

Protein Structure

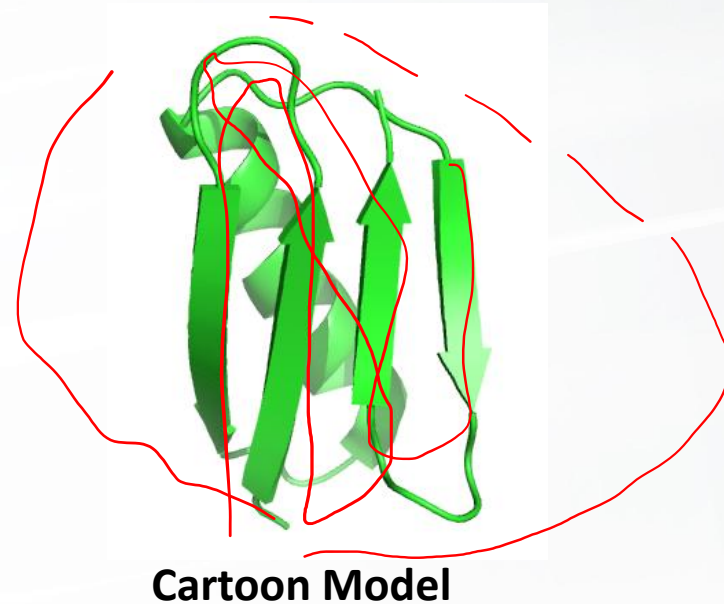
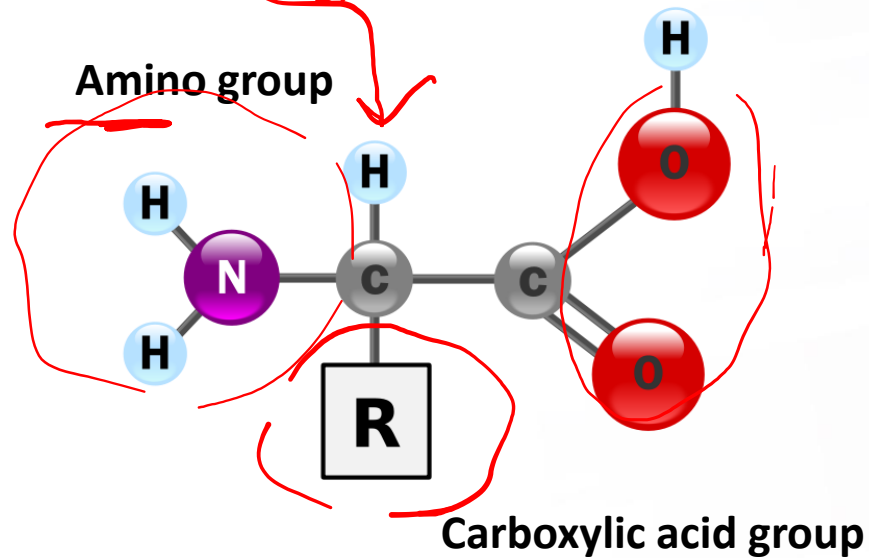
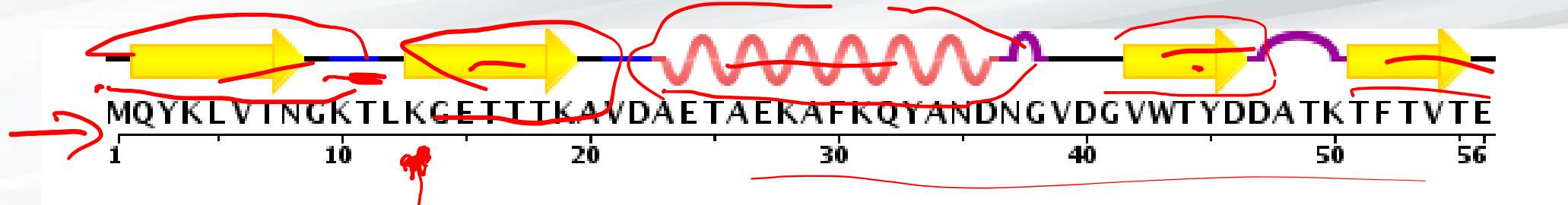
Biochemistry Boot Camp 2023

Session #1

Matthew Thompson

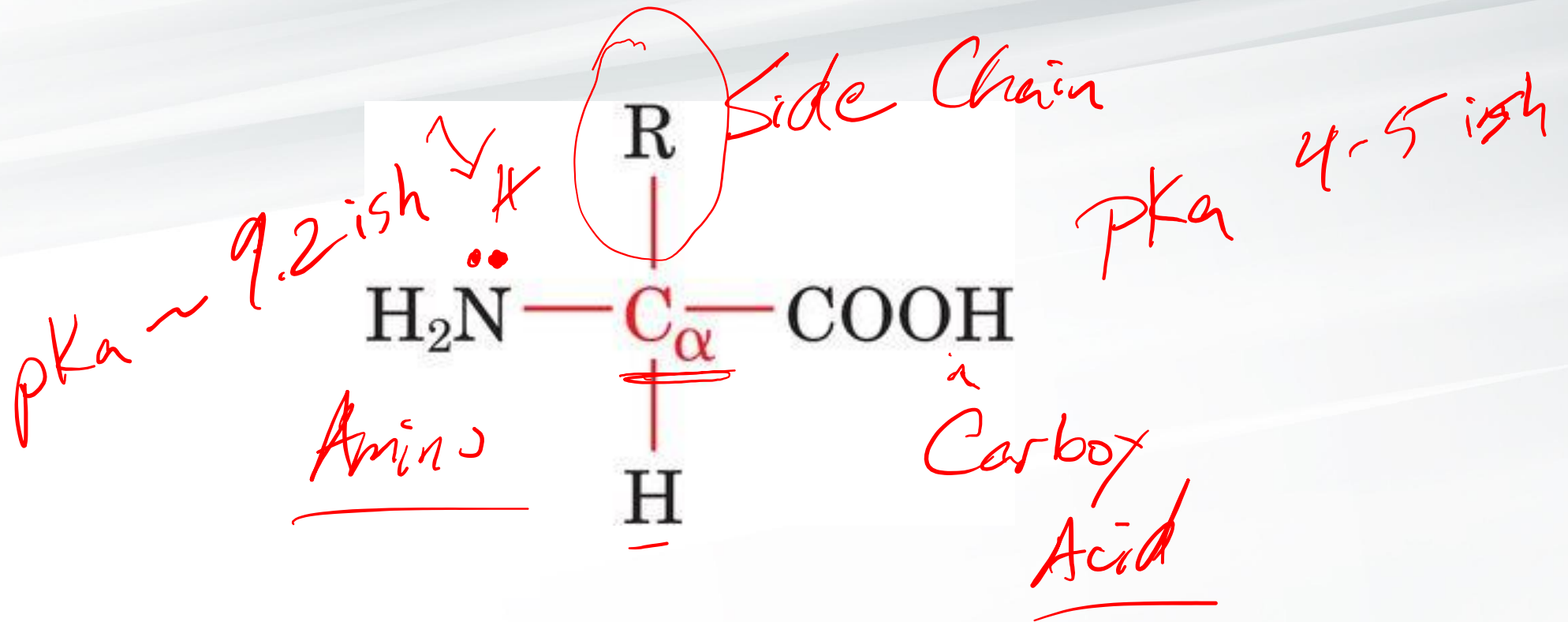
mthompson10@ua.edu

What is a Protein?



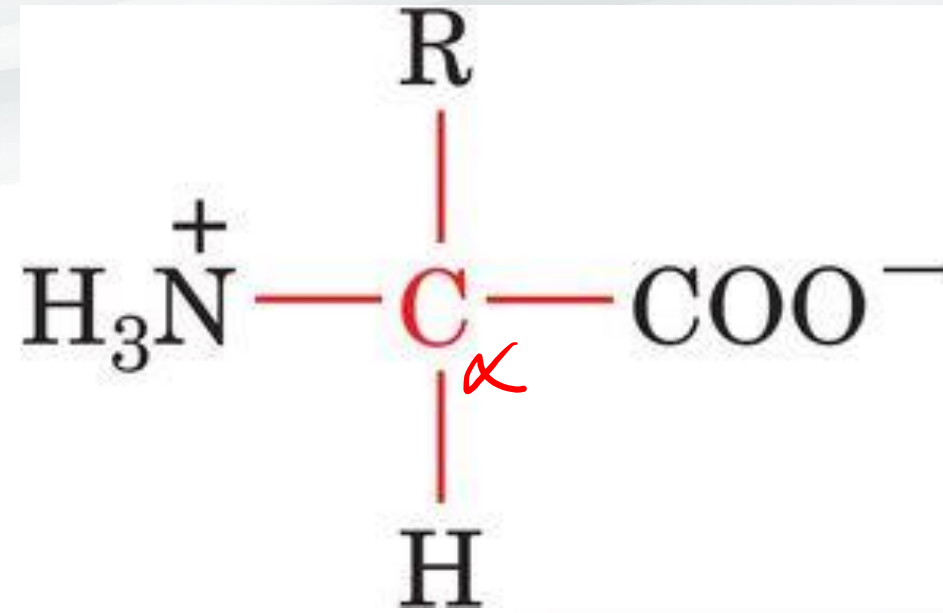
- **Answer:** Polymer of amino acids connected by peptide bonds
 - Multiple representations, including sequence, chemical structure, simplifying cartoons
- PyMOL Download: Staphylococcal Nuclease (sticks and cartoon)

General Structure of α -Amino Acid

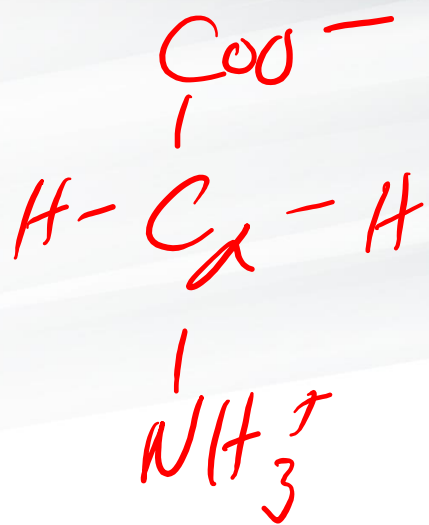


$$pH = pK_a + \log \frac{[base]}{[acid]}$$

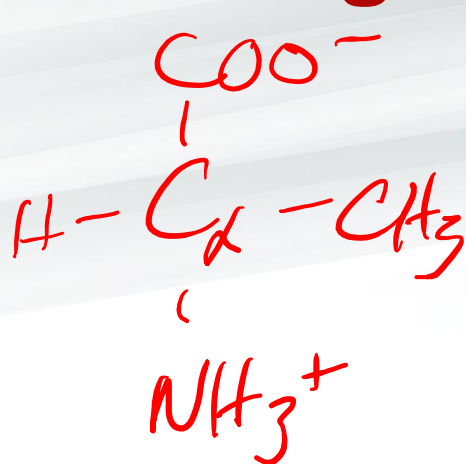
Dipolar Amino Acid: Zwitterion



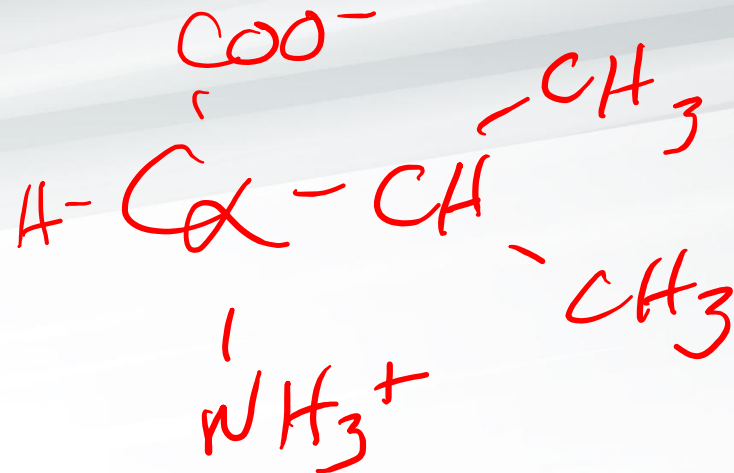
Drawing Amino Acids 101



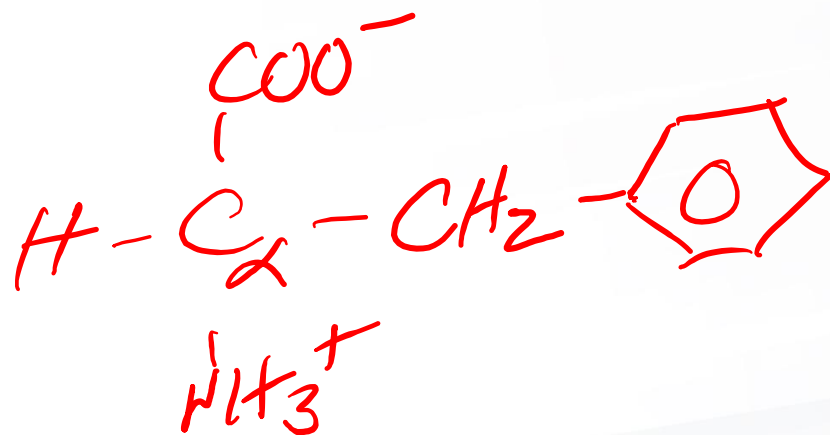
Glycine, Gly
G



Alanine, Ala
A



Valine, Val
V



phenylalanine
Phe, F

Structure of Amino Acids

TABLE 4-1 Covalent Structures and Abbreviations of the "Standard" Amino Acids of Proteins, Their Occurrence, and the pK Values of Their Ionizable Groups

Name, Three-letter Symbol, and One-letter Symbol	Structural Formula ^a	Residue Mass (D) ^b	Average Occurrence in Proteins (%) ^c	pK ₁ α-COOH ^d	pK ₂ α-NH ₃ ^{+,nd}	pK _R Side Chain ^d
<i>Amino acids with nonpolar side chains</i>						
Glycine Gly G		57.0	7.1	2.35	9.78	
Alanine Ala A		71.1	8.2	2.35	9.87	
Valine Val V		99.1	6.9	2.29	9.74	
Leucine Leu L		113.2	9.7	2.33	9.74	
Isoleucine Ile I		113.2	6.0	2.32	9.76	
Methionine Met M		131.2	2.4	2.13	9.28	
Proline Pro P		97.1	4.7	1.95	10.64	
Phenylalanine Phe F		147.2	3.9	2.20	9.31	
Tryptophan Trp W		186.2	1.1	2.46	9.41	

↑↑↑

pKa Side Chain

9 AAs

Nonpolar

^aThe ionic forms shown are those predominating at pH 7.0 (except for that of histidine^e) although residue mass is given for the neutral compound. The C_α atoms, as well as atoms marked with an asterisk, are chiral centers with configurations as indicated according to Fischer projection formulas (Section 4-2). The standard organic numbering system is provided for heterocycles.

^bThe residue masses are given for the neutral residues. For the molecular masses of the parent amino acids, add 18.0 D, the molecular mass of H₂O, to the residue masses. For side chain masses, subtract 56.0 D, the formula mass of a peptide group, from the residue masses.

^cThe average amino acid composition in the complete SWISS-PROT database (<http://www.expasy.ch/sprot/relnotes/relstat.html>), Release 2013_13. Individual proteins may exhibit large deviations from these quantities.

^dData from Dawson, R.M.C., Elliott, D.C., Elliott, W.H., and Jones, K.M., *Data for Biochemical Research* (3rd ed.), pp. 1–31, Oxford Science Publications (1986).

Structures of Standard Amino Acids

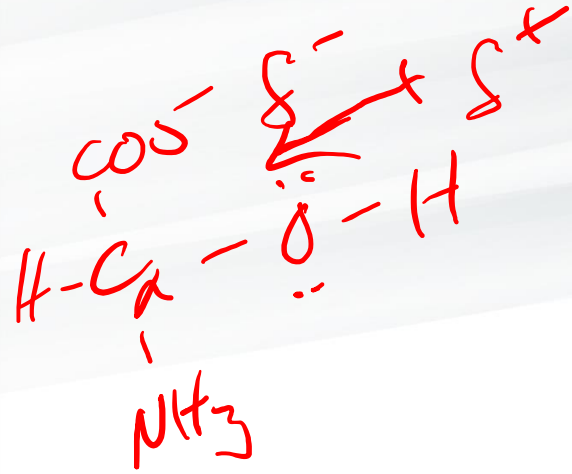


TABLE 4-1 (Continued)

Name, Three-letter Symbol, and One-letter Symbol	Structural Formula ^a	Residue Mass (D) ^b	Average Occurrence in Proteins (%) ^c	pK ₁ α-COOH ^d	pK ₂ α-NH ₃ ⁺ ^d	pK _R Side Chain ^d
<i>Amino acids with uncharged polar side chains</i>						
Serine Ser S		87.1	6.6	2.19	9.21	
Threonine Thr T		101.1	5.3	2.09	9.10	
Asparagine ^e Asn N		114.1	4.1	2.14	8.72	
Glutamine ^e Gln Q		128.1	3.9	2.17	9.13	
Tyrosine Tyr Y		163.2	2.9	2.20	9.21	10.46 (phenol)
Cysteine Cys C		103.1	1.4	1.92	10.70	8.37 (sulfhydryl)
<i>Amino acids with charged polar side chains</i>						
Lysine Lys K		128.2	5.9	2.16	9.06	10.54 (ε-NH ₃ ⁺)
Arginine Arg R		156.2	5.5	1.82	8.99	12.48 (guanidino)
Histidine ^f His H		137.1	2.3	1.80	9.33	6.04 (imidazole)
Aspartic acid ^e Asp D		115.1	5.4	1.99	9.90	3.90 (β-COOH)
Glutamic acid ^e Glu E		129.1	6.8	2.10	9.47	4.07 (γ-COOH)

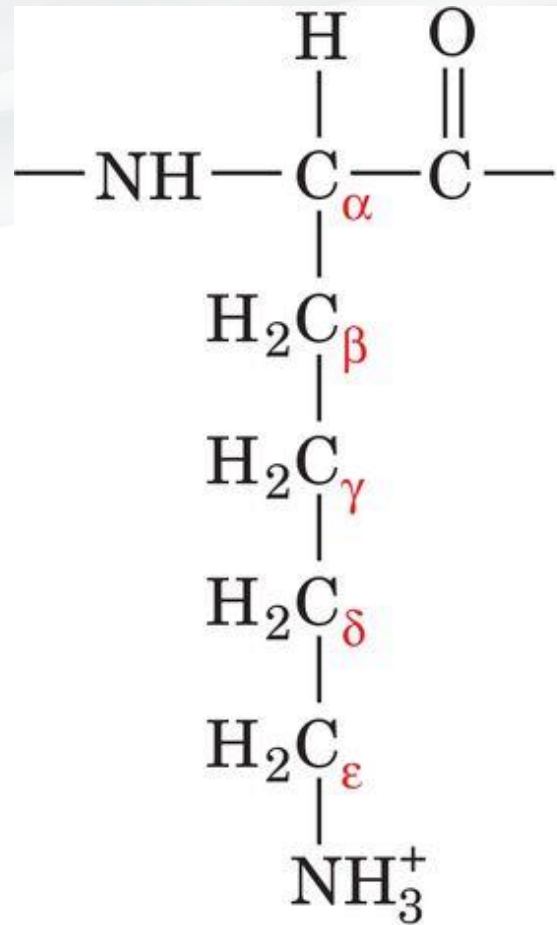
6 polar Uncharged

5 polar Charged

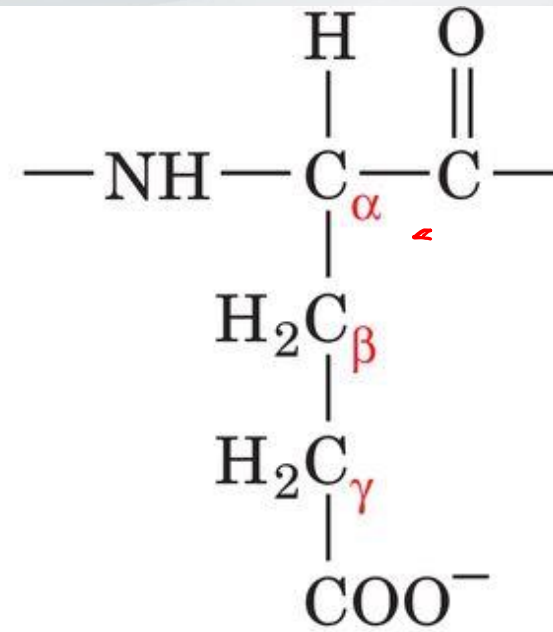
Exchangeable Proton

^aThe three- and one-letter symbols for asparagine or aspartic acid are Asx and B, whereas for glutamine or glutamic acid they are Glx and Z. The one-letter symbol for an undetermined or "nonstandard" amino acid is X.
^bBoth neutral and protonated forms of histidine are present at pH 7.0, since its pK_R is close to 7.0.

Amino Acid Nomenclature

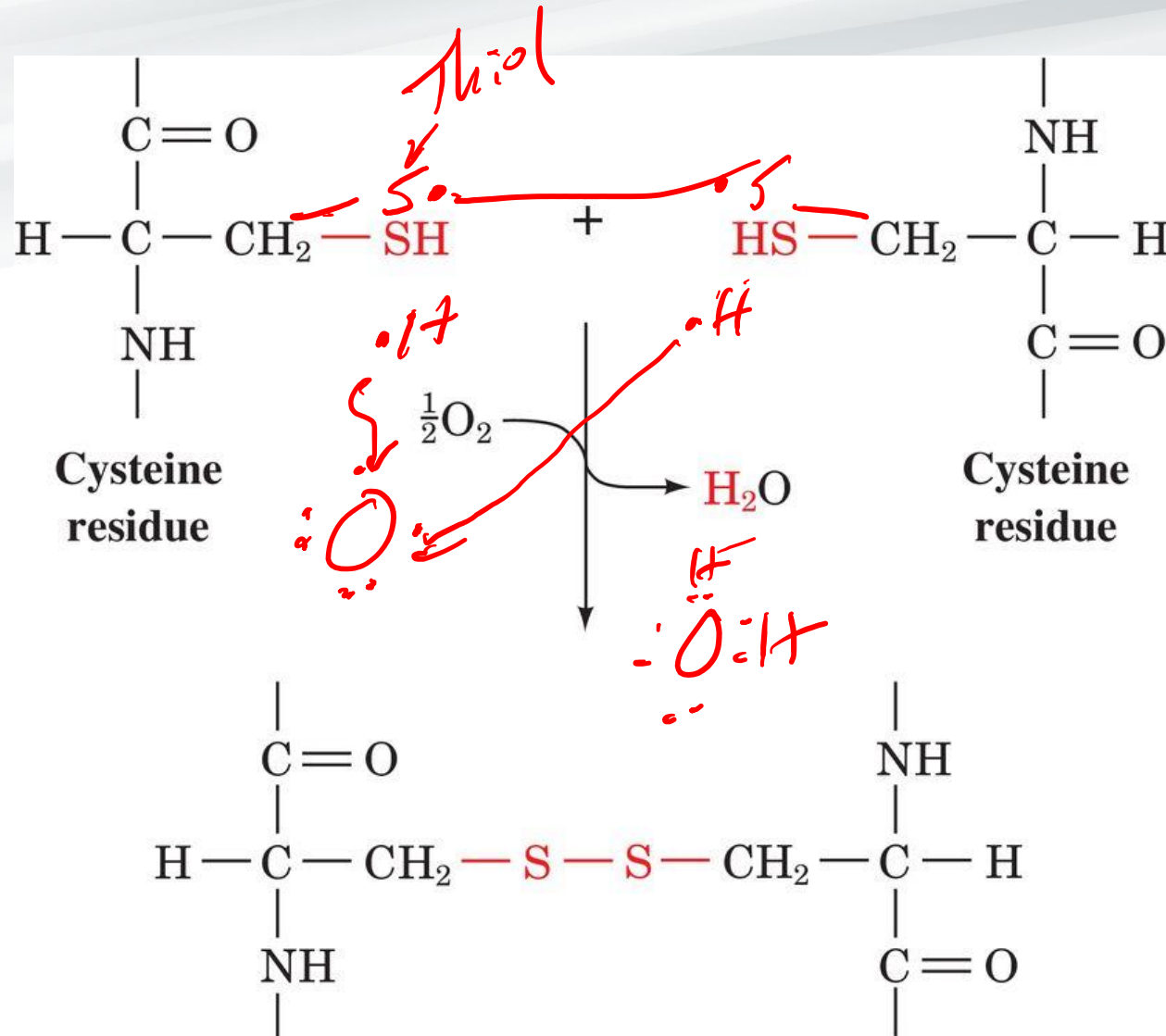


Lys

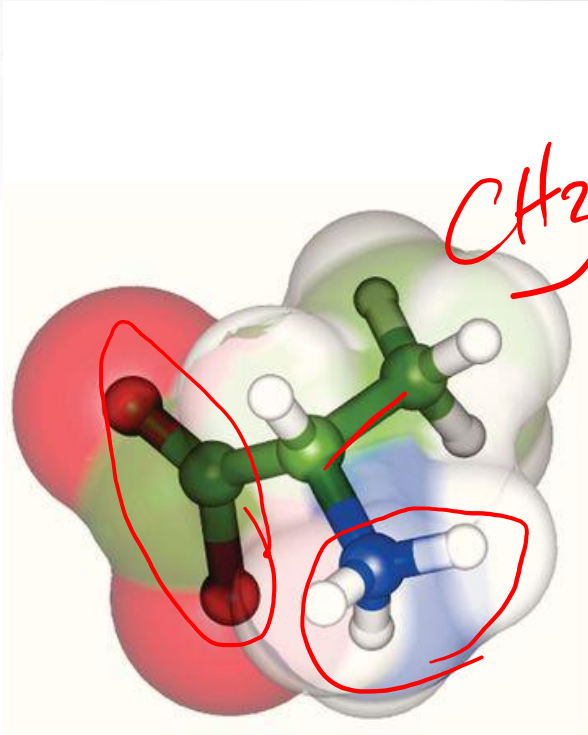


Glu

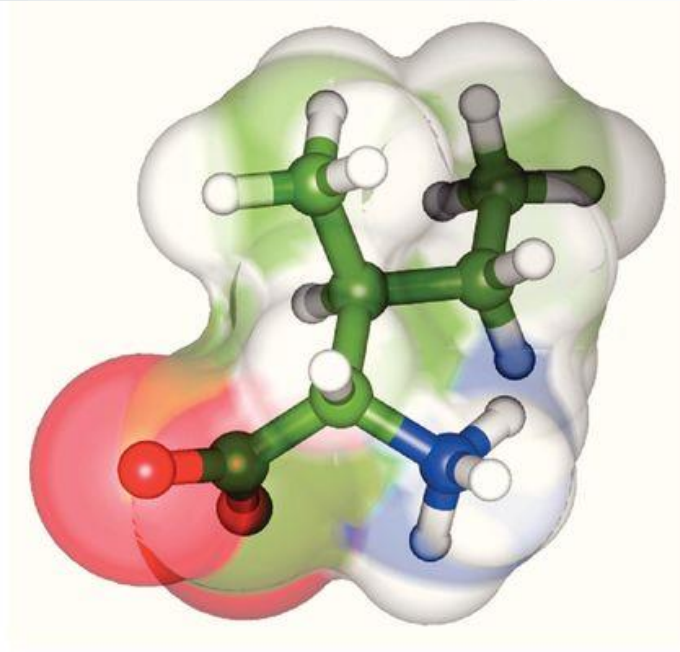
Cysteine Can Form Disulfide Bonds



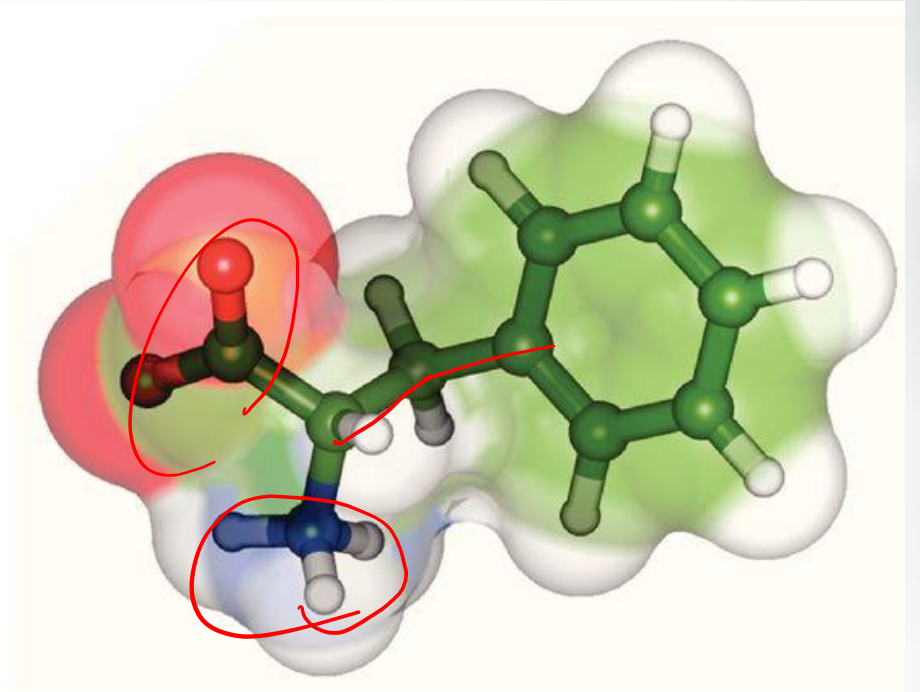
Nonpolar Side Chains



Alanine

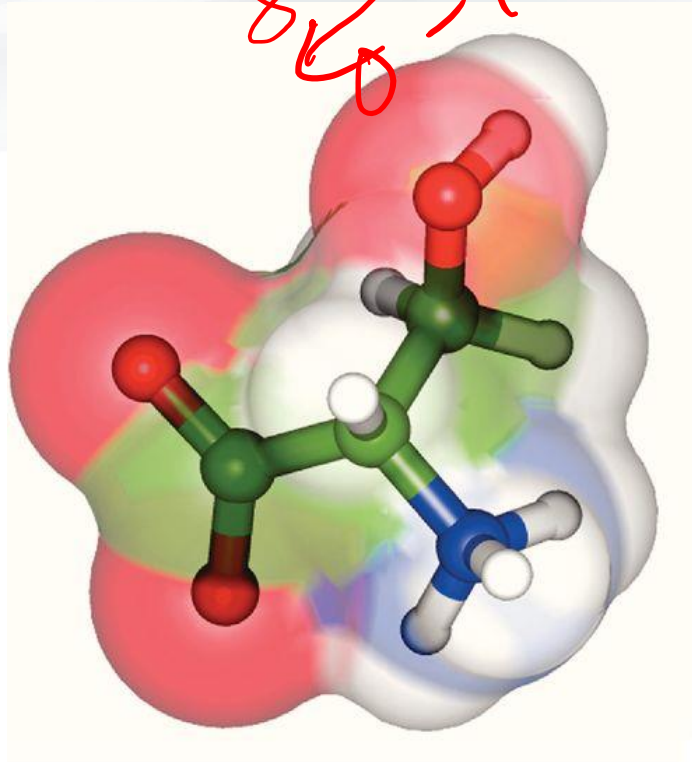


Isoleucine

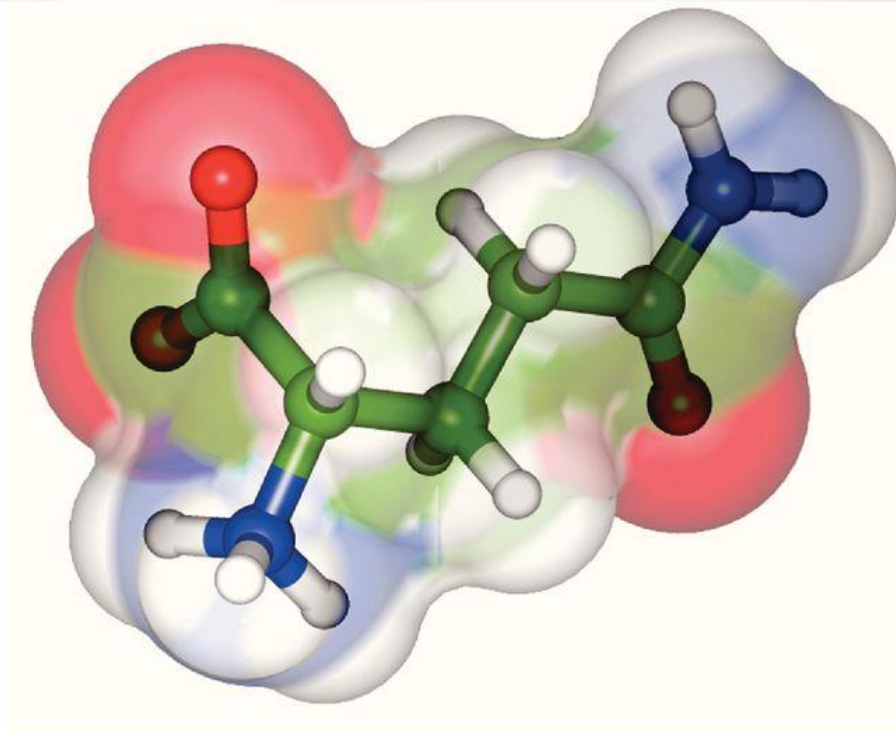


Phenylalanine

Uncharged Polar Side Chains

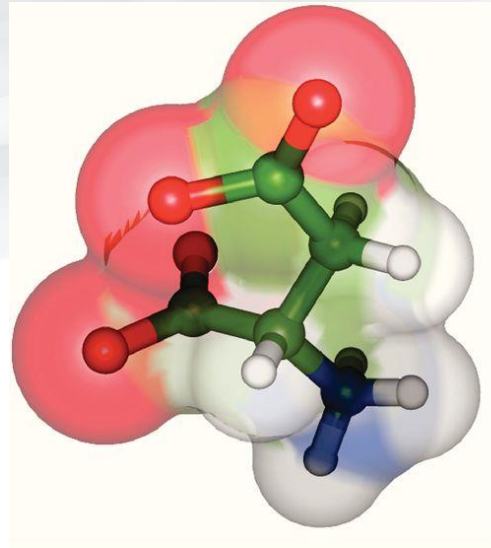


Serine

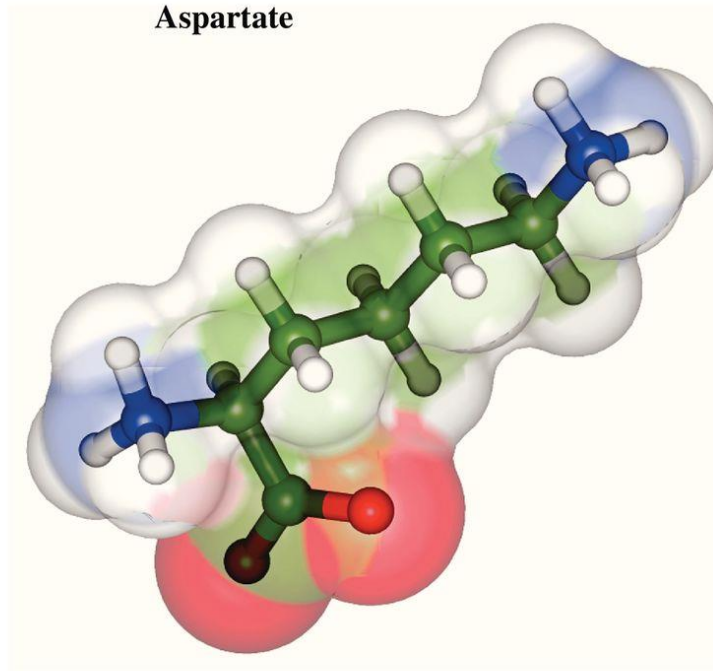


Glutamine

Charged Polar Side Chains

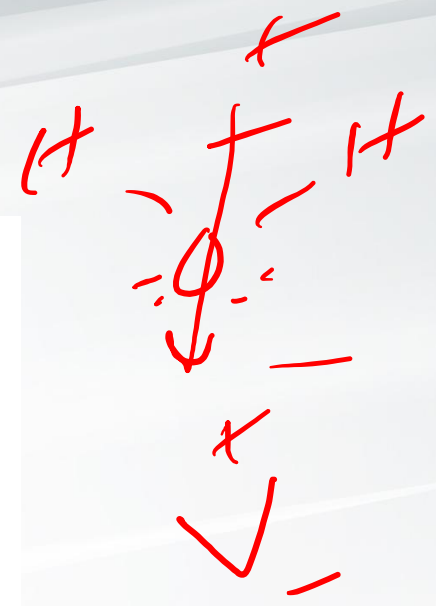
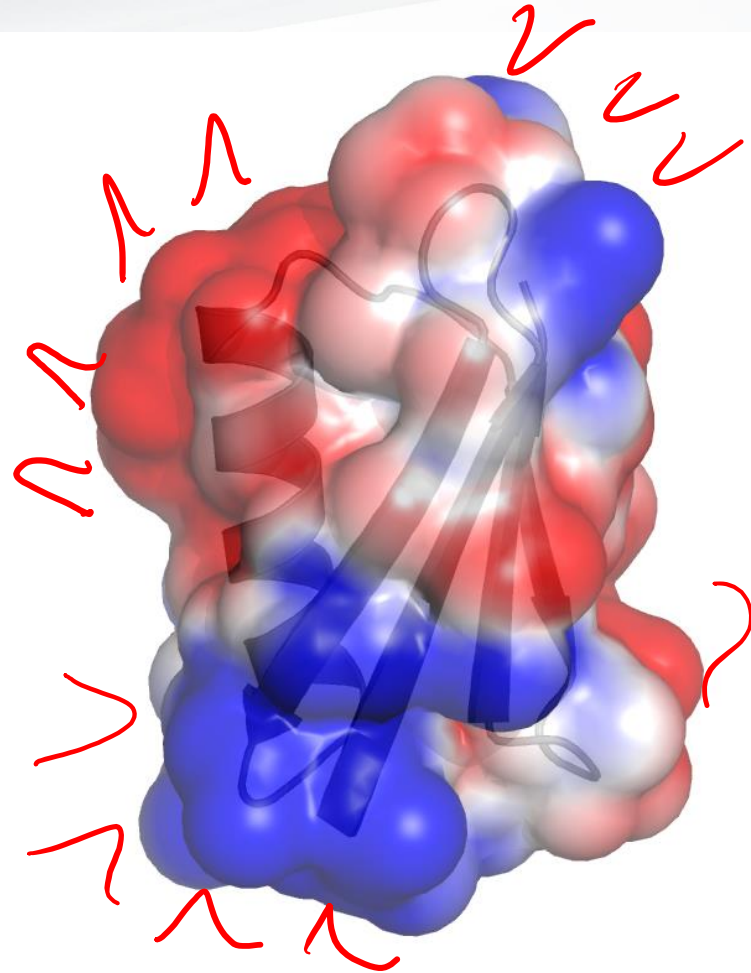


Aspartate



Lysine

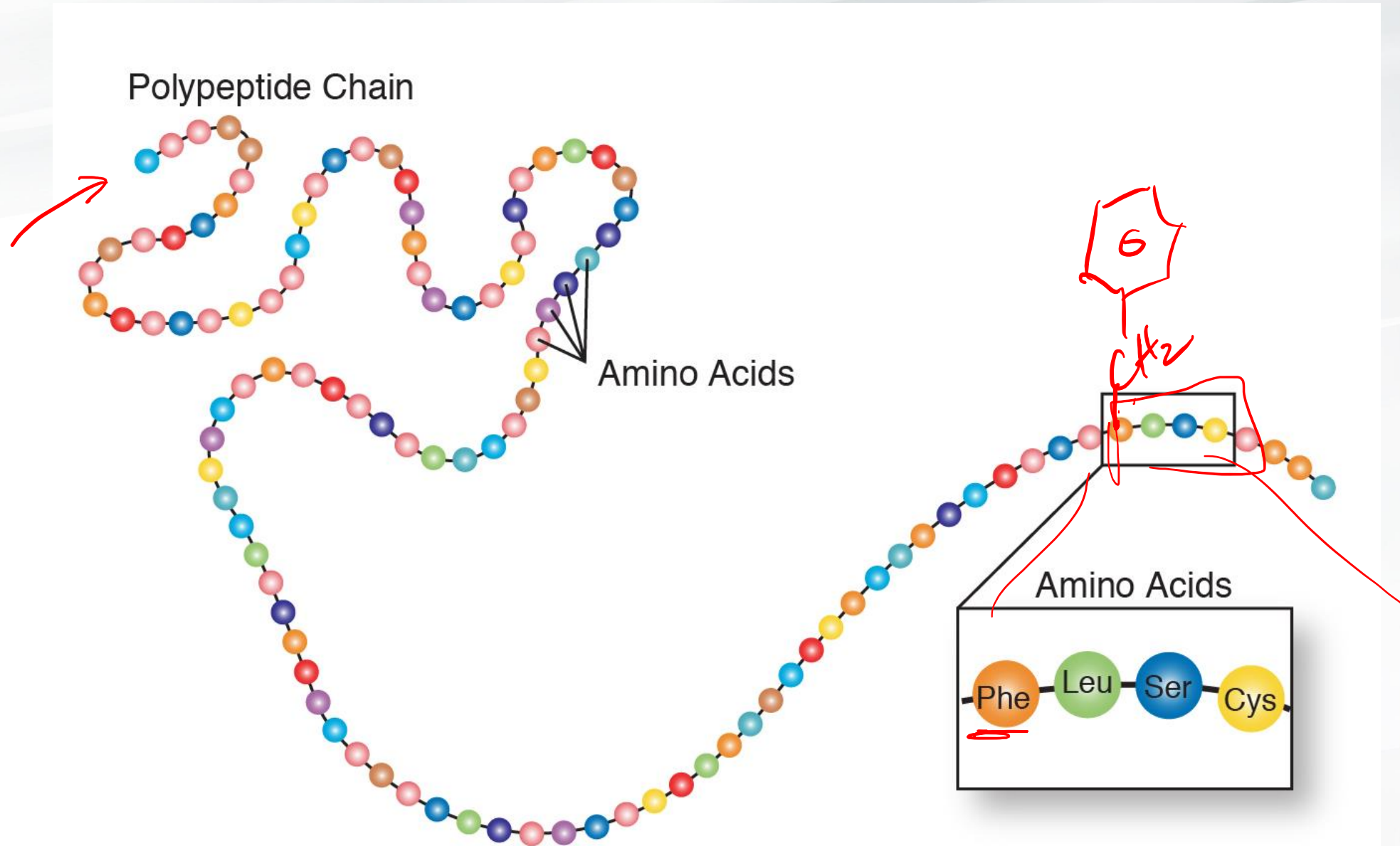
Electrostatic surface



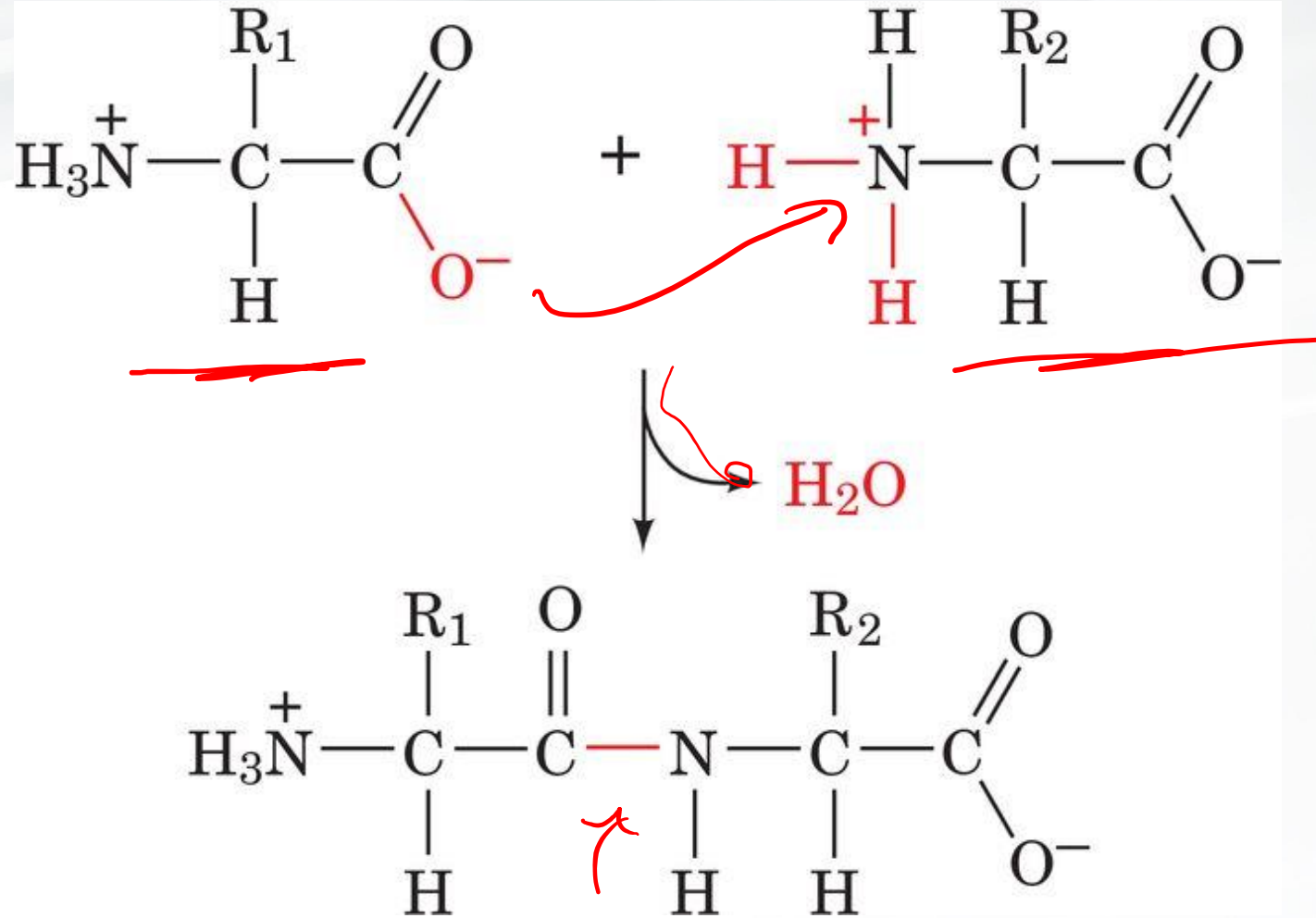
Red – negative charge
Blue – positive charge



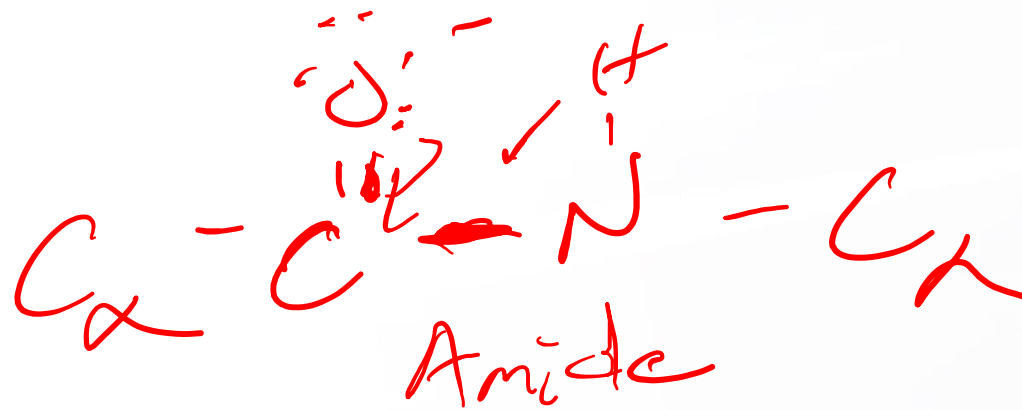
Proteins Are Composed of a “Chain” of Amino Acids



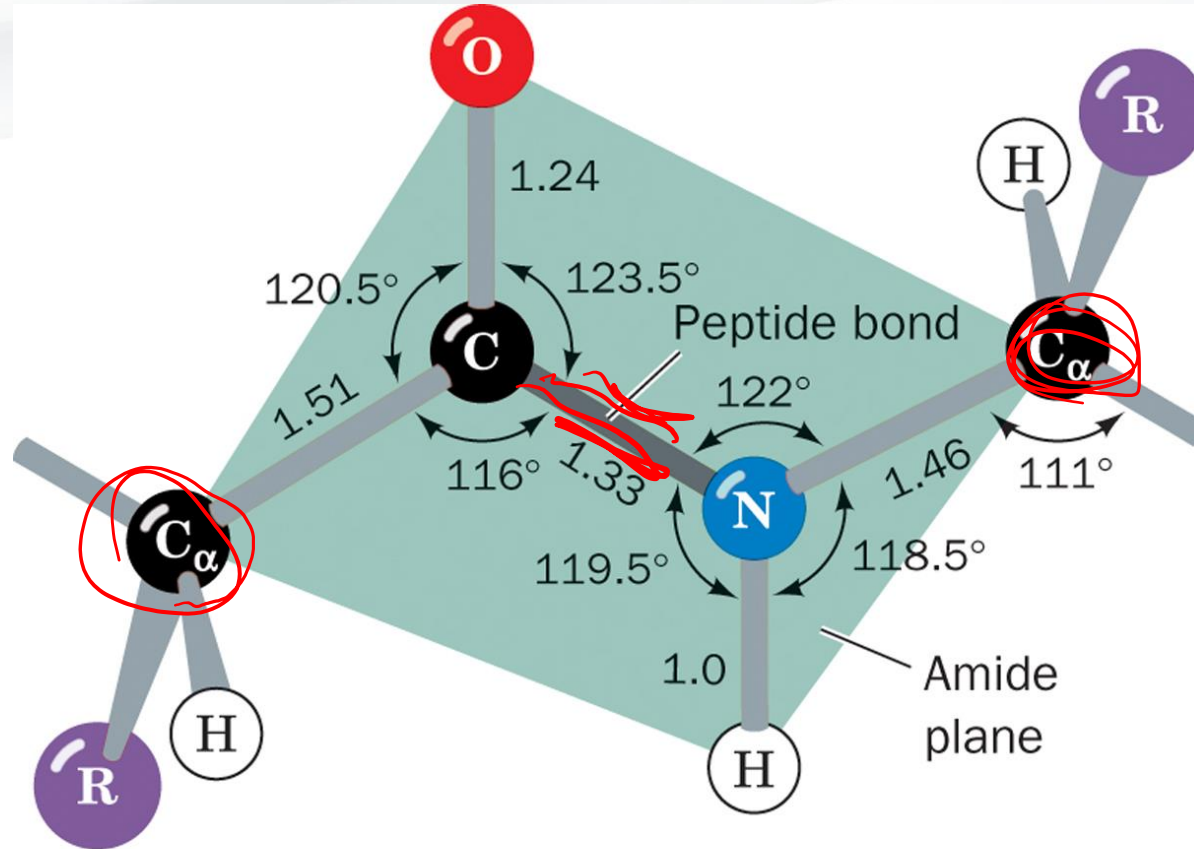
Condensation of Two Amino Acids



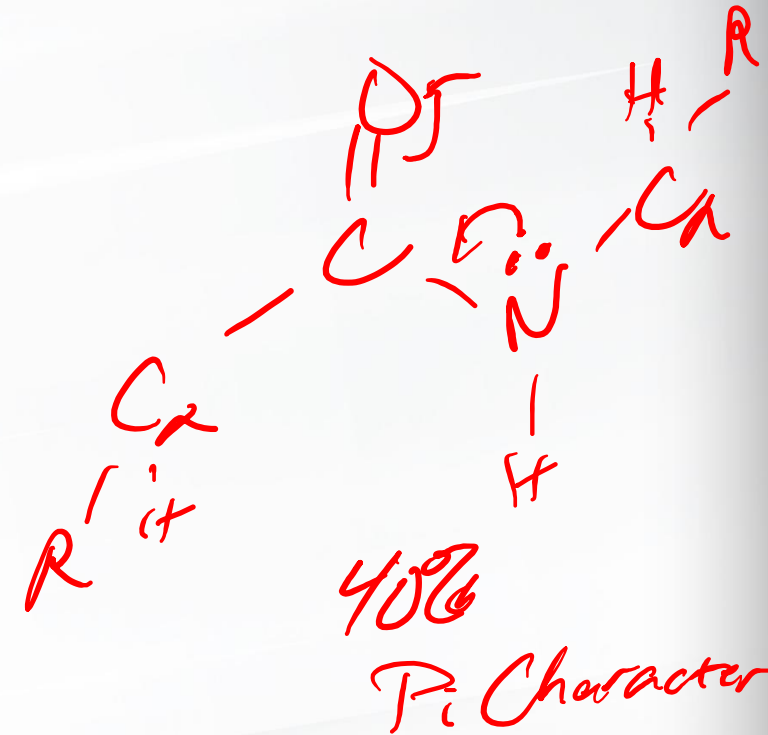
Condensation of Two Amino Acids



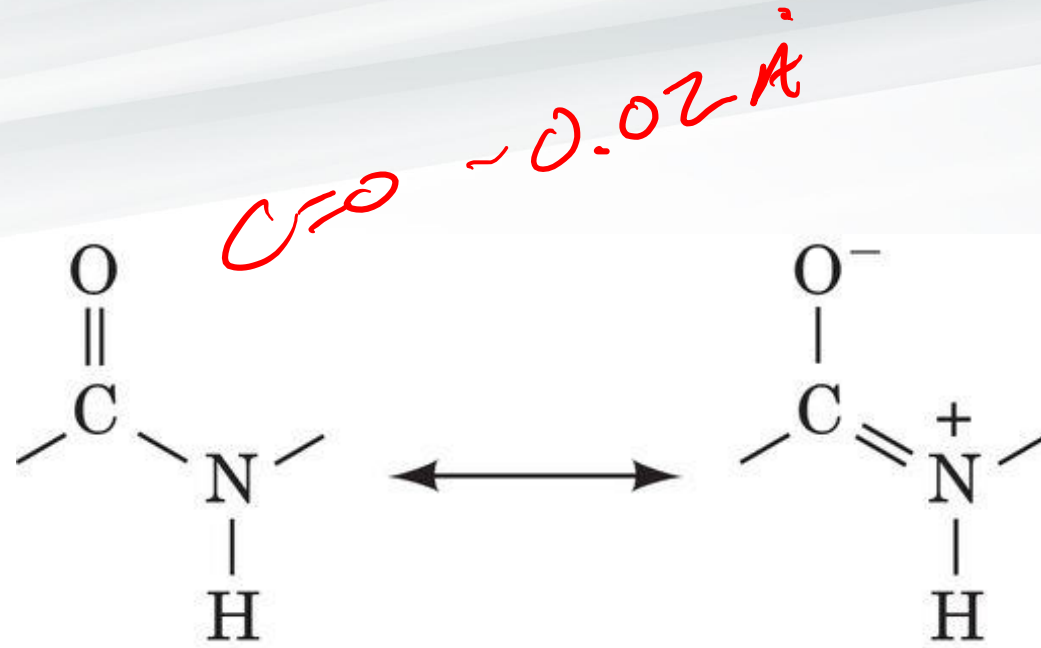
Peptide Bonds Assume Trans Conformation



After Marsh, R.E. and Donohue, J., Adv. Protein Chem. 22, 249 (1967).

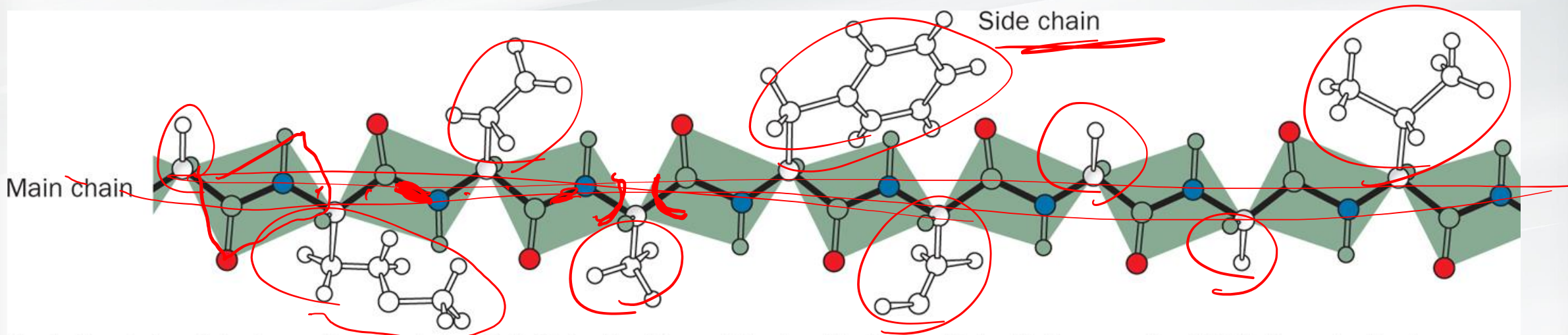


Resonance of Peptide Bond

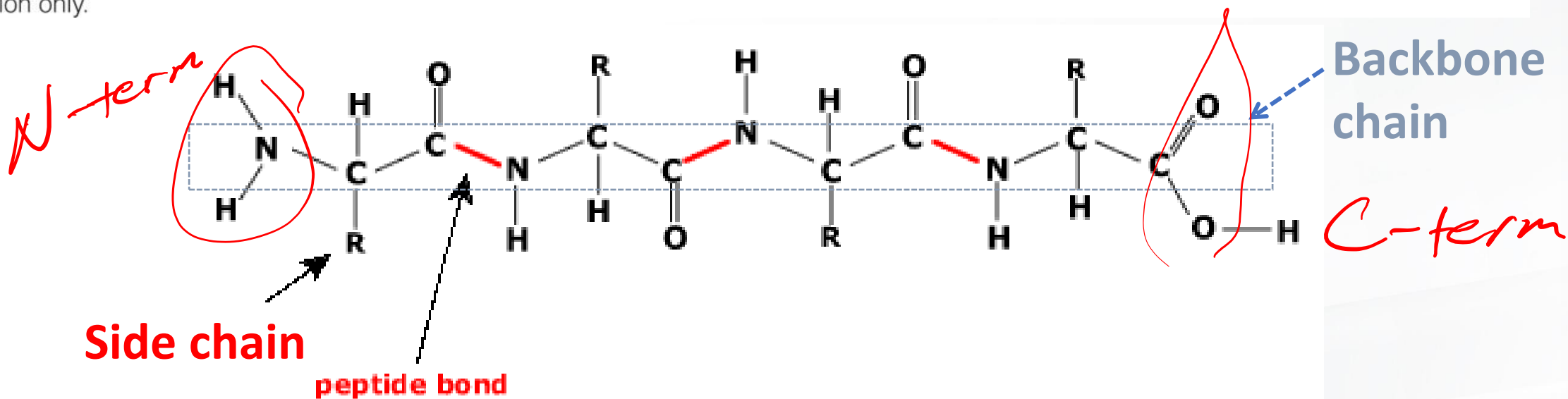


C-N ~ 0.13 Å Shorter

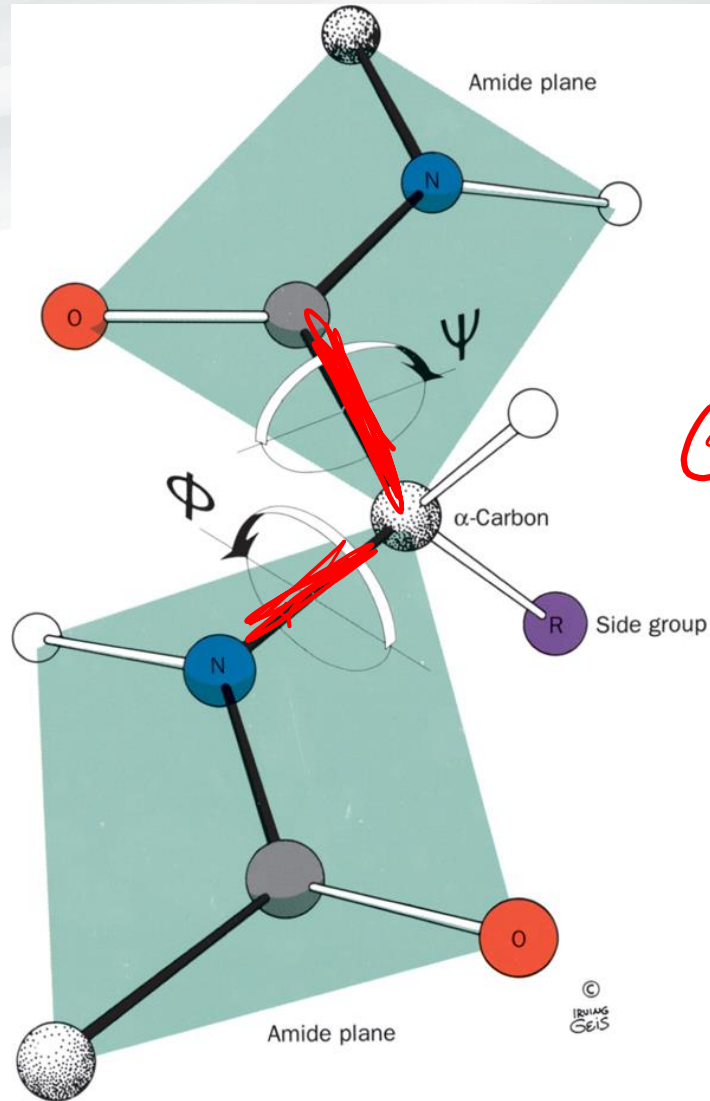
Extended Conformation of Polypeptide The Protein “Backbone”



Illustration, Irving Geis. Image from the Irving Geis Collection/Howard Hughes Medical Institute. Rights owned by HHMI. Reproduction by permission only.

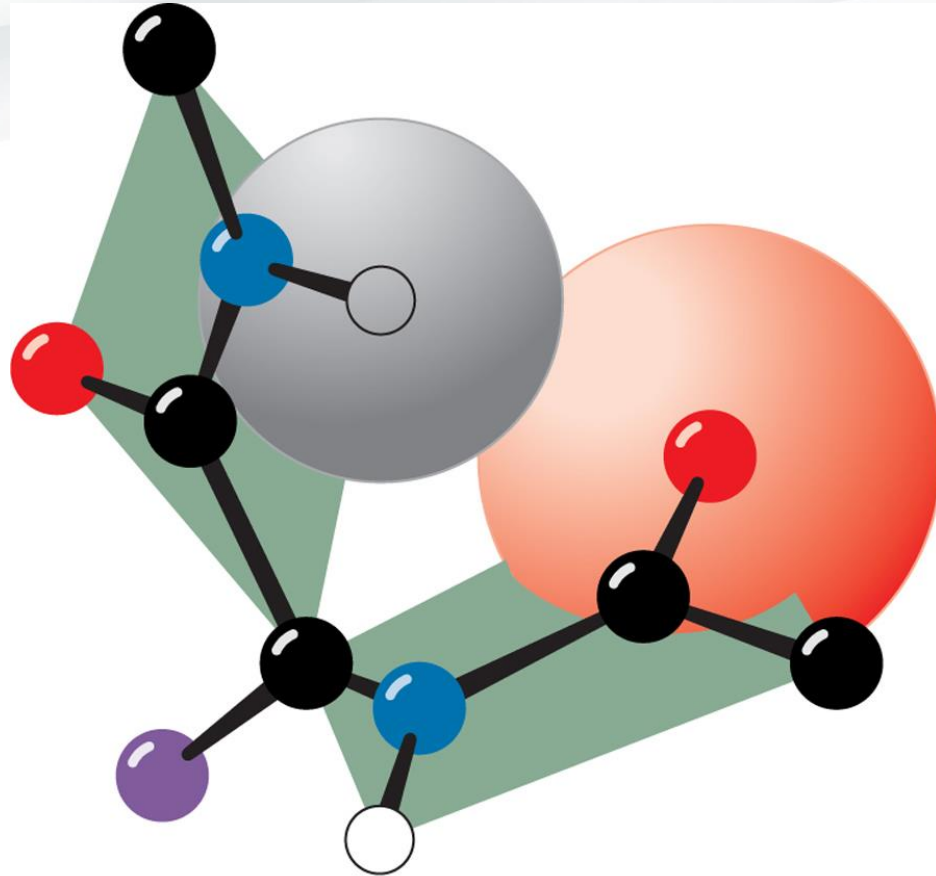


Torsion Angles of Polypeptide Backbone



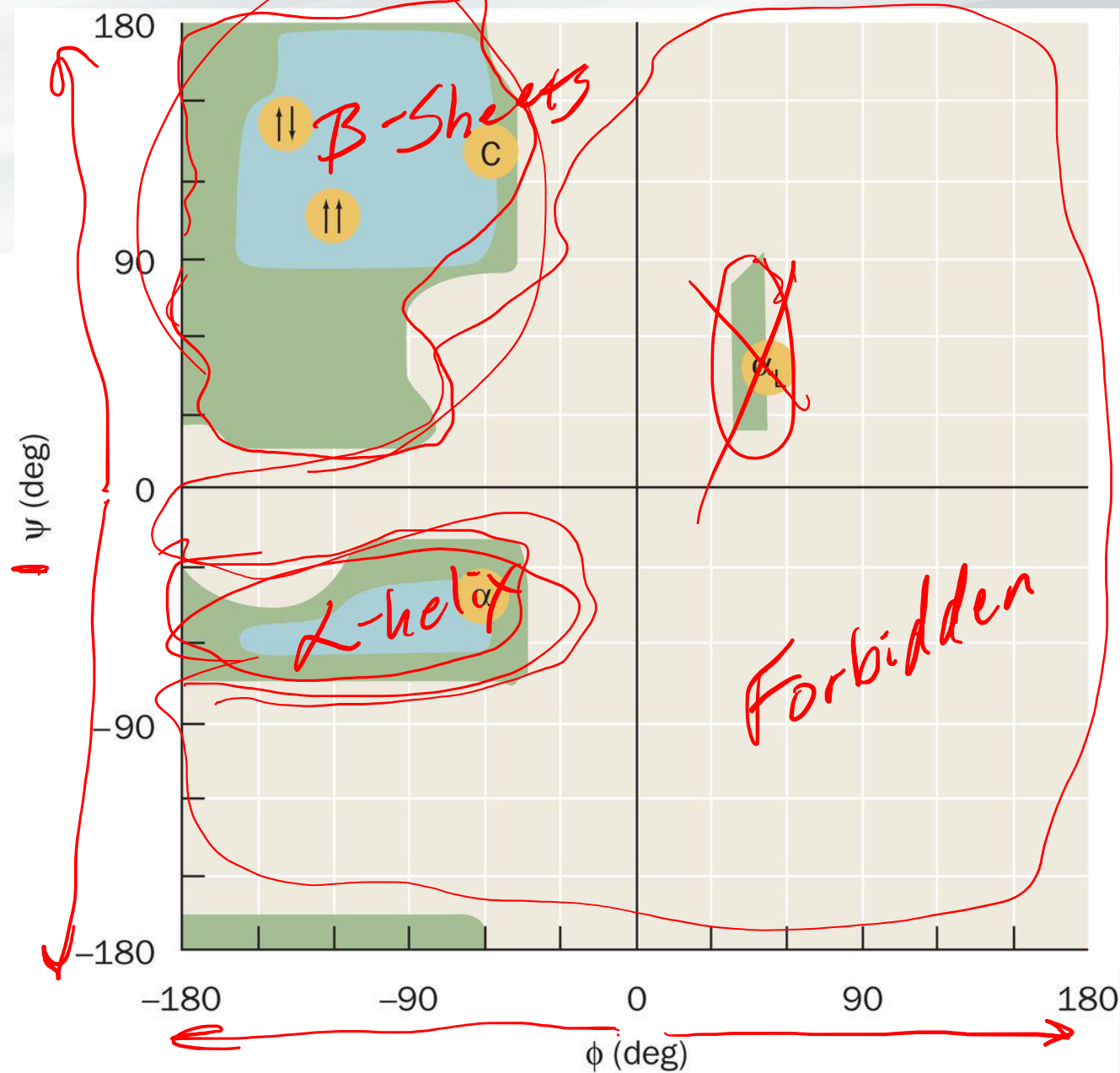
Illustration, Irving Geis. Image from the Irving Geis Collection/Howard Hughes Medical Institute. Rights owned by HHMI. Reproduction by permission only.

Steric Interference of Adjacent Peptide Groups



Illustration, Irving Geis. Image from the Irving Geis Collection/Howard Hughes Medical Institute. Rights owned by HHMI. Reproduction by permission only.

Ramachandran Diagram



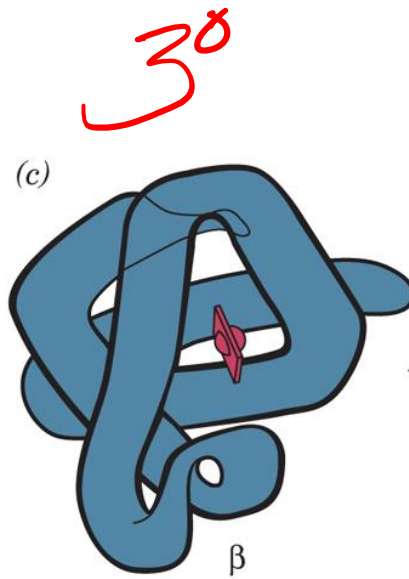
Levels of Protein Structure

(a) – Lys – Ala – His – Gly – Lys – Lys – Val – Leu – Gly – Ala –
Primary structure (amino acid sequence in a polypeptide chain)

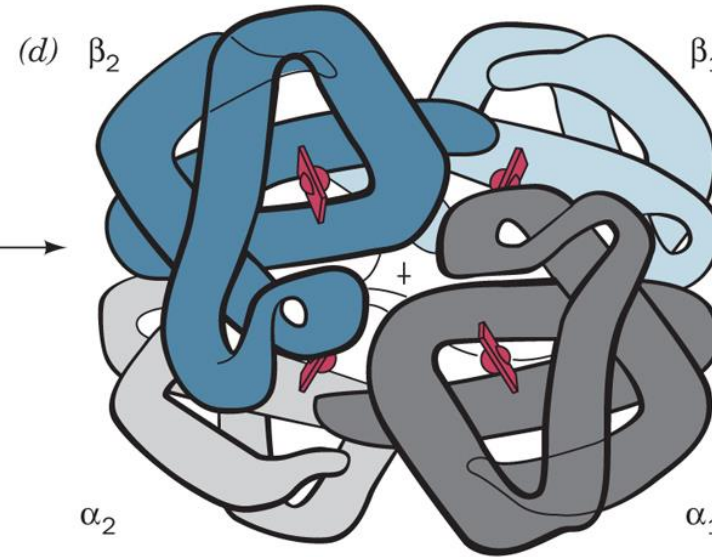
Primary 1^o Sequence



2^o



3^o



4^o

1° Structure

(a) – Lys – Ala – His – Gly – Lys – Lys – Val – Leu – Gly – Ala –
Primary structure (amino acid sequence in a polypeptide chain)

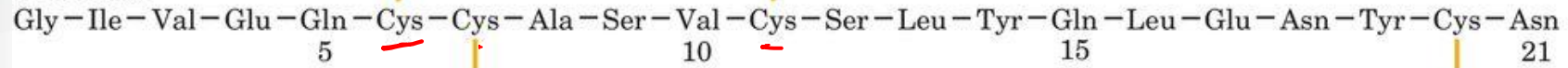
Illustration, Irving Geis. Image from the Irving Geis Collection/Howard Hughes Medical Institute. Rights owned by HHMI. Reproduction by permission only.

Kitchen

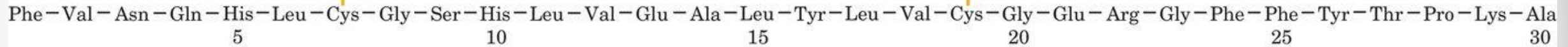
thicken

Bovine Insulin: 1° Structure

A chain

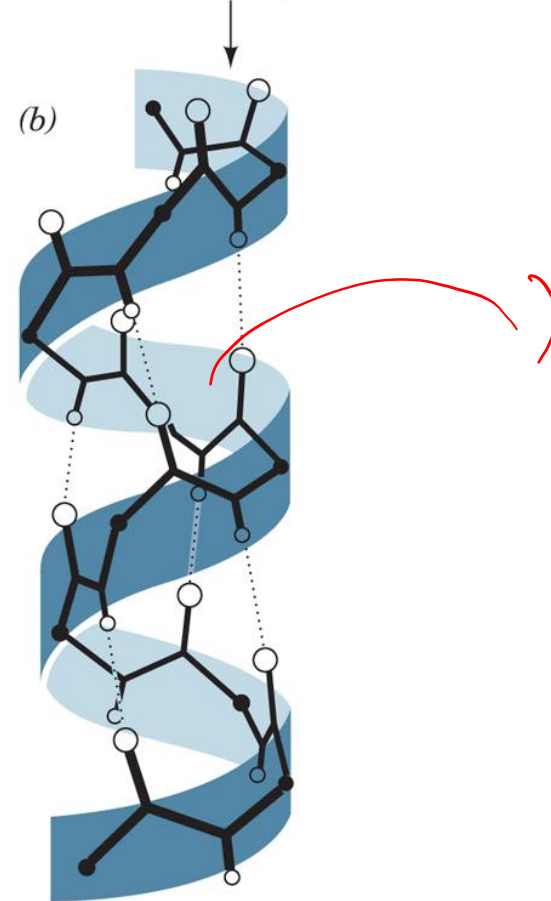


B chain



1° & 2° Structure

(a) – Lys – Ala – His – Gly – Lys – Lys – Val – Leu – Gly – Ala –
Primary structure (amino acid sequence in a polypeptide chain)

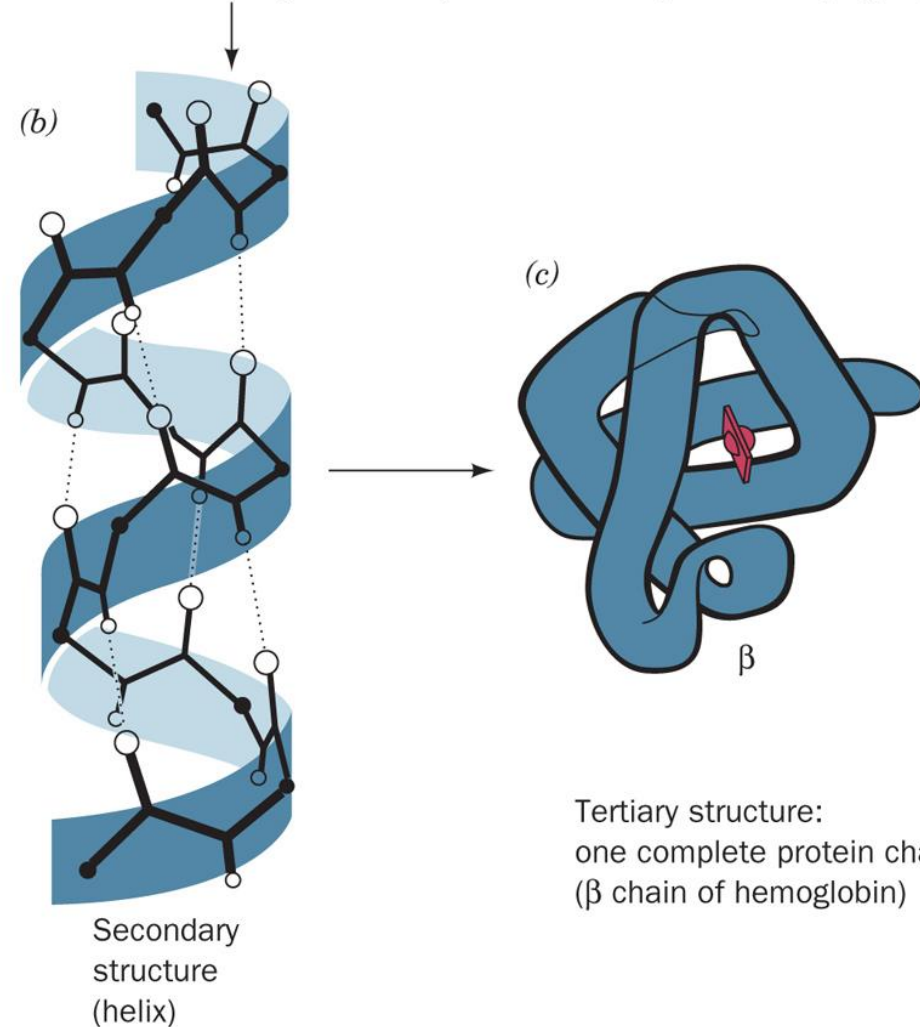


Secondary
structure
(helix)

Illustration, Irving Geis. Image from the Irving Geis Collection/Howard Hughes Medical Institute. Rights owned by HHMI. Reproduction by permission only.

1°, 2°, & 3° Structure

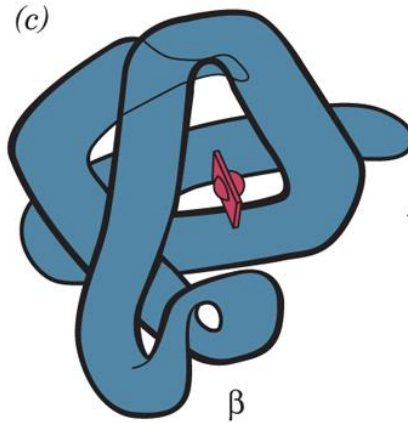
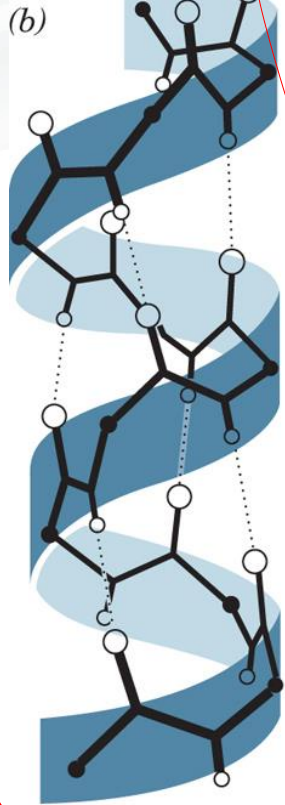
(a) – Lys – Ala – His – Gly – Lys – Lys – Val – Leu – Gly – Ala –
Primary structure (amino acid sequence in a polypeptide chain)



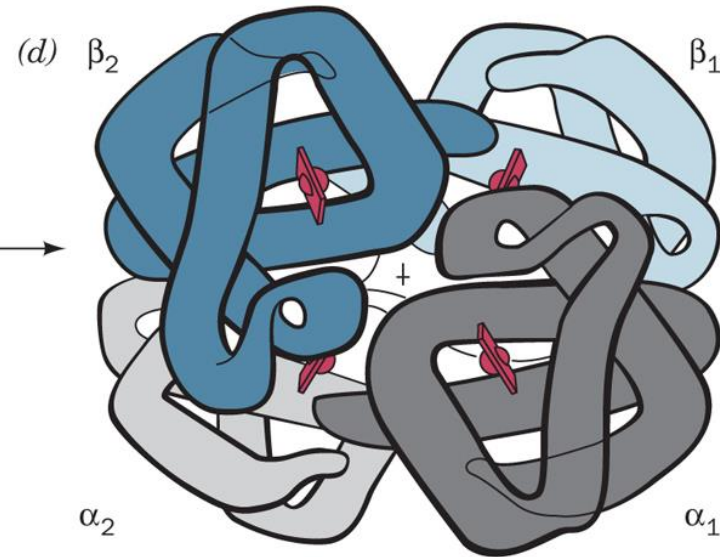
Illustration, Irving Geis. Image from the Irving Geis Collection/Howard Hughes Medical Institute. Rights owned by HHMI. Reproduction by permission only.

1°, 2°, 3° & 4° Structure

(a) – Lys – Ala – His – Gly – Lys – Lys – Val – Leu – Gly – Ala –
Primary structure (amino acid sequence in a polypeptide chain)



Tertiary structure:
one complete protein chain
(β chain of hemoglobin)

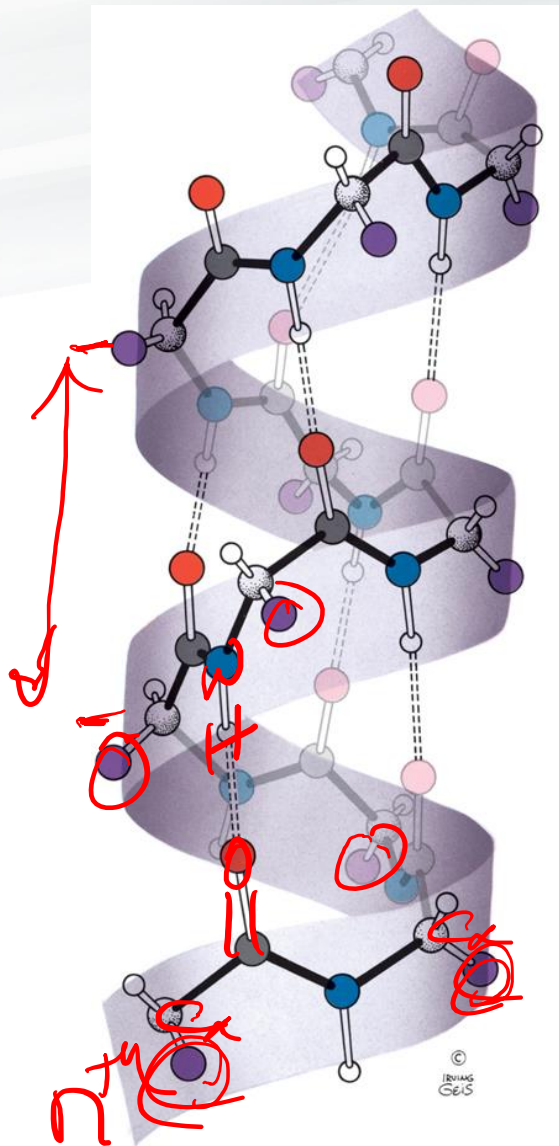


Quaternary structure:
the four separate chains
of hemoglobin assembled
into an oligomeric protein

©
IRVING
GEIS

Illustration, Irving Geis. Image from the Irving Geis Collection/Howard Hughes Medical Institute. Rights owned by HHMI. Reproduction by permission only.

The α Helix



Illustration, Irving Geis. Image from the Irving Geis Collection/Howard Hughes Medical Institute. Rights owned by HHMI. Reproduction by permission only.

Right-Handed α -Helix

Optimal $\phi + \psi$

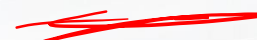
$$\phi = -57^\circ$$

$$\psi = -47^\circ$$

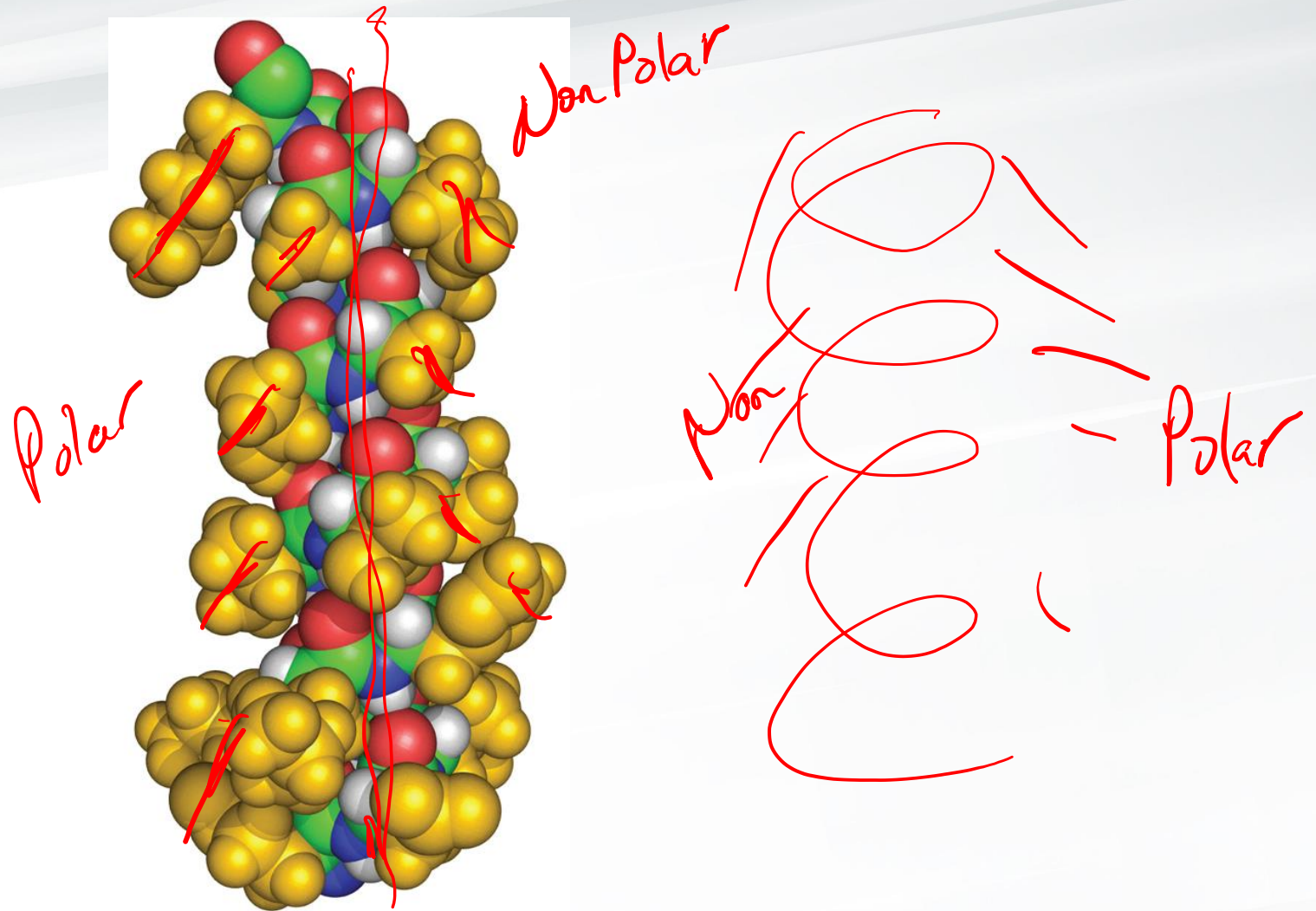
3.6 residues₀

pitch $\sim 5.4 \text{ \AA}$

$$n^{\text{th}} - n+4$$

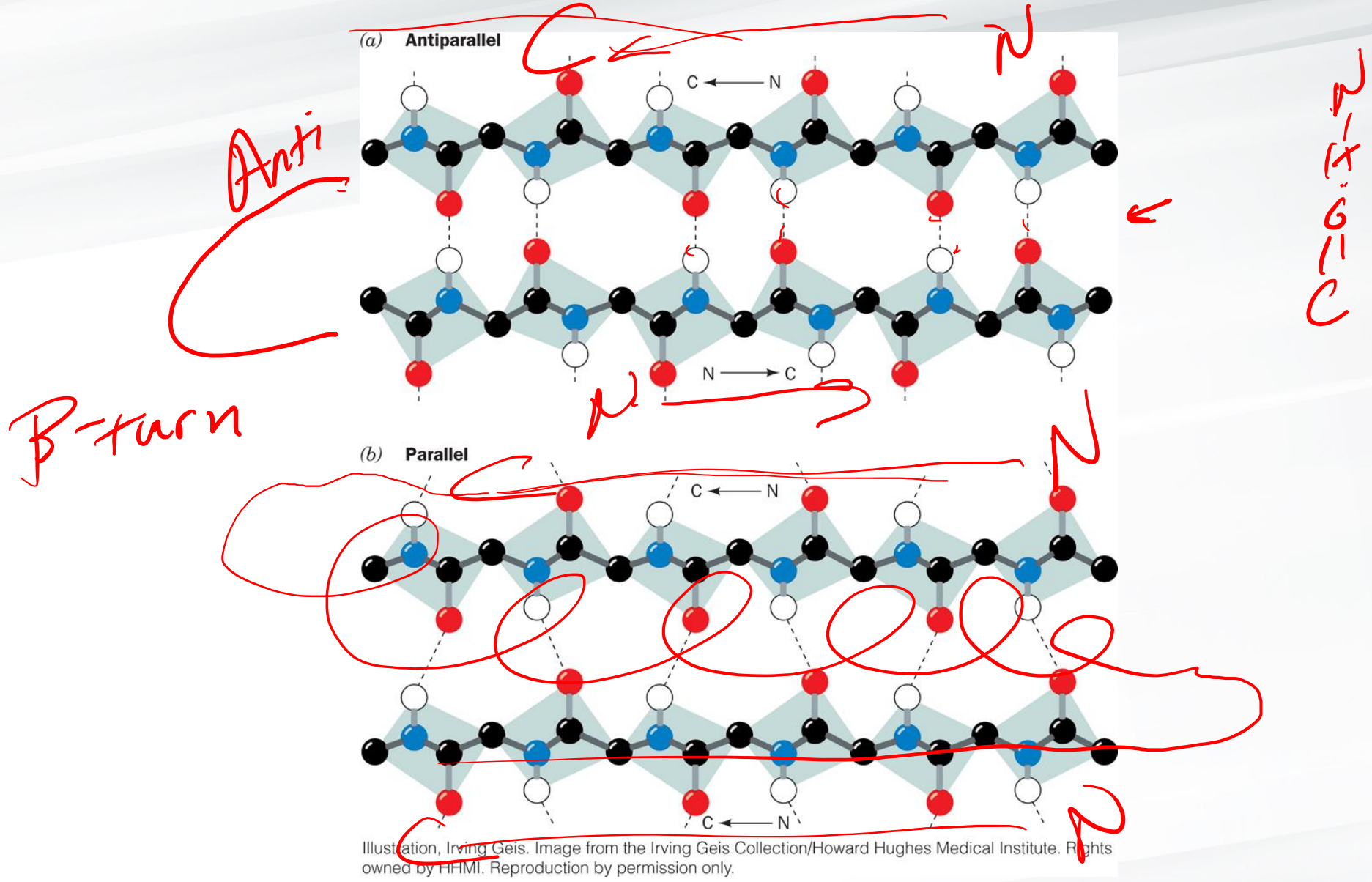


The α Helix: Space Filling Model

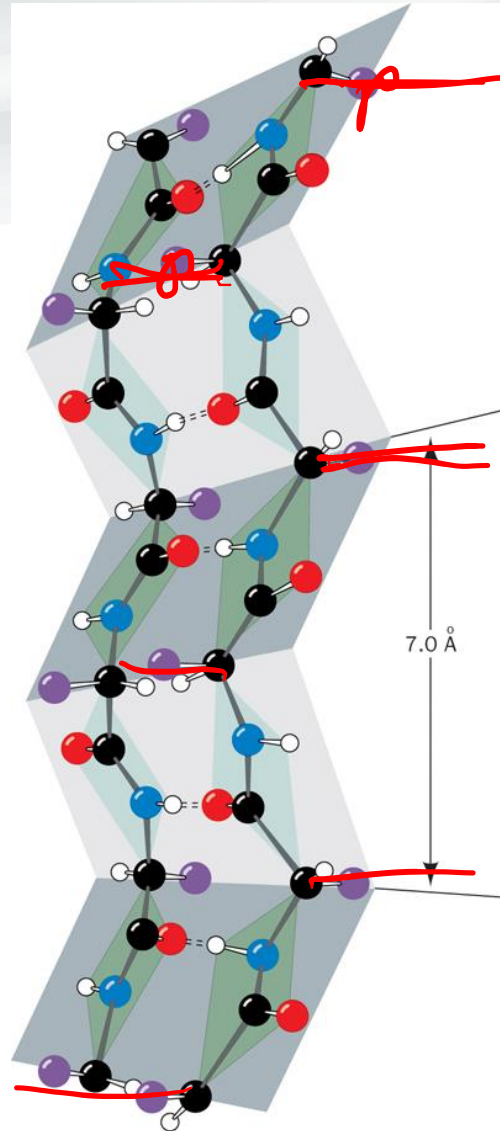


Based on an X-ray structure by Ilme Schlichting, Max Planck Institut für Molekulare Physiologie, Dortmund, Germany. PDBid 1A6M (for the definition of "PDBid" see Section 6-2E).

β Sheets

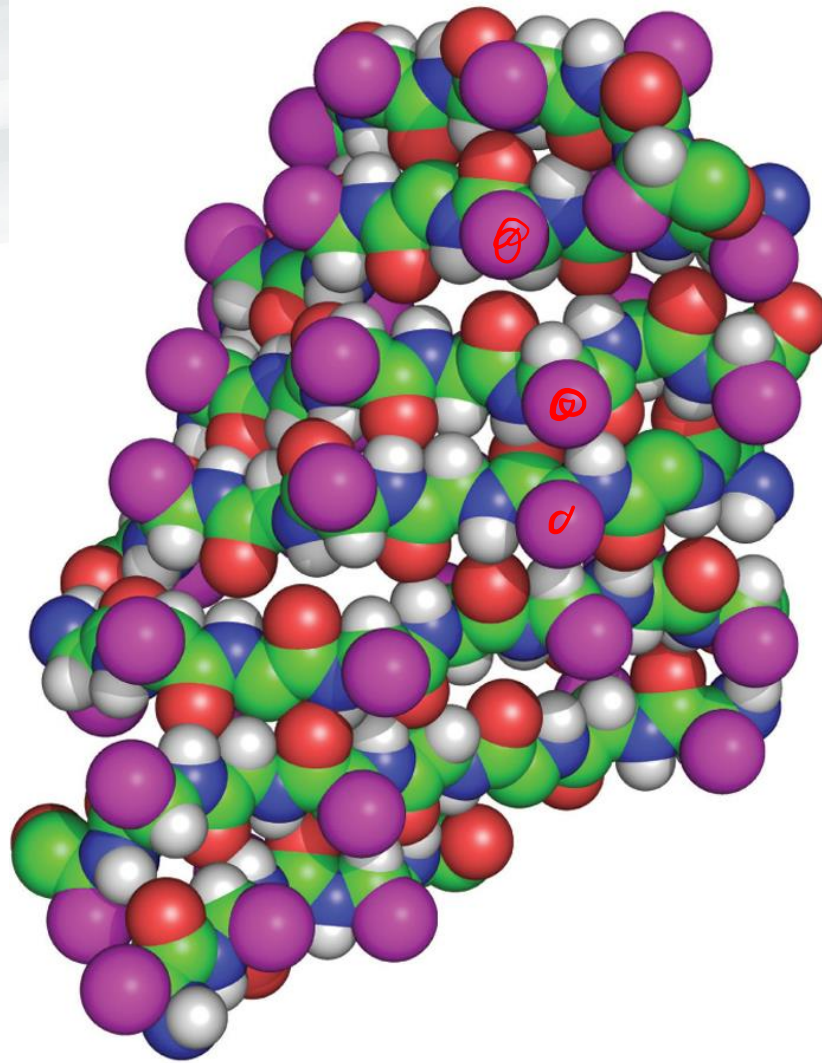


Pleating of β Sheet



Illustration, Irving Geis. Image from the Irving Geis Collection/Howard Hughes Medical Institute. Rights owned by HHMI. Reproduction by permission only.

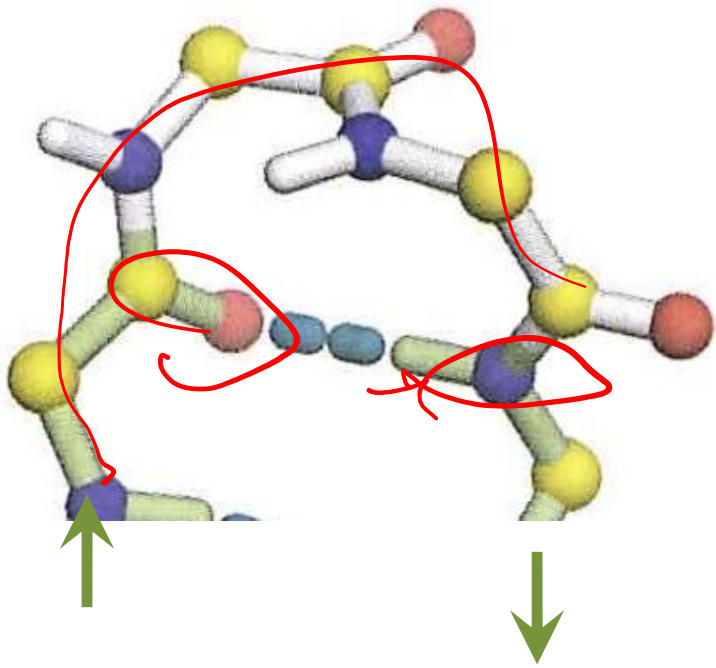
β Sheet: Space Filling Model



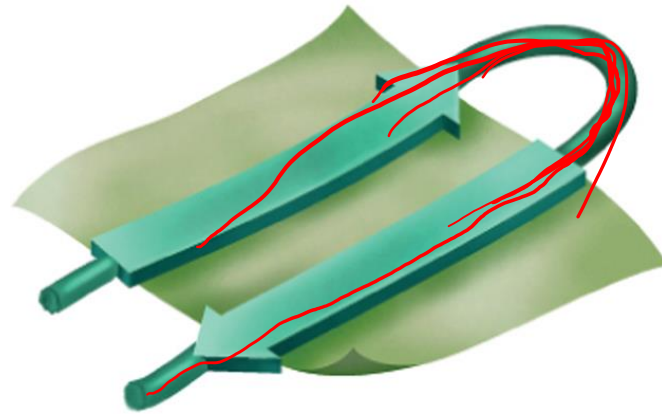
Based on an X-ray structure by Gerald Edelman,
The Rockefeller University. PDBid 2CNA.

Connecting Adjacent β Strands

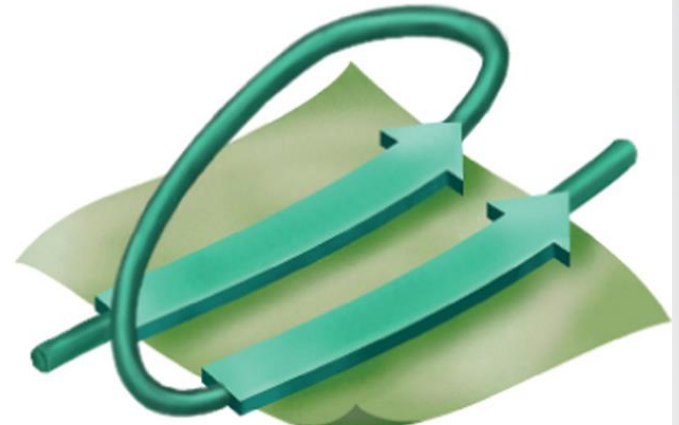
Beta Turns Reverse the Direction of the Backbone



(a)

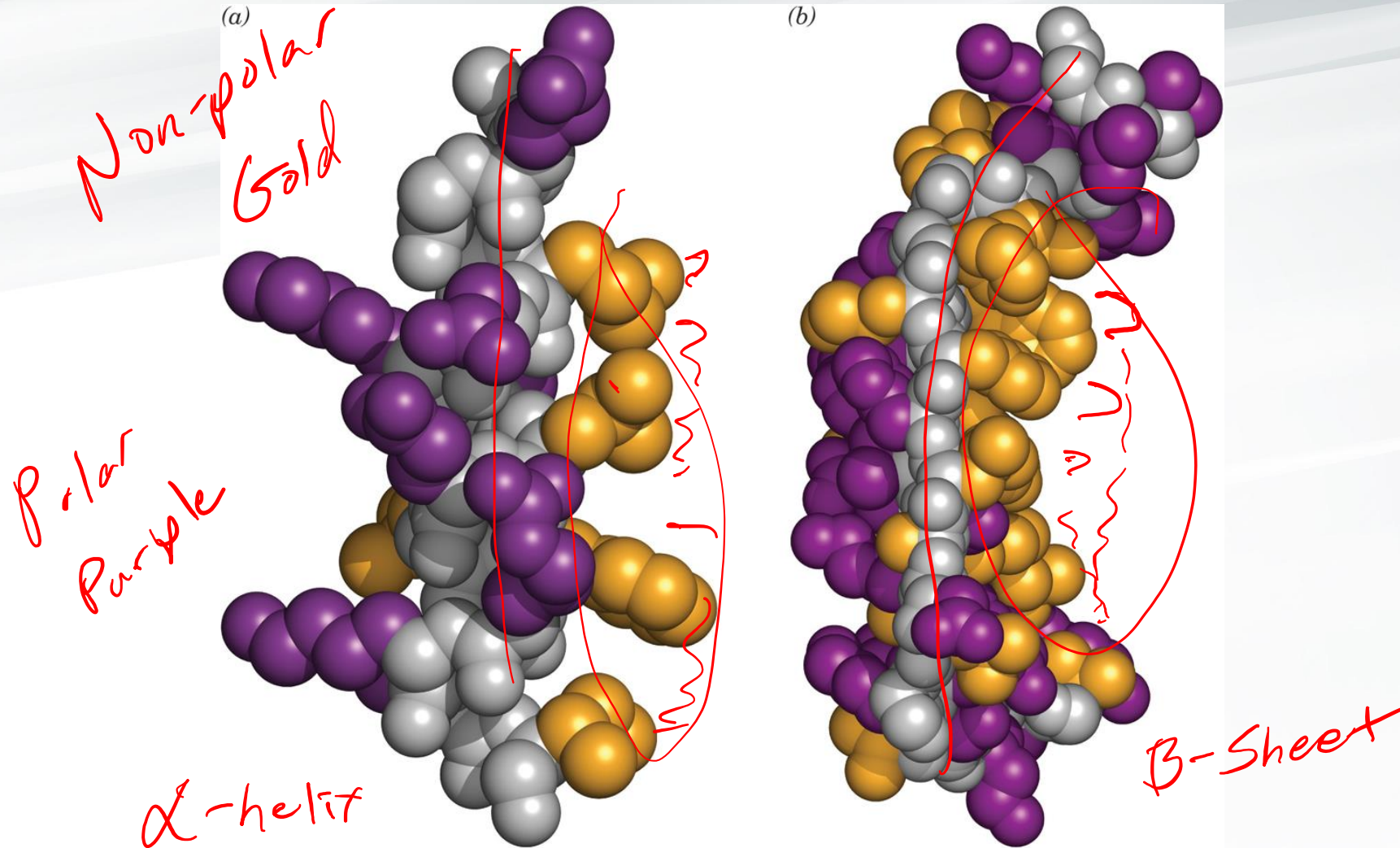


(b)



After Richardson, J.S., *Adv. Protein Chem.* 34, 196 (1981)

Side Chain Location

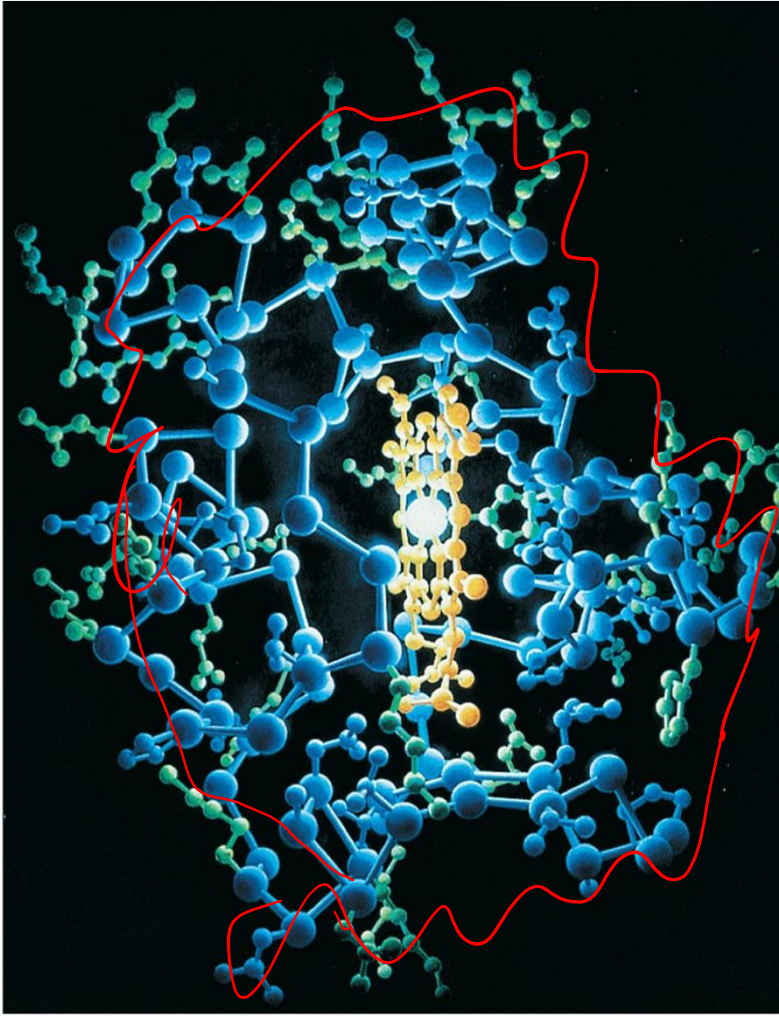


Based on X-ray structures by Ilme Schlichting, Max Planck Institut für Molekulare Physiologie, Dortmund, Germany, and Gerald Edelman, The Rockefeller University. PDBids 1A6M and 2CNA.

Oxy-Myoglobin and Concanavalin A
PDBids [1A6M](#) and [2CNA](#)

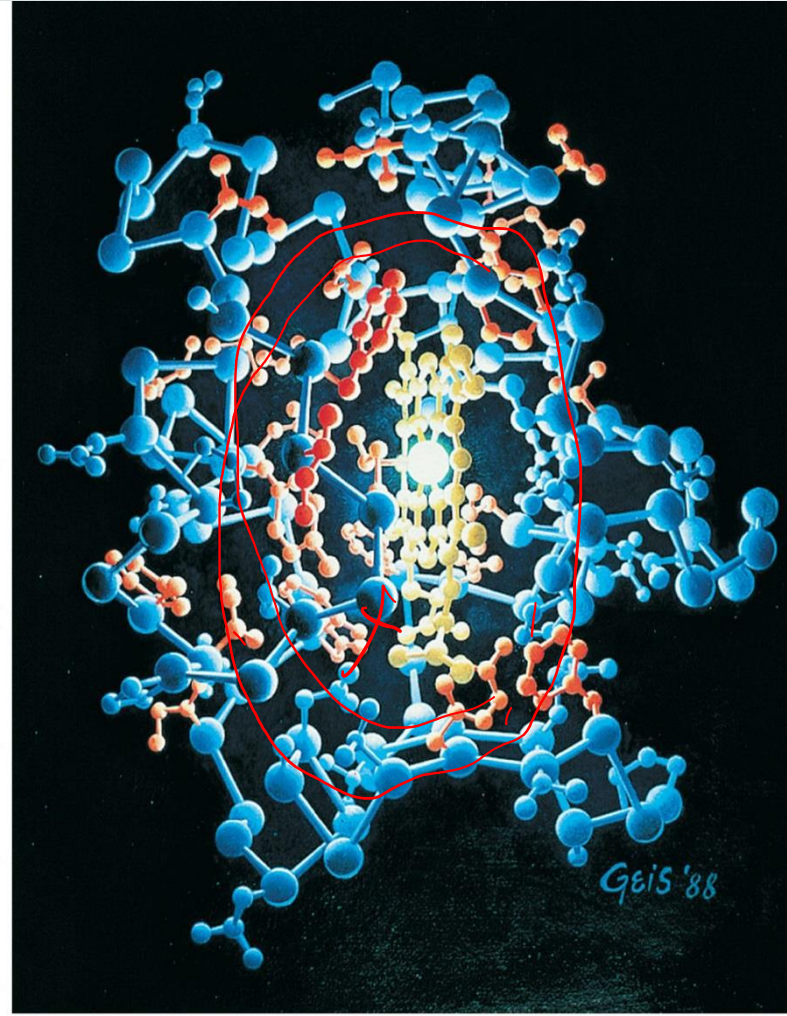
Side Chain Distribution in Cytochrome

Illustration, Irving Geis. Image from the Irving Geis Collection/Howard Hughes Medical Institute. Rights owned by HHMI. Reproduction by permission only



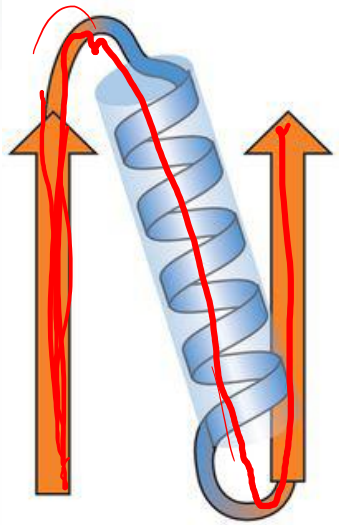
(a)

Illustration, Irving Geis. Image from the Irving Geis Collection/Howard Hughes Medical Institute. Rights owned by HHMI. Reproduction by permission only

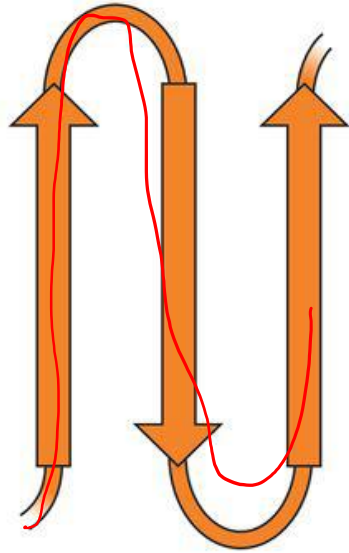


(b)

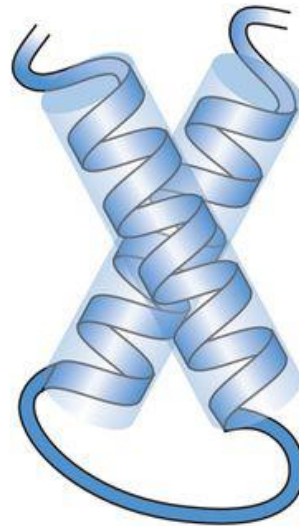
Motifs: Supersecondary Structures



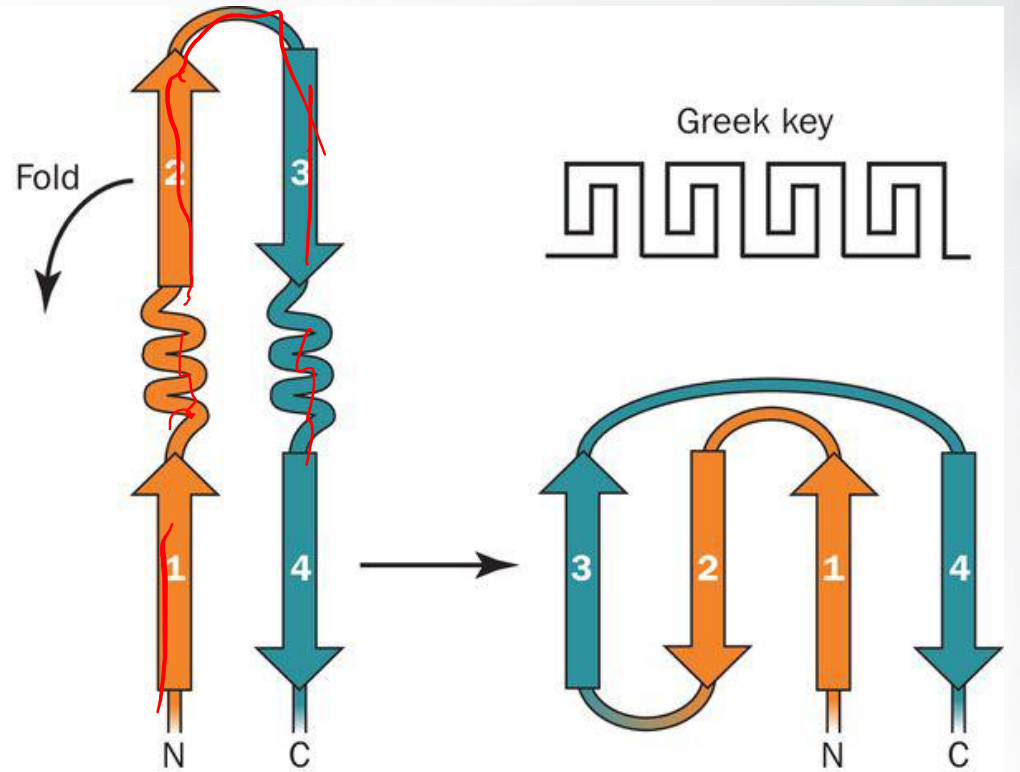
(a) *B₂B*



(b) *B₂B*

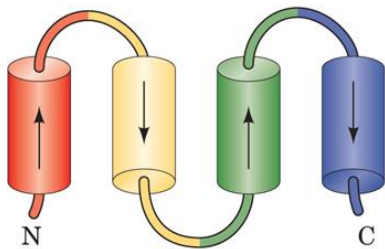
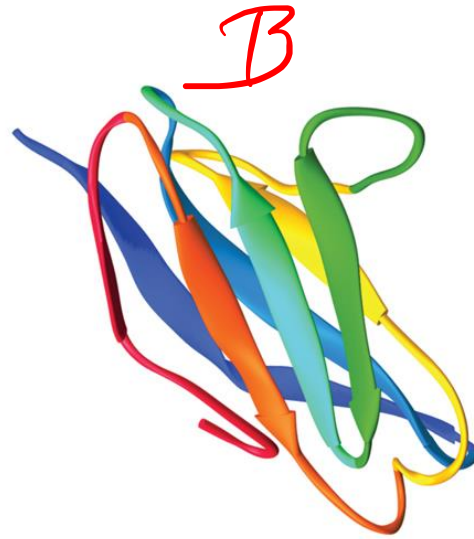
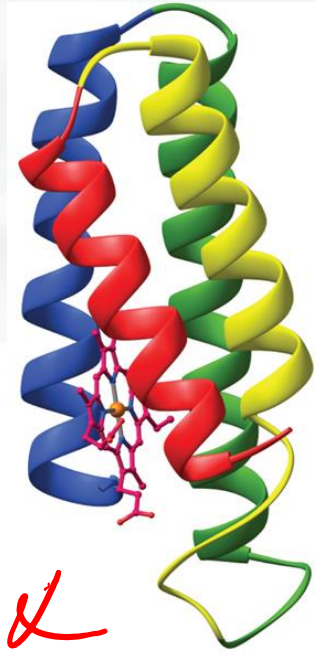


(c) *α*



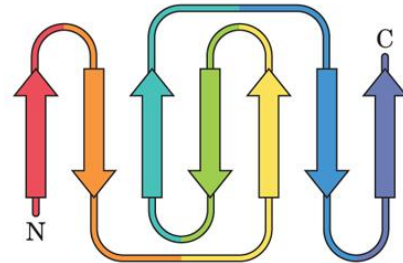
Greek Key

Protein Classification: α , β , or α/β

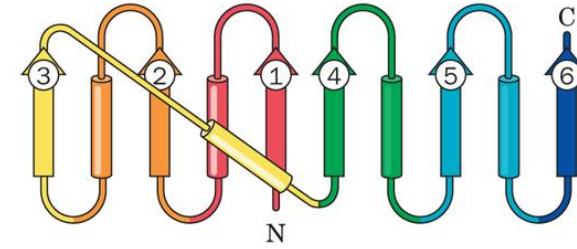


(a)

Based on X-ray structures by (a) F. Scott Matthews, Washington University School of Medicine; (b) Roberto Poljak, The Johns Hopkins School of Medicine; and (c) Michael Rossmann, Purdue University. PDBids (a) 256B, (b) 7FAB, and (c) 6LDH.



(b)



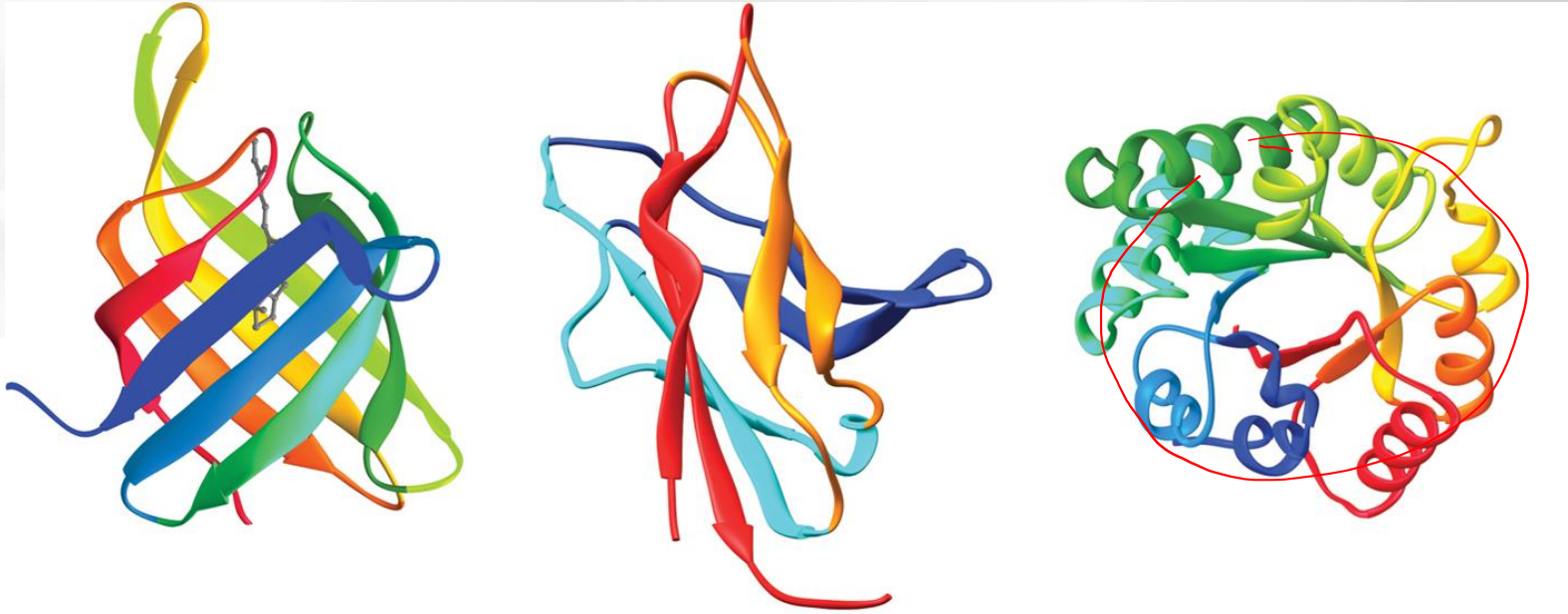
(c)

Cytochrome *b562*
PDBid [256B](#)

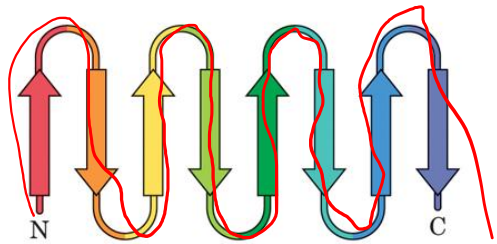
Human immunoglobulin fragment
PDBid [7FAB](#)

Dogfish lactate dehydrogenase
PDBid [6LDH](#)

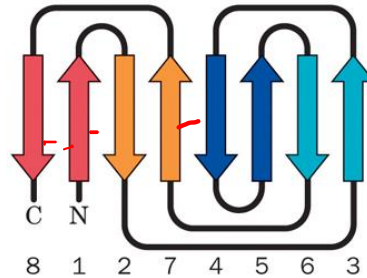
Protein Topology: 8-Stranded β Barrels



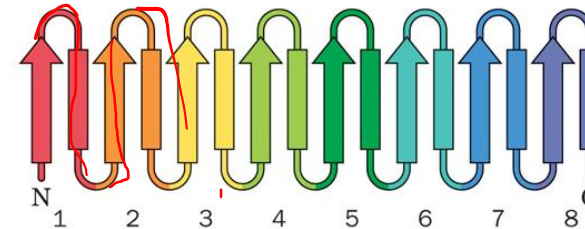
2/B



(a)



(b)



(c)

Based on X-ray structures by (a) T. Alwyn Jones, Biomedical Center, Uppsala, Sweden; (b) Patrick Van Roey, New York State Department of Health, Albany, New York; and (c) David Phillips, Oxford University, Oxford, U.K. PDBids (a) 1RBP, (b) 1PNG, and (c) 1TIM.

Peptide-N⁴-(N-acetyl- β -D-glucosaminyl) asparagine amidase

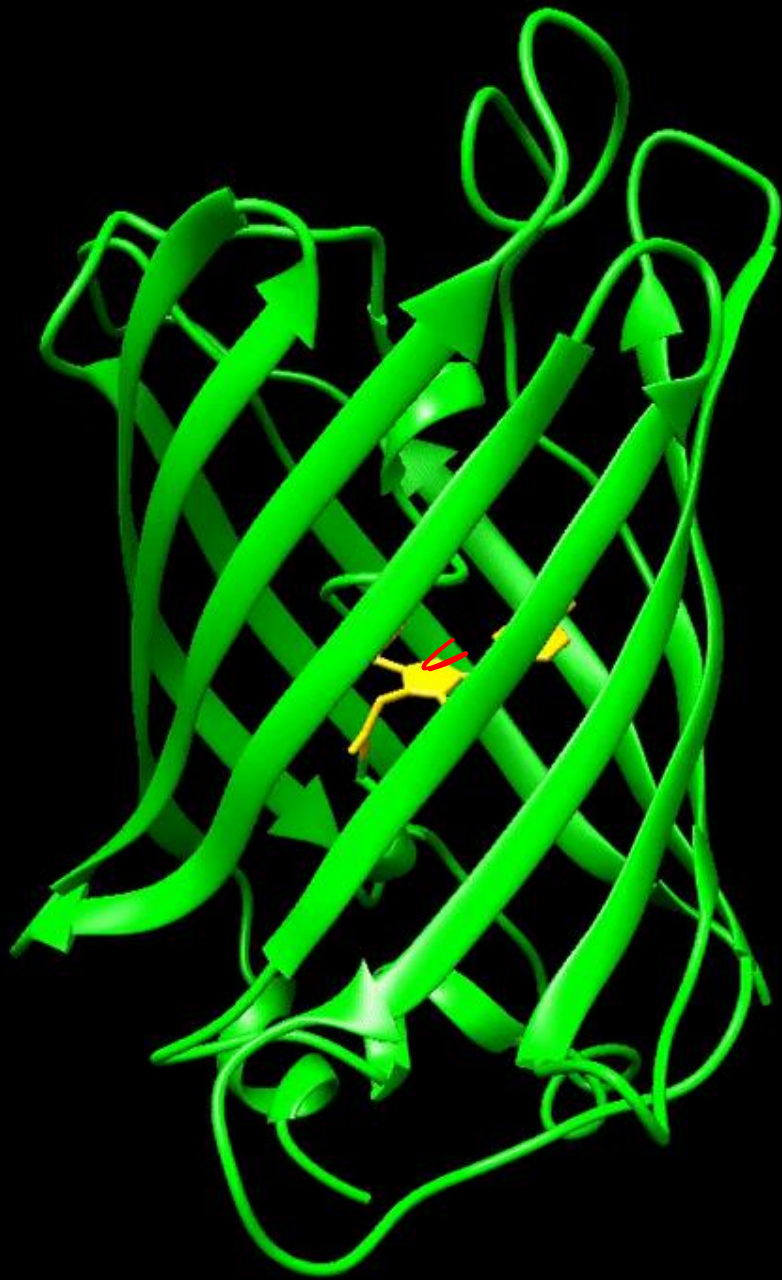
PDBid [1PNG](#)

Human retinol binding protein

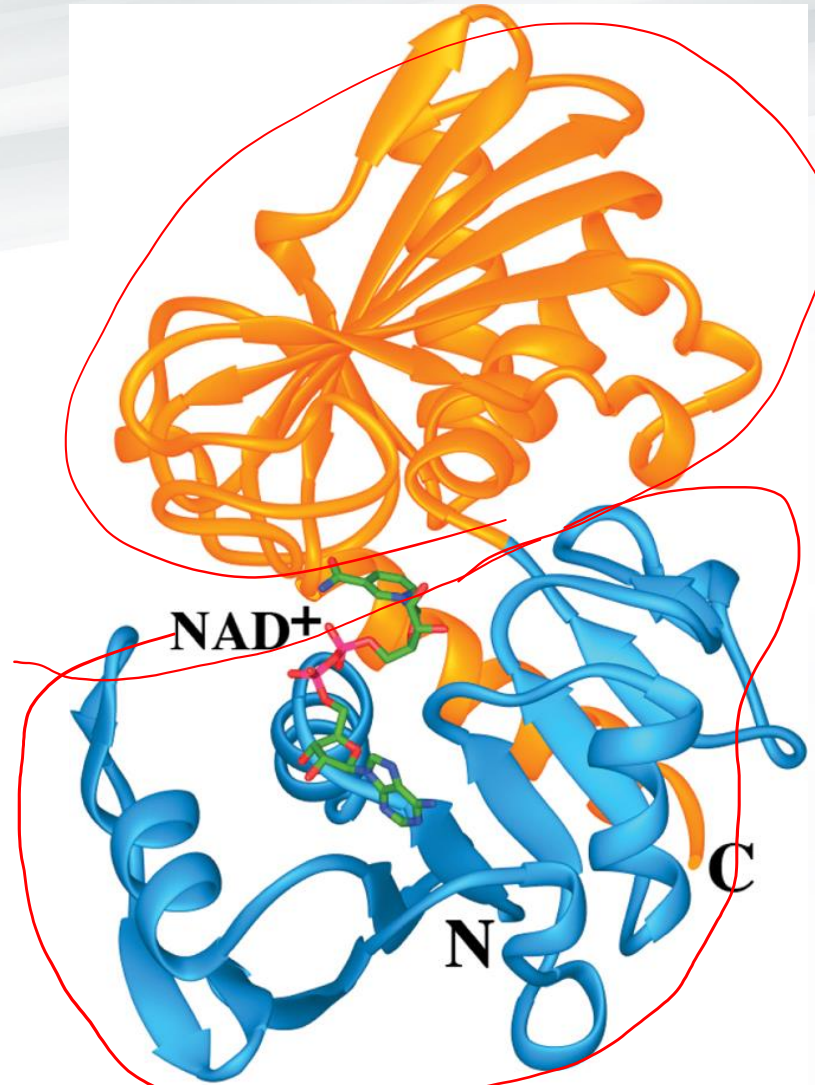
PDBid [1RBP](#)

Triose phosphate isomerase

PDBid [1TIM](#)



2-Domain Protein : GAPDH

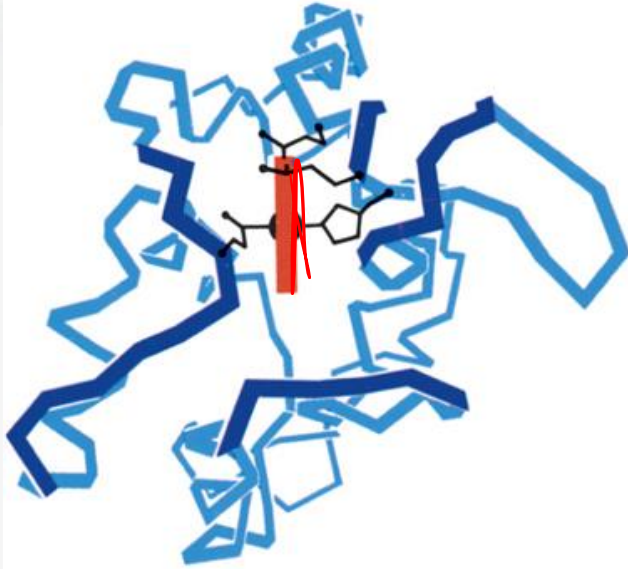


Based on an X-ray structure by Alan Wonacott, Imperial College, London, U.K. PDBid 1GD1.

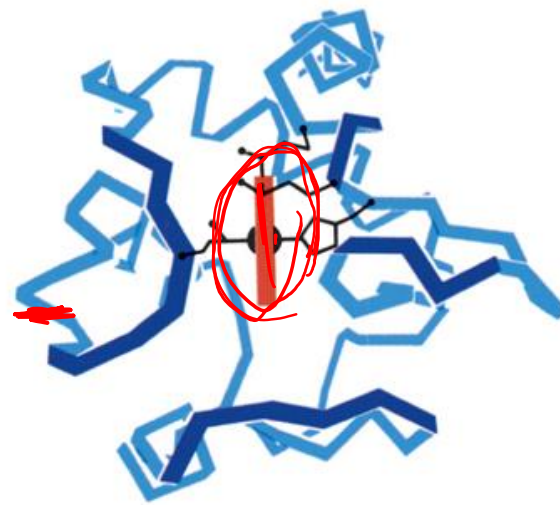
Glyceraldehyde-3-phosphate dehydrogenase
PDBid [1GD1](#)

Structure Conserved More Than Sequence

(a) *Paracoccus c₅₅₀*
134 amino acid residues



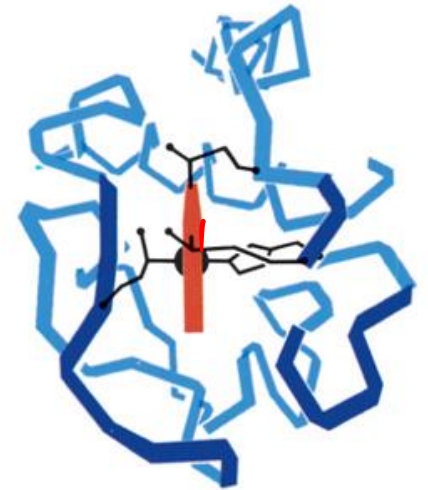
(b) *Rhodospirillum c₂*
112 amino acid residues



(c) Tuna *c*
103 amino acid residues

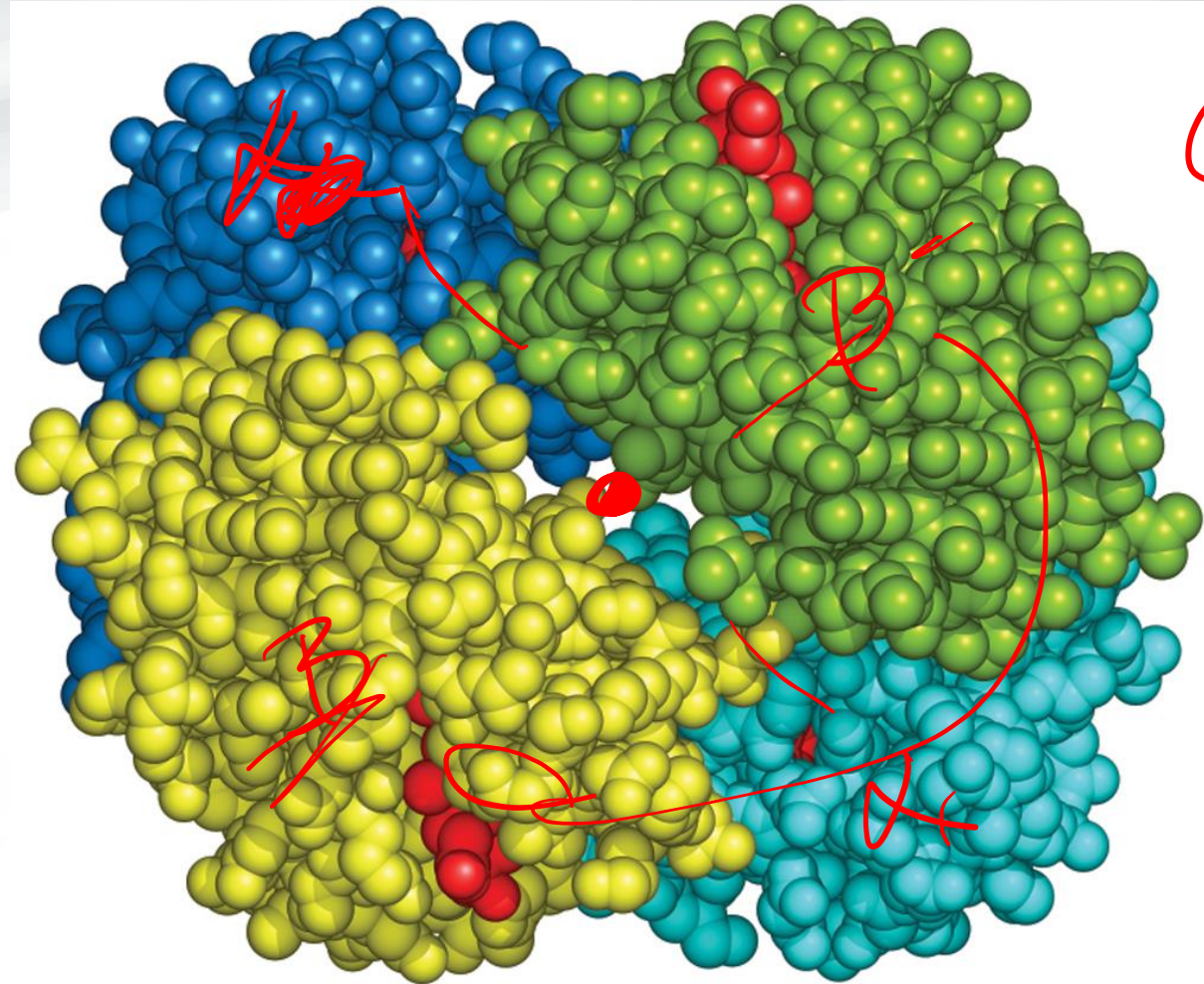
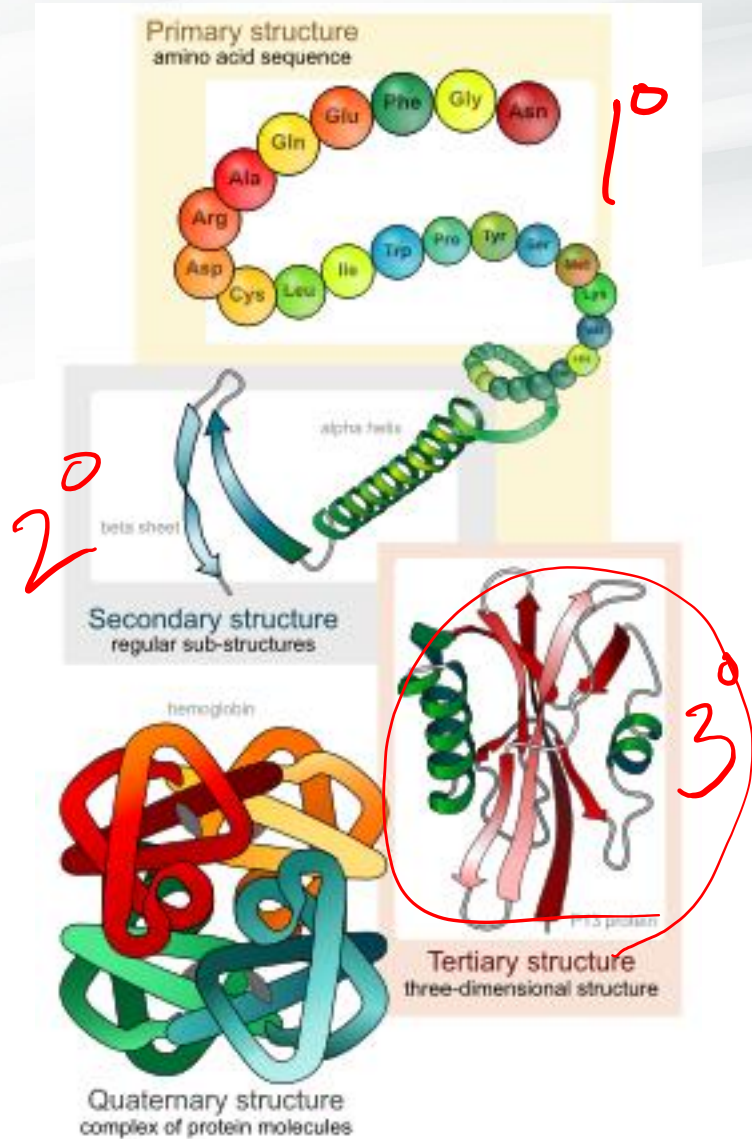


(d) *Chlorobium c₅₅₅*
86 amino acid residues



Illustration, Irving Geis. Image from the Irving Geis Collection/Howard Hughes Medical Institute. Rights owned by HHMI. Reproduction by permission only.

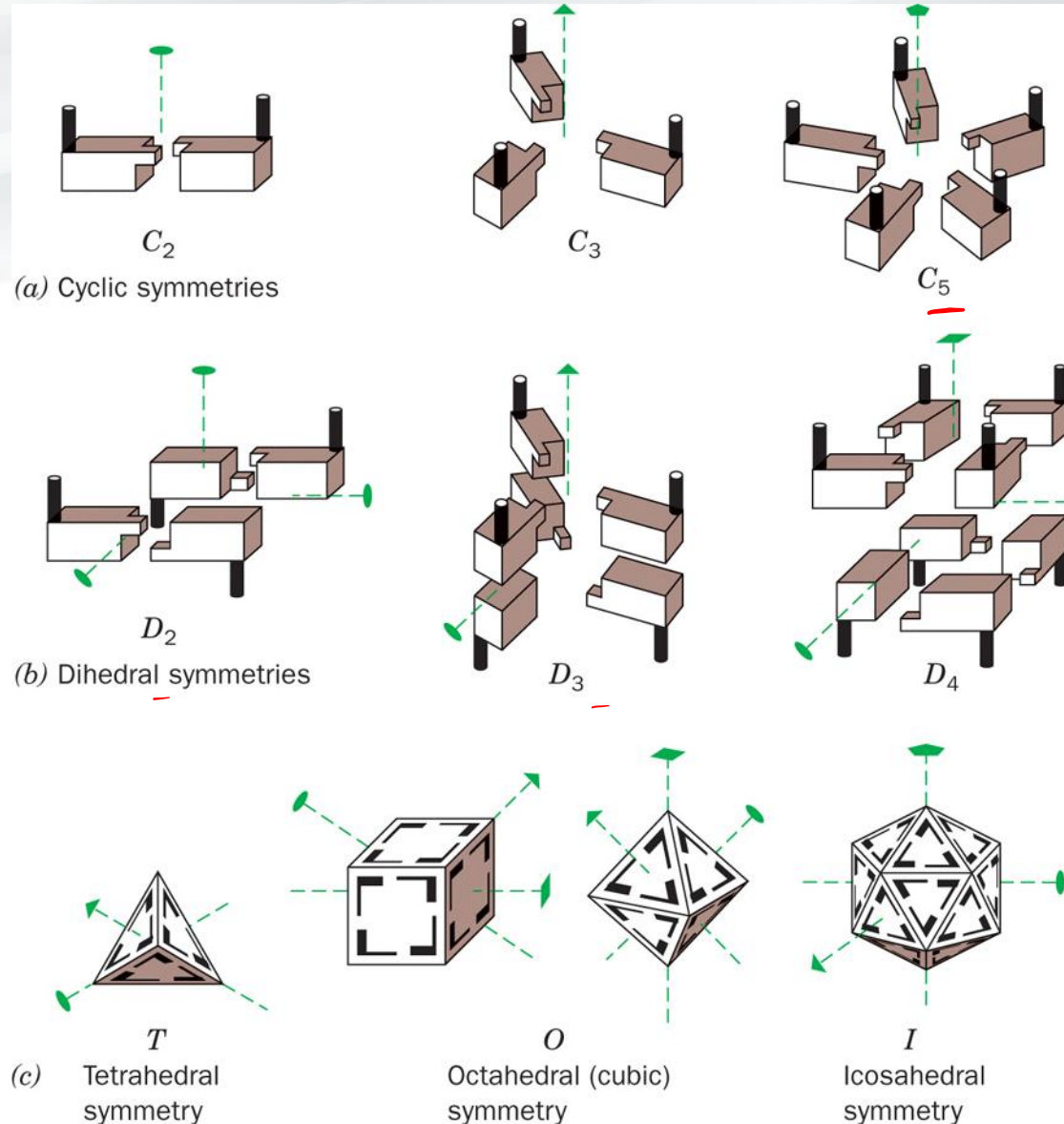
4^o Structure of Hemoglobin



Based on an X-ray structure by Max Perutz, MRC Laboratory of Molecular Biology, Cambridge, U.K. PDBid 2DHB.

Deoxyhemoglobin
PDBid [2DHB](#)

Symmetries of Oligomeric Proteins



Illustration, Irving Geis. Image from the Irving Geis Collection/Howard Hughes Medical Institute. Rights owned by HHMI. Reproduction by permission only.

Summary

- Proteins are chemical chains made up of 20 different amino acids
- The chemical structure of the backbone has consequences: helices, sheets, and turns
- Protein structure is classified into a four-level hierarchy