

Protein Modification, targeting and degradation

Protein modification

- Proteins undergo a variety of modifications that are critical for function. There are numerous amino acid modifications such as collagen.



Three collagen molecules.

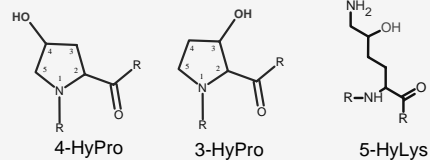
Each molecule is composed of a left-handed helix.

These are not α -helices.

Three such helices are coiled to form a right-handed superhelix

Collagen's unique blend of amino acids

- ~ 30% of residues are Glycine
- ~ 30% of residues are Proline or Hydroxyproline (HyPro)
- 5-hydroxylysine (HyLys) also occurs; a site for glycosylation
- Hydroxylation of Pro, Lys is a post-translational modification, requires vitamin C as a reactant
- The sequence of collagen bears long stretches of Gly--Pro/HyPro-X repeats



Structure of Collagen

- Collagen makes up 25 to 35% of the total protein of mammals.
- It is found in all forms of connective tissue.
- Collagen itself is an insoluble protein because of extensive cross-linking.

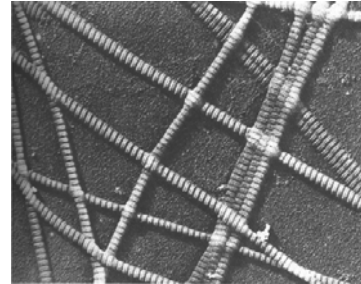
Structure of Collagen (continued)

- The structural unit of collagen is a tropocollagen, a supercoil made up of 3 helices, with a molecular mass of ~285 kdal.
- Each collagen helix consists of ~ 1000 amino acid residues. The helix is left-handed. **It is not an α -helix.**
- The helix contains 3 amino acids per turn, with a pitch of 0.94 nm.

Structure of Collagen (continued)

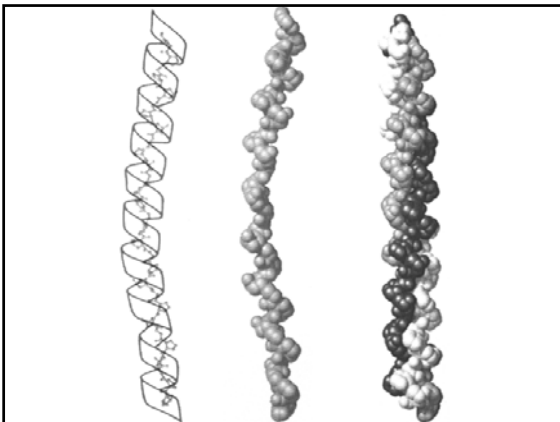
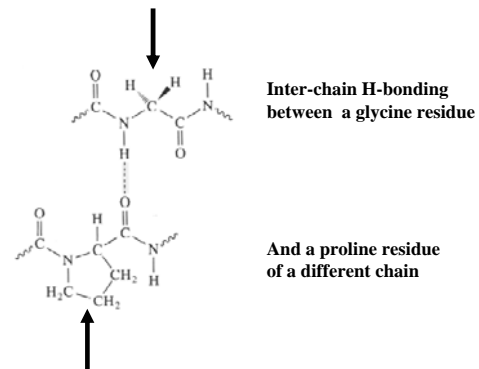
- Each unit of tropocollagen is about 1.5 nm wide and 300 nm long. Bundles of these 3-stranded supercoils can be seen as collagen fibres in the electron microscope.
- Mature collagen has extensive covalent crosslinkages between individual collagen molecules.

An electron micrograph of collagen from skin



Structure of Collagen (continued)

- There is no intra-helical H-bonding in collagen helices.
- Rather, H-bonding occurs between the amide N of glycine residues in the central axis and the carbonyls of other residues in the adjacent chains. Often proline and hydroxyproline are involved.



Amino acid sequence of collagen

H_3N^+ -Glu - Met - Ser - Tyr - Gly - Tyr - Asp - Glu - Lys - Ser - Ala - Gly - Val - Ser - Val - 15
 Pro - Gly - Pro - Met - Gly - Pro - Ser - Gly - Pro - Arg - Gly - Leu - Hyp - Gly - Pro - 30
 Hyp - Gly - Ala - Hyp - Gly - Pro - Gln - Gly - Phe - Gln - Gly - Pro - Hyp - Gly - Glu - 45
 Hyp - Gly - Glu - Hyp - Gly - Ala - Ser - Gly - Pro - Met - Gly - Pro - Arg - Gly - Pro - 60
 Hyp - Gly - Pro - Hyp - Gly - Lys - Asn - Gly - Asp - Asp - Gly - Glu - Ala - Gly - Lys - 75
 Pro - Gly - Arg - Hyp - Gly - Gln - Arg - Gly - Pro - Hyp - Gly - Pro - Gln - Gly - Ala - 90
 Arg - Gly - Leu - Hyp - Gly - Thr - Ala - Gly - Leu - Hyp - Gly - Met - Hyl - Gly - His - 105
 Arg - Gly - Phe - Ser - Gly - Leu - Asp - Gly - Ala - Lys - Gly - Asn - Thr - Gly - Pro - 120
 Ala - Gly - Pro - Lys - Gly - Glu - Hyp - Gly - Ser - Hyp - Gly - Glx - Asx - Gly - Ala - 135
 Hyp - Gly - Gln - Met -

Defective Hydroxylation Is One Of The Biochemical Lesions in Scurvy

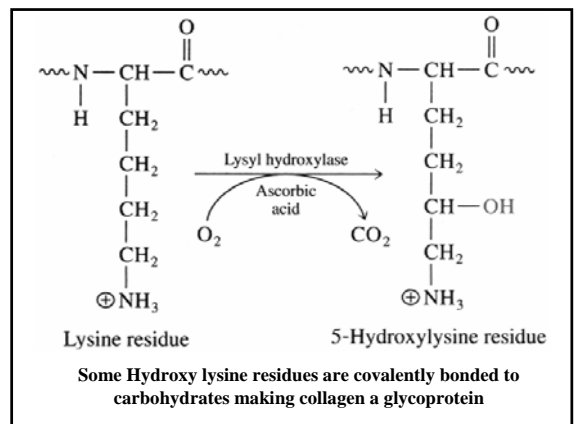
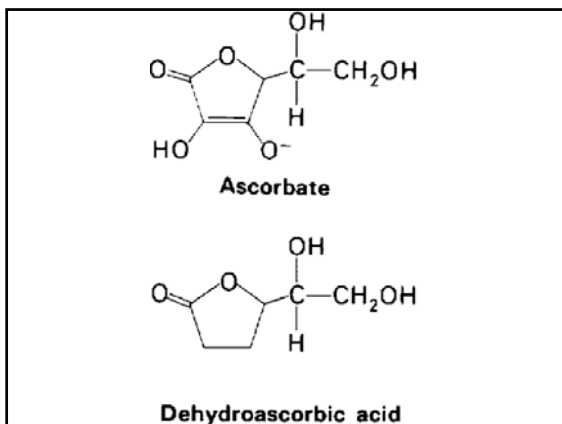
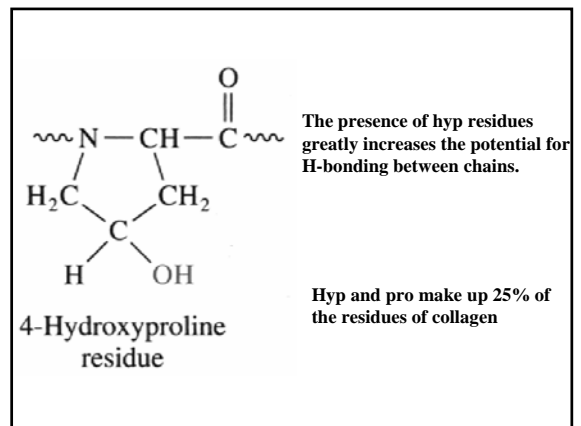
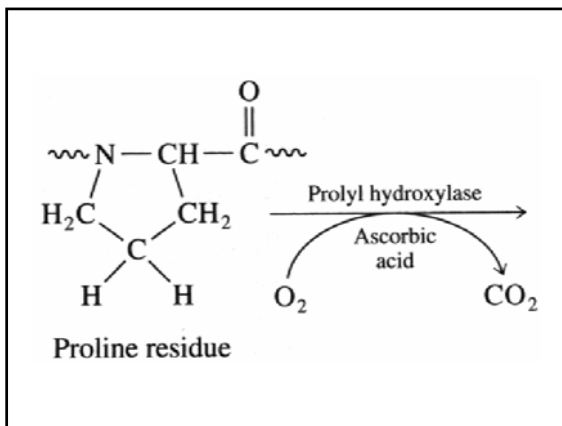
The importance of the hydroxylation of collagen becomes evident in *scurvy*. A vivid description of this disease was given by Jacques Cartier in 1536, when it afflicted his men as they were exploring the Saint Lawrence River:

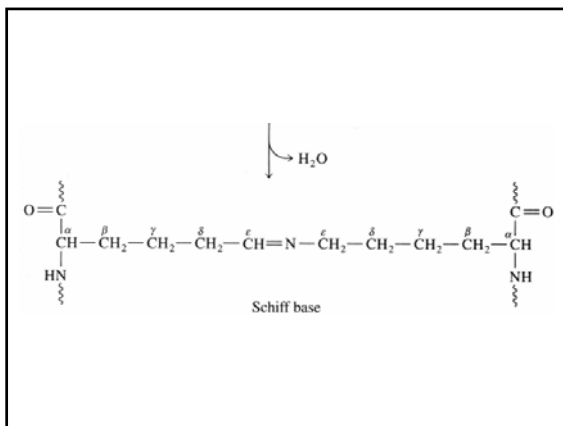
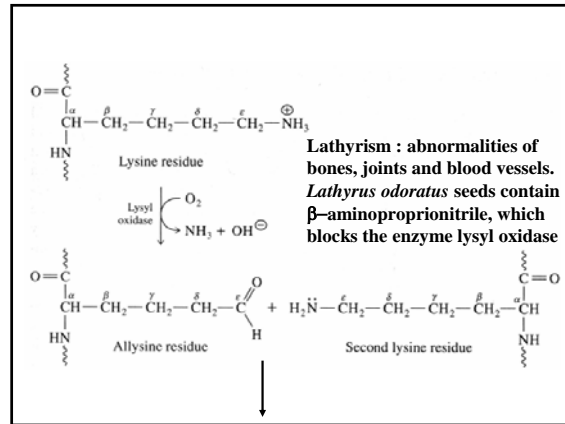
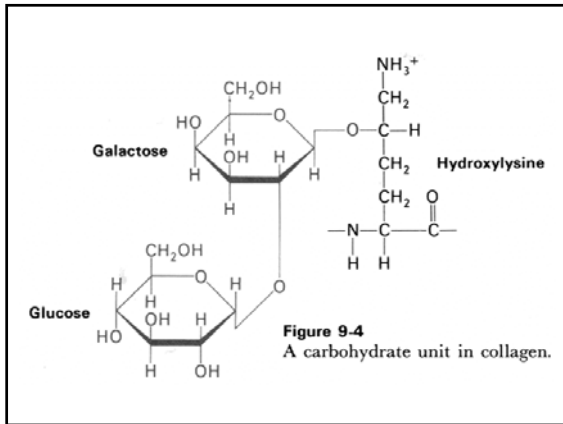
Some did lose all their strength, and could not stand on their feet... Others also had all their skins spotted with spots of blood of a purple colour: then did it ascend up to their ankles, knees, thighs, shoulders, arms, and necks. Their mouths became stinking, their gums so rotten, that all the flesh did fall off, even to the roots of the teeth, which did also almost all fall out.

The means of preventing scurvy was succinctly stated by James Lind, a Scottish physician, in 1753:

Experience indeed sufficiently shows that as greens or fresh vegetables, with ripe fruits, are the best remedies for it, so they prove the most effectual preservatives against it.

Lind urged the inclusion of lemon juice in the diet of sailors. His advice was adopted by the British Navy some forty years later.





Genetic disorders of collagen are numerous :

One such group of 10 different collagen deficiency diseases is the Ehlers-Danlos (E-D) syndrome.

The "India-rubber man" of circus fame had an E-D syndrome.

Osteogenesis imperfecta : abnormal (fragile) bone formation in human babies

Paganini may also have suffered from an E-D condition known as Marfan's syndrome.

Osteogenesis imperfecta (brittle bone disease)

- Type II :
- Most severe form, frequently lethal at or shortly after birth, often due to respiratory problems.
- In recent years, some people with Type II have lived into young adulthood.
- Numerous fractures and severe bone deformity; small stature with underdeveloped lungs.
- Collagen is improperly formed.

Proteins can also be cleaved into smaller functional units

Activation of Proinsulin

Insulin consists of two short, disulfide linked peptides. Synthesized in the pancreas as preproinsulin. An N-terminal signal peptide directs preproinsulin into the secretory pathway. The signal peptide is cleaved off in the ER, leaving proinsulin. Proinsulin folds into a specific tertiary structure with the correct disulfide bonds in secretory granules of β cells. Elevated blood glucose levels trigger insulin secretion. Proinsulin is converted into active insulin by proteolytic removal of the 33 residue C-peptide.

How do molecules get in and out of the nucleus?

Protein transport into the nucleus (NLS signal)

- Two major types of signals have been identified for the nuclear import of proteins: SV40 type and bipartite type. The former was first found in the large T antigen of the SV40 virus. It has the following sequence
- PKKKRKV
- This type of signal is characterized by a few consecutive basic residues and in many cases also contains a proline residue.
- The bipartite type was first identified in *Xenopus* nucleoplasmin with the following NLS:
- **KRPAATKKAGQAKKKK**
- Its characteristic pattern is: two basic residues, 10 spacer residues, and another basic region consisting of at least 3 basic residues out of 5 residues.

Importins and exportins

- After RNA molecules (mRNA, tRNA and rRNA) are produced in the nucleus, they must be exported to the cytoplasm for protein synthesis. In addition, proteins operating in the nucleus must be imported from the cytoplasm. The traffic through the nuclear envelope is mediated by a protein family which can be divided into **exportins** and **importins**. Binding of a molecule (a "cargo") to exportins facilitates its export to the cytoplasm. Importins facilitate import into the nucleus.

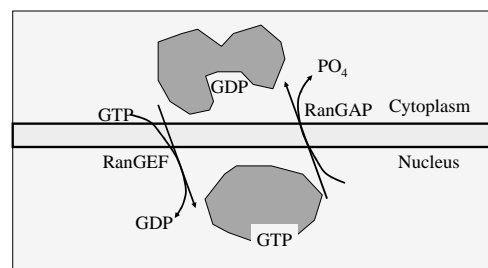
Importins and exportins are regulated by Ran

- Like other G proteins, Ran can switch between GTP-bound and GDP-bound states. Transition from the GTP-bound to the GDP-bound state is catalyzed by a GTPase-activating protein (**GAP**) which induces hydrolysis of the bound GTP. The reverse transition is catalyzed by guanine nucleotide exchange factor (**GEF**) which induces exchange between the bound GDP and the cellular GTP.

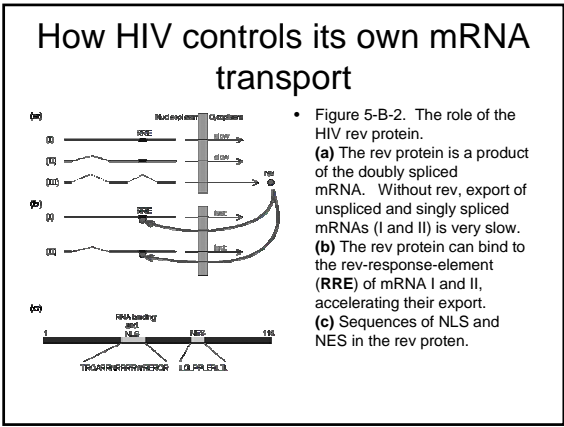
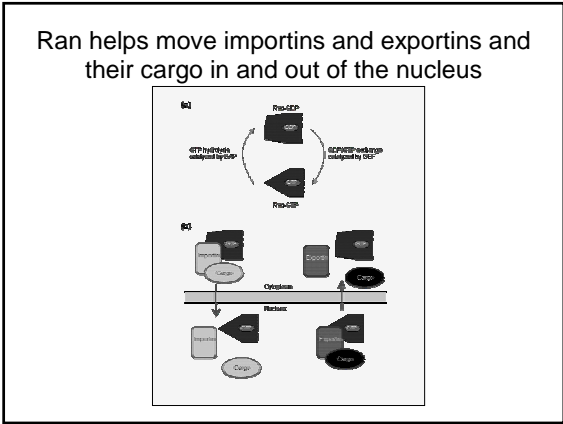
How Ran works

- The GEF of Ran (denoted by RanGEF) is located predominantly in the nucleus while RanGAP is located almost exclusively in the cytoplasm. Therefore, **in the nucleus Ran will be mainly in the GTP-bound state** due to the action of RanGEF while **cytoplasmic Ran will be mainly loaded with GDP**. This asymmetric distribution has led to the following model for the function of exportins and importins.

Ran switches between GDP bound and GTP bound



- **RanGTP enhances binding between an exportin and its cargo but stimulates release of importin's cargo; RanGDP has the opposite effect**, namely, it stimulates the release of exportin's cargo, but enhances the binding between an importin and its cargo. Therefore, the exportin and its cargo may move together with RanGTP inside the nucleus, but the cargo will be released as soon as the complex moves into the cytoplasm (through nuclear pores), since RanGTP will be converted to RanGDP in the cytoplasm. By contrast, the importin and its cargo may move together with RanGDP in the cytoplasm, but the cargo will be released in the nucleus since RanGDP will be converted to RanGTP in the nucleus.

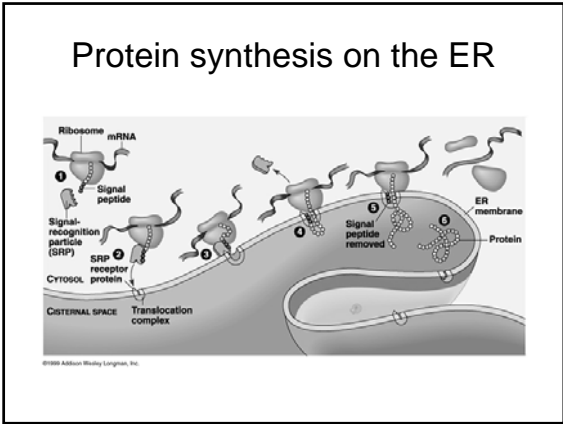


How do proteins find their way to the endoplasmic reticulum

- 1. An mRNA encoding the secretory protein binds to a cytosolic ribosome.
- 2. The first 70 or so amino acids are translated. Since approximately 30 amino acids of the protein remain buried in the ribosome at any one time this leaves approximately 30 amino acids on the N-terminus of the protein sticking out of the ribosome.
- These 30 amino acids encode a signal sequence, which binds with the signal recognition particle or SRP in the ribosome.

Then

- Initially the signal recognition protein (SRP) is bound to GDP when it binds the signal sequence this triggers release of GDP and binding of GTP by the SRP. This blocks further protein synthesis.
- The complex of SRP, ribosome and GTP binds to an SRP receptor on the ER which is also bound to GTP.
- Both GTPs (one on SRP and one on the SRP receptor) are hydrolyzed and this powers transfer of the polypeptide to the translocon (a channel made of several proteins). The translocon gate opens allowing entry of the polypeptide.



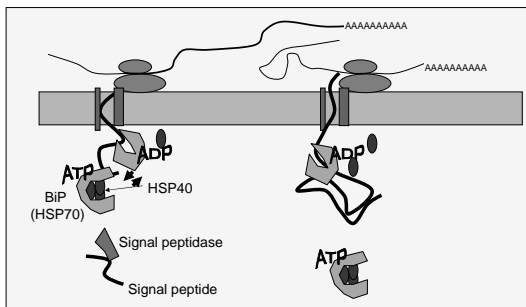
Protein folding

- 1) The signal peptide is cleaved within lumen by **signal peptidase**
- 2) **BiP** (a chaperonin) helps protein fold correctly. It is a member of the HSP70 family of heat shock proteins. When bound to ATP it is in the open state and weakly binds to target protein. But with the help of HSP40 proteins it hydrolyzes ATP to ADP. This leads to a conformational change that causes BiP to clamp tightly to hydrophobic regions of the protein. This process is repeated over and over until protein is folded into its final form.
- 3) protein is soluble inside lumen where it can be further modified

Chaperones

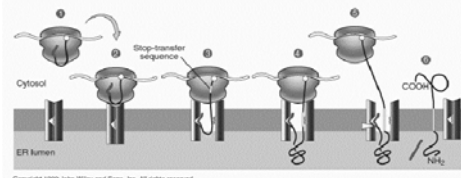
- Chaperones are proteins in the cell which function to help proteins to fold correctly
 - molecular chaperones guide folding of proteins present in the cytosol, lumen of RER, mitochondria, etc.
 - chaperones also promote the assembly of protein complexes from subunits
 - chaperones prevent the aggregation of unfolded proteins
- heat shock proteins**
 - a set of proteins induced by a brief exposure of cells to elevated temperature (42° C)
 - many of molecular chaperones are heat shock proteins (hsp)
 - the heat shock causes many proteins to unfold or misfold and the hsp are induced to help refold these proteins correctly
- examples of heat shock proteins:
 - hsp 70 - cytosol hsp 60 - cytosol BiP - in lumen of RER (BiP is an abbreviation for**
 - binding protein) mitochondrial hsp - (mhsp 60 and mhsp 70)**
- mechanism of action:
 - bind to exposed hydrophobic regions of proteins to achieve proper folding ATP requirement for action

Protein folding by BiP

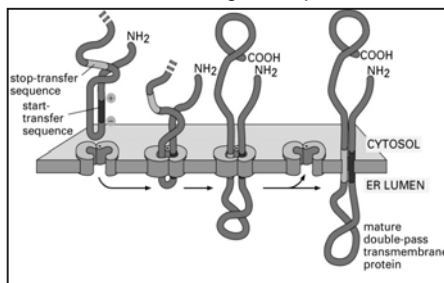


Membrane proteins

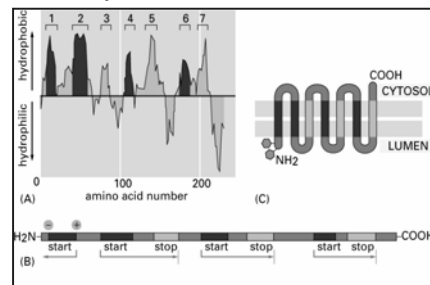
- Complications:** proteins embedded in membranes
- protein contains a **stop-transfer** sequence which is **too hydrophobic** to emerge into aqueous environment of ER lumen
- stop-transfer sequence therefore gets stuck in membrane
- ribosome lets go of translocon, finishes job in cytoplasm
- translocon dissociates, leaves protein embedded in membrane
- example = LDL receptor



Further complication of membrane proteins.
Insertion of double pass transmembrane protein with internal signal sequence



Multiple transmembrane sequences are common



Disulfide bonds form between cysteines

- PDI protein disulfide isomerase works in the ER. In the cytosol most Cysteines are in the reduced state partly because of active oxygen radical scavengers. In the ER PDI works by forming disulfide bonds with the target protein and then transferring that bond to another cysteine within the target protein.

Further protein modification

Why glycosylation?

- Aids in proper protein folding.
 - Provides protection against proteases (e.g. lysosomal membrane proteins)
 - Employed for signaling.
- Most soluble and membrane-bound proteins made in the ER are glycoproteins, in contrast to cytosolic proteins.
 - Glycoprotein synthesis is a 3-part process:
 1. Assembly of the precursor oligosaccharide
 2. En-bloc transfer to the protein
 3. Modification of the oligosaccharide by removal of sugars

Protein glycosylation

1. Assembly of the precursor oligosaccharide...

IMPORTANT POINTS:

- Assembly takes place on the carrier lipid dolichol, anchored in the ER membrane.
- A pyrophosphate bridge joins the 1st sugar to the dolichol.
- Sugars are added singly and sequentially.
- After the two N-acetylglucosamines are added, the assembly flips from the cytosolic side to the ER lumen.
- Nine mannose and three glucose molecules are added, totaling 14 sugars.

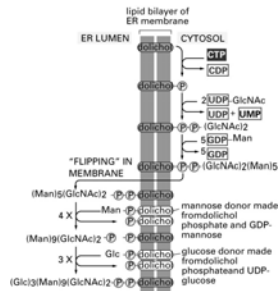


Figure 12-53. Molecular Biology of the Cell, 4th Edition.

Now in the second step, attachment

2. En-bloc transfer of the oligosaccharide to the protein

- One step transfer, catalyzed by oligosaccharyl transferase, which is bound to the membrane at the translocator.
- Covalently attached to certain asparagines in the polypeptide chain (said to be "N-linked" glycosylation).
- Attaches to NH₂ side chain of Asn but only in the context: **Asn-x-Ser or Asn-x-Thr**

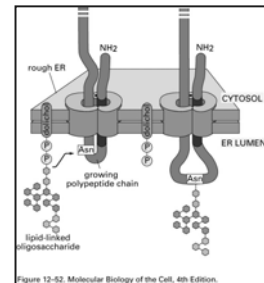


Figure 12-52. Molecular Biology of the Cell, 4th Edition.