PROTEIN LOCALIZATION

2.8.1 Introduction

Proteins are synthesized in two types of location:

- The vast majority of proteins are synthesized by ribosomes in the cytosol.
- A small minority are synthesized by ribosomes within organelles (mitochondria or chloroplasts).

Proteins synthesized in the cytosol can be divided into two general classes with regard to localization: those that are not associated with membranes; and those that are associated with membranes (see *Molecular Biology Supplement 32.5 Membranes and membrane proteins*). **Figure 8.1** maps the cell in terms of the possible ultimate destinations for a newly synthesized protein and the systems that transport it:

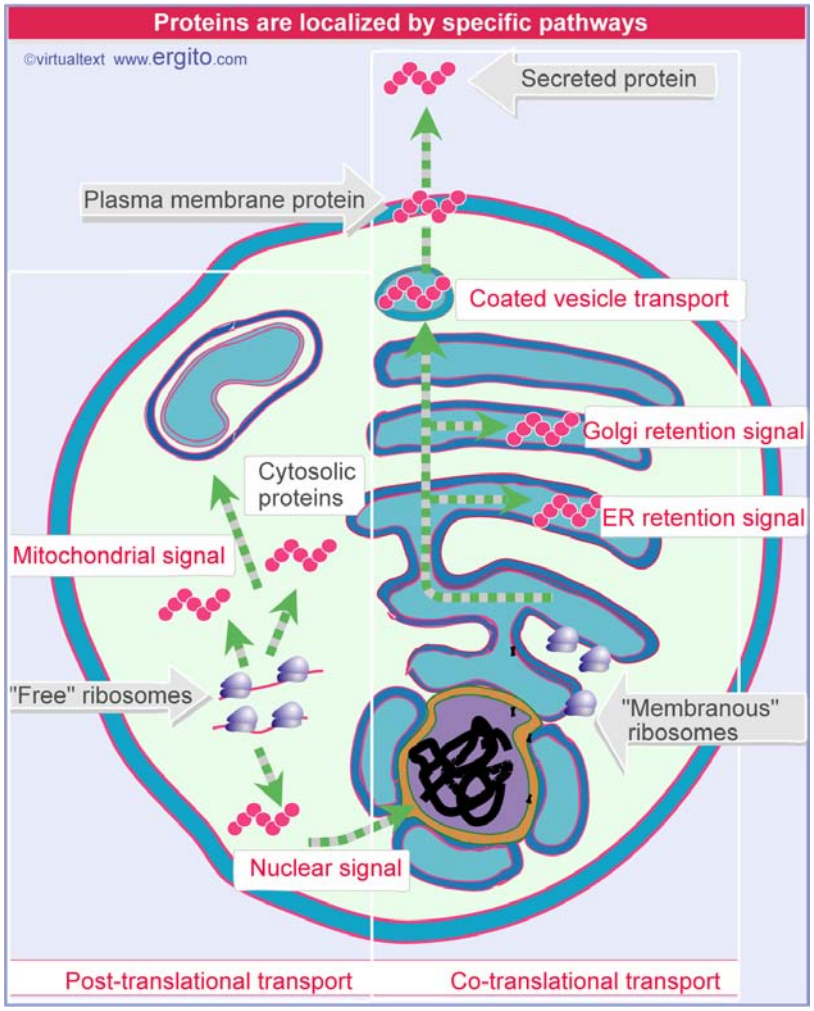


Figure 8.1 Overview: proteins that are localized post-translationally are released into the cytosol after synthesis on free ribosomes. Some have signals for targeting to organelles such as the nucleus or mitochondria. Proteins that are localized cotranslationally associate with the ER membrane during synthesis, so their ribosomes are "membrane-bound". The proteins pass into the endoplasmic reticulum, along to the Golgi, and then through the plasma membrane, unless they have signals that cause retention at one of the steps on the pathway. They may also be directed to other organelles, such as endosomes or lysosomes.

- Cytosolic (or "soluble") proteins are not localized in any particular organelle. They are synthesized in the cytosol, and remain there, where they function as individual catalytic centers, acting on metabolites that are in solution in the cytosol.
- Macromolecular structures may be located at particular sites in the cytoplasm; for example, centrioles are associated with the regions that become the poles of the mitotic spindle.
- Nuclear proteins must be transported from their site of synthesis in the cytosol through the nuclear envelope into the nucleus.

- Most of the proteins in cytoplasmic organelles are synthesized in the cytosol and transported specifically to (and through) the organelle membrane, for example, to the mitochondrion or peroxisome or (in plant cells) to the chloroplast. (Those proteins that are synthesized within the organelle remain within it.)

- The cytoplasm contains a series of membranous bodies, including endoplasmic reticulum (ER), Golgi apparatus, endosomes, and lysosomes. This is sometimes referred to as the "reticuloendothelial system." Proteins that reside within these compartments are inserted into ER membranes, and then are directed to their particular locations by the transport system of the Golgi apparatus. (For an introduction see *Molecular Biology Supplement 32.6 ER and Golgi*).

- Proteins that are secreted from the cell are transported to the plasma membrane and then must pass through it to the exterior. They start their synthesis in the same way as proteins associated with the reticuloendothelial system, but pass entirely through the system instead of halting at some particular point within it.

PROTEIN LOCALIZATION

2.8.2 Passage across a membrane requires a special apparatus

Key Terms

Protein **translocation** describes the movement of a protein across a membrane. This occurs across the membranes of organelles in eukaryotes, or across the plasma membrane in bacteria. Each membrane across which proteins are translocated has a channel specialized for the purpose.

Key Concepts

- Proteins pass across membranes through specialized protein structures embedded in the membrane.
- Substrate proteins interact directly with the transport apparatus of the ER or mitochondria or chloroplasts, but require carrier proteins to interact with peroxisomes.
- A much larger and complex apparatus is required for transport into the nucleus.

The process of inserting into or passing through a membrane is called protein **translocation**. The same dilemma must be solved for every situation in which a protein passes through a membrane. The protein presents a hydrophilic surface, but the membrane is hydrophobic. Like oil and water, the two would prefer not to mix. The solution is to create a special structure in the membrane through which the protein can pass. There are three different types of arrangements for such structures.

The endoplasmic reticulum, mitochondria, and chloroplasts contain proteinaceous structures embedded in their membranes that allow proteins to pass through without contacting the surrounding hydrophobic lipids. **Figure 8.2** shows that a substrate protein binds directly to the structure, is transported by it to the other side, and then released.

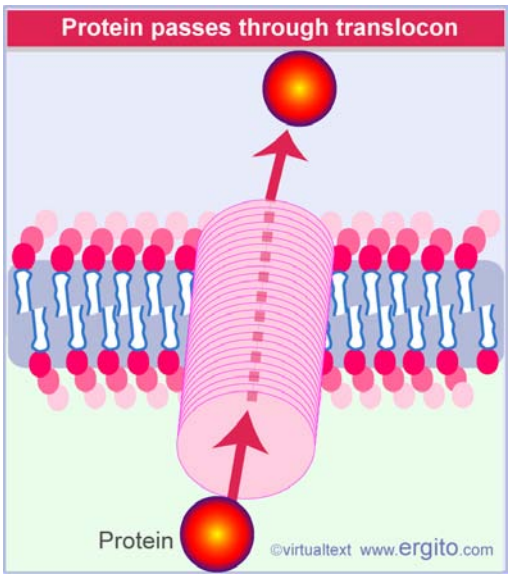


Figure 8.2 Proteins enter the ER or a mitochondrion by binding to a translocon that transports them across the membrane.

Peroxisomes also have such structures in their membranes, but the substrate proteins do not bind directly to them. **Figure 8.3** shows that instead they bind to carrier proteins in the cytosol, the carrier protein is transported through the channel into the peroxisome, and then the substrate protein is released.

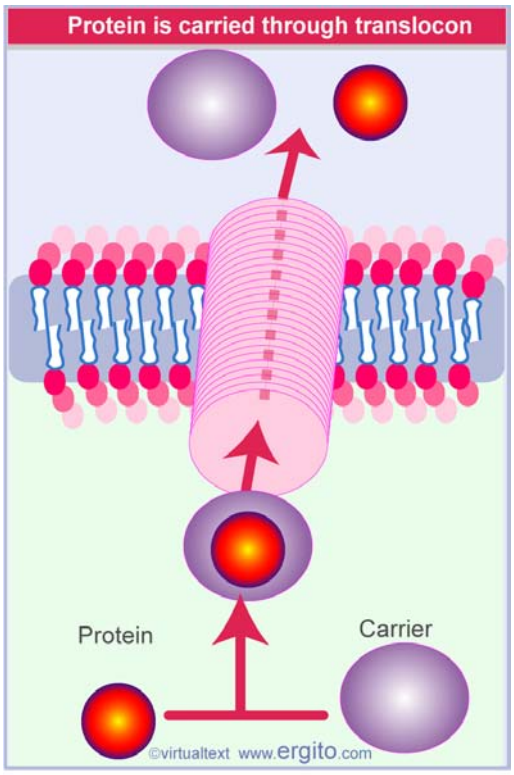


Figure 8.3 Proteins are transported into peroxisomes by a carrier protein that binds them in the cytosol, passes with them through the membrane channel, and releases them on the other side.

For transport into the nucleus, a much larger and more complex structure is employed. This is the nuclear pore. **Figure 8.4** shows that, although the pore provides the environment that allows a substrate to enter (or to leave) the nucleus, it does not actually provide the apparatus that binds to the substrate proteins and moves them through. Included in this apparatus are carrier proteins that bind to the substrates and transport them through the pore to the other side.

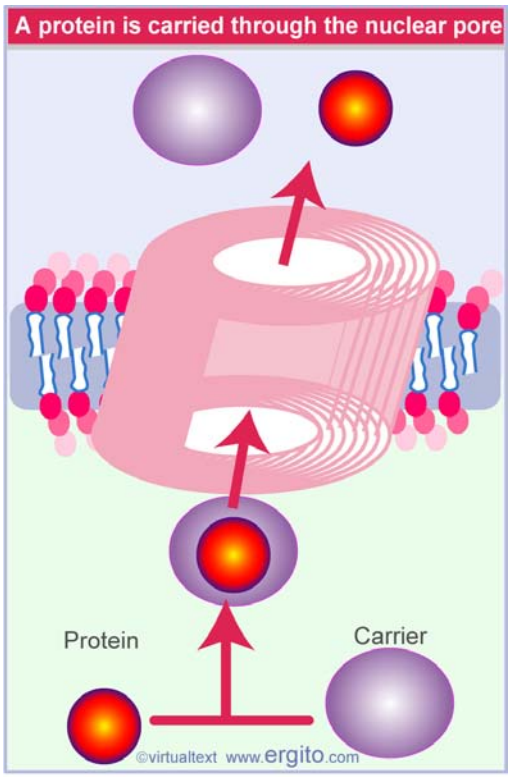


Figure 8.4 Proteins enter the nucleus by passage through very large nuclear pores. The transport apparatus is distinct from the pore itself and includes components that carry the protein through the pore.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.2.8.2

PROTEIN LOCALIZATION

2.8.3 Protein translocation may be post-translational or co-translational

Key Terms

The **leader** of a protein is a short N-terminal sequence responsible for initiating passage into or through a membrane.

A protein to be imported into an organelle or secreted from bacteria is called a "**preprotein**" until its signal sequence has been removed.

Key Concepts

- Proteins that are imported into cytoplasmic organelles are synthesized on free ribosomes in the cytosol.
- Proteins that are imported into the ER-Golgi system are synthesized on ribosomes that are associated with the ER.
- Proteins associate with membranes by means of specific amino acid sequences called signal sequences.
- Signal sequences are most often leaders that are located at the N-terminus.
- N-terminal signal sequences are usually cleaved off the protein during the insertion process.

There are two ways for a protein to make its initial contact with a membrane:

- The nascent protein may associate with the translocation apparatus while it is still being synthesized on the ribosome. This called **co-translational translocation**.
- The protein may be released from a ribosome after translation has been completed. Then the completed protein diffuses to the appropriate membrane and associates with the translocation apparatus. This is called **post-translational translocation**.

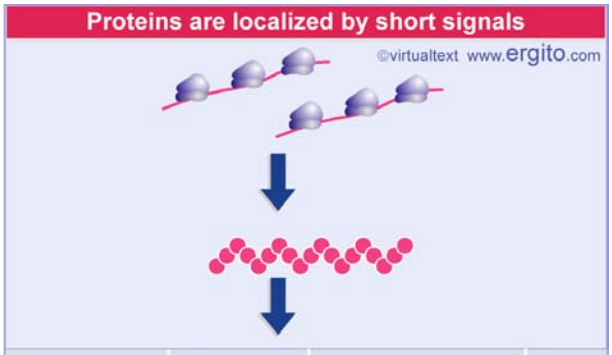
The location of a ribosome depends on whether the protein under synthesis is associating with a membrane co-translationally:

- Co-translational translocation is used for proteins that enter the endoplasmic reticulum. The consequence of this association is that the ribosome is localized to the surface of the endoplasmic reticulum. Because the ribosomes are associated with the ER membranes during synthesis of these proteins, and are therefore found in membrane fractions of the cell, they are sometimes described as "membrane-bound".

- All other ribosomes are located in the cytosol; because they are not associated with any organelle, and fractionate separately from membranes, they are sometimes called "free ribosomes". The free ribosomes synthesize all proteins except those that are translocated co-translationally. The proteins are released into the cytosol when their synthesis is completed. Some of these proteins remain free in the cytosol in quasi-soluble form; others associate with macromolecular cytosolic structures, such as filaments, microtubules, centrioles, etc., or are transported to the nucleus, or associate with membrane-bound organelles by post-translational translocation.

To associate with a membrane (or any other type of structure), a protein requires an appropriate signal, typically a sequence motif that causes it to be recognized by a translocation system (or to be assembled into a macromolecular structure).

Figure 8.5 summarizes some signals used by proteins released from cytosolic ribosomes. Import into the nucleus results from the presence of a variety of rather short sequences within proteins. These "nuclear localization signals" enable the proteins to pass through nuclear pores. One type of signal that determines transport to the peroxisome is a very short C-terminal sequence. Mitochondrial and chloroplast proteins are synthesized on "free" ribosomes; after their release into the cytosol they associate with the organelle membranes by means of N-terminal sequences of ~25 amino acids in length that are recognized by receptors on the organelle envelope.



Organelle	Signal location	Type	Signal length
Mitochondrion	N-terminal	Amphipathic helix	12-30
Chloroplast	N-terminal	Charged	>25
Nucleus	Internal	Basic or bipartite	4-9
Peroxisome	C-terminal	Short tetrapeptide	3-4

Figure 8.5 Proteins synthesized on free ribosomes in the cytosol are directed after their release to specific destinations by short signal motifs.

Proteins that reside within the reticuloendothelial system enter the endoplasmic reticulum while they are being synthesized. The principle of co-translational translocation is summarized in **Figure 8.6**. An important feature of this system is that the nascent protein is responsible for recognizing the translocation apparatus. This requires the signal for co-translational translocation to be part of the protein that is first synthesized, and, in fact, it is usually located at the N-terminus.

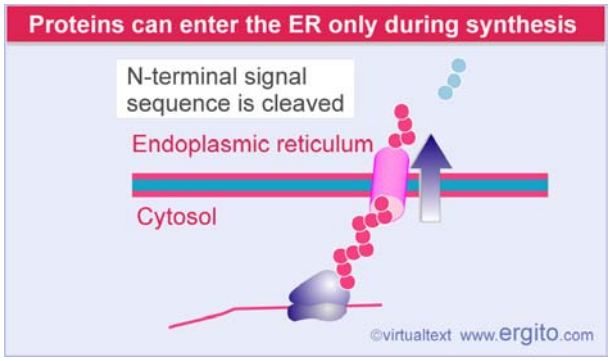


Figure 8.6 Proteins can enter the ER-Golgi pathway only by associating with the endoplasmic reticulum while they are being synthesized.

A common feature is found in proteins that use N-terminal sequences to be transported co-translationally to the ER or post-translationally to mitochondria or chloroplasts. The N-terminal sequence comprises a **leader** that is not part of the mature protein. The protein carrying this leader is called a **preprotein**, and is a transient precursor to the mature protein. The leader is cleaved from the protein during protein translocation.

Last updated on 2-20-2002

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.2.8.3

PROTEIN LOCALIZATION

2.8.4 Chaperones may be required for protein folding

Key Terms

Self-assembly refers to the ability of a protein (or of a complex of proteins) to form its final structure without the intervention of any additional components (such as chaperones). The term can also refer to the spontaneous formation of any biological structure that occurs when molecules collide and bind to each other.

Chaperones are a class of proteins which bind to incompletely folded or assembled proteins in order to assist their folding or prevent them from aggregating.

Key Concepts

- Proteins that can acquire their conformation spontaneously are said to self-assemble.
- Proteins can often assemble into alternative structures.
- A chaperone directs a protein into one particular pathway by excluding alternative pathways.
- Chaperones prevent the formation of incorrect structures by interacting with unfolded proteins to prevent them from folding incorrectly

Some proteins are able to acquire their mature conformation spontaneously. A test for this ability is to denature the protein and determine whether it can then renature into the active form. This capacity is called **self-assembly**. A protein that can self-assemble can fold or refold into the active state from other conformations, including the condition in which it is initially synthesized. This implies that the internal interactions are intrinsically directed toward the right conformation. The classic case is that of ribonuclease; it was shown in the 1970s that, when the enzyme is denatured, it can renature *in vitro* into the correct conformation (1115). More recently the process of intrinsic folding has been described in detail for some small proteins.

When correct folding does not happen, and alternative sets of interactions can occur, a protein may become trapped in a stable conformation that is not the intended final form. Proteins in this category cannot self-assemble. Their acquisition of proper structure requires the assistance of a **chaperone**. (For a description of the discovery that proteins require assistance to fold *in vivo*, see *Great Experiments 5.2 The discovery of protein folding by chaperonins*. For an introduction to the other activities that are involved, see *Molecular Biology Supplement 32.4 Protein folding*.)

Protein folding takes place by interactions between reactive surfaces. Typically these surfaces consist of exposed hydrophobic side chains. Their interactions form a hydrophobic core. The intrinsic reactivity of these surfaces means that incorrect

interactions may occur unless the process is controlled. **Figure 8.7** illustrates what would happen. As a newly synthesized protein emerges from the ribosome, any hydrophobic patch in the sequence is likely to aggregate with another hydrophobic patch. Such associations are likely to occur at random and therefore will probably not represent the desired conformation of the protein. (For an account of the molecular interactions involved in folding see 2389).

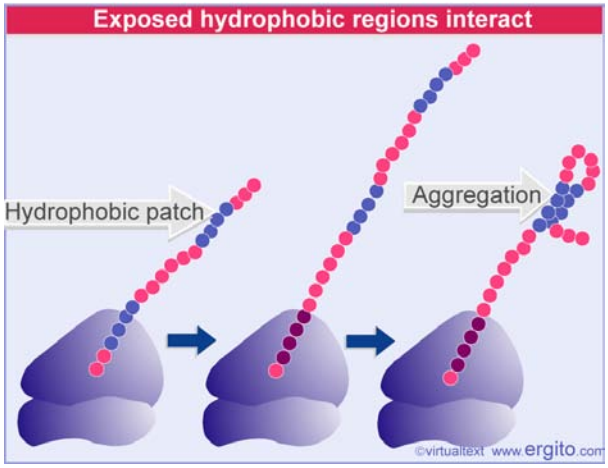


Figure 8.7 Hydrophobic regions of proteins are intrinsically interactive, and will aggregate with one another when a protein is synthesized (or denatured) unless prevented.

Chaperones are proteins that mediate correct assembly by causing a target protein to acquire one possible conformation instead of others (for review see 50; 2388). This is accomplished by binding to reactive surfaces in the target protein that are exposed during the assembly process, and preventing those surfaces from interacting with other regions of the protein to form an incorrect conformation. Chaperones function by preventing formation of incorrect structures rather than by promoting formation of correct structures. **Figure 8.8** shows an example in which a chaperone in effect sequesters a hydrophobic patch, allowing interactions to occur that would not have been possible in its presence, as can be seen by comparing the result with **Figure 8.7**.

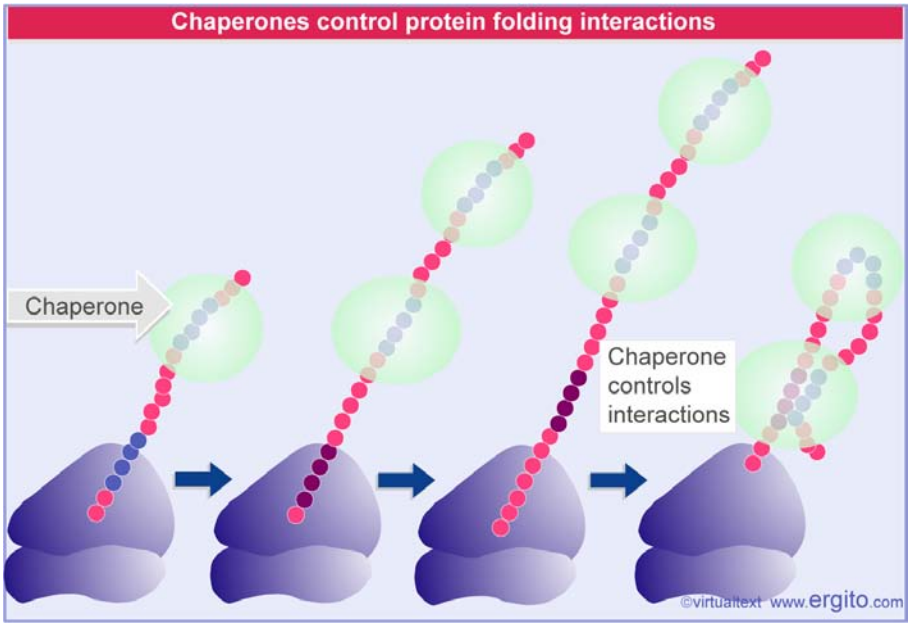


Figure 8.8 Chaperones bind to interactive regions of proteins as they are synthesized to prevent random aggregation. Regions of the protein are released to interact in an orderly manner to give the proper conformation.

An incorrect structure may be formed either by misfolding of a single protein or by interactions with another protein. The density of proteins in the cytosol is high, and "macromolecular crowding" can increase the efficiencies of many reactions compared to the rates observed *in vitro*. Crowding can cause folding proteins to aggregate, but chaperones can counteract this effect (2336). So one role of chaperones may be to protect a protein so that it can fold without being adversely affected by the crowded conditions in the cytosol.

We do not know what proportion of proteins can self-assemble as opposed to those that require assistance from a chaperone. (It is not axiomatic that a protein capable of self-assembly *in vitro* actually self-assembles *in vivo*, because there may be rate differences in the two conditions, and chaperones still could be involved *in vivo*. However, there is a distinction to be drawn between proteins that can in principle self-assemble and those that in principle must have a chaperone to assist acquisition of the correct structure.)

Last updated on 3-20-2002

Reviews

50. Ellis, R. J. and van der Vies, S. M. (1991). *Molecular chaperones.* Annu. Rev. Biochem. 60, 321-347.

2388. Hartl, F. U. and Hayer-Hartl, M. (2002). *Molecular chaperones in the cytosol: from nascent chain to folded protein.* Science 295, 1852-1858.

2389. Fersht, A. R. and Daggett, V. (2002). *Protein folding and unfolding at atomic resolution.* Cell 108, 573-582.

References

1115. Anfinsen, C. B. (1973). *Principles that govern the folding of protein chains.* Science 181, 223-230.

2336. van den Berg, B., Ellis, R. J., and Dobson, C. M. (1999). *Effects of macromolecular crowding on protein folding and aggregation.* EMBO J. 18, 6927-6933.

PROTEIN LOCALIZATION

2.8.5 Chaperones are needed by newly synthesized and by denatured proteins

Key Concepts

- Chaperones act on newly synthesized proteins, proteins that are passing through membranes, or proteins that have been denatured.
- Hsp70 and some associated proteins form a major class of chaperones that act on many target proteins.
- Group I and group II chaperonins are large oligomeric assemblies that act on target proteins they sequester in internal cavities.
- Hsp90 is a specialized chaperone that acts on proteins of signal transduction pathways.

The ability of chaperones to recognize incorrect protein conformations allows them to play two related roles concerned with protein structure:

- When a protein is initially synthesized, that is to say, as it exits the ribosome to enter the cytosol, it appears in an unfolded form. Spontaneous folding then occurs as the emerging sequence interacts with regions of the protein that were synthesized previously. Chaperones influence the folding process by controlling the accessibility of the reactive surfaces. This process is involved in initial acquisition of the correct conformation.
- When a protein is denatured, new regions are exposed and become able to interact. These interactions are similar to those that occur when a protein (transiently) misfolds as it is initially synthesized. They are recognized by chaperones as comprising incorrect folds. This process is involved in recognizing a protein that has been denatured, and either assisting renaturation or leading to its removal by degradation.

Chaperones may also be required to assist the formation of oligomeric structures and for the transport of proteins through membranes. A persistent theme in membrane passage is that control (or delay) of protein folding is an important feature. **Figure 8.9** shows that it may be necessary to maintain a protein in an unfolded state before it enters the membrane because of the geometry of passage: the mature protein could simply be too large to fit into the available channel. Chaperones may prevent a protein from acquiring a conformation that would prevent passage through the membrane; in this capacity, their role is basically to maintain the protein in an unfolded, flexible state. Once the protein has passed through the membrane, it may require another chaperone to assist with folding to its mature conformation in much the same way that a cytosolic protein requires assistance from a chaperone as it emerges from the ribosome. The state of the protein as it emerges from a membrane is probably similar to that as it emerges from the ribosome – basically extended in a

more or less linear condition.

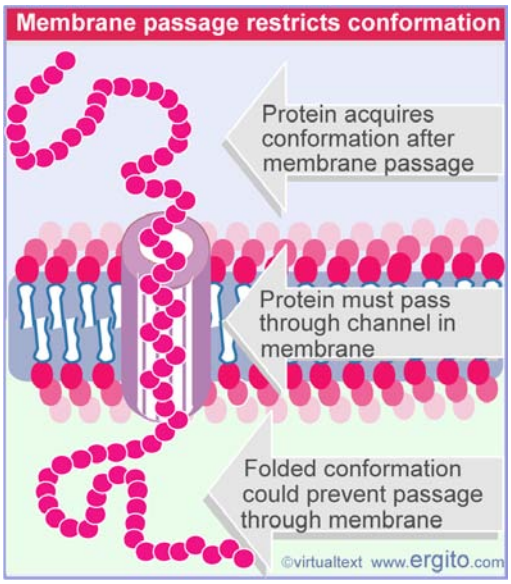


Figure 8.9 A protein is constrained to a narrow passage as it crosses a membrane.

Two major types of chaperones have been well characterized (for review see 2284). They affect folding through two different types of mechanism:

- **Figure 8.10** shows that the *Hsp70 system consists of individual proteins that bind to, and act on, the substrates whose folding is to be controlled.* It recognizes proteins as they are synthesized or emerge from membranes (and also when they are denatured by stress). Basically it controls the interactions between exposed reactive regions of the protein, enabling it to fold into the correct conformation *in situ*. The components of the system are Hsp70, Hsp40, and GrpE. The name of the system reflects the original identification of Hsp70 as a protein induced by heat shock. The Hsp70 and Hsp40 proteins bind individually to the substrate proteins. They use hydrolysis of ATP to provide the energy for changing the structure of the substrate protein, and work in conjunction with an exchange factor that regenerates ATP from ADP.

- **Figure 8.11** shows that *a chaperonin system consists of a large oligomeric assembly (represented as a cylinder).* This assembly forms a structure into which unfolded proteins are inserted. The protected environment directs their folding. There are two types of chaperonin system. GroEL/GroES is found in all classes of organism. TRiC is found in eukaryotic cytosol.

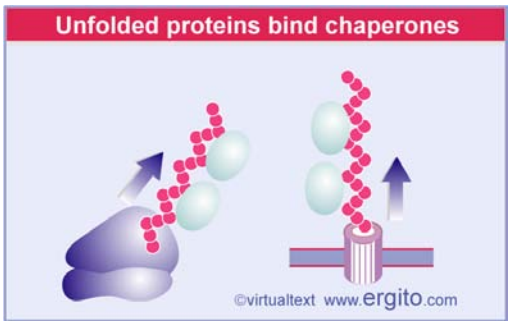


Figure 8.10 Proteins emerge from the ribosome or from passage through a membrane in an unfolded state that attracts chaperones to bind and protect them from misfolding.

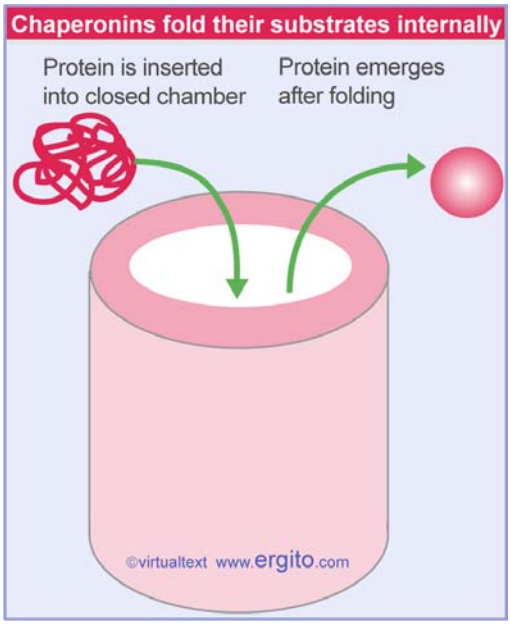


Figure 8.11 A chaperonin forms a large oligomeric complex and folds a substrate protein within its interior.

The components of the systems are summarized in **Figure 8.12**. The Hsp70 system and the chaperonin systems both act on many different substrate proteins. Another system, the Hsp90 protein, functions in conjunction with Hsp70, but is directed against specific classes of proteins that are involved in signal transduction, especially the steroid hormone receptors and signaling kinases (for review see 2506). Its basic function is to maintain its targets in an appropriate conformation until they are stabilized by interacting with other components of the pathway (2507; 2508).

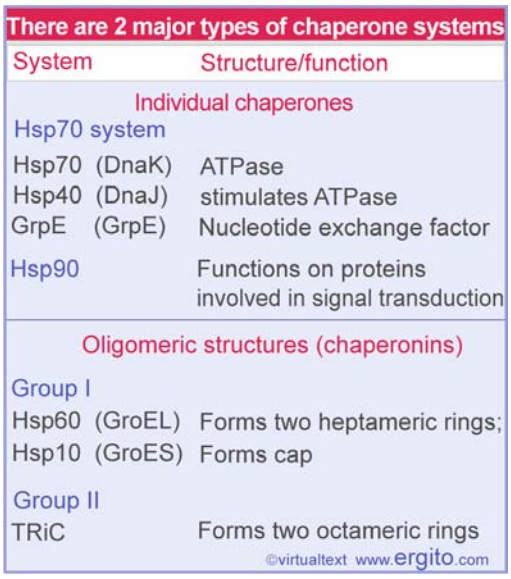


Figure 8.12 Chaperone families have eukaryotic and bacterial counterparts (named in parentheses).

(The reason many of these proteins are named "hsp", which stands for "heat shock protein" is that increase in temperature causes production of heat shock proteins whose function is to minimize the damage caused to proteins by heat denaturation. Many of the heat shock proteins are chaperones and were first discovered, and named, as part of the heat shock response.)

Last updated on 6-20-2002

Reviews

2284. Frydman, J. (2001). *Folding of newly translated proteins in vitro: the role of molecular chaperones.* Annu. Rev. Biochem. 70, 603-647.

2506. Moarefi, I. and Hartl, F. U. (2001). *Hsp90: a specialized but essential protein-folding tool.* J. Cell Biol. 154, 267-273.

References

2507. Rutherford, S. L. and Lindquist, S. (1998). *Hsp90 as a capacitor for morphological evolution.* Nature 396, 336-342.

2508. Queitsch, C., Sangster, T. A., and Lindquist, S. (2002). *Hsp90 as a capacitor of phenotypic variation.* Nature 417, 618-624.

PROTEIN LOCALIZATION

2.8.6 The Hsp70 family is ubiquitous

Key Concepts

- Hsp70 is a chaperone that functions on target proteins in conjunction with DnaJ and GrpE.
- Members of the Hsp70 family are found in the cytosol, in the ER, and in mitochondria and chloroplasts.

The Hsp70 family is found in bacteria, eukaryotic cytosol, in the endoplasmic reticulum, and in chloroplasts and mitochondria (for review see 2284). A typical Hsp70 has two domains: the N-terminal domain is an ATPase; and the C-terminal domain binds the substrate polypeptide (2337; 2338). When bound to ATP, Hsp70 binds and releases substrates rapidly; when bound to ADP, the reactions are slow. Recycling between these states is regulated by two other proteins, Hsp40 (DnaJ) and GrpE.

Figure 8.13 shows that Hsp40 (DnaJ) binds first to a nascent protein as it emerges from the ribosome. Hsp40 contains a region called the J domain (named for DnaJ), which interacts with Hsp70. Hsp70 (DnaK) binds to both Hsp40 and to the unfolded protein. In effect, two interacting chaperones bind to the protein. The J domain accounts for the specificity of the pairwise interaction, and drives a particular Hsp40 to select the appropriate partner from the Hsp70 family.

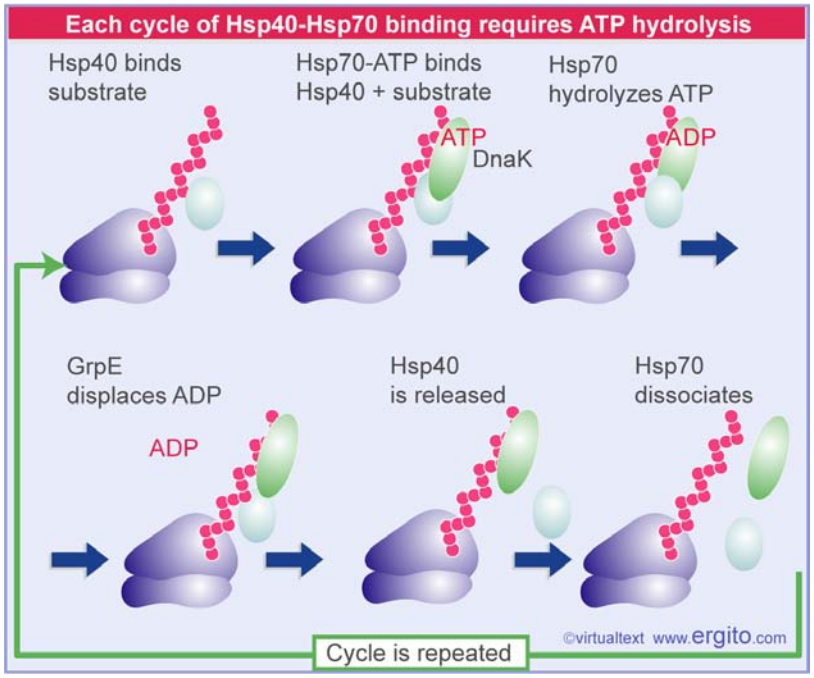


Figure 8.13 Hsp40 binds the substrate and then Hsp70. ATP hydrolysis drives conformational change. GrpE displaces the ADP; this causes the chaperones to be released. Multiple cycles of association and dissociation may occur during the folding of a substrate protein.

The interaction of Hsp70 (DnaK) with Hsp40 (DnaJ) stimulates the ATPase activity of Hsp70. The ADP-bound form of the complex remains associated with the protein substrate until GrpE displaces the ADP. This causes loss of Hsp40 followed by dissociation of Hsp70. The Hsp70 binds another ATP and the cycle can be repeated. GrpE (or its equivalent) is found only in bacteria, mitochondria, and chloroplasts; in other locations, the dissociation reaction is coupled to ATP hydrolysis in a more complex way.

Protein folding is accomplished by multiple cycles of association and dissociation. As the protein chain lengthens, Hsp70 (DnaK) may dissociate from one binding site and then reassociate with another, thus releasing parts of the substrate protein to fold correctly in an ordered manner. Finally, the intact protein is released from the ribosome, folded into its mature conformation (for review see 54; 63; 66).

Different members of the Hsp70 class function on various types of target proteins. Cytosolic proteins (the eponymous Hsp70 and a related protein called Hsc70) act on nascent proteins on ribosomes. Variants in the ER (called BiP or Grp78 in higher eukaryotes, called Kar2 in *S. cerevisiae*), or in mitochondria or chloroplasts, function in a rather similar manner on proteins as they emerge into the interior of the organelle on passing through the membrane.

What feature does Hsp70 recognize in a target protein? It binds to a linear stretch of amino acids embedded in a hydrophobic context (2339; 2340). This is precisely the sort of motif that is buried in the hydrophobic core of a properly folded, mature protein. Its exposure therefore indicates that the protein is nascent or denatured. Motifs of this nature occur about every 40 amino acids. Binding to the motif prevents

it from misaggregating with another one.

This mode of action explains how the Hsp70 protein Bip can fulfill two functions: to assist in oligomerization and/or folding of newly translocated proteins in the ER; and to remove misfolded proteins. Suppose that BiP recognizes certain peptide sequences that are inaccessible within the conformation of a mature, properly folded protein. These sequences are exposed and attract BiP when the protein enters the ER lumen in an essentially one-dimensional form. And if a protein is misfolded or denatured, they may become exposed on its surface instead of being properly buried.

Last updated on 2-12-2002

Reviews

54. Georgopoulos, C. and Welch, W. J. (1993). *Role of the major heat shock proteins as molecular chaperones.* Annu. Rev. Cell Biol. 9, 601-634.

63. Hartl, F.-U. (1966). *Molecular chaperones in cellular protein folding*. Nature 381, 571-580.

66. Bukau, B. and Horwich, A. L. (1998). *The Hsp70 and Hsp60 chaperone machines*. Cell 92, 351-366.

2284. Frydman, J. (2001). *Folding of newly translated proteins in vitro: the role of molecular chaperones.* Annu. Rev. Biochem. 70, 603-647.

References

2337. Flaherty, K. M., DeLuca-Flaherty, C., and McKay, D. B. (1990). *Three-dimensional structure of the ATPase fragment of a 70K heat-shock cognate protein.* Nature 346, 623-628.

2338. Zhu, X., Zhao, X., Burkholder, W. F., Gragerov, A., Ogata, C. M., Gottesman, M. E., Hendrickson, and, Gottesman, M. E. (1996). *Structural analysis of substrate binding by the molecular chaperone DnaK.* Science 272, 1606-1614.

2339. Flynn, G. C., Pohl, J., Flocco, M. T., and Rothman, J. E. (1991). *Peptide-binding specificity of the molecular chaperone BiP.* Nature 353, 726-730.

2340. Blond-Elguindi, S., Cwirla, S. E., Dower, W. J., Lipshutz, R. J., Sprang, S. R., Sambrook, J. F., and Gething, M. J. (1993). *Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of BiP.* Cell 75, 717-728.

PROTEIN LOCALIZATION

2.8.7 Hsp60/GroEL forms an oligomeric ring structure

Key Concepts

- Hsp60/GroEL forms an oligomeric structure consisting of 14 subunits arranged in two inverted heptameric rings.
- A GroES heptamer forms a dome that caps one end of the double ring.
- A substrate protein undergoes a cycle of folding in the cavity of one of the Hsp60/GroEL rings. It is released and rebound for further cycles until it reaches mature conformation.
- Hydrolysis of ATP provides energy for the folding cycles.

Large (oligomeric) structures with hollow cavities are often used for handling the folding or degradation of proteins (for review see 968). The typical structure is a ring of many subunits, forming a doughnut or cylinder. **Figure 8.14** shows that the target protein is in effect placed in a controlled environment – the cavity – where it is closely associated with the surrounding protein. This creates a high local concentration of binding sites and supports cooperative interactions. In the case of folding, the closed environment prevents the target protein from forming wrongful interactions with other proteins, which may be important in driving folding along the proper pathway. In the case of degradation, isolation presumably makes for a more controlled process than would be possible in open cytosol (see *Molecular Biology Supplement 8.32 The proteasome is a large machine that degrades ubiquitinated proteins*). The energy for these processes is provided by hydrolysis of ATP – typically the subunits of the ring are ATPases.

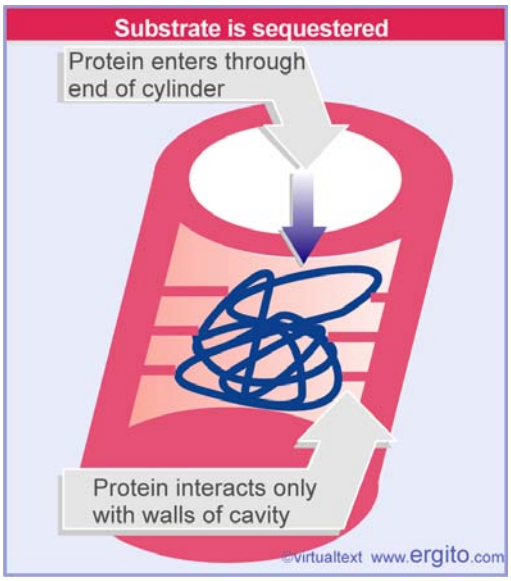


Figure 8.14 A protein may be sequestered within a controlled environment for folding or degradation.

The Hsp60 class of chaperones forms a large apparatus that consists of two types of subunit. **Figure 8.15** illustrates the structure schematically. Hsp60 itself (known as GroEL in *E. coli*) forms a structure consisting of 14 subunits that are arranged in two heptameric rings stacked on top of each other in inverted orientation. This means that the top and bottom surfaces of the double ring are the same. The central hole is blocked at the equator of each ring by the COOH ends of the subunits, which protrude into the interior. So the ends of the double cylinder form symmetrical cavities extending about half way into each unit.

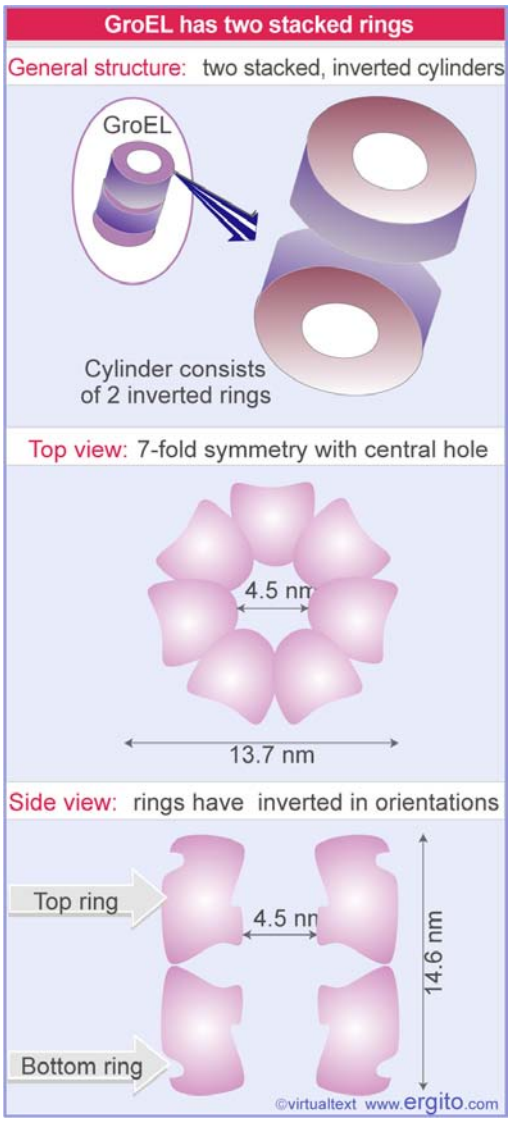


Figure 8.15 GroEL forms an oligomer of two rings, each comprising a hollow cylinder made of 7 subunits.

This structure associates with a heptamer formed of subunits of Hsp10 (GroES in *E. coli*). A single GroES heptamer forms a dome that associates with one surface of the double ring, as shown in **Figure 8.16**. The dome sits over the central cavity, thus capping one opening of the cylinder. The dome is hollow and in effect extends the cavity into the closed surface. We can distinguish the two rings of GroEL as the proximal ring (bound to GroES) or the distal ring (not bound to GroES). The entire GroEL/GroES structure has a mass ~10^6 daltons, comparable to a small ribosomal subunit. GroEL is sometimes called a chaperonin, and GroES is called a co-chaperonin, because GroEL plays the essential role in guiding folding, but GroES is required for its activity (459; 463; 465).

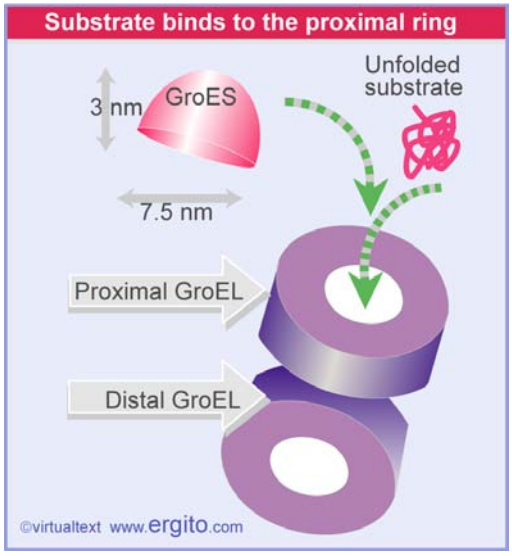


Figure 8.16 Two rings of GroEL associate back to back to form a hollow cylinder. GroES forms a dome that covers the central cavity on one side. Protein substrates bind to the cavity in the proximal ring.

GroEL binds to many unfolded proteins, probably by recognizing a condensed "molten globule" state. Interaction with the substrate is based on hydrophobic interactions between surfaces of the substrate and residues of GroEL that are exposed in its central cavity. Substrates may be provided by proteins that have become denatured; or they may be transferred to GroEL by other chaperones – for example, Hsp70 may assist a nascent protein in folding, but then passes it on to GroEL for the process to be completed when it is released from the ribosome.

The key reactions in substrate binding and folding are illustrated in **Figure 8.17**. The reaction starts when substrate and ATP are bound to the same ring of GroEL. This defines the proximal ring. Then GroES caps this ring. Binding of GroES induces a conformational change in the proximal GroEL ring, increasing the volume of the central cavity. This also changes the environment for the substrate. The hydrophobic residues in GroEL that had previously bound substrate are involved in binding to GroES. The result is that the substrate now finds itself in a hydrophilic environment that forces a change in its conformation (460; 464).

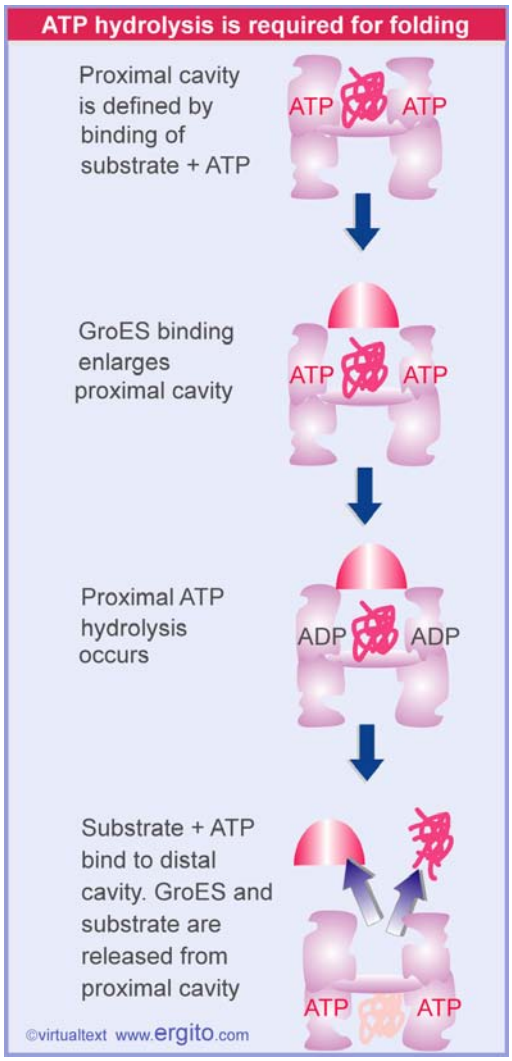


Figure 8.17 Protein folding occurs in the proximal GroEL ring and requires ATP. Release of substrate and GroES requires ATP hydrolysis in the distal ring. *This is a static version of an interactive figure; see http://www.ergito.com/main.jsp?bcs=MBIO.2.8.7 to view properly.*

ATP plays an important role in GroEL function. Each subunit of GroEL has a molecule of ATP. The presence of ATP on the subunits of the proximal ring is required for folding to occur. Hydrolysis is required for the transition to the next stage. Hydrolysis of the ATP in the proximal ring changes the properties of the distal ring in such a way as to allow substrate and ATP to bind to it. This in turn triggers the dissociation of the substrate and GroES from the proximal ring. Now the situation at the start of the cycle has been restored. The ring that was the distal ring in the previous cycle is bound to substrate and ATP, and becomes the proximal ring for the next cycle. So the rings of GroEL alternate as proximal and distal (466; 967).

An important question in the action of this (and other macromolecular) chaperones is whether their action is processive. Does a substrate enter the central cavity, undergo

multiple cycles of folding within it, and become released in mature form? Or does it undergo a single folding cycle, after which it is released; typically it will still have improperly folded regions, and therefore will be bound again for another folding cycle. This process will continue until the protein has reached a mature conformation that does not offer a substrate to the chaperone.

These models have been tested by using a mutant GroEL that can bind unfolded proteins but cannot release them. When this "trap GroEL" is added to wild-type GroEL that is actively engaged with a substrate, it blocks the appearance of mature protein. This suggests that the substrate has been released before folding was completed. The simplest explanation is that substrate protein is released after each folding cycle. One cycle of folding, ATP hydrolysis, and release takes about 15 sec *in vitro* (461; 462).

Last updated on 5-3-2000

Reviews

968. Horwich, A.L., Weber-Ban, E.U., and Finley, D. (1999). *Chaperone rings in protein folding and degradation.* Proc. Natl. Acad. Sci. USA 96, 11033-11040.

References

459. Braig, K. et al. (1994). *The crystal structure of the bacterial chaperonin GroEL at 28* Å. Nature 371, 578-586.

460. Chen, S. et al. (1994). *Location of a folding protein and shape changes in GroEL-GroES complexes imaged by cryo-electron microscopy*. Nature 371, 261-264.

461. Weissman, J. S. et al. (1994). *GroEL-mediated protein folding proceeds by multiple rounds of binding and release of nonnative forms*. Cell 78, 693-702.

462. Weissman, J. A. et al. (1995). *Mechanism of GroEL action: productive release of polypeptide from a sequestered position under GroES*. Cell 83, 577-587.

463. Hunt, J. F. et al. (1996). *The crystal structure of the GroES co-chaperonin at 28* Å *resolution*. Nature 379, 37-45.

464. Mayhew, M. et al. (1996). *Protein folding in the central cavity of the GroEL-GroES chaperonin complex*. Nature 379, 420-426.

465. Xu, Z., Horwich, A. L., and Sigler, P. B. (1997). *The crystal structure of the asymmetric GroEL-GroES-(ADP)7 chaperonin complex*. Nature 388, 741-750.

466. Rye, H. S. et al. (1997). *Distinct actions of cis and trans ATP within the double ring of the chaperonin GroEL*. Nature 388, 792-798.

967. Rye, H. S., Roseman, A. M., Chen, S., Furtak, K., Fenton, W. A., Saibil, H. R., and Horwich, A. L. (1999). *GroEL-GroES cycling: ATP and nonnative polypeptide direct alternation of folding-active rings.* Cell 97, 325-338.

PROTEIN LOCALIZATION

2.8.8 Signal sequences initiate translocation

Key Terms

Protein sorting (targeting) is the direction of different types of proteins for transport into or between specific organelles.

A **signal sequence** is a short region of a protein that directs it to the endoplasmic reticulum for co-translational translocation.

Key Concepts

- Proteins associate with the ER system only co-translationally.
- The signal sequence of the substrate protein is responsible for membrane association.

Proteins that associate with membranes via N-terminal leaders use a hierarchy of signals to find their final destination. In the case of the reticuloendothelial system, the ultimate location of a protein depends on how it is directed as it transits the endoplasmic reticulum and Golgi apparatus. The leader sequence itself introduces the protein to the membrane; the intrinsic consequence of the interaction is for the protein to pass through the membrane into the compartment on the other side. For a protein to reside within the membrane, a further signal is required to stop passage through the membrane. Other types of signals are required for a protein to be sorted to a particular destination, that is, to remain within the membrane or lumen of some particular compartment. The general process of finding its ultimate destination by transport through successive membrane systems is called **protein sorting** or **targeting**, and is discussed in *Molecular Biology 6.27 Protein trafficking*.

The overall nature of the pathway is summarized in **Figure 8.18**. The "default pathway" takes a protein through the ER, into the Golgi, and on to the plasma membrane. Proteins that reside in the ER possess a C-terminal tetrapeptide (KDEL, which actually provides a signal for them to return to the ER from the Golgi). The signal that diverts a protein to the lysosome is a covalent modification: the addition of a particular sugar residue. Other signals are required for a protein to become a permanent constituent of the Golgi or the plasma membrane. We discuss direction to these locations in *Molecular Biology 6.27 Protein trafficking*.

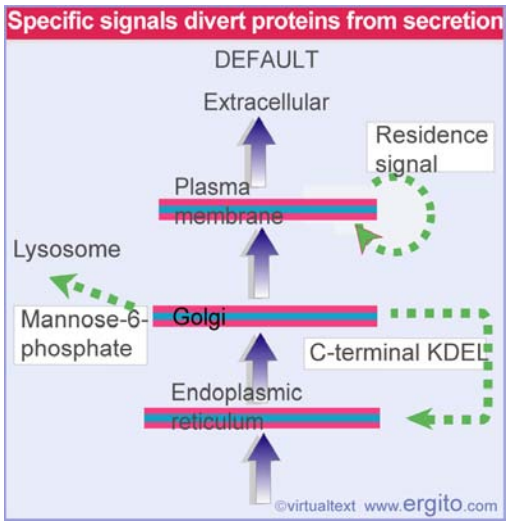


Figure 8.18 Proteins that enter the ER-Golgi pathway may flow through to the plasma membrane or may be diverted to other destinations by specific signals.

There is a common starting point for proteins that associate with, or pass through, the reticuloendothelial system of membranes. *These proteins can associate with the membrane only while they are being synthesized.* The ribosomes synthesizing these proteins become associated with the endoplasmic reticulum, enabling the nascent protein to be co-translationally transferred to the membrane. Regions in which ribosomes are associated with the ER are sometimes called the "rough ER," in contrast with the "smooth ER" regions that lack associated polysomes and which have a tubular rather than sheet-like appearance (for review see 1945). **Figure 8.19** shows ribosomes in the act of transferring nascent proteins to ER membranes.

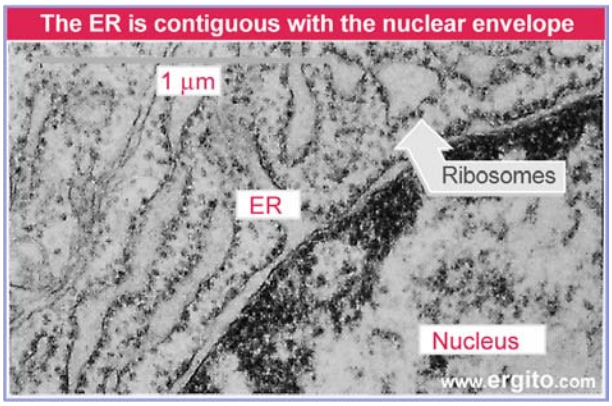


Figure 8.19 The endoplasmic reticulum consists of a highly folded sheet of membranes that extends from the nucleus. The small objects attached to the outer surface of the membranes are ribosomes. Photograph kindly provided by Lelio Orci.

The proteins synthesized at the rough endoplasmic reticulum pass from the ribosome directly to the membrane. Then they are transferred to the Golgi apparatus, and finally are directed to their ultimate destination, such as the lysosome or secretory

vesicle or plasma membrane. The process occurs within a membranous environment as the proteins are carried between organelles in small membrane-coated vesicles (see *Molecular Biology 6.27 Protein trafficking*.)

Co-translational insertion is directed by a **signal sequence**. Usually this is a cleavable leader sequence of 15-30 N-terminal amino acids. At or close to the N-terminus are several polar residues, and within the leader is a hydrophobic core consisting exclusively or very largely of hydrophobic amino acids (for review see 2356). There is no other conservation of sequence. **Figure 8.20** gives an example.

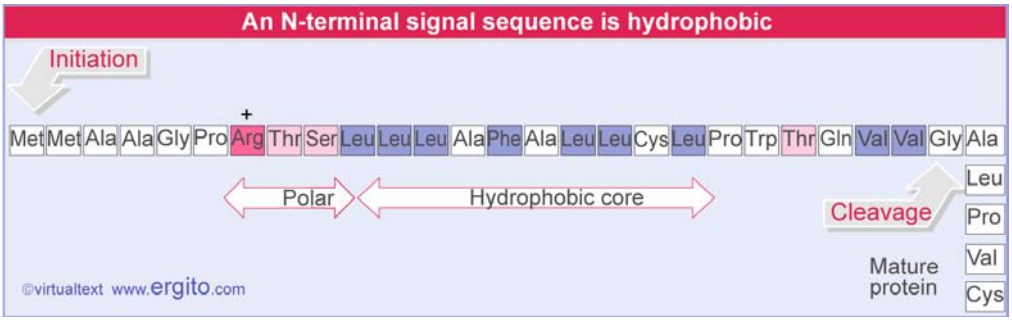


Figure 8.20 The signal sequence of bovine growth hormone consists of the N-terminal 29 amino acids and has a central highly hydrophobic region, preceded or flanked by regions containing polar amino acids.

The signal sequence is both necessary and sufficient to sponsor transfer of any attached polypeptide into the target membrane. A signal sequence added to the N-terminus of a globin protein, for example, causes it to be secreted through cellular membranes instead of remaining in the cytosol (2790).

The signal sequence provides the connection that enables the ribosomes to attach to the membrane. There is no intrinsic difference between free ribosomes (synthesizing proteins in the cytosol) and ribosomes that are attached to the ER. A ribosome starts synthesis of a protein without knowing whether the protein will be synthesized in the cytosol or transferred to a membrane. It is the synthesis of a signal sequence that causes the ribosome to associate with a membrane (467; for review see 47).

Last updated on 2-20-2002

Reviews

47. Lee, C. and Beckwith, J. (1986). *Cotranslational and posttranslational protein translocation in prokaryotic systems.* Annu. Rev. Cell Biol. 2, 315-336.

1945. Palade, G. (1975). *Intracellular aspects of the process of protein synthesis.* Science 189, 347-358.

References

467. Blobel, G. and Dobberstein, B. (1975). *Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma*. J. Cell Biol. 67, 835-851.

2356. von Heijne, G. (1985). *Signal sequences. The limits of variation.* J. Mol. Biol. 184, 99-105.

2790. Lingappa, V. R., Chaidez, J., Yost, C. S., and Hedgpeth, J. (1984). *Determinants for protein localization: beta-lactamase signal sequence directs globin across microsomal membranes.* Proc. Natl. Acad. Sci. USA 81, 456-460.

PROTEIN LOCALIZATION

2.8.9 The signal sequence interacts with the SRP

Key Terms

Protein **translocation** describes the movement of a protein across a membrane. This occurs across the membranes of organelles in eukaryotes, or across the plasma membrane in bacteria. Each membrane across which proteins are translocated has a channel specialized for the purpose.

The **signal recognition particle (SRP)** is a ribonucleoprotein complex that recognizes signal sequences during translation and guides the ribosome to the translocation channel. SRPs from different organisms may have different compositions, but all contain related proteins and RNAs.

Signal peptidase is an enzyme within the membrane of the ER that specifically removes the signal sequences from proteins as they are translocated. Analogous activities are present in bacteria, archaebacteria, and in each organelle in a eukaryotic cell into which proteins are targeted and translocated by means of removable targeting sequences. Signal peptidase is one component of a larger protein complex.

Key Concepts

- The signal sequence binds to the SRP (signal recognition particle).
- Signal-SRP binding causes protein synthesis to pause.
- Protein synthesis resumes when the SRP binds to the SRP receptor in the membrane.
- The signal sequence is cleaved from the translocating protein by the signal peptidase located on the "inside" face of the membrane.

Protein **translocation** can be divided into two general stages: first ribosomes carrying nascent polypeptides associate with the membranes; and then the nascent chain is transferred to the channel and translocates through it.

The attachment of ribosomes to membranes requires the **signal recognition particle** (**SRP**). The SRP has two important abilities:

- It can bind to the signal sequence of a nascent secretory protein.
- And it can bind to a protein (the SRP receptor) located in the membrane.

The SRP and SRP receptor function catalytically to transfer a ribosome carrying a nascent protein to the membrane. The first step is the recognition of the signal sequence by the SRP. Then the SRP binds to the SRP receptor and the ribosome binds to the membrane. The stages of translation of membrane proteins are

summarized in **Figure 8.21**.

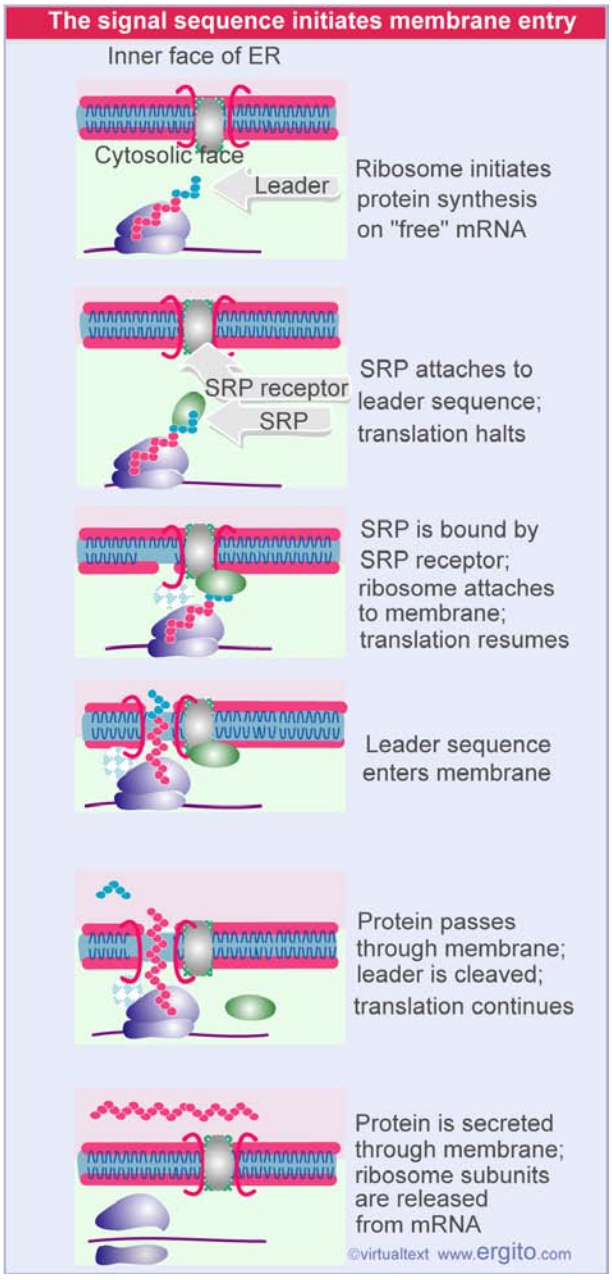


Figure 8.21 Ribosomes synthesizing secretory proteins are attached to the membrane via the signal sequence on the nascent polypeptide.
This is a static version of an interactive figure; see http://www.ergito.com/main.jsp?bcs=MBIO.2.8.9 to view properly.

The role of the SRP receptor in protein translocation is transient. When the SRP binds to the signal sequence, it arrests translation (468). This usually happens when ~70 amino acids have been incorporated into the polypeptide chain (at this point the 25 residue leader has become exposed, with the next ~40 amino acids still buried in

the ribosome).

Then when the SRP binds to the SRP receptor, the SRP releases the signal sequence. The ribosome becomes bound by another component of the membrane. At this point, translation can resume. When the ribosome has been passed on to the membrane, the role of SRP and SRP receptor has been played. They now recycle, and are free to sponsor the association of another nascent polypeptide with the membrane (for review see 56).

This process may be needed to control the conformation of the protein. If the nascent protein were released into the cytoplasm, it could take up a conformation in which it might be unable to traverse the membrane. The ability of the SRP to inhibit translation while the ribosome is being handed over to the membrane is therefore important in preventing the protein from being released into the aqueous environment.

The signal peptide is cleaved from a translocating protein by a complex of 5 proteins called the **signal peptidase**. The complex is several times more abundant than the SRP and SRP receptor. Its amount is equivalent roughly to the amount of bound ribosomes, suggesting that it functions in a structural capacity. It is located on the lumenal face of the ER membrane, which implies that the entire signal sequence must cross the membrane before the cleavage event occurs. Homologous signal peptidases can be recognized in eubacteria, archaea, and eukaryotes (3066).

Last updated on 2-20-2002

Reviews

56. Walter, P. and Johnson, A. E. (1994). *Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane*. Annu. Rev. Cell Biol. 10, 87-119.

References

468. Walter, P. and Blobel, G. (1981). *Translocation of proteins across the ER III SRP causes signal sequence and site specific arrest of chain elongation that is released by microsomal membranes*. J. Cell Biol. 91, 557-561.

3066. Tjalsma, H., Bolhuis, A., van Roosmalen, M. L., Wiegert, T., Schumann, W., Broekhuizen, C. P., Quax, W. J., Venema, G., Bron, S., and van Dijl, M. (1998). *Functional analysis of the secretory precursor processing machinery of Bacillus subtilis: identification of a eubacterial homolog of archaeal and eukaryotic signal peptidases.* Genes Dev. 12, 2318-2331.

PROTEIN LOCALIZATION

2.8.10 The SRP interacts with the SRP receptor

Key Terms

The **Alu domain** comprises the parts of the 7S RNA of the SRP that are related to Alu RNA.

The **S domain** is the sequence of 7S RNA of the SRP that is not related to Alu RNA.

Key Concepts

- The SRP is a complex of 7S RNA with 6 proteins.
- The bacterial equivalent to the SRP is a complex of 4.5S RNA with two proteins.
- The SRP receptor is a dimer.
- GTP hydrolysis releases the SRP from the SRP receptor after their interaction.

The interaction between the SRP and the SRP receptor is the key event in eukaryotic translation in transferring a ribosome carrying a nascent protein to the membrane (for review see 2285). An analogous interacting system exists in bacteria, although its role is more restricted.

The SRP is an 11S ribonucleoprotein complex, containing 6 proteins (total mass 240 kD) and a small (305 base, 100 kD) 7S RNA. **Figure 8.24** shows that the 7S RNA provides the structural backbone of the particle; the individual proteins do not assemble in its absence (468; 469).

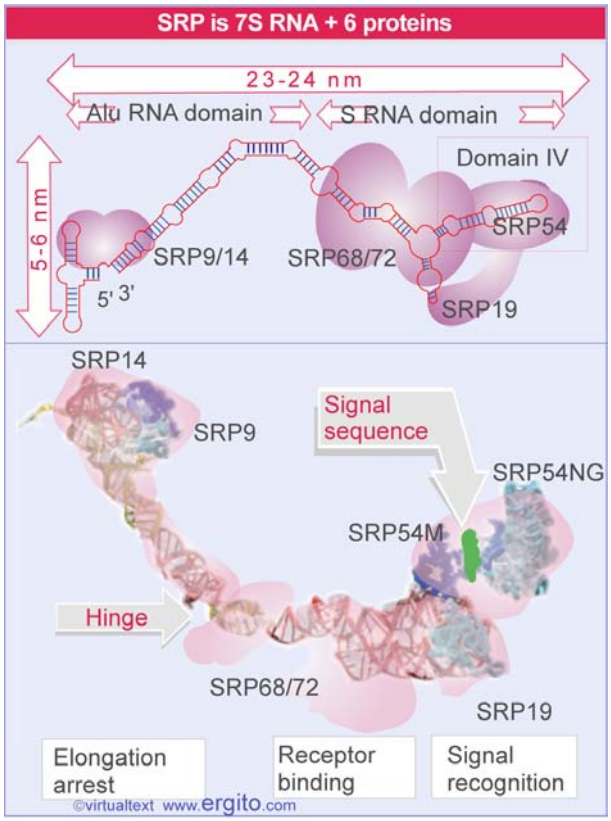


Figure 8.24 7S RNA of the SRP has two domains. Proteins bind as shown on the two dimensional diagram above to form the crystal structure shown below. Each function of the SRP is associated with a discrete part of the structure.

The 7S RNA of the SRP particle is divided into two parts. The 100 bases at the 5′ end and 45 bases at the 3′ end are closely related to the sequence of Alu RNA, a common mammalian small RNA. They therefore define the **Alu domain**. The remaining part of the RNA comprises the **S domain**.

Different parts of the SRP structure depicted in **Figure 8.24** have separate functions in protein targeting. SRP54 is the most important subunit. It is located at one end of the RNA structure, and is directly responsible for recognizing the substrate protein by binding to the signal sequence(2354). It also binds to the SRP receptor in conjunction with the SRP68-SRP72 dimer that is located at the central region of the RNA. The SRP9-SRP14 dimer is located at the other end of the molecule; it is responsible for elongation arrest (470).

The SRP is a flexible structure. In its unengaged form (not bound to signal sequence), it is quite extended, as can be seen from the crystal structure of **Figure 8.24**. **Figure 8.25** shows that binding to a signal sequence triggers a change of conformation, and the protein bends at a hinge to allow the SRP54 end to contact the ribosome at the protein exit site, while the SRP19 swings around to contact the ribosome at the elongation factor binding site (4767). This enables it to cause the elongation arrest that gives time for targeting to the translocation site on the membrane.

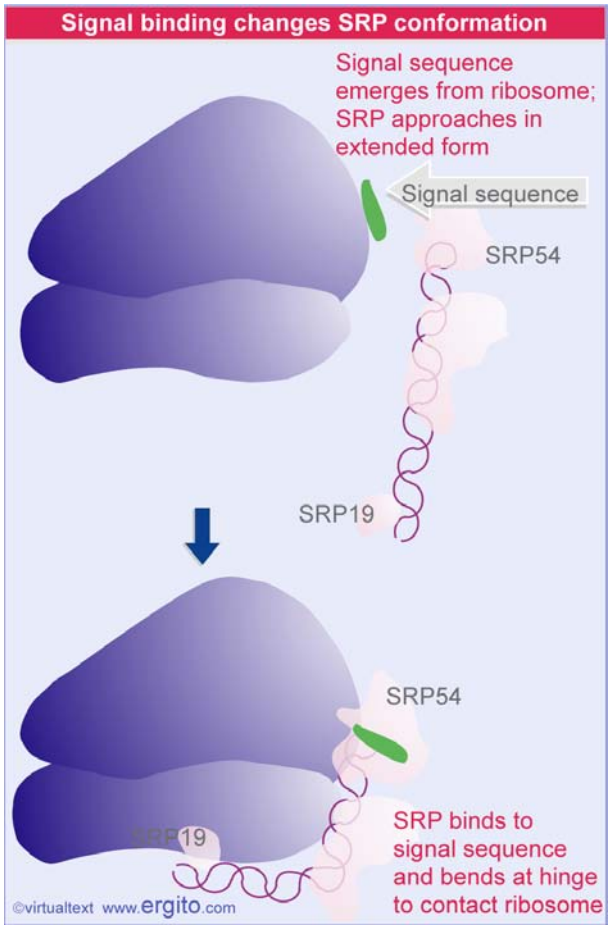


Figure 8.25 SRP binds to a signal sequence as it emerges from the ribosome. The binding causes the SRP to change conformation by bending at a "hinge", allowing SRP54 to contact th e ribosome at the protein exit siter, while SRP19 makes a second set of contacts.

The SRP receptor is a dimer containing subunits SR α (72 kD) and SR β (30 kD) (2355). The β subunit is an integral membrane protein. The amino-terminal end of the large α subunit is anchored by the β subunit. The bulk of the α protein protrudes into the cytosol. A large part of the sequence of the cytoplasmic region of the protein resembles a nucleic acid-binding protein, with many positive residues. This suggests the possibility that the SRP receptor recognizes the 7S RNA in the SRP.

There is a counterpart to SRP in bacteria, although it contains fewer components. *E. coli* contains a 4.5S RNA that associates with ribosomes and is homologous to the 7S RNA of the SRP. It associates with two proteins: Ffh is homologous to SRP54. FtsY is homologous to the α subunit of the SRP receptor. In fact, FtsY replaces the functions of both the α and β SRP subunits; its N-terminal domain substitutes for SRP β in membrane targeting, and the C-terminal domain interacts with the target protein. The role of this complex is more limited than that of SRP-SRP receptor. It is probably required to keep some (but not all) secreted proteins in a conformation that enables them to interact with the secretory apparatus. This could be the original connection between protein synthesis and secretion; in eukaryotes the SRP has acquired the additional roles of causing translational arrest and targeting to the

membrane.

Why should the SRP have an RNA component? The answer must lie in the evolution of the SRP: it must have originated very early in evolution, in an RNA-dominated world, presumably in conjunction with a ribosome whose functions were mostly carried out by RNA. The crystal structure of the complex between the protein-binding domain of 4.5S RNA and the RNA-binding domain of Ffh suggests that RNA continues to play a role in the function of SRP.

The 4.5S RNA has a region (domain IV) that is very similar to domain IV in 7S RNA (see **Figure 8.24**). Ffh consists of three domains (N, G, and M). The M domain (named for a high content of methionines) performs the key binding functions (954). It has a hydrophobic pocket that binds the signal sequence of a target protein. The hydrophobic side chains of the methionine residues create the pocket by projecting into a cleft in the protein structure. Next to the pocket is a helix-turn-helix motif that is typical of DNA-binding proteins (see *Molecular Biology 3.12.12 Repressor uses a helix-turn-helix motif to bind DNA*).

The crystal structure shows that the helix-loop-helix of the M domain binds to a duplex region of the 4.5S RNA in domain IV (953). The negatively charged backbone of the RNA is adjacent to the hydrophobic pocket. This raises the possibility that a signal sequence actually binds to both the protein and RNA components of the SRP. The positively charged sequences that start the signal sequence (see **Figure 8.20**) could interact with the RNA, while the hydrophobic region of the signal sequence could sit in the pocket.

GTP hydrolysis plays an important role in inserting the signal sequence into the membrane. Both the SRP and the SRP receptor have GTPase capability. The signal-binding subunit of the SRP, SRP54, is a GTPase. And both subunits of the SRP receptor are GTPases. All of the GTPase activities are necessary for a nascent protein to be transferred to the membrane. **Figure 8.23** shows that the SRP starts out with GDP when it binds to the signal sequence. The ribosome then stimulates replacement of the GDP with GTP. The signal sequence inhibits hydrolysis of the GTP. This ensures that the complex has GTP bound when it encounters the SRP receptor.

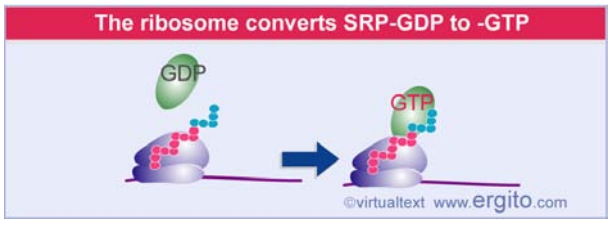


Figure 8.23 The SRP carries GDP when it binds the signal sequence. The ribosome causes the GDP to be replaced with GTP.

For the nascent protein to be transferred from the SRP to the membrane, the SRP must be released from the SRP receptor. **Figure 8.22** shows that this requires hydrolysis of the GTPs of both the SRP and the SRP receptor. The reaction has been characterized in the bacterial system, where it has the unusual feature that Ffh

activates hydrolysis by FtsY, and FtsY reciprocally activates hydrolysis by Ffh (2335).

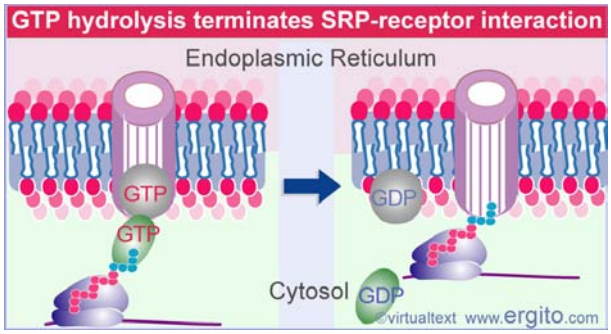


Figure 8.22 The SRP and SRP receptor both hydrolyze GTP when the signal sequence is transferred to the membrane.

Last updated on March 4, 2004

Reviews

2285. Keenan, R. J., Freymann, D. M., Stroud, R. M., and Walter, P. (2001). *The signal recognition particle.* Annu. Rev. Biochem. 70, 755-775.

References

468. Walter, P. and Blobel, G. (1981). *Translocation of proteins across the ER III SRP causes signal sequence and site specific arrest of chain elongation that is released by microsomal membranes*. J. Cell Biol. 91, 557-561.

469. Walter, P. and Blobel, G. (1982). *Signal recognition particle contains a 7S RNA essential for protein translocation across the ER.* Nature 299, 691-698.

470. Siegel, V. and Walter, P. (1988). *Each of the activities of SRP is contained within a distinct domain: analysis of biochemical mutants of SRP*. Cell 52, 39-49.

953. Batey, R. T., Rambo, R. P., Lucast, L., Rha, B., and Doudna, J. A. (2000). *Crystal structure of the ribonucleoprotein core of the signal recognition particle.* Science 287, 1232-1239.

954. Keenan, R. J., Freymann, D. M., Walter, P., and Stroud, R. M. (1998). *Crystal structure of the signal sequence-binding subunit of the signal recognition particle.* Cell 94, 181-191.

2335. Powers, T. and Walter, P. (1995). *Reciprocal stimulation of GTP hydrolysis by two directly interacting GTPases.* Science 269, 1422-1424.

2354. Zopf, D., Bernstein, H. D., Johnson, A. E., and Walter, P. (1990). *The methionine-rich domain of the 54 kd protein subunit of the signal recognition particle contains an RNA binding site and can be crosslinked to a signal sequence.* EMBO J. 9, 4511-4517.

2355. Tajima, S., Lauffer, L., Rath, V. L., and Walter, P. (1986). *The signal recognition particle receptor is a complex that contains two distinct polypeptide chains.* J. Cell Biol. 103, 1167-1178.

4767. Halic, M., Becker, T., Pool, M. R., Spahn, C. M., Grassucci, R. A., Frank, J., and Beckmann, R. (2004). *Structure of the signal recognition particle interacting with the elongation-arrested ribosome.* Nature 427, 808-814.

PROTEIN LOCALIZATION

2.8.11 The translocon forms a pore

Key Terms

A **translocon** is a discrete structure in a membrane that forms a channel through which (hydrophilic) proteins may pass.

Key Concepts

- The Sec61 trimeric complex provides the channel for proteins to pass through a membrane.
- A translocating protein passes directly from the ribosome to the translocon without exposure to the cytosol.

There is a basic problem in passing a (largely) hydrophilic protein through a hydrophobic membrane. The energetics of the interaction between the charged protein and the hydrophobic lipids are highly unfavorable. However, a protein in the process of translocation across the ER membrane can be extracted by denaturants that are effective in an aqueous environment. The same denaturants do not extract proteins that are resident components of the membrane. This suggests the model for translocation illustrated in **Figure 8.26**, in which proteins that are part of the ER membrane form an aqueous channel through the bilayer. A translocating protein moves through this channel, interacting with the resident proteins rather than with the lipid bilayer. The channel is sealed on the lumenal side to stop free transfer of ions between the ER and the cytosol.

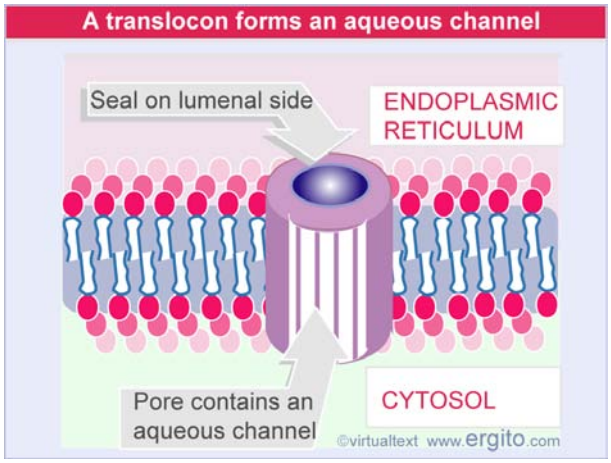


Figure 8.26 The translocon is a trimer of Sec61 that forms a channel through the membrane. It is sealed on the lumenal (ER) side.

The channel through the membrane is called the **translocon**. Its components have been identified in two ways. Resident ER membrane proteins that are crosslinked to translocating proteins are potential subunits of the channel (2822). And *sec* mutants

in yeast (named because they fail to secrete proteins) include a class that cause precursors of secreted or membrane proteins to accumulate in the cytosol (2791; 2821). These approaches together identified the *Sec61 complex*, which consists of three transmembrane proteins: Sec61 α, β, γ. Sec61 is the major component of the translocon. In detergent (which provides a hydrophobic milieu that mimics the effect of a surrounding membrane), Sec61 forms cylindrical oligomers with a diameter of ~85Å and a central pore of ~20Å. Each oligomer consists of 4 heterotrimers (476).

Is the channel a preexisting structure (as implied in the figure) or might it be assembled in response to the association of a hydrophobic signal sequence with the lipid bilayer? Channels can be detected by their ability to allow the passage of ions (measured as a localized change in electrical conductance). Ion-conducting channels can be detected in the ER membrane, and their state depends on protein translocation (472; 474). This demonstrates that the channel is a permanent feature of the membrane.

A channel opens when a nascent polypeptide is transferred from a ribosome to the ER membrane. The translocating protein fills the channel completely, so ions cannot pass through during translocation. But if the protein is released by treatment with puromycin, then the channel becomes freely permeable. If the ribosomes are removed from the membrane, the channel closes, suggesting that the open state requires the presence of the ribosome. This suggests that the channel is controlled in response to the presence of a translocating protein.

Measurements of the abilities of fluorescence quenching agents of different sizes to enter the channel suggest that it is large, with an internal diameter of 40-60Å. This is much larger than the diameter of an extended α-helical stretch of protein. It is also larger than the pore seen in direct views of the channel; this discrepancy remains to be explained (472).

The aqueous environment of an amino acid in a protein can be measured by incorporating variant amino acids that have photoreactive residues. The fluorescence of these residues indicates whether they are in an aqueous or hydrophobic environment. Experiments with such probes show that when the signal sequence is first synthesized in the ribosome, it is in an aqueous state, but is not accessible to ions in the cytosol. It remains in the aqueous state throughout its interaction with a membrane. This suggests that the translocating protein travels directly from an enclosed tunnel in the ribosome into an aqueous channel in the membrane.

In fact, access to the pore is controlled (or "gated") on *both* sides of the membrane. Before attachment of the ribosome, the pore is closed on the lumenal side. **Figure 8.27** shows that when the ribosome attaches, it seals the pore on the cytosolic side. When the nascent protein reaches a length of ~70 amino acids, that is, probably when it extends fully across the channel, the pore opens on the lumenal side. So at all times, the pore is closed on one side or the other, maintaining the ionic integrities of the separate compartments (474; 475).

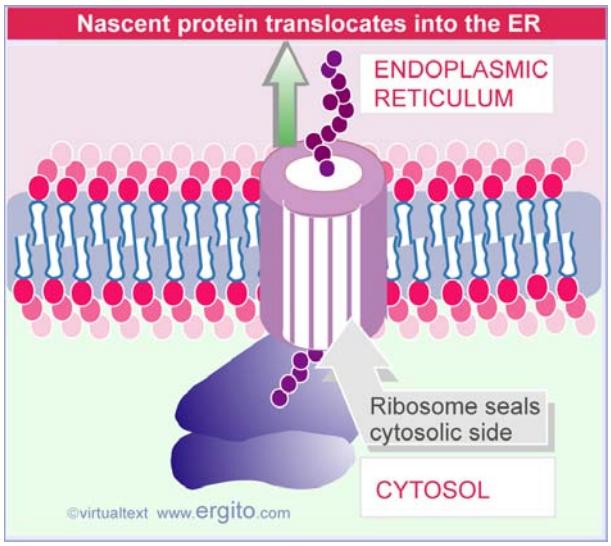


Figure 8.27 A nascent protein is transferred directly from the ribosome to the translocon. The ribosomal seals the channel on the cytosolic side.

The translocon is versatile, and can be used by translocating proteins in several ways:

- It is the means by which nascent proteins are transferred from cytosolic ribosomes to the lumen of endoplasmic reticulum (see *Molecular Biology 2.8.12 Translocation requires insertion into the translocon and (sometimes) a ratchet in the ER*).
- It is also the route by which integral membrane proteins of the ER system are transferred to the membrane; this requires the channel to open or disaggregate in some unknown way so that the protein can move laterally into the lipid bilayer (see *Molecular Biology 2.8.16 How do proteins insert into membranes?*).
- Proteins can also be transferred from the ER back to the cytosol; this is known as reverse translocation (see *Molecular Biology 2.8.13 Reverse translocation sends proteins to the cytosol for degradation*).

Last updated on January 7, 2004

References

472. Simon, S. M. and Blobel, G. (1991). *A protein-conducting channel in the endoplasmic reticulum*. Cell 65, 371-380.

474. Crowley, K. S. (1994). *Secretory proteins move through the ER membrane via an aqueous, gated pore*. Cell 78, 461-471.

475. Liao, S., Lin, J., Do, H., and Johnson, A. E. (1997). *Both lumenal and cytosolic gating of the aqueous ER translocon pore are regulated from inside the ribosome during membrane protein integration*. Cell 90, 31-41.

476. Hanein, D. et al. (1996). *Oligomeric rings of the Sec61p complex induced by ligands required for protein translocation*. Cell 87, 721-732.

2791. Deshaies, R. J. and Schekman, R. (1987). *A yeast mutant defective at an early stage in import of secretory protein precursors into the endoplasmic reticulum.* J. Cell Biol. 105, 633-645.

2821. Esnault, Y., Blondel, M. O., Deshaies, R. J., Scheckman, R., and Kepes, F. (1993). *The yeast SSS1 gene is essential for secretory protein translocation and encodes a conserved protein of the endoplasmic reticulum.* EMBO J. 12, 4083-4093.

2822. Mothes, W., Prehn, S., and Rapoport, T. A. (1994). *Systematic probing of the environment of a translocating secretory protein during translocation through the ER membrane.* EMBO J. 13, 3973-3982.

4522. Van den Berg, B., Clemons, W. M., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C., and Rapoport, T. A. (2004). *X-ray structure of a protein-conducting channel.* Nature 427, 36-44.

PROTEIN LOCALIZATION

2.8.12 Translocation requires insertion into the translocon and (sometimes) a ratchet in the ER

Key Concepts

- The ribosome, SRP, and SRP receptor are sufficient to insert a nascent protein into a translocon.
- Proteins that are inserted post-translationally require additional components in the cytosol and Bip in the ER.
- Bip is a ratchet that prevents a protein from slipping backward.

The translocon and the SRP receptor are the basic components required for co-translational translocation. When the Sec61 complex is incorporated into artificial membranes together with the SRP receptor, it can support translocation of some nascent proteins. Other nascent proteins require the presence of an additional component, TRAM, which is a major protein that becomes crosslinked to a translocating nascent chain. TRAM stimulates the translocation of all proteins.

The components of the translocon and their functions are summarized in **Figure 8.28**. The simplicity of this system makes several important points. We visualize Sec61 as forming the channel and also as interacting with the ribosome. The initial targeting is made when the SRP recognizes the signal sequence as the newly synthesized protein begins to emerge from the ribosome. The SRP binds to the SRP receptor, and the signal sequence is transferred to the translocon. When the signal sequence enters the translocon, the ribosome attaches to Sec61, forming a seal so that the pore is not exposed to the cytosol. Cleavage of the signal peptide does not occur in this system, and therefore cannot be necessary for translocation *per se*. In this system, components on the lumenal side of the membrane are not needed for translocation.

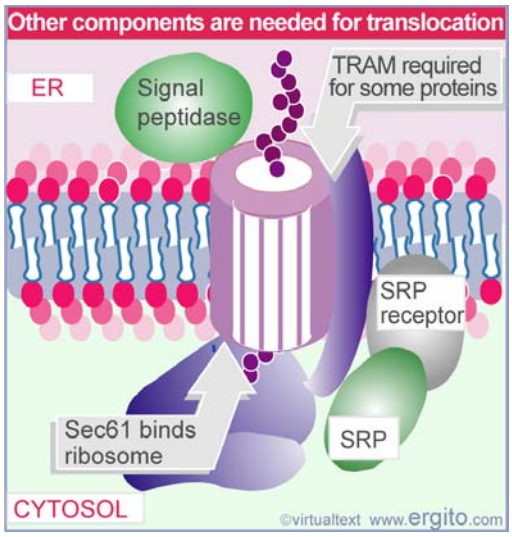


Figure 8.28 Translocation requires the translocon, SRP, SRP receptor, Sec61, TRAM, and signal peptidase.

Of course, the efficiency of the *in vitro* system is relatively low. Additional components could be required *in vivo* to achieve efficient transfer or to prevent other cellular proteins from interfering with the process (473; for review see 48; 60).

A more complex apparatus is required in certain cases in which a protein is inserted into a membrane post-translationally. The same Sec61 complex forms the channel, but four other Sec proteins are also required, and in addition the chaperone BiP (a member of the Hsp70 class) and a supply of ATP are required on the lumenal side of the membrane. **Figure 8.29** shows that BiP behaves as ratchet (936). In the absence of BiP, Brownian motion allows the protein to slip back into the cytosol. But BiP grabs the protein as it exits the pore into the endoplasmic reticulum. This stops the protein from moving backward. BiP does not pull the protein through; it just stops it from sliding back. (The reason why BiP is required for post-translational translocation but not for co-translational translocation may be that a newly synthesized protein is continuously extruded from the ribosome and therefore cannot slip backward.)

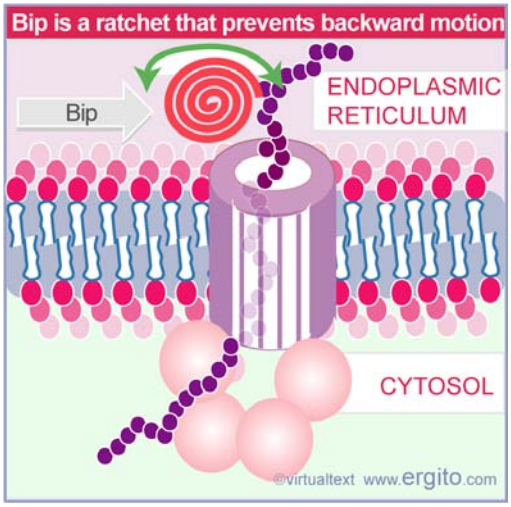


Figure 8.29 BiP acts as a ratchet to prevent backward diffusion of a translocating protein.
This is a static version of an interactive figure; see http://www.ergito.com/main.jsp?bcs=MBIO.2.8.12 to view properly.

Last updated on 11-13-2001

Reviews

48. Walter, P. and Lingappa, V. (1986). *Mechanism of protein translocation across the endoplasmic reticulum membrane.* Annu. Rev. Cell Biol. 2, 499-516.

60. Rapoport, T. A., Jungnickel, B., and Kutay, U. (1996). *Protein transport across the eukaryotic endoplasmic reticulum and bacterial inner membranes.* Annu. Rev. Biochem. 65, 271-303.

References

473. Gorlich, D. and Rapoport, T. A. (1993). *Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane*. Cell 75, 615-630.

936. Matlack, K. E., Misselwitz, B., Plath, K., and Rapoport, T. A. (1999). *BiP acts as a molecular ratchet during posttranslational transport of prepro-alpha factor across the ER membrane.* Cell 97, 553-564.

PROTEIN LOCALIZATION

2.8.13 Reverse translocation sends proteins to the cytosol for degradation

Key Terms

Retrograde translocation (Reverse translocation) is the translocation of a protein from the lumen of the ER to the cytoplasm. It usually occurs to allow misfolded or damaged proteins to be degraded by the proteasome.

Key Concepts

- Sec61 translocons can be used for reverse translocation of proteins from the ER into the cytosol.

Several important activities occur within the endoplasmic reticulum. Proteins move through the ER en route to a variety of destinations (see *Molecular Biology 6.27 Protein trafficking*). They are glycosylated and folded into their final conformations. The ER provides a "quality control" system in which misfolded proteins are identified and degraded. However, the degradation itself does not occur in the ER, but may require the protein to be exported back to the cytosol.

The first indication that ER proteins are degraded in the cytosol and not in the ER itself was provided by evidence for the involvement of the proteasome, a large protein aggregate with several proteolytic activities (see *Molecular Biology Supplement 8.32 The proteasome is a large machine that degrades ubiquitinated proteins*). Inhibitors of the proteasome prevent the degradation of aberrant ER proteins. Proteins are marked for cleavage by the proteasome when they are modified by the addition of ubiquitin, a small polypeptide chain (see *Molecular Biology 2.8.31 Ubiquitination targets proteins for degradation*). The important point to note now is that ubiquitination and proteasomal degradation both occur in the cytosol (with a minor proportion in the nucleus).

Transport from the ER back into the cytosol occurs by a reversal of the usual process of import (for review see 2979). This is called **reverse translocation**. The Sec61 translocon is used. The conditions are different; for example, the translocon is not associated with a ribosome. Some mutations in Sec61 prevent reverse translocation, but do not prevent forward translocation (2179; 2180). This could be either because there is some difference in the process or (more likely) because these regions interact with other components that are necessary for reverse translocation.

Figure 8.30 points out that we do not know how the channel is opened to allow insertion of the protein on the ER side. Special components are presumably involved. One model is that misfolded or misassembled proteins are recognized by chaperones, which transfer them to the translocon (for review see 2178). In one particular case, human cytomegalovirus (CMV) codes for cytosolic proteins that destroy newly synthesized MHC class I (cellular major histocompatibility complex) proteins. This requires a viral protein product (US2), which is a membrane protein that functions in

the ER. It interacts with the MHC proteins and probably conveys them into the translocon for reverse translocation.

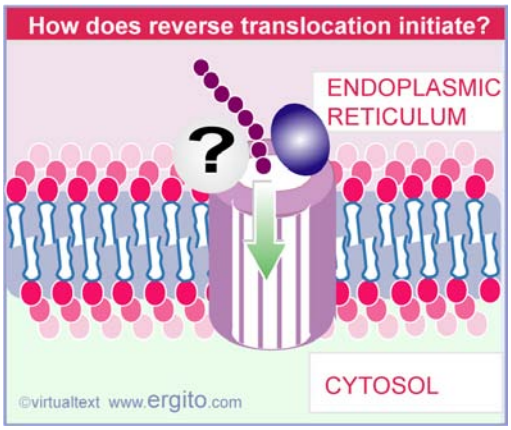


Figure 8.30 Reverse translocation uses the translocon to send an unfolded protein from the ER to the cytosol, where it is degraded. The mechanism of putting the translocon into reverse is not known.

The system involved in the degradation of aberrant ER proteins can be identified by mutations (in yeast) that lead to accumulation of aberrant proteins. Usually a protein that misfolds (produced by a mutated gene) is degraded instead of being transported through the ER. Yeast mutants that cannot degrade the substrate fall into two classes: some identify components of the proteolytic apparatus, such as the enzymes involved in ubiquitination; other identify components of the transport apparatus, including Sec61, BiP, and Sec63. There is also a protein in the ER membrane that functions on the cytosolic side to localize the ubiquitination enzymes at the translocon. In fact, retrograde transport into the cytosol cannot occur in the absence of this protein, which suggests that there is a mechanical link between retrograde transport and degradation (494).

Last updated on 11-13-2001

Reviews

2178. Johnson, A. E. and Haigh, N. G. (2000). *The ER translocon and retrotranslocation: is the shift into reverse manual or automatic?* Cell 102, 709-712.

2979. Tsai, B., Ye, Y., and Rapoport, T. A. (2002). *Retro-translocation of proteins from the endoplasmic reticulum into the cytosol.* Nat. Rev. Mol. Cell Biol. 3, 246-255.

References

494. Wiertz, E. J. H. J. et al. (1996). *Sec61-mediated transfer of a membrane protein from the endoplasmic reticulum to the proteasome for destruction*. Nature 384, 432-438.

2179. Zhou, M. and Schekman, R. (1999). *The engagement of Sec61p in the ER dislocation process.* Mol. Cell 4, 925-934.

2180. Wilkinson, B. M., Tyson, J. R., Reid, P. J., and Stirling, C. J. (2000). *Distinct domains within yeast Sec61p involved in post-translational translocation and protein dislocation.* J. Biol. Chem. 275, 521-529.

PROTEIN LOCALIZATION

2.8.14 Proteins reside in membranes by means of hydrophobic regions

Key Terms

The **transmembrane region (transmembrane domain)** is the part of a protein that spans the membrane bilayer. It is hydrophobic and in many cases contains approximately 20 amino acids that form an α -helix. It is also called the transmembrane domain.

A **transmembrane protein (Integral membrane protein)** extends across a lipid bilayer. A hydrophobic region (typically consisting of a stretch of 20-25 hydrophobic and/or uncharged aminoa acids) or regions of the protein resides in the membrane. Hydrophilic regions are exposed on one or both sides of the membrane.

Key Concepts

- Group I proteins have the N-terminus on the far side of the membrane; group II proteins have the opposite orientation.
- Some proteins have multiple membrane-spanning domains.

All biological membranes contain proteins, which are held in the lipid bilayer by noncovalent interactions. The operational definition of an **integral membrane protein** is that it requires disruption of the lipid bilayer in order to be released from the membrane. A common feature in such proteins is the presence of at least one **transmembrane domain**, consisting of an α-helical stretch of 21-26 hydrophobic amino acids. A sequence that fits the criteria for membrane insertion can be identified by a hydropathy plot, which measures the cumulative hydrophobicity of a stretch of amino acids. A protein that has domains exposed on both sides of the membrane is called a **transmembrane protein**. The association of a protein with a membrane takes several forms (see *Molecular Biology Supplement 32.5 Membranes and membrane proteins*). The topography of a membrane protein depends on the number and arrangement of transmembrane regions.

When a protein has a single transmembrane region, its position determines how much of the protein is exposed on either side of the membrane. A protein may have extensive domains exposed on both sides of the membrane or may have a site of insertion close to one end, so that little or no material is exposed on one side. The length of the N-terminal or C-terminal tail that protrudes from the membrane near the site of insertion varies from insignificant to quite bulky.

Figure 8.31 shows that proteins with a single transmembrane domain fall into two classes. Group I proteins in which the N-terminus faces the extracellular space are more common than group II proteins in which the orientation has been reversed so that the N-terminus faces the cytoplasm. Orientation is determined during the insertion of the protein into the endoplasmic reticulum.

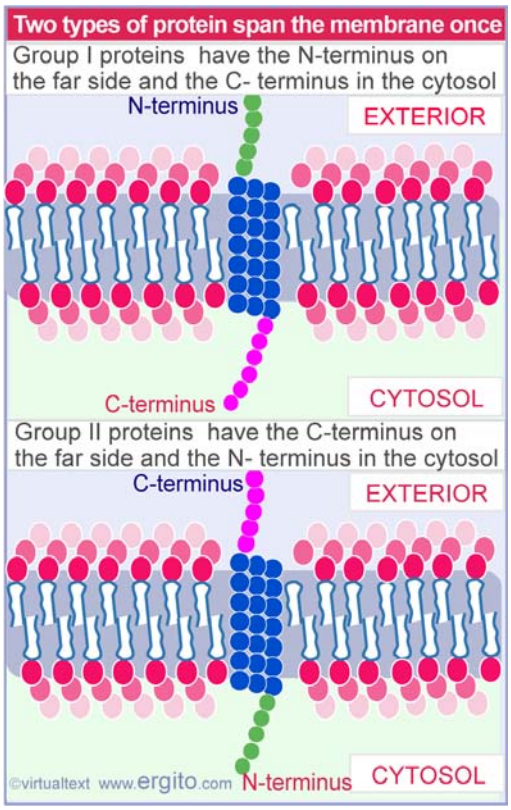


Figure 8.31 Group I and group II transmembrane proteins have opposite orientations with regard to the membrane.

Figure 8.32 shows orientations for proteins that have multiple membrane-spanning domains. An odd number means that both termini of the protein are on opposite sides of the membrane, whereas an even number implies that the termini are on the same face. The extent of the domains exposed on one or both sides is determined by the locations of the transmembrane domains. Domains at either terminus may be exposed, and internal sequences between the domains "loop out" into the extracellular space or cytoplasm. One common type of structure is the 7-membrane passage or "serpentine" receptor; another is the 12-membrane passage component of an ion channel.

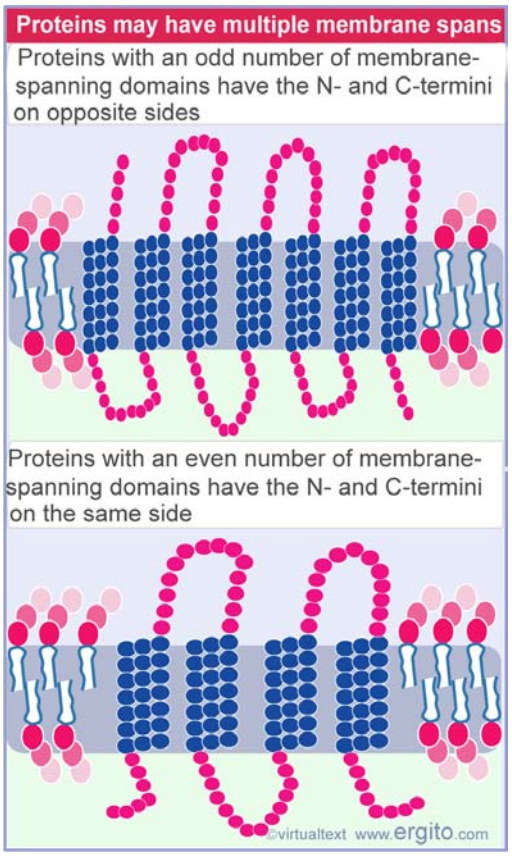


Figure 8.32 The orientations of the termini of multiple membrane-spanning proteins depends on whether there is an odd or even number of transmembrane segments.

Does a transmembrane domain itself play any role in protein function besides allowing the protein to insert into the lipid bilayer? In the simple group I or II proteins, it has little or no additional function; often it can be replaced by any other transmembrane domain. However, transmembrane domains play an important role in the function of proteins that make multiple passes through the membrane or that have subunits that oligomerize within the membrane. The transmembrane domains in such cases often contain polar residues, which are not found in the single membrane-spanning domains of group I and group II proteins. Polar regions in the membrane-spanning domains do not interact with the lipid bilayer, but instead interact with one another. This enables them to form a polar pore or channel within the lipid bilayer. Interaction between such transmembrane domains can create a hydrophilic passage through the hydrophobic interior of the membrane. This can allow highly charged ions or molecules to pass through the membrane, and is important for the function of ion channels and transport of ligands. Another case in which conformation of the transmembrane domains is important is provided by certain receptors that bind lipophilic ligands. In such cases, the transmembrane domains (rather than the extracellular domains) bind the ligand within the plane of the membrane.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.2.8.14

PROTEIN LOCALIZATION

2.8.15 Anchor sequences determine protein orientation

Key Terms

An **anchor (stop-transfer)** (often referred to as a "transmembrane anchor") is a segment of a transmembrane protein which resides in the membrane.

Key Concepts

- An anchor sequence halts the passage of a protein through the translocon. Typically this is located at the C-terminal end and results in a group I orientation in which the N-terminus has passed through the membrane.
- A combined signal-anchor sequence can be used to insert a protein into the membrane and anchor the site of insertion. Typically this is internal and results in a group II orientation in which the N-terminus is cytosolic.

Proteins that are secreted from the cell pass through a membrane while remaining in the aqueous channel of the translocon. By contrast, proteins that reside in membranes start the process in the same way, but then transfer from the aqueous channel into the hydrophobic environment. The challenge in accounting for insertion of proteins into membranes is to explain what distinguishes transmembrane proteins from secreted proteins, and causes this transfer. The pathway by which proteins of either type I or type II are inserted into the membrane follows the same initial route as that of secretory proteins, relying on a signal sequence that functions co-translationally. But proteins that are to remain within the membrane possess a second, **stop-transfer** signal. This takes the form of a cluster of hydrophobic amino acids adjacent to some ionic residues. The cluster serves as an **anchor** that latches on to the membrane and stops the protein from passing right through.

A surprising property of anchor sequences is that they can function as signal sequences when engineered into a different location. When placed into a protein lacking other signals, such a sequence may sponsor membrane translocation. One possible explanation for these results is that the signal sequence and anchor sequence interact with some common component of the apparatus for translocation. Binding of the signal sequence initiates translocation, but the appearance of the anchor sequence displaces the signal sequence and halts transfer.

Membrane insertion starts by the insertion of a signal sequence in the form of a hairpin loop, in which the N-terminus remains on the cytoplasmic side. Two features determine the position and orientation of a protein in the membrane: whether the signal sequence is cleaved; and the location of the anchor sequence.

The insertion of type I proteins is illustrated in **Figure 8.33**. The signal sequence is N-terminal. The location of the anchor signal determines when transfer of the protein is halted. When the anchor sequence takes root in the membrane, domains on the

N-terminal side will be located in the lumen, while domains on the C-terminal side are located facing the cytosol.

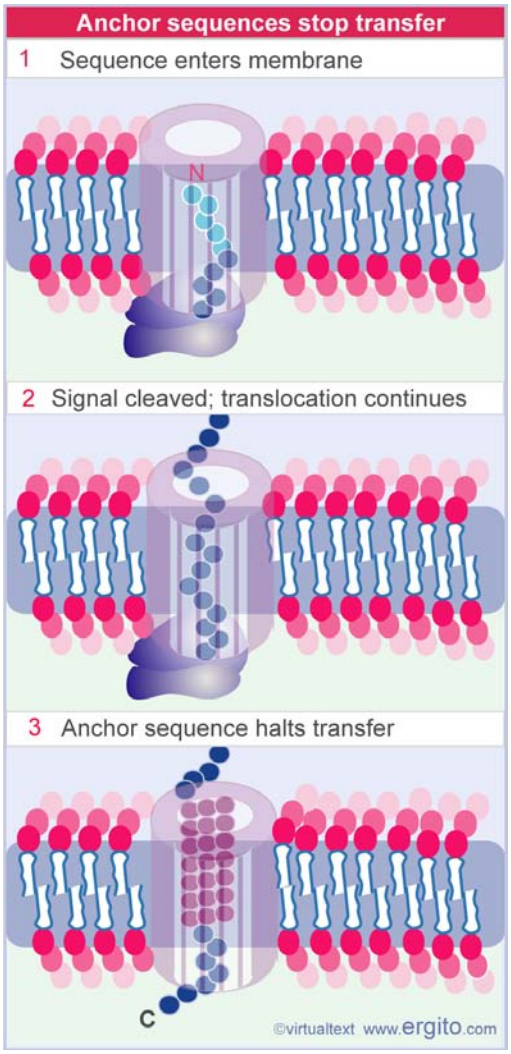


Figure 8.33 Proteins that reside in membranes enter by the same route as secreted proteins, but transfer is halted when an anchor sequence passes into the membrane. If the anchor is at the C-terminus, the bulk of the protein passes through the membrane and is exposed on the far surface.

A common location for a stop-transfer sequence of this type is at the C-terminus. As shown in the figure, transfer is halted only as the last sequences of the protein enter the membrane. This type of arrangement is responsible for the location in the membrane of many proteins, including cell surface proteins. Most of the protein sequence is exposed on the lumenal side of the membrane, with a small or negligible tail facing the cytosol.

Type II proteins do not have a cleavable leader sequence at the N-terminus. Instead the signal sequence is combined with an anchor sequence. We imagine that the general pathway for the integration of type I proteins into the membrane involves the

steps illustrated in **Figure 8.34**. The signal sequence enters the membrane, but the joint signal-anchor sequence does not pass through. Instead it stays in the membrane (perhaps interacting directly with the lipid bilayer), while the rest of the growing polypeptide continues to loop into the endoplasmic reticulum.

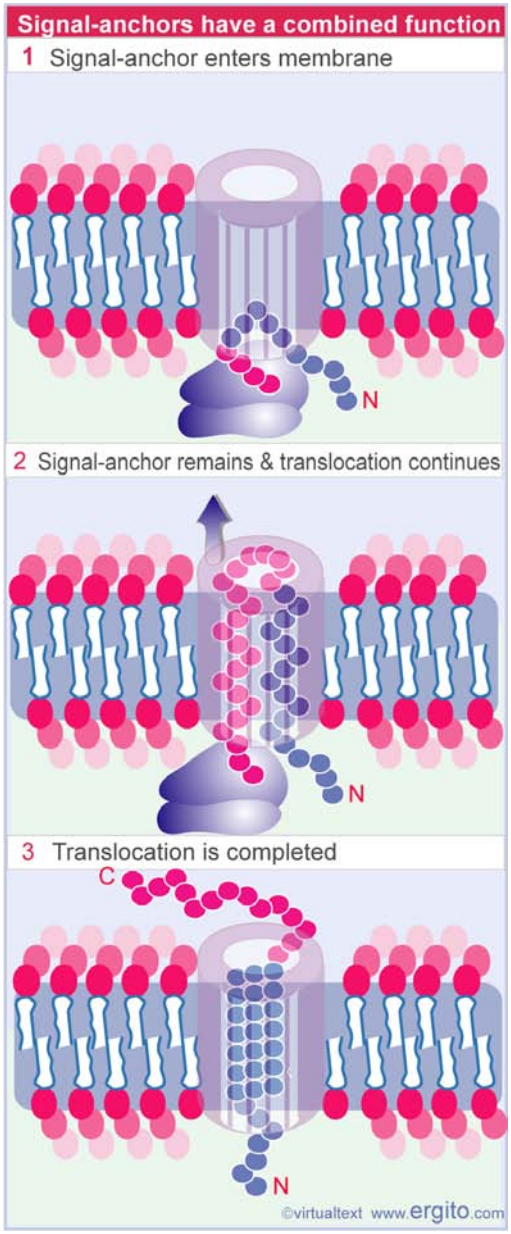


Figure 8.34 A combined signal-anchor sequence causes a protein to reverse its orientation, so that the N-terminus remains on the inner face and the C-terminus is exposed on the outer face of the membrane.

The signal-anchor sequence is usually internal, and its location determines which parts of the protein remain in the cytosol and which are extracellular. Essentially all the N-terminal sequences that precede the signal-anchor are exposed to the cytosol.

Usually the cytosolic tail is short, ~6-30 amino acids. In effect the N-terminus remains constrained while the rest of the protein passes through the membrane. This reverses the orientation of the protein with regard to the membrane.

The combined signal-anchor sequences of type II proteins resemble cleavable signal sequences. **Figure 8.35** gives an example. Like cleavable leader sequences, the amino acid composition is more important than the actual sequence. The regions at the extremities of the signal-anchor carry positive charges; the central region is uncharged and resembles a hydrophobic core of a cleavable leader. Mutations to introduce charged amino acids in the core region prevent membrane insertion; mutations on either side prevent the anchor from working, so the protein is secreted or located in an incorrect compartment.

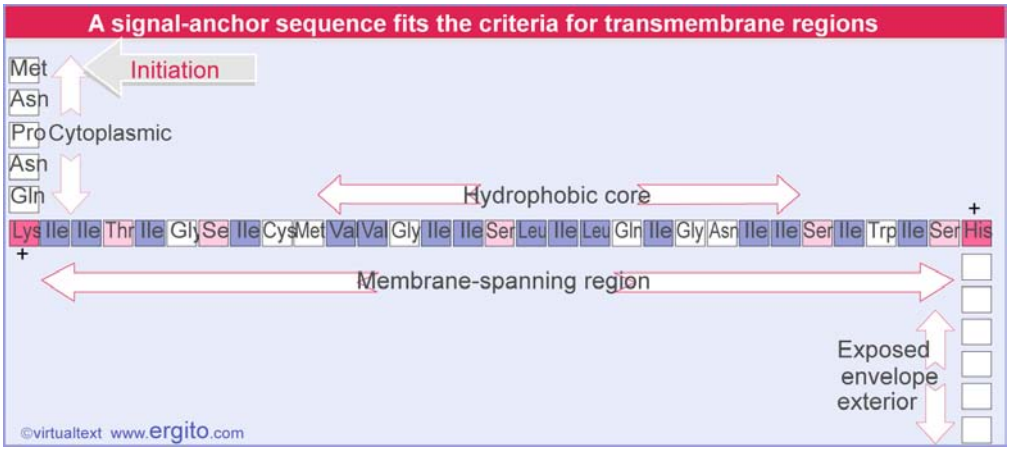


Figure 8.35 The signal-anchor of influenza neuraminidase is located close to the N-terminus and has a hydrophobic core.

The distribution of charges around the anchor sequence has an important effect on the orientation of the protein. More positive charges are usually found on the cytoplasmic side (N-terminal side in type II proteins). If the positive charges are removed by mutation, the orientation of the protein can be reversed. The effect of charges on orientation is summarized by the "positive inside" rule, which says that the side of the anchor with the most positive charges will be located in the cytoplasm. The positive charges in effect provide a hook that latches on to the cytoplasmic side of the membrane, controlling the direction in which the hydrophobic region is inserted, and thus determining the orientation of the protein.

This content is available online at http://www.ergito.com/main.jsp?bcs=MBIO.2.8.15

PROTEIN LOCALIZATION

2.8.16 How do proteins insert into membranes?

Key Concepts

- Transfer of transmembrane domains from the translocon into the lipid bilayer is triggered by the interaction of the transmembrane region with the translocon.

We have a reasonable understanding of the processes by which secreted proteins pass through membranes and of how this relates to the insertion of the single-membrane spanning group I and group II proteins. We cannot yet explain the details of insertion of proteins with multiple membrane-spanning domains.

We understand how a secreted protein passes through a membrane without any conflict, but it is difficult to apply the same model to a protein that resides in the membrane. **Figure 8.36** illustrates the difference between the organization of a translocating protein, which is protected from the lipid bilayer by the aqueous channel, and a transmembrane protein, which has a hydrophobic segment directly in contact with the membrane. We do not know how a protein is transferred from its passage through the proteinaceous channel into the lipid bilayer itself. One possibility is that there is some mechanism for transferring hydrophobic transmembrane domains directly from the channel into the membrane, as suggested in **Figure 8.37**. This idea is supported by observations of an *in vitro* system which measured transfer into a lipid environment for proteins with different transmembrane domains. When the domain passed a threshold of hydrophobicity, the protein could pass from a channel consisting of Sec61 and TRAM into the lipid bilayer (1205). The simplest explanation is that the structure of the channel allows the translocating protein to contact the lipid bilayer, so that a sufficiently hydrophobic segment can simply partition directly into the lipid. An alternative is that hydrophobic domains cause the channel to disaggregate, exposing the hydrophobic amino acids to the lipid bilayer.

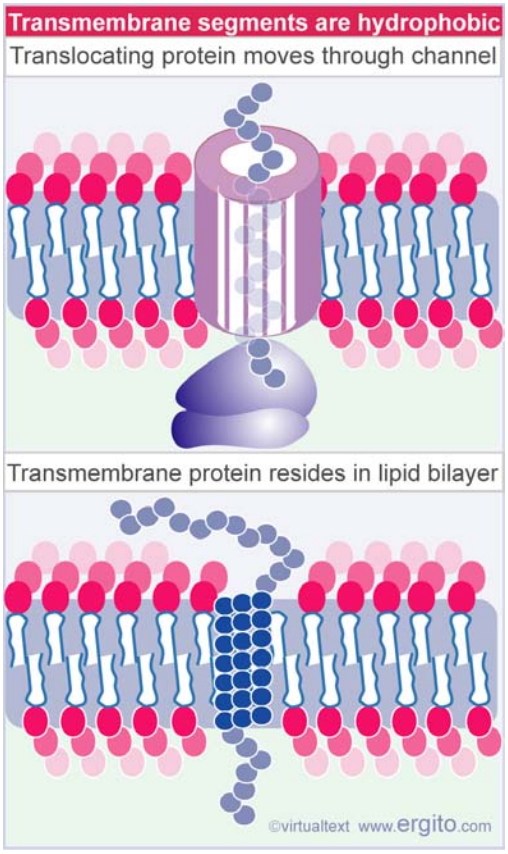


Figure 8.36 How does a transmembrane protein make the transition from moving through a proteinaceous channel to interacting directly with the lipid bilayer?

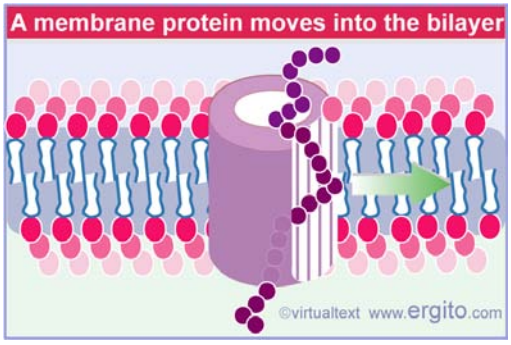


Figure 8.37 Newly synthesized membrane proteins are able to transfer laterally from the translocon into the lipid bilayer. The mechanism of transfer is not known.

It has always been a common assumption that, whatever the exact mechanism for transferring the transmembrane segment into the membrane, it is triggered by the presence of the transmembrane sequence in the pore. However, changes in the pore occur earlier in response to the synthesis of the transmembrane sequence in the ribosome (475). When a secreted protein passes through the pore, the channel

remains sealed on the cytosolic side but opens on the lumenal side after synthesis of the first 70 residues. But as soon as a transmembrane sequence has been fully synthesized, that is, while it is still entirely within the ribosome, the pore closes on the lumenal side. How this change relates to the transfer of the transmembrane sequence into the membrane is not clear.

The process of insertion into a membrane has been characterized for both type I proteins and type II proteins, in which there is a single transmembrane domain. How is a protein with multiple membrane-spanning regions inserted into a membrane? Much less is known about this process, but we assume that it relies on sequences that provide signal and/or anchor capabilities. One model is to suppose that there is an alternating series of signal and anchor sequences. Translocation is initiated at the first signal sequence and continues until stopped by the first anchor. Then it is reinitiated by a subsequent signal sequence, until stopped by the next anchor. It is possible that there are multiple pathways for integration into the membrane, because in some cases a transmembrane domain seems to move into the lipid bilayer as soon as it enters the translocon, but in other cases there can be a delay until other transmembrane regions have been synthesized (2736; 2737; 2772; 2779; for review see 49; 2739).

Last updated on 11-13-2001

Reviews

49. Wickner, W. T. and Lodish, H. (1985). *Multiple mechanisms of protein insertion into and across membranes.* Science 230, 400-407.

2739. Hegde, R. S. and Lingappa, V. R. (1997). *Membrane protein biogenesis: regulated complexity at the endoplasmic reticulum.* Cell 91, 575-582.

References

475. Liao, S., Lin, J., Do, H., and Johnson, A. E. (1997). *Both lumenal and cytosolic gating of the aqueous ER translocon pore are regulated from inside the ribosome during membrane protein integration*. Cell 90, 31-41.

1205. Heinrich, S. U., Mothes, W., Brunner, J., and Rapoport, T. A. (2000). *The Sec61p complex mediates the integration of a membrane protein by allowing lipid partitioning of the transmembrane domain.* Cell 102, 233-244.

2736. Mothes, W., Heinrich, S. U., Graf, R., Nilsson, I., von Heijne, G., Brunner, J., and Rapoport, T. A. (1997). *Molecular mechanism of membrane protein integration into the endoplasmic reticulum.* Cell 89, 523-533.

2737. Do, H., Do, H., Lin, J., Andrews, D. W., and Johnson, A. E. (1996). *The cotranslational integration of membrane proteins into the phospholipid bilayer is a multistep process.* Cell 85, 369-378.

2772. Borel, A. C., and Simon, S. M. (1996). *Biogenesis of polytopic membrane proteins: membrane segments assemble within translocation channels prior to membrane integration.* Cell 85, 379-389.

2779. Kim, P. K., Janiak-Spens, F., Trimble, W. S., Leber, B., and Andrews, D. W. (1997). *Evidence for multiple mechanisms for membrane binding and integration via carboxyl-terminal insertion sequences.* Biochemistry 36, 8873-8882.

PROTEIN LOCALIZATION

2.8.17 Post-translational membrane insertion depends on leader sequences

Key Concepts

- N-terminal leader sequences provide the information that allows proteins to associate with mitochondrial or chloroplast membranes

Mitochondria and chloroplasts synthesize only some of their proteins. Mitochondria synthesize only ~10 organelle proteins; chloroplasts synthesize ~50 proteins. The majority of organelle proteins are synthesized in the cytosol by the same pool of free ribosomes that synthesize cytosolic proteins. They must then be imported into the organelle.

Many proteins that enter mitochondria or chloroplasts by a post-translational process have leader sequences that are responsible for primary recognition of the outer membrane of the organelle. As shown in the simplified diagram of **Figure 8.38**, the leader sequence initiates the interaction between the precursor and the organelle membrane. The protein passes through the membrane, and the leader is cleaved by a protease on the organelle side.

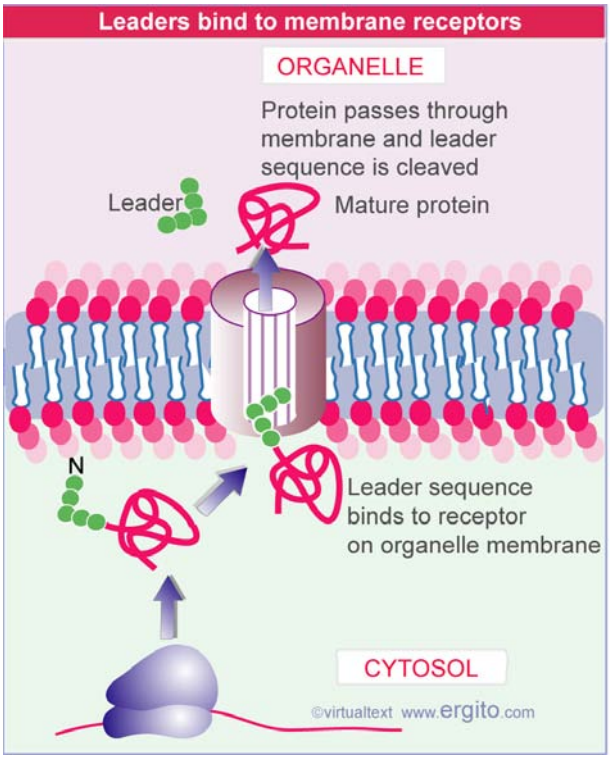


Figure 8.38 Leader sequences allow proteins to recognize mitochondrial or chloroplast surfaces by a post-translational process.

The leaders of proteins imported into mitochondria and chloroplasts usually have both hydrophobic and basic amino acids. They consist of stretches of uncharged amino acids interrupted by basic amino acids, and they lack acidic amino acids. There is little other homology. An example is given in **Figure 8.39**. Recognition of the leader does not depend on its exact sequence, but rather on its ability to form an amphipathic helix, in which one face has hydrophobic amino acids, and the other face presents the basic amino acids.

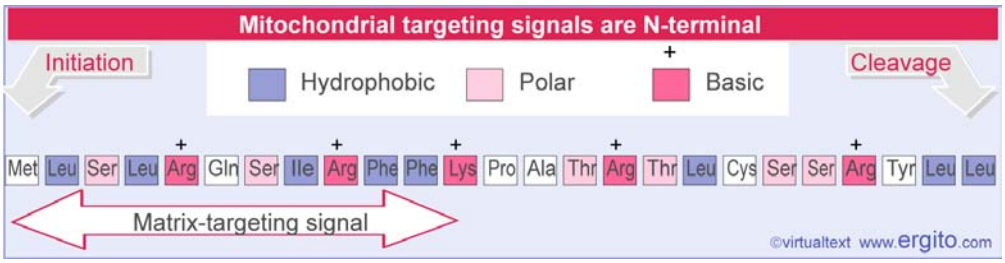


Figure 8.39 The leader sequence of yeast cytochrome *c* oxidase subunit IV consists of 25 neutral and basic amino acids. The first 12 amino acids are sufficient to transport any attached polypeptide into the mitochondrial matrix.

The leader sequence contains all the information needed to localize an organelle protein. The ability of a leader sequence can be tested by constructing an artificial protein in which a leader from an organelle protein is joined to a cytosolic protein. The experiment is performed by constructing a hybrid gene, which is then translated

into the hybrid protein.

Several leader sequences have been shown by such experiments to function independently to target any attached sequence to the mitochondrion or chloroplast. For example, if the leader sequence given in **Figure 8.39** is attached to the cytosolic protein DHFR (dihydrofolate reductase), the DHFR becomes localized in the mitochondrion.

The leader sequence and the transported protein represent domains that fold independently. Irrespective of the sequence to which it is attached, the leader must be able to fold into an appropriate structure to be recognized by receptors on the organelle envelope. The attached polypeptide sequence plays no part in recognition of the envelope (for review see 51; 59).

What restrictions are there on transporting a hydrophilic protein through the hydrophobic membrane? An insight into this question is given by the observation that methotrexate, a ligand for the enzyme DHFR, blocks transport into mitochondria of DHFR fused to a mitochondrial leader. The tight binding of methotrexate prevents the enzyme from unfolding when it is translocated through the membrane. So although the sequence of the transported protein is irrelevant for targeting purposes, in order to follow its leader through the membrane, it requires the flexibility to assume an unfolded conformation (455).

Hydrolysis of ATP is required both outside and inside for translocation across the membrane. It may be involved with pushing the protein from outside and pulling from inside. In the cases of mitochondrial import and bacterial export, there is also a requirement for an electrochemical potential across the inner membrane to transfer the amino terminal part of the leader.

Last updated on 8-8-2002

Reviews

51. Baker, K. P. and Schatz, G. (1991). *Mitochondrial proteins essential for viability mediate protein import into yeast mitochondria*. Nature 349, 205-208.

59. Schatz, G. and Dobberstein, B. (1996). *Common principles of protein translocation across membranes.* Science 271, 1519-1526.

References

455. Eilers, M. and Schatz, G. (1986). *Binding of a specific ligand inhibits import of a purified precursor protein into mitochondria*. Nature 322, 228-232.

PROTEIN LOCALIZATION

2.8.18 A hierarchy of sequences determines location within organelles

Key Concepts

- The N-terminal part of a leader sequence targets a protein to the mitochondrial matrix or chloroplast lumen.
- An adjacent sequence can control further targeting, to a membrane or the intermembrane spaces.
- The sequences are cleaved successively from the protein

The mitochondrion is surrounded by an envelope consisting of two membranes. Proteins imported into mitochondria may be located in the outer membrane, the intermembrane space, the inner membrane, or the matrix. A protein that is a component of one of the membranes may be oriented so that it faces one side or the other.

What is responsible for directing a mitochondrial protein to the appropriate compartment? The "default" pathway for a protein imported into a mitochondrion is to move through both membranes into the matrix. This property is conferred by the N-terminal part of the leader sequence. A protein that is localized within the intermembrane space or in the inner membrane itself requires an additional signal, which specifies its destination within the organelle. A multipart leader contains signals that function in a hierarchical manner, as summarized in **Figure 8.40**. The first part of the leader targets the protein to the organelle, and the second part is required if its destination is elsewhere than the matrix. The two parts of the leader are removed by successive cleavages.

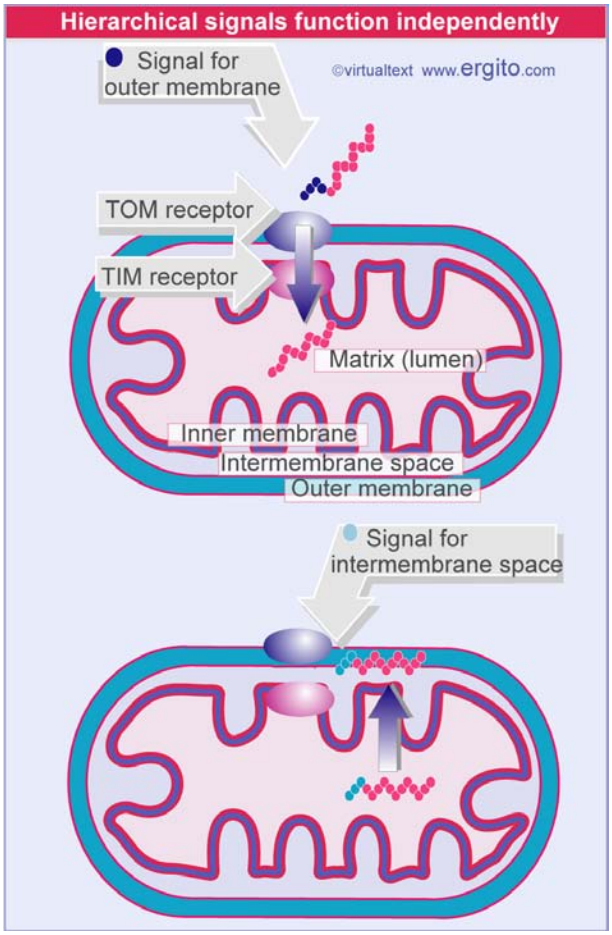


Figure 8.40 Mitochondria have receptors for protein transport in the outer and inner membranes. Recognition at the outer membrane may lead to transport through both receptors into the matrix, where the leader is cleaved. If it has a membrane-targeting signal, it may be re-exported.

Cytochrome *c1* is an example. It is bound to the inner membrane and faces the intermembrane space. Its leader sequence consists of 61 amino acids, and can be divided into regions with different functions. The sequence of the first 32 amino acids alone, or even the N-terminal half of this region, can transport DHFR all the way into the matrix. So the first part of the leader sequence (32 N-terminal amino acids) comprises a matrix-targeting signal. But the intact leader transports an attached sequence – such as murine DHFR – into the intermembrane space.

What prevents the protein from proceeding past the intermembrane space when it has an intact leader? The region following the matrix-targeting signal (comprising 19 amino acids of the leader) provides another signal that localizes the protein at the inner membrane or within the intermembrane space. For working purposes, we call this the membrane-targeting signal.

The two parts of a leader that contains both types of signal have different compositions. As indicated in **Figure 8.41**, the 35 N-terminal amino acids resemble other organelle leader sequences in the high content of uncharged amino acids,

punctuated by basic amino acids. The next 19 amino acids, however, comprise an uninterrupted stretch of uncharged amino acids, long enough to span a lipid bilayer. This membrane-targeting signal resembles the sequences that are involved in protein translocation into membranes of the endoplasmic reticulum (see *Molecular Biology 2.8.8 Signal sequences initiate translocation*).

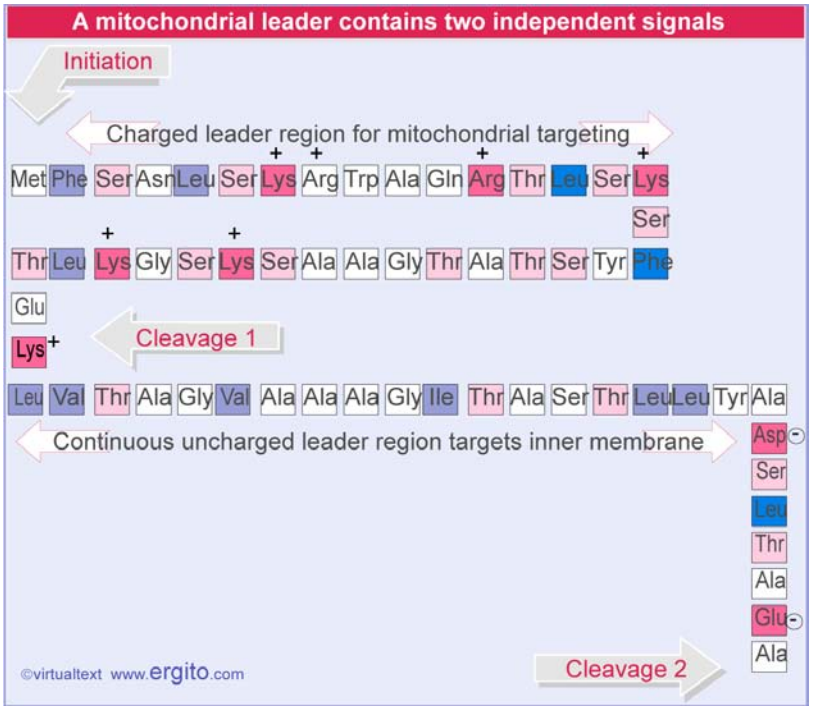


Figure 8.41 The leader of yeast cytochrome *c*1 contains an N-terminal region that targets the protein to the mitochondrion, followed by a region that targets the (cleaved) protein to the inner membrane. The leader is removed by two cleavage events.

Cleavage of the matrix-targeting signal is the sole processing event required for proteins that reside in the matrix. This signal must also be cleaved from proteins that reside in the intermembrane space; but following this cleavage, the membrane-targeting signal (which is now the N-terminal sequence of the protein) directs the protein to its destination in the outer membrane, intermembrane space, or inner membrane. Then it in turn is cleaved.

The N-terminal matrix-targeting signal functions in the same manner for all mitochondrial proteins. Its recognition by a receptor on the outer membrane leads to transport through the two membranes. And the same protease is involved in cleaving the matrix-targeting signal, irrespective of the final destination of the protein. This protease is a water soluble, Mg^{2+}-dependent enzyme that is located in the matrix. So the N-terminal sequence must reach the matrix, even if the protein ultimately will reside in the intermembrane space.

Residence in the matrix occurs in the absence of any other signal. If there is a membrane-targeting signal, however, it is activated by cleavage of the matrix-targeting signal. Then the remaining part of the leader (which is now N-terminal) causes the protein to take up its final destination.

The nature of the membrane-targeting signal is controversial. One model holds that the entire protein enters the matrix, after which the membrane-targeting signal causes it to be re-exported into or through the inner membrane. An alternative model proposes that the membrane-targeting sequence simply prevents the rest of the protein from following the leader through the inner membrane into the matrix. Whichever model applies, another protease (located within the intermembrane space) completes the removal of leader sequences (453; 454).

Passage through chloroplast membranes is achieved in a similar manner. **Figure 8.42** illustrates the variety of locations for chloroplast proteins. They pass the outer and inner membranes of the envelope into the stroma, a process involving the same types of passage as into the mitochondrial matrix. But some proteins are transported yet further, across the stacks of the thylakoid membrane into the lumen. Proteins destined for the thylakoid membrane or lumen must cross the stroma en route.

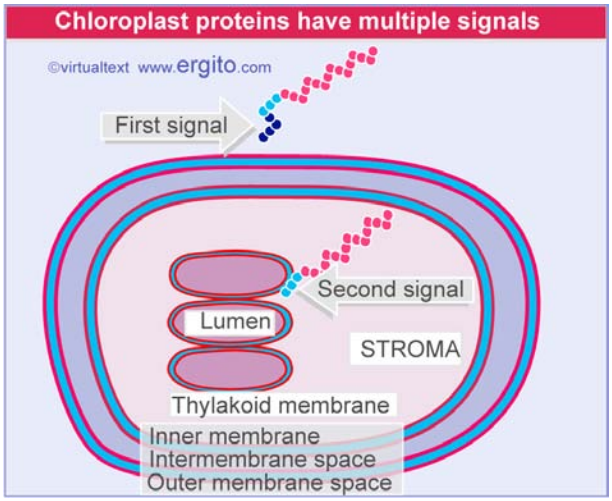


Figure 8.42 A protein approaches the chloroplast from the cytosol with a ~50 residue leader. The N-terminal half of the leader sponsors passage into the envelope or through it into the stroma. Cleavage occurs during envelope passage.

Chloroplast targeting signals resemble mitochondrial targeting signals. The leader consists of ~50 amino acids, and the N-terminal half is needed to recognize the chloroplast envelope. A cleavage between positions 20-25 occurs during or following passage across the envelope, and proteins destined for the thylakoid membrane or lumen have a new N-terminal leader that guides recognition of the thylakoid membrane. There are several (at least four) different systems in the chloroplast that catalyze import of proteins into the thylakoid membrane (for review see 64).

The general principle governing protein transport into mitochondria and chloroplasts therefore is that the N-terminal part of the leader targets a protein to the organelle matrix, and an additional sequence (within the leader) is needed to localize the protein at the outer membrane, intermembrane space, or inner membrane.

Reviews

64. Cline, K. and Henry, R. (1996). *Import and routing of nucleus-encoded chloroplast proteins*. Annu. Rev. Cell Dev. Biol. 12, 1-26.

References

453. van Loon, A. P. G. M. et al. (1986). *The presequences of two imported mitochondrial proteins contain information for intracellular and intramitochondrial sorting*. Cell 44, 801-812.

454. Hartl, F. U., Ostermann, J., Guiard, B., and Neupert, W. (1987). *Successive translocation into and out of the mitochondrial matrix: targeting of proteins to the intermembrane space by a bipartite signal peptide*. Cell 51, 1027-1037.

PROTEIN LOCALIZATION

2.8.19 Inner and outer mitochondrial membranes have different translocons

Key Terms

The **TOM complex (TOM)** resides in the outer membrane of the mitochondrion and is responsible for importing proteins from the cytosol into the space between the membranes.

The **TIM complex (TIM)** resides in the inner membrane of mitochondria and is responsible for transporting proteins from the intermembrane space into the interior of the organelle.

Key Concepts

- Transport through the outer and inner mitochondrial membranes uses different receptor complexes.
- The TOM (outer membrane) complex is a large complex in which substrate proteins are directed to the Tom40 channel by one of two subcomplexes.
- Different TIM (inner membrane) complexes are used depending on whether the substrate protein is targeted to the inner membrane or to the lumen.
- Proteins pass directly from the TOM to the TIM complex.

There are different receptors for transport through each membrane in the chloroplast and mitochondrion. In the chloroplast they are called TOC and TIC, and in the mitochondrion they are called **TOM** and **TIM**, referring to the outer and inner membranes, respectively (for review see 2353).

The TOM complex consists of ~9 proteins. many of which are integral membrane proteins. A general model for the complex is shown in **Figure 8.43**. The TOM aggregate has a size of >500 kD, with a diameter of ~138 Å, and forms an ion-conducting channel. A complex contains 2-3 individual rings of diameter 75 Å, each with a pore of diameter 20 Å.

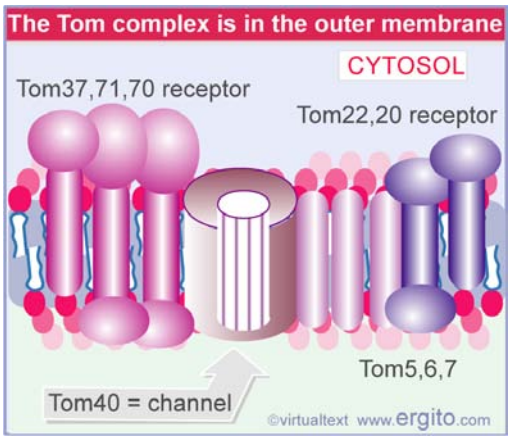


Figure 8.43 TOM proteins form receptor complex(es) that are needed for translocation across the mitochondrial outer membrane.

Tom40 is deeply imbedded in the membrane and provides the channel for translocation. It contacts preproteins as they pass through the outer membrane. It binds to three smaller proteins, Tom5,6,7, which may be components of the channel or assembly factors. There are two subcomplexes that provide surface receptors. Tom20,22 form a subcomplex with exposed domains in the cytosol. Most proteins that are imported into mitochondria are recognized by the Tom20,22 subcomplex, which is the primary receptor and recognizes the N-terminal sequence of the translocating protein. Tom37,70,71 provides a receptor for a smaller number of proteins that have internal targeting sequences.

When a protein is translocated through the TOM complex, it passes from a state in which it is exposed to the cytosol into a state in which it is exposed to the intermembrane space. However, it is not usually released, but instead is transferred directly to the TIM complex. It is possible to trap intermediates in which the leader is cleaved by the matrix protease, while a major part of the precursor remains exposed on the cytosolic surface of the envelope. This suggests that a protein spans the two membranes during passage. The TOM and TIM complexes do not appear to interact directly (or at least do not form a detectable stable complex), and they may therefore be linked simply by a protein in transit. When a translocating protein reaches the intermembrane space, the exposed residues may immediately bind to a TIM complex, while the rest of the protein continues to translocate through the TOM complex (for review see 65).

There are two TIM complexes in the inner membrane.

The Tim17-23 complex translocates proteins to the lumen. Substrates are recognized by their possession of a positively charged N-terminal signal. Tim17-23 are transmembrane proteins that comprise the channel. **Figure 8.44** shows that they are associated with Tim44 on the matrix side of the membrane. Tim 44 in turn binds the chaperone Hsp70. This is also associated with another chaperone, Mge, the counterpart to bacterial GrpE. This association ensures that when the imported protein reaches the matrix, it is bound by the Hsp70 chaperone. The high affinity of Hsp70 for the unfolded conformation of the protein as it emerges from the inner membrane helps to "pull" the protein through the channel (for review see 3436).

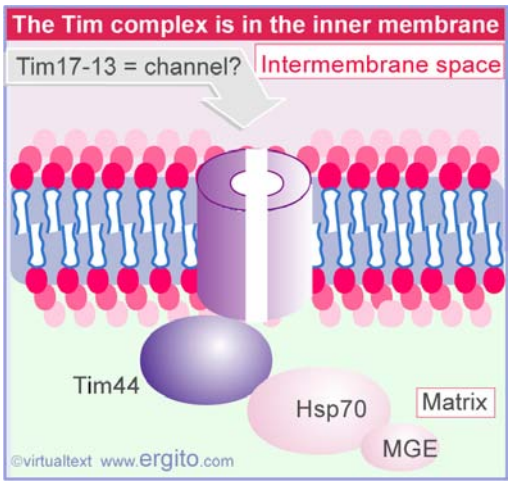


Figure 8.44 Tim proteins form the complex for translocation across the mitochondrial inner membrane.

A major chaperone activity in the mitochondrial matrix is provided by Hsp60 (which forms the same sort of structure as its counterpart GroEL). Association with Hsp60 is necessary for joining of the subunits of imported proteins that form oligomeric complexes. An imported protein may be "passed on" from Hsp70 to Hsp60 in the process of acquiring its proper conformation (458).

The Tim22-54 complex translocates proteins that reside in the inner membrane.

How does a translocating protein finds its way from the TOM complex to the appropriate TIM complex? Two protein complexes in the intermembrane space escort a translocating protein from TIM to TOM. The Tim9-10 and Tim8-13 complexes act as escorts for different sets of substrate proteins (944). Tim9-10 may direct its substrates to either Tim22-54 or Tim23-17, while Tim8-13 directs substrates only to Tim22-54. Some substrates do not use either Tim9-10 or Tim8-13, so other pathways must also exist. The pathways are summarized in **Figure 8.45**.

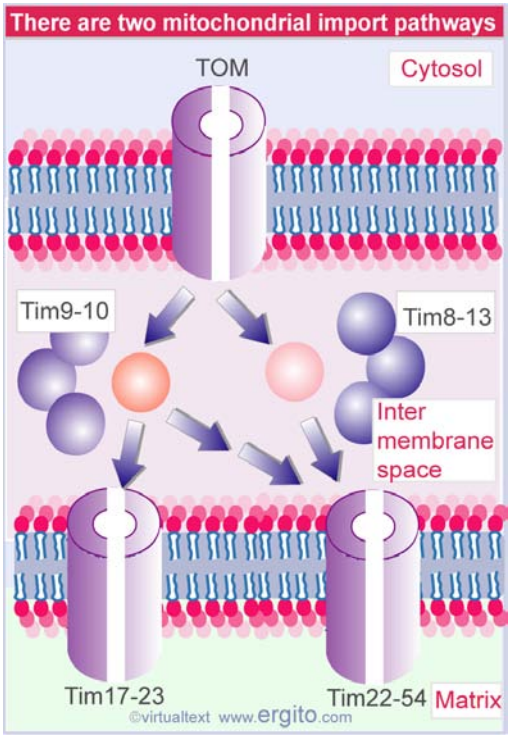


Figure 8.45 Tim9-10 takes proteins from TOM to either TIM complex, and Tim8-13 takes proteins to Tim22-54.

What is the role of the escorting complexes? They may be needed to help the protein exit from the TOM complex as well as for recognizing the TIM complex. **Figure 8.46** shows that a translocating protein may pass directly from the TOM channel to the Tim9,10 complex, and then into the Tim22-54 channel

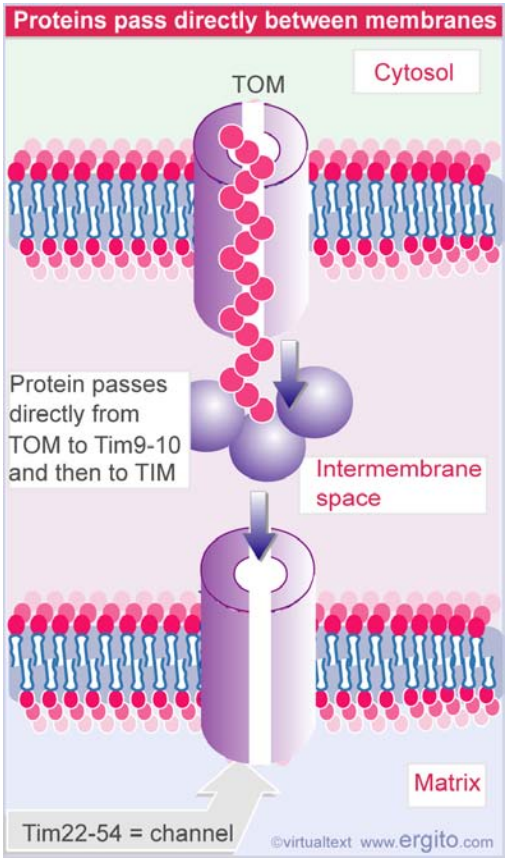


Figure 8.46 A translocating protein may be transferred directly from TOM to Tim22-54.
This is a static version of an interactive figure; see http://www.ergito.com/main.jsp?bcs=MBIO.2.8.19 to view properly.

A mitochondrial protein folds under different conditions before and after its passage through the membrane. Ionic conditions and the chaperones that are present are different in the cytosol and in the mitochondrial matrix. It is possible that a mitochondrial protein can attain its mature conformation *only* in the mitochondrion.

Last updated on 2-2-2000

Reviews

65. Neupert, W. (1997). *Protein import into mitochondria*. Annu. Rev. Biochem. 66, 863-917.

2353. Dalbey, R. E. and Kuhn, A. (2000). *Evolutionarily related insertion pathways of bacterial, mitochondrial, and thylakoid membrane proteins.* Annu. Rev. Cell Dev. Biol. 16, 51-87.

3436. Neupert, W. and Brunner, M. (2002). *The protein import motor of mitochondria.* Nat. Rev. Mol. Cell Biol. 3, 555-565.

References

458. Ostermann, J., Horwich, A. L., Neupert, W., and Hartl, F. U. (1989). *Protein folding in mitochondria requires complex formation with hsp60 and ATP hydrolysis.* Nature 341, 125-130.

944. Leuenberger, D., Bally, N. A., Schatz, G., and Koehler, C. M. (1999). *Different import pathways through the mitochondrial intermembrane space for inner membrane proteins.* EMBO J. 18, 4816-4822.

PROTEIN LOCALIZATION

2.8.20 Peroxisomes employ another type of translocation system

Key Terms

The **peroxisome** is an organelle in the cytoplasm enclosed by a single membrane. It contains oxidizing enzymes.

Peroxins are the protein components of the peroxisome.

Key Concepts

- Proteins are imported into peroxisomes in their fully folded state.
- They have either a PTS1 sequence at the C-terminus or a PTS2 sequence at the N-terminus.
- The receptor Pex5p binds the PTS1 sequence, and the receptor Pex7p binds the PTS2 sequence.
- The receptors are cytosolic proteins that shuttle into the peroxisome carrying a substrate protein and then return to the cytosol.

Peroxisomes are small bodies (05-1.5 μm diameter) enclosed by a single membrane. They contain enzymes concerned with oxygen utilization, which convert oxygen to hydrogen peroxide by removing hydrogen atoms from substrates. Catalase then uses the hydrogen peroxide to oxidize a variety of other substrates. Their activities are crucial for the cell. Since the fatal disease of Zellweger syndrome was found to be caused by lack of peroxisomes (2316), >15 human diseases have been linked to disorders in peroxisome function.

All of the components of the peroxisome are imported from the cytosol. Proteins that are required for peroxisome formation are called **peroxins**. 23 genes coding for peroxins have been identified, and human peroxisomal diseases have been mapped to 12 complementation groups, most identified with specific genes. Peroxisomes appear to be absent from cells that have null mutations in some of these genes. In some of these cases, introduction of a wild-type gene leads to the reappearance of peroxisomes (2317; 2318). It has generally been assumed that, like other membrane-bounded organelles, peroxisomes can arise only by duplication of pre-existing peroxisomes. But these results raised the question of whether it might be possible to assemble them *de novo* from their components. In at least some cases, however, the absence of peroxins leaves the cells with peroxisomal ghosts – empty membrane bodies. Even when they cannot be easily seen, it is hard to exclude the possibility that there is some remnant that serves to regenerate the peroxisomes (for review see 2305).

Transport of proteins to peroxisomes occurs post-translationally. Proteins that are imported into the matrix have either of two short sequences, called PTS1 and PTS2.

The PTS1 signal is a tri- or tetrapeptide at the C-terminus. It was originally characterized as the sequence SKL (Ser-Lys-Leu) (2319), but now a large variety of sequences have been shown to act as a PTS1 signal (2320). The addition of a suitable sequence to the C-terminus of cytosolic proteins is sufficient to ensure their import into the organelle. The PTS2 signal is a sequence of 9 amino acids, again with much diversity, and this can be located near the N-terminus or internally (for review see 1962). It is possible there may be a third type of sequence called PTS3.

Several peroxisomal proteins are necessary for the import of proteins from the cytosol. The peroxisomal receptors that bind the two types of signals are called Pex5p and Pex7p, respectively. The other proteins are part of membrane-associated complexes concerned with the translocation reaction.

Transport into the peroxisome has unusual features that mark important differences from the system used for transport into other organelles.

Proteins can be imported into the peroxisome in their mature, fully-folded state (1961). This contrasts with the requirement to unfold a protein for passage into the ER or mitochondrion, where it passes through a channel in the membrane into the organelle in something akin to an unfolded thread of amino acids. It is not clear how the structure of a preexisting channel could expand to permit this. One possibility is to resurrect an old idea and to suppose that the channel assembles around the substrate protein when it associates with the membrane.

The Pex5p and Pex7p receptors are not integral membrane proteins, but are largely cytosolic, with only a small proportion associated with peroxisomes (2321; 2322). They behave in the same way, cycling between the peroxisome and the cytosol. **Figure 8.47** shows that the receptor binds a substrate protein in the cytosol, takes it to the peroxisome, moves with it through the membrane into the interior, and then returns to the cytosol to undertake another cycle. This shuttling behavior resembles the carrier system for import into the nucleus (see *Molecular Biology 2.8.28 Transport receptors carry cargo proteins through the pore*).

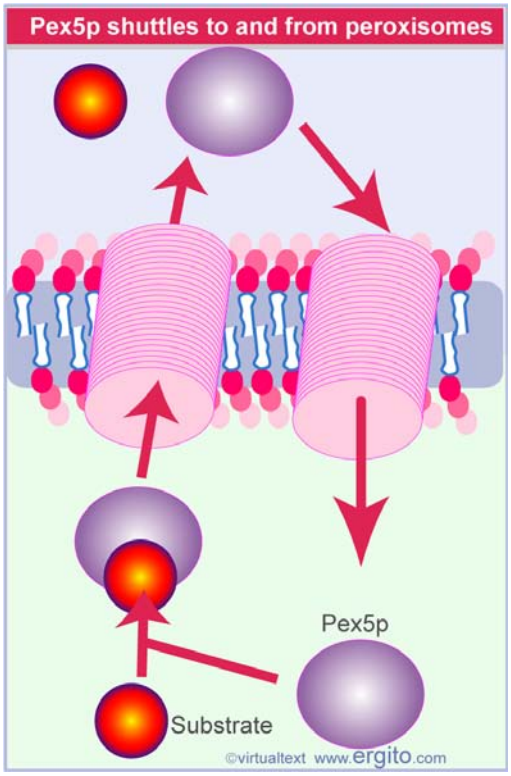


Figure 8.47 The Pex5p receptor binds a substrate protein in the cytosol, carries it across the membrane into the peroxisome, and then returns to the cytosol.

The import pathways converge at the perixosomal membrane, where Pex5p and Pex7p both interact with the same membrane protein complex, consisting of Pex14p and Pex13p. The receptors dock with this complex, and then several other peroxins are involved with the process of transport into the lumen. The details of the transport process are not yet clear.

Proteins that are incorporated into the peroxisomal membrane have a sequence called the mPTS, but little is known about the process of integration. Pex3p may be a key protein, because in its absence other proteins are not found in peroxisomal membranes. Pex3p has its own mPTS, which raises the question of how it enters the membrane. Perhaps it interacts with Pex3p that is already in the membrane (for review see 2305). This bears on the question of whether peroxisomes can ever assemble *de novo*.

Last updated on 1-22-2002

Reviews

1962. Subramani, S., Koller, A., and Snyder, W. B. (2000). *Import of peroxisomal matrix and membrane proteins.* Annu. Rev. Biochem. 69, 399-418.

2305. Purdue, P. E. and Lazarow, P. B. (2001). *Peroxisome biogenesis.* Annu. Rev. Cell Dev. Biol. 17, 701-752.

References

1961. Walton, P. A., Hill, P. E., and Hill, S. (1995). *Import of stably folded proteins into peroxisomes.* Mol. Biol. Cell 6, 675-683.

2316. Goldfischer, S., Moore, C. L., Johnson, A. B., Spiro, A. J., Valsamis, M. P., Wisniewski, H. K., Ritch, R. H., Norton, W. T., Rapin, I., and Gartner, L. M. (1973). *Peroxisomal and mitochondrial defects in the cerebro-hepato-renal syndrome.* Science 182, 62-64.

2317. Matsuzono, Y., Kinoshita, N., Tamura, S., Shimozawa, N., Hamasaki, M., Ghaedi, K., Wanders, R. J., Suzuki, Y., Kondo, N., and Fujiki, Y. (1999). *Human PEX19: cDNA cloning by functional complementation, mutation analysis in a patient with Zellweger syndrome, and potential role in peroxisomal membrane assembly.* Proc. Natl. Acad. Sci. USA 96, 2116-2121.

2318. South, S. T. and Gould, S. J. (1999). *Peroxisome synthesis in the absence of preexisting peroxisomes.* J. Cell Biol. 144, 255-266.

2319. Gould, S. J., Keller, G. A., Hosken, N., Wilkinson, J., and Subramani, S. (1989). *A conserved tripeptide sorts proteins to peroxisomes.* J. Cell Biol. 108, 1657-1664.

2320. Elgersma, Y., Vos, A., van den Berg, M., van Roermund, C. W., van der Sluijs, P., Distel, B., and Tabak, H. F. (1996). *Analysis of the carboxyl-terminal peroxisomal targeting signal 1 in a homologous context in S. cerevisiae.* J. Biol. Chem. 271, 26375-26382.

2321. Dodt, G. and Gould, S. J. (1996). *Multiple PEX genes are required for proper subcellular distribution and stability of Pex5p, the PTS1 receptor: evidence that PTS1 protein import is mediated by a cycling receptor.* J. Cell Biol. 135, 1763-1774.

2322. Elgersma, Y., Elgersma-Hooisma, M., Wenzel, T., McCaffery, J. M., Farquhar, M. G., and Subramani, S. (1998). *A mobile PTS2 receptor for peroxisomal protein import in Pichia pastoris.* J. Cell Biol. 140, 807-820.

PROTEIN LOCALIZATION

2.8.21 Bacteria use both co-translational and post-translational translocation

Key Terms

The **periplasm** (or periplasmic space) is the region between the inner and outer membranes in the bacterial envelope.

Signal peptidase is an enzyme within the membrane of the ER that specifically removes the signal sequences from proteins as they are translocated. Analogous activities are present in bacteria, archaebacteria, and in each organelle in a eukaryotic cell into which proteins are targeted and translocated by means of removable targeting sequences. Signal peptidase is one component of a larger protein complex.

Key Concepts

- Bacterial proteins that are exported to or through membranes use both post-translational and co-translational mechanisms.

The bacterial envelope consists of two membrane layers. The space between them is called the **periplasm**. Proteins are exported from the cytoplasm to reside in the envelope or to be secreted from the cell. The mechanisms of secretion from bacteria are similar to those characterized for eukaryotic cells, and we can recognize some related components. **Figure 8.48** shows that proteins that are exported from the cytoplasm have one of four fates:

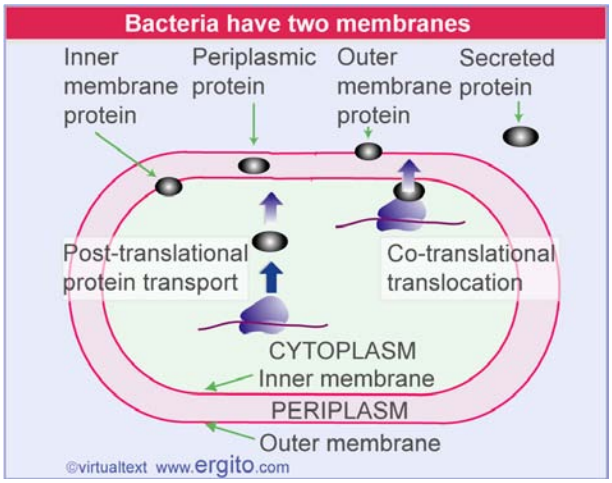


Figure 8.48 Bacterial proteins may be exported either post-translationally or co-translationally, and may be located within either membrane or the periplasmic space, or may be secreted.

- to be inserted into the inner membrane
- to be translocated through the inner membrane to rest in the periplasm
- to be inserted into the outer membrane
- to be translocated through the outer membrane into the medium.

Different protein complexes in the inner membrane are responsible for transport of proteins depending on whether their fate is to pass through or stay within the inner membrane. This resembles the situation in mitochondria, where different complexes in each of the inner and outer membranes handle different subsets of protein substrates depending on their destinations (see *Molecular Biology 2.8.17 Post-translational membrane insertion depends on leader sequences*) A difference from import into organelles is that transfer in *E. coli* may be either co- or post-translational. Some proteins are secreted both co-translationally and post-translationally, and the relative kinetics of translation versus secretion through the membrane could determine the balance.

Exported bacterial proteins have N-terminal leader sequences, with a hydrophilic N-terminus and an adjacent hydrophobic core. The leader is cleaved by a **signal peptidase** that recognizes precursor forms of several exported proteins. The signal peptidase is an integral membrane protein, located in the inner membrane. Mutations in N-terminal leaders prevent secretion; they are suppressed by mutations in other genes, which are thus defined as components of the protein export apparatus. Several genes given the general description *sec* are implicated in coding for components of the secretory apparatus by the occurrence of mutations that block secretion of many or all exported proteins.

PROTEIN LOCALIZATION

2.8.22 The *Sec* system transport proteins into and through the inner membrane

Key Concepts

- The bacterial SecYEG translocon in the inner membrane is related to the eukaryotic Sec61 translocon.
- Various chaperones are involved in directing secreted proteins to the translocon.

There are several systems for transport through the inner membrane. The best characterized is the Sec system, whose components are shown in **Figure 8.49**. The translocon that is embedded in the membrane consists of three subunits that are related to the components of mammalian/yeast Sec61 (for review see 45; 47). Each of the subunits is an integral transmembrane protein. (SecY has 10 transmembrane segments and SecE has 3 transmembrane segments.) The functional translocon is a trimer with one copy of each subunit (1066). The major pathway for directing proteins to the translocon consists of SecB and SecA. SecB is a chaperone that binds to the nascent protein to control its folding. It transfers the protein to SecA, which in turn transfers it to the translocon.

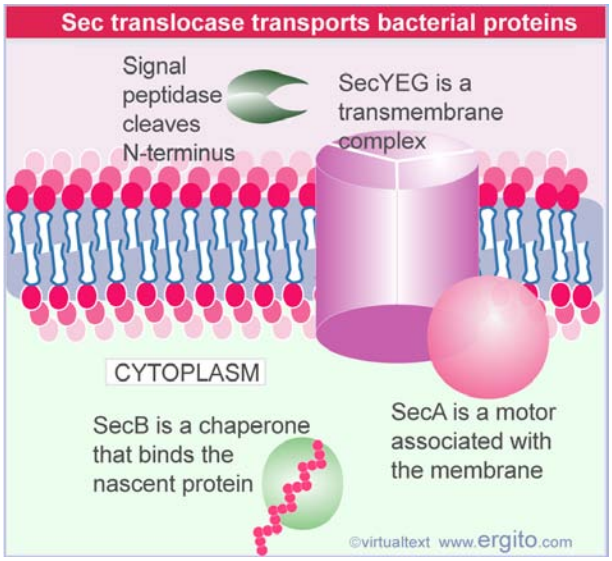


Figure 8.49 The Sec system has the SecYEG translocon embedded in the membrane, the SecA associated protein that pushes proteins through the channel, the SecB chaperone that transfers nascent proteins to SecA, and the signal peptidase that cleaves the N-terminal signal from the translocated protein.

Figure 8.50 shows that there are two predominant ways of directing proteins to the Sec channel:

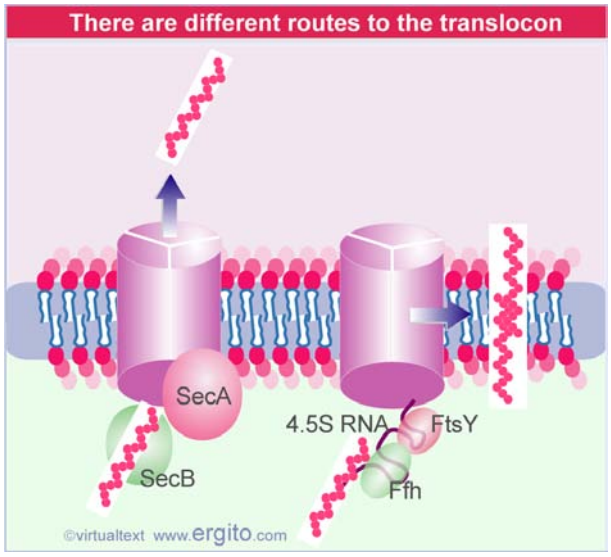


Figure 8.50 SecB/SecA transfer proteins to the translocon in order for them to pass through the membrane. 4.4S RNA transfers proteins that enter the membrane.

- the SecB chaperone;
- and the 4.5S RNA-based SRP (2357).

Several chaperones can increase the efficiency of bacterial protein export by preventing premature folding; they include "trigger factor" (characterized as a chaperone that assists export), GroEL (see earlier), and SecB (identified as the product of one of the *sec* mutants). Although SecB is the least abundant of these proteins, it has the major role in promoting export. It has two functions. First, it behaves as a chaperone and binds to a nascent protein to retard folding. It cannot reverse the change in structure of a folded protein, so it does not function as an unfolding factor. Its role is therefore to inhibit improper folding of the newly synthesized protein. Second, it has an affinity for the protein SecA. This allows it to target a precursor protein to the membrane (456; 457). The SecB-SecYEG pathway is used for translocation of proteins that are secreted into the periplasm and is summarized in**Figure 8.51**.

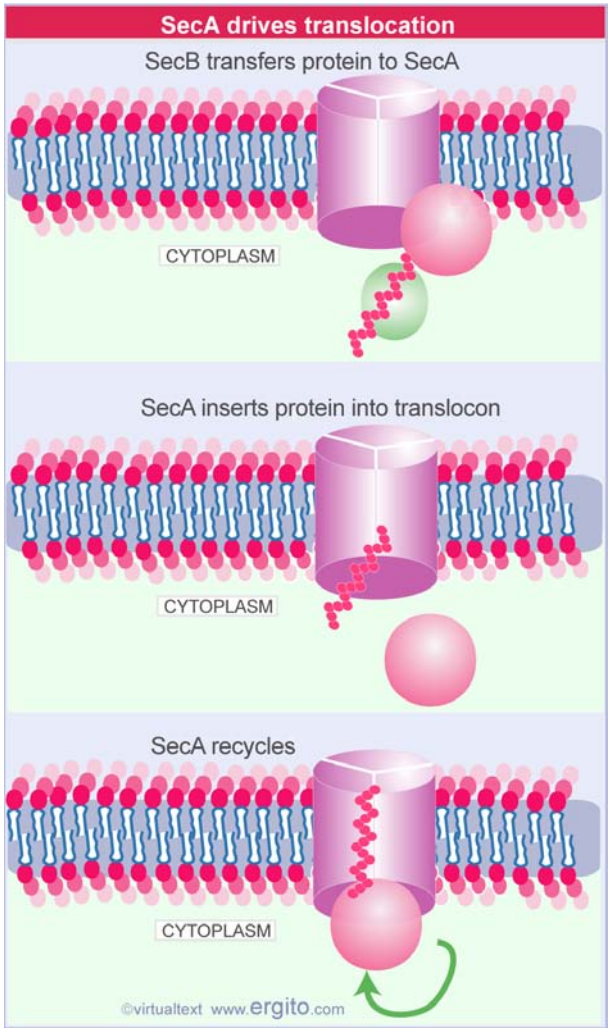


Figure 8.51 SecB transfers a nascent protein to SecA, which inserts the protein into the channel. Translocation requires hydrolysis of ATP and a protonmotive force. SecA undergoes cycles of association and dissociation with the channel and provides the motive force to push the protein through.

SecA is a large peripheral membrane protein that has alternative ways to associate with the membrane. As a peripheral membrane protein, it associates with the membrane by virtue of its affinity for acidic lipids and for the SecY component of the translocon., which are part of a multisubunit complex that provides the translocase function. However, in the presence of other proteins (SecD and SecF), SecA can be found as a membrane-spanning protein. It probably provides the motor that pushes the substrate protein through the SecYEG translocon.

SecA recognizes both SecB and the precursor protein that it chaperones; probably features of the mature protein sequence as well as its leader are required for recognition. SecA has an ATPase activity that depends upon binding to lipids, SecY, and a precursor protein. The ATPase functions in a cyclical manner during translocation. After SecA binds a precursor protein, it binds ATP, and ~20 amino acids are translocated through the membrane. Hydrolysis of ATP is required to

release the precursor from SecA. Then the cycle may be repeated. Precursor protein is bound again to provide the spur to bind more ATP, translocate another segment of protein, and release the precursor. SecA may alternate between the peripheral and integral membrane forms during translocation; with each cycle, a 30 kD domain of SecA may insert into the membrane and then retract (471).

Another process can also undertake translocation. When a precursor is released by SecA, it can be driven through the membrane by a protonmotive force (that is, an electrical potential across the membrane). This process cannot initiate transfer through the membrane, but can continue the process initiated by a cycle of SecA ATPase action. So after or between cycles of the SecA-ATP driven reaction, the protonmotive force can drive translocation of the precursor.

The *E. coli* ribonucleoprotein complex of 4.5S RNA with Ffh and FtsY proteins is a counterpart to the eukaryotic SRP (see *Molecular Biology 2.8.10 The SRP interacts with the SRP receptor*). It probably plays the role of keeping the nascent protein in an appropriate conformation until it interacts with other components of the secretory apparatus. It is needed for the secretion of some, but not all, proteins. As we see in **Figure 8.50**, its substrates are integral membrane proteins (1065). The basis for differential selection of substrates is that the *E. coli* SRP recognizes an anchor sequence in the protein (anchor sequences by definition are present only in integral membrane proteins). Chloroplasts have counterparts to the Ffh and FtsY proteins, but do not require an RNA component.

Last updated on 2-22-2002

Reviews

45. Oliver, D. (1985). *Protein secretion in E. coli.* Annu. Rev. Immunol. 39, 615-648.

47. Lee, C. and Beckwith, J. (1986). *Cotranslational and posttranslational protein translocation in prokaryotic systems.* Annu. Rev. Cell Biol. 2, 315-336.

References

456. Collier, D. N. et al. (1988). *The antifolding activity of SecB promotes the export of the E. coli maltose-binding protein*. Cell 53, 273-283.

457. Crooke, E. et al. (1988). *ProOmpA is stabilized for membrane translocation by either purified E. coli trigger factor or canine signal recognition particle*. Cell 54, 1003-1011.

471. Brundage, L. et al. (1990). *The purified E. coli integral membrane protein SecY/E is sufficient for reconstitution of SecA-dependent precursor protein translocation*. Cell 62, 649-657.

1065. Beck, K., Wu, L. F., Brunner, J., and Muller, M. (2000). *Discrimination between SRP- and SecA/SecB-dependent substrates involves selective recognition of nascent chains by SRP and trigger factor.* EMBO J. 19, 134-143.

1066. Yahr, T. L. and Wickner, W. T. (2000). *Evaluating the oligomeric state of SecYEG in preprotein translocase.* EMBO J. 19, 4393-4401.

2357. Valent, Q. A., Scotti, P. A., High, S., von Heijne, G., Lentzen, G., Wintermeyer, W., Oudega, B., and Luirink, J. (1998). *The E. coli SRP and SecB targeting pathways converge at the translocon.* EMBO J. 17, 2504-2512.

PROTEIN LOCALIZATION

2.8.23 *Sec*-independent translocation systems in *E. coli*

Key Concepts

- *E. coli* and organelles have related systems for protein translocation.
- One system allows certain proteins to insert into membranes without a translocation apparatus.
- YidC is homologous to a mitochondrial system for transferring proteins into the inner membrane.
- The tat system transfers proteins with a twin arginine motif into the periplasmic space.

The most striking alternative system for protein translocation in *E. coli* is revealed by the coat protein of phage M13. **Figure 8.52** shows that this does not appear to require any translocation apparatus! It can insert post-translationally into protein-free liposomes (2358). Targeting the protein to the membrane requires specific sequences (comprising basic residues) in the N- and C-terminal regions of the protein. They may interact with negatively charged heads of phospholipids. Then the protein enters the membrane by using hydrophobic groups in its N-terminal leader sequence and an internal anchor sequence. Hydrophobicity is the main driving force for translocation, but it can be assisted by a protonmotive force that is generated between the positively charged periplasmic side of the membrane and an acidic region in the protein. This drives the protein through the membrane, and leader peptidase can then cleave the N-terminal sequence. The generality of this mechanism in bacteria is unclear; it may apply only to the special case of bacteriophage coat proteins (for review see 2353). Some chloroplast proteins may insert into the thylakoid membrane by a similar pathway.

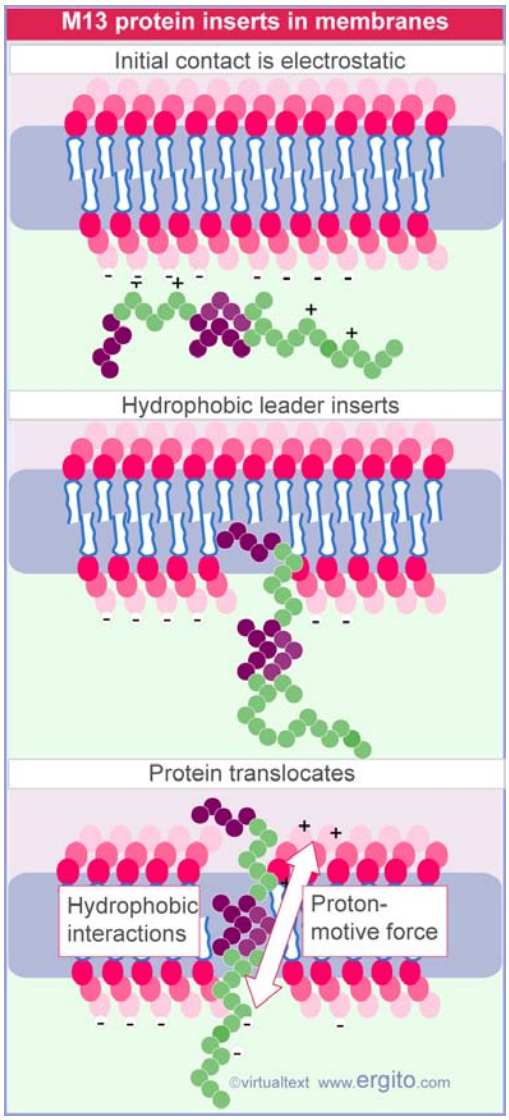


Figure 8.52 M13 coat protein inserts into the inner membrane by making an initial electrostatic contact, followed by insertion of hydrophobic sequences. Translocation is driven by hydrophobic interactions and a protonmotive force until the anchor sequence enters the membrane.

Mutations in the gene *yidC* block insertion of proteins into the inner membrane (1065; 1088). YidC is homologous to the protein Oxa1p that is required when proteins are inserted into the inner mitochondrial membrane from the matrix. It can function either independently of SecYEG or in conjunction with it. The insertion of some of the YidC-dependent proteins requires SecYEG, suggesting that YidC acts in conjunction with the translocon to divert the substrate into membrane insertion as opposed to secretion (2360). Other proteins whose insertion depends on YidC do not require SecYEG: it seems likely that some other (unidentified) functions are required instead of the translocon.

The *Tat* system is named for its ability to transport proteins bearing a twin arginine

targeting motif. It is responsible for translocation of proteins that have tightly bound cofactors. This may mean that they have limitations on their ability to unfold for passage through the membrane. This would be contrary to the principle of most translocation systems, where the protein passes through the membrane in an unfolded state, and then must be folded into its mature conformation after passage (for review see 1064). This system is related to a system in the chloroplast thylakoid lumen called Hcf106 (for review see 1064). Both of these systems transport proteins into the periplasm.

Last updated on 2-22-2002

Reviews

1064. Dalbey, R. E. and Robinson, C. (1999). *Protein translocation into and across the bacterial plasma membrane and the plant thylakoid membrane.* Trends Biochem. Sci. 24, 17-22.

2353. Dalbey, R. E. and Kuhn, A. (2000). *Evolutionarily related insertion pathways of bacterial, mitochondrial, and thylakoid membrane proteins.* Annu. Rev. Cell Dev. Biol. 16, 51-87.

References

1065. Beck, K., Wu, L. F., Brunner, J., and Muller, M. (2000). *Discrimination between SRP- and SecA/SecB-dependent substrates involves selective recognition of nascent chains by SRP and trigger factor.* EMBO J. 19, 134-143.

1088. Samuelson, J. C., Chen, M., Jiang, F., Moller, I., Wiedmann, M., Kuhn, A., Phillips, G. J., and Dalbey, R. E. (2000). *YidC mediates membrane protein insertion in bacteria.* Nature 406, 637-641.

2358. Soekarjo, M., Eisenhawer, M., Kuhn, A., and Vogel, H. (1996). *Thermodynamics of the membrane insertion process of the M13 procoat protein, a lipid bilayer traversing protein containing a leader sequence.* Biochemistry 35, 1232-1241.

2360. Scotti, P. A., Urbanus, M. L., Brunner, J., de Gier, J. W., von Heijne, G., van der Does, C., Driessen, A. J., Oudega, B., and Luirink, J. (2000). *YidC, theE. coli homologue of mitochondrial Oxa1p, is a component of the Sec translocase.* EMBO J. 19, 542-549.

PROTEIN LOCALIZATION

2.8.24 Pores are used for nuclear import and export

Key Terms

The **nuclear envelope** is a layer of two concentric membranes (inner and outer nuclear membranes) that surrounds the nucleus and its underlying intermediate filament lattice, the nuclear lamina. The nuclear envelope is penetrated by nuclear pores. The outer membrane is continuous with the membrane of the rough endoplasmic reticulum.

A **nuclear pore complex (NPC)** is a very large, proteinaceous structure that extends through the nuclear envelope, providing a channel for bidirectional transport of molecules and macromolecules between the nucleus and the cytosol.

Key Concepts

- The same nuclear pores are used for importing proteins into the nucleus and for exporting proteins and RNA from the nucleus.

The nucleus is segregated from the cytoplasm by a layer of two membranes that constitute the **nuclear envelope**. The inner membrane contacts the nuclear lamina, providing in effect a surface layer for the nucleus. The outer membrane is continuous with the endoplasmic reticulum in the cytosol. The space between the two membranes is continuous with the lumen of the endoplasmic reticulum. The two membranes come into contact at openings called **nuclear pore complexes**. At the center of each complex is a pore that provides a water-soluble channel between nucleus and cytoplasm. This means that the nucleus and cytosol have the same ionic milieu. There are ~3000 pore complexes on the nuclear envelope of an animal cell.

Transport between nucleus and cytoplasm proceeds in both directions. Since all proteins are synthesized in the cytosol, any proteins required in the nucleus must be transported there. Since all RNA is synthesized in the nucleus, the entire cytoplasmic complement of RNA (mRNA, rRNA, tRNA, and other small RNAs) must be derived by export from the nucleus. The nuclear pores are used for both import and export of material. **Figure 8.53** summarizes the frequency with which the pores are used for some of the more prominent substrates.

Nuclear pores work in both directions

Direction	Substrate	Passages /pore/min
Import	Histones	100
	Nonhistone proteins	100
	Ribosomal proteins	150
Export	Ribosomal subunits	~5
	mRNA	<1

Figure 8.53 Nuclear pores are used for import and export.

We can form an impression of the magnitude of import by considering the histones, the major protein components of chromatin. In a dividing cell, enough histones must be imported into the nucleus during the period of DNA synthesis to associate with a diploid complement of chromosomes. Since histones form about half the protein mass of chromatin, we may conclude that overall about 200 chromosomal protein molecules must be imported through each pore per minute.

Uncertainties about the processing and stability of mRNA make it more difficult to calculate the number of mRNA molecules exported, but to account for the ~250,000 molecules of mRNA per cell probably requires ~1 event per pore per minute. The major RNA synthetic activity of the nucleus is of course the production of rRNA, which is exported in the form of assembled ribosomal subunits. Just to double the number of ribosomes during one cell cycle would require the export of ~5 ribosomal subunits (60S and 40S) through each pore per minute.

For ribosomal proteins to assemble with the rRNA, they must first be imported into the nucleus. So ribosomal proteins must shuttle into the nucleus as free proteins and out again as assembled ribosomal subunits. Given ~80 proteins per ribosome, their import must be comparable in magnitude to that of the chromosomal proteins.

PROTEIN LOCALIZATION

2.8.25 Nuclear pores are large symmetrical structures

Key Concepts

- The nuclear pore is an annular structure with 8-fold symmetry.

How does a nuclear pore accommodate the transit of material of varied sizes and characteristics in either direction? Nuclear pore complexes have a uniform appearance when examined by microscopy. The pores can be released from the nuclear envelope by detergent, and **Figure 8.54** shows that they appear as annular structures, consisting of rosettes made of 8 spokes. **Figure 8.55** shows a model for the pore based on three-dimensional reconstruction of electron microscopic images. It consists of an upper ring and a lower ring, connected by a lattice of 8 structures.

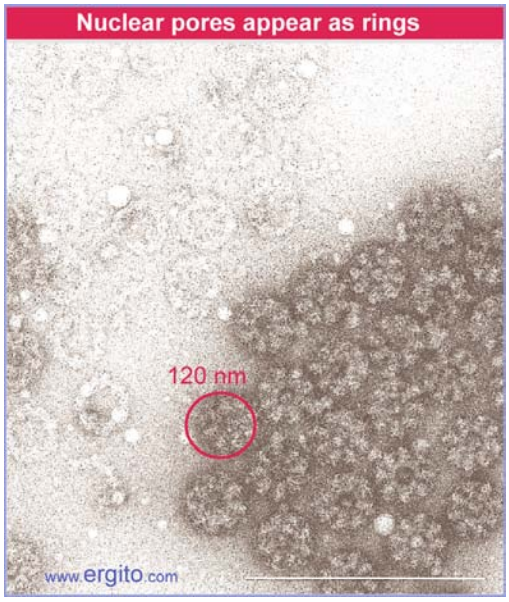


Figure 8.54 Nuclear pores appear as annular structures by electron microscopy. The circle around one pore has a diameter of 120 nm. Photograph kindly provided by Ronald Milligan.

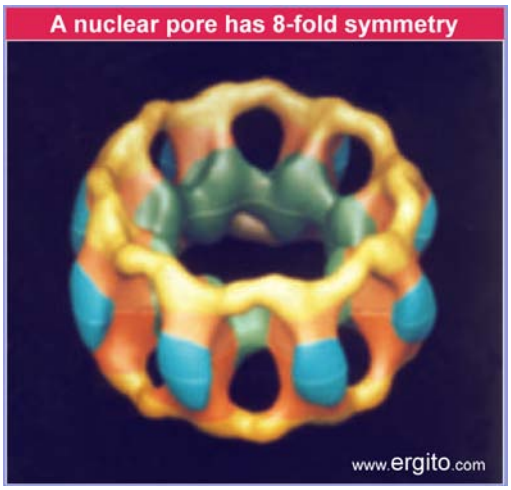


Figure 8.55 A model for the nuclear pore shows 8-fold symmetry. Two rings form the upper and lower surfaces (shown in yellow); they are connected by the spokes (shown in green on the inside and blue on the outside). Photograph kindly provided by Ronald Milligan.

The basis for the 8-fold symmetry is explained in terms of individual components in the schematic view from above shown in **Figure 8.56**. This includes the central structure of **Figure 8.55**, and extends it with an internal transporter and surrounding radial arms. The outside of the pore complex as such consists of a ring of diameter ~120 nm. The ring itself consists of 8 subunits. The 8 radial arms outside the ring may be responsible for anchoring the pore complex in the nuclear envelope; they penetrate the membrane. The 8 interior spokes project from the ring, closing the opening to a diameter of ~48 nm. Within this region is the transporter, which contains a pore that approximates a cylinder <10 nm in diameter (482).

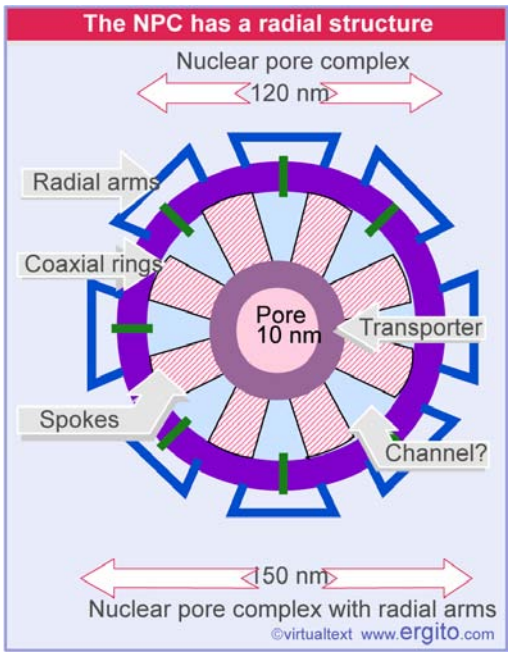


Figure 8.56 The outsides of the nuclear coaxial (cytoplasmic and nucleoplasmic) rings are connected to radial arms. The interior is connected to spokes that project towards the transporter that contains the central pore.

The pore provides a passage across the outer and inner membranes of the nuclear envelope. As illustrated in **Figure 8.57**, the side view has two-fold symmetry about a horizontal axis in the plane of the nuclear envelope. There are matching annuli at the outer and inner membranes, comprising the surfaces that project into the cytosol and into the nucleus, and each is connected to the spokes, which form a central ring. (Only 2 of the 8 spokes are seen in this side view.) The spokes are symmetrical about the horizontal axis. The central pore projects for the distance across the envelope. Sometimes material can be seen within the pore, but it has been difficult to equate such views with the transport of any particular material.

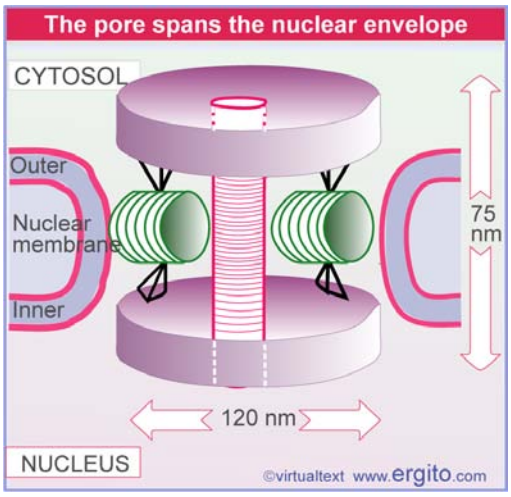


Figure 8.57 The nuclear pore complex spans the nuclear envelope by means of a triple ring structure. The side view shows two-fold symmetry from either horizontal or perpendicular axes.

The size of the nuclear pore complex corresponds to a total mass ~50 × 10^6 daltons (compare this with the 80S ribosome at 4 × 10^6 daltons). We can identify the smallest repeating component by using the 8-fold symmetry as seen in cross-section (see **Figure 8.56**) and the 2-fold symmetry seen from the side (see **Figure 8.57**). This divides the scaffold into 16 identical units. Each of these units consists of ~30 different proteins, most often each present in 1-2 copies per unit (959). The central pore constitutes only a small part of the overall complex (for review see 52; 57).

Individual protein components of the nuclear pore can be localized by immunoelectron microscopy. Most proteins are found on both sides of the pore; very few are found on only one side (2705). This supports the view of the complex as a symmetrical structure built from identical assemblies. Some of the pore complex proteins are transmembrane proteins; they probably help to anchor the complex in the envelope.

Last updated on 11-6-2002

Reviews

52. Forbes, D. J. (1992). *Structure and function of the nuclear pore complex*. Annu. Rev. Cell Biol. 8, 495-527.

57. Davis, L. I. (1995). *The nuclear pore complex.* Annu. Rev. Biochem. 64, 865-896.

References

482. Hinshaw, J. E., Carragher, B. O., and Milligan, R. A. (1992). *Architecture and design of the nuclear pore complex*. Cell 69, 1133-1141.

959. Rout, M. P., Aitchison, J. D., Suprapto, A., Hjertaas, K., Zhao, Y., and Chait, B. T. (2000). *The yeast nuclear pore complex: composition, architecture, and transport mechanism.* J. Cell Biol. 148, 635-651.

2705. Yang, Q., Rout, M. P., and Akey, C. W. (1998). *Three-dimensional architecture of the isolated yeast nuclear pore complex: functional and evolutionary implications*. Mol. Cell 1, 223-234.

PROTEIN LOCALIZATION

2.8.26 The nuclear pore is a size-dependent sieve for smaller material

Key Concepts

- The central channel is large enough to allow proteins of <50 kD to pass.
- Larger proteins must require the channel to open wider for passage.

The ability of compounds to diffuse freely through the pores is limited by their size. **Figure 8.58** summarizes the results of two sets of experiments in which material was injected into the cytoplasm, and its entry into the nucleus was followed over 24 hours (3068). Using dextrans (large saccharides) of different sizes shows that the smallest size equilibrates very rapidly, with just over half of the material localized in the nucleus within minutes. As the size of the dextran increases, entry into the nucleus becomes limited. By a diameter of 7 nm, virtually no dextran can pass through the nuclear pore. Analogous results are obtained with proteins (see panel on right of figure), where there is a progressive reduction in the proportion in the nucleus as the size increases.

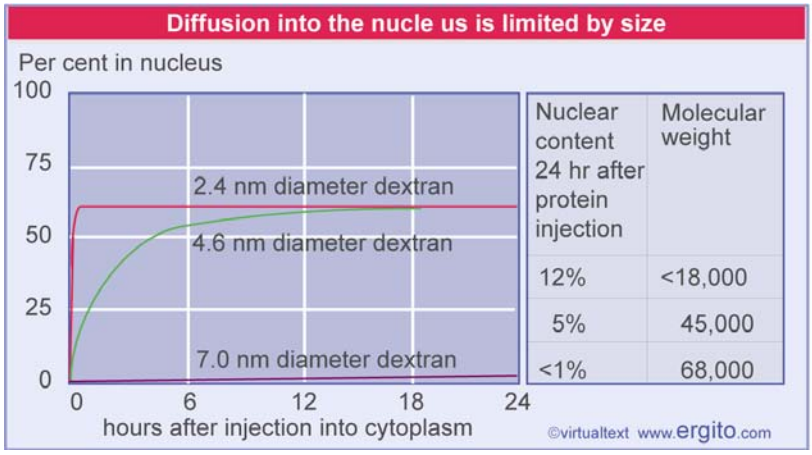


Figure 8.58 Small molecules and macromolecules can enter the nucleus freely, but diffusion is limited by size.

It is convenient to consider the material in three size classes:

- Molecules of <5000 daltons that are injected into the cytoplasm appear virtually instantaneously in the nucleus: we may conclude that the nuclear envelope is freely permeable to ions, nucleotides, and other small molecules.
- Proteins of 5-50 kD diffuse at a rate that is inversely related to their size, presumably determined by random contacts with, and passage through, the pore. It takes a few hours for the levels of an injected protein to equilibrate between

cytoplasm and nucleus. We may conclude that small proteins can enter the nucleus by passive diffusion (but they may also be actively transported). The nuclear envelope in effect provides a mesh or molecular sieve that permits passage of material <50 kD. A globular protein of 50 kD in mass would have a diameter of ~5 nm if it were spherical.

- Proteins >50 kD in size do not enter the nucleus by passive diffusion; a mechanism of active transport must be required for their passage (480; for review see 46).

Proteins are transported through the central pore. Electron-dense objects can sometimes be seen in the pore, and are usually assumed to represent material being transported (although this has not been proven). It is likely that the size limits on diffusion into the nucleus are determined by the diameter of the central channel. However, surrounding this channel are eight smaller openings, and it remains possible that smaller material can pass through them.

The ability of small proteins to diffuse through the pore means that, in the absence of any intervention, they will equilibrate between cytosol and nucleus. However, the distribution will be influenced by other interactions; for example, a small protein that is a component of chromatin could be bound to chromatin after it has diffused into the nucleus, and therefore will be largely localized in the nucleus. Larger proteins must use an active transport mechanism that overcomes the apparent size restriction of the pores. Also, active transport must be used for any protein that requires transport against a concentration gradient (for example, a protein that is localized freely in the nucleoplasm).

Transport through the pore has been characterized by using colloidal gold particles coated with a nuclear protein. When these particles are injected into the cytoplasm, they cluster at the nuclear pores, and then accumulate in the nucleus (477). This suggests that the pore structure can widen to accommodate objects of the size of the coated gold particles (~20 nm). Similar experiments have shown that gold particles coated with polynucleotides can be exported from the nucleus via pores. Following a simultaneous injection of RNA-gold particles of one size together with protein-gold particles of another size, pores can be seen to have both sizes of particles, which suggests that the same pores can be used for export and import (2604).

The rigidity of the gold particle excludes the possibility that transport through the pore requires the protein to change into a conformation with a diameter physically smaller than the pore. We conclude that the nuclear pore has a "gating" mechanism that allows the interior to expand as material passes through. Pores engaged in transporting material appear to be opened to a diameter of ~20 nm, possibly by a mechanism akin to the iris of a camera lens. It is possible that two irises, one connected with the cytoplasmic ring and one connected with the nucleoplasmic ring, open in turn as material proceeds through the pore. Very large substrates, such as exported ribonucleoprotein particles, may have to change their conformation to conform with the limit of 20 nm.

We believe that all pores are identical. The nuclear pore complex provides a structural framework that supports the proteins actually responsible for binding and transporting material into (or out of) the nucleus. However, it does not include all of

the active components that are involved in binding and translocation from one side to the other. Accessory factors that associate with the pore are responsible for the actual transport process.

Reviews

46. Dingwall, C. and Laskey, R. A. (1986). *Protein import into the cell nucleus.* Annu. Rev. Cell Biol. 2, 367-390.

References

477. Feldherr, C. M., Kallenbach, E., and Schultz, N. (1984). *Movement of a karyophilic protein through the nuclear pores of oocytes*. J. Cell Biol. 99, 2216-2222.

480. Akey, C. W. and Goldfarb, D. S. (1989). *Protein import through the nuclear pore complex is a multistep process*. J. Cell Biol. 109, 971-982.

2604. Dworetzky, S.I. and Feldherr, C.M. (1988). *Translocation of RNA-coated gold particles through the nuclear pores of oocytes*. J. Cell Biol. 106, 575-584.

3068. Paine, P. L. (1975). *Nucleocytoplasmic movement of fluorescent tracers microinjected into living salivary gland cells.* J. Cell Biol. 66, 652-657.

PROTEIN LOCALIZATION

2.8.27 Proteins require signals to be transported through the pore

Key Terms

A **nuclear localization signal (NLS)** is a domain of a protein, usually a short amino acid sequence, that interacts with an importin, allowing the protein to be transported into the nucleus.

A **nuclear export signal (NES)** is a domain of a protein, usually a short amino acid sequence, which interacts with an exportin, resulting in the transport of the protein from the nucleus to the cytoplasm.

Key Concepts

- The NLS and NES consist of short sequences that are necessary and sufficient for proteins to be transported through the pores into or out of the nucleus.

To be transported through the nuclear pore, a protein must have a special signal in its sequence. The most common motif responsible for import into the nucleus is called the *nuclear localization signal* (**NLS**). Its presence in a cytosolic protein is necessary and sufficient to sponsor import into the nucleus (2601). Mutation of the signal can prevent the protein from entering the nucleus (1014).

The summary of nuclear localization signals in **Figure 8.59** shows that there is no apparent conservation of sequence of NLS signals; perhaps the shape of the region or its basicity are the important features. Many NLS sequences take the form of a short, rather basic stretch of amino acids (2639). Often there is a proline residue to break α-helix formation upstream of the basic residues. Hydrophobic residues are rare. Some NLS signals are bipartite and require two separate short clusters (481). Competition studies suggest that NLS sequences are interchangeable, suggesting that they are all recognized by the same import system.

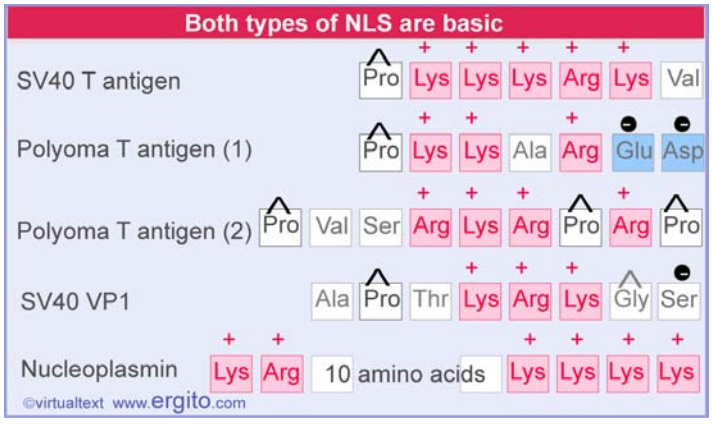


Figure 8.59 Nuclear localization signals have basic residues.

Many exported proteins have a common type of signal that is necessary and sufficient for the protein to move from the nucleus to the cytosol. It is called an **NES** (nuclear export signal), and typically consists of an ~10 amino acid sequence. The only common feature in the NES sequences in different proteins is a pattern of conserved leucines (1016; 1017).

A protein may have both an NLS and an NES, the former used for its import into the nucleus, and the latter for its export. They may function constitutively, or their use can be regulated, for example, by association with other proteins that obscure or expose the relevant sequences (486).

References

481. Robbins, J. et al. (1991). *Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequences*. Cell 64, 615-623.

486. Michael, W. M., Choi, M. and Dreyfuss, G. (1995). *A nuclear export signal in hnRNP A1: a signal-mediated, temperature-dependent nuclear protein export pathway*. Cell 83, 415-422.

1014. Kalderon, D., Richardson, W. D., Markham, A. F., and Smith, A. E. (1984). *Sequence requirements for nuclear location of simian virus 40 large-T antigen.* Nature 311, 33-38.

1016. Fischer, U., Huber, J., Boelens, W. C., Mattaj, I. W., and Luhrmann, R. (1995). *The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs.* Cell 82, 475-483.

1017. Bogerd, H. P., Fridell, R. A., Benson, R. E., Hua, J., and Cullen, B. R. (1996). *Protein sequence requirements for function of the human T-cell leukemia virus type 1 Rex nuclear export signal delineated by a novel in vitro randomization-selection assay.* Mol. Cell Biol. 16, 4207-4214.

2601. Dingwall, C., Sharnick, S. V., and Laskey, R. A. (1982). *A polypeptide domain that specifies migration of nucleoplasmin into the nucleus.* Cell 30, 449-458.

2639. Kalderon, D., Roberts, B. L., Richardson, W. D., and Smith, A. E. (1984). *A short amino acid sequence able to specify nuclear location.* Cell 39, 499-509.

PROTEIN LOCALIZATION

2.8.28 Transport receptors carry cargo proteins through the pore

Key Terms

Importins are transport receptors that bind cargo molecules in the cytoplasm and translocate into the nucleus, where they release the cargo.

Exportins are transport receptors that bind their cargo and associate with RanGTP in the nucleus. The trimeric complex translocates across the nuclear envelope into the cytoplasm, where hydrolysis of GTP bound to Ran results in release of cargo.

The initial association of a translating ribosome with the translocation channel in the membrane of the ER is called **docking**.

Translocation describes the stage of nuclear import or export when a protein or RNA substrate moves through the nuclear pore.

Nucleoporin was originally defined to describe the components of the nuclear pore complex that bind to the inhibitory lectins, but now is used to mean any component of the basic nuclear pore complex.

Key Concepts

- Transport receptors have the dual properties of recognizing NLS or NES sequences and binding to the nuclear pore.
- Exportins transport substrates from nucleus to cytoplasm; importins transport substrates from cytoplasm to nucleus.
- Exportins and importins interact with nucleoporins in the pore.

The basic principle of import and export is illustrated in **Figure 8.60**. A carrier protein (or transport receptor) takes the substrate through the pore. The transport receptor must then be returned across the membrane to function in another cycle. The transport receptors are classified according to the direction in which they transport the cargo. **Importins** bind the cargo in the cytoplasm and release it in the nucleus. **Exportins** bind the cargo in the nucleus and release it in the cytoplasm.

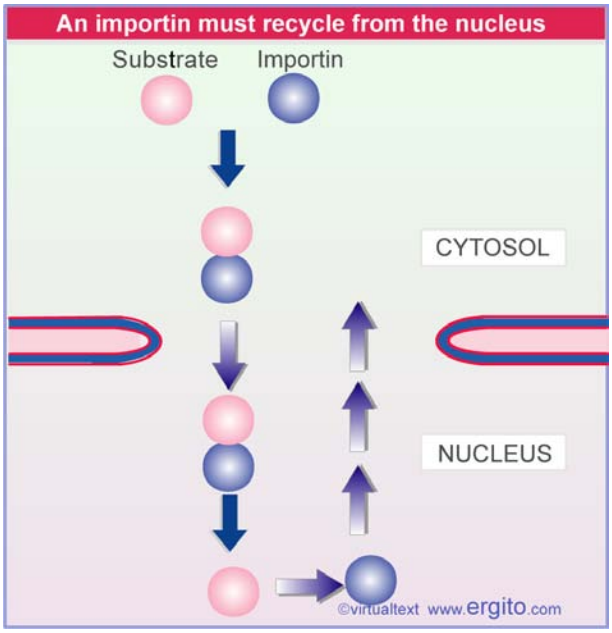


Figure 8.60 A carrier protein binds to a substrate, moves with it through the nuclear pore, is released on the other side, and must be returned for reuse.
This is a static version of an interactive figure; see http://www.ergito.com/main.jsp?bcs=MBIO.2.8.28 to view properly.

There are multiple pathways for import and export. Transport for all of the substrates for any particular pathway can be inhibited by saturating that pathway with one of its substrates. **Figure 8.61** summarizes the independent pathways. At least two different types of pathways exist for import of proteins; and each class of RNA is exported by a different system. Each transport receptor recognizes a particular type of sequence in its substrate, thus defining the specificity of the system (for review see 1004).

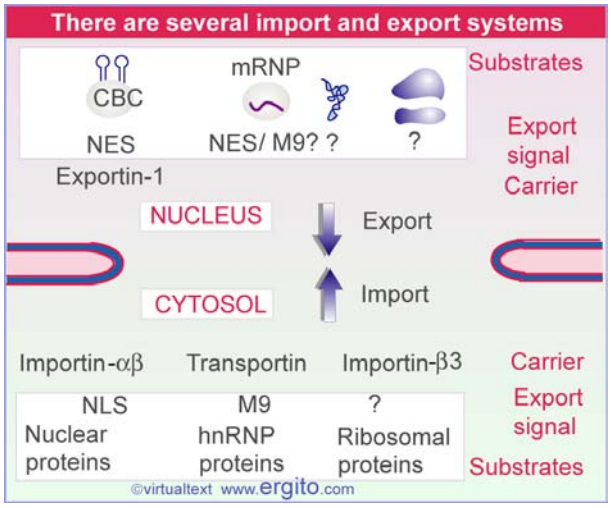


Figure 8.61 There are multiple pathways for nuclear export and import.

The first handle on the process of import was provided by systems for transport that depend upon the presence of an NLS in the substrate protein. An *in vitro* assay for nuclear pore import has been developed by using permeabilized cells. When cells are treated with digitonin, the plasma membrane becomes permeable, but the nuclear envelope remains intact. Labeled proteins can be imported into the nucleus in a process that is dependent upon the provision of cytosolic components (1015). **Figure 8.62** shows how this system has been used to characterize the transport process.

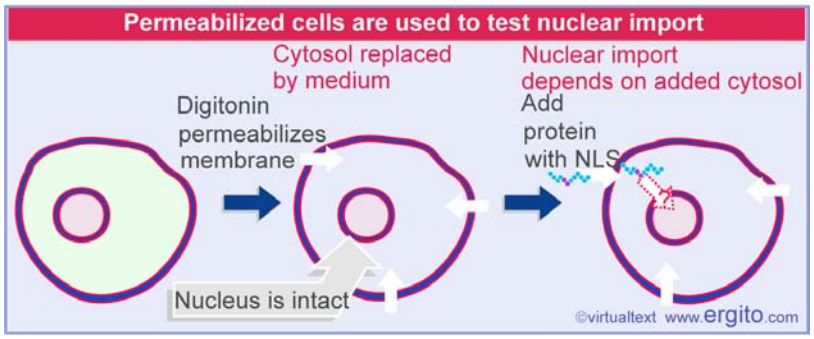


Figure 8.62 The assay for nuclear pore function uses permeabilized cells.

Transport can be divided into two stages: **docking**, which consists of binding to the pore; and **translocation**, which consists of movement through it. In the absence of ATP, proteins containing a nuclear import signal can bind at the pore, but they remain at the cytoplasmic face. A cytosolic fraction is needed for binding. When ATP is provided, proteins can be translocated through the pore. A different cytosolic fraction is needed to support translocation. The need for cytosolic fractions at both stages reinforces the view of the pore as a structure that provides the framework for transport, but that does not provide all of the necessary facilities for handling the substrates (478; 479; 483).

Figure 8.63 summarizes the functions of the components involved in nuclear import. The transport receptor is the key intermediate in the docking reaction. It can bind to both the nuclear pore and the cargo protein. Some transport receptors are single proteins, such as transportin or importin- β3, which undertake both binding reactions (484). Others are dimers in which one subunit binds to the pore, and the other is an adaptor that binds to the cargo protein. In the best characterized case, importin- α β has a β subunit that binds to the nuclear pore. The α subunit binds proteins that have an NLS sequence. The single protein receptors transportin and importin- β3 are related to the β subunit of importin- α β.

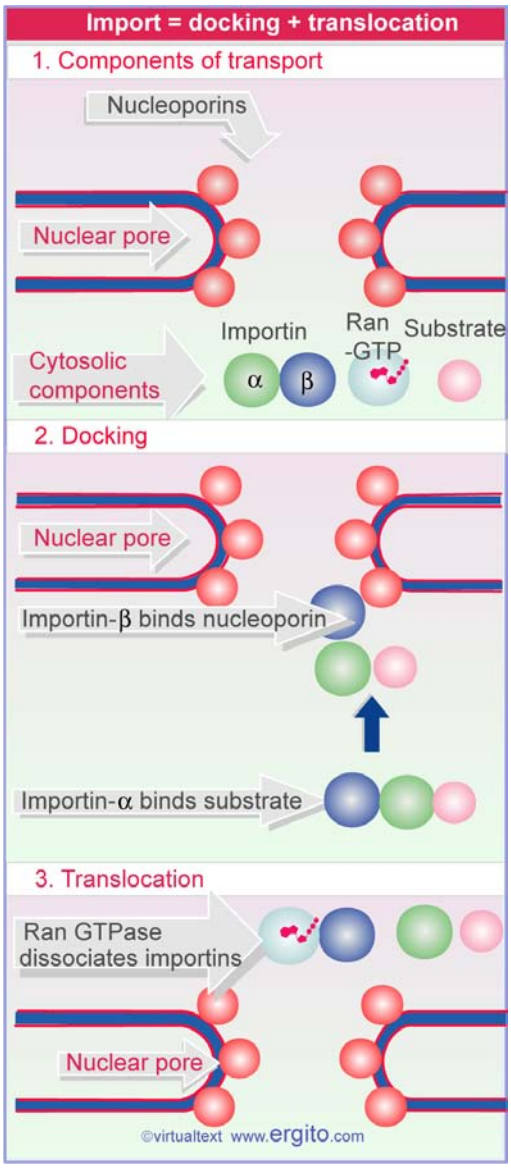


Figure 8.63 Nuclear import takes place in two stages. Both docking and translocation depend on cytosolic components. Translocation requires nucleoporins.

Translocation through the pore is inhibited by wheat germ agglutinin, a lectin (glycoprotein). The component proteins of the pore that bind to lectins were originally called **nucleoporins**. Note that nucleoporin has since come to be used to mean any component of the nuclear pore complex. The lectin-binding components are localized at or near the region of the central pore, and appear to be located on both sides of the nuclear envelope. When they are removed, pore complexes remain normal in appearance, but can no longer function to transport large material. Material smaller than the pore size continues to be able to move through by diffusion. When the lectin-binding proteins are added back, they restore full activity to the deficient pores. This suggests that they are needed for active transport of material larger than the resting diameter. Some of these proteins have some simple peptide repeating

motifs (GKFG, FG, FXFG), and it is probably these motifs that bind the importin- β carrier proteins (485). They are called FG-nucleoporins.

Last updated on 5-22-2000

Reviews

1004. Gorlich, D. and Kutay, U. (1999). *Transport between the cell nucleus and cytoplasm.* Annu. Rev. Cell Dev. Biol. 15, 607-660.

References

478. Newmeyer, D. D. and Forbes, D. J. (1988). *An NEM-sensitive cytosolic factor necessary for nuclear protein import: requirement in a signal-mediated binding to the nuclear pore*. Cell 52, 641-653.

479. Richardson, W. D. et al. (1988). *Nuclear protein migration involves two steps: rapid binding at the nuclear envelope followed by slower translocation through nuclear pores*. Cell 52, 655-664.

483. Moore, M. S. and Blobel, G. (1992). *The two steps of nuclear import, targeting to the nuclear envelope and translocation through the nuclear pore, require different cytosolic factors*. Cell 69, 939-950.

484. Moroianu, J., Blobel, G., and Radu, A. (1995). *Previously identified protein of uncertain function is importin α docks import substrate at nuclear pore complexes*. Proc. Natl. Acad. Sci. USA 92, 2008-2011.

485. Radu, A., Moore, M. S., and Blobel, G. (1995). *The peptide repeat domain of nucleoporin Nup98 functions as a docking site in transport across the nuclear pore complex*. Cell 81, 215-222.

1015. Adam, S. A., Marr, R. S., and Gerace, L. (1990). *Nuclear protein import in permeabilized mammalian cells requires soluble cytoplasmic factors.* J. Cell Biol. 111, 807-816.

PROTEIN LOCALIZATION

2.8.29 Ran controls the direction of transport

Key Concepts

- The nucleus contains Ran-GTP, which stabilizes export complexes, while the cytosol contains Ran-GDP, which stabilizes import complexes.
- Movement through the nuclear pore does not involve a motor.

How does material move through the pore? One of the striking features of the composition of the nuclear pore complex is the lack of any protein with motor activity (959). This suggests that movement must depend on the affinity of the carrier proteins for the pore itself. But now we have to explain how transport can work for both import and export. The answer lies in the properties of the monomeric G-protein, Ran.

The cytosolic fraction that supports translocation has two active components. One is Ran; the other may be involved in targeting Ran to the nuclear pore. Ran is a typical monomeric G-protein that can exist in either the GTP-bound or GDP-bound state (see *Molecular Biology Supplement 32.10 G proteins*). Its GTPase activity generates Ran-GDP. Then an exchange factor is needed to replace GDP with GTP to regenerate Ran-GTP.

The directionality of nuclear import is controlled by the state of Ran (for review see 1004; 2293). **Figure 8.64** shows that conditions in the nucleus and cytosol differ so that typically there is Ran-GTP in the nucleus, but there is Ran-GDP in the cytosol. The reason for this difference is an asymmetric distribution of two proteins that act on Ran. The nucleus contains Ran-GEF, which stimulates replacement of GDP by GTP, thus converting Ran-GDP to Ran-GTP. (In fact, this protein, also known as Rcc1, is localized to chromatin.) The cytoplasm contains Ran-GAP, which causes the GTP to be hydrolyzed to GDP. The Ran-GAP is localized on the surface of the cytoplasmic side of the nuclear pore complex, together with a Ran-binding protein (RanBP1) that stimulates its activity.

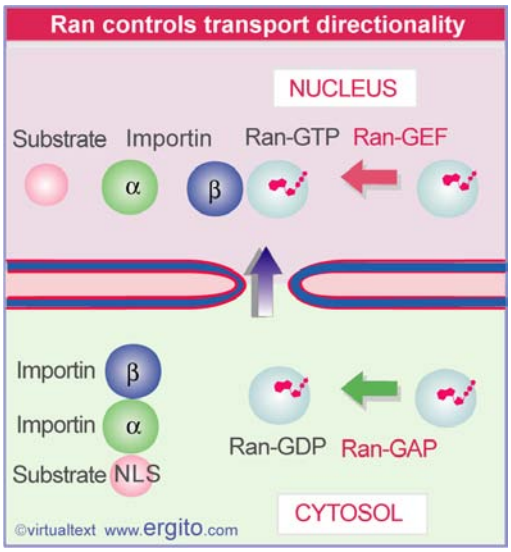


Figure 8.64 The state of the guanine nucleotide bound to Ran controls directionality of nuclear import and export.

Export complexes are stable in the presence of Ran-GTP, whereas import complexes are stable in the presence of Ran-GDP. So export complexes are driven to form in the nucleus and dissociate in the cytosol, whereas the reverse is true of import complexes. The reaction has been best characterized for the complex of importin- α β with an NLS-containing protein. The triple complex is stable in the presence of Ran-GDP, and thus can form in the cytosol. However, Ran-GTP causes importin- α to dissociate from importin- β. This leads to release of the substrate protein in the nucleus (1018; 1019). The effect of Ran-GTP in causing the importin dimer to release its substrate is also important at mitosis, when importins release proteins that trigger the attachment of microtubules to the spindle (see *Molecular Biology 6.29.23 A monomeric G protein controls spindle assembly*).

The crystal structures of complexes containing importin- β show how it binds to importin- α and to Ran. Importin- β consists of a series of repeating units coiled into a superhelix. The individual repeating unit (called HEAT) itself consists of two α-helices (HEAT-A and HEAT-B). Importin- α has a similar structure, but in addition has a domain (IBB for importin-binding domain) that binds importin- β. **Figure 8.65** shows that importin- β winds around the IBB of importin- α, making a tightly integrated dimer (1020).

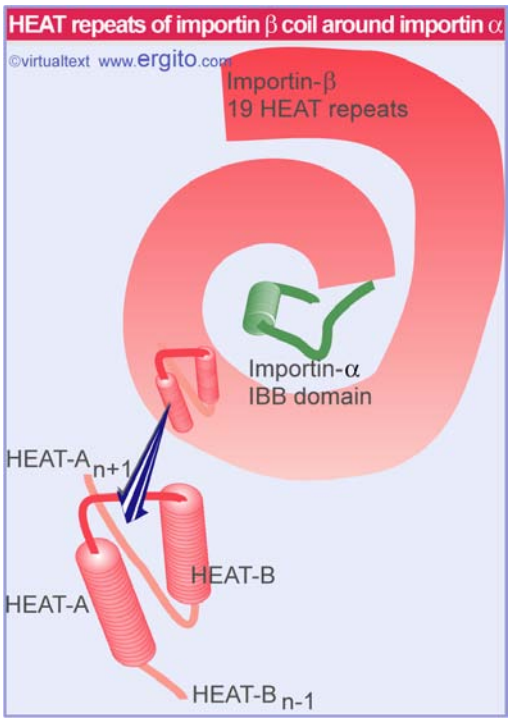


Figure 8.65 Importin- β consists of 19 HEAT repeats organized in a right-handed superhelix. Each HEAT unit consists of two α -helices (A and B) lying at an angle to one another. Importin- β is folded tightly around the IBB domain of importin- α .

Figure 8.66 shows that Ran-GTP binds tightly to two of the HEAT repeats in importin- β. Its binding site partially overlaps the binding site of importin- α (1021). This explains why binding of Ran-GTP displaces importin- α (and the NLS of the cargo protein) from importin- β. There is a large structural change in Ran when GTP is hydrolyzed to GDP (involving the unfolding of a helical region), which explains why the displacement reaction is specific for Ran-GTP.

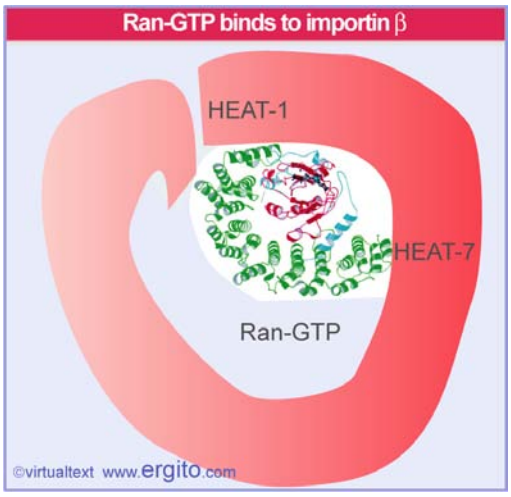


Figure 8.66 Importin- β binds Ran-GTP through close contacts to the N-terminal HEAT repeats and to repeats 7-8. The structure is significantly different from the importin- β importin- α structure.

As pointed out in **Figure 8.60**, in order to function more than once, an importin must return to the cytosol after taking its substrate into the nucleus. In effect, when an importin is released in the nucleus, it must become a substrate for an exportin! Such a reaction has been characterized for importin- α, which is bound in the presence of Ran-GTP by a protein called CAS (489). CAS behaves in a similar manner to importin- β, except that it moves in the opposite direction. Like exportins, it dissociates from its substrate (importin- α) when Ran-GTP is hydrolyzed in the cytosol.

The NES binds to exportin-1, which is related in sequence to importin- β. Exportin-1 is needed for the export of U snRNAs, some proteins, and possibly (some) mRNAs. It binds the NES motif; it binds nucleoporins; and it binds Ran-GTP. The complex forms in the presence of Ran-GTP; hydrolysis of the GTP to generate Ran-GDP is accompanied by dissociation of the complex. So Ran controls directionality of export in the reverse sense from its control of import: because Ran-GTP is high in the nucleus, the complex forms there; because it becomes Ran-GDP in the cytosol, the complex dissociates there.

It was thought for a long time that energy for transport would be provided by the hydrolysis of GTP by Ran. However, this is not exactly true. Transport through the pore does not itself require energy (2330). But the hydrolysis of GTP is required for a second cycle of transport, because it is necessary to generate Ran-GDP.

How does the importin-substrate complex cross the nuclear pore? It has a considerable distance to travel, ~200 nm. The pore itself has a symmetrical structure, and very few of its component proteins are concentrated on one side or the other (959). There are no obvious markers that might be used to indicate directionality. However, it is controversial whether the directionality of transport is intrinsic to the pore. At one time it was thought that it could be reversed by reversing the relative concentrations of Ran-GDP and Ran-GTP on either side, but more recent experiments suggest that it cannot (2330; 2332). This leaves us with two types of model: there is some unknown intrinsic directionality resulting from an uneven

distribution of proteins within the pore; or transport is stochastic, with receptor-substrate complexes bumbling through the pore until they emerge on the other side, where conditions ensure dissociation.

Importins and exportins move through the pore by a process of facilitated diffusion, but the details are still unclear. The crucial property that enables the importins and exportins to translocate through the pore is their ability to interact with the FG-nucleoporins by binding to the hydrophobic (Phe-Gly) repeats. The various transport receptors compete with one another to bind to the FG-nucleoporins, which argues that there is a common mechanism of translocation. FG-nucleoporins are distributed throughout the central pore, with a concentration at the midpoint. One model suggests that the hydrophobic repeats interact in this region to form a large mesh (2679). The ability of transport receptors to interact with the repeats allows them to "dissolve" in the mesh and thus to pass through it. The different rates at which different substrates are translocated by the pore is determined by how easily they can be incorporated into the mesh. Some crucial questions remain unanswered, especially what determines directionality within the pore itself, since the process seems to be too rapid to be the result of a random walk that is ended by dissociation of the complex on the appropriate side of the pore.

Last updated on 11-6-2002

Reviews

1004. Gorlich, D. and Kutay, U. (1999). *Transport between the cell nucleus and cytoplasm.* Annu. Rev. Cell Dev. Biol. 15, 607-660.

2293. Komeili, A. and O'Shea, E. K. (2001). *New perspectives on nuclear transport.* Annu. Rev. Genet. 35, 341-364.

References

489. Kutay, U. et al. (1997). *Export of importin α from the nucleus is mediated by a specific nuclear transport factor*. Cell 90, 1061-1071.

959. Rout, M. P., Aitchison, J. D., Suprapto, A., Hjertaas, K., Zhao, Y., and Chait, B. T. (2000). *The yeast nuclear pore complex: composition, architecture, and transport mechanism.* J. Cell Biol. 148, 635-651.

1018. Gorlich, D., Pante, N., Kutay, U., Aebi, U., and Bischoff, F. R. (1996). *Identification of different roles for RanGDP and RanGTP in nuclear protein import.* EMBO J. 15, 5584-5594.

1019. Rexach, M. and Blobel, G. (1995). *Protein import into nuclei: association and dissociation reactions involving transport substrate, transport factors, and nucleoporins.* Cell 83, 683-692.

1020. Cingolani, G., Petosa, C., Weis, K., Muller, C. W., Cingolani, G., Petosa, C., Weis, K., and Muller, C. W. (1999). *Structure of importin-beta bound to the IBB domain of importin-alpha*. Nature 399, 221-229.

1021. Chook, Y. M. and Blobel, G. (1999). *Structure of the nuclear transport complex karyopherin-beta2-Ran x GppNHp*. Nature 399, 230-237.

2330. Nachury, M. V. and Weis, K. (1999). *The direction of transport through the nuclear pore can be inverted.* Proc. Natl. Acad. Sci. USA 96, 9622-9627.

2332. Keminer, O., Siebrasse, J. P., Zerf, K., and Peters, R. (1999). *Optical recording of signal-mediated protein transport through single nuclear pore complexes.* Proc. Natl. Acad. Sci. USA 96, 11842-11847.

2679. Ribbeck, K. and Gorlich, D. (2001). *Kinetic analysis of translocation through nuclear pore complexes.* EMBO J. 20, 1320-1330.

PROTEIN LOCALIZATION

2.8.30 RNA is exported by several systems

Key Concepts

- There are (at least) three export systems for RNA.
- Each consists of an exportin that binds particular types of RNA.

Export systems have similar components to import systems. **Figure 8.67** illustrates some examples of export systems. The major substrates for export from the nucleus are ribonucleoproteins – ribosomes, mRNPs, snRNPs, and tRNA-protein complexes. In the first three cases, one of the protein components of the complex may be responsible for export (for example, for snRNAs it is the cap-binding complex (CBC). tRNA is bound directly by a specific export protein (487; 488).

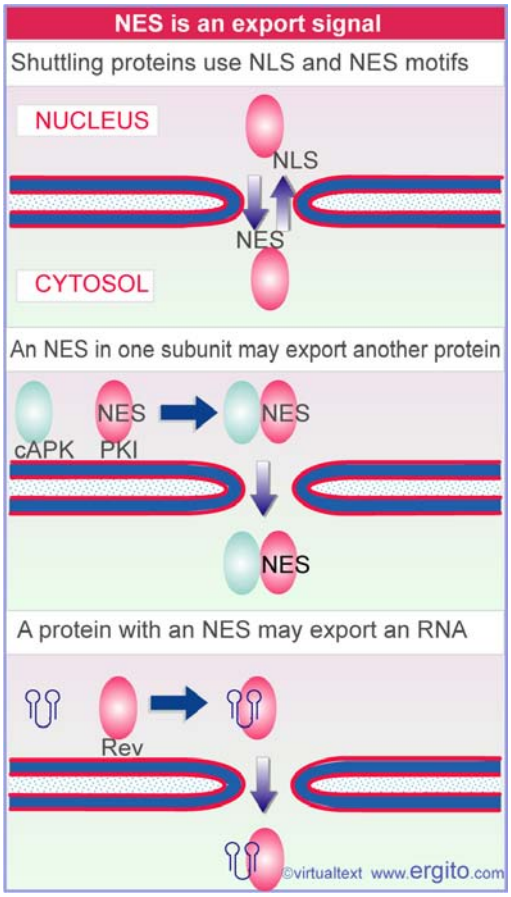


Figure 8.67 The common feature in proteins that are exported from the nucleus to the cytosol is the presence of an NES.

Export of mRNA has similar requirements to import. Mutations in a yeast FG-nucleoporin block export of RNA from the nucleus without affecting import of proteins. This suggests the possibility that the apparatus is similar for both import and export, but could have components that confer directionality or specificity for particular substrates. There is evidence for diversity in the export apparatus; using an assay for export of microinjected RNAs from the nucleus of the Xenopus oocyte, tRNAs, other small RNAs, and mRNAs each saturate transport only of their own class. This suggests that there are at least three groups of exported RNAs.

Some proteins "shuttle" between the cytoplasm and nucleus; they remain only briefly in either compartment before cycling back to the other. This behavior is characteristic of certain proteins that are bound to poly(A)$^+$ RNA in both the nucleus and the cytoplasm. The motif responsible for transport in one such protein (M9) has a single amino acid stretch that functions as both an import and export signal, and is therefore responsible for movement in both directions (for review see 58).

One particular issue with the export of mRNA is how to distinguish the final, processed mRNA from precursors that are not fully processed (for example, which retain some introns). Part of the answer may lie in the relative timing of events. Processing is connected with transcription in such a way that it is likely to be completed by the time the RNA is released from DNA. However, there are also some specific links that may connect export to the preceding events. One of the mRNA-binding proteins that is exported from the nucleus bound to mRNA attaches to the RNA via an interaction with the transcription apparatus when transcription is initiated (2246). This suggests that the proteins involved in exporting mRNA may become complexed with it at a very early stage of its production. And then mRNAs that are spliced may require splicing to occur in order for other components of the export apparatus to bind to the mRNA (see *Molecular Biology 5.24.10 Splicing is connected to export of mRNA*). The proteins bound to the mRNA then interact with a protein called Mex67 (in yeast) or Tap (in animal cells), which is unrelated to the exportin or importin families, but can interact directly with nucleoporins (for review see 2422).

Last updated on 4-23-2002

Reviews

58. Gorlich, D. and Mattaj, I. W. (1996). *Nucleocytoplasmic transport.* Science 271, 1513-1518.

2422. Reed, R. and Hurt, E. (2002). *A conserved mRNA export machinery coupled to pre-mRNA splicing.* Cell 108, 523-531.

References

487. Stade, K., Ford, C. S., Guthrie, C., and Weis, K. (1997). *Exportin 1 (Crm1p) is an essential nuclear export factor*. Cell 90, 1041-1050.

488. Fornerod, M. et al. (1997). *CRM1 is an export receptor for leucine-rich nuclear export signals*. Cell 90, 1051-1060.

2246. Lei, E. P., Krebber, H., and Silver, P. A. (2001). *Messenger RNAs are recruited for nuclear export during transcription.* Genes Dev. 15, 1771-1782.

PROTEIN LOCALIZATION

2.8.31 Ubiquitination targets proteins for degradation

Key Terms

Ubiquitin has a highly conserved sequence of 76 amino acids. It is linked via its COOH group to the ε NH_2 group of a lysine residue in a target protein.

Key Concepts

- Ubiquitin is added to proteins that are targeted for degradation by an apparatus consisting of three components.

A major pathway for protein degradation involves two stages: first the protein is targeted; and then it is proteolysed by a large complex that we describe in the next section. A small polypeptide called **ubiquitin** is connected by a covalent link to the substrate protein that is to be degraded (490; for review see 55; see *Great Experiments 5.3 Ubiquitin conjugation as a proteolytic signal: The first experiments*). **Figure 8.68** shows that there are three components of the ubiquitination system (for review see 1818).

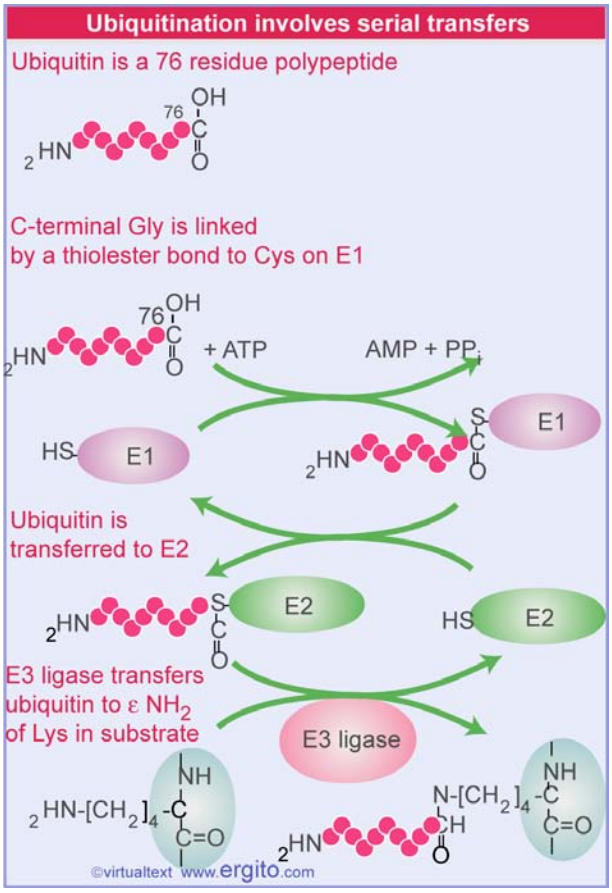


Figure 8.68 The ubiquitin cycle involves three activities. E1 is linked to ubiquitin. E3 binds to the substrate protein. E2 transfers ubiquitin from E1 to the substrate. Further cycles generate polyubiquitin.

- The ubiquitin-activating enzyme, E1, utilizes the cleavage of ATP to link itself via a high energy thiolester bond from a Cys residue to the C-terminal Gly residue of ubiquitin (for review see 53).
- The ubiquitin is then transferred to the ubiquitin-conjugating enzyme, E2.
- The ubiquitin ligase, E3, transfers the ubiquitin from E2 to form an isopeptide bond to the ε NH_2 group of a Lys in the substrate protein (for review see 2283).

Ubiquitin is released from a degraded substrate by an isopeptidase.

Responsibility for choosing substrate proteins to be ubiquitinated lies with both E2 and E3. In many cases, E3 selects the substrate protein by binding to it before transfer of ubiquitin is initiated. A cell may contain several E3 proteins that use different criteria for selecting substrates. There are also multiple varieties of E2, and they also may play a role in targeting substrate proteins, sometimes independently of E3 (for review see 1813).

The addition of a single ubiquitin residue to a substrate protein is not sufficient to

cause its degradation. **Figure 8.69** shows that further ubiquitin residues are added to form a polyubiquitin chain, in which each additional ubiquitin is added to the Lys at position 46 of the preceding ubiquitin. The formation of polyubiquitin is a signal for the proteasome to degrade the protein (491).

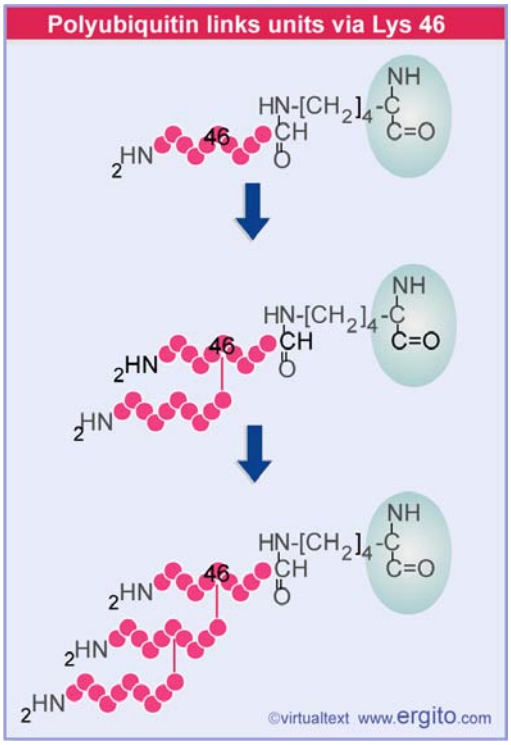


Figure 8.69 Polyubiquitin is generated by successive cycles of adding ubiquitin.

Targeting for degradation by the proteasome in the cytosol is the major function of ubiquitination. However, it also targets plasma membrane proteins for degradation in lysosomes, and possibly may have other regulatory effects.

In addition to ubiquitin, there are ubiquitin-like proteins that modify target proteins in a similar way. The best characterized of these is SUMO (also known as Sentrin). A difference between the ubiquitin system and the SUMO system is that only a single SUMO residue is added to a target protein, compared with polyubiquitination. The consequences of sumoylation are not entirely clear; it may be concerned with protein localization or with protection against ubiquitination. SUMO has fewer targets than ubiquitin, but they often include important cellular proteins (for review see 3451).

Ubiquitination (or sumoylation) can be reversed by proteases that cleave the conjugate from the target protein, so there is the potential for significant complexity in regulation.

Last updated on 1-30-2003

Reviews

53. Jentsch, S. (1992). *The ubiquitin-conjugation system.* Annu. Rev. Genet. 26, 179-207.

55. Ciechanover, A. (1994). *The ubiquitin-proteasome proteolytic pathway.* Cell 79, 13-21.

1813. Hershko, A., and Ciechanover, A. (1998). *The ubiquitin system.* Annu. Rev. Biochem. 67, 425-479.

1818. Weissman, A. M. (2001). *Themes and variations on ubiquitylation.* Nat. Rev. Mol. Cell Biol. 2, 169-178.

2283. Pickart, C. M. (2001). *Mechanisms underlying ubiquitination.* Annu. Rev. Biochem. 70, 503-533.

3451. Muller, S., Hoege, C., Pyrowolakis, G., and Jentsch, S. (2001). *SUMO, ubiquitin's mysterious cousin.* Nat. Rev. Mol. Cell Biol. 2, 202-210.

References

490. Ciechanover, A. et al. (1980). *ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation*. Proc. Natl. Acad. Sci. USA 77, 1365-1368.

491. Chau, V., Tobias, J. W., Bachmair, A., Marriott, D., Ecker, D. J., Gonda, D. K., and Varshavsky, A. (1989). *A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein.* Science 243, 1576-1583.

PROTEIN LOCALIZATION

2.8.32 The proteasome is a large machine that degrades ubiquitinated proteins

Key Terms

The **proteasome** is a large complex with an interior cavity that degrades cytosolic proteins previously marked by covalent addition of ubiquitin.

Key Concepts

- Protein degradation in the cytosol is catalyzed by the proteasome.
- The 20S proteasome consists of four stacked rings each containing 7 subunits.
- The 26S proteasome is formed in eukaryotes when caps associate with one or both ends of the 20S proteasome.
- Caps pass ubiquitinated proteins to the core for proteolysis.
- The proteasome contains several different protease activities.

What determines the stability of proteins? A cell contains many proteases, with varying specificities. We may divide them into three general groups:

- Some proteases are involved in specific processing events to generate mature proteins that are smaller than the precursors. Such proteases are involved in a variety of activities, including cleaving the signal sequence from a secreted protein, and cleaving cytosolic enzymes into their mature forms. The caspase group of proteases are involved a pathway that leads to cell death (see *Molecular Biology 6.29.27 A common pathway for apoptosis functions via caspases*).
- Lysosomes are membrane-bounded organelles that degrade proteins imported into the cell; we discuss this process in the context of protein transport through membranes in *Molecular Biology 6.27.15 Receptors recycle via endocytosis*.
- The **proteasome** is a large complex that degrades cytosolic proteins. It is involved in both general degradation (the complete conversion of a protein into small fragments) and in certain specific processing events. The major substrates for complete degradation are proteins that have been misfolded – this is basically a quality control system – and certain proteins whose degradation is a regulatory event, for example, to allow progress through the cell cycle.

The proteasome was originally discovered as a large complex that degrades proteins conjugated to ubiquitin. It exists in two forms. A 20S complex of ~700 kD has protease activity. Additional proteins convert the complex to a 26S form of ~2000 kD; they are regulatory subunits that confer specificity – for example, for binding to ubiquitin conjugates. ATP cleavage is required for the conversion from 20S to 26S,

and is also required later in the reaction for cleaving peptide bonds, releasing the products, etc (for review see 996).

The 20S complex takes the form of a hollow cylinder, and the additional components of the 26S complex are attached to the ends of the cylinder, making a dumbbell. Basically, the active sites are contained in the interior of a barrel, and access is obtained through relatively narrow channels, typically allowing only access only to unfolded proteins. This protects normal, mature proteins from adventitious degradation (492; for review see 67).

This general type of structure is common to ATP-dependent proteases. For example, the ClpAP protease in *E. coli*, which is not related by sequence to the proteasome, has a structure in which the ClpP protease forms two rings of 7 subunits each, with the proteolytic activities contained in a central cavity (2226). ClpA is hexameric and stacks on to the ClpP complex (which implies an interesting symmetry mismatch in the ClpAP complex). ClpA provides the ATPase activity. It unfolds the substrates and translocates them in denatured form through a narrow passage into the ClpP cavity, where they are degraded (2209; 2210). Degradation is processive; once a substrate has been admitted to the central cavity, the reaction proceeds to its end.

The simplest proteasome is found in the archaea. **Figure 8.70** shows the top view of the crystal structure of the 20S assembly. It consists of two types of subunits, organized in the form $\alpha_7 - \beta_7 - \beta_7 - \alpha_7$, where each septamer forms a ring. **Figure 8.71** shows the side view of the backbone. The α subunits form the two outer rings (on top and on the bottom), and the β subunits form the two inner rings. The β subunits have the protease activities, and the active sites are located at the N-terminal ends that project into the interior. The opening of ~20Å restricts the entrance for substrates. A yet simpler structure is found in *E. coli*, where a protein related to the β subunit, HslV, forms a structure of two six-member rings with a proteolytic core (493).

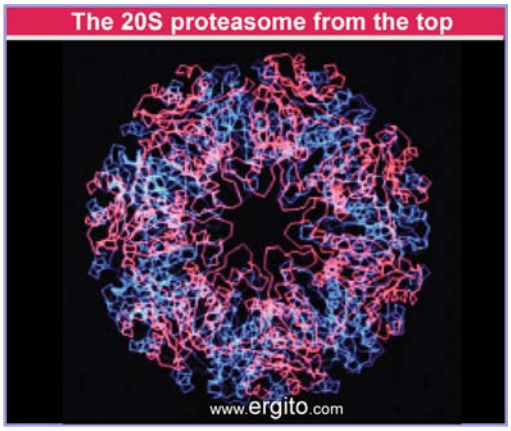


Figure 8.70 The top view of the archaeal 20S proteasome shows a hollow cylinder consisting of heptameric rings of α subunits (red) and β subunits (blue). Photograph kindly provided by Robert Huber.

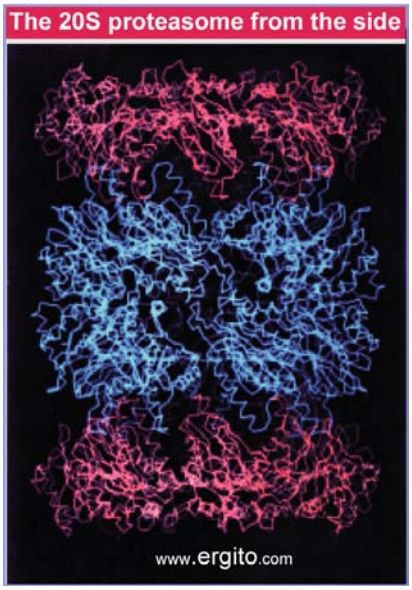


Figure 8.71 The side view of the archaeal 20S proteasome shows the rings of α subunits (red) and β subunits (blue). Photograph kindly provided by Robert Huber.

The eukaryotic 20S proteasome is more complex, consisting of 7 different α subunits and 7 different β subunits. **Figure 8.72** shows that it has the same general structure of α– β– β– α rings. The rings in each half of the structure are organized in the opposite rotational sense. A significant structural difference with the archaeal proteasome is that the central hole is occluded, so that there is no obvious entrance from the ends of the cylinder. This probably means that the structure is rearranged at some point to allow entrance from the ends (495).

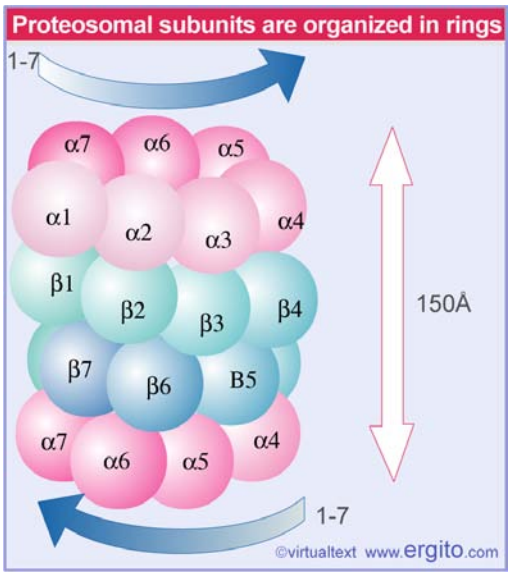


Figure 8.72 The eukaryotic 20S proteasome consists of two dimeric rings organized in counter-rotation.

The eukaryotic 26S proteasome is formed when the 19S caps associate with the 20S core, binding to one or both ends, to form an elongated structure of ~45 nm in length. The 19S caps are found only in eukaryotic (not archaeal or bacterial) proteasomes. The caps recognize ubiquitinated proteins, and pass them to the 20S core for proteolysis. The 19S caps contain ~18 subunits, several of which are ATPases; presumably the hydrolysis of ATP provides energy for handling the substrate proteins (for review see 61).

The hydrolytic mechanism of the proteasome is different from that of other proteases. The active site of a catalytic β-subunit is an N-terminal threonine; the hydroxyl group of the threonine attacks the peptide bond of the substrate. The proteasome contains several protease activities, with different specificities, for example, for cleaving after basic, acidic, or hydrophobic amino acids, allowing it to attack. a variety of types of targets. Proteolytic activities with different substrate specificities may be provided by different β subunits. More than one β subunit may be needed for a particular enzymatic activity. The peptide products typically are octa- and nona-peptides. Proteasomes function processively, that is, a substrate is degraded to completion within the cavity, without any intermediates being released. Basically the central chamber traps proteins until they have been degraded to fragments below a certain size.

Inhibitors of the proteasome block the degradation of most cellular proteins, showing that it is responsible for bulk degradation. In fact, a significant proportion of newly synthesized proteins are immediately degraded by the proteasome (which casts a light on the efficiency of the production of proteins) (1024; 1025). It is also responsible for cleaving antigens in cells of the immune system to generate the small peptides that are presented on the surface of the cell to provoke the immune response (see *Molecular Biology 5.25.18 T cell receptors are related to immunoglobulins*). The peptide fragments are then transported by TAP (the transporter associated with antigen processing) from the cytosol into the ER, where they are bound by MHC

molecules. Other reactions in which target proteins are completely degraded include the removal of cell cycle regulators; in particular, cyclins are degraded during mitosis (see *Molecular Biology 6.29.18 Protein degradation is important in mitosis*), and replication control proteins are degraded during the phase of DNA synthesis. In addition to these reactions, the proteasome may undertake specific processing events, for example, cleaving a precursor to a transcription factor to generate the active protein. The means by which these activities are regulated remain to be discovered.

Last updated on 12-14-2001

Reviews

61. Coux, O., Tanaka, K., and Goldberg, A. L. (1996). *Structure and functions of the 20S and 26S proteasomes.* Annu. Rev. Biochem. 65, 801-847.

67. Baumeister, W. et al. (1998). *The proteasome: paradigm of a self-compartmentalizing protease*. Cell 92, 367-380.

996. Voges, D., Zwickl, P., and Baumeister, W. (1999). *The 26S proteasome: a molecular machine designed for controlled proteolysis.* Annu. Rev. Biochem. 68, 1015-1068.

References

492. Eytan, E. et al. (1989). *ATP-dependent incorporation of 20S protease into the 26S complex that degrades proteins conjugated to ubiquitin*. Proc. Natl. Acad. Sci. USA 86, 7751-7755.

493. Lowe, J. et al. (1995). *Crystal structure of the 20S proteasome from the archaeon T. acidophilum at 34* Å *resolution*. Science 268, 533-539.

495. Groll, M. et al. (1997). *Structure of 20S proteasome from yeast at 24* Å *resolution*. Nature 386, 463-471.

1024. Schubert, U., Anton, L. C., Gibbs, J., Norbury, C. C., Yewdell, J. W., and Bennink, J. R. (2000). *Rapid degradation of a large fraction of newly synthesized proteins by proteasomes.* Nature 404, 770-774.

1025. Reits, E. A., Vos, J. C., Gromme, M., and Neefjes, J. (2000). *The major substrates for TAP in vitro are derived from newly synthesized proteins.* Nature 404, 774-778.

2209. Ishikawa, T., Beuron, F., Kessel, M., Wickner, S., Maurizi, M. R., and Steven, A. C. (2001). *Translocation pathway of protein substrates in ClpAP protease.* Proc. Natl. Acad. Sci. USA 98, 4328-4333.

2210. Reid, B. G., Fenton, W. A., Horwich, A. L., and Weber-Ban, E. U. (2001). *ClpA mediates directional translocation of substrate proteins into the ClpP protease.* Proc. Natl. Acad. Sci. USA 98, 3768-3772.

2226. Wang, J., Hartling, J. A., and Flanagan, J. M. (1997). *The structure of ClpP at 2.3* Å *resolution suggests a model for ATP-dependent proteolysis.* Cell 91, 447-456.

PROTEIN LOCALIZATION

2.8.33 Summary

A protein that is inserted into, or passes through, a membrane has a signal sequence that is recognized by a receptor that is part of the membrane or that can associate with it. The protein passes through an aqueous channel that is created by transmembrane protein(s) that reside in the membrane. In almost all cases, the protein passes through the channel in an unfolded form, and association with chaperones when it emerges is necessary in order to acquire the correct conformation. The major exception is the peroxisome, where an imported protein in its mature conformation binds to a cytosolic protein that carries it through the channel in the membrane.

Synthesis of proteins in the cytosol starts on "free" ribosomes. Proteins that are secreted from the cell or that are inserted into membranes of the reticuloendothelial system start with an N-terminal signal sequence that causes the ribosome to become attached to the membrane of the endoplasmic reticulum. The protein is translocated through the membrane by co-translational transfer. The process starts when the signal sequence is recognized by the SRP (a ribonucleoprotein particle), which interrupts translation. The SRP binds to the SRP receptor in the ER membrane, and transfers the signal sequence to the Sec61/TRAM receptor in the membrane. Synthesis resumes, and the protein is translocated through the membrane while it is being synthesized, although there is no energetic connection between the processes. The channel through the membrane provides a hydrophilic environment, and is largely made of the protein Sec61.

A secreted protein passes completely through the membrane into the ER lumen. Proteins that are integrated into membranes can be divided into two general types based on their orientation. For type I integral membrane proteins, the N-terminal signal sequence is cleaved, and transfer through the membrane is halted later by an anchor sequence. The protein becomes oriented in the membrane with its N-terminus on the far side and its C-terminus in the cytosol. Type II proteins do not have a cleavable N-terminal signal, but instead have a combined signal-anchor sequence, which enters the membrane and becomes embedded in it, causing the C-terminus to be located on the far side, while the N-terminus remains in the cytosol. The orientation of the signal-anchor is determined by the "positive inside" rule that the side of the anchor with more positive charges will be located in the cytoplasm. Proteins that have single transmembrane spanning regions move laterally from the channel into the lipid bilayer. Proteins may have multiple membrane-spanning regions, with loops between them protruding on either side of the membrane. The mechanism of insertion of multiple segments is unknown.

In the absence of any particular signal, a protein is released into the cytosol when its synthesis is completed. Proteins are imported post-translationally into mitochondria or chloroplasts. They possess N-terminal leader sequences that target them to the outer membrane of the organelle envelope; then they are transported through the outer and inner membranes into the matrix. Translocation requires ATP and a potential across the inner membrane. The N-terminal leader is cleaved by a protease within the organelle. Proteins that reside within the membranes or intermembrane

space possess a signal (which becomes N-terminal when the first part of the leader is removed) that either causes export from the matrix to the appropriate location or which halts transfer before all of the protein has entered the matrix. Control of folding, by Hsp70 and Hsp60 in the mitochondrial matrix, is an important feature of the process.

Mitochondria and chloroplasts have separate receptor complexes that create channels through each of the outer and inner membranes. All imported proteins pass directly from the TOM complex in the outer membrane to a TIM complex in the inner membrane. Proteins that reside in the inter-membrane space or in the outer membrane are re-exported from the TIM complex after entering the matrix. The TOM complex uses different receptors for imported proteins depending on whether they have N-terminal or internal signal sequences, and directs both types into the Tom40 channel. There are two TIM receptors in the inner membrane, one used for proteins whose ultimate destination is the inner matrix, the other used for proteins that are re-exported to the inter-membrane space or outer membrane.

Bacteria have components for membrane translocation that are related to those of the co-translational eukaryotic system, but translocation often occurs by a post-translational mechanism. SecY/E provide the translocase, and SecA associates with the channel and is involved in inserting and propelling the substrate protein. SecB is a chaperone that brings the protein to the channel. Some integral membrane proteins are inserted into the channel by an interaction with an apparatus resembling the SRP, consisting of 4.5S RNA and the Fth and FtsY proteins. The protein YidC is homologous to a mitochondrial protein and is required for insertion of some membrane proteins.

Nuclear pore complexes are massive structures embedded in the nuclear membrane, and are responsible for all transport of protein into the nucleus and RNA out of the nucleus. They have 8-fold symmetry seen in cross section and 2-fold symmetry viewed from the side. Each nuclear pore complex contains a central pore, which forms a channel of diameter <10 nm. The central channel can be opened to a diameter of ~20 nm to allow passage of larger material, some of which may need to undergo conformational changes to fit. The proteins of the complex are called nucleoporins; a subset called FG-nucleoporins have hydrophobic repeated sequences of Phe-Gly, are found in the central pore, and may be important in the translocation process.

Proteins that are actively transported into the nucleus require specific NLS sequences, which are short, but do not seem to share common features except for their basicity. Proteins that are exported from the nucleus have specific NES sequences, which share a pattern of leucine residues. Transport is a two stage process, involving docking followed by translocation. The docking reaction is undertaken by a transport receptor. Importins carry proteins into the nucleus, and exportins carry proteins out of the nucleus. The best characterized transport receptor is importin- α β, which has subunits that bind to the substrate protein and to a nucleoporin protein in the pore, respectively. Other transport receptors consist of single proteins that have both functions. The direction of translocation is controlled by Ran. The presence of Ran-GDP in the cytosol destabilizes export complexes. The presence of Ran-GTP in the nucleus destabilizes import complexes. This ensures release of substrate on the appropriate side of the nuclear envelope. ATP is required for translocation only in order to support the regeneration of Ran-GTP from

Ran-GDP; energy is not required for the translocation process itself. The mechanism of translocation is not understood in detail, but is likely to involve interactions of the transport receptors with the FG-nucleoporins.

The major system responsible for bulk degradation of proteins, but also for certain specific processing events, is the proteasome, a large complex that contains several protease activities. It acts upon substrate proteins that have been conjugated to ubiquitin through an isopeptide bond, and upon which a polyubiquitin chain has formed.