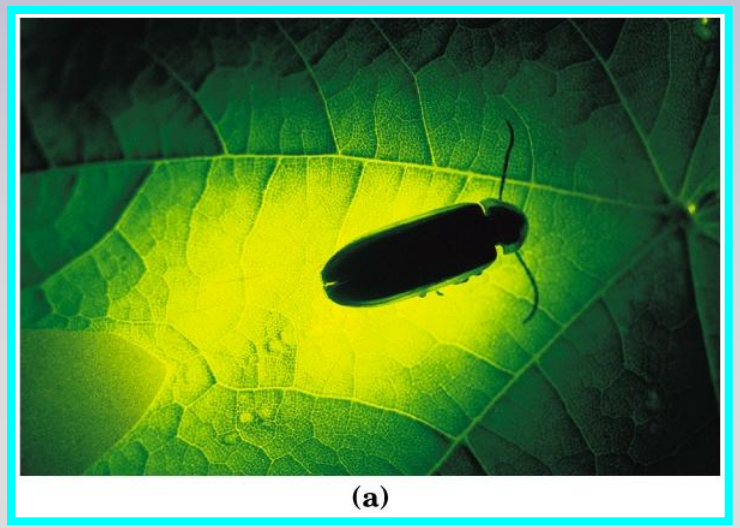
Amino Acids, Peptides , And Proteins

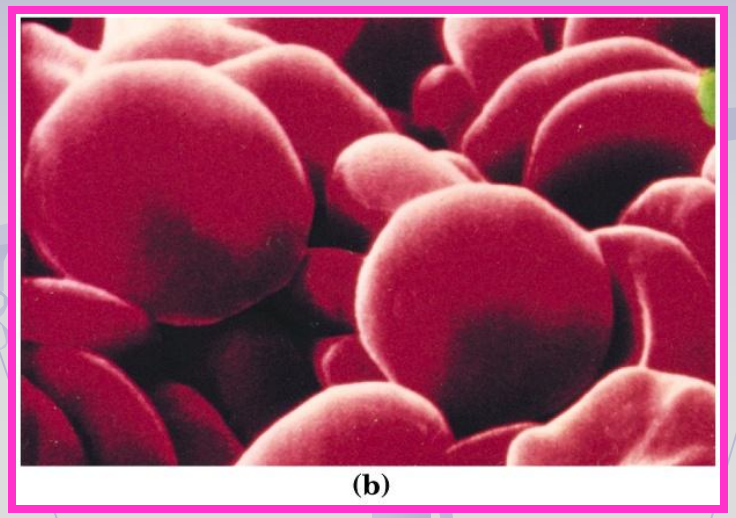
Amino Acids, Peptides , And Proteins

Dr.Sulieman Al-Khalil

Great Variety of Protein Functions



(a)

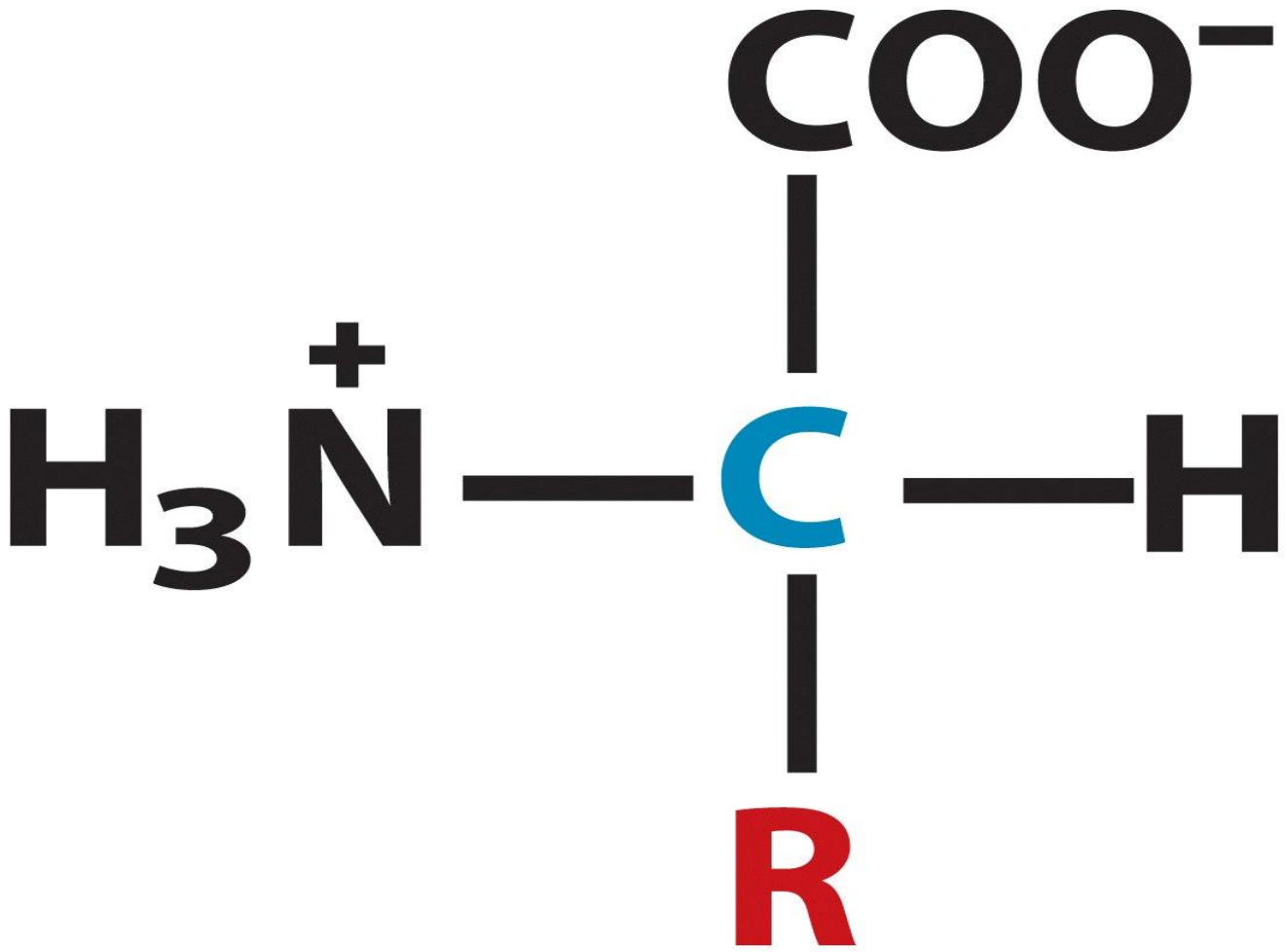


(b)

Amino acids

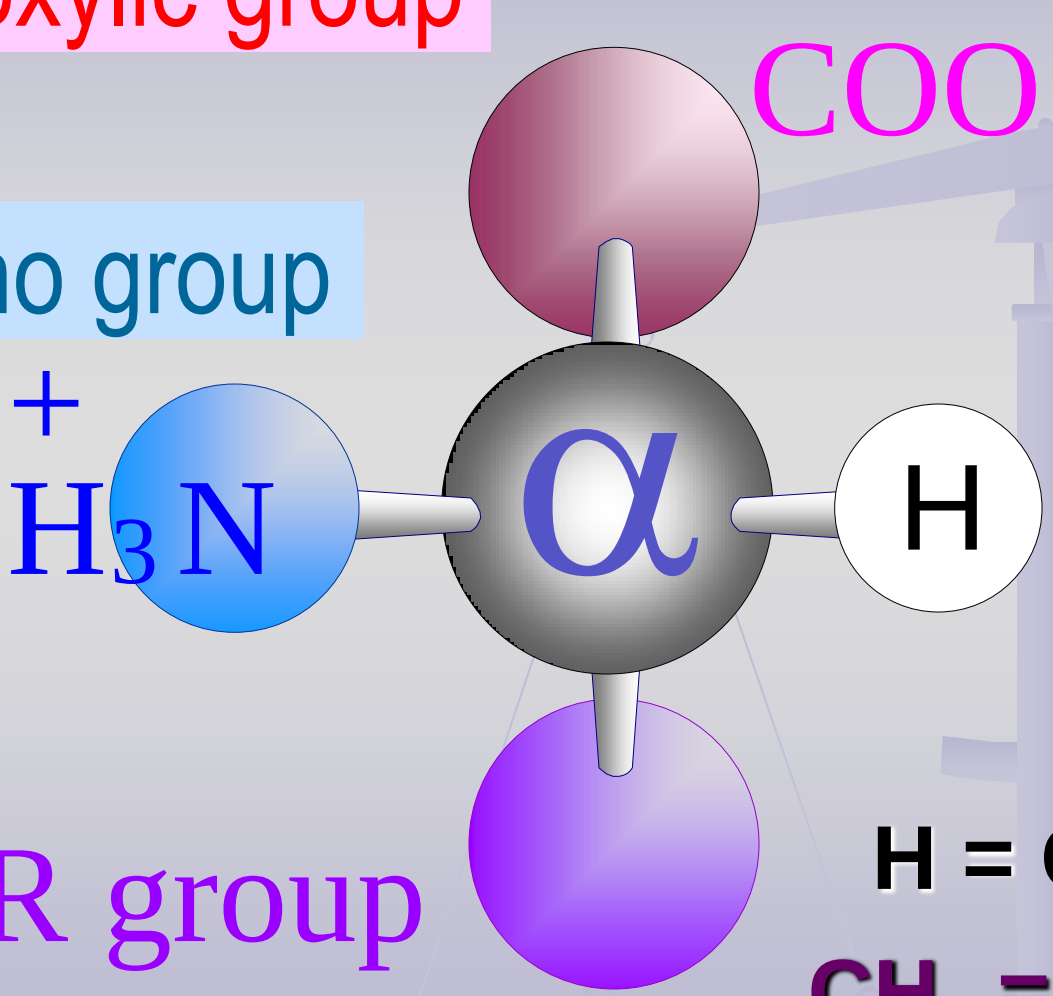
- **Proteins** : polymers consist of amino acids (a.a's) as functional structural units and easily demonstrated by hydrolysis , i.e. Chemically or enzymetically .
- **Amino acids** : group of organic Cps contain 2 functional groups → Amino group & Carboxylic group , and both are attached to the same α Carbon in ionized form .
- A .A's in nature : 300 a.a , 20 of them are found in proteins and others are not isolated yet .

General Structure of an Amino Acid



L-Form Amino Acid Structure

Carboxylic group



H = Glycine

CH_3 = Alanine

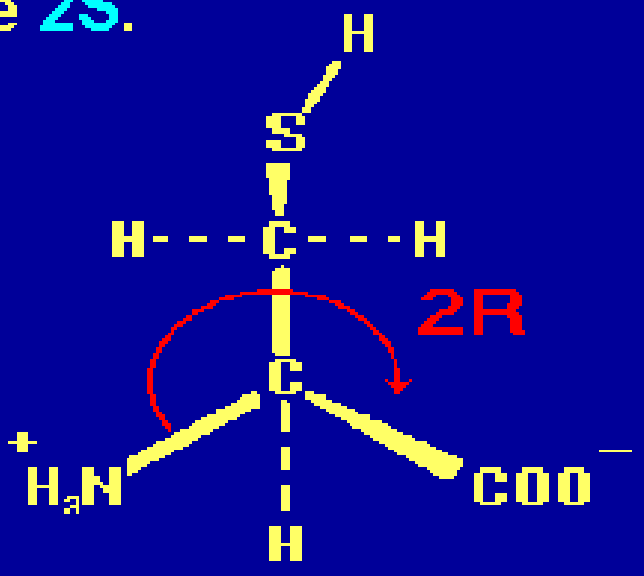
Juang RH (2004) BCbasics

- Under complete hydrolysis of proteins , 20 a.a's are found & have the previous common structure.
- When R= H → this a.a is Glycine (Gly ; Greek glykos , "sweet") .
- Gly is optically inactive [Achiral Cp] ; α-Carbon lose its chirality → It has symmetry .
- All other a.a's have chiral Carbon.
- At pH=7 , some a.a's are found in protein at Levo or Dextro level.
- D- Glyceraldehyde is the reference Cp for all a.a stereochemistry .
- A.A's have L- configuration in protein against D- Glyceraldehyde configuration .

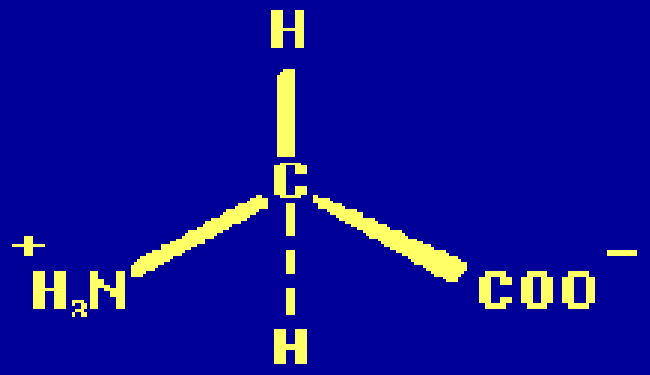
Note about,

D, L Configuration Notation versus R, S

The common naturally occurring amino acids are all L configuration and all but cysteine and glycine are 2S.

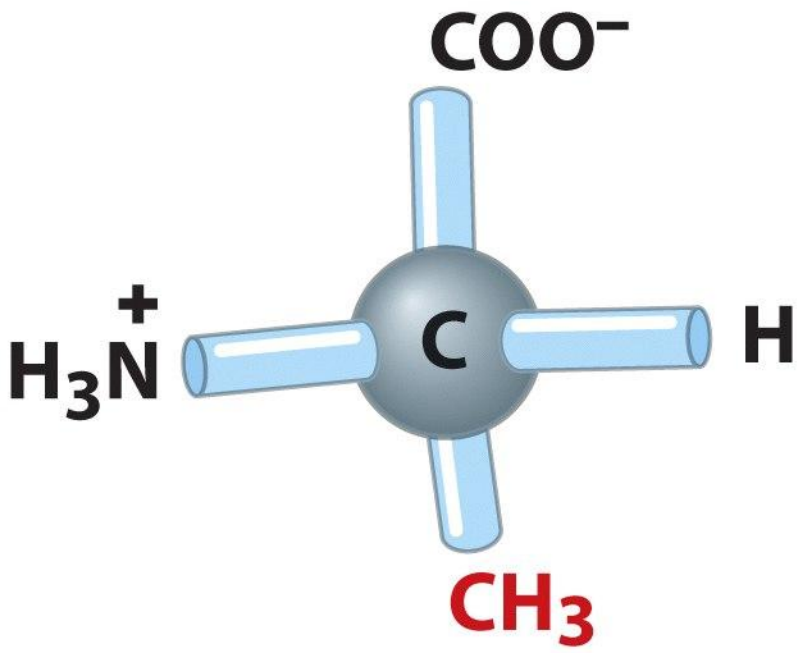


L-Cysteine

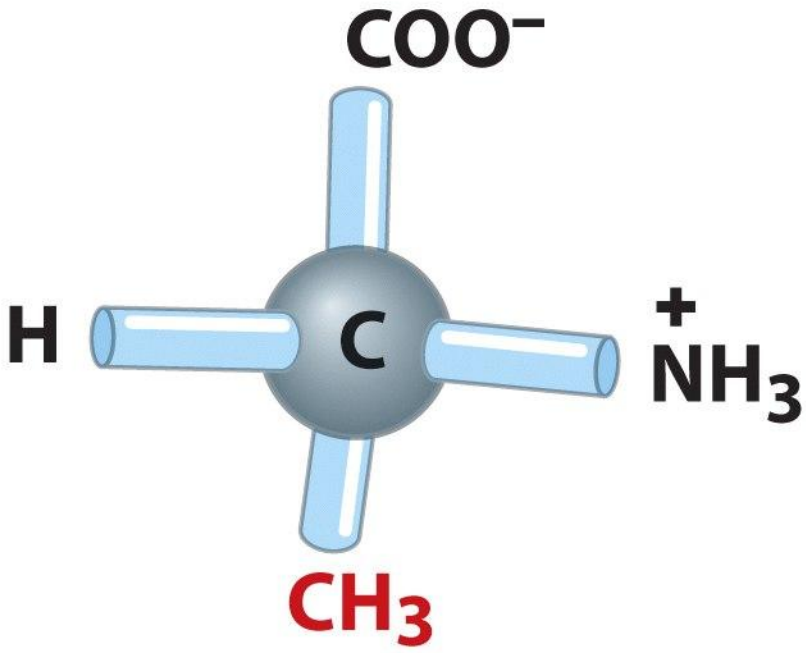


Glycine

All Amino Acids (Except G) Have Two Stereoisomers

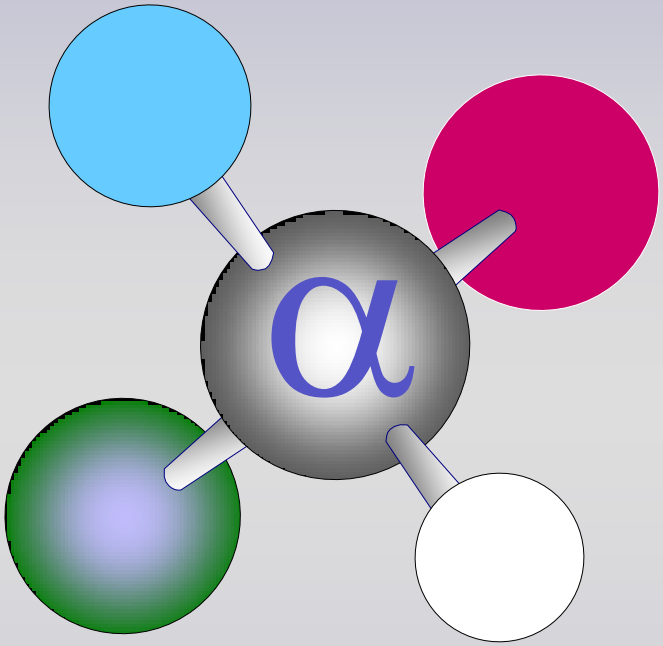


L-Alanine

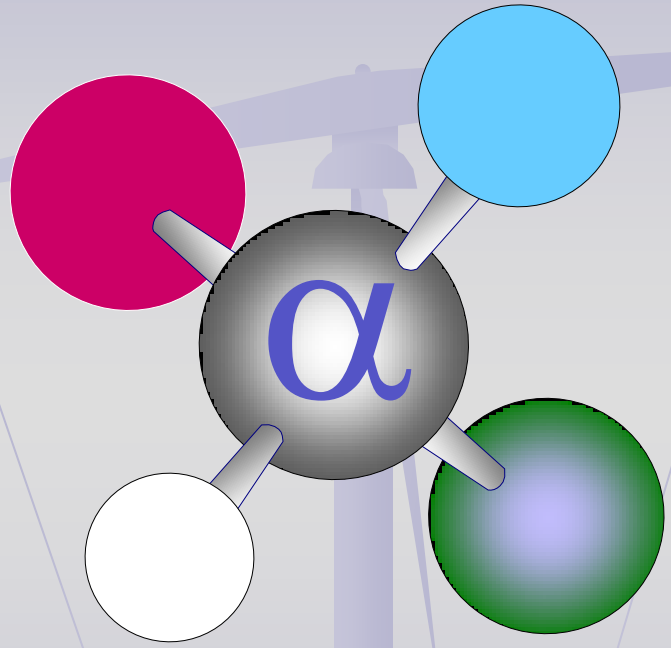


D-Alanine

Mirror Images of Amino Acid



Mirror image

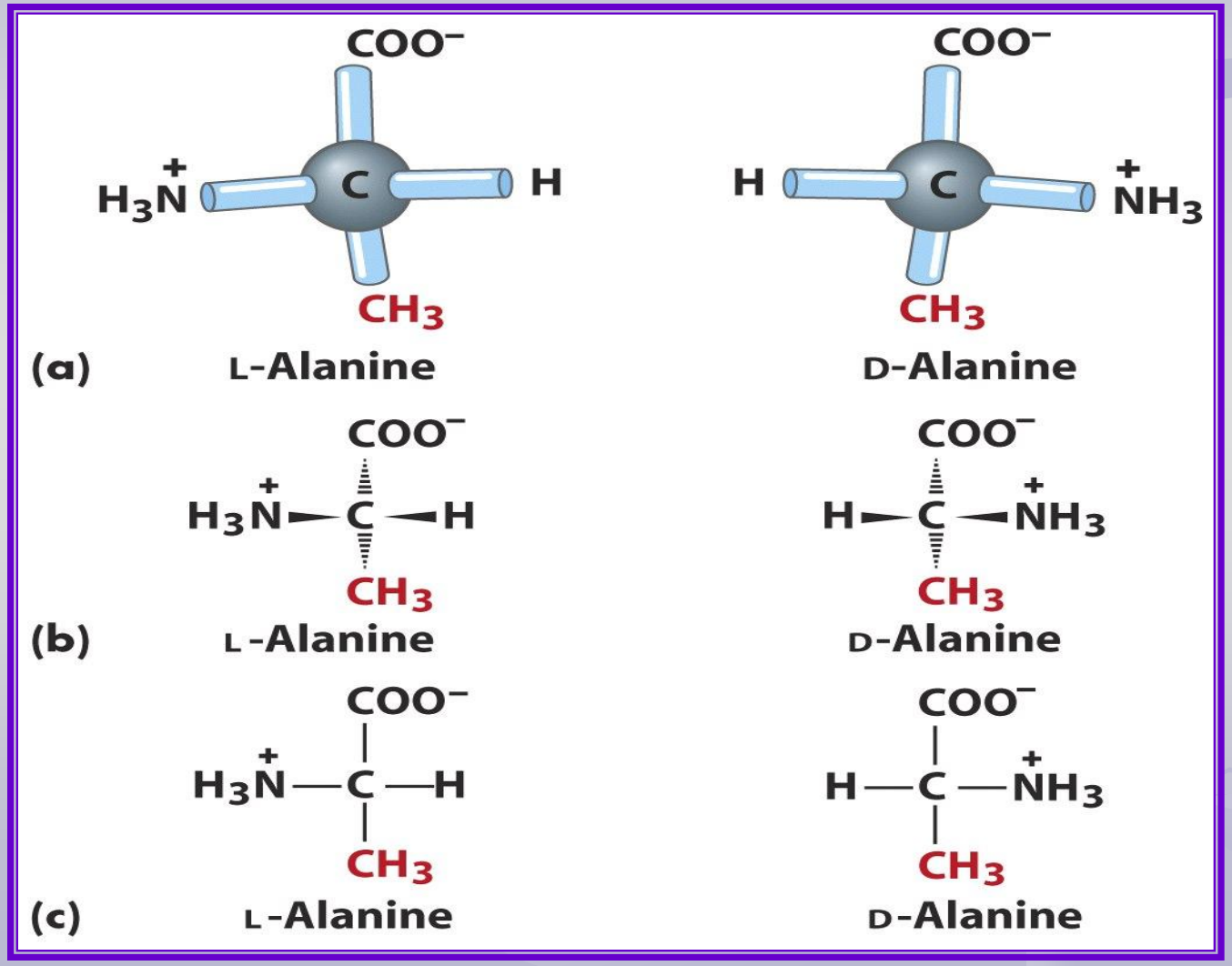


Same chemical properties

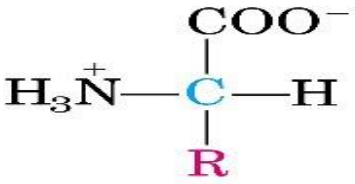
Stereo isomers

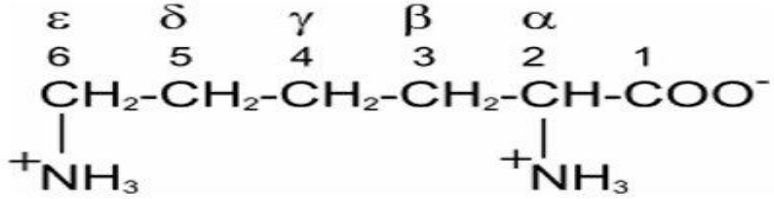
Juang RH (2004) BCbasics

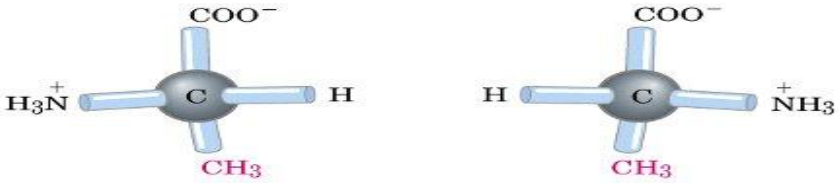
All Amino Acids (Except G) Have Two Stereoisomers



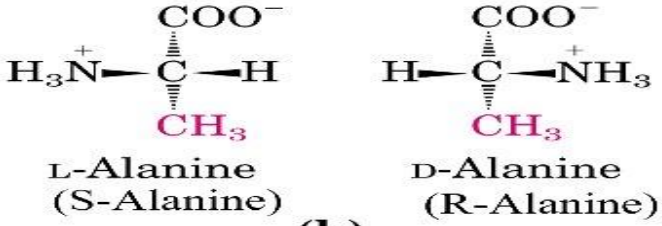
Amino Acids



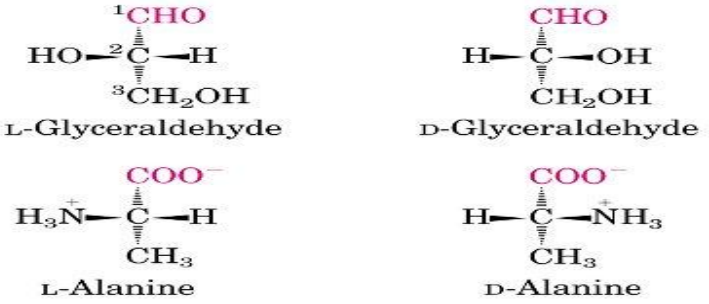


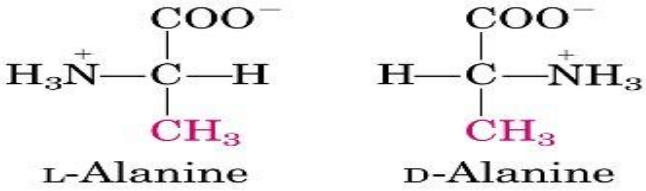


(a)



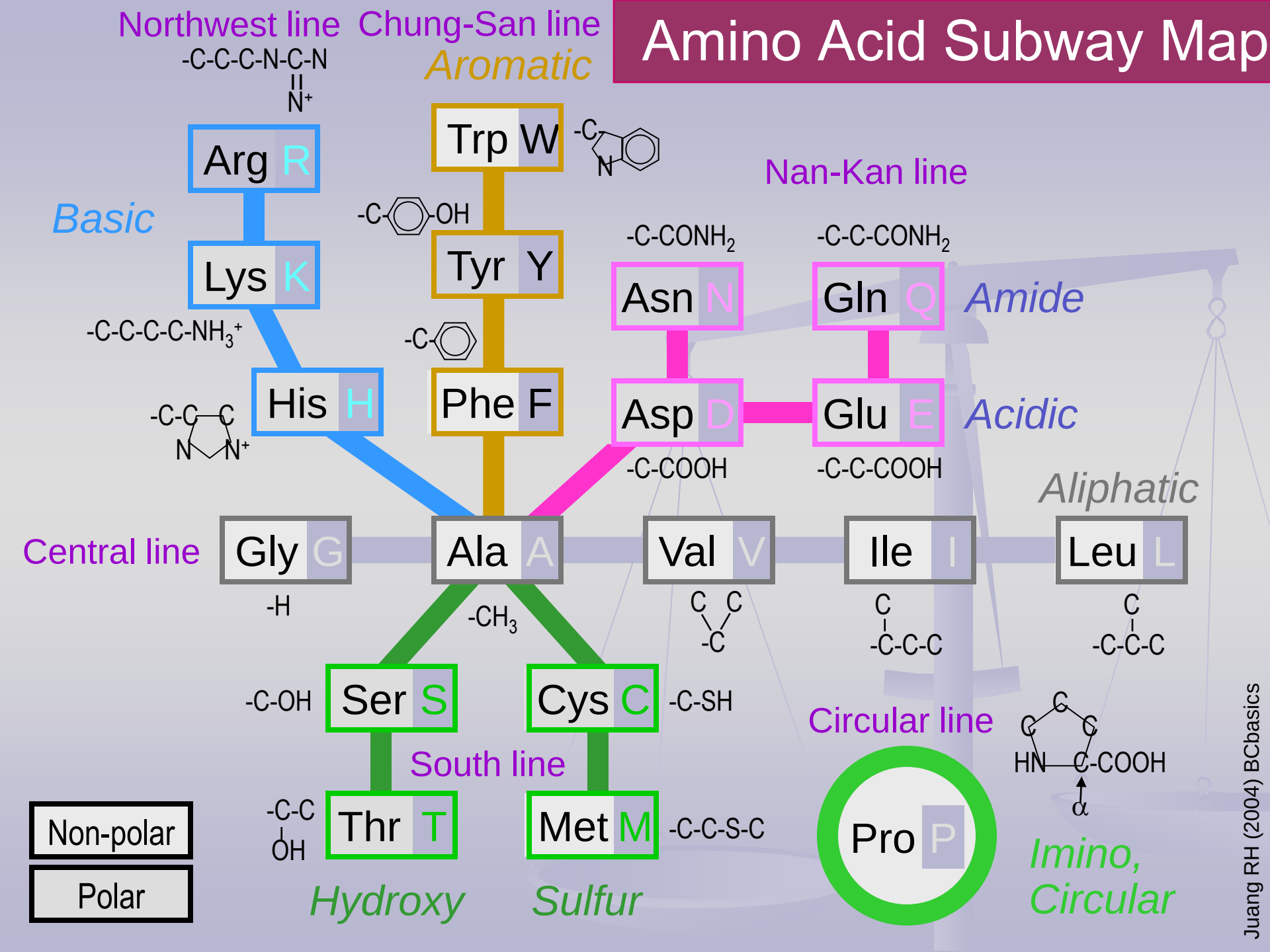
(b)



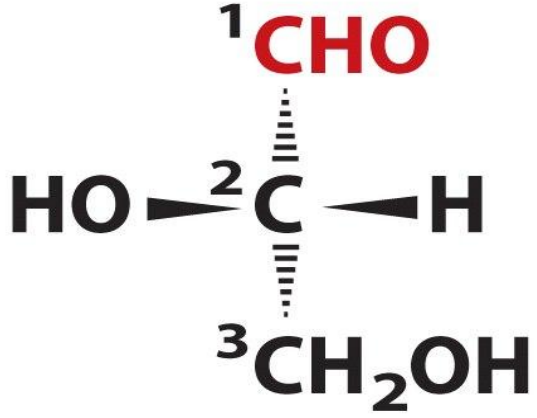


(c)

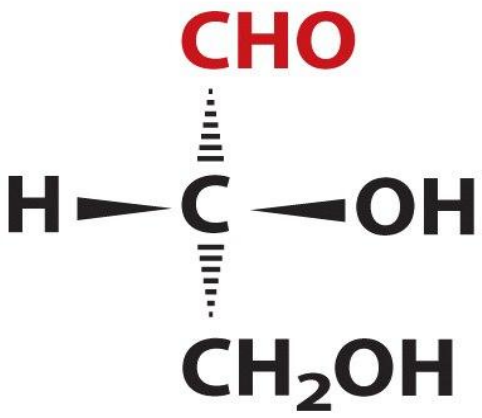
Amino Acid Subway Map



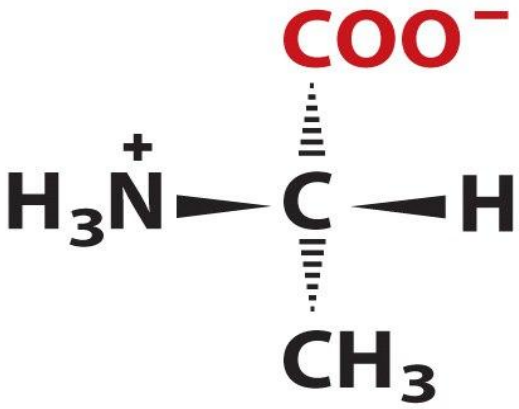
Convention Based on Glyceraldehyde



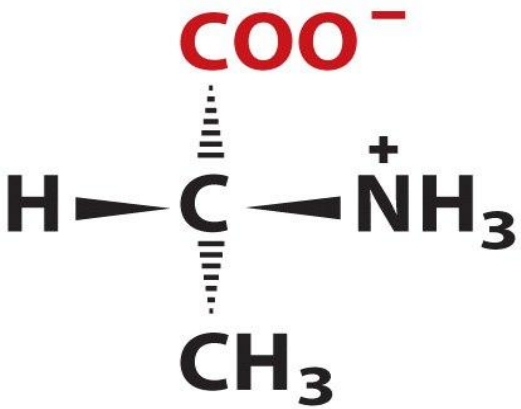
L-Glyceraldehyde



D-Glyceraldehyde



L-Alanine



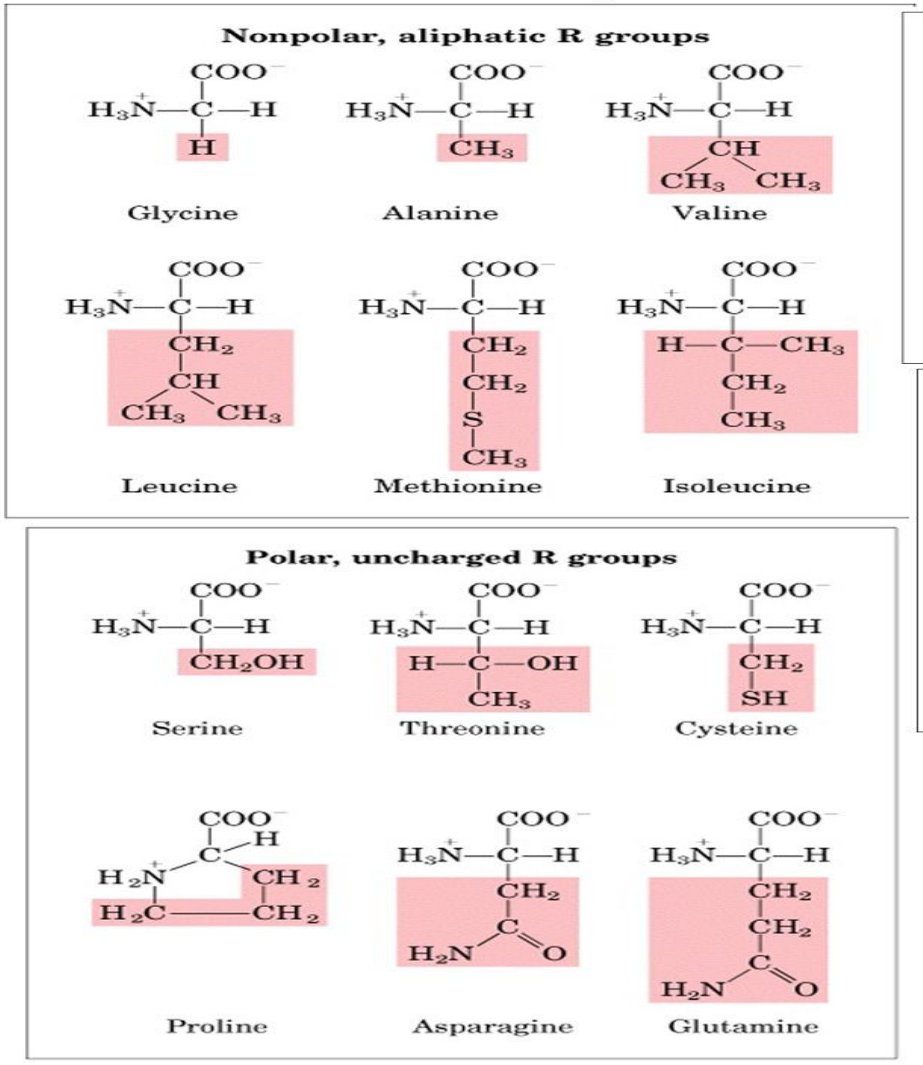
D-Alanine

Classification of a.a

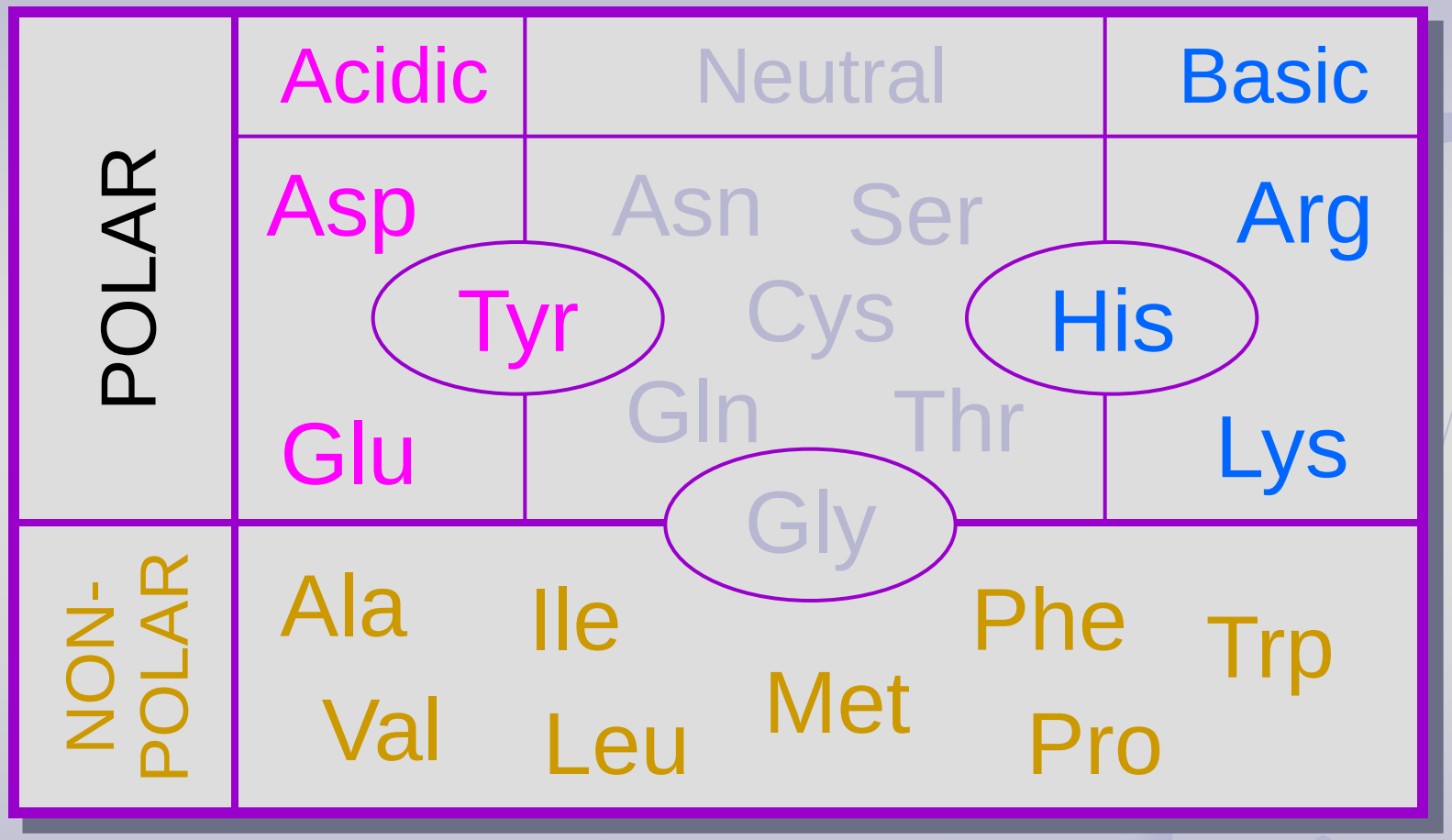
- They are classified according to R group :-

1. Non polar or Hydrophobic a.a:-
   - Aliphatic.
   - Aromatic.
2. A.A with polar , uncharged R group :-
   - OH containing a.a .
   - SH- containing a.a .
   - Amide group.
3. A .A with positively charged R group.
4. A .A with Negatively charged R group.

Twenty standard Amino Acids



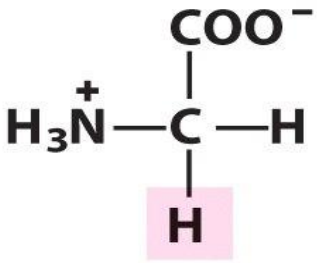
Classification of Amino Acids by Polarity



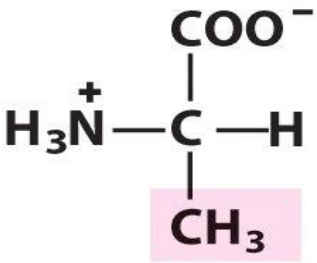
Polar or non-polar, it is the bases of the amino acid properties.

Juang RH (2003) Biochemistry

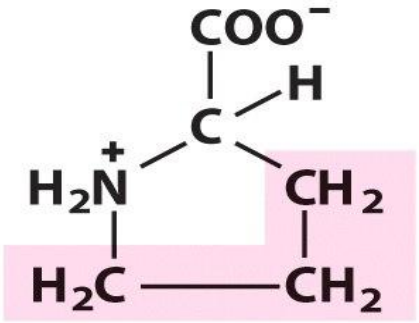
Nonpolar, aliphatic R groups



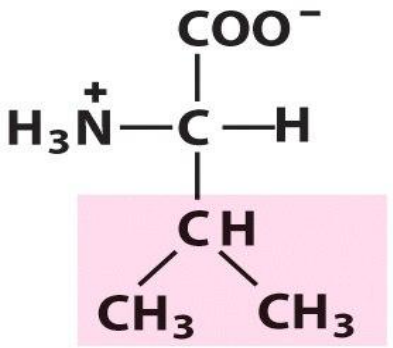
Gly
G
Glycine



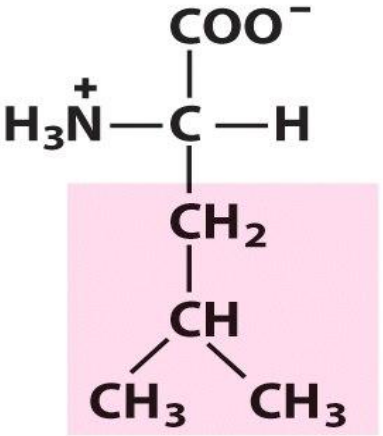
Ala
A
Alanine



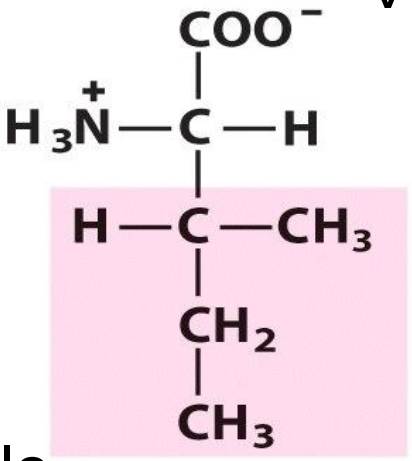
Pro
P
Proline



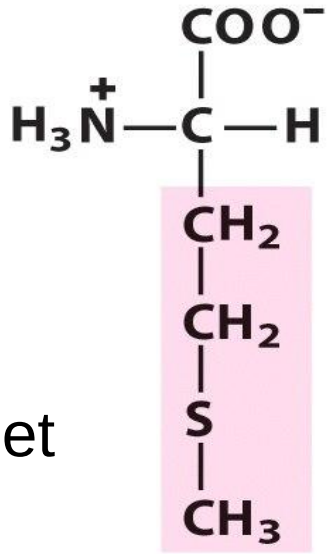
Val
V
Valine



Leu
L
Leucine

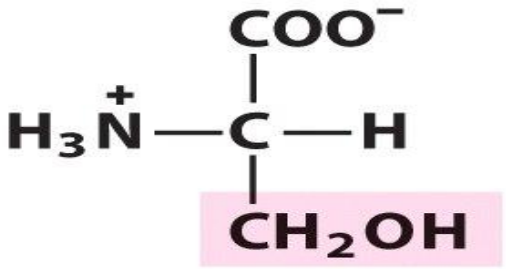


Ile
I
Isoleucine

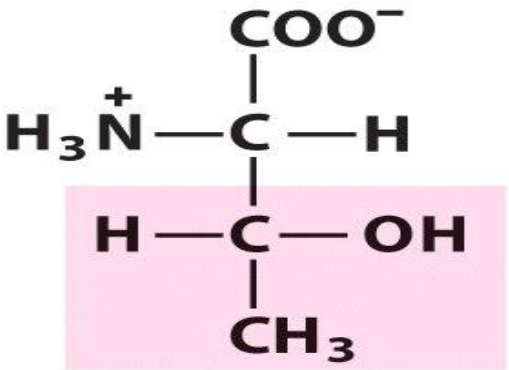


Met
M
Methionine

Polar, uncharged R groups

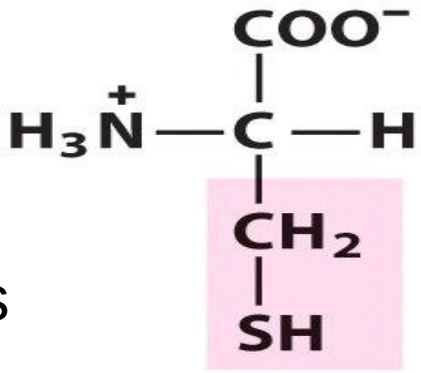


Ser
S **Serine**



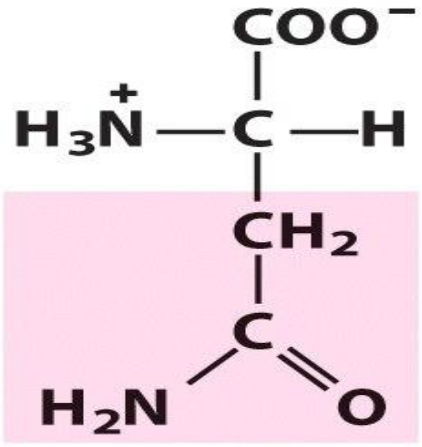
Thr
T **Threonine**

Cys
C



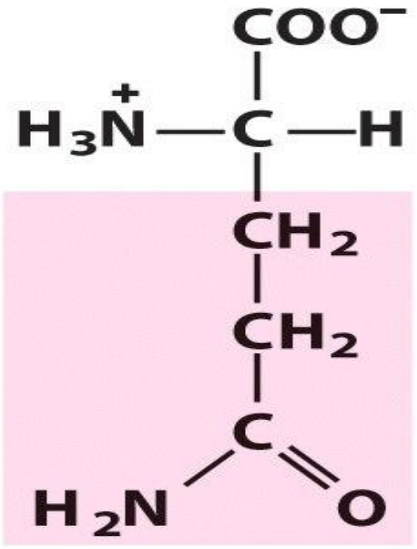
Cysteine

Asn
N



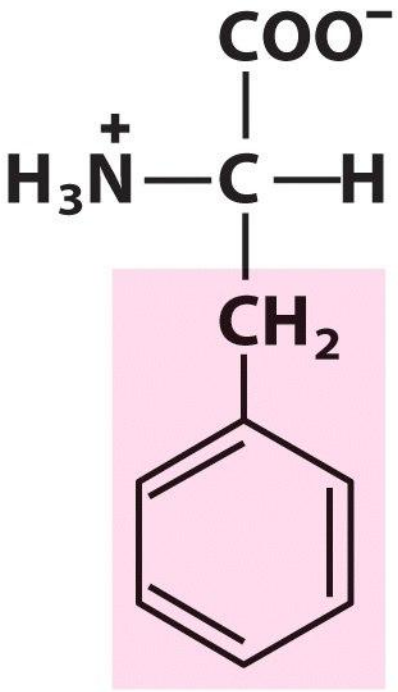
Asparagine

Gln
Q

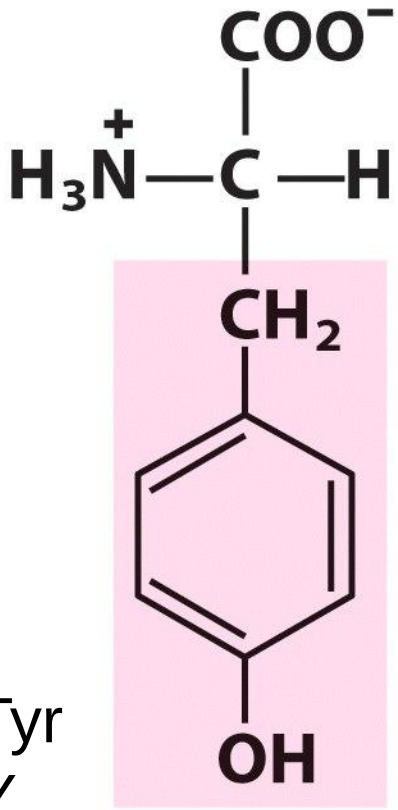


Glutamine

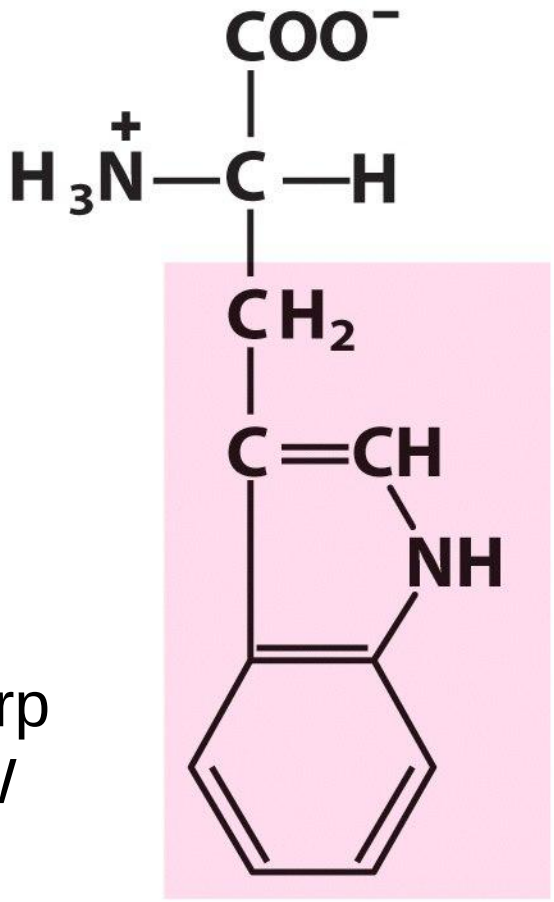
Aromatic R groups



Phe
F



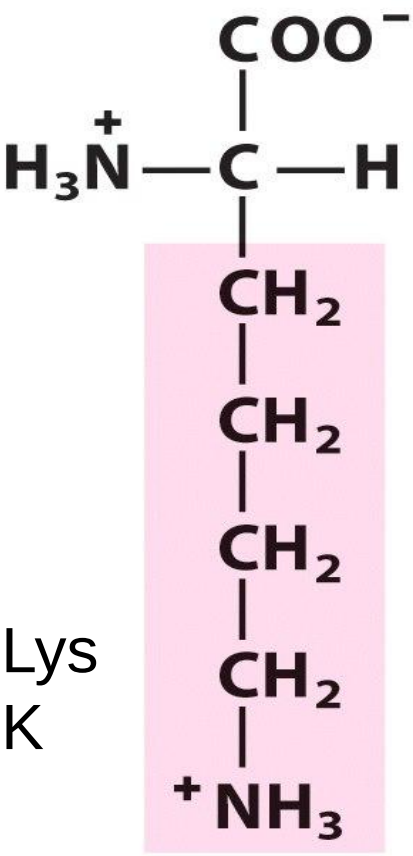
Tyr
Y



Trp
W

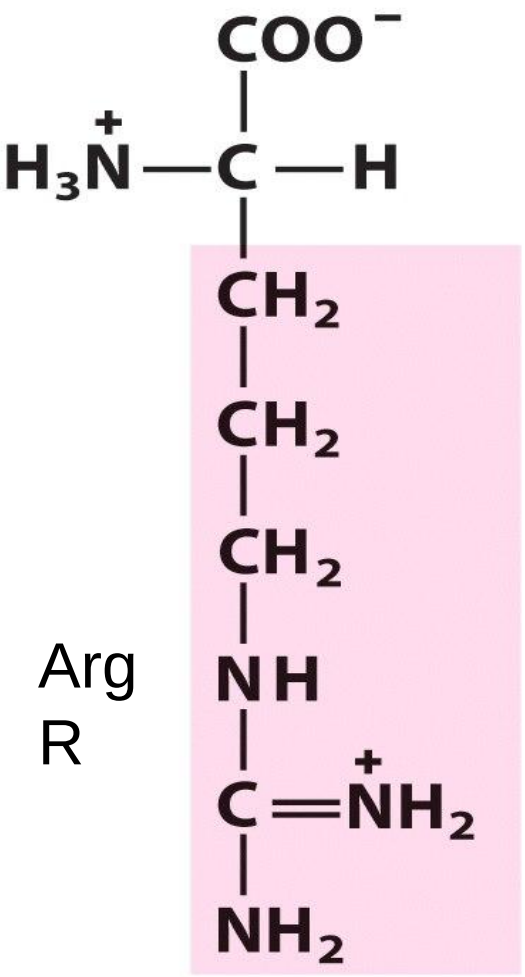
Phenylalanine **Tyrosine** **Tryptophan**

Positively charged R groups



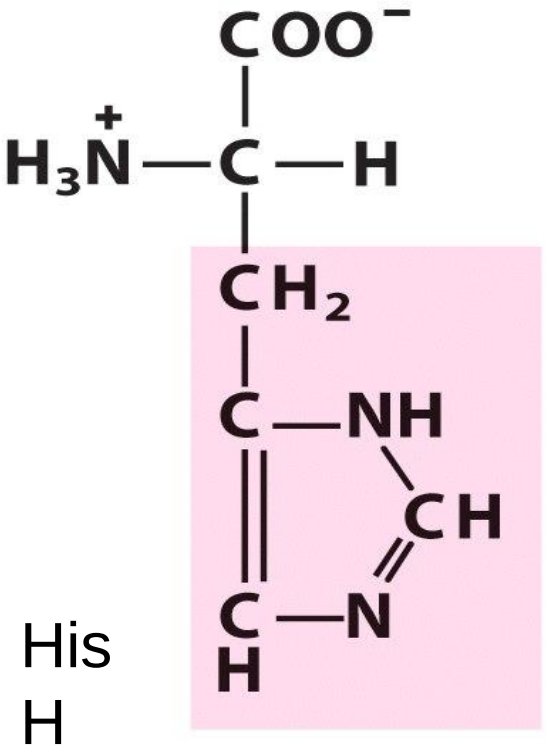
Lys
K

Lysine



Arg
R

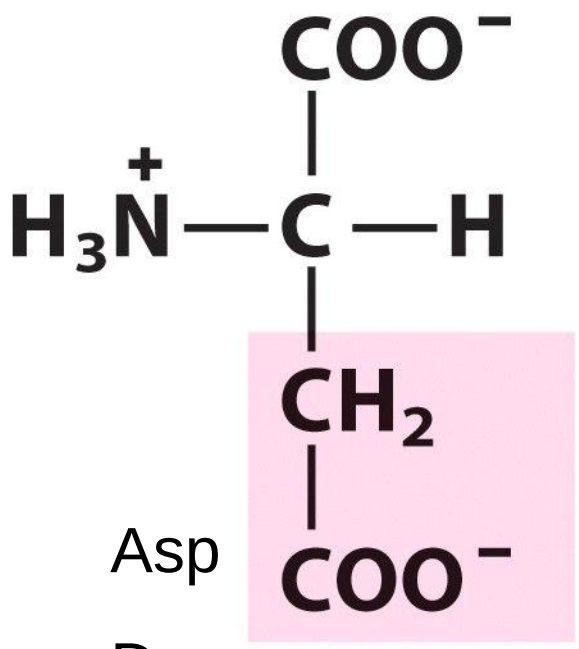
Arginine



His
H

Histidine

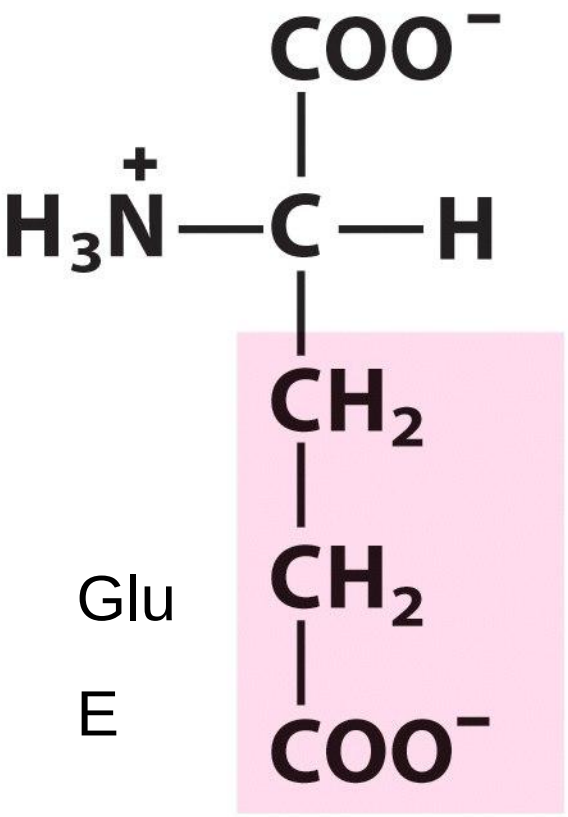
Negatively charged R groups



Asp

D

Aspartate



Glu

E

Glutamate

TABLE 3–1 Properties and Conventions Associated with the Common Amino Acids Found in Proteins

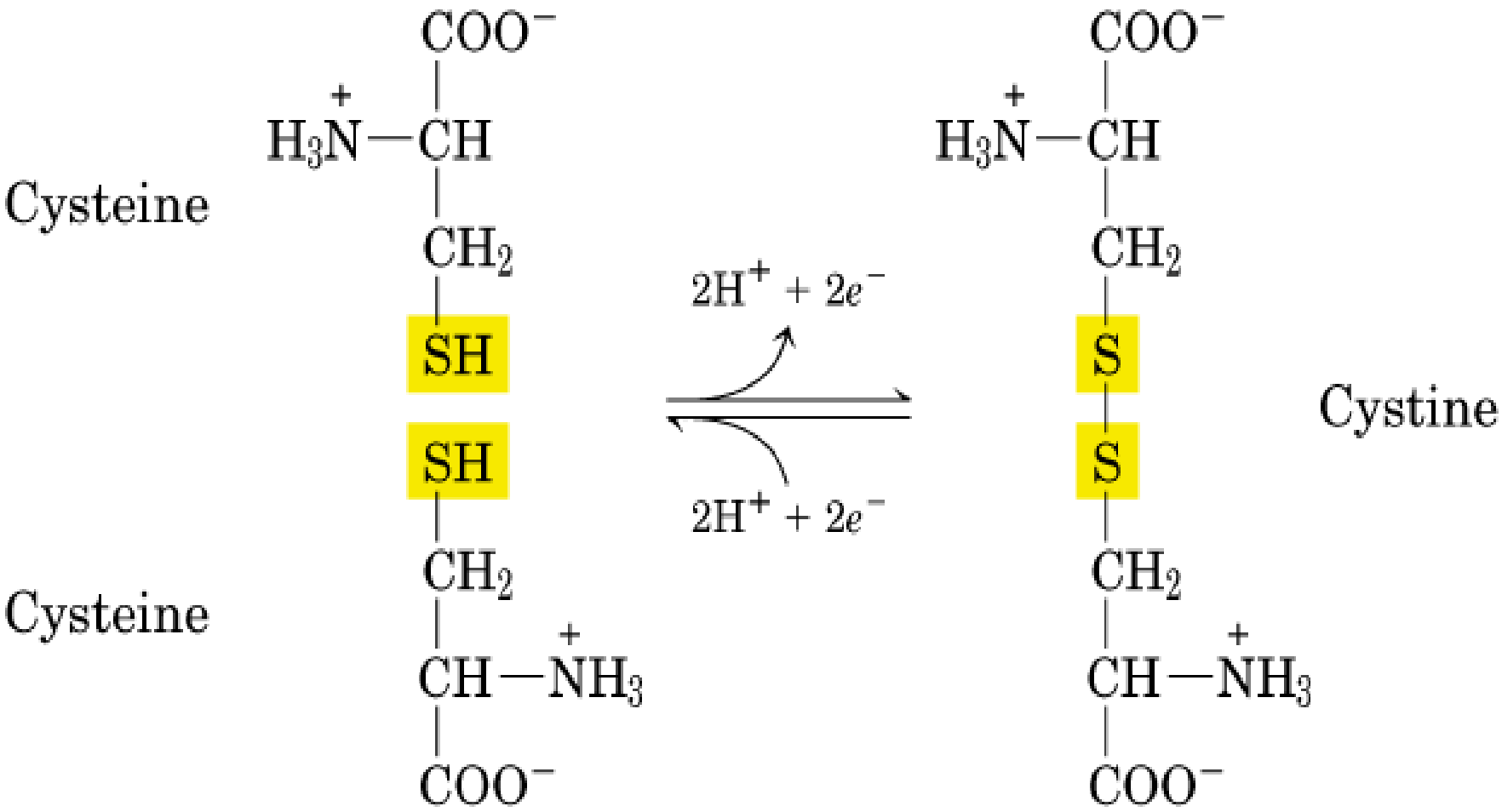
Amino acid	Abbreviation/ symbol	M_r	pK_a values: pK_1 (—COOH)	pK_2 ($—NH_3^+$)	pK_R (R group)	pI	Hydropathy index*	Occurrence in proteins (%)†
Nonpolar, aliphatic R groups								
Glycine	Gly G	75	2.34	9.60		5.97	−0.4	7.2
Alanine	Ala A	89	2.34	9.69		6.01	1.8	7.8
Proline	Pro P	115	1.99	10.96		6.48	1.6	5.2
Valine	Val V	117	2.32	9.62		5.97	4.2	6.6
Leucine	Leu L	131	2.36	9.60		5.98	3.8	9.1
Isoleucine	Ile I	131	2.36	9.68		6.02	4.5	5.3
Methionine	Met M	149	2.28	9.21		5.74	1.9	2.3
Aromatic R groups								
Phenylalanine	Phe F	165	1.83	9.13		5.48	2.8	3.9
Tyrosine	Tyr Y	181	2.20	9.11	10.07	5.66	−1.3	3.2
Tryptophan	Trp W	204	2.38	9.39		5.89	−0.9	1.4
Polar, uncharged R groups								
Serine	Ser S	105	2.21	9.15		5.68	−0.8	6.8
Threonine	Thr T	119	2.11	9.62		5.87	−0.7	5.9
Cysteine	Cys C	121	1.96	10.28	8.18	5.07	2.5	1.9
Asparagine	Asn N	132	2.02	8.80		5.41	−3.5	4.3
Glutamine	Gln Q	146	2.17	9.13		5.65	−3.5	4.2
Positively charged R groups								
Lysine	Lys K	146	2.18	8.95	10.53	9.74	−3.9	5.9
Histidine	His H	155	1.82	9.17	6.00	7.59	−3.2	2.3
Arginine	Arg R	174	2.17	9.04	12.48	10.76	−4.5	5.1
Negatively charged R groups								
Aspartate	Asp D	133	1.88	9.60	3.65	2.77	−3.5	5.3
Glutamate	Glu E	147	2.19	9.67	4.25	3.22	−3.5	6.3

*A scale combining hydrophobicity and hydrophilicity of R groups; it can be used to measure the tendency of an amino acid to seek an aqueous environment (− values) or a hydrophobic environment (− values). See Chapter 11. From Kyte, J. & Doolittle, R.F. (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157,** 105–132.

†Average occurrence in more than 1,150 proteins. From Doolittle, R.F. (1989) Redundancies in protein sequences. In *Prediction of Protein Structure and the Principles of Protein Conformation* (Fasman, G.D., ed.), pp. 599–623, Plenum Press, New York.

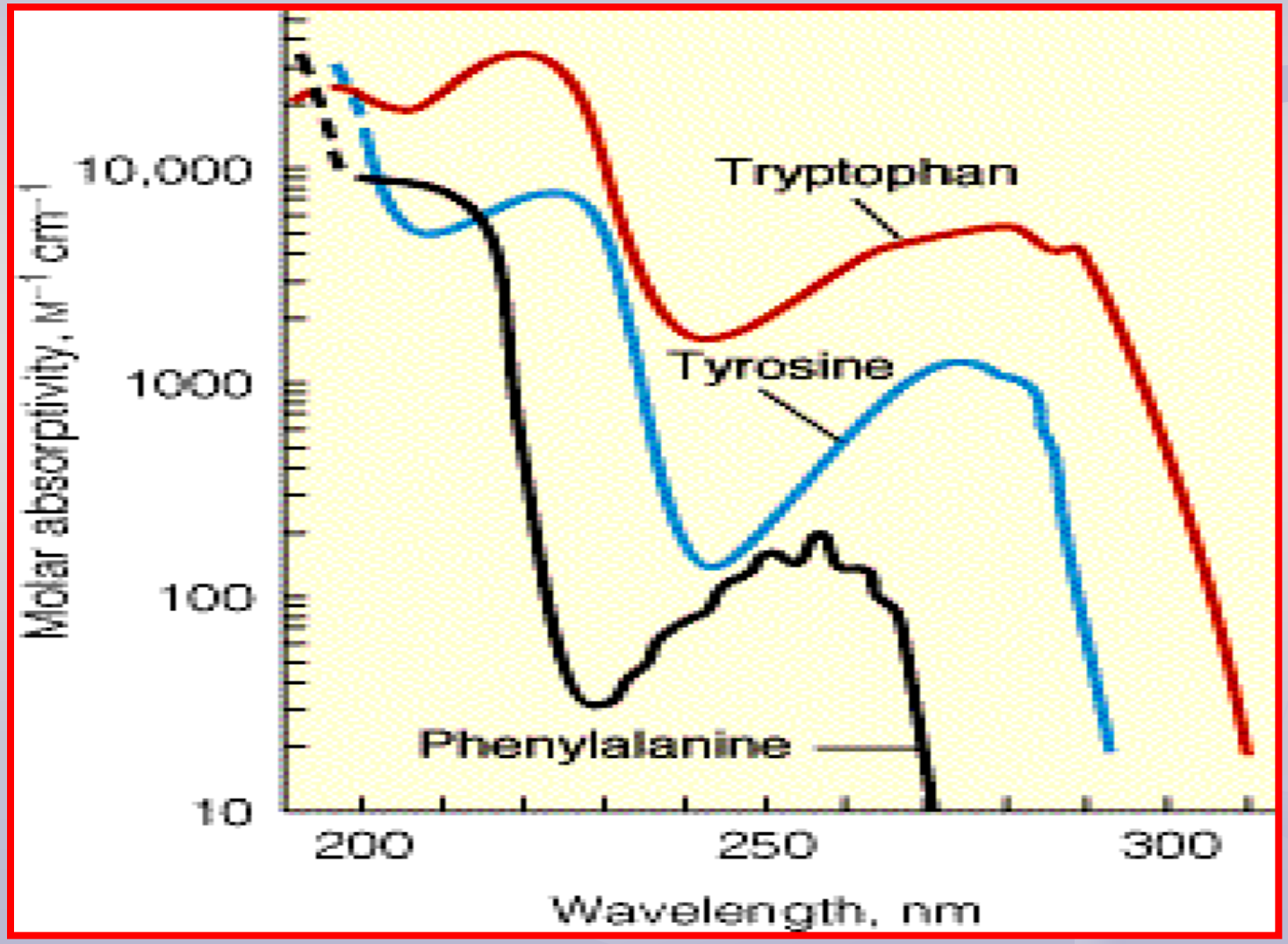
- All a.a's that are nonpolar , are : neutral and **Essential** ; have animal or plant sources, except L- Gly that is non essential.
- According to a.a with polar , uncharged R group:
  - OH containing a.a :-
    - L-Ser → Polar , Neutral and Non essential .
    - L-Tyr → Polar , weakly acidic & Non essential.
    - L- Thr → Polar , Neutral & essential.
  - SH- containing a.a :-
    - L-Cys→Polar,weakly acidic & Non essential.
    - L-Met → Non polar , Neutral & Essential .
  - Amide group → L- Gln & L- Asn are Polar , Neutral & Non essential .

Reversible formation of Disulfide bond



- A .A with positively charged R group:-
  - L- Lys → Polar , Essential & Basic (pKa = 10.5).
  - L- Arg → Polar , essential & Basic (pKa = 12.5).
  - L-His → Polar , Essential & Weakly Basic (pKa = 6.0) .This is the only a.a which can dissociate in neutral pH , it plays an important role in catalytic activity of some enzymes.
- A .A with Negatively charged R group :- L-Asp & L-Glu are Polar , Acidic & Non Essential .

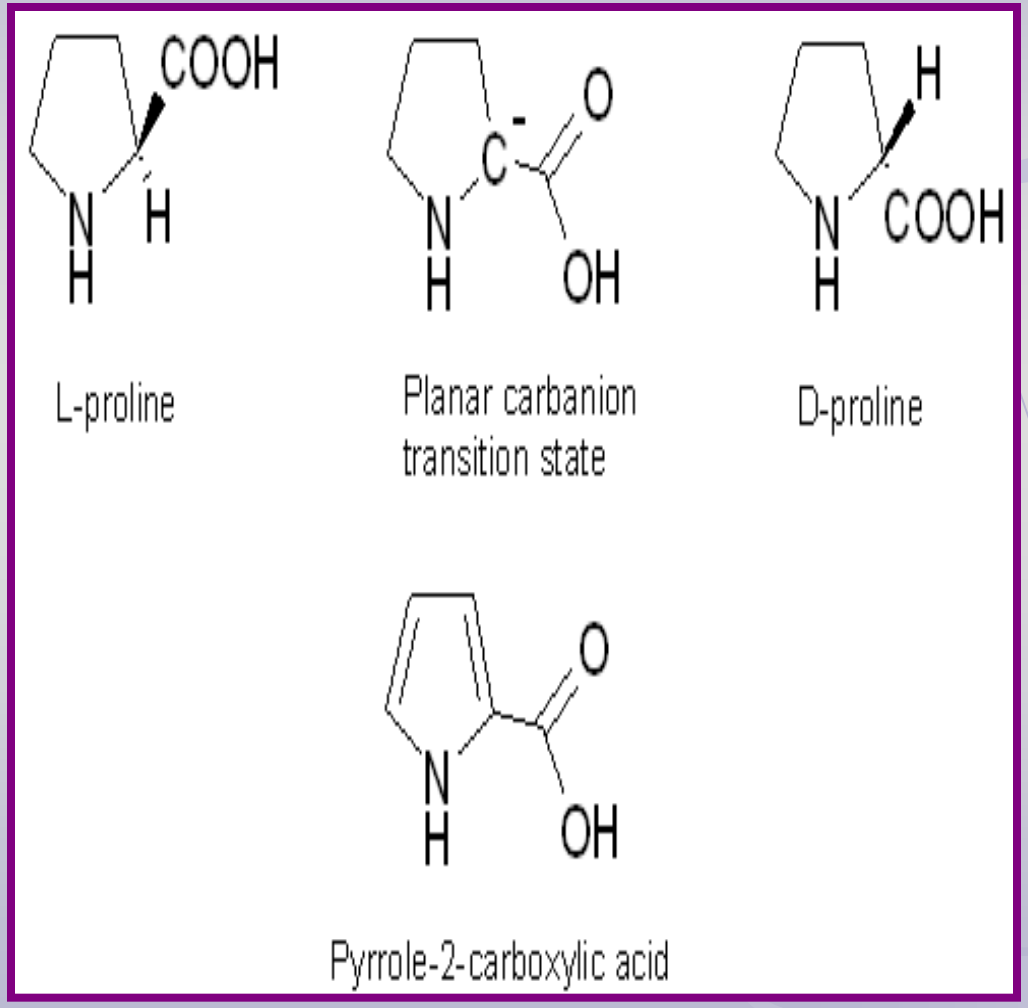
Absorption of UV by Aromatic A.A

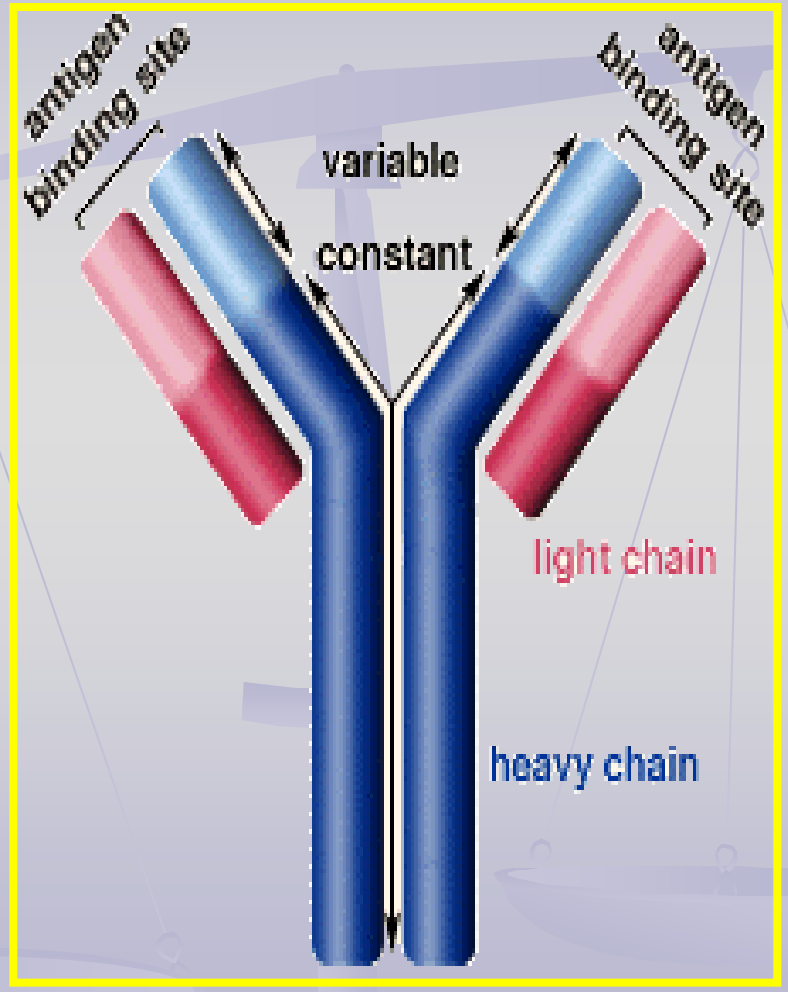


Note :-

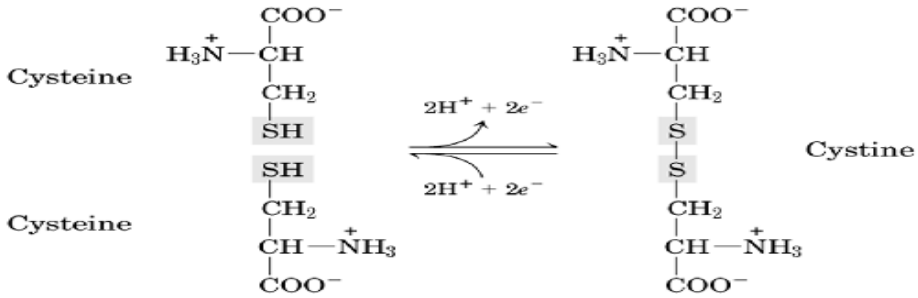
- **Proline** has an unusual imine ring structure (a secondary amine), where the terminal amine group is actually incorporated into the side chain.
- This causes changes to the secondary structure of a protein.
- Hydrophobic residues are often found in membrane bound proteins, and the aromatic ones contribute to protein absorbance at 280 nm, which is an important method of protein quantification.
- Proline is 2^{nd} a.a → Flexibility to protein structure of antibody .

Proline and Antibody

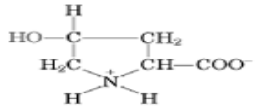


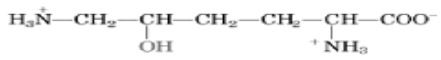


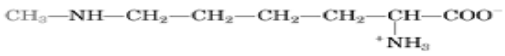
Disulfide bond formation

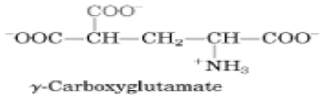


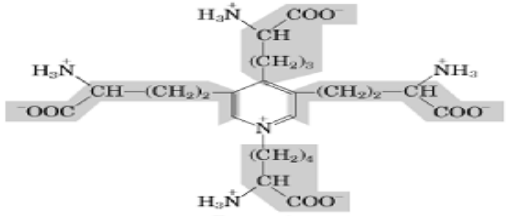
Nonstandard amino acids

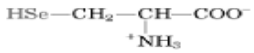

4-Hydroxyproline

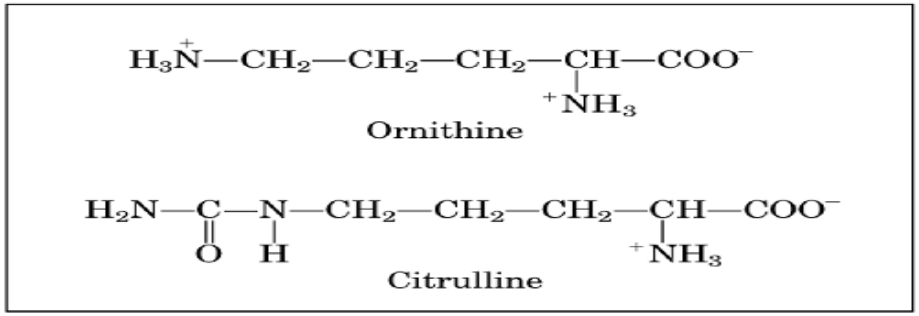

5-Hydroxylysine


6-*N*-Methyllysine


γ-Carboxyglutamate


Desmosine


Selenocysteine


Ornithine

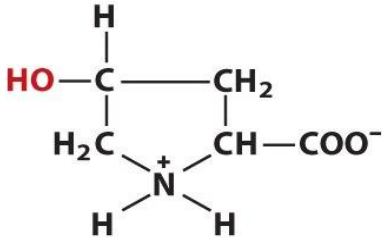
Citrulline

(b)

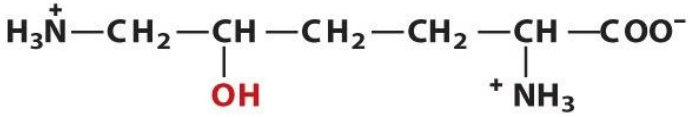
Uncommon A.A also have Important Function

- **4-Hydroxy Proline:** derivative of Proline .
- **5-hydroxy Proline :** derivative of Lysine .
- Both are found in Collagen in C.T .
- **N-Methyl lysine :** constituent of Myosin in muscle.
- **γ- Carboxy glutamate**: found in blood clotting protein prothrombin & certain other proteins that binds with Ca^{2+} as part of their biological function.
- **Desmosine :-** more complex a.a , a derivative of 4 Lys residues , found in fibrous protein elastin.

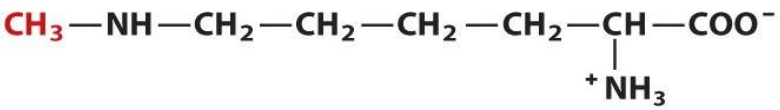
Some non-standard amino acids



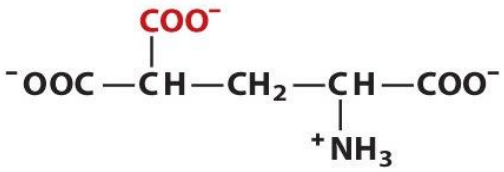
4-Hydroxyproline



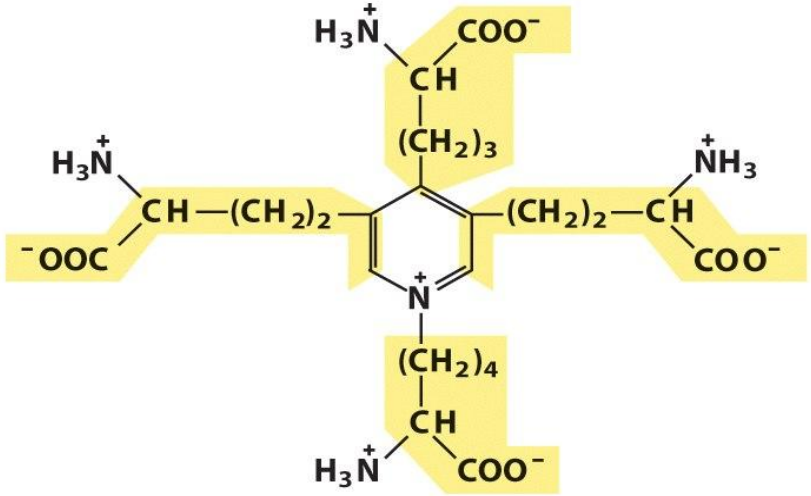
5-Hydroxylysine



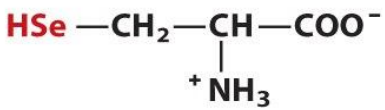
6-*N*-Methyllysine



γ-Carboxyglutamate

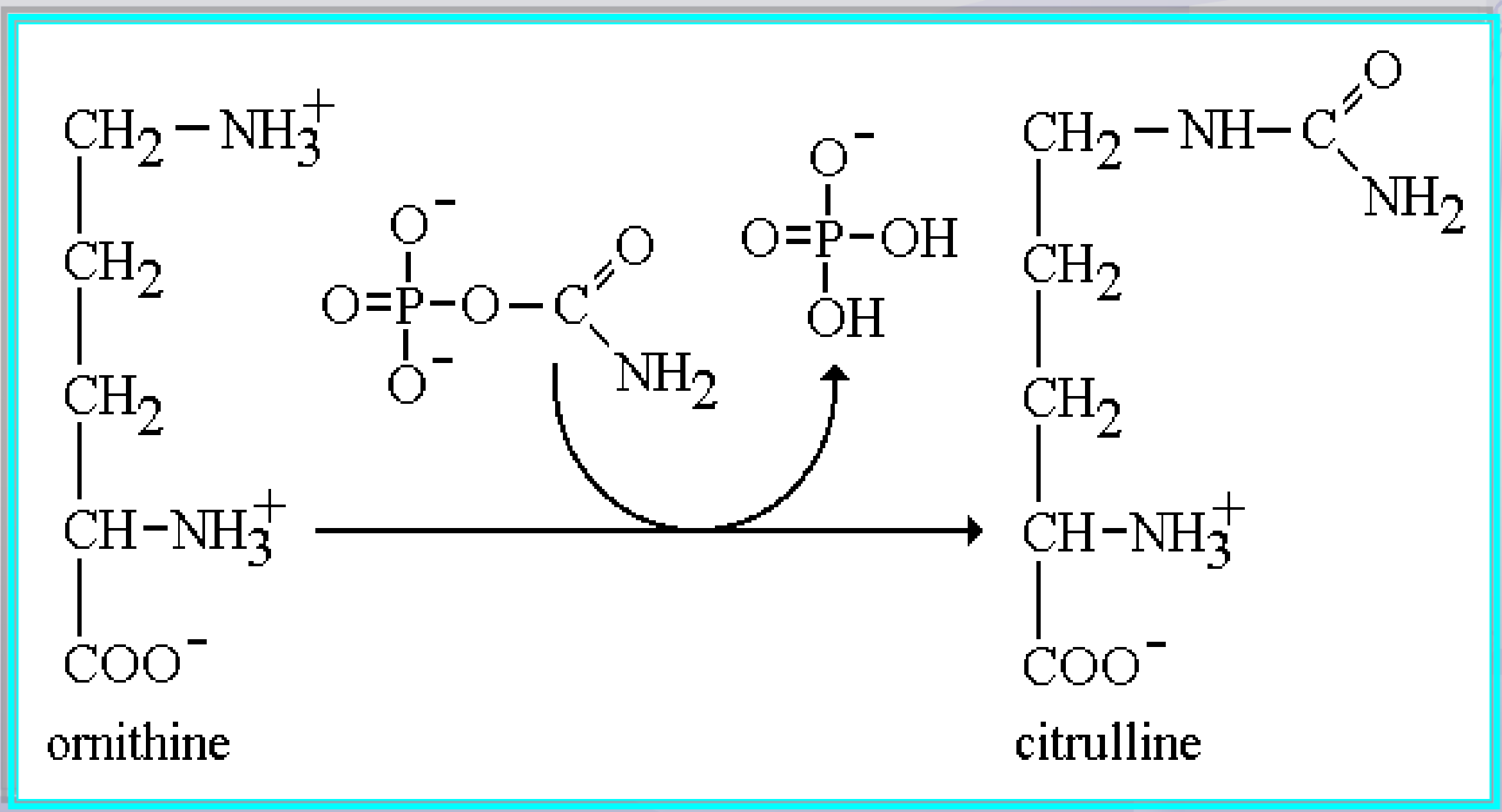


Desmosine

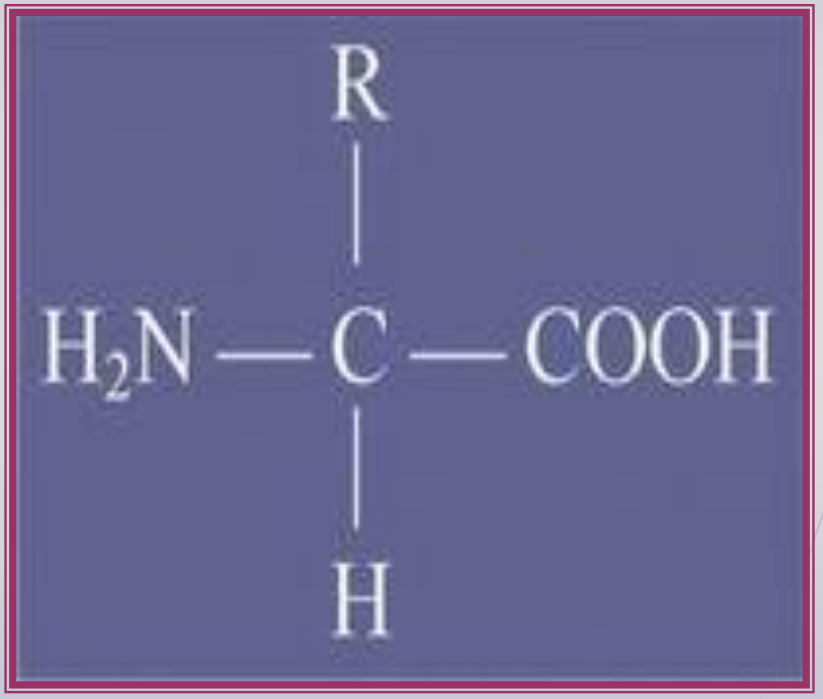


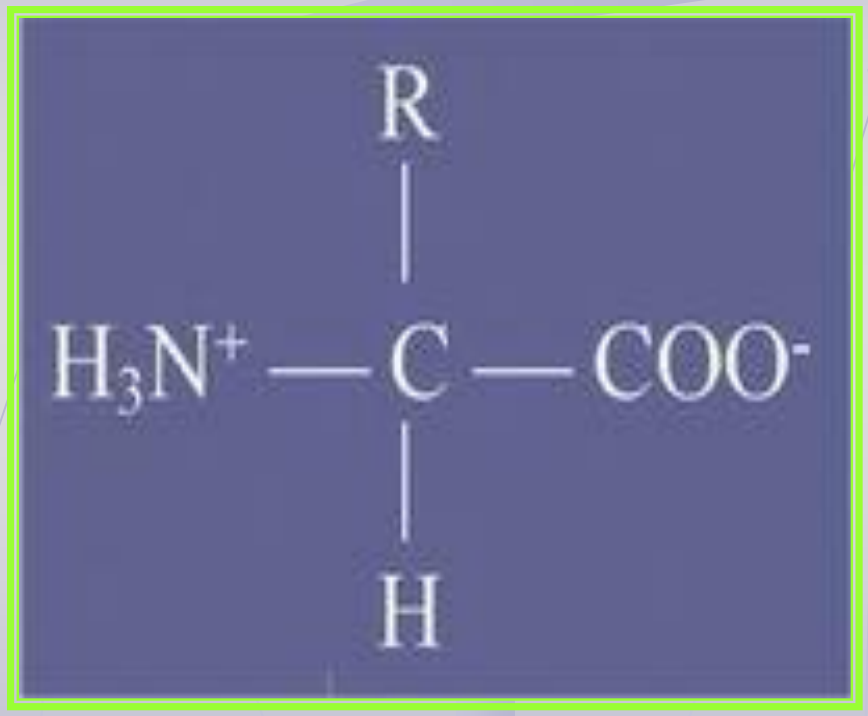
Selenocysteine

- **Selenocysteine:-** rare a.a residue , is introduced during protein synthesis , derived from Ser .
- **Ornithine & Citrulline →** both are metabolites in Arginine & in Urea cycle

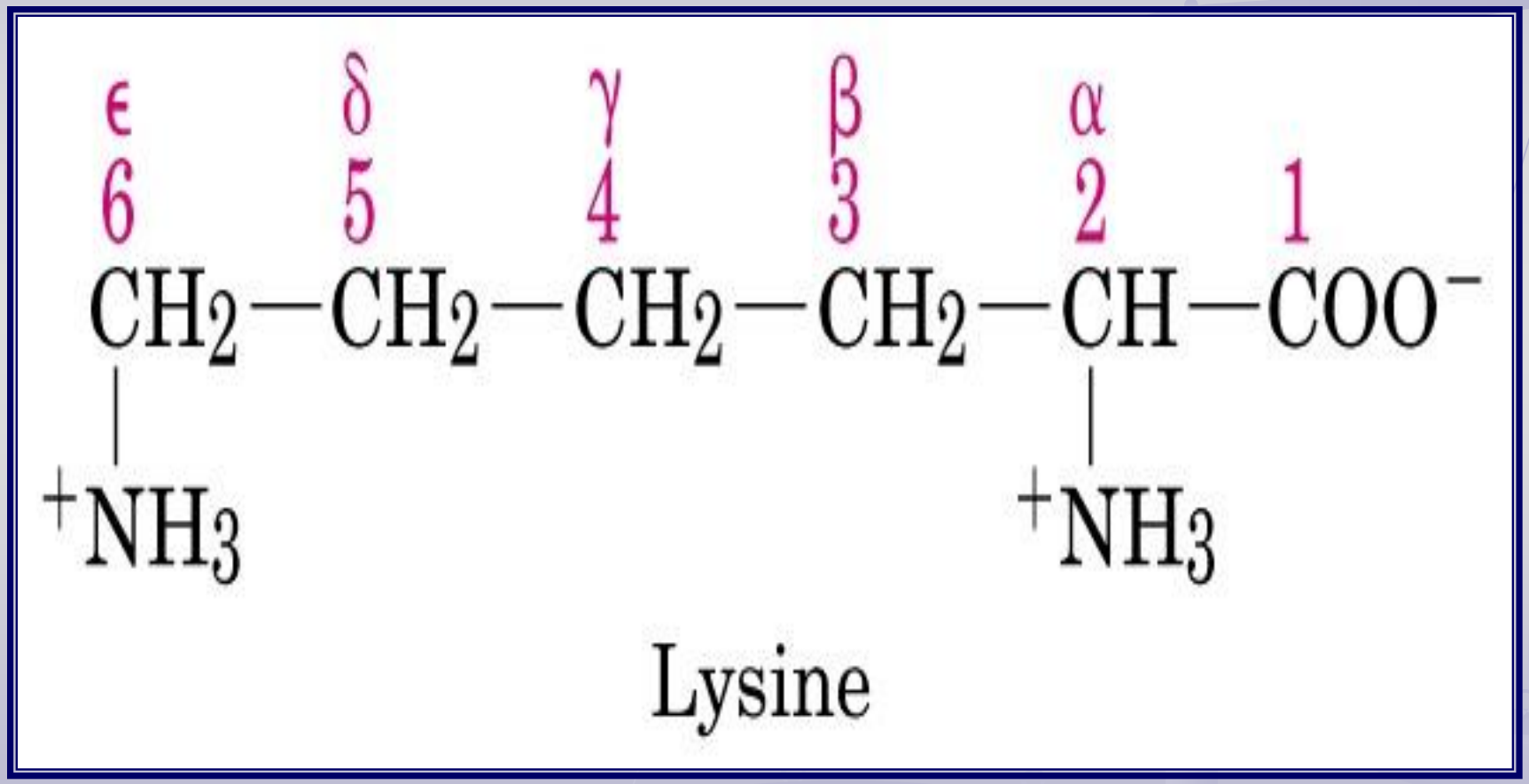


Nonionic and Zwitterionic forms of amino acids



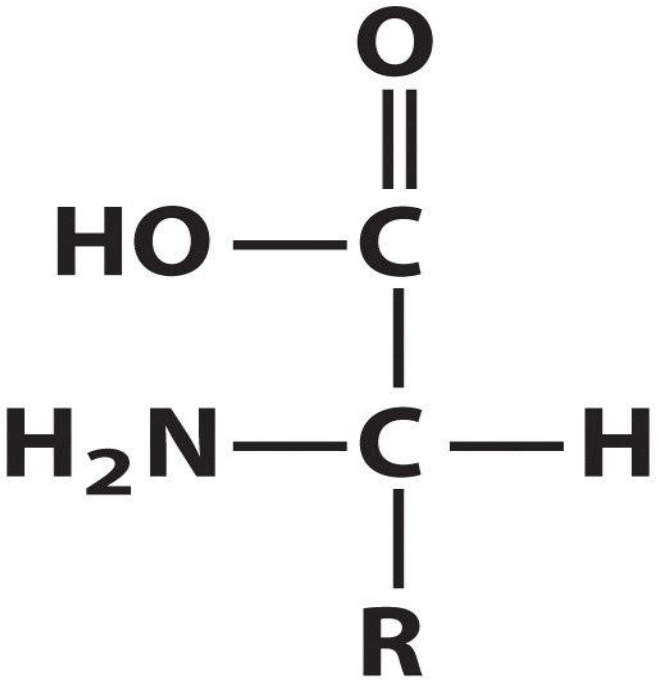


Side-chain Carbons are labeled Starting from Main Chain

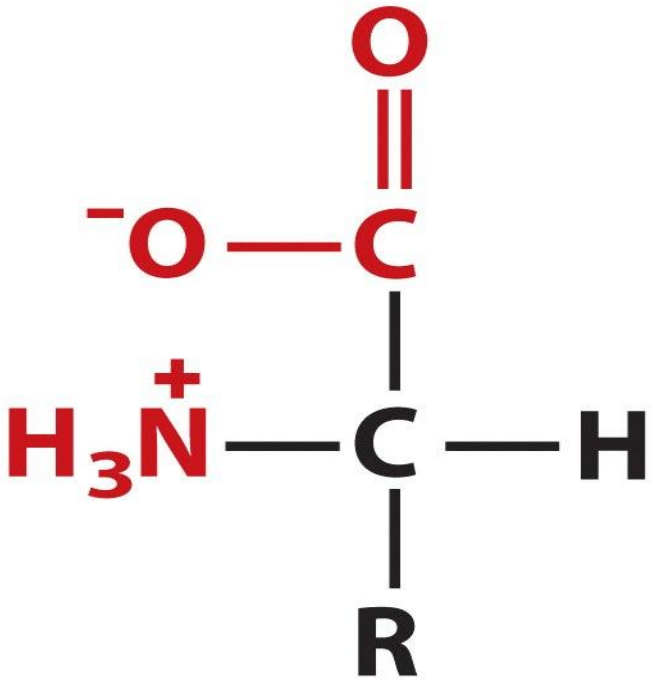


Lysine

Amino Acids are Zwitterions at Neutral pH

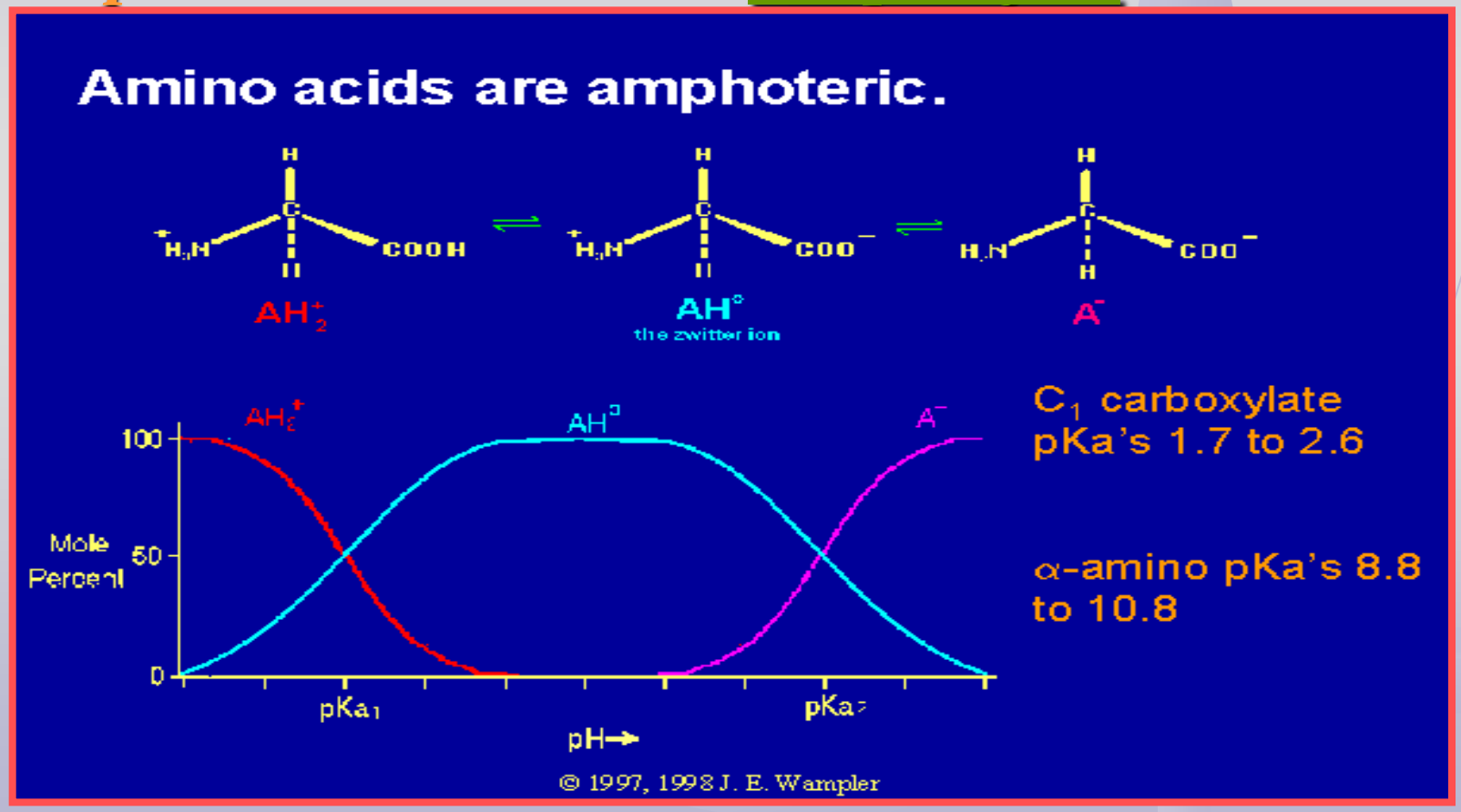


Nonionic form



Zwitterionic form

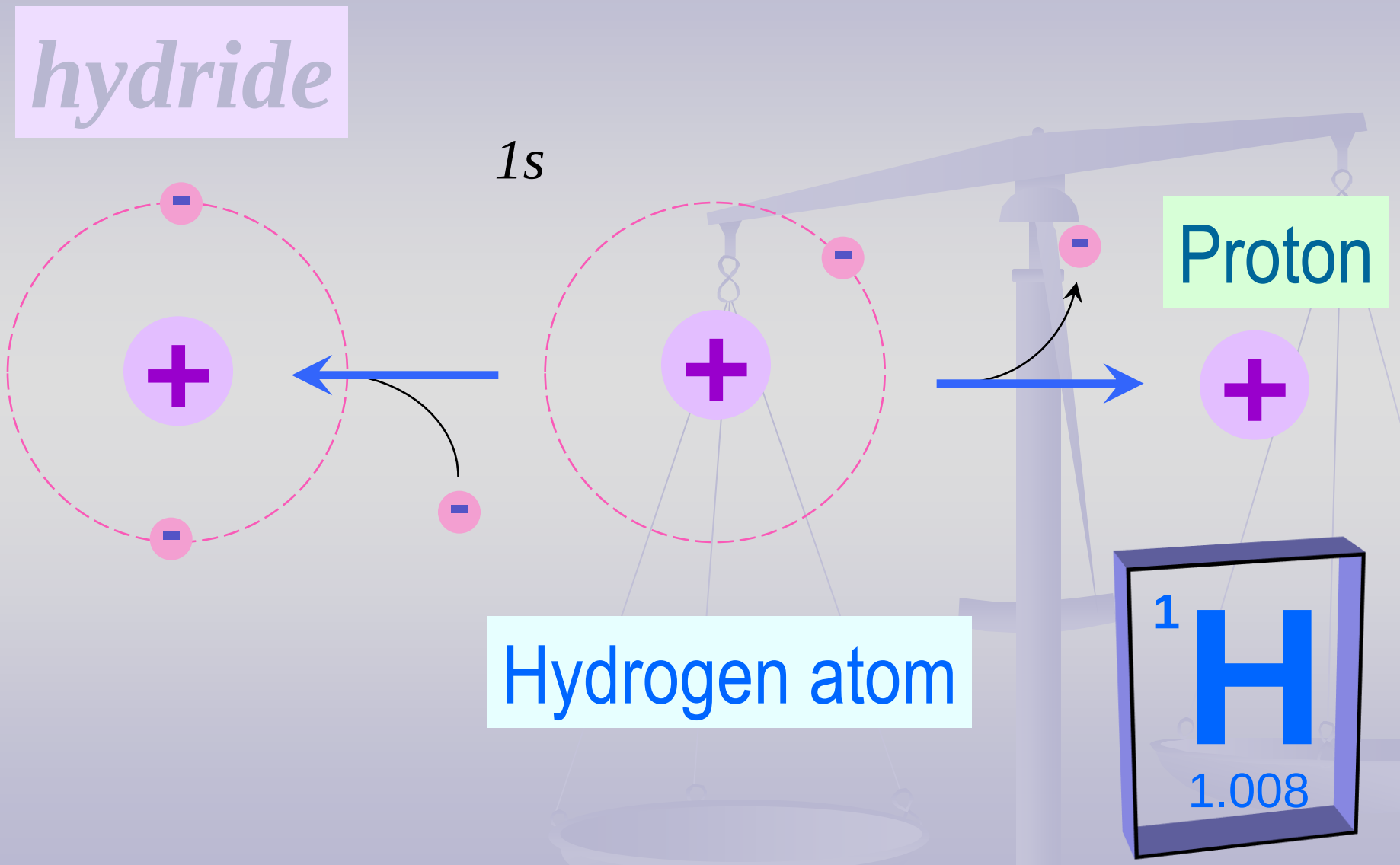
- A.A has dipolar ion form "**Zwitterion**" → (hybrid ion) that act as acid [Proton Donor] or as base [Proton Acceptor] .
- The substance that has this dual nature are ***Amphoteric*** & are called **Ampholytes .**



Titration Curve for A.A

- pK_1 for Carboxyl group COO^- .
- pK_2 for amine group NH_3^+ .
- Gly presents as dipolar form & fully ionized , but with no net electric charge (net –ve charge) .
- The characteristic - his net electric charge = 0 – is called **Isoelectric point (pI) Isoelectric point (pI)** or **Isoelectric pH**.
- At pH = 5.97 , pI is found between 2 stages in its titration curve .
- When pI = pH → no ionizable group at side chain → **$pI = pK_1 + pK_2 / 2$** .

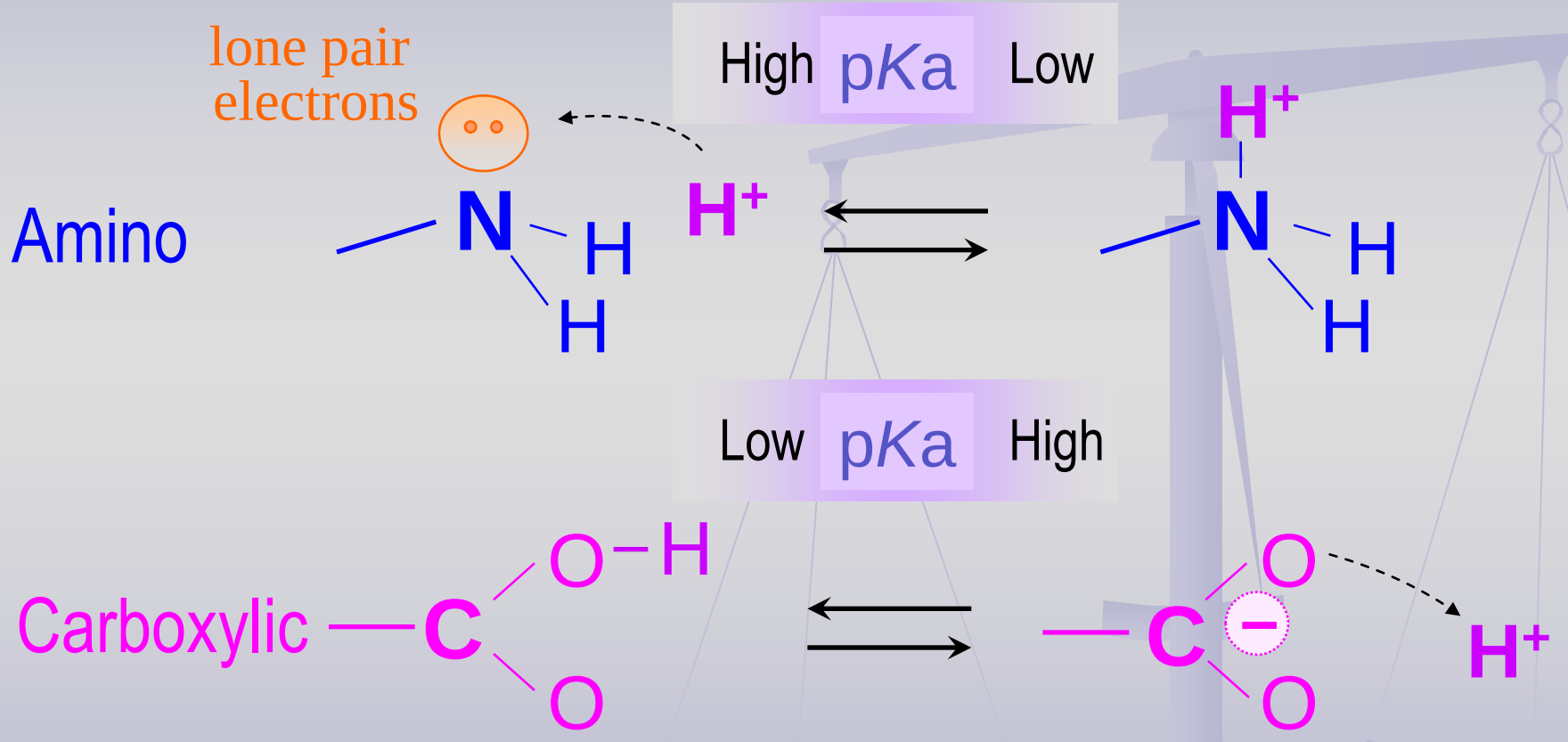
Hydride, Hydrogen and Proton



Juang RH (2004) BCbasics

Proton Is Adsorbed or Desorbed

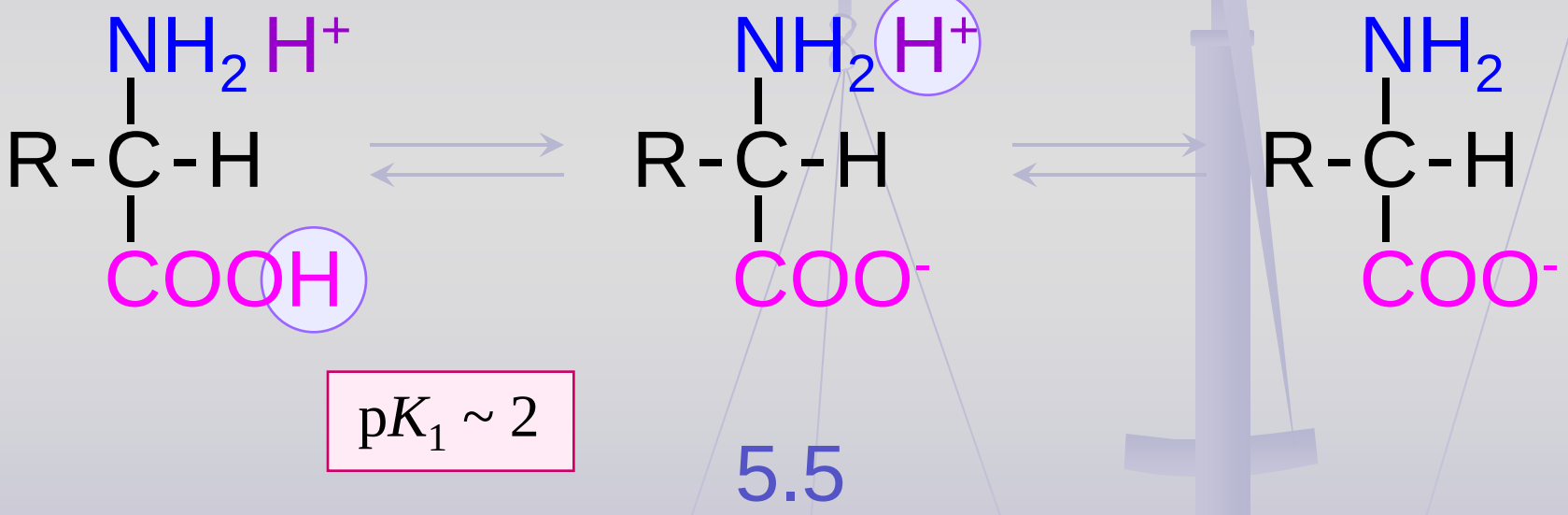
Proton： abundant and small, affects the charge of a molecule



Ampholyte contains both positive and negative groups on its molecule

Juang RH (2004) BCbasics

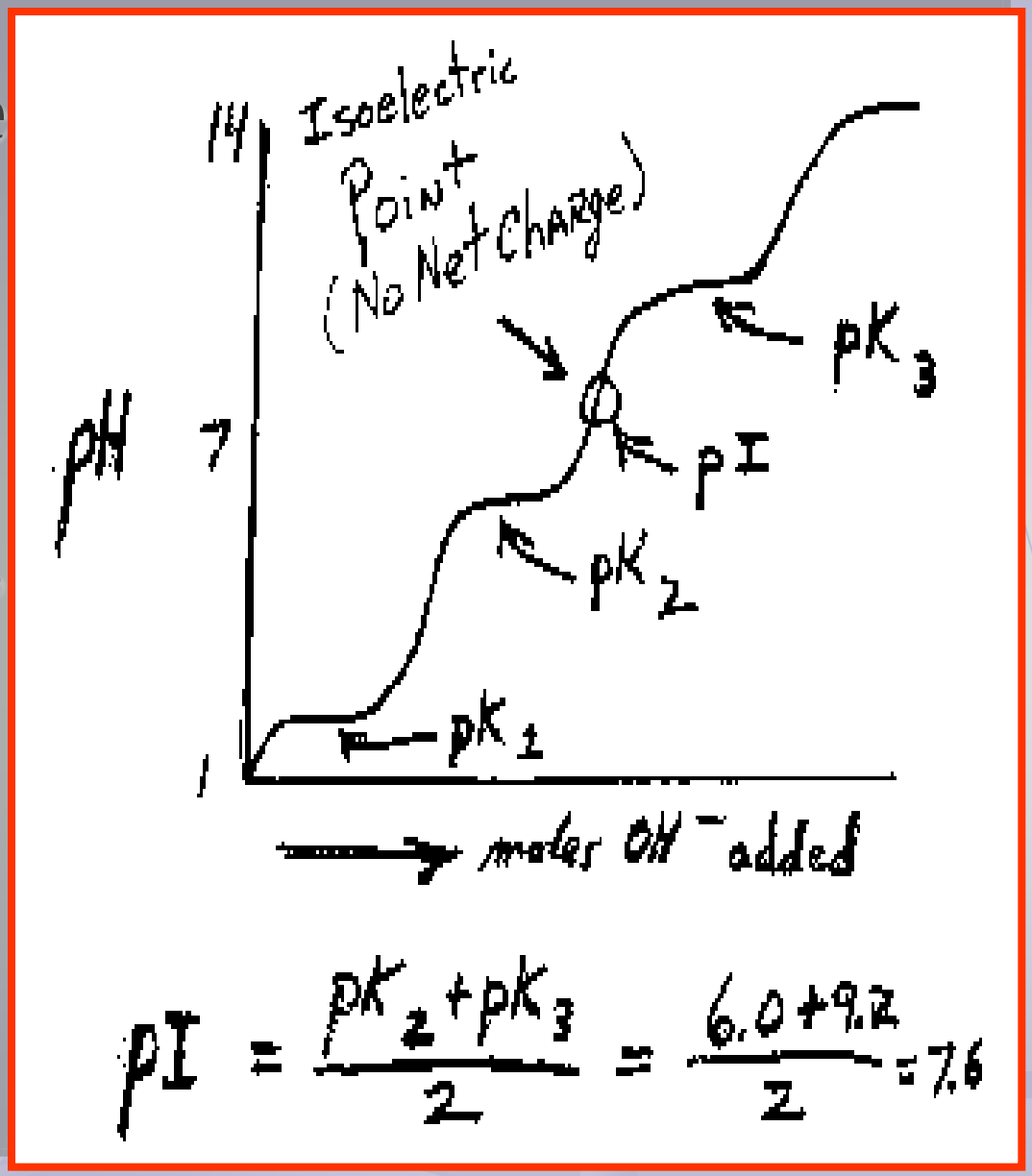
Acidic environment Neutral environment Alkaline environment



+1 0 -1

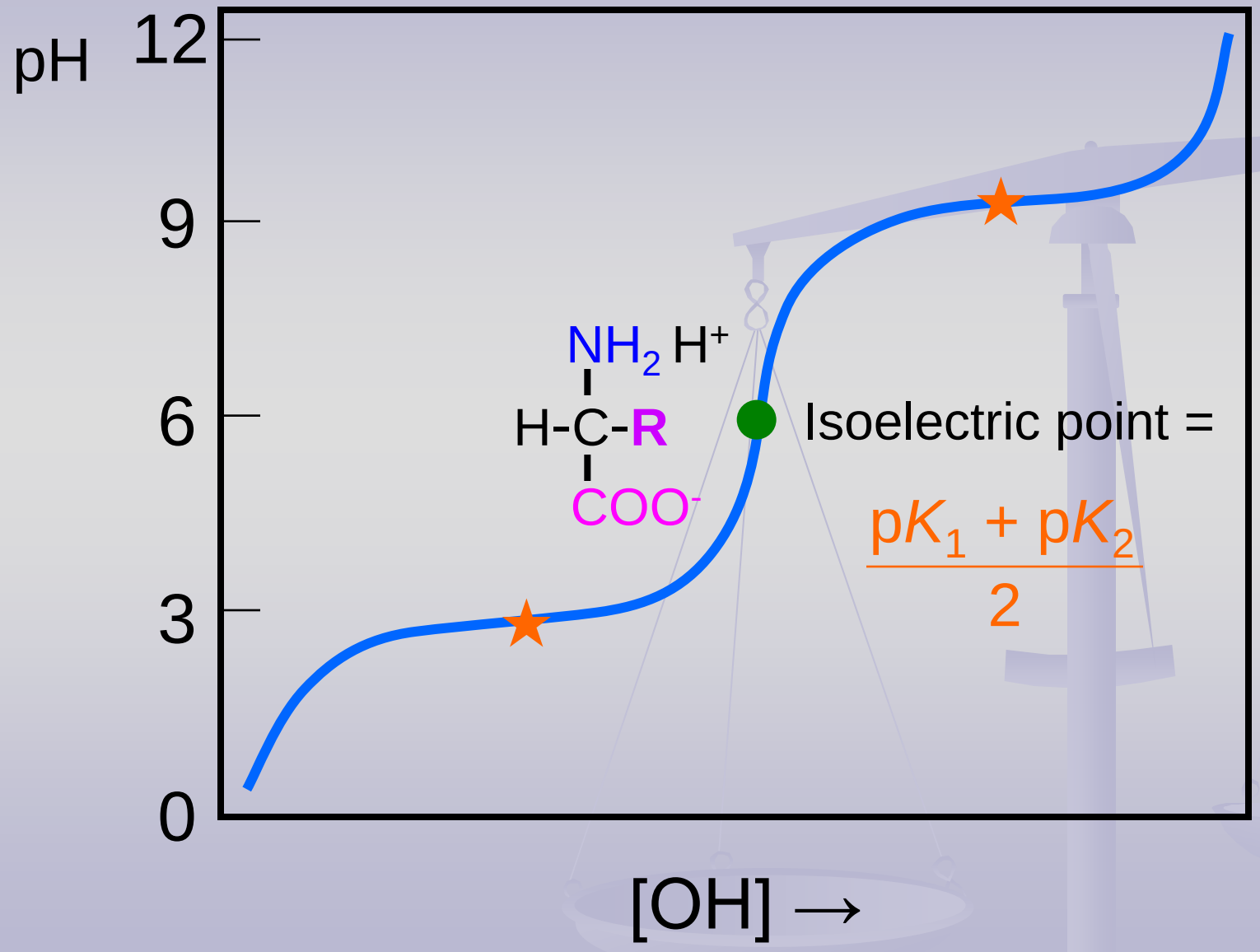
Isoelectric point

- For Gly :pI =5.97
  → Gly have net –ve charge at any pH > pI & will move towards the +ve electrode or anode in the electric field , and visa versa .
- pI for acidic a.a = **$pK_{a1} + pK_{a2} / 2$** .
- pI for basic a.a = **$pK_{a2} + pK_{a3} / 2$** .



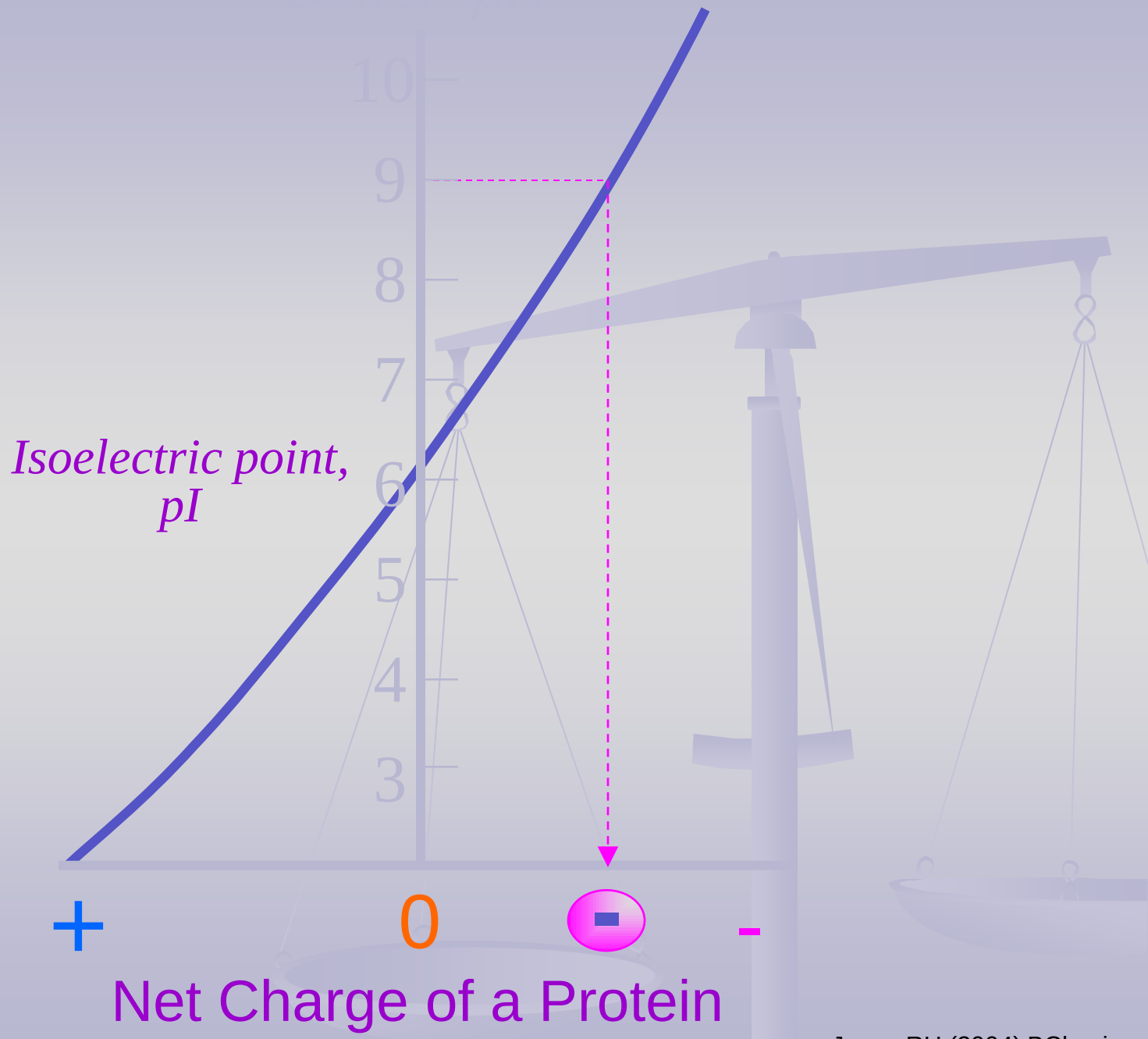
pI calculation for His

Amino Acids Have Buffering Effect



Juang RH (2004) BCbasics

Environment pH vs Protein Charge



Juang RH (2004) BCbasics

p*Ka* of Amino Acid Residues

Residues on amino acids can release or accept protons

α-COOH	⇌	α-COO$^-$ + H$^+$	p*Ka* =	1.8~2.4
R-COOH	⇌	R-COO$^-$ + H$^+$	p*Ka* =	3.9~4.3
His-Imidazole·H$^+$	⇌	His-Imidazole + H$^+$	p*Ka* =	6.0
Cys-SH	⇌	Cys-S$^-$ + H$^+$	p*Ka* =	8.3
Tyr-OH	⇌	Tyr-O$^-$ + H$^+$	p*Ka* =	10
α-NH$_3^+$	⇌	α-NH$_2$ + H$^+$	p*Ka* =	8.8~11
R-NH$_3^+$	⇌	R-NH$_2$ + H$^+$	p*Ka* =	10~12.5

Smaller p*Ka* releases proton easier

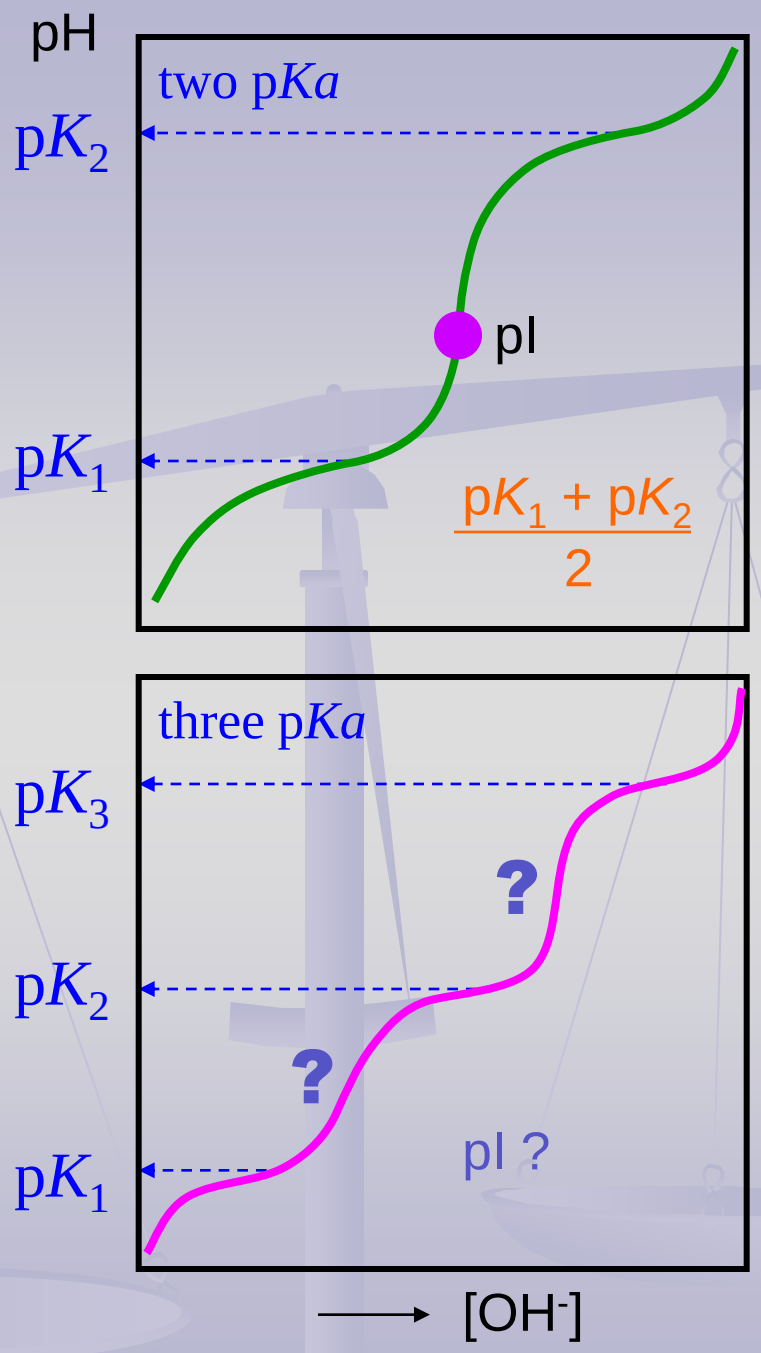
Only His has the residue with a neutral p*Ka* (imidazole)

p*Ka* of α carboxylic or amino groups is lower than p*Ka* of the R residues

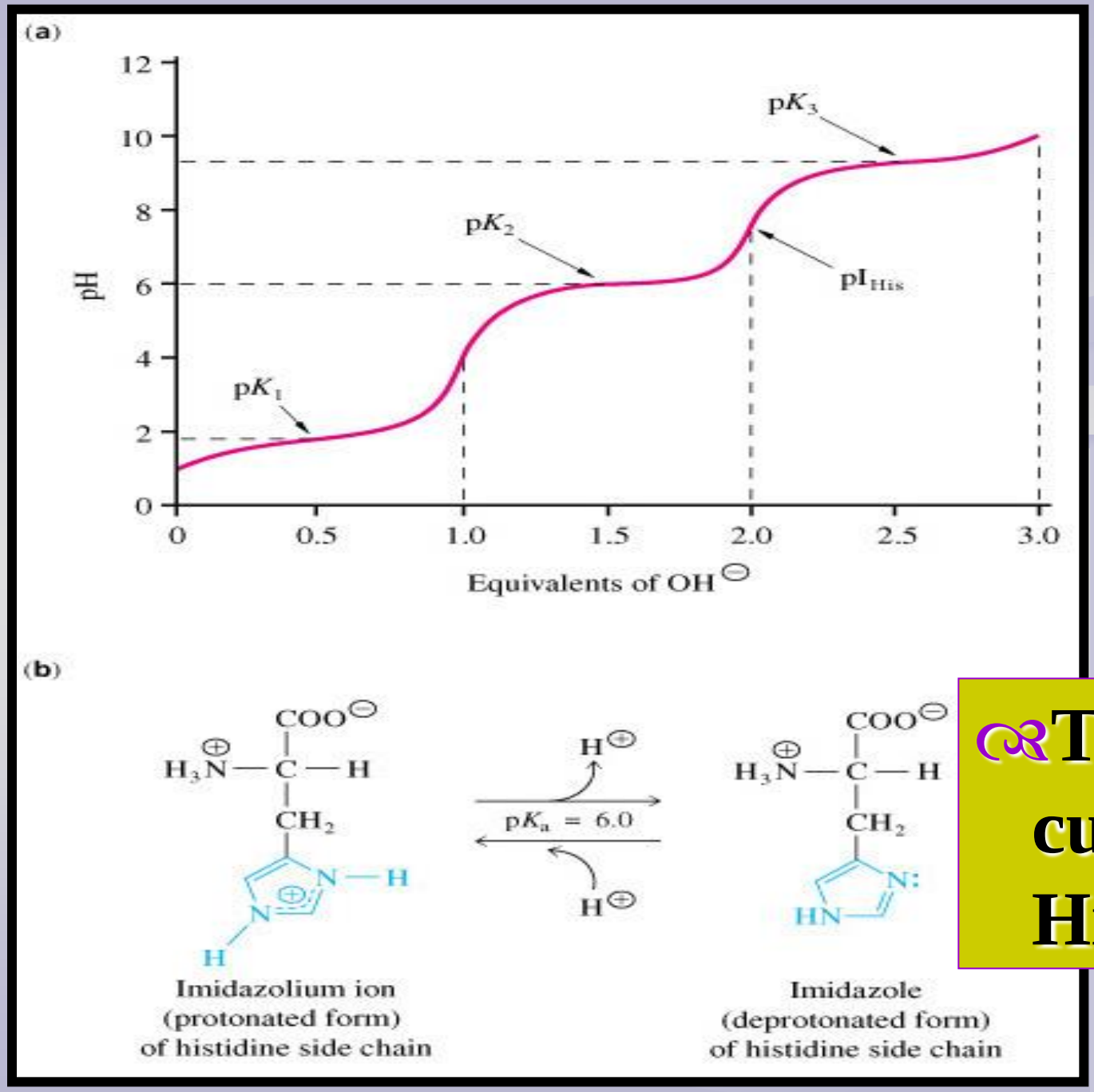
Juang RH (2004) BCbasics

pKa of Amino Acids

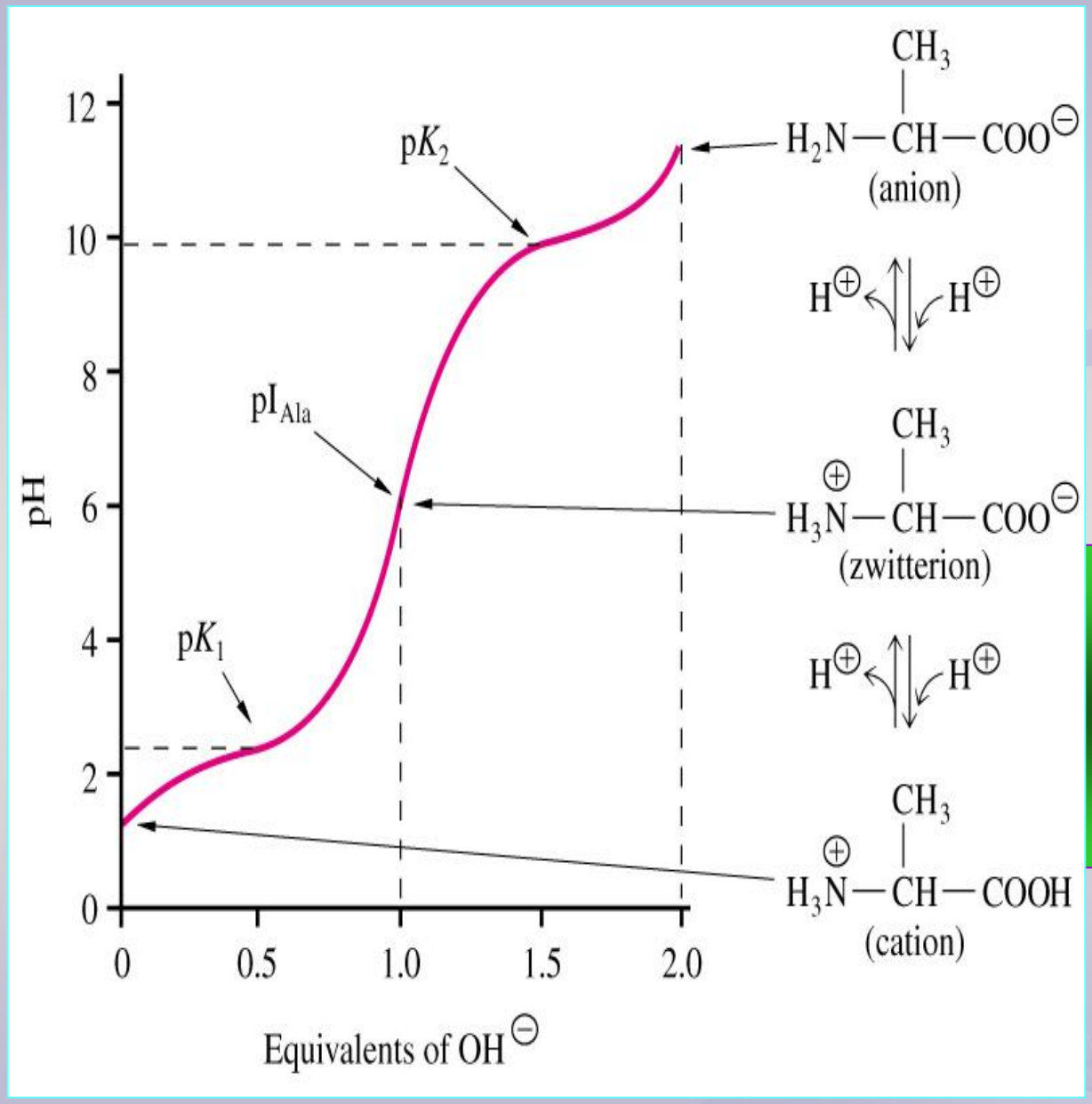
Amino acids		-COOH	-NH_2	-R
Gly	G	2.34	9.60	
Ala	A	2.34	9.69	
Val	V	2.32	9.62	
Leu	L	2.36	9.68	
Ile	I	2.36	9.68	
Ser	S	2.21	9.15	
Thr	T	2.63	10.4	
Met	M	2.28	9.21	
Phe	F	1.83	9.13	
Trp	W	2.38	9.39	
Asn	N	2.02	8.80	
Gln	Q	2.17	9.13	
Pro	P	1.99	10.6	
Asp	D	2.09	9.82	3.86
Glu	E	2.19	9.67	4.25
His	H	1.82	9.17	6.0
Cys	C	1.71	10.8	8.33
Tyr	Y	2.20	9.11	10.07
Lys	K	2.18	8.95	10.53
Arg	R	2.17	9.04	12.48



Juang RH (2004) BCbasics

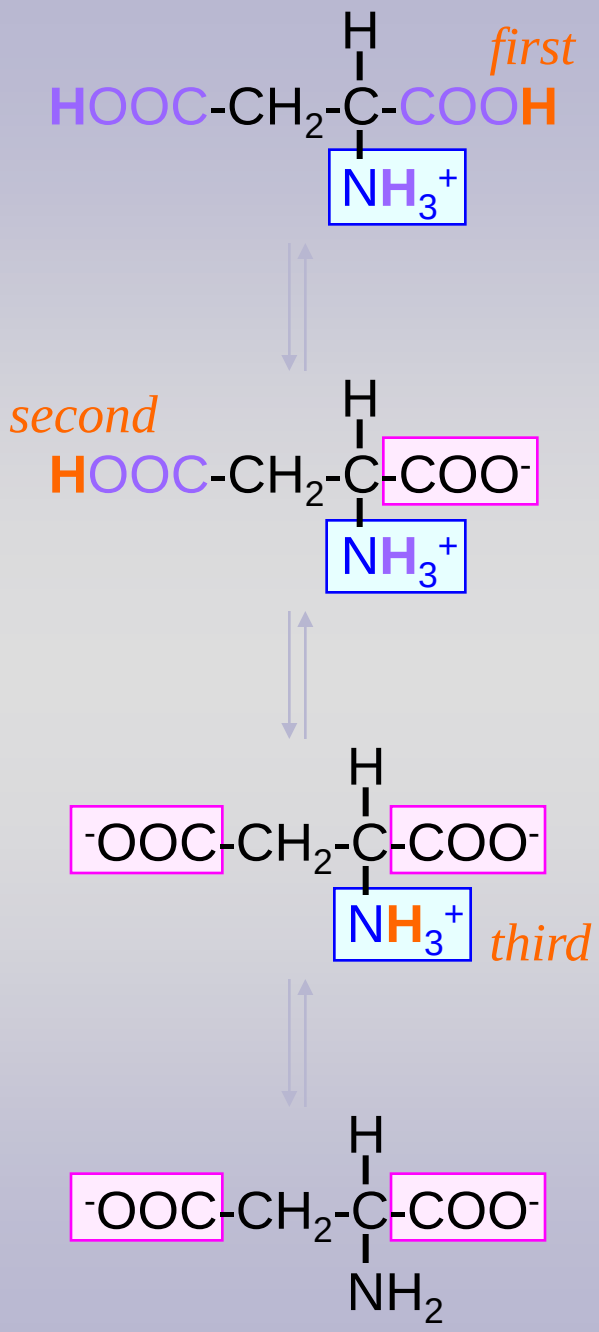


Titration curve for His



Titration Curve for Ala

Aspartic acid



+1

$pK_1 = 2.1$

0

$pK_2 = 3.9$

-1

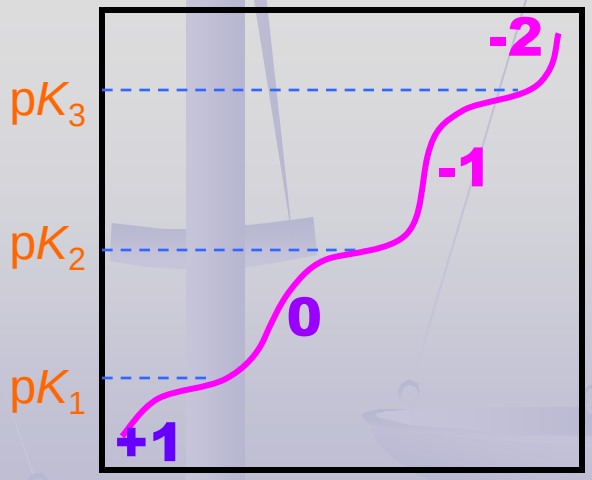
$pK_3 = 9.8$

-2

Isoelectric point is the average of the two p*K*a flanking the zero net-charged form

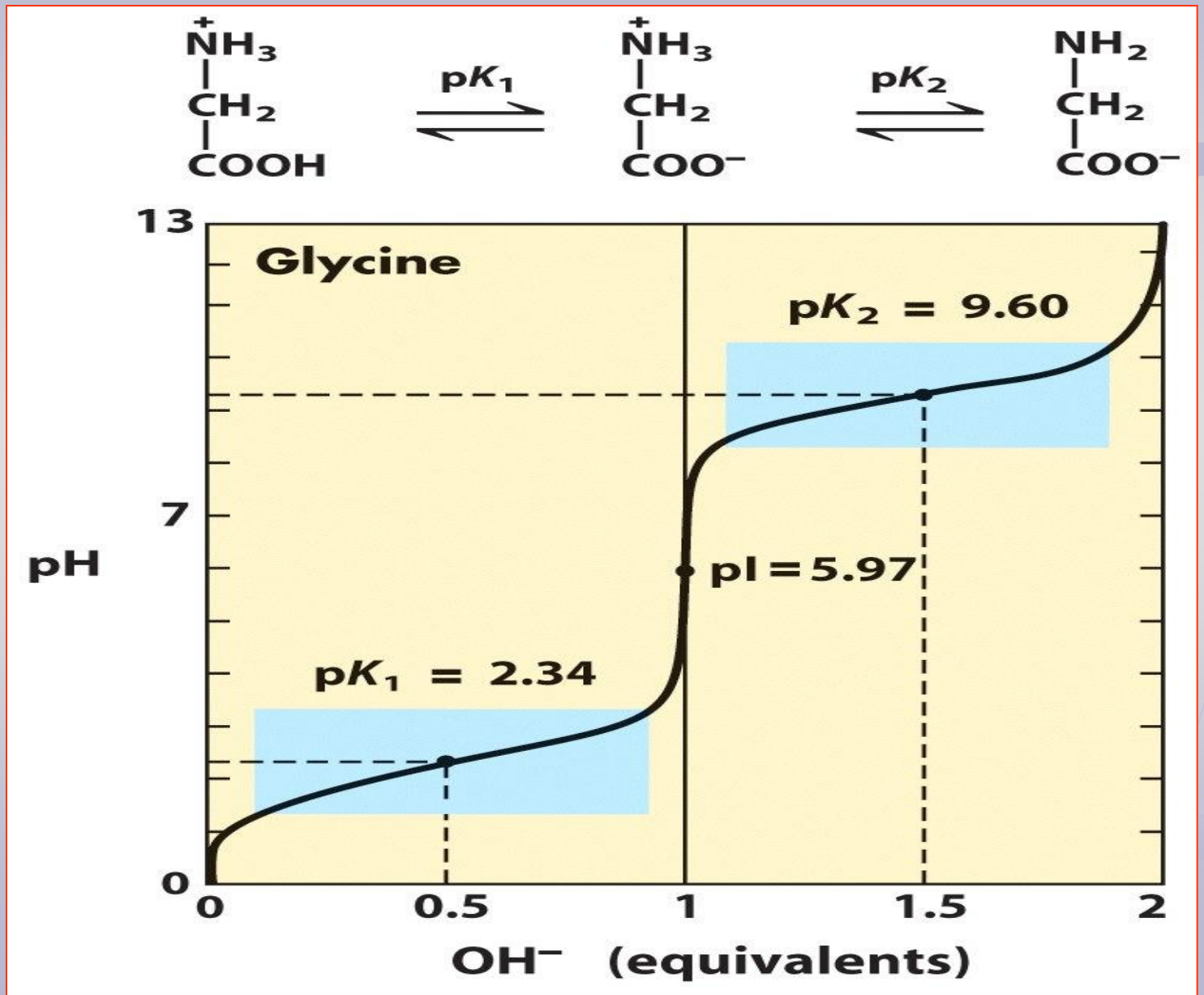
$$\frac{2.1 + 3.9}{2} = 3.0$$

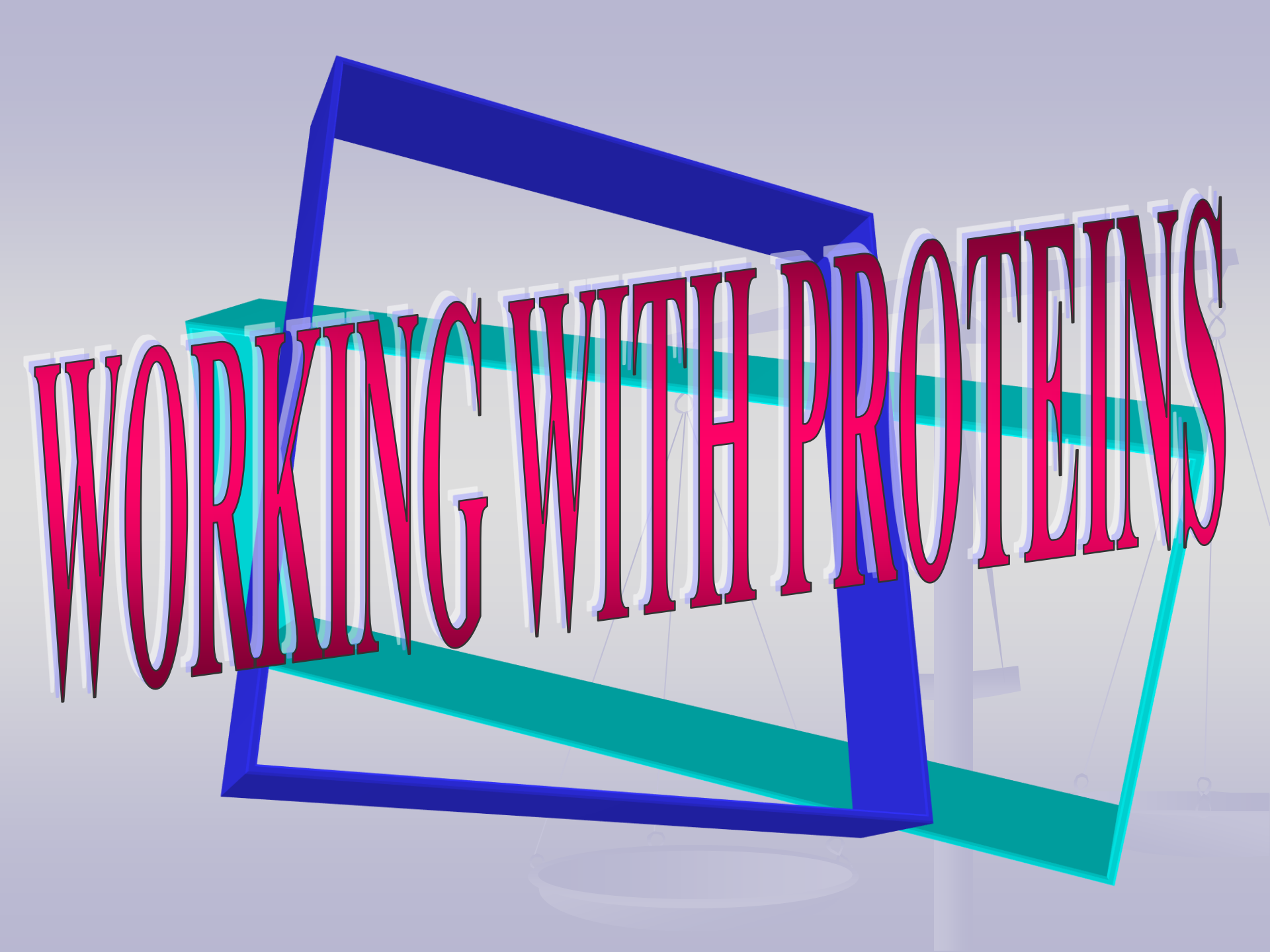
Isoelectric point



Juang RH (2004) BCbasics

Titration of an Amino Acid

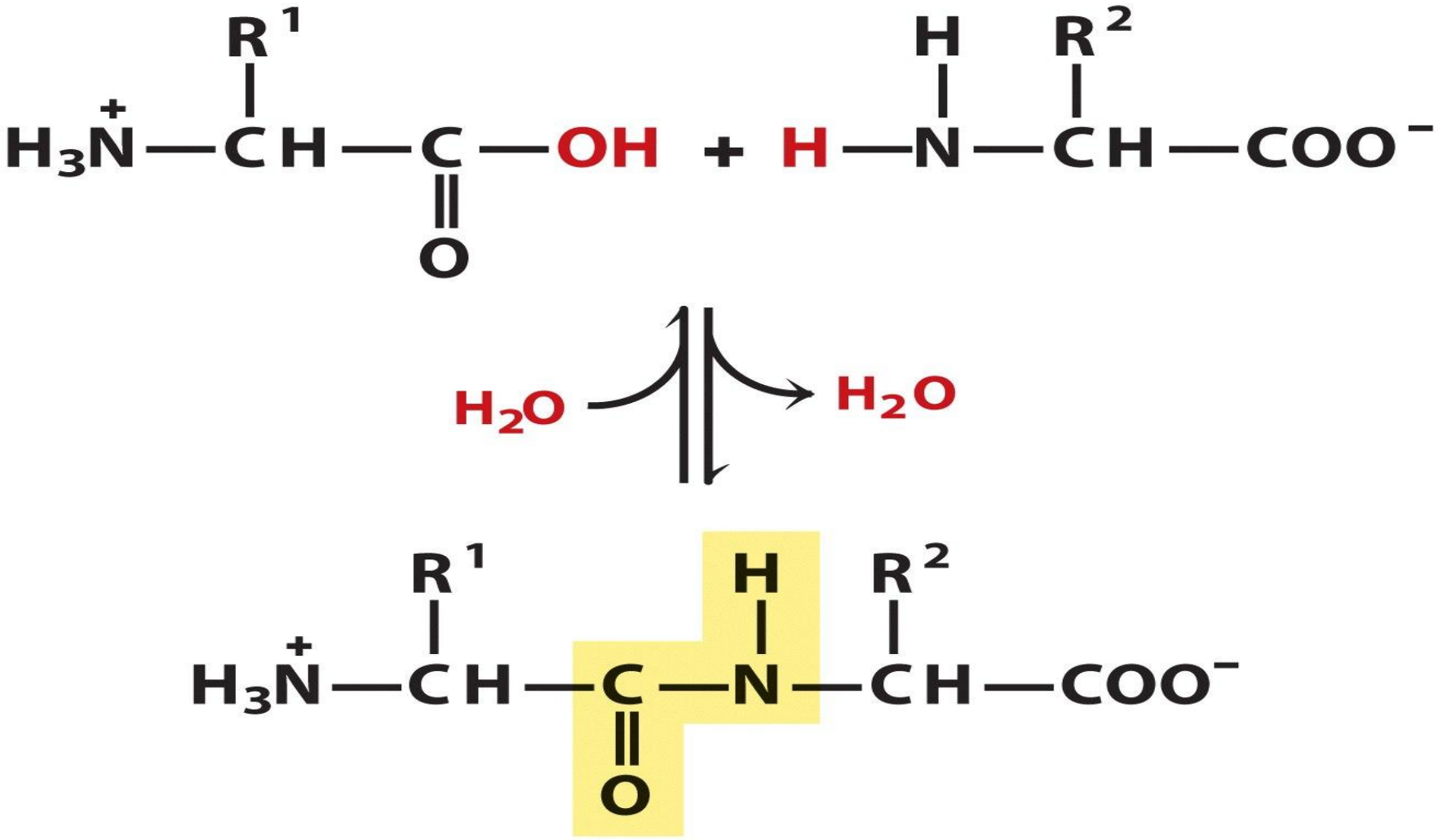



WORKING WITH PROTEINS

The Covalent structure of Proteins

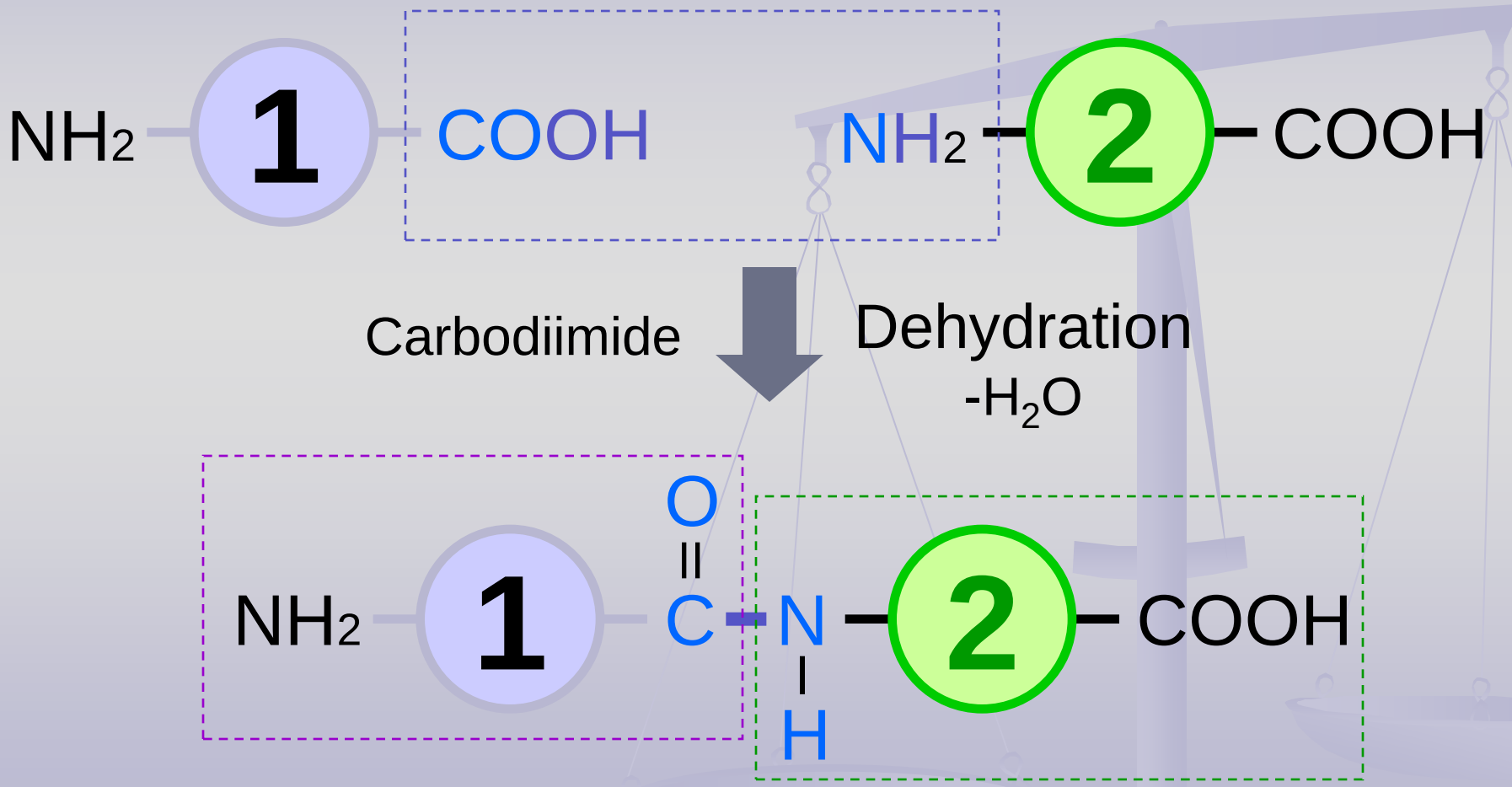
- The function of a protein depends on it's a.a sequence .
- In human , some proteins are polymorphic ; having a.a sequence variants in the human population.
- A.A sequences of millions of Proteins have been determined .

Amino Acids Are Joined Through Peptide Bonds

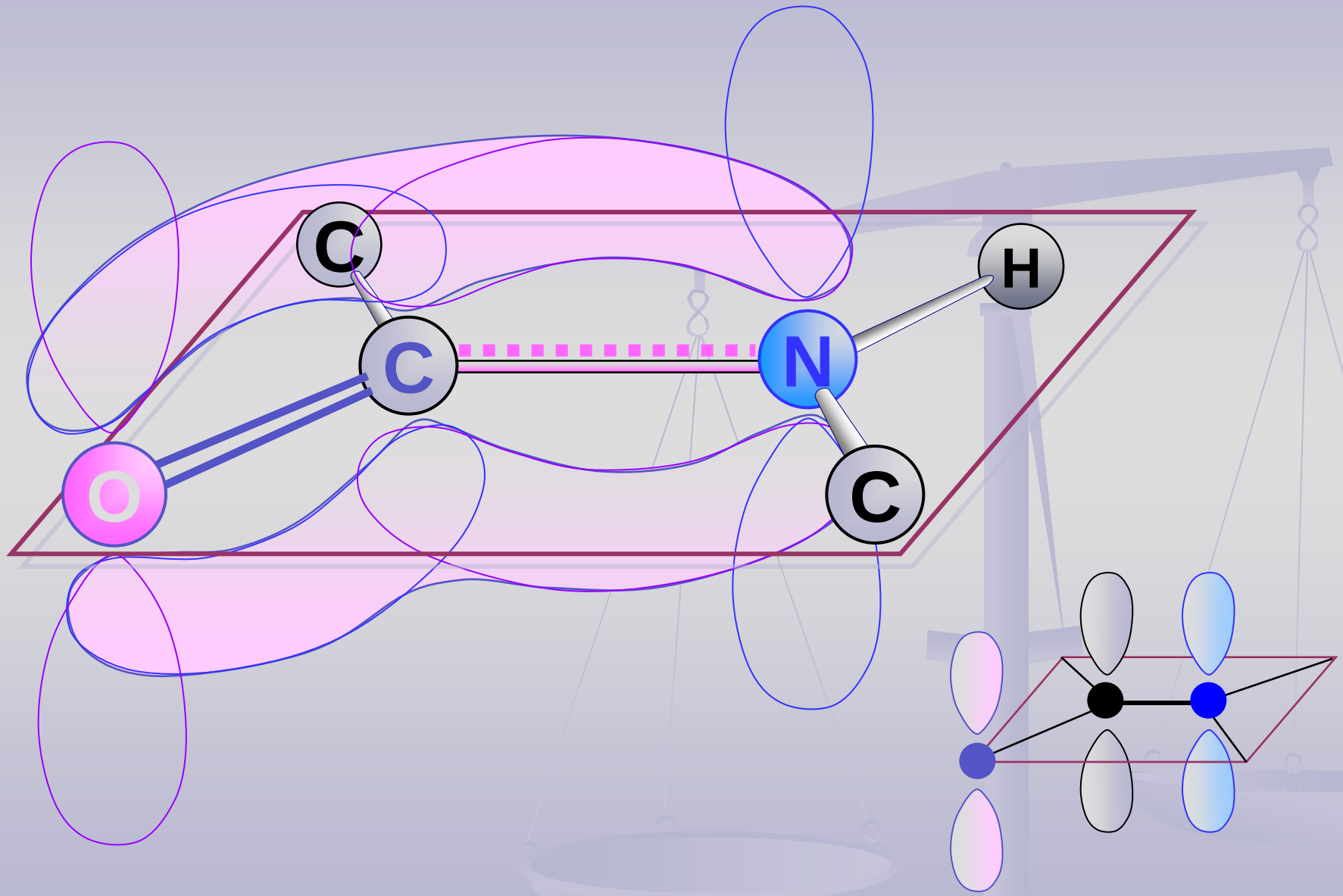


Formation of Peptide Bonds by Dehydration

Amino acids are connected head to tail



Peptide Bond Is Rigid and Planar



Juang RH (2004) BCbasics

Reaction of a.a side chains

1. *Polar – Dervitizing reagent*

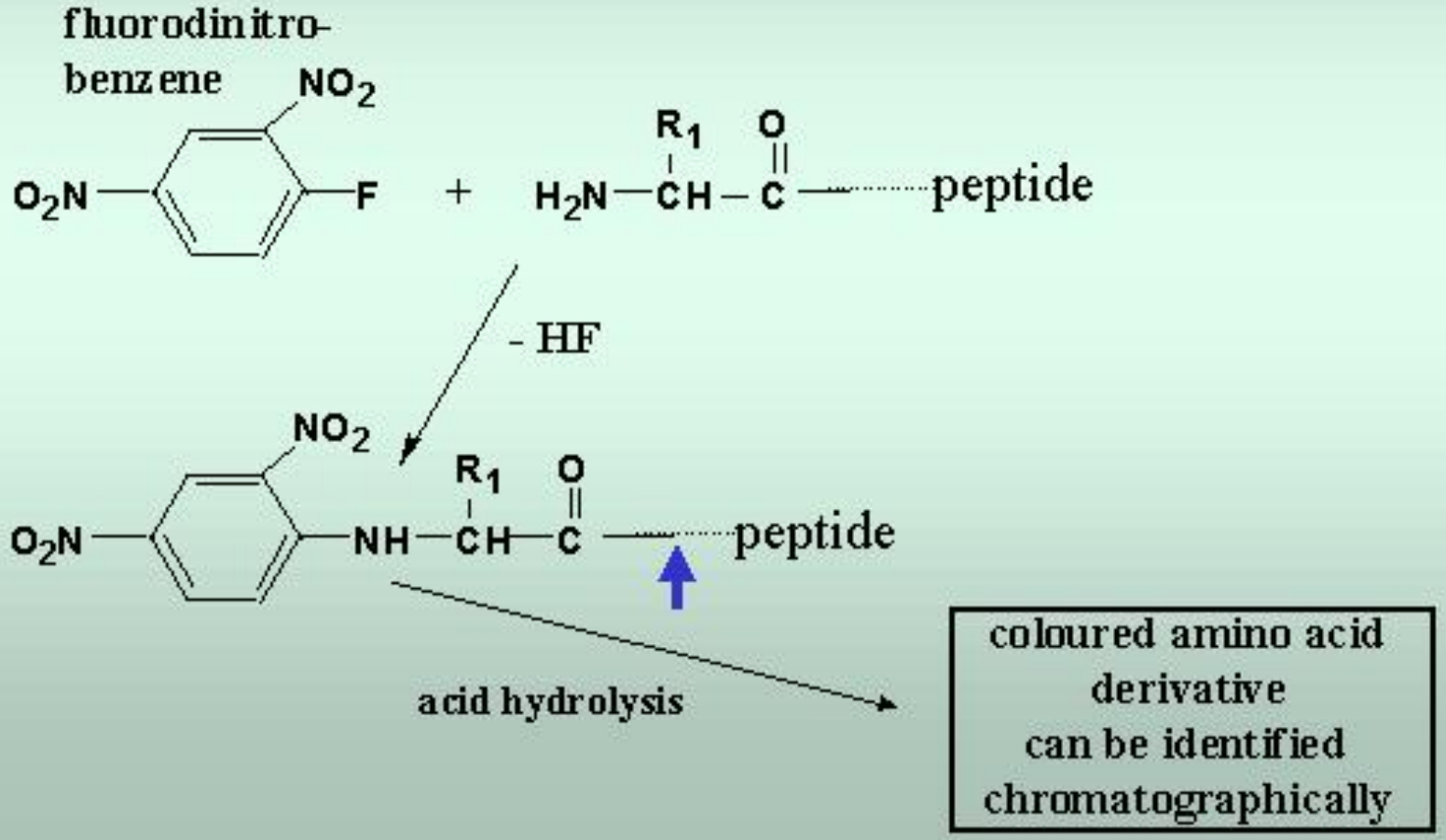
- Sanger developed 1- F- 2,4 – Dinitrobenzene **(FDNB)** & worked out the sequence of a.a residues in the polypeptide chains of human insulin.
- **Identification of C & N terminal residues :-**
- It is also called “End Group Analysis ” .
- It reveals the identity of N terminal & C terminal residues.
- Also reveals how many ends are there , & so how many polypeptide chains there are.

Bovine insulin sequence

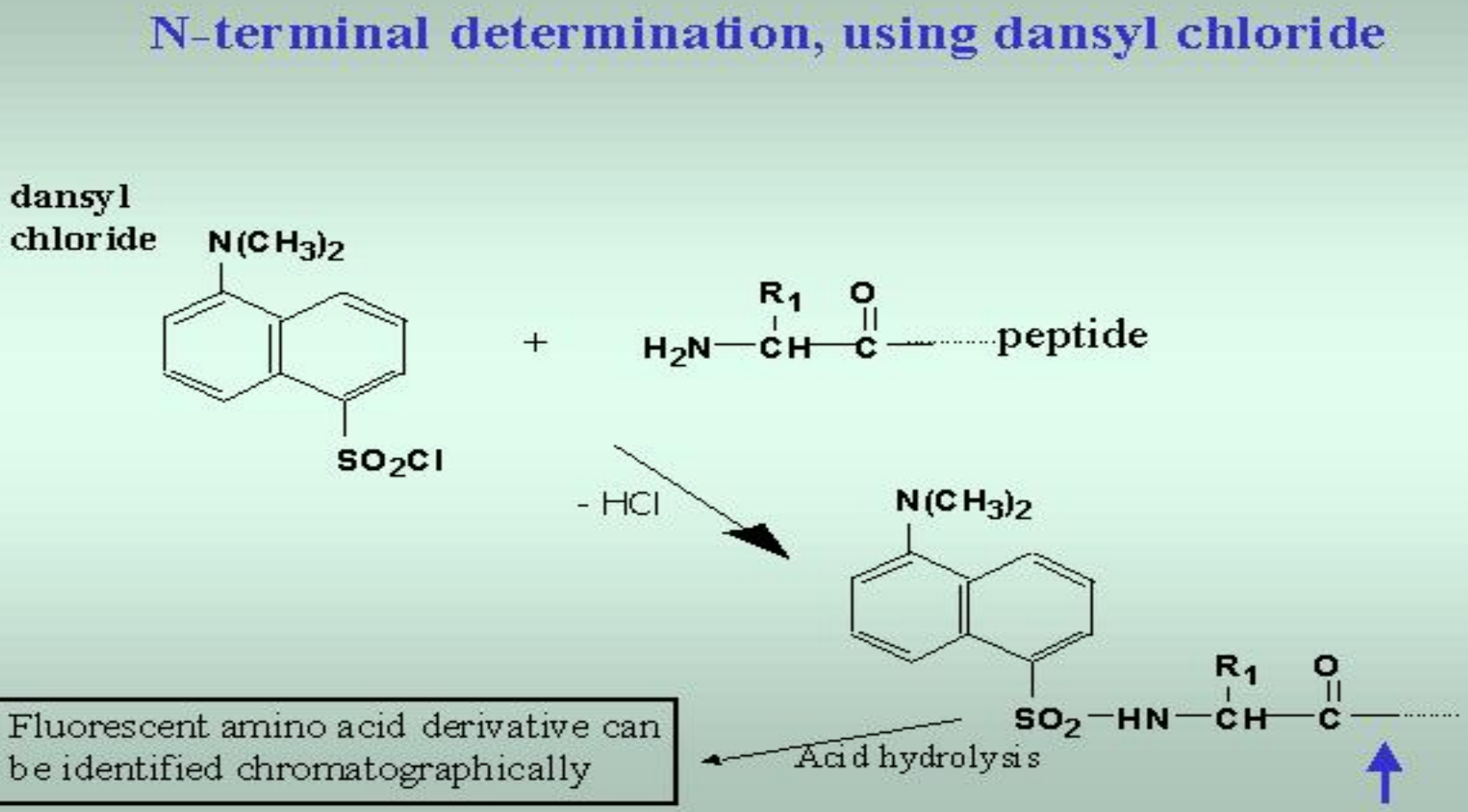
N-terminal analysis

- react with fluorodinitrobenzene (FDNB) or with dansyl chloride.
- identify hydrolyzed end group on paper chromatography or electrophoresis.
    - FDNB derivative yellow
    - dansyl derivative fluorescent (more sensitive)
- For identification of other amino acids, *Edman degradation* is used.

Reaction with FDNB



2. Reaction with Dansyl Chloride



3. Edman's Reagent

Edman degradation

- The most widely used method for determining the sequence of amino acids in a protein.
- **Amino acids are removed one residue at a time, sequentially from the N-terminal end.**
- The liberated amino acid derivative is then identified by chromatography
- The cycle can be repeated many times to obtain the sequence of a peptide.

Edman degradation

- Edman degradation uses Edman's reagent:
  - **Phenyl-N=C=S (phenylisothiocyanate, PTC)**
- PTC combines with the free N terminal amino acid.
- The N terminal amino acid is excised and separated as a **PTH derivative** (**phenylthiohydantoin** derivative).
- Amino acid sequencing can be done automatically in a gas phase sequenator;
  - HPLC separation of PTH derivatives
- Can determine the sequence of polypeptides of 30–60 residues.