Lecture 7

Post-Translational Processing

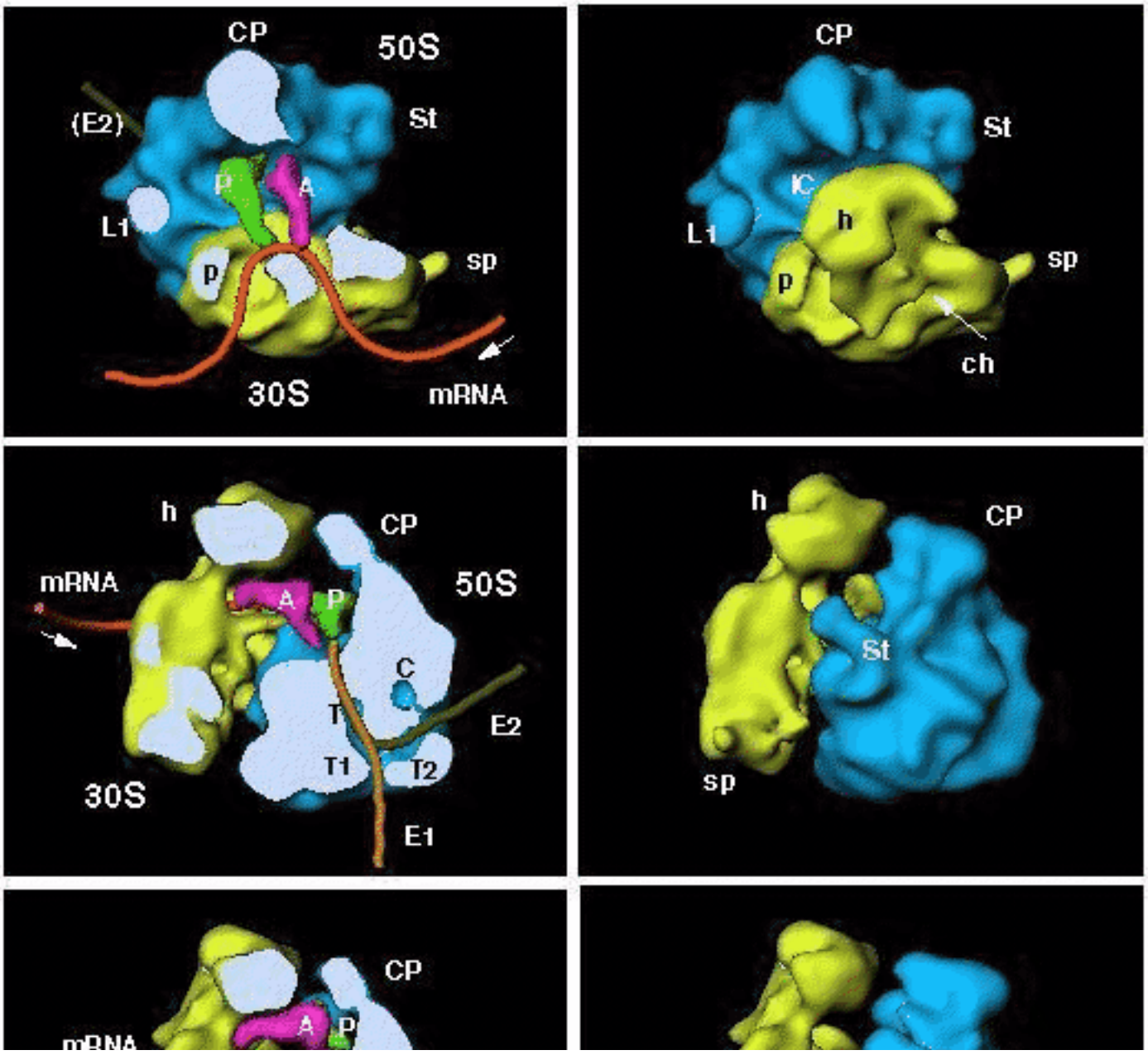
Protein synthesis consists of several steps: from the translation of the information from mRNA to the folded and fully processed, active protein in its proper compartment of action.

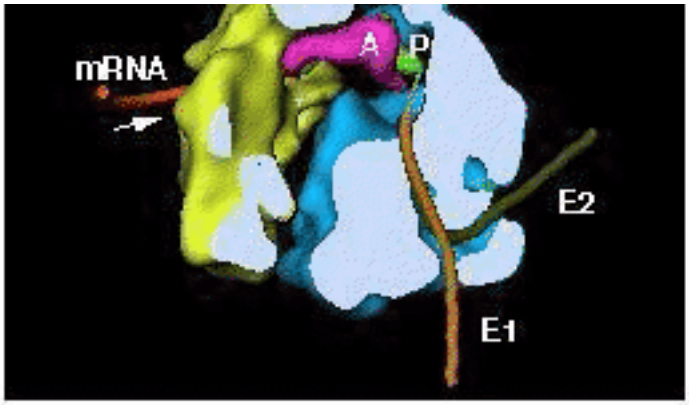
The mRNA sequence predicts a specific length polypeptide chain made up of the primary 20 amino acids.

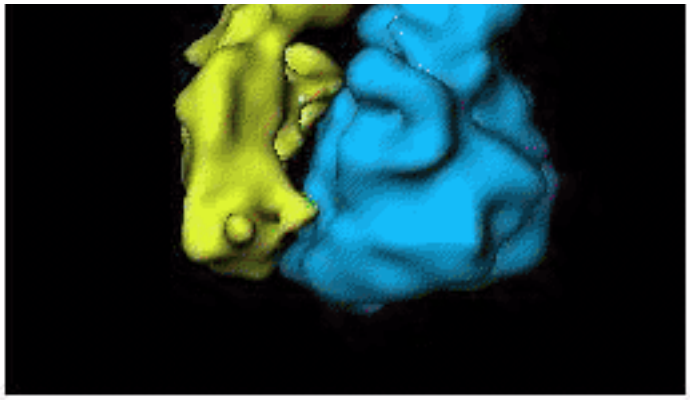
Fully processed protein products are almost always shorter than their mRNA would predict, and globally contain about 200 different amino acids.

(determined by sequencing, biochemistry, X-ray crystallography)

During translation, about 30-40 polypeptide residues are relatively protected by the ribosome (tunnel T and exit sites E1 and E2 in the large subunit). Once the polypeptide chain emerges from the ribosome it starts to fold and can be subject to post-translational modifications.







So after translation several additional steps must be considered as part of the complete protein biosynthetic process:

1. Covalent modification of

    a: peptide bonds
    b: the N-terminus
    c: the C-terminus
    d: amino acid residues (side chains).

2. Noncovalent modifications: folding, addition of co-factors.

3. Translocation: compartment selection and transport (Trafficking/Targeting).

4. Involvement of molecular chaperones in 1, 2, and 3.

Why post-translational processing?

- adds functionality

- effects targeting
- regulates activity
- increases mechanical strength
- changes recognition

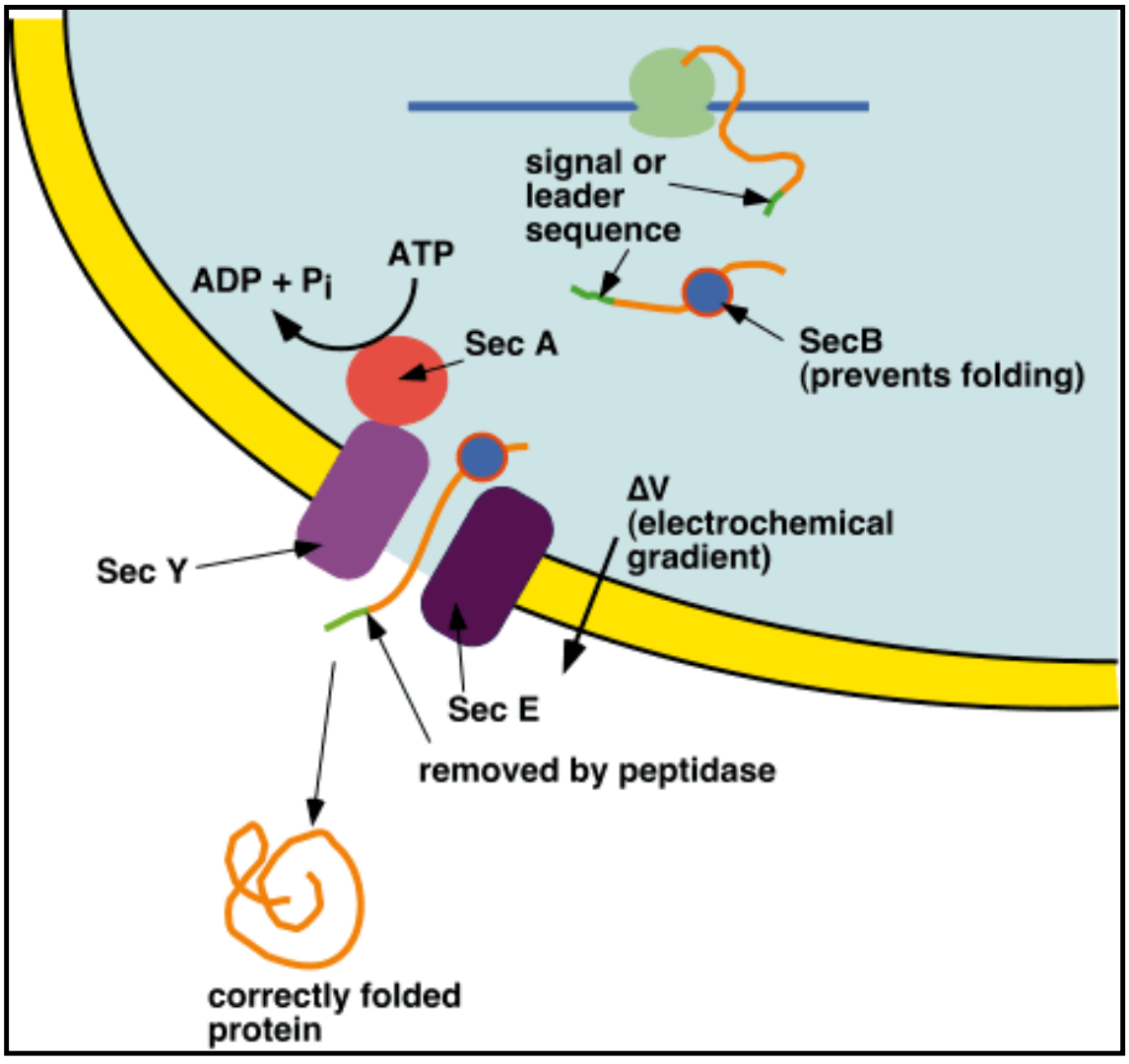
Next: Covalent Modifications

Please report typos, errors etc. by EMAIL (mention the title of this page).

Covalent Modifications

Modifications involving the peptide bond (*peptide bond cleavage* or *limited proteolysis*):

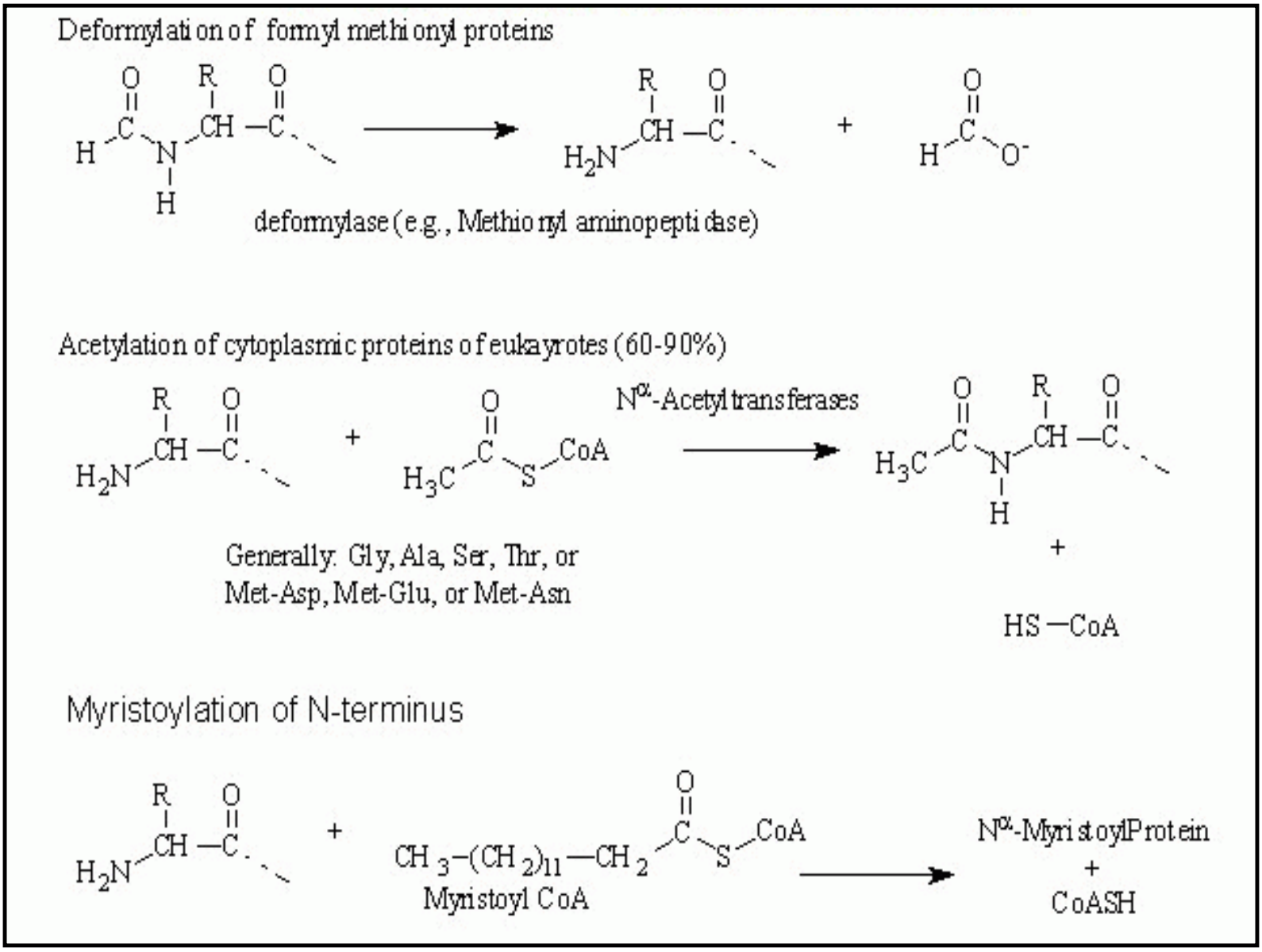
- usually carried out by enzymes called ***peptidases*** or ***proteases:***
    - activation of proenzymes (digestive enzymes, blood clotting cascade, complement activation etc.) and prohormones (insulin)
    - production of active neuropeptides and peptide hormones from high molecular weight precursors
    - macromolecular assembly in virus particles (e.g. HIV protease)
    - removal of signal sequences



These reactions are often exquisitely specific for only one or a few peptide bonds.

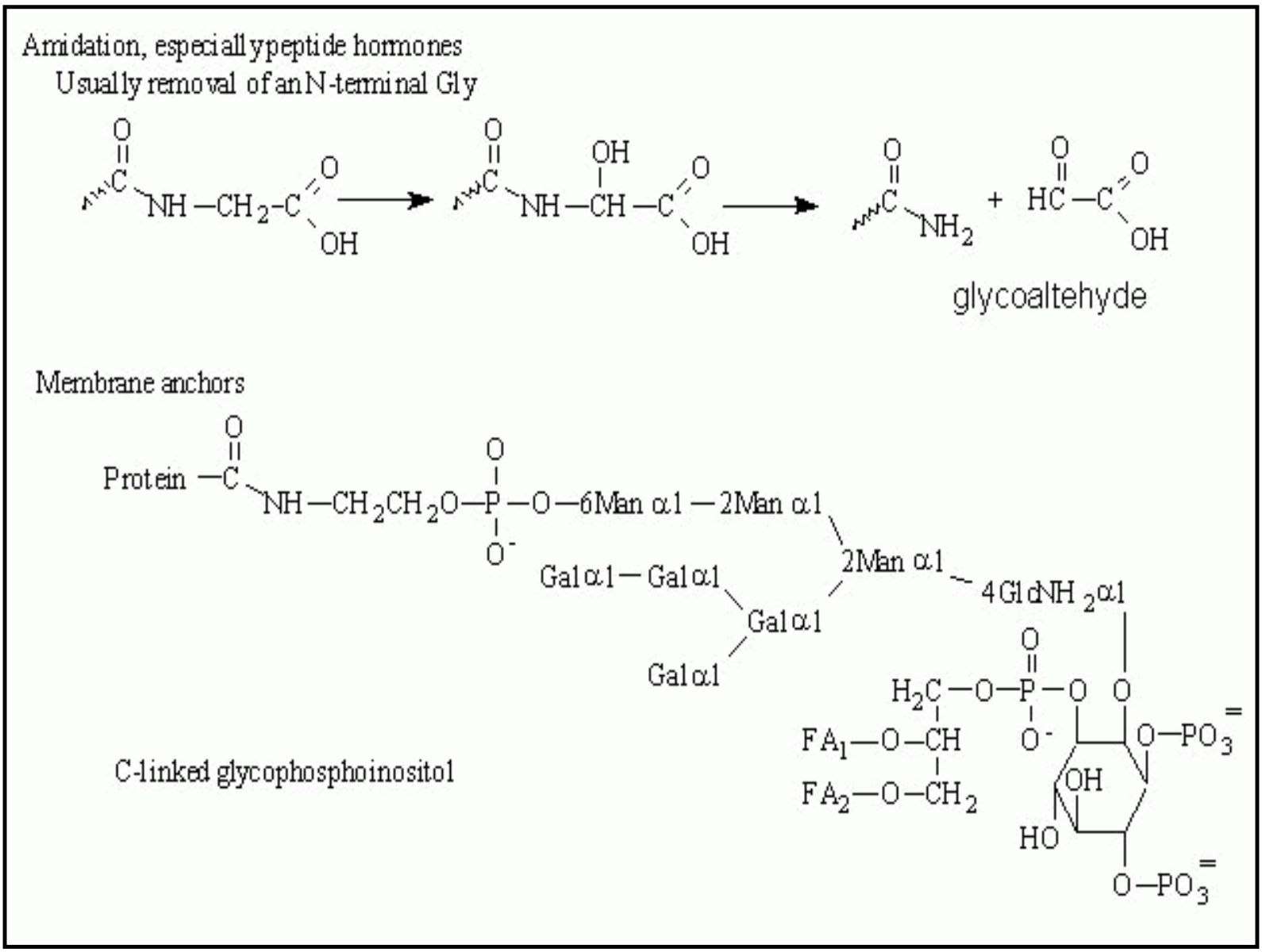
Modifications involving the amino terminus:

- trimming of formyl group from formyl-Met
- proteolytic removal of N-terminal Met by aminopeptidases
- acetylation
- lipidation (myristoylation)



Modifications involving the carboxy terminus:

- amidation of C-terminal glycine
- attachment of membrane anchors



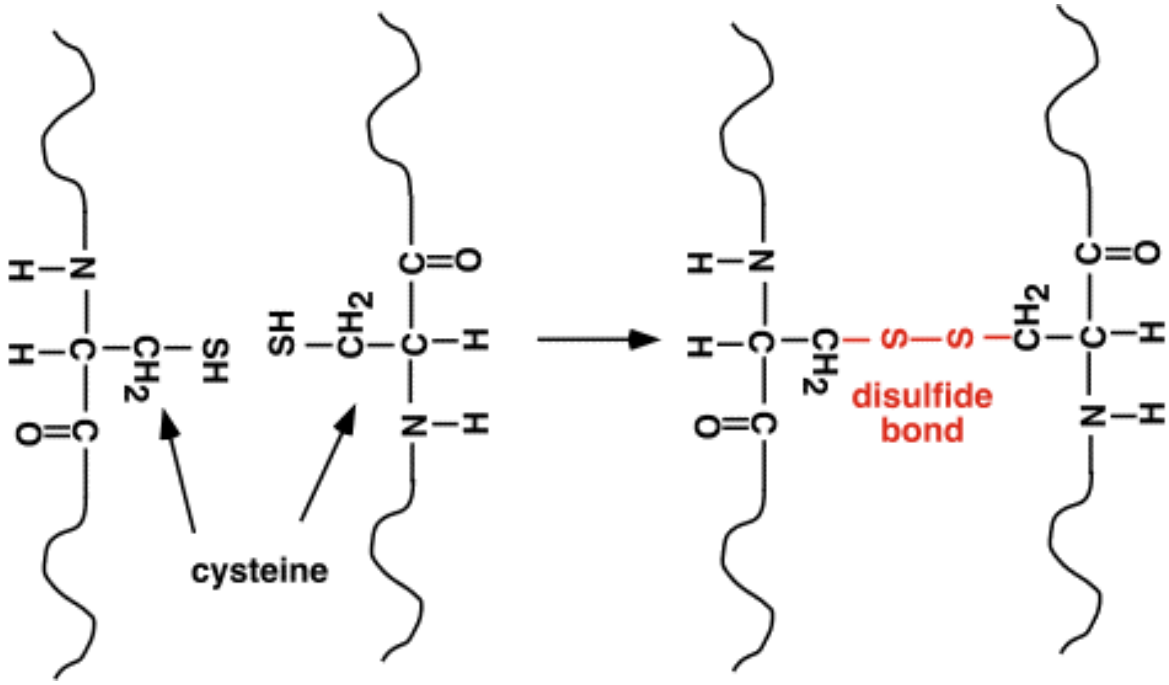
Next: Side Chain Modifications

Please report typos, errors etc. by EMAIL (mention the title of this page).

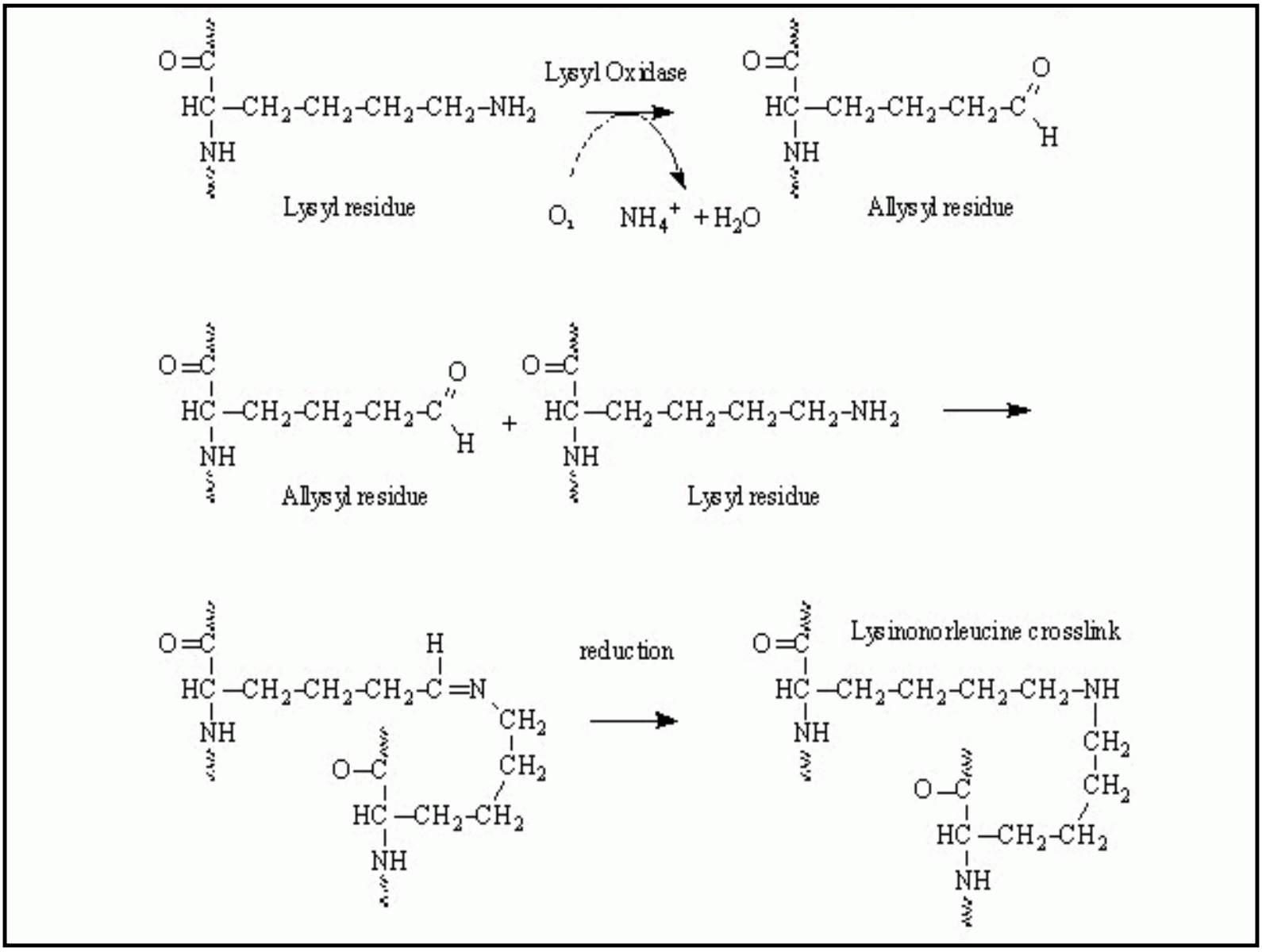
Side Chain Modifications

Modifications involving amino acid side chains:

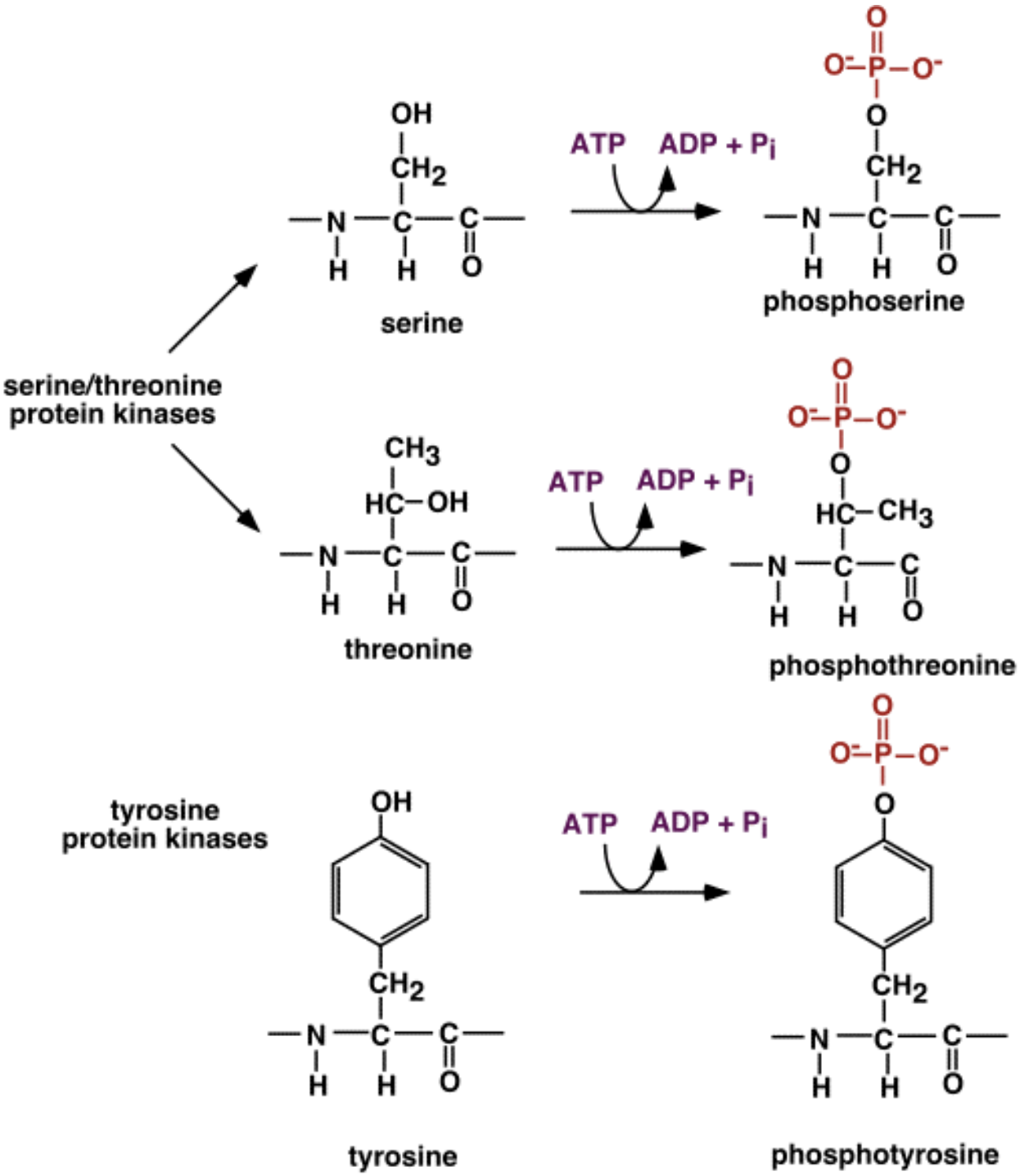
- disulfide cross-linking

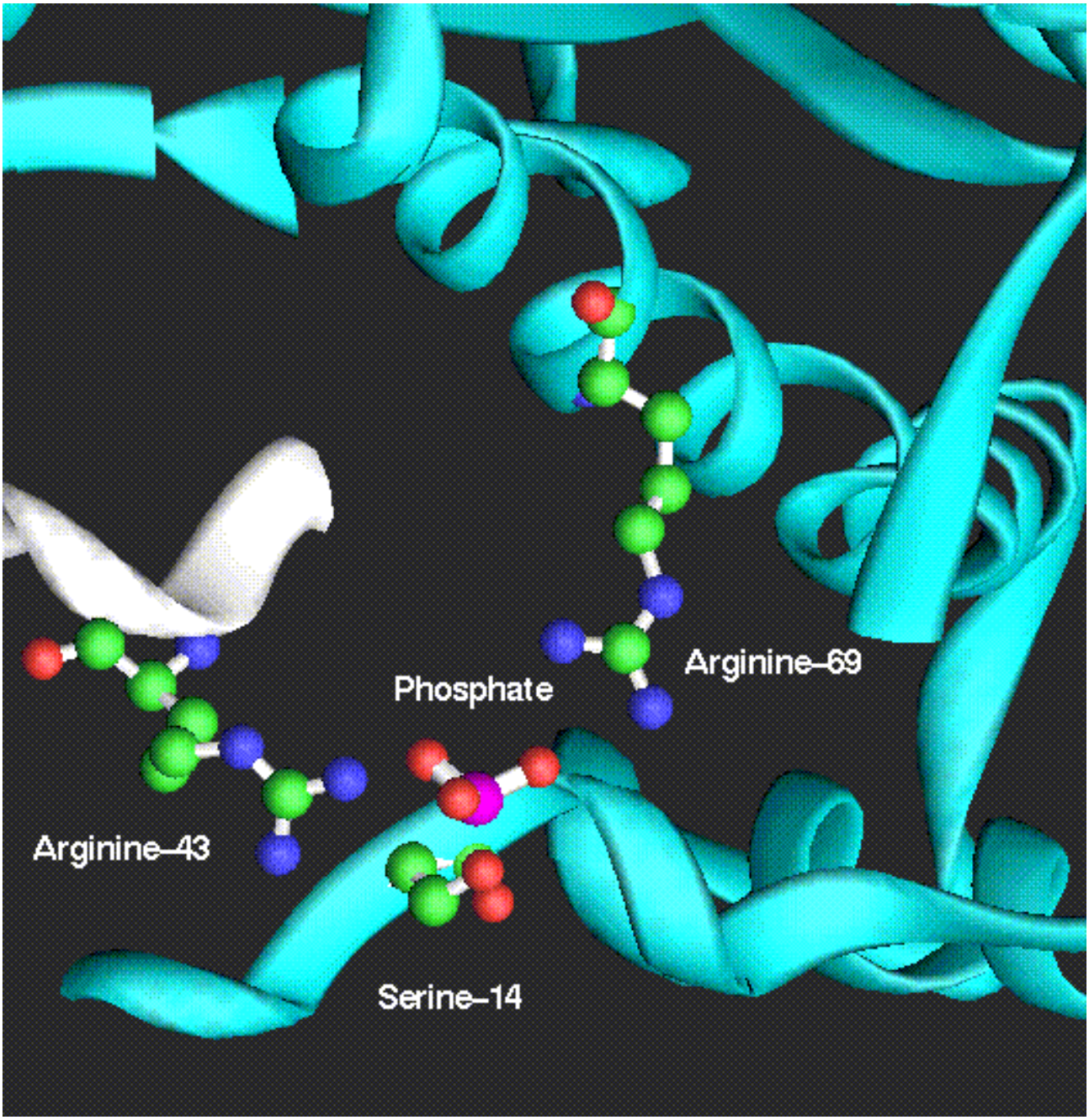


- lysinonorleucine cross-linking (collagen)



- Phosphorylation of hydroxyls by kinases (serine, threonine, tyrosine):


serine/threonine
protein kinases
serine
ATP
ADP + Pi
phosphoserine
threonine
ATP
ADP + Pi
phosphothreonine
tyrosine
protein kinases
tyrosine
ATP
ADP + Pi
phosphotyrosine

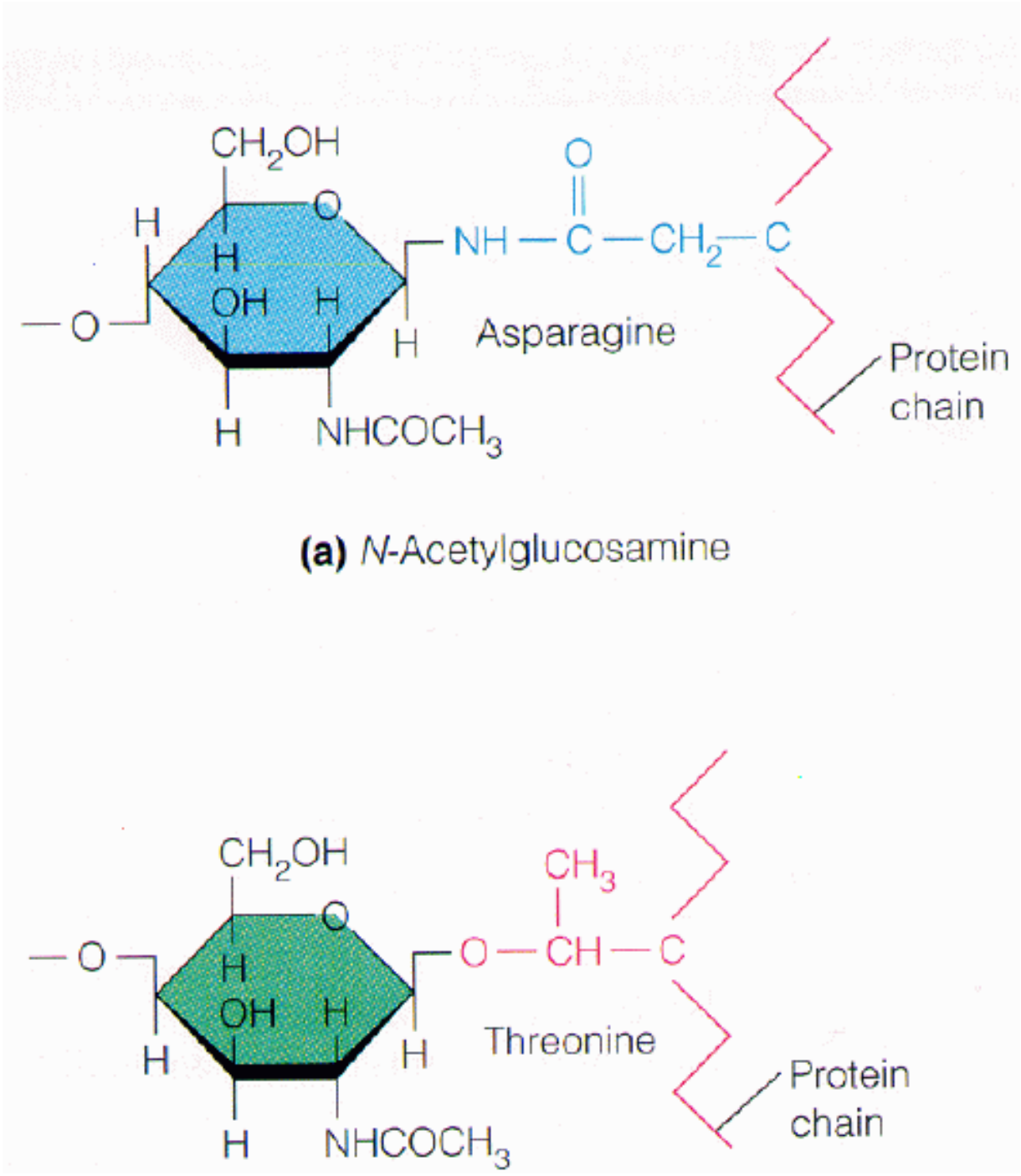


Glycogen phosphorylase is the ultimate enzyme in a cascade which catalyzes the degradation of glycogen to glucose-1-phosphate. Phosphorylation of phosphorylase occurs on serine-14 and converts the inactive *phosphorylase b* to the active *phosphorylase a*.

Phosphorylation is *reversible* and is used in many pathways to control activity. Enzymes that *add* a phosphate to a hydroxyl side chain are commonly called *kinases*. Enzymes that *remove* a phosphate from a phosphorylated side chain are called *phosphatases*.

Glycosylation

- There are two basic types of glycosylation which occur on:
  - asparagines (N-linked, see **(a)** below) and
  - serines and threonines (O-linked, see **(b)** below)



- Covalently attached to the polypeptide as oligosaccharide chains containing 4 to 15 sugars
- Sugars frequently comprise 50% or more of the total molecular weight of a glycoprotein

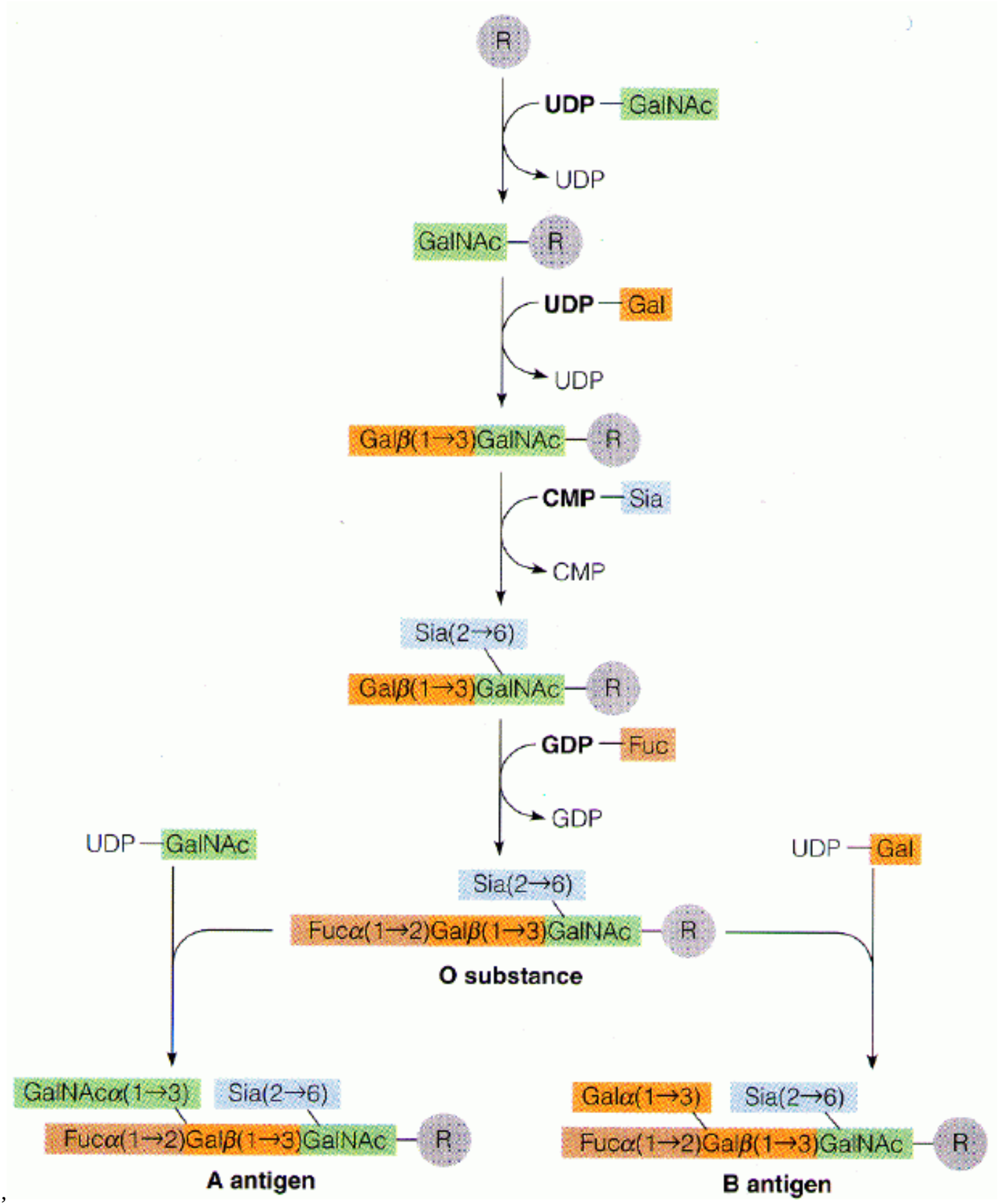
- Most glycosylated proteins are either ***secreted*** or ***remain membrane-bound***
- Glycosylation is the **most abundant form** of post-translational modification
- Glycosylation confers resistance to protease digestion by steric protection
- Important in ***cell-cell recognition***

N-linked glycosylation on asparagine (Asn) side chains:

- an alkali-stable bond between the amide nitrogen of asparagine and the C-1 of an amino sugar residue
- occurs co-translationally in the endoplasmic reticulum (ER) during synthesis
- lipid-linked oligosaccharide complex is transferred to polypeptide by ***oligosaccharyl transferase***
- ***target sequence*** or ***consensus site*** on protein is **Asn**-X-Ser/Thr
- further processing in Golgi apparatus
- Examples:
    - Heavy chain of immunoglobulin G (IgG)
    - Hen ovalbumin
    - Ribonuclease B

O-linked glycosylation on serine (Ser) or threonine (Thr) side chains

- an alkali-labile bond between the hydroxyl group of serine or threonine and an amino sugar
- carried out by a class of membrane-bound enzymes called ***glycosyl transferases*** which reside in the endoplasmic reticulum (ER) or the Golgi apparatus
- nucleotide-linked monosaccharides added to protein side chain one at a time
- Example: Blood group antigens on erythrocyte surface:
    - The ***A antigen*** and ***B antigen*** are pentasaccharides which differ in composition of the 5th sugar residue
    - The ***O substance*** is a tetrasaccharide which is missing the 5^{th} residue and does not elecit an antibody response (non-antigenic).


R
UDP GalNAc
UDP
GalNAc R
UDP Gal
UDP
Galβ(1→3)GalNAc R
CMP Sia
CMP
Sia(2→6)
Galβ(1→3)GalNAc R
GDP Fuc
GDP
UDP GalNAc
UDP Gal
Sia(2→6)
Fucα(1→2)Galβ(1→3)GalNAc R
O substance
GalNAcα(1→3)
Sia(2→6)
Fucα(1→2)Galβ(1→3)GalNAc R
A antigen
Galα(1→3)
Sia(2→6)
Fucα(1→2)Galβ(1→3)GalNAc R
B antigen

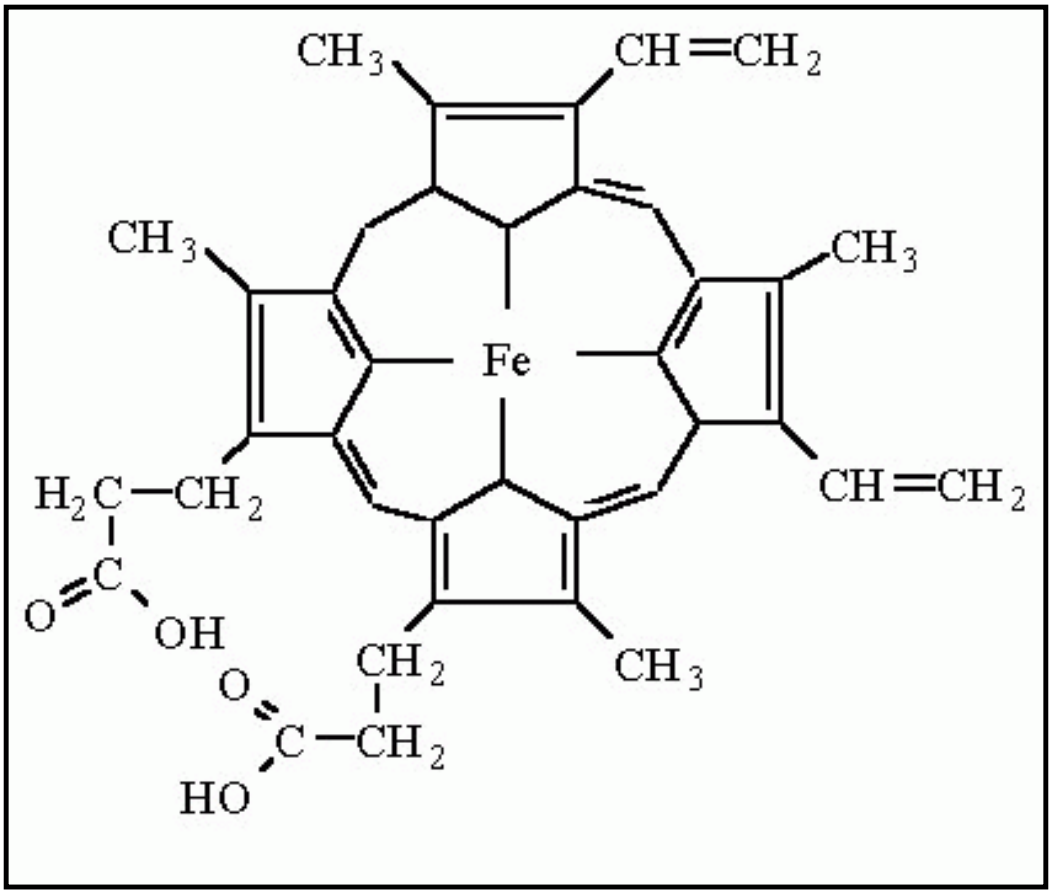
,

Blood Type (Antigen)	Glycolsyl transferase	Antibodies Against	Can Safely Receive Blood from	Can Safely Donate Blood to
O	neither	A & B	O	O, A, B & AB
A	UDP-GalNAc	B	O & A	A & AB
B	UDP-Gal	A	O & B	B & AB
AB	both	neither	O, A, B, & AB	AB

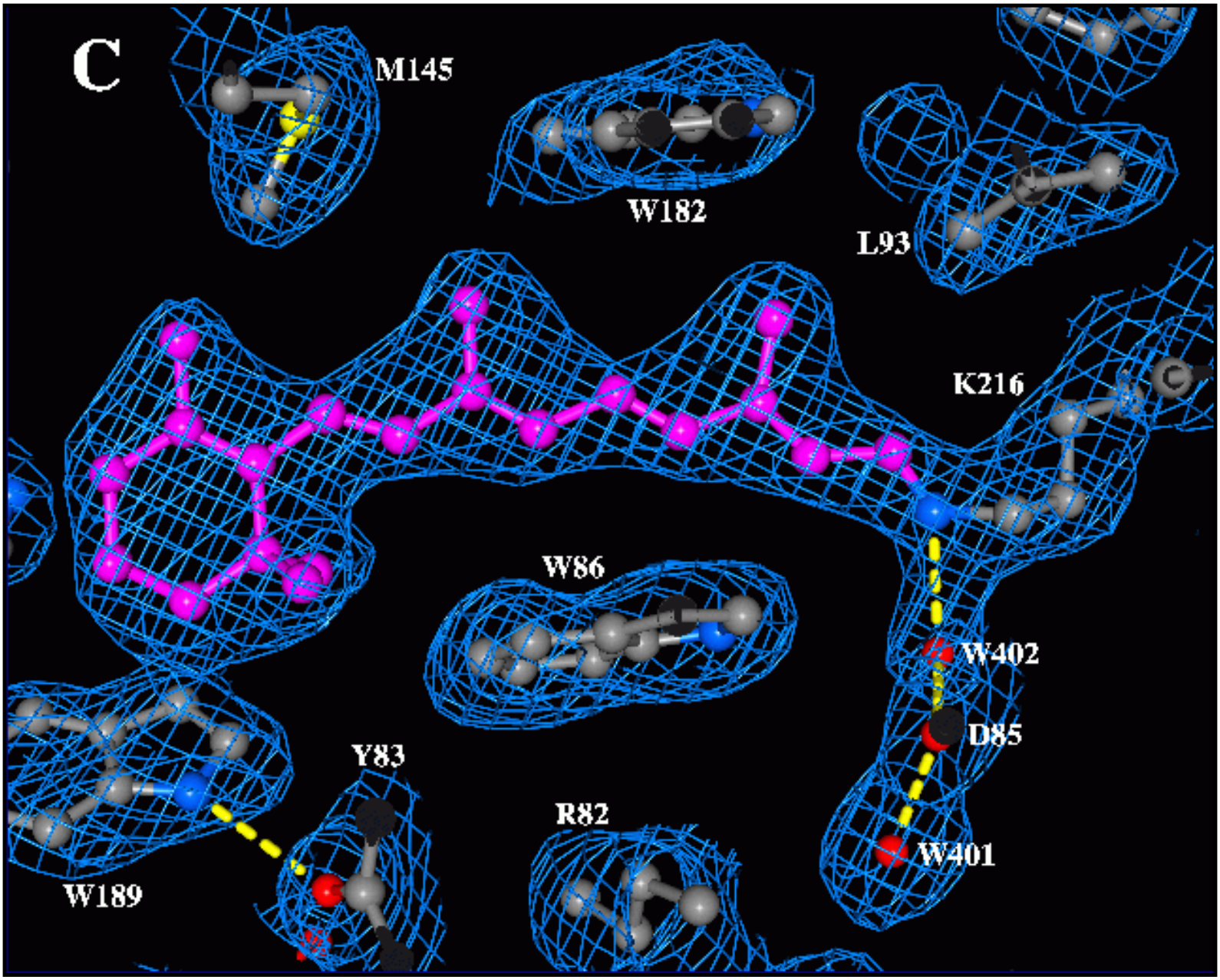
Examples of monosaccharides used in glycosylation

Hexoses	Galactose
	Mannose
	Glucose
Deoxyhexoses	L-Fucose
Hexosamines	N-Acetylglucosamine
	N-Acetylgalactosamine
Sialic Acid	Acylneuraminic Acid
Pentoses	Xylose
	L-Arabinose

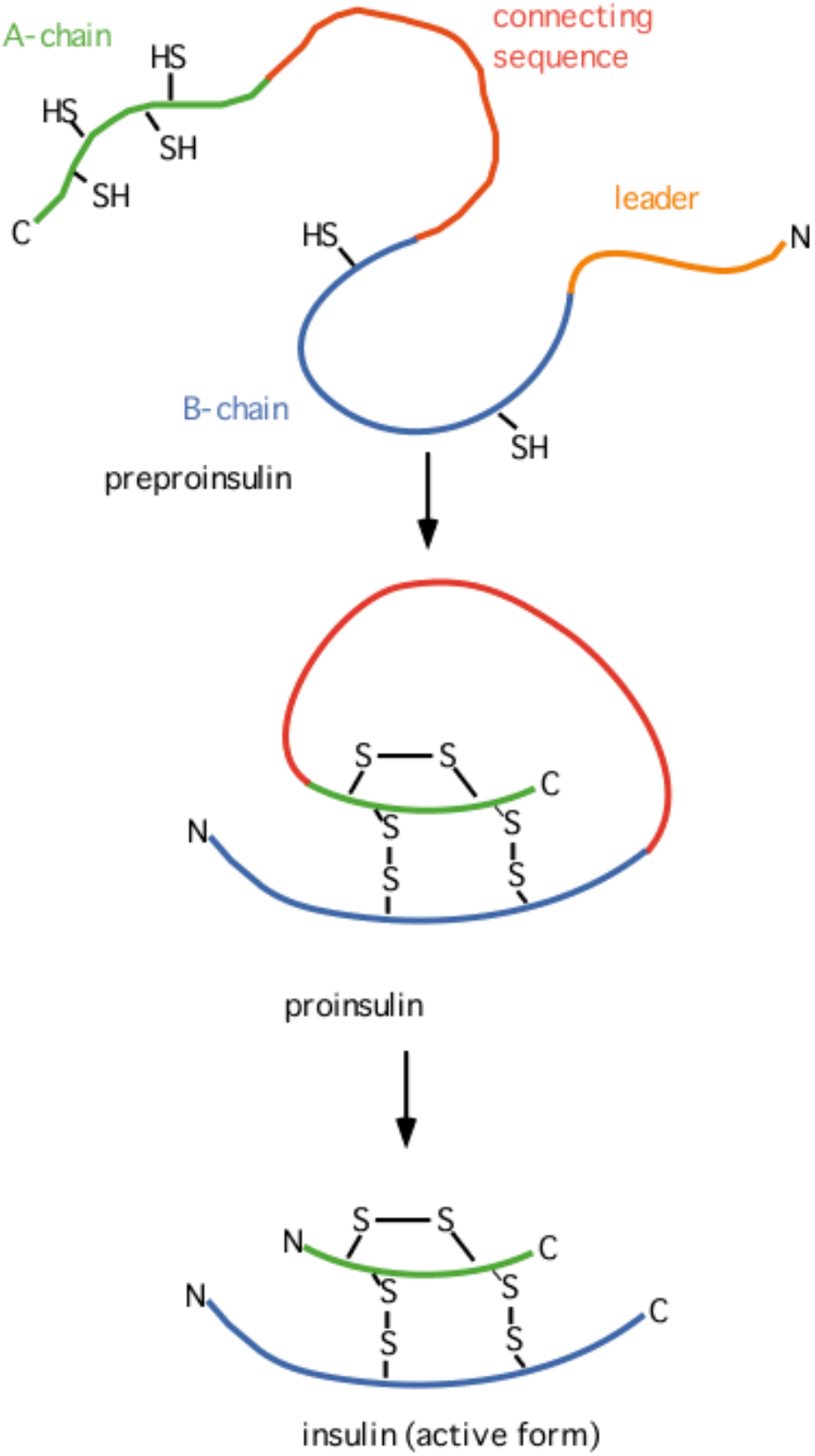
- Prosthetic group attachment (heme, retinal etc.)



Heme group, attached to histidine side chain via Fe.



Retinal attached to lysine side chain via covalent Schiff base linkage.



Processing of pre-pro-insulin to active insulin

- Pre-pro-insulin is synthesized as a random coil on membrane-associated ribosomes
- After membrane-transport the leader sequence (yellow) is cleaved off by a protease and the resulting pro-insulin folds into a stable conformation.
- Disulfide bonds form between cysteine side chains.
- The connecting sequence (red) is cleaved off to form the mature and active insulin molecule.

Next: Noncovalent Modifications

Please report typos, errors etc. by EMAIL (mention the title of this page).

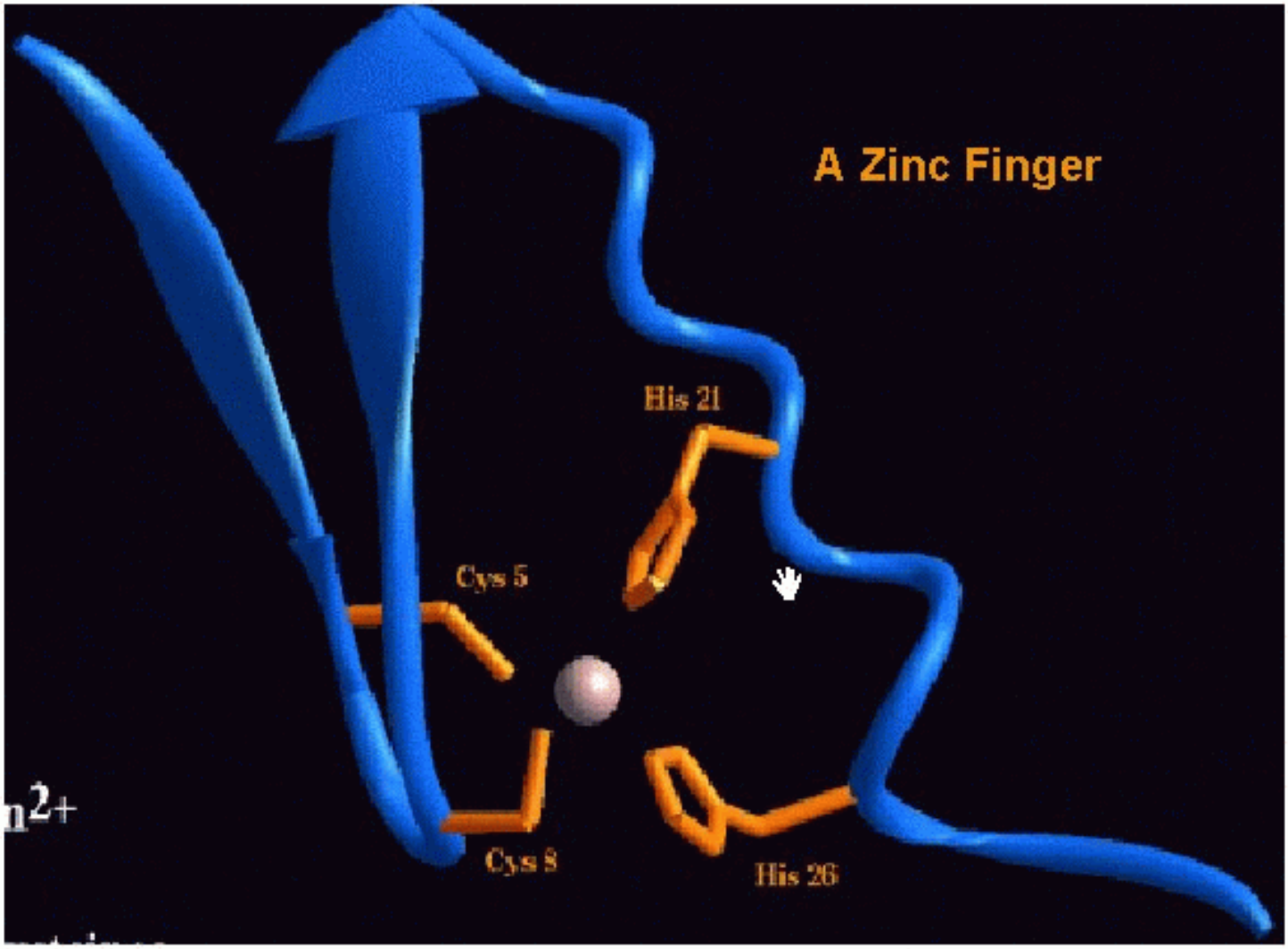
Noncovalent Modifications

Addition of metal ions and co-factors:

Nearly 50% of all proteins contain metal ions

Metal ions play regulatory as well as structural roles

- Calcium (Ca^{++}): very important intra-cellular messenger, i.e. calmodulin
- Magnesium (Mg^{++}): ATP enzymes
- Copper (Cu^{++}), Nickel (Ni^{+}), Iron (Fe^{++})
- Zinc (Zn^{++}): Zinc finger domains are used for DNA recognition:



A zinc finger domain: Zn^{++} is bound by two cysteine and two histidine residues. Zinc finger domains interact in the major groove with three consecutive bases from one strand of duplex B-form DNA.

Modifications involving tertiary structure (protein fold)

Enzymes called ***molecular chaperones*** are responsible for detecting mis-folded proteins.

Chaperones only bind mis-folded proteins that exhibit large hydrophobic patches on their surfaces.

Subunit multimerization

Many enzymes are only functional as multimeric units, either as homo- or hetero-oligomers. Example: ***ribosomes***!

Next: Chaperone-Assisted Protein Folding

Please report typos, errors etc. by EMAIL (mention the title of this page).

Chaperone-Assisted Protein Folding

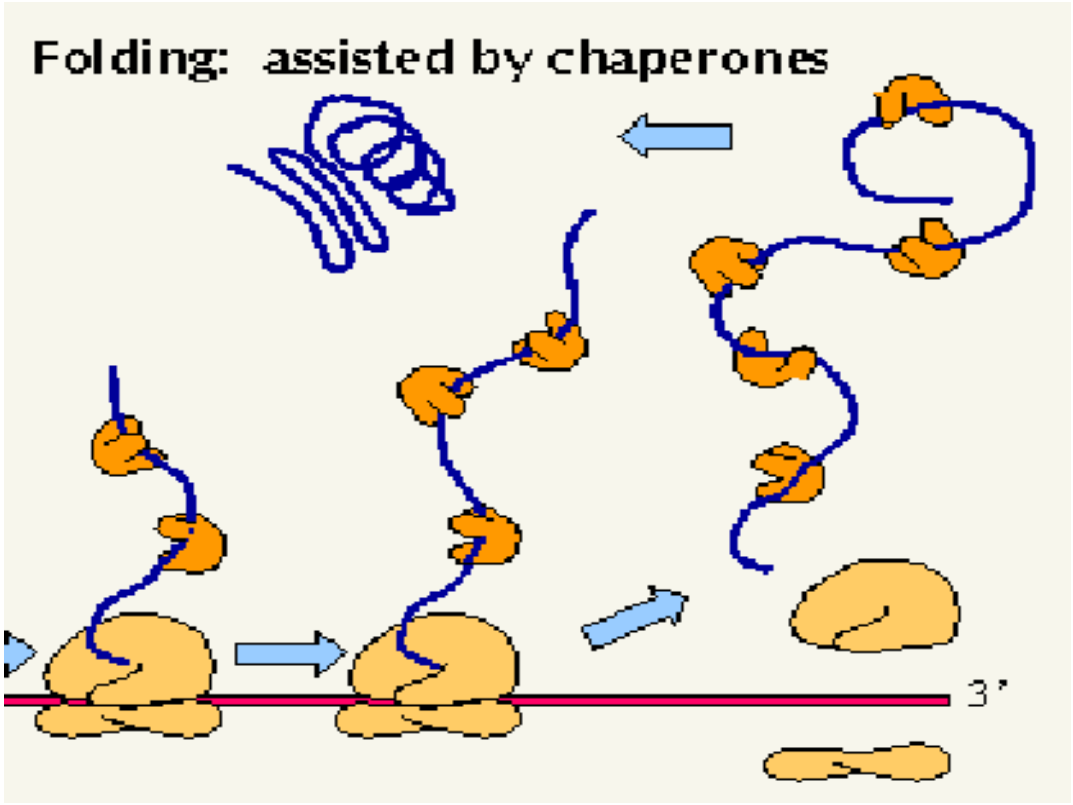
Quote: "The traditional role of the human chaperone described in biochemical terms is to prevent ***improper interactions*** between potentially complementary surfaces, and to disrupt any ***improper liaisons*** which occur."

Chaperones:

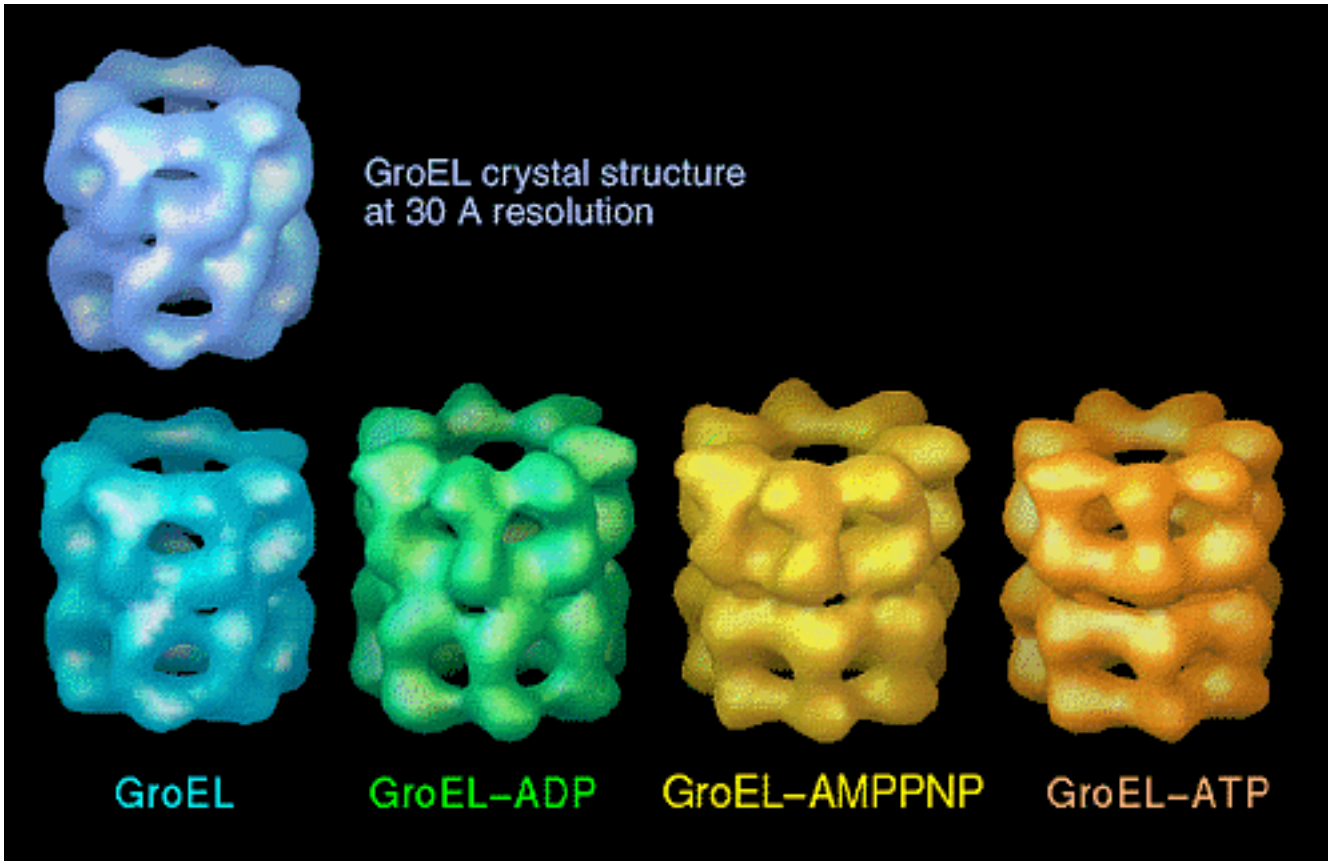
- Mediate folding and assembly.
- Do not convey steric information.
- Do not form part of the final structure.
- Suppress non-productive interactions by binding to transiently exposed portions of the polypeptide chain.
- First identified as ***heat shock proteins*** (Hsp).
- Hsp expression is elevated when cells are grown at higher-than-normal temperatures.
- Stabilize proteins during synthesis.
- Assist in protein folding by binding and releasing unfolded/mis-folded proteins.
- Use an ATP-dependent mechanism.

Major types of chaperones:

- Hsp70 (cytoplasm, ER, chloroplasts, mitochondria):
    - thought to bind and stabilize the nascent polypeptide chain as it is being extruded from the ribosome.
    - also involved in "pulling" newly synthesized polypeptide into ER lumen.



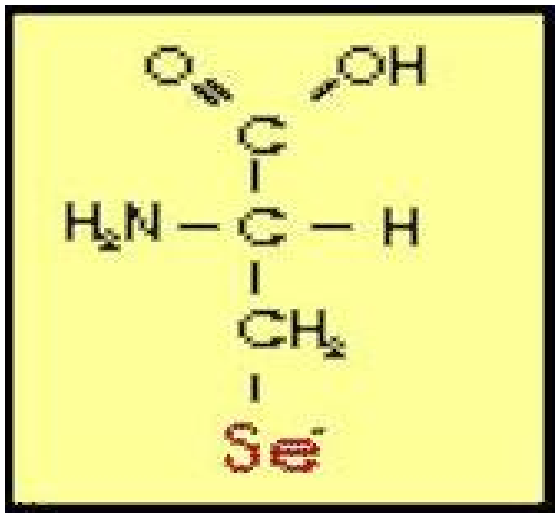
- Hsp60 (mitochondria, chloroplasts):
  - forms large 28-subunit complexes called GroEL



Next: Selenocysteine

Please report typos, errors etc. by EMAIL (mention the title of this page).

Selenocysteine: the 21st Amino Acid?



Although the element selenium was discovered in 1817 by Berzelius it was only shown over 100 years later to be an essential micronutrient in all three lines of decent. Subsequent analysis of several enzymes that catalyze oxidation-reduction reactions showed that selenium occurs in the form of the unusual amino acid ***selenocysteine***. How this amino acid is incorporated into the protein was unclear in these days.

Today it is well established that the incorporation of selenocysteine is *co-translational*. Interestingly, the base triplet encoding this amino acid is **UGA**, a codon that normally functions as a ***STOP signal*** in translation. Since this codon has still retained its "normal" function in all organisms known to synthesize selenocysteine-containing proteins, the incorporation of this amino acid requires a specific pathway. This

pathway has been elucidated for the bacterium *E. coli* by August Böck and coworkers.

General pathway

There are one *cis*-acting element and four *trans*-acting factors involved in this incorporation. The *cis*-acting element is a specific *stem-loop structure* of the mRNA directly 3' of the **UGA** codon. The four *trans*-acting factors are:

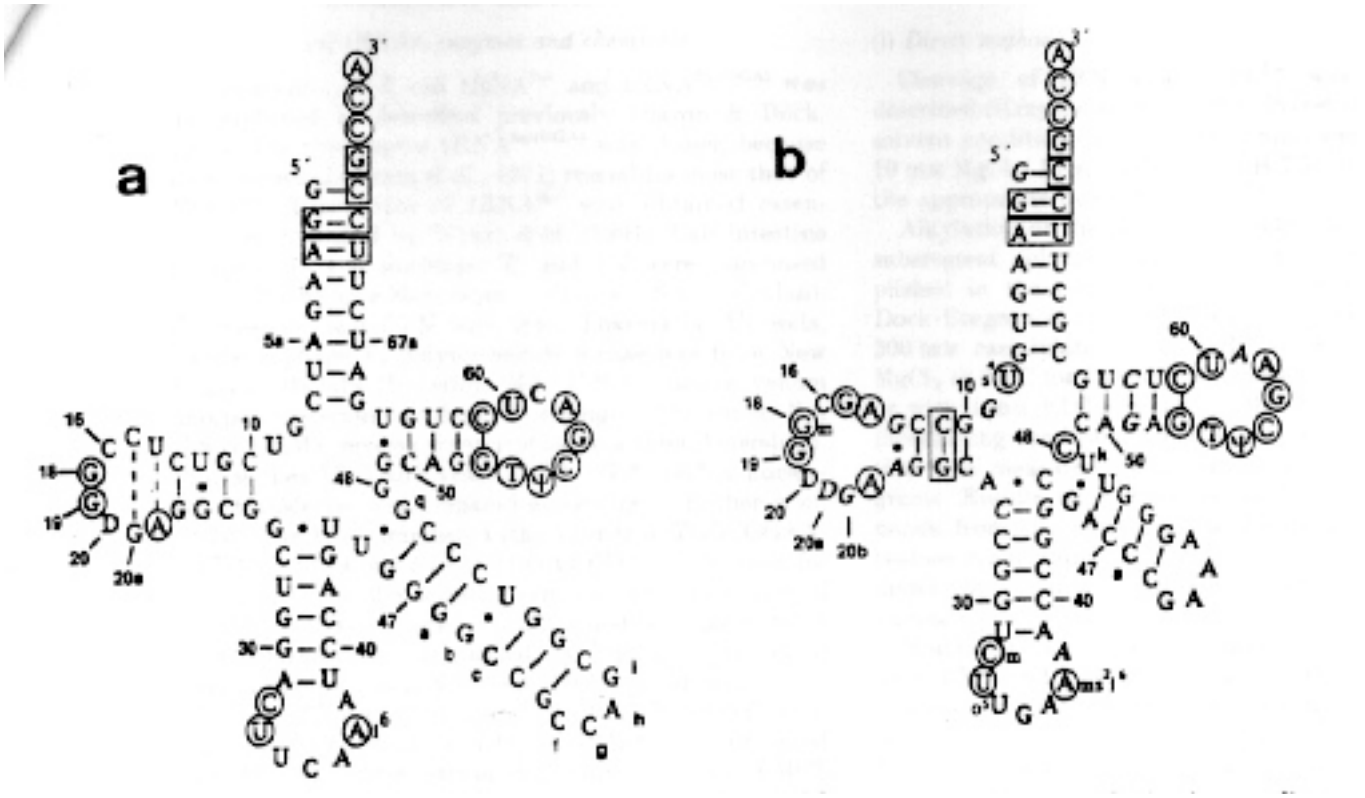
- a *specific* $tRNA^{Sec}$ which is charged with serine by serinyl-tRNA synthetase
- the enzyme selenophosphate synthetase which generates inorganic selenophosphate
- the enzyme selenocysteine synthetase which uses selenophosphate to convert seryl-$tRNA^{Sec}$ to selenocysteyl-$tRNA^{Sec}$
- a specific translation factor (SelB) that substitutes for EF-Tu ***and*** recognizes the *cis*-acting element

In the first step the specific *$tRNA^{Sec}$* is charged by the normal seryl-tRNA synthetase with *serine*, and that serine is subsequently converted to selenocysteine by the enzyme *selenocysteine synthetase*. The low molecular weight selenium donor - selenophosphate - is provided by the action of the *selenophosphate synthetase.* Finally, the selenocysteyl-$tRNA^{Sec}$ is recognized by a specific translation factor

SelB that delivers it to a **UGA** codon at the ribosomal A-site ***only*** in the presence of the stem-loop structure downstream of the codon on the mRNA.

tRNA^{Sec}

tRNA^{Sec} differs in several aspects from the consensus for canonical tRNAs. In particular it includes an acceptor stem elongated by extra one base pair, an elongated D-arm with only four nucleotides in the loop, and a **UCA** anticodon. These differences ensure that tRNA^{Sec} is not recognized by the normal translation factor EF-Tu thus preventing mis-incorporation.



The two structures shown here are (a) $tRNA^{Sec}$ and (b) $tRNA^{Ser}$ of *E. coli*. Residues which are normally conserved in tRNA are shown circled. Identity nucleotides in $tRNA^{Ser}$ and their counterparts in $tRNA^{Sec}$ are shown boxed. $tRNA^{Sec}$ is the longest tRNA species known consisting of 95 nucleotides.

Selenophosphate synthetase

This enzyme catalyses the synthesis of the low molecular weight donor *selenophosphate* using ATP as the phosphate donor. Interestingly, the gamma-phosphate is transferred to selenium whereas the beta-phosphate is released, leaving AMP.

Selenocysteine synthetase

This enzyme possesses a prosthetic group (pyridoxal phosphate). The active enzyme is composed of ten identical subunits arranged as a stack of two five-membered rings. It converts the serine attached to $tRNA^{Sec}$ to selenocysteine.

Translation factor SelB

This translation factor substitutes for elongation factor EF-Tu in the specific incorporation of selenocysteine. This protein is considerable longer than its regular counterpart. The N-terminal half of the protein resembles EF-Tu in structure and function - GTP and tRNA binding - whereas the C-terminal half is responsible for the *recognition of the*

specific stem-loop structure adjacent to the **UGA** codon. A complex between SelB, GTP, selenocysteyl-tRNA^{Sec} and the stem-loop structure of the mRNA has been detected and shown to be functionally necessary.

Next: Summary

Please report typos, errors etc. by EMAIL (mention the title of this page).

Summary

- Post-translational processing controls ***folding, targeting***, ***activation*** and ***stability*** of proteins
- Co- and post-translational modifications increase ***diversity*** and ***functionality*** of proteins
- Common forms of co- or post-translational modifications are:
  - proteolysis
  - phosphorylation
  - glycosylation
  - metal binding
- Chaperones mediate folding and assembly of newly synthesized proteins
- Selenocysteine is sometimes called the 21st amino acid
  - uses one of the three STOP codons
  - has its own tRNA^{Sec}
  - requires other factors, including a *cis*-acting element on the mRNA

Next Lecture: Protein Trafficking/Targeting

Please report typos, errors etc. by EMAIL (mention the title of this page).