
Gibbs Free Energy Formula for Protein Folding

Yi Fang

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/52410>

1. Introduction

Proteins are life's working horses and nature's robots. They participate in every life process. They form supporting structures of cell, fibre, tissue, and organs; they are catalysts, speed up various life critical chemical reactions; they transfer signals so that we can see, hear, and smell; they protect us against intruders such as bacteria and virus; they regulate life cycles to keep that everything is in order; etc., just mention only a few of their functions.

The first thing drawing our attention of proteins are their size. Proteins are macromolecules, that is, large molecules. Non-organic molecules usually are small, consisting of from a couple of atoms to a couple of dozen atoms. A small protein will have thousands of atoms, large ones have over ten thousand atoms. With their huge number of atoms, one can imagine that how complicated should be of a protein molecule. Fortunately, there are some regularities in these huge molecules, i.e., proteins are polymers building up by monomers or smaller building blocks. The monomers of proteins are amino acids, life employs 20 different amino acids to form proteins. In cell, a series of amino acids joined one by one into amino acids sequences. The order and length of this amino acid sequence is translated from DNA sequences by the universal genetic code. The bond joining one amino acid to the next one in sequence is peptide bond (a covalent bond) with quite regular specific geometric pattern. Thus amino acids sequences are also called peptide chains. But the easy translation and geometric regularities stop here. The peptide chain has everything required to a molecule, all covalent bonds are correctly formed. But to perform a protein's biological function, the peptide chain has to form a specific shape, called the protein's **native structure**. Only in this native structure a protein performs its biological function. Proteins fall to wrong shapes not only will not perform its function, but also will cause disasters. Many disease are known to be caused by some proteins taking wrong structure.

How the peptide chain take its native structure? Is there another genetic code to guide the process of taking to the native structure? In fact, at this stage, life's most remarkable drama

takes stage. Once synthesized, the peptide chain of a protein spontaneously (some need the help of other proteins and molecules) fold to its native structure. This process is called **protein folding**. At this stage, everything is governed by simple but fundamental physical laws.

The **protein folding problem** then can be roughly divided into three aspects: 1. folding process: such that how fast a peptide chain folds, what are the intermediate structures between the initial shape and the native structure. 2. the mechanics of the folding, such as what is the deriving force. 3. the most direct application to biological study is the prediction of the native structure of a protein from its peptide chain. All three parts of the protein folding problem can have a unified treatment: writing down the Gibbs free energy formula $G(\mathbf{X})$ for any conformation $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_M) \in \mathbb{R}^{3M}$ of protein, where $\mathbf{x}_i \in \mathbb{R}^3$ is the atom \mathbf{a}_i 's atomic center.

The fundamental law for protein folding is the **Thermodynamic Principle**: the amino acid sequence of a protein determines its native structure and the native structure of the protein has the minimum Gibbs free energy among all possible conformations as stated in Anfinsen (1973). Let \mathbf{X} be a conformation of a protein, is there a natural Gibbs free energy function $G(\mathbf{X})$? The answer must be positive, as G. N. Lewis said in 1933: "There can be no doubt but that in quantum mechanics one has the complete solution to the problems of chemistry." (quoted from Bader (1990), page 130.) Protein folding is a problem in biochemistry, why such a formula $G(\mathbf{X})$ has not been found and what is the formula? This chapter is trying to give the answers.

First, the Gibbs free energy formula is given, it has two versions, the chemical balance version (1) and the geometric version (2).

1.1. The formula

Atoms in a protein are classified into classes H_i , $1 \leq i \leq H$, according to their levels of hydrophobicity. The formula has two versions, the chemical balance version is:

$$G(\mathbf{X}) = \mu_e N_e(\mathbf{X}) + \sum_{i=1}^H \mu_i N_i(\mathbf{X}), \quad (1)$$

where $N_e(\mathbf{X})$ is the mean number of electrons in the space included by the first hydration shell of \mathbf{X} , μ_e is its chemical potential. $N_i(\mathbf{X})$ is the mean number of water molecules in the first hydration layer that directly contact to the atoms in H_i , μ_i is the chemical potential.

Let $M_{\mathbf{X}}$ (see FIGURE 3) be the molecular surface for the conformation \mathbf{X} , defining $M_{\mathbf{X}_i} \subset M_{\mathbf{X}}$ as the set of points in $M_{\mathbf{X}}$ that are closer to atoms in H_i than to any atoms in H_j , $j \neq i$. Then the geometric version of $G(\mathbf{X})$ is:

$$G(\mathbf{X}) = v_e \mu_e V(\Omega_{\mathbf{X}}) + d_w v_e \mu_e A(M_{\mathbf{X}}) + \sum_{i=1}^H v_i \mu_i A(M_{\mathbf{X}_i}), \quad v_e, v_i > 0, \quad (2)$$

where $V(\Omega_{\mathbf{X}})$ is the volume of the domain $\Omega_{\mathbf{X}}$ enclosed by $M_{\mathbf{X}}$, d_w the diameter of a water molecule, and $A(M_{\mathbf{X}})$ and $A(M_{\mathbf{X}_i})$ the areas of $M_{\mathbf{X}}$ and $M_{\mathbf{X}_i}$, $v_e [V(\Omega_{\mathbf{X}}) + d_w A(M_{\mathbf{X}})] = N_e$, $v_i A(M_{\mathbf{X}_i}) = N_i(\mathbf{X})$, $1 \leq i \leq H$. The v_e and v_i are independent of \mathbf{X} , they are the average numbers of particles per unit volume and area.

Before the actual derivation is given, some basic facts should be stated, such as hydrophobicity, protein structures, and the environment in which the protein folds. Brief description of the methods in the experimental measurements and theoretical derivation of the Gibbs free energy of the protein folding is introduced to give the motivation and idea of the derivation. By making critics on the previous derivation, the necessary concepts would be clarified, what are important in the derivation would be identified, and would set the thermodynamic system that most fit the reality currently known about the protein folding process. Then both classical and quantum statistical derivations were given, the only difference is that in the classical statistically derived formula, the volume and the whole surface area terms in formula (2) are missing. Thus it is that only quantum statistical method gives us the volume and whole surface terms in formula (2). After the derivations, some remarks are made. A direct application of the Gibbs free energy formula (2) is the *ab initio* prediction of proteins' natives structures. Gradient formulas of $G(\mathbf{X})$ are given to be able to apply the Newton's fastest descending method. Finally, it should be emphasized that the gradient $\nabla G(\mathbf{X})$ not only can be used to predict the native structure, it is actually the force that forces the proteins to fold as stated in Ben-Naim (2012). In Appendix, integrated gradient formulas of $G(\mathbf{X})$ on the molecular surface are given.

2. Proteins

2.1. Amino acids

There are 20 different amino acids that appear in natural proteins. All amino acids have a common part, or the **back bone** consisting of 9 atoms in FIGURE 1 (except the R).

NH_2 is the amino group and COOH is the carboxyl group of the back bone. Single amino acid is in polar state, so the amino group gains one more hydrogen from the carboxyl group, or perhaps the amino group losses one electron to the carboxyl group. Geometrically it is irrelevant since after forming peptide bonds the amino group will loss one H to become NH and the carboxyl group will loss one OH to become CO. Thus an amino acid in the sequence is also called a **residue**.

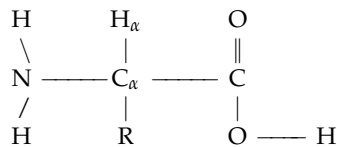


Figure 1. An generic amino acid.

The group R in FIGURE 1 is called **side chain**, it distinguishes the 20 different amino acids. A side chain can be as small as a single hydrogen atom as in Glycine, or as large as consisting of 18 atoms including two rings as in Tryptophan. 15 amino acids have side chains that contain more than 7 atoms, i.e., more atoms than that of the back bone in an amino acid sequence. Except Glycine, a C_β carbon in a side chain forms a covalent bond with the **central carbon** C_α of the back bone.

2.2. Hydrogen bonds

A hydrogen bond is the attractive interaction of a hydrogen atom with an electronegative atom (the **accepter**), like nitrogen, oxygen or fluorine (thus the name “hydrogen bond”, which must not be confused with a covalent bond to hydrogen). The hydrogen must be covalently bonded to another electronegative atom (forming a **donor group**) to create the hydrogen bond. These bonds can occur between molecules (intermolecular), or within different parts of a single molecule (intramolecular). The hydrogen bond is stronger than the van der Waals interaction, but weaker than covalent or ionic bond. Hydrogen bond occurs in both inorganic molecules such as water and organic molecules such as DNA, RNA, and proteins.

Some amino acids’ side chains contain hydrogen bond donors or acceptors that can form hydrogen bond with either other side chains in the same protein (intramolecular hydrogen bond) or with surrounding water molecules (inter-molecular hydrogen bond). Those amino acids whose side chains do not contain either donors or acceptors of hydrogen bond are classified as hydrophobic.

2.3. Hydrophobicity levels

Every atom in a protein belongs to a moiety or atom group, according to the moiety’s level of ability to form hydrogen bond, the atom is assigned a hydrophobicity level. All the hydrophobicity scales are tested or theorized in some aspects of individual amino acid, either as a independent molecule or as a residue in a protein, in various environments such as solvent, PH value, temperature, pressure, etc. That is just like taking a snap shot of an object with complicated shape. All snap shots are different if taking from different angles of view. Therefore, there are many different classifications of hydrophobicity, for example, in Eisenberg and McLachlan (1986) there are five classes, C, O/N, O⁻, N⁺, S. Let a protein have M atoms $\{\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_M\}$. One can assume that there are H hydrophobic classes, such that $\{\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_M\} = \cup_{i=1}^H H_i$.

2.4. Protein structures

Let a molecule have M atoms, listed as $(\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_M)$. A presentation of a structure \mathbf{X} of this molecule is a series atomic centers (nuclear centers) of the atoms \mathbf{a}_i , $\mathbf{x}_i \in \mathbb{R}^3$. Hence it can be written as a point in \mathbb{R}^{3M} , $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_M)$. The space \mathbb{R}^{3M} then is called the *control space*. The real shape of the structure \mathbf{X} is realized in \mathbb{R}^3 , called the *behavior space* as defined in Bader (1990), it is a bunch of overlapping balls (spheres), $P_{\mathbf{X}} = \cup_{i=1}^M B(\mathbf{x}_i, r_i)$, where r_i is the van der Waals radius of the atom \mathbf{a}_i and $B(\mathbf{x}, r)$ is the closed ball $\{\mathbf{y} : |\mathbf{y} - \mathbf{x}| \leq r\} \subset \mathbb{R}^3$, of center \mathbf{x} and radius r .

Protein native structures are complicated. Unlike the famous double-helix structure of DNA structure, the only general pattern for protein structure is no pattern at all. To study the native structures of proteins people divide the structures in different levels and make structure classifications.

The amino acid sequence of a protein is called its **primary structure**. Regular patterns of local (along the sequences) structures such as helix, strand, and turn are called the **secondary structure** which contain many intramolecular hydrogen bonds in regular patterns. The global

assembly of these secondary structures, connected by turns and irregular loops, is called the **tertiary structure**. For proteins having multiple amino acid sequences or structurally associated with other molecules there are also **quaternary structures**, see Branden and Tooze (1999) and Finkelstein and Ptitsyn (2002).

The secondary structures are local structures, they are usually in helix, strand, and turn. A common feature of them is that they have regular geometric arrangement of their main chain atoms, such that there are good opportunities to form hydrogen bonds between different residues. Several strands may form sheet, stabilized by regular pattern of hydrogen bonds. Turns and loops are necessary for the extended long chain to transfer to a sphere like shape. Turns are short, 3 or 4 residues long. Loops involves many residues, but without any regular pattern of hydrogen bonds. Loops often form the working place of the protein, therefore appear on the out surface of the native structure.

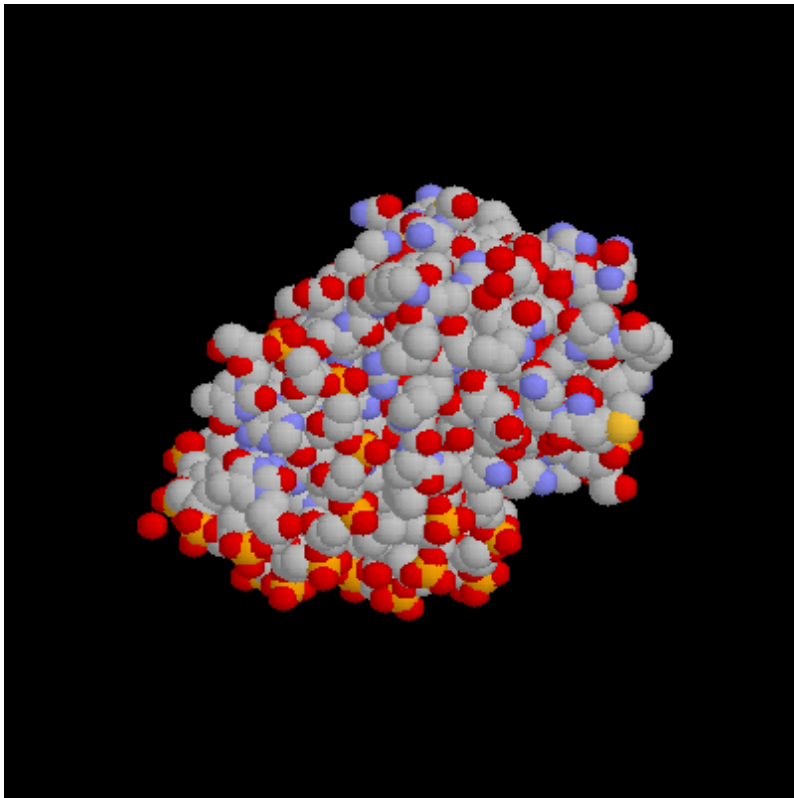


Figure 2. P_X is a bunch of overlapping balls, called the space-filling model, or CPK model.

3. Some functions in thermodynamics

A thermodynamic system consists of particles in a region $\Gamma \subset \mathbb{R}^3$ and a bath or environment surrounding it. A wall, usually the boundary $\partial\Gamma$ separating the system with its surrounding.

If no energy and matter can be exchanged through the wall, the system is an **isolated system**. If only energy can be exchanged, the system is a **closed system**. If both energy and matter can be exchanged with the surrounding, the system is an **open system**.

For an open system Γ of variable particles contacting with surrounding thermal and particle bath, let U, T, S, P, V, μ and N be the inner energy, temperature, entropy, pressure, volume, chemical potential, and the number of particles of the system Γ respectively, then

$$dU = TdS - PdV + \mu dN, \quad (3)$$

By Legendre transformations various extensive quantities can be derived,

$$F = U - TS, \quad G = U - TS + PV, \quad \phi = F - \mu N = U - TS - \mu N \quad (4)$$

where $F, G,$ and ϕ are Helmholtz, Gibbs free energies, and thermodynamic potential respectively. Then

$$dF = -SdT - PdV + \mu dN, \quad dG = -SdT + VdP + \mu dN, \quad d\phi = -SdT - PdV - Nd\mu. \quad (5)$$

Which shows that $U = U(S, V, N), F = F(T, V, N), G = G(T, P, N), \phi = \phi(T, V, \mu)$. All **extensive** quantities satisfy a linear homogeneous relation, i.e., consider a scaling transformation which enlarges the actual amount of matter by a factor λ , then all extensive quantities are multiplied by a factor λ . U, S, V, N, F, G, ϕ are extensive, while T, P, μ are **intensive**. Thus

$$\lambda U = U(\lambda S, \lambda V, \lambda N), \quad \lambda F = F(T, \lambda V, \lambda N), \quad \lambda G = G(T, P, \lambda N), \quad \lambda \phi(T, V, \mu) = \phi(T, \lambda V, \mu). \quad (6)$$

From equations in (5) $(\frac{\partial \phi}{\partial V})_{T, \mu} = -P$. By equations in (6)

$$\phi = \frac{d(\lambda \phi)}{d\lambda} = V \left(\frac{\partial \phi}{\partial V} \right)_{T, \mu} = -PV \quad (7)$$

and

$$\phi(T, V, \mu) = -PV. \quad (8)$$

Equation (8) is true for any open thermodynamics system.

4. Statistical mechanics

Thermodynamics is a phenomenological theory of macroscopic phenomena that neglects the individual properties of particles in a system. Statistical mechanics is the bridge between the macroscopic and microscopic behavior. In statistical mechanics, the particles in a system obey either classical or quantum dynamic laws, and the macroscopic quantities are statistical averages of the corresponding microscopic quantities. If the particles obey classical dynamical law, it is the classical statistical mechanics. If the particles obey quantum dynamical law, it is the quantum statistical mechanics. But the averaging to get macroscopic quantities from microscopic ones are in the same principle and formality.

Protein folding studies the structure of the protein molecule, what is the native structure and why and how the protein folds to it. All these aspects are specific properties of a particle, the protein molecule. To get the Gibbs free energy formula $G(\mathbf{X})$ for each conformation \mathbf{X} , statistical mechanics is needed with careful specification of the thermodynamic system.

4.1. The canonic ensemble

Statistical mechanics uses ensembles of all microscopic states under the same macroscopic character, for example, all microscopic states corresponding to the same energy E . The probability of this ensemble then is proportional to

$$p_E \propto \exp(-\beta E), \quad (9)$$

where $\beta = 1/kT$, k the Boltzmann constant and T the temperature. If there are only a series energy levels E_1, E_2, \dots , then the probability distribution for canonic ensemble is

$$p_i = \frac{\exp(-\beta E_i)}{\sum_{n=1}^{\infty} \exp(-\beta E_n)}. \quad (10)$$

Various of thermodynamic quantities, such as the inner energy of the system, can be put as the means:

$$U = \langle E_i \rangle = \frac{\sum_{i=1}^{\infty} E_i \exp(-\beta E_i)}{\sum_{n=1}^{\infty} \exp(-\beta E_n)}. \quad (11)$$

If only the Hamiltonian $H(\mathbf{q}, \mathbf{p})$ is known, where $\mathbf{q} = (\mathbf{q}_1, \dots, \mathbf{q}_i, \dots, \mathbf{q}_N) \in \Gamma^N$ is the position of the N particles in the thermodynamic system $\Gamma \subset \mathbb{R}^3$ under study, and $\mathbf{p} = (\mathbf{p}_1, \dots, \mathbf{p}_i, \dots, \mathbf{p}_N)$ momentums of these particles, the *canonical phase-space density* of the system then is

$$p_c(\mathbf{q}, \mathbf{p}) = \frac{\exp[-\beta H(\mathbf{q}, \mathbf{p})]}{N! h^{3N} \int_{\Gamma^N} d\mathbf{q}^N \int_{\mathbb{R}^{3N}} \exp[-\beta H(\mathbf{q}, \mathbf{p})] d\mathbf{p}^N} = \frac{\exp[-\beta H(\mathbf{q}, \mathbf{p})]}{\mathcal{Z}(T, V, N)}. \quad (12)$$

where $N!$ is the Gibbs corrector because that the particles in the system is indistinguishable. $\mathcal{Z}(T, V, N)$ is called the *canonic partition function*, it depends on the system's temperature T , volume V , and particle number N . Note that under the assumption of the canonic ensemble, they are all fixed for the fixed thermodynamical system Γ . Especially, $V = V(\Gamma) = \int_{\Gamma} d\mathbf{q}$ implicitly set that $\Gamma \subset \mathbb{R}^3$ has a volume.

Then the entropy S is

$$\begin{aligned} S &= S(\Gamma) = \langle -k \ln p_c \rangle = \frac{k}{N! h^{3N}} \int_{\Gamma^N} d\mathbf{q}^N \int_{\mathbb{R}^{3N}} [\beta H(\mathbf{q}, \mathbf{p}) + \ln \mathcal{Z}(T, V, N)] p_c(\mathbf{q}, \mathbf{p}) d\mathbf{p}^N \\ &= \frac{1}{T} [\langle H \rangle + kT \ln \mathcal{Z}(T, V, N)]. \end{aligned} \quad (13)$$

From which the Helmholtz free energy $F = F(\Gamma)$ and the Gibbs free energy $G = G(\Gamma)$ are obtained,

$$F = U - TS = -kT \ln \mathcal{Z}(T, V, N), \quad G = PV + F = PV - kT \ln \mathcal{Z}(T, V, N). \quad (14)$$

Therefore, to obtain the Gibbs free energy one has to really calculate $\ln \mathcal{Z}(T, V, N)$, a task that often cannot be done.

4.2. The grand canonic ensemble

The grand canonic ensemble or macroscopic ensemble deals with an open thermodynamic system Γ , i.e., not only energy can be exchanged, matter particles can also be exchanged between Γ and environment. Therefore, the particle number N in Γ is variable.

In classical mechanics, suppose that the phase space is $(\mathbf{q}, \mathbf{p}) \in \Gamma^N \times \mathbb{R}^{3N}$. Let H be the Hamiltonian, the grand canonic phase-space density is

$$p_{gc}(\mathbf{q}, \mathbf{p}, N) = \frac{\exp[-\beta(H - \mu N)]}{\sum_{N=0}^{\infty} \frac{1}{N! h^{3N}} \int_{\Gamma^N} d\mathbf{q}^N \int_{\mathbb{R}^{3N}} \exp[-\beta(H(\mathbf{q}, \mathbf{p}) - \mu N)] d\mathbf{p}^N} = \frac{\exp[-\beta(H - \mu N)]}{\mathcal{Z}(T, V, \mu)}, \quad (15)$$

where $V = V(\Gamma)$ is the volume of the system. By definition the entropy is

$$\begin{aligned} S(\Gamma) &= \langle -k \ln p_{gc} \rangle = k \sum_{N=0}^{\infty} \int_{\Gamma^N} d\mathbf{q}^N \int_{\mathbb{R}^{3N}} \{\beta[H(\mathbf{q}, \mathbf{p}) - \mu N] + \ln \mathcal{Z}\} p_{gc}(\mathbf{q}, \mathbf{p}) d\mathbf{p}^N \\ &= \frac{1}{T} [\langle H \rangle - \mu \langle N \rangle + kT \ln \mathcal{Z}]. \end{aligned} \quad (16)$$

Here $\langle H \rangle = U$ is the inner energy of the system Γ , $\langle N \rangle = N(\Gamma)$ is the mean number of particles in Γ . More importantly, the function $-kT \ln \mathcal{Z}(T, V, \mu)$ is nothing but the grand canonic potential ϕ , from equation (8) it is just $-PV$. Thus

$$G = U + PV - TS = \mu \langle N \rangle. \quad (17)$$

5. Experimental measuring and theoretical derivation of the Gibbs free energy of protein folding

The newly synthesized peptide chain of a protein automatically folds to its native structure in the physiological environment. Change of environment will make a protein denatured, i.e., the protein no longer performs its biological function. The facts that denaturation does not change the protein molecule, that the only thing changed is its structure, was first theorized by Hisen Wu based on his own extensive experiments, Hisen Wu (1931). It was found that after removing the agents that caused the change of environment, some protein can automatically retake its native structure, this is called renaturation or refolding. After many experiments in denaturation and renaturation, Anfinsen summarized the Thermodynamic Principle as the fundamental law of the protein folding, Anfinsen (1973). Anfinsen's work actually show that protein refolds spontaneously after removing denaturation agents. Therefore, in the physiological or similar environment, the native structure has the minimum Gibbs free energy; and in a changed environment, the denatured structure(s) will have the smaller Gibbs free energy. The Thermodynamic Principle of protein folding then is the general thermodynamics law, if a change happens spontaneously, then the end state will have smaller Gibbs free energy than the initial state.

To apply the Thermodynamic Principle in the research of protein folding, it is necessary to know the Gibbs free energy formula $G(\mathbf{X})$ for each conformation \mathbf{X} . Until now, theoretical derivation of $G(\mathbf{X})$ is unsuccessful and rarely being tried. Most knowledge of the Gibbs free energy of protein folding comes from experiment observations.

5.1. Experimental measuring of ΔG

The basic principle of experimentally measuring ΔG , the difference in Gibbs free energy between the native and the denatured structures of a protein is as follows. For protein molecules in a solution, the criterion of the protein is in the native structure is that it performs its biological function, otherwise the protein is denatured or not in the native structure. The level of biological function indicates the degree of the denaturation. Let B be the native structure, denote its molar concentration as $[B]$. Denote A as a non-native structure of the same protein in the solution and $[A]$ its molar concentration.

Three things to be borne in mind: 1. the environment is the physiological environment or similar one such that the protein can spontaneously fold; 2: individual molecule cannot be directly measured, so the measuring is in per mole term, $R = N_A k$ instead of k should be used, where N_A is the Avogadro's number; 3: the environment in reality has constant pressure P , hence the enthalpy $H = U + PV$ can replace the inner energy U , where V is the volume of the system (it is a subset of the whole \mathbb{R}^3).

As expressed in (9), the probabilities of the protein takes the conformations A and B are

$$p_A \propto W_A \exp\left(-\frac{H_A}{RT}\right), \quad p_B \propto W_B \exp\left(-\frac{H_B}{RT}\right), \quad (18)$$

where $H_A = U_A + PV$ and $H_B = U_B + PV$ are the enthalpy per mole for A and B , W_A (W_B) is the number of ways of the enthalpy H_A (H_B) can be achieved by microscopic states. The quantities $[A]$ and $[B]$ are assumed to be measurable in experiment. Therefore their ration $K = [A]/[B]$ is also measurable. Then

$$\Delta G^o = -RT \ln K. \quad (19)$$

To see that equation (19) is true, note that the ratio K is equal to the ratio p_A/p_B and the entropies per mole are $S_A = R \ln W_A$, $S_B = R \ln W_B$, therefore

$$\begin{aligned} -RT \ln K &= -RT \ln \frac{p_A}{p_B} = -RT \left(\frac{H_B}{RT} - \frac{H_A}{RT} \right) - RT(\ln W_A - \ln W_B) \\ &= H_A - H_B - T(R \ln W_A - R \ln W_B) = H_A - H_B - T(S_A - S_B) \\ &= H_A - TS_A - (H_B - TS_B) = G_A - G_B = \Delta G^o. \end{aligned} \quad (20)$$

But in reality, the ratio K is measurable in experiment is only theoretical, since in physiological environment $K \cong 0$, i.e., almost all protein molecules take the native structure B .

There is no way to change the native structure B to A while keeping the environment unchanged. In experiments, one has to change the environment to get the protein denatured,

that is, to change its shape from the native structure B to another conformation A . Heating the solution is a simple way to change the environment, during the heating, the system absorbs an amount of heat H , the system's temperature increased from T_0 to T_1 . Then

$$G(A, T_1) - G(B, T_0) = f(H), \quad (21)$$

where $f(H)$ is a function depending on H and its value is obtained from experiment. What really needed is

$$\Delta G = G(A, T_0) - G(B, T_0). \quad (22)$$

To get ΔG , interpolation to equation (21) is used to estimate the value in T_0 . Other methods of changing environment face the same problem, i.e., interpolation has neither theoretical nor observation basis.

Equation (19) may give the reason why ΔG is used whenever referring the Gibbs free energy. For experiment, only ΔG can be got. In theoretical derivation, this rule no longer to be followed and moreover, without a base structure to compare to, the notation ΔG will look strange.

More importantly, it should be emphasized again, that the Thermodynamic Principle really says that in the physiological environment the native structure has the minimum Gibbs free energy; and in other environment, the native structure no longer has the minimum Gibbs free energy. Summarizing, it is

$$G(B, T_0) < G(A, T_0), \quad G(A, T_1) < G(B, T_1). \quad (23)$$

It should always keep in mind that before comparison, first clarify the environment.

When deriving the Gibbs free energy formula, the first thing is also to make clear what is the environment. Another reality that should be borne in mind is that during the protein folding process, the environment does not change.

Remember that after removing the denaturation agent some proteins will spontaneously refold to their native structure, this is called the refolding or renaturation. Distinguish the original protein folding problem and protein refolding problem is another important issue. Only in the refolding case, a theoretical derivation can make the environment change, for example, lower the temperature to the room temperature (around 300K). Some discussions on protein folding are really talking about refolding, because they start from changing the environment from nonphysiological to physiological.

While experiment has no way to change the native structure without disturbing the environment, theory can play a role instead. Formulas (1) and (2) give us the chance to compare ΔG , as long as the accurate chemical potentials' values are known.

5.2. Theoretical consideration of the protein folding problem

Protein folding is a highly practical field. Very few attention was paid to its theoretical part. For example, almost nobody has seriously considered the Gibbs free energy formula. Instead, all kinds of empirical models are tried in computer simulation, without any justification in fundamental principle.

One attempt to theoretically get the Gibbs free energy formula from canonic ensemble is summarized by Lazaridis and Karplus (2003), the theoretical part of it is reported below and why it is not successful will be briefly pointed out. Their notations such as $\mathbf{R} = \mathbf{X}$ as conformation, $A = F$ as the Helmholtz free energy, $Q = Z$ as the partition function, $\Lambda = h$, etc., will be kept in this section.

Treating the protein folding system as the set of all conformations plus surrounding water molecules with a phase point (\mathbf{R}, \mathbf{r}) , where \mathbf{r} are coordinates of N water molecules plus their orientations. The Hamiltonian H can be decomposed as

$$H = H_{mm} + H_{mw} + H_{ww}, \quad (24)$$

where mm means interactions inside the protein, mw between protein and water molecules, and ww water to water, all in the atomic level. Triplet interactions mmm , mmw , etc., can also be considered, but for simplicity only take the pairwise atomic interactions.

Applying the canonic ensemble, the canonic partition function is

$$Q = \frac{\int \exp(-\beta H) d\mathbf{r}^N d\mathbf{R}^M}{N! \Lambda^{3M} \Lambda^{3N}} = \frac{Z}{N! \Lambda^{3M} \Lambda^{3N}},$$

and the Helmholtz free energy is given by

$$A = -kT \ln Q = -kT \ln \left[\int \exp(-\beta H) d\mathbf{r}^N d\mathbf{R}^M \right] + kT \ln(N! \Lambda^{3M} \Lambda^{3N}). \quad (25)$$

To separate the contributions made by water molecules and the conformations, the *effective energy* W is defined,

$$\exp(-\beta W) = \exp(-\beta H_{mm}) \frac{\int \exp(-\beta H_{mw} - \beta H_{ww}) d\mathbf{r}^N}{\int \exp(-\beta H_{ww}) d\mathbf{r}^N} = \exp(-\beta H_{mm}) \exp(-\beta X), \quad (26)$$

Define

$$\langle \exp(-\beta H_{mw}) \rangle_o = \frac{\int \exp(-\beta H_{mw}) \exp(-\beta H_{ww}) d\mathbf{r}^N}{\int \exp(-\beta H_{ww}) d\mathbf{r}^N}. \quad (27)$$

The effective energy $W(\mathbf{R})$ is:

$$W(\mathbf{R}) = H_{mm}(\mathbf{R}) + X(\mathbf{R}) = H_{mm}(\mathbf{R}) - kT \ln \langle \exp(-\beta H_{mw}) \rangle_o \equiv H_{mm}(\mathbf{R}) + \Delta G^{\text{slv}}(\mathbf{R}). \quad (28)$$

The term $\Delta G^{\text{slv}}(\mathbf{R})$ is called the *solvation free energy* while H_{mm} is the *intra-macromolecular energy*.

After changing \mathbf{R} to interior coordinates \mathbf{q} , it is stated that

$$Z = V 8\pi^2 \int \exp(-\beta H_{ww}) d\mathbf{r}^N \int \exp(-\beta W) d\mathbf{q}, \quad (29)$$

because the interior coordinates has only $3M - 6$ dimension, the integration of the remaining 6 dimension over the system getting the value $V 8\pi^2$, implying that each \mathbf{x}_i in \mathbf{R} can be any point in the system that has volume V . As usual, the probability of finding the system at the

configuration (\mathbf{q}) is:

$$p(\mathbf{q}) = \frac{\exp[-\beta W(\mathbf{q})]}{\int \exp[-\beta W(\mathbf{q})] d\mathbf{q}}. \quad (30)$$

Consequently,

$$\int p(\mathbf{q}) \ln p(\mathbf{q}) d\mathbf{q} = -\ln Z + \ln \int \exp(-\beta H_{ww}) d\mathbf{r}^N + \ln(V8\pi^2) - \beta \int p(\mathbf{q}) W(\mathbf{q}) d\mathbf{q}, \quad (31)$$

From equation (25),

$$\begin{aligned} A &= -kT \int \exp(-\beta H_{ww}) d\mathbf{r}^N + kT \ln \left(\frac{\Lambda^{3M}}{V8\pi^2} \right) + \int p(\mathbf{q}) W(\mathbf{q}) d\mathbf{q} + kT \int p(\mathbf{q}) \ln p(\mathbf{q}) d\mathbf{q} \\ &= A^o + kT \ln \left(\frac{\Lambda^{3M}}{V8\pi^2} \right) + \langle W \rangle - TS^{\text{conf}}, \end{aligned} \quad (32)$$

where $A^o = -kT \int \exp(-\beta H_{ww}) d\mathbf{r}^N$ is the pure Helmholtz free energy of pure solvent; the term $-TS^{\text{conf}} = kT \int p(\mathbf{q}) \ln p(\mathbf{q}) d\mathbf{q}$ is the contribution of the configurational entropy of the macromolecule to the free energy.

The Gibbs free energy is $G = A + PV$. Since the volume is thought negligible under ambient conditions so Gibbs and Helmholtz free energies are considered identical.

Now for any subset of $A \subset \Gamma$, integrals restricted on A gives the Helmholtz energy A_A , i.e.,

$$A_A = A^o + kT \ln \left(\frac{\Lambda^{3M}}{V8\pi^2} \right) + \langle W \rangle_A - TS_A^{\text{conf}}. \quad (33)$$

Thus for two different subsets A and B , the difference in the Helmholtz free energy is

$$\begin{aligned} \Delta A &= A_B - A_A = \langle W \rangle_B - \langle W \rangle_A - T(S_B^{\text{conf}} - S_A^{\text{conf}}) \\ &= \Delta \langle H_{mm} \rangle + \Delta \langle \Delta G^{\text{slv}} \rangle - T \Delta S^{\text{conf}}. \end{aligned} \quad (34)$$

Especially, "If A is the denatured state and B the native state, both of which have to be defined in some way and both of which include many configurations, Eq. (34) gives the free energy of folding."

5.3. Critics of the derivation in Lazaridis and Karplus (2003)

Protein folding is considered a very practical research field, dominating activities are computer simulations with empirical models. There are very few theoretical discussions about protein folding. This derivation in Lazaridis and Karplus (2003) is a rare example deserving an analysis to see why for decades there has been no theoretic progress in this field. Many lessons can be learned from this example.

One important lesson from the derivation Lazaridis and Karplus (2003) is that when dealing with thermodynamics and statistical mechanics, the thermodynamic system must be clearly defined. The system will occupy a space in \mathbb{R}^3 , what is it? How to delimit it?

More importantly, it is not just one conformation \mathbf{R} , but all conformations of a single protein are considered in the derivation. As a single point $\mathbf{R} \in \mathbb{R}^{3M}$, no structural features of the conformation \mathbf{R} are considered, i.e., this particle is structureless. Remember that the research object is the conformation of the protein, we cannot treat them as structureless particles. Yes, classical derivations such as the ideal gas system are defined this way, that is because that the interest is not in the individual particle's structure but the macroscopic properties of the idea gas. The lesson then is that instead of considering all conformations together in a system, specific thermodynamic system has to be tailored for each individual conformation \mathbf{R} . And such a system contains only one conformation \mathbf{R} , with its structure geometry, and other particles such as water molecules, thus the Gibbs free energy of such a system will be indexed by \mathbf{R} , $G = G(\mathbf{R})$.

Perhaps the biggest lesson to be learned is that when solving a problem, one should concentrate on the specific features of the problem to design the ways to attack it, not just imitate successful classical examples.

The derivation of Lazaridis and Karplus (2003) gives the effective energy $W(\mathbf{R})$ as some substitute of the Gibbs free energy without theoretic basis for its relation to the Thermodynamic Principle. Moreover, the formula $W(\mathbf{R})$ tells us nothing of how to calculate it, all are buried in multiple-integrations without clear delimitation. Being the only function for individual conformation \mathbf{R} , it was pointed out in Lazaridis and Karplus (2003) that "The function W defines a hyper-surface in the conformation space of the macromolecule in the presence of equilibrated solvent and, therefore, includes the solvation entropy. This hyper-surface is now often called an 'energy landscape'. It determines the thermodynamics and kinetics of macromolecular conformational transitions." From this comment it can be seen that the authors are not against individual quantities such as $W(\mathbf{R})$ and think they are important to the study of protein folding. Changing the "effective energy" $W(\mathbf{R})$ to the Gibbs free energy $G(\mathbf{R})$, the comment really makes sense. The lesson should be learned is that never invent theoretical concepts without firm theoretical basis. Another one is that always keep in mind that useful Gibbs free energy formula should be calculable.

From now on, the notation $\mathbf{X} = \mathbf{R}$ will be used to represent a conformation. To put the Thermodynamic Principle in practice, not merely as a talking show, what really needed is $G(\mathbf{X})$, the Gibbs free energy of each individual conformation \mathbf{X} , not the effective energy $W(\mathbf{R})$. One hopes that the formula $G(\mathbf{X})$ should be calculable, not buried in multiple integrations. To get such a formula, the grand canonic ensemble and eventually the quantum statistics have to be applied.

6. Necessary preparations for the derivation of the Gibbs free energy formula

Summarizing what have learned from the critics of the derivation in Lazaridis and Karplus (2003), in any attempt of derivation of the Gibbs free energy formula one has to: 1. clearly state all assumptions used in the derivation; and 2. for each conformation \mathbf{X} , set a thermodynamic system $\mathcal{T}_{\mathbf{X}}$ associated with \mathbf{X} ; 3. use the grand canonic ensemble.

6.1. The assumptions

All assumptions here are based on well-known facts of consensus among protein folding students. Let \mathcal{U} be a protein with M atoms $(\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_M)$. A structure (conformation) of \mathcal{U} is a point $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_M) \in \mathbb{R}^{3M}$, $\mathbf{x}_i \in \mathbb{R}^3$ is the atomic center (nuclear) position of \mathbf{a}_i . Alternatively, the conformation \mathbf{X} corresponds to a subset in \mathbb{R}^3 , $P_{\mathbf{X}} = \cup_{i=1}^M B(\mathbf{x}_i, r_i) \subset \mathbb{R}^3$ where r_i 's are van der Waals radii.

1. The proteins discussed here are monomeric, single domain, self folding globular proteins.
2. Therefore, in the case of our selected proteins, the environment of the protein folding, the physiological environment, is pure water, there are no other elements in the environment, no chaperonins, no co-factors, etc. This is a rational simplification, at least when one considers the environment as only the first hydration shell of a conformation, as in our derivation of the $G(\mathbf{X})$.
3. During the folding, the environment does not change.
4. Anfinsen (1973) showed that before folding, the polypeptide chain already has its main chain's and each residue's covalent bonds correctly formed. Hence, our conformations should satisfy the following steric conditions set in Fang (2005) and Fang and Jing (2010): there are $\epsilon_{ij} > 0$, $1 \leq i < j \leq M$ such that for any two atoms \mathbf{a}_i and \mathbf{a}_j in $P_{\mathbf{X}} = \cup_{k=1}^M B(\mathbf{x}_k, r_k)$,

$$\begin{aligned} \epsilon_{ij} &\leq |\mathbf{x}_i - \mathbf{x}_j|, && \text{no covalent bond between } \mathbf{a}_i \text{ and } \mathbf{a}_j; \\ d_{ij} - \epsilon_{ij} &\leq |\mathbf{x}_i - \mathbf{x}_j| \leq d_{ij} + \epsilon_{ij}, && d_{ij} \text{ is the standard bond length between } \mathbf{a}_i \text{ and } \mathbf{a}_j. \end{aligned} \quad (35)$$

All conformations satisfying the steric conditions (35) will be denoted as \mathfrak{X} and in this chapter only $\mathbf{X} \in \mathfrak{X}$ will be considered.

5. A water molecule is treated as a single particle centered at the oxygen nuclear position $\mathbf{w} \in \mathbb{R}^3$, and the covalent bonds in it are fixed. In the Born-Oppenheimer approximation, only the conformation \mathbf{X} is fixed, all particles, water molecules or electrons in the first hydration shell of $P_{\mathbf{X}}$, are moving.
6. As in section 2.3, there are H hydrophobic levels H_i , $i = 1, \dots, H$, such that $\cup_{i=1}^H H_i = (\mathbf{a}_1, \dots, \mathbf{a}_i, \dots, \mathbf{a}_M)$.

6.2. The thermodynamic system $\mathcal{T}_{\mathbf{X}}$

Let d_w be the diameter of a water molecule and $M_{\mathbf{X}}$ be the molecular surface of $P_{\mathbf{X}}$ as defined in Richards (1977) with the probe radius $d_w/2$, see FIGURE 3. Define

$$\mathcal{R}_{\mathbf{X}} = \{\mathbf{x} \in \mathbb{R}^3 : \text{dist}(\mathbf{x}, M_{\mathbf{X}}) \leq d_w\} \setminus P_{\mathbf{X}} \quad (36)$$

as the first hydration shell surrounding $P_{\mathbf{X}}$, where $\text{dist}(\mathbf{x}, S) = \inf_{\mathbf{y} \in S} |\mathbf{x} - \mathbf{y}|$. Then $\mathcal{T}_{\mathbf{X}} = P_{\mathbf{X}} \cup \mathcal{R}_{\mathbf{X}}$ will be our thermodynamic system of protein folding at the conformation \mathbf{X} .

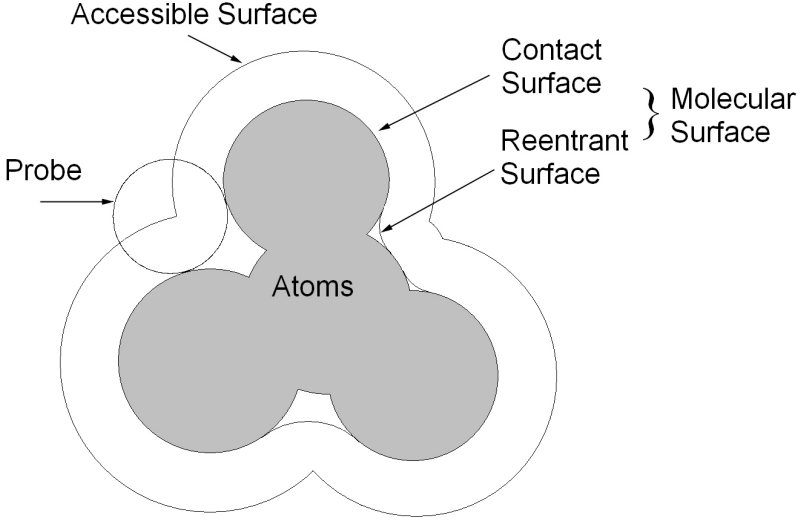


Figure 3. Two dimensional presenting of molecular surface Richards (1977) and solvent accessible surface Lee and Richards (1971). This figure was originally in Fang and Jing (2010).

Let $I_i \subset \{1, 2, \dots, M\}$ be the subset such that $\mathbf{a}_j \in H_i$ if and only if $j \in I_i$. Define $P_{X_i} = \cup_{j \in I_i} B(\mathbf{x}_j, r_j) \subset P_X$ and as shown in FIGURE 4,

$$\mathcal{R}_{X_i} = \{\mathbf{x} \in \mathcal{R}_X : \text{dist}(\mathbf{x}, P_{X_i}) \leq \text{dist}(\mathbf{x}, P_X \setminus P_{X_i})\}, \quad 1 \leq i \leq H. \quad (37)$$

Let $V(\Omega)$ be the volume of $\Omega \subset \mathbb{R}^3$, then

$$\mathcal{R}_X = \cup_{i=1}^H \mathcal{R}_{X_i}, \quad V(\mathcal{R}_X) = \sum_{i=1}^H V(\mathcal{R}_{X_i}), \text{ and for } i \neq j, \quad V(\mathcal{R}_{X_i} \cap \mathcal{R}_{X_j}) = 0. \quad (38)$$

Since M_X is a closed surface, it divides \mathbb{R}^3 into two regions Ω_X and Ω'_X such that $\partial\Omega_X = \partial\Omega'_X = M_X$ and $\mathbb{R}^3 = \Omega_X \cup M_X \cup \Omega'_X$. Note that $P_X \subset \Omega_X$ and all nuclear centers of atoms in the water molecules in \mathcal{R}_X are contained in Ω'_X . Moreover, Ω_X is bounded, therefore, has a volume $V(\Omega_X)$. For $S \subset \mathbb{R}^3$, denote \bar{S} as the closure of S . Define the hydrophobicity subsurface M_{X_i} , $1 \leq i \leq H$, as

$$M_{X_i} = M_X \cap \overline{\mathcal{R}_{X_i}}. \quad (39)$$

Let $A(S)$ be the area of a surface $S \subset \mathbb{R}^3$, then

$$M_X = \cup_{i=1}^H M_{X_i}, \quad A(M_X) = \sum_{i=1}^H A(M_{X_i}), \text{ and if } i \neq j, \text{ then } A(M_{X_i} \cap M_{X_j}) = 0. \quad (40)$$

Although the shape of each atom in a molecule is well defined by the theory of atoms in molecules as in Bader (1990) and Popelier (2000), what concerning us here is the overall

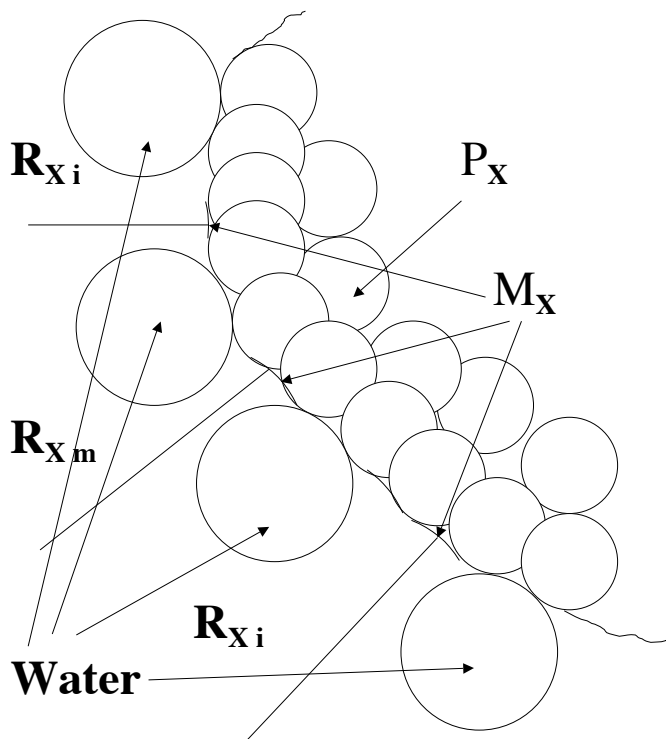


Figure 4. Note that \mathcal{R}_{X_i} generally are not connected, i.e., having more than one block.

shape of the structure P_X . The cutoff of electron density $\rho \geq 0.001\text{au}$ in Bader (1990) and Popelier (2000) gives the overall shape of a molecular structure that is just like P_X , a bunch of overlapping balls. Moreover, the boundary of the $\rho \geq 0.001\text{au}$ cutoff is very similar to the molecular surface M_X which was defined by Richards (1977) and was shown has more physical meaning as the boundary surface of the conformation P_X in Tuñón *et. al.* (1992) and Jackson and Sternberg (1993).

7. Gibbs free energy formula: Classical statistical mechanics derivation

The grand canonic ensemble or macroscopic ensemble will be applied to derive the desired Gibbs free energy formula $G(X)$. In addition to let the number of water molecules vary, the assumptions is that the chemical potential μ will be different for water molecules contacting to different hydrophobicity levels H_i (or falling in \mathcal{R}_{X_i}). Counting the numbers N_i of water molecules that contact to atoms in H_i , the N and μ in equation (15) should be modified to $N = (N_1, \dots, N_i, \dots, N_H)$, $\mu = (\mu_1, \dots, \mu_i, \dots, \mu_H)$. Let $(\mathbf{q}, \mathbf{p}) \in \mathcal{R}_X^M \times \mathbb{R}^{3M}$ be the water molecules' phase space for a fixed N , where $M = \sum_{i=1}^H N_i$. Let $H_X = H_X(\mathbf{q}, \mathbf{p})$ be the Hamiltonian. The grand canonic phase density function will be

$$\begin{aligned}
p_{\mathbf{X}}(\mathbf{q}, \mathbf{p}, N) &= \frac{\exp\{-\beta[H_{\mathbf{X}}(\mathbf{q}, \mathbf{p}) - \sum_{i=1}^H \mu_i N_i]\}}{\sum_{M=0}^{\infty} \frac{1}{M! h^{3M}} \sum_{\sum N_i=M} \prod_{i=1}^H \int_{\mathcal{R}_{\mathbf{X}_i}^{N_i}} d\mathbf{q}^{N_i} \int_{\mathbb{R}^{3M}} \exp\{-\beta[H_{\mathbf{X}} - \sum_{i=1}^H \mu_i N_i]\} d\mathbf{p}^{3M}} \\
&= \frac{\exp\{-\beta[H_{\mathbf{X}}(\mathbf{q}, \mathbf{p}) - \sum_{i=1}^H \mu_i N_i]\}}{\mathcal{Z}(T, V, \mu)}. \tag{41}
\end{aligned}$$

The entropy $S(\mathbf{X}) = S(\mathcal{T}_{\mathbf{X}})$ is

$$\begin{aligned}
S &= \langle -k \ln p_{\mathbf{X}} \rangle = -k \sum_{M=0}^{\infty} \sum_{N_1+\dots+N_H=M} \prod_{i=1}^H \int_{\mathcal{R}_{\mathbf{X}_i}^{N_i}} d\mathbf{q}^{N_i} \int_{\mathbb{R}^{3M}} \ln p_{\mathbf{X}} p_{\mathbf{X}} d\mathbf{p}^M \\
&= -k \sum_{M=0}^{\infty} \sum_{N_1+\dots+N_H=M} \prod_{i=1}^H \int_{\mathcal{R}_{\mathbf{X}_i}^{N_i}} d\mathbf{q}^{N_i} \left[\beta \sum_{i=1}^H \mu_i N_i - \beta H_{\mathbf{X}}(\mathbf{q}, \mathbf{p}) - \ln \mathcal{Z} \right] p_{\mathbf{X}} d\mathbf{p}^M \tag{42}
\end{aligned}$$

$$= \frac{1}{T} \left[\langle H \rangle - \sum_{i=1}^H \mu_i \langle N_i \rangle + kT \ln \mathcal{Z}(T, V, \mu) \right] \tag{43}$$

$$= \frac{1}{T} \left[U(\mathcal{T}_{\mathbf{X}}) - \sum_{i=1}^H \mu_i N_i(\mathcal{T}_{\mathbf{X}}) + kT \ln \mathcal{Z}(T, V, \mu) \right]$$

where $U(\mathbf{X}) = U(\mathcal{T}_{\mathbf{X}}) = \langle H \rangle$ is the inner energy, $N_i(\mathbf{X}) = \langle N_i \rangle$ the mean number of water molecules in $\mathcal{R}_{\mathbf{X}_i}$. By equation (8), $kT \ln \mathcal{Z}(T, V, \mu) = -\phi(T, V, \mu) = PV(\mathcal{T}_{\mathbf{X}})$. Therefore, from $G = U + PV - TS$,

$$G(\mathbf{X}) = G(\mathcal{T}_{\mathbf{X}}) = U(\mathbf{X}) + PV(\mathcal{T}_{\mathbf{X}}) - TS(\mathcal{T}_{\mathbf{X}}) = \sum_{i=1}^H \mu_i N_i(\mathbf{X}). \tag{44}$$

The Gibbs free energy given in formula (44) does not involve any integration at all, just counting the number of water molecules contacting atoms in H_i . Furthermore, against the effective energy, potential function H_{mm} plays no role at all, a surprise indeed. But formula (44) also is not easy to calculate, counting the number of water molecules actually need more knowledge of the conformation's boundary, the molecular surface $M_{\mathbf{X}}$. Formula (44) can be directly transferred into a geometric version.

7.1. Converting formula (44) to a geometric version

Since every water molecule in $\mathcal{R}_{\mathbf{X}_i}$ has contact with the surface $M_{\mathbf{X}_i}$, $N_i(\mathbf{X})$ is proportional to the area $A(M_{\mathbf{X}_i})$. Therefore, there are $v_i > 0$, such that

$$v_i A(M_{\mathbf{X}_i}) = N_i(\mathbf{X}), \quad 1 \leq i \leq H. \tag{45}$$

Substitute in (44),

$$G(\mathbf{X}) = G(\mathcal{T}_{\mathbf{X}}) = \sum_{i=1}^H \mu_i v_i A(M_{\mathbf{X}_i}). \tag{46}$$

For each conformation \mathbf{X} , the molecular surface $M_{\mathbf{X}}$ is calculable, see Connolly (1983). The areas $A(M_{\mathbf{X}})$ and $A(M_{\mathbf{X}_i})$ are also calculable. Therefore, unlike the formula given in (34), this formula is calculable. Moreover, our derivation theoretically justified the surface area models that will be discussed later, only difference is that the molecular surface area is used here instead of the solvent accessible surface area.

But still something is missing. That is, the volume $V(\mathcal{T}_{\mathbf{X}})$, an important thermodynamic quantity, does not show here at all. It seems that no way to put the $V(\mathcal{T}_{\mathbf{X}})$ here in the classical statistical mechanics. To resolve this, the quantum statistical mechanics is necessary.

8. A quantum statistical theory of protein folding

In 1929 Dirac wrote: “The underlying physical laws necessary for the mathematical theory of ... the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.” (quoted from Bader (1990), page 132). Yes, the multidimensional Schrödinger equation for protein folding is beyond our ability to solve, no matter how fast and how powerful our computers are. But mathematical theory guarantees that there is a complete set of eigenvalues (energy levels) and eigenfunctions to the Schrödinger equation in the Born-Oppenheimer approximation. Then consider that in the statistical mechanics, ensembles collect all (energy) states of the same system. Although one cannot have exact solutions to the Schrödinger equation, the eigenvalues of it are theoretically known. Thus one can apply the grand canonical ensemble to obtain the desired Gibbs free energy formula $G(\mathbf{X})$. This is the main idea of the derivation.

8.1. The Schrödinger equation

For any conformation $\mathbf{X} \in \mathfrak{X}$, let $\mathbf{W} = (\mathbf{w}_1, \dots, \mathbf{w}_i, \dots, \mathbf{w}_N) \in \mathbb{R}^{3N}$ be the nuclear centers of water molecules in $\mathcal{R}_{\mathbf{X}}$ and $\mathbf{E} = (\mathbf{e}_1, \dots, \mathbf{e}_i, \dots, \mathbf{e}_L) \in \mathbb{R}^{3L}$ be electronic positions of all electrons in $\mathcal{T}_{\mathbf{X}}$. Then the Hamiltonian for the system $\mathcal{T}_{\mathbf{X}}$ is

$$\hat{H} = \hat{T} + \hat{V} = - \sum_{i=1}^M \frac{\hbar^2}{2m_i} \nabla_i^2 - \frac{\hbar^2}{2m_w} \sum_{i=1}^N \nabla_i^2 - \frac{\hbar^2}{2m_e} \sum_{i=1}^L \nabla_i^2 + \hat{V}(\mathbf{X}, \mathbf{W}, \mathbf{E}), \quad (47)$$

where m_i is the nuclear mass of atom \mathbf{a}_i , m_w and m_e are the masses of water molecule and electron; ∇_i^2 is Laplacian in corresponding \mathbb{R}^3 ; and V the potential.

8.2. The first step of the Born-Oppenheimer approximation

Depending on the shape of $P_{\mathbf{X}}$, for each i , $1 \leq i \leq H$, the maximum numbers $N_{\mathbf{X}_i}$ of water molecules contained in $\mathcal{R}_{\mathbf{X}_i}$ vary. Theoretically all cases are considered, i.e., there are $0 \leq N_i \leq N_{\mathbf{X}_i}$ water molecules in $\mathcal{R}_{\mathbf{X}_i}$, $1 \leq i \leq H$. Let $M_0 = 0$ and $M_i = \sum_{j \leq i} N_j$ and $\mathbf{W}_i = (\mathbf{w}_{M_{i-1}+1}, \dots, \mathbf{w}_{M_{i-1}+j}, \dots, \mathbf{w}_{M_i}) \in \mathbb{R}_{\mathbf{X}_i}^{N_i}$, $1 \leq i \leq H$, and $\mathbf{W} = (\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_{M_H}) \in \prod_{i=1}^H \mathbb{R}_{\mathbf{X}_i}^{N_i}$ denote the nuclear positions of water molecules in $\mathcal{R}_{\mathbf{X}}$. As well, there will be all possible numbers $0 \leq N_e < \infty$ of electrons in $\mathcal{T}_{\mathbf{X}}$. Let $\mathbf{E} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{N_e}) \in \mathbb{R}^{3N_e}$ denote their nuclear positions. For each fixed $\mathbf{X} \in \mathfrak{X}$ and $N = (N_1, \dots, N_H, N_e)$, the Born-Oppenheimer

approximation has the Hamiltonian

$$\hat{H}_X = -\frac{\hbar^2}{2} \left\{ \frac{1}{m_w} \sum_{j=1}^{M_H} \nabla_j^2 + \frac{1}{m_e} \sum_{v=1}^{N_e} \nabla_v^2 \right\} + \hat{V}(\mathbf{X}, \mathbf{W}, \mathbf{E}). \quad (48)$$

The eigenfunctions $\psi_i^{\mathbf{X},N}(\mathbf{W}, \mathbf{E}) \in L_0^2(\prod_{i=1}^H \mathcal{R}_{\mathbf{X}_i}^{N_i} \times \mathcal{T}_{\mathbf{X}}^{N_e}) = \mathcal{H}_{\mathbf{X},N}$, $1 \leq i < \infty$, comprise an orthonormal basis of $\mathcal{H}_{\mathbf{X},N}$. Denote their eigenvalues (energy levels) as $E_{\mathbf{X},N}^i$, then $\hat{H}_X \psi_i^{\mathbf{X},N} = E_{\mathbf{X},N}^i \psi_i^{\mathbf{X},N}$.

8.3. Grand partition function and grand canonic density operator

Since the numbers N_i and N_e vary, the grand canonic ensemble should be adopted. Let μ_i be the chemical potentials, that is, the Gibbs free energy per water molecule in $\mathcal{R}_{\mathbf{X}_i}$. Let μ_e be electron chemical potential. The grand canonic density operator is like in equation (15), or see Greiner *et. al.* (1994) and Dai (2007)

$$\hat{\rho}_X = \exp \left\{ -\beta \left[\hat{H}_X - \sum_{i=1}^H \mu_i \hat{N}_i - \mu_e \hat{N}_e - \phi(\mathbf{X}) \right] \right\}. \quad (49)$$

where $\phi(\mathbf{X})$ is the grand canonic potential ϕ in equation (8) with the index \mathbf{X} and the grand partition function is

$$\begin{aligned} \exp[-\beta\phi(\mathbf{X})] &= \text{Trace} \left\{ \exp \left[-\beta \left(\hat{H}_X - \sum_{i=1}^H \mu_i \hat{N}_i - \mu_e \hat{N}_e \right) \right] \right\} \\ &= \sum_{i,N} \exp \left\{ -\beta \left[E_{\mathbf{X},N}^i - \sum_{i=1}^H \mu_i N_i - \mu_e N_e \right] \right\}. \end{aligned} \quad (50)$$

8.4. The Gibbs free energy $G(\mathbf{X})$

As in equation (16), under the grand canonic ensemble the entropy $S(\mathbf{X}) = S(\mathcal{T}_X)$ of the system \mathcal{T}_X is

$$\begin{aligned} S(\mathbf{X}) &= -k \text{Trace}(\hat{\rho}_X \ln \hat{\rho}_X) = -k \langle \ln \hat{\rho}_X \rangle = k\beta \left\langle \hat{H}_X - \phi(\mathbf{X}) - \sum_{i=1}^H \mu_i \hat{N}_i - \mu_e \hat{N}_e \right\rangle \\ &= \frac{1}{T} \left[\langle \hat{H}_X \rangle - \langle \phi(\mathbf{X}) \rangle - \sum_{i=1}^H \mu_i \langle \hat{N}_i \rangle - \mu_e \langle \hat{N}_e \rangle \right] \\ &= \frac{1}{T} \left[U(\mathbf{X}) - \phi(\mathbf{X}) - \sum_{i=1}^H \mu_i N_i(\mathbf{X}) - \mu_e N_e(\mathbf{X}) \right]. \end{aligned} \quad (51)$$

Denote $\langle \hat{N}_i \rangle = N_i(\mathbf{X})$ as the mean number of water molecules in $\mathcal{R}_{\mathbf{X}i}$, $1 \leq i \leq H$, and $\langle \hat{N}_e \rangle = N_e(\mathbf{X})$ the mean number of electrons in $\mathcal{T}_{\mathbf{X}}$. The inner energy $\langle \hat{H}_{\mathbf{X}} \rangle$ of the system $\mathcal{T}_{\mathbf{X}}$ is denoted as $U(\mathbf{X}) = U(\mathcal{T}_{\mathbf{X}})$. By equation (8) and the remark after it $\phi(\mathbf{X})(T, V, \mu_1, \dots, \mu_H, \mu_e) = -PV(\mathbf{X})$, where $V(\mathbf{X}) = V(\mathcal{T}_{\mathbf{X}})$ is the volume of the thermodynamic system $\mathcal{T}_{\mathbf{X}}$. Thus by equation (51) the Gibbs free energy $G(\mathbf{X}) = G(\mathcal{T}_{\mathbf{X}})$ in formula (1) is obtained:

$$G(\mathbf{X}) = G(\mathcal{T}_{\mathbf{X}}) = PV(\mathbf{X}) + U(\mathbf{X}) - TS(\mathbf{X}) = \sum_{i=1}^H \mu_i N_i(\mathbf{X}) + \mu_e N_e(\mathbf{X}). \quad (52)$$

8.5. Converting formula (1) to geometric form (2)

As in the classical statistical mechanics case,

$$v_i A(M_{\mathbf{X}i}) = N_i(\mathbf{X}), \quad 1 \leq i \leq H. \quad (53)$$

Similarly, there will be a $v_e > 0$ such that $v_e V(\mathcal{T}_{\mathbf{X}}) = N_e(\mathbf{X})$. By the definition of $\mathcal{T}_{\mathbf{X}}$ and $\Omega_{\mathbf{X}}$, it is roughly $V(\mathcal{T}_{\mathbf{X}} \setminus \Omega_{\mathbf{X}}) = d_w A(M_{\mathbf{X}})$. Thus

$$N_e(\mathbf{X}) = v_e V(\mathcal{T}_{\mathbf{X}}) = v_e [V(\Omega_{\mathbf{X}}) + V(\mathcal{T}_{\mathbf{X}} \setminus \Omega_{\mathbf{X}})] = v_e V(\Omega_{\mathbf{X}}) + d_w v_e A(M_{\mathbf{X}}). \quad (54)$$

Substitute equations (45) and (54) into formula (1), formula (2) is obtained.

9. Some remarks

The question to applying fundamental physical laws directly to the protein folding problem is, can it be done? It should be checked that how rigorous is the derivation and be asked that are there any fundamental errors? Possible ways to modify the formula or the derivation will also be discussed.

By applying quantum statistics the protein folding problem is theoretically treated. A theory is useful only if it can make explanations to the observed facts and if it can simplify and improve research methods as well as clarify concepts. It will be shown that $G(\mathbf{X})$ can do exactly these.

If the same theoretical result can be derived from two different disciplines, it is often not just by chance. An early phenomenological mathematical model Fang (2005), starting from purely geometric reasoning, has achieved formula (2), with just two hydrophobic levels, hydrophobic and hydrophilic.

A theory also has to be falsifiable, that is making a prediction to be checked. The fundamental prediction is that minimizing formula (1) or (2) the native structures will be obtained for the amino acid sequences of proteins considered in the assumptions of the formulas. That can only be done after the actual values of the chemical potentials appear in the formulas, for the physiological environment, are determined.

9.1. How rigorous is the derivation?

Two common tools in physics, the first step of the Born-Oppenheimer approximation in quantum mechanics and the grand canonic ensemble in statistical physics, are applied to obtain formula (1).

9.1.1. The Born-Oppenheimer approximation

The Born-Oppenheimer approximation “treats the electrons as if they are moving in the field of fixed nuclei. This is a good approximation because, loosely speaking, electrons move much faster than nuclei and will almost instantly adjust themselves to a change in nuclear position.” Popelier (2000). Since the mass of a water molecule is much less than the mass of a protein, this approximation can be extended to the case of when \mathbf{X} changes the other particles, electrons and water molecules, will quickly adjust themselves to the change as well.

9.1.2. The statistical physics in general and the grand canonic ensemble in particular

“Up to now there is no evidence to show that statistical physics itself is responsible for any mistakes,” the Preface of Dai (2007). Via the ensemble theory of statistical mechanics only one protein molecule and particles in its immediate environment are considered, it is justified since as pointed out in Dai (2007) page 10, “When the duration of measurement is short, or the number of particles is not large enough, the concept of ensemble theory is still valid.” And among different ensembles, “Generally speaking, the grand canonic ensemble, with the least restrictions, is the most convenient in the mathematical treatment.” Dai (2007) page 16. In fact, the canonic ensemble has been tried and ended with a result that the eigenvalues of the quantum mechanics system have to be really calculated, to do it accurately is impossible.

The derivations in this chapter only puts together the two very common and sound practices: the Born-Oppenheimer approximation (only the first step) and the grand canonic ensemble, and apply them to the protein folding problem. As long as protein folding obeys the fundamental physical laws, there should not be any serious error with the derivation.

9.2. Equilibrium and quasi-equilibrium

A protein’s structure will never be in equilibrium, in fact, even the native structure is only a snapshot of the constant vibration state of the structure. The best description of conformation \mathbf{X} is given in Chapter 3 of Bader (1990). Simply speaking, a conformation \mathbf{X} actually is any point \mathbf{Y} such that all \mathbf{y}_i are contained in a union of tiny balls centered at \mathbf{x}_i , $i = 1, \dots, M$. In this sense, it can only be anticipated that a quasi-equilibrium description (such as the heat engine, Baily (1994) page 94) of the thermodynamic states of the protein folding. This has been built-in in the Thermodynamical Principle of Protein Folding. So the quantities such as $S(\mathbf{X})$, $\phi(\mathbf{X})$, and $G(\mathbf{X})$ can only be understood in this sense. That is, observing a concrete folding process one will see a series of conformations \mathbf{X}_i , $i = 1, 2, 3, \dots$. The Thermodynamic Principle then says that measuring the Gibbs free energy $G(\mathbf{X}_i)$ one will observe that eventually $G(\mathbf{X}_i)$ will converge to a minimum value and the \mathbf{X}_i will eventually approach to the native structure. While all the time, no conformation \mathbf{X}_i and thermodynamic system $\mathcal{T}_{\mathbf{X}_i}$ are really in equilibrium state.

9.3. Potential energy plays no role in protein folding

Formulas (1) and (2) theoretically show that hydrophobic effect is the driving force of protein folding, it is not just solvent free energy besides the pairwise interactions such as the Coulombs, etc., as all force fields assumed. Only in the physiological environment the

hydrophobic effect works towards to native structure, otherwise it will push denaturation as discussed in explanation of folding and unfolding. Formulas (1) and (2) show that the Gibbs free energy is actually independent of the potential energy, against one's intuition and a bit of surprising. The explanation is that during the folding process, all covalent bonds in the main chain and each side chain are kept invariant, the potential energy has already played its role in the synthesis process of forming the peptide chain, which of course can also be described by quantum mechanics. According to Anfinsen (1973), protein folding is after the synthesis of the whole peptide chain, so the synthesis process can be skipped and the concentration can be focused on the folding process.

The steric conditions (35) will just keep this early synthesis result, not any $\mathbf{X} = (x_1, \dots, x_i, \dots, x_M)$ is eligible to be a conformation, it has to satisfy the steric conditions (35). The steric conditions not only pay respect to the bond length, it also reflect a lot of physic-chemical properties of a conformation: They are defined via the allowed minimal atomic distances, such that for non-bonding atoms, the allowed minimal distances are: shorter between differently charged or polarized atoms; a little longer between non-polar ones; and much longer (generally greater than the sum of their radii) between the same charged ones, etc. For example, minimal distance between sulfur atoms in Cysteine residues to form disulfide bonds is allowed. And for any newly found intramolecular covalent bond between side chains, such as the isopeptide bonds in Kang and Baker (2011), the steric conditions can be easily modified to allow the newly found phenomenon.

The drawback of the steric conditions is that the minimization in equation (57) becomes a constrained minimization.

9.4. Unified explanation of folding and denaturation

Protein denaturation is easy to happen, even if the environment is slightly changed, as described by Hsien Wu (1931). (Hsien Wu (1931) is the 13th article that theorizes the results of a series experiments, and a preliminary report was read before the XIIIth International Congress of Physiology at Boston, August 19-24, 1929, and published in the *Am. J. Physiol.* for October 1929. In which Hsien Wu first suggested that the denatured protein is still the same molecule, only structure has been changed.) Anfinsen in various experiments showed that after denaturation by changed environment, if removing the denature agent, certain globular proteins can spontaneously refold to its native structure, Anfinsen (1973). The spontaneous renaturation suggests that protein folding does not need outside help, at least to the class of proteins in this chapter. Therefore, the fundamental law of thermodynamics asserts that in the environments in which a protein can fold, the native structure must have the minimum Gibbs free energy. The same is true for denaturation, under the denatured environment, the native structure no longer has the minimum Gibbs free energy, some other structure(s), will have the minimum Gibbs free energy. Thus let En present environment, any formula of Gibbs free energy should be stated as $G(\mathbf{X}, En)$ instead of just $G(\mathbf{X})$, unless the environment is specified like in this chapter. Let En_N be the physiological environment and En_U be some denatured environment, \mathbf{X}_N be the native structure and \mathbf{X}_U be one of the denatured stable structure in En_U , then the thermodynamic principle for both of protein folding and unfolding should be that

$$G(\mathbf{X}_N, En_N) < G(\mathbf{X}_U, En_N), \quad G(\mathbf{X}_N, En_U) > G(\mathbf{X}_U, En_U). \quad (55)$$

To check this, an experiment should be designed that can suddenly put proteins in a different environment. Formulas (1) and (2) should be written as $G(\mathbf{X}, En_N)$. Indeed, the chemical potentials μ_e and μ_i 's are Gibbs free energies per corresponding particles, $\mu = u + Pv - Ts$. Two environment parameters, temperature T and pressure P , explicitly appear in μ , the inner energy u and entropy s may also implicitly depend on the environment. According to formulas (1) and (2), if $\mu_i < 0$, then make more H_i atoms to expose to water (make larger $A(M_{X_i})$) will reduce the Gibbs free energy. If $\mu_i > 0$, then the reverse will happen. Increase or reduce the H_i atoms' exposure to water ($A(M_{X_i})$), the conformation has to change. The conformation changes to adjust until a conformation \mathbf{X}_N is obtained, such that the net effect of any change of the conformation will either increase some H_i atoms' exposure to water while $\mu_i > 0$ or reduce H_i atoms' exposure to water while $\mu_i < 0$. In other words, the $G(\mathbf{X}, En_N)$ achieves its minimum at $G(\mathbf{X}_N, En_N)$. Protein folding, at least for the proteins considered in the assumptions, is explained very well by formulas (1) and (2).

In changed environment, the chemical potentials μ_e and μ_i 's in formulas (1) and (2) changed their values. With the changed chemical potentials, $G(\mathbf{X}, En_U)$ has the same form as $G(\mathbf{X}, En_N)$ but different chemical potentials. Therefore, the structure \mathbf{X}_U will be stable, according to the second inequality in (55), the process is exactly the same as described for the protein folding if the changing environment method does not include introducing new kinds (non-water) of particles, for example, if only temperature or pressure is changed.

Even in the new environment including new kinds of particles, formulas (1) and (2) can still partially explain the denaturation, only that more obstructs prevent the protein to denature to \mathbf{X}_U , but any way it will end in some structure other than the \mathbf{X}_N , the protein is denatured. Actually, this is a hint of how to modify the current formulas to extend to general proteins.

9.5. Explain hydrophobic effect and the role played by hydrogen bonding

In 1959, by reviewing the literature Kauzmann concluded that the hydrophobic effect is the main driving force in protein folding, Kauzmann (1959). Empirical correlation between hydrophobic free energy and aqueous cavity surface area was noted as early as by Reynolds *et.al.* (1974), giving justification of the hydrophobic effect. Various justifications of hydrophobic effect were published, based on empirical models of protein folding, for example, Dill (1990). But the debate continues to present, some still insist that it is the hydrogen bond instead of hydrophobic effect plays the main role of driving force in protein folding, for example, Rose *et. al* (2006). The theoretically derived formulas (1) and (2) can explain why the hydrophobic effect is indeed the driving force. A simulation of reducing hydrophobic area alone by Fang and Jing (2010) shows that the result is the appearance of regularly patterned intramolecular hydrogen bonds associated to the secondary structures.

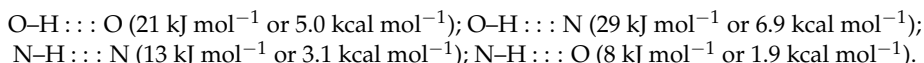
In fact, according to formulas (1) and (2), if $\mu_i < 0$, then make more H_i atoms to appear in the boundary of P_X will reduce the Gibbs free energy. If $\mu_i > 0$, then the reverse will happen, reducing the exposure of H_i atoms to water will reduce the Gibbs free energy. This gives a theoretical explanation of the hydrophobic effect. The kinetic formulas $\mathbf{F}_i = -\nabla_{\mathbf{x}_i} G(\mathbf{X})$ (will be discussed later) is the force that push the conformation to change to the native structure.

The mechanics stated above works through the chemical potentials μ_i for various levels of hydrophobicity. In physiological environment, all hydrophobic H_i 's will have positive μ_i , all

hydrophilic H_i 's will have negative μ_i . Thus changing conformation P_X such that the most hydrophilic H_i ($\mu_i = \min(\mu_1, \dots, \mu_H)$) gets the first priority to appear on the boundary, and the most hydrophobic H_i ($\mu_i = \max(\mu_1, \dots, \mu_H)$) gets the first priority to hide in the hydrophobic core to avoid contacting with water molecules, etc. One should keep in mind that all the time, the steric conditions (35) have to be obeyed.

But the hydrophobic effect is actually partially working through hydrogen bond formation. This is well presented in the chemical potentials in formulas (1) and (2). In fact, the values of the chemical potentials reflect the ability of the atoms or atom groups to form hydrogen bond, either with another atom group in the protein or with water molecules. This gives a way to theoretically or experimentally determine the values of hydrophilic chemical potentials: checking the actual energy value of the hydrogen bond.

According to Fikelstein and Ptitsyn (2002), energies of hydrogen bonds appearing in protein (intermolecular or intramolecular) are (the positive sign means that to break it energy is needed) and their energies are:



For hydrophobic ones, it will be more complicated, common sense is that it reduces the entropy that certainly comes from the inability of forming hydrogen bonds with water molecules. Hence although hydrophobic effect is the driving force of protein folding, it works through the atom's ability or inability to form hydrogen bonds with water molecules.

How to explain the intramolecular hydrogen bonds? It seems that formulas (1) and (2) do not address this issue. The possible theory is that the amino acid sequence of a protein is highly selectable in evolution, in fact only a tiny number of amino acid sequences can really become a protein.

Indeed, suppose in average each species (or "kind" of prokaryote) has 10^5 proteins (*Homo sapiens* has around 3×10^5), and assume that per protein has 100 variants (versions with tiny difference in the peptide sequence of the protein), then there are at most 10^{47} peptide sequences that can really produce a natural protein. Now further suppose that only one in 10^{13} theoretically protein producing peptide sequences on the earth get a chance to be realized, then there will be at most 10^{60} possible protein producing peptide sequence. A huge number! The number of peptide sequences of length less than or equal to n is

$$N(n) = \sum_{i=1}^n 20^i = \frac{20^{n+1} - 20}{19} = \frac{20}{19}(20^n - 1) \cong 20^{n+0.0171} \cong 10^{1.301(n+0.0171)}. \quad (56)$$

The longest amino acid sequence in the record of ExPASy Proteomics Server has 35,213 residues. Then $N(35,213) > 10^{1.3 \times 35,213} > 10^{45060}$ and the ratio of the number of potentially protein producing peptide sequences to the number of all possible sequence of length up to 35,213 is less than $10^{60} / 10^{45060} = 10^{-45000}$, so tiny a number that it is undistinguishable from zero. Even assuming that the longest peptide sequence is only 400, the ratio is still less than 10^{-460} . How small a chance that a random peptide sequence happens to be a protein's peptide sequence!

With these highly specially selected peptide sequences, one can assume that while shrinking the various hydrophobic surfaces to form a hydrophobic core, residues are put in positions to form secondary structures and their associated hydrogen bonds. This sounds a little bit too arbitrary. But the huge number of candidate peptide sequences makes the evolutionary selection not only possible but also probable. Moreover, a simulation of shrinking hydrophobic surface area alone indeed produced secondary structures and hydrogen bonds. The simulation was reported by Fang and Jing (2010). Without calculating any dihedral angles or electronic charges, without any arbitrary parameter, paying no attention to any particular atom's position, by just reducing hydrophobic surface area (there it was assumed that there are only two kinds of atoms, hydrophobic and hydrophilic), secondary structures and hydrogen bonds duly appeared. The proteins used in the simulation are 2i9c, 2hng, and 2ib0, with 123, 127, and 162 residues. No simulation of any kind of empirical or theoretical models had achieved such a success. More than anything, this simulation should prove that hydrophobic effect alone will give more chance of forming intramolecular hydrogen bonds. Indeed, pushing hydrophilic atoms to make hydrogen bonds with water molecules will give other non-boundary hydrophilic groups more chance to form intramolecular hydrogen bonds.

Again formula (2) can partly explain the success of this simulation, when there are only two hydrophobic classes in formula (2), the hydrophobic area presents the main positive part of the Gibbs free energy, reducing it is reducing the Gibbs free energy, no matter what is the chemical potential's real value.

9.6. Explanation of the successes of surface area models

In 1995, Wang *et al* (1995) compared 8 empirical energy models by testing their ability to distinguish native structures and their close neighboring compact non-native structures. Their models WZS are accessible surface area models with 14 hydrophobicity classes of atoms, $\sum_{i=1}^{14} \sigma_i A_i$. Each two combination of three targeting proteins were used to train WZS to get σ_i , hence there are three models WZS1, WZS2, and WZS3. Among the 8 models, all WZS's performed the best, distinguishing all 6 targeting proteins. The worst performer is the force field AMBER 4.0, it failed in distinguishing any of the 6 targets.

These testing and the successes of various surface area models such as Eisenberg and MacLachlan (1986), showed that instead of watching numerous pairwise atomic interactions, the surface area models, though looking too simple, have surprising powers. Now the formula (2) gives them a theoretic justification. On the other hand, the successes of these models also reinforce the theoretical results.

There is a gap between the accessible surface area model in Eisenberg and Maclanchlan (1986) and the experiment results (surface tension), as pointed out in Tuñón *et. al.* (1992). The gap disappeared when one uses the molecular surface area to replace the accessible surface area, in Tuñón *et. al.* (1992) it was shown that molecular surface area assigned of 72-73 cal/mol/Å² perfectly fits with the macroscopic experiment data. Later it was asserted that the molecular surface is the real boundary of protein in its native structure by Jackson and Sternberg (1993).

By the definition of Ω'_X , as shown in FIGURE 3 and FIGURE 4, water molecules contact to P_X must be outside the molecular surface M_X . Since the assessable surface is in the middle of the first hydration shell, it is better to use the molecular surface M_X as the boundary of

the conformation P_X . Moreover, the conversion of the mean numbers $N_i(\mathbf{X})$ to surface area, $N_i(\mathbf{X}) = v_i A(M_{X,i})$, only works for the molecular surface, not for the accessible surface. This can explain the conclusions that molecular surface is a much better boundary than accessible surface as stated in Tuñón *et. al.* (1992) and Jackson and Sternberg (1993).

In fact, the advantage of the solvent accessible surface is that by definition of it one knows exactly each atom occupies which part of the surface, therefore, one can calculate its share in surface area. This fact may partly account why there are so many models based on the solvent accessible surface, even people knew the afore mentioned gap. For other surfaces, one has to define the part of surface that belongs to a specific hydrophobicity class. This was resolved in Fang (2005) via the distance function definition as is used here.

All surface area models neglected one element, the volume of the structure. As early as in the 1970's, Richards and his colleagues already pointed out that the native structure of globular proteins is very dense, or compact, (density = 0.75, Richards (1977)). To make a conformation denser, obviously we should shrink the volume $V(\Omega_X)$. The model in Fang (2005) introduced volume term but kept the oversimplification of all atoms are either hydrophobic or hydrophilic. The derivation of formulas (1) and (2) shows that volume term should be counted, but it may be that $v_e \mu_e$ is very small, in that case, volume maybe really is irrelevant.

9.7. Coincidence with phenomenological mathematical model

If a theoretical result can be derived from two different disciplines, its possibility of correctness will be dramatically increased. Indeed, from a pure geometric consideration, a phenomenological mathematical model, $G(\mathbf{X}) = aV(\Omega_X) + bA(M_X) + cA(M_{X1})$, $a, b, c > 0$ (it was assumed that there are only two hydrophobicity levels, hydrophobic and hydrophilic, the hydrophilic surface area $A(M_{X2})$ is absorbed in $A(M_X)$ by $A(M_{X2}) = A(M_X) - A(M_{X1})$), was created in Fang (2005). It was based on the well-known global geometric characteristics of the native structure of globular proteins: 1. high density; 2. smaller surface area; 3. hydrophobic core, as demonstrated and summarized in Richards (1977) and Novotny *et.al* (1984). So that to obtain the native structure, one should shrink the volume (increasing the density) and surface area, and form better hydrophobic core (reducing the hydrophobic surface area $A(M_{X1})$) simultaneously and cohesively.

The coincidence of formula (2) and the phenomenological mathematical model of Fang (2005) cannot be just a coincidence. Most likely, it is the same natural law reflected in different disciplines. The advantage of formula (2) is that everything there has its physical meaning.

10. Applications

After the derivation it is suitable to point out some immediate applications of the formula $G(\mathbf{X})$.

10.1. Energy surface or landscape

An obvious application is the construction of Gibbs free energy surface or landscape. Empirical estimate is no longer needed, the Gibbs free energy formula $G : \mathfrak{X} \rightarrow \mathbb{R}$ gives

a graph $(\mathbf{X}, G(\mathbf{X}))$ over the space \mathfrak{X} (all eligible conformations for a given protein), and this is nothing but the Gibbs free energy surface. Mathematically it is a $3M$ dimensional hyper-surface. Its characteristics concerned by students of energy surface theory, such as how rugged it is? how many local minimums are there? is there a funnel? etc., can be answered by simple calculations of the formula.

Since the function G is actually defined on the whole \mathbb{R}^{3M} (on an domain of \mathbb{R}^{3M} containing all \mathfrak{X} is enough), mathematical tools can be explored to study its graph, and compare the results with the restricted conformations. One important question is: Does the absolute minimum structure belongs to \mathfrak{X} ?

10.2. Structure prediction

Prediction of protein structures is the most important method to reveal proteins' functions and working mechanics, it becomes a bottle neck in the rapidly developing life science. With more and more powerful computers, this problem is attacked in full front. Various models, homologous or *ab initio*, full atom model or coarse grained, with numerous parameters of which many are quite arbitrary, are used to achieve the goal. Although our computer power grows exponentially, prediction power does not follow that way. At this moment, one should take a deep breath and remind what the great physicist Fermi said: "There are two ways of doing calculations in theoretical physics. One way, and this is the way I prefer, is to have a clear physical picture of the process that you are calculating. The other way is to have a precise and self consistent mathematical formalism." And "I remember my friend Johnny von Neumann used to say, with four parameters I can fit an elephant, and with five I can make him wiggle his trunk." Quoted from Dyson (2004).

These remarks should also apply to any scientific calculation, not just theoretical physics. Look at the current situation, all *ab initio* prediction models are actually just empirical with many parameters to ensure some success. Fermi's comments remind us that a theory should be based on fundamental physical laws, and contain no arbitrary parameters. Look at formulas (1) and (2), one sees immediately that they are neat, precise and self consistent mathematical formulas. Furthermore, they including no arbitrary parameter, all terms in them have clear physical meanings. Chemical potentials μ_e and μ_i 's, geometric constants v_e and v_i 's, can be evaluated by theory or experiments, they are not arbitrary at all.

But a theory has to be developed, tested, until justified or falsified. For interested researchers, the tasks are to determine the correct values of the chemical potentials in formula (1) and the geometric ratios v_e and v_i in formula (2). There are many estimates to them, but they are either for the solvent accessible surface area such as in Eisenberg and MaLachlan (1986) hence not suit to the experiment data as pointed out in Tuñón *et. al.*, or do not distinguish different hydrophobicity levels as in Tuñón *et. al.* (1992). To get the correct values of the chemical potentials and geometric constants, commonly used method of training with data can be employed, in which one can also test the formulas' ability of discriminating native and nearby compact non-native structures. After that, a direct test is to predict the native structure from the amino acid sequence of a protein by minimizing the following:

$$G(\mathbf{X}_N) = \inf_{\mathbf{X} \in \mathfrak{X}} G(\mathbf{X}). \quad (57)$$

This is the first time that a theoretically derived formula of the Gibbs free energy is available. Before this, all *ab initio* predictions are not really *ab initio*. A combined (theoretical and experimental) search for the values of chemical potentials will be the key for the success of the *ab initio* prediction of protein structure.

10.3. Gradient

With formula (2) as the Gibbs free energy, the minimization in equation (57) can be pursued by Newton's fastest descending method. To state the result, some definitions are necessary.

10.3.1. Molecular graphs

Given a molecule U , let V be the set of atoms in U and $N = |V|$ be the number of atoms and label the atoms as $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_N$. For $1 \leq i, j \leq N$, define $B_{ij} = n$ if atoms i and j are connected by a bond with valency n (one can imagine that n is not necessarily a whole number), if i and j do not form a bond, then $B_{ij} = 0$. The molecule formula of U in chemistry can be seen as a graph $G(U) = (V, E)$, where V acts as the vertex set of $G(U)$ and E is the edge set of $G(U)$. An edge in E is denoted by $\{i, j\}$. If two atoms \mathbf{a}_i and \mathbf{a}_j are connected by a covalent bond, i.e., $B_{ij} = n \geq 1$, then $\{i, j\} \in E$ is an edge. Call $G(U)$ the **molecular graph** of U . FIGURE 1 is a molecular graph if the side chain R consisting of only one atom, such as in the amino acid Glycine.

A graph G is connected if from any vertex v one can follow the edges in the graph to arrive any other vertex. If a graph is not connected, then it has several **connected components**, each is itself a connected graph. All molecular graphs are connected.

10.3.2. Rotatable bonds

Let $\mathbf{b} = \mathbf{a}_\alpha \mathbf{a}_\beta$ be a covalent bond in the molecule U connecting two atoms \mathbf{a}_α and \mathbf{a}_β . The bond \mathbf{b} is rotatable if and only if: 1. the valency of \mathbf{b} is not greater than 1; 2. in the molecular graph $G(U)$, if one deletes $\{\alpha, \beta\}$, the remaining graph $G(U) \setminus \{\alpha, \beta\} = (V, E \setminus \{\alpha, \beta\})$ has exactly two connected components and neither component has rotational symmetry around the bond \mathbf{b} .

10.3.3. Derivatives of $G(\mathbf{X})$

Let $\mathbf{x}_i = (x_i, y_i, z_i)$, write $\mathbf{F} = -\nabla_{\mathbf{x}_i} G(\mathbf{X}) = -(G_{x_i}, G_{y_i}, G_{z_i})(\mathbf{X})$. The calculation of $G_{x_i}(\mathbf{X})$, for example, is via Lie vector field induced by moving the atomic position \mathbf{x}_i . In fact, any infinitesimal change of structure \mathbf{X} will induce a Lie vector field $\vec{L} : \mathbf{X} \rightarrow \mathbb{R}^3$. For example, moving \mathbf{x}_i from \mathbf{x}_i to $\mathbf{x}_i + (\Delta x_i, 0, 0)$ while keep other nuclear center fixed, will induce $L_{x_i} : \mathbf{X} \rightarrow \mathbb{R}^3$, such that $\vec{L}_{x_i}(\mathbf{x}_i) = (1, 0, 0)$ and $\vec{L}_{x_i}(\mathbf{x}_j) = (0, 0, 0)$ for $j \neq i$. Similarly \vec{L}_{y_i} and \vec{L}_{z_i} can be described as well. Then write $G_{x_i} = G_{\vec{L}_{x_i}}$, etc. and

$$\nabla_{\mathbf{x}_i} G(\mathbf{X}) = (G_{\vec{L}_{x_i}}, G_{\vec{L}_{y_i}}, G_{\vec{L}_{z_i}})(\mathbf{X}), \quad (58)$$

Rotating around a covalent bond b_{ij} also induce a Lie vector field $L_{b_{ij}} : \mathbf{X} \rightarrow \mathbb{R}^3$. In fact if $\mathbf{a}_i \mathbf{a}_j$ form the covalent bond b_{ij} , then the bond axis is

$$\mathbf{b}_{ij} = \frac{\mathbf{x}_j - \mathbf{x}_i}{|\mathbf{x}_j - \mathbf{x}_i|}. \quad (59)$$

If \mathbf{b}_{ij} is rotatable, denoting all nuclear centers in one component by $R_{b_{ij}}$ and others in $F_{b_{ij}}$. One can rotate all centers in $R_{b_{ij}}$ around \mathbf{b}_{ij} for certain angle while keep all centers in $F_{b_{ij}}$ fixed. The induced Lie vector field $\vec{L}_{b_{ij}}$ will be

$$\vec{L}_{b_{ij}}(\mathbf{x}_k) = (\mathbf{x}_k - \mathbf{x}_i) \wedge \mathbf{b}_{ij}, \text{ if } \mathbf{x}_k \in R_{b_{ij}}; \quad (60)$$

$$\vec{L}_{b_{ij}}(\mathbf{x}_k) = \vec{0}, \text{ if } \mathbf{x}_k \in F_{b_{ij}}. \quad (61)$$

Any such a Lie vector field \vec{L} will generate a family of conformations $\mathbf{X}_t = (x_{1t}, \dots, x_{it}, \dots, x_{Mt})$, where $\mathbf{x}_{kt} = \mathbf{x}_k + t\vec{L}(\mathbf{x}_k)$, $k = 1, \dots, M$. Moreover, the Lie vector field \vec{L} can be generated to the molecular surface $M_{\mathbf{X}}$, as shown in Appendix A.

The derivative $G_{\vec{L}}(\mathbf{X})$ is given by

$$G_{\vec{L}}(\mathbf{X}) = v_e \mu_e V_{\vec{L}}(\Omega_{\mathbf{X}}) + d_w v_e \mu_e A_{\vec{L}}(M_{\mathbf{X}}) + \sum_{i=1}^H v_i \mu_i A_{\vec{L}}(M_{\mathbf{X}_i}), \quad (62)$$

with

$$V_{\vec{L}}(\Omega_{\mathbf{X}}) = - \int_{M_{\mathbf{X}}} \vec{L} \cdot \vec{N} d\mathcal{H}^2, \quad A_{\vec{L}}(M_{\mathbf{X}}) = -2 \int_{M_{\mathbf{X}}} H(\vec{L} \cdot \vec{N}) d\mathcal{H}^2, \quad (63)$$

where \vec{N} is the outer unit normal of $M_{\mathbf{X}}$, H the mean curvature of $M_{\mathbf{X}}$, and \mathcal{H}^2 the Hausdorff measure. Define $f_{t,i} : \mathbb{R}^3 \rightarrow \mathbb{R}$ as $f_{t,i}(\mathbf{x}) = \text{dist}(\mathbf{x}, M_{\mathbf{X}_i}) - \text{dist}(\mathbf{x}, M_{\mathbf{X}_i} \setminus M_{\mathbf{X}_i})$, and define on $M_{\mathbf{X}}$

$$\nabla_{M_{\mathbf{X}}} f_{0,i} = \nabla f_{0,i} - (\nabla f_{0,i} \cdot \vec{N}) \vec{N}, \quad f'_{0,i} = \left. \frac{\partial f_{t,i}}{\partial t} \right|_{t=0}, \quad \frac{df_{0,i}}{dt} = \vec{L} \cdot \nabla f_{0,i} + f'_{0,i}, \quad (64)$$

then let $\vec{\eta}$ be the unit outward conormal vector of $\partial M_{\mathbf{X}_i}$ (normal to $\partial M_{\mathbf{X}_i}$ but tangent to $M_{\mathbf{X}}$),

$$A_{\vec{L}}(M_{\mathbf{X}_i}) = -2 \int_{M_{\mathbf{X}_i}} H(\vec{L} \cdot \vec{N}) d\mathcal{H}^2 + \int_{\partial M_{\mathbf{X}_i}} \left[\vec{L} \cdot \vec{\eta} - \frac{\frac{df_{0,i}}{dt}}{|\nabla_{M_{\mathbf{X}}} f_{0,i}|} \right] d\mathcal{H}^1. \quad (65)$$

The \mathbf{X}_t is all the information needed in calculating the molecular surface $M_{\mathbf{X}_t}$, see Connolly (1983). To calculate, the above formulas have to be translated into formulas on the molecular surface $M_{\mathbf{X}}$. These translations are given in Appendix A, they are calculable (all integrals are integrable, i.e., can be expressed by analytic formulas with variables \mathbf{X}) and were calculated piecewise on $M_{\mathbf{X}}$.

10.3.4. The gradient

Let a protein \mathcal{U} have L rotatable bonds $(\mathbf{b}_1, \dots, \mathbf{b}_i, \dots, \mathbf{b}_L)$. Let θ_i denote the dihedral angle around the rotatable bond \mathbf{b}_i . A conformation \mathbf{X} of \mathcal{U} can be expressed in terms of these rotatable dihedral angles $\Theta = (\theta_1, \dots, \theta_i, \dots, \theta_L)$, then

$$G(\mathbf{X}) = G(\Theta), \quad (66)$$

and the gradient of G can be written as

$$\nabla G(\Theta) = \left(\frac{\partial G}{\partial \theta_1}, \dots, \frac{\partial G}{\partial \theta_i}, \dots, \frac{\partial G}{\partial \theta_L} \right) (\Theta) = (G_{\bar{L}_{b_1}}, \dots, G_{\bar{L}_{b_i}}, \dots, G_{\bar{L}_{b_L}})(\mathbf{X}). \quad (67)$$

If the rotation around \mathbf{b}_i with rotating angle $-sG_{\bar{L}_{b_i}}(\mathbf{X})$ on R_{b_i} and fix atoms in F_{b_i} be denoted as M_i , new conformation $\mathbf{Y}_s = M_L \circ M_{L-1} \circ \dots \circ M_1$ will be obtained, where $s > 0$ is a suitable step length. That is to say, the dihedral angles of \mathbf{Y}_s are

$$[\theta_1 - sG_{\bar{L}_{b_1}}(\mathbf{X}), \dots, \theta_i - sG_{\bar{L}_{b_i}}(\mathbf{X}), \dots, \theta_L - sG_{\bar{L}_{b_L}}(\mathbf{X})].$$

The order of rotations in fact is irrelevant, i.e., by any order, the same conformation \mathbf{Y}_s will always be obtained, as proved in Fang and Jing (2008) and Appendix A. This way one can fast change the structure by simultaneous rotate around all rotatable bonds.

This actually is the Newton's fastest descending method, it reduces the Gibbs free energy $G(\mathbf{X})$ most efficiently. Afore mentioned simulations of Fang and Jing (2010) used this method.

10.4. Kinetics

There are evidence that some protein's native structure is not the global minimum of the Gibbs free energy, but only a local minimum. If the native structure of a protein achieves the global minimal value of the Gibbs free energy, the folding process is **thermodynamic**; if it is only a local minimum, the folding process is **kinetic**, Lazaridis and Karplus (2003).

With the formula (2) and the gradient just obtained, one actually has the kinetic in hand. In fact, for any atomic position \mathbf{x}_i , the kinetic force is $\mathbf{F}_i(\mathbf{X}) = -\nabla_{\mathbf{x}_i} G(\mathbf{X})$, Dai (2007). With formula (2) these quantities are readily calculable as mentioned above. The resulting Newton's fastest descending method will help us find the native structure, either in the thermodynamic case or in the kinetic case, here the thermodynamic and kinetic cases are combined by the Gibbs free energy formula (2) and its derivatives.

The moving along $-\nabla G$ method was used in the simulation in Fang and Jing (2010).

11. Conclusion

A quantum statistical theory of protein folding for monomeric, single domain, self folding globular proteins is suggested. The assumptions of the theory fit all observed realities of protein folding. The resulting formulas (1) and (2) do not have any arbitrary parameters and all terms in them have clear physical meaning. Potential energies involving pairwise interactions between atoms do not appear in them.

Formulas (1) and (2) have explanation powers. They give unified explanation to folding and denaturation, to the hydrophobic effect in protein folding and its relation with the hydrogen bonding. The formulas also explain the relative successes of surface area protein folding models. Relation between kinetic and thermodynamic of protein folding is discussed, driving force formula comes from the Gibbs free energy formula (2) are also given. Energy surface theory will be much easier to handle. The concept of ΔG is clarified.

Appendix

A. Calculations on the molecular surface

A.1. Rotation order

Let $P_X = \cup_{i=1}^N B(\mathbf{x}_i, r_i)$ and $\mathbf{x}_{\alpha 0}$ and $\mathbf{x}_{\alpha 1}$ be bonded by b_α , the rotation line of b_α is $\mathbf{x}_{\alpha 0} + t \frac{\mathbf{x}_{\alpha 1} - \mathbf{x}_{\alpha 0}}{|\mathbf{x}_{\alpha 1} - \mathbf{x}_{\alpha 0}|} = \mathbf{x}_{\alpha 0} + t \mathbf{b}_\alpha$. Each b_α divides $\{\mathbf{x}_1, \dots, \mathbf{x}_M\}$ into two groups F_α and R_α , balls in R_α will be rotated while balls in F_α will be fixed. Note that these partitions are independent of P_X , they only depend on the molecular graph of the protein molecule. Let M_α be this rotation-fixation, it will be shown that

$$M_\alpha \circ M_\beta(\mathbf{X}) = M_\beta \circ M_\alpha(\mathbf{X}), \mathbf{X} \in (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M), 1 \leq \alpha, \beta \leq L. \quad (1)$$

The formula of rotating a point X around a line $L: \mathbf{y} = \mathbf{x} + t\mathbf{b}$ ($|\mathbf{b}| = 1$) by an angle ω is $R(X) = \mathbf{x} + \mathbf{A}(\omega)(X - \mathbf{x})$. Let I be the identity matrix, $B = \mathbf{b}\mathbf{b}^T$ and Z_b the matrix such that the outer product $\mathbf{b} \wedge X = Z_b X$, then the orthonormal matrix $\mathbf{A}(\omega) = (1 - \cos \omega)B + \cos \omega I + \sin \omega Z_b$.

The topology of a protein molecule guarantees that if two bonds b_α and b_β such that $R_\alpha \subset R_\beta$, then $\{\mathbf{x}_{\alpha 0}, \mathbf{x}_{\alpha 1}\} \subset R_\beta$. Let b_1 and b_2 be two bonds and $L_1: \mathbf{x} = \mathbf{x}_1 + t\mathbf{b}_1$ and $L_2: \mathbf{x} = \mathbf{x}_2 + t\mathbf{b}_2$ be the two rotating lines and $X \in (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$. To prove equation (1), there are only two cases to consider: $R_1 \subset R_2$ and $R_1 \cap R_2 = \emptyset$. In any case, if $X \in F_1 \cap F_2$, then $M_1 \circ M_2(X) = M_2 \circ M_1(X) = X$. If $X \in R_1 \subset R_2$, then

$$M_2 \circ M_1(X) = \mathbf{x}_2 + \mathbf{A}_2(\omega_2)(\mathbf{x}_1 - \mathbf{x}_2) + \mathbf{A}_2(\omega_2)\mathbf{A}_1(\omega_1)(X - \mathbf{x}_1). \quad (2)$$

On the other hand b_1 and hence L_1 itself will be rotated by M_2 , $L_3 = M_2(L_1) = \mathbf{x}_3 + t\mathbf{b}_3$, where $\mathbf{x}_3 = \mathbf{x}_2 + \mathbf{A}_2(\omega_2)(\mathbf{x}_1 - \mathbf{x}_2)$, $\mathbf{b}_3 = \mathbf{A}_2(\omega_2)\mathbf{b}_1$. Since $X \in R_1 \subset R_2$ and $M_2(X) \in R_1$ (in the new conformation $M_2(P)$ where rotation around b_1 is rotation around L_3), $M_1 \circ M_2(X)$ will be the rotation R_3 around L_3 of $M_2(X)$ by angle ω_1 , thus

$$M_1 \circ M_2(X) = \mathbf{x}_2 + \mathbf{A}_2(\omega_2)(\mathbf{x}_1 - \mathbf{x}_2) + \mathbf{A}_3(\omega_1)\mathbf{A}_2(\omega_2)(X - \mathbf{x}_1). \quad (3)$$

Let $\mathbf{v} \in \mathbb{R}^3$ be an arbitrary vector, writing $\mathbf{A}_1(\omega_1) = \mathbf{A}_1$, $\mathbf{A}_2(\omega_2) = \mathbf{A}_2$, and $\mathbf{A}_3(\omega_1) = \mathbf{A}_3$, then

$$\mathbf{A}_2\mathbf{A}_1\mathbf{v} = (1 - \cos \omega_1)(\mathbf{b}_1 \bullet \mathbf{v})\mathbf{A}_2\mathbf{b}_1 + \cos \omega_1\mathbf{A}_2\mathbf{v} + \sin \omega_1\mathbf{A}_2(\mathbf{b}_1 \wedge \mathbf{v}). \quad (4)$$

For any orthonormal matrix O , $(O\mathbf{b}_1) \bullet (O\mathbf{v}) = \mathbf{b}_1 \bullet \mathbf{v}$, $O(\mathbf{b}_1 \wedge \mathbf{v}) = (O\mathbf{b}_1) \wedge (O\mathbf{v})$. Then by $\mathbf{b}_3 = \mathbf{A}_2(\omega_2)\mathbf{b}_1$,

$$\begin{aligned} \mathbf{A}_3\mathbf{A}_2\mathbf{v} &= (1 - \cos \omega_1)[\mathbf{b}_3 \bullet (\mathbf{A}_2\mathbf{v})]\mathbf{b}_3 + \cos \omega_1(\mathbf{A}_2\mathbf{v}) + \sin \omega_1\mathbf{b}_3 \wedge (\mathbf{A}_2\mathbf{v}) \\ &= (1 - \cos \omega_1)(\mathbf{b}_1 \bullet \mathbf{v})\mathbf{A}_2\mathbf{b}_1 + \cos \omega_1\mathbf{A}_2\mathbf{v} + \sin \omega_1\mathbf{A}_2(\mathbf{b}_1 \wedge \mathbf{v}). \end{aligned} \quad (5)$$

Since \mathbf{v} was arbitrary, equations (2) to (5) show equation (1) is true.

If $R_1 \cap R_2 = \emptyset$ and $X \in R_2$, then X and $M_2(X) \in F_1$ hence $M_1 \circ M_2(X) = M_2(X) = M_2 \circ M_1(X)$.

The molecular surface is consisted of faces. Thus all integrals can be integrated piecewise on faces. There are three kinds of faces, convex, concave, and saddle, Connolly (1983). The formulas on each kind of face are given below. The notation $\dot{\mathbf{x}}$ means $L(\mathbf{x})$ with L the corresponding Lie vector field. All van der Waals radii r_i , as well as the probe radius r_p , are constants.

A.2. Convex face

A convex face is a piece of spherical surface lying on some $S_i = \partial B(\mathbf{x}_i, r_i)$ and bounded by circular arcs $\gamma_\nu, \nu = 1, \dots, n_F$, let \mathbf{v}_ν^0 and \mathbf{v}_ν^1 be γ_ν 's vertices and \mathbf{c}_ν and r_ν the center and radius of γ_ν 's circle, $r_\nu \phi_\nu$ the arc length of γ_ν , $\mathbf{e}_3^\nu = (z_{\nu 1}, z_{\nu 2}, z_{\nu 3})$ be the unit vector in the direction of $(\mathbf{v}_\nu^0 - \mathbf{c}_\nu) \wedge (\mathbf{v}_\nu^1 - \mathbf{c}_\nu)$, $d_\nu = \mathbf{e}_3^\nu \bullet (\mathbf{c}_\nu - \mathbf{x}_i)$, $\mathbf{e}_1^\nu = \frac{\mathbf{v}_\nu^0 - \mathbf{c}_{i\nu}}{r_\nu} = (x_{\nu 1}, x_{\nu 2}, x_{\nu 3})$, $\mathbf{e}_2^\nu = \mathbf{e}_3^\nu \wedge \mathbf{e}_1^\nu = (y_{\nu 1}, y_{\nu 2}, y_{\nu 3})$, $1 \leq \nu \leq n_F$. A point \mathbf{x} on F has the form $\mathbf{x} = \mathbf{x}_i - r_i \mathbf{N}$ and $X_\alpha(\mathbf{x}) = \dot{\mathbf{x}} - r_i \dot{\mathbf{N}}$, by $\mathbf{N} \bullet \dot{\mathbf{N}} \equiv 0$ and the general divergence formula on sphere,

$$r_i \int_F (X_\alpha \bullet \mathbf{N}) H d\mathcal{H}^2 = \int_F X_\alpha \bullet \mathbf{N} d\mathcal{H}^2 = \frac{-1}{r_i} \dot{\mathbf{x}}_i \bullet \sum_{\nu=1}^M (X_\nu, Y_\nu, Z_\nu), \quad (6)$$

where

$$X_\nu = \frac{r_\nu^2}{2} \{ \phi_\nu z_{\nu 1} + \sin \phi_\nu [\cos \phi_\nu (x_{\nu 2} y_{\nu 3} + x_{\nu 3} y_{\nu 2}) + \sin \phi_\nu (y_{\nu 2} y_{\nu 3} - x_{\nu 2} x_{\nu 3})] \\ + r_\nu d_\nu z_{\nu 2} [y_{\nu 3} \sin \phi_\nu - x_{\nu 3} (1 - \cos \phi_\nu)] \}, \quad (7)$$

$$Y_\nu = \frac{r_\nu^2}{2} \{ \phi_\nu z_{\nu 2} + \sin \phi_\nu [\cos \phi_\nu (x_{\nu 3} y_{\nu 1} + x_{\nu 1} y_{\nu 3}) + \sin \phi_\nu (y_{\nu 1} y_{\nu 3} - x_{\nu 1} x_{\nu 3})] \\ + r_\nu d_\nu z_{\nu 3} [y_{\nu 1} \sin \phi_\nu - x_{\nu 1} (1 - \cos \phi_\nu)] \}, \quad (8)$$

$$Z_\nu = \frac{r_\nu^2}{2} \{ \phi_\nu z_{\nu 3} + \sin \phi_\nu [\cos \phi_\nu (x_{\nu 1} y_{\nu 2} + x_{\nu 2} y_{\nu 1}) + \sin \phi_\nu (y_{\nu 1} y_{\nu 2} - x_{\nu 1} x_{\nu 2})] \\ + r_\nu d_\nu z_{\nu 1} [y_{\nu 2} \sin \phi_\nu - x_{\nu 2} (1 - \cos \phi_\nu)] \}. \quad (9)$$

A.3. Concave face

A concave face F is a spherical polygon on the probe sphere S when S is simultaneously tangent to 3 balls $B(\mathbf{x}_i, r_i)$, $1 \leq i \leq 3$. F is expressed by parameters $t_i \geq 0$, $i = 1, 2, 3$,

$$\mathbf{x} = \mathbf{p} + r\mathbf{N} = \mathbf{p} + r \frac{t_1 \mathbf{x}_1 + t_2 \mathbf{x}_2 + t_3 \mathbf{x}_3 - \mathbf{p}}{|t_1 \mathbf{x}_1 + t_2 \mathbf{x}_2 + t_3 \mathbf{x}_3 - \mathbf{p}|}, \quad t_1 + t_2 + t_3 = 1. \quad (10)$$

$$\phi_t(\mathbf{x}) = \mathbf{p}(t) + r\mathbf{N}(t) = \mathbf{p}(t) + r \frac{t_1 \mathbf{x}_1(t) + t_2 \mathbf{x}_2(t) + t_3 \mathbf{x}_3(t) - \mathbf{p}(t)}{|t_1 \mathbf{x}_1(t) + t_2 \mathbf{x}_2(t) + t_3 \mathbf{x}_3(t) - \mathbf{p}(t)|}, \quad (11)$$

$$X_\alpha(\mathbf{x}) = \left. \frac{d\phi_t(\mathbf{x})}{dt} \right|_{t=0} = \dot{\mathbf{p}} + r\dot{N}. \text{ Using } |\mathbf{p}(t) - \mathbf{x}_i(t)| = r_i + r = \text{constant, let } b_i = (\mathbf{x}_i - \mathbf{p}) \bullet \dot{\mathbf{x}}_i, \\ \mathbf{b} = (b_1, b_2, b_3)^T, A = \begin{pmatrix} \mathbf{x}_1 - \mathbf{p} \\ \mathbf{x}_2 - \mathbf{p} \\ \mathbf{x}_3 - \mathbf{p} \end{pmatrix}, \text{ then } \det A \neq 0, \dot{\mathbf{p}} = A^{-1}\mathbf{b}. \text{ By } X_\alpha \bullet N = \dot{\mathbf{p}} \bullet N,$$

$$r \int_F (X_\alpha \bullet N) H d\mathcal{H}^2 = - \int_F X_\alpha \bullet N d\mathcal{H}^2 = \frac{1}{r} \dot{\mathbf{p}} \bullet \sum_{i=1}^3 (X_i, Y_i, Z_i). \quad (12)$$

Here the $X_i, Y_i,$ and Z_i are the same as in equations (7) to (9).

Assume that \mathbf{x}_1 has different water association with \mathbf{x}_2 and \mathbf{x}_3 , let $R_i = r_i + r, d_{ij} = |\mathbf{x}_i - \mathbf{x}_j|, y_{ij} = (R_i^2 - R_j^2)/2d_{ij}$. Then $f_P(\mathbf{x}) = (\mathbf{x} - \mathbf{p}) \bullet \mathbf{n}_k$, where $\mathbf{n}_k = (\mathbf{x}_k - \mathbf{x}_1)/d_{1k}$ is the directed unit normal of the dividing plane P_k (passing through \mathbf{p} and $\mathbf{t}_{1k} = \frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_k) + y_{1k}\mathbf{n}_k$ and perpendicular to it), $k = 2, 3$. The projection of $\partial W \cap F$ on the $\mathbf{x}_1\mathbf{x}_2\mathbf{x}_3$ plane is in the form of one or two curves $\gamma_k, (\{j, k\} = \{2, 3\})$

$$t_k = A_k t_j + B_k, 0 \leq t_j \leq z_j, A_k = \frac{d_{1j} \cos \omega}{-d_{1k}}, B_k = \frac{d_{1k} + 2y_{1k}}{2d_{1k}}, \quad (13)$$

where $\cos \omega = \frac{(\mathbf{x}_2 - \mathbf{x}_1) \bullet (\mathbf{x}_3 - \mathbf{x}_1)}{d_{12}d_{13}}$. $F \cap W$ is a spherical polygon with arcs $\gamma_\nu, 1 \leq \nu \leq n$, including some γ_k as above, so $\int_{W \cap F} (X_\alpha \bullet N) H d\mathcal{H}^2$ has the similar form as that in equation (12).

Let $\mathbf{A}_k = \mathbf{x}_j - \mathbf{x}_1 + A_k(\mathbf{x}_k - \mathbf{x}_1), \mathbf{B}_k = B_k(\mathbf{x}_k - \mathbf{x}_1) + (\mathbf{x}_1 - \mathbf{p}), \mathbf{C}_k = \mathbf{B}_k \wedge \mathbf{A}_k$. Treating A_k and B_k as constants and setting $H_k = \dot{\mathbf{p}} \bullet \mathbf{C}_k, J_k = \dot{\mathbf{A}}_k \bullet \mathbf{C}_k,$ and $K_k = \dot{\mathbf{B}}_k \bullet \mathbf{C}_k$. Let $a_k t_j^2 + b_k t_j + c_k = |\mathbf{A}_k t_j + \mathbf{B}_k|^2 > 0$, then $\Delta_k = 4a_k c_k - b_k^2 > 0$. By $\eta = N'_{t_j} \wedge N / |N'_{t_j}|$ and $d\mathcal{H}^1 = r |N'_{t_j}| dt_j,$

$$\int_{(\partial W \cap F) \cap \gamma_k} X_\alpha \bullet \eta d\mathcal{H}^1 = \frac{2rH_k}{\sqrt{\Delta_k}} \left(\arctan \frac{2a_k z_j + b_k}{\sqrt{\Delta_k}} - \arctan \frac{b_k}{\sqrt{\Delta_k}} \right) \\ + \frac{2r^2 J_k}{\Delta_k} \left(2\sqrt{c_k} - \frac{b_k z_j + 2c_k}{\sqrt{a_k z_j^2 + b_k z_j + c_k}} \right) \\ + \frac{2r^2 K_k}{\Delta_k} \left(\frac{2a_k z_j + b_k}{\sqrt{a_k z_j^2 + b_k z_j + c_k}} - \frac{b_k}{\sqrt{c_k}} \right). \quad (14)$$

Let $U_k = (\mathbf{A}_k \bullet \mathbf{n}_k), V_k = (\mathbf{B}_k \bullet \mathbf{n}_k), W_k = |\mathbf{C}_k \bullet \mathbf{n}_k| > 0$, then

$$\int_{\gamma_k} \frac{\frac{df_P}{dt}}{|\nabla_{M_P} f_P|} d\mathcal{H}^1 = \frac{\pm 2r^2}{W_k} \left(\frac{(2a_k z_j + b_k)V_k}{\sqrt{a_k z_j^2 + b_k z_j + c_k}} - \frac{b_k V_k}{\sqrt{c_k}} \right. \\ \left. + 2\sqrt{c_k} U_k - \frac{(b_k z_j + 2c_k)U_k}{\sqrt{a_k z_j^2 + b_k z_j + c_k}} \right), \quad (15)$$

where the sign is determined by orientation.

A.4. Saddle face

A saddle face F is generated when the probe S simultaneously tangents to two balls $B(\mathbf{x}_1, r_1)$ and $B(\mathbf{x}_2, r_2)$, and rolls around the axis $\mathbf{e}_2 = \frac{\mathbf{x}_2 - \mathbf{x}_1}{d_{12}}$. The starting and stopping positions of the probe center is \mathbf{p} and \mathbf{q} . Let $y = [(r_1 + r)^2 - (r_2 + r)^2]/2d_{12}$ and $\mathbf{t} = \frac{1}{2}(\mathbf{x}_1 + \mathbf{x}_2) + y\mathbf{e}_2$, $R = |\mathbf{p} - \mathbf{t}|$, $\mathbf{e}_1 = (\mathbf{p} - \mathbf{t})/R$, $\mathbf{e}_3 = \mathbf{e}_1 \wedge \mathbf{e}_2$, then F is parametrized by $0 \leq \psi \leq \psi_s$, $\theta_1 \leq \theta \leq \theta_2$,

$$\mathbf{x}(\psi, \theta) = \mathbf{t} + (R - r \cos \theta)(\cos \psi \mathbf{e}_1 + \sin \psi \mathbf{e}_3) + r \sin \theta \mathbf{e}_2, \quad (16)$$

where let $\omega_s = \arccos[(\mathbf{p} - \mathbf{t}) \bullet (\mathbf{q} - \mathbf{t})/R^2]$, then $\psi_s = \omega_s$ or $2\pi - \omega_s$. $\theta_1 = \arctan[-(d_{12} + 2y)/2R]$, $\theta_2 = \arctan[(d_{12} - 2y)/2R]$. These data are uniquely determined by the conformation P , see Connolly (1983). Let $\theta_k(t)$ and $\phi_s(t)$ be similarly defined for the conformation P_t , one can define $\phi_t(\psi) = \frac{\psi \psi_s(t)}{\psi_s}$, $\phi_t(\theta) = \frac{\theta_1(t)(\theta_2 - \theta) + \theta_2(t)(\theta - \theta_2)}{\theta_2 - \theta_1}$, and $U(\psi) = \cos \psi \mathbf{e}_1 + \sin \psi \mathbf{e}_3$, then for the same $0 \leq \phi \leq \phi_s$ and $\theta_1 \leq \theta \leq \theta_2$,

$$\phi_t(\mathbf{x}) = \mathbf{t}(t) + [R - r \cos \phi_t(\theta)]U(\phi_t(\psi)) + r \sin \phi_t(\theta) \mathbf{e}_2(t). \quad (17)$$

Let $\dot{U} = \cos \psi \dot{\mathbf{e}}_1 + \sin \psi \dot{\mathbf{e}}_3$, $U' = -\sin \psi \mathbf{e}_1 + \cos \psi \mathbf{e}_3$, then

$$\begin{aligned} X_\alpha(\mathbf{x}) &= \dot{\mathbf{t}} + (\dot{R} + r\dot{\theta} \sin \theta)U + (R - r \cos \theta)(\dot{U} + \psi U') \\ &\quad + r\dot{\theta} \cos \theta \mathbf{e}_2 + r \sin \theta \dot{\mathbf{e}}_2. \end{aligned} \quad (18)$$

On F , $N = -\cos \theta U(\psi) + \sin \theta \mathbf{e}_2$, $d\mathcal{H}^2 = r(R - r \cos \theta)d\theta d\psi$, $2H = \frac{R - 2r \cos \theta}{r(R - r \cos \theta)}$. Let $J = \int_0^{\psi_s} U(\psi) d\psi$, then

$$\begin{aligned} 4 \int_F X_\alpha \bullet N d\mathcal{H}^2 &= 4rR(\phi_s \dot{\mathbf{t}} \bullet \mathbf{e}_2 - RJ \bullet \dot{\mathbf{e}}_2)(\cos \theta_1 - \cos \theta_2) \\ &\quad + 4rR(\phi_s \dot{R} + \dot{\mathbf{t}} \bullet J)(\sin \theta_1 - \sin \theta_2) \\ &\quad - r^2(\phi_s \dot{\mathbf{t}} \bullet \mathbf{e}_2 - RJ \bullet \dot{\mathbf{e}}_2)(\cos 2\theta_1 - \cos 2\theta_2) \\ &\quad + r^2(\phi_s \dot{R} + \dot{\mathbf{t}} \bullet J)[2(\theta_1 - \theta_2) + \sin 2\theta_1 - \sin 2\theta_2], \end{aligned} \quad (19)$$

$$\begin{aligned} 2 \int_F (X_\alpha \bullet N) H d\mathcal{H}^2 &= 2R(\phi_s \dot{\mathbf{t}} \bullet \mathbf{e}_2 - R \bullet \dot{\mathbf{e}}_2)(\cos \theta_1 - \cos \theta_2) \\ &\quad + 2R(\phi_s \dot{R} + \dot{\mathbf{t}} \bullet J)(\sin \theta_1 - \sin \theta_2) \\ &\quad - r(\phi_s \dot{\mathbf{t}} \bullet \mathbf{e}_2 - RJ \bullet \dot{\mathbf{e}}_2)(\cos 2\theta_1 - \cos 2\theta_2) \\ &\quad + r(\phi_s \dot{R} + \dot{\mathbf{t}} \bullet J)[2(\theta_1 - \theta_2) + \sin 2\theta_1 - \sin 2\theta_2]. \end{aligned} \quad (20)$$

Assume that \mathbf{x}_1 is hydrophobic and \mathbf{x}_2 is not, then the dividing plane P passing through \mathbf{p} and \mathbf{t} and is perpendicular to \mathbf{e}_2 . The curve $\partial W \cap F$ is given by $\mathbf{x}(\psi) = \mathbf{t} + (R - r)U(\psi)$, $0 \leq \psi \leq \psi_s$, on which $d\mathcal{H}^1 = (R - r)d\psi$. The hydrophobic surface integral on F then is the same as in equation (20), except $\theta_1 = 0$. Since on $\partial W \cap F$, $\eta = N' \wedge N = \mathbf{e}_2$, $\frac{d\theta(t)}{dt}|_{t=0} = \dot{\theta}_0 = \frac{\dot{\theta}_1 \theta_2 - \dot{\theta}_2 \theta_1}{\theta_2 - \theta_1}$, by equation (18),

$$\int_{\partial W \cap F} X_\alpha \bullet \eta d\mathcal{H}^1 = (R - r)\phi_s(r\dot{\theta}_0 + \dot{\mathbf{t}} \bullet \mathbf{e}_2) - (R - r)^2 \dot{\mathbf{e}}_2 \bullet J, \quad (21)$$

Let $\mathbf{n}_j = \mathbf{e}_2$, then $f_{P_t}(\phi_t(\mathbf{x})) = [\phi_t(\mathbf{x}) - \mathbf{t}(t)] \bullet \mathbf{n}_j(t)$, $|\nabla_{M_P} f_P| = \mathbf{n}_j \bullet \boldsymbol{\eta} = 1$, and $\left. \frac{df_{P_t}(\phi_t(X))}{dt} \right|_{t=0} = [\phi_t(X) - \mathbf{t}(t)] \bullet \mathbf{e}_2 + [(R - r)U] \bullet \dot{\mathbf{e}}_2 = r\dot{\theta}_0$.

$$\int_{\partial W_P \cap F} \frac{\frac{df_P}{dt}}{|\nabla_{M_P} f_P|} d\mathcal{H}^1 = (R - r)\phi_s r\dot{\theta}_0. \quad (22)$$

Author details

Yi Fang

Department of Mathematics, Nanchang University, 999 Xuefu Road, Honggutan New District, Nanchang, 330031, China

12. References

- [1] Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science* 181, 223-230.
- [2] Bader, R. F. W. (1990) *Atoms in Molecules: A Quantum Theory*. (Clarendon Press · Oxford).
- [3] Baily, M. (1994) *A Survey of Thermodynamics*. American Institute of Physics New York.
- [4] Ben-Naim, A. (2012) Levinthal's question revisited, and answered. *Journal of Biomolecular Structure and Dynamics* 30(1), 113-124 (2012).
- [5] Branden, C. and J. Tooze, J. (1999) *Introduction to Protein Structure*. (Second Edition, Garland).
- [6] Connolly, M. L. (1983) Analytical molecular surface calculation. *J. Appl. Cryst.*, 16:548-558.
- [7] Dai, X. (2007) *Advanced Statistical Physics*. (Fudan University Press, Shanghai).
- [8] Dill, K. A. (1990) Dominant forces in protein folding. *Biochemistry*, 29 7133-7155.
- [9] Dyson, F. (2004) A meeting with Enrico Fermi: How an intuitive physicist rescued a team from fruitless research. *Nature* 427, 297.
- [10] Eisenberg, D and McLachlan, A. D. Solvation energy in protein folding and binding. *Nature* 319, 199-203.
- [11] ExPASy Proteomics Server. <http://au.expasy.org/sprot/relnotes/relstat.html>
- [12] Fang, Y. (2005) Mathematical protein folding problem. In: D. Hoffman, Ed, *Global Theory of Minimal Surfaces. (Proceedings of the Clay Mathematical Proceedings, 2 2005)* pp. 611-622.
- [13] Fang, Y. and Jing, J. (2008) Implementation of a mathematical protein folding model. *International Journal of Pure and Applied Mathematics*, 42(4), 481-488.
- [14] Fang, Y. and Jing, J. (2010) Geometry, thermodynamics, and protein. *Journal of Theoretical Biology*, 262, 382-390.
- [15] Finkelstein, A. V. and Ptitsyn, O. B. (2002) *Protein Physics: A Course of Lectures*. Academic Press, An imprint of Elsevier Science, Amsterdam
- [16] Greiner, W., Neise, L., and Stöker, H. (1994) *Thermodynamics and Statistical Mechanics*. (Springer-Verlag, New York, Berlin, ...).
- [17] Jackson, R. M., and Sternberg, M. J. E. (1993) Protein surface area defined. *Nature* 366, 638.
- [18] Kang, H. J. and Baker, E. N. (2011) Intramolecular isopeptide bonds: protein crosslinks built for stress? *TIBS*, 36(4), 229-237.

- [19] Kauzmann, W. (1959) Some factors in the interpretation of protein denaturation. *Adv. Protein Chem.* 14, 1-63 (1959).
- [20] Lazaridis, T., and Karplus, M. (2003) Thermodynamics of protein folding: a microscopic view. *Biophysical Chemistry*, 100, 367-395.
- [21] Lee, B., and Richards, F. M. (1971) The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* 55:379-400.
- [22] Novotny, J., Bruccoleri, R., and Karplus, M. (1984) An analysis of incorrectly folded protein models. Implications for structure predictions. *J. Mol. Biol.* 177, 787-818.
- [23] Popelier, P. (2000) *Atoms in Molecules: An Introduction*. (Prentice Hall).
- [24] Reynolds, J. A., Gilbert, D. B., and Tanford, C. (1974) Empirical correlation between hydrophobic free energy and aqueous cavity surface area. *Proc. Natl. Acad. Sci. USA* 71(8), 2925-2927.
- [25] Richards, F. M. (1977) Areas, volumes, packing, and protein structure. *Ann. Rev. Biophys. Bioeng.* 6, 151-176.
- [26] Rose, G. D., Fleming, P. J., Banavar, J. R., and Maritan, A. (2006) A backbone based theory of protein folding. *PNAS* 103(45), 16623-16633.
- [27] Tuñón, I., Silla, E., and Pascual-Ahuir, J. L. (1992) Molecular surface area and hydrophobic effect. *Protein Engineering* 5(8), 715-716.
- [28] Wang, Y., Zhang, H., Li, W., and Scott, R. A. (1995) Discriminating compact nonnative structures from the native structure of globular proteins. *PNAS*, 92, 709-713.
- [29] Wu, H. (1931) Studies on denaturation of proteins XIII. A theory of denaturation. *Chinese Journal of Physiology* 4, 321-344 (1931). A preliminary report was read before the XIIIth International Congress of Physiology at Boston, Aug. 19-24, 1929 and published in the *Am. J. Physiol.* for Oct. 1929. Reprinted in *Advances in Protein Chemistry* 46, 6-26 (1995).