CHAUDHARY MAHADEO PRASAD COLLEGE

(A CONSTITUENT PG COLLEGE OF UNIVERSITY OF ALLAHABAD)

E-Learning Module

Subject: Botany

(Study material for Post Graduate Students)

M.Sc. IV Sem

COURSE CODE: BOT 605

Molecular Biology and Molecular techniques

Unit II: Protein Sorting

Developed by

Dr. Sarita Srivastava

Associate Professor

Department of Botany

Protein targeting or Protein Sorting

How molecular labels are used to direct proteins to different parts of the cell (and to the cell exterior).

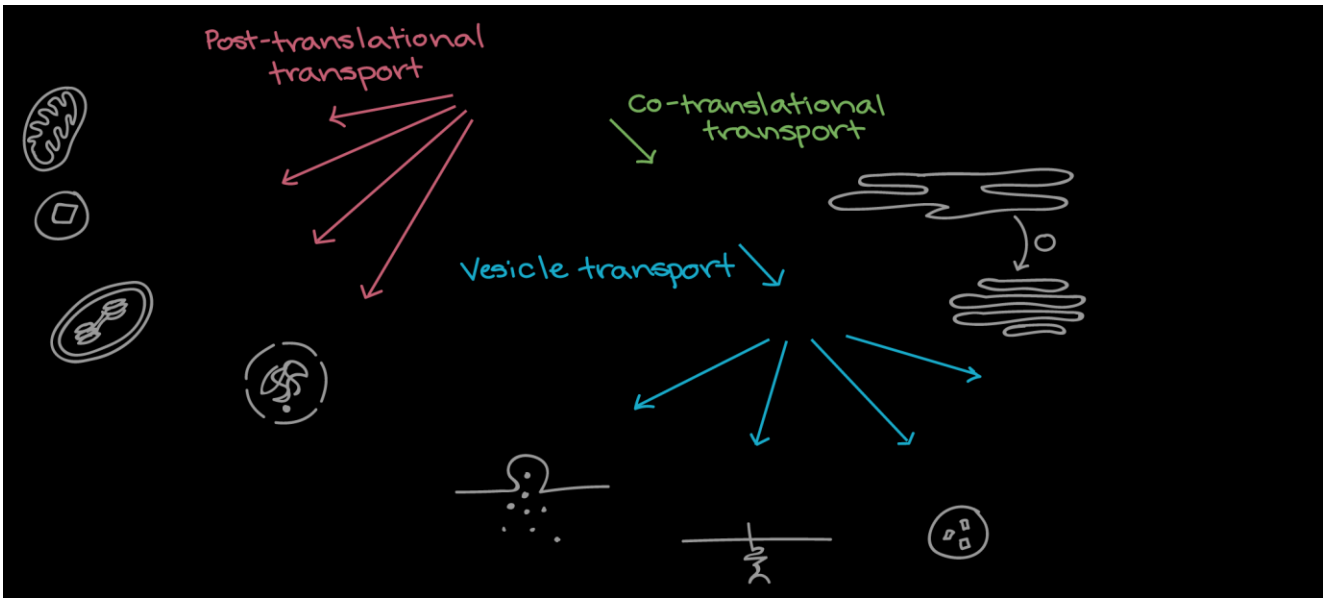
Protein targeting or protein sorting is the biological mechanism by which proteins are transported to their appropriate destinations in the cell or outside it. Proteins can be targeted to the inner space of an organelle, different intracellular membranes, plasma membrane, or to exterior of the cell via secretion. This delivery process is carried out based on information contained in the protein itself. Correct sorting is crucial for the cell; errors can lead to diseases.

Introduction:

Different proteins need to be sent to different parts of a eukaryotic cell, or, in some cases, exported out of the cell and into the extracellular space. How do the right proteins get to the right places?
Cells have various shipping systems, kind of like molecular versions of the postal service, to make sure that proteins arrive at their correct destinations. In these systems, molecular labels (often, amino acid sequences) are used to "address" proteins for delivery to specific locations. Let's take a look at how these shipping systems work.

Overview of cellular shipping routes

Translation of all proteins in a eukaryotic cell begins in the cytosol (except for a few proteins made in mitochondria and chloroplasts). As a protein is made, it passes step by step through a shipping "decision tree." At each stage, the protein is checked for molecular tags to see if it needs to be re-routed to a different pathway or destination.



The first major branch point comes shortly after translation starts. At this point, the protein will either remain in the cytosol for the rest of translation, or be fed into the endoplasmic reticulum (ER) as it is translated.

- Proteins are fed into the ER during translation if they have an amino sequence called a **signal peptide**. In general, proteins bound for organelles in the endomembrane system (such as the ER, Golgi apparatus, and lysosome) or for the exterior of the cell must enter the ER at this stage.
- Proteins that do not have a signal peptide stay in the cytosol for the rest of translation. If they lack other "address labels," they'll stay in the cytosol permanently. However, if they have the right labels, they can be sent to the mitochondria, chloroplasts, peroxisomes, or nucleus after translation.

The endomembrane system and Secretory pathway: Proteins destined for any part of the endomembrane system (or the outside of the cell) are brought to the ER during translation and fed in as they're made.

Targeting signals

Targeting signals are the pieces of information that enable the cellular transport machinery to correctly position a protein inside or outside the cell. This information is contained in the polypeptide chain or in the folded protein. The continuous stretch of amino acid residues in the chain that enables targeting are called **signal peptides** or **targeting peptides**.

There are two types of targeting peptides, the **presequences** and the **internal targeting peptides.** The presequences of the targeting peptide are often found at the N-terminal extension and is composed of between 6-136 basic and hydrophobic amino acids. In case of peroxisomes the targeting sequence is on the C-terminal extension mostly. Other signals, known as **signal patches**, are composed of parts which are separate in the primary sequence. They become functional when folding brings them together on the protein surface. In addition, protein modifications like glycosylations can induce targeting.

Signal Peptides

The **signal peptide** that sends a protein into the endoplasmic reticulum during translation is a series of hydrophobic (“water-fearing”) amino acids, usually found near the beginning (N-terminus) of the protein. When this sequence sticks out of the ribosome, it’s recognized by a protein complex called the **signal-recognition particle (SRP)**, which takes the ribosome to the ER. There, the ribosome feeds its amino acid chain into the ER lumen (interior) as it's made. In some cases, the signal peptide is snipped off during translation and the finished protein is released into the interior of the ER (as shown above). In other cases, the signal peptide or another stretch of hydrophobic amino acids gets embedded in the ER membrane. This creates a transmembrane (membrane-crossing) segment that anchors the protein to the membrane

Transport through the endomembrane system

In the ER, proteins fold into their correct shapes, and may also get sugar groups attached to them. Most proteins are then transported to the Golgi apparatus in membrane vesicles. Some proteins, however, need to stay in the ER and do their jobs there. These proteins have amino acid tags that ensure they are shipped back to the ER if they "escape" into the Golgi. In the Golgi apparatus, proteins may undergo more modifications (such as addition of sugar groups) and before going on to their final destinations. These destinations include lysosomes, the plasma membrane, and the cell exterior. Some proteins need

to do their jobs in the Golgi (are "Golgi-resident), and a variety of molecular signals, including amino acid tags and structural features, are used to keep them there or bring them back If they don't have any specific tags, proteins are sent from the Golgi to the cell surface, where they're secreted to the cell exterior (if they're free-floating) or delivered to the plasma membrane (if they're membrane-embedded). This default pathway is shown in the diagram above for a membrane protein, colored in green, that bears sugar groups, colored in purple.
Proteins are shipped to other destinations if they contain the right molecular labels. For example, proteins destined for the lysosome have a molecular tag consisting of a sugar with a phosphate group attached. In the Golgi apparatus, proteins with this tag are sorted into vesicles bound for the lysosome.

Targeting to non-endomembrane organelles

Proteins that are made in the cytosol (don't enter ER during translation) may stay permanently in the cytosol. However, they may also be shipped to other, non-endomembrane destinations in the cell. For instance, proteins bound for the mitochondria, chloroplasts, peroxisomes, and nucleus are usually made in the cytosol and delivered after translation is complete.
to be delivered to one of these organelles after translation, a protein must contain a specific amino acid "address label." The label is recognized by other proteins in the cell, which help transport the protein to the right destination.

As an example, let's consider delivery to the peroxisome, an organelle involved in detoxification. Proteins needed in the peroxisome have a specific sequence of amino acids called a **peroxisomal targeting signal**. The classic signal consists of just three amino acids, serine-lysine-leucine, found at the very end (C-terminus) of a protein. This pattern of amino acids is recognized by a helper protein in the cytosol, which brings the protein to the peroxisome.

Mitochondrial, chloroplast, and nuclear targeting are generally similar to peroxisomal targeting. That is, a certain amino acid sequence sends the protein to its target organelle (or a compartment inside that organelle). However, the nature of the "address labels" is different in each case.

Protein translocation

In 1970, Günter Blobel conducted experiments on the translocation of proteins across membranes. He was awarded the 1999 Nobel prize for his findings. He discovered that many proteins have a signal sequence, that is, a short amino acid sequence at one end that functions like a postal code for the target organelle. The translation of mRNA into protein by a ribosome takes place within the cytosol. If the synthesized proteins "belong" in a different organelle, they can be transported there in either of two ways depending on the protein: Co-translational translocation (translocation during the process of translation), and post-translational translocation (translocation after the process of translation is complete).

Co-translational translocation

Most proteins that are secretory, membrane-bound, or reside in the endoplasmic reticulum (ER), golgi or endosomes use the **co-translational translocation pathway.** This process

begins with the N-terminal signal peptide of the protein being recognized by a signal recognition particle (SRP) while the protein is still being synthesized on the ribosome. The synthesis pauses while the ribosome-protein complex is transferred to an SRP receptor on the ER in eukaryotes, and the plasma membrane in prokaryotes. There, the nascent protein is inserted into the translocon, a membrane-bound protein conducting channel composed of the Sec61 translocation complex in eukaryotes, and the homologous SecYEG complex in prokaryotes. In secretory proteins and type I transmembrane proteins, the signal sequence is immediately cleaved from the nascent polypeptide once it has been translocated into the membrane of the ER (eukaryotes) or plasma membrane (prokaryotes) by signal peptidase. The signal sequence of type II membrane proteins and some polytopic membrane proteins are not cleaved off and therefore are referred to as signal anchor sequences. Within the ER, the protein is first covered by a chaperone protein to protect it from the high concentration of other proteins in the ER, giving it time to fold correctly. Once folded, the protein is modified as needed (for example, by glycosylation), then transported to the Golgi for further processing and goes to its target organelles or is retained in the ER by various ER retention mechanisms.

The amino acid chain of transmembrane proteins, which often are transmembrane receptors, passes through a membrane one or several times. They are inserted into the membrane by translocation, until the process is interrupted by a stop-transfer sequence, also called a membrane anchor or signal-anchor sequence. These complex membrane proteins are at the moment mostly understood using the same model of targeting that has been developed for secretory proteins. However, many complex multi-transmembrane proteins contain structural aspects that do not fit the model. Seven transmembrane G-protein coupled receptors (which represent about 5% of the genes in humans) mostly do not have an amino-terminal signal sequence. In contrast to secretory proteins, the first transmembrane domain acts as the first signal sequence, which targets them to the ER membrane. This also results in the translocation of the amino terminus of the protein into the ER membrane lumen. This would seem to break the rule of "co-translational" translocation which has always held for mammalian proteins targeted to the ER. This has been demonstrated with opsin with in vitro experiments.[1][2] A great deal of the mechanics of transmembrane topology and folding remains to be elucidated.

Post-translational translocation

Even though most secretory proteins are co-translationally translocated, some are translated in the cytosol and later transported to the ER/plasma membrane by a post-translational system. In prokaryotes this requires certain cofactors such as SecA and SecB. This pathway is poorly understood in eukaryotes, but is facilitated by Sec62 and Sec63, two membrane-bound proteins.

In addition, proteins targeted to other destinations, such as mitochondria, chloroplasts, or peroxisomes, use specialized post-translational pathways. Also, proteins targeted for the nucleus are translocated post-translation. They pass through the nuclear envelope via nuclear pores.

Sorting of proteins to different organelles

Mitochondria

Most mitochondrial proteins are synthesized as cytosolic precursors containing uptake peptide signals. Cytosolic chaperones deliver preproteins to channel linked receptors in the mitochondrial membrane. The preprotein with presequence targeted for the mitochondria is bound by receptors and the General Import Pore (GIP) (Receptors and GIP are collectively known as Translocase of Outer Membrane or TOM) at the outer membrane. The preprotein is translocated through TOM as hairpin loops. The preprotein is transported through the intermembrane space by small TIMs (which also acts as molecular chaperones) to the TIM23 or 22 (Translocase of Inner Membrane) at the inner membrane. Within the matrix the targeting sequence is cleaved off by mtHsp70.

Three mitochondrial outer membrane receptors are known:

TOM70 Binds to internal targeting peptides and acts as a docking point for cytosolic chaperones.
TOM20 Binds presequences
TOM22 Binds both presequences and internal targeting peptides

The TOM channel (TOM40) is a cation specific high conductance channel with a molecular weight of 410 kDa and a pore diameter of 21Å.

The presequence translocase23 (TIM23) is localized to the mitochondrial inner membrane and acts a pore forming protein which binds precursor proteins with its N-terminus. TIM23 acts a translocator for preproteins for the mitochondrial matrix, the inner mitochondrial membrane as well as for the intermembrane space. TIM50 is bound to TIM23 at the inner mitochondrial side and found to bind presequences. TIM44 is bound on the matrix side and found binding to mtHsp70. The presequence translocase22 (TIM22) binds preproteins exclusively bound for the inner mitochondrial membrane.

Mitochondrial matrix targeting sequences are rich in positively charged amino acids and hydroxylated ones. Proteins are targeted to submitochondrial compartments by multiple signals and several pathways.

Targeting to the outer membrane, intermembrane space, and inner membrane often requires another signal sequence in addition to the matrix targeting sequence.

Chloroplasts

The preprotein for chloroplasts may contain a stromal import sequence or a stromal and thylakoid targeting sequence. The majority of preproteins are translocated through the Toc and Tic complexes located within the chloroplast envelope. In the stroma the stromal import sequence is cleaved off and folded as well as intra-chloroplast sorting to thylakoids continues. Proteins targeted to the envelope of chloroplasts usually lack cleavable sorting sequence.

Many proteins are needed in both mitochondria and chloroplasts. In general, the targeting peptide is of intermediate character to the two specific ones. The targeting peptides of these proteins have a high content of basic and hydrophobic amino acids, a low content of negatively charged amino acids. They have a lower content of alanine and a higher content of leucine and phenylalanine. The dual targeted proteins have a more hydrophobic targeting peptide than both mitochondrial and chloroplastic ones.

Peroxisomes

All peroxisomal proteins are encoded by nuclear genes.

There are two types of known **Peroxisome Targeting Signals (PTS)**:

Peroxisome targeting signal 1 (PTS1): a C-terminal tripeptide with a consensus sequence (S/A/C)-(K/R/H)-(L/A). The most common PTS1 is serine-lysine-leucine (SKL). Most peroxisomal matrix proteins possess a PTS1 type signal.

Peroxisome targeting signal 2 (PTS2): a nonapeptide located near the N-terminus with a consensus sequence (R/K)-(L/V/I)-XXXXX-(H/Q)-(L/A/F) (where X can be any amino acid).

There are also proteins that possess neither of these signals. Their transport may be based on a so-called **"piggy-back"** mechanism: such proteins associate with PTS1-possessing matrix proteins and are translocated into the peroxisomal matrix together with them.

Diseases

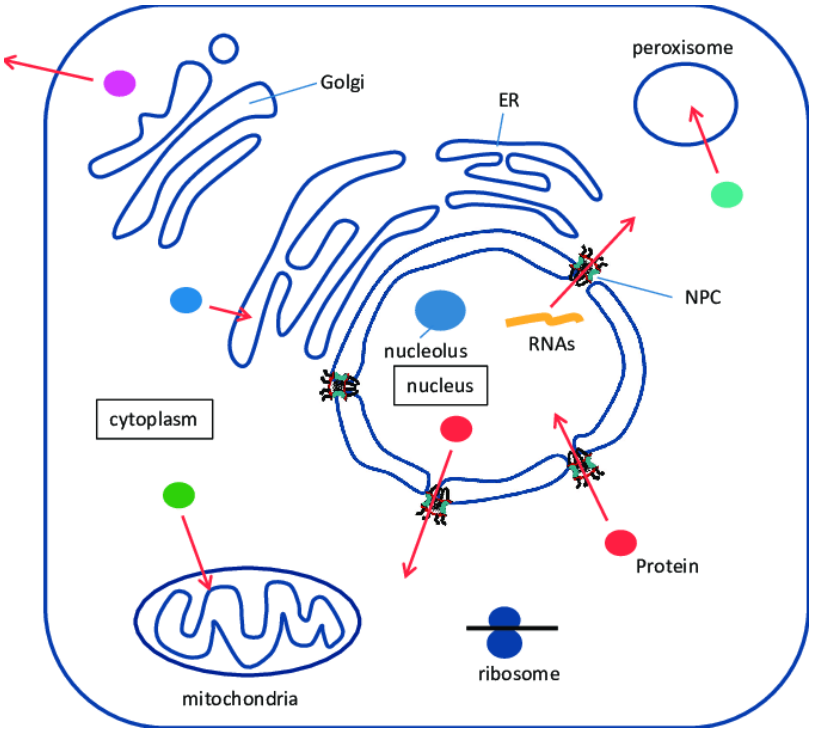
Peroxisomal protein transport is defective in the following genetic diseases:

- Zellweger syndrome.
- Adrenoleukodystrophy (ALD).
- Refsum disease

In bacteria and archaea As discussed above (see protein translocation), most prokaryotic membrane-bound and secretory proteins are targeted to the plasma membrane by either a co-translation pathway that uses bacterial SRP or a post-translation pathway that requires SecA and SecB. At the plasma membrane, these two pathways deliver proteins to the SecYEG translocon for translocation. Bacteria may have a single plasma membrane (Gram-positive bacteria), or an inner membrane plus an outer membrane separated by the periplasm (Gram-negative bacteria). Besides the plasma membrane the majority of prokaryotes lack membrane-bound organelles as found in eukaryotes, but they may assemble proteins onto various types of inclusions such as gas vesicles and storage granules.

Bioinformatics: Identifying protein targeting motifs in proteins **Minimotif Miner** is a bioinformatics tool that searches protein sequence queries for a known protein targeting sequence motifs.

Organelles surrounded by lipid bilayers and intracellular trafficking in eukaryotic cells. The transcribed RNAs are exported from the nucleus to the cytoplasm. After translation, proteins have to be targeted to their destination. The mechanism of protein delivery varies depending on the destination organelle. Proteins and RNAs are efficiently transported through the NPC, bi-directionally. NPC; nuclear pore complex, ER; endoplasmic reticulum.


peroxisome
Golgi
ER
NPC
nucleolus
RNAs
nucleus
cytoplasm
Protein
ribosome
mitochondria