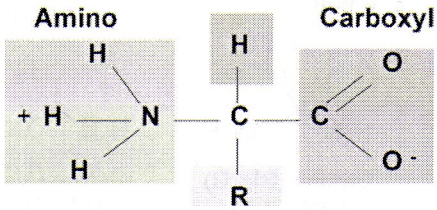
1. AMINO ACIDS CONVENTION



- All amino acids have the generic backbone structure consisting of:
  - A carboxylic acid
  - An amino group
  - A variable R group
  - All connected by a chiral α-Carbon
- All biological amino acids are L-configured (unlike sugars, which are D)
- Exception to convention: Proline
  - The α-amino group interacts with the R-group in a cyclic fashion.
  - They make up a singular ring structure.

- Generally:
  - pK_a of a generic primary amino group is ~ 9.4.
  - pK_a of a generic primary carboxylic group is ~2.2
  - You need to memorize individual pK_a's of the R groups on your own.
  - Each R group distorts the pKa of all the other groups, but they are relatively similar.
- The pKa of the carboxylic and amine group are lower than they should be in amino acids
  - (generally carboxyl groups have pKa of 5)
  - This is due to a mutual negative inductive effect as the electronegative atoms pull on the electrons from the other group – helping stabilize the charged form.

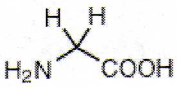
- Ionization: conversion of an atom or molecule into an ion (a charged particle).
  - This is achieved by adding or removing charged particles (i.e. electrons or protons (H^+))
- Ampholyte: a molecule that holds both pos. and neg. charges.
  - They do not necessarily need to be net neutral.
- Zwitterion: a molecule that has both a pos. and neg. charge at neutral pH; a net neutrality is achieved as a result.
  - All zwitterions are ampholytes, but not all ampholytes are zwitterions.
- Isoelectric point (pI): the pH at which a molecule theoretically has net neutrality – pH at which a zwitterion exists.
  - Molecules below this pH generally have a net pos. charge.
  - Molecules above this pH generally have a net neg. charge.
  - The isoelectric point is the midpoint between 2 pKa's, in which there is one pos. and one neg. charge.
- Some of the amino acids have different groups that can also be ionized (charged polar a.a.'s)
  - These ampholytes typically have an isoelectric point not equal to the 5.5ish range as seen in some other a.a.'s.
  - E.g. Lys has one α-carboxylic acid, an α-amino group, and an ε-amino group.

1.1 AMINO ACID CODES:

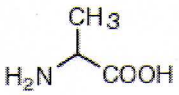
- You need to know the name, structure, one/three letter code, pI and R-pKa for each amino acids used in humans.
  - By convention, 3-ltr codes are used for short peptides
  - Long peptides and proteins use 1-ltr codes.
- A.a.'s are synthesized by pathways present in plants, and microorganisms – Plants can synthesize all their a.a.'s
- Humans cannot synthesize 10 a.a.'s – we lack the biosynthesis capabilities.
  - These are grouped into a category known as essential a.a.'s
  - The only way to get this in our system is to consume them with our food.
  - Essential a.a.'s are:
    - R, H, I, L, K, M, F, T, W, and V.
- Failure to obtain any of these 10 ess.a.a.'s results in catabolism of bodily proteins to use as a source of them = bad.
- The rest of the a.a.'s (A, N, D, C, E, Q, G, P, S, Y) are non-essential a.a.'s and can be made by humans.
  - They can be a *denovo* synthesis – made without use of a precursor.
  - Or, they are a semi-synthetic product – made using an ess.a.a. as a precursor.

1.2 AMINO ACIDS TO KNOW:

Small

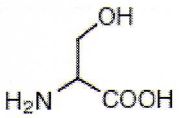


Glycine (Gly, G)
MW: 57.05

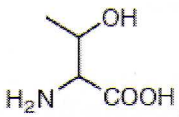


Alanine (Ala, A)
MW: 71.09

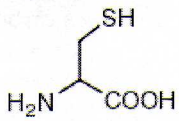
Nucleophilic



Serine (Ser, S)
MW: 87.08, pK_a ~ 16

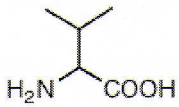


Threonine (Thr, T)
MW: 101.11, pK_a ~ 16

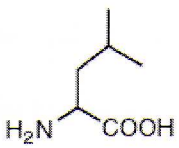


Cysteine (Cys, C)
MW: 103.15, pK_a = 8.35

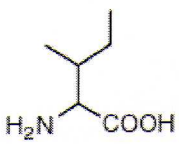
Hydrophobic



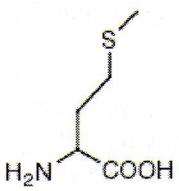
Valine (Val, V)
MW: 99.14



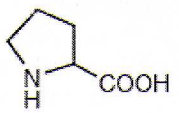
Leucine (Leu, L)
MW: 113.16



Isoleucine (Ile, I)
MW: 113.16

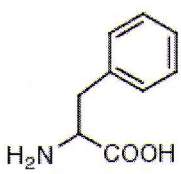


Methionine (Met, M)
MW: 131.19



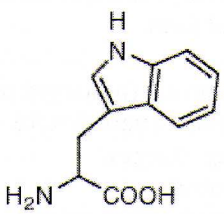
Proline (Pro, P)
MW: 97.12

Aromatic



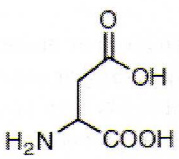
Phenylalanine (Phe, F)
MW: 147.18

Tyrosine (Tyr, Y)
MW: 163.18

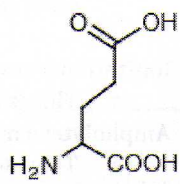


Tryptophan (Trp, W)
MW: 186.21

Acidic

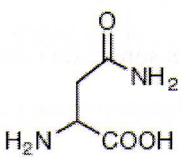


Aspartic Acid (Asp, D)
MW: 115.09, pK_a = 3.9

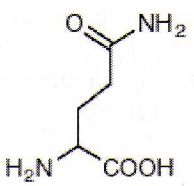


Glutamic Acid (Glu, E)
MW: 129.12, pK_a = 4.07

Amide



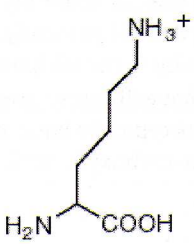
Asparagine (Asn, N)
MW: 114.11



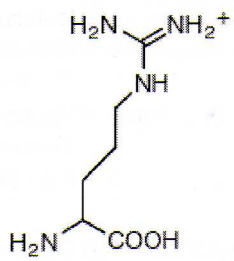
Glutamine (Gln, Q)
MW: 128.14

Basic

Histidine (His, H)
MW: 137.14, pK_a = 6.04



Lysine (Lys, K)
MW: 128.17, pK_a = 10.79



Arginine (Arg, R)
MW: 156.19, pK_a = 12.48

Amino Acid Nomenclature

- All the carbons following the α-carbon in the R group is named based on the order of the Greek alphabet.
  - 2nd carbon = β-carbon (beta)
  - 3rd carbon = γ-carbon (gamma)
  - 4th carbon = δ-carbon (delta)
  - 5th carbon = ε-carbon (epsilon)
  - 6th carbon = ζ-carbon (zeta), and so on...

2. AMINO ACID CATEGORIES

- The 20 standard a.a.'s are classified according to polarity of their side chains (R-groups)

2.1 NON-POLAR AMINO ACIDS

Aliphatic R groups:

- The R groups are non-polar and therefore hydrophobic in nature.
- Aliphatic R-groups such as Gly, Ala, Val, Leu, and Ile are non-polar.
    - Aliphatic groups are fairly stable and non-reactive and are rarely involved in post-translational mods.
- These a.a.'s do not carry a dipole moment within their R-groups.
- Hydrophobicity increases with the number of C atoms present on the hydrocarbon chain in the R-group.
    - As length of carbon chain increases, it becomes more hydrophobic.
    - So, the increasing order of hydrophobicity is: G < A < V < L = I
        - Note: Leucine and Isoleucine have comparable hydrophobicity
- Typically, it is more stable for aliphatic a.a. to be oriented within the structure of a protein.
    - These a.a.'s will often be positioned in a way to avoid water molecules and form hydrophobic regions.
- Gly and Ala are "ambivalent" due to their small R groups, and can appear both inside and outside of the protein.
    - Their weak hydrophobicity allows for this property.
- Note: Ile has a chiral centre found within their R-group as well.
- Contribution of pI of these amino acids is 1º-ily determined by α-carboxylic and α-amino group.
    - Their pI values range from 5.97 to 6.02 as a result.
- Mnemonic: **G**et **A** **V**acation of **L**ovely **I**slands

Sulfuric R groups:

- There are only 2 sulfur containing a.a.'s (M and C).
- It is also important to know that sulfur has an electronegative value close to that of Carbon.
- Cysteine does have ionizable capacity to yield a thiolate anion, and can also be considered an uncharged polar group.
    - However, realistically, Cys is usually involved in disulfide bridges, which are a lot less reactive
    - The bond between 2 same atoms eliminate the dipole moment, yielding a non-polar property in most practical applications.
    - Cys has a pI of 5.07 as its sulfhydryl group (-SH) has a pKa of 8.37
        - This makes it slightly acidic when it is free in solution.
- Met contains a thiol ether and is the most hydrophobic of all common a.a.'s, and a pI of 5.74.
    - Therefore, almost exclusively localized to protein interiors.
    - It is the only sulfur containing a.a. with a highly non-polar side chain.
    - Fun fact: During mRNA translation, Met is the initiating a.a. of the peptide chain.
- Sulfur containing groups are more reactive than aliphatic groups and do participate in many post-trascriptional mods
    - E.g. Met and Cys can be oxidized to sulfoxyde or even as sulfone.
    - Cysteine almost always dimerizes with another cysteine to form Cystine – a disulfide bridge.
        - Disulfide bridges can occur intramolecularly, or intermolecularly.
- Mnemonic: **M**arvelous **C**arribean – Get A Vacation on Lovely Islands.

Cyclic R groups:

- Proline is the only cyclic secondary amino acid.
- Conformational constraints are imposed by the cyclic pyrrolidine side group merging with the α-amino group.
- It is also ambivalent and can be found equally inside and outside of protein structures.
- Pros are highly used in proteins that have sharp bends, as it can cause kinks and bends in primary sequence.
- Mnemonic: Marvelous Carribean – Get A Vacation on Lovely **P**aradise Islands.

Aromatic R groups:

- There are 3 aromatic amino acids (Trp, Phe, Tyr), 2 of which are considered non-polar (Trp and Phe)
- A.a.'s with aromatic rings (regardless of hydrophobicity) can participate in:
  - π stacking
  - H → πelectron interactions
  - Cations → πelectron interactions
- Aromatic rings (like benzene) have electron cloud densities above and below the plane of the molecule, and stabilized by the slight positive dipoles on the plane of the ring (H and nodal C's)
  - This property attracts pos. charges to the electron cloud, causing an interaction (H and cation interactions)
  - The H interaction can also be thought of as a weaker kind of H-bond.
- π stacking can also occur, where phenyl groups lie on top of one another, like a multi-layer sandwich.
  - The mechanism as to how this occurs is currently unknown, and is intensively being researched.
  - Fun fact: π stacking also occurs in nucleic acids to stabilize and hold N. bases in a certain orientation.
- Mnemonic: Marvelous Carribean – Get A Vacation on Lovely Paradise Islands **W**ith **F**riends.

2.2 UNCHARGED POLAR AMINO ACIDS

Hydroxyl containing R groups:

- Tyr, Ser, and Thr are a.a.'s with a hydroxyl (-OH) group attached – Tyr is aromatic.
- The R pKa of Tyr is 10.46, so it remains an uncharged polar group under physiological conditions.
- The R pKa of Ser and Thr are very high (>14), and can be considered non-ionizing.
- Note: Thr also contains a chiral centre within its R group.

Amide containing R groups:

- Asn and Gln are the only a.a.'s with an amide R-group.
  - Amides are highly stable, non reactive functional groups, so they are not ionizable.
  - They are also highly hydrophilic and will almost always face the aqueous environment.

Sulfur containing R groups:

- Cys is an ambiguous a.a., in which some texts say its non-polar, and some say it is polar.
- Free-floating Cys in an aqueous solution IS POLAR.
- Cys involved in peptide links almost always have a disulfide bridge interactions making it NON-POLAR.
  - $Cys + Cys \xrightarrow{Oxidation} Cystine$
- Again, disulfide bridges can link 2 different pptd chains or cross-link Cys intramolecularly.

2.3 CHARGED POLAR AMINO ACIDS

Basic R groups:

- Lys, Arg, and His are polar, basic amino acids, that contain an ionizable amine in their R group.
- Due to their highly hydrophilic nature, they are most often found on protein surfaces.
- The hydrocarbon character of Lys locates it close to protein surfaces with its amino group in aqueous contact.
- Basic a.a. are pos. charged when pH is below the R-pKa value.

- At physiological pH, they are almost always found in their charged state.
  - Exception is His – it is neutral in the alkaline ranges of physiological pH (>pH7.0).
  - Lys and Arg remain fully protonated (charged) at pH 7, but only 10% of His's are protonated.
  - His often participates in enzymatic catalytic reactions
    - R pKa is nearly neutral, and so it can readily act as proton donors and acceptors.
- When Arg is protonated on its ε-carbon, the pos. charge is stabilized by resonance, between the N's.

- R and K side chains are protonated under physiological conditions, and often participate in electrostatic interactions

Acidic R groups:

- Asp and Glu are the only polar, acidic amino acids, containing an extra set of carboxylic acids in their R group.
- They are neg. charged above the pH of 3, and usually always carry a negative charge in physiological conditions.

3. INDUSTRIAL USE OF AMINO ACIDS

- Glycine is used in industries and factories for:
  - Animal/Pet food flavor enhancer, covering the bad tastes.
  - pH buffering in antiperspirants and deodorants.
- Glutamate is used as MSG (mono-sodium glutamate) in foods as flavor enhancers.
- Phenylalanine is needed for the production of aspartame, a main ingredient in artificial sweeteners.
- Arginine is used for dentin hypersensitivity.
- Tryptophan can be used as a sleep inducer in pharmaceutical products.
- These are just a few examples, and may appear on the midterm (fair game)

- The table on the right shows how much of an amino acid is produced on a yearly basis, as well as a common use for them.

Table 1. Estimated global production of amino acids (1996)*

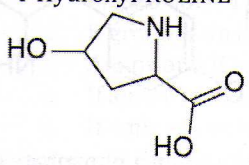
Amino acid	Amount (ton/y)	Process	Uses
L-glutamate	1,000,000	Ferm.	Flavor enhancer
D, L. Methionine	350,000	Chemical	Food , Feed Pharm.
L-Lysine HCL	250,000	Ferm.	Feed Supplement
Glycine	22,000	Chemical	Pharm., soy sauce
L-Phenylalanine	8,000	Ferm., Synthesis	Aspartame
L-Aspartic acid	7,000	Enzymatic	Aspartame, Pharm.
L-Threonine	4,000	Ferm.	Feed supplement
L-Cysteine	1,500	Extraction, Enzyma.	Pharm.
D, L –Alanine	1,500	Chemical	Flavor, sweetener
L- Glutamine	1,300	Ferm.	Pharmaceuticals
L-Arginine	1,200	Ferm.	Flavor, pharm.
L- Tryptophan	500	Ferm., Enzymatic	Feed suppl., Pharm.
L – Valine	500	Ferm.	Pharmaceuticals
L –Leucine	500	Ferm., Extraction	Pharmaceuticals
L –Alanine	500	Enzymatic	Pharm.
L –Isoleucine	400	Ferm.	Pharmaceuticals
L – Histidine	400	Ferm.	Pharmaceuticals
L – Proline	350	Ferm.	Pharmaceuticals
L – Serine	200	Ferm.	Pharmaceuticals
L - Tyrosine	120	Extraction	Pharmaceuticals

*From Ikeda, M. 2003. Adv. Biochem. Eng. Biotech. 79:1-35.

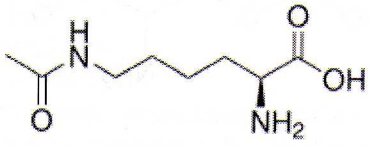
4. LESS COMMON AMINO ACIDS TO KNOW

- There are some amino acids that are found in human proteins, but are relatively rare.
- They are typically formed as a product of post-translational modification (ptm), catalyzed by specific enzymes.
  - Common ptm's include: glycosylation, oxidation, hydroxylation, methylation, acetylation, phosphorylation.
- The ptm-ed a.a. to know are:

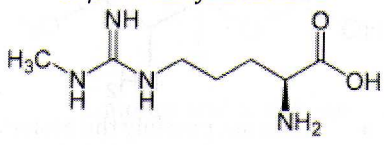
4-HydroxyPROLINE



ε-N-AcetylLYSINE

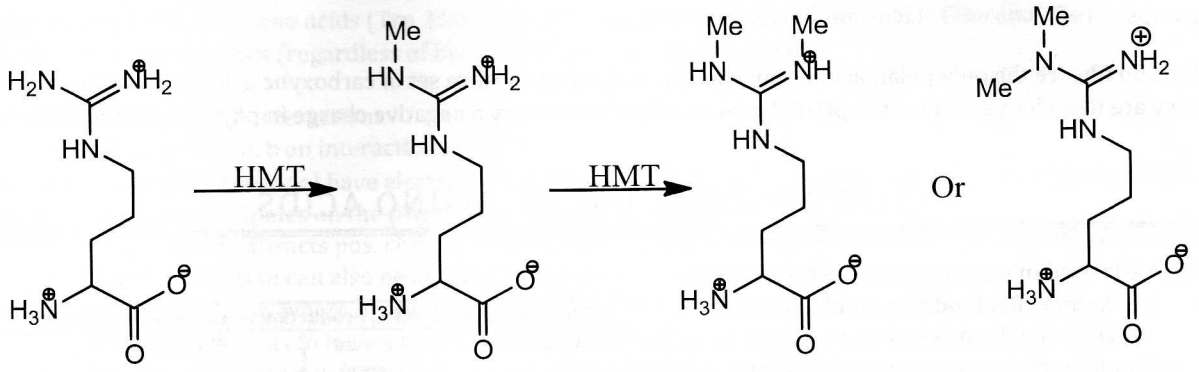


ω-β-N-MethylARGININE



- For these amino acids, only knowing the name and structures are enough.

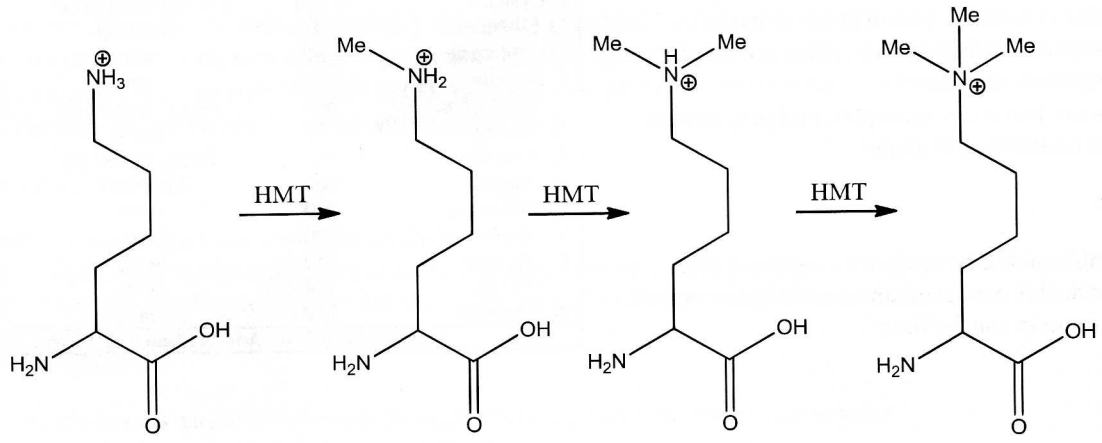
4.1 ARGININE DERIVATIVES BY PTM



- Structures above are:

Arg → Histone Methyl Transferase → Mono-methyl Arg → HMT →Symmetrical di-methyl Arg Or Asymmetrical di-methyl Arg

4.2 LYSINE DERIVATIVES BY PTM

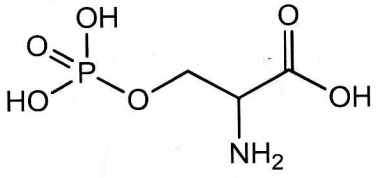


- Structures above are:

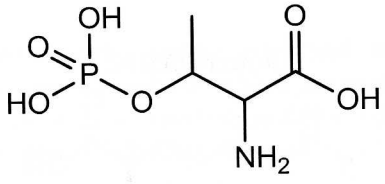
Lys → HMT → Mono-methyl Lys → HMT → Di-methyl Lys → Tri-methyl Lys

4.3 PHOSPHORYLATED AMINO ACIDS

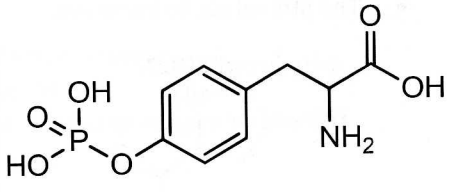
PhosphoSERINE



PhosphoTHREONINE



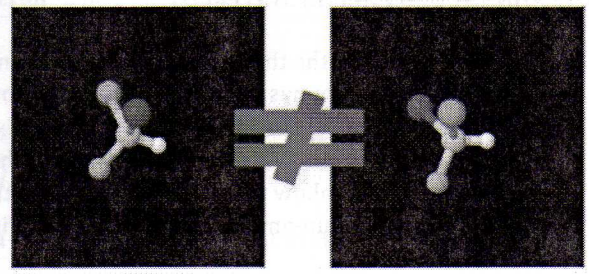
PhosphoTYROSINE



- These are possibly the easier modified a.a.'s to remember.
- The same basic structure as their common hydroxyl counterpart, except H has been replaced by a phosphate group.

5. ENANTIOMERS

- Enantiomers are a pair of molecules that are mirror images of each other.
- Enantiomers are therefore not-superposable – you cannot fit one on top of the other exactly.
- Even by rotating them, the image on the right would not match the one on the left.
    - They are optically active, where each enantiomer will refract plane polarized light by a certain angle.
- Gly is the only a.a. that is not optically active (R = H)
    - It can be superposed with its mirror image.
    - Therefore, there is no stereoisomer, there is no such thing as D-Gly or L-Gly
- All the other amino acids we need to know are optically active and have enantiomers.
- All biological proteins use L-configured amino acids only.
- Some amino acids have multiple chiral centres, such as Thr – which has 2.
    - Therefore, there are 4 possible stereoisomers of Thr that can exist.
    - The only used in the body is L-(2S, 3R) Threonine



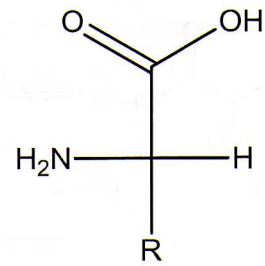
(S)-bromo(chloro)fluoromethane (R)-bromo(chloro)fluoromethane

5.1 POLARIMETRY

- Back in the days of old, before the molecular structures of compounds were known, the only way to differentiate between enantiomers was to use plane polarized light.
- A solution of one pure form of an enantiomer would refract and bend light to a certain degree when passed through a polarimeter tube.
- A solution of the other form would refract light in the opposite direction to the same degree.
- So scientists at the time assigned "d" to solutions that angled light to the right, and "l" to compounds that angled left.
- Note: there is NO CORRELATION b/w the Polarimetric (d,l) system, Fisher (D, L) system, and Prelog (R,S) system
    - E.g. L-Glutamic acid actually refracts light to the right, therefore, it is polarimetrically 'd'

5.2 FISHER-ROSANOFF PROJECTION

- The Fischer projections are characterized by one or more chiral centres surrounded by 4 bonds in a + pattern.
- The adjacent chiral centres are aligned vertically.
    - It was originally only intended for carbohydrates, but could be extrapolated to amino acids.
        - I.e. It can be ambiguous though (There is no way to differentiate the 2 L threonines using Fisher.
- Rules for Fisher Projection:
    - Chiral carbon is the centre/intersection of the 2 lines.
    - If there are multiple chiral carbons, they are aligned vertically.
    - Vertical lines always point into the plane of the page
    - Horizontal lines always point out of the plane of the page.
    - Most oxidized group (by convention) goes on top (so COOH)
    - R group is on the bottom
    - Amine and H are on the horizontal lines
    - If amine is on the left, then its L-amino acid.
    - If amine is on the right, then its D-amino acid.



L-amino acid backbone

- Geometry of Fisher projections:
    - 180º rotation of a Fisher projection does not change the enantiomer, keeps the structure constant.
    - 90º rotation of a Fisher projection leads to the mirrored enantiomer.
- The Fisher projection is ambiguous and its nomenclature fails when dealing with multiple chiral centres
    - It has therefore been retired from use in the scientific community.

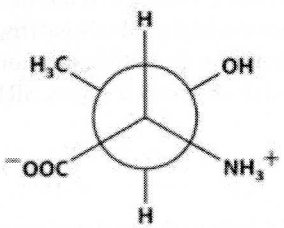
- The only reason it's still taught is because its "easier" for students to learn – R/S is the modern convention

5.3 CAHN-INGOLD PRELOG SYSTEM

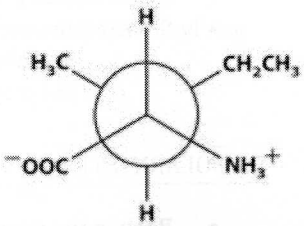
- This is based on the theory that a chiral carbon has 4 distinct groups.
  - Each group is prioritized based on atomic weight of the atom just immediately adjacent to the chiral carbon.
- A numerical rank is assigned based on the weight, with the lightest atom (oft H) oriented to the back of the molecule.
- If the numbers follow a clockwise turn, then the configuration of that chiral carbon is R
- If the numbers follow an anti-clockwise turn, then the configuration of the chiral carbon is S.
- This is far more un-ambiguous than the D/L Fisher system.

- In biological system, almost all L-amino acids are S configured in the Prelog system.
- Exception: L-Cysteine – it is the only R configured L-amino acid used biologically.
- This is because in most cases, the priority rank follows: NH > COO > R > H for amino acids.
  - For L-Cys, where R = $(-CH_2 - SH)$, it changes the rank to: NH > R > COOH > H.

- The Fisher projections to the right show the Newman Projections, and the R/S configuration of the amino acids with more than 1 Chiral centre.
- We need to memorize this for the midterm for sure.
- The 2 refers to the α-Carbon, and 3 to the β-Carbon.



(2*S*,3*R*)-Threonine

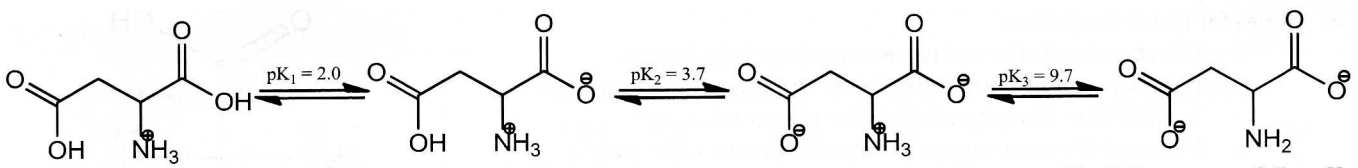


(2*S*,3*S*)-Isoleucine

6. IONIZATION OF AMINO ACIDS

- The ionization of α-amino and α-carboxylic groups of amino acids will allow for 3 distinct states.
- At low pH, the acidic environment provides excess H^+ molecules to the a.a.'s leading to protonation.
  - This gives the a.a. a net cationic charge.
- At neutral pH, the zwitterionic state predominates.
  - There is a net neutrality experienced by the solution.
- At high pH values, the basic environment steals the H^+ from the a.a.'s, leading to deprotonation.
  - This yields the anionic form of the a.a.

- One way to think about this is to think each pKa is a threshold, at which the molecule loses 1 proton (H^+)



pH < 2 | pH = 2 | 2 < pH < 3.7 (pI region) | pH = 3.7 | 3.7 < pH < 9.7 | pH = 9.7 | 9.7 < pH

- Example of a single amino acid: Aspartic Acid
- So starting at pH 0, the amino acid has an overall pos. charge.
- When surrounding pH matches a pKa value (pH = 2), we can think of it as a toll, where one H has to be paid.
  - The first pKa corresponds to alpha carboxylic – it gets deprotonated.
  - In the case of Asp acid, this is the place where the isoelectric point occurs.
    - Theoretically all of the amino acid carries both a pos. and neg. charge in between the 2 pKa values.
- As the pH keeps increasing, the pH eventually matches the next pKa value of 3.7.

- This corresponds to the R carboxylic group, which gets deprotonated at this stage.
- It now exists as an ampholyte, carrying a net neg. charge.

- When pH = 9.7, the alpha amino gets deprotonated, giving an overall -2 neg. charge.

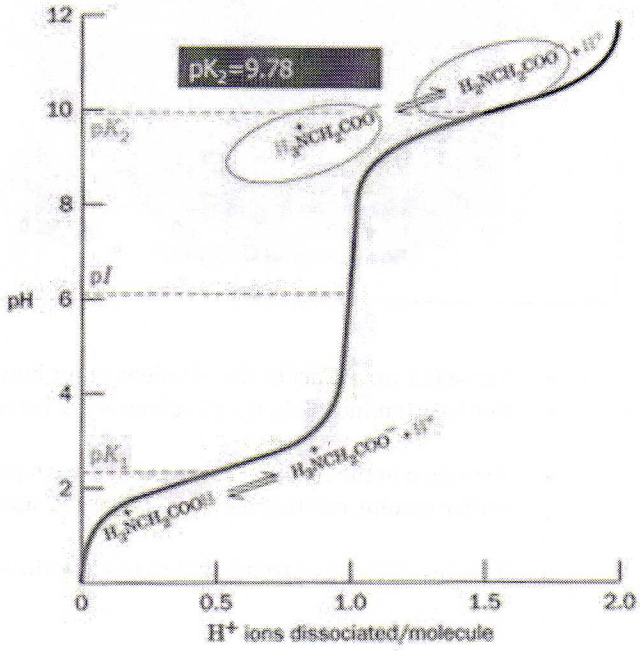
6.1 HENDERSON-HASSELBACH EQUATION

$$pH = pK_a + \log\frac{[A^-]}{[HA]}$$

- Using the H-H eqn, the pH will be equal to pK when both the acid and conj. base is present in equal proportions.
  - When $[A^-] = [HA]$, $\log[1] = 0$, and therefore, pH = pK
- So, from this, we can also infer that when pH is exactly equal to a pK value,
  - Both molecules across the arrow will exist in equal proportions.
  - E.g. In the Asp example above, when aqueous pH = 2, both the pos. charged specie, and zwitterion specie will exist in equal amounts (50% each)

6.2 TITRATION OF GLYCINE

- At low pH, both amino and the carboxyl acid-base groups are fully protonated, with the amino acid as the cationic form.
- Adding a base changes equilibrium towards the zwitterion.
- When pH = pK_1, the 2 charged species exist in equilibrium in equal amounts.
  - Also note: pH values do not change drastically when pH = pK
  - This allows for a.a.'s to have a buffering capacity.
  - Generally, amino acids are good buffers at a range of:
    - $pH = pK \pm 1$
    - $pH = 2.3 \pm 1$



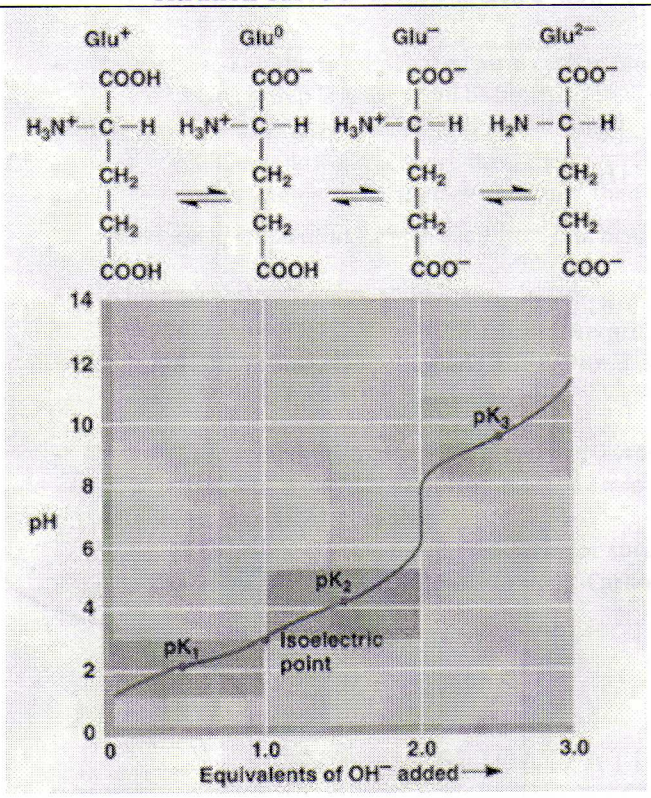
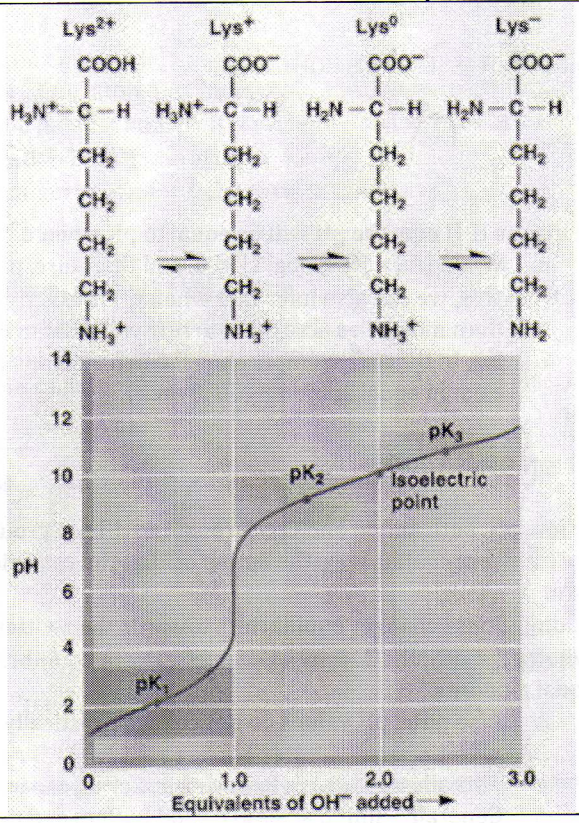
- With a base equivalent of 1.0 (1 eqv. NaOH added), pH increases quickly, at which carboxylic acid is fully ionized, and cannot provide protons to neutralize OH^-'s, drastically increasing pH.
  - At this state, the isoelectric point also occurs.
  - At the isoelectric point, theoretically, every single molecule of the a.a. exists in its zwitterionic form.

- As the pH nears the second pK of Gly, buffering capacity is achieved again
- When pH = pK_2,both charged specie exists in solution: 50% zwitterion and 50% anion.
- Any more addition of base, and Gly assumes a pure anionic form.

- The trend described above applies to all amino acids that do not have an ionizable R group.

- In the case for Gly, the isoelectric point was directly in between the 2 amino acids.
  - The exact pI is calculated as the midpoint between the 2 surrounding pK values.
  - So for Gly, the pI value would be 6.04
- For charge-able amino acids, their pI values also occur where a zwitterion exists.
- For acidic amino acids, the pI values occur between the 2 most acidic pKa's
- For basic amino acids, the pI values occur between the 2 most basic pKa's.

6.3 TITRATION OF CHARGED AMINO ACIDS.

Titration Curve of Glutamic Acid	Titration Curve of Lysine
	

- For acidic amino acids, the pI values occur between the 2 most acidic pKa's
- For basic amino acids, the pI values occur between the 2 most basic pKa's.

- Also seen in the titration curves above, each pK values provides a "buffer zone," where that amino acid can act as a buffer system, existing between 2 charged states.

- This titration curve trend applies to all acidic and basic amino acids.