

Charles Martin
Cell Biology and Neuroscience
B323 Nelson Labs
martin@biology.rutgers.edu

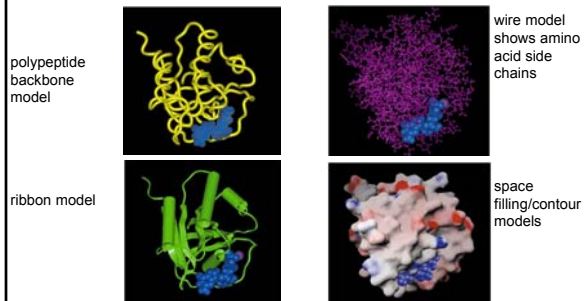
Protein Structure and Function

Protein Central Dogma (revised)

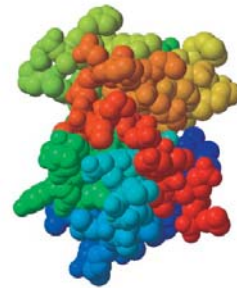
- Proteins are large molecules that are formed as single, unbranched chains of amino acid monomers
 - But, proteins can be turned into branched structures by ubiquitin and other ubiquitin-like molecules
- There are 20 different amino acids commonly found in proteins
- A protein's amino acid sequence determines its three-dimensional structure (conformation)
 - Well, sort of...
- A protein's 3-dimensional structure determines its chemical function(s)
 - (along with a whole lot of different post-translational modifications that can alter parts of its structure and change its functions)

Detailed structures of proteins are complex, so we have to look at them in different ways.

Different graphical representations of the same protein



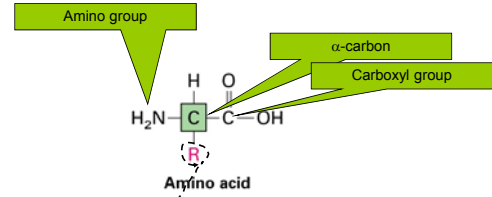
SH2 domain space filling model



Opening Image from Chapter 3

Let's Review

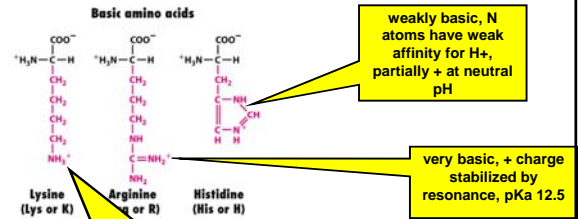
All amino acids have the same general structure



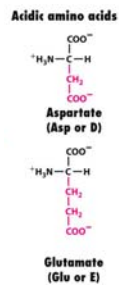
The functional group (or "side chain") determines the unique chemical and physical properties of each amino acid

Eleven of the most common amino acids are hydrophilic. Seven of them have ionizable side chains.

Basic amino acids have a positive charge at pH 7.0

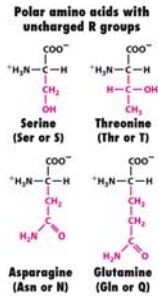


Acidic amino acids have a negative charge at pH 7



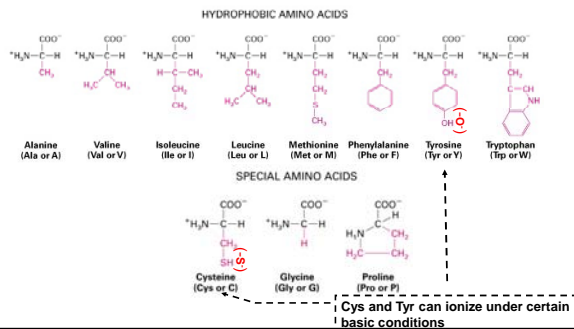
4 of the hydrophilic amino acids are polar, but uncharged

Uncharged polar amino acids

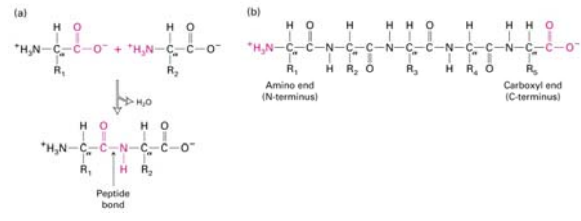


The remaining amino acids have hydrophobic and “special” functional groups

Hydrophobic and “special” amino acids



Amino acids are linked by an amide linkage, called a peptide bond, to form polypeptide chains



Peptide bonds and the α carbon atoms form the linear backbone of proteins, which is a regular, repeating unit

- The functional groups of amino acids form “side chains” that are connected to the backbone

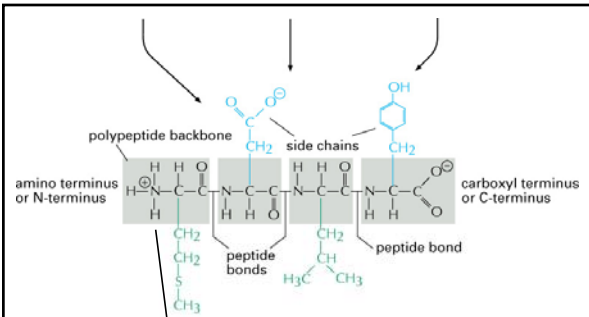
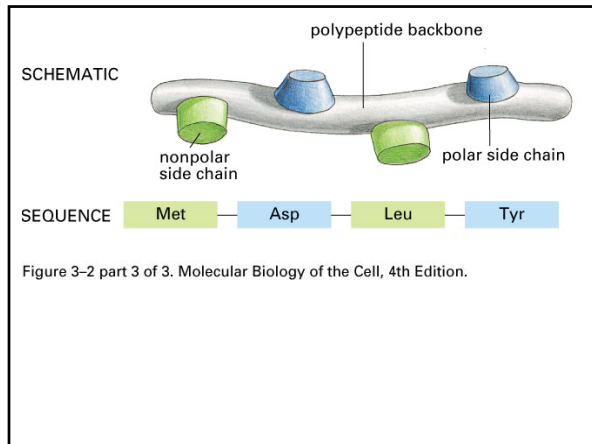
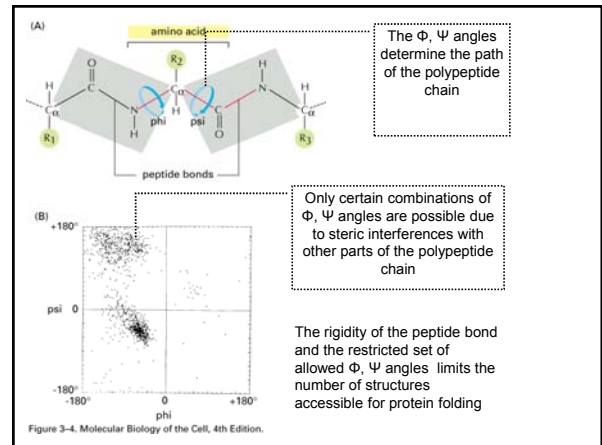
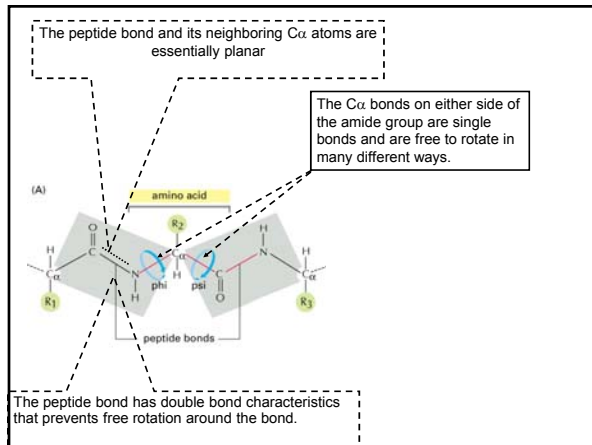


Figure 3-2 part 2 of 3. Molecular Biology of the Cell, 4th Edition.

Note that the backbone has polarity and the amino terminal residue is always written on the left



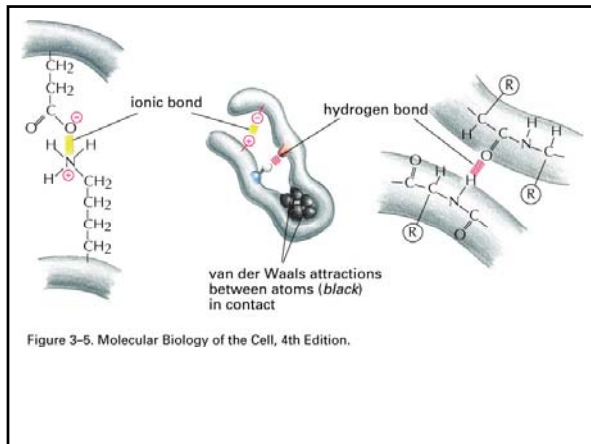
Polypeptide chains are flexible, but conformationally restricted



The shape of proteins is determined through 4 levels of structure

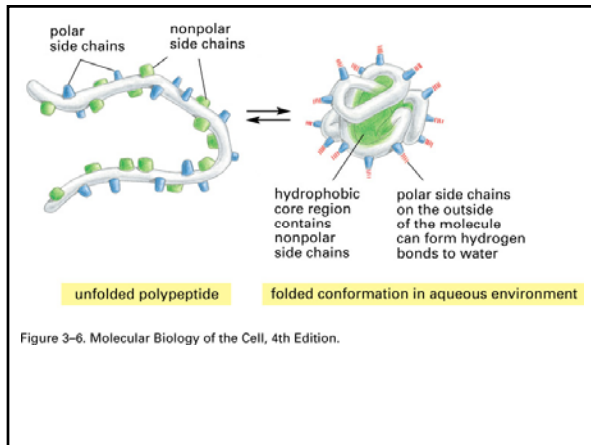
- **Primary:** the linear sequence of amino acids
- **Secondary:** the localized organization of parts of a polypeptide chain (e.g., the α helix or β sheet)
- **Tertiary:** the overall, three-dimensional arrangement of the polypeptide chain
- **Quaternary:** the association of two or more polypeptides into a multi-subunit complex
- The final, 3-dimensional, folded structure is generally one in which the free energy of the molecule is minimized

Three types of weak, non-covalent bonds also constrain the folding of proteins into their energy-minimized 3-D structures



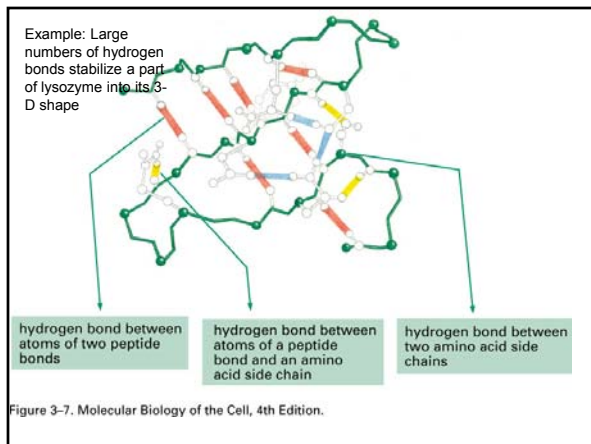
Hydrophobic interactions also play a role in determining protein shape

- Residues with hydrophobic side chains tend to cluster in the interior of the protein molecule, avoiding contact with water
- Polar side chains tend to be arranged on the outsides of proteins in contact with the aqueous medium



All of these bonds are about 30-300 times weaker than covalent bonds

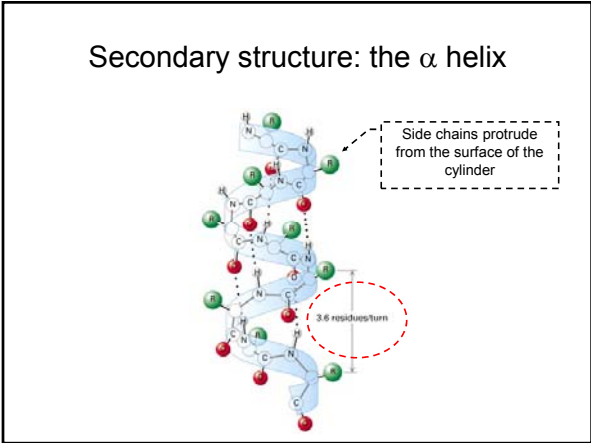
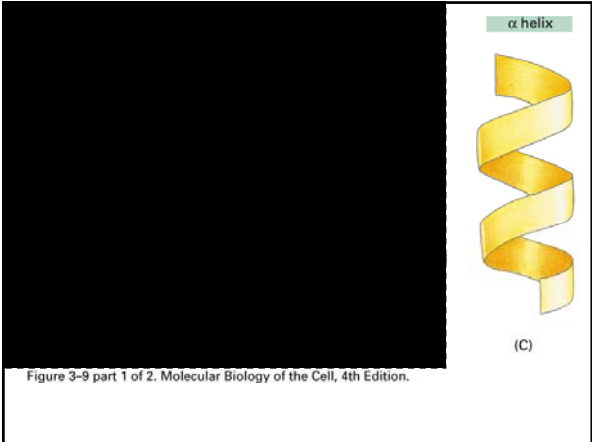
- So why are they important?
- Many weak bonds applied together can produce a large force. The stability of a protein is determined by the combined strength of many non-covalent bonds



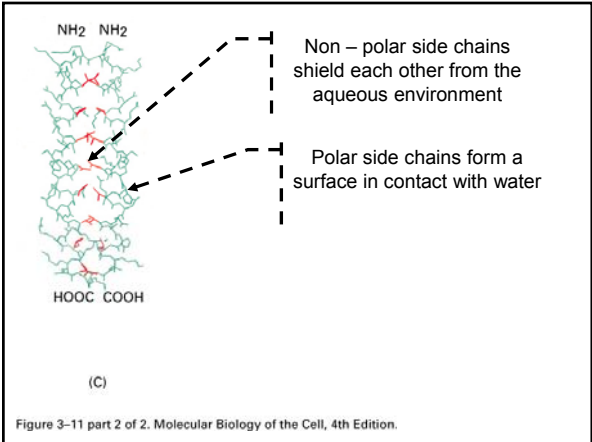
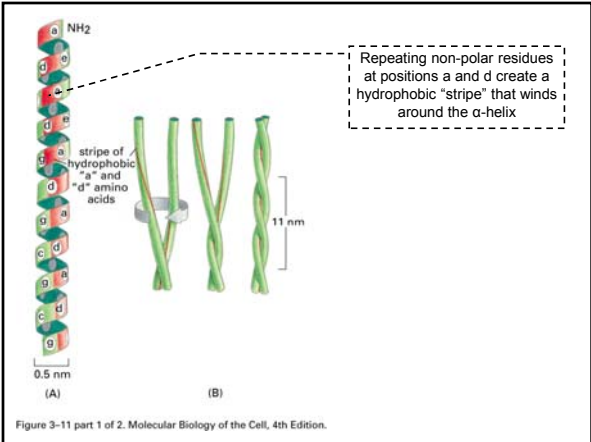
Secondary Structure

- The α -helix and the β -sheet are two regular folding patterns found in almost all proteins
- What produces these structures and why are they so common?
 - They result from hydrogen bonding between multiple N-H and C=O groups *in the backbone*.
 - Side chains are not involved in these structures.

The α -helical backbone is a rigid cylinder with the amino acid side chains protruding from its surface

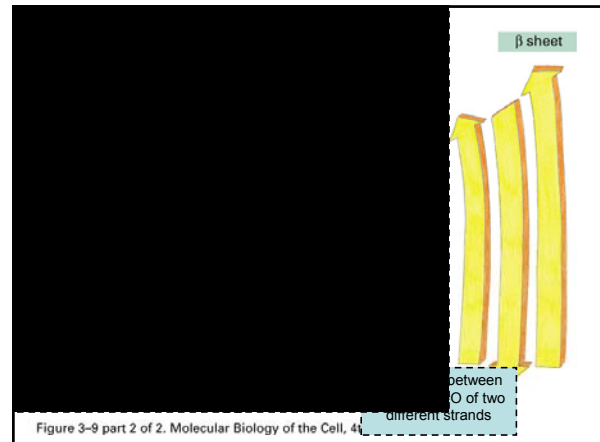


α -helices can form very stable **coiled-coil** structures through hydrophobic interactions between non-polar side chains

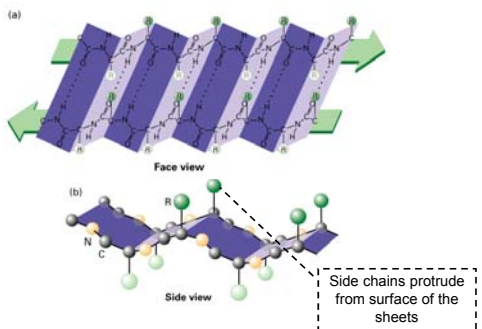


β -sheets are found in the core of many proteins

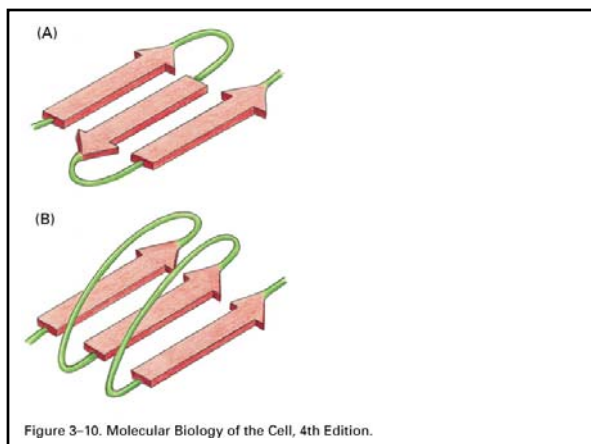
- β -sheets are rigid, relatively flat and extended structures that are stabilized by hydrogen bonds **between neighboring polypeptide strands**



Secondary structure: the beta sheet
Where do the side chains go?



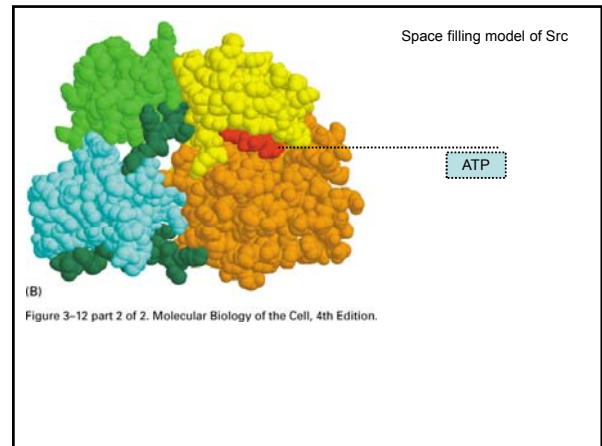
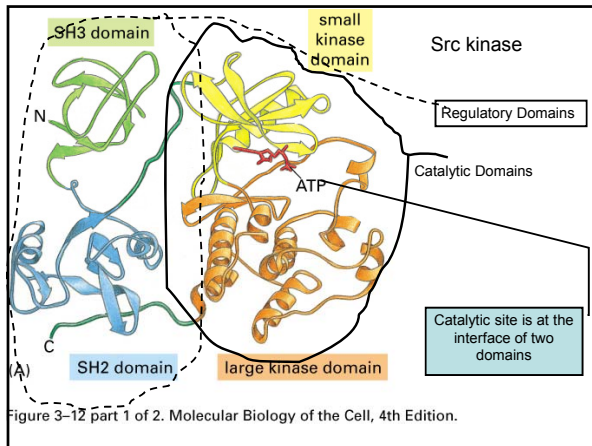
β -sheets can be in either a parallel or antiparallel orientation



Protein domains represent another important unit of organization

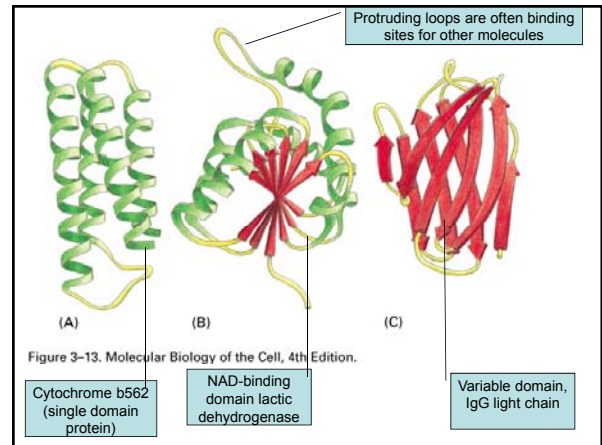
Domains are parts of a polypeptide chain that can fold independently into a compact, stable structure

These are the modules from which most larger proteins are built



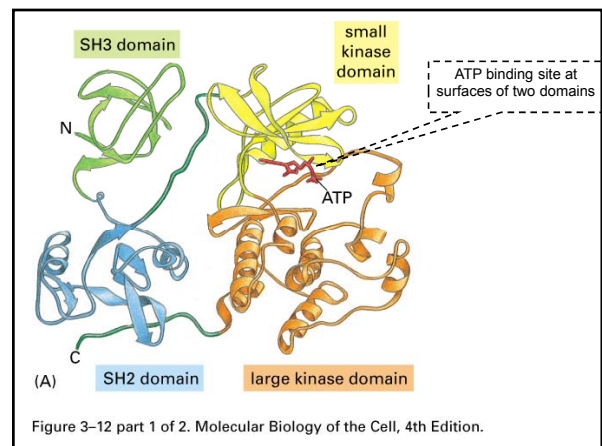
Domains are constructed from different combinations of α -helices and β -sheets at their core

- Each combination is called a *protein fold*



Most large multi-domain proteins have evolved by recombination and joining of preexisting domains in new combinations (*Domain Shuffling*)

Many small molecule binding sites in proteins are created at the surfaces between new combinations of domains



Proteins derived from Domain Shuffling

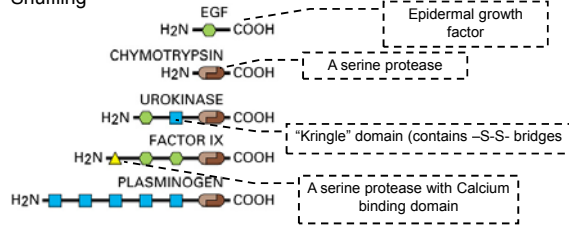


Figure 3-18. Molecular Biology of the Cell, 4th Edition.

Large proteins often contain more than one polypeptide chain

Binding between two protein surfaces generally involves sets of non-covalent bonds

Cro repressor is a symmetrical head to head dimer held together by hydrophobic interactions and hydrogen bonds



Figure 3-21. Molecular Biology of the Cell, 4th Edition.

Other symmetric protein complexes can consist of multiple subunits

Neuraminidase forms a ring of 4 identical polypeptide chains

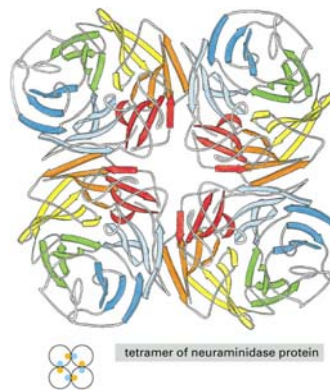
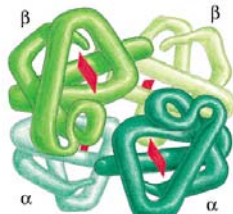


Figure 3-22. Molecular Biology of the Cell, 4th Edition.

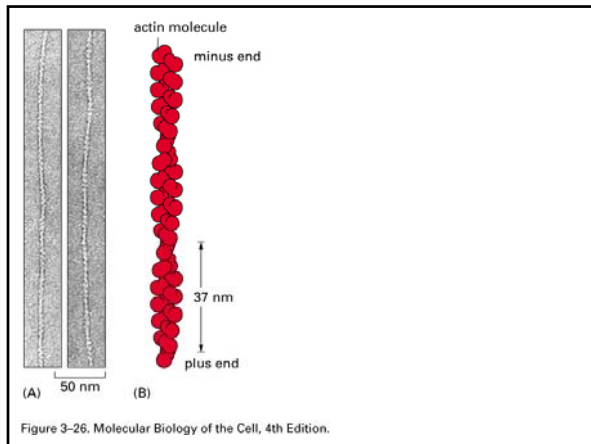


Hemoglobin has two different subunits and is an $\alpha_2\beta_2$ tetramer

Figure 3-23. Molecular Biology of the Cell, 4th Edition.

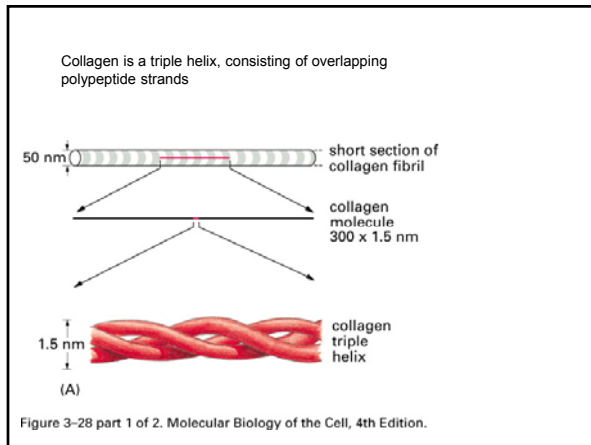
Some globular proteins can form long helical filaments

- Globular proteins fold into a compact, ball-like shape with irregular surfaces
- Example: Actin filaments form in a helical arrangement that can be the length of the cell



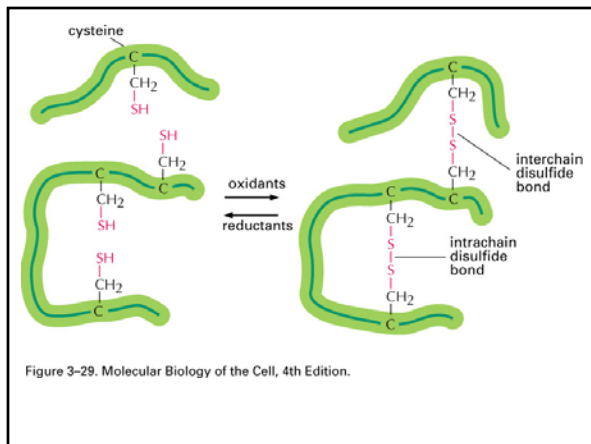
Fibrous proteins can also span long distances. These tend to have simple elongated structures

- Many fibrous proteins are components of extracellular matrix that binds cells together to form tissues



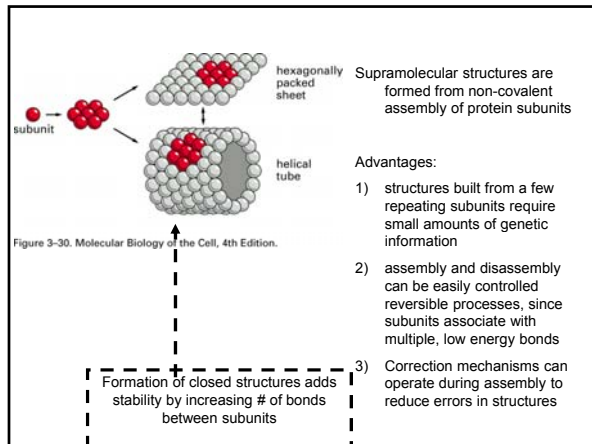
Most extracellular proteins are stabilized by covalent -S-S- cross-links

Disulfide bond formation is catalyzed in the ER prior to export



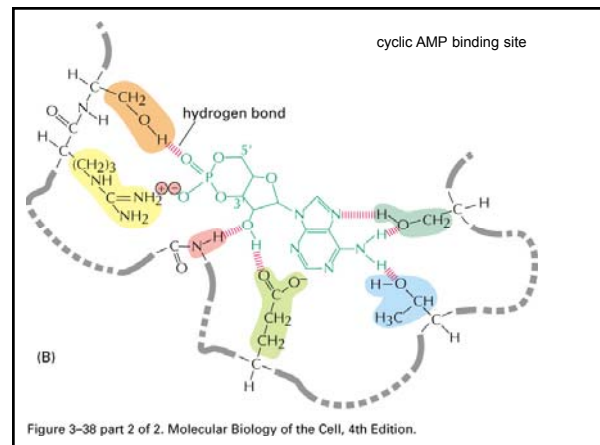
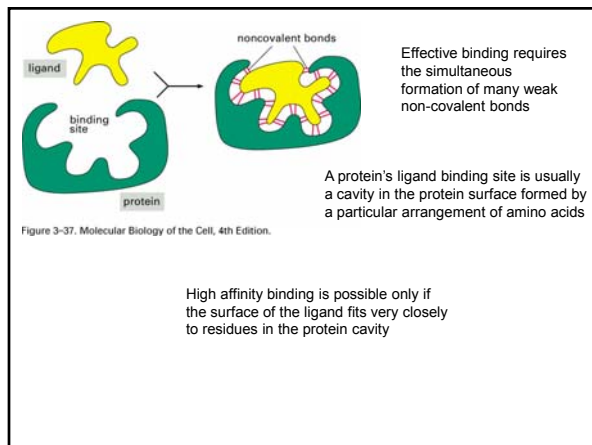
Proteins can be subunits for the assembly of large structures

- enzyme complexes
- ribosomes
- proteasomes
- filamentous structures (nuclear lamina)
- viruses
- membranes



Protein Function – Some General Principles

- All proteins bind to other molecules
- Protein binding has a high degree of specificity for its *ligands* (binding partners)
- Ligand specificity and affinity depends on formation of sets of weak non-covalent bonds and hydrophobic interactions.

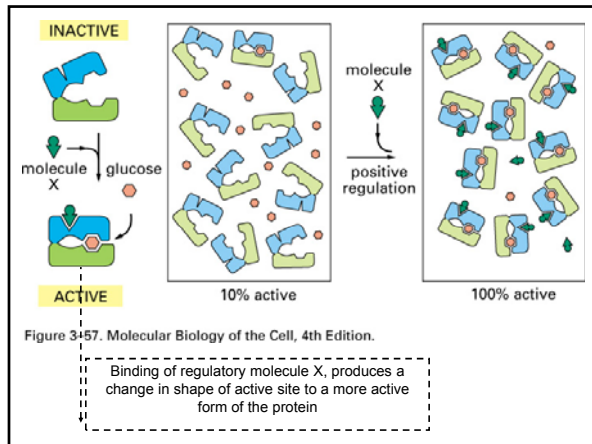


Enzymes are highly specific catalysts

- Enzymes speed reactions by selectively stabilizing unstable *transition states* (conformations) of their ligands
- This lowers the activation energy of the reaction
- so endeth the lesson on enzymology...

The catalytic activities of many enzymes are highly regulated through small molecule binding sites

- **Allosteric enzymes** have two or more binding sites that interact with other molecules
 - an *active site* that recognizes substrates
 - a *regulatory site* that recognizes a regulatory molecule
- binding of a regulatory molecule at one site on the protein causes a conformational change in the polypeptide that can switch the active site conformation “On” or “Off”



In cells, the function of many proteins is to bind to other proteins

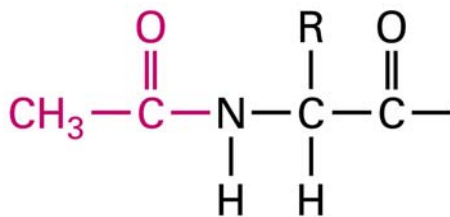
- Binding can occur through several types of interfaces
 - contact with an extended loop of a polypeptide chain on another protein
 - coiled coil interactions between two α -helices
 - precise matching of complementary rigid surfaces between two proteins

How is protein function regulated?

- Cells can regulate the steady state levels of proteins through synthesis or degradation
 - Control of mRNA levels through transcription or mRNA stability regulation
 - Control of translation of a protein's mRNA
 - Targeted degradation of a protein through proteolysis
- Changing the activity of a protein through conformational changes
- Changing the location of a protein by moving it to a different part of the cell

Protein structure and function can also be regulated by covalent modifications of exposed residues.

N-terminal acetylation stabilizes proteins
 – non-acetylated proteins are degraded rapidly by proteases

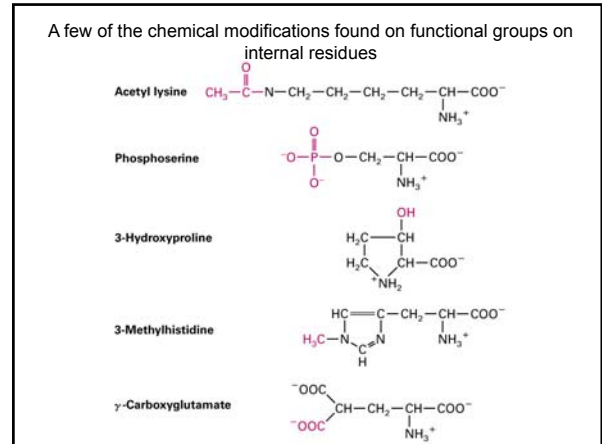
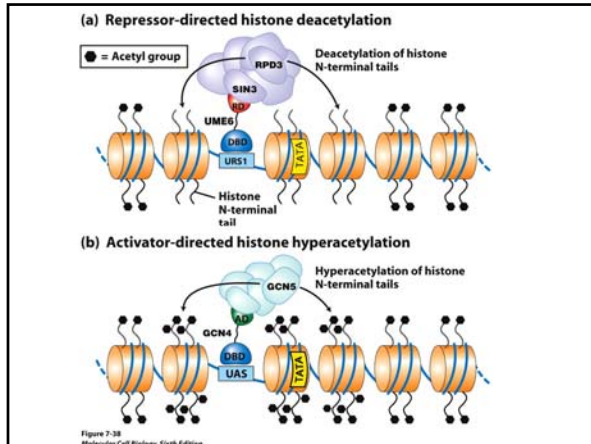


Acetylated N-terminus

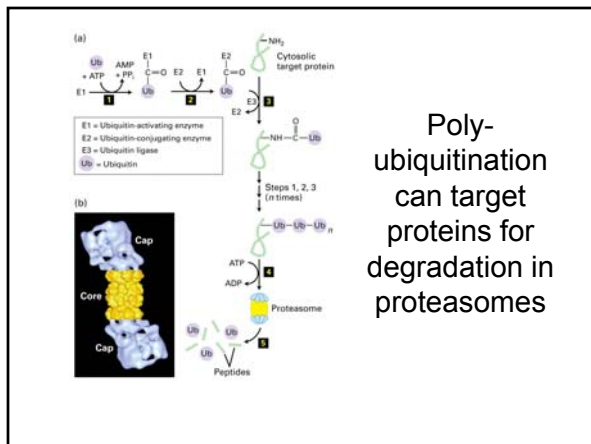
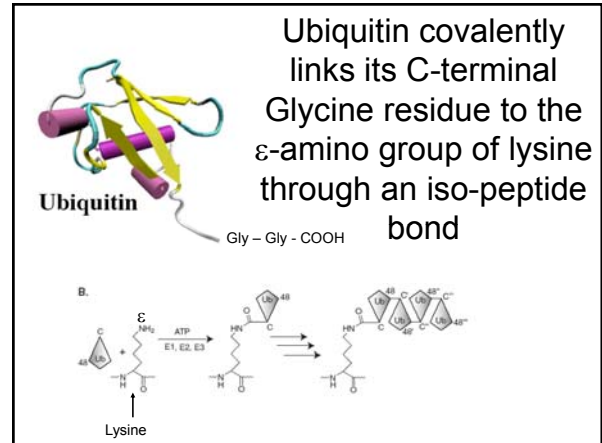
Acetylation and Deacetylation of Histone Tails Control Transcription Activity

Deacetylation inhibits binding of transcription factors to the TATA box, repressing gene expression

Hyperacetylation of histone N-terminal tails facilitates access of general transcription factors needed for transcription initiation



Ubiquitin is one of a family of small proteins that can be covalently linked to the ϵ -amino group of exposed lysine residues



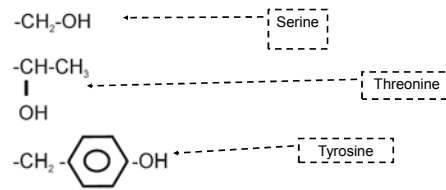
Monoubiquitination, or Linkage of Small Ubiquitin-Like Molecules (SUMOs) Can Regulate Protein Structure and Activity

In cells, many changes in protein binding / catalytic functions are driven by phosphorylation

Cells contain a large collection of protein kinases and phosphatases

What amino acids are phosphorylated?

Phosphorylated amino acids



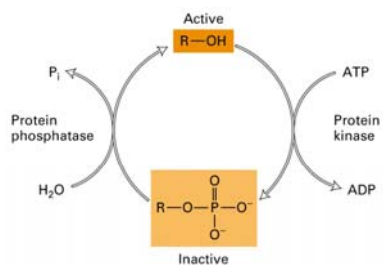
Protein phosphorylation and dephosphorylation play a major role in *regulating enzyme activity* and in *driving the regulated assembly and disassembly of protein complexes*

- Addition of a phosphate group (2 – charges) to a residue can attract + charged side chains, causing major conformation changes
- Attached PO₄⁻ groups can form structures that can be recognized as binding sites by other proteins

Protein phosphorylation is reversible and can act as a molecular switch

- dephosphorylation can restore original conformation and activity of the protein

Phosphorylation / dephosphorylation

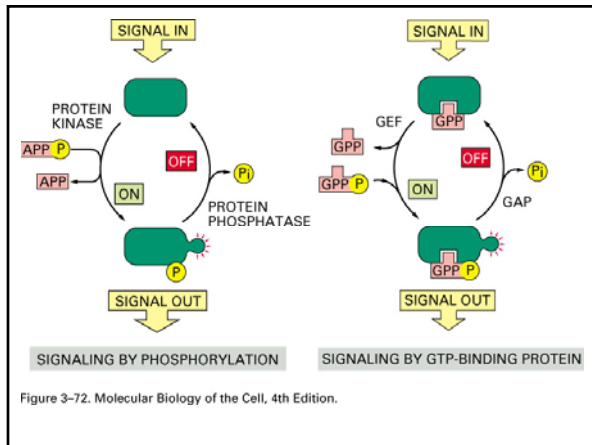
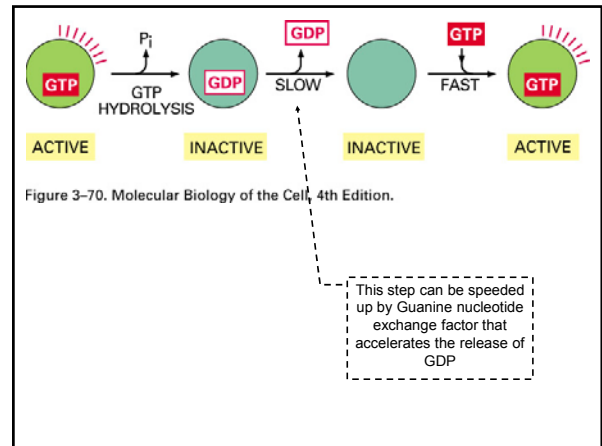
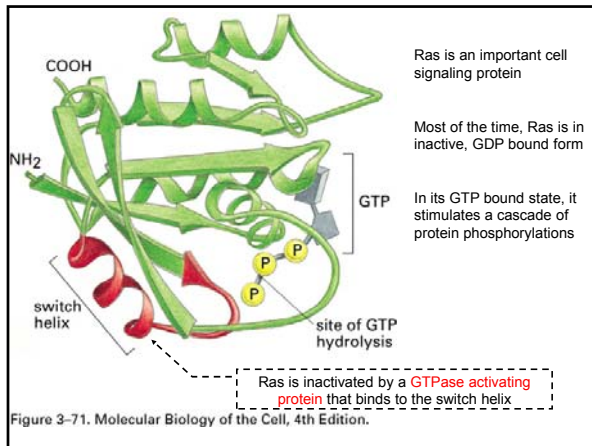


Note: In this case the *dephosphorylated* form of the protein is active

figure

Other proteins bind and hydrolyze GTP to act as a molecular switch (GTP binding proteins or GTPases)

- Actually another form of phosphorylation/ dephosphorylation
- GTP binds tightly to protein, usually activating it
- Protein can self-catalyze conversion from GTP to GDP
- Conformational change converts protein to inactive form



Phosphoproteins can serve as signal integrators for a molecular switch

- **Example:** Activation of a protein requires the input of multiple signals from different parts of the cell
- cdk kinase (cyclin dependent kinase) – involved in cell division

