Published in final edited form as:

Prog Biophys Mol Biol. 2010 ; 102(2-3): 73–84. doi:10.1016/j.pbiomolbio.2010.01.003.

The Protein Kingdom Extended: Ordered and Intrinsically Disordered Proteins, Their Folding, Supramolecular Complex Formation, and Aggregation

Konstantin K. Turoverov^{1,*}, Irina M. Kuznetsova¹, and Vladimir N. Uversky^{2,3,*}

¹Laboratory of structural dynamics, stability and folding of proteins, Institute of cytology RAS, S.Petersburg, 194064, Tikhoretsky ave, 4

²Institute for Intrinsically Disordered Protein Research, Center for Computational Biology and Bioinformatics, Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN 46202, USA

³Institute for Biological Instrumentation, Russian Academy of Sciences, 142290 Pushchino, Moscow Region, Russia

Abstract

The native state of a protein is usually associated with a compact globular conformation possessing a rigid and highly ordered structure. At the turn of the last century certain studies arose which concluded that many proteins cannot, in principle, form a rigid globular structure in an aqueous environment, but they are still able to fulfill their specific functions — i.e., they are native. The existence of the disordered regions allows these proteins to interact with their numerous binding partners. Such interactions are often accompanied by the formation of complexes that possess a more ordered structure than the original components. The functional diversity of these proteins, combined with the variability of signals related to the various intra- and intercellular processes handled by these proteins and their capability to produce multi-variant and multi-directional responses allow them to form a unique regulatory net in a cell. The abundance of disordered proteins inside the cell is precisely controlled at the synthesis and clearance levels as well as via interaction with specific binding partners and posttranslational modifications. Another recently recognized biologically active state of proteins is the functional amyloid. The formation of such functional amyloids is tightly controlled and therefore differs from the uncontrolled formation of pathogenic amyloids which are associated with the pathogenesis of several conformational diseases, the development of which is likely to be determined by the failures of the cellular regulatory systems rather than by the formation of the proteinaceous deposits and/or by the protofibril toxicity.

© 2010 Elsevier Ltd. All rights reserved.

*Corresponding authors: Konstantin K. Turoverov, Ph.D., D.Sc. Laboratory of structural dynamics, stability and folding of proteins, Institute of cytology RAS, S.Petersburg, 194064, Tikhoretsky ave., 4, Russia. kkt@mail.cytspb.rssi.ru; Vladimir N. Uversky, Ph.D., D.Sc. Center for Computational Biology and Bioinformatics, Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN 46202, USA. vuversky@iupui.edu.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

protein folding; globular proteins; natively disordered proteins; protein-protein and DNA-protein complexes; amorphous aggregates; amyloid fibrils; functional amyloid; inter- and intramolecular contacts

Introduction

The understanding of what a native protein is, how and why it is formed, and how the processes of protein denaturation and renaturation occur has changed dramatically over the past 50 years. For a long time it was postulated that protein denaturation is a reversible "all-or-none" transition between the native and denatured states. The denatured, biologically inactive form was considered to be equivalent to a random coil state of a synthetic polymer. The native state was viewed as a compact globular form with a specific, rigid, and well-defined structure. This structural rigidity determined the ability of many proteins to form crystals, which allowed the X-ray-based determination of 3D structure of many proteins down to the atomic resolution (Berman et al., 2000).

Even in these highly ordered structures, all the atoms were surely subjected to thermal motions of different amplitudes, where atoms located at the chain termini, in the loop regions, or in active sites of enzyme were more mobile than other atoms. Although a significant functional role of these thermal oscillations for protein function was proposed (Varshavsky, 1979), they were assumed to happen around equilibrium position, have relatively small amplitudes, and preserve the integrity of a protein globule. Numerous exceptions of this solid-state view of a functional protein were observed for a number of years, but were mostly ignored.

The situation changed at the turn of the century when it was systematically shown that many partially or completely disordered proteins are functional [e.g., see (Dunker et al., 2001; Tompa, 2002; Uversky et al., 2000; Wright and Dyson, 1999)] for early reviews on this topic). Some of these proteins become ordered in the complex with their binding partners, whereas other remain substantially disordered during their action (Dunker et al., 2008a). As the main criterion of a native protein is its ability to perform a biological function, these partially or completely disordered proteins must be regarded as native entities.

Before the definition of the native protein was changed so drastically, the mechanisms of protein folding-unfolding processes underwent dramatic reevaluation themselves. The experimental data, accumulated by the end of the 1960th, consistently showed that the transition between the native and the completely unfolded states of a globular protein did not always follow the "all-or-none" principle, as stable equilibrium or kinetic intermediates were observed (Tanford, 1968). One of the folding intermediates, now known as the molten globule (Ptitsyn, 1995), was structurally characterized in early 1980th, and later it became evident that there can be several partially folded intermediates (Uversky, 2003). Some of these intermediates were located on the folding pathway, as they had the structural elements inherent to the final folded state. Some other intermediates were clearly located outside the folding pathway, representing a kind of folding trap (Hamada et al., 1996).

Many biologically active proteins act as specific oligomers, while non-specific aggregation was commonly viewed as an experimental artifact hampering the detailed structural characterization of proteins. Since the pathogenesis of many known grave diseases was related to protein misfolding and aggregation [see e.g. (Dobson, 1999)], the research

priorities have changed, generating the burst of studies on protein misfolding and non-specific aggregation.

In this work, we represent a unified model for the existence of native globular and intrinsically disordered proteins (IDPs), the formation of the intermolecular complexes, and nonspecific aggregation. This unified model is based on the folding energy landscape where:

- i. For a globular protein, the global minimum of the free energy determines a thermodynamically stable state that corresponds to the compact and unique 3D structure stabilized by strong intramolecular interactions.
- ii. For an IDP, the energy landscape is characterized by the lack of such a global minimum and by the presence of multiple local minima separated by low energy barriers. As a consequence, such a protein exists as a dynamic ensemble of interconverting structures.
- iii. Functional complexes of IDPs with their partners are formed via the specific intermolecular interactions that change the energy landscape creating more defined free energy minima.
- iv. A similar mechanism underlies the formation of the non-productive protein complexes (oligomers, amorphous aggregates, amyloid-like fibrils, etc.). Here, interactions of the disordered parts of IDPs or interactions of the denatured proteins result in the appearance of profound free energy minima.

Folding of globular proteins

Folding code

In protein biosynthesis, the information encoded in the DNA/mRNA nucleotide sequence is read step-by-step, and the corresponding amino acids are gathered one after another into the polypeptide chain. Therefore, the one-dimensional information encoded in the DNA nucleotide sequence is sequentially transformed into the one-dimensional information of the protein amino acid sequence. As the interactions between remote amino acid residues play a crucial role in protein folding, this process obviously deviates from the linear information transduction. Only some amino acid residues are crucial for protein folding. Therefore, proteins with very low sequence homology can have similar structures, whereas a single amino acid replacement can significantly affect the rate of protein folding, or in some extreme cases, can completely halt the correct protein folding.

Many proteins have rigid globular structures in aqueous solutions and are functional only in this state. The native state of these proteins is a unique conformation, which is entropically unfavorable since it has significant restrictions of the conformational freedom. On the other hand, the unfolded state of a polypeptide chain is entropically favorable, representing a dynamic ensemble of a large number of conformations originating from the main chain rotational isomerization around Φ and Ψ angles. Therefore, the possibility of a given polypeptide chain to fold into a compact state is determined by its ability to form numerous intramolecular contacts of different physical nature, in order to compensate the free energy increase due to the decrease in the entropy component (Finkelstein and Ptitsyn, 2002). Although a native state of a globular protein has a clearly-defined and unique 3D structure, the folding and ordering degree can greatly vary for the different parts of a given protein. In X-ray data, this is seen from the *B*-factor values, which characterize the mobility of separate atoms (Berman et al., 2000). For example, the atoms in the active center of an enzyme typically have high *B*-factors. Furthermore, some globular proteins are shown to have unstructured, highly dynamic fragments (termini, loops, etc.), which could not be detected by X-ray analysis, thus corresponding to the regions of missing electron density.

The first direct evidence that all the information necessary for a given polypeptide chain to fold into a unique tertiary structure is encoded in its amino acid sequence was obtained by Anfinsen's group (Anfinsen, 1973). The reduced and urea-denatured ribonuclease A was shown to completely restore its native structure and functional state after the removal of the denaturant and the reducing agent. Later, the capability to regain the native structure *in vitro* was demonstrated for a variety of proteins. In essence, protein folding can be regarded as a second part of the genetic code, as the protein amino acid sequence contains information about its functional 3D structure.

The folding of a typical globular protein occurs at the millisecond-to-second time scale. However, for a small protein consisting of 100 amino acid residues, a simple search for a native state (which meets the requirements of the free energy minimum) among all the alternative conformations would take a billion years. This contradiction represents the essence of the Levinthal's paradox (Levinthal, 1968), which is resolved by the fact that amino acid sequences bear the information not only related to their native structures, but also to the pathways of their formation.

Some globular proteins fold into a unique globular structure only after ligand binding. The typical example of such proteins is a globular actin, which denature after the removal of ligands (Ca^{2+} and ATP) (Altschuler and Willison, 2008; Kuznetsova et al., 1999). It is very likely that such proteins could not attain ordered structured spontaneously, and should therefore be classified as IDPs (see below).

Protein folding models

Several models of protein folding have been developed. The "nucleation and growth" model is based on the assumption that protein folding is similar to the crystallization process, and that the limiting step in the folding process is the nucleus formation (Radford, 2000). This model describes the folding of small single-domain proteins that follow the "all-or-none" principle.

The "sequential protein folding" model, also known as the "framework" or "hierarchic" model, was proposed in 1973 by O.B. Ptitsyn (Ptitsyn, 1973). It suggests that folding starts with the backbone forming secondary structure elements, which then interact to form a more advanced folding intermediate; the specific packing of the side chains concludes the process. Each stage of the folding process stabilizes the major structural elements formed at the previous state, suggesting the existence of several folding intermediates. Therefore, long before the appearance of convincing experimental data, O.B. Ptitsyn put forward the idea of a partially folded conformation that serves as a universal folding intermediate. Eight years later, such a folding intermediate was found in a test tube (Dolgikh et al., 1981) and named the "molten globule state" (Ohgushi and Wada, 1983). Other partially folded intermediates (e.g., pre-molten globule and highly ordered molten globule) were later found (Uversky, 2003).

According to the current view, protein folding is realized via different pathways that are determined by the protein's energy landscape (Jahn and Radford, 2005; Radford, 2000). This landscape describes the dependence of the free energy on all the coordinates determining the protein conformation. The number of conformational states accessible by a polypeptide chain is reduced while approaching the native state. Therefore, this energetic surface is often called the "energy funnel" (Figure 1). The unfolded polypeptide chain's free energy represents a large "hilly plateau" describing the dynamic ensemble of a large number of conformations. Hills on the plateau correspond to the forbidden conformations, and the plateau is separated from the entrance to the folding funnel by high energetic barrier(s) corresponding to the transitional state(s) (Finkelstein and Ptitsyn, 2002). This barrier is of

great importance for the proper protein functioning, as its existence guarantees the structural identity of all the native protein molecules. The ability of native globular proteins to form crystals is the major proof of this hypothesis.

The transition from the unfolded state to the uniquely-folded native state can, in principle, be realized via different pathways. Frequently it seems that the experimental data prove the validity of this hypothesis, but in fact the rates of the various folding-unfolding stages might dramatically change depending on the nature of the denaturing conditions (or the concentration of a given denaturant). As a result, not all of the folding or unfolding steps can be registered experimentally, suggesting that the character of the unfolding pathway depends upon the choice of the denaturant (or denaturant concentration). At the same time, our experimental data on several proteins, including actin, showed that the number and the order of appearance of intermediate states is not dictated by denaturing agents (Kuznetsova et al., 2005; Povarova et al., 2007).

The energy landscape model not only elucidates the mechanisms of the globular protein folding, but also explains the nature of the IDPs, describes the formation of their supramolecular complexes (see next Section), and delineates the formation of potentially pathogenic oligomers, amorphous aggregates, and amyloid-like fibrils (see Section “Amyloid and amyloid-like fibrils”).

Intrinsically disordered proteins

Protein non-folding code

Although the idea that the behavior of functional proteins is generally described by the sequence – to structure – to function paradigm (Fischer, 1894) was believed for a long time, numerous functional proteins have been found to be either completely disordered or to contain lengthy disordered segments. These proteins were observed during the course of many decades, but were typically considered as outliers and were mostly ignored. At the turn of the century it became clear that proteins do not necessarily have to be tightly folded into specific 3D structures to be functional (for recent reviews see (Dunker et al., 2005; Dunker et al., 2001; Dyson and Wright, 2005; Iakoucheva et al., 2002; Radivojac et al., 2007; Tompa, 2002; Uversky, 2003; Uversky et al., 2000; Uversky et al., 2005; Wright and Dyson, 1999)). Such disordered functional proteins can be described by various techniques including X-ray crystallography, NMR, circular dichroism, hydrodynamic methods, proteolytic sensitivity, etc. (Receveur-Brechot et al., 2006). They are currently known by different names, including “natively denatured”, “natively unfolded”, “intrinsically unstructured”, etc., with the most frequently used term being “intrinsically disordered proteins (IDPs)”. These terms underline the intrinsic amino acid sequence-determined inability of these proteins to form ordered structures. These proteins, despite being unable to form ordered structures, are nonetheless native since they have important biological functions. Therefore, they can be also named “native disordered” proteins.

The inability of IDPs to form rigid globular structures is linked to the peculiarities of their amino acid sequences. For example, the smaller the content of hydrophobic amino acid residues and the higher the net charge of a polypeptide chain, the smaller the probability of this chain to fold into a compact globule (Uversky et al., 2000). These observations provided a basis for the charge-hydrophathy (CH-plot) algorithm for the protein foldability evaluation (Uversky et al., 2000). In comparison with typical globular proteins, IDPs contain less tryptophan, phenylalanine, tyrosine, cysteine, valine, leucine, and histidine, as they are enriched in polar and charged residues (such as lysine, arginine, glutamine, asparagine, glutamic and aspartic acid, serine and threonine) as well as in proline (Radivojac et al., 2007; Romero et al., 1997). These differences were utilized to create the first algorithm for

the per residue prediction of disorder in proteins, PONDR[®] (Li et al., 1999). More than 60 disorder predictors are known now. Major algorithms for IDP/IDR prediction together with the summary of the basic concepts of various predictors and the analysis of the strengths and shortcomings of many of the prediction methods are summarized in several recent reviews [reviewed in (Ferron et al., 2006; He et al., 2009; Radivojac et al., 2007)]. Currently, predictions of disorder are playing major roles in directing laboratory experiments. These prediction-driven experiments often discover more disordered proteins or disordered regions, thus creating a positive feedback loop in the investigation of IDPs and IDRs.

Abundance of IDPs and their functions

The application of disorder predictors revealed that IDPs are widely spread in nature, being more common in eukaryotes than in prokaryote and archae, likely due to the more complicated regulation and signaling systems in higher organisms (Dunker et al., 2000; Ward et al., 2004). The structural variability of IDPs is very high, and native coils, native pre-molten globules, and native molten globules have been described in literature (Dunker and Obradovic, 2001; Uversky, 2002a; Uversky, 2002b; Uversky, 2003). The protein can be completely unstructured, contain some elements of secondary structure, or have long disordered loops or tails. In multi-domain proteins, the domains might be connected by highly flexible linkers, and one or several domains might be completely unstructured. Because of this great variability, there is no strict boundary between globular and intrinsically disordered proteins.

IDPs cannot spontaneously fold into the compact globular structures. Some of them fold as a whole while interacting with their partners [see e.g. (Uversky, 2002a; Uversky, 2002b; Uversky, 2003; Uversky, 2009b; Uversky et al., 2000)], if the free energy of complex is lower than the free energies of IDP and its partner before their interaction (Figure 1B). More often, however, only a part of an IDP, a specific recognition element, which is a relatively short amphipathic linear motif contained within long disordered sequence, is involved in the coupled folding and binding events (Cheng et al., 2007; Dyson and Wright, 2002; Dyson and Wright, 2005; Fuxreiter et al., 2004; Lacy et al., 2004; Mohan, 2006; Oldfield et al., 2005; Vacic et al., 2007; Wright and Dyson, 1999). There are two major models describing the coupled folding and binding process in proteins, the conformational selection model and the simultaneous binding and folding model, also known as the induced folding model [see (Wright and Dyson, 2009) for an excellent recent review]. The conformational selection mechanism is based on the hypothesis that when free in solution, IDP populates the ensemble of conformations and, from this ensemble, the binding partner “chooses” or “selects” a specific conformation which closely approximates that of the bound form. An illustrative example of the conformational selection mechanism in the recognition process is so-called “preformed structural elements” or elements of local residual structure, which are frequently observed in IDPs and which are crucial for the IDP interactions with its specific partners (Fuxreiter et al., 2004). Induced folding model postulates that IDP associates with its binding partner in a fully disordered state and subsequently folds in association with the target protein. In molecular recognition, this model is exemplified by so-called molecular recognition elements or features, which are short (around 20 residues) structural elements which are found within the regions of disorder and which mediates certain classes of binding events of disordered regions undergoing a disorder-to-order transition into a specific structure that is stabilized by binding to its partner (Cheng et al., 2007; Garner et al., 1999; Mohan, 2006; Oldfield et al., 2005; Vacic et al., 2007). These recognition motifs can fold into α -helix, β -strand, or form irregular structure on binding to a target protein (Mohan, 2006; Vacic et al., 2007). Obviously, in reality, either one of the outlined above mechanisms, the conformational selection model or the induced folding model, or some combination of the two can be used (Wright and Dyson, 2009).

The intrinsic lack of structure and function-related disorder-to-order transitions provide IDPs with various functional advantages, such as: (i) Decoupled specificity and strength of binding provides for high-specificity-low-affinity interactions; (ii) Increased speed of interaction due to greater capture radius and the ability to spatially search through interaction space; (iii) Increased interaction (surface) area per residue; (iv) The ability for one-to-many and many-to-one interactions; (v) Increased capture radius for a specific binding site in comparison with that of ordered protein with its restricted conformational freedom, (vi) Fast binding kinetics, etc. The mentioned fast binding kinetics and increased capture radius represent the essence of the so-called “fly-casting mechanism” of protein binding, according to which the unfolded polypeptide first binds weakly at a relatively large distance from the actual binding site and then folds as the protein approaches the binding site (Shoemaker et al., 2000). This “fly-casting mechanism” of protein binding was recently checked in a detailed two-step analysis (Huang and Liu, 2009). First, the available experimental data on the binding kinetics of ordered proteins and of IDPs were compared to show that generally IDPs bind faster than ordered proteins. Second, coarsegrained molecular dynamics simulations of the pKID–KIX complex were performed to show that interactions of IDPs with binding partners were characterized by high on- and high off-rates resulting from the lower binding free-energy barriers. Intriguingly, contrarily to the expectations, the capture rates of IDPs were comparable to those of ordered proteins. This is because of the fact that the actual role of the greater capture radii of IDPs in the increase of the binding rates is almost completely eliminated by the slower diffusion of extended IDPs in the encounter processes. Thus, the major kinetic advantage of IDPs is their lower free-energy barriers which provide encounter complexes with a greater probability to evolve into the final bound states (Huang and Liu, 2009).

This predisposition to form complexes with specific partners represents a molecular basis for the functions of IDPs in signaling, recognition, and regulation. Although many proteins are involved in such processes, special attention was paid to hubs or network concentrators that serve as “conductors” of biological processes. Many of the hubs are IDPs (Dunker et al., 2005), including α -synuclein (Uversky, 2008), p53 (Chumakov, 2007), HMG proteins (Reeves and Beckerbauer, 2001), and others.

Interestingly, bioinformatics analyses revealed that many proteins associated with such maladies as cancer (Iakouchcheva et al., 2002), cardiovascular disease (Cheng et al., 2006), diabetes (Uversky et al., 2008), neurodegenerative diseases (Uversky, 2009a), and many others belong to the IDP family (Uversky et al., 2008). Therefore, the intensive studies on IDPs were in part stimulated by their association with various conformational diseases.

Amyloids and amyloid-like fibrils

Protein misfolding code

Protein misfolding is a wide-spread phenomenon. Any protein with changes in native structure which affect its normal function is misfolded. Aggregation and formation of amyloid-specific cross- β structure is only one of the forms of misfolding. For example, vast majority of cancer-related mutations in p53 are known to occur in its DNA-binding domain, affecting interaction of this protein with DNA (Hollstein et al., 1991; Joerger and Fersht, 2007a; Joerger and Fersht, 2007b; Joerger and Fersht, 2008). These mutations result in the dramatic destabilization of this domain in such a way that it becomes disordered at physiological conditions (Bullock et al., 1997). Therefore, the terms “misfolded” and “aggregated” are not equivalent. The ability of a protein to aggregate and to form fibrils depends on many factors, including protein sequence and environment. Typically, there is equilibrium between native (ordered or intrinsically disordered) and some aggregation-prone (typically partially folded) conformations even under physiological conditions. Normally,

this equilibrium is shifted toward the native state, and any changes in protein sequence (mutations) or environment that affect this equilibrium will inevitably contribute to the propensity to aggregate. Not all the partially unfolded conformations have increased propensity to aggregate. In fact, only a small fraction of such conformations is aggregation prone.

Protein aggregation represents a serious problem. It is estimated that 26 million people worldwide have Alzheimer's disease (AD, <http://www.ahaf.org/alzdis/about/adabout.htm>), and that about 6 million people suffer from Parkinson's disease (PD). These numbers are likely to double in twenty years. A common theme in both diseases (as well as for the number of other neurodegenerative diseases and various amyloidoses) is the occurrence of amyloid protein aggregates in the brain tissue of affected patients with neurodegenerative diseases or in any organ or tissue of patients suffering from amyloidosis. The origin and detailed mechanisms of these diseases are not understood, but a universal factor is the failure of critical proteins to fold into their correct structures and the resultant aggregation of these misfolded proteins. Protein misfolding and aggregation is not limited to those proteins involved in neurodegenerative diseases or amyloidoses. Recent work has shown that proteins and small peptides that are completely unrelated to disease also exhibit a propensity to aggregate under appropriate conditions, suggesting that the propensity to aggregate into fibrillar aggregates may be a generic feature of all polypeptide chains (Chiti et al., 1999; Fandrich et al., 2001; Goers et al., 2002; Guijarro et al., 1998; Munishkina et al., 2004; Pavlov et al., 2002). Protein aggregation is also a key issue in the food, biotechnology, and pharmaceutical industries (Frokjaer and Otzen, 2005; Wang, 2005). Although the aggregation of proteins is an important method of providing texture in food products, often in the large scale production of proteins and peptides for biotechnological or therapeutic purposes the same phenomenon is completely undesirable with significant economic consequences.

Amyloid fibril is a relatively recent discovery, but seems to be the universal state of a polypeptide chain, as the number of proteins shown to form such structures *in vitro* is constantly increasing. Amyloid fibrils of different origin have similar morphology, consisting of 2–6 unbranched protofilaments 2–5 nm in diameter associated laterally or twisted together to form fibrils with 4–13 nm diameter (Dobson, 1999) and displaying many common properties, including a core cross- β -sheet structure in which continuous β -sheets are formed with β -strands running perpendicular to the long axis of the fibrils (Serpell et al., 1997). In addition to amyloid fibrils, proteins can self-polymerize to form several other types of aggregates, e.g. soluble oligomers and amorphous aggregates. Amorphous aggregates are typically formed essentially faster than fibrils. There is no special conformational prerequisite for amorphous aggregation to occur, and many destabilized and partially unfolded proteins precipitate out of solution in a form of amorphous aggregates. On the other hand, fibrillation requires special conditions promoting formation of the specific amyloidogenic conformations (Uversky et al., 2006).

The energy landscape model suggests that some folding intermediates might have structural elements not present in the final folded state. The appearance of such misfolded intermediates might initiate protein oligomerization or aggregation. This model provides a visual (but not an overly strict) explanation of the molecular mechanisms underlying conformational diseases such as transmissible spongiform encephalopathy (TSE). Prion protein (PrP), the causative TSE agent, can exist in two forms: a normal or cellular form (PrP^C) normally expressed at low levels in neurons and other cell types, and an abnormal or scrapie form (PrP^{Sc}) built-up in diseased brain (Cohen and Prusiner, 1998). PrP^C is a cell-surface glycoprotein, the C-terminal domain of which has α -helical structure. This conformation corresponds to the first free energy minimum, and can also can be transformed

to the β -structure-rich conformation, PrP^{Sc} , which has a high tendency to aggregate and easily forms amyloid fibrils. As PrP^{Sc} corresponds to the lower free energy minimum than PrP^{C} , the transition from PrP^{C} to PrP^{Sc} is irreversible. The aggregated PrP^{Sc} is resistant to heat and to the digestion by proteases. The interaction of PrP^{C} with PrP^{Sc} amyloid fibrils transforms this soluble α -helical protein to the β -sheet structure, thus removing it from the ensemble of normal molecules. This initiates a chain reaction which leads to the accumulation of enormous amounts of PrP^{Sc} to levels that are not manageable by the cell defense systems, which finally results in the brain tissue damage.

Recently, an atomic level analysis of fibril-like structures of segments from proteins known to fibrillate revealed that these amyloid-like structures consisted of pairs of tightly packed, highly complementary β -sheets, termed "steric zippers", each including short segments of protein molecules stacked into β -sheets that run the entire length of the amyloid-like fibrils (Nelson et al., 2005; Sawaya et al., 2007; Wiltzius et al., 2009). Detailed analysis of crystal structures of steric zippers formed by segments of prion and other amyloid proteins provided a potential explanation for the intriguing phenomenon of prion strains (i.e., phenotypic variants encoded by protein conformations) (Wiltzius et al., 2009), in which structural conversions of the same protein give rise to different disease characteristics or phenotypes (Chien et al., 2004). In fact, two different types of packing polymorphism were discovered, the registration polymorphism based on the alternative packing arrangements of the β -sheets of the same segment of a protein, and the segmental polymorphism, where prion strains were encoded by distinct β -sheets built from different segments of a protein (Chien et al., 2004).

Several human diseases are associated with the pathogenic conformational changes in corresponding proteins and the subsequent accumulation of proteinaceous deposits in brain (Kelly, 1998). The pathogenesis of PD and AD is determined (at least in part) by the age-related failure of the cellular and organism defense systems which aim at the clearance of misfolded proteins. Interestingly, AD and PD are associated with the aggregation of IDPs – $\text{A}\beta$ and tau in AD and α -synuclein in PD (Uversky, 2009a).

Since many completely or partially disordered proteins, especially hubs, play key roles in regulation and coordination of almost all the biochemical processes inside the cell, their misfolding can be pathogenic. This fact provides a new view on the molecular mechanisms of the diseases associated with amyloid fibril formation. In contrast to the globular proteins, which have to unfold prior to aggregation (Jahn and Radford, 2005), IDPs are always ready for intermolecular interactions. An unbound fragment of an IDP possesses a strong ability to interact, and therefore can bind either to natural partners forming native complexes or to similar molecules forming various aggregates. This raises the question of why IDPs do not always form aggregates in the norm. One of the potential answers is the fact that inside the cell, the IDPs typically form complexes with natural partners (Dunker et al., 2002a). Furthermore, nature developed very sophisticated protection mechanisms (chaperones, proteasome, etc.) for the effective prevention of dangerous consequences of misfolding. In other words, in the cell, any given protein is not acting in the isolation, is never alone and is constantly "watched" by the protective machinery. This machinery is rather robust and can tolerate significant loads. Obviously, factors that affect these protective mechanisms will contribute to the probability of disease development (see below). Finally, based on the analysis of the IDP amino acid composition, it would be clearly a mistake to assume that an averaged IDP possesses higher propensity towards aggregation than an averaged ordered protein. In fact, many IDPs contain large number of charged and polar residues. In addition to the hydrophobic interactions, the net charge is one of the major factors determining aggregation behavior of a protein. Many proteins tend to aggregate when the net charge is zero (at their respective pIs); i.e., under the conditions where the repulsion between the

various protein molecules is minimized (Chiti, 2006). However, since IDPs are often characterized by extreme pIs (which can be as low as pH 3.0 and as high as pH 12.0) it would be very difficult for them to reach the conditions of zero net charge physiologically. This means that evolution worked against misfolding and aggregation, and the sequences of IDPs were selected so that they do not misfold and aggregate easily.

The fibril formation and the accumulation of proteinaceous deposits in the affected organs are not necessarily the causative agents of the conformational diseases, as various protofibrils can be cytotoxic (Bucciantini et al., 2002). We believe that the true mechanisms of the development of these diseases are not associated with early or late fibrillogenesis stages, but are determined by some deep distortions of the vital regulatory processes in the cells. If this hypothesis is correct, then the formation of protofibrils and fibrils represents consequence rather than the cause of the disease.

Functional amyloids

Although the fibrous amyloid-like forms of proteins are usually considered as pathological hallmarks of various diseases, recent studies revealed that amyloid-like proteins occur naturally in many organisms and may have crucial biological functions, e.g. playing important roles in adhesion to surfaces, cell aggregation, biofilm formation, molecular sequestration and storage, etc. (Barnhart and Chapman, 2006; Chapman et al., 2002; Chiti and Dobson, 2006; Maji et al., 2009; Mostaert et al., 2006). The most studied of such functional amyloids are the yeast prions, $[PSI^+]$, $[URE3]$, $[Het-s]$, and $[RNQ1^+]$, where the heritable prion strains result from self-propagating conformational differences within the corresponding prion protein (Chien and Weissman, 2001). The $[PSI^+]$ prion element of the yeast *S. cerevisiae* (the [] in this designation indicates that the determinant is extra-chromosomal, and the capital letters indicate that $[PSI^+]$ is dominant) is transmitted through self-propagating aggregates of Sup35 and might help $[PSI^+]$ carriers to survive at the unfavorable conditions. Sup35 is a component of the eukaryotic translation release factor (Stansfield et al., 1995) and is considered as an epigenetic modulator of the fidelity of protein synthesis (Serio and Lindquist, 1999). Being required for the faithful termination of translation at stop codons in messenger RNAs, Sup35p can exist in at least two different stable physical states. The “normal” state is associated with accurate termination, whereas the “altered” state allows nonsense suppression. A self-propagating change in Sup35p from the normal (active) to the altered (inactive) state is similar to mammalian prions and is proposed to give rise to the $[PSI^+]$ nonsense suppressor determinant. When Sup35 is in the $[PSI^+]$ state, ribosomes often fail to release polypeptides at stop codons, causing a non-Mendelian trait which is easily detected by nonsense suppression (Serio and Lindquist, 1999). Therefore, it is believed now that $[PSI^+]$ represents a dominant, epigenetic loss-of-function in Sup35p, since the reduction in termination factor activity can result in increased nonsense suppression by mutant tRNAs. $[PSI^+]$ might act as a buffer for nonsense mutations, thereby increasing the genomic flexibility of the organism and influencing the rate at which it evolves (Serio and Lindquist, 1999). Interestingly, although $[PSI^+]$ is not typically required under laboratory growth conditions, one obligate $[PSI^+]$ strain is known, where a nonsense fatal mutation in the essential gene, heat shock transcription factor (HSF) is present, and where $[PSI^+]$ allows synthesis of functional HSF and, therefore, maintains cell viability (Lindquist et al., 1995). At the molecular level, it is known that the propagation of $[PSI^+]$ is mediated by the (glutamine/asparagine)-rich amino-terminal prion-determining domain (PrD) (DePace et al., 1998; Ter-Avanesyan et al., 1994). Transient overexpression of this domain was shown to lead to the formation of *de novo* Sup35p prion aggregates *in vivo* (Patino et al., 1996), whereas the purified PrD is known to form self-propagating amyloid fibres *in vitro* (Glover et al., 1997; King et al., 1997).

The underlying determinant of another well-studied prion-like trait [URE3] is the Ure2p protein (Aigle and Lacroute, 1975; Wickner, 1994). Ure2p is involved in the nitrogen regulation mediating the shut-off of production of proteins involved in assimilation of poor nitrogen sources when yeast is grown in the presence of a good nitrogen source such as glutamine, ammonia, or glutamate (Coschigano and Magasanik, 1991; Courchesne and Magasanik, 1988). Interestingly, Ure2p and Sup35p being completely unrelated share many similarities: their N-terminal domains are involved in propagation or generation of the prion state and their C-terminal domains are responsible for the normal function (Wickner, 1994).

The [Het-s] infectious element of the filamentous fungus *Podospora anserina* represents the prion form of the HET-s protein (Coustou et al., 1997). [Het-s] strains differ from the prion-free strains by their reactivity toward strains expressing the HET-S protein, which is a polymorphic variant that differs from HET-s by 13 amino acid residues (Deleu et al., 1993). The prion form of HET-S spreads rapidly through the colony and can convert the non-prion form of the protein to a prion state after compatible colonies have merged (Maddelein et al., 2002). Cell fusions between the prion-free and [Het-S] strains are viable, whereas fusions between [Het-s] and [Het-S] are fatal. This mechanism prevents promiscuous sharing between unrelated colonies and ensures that only related colonies obtain the benefit of sharing resources, therefore representing a natural system of the protective "incompatibility".

A yeast protein Rnq1 from *Saccharomyces cerevisiae* is required for another yeast prion [*PIN*⁺] (also known as [*RNQI*⁺]), which is necessary for the *de novo* induction of a second prion, [*PSI*⁺] (Kurahashi et al., 2008). Rnq1 is a yeast protein of unknown function which contains a QN-rich prion domain, hence named so for rich in asparagine (N) and glutamine (Q) (Sondheimer and Lindquist, 2000).

Overall, the evidence has accumulated that at least some prions are beneficial for yeast allowing cells to survive certain stress conditions (True et al., 2004; True and Lindquist, 2000). Similar to yeast, beneficial prion-like structural conversions and formation of cell-surface proteinaceous filaments known as curli have been proposed to be involved in the biofilm formation in bacteria (Chapman et al., 2002). Curli can promote colonization of an epithelial surface, entry into host cells, exchange of DNA between bacteria, and development of bacterial communities organized as biofilms, colonies, or multicellular fruiting bodies. The formation of curli is a highly controlled process that involves CsgA curlin subunit, the purified form of which adopts a soluble, unstructured form that upon prolonged incubation assembles into fibers morphologically similar to curli, the CsgB nucleator, and the nucleation-precipitation machinery consisting of the CsgE chaperone and the CsgF nucleator, which may work independently or in concert with CsgB to guide in vivo extracellular nucleation of CsgA (Chapman et al., 2002).

Finally, functional amyloids were found even in mammals, where the amyloid structures were suggested to play a normal role in the melanin formation (Fowler et al., 2006; Fowler et al., 2007). It has been shown that mammalian melanosomes, highly abundant cellular organelles generated in developmentally specialized cells including melanocytes and retinal pigment epithelium of the skin and eyes, contain insoluble, luminal fibers made of the Pmel17 protein. Intriguingly, these structures possessed all the properties of typical amyloids. Similar to the described above curli morphogenesis, the formation of the Pmel17 amyloids is tightly controlled by the secretory pathway which utilizes the membrane sequestration and proteolytic steps to protect the cell from amyloid and amyloidogenic intermediates that can be toxic (Fowler et al., 2006). Pmel17 amyloid templated and accelerated the covalent polymerization of reactive small molecules into melanin, which is a critically important biopolymer that protects against a broad range of cytotoxic insults

including UV and oxidative damage. Pmel17 amyloid also was shown to play a role in mitigating the toxicity associated with melanin formation by sequestering and minimizing diffusion of highly reactive, toxic melanin precursors out of the melanosome (Fowler et al., 2006).

Thus, these data clearly show that in addition to the well-known harmful amyloids, formation of which is associated with the pathology of various diseases, there are biologically active, native amyloids. Since these functional amyloids are found in a broad range of organisms, starting from bacteria and ending with human, they may represent an ancient, evolutionarily conserved protein quaternary structure, reemphasizing the importance of studies on the molecular mechanisms of amyloid fibril formation.

Protein folding in vivo

Folding of globular proteins inside the cell

Newly synthesized protein finds itself in the "overcrowded" physiological cell medium, where concentration of proteins, nuclear acids, and polysaccharides are as high as 400 mg/ml, and where macromolecules occupy a significant part of the medium volume – up to 40% (Ellis, 2001; Zimmerman and Minton, 1993). Such conditions can greatly affect all the biological processes including protein folding, misfolding, and aggregation (Chebotareva et al., 2004; Minton, 2000; Uversky et al., 2002; van den Berg et al., 2000). The folding of proteins in the living cell is complicated by at least two factors: the realm of unfavorable contacts with "neighbors," and the appearance of the incorrect intramolecular contacts at a co-translational folding. Therefore, in order for the correct folding to occur, a set of special protein-helpers provide assistance – chaperones and enzymes regulating cis-trans proline isomerization and the formation of the disulfide bridges (Bader and Bardwell, 2001; Fink, 1999; Gilbert, 1994; Schmid, 2001), which prevent protein aggregation and misfolding, accelerate folding, and participate in protein transport (e.g., in protein translocation through the membranes). Chaperone-regulated co-translational folding is especially important for multidomain proteins.

Chaperones constitute a wide family of proteins with various molecular masses, structures, and functions. For example, Hsp70 chaperones, with the assistance of the co-chaperones of the DnaJ/Hsp40 family, interact non-specifically with small hydrophobic clusters at the newly synthesized polypeptide chain (Feldman and Frydman, 2000). The major role of Hsp70 is to prevent the undesirable interactions of nascent polypeptides with other molecules. For some proteins, interaction with Hsp70 is sufficient for correct folding, but the folding of multidomain proteins requires the participation of other helpers. For example, a correct folding of actin relies on its interaction with the prefolding (PFD; (Martin-Benito et al., 2002)), which participates in the translocation of the partially folded actin to the CCT chaperonin (chaperon containing TCP-1) (Feldman and Frydman, 2000; Neirynck et al., 2006). In contrast to typical Hsp70, CCT has specific sites for interaction with actin and tubulin and also participates in the folding of a number of other proteins non-homologous to actin or tubulin (Willison and Grantham, 2001).

CCT consist of two stacked toroids, each containing eight three-domain proteins. The equatorial domains are responsible for the intertoroid interactions and for the interaction with ATP, whereas the apical domains realize the interaction with the substrate and provide the passage of the substrate to the central cavity. The folding of actin is a complex multi-stage, ATP-dependent process controlled by CCT (Figure 2) (Altschuler and Willison, 2008; Neirynck et al., 2006).

It is important to remember that despite their crucial role in the folding of globular protein *in vivo*, chaperones can only assist protein folding. They do not carry structural information, which is necessary for the newly synthesized polypeptide chain to fold. It is very likely that interactions with chaperones and other proteins are even more important for IDPs. Such interactions will definitely prevent these proteins from aggregation and proteolysis (see below).

IDPs in living cells: Controlled chaos

It is clear now that the IDPs are real, abundant, diversified, and vital. Functions of IDPs are complementary to the catalytic activities of ordered proteins (Dunker et al., 2002a; Dunker et al., 2002b; Dunker et al., 2005; Dunker et al., 2001; Dunker and Obradovic, 2001; Dunker et al., 2008a; Dunker et al., 2008b; Dunker Uversky, 2008; Dyson and Wright, 2005; Iakoucheva et al., 2002; Radivojac et al., 2007; Uversky, 2002a; Uversky, 2002b; Uversky et al., 2000; Uversky et al., 2005; Vucetic et al., 2007; Wright and Dyson, 1999; Xie et al., 2007a; Xie et al., 2007b). Many disorder-related functions (e.g., signaling, control, regulation and recognition) are incompatible with well-defined, stable 3-D structures (Dunker et al., 2005; Dunker et al., 2001; Dunker and Obradovic, 2001; Dunker et al., 2008a; Dunker and Uversky, 2008; Dyson and Wright, 2005; Radivojac et al., 2007; Tompa, 2002; Tompa, 2005; Uversky, 2002a; Uversky, 2002b; Uversky et al., 2000; Uversky et al., 2005; Vucetic et al., 2007; Wright and Dyson, 1999; Xie et al., 2007a; Xie et al., 2007b). Intrinsic disorder is assumed to provide several functional advantages for its carriers including increased interaction surface area, structural plasticity to interact with several targets, high specificity for given partners combined with high k_{on} and k_{off} rates that enable rapid association with the partner without an excessive binding strength, the ability to fold upon binding and accessible post-translational modification sites. Structurally, IDPs range from completely unstructured polypeptides (native coils, that resemble the highly unfolded states of globular proteins) to extended partially structured forms (native pre-molten globules) or even to compact disordered ensembles that may contain significant secondary structure (native molten globules) (Dunker and Obradovic, 2001; Uversky, 2002a; Uversky, 2002b; Uversky, 2003). These proteins are highly abundant in nature (~55% of eukaryotic proteins are predicted to contain at least one disordered region that is at least 40 amino acids in length (Dunker et al., 2001) and are often associated with human diseases (Midic et al., 2009; Uversky et al., 2008; Uversky et al., 2009).

Unique structural and functional plasticity of IDPs defines their wide involvement in various biological processes, especially those where a complex interplay between different cellular levels, compartments, and structures is required (Gsponer and Babu, 2009). Such interplay constitutes a core of the effective cellular response to different stimuli and external conditions, and IDPs with their ability to sample broad repertoire of structural and functional states provide wider response variability than do structured proteins (Gsponer and Babu, 2009).

The highly dynamic nature of IDPs is a visual illustration of the chaos. However, the evolutionary persistence of these highly dynamic proteins, their unique functionality and involvement in all the major cellular processes evidence that this chaos is tightly controlled (Uversky and Dunker, 2008). To answer the question on how these proteins are governed and regulated inside the cell Gsponer *et al.* conducted a detailed study focused at the intricate mechanisms of the IDP regulation (Gsponer et al., 2008). To this end, all the *Saccharomyces cerevisiae* proteins were grouped into three classes using one of the available disorder predictors, DisoPred2 (Ward et al., 2004): (i) 1971 highly ordered proteins containing 0 – 10% of the predicted disorder; (ii) 2711 moderately disordered proteins with 10 – 30% predicted disordered residues; and (iii) 2020 highly disordered proteins containing 30 – 100% of the predicted disorder. Then, the correlations between

intrinsic disorder and the various regulation steps of protein synthesis and degradation were evaluated.

To examine the transcription of genes encoding IDPs and ordered proteins, the transcriptional rates and the degradation rates of the corresponding transcripts were compared (Gsponer et al., 2008). This analysis revealed that the transcriptional rates of mRNAs encoding IDPs and ordered proteins were comparable. However the IDP-encoding transcripts were generally less abundant than transcripts encoding ordered proteins due to the increased decay rates of the formers.

The existence of tight regulation of the IDP abundance was also established at the protein level. In fact, IDPs were shown to be less abundant than ordered proteins due to the lower rate of protein synthesis and shorter protein half-lives. As the abundance and half-life in a cell of certain proteins can be further modulated via their post-translational modifications such as phosphorylation (Grimmler et al., 2007), the experimentally determined yeast kinase-substrate network was analyzed next. IDPs were shown to be substrates of twice as many kinases as were ordered proteins. Furthermore, the vast majority of kinases whose substrates were IDPs were either regulated in a cell-cycle dependent manner, or activated upon exposure to particular stimuli or stress (Gsponer et al., 2008). Therefore, post-translational modifications may not only serve as important mechanism for the fine-tuning of the IDP functions but possibly they are necessary to tune the IDP availability under the different cellular conditions.

In addition to *Saccharomyces cerevisiae*, similar regulation trends were also found in *Schizosaccharomyces pombe* and *Homo sapiens* (Gsponer et al., 2008). Based on these observations it has been concluded that both unicellular and multicellular organisms appear to use similar mechanisms for regulation of the IDP availability. Overall, this study clearly demonstrated that there is an evolutionarily conserved tight control of synthesis and clearance of most IDPs. This tight control is directly related to the major roles of IDPs in signaling, where it is crucial to be available in appropriate amounts and not to be present longer than needed (Gsponer et al., 2008). It has been also pointed out that although the abundance of many IDPs is under the strict control, some IDPs could be present in cells in large amounts or/and for long periods of time due to either specific post-translational modifications or via interactions with other factors, which could promote changes in cellular localization of IDPs or protect them from the degradation machinery (Dunker et al., 2001; Grimmler et al., 2007; Iakoucheva et al., 2004; Tompa, 2005; Xie et al., 2007a). Overall, this study clearly showed that the chaos seemingly introduced into the protein world by the discovery of IDPs is under the tight control (Uversky and Dunker, 2008).

Concluding remarks

Until quite recently the interests of physicists in studying the structure and folding of native proteins was mostly limited by globular proteins, whose molecules form identical structures in their native states. The unique structures of such proteins can be determined by X-ray crystallography. Studies on globular proteins were rather attractive for physicists, as the subject was quite adequate to the methods which they can suggest for its analysis. On the other hand, researchers were intrigued by the scale of the problem, attempting to understand why and how a polypeptide chain (in contrast to any synthetic polymer) is able to attain a unique, compact, and functional conformation, which is identical for all macromolecules. Although this problem is far from being completely resolved, significant progress was achieved in understanding globular protein folding and functioning.

Recently, it was recognized that the biological functions can be carried out not only by ordered proteins but also by proteins lacking rigid tertiary structure. Such IDPs constitute a significant portion (10–50%) of any given proteome. To fully apprehend the IDP structure and function it is necessary to investigate protein complexes. Although this task is much more complex than the structural and functional analysis of ordered proteins, the success in the resolving of this problem relies on the knowledge gained from the analysis of globular protein folding. Overall, recognizing that the majority of proteins cannot in principle form folded globular structures, together with the understanding that these IDPs, play a number of key roles in all cellular processes brought the analysis of these proteins to the front of modern protein science.

Another recent discovery is the establishing that in addition to the well-known pathological amyloids, which represent characteristic hallmarks of several conformational diseases, some amyloid proteins might have crucial biological functions. The illustrative examples of such functional amyloids include a set of yeast prions, curli in bacteria, and specific Pmel17 amyloids in melanosomes of mammals. The wide spread of such functional amyloids between different kingdoms of life and a broad spectrum of their biological functions suggest they represent a type of evolutionary conserved supramolecular structures.

Acknowledgments

The authors are thankful to Alexey Uversky for careful reading and editing the manuscript. This work was supported in part by grants from Russian Foundation of Basic Research (grants 07–04–01454 and 07–04–01038 to K.K.T. and I.M.K.), Program "Leading Scientific Schools of Russia" (grant #1961.2008.4 to K.K.T. and I.M.K.), Program "Molecular and Cell Biology" RAS (to K.K.T., I.M.K. and V.N.U.), the grants R01 LM007688-01A1 (to V.N.U.) and GM071714-01A2 (to V.N.U.) from the National Institutes of Health, and the grant EF 0849803 (to V.N.U.) from the National Science Foundation. We gratefully acknowledge the support of the IUPUI Signature Centers Initiative.

References

- Aigle M, Lacroute F. Genetical aspects of [URE3], a non-mitochondrial, cytoplasmically inherited mutation in yeast. *Mol Gen Genet* 1975;136:327–335. [PubMed: 16095000]
- Altschuler GM, Willison KR. Development of free-energy-based models for chaperonin containing TCP-1 mediated folding of actin. *J R Soc Interface* 2008;5:1391–1408. [PubMed: 18708324]
- Anfinsen CB. Principles that govern the folding of protein chains. *Science* 1973;181:223–230. [PubMed: 4124164]
- Bader MW, Bardwell JC. Catalysis of disulfide bond formation and isomerization in *Escherichia coli*. *Adv Protein Chem* 2001;59:283–301. [PubMed: 11868275]
- Barnhart MM, Chapman MR. Curli biogenesis and function. *Annu Rev Microbiol* 2006;60:131–147. [PubMed: 16704339]
- Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Res* 2000;28:235–242. [PubMed: 10592235]
- Bucciantini M, Giannoni E, Chiti F, Baroni F, Formigli L, Zurdo J, Taddei N, Ramponi G, Dobson CM, Stefani M. Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* 2002;416:507–511. [PubMed: 11932737]
- Bullock AN, Henckel J, DeDecker BS, Johnson CM, Nikolova PV, Proctor MR, Lane DP, Fersht AR. Thermodynamic stability of wild-type and mutant p53 core domain. *Proc Natl Acad Sci U S A* 1997;94:14338–14342. [PubMed: 9405613]
- Chapman MR, Robinson LS, Pinkner JS, Roth R, Heuser J, Hammar M, Normark S, Hultgren SJ. Role of *Escherichia coli* curli operons in directing amyloid fiber formation. *Science* 2002;295:851–855. [PubMed: 11823641]
- Chebotaeva NA, Andreeva IE, Makeeva VF, Livanova NB, Kurganov BI. Effect of molecular crowding on self-association of phosphorylase kinase and its interaction with phosphorylase b and glycogen. *J Mol Recognit* 2004;17:426–432. [PubMed: 15362101]

- Cheng Y, LeGall T, Oldfield CJ, Dunker AK, Uversky VN. Abundance of intrinsic disorder in protein associated with cardiovascular disease. *Biochemistry* 2006;45:10448–10460. [PubMed: 16939197]
- Cheng Y, Oldfield CJ, Meng J, Romero P, Uversky VN, Dunker AK. Mining alpha-helix-forming molecular recognition features with cross species sequence alignments. *Biochemistry* 2007;46:13468–13477. [PubMed: 17973494]
- Chien P, Weissman JS. Conformational diversity in a yeast prion dictates its seeding specificity. *Nature* 2001;410:223–227. [PubMed: 11242084]
- Chien P, Weissman JS, DePace AH. Emerging principles of conformation-based prion inheritance. *Annu Rev Biochem* 2004;73:617–656. [PubMed: 15189155]
- Chiti F. Relative importance of hydrophobicity, net charge and secondary structure propensities in protein aggregation. In: Uversky, VN.; Fink, AL., editors. *Protein Misfolding, Aggregation and Conformational Diseases*, Vol. I: Protein Aggregation and Conformational Diseases. New York: Springer Science+Business Media, LLC; 2006. p. 43-59.
- Chiti F, Dobson CM. Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem* 2006;75:333–366. [PubMed: 16756495]
- Chiti F, Webster P, Taddei N, Clark A, Stefani M, Ramponi G, Dobson CM. Designing conditions for in vitro formation of amyloid protofilaments and fibrils. *Proc Natl Acad Sci U S A* 1999;96:3590–3594. [PubMed: 10097081]
- Chumakov PM. Versatile functions of p53 protein in multicellular organisms. *Biochemistry (Mosc)* 2007;72:1399–1421. [PubMed: 18282133]
- Cohen FE, Prusiner SB. Pathologic conformations of prion proteins. *Annu Rev Biochem* 1998;67:793–819. [PubMed: 9759504]
- Coschigano PW, Magasanik B. The URE2 gene product of *Saccharomyces cerevisiae* plays an important role in the cellular response to the nitrogen source and has homology to glutathione s-transferases. *Mol Cell Biol* 1991;11:822–832. [PubMed: 1990286]
- Courchesne WE, Magasanik B. Regulation of nitrogen assimilation in *Saccharomyces cerevisiae*: roles of the URE2 and GLN3 genes. *J Bacteriol* 1988;170:708–713. [PubMed: 2892826]
- Coustou V, Deleu C, Saupe S, Begueret J. The protein product of the het-s heterokaryon incompatibility gene of the fungus *Podospora anserina* behaves as a prion analog. *Proc Natl Acad Sci U S A* 1997;94:9773–9778. [PubMed: 9275200]
- Deleu C, Clave C, Begueret J. A single amino acid difference is sufficient to elicit vegetative incompatibility in the fungus *Podospora anserina*. *Genetics* 1993;135:45–52. [PubMed: 8224826]
- DePace AH, Santoso A, Hillner P, Weissman JS. A critical role for amino-terminal glutamine/asparagine repeats in the formation and propagation of a yeast prion. *Cell* 1998;93:1241–1252. [PubMed: 9657156]
- Dobson CM. Protein misfolding, evolution and disease. *Trends Biochem Sci* 1999;24:329–332. [PubMed: 10470028]
- Dolgikh DA, Gilmanshin RI, Brazhnikov EV, Bychkova VE, Semisotnov GV, Venyaminov S, Ptitsyn OB. Alpha-Lactalbumin: compact state with fluctuating tertiary structure? *FEBS Lett* 1981;136:311–315. [PubMed: 7327267]
- Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradovic Z. Intrinsic disorder and protein function. *Biochemistry* 2002a;41:6573–6582. [PubMed: 12022860]
- Dunker AK, Brown CJ, Obradovic Z. Identification and functions of usefully disordered proteins. *Adv Protein Chem* 2002b;62:25–49. [PubMed: 12418100]
- Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN. Flexible nets: The roles of intrinsic disorder in protein interaction networks. *FEBS Journal* 2005;272:5129–5148. [PubMed: 16218947]
- Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CM, Hipps KW, Ausio J, Nissen MS, Reeves R, Kang C, Kissinger CR, Bailey RW, Griswold MD, Chiu W, Garner EC, Obradovic Z. Intrinsically disordered protein. *J Mol Graph Model* 2001;19:26–59. [PubMed: 11381529]
- Dunker AK, Obradovic Z. The protein trinity--linking function and disorder. *Nat Biotechnol* 2001;19:805–806. [PubMed: 11533628]

- Dunker AK, Obradovic Z, Romero P, Garner EC, Brown CJ. Intrinsic protein disorder in complete genomes. *Genome Inform Ser Workshop Genome Inform* 2000;11:161–171.
- Dunker AK, Oldfield CJ, Meng J, Romero P, Yang JY, Chen JW, Vacic V, Obradovic Z, Uversky VN. The unfoldomics decade: an update on intrinsically disordered proteins. *BMC Genomics* 2008a;9 Suppl 2:S1.
- Dunker AK, Silman I, Uversky VN, Sussman JL. Function and structure of inherently disordered proteins. *Curr Opin Struct Biol* 2008b;18:756–764. [PubMed: 18952168]
- Dunker AK, Uversky VN. Signal transduction via unstructured protein conduits. *Nat Chem Biol* 2008;4:229–230. [PubMed: 18347590]
- Dyson HJ, Wright PE. Coupling of folding and binding for unstructured proteins. *Curr Opin Struct Biol* 2002;12:54–60. [PubMed: 11839490]
- Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 2005;6:197–208. [PubMed: 15738986]
- Ellis RJ. Macromolecular crowding: obvious but underappreciated. *Trends Biochem Sci* 2001;26:597–604. [PubMed: 11590012]
- Fandrich M, Fletcher MA, Dobson CM. Amyloid fibrils from muscle myoglobin. *Nature* 2001;410:165–166. [PubMed: 11242064]
- Feldman DE, Frydman J. Protein folding in vivo: the importance of molecular chaperones. *Curr Opin Struct Biol* 2000;10:26–33. [PubMed: 10679467]
- Ferron F, Longhi S, Canard B, Karlin D. A practical overview of protein disorder prediction methods. *Proteins* 2006;65:1–14. [PubMed: 16856179]
- Fink AL. Chaperone-mediated protein folding. *Physiol Rev* 1999;79:425–449. [PubMed: 10221986]
- Finkelstein, AV.; Ptitsyn, OB. *Protein Physics: A Course of Lectures*. Academic Press; 2002. p. 354
- Fischer E. Einfluss der Configuration auf die Wirkung der Enzyme. *Ber. Dt. Chem Ges* 1894;27:2985–2993.
- Fowler DM, Koulov AV, Alory-Jost C, Marks MS, Balch WE, Kelly JW. Functional amyloid formation within mammalian tissue. *PLoS Biol* 2006;4:e6. [PubMed: 16300414]
- Fowler DM, Koulov AV, Balch WE, Kelly JW. Functional amyloid--from bacteria to humans. *Trends Biochem Sci* 2007;32:217–224. [PubMed: 17412596]
- Frokjaer S, Otzen DE. Protein drug stability: a formulation challenge. *Nat Rev Drug Discov* 2005;4:298–306. [PubMed: 15803194]
- Fuxreiter M, Simon I, Friedrich P, Tompa P. Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *J Mol Biol* 2004;338:1015–1026. [PubMed: 15111064]
- Garner E, Romero P, Dunker AK, Brown C, Obradovic Z. Predicting Binding Regions within Disordered Proteins. *Genome Inform Ser Workshop Genome Inform* 1999;10:41–50.
- Gilbert HF. Protein chaperones and protein folding. *Curr Opin Biotechnol* 1994;5:534–539. [PubMed: 7765469]
- Glover JR, Kowal AS, Schirmer EC, Patino MM, Liu JJ, Lindquist S. Self-seeded fibers formed by Sup35, the protein determinant of [PSI⁺], a heritable prion-like factor of *S. cerevisiae*. *Cell* 1997;89:811–819. [PubMed: 9182769]
- Goers J, Permyakov SE, Permyakov EA, Uversky VN, Fink AL. Conformational prerequisites for alpha-lactalbumin fibrillation. *Biochemistry* 2002;41:12546–12551. [PubMed: 12369846]
- Grimmler M, Wang Y, Mund T, Cilensek Z, Keidel EM, Waddell MB, Jakel H, Kullmann M, Kriwacki RW, Hengst L. Cdk-inhibitory activity and stability of p27Kip1 are directly regulated by oncogenic tyrosine kinases. *Cell* 2007;128:269–280. [PubMed: 17254966]
- Gsponer J, Babu MM. The rules of disorder or why disorder rules. *Prog Biophys Mol Biol* 2009;99:94–103. [PubMed: 19344736]
- Gsponer J, Futschik ME, Teichmann SA, Babu MM. Tight regulation of unstructured proteins: from transcript synthesis to protein degradation. *Science* 2008;322:1365–1368. [PubMed: 19039133]
- Guijarro JI, Sunde M, Jones JA, Campbell ID, Dobson CM. Amyloid fibril formation by an SH3 domain. *Proc Natl Acad Sci U S A* 1998;95:4224–4228. [PubMed: 9539718]

- Hamada D, Segawa S, Goto Y. Non-native alpha-helical intermediate in the refolding of beta-lactoglobulin, a predominantly beta-sheet protein. *Nat Struct Biol* 1996;3:868–873. [PubMed: 8836104]
- He B, Wang K, Liu Y, Xue B, Uversky VN, Dunker AK. Predicting intrinsic disorder in proteins: an overview. *Cell Res* 2009;19:929–949. [PubMed: 19597536]
- Hollstein M, Sidransky D, Vogelstein B, Harris CC. p53 mutations in human cancers. *Science* 1991;253:49–53. [PubMed: 1905840]
- Huang Y, Liu Z. Kinetic advantage of intrinsically disordered proteins in coupled folding-binding process: a critical assessment of the "fly-casting" mechanism. *J Mol Biol* 2009;393:1143–1159. [PubMed: 19747922]
- Iakoucheva LM, Brown CJ, Lawson JD, Obradovic Z, Dunker AK. Intrinsic disorder in cell-signaling and cancer-associated proteins. *J Mol Biol* 2002;323:573–584. [PubMed: 12381310]
- Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK. The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res* 2004;32:1037–1049. [PubMed: 14960716]
- Jahn TR, Radford SE. The Yin and Yang of protein folding. *Febs J* 2005;272:5962–5970. [PubMed: 16302961]
- Joerger AC, Fersht AR. Structural biology of the tumor suppressor p53 and cancer-associated mutants. *Adv Cancer Res* 2007a;97:1–23. [PubMed: 17419939]
- Joerger AC, Fersht AR. Structure-function-rescue: the diverse nature of common p53 cancer mutants. *Oncogene* 2007b;26:2226–2242. [PubMed: 17401432]
- Joerger AC, Fersht AR. Structural biology of the tumor suppressor p53. *Annu Rev Biochem* 2008;77:557–582. [PubMed: 18410249]
- Kelly JW. The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Curr Opin Struct Biol* 1998;8:101–106. [PubMed: 9519302]
- King CY, Tittmann P, Gross H, Gebert R, Aepli M, Wuthrich K. Prion-inducing domain 2–114 of yeast Sup35 protein transforms in vitro into amyloid-like filaments. *Proc Natl Acad Sci U S A* 1997;94:6618–6622. [PubMed: 9192614]
- Kurahashi H, Ishiwata M, Shibata S, Nakamura Y. A regulatory role of the Rnq1 nonprion domain for prion propagation and polyglutamine aggregates. *Mol Cell Biol* 2008;28:3313–3323. [PubMed: 18332119]
- Kuznetsova IM, Biktashev AG, Khaitlina SY, Vassilenko KS, Turoverov KK, Uversky VN. Effect of self-association on the structural organization of partially folded proteins: inactivated actin. *Biophys J* 1999;77:2788–2800. [PubMed: 10545377]
- Kuznetsova IM, Forge V, Turoverov KK. [Structural dynamics, stability and folding of proteins]. *Tsitologiya* 2005;47:943–952. [PubMed: 16706198]
- Lacy ER, Filippov I, Lewis WS, Otieno S, Xiao L, Weiss S, Hengst L, Kriwacki RW. p27 binds cyclin-CDK complexes through a sequential mechanism involving binding-induced protein folding. *Nat Struct Mol Biol* 2004;11:358–364. [PubMed: 15024385]
- Levinthal C. Are there pathways for protein folding? *Journal de Chimie Physique* 1968;65:44–45.
- Li X, Romero P, Rani M, Dunker AK, Obradovic Z. Predicting Protein Disorder for N-, C-, and Internal Regions. *Genome Inform Ser Workshop Genome Inform* 1999;10:30–40.
- Lindquist S, Patino MM, Chernoff YO, Kowal AS, Singer MA, Liebman SW, Lee KH, Blake T. The role of Hsp104 in stress tolerance and [PSI⁺] propagation in *Saccharomyces cerevisiae*. *Cold Spring Harb Symp Quant Biol* 1995;60:451–460. [PubMed: 8824419]
- Maddelein ML, Dos Reis S, Duvezin-Caubet S, Coulary-Salin B, Saupe SJ. Amyloid aggregates of the HET-s prion protein are infectious. *Proc Natl Acad Sci U S A* 2002;99:7402–7407. [PubMed: 12032295]
- Maji SK, Perrin MH, Sawaya MR, Jessberger S, Vadodaria K, Rissman RA, Singru PS, Nilsson KP, Simon R, Schubert D, Eisenberg D, Rivier J, Sawchenko P, Vale W, Riek R. Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science* 2009;325:328–332. [PubMed: 19541956]

- Martin-Benito J, Boskovic J, Gomez-Puertas P, Carrascosa JL, Simons CT, Lewis SA, Bartolini F, Cowan NJ, Valpuesta JM. Structure of eukaryotic prefoldin and of its complexes with unfolded actin and the cytosolic chaperonin CCT. *Embo J* 2002;21:6377–6386. [PubMed: 12456645]
- Midic U, Oldfield CJ, Dunker AK, Obradovic Z, Uversky VN. Protein disorder in the human diseasome: unfoldomics of human genetic diseases. *BMC Genomics* 2009;10 Suppl 1:S12. [PubMed: 19594871]
- Minton AP. Implications of macromolecular crowding for protein assembly. *Curr Opin Struct Biol* 2000;10:34–39. [PubMed: 10679465]
- Mohan, A. The School of Informatics. Indianapolis: Indiana University; 2006. MoRFs: A dataset of Molecular Recognition Features; p. 59
- Mostaert AS, Higgins MJ, Fukuma T, Rindi F, Jarvis SP. Nanoscale mechanical characterisation of amyloid fibrils discovered in a natural adhesive. *J Biol Phys* 2006;32:393–401. [PubMed: 19669445]
- Munishkina LA, Fink AL, Uversky VN. Conformational prerequisites for formation of amyloid fibrils from histones. *J Mol Biol* 2004;342:1305–1324. [PubMed: 15351653]
- Neiryneck K, Waterschoot D, Vandekerckhove J, Ampe C, Rommelaere H. Actin interacts with CCT via discrete binding sites: a binding transition-release model for CCT-mediated actin folding. *J Mol Biol* 2006;355:124–138. [PubMed: 16300788]
- Nelson R, Sawaya MR, Balbirnie M, Madsen AO, Riekel C, Grothe R, Eisenberg D. Structure of the cross-beta spine of amyloid-like fibrils. *Nature* 2005;435:773–778. [PubMed: 15944695]
- Ohgushi M, Wada A. 'Molten-globule state': a compact form of globular proteins with mobile side-chains. *FEBS Lett* 1983;164:21–24. [PubMed: 6317443]
- Oldfield CJ, Cheng Y, Cortese MS, Romero P, Uversky VN, Dunker AK. Coupled folding and binding with alpha-helix-forming molecular recognition elements. *Biochemistry*. 2005 In press.
- Patino MM, Liu JJ, Glover JR, Lindquist S. Support for the prion hypothesis for inheritance of a phenotypic trait in yeast. *Science* 1996;273:622–626. [PubMed: 8662547]
- Pavlov NA, Cherny DI, Heim G, Jovin TM, Subramaniam V. Amyloid fibrils from the mammalian protein prothymosin alpha. *FEBS Lett* 2002;517:37–40. [PubMed: 12062405]
- Povarova OI, Kuznetsova IM, Turoverov KK. Different disturbances--one pathway of protein unfolding. Actin folding-unfolding and misfolding. *Cell Biol Int* 2007;31:405–412. [PubMed: 17336100]
- Ptitsyn OB. [Stages in the mechanism of self-organization of protein molecules]. *Dokl Akad Nauk SSSR* 1973;210:1213–1215. [PubMed: 4721708]
- Ptitsyn OB. Molten globule and protein folding. *Adv Protein Chem* 1995;47:83–229. [PubMed: 8561052]
- Radford SE. Protein folding: progress made and promises ahead. *Trends Biochem Sci* 2000;25:611–618. [PubMed: 11116188]
- Radivojac P, Iakoucheva LM, Oldfield CJ, Obradovic Z, Uversky VN, Dunker AK. Intrinsic disorder and functional proteomics. *Biophys J* 2007;92:1439–1456. [PubMed: 17158572]
- Receveur-Brechot V, Bourhis JM, Uversky VN, Canard B, Longhi S. Assessing protein disorder and induced folding. *Proteins* 2006;62:24–45. [PubMed: 16287116]
- Reeves R, Beckerbauer L. HMGI/Y proteins: flexible regulators of transcription and chromatin structure. *Biochim Biophys Acta* 2001;1519:13–29. [PubMed: 11406267]
- Romero P, Obradovic Z, Kissinger CR, Villafranca JE, Dunker AK. Identifying disordered regions in proteins from amino acid sequences. *IEEE Int. Conf Neural Netw* 1997;1:90–95.
- Sawaya MR, Sambashivan S, Nelson R, Ivanova MI, Sievers SA, Apostol MI, Thompson MJ, Balbirnie M, Wiltzius JJ, McFarlane HT, Madsen AO, Riekel C, Eisenberg D. Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature* 2007;447:453–457. [PubMed: 17468747]
- Schmid FX. Prolyl isomerases. *Adv Protein Chem* 2001;59:243–282. [PubMed: 11868274]
- Schultz CP. Illuminating folding intermediates. *Nat Struct Biol* 2000;7:7–10. [PubMed: 10625415]
- Serio TR, Lindquist SL. [PSI+]: an epigenetic modulator of translation termination efficiency. *Annu Rev Cell Dev Biol* 1999;15:661–703. [PubMed: 10611975]

- Serpell LC, Sunde M, Blake CC. The molecular basis of amyloidosis. *Cell Mol Life Sci* 1997;53:871–887. [PubMed: 9447239]
- Shoemaker BA, Portman JJ, Wolynes PG. Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc Natl Acad Sci U S A* 2000;97:8868–8873. [PubMed: 10908673]
- Sondheimer N, Lindquist S. Rnq1: an epigenetic modifier of protein function in yeast. *Mol Cell* 2000;5:163–172. [PubMed: 10678178]
- Stansfield I, Jones KM, Kushnirov VV, Dagkesamanskaya AR, Poznyakovski AI, Paushkin SV, Nierras CR, Cox BS, Ter-Avanesyan MD, Tuite MF. The products of the SUP45 (eRF1) and SUP35 genes interact to mediate translation termination in *Saccharomyces cerevisiae*. *Embo J* 1995;14:4365–4373. [PubMed: 7556078]
- Tanford C. Protein denaturation. *Adv Protein Chem* 1968;23:121–282. [PubMed: 4882248]
- Ter-Avanesyan MD, Dagkesamanskaya AR, Kushnirov VV, Smirnov VN. The SUP35 omnipotent suppressor gene is involved in the maintenance of the non-Mendelian determinant [psi+] in the yeast *Saccharomyces cerevisiae*. *Genetics* 1994;137:671–676. [PubMed: 8088512]
- Tomba P. Intrinsically unstructured proteins. *Trends Biochem Sci* 2002;27:527–533. [PubMed: 12368089]
- Tomba P. The interplay between structure and function in intrinsically unstructured proteins. *FEBS Lett* 2005;579:3346–3354. [PubMed: 15943980]
- True HL, Berlin I, Lindquist SL. Epigenetic regulation of translation reveals hidden genetic variation to produce complex traits. *Nature* 2004;431:184–187. [PubMed: 15311209]
- True HL, Lindquist SL. A yeast prion provides a mechanism for genetic variation and phenotypic diversity. *Nature* 2000;407:477–483. [PubMed: 11028992]
- Uversky VN. Natively unfolded proteins: a point where biology waits for physics. *Protein Sci* 2002a;11:739–756. [PubMed: 11910019]
- Uversky VN. What does it mean to be natively unfolded? *Eur J Biochem* 2002b;269:2–12. [PubMed: 11784292]
- Uversky VN. 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. [PubMed: 14523548]
- Uversky VN. Alpha-synuclein misfolding and neurodegenerative diseases. *Curr Protein Pept Sci* 2008;9:507–540. [PubMed: 18855701]
- Uversky VN. Intrinsic disorder in proteins associated with neurodegenerative diseases. *Frontiers in Bioscience*. 2009a In press.
- Uversky VN. Intrinsically disordered proteins and their environment: Effects of strong denaturants, temperature, pH, counter ions, membranes, binding partners, osmolytes, and macromolecular crowding. *Protein J*. 2009b
- Uversky VN, Cooper EM, Bower KS, Li J, Fink AL. Accelerated alpha-synuclein fibrillation in crowded milieu. *FEBS Lett* 2002;515:99–103. [PubMed: 11943202]
- Uversky VN, Dunker AK. Biochemistry. Controlled chaos. *Science* 2008;322:1340–1341. [PubMed: 19039128]
- Uversky, VN.; Fernandez, A.; Fink, AL. Structural and Conformational prerequisites of amyloidogenesis. In: Uversky, VN.; Fink, AL., editors. *Protein Misfolding, Aggregation and Conformational Diseases, Vol. I: Protein Aggregation and Conformational Diseases*. New York: Springer Science+Business Media, LLC; 2006. p. 1-20.
- Uversky VN, Gillespie JR, Fink AL. Why are "natively unfolded" proteins unstructured under physiologic conditions? *Proteins* 2000;41:415–427. [PubMed: 11025552]
- Uversky VN, Oldfield CJ, Dunker AK. Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *J Mol Recognit* 2005;18:343–384. [PubMed: 16094605]
- Uversky VN, Oldfield CJ, Dunker AK. Intrinsically disordered proteins in human diseases: introducing the D2 concept. *Annu Rev Biophys* 2008;37:215–246. [PubMed: 18573080]
- Uversky VN, Oldfield CJ, Midic U, Xie H, Xue B, Vucetic S, Iakoucheva LM, Obradovic Z, Dunker AK. Unfoldomics of human diseases: linking protein intrinsic disorder with diseases. *BMC Genomics* 2009;10 Suppl 1:S7. [PubMed: 19594884]

- Vacic V, Oldfield CJ, Mohan A, Radivojac P, Cortese MS, Uversky VN, Dunker AK. Characterization of molecular recognition features, MoRFs, and their binding partners. *J Proteome Res* 2007;6:2351–2366. [PubMed: 17488107]
- van, den; Berg, B.; Wain, R.; Dobson, CM.; Ellis, RJ. Macromolecular crowding perturbs protein refolding kinetics: implications for folding inside the cell. *EMBO J* 2000;19:3870–3875. [PubMed: 10921869]
- Varshavsky, YM. On the dynamic aspects of the spacial structure of biopolymers. In: Varshavsky, YM., editor. *Physical Methods for the Analysis of Molecular and Supramolecular Structures*. Moscow: Nauka; 1979. p. 58-71.
- Vucetic S, Xie H, Iakoucheva LM, Oldfield CJ, Dunker AK, Obradovic Z, Uversky VN. Functional anthology of intrinsic disorder. 2. Cellular components, domains, technical terms, developmental processes, and coding sequence diversities correlated with long disordered regions. *J Proteome Res* 2007;6:1899–1916. [PubMed: 17391015]
- Wang W. Protein aggregation and its inhibition in biopharmaceutics. *Int J Pharm* 2005;289:1–30. [PubMed: 15652195]
- Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT. Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J Mol Biol* 2004;337:635–645. [PubMed: 15019783]
- Wickner RB. [URE3] as an altered URE2 protein: evidence for a prion analog in *Saccharomyces cerevisiae*. *Science* 1994;264:566–569. [PubMed: 7909170]
- Willison, KR.; Grantham, J. The role of cytosolic chaperonin, CCT, in normal eukaryotic cell growth. In: Lund, P., editor. *Molecular chaperones in the cell: frontiers in molecular biology*. Oxford: Oxford University Press; 2001. p. 90-118.
- Wiltzius JJ, Landau M, Nelson R, Sawaya MR, Apostol MI, Goldschmidt L, Soriaga AB, Cascio D, Rajashankar K, Eisenberg D. Molecular mechanisms for protein-encoded inheritance. *Nat Struct Mol Biol* 2009;16:973–978. [PubMed: 19684598]
- Wright PE, Dyson HJ. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J Mol Biol* 1999;293:321–331. [PubMed: 10550212]
- Wright PE, Dyson HJ. Linking folding and binding. *Curr Opin Struct Biol* 2009;19:31–38. [PubMed: 19157855]
- Xie H, Vucetic S, Iakoucheva LM, Oldfield CJ, Dunker AK, Obradovic Z, Uversky VN. Functional anthology of intrinsic disorder. 3. Ligands, post-translational modifications, and diseases associated with intrinsically disordered proteins. *J Proteome Res* 2007a;6:1917–1932. [PubMed: 17391016]
- Xie H, Vucetic S, Iakoucheva LM, Oldfield CJ, Dunker AK, Uversky VN, Obradovic Z. Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. *J Proteome Res* 2007b;6:1882–1898. [PubMed: 17391014]
- Zimmerman SB, Minton AP. Macromolecular crowding: biochemical, biophysical, and physiological consequences. *Annu Rev Biophys Biomol Struct* 1993;22:27–65. [PubMed: 7688609]

Figure 1.

The energy landscape model illustrating the formation of native globular and intrinsically disordered proteins, supramolecular complexes, amorphous aggregates and amyloid fibrils. **A.** Globular proteins. In the globular protein folding, the increase in the free energy associated with the folding-induced entropy decrease is compensated by the formation of specific intramolecular contacts. Local free energy minima at the energy landscape correspond to the formation of partially folded intermediates (1, 2). Intermolecular contacts of partially folded protein molecules can result in the formation of oligomers, amorphous aggregates or amyloid fibrils

B. Intrinsically disordered proteins. Many native proteins with distinctive biological functions lack compact globular structure in aqueous solutions. Disordered segments of these proteins can gain ordered structure at the interaction with specific binding partners in a case if the free energy of such complexes is lower than the free energies of the intrinsically disordered protein and its partner. The propensity of native completely or partially disordered proteins to interact with various partners determines their biological functions in recognition of various binding partners (Ligands, nucleic acids and other proteins), in regulation of almost all cellular processes, and in signal transduction. In contrast to the folded globular proteins which have to unfold to become amyloidogenic, disordered proteins seem to be always ready for such intermolecular interactions. 1, 2, and 3 represents native complexes of intrinsically disordered proteins with various partners. This figure is based on the energy funnel model developed for globular proteins (Jahn and Radford, 2005; Schultz, 2000).

Figure 2.

The model of sequential actin folding on the chaperonin CCT. Greek letters denote CCT subunits; small (S) and large (L) domains of actin are shown as cylinders with the corresponding letters. Actin structure corresponding to the each folding step is shown to the right. Sites of interaction with CCT are shown in green. Green arrows indicate actin site(s) that work at the given stage. PFD-actin is a complex of actin with prefolding. Reproduced from (Neirynek et al., 2006) with the permission from authors.