Thermodynamics of peptide aggregation processes: An analysis from perspectives of three statistical ensembles

Christoph Junghans,[1,2,a)] Michael Bachmann,[1,3,b)] and Wolfhard Janke[1,c)]
[1]*Institut für Theoretische Physik and Centre for Theoretical Sciences (NTZ), Universität Leipzig, Postfach 100920, D-04009 Leipzig, Germany*
[2]*Max-Planck-Institut für Polymerforschung, Ackermannweg 10, D-55128 Mainz, Germany*
[3]*Computational Biology and Biological Physics Group, Department of Theoretical Physics, Lunds Universitet, Sölvegatan 14A, SE-223 62 Lund, Sweden*

(Received 6 August 2007; accepted 6 December 2007; published online 28 February 2008)

We employ a mesoscopic model for studying aggregation processes of proteinlike hydrophobic-polar heteropolymers. By means of multicanonical Monte Carlo computer simulations, we find strong indications that peptide aggregation is a phase separation process, in which the microcanonical entropy exhibits a convex intruder due to non-negligible surface effects of the small systems. We analyze thermodynamic properties of the conformational transitions accompanying the aggregation process from the multicanonical, canonical, and microcanonical perspective. It turns out that the microcanonical description is particularly advantageous as it allows for unraveling details of the phase-separation transition in the thermodynamic region, where the temperature is not a suitable external control parameter anymore. © *2008 American Institute of Physics*.
[DOI: 10.1063/1.2830233]

I. INTRODUCTION

Beside receptor-ligand binding mechanisms, folding and aggregation of proteins belong to the biologically most relevant molecular structure formation processes. While the specific binding between receptors and ligands is not necessarily accompanied by global cooperative structural changes, protein folding and oligomerization of peptides are typically accompanied by conformational transitions.[1] Proteins and their aggregates are comparatively small systems. A typical protein consists of a sequence of some hundreds of amino acids and aggregates are often formed by only a few peptides. A very prominent example is the extracellular aggregation of the $A\beta$ peptide, which is associated with Alzheimer's disease. Following the amyloid hypothesis, it is believed that these aggregates (which can also take fibrillar forms) are neurotoxic, i.e., they are able to fuse into cell membranes of neurons and open calcium ion channels. It is known that extracellular Ca^{2+} ions intruding into a neuron can promote its degeneration.[2–4]

Conformational transitions that proteins experience during structuring and aggregation are not phase transitions in the strict thermodynamic sense, and their statistical analysis is usually based on studies of signals exposed by energetic and structural fluctuations, as well as system-specific "order" parameters. In these studies, the temperature T is considered as an adjustable, external control parameter and, for the analysis of the pseudophase transitions, the peak structure of quantities such as the specific heat and the fluctuations of the gyration tensor components or "order" parameter as functions of the temperature are investigated. The natural ensemble for this kind of analysis is the canonical ensemble, where the possible states of the system with energies E are distributed according to the Boltzmann probability $\exp(-E/k_BT)$, where k_B is the Boltzmann constant. However, phase separation processes of small systems as, e.g., droplet condensation, are accompanied by surface effects at the interface between the pseudophases.[5–15] This is reflected by the behavior of the microcanonical entropy $\mathcal{S}(E)$, which exhibits a *convex* monotony in the transition region. Consequences are the backbending of the caloric temperature $T(E)=(\partial\mathcal{S}/\partial E)^{-1}$, i.e., the *decrease* of temperature with increasing system energy, and the negativity of the microcanonical specific heat $C_V(E)=(\partial T(E)/\partial E)^{-1}=-(\partial\mathcal{S}/\partial E)^2/(\partial^2\mathcal{S}/\partial E^2)$. The physical reason is that the free energy balance in phase equilibrium requires the minimization of the interfacial surface and, therefore, the loss of entropy.[16] A reduction of the entropy can, however, only be achieved by transferring energy into the system. Recently, we have shown that, employing a minimalistic heteropolymer model, the aggregation of two small peptides is such a phase separation process, where we observed the mentioned peculiar small-system effects.[17] Here, we consider the aggregation process from the multicanonical, canonical, and microcanonical perspectives. Our results were obtained from multicanonical computer simulations of a mesoscopic hydrophobic-polar heteropolymer model for aggregation, which is based on a simple off-lattice model, originally introduced to study tertiary folding of proteins from a coarse-grained point of view.

The paper is organized as follows. In Sec. II, we define the aggregation model employed in our computational study, where we primarily used multicanonical sampling. This method is also briefly described here as well as the aggregation order parameter needed to discriminate the

a)Electronic mail: junghans@mpip-mainz.mpg.de.
b)Electronic mail: michael.bachmann@itp.uni-leipzig.de.
c)Electronic mail: wolfhard.janke@itp.uni-leipzig.de. URL: http://www.physik.uni-leipzig.de/CQT.html

pseudophases. Section III is devoted to the main part of the paper: the presentation of the results for the aggregation of two small peptides obtained from multicanonical, canonical, and microcanonical views. The comparison with the results obtained for larger systems is performed in Sec. IV. The paper is concluded by a summary of our results in Sec. V.

II. MODEL AND METHODS

For our aggregation study on mesoscopic scales, we employ a novel model that is based on a known hydrophobic-polar single-chain approach, originally introduced for heteropolymer chains in two dimensions.[18] In this section, we define this model, describe the simulation methods, and introduce a suitable order parameter that allows for the discrimination of the macrostates or "pseudophases" the multiple-chain system can reside in.

A. Mesoscopic hydrophobic-polar aggregation model

For our aggregation study of proteinlike heteropolymers, we assume that the tertiary folding process of the individual chains is governed by hydrophobic-core formation in an aqueous environment. A comparatively simple but powerful model is the AB model,[18] where only two types of amino acids are considered: hydrophobic residues (A), which avoid contact with the polar environment and polar residues (B) being favorably attracted by the solvent. The model is a C^{α} type model in that each residue is represented by only a single interaction site (the "C^{α} atom"). Thus, the natural dihedral torsional degrees of freedom of realistic protein backbones are replaced by virtual bond and torsion angles between consecutive interaction sites. The large torsional barrier of the peptide bond between neighboring amino acids is in the AB model effectively taken into account by introducing a bending energy. Nonbonded residues experience weak pairwise long-range attraction (AA and BB pairs) or repulsion (AB pairs), respectively. Although this coarse-grained picture is obviously not capable to reproduce microscopic properties of specific realistic proteins, it qualitatively exhibits, however, sequence-dependent features observed in nature, as, for example, tertiary folding pathways known from two-state folding, folding through intermediates, and metastability.[19]

For our systems of more than one chain, we further assume that the interaction strength between nonbonded residues is independent of the individual properties of the chains the residues belong to. Therefore, we use the same parameter sets as in the AB model for the pairwise interactions between residues of different chains. Our aggregation model reads[17]

$$E = \sum_{\mu} E_{\mathrm{AB}}^{(\mu)} + \sum_{\mu<\nu} \sum_{i_\mu, j_\nu} \Phi(r_{i_\mu j_\nu}; \sigma_{i_\mu}, \sigma_{j_\nu}), \tag{1}$$

where μ, ν label the M polymers interacting with each other, and i_μ, j_ν index the $N_{\mu,\nu}$ monomers of the respective μth and νth polymer. The intrinsic single-chain energy is given by

$$E_{\mathrm{AB}}^{(\mu)} = \frac{1}{4} \sum_{i_\mu} (1 - \cos \vartheta_{i_\mu}) + \sum_{j_\mu > i_\mu + 1} \Phi(r_{i_\mu j_\mu}; \sigma_{i_\mu}, \sigma_{j_\mu}), \tag{2}$$

with $0 \leqslant \vartheta_{i_\mu} \leqslant \pi$ denoting the bending angle between monomers i_μ, $i_\mu+1$, and $i_\mu+2$. The nonbonded inter-residue pair potential

$$\Phi(r_{i_\mu j_\nu}; \sigma_{i_\mu}, \sigma_{j_\nu}) = 4[r_{i_\mu j_\nu}^{-12} - C(\sigma_{i_\mu}, \sigma_{j_\nu}) r_{i_\mu j_\nu}^{-6}] \tag{3}$$

depends on the distance $r_{i_\mu j_\nu}$ between the residues, and on their type, $\sigma_{i_\mu} = A, B$. The long-range behavior is attractive for like pairs of residues $[C(A,A)=1,\ C(B,B)=0.5]$ and repulsive otherwise $[C(A,B)=C(B,A)=-0.5]$. The lengths of all virtual peptide bonds are set to unity.

In this paper, we report on the results obtained from statistical mechanics studies of the aggregation processes of short polymers. Our primary interest is devoted to the heteropolymer with the sequence S1: $AB_2AB_2ABAB_2AB$, which is a Fibonacci sequence,[18] whose single-chain properties are already known.[20] Throughout the paper, we are going to study the thermodynamics of systems with up to four chains of this sequence over the whole energy and temperature regime.

B. Simulation methods

We have used generalized-ensemble Markovian Monte Carlo algorithms to sample the conformational space of the systems studied. The powerful error-weighted multicanonical method[21–23] proved to be particularly useful as it makes it possible to scan the whole phase space with very high accuracy.[20] The principle idea is to deform the Boltzmann energy distribution

$$p_{\mathrm{can}}(E;T) \propto g(E) \exp(-E/k_B T), \tag{4}$$

where $g(E)$ is the density of states with energy E and $k_B T$ is the thermal energy at temperature T, in such a way that the notoriously difficult sampling of the tails is increased and—particularly useful—the sampling rate of the entropically strongly suppressed lowest-energy conformations is improved. In order to achieve this, the canonical Boltzmann distribution is modified by the multicanonical weight $W_{\mathrm{muca}}(E;T)$ which, in the ideal case, flattens the energy distribution:

$$p_{\mathrm{muca}}(E) = W_{\mathrm{muca}}(E;T) p_{\mathrm{can}}(E;T) = \mathrm{const}_{E;T}. \tag{5}$$

As the canonical distribution is, of course, not known in the beginning and $W_{\mathrm{muca}}(E;T) \sim p_{\mathrm{can}}^{-1}(E;T)$, the multicanonical weights have to be determined recursively, which can be done in an efficient way.[23,24] Recall that the simulation temperature T does not possess any meaning in the multicanonical ensemble as, according to Eq. (5), the energy distribution is always constant, independently of temperature. Actually, it is convenient to set it to infinity in which case $\lim_{T\to\infty} p_{\mathrm{can}}(E;T) \sim g(E)$ and, thus, $\lim_{T\to\infty} W_{\mathrm{muca}}(E;T) \sim g^{-1}(E)$. The latter expression is sometimes parametrized as $W_{\mathrm{muca}}(E) \sim \exp[-\beta(E)E + \alpha(E)]$, where, for a suitable choice of $\alpha(E)$, $\beta(E)$ can be identified with the microcanonical temperature.[24]

In our simulations, conformational changes of the individual chains included spherical updates[20] and semilocal crankshaft moves, i.e., rotations around the axis between the nth and $(n+2)$th residue. A typical multicanonical run contained of the order of 10^{10} single updates. The polymer chains were embedded into a cubic box with edge lengths L and periodic boundary conditions were used. In our simulations, the edge lengths of the simulation box were chosen to be $L=40$ which is sufficient to reduce undesired finite-size effects.

For cross-checks, we have also performed replica-exchange (parallel tempering) simulations.[25,26] Verifying the lowest-energy conformations found in the multicanonical simulations, we have also performed optimization runs using the energy-landscape paving method.[27]

C. Order parameter of aggregation and fluctuations

In order to distinguish between the fragmented and the aggregated regime, we introduce the order parameter

$$\Gamma^2 = \frac{1}{2M^2} \sum_{\mu,\nu=1}^{M} \mathbf{d}_{\mathrm{per}}^2(\mathbf{r}_{\mathrm{COM},\mu}, \mathbf{r}_{\mathrm{COM},\nu}), \tag{6}$$

where the summations are taken over the minimum distances $\mathbf{d}_{\mathrm{per}}=(d_{\mathrm{per}}^{(1)}, d_{\mathrm{per}}^{(2)}, d_{\mathrm{per}}^{(3)})$ of the respective centers of mass of the chains (or their periodic continuations). The center of mass of the μth chain in a box with periodic boundary conditions is defined as $\mathbf{r}_{\mathrm{COM},\mu}=\Sigma_{i_\mu=1}^{N_\mu}[\mathbf{d}_{\mathrm{per}}(\mathbf{r}_{i_\mu}, \mathbf{r}_{1_\mu})+\mathbf{r}_{1_\mu}]/N_\mu$, where $\mathbf{r}_{1_\mu}$ is the coordinate vector of the first monomer and serves as a reference coordinate in a local coordinate system.

Our aggregation parameter is to be considered as a qualitative measure; roughly, fragmentation corresponds to large values of Γ, aggregation requires the centers of masses to be in close distance, in which case, Γ is comparatively small. Despite its qualitative nature, it turns out to be a surprisingly manifest indicator for the aggregation transition and allows even a clear discrimination of different aggregation pathways, as will be seen later on.

According to the Boltzmann distribution [Eq. (4)], we define canonical expectation values of any observable O by

$$\langle O \rangle(T) = \frac{1}{Z_{\mathrm{can}}(T)} \prod_{\mu=1}^{M} \left[\int \mathcal{D}\mathbf{X}_\mu \right] O(\{\mathbf{X}_\mu\}) e^{-E(\{\mathbf{X}_\mu\})/k_B T}, \tag{7}$$

where the canonical partition function Z_{can} is given by

$$Z_{\mathrm{can}}(T) = \prod_{\mu=1}^{M} \left[\int \mathcal{D}\mathbf{X}_\mu \right] e^{-E(\{\mathbf{X}_\mu\})/k_B T}. \tag{8}$$

Formally, the integrations are performed over all possible conformations $\mathbf{X}_\mu$ of the M chains.

Similar to the specific heat per monomer $c_V(T) = d\langle E \rangle / N_{\mathrm{tot}} dT = (\langle E^2 \rangle - \langle E \rangle^2)/N_{\mathrm{tot}} k_B T^2$ (with $N_{\mathrm{tot}}=\Sigma_{\mu=1}^{M} N_\mu$), which expresses the thermal fluctuations of the energy, the temperature derivative of $\langle \Gamma \rangle$ per monomer, $d\langle \Gamma \rangle / N_{\mathrm{tot}} dT = (\langle \Gamma E \rangle - \langle \Gamma \rangle \langle E \rangle)/N_{\mathrm{tot}} k_B T^2$, is a useful indicator for cooperative behavior of the multiple-chain system. Since the system size is small—the number of monomers N_{tot} as well as the number of chains M—aggregation transitions, if any, are expected to be signalized by the peak structure of the fluctuating quantities as functions of the temperature. This requires the temperature to be a unique external control parameter which is a natural choice in the canonical statistical ensemble. Furthermore, this is a typically easily adjustable and, therefore, convenient parameter in experiments. As we have stressed recently,[17] however, aggregation is a phase separation process and, since the system is small, there is no uniform mapping between temperature and energy. For this reason, the total system energy is the more appropriate external parameter. Thus, the microcanonical interpretation will turn out to be the more favorable description, at least in the transition region. We will discuss this in detail in the following section.

III. STATISTICS OF THE TWO-CHAIN HETEROPOLYMER SYSTEM IN THREE ENSEMBLES

For the qualitative description of the aggregation and the accompanied conformational cooperativity within the whole system, it is sufficient to consider a very small system which is computationally reliably tractable and, thus, yields precise results for all energies and temperatures. Our heteropolymer system consists of two identical chains with the amino acid composition S1 and will be denoted as $2\times$S1. In the following, we discuss the aggregation behavior of this system from the multicanonical, the canonical, and the microcanonical point of view.

A. Multicanonical results

In a multicanonical simulation, the phase space is sampled in such a way that the energy distribution gets as flat as possible. Thermodynamically, this means that the sampling of the phase space is performed for all temperatures within a single simulation.[21–24] The desired information for the thermodynamic behavior of the system at a certain temperature is then obtained by simply reweighting the multicanonical into the respective canonical distribution, according to Eq. (5). Since the multicanonical ensemble contains all thermodynamic informations, including the conformational transitions, it is quite useful to measure within the simulation the multicanonical histogram

$$h_{\mathrm{muca}}(E_0, \Gamma_0) = \sum_{t_{\mathrm{muca}}} \delta_{E,E_0} \delta_{\Gamma,\Gamma_0}, \tag{9}$$

where t_{muca} labels the Monte Carlo "time" steps. More formally, this distribution can be expressed as a conformation-space integral

$$\begin{aligned} h_{\mathrm{muca}}(E_0,\Gamma_0) &\propto \langle \delta(E-E_0)\delta(\Gamma-\Gamma_0)\rangle_{\mathrm{muca}} \\ &= \frac{1}{Z_{\mathrm{muca}}} \prod_{\mu=1}^{M} \left[\int \mathcal{D}\mathbf{X}_\mu \right] \delta(E(\{\mathbf{X}_\mu\})-E_0)\delta(\Gamma(\{\mathbf{X}_\mu\})-\Gamma_0) \\ &\quad \times e^{-\mathcal{H}_{\mathrm{muca}}(E(\{\mathbf{X}_\mu\}))/k_B T} \propto e^{-\mathcal{F}_{\mathrm{muca}}(E_0,\Gamma_0)/k_B T} \end{aligned} \tag{10}$$

with the multicanonical energy $\mathcal{H}_{\mathrm{muca}}(E)=E - k_B T \ln W_{\mathrm{muca}}(E;T)$ which is independent of temperature.

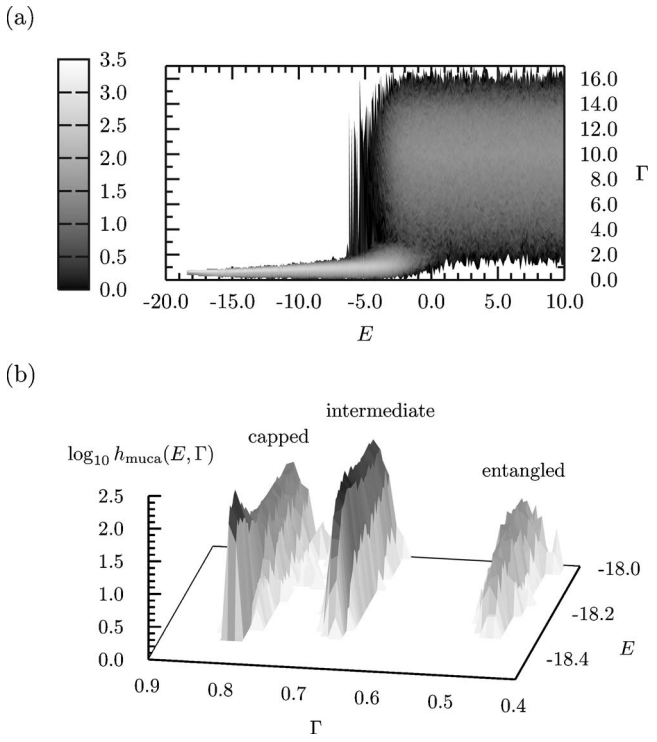

FIG. 1. (a) Multicanonical histogram $\log_{10} h_{\mathrm{muca}}$ as a function of energy E and aggregation parameter Γ, (b) section of $\log_{10} h_{\mathrm{muca}}$ in the low-energy tail.

The multicanonical partition function is also trivially a constant in temperature,

$$Z_{\mathrm{muca}} = \prod_{\mu=1}^{M} \left[\int \mathcal{D}\mathbf{X}_\mu \right] e^{-\mathcal{H}_{\mathrm{muca}}(E(\{\mathbf{X}_\mu\}))/k_B T} = \mathrm{const}_T. \quad (11)$$

It is obvious that integrating $h_{\mathrm{muca}}(E,\Gamma)$ over Γ recovers the uniform multicanonical energy distribution

$$\int_0^\infty d\Gamma h_{\mathrm{muca}}(E,\Gamma) \sim p_{\mathrm{muca}}(E). \quad (12)$$

The canonical distribution of energy and Γ parameter at temperature T can be retained, similar to inverting Eq. (5), by performing the simple reweighting

$$h_{\mathrm{can}}(E,\Gamma;T) = h_{\mathrm{muca}}(E,\Gamma) W^{-1}_{\mathrm{muca}}(E;T), \quad (13)$$

which is, due to the restriction to a certain temperature, less favorable to gain an overall impression of the phase behavior (i.e., the transition pathway) of the system, compared to the multicanonical analog $h_{\mathrm{muca}}(E,\Gamma)$.

In Fig. 1(a), $h_{\mathrm{muca}}(E,\Gamma)$ is shown for the two-peptide system $2\times S1$ as a color-coded projection onto the E-Γ plane, which is the direct output obtained in the multicanonical simulation. Qualitatively, we observe two separate main branches (which are "channels" in the corresponding free-energy landscape), between which a noticeable transition occurs. In the vicinity of the energy $E_{\mathrm{sep}} \approx -3.15$, both channels overlap, i.e., the associated macrostates coexist. Since Γ is an effective measure for the spatial distance between the two peptides, it is obvious that conformations with separated or fragmented peptides belong to the dominating channel in the regime of high energies and large Γ values, whereas the aggregates are accumulated in the narrow low-energy and small-Γ channel. Thus, the main observation from the multicanonical, comprising point of view is that the aggregation transition is a phase-separation process which, even for this small system, already appears in a surprisingly clear fashion.

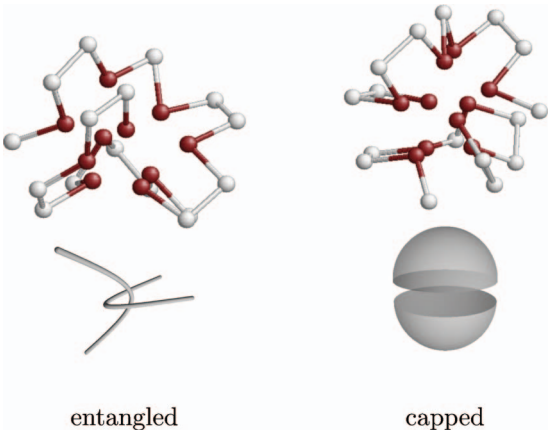

FIG. 2. (Color) Representatives and schematic characteristics of entangled and spherically capped conformations dominating the lowest-energy branches in the multicanonical histogram shown in Fig. 1(b). Dark spheres correspond to hydrophobic (A), light ones to polar (B) residues.

The high precision of the multicanonical method allows us even to reveal further details in the lowest-energy aggregation regime, which is usually a notoriously difficult sampling problem. Figure 1(b) shows that the tight aggregation channel splits into three separate, almost degenerate subchannels at lowest energies. From the analysis of the conformations in this region, we find that representative conformations with smallest Γ values, $\Gamma \approx 0.45$, are typically entangled, while those with $\Gamma \approx 0.8$ have a spherically capped shape. This is the subchannel connected to the lowest-energy states. Examples are shown in Fig. 2. The also highly compact conformations belonging to the intermediate subphase do not exhibit such characteristic features and are rather globules without noticeable internal symmetries. In all cases, the aggregates contain a single compact core of hydrophobic residues. Thus, the aggregation is not a simple docking process of two prefolded peptides, but a complex cooperative folding-binding process. This is a consequence of the energetically favored hydrophobic inter-residue contacts which, as the results show, overcompensate the entropic steric constraints. The story is, however, even more interesting, as also non-negligible surface effects come into play. After the following standard canonical analysis, this will be discussed in more detail in the subsequent microcanonical interpretation of our results.

B. Canonical perspective

Phase transitions are typically described in the canonical ensemble with the temperature kept fixed. This is also natural from an experimentalist's point of view, since the temperature is a convenient external control parameter. The macrostates are weighted according to the Boltzmann distribution [Eq. (4)]. A nice feature of the canonical ensemble is that the temperature dependence of fluctuations of thermody-

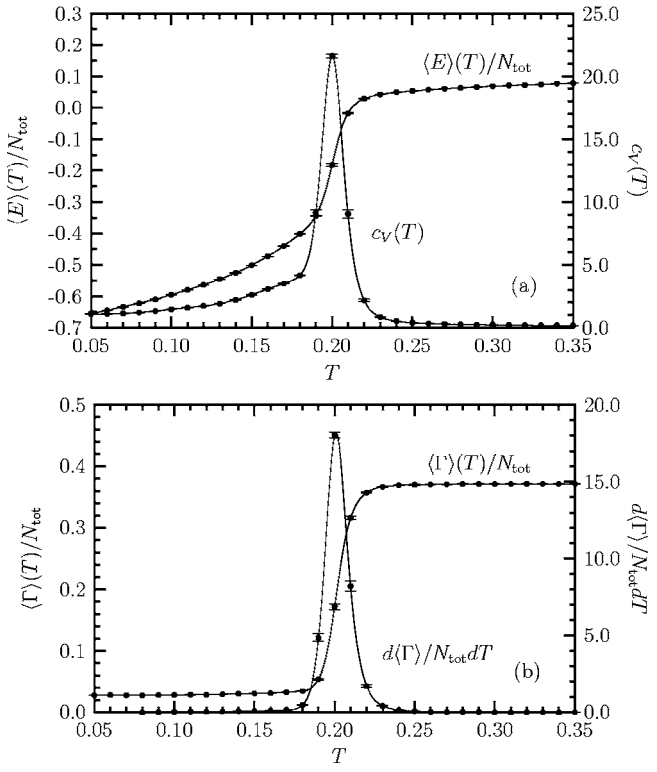

FIG. 3. (a) Mean energy $\langle E\rangle/N_{\rm tot}$ and specific heat per monomer c_V, and (b) $\langle\Gamma\rangle/N_{\rm tot}$ and $d\langle\Gamma\rangle/N_{\rm tot}dT$ as functions of the temperature.

namic quantities is usually a very useful indicator for phase or pseudophase transitions. This cooperative thermodynamic activity is typically signalized by peaks or, in the thermodynamic limit (if it exists), by divergences of these fluctuations. Even for small systems, peak temperatures can frequently be identified with transition temperatures. Although in these cases peak temperatures typically depend on the fluctuating quantities considered, in most cases associated pseudophase transitions are doubtlessly manifest. In such cases, the transition ranges over an extended temperature interval, as, e.g., in the folding process of proteins or heteropolymers.[28]

In our aggregation study of the $2\times S1$ system, however, we obtain from the canonical analysis a surprisingly clear picture of the aggregation transition. Figure 3(a) shows the canonical mean energy $\langle E\rangle$ and the specific heat per monomer c_V, plotted as functions of the temperature T. In Fig. 3(b), the temperature dependence of the mean aggregation order parameter $\langle\Gamma\rangle$ and the fluctuations of Γ are shown. The aggregation transition is signalized by very sharp peaks and from both figures we read off peak temperatures close to $T_{\rm agg}\approx 0.20$. The aggregation of the two peptides is a single-step process, in which the formation of the aggregate with a common compact hydrophobic core governs the folding behavior of the individual chains. Folding and binding are not separate processes.

The dominance of the interchain binding interaction can also be seen by considering the lowest-energy conformation found in our simulations. The energy of this conformation, which is shown in Fig. 4, is $E_{\rm min}\approx -18.4$ in our energy units. The peptide-peptide binding energy [i.e., the second term in Eq. (1)] is with $E^{(1,2)}_{AB,\rm min}\approx -11.4$ much stronger than

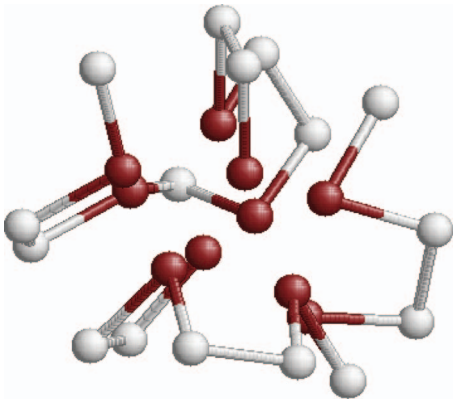

FIG. 4. (Color) The minimum-energy $2\times S1$ complex with $E_{\rm min}\approx -18.4$ as found in our simulations is a capped aggregate.

the intrinsic single-chain energies $E^{(1)}_{AB,\rm min}\approx -3.2$ and $E^{(2)}_{AB,\rm min}\approx -3.8$, respectively. The single-chain minimum energy is with $E^{\rm single}_{AB,\rm min}\approx -5.0$ noticeably smaller.[20]

The comparatively strong interchain interaction and the strength of the aggregation transition despite the smallness of the system lead to the conclusion that surface effects are of essential importance for the aggregation of the peptides. This is actually confirmed by a detailed microcanonical analysis which is performed in the next subsection.

C. Microcanonical interpretation

In the microcanonical analysis, the system energy E is kept (almost) fixed and treated as an external control parameter. The system can only take macrostates with energies in the interval $(E,E+\Delta E)$ with ΔE being sufficiently small to satisfy $\Delta G(E)=g(E)\Delta E$, where $\Delta G(E)$ is the phase-space volume of this energetic shell. In the limit $\Delta E\to 0$, the total phase-space volume up to the energy E can thus be expressed as

$$G(E)=\int_{E_{\rm min}}^{E} dE'\,g(E'). \tag{14}$$

Since $g(E)$ is positive for all E, $G(E)$ is a monotonically increasing function and this quantity is suitably related to the microcanonical entropy $\mathcal{S}(E)$ of the system. In the definition of Hertz,

$$\mathcal{S}(E)=k_B\ln G(E). \tag{15}$$

Alternatively, the entropy is often directly related to the density of states $g(E)$ and defined as

$$S(E)=k_B\ln g(E). \tag{16}$$

The density of states exhibits a decrease much faster than exponential toward the low-energy states. For this reason, the phase-space volume at energy E is strongly dominated by the number of states in the energy shell ΔE. Thus, $G(E)\approx\Delta G(E)\sim g(E)$ is directly related to the density of states. This virtual identity breaks down in the higher-energy region, where $\ln g(E)$ is getting flat—in our case far above the energetic regions being relevant for the discussion of the aggregation transition (i.e., for energies $E\gg E_{\rm frag}$, see Fig. 5). Actually, both definitions of the entropy led in our study to

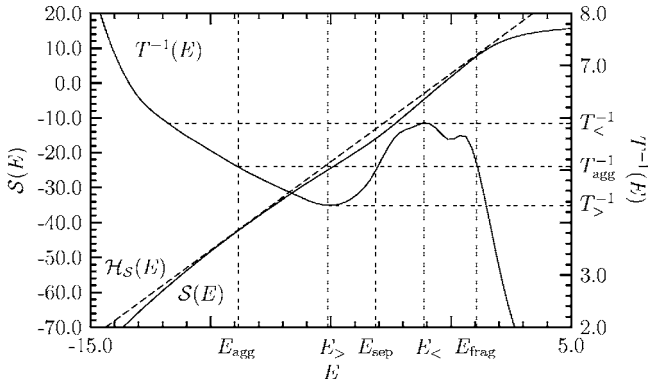

FIG. 5. Microcanonical Hertz entropy $\mathcal{S}(E)$ of the $2\times$S1 system, concave Gibbs hull $\mathcal{H}_\mathcal{S}(E)$, and inverse caloric temperature $T^{-1}(E)$ as functions of energy. The phase separation regime ranges from E_{agg} to E_{frag}. Between $T_<^{-1}$ and $T_>^{-1}$, the temperature is no suitable external control parameter and the canonical interpretation is not useful: The inverse caloric temperature $T^{-1}(E)$ exhibits an obvious backbending in the transition region. Note the second, less-pronounced backbending in the energy range $E_<<E<E_{frag}$.

virtually identical results in the analysis of the aggregation transition.[17] The (reciprocal) slope of the microcanonical entropy fixes the temperature scale and the corresponding caloric temperature is then defined via $T(E)=(\partial\mathcal{S}(E)/\partial E)^{-1}$ for fixed volume V and particle number N_{tot}.

As long as the mapping between the caloric temperature T and the system energy E is bijective, the canonical analysis of crossover and phase transitions is suitable since the temperature can be treated as external control parameter. For systems, where this condition is not satisfied, however, in a standard canonical analysis one may easily miss a physical effect accompanying condensation processes: Due to surface effects (the formation of the contact surface between the peptides requires a rearrangement of monomers in the surfaces of the individual peptides), additional energy does not necessarily lead to an increase of temperature of the condensate. Actually, the aggregate can even become colder. The supply of additional energy supports the fragmentation of parts of the aggregate, but this is overcompensated by cooperative processes of the particles aiming to reduce the surface tension. Condensation processes are phase-separation processes and as such aggregated and fragmented phases coexist. Since in this phase-separation region T and E are not bijective, this phenomenon is called the "backbending effect." The probably most important class of systems exhibiting this effect is characterized by their smallness and the capability to form aggregates, depending on the interaction range. The fact that this effect could be indirectly observed in sodium clustering experiments[8] gives rise to the hope that backbending could also be observed in aggregation processes of small peptides.

Since the $2\times$S1 system apparently belongs to this class, the backbending effect is also observed in the aggregation/fragmentation transition of this system. This is shown in Fig. 5, where the microcanonical entropy $\mathcal{S}(E)$ is plotted as function of the system energy. The phase-separation region of aggregated and fragmented conformations lies between $E_{agg}\approx-8.85$ and $E_{frag}\approx1.05$. Constructing the concave Gibbs hull $\mathcal{H}_\mathcal{S}(E)$ by linearly connecting $\mathcal{S}(E_{agg})$ and $\mathcal{S}(E_{frag})$ (straight dashed line in Fig. 5), the entropic deviation due to surface effects is simply $\Delta\mathcal{S}(E)=\mathcal{H}_\mathcal{S}(E)-\mathcal{S}(E)$. The deviation is maximal for $E=E_{sep}$ and $\Delta\mathcal{S}(E_{sep})\equiv\Delta\mathcal{S}_{surf}$ is the surface entropy. The Gibbs hull also defines the aggregation transition temperature

$$T_{agg}=(\partial\mathcal{H}_\mathcal{S}(E)/\partial E)^{-1}. \tag{17}$$

For the $2\times$S1 system, we find $T_{agg}\approx0.198$, which is virtually identical with the peak temperatures of the fluctuating quantities discussed in Sect. III B.

The inverse caloric temperature $T^{-1}(E)$ is also plotted into Fig. 5. For a fixed temperature in the interval $T_<<T<T_>$ ($T_<\approx0.169$ and $T_>\approx0.231$), different energetic macrostates coexist. This is a consequence of the backbending effect. Within the backbending region, the temperature decreases with increasing system energy. The horizontal line at $T_{agg}^{-1}\approx5.04$ is the Maxwell construction, i.e., the slope of the Gibbs hull $\mathcal{H}_\mathcal{S}(E)$. Although the transition seems to have similarities with the van der Waals description of the condensation/evaporation transition of gases—the "overheating" of the aggregate between T_{agg} and $T_>$ (within the energy interval $E_{agg}<E<E_>\approx-5.13$) is as apparent as the "undercooling" of the fragments between $T_<$ and T_{agg} (in the energy interval $E_{frag}>E>E_<\approx-1.13$)—it is important to notice that in contrast to the van der Waals picture, the backbending effect in-between is a real physical effect. Another essential result is that in the transition region the temperature is not a suitable external control parameter: The macrostate of the system cannot be adjusted by fixing the temperature. The better choice is the system energy which is unfortunately difficult to control in the experiments. Another direct consequence of the energetic ambiguity for a fixed temperature between $T_<$ and $T_>$ is that the canonical interpretation is not suitable for detecting the backbending phenomenon. It should also be noted that in this region, the microcanonical specific heat $c_V(E)$ can become negative,[17] which is a remarkable, but somehow "exotic" side effect.

The precise microcanonical analysis reveals also a further detail of the aggregation transition. Close to $E_{pre}\approx-0.32$, the T^{-1} curve in Fig. 5 exhibits another "backbending" which signalizes a second, but unstable transition of the same type. The associated transition temperature $T_{pre}\approx0.18$ is smaller than T_{agg}, but this transition occurs in the energetic region where fragmented states dominate. Thus, this transition can be interpreted as the premelting of aggregates by forming intermediate states. These intermediate structures are rather weakly stable: the population of the premolten aggregates never dominates. In particular, at T_{pre}, where premolten aggregates and fragments coexist, the population of compact aggregates is much larger. This can nicely be seen in the canonical energy histograms at these temperatures plotted in Fig. 6, where the second backbending is only signalized by a small cusp in the coexistence region. Since both transitions are phase-separation processes, structure formation is accompanied by releasing latent heat which can be defined as the energetic widths of the phase coexistence regimes, i.e., $\Delta Q_{agg}=E_{frag}-E_{agg}=T_{agg}[\mathcal{S}(E_{frag})-\mathcal{S}(E_{agg})]\approx9.90$ and $\Delta Q_{pre}=E_{frag}-E_{pre}=T_{pre}[\mathcal{S}(E_{frag})-\mathcal{S}(E_{pre})]\approx1.37$.

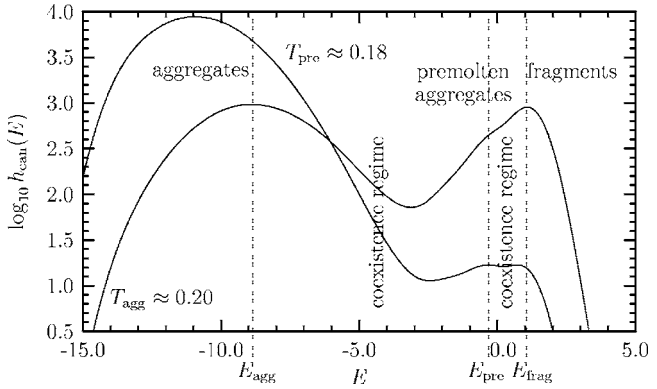

FIG. 6. Logarithmic plots of the canonical energy histograms (not normalized) at $T \approx 0.18$ and $T \approx 0.20$, respectively.

Obviously, the energy required to melt the premolten aggregate is much smaller than to dissolve a compact (solid) aggregate.

For the comparison of the surface entropies, we use the definition (16) of the entropy. In the case of the aggregation transition, the surface entropy is $\Delta S_{surf}^{agg} \approx \Delta S_{surf}^{agg} = H_S(E_{sep}) - S(E_{sep})$, where $H_S(E) \approx \mathcal{H}_S(E)$ is the concave Gibbs hull of $S(E)$. Since $H_S(E_{sep}) = H_S(E_{frag}) - (E_{frag} - E_{sep})/T_{agg}$ and $H_S(E_{frag}) = S(E_{frag})$, the surface entropy is

$$\Delta S_{surf}^{agg} = S(E_{frag}) - S(E_{sep}) - \frac{1}{T_{agg}}(E_{frag} - E_{sep}). \tag{18}$$

Yet utilizing that the canonical distribution $h_{can}(E)$ at T_{agg} shown in Fig. 6 is $h_{can}(E) \sim g(E)\exp(-E/k_B T_{agg})$, the surface entropy can be written in the simple and computationally convenient form[11]

$$\Delta S_{surf}^{agg} = k_B \ln \frac{h_{can}(E_{frag})}{h_{can}(E_{sep})}. \tag{19}$$

A similar expression is valid for the coexistence of premolten and fragmented states at T_{pre}. The corresponding canonical distribution is also shown in Fig. 6. Thus, we obtain (in units of k_B) for the surface entropy of the aggregation transition $\Delta S_{surf}^{agg} \approx 2.48$ and for the premelting $\Delta S_{surf}^{pre} \approx 0.04$, confirming the weakness of the interface between premolten aggregates and fragmented states.

IV. AGGREGATION TRANSITION IN LARGER HETEROPOLYMER SYSTEMS

In order to verify the general validity of the statements in the previous section for the $2 \times S1$ system, we have also performed simulations of systems consisting of three (in the following referred to as $3 \times S1$) and four ($4 \times S1$) identical peptides with sequence S1.

Although the formation of compact hydrophobic cores is more complex in larger compounds of our exemplified sequence S1, the aggregation transition is little influenced by this. This is nicely seen in Figs. 7(a) and 7(b), where the temperature dependence of the canonical expectation values of Γ and E, as well as for their fluctuations, are shown for the $3 \times S1$ system. For comparison, also results for the $4 \times S1$ system are plotted into the same figures. Note that for the

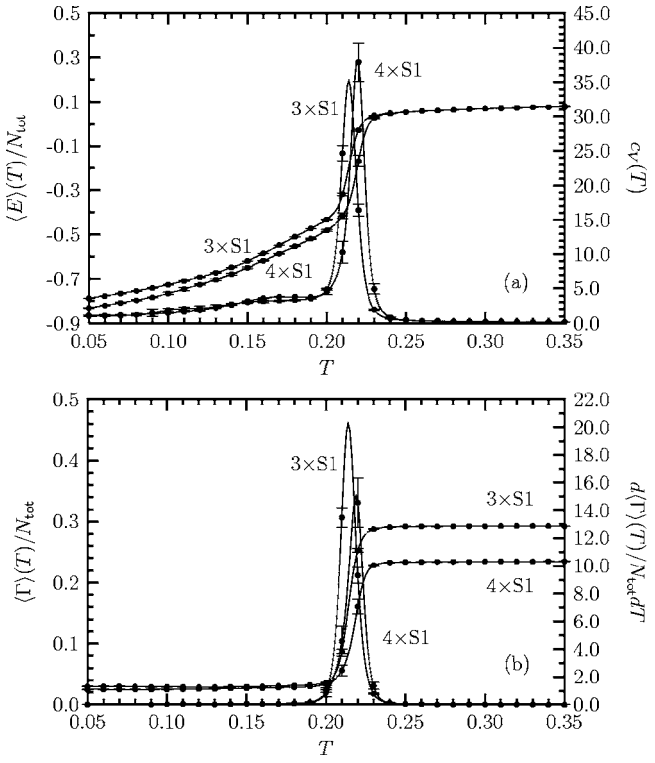

FIG. 7. (a) Mean energy $\langle E \rangle / N_{tot}$ and specific heat per monomer c_V, (b) mean aggregation parameter $\langle \Gamma \rangle / N_{tot}$ and its fluctuations $d\langle \Gamma \rangle / N_{tot} dT$ as functions of the temperature for the $3 \times S1$ and $4 \times S1$ heteropolymer systems.

$4 \times S1$ system finite-size effects are larger since, for computational reasons, we have kept the edge length of the simulation box $L=40$, which is smaller than the successive arrangement of four straight chains with 13 monomers. This influences primarily the entropy in the high-energy regime far above the aggregation transition energy. Nonetheless, in the canonical interpretation, it acts back onto the transition as undesired states (chain ends overlapping due to the periodic boundary conditions) are (weakly) populated at the transition temperature, whereas others are suppressed. We have performed a detailed analysis of the box size dependence (results not shown) and found that the canonical transition temperature scales slightly, but noticeably with the box size. Thus, the results obtained by *canonical* statistics for the $4 \times S1$ system should not quantitatively be compared to the canonical results for the $2 \times S1$ and $3 \times S1$ systems.

As has already been discussed for the $2 \times S1$ system, there are also for the larger systems no obvious signals for separate aggregation and hydrophobic-core formation processes. Only weak activity in the energy fluctuations in the temperature region below the aggregation transition temperature indicates that local restructuring processes of little cooperativity (comparable with the discussion of the premolten aggregates in the discussion of the $2 \times S1$ system) are still happening. The strength of the aggregation transition is also documented by the fact that the peak temperatures of energetic *and* aggregation parameter fluctuations are virtually identical for the $3 \times S1$ system, i.e., the aggregation temperature is $T_{agg} \approx 0.21$ (for $4 \times S1$ $T_{agg} \approx 0.22$).

For homogeneous multiple-chain systems, two variants

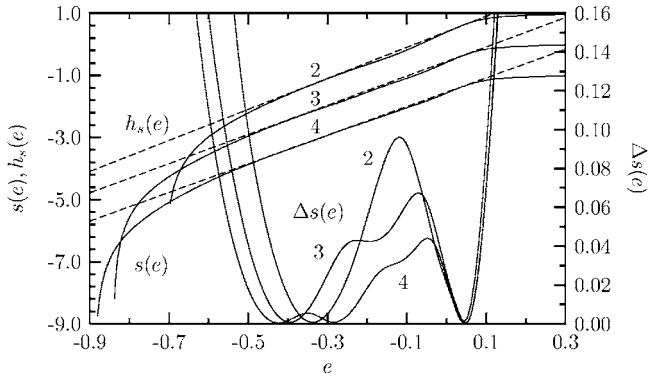

FIG. 8. Microcanonical entropies per monomer $s(e)$, respective Gibbs constructions $h_s(e)$ (left scale), and deviations $\Delta s(e)=h_s(e)-s(e)$ (right scale) for $2\times S1$ (labeled as "2"), $3\times S1$ ("3"), and $4\times S1$ ("4") as functions of the energy per monomer e.

of thermodynamic limits are of particular interest: (i) $M\to\infty$, while $N_\mu=$const, (ii) $N_\mu\to\infty$ with $M=$const; both limits considered for constant polymer density. Since for proteins the sequence of amino acids is fixed, in this case only (i) is relevant and it is future work to perform a scaling analysis for multiple-peptide systems in this limit. A particularly interesting question is to what extent remnants of the finite-system effects, as discussed in this paper, survive in the limit of an infinite number of chains, dependent on the peptide density. Since we have focused our study on the precise analysis of systems of few peptides for all energies and temperatures, it was computationally inevitable to restrict ourselves to small systems, for which a scaling analysis is not very useful. Nonetheless, we would like to devote a few interesting remarks to the comparison of, once more, microcanonical aspects of the aggregation transition in dependence of the system size.

In Fig. 8, the microcanonical entropies per monomer $s(e)=\mathcal{S}(e)/N_{\text{tot}}$ (shifted by an unimportant constant for clearer visibility) and the corresponding Gibbs hulls $h_s(e)=\mathcal{H}_{\mathcal{S}}(e)/N_{\text{tot}}$ are shown for $2\times S1$ (in the figure denoted by "2"), $3\times S1$ ("3"), and $4\times S1$ ("4"), respectively, as functions of the energy per monomer $e=E/N_{\text{tot}}$. Although the convex entropic "intruder" is apparent for larger systems as well, its relative strength decreases with increasing number of chains. The slopes of the respective Gibbs constructions determine the aggregation temperature [Eq. (17)], which are found to be $T_{\text{agg}}^{3\times S1}\approx 0.212$ and $T_{\text{agg}}^{4\times S1}\approx 0.217$, confirming the peak temperatures of the fluctuation quantities plotted in Fig. 7.

The existence of the interfacial boundary entails a transition barrier whose strength is characterized by the surface entropy $\Delta\mathcal{S}_{\text{surf}}$. In Fig. 8, the individual entropic deviations per monomer, $\Delta s(e)=\Delta\mathcal{S}(e)/N_{\text{tot}}$ are also shown and the maximum deviations, i.e., the surface entropies $\Delta\mathcal{S}_{\text{surf}}$ and relative surface entropies per monomer $\Delta s_{\text{surf}}=\Delta\mathcal{S}_{\text{surf}}/N_{\text{tot}}$ are listed in Table I. There is no apparent difference between the values of $\Delta\mathcal{S}_{\text{surf}}$ that would indicate a trend for a vanishing of the *absolute* surface barrier in larger systems. However, the *relative* surface entropy Δs_{surf} obviously decreases. Whether or not it vanishes in the thermodynamic limit cannot be decided from our results and is a study worth in its own right.

It is also interesting that subleading effects increase and the double-well form found for $2\times S1$ changes by higher-order effects, and it seems that for larger systems the almost single-step aggregation of $2\times S1$ is replaced by a multiple-step process.

Not surprisingly, the fragmented phase is hardly influenced by side effects and the right-most minimum in Fig. 8 lies well at $e_{\text{frag}}=E_{\text{frag}}/N_{\text{tot}}\approx 0.04-0.05$. Since the Gibbs construction covers the whole convex region of $s(e)$, the aggregation energy per monomer $e_{\text{agg}}=E_{\text{agg}}/N_{\text{tot}}$ corresponds to the left-most minimum and its value changes noticeably with the number of chains. As consequence, the latent heat per monomer $\Delta q=\Delta Q/N_{\text{tot}}=T_{\text{agg}}[\mathcal{S}(E_{\text{frag}})-\mathcal{S}(E_{\text{agg}})]/N_{\text{tot}}$ that is required to fragment the aggregate increases from two to four chains in the system (see Table I). Although the systems under consideration are too small to extrapolate phase transition properties in the thermodynamic limit, it is obvious that the aggregation/fragmentation transition exhibits strong similarities to condensation/evaporation transitions of colloidal systems. This suggests that the entropic transition barrier $\Delta q/T_{\text{agg}}$, which increases with the number of chains (cf. the values in Table I), would survive in the thermodynamic limit and the transition was first-order-like. More surprising would be, however, if the convex intruder would not disappear, i.e., if the absolute and relative surface entropies $\Delta\mathcal{S}_{\text{surf}}$ and Δs_{surf} do not vanish. This is definitely a question of fundamental interest as the common claim is that pure surface effects typically exhibited only by "small" systems are irrelevant in the thermodynamic limit. This requires, however, studies of much larger systems. It should clearly be noted, however, that protein aggregates forming themselves in biological systems often consist only of a few peptides and are definitely of small size and the surface effects are responsible for structure formation and are not unimportant side effects. One should keep in mind that standard thermodynamics and the thermodynamic limit are somewhat theoretical constructs valid only for very large systems. The increasing interest in

TABLE I. Aggregation temperatures T_{agg}, surface entropies $\Delta\mathcal{S}_{\text{surf}}$, relative surface entropies per monomer Δs_{surf}, relative aggregation and fragmentation energies per monomer, e_{agg} and e_{frag}, respectively, latent heat per monomer Δq, and phase-separation entropy per monomer $\Delta q/T_{\text{agg}}$. All quantities for systems consisting of two, three, and four 13mers with sequence S1.

System	T_{agg}	$\Delta\mathcal{S}_{\text{surf}}$	Δs_{surf}	e_{agg}	e_{frag}	Δq	$\Delta q/T_{\text{agg}}$
$2\times S1$	0.198	2.48	0.10	−0.34	0.04	0.38	1.92
$3\times S1$	0.212	2.60	0.07	−0.40	0.05	0.45	2.12
$4\times S1$	0.217	2.30	0.04	−0.43	0.05	0.48	2.21

physical properties of small systems, in particular, in conformational transitions in molecular systems, requires in part a revision of dogmatic thermodynamic views. Indeed, by means of today's chemo-analytical and experimental equipment, effects such as those described throughout the paper, should actually experimentally be verifiable as these are real physical effects. For studies of the condensation of atoms, where a similar behavior occurs, such experiments have actually already been performed.[8]

V. SUMMARY

In this paper, we have extended the microcanonical analysis of the aggregation of an exemplified two-peptide system[17] by interpreting the results from the multicanonical and the canonical perspective as well. In addition, these results are compared with aggregation properties of larger systems consisting of three and four peptides, each of which with the same sequence. From the conventional canonical analysis of statistical fluctuations of energy and a suitably chosen order parameter—the root mean square distance of the centers of masses of the individual polymers—we obtain the typical small-system indications of a thermodynamic phase transition: Sharp peaks in the specific heat and in the order parameter fluctuations at almost the same temperature signalize a strong transition, which we clearly identify as the aggregation transition. For all systems considered, the general behavior is similar. There is only this single transition which also indicates that conformational changes of the polymers accompany the aggregation process and are not separate transitions. We expect that this coincidence is sequence-dependent and a comparison between different sequences would be a study in its own right. At least for the semiflexible homopolymer of same size which in our notation would have the sequence A_{13} (or also B_{13}), we find that aggregation and collapse are separate processes.[29]

A quite remarkable result of the exemplified heteropolymer study presented in this paper is that the aggregation process of a small number of peptides is a phase-separation process, where interfacial surface effects entail a loss of entropy. This loss must be compensated by additional energy delivered to the system. In consequence, the caloric temperature decreases, i.e., the aggregate is getting colder although its total energy increases. This is known as the temperature backbending which is a real thermodynamic effect and not an artefact of the theory. In the systems considered throughout the paper, the relative influence of the surface effects seems to decrease with the number of chains in the system. Since the length of the peptides is fixed by their hydrophobic-polar monomer composition, a thermodynamic limit towards infinite chain lengths is, however, not existing. It is just the smallness of such molecular systems that allows these to trigger biological interchange processes which are inevitably connected with conformational activity.

ACKNOWLEDGMENTS

This work is partially supported by the DFG (German Science Foundation) under Grant No. JA 483/24-1/2. M.B. thanks the DFG and the Wenner-Gren Foundation for research fellowships. Support by the DAAD-STINT Personnel Exchange Programme with Sweden is gratefully acknowledged. We also thank the John von Neumann Institute for Computing (NIC), Forschungszentrum Jülich, for supercomputer time under Grant No. hlz11.

[1] J. Gsponer and M. Vendruscolo, Prot. Pept. Lett. **13**, 287 (2006).
[2] H. Lin, R. Bhatia, and R. Lal, FASEB J. **15**, 2433 (2001).
[3] A. Quist, I. Doudevski, H. Lin, R. Azimova, D. Ng, B. Frangione, B. Kagan, J. Ghiso, and R. Lal, Proc. Natl. Acad. Sci. U.S.A. **102**, 10427 (2005).
[4] H. A. Lashuel and P. T. Lansbury, Jr., Q. Rev. Biophys. **39**, 167 (2006).
[5] D. H. E. Gross, *Microcanonical Thermodynamics* (World Scientific, Singapore, 2001).
[6] D. H. E. Gross and J. F. Kenney, J. Chem. Phys. **122**, 224111 (2005).
[7] W. Thirring, Z. Physik **235**, 339 (1970).
[8] M. Schmidt, R. Kusche, T. Hippler, J. Donges, W. Kronmüller, B. von Issendorff, and H. Haberland, Phys. Rev. Lett. **86**, 1191 (2001).
[9] M. Pichon, B. Tamain, R. Bougault, and O. Lopez, Nucl. Phys. A **749**, 93c (2005).
[10] O. Lopez, D. Lacroix, and E. Vient, Phys. Rev. Lett. **95**, 242701 (2005).
[11] W. Janke, Nucl. Phys. B (Proc. Suppl.) **63A-C**, 631 (1998).
[12] H. Behringer and M. Pleimling, Phys. Rev. E **74**, 011108 (2006).
[13] D. J. Wales and R. S. Berry, Phys. Rev. Lett. **73**, 2875 (1994); D. J. Wales and J. P. K. Doye, J. Chem. Phys. **103**, 3061 (1995).
[14] S. Hilbert and J. Dunkel, Phys. Rev. E **74**, 011120 (2006); J. Dunkel and S. Hilbert, Physica A **370**, 390 (2006).
[15] A. Nußbaumer, E. Bittner, T. Neuhaus, and W. Janke, Europhys. Lett. **75**, 716 (2006).
[16] D. H. E. Gross, Physica E (Amsterdam) **29**, 251 (2005).
[17] C. Junghans, M. Bachmann, and W. Janke, Phys. Rev. Lett. **97**, 218103 (2006).
[18] 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).
[19] S. Schnabel, M. Bachmann, and W. Janke, Phys. Rev. Lett. **98**, 048103 (2007); J. Chem. Phys. **126**, 105102 (2007).
[20] M. Bachmann, H. Arkın, and W. Janke, Phys. Rev. E **71**, 031906 (2005).
[21] B. A. Berg and T. Neuhaus, Phys. Lett. B **267**, 249 (1991); Phys. Rev. Lett. **68**, 9 (1992).
[22] W. Janke, Physica A **254**, 164 (1998); B. A. Berg, Fields Inst. Commun. **26**, 1 (2000).
[23] W. Janke, in *Computer Simulations of Surfaces and Interfaces*, NATO Advanced Studies Institute, Series II. Mathematics, Physics and Chemistry, edited by B. Dünweg, D. P. Landau, and A. I. Milchev (Kluwer, Dordrecht, 2003), Vol. 114, p. 137.
[24] T. Çelik and B. A. Berg, Phys. Rev. Lett. **69**, 2292 (1992).
[25] K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996); K. Hukushima, H. Takayama, and K. Nemoto, Int. J. Mod. Phys. C **7**, 337 (1996).
[26] C. J. Geyer, in *Computing Science and Statistics*, Proceedings of the 23rd Symposium on the Interface, edited by E. M. Keramidas (Interface Foundation, Fairfax Station, 1991), p. 156.
[27] U. H. E. Hansmann and L. T. Wille, Phys. Rev. Lett. **88**, 068105 (2002).
[28] M. Bachmann and W. Janke, Phys. Rev. Lett. **91**, 208105 (2003); J. Chem. Phys. **120**, 6779 (2004).
[29] C. Junghans, M. Bachmann, and W. Janke (unpublished).