CMLS, Cell. Mol. Life Sci. 60 (2003) 1852–1871
1420-682X/03/091852-20
DOI 10.1007/s00018-003-3096-6

CMLS Cellular and Molecular Life Sciences

Review

Protein folding revisited. A polypeptide chain at the folding – misfolding – nonfolding cross-roads: which way to go?

V. N. Uversky

Institute for Biological Instrumentation, Russian Academy of Sciences, 142292 Pushchino, Moscow Region (Russia) and Department of Chemistry and Biochemistry, University of California, Santa Cruz, California 95064 (USA), Fax: +1 831 459 2935, e-mail: uversky@hydrogen.ucsc.edu

Received 5 March 2003; received after revision 28 March 2003; accepted 31 March 2003

Abstract. The structure-function paradigm claims that a specific function of a protein is determined by its unique and rigid three-dimensional (3D) structure. Thus, following its biosynthesis on the ribosome, a protein must fold to be functional. This idea represents one of the cornerstones of modern biology. Numerous cases when, due to the effect of environmental factors or because of genetic defects (mutations), a polypeptide chain has lost its capability to gain a proper functional 3D structure (i.e. became misfolded), seem to confirm this concept. Consequences of such misfolding are well known and represent lost of function, aggregation, development of conformational disorders and cell death. However, the recent revelation of countless examples of intrinsically disordered proteins has cast doubt on the general validity of the structure-function paradigm and revealed an intriguing route of functional disorder. Thus, in a living cell, a polypeptide chain chooses between three potential fates – functional folding, potentially deadly misfolding and mysterious nonfolding. This choice is dictated by the peculiarities of amino acid sequence and/or by the pressure of environmental factors. The aim of the present review is to outline some interesting features of these three routes.

Key words. Protein folding; misfolding; nonfolding; partially folded intermediate; unfolded protein; random coil; natively unfolded protein; intrinsically disordered protein.

Folding

The ability of proteins to adopt their functional highly structured states in the intracellular environment during and after synthesis is one of the most remarkable evolutionary achievements of biology. In recent years, our understanding of the mechanisms of the protein self-organization process has increased dramatically. This understanding has been achieved due to the application of a variety of novel experimental and theoretical approaches to this complex task [1–6]. In this part of article we will focus on such protein folding-related problems as (i) polymer aspects underlying protein folding and structural peculiarities of the unfolded state and folding intermediates; (ii) the mechanisms of early stages of protein folding and (iii) peculiarities of protein amino acid composition favoring formation of equilibrium partially folded intermediate(s).

Polymer aspects of protein folding

Proteins as polymer homologues

It is generally accepted that the protein molecule has a unique primary sequence which governs its 3D structure and ensures proper biological activity. In this respect, each protein represents a unique case. That is why understanding the effect of sequence variations on biological performance represents a difficult challenge. It is also thought that natural polypeptides have originated as random copolymers of amino acids and were just evolution-

ary adjusted or 'edited' (based on the principle of natural selection) to acquire and refine their various 3D structures and functional properties [7–9]. Since protein molecules are remarkably unique, a serious question arises regarding the existence of any general features in protein self-organization. To understand the common physicochemical principles underlying the protein folding process, it is important to delineate common polymer roots and their impact on protein structures. The traditional way of performing such an analysis is to determine the correlation between different physical characteristics of a polymer (e.g. its molecular density) and its length.

Recently, molecular dimensions of 180 proteins in a variety of conformational states have been analyzed in order to establish a potential correlation between hydrodynamic dimensions and the length of polypeptide chain [10–13]. Protein categories analyzed include native globular proteins with nearly spherical shapes; equilibrium molten globules; compact denatured (or premolten globule) states in the presence of strong denaturants; denaturant-unfolded proteins without cross-links; denaturant-unfolded proteins with preserved disulfide bridges; natively unfolded proteins (see below). Figure 1 represents the results of this analysis and clearly shows that in all cases an excellent correlation between the apparent molecular density (determined as $\rho = M/(4\pi R_S^3/3)$, where M is a molecular mass and R_S is a hydrodynamic radius of a given protein) and molecular mass is observed. This correlation gave rise to a set of the standard equations, $R_S = K_h M^{\varepsilon}$, for a number of conformational states of a polypeptide chain [13]:

$$R_S \text{ (native)} = (0.75 \pm 0.05)\, M^{(0.33 \pm 0.02)} \quad (1)$$
$$R_S \text{ (MG)} = (0.90 \pm 0.10)\, M^{(0.33 \pm 0.02)} \quad (2)$$
$$R_S \text{ (pre-MG)} = (0.60 \pm 0.10)\, M^{(0.40 \pm 0.02)} \quad (3)$$
$$R_S \text{ (coiled)} = (0.28 \pm 0.02)\, M^{(0.49 \pm 0.01)} \quad (4)$$
$$R_S \text{ (8 M urea)} = (0.22 \pm 0.01)\, M^{(0.52 \pm 0.01)} \quad (5)$$
$$R_S \text{ (6 M GdmHCl)} = (0.19 \pm 0.01)\, M^{(0.54 \pm 0.01)} \quad (6)$$

Importantly, analysis showed that for a given conformational state, parameters K_h and ε were invariable over a wide range of chain lengths. Furthermore, statistical analysis has revealed that the relative errors of the recovered approximations exhibit random distribution over the wide range of chain lengths and do not generally exceed 10% [13]. This means that the effective protein dimensions in a variety of conformational states can be predicted based on the chain length with an accuracy of 10%. Thus, regardless of the differences in amino acid sequences and biological functions, protein molecules behave as polymer homologues in a number of conformational states.

It is important to remember that the most unambiguous characteristic of the conformation of a polymer molecule remains the molecular density. For instance, the density of a globule is expected to be independent of chain length, whereas the density of a partially collapsed or swelled macromolecule depends on both the chain length, and therefore on its molecular weight M, and on the nonspecific interactions of the monomer units with the solvent [14]. Keeping this in mind, let us try to analyze the data presented above.

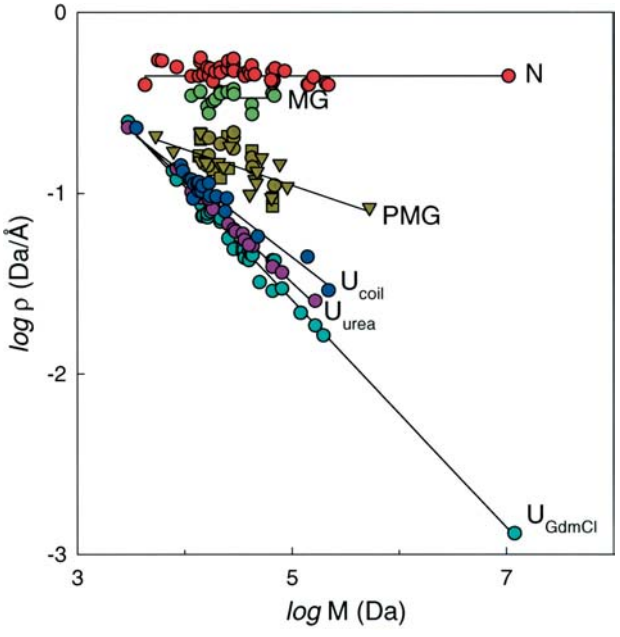

Figure 1. Variation of the density of protein molecules, ρ, with protein molecular weight, M, for a number of thermodynamically stable conformational states: native, molten globule; premolten globules (intermediates accumulated during the unfolding by urea or GdmCl are shown by circles; proteins with intact disulfate bridges in 8 M urea or 6 M GdmHCl are shown as squares; native premolten globules are shown as reversed triangles); native coils; proteins without cross-links or with reduced cross-links unfolded in 8 M urea; proteins without cross-links or with reduced cross-links unfolded in 6 M GdmHCl. The proteins used for this analysis are from [10–13]. The solid lines represent the best fit of the data to the standard function $R_S = K_h M^{\varepsilon}$.

Unfolded state

Obviously, to be able to solve the problem of protein folding, it is necessary to first characterize the unfolded state, which represents the starting point of the folding reaction. It is known that the unfolded state represents an ensemble of rapidly interchanging conformations, some of which are extended, and some more compact. It is possible that when stabilizing interactions occur, they induce a more populated ensemble of chain conformations, and, if such structures exist in the unfolded state, they would probably guide the folding process and function as folding-initiation sites [15]. In this respect, it is important to mention that theoretical studies have shown that small preferences for nativelike interactions in the unfolded state will substantially increase the probability of reaching the native state [16].

Coming back to the polymer roots, under conditions known as 'ideal' or 'Θ conditions', i.e. when the attractions of the macromolecular segments are balanced by those with the solvent, the density of macromolecules is expected to follow $M^{-0.5}$, thereby, the $R_S \sim lN^{0.5}$ with l being a statistical chain length [14, 17, 18]. Here, the polymer is assumed to be in a random coil conformation, and its conformational behavior can be described with the Gaussian statistics [14]. Further, in a good solvent, the macromolecular coil is expanded due to the prevalence of the repulsive interactions between polymer segments, and the molecular dimensions change more significantly with increasing chain length, $R_S \sim (l^2B)^{0.2} N^{0.6}$.

Note that the 'fully' unfolded states induced by the GdmHCl or urea provide $K_h = 0.19$ and 0.22, respectively, whereas $\varepsilon = 0.54$ and 0.52 were obtained [13]. Given that the recovered ε values are less than 0.6, it appears that these unfolded protein chains exhibit features of macromolecular coils in Θ solvents. This observation is in a good agreement with the results of earlier studies [17, 19], where, in particular, it was shown that the size of unfolded proteins in 6 M GdmCl is in reasonable agreement with the random coil model [17]. However, later it was pointed out that the inclusion of 'knots' of collapsed structure into the random coil model would not have a great influence on the hydrodynamic dimensions of a coil. Furthermore, values which Tanford [17] measured for the unfolded proteins correspond better to a model where 20% of the residues are located in the collapsed structures [20].

In agreement with this observation, a great many protein-folding studies revealed the presence of an assured residual structure even under the most severe denaturing conditions, such as high concentrations of strong denaturants. For example, considerable residual structure involving both α-helical and β-structural elements has been detected in the staphylococcal nuclease (SNase) [21–27]. A stretch of continuous nonpolar residues ranging between Ile 95 and Tyr 103 forming β-strand 3 was shown to be involved into the formation of a compact cluster in the unfolded state of the tryptophan synthetase α subunit [28, 29]. Similarly, a contiguous stretch of nonpolar residues comprising Val 54, Val 56, Trp 58, and Leu 59 was found to form a cluster in a urea-unfolded fragment of the protein 434 [30]. Investigations of the SH3 have identified turnlike structure in regions that fold into β strands [31, 32]. Nuclear magnetic resonance (NMR) analysis of urea-induced unfolding of the molten globule state of α-lactalbumin revealed a remarkably stable part of the protein in the core of the helical domain, which comprises interactions in both the N and C terminal parts of the protein [33]. The existence of defined residual structure has been observed in the unfolded state of varnase [34]. The WW domain retains a nativelike core in high concentrations of GdmCl and urea [35]. The unfolded state of bovine pancreatic trypsin inhibitor (BPTI) has been reported to be relatively compact in 6 M GdmCl under reducing conditions [36, 37]. Human carbonic anhydrase II, HCA II, has also been shown to possess a compact unfolded state [38–40]. Unfolded apomyoglobin (8 M urea at pH 2) displays distinct regions with dramatically different backbone mobility, suggesting the existence of residual structure [41]. Extensive clusters of hydrophobic structure exist in the unfolded state of lysozyme even under strongly denaturing conditions [42]. All these facts unambiguously show that the existence of profound residual structure might be a general characteristic of unfolded polypeptide chains under aggressively denaturing conditions [43–47]. In other words, unfolded states of proteins exhibit behavior that is not random coil in nature, which is not surprising considering the complexity of polypeptides. In fact, it has been pointed out that a total lack of intraresidue interactions would be unexpected in the unfolded state because certain (e.g. hydrophobic) side chains have high affinity for each other in a folded protein [40]. In addition, some secondary structure within unfolded proteins could be expected due to the preferential distribution of phi and psi angles [19, 48, 49], and some residual hydrophobic interactions may also be present [40, 50]. All this restricts considerably the conformational space of the unfolded polypeptide chain. Thus, it seems most likely that the polypeptide chains under strong denaturing conditions are still below the critical point (poor solvent conditions), and can be easily transformed to the compact state. For example, small fluctuations of temperature from 25 to 30 °C were shown to encourage cooperative collapse of fully unfolded proteins in 6 M GdmCl [9]. Thus, globular proteins are never random coils without positional correlations, and biological polypeptide chains represent macromolecular coils below a critical point even under the harsh denaturing conditions.

Premolten globule state

Now let us consider a situation when the thermodynamic quality of the solvent worsens. In this case binary interactions between the monomers become mainly attractive [14]. As a result, the polymer chain partially collapses, leading to an increase in the molecular density and bringing many-body interactions into scene. It has been found that under appropriate conditions, many proteins can form a specific compact denatured conformation, a premolten globule state [10, 13, 51–60]. Major structural characteristics of this intermediate are summarized below. This conformation is characterized by considerable secondary structure, although much less than that of the molten globule. The premolten globule state is considerably less compact than the molten globule, but it is still more compact than the corresponding random coil. Premolten globules can interact with the hydrophobic fluo-

rescent probe ANS, although more weakly than in the molten globule state. This means that at least some hydrophobic clusters are already formed in the premolten globule state, although there is no globular structure. Finally, the premolten globule and the molten globule are separated by an all-or-none phase transition, reflecting the fact that these partially folded intermediates represent discrete phase states. Importantly, several structural elements of these squeezed coils may occupy nativelike positions [10]. We assume that this compact denatured state might represent a general intermediate in protein folding. Obviously, the existence of such a state substantially reduces any search through the conformational space, ensuring rapid folding. Given that this state might comprise a specific nativelike core with burial of hydrophobic residues, the transition from this state to the molten globule one or to the native state would not require significant energy changes and could occur quite easily. It may well be the case that protein folding in vivo involves mostly molten globule and denatured compact states. In this occasion, it would require only slight changes to make the coil 'condense onto itself' and form a globule.

Analysis of hydrodynamic data reveals that the molecular dimensions of premolten globules follow the chain length as $R_S = 0.6\ M^{0.40}$ (see above). This indicates poor solvent conditions. Thus, one can conclude that this conformation exhibits behavior which is typical for squeezed macromolecular coils. Therefore, any small variations in the protein environment, i.e. changes in the thermodynamic quality of the solvent or changes induced by proton transfer, interactions with a ligand, fluctuations of temperature and so on, can trigger transition of the compact protein molecule to the more rigid molten globule or native state [14]. Importantly, fig. 1 shows that stabilization of the premolten globule state might be achieved by incorporation of disulfide bridges into the protein sequence. In fact, the hydrodynamic behavior of unfolded proteins with intact disulfide bridges was shown to be close to that of premolten globules. Finally, the preexponential term $K_h = 0.6$ retrieved for the premolten globule state is significantly larger than those retrieved for fully unfolded species, indicating, most likely, developing multiple body interactions.

Globular states – native and molten globule

The theory of 'coil-globule' transition predicts that the overall dimension of a polymer globule, R_S, is anticipated to change with the chain length, N, as $R \sim (C/B)^{1/3}N^{1/3}$. Here, B and C are the second and the third virial coefficients which characterize the pair collisions and three-body interactions of the monomer units of the polymer chain [14]. The density of the globules is expected to show no change with increasing chain length, owing to $\rho \sim N/R^3 \sim (-B/C)$. These results are in excellent agreement with the data obtained for the native and molten globules of proteins (fig. 1, see also [10–13]). In particular, it has been found that for native and molten globule proteins, the parameter K_h has a value of 0.75 and 0.9, respectively, whereas ε equals 0.33 for both conformational states. Given that the overall dimension of globules is determined by the balance of pair collisions and three-body interactions, the difference in preexponential terms K_h observed for the native and molten globules most likely reflects the larger probability of three-body interactions in the molten globule state. This could be because of the compact but flexible nature of the molten globules.

Concluding remarks

Regardless of the differences in primary amino acid sequences, protein molecules behave as polymer homologues in a number of conformational states, allowing speculation about volume interactions being a driving force in a formation of equilibrium structures. For instance, both native and molten globules exhibit key features of polymer globules, where fluctuations in molecular density are expected to be much less than the molecular density itself. Protein molecules in the premolten globule state possess properties of squeezed coils. Furthermore, even high concentrations of strong denaturants are more likely to constitute poor solvents for protein chains. Thus, globular proteins are never random coils without positional correlations, and biological polypeptide chains represent macromolecular coils below a critical point, even under harsh denaturing conditions.

Early stages of protein folding

Although understanding of initial events in protein folding is an important prerequisite to solve the problem of protein self-organization, these early folding events are far from being understood. What is the first step in protein folding, hydrophobic collapse (compaction) or secondary structure formation? Furthermore, it is still not clear whether the major driving force in protein folding is hydrogen bonding or hydrophobic interactions, or both. As early as 1937 it was suggested that protein folding is directed mostly by formation of intramolecular hydrogen bonds [61]. This gave rise to the secondary structure framework model. According to this model, the polypeptide chain undergoes local folding to nativelike elements of secondary structure, which direct subsequent folding through diffusion and collision processes [62, 63]. On the other hand, it has been pointed out that hydrophobic attraction could be the dominant force governing protein folding, due to the fact that hydrogen bonding to the solvent molecules would strongly favor the unfolded state [64]. This hypothesis gave rise to the hydrophobic collapse model, according to which strong hydrophobic attractions should first lead to nonspecific compaction of the polypeptide chain into a structureless globule. This

step significantly reduces the conformational space and consequently facilitates formation of secondary and tertiary structure [65–68].

Equilibrium partially folded conformations are usually considered as stable counterparts of kinetic intermediates transiently populated during protein refolding kinetics [69]. Many globular proteins have been shown to adopt several stable conformations, e.g. the native, molten globule, premolten globule and unfolded states [9, 10, 13, 51–60]. In an attempt to elucidate the mechanism involved in the initial events of protein folding, the literature on the conformational characteristics of 41 globular proteins in native and partially folded conformational states has been analyzed [70]. Figure 2 represents results of this analysis as $(R_S^U/R_S)^3$ (relative compactness) vs. $[\theta]_{222}/[\theta]_{222}{}^U$ (relative content of ordered secondary structure), and clearly shows that the degree of compactness and amount of ordered secondary structure are highly correlated parameters. In other words, no compact equilibrium intermediates lacking secondary structure, or highly ordered noncompact species were found among the 41 proteins from the data set. In the context of the early stages of protein folding this suggests that hydrophobic collapse and formation of secondary structure occur simultaneously, rather than representing two independent and sequential processes.

As has been already mentioned, it is now becoming more and more evident that when a protein unfolds, not all of its structure is lost [26, 30, 43, 47, 71–80]. Moreover, it has been shown that long-range order and nativelike spatial positioning and orientation of chain segments are present, even in concentrated solutions of strong denaturants. Thus, even the unfolded expanded polypeptide chain may have significant nativelike topology under harsh denaturing conditions [10, 26]. This means that the unfolded protein is predisposed to adopt specific backbone conformations rather than to become a random coil, as Flory postulated in his isolated pair hypothesis [81]. Therefore, favored backbone conformations already preexist in the denatured protein, effectively restricting the number of states accessible to the unfolded polypeptide chain [46]. Consequently, protein folding does not represent a random search for the favorable native structure throughout the enormously large conformational space. Instead, it can be considered as a directed run within a rather narrow conformational corridor, in which the 'cementing' of the preexisting short- and long-range contacts occurs. In this view, the discussion of which happens

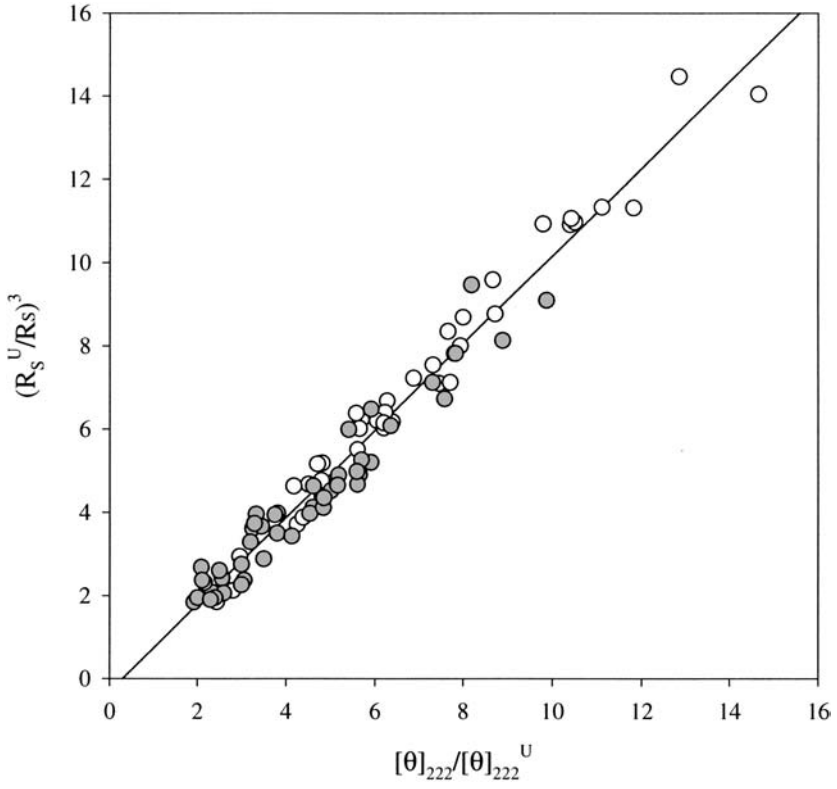

Figure 2. Correlation between the degree of compactness (calculated for different conformational states as the decrease in hydrodynamic volume relative to the volume of the unfolded conformation) and amount of ordered secondary structure (calculated for different conformational states from their far – UV CD spectra as the increase in negative ellipticity at 222 nm, $[\theta]_{222}$, relative to that of the unfolded conformation). Open and gray symbols correspond to the data for native globular proteins and their partially folded intermediates, respectively. Figure is modified from the data presented in [70].

first – compaction or secondary structure formation – is no longer relevant. The only acceptable scenario of the earliest stages of protein folding would be that hydrophobic collapse happens simultaneously with the formation of secondary structure [70].

Interestingly, this conclusion agrees well with the data published recently by Fernandez et al. [82]. In fact, it has been shown that soluble globular proteins of moderate size ($N < 102$) deposited in the protein data bank (PDB) are characterized by a commensurable relationship between hydrophobic surface burial and number of backbone hydrogen bonds. Furthermore, an analysis of ~50,000 conformations from the all-atom simulation of villin headpiece showed that not only the hydrophobic collapse is concurrent with the formation of backbone amide-carbonyl hydrogen bonds, they are also dynamically coupled processes [82]. It has been concluded that hydrophobic clustering of the side chains is inevitably bound to backbone burial, and the latter process becomes thermodynamically too costly and kinetically unfeasible without amide-carbonyl hydrogen-bond formation. Furthermore, the desolvation of most hydrogen bonds is exhaustive along the pathway, implying that such bonds guide the collapse process [82].

Amino acid composition favoring equilibrium partially folded intermediate(s)

A variety of different physicochemical forces play a role in stabilizing the unique 3D structure of a protein. Both the strength and specificity of many of these forces are strongly dependent on environmental conditions in such a way that changes in the environment can reduce or even eliminate part of the conformational interactions, while the remainder are unchanged or even intensified. Under some environmental conditions, native protein structure can be transformed into new conformations with properties intermediate between those of the native and completely unfolded states. Thus, the ability of a protein to adopt different stable partially folded conformations should be considered as an intrinsic property of a polypeptide chain. Since all the necessary and sufficient information to fold into the native, biologically active conformation is thought to be present in protein amino acid sequence [83], the capability of a given protein to adopt equilibrium partially folded conformation(s) may also be encoded in specific features of its amino acid sequence.

Interestingly, it has been shown that not all proteins (even homologous ones) have an identical response to changes in their environment. For example, hen egg white lysozyme represents a textbook illustration for the two-state model of denaturant-induced unfolding [17], whereas accumulation of classical molten globule under different experimental conditions was described for its homologue α-lactalbumin [84]. Analysis of the literature on equilibrium unfolding of globular proteins induced by changes in pH, temperature or strong denaturants (urea or guanidinium chloride) revealed that unfolding in 115 proteins is accompanied by accumulation of equilibrium intermediate states of one sort or another. Another set comprises of 39 proteins, which were shown to unfold according to a simple two-state model; i.e. no equilibrium intermediate of any kind was formed during their unfolding [85].

In an attempt to understand which factors may be responsible for such tremendous differences in the formation of equilibrium partially folded intermediates, the general sequence features of proteins from both groups have been analyzed using a simple method comparing global sequence charge and hydrophobicity [85]. Figure 3 represents the results of this analysis as a plot of mean hydrophobicity vs. mean net charge, i.e. as a distribution of groups within the charge-hydrophobicity phase space. Figure 3 shows that reliable separation of both groups of proteins takes place in such coordinates. These data imply that the competency of a protein to form equilibrium intermediate(s) may be predetermined by the bulk content of hydrophobic and charged amino acid residues [85]. In other words, this competency may be encoded in the charge/hydrophobicity ratio of its polypeptide chain, not its sequence. This may mean that partially folded conformations are stabilized mostly by nonspecific, side chain-side chain interactions of hydrophobic amino acid residues [85]. Interestingly, proteins that do not have equilibrium intermediates are less hydrophobic and have, in general, a larger net charge than those competent to form discrete intermediate states. This may indicate that such proteins are less strengthened by hydrophobic inter-

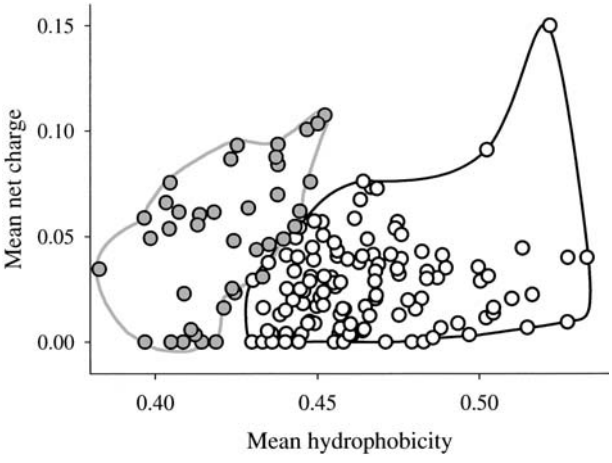

Figure 3. Comparison of mean net charge vs. mean hydrophobicity for the set of 115 proteins able to form equilibrium intermediates (open symbols and black lines) and the set of 39 proteins shown to unfold without accumulation of partially folded conformations (gray symbols and lines). Figure is modified from the data presented in [84].

actions and more disturbed by electrostatic repulsion. Thus, smaller environmental changes may be required to overcome the marginal stabilization energies leading to immediate and complete unfolding of the protein.

Misfolding

Molecular mechanisms of deadly protein misfolding

As has been already noted, the sequences of proteins have evolved in such a way that their unique native states can be found very efficiently even in the complex environment inside a living cell. However, under some conditions, proteins fail to fold properly or to remain correctly folded; this misfolding can lead to the development of different pathological conditions. A number of human diseases, including the amyloidoses and several neurodegenerative disorders, such as Alzheimer's disease, Parkinson's disease, transmissible spongiform encephalopathies and so on, originate from the deposition of stable, ordered and filamentous protein aggregates, commonly known as amyloid fibrils. In each of these pathological states, a specific protein or protein fragment changes from its natural soluble form into insoluble fibrils, which accumulate in a variety of organs and tissues [86–91]. Approximately 20 different proteins are known to be involved in the amyloidoses (extracellular deposits), which are unrelated structurally or at the level of primary structure. In addition, a number of diseases also originate from the deposition of fibrillar proteins, but within cells, i.e. intracellular deposits.

Importantly, prior to fibrillation, amyloidogenic polypeptides may be rich in β sheets, α helices, β-helices or contain both α helices and β sheets. They may be well folded or belong to a newly discovered class of natively unfolded (or intrinsically unstructured) proteins. Despite these differences, the fibrils from different pathologies display 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 [92]. Amyloid fibrils have been shown to form in vitro from disease-associated [93–97], as well as from disease-unrelated proteins and peptides [98–115]. Moreover, there is an increasing belief that the ability to from fibrils is a generic property of the polypeptide chain, i.e. many proteins, perhaps all, are able to form amyloid fibrils under appropriate conditions [87, 106, 115]. If so, this would dramatically extend the structural diversity of polypeptide chains able to fibrillate.

Since all fibrils independent of the original structure of the given amyloidogenic protein have a common cross-β structure, considerable conformational rearrangements have to occur prior to fibrillation. Such changes cannot happen in a native protein, due to its stable and rigid tertiary structure. Thus, protein destabilization favoring partial unfolding and culminating in the formation of a partially unfolded conformation is required. Presumably, such a partially unfolded conformation favors reciprocal and specific intermolecular interactions, including electrostatic attraction, hydrogen bonding and hydrophobic contacts, which are necessary for oligomerization and fibrillation [86–91, 116–119].

Obviously, this model does take into account a class of natively unfolded proteins, as they are devoid of rigid tertiary structure in their native state. The primary step in the fibrillation of these proteins represents partial folding, i.e. stabilization of a partially folded conformation [97, 114, 120–122]. Thus, the general hypothesis of fibrillogenesis states that structural transformation of a polypeptide chain into the partially folded conformation represents an important prerequisite for successful protein fibrillation.

A question then arises about the nature of the amyloidogenic intermediate(s). It has been already mentioned that the conformational space of globular proteins involves four general conformations (see above). These include the native state, unfolded states and two classes of partially folded intermediates. The latter are compact, relatively well folded intermediates with substantial native-like secondary structure but little tertiary structure, often referred to as molten globules, and significantly less compact, relatively unfolded intermediates, with substantially less secondary structure, often known as premolten globules. Potentially, either of these latter conformations may play a role as the crucial amyloidogenic species. Detailed structural analysis of early fibrillation events in several proteins has demonstrated that the amyloidogenic conformation is only slightly folded and shares many structural properties with the premolten globule state.

Fibrillogenesis of globular proteins; requirement for partial unfolding

Data have been reported indicating that the first critical step in protein fibrillogenesis is the partial unfolding of the protein. Due to structural fluctuations (conformational breathing) the structure of a globular protein under physiological conditions represents a mixture of tightly folded and multiple partially unfolded conformations, with great prevalence of the former. Most mutations associated with accelerated fibrillation and protein deposition diseases have been shown to destabilize the native structure, increasing the steady-state concentration of partially folded conformers [86–91, 123–130]. Conversely, it has been shown that the amyloidogenicity of a protein can be significantly reduced by stabilization of the native structure, e.g. via specific binding of ligands [131–133].

The fibrillation-provoking destabilization of a rigid protein is achieved in vitro at low or high pH, high tempera-

tures, low-to-moderate concentrations of strong denaturants, organic solvents and so on. This fact is well illustrated in recent studies on the fibrillogenesis of transthyretin (TTR), also known as prealbumin, which is a homotetramer composed of 127 amino acid subunits characterized by 2,2,2 molecular symmetry. TTR is found in human plasma (0.1–0.4 mg/ml) and cerebral spinal fluid (0.017 mg/ml), with the plasma form being the amyloidogenic precursor. Wild-type TTR amyloidogenesis may cause senile systemic amyloidosis, characterized by deposition and pathology in the heart after age 60 [59]. Early onset amyloid formation (as early as the 2nd decade) by one of more than 80 single-site TTR variants provokes the number of diseases collectively termed familial amyloid polyneuropathy [134]. TTR can be converted into amyloid in vitro by acid-mediated dissociation of the homotetramer into monomers. The pH required for disassembly also results in tertiary structural changes within the monomeric subunits, finally leading to the enhanced fibrillation [135]. Recently, in a quest to understand the relationship between the tertiary structural changes and amyloidogenicity, a monomeric mutant with nativelike structure and stability has been designed, which was nonamyloidogenic, unless partially unfolded [136].

Light chain, or AL, amyloidosis is a pathological condition arising from systemic extracellular deposition of monoclonal immunoglobulin light-chain variable domains in the form of insoluble amyloid fibrils, especially in the kidneys [137]. Structural and fibrillation properties of one of the amyloidogenic light-chain variable domains, SMA, have been analyzed under a variety of conditions [138]. The results of biophysical analysis revealed that a decrease in pH resulted in the accumulation of two partially folded intermediates. A relatively nativelike intermediate, I_N, was observed between pH 4 and 6, and was characterized by little loss of secondary structure, combined with significant changes in tertiary structure and enhanced 1-anilino-8-naphthalenesulfonate (ANS) binding. At pH below 3, a relatively unfolded, but compact, intermediate, I_U, with decreased tertiary and secondary structure, was observed. The I_U intermediate readily forms amyloid fibrils, whereas I_N preferentially leads to amorphous aggregates [138]. Comparable data have been recently reported for another light-chain variable domain, LEN [113, 139].

α-Lactalbumin, or α-LA, is a small acidic protein with a single Ca^{2+} binding site. It is very attractive for studies of partially folded conformations since it adopts the classic molten globule conformation at acidic pH, moderate guanidinium-chloride concentrations or elevated temperatures (apo form) [140]. α-LA comprises of a large α-helix domain and a small β-sheet domain connected by a calcium-binding loop and four disulfide bridges [140]. α-LA forms amyloid fibrils at low pH. *S*-Carboxymethyl-α-lactalbumin, a disordered form of the protein with three of the disulfide bridges reduced, was even more susceptible to fibrillation. *S*-Carboxymethyl-α-lactalbumin exhibits the properties of a premolten globule, and its fibrillation is orders of magnitude faster than when starting with the molten globule conformation [115]. Other partially folded conformations induced in α-LA at neutral pH, either by removal of Ca^{2+} or by binding of Zn^{2+} to the Ca^{2+} protein, did not fibrillate, although Zn^{2+}-loaded α-lactalbumin precipitated out of solution as amorphous aggregates. Based on these data it was concluded that transformation from native state to a substantially unfolded conformation is required for successful fibril formation, whereas less unfolded species may form amorphous aggregates [115].

Finally, fibrillation of bovine β-lactoglobulin represents another illustrative example of the importance to being unfolded [141]. It has been shown that this protein will form fibrils in urea solutions; the process is denaturant concentration dependent, showing the highest efficiency in the vicinity of 5 M urea, which corresponds roughly to the C_m value. Importantly, it was shown that molten globule-like intermediate was not accumulated during the urea-induced unfolding of β-lactoglobulin [141]. Thus, it has been concluded that amyloid fibril formation by bovine β-lactoglobulin is promoted under conditions where significant accumulation of unfolded protein occurs, but is inhibited under conditions where higher denaturant concentrations destabilize intermolecular interactions [141]. Generally speaking, it has been concluded that amyloid formation in vitro can be achieved by destabilizing the native state of the protein under conditions in which noncovalent interactions still remain favorable [98, 99, 103, 106, 141, 142].

Fibrillogenesis of natively unfolded proteins; requirement for partial refolding

We now consider details of the fibrillogenesis of intrinsically unstructured proteins, which constitute a significant portion of known amyloidogenic proteins. It seems reasonable to assume that such proteins are well suited for amyloidogenesis as they lack significant secondary and tertiary structure, as well as many specific intrachain interactions. In the absence of such conformational constraints they would be expected to be substantially more conformationally flexible, and thus able to polymerize more readily than tightly packed globular proteins. Substantial evidence suggests that the earliest stage of fibrillation of these proteins is their partial refolding. The subsequent section represents several illustrative examples of fibril formation of natively unfolded proteins.

α-Synuclein is a small (14 kDa), soluble, intracellular, highly conserved protein that is abundant in various regions of the brain. This protein has been estimated to account for as much as 1% of the total protein in soluble cy-

tosolic brain fractions. Structurally, purified α-synuclein is a typical natively unfolded protein [97]. Several observations led to the conclusion that this presynaptic protein might be involved in the pathogenesis of Parkinson's disease, PD. α-Synuclein was shown to be a major fibrillar component of Lewy bodies and Lewy neurites. Two different missense mutations in the α-synuclein gene, corresponding to A53T and A30P substitutions in α-synuclein, have been identified in two kindreds with autosomal-dominantly inherited, early-onset PD. Furthermore, the production of WT or A53T α-synuclein in transgenic mice, or of WT, A30P and A53T in transgenic flies leads to the motor deficits and neuronal inclusions reminiscent of PD. Interestingly the peptide derived from the central hydrophobic region of α-synuclein represents a second major intrinsic constituent of Alzheimer's plaques. This 35-amino acid peptide, known as NAC, was shown to amount to about 10% of the amyloid plaque. These observations indicate that α-synuclein is a key player in the pathogenesis of several neurodegenerative disorders. The fibrillogenesis of this protein is probably the most precisely studied and best understood among other amyloidogenic members of the family of natively unfolded proteins. In particular, accumulated data strongly suggest that formation of a partially folded intermediate (possessing the major characteristics of the premolten globule) represents the critical first step of α-synuclein fibrillogenesis. This partially folded intermediate can be stabilized by numerous factors, including high temperatures, low pH [97], the presence of several pesticides [143, 144], or metal ions [145], or at moderate concentrations of trimethylamine-*N*-oxide [146] or other organic solvents [147]. Importantly, under all these conditions α-synuclein was shown to undergo significantly enhanced fibrillation. In contrast, fibril formation was considerably slowed or inhibited under conditions favoring formation of more folded conformations [146, 147], or by stabilization of the fully unfolded form, e.g. by oxidation of its methionines [148].

Pancreatic islet β cells produce, in addition to insulin, a peptide called amylin or islet amyloid polypeptide, IAAP [149]. Amylin has several functions associated with the normal regulation of fuel metabolism. Dysfunction of amylin due to mutation, or β-cell dysfunction resulting in amyloid formation, has been associated with the development of non-insulin-dependent diabetes mellitus. Amylin is an unstructured peptide hormone of 37 amino acid residues. Human amylin and its 8–37 fragment were shown to polymerize under physiological conditions. The process of polymerization was very fast (lag times were 100 and 50 min for full-length amylin and its 8–37 fragment, respectively) and resulted in the appearance of amyloid fibrils [93]. Interestingly, both peptides showed formation of a partially folded (premolten globule-like) intermediate(s) early in the fibrillation process. It takes ~90 min for full-length amylin to form such an intermediate, whereas this period was almost twice as short for the truncated peptide, showing excellent agreement with fibrillation lag times [93].

Alzheimer's disease, known as AD, is the most prevalent age-dependent dementia. AD is characterized pathologically by accumulation of extracellular amyloid deposits in the cerebral neuropil and vasculature, and of intracellular neurofibrillary tangles. Amyloid deposits contain the amyloid β protein (Aβ), which is a 40–42 residue peptide produced by endoproteolytic cleavage of the amyloid β-protein precursor (APP). Many lines of evidence support the crucial role of Aβ in AD. Fibrillation of Aβ is associated with the development of the cascade of neuropathogenetic events, ending with the appearance of cognitive and behavioral features typical of AD. Aβ appears to be unfolded at the beginning of the fibrillation under physiological conditions. NMR studies have shown that monomers of Aβ(1–40) or Aβ(1–42) possess no α-helical or β-sheet structure [150]; i.e. they exist predominately as 'random' extended chains. Partial refolding to the premolten globule-like conformation has been detected at the earliest stages of the Aβ amyloidosis [151].

Prothymosin α is a 109-residue protein. It is very acidic, containing ~50% aspartic and glutamic acid, no aromatic or cysteine residues, and very few large hydrophobic aliphatic amino acids [152]. Because of these features, prothymosin α, or pTα, is in a random coil-like conformation with no regular secondary structure under the conditions of neutral pH [152, 153]. However, pTα adopts a partially folded conformation at acidic pH [153]. Interestingly, it has recently been established that at low pH (below pH 3; i.e. under conditions favoring the formation of the premolten globule-like conformation [153]), prothymosin α is capable of fast formation (lag time ~100 min) of regular elongated fibrils with a flat ribbonlike structure 4-5 nm in height and 12–13 nm in width [114].

Conformational prerequisites for amyloidogenesis

Fig. 4 compares parameters of far-ultraviolet (UV) circular dichroism (CD) spectra measured for 'pure' amyloidogenic conformations of 11 proteins (SH3 domain [98], cytocrome c_{552} [110], monellin [100], methionine aminopeptidase [106], SMA [138], α-lactalbumin [115], phosphoglycerate kinase, PGK [105], amylin [93], prothymosin α [114], Aβ [151] and α-synuclein [97]), with those retrieved for the four basic protein conformations, native, molten globule, premolten globule and unfolded states. In this plot data for the premolten globule and unfolded states are taken from [11], and the data for molten globules are from [154]. The set of native proteins comprises two parts: native forms of the globular proteins known to form amyloids readily (SH3 domain [98],

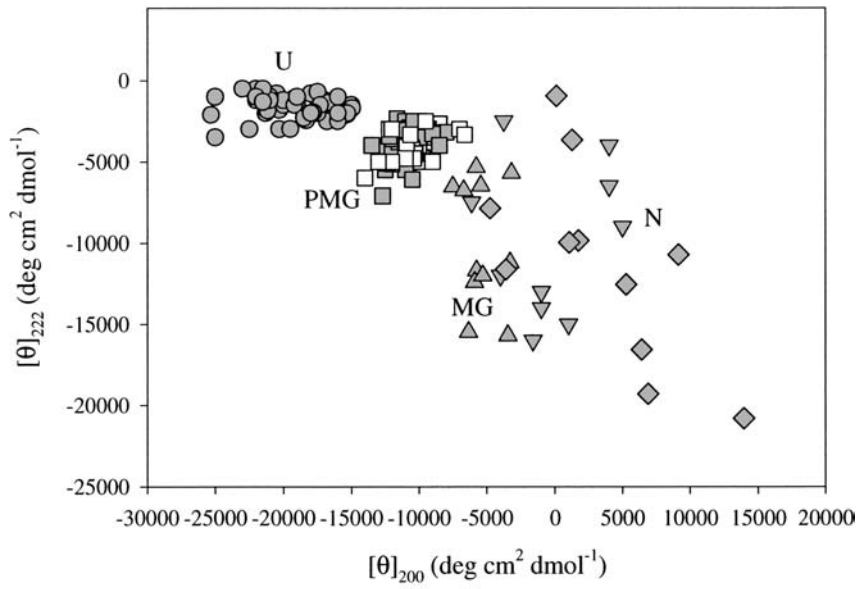

Figure 4. Comparison parameters of far-UV CD spectra measured for 'pure' amyloidogenic conformations of 11 proteins (open squares) with those retrieved for the four basic protein conformations, native, molten globule, premolten globule and unfolded states (see the text). Data for the unfolded states, premolten globules, and molten globules are shown by gray circles, squares and triangles, respectively. The set of native proteins comprises two parts, native forms of the amyloidogenic globular proteins (gray reversed triangles) and native forms of proteins able to form molten globules (gray diamonds). Modified from figure presented in [171].

cytocrome c_{552} [110], monellin [100], methionine aminopeptidase [106], SMA [138], PGK [105], human lysozyme [155], transthyretin [136], α-lactalbumin [115], prion protein [156], β-lactoglobulin [141]) and native forms of proteins able to form molten globules [154]. Figure 4 clearly shows that all amyloidogenic conformations are grouped with the set of data for the premolten globule proteins. This is an extremely important observation, which demonstrates the importance of this conformation for protein fibrillogenesis.

The central model of protein misfolding is illustrated by fig. 5, which shows that the process of fibrillation (which is used as an example of a more general phenomenon – aggregation) can be divided into three major steps: (i) structural transformation of a native soluble protein (rigid globular or flexible natively unfolded) into the 'sticky' amyloidogenic precursor intermediate, (ii) nucleation and (iii) fibril elongation and growth. Thus, transformation into the partially folded conformation is a critical early stage of fibrillogenesis and precedes the appearance of any aggregated material.

Polymer aspects of protein misfolding

Behavior of a given polymer in a given solution is determined by the peculiarities of polymer segment – solvent interactions. For example, the major reason for the appearance of globular conformation (in our particular case, we are talking about the correctly folded form of a 'normal' globular protein) in a poor solvent (water) is that this conformation effectively excludes a portion of segments from unfavorable contacts with the solvent and forms the shielding interface between the polymer interior and solvent. In turn, the stability of globular conformation also depends on the peculiarities of interactions between protein globule and solvent. Obviously, many factors may affect the efficiency of coil-globule transition (i.e. the efficiency and direction of the process of protein folding), as well as change the efficiency of the shield (interface between the polymer and solvent), and thus may modulate stability of a native protein molecule. Basically, point amino acid substitutions, changes in pH, temperature and numerous other environmental circumstances may considerably affect the mode of polymer-solvent interactions. Thus, protein misfolding (aggregation) may originate from the changes in relative quality of solvent, which appear either due to specific changes in protein amino acid composition or because of solvent composition modifications.

Multiple pathways of protein misfolding

Although the data presented above were mostly devoted to the consideration of protein fibrillation, the process of amyloid fibril formation does not represent the only misfolding route. In fact, contrary to the process of productive protein folding, leading to the appearance of rigid conformation with specific function, the end products of misfolding may have a different appearance. The morphology of these end products depends on the particular

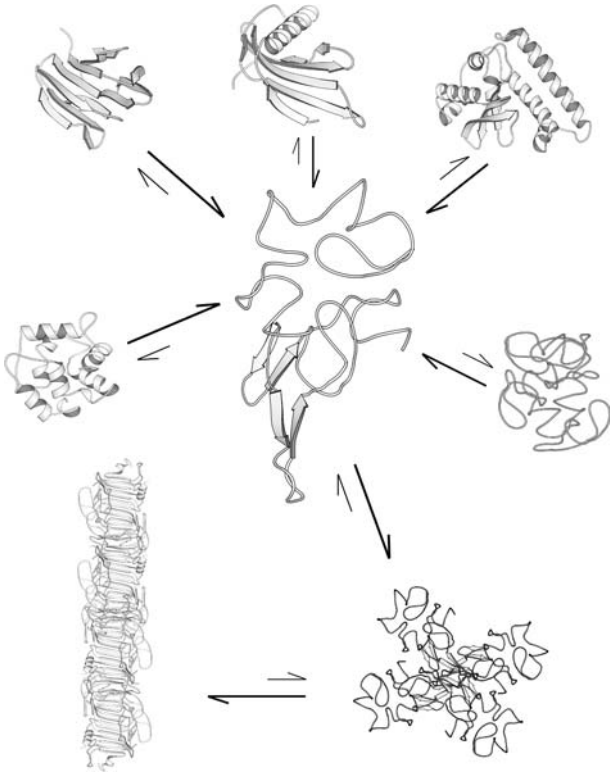

Figure 5. General model for the protein misfolding (fibril formation is considered as an example). Three general stages of the process include structural transformation of a native soluble protein (rigid or natively unfolded) into the 'sticky' amyloidogenic precursor intermediate, nucleation and fibril elongation/growth. Additional conformational changes may occur between the aggregation-competent intermediate and the fibrils. Structures at the top represent natively folded proteins with different structures. A common amyloidogenic intermediate for different rigid and natively unfolded proteins is shown for convenience only. Experimental data support the idea that amyloidogenic intermediates likely fall in the class of premolten globules. However, this still leaves a lot of room for structural diversity. The oligomeric intermediate (representing the nucleus or a soluble aggregate) is shown as a tetramer for convenience only, and could be much larger.

experimental conditions, and misfolded product may appear as soluble oligomers, amorphous aggregates or amyloid-like fibrils. Any of these three species could be cytotoxic, thus giving rise to the development of pathological conditions. The reason for such a morphological difference is potentially connected with the diversity of the partially folded intermediates favoring protein self-association. In fact, multiple environmental factors, such as point mutations, decrease in pH, increase in temperature, the presence of small organic molecules or metal ions, and other charged molecules, might induce structural rearrangements within a protein molecule, shifting equilibrium toward the partially folded conformation(s). As different factors may stabilize slightly different partially folded intermediates, the formation of morphologically different aggregates is expected. This idea is illustrated by fig. 6, which represents a model of α-synuclein misfolding. It has been shown that aggregation of this protein dramatically depends on the experimental conditions and might lead to the appearance of one of the three misfolded forms discussed above (soluble oligomers, amorphous aggregates or amyloid fibrils) or a mixture thereof.

Nonfolding

The phenomenon of natively unfolded proteins

It has been shown that a considerable number of proteins possess some amount of disorder and not rigid structure, under physiological conditions. A special term, 'natively denatured', was introduced in 1994 [122] to emphasize the existence of drastic structural differences between a 'normal' globular protein, with rigid tertiary structure, and an 'abnormal', extremely flexible, tau protein. Two years later, a new term, 'natively unfolded', originated as a result of conformational analysis of α-synuclein, which under physiological conditions appeared to be noncompact and lacked any secondary structure [157]. Two alternative terms, 'intrinsically unstructured' [158] and 'intrinsically disordered' [159], have also been suggested to describe these proteins. Since nonrigid proteins show an extremely wide diversity in their structural properties, the meaning of the above terms should be clarified. The terms 'denatured' and 'disordered' may be considered as synonyms, and indicate any set of nonrigid conformations of polypeptide chains, including different compact partially folded conformations: molten globules, premolten globules and coils. The terms 'unstructured' and 'unfolded', may be considered synonymous, and should only be applied to the subset of disordered proteins characterized by the absence of any (or almost any) ordered structure. Only natively unfolded proteins will be considered below, excluding the native molten globules.

Natively unfolded or intrinsically unstructured proteins constitute a unique tribe of the protein kingdom. The number of proteins and protein domains that have been shown in vitro to have little or no ordered structure under physiological conditions is rapidly amplifying. For example, the first list of natively unfolded proteins published in 2000 [160] contained 91 members (the full-length proteins and their domains with a chain length of more than 50 amino acid residues). Now this set has almost doubled. The intriguing family of intrinsically unstructured proteins has attracted an excessive attention of researchers over the past decade due to several reasons. First of all, the existence of natively unfolded proteins, with their lack of ordered rigid structure under physiological conditions in vitro, disputes the validity of one of the cornerstones in protein biology, chemistry and physics, i.e. the structure-function paradigm. Note that this concept, formulated more than 100 years ago as a lock-and-key model for explaining the amazing specificity of the enzy-

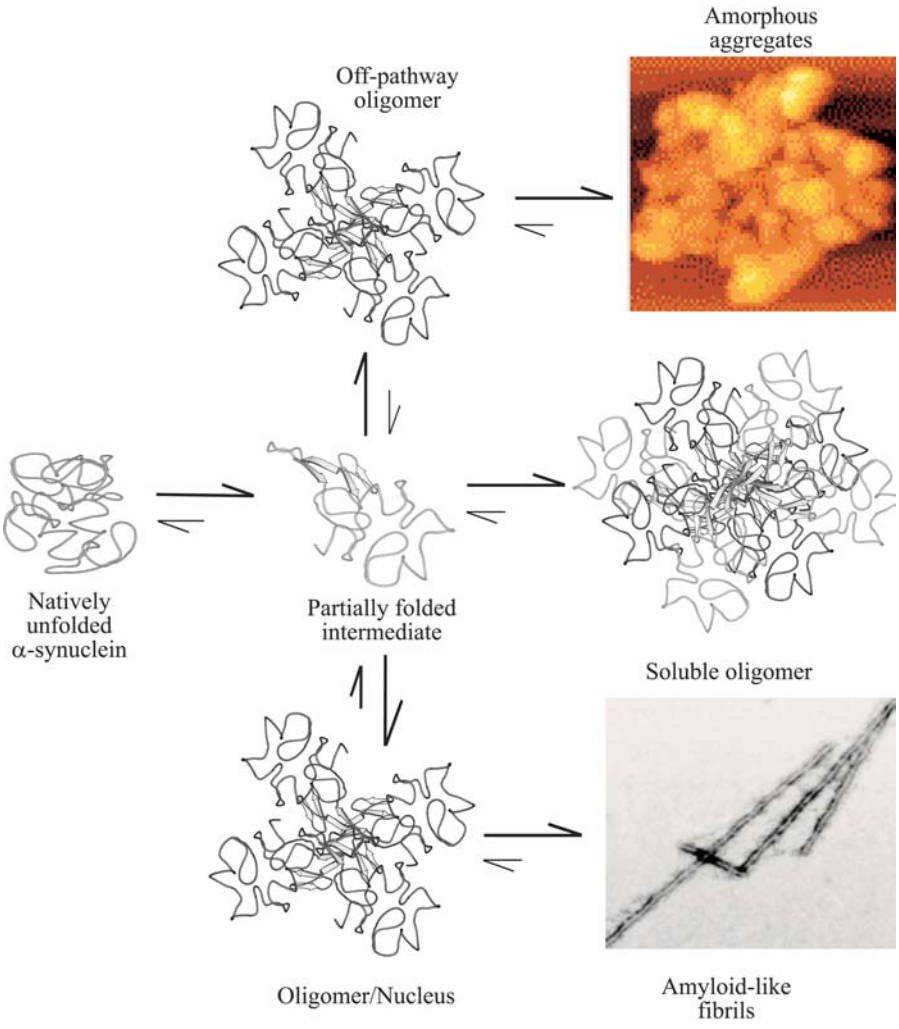

Figure 6. Multiple pathways of human α-synuclein aggregation.

matic hydrolysis of glucosides [161], claims that a specific function of a protein is determined by its unique and rigid 3D structure. Natively unfolded proteins, being unstructured, are involved in countless biological activities (see below), and effectively resist those evolutionary pressures that normally favor stable globular folds. For example, α-synuclein (one of the best-characterized members of the family) isolated from different organisms possesses a high degree of sequence conservation, e.g. mouse and rat α-synucleins are identical throughout the first 93 residues, and human and canary proteins differ from them by only two residues [162]. To overcome the conflict of the existence of functionally active unfolded polypeptides with the structure-function paradigm, it has been suggested that the lack of rigid globular structure under physiological conditions might represent a considerable functional advantage for natively unfolded proteins, as their large plasticity allows them to interact efficiently with several different targets [159, 163–170]. Furthermore, a disorder-to-order transition induced in natively unfolded proteins during the binding of specific targets in vivo might represent a simple mechanism for regulation of numerous cellular processes, including transcription and translation, and cell cycle control (reviewed in [159, 163–170]).

Second, biomedical aspects of intrinsic disorder are of great importance too. It has been established that the deposition of some natively unfolded proteins is related to the development of several neurodegenerative disorders [89, 90, 171]. Some examples include AD (deposition of amyloid-β, tau-protein, α-synuclein fragment NAC[1] [172–175]); Niemann-Pick disease type C, subacute sclerosing panencephalitis, argyrophilic grain disease, myotonic dystrophy and motor neuron disease with neurofibrillary tangles (accumulation of tau protein in form of neurofibrillary tangles [174]); Down's syndrome (nonfilamentous amyloid-β deposits [176]); PD, dementia with Lewy body, diffuse Lewy body disease, Lewy body

variant of AD, multiple system atrophy and Hallervorden-Spatz disease (deposition of α-synuclein in a form of Lewy bodies, or Lewy neurites [177–181]).

Finally, the intrinsically disordered proteins represent an attractive subject for the biophysical characterization of the unfolded polypeptide chain under physiological conditions. A large variety of biophysical and biochemical methods have been applied for the structural description of these proteins. This includes ^{1}H-NMR[1], heteronuclear NMR, CD, optical rotatory dispersion (ORD), Fourier transform infrared spectroscopy (FTIR), intrinsic and extrinsic fluorescence, small-angle X-ray scattering (SAXS), small-angle neutron scattering (SANS), dynamic and static light scattering, gel electrophoresis, gel filtration, sedimentation, viscometry, scanning calorimetry, proteolytic mapping, epitope mapping and electron microscopy (summarized in [160, 165, 167, 182]).

How does an amino acid sequence encode unfoldedness?

As it has been already pointed out, the correct folding of a protein into its rigid biologically active conformation is determined by its amino acid sequence [83]. This gave rise to the reasonable assumption that the absence of rigid structure in natively unfolded proteins may also be encoded somehow in the specific features of their amino acid sequences. In an attempt to understand the relationship between sequence and disorder, Dunker et al. have developed several neuronal network predictors [159, 183–191]. The results of their analysis were more than impressive, as it was established that the disordered regions shared at least some common sequence features between many proteins, and that more than 15,000 proteins in the Swiss Protein database were identified as having long regions of sequence that shared these features [185]. Interestingly, the disordered proteins with highest scores were shown to have low sequence complexity, suggesting that the sequences of natively unfolded proteins may be essentially degenerated. However, it was later established that distribution of the complexity values for ordered and disordered sequences overlapped, suggesting that low sequence complexity did not represent the only characteristic feature of intrinsically disordered proteins [186]. Overall, the sequence of the intrinsically disordered proteins is characterized by amino acid compositional bias and the existence of highly predictable flexibility [188, 192]. Further, the majority of the intrinsically disordered proteins, being substantially depleted in I, L, V, W, F, Y, C and N, are enriched in E, K, R, G, Q, S, P and A [159]. Note that these features may account for the low hydrophobicity and high net charge of natively unfolded proteins, which have been pointed out for a few individual proteins [152, 157, 193].

Recently, it was established that comparison of overall hydrophobicity and net charge of native and natively unfolded protein sequences might be used to predict whether a given amino acid sequence encodes a native (folded) or an intrinsically unstructured protein [160]. In fact, this survey shows that the natively unfolded proteins are specifically localized within a unique region of the charge-hydrophobicity phase space, being completely isolated from the native globular proteins. Obviously, this allows estimation of the 'boundary' mean hydrophobicity value, $\langle H \rangle_b$, below which a polypeptide chain with a given mean net charge $\langle R \rangle$ will be most probably unfolded [160]:

$$\langle H \rangle = \frac{\langle R \rangle + 1.151}{2.785} \qquad (7)$$

The validity of these predictions has been successfully proven for several proteins, including peptides of the α domain of α-lactalbumin, lysozyme and α-lactalbumin/lysozyme chimeras [194], phosphodiesterase γ subunit [195], nucleoporin Nup2p [196] and other FG nucleoporins of *Saccaromyces cerevisiae* [197], the cell cycle inhibitor p57^{Kip2} [198] and several other proteins.

Two flavors of intrinsic unfoldedness

One of the major statements of the first part of this review was a conclusion that globular proteins are never random coils. Accumulated data nonambiguously show that they most likely represent the polymer coils below a critical point, even under harsh denaturing conditions. This means that high concentrations of urea and GdmHCl constitute relatively bad solvents for the polypeptide chains. On the other hand, a considerable number of sequences encode for the intrinsically unstructured proteins without specific structure. These findings raise several compelling biophysical questions related to the structural characteristics of natively unfolded proteins. How unfolded are these proteins? Are they random coils, or do they possess residual structure? If they have residual structure, how should they be classified? Fortunately, the accumulated information on natively unfolded proteins allowed making an initial structural classification of these intriguing members of the polypeptide kingdom [11, 182].

Based on the analysis of the available literature, it has been concluded that intrinsically unstructured proteins do not possess uniform structural properties, as expected for members of a single thermodynamic entity. In fact, they may be divided into two structurally different groups, intrinsic coils and premolten globules. Proteins from the first group have hydrodynamic dimensions typical of considerably unfolded polypeptide chain in poor solvent (see below and fig. 1), and do not possess any (or almost

any) ordered secondary structure. Proteins from the second group are more compact (see below and fig. 1), and exhibit some amount of residual secondary structure. However, they are still less dense than native globular or molten globule proteins [11, 182].

Importantly, several natively unfolded proteins were shown to be characterized by the absence of globular structure, or, in other words, they did not have a tightly packed core under the physiological conditions in vitro [11, 97, 153, 182, 199, 200]. This directly confirms an idea that intrinsically unstructured proteins belong to the class of polymer coils. They may be relatively extended (intrinsic coils), or partially collapsed (intrinsic premolten globules). This point will be considered in more detail below.

Polymer aspects of nonfolding

It is interesting to apply the same formalism of polymer physics represented in the first part of the review to the two classes of intrinsically unstructured proteins mentioned above. This type of analysis is available because considerable data have been accumulated on the hydrodynamic properties of these proteins. Obviously, the behavior of any polypeptide chain in a solution is determined by the interaction of the polymer with the solvent. The fact that natively unfolded proteins, with their depleted hydrophobicity, are noncompact under physiological conditions indicates that 'salted water' (typical 'physiological' buffer contains 100–150 mM NaCl) does not represent for them a poor solvent. In other words, these conditions do not force polymer segments to interact specifically with each other and, thus, do not force them to be effectively excluded from the solvent. On the other hand, it has already been noted that even high concentrations of strong denaturants do not represent a good solvent for a polypeptide chain encoding for a typical globular protein, and a globular protein was assumed to never be a random coil.

In this view, the data on the hydrodynamic analysis of intrinsically unstructured proteins are extremely important. Such an analysis reveals that the molecular dimensions of natively unfolded proteins follow the chain length as $R_S = 0.28\ M^{0.49}$ or $R_S = 0.6\ M^{0.40}$ for the native coils and premolten globules, respectively (see part 1 and fig. 1). Fig. 1 clearly reflects that native coils belong to the class of relatively extended unfolded conformations. Importantly, these coils show the largest K_h and the smallest ε values between different unfolded conformations of a polypeptide chain (cf. eqs 5, 6 and 7). This means that native coils under physiological conditions are in considerably worsened solvent conditions compared with the globular proteins in the urea or the GdmCl solutions (lowest ε value), which gives rise to the increased probability of multiple body interactions (highest K_h value). Finally, fig. 1 shows that native premolten globules follow exactly the same dependence as premolten globules of 'normal' globular proteins. Thus, these proteins may exhibit structural features of a squeezed polymer coil.

Function-induced folding of intrinsically unstructured proteins

An important feature of the intrinsically unstructured proteins is that they are able to undergo a disorder-to-order transition (i.e. partial or complete folding) during or prior to their biological function [12, 158–160, 164–167, 169, 201–207]. In other words, intrinsically unfolded proteins in vivo are likely to be stabilized by functional binding to specific targets and ligands (such as a variety of small molecules, substrates, cofactors, other proteins, nucleic acids, membranes and so on). The functional importance of being disordered has been intensively analyzed, and it has been established that increased intrinsic plasticity represents an important prerequisite for effective molecular recognition. The diapason of biological functions of the natively unfolded proteins is extremely wide, including cell cycle control, transcriptional and translational regulation, modulation of activity and/or assembly of other proteins, and even regulation of nerve cell function (reviewed in [159, 163, 166–168, 201]). It has been suggested that the persistence of natively unfolded proteins throughout evolution may be due to the definite advantages of disorder-to-order transitions accompanying functional performance of flexible structures in comparison with action of rigid proteins [159, 163, 166, 185, 187, 188]. These potential advantages of intrinsic lack of structure and function-related folding are (i) the ability of binding to several different targets, known as one-to-many signaling; (ii) the possibility of high specificity coupled with low affinity; (iii) the precise control and simple regulation of the binding thermodynamic; (iv) the capability to overcome steric restrictions, enabling essentially larger interaction surfaces in the complex than could be obtained for the rigid partners; (v) the increased rates of specific macromolecular association; (vi) the reduced lifetime of intrinsically disordered proteins in the cell, possibly representing a mechanism of rapid turnover of important regulatory molecules.

Finally, a few words should be added about the extent of function-related folding of intrinsically unstructured proteins. It has been shown that the range of conformational changes induced in natively unfolded proteins as a result of their interaction with natural partners is very wide [165, 208]. In fact, examples of all possible conformational transitions have been described, including function-induced transitions of coil to premolten globule, coil to molten globule, coil to ordered conformation, premolten globule to molten globule, premolten globule to

rigid structure and molten globule to ordered, and rigid form [11, 182]. This means that the structure-function paradigm, which emphasizes that ordered 3D structures represent an indispensable prerequisite to effective protein functioning, should be redefined to include intrinsically unstructured proteins [159]. According to this redefined paradigm, native proteins (or their functional regions) can exist in any of the known conformational states, ordered, molten globule, premolten globule and coil. Function can arise from any of these conformations and transitions between them. Thus, not just the ordered state but any of the known polypeptide conformations can be the native state of a protein.

It was suggested that the structure-forming effect of natural partners might be explained by their influence on the hydrophobicity and/or net charge of the natively unfolded polypeptide. Obviously, these parameters could be changed in such a way that they will approach values typical of ordered native proteins forcing a polypetide to fold [160]; i.e. formation of protein-ligand complexes may increase a 'critical mass' of the folding unit, thus initiating a chain of folding reactions.

There is, however, an alternative view to the problem of existence and functionality of intrinsically unstructured proteins. It is possible that these proteins are as folded in the cell as 'normal' proteins, due to the fact that concentration of their ligands is always high enough in vivo to fold them. Thus, evolution never sees them as any different from normal proteins. In this view, the property of being intrinsically unstructured is merely an artifact resulting from the proteins being studied after being purified to homogeneity. Evolutionarily, intrinsically unstructured proteins persist not because there is an advantage to being intrinsically unstructured, but because there is no disadvantage as long as there are enough ligands around in vivo to fold the proteins. The stability of the folded, ligand bound form may only be a few kcal/mol, but this would be enough to keep the vast majority of molecules folded and functional (even most normal, folded proteins are only marginally stable).

Polypeptide at the crossroads: concluding remarks

The data considered in this review unambiguously show that in in vitro experiments and in a living cell, a polypeptide chain chooses between three potential routes – nonfolding, folding and misfolding, with the last two representing competitive routes to higher structural order. For a single-chain protein, nonfolding and folding represent a choice of each individual molecule, whereas misfolding is a fate of the ensemble of molecules.

Multiple factors, stemming from the peculiarities of protein amino acid sequence and/or the features of protein environment, might affect the choice between nonfolding, folding and misfolding. The primary selection between low structural complexity (nonfolding) and increased complexity (folding) is determined only by amino acid composition. An abnormally highly charged polypeptide with low overall hydrophobicity will not fold, giving rise to a natively unfolded protein, whereas a usual polypeptide chain will choose the folding path at identical conditions. On the other hand, some changes in amino acid sequence (point mutations) may favor the misfolding pathway for both the natively unfolded and natively folded proteins.

Importantly, for a given polypeptide chain a chosen fate is not a final one, and a choice may be further modulated by environmental pressure. Thus, intrinsically unstructured proteins may be forced to fold or misfold via modification of their environment (addition of natural binding partners, changes in properties of solvent and so on), whereas a destabilizing environment may push a natively folded protein to the misfolding route. Alternatively, the presence of chaperones may reverse the misfolding route and effectively dissolve small aggregates [209]. Another important point is that misfolding of natively unfolded proteins to some extent resembles the process of protein folding; i.e. it represents a way from a simple, flexible and disordered conformation (unfolded polypeptide chain) somehow via a more ordered partially folded intermediate, to a complex and rigid structure, e.g. amyloid fibril. On the other hand, misfolding of a rigid globular protein involves a step of transient disordering and formation of a partially unfolded intermediate, which is followed by a subsequent increase in order originating from formation of specific protein aggregates.

Acknowledgements. The author expresses his deepest gratitude to Alexey Uversky and Larissa Munishkina for careful reading and editing the manuscript. This research was supported in part by a grant from INTAS (01-2347 CA).

1 Bryngelson J. D., Onuchic J. N., Socci N. D. and Wolynes P. G. (1995) Funnels, pathways and the energy landscape of protein folding – a synthesis. Proteins Structure Function Genetics **21:** 167–195
2 Dill K. A. and Chan H. S. (1997) From Levinthal to pathways to funnels. Nat. Struct. Biol. **4:** 10–19
3 Dobson C. M., Sali A. and Karplus M. (1998) Protein folding: a perspective from theory and experiment. Angew. Chem. Int. Ed. **37:** 868–893
4 Dobson C. M. and Karplus M. (1999) The fundamentals of protein folding: bringing together theory and experiment. Curr. Opin. Struct. Biol. **9:** 92–101
5 Dinner A. R., Sali A., Smith L. J., Dobson C. M. and Karplus M. (2000) Understanding protein folding via free-energy surfaces from theory and experiment. Trends Biochem. Sci. **25:** 331–339
6 Fersht A. R. and Daggett V. (2002) Protein folding and unfolding at atomic resolution. Cell **108:** 573–582
7 Ptitsyn O. B. and Volkenstein M. V. (1986) Protein structure and neutral theory of evolution. J. Biomol. Struct. Dynamics **4:** 137–156
8 Lau K. F. and Dill K. A. (1990) Theory for protein mutability and biogenesis. Proc. Natl. Acad. Sci. USA **87:** 638–642

9 Ptitsyn O. B. (1995) Molten globule and protein folding. Adv. Protein Chem. **47:** 83–229
10 Tcherkasskaya O. and Uversky V. N. (2001) Denatured collapsed states in protein folding: Example of apomyoglobin. Proteins Structure Function Genetics **44:** 244–254
11 Uversky V. N. (2002) Natively unfolded proteins: a point where biology waits for physics. Protein Sci. **11:** 739–756
12 Uversky V. N. (2002) What does it mean to be natively unfolded? Eur. J. Biochem. **269:** 2–12
13 Tcherkasskaya O., Davidson E. A. and Uversky V. N. (2003) Biophysical constraints for protein structure prediction. J. Proteome Res. **2:** 37–42
14 Grossberg A.Yu. and Khohlov A. R. (1989) Statistical Physics of Macromolecules, Nauka, Moscow
15 Baldwin R. L. (1986) Seeding protein folding. Trends Biochem. Sci. **11:** 6–9
16 Finkelstein A. V., Galzitskaya O. V. and Badretdinov A. Y. (1996) A folding pathway solving Levinthal's paradox. Progr. Biophys. Mol. Biol. **65:** A412
17 Tanford C. (1968) Protein denaturation. Adv. Protein Chem. **23:** 121–282
18 Tanford C. (1961) Physical Chemistry of Macromolecules, Wiley, New York
19 Smith L. J., Fiebig K. M., Schwalbe H. and Dobson C. M. (1996) The concept of a random coil – residual structure in peptides and denatured proteins. Folding Design **1:** R95– R106
20 Miller W. G. and Goebel C. V. (1968) Dimensions of protein random coils. Biochemistry **7:** 3925–3935
21 Alexandrescu A. T., Abeygunawardana C. and Shortle D. (1994) Structure and dynamics of a denatured 131-residue fragment of staphylococcal nuclease: a heteronuclear NMR study. Biochemistry **33:** 1063–1072
22 Wang Y. and Shortle D. (1995) The equilibrium folding pathway of staphylococcal nuclease: identification of the most stable chain-chain interactions by NMR and CD spectroscopy. Biochemistry **34:** 15895–15905
23 Wang Y. and Shortle D. (1997) Residual helical and turn structure in the denatured state of staphylococcal nuclease: analysis of peptide fragments. Folding Design **2:** 93–100
24 Wang Y. and Shortle D. (1996) A dynamic bundle of four adjacent hydrophobic segments in the denatured state of staphylococcal nuclease. Protein Sci. **5:** 1898–1906
25 Gillespie J. R. and Shortle D. (1997) 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. **268:** 170–184
26 Shortle D. and Ackerman M. S. (2001) Persistence of native-like topology in a denatured protein in 8 M urea. Science **293:** 487–489
27 James E., Wu P. G., Stites W. and Brand L. (1992) Compact denatured state of a staphylococcal nuclease mutant by guanidinium as determined by resonance energy-transfer. Biochemistry **31:** 10217–10225
28 Saab-Rincon G., Froebe C. L. and Matthews C. R. (1993) Urea-induced unfolding of the alpha subunit of tryptophan synthase: one-dimensional proton NMR evidence for residual structure near histidine-92 at high denaturant concentration. Biochemistry **32:** 13981–13990
29 Saab-Rincon G., Gualfetti P. J. and Matthews C. R. (1996) Mutagenic and thermodynamic analyses of residual structure in the alpha subunit of tryptophan synthase. Biochemistry **35:** 1988–1994
30 Neri D., Billeter M., Wider G. and Wuthrich K. (1992) NMR determination of residual structure in a urea-denatured protein, the 434-repressor. Science **257:** 1559–1563
31 Zhang O. and Forman-Kay J. D. (1995) Structural characterization of folded and unfolded states of an SH3 domain in equilibrium in aqueous buffer. Biochemistry **34:** 6784–6794
32 Zhang O. and Forman-Kay J. D. (1997) NMR studies of unfolded states of an SH3 domain in aqueous solution and denaturing conditions. Biochemistry **36:** 3959–3970
33 Schulman B. A., Kim P. S., Dobson C. M. and Redfield C. (1997) A residue-specific NMR view of the non-cooperative unfolding of a molten globule. Nat. Struct. Biol. **4:** 630–634
34 Nolting B., Golbik R., Solergonzalez A. S. and Fersht A. R. (1997) Circular dichroism of denatured barstar suggests residual structure. Biochemistry **36:** 9899–9905
35 Koepf E. K., Petrassi H. M., Sudol M. and Kelly J. W. (1999) WW: an isolated three-stranded antiparallel beta-sheet domain that unfolds and refolds reversibly; evidence for a structured hydrophobic cluster in urea and GdnHCl and a disordered thermal unfolded state. Protein Sci. **8:** 841–853
36 Gottfried D. S. and Haas E. (1992) Nonlocal interactions stabilize compact folding intermediates in reduced unfolded bovine pancreatic trypsin inhibitor. Biochemistry **31:** 12353–12362
37 Amir D., Krausz S. and Haas E. (1992) Detection of local structures in reduced unfolded bovine pancreatic trypsin inhibitor. Proteins **13:** 162–173
38 Svensson M., Jonasson P., Freskgard P. O., Jonsson B. H., Lindgren M., Martensson L. G. et al. (1995) Mapping the folding intermediate of human carbonic anhydrase II. Probing substructure by chemical reactivity and spin and fluorescence labeling of engineered cysteine residues. Biochemistry **34:** 8606–8620
39 Martensson L. G., Jonsson B. H., Freskgard P. O., Kihlgren A., Svensson M. and Carlsson U. (1993) Characterization of folding intermediates of human carbonic anhydrase II: probing substructure by chemical labeling of SH groups introduced by site-directed mutagenesis. Biochemistry **32:** 224–231
40 Hammarstrom P. and Carlsson U. (2000) Is the unfolded state the Rosetta Stone of the protein folding problem? Biochem. Biophys. Res Commun. **276:** 393–398
41 Schwarzinger S., Wright P. E. and Dyson H. J. (2002) Molecular hinges in protein folding: the urea-denatured state of apomyoglobin. Biochemistry **41:** 12681–12686
42 Klein-Seetharaman J., Oikawa M., Grimshaw S. B., Wirmer J., Duchardt E., Ueda T. et al. (2002) Long-range interactions within a nonnative protein. Science **295:** 1719–1722
43 Shortle (1996) Structural analysis of non-native states of proteins by nmr methods. Curr. Opin. Struct. Biol. **6:** 24–30
44 Shortle D. (1996) The denatured state (the other half of the folding equation) and its role in protein stability. FASEB J. **10:** 27–34
45 Dill K. A. and Shortle D. (1991) Denatured states of proteins. Annu. Rev. Biochem. **60:** 795–825
46 Baldwin R. L. and Zimm B. H. (2000) Are denatured proteins ever random coils? Proc. Natl. Acad. Sci. USA **97:** 12391–12392
47 Plaxco K. W. and Gross M. (2001) Unfolded, yes, but random? Never! Nat. Struct. Biol. **8:** 659–660
48 Ramachandran G. N. and Sasisekharan V. (1968) Conformation of polypeptides and proteins. Adv. Protein Chem. **23:** 283–438
49 Shortle D. (2002) The expanded denatured state: an ensemble of conformations trapped in a locally encoded topological space. Adv. Protein Chem. **62:** 1–23
50 Shimizu S. and Chan H. S. (2002) Origins of protein denatured state compactness and hydrophobic clustering in aqueous urea: inferences from nonpolar potentials of mean force. Proteins **49:** 560–566
51 Palleros D. R., Shi L., Reid K. L. and Fink A. L. (1993) Three-state denaturation of DnaK induced by guanidine hydrochloride. Evidence for an expandable intermediate. Biochemistry **32:** 4314–4321
52 Uversky V. N. and Ptitsyn O. B. (1994) 'Partly folded' state, a new equilibrium state of protein molecules: four-state guani-

dinium chloride-induced unfolding of beta-lactamase at low temperature. Biochemistry **33:** 2782–2791
53 Ptitsyn O. B., Bychkova V. E. and Uversky V. N. (1995) Kinetic and equilibrium folding intermediates. Phil. Trans. Roy. Soc. London B **348:** 35–41
54 Uversky V. N. and Ptitsyn O. B. (1996) 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. **255:** 215–228
55 Zhang Y. L., Zhou J. M. and Tsou C. L. (1996) Sequential unfolding of adenylate kinase during denaturation by guanidine hydrochloride. Biochim. Biophys. Acta **1295:** 239–244
56 Uversky V. N. (1997) Diversity of equilibrium compact forms of denatured globular proteins. Protein Peptide Lett. **4:** 355–367
57 Karnoup A. S. and Uversky V. N. (1997) Sequential compaction of a random copolymer of hydrophilic and hydrophobic amino acid residues. Macromolecules **30:** 7427–7434
58 Uversky V. N., Karnoup A. S., Segel D. J., Seshadri S., Doniach S. and Fink A. L. (1998) Anion-induced folding of staphylococcal nuclease: characterization of multiple equilibrium partially folded intermediates. J. Mol. Biol. **278:** 879–894
59 Bushmarina N. A., Kuznetsova I. M., Biktashev A. G., Turoverov K. K. and Uversky V. N. (2001) Partially folded conformations in the folding pathway of bovine carbonic anhydrase II: A fluorescence spectroscopic analysis. Chembiochem **2:** 813–821
60 Kuznetsova I. M., Stepanenko O. V., Turoverov K. K., Zhu L., Zhou J. M., Fink A. L. et al. (2002) Unraveling multistate unfolding of rabbit muscle creatine kinase. Biochim. Biophys. Acta protein Struct. Mol. Enzymol. **1596:** 138–155
61 Mirsky A. E. and Pauling L. (1937) On the structure of native, denatured and coagulated proteins. Proc. Natl. Acad. Sci. USA **33:** 439–441
62 Karplus M. and Weaver D. L. (1976) Protein-folding dynamics. Nature **260:** 404–406
63 Kim P. S. and Baldwin R. L. (1982) Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. Annu. Rev. Biochem. **51:** 459–489
64 Kauzmann W. (1959) Some factors in the interpretation of protein denaturation. Adv. Protein Chem. **14:** 1–63
65 Dill K. A. (1985) Theory for the folding and stability of globular proteins. Biochemistry **24:** 1501–1509
66 Covell D. G. and Jernigan R. L. (1990) Conformations of folded proteins in restricted spaces. Biochemistry **29:** 3287–3294
67 Chan H. S. and Dill K. A. (1990) Origins of structure in globular proteins. Proc. Natl. Acad. Sci. USA **87:** 6388–6392
68 Gregoret L. M. and Cohen F. E. (1991) Protein folding. Effect of packing density on chain conformation. J. Mol. Biol. **219:** 109–122
69 Fink A. L. (1995) Compact intermediate states in protein folding. Annu. Rev. Biophys. Biomol. Struct. **24:** 495–522
70 Uversky V. N. and Fink A. L. (2002) The chicken-egg scenario of protein folding revisited. FEBS Lett. **515:** 79–83
71 Evans P. A., Topping K. D., Woolfson D. N. and Dobson C. M. (1991) Hydrophobic clustering in nonnative states of a protein: interpretation of chemical shifts in NMR spectra of denatured states of lysozyme. Proteins **9:** 248–266
72 Logan T. M., Theriault Y. and Fesik S. W. (1994) Structural characterization of the FK506 binding protein unfolded in urea and guanidine hydrochloride. J. Mol. Biol. **236:** 637–648
73 Lumb K. J. and Kim P. S. (1994) Formation of a hydrophobic cluster in denatured bovine pancreatic trypsin inhibitor. J. Mol. Biol. **236:** 412–420
74 Arcus V. L., Vuilleumier S., Freund S. M., Bycroft M. and Fersht A. R. (1995) A comparison of the Ph, urea and temperature-denatured states of barnase by heteronuclear NMR: implications for the initiation of protein folding. J. Mol. Biol. **254:** 305–321
75 Miranker A. D. and Dobson C. M. (1996) Collapse and cooperativity in protein folding. Curr. Opin. Struct. Biol. **6:** 31–42
76 Wong K. B., Freund S. M. and Fersht A. R. (1996) Cold denaturation of barstar: 1h, 15n and 13c nmr assignment and characterisation of residual structure. J. Mol. Biol. **259:** 805–818
77 Schwalbe H., Fiebig K. M., Buck M., Jones J. A., Grimshaw S. B., Spencer A. et al. (1997) Structural and dynamical properties of a denatured protein. Heteronuclear 3D NMR experiments and theoretical simulations of lysozyme in 8 M urea. Biochemistry **36:** 8977–8991
78 Blanco F. J., Serrano L. and Forman-Kay J. D. (1998) High populations of non-native structures in the denatured state are compatible with the formation of the native folded state. J. Mol. Biol. **284:** 1153–1164
79 Kortemme T., Kelly M. J., Kay L. E., Forman-Kay J. and Serrano L. (2000) Similarities between the spectrin SH3 domain denatured state and its folding transition state. J. Mol. Biol. **297:** 1217–1229
80 Garcia P., Serrano L., Durand D., Rico M. and Bruix M. (2001) NMR and SAXS characterization of the denatured state of the chemotactic protein CheY: implications for protein folding initiation. Protein Sci. **10:** 1100–1112
81 Flory P. J. (1953) Principles of Polymer Chemistry, Cornell University Press, Ithaca, NY
82 Fernandez A., Kardos J. and Goto Y. (2003) Protein folding: could hydrophobic collapse be coupled with hydrogen-bond formation? FEBS Lett. **536:** 187–192
83 Anfinsen C. B., Haber E., Sela M. and White F. N. (1961) The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. Proc. Natl. Acad. Sci. USA **47:** 1309–1314
84 Kuwajima K. (1996) The molten globule state of alpha-lactalbumin. FASEB J. **10:** 102–109
85 Uversky V. N. (2002) Cracking the folding code. Why do some proteins adopt partially folded conformations, whereas other don't? FEBS Lett. **514:** 181–183
86 Kelly J. W. (1998) The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. Curr. Opin. Struct. Biol. **8:** 101–106
87 Dobson C. M. (1999) Protein misfolding, evolution and disease. Trends Biochem. Sci. **24:** 329–332
88 Bellotti V., Mangione P. and Stoppini M. (1999) Biological activity and pathological implications of misfolded proteins. Cell. Mol. Life Sci. **55:** 977–991
89 Uversky V. N., Talapatra A., Gillespie J. R. and Fink A. L. (1999) Protein deposits as the molecular basis of amyloidosis. Part I. Systemic amyloidoses. Med. Sci. Monitor **5:** 1001–1012
90 Uversky V. N., Talapatra A., Gillespie J. R. and Fink A. L. (1999) Protein deposits as the molecular basis of amyloidosis. II. Localized amyloidosis and neurodegenerative disorders. Med. Sci. Monitor **5:** 1238–1254
91 Rochet J. C. and Lansbury P. T. Jr (2000) Amyloid fibrillogenesis: themes and variations. Curr. Opin. Struct. Biol. **10:** 60– 68
92 Sunde M., Serpell L. C., Bartlam M., Fraser P. E., Pepys M. B. and Blake C. C. (1997) Common core structure of amyloid fibrils by synchrotron X-ray diffraction. J. Mol. Biol. **273:** 729–739
93 Goldsbury C. S., Cooper G. J. S., Goldie K. N., Muller S. A., Saafi E. L., Gruijters W. T. M. et al. (1997) Polymorphic fibrillar assembly of human amylin. J. Struct. Biol. **119:** 17–27
94 Harper J. D., Lieber C. M. and Lansbury P. T. Jr (1997) Atomic force microscopic imaging of seeded fibril formation and fibril branching by the Alzheimer's disease amyloid-beta protein. Chem. Biol. **4:** 951–959

95 Conway K. A., Harper J. D. and Lansbury P. T. Jr (1998) Accelerated in vitro fibril formation by a mutant alpha-synuclein linked to early-onset Parkinson's disease. Nat. Med. **4:** 1318–1320

96 Serpell L. C., Sunde M., Benson M. D., Tennent G. A., Pepys M. B. and Fraser P. E. (2000) The protofilament substructure of amyloid fibrils. J. Mol. Biol. **300:** 1033–1039

97 Uversky V. N., Li J. and Fink A. L. (2001) Evidence for a partially folded intermediate in alpha-synuclein fibril formation. J. Biol. Chem. **276:** 10737–10744

98 Guijarro J. I., Sunde M., Jones J. A., Campbell I. D. and Dobson C. M. (1998) Amyloid fibril formation by an SH3 domain. Proc. Natl. Acad. Sci. USA **95:** 4224–4228

99 Litvinovich S. V., Brew S. A., Aota S., Akiyama S. K., Haudenschild C. and Ingham K. C. (1998) Formation of amyloid-like fibrils by self-association of a partially unfolded fibronectin type III module. J. Mol. Biol. **280:** 245–258

100 Konno T., Murata K. and Nagayama K. (1999) Amyloid-like aggregates of a plant protein: a case of a sweet-tasting protein, monellin. FEBS Lett. **454:** 122–126

101 Schuler B., Rachel R. and Seckler R. (1999) Formation of fibrous aggregates from a non-native intermediate: the isolated P22 tailspike beta-helix domain. J. Biol. Chem. **274:** 18589–18596

102 Gross M., Wilkins D. K., Pitkeathly M. C., Chung E. W., Higham C., Clark A. et al. (1999) Formation of amyloid fibrils by peptides derived from the bacterial cold shock protein CspB. Protein Sci. **8:** 1350–1357

103 Chiti F., Webster P., Taddei N., Clark A., Stefani M., Pamponi G. et al. (1999) Designing conditions for in vitro formation of amyloid protofilaments and fibrils. Proc. Natl. Acad. Sci. USA **96:** 3590–3594

104 Krebs M. R., Wilkins D. K., Chung E. W., Pitkeathly M. C., Chamberlain A. K., Zurdo J. et al. (2000) Formation and seeding of amyloid fibrils from wild-type hen lysozyme and a peptide fragment from the beta-domain. J. Mol. Biol. **300:** 541–549

105 Damaschun G., Damaschun H., Fabian H., Gast K., Krober R., Wieske M. et al. (2000) Conversion of yeast phosphoglycerate kinase into amyloid-like structure. Proteins **39:** 204–211

106 Yutani K., Takayama G., Goda S., Yamagata Y., Maki S., Namba K. et al. (2000) The process of amyloid-like fibril formation by methionine aminopeptidase from a hyperthermophile, *Pyrococcus furiosus*. Biochemistry **39:** 2769–2777

107 Ohnishi S., Koide A. and Koide S. (2000) Solution conformation and amyloid-like fibril formation of a polar peptide derived from a beta-hairpin in the OspA single-layer beta-sheet. J. Mol. Biol. **301:** 477–489

108 Fezoui Y., Hartley D. M., Walsh D. M., Selkoe D. J., Osterhout J. J. and Teplow D. B. (2000) A de novo designed helix-turn-helix peptide forms nontoxic amyloid fibrils. Nat. Struct. Biol. **7:** 1095–1099

109 Fandrich M., Fletcher M. A. and Dobson C. M. (2001) Amyloid fibrils from muscle myoglobin. Nature **410:** 165–166

110 Pertinhez T. A., Bouchard M., Tomlinson E. J., Wain R., Ferguson S. J., Dobson C. M. et al. (2001) Amyloid fibril formation by a helical cytochrome. FEBS Lett. **495:** 184–186

111 Nielsen L., Frokjaer S., Brange J., Uversky V. N. and Fink A. L. (2001) Probing the mechanism of insulin fibril formation with insulin mutants. Biochemistry **40:** 8397–8409

112 Johansson J. (2001) Membrane properties and amyloid fibril formation of lung surfactant protein C. Biochem. Soc. Trans. **29:** 601–606

113 Souillac P. O., Uversky V. N., Millett I. S., Khurana R., Doniach S. and Fink A. L. (2002) Effect of association state and conformational stability on the kinetics of immunoglobulin light chain amyloid fibril formation at physiological pH. J. Biol. Chem. **277:** 12657–12665

114 Pavlov N. A., Cherny D. I., Heim G., Jovin T. M. and Subramaniam V. (2002) Amyloid fibrils from the mammalian protein prothymosin alpha. FEBS Lett. **517:** 37–40

115 Goers J., Permyakov S. E., Permyakov E. A., Uversky V. N. and Fink A. L. (2002) Conformational prerequisites for alpha-lactalbumin fibrillation. Biochemistry **41:** 12546–12551

116 Lansbury P. T. Jr (1999) Evolution of amyloid: what normal protein folding may tell us about fibrillogenesis and disease. Proc. Natl. Acad. Sci. USA **96:** 3342–3344

117 Fink A. L. (1998) Protein aggregation: folding aggregates, inclusion bodies and amyloid. Folding Design **3:** 9–15

118 Dobson C. M. (2001) Protein folding and its links with human disease. Biochem. Soc. Symp. **68:** 1–26

119 Zerovnik E. (2002) Amyloid-fibril formation. Proposed mechanisms and relevance to conformational disease. Eur. J. Biochem. **269:** 3362–3371

120 Teplow D. B. (1998) Structural and kinetic features of amyloid beta-protein fibrillogenesis. Amyloid Int. J. Exp. Clin. Invest. **5:** 121–142

121 Kayed R., Bernhagen J., Greenfield N., Sweimeh K., Brunner H., Voelter W. et al. (1999) Conformational transitions of islet amyloid polypeptide (IAPP) in amyloid formation in vitro. J. Mol. Biol. **287:** 781–796

122 Schweers O., Schonbrunn-Hanebeck E., Marx A. and Mandelkow E. (1994) Structural studies of tau protein and Alzheimer paired helical filaments show no evidence for beta-structure. J. Biol. Chem. **269:** 24290–24297

123 Kelly J. W. (1996) Alternative conformations of amyloidogenic proteins govern their behavior. Curr. Opin. Struct. Biol. **6:** 11–17

124 Wetzel R. (1997) Domain stability in immunoglobulin light chain deposition disorders. Adv. Protein Chem. **50:** 183–242

125 Wetzel R. (1999) Amyloid, Prions and Other Protein Aggregates, Academic Press, San Diego

126 Canet D., Sunde M., Last A. M., Miranker A., Spencer A., Robinson C. V. et al. (1999) Mechanistic studies of the folding of human lysozyme and the origin of amyloidogenic behavior in its disease-related variants. Biochemistry **38:** 6419–6427

127 Lashuel H. A., Wurth C., Woo L. and Kelly J. W. (1999) The most pathogenic transthyretin variant, L55P, forms amyloid fibrils under acidic conditions and protofilaments under physiological conditions. Biochemistry **38:** 13560–13573

128 Saraiva M. J. (2001) Transthyretin amyloidosis: a tale of weak interactions. FEBS Lett. **498:** 201–203

129 Heegaard N. H., Sen J. W., Kaarsholm N. C. and Nissen M. H. (2001) Conformational intermediate of the amyloidogenic protein beta 2-microglobulin at neutral pH. J. Biol. Chem. **276:** 32657–32662

130 Colon W., Lai Z., McCutchen S. L., Miroy G. J., Strang C. and Kelly J. W. (1996) FAP mutations destabilize transthyretin facilitating conformational changes required for amyloid formation. Ciba Found. Symp. **199:** 228–238

131 Chiti F., Taddei N., Stefani M., Dobson C. M. and Ramponi G. (2001) Reduction of the amyloidogenicity of a protein by specific binding of ligands to the native conformation. Protein Sci. **10:** 879–886

132 Miroy G. J., Lai Z., Lashuel H. A., Peterson S. A., Strang C. and Kelly J. W. (1996) Inhibiting transthyretin amyloid fibril formation via protein stabilization. Proc. Natl. Acad. Sci. USA **93:** 15051–15056

133 Nielsen L., Khurana R., Coats A., Frokjaer S., Brange J., Vyas S. et al. (2001) Effect of environmental factors on the kinetics of insulin fibril formation: Elucidation of the molecular mechanism. Biochemistry **40:** 6036–6046

134 Saraiva M. J., Birken S., Costa P. P. and Goodman D. S. (1984) Amyloid fibril protein in familial amyloidotic polyneuropathy, Portuguese type. Definition of molecular abnormality in transthyretin (prealbumin). J. Clin. Invest **74:** 104–119

135 Lai Z., Colón W. and Kelly J. W. (1996) The acid-mediated denaturation pathway of transthyretin yields a conformational intermediate that can self-assemble into amyloid. Biochemistry **35:** 6470–6482

136 Jiang X., Smith C. S., Petrassi H. M., Hammarstrom P., White J. T., Sacchettini J. C. et al. (2001) An engineered transthyretin monomer that is nonamyloidogenic, unless it is partially denatured. Biochemistry **40:** 11442–11452

137 Buxbaum J. (1992) Mechanisms of disease: monoclonal immunoglobulin deposition. Amyloidosis, light chain deposition disease, and light and heavy chain deposition disease. Hematol. Oncol. Clin. North Am. **6:** 323–346

138 Khurana R., Gillespie J. R., Talapatra A., Minert L. J., Ionescu-Zanetti C., Millett I. et al. (2001) Partially folded intermediates as critical precursors of light chain amyloid fibrils and amorphous aggregates. Biochemistry **40:** 3525–3535

139 Souillac P. O., Uversky V. N., Millett I. S., Khurana R., Doniach S. and Fink A. L. (2002) Elucidation of the molecular mechanism during the early events in immunoglobulin light chain amyloid fibrillation – evidence for an off-pathway oligomer at acidic pH. J. Biol. Chem. **277:** 12666–12679

140 Permyakov E. A. and Berliner L. J. (2000) alpha-Lactalbumin: structure and function. FEBS Lett. **473:** 269–274

141 Hamada D. and Dobson C. M. (2002) A kinetic study of beta-lactoglobulin amyloid fibril formation promoted by urea. Protein Sci **11:** 2417–2426

142 Ramirez-Alvarado M., Merkel J. S. and Regan L. (2000) A systematic exploration of the influence of the protein stability on amyloid fibril formation in vitro. Proc. Natl. Acad. Sci. USA **97:** 8979–8984

143 Uversky V. N., Li J. and Fink A. L. (2001) Pesticides directly accelerate the rate of alpha-synuclein fibril formation: a possible factor in Parkinson's disease. FEBS Lett. **500:** 105–108

144 Manning-Bog A. B., McCormack A. L., Li J., Uversky V. N., Fink A. L. and Di Monte D. A. (2002) The herbicide paraquat causes up-regulation and aggregation of alpha-synuclein in mice – paraquat and alpha-synuclein. J. Biol. Chem. **277:** 1641–1644

145 Uversky V. N., Li J. and Fink A. L. (2001) Metal-triggered structural transformations, aggregation and fibrillation of human alpha-synuclein – a possible molecular link between Parkinson's disease and heavy metal exposure. J. Biol. Chem. **276:** 44284–44296

146 Uversky V. N., Li J. and Fink A. L. (2001) Trimethylamine-N-oxide-induced folding of alpha-synuclein. FEBS Lett. **509:** 31–35

147 Munishkina L. A., Phelan C., Uversky V. N. and Fink A. L. (2003) Conformational behavior and aggregation of human α-synuclein in organic solvents. Modeling the effect of membranes. Biochemistry **42:** 2720–2730

148 Uversky V. N., Yamin G., Souillac P. O., Goers J., Glaser C. B. and Fink A. L. (2002) Methionine oxidation inhibits fibrillation of human alpha-synuclein in vitro. FEBS Lett. **517:** 239–244

149 Cooper G. J., Willis A. C., Clark A., Turner R. C., Sim R. B. and Reid K. B. (1987) Purification and characterization of a peptide from amyloid-rich pancreases of type 2 diabetic patients. Proc. Natl. Acad. Sci. USA **84:** 8628–8632

150 Zagorski M. G., Yang J., Shao H., Ma K., Zeng H. and Hong A. (1999) Methodological and chemical factors affecting amyloid beta peptide amyloidogenicity. Methods Enzymol. **309:** 189–204

151 Kirkitadze M. D., Condron M. M. and Teplow D. B. (2001) Identification and characterization of key kinetic intermediates in amyloid beta-protein fibrillogenesis. J. Mol. Biol. **312:** 1103–1119

152 Gast K., Damaschun H., Eckert K., Schulze-Forster K., Maurer H. R., Muller-Frohne M. et al. (1995) Prothymosin alpha: a biologically active protein with random coil conformation. Biochemistry **34:** 13211–13218

153 Uversky V. N., Gillespie J. R., Millett I. S., Khodyakova A. V., Vasiliev A. M., Chernovskaya T. V. et al. (1999) Natively unfolded human prothymosin alpha adopts partially folded collapsed conformation at acidic pH. Biochemistry **38:** 15009–15016

154 Vassilenko K. S. and Uversky V. N. (2002) Native-like secondary structure of molten globules. Biochim. Biophys. Acta Protein Struct. Mol. Enzymol. **1594:** 168–177

155 Morozova-Roche L. A., Zurdo J., Spencer A., Noppe W., Receveur V., Archer D. B. et al. (2000) Amyloid fibril formation and seeding by wild-type human lysozyme and its disease-related mutational variants. J. Struct. Biol. **130:** 339–351

156 Nandi P. K., Leclerc E., Nicole J. C. and Takahashi M. (2002) DNA-induced partial unfolding of prion protein leads to its polymerisation to amyloid. J. Mol. Biol. **322:** 153–161

157 Weinreb P. H., Zhen W., Poon A. W., Conway K. A. and Lansbury P. T. Jr (1996) NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded. Biochemistry **35:** 13709–13715

158 Wright P. E. and Dyson H. J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. J. Mol. Biol. **293:** 321–331

159 Dunker A. K., Lawson J. D., Brown C. J., Williams R. M., Romero P., Oh J. S. et al. (2001) Intrinsically disordered protein. J. Mol. Graphics Modelling **19:** 26–59

160 Uversky V. N., Gillespie J. R. and Fink A. L. (2000) Why are 'natively unfolded' proteins unstructured under physiologic conditions? Proteins Struct. Function Genetics **41:** 415–427

161 Fischer E. (1894) Einfluss der configuration auf die wirkung der enzyme. Ber. Dt. Chem. Ges. **27:** 2993

162 Clayton D. F. and George J. M. (1998) The synucleins: a family of proteins involved in synaptic function, plasticity, neurodegeneration and disease. Trends Neurosci. **21:** 249–254

163 Wright P. E. and Dyson H. J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. J. Mol. Biol. **293:** 321–331

164 Dunker A. K. and Obradovic Z. (2001) The protein trinity – linking function and disorder. Nat. Biotechnol. **19:** 805–806

165 Uversky V. N. (2002) Natively unfolded proteins: a point where biology waits for physics. Protein Sci. **11:** 739–756

166 Dyson H. J. and Wright P. E. (2002) Coupling of folding and binding for unstructured proteins. Curr. Opin. Struct. Biol. **12:** 54–60

167 Tompa P. (2002) Intrinsically unstructured proteins. Trends Biochem. Sci. **27:** 527–533

168 Dunker A. K., Brown C. J., Lawson J. D., Iakoucheva L. M. and Obradovic Z. (2002) Intrinsic disorder and protein function. Biochemistry **41:** 6573–6582

169 Dunker A. K., Brown C. J. and Obradovic Z. (2002) Identification and functions of usefully disordered proteins. Unfolded Proteins **62:** 25–49

170 Iakoucheva L. M., Brown C. J., Lawson J. D., Obradovic Z. and Dunker A. K. (2002) Intrinsic disorder in cell-signaling and cancer-associated proteins. J. Mol. Biol. **323:** 573–584

171 Uversky V. N. and Fink A. L. (2003) Conformational constraints for the amyloid fibrillation: the importance of being unfolded. Biochim. Biophys. Acta, in press

172 Glenner G. G. and Wong C. W. (1984) Alzheimers disease and Downs syndrome – sharing of a unique cerebrovascular amyloid fibril protein. Biochem. Biophys. Res. Commun. **122:** 1131–1135

173 Masters C. L., Multhaup G., Simms G., Pottgiesser J., Martins R. N. and Beyreuther K. (1985) Neuronal origin of a cerebral amyloid – neurofibrillary tangles of Alzheimers disease contain the same protein as the amyloid of plaque cores and blood vessels. EMBO J. **4:** 2757–2763

174 Lee V. M. Y., Balin B. J., Otvos L. and Trojanowski J. Q. (1991) A68 – a major subunit of pPaired helical filaments and derivatized forms of normal tau. Science **251:** 675–678

175 Ueda K., Fukushima H., Masliah E., Xia Y., Iwai A., Yoshimoto M. et al. (1993) Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease. Proc. Natl. Acad. Sci. USA **90:** 11282–11286

176 Wisniewski K. E., Dalton A. J., Mclachlan D. R. C., Wen G. Y. and Wisniewski H. M. (1985) Alzheimers disease in Down syndrome – clinicopathologic studies. Neurology **35:** 957–961

177 Arawaka S., Saito Y., Murayama S. and Mori H. (1998) Lewy body in neurodegeneration with brain iron accumulation type 1 is immunoreactive for *a*-synuclean. Neurology **51:** 887–889

178 Arima K., Uéda K., Sunohara N., Hirai S., Izumiyama Y., Tonozuka-Uehara H. et al. (1998) Immunoelectron-microscopic demonstration of NACP/*a*-synuclein-epitopes on the filamentous component of Lewy bodies in Parkinson's disease and in dementia with Lewy bodies. Brain Res. **808:** 93–100

179 Tu P. H., Galvin J. E., Baba M., Giasson B., Tomita T., Leight S. et al. (1998) Glial cytoplasmic inclusions in white matter oligodendrocytes of multiple system atrophy brains contain insoluble alpha-synuclein. Ann. Neurol. **44:** 415–422

180 Wakabayashi K., Yoshimoto M., Tsuji S. and Takahashi H. (1998) Alpha-synuclein immunoreactivity in glial cytoplasmic inclusions in multiple system atrophy. Neurosci. Lett. **249:** 180–182

181 Galvin J. E., Lee V. M., Schmidt M. L., Tu P. H., Iwatsubo T. and Trojanowski J. Q. (1999) Pathobiology of the Lewy body. Adv. Neurol. **80:** 313–324

182 Uversky V. N. (2002) What does it mean to be natively unfolded? Eur. J. Biochem. **269:** 2–12

183 Romero P., Obradovic Z., Kissinger C., Villafranca J. E. and Dunker A. K. (1997) Identifying disordered regions in proteins from amino acid sequence. Proc. IEEE Int. Conf. Neuronal Networks 1997 **1:** 90–95

184 Romero P., Obradovic Z. and Dunker A. K. (1998) Sequence data analysis for long distorted regions prediction in the calcineurin family. Genome Informatics **8:** 110–124

185 Romero P., Obradovic Z., Kissinger C., Villafranca J. E., Garner E., Guilliot S. et al. (1998) Thousands of proteins likely to have long disordered regions. Pac. Symp. Biocomput. **3:** 437–448

186 Romero P., Obradovic Z., Li X. H., Garner E. C., Brown C. J. and Dunker A. K. (2001) Sequence complexity of disordered protein. Proteins Structure Function and Genetics **42:** 38–48

187 Dunker A. K., Obradovic Z., Romero P., Kissinger C. and Villafranca J. E. (1997) On the importance of being disordered. PDB Newsletter **81:** 3–5

188 Dunker A. K., Garner E., Guilliot S., Romero P., Albercht K., Hart J. et al. (1998) Protein disorder and the evolution of molecular recognition: theory, predictions and observations. Pac. Symp. Biocomput. **3:** 473–484

189 Garner E., Guilliot S., Dunker A. K., Romero P. and Obradovic Z. (1998) Predicting long disordered regions in protein from amino acid sequence. Biophys. J. **74:** A281

190 Li X., Romero P., Rani M., Dunker A. K. and Obradovic Z. (1999) Predicting Protein Disorder for N-, C- and Internal Regions. Genome Inform. Ser. Workshop Genome Inform. **10:** 30–40

191 Li X., Obradovic Z., Brown C. J., Garner E. C. and Dunker A. K. (2000) Comparing predictors of disordered protein. Genome Inform. Ser. Workshop Genome Inform. **11:** 172– 184

192 Garner E., Cannon P., Romero P., Obradovic Z. and Dunker A. K. (1998) Predicting disordered regions from amino acid sequence: common themes despite differing structural characterization. Genome Inform. Ser. Workshop Genome Inform. **9:** 201–213

193 Hemmings H. C. Jr, Nairn A. C., Aswad D. W. and Greengard P. (1984) 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. **4:** 99–110

194 Demarest S. J., Zhou S. Q., Robblee J., Fairman R., Chu B. and Raleigh D. P. (2001) A comparative study of peptide models of the alpha-domain of alpha-lactalbumin, lysozyme and alpha-lactalbumin/lysozyme chimeras allows the elucidation of critical factors that contribute to the ability to form stable partially folded states. Biochemistry **40:** 2138–2147

195 Uversky V. N., Permyakov S. E., Zagranichny V. E., Rodionov I. L., Fink A. L., Cherskaya A. M. et al. (2002) Effect of zinc and temperature on the conformation of the gamma subunit of retinal phosphodiesterase: A natively unfolded protein. J. Proteome Res. **1:** 149–159

196 Denning D. P., Uversky V., Patel S. S., Fink A. L. and Rexach M. (2002) The *Saccharomyces cerevisiae* nucleoporin Nup2p is a natively unfolded protein. J. Biol. Chem. **277:** 33447–33455

197 Denning D. P., Patel S. S., Uversky V. N., Fink A.L. and Rexach M. (2003) Disorder in the pore. The FG nucleoporins of *S. cerevisiae* constitute a family of natively unfolded proteins. Proc. Natl. Acad. Sci. USA **100:** 2450–2455

198 Adkins J. N. and Lumb K. J. (2002) Intrinsic structural disorder and sequence features of the cell cycle inhibitor p57Kip2. Proteins **46:** 1–7

199 Uversky V. N., Li J., Souillac P., Millett I. S., Doniach S., Jakes R. et al. (2002) Biophysical properties of the synucleins and their propensities to fibrillate – inhibition of alpha-synuclein assembly by beta- and gamma-synucleins. J. Biol. Chem. **277:** 11970–11978

200 Konno T., Tanaka N., Kataoka M., Takano E. and Maki M. (1997) A circular dichroism study of preferential hydration and alcohol effects on a denatured protein, pig calpastatin domain I. Biochim. Biophys. Acta **1342:** 73–82

201 Dunker A. K., Brown C. J. and Obradovic Z. (2002) Identification and functions of usefully disordered proteins. Adv. Protein Chem. **62:** 25–49

202 Plaxco K. W. and Gross M. (1997) Cell biology: the importance of being unfolded. Nature **386:** 657–659

203 Schulz G. E. (1979) Nucleotide binding proteins. In: Molecular Mechanism of Biological Recognition, pp. 79–94, Elsevier/North-Holland Biomedical Press, New York

204 Pontius B. W. (1993) Close encounters – why unstructured, polymeric domains can increase rates of specific macromolecular association. Trends Biochem. Sci. **18:** 181–186

205 Spolar R. S. and Record M. T. (1994) Coupling of local folding to site-specific binding of proteins to DNA. Science **263:** 777–784

206 Rosenfeld R., Vajda S. and Delisi C. (1995) Flexible docking and design. Annu. Rev. Biophys Biomol. Struct. **24:** 677–700

207 Namba K. (2001) Roles of partly unfolded conformations in macromolecular self-assembly. Genes Cells **6:** 1–12

208 Uversky V. N. and Narizhneva N. V. (1998) Effect of natural ligands on the structural properties and conformational stability of proteins. Biochemistry **63:** 420–433

209 Ben Zvi A. P. and Goloubinoff P. (2001) Review: mechanisms of disaggregation and refolding of stable protein aggregates by molecular chaperones. J. Struct. Biol. **135:** 84–93