Conformational transitions of a confined lattice protein: A Wang-Landau study

Busara Pattanasiri1,2,3, Ying Wai Li1, David P. Landau1, Thomas Wüst4 and Wannapong Triampo2,5,6

1 Center for Simulational Physics, Department of Physics and Astronomy, University of Georgia, Athens, GA 30602, USA
2 R&D Group of Biological and Environmental Physics, Department of Physics, Faculty of Science, Mahidol University, Bangkok 10400, Thailand
3 Department of Chemistry, Faculty of Science, Mahidol University, Bangkok 10400, Thailand
4 Swiss Federal Research Institute WSL, Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland
5 ThEP Center, CHE, 328 Si Ayutthaya Road, Bangkok 10400, Thailand
6 Institute for Innovative Learning, Mahidol University, 999, Phuttamonthon 4 Road, Salaya, Nakorn Pathom 73170, Thailand

E-mail: bairng@physast.uga.edu, ywli@physast.uga.edu, dlandau@physast.uga.edu, thomas.wuest@wsl.ch, wtriampo@gmail.com

Abstract.

We use Wang-Landau sampling with suitable Monte Carlo trial moves to study a hydrophobic-polar (HP) lattice protein confined between two parallel, attractive walls. The density of states is determined iteratively by a random walk in energy space. Thermodynamic and structural properties, such as specific heat, number of surface contacts and number of H-H monomer pairs, are then calculated. When the surface attraction is comparable to the internal attraction among the hydrophobic monomers in the chain, two conformational “transitions”, adsorption at higher temperature and collapse at lower temperature, have been analyzed based on these properties. This transition behavior depends on the variation of surface separation.

1. Introduction

Proteins are the most important molecules in living cells. They can fold to native structures to perform their own specific biological functions, e.g. cell structure stabilization, transporting chemical compounds, and converting chemical energy into mechanical energy [1, 2]. Misfolding to incorrect three-dimensional conformations produces inactive proteins that lead to toxicity. It can cause several diseases, such as Alzheimer’s, Huntington’s, Parkinson’s, and cancer-related syndromes [3–5]. As a result, there is a wide interest in studying parameters that influence the conformational stability of proteins in solution or cellular environment.

Some experimental and theoretical work indicates that confinement can have a stabilizing effect on protein folding to the native state [6–10], understanding confinement effects on protein folding is thus important. Moreover, protein in confined spaces is also an interesting topic in biotechnology and food processing applications [11, 12].

A number of techniques are proposed to study proteins in vivo and in vitro, e.g. surface plasmon resonance (SPR) imaging [13], ellipsometry [14] and atomic force microscopy [15]. However, monitoring the role of proteins in the living cell is much more difficult to perform and
most of the methods are restricted to indirect observations. The capability of modern computer simulations offers an opportunity to overcome these limitations. A widely used subject of study is a simplified but efficient protein model, hydrophobic-polar (HP) model, introduced by Dill in 1985 [16].

In this work we focus on the thermodynamic and structural properties of a confined HP lattice protein between two parallel, attractive surfaces. We have employed Wang-Landau sampling [17-20] with suitable Monte Carlo trial updates, pull moves [21] and bond-rebridging moves [22], to obtain the density of states. The effect of volumetric spaces between two attractive surfaces is also discussed.

2. HP lattice protein model
The hydrophobic-polar (HP) model is a simplified model for investigating protein folding to a native state. In this model, an amino acid is classified as either a hydrophobic residue (H) or a polar residue (P) due to their affinity with water. The protein sequence is modeled as a self-avoiding walk on a 3-dimensional cubic lattice. To mimic the hydrophobic effect, an attraction with strength $\varepsilon_{HH}$ is assigned between two non-bonded, nearest-neighbor hydrophobic residues so as to produce the hydrophobic core as found in the native structure of a real protein. However, it should be noted that since the protein sequence and interactions are overly simplified, the model introduces an artificial degeneracy to the system. An HP chain behaves similarly as a real protein in the sense that it folds to structures with a minimum free energy, but the “native” state structure is not unique.

In order to study a confined protein in our system, two attractive surfaces are constructed parallel to the xy-plane, at $z = 0$ and $z = h_w$. The energy function of an HP sequence in a confined space can thus be calculated as

$$E = -\varepsilon_{HH}n_{HH} - \varepsilon_S n_S,$$

where $n_{HH}$ is the number of hydrophobic pairs, $n_S$ is the number of monomers adjacent to the attractive surfaces, and $\varepsilon_S$ is the surface attractive strength. A schematic diagram of our system is shown in Figure 1.

3. Wang-Landau sampling and trial moves
Wang-Landau (WL) sampling [17-20] is used to estimate the density of states in energy, $g(E)$. Starting from an arbitrary configuration and an initial guess of $g(E) = 1$ for all energies $E$, a random walk in energy space is then accomplished by generating trial configurations. The

---

**Figure 1.** Schematic diagram of an HP protein between two attractive surfaces. The big orange spheres and the small black spheres represent the P-monomers and H-monomers, respectively. The light spheres are attractive surfaces molecules.

---
acceptance probability from the current energy $E_1$ to a new energy $E_2$ is

$$\begin{equation}
p(E_1 \rightarrow E_2) = \min \left( \frac{g(E_1)}{g(E_2)}, 1 \right).
\end{equation}$$

(2)

Every time a new configuration is accepted, $g(E_2) \rightarrow g(E_2) \times f$; otherwise, $g(E_1) \rightarrow g(E_1) \times f$, where $f$ is the modification factor which is initially set to $e \approx 2.71828$. During the random walk, the histogram of visited states is accumulated, i.e., $H(E) \rightarrow H(E) + 1$. When the histogram is sufficiently “flat”, $H(E)$ is reset and a new random walk is performed with a smaller modification factor $f = \sqrt{f}$. In order to yield an accurate density of states, $H(E)$ is said to be flat when all the entries of $H(E)$ are greater than 80% of their average and the final modification factor $f$ is less than $\exp(10^{-8})$.

During the simulation, two types of trial moves, pull moves [21] and bond-rebridging moves [22], are used for generating new configurations. These moves have been proven to be a powerful for ground state search and accurate estimation of $g(E)$ when combined with WL sampling [23]. Bond-rebridging moves are efficient to reorder the monomers in a highly dense conformation; however, its acceptance rate is rather low. To compensate for this, a higher calling probability (80%) is assigned to bond-rebridging moves, while 20% of the new trial configurations are generated from pull moves.

4. Thermodynamics of physical quantities

After the density of states is obtained from WL sampling, the partition function at a particular temperature, $Z$, with respect to the total energy of the system $E$ can then be calculated as

$$Z = \sum_E g(E)e^{-E/k_B T},$$

(3)

where $T$ is temperature and $k_B$ is the Boltzmann constant. From the partition function, any desired thermodynamic properties can be derived. For instance, the average energy $\langle E \rangle$ and specific heat $C_V$ can be calculated as

$$\begin{equation}
\langle E \rangle = \frac{1}{Z} \sum_E E g(E)e^{-E/k_B T},
\end{equation}$$

(4)

$$C_V = \frac{1}{k_B T^2} \langle (E^2) - (E)^2 \rangle.$$
The partition function of structural quantity $Q$ at a particular temperature, $Z$, and the average of $Q$, $\langle Q \rangle$, can then be obtained:

$$Z = \sum_{E,Q} g(E,Q)e^{-E/k_BT}$$

$$\langle Q \rangle = \frac{1}{Z} \sum_{E,Q} Qg(E,Q)e^{-E/k_BT}.$$

6. Results and Discussion
We performed simulations on a benchmark 48mer sequence (PHPHPPPPHPHPPHHHNNHHPPHHPPHPHPPHPPHPPHPPHHPPPPH), which has been designed by Yue et al [24] for algorithm testing purpose. In this sequence, the numbers of H-monomers and P-monomers are the same. Both surfaces interact with all monomers regardless of their type and the surface strength is set to be the same as the internal H-H interaction, i.e., $\varepsilon_S = \varepsilon_{HH} = 1$.

To understand the folding behavior of a confined protein, we simulated the case of a HP chain located between two attractive surfaces with surface separation $h_w = 49$, i.e., the chain can touch both surfaces only when it is a vertical straight chain. In this case, two peaks in specific heat, corresponding to conformational transitions, are observed as shown in Figure 2. The nature of these two peaks can also be identified by examining $d\langle n_S \rangle/dT$ and $d\langle n_{HH} \rangle/dT$ as shown in the lower panel. The peaks in $C_V$ and $d\langle n_S \rangle/dT$ at higher temperature, $kT/\varepsilon_{HH} \approx 2.15$, correspond to an adsorption transition, where an extended chain-like structure in free space adsorbs on a surface.

Reducing the temperature further makes the chain keep forming surface contacts until most of the monomers are on the surface. We called this process a "flattening" of the structure, and it is signaled by a small bump in $d\langle n_S \rangle/dT$ around $kT/\varepsilon_{HH} \approx 1.0$. In this case, the flattening process is not signaled in $C_V$. The peak at lower temperature, $kT/\varepsilon_{HH} \approx 0.4$, represents the hydrophobic core (H-core) formation in 2-dimensions as $d\langle n_{HH} \rangle/dT$ also shows a peak at the same temperature. Typical configurations are shown in the middle panel.

When the surface separation is reduced, vertical movements of the chain are restricted, resulting in a less pronounced adsorption peak due to a reduction in the number of conformations available. See the cases of $h_w = 49$ and $h_w = 25$ in Figure 3. Further reduction in the surface separation decreases the adsorption peak height until it becomes a shoulder, as in the cases of $h_w = 10$ and $h_w = 6$. However, this transition behavior is different from that for larger surface separations. In this case, the chain can have contacts with both surfaces, forming "bridges" between them even at high temperature. To maximize the number of surface contacts, the chain needs to detach from one surface and adsorb on the other until most of the monomers sit on the same surface. We called it a "debridging" process, which becomes harder to occur when the surface separation is smaller, resulting in a shift of the shoulder in $C_V$ to a lower temperature. The debridging transition is still signaled by a peak in $d\langle n_S \rangle/dT$ as for the adsorption transition.

H-core formation temperature is not affected by surface separation. It forms a 2-dimensional H-core when the chain is already in contact with either surface, except when it is very squeezed by the two surfaces as in the $h_w = 3$ case. The chain transforms from an extended squeezed chain to a 3-dimensional squeezed globule since all monomers touch the surfaces all the time. Therefore, only a single H-core formation peak with an exceptionally large magnitude is observed in $C_V$ and $d\langle n_{HH} \rangle/dT$, respectively.

Figure 4 shows the "adsorption" temperatures for different $h_w$, as obtained from the peak positions in $C_V$ and $d\langle n_S \rangle/dT$ for the adsorption or debridging process. The difference in adsorption temperatures determined by these two measures systematically widens with the reduction of surface separation. It should be noted that there is no finite size scaling in the HP model, as each sequence is unique even with the same chain length [25]. To understand the
Thermodynamics of the 48mer confined between two attractive surfaces with $\varepsilon_S = \varepsilon_{HH} = 1$ and $h_w = 49$. Upper panel is the specific heat; middle panel is the typical configurations at each temperature; lower panel is the thermal derivative of the number of surface contacts, $d\langle n_S \rangle/dT$, and that of the number of H-H pairs, $d\langle n_{HH} \rangle/dT$. The big orange spheres and the small black spheres represent the P-monomers and H-monomers respectively. The light spheres are attractive surfaces molecules. Error bars are smaller than the data points.

The origin of the difference, one may start from the energy function of the system in Eq. 1. The specific heat and its thermal derivative can be expressed as:

$$C_V = \frac{d\langle E \rangle}{dT} = -\varepsilon_{HH} \frac{d\langle n_{HH} \rangle}{dT} - \varepsilon_S \frac{d\langle n_S \rangle}{dT},$$

$$\frac{dC_V}{dT} = -\varepsilon_{HH} \frac{d^2\langle n_{HH} \rangle}{dT^2} - \varepsilon_S \frac{d^2\langle n_S \rangle}{dT^2}.$$  

By considering the lower panel in Figure 3, we see that $d^2\langle n_{HH} \rangle/dT^2 \geq 0$ at the adsorption temperature determined from the $C_V$ peak. This implies $d^2\langle n_S \rangle/dT^2 \leq 0$, since the left hand-side in Eq. 9 is zero at the adsorption temperature. The adsorption peak in $d\langle n_S \rangle/dT$, therefore, occurs at the same or at a higher temperature than that of the peak in $C_V$. 

Figure 2. Thermodynamics of the 48mer confined between two attractive surfaces with $\varepsilon_S = \varepsilon_{HH} = 1$ and $h_w = 49$. Upper panel is the specific heat; middle panel is the typical configurations at each temperature; lower panel is the thermal derivative of the number of surface contacts, $d\langle n_S \rangle/dT$, and that of the number of H-H pairs, $d\langle n_{HH} \rangle/dT$. The big orange spheres and the small black spheres represent the P-monomers and H-monomers respectively. The light spheres are attractive surfaces molecules. Error bars are smaller than the data points.
Figure 3. Thermodynamics of the 48mer confined between two attractive surfaces with different separation distances, $h_w$. $\varepsilon_S = \varepsilon_{HH} = 1$ in all cases. Upper panel is the specific heat; middle panel is the thermal derivative of number of surface contacts, $d\langle n_S \rangle / dT$; lower panel is the thermal derivative of number of H-H pairs, $d\langle n_{HH} \rangle / dT$. Error bars are smaller than the data points.

The systematic changes in the difference in the two adsorption temperatures determined by $C_V$ and $d\langle n_S \rangle / dT$ is caused by the different values of $d^2\langle n_{HH} \rangle / dT^2$. Narrowing the surface separation shifts the adsorption to a lower temperature, where $d^2\langle n_{HH} \rangle / dT^2$ has a larger value. The difference between the adsorption peak positions in $C_V$ and $d\langle n_S \rangle / dT$ are therefore enhanced.

7. Conclusions
In this work, a coarse-grained lattice model, the HP model, and Wang-Landau sampling with suitable Monte Carlo trial moves, pull moves and bond-rebridging moves, have been used to study a protein confined by two attractive surfaces. Narrowing the separation between the two attractive surfaces decreases the vertical degrees of freedom of the chain, resulting in a reduction in the magnitude and a shift to a lower temperature of the adsorption peak, while the collapse transition temperature is not affected. The difference in “adsorption” temperatures
Figure 4. Adsorption temperatures for different separation distances, $h_w$, determined by the adsorption peaks in the specific heat, $C_V$, and the thermal derivative of the number of surface contacts, $d\langle n_S \rangle/dT$, respectively.

determined by the peak positions in the specific heat and the derivative of the number of surface contacts is also observed, and it systematically increases with the reduction of surface separation. However, the details of this “adsorption” behavior might be different. It could simply be that a desorbed chain adsorbs on a surface, or that the bridging structure originally connecting the two surfaces gets destroyed, depending on the surface separation. Further investigation is necessary to distinguish these two conformational transition processes.

Acknowledgments
We thank M. Bachmann, S. Schnabel, and T. Vogel for fruitful discussions. This work is supported by the National Science Foundation (NSF) under grant no. DMR-0810223. B. Pattanasiri is supported by The Royal Thai Government Scholarship under the Development and Promotion of Science and Technology Talent Project (DPST), and the grant from The Institute for the Promotion of Teaching Science and Technology (IPST). W. Triampo is supported by Faculty of Science, Mahidol University, and the Thailand Center of Excellence in Physics (ThEP).

References
[1] Creighton T E 1993 Proteins: Structures and Molecular Properties (New York: W. H. Freeman)
[2] Collarini E J and Oxender D L 1987 Ann. Rev. Nutr. 7 75–90
[3] Taubes G 1996 Science 271 1493–5
[4] Selkoe D J 2004 Nat. Cell. Biol. 6 1054–61
[5] Forloni G et al 2002 Neurobiol Aging 23 957–76
[6] Javidpour L and Sahimi M 2011 J. Chem. Phys. 135 125101–12
[7] Mittal J and Best R B 2008 PNAS 105 20233–38
[8] Lucent D, Vishal V and Pande V S 2007 PNAS 104 10430–34
[9] Tian J and Garcia A E 2011 J. Am. Chem. Soc. 133 15157–64
[10] Sorin E J and Pande V S 2006 J. Am. Chem. Soc. 128 6316–17
[11] Martin C R and Kohli P 2003 Nat. Rev. Drug Discov. 2 29–37
[12] Foegeding E A, Luck P J and Davis J P 2006 Food Hydrocolloids 20 284–92
[13] Jordan C E and Corn R M 1997 Anal. Chem. 69 1449–56
[14] Seitz R, Brings R and Geiger R 2005 Appl. Surf. Sci. 252 154–57
[15] Agnihotri A and Siedlecki C A 2004 Langmuir 20 8846–52
[16] Dill K A 1985 Biochemistry 24 1501–9
[17] Wang F and Landau D P 2001 Phys. Rev. Lett. 86 2050–53
[18] Wang F and Landau D P 2001 Phys. Rev. E 64 056101–16
[19] Wang F and Landau D P 2002 Comput. Phys. Commun. 147 674–77
[20] Landau D P, Tsai S -H and Exler M 2004 Am. J. Phys. 72 1294–302
[21] Lesh N, Mitzenmacher M and Whitesides S 2003 Proc. Seventh Annual Int. Conf. on Computational Molecular Biology, ed T Lengauer et al. (Berlin: RECOMB) p 188
[22] Deutsch J M 1997 J. Chem. Phys. 106 8849–54
[23] Wüst T and Landau D P 2009 Phys. Rev. Lett. 102 178101–4
[24] Yue K, Fiebig K M, Thomas P D, Chan H S, Shakhnovich E I and Dill K A 1995 Proc. Natl. Acad. Sci. USA. 92 325–29
[25] Bachmann M and Janke W 2004 J. Chem. Phys. 120 6779–91