Protein Folding & Stability

PQB November 3, 2017
James Robertson
james.robertson@stonybrook.edu
Protein folding and stability

Proteins have stable equilibrium conformations

Important to the cell
Denaturation experiment and forces that drive folding
Anfinsen’s hypothesis

Protein structure is determined by amino acid sequence, in a given environment

- bovine pancreatic ribonuclease A
- denatured protein
- re-established native conditions, found protein in native state
- first to show reversible folding
- native structures are stable equilibrium states of matter
The denaturation experiment

- where $x$ is some denaturant: temperature, urea, etc.
- measure fraction native and denatured in each beaker
  - native fraction $f_N(x)$
  - denatured $f_D(x) = 1 - f_N(x)$
Proteins are denatured by increasing temperature or denaturant concentration

- equilibrium constant $K = \frac{f_N}{f_D}$
- folding free energy $\Delta G_{\text{fold}} = -RT\ln K$
Free energies of folding from denaturation curves

(A) and (B) show the fraction of denatured protein as a function of [denaturant] (M) and temperature (°C), respectively. (C) and (D) illustrate the change in free energy (ΔG) as a function of [denaturant] and temperature.
Calorimetry measures heat capacity of protein unfolding with temperature.
Native structures provide insight to driving forces

- compact structures
- hydrophobic cores
- hydrogen bonding is key feature of secondary structures
- well-packed by van der Waals
- salt bridges, electrostatic interactions
- chain entropies
Hydrogen bonding

• prevalent in secondary structures
• backbone carbonyls and amides
  • —N-H - - - - O=C—
• important for stabilizing structure
Electrostatics

- salt bridges
- often at surface

\[ \text{NH}_3^+ \ldots \ldots \text{CO}_2^- \]
van der Waals

- local interactions
- $r^6$ attractions - bring atoms together
- $r^{12}$ repulsions - prevent overlap
- important for packing refinement, filling space
Proteins have hydrophobic cores

- hydrophobic interactions important to protein folding
- proteins are denatured by solvents that weaken hydrophobic interactions
- positive heat capacity of unfolding

Table 3.1: The transfer of nonpolar molecules into water at 23°C is unfavorable (the free energy is positive), dominated by the entropy, and associated with a positive change in heat capacity.

<table>
<thead>
<tr>
<th>Substance</th>
<th>Surface Area (Å²)</th>
<th>ΔG^{hyd} (kcal mol⁻¹)</th>
<th>ΔH^{hyd} (kcal mol⁻¹)</th>
<th>ΔS^{hyd} (cal K⁻¹ mol⁻¹)</th>
<th>ΔC_p^{hyd} (cal K⁻¹ mol⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benzene</td>
<td>240</td>
<td>4.64</td>
<td>0.497</td>
<td>-13.88</td>
<td>53.8</td>
</tr>
<tr>
<td>Toluene</td>
<td>275</td>
<td>5.45</td>
<td>0.413</td>
<td>-16.9</td>
<td>62.9</td>
</tr>
<tr>
<td>Ethylbenzene</td>
<td>291</td>
<td>6.26</td>
<td>0.483</td>
<td>-19.4</td>
<td>76.0</td>
</tr>
<tr>
<td>Cyclohexane</td>
<td>273</td>
<td>6.74</td>
<td>-0.024</td>
<td>-22.7</td>
<td>86.0</td>
</tr>
<tr>
<td>Pentane</td>
<td>272</td>
<td>6.86</td>
<td>-0.478</td>
<td>-24.47</td>
<td>95.6</td>
</tr>
<tr>
<td>Hexane</td>
<td>282</td>
<td>7.77</td>
<td>0.00</td>
<td>-26.08</td>
<td>105.0</td>
</tr>
</tbody>
</table>

Hydrophobicity scales of amino acids
Membrane-spanning regions of membrane protein
The HP model
Use statistical mechanics to describe protein stabilities

- obtain free energy in terms of partition function $Q$
  \[ G = -RT \ln Q \]

- $Q$ is sum of relative statistical weights over all $j$ microstates
  \[ Q = \sum_{j=1}^{s} \omega(\epsilon_j) e^{-\beta \epsilon_j} \]

- compute probability $p_j$ of any state $j$ using Boltzmann distribution law
  \[ p_j = \frac{\omega(\epsilon_j) e^{-\beta \epsilon_j}}{Q} \]

- compute ensemble averages of property $A$
  \[ \langle A \rangle = \sum_{j} A_j p_j \]
HP model overview

- represent protein as string of monomer beads on lattice
- 20 different amino acids reduced to binary code of two monomer types
  - H hydrophobic
  - P polar
- HH contacts are favorable; $\varepsilon_0 < 0$
HP sequence HPPHPHPH

<table>
<thead>
<tr>
<th>microstates</th>
<th>contacts</th>
<th>energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>$\varepsilon_0$</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>$2\varepsilon_0$</td>
</tr>
</tbody>
</table>
HP sequence HPPHPH

• calculate the partition function $Q$ for this HP sequence

$$Q = 28 + 7e^{-\varepsilon_0/RT} + e^{-2\varepsilon_0/RT} = 28 + 7x + x^2$$

• we can also calculate the populations of each state

$$p_N = x^2/Q$$
$$p_I = 7x/Q$$
$$p_D = 28/Q$$

• average energy

$$\langle \epsilon \rangle = 0p_D + \varepsilon_0 p_I + 2\varepsilon_0 p_N.$$ 

• free energy of folding $\Delta G_{\text{fold}} = G_N - G_D$

$$\Delta G_{\text{fold}} = -RT \ln \left( \frac{p_N}{p_D} \right) = 2\varepsilon_0 + RT \ln 28$$

• midpoint of folding transition

$$T_m = \frac{-2\varepsilon_0}{R \ln 28}$$
Heating denatures HPPHPH
Proteins have a folding code

- sequence encodes native structure
- HPPHPH folds to native structure that maximizes number of HH contacts
  - different sequences will have different structures
- HP patterns dominate folding
- HP models describes many aspects of folding
Protein chains collapse, secondary structures are stabilized
Protein folding landscapes are funnel-shaped
Protein folding landscapes are funnel-shaped.
Two-state model
Two-state thermodynamics describe simple protein denaturation

- Assume protein is all N or all D
- Midpoint of transition $f_N = f_D = 1/2$ and $\Delta G = 0$
- Alternative model would have intermediate states
- Folding is driven by intra chain contacts and opposed by chain entropy

\[
\Delta G_{\text{fold}}(N) = N(g + RT \ln z)
\]

- $g < 0$ is free energy of transferring aa from water to hydrophobic core
- Approximate chain entropy; $Q = z^N$

\[
S_D = R \ln Q = NR \ln z.
\]
Protein stability depends linearly on [denaturant]

- denaturants or stabilizers linearly weakens or strengthens interactions
  - denaturant $m_1 > 0$
  - stabilizer $m_1 < 0$
- now model predicts linear dependence of folding free energy on denaturant concentration

$$g(c) = g_0 + m_1 c$$

$$\Delta G_{\text{fold}}(c) = N(RT \ln z + g_0 + m_1 c)$$
Protein stability depends nonlinearly on temperature.

\[ \Delta G_{\text{fold}}(T) = \Delta G_0 + \Delta c_p (T - T_h) - T \Delta c_p \ln \left(\frac{T}{T_s}\right) \]

\[ T_m = \frac{-g(T_m)}{R \ln z}. \]
Proteins tend to unfold in acidic or basic solutions.

- myoglobin
- lysozyme
- RNAse A

Histidine is the only amino acid that has a $pK_a$ in the common range of experiments, so, in some locations in protein structures, it exists in charged form and in others it is uncharged.
Protein folding and stability

- proteins fold to compact native states
  - favorable interactions > chain entropy
- HP model describes aspects of protein folding
- two-state model describes how $\Delta G_{\text{folding}}$ depend on denaturants, and chain length
- proteins denature due to temperature, denaturant, and acidic or basic solutions