

20.320, notes for 10/25

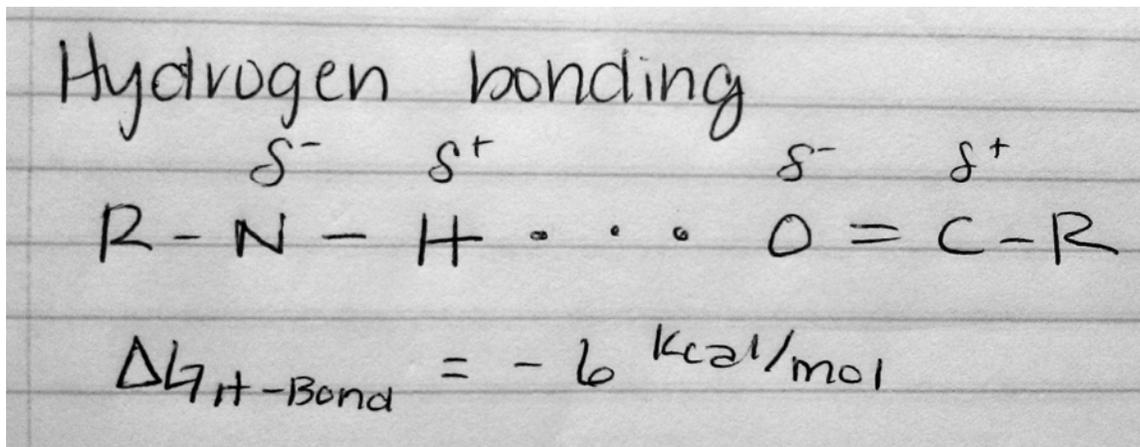
Thursday, October 25, 2012
9:40 AM

Previously, on 20.320...

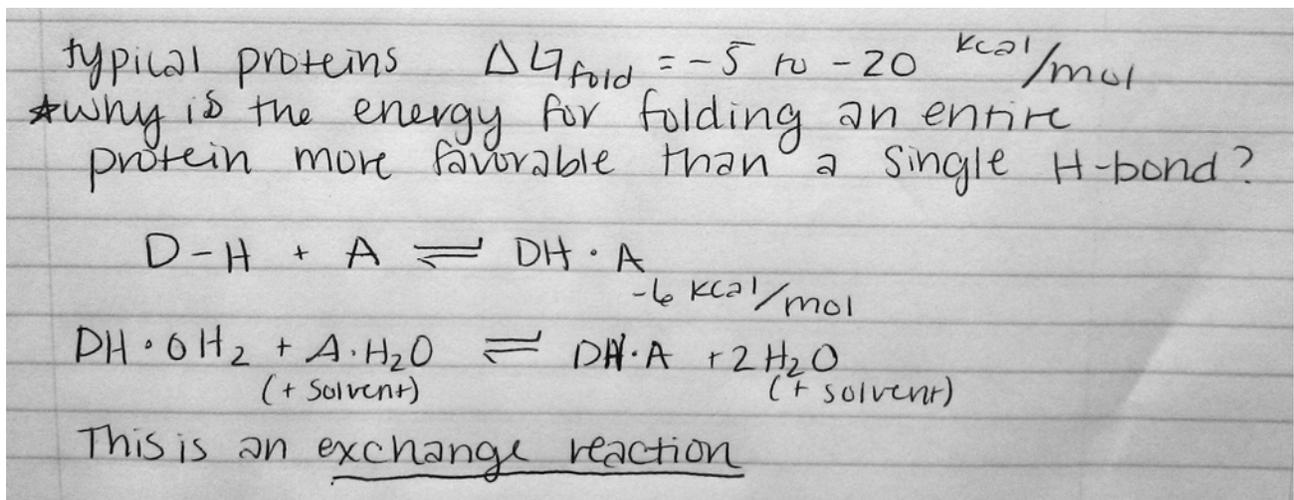
We covered electrostatics. You know, potential energy as a function of charge and separation.

Hydrogen Bonding

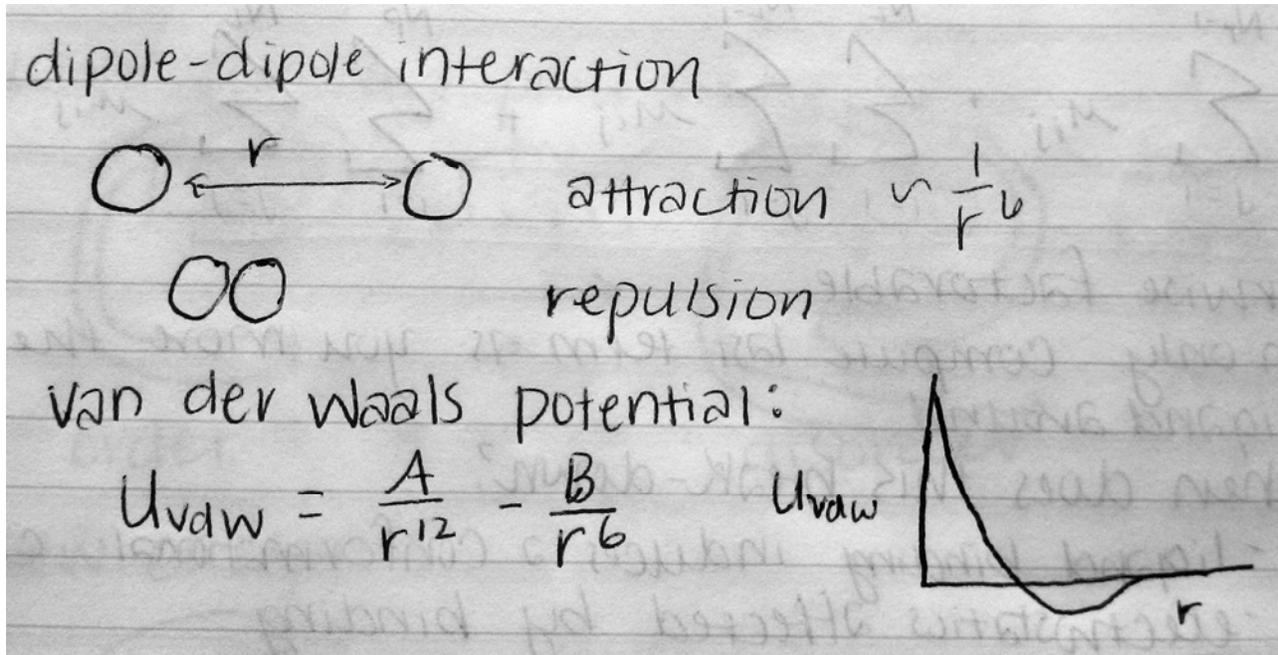
Will happen between a two molecules, one of them holding a highly electronegative group like O or N, and the other holding a hydrogen attached to a group like that. H-bonds are the main contributor to the stability of proteins. The free energy change associated with forming a single H-bond is ~ -6 kcal/mol (in a vacuum). Protein folding is also spontaneous, but its stabilization energy is only in the order of -5 to -20 kcal/mol (even though dozens to hundreds of new H-bonds are being created). The reason why this stabilization energy is so comparatively tiny is that the protein is already forming lots of H-bonds with the water around it, all of which must be broken.



Ours is really an exchange reaction, with exactly the same number of H-bonds formed as destroyed. And yet, folding leads to a just slightly more favorable rearrangement. The reason is that not all H-bonds are the same. The geometry can be more favorable (straight-line alignment of the molecules), and the overall system entropy is increased when water needs only arrange itself around a small, folded protein and not around an unfolded ribbon. All of this makes proteins fold.



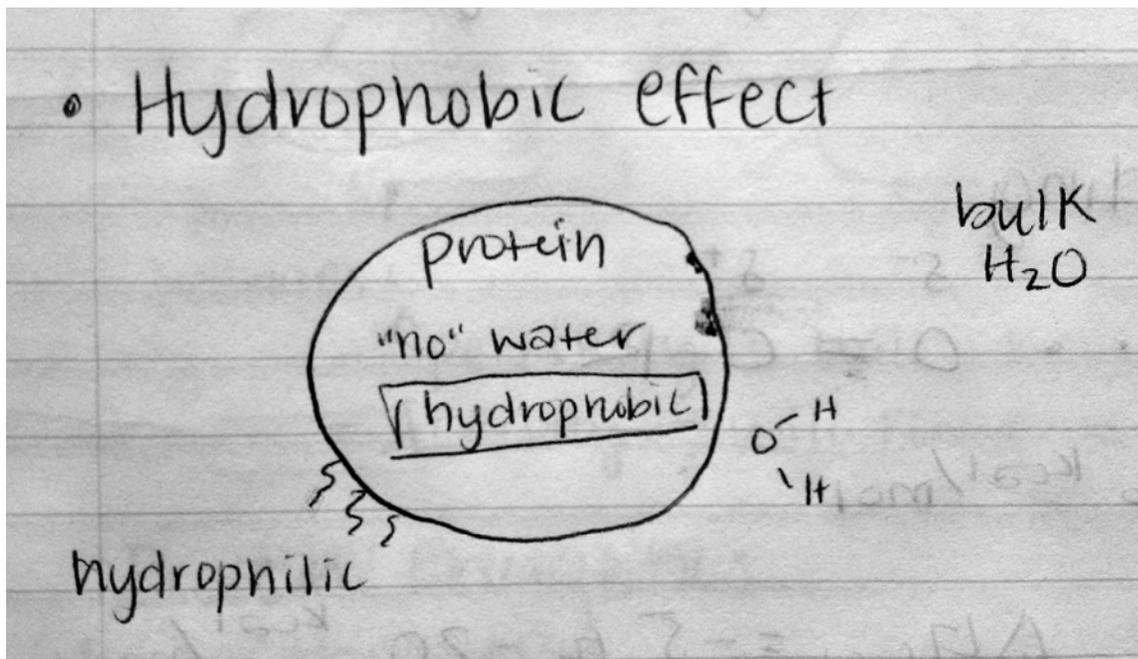
Why don't H-bonds go through each other? Two atoms, even if uncharged, will attract each other beyond a certain distance (until their electron clouds overlap and start repulsing). This is due to the mobility of their electron clouds, and is called dipole-dipole interaction. This attraction scales proportional to $1/r^6$. If we combine this attraction with the much shorter-range repulsion force we mentioned, we get the more complete picture of Van der Waals forces:



The hydrophobic effect

This is perhaps one of the most important effects to consider in biology. The net effect is that non-polar (oily) substances will tend to stick to each other and remain separate from water. The reason has to do with the formation of H-bonds in an aqueous solution.

Hydrophobic substances can't form H-bonds with water, by virtue of being non-polar. Therefore, water arranges itself around these molecules in a highly ordered pattern that maximizes H-bonds among water molecules while avoiding the hydrophobic core. These are called clathrate cages, and are structures of much lower entropy than free water with all of its degrees of freedom unhindered. Water would rather be bonding, and the best way to do it is to put all the hydrophobic stuff together so that the minimum number of water molecules are trapped forming clathrate cages (one large oil drop has less surface than two small ones). The net result, again, is the bunching-up of hydrophobic molecules and bodies in water.

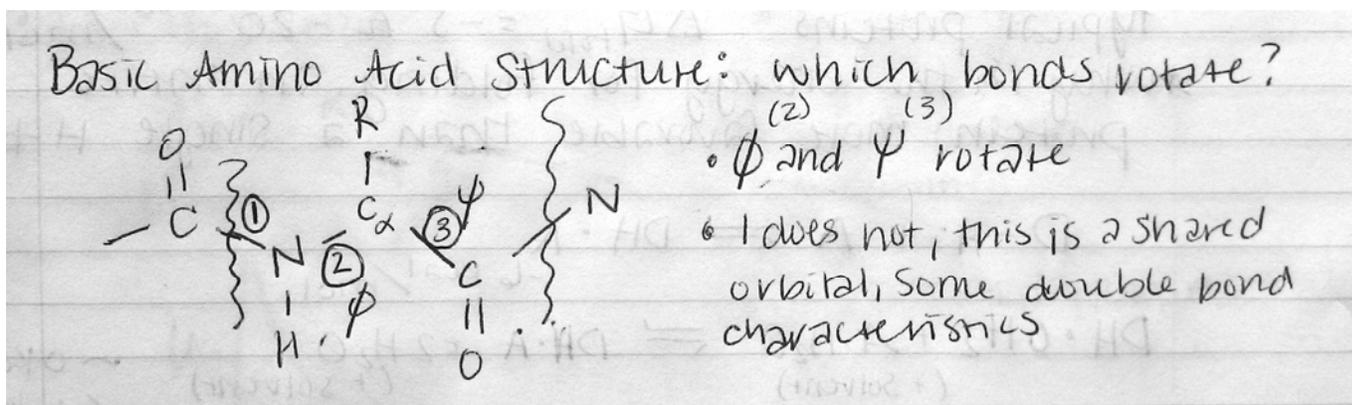


And so, to calculate protein-protein and protein-ligand interactions we will calculate the values of all these interactions in one and another state, trying to determine which is more favorable. We need to do all these calculations for just about every atom in our protein. It's not trivial.

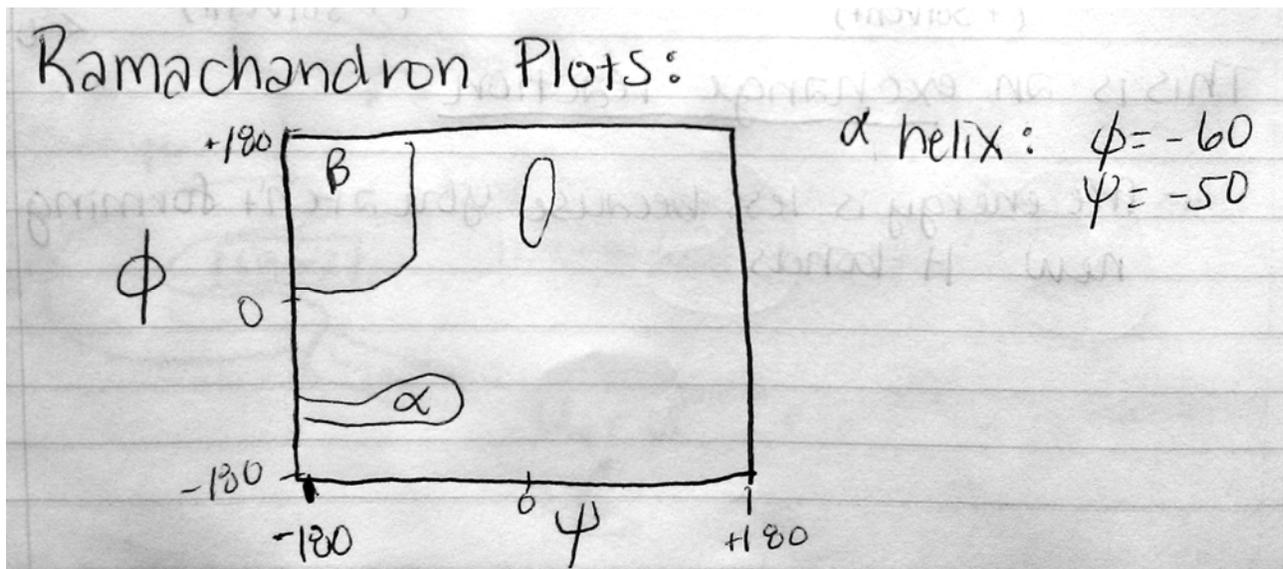
How to represent proteins

We can focus on several levels of abstraction, all the way from the probability cloud of every atom in the protein (and surrounding solvent) to abstract versions of primary, secondary, or tertiary structure.

To represent a protein, we could specify the coordinates of every atom in the structure. Because of the constant length of covalent bonds and their different ability to rotate, we can get away with specifying nothing but the protein's sequence and the rotation angles of each bond. Our protein backbone is best modeled as a chain of flat units, with two angles determining all the important variation. We get the same picture in much fewer words, because rather than needing 3 coordinates per atom we need just two angles per residue. Now, not all combinations of these angles are always possible. Enter the Ramachandran plot:



The Ramachandran plot is a plot of ϕ vs ψ (for a given peptide), showing which combinations of the angles are possible and which lead to the amino acid backbone running onto itself. There are also regions of the plot that are characteristic of particular structures (like alpha-helices or β -barrels).



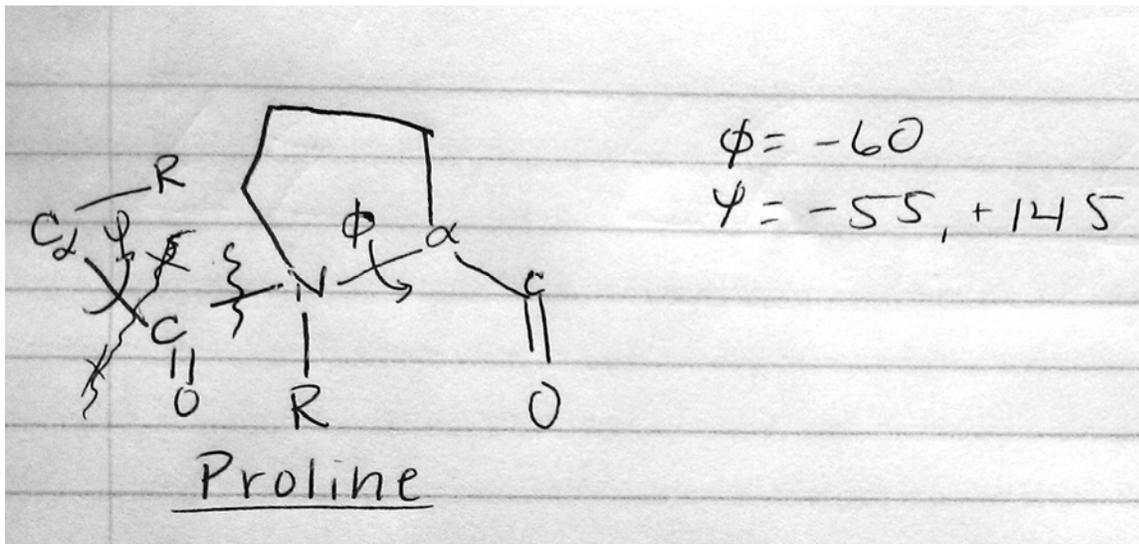
The Ramachandran plot is an example of conformational space, an n -dimensional space ($n = 2$ in this case) that describes the structure of a system. Each dimension in the space is a meaningful measurement of our system, so the plot contains possible states of the system represented by different locations in the space. The most important concepts of conformational space are that :

1. Similar structures are close to each other in the space.
2. We can associate an energy value to each point in the conformational space.

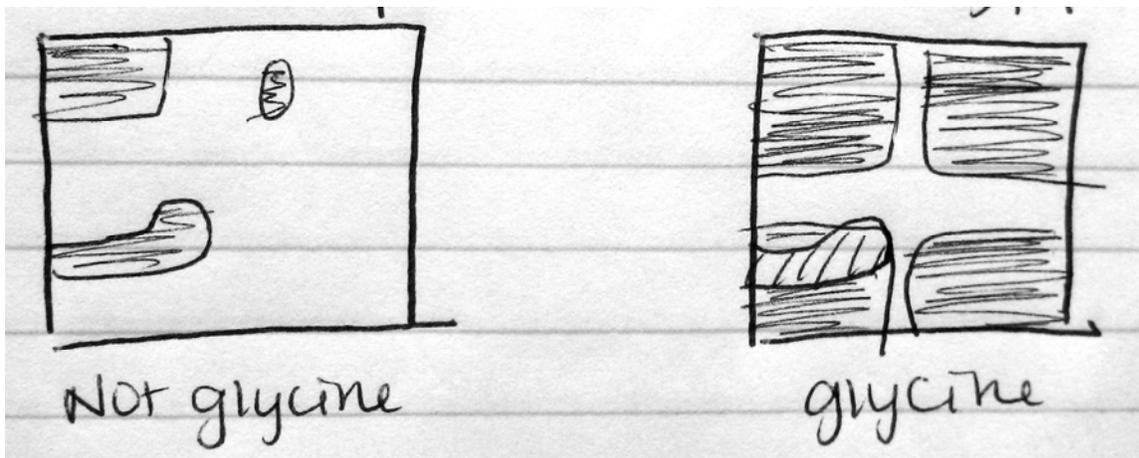
Note that we talk about forbidden regions, but in reality the cutoff is not that binary. It's just that there's an energy penalty associated with forbidden areas. Some of them happen in proteins (structural stress is possible), but never too many without a good compensating region.

In 1951, Linus Pauling predicted that alpha-helices ought to happen inside proteins. The reason is that there's very little water inside the protein, and so the residues need to form H-bonds somehow else if the structure is going to be energetically favorable overall (remember all the bonds that were broken from water when it folded). Alpha-helices are the structures that maximize H-bond formation, compensating for the bonds that can't be formed with water.

On that note, do you know why proline doesn't form alpha-helices? Figuring it out is a nice example of the useful ideas that come from thinking in conformational space. The problem with proline is not because of its own ϕ and ψ angles, which are actually just about right. What happens is that the residue next to proline are restricted, twisted such that their ψ angle cannot be -50° . This is not a problem, though, at the end of the helix.



Glycine is another residue that doesn't go into alpha helices. But why? It's the most flexible of residues, so it can definitely adopt the right conformations. Steric hindrance is not the problem. What happens is the opposite: glycine has immense conformational freedom, much more than any other amino acid. Therefore, going into the very limited conformational space of alpha-helices means restricting itself immensely. There is a big entropic cost to this, which is less intuitive than thinking of prolines and their steric limitations.



Entropy is usually much harder to mentally account for than enthalpy, but it's really important. That's what we're learning to do here.

MIT OpenCourseWare
<http://ocw.mit.edu>

20.320 Analysis of Biomolecular and Cellular Systems
Fall 2012

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.