

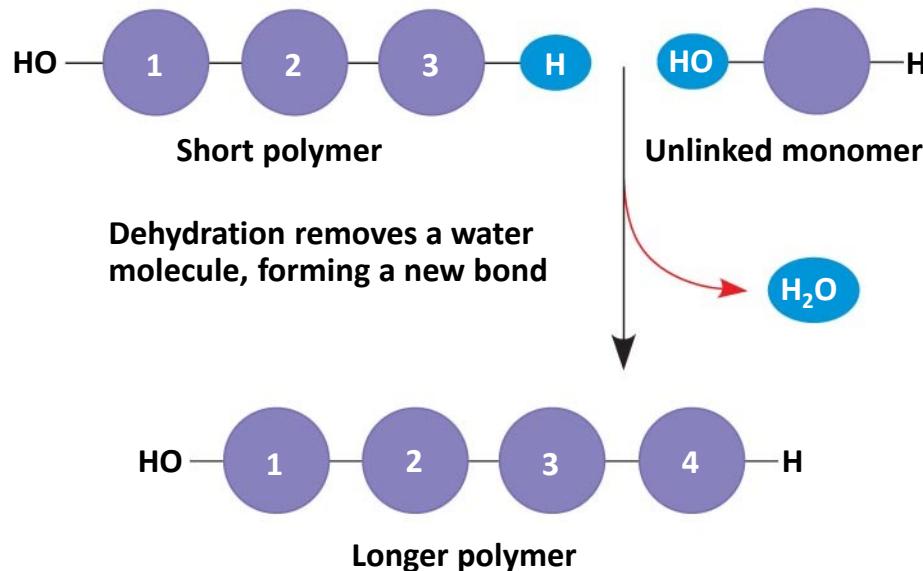
15/01/2023

Overview: The Molecules of Life

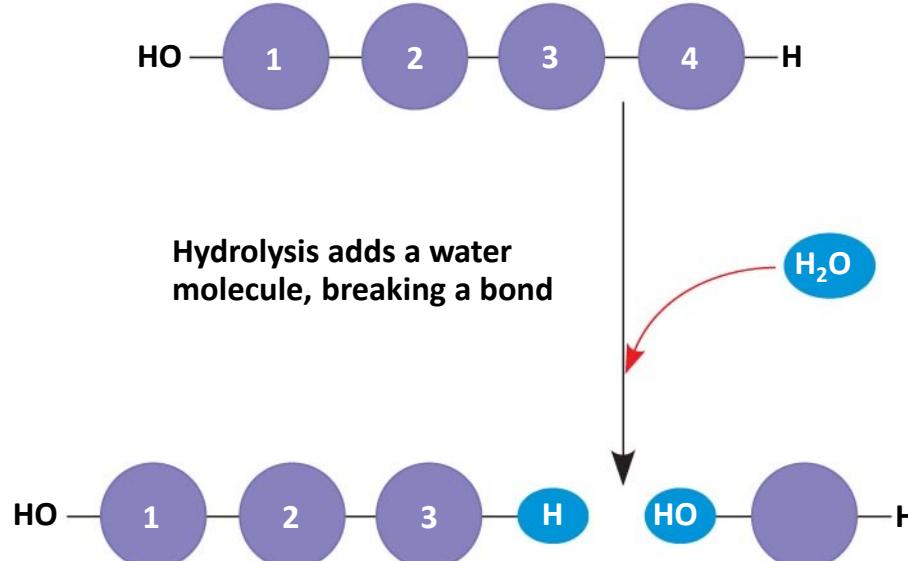
- All living things are made up of four classes of large biological molecules: **carbohydrates, lipids, proteins, and nucleic acids**
- Within cells, small organic molecules are joined together to form larger molecules
- **Macromolecules** are large molecules composed of thousands of covalently connected atoms
- Molecular structure and function are inseparable

The Synthesis and Breakdown of Polymers

- A **condensation reaction** or more specifically a **dehydration reaction** occurs when two monomers bond together through the loss of a water molecule
- **Enzymes** are macromolecules that speed up the dehydration process
- Polymers are disassembled to monomers by **hydrolysis**, a reaction that is essentially the reverse of the dehydration reaction



(a) Dehydration reaction in the synthesis of a polymer



(b) Hydrolysis of a polymer

The Diversity of Polymers

- Each cell has thousands of different kinds of macromolecules
- Macromolecules vary among cells of an organism, vary more within a species, and vary even more between species
- An immense variety of polymers can be built from a small set of monomers

Proteins do most of the work of the cell

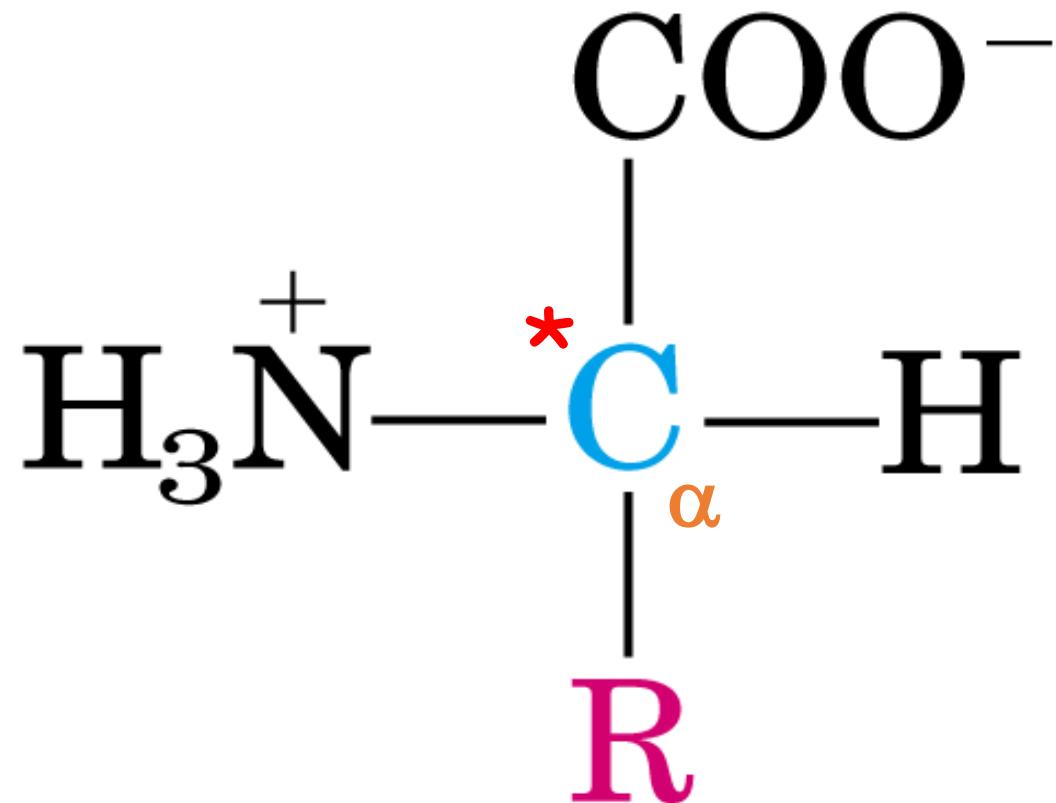
- Each protein is **specialized** to do a certain job.
- Some proteins are **structural**: control shapes of cells and bind cells together. E.g Keratin
- Chemical reactions of the cells are controlled by protein **enzymes**.
- Protein **pumps** move things across the cell membrane.
- Proteins give **mobility**: muscle, flagella, molecular motors

General properties of proteins

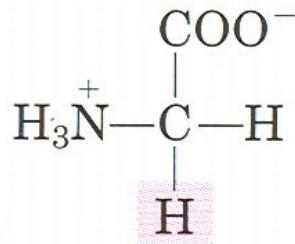
- Most abundant biomolecule; accounts for 50% of dry weight.
- Built by assembling long chains of amino acids (monomers), following by intricate folding.
- Final shape of protein is very specific. Unless correctly folded, is not functional
- Several 1000 different types of proteins in any cell; millions of protein molecules.
- All proteins are composed of 20 "standard" amino acids.

Amino Acids

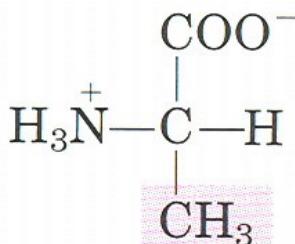
The common amino acids are known as α -amino acids because they have a primary amino group ($-\text{NH}_2$) as a substituent of the α carbon atom, the carbon next to the carboxylic acid group. The 20 standard amino acids differ in the structure of their side chains (R groups).



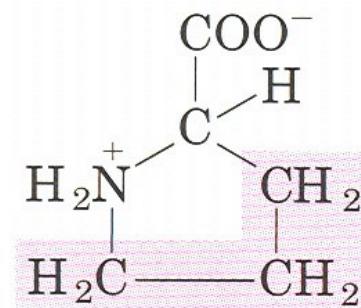
Nonpolar Aliphatic Amino Acids



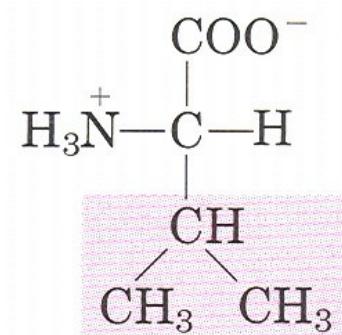
Glycine
Gly, G



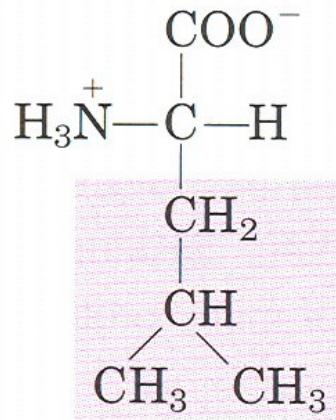
Alanine
Ala, A



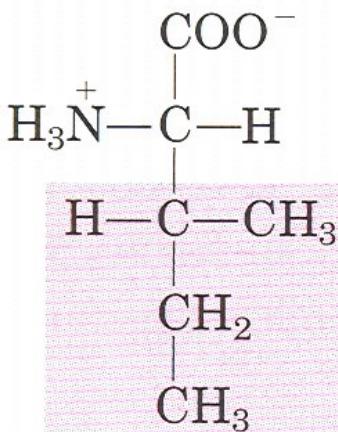
Proline
Pro, P



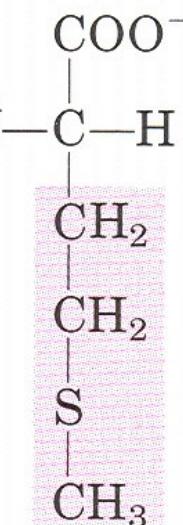
Valine
Val, V



Leucine
Leu, L

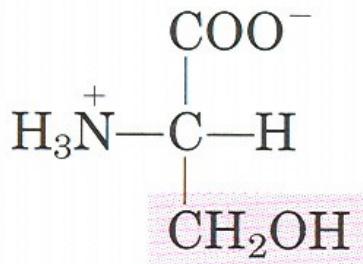


Isoleucine
Ile, I



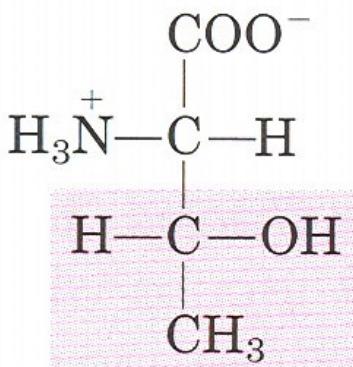
Methionine
Met, M

Polar Uncharged Amino Acids



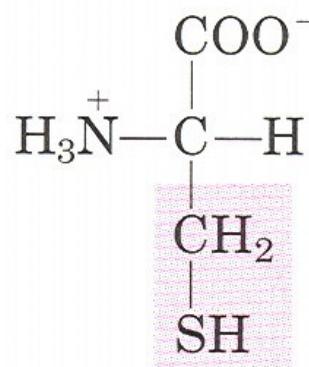
Serine

Ser, S



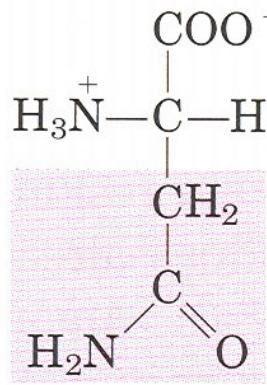
Threonine

Thr, T

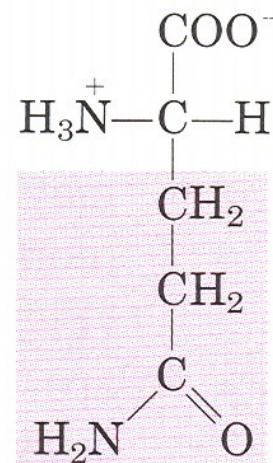


Cysteine

Cys, C

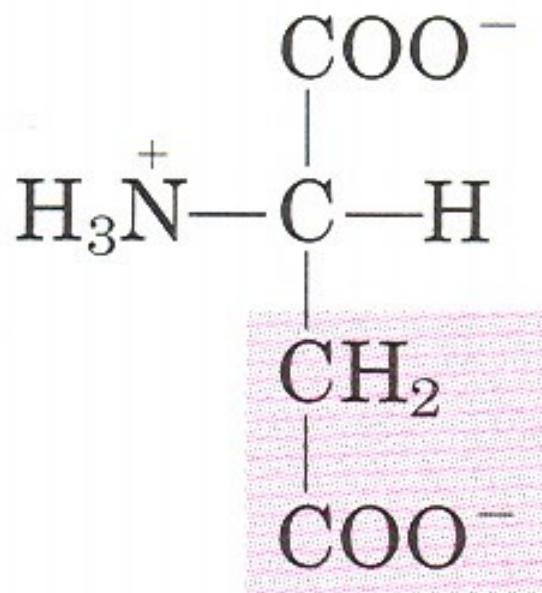


Asn, N Asparagine



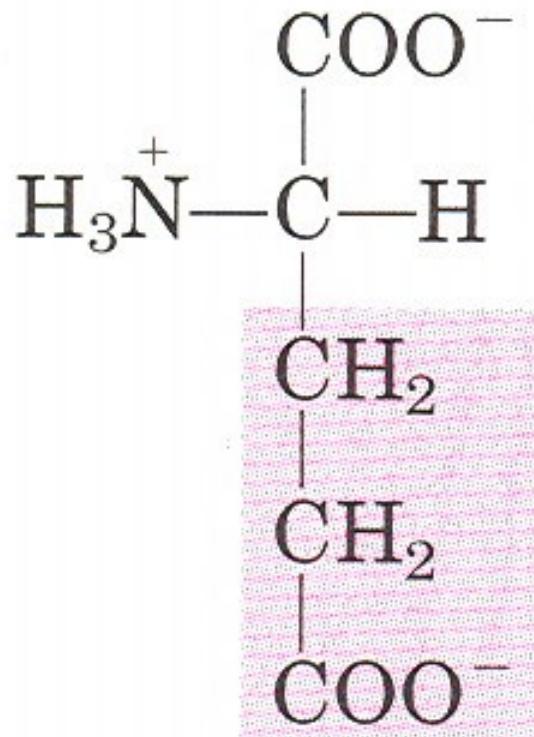
Glutamine *Gln, Q*

Negatively charged Amino Acids



Aspartate

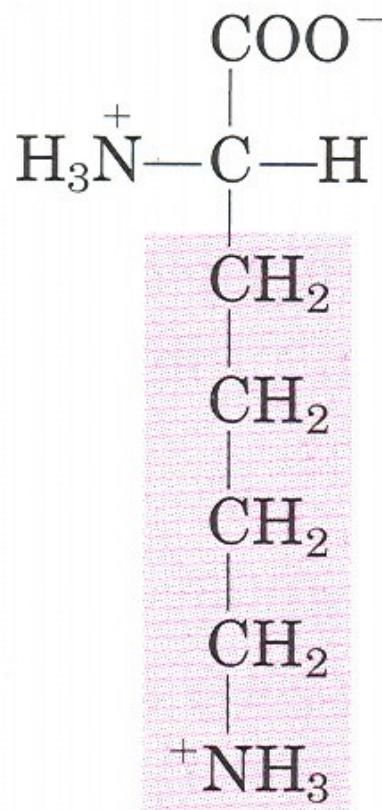
Asp, D



Glutamate

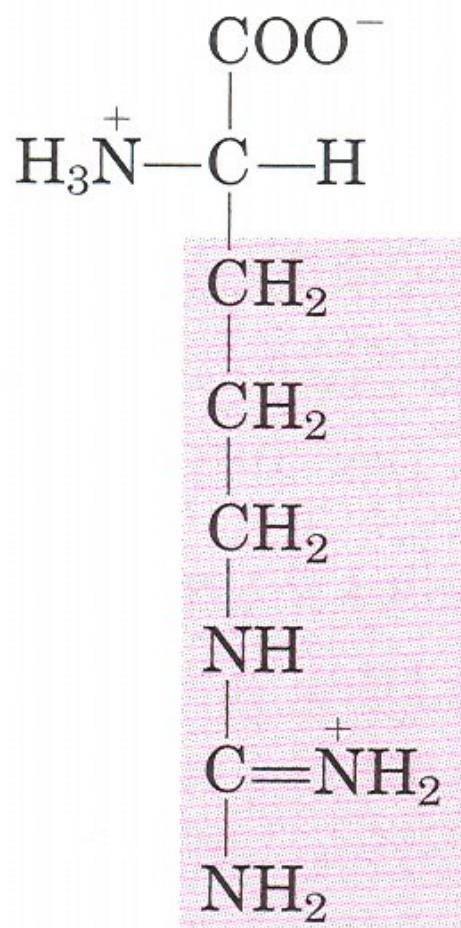
Glu, E

Positively charged Amino Acids



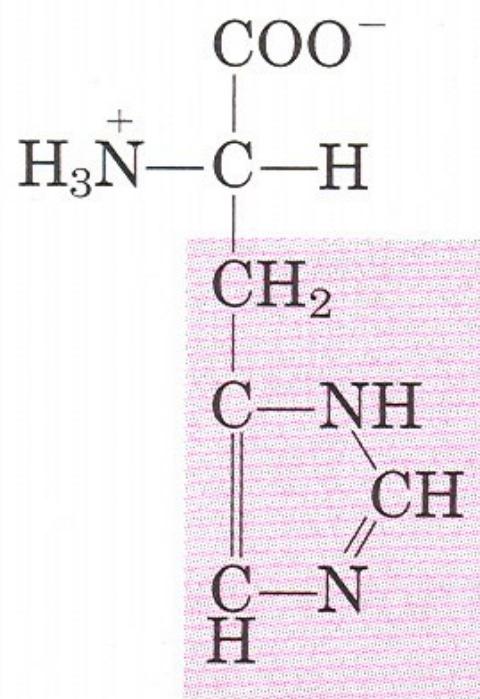
Lysine

Lys, K



Arginine

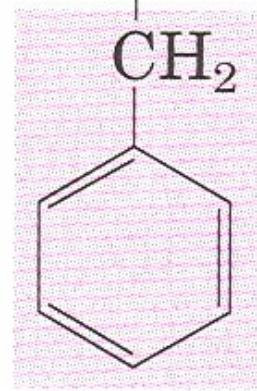
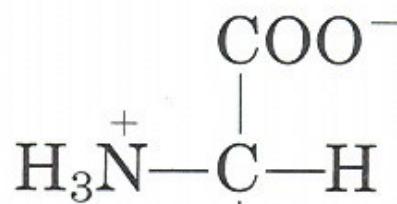
Arg, R



Histidine

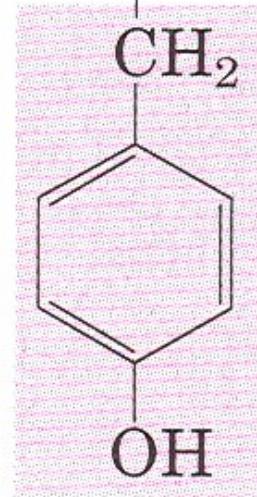
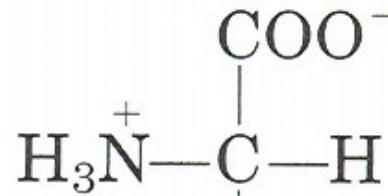
His, H

Aromatic Amino Acids



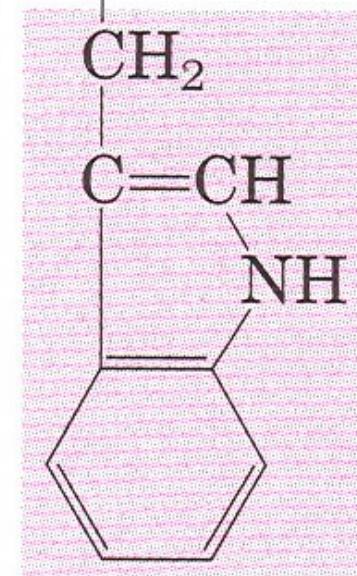
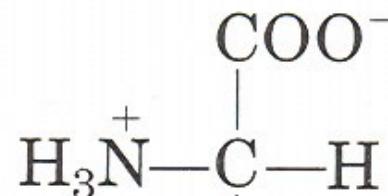
Phenylalanine

Phe, F



Tyrosine

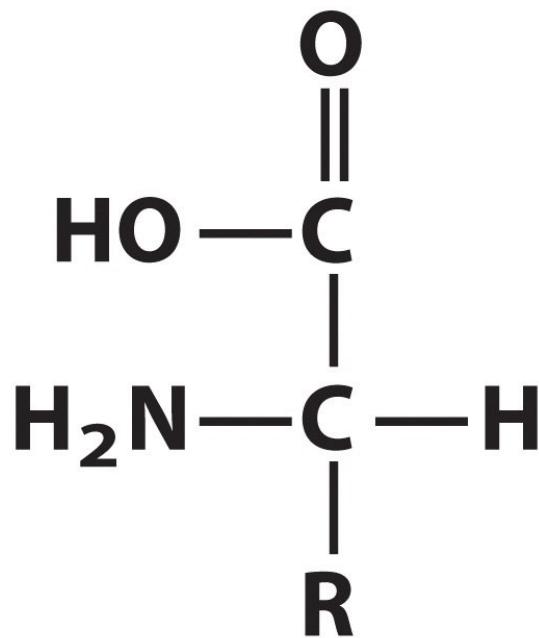
Tyr, Y



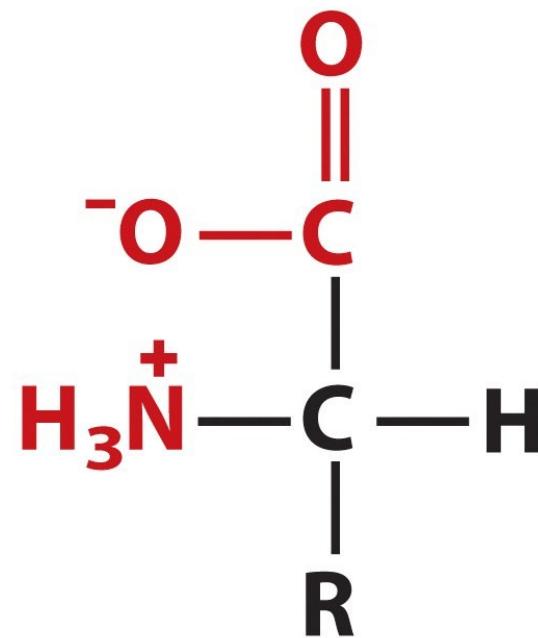
Tryptophan

Trp, W

Amino Acids are Zwitterions (Twin Ions) at Physiological pH



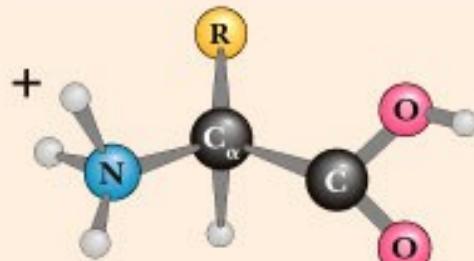
**Nonionic
form**



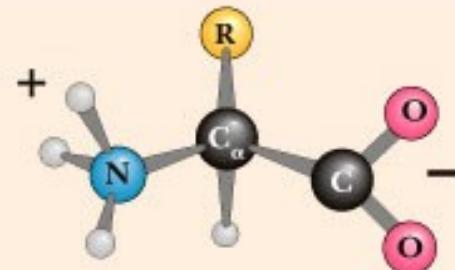
**Zwitterionic
form**

Ionic Forms of Amino Acids

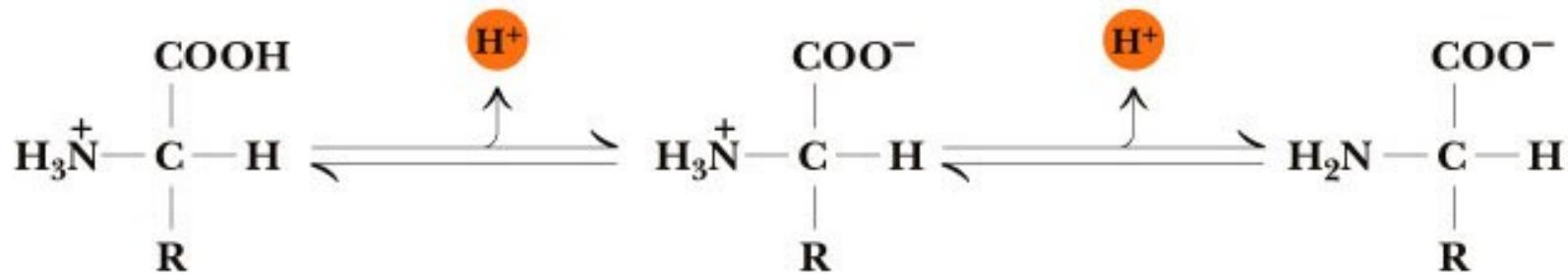
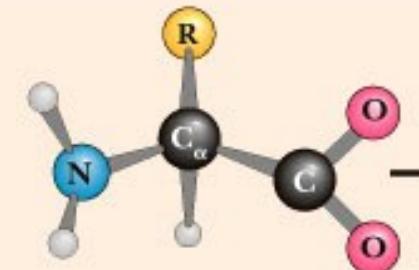
pH 1, Net charge +1



pH 7, Net charge 0



pH 13, Net charge -1



Cationic form

Zwitterion (neutral) form

Anionic form

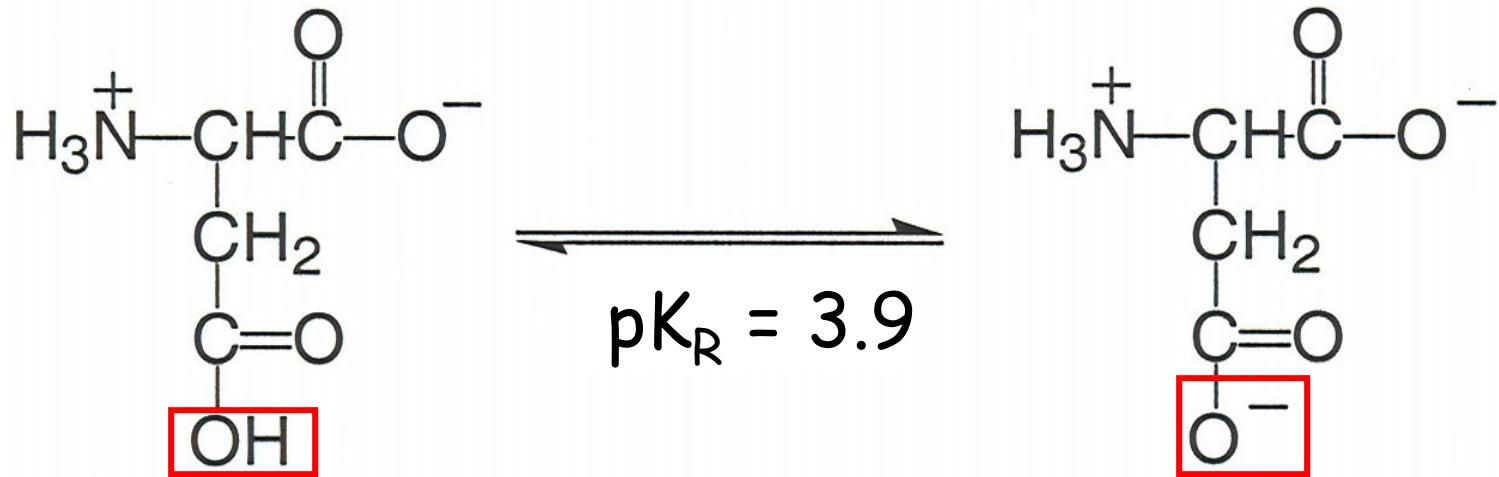
pK_a Values of the Amino Acids

You should know these numbers and
know what they mean!

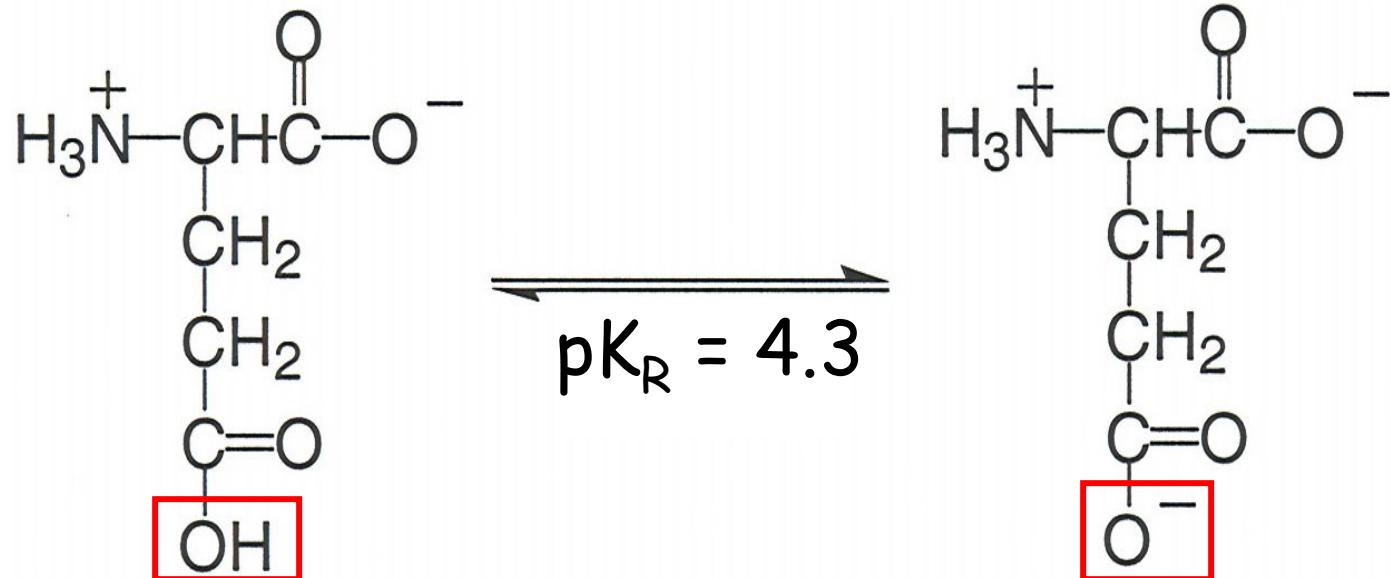
- Alpha carboxyl group: $pK_1 \approx 2$
- Alpha amino group: $pK_2 \approx 9$
- These numbers are approximate, but entirely suitable for our purposes.

Ionization of Side Chains: Acidic aa

asp

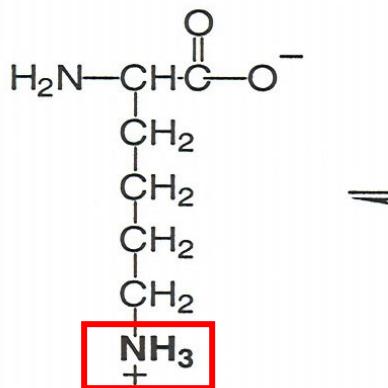


glu

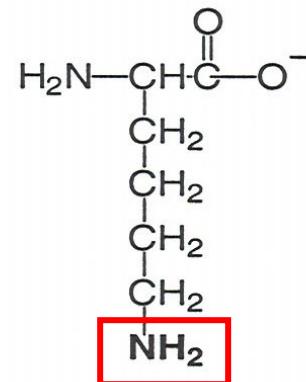


Ionization of Side Chains: Basic aa

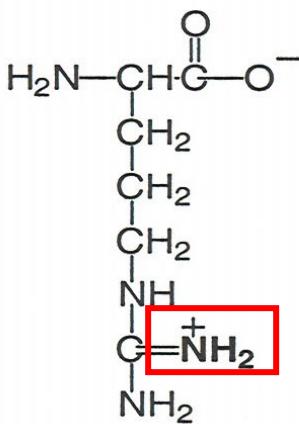
lys



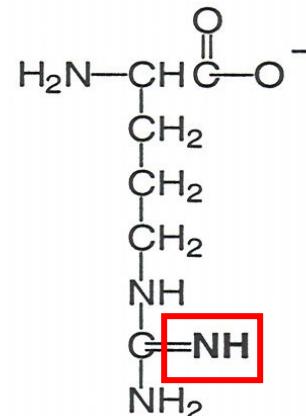
$$\text{pK}_R = 10.5$$



arg

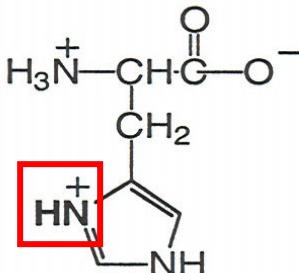


$$\text{pK}_R = 12.5$$

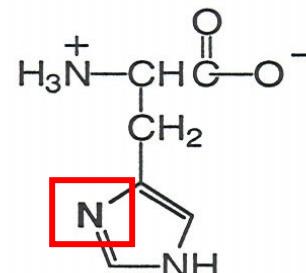


guanidinium

his



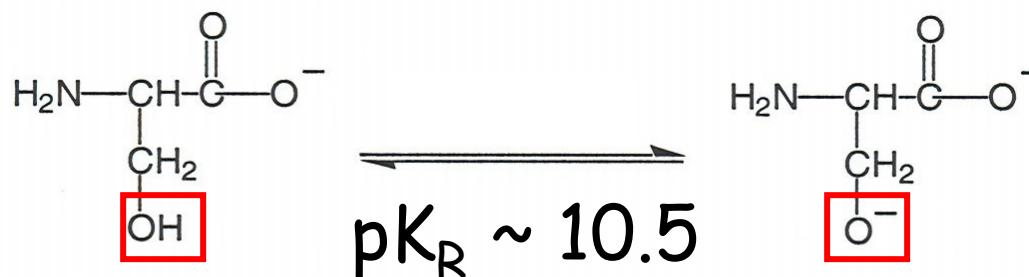
$$\text{pK}_R = 6.0$$



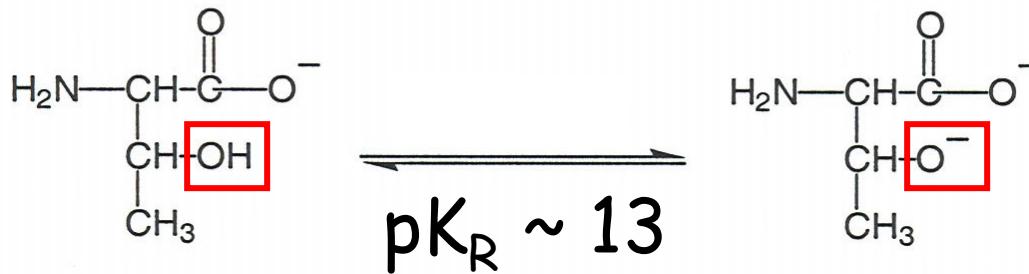
imidazolium

Ionization of Side Chains: -OH Groups

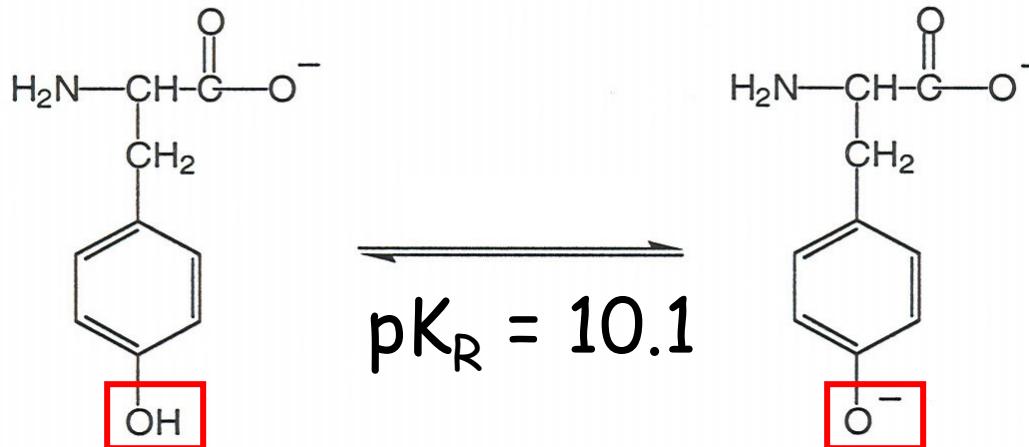
ser



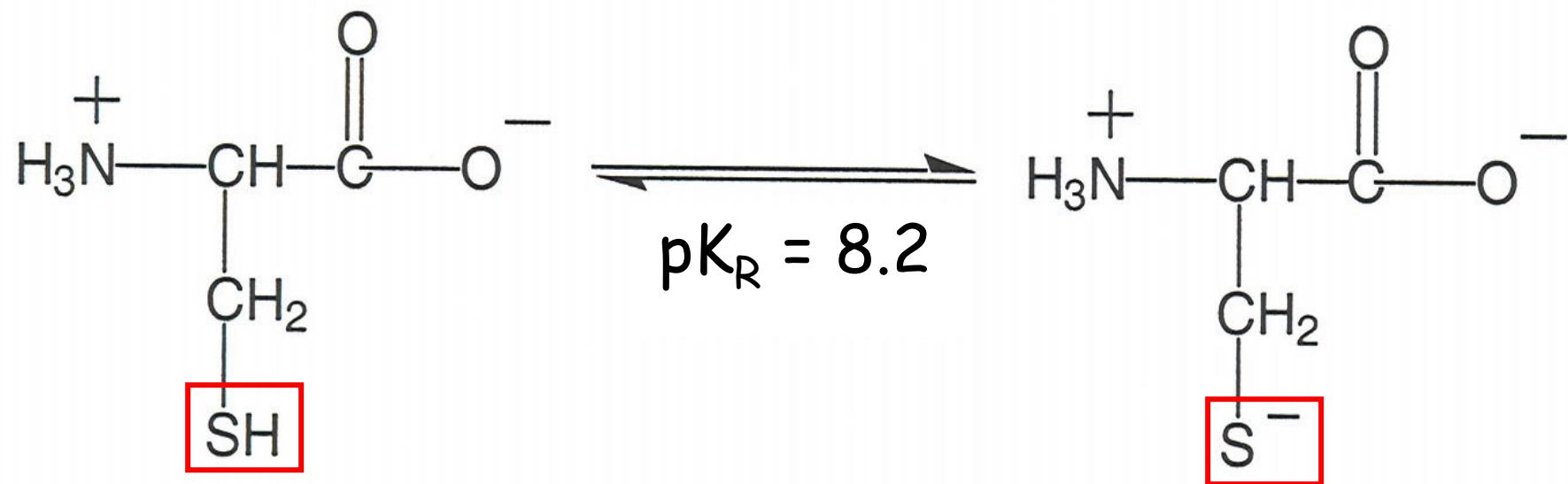
thr



tyr



Ionization of Side Chains: Cysteine



17/01/2023

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

Amino acid	Abbreviation/ symbol	M_r	pK_a values				Hydropathy index*	Occurrence in proteins (%)†
			pK_1 (—COOH)	pK_2 (—NH ₃ ⁺)	pK_R (R group)	pI		
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

*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.

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

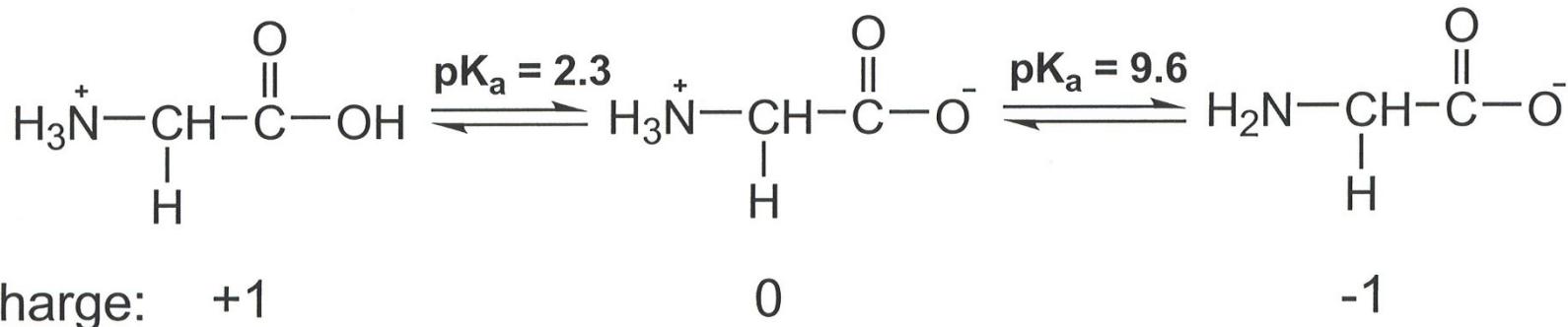
Amino acid	Abbreviation/ symbol	M_r	pK_a values				Hydropathy index*	Occurrence in proteins (%)†				
			pK_1 (—COOH)	pK_2 (—NH ₃ ⁺)	pK_R (R group)	pI						
Polar, uncharged												
R groups												
Serine	Ser S	105	2.21	9.15		5.68	-0.8	6.8				
Threonine	Thr T	119	2.11	9.62		5.87	-0.7	5.9				
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.

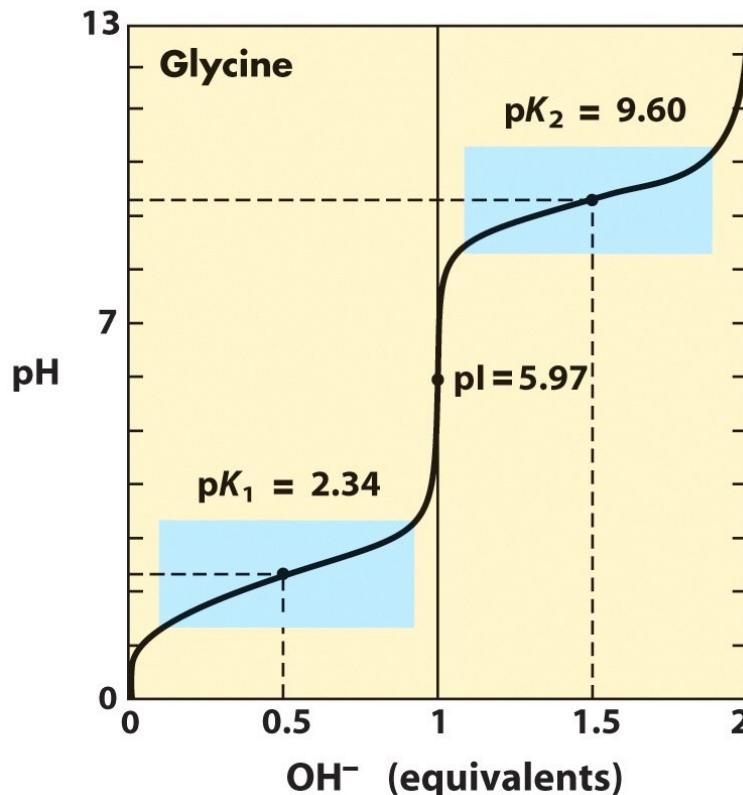
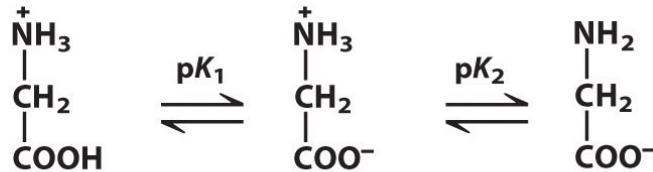
Ampholites and Isoelectric Point

Ampholites are molecules that contain both acidic and basic pK_a values (e.g. amino acids).



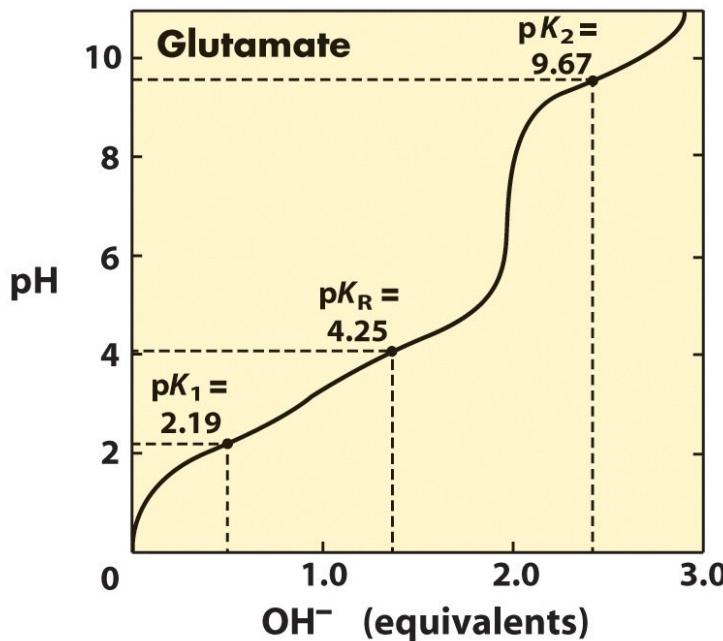
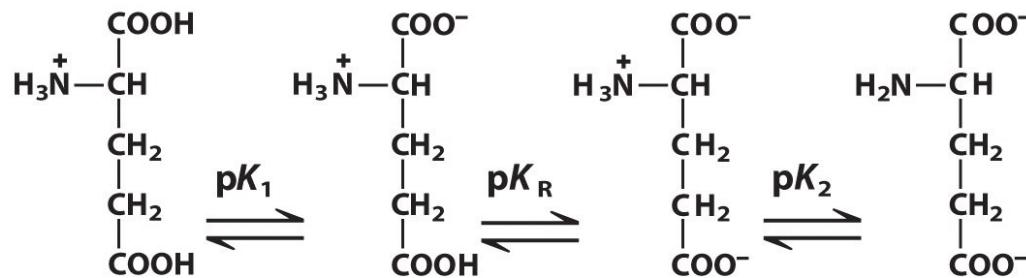
The **isoelectric point (pI)** is the pH where the average charge is zero. The pI lies **midway** between the two pK_a s values that indicate the protonation and deprotonation of the electrically neutral form.

Titration of Glycine



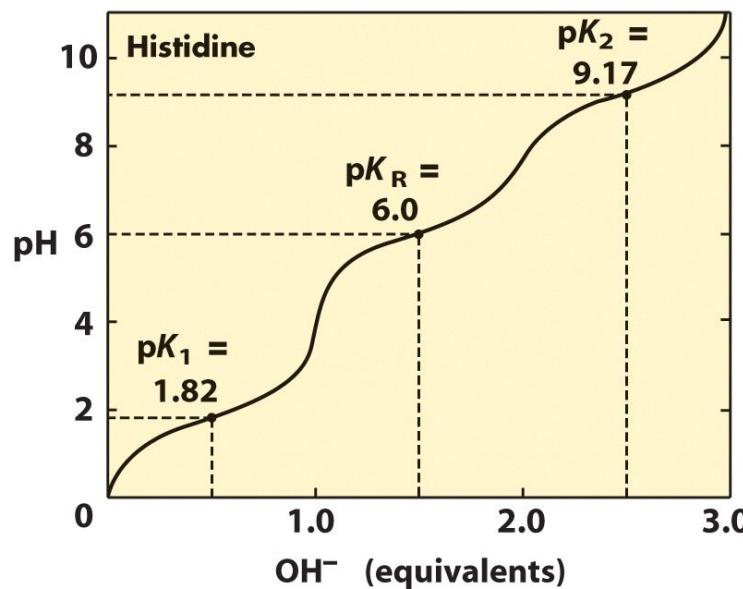
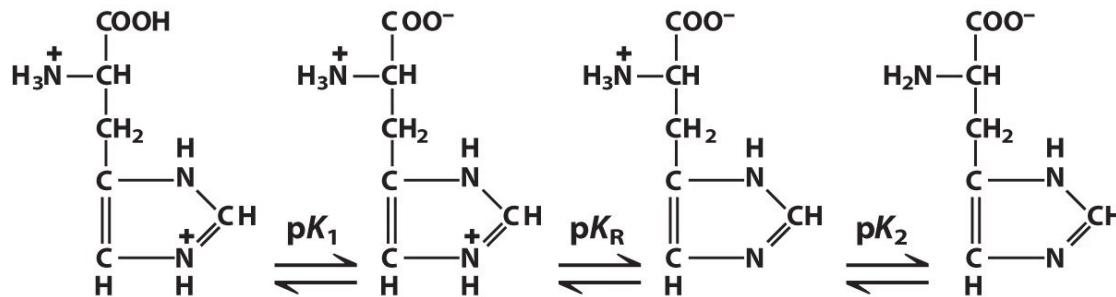
$$pI = \frac{pK_1 + pK_2}{2} = \frac{2.34 + 9.60}{2} = 5.97$$

Titration of Glutamic Acid



$$pI = \frac{pK_{\alpha-\text{COOH}} + pK_{R-\text{COOH}}}{2} = \frac{2.2 + 4.3}{2} = 3.3$$

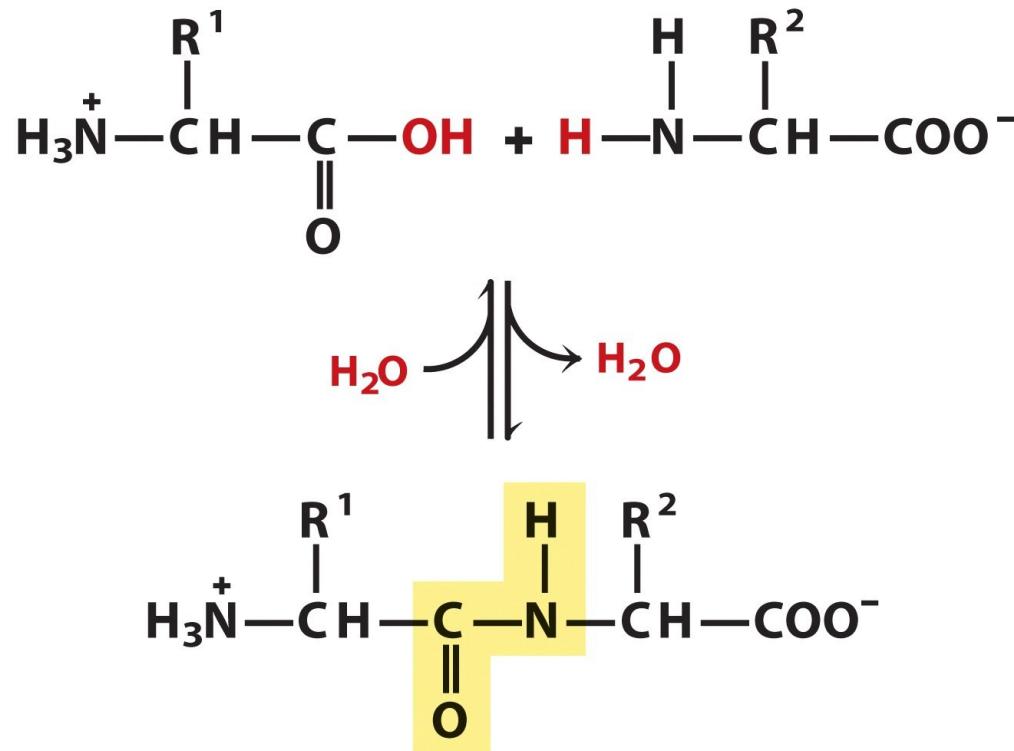
Titration of Histidine



$$pI = \frac{pK_{\alpha-\text{NH}_2} + pK_{R-\text{NH}_2}}{2} = \frac{9.17 + 6.0}{2} = 7.59$$

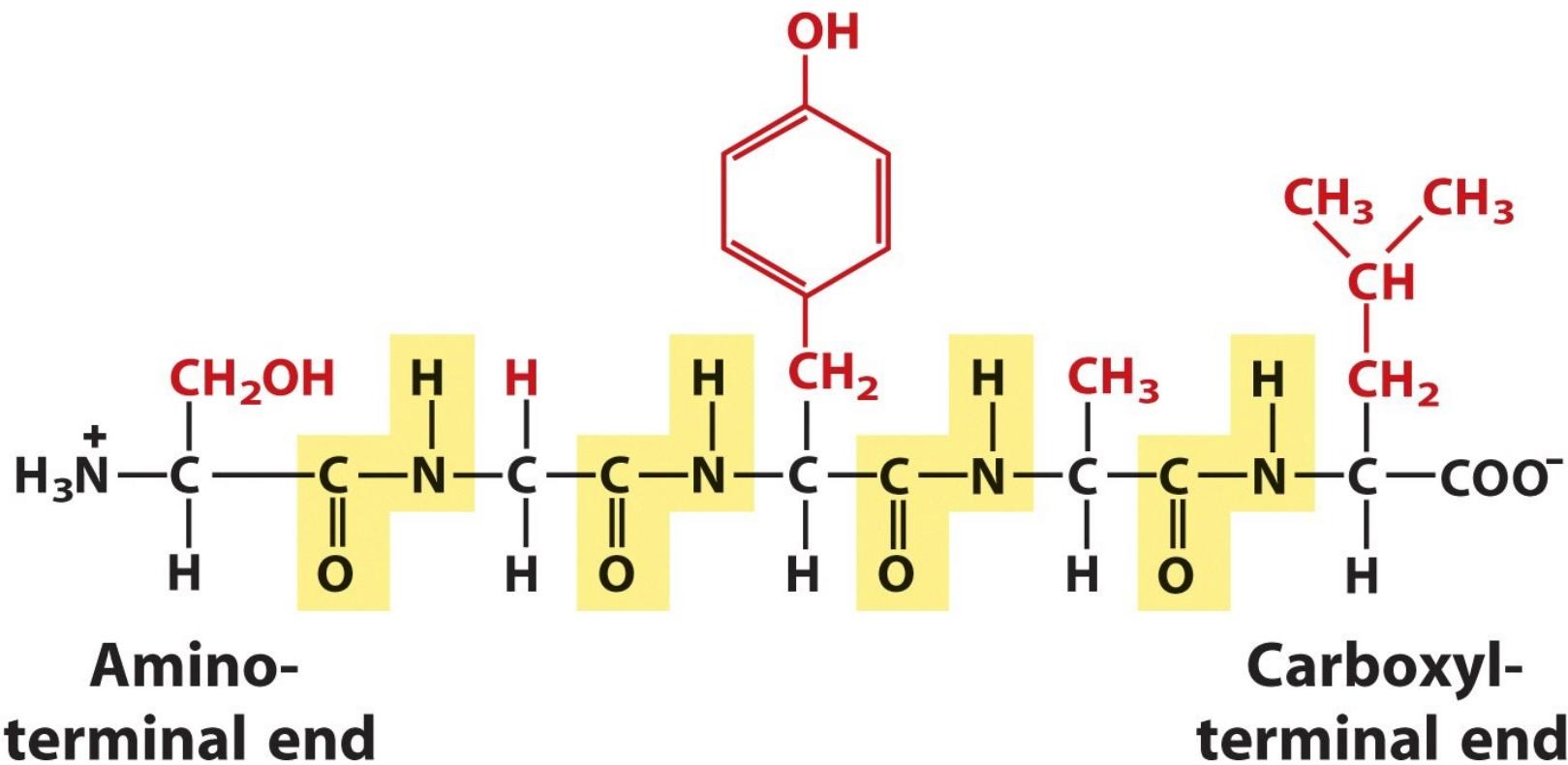
Peptides and proteins

Peptides are chains of amino acids



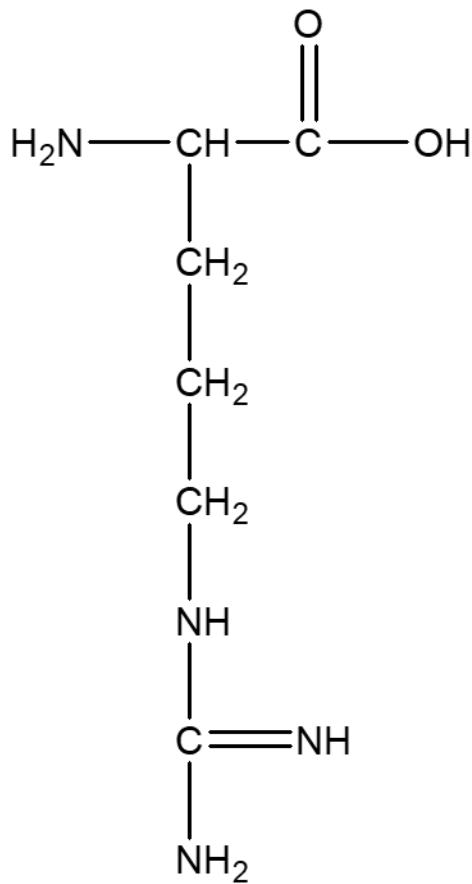
Peptide Bond Formation

Ser-Gly-Tyr-Ala-Lue or SGYAL

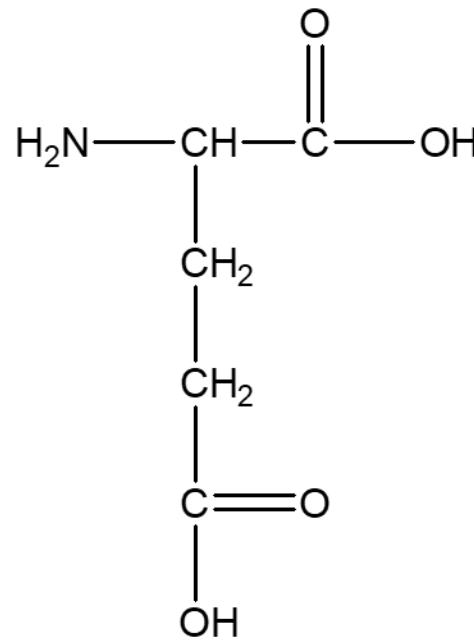


Key convention: display the amino terminal on the left and the carboxyl terminal on the right. The sequence is read left to right.

How many dipeptides can form two of the following AA?

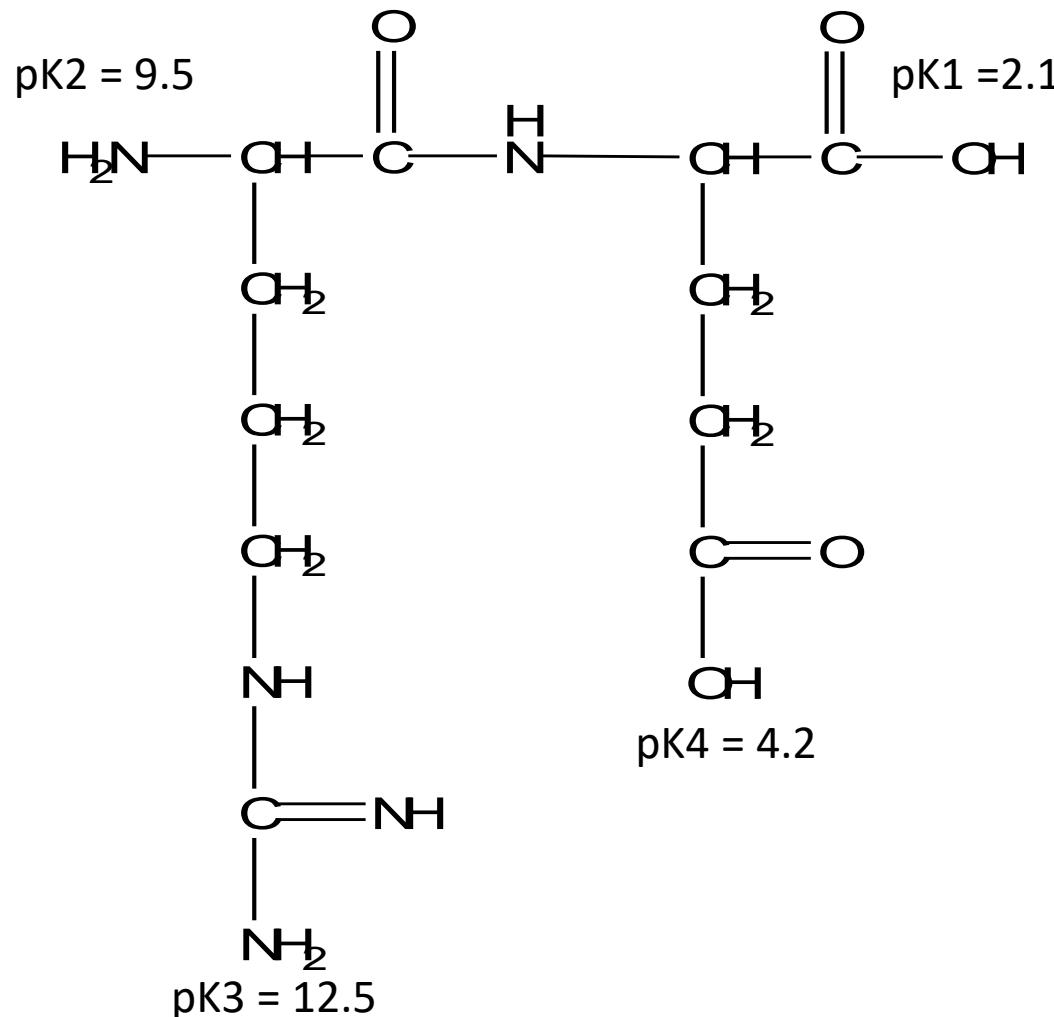


Arg



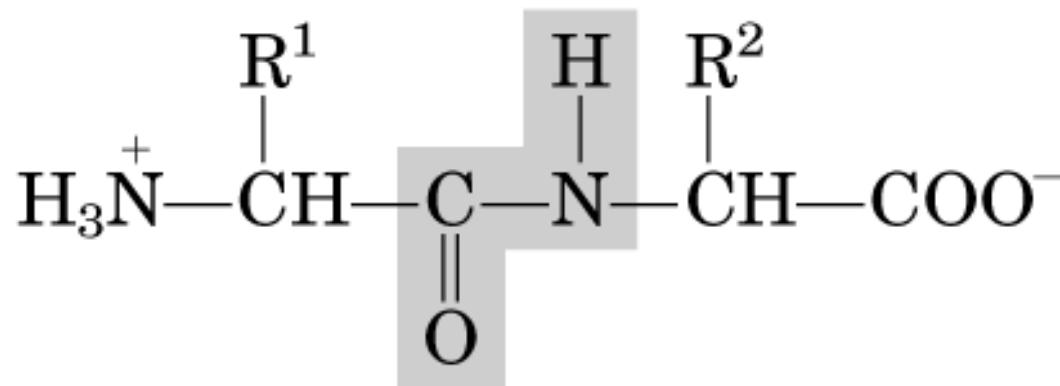
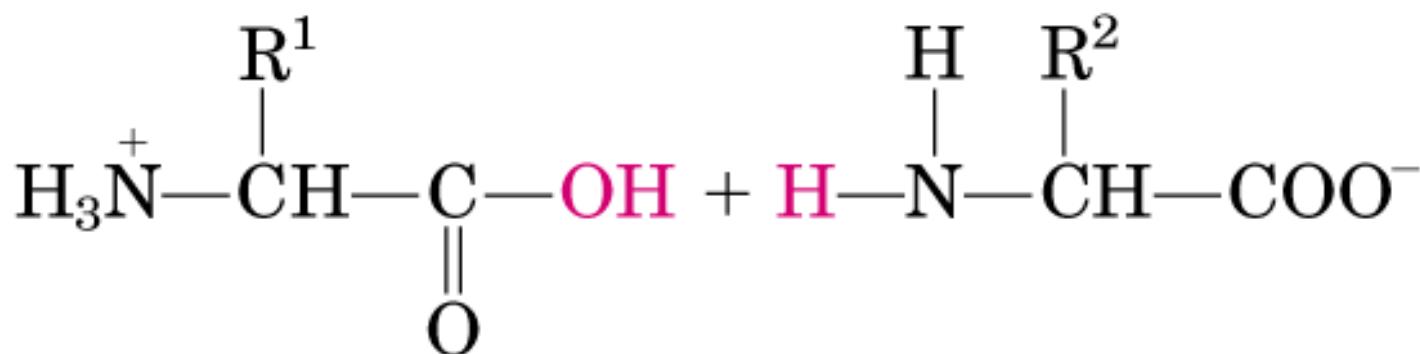
Glu

What are different ionic forms of the dipeptide below in different pHs? What is the pI value of this dipeptide?



19/01/2023

Formation of Peptide Bond



Favorable Interactions in Proteins

Electrostatic interactions

Dipolar interactions

 Dipole-dipole interactions

 Dipole-induced dipole interactions

 Charge-dipole interactions

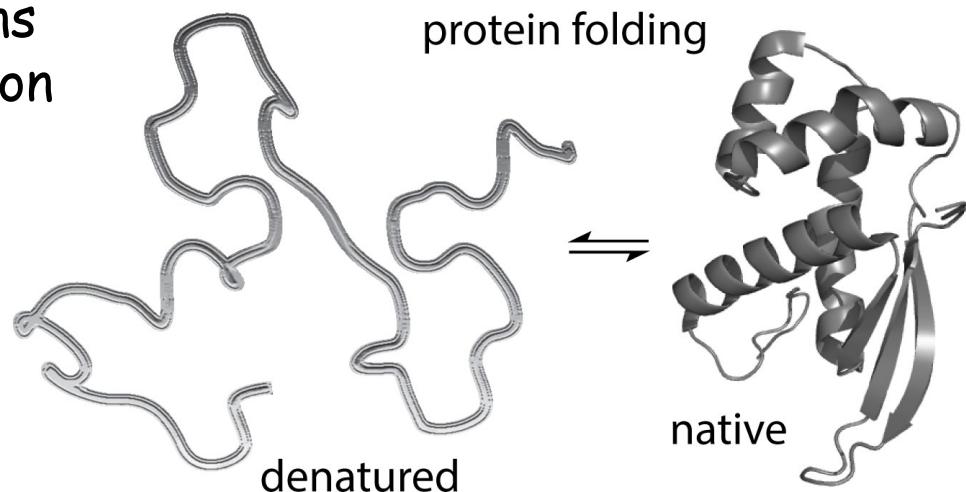
 Fluctuating dipoles (London Dispersion)

Cation- π interactions

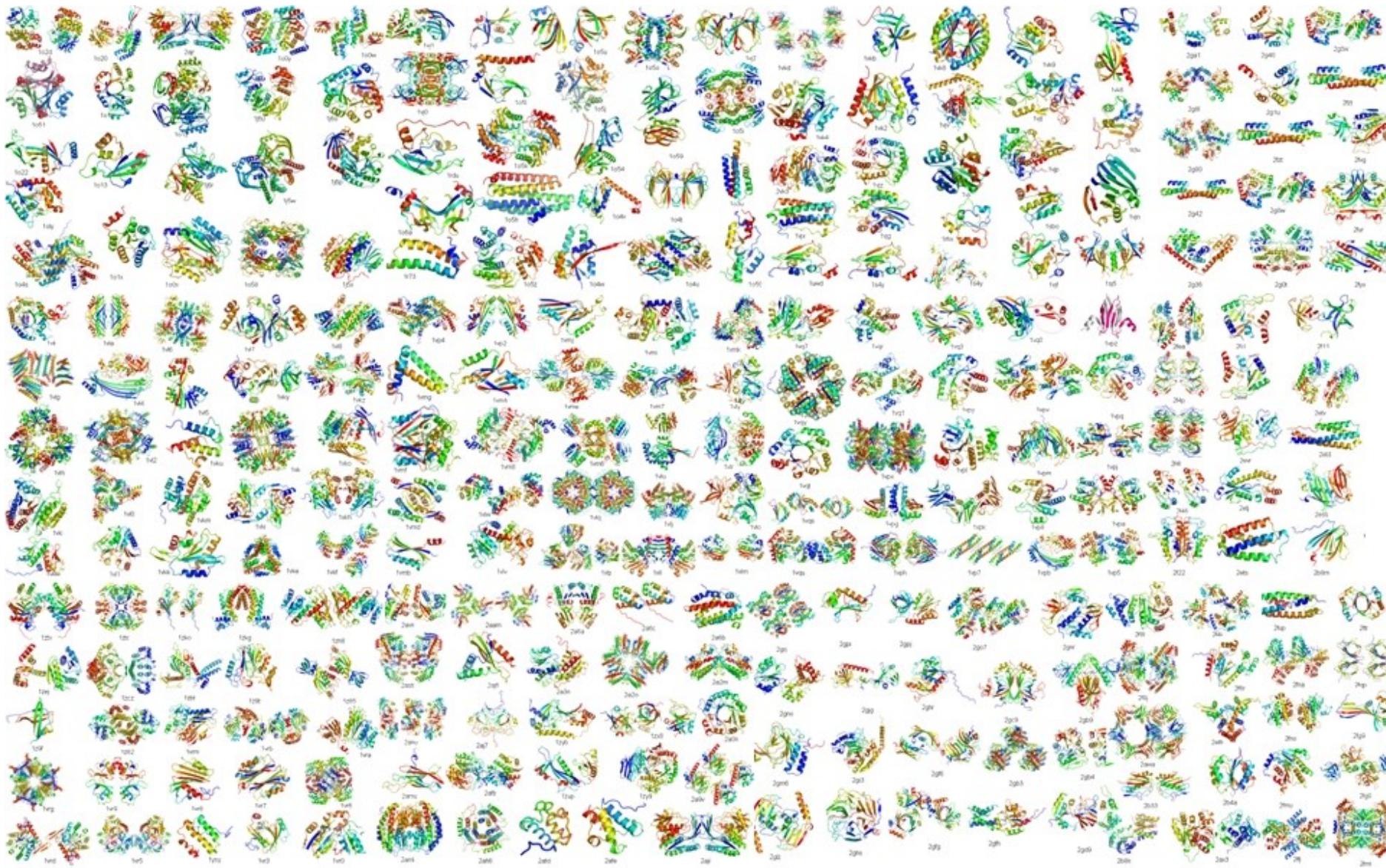
Hydrogen bonding

The hydrophobic effect

Short range repulsion



Proteins are diverse in nature



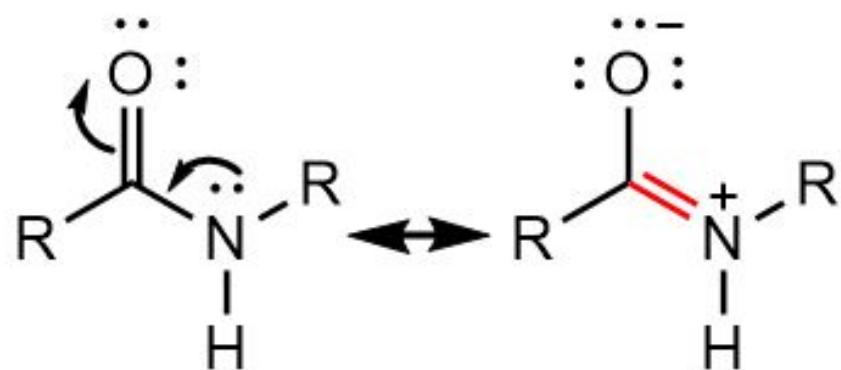
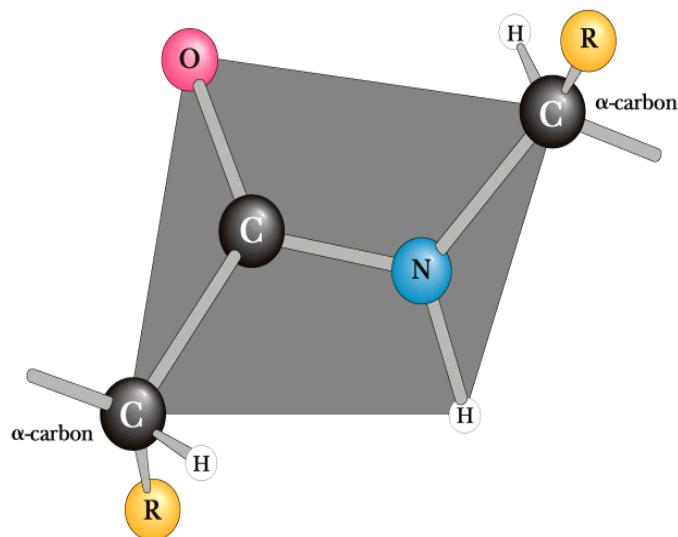
Biologically active peptides and Polypeptides occur in a vast range of sizes and compositions

TABLE 3-2 Molecular Data on Some Proteins

	Molecular weight	Number of residues	Number of polypeptide chains
Cytochrome c (human)	13,000	104	1
Ribonuclease A (bovine pancreas)	13,700	124	1
Lysozyme (chicken 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	609	1
Hexokinase (yeast)	102,000	972	2
RNA polymerase (<i>E. coli</i>)	450,000	4,158	5
Apolipoprotein B (human)	513,000	4,536	1
Glutamine synthetase (<i>E. coli</i>)	619,000	5,628	12
Titin (human)	2,993,000	26,926	1

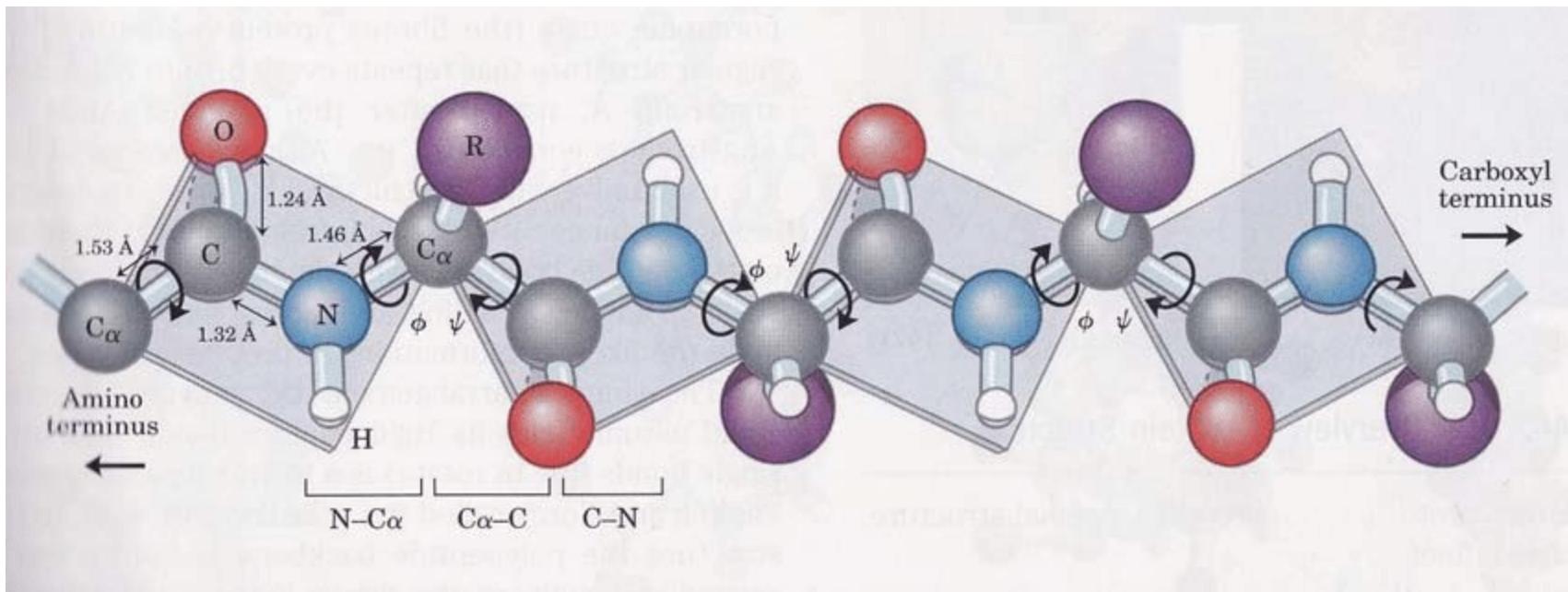
The Coplanar Nature of the Peptide Bond

Six atoms of the peptide group lie in a plane!



The Coplanar Nature of the Peptide Bond

Six atoms of the peptide group lie in a plane!



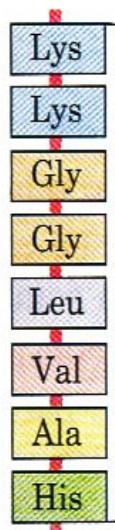
Dihedral angles around C_α-C and N-C_α bonds are named as ψ and ϕ by a common convention. Their values change between -180° and 180° .

Summary of the Peptide Bond

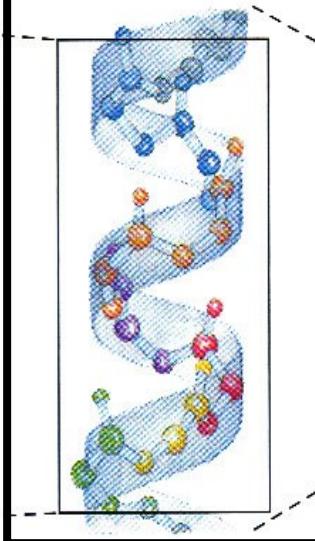
- is usually found in the *trans* conformation
- has partial (40%) double bond character
- due to the double bond character, the six atoms of the peptide bond group are always planar!
- peptide bond is polar: N partially positive; O partially negative

The Four Levels of Protein Structure

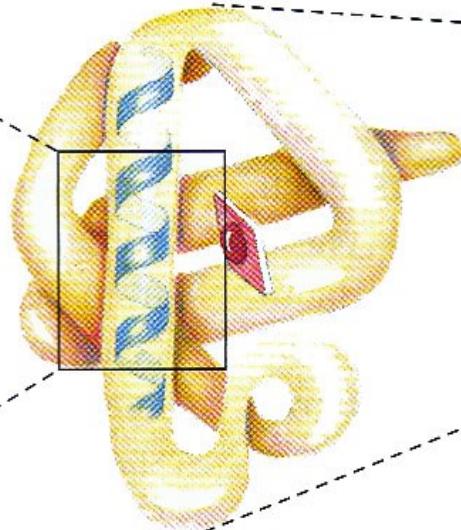
Primary structure



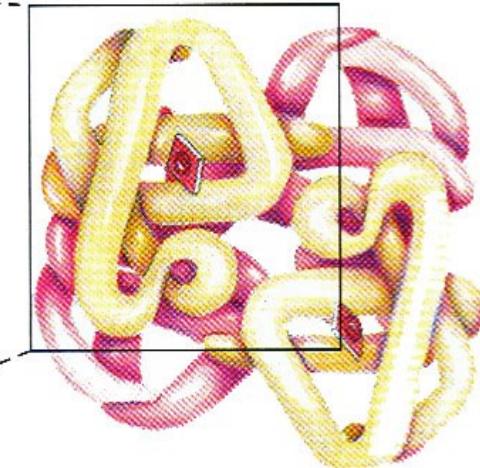
Secondary structure



Tertiary structure



Quaternary structure



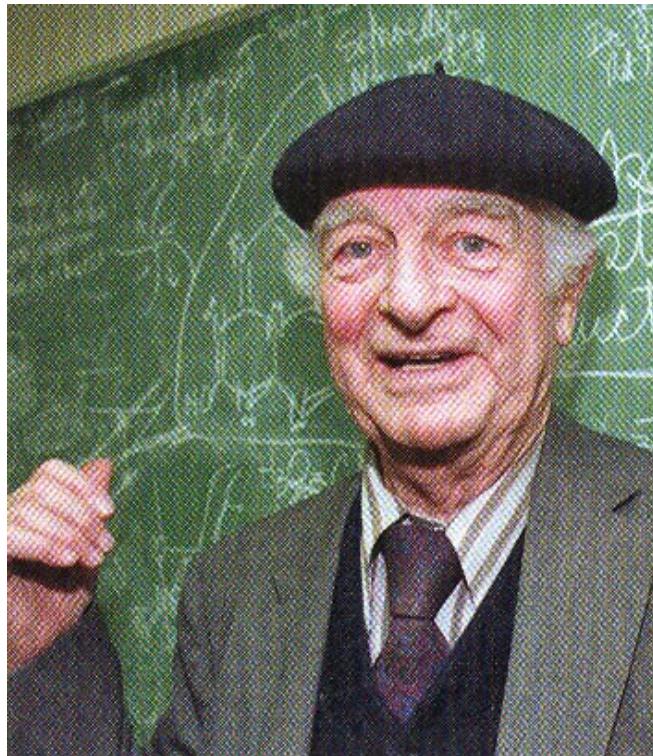
Secondary structure is the local spatial arrangement of a polypeptide's backbone atoms without regard to the conformation of its side chains.

Classes of Secondary Structure

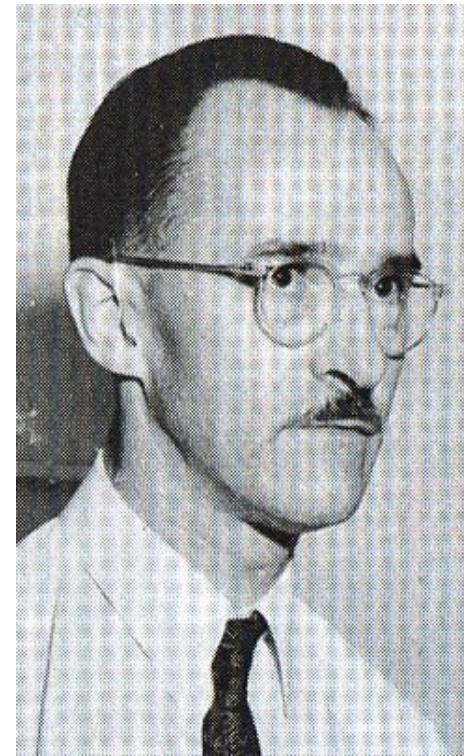
All these are local structures that are
stabilized by hydrogen bonds

- Alpha helix
- Beta sheet
- Beta turns

Papers by Linus Pauling and Robert Corey in 1951 proposed the α -helix and the β -sheet, now known to form the backbones of tens of thousands of proteins.



Linus Pauling (1901-1994), winner of the Nobel Prize in Chemistry in 1954 and the Nobel Peace Prize in 1962.



Robert Corey (1897-1971)

The Alpha Helix

- First proposed by Linus Pauling and Robert Corey in 1951
- Identified in keratin by Max Perutz
- A ubiquitous component of proteins
- Stabilized by H-bonds

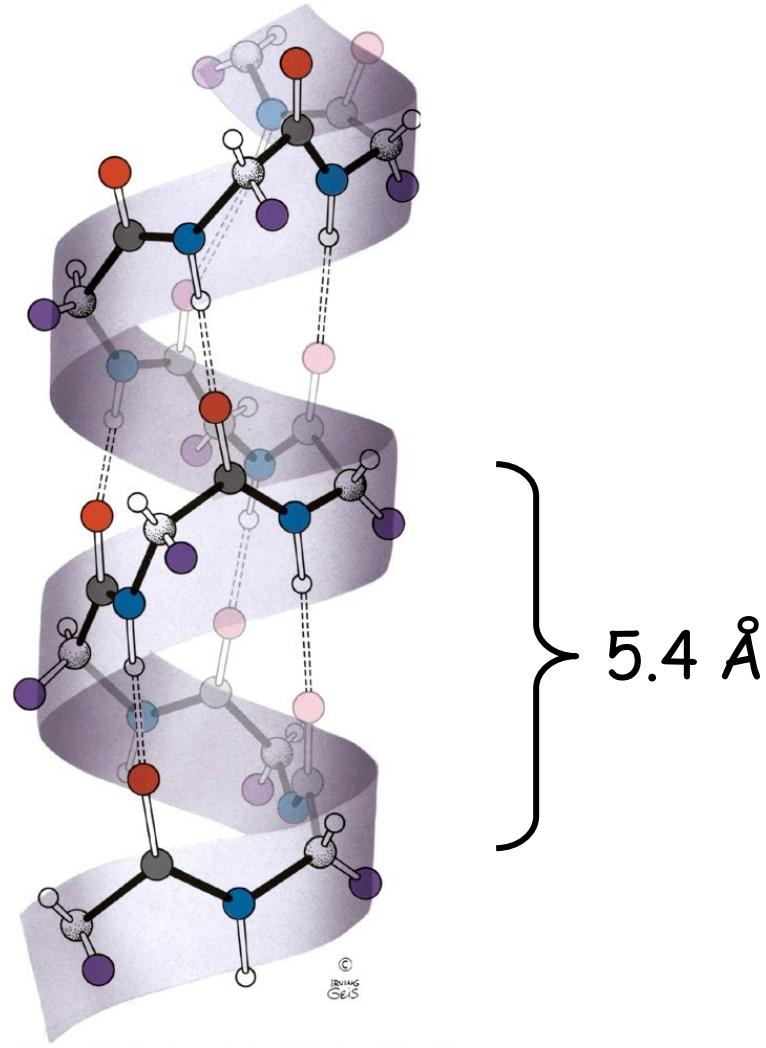
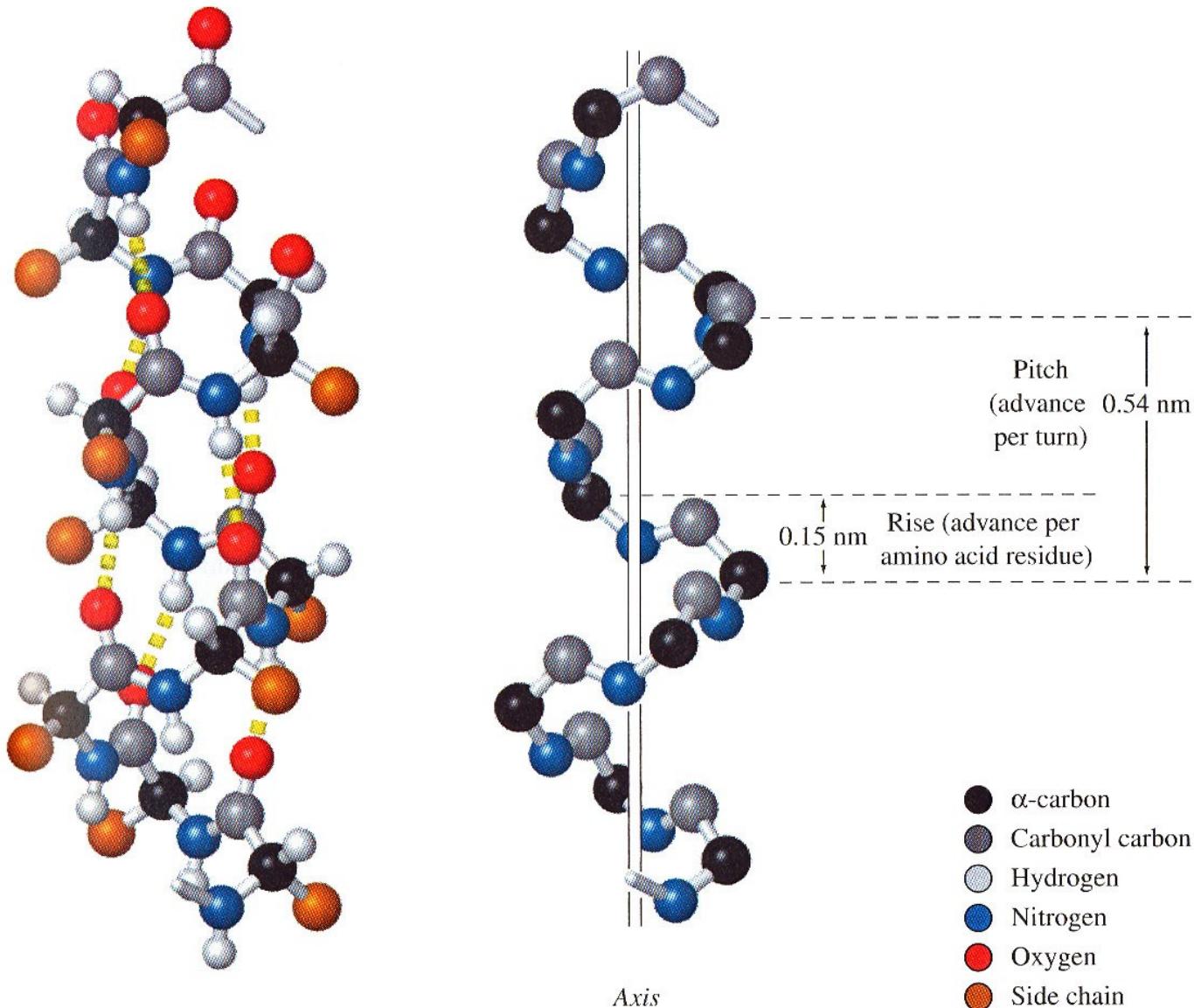


Figure 6-7 Fundamentals of Biochemistry, 2/e

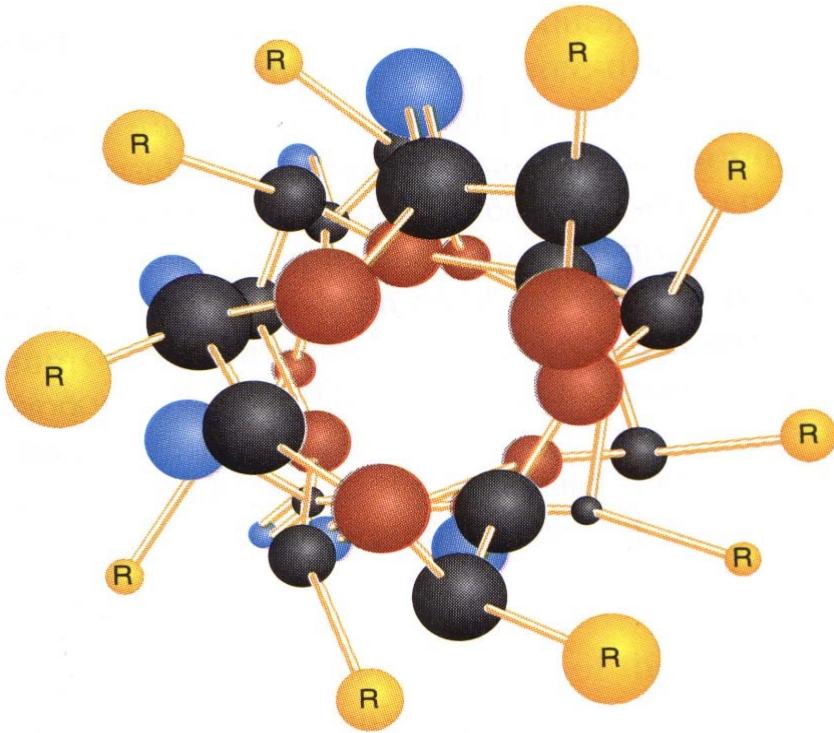
The Alpha-Helix

- Residues per turn: **3.6**
- The non-integral number of residues per turn was a surprise to crystallographers
- Rise per residue: 1.5 \AA
- Rise per turn (pitch): $3.6 \times 1.5 \text{ \AA} = 5.4 \text{ \AA}$

Right-handed Alpha Helix

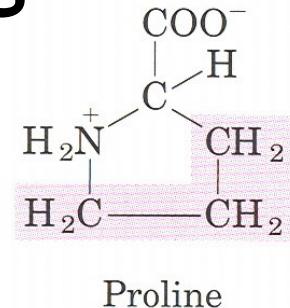


Top View of the Alpha-Helix



The R groups point away from the long axis of the helix

Amino acid sequence affects Alpha-Helix stability

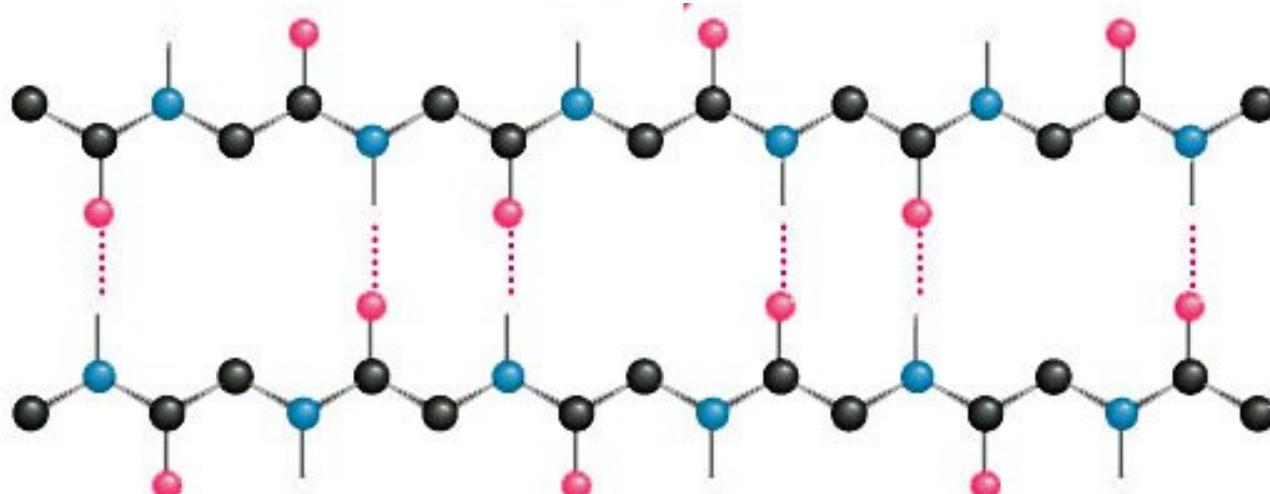


- **proline** creates a bend because
 1. the restricted rotation due to its cyclic structure
 2. its α -amino group has no N-H for hydrogen bonding
- strong electrostatic repulsion caused by the proximity of **several side chains of like charge**, e.g., Lys and Arg or Glu and Asp
- **steric crowding** caused by the proximity of bulky side chains, e.g., Val, Ile, Thr

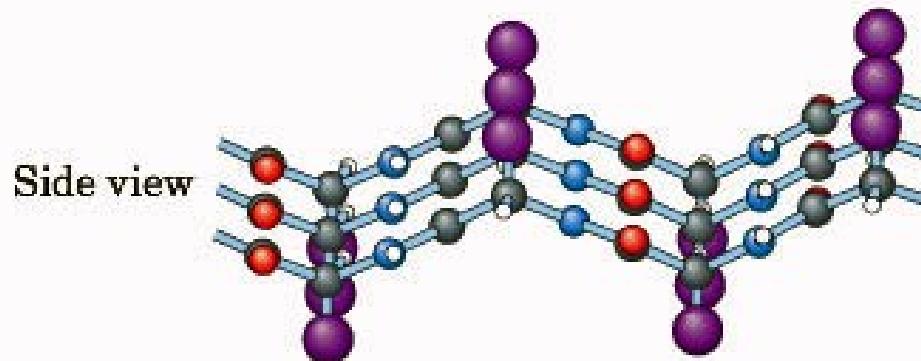
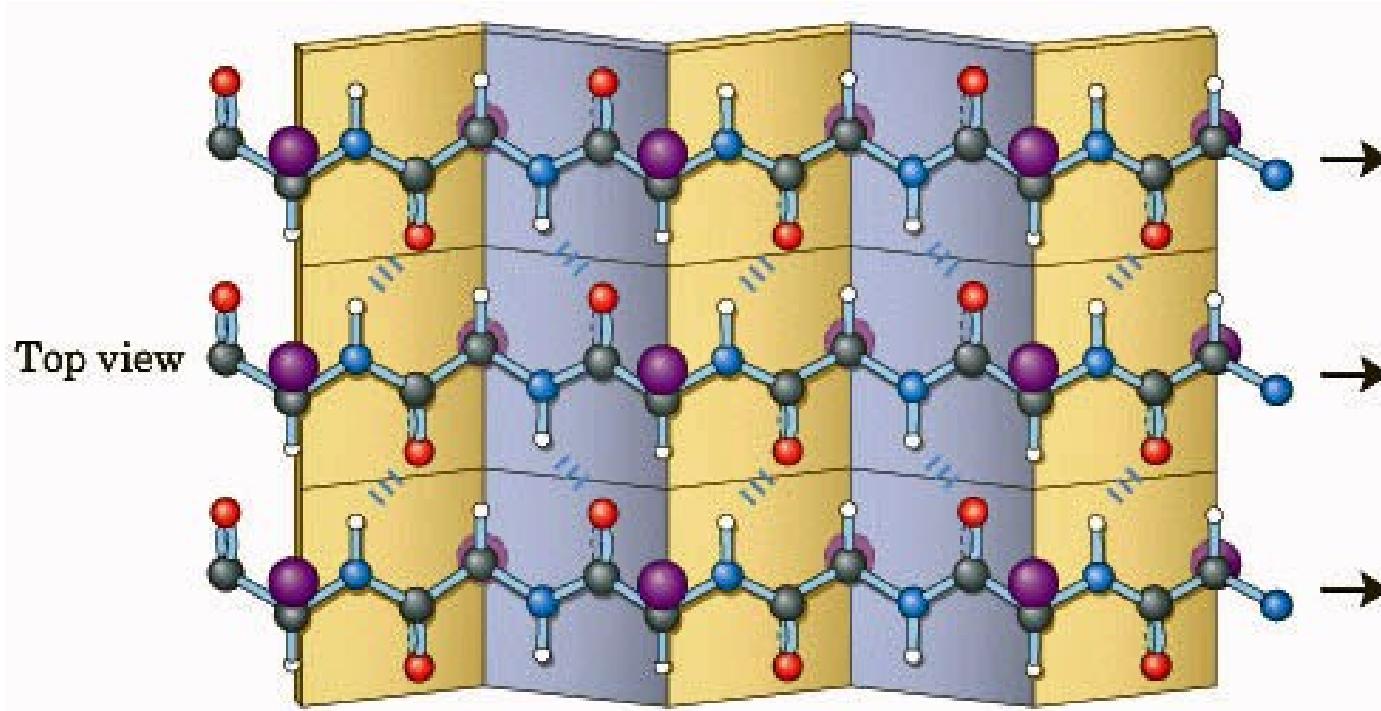
The BetaSheet

Composed of beta strands

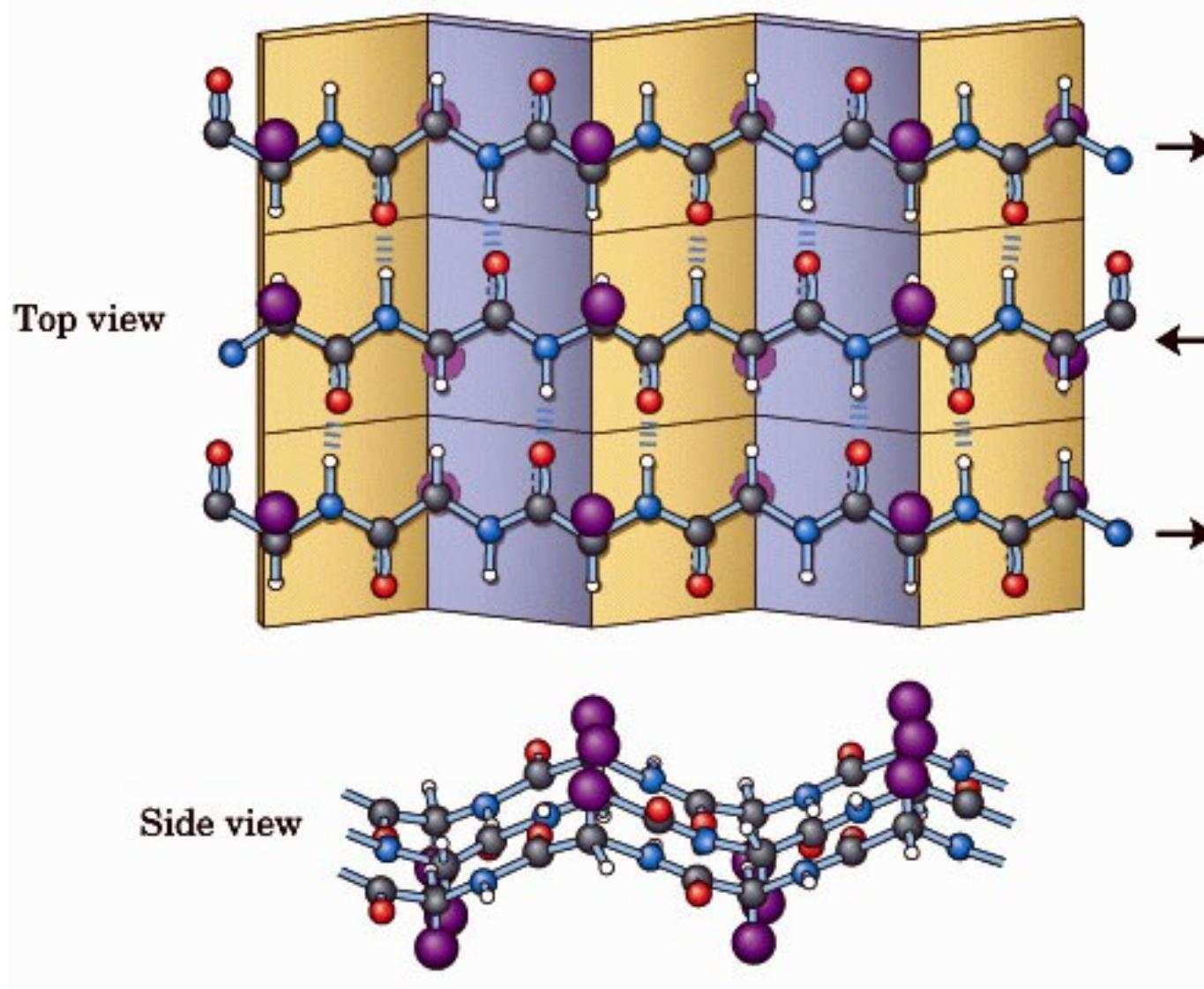
- Also first postulated by Pauling and Corey in 1951
- Strands may be parallel or antiparallel



Parallel Beta-Pleated Sheet

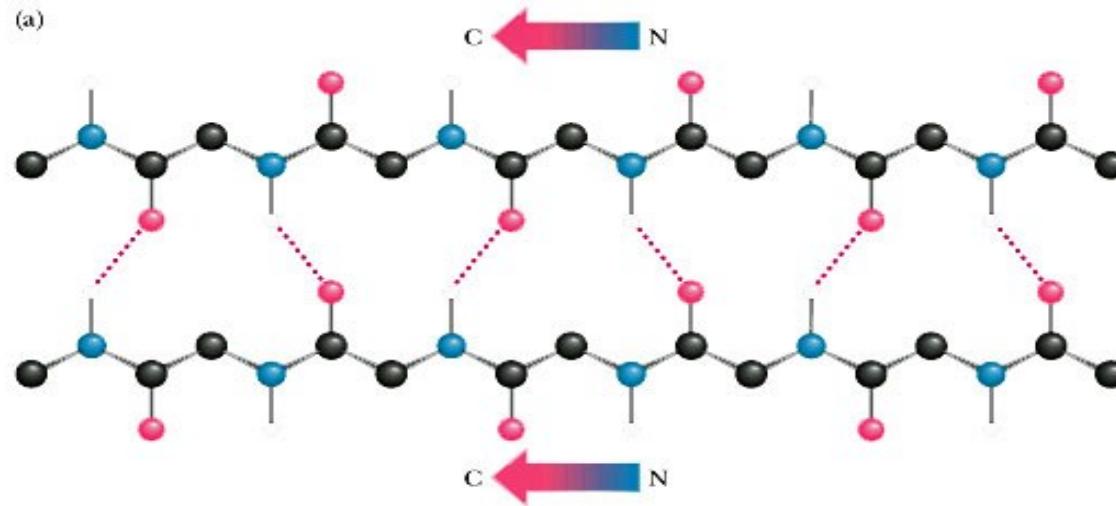


Antiparallel Beta-Pleated Sheet

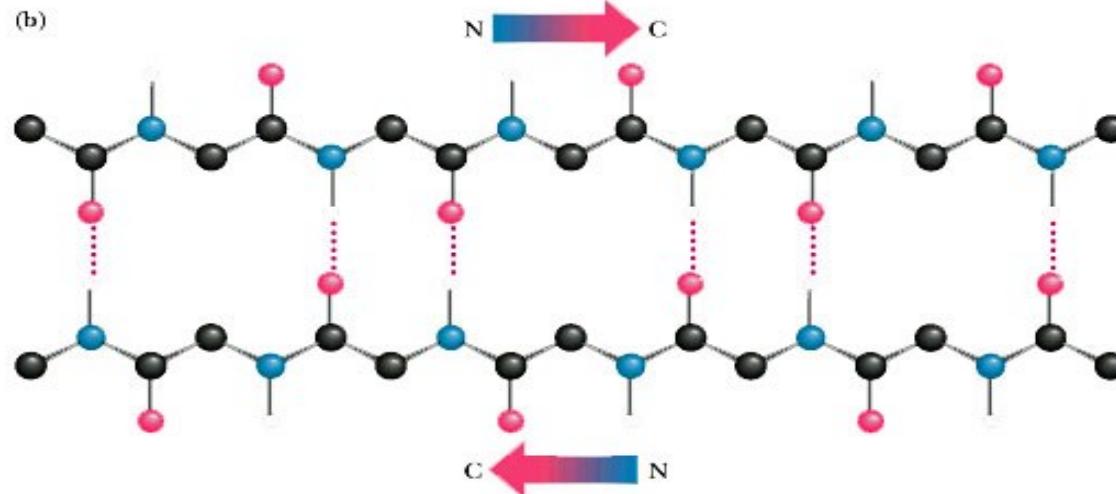


Hydrogen Bond Arrangement in Parallel and Antiparallel Beta-Pleated Sheets

(a)



(b)



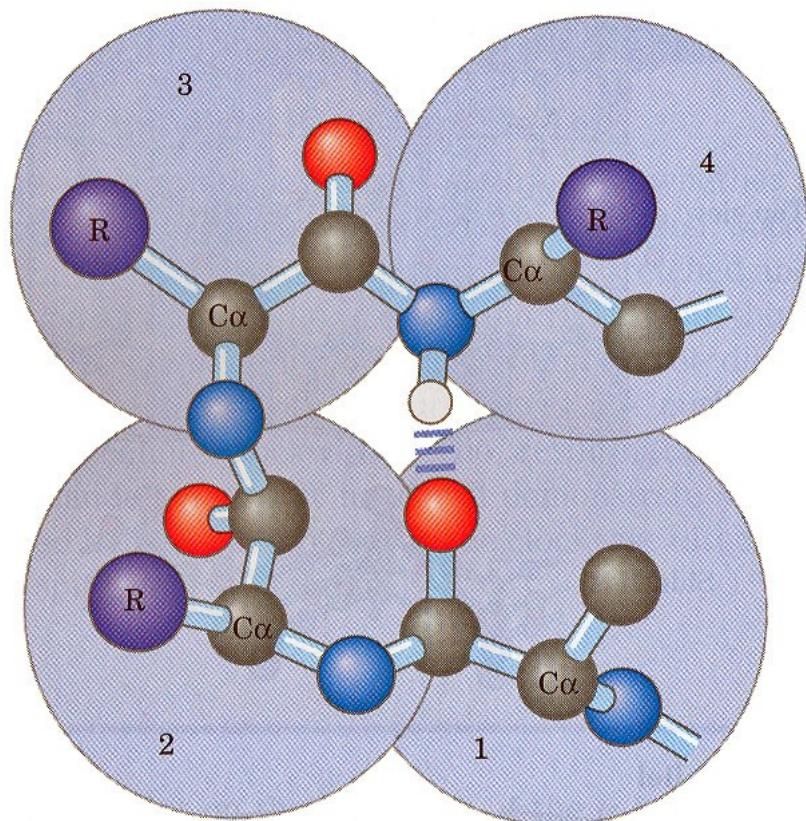
The Beta-Pleated Sheet

- Rise per residue:
 - 3.47 \AA for antiparallel strands
 - 3.25 \AA for parallel strands
 - Compared to 1.5 \AA for α -helix

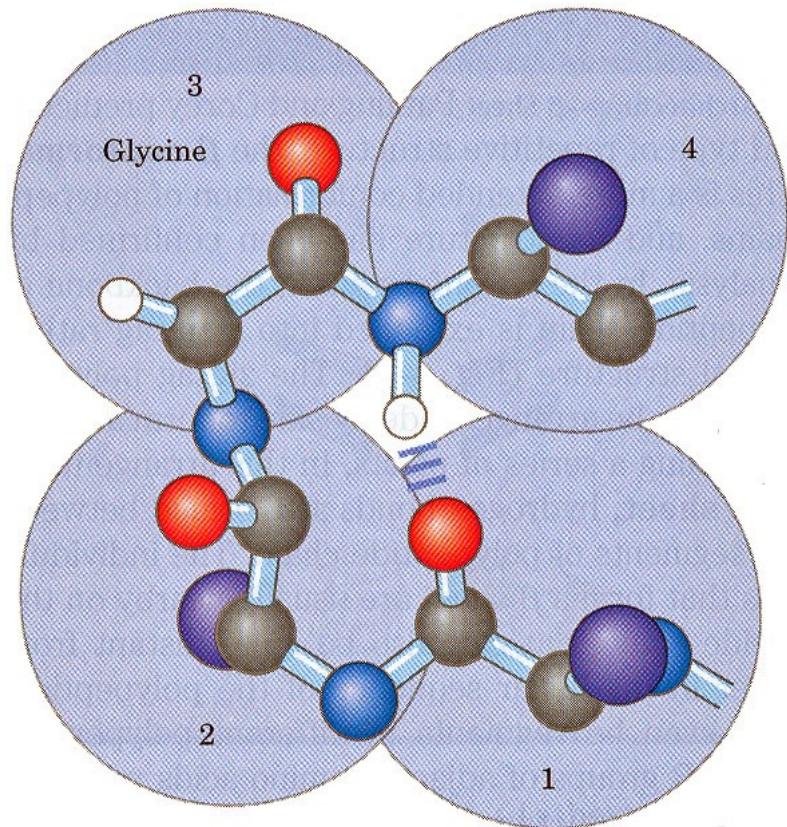
The Beta Turn

- allows the peptide chain to reverse direction
- carbonyl C of one residue is H-bonded to the amide proton of a residue three residues away
- proline and glycine are prevalent in beta turns

Structures of beta-Turns



Type I



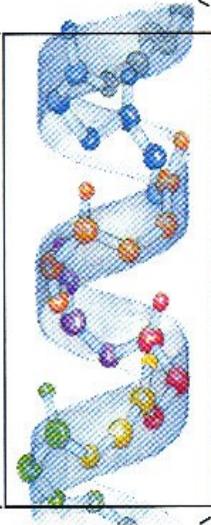
Type II
($\alpha_3 = \text{gly}$)

The Four Levels of Protein Structure

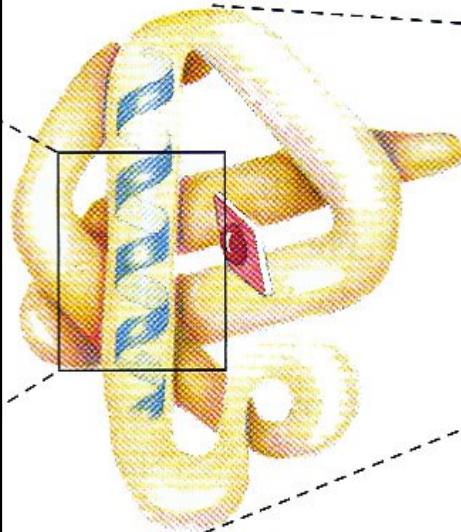
Primary
structure

Lys
Lys
Gly
Gly
Leu
Val
Ala
His

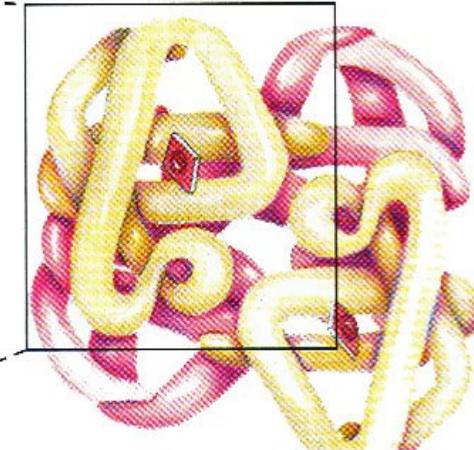
Secondary
structure



Tertiary
structure

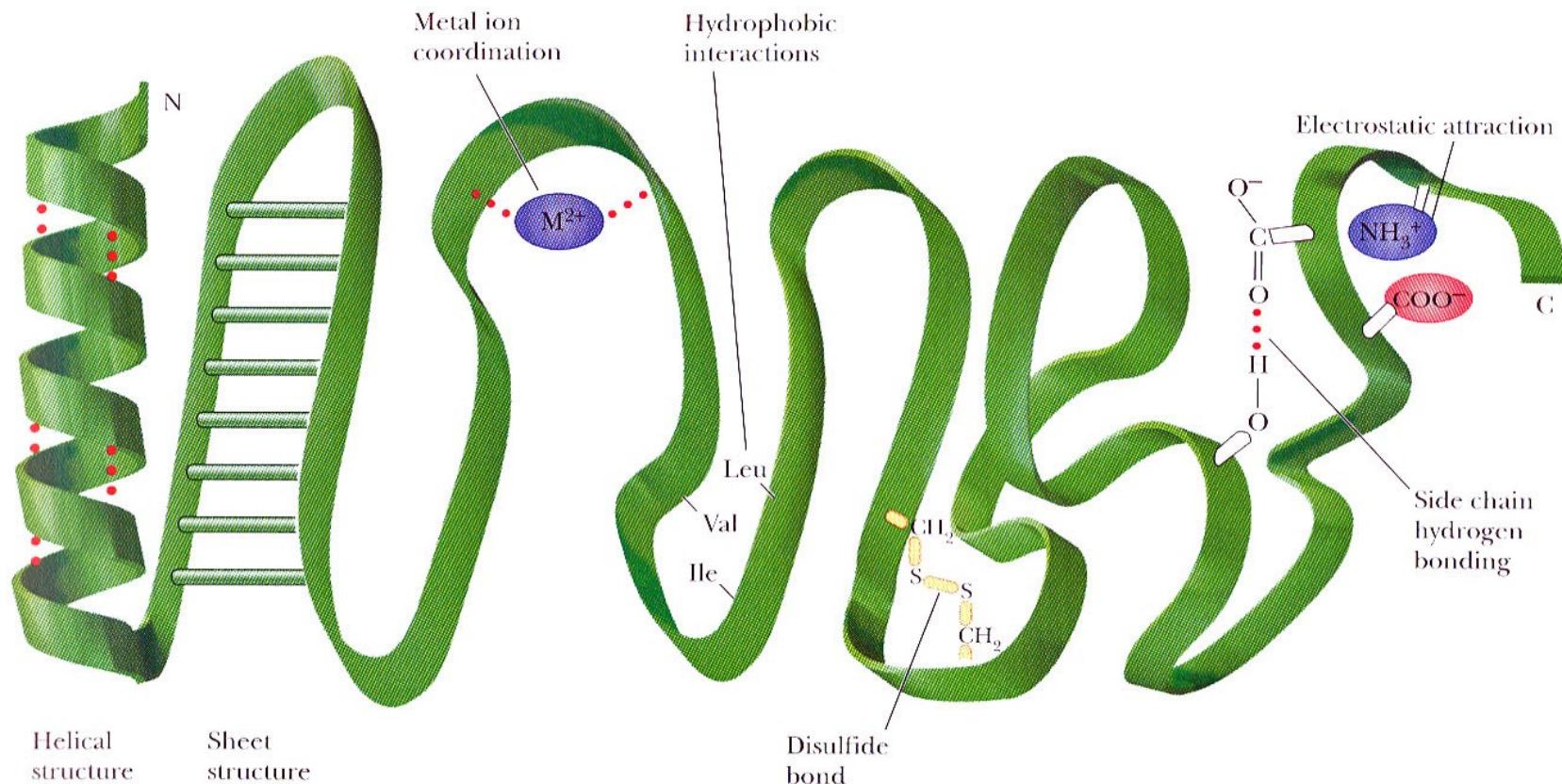


Quaternary
structure



Tertiary structure is the overall three-dimensional arrangement of all atoms in a protein; some proteins contain two or more separate polypeptide chains, or subunits, which maybe identical or different. The arrangement of these protein subunits in three-dimensional complexes constitutes quaternary structure.

Forces that Stabilize the Tertiary Structure of Proteins



Note that the helical and sheet structures are two kinds of backbone hydrogen bonding. Although backbone hydrogen bonding is part of secondary structure, the conformation of the backbone puts constraints on the possible arrangement of the side chains.