Eur. Phys. J. B **10**, 193–196 (1999)

THE EUROPEAN PHYSICAL JOURNAL B

EDP Sciences

Hot and cold denaturation of proteins: Critical aspects

A. Hansen[a], M.H. Jensen, K. Sneppen, and G. Zocchi

Niels Bohr Institute and NORDITA, Blegdamsvej 17, 2100 Copenhagen Ø, Denmark

Received 9 November 1998

Abstract. We argue that the first order hot and cold folding transitions of proteins observed at physiological chemical conditions ends in a critical point at a given temperature and chemical potential of the surrounding water. We investigate the properties of this critical point using a single-pathway scenario for the folding process. This pathway assumption determines the form of a Hamiltonian whose critical properties define a new universality class.

PACS. 05.70.Jk Critical point phenomena – 82.20.Db Statistical theories (including transition state) – 87.15.By Structure, bonding, conformation, configuration, and isomerism of biomolecules

Biologically relevant proteins are macromolecules [1] whose structures are determined by the evolutionary process [2,3]. There have been many attempts to grasp aspects of the protein folding process [4], in enumeration of configurations [5], in description of folding pathways [6] and in discussing the influence of water on protein structure [7]. In the present paper we discuss further consequences of protein folding along predesigned pathways. Such pathways may for example have evolved by a protein evolution where subunits subsequently are added to an already folding protein.

Several models have been proposed which address different aspects of the folding transition. A simple but appealing one is the "zipper model" [8], which was introduced to describe the helix–coil transition. In this model, the relevant degrees of freedom (conformational angles) are modeled through binary variables. Each variable is either matching the ordered structure (helix), or in a "coiled" state. A related parametrization for the 3d folding transition has been proposed by Zwanzig [9], describing it in terms of variables ψ_i, each of which is "true" (1) when there is local match with the correct ground state, or "false" (0) if there is no match. The term "local" is here defined through the parametrization index i. A zipper scenario that deals with the initial pathway of protein folding has been proposed by Dill *et al.* [6]. We can parametrize this model in the same way as done by Zwanzig by assigning the value one to each of the binary variables ψ_i describing closed contacts in the zipper. Built into the model is that opening and closing of contacts occur in a particular order: they behave as the individual locks in a zipper. This ordering is characterized through imposing the constraints

$$\psi_i \geq \psi_{i+1}. \tag{1}$$

The variables ψ_i alone cannot describe the degrees of freedom that become liberated when a portion of the zipper is open. The open part of the zipper may move freely ($\psi_i = 0$) whereas they cannot move in the part of the zipper where the contacts are closed ($\psi_i = 1$). In order to take into account this effect, we introduce a second, independent set of variables ξ_i. For simplicity, we also make these variables binary, taking the values 1 or $-B$. We are now in the position to propose a Hamiltonian for this zipper model,

$$H = -\sum_{i=1}^{N} \psi_i \xi_i, \tag{2}$$

subjected to the constraints (1).

We note that for any finite value of B, parts of the protein may unfold inside the already folded region *i.e.* in the parts of the zipper where $\psi_i = 1$. In order to prevent this, we assume B to be sufficiently large compared to any other energy scale in the system — in particular kT, where T is the temperature — so that the ξ_i variables never assume the value $-B$ as long as $\psi_i = 1$.

We will in the following use this Hamiltonian as a starting point for analyzing the hot and cold denaturation transitions of proteins when dissolved in water [10,11]. It is awkward to work with the Hamiltonian (2) directly because of the constraints (1). We therefore make a transformation to a different set of variables where the constraints (1) are implicitly taken into account and we obtain then the Hamiltonian we have already introduced previously [10,11]. We define a set of binary, unconstrained variables φ_i, by the following relation:

$$\psi_i = \varphi_1 \cdots \varphi_i. \tag{3}$$

[a] *Permanent address*: Department of Physics, Norwegian University of Science and Technology, NTNU, 7491 Trondheim, Norway
e-mail: `Alex.Hansen@phys.ntnu.no`

In particular, $\psi_1 = \varphi_1$. In the limit when $B \to \infty$, the Hamiltonian (2) becomes

$$H = -\varphi_1 - \varphi_1\varphi_2 - \varphi_1\varphi_2\varphi_3 - \cdots - \varphi_1\varphi_2\cdots\varphi_N, \quad (4)$$

where there are no additional constraints [11]. The role of the variables ξ_i — which is to provide entropy to the unfolded part ($\psi_i = 0$) of the zipper — is now played by the degeneracy introduced into the Hamiltonian in the following way: when a particular $\varphi_j = 0$, the Hamiltonian (4) will be degenerate with respect to the variables φ_i where $i > j$.

The interactions between protein and water may be taken into account by adding to (4) a coupling parametrized through water variables $w_1, w_2, ..., w_N$ [11]. Returning for a moment to the original variables ψ_i, we propose an interaction $(1 - \psi_i\xi_i)w_i$. The rationale behind this form is that when a contact is open ($\psi_i = 0$), the part of the protein parametrized by i is exposed to water and interact, while if the contact is closed ($\psi_i = 1$), there is no access to the water and the interaction is zero. Returning to the new variables φ_i, the resulting Hamiltonian is [11]

$$\begin{aligned} H = &-\mathcal{E}_0(\varphi_1 + \varphi_1\varphi_2 + \varphi_1\varphi_2\varphi_3 + \cdots + \varphi_1\varphi_2\cdots\varphi_N) \\ &+ [(1-\varphi_1)w_1 + (1-\varphi_1\varphi_2)w_2 + \cdots \\ &+ (1-\varphi_1\varphi_2\cdots\varphi_N)w_N], \end{aligned} \quad (5)$$

where we have introduced a scale parameter $\mathcal{E}_0$ in order to vary the relative strength of the protein self interactions and the protein-water interactions. In order to model hydrophobicity, we assume the w_i variables take values $\mathcal{E}_{\min} + s\Delta$, $s = 0, 1, ..., g-1$. Here, Δ is the spacing of the energy levels of the water-protein interactions. The equidistant energy levels reflect the experimentally observed approximate constant heat capacity at intermediate temperatures, whereas the finite number of levels g takes into account that protein-water interactions vanish at high temperatures, in practice above 120 °C.

The number of terms in the Hamiltonian (5), N, is the number of contact in the zipper model. This number may be equal to the number of amino acids, but is *a priori* unknown. It is important to realize that if one parametrize the folding with fewer steps N, each unit will be larger and energies and entropies appropriately increased (inversely proportional to N).

The partition function is

$$Z = \left(\mathrm{e}^{\mathcal{E}_0/T}\right)^N \left(\frac{1}{2}\frac{r^{-N}-1}{1-r} + 1\right). \quad (6)$$

The variable r is the ratio of statistical weights of unfolded to folded state, per variable:

$$r = \frac{g}{2}\mathrm{e}^{-\mu/T}\frac{1-\mathrm{e}^{-\Delta/T}}{1-\mathrm{e}^{-g\Delta/T}} \quad (7)$$

with $\mu = -\mathcal{E}_0 - \mathcal{E}_{\min}$ being the chemical potential of the surrounding water.

The physical meaning of this model is that the water molecules in contact with an unfolded portion of the protein has lower entropy than when not in contact (thus in the our model, hydrophobicity is caused by ordering of water and not by repulsive potentials, as is usually believed [13]). In the model one finds that a first order transition takes place when the parameter r switches between $r < 1$ and $r > 1$. Plotting r against T one obtains a non-monotonic function which for small μ values passes $r = 1$ twice, corresponding to unfolding at both low and high temperature, as indeed seen in experiments [14,15]. The mechanism for the transitions is the following. At high temperature the entropy gain of the protein chain causes the unfolding. As temperature is lowered the system gains more entropy by shielding the hydrophobic residues from the water. This leads to folding. As the temperature is lowered even further the cold unfolding transition occurs. Below this transition entropy is insignificant and the dominating effect is the attractive coupling between the water and the unfolded protein.

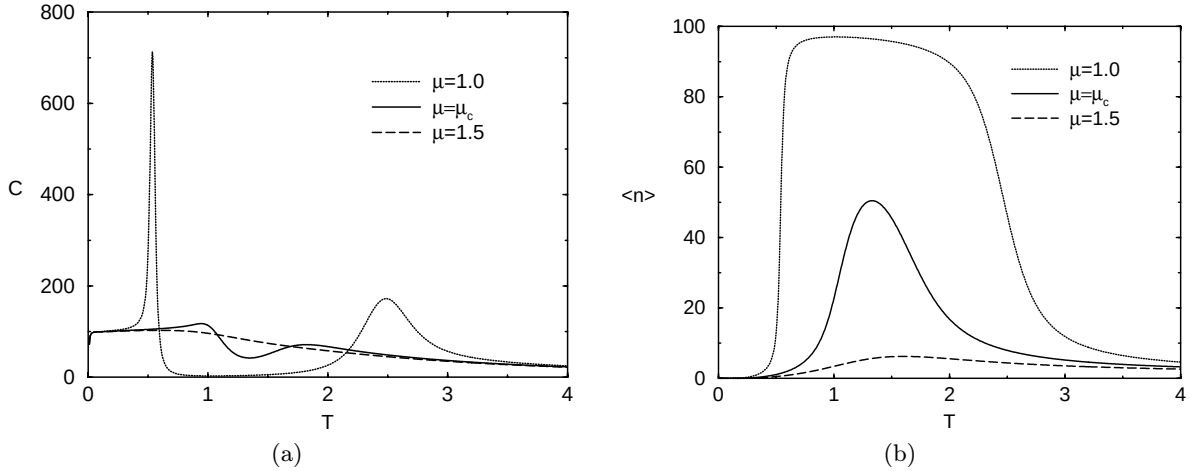
For an intermediate value of the chemical potential, r just touches the line $r = 1$, that is $\mathrm{d}r/\mathrm{d}T = 0$ when $r = 1$, corresponding to a merging of two first order transitions. This defines a critical point. Around this point, r varies quadratically in $T - T_\mathrm{c}$ and linearly in $\mu - \mu_\mathrm{c}$, as seen from expanding equation (7). In experiments of protein folding this point is accessible by changing the pH value of the solution. In fact, Privalov's data on low pH values indeed indicate that such a critical point exists. The scaling properties around this point thus opens for a possibility to gain insight into the nature of the folding process, in particular whether the pathway scheme we suggest can be falsified.

In Figure 1a we show heat capacity as a function of temperature for chemical potential below, at and above the critical value $\mu = \mu_\mathrm{c}$. For the chosen values of $\mathcal{E}_0 = 1$ and level density $\Delta = 0.02$ and $g = 350$ the critical point is situated at $T_\mathrm{c} = 1.33303\ldots$, $\mu_\mathrm{c} = 1.2838\ldots$. That is, it is situated at a *minimum* of the heat capacity curve. This is at first sight surprising, usually heat capacity has a pronounced increase at the critical point. The minimum reflects a partial ordering, as envisioned in Figure 1b where we show the degree of folding, counted by the average number of folded variables $\varphi_i = 1$, $i = 1, ..., n$ from $i = 1$ until the first variable $i = n+1$ which takes value $\varphi_{n+1} = 0$. The average value of this $\langle n\rangle$ is $N/2$ at the critical point, reflecting that the system is on average half ordered at this point. Correspondingly the heat capacity dips to a value in between the value of an unfolded and a completely folded state.

To characterize the functional form of the dip in the heat capacity, we investigate analytically $C_\mathrm{sing}(T) = C(T,\mu) - C(T,\mu_\mathrm{c})$ with $\mu \gg \mu_\mathrm{c}$ for different values of the size N. For finite N we may express the singular part of the heat capacity in the form:

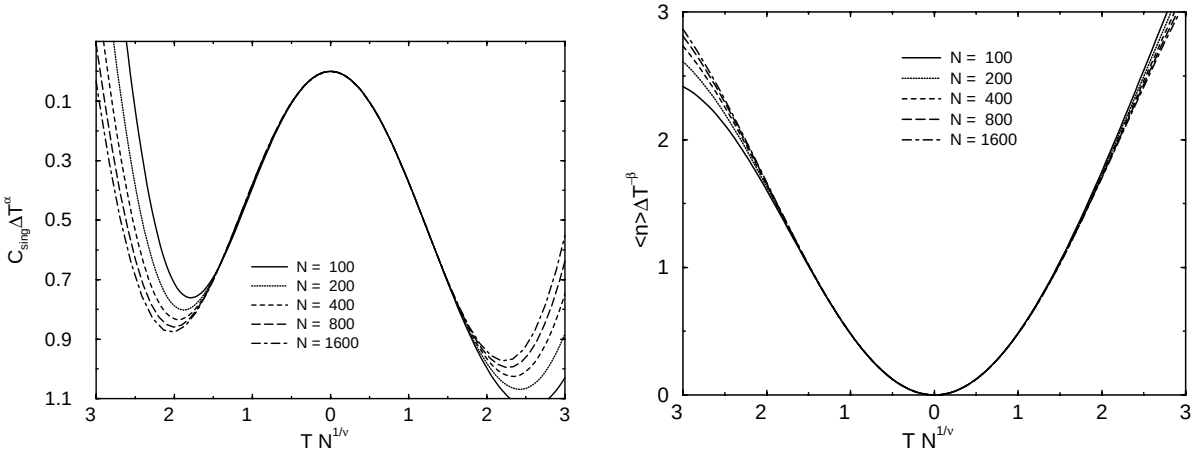
$$C_\mathrm{sing} = |T_\mathrm{c} - T|^{-\alpha} g\left((T_\mathrm{c} - T)N^{1/\nu}\right) \quad (8)$$

where $g(x) \to$ const when $x \to \infty$ and $g(x) \propto x^\alpha$ when $x \to 0$. We find analytically $\alpha = \nu = 2$ from differentiating the partition function (6). Figure 2a demonstrate this finite size scaling. Similarly we in Figure 2b show the behavior of the order parameter $\langle n\rangle$ as function of $T - T_\mathrm{c}$



(a) (b)

Fig. 1. (a) Heat capacity, C, as a function of T. (b) Degree of folding, $\langle n \rangle$, as a function of T. Here $g = 350$, $\Delta = 0.02$ and $N = 100$. The value $N = 100$ has been chosen as to be close to realistic values for this parameter.



(a) (b)

Fig. 2. (a) Finite size scaling of the heat capacity for $\mu = \mu_c$, $g = 350$ and $\Delta = 0.02$. Here $\alpha = 2$ and $\nu = 2$. (b) Finite size scaling of degree of folding, $\langle n \rangle$. Here $\beta = -2$.

and N:

$$\langle n \rangle = |T - T_c|^{\beta} f\left((T - T_c)N^{1/\nu}\right) \tag{9}$$

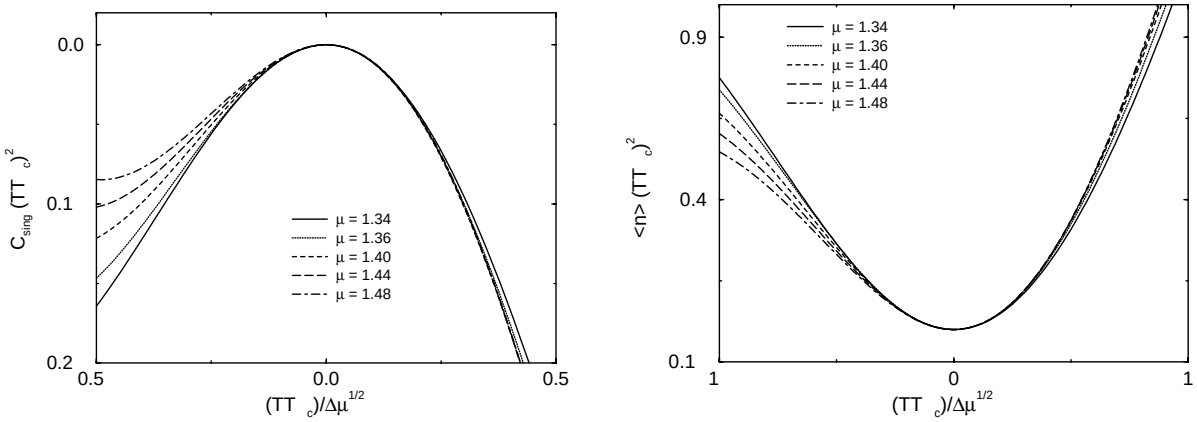
with $f(x) \to$ const when $x \to \infty$ and $f(x) \propto x^{-\beta}$ when $x \to 0$ where exponents $\beta = -2$, also found analytically. At a first glance, it might seem surprising that β is negative. This however reflects in part the unusual use of an extensive (in N) order parameter, and in part, that for $\mu = \mu_c$, the order parameter only becomes non-zero at $T = T_c$ when $N \to \infty$.

Likewise, we find that the susceptibility $\chi = \mathrm{d}\langle n \rangle / \mathrm{d}\mu$ scales as $|T - T_c|^{-\gamma}$ where $\gamma = 4$ and that $\langle n \rangle \propto (\mu - \mu_c)^{1/\delta}$ for $\mu > \mu_c$ where $\delta = -1$. Thus the usual exponent relations, $\alpha + 2\beta + \gamma = 2$, $\alpha + \beta(\delta + 1) = 2$, and $\gamma(\delta + 1) = (2 - \alpha)(\delta - 1)$ are fulfilled [16]. (Note, that the hyperscaling relation $d\nu = 2 - \alpha$, where d is the dimensionality of the system, is not fulfilled. This relation has however no meaning, as there are no spatial degrees of freedom.)

In terms of experiments on proteins, the relevant scaling behaviour is the how the degree of folding (order parameter) and the heat capacity behaves as function of temperature, when one changes the chemical potential away from its critical value. The qualitative prediction is that the width of the singular part of the heat capacity has a minimum at the critical value $\mu = \mu_c$. The broadening of the heat capacity is

$$C_{\mathrm{sing}}(T - T_c)^2 = h\left(\frac{T - T_c}{\Delta\mu^{1/2}}\right) \quad \text{for} \quad \mu > \mu_c \tag{10}$$

where $h(x) \propto x^{-2}$ for $x \to \infty$ and $h(x) =$ const for $x \to 0$ and where $\Delta\mu = \max(\mu - \mu_c, \Delta\mu_{\min})$ with $\Delta\mu_{\min} \propto 1/N$ takes into account the finite size sensitivity of the scaling. We show in Figure 3a, an example of such a data collapse. These predictions are experimentally accessible through the use of standard calorimetric techniques, where one should seek to obtain a data collapse above the critical point, *i.e.* the point of minimal width. The heat



(a) (b)

Fig. 3. (a) $C_{\rm sing}(T-T_{\rm c})^2$ *vs.* $(T-T_{\rm c})/\Delta\mu^{1/2}$. (b) $\langle n\rangle(T-T_{\rm c})^2$ *vs.* $(T-T_{\rm c})/\Delta\mu^{1/2}$. We have chosen $N=100$, $g=350$ and $\Delta=0.02$. Note the good quality of the data collapse in spite of smallness of the system.

capacity below the critical μ is complicated by the merging of two first order transitions. However, the distance between these moves away from each other in T as $\Delta\mu^{1/2}$.

Likewise, we expect the degree of folding $\langle n\rangle$ to show data collapse of the form

$$\langle n\rangle(T-T_{\rm c})^2\mu > \mu_{\rm c} \tag{11}$$

where $k(x)$ behaves asymptotically as h. We show this in Figure 3b. This quantity can be observed experimentally through fluorescence measurements.

In summary, we have proposed that a folding pathways implemented for proteins in water implies a critical point with a diminishing heat capacity at criticality. We have determined all critical exponents, and proposed two experiments that could confirm or falsify the concept of a sequentially ordered folding.

References

1. B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, J.D. Watson, *Molecular biology of the cell* (Garland Publ., New York, 1994).
2. H. Neurath, Science **224**, 350 (1984).
3. W. Gilbert, Science **228**, 823 (1985).
4. K.A. Dill, S. Bromberg, K. Yue, K.M. Fiebig, D.P. Yee, P.D. Thomas, H.S. Chan, Protein Sci. **4**, 561 (1995).
5. P.A. Lindgård, H. Bohr, Phys. Rev. E **56**, 4497 (1997).
6. K.A. Dill, K.M. Fiebig, H.S. Chan, Proc. Natl. Acad. Sci. **90**, 1942 (1993).
7. H. Li, C. Tang, N.S. Wingreen, Proc. Natl. Acad. Sci. **95**, 4987 (1998).
8. J.A. Schellman, J. Phys. Chem. **62**, 1485 (1958); B.H. Zimm, J.K. Bragg, J. Chem. Phys. **31**, 526 (1959); C.R. Cantor, P.R. Schimmel, *Biophysical Chemistry* (W.H. Freeman, San Fransisco, 1980), Chap. 20.
9. R. Zwanzig, Proc. Natl. Acad. Sci. **92**, 9801 (1995).
10. A. Hansen, M.H. Jensen, K. Sneppen, G. Zocchi, Physica A **250**, 355 (1998).
11. A. Hansen, M.H. Jensen, K. Sneppen, G. Zocchi, Eur. Phys. J. B **6**, 157 (1998).
12. B. Nolting, R. Golbik, J.L. Neira, A.S. Soler-Gonzales, G. Schreiber, A.R. Fersht, Proc. Natl. Acad. Sci. **94**, 826 (1997).
13. J.T. Edsall, J. Am. Soc. **57**, 1506 (1935).
14. P.L. Privalov, Biochem. and Molecul. Biol. **25**, 281 (1990).
15. G.I. Makhatadze, P.L. Privalov, Adv. Prot. Chem. **47**, 307 (1995).
16. H.E. Stanley, *Phase Transitions and Critical Phenomena* (Cambridge Univ. Press, Cambridge, 1971).