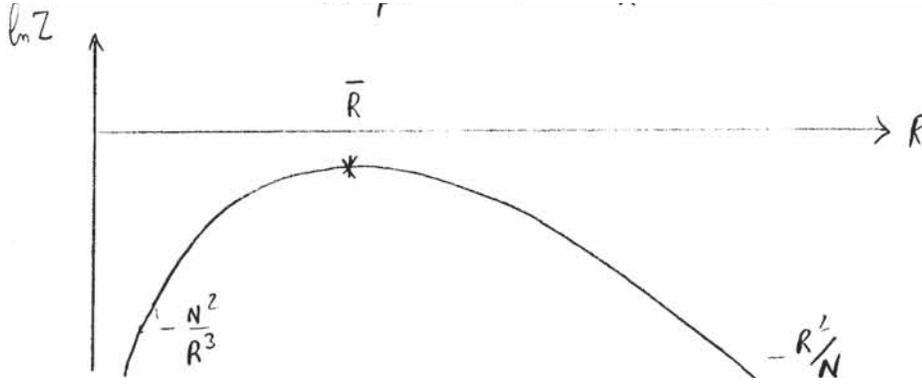


2.3.2 Swollen (coil) polymers in good solvents

Most of the terms in the trial free energy of Eq. (2.49) have definite sign. The exception is the term proportional to $N^2(a/R)^3$ which has opposing contributions from the repulsive and attractive parts of the potential, and is proportional to $(\chi - 1/2)$. The sign of this term determines whether attraction or repulsion is the dominant effect, leading to two different phases. For $\chi < 1/2$ the latter is more important favoring large R and swollen polymers. This tendency is opposed by the reduction of entropy at larger R . Indeed, one can self-consistently check that all other terms are less important in this limit, such that

$$\ln Z(N, R) = \text{constant} - \frac{3R^2}{4Na\xi_p} - \frac{1 - 2\chi}{2} N^2 \left(\frac{a}{R}\right)^3 + \text{higher order terms.} \quad (2.50)$$



Extremizing the above expression leads to

$$\begin{aligned} \frac{\partial \ln Z}{\partial R} &= -\frac{3R}{2Na\xi_p} + \frac{3(1 - 2\chi)}{2} N^2 \left(\frac{a^3}{R^4}\right) \Rightarrow \bar{R}^5 = (1 - 2\chi)a^4\xi_p N^3, \\ \bar{R} &= (1 - 2\chi)^{1/5} (a^4\xi_p)^{1/5} N^{3/5}. \end{aligned} \quad (2.51)$$

In the absence of interactions, the typical size of the polymer grows as $\sqrt{a\xi_p N}$, Eq. (2.39). An interesting consequence of repulsion due to excluded volume is that the scaling of size is changed to $R \propto N^\nu$, with an exponent $\nu > 1/2$. The variational treatment leading to Eq. (2.51) thus predicts the so-called *Flory exponent* of $\nu = 3/5$.

Going beyond the mean-field variational treatment is not trivial, and one of the triumphs of renormalization group theory is to estimate the exact value of $\nu = 0.591\dots$, remarkably close to the Flory approximation of $3/5$. While not directly relevant to real polymers, it is possible to inquire about the exponent ν for self-avoiding walks in d -spatial dimensions— e.g., for polymers confined to a $d = 2$ dimensional surface. Ignoring the attractive part of the interaction, but incorporating the repulsive cores, generalizes Eq. (2.52) to

$$\ln Z(N, R) = \text{constant} - \frac{dR^2}{4Na\xi_p} - \frac{N^2}{2} \left(\frac{a}{R}\right)^d. \quad (2.52)$$

Extremization now gives

$$\frac{\partial \ln Z}{\partial R} = -\frac{dR}{2Na\xi_p} + \frac{d}{2}N^2 \left(\frac{a^d}{R^{d+1}} \right) \Rightarrow \bar{R} = (a^{d+1}\xi_p)^{\frac{1}{d+2}} N^{\frac{3}{d+2}},$$

i.e. a generalized Flory exponent of

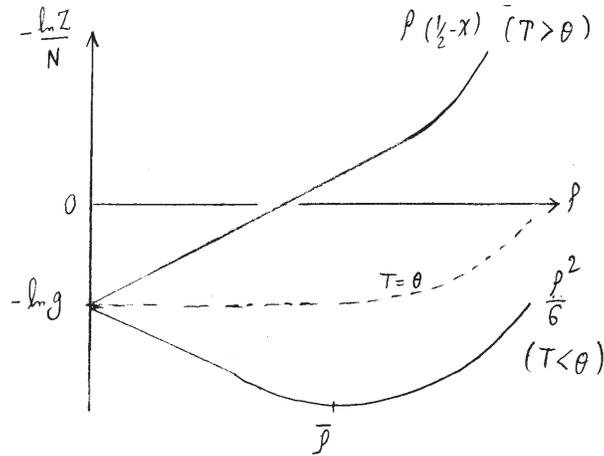
$$\nu_F(d) = \frac{3}{d+2}. \quad (2.53)$$

The predicted values of $\nu = 1, 3/4, 1/2$ in $d = 1, 2, 4$ are in fact exact. Above four dimensions the excluded volume constraint is irrelevant and ν remains fixed at $1/2$.

2.3.3 Compact (globular) polymers in bad solvents

On lowering temperature $\chi(T)$ typically becomes larger, and the coefficient $(1 - 2\chi)$ in Eq. (2.52) changes sign at the so-called θ -point ($\chi(\theta) = 1/2$). At temperatures $T < \theta$ the attractive component of the interaction is more important leading to compact (globular) shapes with a finite number density $\rho = N(a/R)^3$. The leading terms in the expansion of the variational free energy can now be recast as

$$-\frac{\ln Z(\rho)}{N} = -\ln g + \frac{1 - 2\chi}{2}\rho + \frac{\rho^2}{6} + \text{higher order terms}. \quad (2.54)$$



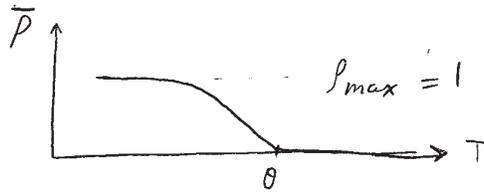
The optimal density for $T < \theta$ is obtained by minimizing the above free energy,

$$-\frac{1}{N} \frac{d \ln Z}{d \rho} = \left(\frac{1}{2} - \chi \right) + \frac{\rho}{3} + \dots,$$

leading to

$$\bar{\rho} = 3 \left(\chi - \frac{1}{2} \right) + \dots, \quad (2.55)$$

i.e. a density that vanishes linearly on approaching the θ -temperature from below. The higher order terms ensure that the density does not exceed the maximum value of unity in a fully compact state.



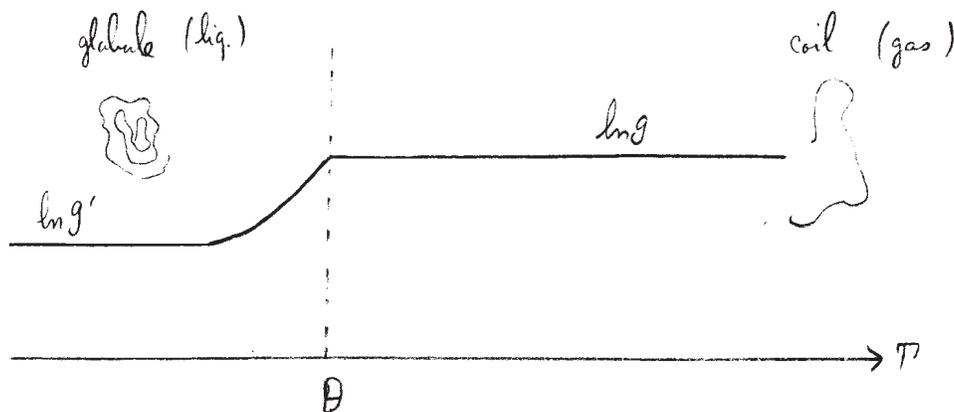
From Eq. (2.48) we note that the energy gain per particle from the attractive interactions satisfies

$$\frac{E_{att.}}{N} = -\bar{\rho}\chi k_B T. \quad (2.56)$$

The entropy per particle is then given by

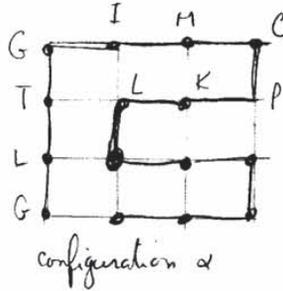
$$\frac{S}{Nk_B} = \frac{-F + E}{Nk_B T} = \frac{\ln Z}{N} - \bar{\rho}\chi = \ln g + \frac{3}{2} \left(\chi - \frac{1}{2} \right)^2 + \dots - \bar{\rho}\chi \approx \ln g - \frac{\bar{\rho}}{2} + \mathcal{O}(\bar{\rho}^2). \quad (2.57)$$

(The final expression includes only the leading linear term as $\chi \rightarrow 1/2$.) Thus close to the θ -temperature the entropy is reduced, initially linearly in temperature, although it also will eventually saturate as does the density. The above analysis is reminiscent of the mean-field analysis of the transition between a gas (low density) and a liquid (high density). The liquid state still encompasses many particle configurations, although fewer than in a gas. Further cooling of liquids typically leads to frozen states with even lower entropy. We may thus inquire if such a freezing transition also exists for polymers.



2.3.4 The Random Energy Model (REM) for compact heteropolymers

Deep in the globular phase, the states of the compact polymer can be visualized as the collection of all maximally compact configurations. In a lattice version, these are self-avoiding walks that visit all sites, leaving no empty ones, and are referred to as *Hamiltonian walks*. The number of Hamiltonian walks also grows exponentially with the number of steps as g'^N , but is much smaller than the number of self-avoiding walks ($g^N \gg g'^N$). For a homopolymer all such configurations are equally likely, but in a heteropolymer the distinct interactions between different monomers leads to variations in energy. Presumably at low temperatures the lower energy states are preferred, and there can potentially be a phase transition to a specific (ground state) configuration. For biological molecules, there are *non-specific* attractive forces that tend to aggregate all monomers, whereas *specific* interactions select a particular (native) shape amongst the manifold of possible compact states.



To explore this scenario, consider all compact configurations for a multi-component heteropolymer such as a protein. The energy of a configuration α is given by

$$E_\alpha = \sum_{\langle ab \rangle} V_{ab}, \quad (2.58)$$

where the sum is over all non-polymeric nearest-neighbor pairs $\langle ab \rangle$, and V_{ab} is the interaction energy assigned to a neighboring pair of monomers a and b . The partition function is obtained by the sum

$$Z = \sum_{\alpha} e^{-\beta E_\alpha}, \quad (2.59)$$

over the g'^N states. To make headway with this hard problem, we make the drastic approximation of assuming that the bond energies V_{ab} are independent random variables. Subject to this assumption, the energies E_α are themselves random variables, and as long as the number of terms N_B in Eq. (2.58) is large, taken from a Gaussian distribution. The mean and variance of the distribution are given by

$$\langle E_\alpha \rangle = N_B \langle V_{ab} \rangle \equiv N \varepsilon_0, \quad \langle E_\alpha^2 \rangle_c = N_B \langle V_{ab}^2 \rangle_c \equiv N \sigma^2, \quad (2.60)$$

where noting that $N_B = (z-1)N$ (of the z per each site of the lattice, one is polymeric), we have folded the proportionality constant into the definitions of ε_0 and σ^2 .

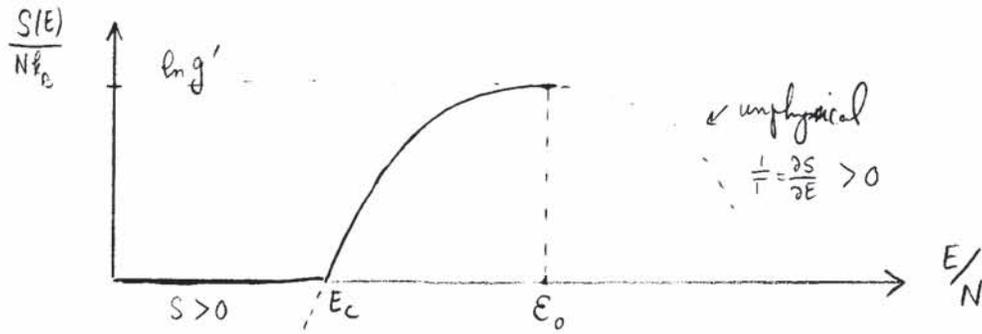
For large N , the probability distribution for the energy will take the Gaussian form

$$p(E) = \frac{1}{\sqrt{2\pi N\sigma^2}} \exp\left[-\frac{(E - N\varepsilon_0)^2}{2N\sigma^2}\right]. \quad (2.61)$$

Since the total number of states is g^N , the density of states is $\Omega(E) = g^N p(E)$, and the entropy of this *random energy model* (REM) is given by

$$S(E) = k_B \ln \Omega(E) = k_B \left[N \ln g' - \frac{(E - N\varepsilon_0)^2}{2N\sigma^2} \right] - \frac{k_B}{2} \ln(2\pi N\sigma^2). \quad (2.62)$$

The last term is not extensive (proportional to N) and can be safely ignored.



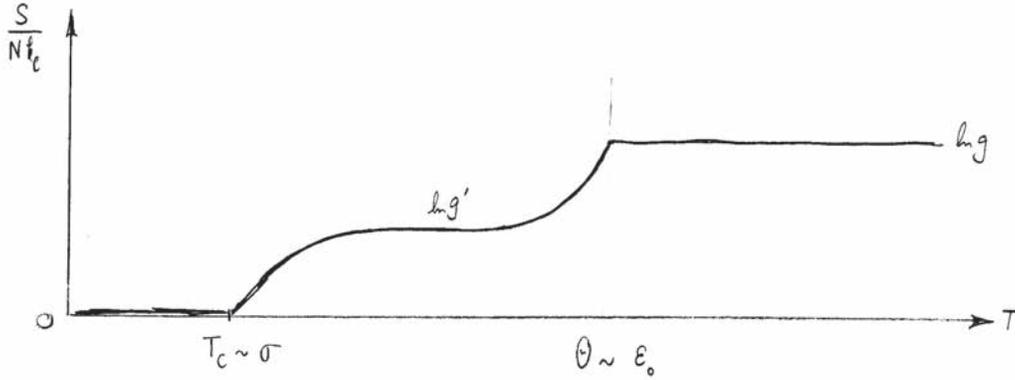
According to Eq. (2.62), $S(E)$ is shaped like a parabola, but thermodynamic constraints imply that only a certain portion of this curve is physical. First, the temperature T is obtained from the slope of the curve via $T^{-1} = dS/dE$. Positive temperatures require the entropy to increase with temperature, and thus only the states with $E < N\varepsilon_0$ are physically accessible. Second, the entropy cannot be negative, and $S(E)$ should thus stick to zero for $E < E_c$, where E_c is easily obtained as

$$S(E_c) = 0 \quad \implies \quad \frac{E_c}{N} = \varepsilon_0 - \sigma \sqrt{2 \ln g'}. \quad (2.63)$$

(Note the connection to the extreme value problem studied earlier: E_c is also the mean value of the lowest of g^N energies randomly selected from $p(E)$.) The singularity of entropy at E_c signifies a phase transition into a glassy state, at a temperature T_c given by

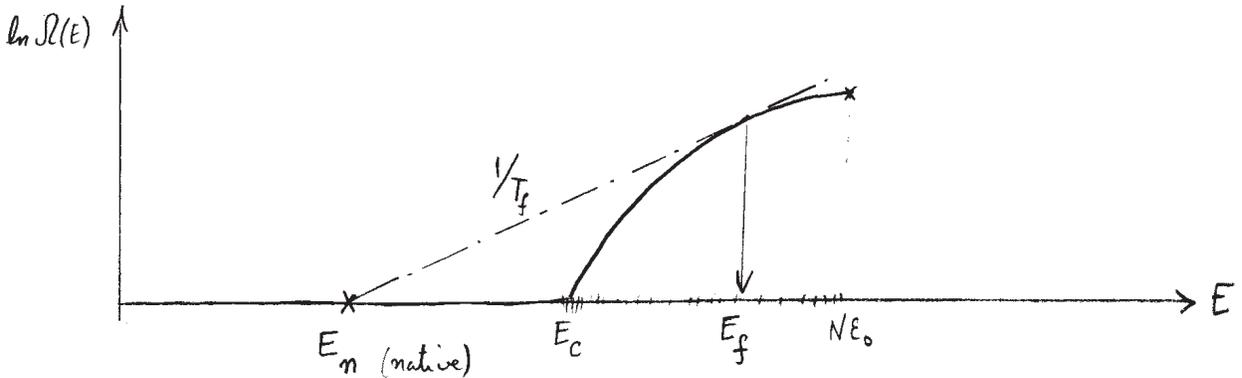
$$\frac{1}{T_c} = \left. \frac{dS}{dE} \right|_{E_c} = -k_B \left(\frac{E_c - N\varepsilon_0}{N\sigma^2} \right) = k_B \frac{\sqrt{2 \ln g'}}{\sigma} \implies k_B T_c = \frac{\sigma}{\sqrt{2 \ln g'}}. \quad (2.64)$$

There are presumably a few low energy states with energy close to E_c , and the system freezes into one of these for $T \leq T_c$.



2.3.5 Designed REM for protein folding

It is tempting to equate the freezing of the heteropolymer with the folding transition separating denatured and folded states of a protein. There is, however, a problem with such an interpretation: As the temperature is lowered towards T_c , the number of states decreases drastically. It is unlikely that the lower energy states of the REM in the vicinity of E_c have much in common. To change its state the polymer will likely have to rearrange many of its monomers, running into high energy barriers in the process. Thus we expect that the kinetics of the REM polymer will slow down significantly on approaching T_c . This contradicts the observation that most proteins fold easily and in a short time. Of course proteins are not typical random heteropolymers, and are presumably “designed” through evolution for both function and ease of folding. Fortunately, we can mimic such “design” by a small modification of the REM; we only need to add to the continuum of random energy states, a single state with low energy ($E_n < E_c$) representing the native configuration.



With the added state at E_n , the system makes a transition to the native state (i.e. folds) at a temperature T_f , high enough that there are still many equivalent states to explore. The location of T_f , and the corresponding energy E_f , can be obtained by equating free energies or Boltzmann weights, and leads to the “tangent construction” whereby T_f and E_f are related

to E_n via

$$\beta_f = \frac{S(E_f)/k_B}{E_f - E_n} = \frac{N \ln g' - (E_f - N\varepsilon_0)^2/(2N\sigma^2)}{E_f - E_n}. \quad (2.65)$$

As depicted in the figure, the above result equates the slope of the tangent line from the point at E_n computed in two different ways.

To justify the above result, note that in the canonical ensemble, the probability of finding the system in the native state is

$$p_n = \frac{e^{-\beta E_n}}{Z(\beta)}, \quad \text{with} \quad Z(\beta) = e^{-\beta E_n} + \int dE \Omega(E) e^{-\beta E}. \quad (2.66)$$

A phase transition in which p_n changes discontinuously from zero to one occurs only in the thermodynamic limit of $N \rightarrow \infty$. For the system to have a well-behaved thermodynamic limit (in which case various thermodynamic identities involving entropy and temperature can be safely used), we must insist that the range of energies as well as $\ln \Omega(E)$ should be proportional to N ; the former implies that $E_n \propto N$. If so, then at a particular value of β a single value of energy E completely dominates the partition function $Z(\beta)$. For the partition function in Eq. (2.66), the dominant value occurs for some $E \geq E_f$ for $\beta \leq \beta_f$, and for $E = E_n$ for $\beta > \beta_f$. The probability to find the system in its native state then jumps discontinuously from 0 to 1 at the point when the corresponding contributions to the partition function are equal, i.e. at

$$e^{-\beta_f E_n} = \Omega(E_f) e^{-\beta_f E_f}, \quad (2.67)$$

which after taking the logarithm leads to the tangent rule in Eq. (2.65).

We can eliminate E_f in terms of β_f by noting that $E = N\varepsilon_0 - N\sigma^2\beta$, and $\ln g' = (\beta_c\sigma)^2/2$. Using these expressions and defining a quantity $\beta_n = (E_n - N\varepsilon_0)/(N\sigma^2)$, the above equation reduces to

$$\beta_f = \frac{\beta_c^2 - \beta_f^2}{-2\beta_f + 2\beta_n}. \quad (2.68)$$

This can be rearranged as a quadratic equation with solution

$$\beta_f = \beta_n - \sqrt{\beta_n^2 - \beta_c^2}. \quad (2.69)$$

The ratio of the folding temperature to the REM freezing temperature is thus

$$\frac{T_f}{T_c} = \frac{\beta_c}{\beta_f} = \frac{\beta_n}{\beta_c} + \sqrt{\left(\frac{\beta_n}{\beta_c}\right)^2 - 1}. \quad (2.70)$$

Faster folding to the native state can be achieved at higher temperatures by increasing the energy difference between E_n and E_c .

MIT OpenCourseWare
<http://ocw.mit.edu>

8.592J / HST.452J Statistical Physics in Biology
Spring 2011

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.