PHYSICAL REVIEW LETTERS

Microcanonical Analyses of Peptide Aggregation Processes

Christoph Junghans,[*] Michael Bachmann,[†] and Wolfhard Janke[‡]
*Institut für Theoretische Physik and Centre for Theoretical Sciences (NTZ), Universität Leipzig,
Augustusplatz 10/11, D-04109 Leipzig, Germany*

We propose the use of microcanonical analyses for numerical studies of peptide aggregation transitions. Performing multicanonical Monte Carlo simulations of a simple hydrophobic-polar continuum model for interacting heteropolymers of finite length, we find that the microcanonical entropy behaves convex in the transition region, leading to a negative microcanonical specific heat. As this effect is also seen in first-order-like transitions of other finite systems, our results provide clear evidence for recent hints that the characterization of phase separation in first-order-like transitions of finite systems profits from this microcanonical view.

PACS numbers: 87.15.Aa, 05.10.−a, 87.15.Cc

Thermodynamic phase transitions in macroscopic, infinitely large systems are typically analyzed in the thermodynamic limit of a canonical ensemble; i.e., the temperature T is treated as an intensive external control parameter adjusted by the heat bath, and the total system energy E is distributed according to the Boltzmann-Gibbs statistics. The probability for a macrostate with energy E is given by $p(E) = g(E) \exp(-E/k_B T)/Z$, where $g(E)$ is the density of states, Z the partition sum, and k_B the Boltzmann constant. As long as the microcanonical entropy $S(E) = k_B \ln g(E)$ is a concave function of E, the microcanonical (caloric) temperature $T(E) = (\partial S(E)/\partial E)^{-1}$ for fixed volume V and particle number $\mathcal{N}$ is a monotonically increasing function of E. Consequently, the microcanonical specific heat $C_V(E) = \partial E/\partial T(E) = -(\partial S/\partial E)^2/(\partial^2 S/\partial E^2)$ is positive. The specific heat can become negative only in an energetic regime, where $S(E)$ is convex. In this region, the caloric $T(E)$ curve exhibits a typical backbending, which means that the system becomes colder with increasing total energy. For this reason, the temperature T is not the most appropriate control parameter, and the analysis of such typically finite systems is more adequately performed in the microcanonical ensemble, where the system energy E is considered as the adjustable external parameter [1,2].

It is a surprising fact that the backbending effect is indeed observed in transitions with phase separation. Although this phenomenon has already been known for a long time from astrophysical systems [3], it has been widely ignored since then as a somehow "exotic" effect. Recently, however, experimental evidence was found from melting studies of sodium clusters by photofragmentation [4]. Bimodality and negative specific heats are also known from nuclei fragmentation experiments and models [5,6], as well as from spin models on finite lattices which experience first-order transitions in the thermodynamic limit [7,8]. This phenomenon is also observed in a large number of other isolated finite model systems for evaporation and melting effects [9,10].

In this Letter, we demonstrate the usefulness of the microcanonical ensemble for studies of the aggregation process of small proteins (peptides), which, due to the fixed inhomogeneous sequence of amino acids, are necessarily systems of finite size. Understanding protein aggregation is essential not only for gaining insights into general mechanisms of protein folding but also for unraveling the reasons for human diseases caused by protein clustering. A well-known example is associated with Alzheimer's disease, where a few identical small fragments of large proteins show the tendency to form fibrils, e.g., the hydrophobic $A\beta_{16-22}$ segment of the β-amyloid peptide $A\beta$ [11].

Our results are based on computer simulations of a simple continuum aggregation model for heteropolymers. Since the hydrophobic force governs the tertiary folding process resulting in a compact hydrophobic core surrounded by a shell of mainly polar residues, in our model the 20 amino acids naturally occurring in proteins are classified as hydrophobic (A) and polar (B) [12]. For the individual peptides, we employ the AB model [13] in three spatial dimensions. At a mesoscopic length scale, this coarse-grained model with virtual peptide bonds and virtual bond angles has proven quite successful in the qualitative characterization of naturally observed protein folding channels [14]. Keeping the same parameter sets for the interaction of monomers of *different* polymers, the model for the aggregate reads:

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

where μ and ν label the M polymers interacting with each other, and i_μ and j_μ index the N monomers of the μth polymer whose intrinsic energy is given by

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

with $0 \leq \vartheta_{i_\mu} \leq \pi$ denoting the bending angle between

218103-1

monomers i_μ, $i_\mu + 1$, and $i_\mu + 2$. The nonbonded inter-residue pair potential

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

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 our aggregation study, we have performed multicanonical simulations [15] for two identical peptides with 13 monomers and the sequence 13.1: $AB_2AB_2ABAB_2AB$ is arbitrarily chosen from the Fibonacci series [13]. For consistency, the simulations were repeated for pairs of identical homopolymers $2 \times A_{13}$ and $2 \times B_{13}$, as well as the larger aggregates 3×13.1 and 4×13.1. In all cases, aggregation behaviors of similar type as for the two-peptide system were identified [16].

For the simulations of the 2×13.1 system, the peptides were confined in a periodic cube with edge lengths $L = 40$. We varied the edge lengths to make sure that effects due to this confinement are negligible. A sequence of spherical-cap updates [17] and three-monomer corner rotations ensured an ergodic scan of the conformational space. After performing 180 multicanonical recursions, a total number of 2×10^{10} updates was generated. The primary result of these simulations is, up to an unimportant constant, the density of states $g(E)$ which has been precisely estimated over about 100 orders of magnitude.

In Fig. 1(a), the microcanonical entropy $S(E) = \ln g(E)$ ($k_B \equiv 1$) [18] is plotted (up to an unimportant additive constant) for the two-peptide system, ranging from the aggregate phase including the lowest energy found in the simulation ($E_{\min} = E_{AB,\min}^{(1)} + E_{AB,\min}^{(2)} + E_{AB,\min}^{(1,2)} \approx -18.407$) to the phase of the fragmented polymers. The conformation of the lowest-energy aggregate has a two-cap-like, globular shape with a compact hydrophobic core jointly formed by the two heteropolymers; see the inset in Fig. 1(a). It should be noted that the individual conformations in the aggregate strongly differ from the single-peptide ground states ($E_{\min}^{single} \approx -4.967$ [17]), and their respective energies in the aggregate are much larger ($E_{AB,\min}^{(1)} \approx -3.197$, $E_{AB,\min}^{(2)} \approx -3.798$). The strongest contribution is due to the interaction between the heteropolymers ($E_{AB,\min}^{(1,2)} \approx -11.412$).

The most interesting region in Fig. 1(a) is the phase coexistence regime $E_{agg} \approx -8.85 \le E \le 1.05 \approx E_{frag}$, where the entropy exhibits a convex intruder. The concave hull $\mathcal{H}_S(E) = S(E_{agg}) + E/T_{agg}$, which is the tangent connecting $S(E_{agg})$ and $S(E_{frag})$, is the Gibbs construction. Its slope defines the inverse of the aggregation temperature $T_{agg} \approx 0.198$. The interval $\Delta Q = E_{frag} - E_{agg} = T_{agg}[S(E_{frag}) - S(E_{agg})] \approx 9.90$ is the latent heat required to release interchain contacts at the aggregation tempera-
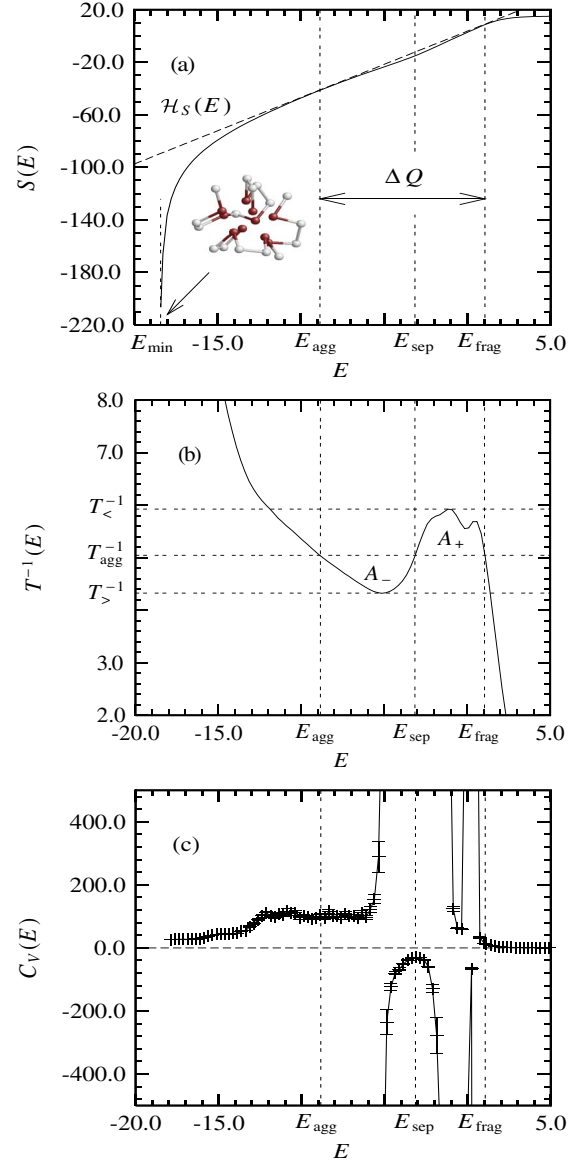

FIG. 1 (color online). Aggregation transition from the microcanonical perspective: (a) microcanonical entropy $S(E)$ (up to a constant) and concave hull $\mathcal{H}_S(E)$, (b) inverse caloric temperature $T^{-1}(E)$, and (c) specific heat $C_V(E)$. The errors are very small and therefore only shown for $C_V(E)$.

ture T_{agg}. The energy, where the difference $\Delta S(E) = \mathcal{H}_S(E) - S(E)$ is maximal, is denoted as E_{sep}, and the associated maximum deviation is the surface entropy $\Delta S_{surf} \equiv \Delta S(E_{sep})$. The derivative of the Gibbs construction gives the Maxwell line $T_{agg}^{-1} = $ const ≈ 5.043 in the reciprocal caloric $T^{-1}(E)$ curve, which is shown in Fig. 1(b). A bijective mapping between T and E is possible only for $T > T_> \approx 0.231$ and $T < T_< \approx 0.169$. This means, for values above $T_>$ and below $T_<$, that the temperature T is a useful control parameter. The two-heteropolymer system forms an aggregate for $T < T_<$, where the separation into individual polymers is not useful because interpolymer attraction dominates over intrinsic

structure formation and the aggregate determines the mesoscopic length and energy scale. On the other hand, for $T > T_>$, the polymers are only weakly interacting fragments; i.e., they can be considered separately: The total system energy is an *extensive* variable ($E \approx E_{AB}^{(1)} + E_{AB}^{(2)}$). *Only in these two temperature regions,* the interpretation of the canonical formalism is generic.

In the transition region $T_< \leq T \leq T_>$, however, the interaction strength between the polymers is as strong as intrinsic monomer-monomer attraction and cannot be neglected. As a consequence of the convexity of $S(E)$ in the interval $E_{agg} < E < E_{frag}$, there is no one-to-one correspondence between temperature and energy in the transition regime which results in the backbending effect: Fragmentation of the aggregate leads to a decrease of temperature, although the system energy increases. The areas $A_+ = T_{agg}^{-1}(E_{sep} - E_{frag}) - [S(E_{frag}) - S(E_{sep})]$ and $A_- = T_{agg}^{-1}(E_{sep} - E_{agg}) - [S(E_{sep}) - S(E_{agg})]$ formed by the Maxwell line and the $T^{-1}(E)$ curve as shown in Fig. 1(b) are identical. These areas determine the interfacial entropy $\Delta S_{surf} = A_+ = A_-$ [7], which is interpreted as the loss of entropy due to the existence of the phase boundary [19] between the aggregate and the fragment macrostates of the polymers. Consequently, as the energy of the total system is not extensive in the transition region, E is the favored control parameter compared with T. Therefore, the aggregation transition is more favorably analyzed in the microcanonical ensemble, at least for such finite systems like the heteropolymers in our study, where an extension towards the thermodynamic limit is not possible.

The most remarkable result is the negativity of the specific heat of the system in the backbending region, as shown in Fig. 1(c). A negative specific heat in the phase separation regime is due to the nonextensivity of the energy of the two subsystems resulting from the interaction between the polymers. "Heating" a *large* aggregate would lead to the stretching of monomer-monomer contact distances; i.e., the potential energy of an exemplified pair of monomers increases, while kinetic energy and, therefore, temperature remain widely constant. In a comparatively *small* aggregate, additional energy leads to cooperative rearrangements of monomers in the aggregate in order to reduce surface tension; i.e., the formation of molten globular aggregates is suppressed. As a consequence, kinetic energy is transferred into potential energy and the temperature decreases. In this regime, the aggregate becomes colder, although the total energy increases.

Figure 2(a) shows the typical bimodal canonical energy distribution $H(E) \sim g(E) \exp(-E/k_B T_{agg})$ close to the transition temperature T_{agg}. The maximum points are identical with the energies of the phase boundaries E_{agg} and E_{frag}, and the minimum is found at E_{sep} [7]. For this reason, the difference of the energies belonging to the maximum points of the canonical distribution is identical with the latent heat ΔQ. The minimum of this distribution coincides
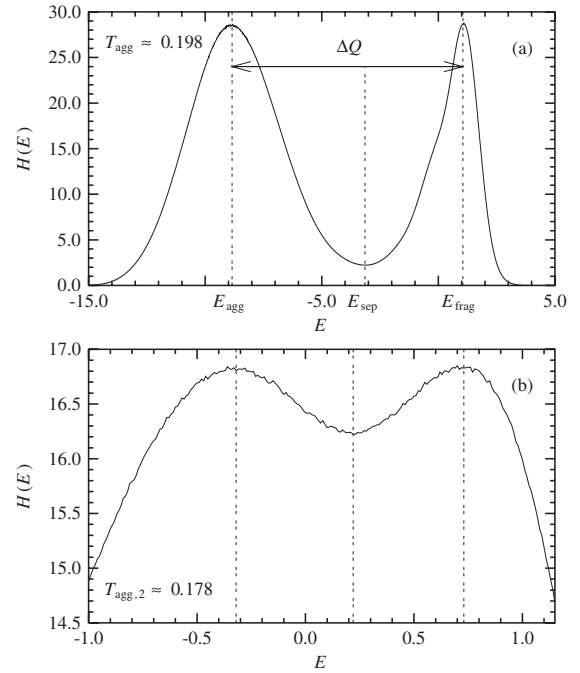

FIG. 2. Bimodal canonical energy distribution close to the (a) aggregation temperature T_{agg} and (b) the subphase transition near $T_{agg,2}$. Vertical dashed lines mark extremal points. The corresponding energies in (a) are identical with those indicating the phase boundaries in Fig. 1 and, in particular, to the crossing points of the reciprocal caloric curve $T^{-1}(E)$ with the Maxwell line in Fig. 1(b).

with the energy $E_{sep} \approx -3.15$, where the Maxwell line crosses the $T^{-1}(E)$ curve in the backbending regime in Fig. 1(b). These identifications are easily proven by setting the logarithmic derivative of $H(E)$ at T_{agg} to zero, which yields $\partial S(E)/\partial E = T_{agg}^{-1}$. The left-hand side is the reciprocal microcanonical temperature and thus $T^{-1}(E) = T_{agg}^{-1}$. As is seen from Fig. 1(b), this equation has *three* solutions, at E_{agg}, E_{sep}, and E_{frag}. Therefore, $H(E)$ possesses three extremal points at exactly these energies. Another expected result of the correspondence between the canonical and microcanonical representations is that the interfacial surface entropy can be written as: $\Delta S_{surf} = k_B \ln(H(E_{agg})/H(E_{sep})) = k_B \ln(H(E_{frag})/H(E_{sep}))$ [7]. These expressions serve as convenient estimators of the surface tension, which can be defined as $\sigma = T_{agg}\Delta S_{surf}/R_{agg}^2$, where R_{agg}^2 is the square radius of gyration of the aggregate.

A short remark shall also be devoted to a second, much weaker transition that accompanies the aggregation transition. It is also of "backbending" type and can be observed in the fragmentation region in Figs. 1(b) and 1(c) close to $E \approx -0.32$. The associated transition temperature is $T_{agg,2} \approx 0.178$ and is, therefore, *smaller* than T_{agg} but happens in the energetic region, where the population of *fragmented* macrostates dominates. In fact, this effect is difficult to understand and requires a system parameter that

allows the structural discrimination between macrostates. A detailed microcanonical analysis of the square relative distance between the centers of masses of the polymers reveals [16] that for energies close to E_{frag} the system is in a fragmented state, and the population of aggregated polymers in this energy region is extremely small. The situation is different for energies $E < 0.22$, where weakly stable aggregated conformations and polymer fragments coexist. Only for much smaller energies ($E < E_{agg}$), compact aggregates dominate. Having this in mind, the transition can also be understood from the canonical view. For temperatures below $T_{agg,2} \approx 0.178$, stable aggregates (solids) of low energies ($E < E_{agg}$) dominate. Approaching $T_{agg,2}$, the system enters the subphase of coexisting unstable premolten aggregates of comparatively high energies ($E \approx -0.32$) and already fragmented peptides. From Fig. 1(b), we see that this process is also accompanied by *cooling* due to monomer arrangements reducing surface tension. These monomer translocations are, however, energetically unfavorable. Eventually, for temperatures larger than T_{agg}, conformations of weakly coupled separate fragments (liquid) dominate. The intermediary subphase is never dominating and is, therefore, unstable. After these remarks, this transition is already visible in the cusplike behavior of the left, inner wing of the right fragmentation peak in Fig. 2(a). Reweighting to the subphase transition temperature $T_{agg,2}$, the bimodal structure of the canonical energy distribution in this energy range is clearly revealed in Fig. 2(b). Compared with the distribution at the aggregation transition in Fig. 2(a), the ratio between maximum and minimum is small and, therefore, the surface tension is also. In consequence, the transition between the solid and the premolten, unstable aggregates is, compared with the aggregation transition, negligibly weak. Note that the aggregation peak, not shown in Fig. 2(b), is much more pronounced than the peaks of the premolten aggregates at $E \approx -0.32$ and fragments at $E \approx 0.73$.

In this Letter, we have shown by employing a mesoscopic hydrophobic-polar heteropolymer aggregation model that the aggregation transition is a phase separation process, where the loss of entropy due to the existence of the phase boundary results in negative specific heat. This is an effect which is guided by changes of the interfacial entropy as a result of surface effects. Therefore, this effect is expected to disappear in the thermodynamic limit of macroscopic systems. It should strongly be emphasized, however, that peptides and proteins, like the exemplified model heteropolymers used in our study, are *necessarily* systems of *finite* length and a thermodynamic limit cannot be defined. For this reason, standard canonical formalisms for the analysis of conformational pseudophase transitions with phase separation are not suitable for these systems, since the temperature is not a unique control parameter and the total system energy measured in units of energy scales of mesoscopic particles (e.g., aggregates or single polymers) is not an extensive, separable quantity. In such cases,

microcanonical thermodynamics with the energy itself as the external control parameter provides a more favorable basis for the study of first-order-like transitions. The interesting phenomenon of the negativity of the microcanonical specific heat in peptide aggregation should be motivation for an experimental verification which is still pending.

*Email address: Christoph.Junghans@itp.uni-leipzig.de
†Email address: Michael.Bachmann@itp.uni-leipzig.de
‡Electronic addresses: Wolfhard.Janke@itp.uni-leipzig.de
http://www.physik.uni-leipzig.de/CQT.html

[1] D. H. E. Gross, *Microcanonical Thermodynamics* (World Scientific, Singapore, 2001).

[2] D. H. E. Gross and J. F. Kenney, J. Chem. Phys. **122**, 224111 (2005).

[3] W. Thirring, Z. Phys. **235**, 339 (1970).

[4] M. Schmidt, R. Kusche, T. Hippler, J. Donges, W. Kronmüller, B. von Issendorff, and H. Haberland, Phys. Rev. Lett. **86**, 1191 (2001).

[5] M. Pichon, B. Tamain, R. Bougault, and O. Lopez, Nucl. Phys. **A749**, 93c (2005).

[6] O. Lopez, D. Lacroix, and E. Vient, Phys. Rev. Lett. **95**, 242701 (2005).

[7] W. Janke, Nucl. Phys. B, Proc. Suppl. **63A–C**, 631 (1998).

[8] H. Behringer and M. Pleimling, Phys. Rev. E **74**, 011108 (2006).

[9] 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).

[10] S. Hilbert and J. Dunkel, Phys. Rev. E **74**, 011120 (2006); J. Dunkel and S. Hilbert, cond-mat/0511501.

[11] G. Favrin, A. Irbäck, and S. Mohanty, Biophys. J. **87**, 3657 (2004).

[12] K. A. Dill, Biochemistry **24**, 1501 (1985); K. F. Lau and K. A. Dill, Macromolecules **22**, 3986 (1989).

[13] 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).

[14] S. Schnabel, M. Bachmann, and W. Janke (unpublished).

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

[16] C. Junghans, M. Bachmann, and W. Janke (unpublished).

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

[18] The Hertz definition of the entropy is $\mathcal{S}(E) = k_B \ln\Gamma(E)$, where $\Gamma(E) = \int_{E'<E} dE' g(E')$ is the phase-space volume. In the thermodynamic limit $\mathcal{S}(E) = S(E)$, but for finite systems, as in our case, these two entropies are not necessarily identical [10]. We have, however, checked our results using both entropy definitions and did not find noticeable deviations in the transition regimes. In particular, the transition temperatures are virtually identical within the numerical precision of our data.

[19] D. H. E. Gross, cond-mat/0509202.