J|A|C|S COMMUNICATIONS

Published on Web 09/07/2010

Interplay between Secondary and Tertiary Structure Formation in Protein Folding Cooperativity

Tristan Bereau,[†] Michael Bachmann,[‡] and Markus Deserno*,[†]

Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, and Institut für Festkörperforschung, Theorie II, Forschungszentrum Jülich, 52425 Jülich, Germany

Received June 14, 2010; E-mail: deserno@andrew.cmu.edu

Abstract: Protein folding cooperativity is defined by the nature of the finite-size thermodynamic transition exhibited upon folding: two-state transitions show a free-energy barrier between the folded and unfolded ensembles, while downhill folding is barrierless. A microcanonical analysis, where the energy is the natural variable, has proved to be better suited than its canonical counterpart to unambiguously characterize the nature of the transition. Replica-exchange molecular dynamics simulations of a high-resolution coarse-grained model allow for the accurate evaluation of the density of states in order to extract precise thermodynamic information and measure its impact on structural features. The method has been applied to three helical peptides: a short helix shows sharp features of a two-state folder, while a longer helix and a three-helix bundle exhibit downhill and two-state transitions, respectively. Extending the results of lattice simulations and theoretical models, we have found that it is the interplay between secondary structure and the loss of non-native tertiary contacts that determines the nature of the transition.

The folding cooperativity of proteins is characterized by the relative population of intermediate states at the transition temperature: while two-state transitions exhibit two energetic peaks characterizing the folded and unfolded ensembles, downhill folders show a unimodal distribution of energetic states without any barrier.[1,2] This aspect of finite-size thermodynamic transitions can provide insight into the folding mechanism, but energetic populations can be difficult to measure. Therefore, protein folding cooperativity is often probed using the calorimetric criterion,[3] which quantifies the sharpness of the specific heat curve. From a computer simulation point of view, however, evaluating the probability density $p(E)$ remains an appealing idea, as it would provide an unambiguous description of the thermodynamic transition. Although this is currently untractable atomistically because of sampling limitations, high-resolution coarse-grained models offer an alternative approach.[4] While cutting down significantly on computational time, they can retain much chemical detail, and some are even able to model the folding of simple peptides with no prior knowledge of the native state. In this communication, we study the link between thermodynamics and structure for helical peptides using such a coarse-grained model,[5] details of which can be found in the Supporting Information.

To characterize the thermodynamics of finite-size systems, it has been shown that a microcanonical analysis based on the entropy $S(E)$ is often more informative than a canonical analysis.[6,7] Microcanonically, $S(E) = k_B \ln \Omega(E)$, where $\Omega(E)$ is the density of states. One remarkable feature of such a description is its ability to unambiguously distinguish between discontinuous (i.e., two-state) and continuous (i.e., downhill) transitions. Indeed, two-state transitions exhibit a depletion of intermediate energetic states that leads to local convexity in the entropy. This can best be observed by defining the quantity $\Delta S(E) = \mathscr{H}(E) - S(E)$, where the first term is the (double-)tangent to $S(E)$ in the transition region.[8−10] The method relies on accurate measurements of the density of states, which were calculated here using the weighted histogram analysis method.[11] All of the order parameters were analyzed as a function of energy.

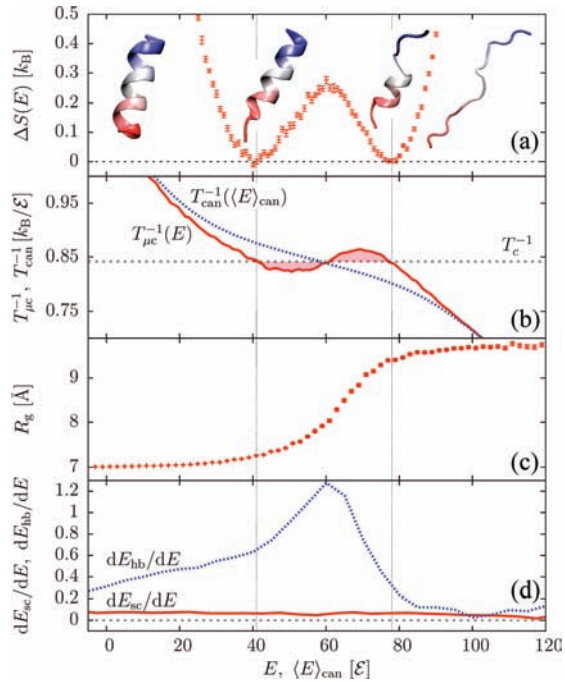

Figure 1. Results for $(AAQAA)_3$. (a) $\Delta S(E)$; error bars reflect the variance of the data points (1σ interval). (b) Inverse temperatures from canonical [$T_{can}^{-1}(\langle E\rangle_{can})$, blue] and microcanonical [$T_{\mu c}^{-1}(E) = \partial S/\partial E$, red] analyses, where $\langle E\rangle_{can}$ is the canonical average energy. (c) Radius of gyration $R_g(E)$ with the error of the mean. (d) Rates of H-bond and side-chain energies dE_{hb}/dE and dE_{sc}/dE. Vertical lines delimit the transition region, whose width corresponds to the microcanonical latent heat ΔQ.

We first examined a short α-helix having the sequence $(AAQAA)_3$.[12] The density of states reveals a discontinuous transition between the folded and unfolded ensembles with nonzero latent heat ΔQ (Figure 1a; representative conformations at different energies are shown). Phase coexistence is associated with a backbending[6] of the microcanonical inverse temperature $T_{\mu c}^{-1} = \partial S/\partial E$, while the corresponding canonical relation $T^{-1}(\langle E\rangle_{can})$, where $\langle E\rangle_{can}$ is the average energy, is monotonic (Figure 1b). The radius of

[†] Carnegie Mellon University.
[‡] Forschungszentrum Jülich.

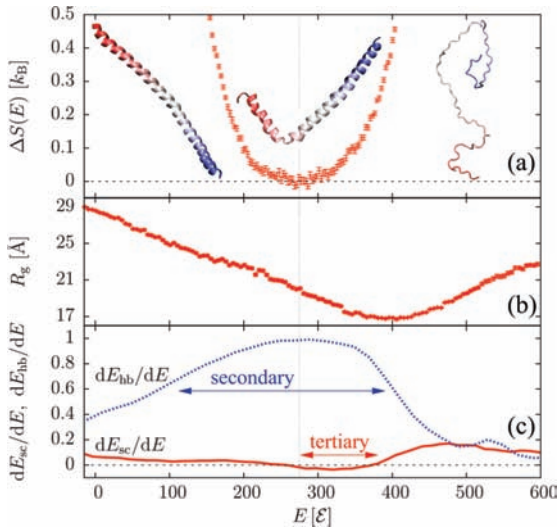

Figure 2. Results for $(AAQAA)_{15}$. (a) $\Delta S(E)$. (b) Radius of gyration $R_g(E)$. (c) Rates of H-bond and side-chain energies dE_{hb}/dE and dE_{sc}/dE. Horizontal arrows indicate where most of the secondary structure forms and where non-native tertiary contacts dissolve. The vertical line marks the transition point.

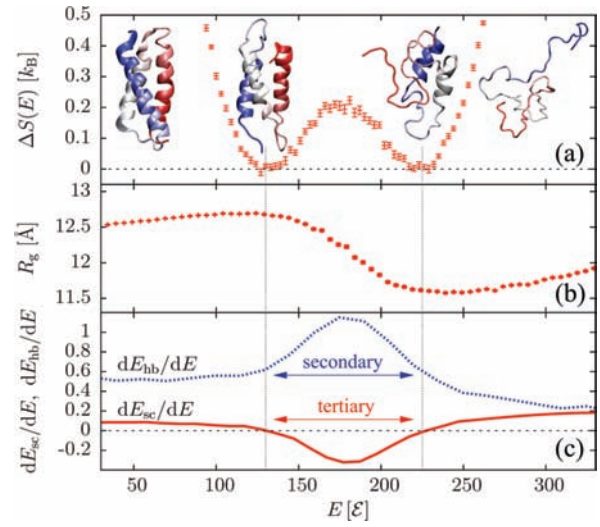

Figure 3. Results for the three-helix bundle α3D. (a) $\Delta S(E)$. (b) Radius of gyration $R_g(E)$. (c) Rates of H-bond and side-chain energies dE_{hb}/dE and dE_{sc}/dE.

gyration (R_g) quickly changes inside the coexistence region (Figure 1c), indicating that most of the structural rearrangements happen within this energy interval. We will assume the hydrogen-bond and side-chain energies E_{hb} and E_{sc} to be suitable proxies of secondary structure and tertiary contacts, repectively. It proves instructive to look at their energetic rates, dE_{hb}/dE and dE_{sc}/dE (Figure 1d): even though dE_{sc}/dE stays virtually flat over the energy range considered, the sharp peak in dE_{hb}/dE indicates that most of the secondary structure forms within the coexistence region.

Elongating the sequence to $(AAQAA)_{15}$ led to a qualitative change in the folding mechanism. The ground state again forms a single α-helix, but the transition is now continuous: as shown in Figure 2a and Figure S2 in the Supporting Information, there is a single transition point, and the latent heat is zero. The radius of gyration (Figure 2b) features a sharp minimum *above* the transition point, indicative of chain collapse into "maximally compact non-native states".[13] Upon a further decrease in the energy, the chain reorganizes from such non-native states into the helical state. In doing so, the rate of tertiary contact formation dE_{sc}/dE dips below zero (Figure 2c), so there is an energetic penalty associated with tertiary rearrangements. Hydrogen-bond formation occurs over a large energetic interval, as indicated by the broad maximum in dE_{hb}/dE. The absence of any two-state signal is consistent with theoretical models of the helix−coil transition:[14] the energetic cost of breaking a hydrogen bond is outweighed by the conformational entropy gained. Further analysis indicates two helices on average at the transition point.

While of similar length, the 73 amino acid de novo three-helix bundle α3D (PDB entry 2A3D)[15] does show a discontinuous transition (Figure 3a). Representative conformations sampled in the two coexisting ensembles stand as good proxies of the ground and unfolded states, unlike the case of the downhill-folding transition of $(AAQAA)_{15}$. The radius of gyration again shows a minimum above the transition (Figure 3b), and folding once more starts from maximally compact non-native states. Notably, secondary structure formation and the loss of non-native tertiary contacts (Figure 3c) are sharp and predominantly localized within the coexistence region. The three helices form inside the same energetic interval because of the interhelical cooperativity imprinted in the sequence.[16] Chain compaction is due to strong side chain−side chain interactions.

Overall, we can correlate thermodynamic features with structural information from the three peptides studied here. While $(AAQAA)_3$ is too short for tertiary interactions to play any role, the transitions associated with $(AAQAA)_{15}$ and the bundle α3D are both associated with tertiary rearrangements. These two examples suggest that independent of its nature, the folding transition is driven by the loss of non-native tertiary contacts (i.e., the region where $dE_{sc}/dE < 0$), which is reminiscent of the heteropolymer collapse model.[13] On the other hand, secondary structure formation shows very different signals: $(AAQAA)_3$ and α3D exhibit sharp peaks, whereas $(AAQAA)_{15}$ displays a broad maximum. As shown in Figure 2c, secondary structure formation in a downhill-folding peptide occurs over a much broader interval than for the loss of non-native tertiary contacts, whereas these two quantities are contained within the same narrow interval for a two-state peptide (Figure 3c). Cooperative secondary and tertiary structure formation has been proposed as a mechanism for two-state folding on the basis of lattice simulations[17] and theoretical models.[16] Beyond this, our results also highlight the interplay between secondary structure *formation* and the *loss* of non-native tertiary contacts. Our conclusions about the thermodynamics of the short, long, and bundled helixes are compatible with the calorimetric criterion; we find $\delta = 0.78,\ 0.52,$ and 0.78, respectively, for the calorimetric ratio.[3] It should be noted that Ghosh and Dill[16] predicted $\delta = 0.72$ for the similar bundle α3C. However, the main strength of a microcanonical analysis stems from its access to fine aspects of thermodynamic information that are otherwise difficult to obtain either canonically or from experiments. It thus stands as a complementary tool for gaining further insight.

Acknowledgment. We acknowledge stimulating discussions with K. Binder, W. Paul, R. H. Swendsen, and M. Taylor and funding through NIH Grant P01AG032131. M.B. thanks the Forschungszentrum Jülich for supercomputer time grants jiff39 and jiff43. T.B. acknowledges support from the Astrid and Bruce McWilliams Fellowship.

Supporting Information Available: Peptide model and simulation and analysis methods. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Privalov, P. *Adv. Protein Chem.* **1979**, *33*, 167–241.
(2) Privalov, P. *Adv. Protein Chem.* **1982**, *35*, 1–104.
(3) Kaya, H.; Chan, H. *Proteins* **2000**, *40*, 637–661.
(4) Voth, G. *Coarse-Graining of Condensed Phase and Biomolecular Systems*; Taylor & Francis: New York, 2008.
(5) Bereau, T.; Deserno, M. *J. Chem. Phys.* **2009**, *130*, 235106.
(6) Gross, D. H. E. *Microcanonical Thermodynamics: Phase Transitions in "Small" Systems*; World Scientific: Singapore, 2001.
(7) Hüller, A. *Z. Phys. B* **1994**, *93*, 401–405.
(8) Deserno, M. *Phys. Rev. E* **1997**, *56*, 5204–5210.
(9) Junghans, C.; Bachmann, M.; Janke, W. *Phys. Rev. Lett.* **2006**, *97*, 218103.
(10) Junghans, C.; Bachmann, M.; Janke, W. *J. Chem. Phys.* **2008**, *128*, 085103.
(11) Kumar, S.; Rosenberg, J.; Bouzida, D.; Swendsen, R.; Kollman, P. *J. Comput. Chem.* **1995**, *16*, 1339–1350.
(12) Scholtz, J.; York, E.; Stewart, J.; Baldwin, R. *J. Am. Chem. Soc.* **1991**, *113*, 5102–5104.
(13) Dill, K.; Stigter, D. *Adv. Protein Chem.* **1995**, *46*, 59–104.
(14) Zimm, B.; Bragg, J. *J. Chem. Phys.* **1959**, *31*, 526–535.
(15) Walsh, S.; Cheng, H.; Bryson, J.; Roder, H.; Degrado, W. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 5486–5491.
(16) Ghosh, K.; Dill, K. *J. Am. Chem. Soc.* **2009**, *131*, 2306–2312.
(17) Kaya, H.; Chan, H. *Phys. Rev. Lett.* **2000**, *85*, 4823–4826.

JA105206W