PREFACE

This book presents recent advances in proteomics research. Topics discussed include redox proteomics approaches to the evaluation of UV-induced protein oxidative damage; decoding amino acid sequences of proteins using inter-residue average distance statistics to extract information on protein folding mechanisms and modelling of protein folding and prediction of rate based on nucleation mechanism.

Chapter 1 - UV-induced photomodification of proteins has been implicated in damage toa broad range of different protein-based substrates. Protein photo-oxidation has been linked to such diverse degenerative processes as human hair and skin damage, degradation of pharmaceuticals, reduced food quality and nutrition, discoloration of natural fibres, eye lens opacification, loss of enzymatic activity and crop damage. Ongoing ozone reduction in the stratosphere, with resultant increased exposure to UVB, hasgenerated further impetus for advancing the understanding and controlling the complex protein degradation mechanisms and pathways underlying these damage processes.

For proteins, photo-oxidative damage is generally attributable to the generation and attack of reactive oxygen species (ROS) on amino acid residue sidechains and the protein backbone. Singlet oxygen and hydroxyl radicals, in particular, play a significant role in initiating protein modification.

This paper explores and discusses the complex cascade of modifications induced within proteins through exposure to irradiation, the effect of photo-oxidation on protein substrate properties from the protein primary level through to higher order structure, and the mechanisms underpinning protein photo-oxidation. It will also discuss exciting new approaches to the

characterisation, location, tracking, and ultimately control of photo-oxidation within proteins.

Chapter 2 - One of the ultimate goals of bioinformatics is to discover the principles governing protein folding. However, information on folding mechanisms of proteins cannot be easily decoded from amino acid sequences using standard bioinformatics techniques such as multiple sequence alignment and so on. The authors do not sufficiently understand how the determinants of protein folding are encoded in amino acid sequences. In this review, the authors summarize how information about protein folding mechanisms can be extracted from amino acid sequences using inter-residue average distance statistics. The authors' kind of predicted contact map (Average Distance Map, ADM) pinpoints regions of possible folding nuclei for proteins. The amino acid sequence of a given protein suffice for construction of the ADM of a protein. An inter-residue effective potential is defined from the average distance statistics and is utilized to predict folding nucleation sites of a protein sequence. The authors review examples of the applications of these methods to the analyses of the protein folding of members of several protein families. The general possibility of predicting protein folding mechanisms from only sequences is also discussed.

Chapter 3 - The problem of protein self-organization is one of the most important problems of molecular biology nowadays. Despite the recent success in the understanding of general principles of protein folding, details of this process are yet to be elucidated. Moreover, the prediction of protein folding rates has its own practical value due to the fact that aggregation directly depends on the rate of protein folding. The time of folding and transition state ensembles for 67 proteins with known experimental data at the point of thermodynamic equilibrium between unfolded and native state have been calculated using a Monte Carlo method and Dynamic programming one where each residue is considered to be either folded as in the native state or completely disordered. The times of folding for 67 proteins which reach the native state within a limit of 10^8 Monte Carlo steps are in a good correlation with experimentally measured folding time at mid-transition point (the correlation coefficient is -0.82). A lower correlation was obtained if to use Dynamic programming approach (the correlation coefficient is -0.72). The capillarity model allows us to predict the folding rate at the same level of correlation as by Monte-Carlo simulations. The calculated model entropy capacity (conformational entropy per residue divided by the average contact energy per residue) for 67 proteins correlates by about 78% with the experimentally measured folding rate at the mid-transition point. Theoretical

consideration of a capillarity model for the process of protein folding demonstrates that the difference in the folding rate for proteins sharing more ball-like and less ball-like folds is the result of differences in the conformational entropy due to a larger surface of the boundary between folded and unfolded phases in the transition state for proteins with more ball-like fold.

Chapter 4 - Despite much progress in molecular biology, biophysics, and biochemistry, the mechanisms by which unfolded polypeptide chains gain their native conformation are not yet fully understood. Likewise, the network of forces that stabilize folded proteins and prevent them from unfolding elude a thorough accounting so far. Such an accounting is, however, necessary to engineer protein stability, which is often required for enzymes in technical applications.

The folding process of proteins starts with a hydrophobic collapse or local events at so-called 'nuclei' or 'initiation sites' and proceeds stepwise to the natively folded protein. The occurrence of rate-limiting steps often leads to the population of folding intermediates. In contrast to the successive protein folding process, protein *un*folding is highly cooperative. Due to the principle of microscopic reversibility, unfolding must be a progressive process as well, which is postulated to start at a confined region—'unfolding region', 'critical region' or similar—of the tertiary structure of the native protein molecule.

The rate constants of the folding (k_f) and unfolding reaction (k_U), both of which determine the respective free energy of activation $\Delta G_f^{\ddagger}$ and $\Delta G_U^{\ddagger}$, provide information on the contribution of these reactions to the thermodynamic stability (Gibbs free energy $\Delta G = \Delta G_U^{\ddagger} - \Delta G_f^{\ddagger}$). The comparison of proteins with their chemically modified or, preferably, genetically engineered variants yields information on the involvement of specific residues in the folding or unfolding reaction.

RibonucleaseA (RNaseA) has been a model protein for half a century and is one of the most thoroughly studied examples concerning the protein folding problem. Whereas the folding reaction is complex due to proline isomerization reactions, the unfolding reaction can be described by a single-exponential reaction. Modifications by, *e.g.*, genetic engineering have revealed the importance of confined regions for the unfolding and refolding reaction, respectively.

Starting from basic principles of thermodynamics and kinetics, the effect of protein modifications on the stability and tools for protein stabilization are presented with special focus on work on RNase A.

Chapter 5 - Yeast is an important platform for the production of therapeutic proteins. It has principally been used for the production of relatively simple proteins that lack *N*-linked glycosylation. However, with recent developments in technologies focused on glycoengineering yeast to produce humanized *N*-glycans, its application to the field of human therapeutics becomes greatly potentiated. A constant challenge in the biotechnology industry is the production of recombinant proteins at high levels. Like other protein expression platforms, productivities in yeast depend on the host strain, its process of cultivation and the type of protein being expressed. In addition, it has been demonstrated that yeast lack some molecular chaperones found in higher eukaryotic cells. Protein folding in the endoplasmic reticulum is a key process that can influence the fate of a protein, directing it either to secretion or ER-associated degradation. Some proteins require specific molecular chaperones for their protein maturation and others require general protein folding chaperones. In this review, yeast folding machinery as well as attempts made to improve protein folding will be discussed.

Chapter 6 - Treatment options for cancer have increased in the last years due to the ongoing research and development of new therapeutics. A focus shift from formerly undirected cytostatic agents to more targeted therapies, mainly antibodies and protein kinase inhibitors has taken place. New agents directed against receptors which control proliferation, migration and angiogenesis of tumor cells have been developed. Mainly specific antibodies directed against the extracellular ligand binding domain of various receptor tyrosine kinases, and protein kinase inhibitors targeting the intracellular protein kinase domain of receptors or cytoplasmic signalling proteins have been increasingly developed in the past years. In this chapter the authors will address the current treatment options with protein kinase inhibitors exemplarily for hepatocellular carcinoma and renal cell cancer. The authors will further give an insight into the actual research perspectives and the conception of incorporating new drugs into clinical trials as single agent treatment or combined with "classical" chemotherapy. The first part elucidates the molecular pathways of carcinogenesis and depicts the mode of action of these new agents in the treatment of hepatocellular carcinoma, the second part illustrates the pathogenesis of renal cell cancer and the resulting therapeutic options by targeting these pathways.

Chapter 7 - The most abundant molecules in nature other than water are proteins, which are involved to stimulate or control virtually every chemical process. Majority of proteins must be converted into tightly folded compact

structures in order to function and their ability to fold to a unique structure generate enormous selectivity and diversity in their functions. Conformational changes in proteins result in misfolding, aggregation, and intra- or extraneuronal accumulation of amyloid fibrils in many neurodegenerative disorders. Proteins appear to fold by diverse pathways, the rates of folding are largely determined by the topology of the native structure. In general, secondary structure is inherently unstable and its stability is enhanced by tertiary interactions. The most conspicuous feature of many neurodegenerative disorders, including Alzheimer's, Parkinson's, and Huntington's disease, is the occurrence of protein aggregates in ordered fibrillar structures known as amyloid found inside and outside of brain cells. The appearance of aggregates in diseased brains implies an underlying incapacity in the cellular machinery of molecular chaperones that normally functions to prevent the accumulation of misfolded proteins. Molecular chaperones provide a first line of defense against misfolded, aggregation-prone proteins, and are among the most potent suppressors of neurodegeneration known for animal models of human disease. A detailed understanding of the molecular basis of protection by chaperones against neurodegeneration might lead to the development of therapies for neurodegenerative disorders that are associated with protein misfolding and aggregation.

In this review, the advances on the protein folding in neurodegenrative disorder and unfolding the role of molecular chaperones in neuroprotection has been discussed.

Versions of these chapters were also published in *Proteomics Research Journal,* Volume 2, Numbers 1/2, published by Nova Science Publishers, Inc. They were submitted for appropriate modifications in an effort to encourage wider dissemination of research.