Physical Principles in Biology
Biology 3550
Spring 2024

Lecture 34

Protein Folding Thermodynamics and Mechanisms, and Structure Prediction

Monday, 8 April 2024

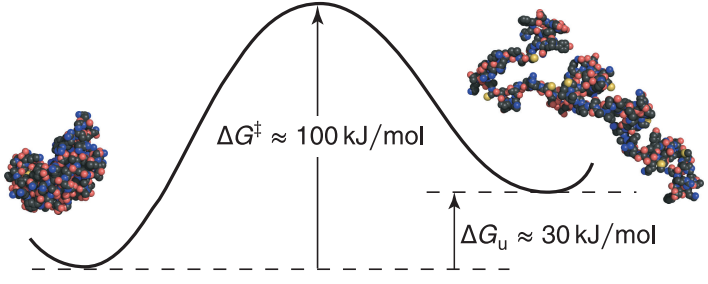
©David P. Goldenberg
University of Utah
goldenberg@biology.utah.edu

Announcements

- Problem Set 5:
  - Due Monday, 15 April at 11:59 PM
  - Submit pdf file on Gradescope
- Quiz 5:
  - Friday, 12 April
  - 25 min, second half of class
- Review Session:
  - 5:15 PM, Thursday, 11 April
  - HEB 2002
  - Come with questions!

Protein Unfolding/Refolding: A Simplified Summary

- Free energy profile for unfolding and refolding:



- What determines the overall equilibrium between native and unfolded states?

Conformational Entropy Change for Protein Unfolding

- From the previous lecture:

$$\Delta S_{\text{conf}} = k \ln 10^n$$

n is the number of amino acid residues. Assumes 1 conformation for the native state and 10 conformations for each residue in the unfolded state.

- On a molar basis for $n = 100$

$$\Delta S_{\text{conf}} = R \ln 10^{100} = 8.314\,\text{J}/(\text{mol}\cdot\text{K}) \times \ln 10^{100}$$

$$= 2 \times 10^3\,\text{J}/(\text{mol}\cdot\text{K})$$

- Corresponding free energy change at 298 K:

$$-T\Delta S_{\text{conf}} = -5.7 \times 10^5\,\text{J/mol} = -570\,\text{kJ/mol}$$

- Compare with the overall free energy change for unfolding, on the order of 30 kJ/mol

Observed Thermodynamics for Protein Folding

For a typical small protein at room temperature (300 K):

- Measured experimentally for unfolding:

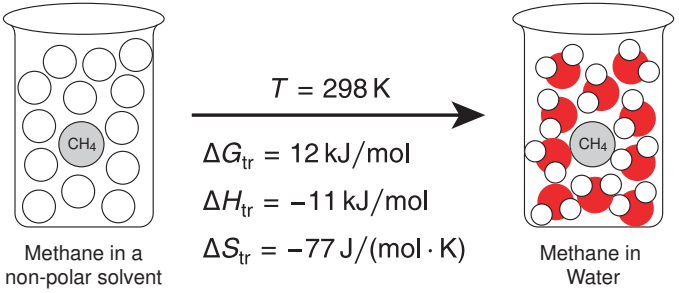
$$\Delta G_{\mathrm{u}} = 30\,\mathrm{kJ/mol}$$
$$\Delta H_{\mathrm{u}} = 100\,\mathrm{kJ/mol}$$
$$\Delta S_{\mathrm{u}} = 230\,\mathrm{J/(mol \cdot K)}$$

- Estimated change in conformational entropy:

$$\Delta S_{\mathrm{conf}} = 2 \times 10^3\,\mathrm{J/(mol \cdot K)}$$
$$-T\Delta S_{\mathrm{conf}} = -570\,\mathrm{kJ/mol}$$

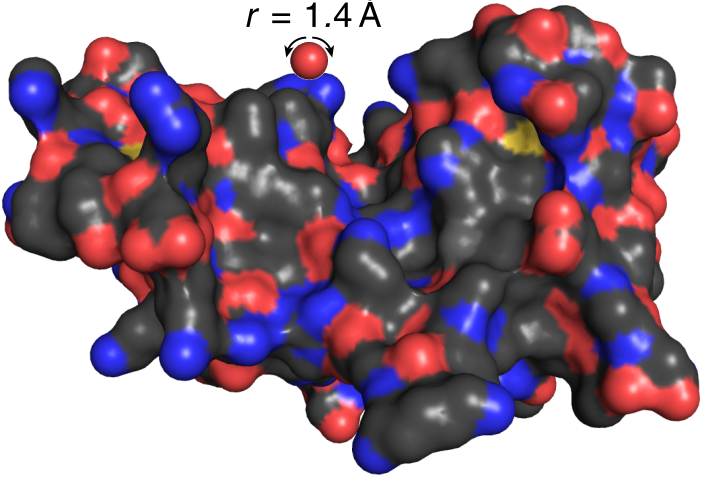
- What we need to explain:
  - Why is $\Delta S_{\mathrm{u}} \ll \Delta S_{\mathrm{conf}}$?
  - Why is $\Delta G_{\mathrm{u}} \gg -T\Delta S_{\mathrm{conf}}$?

Thermodynamics of Transfer of a Non-polar Molecule to Water



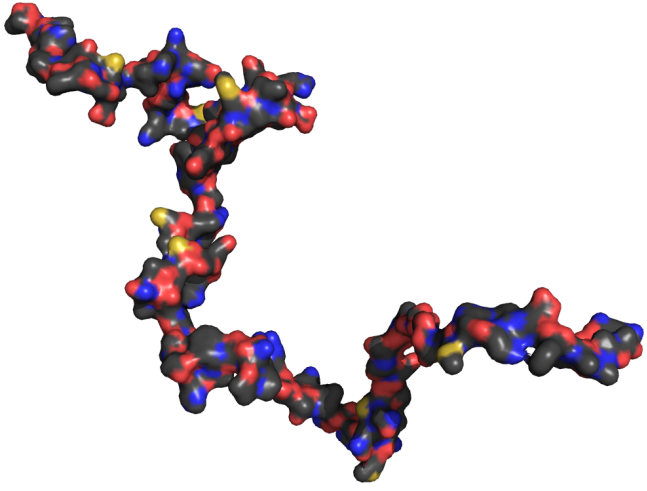
- $\Delta G_{\text{tr}} = \Delta H_{\text{tr}} - T\Delta S_{\text{tr}}$
- ΔG_{tr} is positive because ΔS_{tr} is negative! (an “entropically driven” process).
- Water molecules become more ordered when a non-polar molecule is introduced.
- Non-polar groups buried in the interior of folded proteins become exposed to water on unfolding.

Solvent-accessible Surface of a Small Protein: Ribonuclease A



Solvent-accessible Surface of Unfolded Ribonuclease A

(one representative conformation)



Change in Accessible Surface Area for Unfolding for a Protein of About 100 Residues

	Folded ($Å^2$)	Unfolded ($Å^2$)	Difference ($Å^2$)
Total	7,000	14,700	7,700
Non-polar	3,800	8,800	5,000
Polar	3,200	5,900	2,700

Thermodynamic Consequence of Non-polar Surface Area Exposed Upon Unfolding (Hydrophobic Effect)

For 5000 Å^2 at 300 K

- $\Delta H_{\text{hyd}} = 35\,\text{kJ/mol}$
- $\Delta S_{\text{hyd}} = -1,500\,\text{J/(mol} \cdot \text{K)}$
- $\Delta G_{\text{hyd}} = 480\,\text{kJ/mol}$

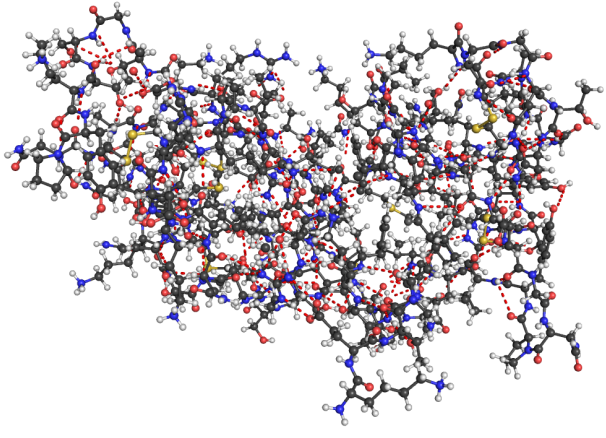
Contributions to Protein Unfolding Thermodynamics

- For protein of 100 amino-acid residues at 300 K:

	ΔH kJ/mol	ΔS J/(mol · K)	ΔG kJ/mol
Conformational entropy		2,000	-570
Hydrophobic effect	35	-1,500	480
Other	65	-270	120
Overall, experimental	100	230	30

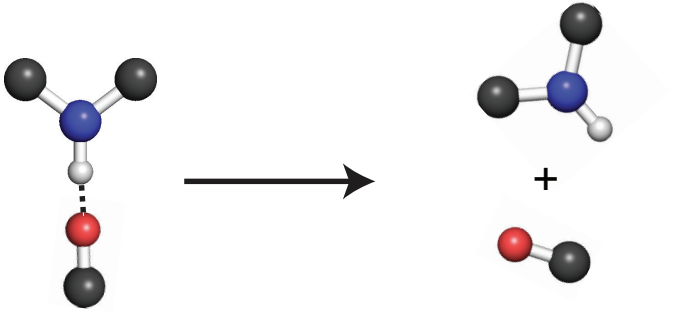
- Increase in conformational entropy is largely compensated for by decrease in water entropy associated with hydrophobic effect.
- What might "other" contributions to ΔH be?
  - Breaking protein hydrogen bonds.
  - Exposure of polar surface area to water.

Hydrogen Bonds in Folded Ribonuclease A



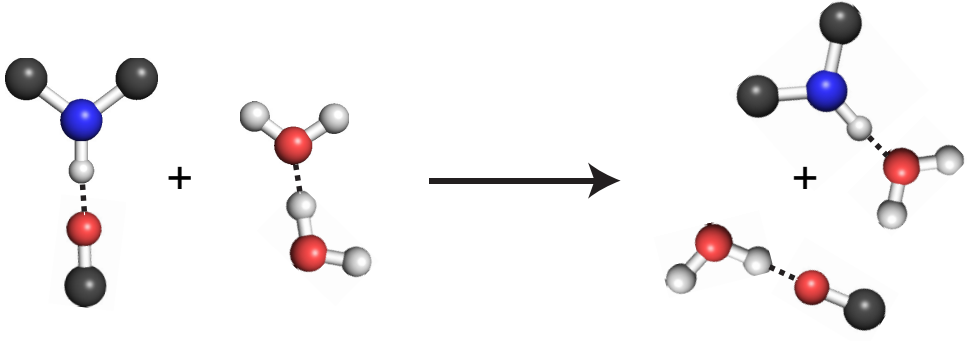
- Red dashes indicate hydrogen bonds.

Breaking a Hydrogen Bond *in vacuo*



- $\Delta H \approx 50\,\mathrm{kJ/mol}$

Breaking a Hydrogen Bond in Water



- $\Delta H = ?$

  $\Delta S = ?$

  $\Delta G = ?$

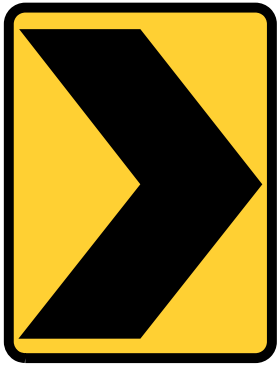
Contributions to Protein Unfolding Thermodynamics

- For protein of 100 amino-acid residues at 300 K:

	ΔH kJ/mol	ΔS J/(mol · K)	ΔG kJ/mol
Conformational entropy		2,000	-570
Hydrophobic effect	35	-1,500	480
Other	65	-270	120
Overall	100	230	30

- Increase in conformational entropy is largely compensated for by decrease in water entropy associated with hydrophobic effect.
- Breaking hydrogen bonds likely represents much of the “other” contributions.

Warning!

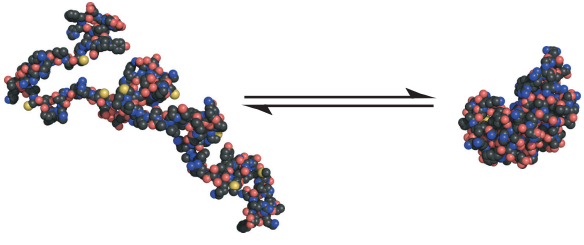


Direction Change

HOW does the folded structure form?

Clicker Question #1

For a protein of 100 amino-acid residues, how long would it take for the chain to randomly sample all of the possible conformations to find the native structure?



A) Less than 1 second

B) $\approx$ 1 minute

C) $\approx$ 1 hour

D) $\approx$ 1 day

E) More than 1 year

All answers count for now.

The Levinthal paradox:

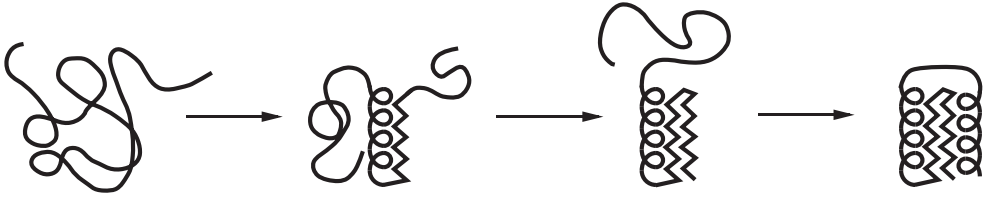
- Consider a polypeptide of 100 amino-acid residues.
- If conformations of individual residues are independent: 10^{100} possible conformations.
- Assume that only 1 in 10^{10} of these conformations is actually possible, because of steric conflicts, leaving 10^{90} conformations.
- The fastest interconversions between conformations is on the order of 10^{-13} s.

$$\begin{aligned}\text{time} &= 10^{90}\,\text{conformations} \times 10^{-13}\,\text{s/conformation} = 10^{77}\,\text{s} \\ &= 10^{77}\,\text{s} \div 3600\,\text{s/hr} \div 24\,\text{h/day} \div 365\,\text{days/year} \\ &\approx 10^{70}\,\text{years}\end{aligned}$$

- How does a polypeptide find it's folded conformation in seconds or minutes?

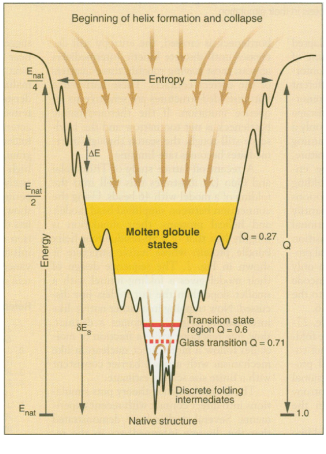
Levinthal, C. (1969). How to fold graciously. In *Mossbauer Spectroscopy in Biological Systems* (DeBrunner, J. & Munck, E., eds.), pp. 22–24. Univ. of Illinois Press, Urbana, IL.
http://www.cc.gatech.edu/~turk/bio_sim/articles/proteins_levinthal_1969.pdf

Protein Folding as a Pathway



- Folding begins with a “nucleus” of local structure.
- Additional structure adds and increases stability.
- Rate-limiting step (transition state) might occur early or late in the pathway.

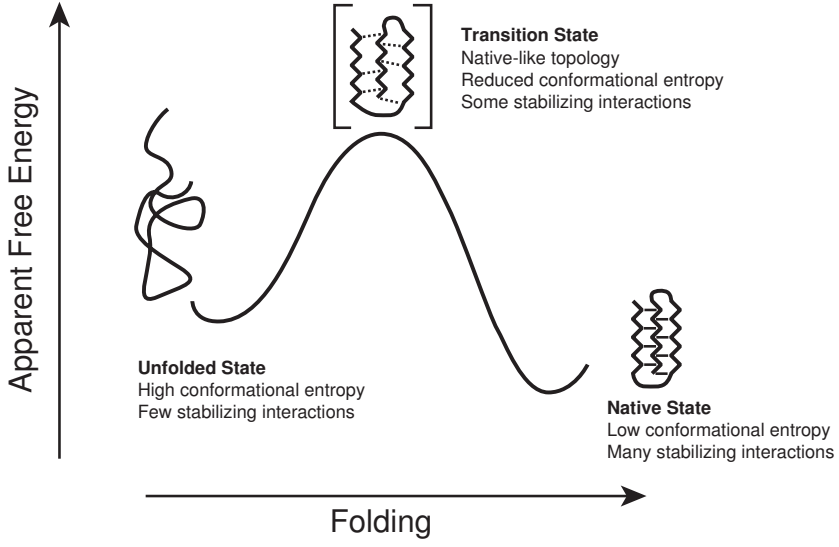
Protein Folding as a Funnel



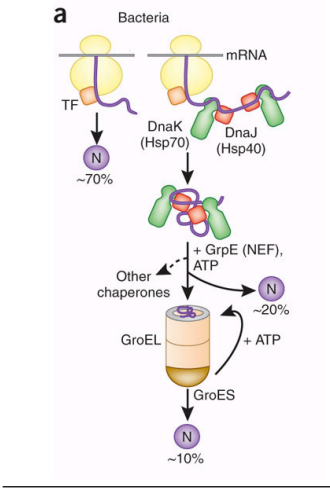
- Folding is viewed as a convergence of many possible starting conformations.
- Top of funnel represents unfolded state.
- Bottom of funnel represents native state.
- Width of funnel represents number of conformations (S_{conf}).
- Distance from top to bottom represents number of stabilizing interactions.

Bryngelson, J. D., Onuchic, J. N., Socci, N. D. & Wolynes, P. G. (1995). Funnels, pathways, and the energy landscape of protein folding: A synthesis. *Proteins*, 21, 167–195. http://dx.doi.org/10.1002/prot.340210302

A Plausible Picture of the Transition State for Protein Folding



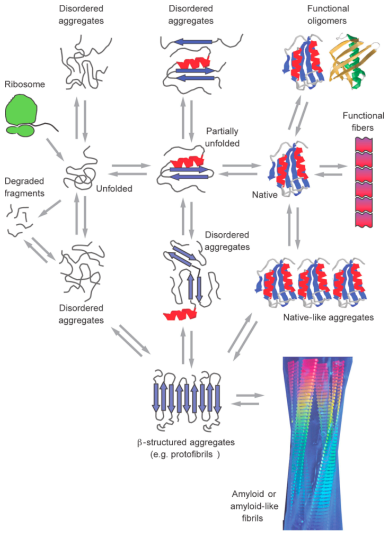
Protein Folding *in vivo*



- Polypeptides are synthesized on ribosomes, starting with the N-terminus.
- Folding may begin on ribosomes.
- Molecular chaperones (Hsp70 and Hsp40) may *limit* folding before synthesis is complete.
- Other chaperones (GrpE and GroE) facilitate correct folding after synthesis.
- Chaperones have a largely negative role: preventing improper interactions.
- Some chaperones are ATP-driven machines that modify structures.

Hartl, F. U. & Hayer-Hartl, M. (2009). Converging concepts of protein folding *in vitro* and *in vivo*. *Nat. Struct. Mol. Biol.*, 16, 574–581. http://dx.doi.org/10.1038/nsmb.1591

Good Pathways and Bad



- Proteins are inherently “sticky”.
- Many folded proteins assemble into functional oligomers and fibers.
- Unfolded or partially folded proteins are especially sticky.
- Unfolded or partially folded proteins tend to form aggregates or abnormal fibers (amyloids).
- Many diseases are associated with amyloid fibers. (Parkinson’s disease, Alzheimer’s disease, prion diseases).

Chiti, F. & Dobson, C. M. (2006). Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.*, 75, 333–366. http://dx.doi.org/10.1146/annurev.biochem.75.101304.123901

Some Approaches to Predicting Protein Structures

- Hierarchical approach:
  - Determine propensities of different amino acids to form α-helices and β-strands.
  - Use propensities to predict segments of polypeptide chain that will form α-helices and β-strands.
  - Assemble secondary-structure elements into overall fold.
  - Doesn't really work!
- Template-based modeling:
  - Identify a protein with a sequence very similar to the protein of interest, and with a known three-dimensional structure.
  - Adjust the known structure to accommodate the sequence of the protein of interest.
  - Works pretty well when the template structure is 50% or more identical to the unknown structure, but accuracy is limited.

Some Approaches to Predicting Protein Structures

- Physics-based modeling:
  - Build a computer model of the polypeptide chain, in arbitrary conformation.
  - Apply mathematical functions that describe all of the forces acting on individual atoms.
  - Simulate process of sampling conformations to find those with minimum energies.
  - Provides information about the folding mechanism AND predicts structure!
  - Now feasible with very small proteins, but with high computational cost.

Lindorff-Larsen, K., Piana, S., Dror, R. O. & Shaw, D. E. (2011). How fast-folding proteins fold. *Science*, 334, 517–520. http://dx.doi.org/10.1126/science.1208351