Protein dynamics implications of the low- and high-temperature denaturation of myoglobin

Ramzi R. Khuri, Trung V. Phan, and Robert H. Austin

Phys. Rev. E **104**, 034414 — Published 24 September 2021

DOI: 10.1103/PhysRevE.104.034414

The Dark Side of the Low and High Temperature Denaturation of Myoglobin

Ramzi R. Khuri*
Department of Natural Sciences, Baruch College, City University of New York, New York, NY 10010

Trung V. Phan† and Robert H. Austin‡
Department of Physics, Princeton University, Princeton, NJ, 08544

We re-investigate a simple model used in the literature concerning the thermodynamic analysis of protein cold denaturation. We derive an exact thermodynamic expression for cold denaturation and give a better approximation than exists in the literature for predicting cold denaturation temperatures in the two-state model. We discuss the "dark-side" implications of this work for previous temperature dependent protein dynamics experiments and discuss micro-fluidic experimental technologies which could explore the thermal stability range of proteins below the bulk freezing point of water.

I. INTRODUCTION

Water soluble proteins typically are unstable to unfolding at both the obvious upper temperature T_H (heat denaturation) but also, and paradoxically, at a low temperature T_C, a phenomena called cold denaturation [1].

Fig. 1 shows the basic phenomena, taken from [1]. The water soluble protein myoglobin (Mb) denatures as a function of pH at both high and low temperatures. Of particular note in Fig. 1 is not only the sharpness of the hot denaturation of Mb at pH4.0 but also the step up in the specific heat C_p of the hot denatured protein, indicating that temperature dependent degrees of freedom have been exposed by hot denaturation.

Typically, the remarkable low-temperature denaturation occurs below the freezing point of water, and so is hidden from view; destabilizing conditions such as extreme pH or hydrogen bonding reagents such as urea are required to bring it into view, in the case of Fig. 1, lowering of the pH. However, the fact that the low temperature denaturation process is often "invisible" does not mean it is irrelevant in obtaining a deeper understanding of protein dynamics and stability [2], as we hope to explain quantitatively here.

There is a deep literature on cold denaturation of proteins, although it is primarily based on thermodynamics. An excellent thermodynamics review can be found in Dias *et al.* [3]. In this paper we wish to stress the statistical mechanics of the limited thermal stability ranges of most proteins, except possibly intrinsically disordered proteins, which apparently have no basic order parameter, present a careful derivation of the phenomena, emphasize the important role that cold denaturation may play in understanding protein dynamics. and suggest experimental approaches to probe what would seem to be an inaccessible region below the bulk freezing point of water.

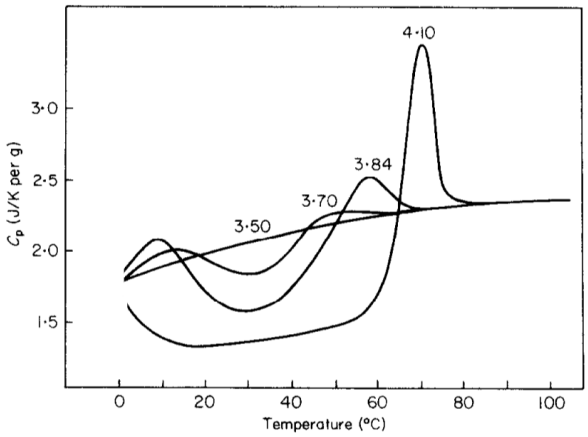

FIG. 1. Temperature dependence of metmyoglobin partial specific heat capacity in sodium acetate solutions of differing pH values. Taken from [1]

We re-investigate the purely mathematical thermodynamic calculation in which we demonstrate that, for each hot denaturation temperature, under a widely satisfied parametric condition, there exists a unique cold denaturation temperature, which can be calculated in terms of the hot denaturation point and the other parameters. We point out a confusion in the original analysis ([4]) that has propagated in the literature and may hinder theoretical understanding. The more careful analysis we perform gives an exact result and a better approximation for the cold-denaturation temperature given the hot-denaturation temperature.

While we mathematically fit the phenomena of cold denaturation using classical thermodynamics, this does not lead to understanding the physics of the simple experimental observation, evident in Fig. 1, that the slope of the specific heat of a heat-denatured unfolded protein (C_D) as a function of temperature ($r_D = dC_D/dT$) is greater than the slope of the specific heat of a folded protein as a function of temperature ($r_N = dC_N/dT$),

* ramzi.khuri@baruch.cuny.edu
† tvphan@princeton.edu
‡ austin@princeton.edu

which leads to the thermodynamic observations.

If $r_D > r_N$, within certain bounds which we discuss below, basic thermodynamics predicts that there must be a second denaturation (unfolding) temperature at low temperatures. The physics behind this involves the protein+water system above the high temperature denaturation transition and below the low temperature denaturation transition. Paradoxically, coming up from below the cold denaturation temperature by increasing the temperature results in a *decrease* in the entropy of the protein-water system, not an increase, and in an increase in the fraction of folded proteins.

The physical reason for cold denaturation must be intimately connected to the entropy of the solvent (water) in which the protein is dissolved, since presumably an isolated folded protein *in vacuo* is in a state of minimum free energy [5] and cannot unfold with decreasing temperature. If the protein + water system can access *more* entropy as it cools by unfolding, the cold denatured system wins thermodynamically at low temperatures, although the conformation of this cold denatured protein must be different than the heat denatured conformation. No doubt there is a connection between the low temperature and the high temperature denaturation phenomena, and that connection must be the physics of the water+protein system.

While we know the structures of a great many proteins in the *folded* state, we have very little knowledge of the structures of *unfolded* proteins as a function of temperature, because they typically, by definition, do NOT fold into the ordered crystals needed for X-ray diffraction. In fact, we know of no structures of a low temperature reversibly denatured protein.

This may seem to be a minor issue if one believes that only correctly folded proteins are functionally active, but the basic fact that low temperature denaturation even exists leads us to speculate that extrapolating *up* ($\uparrow$) from the low-temperature denatured configurations may reveal states that also play a role in protein activity.

We note that biological systems also take account of this two-faced nature of protein stability. We now know that for water soluble proteins at normal conditions, folding into active conformations after synthesis from the ribosome is a formidable problem; proteins probably need the assistance of chaperone proteins to assume the functional conformation with the further expenditure of free-energy via ATP hydrolysis [6].

Further, even when folded into a basin of biologically active conformations [7] (which may not be the conformation basin which is lowest in the free-energy landscape) a high temperature excursion results in incorrectly folded proteins which need to be re-folded by heat-shock chaperones. In a similar manner, a low temperature excursion also results in the generation of mis-folded proteins. There is a separate family of chaperons called cold-shock chaperones [8] whose task is to refold proteins which have been mis-folded by cold temperatures.

II. ICE ENTROPY

One possible hidden source of *increasing* entropy with *decreasing* temperature may be the interaction of the water solvent with the exposed hydrophobic residues upon cold denaturation. The water molecule in the *sp* hybridization approximation has 4 bonding orbitals forming a tetrahedral symmetry around the oxygen atom. In the case of ice, a dynamic kind of lattice is formed [9]. Any two of the four lobes contain positively charged hydrogen atoms, while the other two lobes contain negatively charged lobes of excess electron density. Because of the high charge separation of the water molecule, the static dielectric constant of water, κ, is extremely high: $\kappa = 88$ at 273 K [10].

Note that there is an intrinsic amount of disorder contained in such a lattice. Charge neutrality requires that any oxygen atom can have at most 2 hydrogen atoms near it. The lattice that forms has a residual amount of entropy due to the possible ways of arranging the hydrogen atoms and still obey the so-called "ice rule": *For the four nearest neighbor hydrogen atoms surrounding the oxygen atom, two are close to it and two are removed from it.*. A particularly lucid discussion of this problem can be found in Baxter, Exactly Solved Models of Statistical Mechanics [11].

Pauling did one of the first quantum calculations of the first residual quantum entropies of ice [12] based on this reasoning. Simply put, there are 6 possible tetrahedral arrangements of two hydrogen atoms around a given oxygen atom, and 1/2 probability for the next adjacent oxygen that one of the orbitals is occupied, yielding at 0K a quantum entropy for a mole of N water molecules:

$$S = k_B \ln W = k_B \ln(6 \times 1/4) \approx 0.4 k_B \mathrm{J/K} \ . \quad (1)$$

We will discuss below that the stability of a protein is quite low at room temperature, with the Gibbs free energy only about $+0.2 k_B T$ per residue. The Pauling enthalpy $H = 0.4 k_B T$ of ice is thus about the same order of magnitude as the Gibbs energy, and thus the entropy increase of water as it penetrates the protein can easily destabilize a protein at low temperatures.

III. PROTEIN-WATER HAMILTONIANS AND PHASE TRANSITIONS

Before we do a thermodynamic analysis, we would like to examine the physics of the water-protein system, since that is a deeper way to understand the system.

Simple water soluble proteins consist of a chain of amino acids. The amino acids interact with each other and with the solvent (water). Our mission here is not to solve the terrifically difficult protein folding problem, but rather to explore an extremely simple model which can elucidate the basic physics of the bistablity of protein structures.

An elegant discussion of simple protein-water Hamiltonians can be found in Sneppen and Zollchi [13], which we follow closely. At one extreme we can assume that there is only one state for a "correctly" folded protein, in which all the amino acids are correctly positioned next to the correct neighbor. This is called the golf-course Hamiltonian H_{gf}, since it has no folding pathway, just a single deep free energy hole. This Hamiltonian can be written in terms of amino acid local order-parameters $\phi_j \in \{0, 1\}$ (0 for wrong-ordering and 1 for correct-ordering). The golf-course Hamiltonian (no water yet!) is then given by:

$$H_{gf} = -NE_0\phi_1\phi_2\phi_3...\phi_N \tag{2}$$

The other possibility for a protein Hamiltonian, and the more biological one, is the "zipper" Hamiltonian, which assumes a folding pathway and thus a sequential series of increasingly correct amino acid neighbors. The zipper Hamiltonian is given by:

$$H_{zp} = -E_0(\phi_1 + \phi_1\phi_2 + ... + \phi_1...\phi_N) \tag{3}$$

Of course, this so far assumes a vacuum. As we discussed above, bulk water has an intrinsic low temperature entropy, and is highly polar. In this exceedingly simple model, we can assume that the highly polar water molecules interact with interior amino acids in the unfolded state. Further, we can assume that in the case of the entropy increase of water that a large number, g, of water molecules participate per exposed amino acid.

Note that we follow here the work of Sneppen and Zollchi because of its simplicity. Obviously, the model below cannot be used to predict *ab initio* the phase states of real proteins, a terrifically hard problem and an area of active interest [14]. However, one advantage of this approach is that it sees the forest if not the trees, and has the potential to explain the sharpness of the folding and unfolding phase transitions in terms of the variance in the states the protein is moving over, as we will discuss in Section V.

Thus, following Sneppen and Zollchi [13] we write the full Hamiltonian for a water-protein system, not exactly identical for golf-course and zipper models but both have the same form:

$$H = -E_o\phi + (1 - \phi)w \tag{4}$$

where ϕ is the protein variable (which could be either the golf-course Hamiltonian or the zipper Hamiltonian, and it includes the presumably constant normally exposed to water amino acids), w is the water-amino acid Hamiltonian for buried amino acids which are exposed when the protein unfolds.

In this model, if the protein is folded ($\phi = 1$) the internal water term w vanishes, while if the protein is unfolded ($\phi = 0$) hydrophobic groups are exposed. Presumably the protein structure swells due to the penetration of the water into the protein core and the water entropy is raised.

The water part of the Hamiltonian w, while simple, is quite subtle. It consists of several parts. An obvious term is the physical interaction of the hydrophobic amino acid with a water molecule, μ. There is an energy term δ which represents the interaction of the g water molecules perturbed by the hydrophobic amino acid, and there must be an entropic term $S_w = k_b \ln(g)$ which represents the entropy gain spread amongst g water molecules. Actually, g includes both the product of the $6 \times 1/4$ water orbitals times some (unknown) number of perturbed water molecules, which we lump into essentially a fitting parameter.

The partition function Z for *each* amino acid becomes:

$$\begin{aligned} Z &= ge^{E_o/k_BT} + e^{-\mu/k_BT}\sum_{j=0}^{g-1} e^{-j\delta/k_bT} \\ &= A + \left(1 - 2^{-1}\right)B \ , \end{aligned} \tag{5}$$

where we do the geometric sum:

$$\begin{aligned} A &= g\exp\left(\frac{E_0}{k_BT}\right) \ , \\ B &= 2\exp\left(-\frac{\mu}{k_BT}\right)\left(\frac{1 - \exp\left(-\frac{g\delta}{k_BT}\right)}{1 - \exp\left(-\frac{\delta}{k_BT}\right)}\right) \ , \end{aligned} \tag{6}$$

In the case of the protein having N amino acids, the partition function $Z = \sum_{\{\phi\}}\sum_{\{w\}}\exp(H/k_BT)$ for those Hamiltonians and the combinatorics can be calculated analytically:

$$\begin{aligned} Z_{gf} &= A^N + \left(1 - 2^{-N}\right)B^N \ , \\ Z_{zp} &= A^N + \frac{1}{2}\frac{B}{A - B}\left(A^N - B^N\right) \ . \end{aligned} \tag{7}$$

The specific heat capacity

$$C = \frac{1}{N}\frac{\partial}{\partial T}\left(k_BT^2\frac{\partial \ln Z}{\partial T}\right) \tag{8}$$

for these two models is surprisingly identical when the number of amino acids in the protein is infinite $N \to \infty$. However, it is clear that finite size effects are quite important in protein phase states and it is the simplicity of the model which somewhat obscures this [15]. The model also ignores pressure effects, these are known to be very important in cold denaturation [16], again we can only apologize for looking at the simplest possible model.

Both protein Hamiltonians exhibit two phase transitions corresponding to cold-denaturation and hot-denaturation. Fig. 2 shows an example of how the specific heat capacity depends on temperature for $N \to \infty$ and $N = 100$. The peaks at transition temperatures (sharper as N increases) indicate that there are finite specific calorimetric enthalpy changes ΔH associated with the cold and hot denaturation (equal to the areas under

the peaks and above the $N \to \infty$ line excluding singularity $C \to \infty$ at the transition temperatures). The assumption $\Delta C \approx \text{const} > 0$ is reasonable when the parameters satisfy $\delta \ll 1$ and $g\delta \gg 1$.

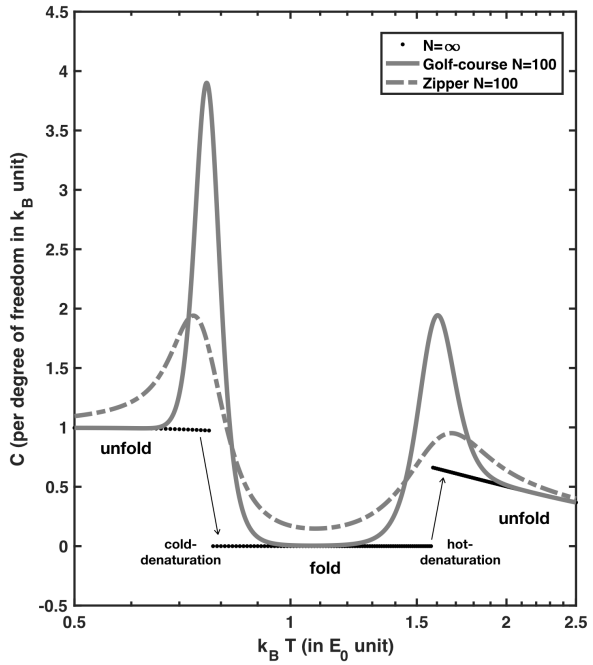

FIG. 2. The specific heat capacity for $N \to \infty$ and $N = 100$ as function of temperature for golf-course model and zipper model using parameters $E_0 = 1$, $\mu = -2$, $\delta = 0.1$ and $g = 60$ satisfy $\delta \ll 1$ and $g\delta \gg 1$).

IV. DETAILED THERMODYNAMIC ANALYSIS

We assume that the protein can exist in two discrete states, the native folded one, and a denatured (unfolded) one which can exists both at low and high temperatures. At thermal equilibrium temperature T, denote the changes in the Gibbs free energy $\Delta G(T)$, the enthalpy $\Delta H(T)$ and the entropy $\Delta S(T)$ between the two states as:

$$\begin{aligned} \Delta G(T) &= G^{(D)}(T) - G^{(N)}(T) \ , \\ \Delta H(T) &= H^{(D)}(T) - H^{(N)}(T) \ , \\ \Delta S(T) &= S^{(D)}(T) - S^{(N)}(T) \ . \end{aligned} \tag{9}$$

Let T_H be the hot-denaturation temperature. Then from the relations between thermodynamics variables:

$$\begin{aligned} \Delta S(T_H) &= \frac{\Delta H(T_H)}{T_H} \ , \\ \Delta G(T_H) &= \Delta H(T_H) - T_H \Delta S(T_H) \ . \end{aligned} \tag{10}$$

From the assumption that the difference in heat capacity is a constant:

$$\begin{aligned} \Delta C_P &= \left(\frac{\partial \Delta H(T)}{\partial T} \right)_P \\ \Rightarrow \ \Delta H(T) &= \Delta H(T_H) + \Delta C_P (T - T_H) \ . \end{aligned} \tag{11}$$

The quantities ΔC_P and $\Delta H(T_H)$ can be read-off directly from the $C_P(T)$ near hot-denaturation temperature T_H, as shown in Fig. 3.

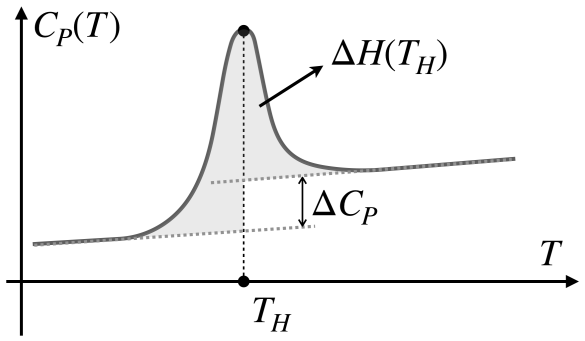

FIG. 3. How to obtain ΔC_P and $\Delta H(T_H)$ from the $C_P(T)$ plot near hot-denaturation temperature T_H.

From the thermodynamics relation:

$$\left(\frac{\partial S}{\partial T} \right)_P = \left(\frac{\partial S}{\partial H} \right)_P \left(\frac{\partial H}{\partial T} \right)_P \ , \ dS = \frac{\Delta C_P}{T} dT \ , \tag{12}$$

it follows that:

$$\begin{aligned} \Delta S(T) &= \Delta S(T_H) + \int_{T_H}^{T} \left(\frac{\partial \Delta S(T)}{\partial T} \right) dT \\ &= \frac{\Delta H(T_H)}{T_H} + \Delta C_P \int_{T_H}^{T} \frac{dT}{T} \\ &= \frac{\Delta H(T_H)}{T_H} + \Delta C_P \ln \left(\frac{T}{T_H} \right) \ . \end{aligned} \tag{13}$$

This, together with (11), gives:

$$\begin{aligned} \Delta G(T) &= \Delta H(T) - T \Delta S(T) \\ &= \Delta H(T_H) \left(\frac{T_H - T}{T_H} \right) + \Delta C_P (T - T_H) \\ &\quad - \Delta C_P T \ln \left(\frac{T}{T_H} \right) \ . \end{aligned} \tag{14}$$

Here is where some confusion in the literature has originated: the second term in the sum in the last line of (14) has the opposite sign to that shown in equation (8) of reference [4], and while some references do realize and fix it, they do not go further to re-examine the theoretical consequences [13, 17] in depth. This is what we will explore below.

We first calculate the first two derivatives of $\Delta G(T)$ with respect to T:

$$\begin{aligned}\frac{\partial \Delta G(T)}{\partial T} &= -\frac{\Delta H(T_H)}{T_H} - \Delta C_P \ln\left(\frac{T}{T_H}\right) \\ &= -\Delta S(T) \ ,\end{aligned} \tag{15}$$

and

$$\frac{\partial^2 \Delta G(T)}{\partial T^2} = -\frac{\Delta C_P}{T} \ . \tag{16}$$

The graph of $\Delta G(T)$ has a zero at $T = T_H$, which corresponds to the hot-denaturation temperature. Its first derivative vanishes only at temperature

$$T_S = T_H \exp\left(-\frac{\Delta H(T_H)}{\Delta C_P T_H}\right) \ . \tag{17}$$

For $\Delta H(T_H) > 0$ and $\Delta C_P > 0$, $T_S < T_H$. Since for these conditions $\partial^2 \Delta G(T)/\partial^2 T < 0$, the curve is concave downwards thus T_S is a maximum. It follows that $\Delta G(T) = 0$ at another point, $T_C < T_S < T_H$. This is the cold denaturation point, whose properties we investigate below. See Fig. 4 for the relation between ΔG and T.

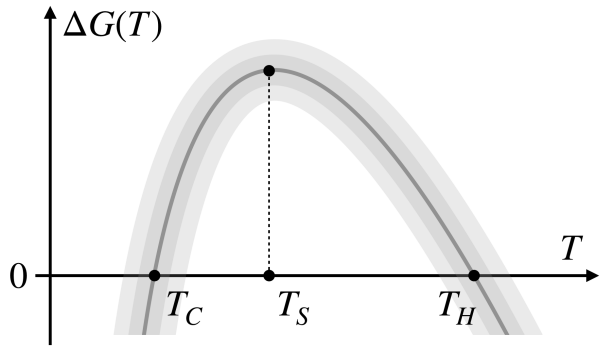

FIG. 4. The relation between ΔG and T from equation (14) with the relevant temperatures: hot-denaturation happens at T_H, cold-denaturation happens at T_C and the protein in its native state is most stable at T_S

We would like to solve the equation $\Delta G(T) = 0$ for a positive value of T_C. To simplify the analysis, let

$$x = \frac{T}{T_H} \ , \tag{18}$$

then from (14), the equation can be written as:

$$(1-x)\left(\frac{\Delta H(T_H)}{\Delta C_P T_H} - 1\right) - x \ln x = 0 \ . \tag{19}$$

We are looking for a solution to (19) for $x < 1$. Clearly, the second term in this equation is always positive, so there is no solution (and therefore no cold denaturation) unless:

$$\frac{\Delta H(T_H)}{\Delta C_P T_H} < 1 \ . \tag{20}$$

This is a specific criterion for protein unfolding at low-temperature depending directly on physical quantities measured at the hot-denaturation transition.

Cold-denaturation can still emerge at low-temperature even though the enthalpy change is positive $\Delta H(T_H) > 0$ is positive, as long as it's small enough $\Delta H(T_H) < \Delta C_P T_H$. We note that the mistake made in [4] leads to the wrong conclusion that for cold-denaturation to happen, the enthalpy change has to be strictly negative $\Delta H(T_H) < 0$.

To further simplify the notation, let

$$y = \frac{\Delta H(T_H)}{\Delta C_P T_H} \ . \tag{21}$$

Then equation (19) can now be written as

$$(1-x)(y-1) - x \ln x = 0 \ , \tag{22}$$

which is equivalent to

$$y = 1 + \frac{x \ln x}{1-x} \ . \tag{23}$$

We will show that, for any $y < 1$, there is a unique solution $x < 1$. It is useful to first calculate the first derivative:

$$\frac{dy}{dx} = \frac{\ln x + 1 - x}{(1-x)^2} \ . \tag{24}$$

It is straightforward to show that, as $x \to 0$, $y \to 1$ and $dy/dx \to -\infty$. For $x \to 1$ $y \to 0$ and (via twice applying l'Hopital's rule) $dy/dx \to -1/2$. It is also straightforward to show that $dy/dx < 0$ throughout the domain $0 < x < 1$ and that $d^2y/dx^2 > 0$. The function $y(x)$ is therefore a monotonically decreasing function with upward concavity. Ergo, $x(y)$ is also a monotonically decreasing function. For each $y < 1$, there is a unique value of $x < 1$, which is what we set out to show. So there is a unique value of $T_C < T_H$ at which denaturation occurs.

To plot $x(y)$, we simply take the mirror image of $y(x)$ given by the $y = x$ line in the same domain of $0 < x < 1$ and $0 < y < 1$. Of course, for $y \to 0$, $x \to 1$ and $dx/dy \to -2$ (the reciprocal of $-1/2$). As $y \to 1$, $x \to 0$ and $dx/dy \to 0$. The relation between x and y can be seen in Fig. 5.

So what happens to the entropy at the denaturation temperatures? Clearly, for $\Delta S(T_H) = \Delta H(T_H)/T_H > 0$. The entropy increases with increasing temperature (hot denaturation) as the protein unfolds. At T_C, however,

$$\Delta S(T_C) = \Delta C_P (y + \ln x) = \Delta C_P \left(\frac{1 - x + \ln x}{1-x}\right) \ . \tag{25}$$

The numerator $1-x+\ln x$ is also the numerator of dy/dx, which was already shown to be negative for $0 < x < 1$. For $x < 1$, therefore, $\Delta S(T_C)$, here defined as the change in entropy as the temperature is increased at the cold denaturation point is, also negative. This is because the

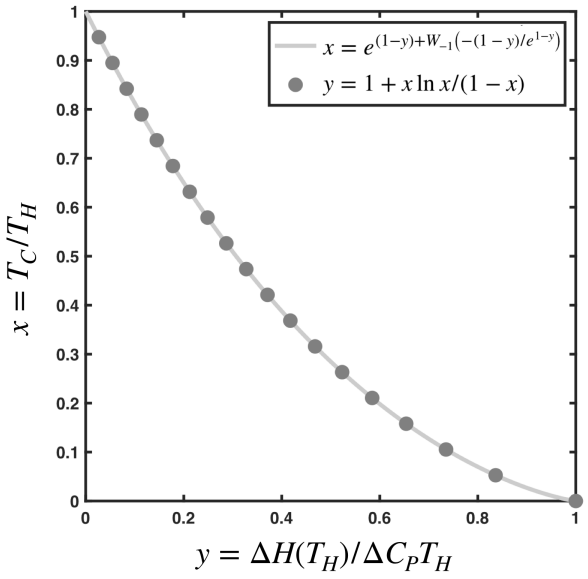

FIG. 5. The relation between x and y, from the equivalent analytical forms of $y(x)$ and $x(y)$, where W_{-1} is a Lambert W-function.

protein folds into the lower entropy state as it is heated past the cold denaturation point. Equivalently, the entropy increases as the temperature goes below the cold denaturation point. This is the paradoxical feature of cold denaturation.

In fact, an explicit solution for $x(y)$ can be written in terms of a Lambert W-function [18]:

$$x(y) = \exp\left((1-y) + W_{-1}\left(-\frac{(1-y)}{\exp(1-y)}\right)\right) . \quad (26)$$

In other words, the temperature T_C can be represented analytically in terms of T_H, ΔH and ΔC_P:

$$T_C = T_H \exp\left(\left(1 - \frac{\Delta H(T_H)}{\Delta C_P T_H}\right) + W_{-1}\left(-\frac{\left(1 - \frac{\Delta H(T_H)}{\Delta C_P T_H}\right)}{\exp\left(1 - \frac{\Delta H(T_H)}{\Delta C_P T_H}\right)}\right)\right) , \quad (27)$$

which can be approximated as a quadratic polynomial from the values x and the first derivatives dx/dy, at $y = 0$ and $y = 1$:

$$T_C \approx T_H\left(1 - \frac{\Delta H(T_H)}{\Delta C_P T_H}\right)^2 . \quad (28)$$

It should be noted that this approximation is an improvement over the standard one widely used in the literature [4, 17] (see Fig. 6), and makes the cold-denaturation condition (20) visible from the constraint that $T_C > 0$ for physical thermodynamic temperature.

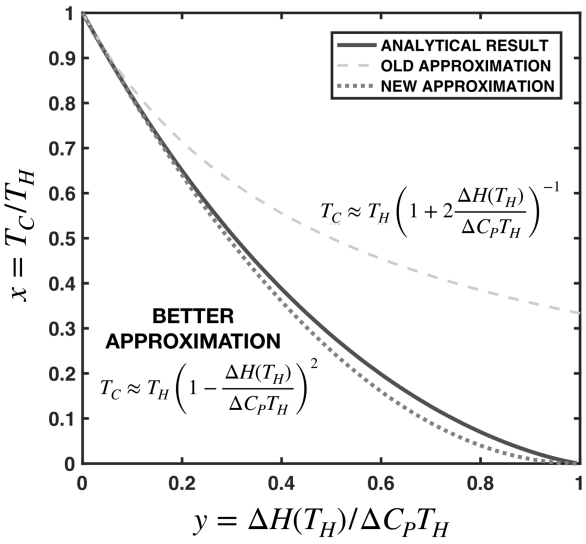

FIG. 6. Comparison between the analytical exact result (27), the standard approximation as given in [4, 13, 17] and our approximation (28).

The change in entropy at cold-denaturation $\Delta S(T_C)$ is given by:

$$\Delta S(T_C) = \Delta C_P\left(1 + W_{-1}\left(-\frac{(1-y)}{\exp(1-y)}\right)\right) = \Delta C_P\left(1 + W_{-1}\left(-\frac{\left(1 - \frac{\Delta H(T_H)}{\Delta C_P T_H}\right)}{\exp\left(1 - \frac{\Delta H(T_H)}{\Delta C_P T_H}\right)}\right)\right) < 0 . \quad (29)$$

The negative sign indicates that the transition from the denatured unfolded basin to the native folded state results in a decrease in entropy as the temperature is raised at the cold denaturation point, and therefore an increase in entropy as the temperature is cooled down at the cold denaturation point. This also follows from equation (15) and the graph for ΔG in Figure 3. The relation between $\Delta S(T_C)$, ΔC_P, $\Delta H(T_H)$ and T_H can be seen in Figure 7.

For completeness, we can redo the analysis above starting from Eq. (10) at cold-denaturation to obtain T_H from T_C and relevant information from $C_P(T)$ around

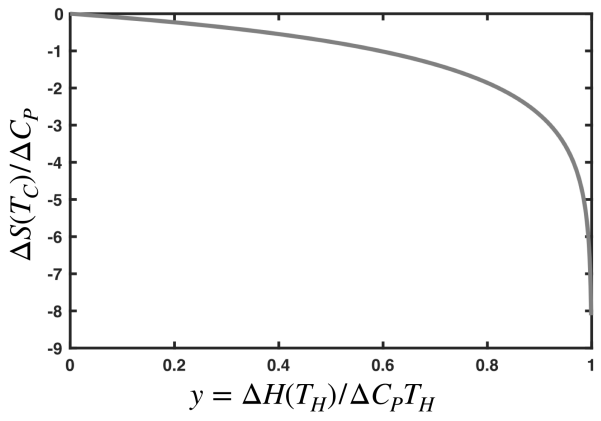

FIG. 7. The relation between $\Delta S(T_C)$, ΔC_P and y shows that $\Delta S(T_C) < 0$ for $0 < y < 1$ when $\Delta C_P > 0$.

that temperature:

$$T_H = T_C \exp\left(\left(1 - \frac{\Delta H(T_C)}{\Delta C_P T_C}\right) + W_0\left(-\frac{\left(1 - \frac{\Delta H(T_C)}{\Delta C_P T_C}\right)}{\exp\left(1 - \frac{\Delta H(T_C)}{\Delta C_P T_C}\right)}\right)\right), \tag{30}$$

note that $\Delta H(T_C) < 0$ which follows from Eq. (29). This can be approximated by replacing $\Delta H(T_H)$ with $\Delta H(T_C) + \Delta C_P(T_C - T_H)$ in Eq. (28):

$$T_H \approx T_C\left(1 - \frac{\Delta H(T_C)}{\Delta C_P T_C}\right)^2 \tag{31}$$

Since $\Delta H(T_C) = T_C \Delta S(T_C)$ and $\Delta S(T_C) < 0$ as we find in Eq. (29), it follows that $\Delta H(T_C) < 0$.

V. THE IMPORTANCE OF THE SHARPNESS OF THE DENATURATION PHASE TRANSITIONS

The zipper Hamiltonian is vastly favored by protein folding experts over the golf-course Hamiltonian because it naturally leads to a folding pathway approach to protein self-assembly, while the golf-course Hamiltonian would seem to pose an existential problem to protein self-assembly, as was pointed out by Levinthal [19].

However, a brief inspection of either the experimental melting curves of Mb in Fig. 1 or the simulated heat capacity calculations of the golf-course and zipper Hamiltonians in Fig. 2 shows that the melting curves of Mb at physiological conditions are surprisingly sharp, and that the golf-course Hamiltonian yields a much sharper melting profile than a zipper Hamiltonian. This indicates that actually the golf-course Hamiltonian would appear to be the better model for understanding cold and hot denaturation.

This observation can be made quantitative. At the peak of the melting curve, the specific heat has a peak value C_P. We can imagine two scenarios at the denaturation temperature. (1) The system can fluctuate between 2 discreet states, folded (F) and unfolded (U), or (2) it can move over a continuum of states, most broadly as a flat line between the F and U extrema. Clearly, the golf-course Hamiltonian represents the 2 discreet state possibility, while the the zipper model represents the continuum flat, broad distribution.

Following Bakk et al. [20], we note that the normalized inverse of the statistical variance α of a system of N amino acids can be written in terms of the ratio of the square of the total number of amino acids in the system N^2 divided by the difference between the average value of the square $\langle n^2\rangle$ and the square of the average value $\langle n\rangle^2$ of the number of amino acids in excited states:

$$\alpha = \frac{N^2}{\langle n^2\rangle - \langle n\rangle^2}\ , \tag{32}$$

which is evaluated in the vicinity of the phase transition temperatures (hot and cold denaturat.

The bigger the variance, the smaller is α, and the sharper is the transition; a 2-state system has the largest variance possible. That is, for a 2 state (golf-course) Hamiltonian there are only 2 possible states, 0 and 1, and so $\alpha_{gc} = 4$, and this is the smallest value that α can have. For the zipper Hamiltonian there is a flat continuum between 0 and 1 which yields an $\alpha_{zip} = 12$. Since the golf-course Hamiltonian must have the maximum variance for any distribution between 0 and 1, it also has the sharpest phase transition. The zipper Hamiltonian does not necessarily have the minimum variance, but it certainly will have a much broader melting peak than a 2-state system.

This statistical result can be recast into thermodynamic terms by noting that the specific heat at constant pressure C_P can be written in terms of the variance in the energy of the system (this is easily shown using the partition function Eq. 5):

$$C_P = \frac{\langle E^2\rangle - \langle E\rangle^2}{k_B T^2} \tag{33}$$

This allows us to write α in terms of measured thermodynamic quantities C_P^{peak} at the peak of the transition, and ΔH the latent enthalpy of the transition:

$$\alpha = \frac{\Delta H(T_C)^2}{k_B T_C^2 C_p^{peak}} \tag{34}$$

The smaller α is, the sharper the transition, with α =4 being the sharpest possible value (a 2 state transition), and $\alpha = 12$ for an evenly broad transition with all intermediate states equally populated, although α can be greater than 12 for other distributions.

As was long ago pointed out by Privalov [1] and emphasized by Bakk et al. [20], most water-soluble proteins have values of α very close to 4, with myoglobin perhaps the sharpest 2-state phase transition of all proteins studied. This indicates, rather shockingly, that proteins in their native state are actually crystalline like in their structure, and that the zipper model fails miserably. How then can proteins fold quickly, if Levinthal's Paradox is actually supported by the statistical mechanics of the calorimetry we have explored here?

While Bakk et al. want to explain this paradox by assuming that proteins fold through a series of crystalline sub-states (multiple pathways and hierarchical Hamiltonians), and this may be correct as far as it goes, we feel that they really should consider far more carefully the role that protein chaperones play in guiding proteins over a golf-course landscape rather than the putative funnels. Some combination of these 2 processes must be at work.

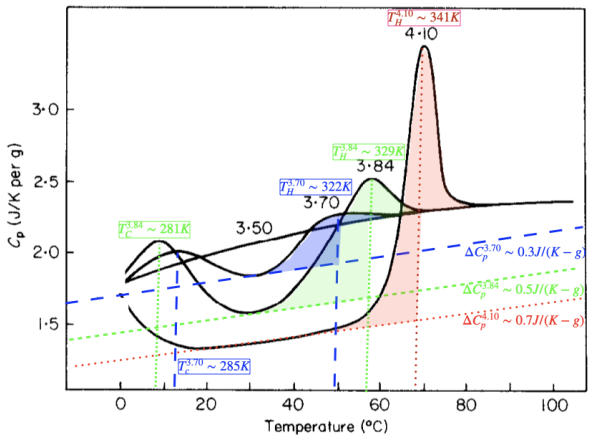

FIG. 8. Analysis of the metMB denaturation curves. 4 pH differential calorimetery scans were reported, at pH 4.10, 3.84, 3.70 and 3.50.

VI. COMPARISON WITH EXPERIMENTS

It should be no surprise that our better approximation for predicting cold denaturation temperatures, Eq. (28), yields good agreement with known cold denaturation temperatures.

For the numbers used in [4], $T_H = 60$C, $\Delta H(T_H) = 500$ kJ/mol and $\Delta C_P = 10$kJ/Kmol, our approximation (28) gives $T_C = -32$C which agrees with the numerical plot in [4], while the approximation used in the same paper results in $T_C = -17$C.

This equation is in quite good agreement with actual experimental results from the literature. For example, Chymotrypsinogen (pH 2.07 under 0.3GPa) has a measured cold-denaturation temperature at -9C [17, 21], while equation (28) yields a temperature of -10C. Similarly, Ribonuclease (pH 2.0 under 0.3GPa) has a cold-denaturation temperature measured at -32C, with a value of -29C from (28).

We are now equipped to understand at a deeper level the denaturation curves of met-Mb shown first in Fig. 1 and now presented again in Fig. 8 with some analysis. We can use (28) to predict the steps in the low temperature denaturation temperature seen in Fig. 1, and we can extract values for α which give measurements of the distribution of protein states at the phase transitions.

Fig. 8 shows some of the extracted thermodynamic values necessary to compute the predicted low temperature denaturation temperatures of Mb for 3 different pH values. In principle, we should be able to go backwards, that is, predict the high temperature denaturation temperatures from the low temperature thermodynamics but unfortunately since the data cut off at 0^oC it is not possible to get the critical values for the thermodynamics variables from this data set. Table I. presents predicted and measured values of T_c.

Next we can determine how the variance in the states given by α changes as one approaches presumably a fully denatured Mb at pH 3.50, which has NO phase transitions.

pH	T_H	ΔC_p(J/gK)	ΔH(J/g)	$T_c^{predicted}$	$T_c^{measure}$	α
4.10	341 K	0.7	26	270 K	N/A	4
3.84	329 K	0.5	14	275 K	281 K	11
3.70	322 K	0.3	4.5	293 K	285 K	16

TABLE I. Thermodynamic analysis of predicted cold denaturation temperatures and high T_c values, and α from Fig. 8.

Our analysis shows (roughly) that we start with a clean 2-state denaturation process at pH 4.10, to an almost flat distribution of states at pH 3.84, with a flat or centered intermediate unfolded distribution at pH 3.70. This analysis suggests that indeed proteins at normal pH ranges are well described by the golf-course Hamiltonian, but progress towards a zipper-like Hamiltonian as the pH (in this case) is lowered until the protein is completely denatured at all temperatures.

VII. A MICROFLUIDIC TECHNIQUE FOR HIGH AND LOW TEMPERATURE DENATURATION

In order to probe the temperature and pressure dependent stability and possible metastability of proteins we have developed a unique microfluidic technology which will allow studies of protein stability over an unprecedented range of temperature and pressures using a microfluidic bubble technology coupled with optical and X-ray techniques.

The basic idea is simple: since we want to study protein conformations over a temperature range from -10 C to +60 C, we will have to supercool proteins dissolved in water to access regions where the ΔG clicks below zero. Although water can be super-cooled because of its

high surface tension, bulk water cannot be stably cooled below 0 C because small fluctuations can nucleate microcrystals [22]. However, in a water drop, the smaller the volume of the water drop V_D the less likely is the formation of a random fluctuation in that drop. Hence, highly purified water, when made into a fine mist, can consist of super-cooled water droplets (also known as clouds).

It is possible, using microfabrication techniques, to generate arbitrarily small water drops using hydrodynamic focusing of a central water jet and side jets of a water-immiscible oil [23]. A fundamental task is finding a water immiscible hydrocarbon of high enough molecular weight that it has a low vapor pressure, and hence will not partition into the water nanodrops, but still have a low enough melting point that we can supercool to at least -10 C with low viscosity. We have found that the simple aliphatic hydrocarbon undecane ($C_{11}H_{24}$) has the desired properties. Unfortunately undecane has undesirable swelling properties for many elastomers (such as polydimethylsiloxane, PDMS) which has required us to develop all-glass and silicon based microfludics, and epoxy-based sealing of microfluidic sealing gaskets to the external pumps.

The chip we have developed to do hydrodynamic focusing uses deep reactive ion etching of channels into silicon, as shown in the SEM image in Fig. 9. The thin (100 micron) pyrex glass coverslip is anodically bonded to the silicon chip. A computer controlled peltier-effect device is placed against the down-stream channel of the nanodrop generator and can be run either as a heater or a cooler to explore the heat denaturation and cold denaturation of the protein molecules within the nanodrops.

Optical studies of the FRET distances of labels in the protein can be done through the thin pyrex window at the single-molecule level [24]. Back-etching of the chip allows us to thin the wafer to under 20 microns over the channel to allow ultimately for small-angle X-ray scattering (SAXS) measurements of the protein radial distribution moments as a function time and temperature after nanodrop generation [25].

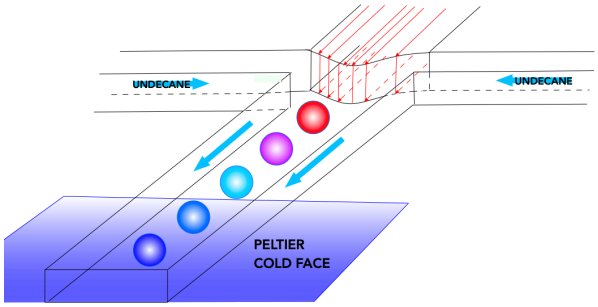

FIG. 9. Schematic of a microdrop generator for supercooled water-protein studies. Undecane comes in from the 2 side channels and a protein solution in water comes in from the top. Adjustment of the flow rates yields nanodrops of water which then pass over a temperature controlled peltier plate.

With this technology, it should be possible to explore in a more systematic way various water soluble proteins and their high and low denaturation points, as a test of the utility of the analysis developed here.

VIII. CONCLUSIONS

In this paper, we showed that, under widely satisfied conditions in globular proteins, for each hot denaturation point, T_H, there exists a unique cold denaturation temperature T_C for a protein. We explicitly calculate T_C in terms of T_H and the other parameters. We also calculated the entropy change $\Delta S(T_C)$, which is shown to be positive with unfolding. We demonstrate agreement with experiments, and correct some fundamental mistakes in the literature.

We used the golf course and zipper Hamiltonians, coupled with their water-protein Hamiltonian of Sneppen and his colleagues, to verify that their elegantly simple physics can reproduce the phenomena of low and high temperature denaturation of globular proteins. This can capture the 2-state nature of the pseudo-first order phase transitions observed, implying, after all these years, that the 2-state model should be taken more seriously.

There are two dark sides to this paper we have to discuss. One dark side, as we have mentioned, is that the sharpness of the low and high denaturation temperatures as evidenced by the state variance given by α implies a discrete 2-state distribution of states under biological conditions, not a continuum, supporting the golf-course Hamiltonian as opposed to the zipper Hamiltonian.

The second dark side of this paper calls into question the interpretation, but not the data, of the work on the dynamics of Mb-CO at low temperatures [26]. Since the authors of that paper cooled myoglobin slowly in a glycerol-water solvent, there is a distinct possibility that the low temperature structure dynamics studied was not of the native Mb structure at room temperature, but rather of a cold-denatured state. Given that presumably the cold-denatured state is a high entropy state consisting basically of a random coil (although that is not known), perhaps it would not be surprising to see glassy-like power law activation energy distributions [27]. This is purely speculative on our part at present since nobody has done differential scanning calorimetry of Mb in a glycerol-water solvent to verify that there is a cold denaturation in this solvent.

One obvious next step is to better understand the broken symmetry of the quantum entropy states of water at cold denaturation, which clearly results in a different change in conformational degrees of freedom than occurs in hot denaturation, and drives the strange phenomenon of cold denaturation. Next, we believe that the actual structure of the cold denatured and hot denatured states need to be measured since they are so critically linked to the conformation of proteins weakly held together between the two denaturation phase changes.

IX. ACKNOWLEDGMENTS

T.P. thanks Tuan K. Do for discussion of models of proteins in water. This work was supported by the US National Science Foundation PHY-1659940, the National Science Foundation, through the Center for the Physics of Biological Function (PHY-1734030) and the Princeton Catalysis Initiative.

[1] P. L. Privalov, V. Griko Yu, S. Venyaminov, and V. P. Kutyshenko, J Mol Biol **190**, 487 (1986).
[2] H. Frauenfelder, P. G. Wolynes, and R. H. Austin, Reviews of Modern Physics **71**, S419 (1999).
[3] C. L. Dias, T. Ala-Nissila, J. Wong-ekkabut, I. Vattulainen, M. Grant, and M. Karttunen, Cryobiology **60**, 91 (2010).
[4] S. Kunugi and N. Tanaka, Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology **1595**, 329 (2002).
[5] J. N. Onuchic, Z. LutheySchulten, and P. G. Wolynes, Annual Review of Physical Chemistry **48**, 545 (1997).
[6] E. C. de Macario and A. J. L. Macario, Frontiers in Bioscience-Landmark **5**, D780 (2000).
[7] K. Swiderek, I. Tunon, S. Marti, and V. Moliner, Acs Catalysis **5**, 1172 (2015).
[8] K. Ghosh, A. M. R. de Graff, L. Sawle, and K. A. Dill, Journal of Physical Chemistry B **120**, 9549 (2016).
[9] M. L. Glasser, E. H. Lieb, and D. B. Abraham, Journal of Mathematical Physics **13**, 887 (1972).
[10] D. P. Fernandez, Y. Mulev, A. R. H. Goodwin, and J. M. H. L. Sengers, Journal of Physical and Chemical Reference Data **24**, 33 (1995).
[11] R. Baxter, *Exactly Solved Models in Statistical Mechanics* (Dover, 2007).
[12] L. Pauling, Journal of the American Chemical Society **57**, 2680 (1935).
[13] K. Sneppen and G. Zocchi, *Physics in molecular biology* (Cambridge University Press, 2005).
[14] B. J. Sirovetz, N. P. Schafer, and P. G. Wolynes, Journal of Physical Chemistry B **119**, 11416 (2015).
[15] M. S. Li, D. K. Klimov, and D. Thirumalai, Physical Review Letters **93**, ARTN 268107 10.1103/PhysRevLett.93.268107 (2004).
[16] B. Uralcan and P. G. Debenedetti, Journal of Physical Chemistry Letters **10**, 1894 (2019).
[17] C. L. Dias, T. Ala-Nissila, J. Wong-ekkabut, I. Vattulainen, M. Grant, and M. Karttunen, Cryobiology **60**, 91 (2010).
[18] R. M. Corless, G. H. Gonnet, D. E. Hare, D. J. Jeffrey, and D. E. Knuth, Advances in Computational mathematics **5**, 329 (1996).
[19] C. Levinthal, Journal De Chimie Physique Et De Physico-Chimie Biologique **65**, 44 (1968).
[20] A. Bakk, P. G. Dommersnes, A. Hansen, J. S. Hoye, K. Sneppen, and M. H. Jensen, Computer Physics Communications **147**, 307 (2002).
[21] S. A. Hawley, Biochemistry **10**, 2436 (1971).
[22] F. C. Frank, Nature **157**, 267 (1946).
[23] J. P. Brody, P. Yager, R. E. Goldstein, and R. H. Austin, Biophysical Journal **71**, 3430 (1996).
[24] X. Sun, T. E. Morrell, and H. Yang, Journal of Physical Chemistry B **120**, 10469 (2016).
[25] L. Pollack, M. W. Tate, A. C. Finnefrock, C. Kalidas, S. Trotter, N. C. Darnton, L. Lurio, R. H. Austin, C. A. Batt, S. M. Gruner, and S. G. J. Mochrie, Physical Review Letters **86**, 4962 (2001).
[26] R. H. Austin, K. W. Beeson, L. Eisenstein, H. Frauenfelder, and I. C. Gunsalus, Biochemistry **14**, 5355 (1975).
[27] R. H. Austin, K. Beeson, L. Eisenstein, H. Frauenfelder, I. C. Gunsalus, and V. P. Marshall, Physical Review Letters **32**, 403 (1974).