PROTEIN FOLDING AND MISFOLDING: NEURODEGENERATIVE DISEASES
Protein Folding and Misfolding: Neurodegenerative Diseases

Edited by

JUDIT OVÁDI
Hungarian Academy of Sciences, Biological Research Center, Institute of Enzymology, Hungary

and

FERENC OROSZ
Hungarian Academy of Sciences, Biological Research Center, Institute of Enzymology, Hungary

Springer
It was twenty five years ago this year that for the first time a protein underlying a form of human cerebral amyloidosis, the Icelandic-type hereditary cerebral haemorrhage was identified. This, together with the recognition that an amino acid substitution can transform the wild type cystatin C into a disease-associated amyloid-forming protein in this condition, was only a prelude to a series of important discoveries that followed. As a result, pathologically altered proteins have been brought into the centre stage of research into the pathomechanism of a number of neurodegenerative diseases, which include epidemiologically such important conditions as Alzheimer’s disease or Parkinson’s disease and, among others, also the transmissible spongiform encephalopathies, Huntington’s chorea, spinocerebellar ataxias, frontotemporal lobar degenerations and amyotrophic lateral sclerosis. Despite the diversity in the amino acid sequence of the different proteins involved in these neurological diseases, one of the common themes underlying the pathomechanisms of all these conditions is protein misfolding, aggregation – hence the term protein folding disorders –, which can trigger cascades of events ultimately resulting in synapse loss and neuron death with devastating clinical consequences in many of the most precious spheres of human existence including personality, cognition, memory, skilled movements and affection.

It is always a challenging task to unite the different topics of the individual chapters into a common theme in a multi-author volume, but the current book edited by Judit Ovadi and Ferenc Orosz fits this task admirably. The contributors of the chapters are very well-chosen to cover a good number of topical areas of neurodegenerative research. Without exception the chapters set forth clearly the current understanding of their chosen topics, which will allow both the specialist reader and the novice entering into the field to acquire the information they require to find. I have no hesitation in expecting that this wisely edited book will shortly become a well-thumbed text on the bookshelves of many research libraries and offices.

London July 2008

Tamas Revesz FRCPath
Professor of Neuropathology
Institute of Neurology
University College London
UK
The worldwide ageing of populations has brought the neurodegenerative diseases into the focus of interest. These diseases constitute large variety of pathological conditions originating from the slow, irreversible and systemic loss of cells in different regions of the brain resulting in degenerative problems with distinct clinical symptoms. The pathological behaviors are frequently associated with “proteininopathies”, the non-physiological behavior of a specific protein, affecting its processing, functioning, and/or folding. These proteins do not have stable tertiary and/or secondary structures in vivo; they enter into aberrant interactions affecting their folding state and function. A number of the diverse human neurodeenerative diseases are now recognized as conformational diseases because these are caused by aggregations of unfolded or misfolded proteins. Knowledge on the intrinsically unstructured proteins, a relatively newly recognized family of gene products as well as on the misfolded proteins produced by genetic mutations or environmental effects has been extensively accumulated in the past years. These proteins frequently cause proteolytic stress and/or enter into aberrant, non-physiological protein–protein interactions leading to sequestration of protein aggregates which are assemblies of many not-yet-identified components in addition to the deposition of well-characterized misfolded peptides and proteins. Such fate is known in the cases of Aβ peptide and tau protein in Alzheimer’s disease, α-synuclein in Parkinson’s disease, the extended polyglutamine stretch of mutant huntingtin in Huntington’s disease and the prion protein in prion diseases. These protein assemblies display diverse ultrastructures such aggresomes, fibers, oligomers or amorphous structures, however, the nature of these species concerning their cytoprotective or cytotoxic effects has not been clarified yet. The understanding of the course and pathomechanism of the diseases arising from interactions of the so called malfolded proteins is crucial for finding effective therapeutic interventions. The identification of aberrant protein–protein interaction(s) playing constitutive role in aggregate formation contributes to the development of new pharmacofors that could prevent or circumvent the development of neurodegenerative disorders in human.

The main focus of this issue is to review the molecular events initiated by unfolded or misfolded proteins leading to cell death via the development of pathological inclusions, with special emphasis on the macromolecular associations of the
malfolded proteins into characteristic ultrastructures found in the cases of neurological disorders, some of them are shown in this issue. There are papers which uncover in details the intriguing interconnections between intrinsic disorder and human neurodegenerative diseases; the characterization of the diseases in relation to their hallmark proteins and ultrastructures. Other papers provide conceptual background of the molecular mechanism of the tendency of disordered proteins for aggregation in vitro and in vivo connected with misfolding diseases. Due to the fundamental biological importance of protein aggregates, and our poor knowledge about the molecular basis or specificity of the general phenomenon of protein aggregation, this problem will be specifically discussed. In the light of the protein based neuropathology the classification of the human neurodegenerative diseases is presented. This book also reviews the structural knowledge accumulated for well-studied and for newly discovered proteins involved in paradigmatic conformational disorders with the aim to broaden our understanding of the pathomechanisms of neurodegeneration, which is crucial for finding effective therapeutic interventions that could prevent or circumvent the development of neurodegenerative disorders in humans.

Acknowledgments  We are grateful to the Hungarian National Scientific Research Fund (OTKA) and the European Union FP6-2003-LIFESCIHEALTH-I Biosim Fund for providing many years of valuable support to our research, which has also enabled us to edit this volume.

Budapest
July 2008

Judit Ovádi
Ferenc Orosz
Institute of Enzymology, BRC
Hungarian Academy of Sciences
Budapest, Hungary
Contents

1 Structural Disorder and Its Connection with Misfolding Diseases . . . . 1
   Veronika Csizmok and Peter Tompa

2 Intrinsic Disorder in Proteins Associated with Neurodegenerative
   Diseases .................................................................................. 21
   Vladimir N. Uversky

3 Dynamic Role of Ubiquitination in the Management of Misfolded
   Proteins Associated with Neurodegenerative Diseases .............. 77
   Esther S.P. Wong, Jeanne M.M. Tan and Kah-Leong Lim

4 Protein Misfolding and Axonal Protection in Neurodegenerative
   Diseases .................................................................................. 97
   Haruhisa Inoue, Takayuki Kondo, Ling Lin, Sha Mi, Ole Isacson
   and Ryosuke Takahashi

5 Endoplasmic Reticulum Stress in Neurodegeneration .................. 111
   Jeroen J.M. Hoozemans and Wiep Schepers

6 Involvement of Alpha-2 Domain in Prion Protein
   Conformationally-Induced Diseases ........................................... 133
   Luisa Ronga, Pasquale Palladino, Ettore Benedetti, Raffaele Ragone
   and Filomena Rossi

7 Synuclein Structure and Function in Parkinson’s Disease .......... 159
   David Eliezer

8 Inhibition of α-Synuclein Aggregation by Antioxidants
   and Chaperones in Parkinson’s Disease ...................................... 175
   Jean-Christophe Rochet and Fang Liu
9 Novel Proteins in α-Synucleinopathies ........................................ 207
Christine Lund Kragh and Poul Henning Jensen

10 TPPP/p25: A New Unstructured Protein Hallmarking Synucleinopathies ......................................................... 225
Ferenc Orosz, Attila Lehotzky, Judit Oláh and Judit Ovádi

11 Protein-Based Neuropathology and Molecular Classification of Human Neurodegenerative Diseases......................... 251
Gabor G. Kovacs and Herbert Budka

Index ........................................................................................................ 273
Contributors

Ettore Benedetti  Dipartimento delle Scienze Biologiche &
C.I.R.Pe.B.,Universit`a degli Studi Federico II di Napoli, Napoli, Italy

Herbert Budka  Institute of Neurology, Medical University of Vienna, AKH 4J,
POB 48, A-1097 Wien, Austria

Veronika Csizmok  Institute of Enzymology, Biological Research Center,
Hungarian Academy of Sciences, Karolina út 29, Budapest, H-1113 Hungary,
e-mail: csv@enzim.hu

David Eliezer  Weill Cornell Medical College, 1300 York Avenue, New York,
NY 10065, USA, e-mail: eliezer@med.cornell.edu

Jeroen J.M. Hoozemans  Department of Pathology, VU University Medical
Center, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands, e-mail:
jjm.hoozemans@vumc.nl

Haruhisa Inoue  Department of Neurology, Kyoto University Graduate School
of Medicine, 54 Kawahara-cho, Shogoin, Sakyoku, Kyoto 606-8507, Japan,
e-mail: haruhisa@kuhp.kyoto-u.ac.jp

Ole Isacson  Neuroregeneration Laboratories and Center for Neuroregeneration
Research at McLean Hospital/Harvard Medical School, NINDS Morris K. Udall
Parkinson’s Disease Research Center of Excellence, 115 Mill Street, Belmont, MA
02478, USA, e-mail: isacson@mclean.harvard.edu

Poul Henning Jensen  Institute of Medical Biochemistry, University of Aarhus,
Ole Worms Allé 1.170, DK-8000 Aarhus C, Denmark, e-mail: phj@biokemi.au.dk

Takayuki Kondo  Department of Neurology, Kyoto University Graduate School
of Medicine, 54 Kawahara-cho, Shogoin, Sakyoku, Kyoto 606-8507, Japan

Gabor G. Kovacs  Institute of Neurology, Medical University of Vienna, AKH
4J, POB 48, A-1097 Wien, Austria, e-mail: gabor.kovacs@meduniwien.ac.at
Christine Lund Kragh  Institute of Medical Biochemistry, University of Aarhus, Ole Worms Allé 1.170, DK-8000 Aarhus C, Denmark, e-mail: clp@biokemi.au.dk

Attila Lehotzky  Institute of Enzymology, Biological Research Center, Hungarian Academy of Sciences, Karolina út 29, Budapest, H-1113 Hungary, e-mail: lehotzky@enzim.hu

Kah-Leong Lim  Neurodegeneration Research Laboratory, National Neuroscience Institute, Singapore and Duke-NUS Graduate Medical School, Singapore; 11 Jalan Tan Tock Seng, 308433, Singapore, e-mail: Kah_Leong_Lim@nni.com.sg

Ling Lin  Neuroregeneration Laboratories and Center for Neuroregeneration Research at McLean Hospital/Harvard Medical School, NINDS Morris K. Udall Parkinson’s Disease Research Center of Excellence 115 Mill Street, Belmont, MA 02478, USA, e-mail: llin@mclean.harvard.edu

Fang Liu  Department of Medicinal Chemistry and Molecular Pharmacology, Purdue University, West Lafayette, IN 47907-2091, USA

Sha Mi  Department of DiscoveryNeurobiology, Biogen Idec Inc., 14 Cambridge Center, Cambridge, MA 02142, United States, e-mail:sha.mi@biogenidec.com

Judit Oláh  Institute of Enzymology, Biological Research Center, Hungarian Academy of Sciences, Karolina út 29, Budapest, H-1113 Hungary, e-mail: olju@enzim.hu

Ferenc Orosz  Institute of Enzymology, Biological Research Center, Hungarian Academy of Sciences, Karolina út 29, Budapest, H-1113 Hungary, e-mail: orosz@enzim.hu

Judit Ovádi  Institute of Enzymology, Biological Research Center, Hungarian Academy of Sciences, Karolina út 29, Budapest, H-1113 Hungary, e-mail: ovadi@enzim.hu

Pasquale Palladino  Dipartimento delle Scienze Biologiche & C.I.R.Pe.B., Università degli Studi Federico II di Napoli, Napoli, Italy

Raffaele Ragone  Dipartimento di Biochimica e Biofisica & CRISCEB, Seconda Università di Napoli, Napoli, Italy

Tamas Revesz  Queen Square Brain Bank, Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, WC1N 3BG, UK, e-mail: T.Revesz@ion.ucl.ac.uk

Jean-Christophe Rochet  Department of Medicinal Chemistry and Molecular Pharmacology, Purdue University, West Lafayette, IN 47907-2091, USA, e-mail: rochet@pharmacy.purdue.edu.
Luisa Ronga  Dipartimento delle Scienze Biologiche & C.I.R.Pe.B., Università degli Studi Federico II di Napoli, Napoli, Italy

Filomena Rossi  Dipartimento delle Scienze Biologiche-Sez. Biostrutture, Università degli Studi “Federico II” di Napoli, Via Mezzocannone, 16, I-80134 Napoli, Italy, e-mail: filomena.rossi@unina.it

Wiep Scheper  Neurogenetics Laboratory and Department of Neurology, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands

Ryosuke Takahashi  Department of Neurology, Kyoto University Graduate School of Medicine, 54 Kawahara-cho, Shogoin, Sakyo-ku, Kyoto 606-8507, Japan, e-mail: ryosuket@kuhp.kyoto-u.ac.jp

Jeanne M.M. Tan  Neurodegeneration Research Laboratory, National Neuroscience Institute, Singapore, Singapore

Peter Tompa  Institute of Enzymology, Biological Research Center, Hungarian Academy of Sciences, Karolina út 29, Budapest, H-1113 Hungary, e-mail: tompa@enzim.hu

Esther S.P. Wong  Neurodegeneration Research Lab, National Neuroscience Institute, Singapore and Department of Anatomy and Structural Biology, Marion Bessin Liver Research Center, Albert Einstein College of Medicine, USA

Vladimir N. Uversky  Center for Computational Biology and Bioinformatics, Department of Biochemistry and Molecular Biology, Institute for Intrinsically Disordered Protein Research, Indiana University School of Medicine, Indianapolis, Indiana 46202, USA and Institute for Biological Instrumentation, Russian Academy of Sciences 142290 Pushchino, Moscow Region, Russia And Molecular Kinetics, Inc., Indianapolis, IN 46268, USA, e-mail: vuversky@iupui.edu
Chapter 1
Structural Disorder and Its Connection with Misfolding Diseases

Veronika Csizmók and Peter Tompa

Abstract  Intrinsically disordered proteins or regions of proteins lack a well-defined structure, yet they carry out important functions often associated with the regulation of cell cycle and transcription. Due to these central roles in key cellular processes, their mutations are frequently involved in neurodegenerative diseases. These diseases are usually caused by the structural transition of disordered proteins to insoluble, highly ordered deposits termed amyloids, such a fate has been described in the case of Aβ peptide and tau protein in Alzheimer’s disease, α-synuclein in Parkinson’s disease or the polyglutamin stretch of huntingtin in Huntington’s disease and the prion protein in prion diseases. Due to the involvement of critical conformational change, these diseases are often denoted as “protein misfolding” diseases. Here we provide a brief overview of the rapidly expanding field of protein disorder to provide a conceptual background for the discussion of the essence of molecular mechanisms of these diseases. We will provide a brief overview of the field in general, directing focus on the tendency of disordered proteins for aggregation in vitro and also in vivo. We will provide some details on neurodegenerative diseases and the proteins involved. It will be shown that the underlying phenomenon of “misfolding” may also result in altering the normal function of proteins (physiological prions). We will wrap up the story by showing that the conformational transition occurs via partially ordered intermediates, which lead to a highly structured cross-β state in amyloids.

1.1 The Concept of Protein Disorder

The classical structure-function paradigm that appeared unshakable for decades rested on the correspondence between function and a well-folded 3D structure. The basic insight provided by this notion into the function of enzymes, receptors...
and structural proteins has precluded alternative views. The spectacular advance of structural biology crowned recently by the success of structural genomics programs has made it the central dogma of molecular biology that a unique structure encoded by sequence is the prerequisite of function. More than 50,000 structures deposited in the protein data bank (PDB, www.pdb.org) bear witness to the power of this paradigm. Its generality, however, does not infer universality, as indicated by the recent recognition that many proteins or regions of proteins lack a well-defined three-dimensional structure under native, physiological conditions [1–4]. The recognition that intrinsic disorder is the native, functional state of these proteins, has brought about the demand of re-assessing the structure-function paradigm [5].

The polypeptide chain of intrinsically disordered, or unstructured, proteins (IDPs/ IUPs) assumes a fluctuating ensemble of alternative conformations, which is the prerequisite of their functions. In a structural sense, IDPs occupy a continuum of states from a fully disordered state devoid of either short- or long-range intrachain interactions (random coil) to a compact state of significant secondary and tertiary contacts (molten globule) [6, 7]. These states in many aspects resemble those attained by globular proteins under highly denaturing conditions. Unlike globular proteins, however, which most often carry out their function as enzymes, small-ligand binding receptors or structural proteins, IDP functions stem from the unfolded states, and are mostly involved in regulating processes of signal transduction and transcription regulation [8–10]. Functional classification of IDPs into six categories is based on that in one category (entropic chains) their function directly stems from disorder, whereas in the other five categories their function stems from transient (display sites, chaperones) or permanent (effectors, assemblers, scavengers) binding to partner molecules [2, 3, 11].

The prevalence of structural disorder in regulatory functions results from the functional advantages structural disorder provides. Among many advantages, most often mentioned and discussed are the separation of specificity from binding strength [5], adaptability to various partners [12] and frequent involvement in post-translational modifications [13]. These and other advantages explain the advance of protein disorder in evolution, i.e. its much higher frequency in eukaryotes than prokaryotes [8–10], and its high proportion/dominance in functionally important proteins also noted in disease, such as tau protein [14], p53 [15], α-synuclein [16], prion protein [17], or BRCA1 [18]. The current most complete collection of IDPs is in the DisProt database (www.disprot.org), which contains about 500 proteins, in which biophysical evidence points to the structural disorder of about 1100 regions [19]. DisProt, and previous less-complete collections of disordered proteins enabled the development of about 25 bioinformatics predictors [20, 21]. The application of such predictors to whole genomes and/or proteomes has suggested that about 5–15% of proteins are fully disordered, and 30–50% of proteins contain at least one long disordered region in higher organisms [8–10].
1.2 Biophysical and Bioinformatics Characterization of Disorder

1.2.1 Biophysical Techniques

The primary observation of the unusual behavior of proteins of heat-stability and anomalous SDS-PAGE mobility, circular dichroism (CD) and NMR spectra suggesting a “denatured” state, as well as the frequent observation of missing coordinates from X-ray structures, have led to the formulation of the concept of protein disorder. The first collection of disorder datasets then led to the creation of bioinformatics tools which brought about the recognition of the generality of protein disorder. To respect this historical order of events, we first survey the most important biophysical methods used for recognition and characterization of disorder, followed by a brief overview of the bioinformatics methods. The physical characteristics of IDPs contrasting globular proteins is apparent with all possible approaches, which explains the multiplicity of methods that can be applied for studying protein disorder [2, 7, 20, 22].

Observations by indirect techniques may provide the first line of evidence for the disorder of a protein. IDPs are resistant to heat and low pH, which form the basis of enrichment strategies employed for their proteomic identification [23, 24]. Their aberrant mobility on SDS-PAGE, suggestive of an unusual amino acid composition, has also been frequently noted in the literature [2]. The open and exposed structural character of their unfolded polypeptide chain is also signaled by an extreme proteolytic sensitivity, which also manifests itself in their ubiquitination-independent degradation by the 20S proteasome, termed “degradation by default” [25]. Proteolytic sensitivity can not only provide a binary classification in terms of order/disorder, but the application of proteases at very low concentrations can also provide low-resolution structural information on the topology of IDPs [26]. Another indirect technique, differential scanning calorimetry provides information on the lack of a globularity, i.e., the absence of compact, cooperative structure of IDPs [27].

Hydrodynamic approaches constitute the most coherent group of techniques for the structural characterization of IDPs. The primary observables are the radius of gyration (R_g) or Stokes radius (R_S), which translate into a large apparent molecular weight (M_W). Such behavior is apparent by size-exclusion chromatography (gel-filtration), dynamic light scattering, and analytical ultracentrifugation. More thorough characterization of hydrodynamic behavior can be attained by small-angle X-ray scattering, which not only enables to determine overall dimensions of the protein, but by careful analysis of scattering intensities it provides low-resolution structural topology-model of the molecule [28]. Thus, hydrodynamic techniques not only provide evidence for disorder, but they also enable description of its type and the overall structural topology of an IDP.

Description of disorder in most detail can be achieved by spectroscopic techniques. UV fluorescence, sensitive to the exposure of Trp residues, enables the rapid
identification of IDPs. The application of a quencher, such as iodine or acrylamide, provides further evidence for the exposure of aromatic residues. CD spectroscopy is sensitive to repetitive secondary structural elements (α-helix and β-strand), or coil conformation, the latter being abundant in IDPs. Even more structural detail can be obtained by a less well-known technique, Raman optical activity measurement, which provides information on details of structure and dynamics of IDPs [29]. The most powerful spectroscopic technique for studying IDPs is NMR, which enables their atomic-level structural characterization [22]. A range of NMR observables, such as secondary chemical shifts, relaxation rates and residual dipolar coupling enable detailed description of equilibrium structural features and also dynamic characteristics of IDPs. To mention just a few examples, NMR has been used to provide evidence for the overall disorder of proteins [30–32], it enabled characterizing residual structure within IDPs [33–35], and also detailed analysis of the mechanism of binding of an IDP to its partner [36]. Recently, NMR even made possible the in vivo characterization of IDPs by the application of in-cell NMR techniques [37].

1.2.2 Bioinformatics Techniques

Followed by the recognition of protein disorder, several bioinformatics algorithms have been developed in rapid succession, which can be used to approach disorder at the residue level [20, 21]. The application of predictors to studying single proteins and/or entire proteomes has contributed basically to the development of the field. Although the predictors are based on different principles, they all rely on common attributes of IDPs, namely, that they are depleted in order-promoting amino acids (WCFIYVL) and are enriched in disorder promoting amino acids (KEPSQRA) [38]. There are more than 20 predictors of disorder, and they can be roughly classified into three groups, such as (i) predictors relying on simple statistics, (ii) predictors applying machine-learning algorithms, and (iii) predictors applying some structural considerations.

The most straightforward approach relies on simple statistics of amino acid propensities, as implemented in the charge-hydropathy plot, by plotting net charge of proteins vs. their net hydrophobicity [6]. IDPs are found in the high net charge – low mean hydrophobicity half of this 2D plane, which suggests a clear interpretation of the physical factors underlying disorder. By calculating the distribution of these features for a pre-defined sequence window, this approach can be made sequence-specific [39].

Arguably the most advanced predictors are those which rely on machine learning approaches, i.e. neural networks and support vector machines. These are trained on datasets of disorder and order, and capture the inherent differences in implicit ways. The classical neural network predictor, PONDR [40], has recently been developed to be able to distinguish between short and long disorder (VSL2) [41]. In other cases, the input data can be generated by sequence alignment, as in the case of DISOPRED2 [9], which relies on a support vector machine algorithm. The power of these
methods also comes from that they can readily accommodate other factors, such as predicted secondary structure or solvent accessibility of the polypeptide chain. Although these algorithms usually perform the best when performance of predictors is compared in the community-wide experiment “critical assessment of structure prediction algorithms” (CASP) [42], their limitation may come from uncertainties inherent in the underlying databases.

A completely different principle has been exploited in the construction of the IUPred algorithm [43, 44]. This approach uses low-resolution force-fields to estimate the total pairwise interaction energy of a (segment of a) protein. The underlying idea is that IDPs lack stable structure because their amino acid composition is not compatible with the formation of interresidue interactions in numbers sufficiently large to overcome the large unfavorable decrease in conformational entropy that accompanies folding. Because IUPred and other similar algorithms, such as Fold Unfold [45] and Ucon [46] do not rely on actual data on protein disorder, their assessment of the structural status of a protein as disordered may be considered as an independent evidence for disorder, and actually for the very existence of intrinsically disordered proteins.

1.3 Disorder In Vivo, the Effect of Crowding?

The structural ensemble of IDPs is very sensitive to variations in environmental conditions, which makes it rather difficult to appreciate the actual structural state of these proteins. Among the variety of factors, crowding caused by extremely high macromolecular concentrations is of special interest, because it may basically influence the structural state of IDPs [47, 48]. Typical concentrations of macromolecules in the cell are on the order of 300–400 mg/ml, which gives rise very large excluded volume effect that favors reactions accompanied by reduction of volume, such as folding and aggregation. In the case of denatured globular proteins crowding does promote them to assume their native-like compact states and to regain at least partial activity [49, 50]. This issue of promoting native structure is of particular importance in the case of IDPs, because it would be logical to assume that crowding may enforce them to fold, and behave as globular proteins, in vivo.

Studies addressing this issue either apply high concentrations of macromolecular crowding agents, such as Dextran or Ficoll 70 (occasionally a small molecular osmolyte, TMAO), or actually follow the behavior of the IDP within a living cell. The results are rather mixed, and they overall suggest that crowding makes IDPs to locally fold or assume more compact structural states, but never to transformation to a unique ordered state. For example, crowding had no effect on two IDPs, the KID domain of p27Kip1, and the trans-activator domain of c-Fos [51], but leads to some compaction of α-synuclein [52]. Under real in vivo conditions, i.e. within a living cell, some IDPs, such as FlgM [53] or tau protein [54] undergo partial ordering, whereas others, such as α-synuclein [37], retain their fully disordered character.
In principle, aggregation, being a second- or higher-order reaction, is particularly sensitive to the effect of crowding. The formation of aggregates is sensitive to the concentration – in fact the chemical activity – of interacting chains, which is basically influenced by the excluded volume effect [47, 55]. As shown by experiment and also theoretical considerations, crowding may increase the rate of aggregation orders of magnitude. For example, the formation of α-synuclein fibrils has a lag time of 80–90 days for a concentration of 300 μM, but addition of polyethylene glycol, Dextran or Ficoll 70 reduces this lag time to 8–10 days [56].

1.4 Disorder and Aggregation

Early on after the recognition of protein disorder, it has been realized that the open and extended conformation of IDPs may be particularly adapted to interactions leading to aggregation, making them, in principle, particularly prone to aggregation [2]. Although several of the proteins involved in amyloid diseases are IDPs, most IDPs are not known for their involvement in aggregation, which suggests that these proteins use some countermeasures against aggregation [57]. Studies of sequences of proteins involved in amyloid diseases unveil that certain features are directly related with disorder. Because the key structural feature of amyloids is an extended H-bonding network of backbone amides in a cross-β scaffold, the exposure of these moieties is key to the misfolding reaction leading to the amyloid state. In accord, deficient local shielding (under-wrapping) of backbone H-bonds is a critical factor in the amyloidogenicity of proteins [58]. Due to their total structural exposure, IDPs are inherently more prone to form amyloids than globular proteins.

In studies of protein aggregation of a range of mutants under conditions favoring the unfolded states of globular proteins, it was found that amyloidogenicity shows a significant positive correlation with hydrophobicity and β-sheet forming potential, and negative correlation with total charge [59]. These results are entirely relevant with respect to how IDPs remain soluble despite their exposed polypeptide chain. As suggested above, IDPs possess high mean net charge and low mean hydrophobicity [6], they are depleted in order-promoting amino acids (WCFIYVL) and are enriched in disorder promoting amino acids (KEPSQRA) [38]. These biases in composition act strongly against amyloid formation. By applying TANGO, the algorithm developed to assess β-aggregation propensity of proteins [60] it was found that globular proteins contain almost three times as much aggregation nucleating regions as IDPs, and formation of the ordered structure of globular proteins can only be achieved at the expense of a higher β-aggregation propensity [61]. In general, formation of structure and aggregates rely on very similar physico-chemical characteristics.

Thus, amino acid sequences of IDPs appear to significantly counteract the inherent propensity of their open structure for aggregation and amyloid formation. Limiting the occurrence of amino acids of significant β-sheet forming potential [62] is probably also of significant inhibitory potential. This may also rationalize the presence of conspicuous conserved Pro or Gly residues in proteins [63], which are inhibitory to the formation of extended β-structures, either due to their restricted
(Pro) or unrestricted (Gly) conformational freedom, serving as “guardians” against aggregation [64]. These considerations can also explain the presence of residues in IDPs known for significant β-breaking potential, such as Pro, Gln and Ser [2]. In a related study, it was observed [65] that the positions flanking aggregating stretches are enriched with residues such as Pro, Lys, Arg, Glu and Asp. These residues are either β-breakers, or are located in the bottom of the aggregation propensity scales. In the E. coli proteome, at least one of these five residues occur at the first position on either side of an aggregation-prone segment, and thus appear to act as “gatekeepers” against aggregation [65].

1.5 Disorder in Neurodegenerative Diseases

Despite these effective countermeasures, disordered proteins do show significant association with aggregation involved in neurodegenerative diseases. Because they are caused by formation of insoluble aggregates, they belong to the family of amyloidoses. Amyloids are highly ordered deposits of misfolded protein, which often originate from full-length proteins, but sometimes from processed segments. Since the diseases and the proteins involved are discussed in detail in the next chapter (See Sect. 1.4), here we concentrate on the underlying structural principles. Amyloidoses are caused by the deposition of insoluble, highly ordered fibrillar aggregates of proteins, which are not related in any aspect [57]. Amyloid diseases are classified by the causative protein that forms the amyloid (Table 1.1), the manner of deposition of the aggregate (systemic vs. tissue-specific cases, the latter primarily meaning neurodegeneration), and the cause of aggregation (sporadic, inherited, or infectious, the latter meaning prion diseases). As seen, the proteins involved might be globular (lysozyme, transthyretin or immunoglobulin), but often they are intrinsically disordered. The diseases are usually intractable and progressive, i.e. they cannot be cured, and either are caused by organ failure (primarily in systemic cases) or impairment in higher-order brain function (cognitive disorder, psychological problems, impairment of movements).

Due to their involvement in diseases, proteins of neurodegenerative diseases have been studied in great detail, and were among the first for which structural disorder

<table>
<thead>
<tr>
<th>Disease</th>
<th>Protein</th>
<th>Region</th>
<th>Structural status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alzheimer’s</td>
<td>APP</td>
<td>Aβ peptide</td>
<td>Disordered</td>
</tr>
<tr>
<td>Huntington’s</td>
<td>Huntingtin</td>
<td>polyQ region</td>
<td>Disordered</td>
</tr>
<tr>
<td>Parkinson’s</td>
<td>α-Synuclein</td>
<td>Whole protein</td>
<td>Disordered</td>
</tr>
<tr>
<td>Prion diseases</td>
<td>Prion</td>
<td>Whole protein</td>
<td>Half disordered</td>
</tr>
<tr>
<td>Lysozyme amyloidosis</td>
<td>Lysozyme</td>
<td>Whole protein</td>
<td>Ordered</td>
</tr>
<tr>
<td>Senile systemic amyloidosis</td>
<td>Transthyretin</td>
<td>Whole protein</td>
<td>Ordered</td>
</tr>
<tr>
<td>AL amyloidosis</td>
<td>Ig light chain</td>
<td>Whole protein</td>
<td>Ordered</td>
</tr>
</tbody>
</table>

The table enlists some of the best known amyloid diseases. The major point is that amyloid deposits may be formed from either ordered or disordered proteins, but in neurodegenerative disorders mostly IDPs are involved.
has been established. In Alzheimer’s disease (AD), extracellular protein deposits (senile plaques) are formed from the 40–42 amino-acid long fragment of amyloid precursor protein (APP), termed amyloid-β peptide (Aβ), which is disordered [66]. Intracellular inclusions also form in AD, from the microtubule-associated protein tau, which was among the first proteins described as disordered [14]. The causative agent of Parkinson’s disease is α-synuclein, also termed “non-A beta component of Alzheimer’s disease amyloid plaque (NACP)”, because it is a minor peptide component of the insoluble fibrillar core of AD plaque. The protein was shown to be disordered by a battery of techniques, such as heat-stability, sedimentation, CD, Fourier-transform infrared spectroscopy (FTIR) and UV spectroscopy [16]. The protein has been at the focus of intense interest ever since, which has resulted in ample detail on the structural ensemble of α-synuclein structure in vitro [67] and also in vivo [37]. Huntington’s disease, and a range of other diseases are caused by the pathologic expansion of glutamin-repeats in proteins. In Huntington’s disease, the CAG-repeat region in exon1 of huntingtin encodes for a run of Gln residues less than 40 in healthy individuals, which undergoes expansion to above 40 residues under pathologic conditions (thus the diseases are also termed CAG-repeat diseases). The repeat region is intrinsically disordered [68], and can undergo transition to the amyloid state.

A special case of amyloidoses is prion diseases, in which the transmission of the amyloid can elicit infectious propagation of the amyloid state. Prions have been first noted as non-conventional infectious entities in mammals, which were shown later to be proteins which may exist in two different structural states, a cellular state and a prion state. The prion state is contagious, and is implicated in a variety of diseases, such as kuru and Creutzfeldt-Jakob disease of humans, bovine spongiform encephalopathy, and scrapie of sheep [69]. Transmission of prions results from that the scrapie state can convert the cellular form to the scrapie form in a self-sustaining, autocatalytic reaction. The two forms are identical at the level of sequence or post-translational modifications [70], and thus the only information that distinguishes them is protein conformation. The structure of the cellular form solved by NMR has an N-terminal disordered and a C-terminal ordered half [17, 71]. Since the cellular form is dominated by disorder and α-helical regions, whereas the scrapie state is largely β-strand, the prion disease constitutes a special class of transmissible protein misfolding diseases [72].

1.6 Physiological Prions

As suggested in the previous section, prions have the dreadful connotation of causing lethal and somewhat mysterious diseases. It is generally held that the propagation of prion diseases results from the conversion of the cellular form to the scrapie state in an autocatalytic reaction [57]. In this section we will discuss that the above structural principle of prion propagation, i.e., the autocatalytic structural conversion from a soluble form to the highly ordered amyloid state, may also serve
the physiological function of proteins. In the case of about 10 proteins it has been shown that their normal cellular function results from their capacity to undergo prion-like structural conversion. Unlike their pathological counterparts, these physiological amyloids/prions do no harm to their host cells, but may confer adaptive advantages under certain conditions [73]. The variety of cases can be exemplified by the curli protein of bacteria, involved in biofilm formation and host invasion, URE2p of yeast involved in the regulation of nitrogen catabolism, or Pmel17 of humans, which functions in scaffolding and sequestration of toxic intermediates during melanin synthesis. Often these proteins are noted for the presence of Q/N-rich, disordered, portable prion domains [74]. Their function and action can be best illustrated by two interesting well-characterized examples, Sup35p of yeast and cytoplasmic polyadenylation element-binding protein (CPEB) of Drosophila melanogaster.

Sup35p in yeast is a protein component of the translational termination complex. Intriguingly, it has been discovered as a non-Mendelian genetic element, [PSI+], which causes translational read-through in yeast cells [75]. Later, it has been recognized that the genetic element corresponds to the altered structural state of the cellular protein, Sup35p, which is composed of a Q/N-rich disordered amino-terminal domain and a globular carboxy-terminal domain. When the amino-terminal domain attains an amyloid-like prion conformation [76], it prevents the globular domain from taking part in the translation termination complex. The physiological readout of this change is the inability of the cell to terminate translation at stop codons, and the resulting read-through might be functionally advantageous under some circumstances [77].

A completely different example is the CPEB protein of D. melanogaster. This neuronal protein regulates mRNA translation by promoting polyadenylation and activation of mRNA localized in the cytoplasm [78, 79]. Its amino-terminal Q/N-rich domain has the capacity to undergo a transition to a prion state, as demonstrated by fusion constructs in yeast. In the activated synapses of fruit-fly it converts to the prion state and provides a molecular marker of the synapses. Its activated prion-like form stimulates translation of CPEB-regulated mRNA, and promotes synaptic growth associated with the maintenance of long-term facilitation. In all, this prion/amyloid functions in synaptic communication and memory formation [78, 79].

1.7 Structural Transition to Amyloid: Partially Folded Intermediates

Although the proteins involved in amyloid formation have practically nothing in common [57], their structural transitions to the amyloid state share common characteristics, both in terms of their kinetics, the mechanism of the structural transition and the final structure attained.
The kinetics of amyloid formation shows two characteristic features, i.e., i) the process involves a lag-phase, i.e. the rate-limiting formation of a critical seed, followed by an exponentially accelerating growth phase, and ii) the lag-phase can be abolished by the addition of small pieces of amyloid, i.e., pre-formed seeds. These features are reminiscent of the process of crystallization, and amyloid formation can be considered as one-dimensional crystal growth. To account for these observations, two models have been developed. The model of “nucleation-polymerization” and “template-assistance” [80, 81] differ in the thermodynamic nature of the critical step. In nucleation-polymerization, it is assumed that the structurally altered molecule is less stable than the original protein species, and it only becomes stable when incorporated into an oligomeric (amyloid) form. Thus, the rate-limiting step is the assembly of a seed of critical size, followed by the practically uninhibited transformation of further molecules upon interaction with the seed. The key assumption of the other model, template-assistance, is that the transformed state is inherently more stable than the solution state, but it is kinetically inaccessible due to a high energy barrier. Molecules already transformed can lower the energy barrier, and bring about conversion in an autocatalytic conversion. In this case, the rate-limiting step is the formation of an effective catalytic molecule. Both models adequately describe the kinetic course of the reaction, and they actually mechanistically converge if the seed size in the template-assistance model is thought to be a monomer.

There are also mechanistic parallels in the misfolding process that leads to the formation of amyloids [82], which appear to apply to both globular and disordered proteins. In the case of globular proteins, fibrillation occurs when the native structure is partially destabilized, because to arrive at the common cross-β structure profound conformational rearrangements have to occur, which cannot take place within the structural confines of the native globular state. In accord, most mutations associated with accelerated fibrillation of globular proteins destabilize the native structure, as demonstrated in the case of lysozyme [83], transthyretin [84], and immunoglobulin light chains [85]. Destabilization of structure by non-native conditions, such as low- or high pH, high temperatures, or the presence of denaturants, also lead to an increased fibrillation, as shown in the case of the SH3 domain of PI3K [86] and the Fn III module of murine fibronectin [87]. The generality of this relation and the importance of an increase in the concentration of partially folded conformers [82, 88] is also underscored by that amyloidogenicity of proteins can be significantly reduced by stabilization of the native structure by ligand binding, for example [89]. The critical involvement of partially structured intermediates is also apparent in the case of IDPs, where the primary step of fibrillogenesis is the stabilization of a partially folded conformation. It has been shown in the case of α-synuclein [90], or islet amyloid polypeptide (IAPP) [91] that the presence of amyloidogenic intermediates is strongly correlated with the enhanced formation of fibrils (Fig. 1.1). In all, it appears that the structural prerequisite of amyloid formation is the transformation of a polypeptide chain into a partially folded conformation.
Fig. 1.1 Transition to the amyloid state via partially ordered intermediate. The experiment shows aggregation of α-synuclein at increasing temperatures from 3.0°C up to 92.0°C in equal increments (11 temperatures). (A) The kinetics of fibril formation monitored by the enhancement of thioflavin T fluorescence, at four temperatures, 27°C (circles), 37°C (inverted triangles), 47°C (squares), and 57°C (diamonds). (B) CD spectra at all temperatures, which show that upon increasing the temperature, in parallel with shortening the lag phase, there is an increase in residual structure of the protein. Adapted with permission from Ref. [90] (Uversky et al. 2001, J Biol Chem 276:10737–10744)

1.8 The Structure of Amyloid: Cross-Beta Models and Flexibility

The third unifying feature of amyloids involved in misfolding diseases is that in spite of the great variety of protein precursor, the resulting amyloid is structurally very similar in most of the cases. Amyloid fibrils visualized by transmission electron microscopy or atomic force microscopy usually consist of 2–6 protofilaments, each about 2–5 nm in diameter, twisted together to form rope-like fibrils typically
7–13 nm wide [92]. As shown by X-ray fiber diffraction, the polypeptide chain runs perpendicular to the axis of the fiber in a β-strand conformation, thus forming an extended β-sheet along the fiber. The structure is highly ordered, and in this sense clearly differs from general protein aggregates [93]. Structural uniformity of amyloids is also signaled by their common tinctorial properties, because they all can be stained by specific dyes such as thioflavin T and Congo red.

Whereas gross similarities of amyloid obtained from different proteins have been apparent for long, establishing its structure at high-resolution remained intractable for many years. Recently, the combined application of solid-state NMR, X-ray crystallography, electron microscopy and electron paramagnetic resonance (EPR) spectroscopy, have begun to provide atomic-level structural information on the structure of amyloids fibers [57, 93]. For example, the X-ray crystal structure of a model amyloid, the heptapeptide Gly-Asn-Asn-Gln-Gln-Asn-Tyr of yeast prion Sup35p [76], suggested a pair of parallel β-sheets composed of β-strands contributed by individual peptide molecules. The strands are stacked, parallel and are located in register in both sheets. The side-chains of the two sheets interdigitate so tightly that water is excluded from the interface, which lead to suggesting the model “steric zipper” (Fig. 1.2), to extend on the previous model “polar zipper” of the extended

![Fibril axis](image)

**Fig. 1.2 Steric-zipper structural model of amyloids.** The fibril formed by a heptapeptide segment (GNNQQNY) from the N-terminus of the yeast prion protein Sup35p. Reproduced with permission from Ref. [76] (Nelson et al. 2005, Nature 435:773–778)
H-bond network supporting the β-sheet structure of polyQ amyloids [94]. Similar results were obtained in the case of the fast-folding WW domain FBP28 [95].

Solid-state NMR results were combined with computational energy minimization procedures to obtain a detailed picture of the amyloid fibrils formed from the Aβ(1–40) peptide of AD [96]. The molecule makes up two β-strands, connected by a short loop, stacked upon each other, parallel and are in register. The two strands participate in the formation of two distinct sheets within the same protofilament. EPR spectroscopy, in which spectra from a series of labeled molecules have been obtained, also lent support to the highly structured, parallel and in-register arrangement of the strands [97]. A similar arrangement, i.e., single-molecule layers that stack on top of one another with parallel, in-register alignment of β-strands has been observed by the same technique in the case of fibrils formed from α-synuclein [98] and human prion protein [99].

These and many other studies have corroborated that amyloid fibrils have a tightly-packed cross-β core region, which lends stability to the structure. Outside of the core the structure is much less defined, the polypeptide chain is exposed and rather flexible, and is often explicitly stated as disordered. For example, in the case of the Aβ(1–42) molecule in AD, residues 13–21 and 30–39 are highly structured in the fibrils by EPR spectroscopy, whereas high flexibility and exposure to the solvent within the N-terminal region is apparent [97], also corroborated by hydrogen-deuterium (H/D) exchange, limited proteolysis and Pro-scanning mutagenesis [57]. In the case of α-synuclein, restricted motility is apparent in the segment 35–97, which roughly corresponds to the NAC region, whereas outside this region the chain is rather flexible in the fibril by EPR [98]. H/D exchange has also shown that in β2-microglobulin fibrils most residues in the middle segment form a rigid beta-sheet core, whereas the N- and C-termini are excluded from this core [100]. In the case of amylin, residues 12–17, 22–27, and 31–37 form stacked β-sandwiches, whereas the N-terminal tail is disordered with a disulfide bridge between Cys2 and Cys7 [101]. The fibril formed by the NM region of Sup35p is largely stabilized by interactions between residues that belong to region N, whereas regions 1–20 and 158–250 (the latter being M) remain largely disordered, as shown by fluorescence proximity analysis [102]. Similar conclusions can be drawn in the case of the human prion, in which the protease-resistant core corresponds to about 140 amino acids encompassing region 90–230, whereas the N-terminal 90 amino acids are largely disordered, as shown by limited proteolysis [69] and electron crystallography [103].

1.9 Conclusions

Neurodegenerative diseases are a major health problem in developed countries, resulting primarily from the deposition of neuronal proteins in the form of insoluble, highly ordered protein aggregates in the brain. Often, the protein or segment of protein involved in the disease is intrinsically disordered, and undergoes a major structural transmission toward the cross-β structure characteristic of the amyloid
fibrils. Understanding the cause of transition and the structural details of both the soluble and aggregated states is a long way ahead, but inevitably will provide the insight required for designing successful intervention strategies against these debilitating diseases.

Acknowledgments This work was supported by grants K60694 and NK71582 from the Hungarian Scientific Research Fund (OTKA), and ETT 245/2006 from the Hungarian Ministry of Health.

Abbreviations

Aβ amyloid β-peptide  
AD Alzheimer’s disease  
APP amyloid precursor protein  
CD circular dichroism  
CPEB cytoplasmic polyadenylation element-binding protein  
EPR electron paramagnetic resonance  
IDP intrinsically disordered protein

References


