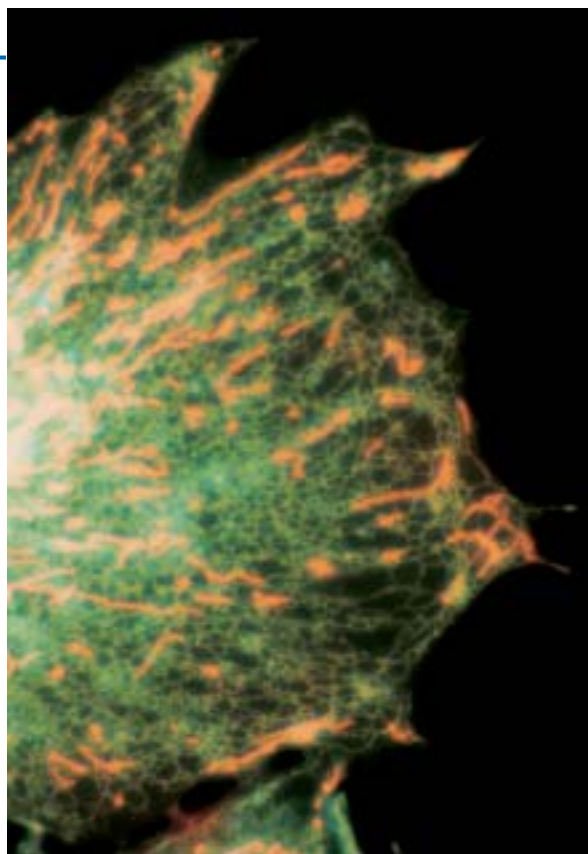


16

MOVING PROTEINS INTO MEMBRANES AND ORGANELLES



A live bovine endothelial cell stained to reveal different intracellular compartments. The lacelike membranes of the endoplasmic reticulum were stained with a green fluorescent dye, and the wormlike mitochondria were stained with an orange fluorescent dye. [Molecular Probes, Inc.]

A typical mammalian cell contains up to 10,000 different kinds of proteins; a yeast cell, about 5000. The vast majority of these proteins are synthesized by cytosolic ribosomes, and many remain within the cytosol. However, as many as half the different kinds of proteins produced in a typical cell are delivered to a particular cell membrane, an aqueous compartment other than the cytosol, or to the cell surface for secretion. For example, many hormone receptor proteins and transporter proteins must be delivered to the plasma membrane, some water-soluble enzymes such as RNA and DNA polymerases must be targeted to the nucleus, and components of the extracellular matrix as well as polypeptide signaling molecules must be directed to the cell surface for secretion from the cell. These and all the other proteins produced by a cell must reach their correct locations for the cell to function properly.

The delivery of newly synthesized proteins to their proper cellular destinations, usually referred to as *protein targeting* or *protein sorting*, encompasses two very different kinds of processes. The first general process involves targeting of a protein to the membrane of an intracellular organelle and can occur either during or soon after synthesis of the pro-

tein by translation at the ribosome. For membrane proteins, targeting leads to insertion of the protein into the lipid bilayer of the membrane, whereas for water-soluble proteins, targeting leads to translocation of the entire protein across the membrane into the aqueous interior of the organelle.

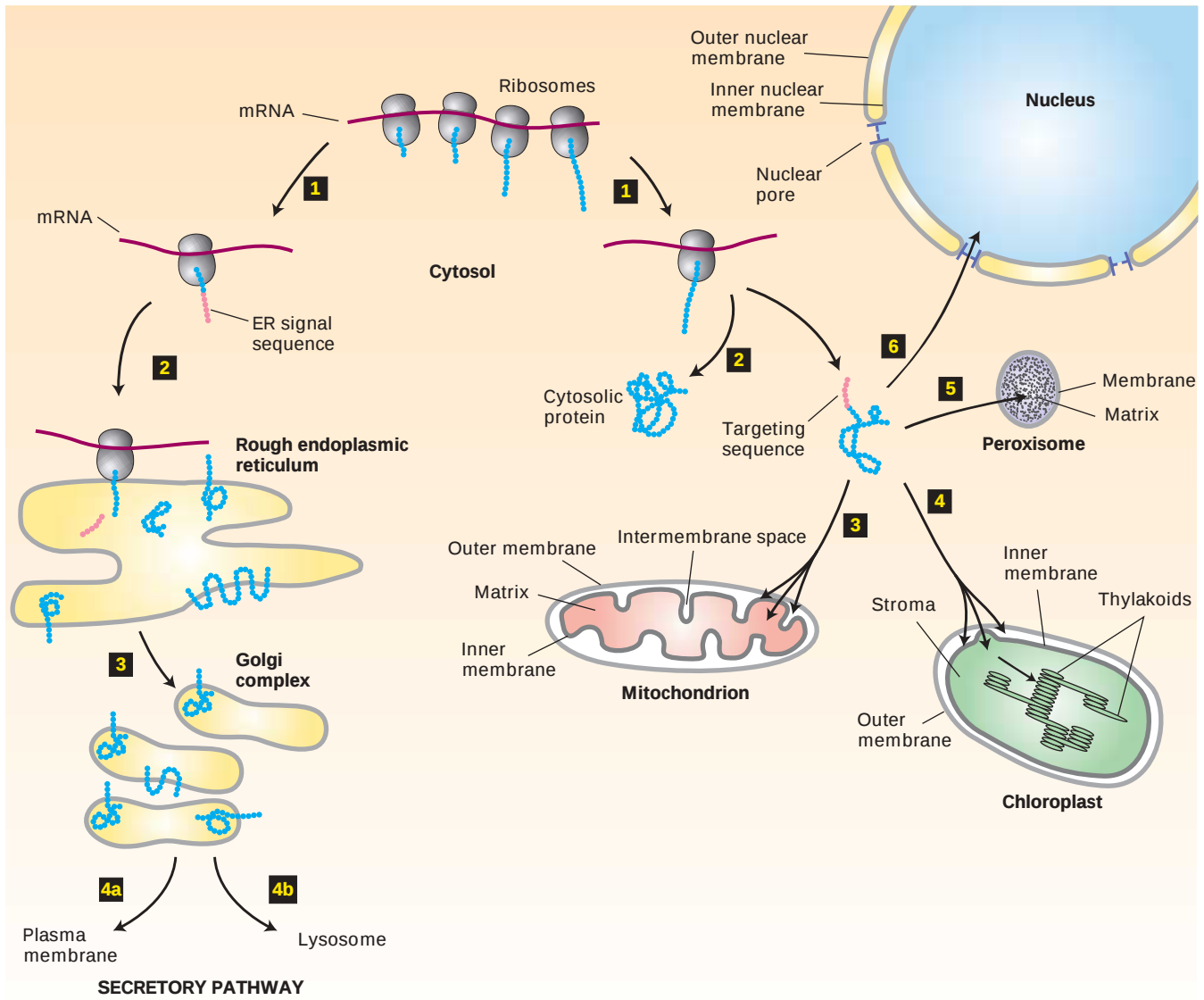
OUTLINE

- 16.1 Translocation of Secretory Proteins Across the ER Membrane
- 16.2 Insertion of Proteins into the ER Membrane
- 16.3 Protein Modifications, Folding, and Quality Control in the ER
- 16.4 Export of Bacterial Proteins
- 16.5 Sorting of Proteins to Mitochondria and Chloroplasts
- 16.6 Sorting of Peroxisomal Proteins

Proteins are sorted to the endoplasmic reticulum (ER), mitochondria, chloroplasts, peroxisomes, and the nucleus by this general process (Figure 16-1).

A second general sorting process applies to proteins that initially are targeted to the ER membrane, thereby entering

the **secretory pathway**. These proteins include not only soluble and membrane proteins that reside in the ER itself but also proteins that are secreted from the cell, enzymes and other resident proteins in the lumen of the Golgi complex and lysosomes, and integral proteins in the membranes of



▲ FIGURE 16-1 Overview of major protein-sorting pathways in eukaryotic cells. All nuclear-encoded mRNAs are translated on cytosolic ribosomes. *Left (secretory pathway):* Ribosomes synthesizing nascent proteins in the secretory pathway are directed to the rough endoplasmic reticulum (ER) by an ER signal sequence (pink; steps **1**, **2**). After translation is completed on the ER, these proteins can move via transport vesicles to the Golgi complex (step **3**). Further sorting delivers proteins either to the plasma membrane or to lysosomes (steps **4a**, **4b**). *Right (nonsecretory pathways):* Synthesis of proteins lacking an ER signal sequence is completed on free ribosomes

(step **1**). Those proteins that contain no targeting sequence are released into the cytosol and remain there (step **2**). Proteins with an organelle-specific targeting sequence (pink) first are released into the cytosol (step **2**) but then are imported into mitochondria, chloroplasts, peroxisomes, or the nucleus (steps **3–6**). Mitochondrial and chloroplast proteins typically pass through the outer and inner membranes to enter the matrix or stromal space, respectively. Other proteins are sorted to other subcompartments of these organelles by additional sorting steps. Nuclear proteins enter through visible nuclear pores by processes discussed in Chapter 12.

these organelles and the plasma membrane. Targeting to the ER generally involves *nascent* proteins still in the process of being synthesized. Proteins whose final destination is the Golgi, lysosome, or cell surface are transported along the secretory pathway by small vesicles that bud from the membrane of one organelle and then fuse with the membrane of the next organelle in the pathway (see Figure 16-1, *left*). We discuss vesicle-based protein sorting in the next chapter because mechanistically it differs significantly from protein targeting to the membranes of intracellular organelles.

In this chapter, we examine how proteins are targeted to the membrane of intracellular organelles and subsequently inserted into the organelle membrane or moved into the interior. Two features of this protein-sorting process initially were quite baffling: how a given protein could be targeted to only one specific membrane, and how relatively large protein molecules could be translocated across a membrane without disrupting the bilayer as a barrier to ions and small molecules. Using a combination of biochemical purification methods and genetic screens for identifying mutants unable to execute particular translocation steps, cell biologists have identified many of the cellular components required for translocation across each of the different intracellular membranes. In addition, many of the major translocation processes in the cell have been reconstituted using *in vitro* systems, which can be freely manipulated experimentally.

These studies have shown that despite some variations, the same basic mechanisms govern protein sorting to all the various intracellular organelles. We now know, for instance, that the information to target a protein to a particular organelle destination is encoded within the amino acid sequence of the protein itself, usually within sequences of 20–50 amino acids, known generically as **signal sequences**, or *uptake-targeting sequences* (see Figure 16-1). Each organelle carries a set of receptor proteins that bind only to specific kinds of signal sequences, thus assuring that the information encoded in a signal sequence governs the specificity of targeting. Once a protein containing a signal sequence has interacted with the corresponding receptor, the protein chain is transferred to some kind of *translocation channel* that allows the protein to pass through the membrane bilayer. The unidirectional transfer of a protein into an organelle, without sliding back out into the cytoplasm, is usually achieved by coupling translocation to an energetically favorable process such as hydrolysis of ATP. Some proteins are subsequently sorted further to reach a subcompartment within the target organelle; such sorting depends on yet other signal sequences and other receptor proteins. Finally, signal sequences often are removed from the mature protein by specific proteases once translocation across the membrane is completed.

For each of the protein-targeting events discussed in this chapter, we will seek to answer four fundamental questions:

1. What is the nature of the *signal sequence*, and what distinguishes it from other types of signal sequences?

2. What is the *receptor* for the signal sequence?

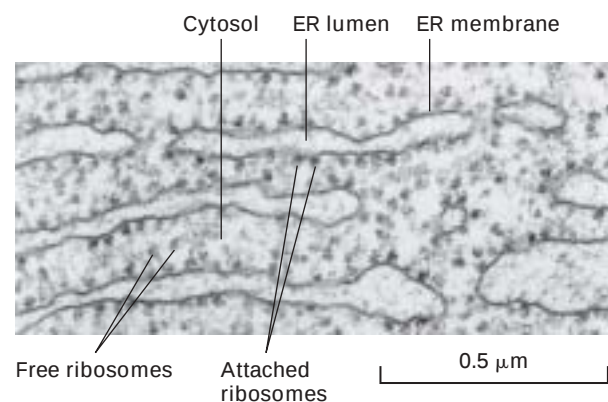
3. What is the structure of the *translocation channel* that allows transfer of proteins across the membrane bilayer? In particular, is the channel so narrow that proteins can pass through only in an unfolded state, or will it accommodate folded protein domains?

4. What is the source of *energy* that drives unidirectional transfer across the membrane?

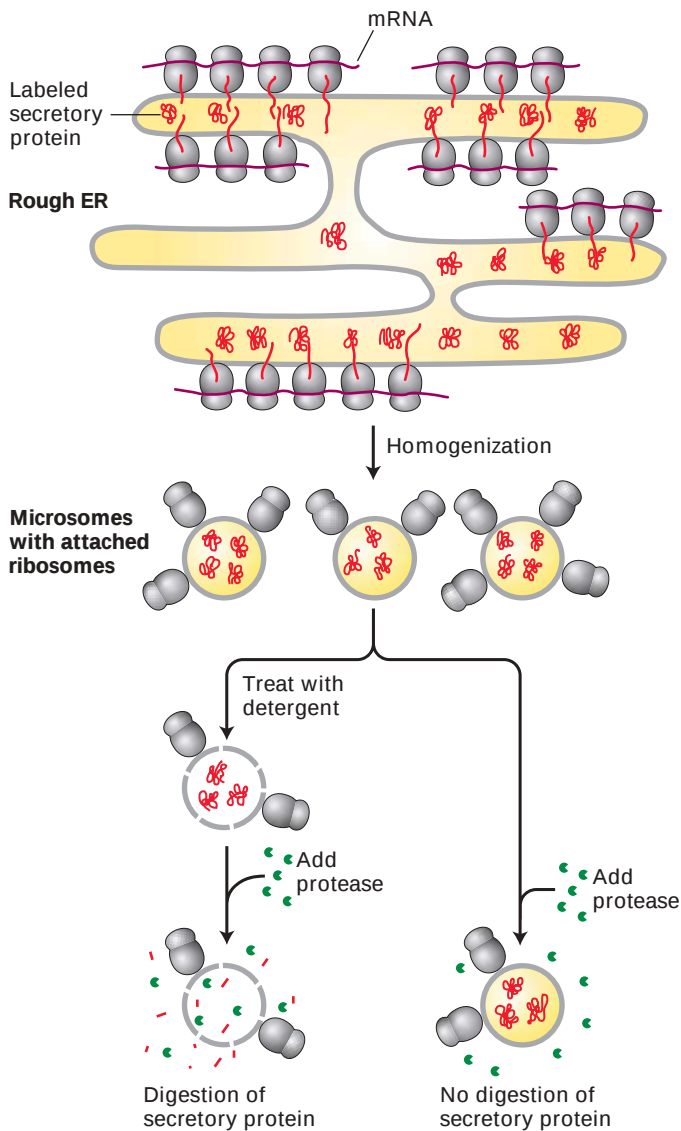
In the first part of the chapter, we cover targeting of proteins to the ER, including the post-translational modifications that occur to proteins as they enter the secretory pathway. We then look at several mechanisms for exporting proteins from bacteria, some of which are similar to protein sorting in eukaryotic cells. The last two sections describe targeting of proteins to mitochondria, chloroplasts, and peroxisomes. We cover the transport of proteins in and out of the nucleus through nuclear pores in Chapter 12 because nuclear transport is intimately related to post-transcriptional events and involves several nucleus-specific variations on the mechanisms discussed in this chapter.

16.1 Translocation of Secretory Proteins Across the ER Membrane

All eukaryotic cells use essentially the same secretory pathway for synthesizing and sorting secreted proteins and soluble luminal proteins in the ER, Golgi, and lysosomes (see Figure 16-1, *left*). For simplicity, we refer to these proteins collectively as *secretory proteins*. Although all cells secrete a variety of proteins (e.g., extracellular matrix proteins), certain types of cells are specialized for secretion of large



▲ **FIGURE 16-2 Electron micrograph of ribosomes attached to the rough ER in a pancreatic acinar cell.** Most of the proteins synthesized by this type of cell are to be secreted and are formed on membrane-attached ribosomes. A few membrane-unattached (free) ribosomes are evident; presumably, these are synthesizing cytosolic or other nonsecretory proteins. [Courtesy of G. Palade.]



◀ **EXPERIMENTAL FIGURE 16-3 Labeling experiments demonstrate that secretory proteins are localized to the ER lumen shortly after synthesis.** Cells are incubated for a brief time with radiolabeled amino acids, so that only newly synthesized proteins become labeled. The cells then are homogenized, fracturing the plasma membrane and shearing the rough ER into small vesicles called microsomes. Because they have bound ribosomes, *microsomes* have a much greater buoyant density than other membranous organelles and can be separated from them by a combination of differential and sucrose density-gradient centrifugation (Chapter 5). The purified microsomes are treated with a protease in the presence or absence of a detergent. The labeled secretory proteins associated with the microsomes are digested by added proteases only if the permeability barrier of the microsomal membrane is first destroyed by treatment with detergent. This finding indicates that the newly made proteins are inside the microsomes, equivalent to the lumen of the rough ER.

which microsomes isolated from pulse-labeled cells are treated with a protease, demonstrate that although secretory proteins are synthesized on ribosomes bound to the cytosolic face of the ER membrane, they become localized in the lumen of ER vesicles during their synthesis.

A Hydrophobic N-Terminal Signal Sequence Targets Nascent Secretory Proteins to the ER

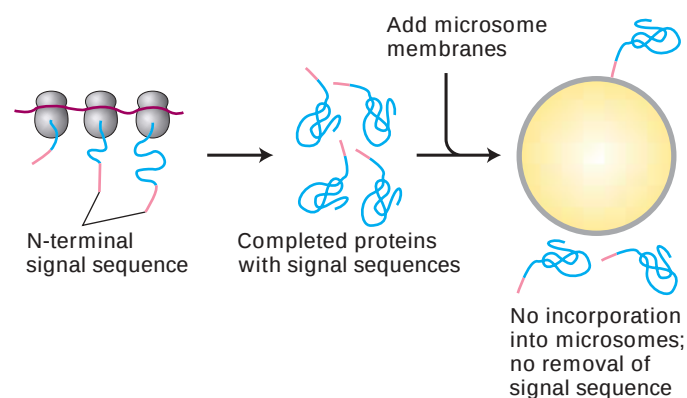
After synthesis of a secretory protein begins on free ribosomes in the cytosol, a 16- to 30-residue ER signal sequence in the nascent protein directs the ribosome to the ER membrane and initiates translocation of the growing polypeptide across the ER membrane (see Figure 16-1, *left*). An ER signal sequence typically is located at the N-terminus of the protein, the first part of the protein to be synthesized. The signal sequences of different secretory proteins contain one or more positively charged amino acids adjacent to a continuous stretch of 6–12 hydrophobic residues (the core), but otherwise they have little in common. For most secretory proteins, the signal sequence is cleaved from the protein while it is still growing on the ribosome; thus, signal sequences are usually not present in the “mature” proteins found in cells.

The hydrophobic core of ER signal sequences is essential for their function. For instance, the specific deletion of several of the hydrophobic amino acids from a signal sequence, or the introduction of charged amino acids into the hydrophobic core by mutation, can abolish the ability of the N-terminus of a protein to function as a signal sequence. As a consequence, the modified protein remains in the cytosol, unable to cross the ER membrane into the lumen. Using recombinant DNA techniques, researchers have produced cytosolic proteins with added N-terminal amino acid sequences. Provided the added sequence is sufficiently long and hydrophobic, such a modified cytosolic protein is translocated to the ER lumen. Thus the hydrophobic residues

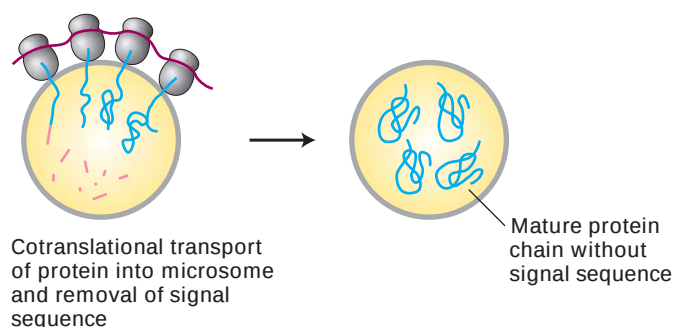
amounts of specific proteins. Pancreatic acinar cells, for instance, synthesize large quantities of several digestive enzymes that are secreted into ductules that lead to the intestine. Because such secretory cells contain the organelles of the secretory pathway (e.g., ER and Golgi) in great abundance, they have been widely used in studying this pathway.

Early pulse-labeling experiments with pancreatic acinar cells showed that radioactively labeled amino acids are incorporated primarily into newly synthesized secretory proteins. The ribosomes synthesizing these proteins are actually bound to the surface of the ER. As a consequence, the portion of the ER that receives proteins entering the secretory pathway is known as the *rough ER* because these membranes are densely studded with ribosomes (Figure 16-2). When cells are homogenized, the rough ER breaks up into small closed vesicles, termed *rough microsomes*, with the same orientation (ribosomes on the outside) as that found in the intact cell. The experiments depicted in Figure 16-3, in

(a) Cell-free protein synthesis; no microsomes present



(b) Cell-free protein synthesis; microsomes present



in the core of ER signal sequences form a binding site that is critical for the interaction of signal sequences with receptor proteins on the ER membrane.

Biochemical studies utilizing a cell-free protein-synthesizing system, mRNA encoding a secretory protein, and microsomes stripped of their own bound ribosomes have clarified the function and fate of ER signal sequences. Initial experiments with this system demonstrated that a typical secretory protein is incorporated into microsomes and has its signal sequence removed only if the microsomes are present during protein synthesis (Figure 16-4). Subsequent experiments were designed to determine the precise stage of protein synthesis at which microsomes must be present in order for translocation to occur. In these experiments, a drug that prevents initiation of translation was added to protein-synthesizing reactions at different times after protein synthesis had begun, and then stripped microsomes were added to the reaction mixtures. These experiments showed that microsomes must be added before the first 70 or so amino acids are linked together in order for the completed secretory protein to be localized in the microsomal lumen. At this point, the first 40 amino acids or so protrude from the ribosome, including the signal sequence that later will be cleaved off, and the next 30 or so amino acids are still buried within a channel in the ribosome. Thus the transport of most secretory proteins into the ER lumen occurs while the nascent protein is still bound to the ribosome and being elongated, a process referred to as **cotranslational translocation**.

◀ EXPERIMENTAL FIGURE 16-4 Cell-free experiments demonstrate that translocation of secretory proteins into microsomes is coupled to translation.

Treatment of microsomes with EDTA, which chelates Mg^{+2} ions, strips them of associated ribosomes, allowing isolation of ribosome-free microsomes, which are equivalent to ER membranes (see Figure 16-3). Synthesis is carried out in a cell-free system containing functional ribosomes, tRNAs, ATP, GTP, and cytosolic enzymes to which mRNA encoding a secretory protein is added. The secretory protein is synthesized in the absence of microsomes (a), but is translocated across the vesicle membrane and loses its signal sequence only if microsomes are present during protein synthesis (b).

Cotranslational Translocation Is Initiated by Two GTP-Hydrolyzing Proteins

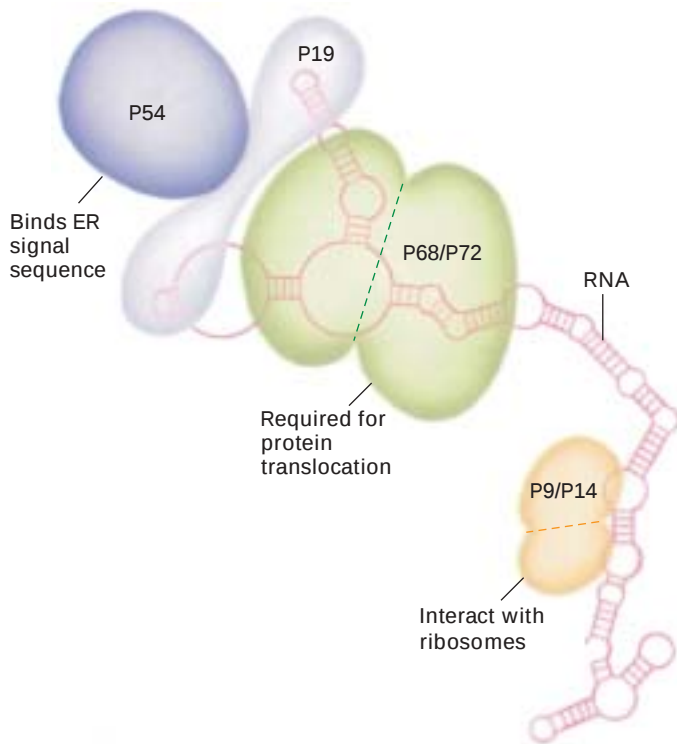
Since secretory proteins are synthesized in association with the ER membrane but not with any other cellular membrane, a signal-sequence recognition mechanism must target them there. The two key components in this targeting are the **signal-recognition particle (SRP)** and its receptor located in the ER membrane. The SRP is a cytosolic ribonucleoprotein particle that transiently binds simultaneously to the ER signal sequence in a nascent protein, to the large ribosomal unit, and to the SRP receptor.

Six discrete polypeptides and a 300-nucleotide RNA compose the SRP (Figure 16-5a). One of the SRP proteins (P54) can be chemically cross-linked to ER signal sequences, evidence that this particular protein is the subunit that binds to the signal sequence in a nascent secretory protein. A region of P54 containing many amino acid residues with hydrophobic side chains is homologous to a bacterial protein known as Ffh, which performs an analogous function to P54 in the translocation of proteins across the inner membrane of bacterial cells. The structure of Ffh contains a cleft whose inner surface is lined by hydrophobic side chains (Figure 16-5b). The hydrophobic region of P54 is thought to contain an analogous cleft that interacts with the hydrophobic N-termini of nascent secretory proteins and selectively targets them to the ER membrane. Two of the SRP proteins, P9 and P14, interact with the ribosome, while P68 and P72 are required for protein translocation.

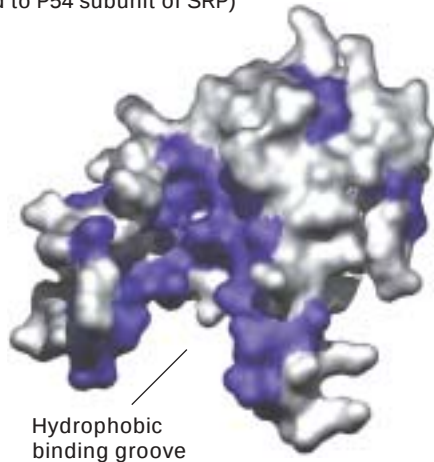
In the cell-free translation system described previously, the presence of SRP slows elongation of a secretory protein when microsomes are absent, thereby inhibiting synthesis of the complete protein (see Figure 16-4). This finding suggests that interaction of the SRP with both the nascent chain of a secretory protein and with the free ribosome prevents the nascent chain from becoming too long for translocation into the ER. Only after the SRP/nascent chain/ribosome complex has bound to the SRP receptor in the ER membrane does SRP release the nascent chain, allowing elongation at the normal rate.

Figure 16-6 summarizes our current understanding of secretory protein synthesis and the role of the SRP and its

(a) Signal-recognition particle (SRP)



(b) Ffh signal sequence-binding domain (related to P54 subunit of SRP)



◀ FIGURE 16-5 Structure of the signal-recognition particle (SRP).

(a) The SRP comprises one 300-nucleotide RNA and six proteins designated P9, P14, P19, P54, P68, and P72. (The numeral indicates the molecular weight $\times 10^3$.) All proteins except P54 bind directly to the RNA. (b) The bacterial Ffh protein is homologous to the portion of P54 that binds ER signal sequences. This surface model shows the binding domain in Ffh, which contains a large cleft lined with hydrophobic amino acids (purple) whose side chains interact with signal sequences. [Part (a) see K. Strub et al., 1991, *Mol. Cell Biol.* **11**:3949; and S. High and B. Dobberstein, 1991, *J. Cell Biol.* **113**:229. Part (b) adapted from R. J. Keenan et al., 1998, *Cell* **94**:181]

nascent secretory protein with the ER membrane but also act together to permit elongation and synthesis of complete proteins only when ER membranes are present.

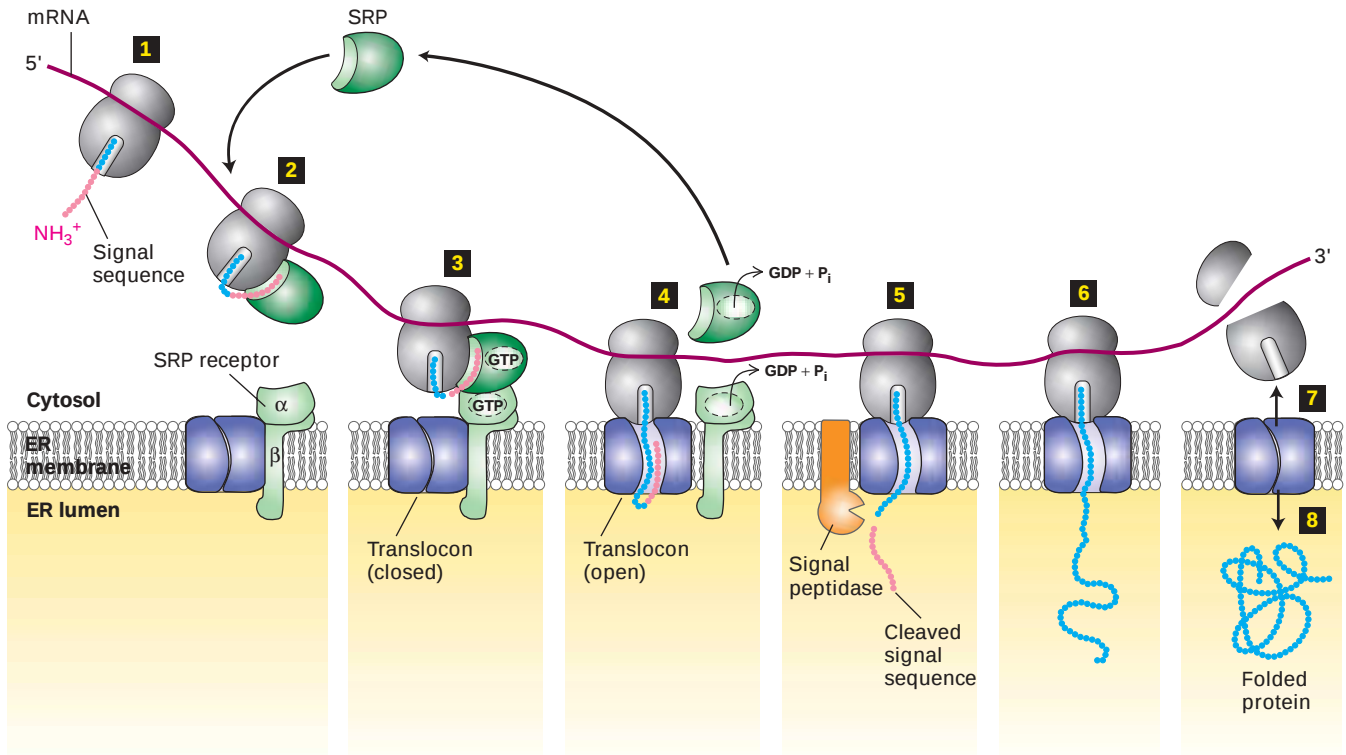
Ultimately, the SRP and SRP receptor function to bring ribosomes that are synthesizing secretory proteins to the ER membrane. The coupling of GTP hydrolysis to this targeting process is thought to contribute to the fidelity by which signal sequences are recognized. Probably the energy from GTP hydrolysis is used to release proteins lacking proper signal sequences from the SRP and SRP receptor complex, thereby preventing their mistargeting to the ER membrane. (A similar coupling of GTP hydrolysis with binding of translation elongation factors to ribosomes increases the fidelity of translation by ejecting aminoacyl-tRNA molecules that cannot form correct base pairs with the codons in mRNA.) Interaction of the SRP/nascent chain/ribosome complex with the SRP receptor is promoted when GTP is bound by both the P54 subunit of SRP and the α subunit of the SRP receptor (see Figure 16-6). Subsequent transfer of the nascent chain and ribosome to a site on the ER membrane where translocation can take place allows hydrolysis of the bound GTP. After dissociating, SRP and its receptor release the bound GDP and recycle to the cytosol ready to initiate another round of interaction between ribosomes synthesizing nascent secretory proteins with the ER membrane.

Passage of Growing Polypeptides Through the Translocon Is Driven by Energy Released During Translation

receptor in this process. The SRP receptor is an integral membrane protein made up of two subunits: an α subunit and a smaller β subunit. Treatment of microsomes with very small amounts of protease cleaves the α subunit very near its site of attachment to the membrane, releasing a soluble form of the SRP receptor. Protease-treated microsomes are unable to bind the SRP/nascent chain/ribosome complex or support cotranslational translocation. The soluble SRP receptor fragment, however, retains its ability to interact with the SRP/nascent chain/ribosome complex, causing release of SRP and allowing chain elongation to proceed. Thus the SRP and SRP receptor not only help mediate interaction of a

Once the SRP and its receptor have targeted a ribosome synthesizing a secretory protein to the ER membrane, the ribosome and nascent chain are rapidly transferred to the **translocon**, a protein-lined channel within the membrane. As translation continues, the elongating chain passes directly from the large ribosomal subunit into the central pore of the translocon. The 60S ribosomal subunit is aligned with the pore of the translocon in such a way that the growing chain is never exposed to the cytoplasm and does not fold until it reaches the ER lumen (see Figure 16-6).

The translocon was first identified by mutations in the yeast gene encoding Sec61 α , which caused a block in the



▲ FIGURE 16-6 Synthesis of secretory proteins and their cotranslational translocation across the ER membrane. Steps **1**, **2**: Once the ER signal sequence emerges from the ribosome, it is bound by a signal-recognition particle (SRP). Step **3**: The SRP delivers the ribosome/nascent polypeptide complex to the SRP receptor in the ER membrane. This interaction is strengthened by binding of GTP to both the SRP and its receptor. Step **4**: Transfer of the ribosome/nascent polypeptide to the translocon leads to opening of this translocation channel and insertion of the signal sequence and adjacent segment of the growing polypeptide into the central pore. Both the SRP and SRP receptor, once dissociated from the translocon, hydrolyze their bound GTP and then are ready to

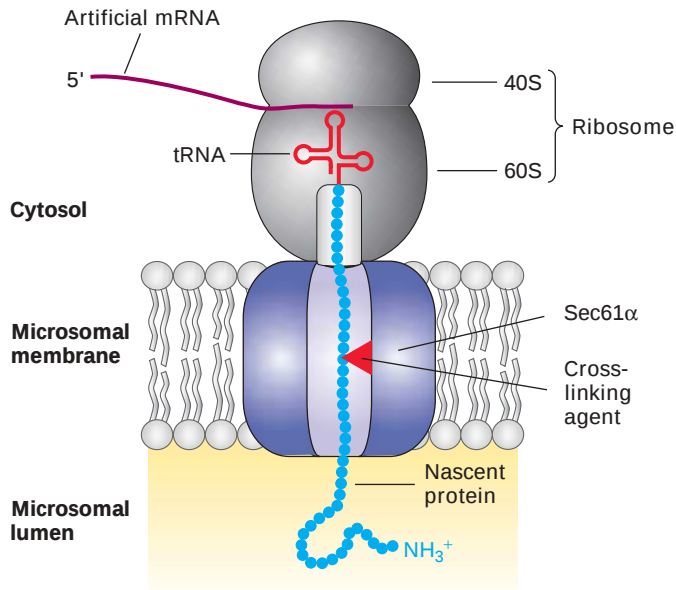
initiate the insertion of another polypeptide chain. Step **5**: As the polypeptide chain elongates, it passes through the translocon channel into the ER lumen, where the signal sequence is cleaved by signal peptidase and is rapidly degraded. Step **6**: The peptide chain continues to elongate as the mRNA is translated toward the 3' end. Because the ribosome is attached to the translocon, the growing chain is extruded through the translocon into the ER lumen. Steps **7**, **8**: Once translation is complete, the ribosome is released, the remainder of the protein is drawn into the ER lumen, the translocon closes, and the protein assumes its native folded conformation.

translocation of secretory proteins into the lumen of the ER. Subsequently, three proteins called the *Sec61 complex* were found to form the mammalian translocon: Sec61 α , an integral membrane protein with 10 membrane-spanning α helices, and two smaller proteins, termed Sec61 β and Sec61 γ . Chemical cross-linking experiments demonstrated that the translocating polypeptide chain comes into contact with the Sec61 α protein in both yeast and mammalian cells, confirming its identity as a translocon component (Figure 16-7).

When microsomes in the cell-free translocation system were replaced with reconstituted phospholipid vesicles containing only the SRP receptor and Sec61 complex, nascent

secretory protein was translocated from its SRP/ribosome complex into the vesicles. This finding indicates that the SRP receptor and the Sec61 complex are the only ER-membrane proteins absolutely required for translocation. Thus the energy derived from chain elongation at the ribosome appears to be sufficient to push the polypeptide chain across the membrane in one direction.

Multiple copies of the Sec61 complex assemble to form a translocon channel, which can be visualized by electron microscopy. Images of the channel have been generated by computer averaging of electron micrographs of purified Sec61 channels bound to ribosomes (Figure 16-8). These show the channel as a cylinder, 5–6 nm high and 8.5 nm in

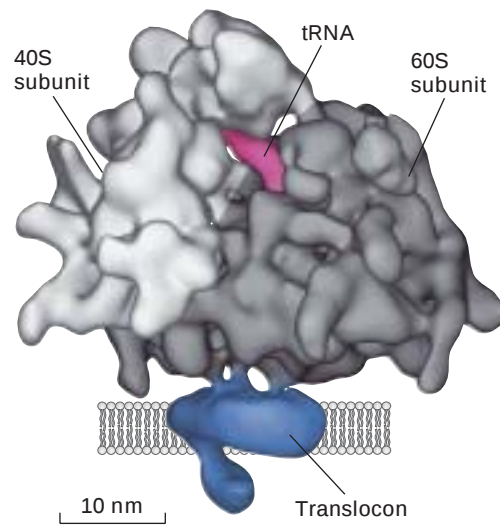


▲ **EXPERIMENTAL FIGURE 16-7 Cross-linking experiments show that Sec61 α is a translocon component that contacts nascent secretory proteins as they pass into the ER lumen.**

An mRNA encoding the N-terminal 70 amino acids of the secreted protein prolactin was translated in a cell-free system containing microsomes (see Figure 16-4b). The mRNA lacked a chain-termination codon and contained one lysine codon, near the middle of the sequence. The reactions contained a chemically modified lysyl-tRNA in which a light-activated cross-linking reagent was attached to the lysine side chain. Although the entire mRNA was translated, the completed polypeptide could not be released from the ribosome and thus became “stuck” crossing the ER membrane. The reaction mixtures then were exposed to an intense light, causing the nascent chain to become covalently bound to whatever proteins were near it in the translocon. When the experiment was performed using microsomes from mammalian cells, the nascent chain became covalently linked to Sec61 α . Different versions of the prolactin mRNA were used to place the modified lysine residue at different distances from the ribosome; cross-linking to Sec61 α was observed only when the modified lysine was positioned within the translocation channel. [Adapted from T. A. Rapoport, 1992, *Science* **258**:931, and D. Görlich and T. A. Rapoport, 1993, *Cell* **75**:615.]

diameter, with a central pore, roughly 2 nm in diameter, perpendicular to the plane of the membrane.

If the translocon were always open in the ER membrane, especially in the absence of attached ribosomes and a translocating polypeptide, small molecules such as ATP and amino acids would be able to diffuse freely through the central pore. To maintain the permeability barrier of the ER membrane, the translocon is regulated so that it is open only when a ribosome–nascent chain complex is bound. Thus the translocon is a gated channel analogous to the gated ion channels described in Chapter 7. When the translocon first opens, a loop of the nascent chain, containing the signal



▲ **EXPERIMENTAL FIGURE 16-8 Electron microscopy reconstruction reveals that a translocon associates closely with a ribosome.** Purified Sec61 complexes were solubilized by treatment of ER membranes with detergents. When ribosomes were added, translocons (blue) reassembled in artificial phospholipid bilayers. The resulting particles were frozen, and electron micrographs of a large number of particles were generated, stored in a computer, and then averaged to produce a single image. A representation of the approximate size and position of the ER lipid bilayer has been added. Note that although the ribosome is firmly attached to the translocon, there is a gap between the two structures. The fingerlike appendage below the translocon channel is thought to be formed from a protein complex that associates with the translocon. [Courtesy Dr. Christopher Akey and Jean-Francois Menetret, Boston University School of Medicine.]

sequence and approximately 30 adjacent amino acids, can insert into the translocon pore (see Figure 16-6).

The mechanism by which the translocon channel opens and closes is controversial at this time. Some evidence suggests that a protein within the ER lumen blocks the translocon pore when a ribosome is not bound to the cytosolic side of the translocon. Other observations, however, indicate that Sec61 complexes may normally reside in the ER membrane in an unassembled state and that the gating process involves the assembly of a translocon channel at the site where the ribosome and nascent chain are brought to the membrane by the SRP and SRP receptor.

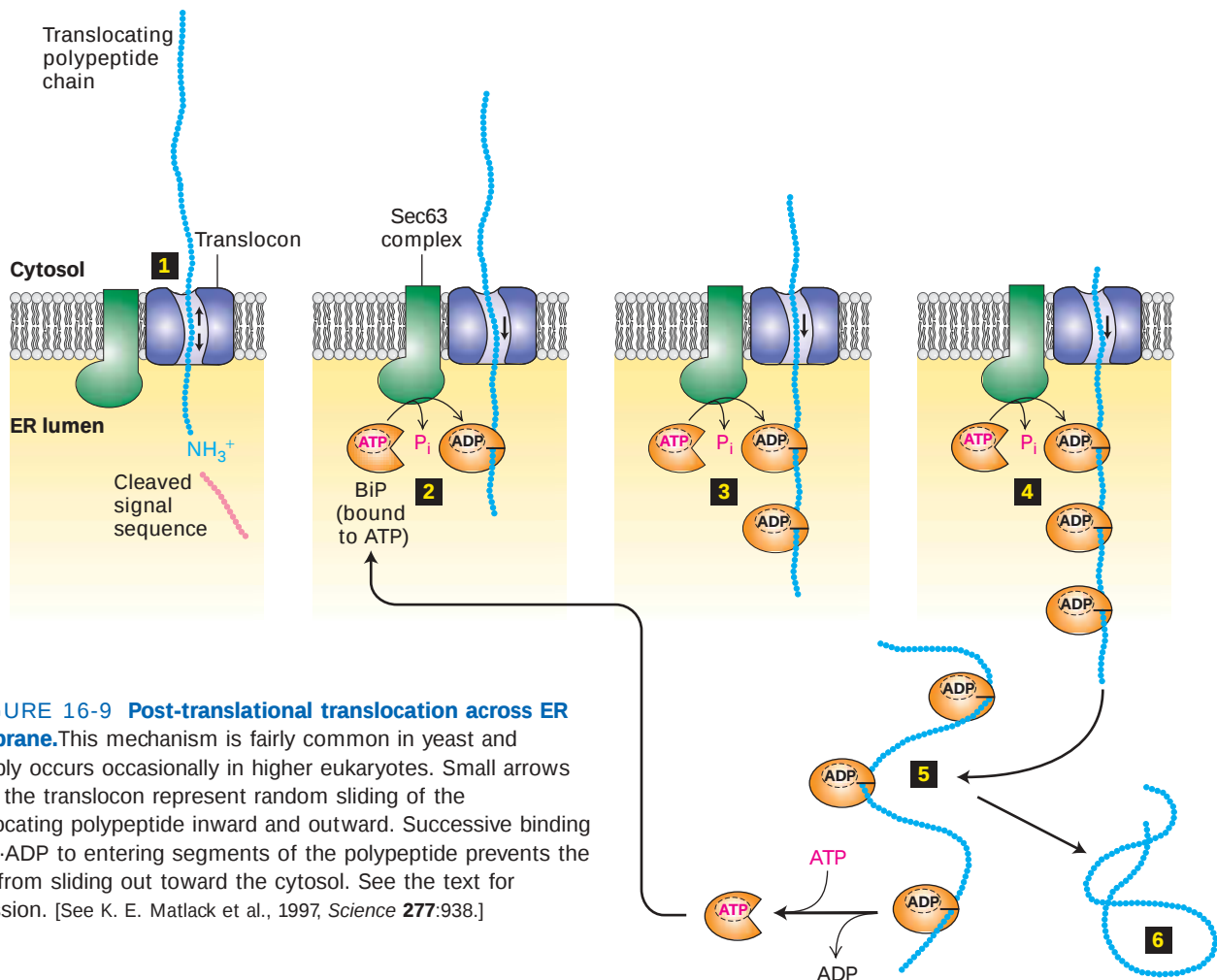
As the growing polypeptide chain enters the lumen of the ER, the signal sequence is cleaved by *signal peptidase*, which is a transmembrane ER protein associated with the translocon (see Figure 16-6). This protease recognizes a sequence on the C-terminal side of the hydrophobic core of the signal peptide and cleaves the chain specifically at this sequence once it has emerged into the luminal space of the ER. After the signal sequence has been cleaved, the growing polypeptide moves through the translocon into the ER lumen. The translocon remains open until translation is completed and the entire polypeptide chain has moved into the ER lumen.

ATP Hydrolysis Powers Post-translational Translocation of Some Secretory Proteins in Yeast

In most eukaryotes, secretory proteins enter the ER by cotranslational translocation, using energy derived from translation to pass through the membrane, as we've just described. In yeast, however, some secretory proteins enter the ER lumen after translation has been completed. In such *post-translational translocation*, the translocating protein passes through the same Sec61 translocon that is used in cotranslational translocation. However, the SRP and SRP receptor are not involved in post-translational translocation, and in such cases a direct interaction between the translocon and the signal sequence of the completed protein appears to be sufficient for targeting to the ER membrane. In addition, the driving force for unidirectional translocation across the ER membrane is provided by an additional protein complex known as the Sec63 complex and a member of the Hsc70 family of molecular **chaperones** known as *BiP*. The tetrameric Sec63 complex is embedded in the ER membrane in the vicinity of the translocon, while *BiP* is localized to the ER lumen. Like other members of the Hsc70 family, *BiP* has

a peptide-binding domain and an ATPase domain. These chaperones bind and stabilize unfolded or partially folded proteins (see Figure 3-11).

The current model for post-translational translocation of a protein into the ER is outlined in Figure 16-9. Once the N-terminal segment of the protein enters the ER lumen, signal peptidase cleaves the signal sequence just as in cotranslational translocation (step 1). Interaction of *BiP*·ATP with the luminal portion of the Sec63 complex causes hydrolysis of the bound ATP, producing a conformational change in *BiP* that promotes its binding to an exposed polypeptide chain (step 2). Since the Sec63 complex is located near the translocon, *BiP* is thus activated at sites where nascent polypeptides can enter the ER. Certain experiments suggest that in the absence of binding to *BiP*, an unfolded polypeptide slides back and forth within the translocon channel. Such random sliding motions rarely result in the entire polypeptide's crossing the ER membrane. Binding of a molecule of *BiP*·ADP to the luminal portion of the polypeptide prevents backsliding of the polypeptide out of the ER. As further inward random sliding exposes more of the polypeptide on the luminal side of the ER membrane, successive binding of



▲ **FIGURE 16-9 Post-translational translocation across ER membrane.** This mechanism is fairly common in yeast and probably occurs occasionally in higher eukaryotes. Small arrows inside the translocon represent random sliding of the translocating polypeptide inward and outward. Successive binding of *BiP*·ADP to entering segments of the polypeptide prevents the chain from sliding out toward the cytosol. See the text for discussion. [See K. E. Matlack et al., 1997, *Science* **277**:938.]

BiP·ADP molecules to the polypeptide chain acts as a ratchet, ultimately drawing the entire polypeptide into the ER within a few seconds (steps 3 and 4). On a slower time scale, the BiP molecules spontaneously exchange their bound ADP for ATP, leading to release of the polypeptide, which can then fold into its native conformation (steps 5 and 6). The recycled BiP·ATP then is ready for another interaction with Sec63.

The overall reaction carried out by BiP is an important example of how the chemical energy released by the hydrolysis of ATP can power the mechanical movement of a protein across a membrane. Some bacterial cells also use an ATP-driven process for translocating completed proteins across the plasma membrane. However, the mechanism of post-translational translocation in bacteria differs somewhat from that in yeast, as we describe in Section 16.4.

KEY CONCEPTS OF SECTION 16.1

Translocation of Secretory Proteins Across the ER Membrane

- Synthesis of secreted proteins; enzymes destined for the ER, Golgi complex, or lysosome; and integral plasma-membrane proteins begins on cytosolic ribosomes, which become attached to the membrane of the ER, forming the rough ER (see Figure 16-1, *left*).
- The ER signal sequence on a nascent secretory protein consists of a segment of hydrophobic amino acids, generally located at the N-terminus.
- In cotranslational translocation, the signal-recognition particle (SRP) first recognizes and binds the ER signal sequence on a nascent secretory protein and in turn is bound by an SRP receptor on the ER membrane, thereby targeting the ribosome/nascent chain complex to the ER.
- The SRP and SRP receptor then mediate insertion of the nascent secretory protein into the translocon. Hydrolysis of GTP by the SRP and its receptor drive this docking process (see Figure 16-6). As the ribosome attached to the translocon continues translation, the unfolded protein chain is extruded into the ER lumen. No additional energy is required for translocation.
- In post-translational translocation, a completed secretory protein is targeted to the ER membrane by interaction of the signal sequence with the translocon. The polypeptide chain is then pulled into the ER by a ratcheting mechanism that requires ATP hydrolysis by the chaperone BiP, which stabilizes the entering polypeptide (see Figure 16-9).
- In both cotranslational and post-translational translocation, a signal peptidase in the ER membrane cleaves the ER signal sequence from a secretory protein soon after the N-terminus enters the lumen.

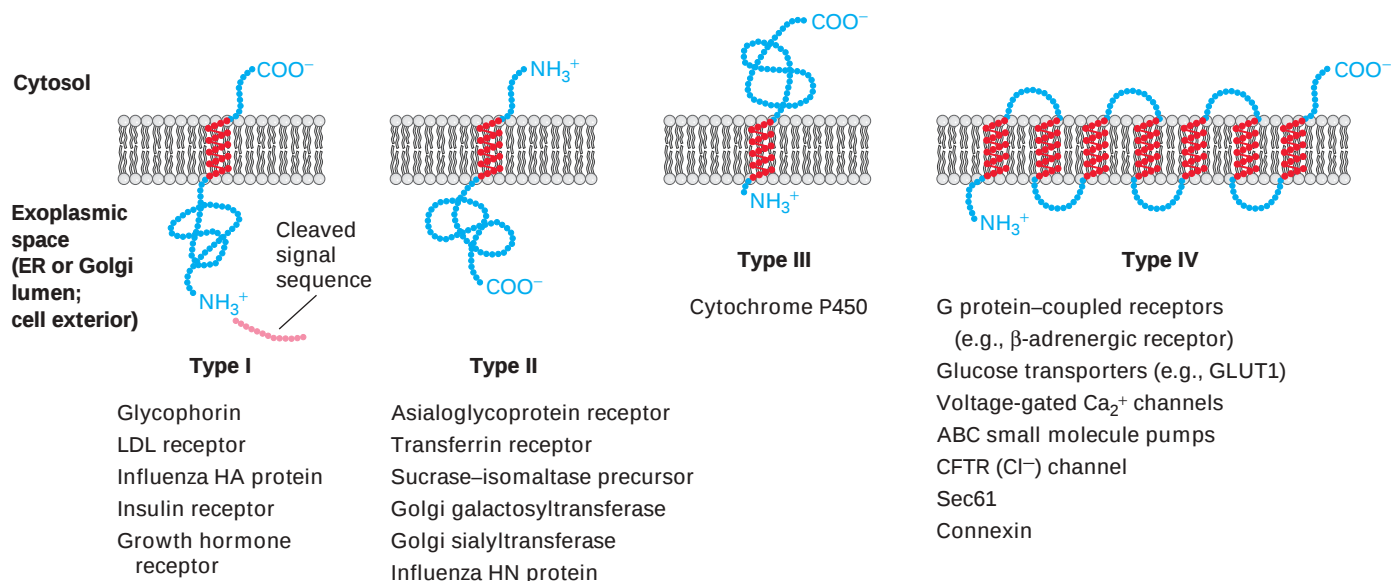
16.2 Insertion of Proteins into the ER Membrane

In previous chapters we have encountered many of the vast array of integral (transmembrane) proteins that are present in the plasma membrane and other cellular membranes. Each such protein has a unique orientation with respect to the membrane's phospholipid bilayer. Integral proteins located in ER, Golgi, and lysosomal membranes and in the plasma membrane, which are synthesized on the rough ER, remain embedded in the membrane as they move to their final destinations along the same pathway followed by soluble secretory proteins (see Figure 16-1). During this transport, the orientation of a membrane protein is preserved; that is, the same segments of the protein always face the cytosol, while other segments always face in the opposite direction. Thus the final orientation of these membrane proteins is established during their biosynthesis on the ER membrane. In this section, we first see how integral proteins can interact with membranes. Then we examine how several types of sequences, collectively known as **topogenic sequences**, direct the insertion and orientation of various classes of integral proteins into the membrane. These processes build on and adapt the basic mechanism used to translocate soluble secretory proteins across the ER membrane.

Several Topological Classes of Integral Membrane Proteins Are Synthesized on the ER

The *topology* of a membrane protein refers to the number of times that its polypeptide chain spans the membrane and the orientation of these membrane-spanning segments within the membrane. The key elements of a protein that determine its topology are membrane-spanning segments themselves, which usually contain 20–25 hydrophobic amino acids. Each such segment forms an α helix that spans the membrane, with the hydrophobic amino acid residues anchored to the hydrophobic interior of the phospholipid bilayer.

Scientists have found it useful to categorize integral membrane proteins into the four topological classes illustrated in Figure 16-10. Topological classes I, II, and III comprise *single-pass* proteins, which have only one membrane-spanning α -helical segment. Type I proteins have a cleaved N-terminal signal sequence and are anchored in the membrane with their hydrophilic N-terminal region on the luminal face (also known as the **exoplasmic face**) and their hydrophilic C-terminal region on the cytosolic face. Type II proteins do not contain a cleavable signal sequence and are oriented with their hydrophilic N-terminal region on the cytosolic face and their hydrophilic C-terminal region on the exoplasmic face (i.e., opposite to type I proteins). Type III proteins have the same orientation as type I proteins, but do not contain a cleavable signal sequence. These different topologies reflect distinct mechanisms used by the cell to



▲ **FIGURE 16-10 Major topological classes of integral membrane proteins synthesized on the rough ER.** The hydrophobic segments of the protein chain form α helices embedded in the membrane bilayer; the regions outside the membrane are hydrophilic and fold into various conformations. All type IV proteins have multiple transmembrane α helices. The type IV topology depicted here corresponds to that of G

protein-coupled receptors: seven α helices, the N-terminus on the exoplasmic side of the membrane, and the C-terminus on the cytosolic side. Other type IV proteins may have a different number of helices and various orientations of the N-terminus and C-terminus. [See E. Hartmann et al., 1989, *Proc. Nat'l. Acad. Sci. USA* **86**:5786, and C. A. Brown and S. D. Black, 1989, *J. Biol. Chem.* **264**:4442.]

establish the membrane orientation of transmembrane segments, as discussed in the next section.

The proteins forming topological class IV contain multiple membrane-spanning segments (see Figure 16-10). Many of the membrane transport proteins discussed in Chapter 7 and the numerous G protein-coupled receptors covered in Chapter 13 belong to this class, sometimes called *multipass* proteins. A final type of membrane protein lacks a hydrophobic membrane-spanning segment altogether; instead, these proteins are linked to an amphipathic phospholipid anchor that is embedded in the membrane.

Internal Stop-Transfer and Signal-Anchor Sequences Determine Topology of Single-Pass Proteins

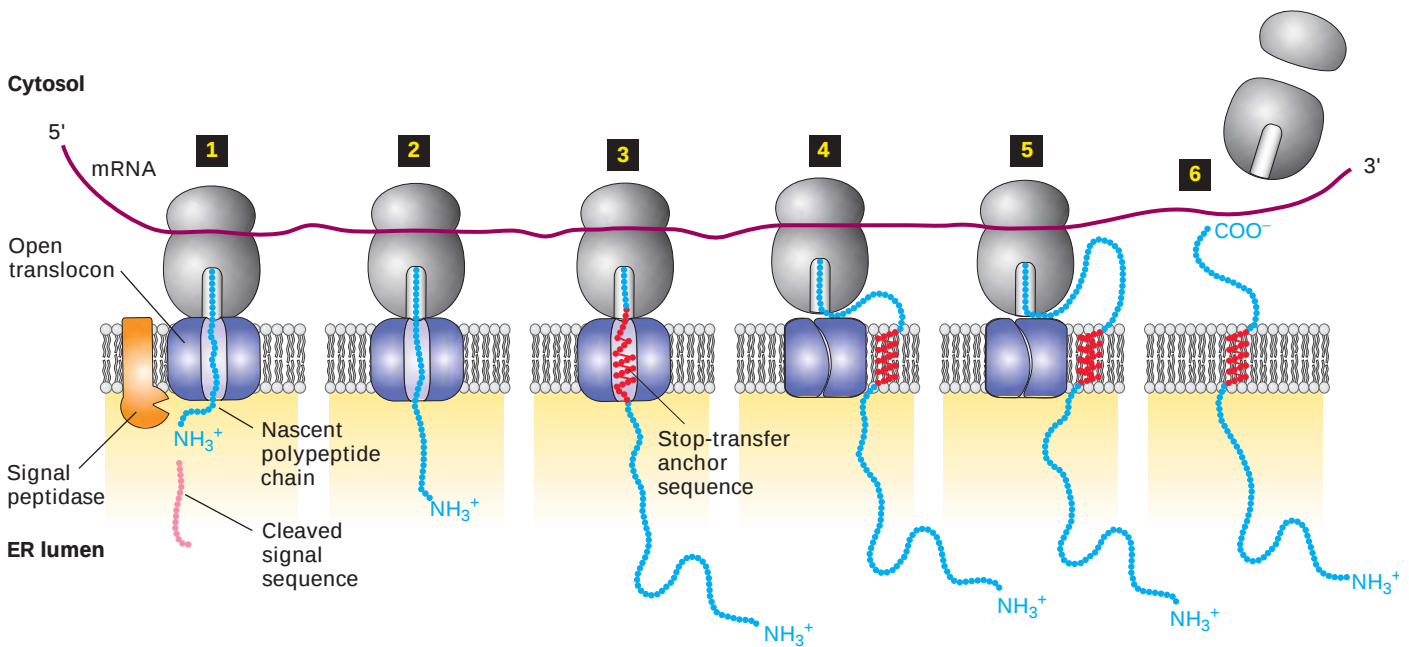
We begin our discussion with the membrane insertion of integral proteins that contain a single, hydrophobic membrane-spanning segment. Two sequences are involved in targeting and orienting type I proteins in the ER membrane, whereas type II and type III proteins contain a single, internal topogenic sequence.

Type I Proteins All type I transmembrane proteins possess an N-terminal signal sequence that targets them to the ER and an internal hydrophobic sequence that becomes the membrane-spanning α helix. The N-terminal signal sequence on a nascent type I protein, like that of a secretory protein,

initiates cotranslational translocation of the protein through the combined action of the SRP and SRP receptor. Once the N-terminus of the growing polypeptide enters the lumen of the ER, the signal sequence is cleaved, and the growing chain continues to be extruded across the ER membrane. However, unlike the case with secretory proteins, a sequence of about 22 hydrophobic amino acids in the middle of a type I protein stops transfer of the nascent chain through the translocon (Figure 16-11). This internal sequence, because of its hydrophobicity, can move laterally between the protein subunits that form the wall of the translocon and become anchored in the phospholipid bilayer of the membrane, where it remains. Because of its dual function, this sequence is called a *stop-transfer anchor sequence*.

Once translocation is interrupted, translation continues at the ribosome, which is still anchored to the now unoccupied and closed translocon. As the C-terminus of the protein chain is synthesized, it loops out on the cytosolic side of the membrane. When translation is completed, the ribosome is released from the translocon and the C-terminus of the newly synthesized type I protein remains in the cytosol.

Support for this model, depicted in Figure 16-11, has come from studies in which cDNAs encoding various mutant receptors for human growth hormone (HGH) are expressed in cultured mammalian cells. The wild-type HGH receptor, a typical type I protein, is transported normally to the plasma membrane. However, a mutant receptor that has charged residues inserted into the α -helical membrane-spanning segment, or that is missing most of this segment, is translocated



▲ **FIGURE 16-11 Synthesis and insertion into the ER membrane of type I single-pass proteins.** Step **1**: After the ribosome/nascent chain complex becomes associated with a translocon in the ER membrane, the N-terminal signal sequence is cleaved. This process occurs by the same mechanism as the one for soluble secretory proteins (see Figure 16-6). Steps **2**, **3**: The chain is elongated until the hydrophobic stop-transfer anchor sequence is synthesized and enters the translocon, where it prevents the nascent chain from extruding farther into the ER

lumen. Step **4**: The stop-transfer anchor sequence moves laterally between the translocon subunits and becomes anchored in the phospholipid bilayer. At this time, the translocon probably closes. Step **5**: As synthesis continues, the elongating chain may loop out into the cytosol through the small space between the ribosome and translocon (see Figure 16-8). Step **6**: When synthesis is complete, the ribosomal subunits are released into the cytosol, leaving the protein free to diffuse in the membrane. [See H. Do et al., 1996, *Cell* **85**:369, and W. Mothes et al., 1997, *Cell* **89**:523.]

entirely into the ER lumen and is eventually secreted from the cell. These findings establish that the hydrophobic membrane-spanning α helix of the HGH receptor and of other type I proteins functions both as a stop-transfer sequence and a membrane anchor that prevents the C-terminus of the protein from crossing the ER membrane.

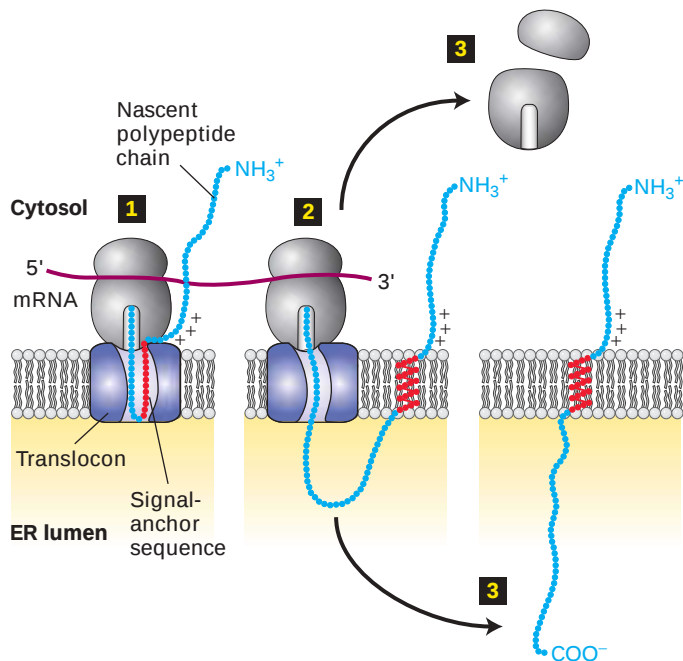
Type II and Type III Proteins Unlike type I proteins, type II and type III proteins lack a cleavable N-terminal ER signal sequence. Instead, both possess a single internal hydrophobic *signal-anchor sequence* that functions as both an ER signal sequence and membrane-anchor sequence. Recall that type II and type III proteins have opposite orientations in the membrane (see Figure 16-10); this difference depends on the orientation that their respective signal-anchor sequences assume within the translocon.

The internal signal-anchor sequence in type II proteins directs insertion of the nascent chain into the ER membrane so that the N-terminus of the chain faces the cytosol (Figure 16-12). The internal signal-anchor sequence is *not* cleaved and remains in the translocon while the C-terminal region of the growing chain is extruded into the ER lumen by co-translational translocation. During synthesis, the signal-anchor sequence moves laterally between the protein sub-

units forming the translocon wall into the phospholipid bilayer, where it functions as a membrane anchor. Thus this function is similar to the anchoring function of the stop-transfer anchor sequence in type I proteins.

In the case of type III proteins, the signal-anchor sequence, which is located near the N-terminus, inserts the nascent chain into the ER membrane with its N-terminus facing the lumen, just the opposite of type II proteins. The signal-anchor sequence of type III proteins also prevents further extrusion of the nascent chain into the ER lumen, functioning as a stop-transfer sequence. Continued elongation of the chain C-terminal to the signal-anchor/stop-transfer sequence proceeds as it does for type I proteins, with the hydrophobic sequence moving laterally between the translocon subunits to anchor the polypeptide in the ER membrane (see Figure 16-11).

One of the features of signal-anchor sequences that appears to determine their insertion orientation is a high density of positively charged amino acids adjacent to one end of the hydrophobic segment. For reasons that are not well understood these positively charged residues tend to remain on the cytosolic side of the membrane, thereby dictating the orientation of the signal-anchor sequence within the translocon. Thus type II proteins tend to have positively charged



▲ FIGURE 16-12 Synthesis and insertion into the ER membrane of type II single-pass proteins. Step **1**: After the internal signal-anchor sequence is synthesized on a cytosolic ribosome, it is bound by an SRP (not shown), which directs the ribosome/nascent chain complex to the ER membrane. This is similar to targeting of soluble secretory proteins except that the hydrophobic signal sequence is not located at the N-terminus and is not subsequently cleaved. The nascent chain becomes oriented in the translocon with its N-terminal portion toward the cytosol. This orientation is believed to be mediated by the positively charged residues shown N-terminal to the signal-anchor sequence. Step **2**: As the chain is elongated and extruded into the lumen, the internal signal-anchor moves laterally out of the translocon and anchors the chain in the phospholipid bilayer. Step **3**: Once protein synthesis is completed, the C-terminus of the polypeptide is released into the lumen, and the ribosomal subunits are released into the cytosol. [See M. Spiess and H. F. Lodish, 1986, *Cell* **44**:177, and H. Do et al., 1996, *Cell* **85**:369.]

residues on the N-terminal side of their signal-anchor sequence, whereas type III proteins tend to have positively charged residues on the C-terminal side of their signal-anchor sequence.

A striking experimental demonstration of the importance of the flanking charge in determining membrane orientation is provided by neuraminidase, a type II protein in the surface coat of influenza virus. Three arginine residues are located just N-terminal to the internal signal-anchor sequence in neuraminidase. Mutation of these three positively charged residues to negatively charged glutamate residues causes neuraminidase to acquire the reverse orientation. Similar experiments have shown that other proteins, with either type II or type III orientation, can be made to “flip” their orientation in the ER membrane by mutating charged residues that flank the internal signal-anchor segment.

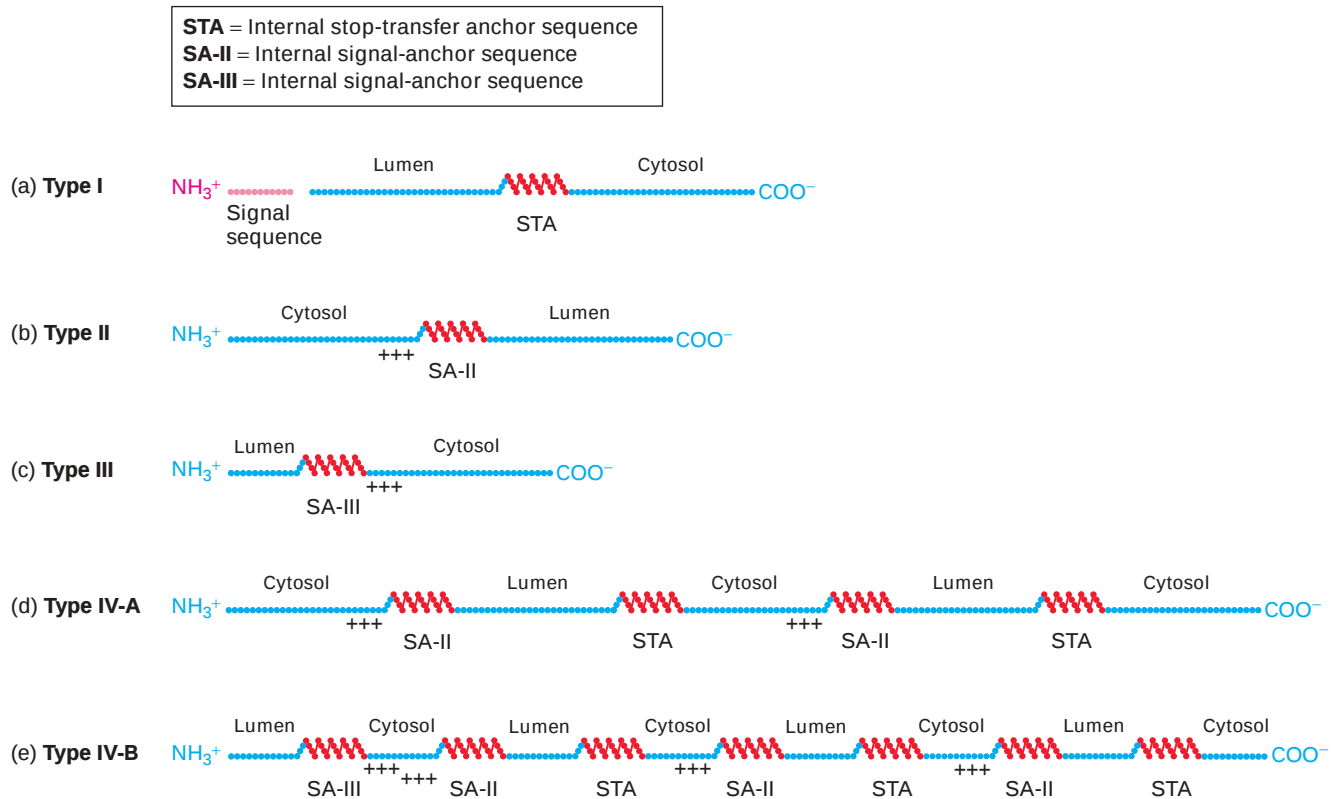
Multipass Proteins Have Multiple Internal Topogenic Sequences

Figure 16-13 summarizes the arrangements of topogenic sequences in single-pass and multipass transmembrane proteins. In multipass (type IV) proteins, each of the membrane-spanning α helices acts as a topogenic sequence in the ways that we have already discussed. Multipass proteins fall into one of two types depending on whether the N-terminus extends into the cytosol or the exoplasmic space (i.e., the ER lumen, cell exterior). This N-terminal topology usually is determined by the hydrophobic segment closest to the N-terminus and the charge of the sequences flanking it. If a type IV protein has an *even* number of transmembrane α helices, both its N-terminus and C-terminus will be oriented toward the same side of the membrane (see Figure 16-13d). Conversely, if a type IV protein has an *odd* number of α helices, its two ends will have opposite orientations (see Figure 16-13e).

Proteins with N-Terminus in Cytosol (Type IV-A) Among the multipass proteins whose N-terminus extends into the cytosol are the various glucose transporters (GLUTs) and most ion-channel proteins discussed in Chapter 7. In these proteins, the hydrophobic segment closest to the N-terminus initiates insertion of the nascent chain into the ER membrane with the N-terminus oriented toward the cytosol; thus this α -helical segment functions like the internal signal-anchor sequence of a type II protein (see Figure 16-12). As the nascent chain following the first α helix elongates, it moves through the translocon until the second hydrophobic α helix is formed. This helix prevents further extrusion of the nascent chain through the translocon; thus its function is similar to that of the stop-transfer anchor sequence in a type I protein (see Figure 16-11).

After synthesis of the first two transmembrane α helices, both ends of the nascent chain face the cytosol and the loop between them extends into the ER lumen. The C-terminus of the nascent chain then continues to grow into the cytosol, as it does in synthesis of type I and type III proteins. According to this mechanism, the third α helix acts as another type II signal-anchor sequence, and the fourth as another stop-transfer anchor sequence (see Figure 16-13d). Apparently, once the first topogenic sequence of a multipass polypeptide initiates association with the translocon, the ribosome remains attached to the translocon, and topogenic sequences that subsequently emerge from the ribosome are threaded into the translocon without the need for the SRP and the SRP receptor.

Experiments that use recombinant DNA techniques to exchange hydrophobic α helices have provided insight into the functioning of the topogenic sequences in type IV-A multipass proteins. These experiments indicate that the order of the hydrophobic α helices relative to each other in the growing chain largely determines whether a given helix functions as a signal-anchor sequence or stop-transfer anchor sequence.



▲ **FIGURE 16-13 Arrangement of topogenic sequences in single-pass and multipass membrane proteins inserted into the ER membrane.** Topogenic sequences are shown in red; soluble, hydrophilic portions, in blue. The internal topogenic sequences form transmembrane α helices that anchor the proteins or segments of proteins in the membrane. (a) Type I proteins contain a cleaved signal sequence and a single internal stop-transfer anchor (STA). (b, c) Type II and type III proteins contain a single internal signal-anchor (SA) sequence. The difference in the orientation of these proteins depends largely on

whether there is a high density of positively charged amino acids (+++) on the N-terminal side of the SA sequence (type II) or on the C-terminal side of the SA sequence (type III). (d, e) Nearly all multipass proteins lack a cleavable signal sequence, as depicted in the examples shown here. Type IV-A proteins, whose N-terminus faces the cytosol, contain alternating SA-II sequences and STA sequences. Type IV-B proteins, whose N-terminus faces the lumen, begin with a SA-III sequence followed by alternating SA-II and STA sequences. Proteins of each type with different numbers of α helices (odd or even) are known.

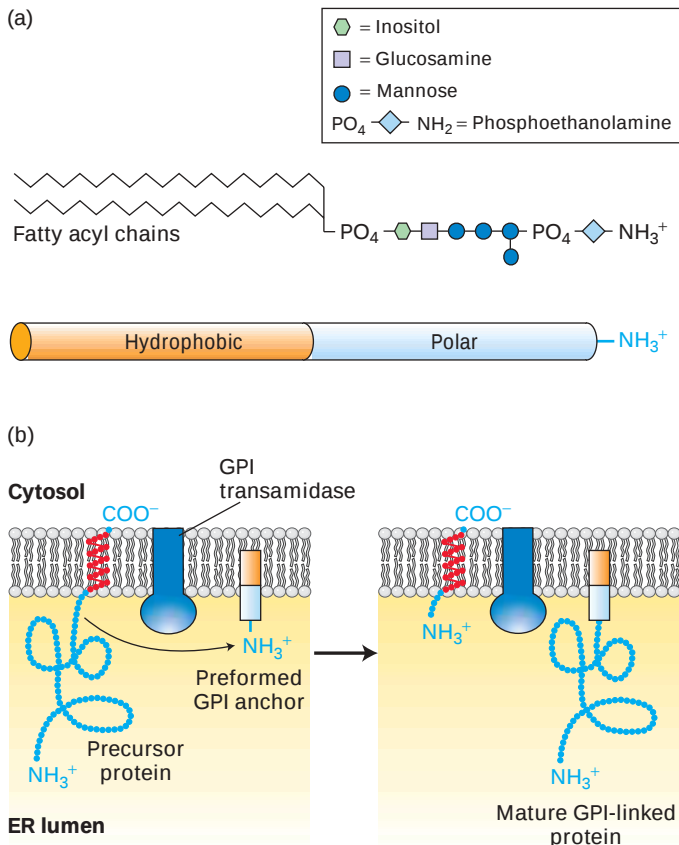
Other than its hydrophobicity, the specific amino acid sequence of a particular helix has little bearing on its function. Thus the first N-terminal α helix and the subsequent odd-numbered ones function as signal-anchor sequences, whereas the intervening even-numbered helices function as stop-transfer anchor sequences.

Proteins with N-Terminus in the Exoplasmic Space (Type IV-B) The large family of G protein-coupled receptors, all of which contain seven transmembrane α helices, constitute the most numerous type IV-B proteins, whose N-terminus extends into the exoplasmic space. In these proteins, the hydrophobic α helix closest to the N-terminus often is followed by a cluster of positively charged amino acids, similar to a type III signal-anchor sequence. As a result, the first α helix inserts the nascent chain into the translocon with the N-terminus extending into the lumen (see Figure 16-13e). As the chain is elongated, it is inserted into the ER membrane

by alternating type II signal-anchor sequences and stop-transfer sequences, as just described for type IV-A proteins.

A Phospholipid Anchor Tethers Some Cell-Surface Proteins to the Membrane

Some cell-surface proteins are anchored to the phospholipid bilayer not by a sequence of hydrophobic amino acids but by a covalently attached amphipathic molecule, *glycosylphosphatidylinositol (GPI)* (Figure 16-14a). These proteins are synthesized and initially anchored to the ER membrane exactly like type I transmembrane proteins, with a cleaved N-terminal signal sequence and internal stop-transfer anchor sequence directing the process (see Figure 16-11). However, a short sequence of amino acids in the luminal domain, adjacent to the membrane-spanning domain, is recognized by a transamidase located within the ER membrane. This enzyme simultaneously cleaves off the original



▲ **FIGURE 16-14 GPI-anchored proteins.** (a) Structure of a glycosylphosphatidylinositol (GPI) from yeast. The hydrophobic portion of the molecule is composed of fatty acyl chains, whereas the polar (hydrophilic) portion of the molecule is composed of carbohydrate residues and phosphate groups. In other organisms, both the length of the acyl chains and the carbohydrate moieties may vary somewhat from the structure shown. (b) Formation of GPI-anchored proteins in the ER membrane. The protein is synthesized and initially inserted into the ER membrane as shown in Figure 16-11. A specific transamidase simultaneously cleaves the precursor protein within the exoplasmic-facing domain, near the stop-transfer anchor sequence (red), and transfers the carboxyl group of the new C-terminus to the terminal amino group of a preformed GPI anchor. [See C. Abeijon and C. B. Hirschberg, 1992, *Trends Biochem. Sci.* **17**:32, and K. Kodukula et al., 1992, *Proc. Nat'l. Acad. Sci. USA* **89**:4982.]

stop-transfer anchor sequence and transfers the remainder of the protein to a preformed GPI anchor in the membrane (Figure 16-14b).

Why change one type of membrane anchor for another? Attachment of the GPI anchor, which results in removal of the cytosol-facing hydrophilic domain from the protein, can have several consequences. Proteins with GPI anchors, for example, can diffuse in the plane of the phospholipid bilayer membrane. In contrast, many proteins anchored by membrane-spanning α helices are immobilized in the membrane because their cytosol-facing segments interact with

the cytoskeleton. In addition, the GPI anchor targets the attached protein to the apical domain of the plasma membrane in certain polarized epithelial cells, as we discuss in Chapter 17.

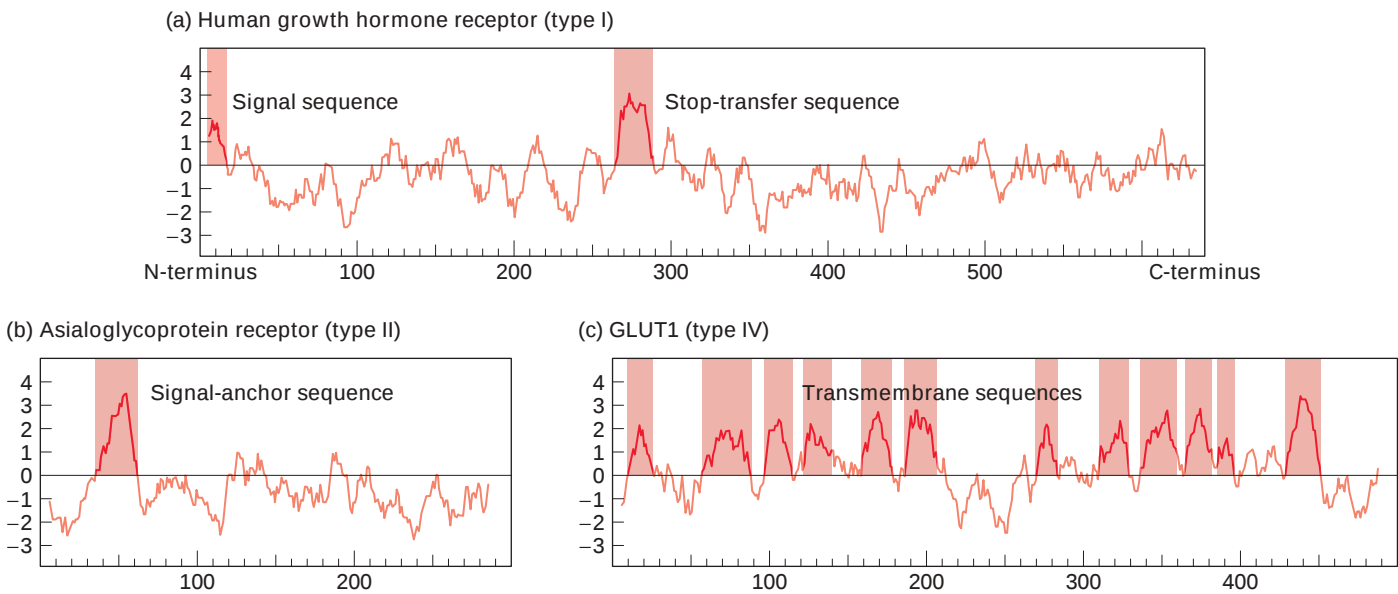
Some membrane proteins are tethered to the cytosolic face of the plasma membrane by other types of lipid anchors (see Figure 5-15). An important example is Ras protein, which plays a key role in intracellular signaling pathways discussed in Chapter 14.

The Topology of a Membrane Protein Often Can Be Deduced from Its Sequence

As we have seen, various topogenic sequences in integral membrane proteins synthesized on the ER govern interaction of the nascent chain with the translocon. When scientists begin to study a protein of unknown function, the identification of topogenic sequences within the corresponding gene sequence can provide important clues about the protein's topological class and function. Suppose, for example, that the gene for a protein known to be required for a cell-to-cell signaling pathway contains nucleotide sequences that encode an apparent N-terminal signal sequence and an internal hydrophobic sequence. These findings would suggest that the protein is a type I integral membrane protein and therefore may be a cell-surface receptor for an extracellular ligand.

Identification of topogenic sequences requires a way to scan sequence databases for segments that are sufficiently hydrophobic to be either a signal sequence or a transmembrane anchor sequence. Topogenic sequences can often be identified with the aid of computer programs that generate a *hydropathy profile* for the protein of interest. The first step is to assign a value known as the *hydropathic index* to each amino acid in the protein. By convention, hydrophobic amino acids are assigned positive values, and hydrophilic amino acids negative values. Although different scales for the hydropathic index exist, all assign the most positive values to amino acids with side chains made up of mostly hydrocarbon residues (e.g., phenylalanine and methionine) and the most negative values to charged amino acids (e.g., arginine and aspartate). The second step is to identify longer segments of sufficient overall hydrophobicity to be N-terminal signal sequences or internal stop-transfer sequences and signal-anchor sequences. To accomplish this, the total hydropathic index for each successive sliding “window” of 20 consecutive amino acids is calculated along the entire length of the protein. Plots of these calculated values against position in the amino acid sequence yield a hydropathy profile.

Figure 16-15 shows the hydropathy profiles for three different membrane proteins. The prominent peaks in such plots identify probable topogenic sequences, as well as their position and approximate length. For example, the hydropathy profile of the human growth hormone receptor reveals the presence of both a hydrophobic signal sequence at the extreme N-terminus of the protein and an internal hydrophobic stop-transfer sequence (see Figure 16-15a). On the basis



▲ **EXPERIMENTAL FIGURE 16-15** **Hydropathy profiles can identify likely topogenic sequences in integral membrane proteins.** Hydropathy profiles are generated by plotting the total hydrophobicity of each segment of 20 contiguous amino acids along the length of a protein. Positive values indicate relatively hydrophobic portions; negative values, relatively polar portions of

the protein. Probable topogenic sequences are marked. The complex profiles for multipass (type IV) proteins, such as GLUT1 in part (c), often must be supplemented with other analyses to determine the topology of these proteins. See the text for discussion.

of this profile, we can deduce, correctly, that the human growth hormone receptor is a type I integral membrane protein. The hydropathy profile of the asialoglycoprotein receptor reveals a prominent internal hydrophobic signal-anchor sequence but gives no indication of a hydrophobic N-terminal signal sequence (see Figure 16-15b). Thus we can predict that the asialoglycoprotein receptor is a type II or type III membrane protein. The distribution of charged residues on either side of the signal-anchor sequence often can distinguish between these possibilities, as positively charged amino acids flanking a membrane-spanning segment usually are oriented toward the cytosolic face of the membrane. For instance, in the case of the asialoglycoprotein receptor, a type II protein, the residues on the N-terminal side of the hydrophobic peak carry a net positive charge.

The hydropathy profile of the GLUT1 glucose transporter, a multipass membrane protein, reveals the presence of many segments that are sufficiently hydrophobic to be membrane-spanning helices (Figure 16-15c). The complexity of this profile illustrates the difficulty both in unambiguously identifying all the membrane-spanning segments in a multipass protein and in predicting the topology of individual signal-anchor and stop-transfer sequences. More sophisticated computer algorithms have been developed that take into account the presence of positively charged amino acids adjacent to hydrophobic segments, as well as the length of and spacing between segments. Using all this information, the best algorithms can predict the complex topology of multipass proteins with an accuracy greater than 75 percent.

Finally, sequence homology to a known protein may permit accurate prediction of the topology of a newly discovered multipass protein. For example, the genomes of multicellular organisms encode a very large number of multipass proteins with seven transmembrane α helices. The similarities between the sequences of these proteins strongly suggest that all have the same topology as the well-studied G protein-coupled receptors, which have the N-terminus oriented to the exoplasmic side and the C-terminus oriented to the cytosolic side of the membrane.

KEY CONCEPTS OF SECTION 16.2

Insertion of Proteins into the ER Membrane

- Integral membrane proteins synthesized on the rough ER fall into four topological classes (see Figure 16-10).
- Topogenic sequences—N-terminal signal sequences, internal stop-transfer anchor sequences, and internal signal-anchor sequences—direct the insertion and orientation of nascent proteins within the ER membrane. This orientation is retained during transport of the completed membrane protein to its final destination.
- Single-pass membrane proteins contain one or two topogenic sequences. In multipass membrane proteins, each α -helical segment can function as an internal topogenic sequence, depending on its location in the polypeptide chain

and the presence of adjacent positively charged residues (see Figure 16-13).

- Some cell-surface proteins are initially synthesized as type I proteins on the ER and then are cleaved with their luminal domain transferred to a GPI anchor (see Figure 16-14).
- The topology of membrane proteins can often be correctly predicted by computer programs that identify hydrophobic topogenic segments within the amino acid sequence and generate hydropathy profiles (see Figure 16-15).

16.3 Protein Modifications, Folding, and Quality Control in the ER

Membrane and soluble secretory proteins synthesized on the rough ER undergo four principal modifications before they reach their final destinations: (1) addition and processing of carbohydrates (*glycosylation*) in the ER and Golgi, (2) formation of disulfide bonds in the ER, (3) proper folding of polypeptide chains and assembly of multisubunit proteins in the ER, and (4) specific proteolytic cleavages in the ER, Golgi, and secretory vesicles.

One or more carbohydrate chains are added to the vast majority of proteins that are synthesized on the rough ER; indeed, glycosylation is the principal chemical modification to most of these proteins. Carbohydrate chains in glycoproteins may be attached to the hydroxyl group in serine and threonine residues or to the amide nitrogen of asparagine. These are referred to as *O-linked* and *N-linked oligosaccharides*, respectively. *O-linked* oligosaccharides, such as those found in collagen and glycophorin, often contain only one to four sugar residues. The more common *N-linked* oligosaccharides are larger and more complex, containing several branches in mammalian cells. In this section we focus on *N-linked* oligosaccharides, whose initial synthesis occurs in the ER. After the initial glycosylation of a protein in the ER, the oligosaccharide chain is modified in the ER and commonly in the Golgi, as well.

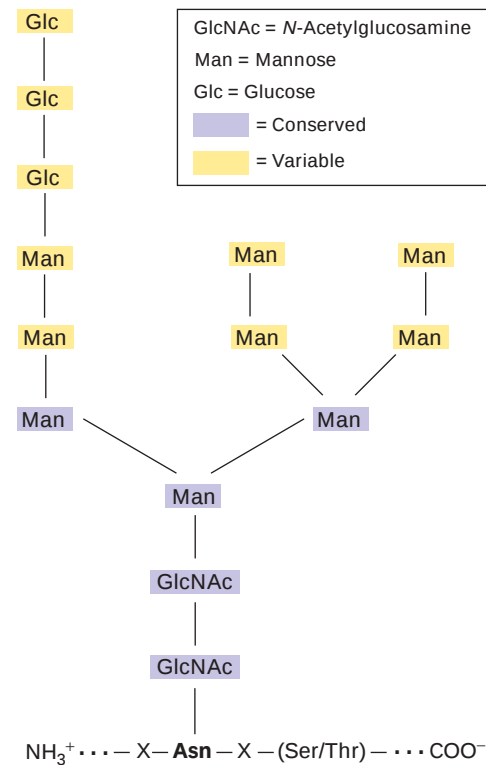
Disulfide bond formation, protein folding, and assembly of multimeric proteins, which take place exclusively in the rough ER, also are discussed in this section. Only properly folded and assembled proteins are transported from the rough ER to the Golgi complex and ultimately to the cell surface or other final destination. Unfolded, misfolded, or partly folded and assembled proteins are selectively retained in the rough ER. We consider several features of such “quality control” in the latter part of this section.

As discussed previously, N-terminal ER signal sequences are cleaved from secretory proteins and type I membrane proteins in the ER. Some proteins also undergo other specific proteolytic cleavages in the Golgi complex or forming secretory vesicles. We cover these cleavages, as well as carbohydrate modifications that occur primarily or exclusively in the Golgi complex, in the next chapter.

A Preformed *N*-Linked Oligosaccharide Is Added to Many Proteins in the Rough ER

Biosynthesis of all *N-linked* oligosaccharides begins in the rough ER with addition of a preformed oligosaccharide precursor containing 14 residues (Figure 16-16). The structure of this precursor is the same in plants, animals, and single-celled eukaryotes—a branched oligosaccharide, containing three glucose (Glc), nine mannose (Man), and two *N*-acetylglucosamine (GlcNAc) molecules, which can be written as $\text{Glc}_3\text{Man}_9(\text{GlcNAc})_2$. This branched carbohydrate structure is modified in the ER and Golgi compartments, but 5 of the 14 residues are conserved in the structures of all *N-linked* oligosaccharides on secretory and membrane proteins.

The precursor oligosaccharide is linked by a pyrophosphoryl residue to *dolichol*, a long-chain polyisoprenoid lipid that is firmly embedded in the ER membrane and acts as a carrier for the oligosaccharide. The dolichol pyrophosphoryl oligosaccharide is formed on the ER membrane in a complex set of reactions catalyzed by enzymes attached to the cytosolic or luminal faces of the rough ER membrane



▲ FIGURE 16-16 Common 14-residue precursor of *N-linked* oligosaccharides that is added to nascent proteins in the rough ER. Subsequent removal and in some cases addition of specific sugar residues occur in the ER and Golgi complex. The core region, composed of five residues highlighted in purple, is retained in all *N-linked* oligosaccharides. The precursor can be linked only to asparagine (Asn) residues that are separated by one amino acid (X) from a serine (Ser) or threonine (Thr) on the carboxyl side.

(Figure 16-17). The final dolichol pyrophosphoryl oligosaccharide is oriented so that the oligosaccharide portion faces the ER lumen.

The entire 14-residue precursor is transferred from the dolichol carrier to an asparagine residue on a nascent polypeptide as it emerges into the ER lumen (Figure 16-18, step 1). Only asparagine residues in the tripeptide sequences Asn-X-Ser and Asn-X-Thr (where X is any amino acid except proline) are substrates for *oligosaccharyl transferase*, the enzyme that catalyzes this reaction. Two of the three subunits of this enzyme are ER membrane proteins whose cytosol-facing domains bind to the ribosome, localizing a third subunit of the transferase, the catalytic subunit, near the growing polypeptide chain in the ER lumen. Not all Asn-X-Ser/Thr sequences become glycosylated; for instance, rapid folding of a segment of a protein containing an Asn-X-Ser/Thr sequence may prevent transfer of the oligosaccharide precursor to it.

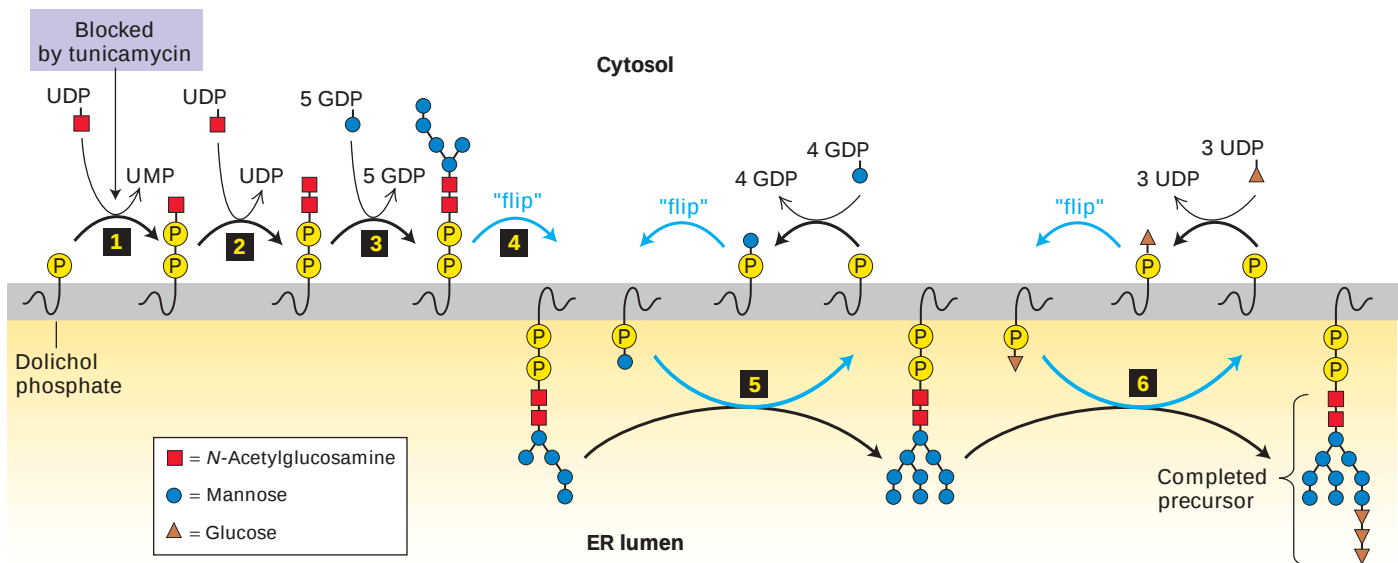
Immediately after the entire precursor, $\text{Glc}_3\text{Man}_9(\text{GlcNAc})_2$, is transferred to a nascent polypeptide, three different enzymes remove all three glucose residues and one particular mannose residue (Figure 16-18, steps 2–4). The three glucose residues, which are the last residues added during synthesis of the precursor on the dolichol carrier, appear to act as a sig-

nal that the oligosaccharide is complete and ready to be transferred to a protein.

Oligosaccharide Side Chains May Promote Folding and Stability of Glycoproteins

The oligosaccharides attached to glycoproteins serve various functions. For example, some proteins require *N*-linked oligosaccharides in order to fold properly in the ER. This function has been demonstrated in studies with the antibiotic tunicamycin, which blocks the first step in formation of the dolichol-linked precursor of *N*-linked oligosaccharides (see Figure 16-17). In the presence of tunicamycin, for instance, the hemagglutinin precursor polypeptide (HA_0) is synthesized, but it cannot fold properly and form a normal trimer; in this case, the protein remains, misfolded, in the rough ER. Moreover, mutation in the HA sequence of just one asparagine that normally is glycosylated to a glutamine residue, thereby preventing addition of an *N*-linked oligosaccharide to that site, causes the protein to accumulate in the ER in an unfolded state.

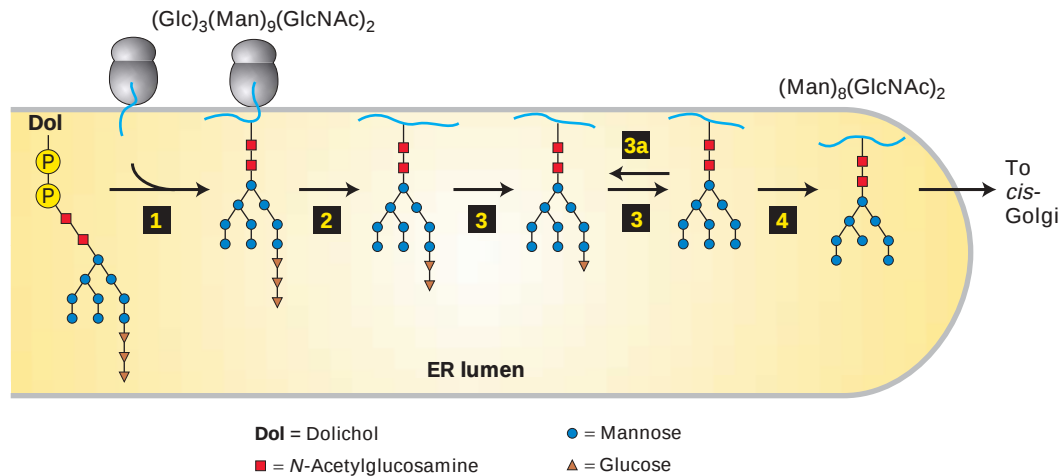
In addition to promoting proper folding, *N*-linked oligosaccharides also confer stability on many secreted gly-



▲ FIGURE 16-17 Biosynthesis of the dolichol pyrophosphoryl oligosaccharide precursor of *N*-linked oligosaccharides.

Dolichol phosphate is a strongly hydrophobic lipid, containing 75–95 carbon atoms, that is embedded in the ER membrane. Two *N*-acetylglucosamine (GlcNAc) and five mannose residues are added one at a time to a dolichol phosphate on the cytosolic face of the ER membrane (steps 1–3). The nucleotide-sugar donors in these and later reactions are synthesized in the cytosol. Note that the first sugar residue is attached to dolichol by a high-energy pyrophosphate linkage. Tunicamycin, which blocks the first enzyme in this pathway, inhibits the synthesis of

all *N*-linked oligosaccharides in cells. After the seven-residue dolichol pyrophosphoryl intermediate is flipped to the luminal face (step 4), the remaining four mannose and all three glucose residues are added one at a time (steps 5, 6). In the later reactions, the sugar to be added is first transferred from a nucleotide-sugar to a carrier dolichol phosphate on the cytosolic face of the ER; the carrier is then flipped to the luminal face, where the sugar is transferred to the growing oligosaccharide, after which the “empty” carrier is flipped back to the cytosolic face. [After C. Abeijon and C. B. Hirschberg, 1992, *Trends Biochem. Sci.* 17:32.]



▲ **FIGURE 16-18 Addition and initial processing of N-linked oligosaccharides in the rough ER of vertebrate cells.** The Glc₃Man₉(GlcNAc)₂ precursor is transferred from the dolichol carrier to a susceptible asparagine residue on a nascent protein as soon as the asparagine crosses to the luminal side of the ER (step **1**). In three separate reactions, first one glucose residue

(step **2**), then two glucose residues (step **3**), and finally one mannose residue (step **4**) are removed. Re-addition of one glucose residue (step **3a**) plays a role in the correct folding of many proteins in the ER, as discussed later. [See R. Kornfeld and S. Kornfeld, 1985, *Ann. Rev. Biochem.* **45**:631, and M. Sousa and A. J. Parodi, 1995, *EMBO J.* **14**:4196.]

coproteins. Many secretory proteins fold properly and are transported to their final destination even if the addition of all N-linked oligosaccharides is blocked, for example, by tunicamycin. However, such nonglycosylated proteins have been shown to be less stable than their glycosylated forms. For instance, glycosylated fibronectin, a normal component of the extracellular matrix, is degraded much more slowly by tissue proteases than is nonglycosylated fibronectin.

Oligosaccharides on certain cell-surface glycoproteins also play a role in cell-cell adhesion. For example, the plasma membrane of white blood cells (leukocytes) contains cell-adhesion molecules (CAMs) that are extensively glycosylated. The oligosaccharides in these molecules interact with a sugar-binding domain in certain CAMs found on endothelial cells lining blood vessels. This interaction tethers the leukocytes to the endothelium and assists in their movement into tissues during an inflammatory response to infection (see Figure 6-30). Other cell-surface glycoproteins possess oligosaccharide side chains that can induce an immune response. A common example is the A, B, O blood-group antigens, which are O-linked oligosaccharides attached to glycoproteins and glycolipids on the surface of erythrocytes and other cell types (see Figure 5-16).

Disulfide Bonds Are Formed and Rearranged by Proteins in the ER Lumen

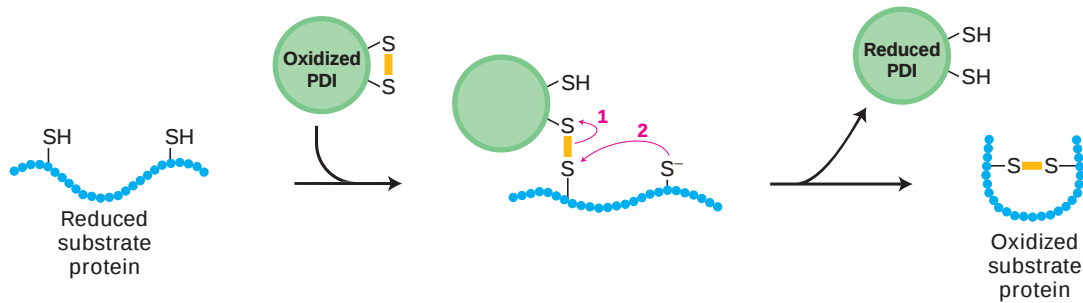
In Chapter 3 we learned that both intramolecular and intermolecular **disulfide bonds** (–S–S–) help stabilize the tertiary and quaternary structure of many proteins. These covalent

bonds form by the oxidative linkage of **sulfhydryl groups** (–SH), also known as **thiol** groups, on two cysteine residues in the same or different polypeptide chains. This reaction can proceed spontaneously only when a suitable oxidant is present. In eukaryotic cells, disulfide bonds are formed only in the lumen of the rough ER; in bacterial cells, disulfide bonds are formed in the periplasmic space between the inner and outer membranes. Thus disulfide bonds are found only in secretory proteins and in the exoplasmic domains of membrane proteins. Cytosolic proteins and organelle proteins synthesized on free ribosomes lack disulfide bonds and depend on other interactions to stabilize their structures.

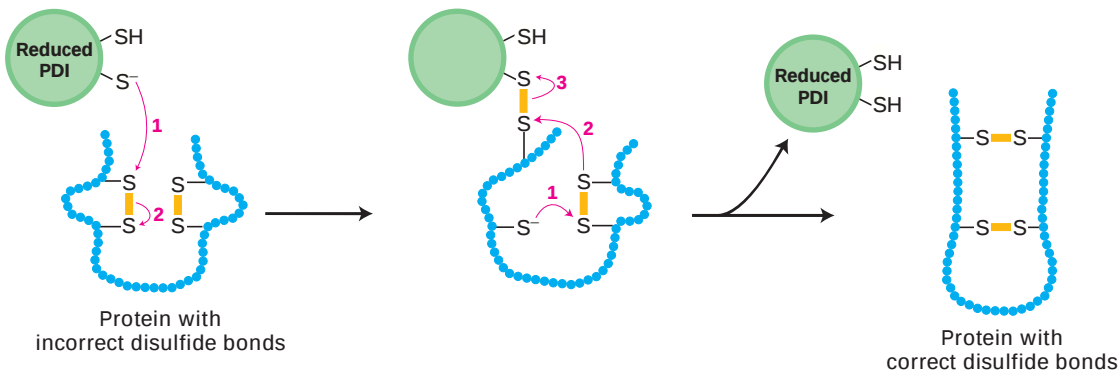
The efficient formation of disulfide bonds in the lumen of the ER depends on the enzyme *protein disulfide isomerase* (PDI), which is present in all eukaryotic cells. This enzyme is especially abundant in the ER of secretory cells in such organs as the liver and pancreas, where large quantities of proteins that contain disulfide bonds are produced. As shown in Figure 16-19a, the disulfide bond in the active site of PDI can be readily transferred to a protein by two sequential thiol-disulfide transfer reactions. The reduced PDI generated by this reaction is returned to an oxidized form by the action of an ER-resident protein, called *Ero1*, which carries a disulfide bond that can be transferred to PDI. It is not yet understood how *Ero1* itself becomes oxidized. Figure 16-20 depicts the organization of the pathway for protein disulfide-bond formation in the ER lumen and the analogous pathway in bacteria.

In proteins that contain more than one disulfide bond, the proper pairing of cysteine residues is essential for normal structure and activity. Disulfide bonds commonly are

(a) Formation of a disulfide bond



(b) Rearrangement of disulfide bonds



▲ **FIGURE 16-19 Formation and rearrangement of disulfide bonds by protein disulfide isomerase (PDI).**

PDI contains an active site with two closely spaced cysteine residues that are easily interconverted between the reduced dithiol form and the oxidized disulfide form. Numbered red arrows indicate the sequence of electron transfers. Yellow bars represent disulfide bonds. (a) In the formation of disulfide bonds, the ionized (S^-) form of a cysteine thiol in the substrate protein reacts with the disulfide ($\text{S}-\text{S}$) bond in oxidized PDI to form a disulfide-bonded

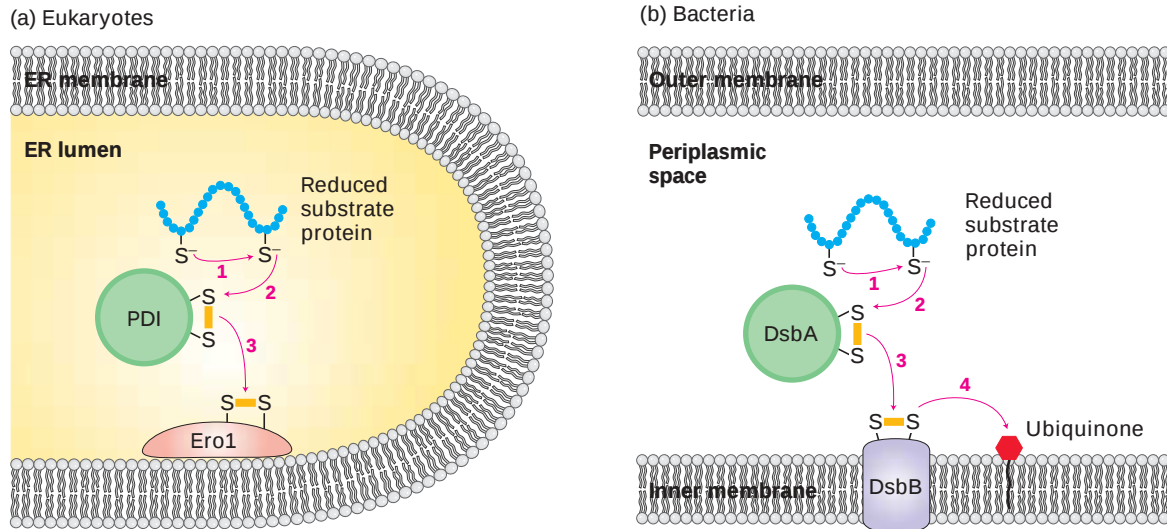
PDI-substrate protein intermediate. A second ionized thiol in the substrate then reacts with this intermediate, forming a disulfide bond within the substrate protein and releasing reduced PDI.

(b) Reduced PDI can catalyze rearrangement of improperly formed disulfide bonds by similar thiol-disulfide transfer reactions. In this case, reduced PDI both initiates and is regenerated in the reaction pathway. These reactions are repeated until the most stable conformation of the protein is achieved. [See M. M. Lyles and H. F. Gilbert, 1991, *Biochemistry* 30:619.]

formed between cysteines that occur sequentially in the amino acid sequence while a polypeptide is still growing on the ribosome. Such sequential formation, however, sometimes yields disulfide bonds between the wrong cysteines. For example, proinsulin has three disulfide bonds that link cysteines 1 and 4, 2 and 6, and 3 and 5. In this case, disulfide bonds initially formed sequentially (e.g., between cysteines 1 and 2) have to be rearranged for the protein to achieve its proper folded conformation. In cells, the rearrangement of disulfide bonds also is accelerated by PDI, which acts on a broad range of protein substrates, allowing them to reach their thermodynamically most stable conformations (Figure 16-19b). Disulfide bonds generally form in a specific order, first stabilizing small domains of a polypeptide, then stabilizing the interactions of more distant segments; this phenomenon is illustrated by the folding of influenza HA protein discussed in the next section.



Most proteins used for therapeutic purposes in humans or animals are secreted glycoproteins stabilized by disulfide bonds. When researchers first tried to synthesize such proteins using plasmid expression vectors in bacterial cells, the results were disappointing. In most cases, the proteins were not secreted (even when a bacterial signal sequence replaced the normal one); instead, they accumulated in the cytosol and often in a denatured state, in part owing to the lack of disulfide bonds. After it became clear that disulfide-bond formation occurs spontaneously only in the ER lumen, biotechnologists eventually developed expression vectors that can be used in animal cells (Chapter 9). Nowadays, such vectors and cultured animal cells are preferred for large-scale production of therapeutic proteins such as tissue plasminogen activator (an anticlotting agent) and erythropoietin, a hormone that stimulates production of red blood cells. ■



▲ **FIGURE 16-20 Pathways for the formation of disulfide bonds in eukaryotes and bacteria.** Numbered red arrows indicate the sequence of electron flow. A disulfide bond (yellow bar) is formed by loss of a pair of electrons from cysteine thiol (–SH) groups. (a) In the ER lumen of eukaryotic cells, electrons from an ionized thiol in a newly synthesized substrate protein are transferred to a disulfide bond in the active site of PDI (see Figure 16-19). PDI, in turn, transfers electrons to a disulfide bond

in the luminal protein Ero1, thereby regenerating the oxidized form of PDI. The mechanism of reoxidation of Ero1 is not known. (b) In the periplasmic space of bacterial cells, the soluble protein DsbA functions like eukaryotic PDI. Reduced DsbA is reoxidized by DsbB, an inner-membrane protein. Finally, reduced DsbB is reoxidized by transferring electrons to oxidized ubiquinone, a lipid cofactor of the electron-transport chain in the inner membrane (see Chapter 8). [See A. R. Frand et al., 2000, *Trends Cell Biol.* **10**:203.]

Chaperones and Other ER Proteins Facilitate Folding and Assembly of Proteins

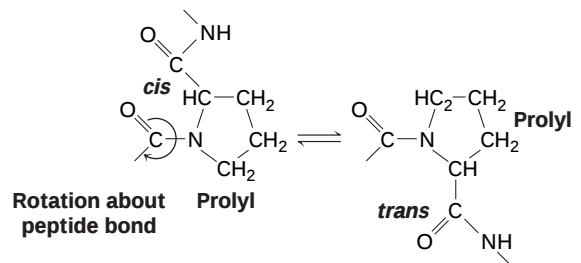
Although many reduced, denatured proteins can spontaneously refold into their native state *in vitro*, such refolding usually requires hours to reach completion. Yet new soluble and membrane proteins produced in the ER generally fold into their proper conformation within minutes after their synthesis. The rapid folding of these newly synthesized proteins in cells depends on the sequential action of several proteins present within the ER lumen.

We have already seen how the chaperone BiP can drive post-translational translocation in yeast by binding fully synthesized polypeptides as they enter the ER (see Figure 16-9). BiP can also bind transiently to nascent chains as they enter the ER during cotranslational translocation. Bound BiP is thought to prevent segments of a nascent chain from misfolding or forming aggregates, thereby promoting folding of the entire polypeptide into the proper conformation. Protein disulfide isomerase (PDI) also contributes to proper folding, which is stabilized by disulfide bonds in many proteins.

As illustrated in Figure 16-21, two other ER proteins, the homologous **lectins** (carbohydrate-binding proteins) *calnexin* and *calreticulin*, bind selectively to certain *N*-linked oligosaccharides on growing nascent chains. The ligand for these two lectins, which contains a single glucose residue, is generated by a specific glucosyltransferase in the ER lumen (see Figure 16-18, step 3a). This enzyme acts only on polypeptide chains

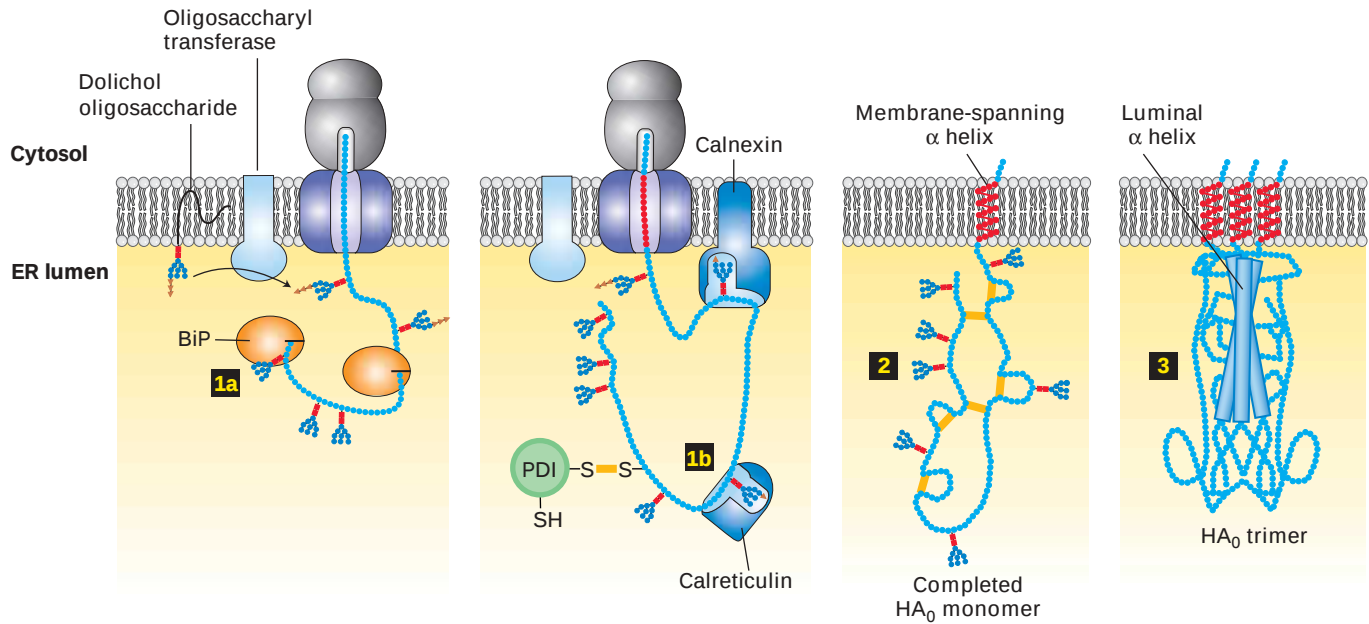
that are unfolded or misfolded. Binding of calnexin and calreticulin to unfolded nascent chains prevents aggregation of adjacent segments of a protein as it is being made on the ER. Thus calnexin and calreticulin, like BiP, help prevent premature, incorrect folding of segments of a newly made protein.

Other important protein-folding catalysts in the ER lumen are *peptidyl-prolyl isomerases*, a family of enzymes that accelerate the rotation about peptidyl-prolyl bonds in unfolded segments of a polypeptide:



Such isomerizations sometimes are the rate-limiting step in the folding of protein domains. Many peptidyl-prolyl isomerases can catalyze the rotation of exposed peptidyl-prolyl bonds indiscriminately in numerous proteins, but some have very specific protein substrates.

Many important secretory and membrane proteins synthesized on the ER are built of two or more polypeptide subunits. In all cases, the assembly of subunits constituting these



▲ **FIGURE 16-21 Folding and assembly of hemagglutinin (HA₀) trimer in the ER.** Transient binding of the chaperone BiP (step **1a**) to the nascent chain and of two lectins, calnexin and calreticulin, to certain oligosaccharide chains (step **1b**) promotes proper folding of adjacent segments. A total of seven *N*-linked oligosaccharide chains are added to the luminal portion of the nascent chain during cotranslational translocation, and PDI catalyzes the formation of six disulfide bonds per monomer. Completed HA₀ monomers are anchored in the membrane by a

single membrane-spanning α helix with their N-terminus in the lumen (step **2**). Interaction of three HA₀ chains with one another, initially via their transmembrane α helices, apparently triggers formation of a long stem containing one α helix from the luminal part of each HA₀ polypeptide. Finally, interactions between the three globular heads occur, generating a stable HA₀ trimer (step **3**). [See U. Tatu et al., 1995, *EMBO J.* **14**:1340, and D. Hebert et al., 1997, *J. Cell Biol.* **139**:613.]

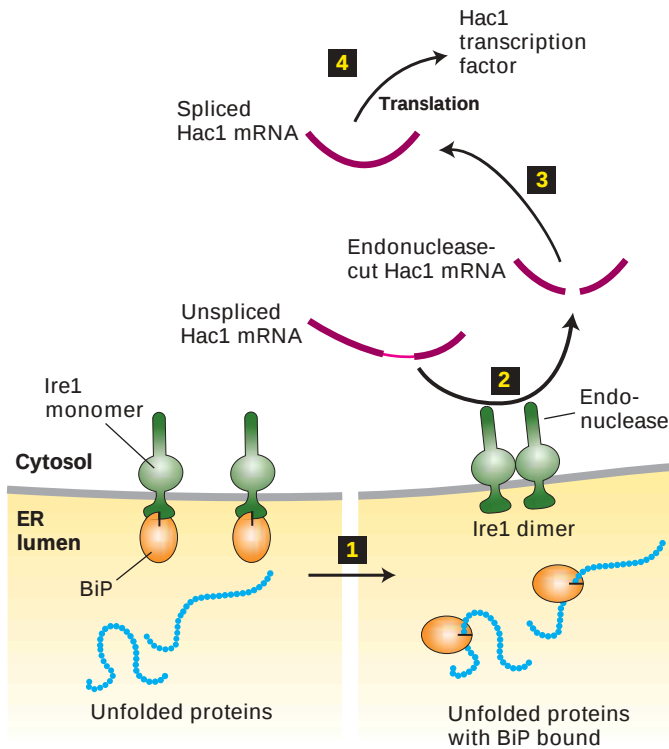
multisubunit (**multimeric**) proteins occurs in the ER. An important class of multimeric secreted proteins is the immunoglobulins, which contain two heavy (H) and two light (L) chains, all linked by intrachain disulfide bonds. Hemagglutinin (HA) is another multimeric protein that provides a good illustration of folding and subunit assembly (see Figure 16-21). This trimeric protein forms the spikes that protrude from the surface of an influenza virus particle. HA is formed within the ER of an infected host cell from three copies of a precursor protein termed HA₀, which has a single membrane-spanning α helix. In the Golgi complex, each of the three HA₀ proteins is cleaved to form two polypeptides, HA₁ and HA₂; thus each HA molecule that eventually resides on the viral surface contains three copies of HA₁ and three of HA₂ (see Figure 3-7). The trimer is stabilized by interactions between the large exoplasmic domains of the constituent polypeptides, which extend into the ER lumen; after HA is transported to the cell surface, these domains extend into the extracellular space. Interactions between the smaller cytosolic and membrane-spanning portions of the HA subunits also help stabilize the trimeric protein.

The time course of HA₀ folding and assembly *in vivo* can be determined by pulse-labeling experiments. In a typical experiment, virus-infected cells are pulse-labeled with a radioactive amino acid; at various times during the subsequent

chase period, membranes are solubilized by detergent and exposed to monoclonal antibodies specific for either HA₀ monomer or trimer. Immediately after the pulse, the monomer-specific antibody is able to immunoprecipitate all the radioactive HA₀ protein. During the chase period, increasing proportions of the total radioactive HA₀ protein instead react with the trimer-specific monoclonal antibody. Such experiments have shown that each newly made HA₀ polypeptide requires approximately 10 minutes to fold and be incorporated into a trimer in living cells.

Improperly Folded Proteins in the ER Induce Expression of Protein-Folding Catalysts

Wild-type proteins that are synthesized on the rough ER cannot exit this compartment until they achieve their completely folded conformation. Likewise, almost any mutation that prevents proper folding of a protein in the ER also blocks movement of the polypeptide from the ER lumen or membrane to the Golgi complex. The mechanisms for retaining unfolded or incompletely folded proteins within the ER probably increase the overall efficiency of folding by keeping intermediate forms in proximity to folding catalysts, which are most abundant in the ER. Improperly folded proteins retained within the ER generally are found permanently bound



to the ER chaperones BiP and calnexin. Thus these luminal folding catalysts perform two related functions: assisting in the folding of normal proteins by preventing their aggregation and binding to irreversibly misfolded proteins.

Both mammalian cells and yeasts respond to the presence of unfolded proteins in the rough ER by increasing transcription of several genes encoding ER chaperones and other folding catalysts. A key participant in this *unfolded-protein response* is Ire1, an ER membrane protein that exists as both a monomer and a dimer. The dimeric form, but not the monomeric form, promotes formation of Hac1, a transcription factor in yeast that activates expression of the genes induced in the unfolded-protein response. As depicted in Figure 16-22, binding of BiP to the luminal domain of monomeric Ire1 prevents formation of the Ire1 dimer. Thus the quantity of free BiP in the ER lumen probably determines the relative proportion of monomeric and dimeric Ire1. Accumulation of unfolded proteins within the ER lumen sequesters BiP molecules, making them unavailable for binding to Ire1. As a result the level of dimeric Ire1 increases, leading to an increase in the level of Hac1 and production of proteins that assist in protein folding.

Mammalian cells contain an additional regulatory pathway that operates in response to unfolded proteins in the ER. In this pathway, accumulation of unfolded proteins in the ER triggers proteolysis of ATF6, a transmembrane protein in the ER membrane. The cytosolic domain of ATF6 released by proteolysis then moves to the nucleus, where it stimulates transcription of the genes encoding ER chaperones. Activa-

◀ **FIGURE 16-22 The unfolded-protein response.** Ire1, a transmembrane protein in the ER membrane, has a binding site for BiP on its luminal domain; the cytosolic domain contains a specific RNA endonuclease. Step **1**: Accumulating unfolded proteins in the ER lumen bind BiP molecules, releasing them from monomeric Ire1. Dimerization of Ire1 then activates its endonuclease activity. Steps **2**, **3**: The unspliced mRNA precursor encoding the transcription factor Hac1 is cleaved by dimeric Ire1, and the two exons are joined to form functional Hac1 mRNA. Current evidence indicates that this processing occurs in the cytosol, although pre-mRNA processing generally occurs in the nucleus. Step **4**: Hac1 is translated into Hac1 protein, which then moves back into the nucleus and activates transcription of genes encoding several protein-folding catalysts. [See U. Ruegsegger et al., 2001, *Cell* **107**:103; A. Bertolotti et al., 2000, *Nature Cell Biol.* **2**:326; and C. Sidrauski and P. Walter, 1997, *Cell* **90**:1031]

tion of a transcription factor by such *regulated intramembrane proteolysis* also occurs in the Notch signaling pathway (see Figure 14-29). We will encounter another example of this phenomenon, the cholesterol-responsive transcription factor SREBP, in Chapter 18.



The hereditary form of emphysema illustrates the detrimental effects that can result from misfolding of proteins in the ER. This disease is caused by a point mutation in α_1 -antitrypsin, which normally is secreted by hepatocytes and macrophages. The wild-type protein binds to and inhibits trypsin and also the blood protease elastase. In the absence of α_1 -antitrypsin, elastase degrades the fine tissue in the lung that participates in the absorption of oxygen, eventually producing the symptoms of emphysema. Although the mutant α_1 -antitrypsin is synthesized in the rough ER, it does not fold properly, forming an almost crystalline aggregate that is not exported from the ER. In hepatocytes, the secretion of other proteins also becomes impaired, as the rough ER is filled with aggregated α_1 -antitrypsin. ■

Unassembled or Misfolded Proteins in the ER Are Often Transported to the Cytosol for Degradation

Misfolded secretory and membrane proteins, as well as the unassembled subunits of multimeric proteins, often are degraded within an hour or two after their synthesis in the rough ER. For many years, researchers thought that proteolytic enzymes within the ER catalyzed degradation of misfolded or unassembled polypeptides, but such proteases were never found. More recent studies have shown that misfolded membrane and secretory proteins are transported from the ER lumen, “backwards” through the translocon, into the cytosol, where they are degraded by the ubiquitin-mediated proteolytic pathway (see Figure 3-13).

Ubiquitinating enzymes localized to the cytosolic face of the ER add ubiquitin to misfolded ER proteins as they exit the ER. This reaction, which is coupled to hydrolysis of ATP,

may provide some of the energy required to drag these proteins back to the cytosol. The resulting polyubiquitinated polypeptides are quickly degraded in proteasomes. Exactly how misfolded soluble and membrane proteins in the ER are recognized and targeted to the translocon for export to the cytosol is not known.

KEY CONCEPTS OF SECTION 16.3

Protein Modifications, Folding, and Quality Control in the ER

- All *N*-linked oligosaccharides, which are bound to asparagine residues, contain a core of three mannose and two *N*-acetylglucosamine residues and usually have several branches (see Figure 16-16).
- *O*-linked oligosaccharides, which are bound to serine or threonine residues, are generally short, often containing only one to four sugar residues.
- Formation of all *N*-linked oligosaccharides begins with assembly of a ubiquitous 14-residue high-mannose precursor on dolichol, a lipid in the membrane of the rough ER (see Figure 16-17). After this preformed oligosaccharide is transferred to specific asparagine residues of nascent polypeptide chains in the ER lumen, three glucose residues and one mannose residue are removed (see Figure 16-18).
- Oligosaccharide side chains may assist in the proper folding of glycoproteins, help protect the mature proteins from proteolysis, participate in cell-cell adhesion, and function as antigens.
- Disulfide bonds are added to many secretory proteins and the exoplasmic domain of membrane proteins in the ER. Protein disulfide isomerase (PDI), present in the ER lumen, catalyzes both the formation and the rearrangement of disulfide bonds (see Figure 16-19).
- The chaperone BiP, the lectins calnexin and calreticulin, and peptidyl-prolyl isomerases work together to assure proper folding of newly made secretory and membrane proteins in the ER. The subunits of multimeric proteins also assemble in the ER.
- Only properly folded proteins and assembled subunits are transported from the rough ER to the Golgi complex in vesicles.
- The accumulation of abnormally folded proteins and unassembled subunits in the ER can induce increased expression of ER protein-folding catalysts via the unfolded-protein response (see Figure 16-22).
- Unassembled or misfolded proteins in the ER often are transported back through the translocon to the cytosol, where they are degraded in the ubiquitin/proteasome pathway.

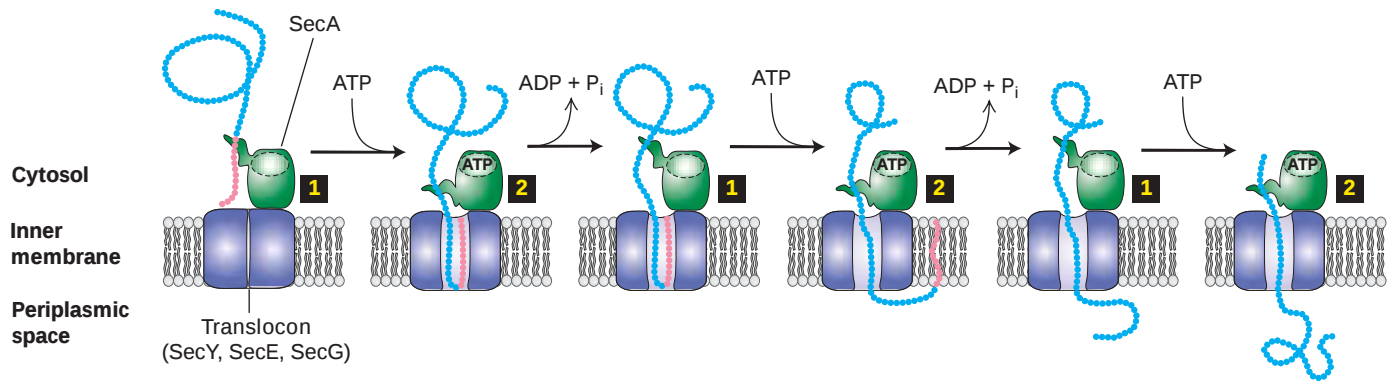
16.4 Export of Bacterial Proteins

The cell wall surrounding gram-negative bacteria comprises an inner membrane, which is the main permeability barrier for the cytoplasm, a periplasmic space containing various proteins and a layer of peptidoglycan, and an outer membrane, which is permeable to small molecules but not proteins. The peptidoglycan layer gives the cell wall its strength, while the periplasmic proteins function in sensing and importing extracellular molecules and in assembling and maintaining the structural integrity of the cell wall. These proteins (like all bacterial proteins) are synthesized on cytosolic ribosomes and then are translocated in the unfolded state across the inner membrane (also called the cytoplasmic membrane). Most proteins translocated across the inner membrane remain associated with the bacterial cell, either as a membrane protein inserted into the outer or inner membrane or trapped within the periplasmic space. Some bacterial species also possess specialized translocation systems that enable proteins to move through both membranes of the cell wall into the extracellular space. In this section, we describe both types of protein secretion in gram-negative bacteria.

Cytosolic SecA ATPase Pushes Bacterial Polypeptides Through Translocons into the Periplasmic Space

The mechanism for translocating bacterial proteins across the inner membrane shares several key features with the translocation of proteins into the ER of eukaryotic cells. First, translocated proteins usually contain an N-terminal hydrophobic signal sequence, which is cleaved by a signal peptidase. Second, bacterial proteins pass through the inner membrane in a channel, or translocon, composed of proteins that are structurally similar to the eukaryotic Sec61 complex. Third, bacterial cells express two proteins, Ffh and its receptor (FtsY), that are homologs of the SRP and SRP receptor, respectively. In bacteria, however, these latter proteins appear to function mainly in the insertion of hydrophobic membrane proteins into the inner membrane. Indeed, all bacterial proteins that are translocated across the inner membrane do so only after their synthesis in the cytosol is completed but before they are folded into their final conformation.

The post-translational translocation of bacterial proteins across the inner membrane cannot involve a ratchet mechanism similar to that mediated by BiP in the ER lumen (see Figure 16-9) because the ATP required for such a mechanism would be lost by diffusion through the outer membrane. Rather, the driving force for translocation of bacterial proteins is generated by SecA, which binds to the cytosolic side of the translocon and hydrolyzes cytosolic ATP. In the model depicted in Figure 16-23, SecA binds to the unfolded translocating polypeptide, and then a conformational change in SecA, driven by the energy released



▲ FIGURE 16-23 Post-translational translocation across inner membrane in gram-negative bacteria. The bacterial inner membrane contains a translocon channel composed of three subunits that are homologous to the components of the eukaryotic Sec61 complex. Translocation of polypeptides from the cytosol to the periplasmic space is powered by SecA, a cytosolic ATPase that binds to the translocon and to the translocating polypeptide. In the model shown here, binding and hydrolysis of ATP cause conformational changes in SecA that push the bound

polypeptide segment through the channel (steps **1**, **2**).

Repetition of this cycle results in movement of the polypeptide through the channel in one direction. Current evidence indicates that the N-terminal signal sequence moves from the channel into the bilayer but at some point is cleaved by a signal peptidase, so that the mature polypeptide enters the periplasmic space. [See A. Economou and W. Wickner, 1994, *Cell* **78**:835, and J. Eichler and W. Wickner, 1998, *J. Bacteriol.* **180**:5776.]

from ATP hydrolysis, acts to push the bound polypeptide segment through the translocon pore toward the periplasmic side of the membrane. Repetition of this cycle eventually pushes the entire polypeptide chain through the translocon into the periplasmic space, where disulfide bonds are formed and the polypeptide folds into its proper conformation.

Several Mechanisms Translocate Bacterial Proteins into the Extracellular Space

Quite different mechanisms from the one shown in Figure 16-23 are used to translocate bacterial proteins from the cytosol across both the inner and outer bacterial membranes to the extracellular space. These secretion mechanisms are particularly important for pathogenic bacteria, which commonly use secreted extracellular proteins to colonize specific tissues within the host and to evade host defense mechanisms. Well-known examples of extracellular proteins that promote the growth and dissemination of pathogenic bacteria include protein toxins (e.g., cholera toxin and tetanus toxin) and pili, which are proteinaceous fibers that project from the outer membrane and assist enteric bacteria in adhering to the epithelium of the gut.

The numerous specialized bacterial secretion systems that have been identified can be classified into four general types based on their mechanism of operation. Both the type I and the type II secretion systems involve two steps. First, substrate proteins are translocated across the inner membrane into the periplasmic space, where they fold and often acquire disulfide bonds. Second, the folded proteins are translocated from the periplasmic space across the outer

membrane by complexes of periplasmic proteins that span the inner and outer membranes. The energy for this translocation comes from hydrolysis of ATP in the cytosol, but the mechanisms that couple ATP hydrolysis and translocation across the outer membrane are not well understood.

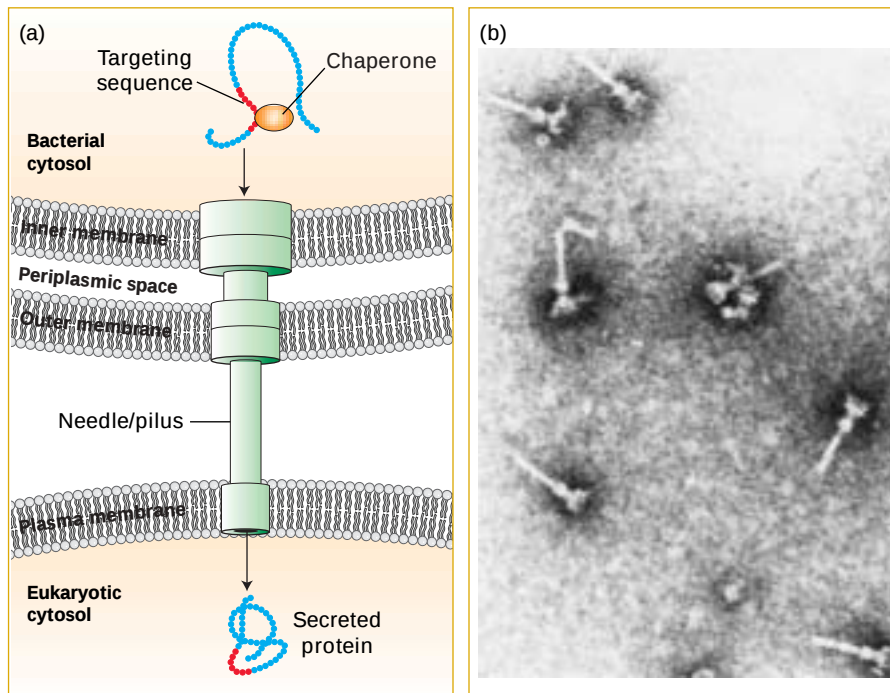
Translocation by the type III and type IV secretion systems, on the other hand, entails a single step. These systems consist of large protein complexes that span both membranes, allowing proteins to be translocated directly from the cytosol to the extracellular environment. The type III system is adapted not only for secreting proteins but also for injecting them into target cells, a very useful property for pathogenic bacteria.

Pathogenic Bacteria Can Inject Proteins into Animal Cells via Type III Secretion Apparatus



Yersinia pestis is the bacterial species responsible for the bubonic plague, one of the deadliest diseases in human history. One reason *Yersinia* is such a virulent pathogen lies in its ability to disable host macrophage cells that might otherwise engulf and destroy the invading bacterial cells. The incapacitating effect of *Yersinia* is mediated primarily by a small set of proteins that the bacterial cells inject into macrophage cells.

Various pathogenic bacteria inject proteins into host cells via a complicated syringe-like machine composed of more than 20 different proteins. This type III secretion apparatus, shown in Figure 16-24, has ringlike components embedded in both the inner and outer membranes of the bacterial cell wall and a hollow needlelike structure (pilus) that projects



◀ **FIGURE 16-24 Type III secretion apparatus for injecting bacterial proteins into eukaryotic cells.**

(a) Schematic diagram of the type III secretion apparatus, which is similar in size and morphology to the bacterial flagellum. Bacterial proteins with targeting sequences (red) that allow interaction with specialized chaperones (orange) enter the cytosol-facing portion of the type III secretion apparatus, travel down the hollow core of the pilus in an ATP-dependent process, and ultimately are delivered to the cytoplasm of the target eukaryotic cell. (b) Electron micrograph of isolated type III secretion apparatuses. Long needlelike pili can be seen extending from the widened basal portions, which are embedded in the outer and inner membranes. See the text for discussion.

[Part (a) adapted from D. G. Thanassi and S. J. Hultgren, 2000, *Curr. Opin. Cell Biol.* **12**:420. Part (b) from T. Kubori et al., 1998, *Science* **280**:602.]

out from the outer membrane. Proteins at the outer (distal) end of the pilus can penetrate the plasma membrane of certain mammalian cells, thus creating a conduit that spans three membranes, providing a connection between the bacterial cytoplasm and that of the target host cell.

A key insight into how the type III secretion apparatus may operate came from the observation that many of the components of the secretion apparatus are homologous to proteins in the base of the bacterial flagellum. The flagellar base is located in the inner membrane and functions as a motor to drive rotation of the attached flagellum, which extends outward from the cell surface. The flagellum, which is a hollow helical tube made up entirely of a repeating polymer of the protein flagellin, grows by addition of new flagellin subunits at the distal tip. Several proteins in the flagellar base are thought to use the energy from ATP hydrolysis to push new flagellin subunits through the central channel of the flagellum toward the distal end. Quite possibly, the type III secretion apparatus uses a similar ATP-driven mechanism to push proteins through the central channel of the pilus into target cells.

Recent experiments have identified the signal sequences that target bacterial proteins for transport through the type III secretion apparatus. For instance, recombinant DNA methods were used to express in *Yersinia* cells chimeric proteins containing adenylate cyclase attached to different portions of YopE, which normally is secreted via the type III apparatus. An amphipathic sequence at the N-terminus of YopE was capable of directing adenylate cyclase, which normally resides in the cytosol, to the type III secretion apparatus for injection into mammalian cells. Other experiments showed that YopE proteins in which this amphipathic tar-

geting sequence is mutated still are secreted normally. This surprising finding eventually led to discovery of a second, independent targeting signal that allows YopE to bind to a small chaperone molecule. The chaperone-YopE complex can be successfully secreted even in the absence of an N-terminal amphipathic sequence. Each of the proteins secreted by the type III apparatus is thought to interact with one of a set of small chaperone proteins (see Figure 16-24). These chaperones may act to keep secreted proteins in a partially unfolded state as they pass through the central channel of the type III apparatus. Once the transported proteins have been released into the target cell, folding can be completed by cytosolic target-cell chaperones. ■

KEY CONCEPTS OF SECTION 16.4

Export of Bacterial Proteins

- Gram-negative bacteria translocate completed proteins across the inner membrane through a translocon related to the ER translocon of eukaryotic cells.
- The driving force for post-translational translocation across the inner membrane of bacteria comes from SecA protein, which uses energy derived from hydrolysis of cytosolic ATP to push polypeptides through the translocon channel (see Figure 16-23).
- Four bacterial secretion systems exist for translocating proteins from the cytosol across both the inner and outer membranes. Hydrolysis of cytosolic ATP drives secretion in all the systems.

■ The type III secretion apparatus, used by pathogenic bacteria to inject proteins into eukaryotic cells, consists of a basal portion that spans both membranes and an extracellular needlelike structure that can penetrate the plasma membrane of a target cell (see Figure 16-24).

16.5 Sorting of Proteins to Mitochondria and Chloroplasts

In the remainder of this chapter, we examine how proteins synthesized on cytosolic ribosomes are sorted to mitochondria, chloroplasts, and peroxisomes (see Figure 16-1). Both mitochondria and chloroplasts are surrounded by a double membrane and have internal subcompartments, whereas peroxisomes are bounded by a single membrane and have a single luminal compartment known as the *matrix*. Because of these and other differences, we consider peroxisomes separately in the last section.

Besides being bounded by two membranes, both mitochondria and chloroplasts also contain similar types of electron-transport proteins and use an F-class ATPase to synthesize ATP (see Figure 8-2). Remarkably, gram-negative bacteria also exhibit these characteristics. Also like bacterial cells, mitochondria and chloroplasts contain their own DNA, which encodes organelle rRNAs, tRNAs, and some proteins (Chapter 10). Moreover, growth and division of mitochondria and chloroplasts are not coupled to nuclear division. Rather, these organelles grow by the incorporation of cellular proteins and lipids, and new organelles form by division of preexisting organelles—both processes occurring

continuously during the interphase period of the cell cycle. The numerous similarities of free-living bacterial cells with mitochondria and chloroplasts have led scientists to hypothesize that these organelles arose by the incorporation of bacteria into ancestral eukaryotic cells, forming endosymbiotic organelles. Striking evidence for this ancient evolutionary relationship can be found in the many proteins of similar sequences shared by mitochondria, chloroplasts, and bacteria, including some of the proteins involved in membrane translocation described in this section.

Proteins encoded by mitochondrial DNA or chloroplast DNA are synthesized on ribosomes within the organelles and directed to the correct subcompartment immediately after synthesis. The majority of proteins located in mitochondria and chloroplasts, however, are encoded by genes in the nucleus and are imported into the organelles after their synthesis in the cytosol. Apparently over eons of evolution much of the genetic information from the ancestral bacterial DNA in these endosymbiotic organelles moved, by an unknown mechanism, to the nucleus. Precursor proteins synthesized in the cytosol that are destined for the matrix of mitochondria or the equivalent space, the stroma, of chloroplasts usually contain specific N-terminal uptake-targeting sequences that specify binding to receptor proteins on the organelle surface. Generally, this sequence is cleaved once it reaches the matrix or stroma. Clearly, these uptake-targeting sequences are similar in their location and general function to the signal sequences that direct nascent proteins to the ER lumen. Although the three types of signals share some common sequence features, their specific sequences differ considerably, as summarized in Table 16-1.

TABLE 16-1 Uptake-Targeting Sequences That Direct Proteins from the Cytosol to Organelles*

Target Organelle	Location of Sequence Within Protein	Removal of Sequence	Nature of Sequence
Endoplasmic reticulum (lumen)	N-terminus	Yes	Core of 6–12 hydrophobic amino acids, often preceded by one or more basic amino acids (Arg, Lys)
Mitochondrion (matrix)	N-terminus	Yes	Amphipathic helix, 20–50 residues in length, with Arg and Lys residues on one side and hydrophobic residues on the other
Chloroplast (stroma)	N-terminus	Yes	No common motifs; generally rich in Ser, Thr, and small hydrophobic residues and poor in Glu and Asp
Peroxisome (matrix)	C-terminus (most proteins); N-terminus (few proteins)	No	PTS1 signal (Ser-Lys-Leu) at extreme C-terminus; PTS2 signal at N-terminus

*Different or additional sequences target proteins to organelle membranes and subcompartments. See Chapter 12 for targeting sequences required for uptake of proteins into the nucleus.

In both mitochondria and chloroplasts, protein import requires energy and occurs at points where the outer and inner organelle membranes are in close contact. Because mitochondria and chloroplasts contain multiple membranes and membrane-limited spaces, sorting of many proteins to their correct location often requires the sequential action of two targeting sequences and two membrane-bound translocation systems: one to direct the protein into the organelle, and the other to direct it into the correct organellar compartment or membrane. As we will see, the mechanisms for sorting various proteins to mitochondria and chloroplasts are related to some of the mechanisms discussed previously.

Amphipathic N-Terminal Signal Sequences Direct Proteins to the Mitochondrial Matrix

All proteins that travel from the cytosol to the same mitochondrial destination have targeting signals that share common motifs, although the signal sequences are generally not identical. Thus the receptors that recognize such signals are able to bind to a number of different but related sequences. The most extensively studied sequences for localizing proteins to mitochondria are the *matrix-targeting* sequences. These sequences, located at the N-terminus, are usually 20–50 amino acids in length. They are rich in hydrophobic amino acids, positively charged basic amino acids (arginine and lysine), and hydroxylated ones (serine and threonine), but tend to lack negatively charged acidic residues (aspartate and glutamate).

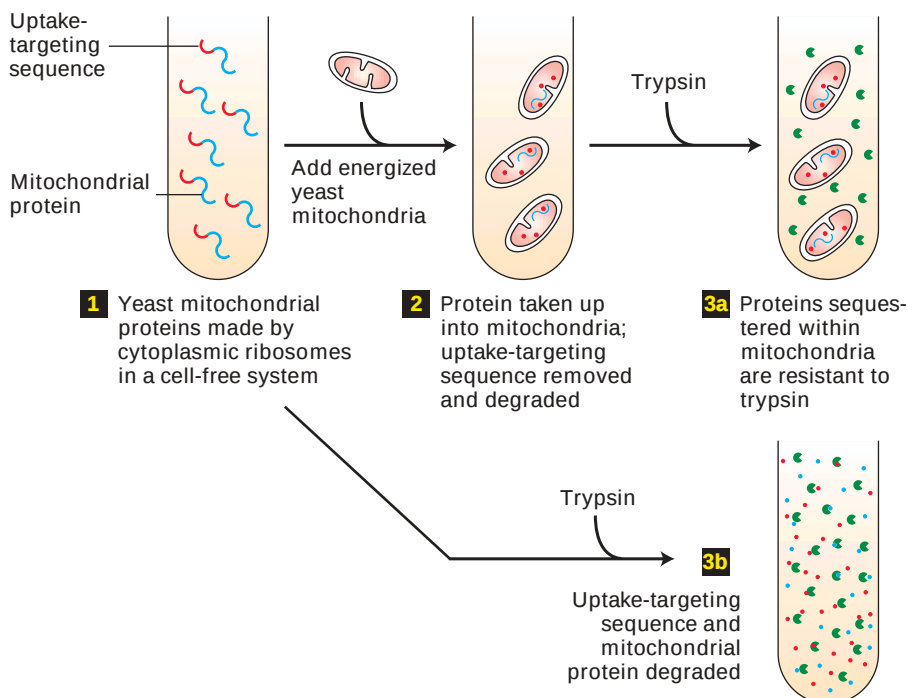
Mitochondrial matrix-targeting sequences are thought to assume an α -helical conformation in which positively

charged amino acids predominate on one side of the helix and hydrophobic amino acids predominate on the other side; thus these sequences are amphipathic. Mutations that disrupt this amphipathic character usually disrupt targeting to the matrix, although many other amino acid substitutions do not. These findings indicate that the amphipathicity of matrix-targeting sequences is critical to their function.

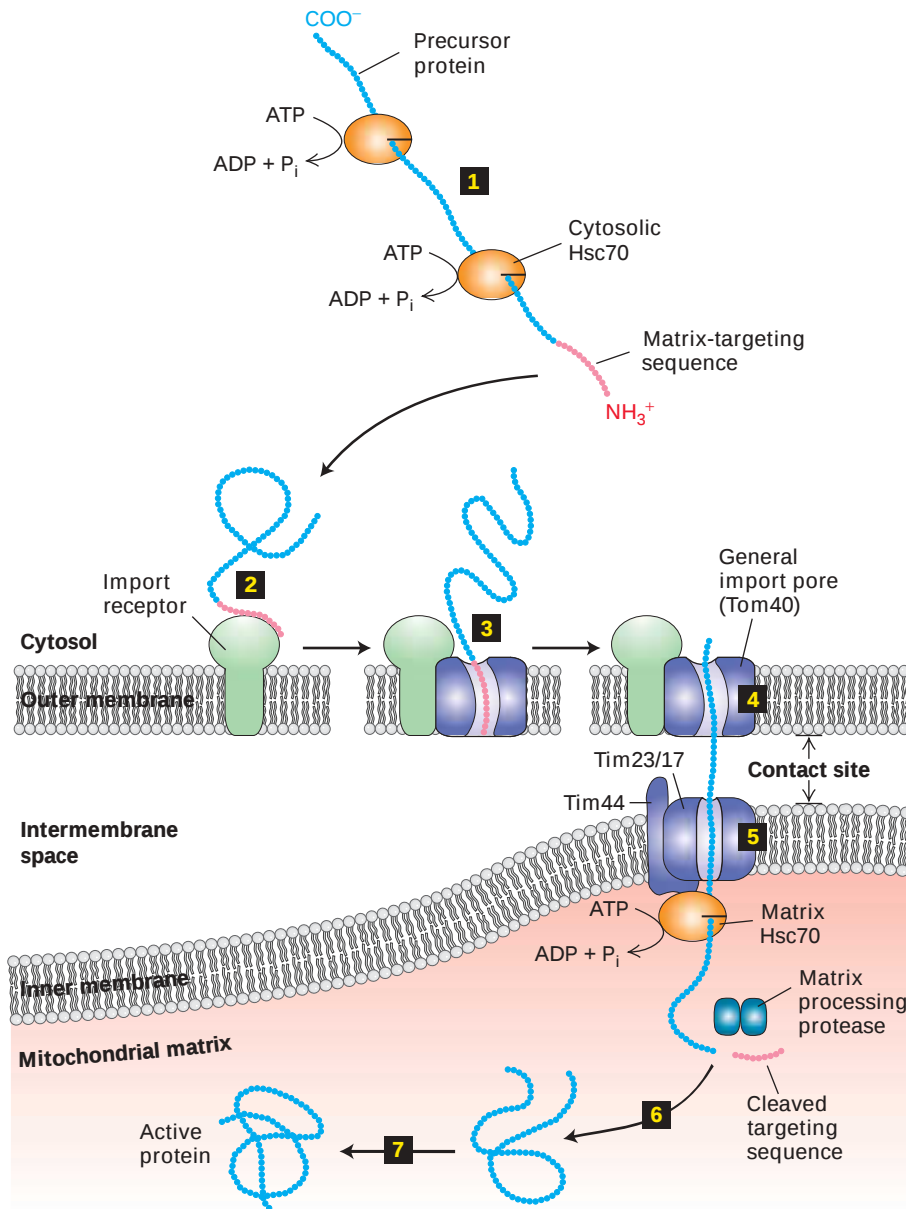
The cell-free assay outlined in Figure 16-25 has been widely used in studies on the import of mitochondrial precursor proteins. In this system, respiring (energized) mitochondria extracted from cells can incorporate mitochondrial precursor proteins carrying appropriate uptake-targeting sequences that have been separately synthesized in the absence of mitochondria. Successful incorporation of the precursor into the organelle can be assayed either by resistance to digestion by an exogenously added protease or, in most cases, by cleavage of the N-terminal targeting sequences by specific proteases. The uptake of completely synthesized mitochondrial precursor proteins by the organelle in this system contrasts with the cell-free cotranslational translocation of secretory proteins, which generally occurs only when microsomal membranes are present during synthesis (see Figure 16-4).

Mitochondrial Protein Import Requires Outer-Membrane Receptors and Translocons in Both Membranes

Figure 16-26 presents an overview of protein import from the cytosol into the mitochondrial matrix, the route into the mitochondrion followed by most imported proteins. We will discuss in detail each step in protein transport into the matrix



◀ **EXPERIMENTAL FIGURE 16-25** The post-translational uptake of precursor proteins into mitochondria can be assayed in a cell-free system. The imported protein must contain an appropriate uptake-targeting sequence. Uptake also requires ATP and a cytosolic extract containing chaperone proteins that maintain the precursor proteins in an unfolded conformation. Protein uptake occurs only with energized (respiring) mitochondria, which have a proton electrochemical gradient (proton-motive force) across the inner membrane. This assay has been used to study targeting sequences and other features of the translocation process.



◀ **FIGURE 16-26 Protein import into the mitochondrial matrix.** Precursor proteins synthesized on cytosolic ribosomes are maintained in an unfolded or partially folded state by bound chaperones, such as Hsc70 (step **1**). After a precursor protein binds to an import receptor near a site of contact with the inner membrane (step **2**), it is transferred into the general import pore (step **3**). The translocating protein then moves through this channel and an adjacent channel in the inner membrane (steps **4**, **5**). Note that translocation occurs at rare “contact sites” at which the inner and outer membranes appear to touch. Binding of the translocating protein by the matrix chaperone Hsc70 and subsequent ATP hydrolysis by Hsc70 helps drive import into the matrix. Once the uptake-targeting sequence is removed by a matrix protease and Hsc70 is released from the newly imported protein (step **6**), it folds into the mature, active conformation within the matrix (step **7**). Folding of some proteins depends on matrix chaperonins. See the text for discussion. [See G. Schatz, 1996, *J. Biol. Chem.* **271**:31763, and N. Pfanner et al., 1997, *Ann. Rev. Cell Dev. Biol.* **13**:25.]

and then consider how some proteins subsequently are targeted to other compartments of the mitochondrion.

After synthesis in the cytosol, the soluble precursors of mitochondrial proteins (including hydrophobic integral membrane proteins) interact directly with the mitochondrial membrane. In general, only unfolded proteins can be imported into the mitochondrion. Chaperone proteins such as cytosolic Hsc70 keep nascent and newly made proteins in an unfolded state, so that they can be taken up by mitochondria. Import of an unfolded mitochondrial precursor is initiated by the binding of a mitochondrial targeting sequence to an *import receptor* in the outer mitochondrial membrane. These receptors were first identified by experiments in which antibodies to specific proteins of the outer mitochondrial membrane were shown to inhibit protein import into

isolated mitochondria. Subsequent genetic experiments, in which the genes for specific mitochondrial outer-membrane proteins were mutated, showed that specific receptor proteins were responsible for the import of different classes of mitochondrial proteins. For example, N-terminal matrix-targeting sequences are recognized by Tom20 and Tom22. (Proteins in the outer mitochondrial membrane involved in targeting and import are designated *Tom* proteins for translocon of the outer membrane.)

The import receptors subsequently transfer the precursor proteins to an import channel in the outer membrane. This channel, composed mainly of the Tom40 protein, is known as the *general import pore* because all known mitochondrial precursor proteins gain access to the interior compartments of the mitochondrion through this channel. When

purified and incorporated into liposomes, Tom40 forms a transmembrane channel with a pore wide enough to accommodate an unfolded polypeptide chain. The general import pore forms a largely passive channel through the outer mitochondrial membrane, and the driving force for unidirectional transport into mitochondria comes from within the mitochondrion. In the case of precursors destined for the mitochondrial matrix, transfer through the outer membrane occurs simultaneously with transfer through an inner-membrane channel composed of the Tim23 and Tim17 proteins. (*Tim* stands for *translocon of the inner membrane*.) Translocation into the matrix thus occurs at “contact sites” where the outer and inner membranes are in close proximity.

Soon after the N-terminal matrix-targeting sequence of a protein enters the mitochondrial matrix, it is removed by a protease that resides within the matrix. The emerging protein also is bound by matrix Hsc70, a chaperone that is localized to the translocation channels in the inner mitochondrial membrane by interacting with Tim44. This interaction stimulates ATP hydrolysis by matrix Hsc70, and together these two proteins are thought to power translocation of proteins into the matrix.

Some imported proteins can fold into their final, active conformation without further assistance. Final folding of many matrix proteins, however, requires a *chaperonin*. As discussed in Chapter 3, chaperonin proteins actively facilitate protein folding in a process that depends on ATP. For instance, yeast mutants defective in Hsc60, a chaperonin in the mitochondrial matrix, can import matrix proteins and cleave their uptake-targeting sequence normally, but the imported polypeptides fail to fold and assemble into the native tertiary and quaternary structures.

Studies with Chimeric Proteins Demonstrate Important Features of Mitochondrial Import

Dramatic evidence for the ability of mitochondrial matrix-targeting sequences to direct import was obtained with chimeric proteins produced by recombinant DNA techniques. For example, the matrix-targeting sequence of alcohol dehydrogenase can be fused to the N-terminus of dihydrofolate reductase (DHFR), which normally resides in the cytosol. In the presence of chaperones, which prevent the C-terminal DHFR segment from folding in the cytosol, cell-free translocation assays show that the chimeric protein is transported into the matrix (Figure 16-27a). The inhibitor methotrexate, which binds tightly to the active site of DHFR and greatly stabilizes its folded conformation, renders the chimeric protein resistant to unfolding by cytosolic chaperones. When translocation assays are performed in the presence of methotrexate, the chimeric protein does not completely enter the matrix. This finding demonstrates that a precursor must be unfolded in order to traverse the import pores in the mitochondrial membranes.

Additional studies revealed that if a sufficiently long spacer sequence separates the N-terminal matrix-targeting se-

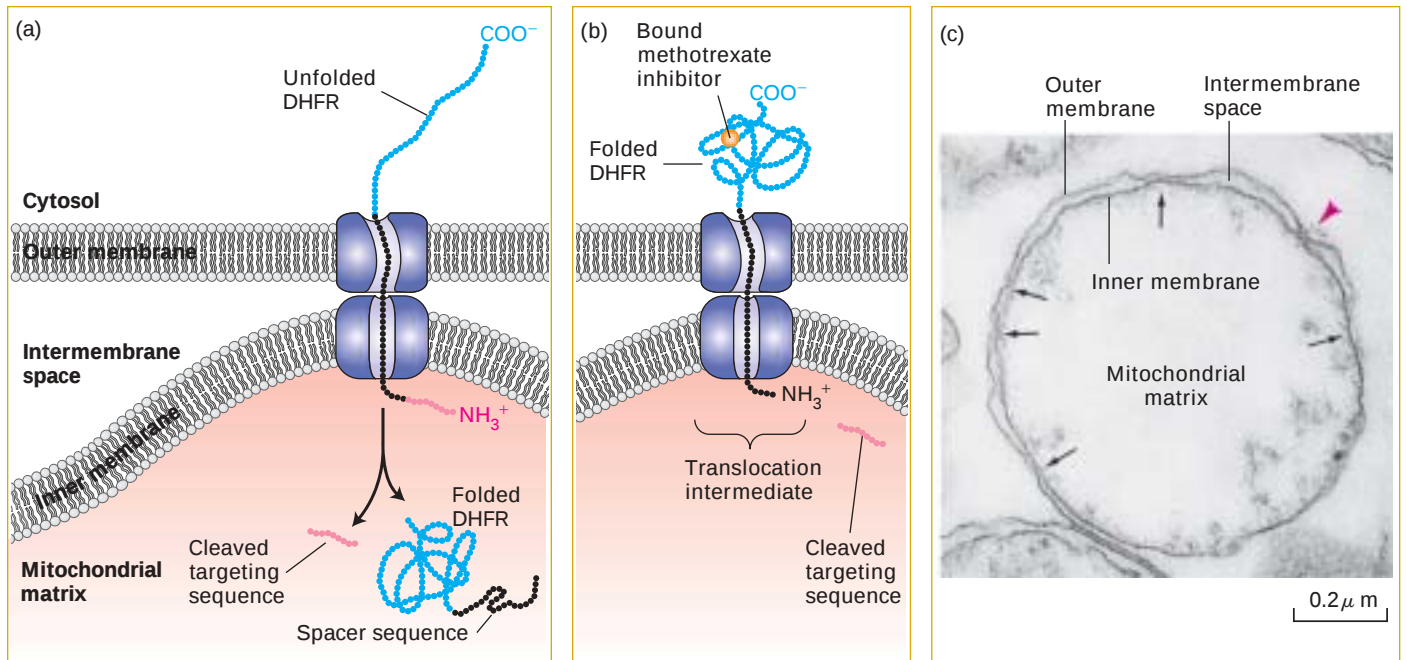
quence and DHFR portion of the chimeric protein, then a stable translocation intermediate forms in the presence of methotrexate (Figure 16-27b). In order for such a stable translocation intermediate to form, the spacer sequence must be long enough to span both membranes; a spacer of 50 amino acids extended to its maximum possible length is adequate to do so. If the chimera contains a shorter spacer—say, 35 amino acids—no stable translocation intermediate is obtained because the spacer cannot span both membranes. These observations provide further evidence that translocated proteins can span both inner and outer mitochondrial membranes and traverse these membranes in an unfolded state.

Microscopic studies of stable translocation intermediates show that they accumulate at sites where the inner and outer mitochondrial membranes are close together, evidence that precursor proteins enter only at such sites (Figure 16-27c). The distance from the cytosolic face of the outer membrane to the matrix face of the inner membrane at these *contact sites* is consistent with the length of an unfolded spacer sequence required for formation of a stable translocation intermediate. Moreover, stable translocation intermediates can be chemically cross-linked to the protein subunits that comprise the translocation channels of both the outer and inner membranes. This finding demonstrates that imported proteins can simultaneously engage channels in both the outer and inner mitochondrial membrane, as depicted in Figure 16-26. Since roughly 1000 stuck translocation intermediates can be observed in a typical yeast mitochondrion, it is thought that mitochondria have approximately 1000 general import pores for the uptake of mitochondrial proteins.

Three Energy Inputs Are Needed to Import Proteins into Mitochondria

As noted previously and indicated in Figure 16-26, ATP hydrolysis by Hsc70 chaperone proteins in both the cytosol and the mitochondrial matrix is required for import of mitochondrial proteins. Cytosolic Hsc70 expends energy to maintain bound precursor proteins in an unfolded state that is competent for translocation into the matrix. The importance of ATP to this function was demonstrated in studies in which a mitochondrial precursor protein was purified and then denatured (unfolded) by urea. When tested in the cell-free mitochondrial translocation system, the denatured protein was incorporated into the matrix in the absence of ATP. In contrast, import of the native, undenatured precursor required ATP for the normal unfolding function of cytosolic chaperones.

The sequential binding and ATP-driven release of multiple matrix Hsc70 molecules to a translocating protein may simply trap the unfolded protein in the matrix. Alternatively, the matrix Hsc70, anchored to the membrane by the Tim44 protein, may act as a molecular motor to pull the protein into the matrix (see Figure 16-26). In this case, the functions of matrix Hsc70 and Tim44 would be analogous to the chaperone BiP and Sec63 complex, respectively, in post-translational translocation into the ER lumen (see Figure 16-9).



▲ **EXPERIMENTAL FIGURE 16-27 Experiments with chimeric proteins show that a matrix-targeting sequence alone directs proteins to the mitochondrial matrix and that only unfolded proteins are translocated across both membranes.** The chimeric protein in these experiments contained a matrix-targeting signal at its N-terminus (red), followed by a spacer sequence of no particular function (black), and then by dihydrofolate reductase (DHFR), an enzyme normally present only in the cytosol. (a) When the DHFR segment is unfolded, the chimeric protein moves across both membranes to the matrix of energized mitochondria and the matrix-targeting signal then is removed. (b) When the C-terminus of the chimeric protein is locked in the folded state by binding of methotrexate, translocation is blocked. If the spacer sequence is long enough

to extend across the transport channels, a stable translocation intermediate, with the targeting sequence cleaved off, is generated in the presence of methotrexate, as shown here. (c) The C-terminus of the translocation intermediate in (b) can be detected by incubating the mitochondria with antibodies that bind to the DHFR segment, followed by gold particles coated with bacterial protein A, which binds nonspecifically to antibody molecules (see Figure 5-51). An electron micrograph of a sectioned sample reveals gold particles (red arrowhead) bound to the translocation intermediate at a contact site between the inner and outer membranes. Other contact site (black arrows) also are evident. [Parts (a) and (b) adapted from J. Rassow et al., 1990, *FEBS Letters* **275**:190. Part (c) from M. Schweiger et al., 1987, *J. Cell Biol.* **105**:235, courtesy of W. Neupert.]

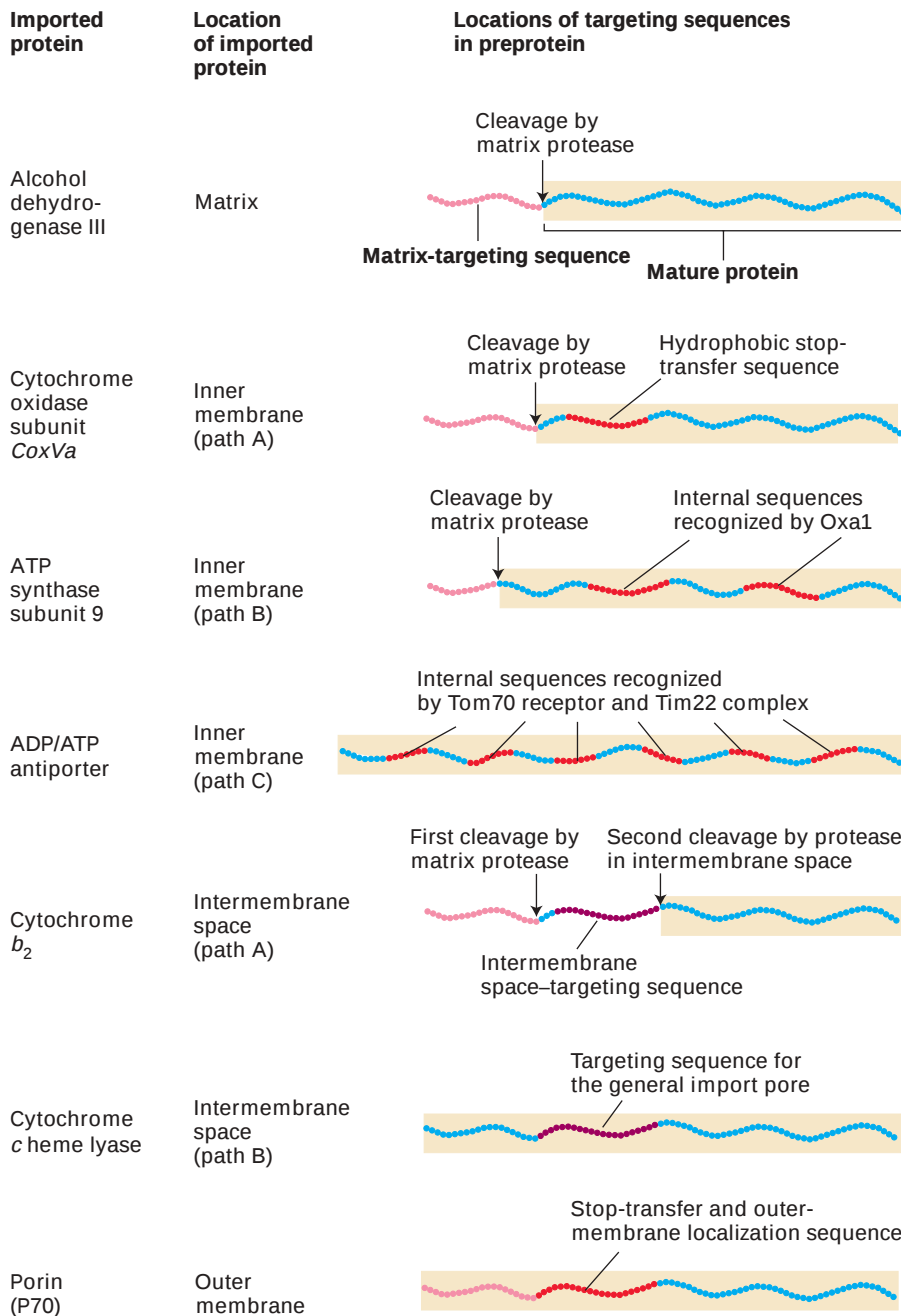
The third energy input required for mitochondrial protein import is a H^+ electrochemical gradient, or **proton-motive force**, across the inner membrane (Chapter 8). In general, only mitochondria that are actively undergoing respiration, and therefore have generated a proton-motive force across the inner membrane, are able to translocate precursor proteins from the cytosol into the mitochondrial matrix. Treatment of mitochondria with inhibitors or uncouplers of oxidative phosphorylation, such as cyanide or dinitrophenol, dissipates this proton-motive force. Although precursor proteins still can bind tightly to receptors on such poisoned mitochondria, the proteins cannot be imported, either in intact cells or in cell-free systems, even in the presence of ATP and chaperone proteins. Scientists do not fully understand how the proton-motive force is used to facilitate entry of a precursor protein into the matrix. Once a protein is partially inserted into the inner membrane, it is subjected to a transmembrane potential of 200 mV (matrix space negative), which is equivalent to an electric gradient of about 400,000

V/cm. One hypothesis is that the positive charges in the amphipathic matrix-targeting sequence could simply be “electrophoresed,” or pulled, into the matrix space by the inside-negative membrane electric potential.

Multiple Signals and Pathways Target Proteins to Submitochondrial Compartments

Unlike targeting to the matrix, targeting of proteins to the intermembrane space, inner membrane, and outer membrane of mitochondria generally requires more than one targeting sequence and occurs via one of several pathways. Figure 16-28 summarizes the organization of targeting sequences in proteins sorted to different mitochondrial locations.

Inner-Membrane Proteins Three separate pathways are known to target proteins to the inner mitochondrial membrane. One pathway makes use of the same machinery that is



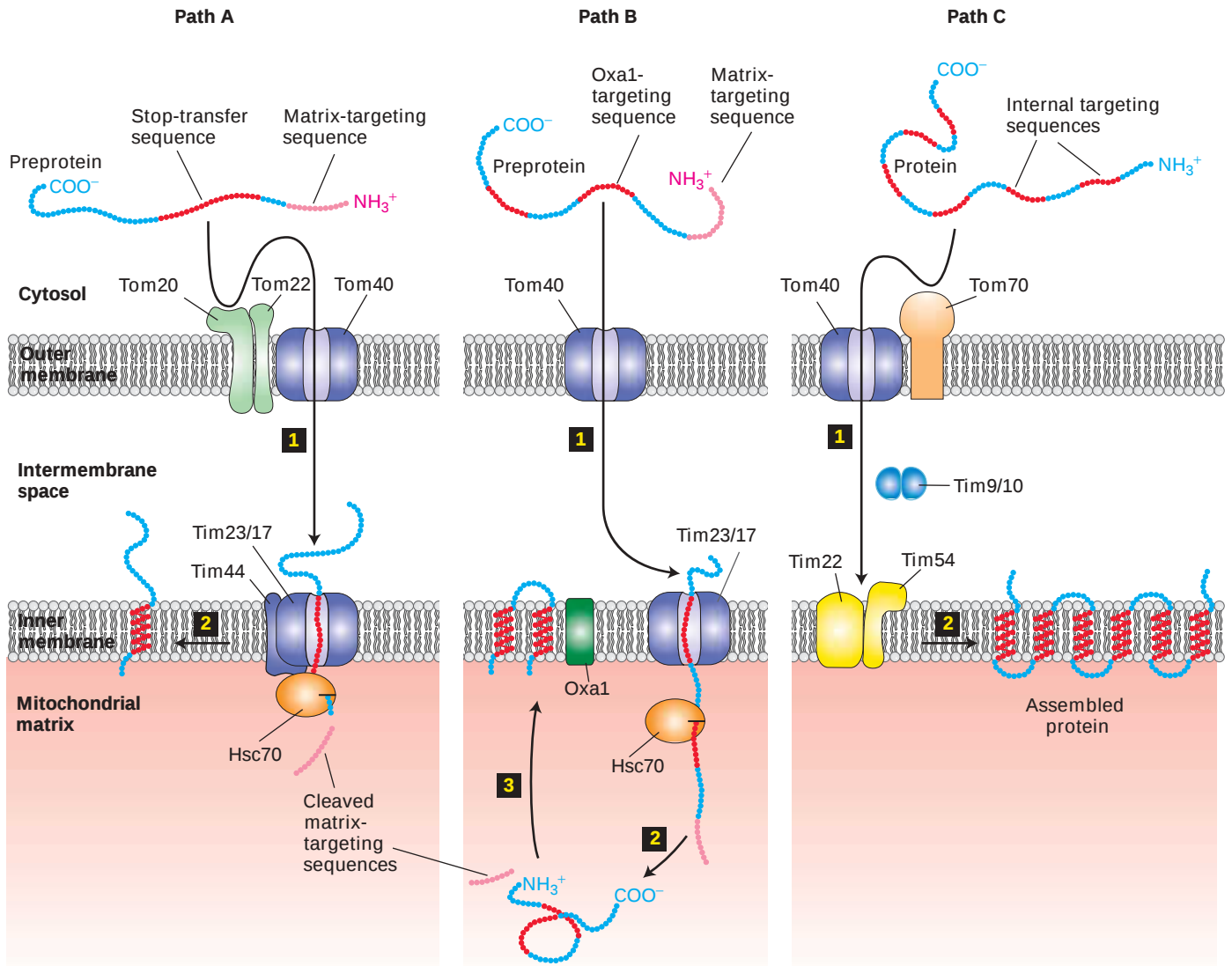
◀ **FIGURE 16-28 Arrangement of targeting sequences in imported mitochondrial proteins.**

Most mitochondrial proteins have an N-terminal matrix-targeting sequence (pink) that is similar but not identical in different proteins. Proteins destined for the inner membrane, the intermembrane space, or the outer membrane have one or more additional targeting sequences that function to direct the proteins to these locations by several different pathways. The designated pathways in parentheses correspond to those illustrated in Figures 16-29 and 16-30). [See W. Neupert, 1997, *Ann. Rev. Biochem.* **66**:863.]

used for targeting of matrix proteins (Figure 16-29, path A). A cytochrome oxidase subunit called CoxVa is a typical protein transported by this pathway. The precursor form of CoxVa, which contains an N-terminal matrix-targeting sequence recognized by the Tom20/22 import receptor, is transferred through the general import pore of the outer membrane and the inner-membrane Tim23/17 translocation complex. In addition to the matrix-targeting sequence, which is cleaved during import, CoxVa contains a hydrophobic stop-transfer sequence. As the protein passes through the Tim23/17 channel, the stop-transfer sequence blocks translocation of the C-terminus across the inner membrane. The

membrane-anchored intermediate is then transferred laterally into the bilayer of the inner membrane much as type I integral membrane proteins are incorporated into the ER membrane (see Figure 16-11).

A second pathway to the inner membrane is followed by proteins (e.g., ATP synthase subunit 9) whose precursors contain both a matrix-targeting sequence and internal hydrophobic domains recognized by an inner-membrane protein termed *Oxa1*. This pathway is thought to involve translocation of at least a portion of the precursor into the matrix via the Tom20/22 and Tim23/17 channels. After cleavage of the matrix-targeting sequence, the protein is in-



▲ **FIGURE 16-29 Three pathways for transporting proteins from the cytosol to the inner mitochondrial membrane.**

Proteins with different targeting sequences are directed to the inner membrane via different pathways. In all three pathways, proteins cross the outer membrane via the Tom40 general import pore. Proteins delivered by pathways A and B contain an N-terminal matrix-targeting sequence that is recognized by the Tom20/22 import receptor in the outer membrane. Although both these pathways use the Tim23/17 inner-membrane channel, they differ in that the entire precursor protein enters the matrix and

then is redirected to the inner membrane in pathway B. Matrix Hsc70 plays a role similar its role in the import of soluble matrix proteins (see Figure 16-26). Proteins delivered by pathway C contain internal sequences that are recognized by the Tom70 import receptor. A different inner-membrane translocation channel (Tim22/54) is used in this pathway. Two intermembrane proteins (Tim9 and Tim10) facilitate transfer between the outer and inner channels. See the text for discussion. [See R. E. Dalbey and A. Kuhn, 2000, *Ann. Rev. Cell Devel. Biol.* **16**:51, and N. Pfanner and A. Geissler, 2001, *Nature Rev. Mol. Cell Biol.* **2**:339.]

serted into the inner membrane by a process that requires interaction with Oxa1 and perhaps other inner-membrane proteins (Figure 16-29, path B). Oxa1 is related to a bacterial protein involved in inserting some inner-membrane proteins in bacteria. This relatedness suggests that Oxa1 may have descended from the translocation machinery in the endosymbiotic bacterium that eventually became the mitochondrion. However, the proteins forming the inner-

membrane channels in mitochondria are not related to the SecY protein in bacterial translocons. Oxa1 also participates in the inner-membrane insertion of certain proteins (e.g., subunit II of cytochrome oxidase) that are encoded by mitochondrial DNA and synthesized in the matrix by mitochondrial ribosomes.

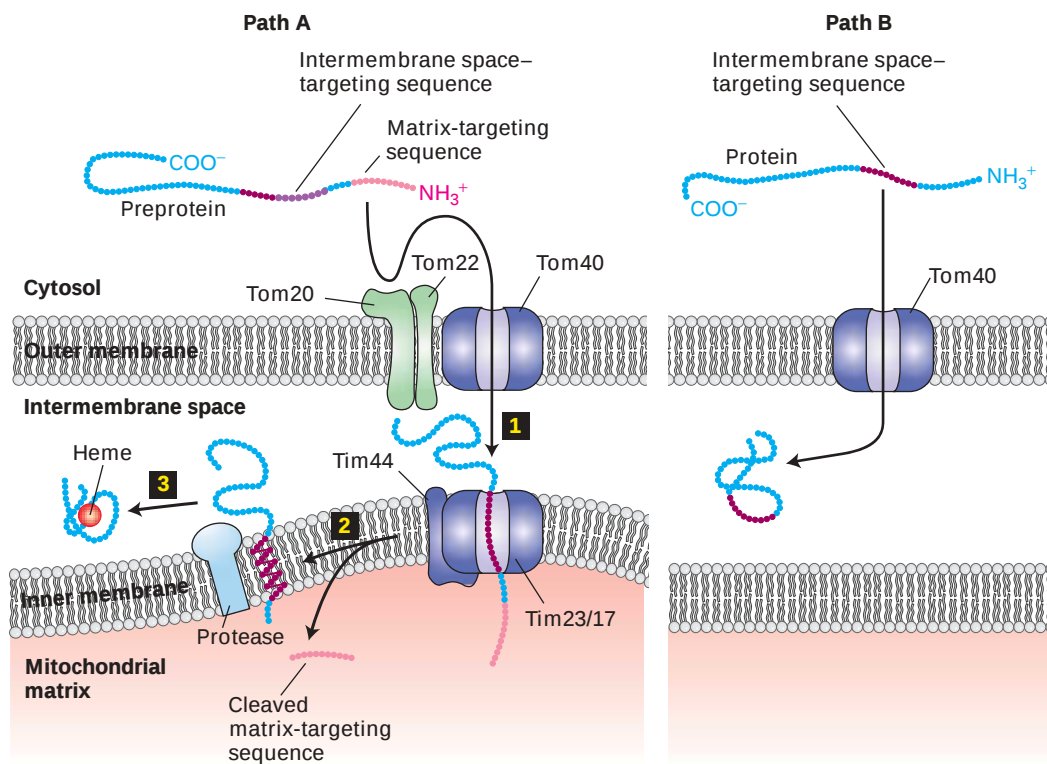
The final pathway for insertion in the inner mitochondrial membrane is followed by multipass proteins that con-

tain six membrane-spanning domains, such as the ADP/ATP antiporter. These proteins, which lack the usual N-terminal matrix-targeting sequence, contain multiple internal mitochondrial targeting sequences. After the internal sequences are recognized by Tom70, a second import receptor located in the outer membrane, the imported protein passes through the outer membrane through the general import pore (Figure 16-29, path C). The protein then is transferred to a second translocation complex in the inner membrane composed of the Tim22 and Tim54 proteins. Transfer to the Tim22/54 complex depends on a multimeric complex of two small proteins, Tim9 and Tim10, that reside in the intermembrane space. These may act as chaperones to guide imported proteins from the general import pore to the Tim22/54 complex in the inner membrane. Ultimately the Tim22/54 complex is responsible for incorporating the multiple hydrophobic segments of the imported protein into the inner membrane.

Intermembrane-Space Proteins Two pathways deliver cytosolic proteins to the space between the inner and outer mitochondrial membranes. The major pathway is followed by

proteins, such as cytochrome b_2 , whose precursors carry two different N-terminal targeting sequences, both of which ultimately are cleaved. The most N-terminal of the two sequences is a matrix-targeting sequence, which is removed by the matrix protease. The second targeting sequence is a hydrophobic segment that blocks complete translocation of the protein across the inner membrane (Figure 16-30, path A). After the resulting membrane-embedded intermediate diffuses laterally away from the Tim23/17 translocation channel, a protease in the membrane cleaves the protein near the hydrophobic transmembrane segment, releasing the mature protein in a soluble form into the intermembrane space. Except for the second proteolytic cleavage, this pathway is similar to that of inner-membrane proteins such as CoxVa (see Figure 16-29, path A).

Cytochrome c heme lyase, the enzyme responsible for the covalent attachment of heme to cytochrome c , illustrates a second pathway for targeting to the intermembrane space. In this pathway, the imported protein is delivered directly to the intermembrane space via the general import pore without involvement of any inner-membrane translocation factors



▲ **FIGURE 16-30 Two pathways for transporting proteins from the cytosol to the mitochondrial intermembrane space.**

Pathway A, the major one for delivery to the intermembrane space, is similar to pathway A for delivery to the inner membrane (see Figure 16-29). The major difference is that the internal targeting sequence in proteins such as cytochrome b_2 destined for the intermembrane space is recognized by an inner-membrane protease, which cleaves the protein on the

intermembrane-space side of the membrane. The released protein then folds and binds to its heme cofactor within the intermembrane space. Pathway B involves direct delivery to the intermembrane space through the Tom40 general import pore in the outer membrane. See the text for further discussion. [See R. E. Dalbey and A. Kuhn, 2000, *Ann. Rev. Cell Devel. Biol.* **16**:51; N. Pfanner and A. Geissler, 2001, *Nature Rev. Mol. Cell Biol.* **2**:339; and K. Diekert et al., 1999, *Proc. Nat'l. Acad. Sci. USA* **96**:11752.]

(Figure 16-30, path B). Since translocation through the Tom40 general import pore does not seem to be coupled to any energetically favorable process such as hydrolysis of ATP or GTP, the mechanism that drives unidirectional translocation through the outer membrane is unclear. One possibility is that cytochrome *c* heme lyase passively diffuses through the outer membrane and then is trapped within the intermembrane space by binding to another protein that is delivered to that location by one of the translocation mechanisms discussed previously.

Outer-Membrane Proteins Experiments with mitochondrial porin (P70) provide clues about how proteins are targeted to the outer mitochondrial membrane. A short matrix-targeting sequence at the N-terminus of P70 is followed by a long stretch of hydrophobic amino acids (see Figure 16-28). If the hydrophobic sequence is experimentally deleted from P70, the protein accumulates in the matrix space with its matrix-targeting sequence still attached. This finding suggests that the long hydrophobic sequence functions as a stop-transfer sequence that both prevents transfer of the protein into the matrix and anchors it as an integral protein in the outer membrane. Normally, neither the matrix-targeting nor stop-transfer sequence is cleaved from the anchored protein. The source of energy to drive outer membrane proteins through the general import pore has not yet been identified.

Targeting of Chloroplast Stromal Proteins Is Similar to Import of Mitochondrial Matrix Proteins

Among the proteins found in the chloroplast stroma are the enzymes of the Calvin cycle, which functions in fixing carbon dioxide into carbohydrates during photosynthesis (Chapter 8). The large (L) subunit of ribulose 1,5-bisphosphate carboxylase (rubisco) is encoded by chloroplast DNA and synthesized on chloroplast ribosomes in the stromal space. The small (S) subunit of rubisco and all the other Calvin cycle enzymes are encoded by nuclear genes and transported to chloroplasts after their synthesis in the cytosol. The precursor forms of these stromal proteins contain an N-terminal *stromal-import* sequence (see Table 16-1).

Experiments with isolated chloroplasts, similar to those with mitochondria illustrated in Figure 16-25, have shown that they can import the S-subunit precursor after its synthesis. After the unfolded precursor enters the stromal space, it binds transiently to a stromal Hsc70 chaperone and the N-terminal sequence is cleaved. In reactions facilitated by Hsc60 chaperonins that reside within the stromal space, eight S subunits combine with the eight L subunits to yield the active rubisco enzyme.

The general process of stromal import appears to be very similar to that for importing proteins into the mitochon-

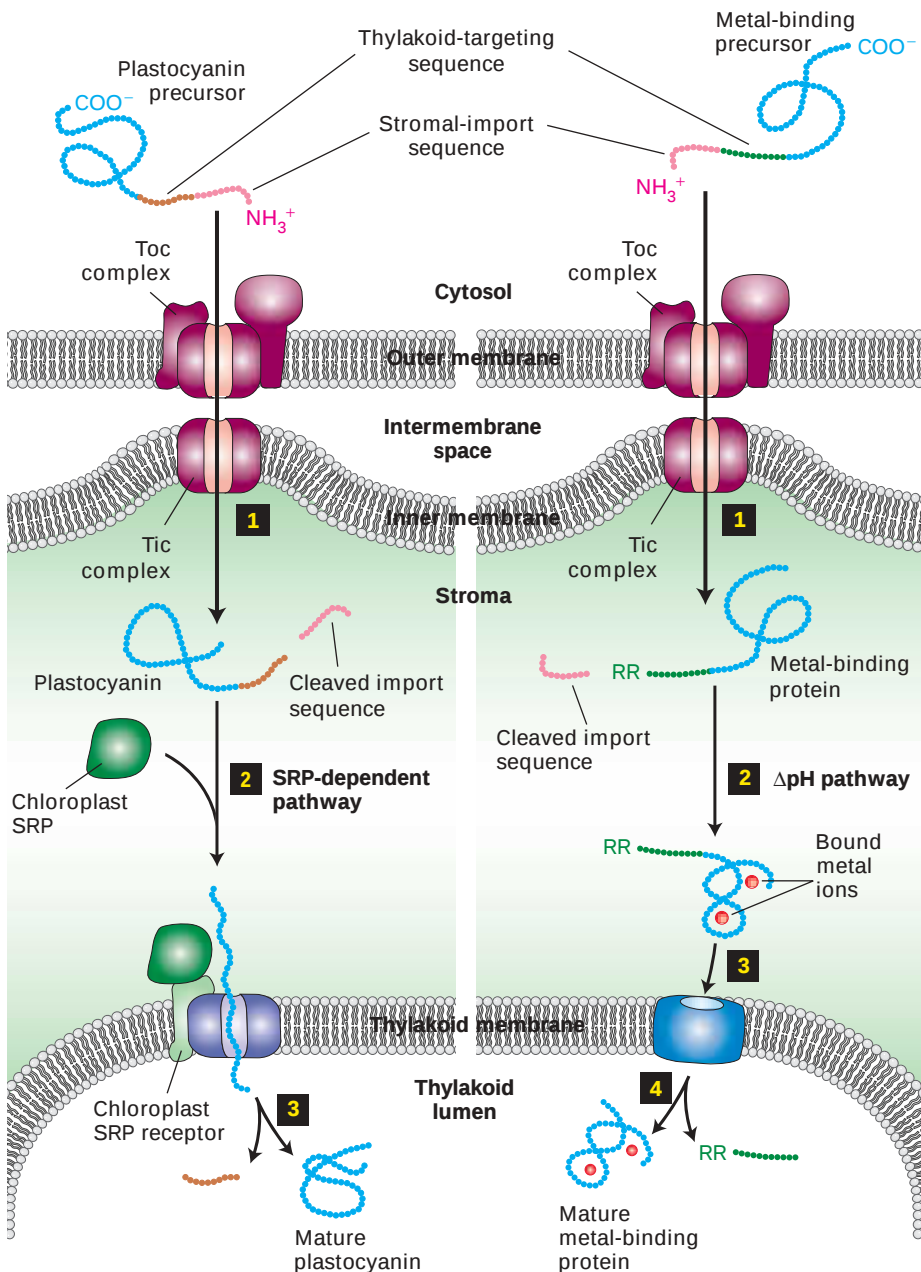
drial matrix (see Figure 16-26). At least three chloroplast outer-membrane proteins, including a receptor that binds the stromal-import sequence and a translocation channel protein, and five inner-membrane proteins are known to be essential for directing proteins to the stroma. Although these proteins are functionally analogous to the receptor and channel proteins in the mitochondrial membrane, they are not structurally homologous. The lack of homology between these chloroplast and mitochondrial proteins suggests that they may have arisen independently during evolution.

The available evidence suggests that chloroplast stromal proteins, like mitochondrial matrix proteins, are imported in the unfolded state. Import into the stroma depends on ATP hydrolysis catalyzed by a stromal Hsc70 chaperone whose function is similar to Hsc70 in the mitochondrial matrix and BiP in the ER lumen. Unlike mitochondria, chloroplasts cannot generate an electrochemical gradient (proton-motive force) across their inner membrane. Thus protein import into the chloroplast stroma appears to be powered solely by ATP hydrolysis.

Proteins Are Targeted to Thylakoids by Mechanisms Related to Translocation Across the Bacterial Inner Membrane

In addition to the double membrane that surrounds them, chloroplasts contain a series of internal interconnected membranous sacs, the **thylakoids** (see Figure 8-30). Proteins localized to the thylakoid membrane or lumen carry out photosynthesis. Many of these proteins are synthesized in the cytosol as precursors containing multiple targeting sequences. For example, plastocyanin and other proteins destined for the thylakoid lumen require the successive action of two uptake-targeting sequences. The first is an N-terminal stromal-import sequence that directs the protein to the stroma by the same pathway that imports the rubisco S subunit. The second sequence targets the protein from the stroma to the thylakoid lumen. The role of these targeting sequences has been shown in cell-free experiments measuring the uptake into chloroplasts of mutant proteins generated by recombinant DNA techniques. For instance, mutant plastocyanin that lacks the thylakoid-targeting sequence but contains an intact stromal-import sequence accumulates in the stroma and is not transported into the thylakoid lumen.

Four separate pathways for transporting proteins from the stroma into the thylakoid have been identified. All four pathways have been found to be closely related to analogous transport mechanisms in bacteria, illustrating the close evolutionary relationship between the stromal membrane and the bacterial inner membrane. Transport of plastocyanin and related proteins into the thylakoid lumen occurs by an SRP-dependent pathway (Figure 16-31, path A). A second pathway for transporting proteins into the thylakoid lumen



◀ **FIGURE 16-31 Two of the four pathways for transporting proteins from the cytosol to the thylakoid lumen.**

In these pathways, unfolded precursors are delivered to the stroma via the same outer-membrane proteins that import stromal-localized proteins. Cleavage of the N-terminal stromal-import sequence by a stromal protease then reveals the thylakoid-targeting sequence. At this point the two pathways diverge. In the SRP-dependent pathway (*left*), plastocyanin and similar proteins are kept unfolded in the stromal space by a set of chaperones (not shown) and, directed by the thylakoid-targeting sequence, bind to proteins that are closely related to the bacterial SRP, SRP receptor, and SecY translocon, which mediate movement into the lumen. After the thylakoid-targeting sequence is removed in the thylakoid lumen by a separate endoprotease, the protein folds into its mature conformation. In the pH-dependent pathway (*right*), metal-binding proteins fold in the stroma, and complex redox cofactors are added. Two arginine residues (RR) at the N-terminus of the thylakoid-targeting sequence and a pH gradient across the inner membrane are required for transport of the folded protein into the thylakoid lumen. The translocon in the thylakoid membrane is composed of at least four proteins related to proteins in the bacterial inner membrane. [See R. Dalbey and C. Robinson, 1999, *Trends Biochem. Sci.* **24**:17; R. E. Dalbey and A. Kuhn, 2000, *Ann. Rev. Cell Devel. Biol.* **16**:51; and C. Robinson and A. Bolhuis, 2001, *Nature Rev. Mol. Cell Biol.* **2**:350.]

utilizes a protein related to bacterial SecA and is thought to utilize a mechanism similar to that depicted in Figure 16-23. A third pathway, which targets proteins to the thylakoid membrane, depends on a protein related to the mitochondrial Oxa1 protein and the homologous bacterial protein (see Figure 16-29, path B). Some proteins encoded by chloroplast DNA and synthesized in the stroma or transported into the stroma from the cytosol are inserted into the thylakoid membrane via this pathway.

Finally, thylakoid proteins that bind metal-containing cofactors follow another pathway into the thylakoid lumen (Figure 16-31, ΔpH pathway). The unfolded precursors of these proteins are first targeted to the stroma, where the N-

terminal stromal-import sequence is cleaved off and the protein then folds and binds its cofactor. A set of thylakoid-membrane proteins assists in translocating the folded protein and bound cofactor into the thylakoid lumen, a process powered by the pH gradient normally maintained across the thylakoid membrane. The thylakoid-targeting sequence that directs a protein to this pH-dependent pathway includes two closely spaced arginine residues that are crucial for recognition. Bacterial cells also have a mechanism for translocating folded proteins with a similar arginine-containing sequence across the inner membrane. The molecular mechanism whereby these large folded globular proteins are transported across the thylakoid membrane is currently under intense study.

KEY CONCEPTS OF SECTION 16.5

Sorting of Proteins to Mitochondria and Chloroplasts

- Most mitochondrial and chloroplast proteins are encoded by nuclear genes, synthesized on cytosolic ribosomes, and imported post-translationally into the organelles.
- All the information required to target a precursor protein from the cytosol to the mitochondrial matrix or chloroplast stroma is contained within its N-terminal uptake-targeting sequence. After protein import, the uptake-targeting sequence is removed by proteases within the matrix or stroma.
- Cytosolic chaperones maintain the precursors of mitochondrial and chloroplast proteins in an unfolded state. Only unfolded proteins can be imported into the organelles. Translocation occurs at sites where the outer and inner membranes of the organelles are close together.
- Proteins destined to the mitochondrial matrix bind to receptors on the outer mitochondrial membrane, and then are transferred to the general import pore (Tom40) in the outer membrane. Translocation occurs concurrently through the outer and inner membranes, driven by the proton-motive force across the inner membrane and ATP hydrolysis by the Hsc70 ATPase in the matrix (see Figure 16-26).
- Proteins sorted to mitochondrial destinations other than the matrix usually contain two or more targeting sequences, one of which may be an N-terminal matrix-targeting sequence (see Figure 16-28).
- Some mitochondrial proteins destined for the inter-membrane space or inner membrane are first imported into the matrix and then redirected; others never enter the matrix but go directly to their final location.
- Protein import into the chloroplast stroma occurs through inner-membrane and outer-membrane translocation channels that are analogous in function to mitochondrial channels but composed of proteins unrelated in sequence to the corresponding mitochondrial proteins.
- Proteins destined for the thylakoid have secondary targeting sequences. After entry of these proteins into the stroma, cleavage of the stromal-targeting sequences reveals the thylakoid-targeting sequences.
- The three known pathways for moving proteins from the chloroplast stroma to the thylakoid closely resemble translocation across the bacterial inner membrane (see Figure 16-31). One of these systems can translocate folded proteins.

encoded by nuclear genes, synthesized on ribosomes free in the cytosol, and then incorporated into preexisting or newly generated peroxisomes. As peroxisomes are enlarged by addition of protein (and lipid), they eventually divide, forming new ones, as is the case with mitochondria and chloroplasts.

The size and enzyme composition of peroxisomes vary considerably in different kinds of cells. However, all peroxisomes contain enzymes that use molecular oxygen to oxidize various substrates, forming hydrogen peroxide (H_2O_2). Catalase, a peroxisome-localized enzyme, efficiently decomposes H_2O_2 into H_2O . Peroxisomes are most abundant in liver cells, where they constitute about 1 to 2 percent of the cell volume.

Cytosolic Receptor Targets Proteins with an SKL Sequence at the C-Terminus into the Peroxisomal Matrix

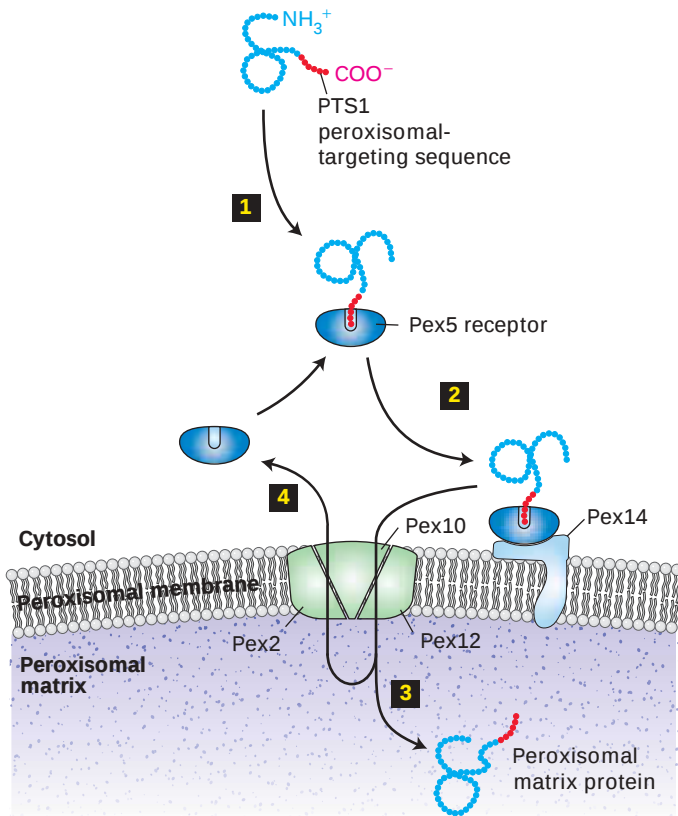
The import of catalase and other proteins into rat liver peroxisomes can be assayed in a cell-free system similar to that used for monitoring mitochondrial protein import (see Figure 16-25). By testing various mutant catalase proteins in this system, researchers discovered that the sequence Ser-Lys-Leu (SKL in one-letter code) or a related sequence at the C-terminus was necessary for peroxisomal targeting. Further, addition of the SKL sequence to the C-terminus of a normally cytosolic protein leads to uptake of the altered protein by peroxisomes in cultured cells. All but a few of the many different peroxisomal matrix proteins bear a sequence of this type, known as *peroxisomal-targeting sequence 1*, or simply *PTS1*.

The pathway for import of catalase and other PTS1-bearing proteins into the peroxisomal matrix is depicted in Figure 16-32. The PTS1 binds to a soluble receptor protein in the cytosol (Pex5), which in turn binds to a receptor in the peroxisome membrane (Pex14). The soluble and membrane-associated peroxisomal import receptors appear to have a function analogous to that of the SRP and SRP receptor in targeting proteins to the ER lumen. Still bound to Pex5, the imported protein then moves through a multimeric translocation channel, a feature that differs from protein import into the ER lumen. At some stage either during or after entry into the matrix, Pex5 dissociates from the peroxisomal matrix protein and is recycled back to the cytoplasm. In contrast to the N-terminal uptake-targeting sequences on proteins destined for the ER lumen, mitochondrial matrix, and chloroplast stroma, the PTS1 sequence is not cleaved from proteins after their entry into a peroxisome. Protein import into peroxisomes requires ATP hydrolysis, but it is not known how the energy released from ATP is used to power unidirectional translocation across the peroxisomal membrane.

The peroxisome import machinery, unlike most systems that mediate protein import into the ER, mitochondria, and chloroplast, can translocate folded proteins across the

16.6 Sorting of Peroxisomal Proteins

Peroxisomes are small organelles bounded by a single membrane. Unlike mitochondria and chloroplasts, peroxisomes lack DNA and ribosomes. Thus all peroxisomal proteins are



◀ **FIGURE 16-32 Import of peroxisomal matrix proteins directed by PTS1 targeting sequence.** Step **1**: Catalase and most other peroxisomal matrix proteins contain a C-terminal PTS1 uptake-targeting sequence (red) that binds to the cytosolic receptor Pex5. Step **2**: Pex5 with the bound matrix protein interacts with the Pex14 receptor located on the peroxisome membrane. Step **3**: The matrix protein–Pex5 complex is then transferred to a set of membrane proteins (Pex10, Pex12, and Pex2) that are necessary for translocation into the peroxisomal matrix by an unknown mechanism. Step **4**: At some point, either during translocation or in the lumen, Pex5 dissociates from the matrix protein and returns to the cytosol, a process that involves the Pex2/10/12 complex and additional membrane and cytosolic proteins not shown. Note that folded proteins can be imported into peroxisomes and that the targeting sequence is not removed in the matrix. [See P. E. Purdue and P. B. Lazarow, 2001, *Ann. Rev. Cell Devel. Biol.* **17**:701; S. Subramani et al., 2000, *Ann. Rev. Biochem.* **69**:399; and V. Dammal and S. Subramani, 2001, *Cell* **105**:187]

are eventually degraded. Genetic analyses of cultured cells from different Zellweger patients and of yeast cells carrying similar mutations have identified more than 20 genes that are required for peroxisome biogenesis. ■

membrane. For example, catalase assumes a folded conformation and binds to heme in the cytoplasm before traversing the peroxisomal membrane. Cell-free studies have shown that the peroxisome import machinery can transport a wide variety of molecules, including very large ones. To explain this unusual ability, scientists have speculated that a translocation channel of variable size may assemble by an unknown mechanism to fit exactly the diameter of the PTS1-bearing substrate molecule and then disassemble once translocation has been completed.

A few peroxisomal matrix proteins such as thiolase are synthesized as precursors with an N-terminal uptake-targeting sequence known as *PTS2*. These proteins bind to a different cytosolic receptor protein, but otherwise import is thought to occur by the same mechanism as for PTS1-containing proteins.

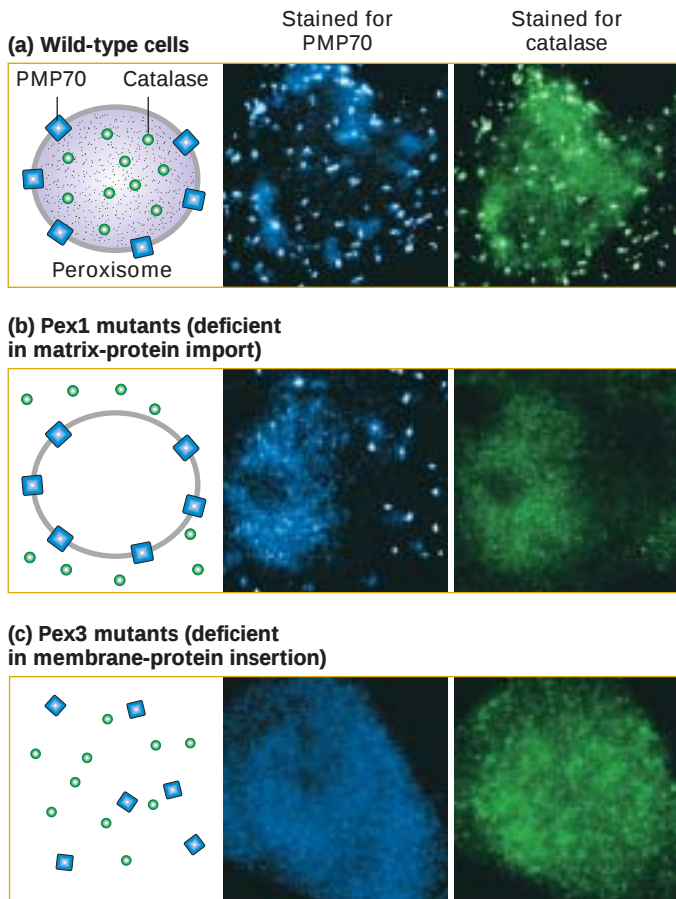
Peroxisomal Membrane and Matrix Proteins Are Incorporated by Different Pathways



Autosomal recessive mutations that cause defective peroxisome assembly occur naturally in the human population. Such defects can lead to severe impairment of many organs and to death. In *Zellweger syndrome* and related disorders, for example, the transport of many or all proteins into the peroxisomal matrix is impaired; newly synthesized peroxisomal enzymes remain in the cytosol and

Studies with peroxisome-assembly mutants have shown that different pathways are used for importing peroxisomal matrix proteins versus inserting proteins into the peroxisomal membrane. For example, analysis of cells from some Zellweger patients led to identification of genes encoding the putative translocation channel proteins Pex10, Pex12, and Pex2. Mutant cells defective in any one of these proteins cannot incorporate matrix proteins such as catalase into peroxisomes; nonetheless, the cells contain empty peroxisomes that have a normal complement of peroxisomal membrane proteins (Figure 16-33b). Mutations in any one of three other genes were found to block insertion of peroxisomal membrane proteins as well as import of matrix proteins (Figure 16-33c). These findings demonstrate that one set of proteins translocates soluble proteins into the peroxisomal matrix but a different set is required for insertion of proteins into the peroxisomal membrane. This situation differs markedly from that of the ER, mitochondrion, and chloroplast, for which, as we have seen, membrane proteins and soluble proteins share many of the same components for their insertion into these organelles.

Although most peroxisomes are generated by division of preexisting organelles, these organelles also can arise *de novo* by the two-stage process depicted in Figure 16-34. In this case, peroxisomal membrane proteins first are targeted to precursor membranes by sequences that differ from both PTS1 and PTS2. Analysis of mutant cells revealed that Pex19 is the receptor protein responsible for targeting of peroxisomal membrane proteins, while Pex3 and Pex16 are necessary for their proper insertion into the membrane. The insertion of peroxisomal membrane proteins generates membranes that have all the components necessary for import of matrix proteins, leading to the formation of mature,



◀ **EXPERIMENTAL FIGURE 16-33 Fluorescent-antibody staining of peroxisomal biogenesis mutants reveals different pathways for incorporation of membrane and matrix proteins.** Cells were stained with antibodies to PMP70, a peroxisomal membrane protein, or with antibodies to catalase, a peroxisomal matrix protein, then viewed in a fluorescent microscope. (a) In wild-type cells, both peroxisomal membrane and matrix proteins are visible as bright foci in numerous peroxisomal bodies. (b) In cells from a Pex12-deficient patient, catalase is distributed uniformly throughout the cytosol, whereas PMP70 is localized normally to peroxisomal bodies. (c) In cells from a Pex3-deficient patient, peroxisomal membranes cannot assemble, and as a consequence peroxisomal bodies do not form. Thus both catalase and PMP70 are mis-localized to the cytosol. [Courtesy of Stephen Gould, Johns Hopkins University.]

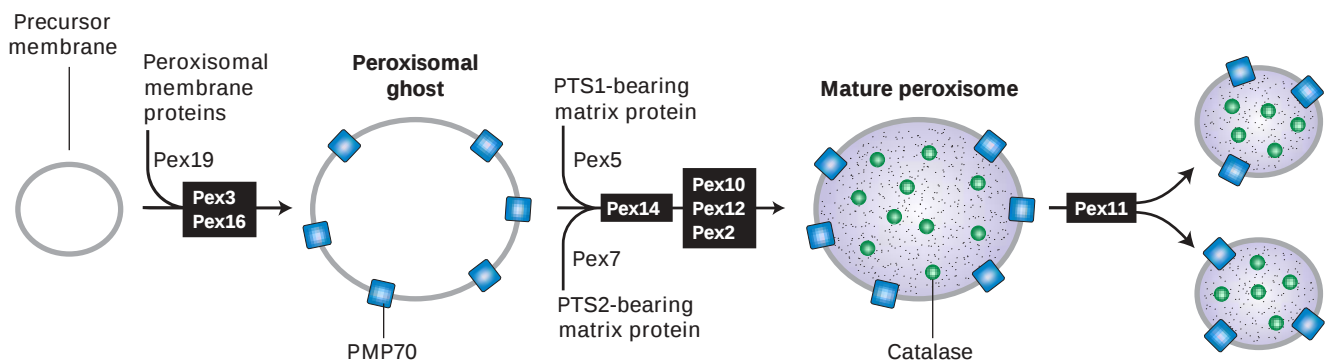
controls the extent of peroxisome division. The small peroxisomes generated by division can be enlarged by incorporation of additional matrix and membrane proteins via the same pathways described previously.

KEY CONCEPTS OF SECTION 16.6

Sorting of Peroxisomal Proteins

- All peroxisomal proteins are synthesized on cytosolic ribosomes and incorporated into the organelle post-translationally.
- Most peroxisomal matrix proteins contain a C-terminal PTS1 targeting sequence; a few have an N-terminal PTS2 targeting sequence. Neither targeting sequence is cleaved after import.
- All proteins destined for the peroxisomal matrix bind to a cytosolic receptor, which differs for PTS1- and PTS2-

functional peroxisomes. Division of mature peroxisomes, which largely determines the number of peroxisomes within a cell, depends on still another protein, Pex11. Over-expression of the Pex11 protein causes a large increase in the number of peroxisomes, suggesting that this protein



▲ **FIGURE 16-34 Model of peroxisomal biogenesis and division.** The first stage in the de novo formation of peroxisomes is the incorporation of peroxisomal membrane proteins into precursor membranes. Pex19 acts as the receptor for membrane-targeting sequences. Pex3 and Pex16 are required for the proper insertion of proteins into the forming peroxisomal membrane. Insertion of all peroxisomal membrane proteins produces a peroxisomal ghost, which is capable of importing proteins

targeted to the matrix. The pathways for importing PTS1- and PTS2-bearing matrix proteins differ only in the identity of the cytosolic receptor (Pex5 and Pex7, respectively) that binds the targeting sequence (see Figure 16-32). Complete incorporation of matrix proteins yields a mature peroxisome. The proliferation of peroxisomes requires division of mature peroxisomes, a process that depends on the Pex11 protein.

bearing proteins, and then are directed to common import receptor and translocation machinery on the peroxisomal membrane (see Figure 16-32).

- Translocation of matrix proteins across the peroxisomal membrane depends on ATP hydrolysis. Many peroxisomal matrix proteins fold in the cytosol and traverse the membrane in a folded conformation.
- Proteins destined for the peroxisomal membrane contain different targeting sequences than peroxisomal matrix proteins and are imported by a different pathway.
- Unlike mitochondria and chloroplasts, peroxisomes can arise *de novo* from precursor membranes, as well as by division of preexisting organelles (see Figure 16-34).

PERSPECTIVES FOR THE FUTURE

As we have seen in this chapter, we now understand many aspects of the basic processes responsible for selectively transporting proteins into the endoplasmic reticulum (ER), mitochondrion, chloroplast, and peroxisome. Biochemical and genetic studies, for instance, have identified *cis*-acting signal sequences responsible for targeting proteins to the correct organelle membrane and the membrane receptors that recognize these signal sequences. We also have learned much about the underlying mechanisms that translocate proteins across organelle membranes, and have determined whether energy is used to push or pull proteins across the membrane in one direction, the type of channel through which proteins pass, and whether proteins are translocated in a folded or an unfolded state. Nonetheless, many fundamental questions remain unanswered, including how fully folded proteins move across a membrane and how the topology of multipass membrane proteins is determined.

The peroxisomal import machinery provides one example of the translocation of folded proteins. It not only is capable of translocating fully folded proteins with bound cofactors into the peroxisomal matrix but can even direct the import of a large gold particle decorated with a (PTS1) peroxisomal targeting peptide. Some researchers have speculated that the mechanism of peroxisomal import may be related to that of nuclear import, the best-understood example of post-translational translocation of folded proteins (Chapter 12). Both the peroxisomal and nuclear import machinery can transport folded molecules of very divergent sizes, and both appear to involve a component that cycles between the cytosol and the organelle interior—the Pex5 PTS1 receptor in the case of peroxisomal import and the Ran-importin complex in the case of nuclear import. However, there also appear to be crucial differences between the two translocation processes. For example, nuclear pores represent large stable macromolecular assemblies readily observed by electron microscopy, whereas analogous porelike structures have not been observed in the peroxisomal membrane.

Moreover, small molecules can readily pass through nuclear pores, whereas peroxisomal membranes maintain a permanent barrier to the diffusion of small hydrophilic molecules. Taken together, these observations suggest that peroxisomal import may require an entirely new type of translocation mechanism. The evolutionarily conserved mechanisms for translocating folded proteins across the cytoplasmic membrane of bacterial cells and across the thylakoid membrane of chloroplasts also are poorly understood. A better understanding of all of these processes for translocating folded proteins across a membrane will likely hinge on future development of *in vitro* translocation systems that allow investigators to define the biochemical mechanisms driving translocation and to identify the structures of trapped translocation intermediates.

Compared with our understanding of how soluble proteins are translocated into the ER lumen and mitochondrial matrix, our understanding of how *cis*-acting sequences specify the topology of multipass membrane proteins is quite elementary. For instance, we do not know how the translocon channel accommodates polypeptides that are oriented differently with respect to the membrane, nor do we understand how local polypeptide sequences interact with the translocon channel both to set the orientation of transmembrane spans and to signal for lateral passage into the membrane bilayer. A better understanding of how the amino acid sequences of membrane proteins can specify membrane topology will be crucial for decoding the vast amount of structural information for membrane proteins contained within databases of genomic sequences.

A more detailed understanding of all translocation processes should continue to emerge from genetic and biochemical studies, both in yeasts and in mammals. These studies will undoubtedly reveal additional key proteins involved in the recognition of targeting sequences and in the translocation of proteins across lipid bilayers. Finally, the now mostly rudimentary structural studies of translocon channels will likely be extended in the future to reveal the structures and conformational states for the channels at resolutions on the atomic scale.

KEY TERMS

chaperones 665	signal-recognition particle (SRP) 661
cotranslational translocation 661	signal (uptake-targeting) sequences 659
general import pore 685	stop-transfer anchor sequence 667
hydropathy profile 671	topogenic sequences 666
<i>N</i> - and <i>O</i> -linked oligosaccharides 673	topology of membrane proteins 666
post-translational translocation 665	translocon 662
signal-anchor sequence 668	unfolded-protein response 679

REVIEW THE CONCEPTS

1. Describe the source or sources of energy needed for unidirectional translocation across the membrane in (a) cotranslational translocation into the endoplasmic reticulum (ER); (b) post-translational translocation into the ER; (c) translocation across the bacterial cytoplasmic membrane; and (d) translocation into the mitochondrial matrix.
2. Translocation into most organelles usually requires the activity of one or more cytosolic proteins. Describe the basic function of three different cytosolic factors required for translocation into the ER, mitochondria, and peroxisomes, respectively.
3. Describe the typical principles used to identify topogenic sequences within proteins and how these can be used to develop computer algorithms. How does the identification of topogenic sequences lead to prediction of the membrane arrangement of a multipass protein? What is the importance of the arrangement of positive charges relative to the membrane orientation of a signal anchor sequence?
4. The endoplasmic reticulum (ER) is an important site of “quality control” for newly synthesized proteins. What is meant by quality control in this context? What accessory proteins are typically involved in the processing of newly synthesized proteins within the ER? Cells generally degrade ER-exit-incompetent proteins. Where within the cell does such degradation occur and what is the relationship of Sec61p protein translocon to the degradation process?
5. Temperature-sensitive yeast mutants have been isolated that block each of the enzymatic steps in the synthesis of the dolichol-oligosaccharide precursor for *N*-linked glycosylation (see Figure 16-17). Propose an explanation for why mutations that block synthesis of the intermediate with the structure dolichol-PP-(GlcNAc)₂Man₅ completely prevent addition of *N*-linked oligosaccharide chains to secretory proteins, whereas mutations that block conversion of this intermediate into the completed precursor—dolichol-PP-(GlcNAc)₂Man₉Glc₃—allow the addition of *N*-linked oligosaccharide chains to secretory glycoproteins.
6. Name four different proteins that facilitate the modification and/or folding of secretory proteins within the lumen of the ER. Indicate which of these proteins covalently modifies substrate proteins and which brings about only conformational changes in substrate proteins.
7. Because you are interested in studying how a particular secretory protein folds within the ER, you wish to determine whether BiP binds to the newly synthesized protein in ER extracts. You find that you can isolate some of the newly synthesized secretory protein bound to BiP when ADP is added to the cell extract but not when ATP is added to the extract. Explain this result based on the mechanism for BiP binding to substrate proteins.

8. Describe how you might use recombinant DNA to engineer a strain of *Yersinia* such that it would be capable of inserting a protein of interest into the cytosol of mammalian macrophage cells.

9. Describe what would happen to the precursor of a mitochondrial matrix protein in the following types of mitochondrial mutants: (a) a mutation in the Tom22 signal receptor; (b) a mutation in the Tom70 signal receptor; (c) a mutation in the matrix Hsc70; and (d) a mutation in the matrix signal peptidase.

10. Describe the similarities and differences between the mechanism of import into the mitochondrial matrix and the chloroplast stroma.

11. Design a set of experiments using chimeric proteins, composed of a mitochondrial precursor protein fused to dihydrofolate reductase (DHFR), that could be used to determine how much of the precursor protein must protrude into the mitochondrial matrix in order for the matrix-targeting sequence to be cleaved by the matrix-processing protease (see Figure 16-27).

12. Protein targeting to both mitochondria and chloroplasts involves the sorting of proteins to multiple sites within the respective organelle. Briefly list these sites. Taking the mitochondrion as an example and the proteins ADP/ATP antiporter and cytochrome *b*₂ as the specific cases, compare and contrast the extent to which a common mechanism is used for the site-specific targeting of these two proteins.

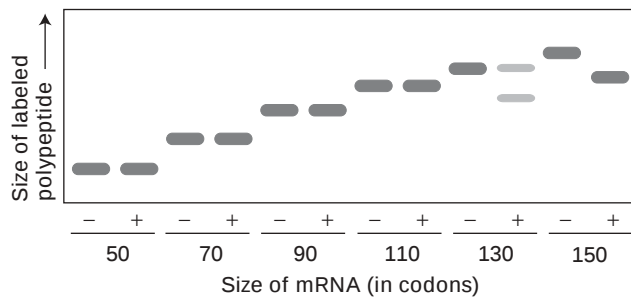
13. Suppose that you have identified a new mutant cell line that lacks functional peroxisomes. Describe how you could determine experimentally whether the mutant is primarily defective for insertion/assembly of peroxisomal membrane proteins or matrix proteins.

ANALYZE THE DATA

Imagine that you are evaluating the early steps in translocation and processing of the secretory protein prolactin. By using an experimental approach similar to that shown in Figure 16-7, you can use truncated prolactin mRNAs to control the length of nascent prolactin polypeptides that are synthesized. When prolactin mRNA that lacks a chain-termination (stop) codon is translated *in vitro*, the newly synthesized polypeptide ending with the last codon included on the mRNA will remain attached to the ribosome, thus allowing a polypeptide of defined length to extend from the ribosome. You have generated a set of mRNAs that encode segments of the N-terminus of prolactin of increasing length, and each mRNA can be translated *in vitro* by a cytosolic translation extract containing ribosomes, tRNAs, aminoacyl-tRNA synthetases, GTP, and translation initiation and elongation factors. When radio-labeled amino acids are included

in the translation mixture, only the polypeptide encoded by the added mRNA will be labeled. After completion of translation, each reaction mixture was resolved by SDS polyacrylamide gel electrophoresis, and the labeled polypeptides were identified by autoradiography.

a. The autoradiogram depicted below shows the results of an experiment in which each translation reaction was carried out either in the presence (+) or the absence (–) of microsomal membranes. Based on the gel mobility of peptides synthesized in the presence or absence of microsomes, deduce how long the prolactin nascent chain must be in order for the prolactin signal peptide to enter the ER lumen and to be cleaved by signal peptidase. (Note that microsomes carry significant quantities of SRP weakly bound to the membranes.)

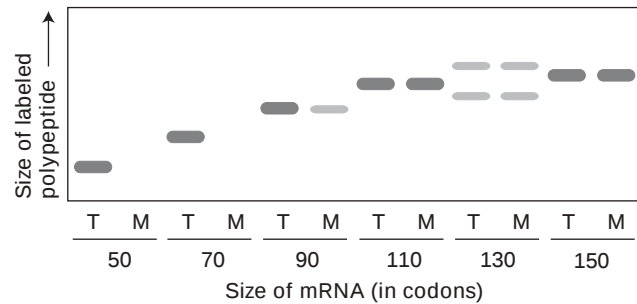


b. Given this length, what can you conclude about the conformational state of the nascent prolactin polypeptide when it is cleaved by signal peptidase? The following lengths will be useful for your calculation: the prolactin signal sequence is cleaved after amino acid 31; the channel within the ribosome occupied by a nascent polypeptide is about 150 Å long; a membrane bilayer is about 50 Å thick; in polypeptides with an α -helical conformation, one residue extends 1.5 Å, whereas in fully extended polypeptides, one residue extends about 3.5 Å.

c. The experiment described in part (a) is carried out in an identical manner except that microsomal membranes are not present during translation but are added after translation is complete. In this case none of the samples shows a difference in mobility in the presence or absence of microsomes. What can you conclude about whether prolactin can be translocated into isolated microsomes post-translationally?

d. In another experiment, each translation reaction was carried out in the presence of microsomes, and then the microsomal membranes and bound ribosomes were separated from free ribosomes and soluble proteins by centrifugation. For each translation reaction, both the total reaction (T) and the membrane fraction (M) were resolved in neighboring gel lanes. Based on the amounts of labeled polypeptide in the membrane fractions in the autoradiogram depicted below, deduce how long the prolactin nascent chain must be in order for ribosomes engaged in

translation to engage the SRP and thereby become bound to microsomal membranes.



REFERENCES

Translocation of Proteins Across the ER Membrane

Dalbey, R. E., and G. Von Heijne. 1992. Signal peptidases in prokaryotes and eukaryotes: a new protease family. *Trends Biochem. Sci.* **17**:474–478.

Johnson, A. E., and M. A. van Waes. 1999. The translocon: a dynamic gateway at the ER membrane. *Ann. Rev. Cell Devel. Biol.* **15**:799–842.

Keenan, R. J., D. M. Freymann, R. M. Stroud, and P. Walter. 2001. The signal recognition particle. *Ann. Rev. Biochem.* **70**:755–775.

Menetret, J. F., A. Neuhof, D. G. Morgan, K. Plath, M. Radermacher, T. A. Rapoport, and C. W. Akey. 2000. The structure of ribosome-channel complexes engaged in protein translocation. *Mol. Cell* **6**:1219–1232.

Rapoport, T. A., K. E. Matlack, K. Plath, B. Misselwitz, and O. Staack. 1999. Post-translational protein translocation across the membrane of the endoplasmic reticulum. *Biol. Chem.* **380**:1143–1150.

Insertion of Proteins into the ER Membrane

Englund, P. T. 1993. The structure and biosynthesis of glycosylphosphatidylinositol protein anchors. *Ann. Rev. Biochem.* **62**:121–138.

Goder, V., and M. Spiess. 2001. Topogenesis of membrane proteins: determinants and dynamics. *FEBS Lett.* **504**:87–93.

Mothes, W., et al. 1997. Molecular mechanism of membrane protein integration into the endoplasmic reticulum. *Cell* **89**:523–533.

von Heijne, G. 1999. Recent advances in the understanding of membrane protein assembly and structure. *Q. Rev. Biophys.* **32**:285–307.

Protein Modifications, Folding, and Quality Control in the ER

Bertolotti, A., Y. Zhang, L. M. Hendershot, H. P. Harding, and D. Ron. 2000. Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response. *Nature Cell Biol.* **2**:326–332.

Ellgaard, L., M. Molinari, and A. Helenius. 1999. Setting the standards: quality control in the secretory pathway. *Science* **286**:1882–1888.

Helenius, A., and M. Aebi. 2001. Intracellular functions of N-linked glycans. *Science* **291**:2364–2369.

Kornfeld, R., and S. Kornfeld. 1985. Assembly of asparagine-linked oligosaccharides. *Ann. Rev. Biochem.* **45**:631–664.

Parodi, A. J. 2000. Protein glucosylation and its role in protein folding. *Ann. Rev. Biochem.* **69**:69–93.

Patil, C., and P. Walter. 2001. Intracellular signaling from the endoplasmic reticulum to the nucleus: the unfolded protein response in yeast and mammals. *Curr. Opin. Cell Biol.* **13**:349–355.

Plempner, R. K., and D. H. Wolf. 1999. Retrograde protein translocation: eradication of secretory proteins in health and disease. *Trends Biochem. Sci.* **24**:266–270.

Sevier, C. S., and C. A. Kaiser. 2002. Formation and transfer of disulphide bonds in living cells. *Nature Rev. Mol. Cell Biol.* **3**:836–847.

Silberstein, S., and R. Gilmore. 1996. Biochemistry, molecular biology, and genetics of the oligosaccharyltransferase. *FASEB J.* **10**:849–858.

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

Export of Bacterial Proteins

Danese, P. N., and T. J. Silhavy. 1998. Targeting and assembly of periplasmic and outer-membrane proteins in *Escherichia coli*. *Ann. Rev. Genet.* **32**:59–94.

Galan, J. E. 2001. Salmonella interactions with host cells: type III secretion at work. *Ann. Rev. Cell Devel. Biol.* **17**:53–86.

Thanassi, D. G., and S. J. Hultgren. 2000. Multiple pathways allow protein secretion across the bacterial outer membrane. *Curr. Opin. Cell Biol.* **12**:420–430.

Wickner, W., and M. R. Leonard. 1996. *Escherichia coli* pre-protein translocase. *J. Biol. Chem.* **271**:29514–29516.

Sorting of Proteins to Mitochondria and Chloroplasts

Chen, K., X. Chen, and D. J. Schnell. 2000. Mechanism of protein import across the chloroplast envelope. *Biochem. Soc. Trans.* **28**:485–491.

Dalbey, R. E., and A. Kuhn. 2000. Evolutionarily related insertion pathways of bacterial, mitochondrial, and thylakoid membrane proteins. *Ann. Rev. Cell Devel. Biol.* **16**:51–87.

Matouschek, A., N. Pfanner, and W. Voos. 2000. Protein unfolding by mitochondria: the Hsp70 import motor. *EMBO Rept.* **1**:404–410.

Neupert, W., and M. Brunner. 2002. The protein import motor of mitochondria. *Nature Rev. Mol. Cell Biol.* **3**:555–565.

Pfanner, N., and A. Geissler. 2001. Versatility of the mitochondrial protein import machinery. *Nature Rev. Mol. Cell Biol.* **2**:339–349.

Robinson, C., and A. Bolhuis. 2001. Protein targeting by the twin-arginine translocation pathway. *Nature Rev. Mol. Cell Biol.* **2**:350–356.

Sorting of Peroxisomal Proteins

Dammai, V., and S. Subramani. 2001. The human peroxisomal targeting signal receptor, Pex5p, is translocated into the peroxisomal matrix and recycled to the cytosol. *Cell* **105**:187–196.

Gould, S. J., and C. S. Collins. 2002. Opinion: peroxisomal-protein import: is it really that complex? *Nature Rev. Mol. Cell Biol.* **3**:382–389.

Gould, S. J., and D. Valle. 2000. Peroxisome biogenesis disorders: genetics and cell biology. *Trends Genet.* **16**: 340–345.

Purdue, P. E., and P. B. Lazarow. 2001. Peroxisome biogenesis. *Ann. Rev. Cell Devel. Biol.* **17**:701–752.

Subramani, S., A. Koller, and W. B. Snyder. 2000. Import of peroxisomal matrix and membrane proteins. *Ann. Rev. Biochem.* **69**:399–418.