Control of Protein Aggregation and Fibril Formation by Small Chemicals

Masahiro Takagi
Japan Advanced Institute of Science and Technology

A protein is a linear polymer of amino acids linked together by peptide bonds. Various, mostly non-covalent, interactions between amino acids in the linear sequence stabilize a specific folded three-dimensional structure for each protein. Proteins fold into a unique native structure even *in vitro* (Anfinsen's dogma). However, they tend to form undesirable and uncontrollable aggregates during unfolding and refolding processes even in the natural environment in living cells. Protein aggregation is a major problem in large-scale production of recombinant proteins, as well as in living cells, which lead to fatal diseases. Amyloidosis, including Alzheimer's and prion diseases, is a heterogeneous group of disorders characterized by the accumulation of extracellular deposition of abnormal protein fibrils (amyloid fibrils). Proteins known to form amyloid fibrils *in vivo* have no homology in their sequences and the three-dimensional structures in their native state forms, but amyloid fibrils have a common core structure. Recently, several non-disease-related fibrils were reported, such as lyozyme at high concentrations of ethanol and myoglobin at high temperatures. Interestingly, fibrils of the non-disease-related proteins are quite similar to those of the disease-related proteins.

In this presentation, we will focus on the following three topics.

(1) How protein native structure is maintained. (Contribution of ion-pairs on thermostability of the protein),
(2) How aggregation formation can be controlled. (Prevention of protein aggregation by small molecules), and
(3) Does all protein form amyloid fibrils? (Formation of amyloid fibrils from non-disease-related proteins.)

Keywords:

Protein: Polymers of amino acids linked by peptide bonds. The specific sequence of amino acids determines the shape and function of the protein.

Folding: Before proteins can carry out their biochemical function, they remarkably assemble themselves, or "fold." The process of protein folding, while critical and fundamental to virtually all of biology, remains a mystery. Moreover, perhaps not surprisingly, when proteins do not fold correctly (i.e. "misfold"), there can be serious effects, including many well known diseases, such as Alzheimer's, Mad Cow (BSE), and Parkinson's disease.

Anfinsen's dogma: The amino-acid sequence of a protein is encoded as genetic information in a nucleotide sequence of DNA. This genetic information is transcribed into messenger RNA and

gets bound to ribosomes where the protein is synthesized. Hence, the tertiary structures of proteins seem to be dependent on various environmental factors within the cell. To the contrary, it was experimentally shown that the three-dimensional structure of a protein is determined solely by amino-acid sequence information. This is called Anfinsen's dogma.

Amyloid: A general term for a variety of different proteins that accumulate as extracellular fibrils of 7-10 nm and have common structural features, including a beta-pleated sheet conformation and the ability to bind such dyes as Congo red and thioflavine. Of particular interest in neuropsychiatry is the beta-amyloid protein, which is the major component of the characteristic senile plaques of Alzheimer's disease and the amyloid precursor protein (APP).