

## \*\*Amino Acids, Peptides, and Proteins\*\*

### Amino Acids Found in Proteins

- Amino acids contain two functional groups, an amino group (-NH<sub>2</sub>) and a carboxyl group (-COOH) which are formed on a carboxylic acid.
  - Alpha amino acids are those where both functional groups are attached to the same carbon.
  - Side chain/R group and hydrogen atom are also attached to alpha carbon.
    - Side chain determines the properties of amino acids and its functions.

### Terminology

- There are a variety of different types of amino acids, however the MCAT only focuses on the 20 alpha-amino acids that are encoded by the human genetic code.
  - Called the **proteinogenic amino acids**

### Stereochemistry of Amino Acids

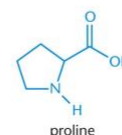
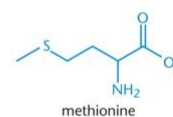
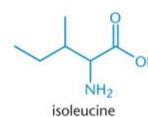
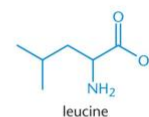
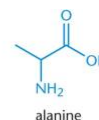
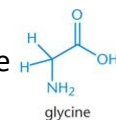
- Alpha carbon is usually a chiral center since it has four different groups attached to it.
  - Thus most amino acids are optically active
    - **Glycine** is the only exception since it has a hydrogen as its R group.
- All chiral amino acids are L-amino acids, which means that the amino group is drawn on the left side for the Fischer projection.
  - Translates to an (*S*) absolute configuration for almost all chiral amino acids.
    - **Cysteine** is the only amino acid that has an L-amino acid configuration but has an (*R*) absolute configuration. This is because the carboxyl group is not the highest priority functional group.

### Structures of the Amino Acids

Amino acids can be classified by the structure of their side chains.

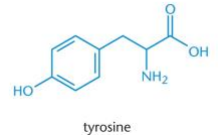
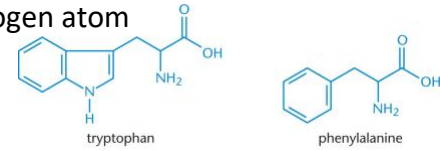
#### Nonpolar, Nonaromatic Side Chains – 7 amino acids

- **Glycine**: has a single hydrogen bond as its side chain and is the smallest amino acid.
- **Alanine, valine, leucine, and isoleucine**: have alkyl side chains containing one to four carbons
- **Methionine**: one of only two amino acids with a sulfur attached to it.
  - Sulfur does not induce polarity since it has a methyl group attached to it.
- **Proline**: forms a *cyclic* amino acid. Amino nitrogen becomes a part of the side chain through the formation of a five-membered ring.
  - Ring places constraints on the flexibility of proline



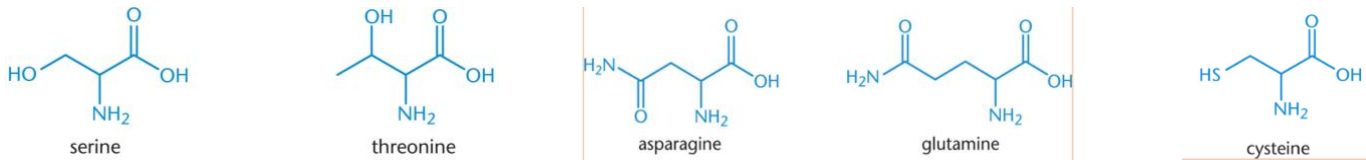
### Aromatic Side Chains – 3 amino acids

- **Tryptophan:** largest and has a double ring system that contains a nitrogen atom
- **Phenylalanine:** smallest and has a benzyl side chain. Non polar.
- **Tyrosine:** -OH group is added to phenylalanine. Polar



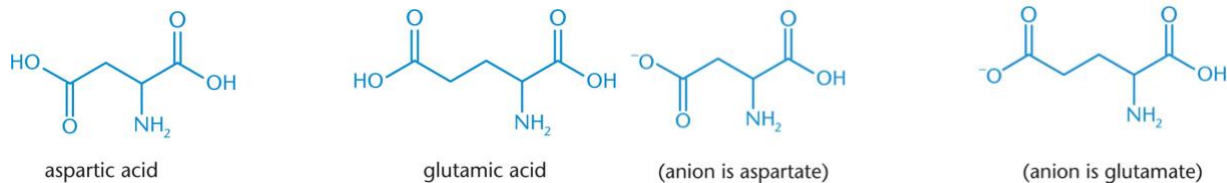
### Polar Side Chains – 5 amino acids

- **Serine and threonine:** have -OH groups which makes them highly polar.
- **Asparagine and glutamine:** amide side chains
  - Amide Nitrogen's do not gain or lose protons with changes in pH, so they do not become charged
    - This is the opposite of what amino group nitrogen's do.
- **Cysteine:** **thiol** (-SH) group in its side chain. Since sulfur is larger and less electronegative than oxygen, the S-H bond is weaker than the O-H bond and is thus more prone to oxidation.



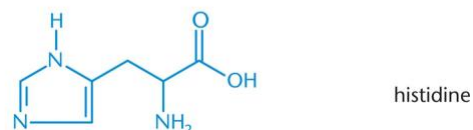
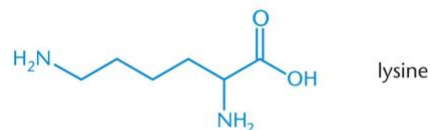
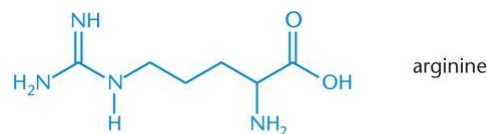
### Negatively Charged (Acidic) Side Chains – 2 Amino acids

- **Aspartic Acid (aspartate):** deprotonated form of aspartic acid
- **Glutamic acid (glutamate):** deprotonated form of glutamic acid



### Positively Charged (Basic) Side Chains – 3 amino acids

- All three have a positively charged nitrogen atom
- **Lysine:** terminal primary amino group
- **Arginine:** three nitrogen atoms in its side chain with the positive charge delocalized over all three.
- **Histidine:** aromatic ring with two nitrogen atoms (ring is called **imidazole**)



## Hydrophobic and Hydrophilic Amino Acids

- Amino acids with long alkyl side chains – alanine, isoleucine, leucine, valine, and phenylalanine – are strongly hydrophobic, and are thus more likely to be found in the interior of proteins
- All amino acids with charged side chains – histidine, arginine, lysine, glutamate, and aspartate – are hydrophilic
  - So are the amides asparagine and glutamine
- Remaining amino acids are somewhere in the middle between being philic and phobic.

## Amino Acid Abbreviations

- Expected to identify an amino acid by the name, three-letter abbreviation and one-letter abbreviations

One letter code	Three letter code	Name
A	ALA	Alanine
R	ARG	Arginine
D	ASP	Aspartic acid
N	ASN	Asparagine
C	CYS	Cysteine
E	GLU	Glutamic acid
Q	GLN	Glutamine
G	GLY	Glycine
H	HIS	Histidine
I	ILE	Isoleucine
L	LEU	Leucine
K	LYS	Lysine
M	MET	Methionine
F	PHE	Phenylalanine
P	PRO	Proline
S	SER	Serine
T	THR	Threonine
W	TRP	Tryptophan
Y	TYR	Tyrosine
V	VAL	Valine

## Acid-Base Chemistry of Amino Acids

- Amino acids are an **amphoteric species** since they have an acidic carboxyl group and a basic amino group.
- Ionizable groups tend to gain protons under acidic conditions and lose them under basic conditions.
  - i.e. at low pH an ionizable group will be protonated

- The  $pK_a$  of a group is the pH at which half of the molecules of the species are deprotonated or  $[HA]=[A^-]$ 
  - If pH is lower than  $pK_a$  then the majority of the species will be protonated.

### Protonation and Deprotonation

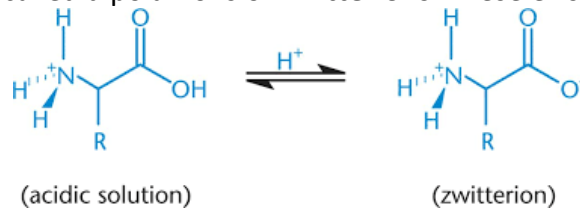
- All amino acids have at least two  $pK_a$  values.
  - $pK_{a1}$  is the  $pK_a$  of the carboxyl group and is usually around 2
  - $pK_{a2}$  is the  $pK_a$  of the amino group and is usually around 9-10
- If the amino acid has an ionizable side chain, then there will be three  $pK_a$  values.

### Positively Charged under Acidic Conditions

- At pH below 1, the  $pK_a$  is far below that of the amino group, so the amino group is fully protonated ( $-NH_3^+$ )
  - Additionally, the carboxylic acid group is fully protonated ( $-COOH$ )
- At very acidic pH values, amino acids tend to be positively charged.

### Zwitterions at Intermediate pH

- At pH of 7.4, carboxylic acid  $pK_a$  has been moved past. As such, you will not find amino acids with their carboxylate group protonated and the amino group unprotonated
  - Amine group stays protonated since the  $pK_a$  still is not high enough.
- Resulting molecule has both a positive charge and a negative charge, but is overall neutral.
  - These are called dipolar ions or **zwitterions**. These exist in water as internal salts.



### Negatively Charged under Basic Conditions

- $pK_a$  of the amino group is below the  $pK_a$  at higher pH's  $>10$ , which means that the amino group deprotonates to  $NH_2$
- So the molecules become negatively charged at high pH

### Titration of Amino Acids

- The titration curve should look like a combination of two monoprotic acid curves or three curves if the side chain is charged.
- Starting at low pH, the amino acid is fully protonated. As the pH approaches  $pK_{a1}$ , the solution begins acting like a buffer and this is characterized by a straight line on the graph.
  - When  $pH = pK_{a1}$ , then  $[HA]=[A^-]$
- **Isoelectric Point (pI)**: the pH at which the molecule is electrically neutral which means that the amino acid exists exclusively in its zwitterion form.
  - Can be calculated for neutral amino acids by the following:

$$pI_{neutral\ amino\ acid} = \frac{pK_{a,1} + pK_{a,2}}{2}$$

- Pass through a secondary buffer phase as more base is added since the amino group begins to deprotonate.

## Amino Acids with Charged Side Chains

- Amino acids with charged side chains – like glutamic acid and lysine – the titration curve has an extra step, but follows the same principles.
- For an amino acid such as glutamic acid, the two carboxyl groups mean that the deprotonated form will still have a charge of +1.
  - The first proton will be deprotonated from the carboxyl groups, while the second proton will instead be protonated from the carboxyl side group.

- Usually results in a **much lower** isoelectric point

$$pI_{\text{acidic amino acid}} = \frac{pK_{a,\text{COOH group}} + pK_{a,\text{R group}}}{2}$$

- Lysine has two amino groups and one carboxyl group, so its charge is +2 in its fully protonated state.
  - First proton is lost from carboxyl group as usual, but second proton to make the molecule neutral is not lost until a pKa of around 9.

- As such, the isoelectric point is usually **much higher**

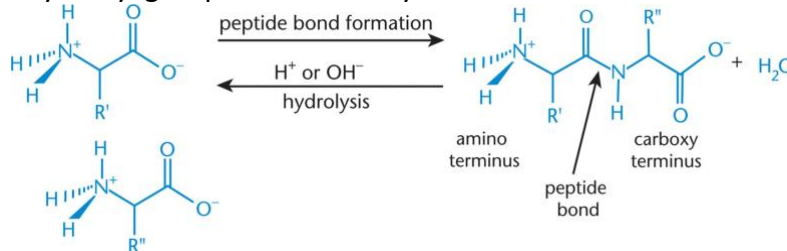
$$pI_{\text{basic amino acid}} = \frac{pK_{a,\text{NH}_3^+} + pK_{a,\text{R group}}}{2}$$

## Peptide Bond Formation and Hydrolysis

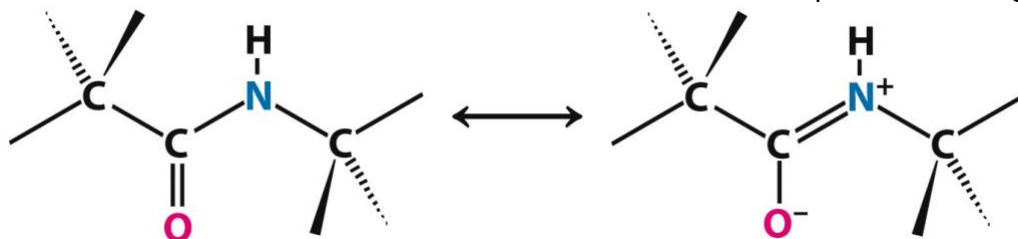
- Peptides: consist of amino acid subunits called **residues**
  - Oligopeptide: used for small peptides with up to 20 residues
  - Polypeptides: longer than 20 chain residues.
- Residues are joined together through **peptide bonds** which is a specialized form of an amide bond that forms between the  $-\text{COO}^-$  group of one amino acid and the  $\text{NH}_3^+$  group of another amino acid.

### Peptide Bond Formation

- **Condensation/dehydration** reaction since it results in the removal of a water molecule.
- Can also be viewed as an **acyl substitution reaction**
- Peptide bond formation is conducted when the electrophilic carbonyl carbon on the first amino acid is attacked by the nucleophilic amino acid on the second amino acid.
  - The hydroxyl group of the carboxylic acid is kicked off



- The amide group can experience resonance since the amide groups have delocalizable pi electrons in the carbonyl and in the lone pair nitrogen.
  - **C-N bond in the amide has partial double bond character**
  - Rotation around this bond is restricted which makes the protein more rigid.



- Free amino end is known as the **N-terminus** and the free carboxyl end is the **C-terminus**.
  - Peptides usually drawn with the N-terminus on the left and C-terminus on the right

### Peptide Bond Hydrolysis

- Amides can be hydrolyzed using acid or base catalysis, but this is usually for the lab
- In living organisms, hydrolysis is catalyzed by hydrolytic enzymes such as *trypsin* and *chymotrypsin*.
  - These work by breaking apart the amide bond by adding a hydrogen atom to the amide nitrogen and an OH group to the carbonyl carbon.

## Primary and Secondary Protein Structures

- Proteins can range in size from a few amino acids in length up to many thousands
- Have four levels of structures: **primary, secondary, tertiary, and quaternary**.

### Primary Structure

- This is the linear arrangement of amino acids coded in an organism's DNA.
- The sequence of amino acids which are listed from the N-Terminus to the C-terminus.
- Encodes all the information needed for folding at all of the higher structural levels.
- Can be determined by **sequencing**

### Secondary Structure

- Local structure of neighboring amino acids
- Result from hydrogen bonding between nearby amino acids
- Two most common:  **$\alpha$ -helices &  $\beta$ -pleated sheets**. This results in added stability due to the hydrogen bonding.

### Alpha-Helices

- Rod-like structure in which the peptide chain coils clockwise around a central axis
- Helix is stabilized by intramolecular hydrogen bonds between a carbonyl oxygen atom and an amide hydrogen atom which is four residues down the chain.
- Side chains of the amino acids point away from the helix core
- This helix is important for the structural integrity of **keratin**.

### Beta-Pleated Sheets

- Peptide chains lie alongside one another which forms rows or strands that are held together by intramolecular hydrogen bonds between carbonyl oxygen atoms on one chain and amide hydrogen atoms in an adjacent chain.
- Pleated shape is used to accommodate as many hydrogen bonds as possible.
- R groups of amino residues point above and below the plane of the beta-pleated sheets.
- *Fibroin* is the primary component used in silk and utilizes beta pleated sheets.

### Secondary Structures and Proline

- Proline has a rigid cyclic structure which will introduce a kink in the peptide chain when it is found in the middle of an alpha helix or pleated sheet
- Proline is thus rarely found in the middle of alpha helices or pleated sheets, unless the helices cross a cell membrane.

- Often found in the turns between chains of the pleated sheet or found as the residue at the start of an alpha helix.

## Tertiary and Quaternary Protein Structures

- Proteins can broadly be divided into **fibrous** proteins (structure that resemble sheets or long strands) or **globular** proteins (spherical structure)
  - Distinctions are caused by tertiary and quaternary protein structures which result from protein folding.

### Tertiary Structure

- The three dimensional shape of a protein. These are mostly determined by the hydrophilic and hydrophobic interactions between R groups of amino acids
  - Hydrophobic residues: tend to be on the interior of proteins since it reduces their proximity of water
  - Hydrophilic Bond: N-H and C=O bonds get pulled by the hydrophobic residues.
    - Can then form electrostatic interaction and hydrogen bonds that stabilize the protein from the inside.
  - Results in most of the amino acids on the surface being hydrophilic (polar or charged) R groups.
- Disulfide Bonds: bond that forms when two **cysteine** molecules become oxidized to form **cystine**
  - Create loops in the protein chain
  - Requires the loss of two protons and two electrons
- Basic idea is that secondary structures form first and then hydrophobic interactions and hydrogen bonds cause the protein to “collapse” into its proper 3-d structure.
  - Molten globules: intermediate states between 2ndary and tertiary structures.
- Denaturation: when a protein loses its tertiary structure, result in loss of functionality.

### Folding and the Solvation Layer

- Solvation Layer: occurs when a solvent dissolve in a solute. The nearby solvent molecules form a layer around the solute.
- When a hydrophobic side chain is placed in aqueous solution, the water molecules cannot form hydrogen bonds with the side chains and the water molecules are thus forced to rearrange themselves in order to maximize bonding.
- This results in an increase in order which leads to a decrease in entropy which then makes the reaction less spontaneous.
- Ultimately a protein achieves maximum stability by putting hydrophobic residues away from water molecules and hydrophilic molecules towards water molecules.

### Quaternary Structure

- Unlike the previous three, not all proteins have a quaternary structure.
- Exists for proteins that contain more than one polypeptide chain
- Quaternary structure is an aggregate of smaller globular peptide or **subunits**, which represents the functional form of the protein.
- Can be more stable since they can reduce the surface area of the protein complex.

- Can reduce the amount of DNA needed to encode the protein complex.
- Can bring catalytic sites close together, which allows intermediates from one reaction to be directly shuttled to a second reaction.
- Can induce **cooperativity or allosteric effects**.
  - One subunit undergoes structural changes which either enhances or reduces the activity of other subunits.

### Conjugated Proteins

- Derive functionality from covalently attached molecules called **prosthetic groups**.
  - Prosthetic groups can be organic molecules or metal ions
  - Lipoproteins: lipid prosthetic group
  - Glycoprotein: carbohydrate prosthetic group
  - Nucleoproteins: nucleic acid prosthetic groups
- These prosthetic groups have a major role in determining the function of their respective proteins.
  - In hemoglobin, each subunit contains a prosthetic group called **heme** which has iron in it to bind to oxygen.
  - Can also direct the protein to be delivered to a certain location

### Denaturation

- Protein loses its three dimensional structure.
- Often irreversible, and these proteins cannot catalyze reactions
- When temperature of a protein increases, its average kinetic energy increases as well
  - If temperatures become high enough, the added energy can be enough to overcome the hydrophobic interactions that hold a protein together
- Solutes such as urea can also denature proteins by directly interfering with the forces that hold the protein together
  - Disrupt the tertiary and quaternary structures by breaking disulfide bridges or by overcoming hydrogen bonds and other side chain interactions