

PROTEIN FOLDING AND MISFOLDING: NEURODEGENERATIVE  
DISEASES

# FOCUS ON STRUCTURAL BIOLOGY

---

Volume 7

---

*Series Editor*

**ROB KAPTEIN**

*Bijvoet Center for Biomolecular Research,  
Utrecht University, The Netherlands*

# 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

*Editors*

Judit Ovádi  
Hungarian Academy of  
Sciences  
Biological Research Center  
Institute of Enzymology  
Karolina út 29  
Budapest 1113  
Hungary  
ovadi@enzim.hu

Ferenc Orosz  
Hungarian Academy of  
Sciences  
Biological Research Center  
Institute of Enzymology  
Karolina út 29  
Budapest 1113  
Hungary  
orosz@enzim.hu

ISBN: 978-1-4020-9433-0

e-ISBN: 978-1-4020-9434-7

DOI 10.1007/978-1-4020-9434-7

Library of Congress Control Number: 2008938904

© Springer Science+Business Media B.V. 2009

No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

# Foreword

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*

## Preface

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 “proteinopathies”, 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 neurodegenerative 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 A $\beta$  peptide and tau protein in Alzheimer’s disease,  $\alpha$ -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 as 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 malformed 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

malformed 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

<b>1 Structural Disorder and Its Connection with Misfolding Diseases . . . .</b>	<b>1</b>
Veronika Csizmók and Peter Tompa	
<b>2 Intrinsic Disorder in Proteins Associated with Neurodegenerative Diseases . . . . .</b>	<b>21</b>
Vladimir N. Uversky	
<b>3 Dynamic Role of Ubiquitination in the Management of Misfolded Proteins Associated with Neurodegenerative Diseases . . . . .</b>	<b>77</b>
Esther S.P. Wong, Jeanne M.M. Tan and Kah-Leong Lim	
<b>4 Protein Misfolding and Axonal Protection in Neurodegenerative Diseases . . . . .</b>	<b>97</b>
Haruhisa Inoue, Takayuki Kondo, Ling Lin, Sha Mi, Ole Isacson and Ryosuke Takahashi	
<b>5 Endoplasmic Reticulum Stress in Neurodegeneration . . . . .</b>	<b>111</b>
Jeroen J.M. Hoozemans and Wiep Scheper	
<b>6 Involvement of Alpha-2 Domain in Prion Protein Conformationally-Induced Diseases . . . . .</b>	<b>133</b>
Luisa Ronga, Pasquale Palladino, Ettore Benedetti, Raffaele Ragone and Filomena Rossi	
<b>7 Synuclein Structure and Function in Parkinson's Disease . . . . .</b>	<b>159</b>
David Eliezer	
<b>8 Inhibition of <math>\alpha</math>-Synuclein Aggregation by Antioxidants and Chaperones in Parkinson's Disease . . . . .</b>	<b>175</b>
Jean-Christophe Rochet and Fang Liu	



<b>9 Novel Proteins in <math>\alpha</math>-Synucleinopathies</b> . . . . .	207
Christine Lund Kragh and Poul Henning Jensen	
<b>10 TPPP/p25: A New Unstructured Protein Hallmarking Synucleinopathies</b> . . . . .	225
Ferenc Orosz, Attila Lehotzky, Judit Oláh and Judit Ovádi	
<b>11 Protein-Based Neuropathology and Molecular Classification of Human Neurodegenerative Diseases</b> . . . . .	251
Gabor G. Kovacs and Herbert Budka	
<b>Index</b> . . . . .	273

# Contributors

**Ettore Benedetti** Dipartimento delle Scienze Biologiche & C.I.R.Pe.B., Università 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 Csizmók** 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, Sakyo-ku, 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, Sakyo-ku, 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 A $\beta$  peptide and tau protein in Alzheimer's disease,  $\alpha$ -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- $\beta$  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

---

V. Csizmók (✉)

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

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](http://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],  $\alpha$ -synuclein [16], prion protein [17], or BRCA1 [18]. The current most complete collection of IDPs is in the DisProt database ([www.disprot.org](http://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 ( $\alpha$ -helix and  $\beta$ -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-hydrophobicity 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 DISO-PRED2 [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 p27<sup>Kip1</sup>, and the trans-activator domain of c-Fos [51], but leads to some compaction of  $\alpha$ -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  $\alpha$ -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  $\alpha$ -synuclein fibrils has a lag time of 80–90 days for a concentration of 300  $\mu$ M, 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- $\beta$  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  $\beta$ -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  $\beta$ -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  $\beta$ -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  $\beta$ -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  $\beta$ -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  $\beta$ -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  $\beta$ -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 1.1** Amyloid diseases

Disease	Protein	Region	Structural status
Alzheimer’s	APP	A $\beta$ peptide	Disordered
Huntington’s	Huntingtin	polyQ region	Disordered
Parkinson’s	$\alpha$ -Synuclein	Whole protein	Disordered
Prion diseases	Prion	Whole protein	Half disordered
Lysozyme amyloidosis	Lysozyme	Whole protein	Ordered
Senile systemic amyloidoses	Transthyretin	Whole protein	Ordered
AL amyloidosis	Ig light chain	Whole protein	Ordered

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- $\beta$  peptide ( $A\beta$ ), 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  $\alpha$ -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  $\alpha$ -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  $\alpha$ -helical regions, whereas the scrapie state is largely  $\beta$ -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<sup>+</sup>], 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].

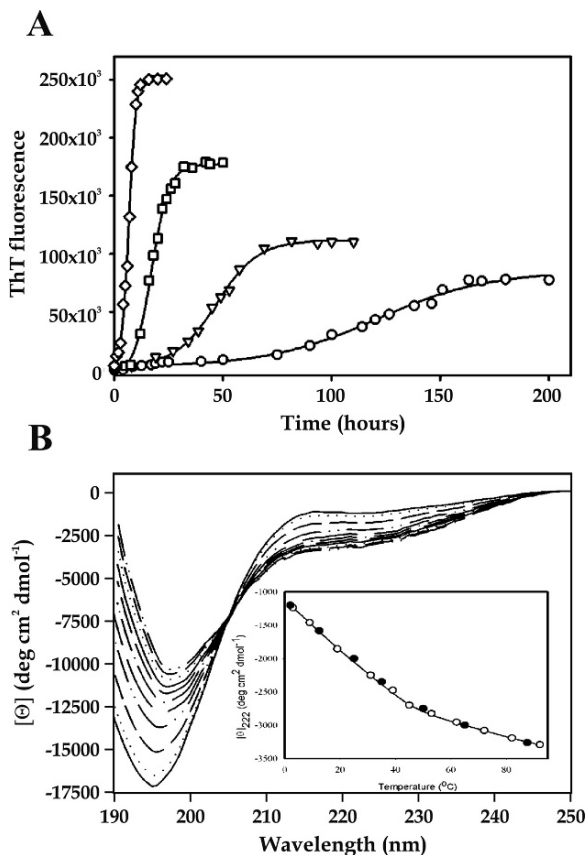
A completely different example is the CPEB protein of *D. melanogaster*. This neuronal protein regulates mRNA translation by promoting polyadenylation and activation of mRNA localized in the cytoplasm [78, 79]. Its amino-terminal Q/N-rich domain has the capacity to undergo a transition to a prion state, 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- $\beta$  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  $\alpha$ -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.



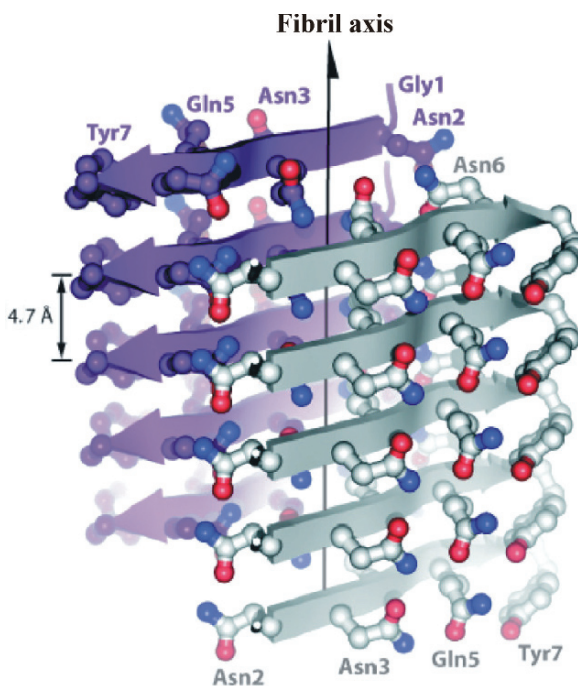
**Fig. 1.1 Transition to the amyloid state via partially ordered intermediate.** The experiment shows aggregation of  $\alpha$ -synuclein at increasing temperatures from 3.0°C up to 92.0°C in equal increments (11 temperatures). (A) The kinetics of fibril formation monitored by the enhancement of thioflavin 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  $\beta$ -strand conformation, thus forming an extended  $\beta$ -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  $\beta$ -sheets composed of  $\beta$ -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



**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  $\beta$ -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 A $\beta$ (1–40) peptide of AD [96]. The molecule makes up two  $\beta$ -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  $\beta$ -strands has been observed by the same technique in the case of fibrils formed from  $\alpha$ -synuclein [98] and human prion protein [99].

These and many other studies have corroborated that amyloid fibrils have a tightly-packed cross- $\beta$  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 A $\beta$ (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  $\alpha$ -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  $\beta$ 2-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  $\beta$ -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 transition toward the cross- $\beta$  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

A $\beta$	amyloid $\beta$ -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

1. Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6:197–208
2. Tompa P (2002) Intrinsically unstructured proteins. *Trends Biochem Sci* 27:527–533
3. Tompa P (2005) The interplay between structure and function in intrinsically unstructured proteins. *FEBS Lett* 579:3346–3354
4. Uversky VN, Oldfield CJ, Dunker AK (2005) Showing your ID: intrinsic disorder as an ID for recognition, regulation and cell signaling. *J Mol Recognit* 18:343–384
5. Wright PE, Dyson HJ (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J Mol Biol* 293:321–331
6. Uversky VN, Gillespie JR, Fink AL (2000) Why are “natively unfolded” proteins unstructured under physiologic conditions? *Proteins* 41:415–427
7. Uversky VN (2002) Natively unfolded proteins: A point where biology waits for physics. *Protein Sci* 11:739–756
8. Iakoucheva L, Brown C, Lawson J, Obradovic Z, Dunker A (2002) Intrinsic disorder in cell-signaling and cancer-associated. *Proteins J Mol Biol* 323:573–584
9. Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT (2004) Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. *J Mol Biol* 337:635–645
10. Tompa P, Dosztányi Z, Simon I (2006) Prevalent structural disorder in *E coli* and *S cerevisiae* proteomes. *J Proteome Res* 5:1996–2000
11. Dunker AK, Brown CJ, Lawson JD, Iakoucheva LM, Obradovic Z (2002) Intrinsic disorder and protein function. *Biochemistry* 41:6573–6582
12. Tompa P, Szász C, Buday L (2005) Structural disorder throws new light on moonlighting. *Trends Biochem Sci* 30:484–489
13. Iakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK (2004) The importance of intrinsic disorder for protein phosphorylation. *Nucleic Acids Res* 32:1037–1049

14. Schweers O, Schonbrunn-Hanebeck E, Marx A, 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
15. Bell S, Klein C, Muller L, Hansen S, Buchner J (2002) p53 contains large unstructured regions in its native state. *J Mol Biol* 322:917–927
16. Weinreb PH, Zhen W, Poon AW, Conway KA, Lansbury PT Jr (1996) NACP, a protein implicated in Alzheimer's disease and learning, is natively unfolded. *Biochemistry* 35:13709–13715
17. Lopez Garcia F, Zahn R, Riek R, Wuthrich K (2000) NMR structure of the bovine prion protein. *Proc Natl Acad Sci USA* 97:8334–8339
18. Mark WY, Liao JC, Lu Y, Ayed A, Laister R, Szymczyna B, Chakrabartty A, Arrowsmith CH (2005) Characterization of segments from the central region of BRCA1: an intrinsically disordered scaffold for multiple protein-protein and protein-DNA interactions? *J Mol Biol* 345:275–287
19. Sickmeier M, Hamilton JA, LeGall T, Vacic V, Cortese MS, Tantos A, Szabo B, Tompa P, Chen J, Uversky VN et al. (2007) DisProt: the database of disordered proteins. *Nucleic Acids Res* 35:D786–D793
20. Ferron F, Longhi S, Canard B, Karlin D (2006) A practical overview of protein disorder prediction methods. *Proteins* 65:1–14
21. Dosztányi Z, Sándor M, Tompa P, Simon I (2007) Prediction of protein disorder at the domain level. *Curr Protein Pept Sci* 8:161–171
22. Dyson HJ, Wright PE (2004) Unfolded proteins and protein folding studied by NMR. *Chem Rev* 104:3607–3622
23. Cortese MS, Baird JP, Uversky VN, Dunker AK (2005) Uncovering the unfoldome: enriching cell extracts for unstructured proteins by acid treatment. *J Proteome Res* 4:1610–1618
24. Galea CA, Pagala VR, Obenauer JC, Park CG, Slaughter CA, Kriwacki RW (2006) Proteomic studies of the intrinsically unstructured mammalian proteome. *J Proteome Res* 5:2839–2848
25. Tsvetkov P, Asher G, Paz A, Reuven N, Sussman JL, Silman I, Shaul Y (2008) Operational definition of intrinsically unstructured protein sequences based on susceptibility to the 20S proteasome. *Proteins* 70:1357–1366
26. Csizmók V, Bokor M, Bánki P, Klement E, Medzihradzky KF, Friedrich P, Tompa K, Tompa P (2005) Primary contact sites in intrinsically unstructured proteins: the case of calpastatin and microtubule-associated protein 2. *Biochemistry* 44:3955–3964
27. Szollosi E, Bokor M, Bodor A, Perczel A, Klement E, Medzihradzky KF, Tompa K, Tompa P (2008) Intrinsic structural disorder of DF31, a Drosophila protein of chromatin decondensation and remodeling activities. *J Proteome Res* 7:2291–2299
28. von Ossowski I, Eaton JT, Czjzek M, Perkins SJ, Frandsen TP, Schulein M, Panine P, Henrissat B, Receveur-Brechot V (2005) Protein disorder: conformational distribution of the flexible linker in a chimeric double cellulase. *Biophys J* 88:2823–2832
29. Zhu F, Kapitan J, Tranter GE, Pudney PD, Isaacs NW, Hecht L, Barron LD (2007) Residual structure in disordered peptides and unfolded proteins from multivariate analysis and ab initio simulation of Raman optical activity data. *Proteins* 70:823–833
30. Watts JD, Cary PD, Sautiere P, Crane-Robinson C (1990) Thymosins: both nuclear and cytoplasmic proteins. *Eur J Biochem* 192:643–651
31. Dobson CM (1993) Flexible friends *Current Biology* 3:530–532
32. Kriwacki RW, Hengst L, Tennant L, Reed SI, Wright PE (1996) Structural studies of p21<sup>Waf1/Cip1/Sdi1</sup> in the free and Cdk2-bound state: conformational disorder mediates binding diversity. *Proc Natl Acad Sci USA* 93:11504–11509
33. Daughdrill GW, Hanely LJ, Dahlquist FW (1998) The C-terminal half of the anti-sigma factor FlgM contains a dynamic equilibrium solution structure favoring helical conformations. *Biochemistry* 37:1076–1082

34. Sivakolundu SG, Bashford D, Kriwacki RW (2005) Disordered p27Kip1 exhibits intrinsic structure resembling the Cdk2/cyclin A-bound conformation. *J Mol Biol* 353: 1118–1128
35. Libich DS, Harauz G (2008) Backbone dynamics of the 18.5 kDa isoform of myelin basic protein reveals transient alpha-helices and a calmodulin-binding site. *Biophys J* 94: 4847–4866
36. Sugase K, Dyson HJ, Wright PE (2007) Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* 447:1021–1025
37. McNulty BC, Young GB, Pielak GJ (2006) Macromolecular crowding in the Escherichia coli periplasm maintains alpha-synuclein disorder. *J Mol Biol* 355:893–897
38. Dunker AK, Lawson JD, Brown CJ, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CM, Hipps KW, Ausio J et al. (2001) Intrinsically disordered protein. *J Mol Graphics Modelling* 19:26–59
39. Prilusky J, Felder CE, Zeev-Ben-Mordehai T, Rydberg EH, Man O, Beckmann JS, Silman I, Sussman JL (2005) FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics* 21:3435–3438
40. Garner E, Cannon P, Romero P, Obradovic Z, Dunker AK (1998) Predicting disordered regions from amino acid sequence: Common Themes despite differing structural characterization. *Genome Inform Ser Workshop Genome Inform* 9:201–213
41. Peng K, Radivojac P, Vucetic S, Dunker AK, Obradovic Z (2006) Length-dependent prediction of protein intrinsic disorder. *BMC Bioinformatics* 7:208
42. Bordoli L, Kiefer F, Schwede T (2007) Assessment of disorder predictions in CASP7. *Proteins* 69 (Suppl 8):129–136
43. Dosztányi Z, Csizmók V, Tompa P, Simon I (2005) IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 21:3433–3434
44. Dosztányi Z, Csizmók V, Tompa P, Simon I (2005) The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. *J Mol Biol* 347:827–839
45. Galzitskaya OV, Garbuzynskiy SO, Lobanov MY (2006) Fold Unfold: web server for the prediction of disordered regions in protein chain. *Bioinformatics* 22:2948–2949
46. Schlessinger A, Punta M, Rost B (2007) Natively unstructured regions in proteins identified from contact predictions. *Bioinformatics* 23:2376–2384
47. Ellis RJ (2001) Macromolecular crowding: obvious but under appreciated. *Trends Biochem Sci* 26:597–604
48. Minton AP (2005) Models for excluded volume interaction between an unfolded protein and rigid macromolecular cosolutes: macromolecular crowding and protein stability revisited. *Biophys J* 88:971–985
49. Baskakov I, Bolen DW (1998) Forcing thermodynamically unfolded proteins to fold. *J Biol Chem* 273:4831–4834
50. Qu Y, Bolen DW (2002) Efficacy of macromolecular crowding in forcing proteins to fold. *Biochem Chem* 101–102:155–165
51. Flaugh SL, Lumb KJ (2001) Effects of macromolecular crowding on the intrinsically disordered proteins c-Fos and p27(Kip1). *Biomacromolecules* 2:538–540
52. Morar AS, Olteanu A, Young GB, Pielak GJ (2001) Solvent-induced collapse of alpha-synuclein and acid-denatured cytochrome c. *Protein Sci* 10:2195–2199
53. Dedmon MM, Patel CN, Young GB, Pielak GJ (2002) FlgM gains structure in living cells. *Proc Natl Acad Sci USA* 99:12681–12684
54. Bodart JF, Wieruszski JM, Amniai L, Leroy A, Landrieu I, Rousseau-Lescuyer A, Vilain JP, Lippens G (2008) NMR observation of Tau in Xenopus oocytes. *J Magn Reson* 192: 252–257
55. Ellis RJ, Minton AP (2006) Protein aggregation in crowded environments. *Biol Chem* 387:485–497

56. Uversky VN, E MC, Bower KS, Li J, Fink AL (2002) Accelerated alpha-synuclein fibrillation in crowded milieu. *FEBS Lett* 515:99–103
57. Chiti F, Dobson CM (2006) Protein misfolding, functional amyloid, and human disease. *Annu Rev Biochem* 75:333–366
58. Fernandez A, Kardos J, Scott LR, Goto Y, Berry RS (2003) Structural defects and the diagnosis of amyloidogenic propensity. *Proc Natl Acad Sci USA* 100:6446–6451
59. Chiti F, Stefani M, Taddei N, Ramponi G, Dobson CM (2003) Rationalization of the effects of mutations on peptide and protein aggregation rates. *Nature* 424:805–808
60. Fernandez-Escamilla AM, Rousseau F, Schymkowitz J, Serrano L (2004) Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat Biotechnol* 22:1302–1306
61. Linding R, Schymkowitz J, Rousseau F, Diella F, Serrano L (2004) A comparative study of the relationship between protein structure and beta-aggregation in globular and intrinsically disordered proteins. *J Mol Biol* 342:345–353
62. Williams RM, Obradovic Z, Mathura V, Braun W, Garner EC, Young J, Takayama S, Brown CJ, Dunker AK (2001) The protein non-folding problem: amino acid determinants of intrinsic order and disorder. *Pac Symp Biocomput* 6:89–100
63. Monsellier E, Chiti F (2007) Prevention of amyloid-like aggregation as a driving force of protein evolution. *EMBO Rep* 8:737–742
64. Parrini C, Taddei N, Ramazzotti M, Degl'Innocenti D, Ramponi G, Dobson CM, Chiti F (2005) Glycine residues appear to be evolutionarily conserved for their ability to inhibit aggregation. *Structure* 13:1143–1151
65. Rousseau F, Serrano L, Schymkowitz JW (2006) How evolutionary pressure against protein aggregation shaped chaperone specificity. *J Mol Biol* 355:1037–1047
66. Kirkitadze MD, Condon MM, Teplow DB (2001) Identification and characterization of key kinetic intermediates in amyloid beta-protein fibrillogenesis. *J Mol Biol* 312:1103–1119
67. Dedmon MM, Lindorff-Larsen K, Christodoulou J, Vendruscolo M, Dobson CM (2005) Mapping long-range interactions in alpha-synuclein using spin-label NMR and ensemble molecular dynamics simulations. *J Am Chem Soc* 127:476–477
68. Vitalis A, Wang X, Pappu RV (2007) Quantitative characterization of intrinsic disorder in polyglutamine: insights from analysis based on polymer theories. *Biophys J* 93:1923–1937
69. Prusiner SB (1998) Prions. *Proc Natl Acad Sci USA* 95:13363–13383
70. Stahl N, Baldwin MA, Teplow DB, Hood L, Gibson BW, Burlingame AL, Prusiner SB (1993) Structural studies of the scrapie prion protein using mass spectrometry and amino acid sequencing. *Biochemistry* 32:1991–2002
71. Donne DG, Viles JH, Groth D, Mehlhorn I, James TL, Cohen FE, Prusiner SB, Wright PE, Dyson HJ (1997) Structure of the recombinant full-length hamster prion protein PrP(29-231): the N terminus is highly flexible. *Proc Natl Acad Sci USA* 94:13452–13457
72. Chien P, Weissman JS, DePace AH (2004) Emerging principles of conformation-based prion inheritance. *Annu Rev Biochem* 73:617–656
73. Fowler DM, Koulov AV, Balch WE, Kelly JW (2007) Functional amyloid—from bacteria to humans. *Trends Biochem Sci* 32:217–224
74. Wickner RB, Taylor KL, Edskes HK, Maddelein ML (2000) Prions: Portable prion domains. *Curr Biol* 10:R335–R337
75. Lindquist S (1997) Mad cows meet psi-chotic yeast: the expansion of the prion hypothesis. *Cell* 89:495–498
76. Nelson R, Sawaya MR, Balbirnie M, Madsen AO, Riekelt C, Grothe R, Eisenberg D (2005) Structure of the cross-beta spine of amyloid-like fibrils. *Nature* 435:773–778
77. Li L, Lindquist S (2000) Creating a protein-based element of inheritance. *Science* 287:661–664
78. Si K, Giustetto M, Etkin A, Hsu R, Janisiewicz AM, Miniaci MC, Kim JH, Zhu H, Kandel ER (2003) A neuronal isoform of CPEB regulates local protein synthesis and stabilizes synapse-specific long-term facilitation in aplysia. *Cell* 115:893–904