Structural properties of small semiconductor-binding synthetic peptides

Gökhan Gökoğlu,[1,2,3,*] Michael Bachmann,[1,3,†] Tarık Çelik,[2,3,‡] and Wolfhard Janke[1,3,§]

[1]*Institut für Theoretische Physik, Universität Leipzig, Augustusplatz 10/11, D-04109 Leipzig, Germany*
[2]*Hacettepe Üniversitesi, Fizik Mühendisliği Bölümü, Beytepe 06800 Ankara, Turkey*
[3]*Centre for Theoretical Sciences (NTZ) of the Centre for Advanced Study (ZHS), Universität Leipzig, Emil-Fuchs-Straße 1, D-04105 Leipzig, Germany*

We have performed exhaustive multicanonical Monte Carlo simulations of three 12-residue synthetic peptides in order to investigate the thermodynamic and structural properties as well as the characteristic helix-coil transitions. In these studies, we employ a realistic model where the interactions between all atoms are taken into account. Effects of solvation are also simulated by using an implicit-solvent model.

I. INTRODUCTION

It is well known that three-dimensional (3D) conformations of peptides and proteins play an important role due to their biological activities. Therefore, many theoretical and experimental studies focus on the determination of the 3D structure of these molecules. In a newly growing field of research, synthetic peptides are investigated for use in hybrid nano-devices, depending on their self-assembly properties [1,2]. In these studies, it is also shown that the binding of peptides on metal and semiconductor surfaces depends on the types of amino acids [3] and on the sequences of the residues in the peptide chain [4–6]. These experiments reveal many different interesting and important problems, which are related to general aspects of the question why and how proteins fold. This regards, for example, the character of the adsorption process—i.e., whether the peptides simply dock to the substrate without noticeable structural changes or whether they perform conformational transitions before binding. A related question is how secondary structures of peptide folds in the bulk influence the binding behavior to substrates. In helical structures, for example, side chains are radially directed and—due to the helical symmetry— residues with a certain distance in the sequence arrange linearly. This could have consequences for docking to a regular crystal surface, where the atoms are also arranged linearly along the main axes. This means that two peptides with the same content of residues, but different sequences, could exhibit completely different binding properties. This behavior was, in fact, observed in a recent experimental adsorption study of peptides in the vicinity of semiconductor substrates [6], although other explanations for this kind of specificity are also conceivable. For these reasons, it is likely that the binding specificity also depends on the thermodynamic and structural properties of peptides in solvent, as has already turned out in investigations of minimalistic models [7,8].

In this paper, we focus on three synthetic peptides, AQNPSDNNTHTH, AQNPSDNNTATA, and TNHDH-

SNAPTNQ [9], whose binding properties were investigated in recent experimental studies [5,6]. The second sequence is a mutated version of the first one, where the histidine residues (H) of the first chain are replaced by alanines (A). The third sequence is a randomly permuted sequence of the first chain. It was shown in the experiments that the first sequence has a strong affinity to adsorb to gallium arsenide (GaAs), whereas the binding to silicon (Si) is very weak. Exchanging the histidines by alanines improves the binding properties to Si, while the adsorption strength to GaAs is noticeably reduced. For the randomly permuted sequence, the binding strength to GaAs is left widely unaffected, while binding to Si is as strong as to GaAs.

Employing multicanonical (MUCA) Monte Carlo sampling [10,11], we analyze single-molecule folds in the bulk and illuminate the thermodynamic and structural properties of these peptides. Generalized-ensemble methods applied to all-atom descriptions of proteins have been very successful in the past—e.g., in revealing the statistical mechanics in the folding process of small proteins [12–14]. For sequences with more than 20 residues, studies of thermodynamics and kinetics employing realistic physical models are computationally extremely demanding. For such systems, reduced all-atom models [15] or models at a higher coarse-grained level [16,17] could be, depending on the particular question, much more promising. Coarse-grained lattice and off-lattice models allow for systematic thermodynamic studies and, at least partly, sequence analyses of heteropolymers with up to 100 monomers [18–20].

First, we have simulated all three molecules in vacuum. Then, in order to see the effects of solvation, we have also performed extensive simulations of the commonly used surface-accessible area solvent model with the Ooi-Obatake-Némethy-Scheraga (OONS) atomic solvation parameter set [21]. The preferential properties of this parameter set compared to others were reported in previous works [22,23].

The rest of the paper is organized as follows. After the description of the peptide model and the computational methods used in this study in Sec. II, we discuss in Sec. III exemplified effects of solvation compared with results obtained in the vacuum simulations. In Sec. IV, we discuss in detail the helix-coil transitions for the three peptides in solvent by means of fluctuations of several energetic and structural quantities. Section V addresses the folding channels in

---

*Electronic address: ggokoglu@hacettepe.edu.tr
†Electronic address: Michael.Bachmann@itp.uni-leipzig.de
‡Electronic address: tcelik@hacettepe.edu.tr
§Electronic address: Wolfhard.Janke@itp.uni-leipzig.de

the free-energy landscape. The paper concludes with a summary in Sec. VI.

II. PEPTIDE MODEL AND SIMULATION METHOD

A. Peptide model

In our simulations a peptide is modeled with all of its atoms. Each atom i, located at the position $\mathbf{r}_i$, carries a partial charge q_i. Covalent bonds between atoms, according to the chemical structure of the amino acids, are considered rigid; i.e., bond lengths are kept constant, as well as bond angles between covalent bonds and certain rigid torsion angles. Distances between nonbonded atoms i and j are defined as $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ and measured in Å in the following. The set of degrees of freedom covers all dihedral torsion angles $\boldsymbol{\xi} = \{\xi_\alpha\}$ of αth residue's backbone (ϕ_α, ψ_α, ω_α) and side chain ($\boldsymbol{\chi} = \chi_\alpha^{(1)}, \chi_\alpha^{(2)}, \dots$). The model incorporates electrostatic Coulomb interactions between the partial atomic charges (all energies in kcal/mol),

$$E_C(\boldsymbol{\xi}) = 332 \sum_{i,j} \frac{q_i q_j}{\varepsilon r_{ij}(\boldsymbol{\xi})}, \tag{1}$$

effective atomic dipole-dipole interaction modeled via Lennard-Jones potentials [24],

$$E_{LJ}(\boldsymbol{\xi}) = \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}(\boldsymbol{\xi})} - \frac{B_{ij}}{r_{ij}^6(\boldsymbol{\xi})} \right), \tag{2}$$

O-H and N-H hydrogen-bond formation,

$$E_{HB}(\boldsymbol{\xi}) = \sum_{i,j} \left(\frac{C_{ij}}{r_{ij}^{12}(\boldsymbol{\xi})} - \frac{D_{ij}}{r_{ij}^{10}(\boldsymbol{\xi})} \right), \tag{3}$$

and considers dihedral torsional barriers (if any):

$$E_{tor}(\boldsymbol{\xi}) = \sum_l U_l[1 \pm \cos(n_l \xi_l)]. \tag{4}$$

The total energy of a conformation, whose structure is completely defined by the set of dihedral angles $\boldsymbol{\xi}$, is

$$E_0(\boldsymbol{\xi}) = E_C(\boldsymbol{\xi}) + E_{LJ}(\boldsymbol{\xi}) + E_{HB}(\boldsymbol{\xi}) + E_{tor}(\boldsymbol{\xi}). \tag{5}$$

The parameters q_i, A_{ij}, B_{ij}, C_{ij}, D_{ij}, U_l, and n_l are taken from the ECEPP/3 (empirical conformational energies for proteins and polypeptides) force field [25], one of the most commonly used all-atom force fields. In all simulations the dielectric constant was set to $\varepsilon = 2$, which is the vacuum value. Proline's ϕ is considered rigid at $-68.8°$. We always used the *trans*-down-puckering conformation of the proline ring. For the implicit-solvent simulations, the model is extended by the solvation-energy contribution, which is given by [26]

$$E_{solv}(\boldsymbol{\xi}) = \sum_i \sigma_i A_i(\boldsymbol{\xi}), \tag{6}$$

where A_i is the solvent-accessible surface area of the ith atom for a given conformation and σ_i is the solvation parameter for the ith atom. The values for σ_i depend on the type of the ith atom and are parametrized according to the suggestions given in Ref. [21]. The total potential energy of the molecule then reads

TABLE I. Peptide sequences studied in this work.

S1	AQNPSDNNTHTH
S2	AQNPSDNNTATA
S3	TNHDHSNAPTNQ

$$E_{tot}(\boldsymbol{\xi}) = E_0(\boldsymbol{\xi}) + E_{solv}(\boldsymbol{\xi}). \tag{7}$$

The described peptide model and the ECEPP/3 parametrization are implemented in the software package SMMP [27], which we used for our study.

In Table I, we list the three sequences investigated with this model.

B. Multicanonical sampling

Multicanonical sampling [10,11,28] is a generalized-ensemble method, in which conformations are ideally sampled according to a flat energy distribution $p_{muca}(E) = \text{const}$; i.e., the Markovian dynamics of the algorithm corresponds to a random walk in energy space. The desired canonical distribution at a certain temperature T is given by $p_{can}(E,T) \sim n(E)\exp(-E/RT)$, where $n(E)$ is the density of states and the gas constant takes the value $R \approx 1.99 \times 10^{-3}$ kcal/(K mol) in the units used in this paper. Since canonical and multicanonical energy distributions are trivially related via $p_{can}(E,T) \sim W_{muca}^{-1}(E)\exp(-E/RT)p_{muca}(E)$, the main task is a precise determination of the multicanonical weights $W_{muca}(E) \sim n^{-1}(E)$.

The implementation of MUCA is not straightforward as the density of states $n(E)$ is unknown *a priori*. Therefore, the weights $W_{muca}(E)$ have to be determined in the first stage of the simulation process by an iterative procedure until the multicanonical histogram $H(E) \approx \text{const}$ in the desired energy interval. An efficient, error-weighted estimation method for the multicanonical weights is described in detail in Refs. [11,28]. We note that the efficiency of the determination of the multicanonical weights usually depends on the choice of the simulation temperature, which was in the present study $T_{sim} = 1000$ K. The reason is that, since the "flat" energy histogram covers a larger region in subsequent recursions, energetic states are hit for the first time, where the multicanonical weights are still undetermined. Because the ratio of the weights controls the acceptance of a conformational update, the dynamics of the recursion part of the algorithm is noticeably influenced. This behavior can be "smoothed" by a careful choice of the simulation temperature.

Eventually, in the second stage of the multicanonical simulation, a long production run is performed based on fixed multicanonical weights. Since the weights for all energetic states in the desired energy range have already been determined, the choice of a certain simulation temperature is unnecessary.

In our concrete implementation, we first carried out Metropolis simulations at relatively high simulation temperatures and MUCA test runs which enabled us to determine the required energy range. This interval was then divided into bins of 1 kcal/mol. At each update step, a trial conformation
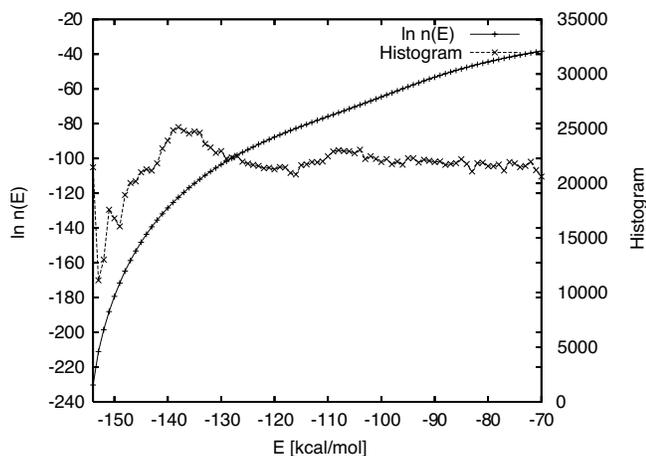
FIG. 1. Natural logarithm of the density of states $n(E)$ and multicanonical histogram from the simulation of the sequence AQNPSDNNTHTH ($S1$) in solvent.

was obtained by changing a dihedral angle $\xi_i \to \xi_i'$ within the range $[-180°, 180°]$, which was accepted according to the transition probability $\omega(\boldsymbol{\xi} \to \boldsymbol{\xi}') = \min\{\exp[S(E(\boldsymbol{\xi})) - S(E(\boldsymbol{\xi}'))], 1\}$, where $S(E(\boldsymbol{\xi})) = -\ln W_{\mathrm{muca}}(E(\boldsymbol{\xi}))$ can be identified with the microcanonical entropy. The dihedral angles were always visited in a predefined, sequential order; i.e., a sweep is a cycle of N Monte Carlo steps (N=total number of dihedral angles).

The weights were built in 200 recursions during a long *single* simulation, where the multicanonical parameters were iterated every 10 000 sweeps. Then, we performed a full simulation of 2×10^6 sweeps with fixed weights, which covers the temperature region up to $T_{\mathrm{max}} = 1000$ K reliably. In Fig. 1, the density of states and the multicanonical histogram for the first sequence considered, AQNPSDNNTHTH ($S1$), is shown. As seen from this figure, the multicanonical histogram is indeed "flat" which is a necessary condition for the multicanonical technique to be reliably working. Statistical expectation values for any thermodynamic quantity $\mathcal{A}$ and all temperatures can finally be calculated from the time series recorded during the multicanonical production run:

$$\langle \mathcal{A} \rangle = \frac{\sum_t \mathcal{A}(\boldsymbol{\xi}^{(t)}) W_{\mathrm{muca}}^{-1}(E(\boldsymbol{\xi}^{(t)})) e^{-\beta E(\boldsymbol{\xi}^{(t)})}}{\sum_t W_{\mathrm{muca}}^{-1}(E(\boldsymbol{\xi}^{(t)})) e^{-\beta E(\boldsymbol{\xi}^{(t)})}}, \quad (8)$$

where $\boldsymbol{\xi}^{(t)}$ labels the conformation at "time" t and $\beta = 1/RT$ is the inverse thermal energy.

The derivative of the quantity $\langle \mathcal{A} \rangle$ with respect to the thermal energy is given by

$$\frac{d\langle \mathcal{A} \rangle}{d(RT)} = \frac{1}{(RT)^2}(\langle E\mathcal{A} \rangle - \langle E \rangle \langle \mathcal{A} \rangle). \quad (9)$$

Expressions like this are typically used to discuss the influence of thermal fluctuations on $\mathcal{A}$.
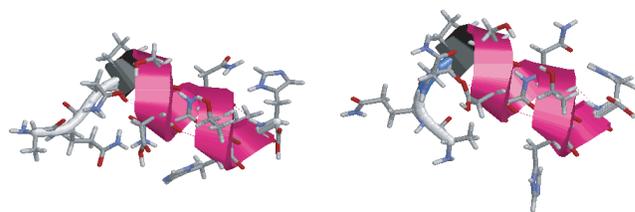
FIG. 2. (Color online) Low-energy reference conformations of the peptide $S1$ in vacuum (left) and solvent (right) for the calculation of the overlap parameter (11).

III. SOLVATION EFFECTS

In this section, we compare the folding behavior of the exemplified peptide $S1$ in vacuum and solvent, respectively. Changes of energy fluctuations—i.e., the widths of the energy distributions—signalize typically a crossover or transition between significantly different macrostates ("phases") of the system considered. For polymers or peptides, the crossover between such macrostates is accompanied by a cooperative conformational transition. Thus it is reasonable to compare the behavior of the peptide in vacuum and solvent with regard to energetic fluctuations, the specific heat (in units of R)

$$C_V = \frac{1}{(RT)^2}(\langle E^2 \rangle - \langle E \rangle^2), \quad (10)$$

and with respect to the angular overlap parameter [29–31] as a structural quantity ("order" parameter), which is suitably defined as

$$Q(\boldsymbol{\xi}^{(t)}, \boldsymbol{\xi}^{\mathrm{ref}}) = 1 - \frac{1}{90° N} \sum_{i=1}^{N} d(\xi_i^{(t)}, \xi_i^{\mathrm{ref}}), \quad (11)$$

where $d(\alpha, \alpha') = \min(|\alpha - \alpha'|, 360° - |\alpha - \alpha'|)$. In this expression, the dihedral angles of the actual conformation $\boldsymbol{\xi}^{(t)}$ are compared with the corresponding torsion angles of a suitable reference conformation, $\boldsymbol{\xi}^{\mathrm{ref}}$. Since the overlap parameter (11) is a measure for the similarity of any conformation and the reference conformation, it can be considered as a system state parameter: $Q = 1$ only if the considered conformation is identical with the reference conformation, which is here chosen to be a typical representative of the (low-energy) helical phase. In Fig. 2, the helical low-energy reference conformations of sequence $S1$ in vacuum (left) and solvent (right) are shown. The main differences regard the side chains and the tail, whereas the helical part is hardly influenced by solvation effects. The overlap Q between these two conformations comparing *all* dihedral angles is 0.595, while considering only the backbone dihedrals, $Q \approx 0.690$. In the latter case, the comparatively still small coincidence is mainly due to the nonhelical tails, which are highly flexible; i.e., there are no stabilizing hydrogen bonds with the backbone. Proline breaks the helix in both cases.

In Fig. 3 we plot the average overlap parameter $\langle Q \rangle$ for this exemplified sequence in vacuum and solvent. For low temperatures, in both cases most of the conformations in the ensemble have similarities with the reference conformations;
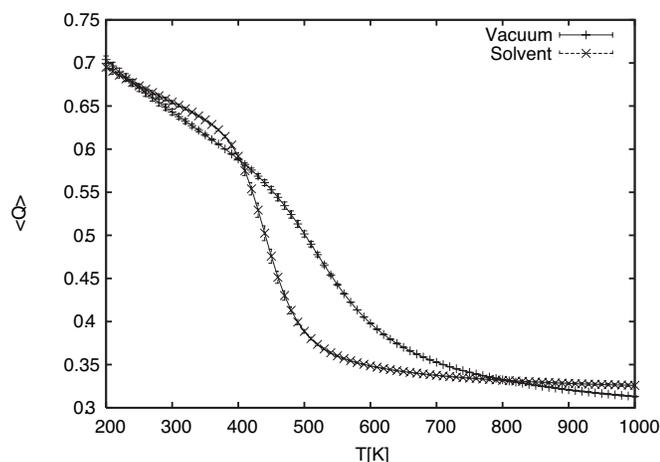
FIG. 3. Average angular overlap Q between ensemble conformations and the reference conformations in Fig. 2 for the peptide $S1$ as a function of temperature in vacuum and solvent.

i.e., the ensemble is dominated by helical conformations. For the peptide in solvent, the average overlap parameter decreases rapidly at about 440 K; the conformations in the high-temperature phase are random coils. The situation is comparable for the peptide in vacuum, with the noticeable difference that $\langle Q \rangle$ decreases much slower at a transition temperature near 530 K. This is confirmed by considering the energetic fluctuations of the system—i.e., the specific heat as shown in Fig. 4. In the vicinity of the peak temperatures, the peptide exhibits conformational activity. The helix-coil transition peak is stronger and sharper for the peptide in solvent; the transition temperature is close to the above-mentioned value.

Summarizing, the main effect of the solvent is the strengthening of the helix-coil transition which is also present in the vacuum case. Furthermore, the transition temperature is shifted by about 100 K towards lower temperatures. These results are as expected, since it is known that solvent stabilizes secondary structures and therefore the barrier to resolve the helix is higher than in the vacuum case and
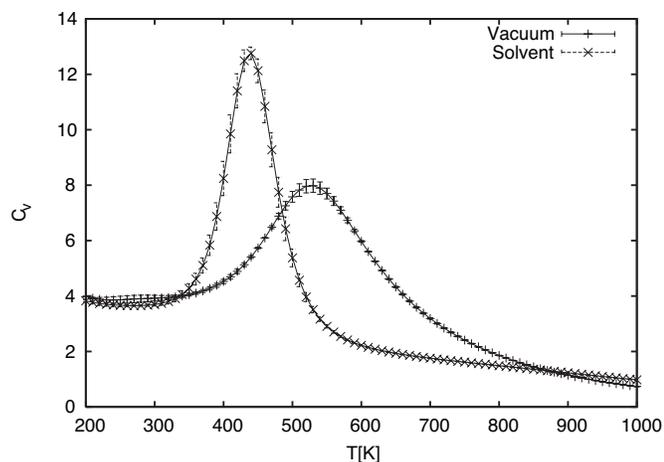
FIG. 4. Specific heat vs temperature for the peptide $S1$ in vacuum and solvent.
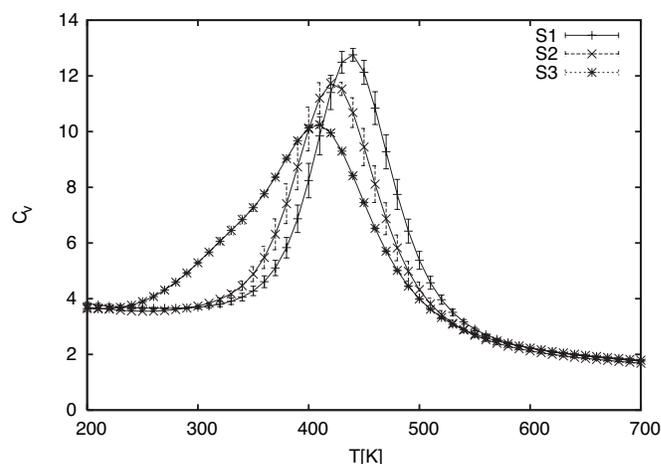
FIG. 5. Specific heat vs temperature for the three peptides in solvent.

the relaxation of the fluctuations of the peptide-solvent coupling degrees of freedom leads to a lower transition temperature. These differences have also been observed in studies using other parameter sets [32].

It should be noted that also in the OONS implicit-solvent model the transition temperature is probably still strongly overestimated as is already known from studies of other helical peptides [33]. One of the reasons is the choice of a temperature-independent solvent-peptide coupling strength and the "smeared," nonlocal, and static polar environment without intrinsic fluid properties.

IV. HELIX-COIL TRANSITIONS OF THE PEPTIDES IN SOLVENT

In the following we discuss the thermodynamic properties of the three synthetic sequences employing the ECEPP/3 force field with OONS implicit-solvent parameter set. Considering several quantities we find strong indications for helix-coil transitions in all three cases. Helix-coil transitions in peptides and nucleic acids were first addressed by Zimm and Bragg [34,35] and have been studied extensively [36].

As a first indication for the conformational transitions, we find strong peaks in the specific heat as shown for the three peptides in solvent in Fig. 5. The peaks are located near 440, 420, and 410 K for the wild-type $S1$, mutant $S2$, and randomly permuted sequence $S3$, respectively. The character of the conformational transition is identified by measuring the temperature dependence of α-helicity and β-sheetness. The helicity is a natural order parameter for the identification of helix-coil transitions in peptides. A residue is defined to be in α-helical state, if its backbone dihedral angles ϕ and ψ are in the range $(-70° \pm 20°)$ and $(-37° \pm 20°)$, respectively. We show the changes of this quantity and its derivative versus temperature in Figs. 6(a) and 6(b), respectively.

In the calculation of the helicity, end residues are not taken into account, because end residues are very flexible and do not conform to any definite state. There is also a proline residue in all chains which is known as helix breaker, because it lacks a primary amine group, and due to the pep-
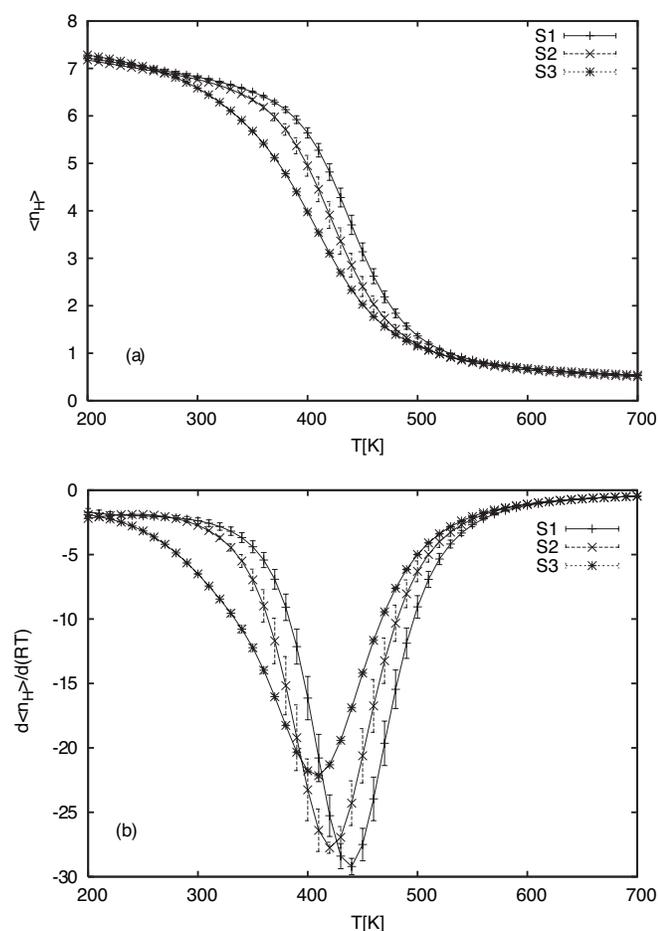
FIG. 6. Average number of helical residues (a) and derivative (b) vs temperature for the systems in solvent.

FIG. 7. Average number of β-sheet structures (a) and derivative (b) vs temperature for the peptides in solvent.

tide bond to the preceding amino acid, there is no H atom allowing for the formation of a hydrogen bond that could stabilize an α-helix or a β-sheet structure. Furthermore, the rigid proline side chain typically forces for steric reasons the ψ angle of the preceding amino acid to take nonhelical values. Hence, the maximum number of residues in a helical segment is 8 for the three sequences under consideration. Rapid decreases of the average helicities [as shown in Fig. 6(a)] are noticed for all three peptides. The thus determined transition temperatures lie in the same temperature region as the peaks of the specific heat, as can be read off from the fluctuations around the average helicity as plotted in Fig. 6(b).

In the force-field parametrization used, β-sheets can be excluded in the low-temperature region. This is confirmed by the plots in Figs. 7(a) and 7(b), where the sheetness $\langle n_\beta \rangle$, which is the average number of Ramachandran angles in a β-sheet state [i.e., $\phi \in (-150° \pm 30°)$ and $\psi \in (150° \pm 30°)$], and the fluctuations of this quantity are shown. No noticeable β-sheet structure is identified in the low-temperature region because the whole ensemble consists of strongly helical conformations. The average sheetness increases slightly above the conformational-transition temperature, but this signal is relatively weak and the high-temperature ensemble is expected to be dominated by random-coil structures.

Another interesting quantity is the probability of the individual residues to become "helical." In Figs. 8(a)–8(c), we plot the grayscale-coded profiles of the probabilities of each residue to be in a helical state. Since it is not intuitive that a single residue can form an α-helix motif, although its dihedral angles are in the range of the α-helical region in the Ramachandran map, we define a residue to be helical only if it is part of a helical segment with at least three successive helical pairs of Ramachandran angles. This allows a clearer view on the helix formation. For the three considered peptides in implicit solvent, the helix-coil transition is a sharp one-step process in the temperature region between 400 and 500 K. The single helical segment of the peptides $S1$ and $S2$ is formed by the residues 5–12 (counting from the N terminus). Proline at position 4 in the sequence breaks the helix and the 1-4 residual tail is coil like. Surprisingly, the randomly permuted sequence $S3$, where proline is located at position 9 (which is, unfortunately, also at the fourth position—counted from the C terminus), exhibits in addition to the strong 1-8 helix a second helical segment between residues 10 and 12. Although this signal is weak, a non-negligible subset of conformations in the low-temperature ensemble contains *two* independent helical segments, broken by proline. The transition temperature for the 10-12 helix is slightly smaller than for the main segment and lies below 400 K. Note that the formation of the second helix for se-

(a)

(b)

(c)
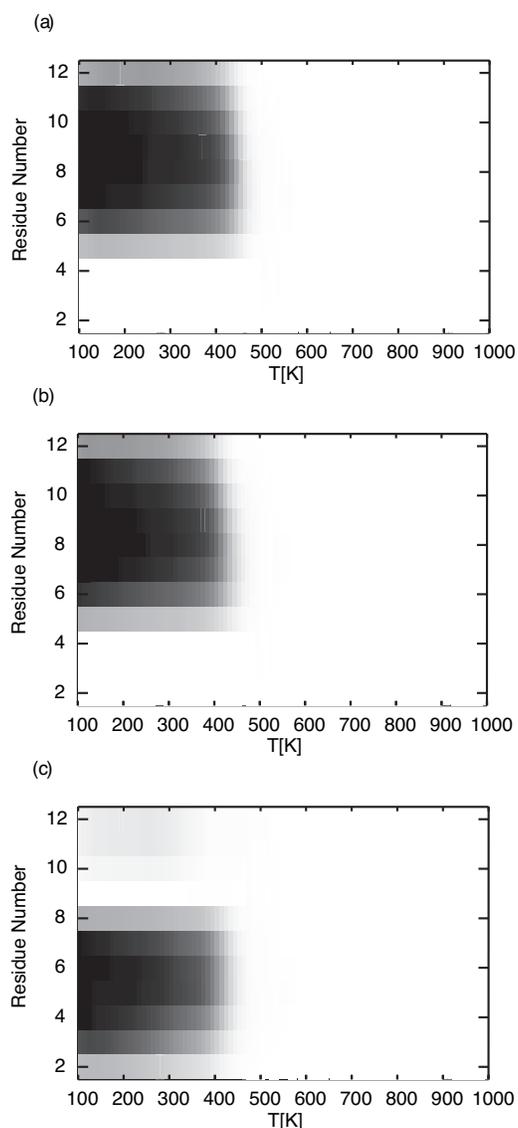
FIG. 8. Probability profiles of each residue to be a part of a three-residue helical segment as a function of temperature for the peptides (a) $S1$, (b) $S2$, and (c) $S3$ in solvent.

quence $S3$ can also be observed as "shoulder" in the corresponding specific heat in Fig. 5.

Hydrogen bonds are mainly responsible for the formation and stability of secondary structures such as α-helices. In Figs. 9(a) and 9(b), the average numbers of hydrogen bonds $\langle n_{\text{H-bond}} \rangle$ and their fluctuations, respectively, are shown for the three sequences. In the helical phase, conformations typically possess approximately four to six hydrogen bonds on average. Hydrogen bonds of the chains with water are obviously not counted employing an implicit peptide-solvent model and averages shown in the figure reflect only intrinsic hydrogen bonds. As can be seen in Fig. 9(a), the peptide with the randomly permuted sequence $S3$ behaves noticeably different than sequences $S1$ and $S2$. For $S3$, $\langle n_{\text{H-bond}} \rangle$ decreases more smoothly with increasing temperature; i.e., the breaking of the individual hydrogen bonds is a process of relatively weak cooperativity. In this case, the hydrogen bonds are comparatively weak and the two independent helical seg-
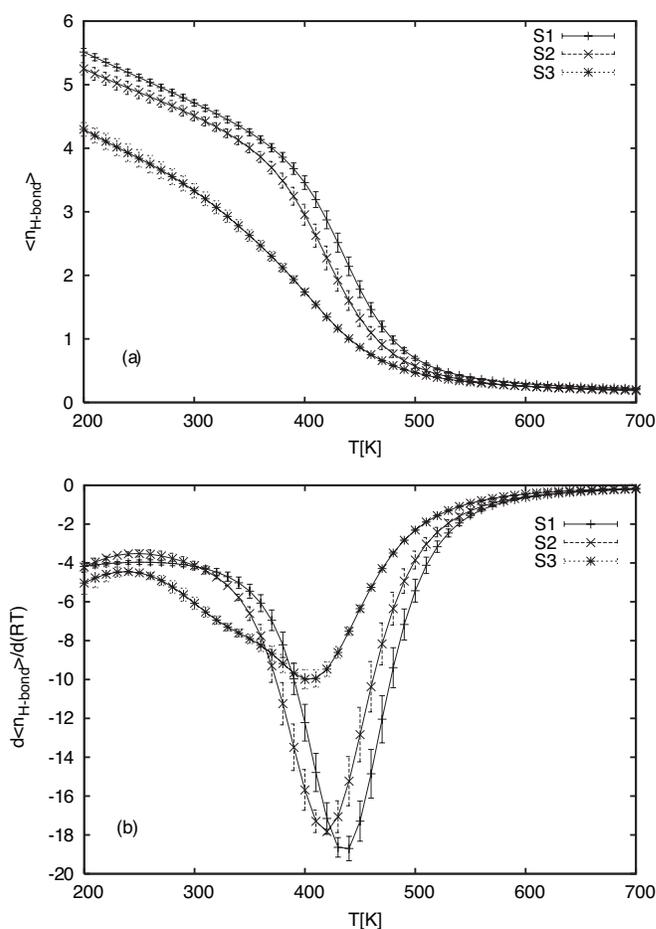
FIG. 9. Average number of hydrogen bonds (a) and derivative (b) vs temperature for the three peptides in solvent.

ments and, more globally, the helical phases are not very stable, as the fluctuations show in Fig. 9(b). This effect is mainly due to the position of the proline in the chain. Although its distance from the ends is identical in all three sequences, the asymmetry with respect to the dihedral constraints destabilizes the larger helical segment of $S3$ (whose average length is smaller than the helices of $S1$ and $S2$) and leads to a noticeable probability of forming a second small and weak helical segment.

Finally, we have also calculated the radius of gyration R_{Gy} as a rather global geometrical quantity, which is mainly useful for quantifying the structural collapse caused by a conformational transition. In contrast to the Θ collapse transition of polymers between nonstructured globular and random-coil conformations, a helix-coil transition is rather a crossover from nonstructured conformations to conformations with highly ordered segments (helices). For this reason, the gyration radius is a too rough measure for the order in the helical phase and is therefore of less importance for the understanding of secondary-structure formation and cannot be deduced from the short-range interactions [37,38].

In Figs. 10(a) and 10(b), the average radius of gyration and its fluctuations are shown. Being a measure of the compactness of the molecule, small values of the gyration radius indicate more tight-packed structures. For all sequences, the
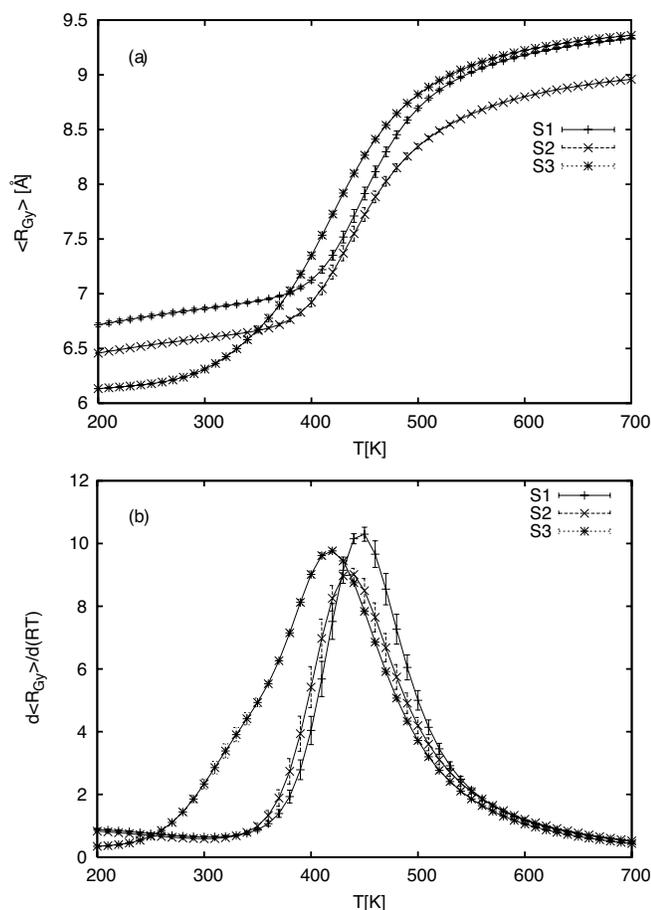
FIG. 10. Average radius of gyration (a) and derivative (b) vs temperature for the systems in solvent.

average radius of gyration changes from 6.0–6.5 Å to 9.0–9.5 Å. The peaks in the derivative of R_{Gy} indicate slightly higher conformational transition temperatures than the identified temperatures from fluctuations of energy and helicity. Sequence $S3$ possesses the most compact conformations in the helical phase—although the average number of hydrogen bonds is the smallest, as can also been seen in Table II, where the differences between maximal and minimal values of average radius of gyration and mean number of hydrogen bonds are given.

In Table III, we list the peak temperatures identified from fluctuations of several thermodynamic quantities discussed in this section for the three peptides $S1$, $S2$, and $S3$. The peak temperatures of the peptides in solvent are compared with

TABLE II. Differences between maximal and minimal values of average gyration radius $\Delta\langle R_{Gy}\rangle = \langle R_{Gy}\rangle_{max} - \langle R_{Gy}\rangle_{min}$ and mean number of hydrogen bonds $\Delta\langle n_{H-bond}\rangle = \langle n_{H-bond}\rangle_{max} - \langle n_{H-bond}\rangle_{min}$ for each sequence over the whole temperature range.

	$\Delta\langle R_{Gy}\rangle$ [Å]	$\Delta\langle n_{H-bond}\rangle$
$S1$	2.6	5.4
$S2$	2.5	5.1
$S3$	3.2	4.1

TABLE III. Helix-coil transition temperatures T_C in K, read off for the three sequences $S1$, $S2$, and $S3$ in vacuum and solvent from the various quantities discussed in the paper. The errors are in the range ±10 K.

		C_V	$d\langle R_{Gy}\rangle/dT$	$d\langle n_H\rangle/dT$	$d\langle n_{H-bond}\rangle/dT$	$d\langle n_\beta\rangle/dT$
$S1$	Vac.	530	600	520	510	550
	Solv.	440	450	440	440	450
$S2$	Vac.	520	580	510	510	550
	Solv.	420	440	420	420	430
$S3$	Vac.	480	580	460	460	520
	Solv.	410	420	410	400	420

the corresponding transition temperatures identified for the systems in vacuum. While for the peptides in solvent the transition temperatures are relatively independent of the fluctuations considered, the deviations for the vacuum systems are noticeable. This is not surprising, as the systems in solvent are stabilized by the environment and the helix-coil transition is a cooperative effect that is accompanied by a strong coupling to the solvent. In vacuum, the conformational freedom is much larger and the helix-coil transition rather an entropic effect. In this case, the finiteness of the systems is more influential than for the peptides in solvent.

V. MULTICANONICAL HISTOGRAMS AND FREE-ENERGY LANDSCAPES

For the study of the helix-coil folding channels, it is useful to investigate the two-dimensional multicanonical histogram of the energy E and a suitable system parameter, which is chosen here to be the angular overlap parameter Q, as defined in Eq. (11). As reference conformations, required for the calculation of Q, we use the lowest-energy conformations found in the multicanonical simulations. These structures are shown in Fig. 2 (right) for $S1$ and in Fig. 11 for $S2$ and $S3$.

The multicanonical histogram is obtained from the multicanonical time series

$$H(E,Q) = \sum_t \delta_{E,E(t)}\delta_{Q,Q(t)}, \qquad (12)$$

where the sum runs over the Monte Carlo steps t. The summation over Q yields the "flat" multicanonical energy distribution. Since the conformational energy E, if replaced by the average energy $\langle E\rangle$, is directly related to the temperature, the
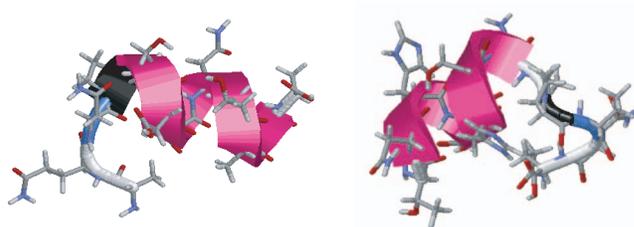
FIG. 11. (Color online) Lowest-energy reference conformations of sequence $S2$ (left) and $S3$ (right) in solvent.
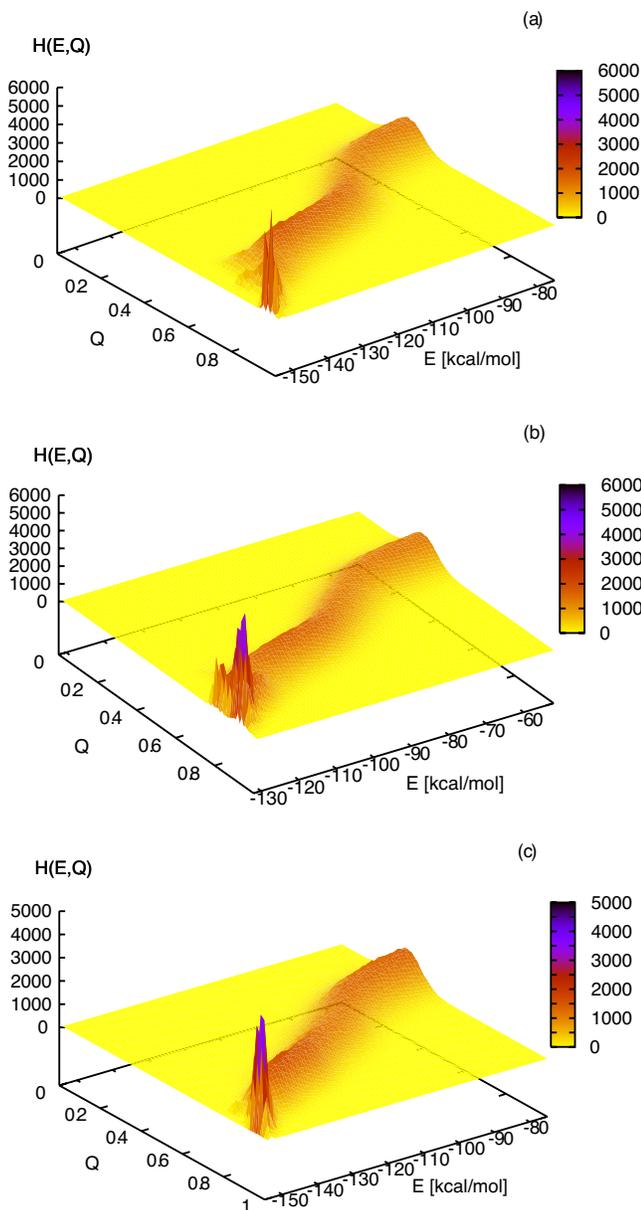
FIG. 12. (Color online) Multicanonical histograms of overlap parameter Q and energy E for the peptides (a) $S1$, (b) $S2$, and (c) $S3$ in solvent.

FIG. 13. (Color online) Free-energy landscapes $F(Q,T)$ for the peptides (a) $S1$, (b) $S2$, and (c) $S3$ in solvent.

histogram $H(E,Q)$ contains sufficient information for the description of the simple folding transition. In Fig. 12, we plot the multicanonical histograms for the three sequences $S1$, $S2$, and $S3$. In all three cases, we find a noticeable turn in the distribution from small values of Q, which correspond to random-coil conformations, to values closer to unity, where the conformations are similar to the helical reference conformation. In correspondence to the interpretation in the previous section, the transition is stronger for the wild-type and mutant sequences $S1$ and $S2$, respectively, and rather a two-step process in the case of the randomly permuted sequence $S3$. The remarkable bifurcations at very low energies are possibly indications for metastable conformations which can be considered as weakly disturbed reference conformations. It should be noted that the angular overlap parameter is cal-

culated by comparing all dihedral angles. This means that deviations in side-chain dihedral angles lead to Q values different from unity, although the backbone dihedral angles could take almost the same values. Helical structures are mainly due to cooperativity along the backbone, but differences in the precise side-chain locations do not destabilize the helical structures. Therefore, these kinds of glassy transitions, which happen under extreme conditions (very low temperatures), are interesting but not in the main focus of this work.

In Figs. 13(a)–13(c), we plot the free-energy landscapes for the three peptides. Here, we assume that the angular overlap parameter Q is a suitable measure for the structural order of the peptides. The free energy as a function of this "order" parameter and temperature is then given by

$$F(Q,T) = -RT \ln p(Q,T), \qquad (13)$$

with the distribution of the overlap parameters

$$p(Q,T) = \int \mathcal{D}\boldsymbol{\xi}\, \delta(Q - Q(\boldsymbol{\xi}))e^{-\beta E(\boldsymbol{\xi})}, \qquad (14)$$

where the integral runs over all possible conformations $\boldsymbol{\xi}$. The free-energy plots in Figs. 13(a)–13(c) confirm that all three peptides experience a conformational transition following a single main folding channel. Although the helix-coil transition separates the pseudophases of random conformations and the long-range-ordered helical phase, no noticeable signal of pseudophase coexistence is observed; i.e., the transition appears rather second order like than first order like.

VI. SUMMARY

We have analyzed thermodynamic properties and folding channels in the free-energy landscape for three synthetic 12-residue peptides which exhibit remarkable adsorption affinities to semiconductors [5,6]. Employing an all-atom model based on the ECEPP/3 force field [25] with OONS implicit solvation parameter set [21] and applying the implementation of the multicanonical Monte Carlo simulation method in the SMMP package [27], we found in all three cases indications for a strong helix-coil transition. Independent of the fluctuations studied, the peak temperatures are very close to each other—despite of the smallness of the peptides. Since experimental verification and biochemical structural analysis of these peptides are still pending, a comparison with experimental data is not yet possible.

Our predictions for the transition temperatures are probably too high, as is to be expected by using implicit-solvent models. Therefore, we expect that the helix-coil transitions could happen under reasonable environmental conditions such that it should be possible to verify our predictions experimentally. This is an important issue, since it is generally expected that selective synthetic peptides and polymers may play an essential role in future nanotechnological applications.

[1] G. M. Whitesides and B. Grzybowski, Science **295**, 2418 (2002).

[2] R. D. Kamien, Science **299**, 1671 (2003).

[3] R. L. Willett, K. W. Baldwin, K. W. West, and L. N. Pfeiffer, Proc. Natl. Acad. Sci. U.S.A. **102**, 7817 (2005).

[4] S. Brown, Nat. Biotechnol. **15**, 269 (1997).

[5] S. R. Whaley, D. S. English, E. L. Hu, P. F. Barbara, and A. M. Belcher, Nature (London) **405**, 665 (2000).

[6] K. Goede, P. Busch, and M. Grundmann, Nano Lett. **4**, 2115 (2004).

[7] M. Bachmann and W. Janke, Phys. Rev. E **73**, 020901(R) (2006).

[8] M. Bachmann and W. Janke, Phys. Rev. Lett. **95**, 058102 (2005); Phys. Rev. E **73**, 041802 (2006).

[9] A, alanine; Q, glutamine; N, asparagine; P, proline; S, serine; D, aspartate; T, threonine; H, histidine.

[10] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991); Phys. Rev. Lett. **68**, 9 (1992).

[11] W. Janke, Physica A **254**, 164 (1998).

[12] G. Gökoğlu, H. Arkın, E. Aktürk, and T. Çelik, Int. J. Mod. Phys. C **16**, 1489 (2005).

[13] G. Gökoğlu, H. Arkın, and T. Çelik, Int. J. Mod. Phys. C **16**, 455 (2005).

[14] H. Arkın and T. Çelik, Eur. Phys. J. B **30**, 577 (2002).

[15] A. Irbäck, B. Samuelsson, F. Sjunnesson, and S. Wallin, Bio-

phys. J. **85**, 1466 (2003).

[16] K. F. Lau and K. A. Dill, Macromolecules **22**, 3986 (1989).

[17] F. H. Stillinger, T. Head-Gordon, and C. L. Hirshfeld, Phys. Rev. E **48**, 1469 (1993); F. H. Stillinger and T. Head-Gordon, Phys. Rev. E **52**, 2872 (1995).

[18] M. Bachmann and W. Janke, Phys. Rev. Lett. **91**, 208105 (2003).

[19] M. Bachmann and W. Janke, J. Chem. Phys. **120**, 6779 (2004).

[20] M. Bachmann, H. Arkın, and W. Janke, Phys. Rev. E **71**, 031906 (2005).

[21] T. Ooi, M. Obatake, G. Némethy, and H. A. Scheraga, Proc. Natl. Acad. Sci. U.S.A. **84**, 3086 (1987).

[22] M. Masuya and Y. Okamoto, Res. Dev. Pure Appl. Chem. **2**, 1 (1998).

[23] B. A. Berg and H. P. Hsu, Phys. Rev. E **69**, 026703 (2004).

[24] Here we use the standard Lennard-Jones parametrization as provided by the ECEPP/3 force field [25]. It should be noted that there exist more refined effective pair potentials for unbonded atomic interactions as given, e. g., in K. Cahill and V. A. Parsegian, J. Chem. Phys. **121**, 10839 (2004).

[25] F. A. Momany, L. A. Carruthers, R. F. McGuire, and H. A. Scheraga, J. Phys. Chem. **78**, 1595 (1974); F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, *ibid.* **79**, 2361 (1975); G. Némethy, M. S. Pottle, and H. A. Scheraga, *ibid.* **87**, 1883 (1983); M. J. Sippl, G. Némethy, and H. A.

Scheraga, *ibid.* **88**, 6231 (1984); G. Némethy, K. D. Gibson, K. A. Palmer, C. N. Yoon, G. Paterlini, A. Zagari, S. Rumsey, and H. A. Scheraga, *ibid.* **96**, 6472 (1992).

[26] D. Eisenberg and A. D. McLachlan, Nature (London) **319**, 199 (1986).

[27] F. Eisenmenger, U. H. E. Hansmann, S. Hayryan, and C. K. Hu, Comput. Phys. Commun. **138**, 192 (2001). The SMMP program package, which contains the ECEPP/3 force field and the OONS solvation parameter set as used in this study, is freely available under www.phy.mtu.edu/biophys/smmp.htm

[28] B. A. Berg, Fields Inst. Commun. **26**, 1 (2000).

[29] U. H. E. Hansmann, Y. Okamoto, and J. N. Onuchic, Proteins: Struct., Funct., Genet. **34**, 472 (1999).

[30] U. H. E. Hansmann and J. N. Onuchic, J. Chem. Phys. **115**,

1601 (2001).

[31] B. A. Berg, H. Noguchi, and Y. Okamoto, Phys. Rev. E **68**, 036126 (2003).

[32] C. A. Schiffer, J. W. Caldwell, P. A. Kollman, and R. M. Stroud, Mol. Simul. **10**, 121 (1993).

[33] U. H. E. Hansmann, J. Chem. Phys. **120**, 417 (2004).

[34] B. H. Zimm and J. K. Bragg, J. Chem. Phys. **31**, 526 (1959).

[35] B. H. Zimm, J. Chem. Phys. **33**, 1349 (1960).

[36] D. Poland and H. A. Scheraga, *Theory of Helix-Coil Transitions in Bioploymers* (Academic Press, New York, 1970).

[37] S. Doniach, Chem. Rev. (Washington, D.C.) **101**, 1763 (2001).

[38] B. Zagrovic, J. Lipfert, E. J. Sorin, I. S. Millett, W. F. van Gunsteren, S. Doniach, and V. S. Pande, Proc. Natl. Acad. Sci. U.S.A. **102**, 11698 (2005).