Amino Acids

- Amino acids are building blocks for proteins
- They have a central α-carbon and α-amino and α-carboxyl groups
- 20 different amino acids
- Same core structure, but different side group (R)
- The α-C is chiral (except glycine); proteins contain only L-isoforms.
- Amino acids are ampholytes, pKa of α-COOH is ~2 and of α-NH_2 is ~ 9
- At physiological pH most aa occur as zwitterions.

Classification of Amino Acids

(based on polarity)

- Hydrophobic / non-polar R group: Glycine, alanine, valine, leucine, isoleucine, methionine, proline, phenylalanine, tryptophan
- Polar R group (net charge 0 at pH 7.4): Serine, threonine, cysteine, tyrosine, asparagine, glutamine, histidine
- Polar R group (Charged ion at pH 7.4): aspartate, glutamate, lysine, arginine

Classification of Amino Acids

(Based on R-group)

- Aliphatic: gly (G), ala (A) , val (V), leu (L), ile (I)
- Aromatic: Trp (W), Phe (F), Tyr (Y), His (H),
- Sulphur : Met (M), Cys (C)
- Hydroxyl: Ser (S), Thr (T), Tyr (Y)
- Cyclic: pro (P)
- Carboxyl: asp (D), glu (E)
- Amine: lys (K), arg (R)
- Amide: asn (N), gln (Q)

Proteins

- Linear polymers of aa via amide linkages form peptides (1-10), polypeptides (11-100) and proteins (>100)
- Eg: Aspartame (2), glutathione (3), vasopressin (9), insulin (51)
- Proteins have a amino-end and carboxyl-end
- In the lab, proteins can be hydrolyzed (to aa) by strong acid treatment
- Physiologic hydrolysis by peptidases and proteases

Protein Structure

- 4 levels of protein structure
- **Primary** structure: aa sequence
- **Secondary** structure: regular chain organization pattern
- **Tertiary** structure: 3D complex folding
- **Quarternary** structure: association between polypeptides

Primary Structure

- Amino acid sequence determines primary structure
- Unique for each protein; innumerable possibilities
- Gene sequence determines aa sequence
- Each aa is called a residue; numbering (& synthesis) always from $–NH_2$ end toward –COOH end
- Amino acids covalently attached to each other by an amide linkage called as a peptide bond.

Peptide Bond

- Peptide bonds are planar (2 α-C and -O=C-N-H- in one plane)
- Partial double bond character due to resonance structures of peptide bond (bond length is 1.32 A^o instead of 1.49 A^o (single) or 1.27 A^o (double)
- Due to steric hindrance, all peptide bonds in proteins are in trans configuration
- The 2 bonds around the α-carbon have freedom of rotation making proteins flexible to bend and fold

Secondary Structure

- Secondary structure is the initial folding pattern (periodic repeats) of the linear polypeptide
- 3 main types of secondary structure: α-helix, β-sheet and bend/loop
- Secondary structures are stabilized by hydrogen bonds

The α-helix

- The α-helix is right-handed or clock-wise (for L-isoforms left-handed helix is not viable due to steric hindrance)
- Each turn has 3.6 aa residues and is 5.4 A^{o} high
- The helix is stabilized by H-bonds between –N-H and –C=O groups of every 4^{th} amino acid
- α-helices can wind around each other to form 'coiled coils' that are extremely stable and found in fibrous structural proteins such as keratin, myosin (muscle fibers) etc

β-Pleated Sheet

- Extended stretches of 5 or more aa are called β-strands
- β-strands organized next to each other make β-sheets
- If adjacent strands are oriented in the same direction (N-end to C-end), it is a parallel β-sheet, if adjacent strands run opposite to each other, it is an antiparallel β-sheet. There can also be mixed β-sheets
- H-bonding pattern varies depending on type of sheet
- β-sheets are usually twisted rather than flat
- Fatty acid binding proteins are made almost entirely of β-sheets

Bend / Loop

- Polypeptide chains can fold upon themselves forming a bend or a loop.
- Usually 4 aa are required to form the turn
- H-bond between the 1^{st} and 4^{th} aa in the turn
- Bends are usually on the surface of globular proteins
- Proline residues frequently found in bends / loops

Tertiary Structure

- 3D folding or ‘bundling up’ of the protein
- Non-polar residues are buried inside, polar residues are exposed outwards to aqueous environment
- Many proteins are organized into multiple ‘domains’
- Domains are compact globular units that are connected by a flexible segment of the polypeptide
- Each domain is contributes a specific function to the overall protein
- Different proteins may share similar domain structures, eg: kinase-, cysteine-rich-, globin-domains

Tertiary Structure

- 5 kinds of bonds stabilize tertiary structure: H-bonds, van der waals interactions, hydrophobic interactions, ionic interactions and disulphide linkages
- In disulphide linkages, the SH groups of two neighboring cysteines form a –S-S- bond called as a disulphide linkage. It is a covalent bond, but readily cleaved by reducing agents that supply the protons to form the SH groups again
- Reducing agents include β-mercaptoethanol and DTT

Quaternary Structure

- association of more than one polypeptides
- Each unit of this protein is called as a subunit and the protein is an oligomeric protein
- Subunits (monomers) can be identical or different
- The protein is homopolymeric or heteropolymeric
- Disulfide bonds usually stabilize the oligomer

AA sequence dictates protein structure

- Each protein has a unique and specific 3D structure that depends on the aa sequence. This is their native conformation.
- Denaturing agents such as urea or guanidinium chloride disrupt the 3D structure. This is called denaturation
- Denaturation is reversible. Removal of denaturants agents and sometimes, presence of a chaperones, is required for refolding
- Protein folding is a cooperative ‘all or none’ process

Prediction of Protein Structure

- Individual aa have a preference for specific 2^o structure
- α-helix (default): A, E, L, M, C
- β-sheets (steric clash): V, T, I, F, W, Y
- Bends: P, G, N
- No definite rules for 3^o structure. Determined by overall sequence and tertiary interactions between remote residues; decrease in free energy.
- Prediction based on computer calculations and comparison to similar domains of known structure

Post-Translational Modification of Proteins

- During synthesis proteins can incorporate only each of the 20 aa
- Many amino acids can be enzymatically modified after incorporation into proteins
- Reversible phosphorylation of S, T, Y serve as regulatory switches
- Amino-terminal aceytlation prevents degradation
- Glycosylation and fatty acylation makes proteins respectively more hydrophilic or hydrophobic
- Protein stability is enhanced by hydroxylation of P in collagen and carboxylation of E in prothrombin