Chemical nature of amino acids

The α-amino acids in peptides and proteins (excluding proline) consist of a carboxylic acid (–COOH) and an amino (–NH₂) functional group attached to the same tetrahedral carbon atom; this carbon is the α-carbon. Distinct R-groups, that distinguish one amino acid from another, also are attached to the alpha-carbon (except in the case of glycine where the R-group is hydrogen).

Draw the structures of the common amino acids
[Draw in structures from page 75 in textbook]

The characteristic properties of different R groups

~ Nonpolar; Aliphatic R Groups:
- Nonpolar and hydrophobic (water-insoluble)
- Alanine, valine, leucine and isoleucine have side chains that cluster together within proteins, stabilizing protein structure.
- Glycine is the simplest amino acid.
- Methionine contains a nonpolar thioether group in its side chain.
- Proline has an aliphatic side chain with a cyclic structure.

~ Aromatic R Groups:
- Nonpolar with all being able to participate in hydrophobic interactions.
- Phenylalanine, tyrosine and tryptophan.

~ Polar, Uncharged R Groups:
- More hydrophilic (soluble in water) because their functional groups form hydrogen bonds in water.
- Serine, threonine, cysteine, asparagine and glutamine.
- Disulfide-linked residues are strongly non polar. These bonds form covalent links between parts of a polypeptide molecule or between 2 polypeptide chains.

~ Positively Charged (Basic) R Groups:
- Extremely hydrophilic.
- Lysine, arginine and histidine have positive charge at pH 7.

~ Negatively Charged (Acidic) R Groups:
- Extremely hydrophilic.
- Aspartate and glutamate have a negative charge at pH 7.

Amino acids can act as an acid or a base

When an amino acid lacking an ionisable R group is dissolved in water at neutral pH, it exists in solution as the dipolar ion; or zwitterion, which can act as either a base or acid.
- Amphoteric substances have a dual acid-base nature.
- These are often called ampholytes

~ Weak acid:
A weak acid is one that does not ionise completely; it is a poor proton donor.

Titration curves reveal the pKₐ of weak acids
- A measured volume of an acid is titrated with a solution of strong base (usually NaOH), of known concentration. The NaOH is added in small increments until the acid is neutralized.
At the mid-point of the titration curve, the pH of the equimolar solution of the acid and its deprotonated form is equal to the $pK_a$ of the weak acid.

Titration curves predict the electric charge of amino acids
- At the point of inflection between the two stages in its titration curve, the amino acid is fully ionised with no net electric charge. (Isoelectric point)
- An amino acid will have a net negative charge at any pH above its $pI$.
- An amino acid will have a net positive charge at any pH below it $pI$.

Buffer systems
- Buffers are aqueous systems that tend to resist changes in pH when small amounts of acid ($H^+$) or base ($OH^-$) are added.
- A buffer system consists of a weak acid (proton donor) and its conjugate base because the solution has the ability of being able to resist dramatic changes in pH upon addition of either a strong acid or a strong base. The conjugate base $A^-$ will react with any strong acid added while the weak acid $HA^+$ will react with any strong base added.
- The buffering power of the system is maximal at the mid-point, where the concentrations of the conjugate base and the remaining weak acid are the same.

~ Amino acids, peptides and proteins have buffering capacity:
- Buffering capacity is the resist in pH change when acid or base is added to the solution, compared to when the acid or base is added to water.
- Amino acids have both amine groups ($-NH_2$) and acid groups ($-COOH$). If acid is added, the protons ($H^+$) will combine with the amine group forming $-NH_3^+$. If base is added a proton will be removed from the acid group, leaving $-COO^-$. Thus, since both an acid and base are on the same molecule pH change will be resisted when acid or base solutions are added.

$pK_a$ and $pK_b$ of amino acids
The -COOH and -NH$_2$ groups in amino acids are capable of ionizing. Thus the $pK_1$ value is of acid group, and $pK_2$ value is of the amino group.

What does $pI$ refer to?
This is the isoelectric point; it is the pH at which a protein carries no net charge.

\[ pI = \frac{(pK_1 + pK_2)}{2} \]

- Tyrosine, cysteine, positively charged R groups and the negatively charged R groups all have a $pK_R$

Net charge on a protein and environmental change
The net charge of any amino acid, peptide or protein, will depend upon the pH of the surrounding aqueous environment. As the pH of a solution of an amino acid or protein changes so too does the net charge. This is because with such a change the amino group may accept a proton or the carboxyl group may donate a proton.
- At low pH: COOH, NH$_3^+$
- At high pH: COO$, NH_2$
Structure of a peptide
- Peptides are chains or amino acids.
- Two amino acid molecules can be covalently joined through a substituted amide linkage, termed a peptide bond, to yield a dipeptide.
- Peptide bond formation is a condensation reaction; formed through the dehydration of water from the carboxyl group of one amino acid and the amino group of another.
- Oligopeptide is the structure of a few amino acids joined together. Whereas polypeptide is the term used to describe many amino acids joined.
- An amino acid unit in a peptide is called a residue; the part left over after losing a hydrogen atom from its amino group and the hydroxyl moiety from its carboxyl group.
- In a peptide, the residue at the end with a free α-amino group is the amino or N-terminal. The residue at the end of the free carboxyl group is the carboxyl or C-terminal.

Describe simple experimental techniques to separate mixtures of peptides and proteins based on charge differences between molecules
~ Ion-exchange chromatography:
- Exploits differences in the sign and magnitude of the net electric charge of proteins at a given pH.
- The column matrix is a synthetic polymer (resin) containing bound charged groups; anionic groups are cation exchangers and cationic groups are anion exchangers.

~ Size-exclusion chromatography (gel filtration):
- Proteins are separated by size, with large proteins emerging from the column sooner than smaller ones.

~ Affinity chromatography:
- Proteins separated based on binding affinity.
- Beads in the column have ligands attached, and any protein with affinity for this ligand binds to the beads, and its migration through the matrix is slowed.

~ Electrophoresis:
- Carried out in gels made up of cross-linked polymer polyacrylamide. This gel acts as a molecular sieve, slowing the migration of proteins approximately in proportion to their charge-to-mass ratio.

2. Protein Structure

Describe protein structure in terms of primary, secondary, tertiary, quaternary levels of protein structure
- Primary structure consists of a sequence of amino acids linked together by peptide and disulfide bonds.
- Secondary structure refers to the stable arrangements of amino acid residues, giving rise to recurring structural patterns.
- Tertiary structure describes all aspects of the 3D folding of a polypeptide.
- Quaternary structure refers to the arrangement of 2 or more polypeptide subunits.

Define how polypeptides can come in a range of sizes and composition
~ Multisubunit protein – Two or more polypeptides that are associated non-covalently. The individual chains may be identical or different
~ Oligomeric protein – At least two are polypeptide chains are identical.
Protomer – Identical units of 1 or more chains
Understand the general principles for determining amino acid sequences of proteins
- 1953, James Watson and Francis Crick deduced the double helix structure of DNA.
- 1953, Frederick Sanger worked out the sequence of amino acid residues in the polypeptide chains of the hormone insulin.

Short polypeptide sequencing:
1. Reagent 1-flouro-2, 4-dinitrobenzene labels and identifies the amino-terminal amino acid residue.
2. The polypeptide is hydrolysed in 6M HCl to its constituent amino acids and the labelled amino acid is identified.
3. Edman degradation sequences an entire polypeptide through moving only the amino-terminal residue from a peptide, leaving all other bonds intact. The reacted peptide converts the amino-terminal amino acid to a phenylthiocarbamoyl (PTC) adduct. The peptide bond next to the PTC adduct is then cleaved in acid, with the removal of the amino-terminal amino acid as an anilinothiazolinone derivative. The derived amino acid is extracted with organic solvents, converted to the more stable phenylthiohydantoin derivative by treatment with aqueous acids, are then identified.

Large polypeptide sequencing:
- Sequenced in smaller segments.
- The accuracy of amino acid sequencing declines as the length of the polypeptide increases.
1. Disulfide bonds interfere with enzymatic and chemical cleavage of the polypeptide. Oxidation of a cysteine residue with performic acid produces 2 cysteic acid residues. Reduction by dithiothreitol or β-mercaptoethanol to form Cys residues must be followed by carboxymethylation by iodoacetate.
2. Cleaving the polypeptide chain through enzyme and chemical methods. Enzymes called proteases catalyse the cleavage of peptide bonds. Some only cleave the peptide bond adjacent to particular amino acid residues. The digestive enzyme trypsin catalyses the peptide bonds in which the carbonyl group is contributed by either a Lys or Arg residue.
3. Ordering of the peptide fragments occurs; with overlapping peptides obtained in 2nd fragmentation yield the correct order of the peptide fragments in the 1st.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Cleavage points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trypsin</td>
<td>Lys, Arg (C)</td>
</tr>
<tr>
<td>Chymotrypsin</td>
<td>Phe, Trp, Tyr (C)</td>
</tr>
<tr>
<td>Asp-N-protease</td>
<td>Asp, Glu (N)</td>
</tr>
<tr>
<td>Pepsin</td>
<td>Leu, Phe, Trp, Tyr (N)</td>
</tr>
<tr>
<td>Endoproteinase</td>
<td>Lys (C)</td>
</tr>
<tr>
<td>Cyanogen bromide</td>
<td>Met (C)</td>
</tr>
</tbody>
</table>

Describe the type of information which can be obtained from comparisons between amino acid sequences and explain why this information is useful.
- Knowledge from amino acid sequences in a protein can offer insights into its 3D structure, its function, cellular location and evolution.
Explain, with examples, the terms 'homologous proteins' and 'isoforms'.
- Homologous proteins are those having similar sequences and functions in different species.
  E.g. Haemoglobin
- Isoforms are different forms of the same protein.
  E.g. G-actin, Creatine Kinase

Describe the interactions that stabilise the conformation of a protein
- Protein structure is stabilized by multiple weak interactions.
- Hydrophobic interactions are largely buried in the protein interior, away from the water.
- Number of hydrogen bonds and ionic interactions within the protein is maximised, thus reducing the number of hydrogen-bonding and ionic groups that are not paired with a suitable partner.

Explain why the peptide group is planar and why this is important for the formation of regular secondary structures
- The α carbons of adjacent amino acid residues are separated by 3 covalent bonds, arranged as Cα—C—N—Cα.
- A small electric dipole exists because of the partial O2−Ve charge and partial N +Ve charge.
- O2 atom is positioned trans to the hydrogen atom of amide nitrogen.
- The peptide C—N bonds cannot rotate freely because of their partial double-bond.
- There is free rotation about the N—Cα and the Cα—C bonds, thus the planar peptide bond has 2 degrees of rotational freedom.
- The 2 rotation angles are known by the greek letters φ (phi) and ψ (psi). φ is the rotation angle about the N—Cα bond and ψ is the rotation angle about the Cα—C bond.
- It is these two degrees of rotational freedom that allows polypeptides to fold up into unique conformations.

Describe the characteristic molecular features of the α-helix, β-sheet (parallel and antiparallel) and β-turns
~ α-helix:
- Polypeptide backbone is tightly wound around an imaginary axis drawn longitudinally through the middle of the helix.
- The R groups protrude outward from the helical backbone.
- The repeating unit is a single turn of the helix, which extends 5.4 Å along the long axis.
- Conformations with φ = -57° and ψ = -47°.
- Each helical turn includes 3.6 amino acid residues and is right-handed.
- Within the helix, every peptide bond participates in hydrogen bonding, and this stabilises the structure.

~ β-sheet:
- Polypeptide backbone is extended into a zigzag shape.
- The zigzag chains can be arranged side by side to form a structure resembling a series of pleats.
- H-bonds form between adjacent segments of the polypeptide chain.
- R groups of adjacent amino acids protrude from the zigzag structure in opposite directions.
- Parallel or antiparallel refers to having the same or opposite amino-to-carboxyl orientations, respectively.
- Repeat period for parallel conformations is shorter (6.5 Å compared to 7 Å).
- Parallel φ = -119° and ψ = 113° and antiparallel φ = -139° and ψ = 135°.

~ β-turns:
- Almost 1/3 of amino acid residues, in a compacted folded structure, are in turns or loops.