Solution Effects and the Folding of an Artificial Peptide

Nelson A. Alves*

Departamento de Física e Matemática, FFCLRP, Universidade de São Paulo. Av. Bandeirantes 3900, CEP 14040-901 Ribeirão Preto, SP, Brazil

Ulrich H. E. Hansmann*

Department of Physics, Michigan Technological University, Houghton, Michigan 49931-1291

Received: April 10, 2003

We explore for a simple artificial peptide, $Ala_{10}-Gly_5-Ala_{10}$, the relation between secondary structure formation and folding and its dependence on the solvent model. Our data rely on multicanonical Monte Carlo simulations where the interactions among all atoms are taken into account. Three implicit solvent models are considered to approximate the interaction between protein and the surrounding water, and the results are compared to previous data from gas-phase simulations. Our data indicate that the final structure and the details of the folding process depend strongly on how the solvent environment affects secondary structure formation.

1. Introduction

Our understanding of the folding process has increased considerably over the past few years. For instance, the energy landscape theory and funnel concept[1,2] have been shown to be powerful tools for characterizing folding not only in minimalistic protein models but also for real proteins.[3−5] However, many questions remain open. For instance, it is still under debate whether the secondary structure forms before, after, or parallel to tertiary contacts, and to what degree secondary structure formation and folding are determined by the intrinsic properties of the protein or by the interaction with the surrounding solvent.

We investigate these questions in the present paper for a simple model, the artificial peptide $Ala_{10}-Gly_5-Ala_{10}$. Polyalanine (Ala_N) exhibits a helix-coil transition in the gas phase and with certain implicit solvent models.[6−10] Hence, our peptide allows us to study the relation between helix formation and folding in a simple model. We have found in previous work[11] that the peptide folds in the gas phase in a two-step process: first, two α-helices are formed that then arrange themselves into a U-like structure. In the present paper, we extend this research to the case where protein−solvent interactions are taken into account and study how the relation between helix formation and folding depends on the details of the solvent model. Preliminary results are published in the proceedings of the 2002 Sanibel meeting.[12]

As in previous work,[11] our results rely on multicanonical simulations of $Ala_{10}-Gly_5-Ala_{10}$ in a detailed representation where the interactions between all atoms are taken into account. Following common practice, we have approximated the protein−water interactions by a solvent-accessible surface term. The three-parameter sets we compare are named by us as OONS,[13] SCH,[14] and WE92,[15] and are described in the respective references. Quantities such as energy, specific heat, and helicity are calculated as functions of temperature. The encountered transitions are further investigated by partition function zeros analysis, which allows one to characterize "phase transitions" in small systems.[16,17] This technique was first introduced to the study of biological molecules in ref 8, and further applications to proteins can be found in refs 11, 18, and 19. Our results show that secondary structure formation, folding, and final structure depend strongly on the solvent environment.

In the next section, we review our simulation techniques and explain the method of our analysis. In Section III, we present and discuss the results of our simulations. Section IV summarizes our conclusions.

2. Methods

Our computer experiments are based on a detailed, all-atom representation of the peptide $Ala_{10}-Gly_5-Ala_{10}$. The intramolecular interactions are described by a standard force field, ECEPP/2[20] (as implemented in the program package SMMP[21]), and given by

$$E_{\mathrm{ECEPP/2}} = E_{\mathrm{C}} + E_{\mathrm{LJ}} + E_{\mathrm{HB}} + E_{\mathrm{tor}} \tag{1}$$

$$E_{\mathrm{C}} = \sum_{(i,j)} \frac{332 q_i q_j}{\epsilon r_{ij}} \tag{2}$$

$$E_{\mathrm{LJ}} = \sum_{(i,j)} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) \tag{3}$$

$$E_{\mathrm{HB}} = \sum_{(i,j)} \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) \tag{4}$$

$$E_{\mathrm{tor}} = \sum_{l} U_l (1 \pm \cos(n_l \chi_l)) \tag{5}$$

Here, r_{ij} (in Å) is the distance between the atoms i and j, and χ_l is the lth torsion angle. The peptide bond angles are set to their common value $\omega = 180°$. We further assume the electrostatic permittivity in the protein interior to be $\epsilon = 2$ (its common value in ECEPP simulations).

The interactions between the peptide and the surrounding water are approximated by adding a solvent-accessible surface

* To whom correspondence should be addressed. E-mail: alves@quark.ffclrp.usp.br; hansmann@mtu.edu.

10.1021/jp034964z CCC: $25.00
Published on Web 08/20/2003

term[13] to the energy function:

$$E = E_{\mathrm{ECEPP/2}} + E_{\mathrm{solv}} \tag{6}$$

with

$$E_{\mathrm{solv}} = \sum_i \sigma_i A_i$$

Here, E_{solv} is the solvation energy. In this approximation, one assumes that the free energy difference between atomic groups immersed in the protein interior and groups exposed to water is proportional to the solvent-accessible surface area A_i of the ith atom with the parameters σ_i as experimentally determined proportionality factors. We have found previously[10] that Ala_{10} has a large helix propensity when the OONS[13] or SCH[14] parameter sets are used, but little tendency to form helices for WE92.[15] Hence, by choosing these three-parameter sets that are all available in SMMP,[21] we can investigate in detail for our peptide the relation between secondary structure formation and folding. Our results are compared to those of gas-phase simulations where the solvation term E_{solv} is set to zero.

Simulations of detailed protein models where the interactions between all atoms are taken into account are extremely difficult. This is because the various competing interactions within the molecule lead to a multitude of local energy minima separated by high barriers. Hence, in the low-temperature region, canonical Monte Carlo or molecular dynamics simulations will get trapped in one of these minima and not thermalize within the available CPU time. Only recently, with the introduction of new and sophisticated algorithms such as generalized-ensemble techniques,[22] was it possible to alleviate this problem in protein simulations.[23] For this reason, our investigations rely again on the use of one of these techniques, multicanonical sampling,[24] where conformations with energy E are assigned a weight $w_{\mathrm{mu}}(E) \propto 1/n(E)$ that is a function of the density of states $n(E)$. A simulation with this weight will generate a 1D random walk in the energy space, allowing itself to escape from any local minimum, and therefore will lead to a uniform distribution of energy:

$$P_{\mathrm{mu}}(E) \propto n(E) w_{\mathrm{mu}}(E) = \mathrm{const.} \tag{7}$$

Because a large range of energies are sampled, one can use the reweighting techniques[25] to calculate thermodynamic quantities over a wide range of temperatures T by

$$\langle \mathcal{A} \rangle_T = \frac{\int \mathrm{d}x\, \mathcal{A}(x) w^{-1}(E(x)) e^{-\beta E(x)}}{\int \mathrm{d}x\, w^{-1}(E(x)) e^{-\beta E(x)}} \tag{8}$$

where x stands for configurations and $\beta = 1/k_{\mathrm{B}}T$ is the inverse temperature. Estimates for the multicanonical weights $w(E) = n^{-1}(E)$ can be found by the iterative procedure described in refs 10 and 26. In our case, we needed 500 000 sweeps for the weight factor calculations. All thermodynamic quantities are then estimated from one production run of 4 000 000 (OONS, WE92) to 8 000 000 (SCH) Monte Carlo sweeps that followed 10 000 sweeps for thermalization. Our simulations start from completely random initial conformations (Hot Start), and one Monte Carlo sweep updates every torsion angle of the peptide once. At the end of every 10th sweep, we store the ECEPP/2 energies E_{Tot}, E_{C}, E_{LJ}, E_{HB}, and E_{tor} of the conformation, the corresponding number n_{H} of helical residues, and the end-to-end distance $d_{\mathrm{e-e}}$. Here, we follow previous work[27] and consider a residue as helical if its backbone angles (ϕ,ψ) are within the range $(-70° \pm 30°, -37° \pm 30°)$. Dividing the time series of the stored quantities into eight bins, we calculate as an estimate for the error bars in our thermal averages the standard deviation of the values obtained separately for each bin.

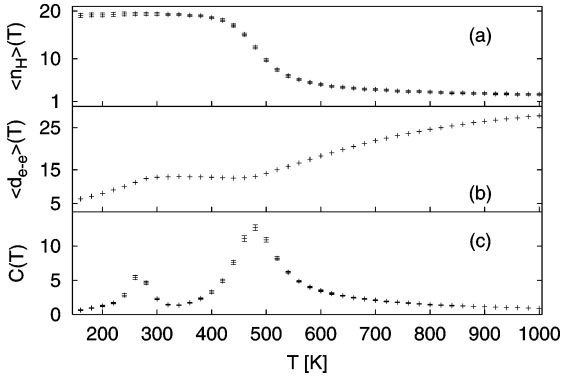

Figure 1. (a) Average number of helical residues $\langle n_{\mathrm{H}} \rangle$ (T), (b) average end-to-end distance $\langle d_{\mathrm{e-e}} \rangle$ (T), and (c) specific heat $C(T)$ as a function of temperature as calculated from simulations of $Ala_{10}-Gly_5-Ala_{10}$ in the gas phase.

We finally use that the multicanonical algorithm allows us to calculate estimates for the spectral density:

$$n(E) = P_{\mathrm{mu}}(E) {w_{\mathrm{mu}}}^{-1}(E) \tag{9}$$

and we can therefore construct the corresponding partition function for our all-atom model of $Ala_{10}-Gly_5-Ala_{10}$ from these estimates by

$$Z(\beta) = \sum_E n(E) e^{-\beta E} \tag{10}$$

The complex zeros of the partition function can be obtained from the complex extension of the temperature variable. Their locations and behavior in the complex temperature plane determine the thermodynamic characteristics of the system.[29] Recently, we have introduced partition function zeros analysis as a tool to characterize structural transitions in biomolecules,[8,30] and we use it here to research the relation between helix formation and folding.

3. Results and Discussion

3.1. Thermodynamics of $Ala_{10}-Gly_5-Ala_{10}$. We start our analysis by presenting first our results for $Ala_{10}-Gly_5-Ala_{10}$ in the gas phase, that is, for simulations of the peptide with only the ECEPP/2 energy term $E_{\mathrm{ECEPP/2}}$ of eq 6. Because polyalanine has a pronounced helix-coil transition in the gas phase, we expect formation of α-helices in our peptide. For this reason, we have measured the average number of helical residues $\langle n_{\mathrm{H}} \rangle$ that is displayed in Figure 1a. At high temperature, few residues have backbone dihedral angles (ϕ,ψ) typical for an α-helix. On the other hand, at low temperatures, helices form, and almost all of the alanine residues are part of an α-helix. The transition between the two temperature regions at $T = 485 \pm 5$ K is sharp and corresponds to a pronounced peak in the specific heat $C(T)$ at $T = 480 \pm 10$ K (Figure 1c). Combining both values, we obtain as our final estimate for the helix-coil transition temperature in the gas phase ${T_{\mathrm{hc}}}^{\mathrm{GP}} = 483 \pm 8$ K. However, we find in Figure 1c a second, smaller peak at the lower temperature ${T_{\mathrm{f}}}^{\mathrm{GP}} = 265 \pm 7$ K, indicating yet another transition. To understand this second peak, we plot in Figure 1b the average end-to-end distance $\langle d_{\mathrm{e-e}} \rangle$ (T) as a function of

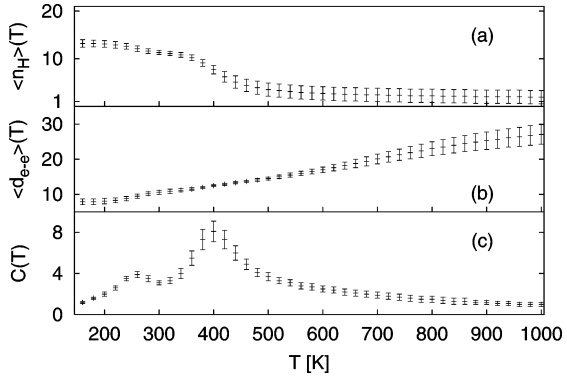

Figure 2. (a) Average number of helical residues $\langle n_H \rangle$ (T), (b) average end-to-end distance $\langle d_{e-e} \rangle$ (T), and (c) specific heat $C(T)$ as a function of temperature as calculated from simulations of $Ala_{10}-Gly_5-Ala_{10}$ with the SCH solvent-accessible surface term.

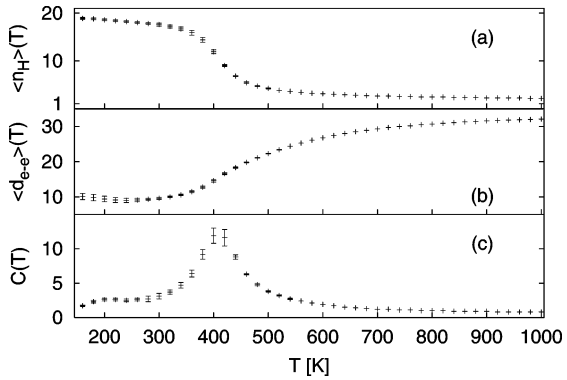

Figure 3. (a) Average number of helical residues $\langle n_H \rangle$ (T), (b) average end-to-end distance $\langle d_{e-e} \rangle$ (T), and (c) specific heat $C(T)$ as a function of temperature as calculated from simulations of $Ala_{10}-Gly_5-Ala_{10}$ with the OONS solvent-accessible surface term.

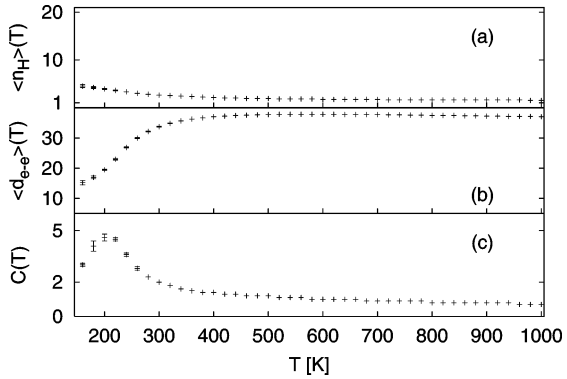

Figure 4. (a) Average number of helical residues $\langle n_H \rangle$ (T), (b) average end-to-end distance $\langle d_{e-e} \rangle$ (T), and (c) specific heat $C(T)$ as a function of temperature as calculated from simulations of $Ala_{10}-Gly_5-Ala_{10}$ with the WE92 solvent-accessible surface term.

temperature. This quantity is a measure for the compactness of a protein conformation and defined here by the distance between N of Ala_1 and O of Ala_{25}. We observe that $\langle d_{e-e} \rangle$ (T) decreases when the temperature is lowered. Below the helix-coil transition T_{hc}, the curve becomes almost flat at a value of $\langle d_{e-e} \rangle \approx 10$ Å, and there is little further change in the compactness of the molecule. However, at the temperature T_f, the end-to-end distance decreases again sharply toward a new value $\langle d_{e-e} \rangle = 6.1$ Å. Hence, T_f marks the folding of the molecule into a defined compact structure with the two terminal ends of the peptide close together. This scenario is supported by Figure 5a where we display the lowest-energy configuration found in our multicanonical gas-phase simulation of 8 000 000 sweeps.[11] It consists of two helixes (made up of the alanine residues) connected by a turn (build out of the flexible glycine residues) to a hairpinlike structure that is consistent with the small value of the end-to-end distance $\langle d_{e-e} \rangle$ at temperatures below T_f in Figure 1b.

The situation does not change qualitatively if we include in our simulation protein−solvent interactions by means of a solvent-accessible surface term with the SCH parameter set.[14] Figure 2a displays the average number of helical residues $\langle n_H \rangle$ (T) as a function of temperature, Figure 2b displays the end-to-end distance $\langle d_{e-e} \rangle$ (T), and Figure 2c displays the specific heat $C(T)$. We find again two peaks in the specific heat, with the one at a higher temperature corresponding to a transition between helical and coil states, and the one at a lower temperature correlated with a change in the end-to-end distance. However, the helix-coil transition temperature T_{hc} is shifted toward lower temperatures: $\langle n_H \rangle$ (T) has a transition at $T = 405 \pm 10$ K and a corresponding peak in specific heat $C(T)$ at $T = 395 \pm 10$ K. Combining both values, we find a value of $T_{hc}^{SCH} = 400 \pm 8$ K that is ∼80 K lower than the corresponding transition temperature in the gas phase. On the other hand, the folding temperature $T_f^{SCH} = 258 \pm 8$ K is within the error bars identical to that in the gas phase. Hence, the SCH solvent lowers the helix-coil transition temperature and decreases in this way the temperature difference between the two transitions. Note that both transitions are weaker than the corresponding transitions in the gas phase. The two maxima in specific heat ($C_{hc}^{SCH} = 8.1(5)$ and $C_f^{SCH} = 3.9(3)$) are smaller than the corresponding values in the gas phase: $C_{hc}^{GP} = 12.7(4)$ and $C_f^{GP} = 5.6(3)$. Similarly, we obtain for the susceptibility

$$\chi = (\langle n_H^2 \rangle - \langle n_H \rangle^2)/N^2 \tag{11}$$

(plots not shown) that is also a measure for the strength of the helix-coil transition $\chi^{SCH} = 0.42(9)$, a value considerably smaller than the gas-phase value $\chi^{GP} = 0.90(3)$. In the low-temperature phase, SCH leads to an average helicity that is smaller than that in the gas phase (at $T = 200$ K, $\langle n_H \rangle^{SCH} = 13.1(7)$ and $\langle n_H \rangle^{GP} = 19.2(4)$), while the end-to-end distance is with $\langle d_{e-e} \rangle^{SCH}$ ($T = 200$ K) $= 8.1(1)$ Å only slightly larger than the gas-phase value $\langle d_{e-e} \rangle^{GP}$ ($T = 200$ K) $= 7.0$ Å. Given the above similarities, it is not surprising that the lowest-energy configuration (Figure 5b) is again a helix-hairpin. However, the two helices are not completely formed at the investigated temperatures. This result suggests that the SCH solvation term hinders helix formation. We remark that we do not find the extended helical configurations that compete in the gas phase with the structure in Figure 5a (see ref 11), indicating that SCH parameters lead to a preference for compact structures.

A slightly different picture appears in simulations with the OONS solvent term. We first note that the use of this implicit solvent term leads to essentially the same lowest-energy structure (displayed in Figure 5c) as in the gas-phase simulations. Both helices are completely formed. However, the folding process differs from that in the gas phase: the specific heat $C(T)$ (Figure 3c) has only a single peak at a temperature $T_c = 408 \pm 10$ K. We display again in Figure 3a and 3b the average number of helical residues $\langle n_H \rangle$ (T) and the average end-to-end distance $\langle d_{e-e} \rangle$ (T), respectively, as functions of temperature. At the temperature T_c, $\langle n_H \rangle$ (T) is increasing rapidly with decreasing temperature, while at the same temperature $\langle d_{e-e} \rangle$ (T) exhibits a sharp drop when the temperature is lowered. Hence, unlike in the case of gas-phase simulations or such with the SCH solvation term, helix formation and folding happen at the same temperature $T_f = T_{hc}$. Note that the corresponding peak in

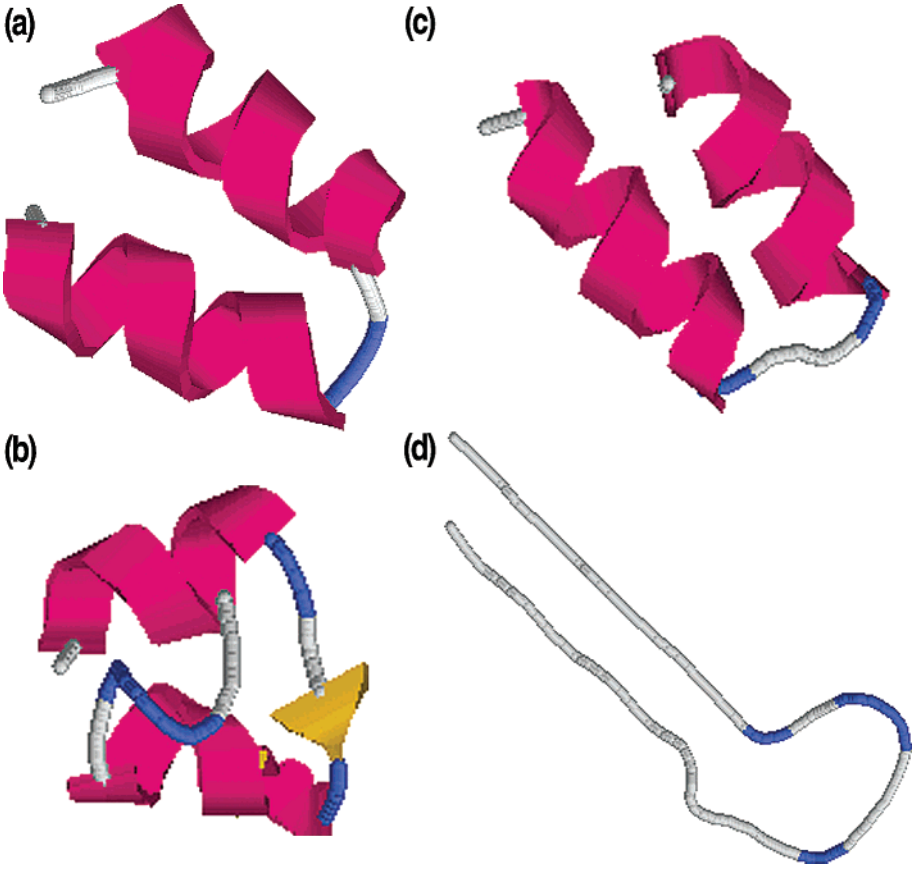

Figure 5. Lowest-energy configuration of $Ala_{10}-Gly_5-Ala_{10}$ as obtained from multicanonical simulations (a) in the gas phase, (b) with the SCH, (c) with the OONS, and (d) with the WE92 solvation term.

specific heat is of comparable height to the peak that characterizes in gas-phase simulations the helix-coil transition: $C_{hc}^{OONS} = 12.3(1.2)$ versus $C_{hc}^{GP} = 12.7(4)$. Similarly, we find that the peak in susceptibility (plot not shown) is for OONS ($\chi^{OONS} = 0.75(4)$) only slightly smaller than the corresponding one for the gas phase: $\chi^{GP} = 0.90(3)$. Both results indicate that the OONS term does not hinder helix formation and the helicity $\langle n_H \rangle^{OONS}$ ($T = 200$ K) = 18.6(2) in the low-temperature phase is indeed comparable to that in the gas phase ($\langle n_H \rangle^{GP}(T = 200$ K) = 19.2(4)) and larger than that with SCH solvent ($\langle n_H \rangle^{SCH}(T = 200$ K) = 13.1(4)). However, the helix-coil transition is shifted by ~70 K toward lower temperatures, and the transition temperature is comparable to that in simulations with a SCH solvent. Again, as with SCH solvent, we do not find the extended single-helix configurations that compete with the helix-hairpin in gas-phase simulations at temperatures between T_{hc} and T_f, while the end-to-end distance $\langle d_{e-e} \rangle^{OONS}(T = 200$ K) = 9.4 Å is larger than the corresponding values in the gas phase and with SCH solvent ($\langle d_{e-e} \rangle^{SCH}(T = 200$ K) = 8.1(1) Å and $\langle d_{e-e} \rangle^{GP}(T = 200$ K) = 7.0 Å).

Our third solvent representation is WE92.[15] Unlike in simulations with the OONS or SCH parameter set, polyalanine shows little tendency to form an α-helix in WE92 solvent.[10] Hence, we do not expect to find a helix-coil transition for $Ala_{10}-Gly_5-Ala_{10}$ in the WE92 solvent. Figure 4a displays the number of helical residues as calculated from multicanonical simulations of our peptide with this parameter set. No indication for a helix-coil transition is found as the numbers of helical residues stay vanishingly small at low temperatures. At $T = 200$ K, that average helicity is with $\langle n_H \rangle^{WE92}$ ($T = 200$ K) = 3.9(1) considerably smaller than the values for the other solvent models. Consequently, we do not observe the strong signal in the specific heat (Figure 4c) that has indicated a helix-coil transition for the gas phase or with OONS and SCH solvent. Although only a single small peak ($C^{WE92} = 4.7(1)$) is observed, its temperature $T_c = 205(5)$ K is much lower than the transition temperatures for the other solvent representations and is related to a pronounced drop of the end-to-end distance $\langle d_{e-e} \rangle^{WE92}$ at this temperature (see Figure 4b). Note, however, that at $T = 200$ K, $\langle d_{e-e} \rangle^{WE92} = 19.2$ Å is still much larger than the values observed in the other models at this temperature: $\langle d_{e-e} \rangle^{GP} = 7.0$ Å, $\langle d_{e-e} \rangle^{SCH} = 8.1$ Å, and $\langle d_{e-e} \rangle^{OONS} = 9.4$ Å. We further find that the average number of residues $\langle n_B \rangle^{WE92} = 9.4(2)$ that have dihedral angles (ϕ,ψ) common in a β-sheet is much larger than that found in GP, SCH, and OONS simulations at $T = 200$ K: $\langle n_B \rangle^{GP} = 0.6(1)$, $\langle n_B \rangle^{SCH} = 2.4(4)$, and $\langle n_B \rangle^{OONS} = 0.5(1)$. This indicates a preference for β-sheet formation in the WE92 solvent, and we find indeed that the lowest-energy configuration in WE92 simulations (displayed in Figure 5d) resembles a β-hairpin. However, as one can see from Figure 4c, the (weak) transition into this state occurs at a very low temperature $T_c = 205(5)$ K, indicating that this state is not as stable as the α-hairpin. Indeed, we find in the WE92 simulation a variety of disordered or only partially ordered structures that have only marginally higher energies (~5 kcal/mol) than the β-hairpin ground-state configuration.

3.2. Partition Function Zeros Analysis of $Ala_{10}-Gly_5-Ala_{10}$. To research in more detail the observed transitions in the four models of $Ala_{10}-Gly_5-Ala_{10}$, we continue with an analysis of the partition function zeros in the complex u-plane ($u = e^{-\beta}$). We display in Table 1 the first complex u-zeros for our peptide in the gas phase, and with SCH, OONS, and WE92 solvents. Following ref 30, we have calculated these zeros by means of the scan procedure,[28] avoiding in this way any

TABLE 1: Partition Function Zeros for $Ala_{10}-Gly_5-Ala_{10}$ in the Gas Phase (GP) or with the Implicit Solvent Term

	Re(u_1)	Im(u_1)	Re(u_2)	Im(u_2)	Re(u_3)	Im(u_3)	Re(u_4)	Im(u_4)
GP	0.3482(17)	0.0458(18)	0.3546(36)	0.0760(20)	0.333(14)	0.107(14)	0.327(15)	0.118(15)
GP	0.1510(33)	0.0407(17)	0.1099(56)	0.0812(33)	0.0772(82)	0.1019(25)	0.0422(81)	0.1189(30)
SCH	0.279(13)	0.058(11)	0.2778(70)	0.091(11)	0.3064(80)	0.134(15)		
SCH	0.132(12)	0.0643(27)						
OONS	0.2898(38)	0.0467(92)	0.2790(92)	0.071(13)	0.337(27)	0.1426(98)		
WE92	0.0829(20)	0.0564(15)	0.0910(57)	0.0960(51)	0.053(11)	0.0752(66)		

approximation from dividing the energy range into small intervals. Error bars are then obtained for the zeros by means of the jackknife method dividing again the time series of energies into eight bins.

Partition function zeros analysis is a common tool in statistical physics, but it relies on an analysis of the limit of infinite large systems. Such finite-size scaling is in general not possible for biological molecules. However, there exist approaches that allow one to define "phase transitions" on small systems by means of partition function zeros analysis. One possibility is the classification scheme by Borrmann et al.,[16] where one computes the discrete line of zeros as an average of the inverse distances between neighboring zeros

$$\phi(\tau_k) = \frac{1}{2}\left(\frac{1}{|\beta_k - \beta_{k-1}|} + \frac{1}{|\beta_{k+1} - \beta_k|}\right) \tag{12}$$

and approximates $\phi(\tau)$ by a simple power law $\phi(\tau) \approx \tau^{\alpha}$. Taking the first four complex zeros, one obtains

$$\alpha = \frac{\ln \phi(\tau_3) - \ln \phi(\tau_2)}{\ln \tau_3 - \ln \tau_2} \tag{13}$$

With a second parameter γ, related to the crossing angle of this line with the real axis,

$$\gamma = [\mathrm{Re}(\beta_2) - \mathrm{Re}(\beta_1)]/(\tau_2 - \tau_1) \tag{14}$$

phase transitions can now be classified according to the values of these two parameters: for $\alpha \le 0$ and $\gamma = 0$, one has a phase transition of first order; it is of second order if $0 < \alpha < 1$ and γ is arbitrary, and for $\alpha > 1$ and arbitrary γ, one has a higher order transition. With the four zeros listed for the gas-phase model of $Ala_{10}-Gly_5-Ala_{10}$ in Table 1, we can calculate these parameters α and γ. For the helix-coil transition at $T = 480$ K, we find $\alpha = 1.1(1.5)$ and $\gamma = -0.4(2)$. The errors reflect large fluctuations in the values of the two parameters α and γ that do not allow us to determine whether the helix-coil transition is a weak first-order or a strong second-order phase transition. However, the strength of the transition indicates that helix formation occurs by means of nucleation mechanism. Our data are more decisive in the case of the collapse and folding transition at $T_f^{GP} = 265$ K. Here, we find $\alpha = 0.32(8)$ and $\gamma = 0.36(2)$, indicating a second-order transition. This is consistent with what one would expect for a transition between extended and compact structures and implies that collapse and folding of the $Ala_{10}-Gly_5-Ala_{10}$ are due to long-range correlations between the residues.

We note that calculating partition function zeros is more difficult for the solvated molecule then for $Ala_{10}-Gly_5-Ala_{10}$ in the gas phase. With our statistics, we are not able to obtain the first four zeros but only the first three. For this reason, we cannot use the Borrmann approach for analyzing the transitions in the solvated molecule. Instead, we have to go on the competing approach by Janke and Kenna (JK)[17] that also allows one to define "phase transitions" for molecules. In the JK approach, one assumes that the zeros condense for large enough systems on a single line,

$$u_j = u_c + r_j \exp(i\varphi) \tag{15}$$

with the cumulative density of zeros given by the average

$$G_N(r_j) = \frac{2j-1}{2N} \tag{16}$$

Here, j labels the complex zeros in order of increasing imaginary part, $u_j = \exp(-\beta_j)$, $j = 1, 2, \ldots$, and N is the number of residues. Janke and Kenna postulate that this density scales for sufficiently large chains as[17]

$$\frac{2j-1}{2N} = a_1(\mathrm{Im}u_j(L))^{a_2} + a_3 \tag{17}$$

In the above three-parameter fit, a value of a_3 consistent with zero is a necessary condition for the existence of a "phase transition". The values of other constants, a_1 and a_2, then characterize the phase transition. For instance, for first-order transitions, the constant a_2 should take values $a_2 \approx 1$ for small r, and in this case the slope of this equation is related to the latent heat.[17] On the other hand, a value of a_2 larger than 1 indicates a second-order transition whose specific heat exponent is given by $\alpha = 2 - a_2$.

Having only three zeros, we cannot calculate meaningful results from the above three-parameter fit. Assuming that the peaks in the specific heat indicate the existence of a "phase transition", we therefore set $a_3 = 0$ in the above fit and restrict ourselves to a two-parameter fit in the JK parameters a_1 and a_2. With this fit, we obtain in the OONS model for the single transition at $T^{OONS} = 408(10)$ K, that marks both helix formation and folding, a characteristic parameter $a_2 = 0.9(5)$. The corresponding specific heat exponent $\alpha = 1.1(5)$ has again a large error bar, but suggests that the molecule has in OONS solvent either a weak first-order or a strong second-order transition. Hence, the critical behavior of the transition that the peptide exhibits in OONS solvent seems to be dominated by the helix-coil transition and is of the same order as the helix-coil transition of the molecule in the gas phase. Again, this indicates that helix formation occurs for $Ala_{10}-Gly_5-Ala_{10}$ through a nucleation mechanism.

For $Ala_{10}-Gly_5-Ala_{10}$ in an SCH solvent, we find for the helix-coil transition at $T_{hc}^{SCH} = 400(8)$ K a value of $a_2 = 1.6(8)$ that corresponds to a specific heat exponent $\alpha = 0.4(8)$. Because of the large error bar, it is difficult to assert the order of this transition, albeit the mean value of α suggests that the helix-coil transition is considerably weaker for the SCH model than for the molecule in the gas phase (or an OONS solvent). Our data are worse for the folding transition at $T_f^{SCH} = 258(8)$ K and do not permit one to draw any conclusion about the order of this transition.

On the other hand, for the WE92 model of $Ala_{10}-Gly_5-Ala_{10}$, we have only a transition at a low temperature $T^{WE92} = 205(5)$ K that is associated with the formation of β-hairpin-like configurations. We obtain for this transition with the JK

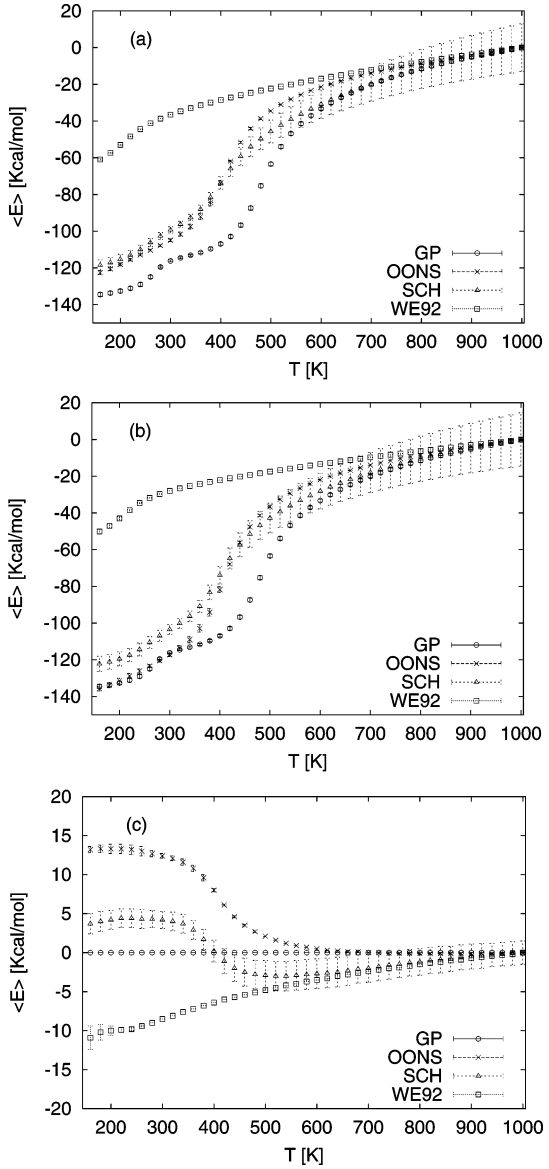

Figure 6. (a) Total energy $\langle E_{\mathrm{Tot}} \rangle = \langle E_{\mathrm{ECEPP/2}} + E_{\mathrm{solv}} \rangle$, (b) potential energy $\langle E_{\mathrm{ECEPP/2}} \rangle$, and (c) solvation energy $\langle E_{\mathrm{solv}} \rangle$ as obtained in multicanonical simulations of $Ala_{10}-Gly_5-Ala_{10}$ in the gas phase (GP) and with SCH, OONS, and WE92 solvent.

approach a negative value for the $a_2 = -0.6(1.4)$ that corresponds to a critical exponent $\alpha = 2.6(1.4)$. Even taking the large error bars into account, we found that this value clearly excludes the possibility of a first-order or strong second-order phase transition and (as one would expect) is inconsistent with a nucleation mechanism for β-sheet formation.

3.3. Helix Formation and Folding. Our above analysis of the thermodynamics of our peptide suggests two different folding mechanisms for $Ala_{10}-Gly_5-Ala_{10}$ that depend on the solvent environment. In the gas phase (GP), with SCH, or with OONS solvent, our peptide folds in a two-step process. The first step is the formation of α-helices by means of a nucleation mechanism and can be characterized by a helix-coil transition temperature T_{hc}. We have shown in previous work[10] that, for polyalanine, helix formation is connected to a large gain in potential energy $E_{\mathrm{ECEPP/2}}$. Consequently, we see in Figure 6b for these three models a large drop in the ECEPP/2-energy at the respective temperatures T_{hc}, but not for the case of WE92 solvent where we do not observe helix formation. Note that we have here (as in Figure 6a and Figure 6c) normalized the

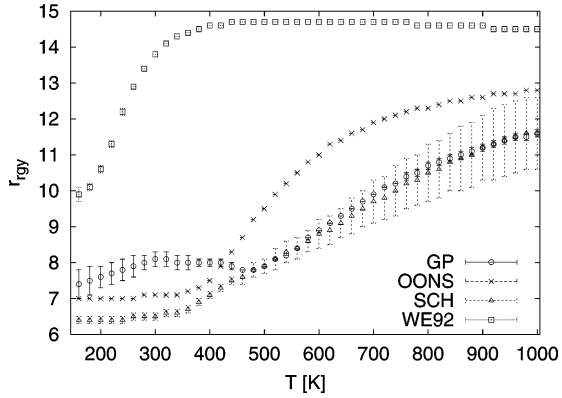

Figure 7. Average radius of gyration $\langle r_{\mathrm{rgy}} \rangle$ as a function of temperature for $Ala_{10}-Gly_5-Ala_{10}$ in the gas phase (GP) and with SCH, OONS, and WE92 solvent.

energies in such a way that their averaged values are zero at $T = 1000$ K. This allows a better comparison of the four models. Note also the considerably larger error bars for the SCH data in Figures 6 and 7 which rely on a simulation with 8 000 000 sweeps instead of the 4 000 000 sweeps of the other solvent models. This indicates that simulations with the SCH solvent are numerically more difficult than those of our peptide in the gas phase, with OONS, or with WE92 solvent. The observed gain in potential energy is partially compensated in SCH and OONS solvents by a loss in solvent energy E_{solv}. This is because the SCH and OONS parameter sets both emphasize the hydrophobic character of the carbon atoms and decrease the hydrophilic character of uncharged oxygen and nitrogen atoms. As is shown in ref 10, the corresponding solvation energies favor coil structures over helical configurations. Hence, the change in total energy E_{Tot} (Figure 6a) is less pronounced for OONS and SCH than for the peptide in the gas phase, indicating a smaller preference for α-helices and leading to a lower helix-coil transition temperature when compared to gas-phase simulations. Note, however, that in both models the overall gain in energy through formation of helices is still larger than the energy gain due to the solvent term in WE92 simulations where no helix formation is observed and where E_{solv} increases with decreasing temperature (Figure 6c).

The formation of α-helices in the gas phase, with SCH, or with OONS solvent then restricts the possible configuration space. Energetically most favorable is the folding of two α-helices (made out of the alanine residues) into a hairpin. This second step can be characterized for GP and SCH by a second, lower folding temperature T_{f}, while for OONS solvent both transitions happen at the same temperature $T_{\mathrm{f}} = T_{\mathrm{hc}}$. Because the SCH and OONS parameter sets both emphasize the hydrophobic character of the carbon atoms and decrease the hydrophilic character of uncharged oxygen and nitrogen atoms, the resulting force makes a close alignment of the two α-helices (built out of the alanine residues) energetically more favorable than they are in the gas phase. Consequently, we do not observe in SCH and OONS simulations the extended single helix configurations that for $T_{\mathrm{f}} < T < T_{\mathrm{hc}}$ compete with the helix hairpin. We conjecture that, for the OONS solvent, this force is so strong that the helices fold immediately after their formation into the hairpin, merging in this way the two transitions into a single one. Note that the above scenario is reminiscent of the well-known framework[32,33] and collision-diffusion model[31] of folding which also propose that local elements of native local secondary structure form independently of tertiary structure. These elements diffuse until they collide and coalesce to give a tertiary structure.

TABLE 2: Energy Differences between the Helical Configuration with the Lowest Energy and the Lowest-Energy β-Sheet Configuration

	GP	SCH	OONS	WE92
ΔE_{Tot}	−69.1	−45.5	−38.4	8.3
$\Delta E_{\mathrm{ECEPP/2}}$	−69.1	−50.2	−49.3	−46.6
ΔE_{solv}		4.7	10.5	54.9

Simulations of the peptide in WE92 solvent differ from the above scenario. In this parameter set, the nitrogen, oxygen, and sulfur atoms are considered hydrophilic, that is, favoring solvent exposure. Hence, this parameter set biases against formation of α-helices, and extended structures are favored. This can be seen in Figure 7 where we display the radius of gyration $\langle r_{\mathrm{rgy}}\rangle$ as a function of temperature for our four models. This quantity is a measure for the compactness of protein configurations. While the SCH and OONS solvents lead to a smaller value of $\langle r_{\mathrm{rgy}}\rangle$ than was observed in gas phase (i.e., to more compact configurations), the WE92 solvent leads to extended configurations and much larger values of $\langle r_{\mathrm{rgy}}\rangle$. Because the WE92 solvent biases against α-helices, the above-described two-step mechanism of folding is suppressed for $Ala_{10}-Gly_5-Ala_{10}$. Instead, β-sheet formation is favored. This observation is supported by Table 2, where we list for the four models the energy differences between the lowest-energy α-helical configuration and the corresponding β-sheet configuration. In the gas phase, α-helices are favored by about ∼70 kcal/mol over β-sheets, and in SCH and OONS solvent they are favored by ∼40 kcal/mol. On the other hand, β-sheets are energetically more favorable by ∼10 kcal/mol in a WE92 simulation. These net-differences are the sum of two terms. For all four models, the ECEPP/2 energy of α-helices is by more than 40 kcal/mol lower than that of β-sheets, while the protein−solvent interaction term favors β-sheets over α-helices for each of the three solvent representations. For SCH and OONS solvents, this term is ∼5 (SCH) and ∼10 kcal/mol (OONS) small as compared to the bias in ECEPP/2 energy toward α-helices (of order >50 kcal/mol). However, for the WE92 solvent, the gain in solvation energy is ∼55 kcal/mol larger than the corresponding loss of ECEPP/2 energy. Consequently, we find that the WE92 ground state of $Ala_{10}-Gly_5-Ala_{10}$ (Figure 5d) is a β-hairpin. However, as the low transition temperature of $T_c = 205(5)$ K indicates, this state is energetically much less stable than the α-hairpin that is observed in the other models (see Figure 6).

How general are the above obtained results? The observed transition temperatures (and especially the helix-coil transition temperatures) are not in the range of physiologically meaningful temperatures and indicate limitations of our force field and the solvent representations. Unfortunately, the authors are not aware of experimental results for the solvated peptide that would allow a comparison. However, an experimental study of $Ala_{10}-Gly_5-Ala_{10}$ in the gas phase is now in preparation. Using the techniques developed by Jarrold and collaborators for examination of gas-phase conformations of proteins,[34] we plan to test the validity of our predicted folding mechanism for $Ala_{10}-Gly_5-Ala_{10}$. Future extension of these experimental tests to solvated variants of the molecule may follow.

4. Conclusion

Summarizing our results, we have performed multicanonical simulations with high statistics of the artificial peptide $Ala_{10}-Gly_5-Ala_{10}$. In the gas phase, the peptide exhibits two characteristic transitions, a helix-coil transition at $T_{\mathrm{hc}}^{\mathrm{GP}} = 483 \pm 8$ K and a second transition at $T_{\mathrm{f}}^{\mathrm{GP}} = 265 \pm 7$ K that leads to an α-hairpin as the most stable structure. Introducing an implicit solvent either lowers the helix-coil transition temperature ($T_{\mathrm{hc}}^{\mathrm{SCH}} = 400(8)$ K for SCH solvent and $T_{\mathrm{hc}}^{\mathrm{OONS}} = 408(10)$ K for the OONS solvent) or inhibits helix formations as in the case of WE92 solvent. If helix formation is not inhibited, the α-hairpin remains the most stable structure with corresponding folding temperatures of $T_{\mathrm{f}}^{\mathrm{SCH}} = 258(8)$ K and $T_{\mathrm{f}}^{\mathrm{OONS}} = T_{\mathrm{hc}}^{\mathrm{OONS}} = 408(10)$ K. In that case, $Ala_{10}-Gly_5-Ala_{10}$ folds in a two-step process where the formation of α-helices predates the folding into the final structure. This folding mechanism is not observed in the case of WE92 where α-helices are not formed and replaced as ground states by a β-hairpin. However, the β-hairpin is only stable below a transition temperature $T_{\mathrm{f}}^{\mathrm{WE92}} = 205(5)$ K. Hence, solvent environments that do not inhibit formation of α-helices seem to lead to more stable tertiary structures for our peptide. Experimental tests of this scenario are now planned. We also intend now to extend our numerical investigations to larger variants of our peptide that allow a direct comparison with recent experimental work by Myers and Oas,[35] where the relation between helix formation and folding was studied for the case of the 58-residues B domain of protein A.

Acknowledgment. N.A.A. acknowledges support by CNPq (Brazil), and U.H.E.H. recognizes support by a research grant from the National Science Foundation (CHE-9981874). The authors are grateful to the Computer Center of Academia Sinica (Taipei, Taiwan) and to DFMA (IFUSP, Brazil) for the computer facilities made available to them. Part of this work was done while U.H.E.H. visited the Ribeirão Preto campus of the University of São Paulo. He thanks FAPESP for a generous travel grant and the Department of Physics and Mathematics for kind hospitality.

References and Notes

(1) Bryngelson, J. D.; Wolynes, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 7524.
(2) Onuchic, J. N.; Luhey-Schulten, Z.; Wolynes, P. G. *Annu. Rev. Phys. Chem.* **1997**, *48*, 545.
(3) Boczko, E. M.; Brooks, C. L., III. *Science* **1995**, *269*, 393.
(4) Hansmann, U. H. E.; Okamoto, Y.; Onuchic, J. N. *Proteins*: *Struct., Funct., Genet.* **1999**, *34*, 472.
(5) Hansmann, U. H. E.; Onuchic, J. N. *J. Chem. Phys.* **2001**, *115*, 1601.
(6) Poland, D.; Scheraga, H. A. *Theory of Helix-Coil Transitions in Biopolymers*; Academic Press: New York, 1970.
(7) Hansmann, U. H. E.; Okamoto, Y. *J. Chem. Phys.* **1999**, *110*, 1267; *111*, 1339(E).
(8) Alves, N. A.; Hansmann, U. H. E. *Phys. Rev. Lett.* **2000**, *84*, 1836.
(9) Mitsutake, A.; Okamoto, Y. *J. Chem. Phys.* **2000**, *112*, 10638.
(10) Peng, Y.; Hansmann, U. H. E. *Biophys. J.* **2002**, *82*, 3269.
(11) Alves, N. A.; Hansmann, U. H. E. *J. Chem. Phys.* **2002**, *117*, 2337.
(12) Hansmann, U. H. E. *Int. J. Quantum Chem.* **2002**, *90*, 1515.
(13) Ooi, T.; Oobatake, M.; Nemethy, G.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 3086.
(14) Schiffer, C. A.; Caldwell, J. W.; Kollman, P. A.; Stroud, R. M. *Mol. Simul.* **1993**, *10*, 121.
(15) Wesson, M.; Eisenberg, D. *Protein Sci.* **1992**, *1*, 227.
(16) Borrmann, P.; Mülken, O.; Harting, J. *Phys. Rev. Lett.* **2000**, *84*, 3511.
(17) Janke, W.; Kenna, R. *J. Stat. Phys.* **2001**, *102*, 1211.
(18) Peng, Y.; Hansmann, U. H. E.; Alves, N. A. *J. Chem. Phys.* **2003**, *118*, 2374.
(19) Wang, F.; Wang, W. *J. Chem. Phys.* **2003**, *118*, 2952.
(20) Sippl, M. J.; Némethy, G.; Scheraga, H. A. *J. Phys. Chem.* **1984**, *88*, 6231 and references therein.
(21) Eisenmenger, F.; Hansmann, U. H. E.; Hayryan, S.; Hu, C. K. *Comput. Phys. Commun.* **2001**, *138*, 192.
(22) Hansmann, U. H. E.; Okamoto, Y. In *Annual Reviews in Computational Annual Reviews of Computational Physics*; Stauffer, D., Ed.; World Scientific: Singapore, 1999; Vol. VI; p 129.
(23) Hansmann, U. H. E.; Okamoto, Y. *J. Comput. Chem.* **1993**, *14*, 1333.
(24) Berg, B. A.; Neuhaus, T. *Phys. Lett. B* **1991**, *267*, 249.

(25) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1988**, *61*, 2635.
(26) Berg, B. A. *J. Stat. Phys.* **1996**, *82*, 323.
(27) Okamoto, Y.; Hansmann, U. H. E. *J. Phys. Chem.* **1995**, *99*, 11276.
(28) Alves, N. A.; de Felicio, J. R. D.; Hansmann, U. H. E. *Int. J. Mod. Phys. C* **1997**, *8*, 1063 and references therein.
(29) Itzykson, C.; Pearson, R. B.; Zuber, J. B. *Nucl. Phys. B* **1983**, *220* [*FS8*], 415.
(30) Alves, N. A.; Ferrite, J. P. N.; Hansmann, U. H. E. *Phys. Rev. E* **2002**, *65*, 036110.
(31) Karplus, M.; Weaver, D. L. *Protein Sci.* **1994**, *3*, 650.
(32) Ptitsyn, O. B. *Protein Eng.* **1994**, *7*, 593.
(33) Kim, P. S.; Baldwin, R. L. *Annu. Rev. Biochem.* **1990**, *59*, 631.
(34) Hudgins, R. R.; Ratner, M. A.; Jarrold, M. F. *J. Am. Chem. Soc.* **1998**, *120*, 12974.
(35) Myers, J. K.; Oas, T. J. *Nat. Struct. Biol.* **2001**, *8*, 552.