

# Protein Biosynthesis

April 11, 2016

1

# DNA the Blueprint

*DNA directs its own replication and its transcription to yield RNA, which in turn, directs its translation to form proteins.*

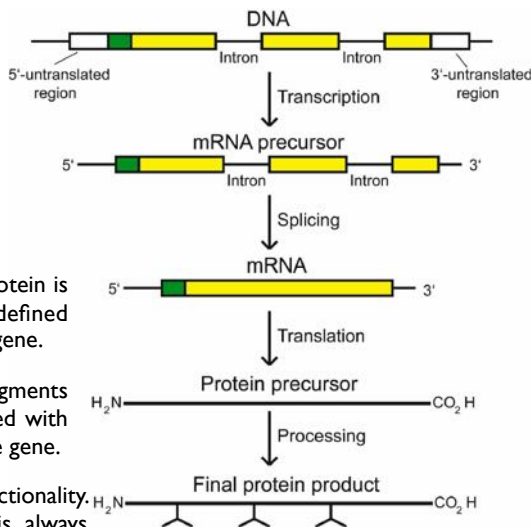
Francis Crick, 1958

- Transcription refers to the process of transferring information from DNA to RNA
- Translation refers to the production of proteins based on the information (template) provided by the RNA.
- The precise control in the amino acid sequences of proteins is due in large part to this template strategy.

2

## From DNA to Proteins

- All of the information for assembling the primary structure of a protein is encoded in the genetic material of the chromosomes.
- Coded in from an alphabet of 4\* nucleotide bases:
  - Adenine (A)
  - Cytosine (C)
  - Guanine (G)
  - Thymine (T)/Uracil (U)
- The primary sequence of a protein is coded for by a specific defined region of DNA/RNA called a gene.
- These genes are flanked by segments of DNA/RNA that are involved with regulation of expression of the gene.
- DNA and genes have directionality. The information in a gene is always read from the 3' end of the coding sequence to the 5' end.



3

## DNA → RNA

- First step in conveying the information in a gene is transcription of the DNA nucleotide sequence into a complementary messenger RNA (**mRNA**) molecule.
- *Note:* some genes do not code for proteins, but instead code for stable RNA molecules such as tRNA and rRNA.
- Gene transcription is highly regulated. Expression of a specific gene can be up-/down-regulated or turned off/on in response to external stimuli or the cellular condition.

4

# DNA → RNA: Regulation

- Gene regulation can be very complex and the details are often not well understood.
- Many regulatory proteins bind to regions upstream of the gene known as promoter regions.
- Other regulatory proteins act at enhancer regions, which may be upstream, downstream or within the gene itself.
- Binding of a regulatory protein can be either a positive or a negative effect on transcription. A variety of mechanisms are used for various genes. These mechanisms often involves the timely binding of several different regulatory proteins.

5

# DNA → RNA: RNA Polymerases

- RNA is transcribed from the gene by RNA polymerases.
- Prokaryotes utilize a single RNA polymerase to transcribe all genes. It is adapted to specific genes by sigma-factor proteins.
- Eukaryotes have multiple RNA polymerases (RNA polymerase I, II, and III).
  - ❖ RNA polymerase I and III transcribe genes that code for stable RNA.
  - ❖ RNA polymerase II transcribes genes that code for proteins.
- The two strands of DNA are unwound near the site where transcription begins, and the polymerase uses only one strand as the template for RNA assembly.
- Transcription is initiated at a site dictated by the specific sequence of the promoter region of the gene, and the template strand is read in a 3' to 5' direction.
- The mRNA strand is assembled sequentially by adding nucleotides (A, C, U and G) to the 3' end of the growing mRNA molecule.

# DNA → RNA: After Assembly

- The growing mRNA is only transiently associated with the DNA template. Only ~12 RNA-DNA base pairs are present lagging the polymerase.
- Transcription continues until the polymerase reaches a termination sequence.
- In eukaryotic cells, the 3' end of the newly transcribed mRNA is trimmed and a poly(A) tail is added (~250 bases).
- In prokaryotes, an mRNA molecule can contain coding regions (**cistrons**) for more than one proteins.
- Eukaryotic cells, an mRNA molecule codes for a single protein.
- In prokaryotic cells a single promoter will regulate the transcription of one or more cistrons (collectively known as an **operon**) into a single mRNA molecule. In eukaryotic cells, each gene has its own promoter.
- While the mRNA in prokaryotic cells is not modified after transcription, the mRNA in eukaryotic cells can be extensively modified.

7

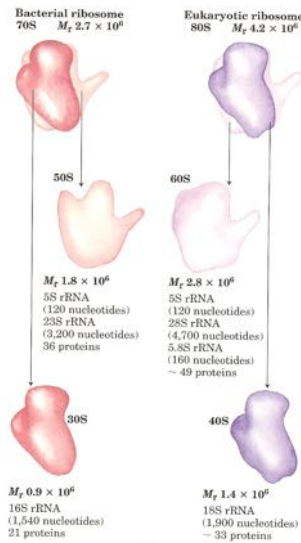
# DNA → RNA: Introns, Exons and Spliceosomes

- The transcribed mRNA in eukaryotic cells is cut and spliced to remove introns (noncoding sequences), splicing together the exons (coding sequences).
- This complicated process takes place in the spliceosome, a large complex consisting of several proteins and small RNA molecules.
- Not all of the sequences that define splicing sites are known, but introns invariable start with sequence GU and end with AG.
- Most known genes signal the removal of all introns from all mRNA molecules (constitutive splicing) which produces a homogeneous population of mature mRNA molecules.
- In the case of certain genes, not all introns are spliced out. This results in new combinations of exons. (alternative splicing). Usually results in variant protein sequences.
- Alternative splicing often results in proteins with different destinations by coding for alternative forms: both with and without signaling sequences or membrane spanning segments.

8

# RNA → Protein: Translation

- Mature mRNA is translated into polypeptide sequences into the cytoplasm.
- Process involves ribosomes, tRNA's and other factors.
- In eukaryotes, the completed mRNA migrates from the nucleus into the cytoplasm. (In prokaryotes, translation occurs while the mRNA is being transcribed.)
- Translation is initiated with complexation of the mRNA, the two ribosomal subunits and three initiation-factor proteins.



9

Lehninger, Principles of Biochemistry, 4th ed.

## Genetic Code

- The code relating the sequence of mRNA to the amino acids is based on codons (nucleotide triplets), and there are 64 different nucleotide triplet combinations.
- Since these codons relate to only 20 different amino acids, the code is highly redundant. Most amino acids are designated by more than one triplet.
  - Only Trp and Met residues have only one codon.
  - Leu, Arg and Ser are each coded by 6 codons.
- Sixty-one codons code for the 20 amino acids. The remaining three codons (UAA, UGA and UAG).
- Many cells have tRNA species (suppressor tRNA) that have anticodons corresponding to the termination codons. These tRNA molecules can insert amino acid residues into the growing peptide, preventing termination and allowing continued chain elongation. They usually do this with low efficiency. Therefore, some genes employ multiple stop codons.
- The genetic code is nearly universal.
- However, a slightly different code may be used in certain cells and organelles.

11

## RNA → Protein: Translation

- In prokaryotes, ribosomes are assembled from 3 different rRNA molecules and over 50 different proteins (eukaryotic ribosomes are larger and contain 4 different RNA molecules).
- In the absence of bound mRNA, the ribosomes disassemble into large and small subunits (50S and 30S subunits in prokaryotes, 60S and 40S in eukaryotes).
- Ribosome binds near the 5' end of the mRNA molecule. Translation is initiated at the point in the sequence that signals to initiate translation.
- mRNA is usually short lived, 2-10 min in bacteria, but longer lived in higher organisms, 0.5-24 h

### Typical Prokaryotic Ribosome

	Ribosome	Small Subunit	Large Subunit
Sed. Coeff.	70S	30S	50S
Mass (KD)	2520	930	1590
Major RNA's		16S=1542 bases	23S=2904 bases
Minor RNA's		5S=120 bases	
RNA mass (KD)	1664	560	1104
RNA proportion	66%	60%	70%
Protein Number		21	31
Protein Mass (KD)	857	370	487
Protein Proportion	34%	40%	30%

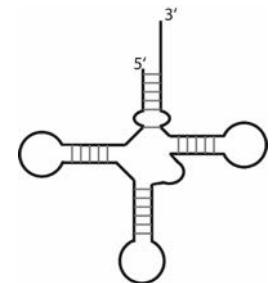
### Typical Eukaryotic Ribosome

	Ribosome	Small Subunit	Large Subunit
Sed. Coeff.	80S	40S	60S
Mass (KD)	4220	1400	2820
RNA mass (KD)	2520	700	1820
RNA proportion	60%	50%	65%
Protein Number		33	49
Protein Mass (KD)	1700	700	1000
Protein Proportion	40%	50%	35%

10

## tRNA and Amino Acid Activation

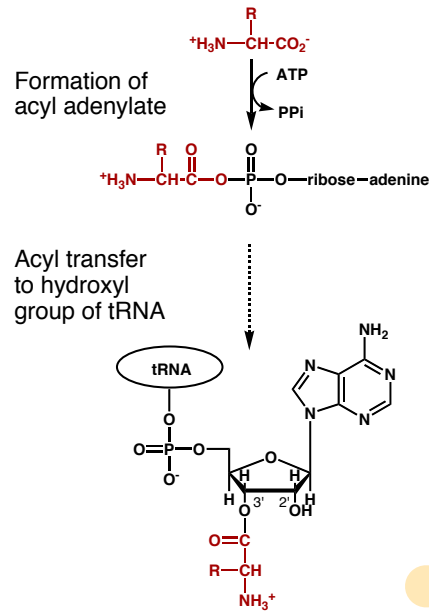
- tRNA's are one of the classes of structural RNA, and have a characteristic "cloverleaf" secondary structure.
- They serve as carriers shuttling amino acid residues to the ribosome-mRNA complex to be incorporated in the growing peptide chain.
- All tRNA molecules are 73-93 nucleotides in length with varied sequences (the sequence of a specific tRNA is defined by a corresponding gene).
- tRNA's are amino acid specific and assure that the amino acids are assembled in the order proscribed by the mRNA template.
- There is at least one class of tRNA molecule in every cell for each type of amino acid.



12

# Charging of tRNA

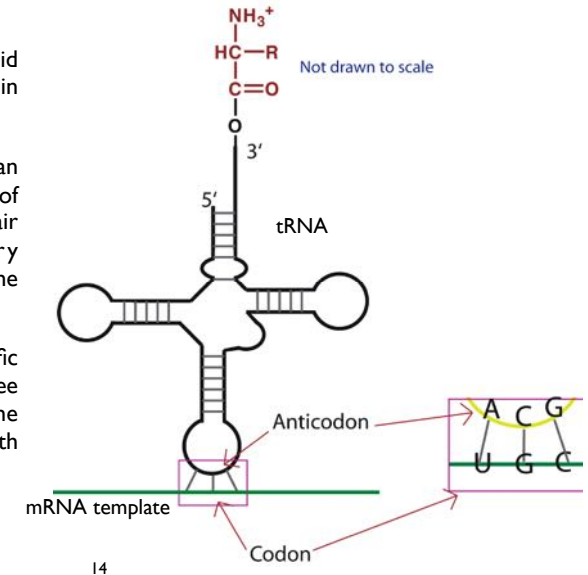
- Coupling of an amino acid to the appropriate t-RNA molecule is carried out by the appropriate aminoacyl tRNA synthetase.
- The amino acid is coupled through the  $\alpha$ -carboxyl group, activating it for peptide bond formation.
- The chemical energy for activation comes from ATP hydrolysis. An aminoacyl adenylate intermediate and pyrophosphate are formed.
- The amino acid moiety is then transferred to the 2' or 3' hydroxyl group of the 3'-terminal A nucleotide of the appropriate tRNA.
- Aminoacyl tRNA synthetases are very specific for both the amino acid and the tRNA being charged.



13

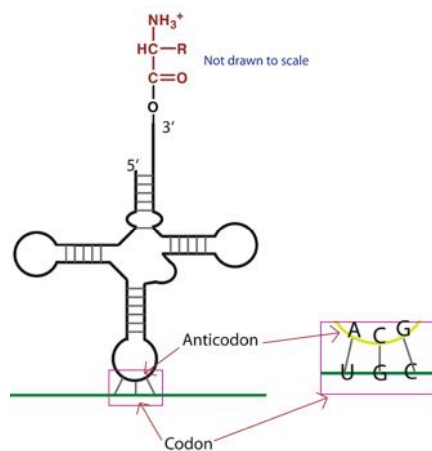
# A Codon by Any Other Name

- Each tRNA links an amino acid to the information contained in the mRNA template.
- Each tRNA molecule contains an **anticodon**, which consists of three nucleotides that base-pair with three complementary nucleotides (a **codon**) in the template mRNA.
- This base-pairing is very specific for at least two of the three sequential nucleotides in the codon. Can be “wobble” with the third nucleotide position.



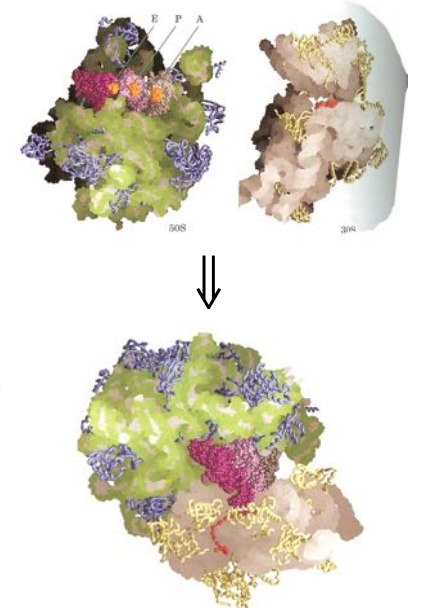
# Initiation of Assembly

- **Initiation of Translation:** requires the presence of a N-formylmethionyl tRNA ( $tRNA_f^{Met}$ ), which interacts with the initiating AUG codon on the mRNA.
- $tRNA_f^{Met}$  is only involved with initiating translation and is unique for other  $tRNA^{Met}$ , which would be used for the incorporation of Met at other positions.
- In prokaryotes, the initiating AUG codon is identified by the ribosome based on interactions between a short upstream segment of mRNA and ribosomal RNA.
- In eukaryotes, the signals that identify the initiating codon are not as clear. In 90-95% of genes, it is the first AUG encountered reading from the 5' end of the mRNA.
- Binding of  $tRNA_f^{Met}$  with the ribosome-mRNA complex is complicated. (In prokaryotes it requires three initiation factors and energy from GTP hydrolysis.)



# Chain Assembly

- The initiating  $tRNA_f^{Met}$  binds to the initiating AUG codon of the mRNA at a site on the ribosome designated P (peptidyl tRNA).
- The growing polypeptide chain is held at the P site.
- An amino acid-charged tRNA then comes in and binds downstream of the initiating codon at the adjacent A site (aminoacyl-tRNA position).
- The initiating Met is then transferred to the free amino group of the amino acid-charged tRNA in the A site.
- The uncharged  $tRNA_f^{Met}$  is then released and the dipeptide-tRNA in the A site is transferred to the P site.
- Uncharged tRNA's may migrate to an E site before they completely disassociate from the ribosome-mRNA complex.

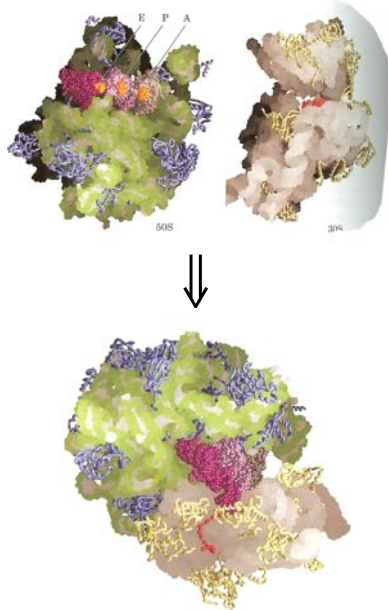


16

Lehninger, Principles of Biochemistry, 4th ed

# Chain Assembly

- Effectively the ribosome moves three nucleotide bases along the mRNA, placing the next codon in A site.
- Migration of the peptide-tRNA to the P site opens up the A site. This allows the next aminoacyl-tRNA in the sequence to bind.
- The process is repeated until the full length peptide has been assembled.
- Via this process, a ribosome can add 1-10 residues/sec to the growing peptide chain.
- The reading frame employed by the ribosome is determined by the position of the initiating AUG.
- The assembly continues until a termination codon is reached (UAA, UGA or UAG). These are recognized by release factors, which cause hydrolysis of the linkage between the polypeptide chain and the last tRNA in the P site.



Lehninger, Principles of Biochemistry, 4th ed

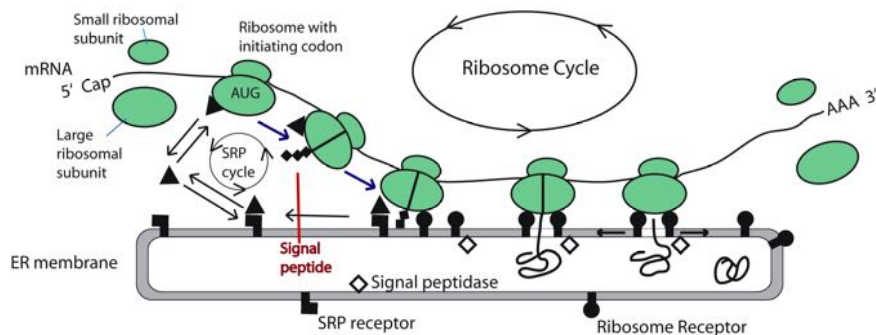
17

# Protein Trafficking and Topogenesis

- Newly synthesized proteins must be targeted to their appropriate locations within (or outside) the cell.
- Directing the trafficking, distribution and targeting of proteins is a means of controlling them and their function.
- Proteins which reside and function in the cytoplasm are usually synthesized by free cytoplasmic ribosomes, and they are released upon completion. A relatively straightforward event.
- For proteins bound for other locations the process is more complicated and necessitates that proteins pass through one or more membrane boundaries.
- Topogenesis** refers to the process by which proteins cross these membranes and the regulation of these processes.

18

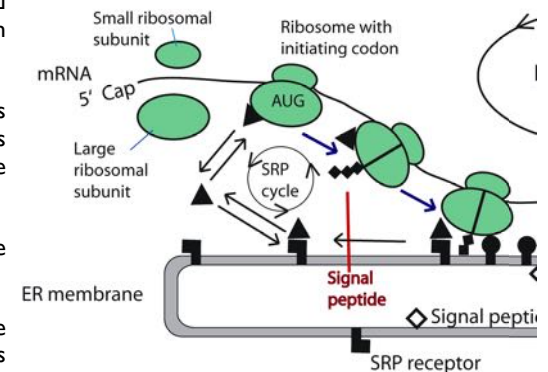
# Proteins Targeted to the ER



19

# Secreted, Lysosomal and Membrane Proteins

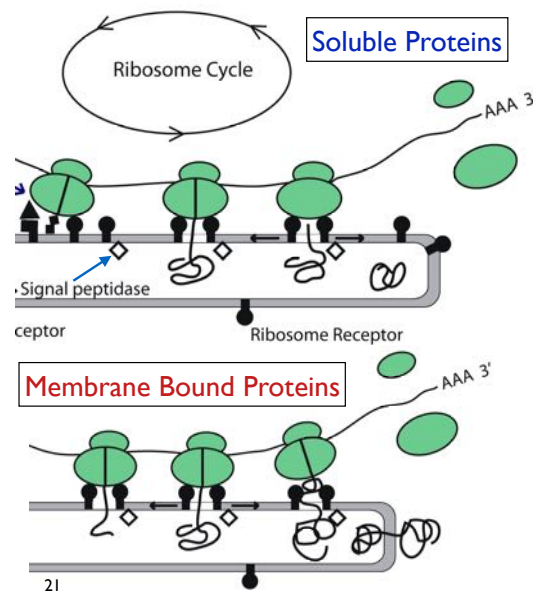
- In eukaryotic cells, proteins bound for secretion from the cell, for lysosomes, other organelles and cellular membranes begin translation on free ribosomes in the cytosol.
- An N-terminal **signal peptide** is synthesized first. This peptide is immediately recognized by the **signal recognition particle (SRP)**.
- SRP binding stops peptide elongation.
- Translation begins again when the ribosome complex interacts with as SRP receptor on the membrane of the ER.



20

## Secreted, Lysosomal and

- The SRP is displaced upon binding of the ribosome complex with the ER membrane, and the block on translation is released and chain extension continues.
- The growing peptide chain is extruded through the membrane into the lumen of the ER.
- The signal peptide is usually cleaved by a specific signal peptidase.
- The new N-terminus is usually assigned residue +1. Residues of the signal peptide are assigned negative numbers counting away from the cleavage site.
- Protein including the signal peptide referred to as the preprotein.
- Upon completion of synthesis, the protein is either released into the lumen or remains in the membrane.



21

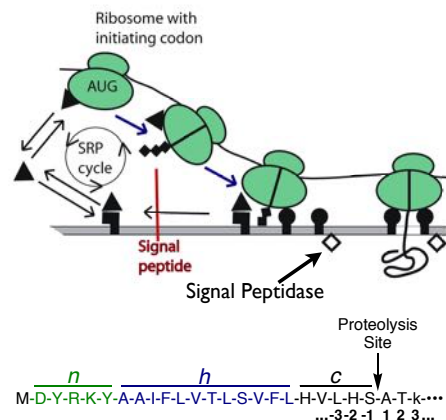
## Signal Peptides and the SRP

- The presence of a 15-26 residue N-terminal signal peptide (sequence) is usually sufficient to cause the protein to be inserted into the ER.
- The sequences of signal peptides are diverse. However they do share some common features, and these characteristics are sufficient to identify signal peptides.
  - There is usually a charged (+) region of 2-5 residues at the N-terminus, after the initiating Met. (region *n*)
  - After the *n* region is a region of 7-15 residues (primarily hydrophobic residues: Leu, Ile, Val, Ala and Phe.) (region *h*)
  - h* region is followed by ~5 predominantly polar residues. (region *c*)
- The positive charge at the N-terminus is believed to assist with insertion into the membrane through electrostatic interactions with the negatively charged head groups of the membrane lipids. The hydrophobic region interacts favorably with the hydrophobic interior of the membrane.
- The signal recognition particle (SRP) acts as an adaptor between the cytoplasmic ribosomal assemblies and the membrane-bound translocation system. The SRP binds to ribosomes with translated signal peptides, halts translation and delivers the ribosomal assembly to the ER.
- The SRP is a complex of 6 polypeptides and one 300 nucleotide RNA.

22

## Binding to the ER

- SRP binding leads to the ribosomal complex delivering the ribosomal complex to the ER. Where the bound SRP is recognized by a receptor; an integral membrane protein, displayed on the ER surface.
- The ribosomal-mRNA-nascent peptide complex is then transferred to other proteins in the ER membrane (i.e. ribosome receptor and peptide translocation complex).
- This is followed by dissociation of the SRP, which restores translation.
- The growing polypeptide chain is guided through the membrane into the lumen of the ER by the peptide translocation complex (ATP driven process).
- The signal peptide is usually cleaved by a specific signal peptidase during or after translocation of the nascent chain across the membrane.



23

The signal peptidase is a large membrane-bound protein on the lumen face of the ER membrane. It consists of a complex of several polypeptide chains.

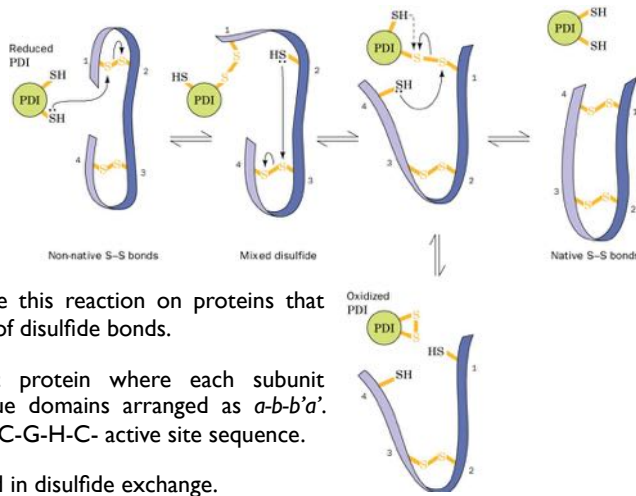
## Release, Folding and Assembly in the ER

- A number of modifications to the nascent polypeptide can occur in the lumen of the ER.
  - Covalent modification: glycosylation.
  - Disulfide formation.
  - At least partial folding.
- In the case of large multi-domain proteins, it appears that individual domains fold sequentially as they are synthesized and cross the membrane entering the lumen.
- The lumen of the ER contains a complement of folding auxiliary proteins: molecular chaperones (i.e. HSP 70 and Chaperonins), protein disulfide isomerase and peptidyl cis-trans isomerase.
- HSP 70 serves in a broad capacity, by binding to regions of unfolded proteins preventing aggregation.
- Some polypeptide chains that transiently associated with HSP 70 are part of larger multimeric proteins. Binding partially stabilizes the bound chains until they encounter other polypeptides that make up the folded multimeric protein.
- The HSP 70 process of binding and releasing polypeptide chains involves several other proteins such as HSP 40 and the hydrolysis of ATP.

24

# Protein Disulfide Isomerases

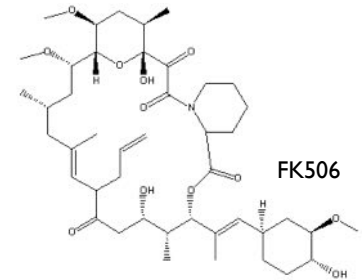
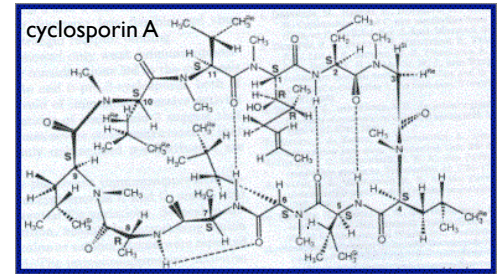
- Protein disulfide isomerases (PDI) catalyze disulfide interchange - facilitating the shuffling of disulfide bonds within a protein.
- Must be able to catalyze this reaction on proteins that denature in the absence of disulfide bonds.
- PDI is a homodimeric protein where each subunit consists of 4 100-residue domains arranged as  $\alpha$ -b-b' $\alpha$ '. PDI-a and PDI-a' have a -C-G-H-C- active site sequence.
- The first Cys participated in disulfide exchange.
- The second Cys functions to release PDI from the disulfide bond formed with the substrate.
- The active site Cys is exposed on the protein surface and is very reactive.



25

# Peptidyl Prolyl *cis-trans* Isomerase

- PPIs catalyze the interconversion of Xaa-Pro peptide bonds between their *cis* and *trans* conformations.
- There are two structurally unrelated families of proteins collectively called immunophilins responsible for this activity:
  - The cyclophilins: *because cyclophilin binds cyclosporin A, a model substrate in the cis conformation, it has been suggested that the enzyme catalyzes the trans to cis conversion.*
  - A family of proteins represented by FK506-binding protein (FKBP12).



26

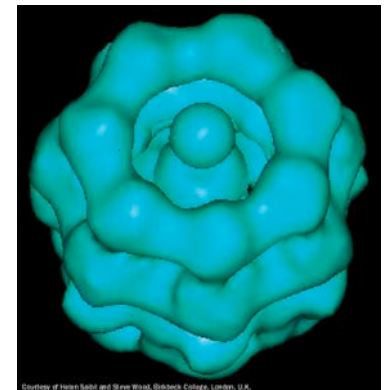
# Molecular Chaperones

- In vivo*, unfolded proteins have a tendency to aggregate.
- Molecular chaperones are proteins that function to prevent or reverse misfolding and improper associations.
- Many of these chaperones are ATPases, using ATP hydrolysis to drive the process.
- Several unrelated classes of chaperones:
  - Heat shock protein 70 (HSP70): highly conserved proteins. These are monomeric 70kD proteins that use ATP hydrolysis to drive the process of reversing protein denaturation and aggregation. Work in association with the chaperone HSP40
  - Chaperonins: pervasive large multi-subunit, cage-like assemblies. Bind improperly folded proteins via exposed hydrophobic surfaces. Use ATP hydrolysis to drive the refolding process within an internal cavity.
  - Heat shock protein 90 (HSP90): Mainly involved with folding of proteins associated with signal transduction. One of the most abundant proteins in eukaryotes.
  - Nucleoplasmins: involved with the assembly of nucleosomes (DNA packaging).

27

# Chaperonins: GroEL/GroES

- Chaperonins consist of two families of proteins that work together: Hsp60 (a.k.a. GroEL) and Hsp10 (a.k.a. GroES).
- HSP60/GroEL consists of 14 identical 60kD subunits arranged in a dimer of large heptameric ring-like structures.
- HSP10/GroES forms single heptameric rings.

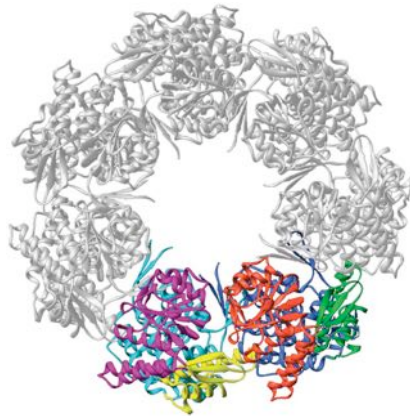


Courtesy of Helen Sahli and Steve Wood, Blacklock College, London, U.K.

28

## Chaperonins: GroEL/GroES

- X-ray crystal structure determined by Arthur Horwich and Paul Sigler.
- GroEL's 14 identical 547-residue monomers associate to form a thick-walled hollow cylinder. With two 7-fold symmetric rings stacked back to back (2 fold symmetry).
- A central channel  $\sim 45\text{\AA}$  in diameter runs through the middle of the complex.
- Each monomer possesses an ATP binding site facing the central chamber.
- GroEL monomers have hydrophobic patches at their apical domains.



29

## Chaperonins: GroEL/GroES

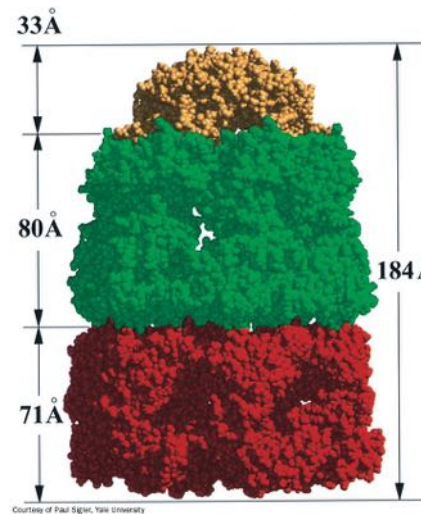
- X-ray structure determined by Lila Gierasch and Johann Deisenhofer.
- A dome like structure is formed by the association of 7 identical 97-residue monomers.
- The central region is "capped" by  $\beta$  hairpins which project from each of the monomers.
- The inner surface of the GroES dome is lined with hydrophilic amino acids.



30

## GroEL/GroES Complex

- The GroES heptamer and 7 ADP molecules bind to the same GroEL ring (the "cis ring" as opposed to the "trans ring").
- The apical and intermediate domains of the cis ring undergo large concerted *en block* movements.
- Results in widening and elongation of the cis cavity ( $85,000$  to  $175,000 \text{\AA}^3$ ).
- Changes allow the chamber to accommodate partially folded substrates of at least 70kD.
- *Trans* ring resembles structure of GroEL alone.



31

## Coordinated Conformational Changes

- The binding of GroES and ATP to the *cis* ring of GroEL strongly inhibits their binding to the *trans* ring.
- This is mediated through concerted small conformational shifts in the GroEL equatorial region.
- The *trans* ring can bind ATP once the *cis* ring has hydrolyzed its bound ATP molecules.
- Binding of ATP by the *trans* ring results in a conformational shift in the *cis* ring which leads to the release of GroES.
- The proper functioning of GroEL requires both the *cis* and *trans* rings.
- The *trans* ring cannot bind substrate, GroES or ATP until the *cis* ring has hydrolyzed its bound ATP.

32

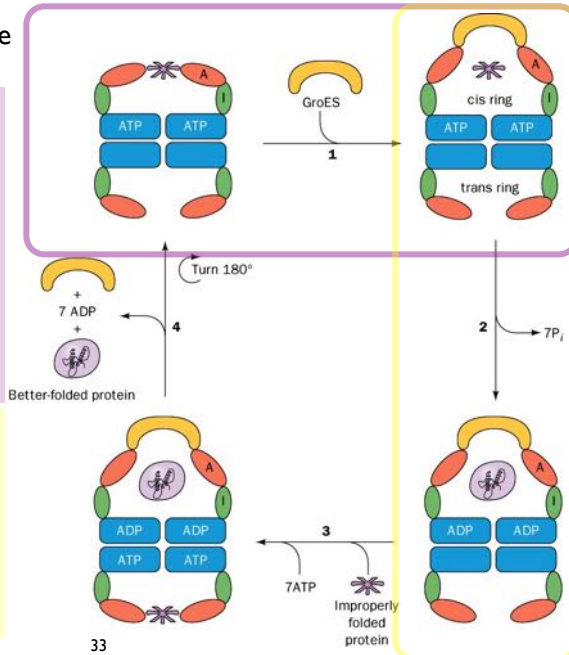


## GroEL/GroES in Action

- GroEL/GroES functions like a two stroke engine.

1. A GroEL ring binds 7 ATPs and substrate, then associates with GroES. Induces a conformational shift in the GroEL ring, releasing substrate into the ring cavity (*substrate starts to fold*). The cavity is now lined with hydrophilic groups to encourage proper folding.

2. Within ~13s the cis ring catalyzes the hydrolysis of the bound ATPs to ADP and P<sub>i</sub>. Hydrolysis of ATP results in weakening of the interactions binding GroES to the GroEL ring.

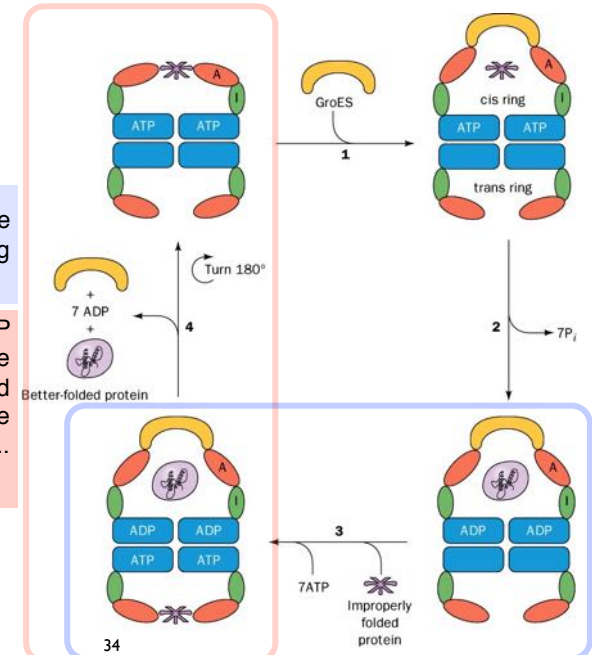


33

## GroEL/GroES in Action

1. A second substrate molecule binds to the *trans* ring followed by ATP.

2. Binding of substrate and ATP to the *trans* ring induces the *cis* ring to release its bound GroES, 7 ADP, and the "better folded" substrate.... Cycle begins again....



34

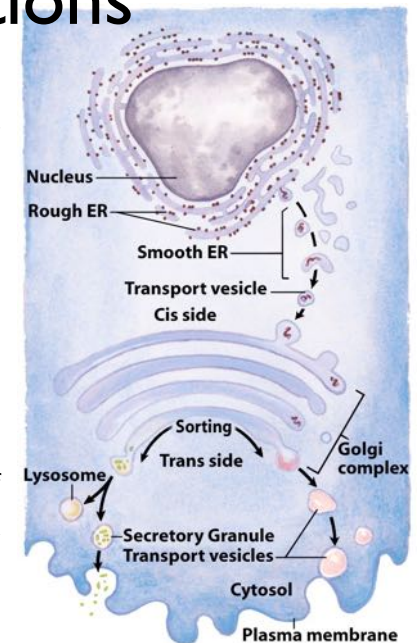
## Sorting in the ER and Golgi Apparatus

- Proteins that are initially directed to the ER must be further directed to their final destination.
- Many proteins are destined for secretion. There are two major pathways: constitutive and regulated.
- Other proteins are bound for cellular compartments: Golgi, lysosomes and other cellular compartments. Still others remain in the ER.
- Signals for these destinations is in the amino acid sequences of the proteins.

35

## Destinations

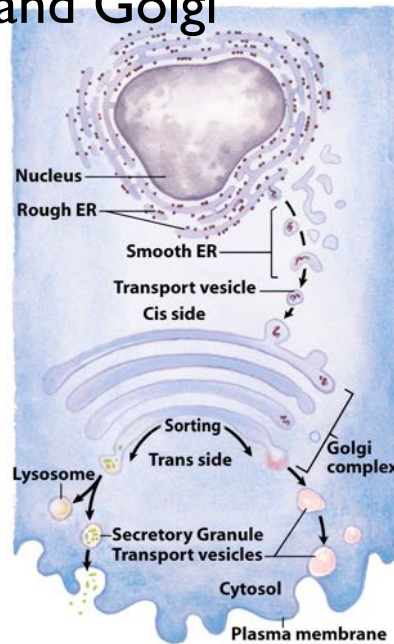
- Constitutive secretion** is believed the default pathway for proteins lacking any other sorting signal.
  - Proteins are incorporated into membrane-bound vesicles and transported to the Golgi apparatus.
  - The Golgi is a complex cellular organelle composed of a characteristic stack of 3-8 flattened membrane compartments. Proteins are believed to move between the compartments via vesicles.
  - Ultimately, proteins are transported in vesicles to the cell membrane for secretion.
- In **regulated secretion**, proteins follow a similar path through the Golgi, but instead of being immediately secreted, the proteins remain in the cell inside of storage vesicles. They remain there until an appropriate stimulus is received.



36

## Lysosomes, ER and Golgi

- **Lysosomes:** proteins destined for lysosomes also pass through the Golgi apparatus.
  - Proteins bound for lysosomes are glycosylated with mannose phosphate in the Golgi.
  - Proteins are then packaged into vesicles which migrate to lysosomes and fuse with them.
- **Endoplasmic Reticulum:** soluble proteins of the ER have similar C-terminal sequences (KDEL). It appears that a receptor recognizes the ER proteins in a compartment or vesicle after leaving the ER. These proteins are then returned to the ER.
- **Golgi Apparatus:** no clear signal has been identified yet. Rather than a contiguous peptide sequence, Golgi proteins may be identified based on properties or patches on the surfaces of folded proteins.



37

## Import into Other Organelles

- While mitochondria and chloroplasts have their own DNA, they only produce 10-20% of their proteins. The rest are synthesized in the cytosol of the cell.
- Importation of cytoplasmic proteins by mitochondria and chloroplasts involves N-terminal signal sequences that direct the proteins to the organelle. As in the ER, the signal sequence is usually proteolyzed after translocation across the membrane. (this is not the case with peroxisomes and glyoxysomes.)
- Translocation across the membrane involves denaturation of the protein.
- The protein refolds after translocation.
- Part of the energy required for translocation is provided by hydrolysis of ATP.

38

## Mitochondria

- **Mitochondria:** proteins destined for mitochondria tend to have N-terminal presequences that direct them to the mitochondria. (not all have the presequence)
  - The presequence is very different from the signal peptide used to target the ER.
  - Presequences ranging from 12-70 amino acid residues are highly variable.
- The N-terminal presequences of mitochondrial peptides interact with a receptor on the surface of the mitochondrion.
- Protein importation occurs at sites where the two membranes (inner and outer) are in close contact, and proteins are imported into the interior matrix. In most cases, the presequence is removed following translocation across the inner membrane.
- Proteins meant for the intermembrane space tend to have a double presequence, which is cleaved in two steps. The first presequence is cleaved in the matrix, the second is removed following translocation into the intermembrane space.
- Proteins intended for the outer membrane also have a double presequence, however in this case the N-terminal signal sequence directs the protein to the mitochondria. The second segment serves to anchor the protein to the membrane. (do not enter matrix)

39

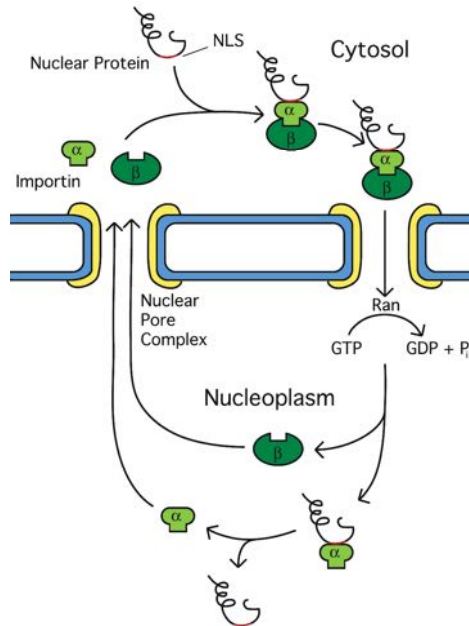
## Chloroplasts, Peroxisomes and Glyoxysomes

- **Chloroplasts:** importation of proteins is very similar to that of mitochondria. But there are significant differences.
  - Proteins have N-terminal presequences (transit peptides), which are cleaved following translocation.
  - Plant cells can distinguish between proteins intended for chloroplasts and those for mitochondria.
  - Protein importation involves ATP hydrolysis.
  - Proteins synthesized in the cytosol are localized in either the single membrane or the lumen of the thylakoid.
  - The proteases that remove the transit peptides are soluble proteins in the stroma of the chloroplasts.
  - Proteins bound for the lumen of the thylakoid must pass through three membranes. They are synthesized with double transit peptides. The N-terminal part is like other transit peptides and the second more closely resembles the signal peptide that targets the ER. The first part is removed in the thylakoid and the second then causes the intermediate protein cross the thylakoid membrane, and is then removed.
- **Peroxisomes and Glyoxysomes:** import all their proteins from the cytoplasm. Import occurs posttranslationally, however in many cases no N-terminal presequence is evident, and no proteolytic processing appears to occur.
  - A targeting sequence for peroxisomes has been identified. It is present at internal positions in the some peroxisomal enzymes.
  - Other peroxisomal and glyoxysomal proteins are synthesized with N-terminal tags, but these signal sequences are removed after translocation.

40

## Nuclear Proteins\*

- Proteins in the nucleus play important roles in DNA replication, maintenance, regulation and expression.
- All nuclear proteins are imported from the cytoplasm.
- Unlike other membrane-bound organelles, the nucleus is disrupted/disassembled during each cell cycle
- The nucleus is surrounded by a nuclear envelope (a double membrane with a perinuclear space).
- The envelope is penetrated by nuclear pore complexes, which form channels spanning the nuclear envelope.
  - A pore with a functional diameter of 90Å that can increase to a diameter of 250Å.
  - The pore complexes are composed of several polypeptide chains.



41 based on Fig 27-37 in Lehninger Principles of Biochemistry, 4th ed.

## Nuclear Proteins\*

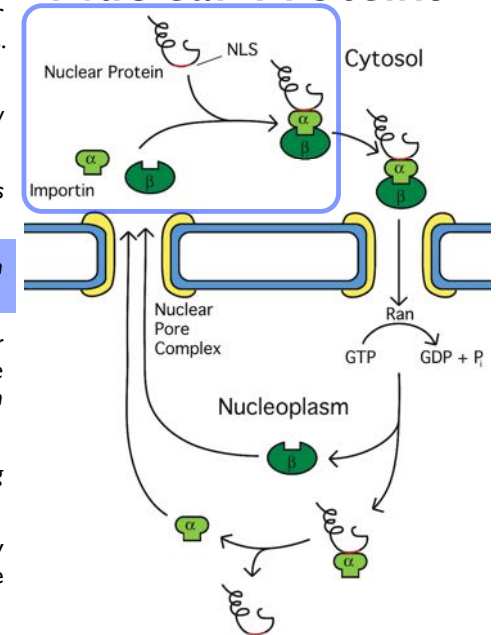
- Small proteins ( $\leq 20,000$  KD) could enter the nucleus readily through the pores. (some selectivity)
- Larger proteins are excluded unless they carry specific signals sequences (NLS).
- Importation is facilitated by proteins such as the importin  $\alpha/\beta$  complex (a heterodimer).

- The complex binds to nuclear proteins in the cytoplasm bearing NLS.

- The complex then binds to the nuclear pore and is translocated into the nucleoplasm with the aid of the Ran GTPase.

- The importin complex disassociates during the translocation process.

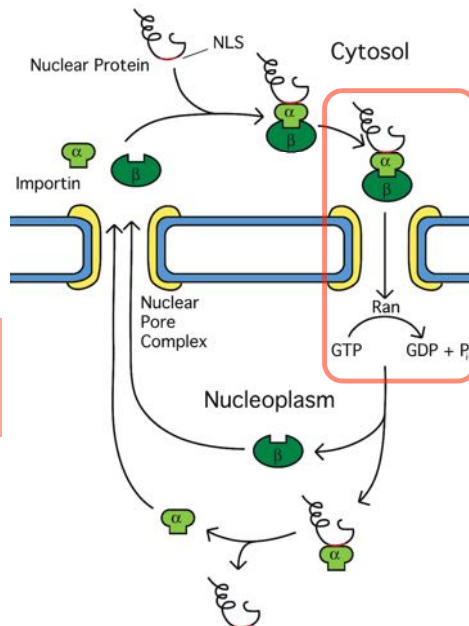
- The signal sequences can vary significantly from one protein to another, and are usually not removed.



42 based on Fig 27-37 in Lehninger Principles of Biochemistry, 4th ed.

## Nuclear Proteins\*

- Small proteins ( $\leq 20,000$  KD) could enter the nucleus readily through the pores. (some selectivity)
- Larger proteins are excluded unless they carry specific signals sequences.
- Importation is facilitated by proteins such as the importin  $\alpha/\beta$  complex (a heterodimer).
  - The complex binds to nuclear proteins in the cytoplasm bearing NLS.
  - The complex then binds to the nuclear pore and is translocated into the nucleoplasm with the aid of the Ran GTPase.
  - The importin complex disassociates during the translocation process.
- The signal sequences can vary significantly from one protein to another, and are usually not removed.



43 based on Fig 27-37 in Lehninger Principles of Biochemistry, 4th ed.

## Nuclear Proteins\*

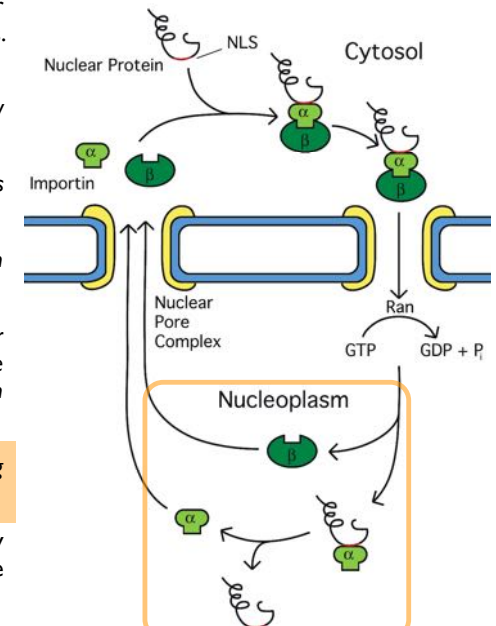
- Small proteins ( $\leq 20,000$  KD) could enter the nucleus readily through the pores. (some selectivity)
- Larger proteins are excluded unless they carry specific signals sequences.
- Importation is facilitated by proteins such as the importin  $\alpha/\beta$  complex (a heterodimer).

- The complex binds to nuclear proteins in the cytoplasm bearing NLS.

- The complex then binds to the nuclear pore and is translocated into the nucleoplasm with the aid of the Ran GTPase.

- The importin complex disassociates during the translocation process.

- The signal sequences can vary significantly from one protein to another, and are usually not removed.



44 based on Fig 27-37 in Lehninger Principles of Biochemistry, 4th ed.

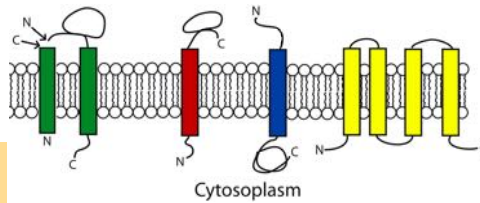
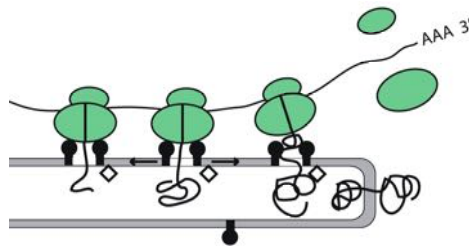
# Membrane Proteins

- Integral membrane proteins often undergo translocation across the membrane of the ER using the same type of presequence as used by soluble proteins.

- In the case of integral proteins, translocation of the growing polypeptide chain across the membrane is interrupted:

- Part of the protein lies on the cytoplasmic side of the membrane.
- One or more segments remain in the membrane.
- The remainder on the lumen side of the membrane.

- Translocation is thought to be halted by a stop-transfer sequence - a cluster of strongly hydrophobic residues flanked by basic residues.

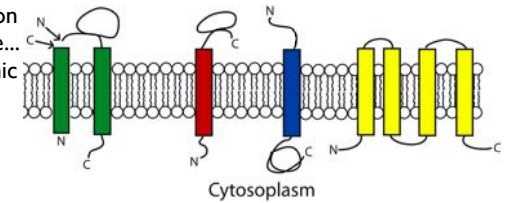
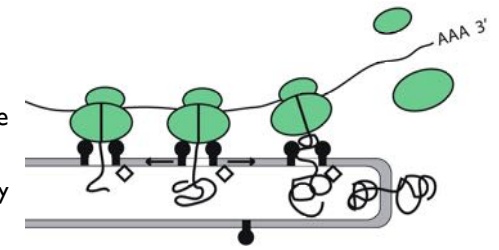


45

# Membrane Proteins

- Membrane proteins have diverse topologies:

- Some integral peptides are anchored by a single membrane-spanning segment.
- Other proteins may have multiple membrane-spanning segments.
- In some cases the N-terminus lies on the cytoplasmic side of the membrane... in others it lies on the extracytoplasmic side.



46

# Membrane Proteins

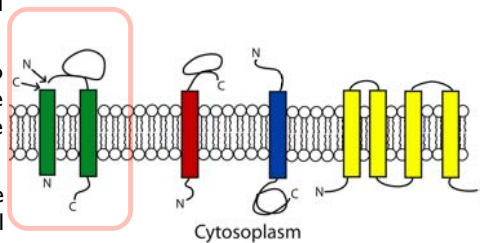
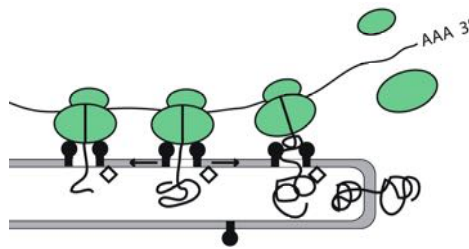
- Assembly of proteins with single transmembrane segments is better understood.

- Cleavage of an N-terminal signal segment results in the N-terminus of the transmembrane segment residing in the ER lumen. (internal stop-transfer segment halts translocation)

- In proteins with the opposite orientation, the N-terminal or internal signal sequences are not cleaved, and provide the transmembrane anchor.

- The N-terminal signal peptide can also serve as a stop-transfer sequence. The amino end is translocated across the membrane.

- Proteins with multiple transmembrane segments are thought to contain a signal sequence before each extracellular loop and a stop transfer after.



47

# Membrane Proteins

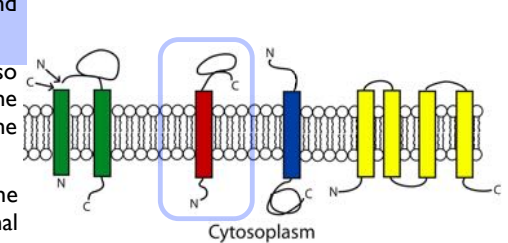
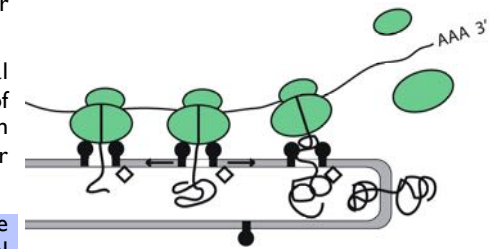
- Assembly of proteins with single transmembrane segments is better understood.

- Cleavage of an N-terminal signal segment results in the N-terminus of the transmembrane segment residing in the ER lumen. (internal stop-transfer segment halts translocation)

- In proteins with the opposite orientation, the N-terminal or internal signal sequences are not cleaved, and provide the transmembrane anchor.

- The N-terminal signal peptide can also serve as a stop-transfer sequence. The amino end is translocated across the membrane.

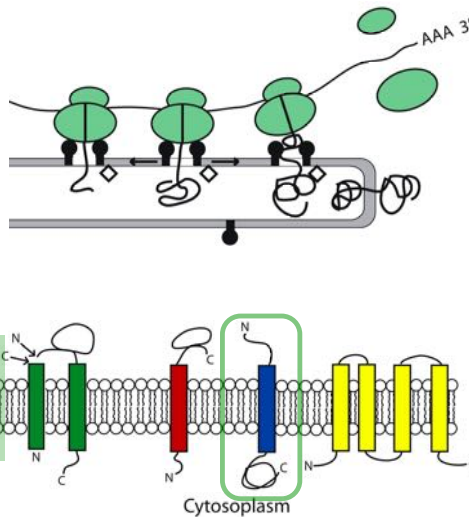
- Proteins with multiple transmembrane segments are thought to contain a signal sequence before each extracellular loop and a stop transfer after.



48

# Membrane Proteins

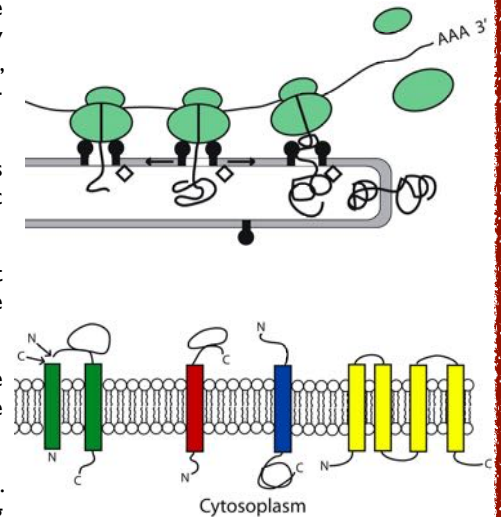
- Assembly of proteins with single transmembrane segments is better understood.
  - Cleavage of an N-terminal signal segment results in the N-terminus of the transmembrane segment residing in the ER lumen. (internal stop-transfer segment halts translocation)
  - In proteins with the opposite orientation, the N-terminal or internal signal sequences are not cleaved, and provide the transmembrane anchor.
  - The N-terminal signal peptide can also serve as a stop-transfer sequence. The amino end is translocated across the membrane.
- Proteins with multiple transmembrane segments are thought to contain a signal sequence before each extracellular loop and a stop transfer after.



49

# Membrane Proteins

- proteins with more than one transmembrane segment are likely threaded through the membrane, employing multiple start- and stop-transfer sequences.
- Eukaryotes: these transfer sequences characterized by groups of hydrophobic residues flanked by basic residues.
- Transfer sequences promote and arrest transfer of residues through the translocon.
- mechanism by which transmembrane proteins exit the translocon to reach the surrounding bilayer is not clear.
- likely exit as single or pairs of helices... with ultimate assembly and helix packing taking place in the bilayer.



50

I. in cases where the N-terminal signal peptides is cleaved -- have new N-terminus in the extracytoplasmic side of membrane. *Original N-terminus likely remains on cytoplasmic face.*

- Internal stop-transfer sequence halts translocation and remaining peptide remains on cytoplasmic face.

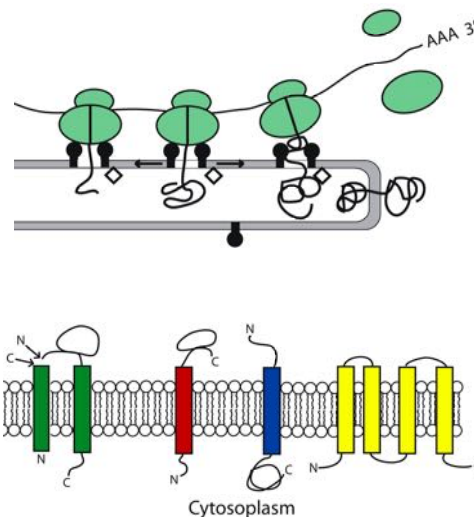
II. Proteins with opposite orientation: the N-terminal or internal signal peptide are not cleaved--signal peptide in such cases tend to be more hydrophobic than other signal sequences.

III. single-span proteins use N-terminal signal sequence, but the N-terminus has been relocated to the other side of the membrane.

polarity of segment influence orientation and which end is translocated across the membrane... basic residues at one end tend to keep that end on the cytoplasmic side.

IV. in cases of multiple membrane spanning segments, thought to contain signal sequence before each extracellular loop and a stop transfer after it.

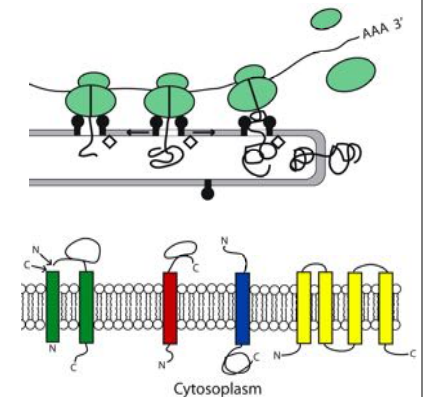
# Membrane Proteins



51

# Membrane Proteins

- In prokaryotes...
  - NO ER... only targeting is for secretion and for membrane proteins.
  - Bacteria have inner and outer membranes.
  - Proteins normally secreted across inner, plasma, membrane... outer membrane not generally permeable to proteins.
  - export of proteins from cytoplasm to periplasm & outer membrane of gram negative bacteria resembles eukaryotic targeting to ER,
  - Involves signal sequence recognized by prok. equivalent of SRP.
  - prok. protein translocation driven by electrochemical potential.
  - there are translocation mechanisms in prok's that are independent of SRP
  - Completed polypeptide chains can be translocated in Prok system more efficiently than in Euk systems.



in prok-release from membrane coincides with folding.

52

# Posttranslational Modification

# Posttranslational Modification

- Assembly of the polypeptide sequence by the ribosome is often not the final step in the formation of a functional protein.
- One or more residues may be removed, added or altered.
- Portions of the polypeptide sequence may be removed through enzymatic proteolysis.
- The polypeptide termini and the side chains of many amino acids can be covalently modified by the addition of various functional groups.
- Posttranslational modifications can serve a vast range of purposes: activating/deactivating proteins, targeting them for specific locations, adding new chemical functionality and targeting the protein for destruction. (to list a few).

54

# Posttranslational Modification

- Proteolytic Processing
- Alteration of Chain Termini
- Side Chain Modification
- Targeting Proteins for Destruction

55

# Proteolytic Processing

- Proteolytic processing of polypeptide chains is relatively common for proteins destined for cellular organelles or for secretion.
- The precursor protein is referred to as the “**pro**” protein.
- Precursor proteins that also bear an N-terminal signal peptide are referred to as the “**prepro**” protein.
- Most proproteins are biologically inactive and proteolysis yields the active protein, providing a means of regulating activity.
- The proteases responsible for cleavage can be very specific for a target sequence.
- Proteolytic processing usually occurs with a folded precursor protein, rather than one that is denatured.
- Whether the target is a signaling protein, an enzyme or involved in blood coagulation, proteolysis provides an excellent trigger for activation. Can initiate a cascade of events.

56

# Proteolytic Processing

- Precursors of proteolytic enzymes are usually inactive.
- Proteolytic enzymes of the digestive tract such as trypsin, chymotrypsin and pepsin are usually produced in inactive forms (zymogens: trypsinogen, chymotrypsinogen and pepsinogen).
- These enzymes are secreted by the pancreas and are activated in the digestive tract.
- This strategy allows cells to accumulate and store the zymogens for secretion when needed.
- The activation of some proteases such as chymotrypsin require the hydrolysis of multiple bonds (3 for chymotrypsin).
- Proteolysis sites are usually determined based on sequence and the folded conformation.

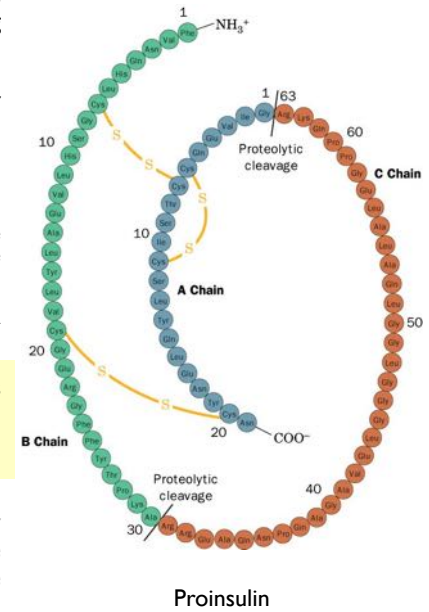
57

# Proteolytic Processing

- Most peptide hormones such as insulin are produced as larger precursor peptides requiring proteolysis for activation and secretion.
- Producing these peptides as part of larger precursor peptides may serve several purposes.
  - Smaller peptides may present problems for ribosomal assembly and translocation to the ER.
  - Many precursors contain several peptide hormones. The incorporated peptide segments may or may not be closely related in sequence and function.
  - Different proteolysis patterns may yield different spectra of active peptides.

- Proinsulin is stored in secretory granules, and is activated by cleavage at two sites. Releasing an internal peptide segment (C-peptide).

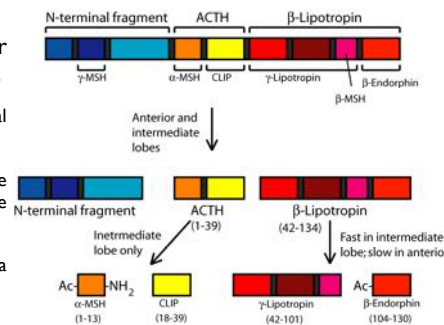
- Proteolytic processing can be very complex and orchestrated. Such is the case for proopiomelanocortin, which contains the sequences for no fewer than 8 peptide hormones.



58

# Proteolytic Processing

- Most peptide hormones such as insulin are produced as larger precursor peptides requiring proteolysis for activation and secretion.
- Producing these peptides as part of larger precursor peptides may serve several purposes.
  - Smaller peptides may present problems for ribosomal assembly and translocation to the ER.
  - Many precursors contain several peptide hormones. The incorporated peptide segments may or may not be closely related in sequence and function.
  - Different proteolysis patterns may yield different spectra of active peptides.
- Proinsulin is stored in secretory granules, and is activated by cleavage at two sites. Releasing an internal peptide segment (C-peptide).



59

# Alterations to Chain Termini

## N-terminal Met:

- The formyl group on the initiating Met is usually removed by a deformylase enzyme.
- A ribosome-associated Met-aminopeptidase removes the Met group on about half the proteins translated in both prokaryotes and eukaryotes. (sequence dependent, and system can be overloaded with high expression)

60

# Alterations to Chain Termini

## Acetylation of N-Terminus

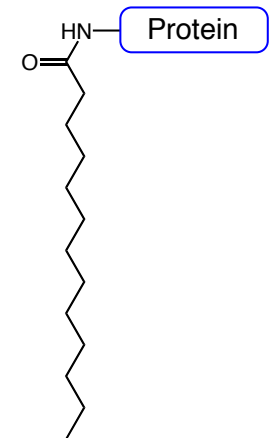
- A number of N<sup>α</sup>-acetyltransferases are believed to catalyze the transfer of an acetyl group from acetyl-CoA to the terminal amino group of a protein. (50-90% of isolated eukaryotic cytoplasmic proteins have their N-termini acetylated.)
- N-terminal acetylation not observed in proteins expressed in mitochondria or chloroplasts.
- Acetylation is somewhat dependent on the nature of the N-terminal amino acid, and can occur during translation of cytoplasmic proteins or can occur postranslationally.
- Although acetylation is the most common N-terminal covalent modification, other modifications have been observed: formyl, pyruvoyl, fatty acyl, glucuronyl, methyl..... etc.
- The state of the amino terminus of a peptide may have a large impact on the rate of its degradation.

61

# Alterations to Chain Termini

## Myristoylation of N-terminus

- A number of cytoplasmic proteins are found with myristic acid attached to their N-terminal amino group.
- Attachment of the myristoyl group occurs either during or shortly after translation. It is transferred to the N-terminal amino group from myristoyl-CoA by the enzyme N-myristoyl transferase.
- It is believed that the attached myristoyl group causes proteins to be loosely associated with a membrane and may help proteins juxtaposition near particular membranes.



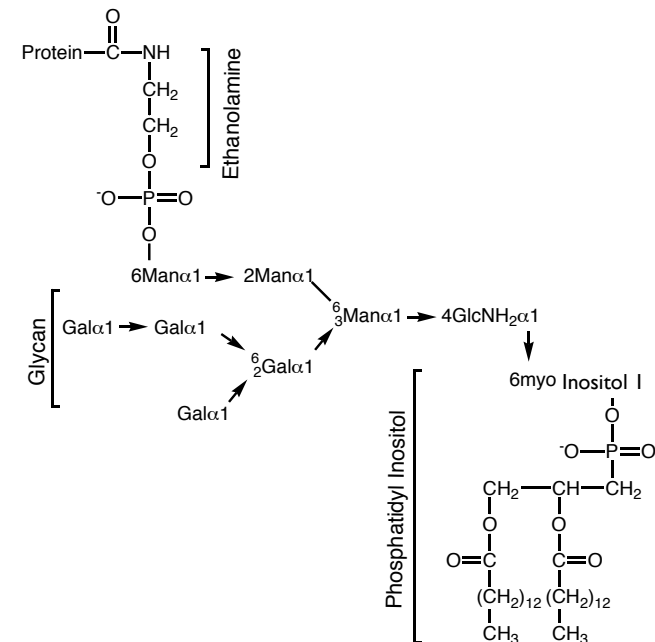
62

# Alterations to Chain Termini

## C-Terminal Glycosyl-Phosphatidylinositol and Farnesyl Membrane Anchors

- Some proteins on the cell surface are anchored to the cell membrane through complex glycosyl-phosphatidylinositols covalently linked to the protein C-terminal carboxyl group through an ethanolamine group.
- The complex glycan and inositol structure is believed to be preassembled prior to attachment to the protein.
- Proteins bearing these groups are synthesized with signal peptides and are directed to the ER. The proteins have a hydrophobic tail at the C-terminal segment of the sequence, which is cleaved from the protein at about the same time that the glycosyl-phosphatidylinositol group is added.
- The glycosyl-phosphatidylinositol group appears to serve as an anchor for otherwise soluble proteins. It permits the protein to diffuse within the membrane more rapidly than a protein anchor would allow.

63

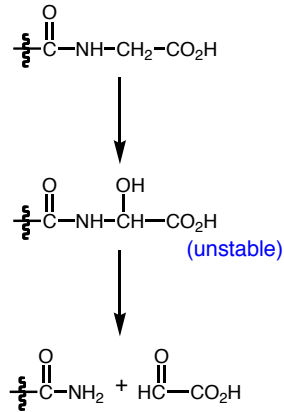


64



# Amidation of the C-terminus

- A C-terminal amide group in place of the carboxyl group is a characteristic feature of many peptide hormones.
- Amidation of the C-terminus may enhance stability of the peptide.
- The amide group is derived from a Gly residue present at the C-terminus.
- The Gly residue is converted to an amide via two reactions catalyzed by two enzymes. (process involves ascorbic acid, copper and oxygen)
- Occurs in secretion granules.
- The Gly residue becomes accessible following proteolytic processing of a parent polypeptide.



65

# Glycosylation

- Attachment of carbohydrate groups is one of the most common posttranslational modifications of eukaryotic proteins. (particularly true for proteins bound for secretion)
- Does not occur in prokaryotes.
- Carbohydrate groups are highly variable and their function appears to be protein specific.
  - Can be involved in activity.
  - Can enhance stability and lifetime.
  - Can play a role in trafficking.

66

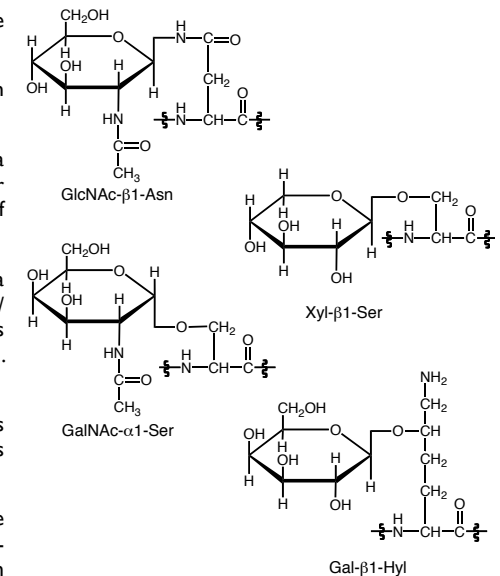
# Glycosylation

- The variability in carbohydrate structure allows them to be tailored to specific functions and interactions.
- The carbohydrate groups can be assembled from at least 8 different sugar monomers: galactose, glucose, fucose, mannose, N-acetylgalactosamine, N-acetylglucosamine, sialic acid and xylose. These sugar monomers can be linked together by a variety of glycosidic linkages and can be highly branched.
- Carbohydrate groups are very hydrophilic.
- They can be large and can significantly impact the surface properties of a protein.
- Two types of glycosylation: N-glycosylation and O-glycosylation.

67

# N-Glycosylation

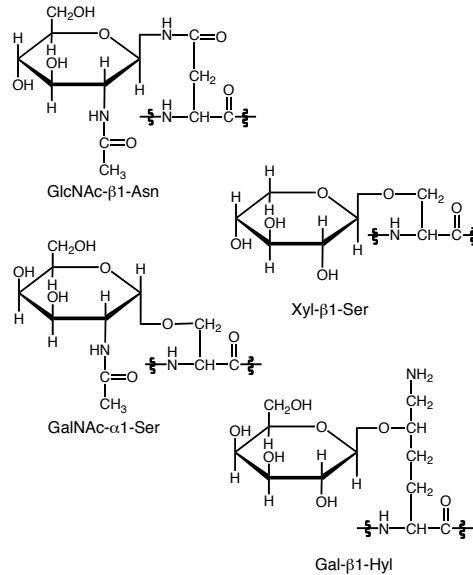
- N-Glycosylation occurs at Asn residues. (side chain)
- N-Glycosylation occurs cotranslationally in the ER.
- Process begins with the attachment of a preassembled core structure. Occurs near the membrane surface on the lumen side of the ER membrane.
- The Asn to be glycosylated lies within a characteristic sequence: -Asn-Xaa-Ser/Thr/Cys-. Xaa can be any residue except Pro. This sequence is not the sole determinant... however additional factors are not obvious.
- After attachment of the core glycan, it is extensively modified as the protein passes through the ER and Golgi.
- The type of processing depends on the protein, cell type and physiological state. N-glycosylated proteins can be heterogeneous in the composition of their attached carbohydrate groups.



68

# O-Glycosylation

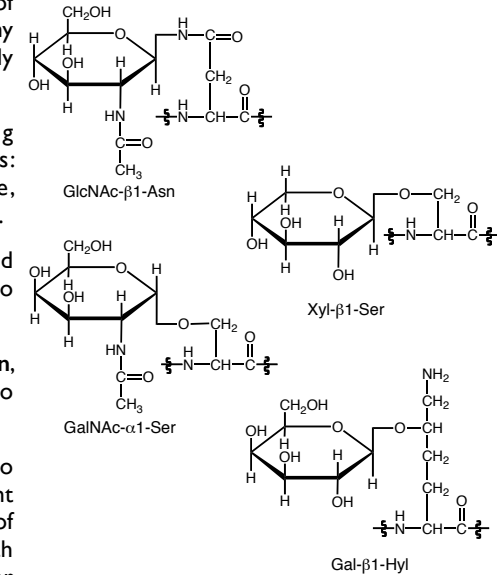
- O-Glycosylation occurs at **side chain hydroxyl groups (Ser and Thr residues)**, and others such as **Hyp and Hyl** residues in collagen).
- O-Glycosylation **occurs primarily posttranslationally** in the Golgi with proteins that are already folded.
- Signals for which Ser and Thr residue is to be glycosylated are not clear. The three dimensional configuration is probably important.
- The carbohydrate group may comprise 65-85% of the mass of the glycosylated protein. Therefore, can dominate the structure.
- O-glycosylation has been observed in cytoplasmic and nuclear proteins. In these cases N-acetylglucosamine groups are attached to Ser and Thr side chains.



69

# Proteoglycan

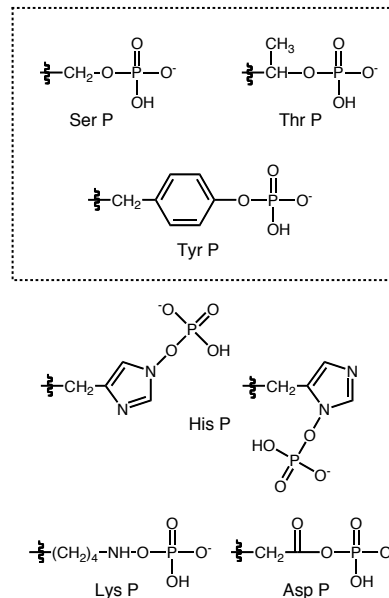
- Proteoglycans are composed of a variety of protein backbones to which one or many glycosaminoglycan chains are covalently attached.
- Glycosaminoglycans are repeating disaccharide chains of three types: chondroitin sulfate/dermatan sulfate, heparan sulfate/heparin or keratan sulfate.
- Tend to be sulfated to varying degrees and are usually attached via a xylose group to the -OH of Ser residues.
- Represent an **extreme in glycosylation**, where the bulk of their structure is due to the carbohydrate.
- The core proteins have been difficult to characterize. They vary in molecular weight and degree of glycosylation. Domains of core proteins may be involved with interactions with membranes, extracellular components and other molecules.



70

# Phosphorylation

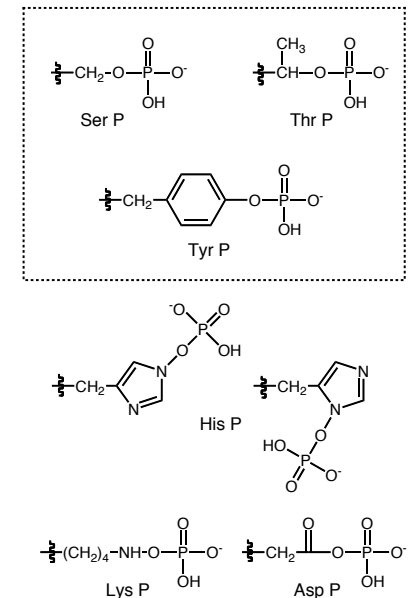
- Protein phosphorylation is very specific and can have dramatic impact on protein function.
- Phosphorylation is reversible.**
- Structurally, the most important aspect of a phosphoryl group is the negative charge.
- Phosphoryl groups are added by specific protein kinases, using ATP as the phosphoryl group donor.
- Similarly, phosphoryl groups are selectively removed by specific phosphatases.



71

# Phosphorylation

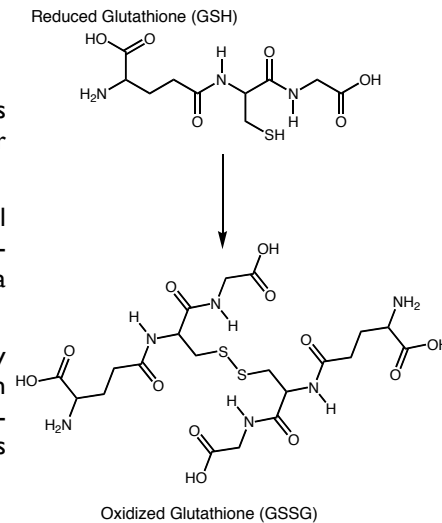
- This process of controlled phosphorylation and dephosphorylation provides a means of regulating the activity of the phosphorylated protein.
- The **main sites for phosphorylation are the -OH groups of Ser, Thr and Tyr**, but the side chains of His, Lys and Asp residues can also be phosphorylated.
- Kinases are selective for both sequence and structure. Therefore, phosphorylation almost invariably occurs with folded protein, and primarily in the cytoplasm.



72

# Disulfide Formation

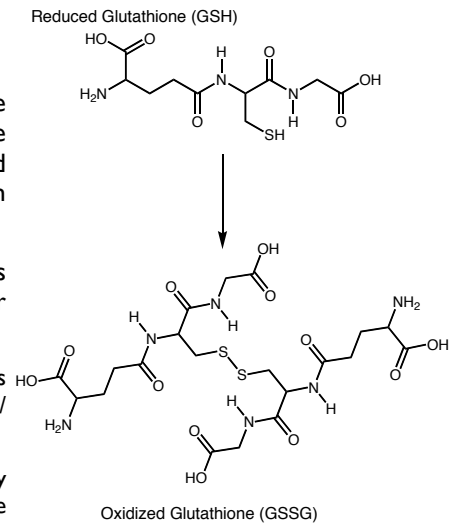
- Disulfide bond formation between Cys residues is a common occurrence for proteins synthesized in the ER.
- Such disulfide bonds often play critical roles in defining and stabilizing the three-dimensional folded conformation of a protein.
- Disulfide formation in the ER likely involves thiol-disulfide exchange between the protein being synthesized and small-molecule disulfide compounds such as glutathione.



73

# Disulfide Formation

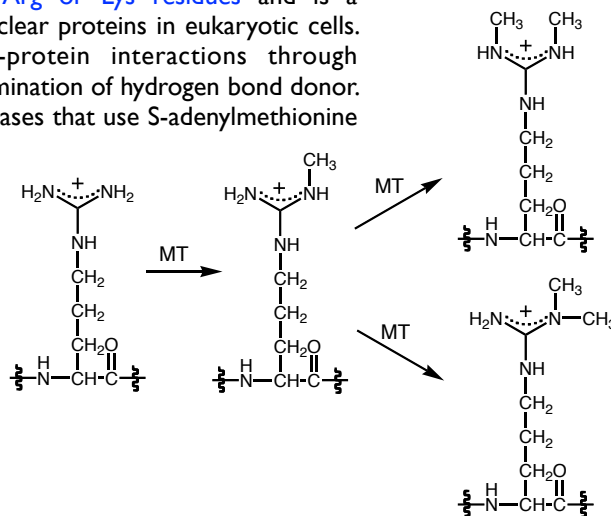
- Formation of a disulfide bond in the course of protein folding would require two such exchanges between oxidized glutathione molecules and Cys residues in the protein.
- In proteins with more than two Cys residues or multiple intramolecular disulfide bonds:
  - Initial formation of disulfide bonds is followed by disulfide exchange/rearrangement.
  - The process is aided by the folding auxiliary protein protein-disulfide-isomerase in the ER.



74

# Methylation and Side Chain Acetylation

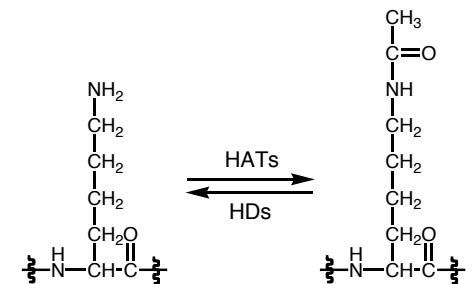
- **Methylation:** occurs at Arg or Lys residues and is a common modification to nuclear proteins in eukaryotic cells. It likely impacts protein-protein interactions through increased steric bulk and elimination of hydrogen bond donor. Catalyzed by methyl transferases that use S-adenylmethionine as a methyl donor.
- Methylation appears to be irreversible. Therefore, regulation must occur via regulation of methylation itself.
- Methylation of histone Lys residues is involved with regulating gene expression, DNA replication and repair.



75

# Methylation and Side Chain Acetylation

- **Acetylation:** the side chain amino groups of Lys residues can be catalyzed by sequence-specific N-acetyltransferases.
- Unlike acetylation of the amino terminus, side chain acetyl groups can be removed by deacetylases.
- Like methylation, acetylation of histone Lys residues plays a major role in the control of gene expression and chromosomal functions. Histone acetylation is associated with chromatin in an active state.



76

# Protein Degradation

77

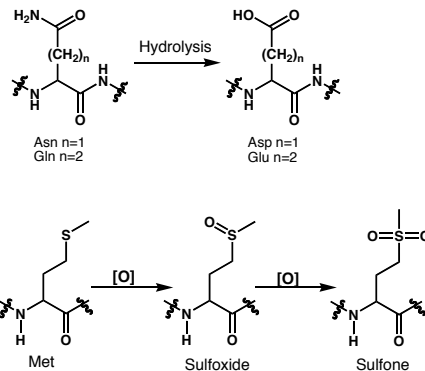
# Protein Degradation

- Proteins are not immortal nor are they immutable.
- Over time, proteins can be subjected to gradual covalent modification as a result of non-enzymatic and enzyme-mediated chemical reactions.
- Cellular proteins are in a state of flux, they are being synthesized, and they are being broken down. The relative rates of these processes determines the amount of protein present.
- Degradation provides a means of eliminating proteins that have become damaged, denatured or misfolded as well as abnormal proteins (errors in biosynthesis).
- Degradation is also used to regulate the levels of various proteins .
- The various protein degradation pathways are highly regulated, and are compartmentalized within organelles (lysosomes) and macromolecular structures (proteasomes).

78

# Chemical Aging

- Proteins age and are susceptible to chemical change and modification (wear and tear).
- Oxidation of sulfur atoms of Cys and Met residues.
- Deamidation of side chains of Asn and Gln residues.
- Racemization of C<sub>α</sub> and epimerization of side chain chiral centers (Thr and Ile).
  - ◆ Asn and Asp residues particularly susceptible to racemization.
  - ◆ **Most proteins are degraded before racemization becomes an issue.**
- Disruption of disulfide bonds.
- Hydrolysis of peptide bonds (especially at Asp residues).
- Enzymes are present that repair some of this damage/modifications (especially Cys and Met oxidation).
- Maintain intracellular glutathione in reduced form.

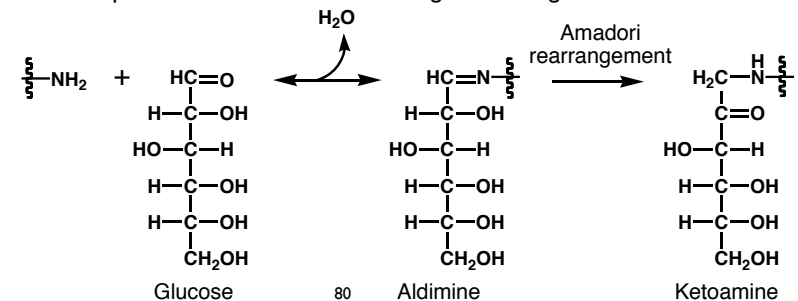


Peptide bonds adjacent to Asp, Ser and Thr are particularly susceptible to hydrolysis.

79

# Chemical Aging

- Proteins are covalently modified by many chemicals present in the environment (many of which are man-made).
- Non-enzymatic reaction between primary amino groups and reducing sugars (such as glucose) results in formation of a Schiff base.
  - Schiff base then rearranges to form a stable ketoamine.
  - Dehydration of ketoamines results in yellow-brown fluorescent products and cross-links between protein molecules.
  - Particular problem when concentration of glucose is high.



80

# Chemical Aging

- Chemical modification can significantly impact protein structure and function.
- Modifications that compromise structure, stability or function usually result in quick degradation of the affected protein.
- Alterations that do not trigger the degradation of a protein tend to accumulate throughout the lifetime of the protein.
- Most proteins turn over at rates that are determined by enzymatic processes.
- The only type of chemical modification that appears to be the primary determinant of the rate of degradation is the oxidation of bound metal centers, which results in the destruction of coordinating side chains.

81

# Factors that Determine the Rate of Degradation

- Originally thought that the rate of degradation of a protein was related to:
  - susceptibility to thermal denaturation
  - absence of stabilizing ligands
  - intrinsic rate of deamidation of Gln and Asn residues
  - susceptibility of Cys, His and Met residues to oxidation
  - glycosylation or phosphorylation
  - presence of free  $\alpha$ -amino group
  - net negative charge
  - flexibility of unfolded state.
- These make sense if susceptibility to enzyme proteolysis were the limiting factor.
- Too limiting... it assumes the same degradation mechanism is used in all cases, but this is likely not the case.

82

# Factors that Determine the Rate of Degradation

- The variety of degradation mechanisms and rates suggest more specific signals are involved.
- Rapidly degraded proteins contain one or more 16-20 residue segments that are rich in Pro, Glu, Ser and Thr residues (PEST regions).
- The N-terminal amino acid can determine the rate of degradation in the ubiquitin-dependent system.
- Folded conformation is important in minimizing rate of degradation. Abnormalities such as chemical modifications, mutations, chain terminations or incorporation of amino acid analogs usually result in rapid degradation through enzymatic proteolysis.

N-terminal residue	Half-life
<b>Stabilizing</b>	
Met, Gly, Ala, Ser, Thr Val	>20 h
<b>Destabilizing</b>	
Ile, Gln	~30 min
Tyr, Glu	~10 min
Pro	~7 min
Leu, Phe, Asp, Lys	~3 min
Arg	~2 min

83

# Proteases Involved in Protein Turnover

- A variety of proteases of varying specificities appear to be involved in the degradation of certain proteins.
- Many intracellular proteases such as the  $\text{Ca}^{2+}$ -dependent calpains appear to be involved with the regulatory cleavage of specific enzymes and not protein degradation.
- The most interesting protein degradation is that which occurs within the cell, whereas most of the well-characterized proteases normally function extracellularly.
- Intracellular proteases tend to be larger than comparable extracellular proteases, with additional domains that function primarily in regulatory capacities.
- While hydrolysis of peptide bonds does not require energy input, the process of protein degradation within the cell usually does. The energy (in the form of ATP) is required for initiation of proteolytic degradation.

84

# Proteases Involved in Protein Turnover

- Several different ATP-dependent proteases have been discovered in a variety of organisms.
- Some of the isolated ATP-dependent proteases are thiol proteases, and are activated by binding of ATP.
- The other group of ATP-dependent proteases are serine proteases that require ATP hydrolysis for proteolytic activity. They also require bound  $Mg^{2+}$  ions. ATPase activity is stimulated by the presence of protein/peptide substrates for degradation.
- Only two proteases, *La* and *Ti* (out of >20 known proteases) in *E. coli* are known to require ATP for enzyme activity.

85

# Proteases Involved in Protein Turnover

- Only two proteases, *La* and *Ti* (out of >20 known proteases) in *E. coli* are known to require ATP for enzyme activity.
- *La* is a tetrameric serine protease, assembled from four identical polypeptide chains (87KD each).
  - Production of *La* is stimulated by environmental stress such as heat shock. It appears to degrade abnormal and denatured proteins that can result under stressful conditions (appears to degrade short lived regulatory proteins as well).
  - *La* requires ATP hydrolysis for proteolytic activity. ATPase activity is stimulated by the presence of protein substrates.
  - ATP appears to play a regulatory role and does not contribute directly to the proteolytic mechanism.
  - Bound ATP is hydrolyzed following proteolysis of a peptide bond in the substrate. (rate limiting step appears to be release of ADP)
  - As a homotetramer, the complex incorporates 4 active sites. All four appear to be under allosteric control.

86

# Proteases Involved in Protein Turnover

- Only two proteases, *La* and *Ti* (out of >20 known proteases) in *E. coli* are known to require ATP for enzyme activity.
- *Ti* is a complex of two polypeptide chains. One chain provides proteolytic site the other the ATPase site. (the isolated proteolytic chain demonstrates ATP independent proteolytic activity in the absence of the ATPase chain).
  - ATP hydrolysis by the ATPase chain activates the proteolytic chain.
  - Like *La*, *Ti* is a heat-shock protein that is produced to degrade damaged proteins.

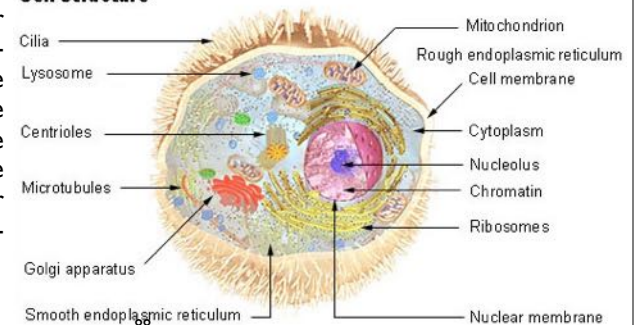
87

# Lysosomes

- Lysosomes are intracellular compartments involved with protein degradation. They contain a variety of proteases that require acidic conditions for enzyme activity (collectively known as cathepsins).
- Some intracellular protein degradation takes place in the lysosomes.
- Cytoplasmic proteins are believed to enter lysosomes when a small volume of cytoplasm is sequestered by cell membrane, which then fuses with the lysosome. (believed to be mainly in response to starvation conditions)

- Alternatively, intracellular proteins may be introduced into the lysosome when invaginations in the lysosome membrane occurs. (believed to be the normal pathway for the degradation of long-lived proteins)

Cell Structure

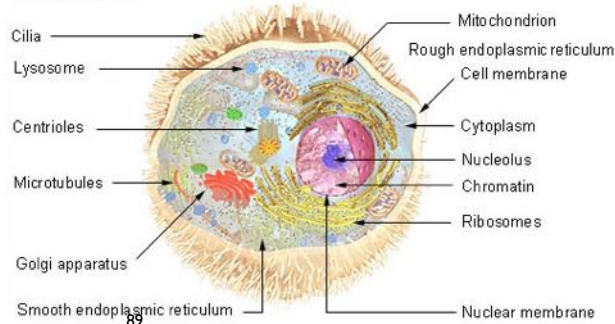


88

# Lysosomes

- It has been proposed that a third pathway exists that is activated under specific conditions. This pathway involves the selective uptake of proteins and peptides based on a specific peptide motif.
- Lysosomes are involved with the uptake and digestion of exogenous proteins, which are taken up from the extracellular medium through endocytosis.

## Cell Structure



# Ubiquitination

- In eukaryotes, proteins destined for degradation are tagged by the addition of many molecules of ubiquitin.
- Ubiquitination is carried out via a multi-enzyme pathway (three enzymes/enzyme complexes E<sub>1</sub>, E<sub>2</sub> and E<sub>3</sub>).
  1. Ubiquitin-activating enzyme (E<sub>1</sub>) is a 105-kD dimer.
  2. Initially, the  $\alpha$ -carboxyl group of the C-terminal Gly residue of ubiquitin is activated by E<sub>1</sub> with formation of an adenylated ubiquitin intermediate (using ATP).
  3. The activated  $\alpha$ -carboxyl group of ubiquitin is then transferred to the thiol side chain of a Cys residue on the ubiquitin-carrier protein E<sub>2</sub> (family of at least 7 different proteins).
  4. E<sub>2</sub> transfers the ubiquitin to the  $\alpha$ - and  $\epsilon$ -amino groups of target proteins.



90

# Ubiquitination

- E<sub>2</sub> transfers ubiquitin to free primary amino groups on proteins that have been selected/bound by ubiquitinating-protein ligase (E<sub>3</sub>).
- Like E<sub>2</sub>, there are a variety of E<sub>3</sub> proteins (each is ~180-kD).
- Once one ubiquitin molecule has been added, additional ubiquitin molecules are added to the attached ubiquitin molecule(s) via their  $\epsilon$ -amino groups, resulting in a ubiquitin chain.
- E<sub>3</sub> plays central role in targeting proteins for degradation. Selects proteins based on nature of N-terminal amino acid:
  - Protein must have free N-terminus.
  - Proteins with N-terminal Met, Ser, Ala, Thr, Val, Gly or Cys are resistant to ubiquitin-mediated degradation.
  - Those with N-terminal Arg, Lys, His, Tyr, Phe, Trp, Leu, Asn, Gln, Asp or Glu are targeted and have half lives of 2-30 minutes.
  - Arg is added at N-terminus of proteins with N-terminal Asp or Glu residues prior to degradation.



91

# Ubiquitination

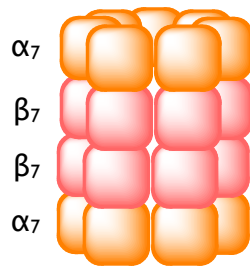
- Most proteins susceptible to degradation through N-terminal degradation are normally secreted proteins, not normally found intracellularly.
- PEST sequence also targets proteins for ubiquitin-mediated degradation.
- Unfolded, partially-unfolded and damaged proteins are also targets for ubiquitination.



92

# Proteasome

- Special case of ATP-dependent protease.
- Proteasomes are large oligomeric structures, with a large central cavity which is site of proteolysis.
- Eukaryotic cells contain two forms of proteasomes: 20S proteasome and 26S proteasome.
  - 20S proteasome is a 700-kD barrel-shaped structure assembled from two different types of polypeptide chain ( $\alpha$  and  $\beta$ ).
  - Assemble into four stacked heptameric rings:  $\alpha_7\beta_7\beta_7\alpha_7$  organization.
  - Proteolytic sites are located within cavity that is 15nm long and 11nm in diameter.
  - Access to cavity mediated by outer  $\alpha_7$  rings (also believed to unfold target proteins).
  - subunits possess proteolytic site (N-terminal Thr residues functions much like Ser in serine proteases.)
  - Targets broken down into oligopeptides ranging from 7-9 residues in length.



Note: ubiquitin not observed in prokaryotes.

# Proteasome

- Eukaryotic cells contain two forms of proteasomes: 20S proteasome and 26S proteasome.
  - Full 26S proteasome is on the order of 2.5 megadaltons, and is 45 nm in length.
  - 26S proteasome incorporates a 20S proteasome plus two additional structures known as 19S regulators (a.k.a. 19S caps or PA700).
  - Unlike archaeal 20S proteasomes, each  $\alpha_7$  ring is assembled from 7 different  $\alpha$  subunits, and each  $\beta_7$  ring composed of 7 different  $\beta$  subunits.
  - Only three of the 7  $\beta$  subunits in each ring have protease active sites.
  - Many of the 19S regulator subunits have ATPase activity, and play a role in substrate specificity.
  - The 19S regulator units play a role in recognizing and selecting ubiquitinated proteins for degradation.
  - The 19S regulator units also carry out unfolding and transport of ubiquitinated proteins into the proteolytic central cavity.

