PREFACE

An important but relatively unknown protein modification is PHBylation. In this book, the authors review the physical properties and molecular characteristics of PHB, and discuss the influences of PHBylation on protein folding. In the second chapter, protein-folding mechanisms, namely the framework model and nucleation-condensation mechanism are discussed. The final chapter examines the fundamental physical law governing protein folding - the Thermodynamic Principle.

An important but relatively unknown protein modification is PHBylation, i.e., covalent attachment of short polymers (< 20 residues) of (R)-3-hydroxybutyrate (cPHB). PHB are water-insoluble and lipid soluble polymers in which hydrophobic methyl groups alternate with hydrophilic ester groups along a flexible linear chain. A cPHB molecule forms a single covalent bond at its ester end to a hydroxy amino acid of the protein (usually serine), and the remainder of the polymer may then form multiple noncovalent hydrophobic, hydrogen and coordinate bonds to neighboring amino acid residues. PHBylated proteins are present in all biological cells – prokaryotic and eukaryotic. They are primarily found in membranes and organelles, wherein they are frequently engaged in the binding and transport of ions or polyanions. Since the flexibility of the PHB backbone decreases markedly below physiological temperatures, the folding and the function(s) of cPHB-modified proteins are highly temperature-sensitive. In Chapter 1, the author review the physical properties and molecular characteristics of PHB, and discuss the influences of PHBylation on protein folding.

Analysis of Φ values provides information on structures of transition states along a folding pathway of a protein via measurement of effects of a mutation on folding kinetics. In Chapter 2, the author discuss the Φ values

calculated for 27 proteins (54–128 amino acid residues long) by means of a simple statistical-mechanical model of protein folding (developed by us); the author compare these data with experimentally observed Φ values. In the model, three-dimensional structure of a protein was taken into consideration via inter-residue contact information of the native structure obtained from the Protein Data Bank. A free-energy profile of the wild-type protein was calculated by assigning a single value to interaction energies of the inter-residue contacts (regardless of the amino acid types) as an approximation. The same procedure was applied to a protein with a single point mutation but with slightly greater or smaller values assigned to the interaction energies between the mutated and other (unmutated) residues. The folding and unfolding rates were individually estimated from the free-energy profiles of the wild-type protein and mutant protein. The Φ value was then calculated for every residue from the differences in the folding or unfolding rates between the wild-type and the mutant. Despite the simplicity of the model, corresponding experimental Φ values were fairly well predicted for most proteins. The successful cases point to the importance of topology of the three-dimensional structure and support validity of assumptions of the model. Conversely, the cases of failure are suggestive of the need for interaction details such as amino acid types and interaction modes during the calculation and revealed limitations of the model. Protein-folding mechanisms, namely the framework model and nucleation-condensation mechanism, are also discussed in terms of the advantages and disadvantages of the proposed model.

The fundamental physical law governing protein folding is the Thermodynamic Principle: the peptide chain of a protein together with the protein's physiological environment determines the protein's native structure; and the native structure has the minimum Gibbs free energy among all conformations of the protein.

Beginning with a thought experiment indicating that protein folding should be considered in the single molecule level and there must be a general folding force that pushes a protein molecule fold along a folding path, Chapter 3 recalls the works leading to the thermodynamics principle. Arguments to support a single molecule understanding of the thermodynamic principle is given. After preparations including the topology of the conformational space $\mathcal{X}_{\mathfrak{A}}$ of a molecule $\mathfrak{A}$ and necessary thermodynamic and quantum statistics tools, a Gibbs free energy function $G(\mathbf{X};\mathcal{E}_N;\mathfrak{A})$ is derived for a globular protein $\mathfrak{A}$, where $\mathbf{X}\in\mathcal{X}_{\mathfrak{A}}$, $\mathcal{E}_N$ is the physiological environment of $\mathfrak{A}$. The

derivation relies on making tailor made, nano sized tiny thermodynamic systems $\mathcal{S}_{X,\mathcal{E}}$, that contains only one protein molecule in the conformation $\mathbf{X}$ and a layer of material particles appear in the environment $\mathcal{E}$.

Many natural phenomena in protein folding and protein structures can be explained with this Gibbs free energy function and its dependence on environment. Among them are a holistic view of protein structure which explains phenomenon such as the chameleon sequences; why just reducing hydrophobic surface area can produce secondary structures and hydrogen bonds; structure stability; allosteric proteins, docking, enzyme-substrate combination; and denaturation and refolding. The new unified explanations are at least as good as current, *ad hoc* ones, in some cases are much better. Testable predictions are given, first among them is the prediction of protein native structure from the only knowledge of the protein's amino acid sequence. It just becomes a pure mathematical problem of minimisation. Another testable prediction is that changed local environment causes amyloid and prion related diseases.