

Monte Carlo simulations for protein denaturation

Author: Neus Pagès Gelabert

Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.*

Advisor: Giancarlo Franzese

Abstract: Proteins are usually in a water solution when they perform their biological function. The interaction between the protein and water is fundamental to understand the protein stability. In this report, we test by Monte Carlo simulations the role played in the context of protein stability by the changes of water compressibility at the interface with the protein.

I. INTRODUCTION

Proteins have a primary structure that describes the linear sequence of amino acids in the polypeptide chain. Within this string there are regions in which the chains are organised into regular moieties that give rise to secondary structure. The tertiary structure is a description of the way the whole string folds into its final three-dimensional shape. Proteins can perform their biological function only if they acquire the correct tertiary structure[1][2].

In biological environments, proteins fold in water solution. Water molecules have a peculiar ability to form hydrogen bonds. This happens because their electronegative and electropositive parts are significantly unshielded and can interact with other charged atoms.[3]

It is generally accepted that water has a fundamental role in determining the proteins-folding rate. One reason why water is relevant for protein folding is the effective attraction that it induces among hydrophobic protein groups[3].

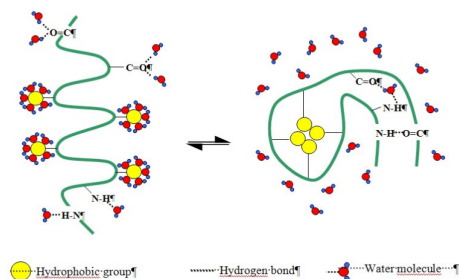


Figure 1: Water is fundamental in protein folding because of the hydrophobic effect[4].

A protein configuration with its hydrophobic groups exposed to water has a higher free energy than a configuration with the same groups buried inside the folded protein, because water around the hydrophobic group cannot form as many hydrogen bonds as in the bulk. Therefore, by decreasing the temperature of the system

from high values, a protein tends to fold to minimize the system free energy. This contributes in a very favourable way to stabilize the folded protein.

By increasing the temperature, a protein unfolds (denatures). Protein's denaturation is normally associated with a structure disorder. This happens when the protein uncoils into a random shape. Then, protein loses its biological function as a consequence of changes in environmental conditions. The heat increases the kinetic energy and causes molecules to vibrate so rapidly and violently that bonds are disrupted. On the other hand, proteins in aqueous solutions can denature upon cooling, this is known as *cold denaturation*. Cold denaturation is thermodynamically justified by the large free-energy gain of water due to the formation of more persistent hydrogen bonds at the interface with the denatured protein.

Proteins can denature not only by increasing or decreasing the temperature, but also as an effect of pressure change. We will discuss more in detail the pressure effect in the following.

II. PROTEIN STABILITY AT LOW TEMPERATURES

A. Close stability region in pressure-temperature plane

Nowadays, it has been proved that the native folded state of many proteins is stable in a limited range of temperatures and pressures. Hawley proposed a theory predicting this stability region[5][6]. Such theory predicts a close stability region with an elliptic shape in the temperature-pressure plane. Hawley's theory is based on two strong hypotheses:

1. Proteins have only two different states: folded and unfolded.
2. Thermodynamic equilibrium holds during denaturation process. In other words, the whole process is reversible.

These two hypotheses do not necessarily hold for many proteins. It is therefore worth testing this prediction with an alternative approach. Here, adopting a coarse-grain

*Electronic address: npagesge@gmail.com

model[7][8][9], we test how the stability region changes when we modify the pressure dependence of the low local compressibility at the water interface. This propriety can indeed, largely change depending on the hydrophobicity at the interface [10][11].

B. A coarse-grain model for a protein solvated in explicit water

We adopt the coarse-grain model for protein solvated in explicit water, introduced in Ref[7][8][9]. The protein's solvent described by the 'many-body' water model. In this model hydration water is particionate into cells that contain at most one water molecule.

The Hamiltonian of the bulk water is:

$$H \equiv U(r) - JN_{HB} - J_{\sigma}N_{coop} \quad (1)$$

which $U(r)$ is:

$$U(r) \equiv \begin{cases} \infty & \text{for } r_{ij} \leq r_0 \\ 4\epsilon \left[\left(\frac{r_0}{r_{ij}} \right)^{12} - \left(\frac{r_0}{r_{ij}} \right)^6 \right] & \text{for } r_0 \leq r_{ij} \leq 6r_0 \end{cases} \quad (2)$$

where r_0 is the hard core volume and r_{ij} is the distance between two water molecules i and j . This term represents the isotropic part of the water molecules interaction due to the van der Walls attraction and the hard core repulsion.

The second term accounts for the directional components hydrogen bonds. Where $N_{HB} \equiv \sum_{(i,j)} n_i n_j \delta_{\sigma_{ij} \sigma_{ij}}$ is the number of hydrogen bonds in the bulk water defined as and the sum extends aver all next neighbour molecules ij . The model assumes that every water molecule i can form up to 4 possible hydrogen bonds. To describe this feature, we define the bonding variable $\sigma_{ij} = 1, \dots, q$, between molecules i and j . If two bonding variables of two facing water molecules are in the same state, a hydrogen bond is created, with an energy gain J . If we assume that the maxim deviation of hydrogen bonds from the linear configuration is 30° , then $q=6$.

The last term describes the cooperative component of the hydrogen bond interaction, due to quantum effect, between the molecule i and the hydrogen bonded molecules in the first hydration shell. Its characteristic energy is J_{σ} . $N_{coop} \equiv J_{\sigma} \sum_i n_i \sum_{(l,k)_i} \delta_{\sigma_{ik} \sigma_{il}}$ is the number of cooperative hydrogen bonds's interactions. The sum runs over the 6 possible couples of bonds that form by each molecule i .

Water molecules which forming an hydrogen bond network have smaller density with respect to water molecules without hydrogen bonds. Therefore, the model assumes that the volume increases when a hydrogen bond is created. As a consequence, the volume V_{bulk} of bulk water is a function of the number of hydrogen bonds.

$$V_{bulk} = Nv_0 + N_{HB}v_{HB} \quad (3)$$

For sake of simplicity the protein here is represented as a self-avoiding hydrophobic chain. We therefore ignore interaction among protein-residue because we want to study the roll of water in the folding process. We do not neglect how the interface effects the water properties. In particular, based on general consideration[7][8][9], we assume that water-water hydrogen bonds near a hydrophobic interface are stronger (more persistent) than the hydrogen bonds between water molecules in the bulk. Furthermore, the interface effects the local compressibility of water[10][11]. We take into account this property by assuming that the proper volume of hydration water dependences on pressure. In Ref.[7][8][9] for simplicity this dependence is assumes to be linear. Here we test how a quadratic dependence would effect the protein stability region.

$$v_{surf} \equiv a_0 + a_1 P + a_2 P^2 \quad (4)$$

in the equation (4), a_0 , a_1 and a_2 are adjustable parameters.

In order to compare our results with the ones of Ref[7][8][9], we chose the following parameters for the model: $v_{HB}/v_0 = 0.5$, $J_{surf}/(4\epsilon) = 0.55$, $J/(4\epsilon) = 0.3$, $J_{\sigma}/(4\epsilon) = 0.05$, $a_0/v_0 = 0.5$ and $a_1(4\epsilon)/v_0^2 = -1$. We vary the parametric $a_2(4\epsilon)^2/v_0^3$ among the following values: 0, -0.1, -0.5 and -1. We chose negative values to a_1 and a_2 because we expect that the proper volume of the hydration water molecules decreases for increasing pressure.

III. MONTE CARLO SIMULATIONS

We study our system with Monte Carlo simulations using a NpT-ensemble (constant N, constant p, constant T)[7][8][9]. The initial configurations of our protein are random. We simulate the system for temperatures, $T(4\epsilon/k_B)$, between 0.1 to 0.9, and pressures, $P(4\epsilon/v_0)$ between -0.3 to 0.9. For each temperature and pressure we simulated so in different initial configurations and collect 5000 independents states.

To evaluate if the protein is folded or denaturated, we calculate the number of contact points between noncontiguous residues of the protein.

IV. RESULTS AND DISCUSSION

A. Linear model, $a_2 = 0$

For comparison, we first consider with previous work[7][8][9] the linear model: $v_{surf} \equiv a_0 + a_1 P$. For five pressure: $-0.3(4\epsilon/v_0)$, $-0.1(4\epsilon/v_0)$, $0.1(4\epsilon/v_0)$, $0.3(4\epsilon/v_0)$, $0.5(4\epsilon/v_0)$, $0.7(4\epsilon/v_0)$ (Fig.2). For each pressure we find that at high temperature the number of contact points is below 30% of the total number of possible contact points. By decreasing the temperature the

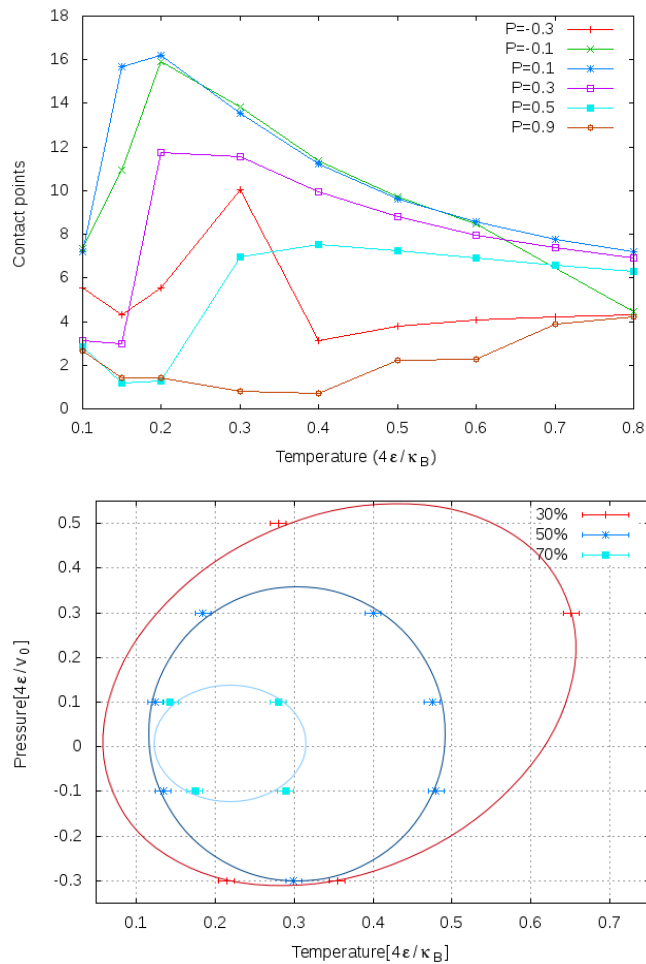


Figure 2: Linear model for $a_2 = 0$. (a) Average number of residue-residue contact points as function of temperature, for different pressures. When the number of contact points increases, the protein folds. The non-monotonic behaviour of the number of contact points shows that the protein is denatured at high and low temperature. (b) Stability region of the protein in the pressure-temperature plane: state points connected by line have 30% (red lines), 50% (blue lines) and 70% (cyan line) stabled contact points.

number of contact points gradually increases up to more than 70%. We project the average number of contact points in the temperature-pressure plane (Fig 2b). We assume that the region with 50% of stabled contact points is the limit between the folded and the unfolded state of the protein.

B. Quadratic model, $a_2 \neq 0$

Next, we considered the quadratic model with $a_2 \neq 0$. Three types of simulations for the quadratic model have been realized: $a_2(4\epsilon)^2/v_o^3 = -0.1$, $a_2(4\epsilon)^2/v_o^3 = -0.5$ and $a_2(4\epsilon)^2/v_o^3 = -1$.

When we compare the average number of contact point

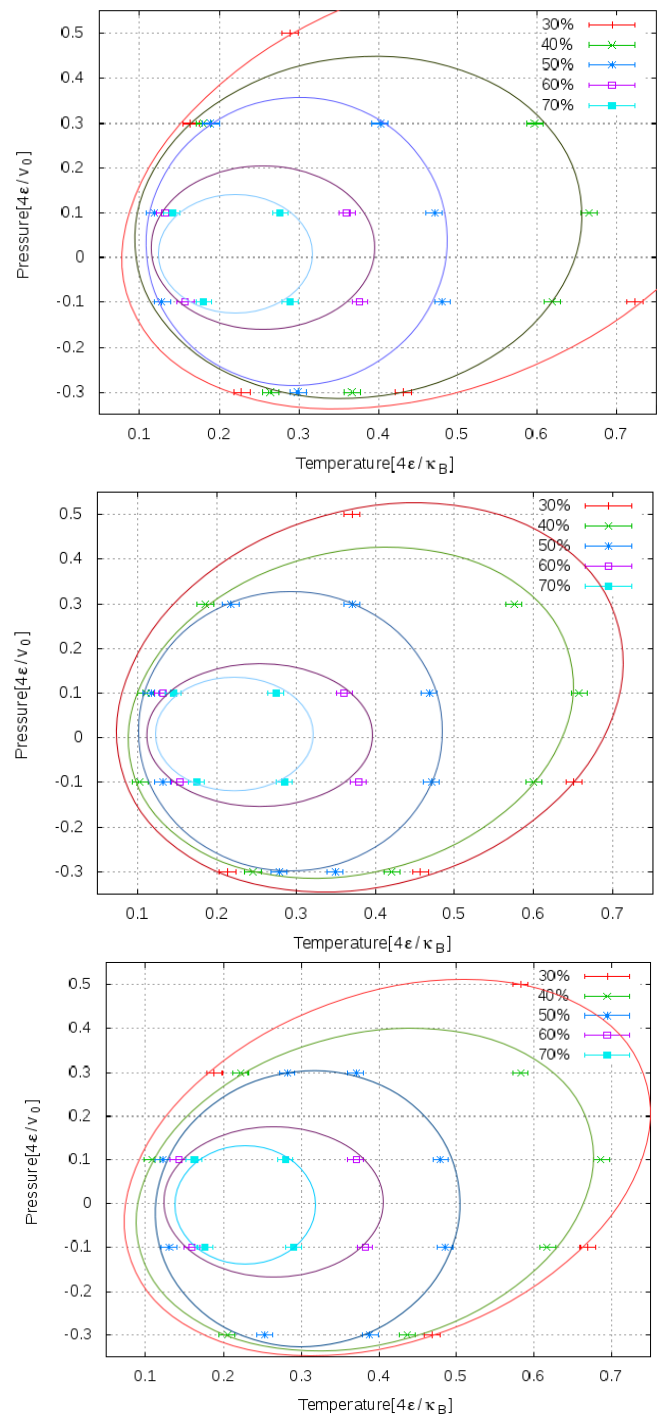


Figure 3: Stability region of the protein using the quadratic model for (a) $a_2 = -0.1$ (b) $a_2 = -0.5$ (c) $a_2 = -1$

as a function of temperature, we observe its behaviour is quantitatively similar to the linear case. However, the stability regions for the quadratic model are characterized by a pressure-range of stability that is reduced with respect to the linear model. The effect is stronger for larger absolute value of the negative parameter a_2 . As expected, different values of a_2 have no relevant effect on

the temperature range of stability. As consequence of these effects, by increasing $|a_2|$ we observe that the stability region becomes more elliptic. In particular, the 50% stability region is almost circular for $a_2 = 0$ which is elliptic for $a_2 = -1$

C. Discussion

We understand our results as the effects of the free-energy changes for the system. In particular, at high temperature, the system maximizes the entropy by unfolding the protein. By decreasing temperature at constant pressure, water tends to form an increasing number of the bulk hydrogen-bonds minimizing the system energy. This implies to reduce the water-protein interface by increasing the number of hydrophobic residue-residue contact points (hydrophobic collapse). However, by further reducing temperature, the system gains further potential energy by establishing as many stronger interface water hydrogen bonds. As consequence, the water-protein interface is maximized and the protein unfolds.

On the other hand, a pressure increase induces the protein denaturation because at the protein-water interface the compressibility of water is larger. Hence, by maximizing the protein-water interface, the system gains enthalpy. By decreasing pressure, the denaturation is associated to a more complex balance between entropy and enthalpy[7][8][9].

We finally observe that Hawley's prediction about elliptic stability regions, would corresponds in our model to a larger negative value for the quadratic parameter a_2 related to the local compressibility of hydrophobic water.

V. CONCLUSIONS

- The model presented here clearly show water has a fundamental roll in folding and unfolding pro-

cess for proteins. This roll is so important, even if we take into account the proprieties hydration water discounting all the contribution that come from protein-protein interaction, we can reproduce the protein stability region.

- By Monte Carlo simulation, we find the stability region has an elliptic shape as predicted by Hawley's theory.
- Building up on previous model, we study the effect of changing the local compressibility of hydration water. To do this goal, we adopt a quadratic relation between the proper volume water molecules at interface and the pressure. We find that modulation the quadratic term, we can reduce the range of pressure for which the folded protein state. In particular, the larger is the quadratic term, the smaller is the pressure range of stability. Therefore, the new model parameter associated to quadratic dependence opens up the possibility to adjust the model to real experimental data.

Acknowledgments

I want to thank my advisor, Giancarlo Franzese, and Valentino Bianco for helping me make the Treball de Fi de Grau. My thanks also to my parents and to Adrià Alfaro.

-
- [1] Ambrish Roy , Alper Kucukural & Yang Zhang. *I-TASSER: a unified platform for automated protein structure and function prediction*. Nature Protocols 5, 725–738 (1 April 2010).
- [2] Nobuyasu Koga, Rie Tatsumi-Koga, Gaohua Liu, Rong Xiao, Thomas B. Acton, Gaetano T. Montelione & David Baker. *Principles for designing ideal protein structures*. Nature 491, 222–227 (08 November 2012).
- [3] Jianmin Gao, Daryl A Bosco, Evan T Powers & Jeffery W Kelly. *Localized thermodynamic coupling between hydrogen bonding and microenvironment polarity substantially stabilizes proteins*. Nature Structural & Molecular Biology 16, 684 - 690 (2009).
- [4] K. Bartik. *The role of water in the structure and function of biological macromolecules*. Société Française d'Exobiologie, French Astrobiology Society (7 Novembre 2005)
- [5] K. Stokely, M. G. Mazza, H. E. Stanley & G. Franzese. *Effect of hydrogen bond cooperativity on the behavior of water*. Proceedings of the National Academy of Sciences (PNAS) 107, 1301-1306 (2010).
- [6] M. G. Mazza, K. Stokely, S. E. Pagnotta, F. Bruni, H. E. Stanley & G. Franzese. *More than one dynamic crossover in protein hydration water*. Proceedings of the National Academy of Sciences of the USA 108, 19873 (2011).
- [7] G. Franzese, V. Bianco & S. Iskov. *Water at interface with proteins*. Food Biophysics 6, 186 (2011).
- [8] V. Bianco, S. Iskov & G. Franzese. *Effect of hydrogen bonds on protein stability*. Journal of Biological Physics, 38, 27 (2012).

- [9] G. Franzese & V. Bianco. *Water at Biological and Inorganic Interfaces*. Food Biophysics, 8, 153 (2013).
- [10] R. Godawat, S. N. Jamadagni & S. Garde. *Characterizing hydrophobicity of interfaces by using cavity formation, solute binding, and water correlations*. Proceedings of the National Academy of Sciences U.S.A. 106, 15119 (2009).
- [11] E. G. Strelakova, J. Luo, H. E. Stanley, G. Franzese & S. V. Buldyrev. *Confinement of Anomalous Liquids in Nanoporous Matrices*. Physical Review Letters, 109, 105701 (2012).