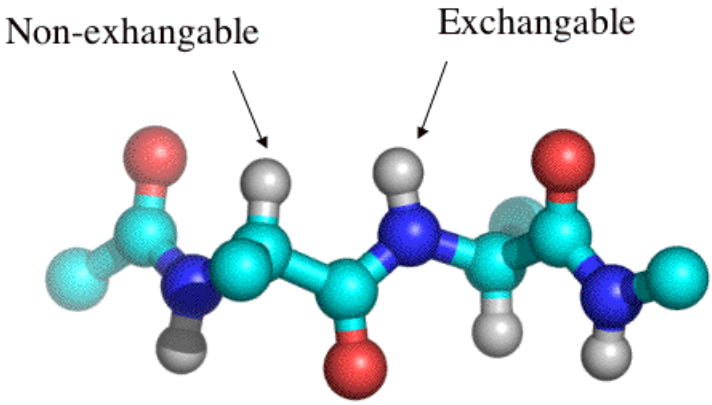
Lecture 4 - Experimental Evidence for Protein Conformational Fluctuations

Molecular simulations attempt to model reality. But do simulations actually accomplish this goal? One of the themes of this course is that you can setup a simulation to do anything you want. But is it garbage or is it useful for interpreting reality? To answer this question one must ask whether or not the simulation agrees with what we already know about the system from experiment. Every molecular simulation provides a variety of metrics which should be compared to experimental data whenever that data exists.

This course is about conformational fluctuations and so we now look at experimental data which bears on this topic. One of the most successful methods for characterizing protein conformational fluctuations is **hydrogen exchange** (frequently abbreviated, HX).

1. Hydrogen Exchange Experiments.
The rate of hydrogen exchange is affected by pH. Hydrogens on backbone amide nitrogens and side chain polar groups are in continual exchange with hydrogens from the solvent. This process occurs on a time scale of ~1 ms. The rates of this intrinsic exchange vary depending on the nature of the neighboring atoms and local geometry. Hydrogens on carbon atoms do not exchange at significant rates. Below is a peptide model illustrating a *non-exchangable* C-H and an *exchangable* backbone N-H. We will ignore side chain hydrogens for this discussion.



At pH values above ~3, hydrogen exchange is catalyzed by OH^- through a process of hydrogen abstraction. At pH 9 the time scale of hydrogen exchange is ~ms. But one can slow the intrinsic rate of hydrogen exchange down to time scales of ~1 hour by changing the pH to 2.5. (Lowering the pH below 2.5 allows the exchange rate to increase again because H^+ in very high concentrations will also catalyze hydrogen exchange).

The rate of hydrogen exchange is also greatly affected by protein structure. Factors such as *hydrogen bonding*, *solvent inaccessiblity* and *steric blocking* all play a part in the decreased rate of hydrogen exchange in structured regions of proteins. The actual rate of exchange on any specific backbone N-H may be much slower than if that hydrogen were freely accessible to water.

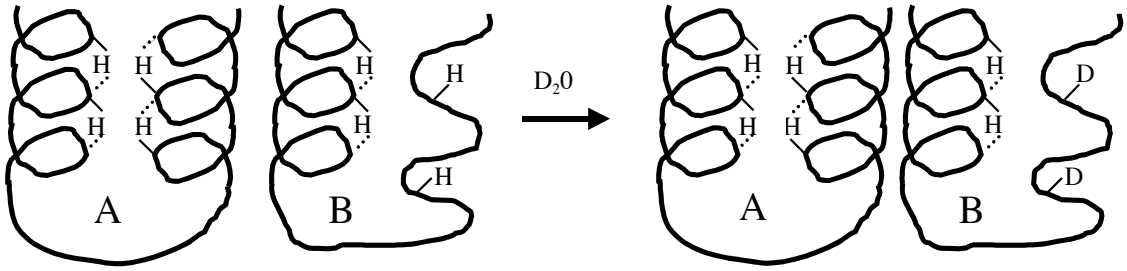
© Patrick Fleming, Johns Hopkins University

To summarize, there exist both an intrinsic *chemical* rate of exchange which occurs with free access to water and an observed *experimental* rate of exchange due to occlusion by neighboring atoms combined with the intrinsic chemical rate.

Both hydrogen atoms and deuterium atoms exchange in a similar manner. This fact is useful experimentally. The idea of a hydrogen exchange experiment is to put a protein into D_20 for a short time at pH ~9, allow the exchange of protein hydrogens with solvent deuterium to occur, then quench the exchange by lowering the pH to ~3. The protein may then be analyzed and wherever hydrogen/deuterium did *not* significantly exchange the protein was structured (i.e. internally hydrogen bonded, packed tightly, etc.); wherever deuterium labeling *did* occur the protein was unstructured (or at least the backbone was exposed to water solvent).

NMR and **mass spectrometry** can identify which positions on the polypeptide are labeled with deuterium. Therefore, one can determine the fraction of polypeptide that becomes accessible to solvent during a protein unfolding experiment. These methods allow one to measure the exchange (or local unfolding) at **each residue** in the protein.

Consider the following figure.



Assume a protein exists in two conformations, **A** with two helices and **B** with only one helix plus a flexible segment. Some hydrogens in conformation A are involved in hydrogen bonds within alpha-helices. In conformation B the helix on the right has unfolded and some hydrogens are unprotected. When the protein containing these two conformations is placed in D_20 the unprotected hydrogens exchange with deuterium.

We can model the exchange of hydrogens on a polypeptide backbone nitrogen, N, by the following reaction scheme,

$$NH_{closed} \underset{k_f}{\overset{k_u}{\rightleftharpoons}} NH_{open} \xrightarrow{k_{ch}} ND \qquad (1)$$

where k_u is the rate of **u**nfolding, k_f is the rate of **f**olding and k_{ch} is the intrinsic ***ch****emical* rate of exchange with free access to water (or D_2O). In almost all cases the rates of opening and closing (protein dynamic flexibility, ps to μs) are faster than the intrinsic exchange rate (ms), *so we assume that the first reaction is at equilibrium*.

The equilibrium constant for unfolding (the relative probabilities (*P*) of open and closed

conformations) is

$$K_{eq} = k_u/k_f = P_u/P_f = [\text{open}]/[\text{closed}] \quad (2)$$

The ***ex****perimentally* observed rate of hydrogen exchange, k_{ex}, is a *product* of the relative probability of being unfolded (K_{eq}) and the intrinsic chemical exchange rates for the unprotected hydrogens,

$$k_{ex} = K_{eq}k_{ch} \quad (3)$$

By rearrangement the equilibrium constant may also be expressed as,

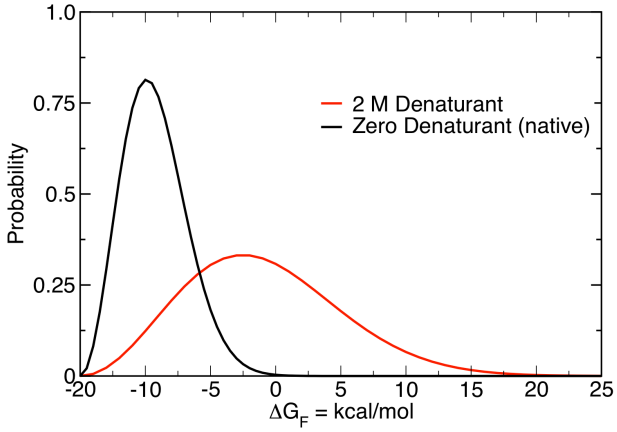
$$K_{eq} = k_{ex}/k_{ch} \quad (4)$$

Using the thermodynamic expression $\Delta G = -RT\ln K$ we can relate the free energy of ***unf****olding* of the helix to hydrogen exchange rates

$$\Delta G_{unf} = -RT\ln K_{eq} = -RT\ln(k_{ex}/k_{ch}) \quad (5)$$

where again k_{ex} is the observed, ***ex****perimental* exchange rate and k_{ch} is the intrinsic ***ch****emical* rate known from model studies. **We now have an expression to calculate the free energy of conformational unfolding from hydrogen exchange data.**

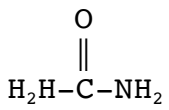
The probability of finding an unfolded segment in a native protein may be finite but very small. (Remember the heat capacity calculation). To obtain better labeling of partially unfolded segments one can destabilize the protein using a denaturant such as guanidinium chloride. A denaturant causes the solvation conditions to become less favorable for the stability of the protein. The whole ensemble of protein states is then shifted such that the unfolded conformation occurs much more frequently as indicated in the figure below. This increased sampling of partially unfolded conformations makes the hydrogen exchange experiment more practical.



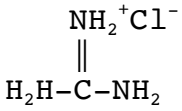
Destabilizing the protein ensemble this way **does not create any new conformations**; it only shifts the probabilities of pre-existing conformations. Remember that in a Boltzmann distribution ALL states are possible, some are more probable than others.

© Patrick Fleming, Johns Hopkins University

2. Protein Unfolding by Denaturants.
A number of agents such as urea and guanidinium chloride (GdmCl) will unfold a protein when added to the aqueous solvent. These are called denaturants.

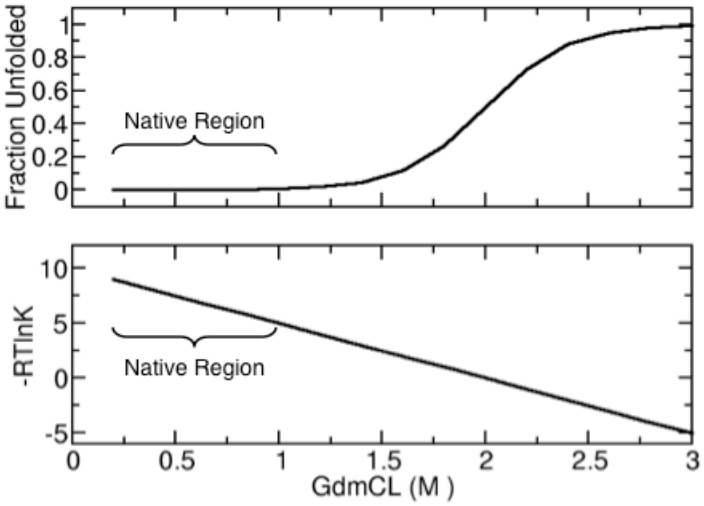


Urea



GdmCl

Below is a typical plot of the fraction unfolded protein versus GdmCl concentration. The unfolding is cooperative as evidenced by the sigmoid shape of the unfolding curve.



A transform of the unfolded experiment data can be linear if the fraction unfolded is converted to –RTlnK where K = [unfolded]/[folded]. This linearization of the data can be used as described below to estimate the free energy of unfolding.

In the region below 1.0 M GdmCL the protein is still mostly folded (as shown in the top panel of the figure above) and this is considered the "native state" region. However, the protein is becoming more and more unfolded with increasing GdmCL concentrations in this region as indicated by the changing value of –RTlnK in the lower plot.

The mechanism of action of these denaturants is still controversial. But it appears that favorable interaction of the denaturant with polypeptide backbone groups that are usually buried in a folded protein is important. This mechanism means that the action of the denaturants is proportional to the interaction surface area exposed during unfolding. Atomic surface area exposed to water is frequently called **accessible surface area (ASA).** We will use this term frequently in the coming weeks.

© Patrick Fleming, Johns Hopkins University

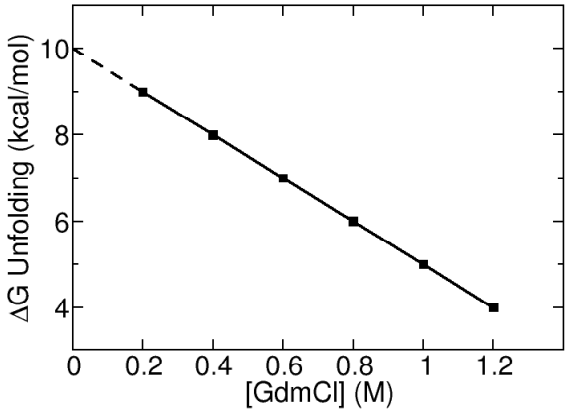
An empirical relationship between the free energy of unfolding and concentration of guanidinium chloride (linear plot above) is

$$\Delta G_{GdmCl} = \Delta G_0 - m[\text{GdmCl}] \quad (6)$$

where the ΔG subscript *(GdmCl)* indicates the condition with denaturant of a specific concentration and the subscript (*0*) indicates at zero denaturant concentration. If the protein exists in only two states, folded and unfolded, then the above relationship is linear with a slope of *m*. If unfolding of the protein exposed a large amount of denaturant interaction surface area (ASA) then *m* is large and *vice versa*.

The above empirical relationship can be used to estimate the energetic stability of a protein under native conditions. Below is shown a typical plot. In this case, the data from a denaturant unfolding experiment was plotted and the free energy of unfolding was estimated by *extrapolation* to zero denaturant.

Linear extrapolation to zero denaturant may be necessary if it is not possible to obtain significant HX data at the experimental condition of zero denaturant.



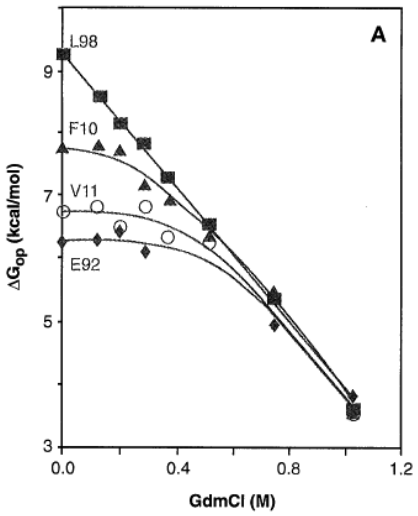
As seen below this type of analysis is used in hydrogen exchange experiments but the results get more interesting.

(Methodological note. If we consider protein folding in the direction of unfolded (*U*) to folded (*F*), ΔG_{fold} is negative (favorable) because proteins typically fold spontaneously. If we consider protein unfolding in the direction of *F* to U, ΔG_{unf} is positive (unfavorable). Thus the values of ΔG_{unf} in the above plot are positive and the values of ΔG_{fold} in the Boltzmann distribution plot showing protein stability are negative. You have to be aware of which convention is in use.)

3. Global and local unfolding.
Consider the following data from a hydrogen exchange experiment on cytochrome *c* taken from a seminal study (Bai, Y., Sosnick, T.R. Mayne, L. and Englander, S.W. *Protein Folding Intermediates: Native-State Hydrogen Exchange*. Science 269:192-197 July, 14, 1995).

© Patrick Fleming, Johns Hopkins University

The plot at right shows the free energy of unfolding of parts of the cytochrome molecule as a function of the denaturant concentration. The part of the protein monitored by leucine 98 (L98) has an unfolding free energy of ~9 kcal/mol in zero denaturant and shows a linear relationship with the denaturant concentration. But other segments of the protein monitored by residues F10, V11 and E92 appear to have a lower free energy of unfolding at zero denaturant, i.e. these segments are more unfolded than L98 in the native protein.



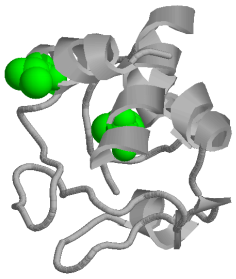
Also, note that the authors of this work were able to measure hydrogen exchange in zero denaturant, i.e. they did not have to extrapolate to zero (not always the case).

The interpretation of this plot is that L98 represents the **global** unfolding of the protein and that the other residues represent **local** segments that are more accessible to denaturant at low denaturant concentrations. Note: This does not mean the L98 is in a region that never unfolds in zero denaturant – remember the Boltzmann distribution of energy levels in a population of molecules. Just the fact that there is enough exchange at L98 to obtain exchange data at zero denaturant is evidence that this region of the protein undergoes occasional dynamic unfolding. In fact, L98 is one of the most slowly exchanging residues, so when L98 unfolds, **all** parts of the protein are undergoing transient unfolding. The proportion of L98 exposing its backbone to solvent at an unfolding energy of 9 kcal/mol (zero denaturant concentration) is 1 in several hundred thousand.

As the concentration of denaturant reaches ~0.5 M even the L98 segment is as destabilized as the other segments (reported by F10, V11, E92) and all segments of the protein are now equally experiencing conformational unfolding (slopes are the same).

Notice that the slope of the line for V11 is horizontal at low denaturant concentrations, i.e. the slope *m* from equation (6) is zero in this range. This would indicate that little new interaction surface area is exposed during the increasing interaction of this residue with denaturant. How can this be?

The answer is as follows: The figure at right shows L98 in green buried in the middle of the protein. Its backbone is only exposed on global unfolding. In contrast V11 in green (upper left) apparently can transiently expose its backbone more frequently and increasing denaturant at low concentrations does not significantly increase this frequency.



The take-home message to this point is that different regions of a protein may have different stabilities (frequencies of unfolding).

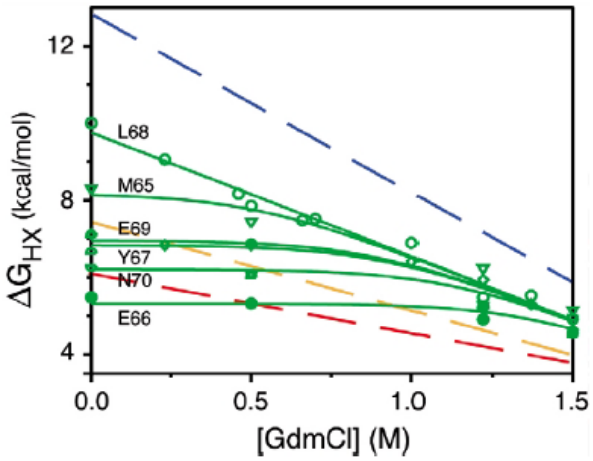
© Patrick Fleming, Johns Hopkins University

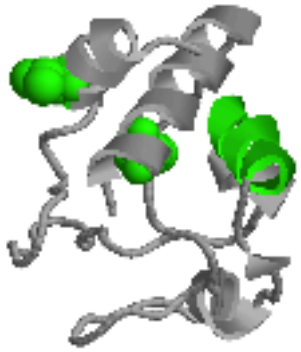
Segments that eventually unfold with the same frequencies as the core of the protein approach the same slope in the plot above. This means that when the core unfolds the whole protein is unfolding cooperatively and exposing the backbone to water.

4. Subglobal unfolding.
When a number of residues close in space (not necessarily in sequence) asymptotically approach the same denaturant sensitivity curve they are considered to unfold as a unit = **subglobal unfolding**.

Consider the data in the left figure below taken from Englander, S.W. Mayne, L. and Rumbley, J.N. *Submolecular cooperativity produces multi-state protein unfolding and refolding*. Biophysical Chemistry 101-102:57-65, 2002. It shows the hydrogen exchange data for the hydrogen bonded amides in the "60's" helix in cytochrome *c* (residues 61-69). This helix is colored green (together with L98 and V11) in the ribbons diagram on the right.

Most of the 60's helix residues in the graph show zero slope at low denaturant concentrations (green data) because the conformational change that allows hydrogen exchange is occurring in the native state. These residues (65, 66, 67, 69, 70) are both close in space and in sequence. From this data, we may say that the helix is a cooperative folding unit. L68 is in the middle of the folding unit and unfolds with the lowest frequency, E66 unfolds with higher frequency. But the entire helix eventually become globally unfolded together.

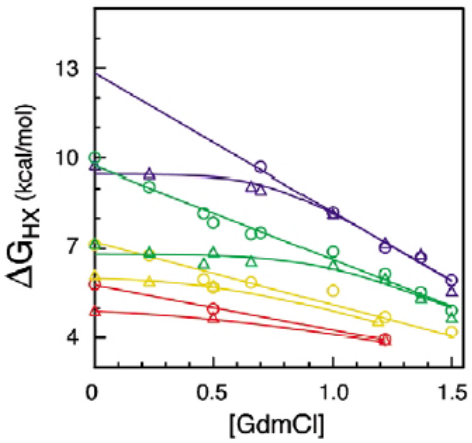


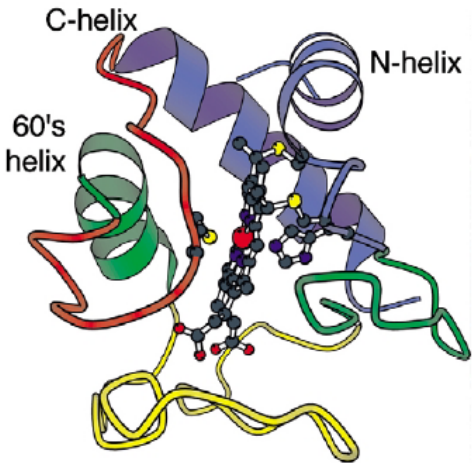


The take-home message from this plot is that different residues in the same unfolding unit may have different stabilities, but at some denaturant concentration they are all unfolding with the same frequency.

Similar experiments show that other segments constitute separate folding units and the data are summarized in the next plot with the corresponding segments in matching colors in the ribbon diagram on the right.

© Patrick Fleming, Johns Hopkins University



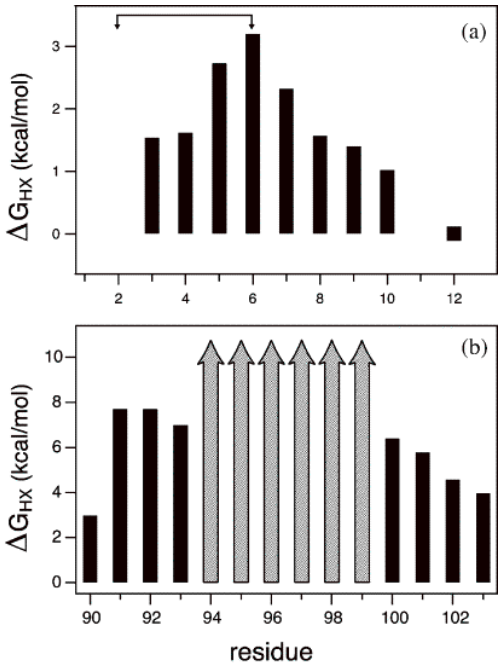


Here the red loop is the least stable, the yellow loop is next in stability, then the green helix and loop and finally the magenta (blue) helices are the most stable.

Summarizing to this point:

- Even the most stable parts of a small globular protein undergo conformational fluctuations that expose the backbone amides to water.
- Different segments of a protein have different degrees of conformational fluctuations which is the same as saying that they have different free energies of unfolding.
- Segments with similar free energies of unfolding and that also are close in space are called cooperative folding units.

Now we look at individual residue differences *within* a cooperative folding unit. The figure below shows free energies of unfolding on a residue by residue basis for two helices, a synthetic peptide helix (top) and the C-terminal helix of cytochrome *c* (bottom).



Note decrease in ΔG_{HX} from residues 6 to 12. (i.e. the rate of HX increases).

Note decrease in ΔG_{HX} from residues 100 to 103. (The rate of HX increases for these residues).

© Patrick Fleming, Johns Hopkins University

In (a) the data for a synthetic 12-residue peptide are shown. The data in part (b) are from residues 90-103 of cytochrome c. This is the C-terminal helix as shown by the blue segment in the figure below

The "core" of the helix unfolds as a unit. But the ends of the helix show "**fraying**" or graduated loss of hydrogen exchange protection. These data demonstrate that the conformational fluctuations along a helix, even in a native protein, are context specific.

In the next computer lab, we will create an energy weighted distribution of peptide helix conformations and demonstrate that such context specific conformational fluctuations are contained in the Boltzmann distribution created by a simulation.

© Patrick Fleming, Johns Hopkins University