PROTEIN FOLDING AND STABILITY

Protein thermodynamics constrain protein evolution

We have seen that the residues on the surface of TEM-1 β-lactamase evolve more rapidly than those in the core, a pattern that is (statistically) universal among proteins with well-defined surfaces and cores. This suggests that the spatial distribution of residues in a protein's folded state dictates a substantial portion of the fitness cost associated with changing those residues. Protein folding constrains evolutionary change. Why?

An easy answer is that changes in the core are more likely to disrupt folding than surface changes, perhaps because the core is tightly packed and thus more sensitive to substitutions than the surface. But can a single amino acid change in a typical protein, made of hundreds of amino acids, actually change the protein in ways visible to natural selection? Proteins are stabilized by elaborate networks of intramolecular interactions, often constituting hundreds of hydrogen-bonding equivalents of energy; how much can one substitution altering at most a handful of hydrogen bonds possibly alter protein stability?

To answer these and other questions, we need a more precise grasp of protein stability and how it is achieved. To position ourselves for understanding deep results in the evolution of protein folding and stability, we need to understand a few concepts which may be unfamiliar: energy, enthalpy, entropy, and free energy.

Thermodynamics and protein stability

We would like to know:

(1) What determines a protein's stability in its native state?
(2) How can we describe how amino-acid substitutions alter protein stability?
(3) Why are proteins marginally stable?

Thermodynamics provides us with a starting point. The most fundamental concept is **energy**, a property of any physical system which is neither created nor destroyed and has been found to obey certain laws. Energy is never observed directly, but shockingly successful physical theories arise from quantitative assertions about its existence and behavior.

The **internal energy** U is the energy intrinsic to a particular system, the energy arising from molecular interactions: covalent bonds, electrostatic interactions, hydrogen bonding, van der Waals interactions, and so on.

The **enthalpy** of a system is $H = U + pV$ where p is pressure and V is volume; this is the intrinsic energy of a system plus the energy required to make room for the system. In most biological systems, p and V will be roughly constant.

A system's **entropy** S measures the number of microsopic configurations compatible with the system's macroscopic state, a notion which statistical mechanics clarifies considerably (see below).

The temperature T should be familiar, though statistical mechanics generalizes its physical meaning.

The **Gibbs free energy** $G = H - TS = U - TS + pV$ provides the core thermodynamic variable from which, for our purposes, all others derive their utility. This is because spontaneous chemical changes—reactions—from A to B imply $G_A \leq G_B$, or equivalently $G_A - G_B = \Delta G_{AB} \leq 0$. From our familiar kinetic starting point

$$\mathrm{A} \underset{k_r}{\overset{k_f}{\rightleftharpoons}} \mathrm{B}$$

we recall the equilibrium constant $K_{\mathrm{eq}} = \frac{k_r}{k_f} = \frac{[A]}{[B]}$ characterizing the equilibrium of this reaction,

$$\mathrm{A} \overset{K_{\mathrm{eq}}}{\rightleftharpoons} \mathrm{B}.$$

Quite often, we know only K_{eq} and have no access to the kinetic rate constants, because we possess information only about the equilibrium state. At equilibrium, the change in free energy ΔG similarly characterizes the reaction:

$$\mathrm{A} \overset{\Delta G}{\rightleftharpoons} \mathrm{B} \tag{1}$$

and the equilibrium constant K_{eq} and change in free energy ΔG are related by

$$\Delta G = -RT \ln K_{\mathrm{eq}} = -RT \ln \frac{[A]}{[B]} \tag{2}$$

where $R = 8.314$ J/(mol·K) is the universal gas constant and ΔG is expressed in kJ/mol or $R = 0.001986$ kcal/(mol·K) and ΔG is expressed in kcal/mol.

Assume for the moment that a protein molecule exists in one of two states, folded and unfolded. Each state has an associated free energy, G_f and G_u, respectively. We'd like to have measure of protein stability, a number that increases with the tendency of the protein to be found in its folded state, and this measure is

$$\Delta G_u := G_u - G_f = -RT \ln \frac{[U]}{[F]} \tag{3}$$

ΔG_u is the **free energy of unfolding** per mole and we will use it and "stability" interchangeably. One also sees the free energy of folding $\Delta G_f = -\Delta G_u$ used often; decreasing ΔG_f means increasing stability. From Eq. 3 we can, given a value for ΔG_u, determine the fraction of protein that is folded at equilibrium:

$$\frac{[F]}{[U]} = e^{\Delta G_u / RT} \tag{4}$$

and thus (because $[F]/[U]$ are the odds of a folded versus an unfolded protein)

$$\Pr(\text{folded}) = \frac{e^{\Delta G_u / RT}}{1 + e^{\Delta G_u / RT}} = \frac{1}{1 + e^{-\Delta G_u / RT}} \tag{5}$$

At $\Delta G_u = 0$, half of the protein molecules are unfolded.

Most proteins require some level of stability to function, and the stability of arbitrarily chosen proteins can be compared, which forms the basis of a common language for talking about protein biophysical properties. (Think about the difficulty in trying to compare the *function* of arbitrarily chosen proteins: how do we compare actin's function to hexokinase's?) Consequently, the effects of mutations are commonly reported as changes in free energy $\Delta\Delta G$:

$$\Delta\Delta G := \Delta G_{\text{mutant}} - \Delta G_{\text{wild-type}} \tag{6}$$

From thermodynamics to molecules: statistical mechanics

Statistical mechanics is the study of how macroscopic (thermodynamic) properties like temperature, internal energy, pressure, and so on arise from the behavior of large collections of individual molecules. A command of basic results from statistical mechanics will allow us to understand key results connecting protein folding and evolution. Coincidentally, and happily, many quantitative tools in molecular evolution directly parallel those in statistical mechanics, meaning we can learn one set of tools and use it in both domains.

Consider a large collection—a system or **ensemble**—of molecules, such as the collection of $\sim 10^6$ pyruvate kinase molecules inside a typical yeast cell, which for a meaningful amount of time exists at constant temperature T, number of molecules N and volume V, while exchanging energy with its surroundings. Each molecule can exist in multiple microscopic states or **microstates**. Statistical mechanics tells us that if state i has energy E_i per mole, then at equilibrium, the probability that a molecule in the ensemble will be in microstate i is

$$p_i = \frac{e^{-E_i/RT}}{\sum_j e^{-E_j/RT}} := \frac{e^{-E_i/RT}}{Z} \tag{7}$$

In statistical mechanics, one is often concerned with energies per molecule rather than per mole, and thus you will see $R/N_A = k_B = 1.38 \times 10^{-23} J/K$ used often, where N_A is Avogadro's number and k_B is **Boltzmann's constant**. Most experimental results we wish to make contact with will be in energies per mole, so we will use R throughout.

The denominator of this proportion we denote $Z := \sum_i e^{-E_i/RT}$, the **partition function**. Because Z is a sum, we can refer to terms corresponding to a particular macroscopic state or states of interest, say "folded" and "unfolded", as partition functions for those states:

$$Z = \sum_i e^{-E_i/RT} = \underbrace{\sum_{i \in \text{folded}} e^{-E_i/RT}}_{Z_{\text{folded}}} + \underbrace{\sum_{i \in \text{unfolded}} e^{-E_i/RT}}_{Z_{\text{unfolded}}}$$

(Recall that "$i \in X$" means "i taking each value in the set X.") Because probabilities for independent microstates add,

$$p_{\text{folded}} = \sum_{i \in \text{folded}} \frac{e^{-E_i/RT}}{Z} = \frac{Z_{\text{folded}}}{Z}$$

$$p_{\text{unfolded}} = \frac{Z_{\text{unfolded}}}{Z}$$

and so on. Many interesting biological states can be productively described in this way: open and closed; bound and unbound; phosphorylated, acetylated and unmodified; *etc.*

A system's internal energy U is its energy over all its possible microstates,

$$U = \sum_i p_i E_i. \tag{8}$$

The **entropy** S of the system is

$$S = -R \sum_i p_i \ln p_i. \tag{9}$$

in energy/(mol·K). Entropy is zero when the system is confined to a single microstate, and is maximized when all microstates are equally likely. The entropy of a macroscopic state thus relates to the number of distinct physical ways that state can be realized—and the more ways there are, the more probable the state. This is the intuitive basis for the second law of thermodynamics, which states that for any closed system, entropy increases: systems move inexorably toward states compatible with larger numbers of microstates, because there are more ways to be in such states.

Using the above expressions for internal energy and entropy, the thermodynamic relation for the Gibbs free energy, $G = U - TS + pV$, can be written

$$\begin{aligned} G &= \sum_i p_i E_i - T\left(-R\sum_i p_i \ln p_i\right) + pV \\ &= \sum_i p_i E_i + RT\sum_i p_i \left(\ln e^{-E_i/RT} - \ln Z\right) + pV \\ &= \sum_i p_i E_i - \sum_i p_i E_i - RT\ln Z + pV \\ &= -RT\ln Z + pV \end{aligned}$$

As above, if we have two states of interest, A and B, then we can write

$$\begin{aligned} G_A &= -RT\ln Z_A + pV \\ G_B &= -RT\ln Z_B + pV \end{aligned}$$

and then the change in free energy between these states is

$$\begin{aligned} \Delta G_{AB} &= G_A - G_B \\ &= -RT\ln Z_A + pV + RT\ln Z_B - pV \\ &= -RT\ln\frac{Z_A}{Z_B} \end{aligned} \tag{10}$$

which bears a strong resemblance to Eq. 2. Indeed, it is the physical *basis* for Eq. 2. The ratio of concentrations in a shared volume is just the ratio molecules in each state, which is the ratio of probabilities of finding molecules in either state, which is also the ratio of partition functions for each state:

$$\frac{Z_A}{Z_B} = \frac{Z_A/Z}{Z_B/Z} = \frac{p_A}{p_B} = \frac{[A]}{[B]}$$

So for protein folding, with $Z_{\text{folded}} =: Z_f$ and $Z_{\text{unfolded}} =: Z_u$, we can rewrite the free energy of unfolding as

$$\Delta G_u = -RT\ln\frac{Z_u}{Z_f} = -RT\ln\left(\frac{\sum_{i\in \text{u}} e^{-E_i/RT}}{\sum_{i\in \text{f}} e^{-E_i/RT}}\right) \tag{11}$$

Eq. 11 connects the macroscopic protein stability ΔG_u to the individual energies of all available protein conformations E_i.

Determinants of protein stability

Christian Anfinsen, who won the Nobel prize for his foundational work on protein folding, proposed the **thermodynamic hypothesis**: that a protein folds into a unique minimum-energy structure dictated by its amino-acid sequence. This hypothesis has proven quite useful, despite the exceptions which have kept it from

becoming law (see sidebar). Moreover, by connecting a precise microstate (the lowest-energy conformation) to a biologically relevant macroscopic state (the folded state), Anfinsen's hypothesis provides an elegant launching point for the analysis of the determinants of protein stability.

Exceptions to the thermodynamic hypothesis include:

(1) Kinetically stable proteins, such as subtilisin, whose route to the lowest-energy state pauses at a long-lived intermediate conformation which is biologically active. If the time to leave such a kinetically paused state is much longer than the protein's half-life, the existence of a lower-energy state becomes biologically irrelevant.
(2) Chaperone-dependent proteins, such as many transcription factors, which, when folding in isolation, become trapped in an intermediate inactive state. Molecular **chaperones** interact with such proteins to bias their folding away from such traps, speeding the attainment of the biologically active (usually lowest-energy) conformation.
(3) Interaction-dependent native states, such as aggregates and **prions**. Even if a protein's lowest-energy state in isolation is the folded, biologically active state, the lowest-energy state amidst a collection of other proteins may be one involving interactions with those proteins. The most common example is protein aggregation.

Following Anfinsen, let us consider the case where the folded state corresponds to a single conformation—a unique microstate—$i = f$ with energy E_f lower than that of any other conformation. Assume, for starters, that the unfolded state similarly corresponds to a single conformation with $E_u > E_f$. (Note that the absolute energies E_i cannot be negative—energy is a physical quantity like mass, volume or temperature, and except in exotic quantum circumstances can be thought of as a substance which is either present in some amount ($E > 0$) or not present ($E = 0$).) In this case, Z_f is simple:

$$Z_f = \sum_{i \in \text{folded}} e^{-E_i/RT} = e^{-E_f/RT}$$

and G_f is even simpler:

$$G_f = -RT \ln Z_f + pV = -RT \ln e^{-E_f/RT} + pV = E_f + pV.$$

The same expressions follow for G_u. Then the free energy of unfolding for this two-microstate system is:

$$\Delta G_u = G_u - G_f = E_u - E_f$$

The stability of the protein is given by the difference between the two energies. Note that, for protein stability, the higher the unfolded-state free energy G_u, the more stable the protein, and vice versa.

This means that the absolute energy of the folded state does not matter for the protein's stability. Only the difference in energies matters. Protein stability reflects a competition between the folded and the unfolded state.

Suppose there is a second unfolded conformation with energy $E > 0$; how does the free energy $\Delta G'_u$ compare with the above ΔG_u? We have $Z'_u = \ln\left(e^{-E_u/RT} + e^{-E/RT}\right) > \ln e^{-E_u/RT} = Z_u$ because each term in the sum is positive; therefore

$$G'_u = -RT\ln(e^{-E_u/RT} + e^{-E/RT}) < G_u$$

and thus

$$\Delta G'_u = G'_u - G_f < G_u - G_f = \Delta G_u$$

An additional unfolded state decreases the stability of the native state *no matter what its energy*. In competitive terms, the addition of any additional member of the unfolded "team," whatever their strength, reduces the likelihood that the folded state will win. An immediate consequence is that a protein can be no more stable than the difference between the free energy of the folded state and the energy of the lowest-energy unfolded conformation:

$$\Delta G_u \leq \min\{E_u\} - G_f \approx \min\{E_u\} - E_f$$

where the approximation reflects Anfinsen's hypothesis that the folded state is unique, such that $G_f = E_f$. The difference between the lowest-energy and next-lowest-energy state, $\min\{E_u\} - E_f$, is called the **energy gap**. Differences in stability arising from these energetic differences—reducing the native-state energy by accumulating more energetically favorable interactions, or raising the lowest unfolded-state energy by disrupting such interactions—are said to be due to **enthalpic effects**.

Now suppose that we have M unfolded conformations, each with energy E_u. It is easy to work out that

$$\Delta G_u = G_u - G_f = (E_u - RT\ln M) - E_f$$

showing that the stability is reduced proportional to the logarithm of the number of alternative unfolded conformations M. Because each alternative conformation i has the same energy, each has the same probability at equilibrium, which must be $p_i = \frac{1}{M}$. And as you may recognize, the reduction in stability

$$-RT\ln M = -T\left(-R\ln\frac{1}{M}\right) = -T\left(-R\sum_i^M \frac{1}{M}\ln\frac{1}{M}\right) = -TS_u$$

where S_u is the entropy of the unfolded state. Reductions in stability of this sort—resulting from the multiplicity, rather than the energy, of the members of the unfolded state—are thus called **entropic effects**. Intuitively, the protein spends more time in the unfolded state because there are simply more ways to be unfolded. (Preview: the hydrophobic effect is an entropic effect, but one in which water molecules are the key actors.)

Evolution of protein stability

The stability ΔG_u of most proteins—the difference between the energy of the native state and the free energy of the unfolded state—ranges from 5 to 20 kcal/mol. For comparison, a single hydrogen bond typically contributes 1–3 kcal/mol of energy. Now we can articulate why a single amino-acid substitution which disrupts a

handful of hydrogen bonds is biologically meaningful: it can devastate the folded state, whose stability depends not on its absolute energy (hundreds of hydrogen bond equivalents) but on the difference between that energy and the free energy of the unfolded state (a handful of hydrogen bond equivalents). In energy-gap terms, a disruptive substitution narrows the energy gap between the native and unfolded states, and the proportion of time spent in the folded state shortens exponentially as that gap shrinks.

Assume that selection acts to maintain a certain level of functional protein, and that the folded state is the functional state. A protein with a stability of 5 kcal/mol at room temperature ($RT = 0.6$) spends virtually all its time in the folded state:

$$\Pr(\text{folded}) = \frac{1}{1 + e^{-5/0.6}} = 0.9998$$

and yet disrupting two or three native-state hydrogen bonds can reduce the stability to 0 kcal/mol (which, recall, meansPr(folded) = 0.5), reducing the amount of folded (functional) protein by half.

What about surface and core substitutions? We can begin to sketch in answers. Surface residues make roughly half their interactions with solvent (water), relying mostly on transient hydrogen bonding to make energetic contributions. Core residues, by contrast, are surrounded with interaction partners capable of coevolving stable, energetically favorable interactions. A persistent hydrogen bond in the core contributes ten times the energy of a transient hydrogen bond formed 10% of the time with water. Elaborate stabilizing networks of hydrogen bonds can evolve within a solvent-shielded protein core that become improbable in the seething solvated milieu at the surface. Close packing—closest in the core—enhances van der Waals interaction energy. Salt bridges become more probable in the core simply because they can form from more directions.

All of these effects make the disruption of energetically favorable folded-state interactions by amino-acid substitutions more likely, on average, in the core than on the surface. A more complete picture must wait for the integration of hydrophobic effects.

But if substitutions are so destructive, why are proteins marginally stable? After all, if proteins had a native-state stability of 50 kcal/mol, a few hydrogen bonds would make little difference.

(Stay tuned...)