

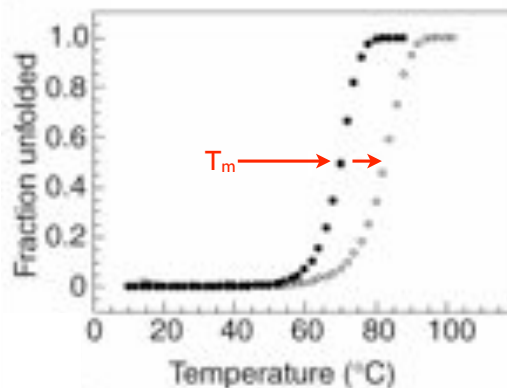
Protein Stability and Folding

Stability of the Folded Conformation

- Folded conformation is only marginally stable and can be disrupted by changes in environment (heat, pH, increased pressure or addition of denaturants).
- Protein denaturation need not involve changes in covalent structure and is usually reversible.
- As the environment changes towards denaturing conditions, initially the structure of small single-domain proteins changes very little.
- This abrupt unfolding is indicative of a very cooperative transition.
- The unfolding of most small single-domain proteins is reversible and equilibrium can be attained.
- Many methods available for visualizing/monitoring unfolding.

Protein Folding: a Two-state Phenomenon

- Protein folding/unfolding is a two-state phenomenon with only fully folded (N) and fully unfolded (U) protein states being present.
- For a two-state transition, the equilibrium constant between N and U can be measured directly from the average fraction of unfolding (α) in the transition region.
- The value of K_{eq} can be determined when α is significantly different from 1 or 0 (in the transition region).
- Allows calculation of ΔG under the set conditions (difference in free energy between U and N states).
- van't Hoff analysis: uses temperature dependence of K_{eq} to estimate ΔH and ΔS .



$$K_{eq} = \frac{[N]}{[U]} = \frac{1 - \alpha}{\alpha} \quad \leftarrow \text{fraction of unfolded protein}$$

$$\Delta G = G_N - G_U = -RT \ln K_{eq}$$

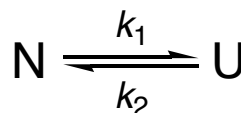
$$\Delta G = \Delta H - T\Delta S$$

$$\Delta G(T_m) = 0 = \Delta H_m - T_m \Delta S_m$$

Protein Folding: a Two-state Phenomenon

Chemical Denaturation:

- Chemical denaturants such as urea, guanidinium chloride or guanidinium thiocyanate can also be used to determine thermodynamic parameters for a protein.
- Chemical denaturants effectively increase the solubility of hydrophobic side-chains, decreasing hydrophobic contribution to stability of the folded protein.
- As the denaturant concentration [denat] increases, K_{eq} shifts towards unfolded.
- ΔG under normal conditions (RT in the absence of denaturant) can be estimated by extrapolation (ΔG^0). [usually between -5 to -10 kcal/mol]
- Parameter m reflects the dependence of ΔG on the denaturant concentration.
- m is dependent on the the denaturant in question.

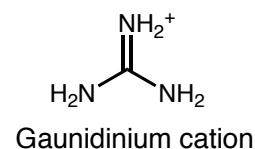
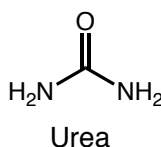


$$K_{eq} = \frac{[N]}{[U]} = \frac{1 - \alpha}{\alpha} \quad \leftarrow \text{fraction of unfolded protein}$$

$$\Delta G = G_N - G_U = -RT \ln K_{eq}$$

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

$$\Delta G = \Delta G^0 + m[\text{denat}]$$



The Unfolded State

- Many proteins under strongly denaturing conditions have been shown to have properties consistent with random coil conformations.
- If interactions between different parts of the polypeptide are preferred over interactions with solvent, then the chain tends to be more compact and less disordered than expected for a random coil.
- While the physical properties of unfolded states produced under different unfolding conditions may differ, they are energetically indistinguishable.
- Difficult to characterize the unfolded state of a protein because many conformations are possible and may be populated.
- The molten-globule state: under certain conditions proteins have been known to demonstrate properties consistent with a molten globule state.

Thermodynamics of Unfolding

- The temperature at which different proteins unfold can vary enormously.
- Most proteins denature at elevated temperatures.
- Thermal denaturation is of intrinsic thermodynamic importance.
- Thermodynamic studies have mostly focused on the two-state transitions of single domain proteins because of their relative simplicity.

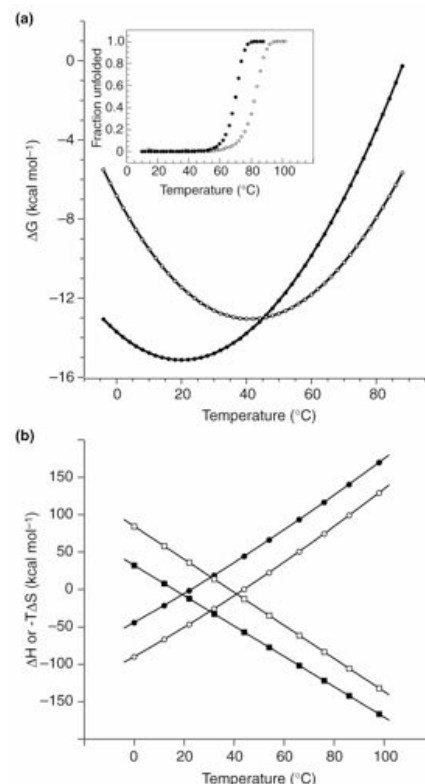


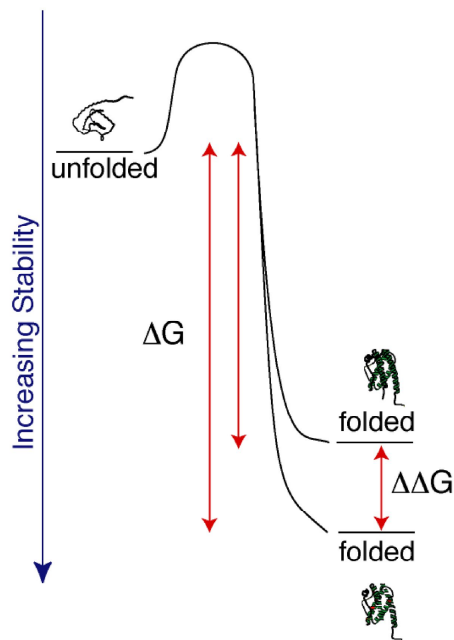
Figure 2. Thermodynamic Comparison of Rop and Ala₁₁le₂-6
(a) Thermal stability profile (ΔG versus T) and representative thermal denaturations (inset) of Rop (solid circles) and Ala₁₁le₂-6 (open circles).
(b) Calculated values of ΔH (squares) and $-T\Delta S$ (circles) as a function of temperature for Rop (solid) and Ala₁₁le₂-6 (open).

Physical Interactions

- The unique properties of proteins is inextricably linked to the complex three-dimensional folded conformations they assume.
- The three-dimensional folded conformation is the result of many simultaneous noncovalent interactions between different parts of the protein and with the environment.
- These interactions are the result of a limited set of fundamental noncovalent forces.
- The complexity of water and an aqueous environment limits our understanding of proteins.

Protein Stability

- Native proteins are only marginally stable under physiological conditions.
- The free energy of denaturation is only ~ 0.1 kcal/mol for each amino amino acid residue.
- The three-dimensional folded conformation arises through a delicate balance of stabilizing and destabilizing forces.
- The observed stability of the folded protein is the result of a very small difference between very large but compensating factors.
- The enthalpic and entropic contributions vary similarly and compensate each other. This results in the free energy being relatively small difference between the two.



Forces Stabilizing Macromolecular Structure

- Noncovalent interactions are key biological forces:
 - ★ Electrostatic Forces:
 - Ionic interactions
 - Van der Waals
 - Hydrogen bonding
 - ★ Hydrophobic interactions
- These forces are transient in nature.
- Individually all are weak (C-C bond ~80 Kcal/mol), but they add up and collectively can be very strong.

Short-Range Repulsions

- Repulsion invariably arises as they molecules/atoms become near enough for their respective electron orbitals begin to overlap.
- Repulsion increases enormously because the electrons on the different molecules cannot occupy the same space at the same time (increasing exponentially with the inverse of distance).
- Because repulsion rises so steeply, it is possible to consider molecules/atoms to have definite dimensions with defined volumes (**van der Waals radius**).
- van der Waals radius is based on smallest distance that can exist between two nonbonded atoms in the crystalline state.

Electrostatic Forces: Point Charges

- All intermolecular forces are thought to be essentially electrostatic in origin.
- The most fundamental noncovalent interaction would therefore be the interaction between electrostatic charges.
- Coulomb's law describes the interaction between two point charges in a vacuum.
- It describes an interaction that is effective over relatively long distances.
- For other environments (such as in solution) the electrostatic interaction is modulated by other interactions.
- In homogenous environments, the electrostatic interaction is diminished by the dielectric constant of the medium.
- At short distances, molecules and atoms cannot be treated as point charges.
- Interactions between very close, oppositely charged groups in proteins usually involve not only electrostatic interactions, but also some degree of hydrogen bonding (**salt bridges**).

Coulomb's Law (vacuum)

$$\Delta E = \frac{Z_A Z_B \epsilon^2}{r_{AB}}$$

ϵ = the charge of an electron

Z_A = number of charges on A

Z_B = number of charges on B

r_{AB} = distance between A and B

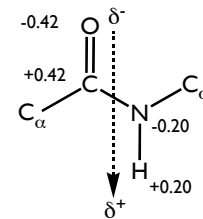
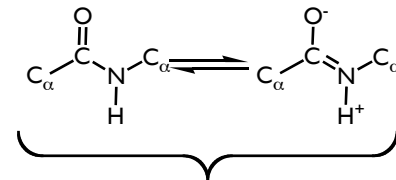
Adjusted for environment

$$\Delta E = \frac{Z_A Z_B \epsilon^2}{D r_{AB}}$$

D = dielectric constant

Electrostatic Forces: van der Waals

- A molecule does not need to have a net charge to participate in electrostatic interactions.
- Electron densities can be localized if covalently linked atoms have different electronegativities.
- The separation of charge in a molecule determines its dipole moment (μ_D), corresponding to the magnitude of the separated charge (Z) and the distance (d) by which it is separated.
- The dipole moment has directionality as well as magnitude.
- The peptide bond which has partial double bond character exemplifies this polarization. The oxygen has a partial negative charge and the -NH- group a partial positive.
- Dipoles interact with point charges, other dipoles and more complex interactions.



$$\mu_D = Zd$$

Dipole Moment

Electronegativities

of common atoms

in proteins:

O = 3.45

N = 2.98

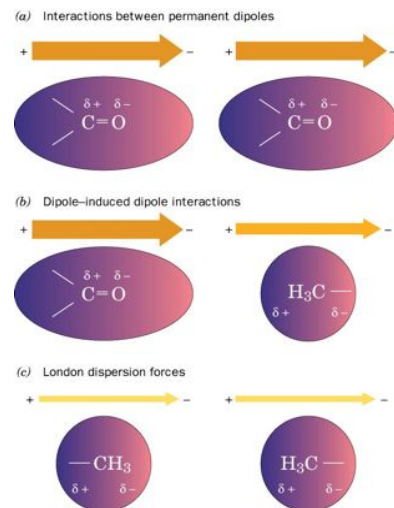
C = 2.55

S = 2.53

H = 2.13

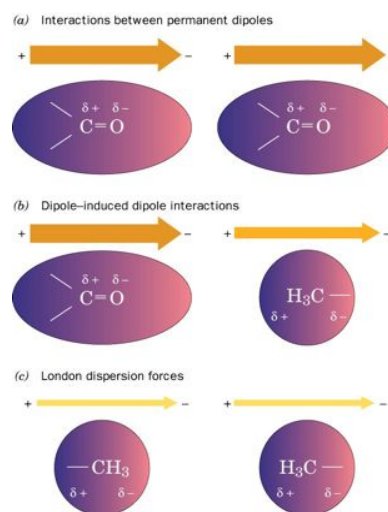
Electrostatic Forces: van der Waals

- Dipolar interactions are weaker than those between ionic groups. This is due to the fact that both attraction and repulsion occur between the two separated charges.
- The strength of dipolar interactions drops much more abruptly with distance than is the case with the interaction with ions (inversely with $\sim d^3$).
- Interactions involving dipoles also effects the dipole charge distribution within the interacting molecules. **Polarizability** describes the disposition of a molecule to have its electron (charge) distribution influenced by an applied electronic field.
- An induced dipole always interacts favorably with the inducing field. However this interaction is only half of that which would have occurred had the dipole already existed.



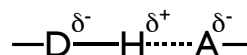
Electrostatic Forces: van der Waals

- All atoms and molecules attract each other even in the absence of charged groups as a result of mutual interactions and induced polarization.
- These are weak interactions only effective at short distances (varying with d^{-6}).
- Can arise from interaction between:
 - Two permanent dipoles.
 - A permanent dipole and an induced dipole.
 - Two induced dipoles (London dispersion forces).
- London dispersion forces: complex interaction. Essentially, an atom or group may have no net dipole, but may have a transient dipole resulting from temporary asymmetry in the distribution of electrons. This transient dipole can similarly polarize nearby neutral atom. (synchronization of electron flow and distribution in neighboring atoms and groups). Interaction becomes insignificant at distances greater than 50Å.
- Van der Waals interactions are often represented by an energy potential as a function of distance (d).
- The optimal distance for the interaction of two atoms is usually 0.3-0.5Å greater than their combined van der Waals radii.



Electrostatic Forces: Hydrogen Bonding

- A hydrogen bond occurs when two electronegative groups compete for the same hydrogen atom.
- In such interactions, the H atom is formally attached to the donor atom via a covalent bond and interacts favorably with the acceptor atom.
- The main component of the hydrogen bond is an electrostatic interaction between the dipole of the covalent D-H bond (H has δ^+) and the δ^- of the acceptor atom.
- The electrostatic and covalent elements of the hydrogen bond make it energetically favorable for the three participating atoms (D,H and A) to be collinear. This is the most common arrangement however some deviation from linearity is observed.
- The lengths and strengths of hydrogen bonds is dependent on the electronegativities of D and A. The greater their electronegativities are, the shorter and stronger the hydrogen bond.

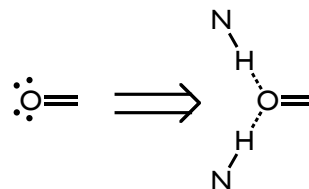


Common
H-donors

N—H
O—H
S—H
C—H

Common
H-acceptors

O=O
—O—
—N=O
—S—
—S—



Hydrophobic Interaction

- Electrostatic, hydrogen bonding and van der Waals interactions between two molecules in an aqueous environment are not particularly favorable due to competing interactions with surrounding water molecules.
- In the case of hydrophobic surfaces, interactions with surrounding water are not as favorable.
- The relative absence of favorable interactions with surrounding water molecules increases the favorability of interactions among nonpolar groups themselves (relative to other solvents).
- The preference of nonpolar molecules and groups for nonpolar environments is known as the hydrophobic interaction.

Hydrophobic Interactions

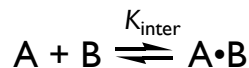
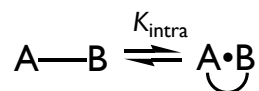
- The magnitude of the hydrophobic interaction is generally measured by the free energy of transfer (ΔG_{tr}) of a nonpolar molecule in the gas, liquid or solid state into water. (positive ΔG_{tr} value indicates that the nonpolar molecule prefers a nonaqueous environment)
- Transferring a solute molecule into a liquid involves:
 - Creating a suitable cavity in the liquid.
 - Introducing the solute molecule into the cavity.
 - Rearranging the solute and liquid molecules to maximize favorable interactions between them.
- The observed thermodynamics of this transition are the net effect of all these factors. Interpretation is not always straightforward.
- At room temperature, the unfavorable transfer of a nonpolar molecule from a nonpolar liquid to water is primarily a result of the unfavorable change in entropy ($\Delta H_{tr} \approx 0$).
- Water molecules cannot form hydrogen bonds with the nonpolar group. Therefore, they are generally believed to satisfy their hydrogen bond potential by forming a hydrogen bonded “iceberg” network among themselves at the nonpolar surface.

Hydrophobic Interactions: to summarize

- Hydrophobic interactions do not result from repulsion between water molecules and nonpolar molecules and surfaces.
- While favorable interactions do occur between water molecules and nonpolar molecules and surfaces, the magnitude of these interactions are less than the favorable van der Waals interactions in a nonpolar environment and the hydrogen bonding in liquid water.
- Hydrophobic interaction results in nonpolar atoms, molecules and groups to interact with each other rather than with water.

Intramolecular Interactions

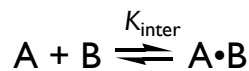
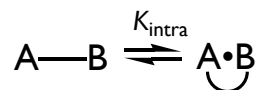
- For molecules to interact with each other they must lose entropy, which is energetically unfavorable.
- The magnitude of the loss in entropy is dependent on the degrees of freedom that become fixed as a result of the interaction.
- In the case of intramolecular interactions, the groups involved are incorporated within the same molecular scaffold, which automatically limits the number of degrees of freedom by fixing the relative distance and orientation of the groups involved.
- Intramolecular and bimolecular interactions can be compared by means of the ratio of their equilibrium constants (the **effective concentration**).



$$\frac{K_{\text{intra}}}{K_{\text{inter}}} = \text{effective concentration of } A-B$$

Intramolecular Interactions

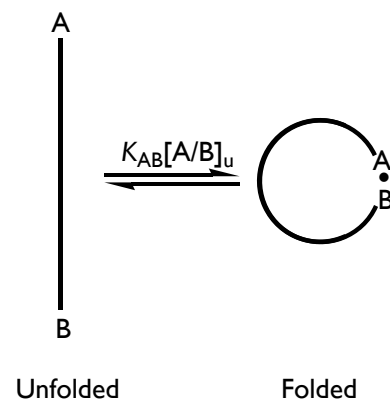
- The maximum effective concentration of two groups in an aqueous solution was believed to be 55M, but much greater values for effective concentration are usually observed.
- Covalently linking the interacting moieties through a bond network results in their concentration relative to each other to be much higher than would be possible were the two groups on separate molecules.
- Interaction between the two groups results in the sacrifice of some fraction of the internal flexibility and conformational freedom of the molecule.
- When there is no entropic difference between molecules with and without interaction between the groups, their effective concentration is at its maximum value.



$$\frac{K_{\text{intra}}}{K_{\text{inter}}} = \text{effective concentration of } A-B$$

Cooperativity of Multiple Interactions

- Multiple groups within a single molecule can behave differently from the same groups in solution.
- The simultaneous presence of multiple interactions within a single molecule results in cooperativity between them. Collectively, these interactions can be much stronger than expected based on their individual strengths.
- **Cooperativity is critical for proteins.**
- Single interactions between groups within a polypeptide chain are not expected to be stable unless these groups lie in close proximity of each other within the covalent structure (resulting in a high effective concentration).
- Due to the size and conformational flexibility of the unfolded protein, groups attached to a moderate sized peptide have effective concentrations in the range of 10^{-2} - 10^{-5} M (depending on proximity).
- Expected values for $K_{obs,u}$ (observed equilibrium constant) for individual hydrogen bonds, salt bridges... etc range from 4×10^{-3} to 10^{-7} .



$$K_{obs,U} = K_{AB}[A/B]_u$$

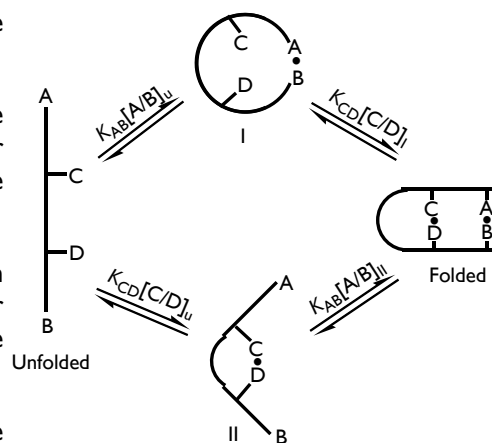
K_{AB} = association constant for free A and B

$[A/B]_u$ = effective concentration of A and B in unfolded peptide

Cooperativity of Multiple Interactions

- Multiple interactions among two or more pairs of groups within the same molecule often do not behave independently, but assist or interfere with each other.
- In the example given: if both interactions $A \cdot B$ and $C \cdot D$ are possible simultaneously, the interactions between one pair of groups constrains the peptide, increasing the effective concentration of the other pair.
- This will proceed in a mutual manner, with both interactions having the same effect on each other (factor Coop is the degree of cooperativity between the interactions).
- Each interaction is more stable in the presence of the other than in its absence.
- In polypeptides containing additional groups that interact simultaneously, equilibria are extended:

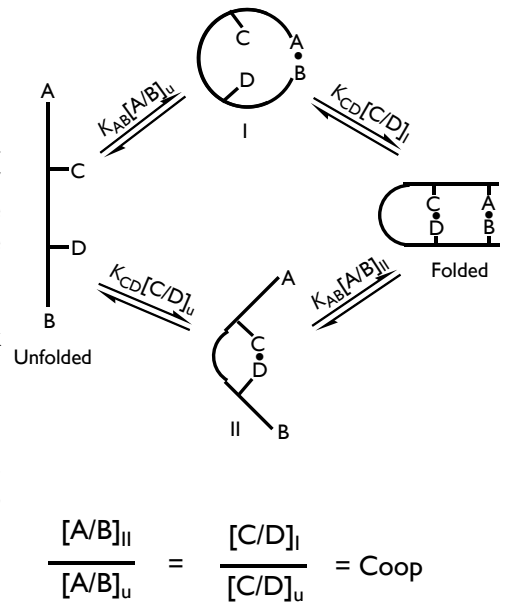
$$K_{net} = (K_{AB}[A/B]_u)(K_{CD}[C/D]_i)(K_{EF}[E/F]_{iii})(K_{GH}[G/H]_{iv}) \dots$$
- The value of K_{net} is pathway independent.
- The final folded conformation is stable-only if the value of K_{net} is greater than unity.



$$\frac{[A/B]_{II}}{[A/B]_u} = \frac{[C/D]_I}{[C/D]_u} = \text{Coop}$$

Cooperativity of Multiple Interactions

- In considering a series of weak interactions, the first will be very weak, with an equilibrium constant of 10^{-3} - 10^{-7} .
- The first interaction increases the effective concentration of the next pair of interacting groups, resulting in a slightly larger equilibrium constant for the interaction (by the factor Coop). If the equilibrium constant of the second interaction is less than unity, the product of the two equilibrium constants is lower than that of the first.
- The net stabilities of conformations with additional weak interactions are even lower than that of the conformation with a single interaction.
- The process continues until the effective concentrations of additional interacting groups are sufficiently high to make the equilibrium constant for each additional interaction greater than unity, with K_{net} increasing with each additional interaction.
- In this manner, a sufficient number of simultaneous weak interactions can make the value of K_{net} greater than unity and provide a stable folded structure.



Mechanism of Protein Folding

Folding Pathways:

- How does a protein fold into its native conformation?
- A protein cannot randomly explore all of the conformational possibilities until it achieves its native conformation.
 - ◆ **Levinthal Paradox:** a 100 residue peptide sampling 10^{13} conformations per second would take 10^{85} sec to fold. (Universe is estimated to ~ 20 billion years or $\sim 6 \times 10^{17}$ seconds old.)
- Therefore, proteins must employ an ordered pathway or set of pathways which ultimately allow the protein to achieve its native fold.
- There is the possibility that the observed folded state may not be the conformation with the lowest possible free energy, but is the most stable of the kinetically accessible conformations.
- Proteins fold to a significant degree within 1 millisecond.

Kinetic Analysis of Complex Reactions

Kinetics of Unfolding

- Protein unfolding is almost always observed to be an all or none process.
- Native protein represents a relatively conformationally homogeneous population, and unfolding generally proceeds with a single kinetic phase and a single rate constant. (no lag phase).

Kinetics of Refolding

- Kinetic complexity is a hallmark of protein folding. Starting with the conformational heterogeneity of the unfolded population.
- Heterogeneity includes *cis-trans* isomerization of peptide bonds (slow process).
- In an unfolded population with the native *cis-trans* isomers, refolding generally occurs with a single rate constant in spite of the conformational heterogeneity of the unfolded state.

Protein Folding Events

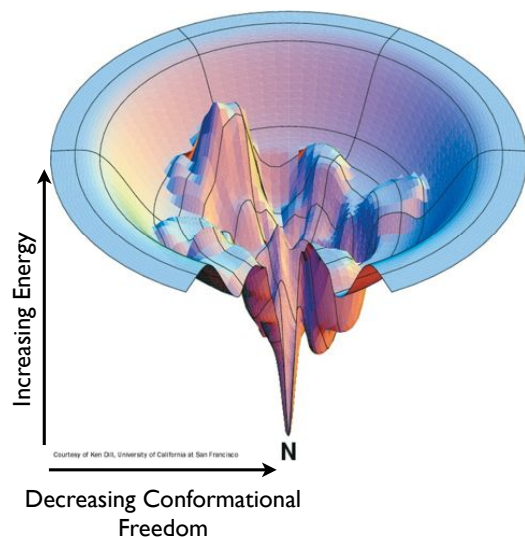
- **Initial folding events (burst phase): [milliseconds]**
 - ❖ For many small-single domain proteins, much of the secondary structure is established.
 - ❖ Much of the driving force attributed to hydrophobic collapse. (hydrophobic groups coalesce and expel water)
 - ❖ Initial collapsed state is molten globular.
 - ❖ Side chains are extensively disordered.
- **Intermediate folding events: [~5-1000 milliseconds]**
 - ❖ Secondary structure stabilizes and native-like tertiary structure appears.
 - ❖ Side chains are still mobile.
- **Final folding events: [\leq several seconds]**
 - ❖ Protein achieves native structure.
 - ❖ Complex motions allow the protein to attain relatively rigid packing and hydrogen bonding.
 - ❖ Remaining interior water molecules are expelled from the core.

Hierarchal Protein Folding

- The folding process begins with the formation of marginally stable local order/ structure.
- These structure elements then interact (locally) to form intermediates of increasing complexity.
- Process continues ultimately yielding the native protein.
- Evidence supporting the premise of hierarchal protein folding:
 - ❖ Many peptide fragments excised from proteins will assume their native conformation.
 - ❖ Observed folding intermediates are consistent with a hierarchal folding process.
 - ❖ Helix boundaries are fixed by their primary sequence, not so much by 3-D interactions.
 - ❖ Secondary structure can be predicted with reasonable accuracy, even when long-range interactions are not accounted for or are suppressed.
- Sequence information defining a specific fold is both distributed throughout the polypeptide chain and is highly overdetermined.

Landscape Theory of Protein Folding

- **Current Thinking:** Protein folding is envisioned to proceed on an energy surface/landscape.
- The landscape represents the conformational energy states available to a polypeptide.
- Polypeptides fold via a series of conformational adjustments that reduce their free energy and entropy until the native folded state is achieved.
- There is no single pathway or closely related set of pathways that a polypeptide must follow in achieving its native conformation.
- Suggests that landscape may include local energy minima and maxima (therefore many possible transient folding intermediates may exist)



Folding of Multidomain and Multimeric Proteins

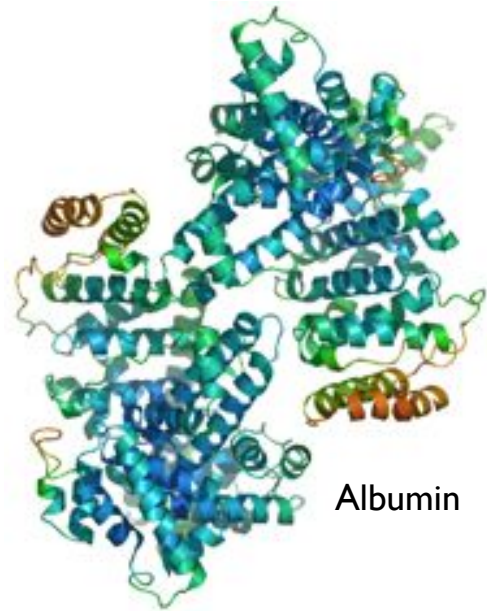
- Large proteins may be composed of multiple **domains** or polypeptide chains.
- Independent domains unfold and refold like single-domain proteins, which can lead to complex unfolding curves for proteins. (in such cases, domains may unfold under different conditions)
- Can also be varying degrees of interaction between the domain. Interactions between domains can effect folding.
- Where the isolated domains are stable, folding of the intact multidomain protein appears to occur by initial folding of the domains, followed by association of the domains.
- Domain association is often the slowest step in the folding process. (domains may not be folded entirely correctly or because small adjustments are required for interaction between the domains.)
- When association is slow step, an intermediate can accumulate where domains are folded but impaired. May lead to intermolecular interactions and precipitation.

Folding of Multidomain and Multimeric Proteins

- Large proteins may be composed of multiple domains or **polypeptide** chains.
- Folding of oligomeric proteins has similar considerations because the polypeptide chains involved often incorporate multiple domains.
- Oligomerization necessitates specific interactions between the polypeptide monomers.
- Polypeptide monomers generally fold to nearly their final conformations before oligomeric association. The specific interactions likely requires that the polypeptide monomers have a folded conformation in order to provide the interaction sites.
- Rate-limiting step may be either intramolecular folding or association of the monomers.
- Final adjustments of the structure appears to have significant energy barrier (it is a relatively slow process)

Flexibility of Protein Structure

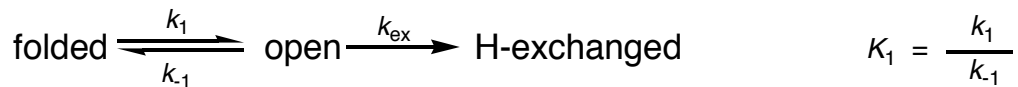
- Protein structures are not static. Both crystal structures and NMR indicate varying degrees of conformational freedom.
- Proteins can be thought of as existing in a range of distinct but closely related microstate conformations that interconvert rapidly at room temperature.
- On a longer time scale, larger backbone conformational movements can occur.
- On the longest time scales, the folded conformation is marginally stable and may transiently sample the unfolded state (10^{-4} - 10^{-12} /s).
- Side chains of residues at the protein surface can have significant conformational freedom.
- Close packing of atoms in the protein interior is constraining and requires coordinated motions.



Conformational Motility

- **Hydrogen Exchange:** Best evidence for extensive structural mobility is that internal groups in proteins react with appropriate reagents in solution. (buried groups either are occasionally at surface or reagent can permeate the protein)
 - Isotopic exchange with water (H_2O , $^2\text{H}_2\text{O}$ and $^3\text{H}_2\text{O}$).
 - Hydrogen atoms covalently attached to various atoms exchange with solvent at different intrinsic rates, depending on tendency of that atom to ionize.
 - Exchange of amide protons most often studied because these hydrogen atoms exchange on a useful time scale.
 - Rates of exchange impacted by temperature, hydrogen bonding, environment and degree of exposure.
 - Rate of exchange of individual H's varies 100-fold.
 - Protons involved in hydrogen bonding in the interior of β -sheets and α -helices tend to exchange least readily.
 - Acid and base catalyzed exchange (via transient protonation of $\text{C}=\text{O}$ and deprotonation of $\text{N}-\text{H}$ respectively).
 - Rates of exchange generally increase at elevated temperatures, but in a complex manner.
 - Classical methods (NMR and MS) only provided insights into average number of protons exchanged.
 - Exchange of individual hydrogen atoms can be followed using ^1H -NMR.

Conformational Motility



Hydrogen exchange continued...

- While proteins may sample unfolded state, this is not likely responsible for the exchange of buried hydrogen atoms. (not all interior hydrogens exchange with same rate)
- Local unfolding or “breathing” is often used to explain the exchange of interior hydrogen atoms. The hypothetical open form is unstable and transient.
- Under most conditions, proteins demonstrate exchange rates consistent with $k_{\text{ex}} < k_{-1}$ and rate = $K_1 k_{\text{ex}}$.
- Alternative explanation involves rare instances of diffusion of solvent into the interior sites in the protein. Supported by exchange in proteins in crystalline state.
- Both require some degree of backbone conformational flexibility.
- Available information suggests that different site in folded proteins likely exchange with solvent by a wide range of different processes, depending on the protein and conditions.

Fluorescent Quenching

- Fluorescence of aromatic groups is instantly quenched by close physical interaction with some small molecules such as O_2 , I^- and acrylamide.
- Aromatic side chains are quenched by diffusion-controlled encounters with such molecules.
- Many internal groups are also quenched, only slightly less efficiently with O_2 .
- Charged and polar quenchers (I^- and acrylamide) are less efficient and likely only act when the protein is in an open conformation.
- Detailed interpretations are complicated by energy transfer between fluorescent groups within the protein, by varied quantum yields and possible localization of quenchers to specific sites in the protein.

Rotations of Side Chains

- Side chains on the protein surface and terminal methyl groups of side chains in the interior tend to have mobilities comparable to those in unfolded proteins (rotating on 10^{-11} - 10^{-8} s time scale).
- Slower motion of interior groups masked by rotation of the entire protein molecule.
- Can study motion of aromatic rings by $^1\text{H-NMR}$.
- Most proteins have Phe and Tyr side chains that give average spectra, suggesting that they rotate on the order of 180° flip $10^4/\text{s}$ even when buried.
- Buried Trp and His residues do not appear to flip.