


Review

Life in Phases: Intra- and Inter- Molecular Phase Transitions in Protein Solutions

Vladimir N. Uversky^{1,2,*}  and Alexei V. Finkelstein^{3,4,5,*}

¹ Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL 33612, USA

² Laboratory of New Methods in Biology, Institute for Biological Instrumentation, Russian Academy of Sciences, Federal Research Center “Pushchino Scientific Center for Biological Research of the Russian Academy of Sciences”, 142290 Pushchino, Moscow, Russia

³ Institute of Protein Research, Russian Academy of Sciences, 142290 Pushchino, Moscow, Russia

⁴ Biology Department, Lomonosov Moscow State University, 119192 Moscow, Russia

⁵ Biotechnology Department, Lomonosov Moscow State University, 142290 Pushchino, Moscow, Russia

* Correspondence: vuvversky@usf.edu (V.N.U.); afinkel@vega.protres.ru (A.V.F.);

Tel.: +1-831-974-5816 (V.N.U.); +7-496-731-8412 (A.V.F.)

Received: 19 November 2019; Accepted: 6 December 2019; Published: 8 December 2019



Abstract: Proteins, these evolutionarily-edited biological polymers, are able to undergo intramolecular and intermolecular phase transitions. Spontaneous intramolecular phase transitions define the folding of globular proteins, whereas binding-induced, intra- and inter- molecular phase transitions play a crucial role in the functionality of many intrinsically-disordered proteins. On the other hand, intermolecular phase transitions are the behind-the-scenes players in a diverse set of macrosystemic phenomena taking place in protein solutions, such as new phase nucleation in bulk, on the interface, and on the impurities, protein crystallization, protein aggregation, the formation of amyloid fibrils, and intermolecular liquid–liquid or liquid–gel phase transitions associated with the biogenesis of membraneless organelles in the cells. This review is dedicated to the systematic analysis of the phase behavior of protein molecules and their ensembles, and provides a description of the major physical principles governing intramolecular and intermolecular phase transitions in protein solutions.

Keywords: protein folding; protein structure; intrinsically-disordered protein; molten globule; secondary structure; coil; phase transition; liquid–liquid phase separation; membraneless organelle; amyloid fibril; crystal

1. Introduction

Protein structures depend on the interplay of chain conformational entropy and the sum of multiple weak interactions of different physico-chemical natures, which can be considered as “conformational forces” defining the free energy change between the folded and unfolded states that are related to protein stability. In this review, we will focus on proteins that exist mainly in the aqueous environment.

Among the weak, noncovalent interactions stabilizing protein structures are hydrogen bonds (having up to 25–40 kJ/mol in vacuum or nonpolar medium, but only 8–10 kJ/mol in aqueous environment), salt bridges (having up to 100 kJ/mol in the absence of water, but only about 5–10 kJ/mol in aqueous environment), long-range electrostatic interactions (which are weaker but more numerous than salt-bridges, and whose free energy depends on the distance between the charges and on their environment), van der Waals interactions (of about 3 kJ/mol for interaction of two methyl groups), and hydrophobic interactions (free energy of which scales with the size of the solute surface as ≈ 10 kJ/mol/nm², which, for a methyl group with a surface area of about 1 nm², would amount to ≈ 10 kJ/mol) [1]. Since these interactions are extremely condition dependent, the presence

(or absence) of a structure in a query protein is condition-dependent too. Furthermore, due to their different physico-chemical natures, various conformational forces can differently react to changes in environmental conditions. In fact, although high concentrations of strong denaturants, such as guanidinium chloride (GdmCl), guanidinium thiocyanate (GTC), or urea can efficiently suppress all (or almost all) intramolecular conformational interactions leading to an almost complete unfolding of a globular protein into a highly-disordered, random, coil-like conformation [2–5], often, environmental alterations can decrease (or even completely eliminate) part of the conformational interactions, whereas the remaining interactions remain unchanged or even strengthen.

Therefore, although it is commonly believed that all the necessary information for a given protein to correctly fold to the specific, unique, and biologically-active conformation is included in its amino acid sequence [1,6–8], this, actually, only concerns distinct proteins in their physiological environment, while, in general, the crucial effect of environment should not be excluded. In fact, changes in the environment of a globular protein can cause a wide spectrum of structural changes, ranging from an almost complete unfolding in the concentrated solutions of a strong denaturant to a more subtle denaturation (which is typically associated with the loss of both the unique 3D structure and the unique biological activity) under some “mild denaturing conditions”. In other words, the complete unfolding of a protein does not necessarily represent the only consequence of denaturation. Instead, some partially-folded conformations can possess properties that are in-between the properties of the folded and the completely unfolded states. As a result, depending on the peculiarities of their environments, the chains of globular proteins may exist in at least four different states in aqueous media, i.e., their own native (ordered) conformation, molten globule, premolten globule, and unfolded [1,9–21] (Figure 1), not to mention other forms that can be induced by nonaqueous environments, such as alcohols, membranes, or other proteins, as well as by post-translational modifications of their chains.

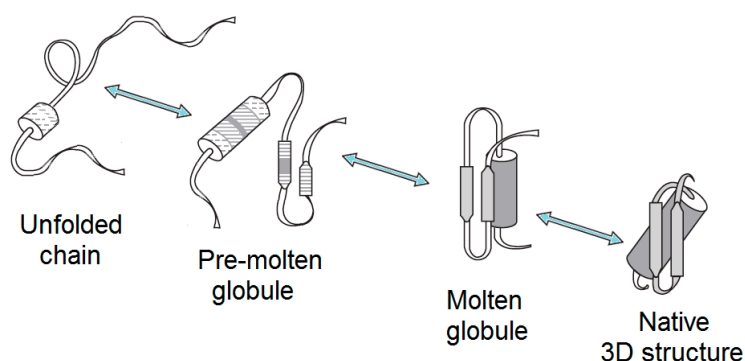


Figure 1. Four main possible stable states of protein molecule: noncompact unfolded chain with, maybe, some traces of secondary structures; swollen “premolten” globule with partly formed secondary structures; compact “molten globule” with almost formed secondary structures and folding pattern, but having no close packing of its mobile side chains; and solid native protein structure [22].

One should keep in mind that although many globular proteins possess clearly defined and unique 3D structures, these structures are rather heterogeneous, with the ordering degrees being greatly diversified in the different parts of a given protein. Such structural heterogeneity is seen in X-ray data as the variability of the values of the *B*-factor characterizing the mobility of separate atoms in a protein [23,24], with the atoms of the active center of an enzyme being typically characterized by the lowest *B*-factor. Additionally, some globular proteins have highly dynamic or even completely unstructured regions (e.g., loops and terminal fragments) that correspond to the regions of missing electron density, being therefore undetectable by X-ray analysis [25–29].

In addition to the “traditional” ordered proteins that “obey” classical function-structure paradigm, where a specific function of a protein is determined by its unique and rigid 3D structure encoded in a unique amino acid sequence encrypted in a corresponding gene, recently, we have witnessed an

increased appreciation of intrinsically-disordered proteins, i.e., biologically-active proteins having no unique structures, at least before interactions with other molecules. In fact, it is recognized now that all organisms from all kingdoms of life (bacteria, archaea, and eukaryotes) and all viruses contain discernible levels of intrinsically-disordered proteins (IDPs) or hybrid proteins containing ordered domains and intrinsically-disordered protein regions (IDPRs) [30–36]. Furthermore, based on computational analyses, it has been concluded that IDPs/IDPRs universally exist in all living organisms [30–32,35,37–51], with the penetrance of disorder in different species increasing with the rise of the organism complexity [30–32,35,52]. Here, the expected fractions of sequences predicted to have long IDPRs (30 residues or longer) are arranged in the following order: Bacteria ~ Archaea << Eukaryotes [30,32,36,38,53]. This increase in the amount of disorder in eukaryotes was attributed to the expansion of the significance of cellular signaling that often relies on IDPs/IDPRs [33,54–60]. Furthermore, only a very small fraction of proteins with known crystal structures in the Protein Data Bank (PDB) is completely devoid of disorder [25,61].

The biological functions of many IDPs/IDPRs are strongly disorder-dependent, and can be described in terms of the entropic chain activities, where an extended random-coil conformation maintaining flexibility while carrying out its function is needed [34,62]. Stochastic machines can serve as illustrative examples of such dynamic signaling complexes with “entropic” chain activities [63]. Structured domains in these stochastic machines are connected by long flexible linkers [63–66]. As a result, the action of such machines does not depend on coordinated conformational changes. Instead, they operate via uncoordinated and stochastic movements of their flexible arms (long, disordered linkers), which, despite being engaged in constant and chaotic movements, can eventually enable productive functionality [63]. However, there is a large set of IDPs/IDPRs that can fold at binding to their partners [55,56,67]. Since various systems may have different degrees of such binding-induced folding, the resulting complexes are characterized by wide structural and functional heterogeneities [68,69]. The existence of such foldable IDPs was used to argue that “a partly disordered polypeptide may be capable of specific recognition through a conformer selection mechanism, but then it is the ordered population that reacts, and the disorder is neither intrinsic nor functional”, and that, therefore, such structureless proteins should be considered as “proteins waiting for a partner” (PWPs), which serve as parts of multi-component complexes that do not fold correctly in the absence of other components [70]. Although based on these arguments it was concluded that molecular disorder is not compatible with protein function [70], and although protein functionality seems to commonly originate from disorder-to-order transitions, the existence of the opposite scenario was emphasized, where the local or even global functional unfolding of ordered proteins represents an important prerequisite for their functionality [71]. This regulated unfolding [72], or dormant or conditional disorder, shows induced [73] and transient character [74]. It can be awoken by a broad range of environmental factors, such as the release of autoinhibition, light exposure, changes in redox potential, mechanical force, various posttranslational modifications, or changes in pH or temperature, or by specific interactions with ligands, membranes, nucleic acids, or other proteins [71]. It was pointed out that the existence of such dormant disorder phenomenon signifies the global importance of intrinsic disorder for protein functionality.

Since flexible protein regions are known to serve as the primary targets for the proteolytic attacks of various proteases [75–80], the fact that many IDPs and IDPRs are not cleaved by cellular proteases represents a very interesting conundrum pertaining to the peculiarities of the cellular life of intrinsically-disordered proteins and hybrid proteins with ordered domains and disordered regions. In fact, it used to be commonly believed that only folding can protect polypeptide chains against proteolysis. In line with this conjecture, it is known that IDPs are exceptionally sensitive to proteolysis *in vitro* [81–83], with this high proteolytic sensitivity being considered one of the characteristic features of IDPs/IDPRs [33,34,55,56,84–86]. Furthermore, for signaling IDPs, it was postulated that fast cellular degradation represents one of the mechanisms of their functional regulation [34,56,57,62,87–89]. However, not all IDPs/IDPRs are degraded rapidly inside the cells, and analyses of the available data on

intracellular protein half-lives suggest that the presence of intrinsic disorder in a protein does not imply a significantly shorter half-life of its carrier [90,91]. One of the potential explanations for this phenomenon is based on the hypothesis that in the cellular milieu, many IDPs/IDPRs exist largely in a bound state, thereby possessing context-dependent resistance to proteolysis [91]. Several different mechanisms were proposed for such binding-induced protection against cellular degradation by default. In fact, some IDPs/IDPRs can always be bound to their specific biological partners. In such cases, binding-induced structural stabilization of IDPs might represent a “by-product” of their functional interactions. For other IDPs/IDPRs, such cellular degradation by default is prevented by interactions with specific binding partners (e.g., proteasome gate keepers and nanny proteins) that interact with IDPs solely or principally to prevent their degradation [91]. Still other IDPs/IDPRs are protected from degradation via interactions with “decoy” DNA binding sites or via intramolecular interactions that minimize the number of relatively unconstrained long IDPRs capable of accessing 20S core proteasome [91]. Importantly, despite the obvious hypothesis that chaperones may offer direct protection from cellular degradation by default by preferential binding IDPs/IDPRs in the cell, comprehensive bioinformatics analysis of chaperone-binding and nonchaperone-binding proteins in three taxonomic groups revealed that there is, in fact, a negative correlation between the intrinsic disorder propensity in proteins and their tendency to be binding partners of chaperones [81]. In other words, the preferential cellular partners of chaperones are ordered proteins, which requires more assistance for folding and protection from misfolding and aggregation than IDPs/IDPRs [81]. Among the other molecular mechanisms used to explain the protection of IDPs/IDPRs from cellular proteases are specific amino acid compositions of disordered proteins and regions and tight regulation of intracellular proteases [81]. Furthermore, the stability of proteins can be regulated by various post-translational modifications (PTMs), and it was shown that IDPs/IDPRs can serve as substrates of twice as many kinases as ordered proteins [92,93]. Since in addition to the conformational stability of proteins, PTMs can affect their activity, folding, interactions, turnover, and localization, and since IDPs/IDPRs are often serve as primary targets for the modifying enzymes [33,56,60,88,94–99], it is likely that PTMs have other roles in protecting IDPs against cellular degradation by default, e.g., by ensuring preferential localization of IDPs and hybrid proteins to the cellular compartments that do not contain proteases. An illustrative example of such compartmentalization is given by some proteinaceous membraneless organelles (PMLOs), where IDPs/IDPRs can be protected from the proteolytic degradation by the preferential protease exclusion from these organelles [100]. In conclusion, although several potential explanations for the remarkable IDP/IDPR protection against cellular degradation by default were given, it seems that there is no universal mechanism for such protection, indicating that protein degradation is not determined by a single characteristic, such as intrinsic disorder propensity, representing a complex, multifactorial process with prominent protein-to-protein variations.

An important feature of a polypeptide chain is its ability to undergo intramolecular and intermolecular phase transitions. The discovery that denaturation of the globular, native, rigid, one-domain protein structure can occur as an “all-or-none” transition [101,102], i.e., a transition without accumulation of visible intermediates, is one of the cornerstones of protein physics. In this article, we consider several aspects of such phase transitions, paying special attention to the spontaneous intramolecular phase transitions in the solutions of ordered proteins undergoing either equilibrium “native state—molten globule—unfolded chain” conformational transitions or kinetic “native state \leftrightarrow unfolded chain” structural transitions, induced intramolecular phase transitions in the solutions of IDPs promoted by their interactions with binding partners, transitions associated with the nucleation in bulk, on the interface, and on the impurities, as well as macrosystemic phenomena of protein aggregation, the formation of amyloid fibrils, and protein crystallization, and intermolecular liquid–liquid or liquid–gel phase transitions that are associated with the biogenesis of membraneless organelles in the cells.

2. Intramolecular Phase Transitions of Ordered Proteins

2.1. Brief Description of Major Partially-Folded States of Globular Proteins

2.1.1. Molten Globule

The structural properties of a protein molecule in the molten globule are well known, and have been outlined in a number of reviews (e.g., [12,13,79,103–117]). Although the molten globular protein completely lacks a rigid, cooperatively-melting, tertiary structure, or has only a trace of such a structure, i.e., is denatured, it preserves high levels of a native-like, secondary structure [12,13,79,103–118]. Molten globules are compact (in comparison with the native state, their hydrodynamic radii are increased by less than 15%, which translates into a ~50% increase in volume) [12,13,18,21,79,103–117,119] and, as evidenced by small-angle X-ray scattering, have a globular structure, which is usually similar to that of native globular proteins [120–124]. The analysis of the molten globules of several proteins by solution NMR spectroscopy revealed that a protein molecule in this intermediate state does not only have a native-like secondary structure, but that it also shows a native-like folding pattern [114,125–135]. This is not surprising, because what is good for the molten globule is probably good for a well-folded protein structure as well. However, there are interesting exceptions; the best known one concerns the molten globule of β -lactoglobulin [136]. Although the native protein is a β -structural molecule, its folding intermediate contains nonnative α -helices, according to the CD spectra.

Despite their secondary-structure-enriched, relatively compact, and globular conformations, molten globules are characterized by the considerable increase in the accessibility of a protein molecule to proteases [75–80]. Finally, one of the most characteristic features of the molten globule is its high affinity to the hydrophobic fluorescence probes (such as 8-anilinonaphthalene-1-sulfonate, ANS or 1,1'-bis(4-anilino-5-naphthalenesulfonic acid), bis-ANS) [137–140].

2.1.2. Premolten Globule

Similar to molten globules, the premolten, globular, partially-folded intermediate represents a denatured conformation with no rigid tertiary structure. Premolten globules are markedly less compact than the molten globular or native states of a protein with a given molecular mass, although these intermediates are still noticeably more compact than random coils. In fact, compared to the native state, the hydrodynamic volumes of the molten globule, premolten globule, and unfolded states are increased 1.5-, ~3-, and ~12-fold, respectively. However, there is no globular structure in a premolten globular protein [15,124], suggesting that this conformation is likely a partially-ordered form of a “squeezed” coil. In line with this hypothesis, premolten globular protein is characterized by the preservation of considerable levels of secondary structure. However, this residual ordering (protein molecule in the premolten globule state usually has ~50% or even less of native secondary structure) is much less pronounced than that of the molten globule, which typically shows a native-like secondary structure. Finally, at least part of the solvent-accessible hydrophobic clusters is already formed in the premolten globule, as evidenced by the ability of this intermediate to interact with the hydrophobic fluorescent probe ANS [11,13–15,17,116].

Finally, it has been shown that the premolten globule (with a relatively large secondary structure content), as well as the unfolded state (with a low content of residual secondary structure), is separated from the molten globule by a sharp transition, which, in some proteins, represents an “all-or-none” transition, i.e., an intramolecular analog of the first order phase transition [11,13–15,17,116,141]. This means that in these cases, the molten and premolten globules represent different thermodynamic (phase) states of a polypeptide chain [10,141]. It seems that the aforementioned all-or-none transition is due to the formation of a secondary structure within a swollen premolten globule [142], especially since it is known that a β -sheet formation is of an “all-or-none” kind [143–145]. No sharp transition from the premolten state to the random coil has been reported as of yet.

2.2. Thermodynamics of the Protein Denaturation. “Wet” and “Dry” Molten Globules

Before the molten globule state was discovered, protein denaturation was typically thought of as the complete decay of the unique protein structure, i.e., a transition to the coil. After the theoretical prediction [22], and then the experimental discovery of the molten globule [146,147], it became clear that the denatured protein can be rather dense, nearly as much as that of the native protein, as well as loose, like the coil, depending on the solvent’s strength and the hydrophobicity of the protein chain.

To understand the molecular basis of protein denaturation, one has to explain why two equally stable phase states of the protein chain can exist, and why they are separated by a free energy barrier (which is required for an “all-or-none” transition). In other words, one has to explain why the protein globule cannot decay by gradual (barrierless [148] or overcoming a very low free-energy barrier [149]) swelling, as typical polymers do due to the persistent connectivity of their chains [150].

In so doing [151–153], one has to consider the major characteristics of proteins defining their difference from “normal” polymers: (i) each globular protein possesses the only chain fold with a peculiar stability; (ii) flexible side groups are linked to a much more rigid protein-chain backbone; and (iii) the packing of a native globular protein is as tight as the packing of a molecular crystal (although with no crystal lattice), where the van der Waals volumes of atoms occupy 70–80% of the volume, whereas only 60–65% of the volume in liquids (melts) is occupied by the van der Waals volumes of atoms [154].

The side chains of the protein can undergo a rotational isomerization. This is done by jumps between the allowed conformations of the side chains. Each jump necessitates some vacant volume near the side chain that jumps. However, since the native protein structure has a tight packing of the chain (which contributes to the enhanced stability of this fold), each jump needs some extra free volume for landing (see inset in Figure 2).

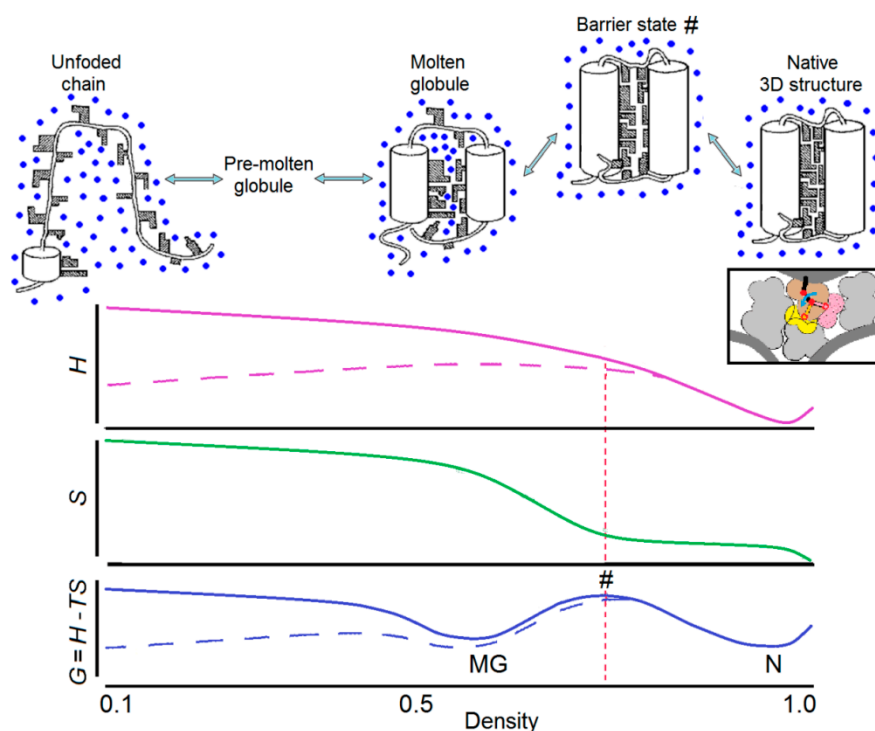


Figure 2. At the top of the picture: The main possible stable states of the protein molecule and the unstable barrier state providing the all-or-none transition between the native structure (N) and the molten globule (MG), and all of the unfolded forms. Blue dots indicate water molecules. Inset: A sketch of a small piece of the close side chain packing. The yellow side chain “head” corresponds to an alternative rotamer of the central side chain, which is forbidden by close packing. At the bottom of the picture: Enthalpy H , entropy S , and free energy G of the protein molecule, depending on its uniform density. T is

the temperature of the MG \leftrightarrow N equilibrium in a “bad” solvent. The dashed lines correspond to a “better” solvent. As is customary in the literature on protein folding theory, the “entropy” S does not include the solvent entropy; correspondingly, “enthalpy” H means, actually, the “free energy of interactions” (also called the “mean force potential”), since, e.g., the hydrophobic, electrostatic and other solvent-mediated forces, with all their solvent entropy, are included in this “enthalpy”. Adapted from [1,153].

Note that the flexible side groups sit at the rigid backbone. The backbone is especially rigid inside the globule, where the α - and β -structures hide H-bonds of their polar peptide groups from the dense hydrophobic environment, and these α - and β -structures are stable, at least until water molecules penetrate into the globule (which requires about the same free volume as the side chain jumps). Therefore, the free volume can be hardly made for a separate jumping side chain, and each of the rigid secondary structure elements, with the entire forest of flexible side chains attached, moves as a whole (at least at the very beginning of the globule’s expansion). Therefore, the expansion of the closely-packed globule, carried out by the moving apart of the rigid α - and β -structures, creates about the same amount of free space near each side group; these spaces are either insufficient for the isomerization of each of the side groups (when the globule expansion is still too small), or are already sufficient for the isomerization of many of them. This means that liberation of the side groups (as well as water penetration) can occur only when the globule expansion crosses a particular threshold, i.e., the “barrier”.

Analysis of the properties of a protein globule at different levels of its uniform expansion [151–153] shows that an expanded state of the protein globule can be as stable as its native (solid) state, but only after the density barrier has been passed. (It should be noted here that this analysis of a uniform globule’s expansion, illustrated by Figure 2, does not aim to model the protein unfolding kinetics, which occurs via intramolecular separation of the native and denatured phases, as shown in Figure 3a below).

Thus, a small expansion of the compact native protein globule is *always* unfavorable [151–153], because it *already* increases the globule’s energy (whose parts already lose their close packing), but does not yet increase the globule’s entropy (since it does not yet liberate the rotational isomerization of the side groups) or allow entry of water into the protein core. That is, the globule’s free energy always increases with a small expansion. In contrast, a large globule’s expansion liberates the rotational isomerization of the side groups and leads (at high enough temperature) to a decrease of the free energy. As a result, protein denaturation occurs not gradually, but as a jump over the free energy barrier, leading to the “all-or-none” kind of transition (Figure 2).

The aforementioned mechanism is related to the transition of a native globular state to any denatured form: molten globule, premolten globule, or coil [141,152]. Therefore, the protein structure tolerates, without significant change, a change of ambient conditions up to a certain limit, and then melts as a whole, like a macroscopic crystal. This provides the reliability of its biological functioning. Put differently, a sudden jump in entropy (mainly entropy of the side chains), which may happen only after the expansion of the globule crosses a particular threshold, explains the origin of the “all-or-none” transition separating the native and denatured state. Such a global entropy jump happens because of the fact that the side chains cannot be liberated one-by-one, since they are held by rigid backbone that coordinates their positions.

The pores in the molten globule are usually “wet,” that is, they are occupied by the solvent, because a water molecule inside the protein is still better than a vacuum [151–153]. Experimentally, the “wetness” of the molten globule is proven by the absence of a visible decrease in the protein floating density [155] after denaturation of any kind.

When the solvent sticks to the protein core (consisting mainly of hydrophobic groups) not too strongly, it only occupies the pores that have been already formed in the molten globule core to ensure side-chain movements, but it does not expand the globule (just as water does not expand a sponge, although it occupies its pores), and does not make new pores. Then, the denatured protein remains the wet molten globule [152,155].

Like the “wet” molten globule, a “dry” molten globule (having no water in its pores) was predicted theoretically [152]; an analysis showed that the dry molten globule should be less stable than the wet one, and therefore, that it is hardly suited to playing the role of a stable, accumulating intermediate in protein melting. However, it has been found [156] that the dry molten globule emerges during fluctuations preceding protein melting.

The molten globule compactness is maintained by the residual hydrophobic interactions of its side groups. They were found to be not very strong. Even in the apomyoglobin, the molten globule (which has a well-developed secondary structure and almost native chain topology of packing of the most of its chain [157]), the residual interactions between the hydrophobic residues appear to be three or four times weaker than those in the native protein [158]; these residual interactions are entirely missing for some hydrophobic residues (which accentuates the heterogeneity of the molten globule; see [13]).

If the residual hydrophobic interactions are weak, i.e., if either the hydrophobicity of the chain is low or the protein chain strongly attracts the solvent, the solvent starts to expand the pores, and the globule swells. The greater the attraction between the solvent and the protein chain, and the smaller the attraction within the protein chain, the greater the chain swelling, leading to the transition to the premolten globule and then to the random coil.

2.3. Kinetics of the “Unfolded Chain ↔ Native State” Transitions

The ability of polypeptide chains of globular proteins to spontaneously form their spatial structures is a long-standing puzzle in molecular biology. Numerous pieces of evidence (such as the independence of the protein structures folded both *in vivo* and *in vitro* on the initial states and configurations of the chains) show that the native protein structure is the most stable of all structures of the chain under physiological conditions [6–8]. Here, it is worth noting that experiments have shown that there is no fundamental difference between the *in vivo* (cotranslational) folding [159–161] and *in vitro* folding of truncated and complete chains [162], at least for small proteins; in both cases, native-like structures emerge only after the entire sequence is available.

The experimentally-measured folding times range from microseconds for small to hours for large single-domain globular proteins; the difference (about 10 orders of magnitude) is the same as that between the life span of a mosquito and the age of the Universe. But these microseconds or even hours are negligible compared to the time necessary to iterate over all possible structures of the protein chain and to find the most stable of them; this requires something like 3^L or even 10^L picoseconds (where L , usually ~ 100 , is the number of amino acid residues in the protein chain), i.e., billions of years [163,164]. Consideration of this “Levinthal’s paradox” led to the idea that the energy landscapes of protein chains must be somehow inclined, like funnels, towards the native protein structures; this would facilitate a sequential protein folding [165–167]. Landscapes of this kind can drastically decrease the time required for the protein chain folding by reducing, for these chains, the free-energy barrier, which, ensuring an “all-or-none” transition, separates their unfolded (U) and natively-folded (N) states. It is noteworthy that the energy landscape is, on average, automatically inclined towards the most stable protein structure, because the interactions present in this structure are, on average, stronger than the other ones. As a result, native-like folding intermediates (possessing a part of these native interactions) are, on average, more stable than the “nonnative-like” folds that do not possess them. In line with earlier analytical estimates [168], computer experiments have shown that a model polymer whose random sequence was slightly “edited” to make the free energy of its most stable fold lower than that of any other fold by at least by a few kcal/mol [169,170] finds this most stable fold in a time, which is many orders of magnitude smaller than the time necessary to iterate over all the possible chain structures.

A physical theory that not only solved the Levinthal’s paradox, but also estimated the dependence of the protein folding time on protein size and shape, was first presented in the second half of 1990s [171–173]. This theory considers overcoming the free-energy barrier separating the natively-folded (N) and unfolded (U) states of protein chains. This barrier occurs in both the uniform (Figure 2) and

nonuniform (Figure 3a) expansions of the globule, but the height of the former barrier is proportional to the protein size, while the height of the latter (occurring via intramolecular separation of the native and denatured phases) is much lower, being proportional to the size of the globule cross-section, i.e., the protein chain length to the power of 2/3. Therefore, the main pathway of the $N \leftrightarrow U$ transition goes via just this lower barrier, the essence of which is an intramolecular phase separation.

The developed theory is applicable to protein and “protein-like” sequences, i.e., those having a distinguished chain fold, the free energy of which is lower than that of any other fold by at least a few $k_B T_{melt}$ [169,170,174,175] (where T_{melt} is its melting temperature). In this theory, a special role is played by the point of thermodynamic (and thus kinetic) equilibrium between the N and U states.

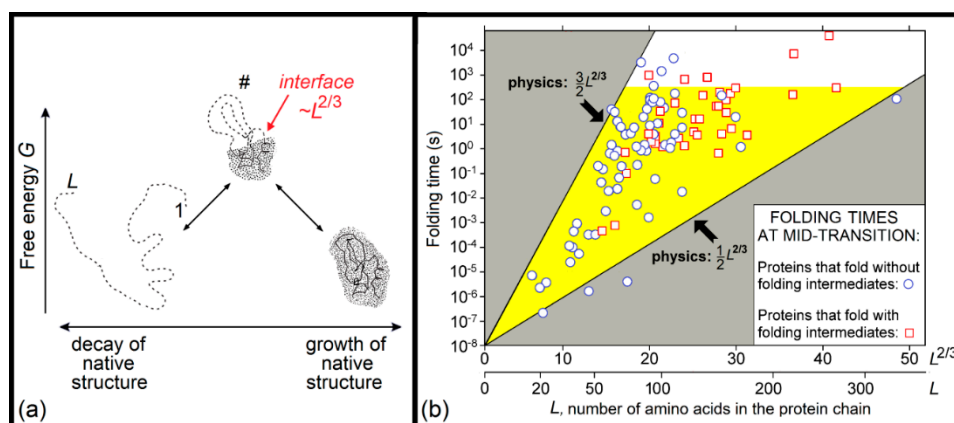


Figure 3. (a) A scheme of the reversible “all-or-none” transition from the unfolded chain to the native globular structure; # marks the rate-determining transition state whose free energy is proportional to the size of the maximal interface of the native and unfolded phases, which scales with the chain length L as $L^{2/3}$. (b) Experimentally-measured in vitro folding times at $N \leftrightarrow U$ equilibrium for 107 single-domain proteins (or separate domains) without SS bonds and covalently bound ligands (although the folding rates for proteins with and without SS bonds are principally the same [163]). Triangle: the region allowed by physics; its golden part corresponds to biologically-reasonable folding times (≤ 10 min) under “biological” ambient conditions; the larger folding times (in the white zone) are observed (for some proteins) only under the equilibrium, i.e., nonbiological conditions. Adapted from [1,176].

Here, the theory obtains the simplest form, because both halves (the native-like and unfolded) of a semi-folded protein have equal free energies, so that the free energy of semi-folded protein is only determined by the interface between these two halves (that is, mainly by the surface free energy of the “native phase”). The maximal unavoidable interface between the N and U states occurring in the course on $U \leftrightarrow N$ transition includes $\approx L^{2/3}$ amino acid residues (L being the number of residues in a protein chain). Therefore, the barrier heights also scale with the protein size as $\approx L^{2/3}$, and, therefore, the corresponding protein folding time scales as $\sim \exp(\approx L^{2/3})$, rather than as $\sim \exp(\approx L)$, appearing in the Levinthal’s paradox. This scaling, $\sim \exp(\approx L^{2/3})$, means that the protein folding time is many orders of magnitude less than the time $\sim \exp(\approx L)$, which is necessary to iterate over all possible chain structures.

A theoretical estimate of the folding time is based on the conventional transition state theory [177–179]. For the $N \leftrightarrow U$ equilibrium, an accurate estimate of the folding (and unfolding) time for a protein chain of L amino acid residues gives

$$\text{TIME} \sim \tau \times \exp[(0.5 \div 1.5)L^{2/3}], \quad (1)$$

where $\tau \approx 10$ ns is the time of the conformational rearrangement of one residue (measured for the helix \leftrightarrow coli transition) [180].

The lower estimate ($\text{TIME} \sim \tau \times \exp[0.5L^{2/3}]$) corresponds to the proteins with “simple” chain folds, which have a transition state (“folding nucleus”) structure where the N–U interface is not covered by the closed unfolded loops; the energy loss for one residue of the phase surface, $\approx 0.5 k_B T_{melt}$, is taken

as $\varepsilon/4$, where $\varepsilon \approx 1.3$ kcal/mol $\approx 2k_B T_{melt}$ is the average heat of native protein melting per residue [102] (this is the first empirical parameter used by the theory, while $\tau \approx 10$ ns is the second and the last used empirical parameter).

The upper estimate (TIME $\sim \tau \times \exp[1.5L^{2/3}]$) corresponds to the proteins with “complicated” chain folds, which have a transition state (“folding nucleus”) structure where the N–U interface is maximally covered by the closed unfolded loops. Strictly speaking, this upper estimate is TIME $\sim \tau \times \exp[(0.5 + \frac{5}{12}\ln(3L^{1/3}))L^{2/3}]$, where the logarithmic term follows [171,181] from averaging the Flory’s estimate for the entropy of closed loops, but for protein chains of a normal size, $\sim 50 \div 200$ residues, this $\frac{5}{12}\ln(3L^{1/3})$ is so close to 1 that there is no need to overcomplicate the simple result given by Equation (1).

Theories developed for the prediction of protein folding nuclei, experimentally studied by Alan Fersht [182] and others that one can find in [183–193], and a more detailed theoretical consideration of folding times for proteins of different sizes, chain folds, and stabilities, are given in [176,181,194–197]. A limited-influence chain knotting and the SS-bonds in the single-domain proteins on the folding rate were estimated in [173,187], and the influence of the native structure stability on the folding rate was estimated in [176] (see also [198]).

The aforementioned estimate (1) of protein folding rates, obtained in 1997, was confirmed by the subsequently obtained experimental data [176,194]; see Figure 3b.

However, one can see that the derived theory of protein folding rates explains Levinthal’s paradox “in *non*-Levinthal’s terms”, i.e., it deals with phase separation and free energy barriers, but gives no estimate as to the number of structures to be iterated over in a search for the most stable chain fold, and offers no explanation as to why such an iteration is feasible, at least for small globular proteins (or domains) of ~ 100 amino acid residues.

Our answer is that the Levinthal’s paradox assumed that the search should be done among *all* conformations of the protein chain (which is indeed impossible), while the search among low-energy folds *only* (i.e., only among compact and well-structured globules), which is done at the level of protein secondary structure assembly (Figure 4), is by many orders of magnitude less voluminous, and is therefore, feasible. A rough estimate [181,199,200] leads to the conclusion that at the level of secondary structure assemblies (or, in other words, at the level of potential molten globules), the search volume does not exceed

$$\sim L^N \sim \exp[\frac{1}{4}\ln(L) L^{2/3}] \quad (2)$$

for a protein chain of L amino acid residues and N secondary structure elements, which, in the main term, scales approximately as the exponent in the aforementioned upper estimate ($\tau \times \exp[(0.5 + \frac{5}{12}\ln(3L^{1/3})) L^{2/3}] \approx \tau \times \exp[1.5L^{2/3}]$) of the protein folding time.

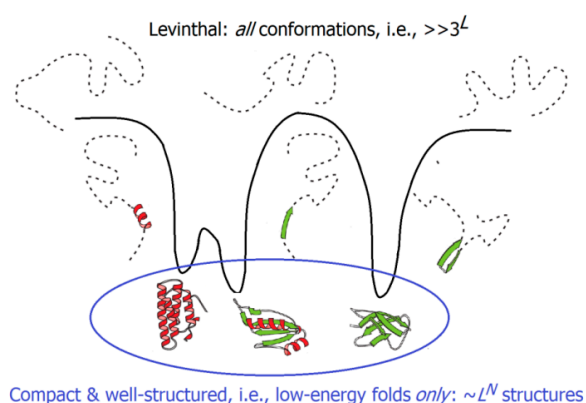


Figure 4. Comparison of a huge search among *all*, for the most part disordered conformations, and a much less voluminous search among *only* compact and well-structured globules, which correspond to the deep energy minima. L is the number of amino acid residues in the protein chain, and N is the number of secondary structure elements in this chain, which can be estimated as $N \approx L/[(3 \div 4.5)L^{1/3}] \approx L^{2/3}/4$ [181].

3. Intramolecular Phase Transitions in Disordered Proteins Induced by Interactions with Binding Partners

The interplay between the amino acid sequence of a protein and environment defines the ability of a polypeptide chain to fold, misfold, or be intrinsically disordered. Although one can induce different degrees of disorder in a molecule of a globular protein by changes in its environment (to generate molten globule, premolten globule, and coil-like states [11–14,16,17,201]), IDPs/IDPRs can be differently disordered under the same physiological conditions and exist as highly-dynamic, conformational ensembles of collapsed (native molten globules) or extended disordered species (native premolten globules and native coils) [17,33,202]. In other words, contrarily to globular proteins that have unique 3D structures under physiological conditions, IDPs/IDPRs exist under the same conditions as dynamic, conformational ensembles with quite different structures that interconvert on a number of timescales. Accumulated data on the structural heterogeneity of IDPs suggest that the representation of these proteins as members of three well-defined structural classes (native molten globules, native premolten globules, and native coils) is an oversimplification. In fact, IDPs/IDPRs might contain foldons (i.e., independently foldable protein units, which should not be mixed with domains, since single-domain proteins might have several foldons [203–207], as was shown for cytochrome *c* [208], apo-cytochrome *b*₅₆₂, ribonuclease H, dimeric triosephosphate isomerase, the OspA protein of *Borrelia* [203], and staphylococcal nuclease [209]), inducible foldons (which are IDPRs capable of at least partial folding, promoted by their interactions with binding partners), morphing inducible foldons (IDPRs with the potential to fold differently due to binding to different partners), semi-foldons (regions that are always in a semi-folded form), and nonfoldons (IDPRs that never fold). On the other hand, the functionality of many ordered proteins depends on the presence of ‘unfoldons’, i.e., regions of ordered proteins that undergo order-to-disorder transitions to make proteins active [210]. Therefore, based on currently available data, one can conclude that intrinsic disorder can have multiple faces, and can affect different levels of protein structural organization, where either whole protein or various regions are disordered to a different degree. Based on these considerations, it has been proposed that functional proteins represent a continuous spectrum of differently-structured/disordered conformations that ranges from fully ordered to completely structureless species and everything in between [210]. It was also pointed out that no boundary is present between the ordered proteins and IDPs. Instead, the structure–disorder space of a protein represents a continuum [210] that defines the protein structure–function continuum [211–214], where instead of the classical “one gene—one protein—one structure—one function” model, any protein represents a dynamic conformational ensemble containing multiple conformational/basic, inducible/modified, and functioning proteoforms. Proteoforms represent a set of distinct protein molecules encoded by a single gene. They originate from allelic variations and various pretranslational mechanisms affecting genes, such as the production of multiple mRNA variants by alternative splicing and mRNA editing. They can also be generated by numerous changes induced in the chemical structures of proteins by various post-translational modifications (PTMs) [215]. Also, some of them can be linked to the presence of IDPRs, or can originate from the functionality [211]. Therefore, the protein structure–function continuum suggests that any protein can be characterized by a broad spectrum of structural features, and can possess various functional potentials [211,212,216]. As a result, IDPs/IDPRs are not homogeneous, but represent a very complex mixture of potentially foldable, partially foldable, differently foldable, or not foldable segments [210,217]. In other words, IDPs/IDPRs behave as highly frustrated systems with no single folded state. This is reflected in their free energy landscapes, which are relatively flat and simple, and do not have a deep energy minimum seen in the free energy landscape of the ordered globular protein, representing instead a kind of ‘hilly plateau’, where hills correspond to forbidden conformations [16,218,219]. Such a simplified and flattened energy landscape is extremely sensitive to different environmental changes that can modify the landscape in a number of different ways, making some energy minima deeper and some energy barriers higher. This explains the conformational plasticity of IDPs/IDPRs, their extreme sensitivity to changes in the environment, and their ability to

specifically interact with many partners of different natures and to fold differently as a result of these interactions [210].

The lack of rigid structures in the IDPs/IDPRs is encoded in the specific features of their amino acid sequences, such as, for proteins/regions with extended disorder, the presence of numerous uncompensated charged groups (often negative) giving rise to their high net charges at neutral pH and extreme pI values [220–222], and a low content of hydrophobic residues [220,221]. On a more global level, amino acid sequences of IDPs/IDPRs have several common features [223,224], such as depletion in the order-promoting residues that would normally form the hydrophobic core of a folded globular protein (e.g., bulky hydrophobic (Ile, Leu, and Val) and aromatic (Trp, Tyr, and Phe) amino acids) and Cys residues. On the other hand, IDPs/IDPRs are noticeably enriched in disorder-promoting amino acids, such as polar residues Arg, Gln, Ser, Glu, and Lys, as well as Gly, Ala, and a hydrophobic structure-breaker Pro [33,225–228]. However, one should keep in mind that although being generally depleted in hydrophobic residues, IDPs/IDPRs still contain some strategically-placed hydrophobic residues, which could be of crucial functional importance. In fact, similar to ordered proteins containing characteristic patterns of hydrophobic and hydrophilic residues which are important for protein folding and function, amino acid sequences of IDPs/IDPRs are also patterned, and these proteins are known to contain so-called molecular recognition features, i.e., regions which are disordered in the unbound form but which can at least partially fold upon interacting with specific partners [55,56,67]. Importantly, since the degree of such binding-induced folding is different for different proteins, the resulting complexes are characterized by broad structural and functional heterogeneity [68,69]. Due to their ability to fill the gaps and cracks between the structural elements of a binding partner [229], IDPs/IDPRs can act as molecular glue or mortar [56]. Since interaction with partners can initiate at least partial conjoint binding-induced folding, IDPs/IDPRs can also serve as molecular epoxy [230–232]. The dynamic ‘on–off’ switch-type interactions commonly found in signaling networks are dependent on intrinsic disorder, since the ability to bind partners with high specificity and low affinity represents one of the specific features of disorder-based interactions [34,233,234]. Many IDPs/IDPRs serve as morphing shape-changers that are able to differently fold as a result of binding to different partners [17,67,87,235–237], with the binding regions of such morphing IDPs/IDPRs being able to adopt completely different structures upon binding to the divergent partners [58,65,238–240].

Under physiological conditions, the capability of a globular protein to gain ordered structure is encoded in its amino acid sequence that contains, so to say, a “blueprint” of a final structure. This “blueprint” can be complete; then, the proteins are foldable, and they fold spontaneously without help from external factors [241–243]. The facts that IDPs/IDPRs cannot spontaneously fold into unique 3D structures and that interactions with specific partners can resolve their foldability problem indicate that some parts of their “blueprints” are missing, and that these missed parts are provided by the binding partners. Although not quite literally, this binding-induced folding of IDPs/IDPRs can be approximated by the interaction-promoted folding and assembly of a globular protein from its polypeptide fragments, as was shown for Trp repressor [244], SH2 domain [245], maltose binding protein [246], oxyanion-translocating ATPase [247], barnase [248], rhodopsin [249,250], B1 domain of streptococcal protein G [251], pig heart CoA transferase [252], *E. coli* thioredoxin [253], bacteriorhodopsin [254], G protein-coupled receptors [255], ubiquitin [256], and *E. coli* aspartate transcarbamoylase (ATCase) [257], to name a few. Curiously, this high efficiency of the functional structure restoration from the peptide fragments prompted Johnsson and Varshavsky to design a ubiquitin split protein sensor (USPS) in order to detect protein–protein interactions in vivo [258]. Here, N- and C-terminal domains of ubiquitin are fused into two proteins, the interactions of which trigger the folding of rationally designed fragments to a functional ubiquitin. This approach was further enhanced by the development of various split reporter proteins which are commonly utilized nowadays in studies of protein–protein interactions, protein localization, intracellular protein dynamics, and protein activity in living cells and animals [259]. Among the split reporter systems used in protein-fragment complementation assays are constructs based on split dihydrofolate

reductase (DHFR) [260], β -galactosidase [261], green fluorescent protein (GFP) [262], firefly and renilla luciferase [263], and β -lactamase [264]. We give these examples here to illustrate the idea of the “blueprint” complementation, where a functional protein with unique structure is produced from inactive fragments as a result of conjoint folding–binding events.

4. Nucleation in Bulk, on the Interface and on the Impurities

In the first order phase transitions like melting or crystallization, as well as in their microscopic analogs, intramolecular “all-or-none” transitions, a key role is played by the nucleation of the new phase [265,266]. Nucleation can be 3-dimensional (that is, in bulk) or 2-dimensional (on the surface or interface).

The free energy of an emerging piece of a new phase consisting of $n > 1$ particles can be estimated as

$$G_3(n) \approx n\Delta\mu + n^{\frac{2}{3}}B_3 \quad (3)$$

for the 3-dimensional case, and

$$G_2(n) \approx n\Delta\mu + n^{\frac{1}{2}}B_2 \quad (4)$$

for the 2-dimensional case [266–268] (Figure 5). Here, $\Delta\mu < 0$ is the chemical potential decrease for the molecule of the “new” phase as compared to the “old” one; $B_3 > 0$ and $B_2 > 0$ stand for the additional free energy of a molecule at the 3- and 2-dimensional phase interfaces.

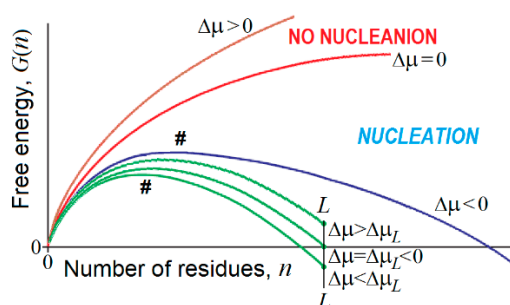


Figure 5. Free energy of a growing piece of a “new” phase at different $\Delta\mu$ values. The symbol # shows the transition state of the process. Three short green lines correspond to “all-or-none” transitions within a protein-like body formed by an L -residue chain (which can be in two phase states) at the equilibrium point ($\Delta\mu = \Delta\mu_L$, see the text), as well as and somewhat above and below of this point.

The free energy of the nucleus (the unavoidable highest-free-energy structure at the pathway of growth of a piece of the emerging new phase) is obtained from equation $dG(n)/dn = 0$, and, in the 3-dimensional case, is

$$G_3^\# = \frac{4B_3}{27} \left[\frac{B_3}{-\Delta\mu} \right]^2. \quad (5)$$

$G_3^\#$ is achieved at

$$n_3^\# = \frac{8}{27} \left[\frac{B_3}{-\Delta\mu} \right]^3, \quad (6)$$

while the “seed”, i.e., the smallest stable piece of the emerging new phase (satisfying the equation $G_3(n) = 0$ at $n > 1$), includes

$$n_3^0 = \left[\frac{B_3}{-\Delta\mu} \right]^3 = \frac{27}{8} n_3^\# \quad (7)$$

particles. Similar relationships can be obtained for the 2-dimensional case.

A few interesting consequences follow from the above (cf. Figure 5) relationships:

- (1) At $-\Delta\mu \rightarrow 0$, $G_3^\# \rightarrow +\infty$ and $G_2^\# \rightarrow +\infty$, which, according to conventional transition state theory, means that the time of the first order phase transition (exponentially dependent on the $G^\#$ value) is infinitely high near the point of thermodynamic equilibrium of the “new” and “old” macroscopic phases. This is a kinetic origin of hysteresis, overcooled liquids, etc., and, by the way, of the enormous time required for the formation of β -sheets in long polypeptides [144,145]. It is worth mentioning that extremely slow nucleation leads to the formation of single and extremely large compact pieces of the erasing phase.
- (2) There is a kind of competition between in bulk and on-surface nucleation of the new phase. At $-\Delta\mu \rightarrow 0$, $G_3^\#$ turns to infinity in proportion to $\left[\frac{B_3}{-\Delta\mu}\right]^2$, while $G_2^\#$ turns to infinity in proportion to only $\left[\frac{B_2}{-\Delta\mu}\right]$, i.e., much more slowly. This means that close to the conditions of phase equilibrium, 3-dimensional (“in bulk”) nucleation becomes kinetically impossible due to the very large $\left[\frac{B_3}{-\Delta\mu}\right]^2$ value, while the 2-dimensional (“on surface”) nucleation can still avoid kinetic problems, and occurs until $\left[\frac{B_2}{-\Delta\mu}\right]$ also becomes too large.
- (3) In contrast, when the phases are far from the equilibrium, that is $-\Delta\mu$ increases and starts to approach B_3 and B_2 (or B_3 and B_2 are small and approach $-\Delta\mu$), the nucleation in bulk should become fast and overcome the on-surface nucleation, because the surface layer is several orders of magnitude smaller than the bulk. It is worth mentioning that fast nucleation leads to the formation of many pieces of the erasing phase that can glue together, forming noncompact, amorphous, or branched aggregates.
- (4) If an all-or-none transition occurs in a microscopic body that includes L particles only, the “new” phase can be stable only if the seed of the arising phase is smaller than L , i.e., $n_3^0 = \left[\frac{B_3}{-\Delta\mu}\right]^3 \leq L$. This means that the new phase can arise only when its stability exceeds some threshold, i.e., $-\Delta\mu \geq -\Delta\mu_L \equiv \frac{B_3}{L^{1/3}}$. At the mid-transition point, where both phases have equal stability, and thus $-\Delta\mu = \frac{B_3}{L^{1/3}}$, the transition state free energy is $G_3^\# = \frac{4}{27}B_3L^{2/3}$, and it includes $n_3^\# = \frac{8}{27}L$ particles. This means that the time of transition to the new phase (which is as stable as that of old one) scales with L in the way given by Equation (1) for formation of the “native phase” of a protein (a microscopic body!), and that the folding nucleus of the new phase includes nearly $1/3$ of the body, i.e., it is not small.
- (5) If the new phase is a little more stable than the old one, that is $-\Delta\mu = -\Delta\mu_0(1 + \delta)$, where $0 < \delta \ll 1$, the free energy of the completely formed new phase is $\Delta G_3(L, \delta) = -\delta \cdot B_3L^{2/3} < 0$, and the transition state free energy $G_3^\#$ of nucleation of this stable phase by $\approx \frac{8}{27}\Delta G_3(L, \delta)$ is lower than the transition state free energy at the mid-transition point. Such an estimate has been used in [176] to describe the decrease in the protein folding time with the increase in protein stability.
- (6) If the new phase is formed around some local “impurity” and interacts with it with the free energy $G_0 < 0$, the free energy of the emerging phase obtains the form $G_3(n) \approx G_0 + n\Delta\mu + n^{2/3}B_3$ (and $G_2(n) \approx G_0 + n\Delta\mu + n^{1/2}B_2$), instead of that given by Equations (3) and (4). This correspondingly (by $G_0 < 0$) decreases the nucleation free energy $G_3^\#$ (as well as $G_2^\#$) of the new phase as compared to that given by Equation (5), and does not change the size $n_3^\#$ (as well as $n_2^\#$) of the critical nucleus given by Equation (6), but decreases the size n_3^0 (as well as n_2^0) of the “seed” relatively to that given by Equation (7).

5. Protein Crystallization, Amorphous Aggregation, and Fibrillation as Intermolecular Phase Transitions

In addition to the spontaneous and binding-induced intramolecular phase transitions described in the previous sections, proteins are able to undergo macrosystemic assembly processes of amorphous aggregation, fibrillation, gelation, crystallization, and liquid–liquid phase separation (see Figure 6). In these cases, insoluble protein ensembles with different degrees of packing orders are formed in

protein solutions as a result of proteins undergoing “soluble–insoluble” changes accompanied by intermolecular phase transitions.

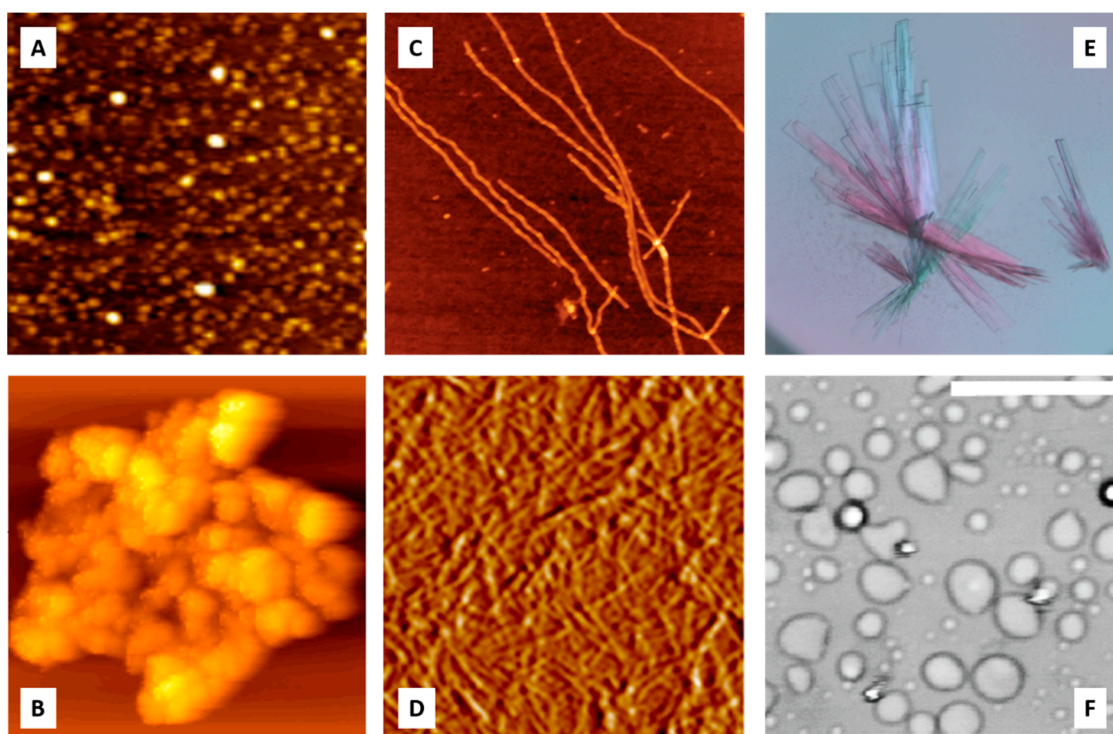


Figure 6. Illustrative examples of phases formed in protein solutions as a result of intermolecular phase transitions: (A). Oligomeric species; (B). Amorphous aggregate; (C). Amyloid fibrils; (D). Protein gel (atomic force microscopy (AFM) image of lysozyme fiber gel); (E). Protein crystal (crystals of slingshot phosphatase 2 in a hanging drop at 20 °C); (F). Liquid droplets (liquid droplets formed by the C-terminal domain of TDP-43 in the presence of RNA as a result of LLPT. Here, 20 μ M TDP-43CTD in presence of 40 μ g/mL RNA forms liquid droplets that coalesce over time. Scale bar represents 20 μ m). Images in plots (A–C) showing different aggregated forms of α -synuclein are AFM images from a personal collection of V.N.U. Image in plot (D) is modified with permission from [269]. Image in plot (E) is a courtesy of Dr. Eric M. Lewandowski and Prof. Yu Chen, University of South Florida. Image shown in plot (F) is a courtesy of Mr. Anukool A. Bhopatkar and Prof. Vijayaraghavan Rangachari, University of Southern Mississippi.

These processes are also different in terms of the degree of structural distortions induced by the environment in a protein molecule that triggers the corresponding transitions. They range from minimal structural changes in crystallization to moderate structural perturbations (denaturation) in amorphous aggregates, and to the large-scale conformational alterations that are typically required for fibrillation.

5.1. Protein Crystallization as a Peculiar Case of Phase Separation of Supersaturated Protein Solutions

An interesting peculiarity of the polypeptide chain of any well-structured globular protein is that its amino acid sequence guarantees the existence of the free energy barrier between the native and denatured (unfolded or partially folded) states [1,102]. This is of great importance for proper protein functioning, as the presence of such a barrier assures the structural identity of native proteins. The ability of native globular proteins to form crystals (known from Hoppe-Zeiler’s works of the 1860s, in which the author describes the method by which crystals of hemoglobin were obtained [270]) is one of the major pieces of evidence supporting this hypothesis. Protein crystallization is a consequence of protein association governed by the details of the protein structure and the peculiarities of its

environment. It represents a phase transition leading to the separation of a solid phase (protein crystal) from a supersaturated protein solution. However, protein crystals are not “dry” (they, in fact, have high solvent contents, i.e., ranging from 27% to 65%, with an average of 43% [271,272]).

A supersaturated protein solution is a metastable system that, by subtle changes in the environment, can be triggered to also undergo liquid–liquid phase separation, gelation, crystallization, or aggregation [273,274]. An important note here is that the aforementioned phenomena of liquid–liquid phase separation, gelation, crystallization, or aggregation occur in the supersaturated solutions of normally-folded proteins with (almost) unperturbed 3D structures. However, since gelation and aggregation commonly come about during protein crystallization, the resulting gels and aggregates are typically considered as disordered phases [274].

Crystallization takes place in supersaturated protein solutions within a rather narrow set of conditions known as a crystallization slot, where the protein solution is characterized by specific molecular interactions involving both solvent and solute molecules, and where protein self-interactions are defined by a specific range of the osmotic second virial coefficient, B_{22} [275–277]. Liquid–liquid phase transition is believed to be driven by the short-range nature of protein interactions [274]. This process can serve as an illustrative example of a spinodal demixing, which represents a transition from one totally unstable thermodynamic phase (supersaturated protein solution) to two coexisting stable or metastable phases (in this case, a liquid relatively depleted in protein and a liquid rich in protein), both containing significant levels of solvent. Furthermore, because of the presence of the high protein concentration phase (when it is metastable in relation to crystallization), crystallization occurs much more rapidly in the concentrated phase-separated state than in the initial solution [273].

It has been pointed out that the aggregation and crystallization in the supersaturated solutions of globular protein can be associated with the universal behavior of the concentration fluctuations taking place in the vicinity of the “spinodal line” in the (T, c) plane (where T is the temperature of the system and c is the protein concentration) [278,279]. This spinodal line represents a boundary showing the limits of the thermodynamic stability of a homogeneous fluid, or, in other words, the line separating a region of instability of the solution against spontaneous, nonnucleated demixing [280], where (unlike nucleation) the interface free energy plays no role. In the proximity of the spinodal line, the critical divergence of amplitudes and the lifetimes of the spontaneous fluctuations of solute concentrations within the stability region occur, and can be viewed as transient demixing [280]. These divergences follow the universal scaling law describing the temperature dependence of parameter ε , that “measures the normalized distance of the representative point of the solution from the spinodal line in a form $\varepsilon = (T - T_s)/T_s$, where T is the actual temperature of the system and T_s is the spinodal temperature” [278–280]. It has also been pointed out that light-scattering experiments in slow temperature scans can provide important information on the existence of anomalous fluctuations, temperature range, where their divergences follow the universal law, and spinodal temperature values [281]. Importantly, since the parameter ε reflects the global effects of all system parameters (such as additives, buffer, concentrations, pH, salts, temperature, etc.) on the solution stability, very different combinations of system parameters can give the same ε value.

A careful analysis of the crystal nucleation of several proteins under a variety of conditions reveals that for all the systems studied, dependencies of the nucleation rates on the ε parameter follow the same universal curve (or “master curve”) that covers a span of several orders of magnitude of nucleation rates [278,279]. The existence of such a master curve represents a direct consequence of the existence of the related critical concentration fluctuations and the universal divergence properties of such fluctuations. It has also been pointed out that the shape of this master curve suggests the existence of a region of thermodynamic instability of the solution against spontaneous demixing, suggesting the applicability of a simple, two-stage model of crystal nucleation. Here, at the first stage, concentration fluctuations generate abundant and relatively long-lived liquid regions where proteins cluster, and then, at the second stage, the rearrangement of clustered proteins into a crystalline form takes place within a characteristic time [278,279]. Importantly, the boundaries of the region with the

universal features of critical fluctuations and crystal nucleation are defined by the presence of minor structural changes in a protein molecule, with such a link between conformation details and universal behavior being triggered by solvent-mediated interactions, and with abrupt changes in the values of solution spinodal temperatures T_S being ascribed to a stepwise change in protein hydration and the related solution thermodynamics [279]. It has also been hypothesized that the correlation between the solution instability and structural changes in a protein undergoing crystallization can be explained by the inability of solvent-induced interactions to distinguish between residues belonging to the same or different proteins if they are located at comparable distances. As a result, at high enough protein concentrations, abrupt changes in the interprotein interactions that affect the solution stability can occur concurrently with the intraprotein interactions that contribute to the conformational stability and structural integrity of individual protein molecules [279].

5.2. Protein Amorphous Aggregation

In contrast to the processes described in the previous section that happen in supersaturated solutions of ordered proteins (such as crystallization and companion to it liquid–liquid phase separation, gelation, and aggregation), the process of amorphous aggregate formation represents one of the consequences of protein misfolding. Importantly, this form of misfolded protein aggregation is rather common, and can take place under particular conditions, even for amyloidogenic proteins when they fail to form amyloid fibrils. Similar to fibrillation (see next section), amorphous aggregate formation is driven by intermolecular interactions, and is linked to protein denaturation, since partially-folded conformations with exposed hydrophobic surfaces exhibit a greater propensity to aggregate [282]. In fact, multiple studies have clearly indicated that amorphous and fibrillar aggregation arises from the intermolecular association of partially-folded intermediates [282–290]. This requirement of the partial unfolding of a protein molecule that precedes intermolecular interactions links amorphous aggregates and amyloid fibril. Furthermore, being large, insoluble intermolecular ensembles, these misfolded aggregates are characterized by high levels of structural stability.

However, these two forms of misfolded aggregates are quite different on many levels. For example, fibrillar aggregates are highly ordered, β -sheeted ensembles, whereas in amorphous aggregation, proteins aggregate/oligomerize without forming specific high-order structures. Furthermore, besides the obvious morphological differences (fibrillar versus amorphous), amyloid fibrils and amorphous aggregates differ from each other by the mechanisms of their formation [291]. A typical fibrillation process is described by a sigmoidal curve characterized by the presence of a lag period that reflects the existence of a high free energy barrier associated with the nucleation of ordered structures of amyloid fibrils; this process can be accelerated by seeding (i.e., by the addition of the fragments of the preformed fibrils). The lag grows and can become huge in the presence of secondary nucleation, that is, the nucleation of branching or fragmentation of fibrils in addition to the primary nucleation of linear protofibrils [292,293]. In other words, the fibrils are formed via a nucleation and growth mechanism, where the overall reaction is rate-limited by the existence of a high free energy barrier associated with the nucleation. On the other hand, the formation of amorphous aggregates represents an instantaneous and spontaneous process that does not have a lag period and a high free energy barrier, and is not accelerated by seeding [291]. Since protein crystals and amyloid fibrils are formed by a nucleation and growth mechanism [294–296], and since crystallization and fibrillation can be accelerated by seeding [297,298], amyloid fibrillation is considered to be similar to crystallization [291,294–296]. On the other hand, amorphous aggregation was proposed to be analogous to the glass transition [291], with the glassy behavior of amorphous aggregates being reflected in the presence of heterogeneous conformations that are fixed by strong attractive forces producing various sites of interaction [291,299].

Importantly, not all amorphous aggregates are formed from denatured protein species. In fact, amorphous aggregation can also happen in the solutions of near-natively-folded proteins [300]. Examples of this phenomenon are given by the aggregation of the cataract-related P23T mutant of γ D-crystallin [301] and bovine pancreatic trypsin inhibitor (BPTI) variant, BPTI-22, containing

22 alanines [302]. In the first case, an analysis of γ D-crystallin aggregated under physiological conditions by solid-state NMR revealed the presence of a well-ordered, native-like conformation [301]. In the second case, the presence of a native-like structure in aggregated BPTI-22 was evidenced by ^{15}N hetero single quantum correlation (HSQC) NMR spectra [302]. Since in many other cases, amorphous aggregates were formed from partially-folded intermediates, it is unclear now how widespread such near-natively-folded amorphous aggregates are.

Recently, it was pointed out that although amorphous aggregation and fibrillation—which can happen for a given protein at similar solution conditions—are often considered as competing processes, or amorphous aggregation is treated as an obligatory intermediate process within the amyloid formation pathway, these two models can be integrated into a single paradigm [303,304]. Here, amorphous aggregation is treated as a liquid–liquid phase transition leading to the formation of the amorphous aggregate that represents a second liquid phase whose liquid-like properties are determined by the intra-phase monomer mobility, and where fibrillation takes place at the interfacial boundary via the heterogeneous growth pathways including the nucleation, growth, and fragmentation of amyloids [292,293,303,304].

5.3. Protein Fibrillation

The interest of researchers in protein misfolding and fibrillation is based upon the involvement of these processes into the pathogenesis of various protein deposition diseases or proteinopathies, such as amyloidoses, and different neurodegenerative disorders. Globally, the molecular mechanisms underlining these different pathological states are the same, where first a transition of specific proteins or protein fragments from a native soluble form into insoluble aggregate/fibrils takes place, with the subsequent accumulation of aggregated material within a variety of organs and tissues [305–312]. Different proteinopathies are caused by unrelated proteins, which, prior to fibrillation, may be rich in α -helices, β -structure, or have an α/β or $\alpha + \beta$ structure. Some of these amyloidogenic proteins are globular proteins with unique 3D structures, whereas others are IDPs with different levels and degrees of disorder. Despite these differences, the fibrils found in different pathologies have many common features, such as similar morphologies, being twisted, rope-like structures, reflecting a filamentous substructure, and the presence of a core cross- β -sheet structure with continuous β -sheets formed by β -strands running perpendicular to the long axis of the fibrils [313]. Importantly, not all amyloids are pathological, and living organisms quite often exploit the inherent ability of proteins to fibrillate in order to generate functional amyloids with novel and diverse biological functionalities [314].

Since amyloid fibrils can be formed *in vitro* from both disease-associated and disease-unrelated proteins and peptides, it is now believed that the ability to fibrillate represents a generic property of a polypeptide chain, with all proteins being potentially able to form amyloid fibrils under the appropriate conditions [307,315–319]. Therefore, contrarily to the globular proteins that can spontaneously fold into unique 3D structures, which are critically dependent on the amino acid composition and sequence of the polypeptide chain [6,7], amyloid fibrils represent a generic phase of any peptide chain stabilized mostly via the main chain-main chain interactions, being therefore rather insensitive to the information encoded in the side chains.

These observations seem to suggest the absence of a specific “aggregation code” defining the formation of amyloid fibrils. However, not all proteins are equally prone to fibrillation; some are prone to aggregate, and can do so (if their concentration is sufficient) even under the physiological or near-physiological conditions, especially if the monomeric protein is deprived of its natural interacting partners, whereas others require rather extreme environmental perturbations (like seeding by fragments of preformed fibrils) to initiate the fibrillation process [307,311,316,320–322]. Furthermore, aggregation can be disallowed by some negative-design features present in the folded state of a protein [323]. Therefore, a selective pressure optimizes the primary sequence to allow a protein to fold into a stable soluble structure. This optimization is needed to prevent the functionally-competent fold from converting to the amyloid phase. Taken together, these findings suggest that not all soluble proteins

have structures that are optimized to the same degree in order to avoid fibrillation, and that the susceptibility of a protein to conversion to the aggregated form is defined by the dependence of its structure on binding partnerships or complexation.

It is important to emphasize here that the formation of amyloid-like fibrils does not embody the only pathological feature of proteinopathies, and that pathological protein deposits can also be in the form of amorphous aggregates, which are cloud-like inclusions without defined morphologies. For many proteins, the aggregation process originating from protein misfolding can generate another alternative final product, i.e., soluble oligomers. It seems that the precursor of soluble aggregates is the most structured, whereas amyloid fibrils are formed from the least-ordered conformation via binding-induced folding. The choice between three pathogenic aggregation pathways, i.e., fibrillation, amorphous aggregation, and oligomerization, is determined by the peculiarities of the amino acid sequence, the protein concentration, and the environment.

5.3.1. Conformational Prerequisites for Amyloidogenesis

The fibrillation of the majority of ordered proteins, which are not able to easily form amyloid fibrils under physiological conditions, requires denaturing conditions [307,311,316,320–322], suggesting that these proteins can fibrillate when their rigid native structure is destabilized and a partially-unfolded conformation is formed [282,306–312,316,320,321,324–327]. Obviously, such a requirement for partial unfolding is not applicable to IDPs with extended disorder, since they do not have stable and well-folded 3D structures in their native states. Instead, a primary step of the fibrillation of such proteins involves partial folding, leading to the stabilization of a partially-folded conformation [317–319,328–330]. Therefore, a general hypothesis of fibrillogenesis which is applicable to ordered proteins and IDPs suggests that protein fibrillation is critically dependent on the structural transformation of a native protein (ordered or intrinsically disordered) into a partially folded, aggregation-prone conformation, enabling the assembly of misfolded aggregates via specific intermolecular interactions of different physico-chemical nature, such as electrostatic attraction, hydrogen bonding, and hydrophobic interactions. Therefore, amyloid fibril formation is promoted when relatively unfolded protein species are formed under conditions whereby noncovalent interactions are still favorable.

5.3.2. Fibrillogenesis of Globular Proteins Depends on Partial Unfolding

Significant evidence supports the idea that the fibrillogenesis of globular proteins requires their partial unfolding [282,306–312,316,320,321,324–327]. One should keep in mind that the unique 3D structures of ordered globular proteins under physiological conditions are not completely immobile, but have structural fluctuations of various degrees and timescales. Due to this conformational breathing, the structure of a globular protein represents a dynamic conformational ensemble containing tightly-folded species and multiple partially-unfolded conformations, with the former greatly predominating [331,332]. The native structures of globular proteins were shown to be destabilized by most mutations associated with the accelerated protein fibrillation and proteinopathies. As a result, these mutations caused the increase in the steady-state levels of partially-folded forms within the conformational ensemble of a mutated protein [306,307,312,316,320,324,326,333–337]. Conversely, the stabilization of a native protein structure via the specific binding of ligands or drugs can significantly reduce the amyloidogenicity of a protein [338–346]. Furthermore, the rate of fibril formation can be significantly accelerated by destabilizing the native structure of a globular protein by utilizing low or high pH, high temperatures, low to moderate concentrations of strong denaturants, organic solvents, etc.

5.3.3. Fibrillogenesis of Extended IDPs Is Driven by Partial Folding

As mentioned, the functionalities of many proteins require a high degree of structural disorder [17,19,20,33,34,54–56,58,59,88,210,234,347]. It seems that due to their lack of significant conformational constraints and substantial conformational mobility, IDPs ought to be better suited to amyloidogenesis than to tightly-packed globular proteins; however, this is not always the case, and many

extended IDPs are rather resilient against aggregation and fibrillation. This is because the fibrillation of such proteins requires their partial folding. Examples of such amyloidogenic extended IDPs undergoing partial folding during their amyloidogenesis are A β [328], tau protein [348], α -synuclein [317], β -synuclein [349], γ -synuclein [350], amylin [351], prothymosin α [352], and histones [319].

5.3.4. Premolten Globule as a Universal Amyloidogenic Intermediate

Based on the analysis of structural events at the early fibrillation stages, it has been concluded that the substantially unfolded conformations of proteins and polypeptides typically serve as fibril precursors [311]. Although any partially-folded species (including the molten globular and the premolten globular species) may potentially play a role of such a crucial fibrillation-prone intermediate, the accumulated evidence indicates that the amyloidogenic species is significantly unfolded, being structurally closer to the premolten globule than to the molten globule state [311]. It seems that among different partially-folded intermediates described for proteins, the most amyloidogenic species is the premolten globule state, which is a relatively swollen conformation lacking a globular structure and possessing a relatively low secondary structure content, i.e., that sums to ~50% or less of the corresponding native value [17].

5.3.5. Sequential Mechanism of Fibril Formation and Morphological Heterogeneity of Amyloid Fibrils

In conclusion, proteins with different types of structures are equally subjected to aggregation [311,354], which represents an extremely complex process consisting of at least three major steps. First, different soluble proteins are transformed into the “sticky” aggregation-prone precursor or intermediate with the premolten globule properties. Since such aggregation-prone intermediates would be structurally different for different proteins, and since even the same protein can be converted into the structurally-different, partially-folded species by different environmental conditions, the variations in the amount of the ordered structure retained in the amyloidogenic precursor is believed to be responsible for the formation of fibrils with distinct morphologies [355]. At the second step, which is usually considered a nucleation step, or the lag period that precedes the formation of the insoluble aggregates, different oligomeric species are formed [354]. The lag phase occurs because the association of monomers is initially unfavorable, however, once a critical nucleus has been generated, and the gain of enthalpy from incorporating additional monomers outweighs the increase in entropy from dissociation. As a result, aggregation becomes energetically favorable and the reaction enters a growth phase [354]. An idealized model of amyloid fibril formation and protein aggregation in general is presented in Figure 7 (see bottom pathway), illustrating the directionality and sequential nature of the aggregation process that includes a series of consecutive steps [353]. Importantly, not all oligomers formed during the protein fibrillation process are productive, i.e., not all of them will eventually “grow” into fibrils. In fact, some metastable oligomers can “crouch”, being able to compete with fibril formation by decreasing the concentration of the fibril-forming free monomers [293]. This scenario is illustrated by the presence of competing misfolded oligomers in the bottom panel of Figure 7. Depending on the peculiarities of the environment, different aggregated forms (oligomers, amyloid fibrils, amorphous aggregates) can be generated via the intermolecular self-assembly of the different partially-folded species of a given protein (see upper pathways in Figure 7). In this model, different oligomers are formed from the structurally-identical monomers. However, since aggregation can cause dramatic structural reorganization of the aggregating protein, monomers at different aggregation stages are not structurally identical [353]. Again, one can expect to find competing misfolded oligomers within all the pathways shown in Figure 7.

Furthermore, the typical aggregation process only rarely results in the formation of a homogeneous intermolecular ensemble, where only one type of aggregates species is present. More often, various aggregated forms appear, giving rise to heterogeneous mixtures of differently-aggregated species (see Figure 8). In addition, for each aggregated form, there is a multitude of different morphologies, with monomers constituting such morphologically-distinctive aggregated forms being potentially

structurally diverse (see Figure 7). All of this suggests that far from being a simple reaction, aggregation represents a very complicated process with multiple related and unrelated pathways, which can be connected or disjointed.

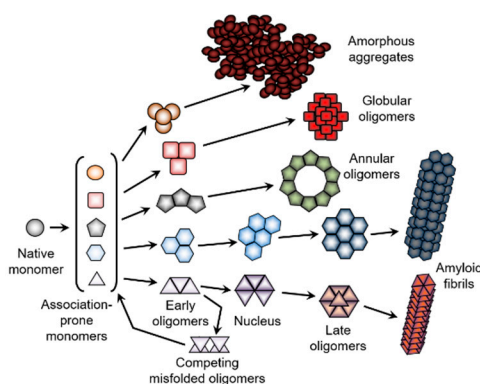


Figure 7. Schematic, oversimplified depiction of the process of protein self-association. Multiple aggregation pathways are generated via the formation of multiple association-prone monomeric forms. The aggregation reaction generates at least three major products, which are amorphous aggregates (top pathway), different soluble oligomers with diverse morphologies (second and third from the top pathways), and morphologically-divergent amyloid fibrils (two bottom pathways). Potential structural changes in the monomers that might happen at each elementary step are shown by color changes. The real situation is more complex, and more different species can be formed during and as a result of aggregation. Various species within the different pathways can interconvert. Modified from [353].

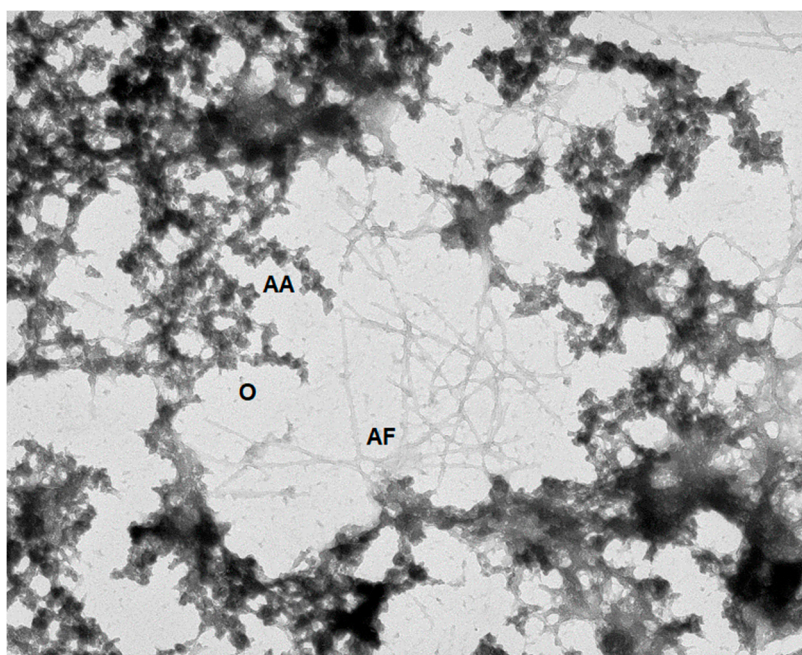


Figure 8. Transmission electron micrograph of a typical end-product of the protein aggregation process (in this case, fibrillation of α -synuclein), where amorphous aggregates (AA), amyloid-like fibrils (AF), and oligomers (O) are present. This image is from the personal collection of V.N.U.

6. Reincarnation of Liquid–Liquid and Liquid–Gel Phase Transitions: Drivers of the Biogenesis of Membraneless Organelles

Although the phenomenon of liquid–liquid phase separation (LLPS) in supersaturated protein solutions has been known in the field of protein crystallography for a long time, it was mostly unknown for the outside world. The situation has changed recently, and we are now witnessing a dramatic

increase in the level of interest in this intriguing phenomenon, not only from crystallographers, but also among researchers working in various fields of protein science, cellular biology, biotechnology, and biomedicine. This is because of the realization that LLPS can drive the cellular compartmentalization and biogenesis of various membraneless organelles (also known as the proteinaceous membraneless organelles (PMLOs), puncta, “biomolecular condensates”, foci, etc.). Curiously, although the existence of such membraneless compartments within the cells has been known to the scientific community for many years (e.g., nucleolus was described as early as in the 1830s [356,357]), the facts that PMLOs are numerous, and that they may have important biological functions, were generally overlooked, mostly due to the inability to isolate them for focused analyses which were in line with a scientific reductionistic approach, i.e., if the functionality of a complex system is the sum of the functions of its constituents, then to understand how such a complex system works, it needs to be taken apart, and individual parts need to be studied separately to understand their structures and functionalities. Although this linear scientific method was successfully utilized for the analysis of the functionality of “traditional” membrane-encapsulated organelles, it obviously failed for PMLOs (no membrane equals no luck with isolation). As a result, for a long time, this inability to be isolated, combined with their transient existence, have placed PMLOs in the category of potential artifacts.

It is recognized now that PMLOs, these highly dynamic protein-based assemblages [358], are often present in cytoplasm, nuclei, the mitochondria of various eukaryotic cells, in chloroplasts of plant cells, as well as in bacterial cells, where they play a number of important roles in the organization of various intracellular processes [358–364]. PMLOs are numerous and very diverse [359,361,365–372]; there are at least 40 different types found in eukaryotic and bacterial cells [373]. Figure 9 illustrates the multiplicity of PMLOs by showing a horde of such phase-separated liquid droplets that can be found in bacterial and eukaryotic cells. A detailed description of eukaryotic PMLOs with illustrative examples is beyond the scopes of this article, and is presented elsewhere [374].

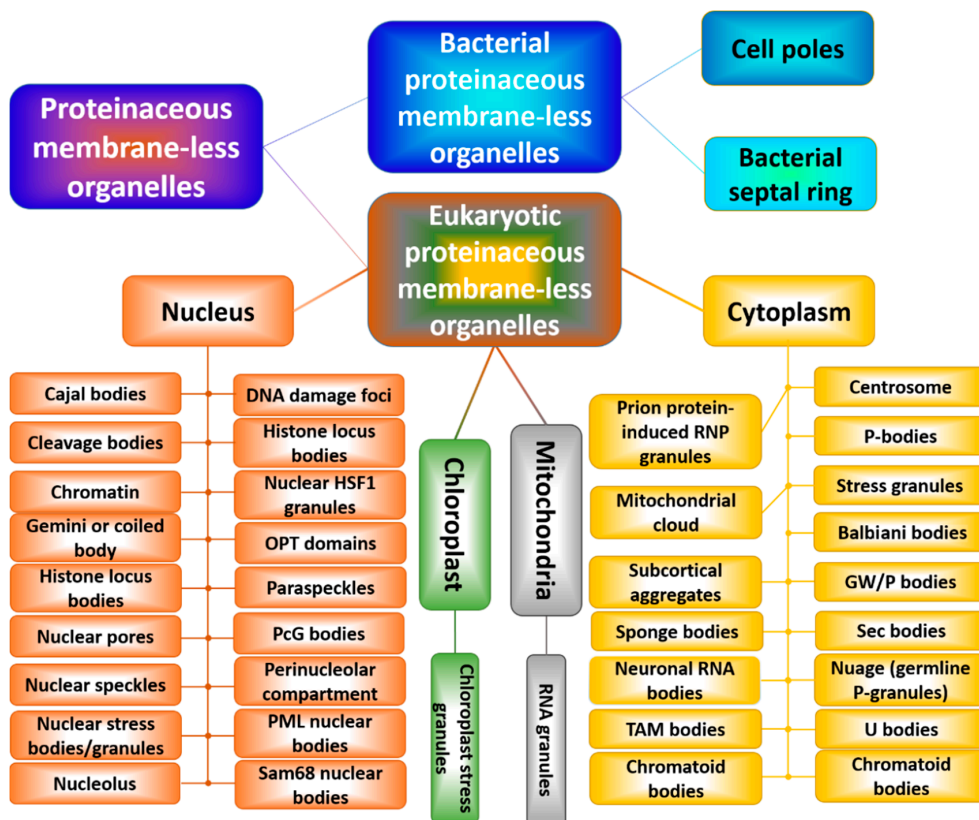


Figure 9. Illustration showing the variability of chloroplast, cytoplasmic, mitochondrial, and nuclear, proteinaceous membraneless organelles (PMLOs) that can be found in eukaryotes and bacterial PMLOs.

This figure was used with permission from Zaslavsky, B.Y., and Uversky, V.N. (2018). In *Aqua Veritas: The Indispensable yet Mostly Ignored Role of Water in Phase Separation and Membraneless Organelles*. *Biochemistry* 57(17), 2437–2451. Copyright (2018) American Chemical Society.

The formation of PMLOs represents a natural way of compartmentalizing various biological processes in different regions of the cell [360]. Since PMLOs are able to respond to, facilitate, regulate, and control different biological functions and stimuli [361], they are now considered to be important controllers of cellular life.

The liquid-like nature of PMLOs and phase-separated droplets can affect and modulate the functions of their components, which remain dynamic and flexible within these droplets, despite being amassed at high concentrations. In line with these considerations, it has been shown that the low-density structure of PMLOs found within the *Xenopus* oocyte nuclei determines the access from the nucleoplasm to the macromolecules within these PMLOs [375]. Due to their increased concentrations of nucleic acids and proteins, PMLOs can accelerate cytoplasmic reactions, thereby behaving as liquid-phase microreactors [376–378]. They can also represent a way recruiting and concentrating specific proteins, as, for example, observed in Negri bodies (NBs), where viral RNAs are synthesized [379]. Since some nuclear PMLOs concentrate specific sets of mRNAs and regulatory proteins, they also can serve as dynamic sensors of localized signals and, thereby, play a dual role in the translation of associated mRNAs, preventing mRNA translation at rest, and promoting local protein synthesis upon activation [380].

There is no doubt that PMLOs are full of mystery. Since they do not have membranes, their biogenesis and structural integrity rely exclusively on protein–protein and/or protein–nucleic acid interactions [381,382], and their components can directly contact and exchange with the exterior environment [383,384]. On the other hand, they have macroscopic dimensions and are detectable by under a microscope. The dimensions of these highly mobile but stable assemblages are dependent on the cell size [376]. PMLOs demonstrate the liquid-like behavior, being capable of dripping, the formation of spherical structures upon fusion, and wetting [385–388].

These intracellular liquid droplets are formed via biological, liquid–liquid phase transitions (LLPTs), also known as intracellular liquid–liquid demixing phase separation [376,389]. Such intracellular LLPTs are concentration-dependent, since PMLO formation is initiated by the colocalization of the participating molecules at high concentrations within a small cellular microdomain [383,384]. The biogenesis of PMLOs is a highly controllable and reversible process, with the formation of PMLOs being initiated by fluctuations in the concentrations of proteins undergoing LLPT, variations in the concentrations of definite small molecules or salts, osmolarity changes, alterations in the solution pH and/or temperature, by alternative splicing and various PTMs of the phase-forming proteins, via binding of these proteins to some specific partners, or by alterations of other environmental conditions affecting protein–protein or protein–nucleic acid interactions [376,389–392].

The fluidity of PMLOs originates from the multivalent interactions between IDPs or proteins containing IDPRs that are not accompanied by noticeable alterations in the structure of proteins undergoing LLPTs [372,393,394]. Therefore, PMLOs represent a special form of disorder-based protein complexes [372,389,393,395], and can be considered as illustrations of the disorder-based emergent behavior of proteins [210,213,396,397]. The lack of noticeable structural changes in IDPs forming PMLOs is supported by the NMR analysis of several PMLOs or liquid droplets, such as the Alzheimer-related protein tau [398,399], elastin-like polypeptides (ELPs) [400], low-complexity domain of the RNA-binding protein fused in sarcoma (FUS) [401], and the heterogeneous nuclear ribonucleoprotein A2 (hnRNPA2) [402], to name a few.

7. Conclusions

This review describes various intra- and inter- molecular phase transitions taking place in protein solutions, thereby representing protein existence as an exciting story of life in phases, where different phase transitions define the structure, function, interactability, aggregation, crystallization,

and compartmentalization of proteins. Although many of these phase transitions are linked to the general polymeric nature of proteins (e.g., intramolecular coil-globule transitions or intermolecular liquid-liquid and liquid-gel phase separation), other phase transitions seems to be rather specific for proteins, which are biological copolymers that were evolutionarily edited to have unique structures and/or functions. This edited polymer nature is related to the ability of globular proteins to undergo intramolecular, globule-globule transitions, giving rise to their unique 3D structures and their ability to be crystallized, as well as being associated with the ability of intrinsically-disordered proteins to undergo binding-induced intramolecular phase transitions originating from their interactions with specific partners. Also, although it seems that amorphous aggregation can take place in supersaturated solutions of various solutes of polymeric and nonpolymeric nature, ordered aggregations, i.e., the formation of amyloid fibrils that requires a dramatic structural rearrangement of protein monomers, might represent a special case of intermolecular phase transitions which is specific to polypeptides.

Author Contributions: Conceptualization, A.V.F. and V.N.U.; Literature collection and analysis, A.V.F. and V.N.U.; Visualization, A.V.F. and V.N.U.; Writing-Original Draft Preparation, V.N.U. and A.V.F.; Writing-Review & Editing, A.V.F. and V.N.U.

Funding: A.V.F. acknowledges support from the RAS Program 9 on Fundamental Research [grant.01201358029]. V.N.U. was supported by the National Institute On Aging of the National Institutes of Health under Award Number RF1AG055088.

Acknowledgments: We are thankful to Eric M. Lewandowski and Yu Chen (University of South Florida) and of Anukool A. Bhopatkar and Vijayaraghavan Rangachari (University of Southern Mississippi) for sharing images of protein crystals and liquid droplets, respectively.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Finkelstein, A.V.; Ptitsyn, O.B. *Protein Physics: A Course of Lectures*, 2nd ed.; Academic Press, An Imprint of Elsevier Science: Amsterdam, The Netherlands; Boston, MA, USA; Heidelberg, Germany; London, UK; New York, NY, USA; Oxford, UK; Paris, France; San Diego, CA, USA; San Francisco, CA, USA; Singapore; Sydney, Australia; Tokyo, Japan, 2016; p. 354.
2. Anson, M.L.; Mirsky, A.E. The effect of denaturation on the viscosity of protein systems. *J. Gen. Physiol.* **1932**, *15*, 341–350. [[CrossRef](#)]
3. Mirsky, A.E.; Pauling, L. On the structure of native, denatured, and coagulated proteins. *Proc. Natl. Acad. Sci. USA* **1936**, *22*, 439–447. [[CrossRef](#)] [[PubMed](#)]
4. Neurath, H.; Greenstein, J.P.; Putnam, F.W.; Erickson, J.O. The chemistry of protein denaturation. *Chem. Rev.* **1944**, *34*, 157–265. [[CrossRef](#)]
5. Tanford, C. Protein denaturation. *Adv. Protein Chem.* **1968**, *23*, 121–282. [[PubMed](#)]
6. Anfinsen, C.B.; Haber, E.; Sela, M.; White, F.H., Jr. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc. Natl. Acad. Sci. USA* **1961**, *47*, 1309–1314. [[CrossRef](#)]
7. Anfinsen, C.B. Principles that govern the folding of protein chains. *Science* **1973**, *181*, 223–230. [[CrossRef](#)]
8. Anfinsen, C.B.; Scheraga, H.A. Experimental and theoretical aspects of protein folding. *Adv. Protein Chem.* **1975**, *29*, 205–300. [[CrossRef](#)]
9. Ptitsyn, O.B.; Pain, R.H.; Semisotnov, G.V.; Zerovnik, E.; Razgulyaev, O.I. Evidence for a Molten Globule State as a General Intermediate in Protein Folding. *FEBS Lett.* **1990**, *262*, 20–24. [[CrossRef](#)]
10. Ptitsyn, O.B.; Uversky, V.N. The molten globule is a third thermodynamical state of protein molecules. *FEBS Lett.* **1994**, *341*, 15–18. [[CrossRef](#)]
11. Uversky, V.N.; Ptitsyn, O.B. “Partly folded” state, a new equilibrium state of protein molecules: Four-state guanidinium chloride-induced unfolding of beta-lactamase at low temperature. *Biochemistry* **1994**, *33*, 2782–2791. [[CrossRef](#)]
12. Ptitsyn, O.B.; Bychkova, V.E.; Uversky, V.N. Kinetic and equilibrium folding intermediates. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **1995**, *348*, 35–41. [[CrossRef](#)] [[PubMed](#)]
13. Ptitsyn, O.B. Molten globule and protein folding. *Adv. Protein Chem.* **1995**, *47*, 83–229. [[CrossRef](#)] [[PubMed](#)]

14. Uversky, V.N.; Ptitsyn, O.B. Further evidence on the equilibrium “pre-molten globule state”: Four-state guanidinium chloride-induced unfolding of carbonic anhydrase B at low temperature. *J. Mol. Biol.* **1996**, *255*, 215–228. [[CrossRef](#)] [[PubMed](#)]
15. Uversky, V.N. Diversity of equilibrium compact forms of denatured globular proteins. *Protein Pept. Lett.* **1997**, *4*, 355–367.
16. Turoverov, K.K.; Kuznetsova, I.M.; Uversky, V.N. The protein kingdom extended: Ordered and intrinsically disordered proteins, their folding, supramolecular complex formation, and aggregation. *Prog. Biophys. Mol. Biol.* **2010**, *102*, 73–84. [[CrossRef](#)]
17. Uversky, V.N. Protein folding revisited. A polypeptide chain at the folding-misfolding-nonfolding cross-roads: Which way to go? *Cell. Mol. Life Sci.* **2003**, *60*, 1852–1871. [[CrossRef](#)]
18. Tcherkasskaya, O.; Uversky, V.N. Polymeric aspects of protein folding: A brief overview. *Protein Pept. Lett.* **2003**, *10*, 239–245. [[CrossRef](#)]
19. Uversky, V.N. Natively unfolded proteins: A point where biology waits for physics. *Protein Sci.* **2002**, *11*, 739–756. [[CrossRef](#)]
20. Uversky, V.N. What does it mean to be natively unfolded? *Eur. J. Biochem.* **2002**, *269*, 2–12. [[CrossRef](#)]
21. Tcherkasskaya, O.; Uversky, V.N. Denatured collapsed states in protein folding: Example of apomyoglobin. *Proteins* **2001**, *44*, 244–254. [[CrossRef](#)]
22. Ptitsyn, O.B. Stages in the mechanism of self-organization of protein molecules. *Dokl. Akad. Nauk SSSR* **1973**, *210*, 1213–1215. [[PubMed](#)]
23. Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242. [[CrossRef](#)] [[PubMed](#)]
24. Berman, H.; Henrick, K.; Nakamura, H. Announcing the worldwide Protein Data Bank. *Nat. Struct. Biol.* **2003**, *10*, 980. [[CrossRef](#)] [[PubMed](#)]
25. DeForte, S.; Uversky, V.N. Resolving the ambiguity: Making sense of intrinsic disorder when PDB structures disagree. *Protein Sci.* **2016**, *25*, 676–688. [[CrossRef](#)] [[PubMed](#)]
26. Ringe, D.; Petsko, G.A. Study of protein dynamics by X-ray diffraction. *Methods Enzymol.* **1986**, *131*, 389–433. [[CrossRef](#)] [[PubMed](#)]
27. Radivojac, P.; Obradovic, Z.; Smith, D.K.; Zhu, G.; Vucetic, S.; Brown, C.J.; Lawson, J.D.; Dunker, A.K. Protein flexibility and intrinsic disorder. *Protein Sci.* **2004**, *13*, 71–80. [[CrossRef](#)]
28. Le Gall, T.; Romero, P.R.; Cortese, M.S.; Uversky, V.N.; Dunker, A.K. Intrinsic disorder in the Protein Data Bank. *J. Biomol. Struct. Dyn.* **2007**, *24*, 325–342. [[CrossRef](#)]
29. Carugo, O.; Djinic-Carugo, K. Criteria to Extract High-Quality Protein Data Bank Subsets for Structure Users. *Methods Mol. Biol.* **2016**, *1415*, 139–152. [[CrossRef](#)]
30. Dunker, A.K.; Obradovic, Z.; Romero, P.; Garner, E.C.; Brown, C.J. Intrinsic protein disorder in complete genomes. *Genome Inform. Ser. Workshop Genome Inform.* **2000**, *11*, 161–171.
31. Uversky, V.N. The mysterious unfoldome: Structureless, underappreciated, yet vital part of any given proteome. *J. Biomed. Biotechnol.* **2010**, *2010*, 568068. [[CrossRef](#)]
32. Ward, J.J.; Sodhi, J.S.; McGuffin, L.J.; Buxton, B.F.; Jones, D.T. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J. Mol. Biol.* **2004**, *337*, 635–645. [[CrossRef](#)] [[PubMed](#)]
33. Dunker, A.K.; Lawson, J.D.; Brown, C.J.; Williams, R.M.; Romero, P.; Oh, J.S.; Oldfield, C.J.; Campen, A.M.; Ratliff, C.M.; Hipps, K.W.; et al. Intrinsically disordered protein. *J. Mol. Graph. Model.* **2001**, *19*, 26–59. [[CrossRef](#)]
34. Uversky, V.N.; Dunker, A.K. Understanding protein non-folding. *Biochim. Biophys. Acta* **2010**, *1804*, 1231–1264. [[CrossRef](#)] [[PubMed](#)]
35. Xue, B.; Dunker, A.K.; Uversky, V.N. Orderly order in protein intrinsic disorder distribution: Disorder in 3500 proteomes from viruses and the three domains of life. *J. Biomol. Struct. Dyn.* **2012**, *30*, 137–149. [[CrossRef](#)]
36. Peng, Z.; Yan, J.; Fan, X.; Mizianty, M.J.; Xue, B.; Wang, K.; Hu, G.; Uversky, V.N.; Kurgan, L. Exceptionally abundant exceptions: Comprehensive characterization of intrinsic disorder in all domains of life. *Cell. Mol. Life Sci.* **2015**, *72*, 137–151. [[CrossRef](#)]
37. Tokuriki, N.; Oldfield, C.J.; Uversky, V.N.; Berezovsky, I.N.; Tawfik, D.S. Do viral proteins possess unique biophysical features? *Trends Biochem. Sci.* **2009**, *34*, 53–59. [[CrossRef](#)]
38. Xue, B.; Williams, R.W.; Oldfield, C.J.; Dunker, A.K.; Uversky, V.N. Archaic chaos: Intrinsically disordered proteins in Archaea. *BMC Syst. Biol.* **2010**, *4* (Suppl. 1), S1. [[CrossRef](#)]

39. Tompa, P.; Dosztanyi, Z.; Simon, I. Prevalent structural disorder in E. coli and S. cerevisiae proteomes. *J. Proteome Res.* **2006**, *5*, 1996–2000. [[CrossRef](#)]
40. Krasowski, M.D.; Reschly, E.J.; Ekins, S. Intrinsic disorder in nuclear hormone receptors. *J. Proteome Res.* **2008**, *7*, 4359–4372. [[CrossRef](#)]
41. Shimizu, K.; Toh, H. Interaction between intrinsically disordered proteins frequently occurs in a human protein-protein interaction network. *J. Mol. Biol.* **2009**, *392*, 1253–1265. [[CrossRef](#)]
42. Pentony, M.M.; Jones, D.T. Modularity of intrinsic disorder in the human proteome. *Proteins* **2010**, *78*, 212–221. [[CrossRef](#)] [[PubMed](#)]
43. Tompa, P.; Kalmar, L. Power law distribution defines structural disorder as a structural element directly linked with function. *J. Mol. Biol.* **2010**, *403*, 346–350. [[CrossRef](#)] [[PubMed](#)]
44. Schad, E.; Tompa, P.; Hegyi, H. The relationship between proteome size, structural disorder and organism complexity. *Genome Biol.* **2011**, *12*, R120. [[CrossRef](#)] [[PubMed](#)]
45. Dyson, H.J. Expanding the proteome: Disordered and alternatively folded proteins. *Q. Rev. Biophys.* **2011**, *44*, 467–518. [[CrossRef](#)] [[PubMed](#)]
46. Pancsa, R.; Tompa, P. Structural disorder in eukaryotes. *PLoS ONE* **2012**, *7*, e34687. [[CrossRef](#)] [[PubMed](#)]
47. Midic, U.; Obradovic, Z. Intrinsic disorder in putative protein sequences. *Proteome Sci.* **2012**, *10* (Suppl. 1), S19. [[CrossRef](#)]
48. Hegyi, H.; Tompa, P. Increased structural disorder of proteins encoded on human sex chromosomes. *Mol. Biosyst.* **2012**, *8*, 229–236. [[CrossRef](#)]
49. Korneta, I.; Bujnicki, J.M. Intrinsic disorder in the human spliceosomal proteome. *PLoS Comput. Biol.* **2012**, *8*, e1002641. [[CrossRef](#)]
50. Kahali, B.; Ghosh, T.C. Disorderiness in Escherichia coli proteome: Perception of folding fidelity and protein-protein interactions. *J. Biomol. Struct. Dyn.* **2013**, *31*, 472–476. [[CrossRef](#)]
51. Di Domenico, T.; Walsh, I.; Tosatto, S.C. Analysis and consensus of currently available intrinsic protein disorder annotation sources in the MobiDB database. *BMC Bioinform.* **2013**, *14* (Suppl. 7), S3. [[CrossRef](#)]
52. Oldfield, C.J.; Cheng, Y.; Cortese, M.S.; Brown, C.J.; Uversky, V.N.; Dunker, A.K. Comparing and combining predictors of mostly disordered proteins. *Biochemistry* **2005**, *44*, 1989–2000. [[CrossRef](#)] [[PubMed](#)]
53. Na, I.; Redmon, D.; Kopa, M.; Qin, Y.; Xue, B.; Uversky, V.N. Ordered disorder of the astrocytic dystrophin-associated protein complex in the norm and pathology. *PLoS ONE* **2013**, *8*, e73476. [[CrossRef](#)] [[PubMed](#)]
54. Wright, P.E.; Dyson, H.J. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **1999**, *293*, 321–331. [[CrossRef](#)]
55. Dunker, A.K.; Brown, C.J.; Obradovic, Z. Identification and functions of usefully disordered proteins. *Adv. Protein Chem.* **2002**, *62*, 25–49.
56. Dunker, A.K.; Brown, C.J.; Lawson, J.D.; Iakoucheva, L.M.; Obradovic, Z. Intrinsic disorder and protein function. *Biochemistry* **2002**, *41*, 6573–6582. [[CrossRef](#)]
57. Dunker, A.K.; Obradovic, Z. The protein trinity—linking function and disorder. *Nat. Biotechnol.* **2001**, *19*, 805–806. [[CrossRef](#)]
58. Dyson, H.J.; Wright, P.E. Coupling of folding and binding for unstructured proteins. *Curr. Opin. Struct. Biol.* **2002**, *12*, 54–60. [[CrossRef](#)]
59. Dyson, H.J.; Wright, P.E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 197–208. [[CrossRef](#)]
60. Tompa, P. The interplay between structure and function in intrinsically unstructured proteins. *FEBS Lett.* **2005**, *579*, 3346–3354. [[CrossRef](#)]
61. Obradovic, Z.; Peng, K.; Vucetic, S.; Radivojac, P.; Brown, C.J.; Dunker, A.K. Predicting intrinsic disorder from amino acid sequence. *Proteins* **2003**, *53*, 566–572. [[CrossRef](#)]
62. Uversky, V.N.; Oldfield, C.J.; Dunker, A.K. Showing your ID: Intrinsic disorder as an ID for recognition, regulation and cell signaling. *J. Mol. Recognit.* **2005**, *18*, 343–384. [[CrossRef](#)] [[PubMed](#)]
63. Xue, B.; Romero, P.R.; Noutsou, M.; Maurice, M.M.; Rudiger, S.G.; William, A.M., Jr.; Mizianty, M.J.; Kurgan, L.; Uversky, V.N.; Dunker, A.K. Stochastic machines as a colocalization mechanism for scaffold protein function. *FEBS Lett.* **2013**, *587*, 1587–1591. [[CrossRef](#)] [[PubMed](#)]
64. Zhang, Y.; Qiu, W.J.; Chan, S.C.; Han, J.; He, X.; Lin, S.C. Casein kinase I and casein kinase II differentially regulate axin function in Wnt and JNK pathways. *J. Biol. Chem.* **2002**, *277*, 17706–17712. [[CrossRef](#)] [[PubMed](#)]

65. Dajani, R.; Fraser, E.; Roe, S.M.; Yeo, M.; Good, V.M.; Thompson, V.; Dale, T.C.; Pearl, L.H. Structural basis for recruitment of glycogen synthase kinase 3beta to the axin-APC scaffold complex. *Embo J.* **2003**, *22*, 494–501. [[CrossRef](#)] [[PubMed](#)]
66. Xing, Y.; Takemaru, K.; Liu, J.; Berndt, J.D.; Zheng, J.J.; Moon, R.T.; Xu, W. Crystal structure of a full-length beta-catenin. *Structure* **2008**, *16*, 478–487. [[CrossRef](#)] [[PubMed](#)]
67. Wright, P.E.; Dyson, H.J. Linking folding and binding. *Curr. Opin. Struct. Biol.* **2009**, *19*, 31–38. [[CrossRef](#)] [[PubMed](#)]
68. Uversky, V.N. Multitude of binding modes attainable by intrinsically disordered proteins: A portrait gallery of disorder-based complexes. *Chem. Soc. Rev.* **2011**, *40*, 1623–1634. [[CrossRef](#)]
69. Uversky, V.N. Intrinsic disorder-based protein interactions and their modulators. *Curr. Pharm. Des.* **2013**, *19*, 4191–4213. [[CrossRef](#)]
70. Janin, J.; Sternberg, M.J. Protein flexibility, not disorder, is intrinsic to molecular recognition. *F1000 Biol. Rep.* **2013**, *5*, 2. [[CrossRef](#)]
71. Jakob, U.; Kriwacki, R.; Uversky, V.N. Conditionally and transiently disordered proteins: Awakening cryptic disorder to regulate protein function. *Chem. Rev.* **2014**, *114*, 6779–6805. [[CrossRef](#)]
72. Mitrea, D.M.; Kriwacki, R.W. Regulated unfolding of proteins in signaling. *FEBS Lett.* **2013**, *587*, 1081–1088. [[CrossRef](#)] [[PubMed](#)]
73. Bardwell, J.C.; Jakob, U. Conditional disorder in chaperone action. *Trends Biochem. Sci.* **2012**, *37*, 517–525. [[CrossRef](#)] [[PubMed](#)]
74. Creamer, T.P. Transient disorder: Calcineurin as an example. *Intrinsically Disord. Proteins* **2013**, *1*, e26412. [[CrossRef](#)]
75. Merrill, A.R.; Cohen, F.S.; Cramer, W.A. On the nature of the structural change of the colicin E1 channel peptide necessary for its translocation-competent state. *Biochemistry* **1990**, *29*, 5829–5836. [[CrossRef](#)]
76. Polverino de Laureto, P.; De Filippis, V.; Di Bello, M.; Zambonin, M.; Fontana, A. Probing the molten globule state of alpha-lactalbumin by limited proteolysis. *Biochemistry* **1995**, *34*, 12596–12604. [[CrossRef](#)]
77. De Filippis, V.; de Laureto, P.P.; Toniutti, N.; Fontana, A. Acid-induced molten globule state of a fully active mutant of human interleukin-6. *Biochemistry* **1996**, *35*, 11503–11511. [[CrossRef](#)]
78. Polverino de Laureto, P.; Frare, E.; Gottardo, R.; Van Dael, H.; Fontana, A. Partly folded states of members of the lysozyme/lactalbumin superfamily: A comparative study by circular dichroism spectroscopy and limited proteolysis. *Protein Sci.* **2002**, *11*, 2932–2946. [[CrossRef](#)]
79. Fontana, A.; de Laureto, P.P.; Spolaore, B.; Frare, E.; Picotti, P.; Zambonin, M. Probing protein structure by limited proteolysis. *Acta Biochim. Pol.* **2004**, *51*, 299–321. [[CrossRef](#)]
80. Campioni, S.; Mossuto, M.F.; Torrassa, S.; Calloni, G.; de Laureto, P.P.; Relini, A.; Fontana, A.; Chiti, F. Conformational properties of the aggregation precursor state of HypF-N. *J. Mol. Biol.* **2008**, *379*, 554–567. [[CrossRef](#)]
81. Hegyi, H.; Tompa, P. Intrinsically disordered proteins display no preference for chaperone binding in vivo. *PLoS Comput. Biol.* **2008**, *4*, e1000017. [[CrossRef](#)]
82. Galea, C.A.; Pagala, V.R.; Obenauer, J.C.; Park, C.G.; Slaughter, C.A.; Kriwacki, R.W. Proteomic studies of the intrinsically unstructured mammalian proteome. *J. Proteome Res.* **2006**, *5*, 2839–2848. [[CrossRef](#)] [[PubMed](#)]
83. Tompa, P. Intrinsically unstructured proteins. *Trends Biochem. Sci.* **2002**, *27*, 527–533. [[CrossRef](#)]
84. Uversky, V.N. Biophysical Methods to Investigate Intrinsically Disordered Proteins: Avoiding an “Elephant and Blind Men” Situation. *Adv. Exp. Med. Biol.* **2015**, *870*, 215–260. [[CrossRef](#)] [[PubMed](#)]
85. Uversky, V.N.; Dunker, A.K. Multiparametric analysis of intrinsically disordered proteins: Looking at intrinsic disorder through compound eyes. *Anal. Chem.* **2012**, *84*, 2096–2104. [[CrossRef](#)] [[PubMed](#)]
86. Receveur-Brechot, V.; Bourhis, J.M.; Uversky, V.N.; Canard, B.; Longhi, S. Assessing protein disorder and induced folding. *Proteins* **2006**, *62*, 24–45. [[CrossRef](#)] [[PubMed](#)]
87. Dunker, A.K.; Cortese, M.S.; Romero, P.; Iakoucheva, L.M.; Uversky, V.N. Flexible nets: The roles of intrinsic disorder in protein interaction networks. *FEBS J.* **2005**, *272*, 5129–5148. [[CrossRef](#)]
88. Dunker, A.K.; Silman, I.; Uversky, V.N.; Sussman, J.L. Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.* **2008**, *18*, 756–764. [[CrossRef](#)]
89. Cortese, M.S.; Uversky, V.N.; Dunker, A.K. Intrinsic disorder in scaffold proteins: Getting more from less. *Prog. Biophys. Mol. Biol.* **2008**, *98*, 85–106. [[CrossRef](#)]

90. Tompa, P.; Prilusky, J.; Silman, I.; Sussman, J.L. Structural disorder serves as a weak signal for intracellular protein degradation. *Proteins* **2008**, *71*, 903–909. [[CrossRef](#)]
91. Suskiewicz, M.J.; Sussman, J.L.; Silman, I.; Shaul, Y. Context-dependent resistance to proteolysis of intrinsically disordered proteins. *Protein Sci.* **2011**, *20*, 1285–1297. [[CrossRef](#)]
92. Gsponer, J.; Futschik, M.E.; Teichmann, S.A.; Babu, M.M. Tight regulation of unstructured proteins: From transcript synthesis to protein degradation. *Science* **2008**, *322*, 1365–1368. [[CrossRef](#)] [[PubMed](#)]
93. Uversky, V.N.; Dunker, A.K. Biochemistry. Controlled chaos. *Science* **2008**, *322*, 1340–1341. [[CrossRef](#)] [[PubMed](#)]
94. Xie, H.; Vucetic, S.; Iakoucheva, L.M.; Oldfield, C.J.; Dunker, A.K.; Obradovic, Z.; Uversky, V.N. Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications, and diseases associated with intrinsically disordered proteins. *J. Proteome Res.* **2007**, *6*, 1917–1932. [[CrossRef](#)]
95. Iakoucheva, L.M.; Radivojac, P.; Brown, C.J.; O'Connor, T.R.; Sikes, J.G.; Obradovic, Z.; Dunker, A.K. The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res.* **2004**, *32*, 1037–1049. [[CrossRef](#)]
96. Daily, K.M.; Radivojac, P.; Dunker, A.K. Intrinsic disorder and protein modifications: Building an SVM predictor for methylation. In Proceedings of the IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology, La Jolla, CA, USA, 15 November 2005. [[CrossRef](#)]
97. Radivojac, P.; Vacic, V.; Haynes, C.; Cocklin, R.R.; Mohan, A.; Heyen, J.W.; Goebel, M.G.; Iakoucheva, L.M. Identification, analysis, and prediction of protein ubiquitination sites. *Proteins* **2010**, *78*, 365–380. [[CrossRef](#)]
98. Reddy, K.D.; Malipeddi, J.; DeForte, S.; Pejaver, V.; Radivojac, P.; Uversky, V.N.; Deschenes, R.J. Physicochemical sequence characteristics that influence S-palmitoylation propensity. *J. Biomol. Struct. Dyn.* **2017**, *35*, 2337–2350. [[CrossRef](#)]
99. Pejaver, V.; Hsu, W.L.; Xin, F.; Dunker, A.K.; Uversky, V.N.; Radivojac, P. The structural and functional signatures of proteins that undergo multiple events of post-translational modification. *Protein Sci.* **2014**, *23*, 1077–1093. [[CrossRef](#)]
100. Ge, X.; Conley, A.J.; Brandle, J.E.; Truant, R.; Filipe, C.D. In vivo formation of protein based aqueous microcompartments. *J. Am. Chem. Soc.* **2009**, *131*, 9094–9099. [[CrossRef](#)]
101. Privalov, P.L.; Khechinashvili, N.N. A thermodynamic approach to the problem of stabilization of globular protein structure: A calorimetric study. *J. Mol. Biol.* **1974**, *86*, 665–684. [[CrossRef](#)]
102. Privalov, P.L. Stability of proteins: Small globular proteins. *Adv. Protein Chem.* **1979**, *33*, 167–241. [[CrossRef](#)]
103. Ohgushi, M.; Wada, A. Liquid-like state of side chains at the intermediate stage of protein denaturation. *Adv. Biophys.* **1984**, *18*, 75–90. [[CrossRef](#)]
104. Kuwajima, K. The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins* **1989**, *6*, 87–103. [[CrossRef](#)] [[PubMed](#)]
105. Goto, Y.; Calciano, L.J.; Fink, A.L. Acid-induced folding of proteins. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 573–577. [[CrossRef](#)] [[PubMed](#)]
106. Christensen, H.; Pain, R.H. Molten globule intermediates and protein folding. *Eur. Biophys. J.* **1991**, *19*, 221–229. [[CrossRef](#)] [[PubMed](#)]
107. Haynie, D.T.; Freire, E. Structural energetics of the molten globule state. *Proteins* **1993**, *16*, 115–140. [[CrossRef](#)] [[PubMed](#)]
108. Dobson, C.M. Protein folding. Solid evidence for molten globules. *Curr. Biol.* **1994**, *4*, 636–640. [[CrossRef](#)]
109. Fink, A.L.; Calciano, L.J.; Goto, Y.; Kurotsu, T.; Palleros, D.R. Classification of acid denaturation of proteins: Intermediates and unfolded states. *Biochemistry* **1994**, *33*, 12504–12511. [[CrossRef](#)]
110. Ptitsyn, O.B. Structures of folding intermediates. *Curr. Opin. Struct. Biol.* **1995**, *5*, 74–78. [[CrossRef](#)]
111. Creighton, T.E. How important is the molten globule for correct protein folding? *Trends Biochem. Sci.* **1997**, *22*, 6–10. [[CrossRef](#)]
112. Arai, M.; Kuwajima, K. Role of the molten globule state in protein folding. *Adv. Protein Chem.* **2000**, *53*, 209–282.
113. Englander, S.W. Protein folding intermediates and pathways studied by hydrogen exchange. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 213–238. [[CrossRef](#)]
114. Redfield, C. Using nuclear magnetic resonance spectroscopy to study molten globule states of proteins. *Methods* **2004**, *34*, 121–132. [[CrossRef](#)] [[PubMed](#)]

115. Baldwin, R.L.; Rose, G.D. Molten globules, entropy-driven conformational change and protein folding. *Curr. Opin. Struct. Biol.* **2013**, *23*, 4–10. [[CrossRef](#)] [[PubMed](#)]
116. Bychkova, V.E.; Semisotnov, G.V.; Balabanov, V.A.; Finkelstein, A.V. The Molten Globule Concept: 45 Years Later. *Biochemistry* **2018**, *83*, S33–S47. [[CrossRef](#)] [[PubMed](#)]
117. Judy, E.; Kishore, N. A look back at the molten globule state of proteins: Thermodynamic aspects. *Biophys. Rev.* **2019**, *11*, 365–375. [[CrossRef](#)]
118. Vassilenko, K.S.; Uversky, V.N. Native-like secondary structure of molten globules. *Biochim. Biophys. Acta* **2002**, *1594*, 168–177. [[CrossRef](#)]
119. Uversky, V.N. Use of fast protein size-exclusion liquid chromatography to study the unfolding of proteins which denature through the molten globule. *Biochemistry* **1993**, *32*, 13288–13298. [[CrossRef](#)]
120. Eliezer, D.; Chiba, K.; Tsuruta, H.; Doniach, S.; Hodgson, K.O.; Kihara, H. Evidence of an Associative Intermediate on the Myoglobin Refolding Pathway. *Biophys. J.* **1993**, *65*, 912–917. [[CrossRef](#)]
121. Kataoka, M.; Hagihara, Y.; Mihara, K.; Goto, Y. Molten Globule of Cytochrome-C Studied by Small-Angle X-Ray-Scattering. *J. Mol. Biol.* **1993**, *229*, 591–596. [[CrossRef](#)]
122. Semisotnov, G.V.; Kihara, H.; Kotova, N.V.; Kimura, K.; Amemiya, Y.; Wakabayashi, K.; Serdyuk, I.N.; Timchenko, A.A.; Chiba, K.; Nikaido, K.; et al. Protein globularization during folding. A study by synchrotron small-angle X-ray scattering. *J. Mol. Biol.* **1996**, *262*, 559–574. [[CrossRef](#)]
123. Kataoka, M.; Kuwajima, K.; Tokunaga, F.; Goto, Y. Structural characterization of the molten globule of alpha-lactalbumin by solution X-ray scattering. *Protein Sci.* **1997**, *6*, 422–430. [[CrossRef](#)]
124. Uversky, V.N.; Karnoup, A.S.; Segel, D.J.; Seshadri, S.; Doniach, S.; Fink, A.L. Anion-induced folding of Staphylococcal nuclease: Characterization of multiple equilibrium partially folded intermediates. *J. Mol. Biol.* **1998**, *278*, 879–894. [[CrossRef](#)] [[PubMed](#)]
125. Baum, J.; Dobson, C.M.; Evans, P.A.; Hanley, C. Characterization of a Partly Folded Protein by Nmr Methods-Studies on the Molten Globule State of Guinea-Pig Alpha-Lactalbumin. *Biochemistry* **1989**, *28*, 7–13. [[CrossRef](#)] [[PubMed](#)]
126. Bushnell, G.W.; Louie, G.V.; Brayer, G.D. High-resolution three-dimensional structure of horse heart cytochrome c. *J. Mol. Biol.* **1990**, *214*, 585–595. [[CrossRef](#)]
127. Jeng, M.F.; Englander, S.W.; Elove, G.A.; Wand, A.J.; Roder, H. Structural description of acid-denatured cytochrome c by hydrogen exchange and 2D NMR. *Biochemistry* **1990**, *29*, 10433–10437. [[CrossRef](#)] [[PubMed](#)]
128. Chyan, C.L.; Wormald, C.; Dobson, C.M.; Evans, P.A.; Baum, J. Structure and stability of the molten globule state of guinea-pig alpha-lactalbumin: A hydrogen exchange study. *Biochemistry* **1993**, *32*, 5681–5691. [[CrossRef](#)] [[PubMed](#)]
129. Wu, L.C.; Laub, P.B.; Elove, G.A.; Carey, J.; Roder, H. A noncovalent peptide complex as a model for an early folding intermediate of cytochrome c. *Biochemistry* **1993**, *32*, 10271–10276. [[CrossRef](#)] [[PubMed](#)]
130. Gillespie, J.R.; Shortle, D. Characterization of long-range structure in the denatured state of staphylococcal nuclease. II. Distance restraints from paramagnetic relaxation and calculation of an ensemble of structures. *J. Mol. Biol.* **1997**, *268*, 170–184. [[CrossRef](#)]
131. Gillespie, J.R.; Shortle, D. Characterization of long-range structure in the denatured state of staphylococcal nuclease. I. Paramagnetic relaxation enhancement by nitroxide spin labels. *J. Mol. Biol.* **1997**, *268*, 158–169. [[CrossRef](#)]
132. Eliezer, D.; Yao, J.; Dyson, H.J.; Wright, P.E. Structural and dynamic characterization of partially folded states of apomyoglobin and implications for protein folding. *Nat. Struct. Biol.* **1998**, *5*, 148–155. [[CrossRef](#)]
133. Bose, H.S.; Whittall, R.M.; Baldwin, M.A.; Miller, W.L. The active form of the steroidogenic acute regulatory protein, StAR, appears to be a molten globule. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 7250–7255. [[CrossRef](#)]
134. van Mierlo, C.P.; Steensma, E. Protein folding and stability investigated by fluorescence, circular dichroism (CD), and nuclear magnetic resonance (NMR) spectroscopy: The flavodoxin story. *J. Biotechnol.* **2000**, *79*, 281–298. [[CrossRef](#)]
135. Bracken, C. NMR spin relaxation methods for characterization of disorder and folding in proteins. *J. Mol. Graph. Model.* **2001**, *19*, 3–12. [[CrossRef](#)]
136. Ikeguchi, M. Transient non-native helix formation during the folding of beta-lactoglobulin. *Biomolecules* **2014**, *4*, 202–216. [[CrossRef](#)] [[PubMed](#)]
137. Ray, S.S.; Singh, S.K.; Balaram, P. An electrospray ionization mass spectrometry investigation of 1-anilino-8-naphthalene-sulfonate (ANS) binding to proteins. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 428–438. [[CrossRef](#)]

138. Uversky, V.N.; Winter, S.; Lober, G. Use of fluorescence decay times of 8-ANS-protein complexes to study the conformational transitions in proteins which unfold through the molten globule state. *Biophys. Chem.* **1996**, *60*, 79–88. [[CrossRef](#)]
139. Shi, L.; Palleros, D.R.; Fink, A.L. Protein conformational changes induced by 1,1'-bis(4-anilino-5-naphthalenesulfonic acid): Preferential binding to the molten globule of DnaK. *Biochemistry* **1994**, *33*, 7536–7546. [[CrossRef](#)]
140. Semisotnov, G.V.; Rodionova, N.A.; Razgulyaev, O.I.; Uversky, V.N.; Gripas, A.F.; Gilmanshin, R.I. Study of the “molten globule” intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers* **1991**, *31*, 119–128. [[CrossRef](#)]
141. Uversky, V.N.; Ptitsyn, O.B. All-or-none solvent-induced transitions between native, molten globule and unfolded states in globular proteins. *Fold. Des.* **1996**, *1*, 117–122. [[CrossRef](#)]
142. Grosberg, A.Y. Collapse and intramolecular phase layering in a polymer in which each unit may be in two states. *Biophysics* **1984**, *29*, 621–626.
143. Adonts, V.G.; Birshtein, T.M.; Elyashevich, A.M.; Skvortsov, A.M. Intramolecular conformational transitions “random coil-helix-folded structure” in polypeptides. *Biopolymers* **1976**, *15*, 1037–1059. [[CrossRef](#)]
144. Finkelstein, A.V. Kinetics of antiparallel beta-structure formation. *Bioorganicheskaya Khimiya* **1978**, *4*, 340–344.
145. Finkelstein, A.V. Rate of beta-structure formation in polypeptides. *Proteins* **1991**, *9*, 23–27. [[CrossRef](#)] [[PubMed](#)]
146. Gil'manshin, R.I.; Dolgikh, D.A.; Ptitsyn, O.B.; Finkel'shtein, A.V.; Shakhnovich, E.I. Protein globule without the unique three-dimensional structure: Experimental data for alpha-lactalbumins and general model. *Biofizika* **1982**, *27*, 1005–1016. [[PubMed](#)]
147. Dolgikh, D.A.; Gilmanshin, R.I.; Brazhnikov, E.V.; Bychkova, V.E.; Semisotnov, G.V.; Venyaminov, S.; Ptitsyn, O.B. Alpha-Lactalbumin: Compact state with fluctuating tertiary structure? *FEBS Lett.* **1981**, *136*, 311–315. [[CrossRef](#)]
148. Kayaman, N.; Guerel, E.E.; Baysal, B.M.; ve Karasz, F. Kinetics of coil-globule collapse in poly(methyl methacrylate) in dilute solutions below theta temperatures. *Macromolecules* **1999**, *32*, 8399–8403. [[CrossRef](#)]
149. Podewitz, M.; Wang, Y.; Quoika, P.K.; Loeffler, J.R.; Schauerl, M.; Liedl, K.R. Coil-Globule Transition Thermodynamics of Poly(N-isopropylacrylamide). *J. Phys. Chem. B* **2019**, *123*, 8838–8847. [[CrossRef](#)]
150. Grosberg, A.Y.; Khokhlov, A.R. *Statistical Physics of Macromolecules*; American Institute of Physics: New York, NY, USA, 1994.
151. Shakhnovich, E.I.; Finkel'shtein, A.V. The theory of cooperative transitions in protein globules. *Dokl. Akad. Nauk SSSR* **1982**, *267*, 1247–1250.
152. Finkelstein, A.V.; Shakhnovich, E.I. Theory of cooperative transitions in protein molecules. II. Phase diagram for a protein molecule in solution. *Biopolymers* **1989**, *28*, 1681–1694. [[CrossRef](#)]
153. Shakhnovich, E.I.; Finkelstein, A.V. Theory of cooperative transitions in protein molecules. I. Why denaturation of globular protein is a first-order phase transition. *Biopolymers* **1989**, *28*, 1667–1680. [[CrossRef](#)]
154. Schulz, G.E.; Schirmer, R.H. *Principles of Protein Structure*, 2nd ed.; Springer: New York, NY, USA, 2013.
155. Kharakoz, D.P.; Bychkova, V.E. Molten globule of human alpha-lactalbumin: Hydration, density, and compressibility of the interior. *Biochemistry* **1997**, *36*, 1882–1890. [[CrossRef](#)]
156. Jha, S.K.; Udgaonkar, J.B. Direct evidence for a dry molten globule intermediate during the unfolding of a small protein. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 12289–12294. [[CrossRef](#)] [[PubMed](#)]
157. Jennings, P.A.; Wright, P.E. Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. *Science* **1993**, *262*, 892–896. [[CrossRef](#)] [[PubMed](#)]
158. Samatova, E.N.; Katina, N.S.; Balobanov, V.A.; Melnik, B.S.; Dolgikh, D.A.; Bychkova, V.E.; Finkelstein, A.V. How strong are side chain interactions in the folding intermediate? *Protein Sci.* **2009**, *18*, 2152–2159. [[CrossRef](#)] [[PubMed](#)]
159. Eichmann, C.; Preissler, S.; Riek, R.; Deuerling, E. Cotranslational structure acquisition of nascent polypeptides monitored by NMR spectroscopy. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 9111–9116. [[CrossRef](#)] [[PubMed](#)]
160. Han, Y.; David, A.; Liu, B.; Magadan, J.G.; Bennink, J.R.; Yewdell, J.W.; Qian, S.B. Monitoring cotranslational protein folding in mammalian cells at codon resolution. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 12467–12472. [[CrossRef](#)]
161. Holtkamp, W.; Kokic, G.; Jager, M.; Mittelstaet, J.; Komar, A.A.; Rodnina, M.V. Cotranslational protein folding on the ribosome monitored in real time. *Science* **2015**, *350*, 1104–1107. [[CrossRef](#)]

162. Flanagan, J.M.; Kataoka, M.; Shortle, D.; Engelman, D.M. Truncated staphylococcal nuclease is compact but disordered. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 748–752. [[CrossRef](#)]
163. Levinthal, C. Are there pathways for protein folding? *J. Chim. Phys.* **1968**, *65*, 44–45. [[CrossRef](#)]
164. Levinthal, C. How to fold graciously. In Proceedings of the Mössbauer Spectroscopy in Biological Systems, Monticello, IL, USA, 17–18 March 1969; pp. 22–24.
165. Go, N.; Abe, H. Noninteracting local-structure model of folding and unfolding transition in globular proteins. I. Formulation. *Biopolymers* **1981**, *20*, 991–1011. [[CrossRef](#)]
166. Leopold, P.E.; Montal, M.; Onuchic, J.N. Protein folding funnels: A kinetic approach to the sequence-structure relationship. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 8721–8725. [[CrossRef](#)]
167. Wolynes, P.G.; Onuchic, J.N.; Thirumalai, D. Navigating the folding routes. *Science* **1995**, *267*, 1619–1620. [[CrossRef](#)] [[PubMed](#)]
168. Zwanzig, R.; Szabo, A.; Bagchi, B. Levinthal's paradox. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 20–22. [[CrossRef](#)] [[PubMed](#)]
169. Abkevich, V.I.; Gutin, A.M.; Shakhnovich, E.I. Specific nucleus as the transition state for protein folding: Evidence from the lattice model. *Biochemistry* **1994**, *33*, 10026–10036. [[CrossRef](#)] [[PubMed](#)]
170. Sali, A.; Shakhnovich, E.; Karplus, M. Kinetics of protein folding. A lattice model study of the requirements for folding to the native state. *J. Mol. Biol.* **1994**, *235*, 1614–1636. [[CrossRef](#)]
171. Finkelstein, A.V.; Badretdinov, A.Y. Physical reason for fast folding of the stable spatial structure of proteins: A solution of the Levinthal paradox. *Mol. Biol.* **1997**, *31*, 391–398.
172. Finkelstein, A.V.; Badretdinov, A. Rate of protein folding near the point of thermodynamic equilibrium between the coil and the most stable chain fold. *Fold. Des.* **1997**, *2*, 115–121. [[CrossRef](#)]
173. Finkelstein, A.V.; Badretdinov, A.Y. Influence of chain knotting on the rate of folding. ADDENDUM to rate of protein folding near the point of thermodynamic equilibrium between the coil and the most stable chain fold. *Fold. Des.* **1998**, *3*, 67–68. [[CrossRef](#)]
174. Finkelstein, A.V.; Badretdinov, A.; Gutin, A.M. Why do protein architectures have Boltzmann-like statistics? *Proteins* **1995**, *23*, 142–150. [[CrossRef](#)]
175. Finkelstein, A.V.; Gutin, A.M.; Badretdinov, A. Boltzmann-like statistics of protein architectures. Origins and consequences. *Subcell. Biochem.* **1995**, *24*, 1–26. [[CrossRef](#)]
176. Garbuzynskiy, S.O.; Ivankov, D.N.; Bogatyreva, N.S.; Finkelstein, A.V. Golden triangle for folding rates of globular proteins. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 147–150. [[CrossRef](#)]
177. Eyring, H. The activated complex in chemical reactions. *J. Chem. Phys.* **1935**, *3*, 107–115. [[CrossRef](#)]
178. Pauling, L. *General Chemistry*; W.H. Freeman & Co.: New York, NY, USA, 1970.
179. Emanuel, N.M.; Knorre, D.G. *The Course in Chemical Kinetics*, 4th ed.; Vysshaja Shkola: Moscow, Russia, 1984.
180. Zana, R. On the rate determining step for helix propagation in the helix-coil transition of polypeptides in solution. *Biopolymers* **1975**, *14*, 2425–2428. [[CrossRef](#)]
181. Finkelstein, A.V.; Badretdin, A.J.; Galzitskaya, O.V.; Ivankov, D.N.; Bogatyreva, N.S.; Garbuzynskiy, S.O. There and back again: Two views on the protein folding puzzle. *Phys. Life Rev.* **2017**, *21*, 56–71. [[CrossRef](#)] [[PubMed](#)]
182. Fersht, A.R. *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding*; W. H. Freeman & Co: New York, NY, USA, 1999.
183. Takada, S. Go-ing for the prediction of protein folding mechanisms. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11698–11700. [[CrossRef](#)]
184. Alm, E.; Baker, D. Prediction of protein-folding mechanisms from free-energy landscapes derived from native structures. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11305–11310. [[CrossRef](#)]
185. Galzitskaya, O.V.; Finkelstein, A.V. A theoretical search for folding/unfolding nuclei in three-dimensional protein structures. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11299–11304. [[CrossRef](#)]
186. Munoz, V.; Eaton, W.A. A simple model for calculating the kinetics of protein folding from three-dimensional structures. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11311–11316. [[CrossRef](#)]
187. Galzitskaya, O.V.; Ivankov, D.N.; Finkelstein, A.V. Folding nuclei in proteins. *FEBS Lett.* **2001**, *489*, 113–118. [[CrossRef](#)]
188. Garbuzynskiy, S.O.; Finkelstein, A.V.; Galzitskaya, O.V. Outlining folding nuclei in globular proteins. *J. Mol. Biol.* **2004**, *336*, 509–525. [[CrossRef](#)]

189. Finkelstein, A.V.; Ivankov, D.N.; Garbuzynskiy, S.O.; Galzitskaya, O.V. Understanding the folding rates and folding nuclei of globular proteins. *Curr. Protein Pept. Sci.* **2007**, *8*, 521–536. [[CrossRef](#)]
190. Garbuzynskiy, S.O.; Kondratova, M.S. Structural features of protein folding nuclei. *FEBS Lett.* **2008**, *582*, 768–772. [[CrossRef](#)] [[PubMed](#)]
191. Ivankov, D.N.; Finkelstein, A.V. Protein folding as flow across a network of folding-unfolding pathways. 2. The “in-water” case. *J. Phys. Chem. B* **2010**, *114*, 7930–7934. [[CrossRef](#)]
192. Ivankov, D.N.; Finkelstein, A.V. Protein folding as flow across a network of folding-unfolding pathways. 1. The mid-transition case. *J. Phys. Chem. B* **2010**, *114*, 7920–7929. [[CrossRef](#)] [[PubMed](#)]
193. Finkelstein, A.V.; Galzitskaya, O.V. Physics of protein folding. *Phys. Life Rev.* **2004**, *1*, 23–56. [[CrossRef](#)]
194. Finkelstein, A.V.; Bogatyreva, N.S.; Garbuzynskiy, S.O. Restrictions to protein folding determined by the protein size. *FEBS Lett.* **2013**, *587*, 1884–1890. [[CrossRef](#)] [[PubMed](#)]
195. Ivankov, D.N.; Finkelstein, A.V. Prediction of protein folding rates from the amino acid sequence-predicted secondary structure. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 8942–8944. [[CrossRef](#)]
196. Ivankov, D.N.; Garbuzynskiy, S.O.; Alm, E.; Plaxco, K.W.; Baker, D.; Finkelstein, A.V. Contact order revisited: Influence of protein size on the folding rate. *Protein Sci.* **2003**, *12*, 2057–2062. [[CrossRef](#)] [[PubMed](#)]
197. Galzitskaya, O.V.; Garbuzynskiy, S.O.; Ivankov, D.N.; Finkelstein, A.V. Chain length is the main determinant of the folding rate for proteins with three-state folding kinetics. *Proteins* **2003**, *51*, 162–166. [[CrossRef](#)]
198. Ivankov, D.N.; Finkelstein, A.V. Solution of the Levinthal’s paradox and a physical theory of protein folding rates. *Biophysics* **2016**, *61*, 1–5.
199. Finkelstein, A.V.; Garbuzynskiy, S.O. Reduction of the Search Space for the Folding of Proteins at the Level of Formation and Assembly of Secondary Structures: A New View on the Solution of Levinthal’s Paradox. *Chemphyschem* **2015**, *16*, 3375–3378. [[CrossRef](#)]
200. Finkelstein, A.V. Some additional remarks to the solution of the protein folding puzzle: Reply to comments on “There and back again: Two views on the protein folding puzzle”. *Phys. Life Rev.* **2017**, *21*, 77–79. [[CrossRef](#)]
201. Fink, A.L. Compact intermediate states in protein folding. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 495–522. [[CrossRef](#)] [[PubMed](#)]
202. Daughdrill, G.W.; Pielak, G.J.; Uversky, V.N.; Cortese, M.S.; Dunker, A.K. Natively disordered proteins. In *Handbook of Protein Folding*; Buchner, J., Kiefhaber, T., Eds.; Wiley-VCH, Verlag GmbH & Co. KGaA: Weinheim, Germany, 2005; pp. 271–353.
203. Maity, H.; Maity, M.; Krishna, M.M.; Mayne, L.; Englander, S.W. Protein folding: The stepwise assembly of foldon units. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 4741–4746. [[CrossRef](#)] [[PubMed](#)]
204. Englander, S.W.; Mayne, L.; Krishna, M.M. Protein folding and misfolding: Mechanism and principles. *Q. Rev. Biophys.* **2007**, *40*, 287–326. [[CrossRef](#)] [[PubMed](#)]
205. Bedard, S.; Krishna, M.M.; Mayne, L.; Englander, S.W. Protein folding: Independent unrelated pathways or predetermined pathway with optional errors. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 7182–7187. [[CrossRef](#)]
206. Krishna, M.M.; Englander, S.W. A unified mechanism for protein folding: Predetermined pathways with optional errors. *Protein Sci.* **2007**, *16*, 449–464. [[CrossRef](#)]
207. Lindberg, M.O.; Oliveberg, M. Malleability of protein folding pathways: A simple reason for complex behaviour. *Curr. Opin. Struct. Biol.* **2007**, *17*, 21–29. [[CrossRef](#)]
208. Maity, H.; Maity, M.; Englander, S.W. How cytochrome c folds, and why: Submolecular foldon units and their stepwise sequential stabilization. *J. Mol. Biol.* **2004**, *343*, 223–233. [[CrossRef](#)]
209. Bedard, S.; Mayne, L.C.; Peterson, R.W.; Wand, A.J.; Englander, S.W. The foldon substructure of staphylococcal nuclease. *J. Mol. Biol.* **2008**, *376*, 1142–1154. [[CrossRef](#)]
210. Uversky, V.N. Unusual biophysics of intrinsically disordered proteins. *Biochim. Biophys. Acta* **2013**, *1834*, 932–951. [[CrossRef](#)]
211. Uversky, V.N. p53 Proteoforms and Intrinsic Disorder: An Illustration of the Protein Structure-Function Continuum Concept. *Int. J. Mol. Sci.* **2016**, *17*, 1874. [[CrossRef](#)]
212. Uversky, V.N. Protein intrinsic disorder and structure-function continuum. *Prog. Mol. Biol. Transl. Sci.* **2019**, *166*, 1–17. [[CrossRef](#)] [[PubMed](#)]
213. Uversky, V.N. Intrinsically disordered proteins and their “mysterious” (meta)physics. *Front. Phys.* **2019**, *7*, 10. [[CrossRef](#)]

214. Uversky, V.N. Dancing Protein Clouds: The Strange Biology and Chaotic Physics of Intrinsically Disordered Proteins. *J. Biol. Chem.* **2016**, *291*, 6681–6688. [[CrossRef](#)] [[PubMed](#)]
215. Smith, L.M.; Kelleher, N.L.; Consortium for Top Down Proteomics. Proteoform: A single term describing protein complexity. *Nat. Methods* **2013**, *10*, 186–187. [[CrossRef](#)]
216. Fonin, A.V.; Darling, A.L.; Kuznetsova, I.M.; Turoverov, K.K.; Uversky, V.N. Multi-functionality of proteins involved in GPCR and G protein signaling: Making sense of structure-function continuum with intrinsic disorder-based proteoforms. *Cell. Mol. Life Sci.* **2019**, *76*, 4461–4492. [[CrossRef](#)] [[PubMed](#)]
217. Uversky, V.N. Paradoxes and wonders of intrinsic disorder: Complexity of simplicity. *Intrinsically Disord. Proteins* **2016**, *4*, e1135015. [[CrossRef](#)] [[PubMed](#)]
218. Uversky, V.N.; Oldfield, C.J.; Dunker, A.K. Intrinsically disordered proteins in human diseases: Introducing the D2 concept. *Annu. Rev. Biophys.* **2008**, *37*, 215–246. [[CrossRef](#)]
219. Fisher, C.K.; Stultz, C.M. Constructing ensembles for intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **2011**, *21*, 426–431. [[CrossRef](#)]
220. Hemmings, H.C., Jr.; Nairn, A.C.; Aswad, D.W.; Greengard, P. DARPP-32, a dopamine- and adenosine 3':5'-monophosphate-regulated phosphoprotein enriched in dopamine-innervated brain regions. II. Purification and characterization of the phosphoprotein from bovine caudate nucleus. *J. Neurosci.* **1984**, *4*, 99–110. [[CrossRef](#)]
221. Gast, K.; Damaschun, H.; Eckert, K.; Schulze-Forster, K.; Maurer, H.R.; Muller-Frohne, M.; Zirwer, D.; Czarnecki, J.; Damaschun, G. Prothymosin alpha: A biologically active protein with random coil conformation. *Biochemistry* **1995**, *34*, 13211–13218. [[CrossRef](#)]
222. Weinreb, P.H.; Zhen, W.; Poon, A.W.; Conway, K.A.; Lansbury, P.T., Jr. NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded. *Biochemistry* **1996**, *35*, 13709–13715. [[CrossRef](#)] [[PubMed](#)]
223. Romero, P.; Obradovic, Z.; Kissinger, C.R.; Villafranca, J.E.; Garner, E.; Guillot, S.; Dunker, A.K. Thousands of proteins likely to have long disordered regions. *Pac. Symp. Biocomput.* **1998**, *3*, 437–448.
224. Garner, E.; Cannon, P.; Romero, P.; Obradovic, Z.; Dunker, A.K. Predicting Disordered Regions from Amino Acid Sequence: Common Themes Despite Differing Structural Characterization. *Genome Inform. Ser. Workshop Genome Inform.* **1998**, *9*, 201–213. [[PubMed](#)]
225. Williams, R.M.; Obradovi, Z.; Mathura, V.; Braun, W.; Garner, E.C.; Young, J.; Takayama, S.; Brown, C.J.; Dunker, A.K. The protein non-folding problem: Amino acid determinants of intrinsic order and disorder. *Pac. Symp. Biocomput.* **2001**, 89–100. [[CrossRef](#)]
226. Romero, P.; Obradovic, Z.; Li, X.; Garner, E.C.; Brown, C.J.; Dunker, A.K. Sequence complexity of disordered protein. *Proteins* **2001**, *42*, 38–48. [[CrossRef](#)]
227. Radivojac, P.; Iakoucheva, L.M.; Oldfield, C.J.; Obradovic, Z.; Uversky, V.N.; Dunker, A.K. Intrinsic disorder and functional proteomics. *Biophys. J.* **2007**, *92*, 1439–1456. [[CrossRef](#)] [[PubMed](#)]
228. Vacic, V.; Uversky, V.N.; Dunker, A.K.; Lonardi, S. Composition Profiler: A tool for discovery and visualization of amino acid composition differences. *BMC Bioinform.* **2007**, *8*, 211. [[CrossRef](#)]
229. Ban, N.; Nissen, P.; Hansen, J.; Moore, P.B.; Steitz, T.A. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* **2000**, *289*, 905–920. [[CrossRef](#)]
230. Teschke, C.M.; King, J. Folding and assembly of oligomeric proteins in *Escherichia coli*. *Curr. Opin. Biotechnol.* **1992**, *3*, 468–473. [[CrossRef](#)]
231. Xu, D.; Tsai, C.J.; Nussinov, R. Mechanism and evolution of protein dimerization. *Protein Sci.* **1998**, *7*, 533–544. [[CrossRef](#)]
232. Gunasekaran, K.; Tsai, C.J.; Nussinov, R. Analysis of ordered and disordered protein complexes reveals structural features discriminating between stable and unstable monomers. *J. Mol. Biol.* **2004**, *341*, 1327–1341. [[CrossRef](#)] [[PubMed](#)]
233. Schulz, G.E. Nucleotide Binding Proteins. In *Molecular Mechanism of Biological Recognition*; Balaban, M., Ed.; Elsevier/North-Holland Biomedical Press: New York, NY, USA, 1979; pp. 79–94.
234. Iakoucheva, L.M.; Brown, C.J.; Lawson, J.D.; Obradovic, Z.; Dunker, A.K. Intrinsic disorder in cell-signaling and cancer-associated proteins. *J. Mol. Biol.* **2002**, *323*, 573–584. [[CrossRef](#)]
235. Meador, W.E.; Means, A.R.; Quijcho, F.A. Modulation of calmodulin plasticity in molecular recognition on the basis of x-ray structures. *Science* **1993**, *262*, 1718–1721. [[CrossRef](#)] [[PubMed](#)]

236. Kriwacki, R.W.; Hengst, L.; Tennant, L.; Reed, S.I.; Wright, P.E. Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: Conformational disorder mediates binding diversity. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 11504–11509. [[CrossRef](#)]
237. Dunker, A.K.; Garner, E.; Guillot, S.; Romero, P.; Albrecht, K.; Hart, J.; Obradovic, Z.; Kissinger, C.; Villafranca, J.E. Protein disorder and the evolution of molecular recognition: Theory, predictions and observations. *Pac. Symp. Biocomput.* **1998**, *3*, 473–484.
238. Hsu, W.L.; Oldfield, C.J.; Xue, B.; Meng, J.; Huang, F.; Romero, P.; Uversky, V.N.; Dunker, A.K. Exploring the binding diversity of intrinsically disordered proteins involved in one-to-many binding. *Protein Sci.* **2013**, *22*, 258–273. [[CrossRef](#)]
239. Oldfield, C.J.; Dunker, A.K. Intrinsically disordered proteins and intrinsically disordered protein regions. *Annu. Rev. Biochem.* **2014**, *83*, 553–584. [[CrossRef](#)]
240. Oldfield, C.J.; Meng, J.; Yang, J.Y.; Yang, M.Q.; Uversky, V.N.; Dunker, A.K. Flexible nets: Disorder and induced fit in the associations of p53 and 14-3-3 with their partners. *BMC Genom.* **2008**, *9* (Suppl. 1), S1. [[CrossRef](#)]
241. Lattman, E.E.; Rose, G.D. Protein folding—what’s the question? *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 439–441. [[CrossRef](#)]
242. Rackovsky, S. On the nature of the protein folding code. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 644–648. [[CrossRef](#)]
243. Alexander, P.A.; Rozak, D.A.; Orban, J.; Bryan, P.N. Directed evolution of highly homologous proteins with different folds by phage display: Implications for the protein folding code. *Biochemistry* **2005**, *44*, 14045–14054. [[CrossRef](#)] [[PubMed](#)]
244. Tasayco, M.L.; Carey, J. Ordered self-assembly of polypeptide fragments to form natively like dimeric trp repressor. *Science* **1992**, *255*, 594–597. [[CrossRef](#)] [[PubMed](#)]
245. Williams, K.P.; Shoelson, S.E. Cooperative self-assembly of SH2 domain fragments restores phosphopeptide binding. *Biochemistry* **1993**, *32*, 11279–11284. [[CrossRef](#)] [[PubMed](#)]
246. Betton, J.M.; Hofnung, M. In vivo assembly of active maltose binding protein from independently exported protein fragments. *Embo J.* **1994**, *13*, 1226–1234. [[CrossRef](#)]
247. Kaur, P.; Rosen, B.P. In vitro assembly of an anion-stimulated ATPase from peptide fragments. *J. Biol. Chem.* **1994**, *269*, 9698–9704.
248. Kippen, A.D.; Fersht, A.R. Analysis of the mechanism of assembly of cleaved barnase from two peptide fragments and its relevance to the folding pathway of uncleaved barnase. *Biochemistry* **1995**, *34*, 1464–1468. [[CrossRef](#)]
249. Ridge, K.D.; Lee, S.S.; Yao, L.L. In vivo assembly of rhodopsin from expressed polypeptide fragments. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 3204–3208. [[CrossRef](#)]
250. Ridge, K.D.; Abdulaev, N.G. Folding and assembly of rhodopsin from expressed fragments. *Methods Enzymol.* **2000**, *315*, 59–70. [[CrossRef](#)]
251. Kobayashi, N.; Honda, S.; Yoshii, H.; Uedaira, H.; Munekata, E. Complement assembly of two fragments of the streptococcal protein G B1 domain in aqueous solution. *FEBS Lett.* **1995**, *366*, 99–103. [[CrossRef](#)]
252. Rochet, J.C.; Oikawa, K.; Hicks, L.D.; Kay, C.M.; Bridger, W.A.; Wolodko, W.T. Productive interactions between the two domains of pig heart CoA transferase during folding and assembly. *Biochemistry* **1997**, *36*, 8807–8820. [[CrossRef](#)]
253. Chaffotte, A.F.; Li, J.H.; Georgescu, R.E.; Goldberg, M.E.; Tasayco, M.L. Recognition between disordered states: Kinetics of the self-assembly of thioredoxin fragments. *Biochemistry* **1997**, *36*, 16040–16048. [[CrossRef](#)] [[PubMed](#)]
254. Marti, T. Refolding of bacteriorhodopsin from expressed polypeptide fragments. *J. Biol. Chem.* **1998**, *273*, 9312–9322. [[CrossRef](#)] [[PubMed](#)]
255. Martin, N.P.; Leavitt, L.M.; Sommers, C.M.; Dumont, M.E. Assembly of G protein-coupled receptors from fragments: Identification of functional receptors with discontinuities in each of the loops connecting transmembrane segments. *Biochemistry* **1999**, *38*, 682–695. [[CrossRef](#)] [[PubMed](#)]
256. Jourdan, M.; Searle, M.S. Cooperative assembly of a natively like ubiquitin structure through peptide fragment complexation: Energetics of peptide association and folding. *Biochemistry* **2000**, *39*, 12355–12364. [[CrossRef](#)]
257. Ni, X.; Schachman, H.K. In vivo assembly of aspartate transcarbamoylase from fragmented and circularly permuted catalytic polypeptide chains. *Protein Sci.* **2001**, *10*, 519–527. [[CrossRef](#)]

258. Johnsson, N.; Varshavsky, A. Split ubiquitin as a sensor of protein interactions in vivo. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 10340–10344. [[CrossRef](#)]
259. Ozawa, T. Designing split reporter proteins for analytical tools. *Anal. Chim. Acta* **2006**, *556*, 58–68. [[CrossRef](#)]
260. Remy, I.; Michnick, S.W. Clonal selection and in vivo quantitation of protein interactions with protein-fragment complementation assays. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 5394–5399. [[CrossRef](#)]
261. Blakely, B.T.; Rossi, F.M.; Tillotson, B.; Palmer, M.; Estelles, A.; Blau, H.M. Epidermal growth factor receptor dimerization monitored in live cells. *Nat. Biotechnol.* **2000**, *18*, 218–222. [[CrossRef](#)]
262. Magliery, T.J.; Wilson, C.G.; Pan, W.; Mishler, D.; Ghosh, I.; Hamilton, A.D.; Regan, L. Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: Scope and mechanism. *J. Am. Chem. Soc.* **2005**, *127*, 146–157. [[CrossRef](#)]
263. Kaihara, A.; Kawai, Y.; Sato, M.; Ozawa, T.; Umezawa, Y. Locating a protein-protein interaction in living cells via split Renilla luciferase complementation. *Anal. Chem.* **2003**, *75*, 4176–4181. [[CrossRef](#)] [[PubMed](#)]
264. Galarneau, A.; Primeau, M.; Trudeau, L.E.; Michnick, S.W. Beta-lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein protein interactions. *Nat. Biotechnol.* **2002**, *20*, 619–622. [[CrossRef](#)] [[PubMed](#)]
265. Landau, L.D.; Lifshitz, E.M. *Statistical Physics (v. 5 of A Course of Theoretical Physics)*, 3rd ed.; Elsevier: London, UK, 1980.
266. Ubbelohde, A.R. *Melting and Crystal Structure*; Clarendon Press: Oxford, UK, 1965.
267. Bagdasarov, K.S.; Givargizov, E.I.; Kuznetsov, V.A.; Demianets, L.N.; Lobachev, A.N.; Chernov, A.A. *Modern Crystallography; Crystal Growth*: Berlin/Heidelberg, Germany; New York, NY, USA; Tokyo, Japan, 1984; Volume 3.
268. Frenkel, J.D. *Kinetic Theory of Liquids*, Reissue ed.; Dover Publications: Mineola, NY, USA, 1955.
269. Wang, R.; Yang, X.; Cui, L.; Yin, H.; Xu, S. Gels of Amyloid Fibers. *Biomolecules* **2019**, *9*, 210. [[CrossRef](#)] [[PubMed](#)]
270. Stahnke, J. Hemoglobin crystals from Reichert to Hoppe-Seyler. *Sudhoffs Arch.* **1979**, *63*, 154–189.
271. Matthews, B.W. Solvent content of protein crystals. *J. Mol. Biol.* **1968**, *33*, 491–497. [[CrossRef](#)]
272. Chruszcz, M.; Potrzebowski, W.; Zimmerman, M.D.; Grabowski, M.; Zheng, H.; Lasota, P.; Minor, W. Analysis of solvent content and oligomeric states in protein crystals—Does symmetry matter? *Protein Sci.* **2008**, *17*, 623–632. [[CrossRef](#)]
273. Muschol, M.; Rosenberger, F. Liquid-liquid phase separation in supersaturated lysozyme solutions and associated precipitate formation/crystallization. *J. Chem. Phys.* **1997**, *107*, 1953–1962. [[CrossRef](#)]
274. 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](#)]
275. Neal, B.L.; Asthagiri, D.; Lenhoff, A.M. Molecular origins of osmotic second virial coefficients of proteins. *Biophys. J.* **1998**, *75*, 2469–2477. [[CrossRef](#)]
276. George, A.; Wilson, W.W. Predicting protein crystallization from a dilute solution property. *Acta Crystallogr. D Biol. Crystallogr.* **1994**, *50*, 361–365. [[CrossRef](#)]
277. Neal, B.L.; Asthagiri, D.; Velev, O.D.; Lenhoff, A.M.; Kaler, E.W. Why is the osmotic second virial coefficient related to protein crystallization? *J. Cryst. Growth* **1999**, *196*, 377–387. [[CrossRef](#)]
278. Pullara, F.; Emanuele, A.; Palma-Vittorelli, M.B.; Palma, M.U. Lysozyme crystallization rates controlled by anomalous fluctuations. *J. Cryst. Growth* **2005**, *58*, 426–438. [[CrossRef](#)]
279. Pullara, F.; Emanuele, A.; Palma-Vittorelli, M.B.; Palma, M.U. Protein aggregation/crystallization and minor structural changes: Universal versus specific aspects. *Biophys. J.* **2007**, *93*, 3271–3278. [[CrossRef](#)] [[PubMed](#)]
280. Debenedetti, P.G. *Metastable Liquids. Concepts and Principles*; Princeton University Press: Princeton, NJ, USA, 1996.
281. Scholte, T.G. Thermodynamic parameters of polymer-solvent systems from light-scattering measurements below the theta temperature. *J. Polym. Sci. B* **1971**, *9*, 1553–1577. [[CrossRef](#)]
282. Fink, A.L. Protein aggregation: Folding aggregates, inclusion bodies and amyloid. *Fold. Des.* **1998**, *3*, R9–R23. [[CrossRef](#)]
283. Mitraki, A.; King, J. Protein folding intermediates and inclusion body formation. *Biotechnology* **1989**, *7*, 690–697. [[CrossRef](#)]
284. Wetzel, R. Mutations and off-pathway aggregation of proteins. *Trends Biotechnol.* **1994**, *12*, 193–198. [[CrossRef](#)]

285. Jaenicke, R. Folding and association versus misfolding and aggregation of proteins. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **1995**, *348*, 97–105. [[CrossRef](#)]
286. Wetzel, R. For protein misassembly, it's the "I" decade. *Cell* **1996**, *86*, 699–702. [[CrossRef](#)]
287. Speed, M.A.; Morshead, T.; Wang, D.I.; King, J. Conformation of P22 tailspike folding and aggregation intermediates probed by monoclonal antibodies. *Protein Sci.* **1997**, *6*, 99–108. [[CrossRef](#)]
288. Speed, M.A.; Wang, D.I.; King, J. Specific aggregation of partially folded polypeptide chains: The molecular basis of inclusion body composition. *Nat. Biotechnol.* **1996**, *14*, 1283–1287. [[CrossRef](#)] [[PubMed](#)]
289. Wetzel, R.; Chrunchy, B.A. Inclusion body formation by interleukin-1 beta depends on the thermal sensitivity of a folding intermediate. *FEBS Lett.* **1994**, *350*, 245–248. [[CrossRef](#)]
290. Haase-Pettingell, C.A.; King, J. Formation of aggregates from a thermolabile in vivo folding intermediate in P22 tailspike maturation. A model for inclusion body formation. *J. Biol. Chem.* **1988**, *263*, 4977–4983. [[PubMed](#)]
291. Yoshimura, Y.; Lin, Y.X.; Yagi, H.; Lee, Y.H.; Kitayama, H.; Sakurai, K.; So, M.; Ogi, H.; Naiki, H.; Goto, Y. Distinguishing crystal-like amyloid fibrils and glass-like amorphous aggregates from their kinetics of formation. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 14446–14451. [[CrossRef](#)]
292. Dovidchenko, N.V.; Finkelstein, A.V.; Galzitskaya, O.V. How to determine the size of folding nuclei of protofibrils from the concentration dependence of the rate and lag-time of aggregation. I. Modeling the amyloid protofibril formation. *J. Phys. Chem. B* **2014**, *118*, 1189–1197. [[CrossRef](#)]
293. Finkelstein, A.V.; Dovidchenko, N.V.; Galzitskaya, O.V. What is Responsible for Atypical Dependence of the Rate of Amyloid Formation on Protein Concentration: Fibril-Catalyzed Initiation of New Fibrils or Competition with Oligomers? *J. Phys. Chem. Lett.* **2018**, *9*, 1002–1006. [[CrossRef](#)]
294. Come, J.H.; Fraser, P.E.; Lansbury, P.T., Jr. A kinetic model for amyloid formation in the prion diseases: Importance of seeding. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 5959–5963. [[CrossRef](#)]
295. Jarrett, J.T.; Lansbury, P.T., Jr. Seeding "one-dimensional crystallization" of amyloid: A pathogenic mechanism in Alzheimer's disease and scrapie? *Cell* **1993**, *73*, 1055–1058. [[CrossRef](#)]
296. Crespo, R.; Rocha, F.A.; Damas, A.M.; Martins, P.M. A generic crystallization-like model that describes the kinetics of amyloid fibril formation. *J. Biol. Chem.* **2012**, *287*, 30585–30594. [[CrossRef](#)]
297. Jarrett, J.T.; Lansbury, P.T., Jr. Amyloid fibril formation requires a chemically discriminating nucleation event: Studies of an amyloidogenic sequence from the bacterial protein OsmB. *Biochemistry* **1992**, *31*, 12345–12352. [[CrossRef](#)]
298. Bergfors, T. Seeds to crystals. *J. Struct. Biol.* **2003**, *142*, 66–76. [[CrossRef](#)]
299. Dima, R.I.; Thirumalai, D. Exploring protein aggregation and self-propagation using lattice models: Phase diagram and kinetics. *Protein Sci.* **2002**, *11*, 1036–1049. [[CrossRef](#)] [[PubMed](#)]
300. Kuroda, Y. Biophysical studies of protein solubility and amorphous aggregation by systematic mutational analysis and a helical polymerization model. *Biophys. Rev.* **2018**, *10*, 473–480. [[CrossRef](#)] [[PubMed](#)]
301. Boatz, J.C.; Whitley, M.J.; Li, M.; Gronenborn, A.M.; van der Wel, P.C.A. Cataract-associated P23T gammaD-crystallin retains a native-like fold in amorphous-looking aggregates formed at physiological pH. *Nat. Commun.* **2017**, *8*, 15137. [[CrossRef](#)]
302. Kato, A.; Maki, K.; Ebina, T.; Kuwajima, K.; Soda, K.; Kuroda, Y. Mutational analysis of protein solubility enhancement using short peptide tags. *Biopolymers* **2007**, *85*, 12–18. [[CrossRef](#)]
303. Hirota, N.; Edskes, H.; Hall, D. Unified theoretical description of the kinetics of protein aggregation. *Biophys. Rev.* **2019**, *11*, 191–208. [[CrossRef](#)]
304. Hirota, N.; Hall, D. Protein Aggregation Kinetics: A Unified Theoretical Description. In *Protein Solubility and Amorphous Aggregation: From Academic Research to Applications in Drug Discovery and Bioindustry*; Kuroda, Y., Arisaka, F., Eds.; CMC Publishing Co., Ltd.: Tokyo, Japan, 2019.
305. Kelly, J.W. Amyloid fibril formation and protein misassembly: A structural quest for insights into amyloid and prion diseases. *Structure* **1997**, *5*, 595–600. [[CrossRef](#)]
306. Kelly, J.W. The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Curr. Opin. Struct. Biol.* **1998**, *8*, 101–106. [[CrossRef](#)]
307. Dobson, C.M. Protein misfolding, evolution and disease. *Trends Biochem. Sci.* **1999**, *24*, 329–332. [[CrossRef](#)]
308. Bellotti, V.; Mangione, P.; Stoppini, M. Biological activity and pathological implications of misfolded proteins. *Cell. Mol. Life Sci.* **1999**, *55*, 977–991. [[CrossRef](#)]

309. Uversky, V.N.; Talapatra, A.; Gillespie, J.R.; Fink, A.L. Protein deposits as the molecular basis of amyloidosis. I. Systemic amyloidosis. *Med. Sci. Monit.* **1999**, *5*, 1001–1012.
310. Uversky, V.N.; Talapatra, A.; Gillespie, J.R.; Fink, A.L. Protein deposits as the molecular basis of amyloidosis. II. Localized amyloidosis and neurodegenerative disorders. *Med. Sci. Monit.* **1999**, *5*, 1238–1254.
311. Uversky, V.N.; Fink, A.L. Conformational constraints for amyloid fibrillation: The importance of being unfolded. *Biochim. Biophys. Acta* **2004**, *1698*, 131–153. [[CrossRef](#)]
312. Rochet, J.C.; Lansbury, P.T., Jr. Amyloid fibrillogenesis: Themes and variations. *Curr. Opin. Struct. Biol.* **2000**, *10*, 60–68. [[CrossRef](#)]
313. Sunde, M.; Serpell, L.C.; Bartlam, M.; Fraser, P.E.; Pepys, M.B.; Blake, C.C. Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol.* **1997**, *273*, 729–739. [[CrossRef](#)]
314. Chiti, F.; Dobson, C.M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **2006**, *75*, 333–366. [[CrossRef](#)] [[PubMed](#)]
315. Fandrich, M.; Fletcher, M.A.; Dobson, C.M. Amyloid fibrils from muscle myoglobin. *Nature* **2001**, *410*, 165–166. [[CrossRef](#)]
316. Dobson, C.M. Protein folding and its links with human disease. *Biochem. Soc. Symp.* **2001**, *68*, 1–26.
317. Uversky, V.N.; Li, J.; Fink, A.L. Evidence for a partially folded intermediate in alpha-synuclein fibril formation. *J. Biol. Chem.* **2001**, *276*, 10737–10744. [[CrossRef](#)]
318. Goers, J.; Permyakov, S.E.; Permyakov, E.A.; Uversky, V.N.; Fink, A.L. Conformational prerequisites for alpha-lactalbumin fibrillation. *Biochemistry* **2002**, *41*, 12546–12551. [[CrossRef](#)]
319. Munishkina, L.A.; Fink, A.L.; Uversky, V.N. Conformational prerequisites for formation of amyloid fibrils from histones. *J. Mol. Biol.* **2004**, *342*, 1305–1324. [[CrossRef](#)] [[PubMed](#)]
320. Dobson, C.M. Getting out of shape. *Nature* **2002**, *418*, 729–730. [[CrossRef](#)] [[PubMed](#)]
321. Dobson, C.M. Protein folding and misfolding. *Nature* **2003**, *426*, 884–890. [[CrossRef](#)] [[PubMed](#)]
322. Dobson, C.M. Protein aggregation and its consequences for human disease. *Protein Pept. Lett.* **2006**, *13*, 219–227. [[CrossRef](#)] [[PubMed](#)]
323. Richardson, J.S.; Richardson, D.C. Natural beta-sheet proteins use negative design to avoid edge-to-edge aggregation. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 2754–2759. [[CrossRef](#)] [[PubMed](#)]
324. Kelly, J.W. Alternative conformations of amyloidogenic proteins govern their behavior. *Curr. Opin. Struct. Biol.* **1996**, *6*, 11–17. [[CrossRef](#)]
325. Lansbury, P.T., Jr. Evolution of amyloid: What normal protein folding may tell us about fibrillogenesis and disease. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 3342–3344. [[CrossRef](#)]
326. Canet, D.; Sunde, M.; Last, A.M.; Miranker, A.; Spencer, A.; Robinson, C.V.; Dobson, C.M. Mechanistic studies of the folding of human lysozyme and the origin of amyloidogenic behavior in its disease-related variants. *Biochemistry* **1999**, *38*, 6419–6427. [[CrossRef](#)]
327. Zerovnik, E. Amyloid-fibril formation. Proposed mechanisms and relevance to conformational disease. *Eur. J. Biochem.* **2002**, *269*, 3362–3371.
328. Teplow, D.B. Structural and kinetic features of amyloid beta-protein fibrillogenesis. *Amyloid Int. J. Exp. Clin. Investig. Off. J. Int. Soc. Amyloidosis* **1998**, *5*, 121–142. [[CrossRef](#)]
329. Schweers, O.; Schonbrunn-Hanebeck, E.; Marx, A.; Mandelkow, E. Structural studies of tau protein and Alzheimer paired helical filaments show no evidence for beta-structure. *J. Biol. Chem.* **1994**, *269*, 24290–24297. [[PubMed](#)]
330. Kaye, R.; Bernhagen, J.; Greenfield, N.; Sweimeh, K.; Brunner, H.; Voelter, W.; Kapurniotu, A. Conformational transitions of islet amyloid polypeptide (IAPP) in amyloid formation in vitro. *J. Mol. Biol.* **1999**, *287*, 781–796. [[CrossRef](#)] [[PubMed](#)]
331. Englander, S.W.; Kallenbach, N.R. Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys.* **1983**, *16*, 521–655. [[CrossRef](#)] [[PubMed](#)]
332. Chamberlain, A.K.; Handel, T.M.; Marqusee, S. Detection of rare partially folded molecules in equilibrium with the native conformation of RNaseH. *Nat. Struct. Biol.* **1996**, *3*, 782–787. [[CrossRef](#)]
333. Lashuel, H.A.; Wurth, C.; Woo, L.; Kelly, J.W. The most pathogenic transthyretin variant, L55P, forms amyloid fibrils under acidic conditions and protofilaments under physiological conditions. *Biochemistry* **1999**, *38*, 13560–13573. [[CrossRef](#)]
334. Wetzel, R. Domain stability in immunoglobulin light chain deposition disorders. *Adv. Protein Chem.* **1997**, *50*, 183–242.

335. Saraiva, M.J. Transthyretin amyloidosis: A tale of weak interactions. *FEBS Lett.* **2001**, *498*, 201–203. [[CrossRef](#)]
336. Nielsen, L.; Frokjaer, S.; Brange, J.; Uversky, V.N.; Fink, A.L. Probing the mechanism of insulin fibril formation with insulin mutants. *Biochemistry* **2001**, *40*, 8397–8409. [[CrossRef](#)]
337. Heegaard, N.H.; Sen, J.W.; Kaarsholm, N.C.; Nissen, M.H. Conformational intermediate of the amyloidogenic protein beta 2-microglobulin at neutral pH. *J. Biol. Chem.* **2001**, *276*, 32657–32662. [[CrossRef](#)]
338. Nielsen, L.; Khurana, R.; Coats, A.; Frokjaer, S.; Brange, J.; Vyas, S.; Uversky, V.N.; Fink, A.L. Effect of environmental factors on the kinetics of insulin fibril formation: Elucidation of the molecular mechanism. *Biochemistry* **2001**, *40*, 6036–6046. [[CrossRef](#)]
339. Peterson, S.A.; Klabunde, T.; Lashuel, H.A.; Purkey, H.; Sacchettini, J.C.; Kelly, J.W. Inhibiting transthyretin conformational changes that lead to amyloid fibril formation. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12956–12960. [[CrossRef](#)]
340. Baures, P.W.; Peterson, S.A.; Kelly, J.W. Discovering transthyretin amyloid fibril inhibitors by limited screening. *Bioorgan. Med. Chem.* **1998**, *6*, 1389–1401. [[CrossRef](#)]
341. Oza, V.B.; Petrassi, H.M.; Purkey, H.E.; Kelly, J.W. Synthesis and evaluation of anthranilic acid-based transthyretin amyloid fibril inhibitors. *Bioorgan. Med. Chem. Lett.* **1999**, *9*, 1–6. [[CrossRef](#)]
342. Klabunde, T.; Petrassi, H.M.; Oza, V.B.; Raman, P.; Kelly, J.W.; Sacchettini, J.C. Rational design of potent human transthyretin amyloid disease inhibitors. *Nat. Struct. Biol.* **2000**, *7*, 312–321. [[CrossRef](#)]
343. Chiti, F.; Taddei, N.; Stefani, M.; Dobson, C.M.; Ramponi, G. Reduction of the amyloidogenicity of a protein by specific binding of ligands to the native conformation. *Protein Sci.* **2001**, *10*, 879–886. [[CrossRef](#)]
344. McCammon, M.G.; Scott, D.J.; Keetch, C.A.; Greene, L.H.; Purkey, H.E.; Petrassi, H.M.; Kelly, J.W.; Robinson, C.V. Screening transthyretin amyloid fibril inhibitors: Characterization of novel multiprotein, multiligand complexes by mass spectrometry. *Structure* **2002**, *10*, 851–863. [[CrossRef](#)]
345. Oza, V.B.; Smith, C.; Raman, P.; Koepf, E.K.; Lashuel, H.A.; Petrassi, H.M.; Chiang, K.P.; Powers, E.T.; Sacchettini, J.; Kelly, J.W. Synthesis, structure, and activity of diclofenac analogues as transthyretin amyloid fibril formation inhibitors. *J. Med. Chem.* **2002**, *45*, 321–332. [[CrossRef](#)]
346. Raghu, P.; Reddy, G.B.; Sivakumar, B. Inhibition of transthyretin amyloid fibril formation by 2,4-dinitrophenol through tetramer stabilization. *Arch. Biochem. Biophys.* **2002**, *400*, 43–47. [[CrossRef](#)]
347. Uversky, V.N. Under-folded proteins: Conformational ensembles and their roles in protein folding, function, and pathogenesis. *Biopolymers* **2013**, *99*, 870–887. [[CrossRef](#)]
348. Chirita, C.N.; Congdon, E.E.; Yin, H.; Kuret, J. Triggers of full-length tau aggregation: A role for partially folded intermediates. *Biochemistry* **2005**, *44*, 5862–5872. [[CrossRef](#)]
349. Yamin, G.; Munishkina, L.A.; Karymov, M.A.; Lyubchenko, Y.L.; Uversky, V.N.; Fink, A.L. Forcing nonamyloidogenic beta-synuclein to fibrillate. *Biochemistry* **2005**, *44*, 9096–9107. [[CrossRef](#)] [[PubMed](#)]
350. Uversky, V.N.; Li, J.; Souillac, P.; Millett, I.S.; Doniach, S.; Jakes, R.; Goedert, M.; Fink, A.L. Biophysical properties of the synucleins and their propensities to fibrillate: Inhibition of alpha-synuclein assembly by beta- and gamma-synucleins. *J. Biol. Chem.* **2002**, *277*, 11970–11978. [[CrossRef](#)]
351. Goldsbury, C.; Goldie, K.; Pellaud, J.; Seelig, J.; Frey, P.; Muller, S.A.; Kistler, J.; Cooper, G.J.; Aebi, U. Amyloid fibril formation from full-length and fragments of amylin. *J. Struct. Biol.* **2000**, *130*, 352–362. [[CrossRef](#)] [[PubMed](#)]
352. Pavlov, N.A.; Cherny, D.I.; Heim, G.; Jovin, T.M.; Subramaniam, V. Amyloid fibrils from the mammalian protein prothymosin alpha. *FEBS Lett.* **2002**, *517*, 37–40. [[CrossRef](#)]
353. Uversky, V.N. Mysterious oligomerization of the amyloidogenic proteins. *FEBS J.* **2010**, *277*, 2940–2953. [[CrossRef](#)]
354. Merlini, G.; Bellotti, V. Molecular mechanisms of amyloidosis. *N. Engl. J. Med.* **2003**, *349*, 583–596. [[CrossRef](#)]
355. Smith, D.P.; Jones, S.; Serpell, L.C.; Sunde, M.; Radford, S.E. A systematic investigation into the effect of protein destabilisation on beta 2-microglobulin amyloid formation. *J. Mol. Biol.* **2003**, *330*, 943–954. [[CrossRef](#)]
356. Wagner, R. Einige Bemerkungen und Fragen über das Keimbläschen (vesicular germinativa). *Müllers Arch. Anat. Physiol. Wiss. Med.* **1835**, *268*, 373–377.
357. Valentin, G.G. *Repertorium für Anatomie und Physiologie*; Verlag von Veit und Comp.: Berlin, Germany, 1836.
358. Toretzky, J.A.; Wright, P.E. Assemblages: Functional units formed by cellular phase separation. *J. Cell Biol.* **2014**, *206*, 579–588. [[CrossRef](#)]

359. Hyman, A.A.; Weber, C.A.; Julicher, F. Liquid-liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.* **2014**, *30*, 39–58. [[CrossRef](#)]
360. Forman-Kay, J.D.; Kriwacki, R.W.; Seydoux, G. Phase Separation in Biology and Disease. *J. Mol. Biol.* **2018**, *430*, 4603–4606. [[CrossRef](#)] [[PubMed](#)]
361. Shin, Y.; Brangwynne, C.P. Liquid phase condensation in cell physiology and disease. *Science* **2017**, *357*, eaaf4382. [[CrossRef](#)] [[PubMed](#)]
362. Alberti, S.; Dormann, D. Liquid-Liquid Phase Separation in Disease. *Annu. Rev. Genet.* **2019**, *53*, 171–194. [[CrossRef](#)] [[PubMed](#)]
363. Alberti, S.; Gladfelter, A.; Mittag, T. Considerations and Challenges in Studying Liquid-Liquid Phase Separation and Biomolecular Condensates. *Cell* **2019**, *176*, 419–434. [[CrossRef](#)]
364. Gomes, E.; Shorter, J. The molecular language of membraneless organelles. *J. Biol. Chem.* **2019**, *294*, 7115–7127. [[CrossRef](#)] [[PubMed](#)]
365. Hnisz, D.; Shrinivas, K.; Young, R.A.; Chakraborty, A.K.; Sharp, P.A. A Phase Separation Model for Transcriptional Control. *Cell* **2017**, *169*, 13–23. [[CrossRef](#)] [[PubMed](#)]
366. 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](#)]
367. Alberti, S. The wisdom of crowds: Regulating cell function through condensed states of living matter. *J. Cell. Sci.* **2017**, *130*, 2789–2796. [[CrossRef](#)]
368. Saha, S.; Weber, C.A.; Nusch, M.; Adame-Arana, O.; Hoegge, C.; Hein, M.Y.; Osborne-Nishimura, E.; Mahamid, J.; Janel, M.; Jawerth, L.; et al. Polar Positioning of Phase-Separated Liquid Compartments in Cells Regulated by an mRNA Competition Mechanism. *Cell* **2016**, *166*, 1572–1584. [[CrossRef](#)]
369. Courchaine, E.M.; Lu, A.; Neugebauer, K.M. Droplet organelles? *Embo J.* **2016**, *35*, 1603–1612. [[CrossRef](#)]
370. Chong, P.A.; Forman-Kay, J.D. Liquid-liquid phase separation in cellular signaling systems. *Curr. Opin. Struct. Biol.* **2016**, *41*, 180–186. [[CrossRef](#)] [[PubMed](#)]
371. Alberti, S.; Hyman, A.A. Are aberrant phase transitions a driver of cellular aging? *Bioessays* **2016**, *38*, 959–968. [[CrossRef](#)] [[PubMed](#)]
372. Uversky, V.N. Supramolecular fuzziness of intracellular liquid droplets: Liquid-liquid phase transitions, membrane-less organelles, and intrinsic disorder. *Molecules* **2019**, *24*, 3265. [[CrossRef](#)] [[PubMed](#)]
373. 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](#)]
374. Uversky, V.N. Protein intrinsic disorder-based liquid-liquid phase transitions in biological systems: Complex coacervates and membrane-less organelles. *Adv. Colloid Interface Sci.* **2017**, *239*, 97–114. [[CrossRef](#)]
375. Handwerker, K.E.; Cordero, J.A.; Gall, J.G. Cajal bodies, nucleoli, and speckles in the *Xenopus* oocyte nucleus have a low-density, sponge-like structure. *Mol. Biol. Cell* **2005**, *16*, 202–211. [[CrossRef](#)]
376. Brangwynne, C.P. Phase transitions and size scaling of membrane-less organelles. *J. Cell Biol.* **2013**, *203*, 875–881. [[CrossRef](#)]
377. Sokolova, E.; Spruijt, E.; Hansen, M.M.; Dubuc, E.; Groen, J.; Chokkalingam, V.; Piruska, A.; Heus, H.A.; Huck, W.T. Enhanced transcription rates in membrane-free protocells formed by coacervation of cell lysate. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 11692–11697. [[CrossRef](#)]
378. Strulson, C.A.; Molden, R.C.; Keating, C.D.; Bevilacqua, P.C. RNA catalysis through compartmentalization. *Nat. Chem.* **2012**, *4*, 941–946. [[CrossRef](#)]
379. Nikolic, J.; Lagaudriere-Gesbert, C.; Scrima, N.; Blondel, D.; Gaudin, Y. Structure and Function of Negri Bodies. *Adv. Exp. Med. Biol.* **2019**, *1140*, 111–127. [[CrossRef](#)]
380. Formicola, N.; Vijayakumar, J.; Besse, F. Neuronal ribonucleoprotein granules: Dynamic sensors of localized signals. *Traffic* **2019**, *20*, 639–649. [[CrossRef](#)] [[PubMed](#)]
381. Dundr, M.; Misteli, T. Biogenesis of nuclear bodies. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a000711. [[CrossRef](#)] [[PubMed](#)]
382. Mao, Y.S.; Zhang, B.; Spector, D.L. Biogenesis and function of nuclear bodies. *Trends Genet.* **2011**, *27*, 295–306. [[CrossRef](#)] [[PubMed](#)]
383. Phair, R.D.; Misteli, T. High mobility of proteins in the mammalian cell nucleus. *Nature* **2000**, *404*, 604–609. [[CrossRef](#)]
384. Pederson, T. Protein mobility within the nucleus—What are the right moves? *Cell* **2001**, *104*, 635–638. [[CrossRef](#)]

385. Brangwynne, C.P.; Eckmann, C.R.; Courson, D.S.; Rybarska, A.; Hoege, C.; Gharakhani, J.; Julicher, F.; Hyman, A.A. Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science* **2009**, *324*, 1729–1732. [[CrossRef](#)]
386. Brangwynne, C.P.; Mitchison, T.J.; Hyman, A.A. Active liquid-like behavior of nucleoli determines their size and shape in *Xenopus laevis* oocytes. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 4334–4339. [[CrossRef](#)]
387. Feric, M.; Brangwynne, C.P. A nuclear F-actin scaffold stabilizes ribonucleoprotein droplets against gravity in large cells. *Nat. Cell Biol.* **2013**, *15*, 1253–1259. [[CrossRef](#)]
388. Wippich, F.; Bodenmiller, B.; Trajkovska, M.G.; Wanka, S.; Aebersold, R.; Pelkmans, L. Dual specificity kinase DYRK3 couples stress granule condensation/dissolution to mTORC1 signaling. *Cell* **2013**, *152*, 791–805. [[CrossRef](#)]
389. Uversky, V.N.; Kuznetsova, I.M.; Turoverov, K.K.; Zaslavsky, B. Intrinsically disordered proteins as crucial constituents of cellular aqueous two phase systems and coacervates. *FEBS Lett.* **2015**, *589*, 15–22. [[CrossRef](#)]
390. Zhu, L.; Brangwynne, C.P. Nuclear bodies: The emerging biophysics of nucleoplasmic phases. *Curr. Opin. Cell Biol.* **2015**, *34*, 23–30. [[CrossRef](#)] [[PubMed](#)]
391. Brangwynne, C.P.; Tompa, P.; Pappu, R.V. Polymer physics of intracellular phase transitions. *Nat. Phys.* **2015**, *11*, 899–904. [[CrossRef](#)]
392. Holehouse, A.S.; Pappu, R.V. Functional Implications of Intracellular Phase Transitions. *Biochemistry* **2018**, *57*, 2415–2423. [[CrossRef](#)] [[PubMed](#)]
393. Uversky, V.N. Intrinsically disordered proteins in overcrowded milieu: Membrane-less organelles, phase separation, and intrinsic disorder. *Curr. Opin. Struct. Biol.* **2017**, *44*, 18–30. [[CrossRef](#)] [[PubMed](#)]
394. Darling, A.L.; Liu, Y.; Oldfield, C.J.; Uversky, V.N. Intrinsically Disordered Proteome of Human Membrane-Less Organelles. *Proteomics* **2018**, *18*, e1700193. [[CrossRef](#)] [[PubMed](#)]
395. Nott, T.J.; Petsalaki, E.; Farber, P.; Jervis, D.; Fussner, E.; Plochowitz, A.; Craggs, T.D.; Bazett-Jones, D.P.; Pawson, T.; Forman-Kay, J.D.; et al. Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles. *Mol. Cell* **2015**, *57*, 936–947. [[CrossRef](#)] [[PubMed](#)]
396. Darling, A.L.; Zaslavsky, B.Y.; Uversky, V.N. Intrinsic Disorder-Based Emergence in Cellular Biology: Physiological and Pathological Liquid-Liquid Phase Transitions in Cells. *Polymers* **2019**, *11*, 990. [[CrossRef](#)]
397. Turoverov, K.K.; Kuznetsova, I.M.; Fonin, A.V.; Darling, A.L.; Zaslavsky, B.Y.; Uversky, V.N. Stochasticity of Biological Soft Matter: Emerging Concepts in Intrinsically Disordered Proteins and Biological Phase Separation. *Trends Biochem. Sci.* **2019**, *44*, 716–728. [[CrossRef](#)]
398. Ambadipudi, S.; Biernat, J.; Riedel, D.; Mandelkow, E.; Zweckstetter, M. Liquid-liquid phase separation of the microtubule-binding repeats of the Alzheimer-related protein Tau. *Nat. Commun.* **2017**, *8*, 275. [[CrossRef](#)]
399. Ambadipudi, S.; Reddy, J.G.; Biernat, J.; Mandelkow, E.; Zweckstetter, M. Residue-specific identification of phase separation hot spots of Alzheimer’s-related protein tau. *Chem. Sci.* **2019**, *10*, 6503–6507. [[CrossRef](#)]
400. Reichheld, S.E.; Muiznieks, L.D.; Keeley, F.W.; Sharpe, S. Direct observation of structure and dynamics during phase separation of an elastomeric protein. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E4408–E4415. [[CrossRef](#)] [[PubMed](#)]
401. Murthy, A.C.; Dignon, G.L.; Kan, Y.; Zerze, G.H.; Parekh, S.H.; Mittal, J.; Fawzi, N.L. Molecular interactions underlying liquid-liquid phase separation of the FUS low-complexity domain. *Nat. Struct. Mol. Biol.* **2019**, *26*, 637–648. [[CrossRef](#)] [[PubMed](#)]
402. Ryan, V.H.; Dignon, G.L.; Zerze, G.H.; Chabata, C.V.; Silva, R.; Conicella, A.E.; Amaya, J.; Burke, K.A.; Mittal, J.; Fawzi, N.L. Mechanistic View of hnRNPA2 Low-Complexity Domain Structure, Interactions, and Phase Separation Altered by Mutation and Arginine Methylation. *Mol. Cell* **2018**, *69*, 465–479. [[CrossRef](#)] [[PubMed](#)]

