13.2 Insertion of Membrane Proteins into the ER

In previous chapters, we have encountered many of the vast array of integral membrane (transmembrane) proteins that are present throughout the cell. Each such protein has a unique orientation with respect to the membrane's phospholipid bilayer. Integral membrane proteins located in the ER, Golgi complex, and lysosomes, as well as in the plasma membrane, which are all synthesized on the rough ER, remain embedded in the membrane as they move to their final destinations along the same pathway that is followed by soluble secretory proteins (see Figure 13-1, *left*). During this transport, the orientation of a membrane protein is preserved; that is, the same segments of the protein always face the cytoplasmic side of the membrane, whereas other segments always face in the opposite direction, which could be considered the exoplasmic side of the membrane. Thus the final orientation of these membrane proteins is established during their biosynthesis on the ER membrane. As we will see in this section, insertion of membrane proteins into the ER membrane employs the same mechanism that has already been described for entry of soluble secretory proteins into the lumen of the ER. Membrane proteins first engage the translocon, made up of the Sec61 complex, by the interaction of a hydrophobic signal sequence with SRP and SRP receptor. Once engaged with the translocon, hydrophilic portions of the membrane protein can enter the ER lumen by translocation, whereas the hydrophobic

portions become embedded in the ER membrane by lateral opening of the translocon gate. The final topological orientation of membrane segments and hydrophilic regions of a membrane protein depends on the arrangement of hydrophobic sequences and neighboring positively charged amino acid residues. These sequences, known collectively as **topogenic sequences**, direct the membrane insertion and orientation of various classes of integral membrane proteins.

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

The **topology** of a membrane protein refers to the number of times its polypeptide chain spans the membrane and the orientation of those membrane-spanning segments within the membrane. The key elements of a protein that determine its topology are the membrane-spanning segments themselves, which are usually α helices containing 20–25 hydrophobic amino acids that contribute to energetically favorable interactions within the hydrophobic interior of the phospholipid bilayer.

Most integral membrane proteins fall into one of the five topological classes illustrated in Figure 13-10. Topological classes I, II, III, and the tail-anchored proteins are **single-pass membrane proteins**, which have only one membrane-spanning α-helical segment. Type I proteins have a cleaved N-terminal ER 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 ER 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 a hydrophobic membrane-spanning segment at their N-terminus and thus have the same orientation as type I proteins, but do not contain a cleavable signal sequence. Finally, tail-anchored proteins have a hydrophobic segment at their C-terminus that spans the membrane. These different topologies reflect distinct mechanisms used by the cell to establish the orientation of transmembrane segments, as we will see shortly.

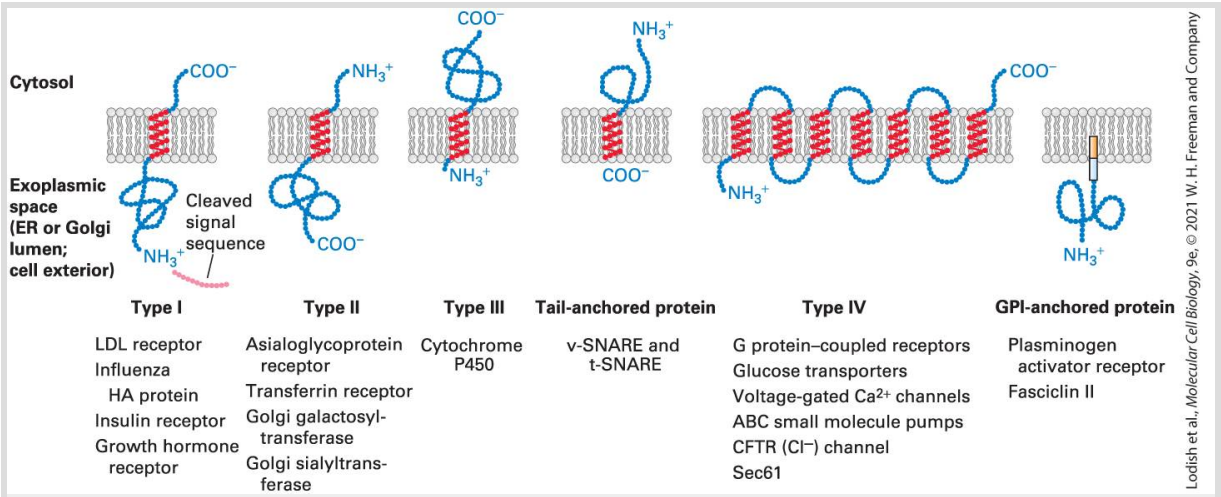

FIGURE 13-10 Classes of ER membrane proteins. Five topological classes of integral membrane proteins are synthesized on the rough ER, as is a sixth type tethered to the membrane by a phospholipid anchor. These membrane proteins are classified by their orientation in the membrane and the types of signals they contain to direct them there. In the integral membrane proteins, 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, *P. Natl. Acad. Sci. USA* **86**:5786; and C. A. Brown and S. D. Black, 1989, *J. Biol. Chem.* **264**:4442.

Description

Each illustration shows a section of E R membrane separating cytosol from exoplasmic space (E R or Golgi lumen; cell exterior). The classes are as follows:

Type 1 which includes L D L receptor, influenza H A protein, insulin receptor, and growth hormone receptor, consists of a helical portion in the membrane with the C-terminal in the cytosol and the N-terminal with folded hydrophilic portion in the exoplasmic space along with a cleaved signal sequence.

Type 2, which includes asialoglycoprotein receptor, transferrin receptor, Golgi galactosyl-transferase, and Golgi sialyltransferase, consists of a helical portion embedded in the membrane with the N-terminal on the cytosolic side and the C-terminal and hydrophilic folded portion in the exoplasmic space.

Type 3, which includes cytochrome p 450, consists of the N-terminal in the exoplasmic space and the C-terminal and folded hydrophilic portion in the cytosol.

Tail-anchored protein, which includes v-SNARE and t-SNARE, consists of a helical portion embedded in the membrane with C-terminal at the exoplasmic face and folded hydrophilic portion and N-terminal in the cytosol.

Type 4, which includes G-coupled receptors, glucose transporters, voltage-gated calcium channels, A B C small molecule pumps, C F T R chloride channels, and sec 61, consists of multiple alpha helices embedded in the membrane, connected by chains alternately on the cytosolic and exoplasmic faces, with the N-terminal on the exoplasmic face and the C-terminal in the cytosol.

G P I-anchored protein, which includes plasminogen activator receptors and fasciclin 2, consists of a hydrophilic portion in the exoplasmic space attached by the C-terminal to glycosylphosphatidylinositol in the membrane.

The proteins forming topological type IV contain two or more membrane-spanning segments and are sometimes called **multipass membrane proteins**. For example, many of the membrane transport proteins discussed in Chapter 11 and the numerous G protein–coupled receptors covered in Chapter 15 belong to this class.

When considering the mechanism of assembly of integral membrane proteins it is often useful to consider the topology of an individual membrane spanning segment. An N_{cyto}-C_{exo} transmembrane segment is oriented with the N-terminal end facing the inside of the cytoplasm and the C-terminal end facing exoplasmic side of the membrane that would be the lumen of the ER or the topological equivalent that is the outside of the cell. The other possible orientation is N_{exo}-C_{cyto}. Thus the transmembrane span of a type I membrane protein is N_{exo}-C_{cyto}, whereas the transmembrane span of a type II protein is N_{cyto}-C_{exo}. Note that for a type IV protein with multiple spans, the orientation strictly alternates between N_{cyto}-C_{exo} and N_{exo}-C_{cyto}.

Some lipid-anchored membrane proteins are also synthesized on the ER. These membrane proteins lack a hydrophobic membrane-spanning segment altogether; instead, they are linked to an amphipathic phospholipid anchor that is embedded in the membrane (Figure 13-10, *right*).

Internal Stop-Transfer Anchor and Signal-Anchor Sequences Determine

Topology of Single-Pass Proteins

We begin our discussion of how membrane protein topology is determined with the insertion of integral membrane proteins that contain a single hydrophobic membrane-spanning segment. As we will see, three main types of topogenic sequences are used to direct proteins to the ER membrane and to orient them within it. We have already introduced one, the N-terminal signal sequence. The other two types of sequences, which we will introduce here, are internal sequences known as *stop-transfer anchor sequences* and *signal-anchor sequences*. Unlike signal sequences, these two types of internal topogenic sequences end up in the mature protein as membrane-spanning segments. However, the two types differ in their final orientation in the membrane.

Type I Proteins

In addition to an N-terminal signal sequence that targets them to the ER, all type I transmembrane proteins possess an internal hydrophobic sequence of approximately 22 amino acids, which becomes the membrane-spanning α helix. The N-terminal signal sequence of a nascent type I protein, like that of a soluble 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 polypeptide chain continues to be extruded across the ER membrane. However, when the sequence that will become a transmembrane domain

enters the translocon, transfer of the protein through the channel stops and the transmembrane domain is allowed to move laterally from the channel into the membrane (Figure 13-11). The gating mechanism that allows lateral movement is the same as that for the opening of the translocon to accept the hydrophobic core of a signal sequence: two five-helix bundles of Sec61α hinge open to allow the hydrophobic transmembrane segment to move laterally past the hydrophobic signal-sequence-binding site through the opened edge of the translocon (see Figure 13-8). When the peptide exits the translocon in this manner, the hydrophobicity of the transmembrane segment anchors it in the hydrophobic interior of the membrane. Because such a sequence functions both to stop passage of the polypeptide chain through the translocon and to become a hydrophobic transmembrane segment in the membrane bilayer, it is called a **stop-transfer anchor sequence (STA)**.

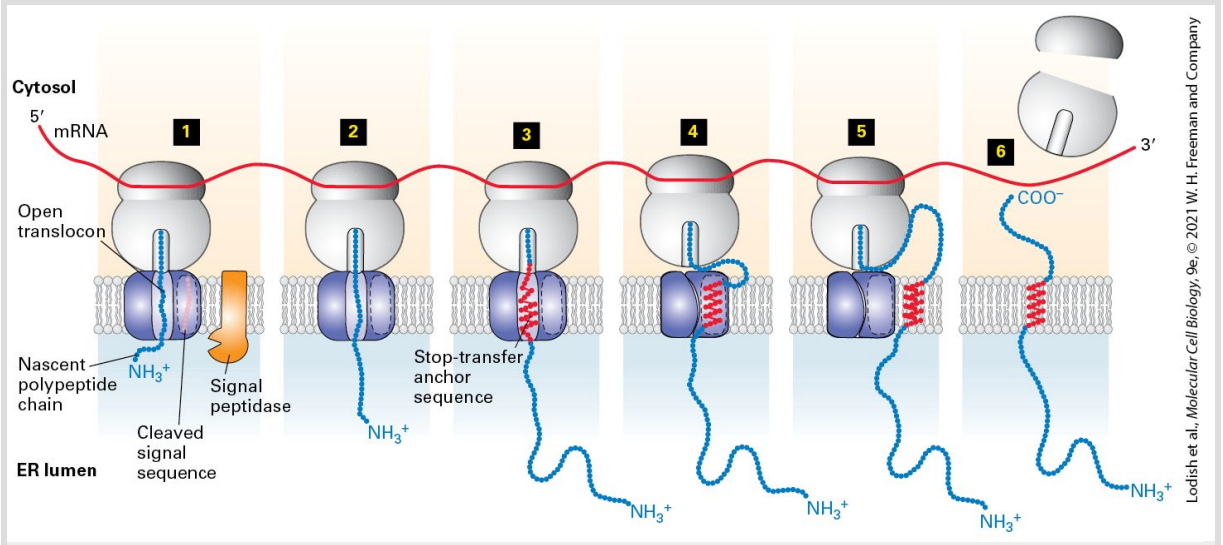

FIGURE 13-11 Membrane insertion and orientation of type I single-pass transmembrane proteins. Step 1: After the nascent polypeptide chain–ribosome complex becomes associated with a translocon in the ER membrane, the N-terminal signal sequence is cleaved. This process occurs by the same mechanism as the one for soluble secretory

proteins (see Figure 13-6). 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 through a hydrophobic cleft between translocon subunits and ultimately 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. Step 6: When synthesis is complete, the ribosomal subunits are released into the cytosol, leaving the protein free to diffuse laterally in the membrane. See H. Do et al., 1996, *Cell* **85**:369; and W. Mothes et al., 1997, *Cell* **89**:523.

Description
The illustrations show a translating ribosome attached to the cytosolic face of translocon embedded in the E R membrane separating cytosol from E R lumen. Step 1. A nascent N-terminal polypeptide chain enters the E R lumen. The signal peptidase is present adjacent to the translocon in the membrane and the signal sequence has been cleaved. Step 2. The polypeptide continues to grow. Step 3. A stop-transfer anchor sequence enters the translocon behind the growing polypeptide chain. Step 4. The stop-transfer anchor sequence stops the further entry of polypeptide chain. The growing chain on the cytosolic side begins to emerge into the cytosol. Step 5. The polypeptide chain in the cytosol continues to grow. Step 6. The ribosome disassociates from the membrane. The type 1 transmembrane protein is now embedded in the membrane with N-terminal chain in lumen, C-terminal in cytosol, and stop-transfer anchor sequence in the membrane.

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 complete, the ribosome is released from the translocon, and the C-terminus of the newly synthesized type I protein remains in the cytosol. Ultimately, the stop-

transfer anchor sequence resides in the membrane in an N_{exo}-C_{cyto} orientation.

Support for this mechanism has come from studies in which cDNAs encoding various mutant receptors for human growth hormone (HGH) were expressed in cultured mammalian cells. The wild-type HGH receptor, a typical type I protein, is normally transported to the plasma membrane. However, a mutant receptor that has charged residues inserted into the single membrane-spanning segment, or that is missing most of this segment altogether, is translocated entirely into the ER lumen and is eventually secreted from the cell as a soluble protein. These kinds of experiments have established that the hydrophobic membrane-spanning segment of the HGH receptor, and of other type I proteins, functions both as a stop-transfer sequence and as 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 (SA)** that functions as both an ER signal sequence and a membrane anchor. Recall that type II and type III proteins have opposite orientations in the membrane (see Figure 13-10); 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 polypeptide chain into the ER membrane so that the N-terminus of the chain faces the

cytosol, using the same SRP-dependent mechanism described for signal sequences (Figure 13-12a). However, the internal signal-anchor sequence does not have a recognition sequence for signal peptidase and therefore is *not* cleaved. Because of its hydrophobicity, the signal-anchor sequence can move laterally from the signal-sequence-binding site at the edge of the translocon directly into the phospholipid bilayer, where it functions as a membrane anchor. As elongation continues, the C-terminal region of the growing chain is extruded through the translocon into the ER lumen by cotranslational translocation.

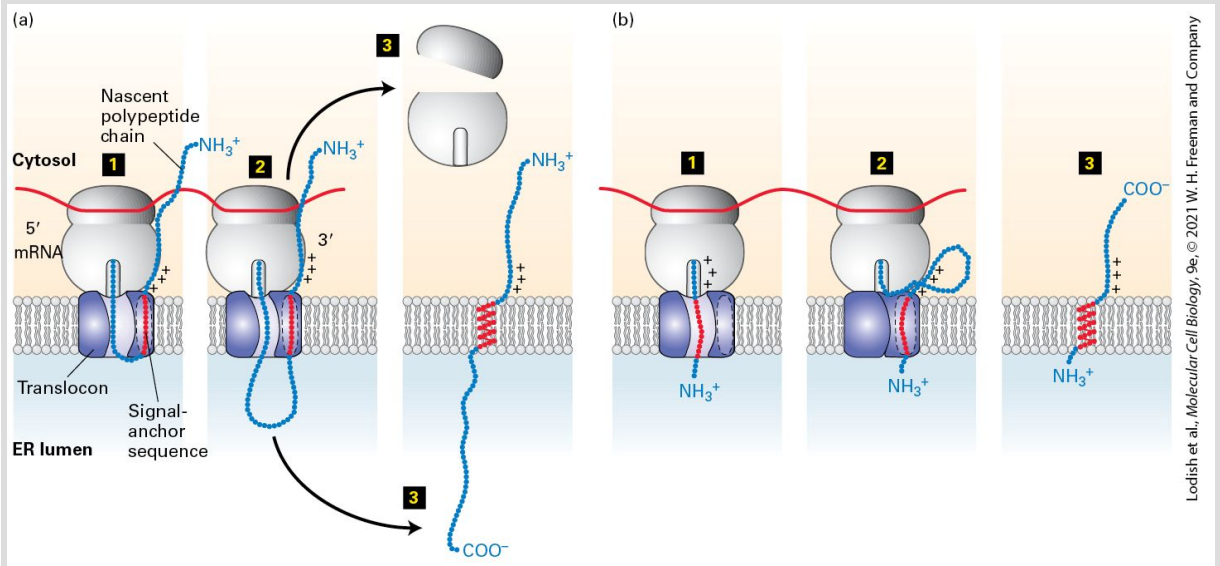

FIGURE 13-12 Membrane insertion and orientation of type II and type III single-pass transmembrane proteins. (a) Type II proteins. Step 1: After the internal signal-anchor sequence is synthesized on a cytosolic ribosome, it is bound by an SRP (not shown), which binds the SRP receptor on the ER membrane. This process is similar to the 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 polypeptide chain becomes oriented in the translocon with its N-terminal portion toward the cytosol. This orientation is dictated 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 sequence moves laterally through a hydrophobic cleft between translocon subunits and

anchors the chain in the phospholipid bilayer. Step 3: Once protein synthesis is complete, the C-terminus of the polypeptide is released into the lumen and the ribosomal subunits are released into the cytosol. (b) Type III proteins. Step 1: Insertion is by a process similar to that of type II proteins, except that positively charged residues on the C-terminal side of the signal-anchor sequence cause the transmembrane segment to be oriented within the translocon with its C-terminal portion toward the cytosol and the N-terminal end in the ER lumen. Note that when the signal-anchor segment engages with the translocon, the preceding hydrophilic segment is short enough to spontaneously pass through the translocon channel. Steps 2–3: Elongation of the C-terminal portion of the polypeptide chain is completed in the cytosol, and the ribosomal subunits are released. See M. Spiess and H. F. Lodish, 1986, *Cell* **44**:177; and H. Do et al., 1996, *Cell* **85**:369.

Description

Both (a) and (b) illustrations show a translating ribosome attached to the cytosolic face of translocon embedded in the E R membrane separating cytosol from E R lumen. The series (a) shows the process of type 2 proteins as follows: Step 1. A growing polypeptide sequence enters and exits the E R lumen through the translocon where the positively-charged signal-anchor sequence gets bound in the side pocket of translocon. Step 2. Polypeptide synthesis continues and the chain grows in the lumen. Step 3. The ribosome dissociates leaving behind the transmembrane protein with the N-terminal in the cytosol, C-terminal in the exoplasm, and signal-anchor sequence in the E R membrane.

The series (b) shows the process of type 3 proteins as follows: Step 1. A growing polypeptide sequence enters the translocon followed by the positively-charged signal-anchor sequence that gets bind into the translocon. Step 2. Polypeptide synthesis continues and the chain grows in the cytosol. Step 3. The ribosome exits leaving behind the transmembrane protein with the N-terminal in the exoplasm, C-terminal in the cytoplasm, and signal-anchor sequence in the E R membrane.

In the case of type III proteins, the signal-anchor sequence, which is located near the N-terminus, directs insertion of the nascent chain into the ER membrane with its N-terminus facing the lumen, in an orientation

opposite to that of the signal-anchor in type II proteins. The signal-anchor sequence of type III proteins also functions like a stop-transfer sequence and prevents further extrusion of the elongating chain into the ER lumen (Figure 13-12b). Continued elongation of the chain C-terminal to the signal-anchor sequence proceeds as it does for type I proteins, with the hydrophobic sequence eventually moving laterally out of the translocon to anchor the polypeptide in the ER membrane (see Figure 13-11).

The key difference between type II and type III proteins is the orientation of the hydrophobic transmembrane segment as it binds to the hydrophobic signal-sequence-binding site at the edge of Sec61α. The membrane span of a type II protein is oriented N_{cyto}-C_{exo} whereas the span of a type III protein is N_{exo}-C_{cyto}. The most important feature that determines whether a protein with a single transmembrane span will assume a type II or type III orientation is the length of the hydrophilic sequence that precedes the transmembrane span. If the hydrophilic segment is more than a few amino acids in length, the energetic cost of transferring this segment across the membrane is too great for the transmembrane segment to assume an N_{exo}-C_{cyto} orientation of a type III protein. The orientation of signal-anchor sequences that are preceded by only a few hydrophilic amino acids is determined by the arrangement of positively charged amino acids adjacent to one end or the other of the hydrophobic segment. These positively charged residues tend to remain on the cytosolic side of the membrane, rather than traversing the membrane into the ER lumen. Thus the position of the charged residues dictates the orientation of the signal-anchor sequence within the translocon as well as whether the rest of the polypeptide chain continues to pass into the ER lumen: type II proteins

tend to have positively charged residues on the N-terminal side of their signal-anchor sequence, orienting the N-terminus in the cytosol and allowing passage of the C-terminal side into the ER (see Figure 13-12a), whereas type III proteins tend to have positively charged residues on the C-terminal side of their signal-anchor sequence, which restrict the C-terminus to the cytosol (see Figure 13-12b). Note that the hydrophobic segment of a type II signal-anchor sequence assumes the same N_{cyto}-C_{exo} orientation as the signal sequence of a secreted protein and that in most respects these signal-anchor sequences behave exactly like signal sequences, although they are not cleaved.

A striking experimental demonstration of the importance of the flanking charge in determining orientation in the membrane is provided by neuraminidase, a type II protein in the surface coat of the 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.

Tail-Anchored Proteins

For all the topological classes of proteins we have considered so far, membrane insertion begins when the SRP recognizes a hydrophobic topogenic sequence as it emerges from the ribosome. Recognition of tail-

anchored proteins, which have a single hydrophobic topogenic sequence at the C-terminus, presents a unique challenge because the hydrophobic C-terminus becomes available for recognition only after translation has been completed and the protein has been released from the ribosome. Insertion of tail-anchored proteins into the ER membrane does not employ an SRP, SRP receptor, or the translocon, but instead depends on a pathway dedicated to this purpose, as depicted in Figure 13-13. This pathway involves a dimer of an ATPase known as *Get3*, which binds to the C-terminal hydrophobic segment of a tail-anchored protein. The Get3 dimer has two ATP-binding sites at the dimer interface in a structure that is similar to the two GTP-binding sites in the heterodimeric interface between SRP and SRP receptor shown in Figure 13-5b. The complex of a Get3 dimer bound to a tail-anchored protein is recruited to the ER by a dimeric integral membrane receptor known as *Get1/Get2*. The tail-anchored protein is released from Get3, and the transmembrane portion of Get1/Get2 participates in the insertion of the tail-anchor into the ER membrane. This process is mechanistically similar to the targeting of type II and type III signal-anchor sequences to the ER by the SRP and SRP receptor; the underlying structural similarity indicates that the mechanism of targeting is fundamentally related, despite the fact that Get3 couples targeting of tail-anchored proteins to ATP hydrolysis, whereas SRP couples protein targeting to GTP hydrolysis.

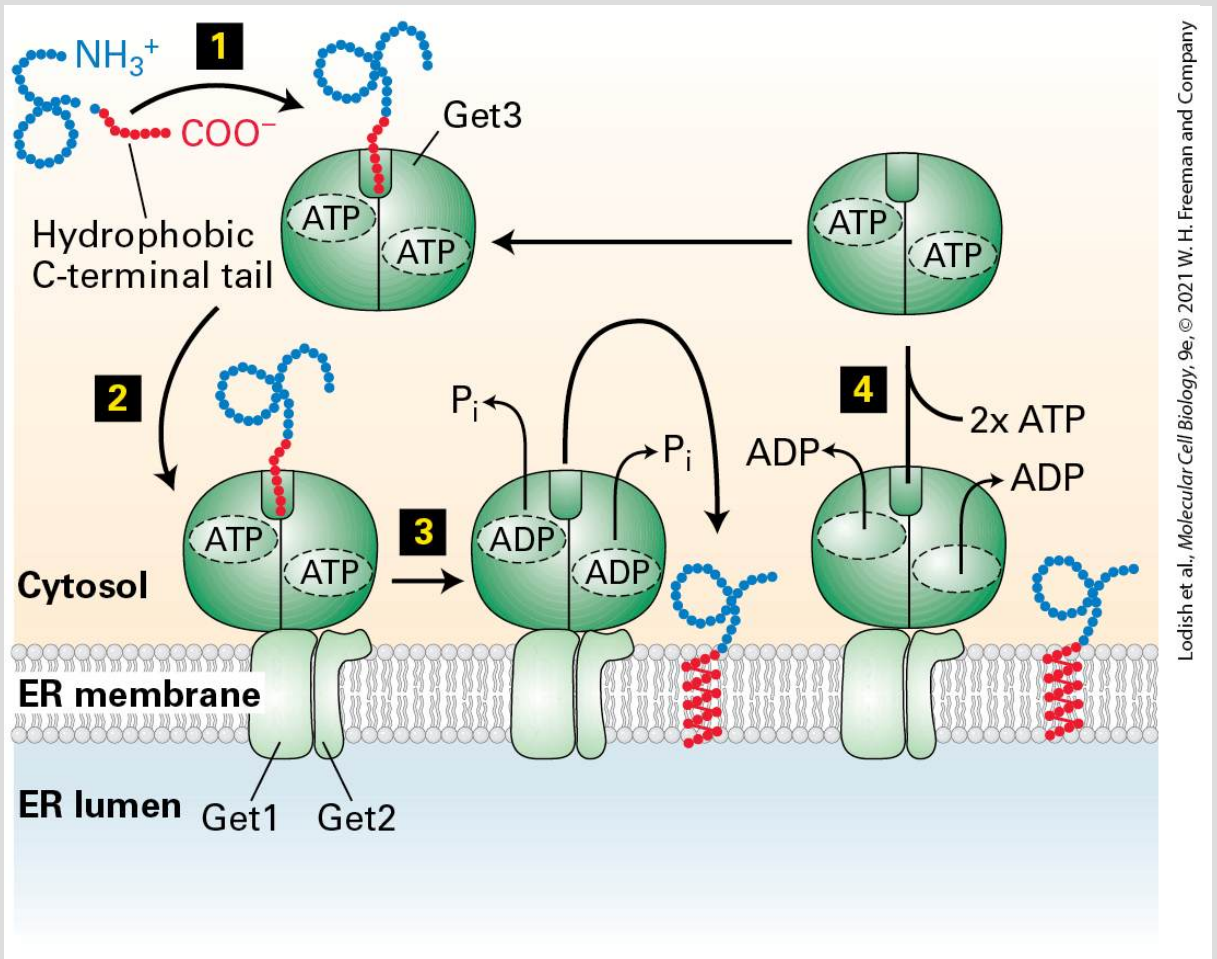

FIGURE 13-13 Insertion of tail-anchored proteins. For C-terminal tail-anchored proteins, the hydrophobic C-terminus is not available for membrane insertion until protein synthesis is complete and the protein has been released from the ribosome. Step 1: Dimeric Get3 contains a hydrophobic binding pocket at the interface between dimer subunits and Get3 in an ATP-bound state binds to the hydrophobic C-terminal tail of the protein. This binding reaction is facilitated by a complex of three other proteins, Sgt2, Get4, and Get5, which sequester the hydrophobic C-terminal tail before transferring it to Get3·ATP (not shown). Step 2: The ternary complex Get3·ATP bound to the protein docks onto the dimeric Get1/Get2 receptor, which is embedded in the ER membrane. Step 3: In succession, ATP is hydrolyzed and ADP is released from Get3. At the same time, the hydrophobic C-terminal tail is released from Get3 and ultimately becomes embedded in the ER membrane in a process that is facilitated by Get1/Get2. Step 4: Get3 binds to ATP and Get3·ATP is released from Get1/Get2 in a soluble form, ready for another round of binding to a hydrophobic C-terminal tail.

Description
The illustrations show a dimeric Get 1 slash Get 2 receptor embedded in the E R membrane separating cytosol from E R lumen. Step 1. A hydrophobic C-terminal tail of polypeptide binds in the pocket of a dimeric Get 3. Each subunit of Get 3 has an A T P attached to it. Step 2. The polypeptide bound dimeric Get 3 binds on the cytosolic surface of the dimeric Get 1 slash Get 2 receptor. Step 3. The polypeptide detaches from dimeric Get 3 as A T Ps are converted to A D P and P i. The C-terminal tail of the detached polypeptide gets embedded in the E R membrane beside the dimeric Get 1 slash Get 2 receptor. Step 4. A D Ps leave the dimeric Get 3, detaches from dimeric Get 1 slash Get 2 receptor, and binds to two A T Ps to get back to its initial structure.

Type IV (Multipass) Proteins

As described in Chapter 11, many physiologically important proteins, such as channel proteins, membrane transporters and pumps, and some receptor proteins, can contain 12 or more membrane-spanning α helices. Although their three-dimensional structure can be quite complex, the membrane topology of multipass proteins can usually be deduced from the same principles used to predict the topology of single-pass proteins. Figure 13-14 summarizes the arrangements of topogenic sequences in single-pass and multipass transmembrane proteins. Two key principles govern the assembly of the vast majority of multipass proteins: (1) the membrane-spanning segments pass from the translocon into the membrane cotranslationally in the sequential order by which they emerge from the ribosome and (2) the first topogenic segment engages the translocon in an SRP and SRP receptor-dependent manner as for single-pass proteins, whereas all subsequent transmembrane segments engage the translocon

independently of SRP. Based on these principles, the first transmembrane segment of a multipass protein acts as a topogenic sequence in the ways that we have already discussed for type I, type II, and type III proteins. Once the orientation of the first transmembrane segment is established (N_{exo}-C_{cyto} as with type I and type III proteins or N_{cyto}-C_{exo} as for type II proteins), each subsequent transmembrane segment assumes the opposite orientation as the one before. Thus an N_{exo}-C_{cyto} transmembrane segment is always followed by an N_{cyto}-C_{exo} transmembrane segment and an N_{cyto}-C_{exo} transmembrane segment is always followed by an N_{exo}-C_{cyto} transmembrane segment. This strict alternation of orientation is a simple consequence of the fact that the hydrophilic segments cannot move across the membrane unless threaded through the channel of the translocon. If a multipass protein has an *even* number of transmembrane α helices, its N-terminus and C-terminus will be oriented toward the same side of the membrane (Figure 13-14d). Conversely, if a type IV protein has an *odd* number of α helices, its two ends will have opposite orientations (Figure 13-14e).

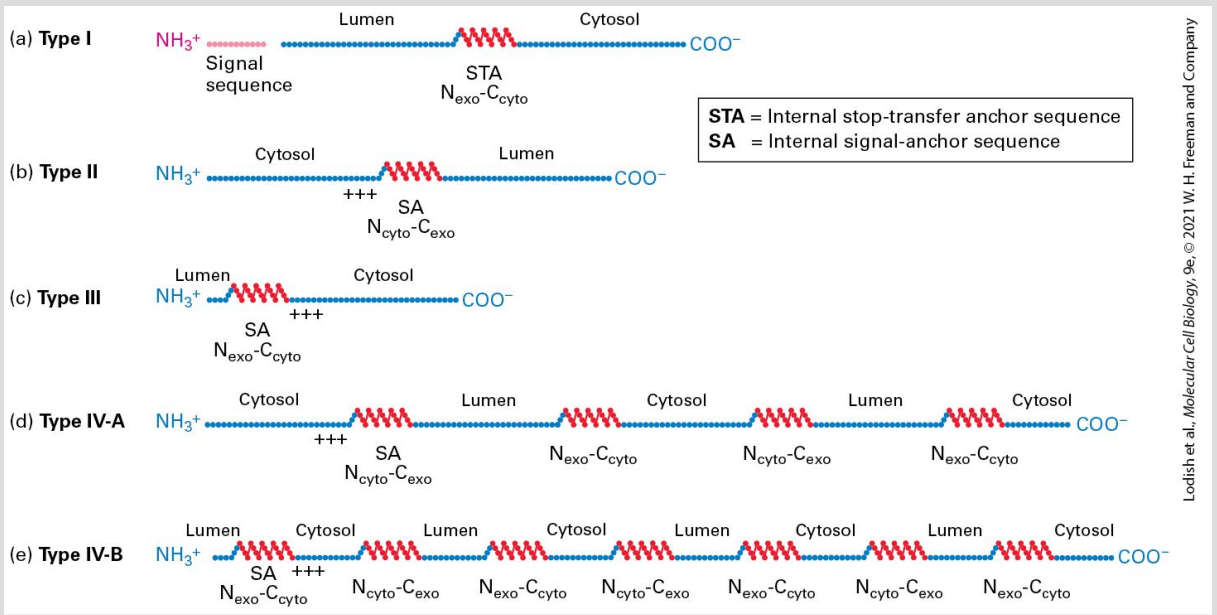

FIGURE 13-14 Topogenic sequences determine the orientation of ER membrane proteins. Topogenic sequences are shown in red; soluble, hydrophilic sequences in blue. The internal topogenic sequences form transmembrane α helices that anchor proteins or segments of proteins in the membrane. (a) Type I proteins contain a cleaved signal sequence and a single internal stop-transfer anchor (STA) sequence. (b, c) Type II and type III proteins contain a single internal signal-anchor (SA) sequence. The difference in the orientation of these protein types depends largely on whether there is a high density of positively charged amino acids $(+++)$ on the N-terminal side (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, begin with an SA sequence similar to a type II protein. Type IV-B proteins, whose N-terminus faces the lumen, begin with an SA sequence similar to a type III protein. After the first topogenic sequence all subsequent transmembrane segments insert into the membrane with the opposite polarity as the segment before.

Description

The sequences have S T A for internal stop-transfer anchor sequence, S A for internal signal anchor sequence, cytosol sequences, and lumen sequences. The topogenic sequences of different classes of transmembrane proteins starting from the N-terminal to C-terminal are as follows:

(a) Type 1: N-terminal, signal sequence, lumen, S T A, cytosol, C-terminal.

(b) Type 2: N-terminal, cytosol, positively charged region, S A, lumen, C-terminal.

(c) Type 3: N-terminal, very small lumen, S A, positively charged region, cytosol, C-terminal.

(d) Type 4-A: N-terminal, cytosol, positively charged region, S A, lumen, anchor sequence, cytosol, anchor sequence, lumen, anchor sequence, cytosol, C-terminal.

(e) Type 4-B: N-terminal, S A, positively charged region, cytosol, anchor sequence, lumen, anchor sequence, cytosol, anchor sequence, lumen, anchor sequence, cytosol, anchor sequence, lumen, anchor sequence, cytosol, C-terminal.

Type IV Proteins with N-Terminus in the Cytosol

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 11. In these proteins, the hydrophobic segment closest to the N-terminus functions like the internal signal-anchor sequence of a type II protein and is inserted into the ER membrane with an N_{cyto}-C_{exo} orientation (see Figure 13-12a). As the nascent chain following the transmembrane segment elongates, it moves through the translocon until the second hydrophobic segment 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 13-11).

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. As the C-terminus of the nascent chain then continues to grow into the cytosol, as it does in synthesis of type III proteins, the ribosome

remains attached to the translocon, and hydrophobic sequences that subsequently emerge from the ribosome are threaded into the translocon without the need for the SRP and the SRP receptor. As new hydrophobic topogenic sequences engage the translocon, the previously engaged sequences move laterally out of the translocon using the same mechanism as for type I, type II, and type III membrane proteins.

Type IV Proteins with N-Terminus in the Exoplasmic Space

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 transmembrane segment closest to the N-terminus is often followed by a cluster of positively charged amino acids, like a type III signal-anchor sequence (see Figure 13-12b). As a result, the nascent polypeptide chain is inserted into the translocon with the N-terminus extending into the lumen (see Figure 13-14e). As the chain is elongated, it is inserted into the ER membrane in alternating orientation, 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 13-15a; see also Chapter 7). 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 an internal stop-transfer anchor sequence directing the process (see Figure 13-11). 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 stop-transfer anchor sequence and transfers the luminal portion of the protein to a preformed GPI anchor in the membrane (Figure 13-15b).

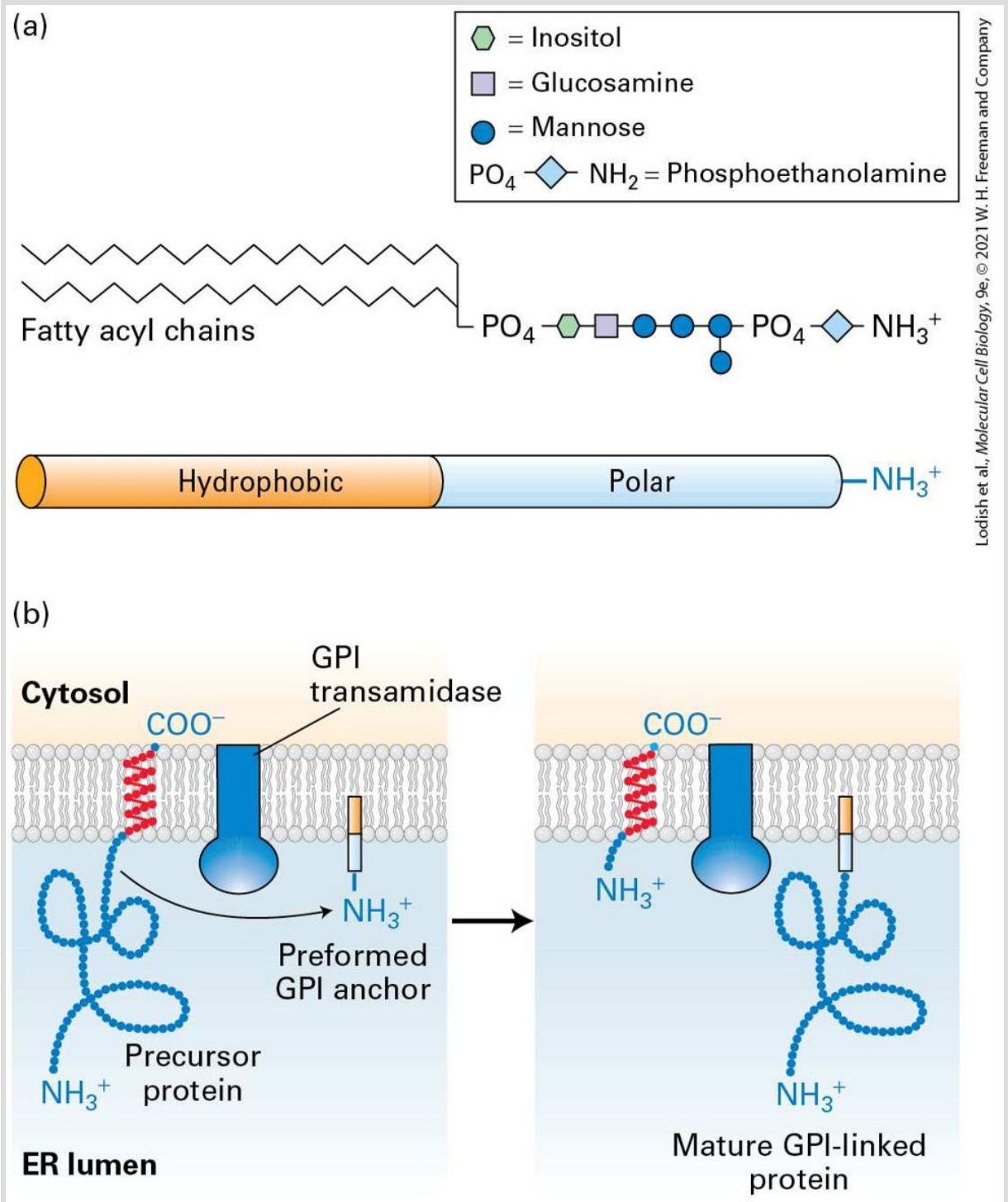

FIGURE 13-15 GPI-anchored proteins. (a) Structure of a glycosylphosphatidylinositol (GPI) molecule from yeast. The hydrophobic portion of the molecule is composed of fatty acyl chains, whereas the polar (hydrophilic) portion is composed of carbohydrate residues and phosphate groups. In other organisms, 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 like a type I transmembrane protein, as shown in Figure 13-11. 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, *P. Natl. Acad. Sci. USA* **89**:4982.

Description

A two-part illustration (a) shows the structure of G P I-anchored proteins. The first part two long wavy fatty acyl chains bound to a phosphate which is further attached to inositol, glucosamine, four mannoses, and phosphoethanolamine at the N H 3 plus end. The second part shows a horizontal rod with left half labeled hydrophobic and right half labeled polar with an N H 3 plus end.

A two-part illustration (b) shows the formation of G P I-anchored proteins on E R membrane. The first part shows a G P I transamidase embedded in the E R membrane adjacent (left side) to a precursor protein having its C-terminal end embedded in E R membrane and N-terminal end with a long chain in E R lumen. A pre-formed G P I anchor is embedded in E R membrane, to the right side of G P I transamidase and have its polar N H 3 plus end protruding in the lumen. An arrow from precursor protein points to pre-formed G P I anchor. The second part shows the lumen precursor protein now bound to the polar N H 3 plus end of the pre-formed G P I anchor which is now labeled, mature G P I-linked protein. The C-terminal end of protein is still anchored in the E R membrane.

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 relatively rapidly in the plane of the phospholipid bilayer. In contrast, many proteins anchored by membrane-

spanning α helices are impeded from moving laterally 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 (instead of the basolateral domain) in certain polarized epithelial cells, as we discuss in Chapter 14.

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

As we have seen, various topogenic sequences in integral membrane proteins synthesized on the ER govern the interaction of the nascent polypeptide chain with the translocon. When scientists begin to study a protein of unknown function, the identification of potential 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 suggest that the protein is a type I integral membrane protein and therefore may be a cell-surface receptor for an extracellular ligand. Furthermore, the implied type I topology suggests that the N-terminal segment that lies between the signal sequence and the internal hydrophobic sequence probably constitutes the extracellular domain with a role in ligand binding, whereas the C-terminal segment that lies after the

internal hydrophobic sequence is probably cytosolic and may play a role in intracellular signaling.

Identification of topogenic sequences requires a way to scan sequence databases for segments that are sufficiently hydrophobic to be either signal sequences or transmembrane anchor sequences. 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 long segments of sufficient overall hydrophobicity to be N-terminal signal sequences or internal stop-transfer anchor sequences and signal-anchor sequences. To accomplish this, the total hydropathic index for each segment 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 13-16 shows the hydropathy profiles for three different membrane proteins. The prominent peaks in such plots identify probable topogenic sequences as well as their positions and approximate lengths. 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 anchor sequence (Figure 13-16a). On the basis of this profile, we can deduce, correctly, that the HGH receptor is a type I integral membrane protein. The hydropathy profile of the asialoglycoprotein receptor, a cell-surface protein that mediates removal of abnormal extracellular glycoproteins, reveals a prominent internal hydrophobic signal-anchor sequence, but gives no indication of a hydrophobic N-terminal signal sequence (Figure 13-16b). Thus we can predict that the asialoglycoprotein receptor is a type II or type III membrane protein. Because the N-terminal segment of the asialoglycoprotein receptor contains 30 hydrophilic residues that could not move across the membrane without a signal sequence to initiate engagement with the translocon, we can deduce that this segment must remain in the cytosol and we can correctly predict that this is a type II protein.

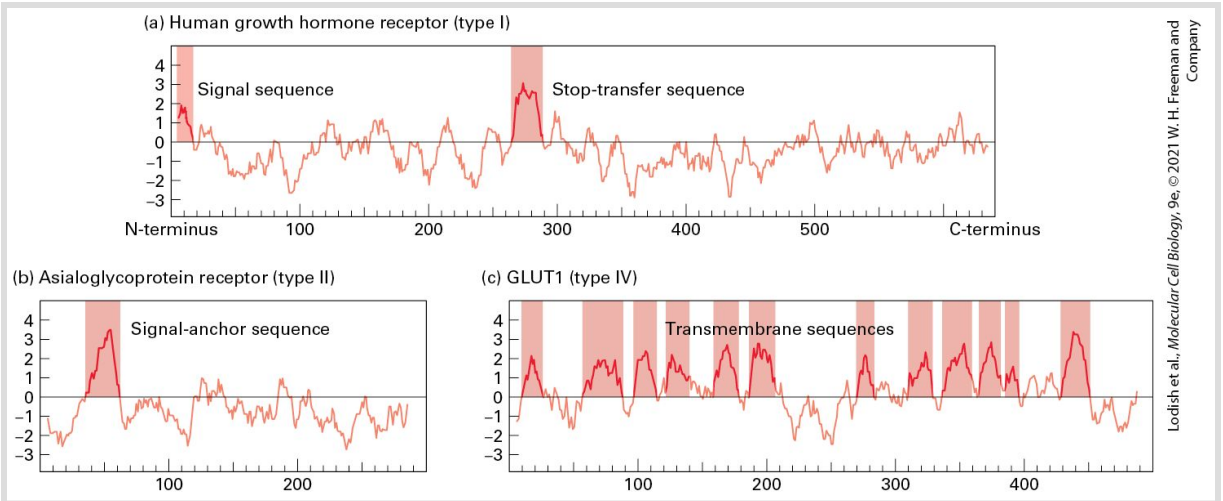

FIGURE 13-16 Hydropathy profiles for three types of proteins. Hydropathy profiles can identify likely topogenic sequences in integral membrane proteins. They are generated by plotting the total hydrophobicity of each segment of 20 contiguous amino acids along the length of a protein. Positive values indicate relatively hydrophobic portions of the protein; negative values relatively hydrophilic portions. Probable hydrophobic topogenic sequences

are marked. (a) The human growth hormone receptor evidently has two topogenic sequences, a signal sequence at the N-terminus and an internal hydrophobic stop-transfer sequence. Note that the membrane-spanning stop-transfer sequence is significantly more hydrophobic than the signal sequence, which is typically the case. From this arrangement it would be possible to correctly deduce that this is a type I membrane protein with an N-terminal domain that binds the growth hormone. (b) The asialoglycoprotein receptor has a single hydrophobic topogenic sequence. This sequence can be correctly predicted to be a signal-anchor sequence for a type II protein because it is preceded by a hydrophilic segment of 30 amino acids. As a type II protein, the C-terminal domain is predicted to be the asialoglycoprotein binding domain. The complex profiles for multipass (type IV) proteins, such as GLUT1 in part (c), must often be supplemented with other analyses to determine the topology of these proteins.

Description

The vertical axis of all three graphs ranges from negative 3 to 4, in increments of 1. Graph (a) shows the human growth hormone receptor (type 1). The horizontal axis is labeled N-terminus on left end and C-terminus on right end and ranges from 100 to 500, in increments of 100. The curve starts at (0 or N-terminus, 1) and ends at (640 or C-terminus, 0). The curve shows several ups and downs with peaks not going beyond 2 and dips not going below negative 3. A shaded region representing signal sequence ranges from N-terminus to 20 with a peak at 2; and another shaded region representing stop-transfer sequence ranges from 260 to 290 with a peak at 3.

Graph (b) shows the asialoglycoprotein receptor (type 2). The horizontal axis ranges from 0 to 300, in increments of 100. The curve starts at (0, 1) and ends at (290, negative 0.5). The curve shows several ups and downs with peaks not going beyond 1 and dips not going below negative 3. A shaded region representing signal-anchor sequence ranges from 40 to 60 with a peak at 3.5.

Graph (c) shows the GLUT 1 (type 4). The horizontal axis ranges from 0 to 500, in increments of 100. The curve starts at (0, negative 1) and ends at (490, 0.3). The curve shows several ups and downs with 12 high peaks going beyond 2. Twenty sequences around the peak regions are shaded and represent transmembrane sequences.

The hydropathy profile of the GLUT1 glucose transporter, a multipass transmembrane protein, shows the presence of many segments that are sufficiently hydrophobic to be membrane-spanning helices (Figure 13-16c). The complexity of this profile illustrates the difficulty both in predicting the orientation of the first topogenic sequence and in unambiguously identifying all the membrane-spanning segments. 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 of greater than 80 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 hydrophobic transmembrane segments. 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 13.2

Insertion of Membrane Proteins into the ER

- Proteins synthesized on the rough ER include five topological classes of integral membrane proteins as well as a lipid-anchored type protein (see Figure 13-10).
- Topogenic sequences — N-terminal signal sequences, internal stop-transfer anchor sequences, and internal signal-anchor sequences — direct the insertion of nascent proteins into the ER membrane and their orientation within it. This orientation is retained during transport of the completed membrane protein to its final destination — for example, the plasma membrane.
- Single-pass membrane proteins that are initiated by interaction with SRP and are synthesized on ribosomes bound to the translocon are of three topological varieties: type I, type II, and type III. The topological type depends on the presence or absence of a signal sequence, the length of the hydrophilic N-terminal segment, and the arrangement of positively charged residues adjacent to the hydrophobic transmembrane segment (see Figure 13-14).
- Tail-anchored proteins that have a single C-terminal transmembrane domain are inserted into the membrane post-translationally. Their insertion depends on a cytosolic receptor, Get3, which is an ATPase with similarity to the GTPase domains of SRP and SRP receptor.
- The topology of multipass (type IV) proteins is established by the orientation of the first transmembrane segment according to the same principles as for single-pass proteins. Each additional transmembrane segment engages the translocon in sequential order as they emerge from the ribosome and thus are inserted in the membrane in alternating orientation.
- Some cell-surface proteins are initially synthesized as type I proteins, but then cleaved, and their luminal domains transferred to a GPI anchor (see Figure 13-15).
- 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 13-16).