Chem 431A-L09-F'07
10//10/07; Wednesday
admin:

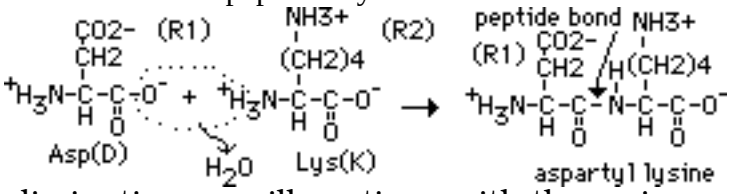
Last time: Finished nucleic acids: physical properties, replication glimpse.

Today: start discussion of amino acids
The lecture notes for Wed & Fri will be posted.
NO CLASS ON FRIDAY

Today: start amino acids (aa).

Introduction to proteins: First of all, by way of introduction, proteins are the most versatile of the bio macromolecules we are going to be able to study. These remarkable molecules do everything from catalysis, to providing support, to covering our body, to aiding in ionic transport, etc…… They can do these because of the versatility of the components that make them up. Recall that proteins are heteropolymers. What are the monomer components? There are 20 different amino acids that make up a protein. To be more precise they are called α-amino acids. These contain an a carbon, ie the carbon that is bound to the carboxyl group, this same α carbon is bound to the amino group as well. So it is an α amino acid. Carbons can have 4 bonds. The other two groups are "H" and finally the fourth is the amino acid "residue". That residue is what confers uniqueness to each of the amino acids.	Let us look at the structure of these aa's. You are to commit to memory all these structures and know their full names as well as their 1- and 3-letter abbreviations. View the structures overhead. Discuss the physical properties of these aa's. glycine, alanine, valine, leu, ile (aliphatic); Pro; Ser. Thr, Cys, Met ("polar"), phe, tyr, trp (aromatic), Lys, Arg, His (basic), Asp, Glu (acidic), Asn, Gln(amides). -H, -c3, cH2-ch- - ch3, -ch(-ch3)-ch2-ch3, Note that these aa's have unique properties which you should master by reading the text book. These aa's are zwitterions. Both basic and acid groups. Can do titration of these aa's and get titration curves. They have various migration patterns if placed under electric field. PH dependent. Each of these aa's are coded = matched to specific tRNAs

Let's conc on the behavior of these acids when they combine to form peptides by removal of water



elimination.we will continue with the amino acids:

1) α-amino acid's are called α-aa's because the amino group is attached to the C_α. C_α is the carbon to which the carboxyl group is attached. If the amino group is attached to the "next" carbon (C_β), it would be a β-aa for instance. There are some β aa's used in the body, but these are not found as part of proteins.

2) the C_α is a *chiral* or asymmetric carbon. that means that the four groups bound to it are different (exc for Gly). the two mirror images (optical isomers) are called *enantiomers* and have a different effect in rotating polarized light passing thru a pure solution of either one of these enantiomers. As such they are distinguished by assigning an "L" or a "D" to them, depending on whether they rotate the polarized light to the left (L) or to the right (D). All the 20 amino acids which are found in life are "L-amino acids". Why? Not clear since left handedness or right handedness are equally probable. There is no necessary preference for either one. It is probable that the evolution just got started on the "left foot". Note that proteins that are L -aa based can act on L proteins. It seems some attribute the advantage of L over D as arising somewhat from the left handedness of β-decay electrons. I don't know if that makes sense to me.

What is clear is that if organisms had to organize their molecules using both D and L aa's they would have to have to significantly multiply the number of metabolic machinery needed.

3) We review the discussion on acids and bases: aa's have both acidic and basic groups. eg G has a pK_a of 2.3 and 9.6. At pH7, we expect it to be ionized and become a zwitterion. Please keep in mind what we learned in Chapt 2 about acids and bases and their pKa's. In this instance, we note that at pH = 2.3, the carboxyl group would be 50% ionized. at pH>2.3, it would be predom ionized while pH<2.3, predom protonated. Now looking at the amino group, we note that at pH2.3, it would be <9.6 so that the amino group would be protonated. until the pH rose to 9.65 …

At what pH would G be predom neutral? at pH= 2.3+9.6)/2=5.95 = pI (isoelectric point). At pI, G would not migrate under an electric field. show the a fraction. In the case of Lys, have 3 pKa's: 2.2, 9.0, 10.0. What is the a fraction profile of Lys? we note that its pI is at 9+10/2=9.5. So if the pH is 5.95, we could separate G from L since G would not migrate and L would go toward the "-" cathode. If the pH was 7, G would be "-" while L would be "+".

3a) note that these aa's have unique properties which you should master by reading the text book. These aa's can be separated in solution based on their charge and partition properties (tendency to associate with one solvent or phase over another).

- electrophoresis (separation according to the charge of the aa) as discussed above.
-chromatography (ion exchange, gc, & hplc)

In summary: These aa's are zwitterions. Both basic and acid groups. Can do titration of these aa's and get titration curves. They

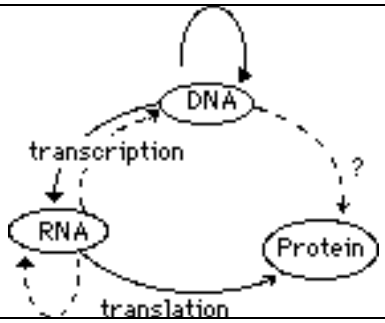
have various migration patterns if placed under electric field. pH dependent.

4) Proteins are long *unbranched* chains of amino acids formed together by peptide bonds. It turns out that the specific sequence of amino acids stringed together to form a protein is very critical to determining what protein's final structure will be and therefore what its function will be in a cell. The length of the *polypeptide* chain is typically over 100 to 1800 aa's (myosin). A muscle protein, called titin, has 17,000 aa's! average is about 270 aa's in eukaryotes (mw=31,700 amu, daltons).

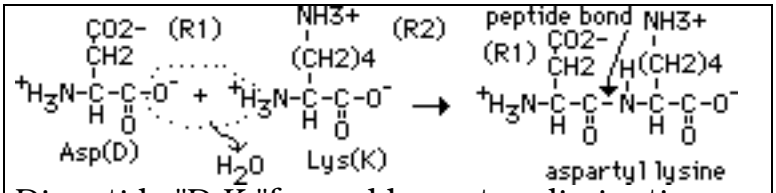
Just as an exercise, how many possible can you make using the 20 aa's? well, there are 20 possibilities for the first and 20 possibilities for the second. so the total is 20x20=400 unique dipeptides. in general, for n-peptide, $(20)^n$! What would it be for the smallest proteins? $(20)^{100}$!

Each protein has a unique sequence of aa's which give it its unique properties. this sequence is the primary structure (1°) of the protein. like DNA right? This sequence is encoded in the genetic material's DNA molecule. By convention, the sequence is read from the N-terminal end of the polypeptide chain through to the C-terminal end of the chain.

Each of these aa's are coded = matched to specific tRNAs ***Show the overhead on the aa's and the *translation* step.

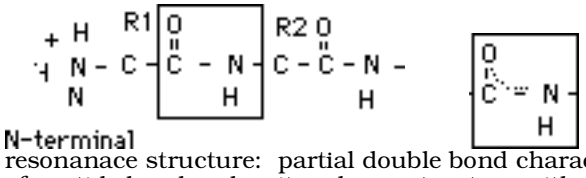


5) How are the monomers bonded to each other? By a *peptide bond*. Consider the reaction of these amino acids when they combine to form peptides by removal of water A metastable bond is formed called the peptide bond. G is negative for its hydrolysis, but actual breakdown of the peptide bond is very slow (can be catalyzed by protease enzymes).



Dipeptide "D K "formed by water elimination.

6) Peptide bond has a *partial double bond* character. Linus Pauling analyzed the X-ray data from proteins (which indicated what the distances from C=O bonds and C-N are supposed to be and what they were in the peptide bond), Led him to conclude that the C-N bond was not a simple sigma-bond, but a partial π bond as well.



resonanace structure: partial double bond character of peptide bond makes it a planar structure with a resonance energy of 88kJ/mol; a twist of 10° involves a twist energy of $88\sin^2(10°)$ kJ/mol

Consequence of this? A double bond character will make the C-N bond hard to rotate (like you can't rotate a double bond).

we talk about peptide bond; 1° structure and translation of mRNA.
Then, 2° structure – α-helix and β-sheet. Also we look at how X-ray diffraction (crystallography) is used to determine structure of molecules. Ramachandran plot (Ψ and ϕ)

It is a planar structure. It is a plane with a twist energy. Resonance energy of 88kJ/mol; a twist of 10° involves a twist energy of 88 $\sin^2$ (10°) kJ/mol

** look at the structure of the protein chain: less flexible due to the presence of the peptide bond.

(Ψ and ϕ)

It means that per residue, there are only 2 degrees of freedom. these are the rotation between C_α carbon and it's N of the preceding peptide bond and C_α carbon and the C of the next peptide bond. This constraint is important in limiting the possible structures that the polypeptide chain can fold into. The angles are Ψ and ϕ. Ψ is the angle of rotation about the C_a-CO and ϕ is the angle of rotation about the C_a-N bond.

** the *trans* form of the peptide is usually favored. The cis form has bulky R groups of adjacent C_α's start to sterically interfere with each other. Except for case of proline where cis is more allowable (1:4 ratio) with trans still more favored.

7) secondary structure of polypeptide chains: we note that the constraints to the peptide bond don't allow it to rotate so easily. There are two regular structures that allow for the constraints and these are the alpha helix and the beta sheet.

OK. α-helix is called (3.6_{13}) 3.6 residues per turn, 13 atoms in the H-bonded ring. 1.5Å rise/residue. so (1.5x3.6 = 5.4Å/turn. ignoring sidechains, diam is 6Å. Sidechains extend outward from the core structure of the helix.
note that the helix is stabilized by the combined effect of the H-bonds. One thing to be said about the DNA helix is that it is stabilized by many factors. when the base pairs are aligned and H-bonded, the DNA double strand is stabilized. but if you heat it up, it starts to vibrate until some T_m wherein, the first bonds are broken and the thing is quickly unravelled. something like a crystal melting.

Linus Pauling (and Corey) (Caltech) proposed the α-helix:

The alpha helix is a repeating structure with the ff parameters:
Draw the structure of a poly peptide and show which H-bond to form the α-helix.

note that each peptide carboxyl is H-bonded to the NH group 4 residues up the chain.

Here, α helix is also stabilized by the H bonds. between the residues which altho they are not neighbors in the 1° sense, are actually neighbors once the α helix is formed. there is stabilization forward and back.

We note that each of these possible 2° structures is characterized by a Ψ, ϕ. Angles Ψ =-45°to-50°, and ϕ =-60° correspond to α-helical structures. These are shown in the Ramachandran Plots which show allowable values.

Max Perutz noted x-ray data that indicated that a helix is found in α keratin (protein in hair) and is of fundamental importance in protein structure in general. It is interesting to view the hair as being composed of helical strands of polypeptide chains. hair is mostly stretchable as would be expected.

Going back to the Ψ and ϕ angles: some of these angles are not favored because of steric crowding. If we could specify these angles for each Ca, we could determine the secondary structure of the polypeptide chain. If the angles are

dipole moment: each peptide bond has a dipole moment. In a helical structure all the dipole moments point along the axis of the helix and there is a net dipole moment for the a helix itself. with + pole in the N-terminus and - pole in the carboxyl terminus end. Often ligands bind to ahelix structures in protein. negativelhy charged ligands often bind to the N terminal end. (but + ligands rarely bind to the c terminus)

Other less common helices: 3_{10} helix, etc.

studies with polyamino acids (ie chains containing only one type of amino acid), show that some aa's tend to form a-helices more than others. such as polyA and polyL. These tend to spontaneously form a helices whereas polyE and polyD form random coils (at pHs where it is neg. charged) but at pHs where sidechains are neutral and protonated, they form a-helices. similarly, when polyK is neutral sidechains, the a helices form spontaneously.

Other common structure is the beta(β) sheet (pleated sheet). Also proposed by pauling and corey in 1951. one can think of the β sheet as if it were a 2-fold helix (helix with only 2 residues/turn). We note that it represents the most extended conformation of the peptide chain. thus proteins with extensive β sheet structure are not very stretchable (like β-keratin in feathers and in in silk).