Abnormal Proteins, Protein-Diseases

PÉTER TOMPA, PÉTER CSERMELY

Introduction

This chapter is focused on protein-diseases, exemplified by the well-known Alzheimer's disease, Parkinson's disease, Creutzfeldt-Jakob disease, and other serious pathological disorders. The chapter first summarizes the most important features and disorders of protein structure, and later describes the genetic and biochemical background, pathophysiology and potential cures of the most prevalent protein diseases.

Misfolded Proteins

Major Features of Protein Structure

Proteins can be described as long chains of their constituent 20 amino acids. A protein of typical length contains 200 to 500 amino acids. These protein chains have to form a unique 3-dimensional pattern. The various pathways leading to this final, native conformation are called together as the protein folding process. Amino acid side chains have special physico-chemical features which are essential for the enzyme activity, binding or complex formation of the hosting protein with other molecules, such as small ligands, proteins, RNA or DNA. The surface of proteins (which includes their binding sites) is defined by the network of interactions between amino acids. The energy which is liberated when the individual interactions between the amino acids are formed helps to overcome the difficulty of making an organized, ordered structure (the loss of entropy) during the folding process. The protein structure is stabilized by hydrophobic, van der Waals, hydrogen-bond and ionic interactions. Protein structure is described by four hierarchical levels. The *primary structure* denotes the amino acid sequence, while the *secondary structure* summarizes the local structural elements. According to the traditional view, α-helices and β-sheets are the most prevalent of the repetitive secondary structures. However, recent data suggest a rather large abundance of polyprolin II-helix as well. The non-repetitive secondary structure elements contain β-turns and coiled (unstructured) conformations. The *tertiary structure* means the complete three-dimensional structure of a polypeptide chain. The *quaternary structure* refers to the complex structure of a protein containing more than one individual polypeptides or other larger constituents (such as heme, or polysaccharide) besides the polypeptide.

Folding of globular proteins starts with hydrophobic collapse in with the internal, hydrophobic core of the protein is formed in a rapid process. This initial phase is followed by a slow re-annealing of the internal structure of folding intermediates, often called molten globules. In this phase, the proteins still have hydrophobic surfaces, which make them very vulnerable to aggregation. At the closing phase, the individual, high-energy bonds (such as ionic interactions or disulfide-bridges, if any) and high-energy active sites are formed. The folding process is often fast (in the range of milliseconds) in case of small proteins (below 30 kDa), but the final rearrangement of the hydrophobic core may require weeks without any extra help.

The evolution of a stable protein structure was not an easy and straightforward process. During this lengthy procedure, special amino-acid sequences had to be selected which are able to fold into unique, native structures. In principle, an amino acid sequence encodes a specific 3-dimensional structure of a protein. According to this view, a protein chain does not require any help to develop the final, native structure. This is indeed true for the dilute solutions of proteins smaller than 30 kDa. However, the folding of larger proteins, especially at the large protein concentration of 300 to 400 mg/ml inside the cell, is not possible without help. Under such conditions (called molecular crowding), the chance of protein aggregation is extremely increased since the hydrophobic surfaces of the molten globule intermediates have a very high chance of getting close to each other. Aggregation is a great danger since in many cases it irreversibly incapacitates the participating proteins.

Cells have developed two methods to prevent, repair or eliminate protein aggregation. Protein-aggregation is prevented by *molecular chaperones* which transiently bind to native, freshly synthesized proteins, or to misfolded, damaged proteins in the cell. During this process chaperones cover hydrophobic surfaces and prevent their binding to each other, which otherwise leads to aggregation. Most types of cellular stress induce an elevated synthesis of several molecular chaperones; therefore, many major chaperones are called heat shock proteins (e.g. Hsp60, Hsp70 or Hsp90, where the numbers refer to the molecular weight of the protein class in kDa). Heat shock pro-