the specific salt and protein involved. (3) Alterations in protein dynamics likely stem from the interplay between the effects of salts on the protein's hydration structure and the direct interactions between salts and protein residues. However, changes in dynamics can influence thermodynamic stability positively or negatively, depending on the specific salt and protein, likely through entropy–enthalpy compensation [132,133]. As a result, a simple and general correlation between the thermodynamic stability and the backbone dynamics of WW4 at both ps-ns and μs-ms time scales is not evident.

The most surprising finding is that $PO_4^{3-}$ can significantly enhance μs-ms backbone dynamics, despite stabilizing the protein and lacking the ability to bind to hydrophobic side chains like $SCN^-$ [124,125]. This raises the question: What could be the underlying mechanism? As illustrated in Figure 4C, phosphate exists in four ionic states in solution, which continuously exchange due to protonation–deprotonation reactions. Notably, these exchanges typically occur at rates on the μs-ms time scale [134]. The phosphate ion's strong ability to interact with water molecules and bind to the amide protons of proteins may increase the frequency of exchange processes among its different ionic states within the hydration shell of the WW4 domain. This could trigger μs-ms conformational exchanges in the WW4 domain. Consequently, because WW4 has many conformational states around the native state, the phosphate anion acts to facilitate the transition from a less stable state to the more stable state, which are separated by relatively low energy barriers (Figure 4B).

However, in the context of amyotrophic lateral sclerosis (ALS)-associated mutants of C71G-PFN1 and nascent hSOD1, which exhibit both folded and unfolded states separated by larger energy barriers, phosphate alone is insufficient to induce folding. As a result, covalently linked polyphosphates, such as diphosphate, triphosphate, or ATP or ADP, are required to facilitate the folding transition [128]. This mechanism may help explain the richness of various phosphate-containing molecules in intracellular environments, where, in addition to approximately 10 mM of free phosphate, there are around 45 mM of phosphocreatine, 3.5 mM of hexose phosphate, and 3–12 mM of ATP [47,48]. While further studies are needed to explore this hypothesis, these molecules may play a crucial role in enhancing protein stability and promoting the folding of a diverse range of proteins.

## 5. ATP Effectively Antagonizes the Crowding-Induced Destabilization

Adenosine triphosphate (ATP) is the universal energy currency in all living cells, driving numerous biological processes through the hydrolysis of its high-energy bonds (I of Figure 5A), which are thermodynamically favorable but kinetically controlled [47,48]. Its central role includes fueling metabolic reactions, such as glycolysis and oxidative phosphorylation, which are essential for producing the energy required for cellular functions. ATP also plays a pivotal role in signal transduction, where it acts as a substrate for kinases in phosphorylation reactions, thereby regulating various cellular pathways. Moreover, ATP is crucial for active transport mechanisms, including the functioning of ATP-binding cassette (ABC) transporters and sodium–potassium pumps, which maintain ion gradients across cell membranes. ATP is also a precursor for nucleic acid synthesis and is involved

in the synthesis of RNA and DNA. Additionally, ATP serves as a signaling molecule in both intracellular and extracellular environments, influencing processes such as muscle contraction and neurotransmission [47,48].
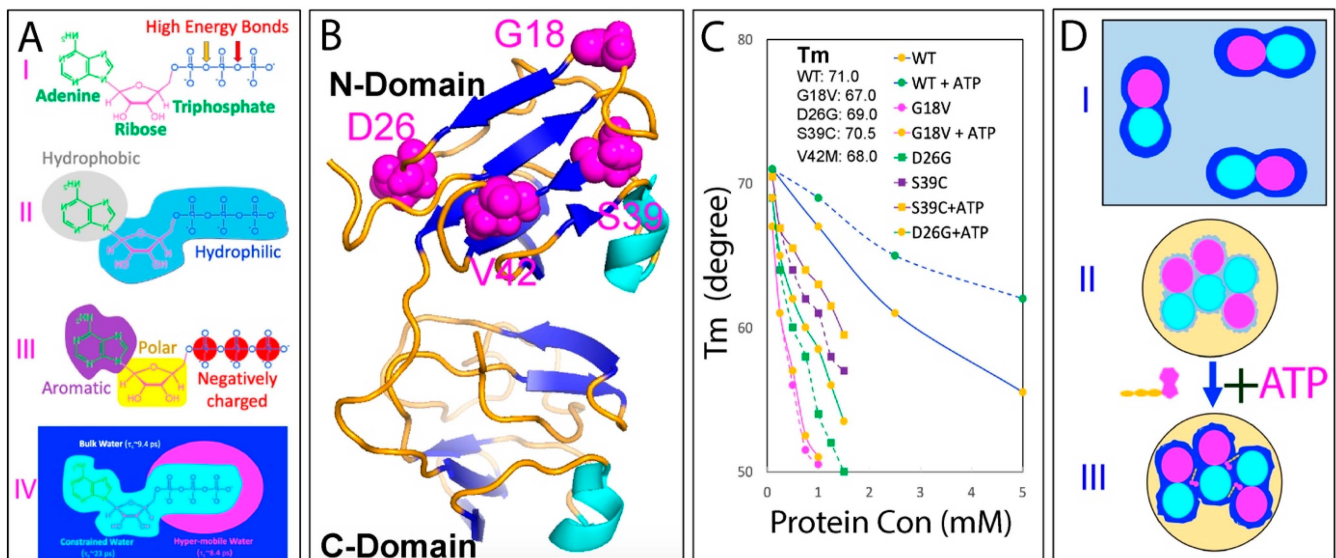


**Figure 5. ATP antagonizes the crowding-induced destabilization of γS-crystallin.** (**A**) ATP has unique structural properties and thus may act as (I) energy currency via the hydrolysis of high-energy bonds; (II) a biological hydrotrope with the presence of hydrophobic adenine and hydrophilic ribose and triphosphate; (III) a bivalent binder via the aromatic purine ring and highly negatively charged triphosphate chain; and (IV) a hydration mediator, resulting from its unique hydration structure, previously derived from the results of microwave dielectric spectroscopy, which was modeled to contain "constrained water" with a dielectric relaxation time (τc) of ~23 ps, as well as "hyper-mobile water" with a τc of ~8.4 ps, even smaller than that of bulk water (9.4 ps). (**B**) Three-dimensional structure of human γS-crystallin with four cataract-causing mutation residues displayed in spheres and labeled. (**C**) Concentration-dependence curves of melting temperature (Tm) of WT, G18V, D26G, and S39C γS-crystallins without and with ATP at 20 mM. (**D**) A speculative model to illustrate that ATP antagonizes the crowding-induced destabilization by targeting protein hydration. Under the extremely crowded condition, the hydration shell (cyan) is disrupted/twisted. However, with the presence of ATP, the disrupted/twisted hydration shell (cyan) is restored to some extent. Figure 5A was adapted from Figure 1b in Ref. [19].

Strikingly, a recent study proposed that ATP may have originated from prebiotic chemical processes even before the polymerization of RNA, DNA, and proteins or the evolution of genetically encoded macromolecular machines. Notably, ATP has been shown to form efficiently only in slightly acidic, unsalted water via the phosphorylation of ADP by acetyl phosphate (AcP), a conserved intermediate linking thioester and phosphate metabolism [96]. Mysteriously, while ATP-dependent enzymes and proteins in modern cells operate at micromolar concentrations, intracellular ATP levels are much higher, typically ranging from 2 to 12 mM depending on the cell type [47,48]. For example, despite being metabolically inactive, the vertebrate lens maintains ATP concentrations between 3 and 7 mM [19,135]. Moreover, although viruses lack ATP [47,48], recent findings reveal that ATP modulates the liquid–liquid phase separation of the SARS-CoV-2 nucleocapsid (N) protein, a critical process in viral infection and replication [136–138]. The ubiquitous presence and conservation of ATP across all biological systems suggest its early emergence in biochemical evolution, potentially influencing fundamental biological processes beyond its traditional roles and shaping essential biological functions and the genome–proteome interface throughout evolutionary history [96,139].

Indeed, ATP has been found to function as a biological hydrotrope at concentrations exceeding 5 mM, dissolving LLPS and protein aggregates or fibrils [140,141]. This newly discovered role is attributed to ATP's amphiphilic nature, comprising a hydrophobic adenine and a hydrophilic triphosphate moiety (II of Figure 5A). Specifically, the aromatic purine ring of ATP is thought to cluster over hydrophobic regions of protein droplets or aggregates, while its triphosphate group interacts strongly with bulk water, thereby facilitating the dissolution of LLPS and aggregates.

Furthermore, NMR analysis has provided residue-specific insights into ATP's interactions with IDRs of FUS, revealing that ATP can biphasically modulate the RGG-rich domain, which lacks large hydrophobic residues and cannot undergo phase separation independently. At low ATP concentrations, it induces LLPS, while at higher concentrations, it dissolves the phase separation [73,142]. This suggests that ATP acts as a bivalent binder, selectively targeting Arg/Lys residues (III of Figure 5A). Additionally, ATP has been shown to bind to conformation-specific pockets in folded proteins, inhibiting amyloid fibrillation either thermodynamically or kinetically [123,143,144]. More recently, even without significant binding or direct effects on the thermodynamic stability at low protein concentrations, ATP has been found to antagonize crowding-induced destabilization of human lens γ-crystallin at extremely high protein concentrations [145–147]. ATP appears to exert these effects by functioning as a hydration mediator (IV of Figure 5A). On the other hand, high-resolution microwave dielectric spectroscopy has revealed that around ATP's triphosphate chain, not only is there constrained water with a dielectric relaxation time longer than that of bulk water, but also an "anomalous" layer of hyper-mobile water with a relaxation time shorter than bulk water [148]. This unique capacity of ATP in increase dynamics of the hydration shell is likely attributed to its triphosphate group, whose different ionic states are continuously exchanging on an s-ms time scale, as proposed for the phosphate group (Figure 4C).

The vertebrate lens is extremely crowded with proteins, reaching concentrations of 200–400 mg/mL in the human lens and exceeding 1000 mg/mL in fish lenses [135,149]. The predominant proteins in the lens are crystallins, which are categorized into three main types—α, β, and γ crystallins—collectively comprising about 90% of the total lens proteins. In the human lens, α-crystallins account for roughly 28%, β-crystallins for 43%, and γ-crystallins for another 28% [149–151]. Functionally, α-crystallins act as molecular chaperones, while β- and γ-crystallins belong to the βγ-crystallin superfamily, serving primarily structural roles. It has been previously shown that α- and β-crystallins can form polydisperse oligomers, whereas γ-crystallins typically exist in a monomeric form in their native state. Notably, because the lens lacks mechanisms for protein repair or recycling, crystallins must remain functional throughout the entire lifespan of an individual. As such, maintaining the solubility and stability of lens crystallins is critical to their physiological role [135,145–147,149–159]. Importantly, genetic mutations or post-translational modifications can lead to the aggregation of crystallins, potentially resulting in cataract formation—an eye disease identified as a priority by the World Health Organization and responsible for 48% of global blindness.

The impact of molecular crowding on protein stability is a fundamental topic in protein science. It has been suggested that the crowded environment within cells, where macromolecules like proteins, nucleic acids, and polysaccharides occupy a significant portion of the available volume, restricts the space available for protein unfolding. This "excluded volume effect" is thought to stabilize the native protein state by limiting the conformational freedom necessary for unfolding [160–164]. However, subsequent studies revealed that macromolecular crowding, both in vitro and in vivo, can also destabilize proteins due to non-specific intermolecular interactions. Thus, the effect of crowding on protein stability is highly context-dependent, influenced by factors such as the specific protein and the conditions within the cellular environment [162–164].

The βγ-crystallins have been evolutionarily optimized for their high solubility, stability, and refractive properties, which are essential for forming the eye lens and its characteristics

of transparency and a refractive index gradient. In the cortex of the human lens, the 178-residue γS-crystallin is the most abundant structural protein. Interestingly, despite γS-crystallin exhibiting weak attractive intermolecular interactions, it remains predominantly monomeric even at very high concentrations. Studies have shown that these attractive interactions are independent of salt concentration, leading to the proposal that this behavior may be attributed to its unique hydration shell [165–167]. The three-dimensional structure of human γS-crystallin, determined by NMR spectroscopy [165], reveals the characteristic four Greek key fold typical of the βγ-crystallin superfamily, organized into distinct N- and C-terminal domains (Figure 5B). Additionally, the packing interactions between these N- and C-domains are thought to play a crucial role in maintaining the protein's high solubility and stability, even in the highly crowded and densely packed environment of the lens.

To date, four cataract-causing mutations have been identified within the N-domain of γS-crystallin: G18V, D26G, S39C, and V42M (Figure 5B) [145–147,166]. Among these, the G18V mutation is associated with childhood-onset cortical cataracts. Interestingly, biophysical studies have shown that the G18V mutant exhibits only a slight reduction in thermodynamic stability compared to the wild-type (WT) protein. However, unlike the WT, the G18V mutant is highly prone to aggregation even at temperatures well below its unfolding threshold [153,167,168]. Structural studies using NMR revealed only localized changes in the G18V γS-crystallin compared to the WT, suggesting that the mutation's aggregation mechanism is more complex than simple mutation-induced denaturation [165]. Notably, recent research has demonstrated that WT γS-crystallin possesses a remarkably robust hydration shell, which allows it to resist the crowding effects even at protein concentrations as high as 400 mg/mL. In contrast, the G18V mutation disrupts this hydration shell, rendering γS-crystallin more susceptible to aggregation, even at low concentrations and temperatures far below its unfolding temperature [153,167,168].

Therefore, understanding the mechanisms that enable crystallins to maintain high solubility and stability throughout life in the crowded environment of the lens is of both fundamental and therapeutic importance. Despite extensive research, several critical aspects remain unresolved. In particular, why does the metabolically inactive lens maintain such high ATP concentrations (3–7 mM), levels that exceed those found in neurons?

Recent studies aimed to investigate whether ATP influences the association, stability, and conformation of both the wild-type (WT) and three cataract-associated mutant forms of human γS-crystallin. These studies characterized the effects of ATP at protein concentrations ranging from 0.1 to 5 mM (~100 mg/mL) using techniques like NMR spectroscopy, dynamic light scattering (DLS), and thermal unfolding [145–147]. Key findings include the following:

1. WT γS-crystallin remains monomeric even at concentrations around 100 mg/mL in buffer containing 150 mM NaCl, mimicking in vivo conditions. At a low protein concentration (0.1 mM), WT is highly stable with a melting temperature (Tm) of 71 °C. However, self-crowding reduces thermal stability, with the Tm dropping to 55.5 °C at 5 mM (Figure 5C).

2. The four cataract-associated mutations (G18V, D26G, S39C, and V42M) significantly increase the tendency for aggregation. Notably, the G18V mutant begins to precipitate at concentrations above 1 mM, while WT only starts forming a gel-like state at concentrations above 5 mM [123]. The other three mutants precipitate at concentrations exceeding 1.5 mM. Despite these aggregation tendencies, at low concentrations (0.1 mM), the mutations only slightly reduce thermal stability, with Tm values of 67.0 °C, 69.0 °C, 70.5 °C, and 68 °C for G18V, D26G, S39C, and V42M, respectively. However, with increasing protein concentrations (self-crowding), the Tm values drop sharply, with G18V falling to 50.5 °C at 1 mM, and D26G, S39C, and V42M decreasing to 50.0 °C, 57.0 °C, and 51.0 °C, respectively, at 1.5 mM (Figure 5C).

3. ATP shows no significant binding to WT or the mutant γS-crystallins, nor does it affect their conformation, even at a high molar ratio of 1:200 (γS-crystallin). ATP only induces minor changes in HSQC peaks, similar to those caused by protein dilution.

4.  ATP also has no detectable effect on αB-crystallin as well as its interactions with WT or mutant γS-crystallins.

5.  Unexpectedly, ATP differentially antagonizes the crowding-induced destabilization of WT and the mutant proteins, even at a molar ratio of 1:1. In WT, the Tm increases from 55.5 °C to 62 °C at 5 mM. Surprisingly, ATP no longer prevents the crowding-induced destabilization of the G18V mutant at 1 mM. For the D26G, S39C, and V42M mutants, ATP still partially antagonizes the crowding-induced destabilization, raising the Tm values from 50.0 °C, 57.0 °C, and 51.0 °C to 53.5 °C, 59.5 °C, and 58.5 °C, respectively, at 1.5 mM (Figure 5C).

These findings suggest that ATP can effectively antagonize the crowding-induced destabilization of WT γS-crystallin, which has a robust hydration shell, but not the G18V mutant, where the hydration shell is disrupted [146,152]. This implies that ATP most likely antagonizes crowding-induced destabilization by interacting directly with the unique hydration shell of human γS-crystallin. The acceleration of the crowding-induced destabilization by cataract-causing mutations appears to result not only from the well-known structural perturbations but also from alterations in the protein's hydration shell. This hydration shell, which plays a central role in protein folding and aggregation, has not been fully appreciated until now.

In light of the current understanding of protein folding and aggregation, a novel mechanism has been proposed to explain how ATP antagonizes the crowding-induced destabilization of γS-crystallin, even at a 1:1 molar ratio. As illustrated in Figure 5D, at low concentrations, γS-crystallin possesses a robust hydration shell that plays a critical role in maintaining its stability (I of Figure 5D). However, in the extremely crowded environment of the lens, even the hydration shell of WT γS-crystallin becomes distorted, with this effect being far more pronounced in cataract-causing mutants (II of Figure 5D). ATP appears to possess the ability to effectively counteract the crowding-induced destabilization by directly interacting with the intact hydration shell of WT γS-crystallin, thereby enhancing its ability to resist the crowding effect. In contrast, the G18V mutation significantly disrupts the intrinsic hydration shell, leaving no possibility for ATP to interact, which explains why ATP loses its stabilizing capacity for this mutation. Thus, in addition to the chaperone activity provided by α-crystallin, ATP at millimolar concentrations is also essential to antagonize the destabilization caused by the extreme crowding of crystallins in the eye lens through this newly identified mechanism. This insight also helps explain the connection between declining ATP levels with aging [169,170] and age-related cataractogenesis, a condition that affects virtually all older individuals. Consequently, restoring normal ATP concentrations in the eye lens could offer a promising therapeutic strategy for treating aggregation-related eye diseases like cataracts.

## 6. ATP Induces Protein Folding with the Highest Capacity

The majority of proteins that require folding from the unfolded state (U) to the folded state (F) for their function are only marginally stable [14,15,171–175]. As a result, genetic mutations and some environmental stresses can easily destabilize these proteins, leading to misfolding and aggregation in cells, particularly in crowded environments [170] and under high salt concentrations [15]. This destabilization is a common pathological feature of aging and neurodegenerative diseases. Currently, it is widely accepted that cells primarily address protein folding and misfolding/aggregation through supramolecular machinery that is energetically powered by ATP [176–178]. However, it remains unclear whether ATP can directly influence the protein folding equilibrium itself, an essential process at the heart of protein homeostasis.

Amyotrophic lateral sclerosis (ALS) is the most common motor neuron disease, first described in 1869, yet its underlying mechanisms remain largely elusive [179]. Approximately 90% of ALS cases are sporadic (SALS), while 10% are classified as familial ALS (FALS). In 1993, human CuZn-superoxide dismutase 1 (hSOD1) was identified as the first causative gene linked to the most common form of FALS [180]. To date, more than

180 mutations have been identified in the 153-residue SOD1 protein [181]. Additionally, misfolding and aggregation of wild-type (WT) hSOD1 have also been extensively associated with SALS [71,180–183]. Subsequently, several mutations of human profilin 1 (hPFN1), a 140-residue protein that plays a crucial role in regulating actin polymerization physiologically [184], have been linked to familial ALS (FALS). The protein adopts a seven-stranded antiparallel β-sheet, which is sandwiched between α-helices at the N- and C-termini on one side and three smaller helical regions on the opposite face (I of Figure 6A). Its C71G mutation with Cys71 replaced by Gly is the most toxic and prone to misfolding and aggregation [107]. As a result, all attempts to determine the crystal structure of this mutant have been unsuccessful [185]. Remarkably, NMR studies have shown that C71G-hPFN1 exists in a dynamic equilibrium between unfolded and folded states, which is characterized by two distinct sets of HSQC peaks (I of Figure 6B). This observation indicates the presence of an energy barrier separating the two states [72,128]. The populations of these states were quantified by NMR, revealing that the folded state comprises approximately 55.2% while the unfolded state accounts for about 44.8%, with an average exchange rate of ~11.7 Hz (~85.5 milliseconds) (II of Figure 6A).

To evaluate the effects of salt ions on the folding of nascent hSOD1 and C71G PFN1, a comprehensive series of common cations and anions were examined using NMR spectroscopy. Among the 12 cations tested, only $Zn^{2+}$ and $Fe^{2+}$ were capable of specifically inducing the folding of nascent hSOD1, resulting in a coexistence of folded and unfolded states; the remaining ions led to precipitation at high concentrations. In fact, $Zn^{2+}$ is a well-recognized cofactor essential for initiating the proper folding of hSOD1, while $Fe^{2+}$ served as the cofactor in the ancestral Fe-SOD1, which shared the same fold as CuZn-SOD1. However, following the Great Oxidation Event, the environmental availability of $Fe^{2+}$ declined, leading to its disappearance in modern organisms. In the case of C71G-hPFN1, a large spectrum of cations and anions, including phosphate, were tested for their ability to promote the conversion of the unfolded state into the folded state. However, none of them succeeded; instead, all led to protein aggregation. These findings suggest that, except for specific cofactors, all salt ions exert non-specific electrostatic screening effects, which in turn promote significant aggregation of nascent hSOD1 and C71G-PFN1.

Very unexpectedly, ATP has been discovered to completely convert C71G-hPFN1 into the folded population even at a ratio of 1:2 (I of Figure 6B) and to transform the nascent hSOD1 into an equilibrium of the folded and unfolded states at 1:8 but without any specific binding to two proteins [128]. To decode the underlying mechanism, the effects of 12 related compounds (Figure 6C) were further visualized by systematic NMR titrations. Notably, ATP and triphosphate were both able to completely convert the unfolded state of C71G-hPFN1 into its folded state at a 1:2 ratio (II of Figure 6B). In contrast, TMAO, a well-known molecule for promoting protein folding [186–188], showed no detectable induction, even at high ratios up to 1:2000, where C71G-hPFN1 instead precipitated. These results suggest that ATP and phosphate exhibits the strongest folding-inducing capacity observed to date.

Although phosphate lacks the ability to induce folding, ADP and diphosphate also exhibit the same folding-inducing capacity, though weaker than that of ATP. The hierarchy of folding-inducing ability is decoded as follows: ATP = ATPP = PPP > ADP = AMP-PNP = AMP-PCP = PP (Figure 6D), while AMP, adenosine, phosphate, TMAO, and NaCl show no inducing effect. Interestingly, results from ATP analogs revealed that very specific structural features are necessary for inducing protein folding. This capacity is not only determined by the number of phosphate groups but also by the atoms linking them. While phosphate cannot induce folding and diphosphate has only weak activity, triphosphate displays the same folding capacity as ATP. However, Adenosine 5′-(pentahydrogen tetraphosphate) (ATPP) shows the same effectiveness as ATP. Moreover, the ability to induce folding also depends on the atoms connecting the beta and gamma phosphates. When the oxygen atom linking these phosphates is replaced with a carbon atom in methyleneadenosine 5-triphosphate (AMP-PCP) or a nitrogen atom in adenylylimidodiphosphate (AMP-PNP),

their folding capacity is reduced to the level of ADP. Most intriguingly, while ATP and ADP do not induce aggregation of C71G-PFN1 even at 20 mM, triphosphate, diphosphate, ATPP, AMP-PCP, and AMP-PNP all trigger aggregation at very low concentrations [128].
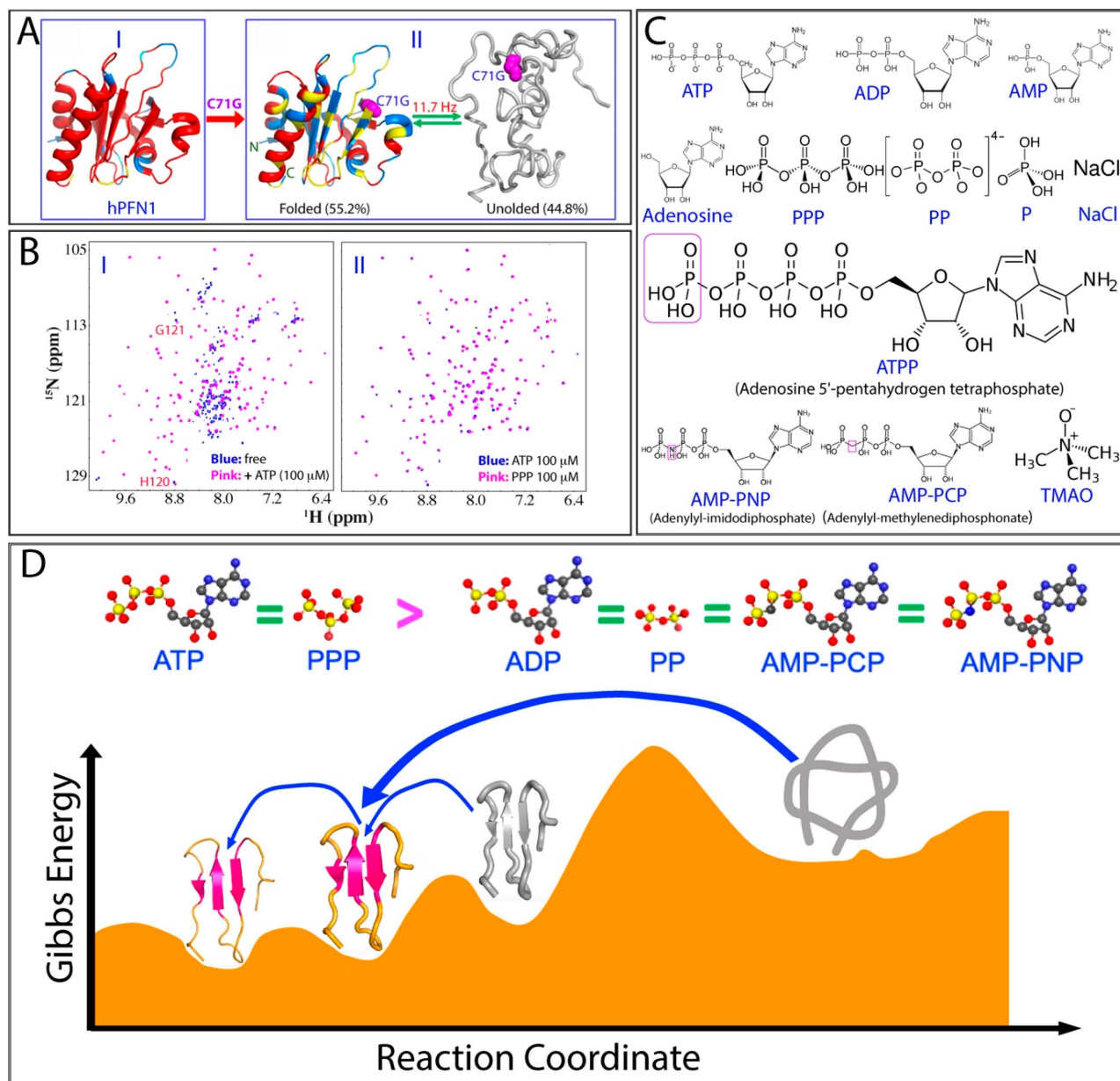


**Figure 6. ATP induces protein folding at the highest efficiency.** (**A**) A diagram to show the ALS-causing C71G mutation destabilizing hPFN1 (I) into the coexistence of the folded (with the average $S^2$ value of 0.73) and unfolded states (II). The populations of the folded and unfolded states have been calculated to be, respectively, 55.2% and 44.8%, which undergo a conformational exchange at 11.7 Hz. (**B**) (I) Superimposition of HSQC spectra of $^{15}$N-labeled C71G-hPFN1 at a concentration of 50 µM in the absence (blue) or in the presence of ATP (pink) at 100 µM. (II) Superimposition of HSQC spectra of C71G-hPFN1 in the presence of ATP at 1:2 (blue) and triphosphate (PPP) (pink) at 1:2. (**C**) Chemical structures of ATP and 11 related compounds. (**D**) Ranking of the capacity of ATP-related compounds in enhancing the conversion of the unfolded state into the folded state of ALS-causing C71G-hPFN1. Figure 6A was adapted from Figure 2a in Ref. [128], while Figure 6C was adapted from Figure 1a in Ref. [128].

So, what could be the mechanism for ATP and triphosphate to induce folding at such a high effectiveness for both hSOD1 and C71G-PFN1, which have unrelated structures and functions? It appears to be very unlikely that the mechanism is just a simple volume exclusion or electrostatic screening effect. At such a low concentration for ATP and triphosphate to induce folding, the volume-excluding mechanism proposed for TMAO at concentrations > M [172,186–188] does not operate at all. The possibility that the folding-inducing capacity arises only from electrostatic screening can also be ruled out, as the ionic strength of sodium triphosphate (PPP) is only 15 times greater than that of sodium chloride. Nevertheless, sodium chloride at 10 mM, with a much higher ionic strength than PPP at 0.1 mM, fails to induce folding but instead triggers the aggregation of C71G-hPFN1 [128].

The most plausible explanation for ATP and triphosphate to induce folding at the highest efficiency is likely their unique and strong ability to interact with protein hydration, which is rooted in phosphate (Figure 4C). Previous NMR studies revealed that ATP and triphosphate exhibit a high capacity to mediate protein hydration, even without strong or specific binding interactions [73,142]. Recent NMR and molecular dynamics (MD) simulation studies indicate that ATP generally has weak and non-specific interactions on folded proteins, even at very high ATP-to-protein ratios. However, these interactions are sufficient to mediate the hydration shell of proteins [189]. On the other hand, it has been suggested that hydrogen bonding with water molecules, particularly involving protein backbone atoms, plays a critical role in protein folding [171,172]. The unfolded state is stabilized when key backbone atoms are hydrogen-bonded to water molecules, whereas the folded state is favored when these atoms form intramolecular hydrogen bonds.

In this context, ATP and triphosphate appear to primarily facilitate protein folding by displacing water molecules that form hydrogen bonds with some key residues of the unfolded state. In the unfolded state (I of Figure 7A), the protein is heavily hydrated, with key residues interacting with water molecules through hydrogen bonds, which creates a kinetic barrier to folding. The introduction of ATP or triphosphate disrupts these water–protein interactions, loosening the water molecules and initiating the folding process (II of Figure 7A). The release of water molecules from the bound state to bulk water during folding also increases entropy, thus energetically driving folding. However, due to the exposure of the hydrophobic patches in such a protein with mutated deficits or highly dynamic tertiary packing, excessive amounts of triphosphate can lead to aggregation, caused by strong electrostatic screening effects. In contrast, ATP not only promotes folding through its triphosphate group but also enhances thermodynamic stability. This may be achieved through dynamic interaction of the aromatic base ring of ATP with the exposed hydrophobic patches or filling structural cavities, thus protecting these patches from exposure to bulk water (III of Figure 7A) [128]. This mechanism underscores that ATP and triphosphate enhance the folding propensity already encoded in the protein sequence. Indeed, ATP was unable to induce folding of intrinsically disordered domains, such as the prion-like domain of TDP-43 [190] and the RGG-rich domain of FUS [73], even at a 1:400 ratio, as these sequences inherently lack the capacity to form folded domains.
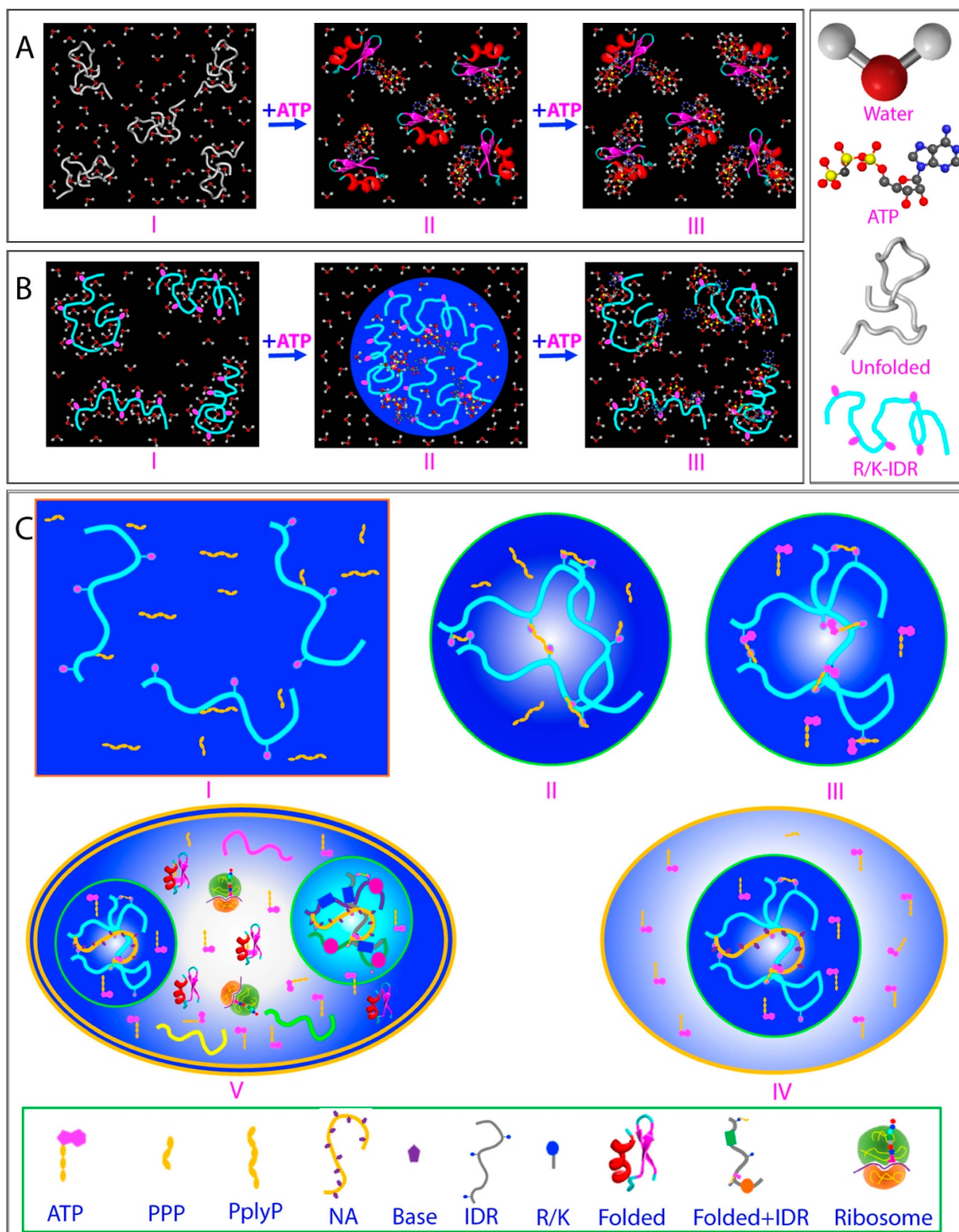
**Figure 7. ATP induces folding, modulates LLPS, and drives the emergence of protocells.** (**A**) ATP induces protein folding entropically, driven by the release of water molecules bound to the unfolded state into bulk solvent. (**B**) ATP induces LLPS of Arg-/Lys-containing IDRs entropically driven by the release of water molecules bound to the unfolded state into bulk solvent, followed by dissolution of LLPS with the excessing binding of ATP. (**C**) The emergence of protocells driven by polyphosphates, ATP, and nucleic acids.

## 7. ATP and Nucleic Acids Competitively Modulate LLPS of Arg/Lys-Containing IDRs

LLPS is now widely recognized as a common mechanism driving the formation of various membrane-less organelles (MLOs) and cellular condensates. Among these MLOs

are nucleoli, Cajal bodies, nuclear speckles, paraspeckles, histone-locus bodies, nuclear gems, and promyelocytic leukemia (PML) bodies within the nucleus, as well as P-bodies, stress granules (SGs), and germ granules in the cytoplasm [191–194]. Earlier studies primarily focused on the phase separation of well-structured proteins, such as lysozyme, showing that this process only occurs at high concentrations (>mM) [195–200]. In contrast, proteins involved in the formation of membrane-less organelles typically contain large IDRs, allowing phase separation at much lower concentrations (~μM) [201–206]. The key to this phenomenon lies in the multivalency of binding sites within IDR-rich proteins, which enables them to interact simultaneously with multiple copies of themselves (homotypic phase separation) or with other biomolecules (heterotypic phase separation). This leads to LLPS, resulting in the separation of a homogeneous solution into two distinct coexisting phases: a dense phase and a dilute phase [201–204]. LLPS is dynamic and reversible, with molecules continuously exchanging between the dense and dilute phases. Despite this exchange, the concentration of molecules in the dense phase is often more than 50 times higher than in the dilute phase.

The major mechanisms for phase separation [195–200,207–209] are often framed by the Flory–Huggins theory [210–214]. Initially developed for polymer solutions [210–212], this theory was later extended by Scott and Tompa to describe polymer mixtures in a common solvent [213,214]. However, its applicability to biological LLPS involving IDRs remains a matter of ongoing debate [204]. One major critique is that even though the Flory–Huggins model fits the experimental phase diagram satisfactorily, it offers little insight into the underlying mechanisms of phase separation and lacks predictive capability [204]. According to the Flory–Huggins model, phase separation is driven by repulsive interactions between distinct macromolecules within an inert solvent. However, this framework may not fully capture the behavior of IDRs, which are highly polar or charged and lack a stable three-dimensional structure. Due to their extended conformation and dynamic nature, nearly all regions of an IDR are exposed to the solvent. This suggests that the hydration properties of intrinsically disordered protein regions (IDPRs) differ significantly from those of well-folded, globular proteins.

Indeed, numerous studies using nuclear magnetic resonance (NMR) relaxation and other biophysical methods have shown that IDRs can bind significantly larger amounts of water than globular proteins. Furthermore, the energy barriers governing water motion in IDPs are notably distinct from those observed in globular proteins [78–82]. Since the hydration shell is involved in various molecular interactions, water cannot simply be treated as an inert solvent in IDR-mediated LLPS. Despite its critical role, the contribution of water to biological phase separation and the formation of proteinaceous MLOs is still largely underappreciated. Currently, it is widely accepted that weak, multivalent interactions help to counterbalance the loss of conformational entropy during phase separation in IDRs. Nevertheless, the strong affinity of IDRs for water molecules suggests that the release of relatively constrained water molecules into the bulk water may also play a significant role as a driving force in biological LLPS.

ATP has been shown to modulate LLPS of full-length proteins and their IDRs in a biphasic manner, including those of FUS [73,142], TDP-43 [190,215], and nucleocapsid (N) proteins [136–138,216,217]. For full-length proteins, LLPS is not only involved in interactions within the IDRs but also in additional interactions from their folded domains. These include the oligomerization of the hydrophobic region in the prion-like domain of TDP-43 and the dimerization of the C-terminal domain of the SARS-CoV-2 N protein. As a result, ATP's regulation of LLPS in these proteins is more complex due to the involvement of multiple structural elements. Here, in order to focus on elucidating how ATP biphasically modulates LLPS specifically within IDRs, only the intrinsically disordered C-terminal domain (CTD) of FUS (residues 371–526), which is rich in arginine–glycine–glycine (RGG) repeats, was reviewed. This region lacks large hydrophobic residues and cannot undergo phase separation on its own under various buffer conditions, even at concentrations as high as 1 mM [73,142].

Despite this, ATP has been found to induce LLPS at low concentrations but disrupt it at higher concentrations. When using NMR to examine residue-specific interactions, ATP was found to modulate LLPS of the FUS CTD, likely by acting as a bivalent binder, specifically interacting with Arg/Lys residues (III of Figure 5D). This mechanism was further validated in studies of the prion-like domain of TDP-43, which contains only six arginine and two lysine residues, and exhibited a biphasic response to ATP [190,215]. Similar behavior has also been observed in the RGG-rich IDR of cold-inducible RNA-binding protein [218]. In this context, a mechanism has been proposed for how ATP biphasically modulates the LLPS of Arg-/Lys-containing IDRs using the FUS C-terminal domain (CTD) as an example (Figure 7B). The FUS CTD, being intrinsically disordered and rich in positively charged Arg and Lys residues, is highly hydrated (I of Figure 7B). When ATP is added at low concentrations, it acts as a bivalent binder, crosslinking FUS CTD molecules through interactions with Arg/Lys residues. This binding also perturbs the hydration shell, leading to condensation of FUS CTD molecules and the release of bound water molecules into the bulk water. The release of these water molecules is expected to energetically favor the formation of large, dynamic complexes, which manifest as phase-separated liquid droplets (II of Figure 7B). At higher ATP concentrations, however, excessive ATP binding occurs. This overwhelming interaction disrupts the formation of these large, dynamic complexes, ultimately leading to the dissolution of the phase-separated droplets (III of Figure 7B).

Strikingly, studies combining semiempirical quantum mechanical (SQM) methods, mean-field theory, and coarse-grained molecular dynamics (CGMD) simulations have demonstrated that ATP can indeed act as a bivalent or even trivalent binder, capable of both enhancing and inhibiting the phase separation of FUS [219]. More recently, a comprehensive investigation using a combination of experimental techniques and simulations uncovered unique properties and mechanisms underlying ATP-induced phase separation in basic intrinsically disordered proteins (bIDPs). Briefly, ATP functions as a bridging agent that crosslinks bIDP chains, forming mesh-like networks that manifest as liquid-like droplets. These ATP-induced droplets display unusual physicochemical properties, including the accumulation of very high ATP concentrations within the droplets, rapid droplet fusion, low interfacial tension, and elevated viscosity. As a result, these droplets exhibit extreme shear-thinning behavior, where viscosity decreases under applied stress, making them highly dynamic and adaptable [220,221].

Like ATP, nucleic acids have also been shown to biphasically modulate LLPS, and most notably, ATP and nucleic acids compete for binding to Arg-/Lys-rich regions. However, nucleic acids differ from ATP in their structure, being composed of multiple covalently linked nucleotides, which allows them to establish multivalent interactions with Arg/Lys residues in a length-dependent manner. This makes nucleic acids act similarly to a polymer of ATP. Nevertheless, ATP possesses a unique triphosphate group that gives it a strong capacity to interact with water molecules and Arg side chains. These differences may influence how ATP and nucleic acids compete for binding to Arg/Lys residues and their respective roles in modulating LLPS [139].

This discovery has significant implications for understanding the regulation of fundamental cellular processes. Since most IDRs are expected to contain multiple Arg/Lys residues, many IDRs, like the 40-residue disordered Nogo-40 (which contains two Arg and three Lys residues), are likely capable of binding both ATP and nucleic acids to drive LLPS, even if they do not have a known function associated with these interactions [222]. In this context, a large portion of the human proteome, approximately half of which consists of IDRs, may be predisposed to undergo LLPS through multivalent interactions with nucleic acids.

However, while nucleic acids can achieve a high binding affinity to IDRs through multivalent but discrete interactions with Arg/Lys residues, this binding appears to be vulnerable to displacement by ATP. Since discrete nucleic acid binding events appear to be largely independent and thus can be simultaneously displaced by ATP, it may explain why only a limited number of MLOs are observed in cells. At high intracellular concentrations,

ATP may play a previously unrecognized role in inhibiting LLPS in many Arg-/Lys-containing IDRs. Even in MLOs that are already formed, such as stress granules (SGs), ATP is likely essential for maintaining their dynamic and reversible nature, preventing them from transitioning into irreversible aggregates, which are linked to a growing number of human diseases, including all known neurodegenerative disorders. This suggests that ATP's modulation of IDR–nucleic acid complexes is critical for cellular homeostasis and may serve to protect against pathological conditions associated with excessive phase separation and aggregation.

## 8. The Roles of ATP and Polyphosphates in the Origin of Protocells

The origin of life on Earth is thought to have resulted from a series of geochemical events, where simple prebiotic inorganic molecules interacted to eventually form biopolymers, leading to the development of protocells [223–228]. Three fundamental elements were crucial to this process: (1) an energy source (biological fuel) to power metabolic and biochemical reactions, (2) molecules capable of encoding, storing, and transmitting genetic information for reproduction, and (3) functional molecules that catalyze biochemical reactions and serve structural roles. These components operated within compartmentalized aqueous environments, facilitating essential processes such as metabolism and replication. Nature seems to have solved the first two requirements by selecting phosphorus as a fundamental building block across all living organisms. For the third challenge, proteins were engineered as the primary functional molecules, possessing an extreme diversity of three-dimensional structures, including folded domains, IDRs with varying properties, and membrane proteins.

Protocells, also known as primitive cells, are thought to represent the earliest life-like entities, exhibiting key functions such as proto-metabolism, compartmentalization, and replication [225–229]. Three main types of protocells have been proposed: membrane-free coacervate droplets, lipid vesicles, and hybrid protocells that arise from a combination of these models [228–232]. Among these, the coacervate droplet model has attracted particular attention due to its basis in LLPS, which allows the spontaneous sequestration and concentration of polyelectrolytes and biomolecules [233–236]. Remarkably, coacervate droplets can mimic several critical cellular functions, including biomolecular crowding and intercellular communication, providing valuable insights into early cell-like behavior.

On the early Earth, where conditions were extreme, synthetic organic polymers or modern biomacromolecules [235,236] likely did not exist. For example, the intensity of terrestrial radioactivity on the primordial Earth is estimated to have been up to 4000 times higher than it is today. Ionizing radiation from both cosmic rays and intrinsic radioisotopes might have played a dual role: while potentially driving key prebiotic chemical reactions, it could also have had harmful effects on emerging biomolecules, such as nucleic acids, which are essential for life [237–240]. However, inorganic polyphosphate (polyP)—a linear, negatively charged polymer made up of tens to hundreds of orthophosphate (Pi) units linked by high-energy phosphoanhydride bonds—emerged long before biological molecules [237,238]. This makes polyP a plausible energy source and phosphate donor for ATP and polynucleotide synthesis during the early stages of life's evolution [241–244]. It is suggested that polyP played a crucial role in the origin and survival of life by providing a flexible, polyanionic scaffold, facilitating the assembly and orientation of biomolecules in prebiotic cells. Remarkably, ATP is also believed to have emerged in the early stages of biochemical evolution [96]. The highest ATP yield was observed in highly pure, unsalted water (such as HPLC-grade water) at an optimal pH of approximately 5.5–6. In contrast, the yield significantly decreased when salts were introduced or when the pH deviated from this range. This specific aqueous environment aligns with conditions that can solubilize otherwise insoluble proteins, and it mirrors the likely chemical environment of prebiotic oceans or early water bodies.

LLPS has been increasingly proposed as a mechanism for concentrating and compartmentalizing biological molecules even at extremely dilute concentrations in aqueous

environments [245]. In light of this, here, a mechanism is proposed to emphasize the role of ATP in the origin of protocells. In the early stages of biochemical evolution (I of Figure 7C), the water body was likely highly unsalted and slightly acidic, conditions under which primordial peptides and proteins were soluble. The abundant presence of polyphosphate and triphosphate could have driven phase separation of primordial peptides and proteins with cationic residues, thereby concentrating these rare biomolecules (II of Figure 7C).

Although polyphosphate and triphosphate have the ability to induce folding and enhance the solubility of well-folded proteins, they can also promote the aggregation of partially folded and disordered proteins with exposed hydrophobic regions, such as C71G-PFN1 and nascent hSOD1. This observation may explain why today, polyphosphates mostly only present at high concentrations within single-celled organisms, primarily acting as primordial protein chaperones [246–250]. In these organisms, most proteins are well folded, with the low increase in protein aggregation potentially induced by polyphosphate/triphosphate.

As ATP emerged and accumulated, its binding to primordial peptides and proteins could have played a key role in driving their phase separation, forming membrane-less compartments (III of Figure 7C). ATP may have also protected these proteins from aggregation even with the rising salt concentrations in the water bodies. As nucleic acids later appeared, they likely bound to primordial peptides and proteins, further promoting phase separation and leading to the formation of dynamic protein–nucleic-acid complexes. Finally, these phase-separated droplets might have become enclosed by primitive membranes, leading to the formation of early cells (IV of Figure 7C).

Within these primordial cells, mechanisms for self-replication of proteins and nucleic acids may have evolved and become established. Over time, with the extensive sampling of protein sequences, some proteins began folding into folded structures as enhanced by ATP, while others remained intrinsically disordered. Under ATP's influence, protein–nucleic-acid complexes likely evolved into two distinct forms: well-folded and tight complexes, exemplified by the ribosome, and dynamic droplets that would form membrane-less organelles (V of Figure 7C). Eventually, cellular compartments became further separated by membranes, primarily composed of phospholipids, which feature both hydrophilic and hydrophobic regions [251–254].

## 9. ATP Acquires the Capacity to Site-Specifically Target the Hydration of Biomolecules

Phosphorus has been selected as a key element for constructing three of the four essential biomolecules of life: ATP, nucleic acids, and phospholipids [255–262]. Its versatility is indispensable for life, also playing extensive roles in signal transduction through phosphorylation; mineralization, with phosphate in hydroxyapatite comprising ~60% of bone mass to ensure skeletal strength; and buffering, stabilizing intracellular pH at ~7.4. Additionally, many other phosphate-containing compounds exist, such as coenzymes (e.g., NAD+/NADP+, FAD, and CoA), phosphosugars (e.g., fructose-1,6-bisphosphate), inositol phosphates, and energy-related molecules like phosphocreatine and phosphoenolpyruvate [47,48]. The prevalence of phosphate-based molecules has led to the adage, "Where there's life, there's phosphorus." A key question thus arises: Why does nature chose phosphates? [259–262]. It is widely believed that phosphorus was selected due to its many unique chemical properties. These include its ability to store and release energy efficiently, its balance between stability and reactivity, and its versatility in supporting a wide range of biochemical functions essential for life.

However, the unexpected findings reviewed here highlight a previously underappreciated characteristic of phosphate: its strong capacity for self-hydration and its role in regulating the hydration dynamics of other biomacromolecules, particularly proteins. Recently, a mechanism underlying phosphate's unique properties was decoded. Unlike the $SCN^-$ anion, which increases the μs-ms backbone dynamics of proteins by disrupting hydrophobic interactions and destabilizing their thermodynamic stability [124], phosphate significantly enhances backbone μs-ms dynamics without disrupting hydrophobic inter-

actions, instead strengthening thermodynamic stability [125]. This ability is likely due to the exchange between different ionic states of phosphate within the hydration shell of proteins, a process that also occurs on the μs-ms time scale [134] (Figure 4C). This mechanism may explain the unique ability of phosphate-containing molecules, such as ATP, ADP, diphosphates, triphosphates, and polyphosphates, to promote protein folding and enhance protein stability. These molecules likely achieve this by facilitating transitions among different protein conformational states. Indeed, ribosomes were initially proposed to act as chaperones for the folding of nascent polypeptides [263–266]. However, it was later discovered that it is actually the rRNA molecules within the ribosome that enhance this folding process. It is now established that nucleic acids, in general, can function as chaperones to assist in protein folding [267–269].

This unique ability of polyphosphates, combined with their capacity for high-energy storage, may serve as a key foundation for their selection as initial scaffolds for LLPS of primitive polypeptides, thus driving the formation of protocells (Figure 7C). Intriguingly, triphosphate exhibits the highest capacity for inducing protein folding. Surprisingly, while phosphate and diphosphate demonstrate weaker abilities, tetraphosphate does not show any improved folding capacity. For instance, adenosine 5′-(pentahydrogen tetraphosphate) (ATPP) has the same capacity to induce folding as ATP but additionally acquires the ability to trigger aggregation. Furthermore, the inducing capacity is also dependent of the atoms that link the phosphates. Replacing the oxygen atom connecting the beta and gamma phosphates of ATP with carbon or nitrogen not only reduces its ability to induce folding to a level comparable to that of ADP but also introduces the ability to promote aggregation, a trait that is absent in both ATP and ADP [128].

In this context, triphosphate appears to be maximized not only for an immediate reactant in probiotic evolution and energy storage [259–262] but also for enhancing protein stability and folding in an energy-independent manner, utilizing its unique feature associated with hydration (Figure 8A). However, triphosphate, diphosphate, and polyphosphates are all inorganic molecules, which come with several drawbacks: (1) they strongly trigger the aggregation of partially folded and disordered proteins with significant exposure of hydrophobic patches through their strong electrostatic screening effects [128]. As a consequence, they primarily function as chaperones in single-celled organisms, where the proteomes are predominantly composed of well-folded proteins. (2) Their interactions mostly with polar or charged surface residues of proteins are relatively weak, resulting in limited effectiveness in enhancing the stability and folding even for folded proteins [125] and in modulating phase separation of IDRs [73]. (3) They lack the ability to interact with hydrophobic surfaces, which are characteristic of biomacromolecules, and cannot effectively modulate hydrophobic interactions, a key driving force in protein folding.

Marvelously, by linking inorganic, charged triphosphate to organic, hydrophobic adenosine, nature has created ATP, which successfully bridges the inorganic and organic worlds existing in water (Figure 8). The unique structure of ATP enables it to interact broadly with hydrophobic patches on various biomacromolecules, particularly proteins, enhancing its affinity through multiple-site interactions. Notably, the aggregation-inducing properties of triphosphate and diphosphate appear to be mitigated by the covalently linked adenosine; ATP and ADP no longer trigger the aggregation of disease-causing mutants, such as C71G-PFN1, which exhibit deficient tertiary packing. Instead, ATP and ADP increase the stability of such proteins. These results suggest that in ATP and ADP, adenosine interacts with triphosphate, preserving the hydration-mediating capacity of triphosphate while simultaneously shielding its aggregation-inducing properties. Importantly, ATP, which closely resembles nucleotides, the building blocks of nucleic acids, has the ability to compete with nucleic acids for interactions with proteins and can even directly influence their conformations [270,271]. As a result, ATP not only serves as the universal energy currency for all living cells but also possesses the capacity to modulate protein homeostasis in an energy-independent manner [19]. Additionally, ATP plays a role in shaping the interfaces between the genome and proteome [139].
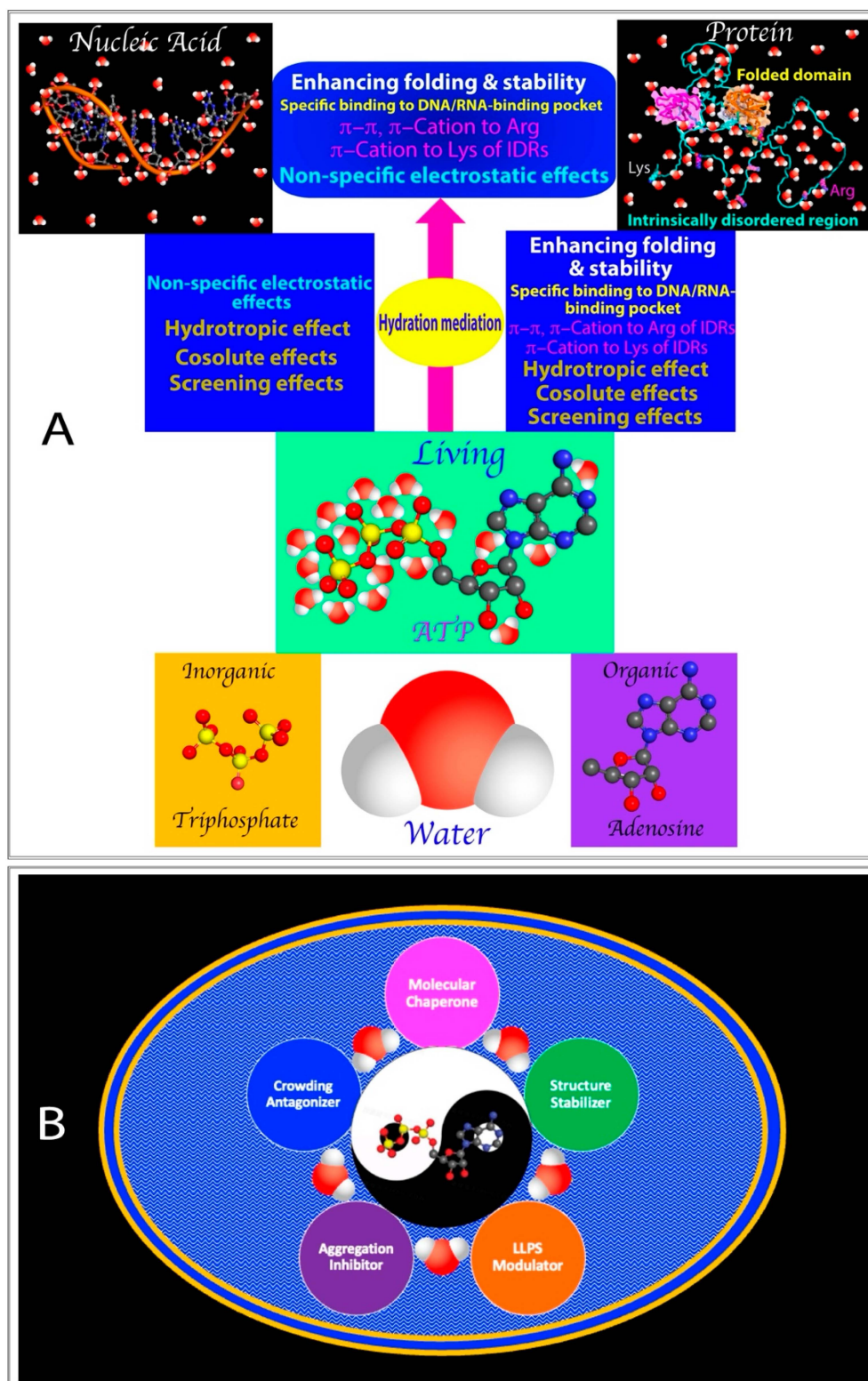
**Figure 8. The emergence and energy-independent functions of ATP.** (**A**) By linking organic, charged triphosphate and organic, hydrophobic adenosine, ATP acquires the capacity to site-specifically and non-specifically interact with proteins and nucleic acids. With a high similarity to nucleic acids, ATP thus has the capacity to shape the interface between the genome and the proteome through the evolutionary trajectory. (**B**) Energy-independent functions of ATP on protein homeostasis in modern cells.

As illustrated in Figure 8A, ATP, proteins, and nucleic acids can all interact with one another in an energy-independent manner. Nucleic acids can specifically bind to the DNA/RNA-binding pockets of proteins, while also exerting non-specific electrostatic effects. They interact with Arg residues in IDRs through $\pi$-$\pi$ and $\pi$-cation interactions, and with Lys residues in IDRs via $\pi$-cation interactions to promote LLPS. Interestingly, nucleic acids can enhance protein folding and stability by modulating hydration. Similarly, ATP can exert comparable effects on proteins but with different specificities and affinities. Additionally, ATP imposes hydrotropic effects, cosolute effects, and screening effects. It also influences nucleic acids through non-specific electrostatic interactions, hydrotropic effects, cosolute effects, and screening effects. Through the complex interplay of these interactions, ATP acquires the energy-independent ability to effectively induce protein folding, inhibit aggregation, and enhance thermodynamic stability.

Consequently, in modern cells, ATP not only serves as the universal energy currency driving all biological processes but also functions as a molecular chaperone in an energy-independent manner [123]. It acts as a structural stabilizer [128,143,144,228,272–275], a modulator of LLPS [73,136–138,142,190,225–231,274–285], an aggregation inhibitor [123, 143,286–291], and a crowding antagonist [145–147] (Figure 8B). In contrast, nucleic acids, with their multiple binding sites compared to ATP, can instigate LLPS in IDRs that contain numerous Arg/Lys residues [190,225]. Notably, although nucleic acids can achieve a high binding affinity to IDRs through multivalent but independent interactions with multiple Arg/Lys residues, this binding is relatively susceptible to displacement by ATP. This susceptibility arises because these dynamic complexes are highly accessible, allowing ATP to simultaneously displace each independent binding event involving nucleic acids. ATP's triphosphate group also possesses a unique ability to interact with water molecules and the side chains of Arg/Lys residues. These distinctions may influence the competition between ATP and nucleic acids for binding to Arg/Lys residues and their respective abilities to modulate LLPS. This phenomenon may explain why a substantial proportion of IDRs contain multiple Arg/Lys residues, making them prone to phase separation upon induction by nucleic acids of typical lengths. However, the occurrence of MLOs is notably limited within cells. It is likely that the high cellular concentrations of ATP inhibit or dissolve most nucleic-acid-induced phase separations of IDRs. Consequently, proteins within MLOs tend to feature folded domains with high nucleic-acid-binding affinities, enabling them to establish interactions with nucleic acids even in the presence of elevated ATP concentrations [139,285].

## 10. Summary and Challenges

All known life forms, including viruses, develop from genetic information compressed in a one-dimensional sequence of nucleic acids, either DNA or RNA (I of Figure 9A) [10,47,48]. The four-nucleotide coding system of nucleic acid molecules offers several advantages over the binary coding (0 and 1) used in computing [292]. These include (1) a higher information density, allowing more data to be stored per unit; (2) built-in error tolerance through redundancy and repair mechanisms; (3) biological efficiency, enabling evolutionary flexibility; and (4) the ability to encode not only protein structures but also regulatory sequences, epigenetic markers, and other complex information. Decoding the compressed genetic information to form a functional cell, and ultimately a fully developed organism, requires multiple complex steps. Among these, protein synthesis is particularly critical, as proteins serve as the primary functional molecules and structural components that drive nearly all cellular processes. Proteins are composed of 20 naturally occurring amino acids, which can be categorized into a simplified five-letter amino acid system (II of Figure 9A) [293]. Proteins spontaneously fold into an extremely diverse spectrum of conformations, including folded domains, IDRs, and membrane proteins, which can further assemble into MLOs or integrate into membrane-enveloped organelles such as mitochondria, the endoplasmic reticulum, and the cell membrane.

**Figure 9. Water "self-decodes" the protein sequence and cellular networks in water.** (**A**) Decoding of genetic information compressed in DNA (I) into the protein sequence (II). In pure water, all proteins are highly hydrated and soluble (III), which could be driven by the release of bound water molecules as well as the interplay of intrinsically encoded interactions and environmental factors to cross the energy barrier to form folded (IV), intrinsically disordered (V), phase-separated (VI), membrane-associated (VII), and aggregated (VIII) states. (**B**) The chemical structures of the phospholipid and DNA, to illustrate the presence of a phosphate group that might provoke μs-ms dynamics. (**C**) A cell constituted by the extremely complex networks, which are operating cooperatively in membrane-compartmentalized aqueous spaces rich in salts and spacetime-specifically energized by ATP.

Protein folding, the process by which linear polypeptides spontaneously self-assemble into functional three-dimensional structures, is recognized as one of biology's greatest mysteries. In his Nobel Lecture, Anfinsen stated that "the native conformation is determined by the totality of interatomic interactions and hence by the amino acid sequence, in a given environment" [12]. However, predicting protein structures from first principles still remains a significant challenge. Nonetheless, AI systems like AlphaFold 2 and 3 [294–296], built using artificial neural network architectures [297–299], have made an unprecedented breakthrough by deep-learning existing protein structures and subsequently achieving accurate structure predictions.

The failure to predict protein structures from first principles has long been attributed to our limited understanding and ability to computationally model protein hydration [300–305]. Despite this, AlphaFold 2 and 3, trained on existing protein structures without hydration information, have accurately predicted protein structures and complexes, also bypassing the need for detailed knowledge of microscopic mechanisms. This success implies that not only the three-dimensional structures, as noted by Anfinsen [12], but also hydration information might be indirectly coded in the protein sequence. In this context, hydration structures and dynamics are also determined by the properties of proteins specified by the sequence. For instance, the cataract-causing G18V mutation in γS-crystallin disrupts its hydration shell, preventing ATP from antagonizing the crowding-induced destabilization of the mutant. In contrast, other cataract mutations like D26G, S39C, and V42M only partially damage the hydration shell [146–148].

All known life forms exist in water, and proteins are also highly hydrated. The unique findings reviewed here, including the effects of salts on protein solubility, membrane interactions, stability, and dynamics, as well as ATP's role in modulating crowding-induced destabilization, protein folding, and phase separation of IDRs, all highlight the central role of protein hydration in these processes. Although the interactions driving protein folding are encoded in the amino acid sequence, it is water that allows the realization of these interactions such as the hydrophobic effect, hydrogen bonding, electrostatic screening, and solvent effects. Water, in this context, acts as the medium through which the folding potential is 'decoded' and expressed. Without water, the forces required for proper protein folding would not manifest correctly, leading to misfolded or non-functional proteins.

Therefore, protein hydration has at least three fundamental roles: (1) It ensures that proteins, regardless of their properties, remain soluble in unsalted water (III of Figure 9A), the medium in which proteins might have originally evolved. (2) Protein–water interactions shape the energy landscape of protein folding. The release of bound water molecules into the bulk solvent entropically drives protein folding, together with the specific interactions encoded in the protein sequence and influenced by environmental factors. This release of water molecules facilitates proteins to overcome energy barriers and to be "self-decoded" into various cellular states, including folded (IV of Figure 9A), intrinsically disordered (V of Figure 9A), phase-separated (VI of Figure 9A), membrane-associated (VII of Figure 9A), and aggregated or amyloid (VIII of Figure 9A) forms. (3) Each of these cellular states of proteins is also uniquely hydrated, with dynamics occurring on different time scales that are essential for their functionality. These dynamics are modulated by various molecules, including universally by salt ions and ATP.

These findings have profound and far-reaching implications, not only for advancing our fundamental understanding of the principles underlying life's molecular architecture and functions but also for driving practical applications across diverse fields such as drug design, biotechnology, and synthetic biology. Additionally, they offer new perspectives on deciphering the mechanisms of protein-aggregation diseases and developing effective therapeutic strategies. For instance, water, ATP, and ions are ubiquitous and indispensable components of all living systems, yet their previously underappreciated contributions may help explain challenges faced in designing drugs with high affinity and specificity or in engineering proteins with desired properties based on first principles. The neglect of these factors in traditional approaches might have led to oversights, particularly in addressing dynamics of molecular interactions and environmental conditions that in fact significantly influence biological processes. Furthermore, disease-causing mutations in the IDRs of proteins such as ALS-causing TDP-43 and FUS might have profoundly affected hydration dynamics, interactions with salts, and ATP binding, potentially altering protein behavior. These changes may cascade to affect the structural organization and functional topologies of associated cellular networks, exacerbating pathological outcomes. Addressing these complexities is expected to pave the way for innovative therapeutic interventions that specifically target these previously unrecognized molecular mechanisms.

The analysis presented here has also highlighted three fundamental challenges that requires future exploration:

1. How does the protein sequence specify hydration, and how is protein hydration modulated, particularly by ATP? Emerging results suggest that the formation of protein hydration by its sequence is a highly specific process. For instance, while ATP is able to effectively antagonize crowding-induced destabilization of the wild-type γS-crystallin even at a 1:1 ratio by mediating its hydration shell [145–147], the cataract-associated G18V mutation is sufficient to dramatically disrupt its hydration shell [165], rendering ATP ineffective in countering the crowding-induced destabilization of the mutant [146]. Intriguingly, other cataract mutations, such as D26G, S39C, and V42M, only partially compromise the hydration shell and remain partially responsive to ATP modulation [147]. So, how can a single mutation, such as G18V, completely disrupt the hydration shell? How is ATP modulation achieved? By altering the shape, thickness, or dynamics of the hydration shell?

In the future, the development and application of emerging experimental and computational techniques will be crucial for overcoming current limitations in understanding protein hydration dynamics. These advancements may include cutting-edge cryo-electron microscopy (cryo-EM) for capturing high-resolution structural insights, artificial intelligence (AI)-based simulations for predicting dynamic interactions, and deep learning tools for molecular modeling that integrate vast amounts of experimental and theoretical data. Such approaches could enable unprecedented precision in mapping hydration networks, understanding their role in protein folding and function, and deciphering complex biomolecular interactions. By bridging gaps in our knowledge, these innovations have the potential to transform how we study and manipulate proteins in biological and therapeutic contexts.

2. In addition to ATP, various phosphate-containing biomolecules exist in cells, including GTP, CTP, TTP, UTP, FAD, NAD/NADP, CoA, and Poly(ADP-ribose) (PAR). Their hydration dynamics and roles in modulating protein hydration remain to be explored, despite their relatively low concentrations in cells. Intriguingly, hydroxyapatite is not only a key structural component for mineralization, forming tusks, horns, bones, and teeth in many organisms, but also plays vital functional roles, such as enabling balance and motion sensing in fish otoliths, crucial for underwater orientation, and serving as a reservoir for calcium and phosphate ions to regulate physiological processes. It is of fundamental interest to assess whether the unique hydration and dynamic behaviors of phosphate contribute to its functions. In particular, two key biomacromolecules, nucleic acids and phospholipids, also contain phosphate groups (Figure 9B). Although most phosphate groups in nucleic acids and phospholipids exhibit only two ionic states (Figure 9B), unlike the more varied ionic states of triphosphate or polyphosphate groups, they still exchange on a relatively slower time scale (μs-ms) across a wide pH range, especially in physiologically relevant conditions [306]. Given their ubiquitous presence in all cells at high densities, the effects of phosphate groups on the dynamics of nucleic acids, phospholipids, and other biomolecules cannot be overlooked. Indeed, nucleic acids have already been shown to act as chaperones in protein folding [267–269]. In contrast, the role of phosphate groups in the dynamics of phospholipids remains almost unknown. How do these groups contribute to the dynamics and functions of phospholipids? Do phospholipid surfaces serve as hubs that influence protein folding and aggregation? If so, what are the underlying mechanisms? Indeed, disease-associated protein mutants have been extensively observed to aggregate on the surfaces of phospholipid membranes [39].

3. The most fundamental mystery in biology is how cellular networks operating at the molecular levels leads to the emergence of functionality and morphogenesis in cells, and ultimately, in whole organisms. It is well-established that cells contain highly complex networks, primarily metabolic, gene regulatory, and protein–protein interaction (PPI) networks [307,308], which involve numerous molecules and environmental factors (Figure 9C). These networks are essential for determining how cells differentiate and develop into

tissues and organs by interpreting the genetic blueprint encoded in nucleic acids to build the organism's physical structure. It was previously hypothesized that biological functions and morphogenesis arise as an 'emergent' phenomenon from these complex networks because networks are universal approximators for almost any functions [309–313]. However, despite extensive research across multiple disciplines, the exact mechanisms cannot be fully understood or reconstructed by analyzing the individual components of these systems or network nodes alone, primarily due to their non-linear nature [314].

While it is well-established that the DNA within an egg encodes all the genetic information necessary to build an organism, recent findings emphasize the crucial role of pre-existing cellular networks, including specific cytoplasmic factors, in decoding the information. This situation can be compared to artificial neural networks in AI, which rely on pre-trained architectures to process data. In a similar manner, pre-existing cellular networks help interpret the developmental information encoded in DNA. Thus, both DNA sequences and the pre-existing cellular networks are required for organismal development. This analogy raises the intriguing possibility that biological information is also stored in a distributed fashion across cellular networks, much like weights and biases are distributed across nodes in generative AI models. If this hypothesis holds, it has several fundamental implications: (1) Data derived from experiments using biochemical, biophysical, and cellular and molecular biology methods might only represent fragments of large cellular networks as well as the weights and biases stored in their nodes. Such data could be unreadable or uninterpretable without the full network context. (2) Proper organismal development depends not only on the genetic information encoded in DNA but also on the pre-existing topological architecture of cellular networks. Both genetic sequences and cellular networks are essential and may be inheritable, implying a more complex model of inheritance than currently thought. (3) Even transient structures, such as water hydrogen-bonding networks and hydration dynamics, occurring on timescales from picoseconds to nanoseconds, might be encoded in cellular networks by modifying the weights or the biases of their nodes, or even reshaping by their topologies. For example, the vast number of phase-separated states of IDRs have been proposed as a mechanism for storing "cellular memory" [15]. Moreover, mutations in IDRs, as seen in prion-like diseases involving proteins such as TDP-43 and FUS, may disrupt protein hydration, leading to alterations in the parameters or/and even architectures of cellular networks.

Cellular networks are far more complex than the artificial neural networks used in current AI systems. First, while AI systems rely on rigid silicon chips with limited flexibility, organisms utilize a diverse range of biomacromolecules and small molecules that exhibit flexibility across various timescales, forming dynamic networks that enhance resilience and adaptability. Second, AI components are rigidly aligned in a vacuum, whereas cellular networks operate in highly hydrated environments, suspended in water. Third, the topological architectures of AI networks are fixed, but cellular networks continuously reorganize and evolve in response to the cell's needs. Fourth, AI networks are uniformly powered by electricity, while cellular networks rely on spacetime-specific energy sources, primarily the chemical energy from ATP. This enables both energy efficiency and precise, localized control over network activity, allowing different regions of the cell to be activated or deactivated as needed. Lastly, multiple distinct networks with unique architectures operate non-uniformly yet cooperatively within a single cell and likely across entire organisms, adding another layer of complexity to biological functionality.

It is increasingly evident that even the artificial neuron, the basic computational unit of artificial neural networks, is not a natural unit for human understanding. This is because many artificial neurons are polysemantic, responding to combinations of seemingly unrelated inputs [315]. The profound complexity of cellular systems, characterized by dynamic architectures, hydration dependence, and flexible interaction architectures, indicates that biological processes can achieve sophistication far surpassing current AI. One intriguing hypothesis is that quantum mechanics may underpin certain biological processes, enabling quantum operations. While the exact mechanisms remain unknown, such explorations

could transform our understanding of biology and drive breakthroughs in cognitive science and AI [316]. For example, consciousness has been proposed to emerge from quantum computations within neurons [316–318].

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** No new data were created or analyzed in this study.

**Conflicts of Interest:** The authors declare no competing interests.

## References

1. Tielens, A.G.G.M. The Molecular Universe. *Rev. Mod. Phys.* **2013**, *85*, 1021–1081. [CrossRef]
2. Ball, P. Water—An enduring mystery. *Nature* **2008**, *452*, 291–292. [CrossRef] [PubMed]
3. Chaplin, M. Water: Its Importance to Life. *Biochem. Mol. Biol. Educ.* **2001**, *29*, 54–59. [CrossRef]
4. Ball, P. Water as an Active Constituent in Cell Biology. *Chem. Rev.* **2008**, *108*, 74–108. [CrossRef]
5. Burgess, J. (Ed.) *Ions in Solution: Basic Principles of Chemical Interactions*, 2nd ed.; Ellis Horwood: Chichester, UK, 1978.
6. Marcus, Y. Effect of Ions on the Structure of Water: Structure Making and Breaking. *Chem. Rev.* **2009**, *109*, 1346–1370. [CrossRef]
7. Israelachvili, J.N. *Intermolecular and Surface Forces*; Academic Press: Cambridge, MA, USA, 2011.
8. Schein, C.H. Solubility as a Function of Protein Structure and Solvent Components. *Nat. Biotechnol.* **1990**, *8*, 308–317. [CrossRef]
9. Bagchi, B. *Water in Biological and Chemical Processes: From Structure and Dynamics to Function*; Cambridge University Press: Cambridge, UK, 2013.
10. Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. *Molecular Biology of the Cell*, 4th ed.; Garland Science: New York, NY, USA, 2002.
11. Wootton, J.C. Sequences with 'unusual' amino acid compositions. *Curr. Opin. Struct. Biol.* **1994**, *4*, 413–421. [CrossRef]
12. Anfinsen, C.B. Principles that govern the folding of protein chains. *Science* **1973**, *181*, 223–230. [CrossRef]
13. Levinthal, C. Are there pathways for protein folding? *J. Chim. Phys.* **1968**, *65*, 44–45. [CrossRef]
14. Dill, K.A. Dominant forces in protein folding. *Biochemistry* **1990**, *29*, 7133–7155. [CrossRef]
15. Song, J. Environment-transformable sequence–structure relationship: A general mechanism for proteotoxicity. *Biophys. Rev.* **2018**, *10*, 503–516. [CrossRef] [PubMed]
16. Wright, P.E.; Dyson, H.J. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 18–29. [CrossRef] [PubMed]
17. Van Der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R.J.; Daughdrill, G.W.; Dunker, A.K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D.T.; et al. Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **2014**, *114*, 6589–6631. [CrossRef] [PubMed]
18. Boulos, I.; Jabbour, J.; Khoury, S.; Mikhael, N.; Tishkova, V.; Candoni, N.; Ghadieh, H.E.; Veesler, S.; Bassim, Y.; Azar, S.; et al. Exploring the World of Membrane Proteins: Techniques and Methods for Understanding Structure, Function, and Dynamics. *Molecules* **2023**, *28*, 7176. [CrossRef]
19. Song, J. ATP energy-independently controls protein homeostasis with unique structure and diverse mechanisms. *Protein Sci.* **2021**, *30*, 1277–1293. [CrossRef]
20. Svergun, D.I.; Richard, S.; Koch, M.H.J.; Sayers, Z.; Kuprin, S.; Zaccai, G. Protein hydration in solution: Experimental observation by X-ray and neutron scattering. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 2267–2272. [CrossRef]
21. Halle, B. Protein hydration dynamics in solution: A critical survey. *Philos. Trans. R. Soc. B Biol. Sci.* **2004**, *359*, 1207–1224. [CrossRef]
22. Chandler, D. Interfaces and the driving force of hydrophobic assembly. *Nature* **2005**, *437*, 640–647. [CrossRef]
23. Levy, Y.; Onuchic, J.N. Water mediation in protein folding and molecular recognition. *Annu. Rev. Biophys.* **2006**, *35*, 389–415. [CrossRef]
24. Dill, K.A.; MacCallum, J.L. The protein-folding problem, 50 years on. *Science* **2012**, *338*, 1042–1046. [CrossRef]
25. Baldwin, R.L. Dynamic hydration shell restores Kauzmann's 1959 explanation of how the hydrophobic factor drives protein folding. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 13052–13056. [CrossRef] [PubMed]
26. Bellissent-Funel, M.-C.; Hassanali, A.; Havenith, M.; Henchman, R.; Pohl, P.; Sterpone, F.; van der Spoel, D.; Xu, Y.; E Garcia, A. Water determines the structure and dynamics of proteins. *Chem. Rev.* **2016**, *116*, 7673–7697. [CrossRef] [PubMed]
27. Bagchi, B. Water Dynamics in the hydration layer around proteins and micelles. *Chem. Rev.* **2005**, *105*, 3197–3219. [CrossRef] [PubMed]
28. Laage, D.; Elsaesser, T.; Hynes, J.T. Water dynamics in the hydration shells of biomolecules. *Chem. Rev.* **2017**, *117*, 10694–10725. [CrossRef]
29. Chong, S.-H.; Ham, S. Distinct role of hydration water in protein misfolding and aggregation revealed by fluctuating thermodynamics analysis. *Acc. Chem. Res.* **2015**, *48*, 956–965. [CrossRef]

30. Tarek, M.; Tobias, D.J. The dynamics of protein hydration water: A quantitative comparison of molecular dynamics simulations and neutron-scattering experiments. *Biophys. J.* **2000**, *79*, 3244–3257. [CrossRef]

31. Doster, W.; Cusack, S.; Petry, W. Dynamical transition of myoglobin revealed by inelastic neutron scattering. *Nature* **1989**, *337*, 754–756. [CrossRef]

32. Feig, M.; Sugita, Y. Variable Interactions between Protein Crowders and Biomolecular Solutes Are Important in Understanding Cellular Crowding. *J. Phys. Chem. B* **2012**, *116*, 599–605. [CrossRef]

33. Dencher, N.A.; Sass, H.J.; Büldt, G. Water and bacteriorhodopsin: Structure, dynamics, and function. *Biochim. Biophys. Acta* **2000**, *1460*, 192–203. [CrossRef]

34. Chiti, F.; Dobson, C.M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **2006**, *75*, 333–366. [CrossRef]

35. Taylor, J.P.; Brown, R.H., Jr.; Cleveland, D.W. Decoding ALS: From genes to mechanism. *Nature* **2016**, *539*, 197–206. [CrossRef] [PubMed]

36. Hartl, F.U. Cellular Homeostasis and Aging. *Annu. Rev. Biochem.* **2016**, *85*, 1–4. [CrossRef] [PubMed]

37. Iadanza, M.G.; Jackson, M.P.; Hewitt, E.W.; Ranson, N.A.; Radford, S.E. A new era for understanding amyloid structures and disease. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 755–773. [CrossRef] [PubMed]

38. Karran, E.; De Strooper, B. The amyloid hypothesis in Alzheimer disease: New insights from new therapeutics. *Nat. Rev. Drug Discov.* **2022**, *21*, 306–318. [CrossRef]

39. Song, J. Molecular mechanisms of phase separation and amyloidosis of ALS/FTD-linked FUS and TDP-43. *Aging Dis.* **2023**, *15*, 2084–2112. [CrossRef]

40. Bohnert, K.A.; Kenyon, C. A lysosomal switch triggers proteostasis renewal in the immortal C. elegans germ lineage. *Nature* **2017**, *551*, 629–633. [CrossRef]

41. Lindner, A.B.; Madden, R.; Demarez, A.; Stewart, E.J.; Taddei, F. Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 3076–3081. [CrossRef]

42. Arakawa, T.; Timasheff, S.N. Theory of protein solubility. *Methods Enzymol.* **1985**, *114*, 49–77.

43. Middaugh, C.R.; Volkin, D.B. Protein solubility. In *Stability of Protein Pharmaceuticals*; Ahern, T.J., Manning, M.C., Eds.; Plenum Press: New York, NY, USA, 1992; pp. 109–134.

44. Schein, C.H. Solubility and secretability. *Curr. Opin. Biotechnol.* **1993**, *4*, 456–461. [CrossRef]

45. Fowler, S.B.; Poon, S.; Zurdo, J. Rational design of aggregation- resistant bioactive peptides: Reengineering human calcitonin. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 10105–10110. [CrossRef]

46. Millero, F.J. *Chemical Oceanography*; CRC Press: Boca Raton, FL, USA, 2013.

47. Guyton, A.C.; Hall, J.E. *Textbook of Medical Physiology*; Elsevier: Amsterdam, The Netherlands, 2011.

48. Leningher, A. *Principles of Biochemistry*; W. H. Freeman and Company: New York, NY, USA, 2005.

49. Hofmeister, F. Zur Lehre von der Wirkung der Salze. *Arch. Exp. Pathol. Pharmakol.* **1888**, *24*, 247–260. [CrossRef]

50. Collins, K.D.; Washabaugh, M.W. The Hofmeister effect and the behavior of water at interfaces. *Q. Rev. Biophys.* **1985**, *18*, 323–422. [CrossRef] [PubMed]

51. Cohn, E.J.; Edsall, J.T.; Kirkwood, J.G.; Mueller, H.; Oncley, J.L.; Scatchard, G. Proteins, Amino Acids and Peptides as Ions and Dipolar Ions. *Am. J. Med. Sci.* **1943**, *205*, 884. [CrossRef]

52. Arakawa, T.; Timasheff, S. The stabilization of proteins by osmolytes. *Biophys. J.* **1985**, *47*, 411–414. [CrossRef]

53. von Hippel, P.H.; Schleich, T. The effects of neutral salts on the structure and conformational stability of macromolecules in solution. In *Structure and Stability of Biological Macromolecules*; Timasheff, S.N., Fasman, G.D., Eds.; Marcel Dekker: New York, NY, USA, 1969; pp. 417–574.

54. Baldwin, R. How Hofmeister ion interactions affect protein stability. *Biophys. J.* **1996**, *71*, 2056–2063. [CrossRef]

55. Christendat, D.; Yee, A.; Dharamsi, A.; Kluger, Y.; Gerstein, M.; Arrowsmith, C.H.; Edwards, A.M. Structural proteomics: Prospects for high throughput sample preparation. *Prog. Biophys. Mol. Biol.* **2000**, *73*, 339–345. [CrossRef]

56. Schubert, U.; Antón, L.C.; Gibbs, J.; Norbury, C.C.; Yewdell, J.W.; Bennink, J.R. Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature* **2000**, *404*, 770–774. [CrossRef]

57. Yewdell, J.W.; David, A. Editorial: Proteostenosis: Cancer's Achilles heel? *J. Leukoc. Biol.* **2012**, *92*, 913–915. [CrossRef]

58. Yewdell, J.W. DRiPs solidify: Progress in understanding endogenous MHC class I antigen processing. *Trends Immunol.* **2011**, *32*, 548–558. [CrossRef]

59. Mulroney, T.E.; Pöyry, T.; Yam-Puc, J.C.; Rust, M.; Harvey, R.F.; Kalmar, L.; Horner, E.; Booth, L.; Ferreira, A.P.; Stoneley, M.; et al. N1-methylpseudouridylation of mRNA causes +1 ribosomal frameshifting. *Nature* **2024**, *625*, 189–194. [CrossRef]

60. Rocklin, G.J.; Chidyausiku, T.M.; Goreshnik, I.; Ford, A.; Houliston, S.; Lemak, A.; Carter, L.; Ravichandran, R.; Mulligan, V.K.; Chevalier, A.; et al. Global analysis of protein folding using massively parallel design, synthesis, and testing. *Science* **2017**, *357*, 168–175. [CrossRef] [PubMed]

61. Song, J. Insight into insoluble proteins with pure water. *FEBS Letts.* **2009**, *583*, 953–959. [CrossRef] [PubMed]

62. Li, M.; Liu, J.; Ran, X.; Fang, M.; Shi, J.; Qin, H.; Goh, J.-M.; Song, J. Resurrecting abandoned proteins with pure water: Cd and nmr studies of protein fragments solubilized in salt-free water. *Biophys. J.* **2006**, *91*, 4201–4209. [CrossRef] [PubMed]

63. Song, J. Why do proteins aggregate? Intrinsically insoluble proteins and dark mediators revealed by studies on insoluble proteins solubilized in pure water. *F1000Research* **2013**, *2*, 94. [CrossRef] [PubMed]

64. Liu, J.; Song, J. Insights into protein aggregation by NMR characterization of insoluble SH3 mutants solubilized in salt-free water. *PLoS ONE* **2009**, *4*, e7805. [CrossRef]

65. Shi, J.; Lua, S.; Tong, J.S.; Song, J. Elimination of the native structure and solubility of the hVAPB MSP domain by the Pro56Ser mutation that causes amyotrophic lateral sclerosis. *Biochemistry* **2010**, *49*, 3887–3897. [CrossRef]

66. Qin, H.; Lim, L.; Wei, Y.; Gupta, G.; Song, J. Resolving the paradox for protein aggregation diseases: NMR structure and dynamics of the membrane-embedded P56S-MSP causing ALS imply a common mechanism for aggregation-prone proteins to attack membranes. *F1000Research* **2013**, *2*, 221. [CrossRef]

67. Qin, H.; Lim, L.-Z.; Wei, Y.; Song, J. TDP-43 N terminus encodes a novel ubiquitin-like fold and its unfolded form in equilibrium that can be shifted by binding to ssDNA. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 18619–18624. [CrossRef]

68. Theng, S.S.; Wang, W.; Mah, W.-C.; Chan, C.; Zhuo, J.; Gao, Y.; Qin, H.; Lim, L.; Chong, S.S.; Song, J.; et al. Disruption of FAT10–MAD2 binding inhibits tumor progression. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E5282–E5291. [CrossRef]

69. Lim, L.; Lee, X.; Song, J. Mechanism for transforming cytosolic SOD1 into integral membrane proteins of organelles by ALS-causing mutations. *Biochim. Biophys. Acta* **2015**, *1848*, 1–7. [CrossRef]

70. Lim, L.; Wei, Y.; Lu, Y.; Song, J. ALS-causing mutations significantly perturb the self-assembly and interaction with nucleic acid of the intrinsically disordered prion-like domain of TDP-43. *PLoS Biol.* **2016**, *14*, e1002338. [CrossRef]

71. Lim, L.; Song, J. SALS-linked WT-SOD1 adopts a highly similar helical conformation as FALS-causing L126Z-SOD1 in a membrane environment. *Biochim. Biophys. Acta* **2016**, *1858*, 2223–2230. [CrossRef] [PubMed]

72. Lim, L.; Kang, J.; Song, J. ALS-causing profilin-1-mutant forms a non-native helical structure in membrane environments. *Biochim. Biophys. Acta Biomembr.* **2017**, *1859*, 2161–2170. [CrossRef] [PubMed]

73. Kang, J.; Lim, L.; Lu, Y.; Song, J. A unified mechanism for LLPS of ALS/FTLD-causing FUS as well as its modulation by ATP and oligonucleic acids. *PLoS Biol.* **2019**, *17*, e3000327. [CrossRef] [PubMed]

74. Delak, K.; Harcup, C.; Lakshminarayanan, R.; Sun, Z.; Fan, Y.; Moradian-Oldak, J.; Evans, J.S. The tooth enamel protein, porcine amelogenin, is an intrinsically disordered protein with an extended molecular configuration in the monomeric form. *Biochemistry* **2009**, *48*, 2272–2281. [CrossRef] [PubMed]

75. Aguado-Llera, D.; Goormaghtigh, E.; de Geest, N.; Quan, X.-J.; Prieto, A.; Hassan, B.A.; Gómez, J.; Neira, J.L. The basic helix-loop-helix region of human neurogenin 1 is a monomeric natively unfolded protein which forms a fuzzy complex. *Biochemistry* **2010**, *49*, 1577–1589. [CrossRef]

76. Futami, J.; Fujiyama, H.; Kinoshita, R.; Nonomura, H.; Honjo, T.; Tada, H.; Matsushita, H.; Abe, Y.; Kakimi, K. Denatured mammalian protein mixtures exhibit unusually high solubility in nucleic acid-free pure water. *PLoS ONE* **2014**, *9*, e113295. [CrossRef]

77. Song, J.; Bai, P.; Luo, L.; Peng, Z.-Y. Contribution of individual residues to formation of the native-like tertiary topology in the α-lactalbumin molten globule. *J. Mol. Biol.* **1998**, *280*, 167–174. [CrossRef]

78. Bokor, M.; Csizmók, V.; Kovács, D.; Bánki, P.; Friedrich, P.; Tompa, P.; Tompa, K. NMR relaxation studies on the hydrate layer of intrinsically unstructured proteins. *Biophys. J.* **2005**, *88*, 2030–2037. [CrossRef]

79. Tompa, K.; Bokor, M.; Tompa, P. Wide-line NMR and protein hydration. *Methods Mol. Biol.* **2012**, *895*, 167–196.

80. Tompa, P.; Han, K.-H.; Bokor, M.; Kamasa, P.; Tantos, A.; Fritz, B.; Kim, D.-H.; Lee, C.; Verebelyi, T.; Tompa, K. Wide-line NMR and DSC studies on intrinsically disordered p53 transactivation domain and its helically pre-structured segment. *BMB Rep.* **2016**, *49*, 497–501. [CrossRef] [PubMed]

81. Tompa, K.; Bokor, M.; Ágner, D.; Iván, D.; Kovács, D.; Verebélyi, T.; Tompa, P. Hydrogen Mobility and Protein–Water Interactions in Proteins in the Solid State. *Chemphyschem* **2017**, *18*, 677–682. [CrossRef] [PubMed]

82. Arya, S.; Singh, A.K.; Khan, T.; Bhattacharya, M.; Datta, A.; Mukhopadhyay, S. Water Rearrangements upon Disorder-to-Order Amyloid Transition. *J. Phys. Chem. Lett.* **2016**, *7*, 4105–4110. [CrossRef]

83. Dalal, V.; Arya, S.; Mukhopadhyay, S. Confined Water in Amyloid-Competent Oligomers of the Prion Protein. *ChemPhysChem* **2016**, *17*, 2804–2807. [CrossRef] [PubMed]

84. Arya, S.; Mukhopadhyay, S. Ordered water within the collapsed globules of an amyloidogenic intrinsically disordered protein. *J. Phys. Chem. B* **2014**, *118*, 9191–9198. [CrossRef] [PubMed]

85. Mulkidjanian, A.Y.; Galperin, M.Y.; Koonin, E.V. Co-evolution of primordial membranes and membrane proteins. *Trends Biochem. Sci.* **2009**, *34*, 206–215. [CrossRef]

86. Mulkidjanian, A.Y.; Bychkov, A.Y.; Dibrova, D.V.; Galperin, M.Y.; Koonin, E.V. Origin of first cells at terrestrial, anoxic geothermal fields. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E821–E830. [CrossRef]

87. Pinti, D.L. The origin and evolution of the oceans. In *Lectures in Astrobiology*; Gargaud, M., Barbier, B., Martin, H., Reisse, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2005; pp. 83–112.

88. Mojzsis, S.J.; Arrhenius, G.; McKeegan, K.D.; Harrison, T.M.; Nutman, A.P.; Friend, C.R.L. Evidence for life on Earth before 3800 million years ago. *Nature* **1996**, *384*, 55–59. [CrossRef]

89. Kasting, J.F.; Catling, D. Evolution of a habitable planet. *Annu. Rev. Astron. Astrophys.* **2003**, *41*, 429–463. [CrossRef]

90. Damer, B.; Deamer, D. Coupled phases and combinatorial selection in fluctuating hydrothermal pools: A scenario to guide experimental approaches to the origin of cellular life. *Life* **2015**, *5*, 872–887. [CrossRef]

91. Sasselov, D.D.; Grotzinger, J.P.; Sutherland, J.D. The origin of life as a planetary phenomenon. *Sci. Adv.* **2020**, *6*, eaax3419. [CrossRef] [PubMed]

92. Toner, J.D.; Catling, D.C. A carbonate-rich lake solution to the phosphate problem of the origin of life. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 883–888. [CrossRef] [PubMed]

93. Cleaves, H.J.; Chalmers, J.H.; Lazcano, A.; Miller, S.L.; Bada, J.L. A reassessment of prebiotic organic synthesis in neutral planetary atmospheres. *Orig. Life Evol. Biosph.* **2013**, *38*, 105–115. [CrossRef] [PubMed]

94. Leman, L.; Orgel, L.; Ghadiri, M.R. Carbonyl sulfide-mediated prebiotic formation of peptides. *Science* **2019**, *306*, 283–286. [CrossRef]

95. Müller, F. Prebiotic chemistry: Starting life from scratch. *Nat. Rev. Chem.* **2014**, *8*, 255–265.

96. Pinna, S.; Kunz, C.; Halpern, A.; Harrison, S.A.; Jordan, S.F.; Ward, J.; Werner, F.; Lane, N. A prebiotic basis for ATP as the universal energy currency. *PLoS Biol.* **2022**, *20*, e3001437. [CrossRef]

97. Taylor, W.R.; Chelliah, V.; Hollup, S.M.; MacDonald, J.T.; Jonassen, I. Probing the dark matter of protein fold space. *Structure* **2009**, *17*, 1244–1252. [CrossRef]

98. Koonin, E.V.; Martin, W. On the origin of genomes and cells within inorganic compartments. *Trends Genet.* **2005**, *21*, 647–654. [CrossRef]

99. Elfrink, K.; Ollesch, J.; Stöhr, J.; Willbold, D.; Riesner, D.; Gerwert, K. Structural changes of membrane-anchored native PrP(C). *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 10815–10819. [CrossRef]

100. Kotler, S.A.; Walsh, P.; Brender, J.R.; Ramamoorthy, A. Differences between amyloid-β aggregation in solution and on the membrane: Insights into elucidation of the mechanistic details of Alzheimer's disease. *Chem. Soc. Rev.* **2014**, *43*, 6692–6700. [CrossRef]

101. Künze, G.; Barré, P.; Scheidt, H.A.; Thomas, L.; Eliezer, D.; Huster, D. Binding of the three-repeat domain of tau to phospholipid membranes induces an aggregated-like state of the protein. *Biochim. Biophys. Acta* **2012**, *1818*, 2302–2313. [CrossRef] [PubMed]

102. Ulmer, T.S.; Bax, A.; Cole, N.B.; Nussbaum, R.L. Structure and dynamics of micelle-bound human alpha-synuclein. *J. Biol. Chem.* **2005**, *280*, 9595–9603. [CrossRef]

103. Auluck, P.K.; Caraveo, G.; Lindquist, S. α-Synuclein: Membrane interactions and toxicity in Parkinson's disease. *Annu. Rev. Cell Dev. Biol.* **2010**, *26*, 211–233. [CrossRef] [PubMed]

104. Kegel, K.B.; Sapp, E.; Yoder, J.; Cuiffo, B.; Sobin, L.; Kim, Y.J.; Qin, Z.-H.; Hayden, M.R.; Aronin, N.; Scott, D.L.; et al. Huntingtin associates with acidic phospholipids at the plasma membrane. *J. Biol. Chem.* **2005**, *280*, 36464–36473. [CrossRef] [PubMed]

105. Brender, J.R.; Salamekh, S.; Ramamoorthy, A. Membrane disruption and early events in the aggregation of the diabetes related peptide IAPP from a molecular perspective. *Acc. Chem. Res.* **2012**, *45*, 454–462. [CrossRef] [PubMed]

106. Gu, J.; Zhou, X.; Sutherland, L.; Kato, M.; Jaczynska, K.; Rizo, J.; McKnight, S.L. Oxidative regulation of TDP-43 self-association by a β-to-α conformational switch. *Proc. Natl. Acad. Sci. USA* **2023**, *120*, e2311416120. [CrossRef]

107. Yang, C.; Danielson, E.W.; Qiao, T.; Metterville, J.; Brown, R.H., Jr.; Landers, J.E.; Xu, Z. Mutant PFN1 causes ALS phenotypes and progressive motor neuron degeneration in mice by a gain of toxicity. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E6209–E6218. [CrossRef]

108. Liu, J.; Lillo, C.; Jonsson, P.A.; Velde, C.V.; Ward, C.M.; Miller, T.M.; Subramaniam, J.R.; Rothstein, J.D.; Marklund, S.; Andersen, P.M.; et al. Toxicity of familial ALS-linked SOD1 mutants from selective recruitment to spinal mitochondria. *Neuron* **2004**, *43*, 5–17. [CrossRef]

109. Vande Velde, C.; Miller, T.M.; Cashman, N.R. Cleveland DW Selective association of misfolded ALS-linked mutant SOD1 with the cytoplasmic face of mitochondria. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 4022–4027. [CrossRef]

110. Sun, S.; Sun, Y.; Ling, S.-C.; Ferraiuolo, L.; McAlonis-Downes, M.; Zou, Y.; Drenner, K.; Wang, Y.; Ditsworth, D.; Tokunaga, S.; et al. Translational profiling identifies a cascade of damage initiated in motor neurons and spreading to glia in mutant SOD1-mediated ALS. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E6993–E7002. [CrossRef]

111. Tipping, K.W.; Karamanos, T.K.; Jakhria, T.; Iadanza, M.G.; Goodchild, S.C.; Tuma, R.; Ranson, N.A.; Hewitt, E.W.; Radford, S.E. pH-induced molecular shedding drives the formation of amyloid fibril-derived oligomers. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 5691–5696. [CrossRef] [PubMed]

112. Shahmoradian, S.H.; Lewis, A.J.; Genoud, C.; Hench, J.; Moors, T.E.; Navarro, P.P.; Castaño-Díez, D.; Schweighauser, G.; Graff-Meyer, A.; Goldie, K.N.; et al. Lewy pathology in Parkinson's disease consists of crowded organelles and lipid membranes. *Nat. Neurosci.* **2019**, *22*, 1099–1109. [CrossRef] [PubMed]

113. Kim, Y.E.; Hipp, M.S.; Bracher, A.; Hayer-Hartl, M.; Hartl, F.U. Molecular chaperone functions in protein folding and proteostasis. *Annu. Rev. Biochem.* **2013**, *82*, 323–355. [CrossRef] [PubMed]

114. Spalding, K.L.; Bhardwaj, R.D.; Buchholz, B.A.; Druid, H.; Frisén, J. Retrospective birth dating of cells in humans. *Cell* **2005**, *122*, 133–143. [CrossRef]

115. Salis, A.; Ninham, B.W. Models and mechanisms of Hofmeister effects in electrolyte solutions, and colloid and protein systems revisited. *Chem. Soc. Rev.* **2014**, *43*, 7358–7377. [CrossRef]

116. Okur, H.I.; Hladílková, J.; Rembert, K.B.; Cho, Y.; Heyda, J.; Dzubiella, J.; Cremer, P.S.; Jungwirth, P. Beyond the Hofmeister Series: Ion-Specific Effects on Proteins and Their Biological Functions. *J. Phys. Chem. B* **2017**, *121*, 1997–2014. [CrossRef]

117. Zhang, Y.; Cremer, P.S. Interactions between macromolecules and ions: The Hofmeister series. *Curr. Opin. Chem. Biol.* **2006**, *10*, 658–663. [CrossRef]

118. Gregory, K.P.; Wanless, E.J.; Webber, G.B.; Craig, V.S.J.; Page, A.J. The electrostatic origins of specific ion effects: Quantifying the Hofmeister series for anions. *Chem. Sci.* **2021**, *12*, 15007–15015. [CrossRef]

119. Carapeto, A.P.; Marcuello, C.; Faísca, P.F.N.; Rodrigues, M.S. Morphological and Biophysical Study of S100A9 Protein Fibrils by Atomic Force Microscopy Imaging and Nanomechanical Analysis. *Biomolecules* **2024**, *14*, 1091. [CrossRef]

120. Ziaunys, M.; Sakalauskas, A.; Mikalauskaite, K.; Smirnovas, V. Polymorphism of Alpha-Synuclein Amyloid Fibrils Depends on Ionic Strength and Protein Concentration. *Int. J. Mol. Sci.* **2021**, *22*, 12382. [CrossRef]

121. Williamson, M.P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16. [CrossRef] [PubMed]

122. Miao, L.; Qin, H.; Koehl, P.; Song, J. Selective and specific ion binding on proteins at physiologically-relevant concentrations. *FEBS Lett.* **2011**, *585*, 3126–3132. [CrossRef] [PubMed]

123. Kang, J.; Lim, L.; Song, J. ATP binds and inhibits the neurodegeneration-associated fibrillization of the FUS RRM domain. *Commun. Biol.* **2019**, *2*, 223. [CrossRef]

124. Lim, L.-Z.; Song, J. NMR Dynamic View of the Destabilization of WW4 Domain by Chaotropic GdmCl and NaSCN. *Int. J. Mol. Sci.* **2024**, *25*, 7344. [CrossRef] [PubMed]

125. Lim, L.-Z.; Song, J. NMR Dynamic View of the Stabilization of the WW4 Domain by Neutral NaCl and Kosmotropic $Na_2SO_4$ and $NaH_2PO_4$. *Int. J. Mol. Sci.* **2024**, *25*, 9091. [CrossRef] [PubMed]

126. Qin, H.; Pu, H.X.; Li, M.; Ahmed, S.; Song, J. Identification and structural mechanism for a novel interaction between a ubiquitin ligase WWP1 and Nogo-A, a key inhibitor for central nervous system regeneration. *Biochemistry* **2008**, *47*, 13647–13658. [CrossRef]

127. Lu, Y.; Lim, L.; Song, J. RRM domain of ALS/FTD-causing FUS characteristic of irreversible unfolding spontaneously self-assembles into amyloid fibrils. *Sci. Rep.* **2017**, *7*, 1043. [CrossRef]

128. Kang, J.; Lim, L.; Song, J. ATP induces folding of ALS-causing C71G-hPFN1 and nascent hSOD1. *Commun. Chem.* **2023**, *6*, 186. [CrossRef]

129. Song, J.; Gilquin, B.; Jamin, N.; Drakopoulou, E.; Guenneugues, M.; Dauplais, M.; Vita, C.; Ménez, A. NMR solution structure of a two-disulfide derivative of charybdotoxin: Structural evidence for conservation of scorpion toxin alpha/beta motif and its hydrophobic side chain packing. *Biochemistry* **1997**, *36*, 3760–3766. [CrossRef]

130. Song, J.; Ménez, A.; Jamin, N.; Gilquin, B.; Vita, C. A gradual disruption of tight side-chain packing: 2D 1H-NMR characterization of acid-induced unfolding of CHABII. *Nat. Struct. Biol.* **1999**, *6*, 129–134. [CrossRef]

131. Wei, Z.; Song, J. Molecular mechanism underlying the thermal stability and pH-induced unfolding of CHABII. *J. Mol. Biol.* **2005**, *348*, 205–218. [CrossRef] [PubMed]

132. Jencks, W.P. On the attribution and additivity of binding energies. *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 4046–4050. [CrossRef] [PubMed]

133. Ryde, U. A fundamental view of enthalpy–entropy compensation. *MedChemComm* **2014**, *5*, 1324–1336. [CrossRef]

134. Van Wazer, J.R.; Callis, C.F.; Shoolery, J.N.; Jones, R.C. Principles of Phosphorus Chemistry. II. Nuclear Magnetic Resonance Measurements. *J. Am. Chem. Soc.* **1956**, *78*, 5715–5726. [CrossRef]

135. Greiner, J.V.; Glonek, T. Hydrotropic function of ATP in the crystalline lens. *Exp. Eye Res.* **2020**, *190*, 107862. [CrossRef]

136. Dang, M.; Li, Y.; Song, J. ATP biphasically modulates LLPS of SARS-CoV-2 nucleocapsid protein and specifically binds its RNA-binding domain. *Biochem. Biophys. Res. Commun.* **2021**, *19*, 50–55. [CrossRef]

137. Dang, M.; Li, T.; Song, J. ATP and nucleic acids competitively modulate LLPS of the SARS-CoV2 nucleocapsid protein. *Commun. Biol.* **2023**, *6*, 80. [CrossRef]

138. Dang, M.; Song, J. CTD of SARS-CoV-2 N protein is a cryptic domain for binding ATP and nucleic acid that interplay in modulating phase separation. *Protein Sci.* **2022**, *31*, 345–356. [CrossRef]

139. Song, J. Adenosine Triphosphate: The Primordial Molecule That Controls Protein Homeostasis and Shapes the Genome–Proteome Interface. *Biomolecules* **2024**, *14*, 500. [CrossRef]

140. Patel, A.; Malinovska, L.; Saha, S.; Wang, J.; Alberti, S.; Krishnan, Y.; Hyman, A.A. ATP as a biological hydrotrope. *Science* **2017**, *356*, 753–756. [CrossRef]

141. Rice, A.M.; Rosen, M.K. ATP controls the crowd. *Science* **2017**, *356*, 701–702. [CrossRef] [PubMed]

142. Kang, J.; Lim, L.; Song, J. ATP enhances at low concentrations but dissolves at high concentrations liquid-liquid phase separation (LLPS) of ALS/FTD-causing FUS. *Biochem. Biophys. Res. Commun.* **2018**, *504*, 545–551. [CrossRef] [PubMed]

143. Dang, M.; Kang, J.; Lim, L.; Li, Y.; Wang, L.; Song, J. ATP is a cryptic binder of TDP-43 RRM domains to enhance stability and inhibit ALS/AD-associated fibrillation. *Biochem. Biophys. Res. Commun.* **2020**, *522*, 247–253. [CrossRef] [PubMed]

144. Dang, M.; Song, J. ALS-causing D169G mutation disrupts the ATP-binding capacity of TDP-43 RRM1 domain. *Biochem Biophys Res Commun.* **2020**, *524*, 459–464. [CrossRef]

145. He, Y.; Kang, J.; Song, J. ATP antagonizes the crowding-induced destabilization of the human eye-lens protein γS-crystallin. *Biochem. Biophys. Res. Commun.* **2020**, *526*, 1112–1117. [CrossRef]

146. He, Y.; Kang, J.; Song, J. Cataract-causing G18V eliminates the antagonization by ATP against the crowding-induced destabilization of human γS-crystallin. *Biochem. Biophys. Res. Commun.* **2020**, *530*, 554–560. [CrossRef]

147. He, Y.; Kang, J.; Song, J. ATP differentially antagonizes the crowding-induced destabilization of human γS-crystallin and its four cataract-causing mutants. *Biochem. Biophys. Res. Commun.* **2020**, *533*, 913–918. [CrossRef]

148. Mogami, G.; Wazawa, T.; Morimoto, N.; Kodama, T.; Suzuki, M. Hydration properties of adenosine phosphate series as studied by microwave dielectric spectroscopy. *Biophys. Chem.* **2011**, *154*, 1–7. [CrossRef]

149. Benedek, G.B. Theory of Transparency of the Eye. *Appl. Opt.* **1971**, *10*, 459–473. [CrossRef]

150. Ghosh, K.S.; Chauhan, P. Crystallins and their complexes. *Subcell. Biochem.* **2019**, *93*, 439–460.

151. Horwitz, J. Alpha-crystallin can function as a molecular chaperone. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 10449–10453. [CrossRef]

152. Horwitz, J. Alpha-Crystallin. *Exp. Eye Res.* **2003**, *76*, 145–153. [CrossRef] [PubMed]

153. Serebryany, E.I.; King, J.A. The βγ-crystallins: Native state stability and pathways to aggregation. *Prog. Biophys. Mol. Biol.* **2014**, *115*, 32–41. [CrossRef] [PubMed]

154. Moreau, K.L.; King, J.A. Protein misfolding and aggregation in cataract disease and prospects for prevention. *Trends Mol. Med.* **2012**, *18*, 273–282. [CrossRef]

155. Tardieu, A.; Vérétout, F.; Krop, B.; Slingsby, C. Protein interactions in the calf eye lens: Interactions between beta-crystallins are repulsive whereas in gammacrystallins they are attractive. *Eur. Biophys. J.* **1992**, *21*, 1–12. [CrossRef] [PubMed]

156. Petitt, P.; Edwards, M.E.; Forciniti, D. A small-angle neutron scattering study of gamma-crystallins near their isoelectric point. *Eur. J. Biochem.* **1997**, *243*, 415–421. [CrossRef] [PubMed]

157. Mills, I.A.; Flaugh, S.L.; Kosinski-Collins, M.S.; King, J.A. Folding and stability of the isolated Greek key domains of the long-lived human lens proteins gamma-crystallin and gammaS-crystallin. *Protein Sci.* **2007**, *16*, 2427–2444. [CrossRef]

158. Zhao, H.; Chen, Y.; Rezabkova, L.; Wu, Z.; Wistow, G.; Schuck, P. Solution properties of γ-crystallins: Hydration of fish and mammal γ-crystallins. *Protein Sci.* **2014**, *23*, 88–99. [CrossRef]

159. Huang, K.-Y.; Kingsley, C.N.; Sheil, R.; Cheng, C.-Y.; Bierma, J.C.; Roskamp, K.W.; Khago, D.; Martin, R.W.; Han, S. Stability of protein-specific hydration shell on crowding. *J. Am. Chem. Soc.* **2016**, *138*, 5392–5402. [CrossRef]

160. Minton, A.P. Influence of macromolecular crowding upon the stability and state of association of proteins: Predictions and observations. *J. Pharm. Sci.* **2005**, *94*, 1668–1675. [CrossRef]

161. Ellis, R. Macromolecular crowding: Obvious but underappreciated. *Trends Biochem. Sci.* **2001**, *26*, 597–604. [CrossRef]

162. Inomata, K.; Ohno, A.; Tochio, H.; Isogai, S.; Tenno, T.; Nakase, I.; Takeuchi, T.; Futaki, S.; Ito, Y.; Hiroaki, H.; et al. High-resolution multi-dimensional NMR spectroscopy of proteins in human cells. *Nature* **2009**, *458*, 106–109. [CrossRef] [PubMed]

163. Sarkar, M.; Smith, A.E.; Pielak, G.J. Impact of reconstituted cytosol on protein stability. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 19342–19347. [CrossRef] [PubMed]

164. Harada, R.; Sugita, Y.; Feig, M. Protein crowding affects hydration structure and dynamics. *J. Am. Chem. Soc.* **2012**, *134*, 4842–4849. [CrossRef] [PubMed]

165. Kingsley, C.N.; Brubaker, W.D.; Markovic, S.; Diehl, A.; Brindley, A.J.; Oschkinat, H.; Martin, R.W. Preferential and specific binding of human αB-crystallin to a cataractrelated variant of γS-crystallin. *Structure* **2013**, *21*, 2221–2227. [CrossRef]

166. Sun, H.; Ma, Z.; Li, Y.; Liu, B.; Li, Z.; Ding, X.; Gao, Y.; Ma, W.; Tang, X.; Li, X.; et al. Gamma-S crystallin gene (CRYGS) mutation causes dominant progressive cortical cataract in humans. *J. Med. Genet.* **2005**, *42*, 706–710. [CrossRef]

167. Ma, Z.; Piszczek, G.; Wingfield, P.T.; Sergeev, Y.V.; Hejtmancik, J.F. The G18V CRYGS mutation associated with human cataracts increases gammaS-crystallin sensitivity to thermal and chemical stress. *Biochemistry* **2009**, *48*, 7334–7341. [CrossRef]

168. Brubaker, W.D.; Freites, J.A.; Golchert, K.J.; Shapiro, R.A.; Morikis, V.; Tobias, D.J.; Martin, R.W. Separating instability from aggregation propensity in γS-crystallin variants. *Biophys. J.* **2011**, *100*, 498–506. [CrossRef]

169. Navarro, A.; Boveris, A. Brain mitochondrial dysfunction in aging. *IUBMB Life* **2008**, *60*, 308–314. [CrossRef]

170. Moran, U.; Phillips, R.; Milo, R. SnapShot: Key Numbers in Biology. *Cell* **2010**, *141*, 1262–1262.e1. [CrossRef]

171. Baldwin, R.L. Early days of studying the mechanism of protein folding. In *Protein Folding Handbook*; Wiley: Hoboken, NJ, USA, 2005; pp. 1–23.

172. Rose, G.D.; Fleming, P.J.; Banavar, J.R.; Maritan, A. A backbone-based theory of protein folding. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 16623–16633. [CrossRef] [PubMed]

173. Kim, D.E.; Gu, H.; Baker, D. The sequences of small proteins are not extensively optimized for rapid folding by natural selection. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 4982–4986. [CrossRef] [PubMed]

174. Baase, W.A.; Liu, L.; Tronrud, D.E.; Matthews, B.W. Lessons from the lysozyme of phage T4. *Protein Sci.* **2010**, *19*, 631–641. [CrossRef] [PubMed]

175. Dyson, H.J.; Wright, P.E. How does your protein fold? Elucidating the apomyoglobin folding pathway. *Acc. Chem. Res.* **2017**, *50*, 105–111. [CrossRef]

176. Pelham, H.R. Speculations on the functions of the major heat shock and glucose-regulated proteins. *Cell* **1986**, *46*, 959–961. [CrossRef]

177. Bukau, B.; Weissman, J.; Horwich, A. Molecular chaperones and protein quality control. *Cell* **2006**, *125*, 443–451. [CrossRef]

178. Kerner, M.J.; Naylor, D.J.; Ishihama, Y.; Maier, T.; Chang, H.-C.; Stines, A.P.; Georgopoulos, C.; Frishman, D.; Hayer-Hartl, M.; Mann, M.; et al. Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli*. *Cell* **2005**, *122*, 209–220. [CrossRef]

179. Bruijn, L.I.; Miller, T.M.; Cleveland, D.W. Unraveling the mechanisms involved in motor neuron degeneration in ALS. *Annu. Rev. Neurosci.* **2004**, *27*, 723–749. [CrossRef]

180. Rosen, D.R.; Siddique, T.; Patterson, D.; Figlewicz, D.A.; Sapp, P.; Hentati, A.; Donaldson, D.; Goto, J.; O'Regan, J.P.; Deng, H.-X.; et al. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* **1993**, *362*, 59–62. [CrossRef]

181. Abel, O.; Powell, J.F.; Andersen, P.M.; Al-Chalabi, A. ALSoD: A user-friendly online bioinformatics tool for amyotrophic lateral sclerosis genetics. *Hum. Mutat.* **2012**, *33*, 1345–1351. [CrossRef]

182. Valentine, J.S.; Doucette, P.A.; Potter, S.Z. Copper-zinc superoxide dismutase and amyotrophic lateral sclerosis. *Annu. Rev. Biochem.* **2005**, *74*, 563–593. [CrossRef] [PubMed]

183. Lim, L.; Kang, J.; Song, J. Extreme diversity of 12 cations in folding ALS-linked hSOD1 unveils novel hSOD1-dependent mechanisms for $Fe^{2+}$/$Cu^{2+}$-induced cytotoxicity. *Sci. Rep.* **2023**, *13*, 19868. [CrossRef] [PubMed]

184. Ferron, F.; Rebowski, G.; Lee, S.H.; Dominguez, R. Structural basis for the recruitment of profilin–actin complexes during filament elongation by Ena/VASP. *EMBO J.* **2007**, *26*, 4597–4606. [CrossRef] [PubMed]

185. Boopathy, S.; Silvas, T.V.; Tischbein, M.; Jansen, S.; Shandilya, S.M.; Zitzewitz, J.A.; Landers, J.E.; Goode, B.L.; Schiffer, C.A.; Bosco, D.A. Structural basis for mutation-induced destabilization of profilin 1 in ALS. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 7984–7989. [CrossRef]

186. Baskakov, I.; Bolen, D.W. Forcing thermodynamically unfolded proteins to fold. *J. Biol. Chem.* **1998**, *273*, 4831–4834. [CrossRef]

187. Liao, Y.-T.; Manson, A.C.; DeLyser, M.R.; Noid, W.G.; Cremer, P.S. Trimethylamine *N* -oxide stabilizes proteins via a distinct mechanism compared with betaine and glycine. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 2479–2484. [CrossRef]

188. Velasquez, M.T.; Ramezani, A.; Manal, A.; Raj, D.S. Trimethylamine N-Oxide: The good, the bad and the unknown. *Toxins* **2016**, *8*, 326. [CrossRef]

189. Nishizawa, M.; Walinda, E.; Morimoto, D.; Kohn, B.; Scheler, U.; Shirakawa, M.; Sugase, K. Effects of weak nonspecific interactions with ATP on proteins. *J. Am. Chem. Soc.* **2021**, *143*, 11982–11993. [CrossRef]

190. Dang, M.; Lim, L.; Kang, J.; Song, J. ATP biphasically modulates LLPS of TDP-43 PLD by specifically binding arginine residues. *Commun. Biol.* **2021**, *4*, 714. [CrossRef]

191. Brangwynne, C.P.; Eckmann, C.R.; Courson, D.S.; Rybarska, A.; Hoege, C.; Gharakhani, J.; Jülicher, F.; Hyman, A.A. Germline Pgranules are liquid droplets that localize by controlled dissolution/condensation. *Science* **2009**, *324*, 1729–1732. [CrossRef]

192. Hyman, A.A.; Weber, C.A.; Jülicher, F. Liquid-liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* **2014**, *30*, 39–58. [CrossRef] [PubMed]

193. Wang, J.; Choi, J.-M.; Holehouse, A.S.; Lee, H.O.; Zhang, X.; Jahnel, M.; Maharana, S.; Lemaitre, R.; Pozniakovsky, A.; Drechsel, D.; et al. A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell* **2018**, *174*, 688–699.e16. [CrossRef] [PubMed]

194. Shin, Y.; Brangwynne, C.P. Liquid phase condensation in cell physiology and disease. *Science* **2017**, *357*, eaaf4382. [CrossRef] [PubMed]

195. Tanaka, T.; Ishimoto, C.; Chylack, L.T. Phase separation of a protein-water mixture in cold cataract in the young rat lens. *Science* **1977**, *197*, 1010–1012. [CrossRef] [PubMed]

196. A Thomson, J.; Schurtenberger, P.; Thurston, G.M.; Benedek, G.B. Binary liquid phase separation and critical phenomena in a protein/water solution. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 7079–7083. [CrossRef]

197. Taratuta, V.G.; Holschbach, A.; Thurston, G.M.; Blankschtein, D.; Benedek, G.B. Liquid–liquid phase separation of aqueous lysozyme solutions: Effects of pH and salt identity. *J. Phys. Chem.* **1990**, *94*, 2140–2472. [CrossRef]

198. Broide, M.L.; Berland, C.R.; Pande, J.; O Ogun, O.; Benedek, G.B. Binary-liquid phase separation of lens protein solutions. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 5660–5664. [CrossRef]

199. Dumetz, A.C.; Chockla, A.M.; Kaler, E.W.; Lenhoff, A.M. Protein phase behavior in aqueous solutions: Crystallization, liquid-liquid phase separation, gels, and aggregates. *Biophys. J.* **2008**, *94*, 570–583. [CrossRef]

200. Overbeek, J.T.G.; Voorn, M.J. Phase separation in polyelectrolyte solutions. Theory of complex coacervation. *J. Cell. Comp. Physiol.* **1957**, *49*, 7–26. [CrossRef]

201. Toretsky, J.A.; Wright, P.E. Assemblages: Functional units formed by cellular phase separation. *J. Cell Biol.* **2014**, *206*, 579–588. [CrossRef]

202. Alberti, S.; Gladfelter, A.; Mittag, T. Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates. *Cell* **2019**, *176*, 419–434. [CrossRef] [PubMed]

203. Brady, J.P.; Farber, P.J.; Sekhar, A.; Lin, Y.-H.; Huang, R.; Bah, A.; Nott, T.J.; Chan, H.S.; Baldwin, A.J.; Forman-Kay, J.D.; et al. Structural and hydrodynamic properties of an intrinsically disordered region of a germ cell-specific protein on phase separation. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E8194–E8203. [CrossRef] [PubMed]

204. Zaslavsky, B.Y.; Uversky, V.N. In Aqua Veritas: The Indispensable yet Mostly Ignored Role of Water in Phase Separation and Membrane-less Organelles. *Biochemistry* **2018**, *57*, 2437–2451. [CrossRef] [PubMed]

205. Brangwynne, C.P.; Tompa, P.; Pappu, R.V. Polymer physics of intracellular phase transitions. *Nat. Phys.* **2015**, *11*, 899–904. [CrossRef]

206. Banani, S.F.; Lee, H.O.; Hyman, A.A.; Rosen, M.K. Biomolecular condensates: Organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **2017**, *18*, 285–298. [CrossRef]

207. Chong, P.A.; Forman-Kay, J.D. Liquid-liquid phase separation in cellular signaling systems. *Curr. Opin. Struct. Biol.* **2016**, *41*, 180–186. [CrossRef]

208. Lyon, A.S.; Peeples, W.B.; Rosen, M.K. A framework for understanding the functions of biomolecular condensates across scales. *Nat. Rev. Mol. Cell Biol.* **2021**, *22*, 215–235. [CrossRef]

209. Berry, J.; Brangwynne, C.P.; Haataja, M.P. Physical principles of intracellular organization via active and passive phase transitions. *Rep. Prog. Phys.* **2018**, *81*, 046601. [CrossRef]

210. Flory, P.J. Thermodynamics of high polymer solutions. *J. Chem. Phys.* **1942**, *10*, 51–61. [CrossRef]

211. Huggins, M.L. Solutions of long chain compounds. *J. Chem. Phys.* **1941**, *9*, 440. [CrossRef]

212. Flory, P.J. *Principles of Polymer Chemistry*; Cornell University Press: Ithaca, NY, USA, 1953.

213. Scott, R.L. The Thermodynamics of High Polymer Solutions 0.5. Phase Equilibria in the Ternary System—Polymer-1 Polymer-2 Solvent. *J. Chem. Phys.* **1949**, *17*, 279–284. [CrossRef]

214. Tompa, H. Phase Relationships in Polymer Solutions. *Trans. Faraday Soc.* **1949**, *45*, 1142–1152. [CrossRef]

215. Dang, M.; Li, T.; Zhou, S.; Song, J. Arg/Lys-containing IDRs are cryptic binding domains for ATP and nucleic acids that interplay to modulate LLPS. *Commun. Biol.* **2022**, *5*, 1315. [CrossRef] [PubMed]

216. Dang, M.; Song, J. Structural basis of anti-SARS-CoV-2 activity of HCQ: Specific binding to N protein to disrupt its interaction with nucleic acids and LLPS. *QRB Discov.* **2021**, *2*, e13. [CrossRef] [PubMed]

217. Dang, M.; Song, J. A review of the effects of ATP and hydroxychloroquine on the phase separation of the SARS-CoV-2 nucleocapsid protein. *Biophys. Rev.* **2022**, *14*, 709–715. [CrossRef]

218. Zhou, Q.; Usluer, S.; Zhang, F.; Lenard, A.J.; Bourgeois, B.M.R.; Madl, T. ATP regulates RNA-driven cold inducible RNA binding protein phase separation. *Protein Sci.* **2021**, *30*, 1438–1453. [CrossRef]

219. Ren, C.-L.; Shan, Y.; Zhang, P.; Ding, H.-M.; Ma, Y.-Q. Uncovering the molecular mechanism for dual effect of ATP on phase separation in FUS solution. *Sci. Adv.* **2022**, *8*, eabo7885. [CrossRef]

220. Kota, D.; Prasad, R.; Zhou, H.-X. Adenosine Triphosphate Mediates Phase Separation of Disordered Basic Proteins by Bridging Intermolecular Interaction Networks. *J. Am. Chem. Soc.* **2024**, *146*, 1326–1336. [CrossRef]

221. Zhou, H.-X.; Kota, D.; Qin, S.; Prasad, R. Fundamental Aspects of Phase-Separated Biomolecular Condensates. *Chem. Rev.* **2024**, *124*, 8550–8595. [CrossRef]

222. Li, M.; Shi, J.; Wei, Z.; Teng, F.Y.; Tang, B.L.; Song, J. Structural characterization of the human Nogo-A functional domains. Solution structure of Nogo-40, a Nogo-66 receptor antagonist enhancing injured spinal cord regeneration. *Eur. J. Biochem.* **2004**, *271*, 3512–3522. [CrossRef]

223. Sahai, N.; Kaddour, H.; Dalai, P. The transition from geochemistry to biogeochemistry. *Elements* **2016**, *12*, 389–394. [CrossRef]

224. Dalai, P.; Sahai, N. Mineral–lipid interactions in the origins of life. *Trends Biochem. Sci.* **2019**, *44*, 331–341. [CrossRef] [PubMed]

225. Schrum, J.P.; Zhu, T.F.; Szostak, J.W. The origins of cellular life. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a002212. [CrossRef] [PubMed]

226. Kempe, S.; Kazmierczak, J. Biogenesis and early life on earth and Europa: Favored by an alkaline ocean? *Astrobiology* **2002**, *2*, 123–130. [CrossRef]

227. Brack, A. From interstellar amino acids to prebiotic catalytic peptides: A review. *Chem. Biodivers.* **2007**, *4*, 665–679. [CrossRef]

228. Deamer, D.W. The first living systems: A bioenergetic perspective. *Microbiol. Mol. Biol. Rev.* **1997**, *61*, 239–261. [CrossRef]

229. Fox, S.W. The evolutionary significance of phase-separated microsystems. *Orig. Life* **1976**, *7*, 49–68. [CrossRef]

230. Tang, T.-Y.D.; Hak, C.R.C.; Thompson, A.J.; Kuimova, M.K.; Williams, D.S.; Perriman, A.W.; Mann, S. Fatty acid membrane assembly on coacervate microdroplets as a step towards a hybrid protocell model. *Nat. Chem.* **2014**, *6*, 527–533. [CrossRef]

231. Lai, Y.-C.; Liu, Z.; Chen, I.A. Encapsulation of ribozymes inside model protocells leads to faster evolutionary adaptation. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2025054118. [CrossRef]

232. Zhu, T.F.; Szostak, J.W. Coupled growth and division of model protocell membranes. *J. Am. Chem. Soc.* **2009**, *131*, 5705–5713. [CrossRef]

233. Koga, S.; Williams, D.S.; Perriman, A.W.; Mann, S. Peptide–nucleotide microdroplets as a step towards a membrane-free protocell model. *Nat. Chem.* **2011**, *3*, 720–724. [CrossRef]

234. Lazcano, A.; Alexandr, I. Oparin and the origin of life: A historical reassessment of the heterotrophic theory. *J. Mol. Evol.* **2016**, *83*, 214–222. [CrossRef]

235. Xu, C.; Martin, N.; Li, M.; Mann, S. Living material assembly of bacteriogenic protocells. *Nature* **2022**, *609*, 1029–1037. [CrossRef] [PubMed]

236. Li, M.; Huang, X.; Tang, T.-Y.D.; Mann, S. Synthetic cellularity based on non-lipid micro-compartments and protocell models. *Curr. Opin. Chem. Biol.* **2014**, *22*, 1–11. [CrossRef] [PubMed]

237. Pastorek, A.; Ferus, M.; Čuba, V.; Šrámek, O.; Ivanek, O.; Civiš, S. Primordial radioactivity and prebiotic chemical evolution: Effect of $\gamma$ radiation on formamide-based synthesis. *J. Phys. Chem. B* **2020**, *124*, 8951–8959. [CrossRef] [PubMed]

238. Dartnell, L.R. Ionizing radiation and life. *Astrobiology* **2011**, *11*, 551–582. [CrossRef] [PubMed]

239. Friedberg, E.C.; Walker, G.C.; Siede, W.; Wood, R.D.; Schultz, R.A.; Ellenberger, T. *DNA Repair and Mutagenesis*; ASM Press: Washington, DC, USA, 1995.

240. Dai, S.; Xie, Z.; Wang, B.; Ye, R.; Ou, X.; Wang, C.; Yu, N.; Huang, C.; Zhao, J.; Cai, C.; et al. An inorganic mineral-based protocell with prebiotic radiation fitness. *Nat. Commun.* **2023**, *14*, 7699. [CrossRef]

241. Kornberg, A.; Rao, N.N.; Ault-Riché, D. Inorganic polyphosphate: A molecule of many functions. *Annu. Rev. Biochem.* **1999**, *68*, 89–125. [CrossRef]

242. Rao, N.N.; Gómez-García, M.R.; Kornberg, A. Inorganic polyphosphate: Essential for growth and survival. *Annu. Rev. Biochem.* **2009**, *78*, 605–647. [CrossRef]

243. Yamagata, Y.; Watanabe, H.; Saitoh, M.; Namba, T. Volcanic production of polyphosphates and its relevance to prebiotic evolution. *Nature* **1991**, *352*, 516–519. [CrossRef]

244. Brown, M.R.W.; Kornberg, A. Inorganic polyphosphate in the origin and survival of species. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16085–16087. [CrossRef] [PubMed]

245. Keating, C.D. Aqueous phase separation as a possible route to compartmentalization of biological molecules. *Acc. Chem. Res.* **2012**, *45*, 2114–2124. [CrossRef] [PubMed]

246. Nguyen, T.Q.; Dziuba, N.; Lindahl, P.A. Isolated Saccharomyces cerevisiae vacuoles contain low-molecular-mass transition-metal polyphosphate complexes. *Metallomics* **2019**, *11*, 1298–1309. [CrossRef] [PubMed]

247. Dahl, J.-U.; Gray, M.J.; Bazopoulou, D.; Beaufay, F.; Lempart, J.; Koenigsknecht, M.J.; Wang, Y.; Baker, J.R.; Hasler, W.L.; Young, V.B.; et al. The anti-inflammatory drug mesalamine targets bacterial polyphosphate accumulation. *Nat. Microbiol.* **2017**, *2*, 16267. [CrossRef]

248. Racki, L.R.; Tocheva, E.I.; Dieterle, M.G.; Sullivan, M.C.; Jensen, G.J.; Newman, D.K. Polyphosphate granule biogenesis is temporally and functionally tied to cell cycle exit during starvation in *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E2440–E2449. [CrossRef]

249. Gray, M.J.; Wholey, W.-Y.; Wagner, N.O.; Cremers, C.M.; Mueller-Schickert, A.; Hock, N.T.; Krieger, A.G.; Smith, E.M.; Bender, R.A.; Bardwell, J.C.; et al. Polyphosphate is a primordial chaperone. *Mol. Cell* **2014**, *53*, 689–699. [CrossRef]

250. Guan, J.; Jakob, U. The Protein Scaffolding Functions of Polyphosphate. *J. Mol. Biol.* **2024**, *436*, 168504. [CrossRef]

251. Van Meer, G.; Voelker, D.R.; Feigenson, G.W. Membrane lipids: Where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* **2008**, *9*, 112–124. [CrossRef]

252. Renne, M.F.; Ernst, R. Membrane homeostasis beyond fluidity: Control of membrane compressibility. *Trends Biochem. Sci.* **2023**, *48*, 963–977. [CrossRef]

253. Zhukov, A.; Popov, V. Eukaryotic Cell Membranes: Structure, Composition, Research Methods and Computational Modelling. *Int. J. Mol. Sci.* **2023**, *24*, 11226. [CrossRef]

254. Duché, G.; Sanderson, J.M. The Chemical Reactivity of Membrane Lipids. *Chem. Rev.* **2024**, *124*, 3284–3330. [CrossRef] [PubMed]

255. Remick, K.A.; Helmann, J.D. The elements of life: A biocentric tour of the periodic table. *Adv. Microb. Physiol.* **2023**, *82*, 1–127. [PubMed]

256. Randolph-Flagg, N.G.; Ely, T.; Som, S.M.; Shock, E.L.; German, C.R.; Hoehler, T.M. Phosphate availability and implications for life on ocean worlds. *Nat. Commun.* **2023**, *14*, 2388. [CrossRef] [PubMed]

257. Nath, S. Phosphorus Chemistry at the Roots of Bioenergetics: Ligand Permutation as the Molecular Basis of the Mechanism of ATP Synthesis/Hydrolysis by $F_OF_1$-ATP Synthase. *Molecules* **2023**, *28*, 7486. [CrossRef]

258. Wimmer, J.L.E.; Kleinermanns, K.; Martin, W.F. Pyrophosphate and Irreversibility in Evolution, or why PP(i) Is Not an Energy Currency and why Nature Chose Triphosphates. *Front Microbiol.* **2021**, *12*, 759359. [CrossRef]

259. Todd, A.R. *Where There's Life There's Phosphorus*; Nakamura, K., Kageyama, M., Oshima, T., Eds.; Science and Society Press: Tokyo, Japan, 1981; p. 275.

260. Westheimer, F.H. Why nature chose phosphates. *Science* **1987**, *235*, 1173–1178. [CrossRef]

261. Bowler, M.W.; Cliff, M.; Walthoand, J.P.; Blackburn, G.M. Why did nature select phosphate for its dominant roles in biology? *New J. Chem.* **2010**, *34*, 784–794. [CrossRef]

262. Kamerlin, S.C.L.; Sharma, P.K.; Prasad, R.B.; Warshel, A. Why nature really chose phosphate. *Q. Rev. Biophys.* **2013**, *46*, 1–132. [CrossRef]

263. Kaiser, C.M.; Goldman, D.H.; Chodera, J.D.; Tinoco, I.; Bustamante, C. The ribosome modulates nascent protein folding. *Science* **2011**, *334*, 1723–1727. [CrossRef]

264. Waudby, C.A.; Dobson, C.M.; Christodoulou, J. Nature and Regulation of Protein Folding on the Ribosome. *Trends Biochem. Sci.* **2019**, *44*, 914–926. [CrossRef]

265. Das, B.; Chattopadhyay, S.; Bera, A.K.; Dasgupta, C. In vitro protein folding by ribosomes from Escherichia coli, wheat germ and rat liver: The role of the 50S particle and its 23S rRNA. *Eur. J. Biochem.* **1996**, *235*, 613–621. [CrossRef] [PubMed]

266. Chattopadhyay, S.; Pal, S.; Pal, D.; Sarkar, D.; Chandra, S.; Das Gupta, C. Protein folding in Escherichia coli: Role of 23S ribosomal RNA. *Biochim. Biophys. Acta* **1999**, *1429*, 293–298. [CrossRef] [PubMed]

267. Horowitz, S.; Bardwell, J.C. RNAs as chaperones. *RNA Biol.* **2016**, *13*, 1228–1231. [CrossRef] [PubMed]

268. Docter, B.E.; Horowitz, S.; Gray, M.J.; Jakob, U.; Bardwell, J.C. Do nucleic acids moonlight as molecular chaperones? *Nucleic Acids Res.* **2016**, *44*, 4835–4845. [CrossRef]

269. Litberg, T.J.; Horowitz, S. Roles of Nucleic Acids in Protein Folding, Aggregation, and Disease. *ACS Chem. Biol.* **2024**, *19*, 809–823. [CrossRef]

270. Hautke, A.; Voronin, A.; Idiris, F.; Riel, A.; Lindner, F.; Lelièvre-Büttner, A.; Zhu, J.; Appel, B.; Fatti, E.; Weis, K.; et al. CAG-repeat RNA hairpin folding and recruitment to nuclear speckles with a pivotal role of ATP as a cosolute. *J. Am. Chem. Soc.* **2023**, *145*, 9571–9583. [CrossRef]

271. Hautke, A.; Ebbinghaus, S. The emerging role of ATP as a cosolute for biomolecular processes. *Biol. Chem.* **2023**, *404*, 897–908. [CrossRef]

272. Wang, L.; Lim, L.; Dang, M.; Song, J. A novel mechanism for ATP to enhance the functional oligomerization of TDP-43 by specific binding. *Biochem. Biophys. Res. Commun.* **2019**, *514*, 809–814. [CrossRef]

273. He, Y.; Kang, J.; Lim, L.; Song, J. ATP binds nucleic-acid-binding domains beyond RRM fold. *Biochem. Biophys. Res. Commun.* **2020**, *522*, 826–831. [CrossRef]

274. Hu, G.; Ou, X.; Li, J. Mechanistic Insight on General Protein-Binding Ability of ATP and the Impacts of Arginine Residues. *J. Phys. Chem. B* **2022**, *126*, 4647–4658. [CrossRef]

275. Ou, X.; Lao, Y.; Xu, J.; Wutthinitikornkit, Y.; Shi, R.; Chen, X.; Li, J. ATP Can Efficiently Stabilize Protein through a Unique Mechanism. *JACS Au* **2021**, *1*, 1766–1777. [CrossRef]

276. Kim, T.H.; Payliss, B.J.; Nosella, M.L.; Lee, I.T.W.; Toyama, Y.; Forman-Kay, J.D.; Kay, L.E. Interaction hot spots for phase separation revealed by NMR studies of a CAPRIN1 condensed phase. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2104897118. [CrossRef] [PubMed]

277. Tian, Z.; Qian, F. Adenosine Triphosphate-Induced Rapid Liquid–Liquid Phase Separation of a Model IgG1 mAb. *Mol. Pharm.* **2021**, *18*, 267–274. [CrossRef] [PubMed]

278. Zhu, Q.; Wu, Y.; Luo, R. Understanding and fine tuning the propensity of ATP-driven liquid–liquid phase separation with oligolysine. *Phys. Chem. Chem. Phys.* **2024**, *26*, 10568–10578. [CrossRef]

279. Kamski-Hennekam, E.R.; Huang, J.; Ahmed, R.; Melacini, G. Toward a molecular mechanism for the interaction of ATP with alpha-synuclein. *Chem. Sci.* **2023**, *14*, 9933–9942. [CrossRef] [PubMed]

280. Mori, T.; Yoshida, N. Tuning the ATP–ATP and ATP–disordered protein interactions in high ATP concentration by altering water models. *J. Chem. Phys.* **2023**, *159*, 035102. [CrossRef] [PubMed]

281. Zalar, M.; Bye, J.; Curtis, R. Nonspecific Binding of Adenosine Tripolyphosphate and Tripolyphosphate Modulates the Phase Behavior of Lysozyme. *J. Am. Chem. Soc.* **2023**, *145*, 929–943. [CrossRef]

282. Chu, X.-Y.; Wang, G.; Zhang, H.-Y. ATP as an anti-aging agent: Beyond the energy reservoir. *Drug Discov. Today* **2021**, *26*, 2783–2785. [CrossRef]

283. Chu, X.-Y.; Xu, Y.-Y.; Tong, X.-Y.; Wang, G.; Zhang, H.-Y. The Legend of ATP: From Origin of Life to Precision Medicine. *Metabolites* **2022**, *12*, 461. [CrossRef]

284. Cvekl, A.; Vijg, J. Aging of the eye: Lessons from cataracts and age-related macular degeneration. *Ageing Res. Rev.* **2024**, *99*, 102407. [CrossRef]

285. Loh, D.; Reiter, R.J. The mitochondria chronicles of melatonin and ATP: Guardians of phase separation. *Mitochondrial Commun.* **2024**, *2*, 67–84. [CrossRef]

286. Dec, R.; Jaworek, M.W.; Dzwolak, W.; Winter, R. Liquid-Droplet-Mediated ATP-Triggered Amyloidogenic Pathway of Insulin-Derived Chimeric Peptides: Unraveling the Microscopic and Molecular Processes. *J. Am. Chem. Soc.* **2023**, *145*, 4177–4186. [CrossRef] [PubMed]

287. Sarkar, S.; Mondal, J. Mechanistic Insights on ATP's Role as a Hydrotrope. *J. Phys. Chem. B* **2021**, *125*, 7717–7731. [CrossRef] [PubMed]

288. Dec, R.; Puławski, W.; Dzwolak, W. Selective and stoichiometric incorporation of ATP by self-assembling amyloid fibrils. *J. Mater. Chem. B* **2021**, *9*, 8626–8630. [CrossRef]

289. Do, T.M.; Horinek, D.; Matubayasi, N. How ATP suppresses the fibrillation of amyloid peptides: Analysis of the free-energy contributions. *Phys. Chem. Chem. Phys.* **2024**, *26*, 11880–11892. [CrossRef]

290. Xu, Z.; Tang, J.; Gong, Y.; Zhang, J.; Zou, Y. Atomistic Insights into the Stabilization of TDP-43 Protofibrils by ATP. *J. Chem. Inf. Model.* **2024**, *64*, 7639–7649. [CrossRef]

291. Sarkar, S.; Gupta, S.; Mahato, C.; Das, D.; Mondal, J. Elucidating ATP's Role as Solubilizer of Biomolecular Aggregate. *eLife* **2024**, *13*, RP99150. [CrossRef]

292. Hinton, G.E.; Salakhutdinov, R.R. Reducing the dimensionality of data with neural networks. *Science* **2006**, *313*, 504–507. [CrossRef]

293. Riddle, D.S.; Santiago, J.V.; Bray-Hall, S.T.; Doshi, N.; Grantcharova, V.P.; Yi, Q.; Baker, D. Functional rapidly folding proteins from simplified amino acid sequences. *Nat. Struct. Biol.* **1997**, *4*, 805–809. [CrossRef]

294. Baek, M.; DiMaio, F.; Anishchenko, I.; Dauparas, J.; Ovchinnikov, S.; Lee, G.R.; Wang, J.; Cong, Q.; Kinch, L.N.; Schaeffer, R.D.; et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **2021**, *373*, 871–876. [CrossRef]

295. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [CrossRef] [PubMed]

296. Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.; Ronneberger, O.; Willmore, L.; Ballard, A.J.; Bambrick, J.; et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **2023**, *630*, 493–500. [CrossRef] [PubMed]

297. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [CrossRef]

298. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]

299. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.

300. Kauzmann, W. Some factors in the interpretation of protein denaturation. *Adv. Protein Chem.* **1959**, *14*, 1–63.

301. Levitt, M.; Warshel, A. Computer simulation of protein folding. *Nature* **1975**, *253*, 694–698. [CrossRef]

302. Roux, B. The calculation of the potential of mean force using computer simulations. *Comput. Phys. Commun.* **1995**, *91*, 275–282. [CrossRef]

303. Jorgensen, W.L.; Chandrasekhar, J.; Madura, J.D.; Impey, R.W.; Klein, M.L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926–935. [CrossRef]

304. Ahlers, J.; Adams, E.M.; Bader, V.; Pezzotti, S.; Winklhofer, K.F.; Tatzelt, J.; Havenith, M. The key role of solvent in condensation: Mapping water in liquid-liquid phase-separated FUS. *Biophys. J.* **2021**, *120*, 1266–1275. [CrossRef]

305. Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249. [CrossRef]

306. Stumm, W.; Morgan, J.J. *Aquatic Chemistry, Chemical Equilibria and Rates in Natural Waters*, 3rd ed.; John Wiley Sons, Inc.: New York, NY, USA, 1996.

307. Barabási, A.L.; Oltvai, Z. Network biology: Understanding the cell's functional organization. *Nat. Rev. Genet.* **2004**, *5*, 101–113. [CrossRef] [PubMed]

308. Watts, D.J.; Strogatz, S.H. Collective dynamics of 'small-world' networks. *Nature* **1998**, *393*, 440–442. [CrossRef] [PubMed]

309. Cybenko, G. Approximation by superpositions of a sigmoidal function. *Math. Control Signals Syst.* **1989**, *2*, 303–314. [CrossRef]

310. Hornik, K.; Stinchcombe, M.; White, H. Multilayer feedforward networks are universal approximators. *Neural Netw.* **1989**, *2*, 359–366. [CrossRef]

311. Telgarsky, M. Benefits of depth in neural networks. In Proceedings of the 29th Annual Conference on Learning Theory, New York, NY, USA, 23–26 June 2016; Volume 49, pp. 1517–1539.

312. Bengio, Y.; Courville, A.C.; Vincent, P. Representation Learning: A Review and New Perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**. [CrossRef]

313. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. In Proceedings of the 27th International Conference on Neural Information Processing Systems (NIPS'14), Montreal, QC, Canada, 8–13 December 2014; Volume 2, pp. 2672–2680.

314. Holland, J.H. *Complexity: A Very Short Introduction, Very Short Introductions*; Oxford Academic: Oxford, UK, 2014.

315. Templeton, A.; Conerly, T.; Marcus, J.; Lindsey, J.; Bricken, T.; Chen, B.; Pearce, A.; Citro, C.; Ameisen, E.; Jones, A.; et al. Scaling Monosemanticity: Extracting Interpretable Features from Claude 3 Sonnet. *Transform. Circuits Thread* **2024**. Available online: https://transformer-circuits.pub/2024/scaling-monosemanticity/index.html (accessed on 9 October 2024).

316. Kinsey, L.J.; Beane, W.S.; Tseng, K.A.-S. Accelerating an integrative view of quantum biology. *Front. Physiol.* **2024**, *14*, 1349013. [CrossRef]

317. Tegmark, M. Importance of quantum decoherence in brain processes. *Phys. Rev. E* **2000**, *61*, 4194–4206. [CrossRef]

318. Penrose, R. *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*; Oxford University Press: Oxford, UK, 1989.