Lecture 3: Contributions to protein stability

Part I. Review of forces

- Covalent bonds
- Non-covalent Interactions
  - Van der Waals Interactions
  - Electrostatics
  - Hydrogen bonds
  - Hydrophobic Interactions

Part II. Review of key concepts from Stat. Mech.

Part III. Contributions to protein stability and binding

Last time, considered case where there were two alternative states. What happens if there are many alternative states?

We know for each state:

Prob $\propto$ exp [S_{tot}/k] $\propto$ exp [-G/kT]

The probabilities of all of the possible states must add up to one. This will be true if we write

Prob(state) = exp [-G(state)/kT]/ Z

where Z is the sum of the exp [G/kT] terms for each of the states

Z is called the partition function

Concrete example

- Suppose a protein has three states:
  - Unfolded state. 1000 different conformations, each with no interactions
  - Intermediate state. 10 different conformations, each with two hydrogen bonds worth 1kcal/M
  - Native state. 1 conformation, with twenty hydrogen bonds worth 1kcal/M

What is the fraction of the protein which is folded as a function of temperature?

Putting everything together: Contributions to protein stability

General Considerations

- Since proteins are surrounded by water molecules, the energetics of protein folding and binding involve considerable tradeoffs between loss of protein-water interactions and gain of protein-protein interactions. For example, van der Waals interactions are gained between protein atoms and lost between protein and water, and similarly, hydrogen bonds formed between donors and acceptors within a protein chain come at the cost of breaking hydrogen bonds between these atoms and water.

- Because of these tradeoffs, the contribution of van der Waals interactions and hydrogen bonding to protein stability is relatively small. However, these interactions do have a very important "negative" influence on protein structures: van der Waals interactions and hydrogen bonds made with water in unfolded or unbound protein chains that are lost during folding or complex formation must be compensated by interactions within the protein or within the complex, or the free energy of folding/ complex formation will have large unfavorable contributions from the lost interactions with water. For this reason, protein structures rarely contain large cavities (which would involve a loss of van der waals interactions) or buried hydrogen bond donors or acceptors that do not make hydrogen bonds.

General Considerations (ctd.)

- In protein folding and binding reactions, the amount of non-polar surface exposed to water may change considerably, and thus the hydrophobic effect plays an important role in these processes. Unfolding, which involves exposure of hydrophobic side chains to water, may be viewed as equivalent to the transfer of these non-polar groups from a non-polar solvent to water.

- The requirements of retaining hydrogen bonding and attractive van der Waals interactions while minimizing the exposure of nonpolar atoms to water give rise to the hallmarks of globular protein structures: hydrophobic cores with few charged or polar atoms that are well packed to avoid loss of van der waals interactions, largely polar surfaces, and extensive alpha helix and beta sheet secondary structure which allow the polypeptide backbone to retain extensive hydrogen bonding while passing back and forth through the protein.

Entropy Loss and the Denatured State

- As noted in the previous lecture, the hydrophobic effect is probably the main driving force for protein folding: there is a large decrease in the amount of exposed hydrophobic surface area during folding. Opposing this large favorable contribution is an almost equally large unfavorable contribution from chain entropy loss. Recall that the free energy change

$$\Delta G = \Delta H - T \Delta S$$

  and that the entropy change ΔS is proportional to the logarithm of the change in the number of accessible states. The native state has much lower entropy than the denatured state because the backbone is relatively fixed and most of the side chains adopt single conformations. There are also contributions from the change in entropy of the solvent when the non-polar groups of the protein become buried in the core of the protein.
- The net ΔS for folding is negative (a reduction in the number of states), and thus *at high temperature proteins unfold* (ΔG for folding becomes positive).
- Denaturants denature proteins by reducing the strength of the hydrophobic effect. (see following table of transfer free energies of the amino acids from water to denaturants)

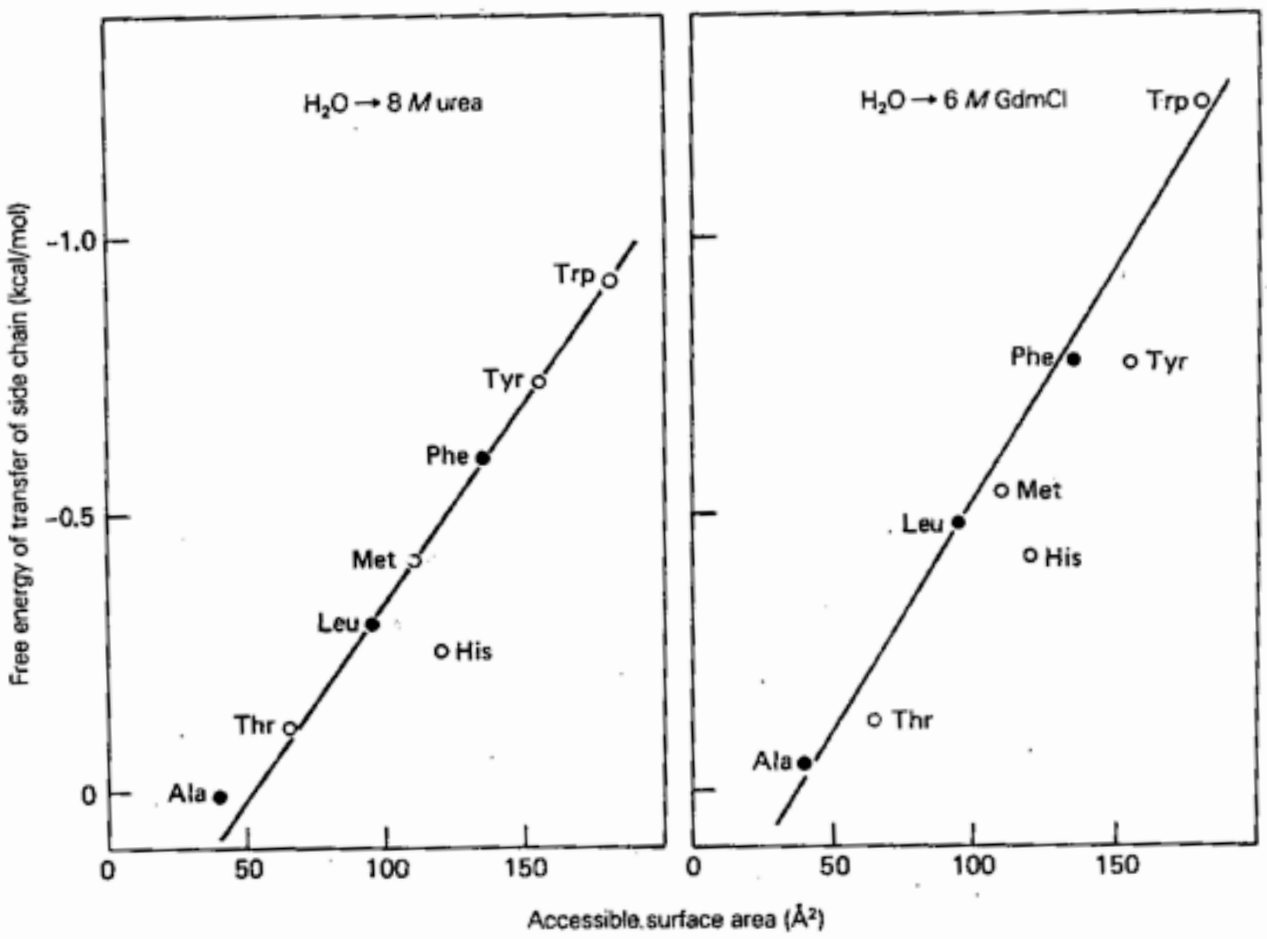


FIGURE 7.16

The denaturants urea and guanidinium chloride (GdmCl) increase the solubilities of both polar and nonpolar amino acid side chains, as measured by the free energy of transfer from water to either denaturant solution (Y. Nozaki and C. Tanford, *J. Biol. Chem.* 238:4074–4081, 1963; 245:1648–1652, 1970). There is a linear correlation of this effect with their accessible surface areas (Table 4.4), although the curves do not extrapolate through the origin. The solid lines have slopes of 7.1 and 8.3 cal/(mol · Å²) for 8 *M* urea and 6 *M* GdmCl, respectively. Residues indicated by open circles have polar groups on side chains. (From T. E. Creighton, *J. Mol. Biol.* 129:235–264, 1979.)

Experimental Data

- How well do these principles account for observed experimental data on protein stability and binding? Unfortunately, because the free energy of protein folding is the difference between two very large contributions: the large chain entropy loss upon folding and the large gain in hydrophobic interactions, it is not currently possible to predict overall protein stability even from high resolution crystal structures.

- Much more amenable to analysis are the changes in protein stability brought about by single amino acid changes. Studies of the effects of a large number of such sequence changes have led to the general conclusions listed on the next slide. It should be kept in mind, however, that while something may be inferred about effects of the mutations in the native state if a high resolution structure is available, there is considerable uncertainty about the effects of mutations on the denatured state.

$$\Delta G = G_{native} - G_{denatured}$$

Given these considerations, what do you expect would be the effect of mutations which

1) Change a non polar residue to a charged residue in the protein core
2) Change a non polar residue to a charged residue on the protein surface
3) Introduce a flexible residue in the protein core
4) Introduce a flexible residue on the protein surface
5) Change a non polar residue to a charged residue at a protein interface
6) Introduce a flexible residue at a protein interface
7) Increase the size of a residue in the protein core
8) Increase the size of a residue on the protein surface

Effect on stability, cont'd

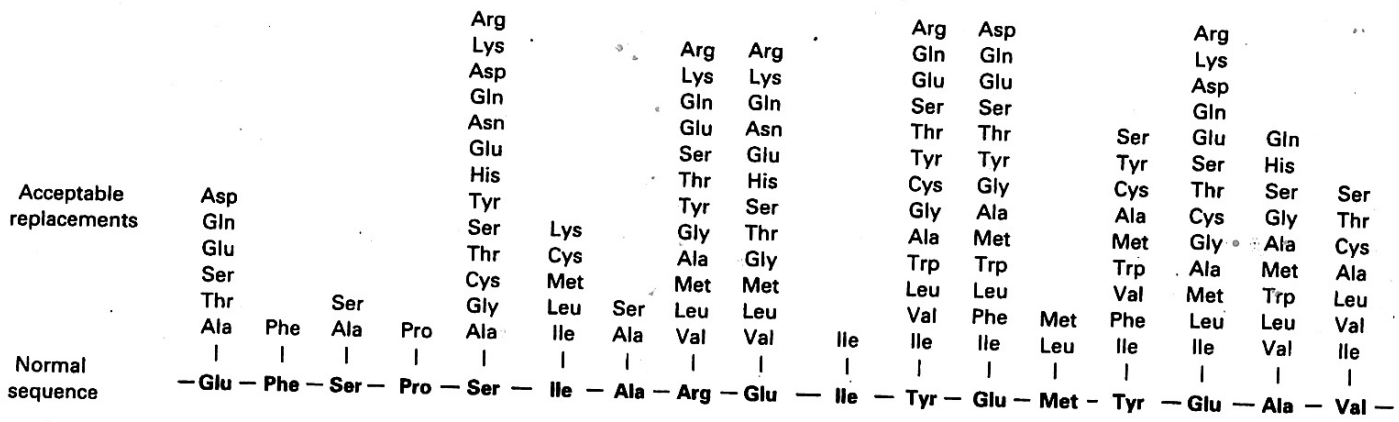
9) Introduction of a disulfide bond between two residues close in the three dimensional structure and close along the amino acid sequence? Or between residues far along the amino acid sequence?

10) How would you design a protein interface that was stable at pH 7 but not pH 5? Stable at pH 5 but not pH 7?

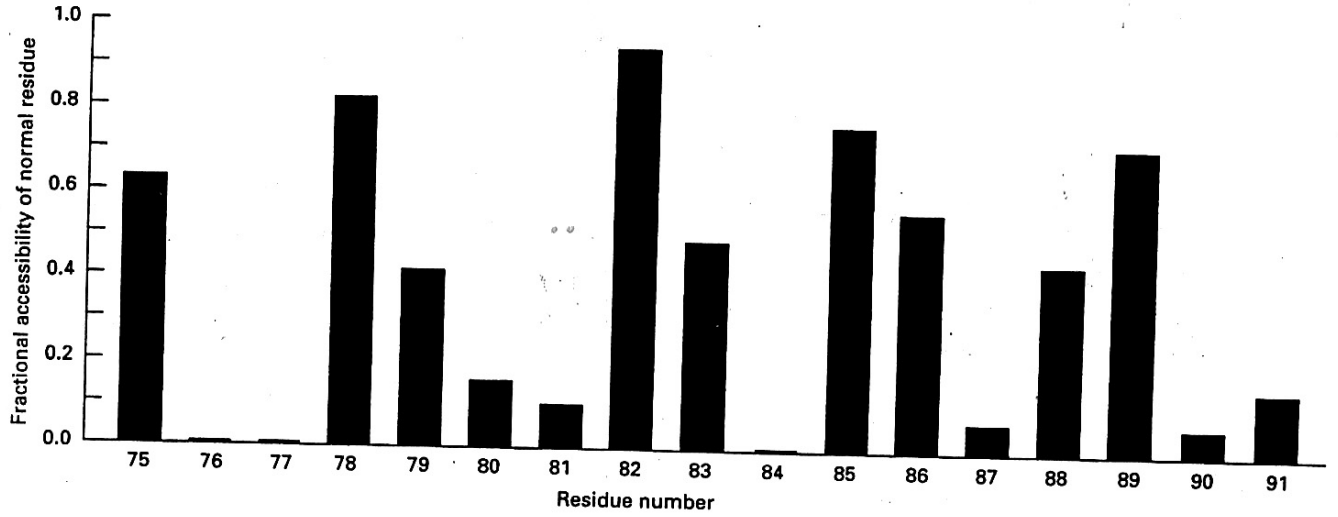
11) Is a completely non-polar environment favorable for a lysine? For a phenylalanine? For a tryptophan?

Conclusions From Studies of Protein Stability

(1) Sequence changes at buried sites almost always have much larger effects on stability than sequence changes at exposed sites. The small change at exposed sites is not surprising given that these residues are likely to have similar environments (ie, largely solvated) in both the denatured and native states. (See figure on λ repressor on next slide)



A



B

FIGURE 7.24

Amino acid substitutions in an α-helical segment that do not substantially lower the net stability of the N-terminal domain of the λ repressor (J. U. Bowie et al., *Science* 247:1306–1310, 1990). *A:* The normal sequence is shown in boldface, and the permitted amino acids are listed above the normal sequence in order of their relative hydrophobicities. *B:* The fractional accessibility of the normal side chain in the folded λ repressor. The most accessible residues generally are most tolerant of amino acid substitutions, although exceptions do occur, such as Ala 90. (Kindly provided by R. Sauer.)

Conclusions From Studies of Protein Stability (ctd.)

(2) Charged residues and to a lesser extent polar residues are disfavored at buried sites. This is expected given the large energetic cost of burying a charge.

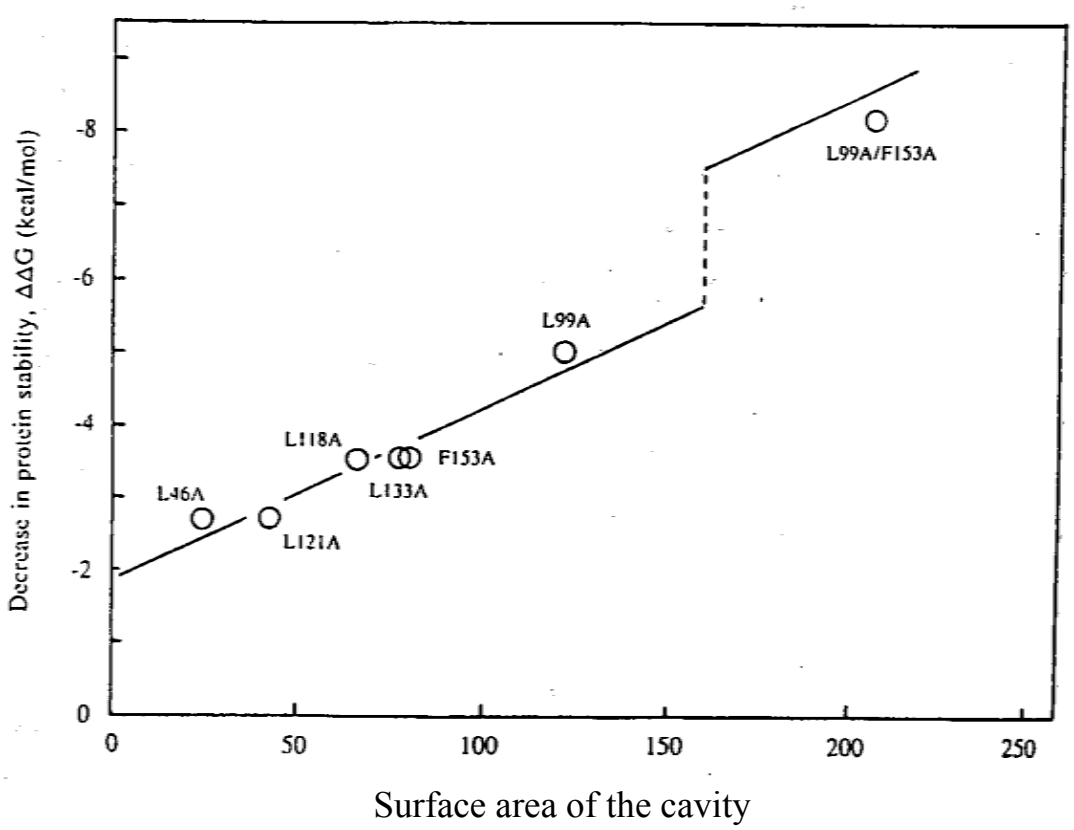
(3) Sequence changes which reduce the amount of hydrophobic burial are destabilizing.
(see table on next slide)

Conclusions From Studies of Protein Stability (ctd.)

Substitution	$\Delta\Delta G$ (kcal/mol) average
Ile -> Val	1.3+/- 0.4
Ile -> Ala	3.8 +/- 0.7
Leu -> Ala	3.5 +/- 1.1
Val -> Ala	2.5 +/- 0.9
-CH2-	1.2 +/- 0.9
Met -> Ala	3.0 +/- 0.9
Phe -> Ala	3.8 +/- 0.3

Conclusions From Studies of Protein Stability (ctd.)

(4) Sequence changes which disrupt side chain packing in the interior or leave large cavities are unfavorable.



(5) Salt bridges between oppositely charged residues on the protein surface contribute relatively little to stability, probably because the more favorable electrostatic interactions are offset by the entropic cost of ordering the sidechains. Repulsive interactions between same charged residues on the surface can be quite destabilizing.

Conclusions From Studies of Protein Stability (ctd.)

(6) Interactions between negatively charged residues at the N termini of alpha helices with the helix dipole formed by the lining up of all of the dipoles in the individual peptide bonds are stabilizing (this was first noted by one of the more senior lecturers in this course!)

Table 4 Substitutions designed to interact with α-helix dipoles in T4 lysozyme[a]

α-helix	Position in helix	Primary mutation	Stability of primary mutant relative to wild-type: pH 2.0 ΔΔG (kcal/mol)	Stability of primary mutant relative to wild-type: pH ~ 6.7 ΔΔG (kcal/mol)	Control mutation	Stability of control relative to wild-type: pH ~ 6.7 ΔΔG (kcal/mol)
39–50	N-cap	Ser38 → Asp	−0.1	0.6	Ser38 → Asn	0.0
108–113	N2	Thr109 → Asp	−0.3	0.6	Thr109 → Asn	0.1
115–123	N1	Thr 115 → Glu	−0.5	0.3	—	—
115–123	N2	Asn116 → Asp	−0.1	0.6	—	—
143–155	N2	Asn144 → Asp	−0.1	0.5	Asn144 → His	0.3

A method for predicting the effect of mutations on stability of proteins/protein interfaces is available on the web: http://robetta.bakerlab.org/ *(Interface Alanine Scanning)*