Generic folding and transition hierarchies for surface adsorption of hydrophobic-polar lattice model proteins

Ying Wai Li, Thomas Wüst, and David P. Landau

Phys. Rev. E **87**, 012706 — Published 11 January 2013

DOI: 10.1103/PhysRevE.87.012706

Generic folding and transition hierarchies for surface adsorption of HP lattice model proteins

Ying Wai Li*
Center for Simulational Physics, University of Georgia, Athens, Georgia 30602, U.S.A.†

Thomas Wüst‡
Swiss Federal Research Institute WSL, Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland

David P. Landau§
Center for Simulational Physics, University of Georgia, Athens, Georgia 30602, U.S.A.

The thermodynamic behavior and structural properties of hydrophobic-polar (HP) lattice proteins interacting with attractive surfaces are studied by means of Wang-Landau sampling. Three benchmark HP sequences (48mer, 67mer and 103mer) are considered with different types of surfaces, each of which attracting either all monomers, only hydrophobic (H) monomers, or only polar (P) monomers, respectively. The diversity of folding behavior in dependence of surface strength is discussed. Analyzing the combined patterns of various structural observables, such as e.g., the derivatives of the numbers of surface contacts, together with the specific heat, we are able to identify generic categories of folding and transition hierarchies. We also infer a connection between these transition categories and the relative surface strengths, i.e., the ratio of the surface attractive strength to the inter-chain attraction among H monomers. The validity of our proposed classification scheme is reinforced by the analysis of additional benchmark sequences. We thus believe that the folding hierarchies and identification scheme are generic for HP proteins interacting with attractive surfaces, regardless of chain length, sequence, or surface attraction.

PACS numbers: 05.10.-a, 82.35.Gh, 87.15.ak, 87.15.Cc

I. INTRODUCTION

Protein adsorption on solid surfaces has attracted strong research interest recently for its numerous applications in nanotechnology and biomaterials [1, 2]. The study of protein functions in experiments often involves the immobilization of proteins [3, 4]. Adhesion of proteins on solid substrates including metals, semiconductors, carbon or silica etc., enables the synthesis of new materials for biosensors or electronic devices [5–8]. It is also the key to reveal the principles of many biological processes and causes of diseases, e.g., when integrating implanted materials with body tissues [9, 10]. Understanding how the functions and conformations of a protein are affected by adsorption and desorption is an important topic in protein drug delivery [11, 12].

It is known that configurational changes of protein molecules upon surface adsorption depend on both the protein's properties (e.g. sequence, size, thermodynamic stability, etc.) and the surface properties (e.g. materials, polarity, surface roughness, etc.); but how large these changes are and where in the protein molecules they occur remain puzzles to be solved [1, 13]. Enormous efforts have been dedicated to unveil the mysteries in protein folding and adsorption mainly by experimental approaches [14, 15]. Nevertheless, the fact that only the "final product" can be obtained and studied in an experiment makes for slow progress in understanding the dynamics and folding processes. From another point of view, the diversity of possible protein sequences and sophisticated interactions among amino acids also complicate theoretical structural prediction in protein folding - not to mention when the protein interacts with solvent molecules or a substrate where an extra level of complexity enters.

With the advances in computer power, numerical simulation has become a promising way to shed more light on the general problem. The study of simplified, coarse-grained protein models in conjunction with Monte Carlo simulations, being able to efficiently explore large conformational phase spaces [16, 17], has been a particularly successful approach in understanding the principles behind protein folding and adsorption. Among these minimalist models, the hydrophobic-polar (HP) model introduced by Dill et al. [18, 19] has been a particularly active, yet difficult, research subject in recent years. The interest in the HP model is twofold: (i) Despite its known limitations [20, 21], the model captures some of the most important qualitative features which drive the folding of proteins and characterize their native states. It thus laid a basis to systematically study many problems in protein folding by means of computer simulations. (ii) The ground state search for an HP sequence is an NP-complete problem [22, 23], and the rough free en-

* ywli@physast.uga.edu
† Current affiliation: National Center for Computational Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, U.S.A.
‡ thomas.wuest@wsl.ch
§ dlandau@physast.uga.edu

ergy landscapes cause traditional Monte Carlo methods, e.g. Metropolis sampling [24], to fail in the low temperature regime [25]. Therefore, the HP model is a well-suited testing ground for a number of emerging numerical methods [26–34]. In addition, the thermodynamic behavior of different HP sequences can vary noticeably, even for the same chain length [32]. Consequently, finite-size scaling cannot be applied to a systematic study of the effect of system size, unlike many other models in statistical physics. Thus, our understanding of the general behavior of the model is still incomplete and the simulation of the HP model remains a challenging computational problem in statistical physics.

Various approaches have been undertaken to better understand the energy landscapes, thermodynamics and conformational transitions of HP proteins adsorbing on an attracting surface [35–40]. In this work, our intent is to identify generic thermodynamic and structural behavior of protein adsorption from a macroscopic perspective using the HP model. We have adopted Wang-Landau sampling [41–43] to obtain the density of states of a system, from which subsequent thermodynamics can be calculated. To the best of our knowledge, it is the first comprehensive analysis of structural transitions for protein adsorption that integrates results from *multiple* HP sequences. Previous work by Bachmann and Janke [39], Swetnam and Allen [40] or Radhakrishna et al. [44] only studied the conformational pseudo-phases based on individual benchmark HP sequences. Comparing the thermodynamic and structural properties of *multiple* sequences allowed us to identify categories of generic transition behavior and to draw a correspondence between these categories and the relative interaction strengths involved. This finding is fundamental as it implies that different HP sequences share certain general, qualitative, characteristics in structural transformations when brought near to an adsorbing substrate.

The article is organized as follows: Section II describes the model employed, Section III explains the sampling method and details, Section IV presents the simulation results, Section V discusses the analysis and interpretation of the results, and Section VI gives a summary and outlook of this work.

II. MODEL

The commonly known 22 amino acids found in proteins can roughly be classified as either hydrophobic or polar depending on the nature of their side chains. The tendency of the non-polar residues to stay away from water molecules has been identified as the key driving "force" in forming tertiary structures. The hydrophobic-polar (HP) model [18, 19] is a coarse-grained lattice model for proteins that captures this hydrophobic effect. In this model, an amino acid is treated as a single monomer and it can only be hydrophobic (H) or polar (P). A protein is thus represented by a heteropolymer which consists of N connected monomers of type H or P. An attractive interaction exists only between a pair of non-bonded nearest neighboring H monomers. This attraction is denoted by ε_{HH} in our discussion, and the magnitude indicates the ability of the H monomers to pull themselves together as determined by the insolubility of the protein in an aqueous environment. In other words, the solvent quality is intrinsically considered by the model. Other factors like charges and acidity of amino acids that also govern protein folding are not handled in this scope.

In addition to the internal interactions within the polymer, the binding of a protein with an attractive substrate contributes to the total energy. We have considered three types of surface fields in view of the setting of the HP model: (i) a surface interacts only with H monomers (a hydrophobic surface) with strength ε_{SH}, (ii) a surface interacts only with P monomers (a polar surface) with strength ε_{SP}, and (iii) a surface interacts with both H and P monomers with equal strength, i.e., $\varepsilon_{SH} = \varepsilon_{SP} \neq 0$. The energy function of an HP sequence interacting with a surface takes the general form:

$$E = -\varepsilon_{HH} n_{HH} - \varepsilon_{SH} n_{SH} - \varepsilon_{SP} n_{SP}, \qquad (1)$$

where n_{HH} denotes the number of interacting pairs between H monomers, n_{SH} the number of surface contacts with H monomers and n_{SP} the number of surface contacts with the P monomers. Besides contributing to the system's energy, these three quantities are also useful "order" parameters that give quantitative measures of the structure of a conformation [45]. The negative signs in front of each term indicate that it is energetically favorable when the monomers interact or come in contact with the surface.

Our simulations were performed on a three-dimensional simple cubic lattice with the attractive surface being the xy-plane placed at $z = 0$. A non-interacting steric wall is placed at $z = N + 1$ to confine the polymer in a way that it can contact both walls with its ends only when it is a vertical straight chain. Periodic boundary conditions are imposed in the x and y directions.

III. METHOD

A. Calculation of thermodynamic quantities

The partition function, Z, at a particular temperature T can be expressed in terms of the energy density of states $g(E)$:

$$Z = \sum_E g(E) e^{-E/k_B T}, \qquad (2)$$

where E is the energy of the system as defined by the energy function, k_B is the Boltzmann constant, and the sum runs over all the energies that the system can take. Since $g(E)$ does not depend on T, one may calculate

Z at any temperature with a single estimate of $g(E)$. All the thermodynamic quantities then follow from the knowledge of Z. For example, the average energy $\langle E\rangle$ and the heat capacity C_V are calculated as:

$$\langle E\rangle = \frac{1}{Z}\sum_E Eg(E)e^{-E/k_BT}, \tag{3}$$

$$C_V = \frac{\langle E^2\rangle - \langle E\rangle^2}{k_BT^2}. \tag{4}$$

The specific heat is then defined as C_V/N accordingly.

B. Wang-Landau sampling and trial moves

Wang-Landau (WL) sampling is a powerful, iterative algorithm to estimate the density of states, $g(E)$, with high accuracy. Details of the algorithm are described in Refs.[41–43]. In our simulations, rather stringent parameters were chosen in order to obtain accurate estimates for $g(E)$: We used a flatness criterion $p = 0.8$ for the 48mer and $p = 0.6$ for the 67mer and 103mer. The final modification factor was set to $\ln(f_{final}) = 10^{-8}$ in all cases.

It has been found that two types of trial moves, pull moves [46] and bond-rebridging moves [47], work particularly well together with WL sampling in search of the global energy minimum conformations and the determination of the density of states for lattice polymers [48–50]. The ability of reaching low energy states allows for a more thorough survey of conformational space, yielding a higher resolution of $g(E)$ and thus more precise thermodynamic quantities especially in the low temperature regime. This is of particular importance for longer chain lengths with more complex energy landscapes [50].

The two trial moves are called with different probabilities. Bond-rebridging moves transform a polymer from one compact state to another without uncoiling, making it more efficient than pull moves in dealing with densely packed polymers. However, since the energy difference resulting from a bond-rebridging move is relatively large, its acceptance rate is rather low. This drawback is compensated for with a higher calling ratio. In our simulations, we used move fractions of 80% and 20% for bond-rebridging moves and pull moves, respectively.

In order to fulfill detailed balance when employing pull moves, an extra factor is added to the acceptance probability of moving from state A (with energy E_A) to state B (with energy E_B):

$$P(A\to B) = \min\left(1, \frac{g(E_A)}{g(E_B)}\frac{n_{B\to A}/n_B}{n_{A\to B}/n_A}\right), \tag{5}$$

where $n_{A\to B}$ is the number of pull moves that transform A to B; n_A is the number of possible pull moves which can be performed from state A; $n_{B\to A}$ and n_B are defined likewise. Because of reversibility of pull moves, $n_{A\to B} = n_{B\to A}$ and the two terms cancel out in Eq. (5). When a pull move is chosen to generate a new configuration, a list of possible moves from state A is first constructed to obtain n_A. A move is then selected randomly from the list, generating state B. n_B is counted and $P(A\to B)$ can then be calculated. This procedure is computationally expensive (it slows down the simulation by approximately an order of magnitude); however, it secures the reliability of our results by taking detailed balance into careful consideration.

We adopted RANLUX as the random number generator (`gsl_rng_ranlxd2`), which is recommended by the GNU Scientific Library (GSL) for its "reliable source of uncorrelated numbers" and "strongest proof of randomness" [51]. It uses a lagged-Fibonacci-with-skipping algorithm [52] with double precision output and a long period of about 10^{171}.

The use of Wang-Landau sampling together with inventive trial moves is *essential* for uncovering the subtle, low temperature thermodynamics of these systems. This is demonstrated in FIG. 1 which shows the specific heat of a 36mer ($P_3H_2P_2H_2P_5H_7P_2H_2P_4H_2P_2HP_2$) interacting with a very weak attractive surface ($\varepsilon_{SH} = \varepsilon_{SP} = 1, \varepsilon_{HH} = 12$). As seen in the figure, all the surface related transitions below $k_BT/\varepsilon_{HH} \approx 0.3$ are completely inaccessible by Metropolis sampling, even with very long runs (10^8 trial moves!) incorporating pull moves and bond rebridging moves! In contrast, Wang-Landau sampling is even able to uncover the shoulder at $k_BT/\varepsilon_{HH} \approx 0.016$ shown in the inset (caused by a rapid change in configurational entropy of the excitation from the ground state to the first few excited states, see Ref. [53] for details). We, therefore, stress that sophisticated Monte Carlo al-

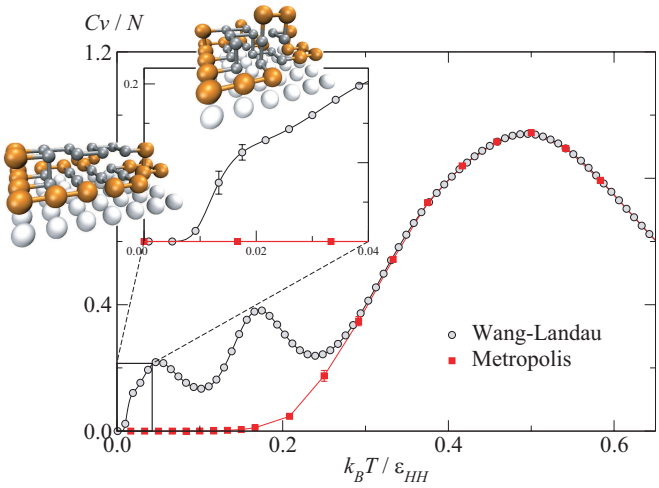


FIG. 1. (Color online) **Main graph:** Specific heat of a 36mer interacting with a very weak attractive surface as obtained by Wang-Landau and Metropolis sampling. Where not shown statistical errors are smaller than the data points. Comparable computational effort was used for each method. **Inset:** Results at extremely low temperature. Typical configurations above and below the shoulder are shown.

gorithms are crucial in simulating systems with delicate low temperature behavior.

Recently, improvements have been proposed to speed up and ensure the convergence of WL sampling in simulating lattice polymers or proteins [40, 44, 50, 54].

C. Calculation of thermodynamics of structural observables

Structural quantities are essential in understanding non-energetic properties of the system. They provide information on the structures and packing of the polymer. Apart from the quantities n_{HH}, n_{SH} and n_{SP} introduced in Section II, structural quantities that are often of interest include the radius of gyration,

$$R_g = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\vec{r}_i - \vec{r}_{cm})^2}, \tag{6}$$

and the end-to-end distance,

$$R_{ee} = |\vec{r}_N - \vec{r}_1|\,, \tag{7}$$

where $\vec{r}_{cm}$ in Eq. (6) is the center of mass of the configuration; $\vec{r}_i$ is the position of monomer i.

To obtain the thermodynamics of a structural observable Q, we estimated the joint density of states, $g(E, Q)$, by multicanonical sampling [55, 56]. Although it is feasible to sample $g(E, Q)$ all over again using a two-dimensional random walk in WL sampling if only one structural quantity is required, it becomes impossible if several of them are of interest. A more efficient way is to make use of the prior knowledge of $g(E)$ and perform a multicanonical sampling. In this process, trial states are accepted or rejected according to $1/g(E)$, where $g(E)$ is the one obtained previously by the one-dimensional WL sampling and is held fixed throughout the whole production procedure. As a new trial state is accepted, any desired structural quantity Q would be calculated and accumulated in a two-dimensional histogram, $H(E, Q)$. The simulation is brought to an end when a sufficiently large number of Monte Carlo steps have been performed. The joint density of states, $g(E, Q)$, is then obtained by reweighing $H(E, Q)$:

$$g(E, Q) = g(E)H(E, Q). \tag{8}$$

As such, we can obtain as many $g(E, Q)$'s for various Q's as desired in a single production run.

The partition function, Z_Q, for observable Q and its expectation value can then be obtained as

$$Z_Q = \sum_{E,Q} g(E, Q)e^{-E/k_BT}, \tag{9}$$

and

$$\langle Q\rangle = \frac{1}{Z_Q}\sum_{E,Q} Qg(E, Q)e^{-E/k_BT}. \tag{10}$$

IV. SIMULATION RESULTS

We have studied three benchmark HP sequences (48mer, 67mer and 103mer) interacting with three types of surfaces (see above). The 48mer ($PHPHP_4HPHPHP_2$-$HPH_6P_2H_3PHP_2HPH_2P_2HPH_3P_4H$) is Seq. 9 among the ten "Harvard sequences" designed originally for algorithm testing purpose [57], where the number of H and P monomers are exactly the same in each sequence, i.e., 24 H monomers and 24 P monomers respectively. The one chosen here for our simulation has the minimum ground state degeneracy in 3D free space [32, 57]. The 67mer ($PHPH_2PH_2PHPP$-$H_3P_3HPH_2PH_2PHP_2H_3P_3HPH_2PH_2PHP_2H_3P_3HP$-$H_2PH_2PHP_2H_3P$) was first introduced to resemble α/β-barrel in real proteins [58], while the 103mer ($P_2H_2P_5H_2P_2H_2PHP_2HP_7HP_3H_2PH_2P_6HP_2HPHP_2H$-$P_5H_3P_4H_2PH_2P_5H_2P_4H_4PHP_8H_5P_2HP_2$) was proposed to model cytochrome c [59].

A. Ground states and limiting behavior

Table I reports the lowest energies found for these three sequences during the estimation of $g(E)$ with different types of attractive surfaces.

TABLE I. Lowest energies found for the 48mer, 67mer and 103mer interacting with different attractive surface types and strengths, which are abbreviated in the surface labels (A, H or P stand for the surface types, the numbers stand for the ratio between ε_{SH} or ε_{SP} and ε_{HH}). Classification of transition categories are denoted by the Roman numbers in parentheses.

surface	ε_{HH}	ε_{SH}	ε_{SP}	48mer	67mer	103mer
Free space without surface:						
2D	1	/	/	−21	−29	−32
3D	1	/	/	−34	−56	−58
Surfaces attract all monomers:						
A1	1	1	1	−69 (I)	−96 (I)	−135 (I)
A2	1	2	2	−117 (I)	−163 (I)	/
A1/2	2	1	1	−93 (II)	−132 (II)	−167 (I/II)
Surfaces attract only H monomers:						
H1	1	1	0	−49 (II)	−72 (II)	−80 (II)
H2	1	2	0	−73 (I)	−108 (I)	/
H1/2	2	1	0	−79 (III)	−118 (II)	−128 (III)
Surfaces attract only P monomers:						
P1	1	0	1	−48 (II)	−69 (II)	−100 (I/II)
P2	1	0	2	−71 (I)	−91 (I)	/
P1/2	2	0	1	−79 (III)	−123 (II)	−150 (II)

To understand the "asymptotic" folding behavior in the limit of a surface with infinite attractive strength, we simulated the case where the HP chain was confined

to a two-dimensional free space. This is equivalent to restricting all monomers of the HP chain on the surface, giving a 2D ground state. Another limiting case is when the surface is totally absent in a three-dimensional space. For more details on these two cases, see [60].

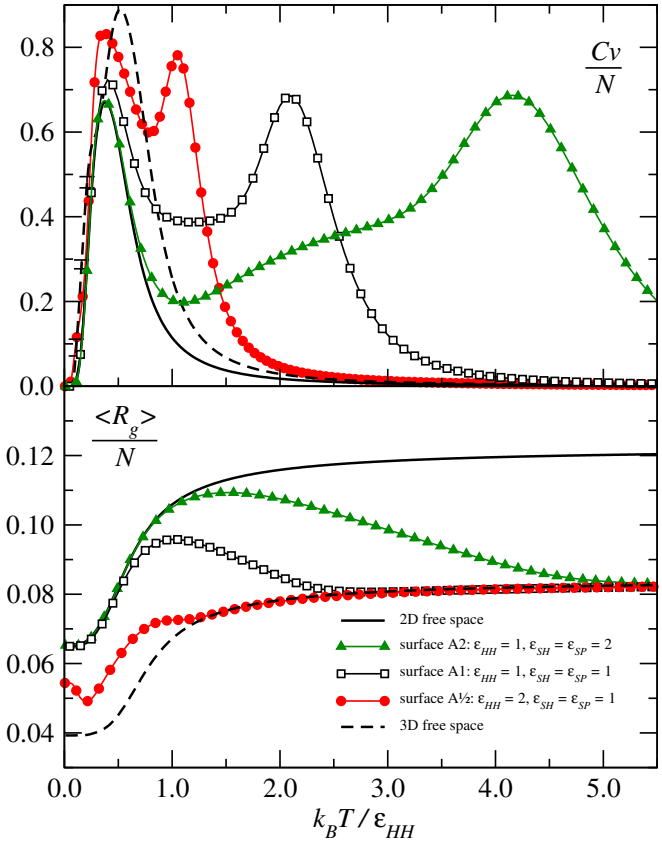


FIG. 2. (Color online) **Upper panel:** Specific heat C_V/N as a function of the effective temperature k_BT/ε_{HH}. **Lower panel:** Averaged radii of gyration per monomer, $\langle R_g\rangle/N$, as a function of k_BT/ε_{HH}, for the 48mer interacting with a surface which attracts all monomers with different strengths (surfaces A1, A2 and A1/2). Note that k_BT is scaled with the internal attraction strength, ε_{HH}, so as to compare different systems in the same energy scale. In this manner, any difference in quantities comes solely from the surface strengths ε_{SH} and ε_{SP}. Errors smaller than the data points are not shown.

These two limiting cases are useful in visualizing upper and lower bounds for thermodynamic observables and they serve as an aid to understand the details of folding behavior. A demonstrative example is the averaged radius of gyration per monomer, $\langle R_g\rangle/N$, shown in FIG. 2 for the 48mer interacting with a surface attracting all monomers (surfaces A1, A2 and A1/2). The radii of gyration of the two limiting cases are plotted on the same figure. Drawing a simple connection to the self-avoiding random walk on square and cubic lattices, it is obvious that $\langle R_g\rangle$ is largest when all the monomers are forced to sit on the surface to form planar structures, yielding the upper bound for $\langle R_g\rangle$. Correspondingly, $\langle R_g\rangle$ is smallest when the HP chain is allowed to fold freely in a three-dimensional space to form 3D structures, giving the lower bound of $\langle R_g\rangle$. Generally speaking, when the HP chain is placed near a surface of finite attractive strength, it remains as an extended coil at high temperature as if the surface is absent. The radii of gyration for all cases thus coincide with the 3D, surface-free one.

As the temperature decreases, the HP chain interacting with a stronger attractive surface (A2) starts the adsorption process the earliest at $k_BT/\varepsilon_{HH} \approx 5.0$ as its $\langle R_g\rangle$ "departs" from the lower bound and begins to approach the upper bound. Such an adsorption "transition" is clearly signaled by the peak in C_V centered at $k_BT/\varepsilon_{HH} \approx 4.25$. At $k_BT/\varepsilon_{HH} \approx 1.0$, $\langle R_g\rangle$ merges with the upper bound signifying a complete adsorption of all monomers. The formation of a hydrophobic core in which the number of inter-chain H-H interactions, n_{HH}, is maximized, then takes place entirely on the surface in this case until the ground state is reached at zero temperature. This process in the low temperature regime is identical to the one in two-dimensional free space, as indicated by the complete agreement in the radii of gyration and the coincidence of the peak at $k_BT/\varepsilon_{HH} \approx 0.5$ observed in C_V.

The thermodynamics for surface A1 is qualitatively similar to that of surface A2 except that it requires a lower adsorption temperature. Since the radii of gyration for both surface types end up with the same value as the upper bound at $T = 0$, one may expect that the ground state conformations for both systems are two-dimensional. This has been confirmed by the number of surface contacts ($n_{SH} = n_{SP} = 24$, meaning the entire chain is in contact with the surface) and the number of H-H interactions ($n_{HH} = 21$, which is the same as the ground state of the 2D limiting case).

While the two peaks in the specific heat of surface A1/2 tend to give the impression that it has the same qualitative folding behavior as the previous cases, the shape of the radius of gyration clearly distinguishes it from the others, apart from showing that the ground state is now three-dimensional. This is the first clue that the specific heat alone does not provide a complete picture of structural transition behavior. Indeed, the transition hierarchy of the 48mer interacting with surface A1/2 is different from surfaces A1 and A2, which can only be verified by examining other structural parameters as we shall see in the following section.

B. Identification of transition categories by canonical analysis of specific heat and structural quantities

Three major transition categories were identified from the 24 systems presented in TABLE I by considering the combined patterns of the specific heat C_V/N and the average radius of gyration $\langle R_g\rangle/N$:

Category I: C_V shows two peaks, a bump between the

peaks might be possible, $\langle R_g \rangle$ shows a maximum between these two peaks;

Category II: C_V shows two peaks, $\langle R_g \rangle$ decreases upon cooling. In the very low temperature regime, it might rise back up a little to form a minimum when the temperature approaches zero;

Category III: C_V shows only one peak with possible shoulders, $\langle R_g \rangle$ decreases upon cooling.

The different combinations of C_V and $\langle R_g \rangle$ in the transition categories are caused by the different order of occurrence in folding processes, which are revealed by further investigation of proper structural parameters and their derivatives. Quantities which are particularly informative for our systems include the derivatives of the average number of H-H interactions, $d\langle n_{HH} \rangle / dT$, and those of the numbers of surface contacts, $d\langle n_{SH} \rangle / dT$ and $d\langle n_{SP} \rangle / dT$. A peak in $d\langle n_{HH} \rangle / dT$ signals the construction of H-H interactions to form a hydrophobic core (H-core formation). Peaks in $d\langle n_{SH} \rangle / dT$ and $d\langle n_{SP} \rangle / dT$ provide information about the formation of surface contacts, which is associated with the adsorption process as well as the "flattening" of structure due to surface attraction.

We observe that $\langle R_{ee} \rangle$ behaves quite similarly as $\langle R_g \rangle$ but is less reliable at low temperature where mainly compact structures are found. For this reason our analysis relies on $\langle R_g \rangle$ rather than on $\langle R_{ee} \rangle$.

1. *Folding behavior with a strong attractive surface: Category I*

FIG. 3 shows a typical transition pattern of Category I demonstrated by the 67mer with surface A2, for which $\varepsilon_{HH} = 1, \varepsilon_{SH} = \varepsilon_{SP} = 2$. It is characterized by two pronounced peaks in C_V, with $\langle R_g \rangle$ attaining its maximum between them as seen in the upper panel of the figure. The nature of transitions to which the two peaks in C_V correspond can be identified by comparing them with $d\langle n_{HH} \rangle / dT$ and $d\langle n_{SH} \rangle / dT$ in the lower panel. Since the surface attracts both types of monomers equally, $d\langle n_{SP} \rangle / dT$ shows similar behavior as $d\langle n_{SH} \rangle / dT$ and thus is not shown in the figure. The C_V peak at $k_BT/\varepsilon_{HH} \approx 4.4$ represents the desorption-adsorption transition where $d\langle n_{SH} \rangle / dT$ peaks at the same temperature. The C_V peak at $k_BT/\varepsilon_{HH} \approx 0.4$ represents the H-core formation as $d\langle n_{HH} \rangle / dT$ also shows a peak at that position. $\langle R_g \rangle$ decreases most rapidly during this latter process when temperature is lowered.

A closer look at C_V in FIG. 3 shows a weak bump between $k_BT/\varepsilon_{HH} \approx 1.0$ and 3.5. The same phenomenon is also observed in $d\langle n_{SH} \rangle / dT$ (and $d\langle n_{SP} \rangle / dT$), suggesting that a subtle process related to the interaction with the surface is taking place in this temperature range. Recalling our previous discussion on the radius of gyration $\langle R_g \rangle$ in Section IV A, this is a region where the HP chain keeps forming contacts with the surface until it adsorbs completely on the surface. We call this process a "flattening" of the structure. When the surface attraction is sufficiently strong, it occurs right after the chain is adsorbed to the surface but before the H-core formation. The top graph in the leftmost column in TABLE II shows a similar case for the 48mer, $\varepsilon_{HH} = 1, \varepsilon_{SH} = \varepsilon_{SP} = 2$.

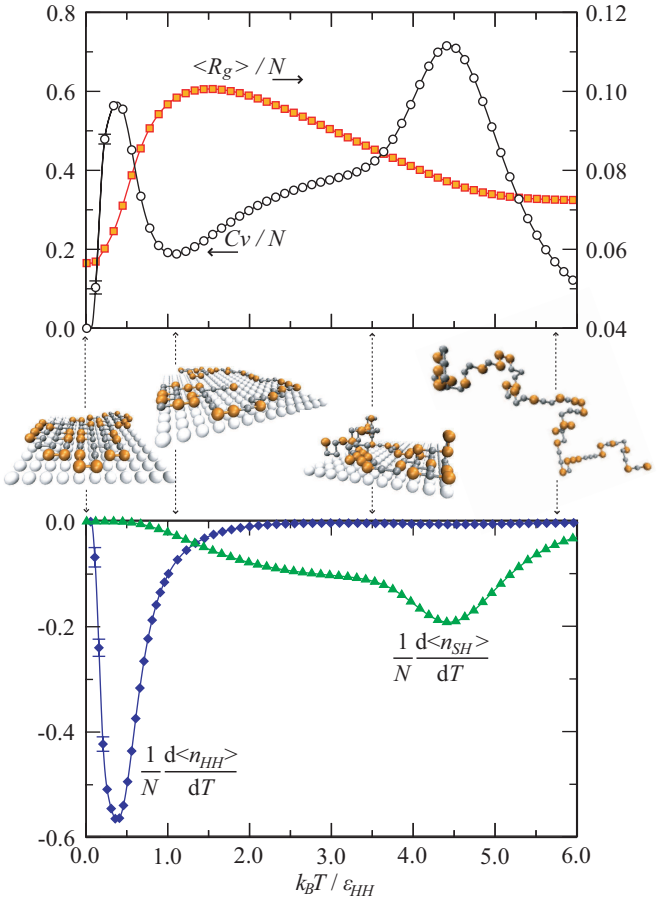


FIG. 3. (Color online) Thermodynamics of the 67mer interacting with surface A2 ($\varepsilon_{HH} = 1, \varepsilon_{SH} = \varepsilon_{SP} = 2$), which shows a typical Category I transition. **Upper panel:** Specific heat, C_V/N, and the average radius of gyration per monomer, $\langle R_g \rangle / N$, as a function of the effective temperature k_BT/ε_{HH}. The horizontal arrows beside the labels indicate the axes to which the quantities refer. **Middle panel:** Typical configurations at different temperatures. **Lower panel:** Derivatives of the average numbers of H-H contacts per monomer, $(1/N)d\langle n_{HH} \rangle / dT$, and that of the average number of surface contacts of H monomers per monomer, $(1/N)d\langle n_{SH} \rangle / dT$, as a function of k_BT/ε_{HH}. Errors smaller than the data points are not shown.

However, there are cases where this "flattening" bump is not observed in C_V, as seen from the other two examples for the 67mer and 103mer with different surface attractions in TABLE II. The flattening process might have been "integrated" within adsorption, or it simply does not cause enough energy fluctuations to give a visible signal in C_V. In the latter case, signals can be found in other structural quantities like $d\langle n_{SH} \rangle / dT$ or

$d\langle n_{SP}\rangle/dT$.

2. *Folding behavior with a moderately attractive surface: Category II*

FIG. 4 shows the thermodynamics for the 103mer with surface P1/2, a typical case in Category II. Similar to Category I, systems in Category II also show two pronounced peaks in C_V and identification of structural transitions depends on the derivatives of $\langle n_{HH}\rangle$, $\langle n_{SH}\rangle$ and $\langle n_{SP}\rangle$. The peak at $k_BT/\varepsilon_{HH} \approx 0.85$ represents the desorption-adsorption transition as identified by the peaks in $d\langle n_{SH}\rangle/dT$ and $d\langle n_{SP}\rangle/dT$. Another peak at $k_BT/\varepsilon_{HH} \approx 0.42$ indicates the H-core formation as signaled by a peak in $d\langle n_{HH}\rangle/dT$.

The feature that differentiates Category II from Category I is the absence of a maximum for $\langle R_g\rangle$ between the two peaks in C_V. It decreases upon cooling until the very low temperature regime. The difference arises from the fact that the flattening of structures occurs at a lower temperature than the H-core formation in the vicinity of a less attractive surface, giving rise to another transition hierarchy than Category I. Two possibilities for $\langle R_g\rangle$ are then observed when the temperature is further lowered: (a) it keeps descending as in FIG. 4; (b) it rises back up until $T = 0$, forming a minimum below the H-core formation temperature as the 48mer does in TABLE II (top graph, middle column for Category II).

Interesting observations at low temperature are revealed by the thermodynamics of $d\langle n_{HH}\rangle/dT$, $d\langle n_{SH}\rangle/dT$ and $d\langle n_{SP}\rangle/dT$ as shown in the lower panel of FIG. 4. During H-core formation at $k_BT/\varepsilon_{HH} \approx 0.42$ where $d\langle n_{HH}\rangle/dT$ peaks at, troughs are observed in $d\langle n_{SH}\rangle/dT$ and $d\langle n_{SP}\rangle/dT$. This is a process of "thickening" during which some of the surface attachments have to be broken to facilitate the construction of H-H interactions.

When the temperature is further lowered to $k_BT/\varepsilon_{HH} \approx 0.25$, a subtle shoulder could barely be seen in C_V and $\langle R_g\rangle$ stays still on cooling; $d\langle n_{SP}\rangle/dT$, however, shows a clear peak. This suggests that surface contacts for the P monomers are established, demonstrating the flattening effect. Eventually the structures with minimal possible energy are attained but they no longer span as many layers vertically as at higher temperature. These structures are not completely planar as in Category I, as forming surface contacts is not always more energetically favorable than forming hydrophobic H-H interactions.

In many other examples of this transition category (see TABLE II), C_V only has two major transition peaks, sometimes with a subtle shoulder or a spike merged into the peaks as a result of a combination of various events. Individual investigation of structural measures is thus essential to segregate different structural changes.

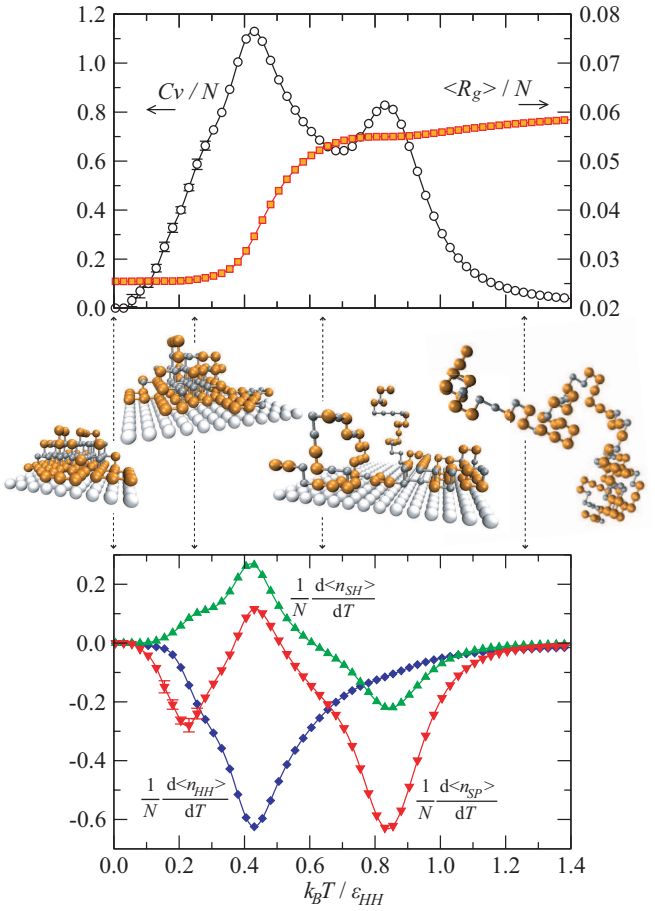


FIG. 4. (Color online) Thermodynamics of the 103mer interacting with surface P1/2 ($\varepsilon_{HH} = 2, \varepsilon_{SH} = 0, \varepsilon_{SP} = 1$), which shows a typical Category II transition. **Upper panel:** Specific heat, C_V/N, and the average radius of gyration per monomer, $\langle R_g\rangle/N$, as a function of the effective temperature k_BT/ε_{HH}. The horizontal arrows beside the labels indicate the axes to which the quantities refer. **Middle panel:** Typical configurations at different temperatures. **Lower panel:** Derivatives of the average numbers of H-H contacts per monomer, $(1/N)d\langle n_{HH}\rangle/dT$, and those of the numbers of surface contacts, $(1/N)d\langle n_{SH}\rangle/dT$ and $(1/N)d\langle n_{SP}\rangle/dT$, as a function of k_BT/ε_{HH}, respectively. Errors smaller than the data points are not shown.

3. *Folding behavior with a weak attractive surface: Category III and beyond*

When the surface attractive strength further reduces, the adsorption and flattening temperatures decrease accordingly. Category III is identified when the adsorption transition coincides with H-core formation, giving a single peak in C_V associated with a shoulder in some cases like the example shown in FIG. 5. The thermodynamics of Category III transitions looks similar to that in 3D free space. In both cases, $\langle R_g\rangle$ decreases upon cooling and C_V peaks at nearly the same temperature, except

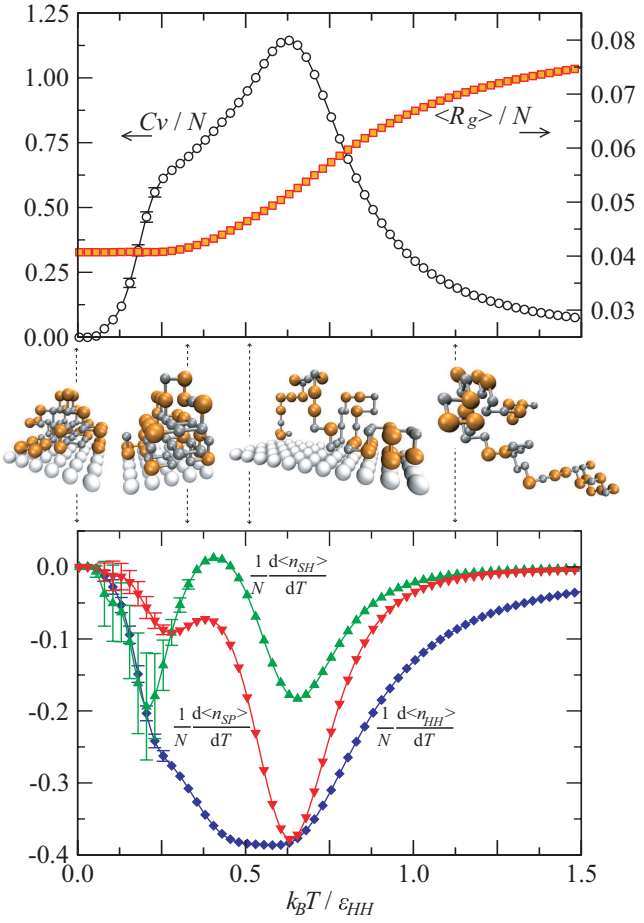


FIG. 5. (Color online) Thermodynamics of the 48mer interacting with surface P1/2 ($\varepsilon_{HH} = 2, \varepsilon_{SH} = 0, \varepsilon_{SP} = 1$), which shows a typical Category III transition. See FIG. 4 for figure explanations.

that a higher peak results for systems of Category III. Since adsorption and H-core formation now occur almost together at nearby temperatures, more conformational degrees of freedom are introduced by the surface interactions and this higher entropy gain results in a larger C_V.

Details of transitions are again provided by $d\langle n_{HH}\rangle/dT$, $d\langle n_{SH}\rangle/dT$ and $d\langle n_{SP}\rangle/dT$. From the positions of peaks in $d\langle n_{SH}\rangle/dT$ and $d\langle n_{SP}\rangle/dT$, one may identify adsorption at $k_BT/\varepsilon_{HH} \approx 0.63$ and flattening at $k_BT/\varepsilon_{HH} \approx 0.24$ respectively. $d\langle n_{HH}\rangle/dT$ manifests a wide peak across the adsorption and flattening temperatures, which suggests that the hydrophobic core is formed roughly in the temperature range $k_BT/\varepsilon_{HH} \approx 0.25 - 0.87$. Instead of producing individual peaks in C_V, the signals of adsorption and flattening are "bridged" and smoothed out by the H-core formation, giving only a peak with a shoulder in C_V, the shape and location of the latter being often subject to some variability. Therefore, it is necessary to rely again on structural quantities to separate signals for the various transitions as illustrated in the previous categories.

For very weak attractive surfaces, adsorption and flattening occur at even lower temperatures. They become distinguishable from the H-core formation which takes place at a higher temperature, forming two or even three distinct peaks in C_V. We generally classify systems with this transition hierarchy as Category IV. For a detailed discussion of an example from this category, see [53, 61].

V. DISCUSSION

A. Remarks on the structural measures and categories

Our results demonstrate that a comprehensive analysis of the specific heat (C_V) combined with a set of appropriate structural quantities is essential to shed light on recognizing structural transformations, especially those subtle ones for which C_V alone provides insufficient information. We also note that in identifying phase transitions, the peaks observed in structural quantities and those in C_V might be slightly off. One possible explanation could be finite size effects [62]. Nevertheless, this does not affect our identification scheme much as these shifts are sufficiently small compared to the difference in temperature scales required to clearly identify distinct categories of transition patterns.

TABLE II further summarizes and compares the thermodynamic features for the three categories discussed in the previous section. We have intentionally chosen examples of various combinations of surface types (i.e. surface strengths) and chain lengths for each category to illustrate that the classification scheme is generic regardless of these variations.

B. Classification of categories using relative surface attraction strengths

Our classification scheme effectively generalized the folding behavior into certain transition hierarchies. We further observe that the dominating factor determining the transition category is closely associated with the relative surface attraction, specifically, the ratio between $\varepsilon_{SH} + \varepsilon_{SP}$ and ε_{HH}.

TABLE III shows the distribution of transition categories against the relative surface attractions for systems with various chain lengths and surface types. Ideally, a perfect correspondence between the transition categories and the relative surface attractions is implied if only the diagonal compartments were filled in Table III. In reality, as thermodynamic subtleties vary from sequence to sequence, some off-diagonal compartments are also occupied (e.g., some systems with $\varepsilon_{SH} + \varepsilon_{SP} < \varepsilon_{HH}$ show Category II behavior). A few systems also reveal "category duality" where the thermodynamics bears properties from both categories (e.g. the 103mer interact-

TABLE II. Characteristic thermodynamics for Categories I, II, and III: Specific heat, C_V/N, and average radius of gyration per monomer, $\langle R_g \rangle /N$, as a function of the effective temperature k_BT/ε_{HH}. Errors smaller than the data points are not shown.

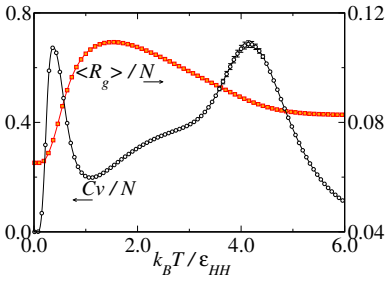
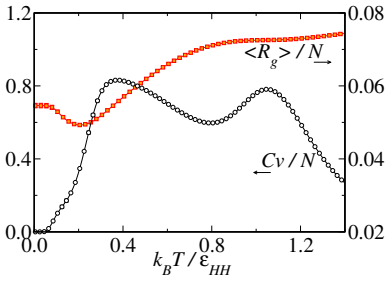
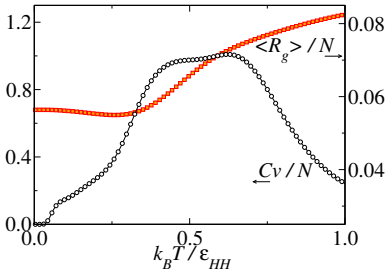
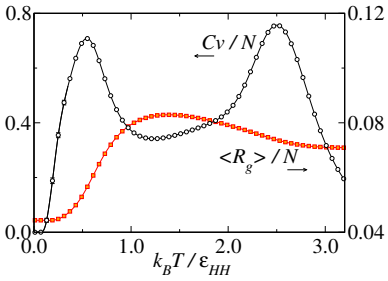
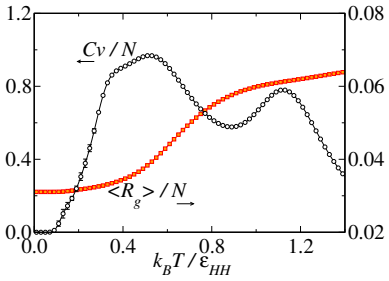
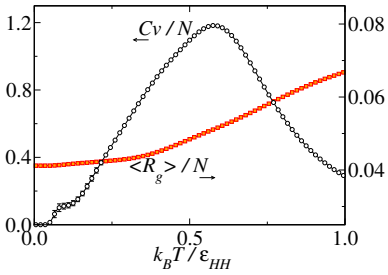
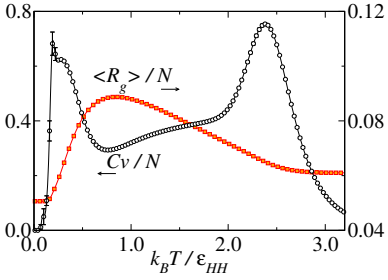
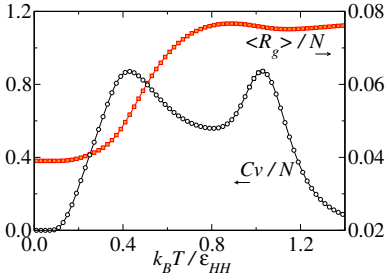
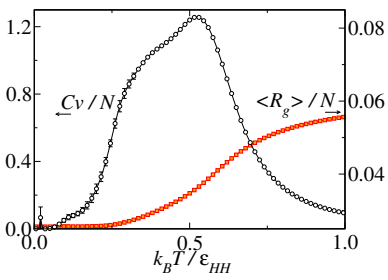
	Category I	**Category II**	**Category III**
	48mer, $\varepsilon_{HH}=1, \varepsilon_{SH}=2, \varepsilon_{SP}=2$ 	48mer, $\varepsilon_{HH}=2, \varepsilon_{SH}=1, \varepsilon_{SP}=1$ 	36mer, $\varepsilon_{HH}=3, \varepsilon_{SH}=1, \varepsilon_{SP}=1$ 
	67mer, $\varepsilon_{HH}=1, \varepsilon_{SH}=2, \varepsilon_{SP}=0$ 	67mer, $\varepsilon_{HH}=1, \varepsilon_{SH}=0, \varepsilon_{SP}=1$ 	48mer, $\varepsilon_{HH}=2, \varepsilon_{SH}=1, \varepsilon_{SP}=0$ 
	103mer, $\varepsilon_{HH}=1, \varepsilon_{SH}=1, \varepsilon_{SP}=1$ 	103mer, $\varepsilon_{HH}=1, \varepsilon_{SH}=1, \varepsilon_{SP}=0$ 	103mer, $\varepsilon_{HH}=2, \varepsilon_{SH}=1, \varepsilon_{SP}=0$ 
C_V	• shows two peaks, one for adsorption at high T, one for H-core formation at low T • might show a bump for flattening between the two peaks	• shows two peaks, one for adsorption at high T, one for H-core formation at low T • might have a shoulder or a spike for flattening below the H-core formation T	• shows only one peak for a combination of adsorption, flattening and H-core formation • peak might have one or more shoulders for processes taking place at nearby T
$\langle R_g \rangle$	• shows a global maximum between the two C_V peaks, signifying the end of flattening and the start of H-core formation upon cooling	• decreases upon cooling, might form a local maximum between the two C_V peaks • might rise back up due to flattening after H-core formation, forming a minimum at low T	• keeps decreasing upon cooling • might rise back up due to flattening after H-core formation, forming a minimum at low T

TABLE III. Distribution of transition categories with respect to the relative surface attractions. The abbreviations refer to the surface types introduced in TABLE I. The numbers are the short forms of the benchmark sequences (e.g. 36 stands for the 36mer, 48.1 and 48.9 correspond to Seq. 1 and 9 among the ten 48mers, respectively).

ε_{SH} + ε_{SP}	Category I	Category II	Category III
$> \varepsilon_{HH}$	A1: 36, 48.1, 48.9, 64, 67, 103 A2: 36, 48.9, 64, 67 H2: 48.1, 48.9, 64, 67 P2: 48.1, 48.9, 64, 67		
$= \varepsilon_{HH}$	A1/2: 103 P1: 103	A1/2: 36, 48.1, 48.9, 64, 67 H1: 48.1, 48.9, 64, 67, 103 P1: 48.1, 48.9, 64, 67	
$< \varepsilon_{HH}$		H1/2: 67 P1/2: 67, 103	H1/2: 48.1, 48.9, 64, 103 P1/2: 48.1, 48.9 A1/3[a]: 36, 64

[a] $\varepsilon_{SH} = \varepsilon_{SP} = 1, \varepsilon_{HH} = 3$

ing with surfaces A1/2 or P1). Nonetheless, the generality of our classification scheme is clearly apparent and allows us to infer the following basic rules: transition Category I occurs for surfaces which are strongly attractive ($\varepsilon_{SH} + \varepsilon_{SP} > \approx \varepsilon_{HH}$); Category II occurs when the hydrophobic internal attraction is approximately comparable to the surface strengths ($\varepsilon_{SH} + \varepsilon_{SP} \approx \varepsilon_{HH}$); Category III can only occur when surface strengths are relatively weak compared to ε_{HH} ($\varepsilon_{SH} + \varepsilon_{SP} < \varepsilon_{HH}$). Besides, we have also investigated the influence of the proportion of monomers by considering the ratio between $(n_{SH}/N)\varepsilon_{SH} + (n_{SP}/N)\varepsilon_{SP}$ and ε_{HH}. However, we did not find an obvious relation between these quantities and the transition categories. We thus believe that $\varepsilon_{SH}+\varepsilon_{SP}$ is more suitable for the classification scheme.

We stress that unlike other existing work, this classification is an inference based on multiple HP sequences of various chain lengths and attributes. In TABLE III, we have also included results from three other sequences, which were used as a "testing set" for the adequacy of the classification scheme: a 36mer ($P_3H_2P_2H_2P_5H_7P_2H_2P_4H_2P_2HP_2$); another 48mer ($HPH_2P_2H_4PH_3P_2H_2P_2HPH_3PHPH_2P_2H_2P_3HP_8H_2$); and a 64mer ($H_{12}PHPHP_2H_2P_2H_2P_2HP_2H_2P_2H_2P_2HP_2H_2P_2H_2P_2HPHPH_{12}$). The 36mer and the 64mer were originally proposed to test a 2D genetic algorithm [26], whereas the 48mer is Seq. 1 of the ten testing sequences in [57]. All results from these additional sequences fall into the diagonal compartments in TABLE III, reinforcing that our classification scheme is applicable to other sequences interacting with an adsorbing substrate without loss of generality. This is thus a breakthrough in our understanding of adsorption properties of lattice proteins: Instead of sequence-dependent individual behavior, the thermodynamics of HP proteins do follow common patterns in structural transitions when they interact with an adsorbing substrate.

Even if ε_{SH} and ε_{SP} have the same magnitude, they are generally not expected to contribute equivalently to the total energy because of the competition between ε_{SH} and ε_{HH} for the H monomers. However, the contrary appears to be the case in our study because of the entropic effects, in which the adsorbing P monomers also hinder the formation of H-H contacts in an indirect manner by dragging the H monomers to the surface. Consider the surface P2 case for instance: As it is energetically more favorable to form surface-P contacts than internal H-H contacts, the chain tends to sit on the surface rather than to form a compact globule. The H monomers are then restricted to sit also on the surface instead of forming H-H contacts, causing both $d\langle n_{SH}\rangle/dT$ and $d\langle n_{SP}\rangle/dT$ to peak at the adsorption temperature.

VI. CONCLUSION

In this work, protein adsorption has been studied with a coarse-grained lattice model, the HP model, interacting with a surface which either attracts all monomers, only hydrophobic monomers or only polar monomers. We have employed Wang-Landau sampling with two effective Monte Carlo trial moves, pull moves and bond-rebridging moves, to obtain the energy density of states and subsequent thermodynamics of structural quantities for a 48mer, 67mer and 103mer. Ground state energies are also reported. Based on the folding and adsorption behavior revealed by a careful, comprehensive analysis of the specific heat, radius of gyration and derivatives of the numbers of surface contacts, we have been able to identify four main types of transition hierarchies (three of which are discussed in detail in this work). We have found that the occurrence of these transition hierarchies is mainly determined by the attractive couplings of the surface relative to the internal hydrophobic attraction, i.e., the ratio between $\varepsilon_{SH} + \varepsilon_{SP}$ and ε_{HH}, regardless of the surface type, chain length or composition of H and P monomers of an HP sequence. Three other benchmark sequences, a 36mer, another 48mer and a 64mer, have confirmed the validity of our classification scheme.

Although other transition categories cannot be excluded from our study, the ones presented here provide a general and representative picture of the thermodynam-

ics of HP proteins interacting with an adsorbing substrate. Our study also demonstrates that classifying transition hierarchies by a combined analysis of the specific heat and appropriate structural parameters provides a powerful route in approaching similar systems of large conformational and sequence spaces, for instance, HP proteins interacting with two confining, attractive surfaces [63, 64].

However, further investigation is necessary to determine if there is a rigorous relation between the proposed transition categories and the relative surface attractions. More statistics from longer chains, or chains of the same length but with different H and P compositions would help clarifying the problem. The next question is whether the same conclusions could be drawn, or what discrepancies would be found, for other lattice models with other energy functions, i.e., different interactions between monomers and with the surface. Another important question is whether similar classification schemes could hold for off-lattice models. In this case, thermodynamics of other structural parameters, e.g., the gyration tensor, density profile, or any suitable ones, could help in identifying such generic transition patterns. All these together are essential in determining the effectiveness of using different simplified protein models in computer simulations to study protein adsorption from a macroscopic perspective.

ACKNOWLEDGMENTS

We thank M. Bachmann for constructive discussions. This work was supported by the National Science Foundation under grant no. DMR-0810223.

[1] V. Hlady and J. Buijs, Curr. Opin. Biotech. **7**, 72 (1996).
[2] J. J. Gray, Curr. Opin. Struc. Biol. **14**, 110 (2004).
[3] G. MacBeath and S. L. Schreiber, Science **289**, 1760 (2000).
[4] E. Phizicky, P. I. Bastiaens, H. Zhu, M. Snyder and S. Fields, Nature **422**, 208 (2003).
[5] S. Brown, Nat. Biotechnol. **15**, 269 (1997).
[6] S. R. Whaley, D. S. English, E. L. Hu, P. F. Barbara and A. M. Belcher, Nature **405**, 665 (2000).
[7] K. Goede, P. Busch and M. Grundmann, Nano Lett. **4**, 2115 (2004).
[8] M. Bachmann, K. Goede, A. G. Beck-Sickinger, M Grundmann, A. Irbäck and W. Janke, Angew. Chem. Int. Ed. **49**, 9530 (2010).
[9] J. D. Andrade and V. Hlady, Adv. Polym. Sci. **79**, 1 (1986).
[10] T. A. Horbett, Surfactant Science Series **110**, 393 (2003).
[11] W. R. Gombotz and D. K. Pettit, Bioconjugate Chem. **6**, 332 (1995).
[12] C. Pinholt, R. A. Hartvig, N. J. Medlicott and L. Jorgensen, Expert Opin. Drug Deliv. **8**, 949 (2011).
[13] M. Rabe, D. Verdes and S. Seeger, Adv. Colloid Interface Sci. **162**, 87 (2011).
[14] J. J. Ramsden, Q. Rev. Biophys. **27**, 41 (1994).
[15] V. Hlady, J. Buijs and H. P. Jennissen, Methods Enzymol. **309**, 402 (1999).
[16] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, 2009), 3rd ed.
[17] F. Ganazzoli and G. Raffaini, Phys. Chem. Chem. Phys. **7**, 3651 (2005).
[18] K. A. Dill, Biochemistry **24**, 1501 (1985).
[19] K. F. Lau and K. A. Dill, Macromolecules **22**, 3986 (1989).
[20] H. S. Chan, Proteins **40**, 543 (2000).
[21] P. De Los Rios and G. Caldarelli, Phys. Rev. E **62** 8449 (2000).
[22] B. Berger and T. Leighton, J. Comput. Biol. **5**, 27 (1998).
[23] P. Crescenzi, D. Goldman, C. Papadimitriou, A. Piccolboni and M. Yannakakis, J. Comput. Biol. **5**, 423 (1998).
[24] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, J. Chem. Phys. **21**, 1087 (1953).
[25] Y. W. Li, T. Wüst and D. P. Landau, JPCS, in press.
[26] R. Unger and J. Moult, J. Mol. Biol. **231**, 75 (1993).
[27] H. Frauenkron, U. Bastolla, E. Gerstner, P. Grassberger and W. Nadler, Phys. Rev. Lett. **80**, 3149 (1998).
[28] G. Chikenji, M. Kikuchi and Y. Iba, Phys. Rev. Lett. **83**, 1886 (1999).
[29] H. -P. Hsu, V. Mehra, W. Nadler and Peter Grassberger, J. Chem. Phys. **118**, 444 (2003).
[30] H. -P. Hsu, V. Mehra, W. Nadler and Peter Grassberger, Phys. Rev. E **68**, 021113 (2003).
[31] T. Prellberg and J. Krawczyk, Phys. Rev. Lett. **92**, 120602 (2004).
[32] M. Bachmann and W. Janke, J. Chem. Phys. **120**, 6779 (2004).
[33] S. C. Kou, J. Oh and W. H. Wong, J. Chem. Phys. **124**, 244903 (2006).
[34] J. Zhang, S. C. Kou and J. S. Liu, J. Chem. Phys. **126**, 225101 (2007).
[35] S. M. Liu and C. A. Haynes, J. Colloid Interface Sci. **284**, 7 (2005).
[36] S. M. Liu and C. A. Haynes, J. Colloid Interface Sci. **282**, 283 (2005).
[37] V. Castells, S. Yang and P. R. Van Tassel, Phys. Rev. E **65**, 031912 (2002).
[38] V. Castells and P. R. Van Tassel, J. Chem. Phys. **122**, 084707 (2005).
[39] M. Bachmann and W. Janke, Phys. Rev. E **73**, 020901 (R) (2006).
[40] A. D. Swetnam and M. P. Allen, Phys. Chem. Chem. Phys. **11**, 2046 (2009).
[41] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).
[42] F. Wang and D. P. Landau, Phys. Rev. E **64**, 056101 (2001).
[43] D. P. Landau, S. -H. Tsai and M. Exler, Am. J. Phys. **72**, 1294 (2004).

[44] M. Radhakrishna, S. Sharma and S. K. Kumar, J. Chem. Phys. **136**, 114114 (2012).
[45] M. Bachmann and W. Janke, Lect. Notes Phys. **736**, 203 (2008).
[46] N. Lesh, M. Mitzenmacher and S. Whitesides, in *RECOMB*, p. 188 (2003).
[47] J. M. Deutsch, J. Chem. Phys. **106**, 8849 (1997).
[48] T. Wüst and D. P. Landau, Comput. Phys. Commun. **179**, 124 (2008).
[49] T. Wüst and D. P. Landau, Phys. Rev. Lett. **102**, 178101 (2009).
[50] T. Wüst and D. P. Landau, J. Chem. Phys. **137**, 064903 (2012).
[51] M. Galassi *et al.*, *GNU Scientific Library Reference Manual* (Network Theory Ltd., 2009), 3rd ed., p. 223.
[52] M. Lüscher, Comput. Phys. Commun. **79**, 100 (1994).
[53] Y. W. Li, T. Wüst and D. P. Landau, Comput. Phys. Commun. **182**, 1896 (2011).
[54] A. D. Swetnam and M. P. Allen, J. Comput. Chem. **32**, 816 (2011).
[55] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991).
[56] B. A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).
[57] K. Yue, K. M. Fiebig, P. D. Thomas, H. S. Chan, E. I. Shakhnovich and K. A. Dill, Proc. Natl. Acad. Sci. USA **92**, 325 (1995).
[58] K. Yue and K. A. Dill, Proc. Natl. Acad. Sci. USA **92**, 146 (1995).
[59] E. E. Lattman, K. M. Fiebig and K. A. Dill, Biochemistry **33**, 6158 (1994).
[60] Y. W. Li, T. Wüst and D. P. Landau, in *Nanophenomena at Surfaces: Fundamentals of Exotic Condensed Matter Properties*, Springer Series in Surface Sciences, edited by M. Michailov (Springer-Verlag, Berlin, Heidelberg, 2011), Vol. 47 , Chap. 7, p. 169.
[61] T. Wüst, Y. W. Li and D. P. Landau, J. Stat. Phys. **144**, 638 (2011).
[62] S. Sharma and S. K. Kumar, J. Chem. Phys. **129**, 134901 (2008).
[63] B. Pattanasiri, Y. W. Li, D. P. Landau and T. Wüst, Int. J. Mod. Phys. C **23**, 1240008 (2012).
[64] B. Pattanasiri, Y. W. Li, D. P. Landau, T. Wüst and W. Triampo, JPCS, in press.