# Lecture 6: Amino acids & Introduction to Protein Structure.

**Pre-lecture Example:**



a) Draw the correct form of glycine at pH=0, pH=7, and pH=12. The pKa of the carboxylate group is 2, and the amino group is 9.

b) At what pH is it more likely for glycine to go through a membrane?

**Solution:**

1) Sketch the curves for fraction protonated versus pH for both the carboxylate (red, solid) and the amino (blue, dotted), using their respective pKa values.

2) Draw the structure based on the fraction protonated at each pH.

|  |  |  |
| --- | --- | --- |
| pH = 0 | pH = 7 | pH = 12 |
|  |  |  |
| *Both the carboxylate and amino are fully protonated at low pH (high [H+])* | *The amino is still full protonated, but the carboxylate is fully deprotonated.* | *Both groups are fully deprotonated at high pH (low [H+ ]).* |

3) Only neutral molecules can go through the membrane, glycine cannot go through a membrane at any pH value – it is always charged.

**Amino Acids (Chapter 6)**

* An amino group attached to a central carbon (α-carbon).
* A carboxylic acid group attached to the α-carbon.
* The amino group, α-carbon, and carboxylic acid will become the “mainchain” of a protein.
* One of twenty different “sidechains” attached to the α-carbon.
* The α-carbon is chiral in all but one amino acid because four different groups are attached to the α-carbon.
* Only one enantiomer (L-form) is present in proteins.

**Protein Structure**



* Amino acids are linked together in linear chains to form proteins.
* The **peptide bond** is formed when two amino acids are linked, releasing a water molecule (**condensation reaction**).
* When incorporated into a protein, an amino acid is called a **residue** (“amino acid residue”).
* Sequence is written starting from the amino terminus to the carboxy-terminus.
* The residue number also begins at the amino terminus: 1-2-3. In this example valine is the second residue.

**Twenty common amino acids found in proteins** (You do **not** need to memorize these).

**Glycine**: side chain is –H, neither polar or non-polar.



**N**on-polar, non-aromatic (Ala, Val, Leu, Ile)

**A**romatic, polar (His) to non-polar (Phe)

**P**olar sidechain (Ser, Thr, Met, Asn, Gln)

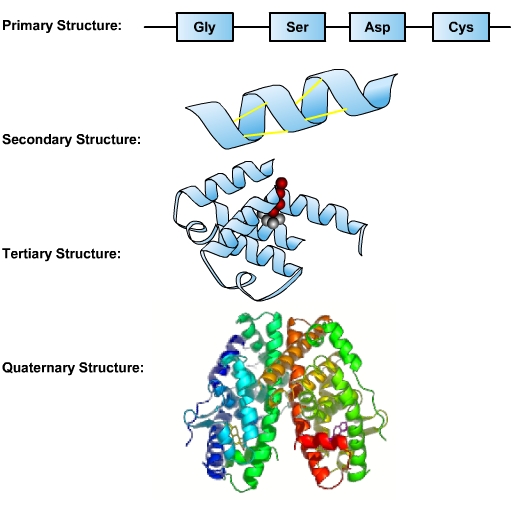
**A**cidic side chain, these groups ionize at pH=7, giving them a negative charge (Asp, Glu).

**B**asic side chains, these groups are protonated at pH=7, giving them a positive charge (Lys, Arg).

Two special amino acids:

Cys – forms crosslinks in proteins.

Pro – its sidechain forms a covalent bond with the nitrogen.

**Protein Structural Hierarchy:**

**1. Primary structure (1°):** The amino acid sequence, written from the amino to the carboxy termini.

**2. Secondary structure (2°):** Configuration of mainchain atoms only.

**3. Tertiary structure (3°):** Entire 3-D structure of one chain, both mainchain and sidechain.

**4. Quaternary structure (4°):** Association of subunits. Subunits can be the same (homo) or different (hetero).

**Secondary Structure**:

**Peptide bond:**

1. It cannot rotate (partial double bond).
2. The carbon is planer because it is bound to three atoms.
3. The nitrogen is also planer, so all four atoms (O, C, N, H) lie in a plane.
4. The C=O can accept a hydrogen bond
5. The N-H can donate a hydrogen bond, but cannot accept one since it is involved in a partial double bond.
6. Trans form is more stable than cis.



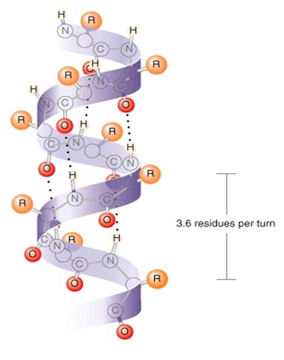
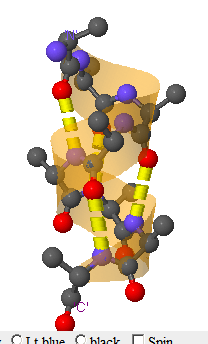
**N-Cα & Cα-CO –** are single bonds and are freely rotatable, giving the unfolded polypeptide chain considerable flexibility.

* The large number of possible conformations stabilizes the unfolded state of proteins because disorder is favored.





Although the mainchain atoms can assume many different conformations, there are two **hydrogen** **bonded** **structures** that are stable. The hydrogen bonds are between main chain atoms, the C=O accepts a hydrogen bond from an N-H.

******α-Helix Structures**

Dimensions, geometry, & H-bonds

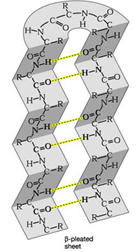
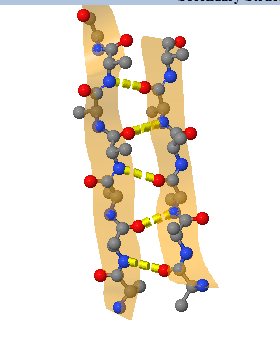
3.6 residues/turn

5.4 Å/turn

H-bonds || to helix axis.

Sidechains point outwards

Right handed twist.

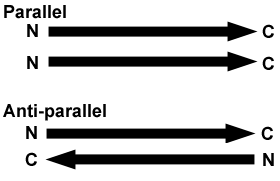
**Beta Structures -** β-Sheets

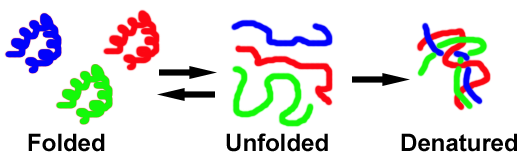
H-bonds perpendicular to direction of strands.

Sidechains point up and down, **above** and **below** the sheet.

a. parallel

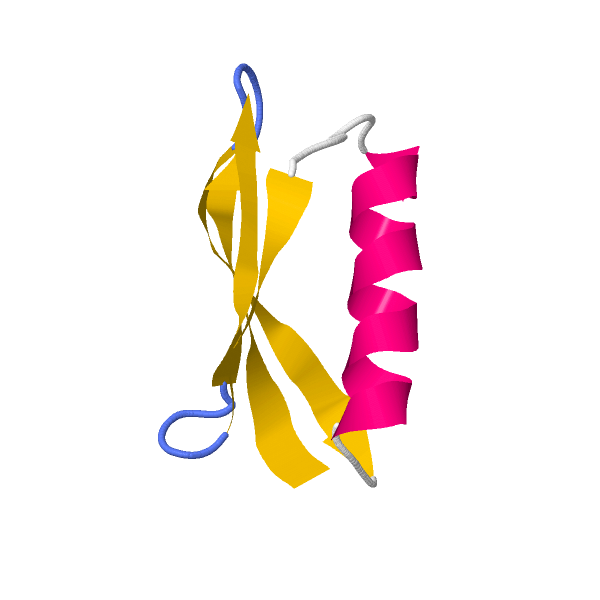
b. antiparallel



**Tertiary Structure – Conformation of all atoms in one chain**

Folded proteins have a well-defined, unique tertiary structure. The tertiary structure is often formed by the assembly of secondary structures. Remember that the tertiary structure refers to the conformation of all of the atoms in the chain, both mainchain and sidechain. Folded proteins are:

i) compact, well packed.

ii) contain H-bonded secondary structures

iii) have the following distribution of amino acids:

|  |  |  |
| --- | --- | --- |
| **Amino Acid type** | **Inside** | **Outside** |
| Polar |  |  |
| Charged |  |  |
| Non-polar |  |  |

To understand protein folding, we need to understand how the different amino acid side chain groups can interact with water, and then with themselves.

**Expectations:**

1. Can you identify the chiral center in an amino acid and know that only one enantiomer (L) is common in natural proteins.

2. Can you predict the ionization state of the sidechain group of amino acids, given the pH and the pKa of that sidechain?

3. Can you describe the formation of the peptide bond – which atoms are involved, what is released?

4. Can you draw a dipeptide, and label the peptide bond, the amino and carboxy terminus.

5. Can you distinguish between mainchain and sidechain atoms?

6. What is the difference between an amino acid and a residue?

7. Do you know how to write the sequence of a protein given its structure?

8. You should know that the peptide bond is planer, trans, and cannot rotate.

9. Can you explain why peptides/proteins can be very flexible when unfolded?

10. Can you describe why helices and sheets are stable, and their properties?

11. Can you identify functional groups on the sidechains of amino acids (polar/non-polar) and discuss their location in folded proteins.