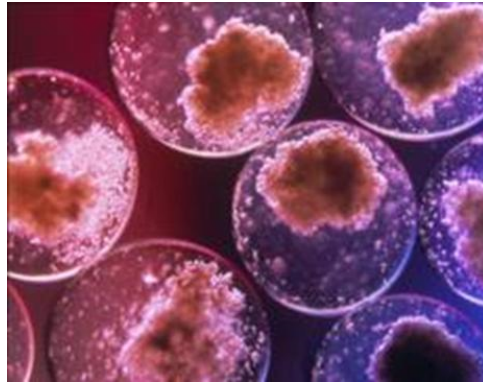
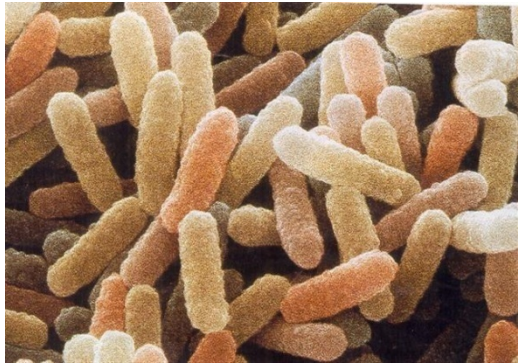




Chemistry and introduction to biochemistry

Macromolecules in living organisms



Aminoacids and sugars

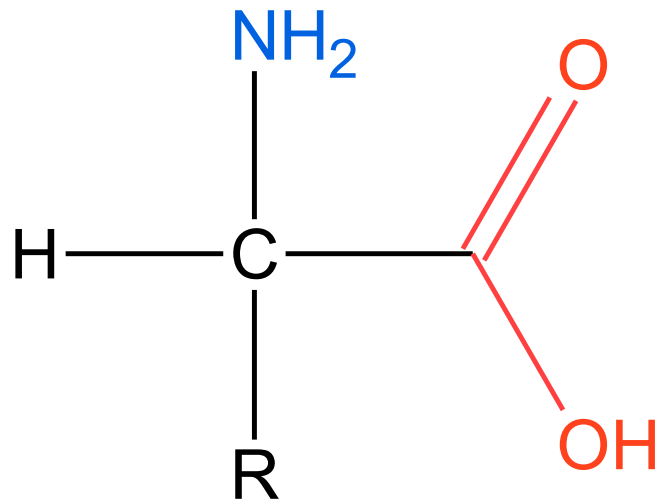
Sugars and polysaccharides

Lipids and membranes

Nucleotides and nucleic acids

Chemico-physical properties of aminoacids

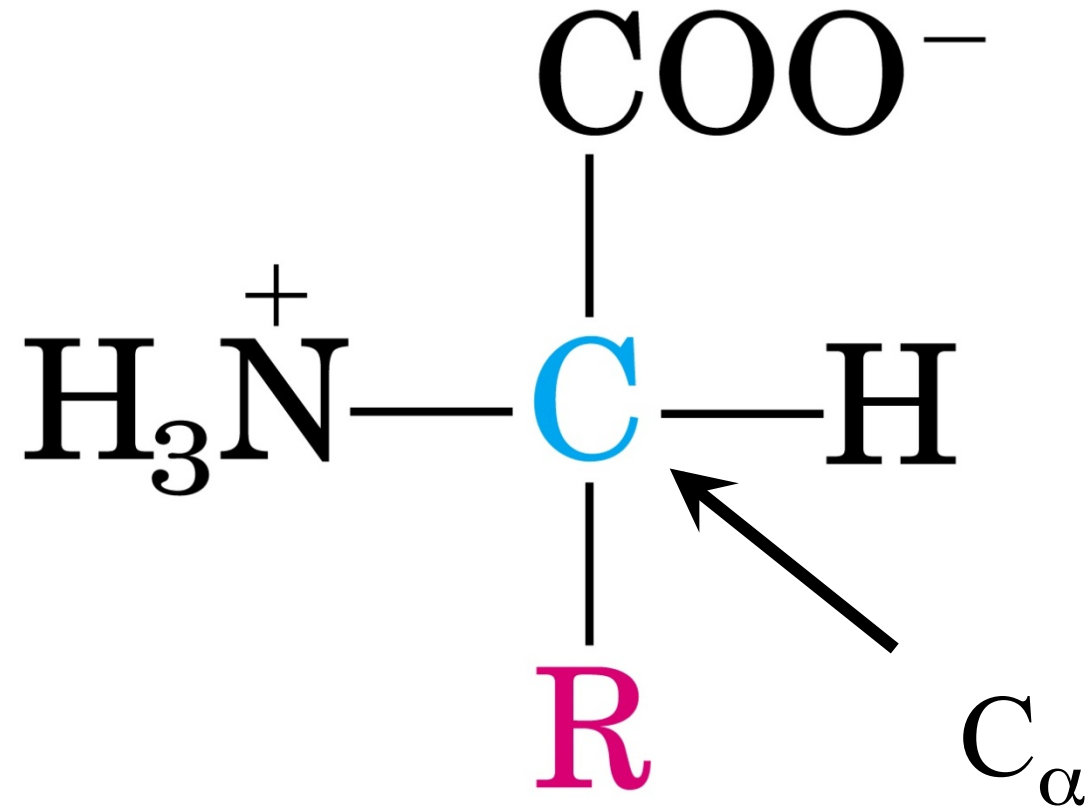
- Protein buiding bocks, also found in hormones and neutransmitters
- They have an **aminic** and a **carboxyl** group



Some of their properties are in contrast with the common behavior of these groups

- a) At difference from aliphatic amines (R-NH_2) and carboxylic acids (R-COOH), aminoacids are **non volatile crystalline solids with high melting points (200-300°C)**.
- b) **Insoluble in apolar solvents** (benzene, ether, etc) and are **soluble in water**
- c) Their water solutions behave as slolution of compounds with **high dipole moment**
- d) Their **acidic constants** (R-COOH) **are higher** with respect to carboxylic acids ($K_A \approx 10^{-2} \text{ M}$) and their **basic constants are lower** with respect to aliphatic amines ($K_B \approx 10^{-4} \text{ M}$)

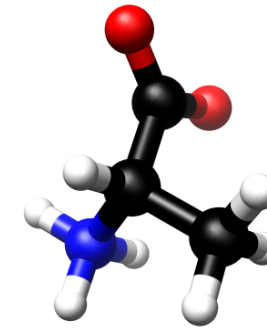
These properties indicate that in solution we have this structure:



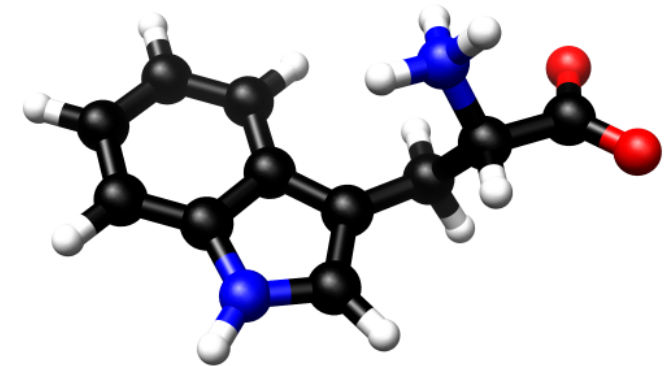
Amino acids are **dipolar ions** or **zwitterions**

NOMENCLATURE

ALANINE		ALA	A
CYSTEINE	*	CYS	C
ASPARTIC ACID		ASP	D
GLUTAMMIC ACIDO		GLU	E
PHENILALANINE	*	PHE	F
GLYCINA		GLY	G
HISTIDINE	*	HIS	H
ISOLEUCINE	*	ILE	I
LYSINA	*	LYS	K
LEUCINE	*	LEU	L
METHIONINE	*	MET	M
ASPARAGINE		ASN	N
PROLINE		PRO	P
GLUTAMINE		GLN	Q
ARGININE		ARG	R
SERINE		SER	S
THREONINE	*	THR	T
VALINE	*	VAL	V
THRIPTOPHANE	*	TRP	W
TYROSINE		TYR	Y



Ala



Trp

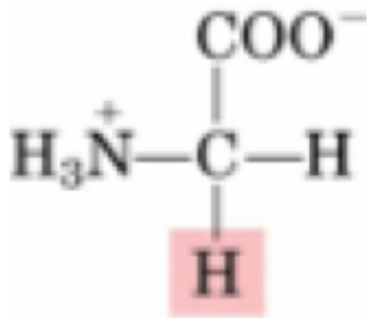
* essential, must be present in the diet.

Amino acids

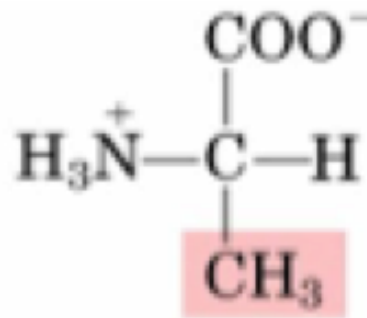
- ◆ Mammals need to assume with the diet 10 out of 20 aa, since they are not able to synthesize them:
- ◆ Cysteine, Phenylalanine, Histidine, Isoleucine, Lysine, Leucine, Methionine, Threonine, Valine and Tryptophan
- ◆ Depending on side chains characteristics, we can classify:
 - Apolar side chains (Gly, Ala, Val, Pro, Leu, Met, Ile)
 - Aromatic side chains (Phe, Tyr, Trp)
 - Uncharged polar side chains (Ser, Thr, Cys, Ans, Gln)
 - Positively charged side chains (His, Lys, Arg)
 - Negatively charged side chains (Glu, Asp)
- ◆ Their names can be shortened in a 3-letters code and in a 1-letter code

NB. Structures at pH 7

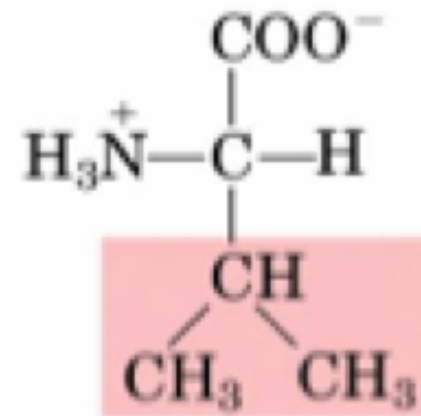
Apolar amino acids



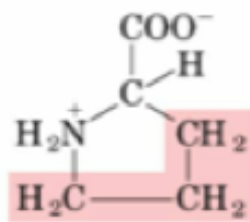
Glycine (Gly, G)



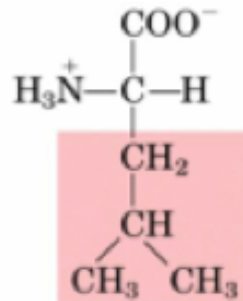
Alanine (Ala, A)



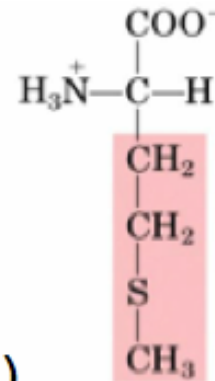
Valine (Val, V)



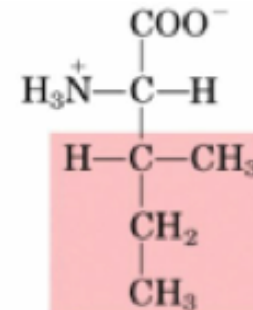
Proline (Pro, P)



Leucine (Leu, L)

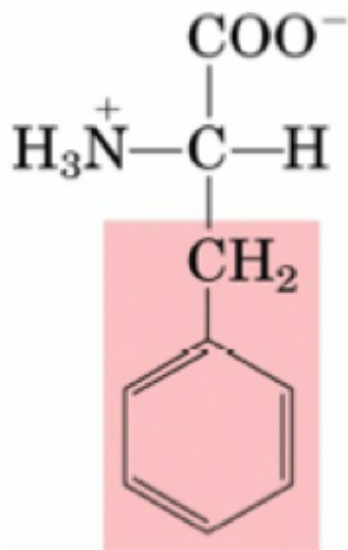


Methionine (Met, M)

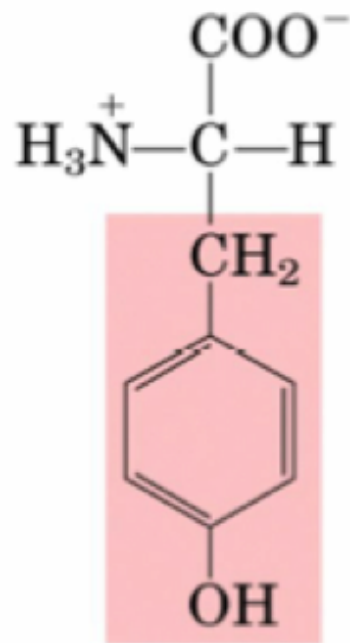


Isoleucine (Ile, I)

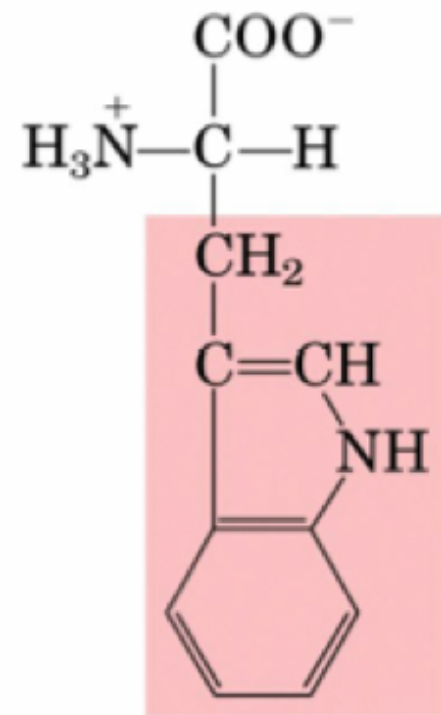
Aromatic amino acids



Phenylalanine (Phe, F)

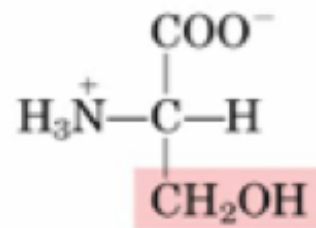


Tyrosine (Tyr, Y)

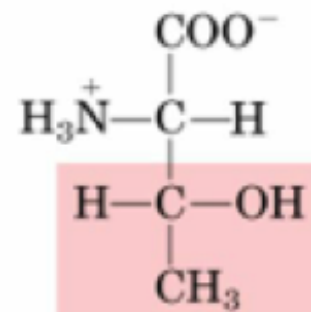


Tryptophan (Trp, W)

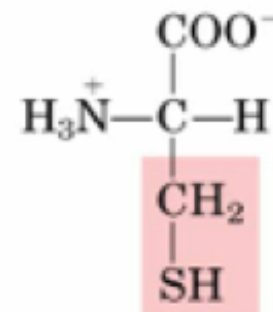
Polar amino acids



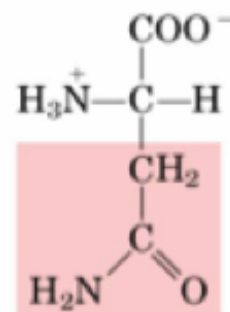
Serine (Ser, S)



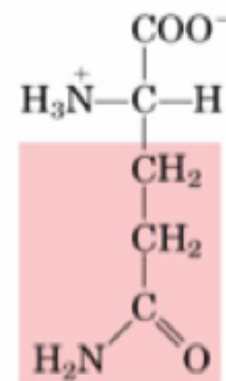
Threonine (Thr, T)



Cysteine (Cys, C)

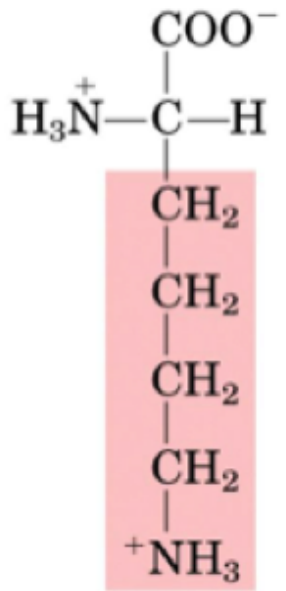


Asparagine (Asn, N)

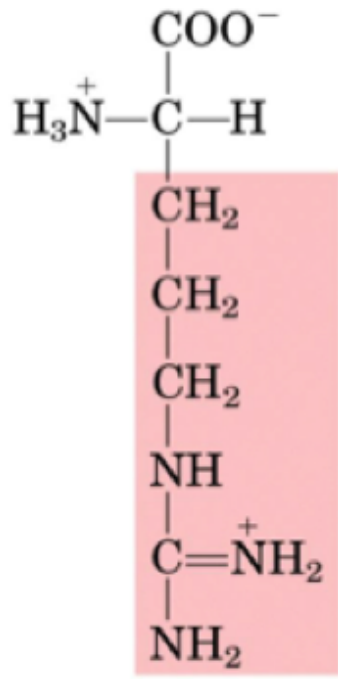


Glutamine (Gln, Q)

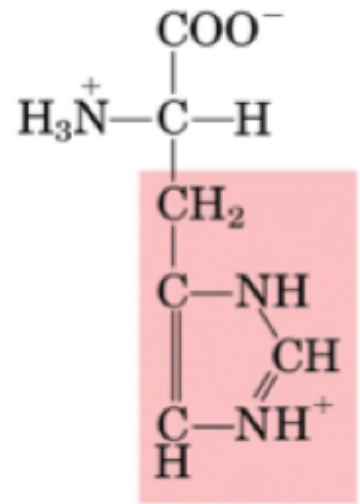
Basic amino acids



Lysine (Lys, K)

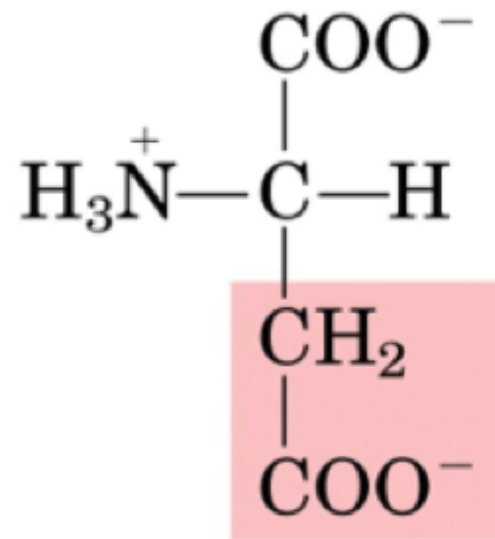


Arginine (Arg, R)

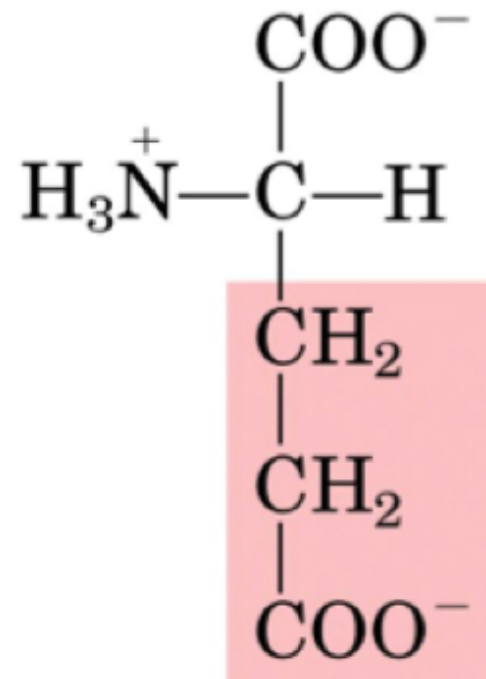


Histidine (His, H)

Acidic amino acids



Aspartic acid (Asp, D)

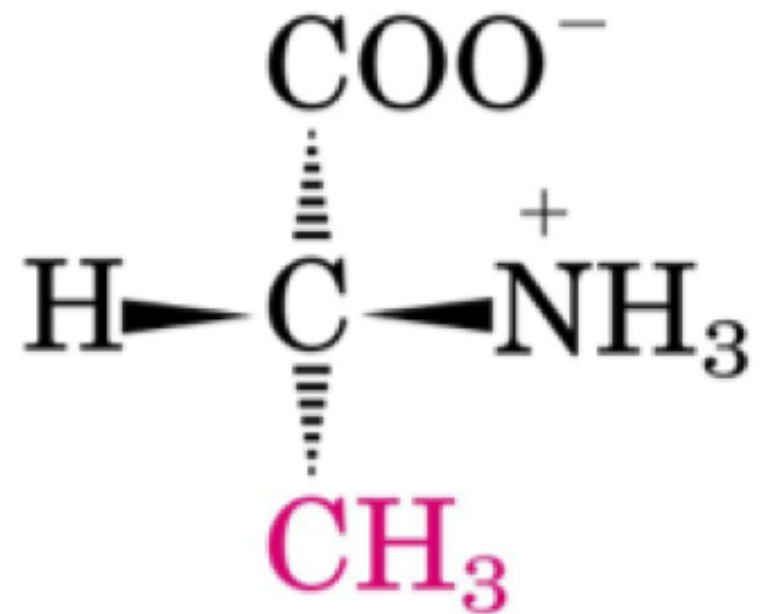


Glutamic acid (Glu, E)

All amino acids but Gly are chiral



L-Alanine



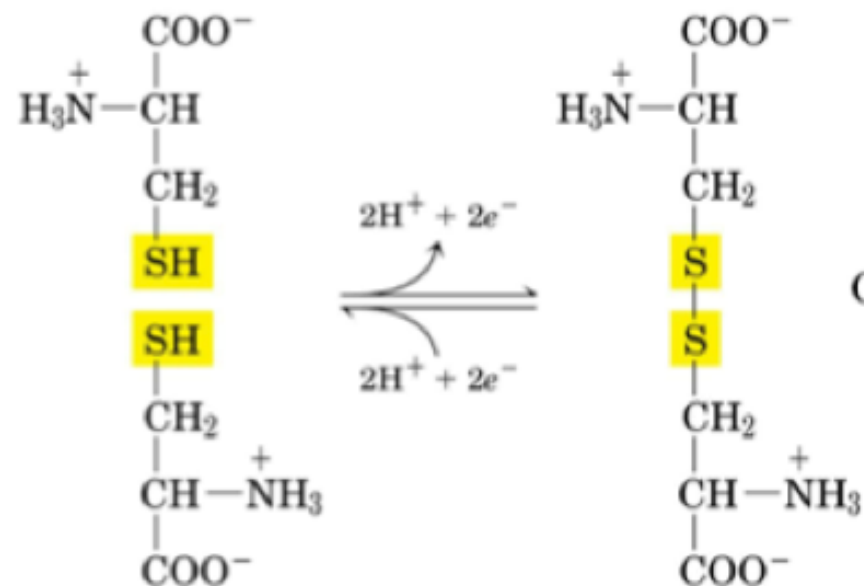
D-Alanine

In living organism there will only be L-amino acids incorporated into proteins

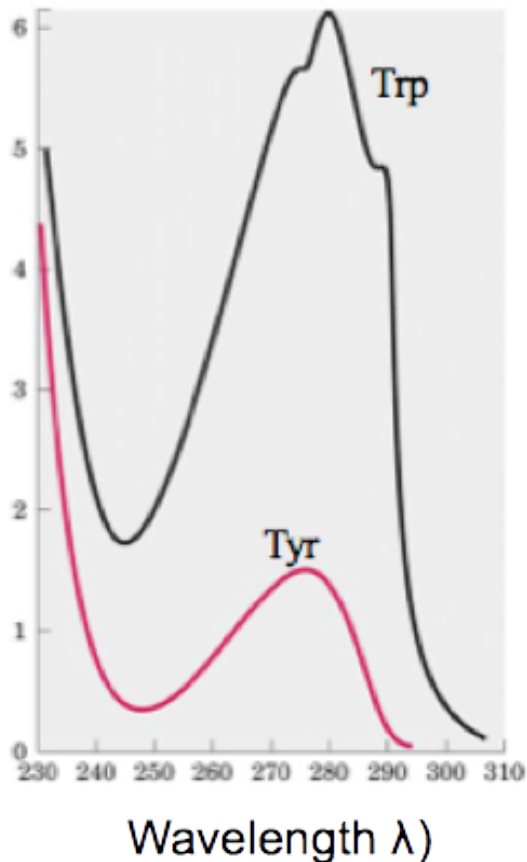
A few bacteria can synthesize D-aa, but they need a special enzyme

Properties of aa: redox reactions

- ◆ Cysteine has a reactive -SH group which can shuttle electrons and protons to acceptor proteins
- ◆ It can also react with a second Cys to form a covalent bond: the disulfide bond
- ◆ Disulfide bonds usually occurs in extracellular proteins, given that the extracellular environment is highly oxidative



Properties of aa: spectroscopic behaviour



- ◆ Aromatic side chains can absorb UV light and can be used to calculate protein quantity in solution
- ◆ The total absorbance of a sample is directly proportional to the number of Trp and/or Tyr and hence to the concentration of a given protein

Acid-base equilibria

	pK _{COOH}	pK _{NH3+}	pK _R	pI
GLY	2.3	9.6	-	6.0
ALA	2.3	9.7	-	6.0
VAL	2.3	9.6	-	6.0
PRO	2.0	10.6	-	6.3
LEU	2.4	9.7	-	6.0
MET	2.3	9.2	-	5.8
ILE	2.4	9.7	-	6.1
PHE	1.8	9.1	-	5.5
TYR	2.2	9.1	10.1	5.7
TRP	2.4	9.4	-	5.9
SER	2.2	9.2	-	5.7
THR	2.6	10.4	-	6.5
CYS	1.8	10.8	8.3	5.0
ASN	2.0	8.8	-	5.4
GLN	2.2	9.1	-	5.7
LYS	2.2	9.0	10.5	9.8
ARG	2.2	9.0	12.5	10.8
HIS	1.8	9.2	6.0	7.6
ASP	2.1	9.8	3.9	3.0
GLU	2.2	9.7	4.3	3.2

ACID-BASE PROPERTIES

pKa ranges from 1.8 and 2.4

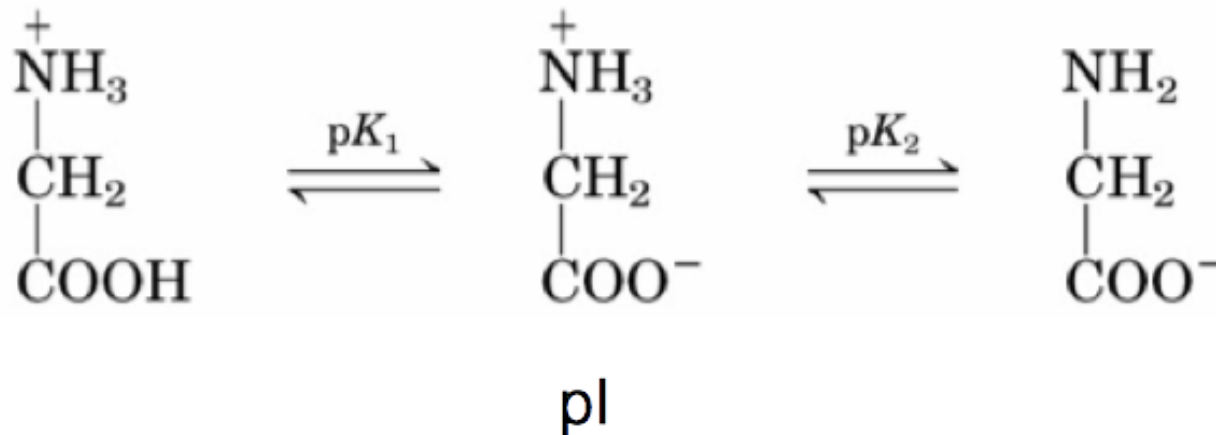
pKb ranges from 9.0 and 10.8

pK_R refers to the properties of the charged side chains

pI is the isoelectric point: the pH at which the aa has reached neutrality (no net charge)

Properties of the aa: acid-base reactions

- ◆ The ionizability of an aa is pH dependent:
- ◆ At $\text{pH} > \text{pI}$ → the aa has a net charge of -1 (anion)
- ◆ At $\text{pH} < \text{pI}$ → the aa has a net charge of +1 (cation)
- ◆ At $\text{pH} = \text{pI}$ → there is the zwitterion, with a net charge of 0



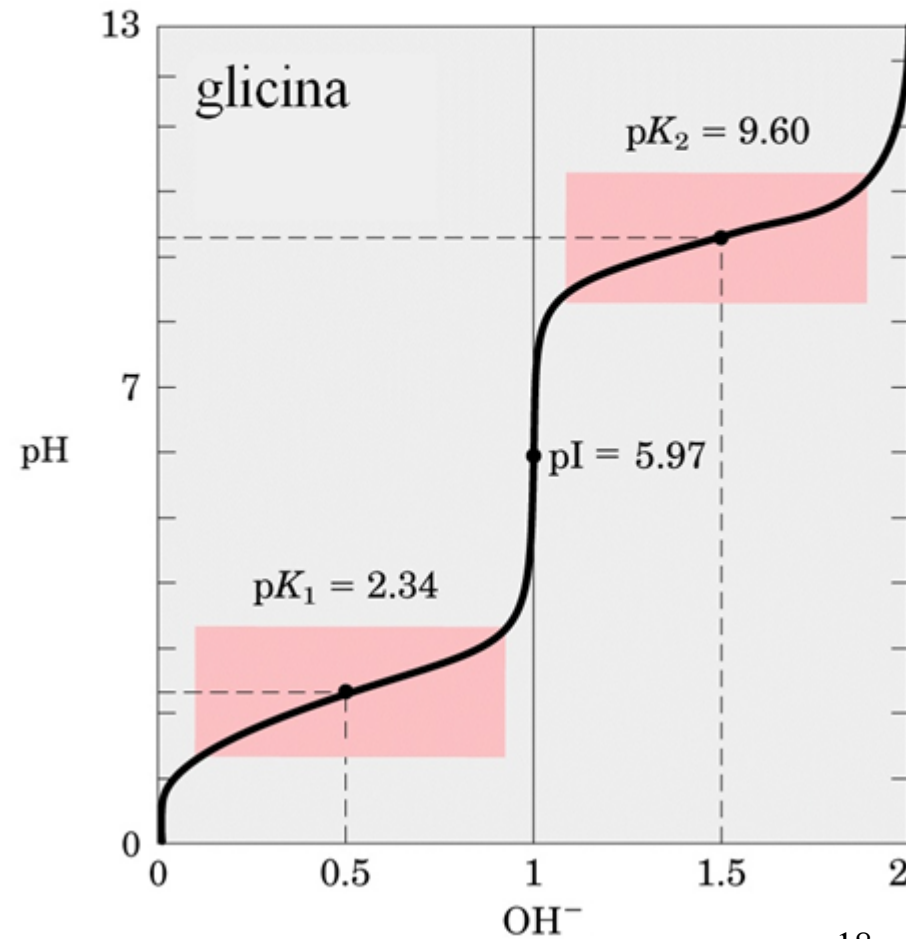
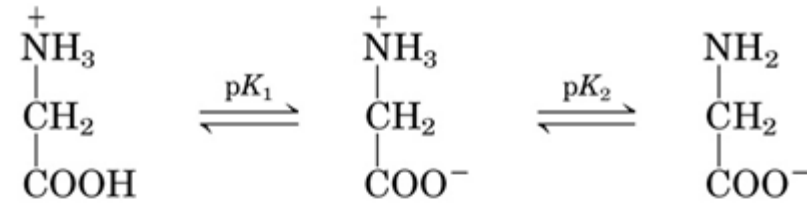
Effect of substituents on pK

ACID	FORMULA	pK
acetic	$\text{CH}_3\text{-COOH}$	4.75
chloroacetic	$\text{ClCH}_2\text{-COOH}$	2.87
dichloroacetic	$\text{Cl}_2\text{CH-COOH}$	1.48
aminoacetic (glycine)	$^+\text{H}_3\text{N-CH}_2\text{-COOH}$	2.35

H. Gutfreund, Enzyme physical properties, Wiley Interscience, 1972, London

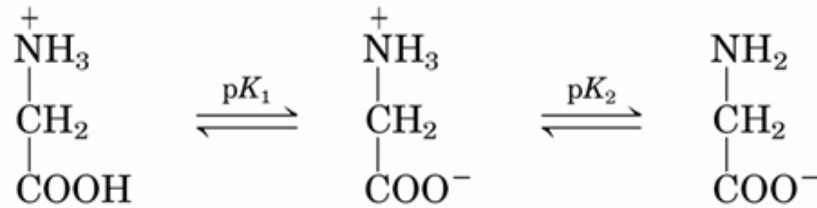
	pK _{COOH}	pK _{NH3+}	pK _R	pI
GLY	2.3	9.6	-	6.0
ALA	2.3	9.7	-	6.0
VAL	2.3	9.6	-	6.0
PRO	2.0	10.6	-	6.3
LEU	2.4	9.7	-	6.0
MET	2.3	9.2	-	5.8
ILE	2.4	9.7	-	6.1
PHE	1.8	9.1	-	5.5
TYR	2.2	9.1	10.1	5.7
TRP	2.4	9.4	-	5.9
SER	2.2	9.2	-	5.7
THR	2.6	10.4	-	6.5
CYS	1.8	10.8	8.3	5.0
ASN	2.0	8.8	-	5.4
GLN	2.2	9.1	-	5.7
LYS	2.2	9.0	10.5	9.8
ARG	2.2	9.0	12.5	10.8
HIS	1.8	9.2	6.0	7.6
ASP	2.1	9.8	3.9	3.0
GLU	2.2	9.7	4.3	3.2

Il pH o punto isoelettrico



The isoelectric point (or pI)

It is the **pH with the highest concentration of the zwitterionic species..**

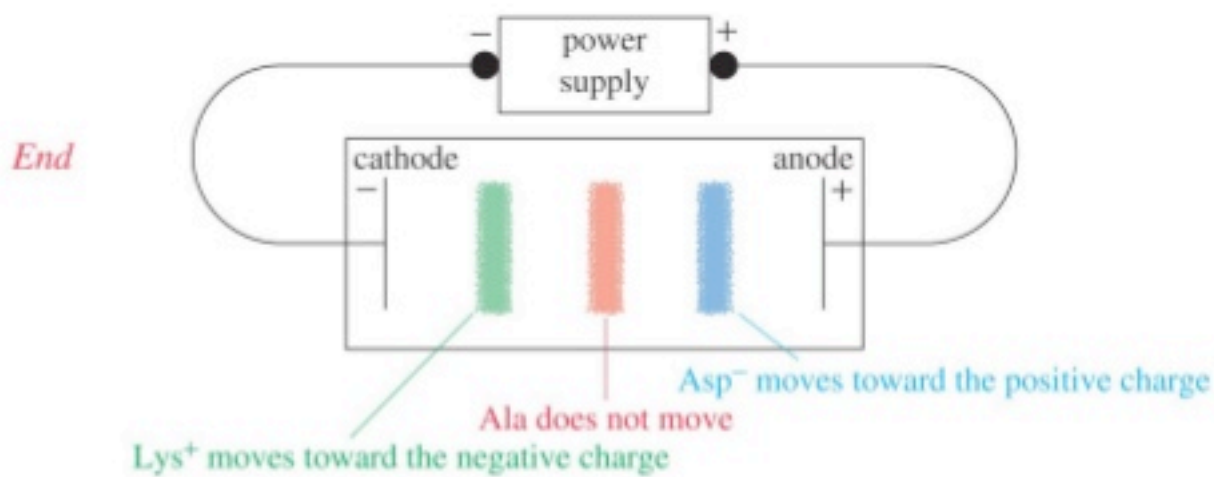
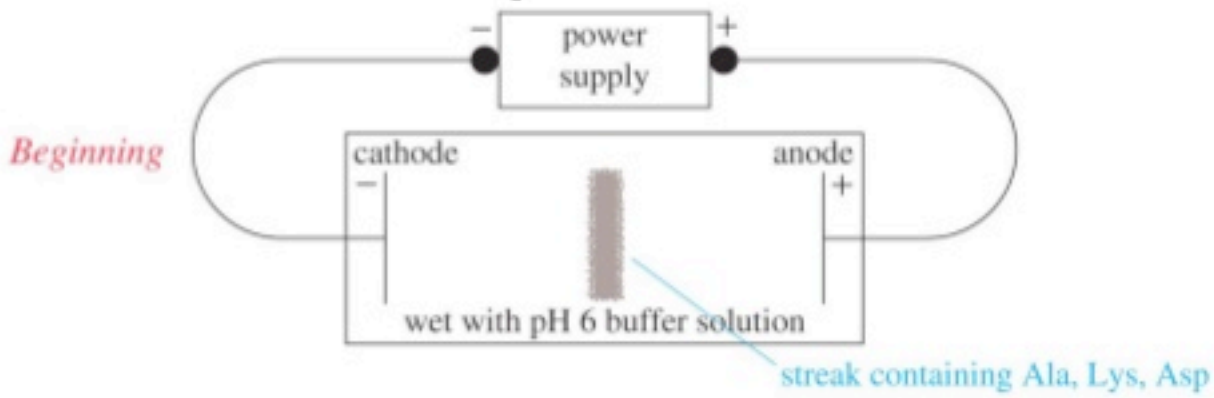


pH < pI → cation

pH = pI → zwitterion

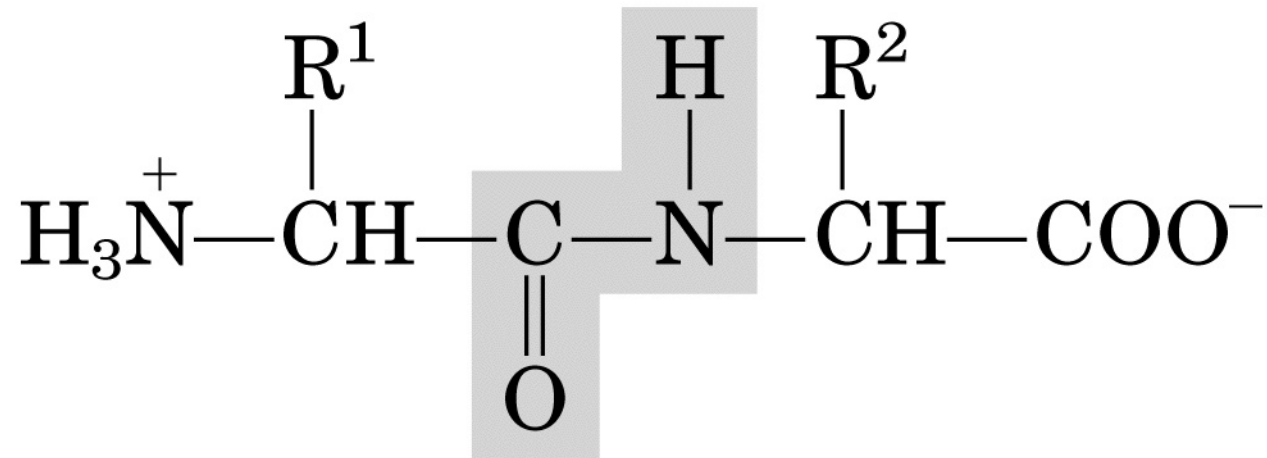
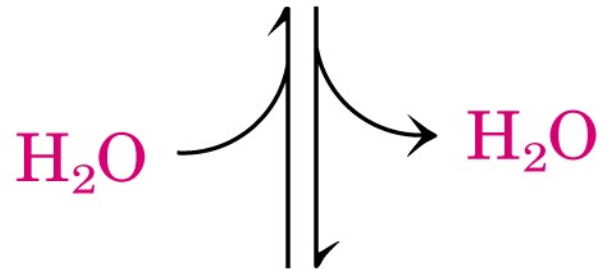
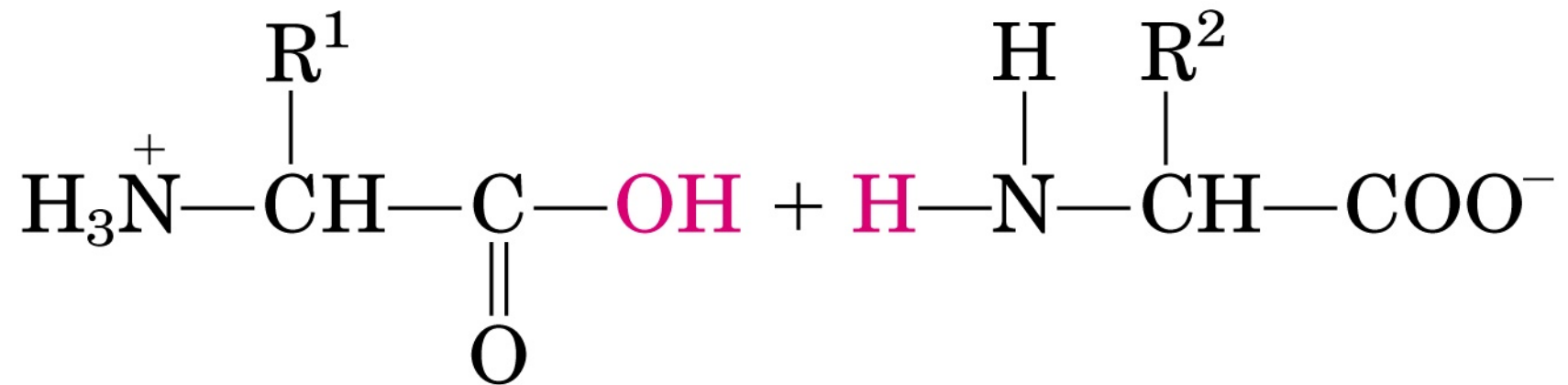
pH > pI → anion

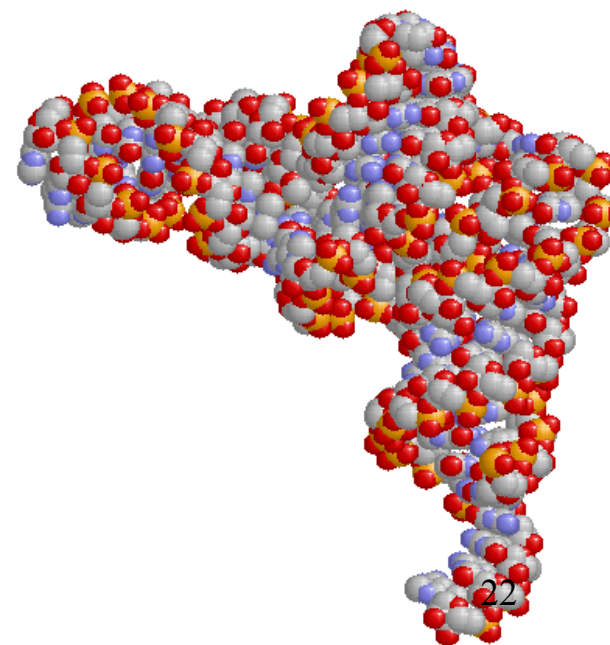
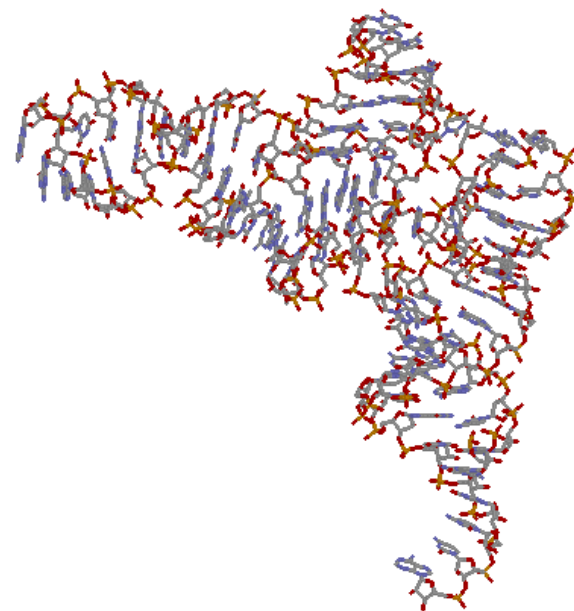
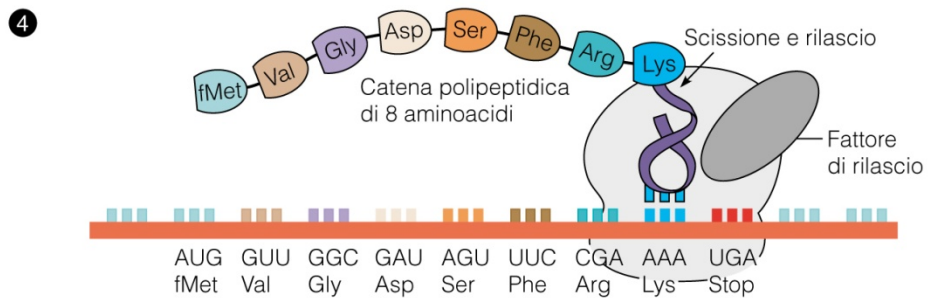
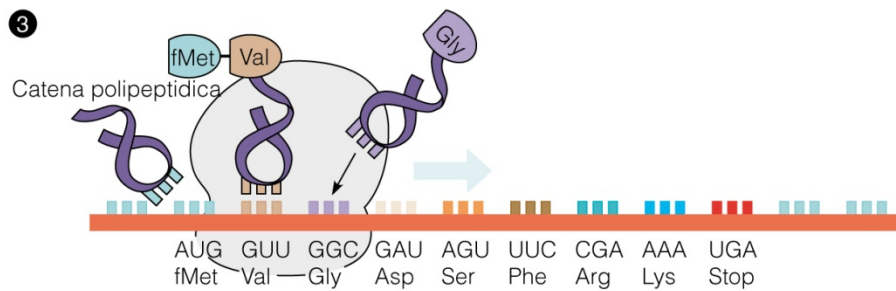
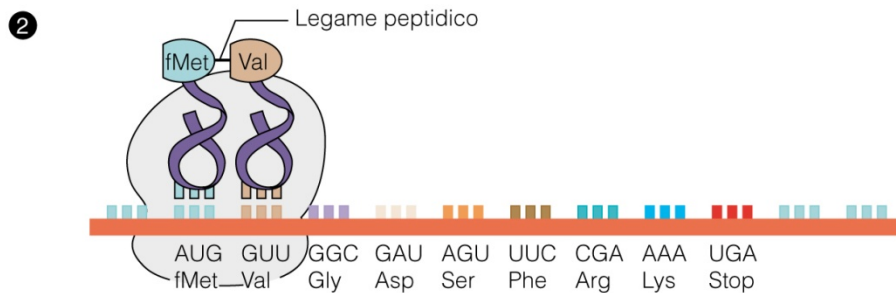
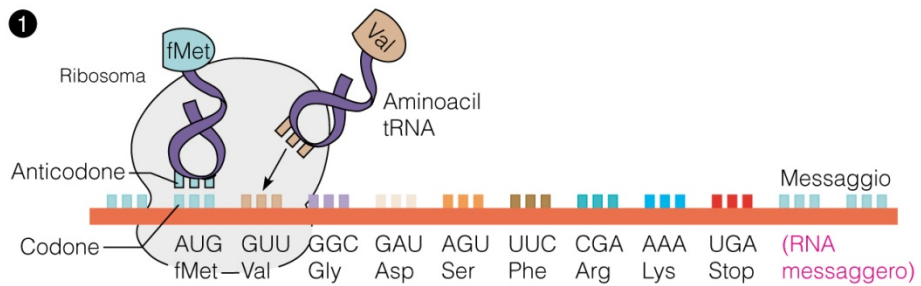
Electrophoresis



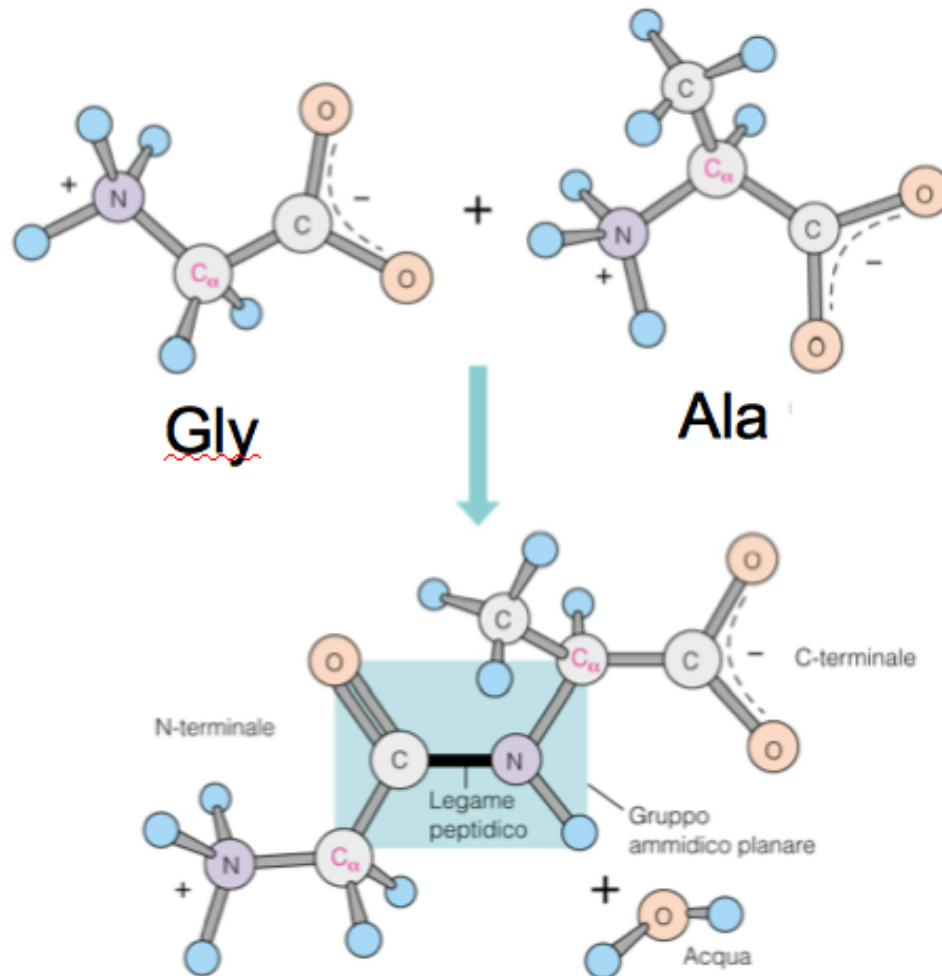
Copyright © 2010 Pearson Prentice Hall, Inc.

Peptide bond.





Polymerization of aa



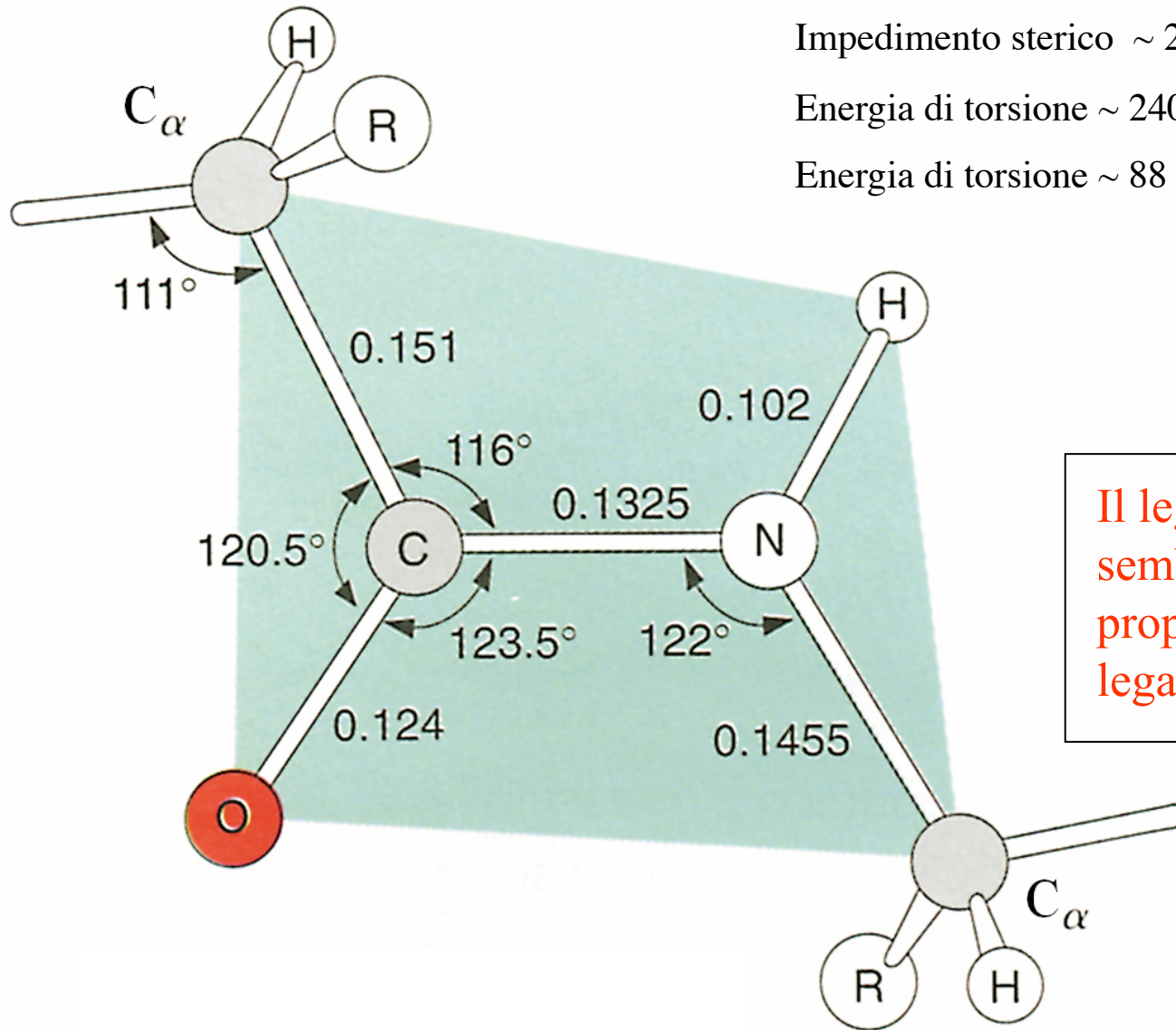
- The peptide bond has 40% character of double bond
- 6 atoms lie in the plane of the bond
- There is a partial negative charge on the carbonyl O and a partial positive charge on the amino N
- All the peptide bonds in proteins are in the *trans* configurations

Energia di torsione $\sim 12 \text{ kJ mol}^{-1}$ (etano)

Impedimento sterico $\sim 24 \text{ kJ mol}^{-1}$ (butano)

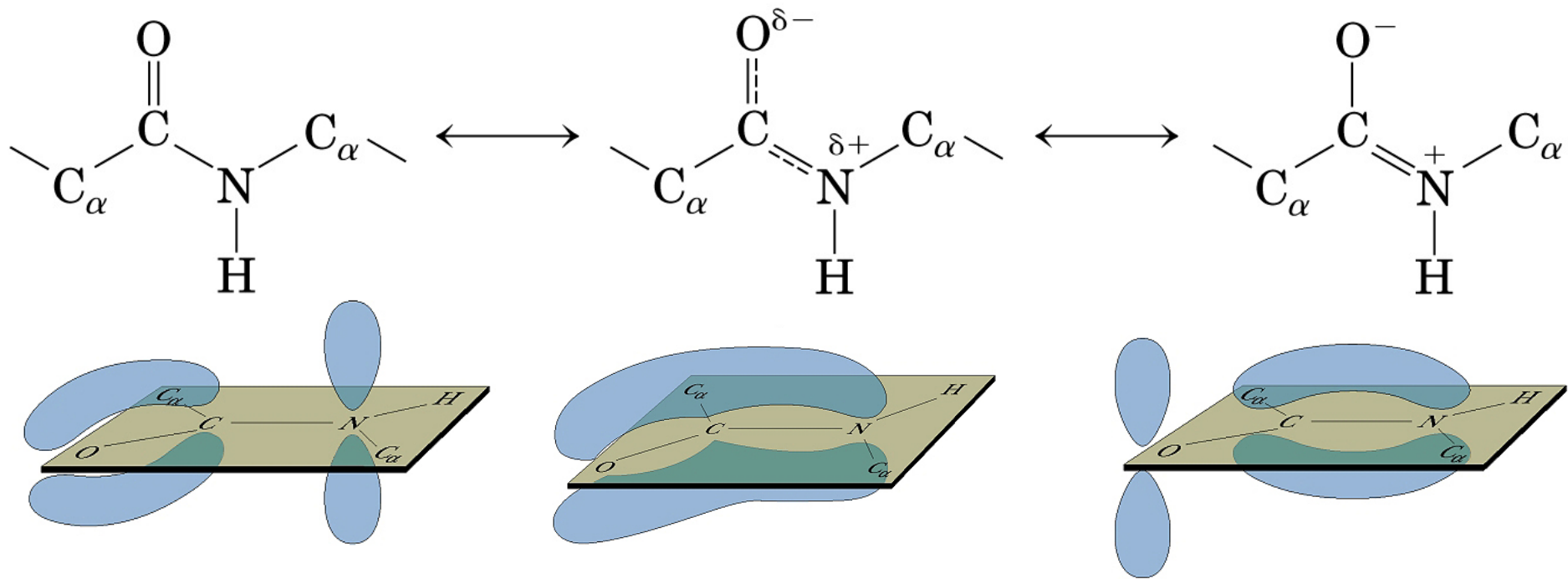
Energia di torsione $\sim 240 \text{ kJ mol}^{-1}$ (2-butene)

Energia di torsione $\sim 88 \text{ kJ mol}^{-1}$ (legame peptidico)



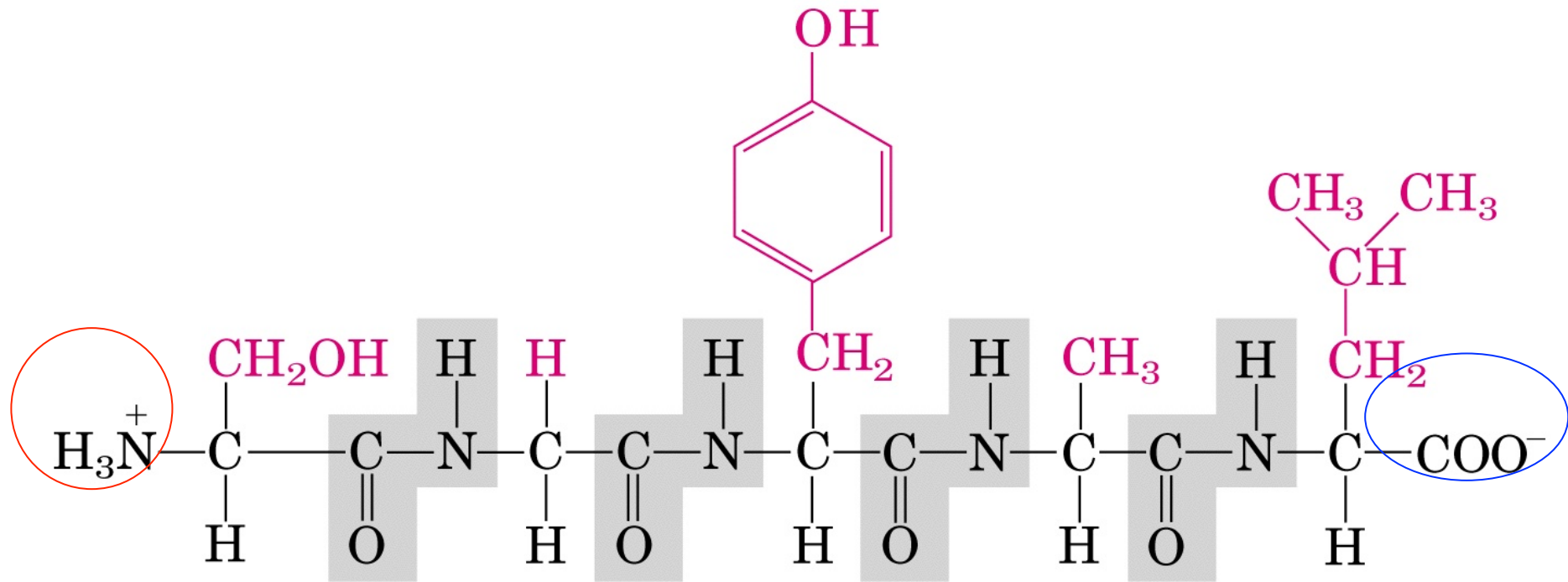
Il legame peptidico sembra non avere le proprietà strutturali di un legame semplice

The peptide bond has a partial double bond character. The carbonyl oxygen atom has a partial negative charge and the amide nitrogen has a partial positive charge. This generates an electric dipole in the hybrid. Virtually all peptide bonds in proteins are present in the trans configuration.



Resonance hybrid

The polypeptide chain (protein) has a direction.




N-terminus

C-terminus

The sequence of a protein, maltoporin.

N-term residue



VAL	ASP	PHE	HIS	GLY	TYR	ALA	ARG	SER	GLY	ILE	GLY	TRP	THR	GLY	SER	GLY	GLY	GLU	GLN	20
GLN	CYS	PHE	GLN	THR	THR	GLY	ALA	GLN	SER	LYS	TYR	ARG	LEU	GLY	ASN	GLU	CYS	GLU	THR	40
TYR	ALA	GLU	LEU	LYS	LEU	GLY	GLN	GLU	VAL	TRP	LYS	GLU	GLY	ASP	LYS	SER	PHE	TYR	PHE	60
ASP	THR	ASN	VAL	ALA	TYR	SER	VAL	ALA	GLN	GLN	ASN	ASP	TRP	GLU	ALA	THR	ASP	PRO	ALA	80
PHE	ARG	GLU	ALA	ASN	VAL	GLN	GLY	LYS	ASN	LEU	ILE	GLU	TRP	LEU	PRO	GLY	SER	THR	ILE	100
TRP	ALA	GLY	LYS	ARG	PHE	TYR	GLN	ARG	HIS	ASP	VAL	HIS	MET	ILE	ASP	PHE	TYR	TYR	TRP	120
ASP	ILE	SER	GLY	PRO	GLY	ALA	GLY	LEU	GLU	ASN	ILE	ASP	VAL	GLY	PHE	GLY	LYS	LEU	SER	140
LEU	ALA	ALA	THR	ARG	SER	SER	GLU	ALA	GLY	GLY	SER	SER	SER	PHE	ALA	SER	ASN	ASN	ILE	160
TYR	ASP	TYR	THR	ASN	GLU	THR	ALA	ASN	ASP	VAL	PHE	ASP	VAL	ARG	LEU	ALA	GLN	MET	GLU	180
ILE	ASN	PRO	GLY	GLY	THR	LEU	GLU	LEU	GLY	VAL	ASP	TYR	GLY	ARG	ALA	ASN	LEU	ARG	ASP	200
ASN	TYR	ARG	LEU	VAL	ASP	GLY	ALA	SER	LYS	ASP	GLY	TRP	LEU	PHE	THR	ALA	GLU	HIS	THR	220
GLN	SER	VAL	LEU	LYS	GLY	PHE	ASN	LYS	PHE	VAL	VAL	GLN	TYR	ALA	THR	ASP	SER	MET	THR	240
SER	GLN	GLY	LYS	GLY	LEU	SER	GLN	GLY	SER	GLY	VAL	ALA	PHE	ASP	ASN	GLU	LYS	PHE	ALA	260
TYR	ASN	ILE	ASN	ASN	ASN	GLY	HIS	MET	LEU	ARG	ILE	LEU	ASP	HIS	GLY	ALA	ILE	SER	MET	280
GLY	ASP	ASN	TRP	ASP	MET	MET	TYR	VAL	GLY	MET	TYR	GLN	ASP	ILE	ASN	TRP	ASP	ASN	ASP	300
ASN	GLY	THR	LYS	TRP	TRP	THR	VAL	GLY	ILE	ARG	PRO	MET	TYR	LYS	TRP	THR	PRO	ILE	MET	320
SER	THR	VAL	MET	GLU	ILE	GLY	TYR	ASP	ASN	VAL	GLU	SER	GLN	ARG	THR	GLY	ASP	LYS	ASN	340
ASN	GLN	TYR	LYS	ILE	THR	LEU	ALA	GLN	GLN	TRP	GLN	ALA	GLY	ASP	SER	ILE	TRP	SER	ARG	360
PRO	ALA	ILE	ARG	VAL	PHE	ALA	THR	TYR	ALA	LYS	TRP	ASP	GLU	LYS	TRP	GLY	TYR	ASP	TYR	380
THR	GLY	ASN	ALA	ASP	ASN	ASN	ALA	ASN	PHE	GLY	LYS	ALA	VAL	PRO	ALA	ASP	PHE	ASN	GLY	400
GLY	SER	PHE	GLY	ARG	GLY	ASP	SER	ASP	GLU	TRP	THR	PHE	GLY	ALA	GLN	MET	GLU	ILE	TRP	420
TRP																				421

C-term residue

Table 6-1 Molecular data on some proteins

	Molecular weight	Number of residues	Number of polypeptide chains
Insulin (bovine)	5,733	51	2
Cytochrome <i>c</i> (human)	13,000	104	1
Ribonuclease A (bovine pancreas)	13,700	124	1
Lysozyme (egg white)	13,930	129	1
Myoglobin (equine heart)	16,890	153	1
Chymotrypsin (bovine pancreas)	21,600	241	3
Chymotrypsinogen (bovine)	22,000	245	1
Hemoglobin (human)	64,500	574	4
Serum albumin (human)	68,500	~550	1
Hexokinase (yeast)	102,000	~800	2
Immunoglobulin G (human)	145,000	~1,320	4
RNA polymerase (<i>E. coli</i>)	450,000	~4,100	5
Apolipoprotein B (human)	513,000	4,536	1
Glutamate dehydrogenase (bovine liver)	1,000,000	~8,300	~40

Type of conjugated protein:

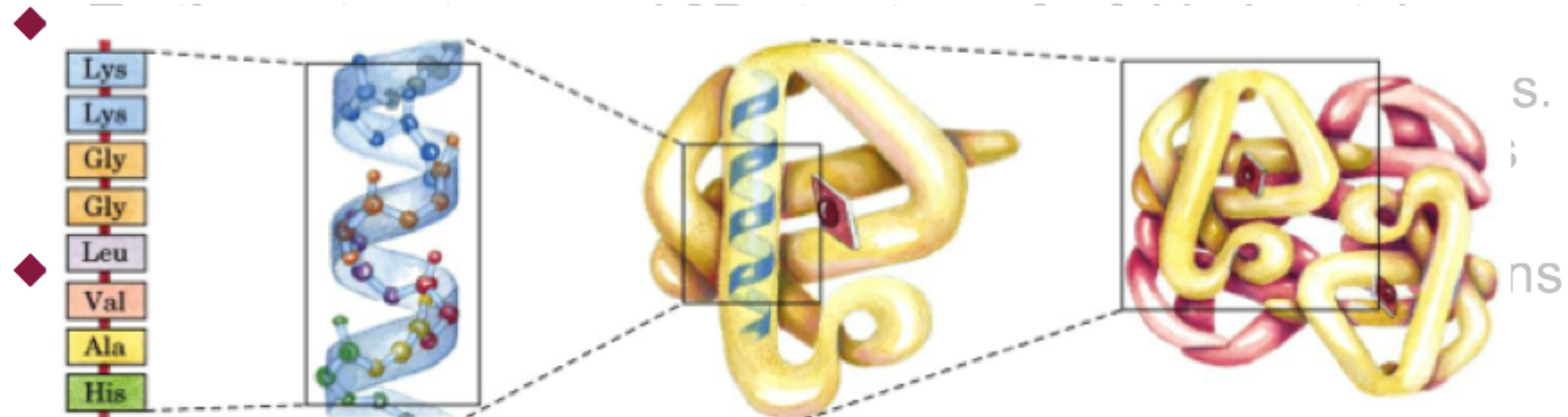
- **Nucleoproteins:** Contains nucleic acids, nitrogen, and phosphorus. Present in chromosomes and in all living forms as a combination of protein with either RNA or DNA.
- **Glycoprotein:** bone (osseomucoid); tendons (tendomucoid); cartilage (chondromucoid).
- **Phosphoproteins:** milk (casein); egg yolk (ovovitellin). Phosphoric acid joined in ester linkage to protein.
- **Lipoproteins:** serum lipoprotein; brain, nerve tissues, milk, and eggs. Water-soluble protein conjugated with lipids.
- **Chromoproteins:** hemoglobin; myoglobin; flavoproteins; respiratory pigments; cytochromes, colored proteins.

Isoelectric Points of Several Common Proteins

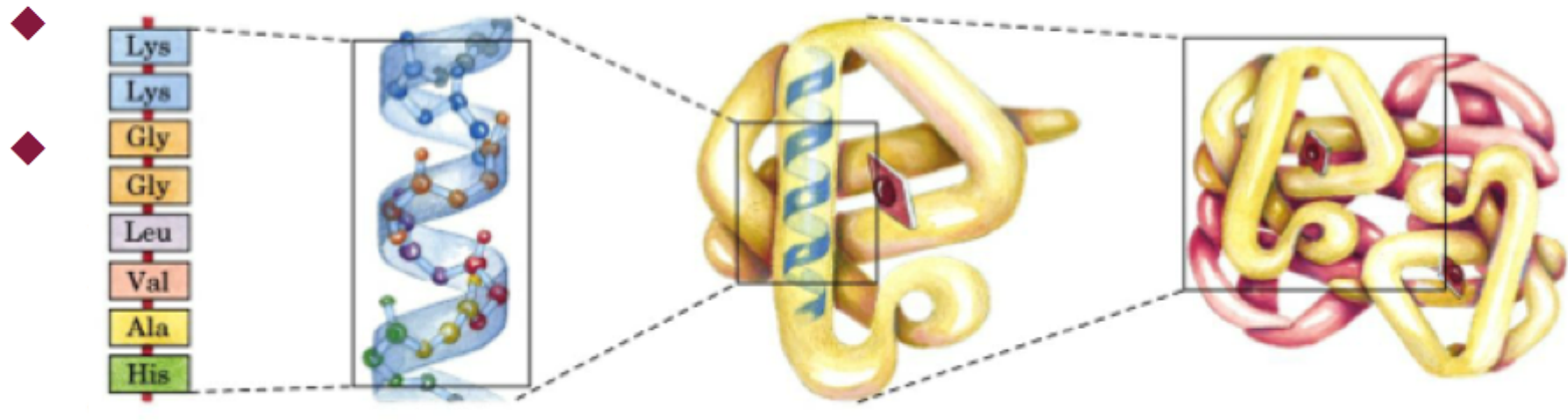
Protein	<i>pI</i>
Pepsin	<1.0
Ovalbumin (hen)	4.6
Serum albumin (human)	4.9
Tropomyosin	5.1
Insulin (bovine)	5.4
Fibrinogen (human)	5.8
γ -Globulin (human)	6.6
Collagen	6.6
Myoglobin (horse)	7.0
Hemoglobin (human)	7.1
Ribonuclease A (bovine)	9.4
Cytochrome <i>c</i> (horse)	10.6
Histone (bovine)	10.8
Lysozyme (hen)	11.0
Salmine (salmon)	12.1

Hierarchical organization of proteins

- ◆ **Primary structure:** the linear sequence of aa bound by peptide bonds
- ◆ **Secondary structures:** regular organization of aa in the space, stabilized by hydrogen bonds among CO and NH of several peptide bonds. It can either be periodic (alpha-helix, beta-sheet) or random. Not all the possible fold are allowed given the restrictions dictated by the peptide bond

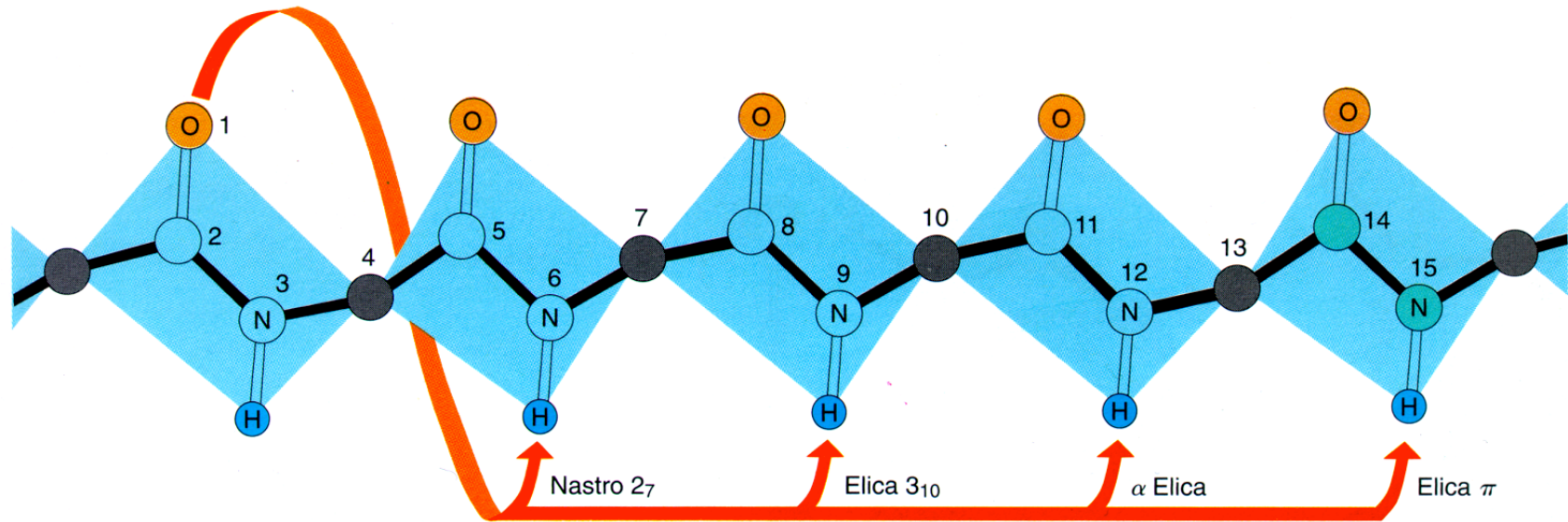


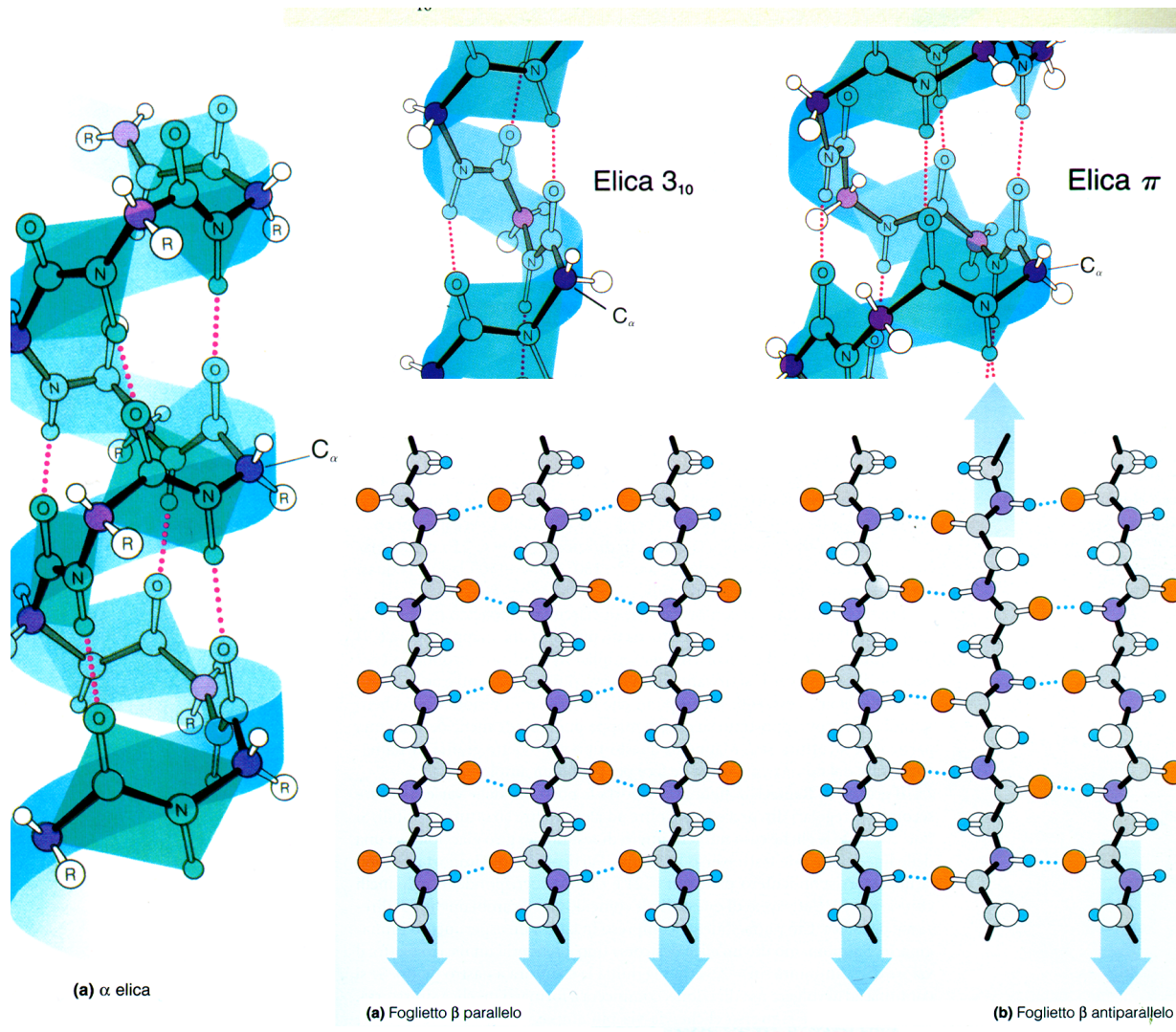
Hierarchical organization of proteins



- ◆ **Tertiary structure:** real 3D structure of a folded protein, derived from the assembly of secondary structure elements. These are stabilized by polar and hydrophobic interactions among the side chains
- ◆ **Quaternary structure:** oligomeric assembly (not all proteins have this arrangement)

Secondary structure





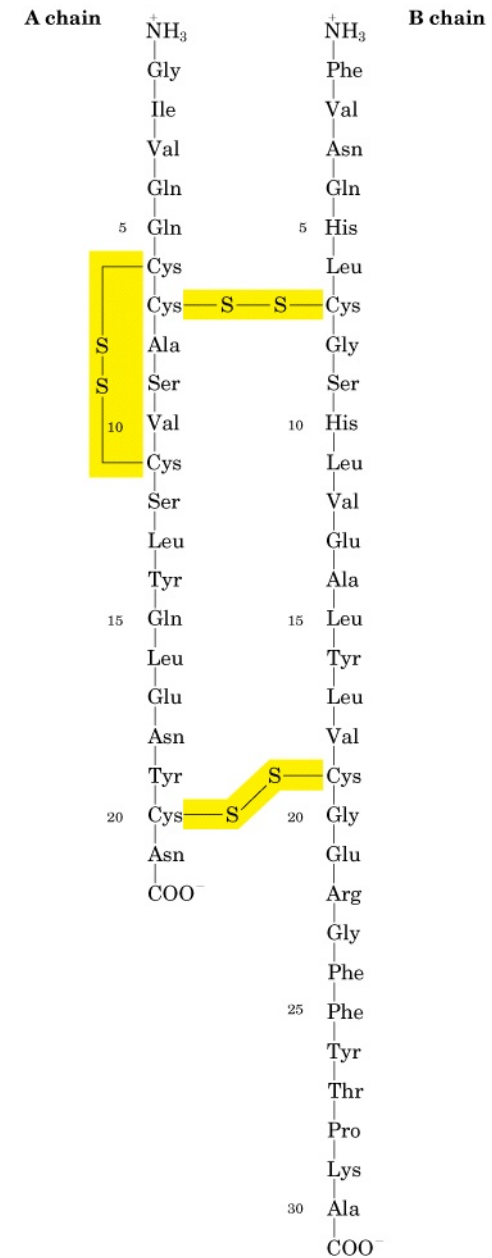
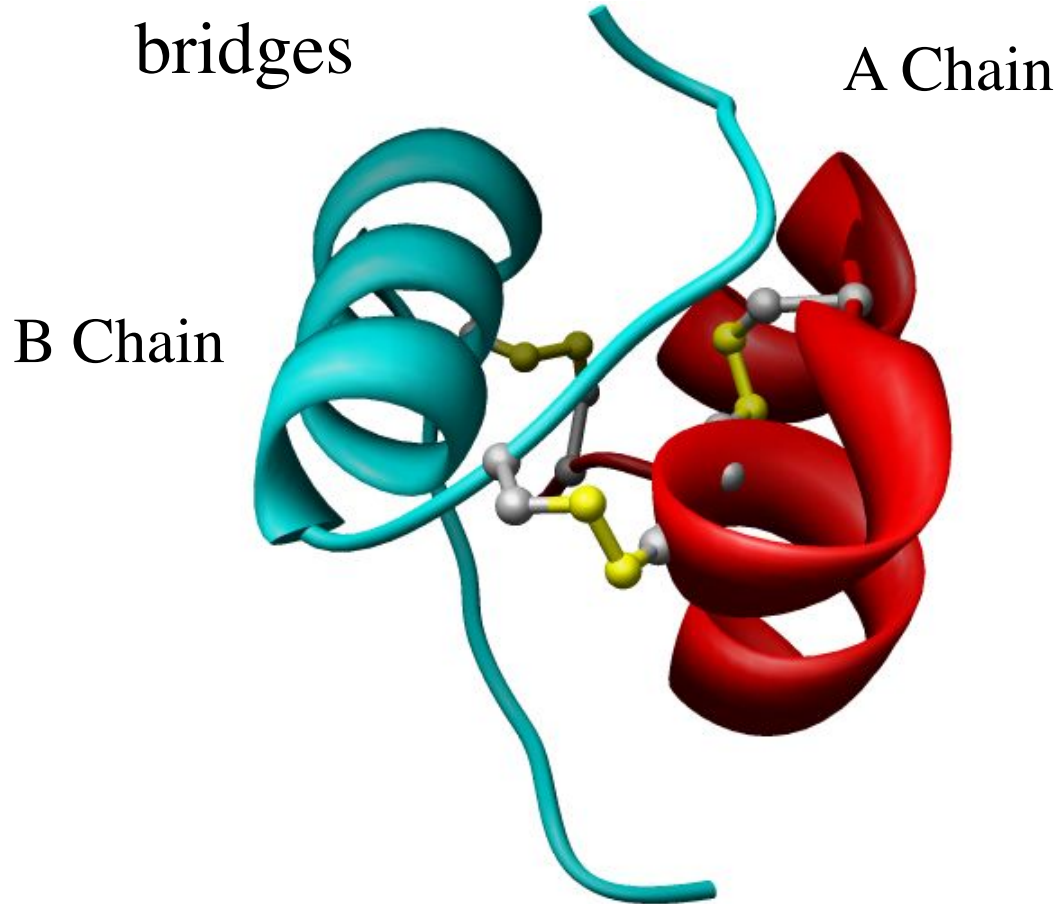
Secondary structure

α Helix

β Sheet

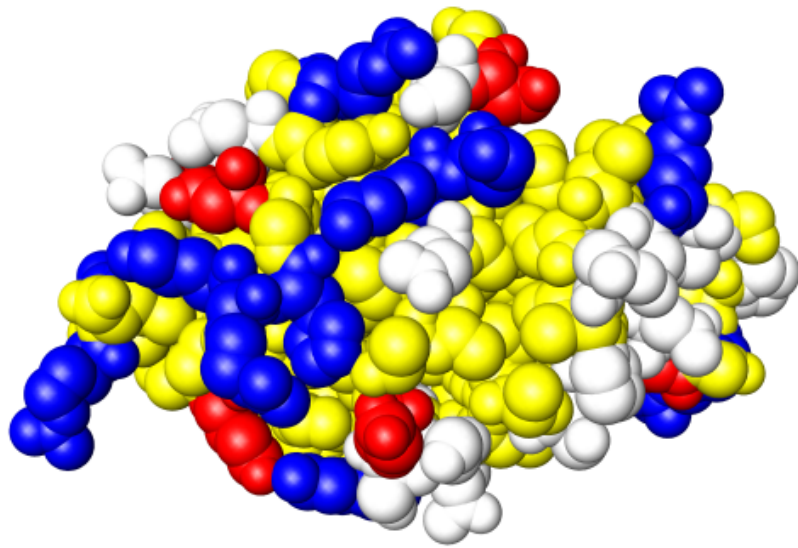
Insulin

Role of disulphide bridges

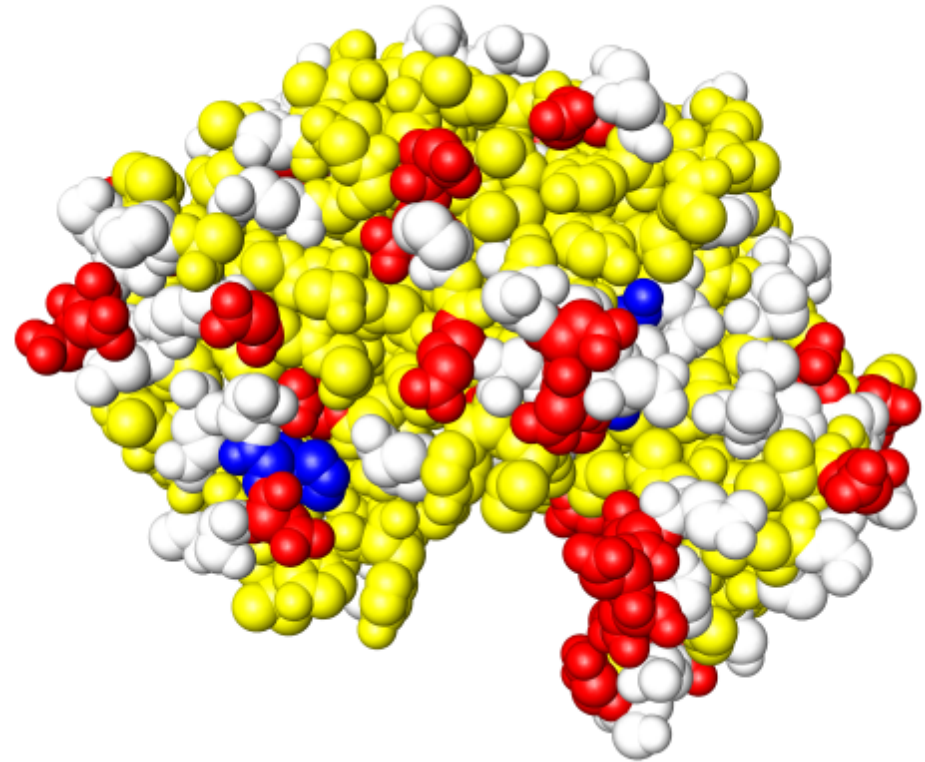


Tertiary structure

Lysozyme (*Gallus gallus*) pI = 11

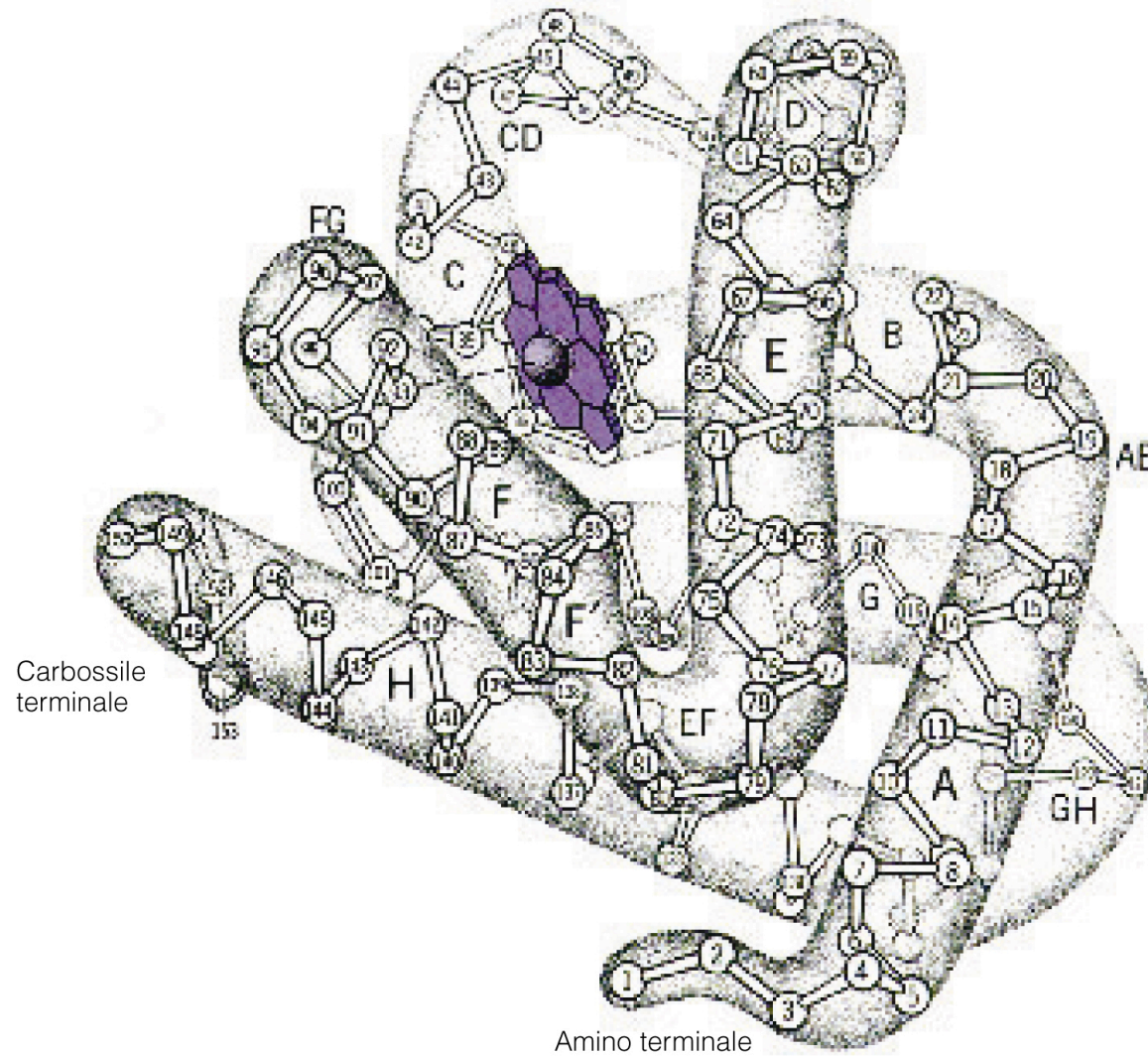


Pepsine (*Sus scropha*) pI = 1

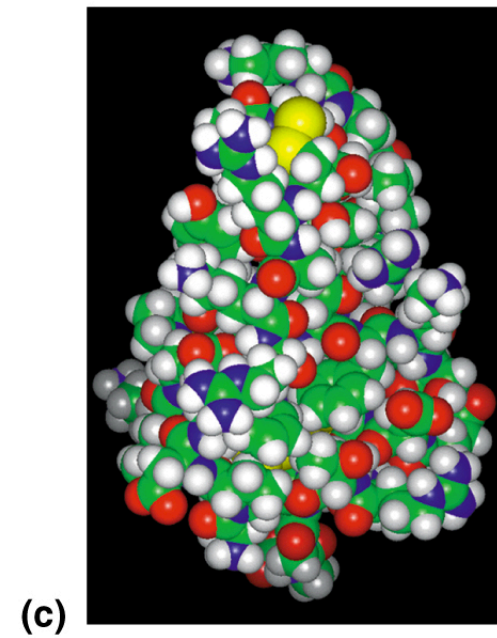
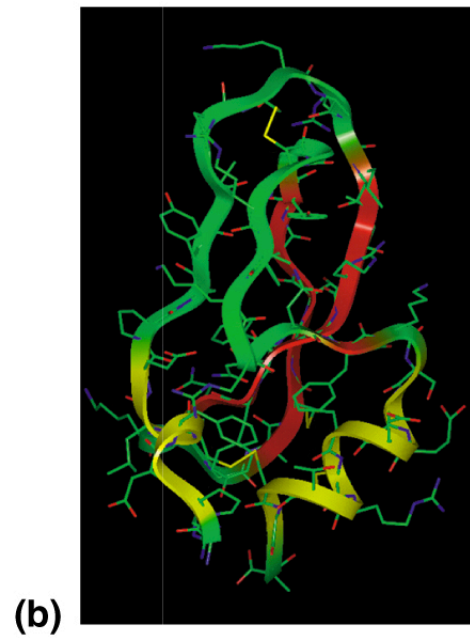
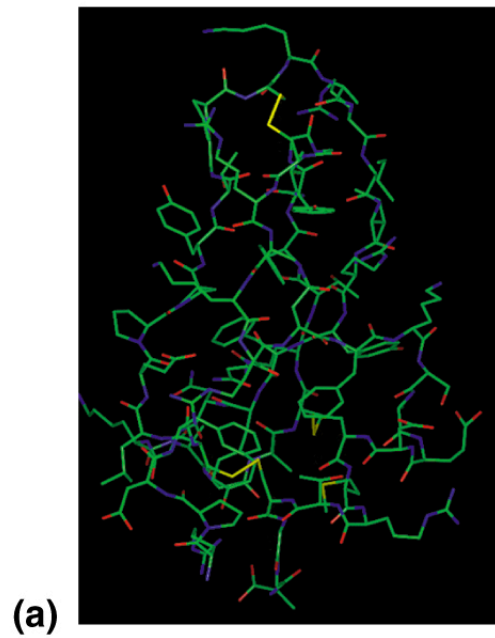


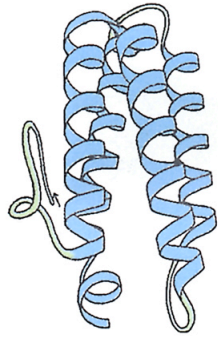
acidic = red; basic = blue; polar = white; hydrophobic = yellow

Myoglobin

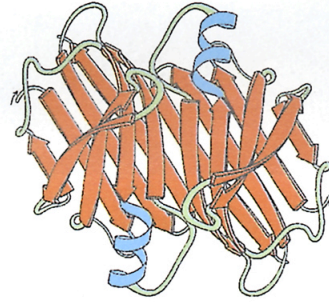


Ways to depict protein structures

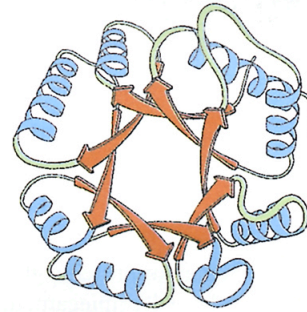




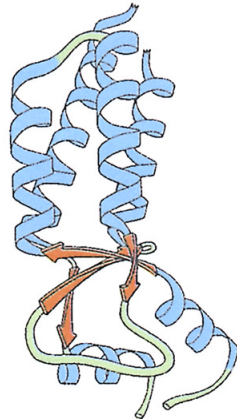
Mioemeritina



Prealbumina

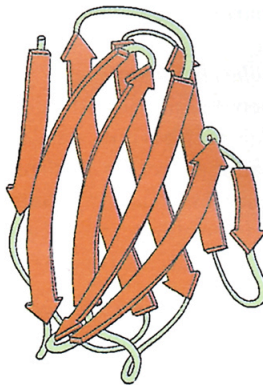


Piruvato chinasi, dominio 1



Proteina del capsido del
mosaico del tabacco

(a) Prevalentemente
ad α elica



Immunoglobulina, dominio V_2

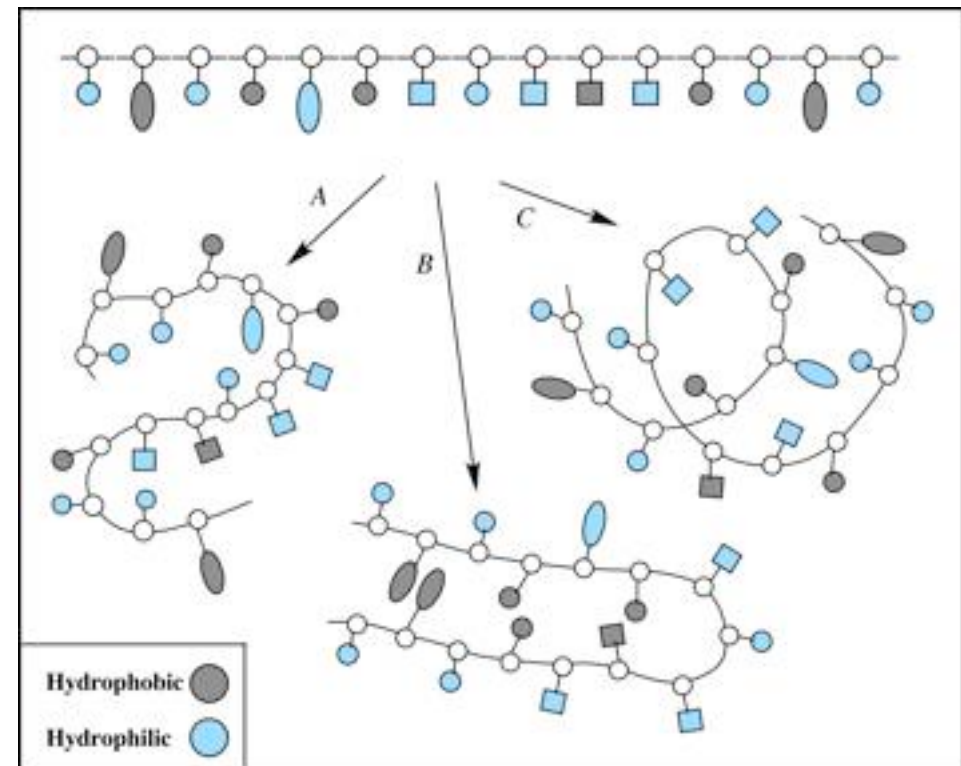
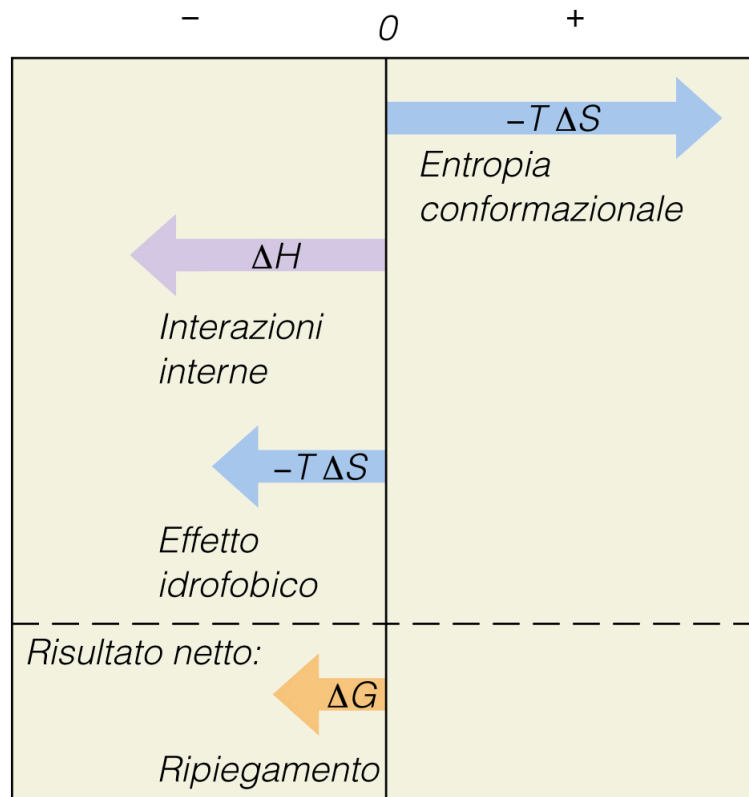
(b) Prevalentemente
a foglietto β



Esochinasi, dominio 2

(c) Struttura mista ad α elica
e a foglietto β

Protein folding



Hydrophobic effect

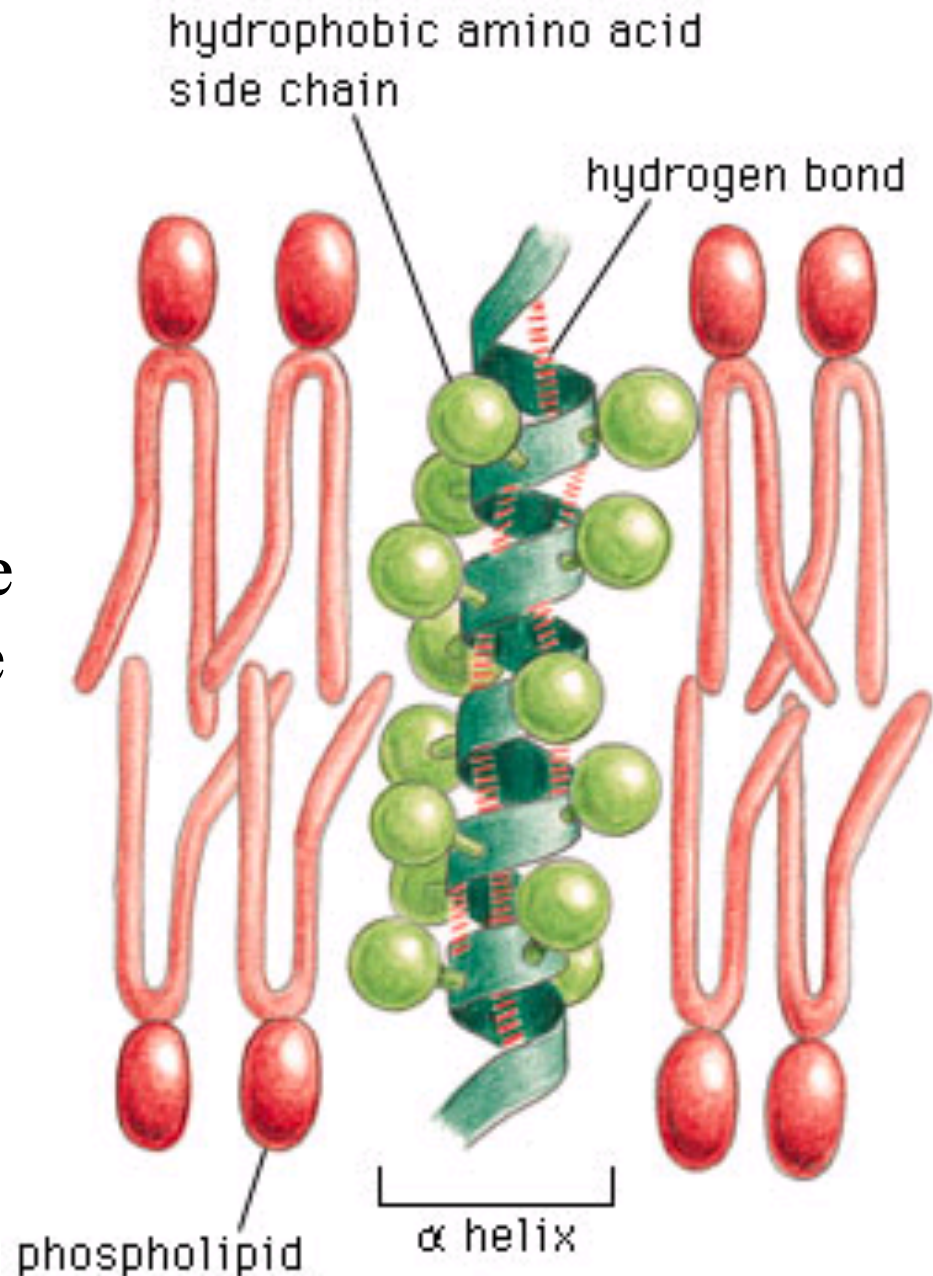
Membrane Proteins

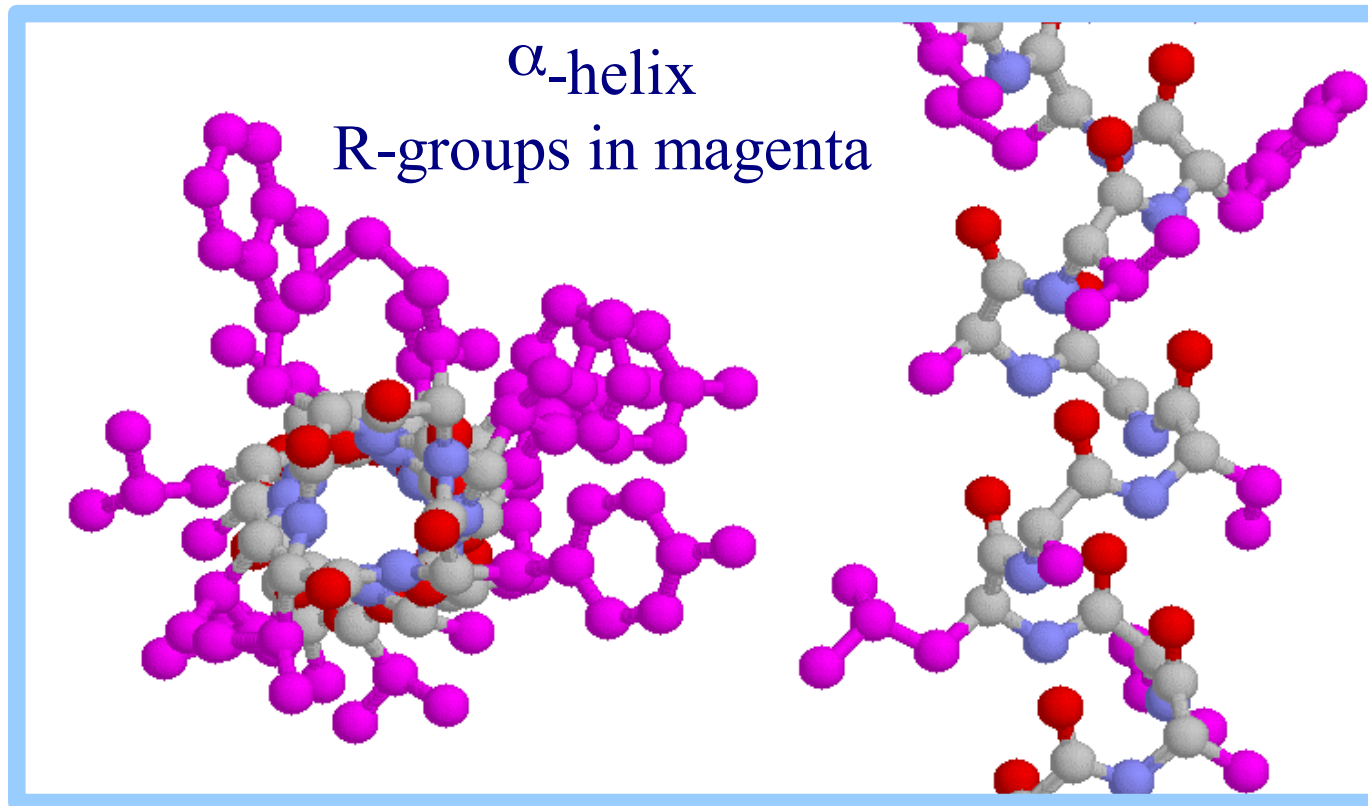
20-30% of the total open reading frames are integral membrane proteins

Membrane proteins are targets for >40% of FDA approved drugs
G-protein coupled receptors, channels, transporters, receptors, proteases (RIP), etc

There are ~1.5% of deposited structures (555 unique)
membrane protein structures

The peptide bonds are hydrogen bonded to each other in the interior while the hydrophobic amino acid side chains contact the lipid chains.





In an α -helix, amino acid R-groups protrude out from the helically coiled polypeptide backbone.

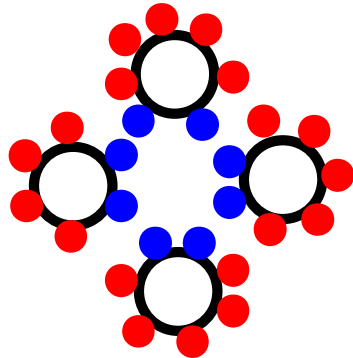
The largely hydrophobic R-groups of a membrane-spanning α -helix contact the hydrophobic membrane core, while the more polar peptide backbone is buried.

Colors: C N O R-group (H atoms not shown).

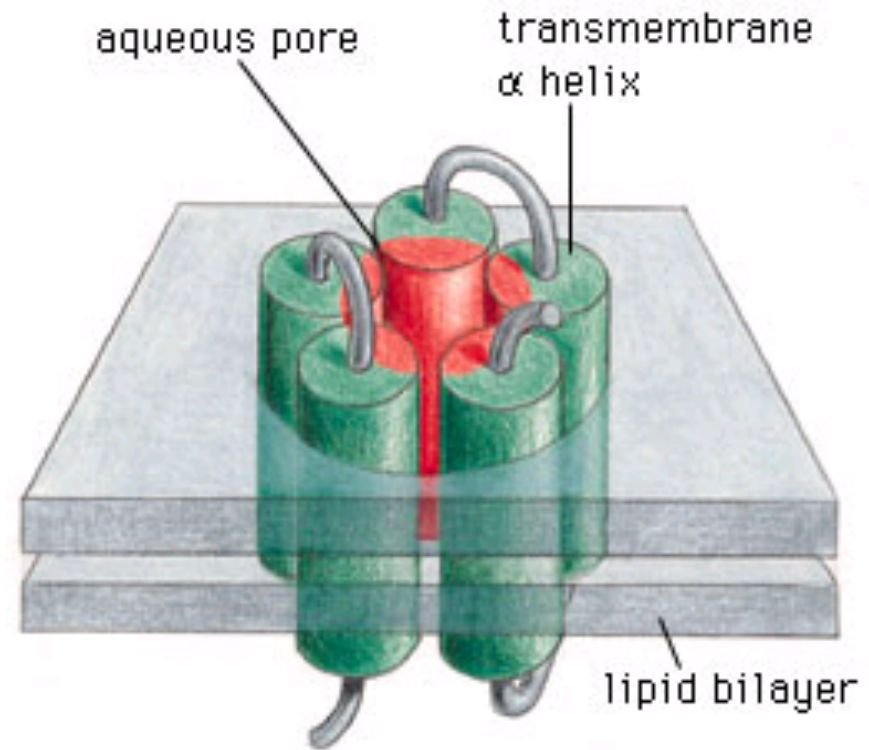
Since a 20-amino acid α -helix just spans a lipid bilayer, one could use **Hydropathy plots to search for 20-amino acid stretches of hydrophobic amino acids in the primary sequence of a protein to identify membrane proteins in open reading frames.**

Except →

Simplified helical wheel diagram of four α -helices lining the lumen of an ion channel.



- Polar amino acid R-group
- Non-polar amino acid R-group



©1998 GARLAND PUBLISHING

An α -helix lining a **water-filled channel** might have polar amino acid R-groups facing the lumen, & non-polar R-groups facing lipids or other hydrophobic α -helices.

Such **mixed polarity** would prevent detection by a hydropathy plot.

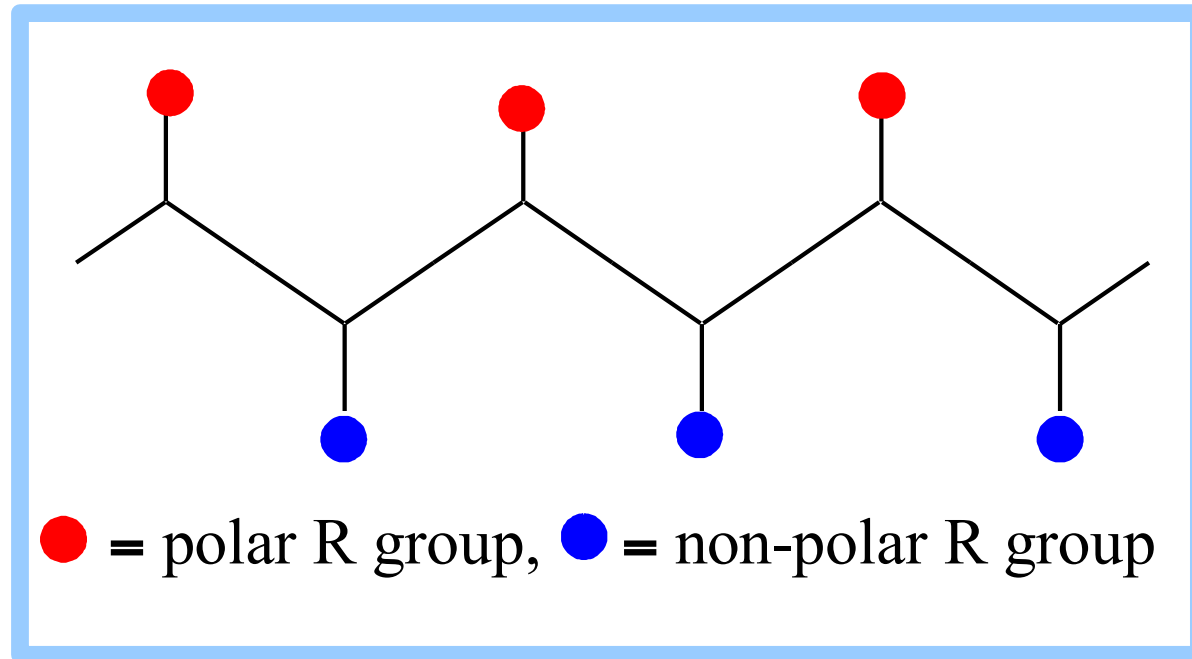
Porin is a β -barrel membrane protein

Porins reside in bacterial outer membrane, have instead β barrel structures.

A β barrel are β sheets rolled up to form a cylindrical pore.

β barrels are less versatile since they can only form wide pores, ~ 20 Å in radius.



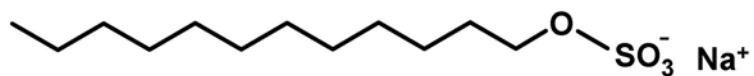


In a β -strand, amino acid R-groups alternately point above & below the strand.

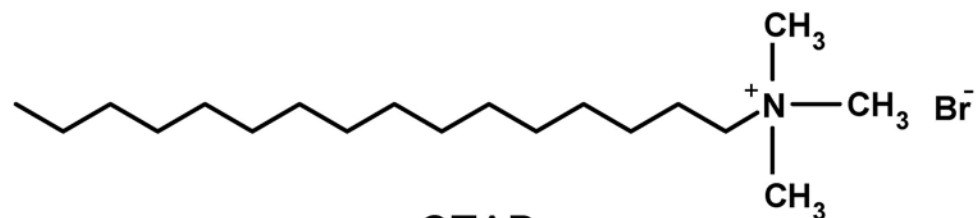
Much of porin primary structure consists of **alternating polar & non-polar amino acids.**

- Polar residues face the aqueous lumen.
- Non-polar residues are in contact with membrane lipids.

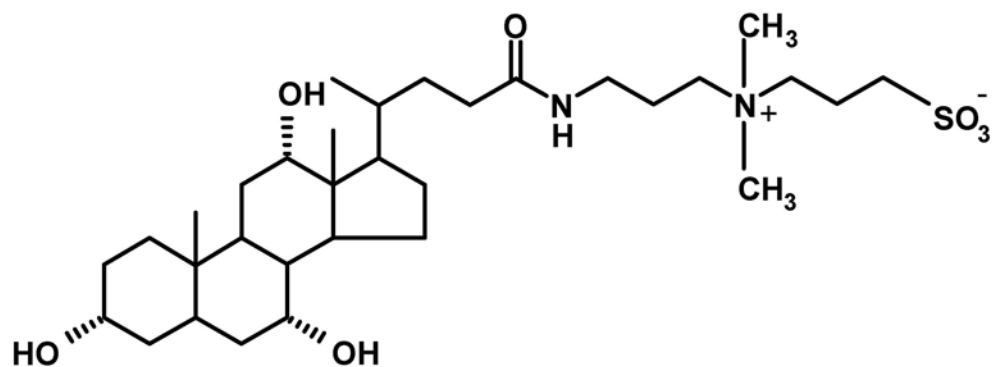
Working with Membrane Proteins



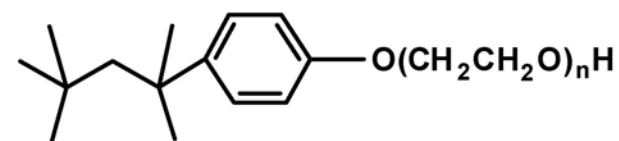
SDS



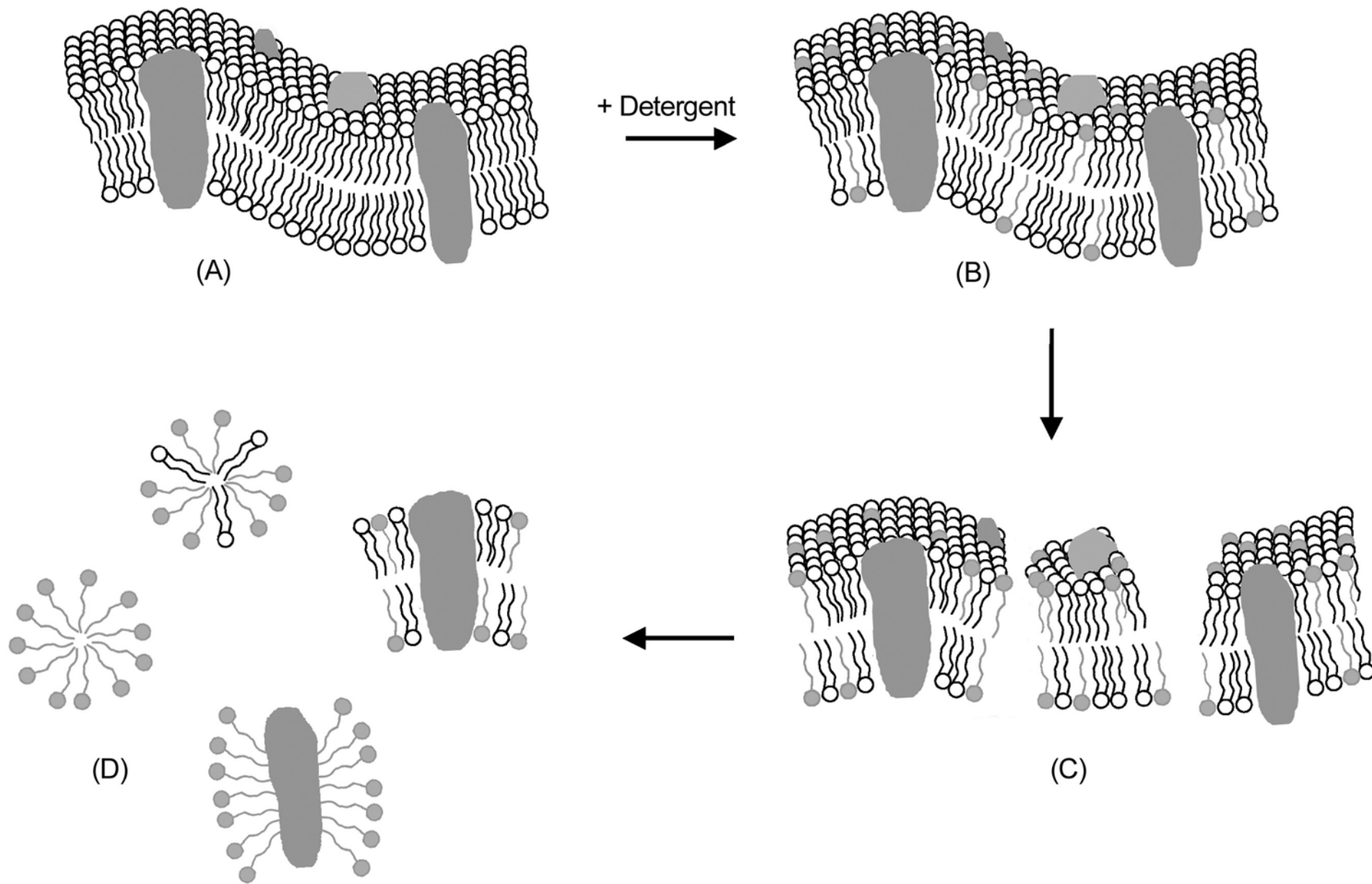
CTAB



CHAPS



Triton X-100 (n=9-10)

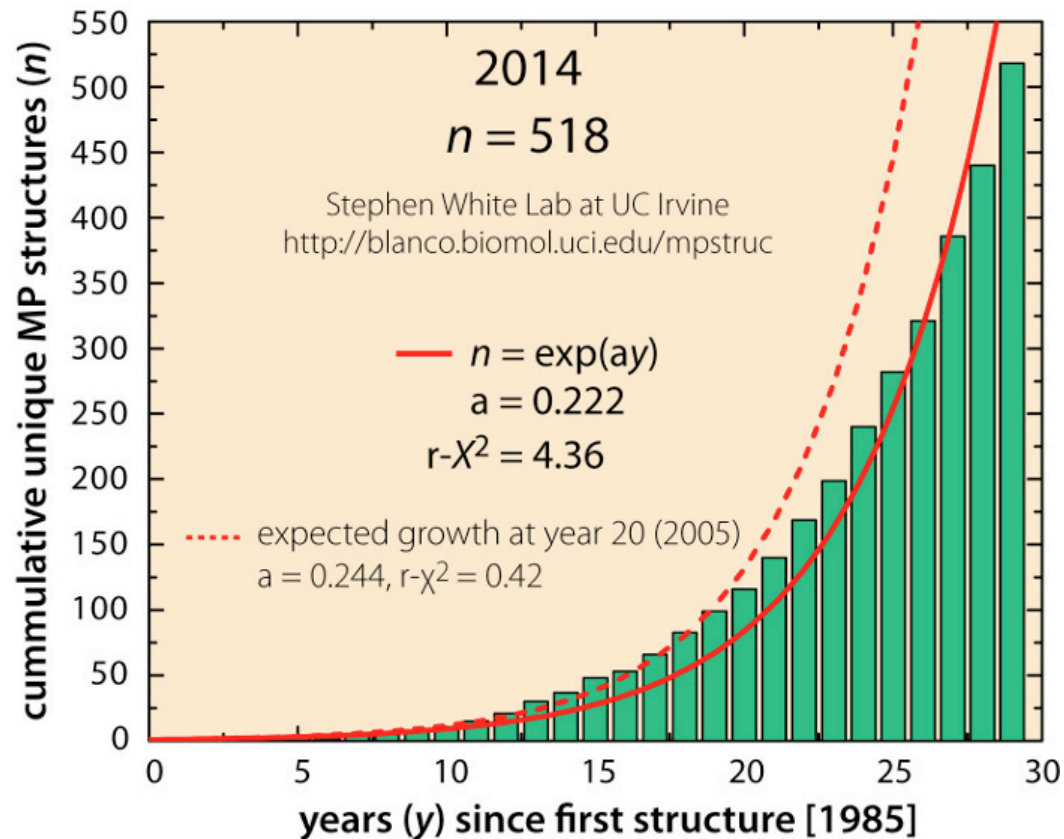


Membrane Protein Structure Determination

~30% of the entire proteome (prokaryotic and eukaryotic)

~50% of all commercial drugs are against membrane protein targets

~1.5% of deposited structures (555 unique)



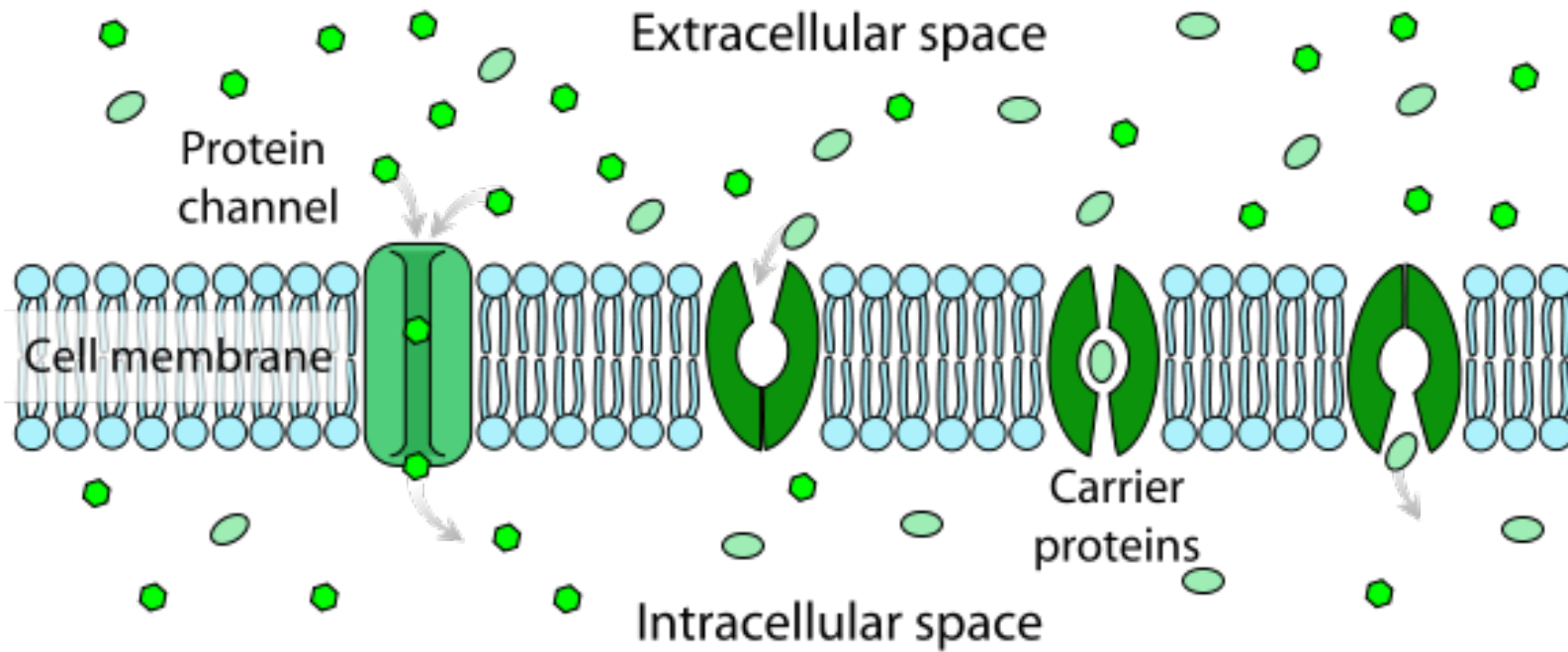
Integral Membrane Proteins

Channels / Transporters

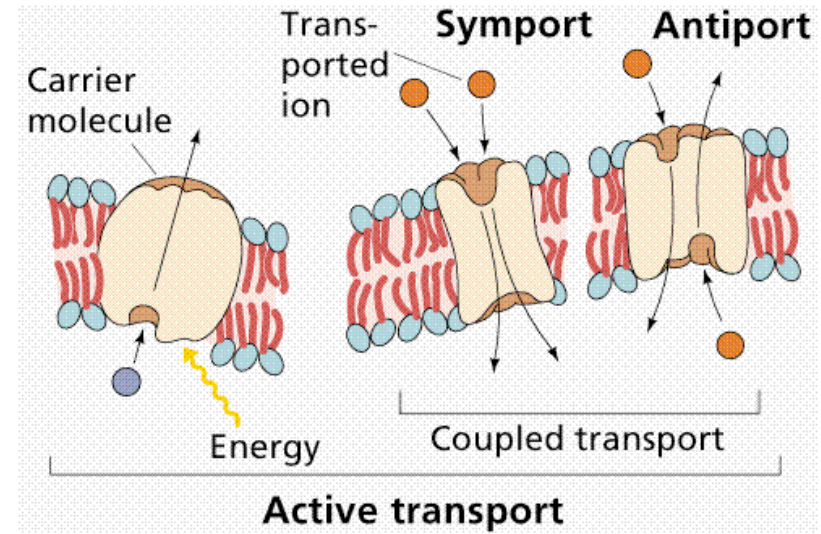
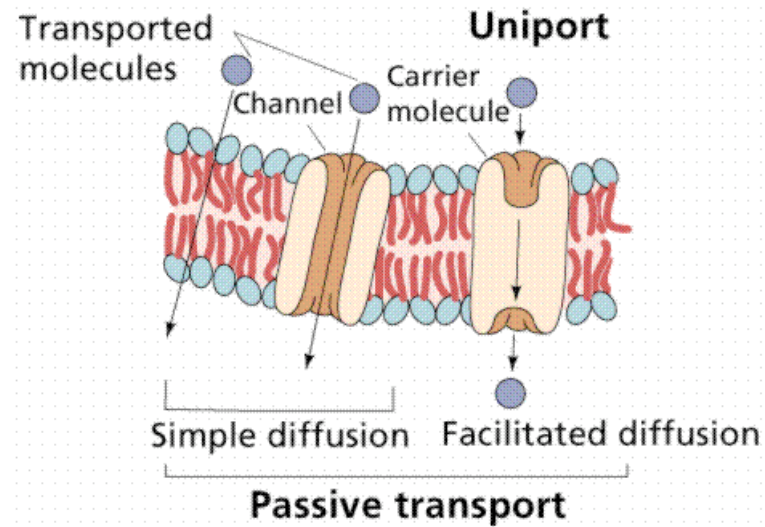
Receptors

Enzymes

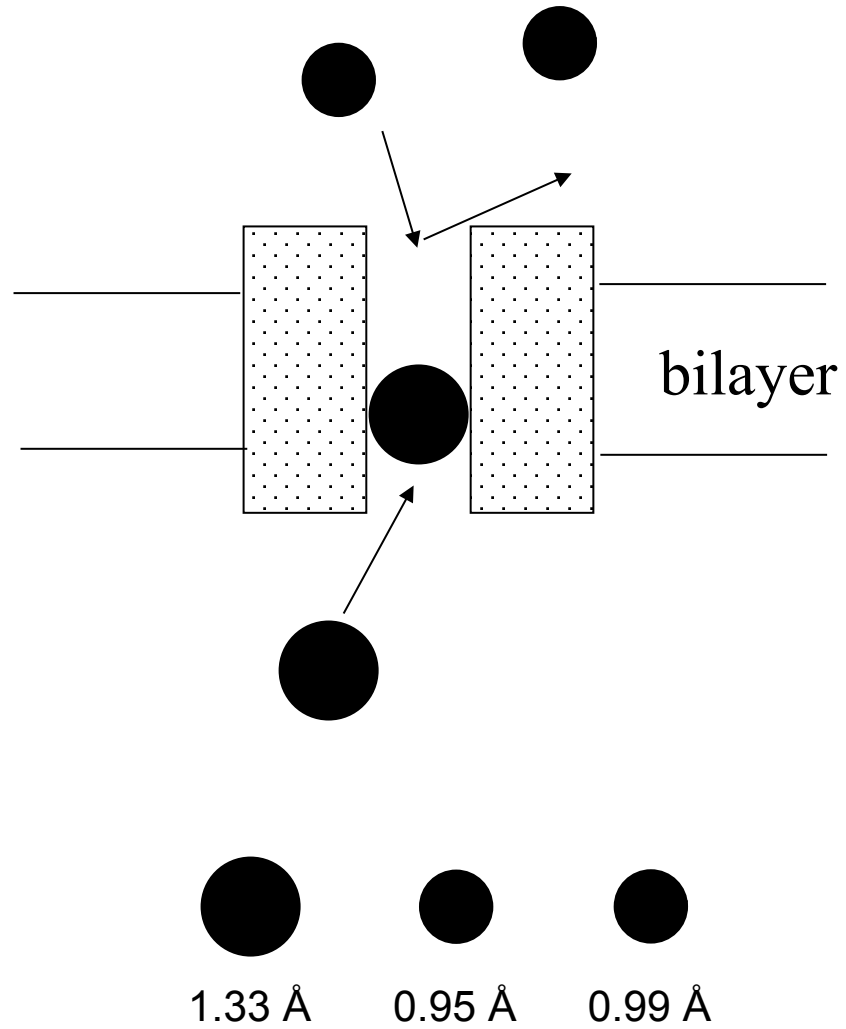
Molecular Physiology of Membrane transport



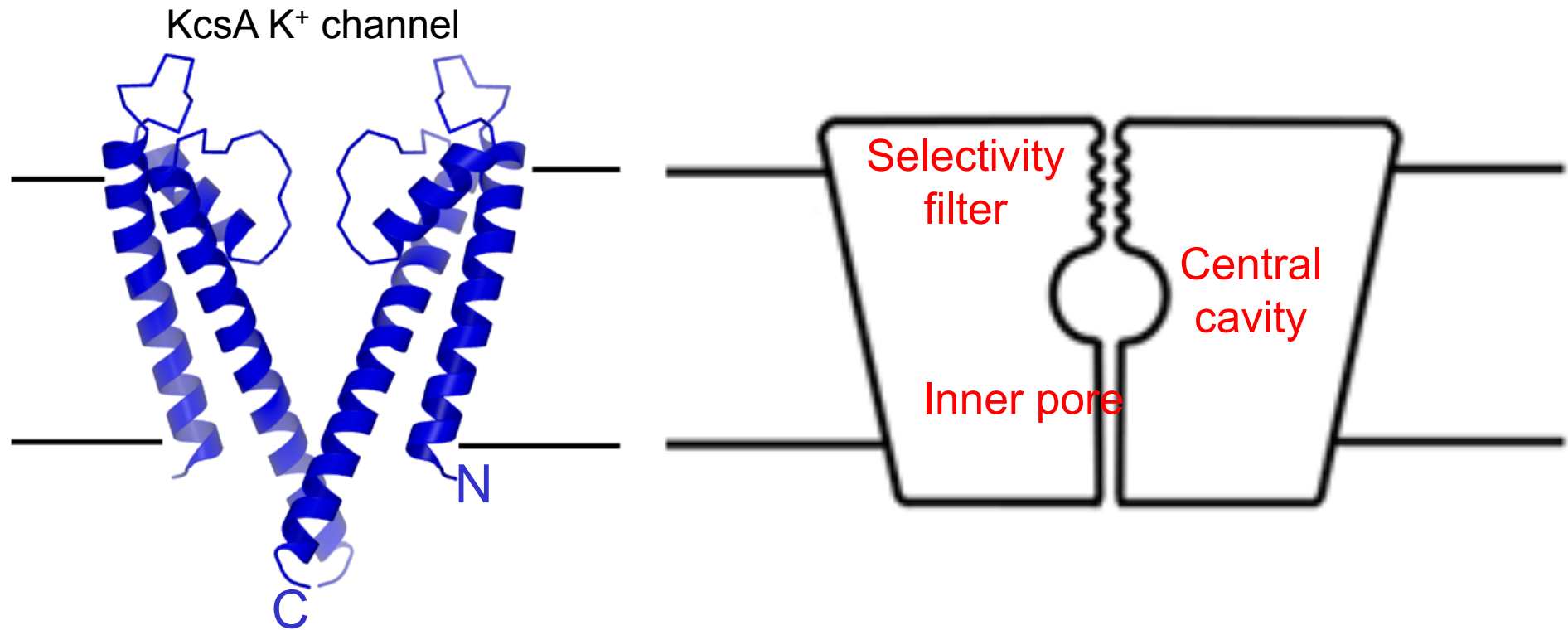
Channels / Transporters



High selectivity and high throughput



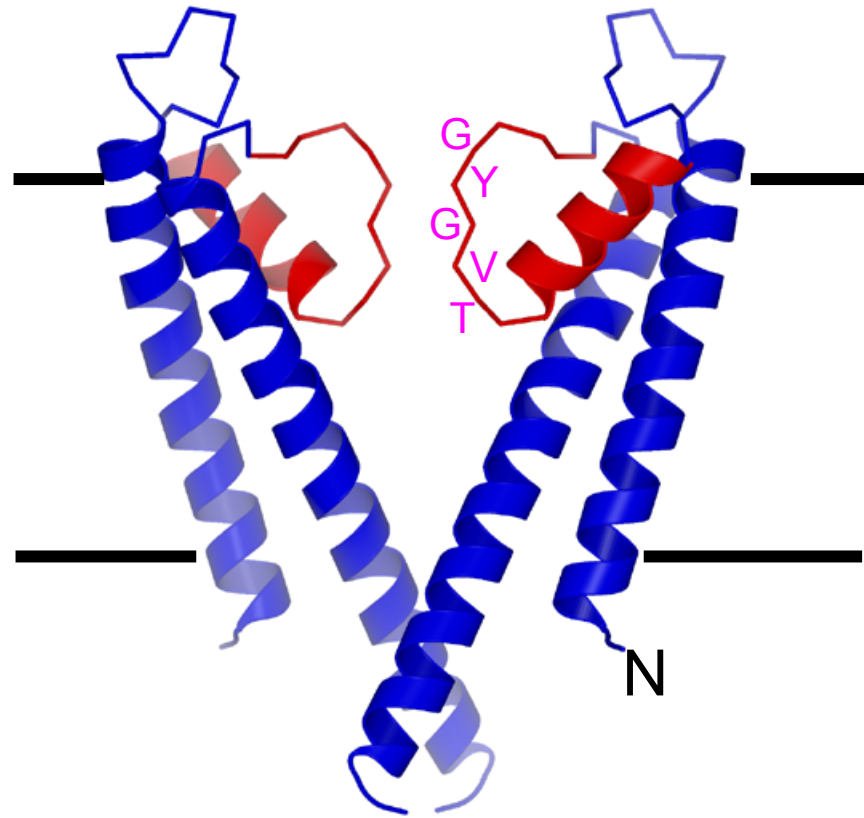
KcsA K⁺ channel is a model pore domain



Zhou Y, Morais-Cabral J, Kaufman A, MacKinnon, R. Nature, 414, 2001

Doyle, D.A., Morais-Cabral J., Pfuetzner, R.A., Kuo, A., Gulbis, J.M., Cohen, S.L., Chait, B.T., MacKinnon, R. Science, 280, 1998

Sequence conservation at the selectivity filter



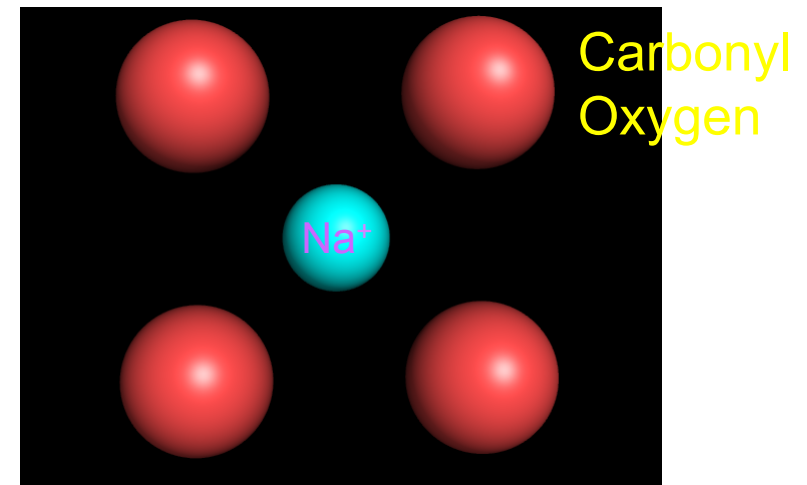
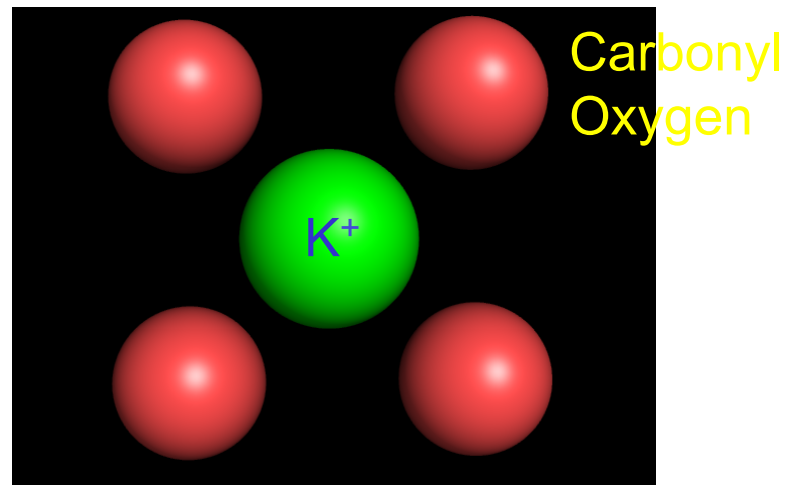
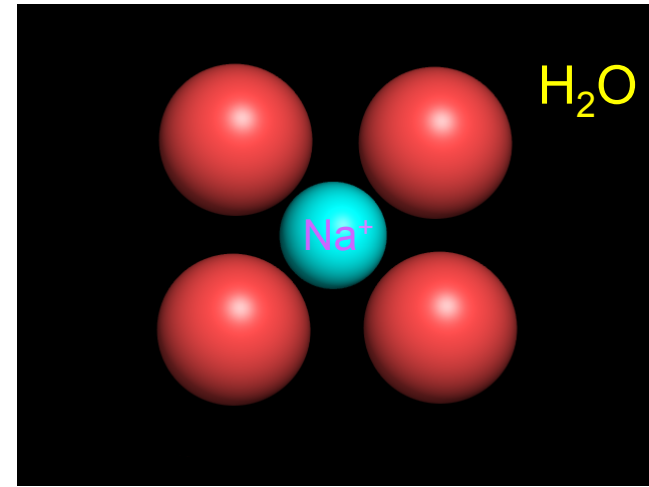
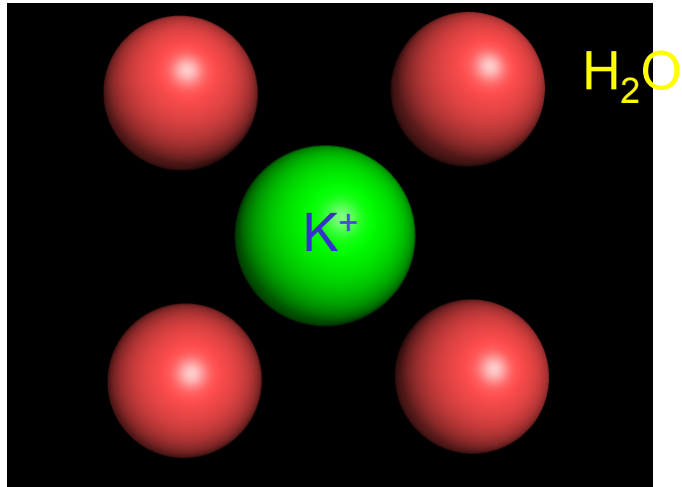
KcsA ALWWSVETATTVGYGDLYP

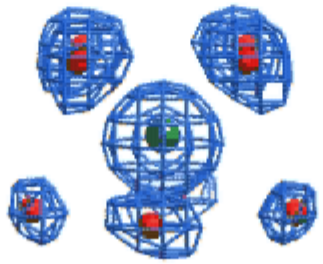
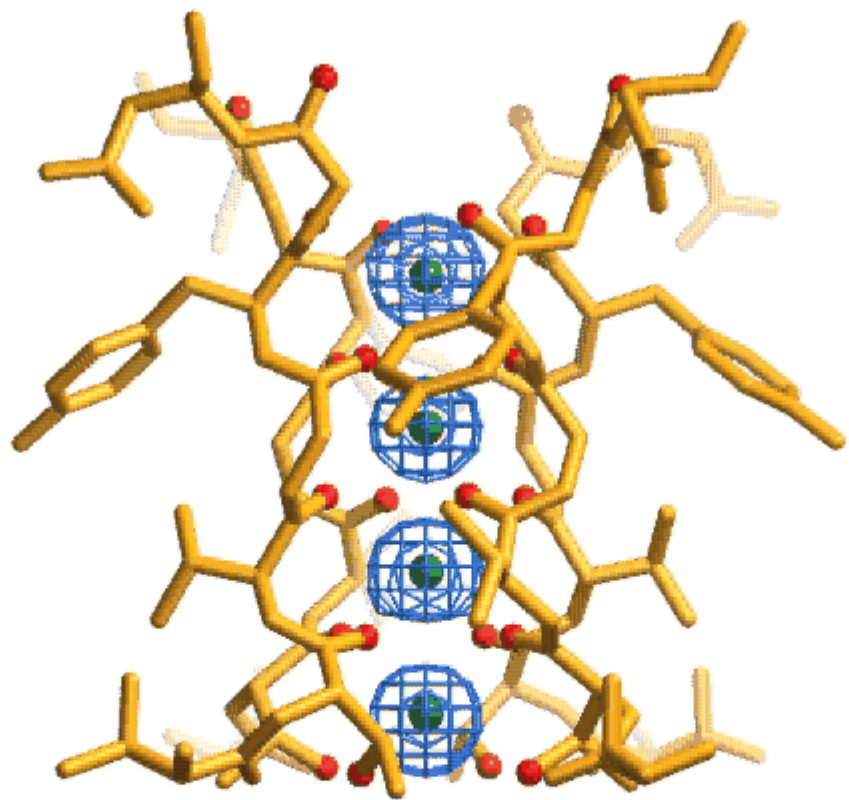
KvAP ALWWAVVTATTVGYGDVVP

Fly Kv

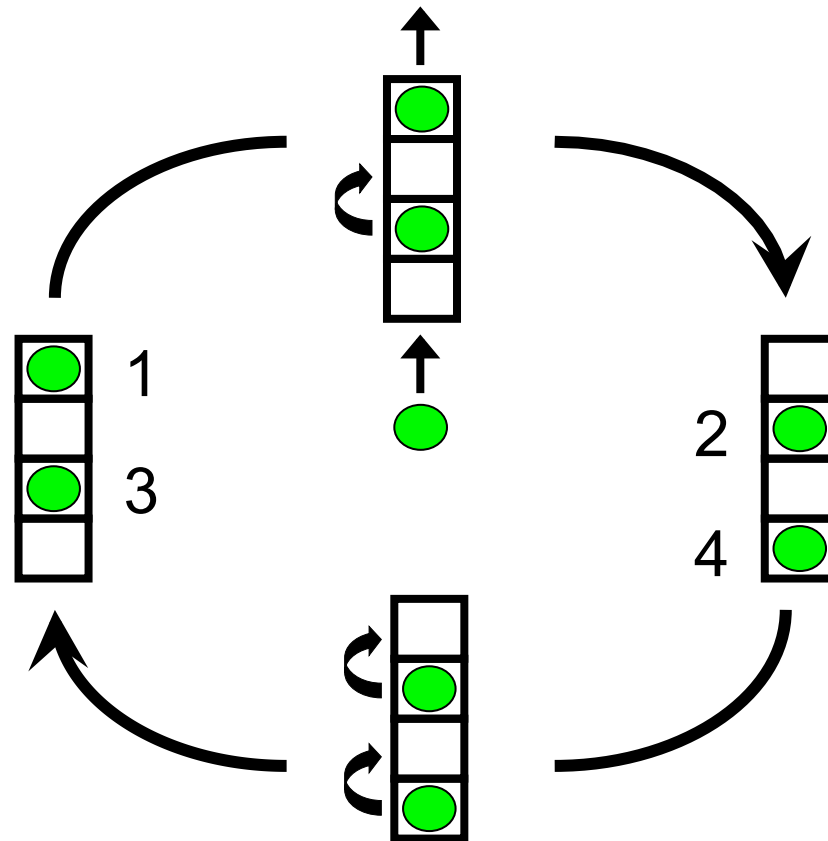
AFWWAVVTMTTVGYGDMTP

K⁺ channel selectivity filter mimics the hydration shell of a K ion

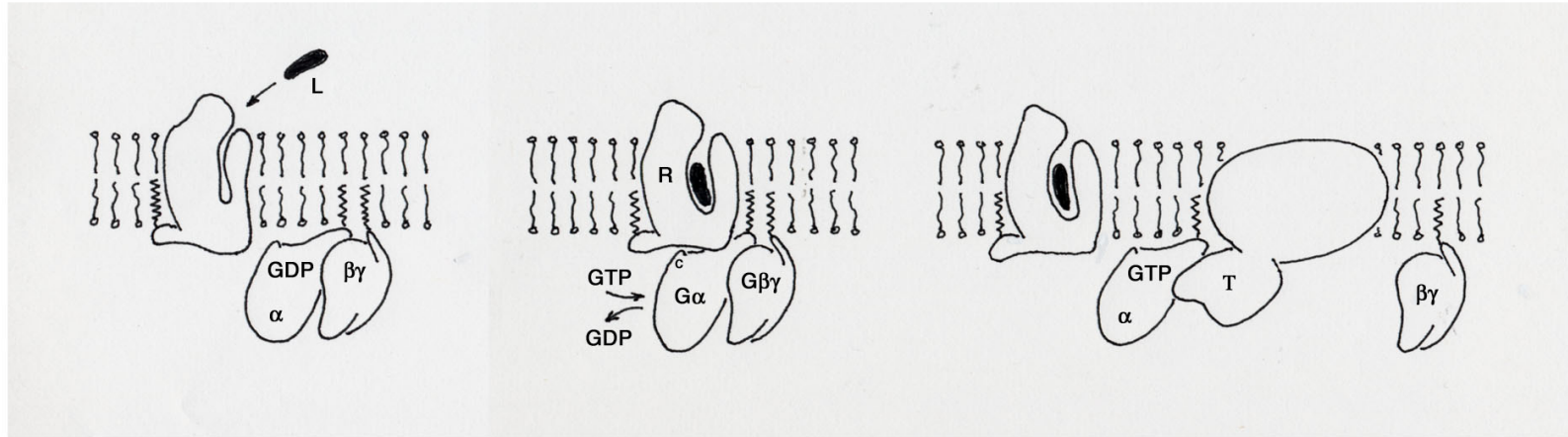




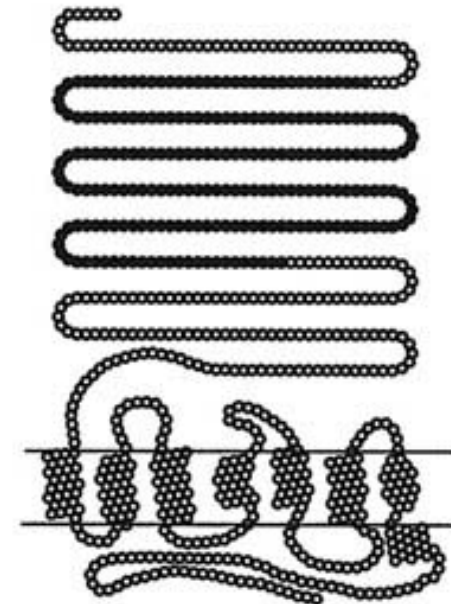
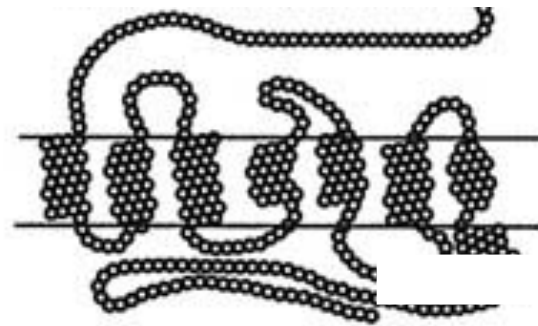
A simple through-put cycle



Receptors – G Protein Coupled Receptors (GPCRs)



- Photons
- Peptides
- Hormones
- Amino Acids
- Small molecules



Structural Studies on GPCRs

