Università degli Studi di Padova

Dipartimento Matematica

DATA SCIENCE
UNIVERSITY OF PADOVA

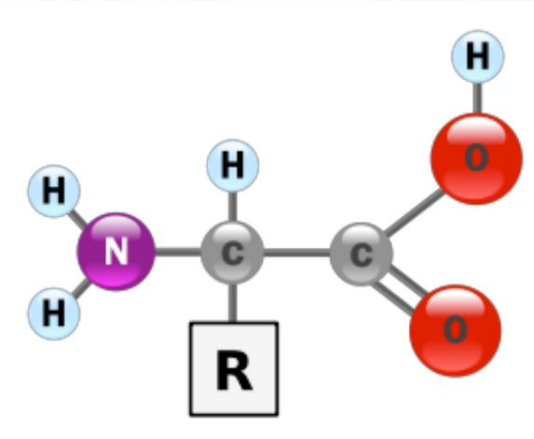
AMINO ACIDS

Master of Science in Data Science

Damiano Piovesan

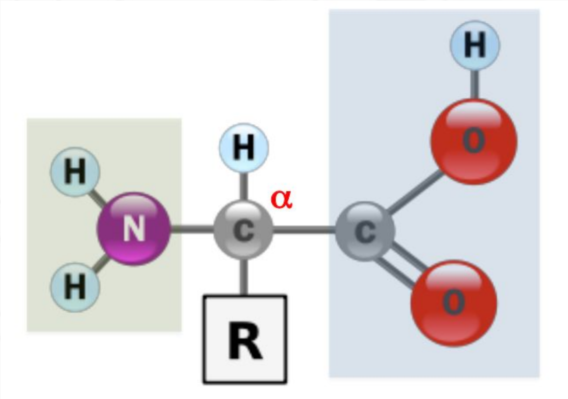
Amino acids

- Amino acids are **monomers** of proteins and other biological macromolecules
- Proteins are **polymers**
- Proteins are chains of amino acids
- They are **linear chains** of different combinations of 20 different amino acids
- Each amino acid has a specific chemical behaviour



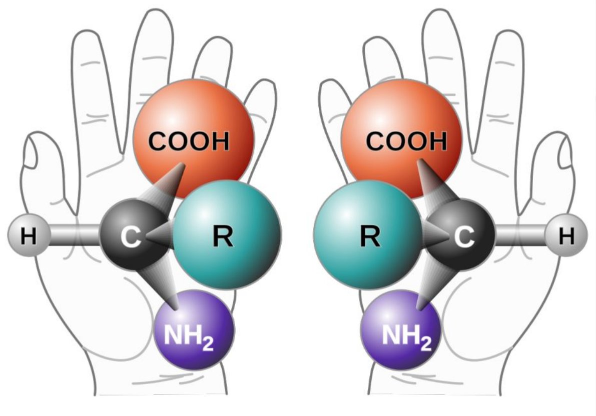
α-amino acids

- Only **20 out of 500** known amino acids appear in the genetic code
- All α-amino acids have a **carboxyl** (-COOH) and an **amminic** (-NH2) group bound to the **α carbon**
- They differ for the **side chain** (R)
- Different side chains have different three dimensional **structure** and **charge**
- **Structure**, **electric charge** and **hydrophobicity** are the principal features

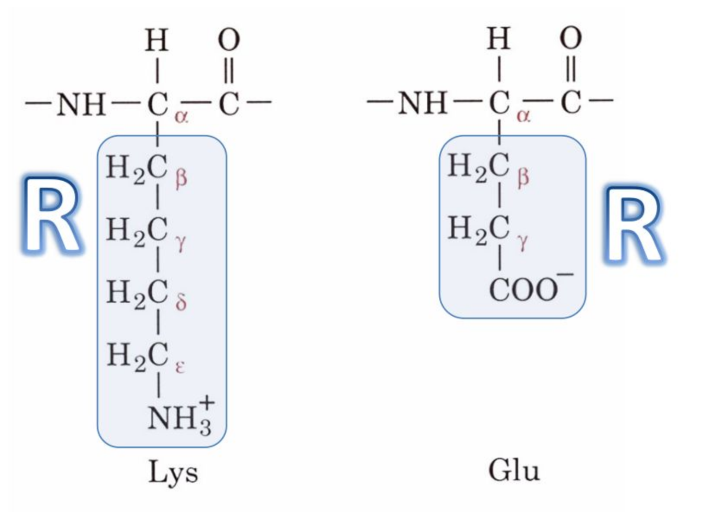


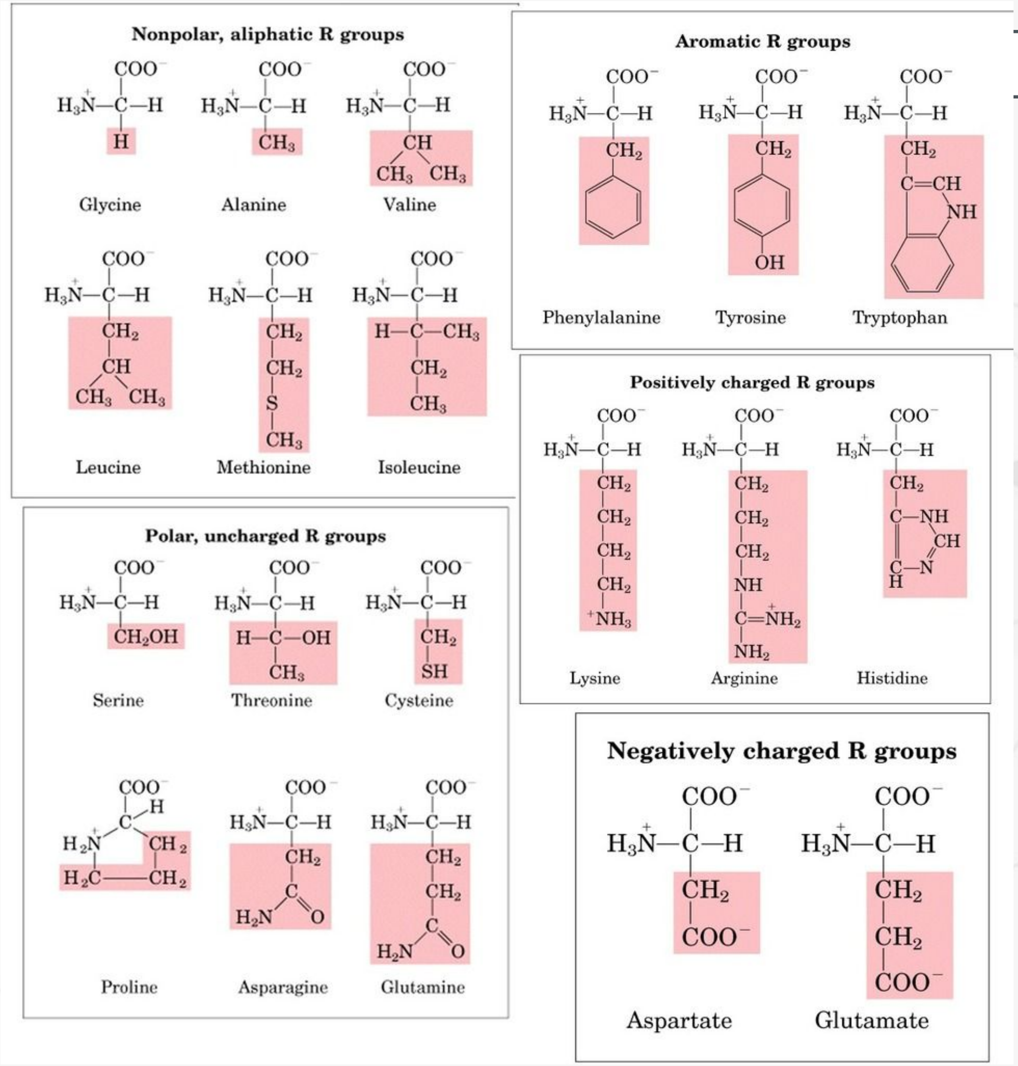
Chirality

- **α-carbon** is always bound to **four different groups** (except Glycine, R → H)
- **Chiral molecules → not superimposable** to mirrored version
- **Translations** and **rotations** not sufficient to superimpose the two versions
- Proteins contain only **L-amino acids**
- Protein **active sites** are asymmetric



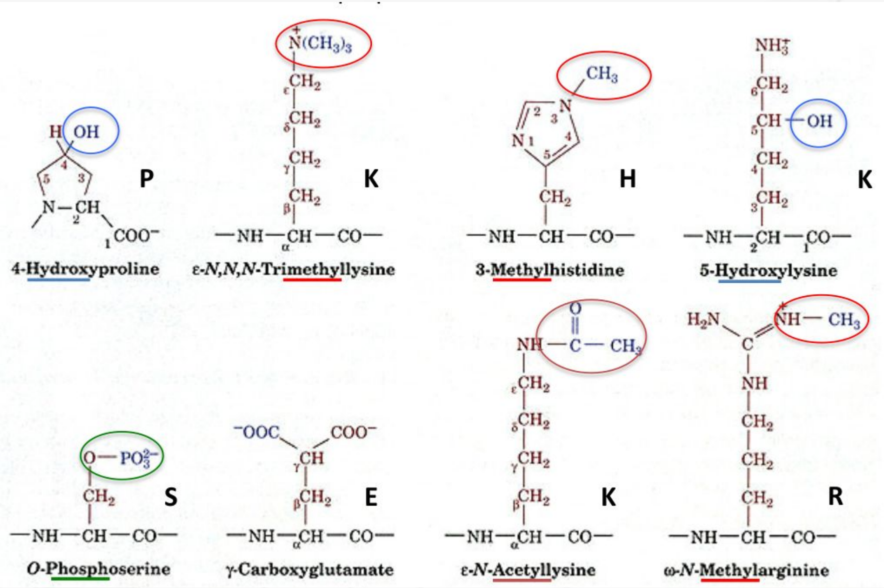
Side chain (R)




Nonpolar, aliphatic R groups
Glycine
Alanine
Valine
Leucine
Methionine
Isoleucine
Aromatic R groups
Phenylalanine
Tyrosine
Tryptophan
Positively charged R groups
Lysine
Arginine
Histidine
Polar, uncharged R groups
Serine
Threonine
Cysteine
Proline
Asparagine
Glutamine
Negatively charged R groups
Aspartate
Glutamate

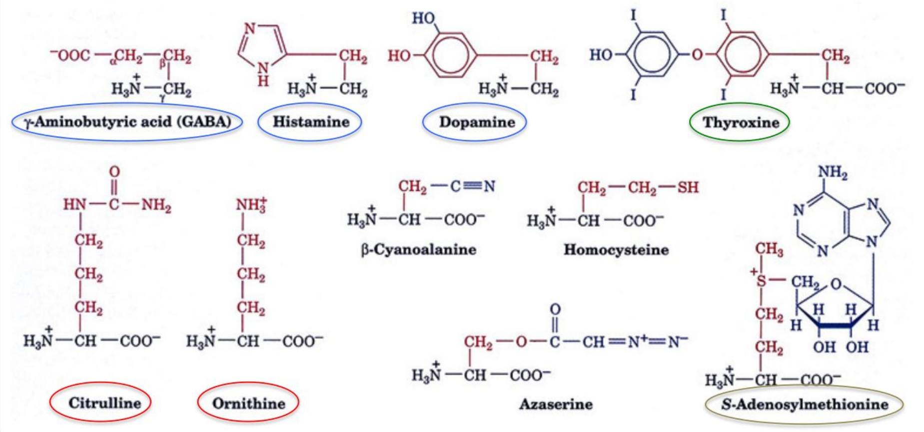
Post translational modifications

- Amino acids can be modified **after** their **synthesis**
- Chemical **groups** can be **added** to the **side chain** conferring new properties
- The same residue can undergo **multiple** modifications



Non protein amino acids

- Some amino acids does not appear in proteins but are produced by specific metabolic reactions
- They are fundamental components of some biological processes
    - Neurotransmitters, Hormones, Urea cycle, Metil group exchangers



UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO MATEMATICA

DATA SCIENCE
UNIVERSITY OF PADOVA

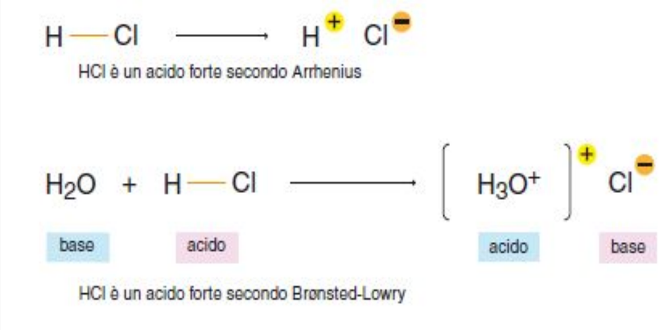
ACID-BASE THEORY

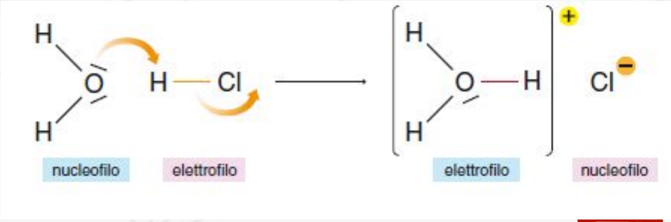
Master of Science in Data Science

Damiano Piovesan

Definitions

- **Arrhenius**
  - acid release H^+
  - base release OH^-
- **BrØnsted-Lowry**
  - acid release H^+
  - base accept H^+
- **Lewis**
  - acid accept electron pairs (electrophile)
  - base release electron pairs (nucleophile)





Acid

$$H_2O + HCl \rightarrow H_3O^+ + Cl^-$$

- React with water → covalent bonds form and break up
- Can be an ion +/-, or a neutral compound
- (BrØnsted-Lowry) Must have 1 H, but only H bound to very electronegative atoms are transferred

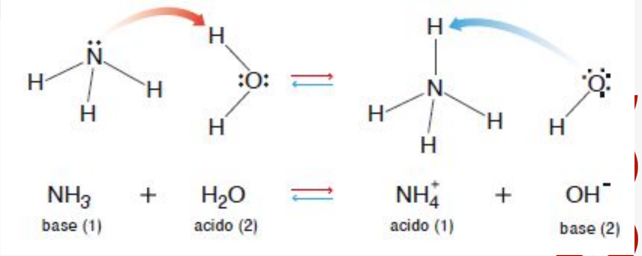
Base

- Only negative ions or neutral molecules
- Metal hydroxides → **Dissolve** in water (do not react), ions get solvatated

$$NaOH(s) \xrightarrow{H_2O} Na^+(aq) + OH^-(aq)$$

- Other bases → **React** with water

$$NH_3(aq) + H_2O(l) \rightleftharpoons NH_4^+(aq) + OH^-(aq)$$



Strong acids and bases

$$HCl(aq) + H_2O(l) \rightarrow Cl^-(aq) + H_3O^+(aq)$$

- React completely with water to form H_3O^+
- Fully **dissociate** or **ionize**
- The strength of an acid (or base) does not depend on its concentration

6 Strong Acids		6 Strong Bases	
$HClO_4$	perchloric acid	LiOH	lithium hydroxide
HCl	hydrochloric acid	NaOH	sodium hydroxide
HBr	hydrobromic acid	KOH	potassium hydroxide
HI	hydroiodic acid	$Ca(OH)_2$	calcium hydroxide
HNO_3	nitric acid	$Sr(OH)_2$	strontium hydroxide
H_2SO_4	sulfuric acid	$Ba(OH)_2$	barium hydroxide

Acid dissociation constant

$$HA + H_2O \rightleftharpoons A^- + H_3O^+$$

$$K = \frac{[A^-]\,[H_3O^+]}{[HA]\,[H_2O]}$$

$$\mathbf{K_a} = K\,[H_2O] = \frac{[A^-]\,[H_3O^+]}{[HA]}$$

$$\mathbf{pK_a = -log\ K_a}$$

- Water (solvent) contraction stays constant in comparison to the acid (solute)
- Water is removed from the denominator of K and included in K_a

Exercise. What is the concentration of pure water?

Weak acids

$$CH_3COOH + H_2O \rightleftharpoons CH_3COO^- + H_3O^+$$

$$\mathbf{K_a} = \frac{[H_3O^+]\,[CH_3COO^-]}{[CH_3COOH]} = \mathbf{1.8 \times 10^{-5}}$$

$$\mathbf{pK_a = ?}$$

Given pK_a how do you calculate the K_a?

$$\mathbf{K_a = 10^{-pKa}}$$

Weak acids

TABLE 9.4 K_A and pK_A Values for Selected Acids			
Name	**Formula**	K_a	pK_a
Hydrochloric acid	HCl	1.0×10^{7}	−7.00
Phosphoric acid	H_3PO_4	7.5×10^{-3}	2.12
Hydrofluoric acid	HF	6.6×10^{-4}	3.18
Lactic acid	$CH_3CH(OH)CO_2H$	1.4×10^{-4}	3.85
Acetic acid	CH_3CO_2H	1.8×10^{-5}	4.74
Carbonic acid	H_2CO_3	4.4×10^{-7}	6.36
Dihydrogenphosphate ion	$H_2PO_4^-$	6.2×10^{-8}	7.21
Ammonium ion	NH_4^+	5.6×10^{-10}	9.25
Hydrocyanic acid	HCN	4.9×10^{-10}	9.31
Hydrogencarbonate ion	HCO_3^-	5.6×10^{-11}	10.25
Methylammonium ion	$CH_3NH_3^+$	2.4×10^{-11}	10.62
Hydrogenphosphate ion	HPO_4^{2-}	4.2×10^{-13}	12.38

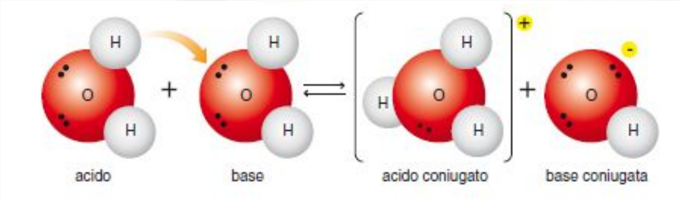
Self-ionization of water

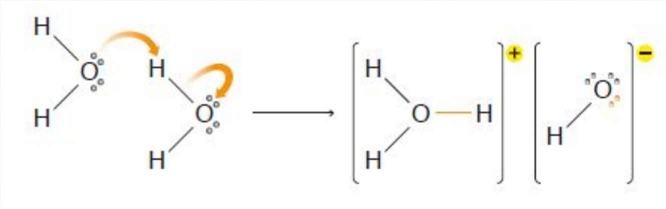
$$2H_2O \rightleftharpoons H_3O^+ + OH^-$$

$$K = \frac{[OH^-]\,[H_3O^+]}{[H_2O]\,[H_2O]}$$

$$K_a = K\,[H_2O] = \frac{[OH\text{-}^-]\,[H_3O^+]}{[H_2O]}$$

$$\mathbf{K_w} = K\,[H_2O]^2 = \mathbf{[OH^-]\,[H_3O^+]}$$





Ionic product of water (K_w)

$$\mathbf{K_w} = K\,[H_2O]^2 = \mathbf{[H_3O^+]\,[OH^-] = 10^{-14}\ M^2}$$

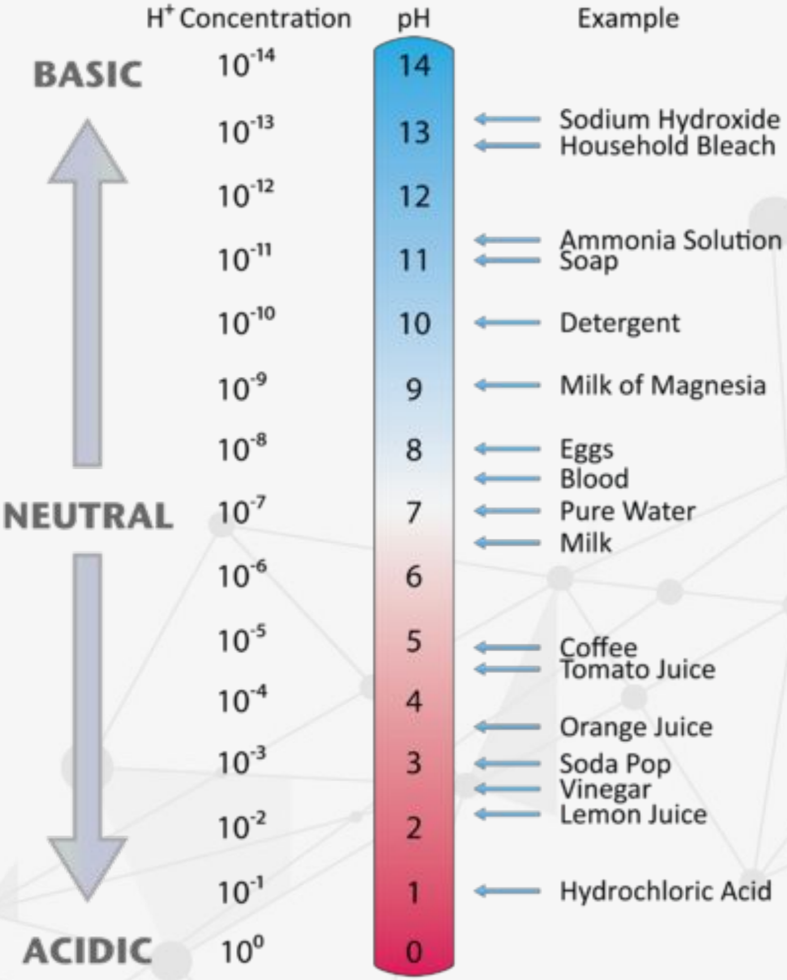
$$[H_3O^+] = [OH^-] = 10^{-7}\ M$$

$$\mathbf{pH} = -\log\,[H_3O^+] = \mathbf{7}$$

$$\mathbf{pK_w} = pH + pOH = \mathbf{14}$$

- K_w is valid for any aqueous solution, not only for water

pH


H+ Concentration
pH
Example
BASIC
NEUTRAL
ACIDIC
10-14
10-13
10-12
10-11
10-10
10-9
10-8
10-7
10-6
10-5
10-4
10-3
10-2
10-1
100
14
13
12
11
10
9
8
7
6
5
4
3
2
1
0
Sodium Hydroxide
Household Bleach
Ammonia Solution
Soap
Detergent
Milk of Magnesia
Eggs
Blood
Pure Water
Milk
Coffee
Tomato Juice
Orange Juice
Soda Pop
Vinegar
Lemon Juice
Hydrochloric Acid

Exercise

Calculate the concentration of hydroxide ions [OH] of an aqueous solution with a concentration of hydrogen ions of [H_3O^+] 10^{-5}.

Exercise

Calculate the concentration of hydroxide ions [OH] of an aqueous solution with a concentration of hydrogen ions of $[H_3O^+]$ 10^{-5}.

$$[OH^-] = K_w / [H_3O^+] = 10^{-14} / 10^{-5} = 10^{-9}$$

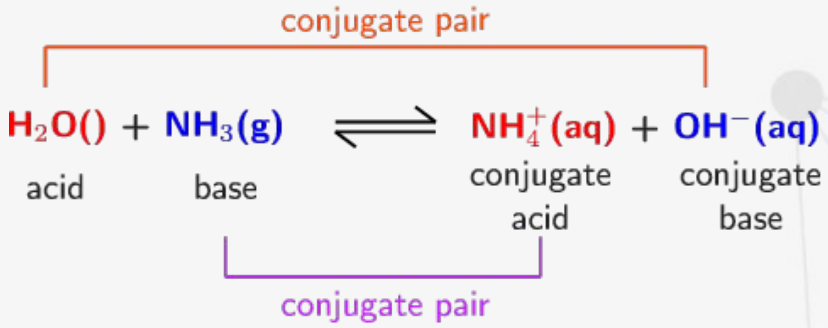
or

$$pH = -\log([H_3O^+]) = -\log(10^{-5}) = 5$$

$$pOH = 14 - 5 = 9$$

$$[OH^-] = 10^{-pOH} = 10^{-9}$$

Conjugate acids and bases (BrØnsted-Lowry)



- Molecules or ions pairs are correlated by gaining or losing a proton
- Stronger the acid, weaker the conjugate base
- Equilibrium is shifted toward the weakest acid (and base) since strong acids (and bases) react more
- It applies to reactions without water as well:

$$CH_3COOH + NH_3 \rightleftharpoons CH_3COO^- + NH_4^+$$

Conjugate acids and bases

Increasing acid strength

Increasing base strength

Acid			Base		
perchloric acid	$HClO_4$	Undergo complete acid ionization in water	Do not undergo base ionization in water	ClO_4^-	perchlorate ion
sulfuric acid	H_2SO_4			HSO_4^-	hydrogen sulfate ion
hydrogen iodide	HI			I^-	iodide ion
hydrogen bromide	HBr			Br^-	bromide ion
hydrogen chloride	HCl			Cl^-	chloride ion
nitric acid	HNO_3			NO_3^-	nitrate ion
hydronium ion	H_3O^+			H_2O	water
hydrogen sulfate ion	HSO_4^-			SO_4^{2-}	sulfate ion
phosphoric acid	H_3PO_4			$H_2PO_4^-$	dihydrogen phosphate ion
hydrogen fluoride	HF			F^-	fluoride ion
nitrous acid	HNO_2			NO_2^-	nitrite ion
acetic acid	CH_3CO_2H			$CH_3CO_2^-$	acetate ion
carbonic acid	H_2CO_3			HCO_3^-	hydrogen carbonate ion
hydrogen sulfide	H_2S			HS^-	hydrogen sulfide ion
ammonium ion	NH_4^+			HN_3	ammonia
hydrogen cyanide	HCN			CN^-	cyanide ion
hydrogen carbonate ion	HCO_3^-			CO_3^{2-}	carbonate ion
water	H_2O			OH^-	hydroxide ion
hydrogen sulfide ion	HS^-	Do not undergo acid ionization in water	Undergo complete base ionization in water	S^{2-}	sulfide ion
ethanol	C_2H_5OH			$C_2H_5O^-$	ethoxide ion
ammonia	NH_3			NH_2^-	amide ion
hydrogen	H_2			H^-	hydride ion
methane	CH_4			CH_3^-	methide ion

Special cases

- **Amphoteric compounds** (react both as an acid and as a base)

Acid-base conjugate pair

$$CH_3COOH + H_2O \rightleftharpoons CH_3COO^- + H_3O^+$$

Acid-base conjugate pair

- **Polyprotic acids**

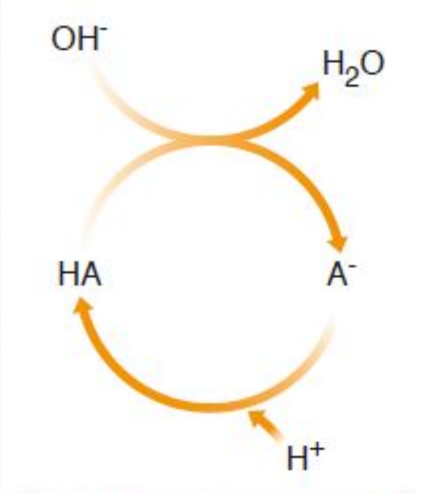
$$H_2CO_3 + H_2O \underset{}{\overset{Ka_1}{\rightleftharpoons}} HCO_3^- + H_3O^+ \qquad Ka_1 = \frac{[HCO_3^-][H_3O^+]}{[H_2CO_3]}$$

$$HCO_3^- + H_2O \underset{}{\overset{Ka_2}{\rightleftharpoons}} CO_3^{2-} + H_3O^+ \qquad Ka_2 = \frac{[CO_3^{2-}][H_3O^+]}{[HCO_3^-]}$$

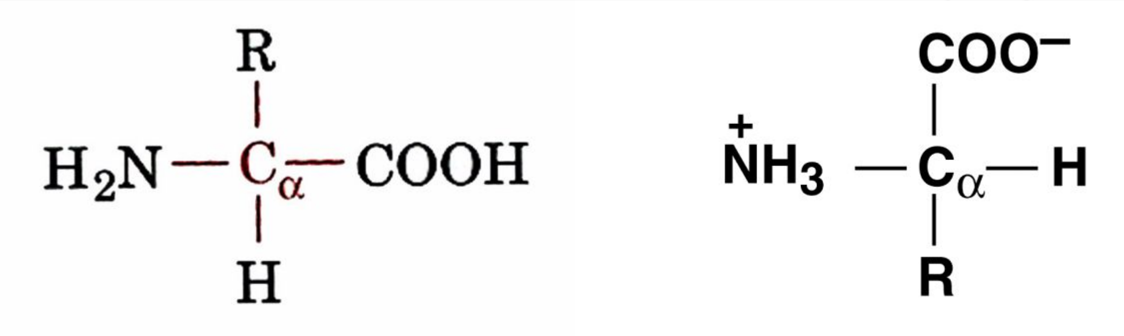
$\boxed{Ka_1 = 4{,}56 \times 10^{-7}}$ $\boxed{Ka_2 = 5{,}61 \times 10^{-11}}$

Buffer solutions

- pH changes very little when a small amount of strong acid or base is added to it (e.g. H_3O^+ o OH^-)
- When H^+ (or OH^-) increases, also the concentration of HA (or A-) rises but the pH does not change
- The generated HA is a weak acid and does not affect the pH, because it is weak and in low amount

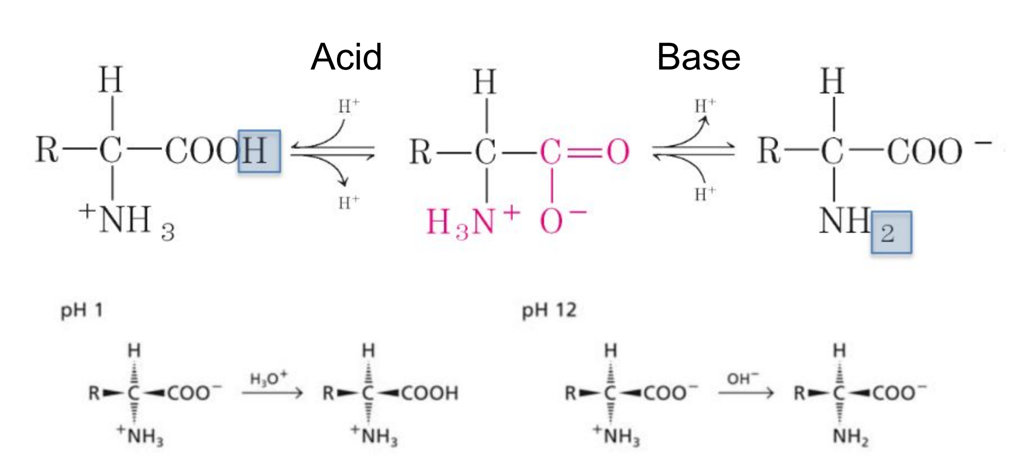


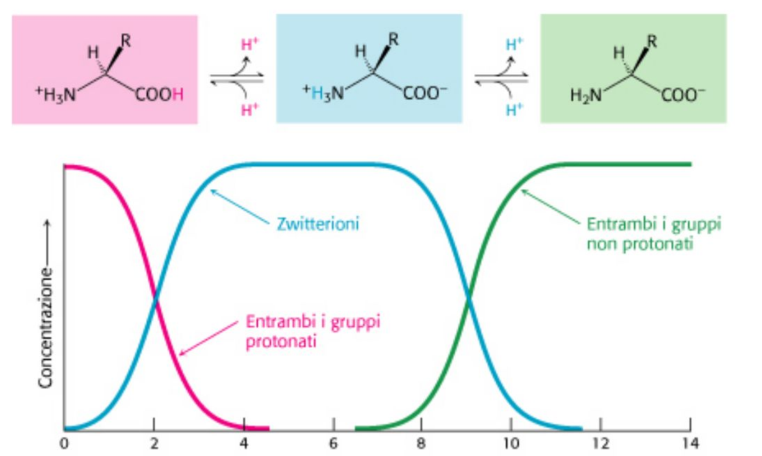
Amino acids as dipolar ions



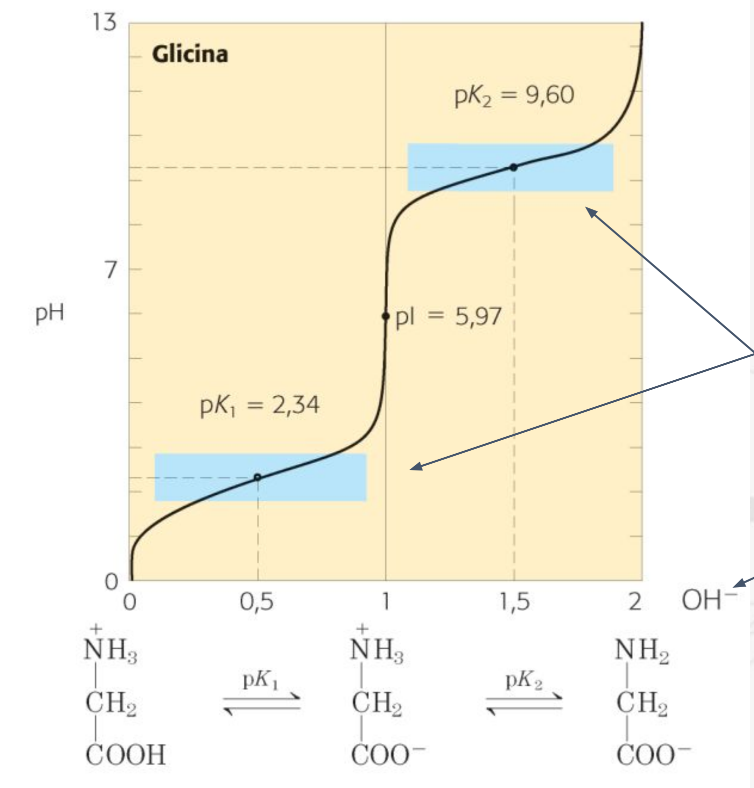
Amino acids as acid-base

- Amino acid at **physiological pH** (ca. 7.0) exist in the **Zwitterionic** form
- The **isoelectric point (pI)** is the **pH value** at which an amino acid is found as **dipolar ion** (**Zwitterionic form**) with null net charge





How to calculate isoelectric point



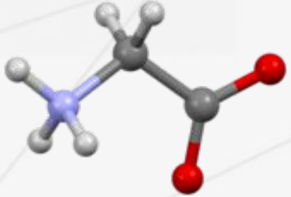
- Two reactions, two pKa

Buffer zones
(pH does not change)
pH = pKa

From a strong base
(eg. NaOH)

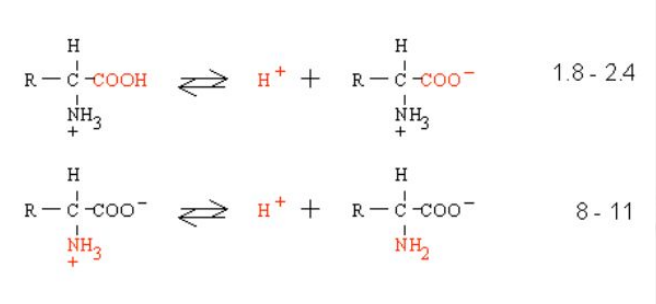
Zwitterion Glycine

$NH_3^+CH_2COO^-$



Side chains dissociation

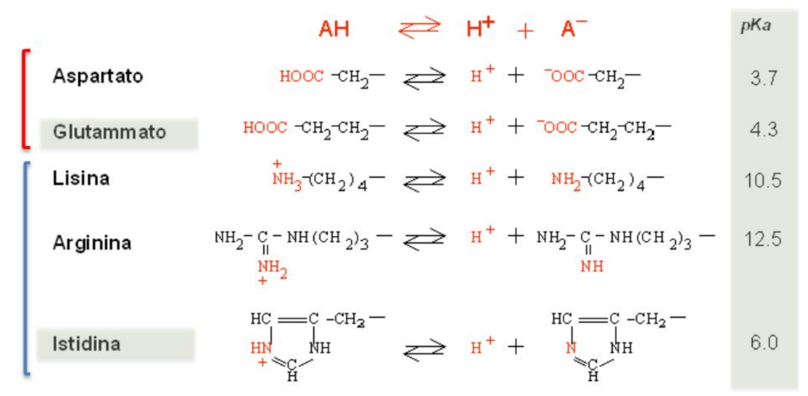
Main chain



Side chain

Acid

Base



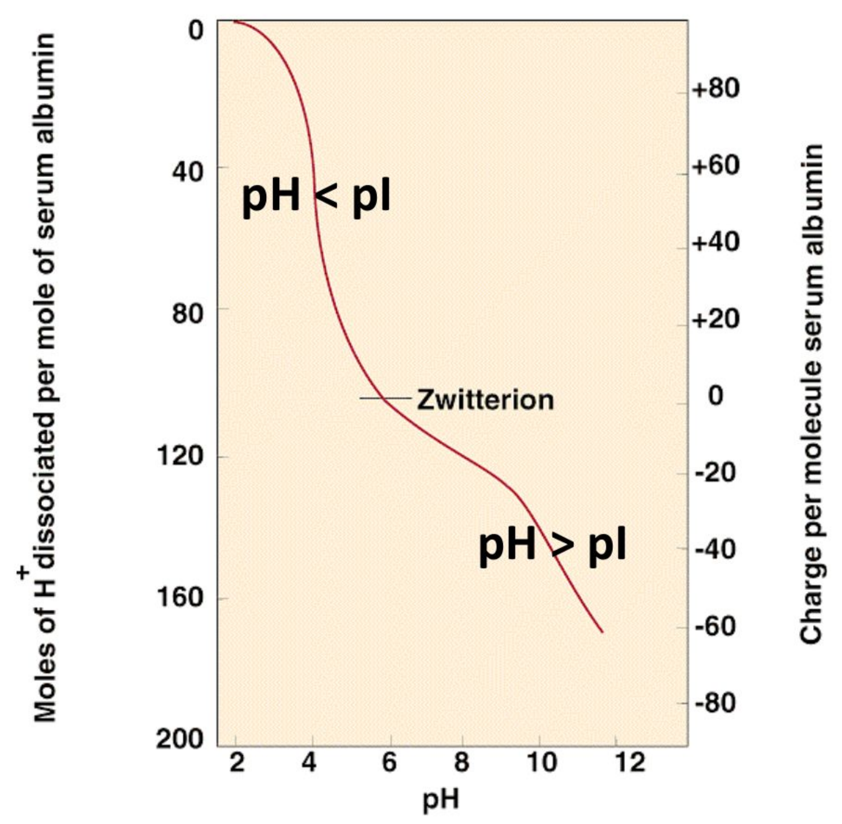
	Amino Acid	pKa Value		
	Name	Alpha Carboxy	+Alpha Amino	Side Chain
Non-Polar Amino Acids	Glycine	2.34	9.60	
	Alanine	2.34	9.69	
	Valine	2.32	9.62	
	Leucine	2.36	9.60	
	Isoleucine	2.36	9.68	
	Methionine	2.28	9.21	
	Phenylalanine	1.83	9.13	
	Tryptophan	2.38	9.39	
	Proline	1.99	10.60	
Polar Amino Acids	Serine	2.21	9.15	
	Threonine	2.63	9.10	
	Cysteine	1.71	10.78	8.33
	Tyrosine	2.2	9.11	10.07
	Asparagine	2.02	8.84	
	Glutamine	2.17	9.13	
Acidic Amino Acids	Aspartic Acid	2.09	9.82	3.86
	Glutamic Acid	2.19	9.67	4.25
Basic Amino acids	Lysine	2.18	8.95	10.79
	Arginine	2.17	9.04	12.48
	Histidine	1.82	9.17	6.04

©Leah4sci.com

Amino Acid Tutorials + Cheat Sheet Leah4sci.com/AminoAcids

Protein pI

- Proteins have specific pI that correspond to the pH when the protein has a null net charge
- pH < pI protein positively charged
- pH > pI protein negatively charged



SIDE CHAINS

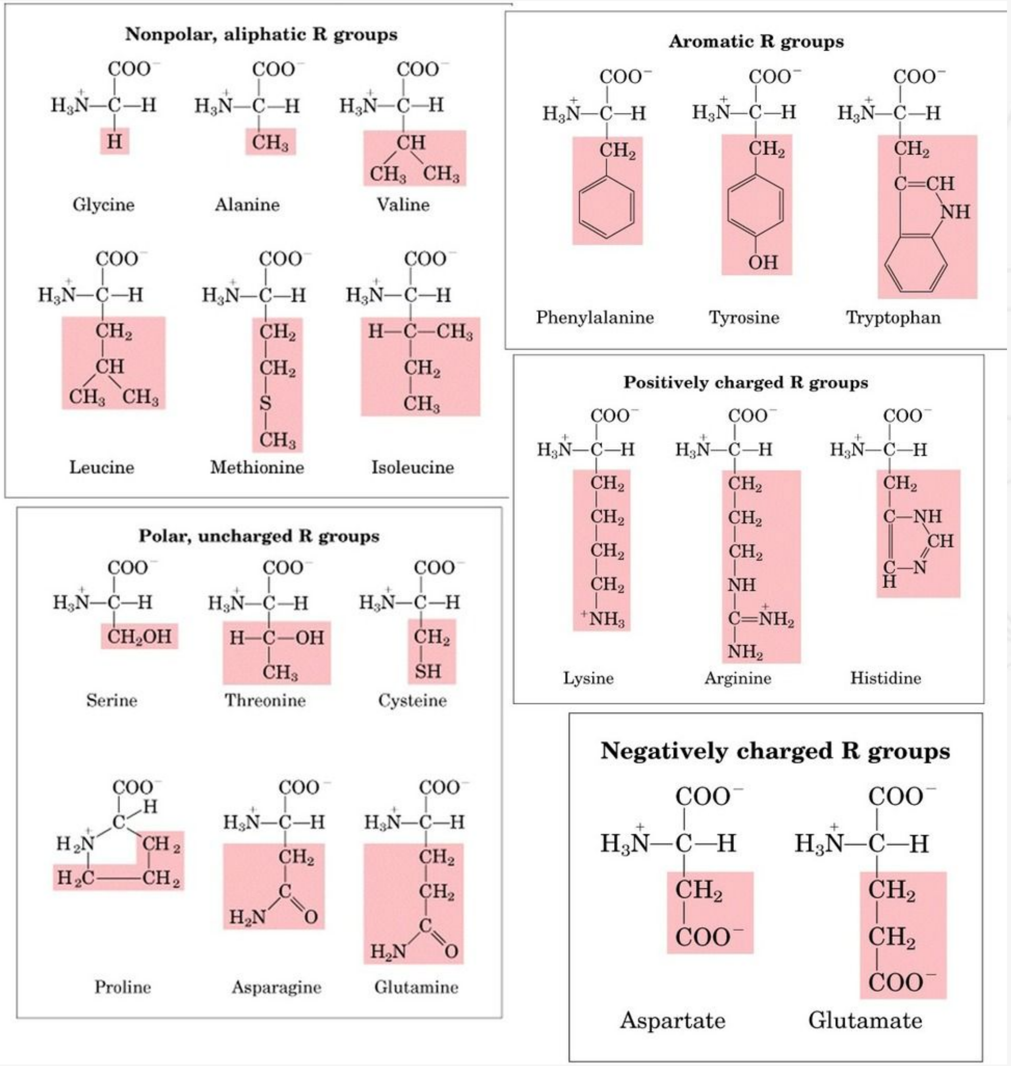
Master of Science in Data Science

Damiano Piovesan

Classification

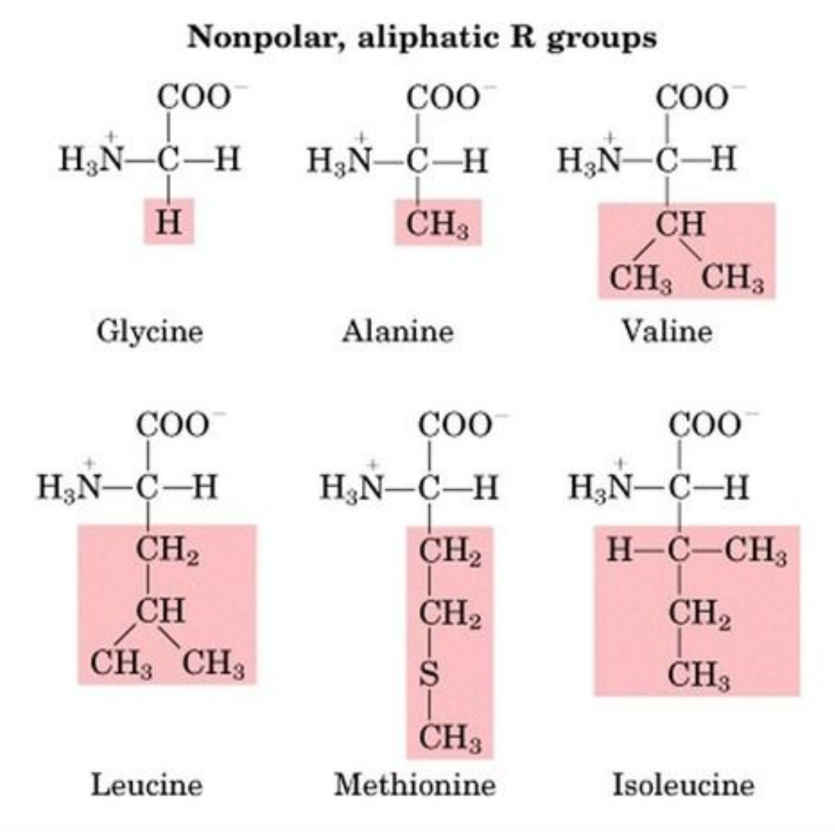
- pKa values
- Charges

At physiological pH → 7.4



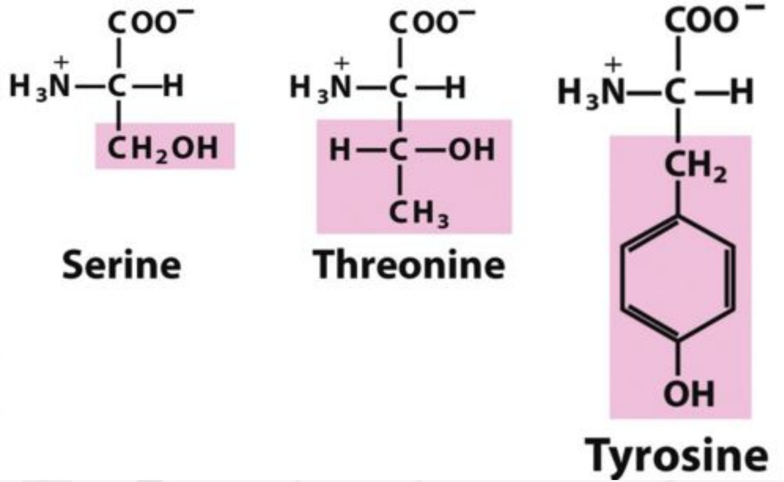
Aliphatic

- Compounds composed solely of **carbon** and **hydrogen**
- **Non polar**, **hydrophobic**
- Tend to stay within the **protein core** (except Glycine)



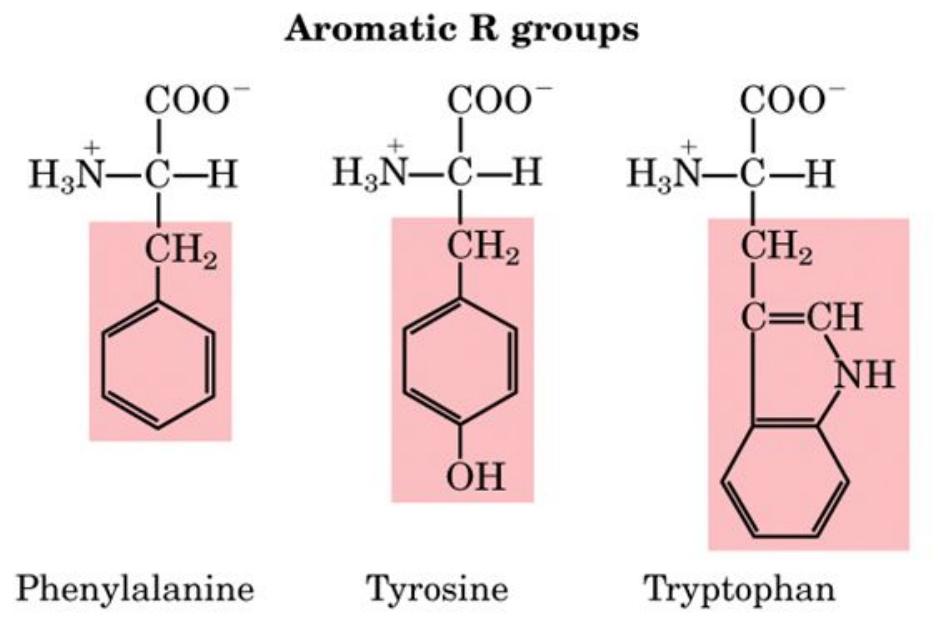
Hydroxyl

- **Polar**, **uncharged** and **hydrophilic**
- The phenolic hydroxyl ionizes with a pK_a of 10 and generally regarded as non ionizing
- **Serine** and **Threonine** can be post-translationally modified (**phosphorylated**)



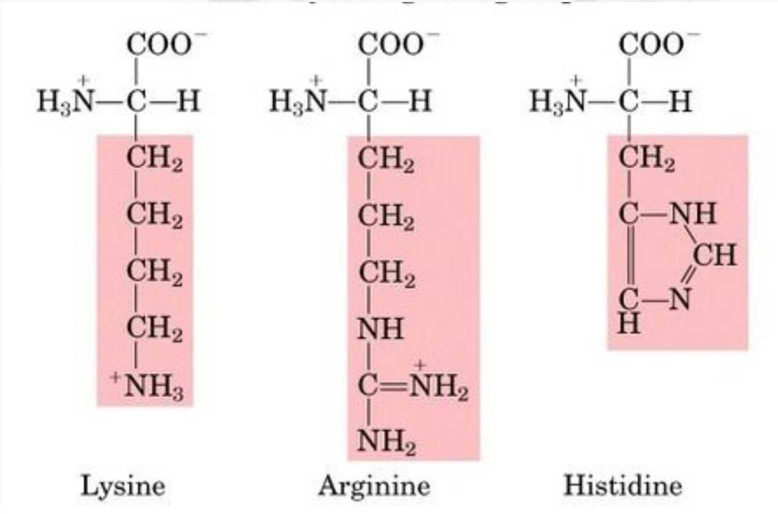
Aromatic

- Relatively **nonpolar**, **hydrophobic**
- **Tyrosine** can form **hydrogen bonds**



Basic

- **Polar** and **positively charged** (at pH < pKa)
- **Very hydrophilic**
- Almost always in contact with the solvent
- Often **histidine** participates in the **active site** of a protein as **proton donor or acceptor**



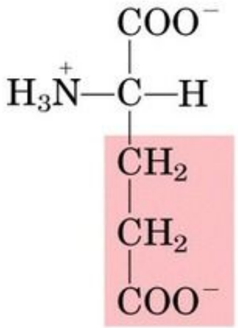
Acidic and their Amides

Acidic

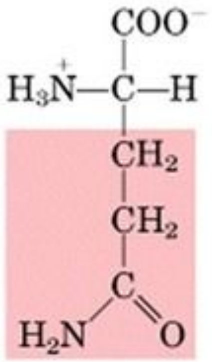
- **Polar** and **negatively charged**
- Have a **second carboxyl group**

Amides

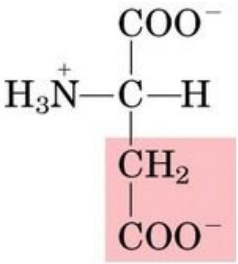
- **Polar** and **uncharged**, and **not ionizable**
- **Very hydrophilic**



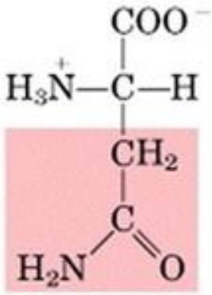
Glutamate



Glutamine



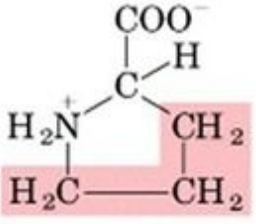
Aspartate



Asparagine

Cyclic

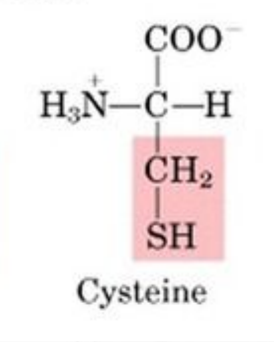
- Proline is the only **cyclic** amino acid
- **Nonpolar**
- Shares many properties with the **aliphatic** group
- Ambivalent amino acid, it can be **inside** or **outside** of a protein molecule
- Due to its unique structure, proline occurs in proteins frequently **in turns or bends**, which are often on the **surface**

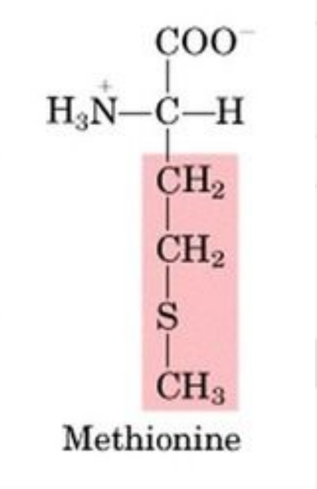


Proline

Sulfur-Containing

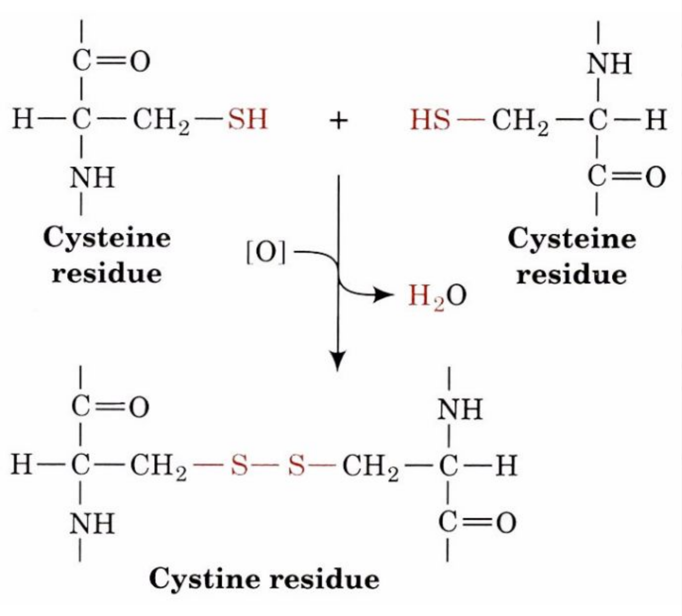
- **Nonpolar** and **hydrophobic**
- **Methionine** almost always found on the interior of proteins
- **Cysteine** ionize to yield the thiolate anion
- Sulfur has a **low propensity to hydrogen bond**, unlike oxygen. H_2S is a gas at room T, H_2O is a liquid
- The **thiol group** of cysteine can react with other thiol groups in an oxidation reaction that yields a **disulfide bond**



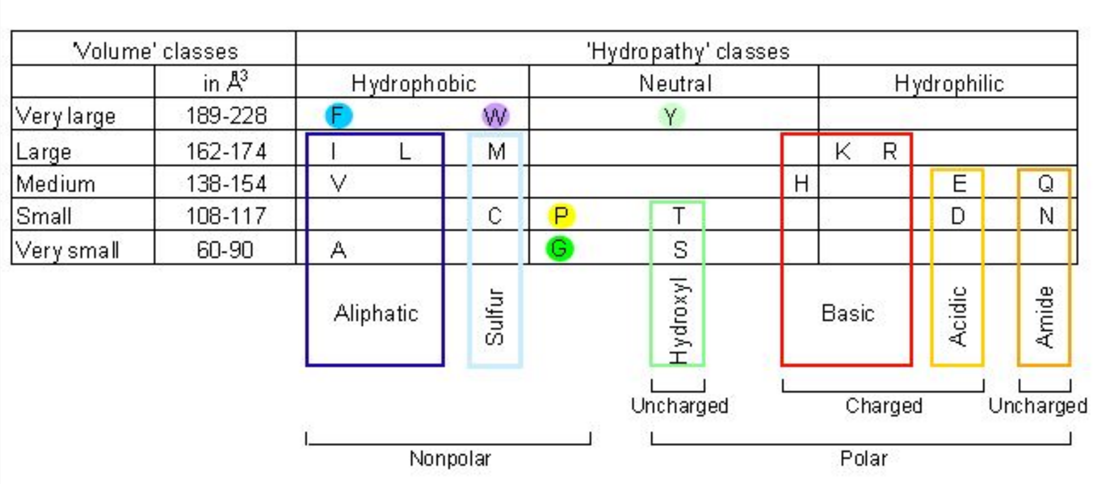


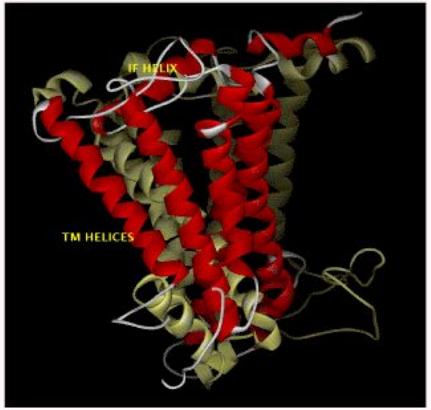
Disulphide bridge

- Two **cysteine** residues can bind by means of an **oxidation** reaction of **-SH** groups
- **S-S** disulfide bridge
- It can connects different regions of a protein, different chains, or stabilize its three dimensional conformation

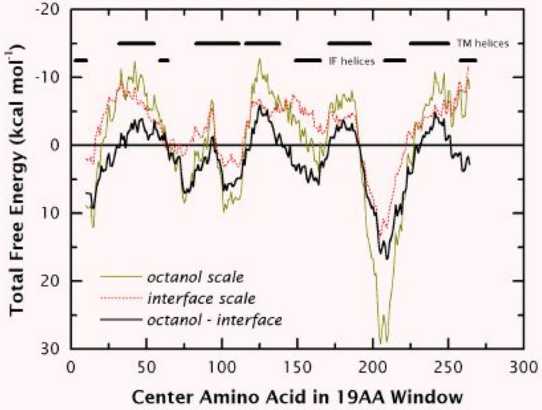


Hydrophobicity





PSRC L-Subunit: *R. sphaeroides*



Università degli Studi di Padova
Dipartimento Matematica
Data Science University of Padova

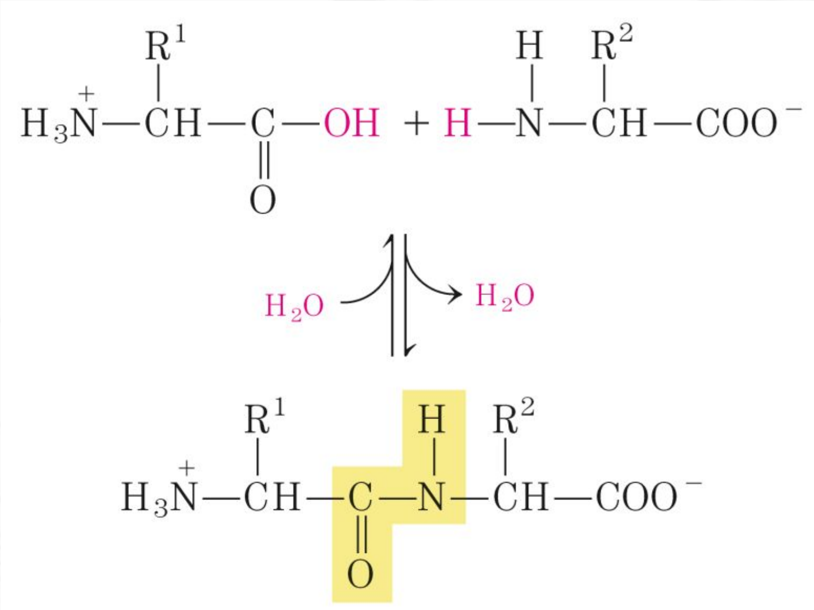
PROTEINS

Master of Science in Data Science

Damiano Piovesan

Proteins

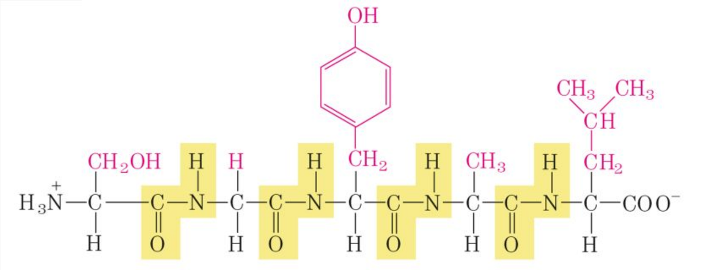
- Two amino acids can form covalent bonds (**peptide bond**) through the carboxyl -COOH and the aminic -NH3 groups
- Proteins are polymers of variable length
- They are made of 20 amino acids covalently bonded with a peptide bond



Peptides and proteins

DIPARTIMENTO MATEMATICA

- **Oligopeptides** are polymers with a low number of amino acids
- **Polypeptides** contains more amino acids (<10 kDa)
- **Proteins** are even larger (>10 kDa)
- 1 Dalton (Da) is the mass of a mole of C atoms
- **N-terminal** (left end)
- **C-terminal** (right end)
- Amino acids in the chain are called **residues**



Pentapeptide (Ser--Gly--Tyr--Ala--Leu o SGYAL)

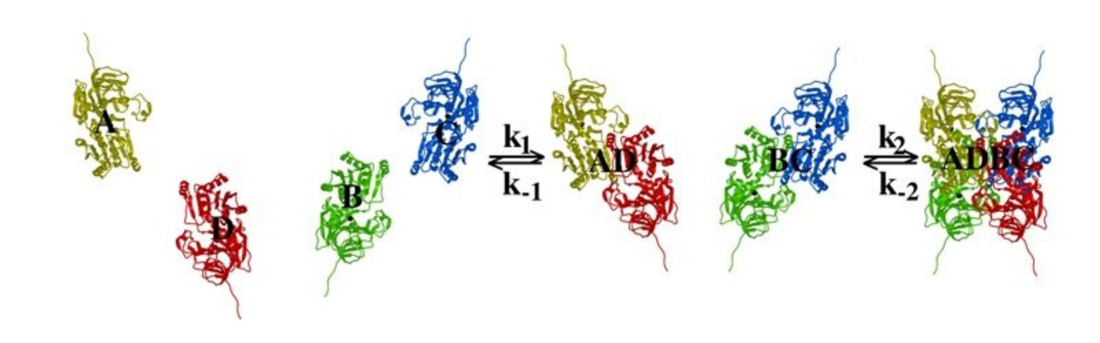
Protein size

- Biologically active peptides and proteins have different size and composition

Protein	**Residues**	**Weight (Da)**
Cytochrome C	104	12,400
Ribonuclease A	124	13,700
Albumin	609	66,000
Apolipoprotein B	4,536	513,000
Titin	26,926	2,293,000

Multi chain proteins

- Functional proteins can have single chains (**monomeric proteins**) or multiple different chains (**multimeric proteins**)
- Chains are not kept together by covalent bonds, but through **intermolecular (weak) interactions**



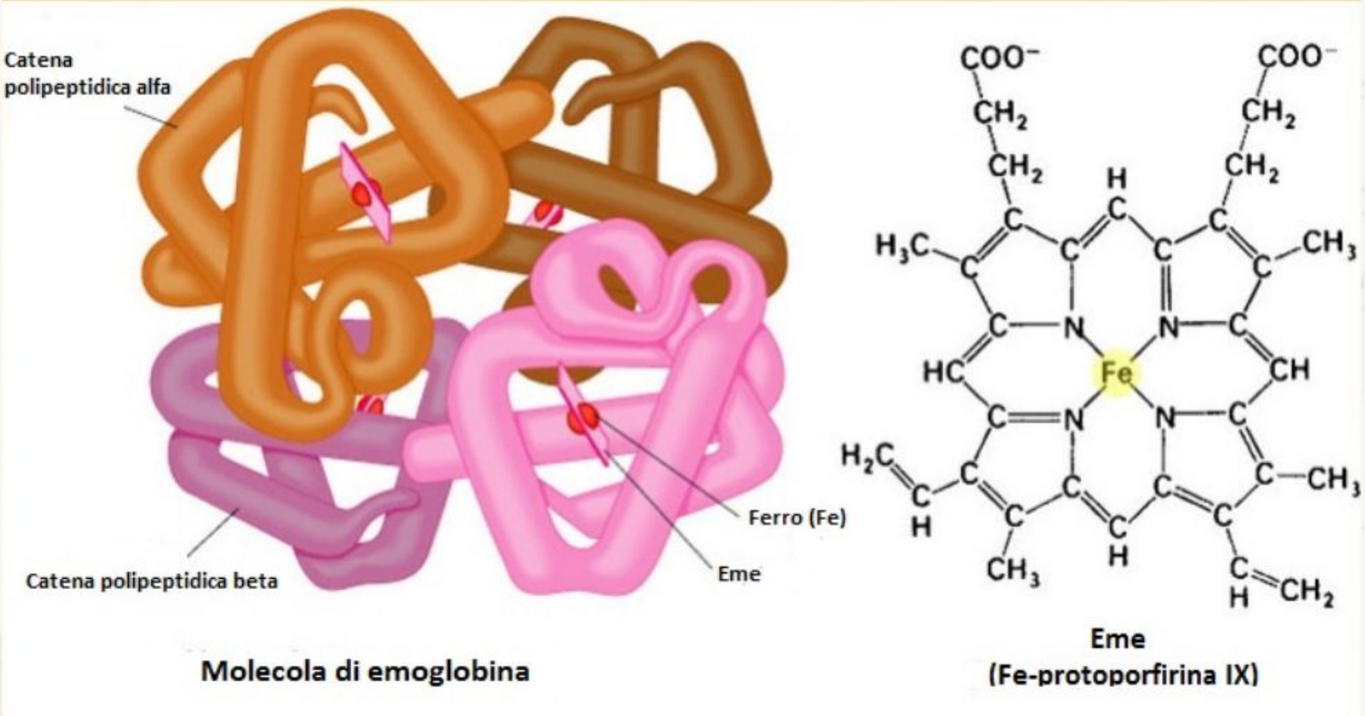
Conjugated proteins

DIPARTIMENTO MATEMATICA

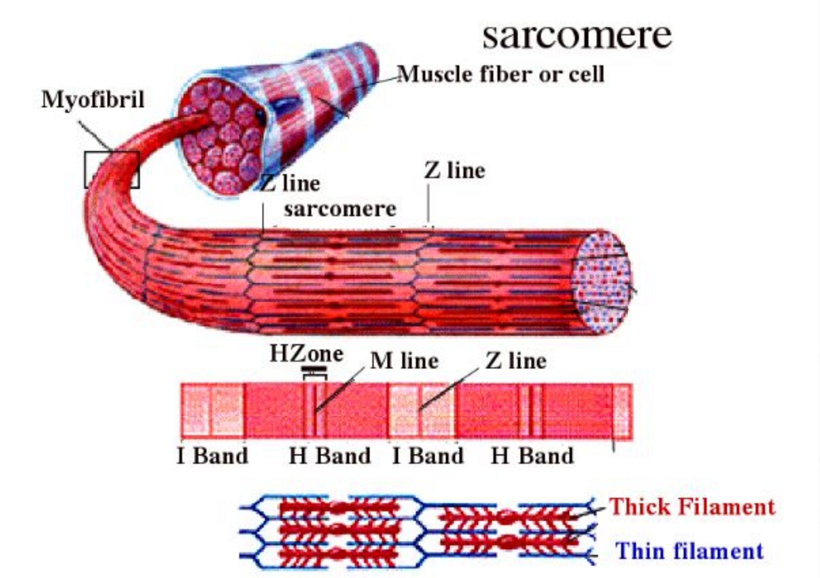
- Some proteins can have chemical groups (prosthetic group) different from amino acids

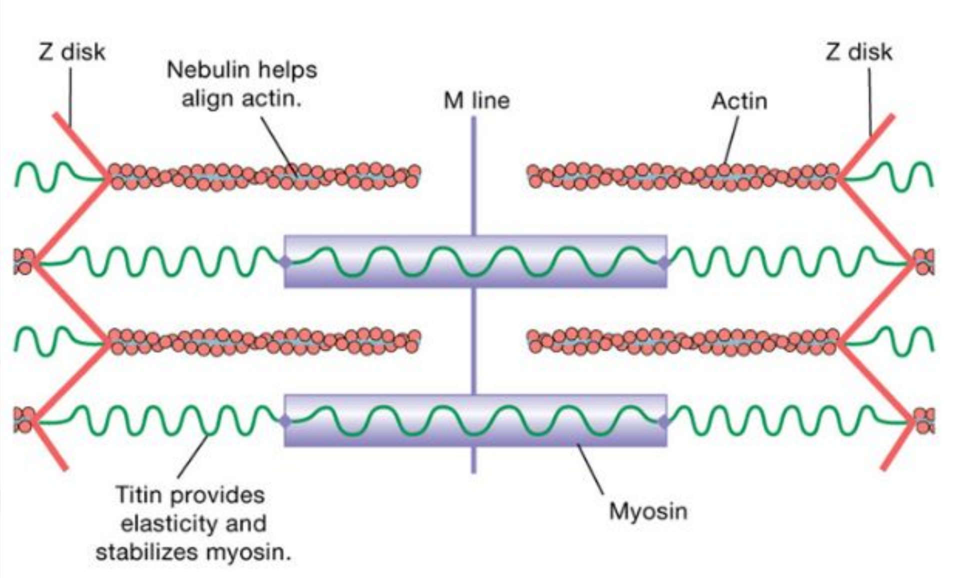
Class	**Prosthetic group**	
Lipoproteins	Lipids	Blood lipoprotein
Glycoproteins	Carbohydrates	Imunoglobulin
Phosphoproteins	Phosphate	Milk casein
Eme-proteins	Eme	Hemoglobine
Flavoproteins	Nucleotides	Succinate dehydrogenase
Metalloproteins	Iron, Zinc, Calcium, Cupper	Ferritine, Alcohol dehydrogenase, Calmodulin Plastocyanin

Hemoglobine



Titin


sarcomere
Muscle fiber or cell
Myofibril
Z line
Z line
sarcomere
HZone
M line
Z line
I Band
H Band
I Band
H Band
Thick Filament
Thin filament


Z disk
Nebulin helps
align actin.
M line
Actin
Z disk
Titin provides
elasticity and
stabilizes myosin.
Myosin