Review II: The Molecules of Life

Judy Wieber

BBSI @ Pitt 2007

Department of Computational Biology
University of Pittsburgh School of Medicine

May 24, 2007

Outline

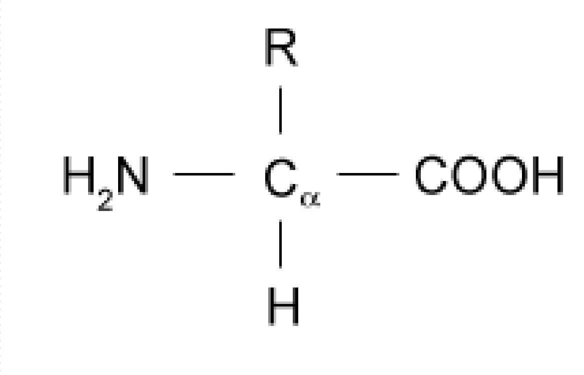
- Introduction
- Proteins
- Carbohydrates
- Lipids
- Nucleic acids
- Special types of proteins

Proteins

- Amino acids
- Peptide bond formation
- Peptides vs. proteins
- Hierarchy of protein structure
- Protein motifs
- Domains

Amino Acids

- Building blocks of proteins
- 20 amino acids
- Linear chain of amino acids form a peptide/protein
- α–carbon = central carbon
- α–carbon = chiral carbon, i.e. mirror images: L and D isomers
- Only L isomers found in proteins
- General structure:

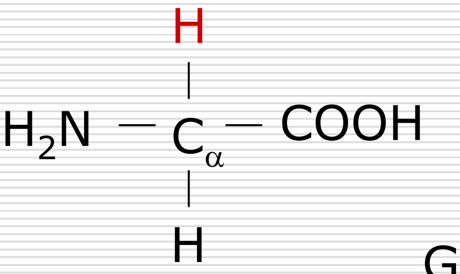


$$H_2N - \overset{\displaystyle R}{\underset{\displaystyle H}{C_\alpha}} - COOH$$

(R = side chain)

Amino Acids (contd.)

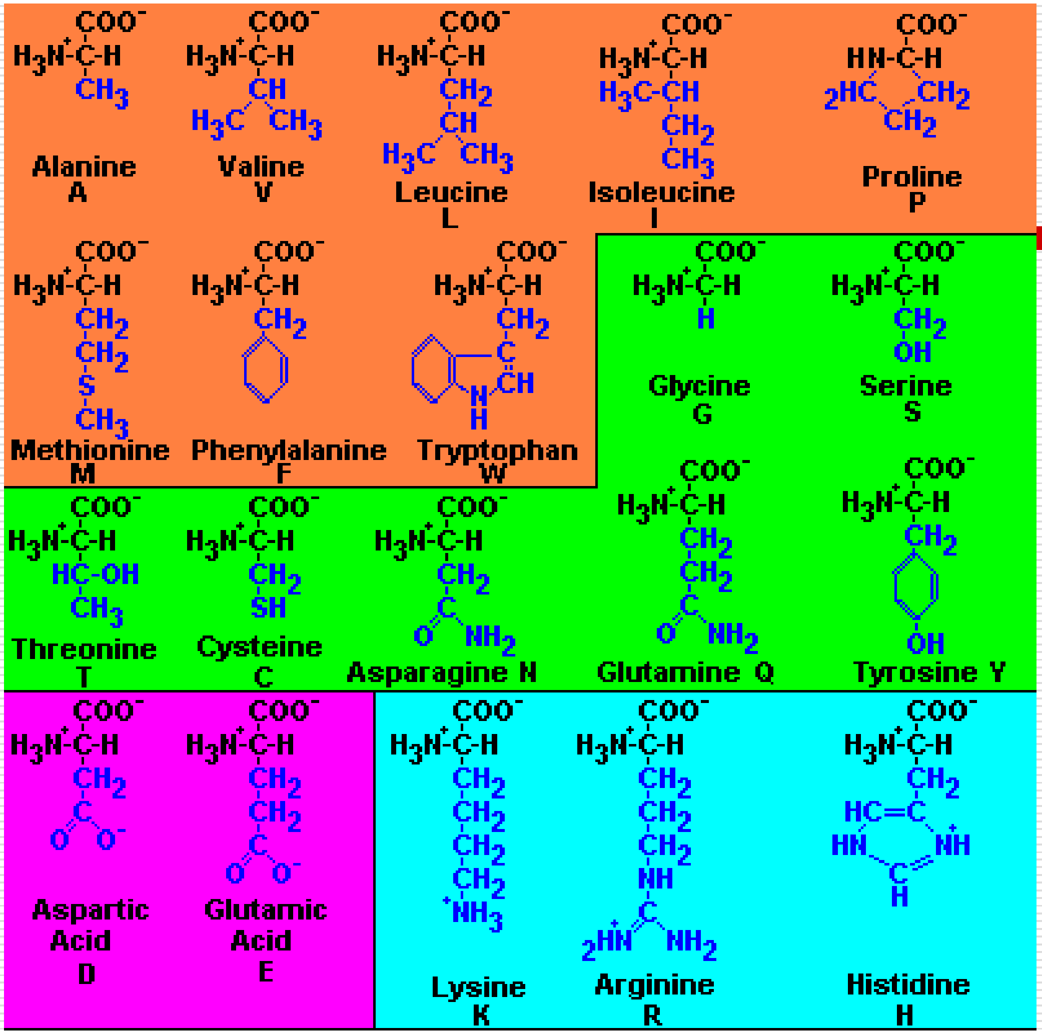
- R group varies
- Thus, can be classified based on R group
- Glycine: simplest amino acid
- Side chain R = H
- Unique because Gly α carbon is *achiral*



$$\mathrm{H_2N} - \overset{\mathrm{H}}{\overset{|}{\underset{\underset{\mathrm{H}}{|}}{\mathrm{C}_\alpha}}} - \mathrm{COOH}$$

Glycine, Gly, G

Amino Acids: Structures



blue = R (side chain)

orange = non-polar, hydrophobic
neutral (uncharged)

green = polar, hydrophilic,
neutral (uncharged)

magenta = polar, hydrophilic,
acidic (- charged)

light blue = polar, hydrophilic,
basic (+ charged)

Amino Acids: Classification

Non-polar, hydrophobic,
neutral (uncharged)

Alanine, Ala, A
Valine, Val, V
Leucine, Leu, L
Isoleucine, Ile, I
Proline, Pro, P
Methionine, Met, M
Phenylalanine, Phe, F
Tryptophan, Trp, W

Polar, hydrophilic,
neutral (uncharged)

Glycine, Gly, G
Serine, Ser, S
Threonine, Thr, T
Cysteine, Cys, C
Asparagine, Asn, N
Glutamine, Gln, Q
Tyrosine, Tyr, Y

Polar, hydrophilic,
Acidic (negatively charged)

Aspartic acid, Asp, D
Glutamic acid, Glu, E

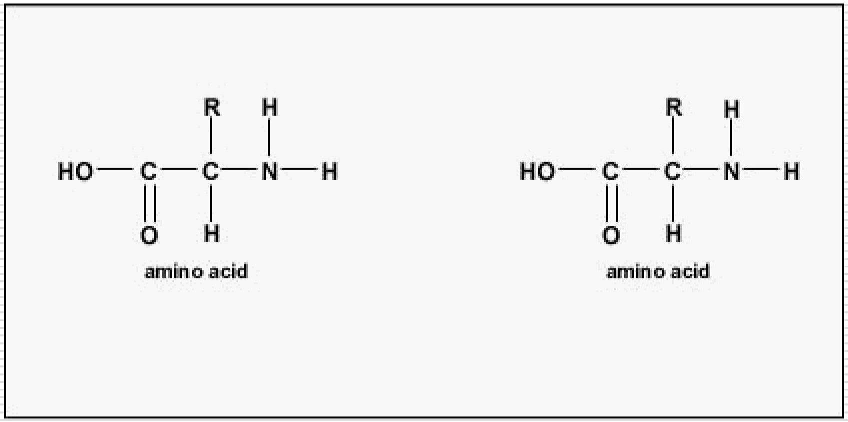
(Amino acid viewer)

Polar, hydrophilic,
basic (positively charged)

Lysine, Lys, K
Arginine, Arg, R
Histidine, His, H

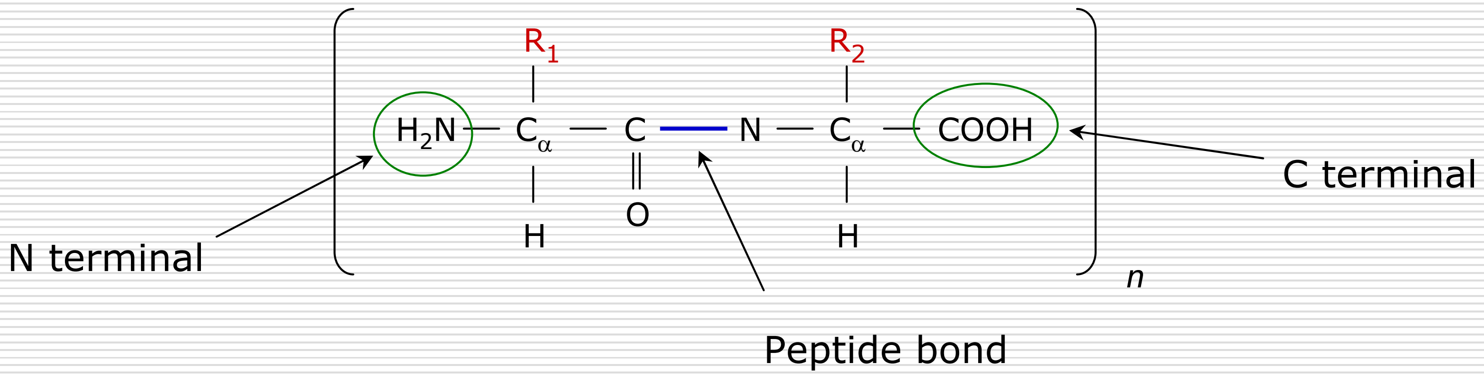
Peptide Bond Formation

- Condensation reaction
- Between –NH_2 of *n* residue and –COOH of *n+1* residue
- Loss of 1 water molecule
- Rigid, inflexible



Peptides/Proteins

- Linear arrangement of *n* amino acid residues linked by peptide bonds
- *n* < 25, generally termed a peptide
- *n* > 25, generally termed a protein
- Peptides have *directionality, i.e. N terminal* → *C-terminal*



Polypeptide Backbone

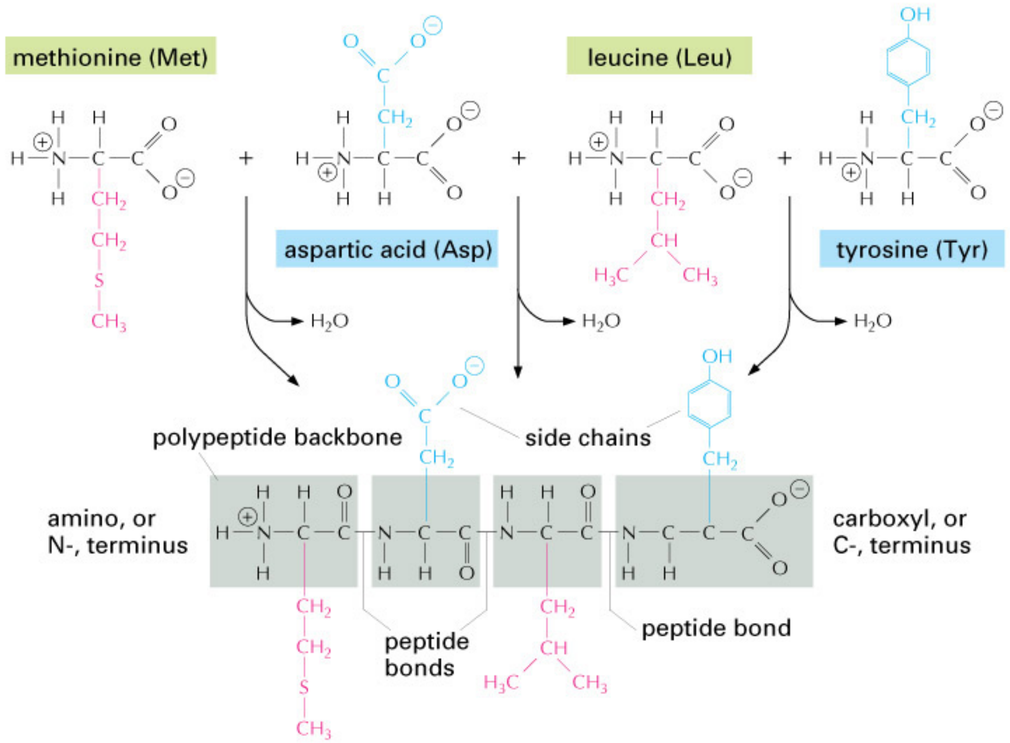


Figure 4-2 Essential Cell Biology, 2/e. (© 2004 Garland Science)

Binding of Ligands

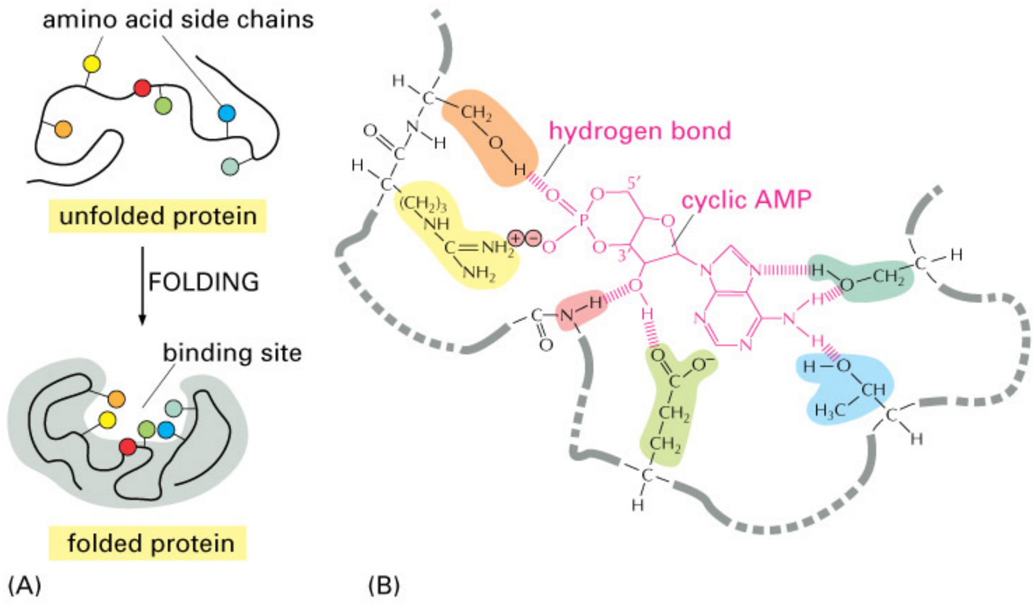


Figure 4-31 Essential Cell Biology, 2/e. (© 2004 Garland Science)

Hierarchy of Protein Structure

- Four levels of hierarchy
- Primary, secondary, tertiary, quarternary

- **Primary structure**: Linear sequence of residues
- e.g: MSNKLVLVLNCGSSSLKFAV …
- e.g: MCNTPTYCDLGKAAKDVFNK …

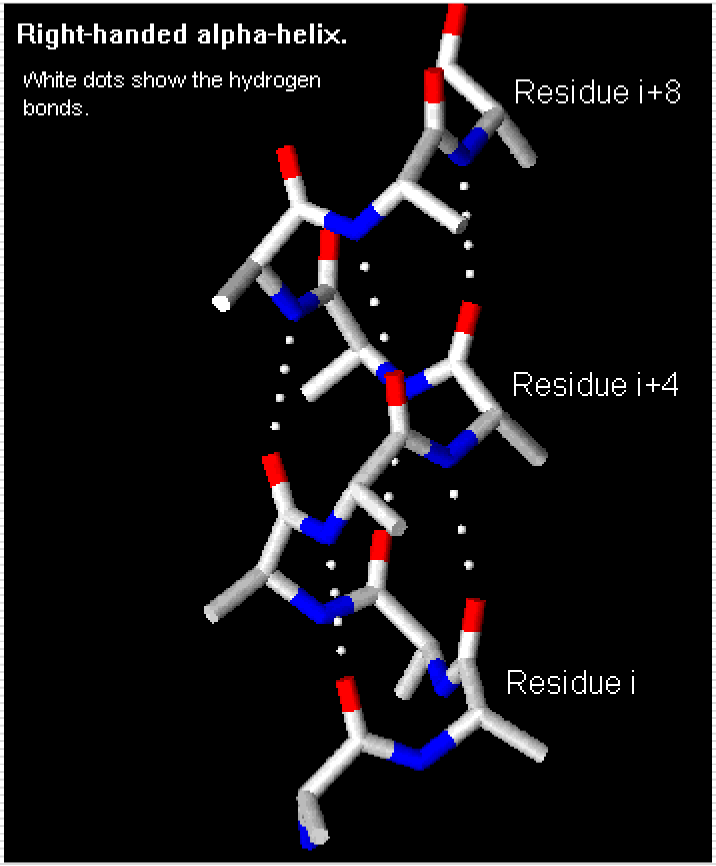
- **Secondary Structure**: Local conformation of the polypeptide backbone
- α-helix, β-strand (sheets), turns, other

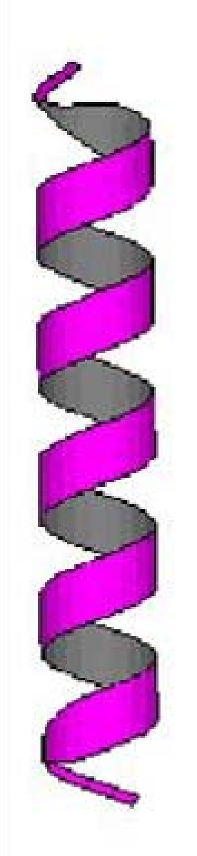
(Amino acids-proteins movie)

Secondary Structure: α-helix

- Most abundant; ~35% of residues in a protein
- *Repetitive secondary structure*
- 3.6 residues per turn; pitch (rise per turn) = 5.4 Å
- C′=O of *i* forms H bonds with NH of residue *i+4*
- *Intra-strand* H bonding
- C′=O groups are parallel to the axis; side chains point away from the axis
- All NH and C′O are H-bonded, except first NH and last C′O
- Hence, polar ends; present at surfaces
- Amphipathic

α-helix (contd.)



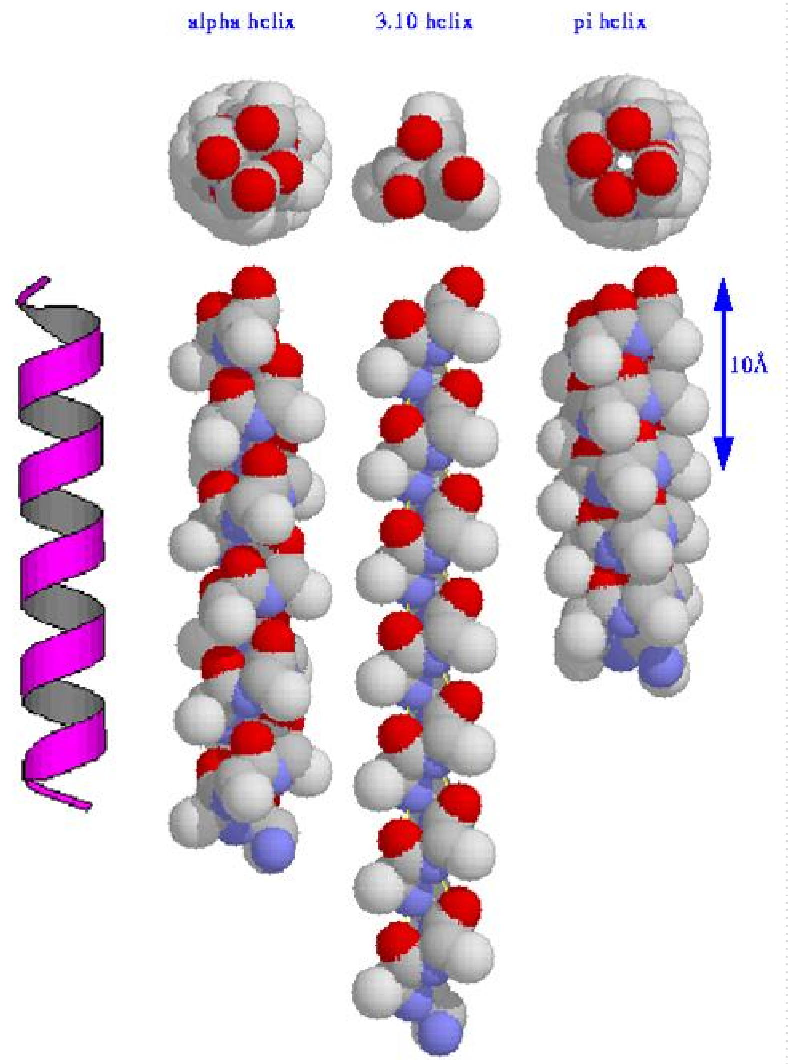


C terminal

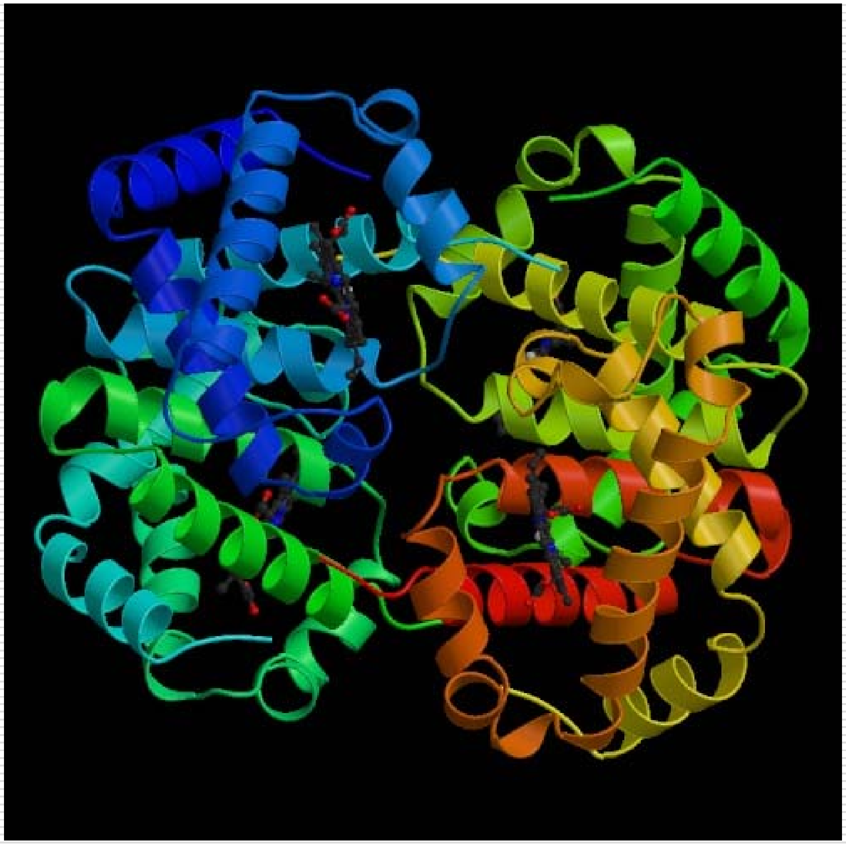
N terminal

α–helix Variations

- Chain is more loosely or tightly coiled
- *3_{10}-helix*: very tightly packed
- *π–helix*: very loosely packed
- Both structures occur rarely
- Occur only at the ends or as single turns



Hemoglobin (PDB 1A3N)



β–sheets

- Other major structural element
- Basic unit is a β-strand
- Usually 5-10 residues
- Can be parallel or anti-parallel based on the relative directions of interacting β-strands
- “Pleated” appearance

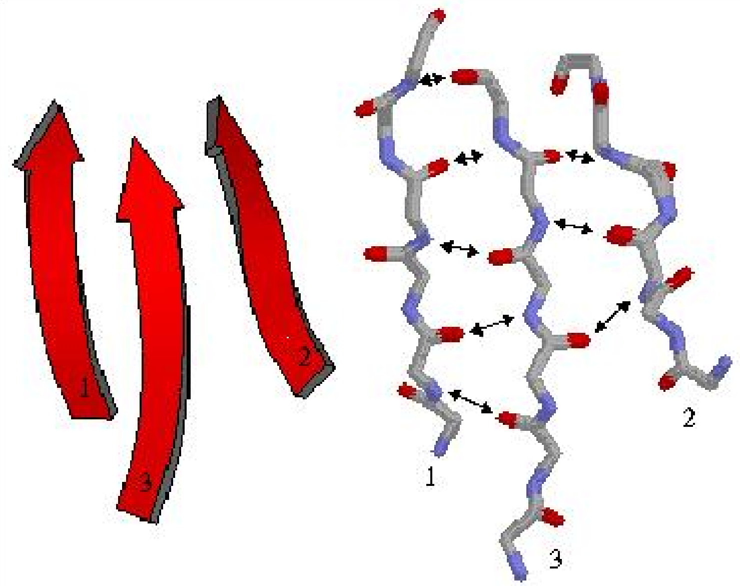
β–sheets

- ***Unlike* α-helices:**
- Are formed with different parts of the sequence
- H-bonding is *inter-strand* (opposed to intra-strand)
- Side chains from adjacent residues are on opposite sides of the sheet and do not interact with one another

- ***Like* α-helices:**
- Repeating secondary structure (2 residues per turn)
- Can be amphipathic

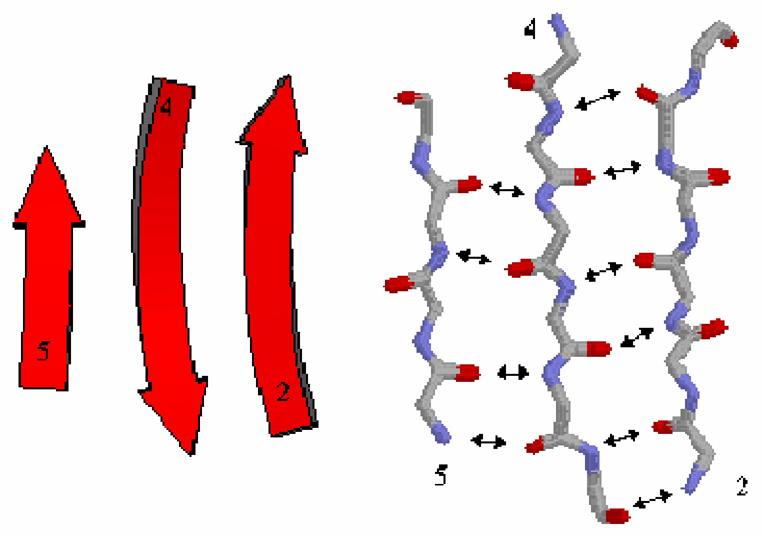
Parallel β–sheets

- The aligned amino acids in the β-strand all run in the same biochemical direction, N- to C-terminal

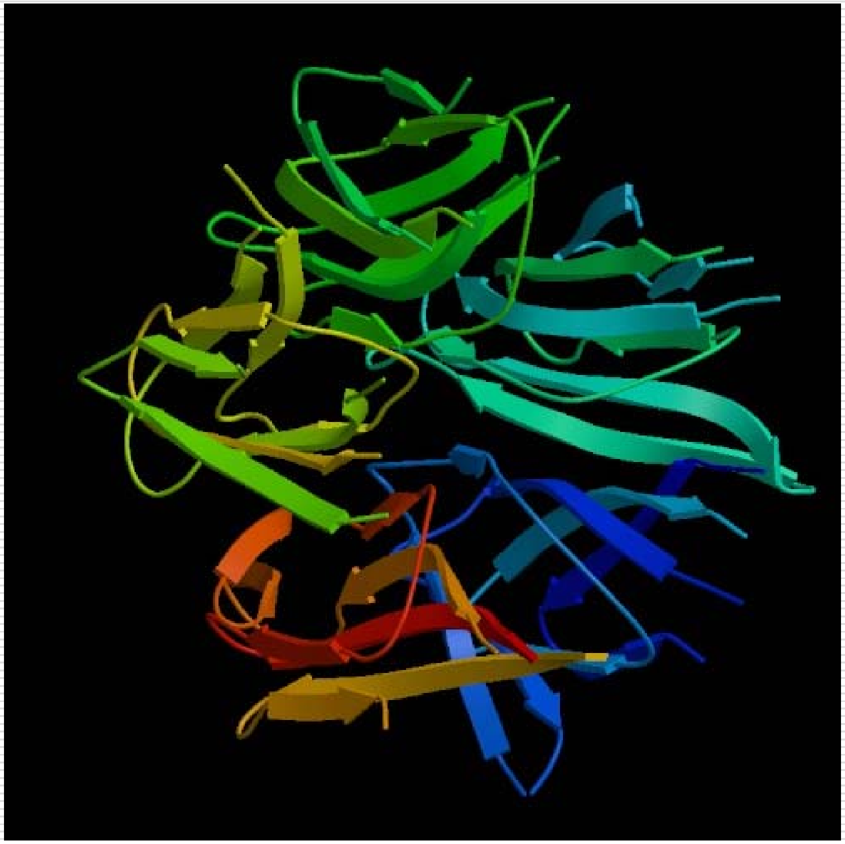


Anti-parallel β–sheets

- The amino acids in successive strands have alternating directions, N-terminal to C-terminal



Nucleoplasmin (PDB 1K5J)



Amino Acid Preferences (1)

- **α-helix forming**
- The amino acid side chain should *cover and protect the backbone H-bonds* in the core of the helix
- Ala, Leu, Met, Glu, Arg, Lys: good helix formers
- Pro, Gly, Tyr, Ser: very poor helix formers

- **β-strand forming**
- Amino acids with large bulky side chains prefer to form β-sheet structures
- Tyr, Trp, Ile, Val, Thr, Cys, Phe

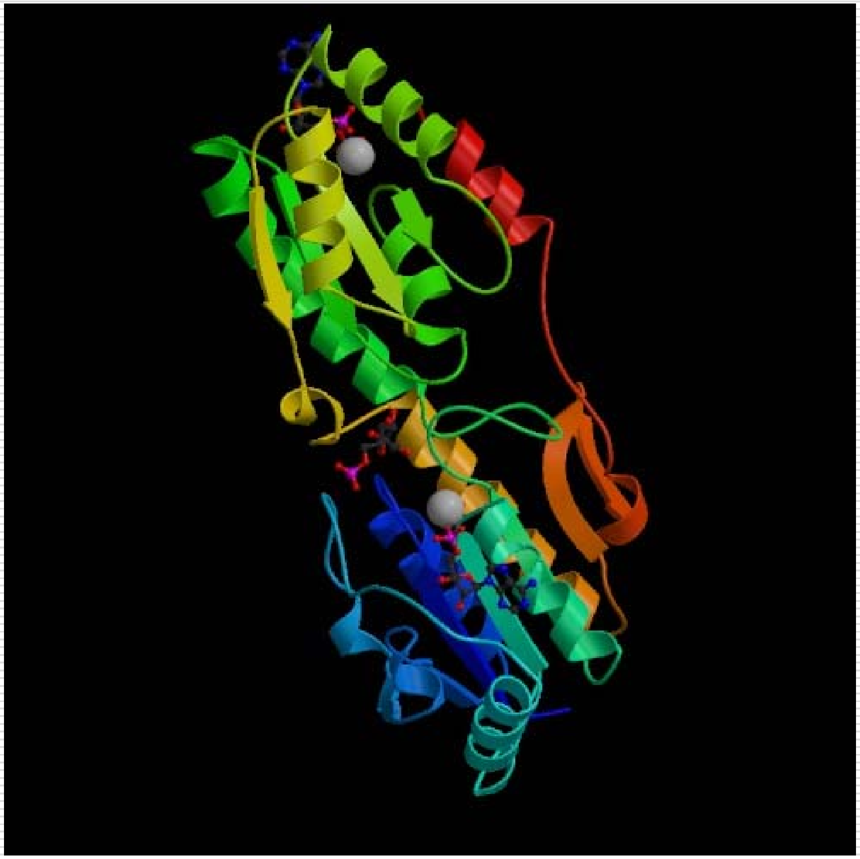
Amino Acid Preferences (2)

- **Secondary structure disruptors**
- Gly: side chain too small
- Pro: side chain linked to α-N, has no N-H to H-bond; rigid structure due to ring
- Asp, Asn, Ser: H-bonding side chains compete directly with backbone H-bonds

Turns/Loops

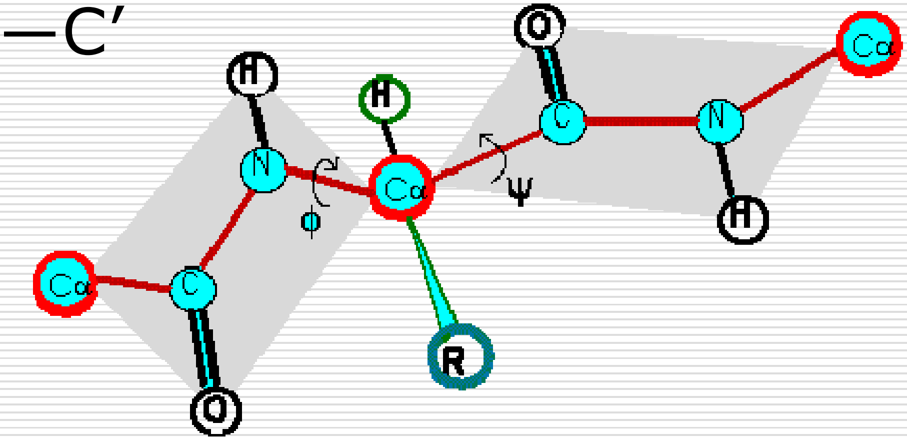
- Third "classical" secondary structure
- Reverses the direction of the polypeptide chain
- Located primarily on protein surface
- Contain polar and charged residues
- Three types: I, II, III

Phosphofructokinase (PDB 4PFK)



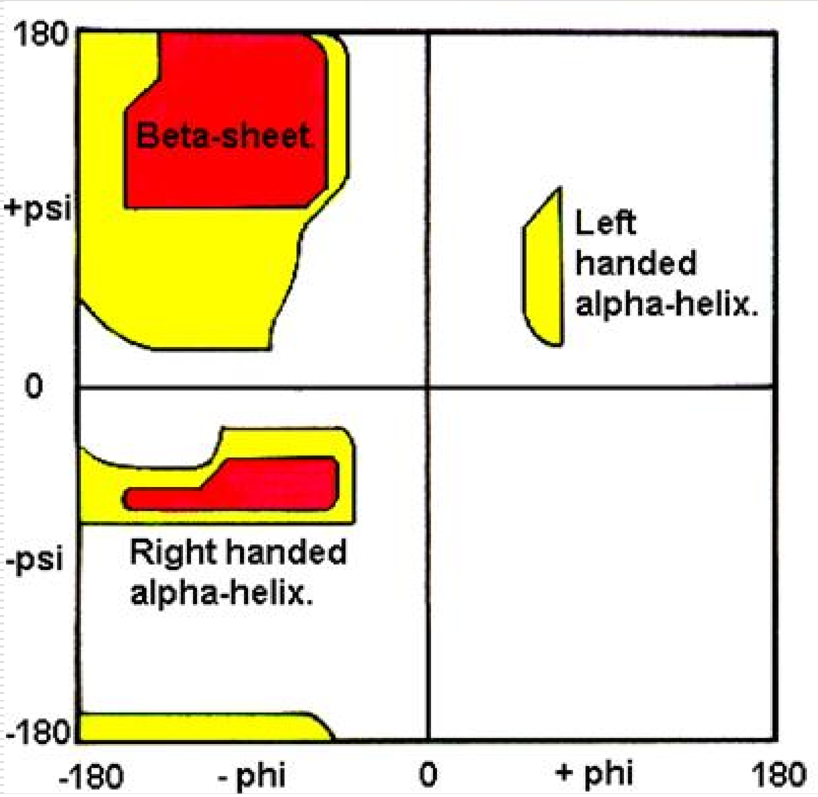
The Torsional Angles: ϕ and ψ

- Each amino acid in a peptide has two degrees of backbone freedom
- These are defined by the ϕ and ψ angles
- ϕ = angle between C_α—N
- ψ = angle between C_α—C′

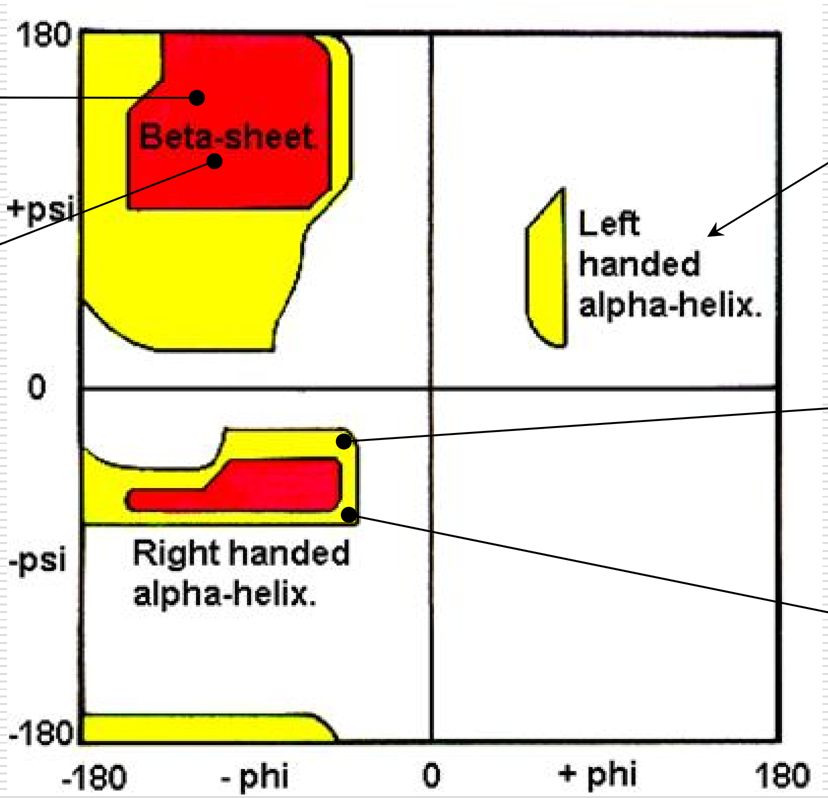


The Ramachandran Plot

- Plot of allowable ϕ and ψ angles
- ϕ and ψ refer to rotations of two rigid peptide units around C_α
- Most combinations produce steric collisions
- Disallowed regions generally involve steric hindrance between side chain C_β methylene group and main chain atoms



The Ramachandran Plot (contd.)



Anti-parallel β-sheet

Parallel β-sheet

Theoretically possible; energetically unstable

3_{10}-helix

π-helix

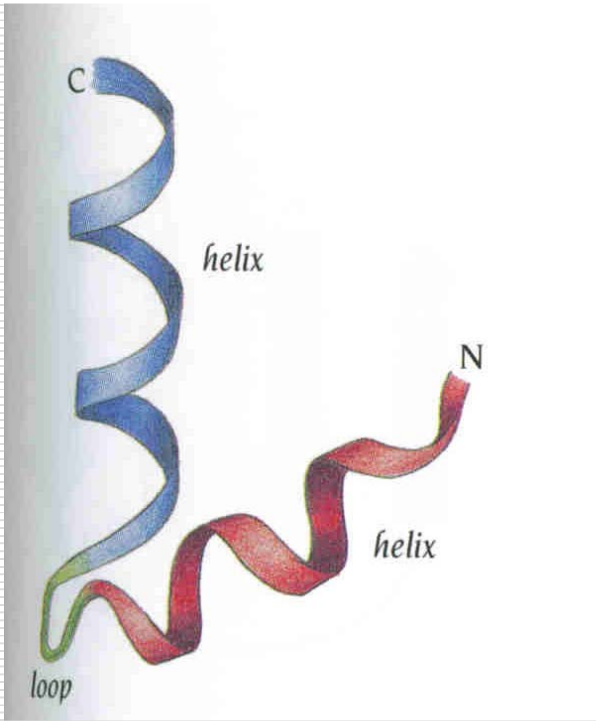
- White: sterically disallowed (except Gly)
- Red: no steric clashes
- Yellow: "allowable" steric clashes

Protein Motifs

- (Simple) combinations of secondary structure elements with a specific geometric arrangement
- “Super-secondary structures”
- Can be associated with a specific function, e.g. DNA binding, or metal ion binding
- Can be part of a larger functional and/or structural assembly (“domain”)

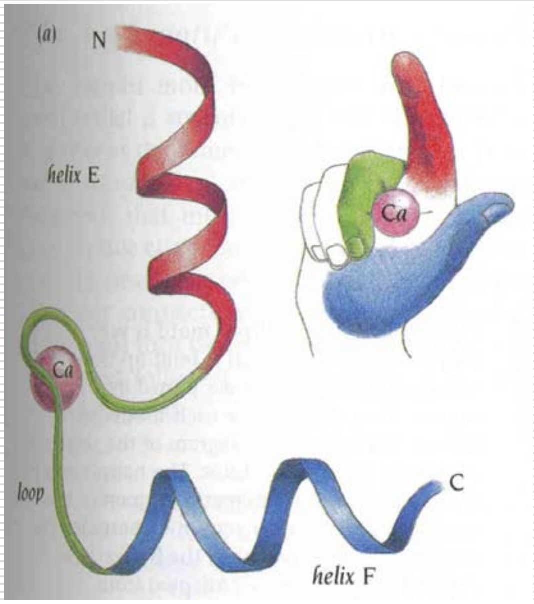
Helix-Turn-Helix (HTH)

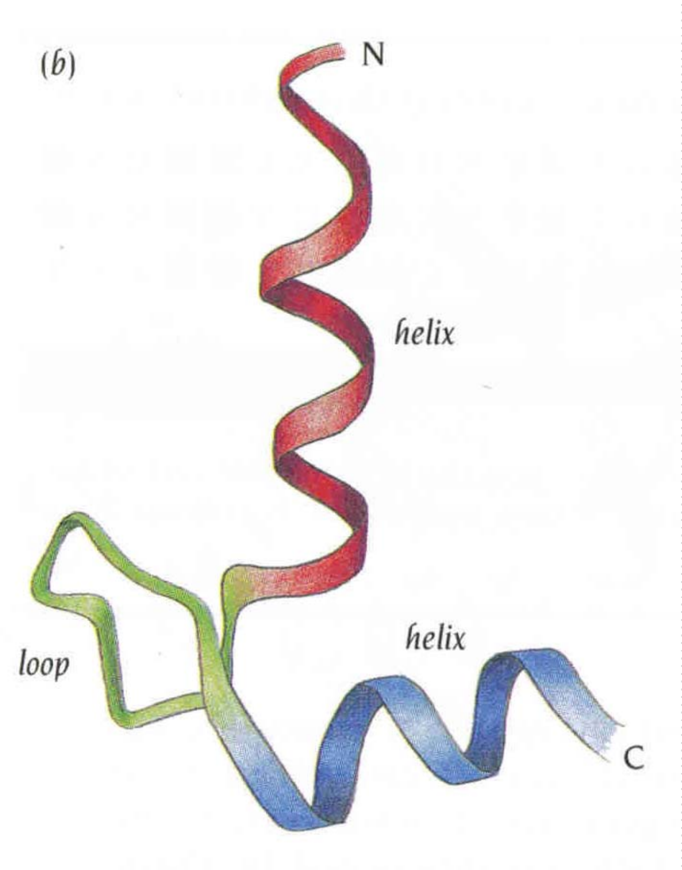
- Simplest α-helix motif: 2 α helices joined by a loop
- Also called *Helix-Loop-Helix, HLH*
- Common structural motif for DNA binding proteins
- One helix recognizes specific sequence of nucleotides and fits into the groove in the DNA double helix; the other stabilizes the bound configuration



EF Hand

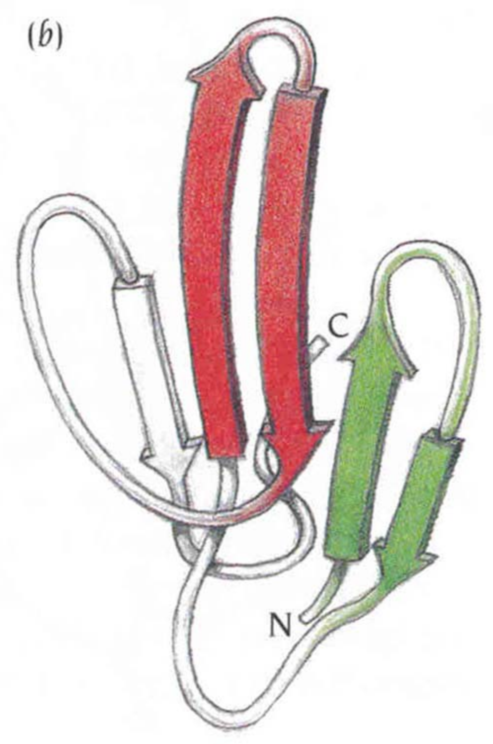
- Specific for several different calcium-binding proteins
- E.g. calmodulin, Troponin-C





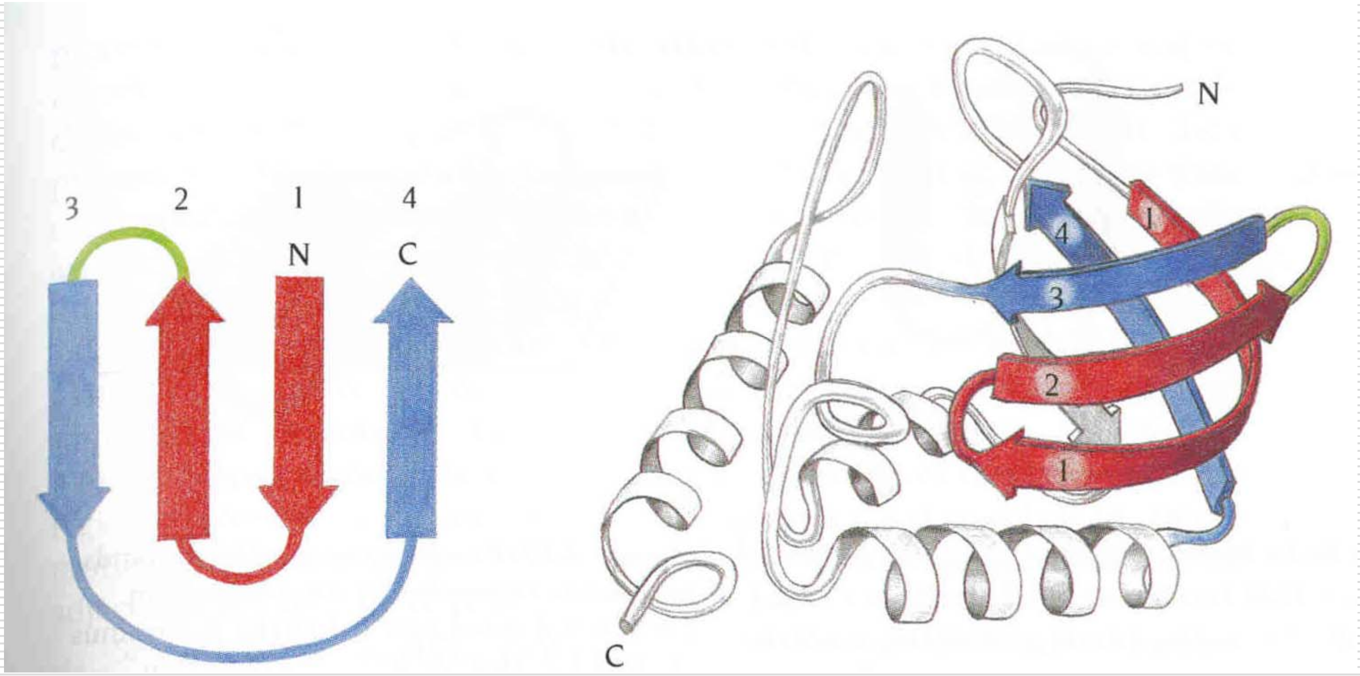
Hairpin β-motif

- Also called *β-hairpin*
- 2 adjacent anti-parallel strands joined by a loop



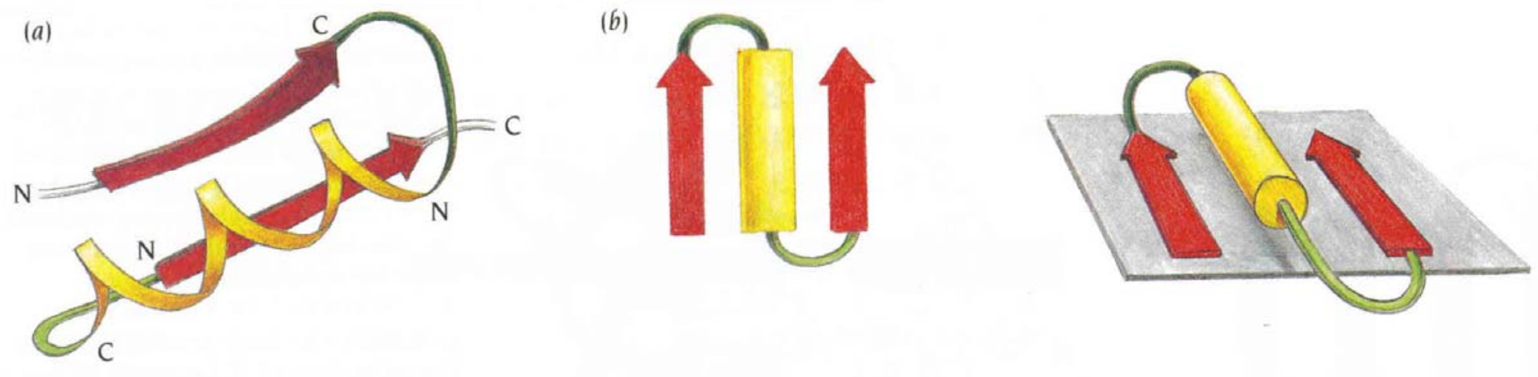
Greek Key motif

- 4 adjacent anti-parallel β-strands



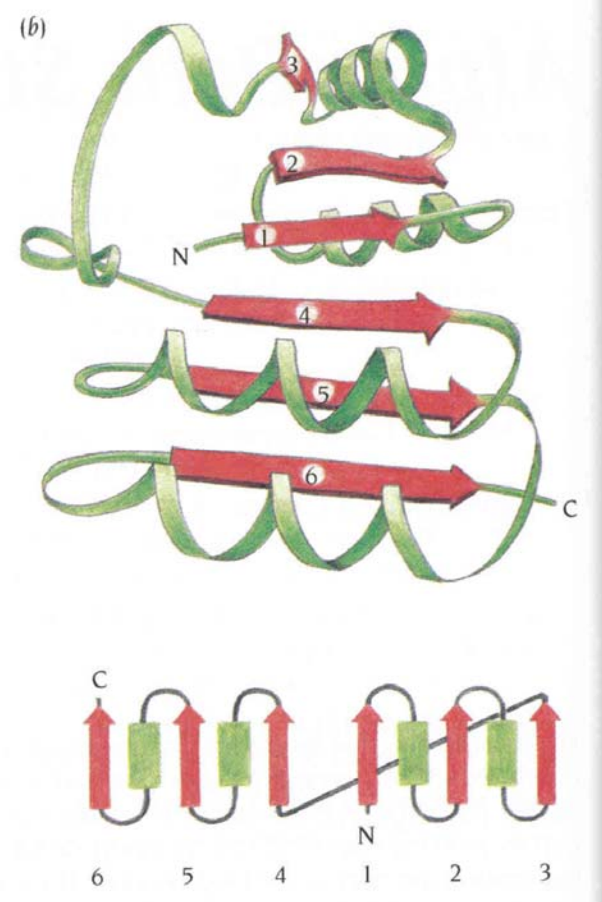
β-α-β Fold

- Parallel β-strands connect by an α helix
- Found in most proteins with parallel β strands. E.g. Triose phosphate isomerase



Rossman Fold

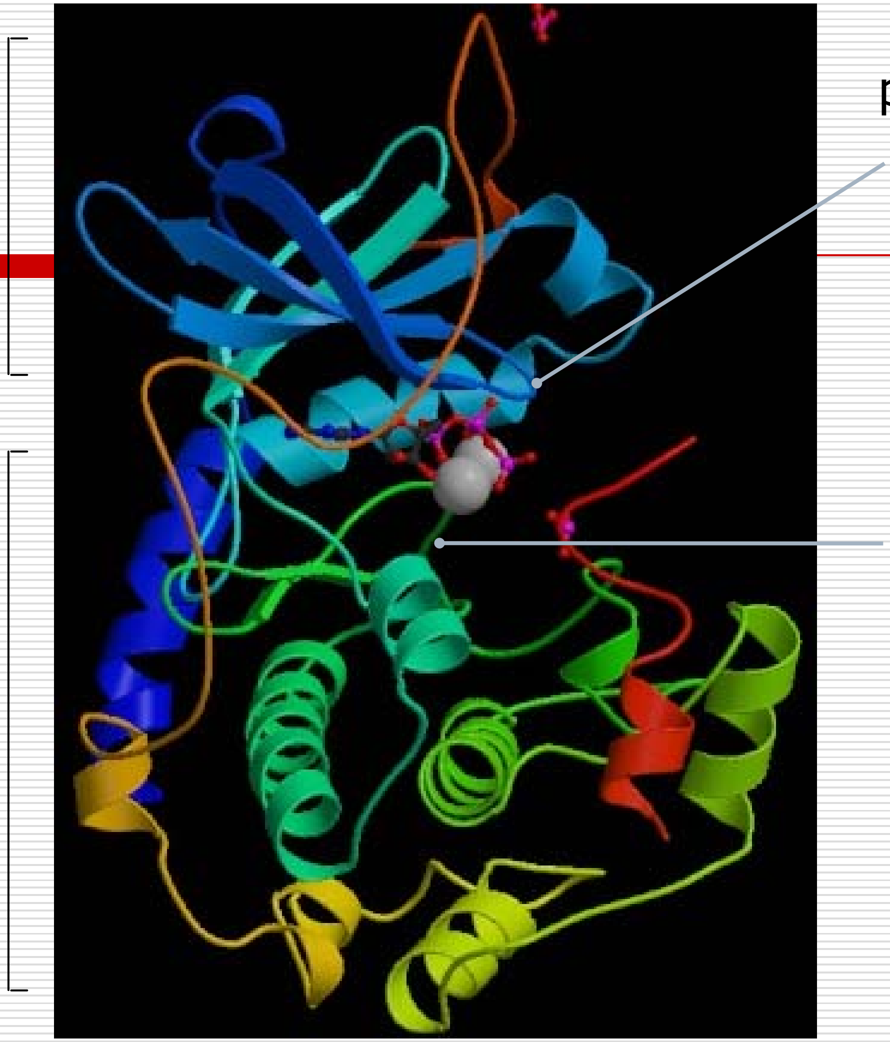
- Unique example of a β–α–β fold
- 3 parallel β-sheets with 2 linking α helices
- Often seen in nucleotide-binding proteins



Domains

- ✓ Primary structure
- ✓ Secondary structure
- ✓ Super-secondary structure
- ☐ *Domains*
- ☐ Fundamental unit of tertiary structure
- ☐ (Part of a) polypeptide chain that can fold independently into a stable tertiary structure
- ☐ E.g. catalytic domain of protein kinase; binding pocket of a ligand

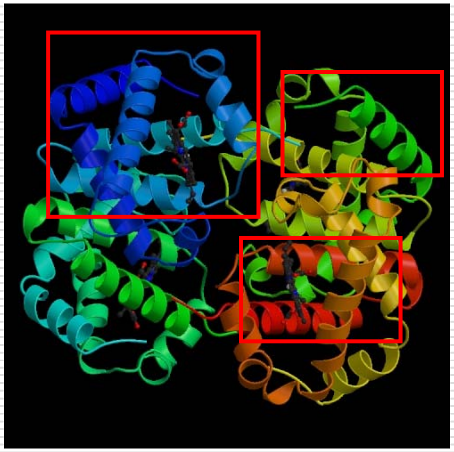
3D Structure of a Protein Kinase Domain



Quarternary Structure

- Spatial organization of subunits to form functional protein
- E.g. Hemoglobin
- 2 α chains, 2 β chains
- Each chain binds heme (Fe)
- Forms an $\alpha_2\beta_2$ tetramer

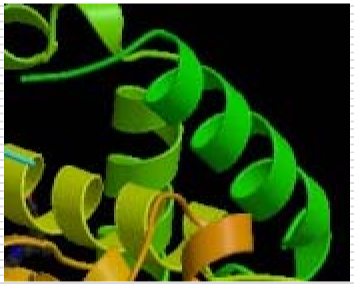
Putting it all together



Primary Structure
… KAAWGKVGAHA …

→

Secondary Structure
α-helix, β-sheets, turns/loops

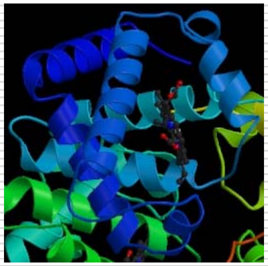


↓

Super-secondary Structure
Heme-binding pocket/domain (His)

←

Tertiary Structure
a single chain (α, β)



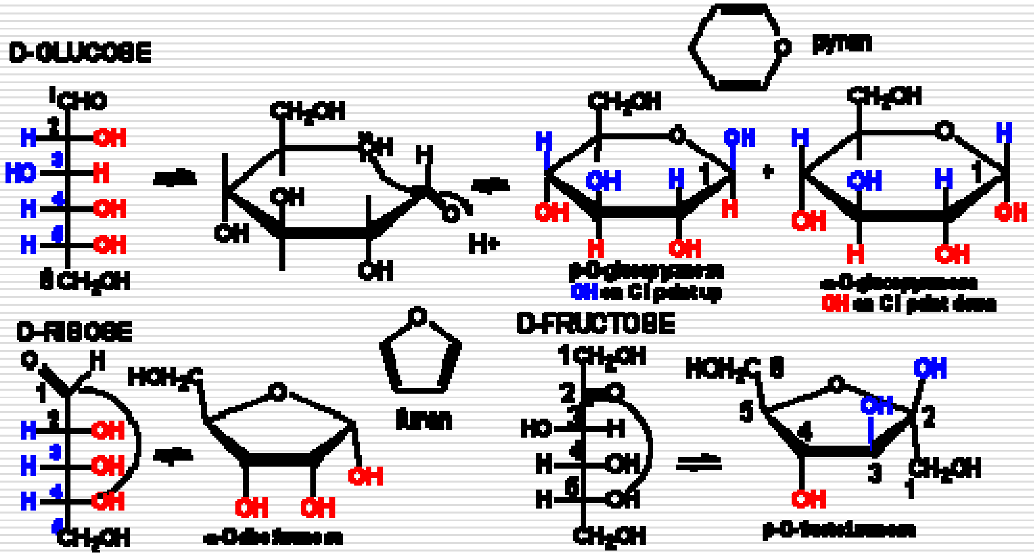
←

Quarternary Structure
$\alpha_2\beta_2$

Carbohydrates

- Sugars
- Polysaccharides

Sugars

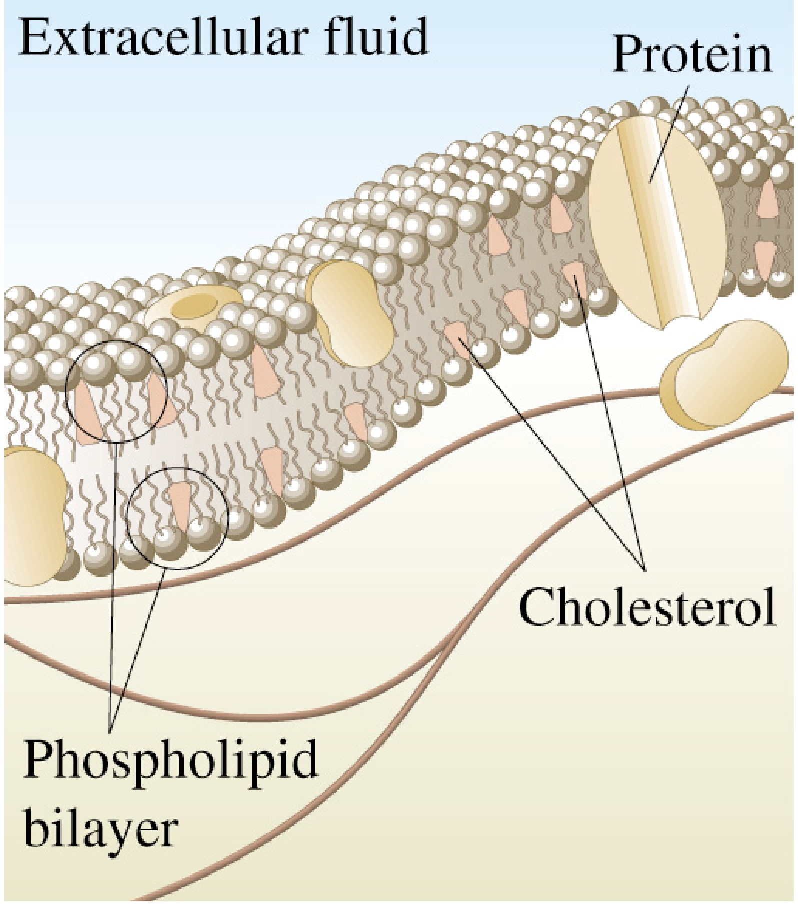

D-GLUCOSE
pyran
H+
β-D-glucopyranose
OH on C1 point up
α-D-glucopyranose
OH on C1 point down
D-RIBOSE
furan
D-FRUCTOSE
α-D-ribofuranose
β-D-fructofuranose

Lipids

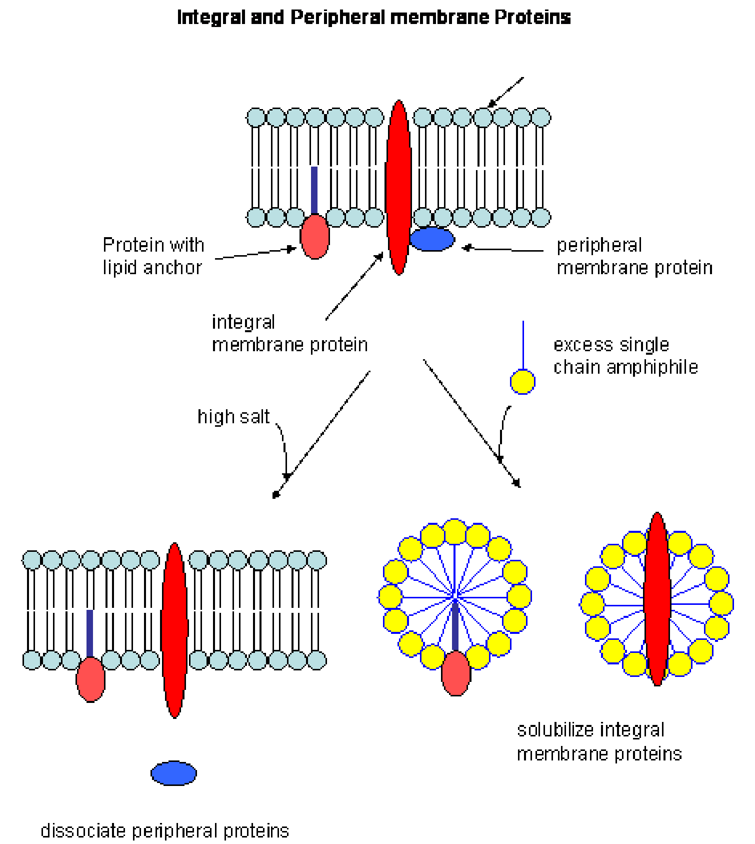
- Fatty acids
- Triglycerides
- Steroids
- Membranes

(Structure of lipids)

Cell Membrane



Types of Membrane Proteins

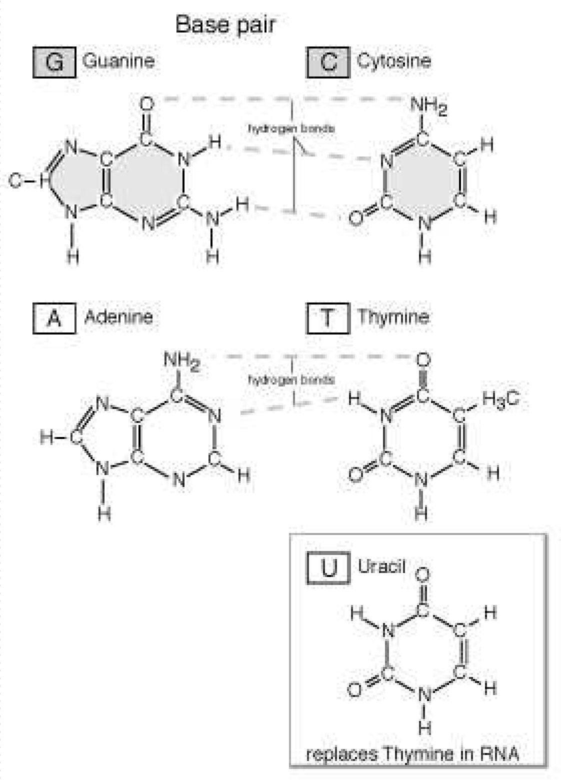


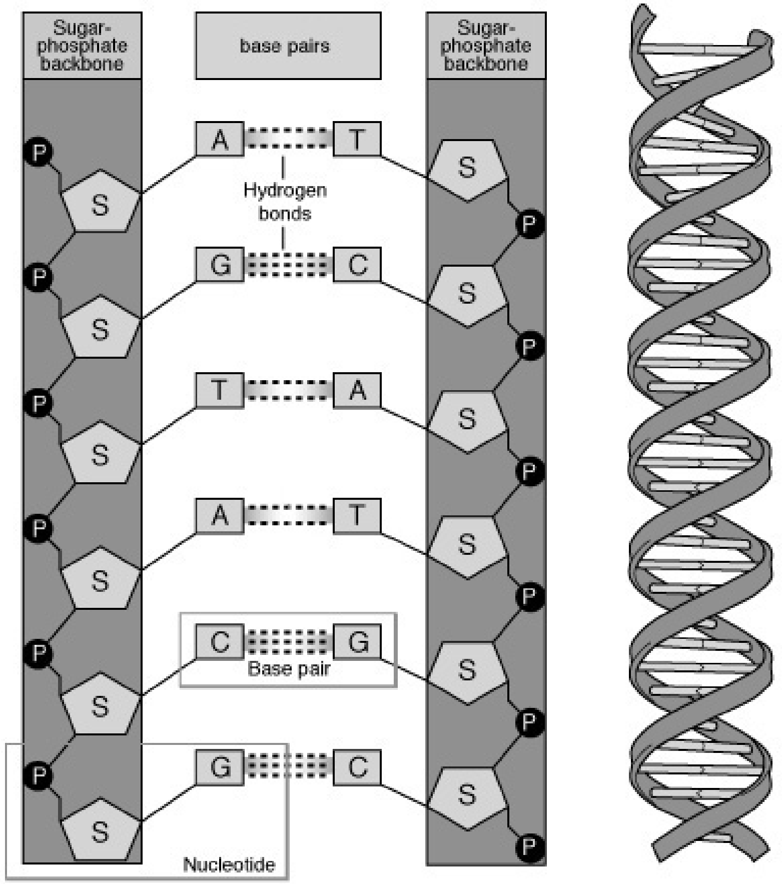
Nucleic Acids

- Nucleotides
- DNA
- RNA

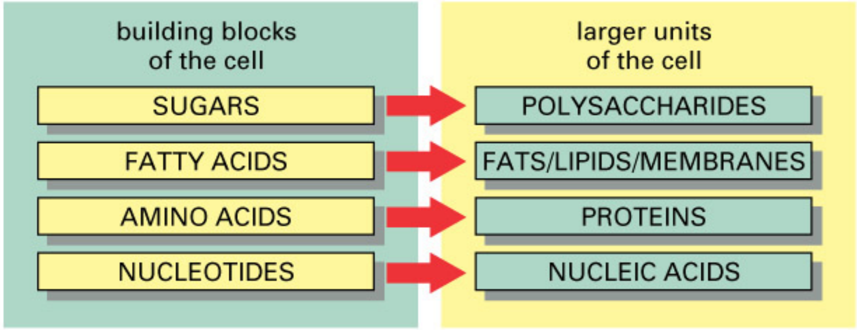
(DNA)

Base Pairing





Biopolymers



building blocks of the cell	larger units of the cell
SUGARS	POLYSACCHARIDES
FATTY ACIDS	FATS/LIPIDS/MEMBRANES
AMINO ACIDS	PROTEINS
NUCLEOTIDES	NUCLEIC ACIDS

Figure 2-15 Essential Cell Biology, 2/e. (© 2004 Garland Science)

Composition of a Cell

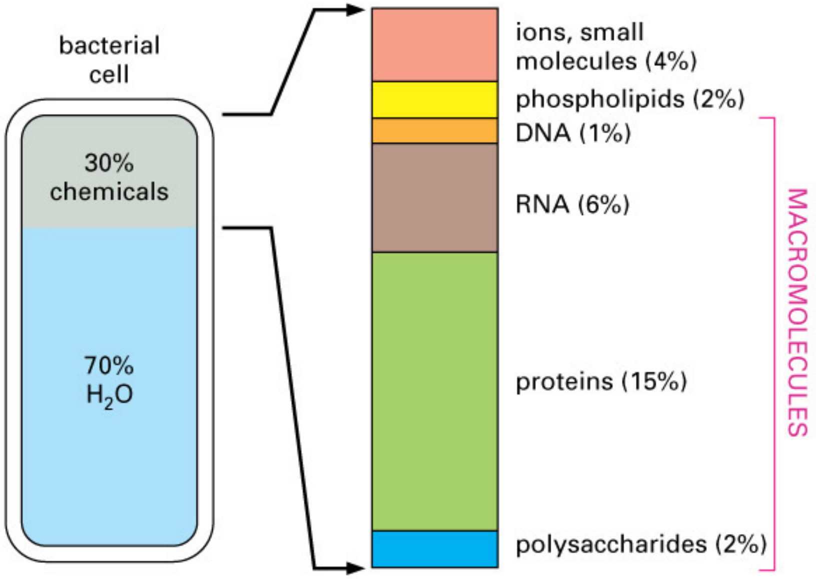
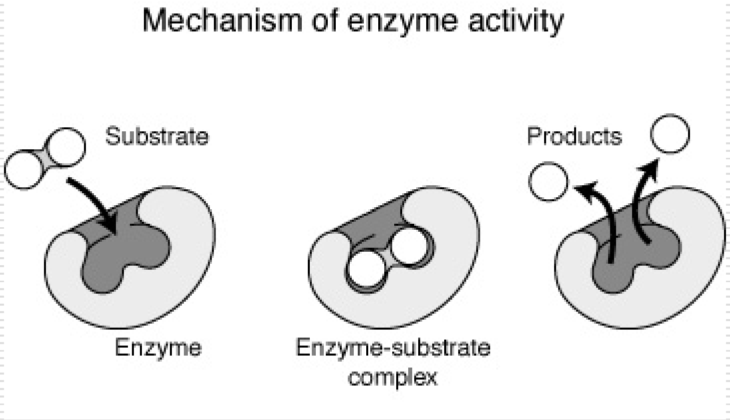


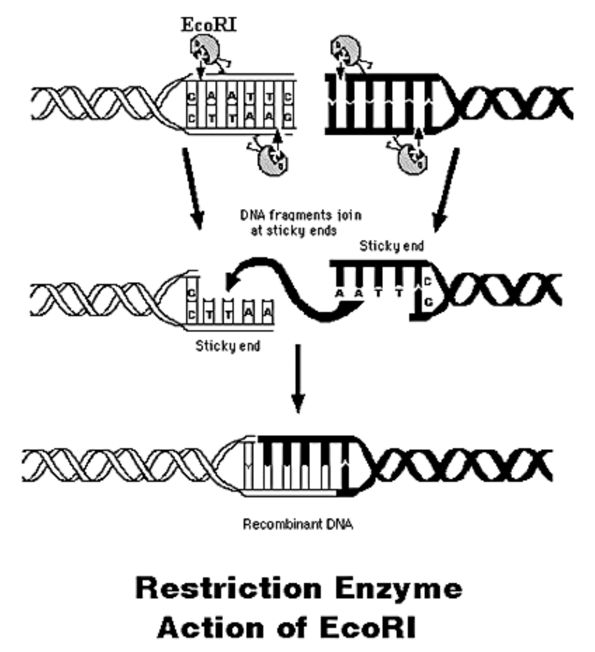
Figure 2-26 Essential Cell Biology, 2/e. (© 2004 Garland Science)

Special Types of Proteins

- Enzymes
- Receptors
- Channel proteins
- Antibodies

Enzymes


Mechanism of enzyme activity



Restriction Enzyme Action of EcoRI

(Ribozymes movie)

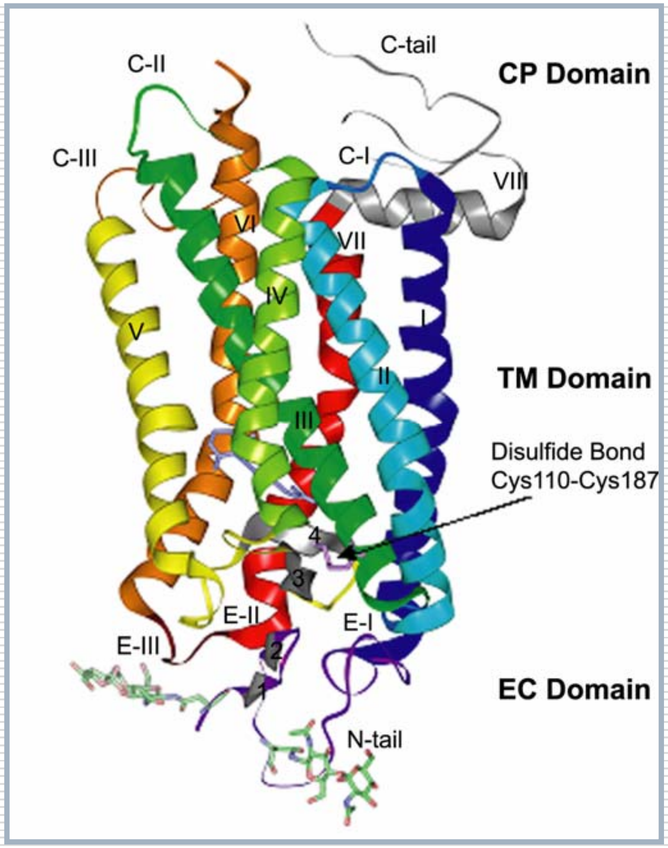
Cell Receptors

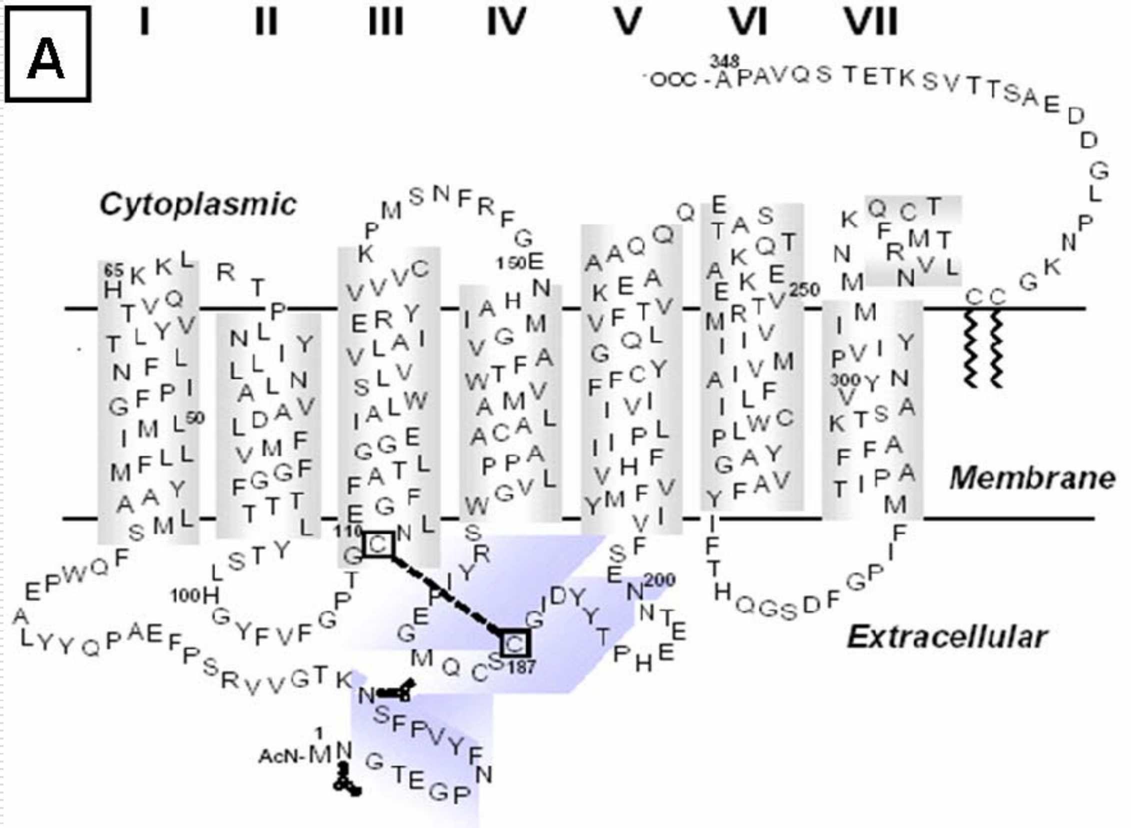
- **Ion-channel linked**: involved in rapid synaptic signaling between excitable cells; mediated by neurotransmitters
- **Enzyme-linked receptors**: when activated, either function directly as enzymes or are associated with enzymes.
- **G-protein coupled receptors (GPCR)**

GPCRs

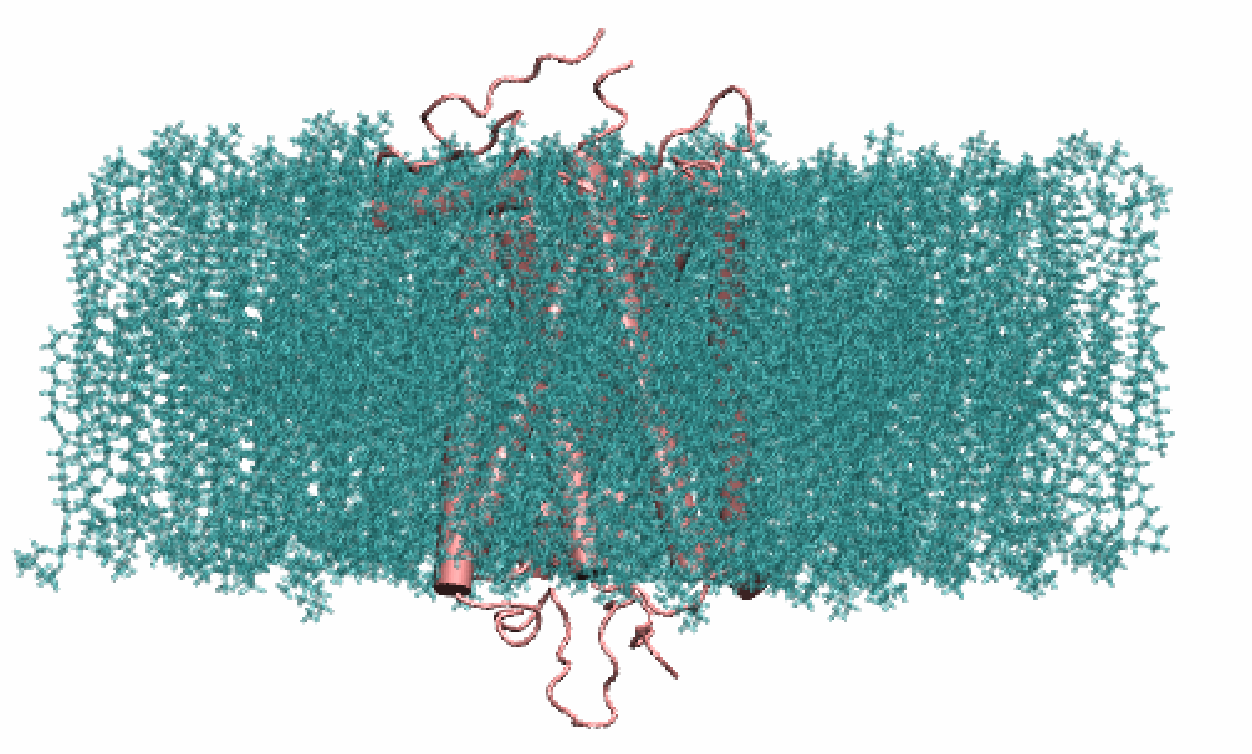
- Largest family of cell-surface receptors
- Biological functions include smell, taste, vision, blood pressure neurotransmission, embryogenesis, cell growth, development
- Rhodopsin is the only GPCR with a known 3D structure
- Contains 7 membrane traversing α helices (7TM)
- N terminal – outside cell, C terminal – inside cell
- Ligand binding outside cell induces conformational change detected inside cell
- Mediating molecule is a G protein (hence the name GPCR)
- Heterotrimeric GTP-binding regulatory protein (α, β, γ)
- Activated G protein transmits signal by binding to other proteins (e.g. adenylate cyclase: converts ATP to cAMP)

GPCR Structure





GPCR Structure (contd.)

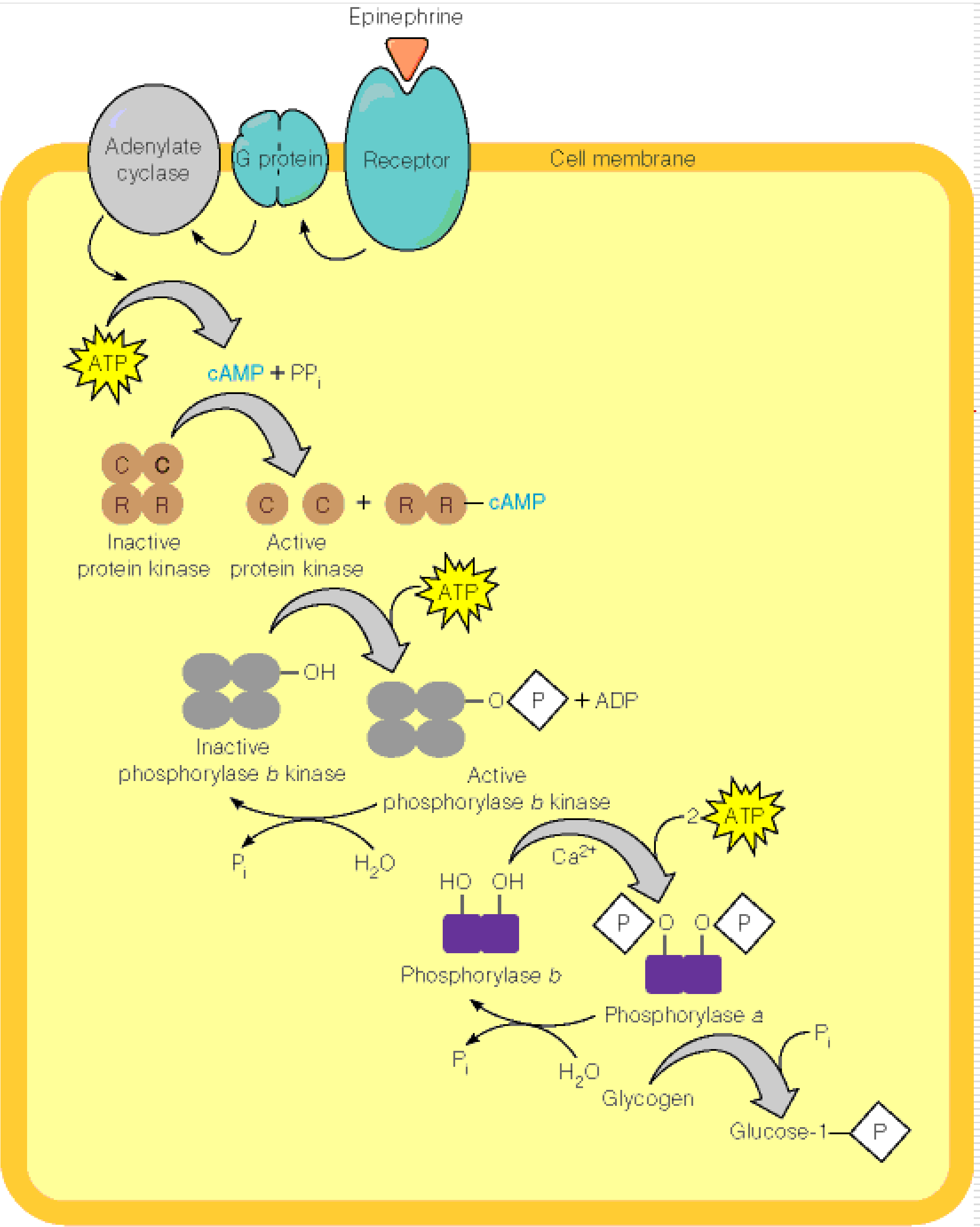


Signal reception

Signal mediation and amplification

Regulation by reversible phosphorylation and dephosphorylation

Signal effects



Channel Proteins

(Structure of Channel Proteins)

Antibody Structure

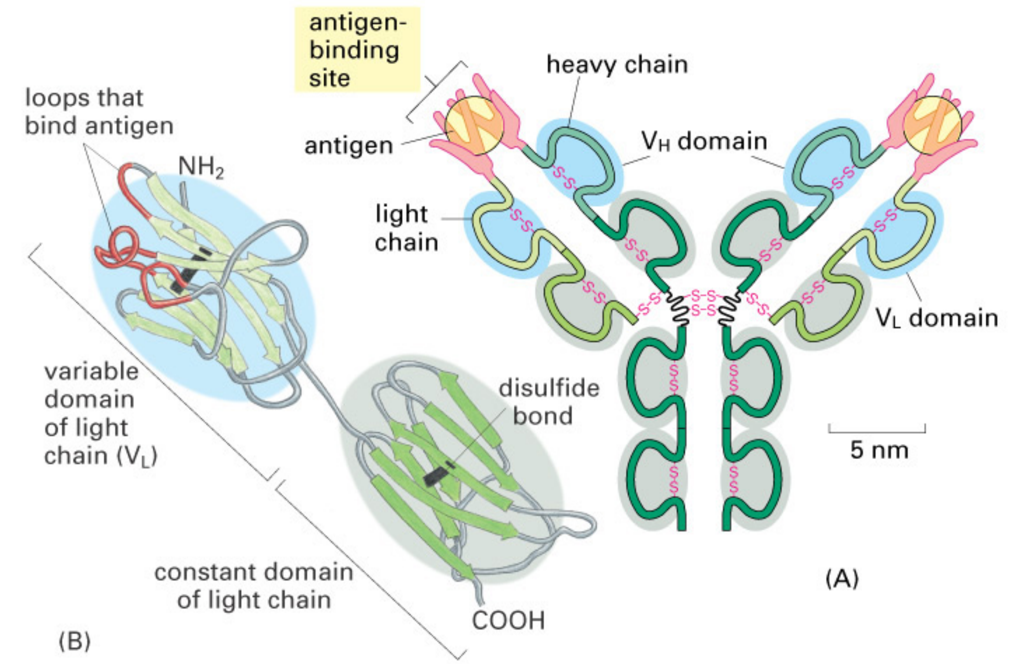


Figure 4-32 **Essential Cell Biology, 2/e. (© 2004 Garland Science)**

Additional Reading

- **General information**
- Biochemistry, 5th ed., Berg, Tymoczko, Stryer
- Biochemistry, 3rd ed., Voet & Voet

- **Detailed information**
- Proteins, 2nd ed., Creighton
- Introduction to Protein Structure, 2nd ed., Branden & Tooze

- **Internet**
- Images: Protein Data Bank (PDB): www.rcsb.org/pdb
- Numerous websites (Google *protein secondary structure)*