Dynamic Insertion of Membrane Proteins at the Endoplasmic Reticulum

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Marie Higy

Aus Colmar, Frankreich

Basel, 2005

Genehmigt von der Philosophisch-
Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Martin Spiess
Prof. Hans-Peter Hauri

Basel, den 6. April 2005

Prof. Dr. Hans-Jakob Wirz
Dekan der Philosophisch-
Naturwissenschaftlichen
Fakultät

Summary

Most eukaryotic membrane proteins are cotranslationally integrated into the endoplasmic reticulum membrane by the Sec61 translocation complex. They are targeted to the translocon by hydrophobic signal sequences which induce the translocation of either their N- or C-terminal sequence. Signal sequence orientation is largely determined by charged residues flanking the apolar sequence (the positive-inside rule), folding properties of the N-terminal segment, and the hydrophobicity of the signal. Recent in vivo experiments suggest that N-terminal signals initially insert into the translocon head-on to yield a translocated N-terminus. Driven by a local electrical potential, the signal may invert its orientation and translocate the C-terminal sequence. Increased hydrophobicity slows down inversion by stabilizing the initial bound state. In vitro crosslinking studies indicate that signals rapidly contact lipids upon entering the translocon. Together with the recent crystal structure of the homologous SecYEβ translocation complex of *Methanococcus jannaschii*, which did not reveal an obvious hydrophobic binding site for signals within the pore, a model emerges in which the translocon allows the lateral partitioning of hydrophobic segments between the aqueous pore and the lipid membrane. Signals may return into the pore for reorientation until translation is terminated. Subsequent transmembrane segments in multispanning proteins behave similarly and contribute to the overall topology of the protein. This thesis was aimed at investigating the integration of single- and double-spanning membrane proteins in mammalian cells. The first part consisted of probing the environment of the signal while its orientation is determined by inserting different hydrophobic residues at various positions throughout a uniform oligo-leucine signal sequence. The resulting topologies revealed a strikingly symmetric position dependence specifically for bulky aromatic amino acids, reflecting the structure of a lipid bilayer. The results support the model that during topogenesis in vivo the signal sequence is exposed to the lipid membrane. The second part consisted of the determination of the kinetics of double-spanning protein topogenesis. The results confirmed that major reorientation of the polypeptide my occur when a second topogenic sequence, conflicting with a first one, enters the translocon. They also showed that the time window for protein reorientation differs for different types of substrate.

Table of Contents

INTRODUCTION

The eukaryotic cell (Figure 1) shows an extraordinary degree of organizational complexity. Macromolecular components that carry out different metabolic processes are segregated in distinct subcellular compartments and these must act in concert to sustain the various cellular functions.

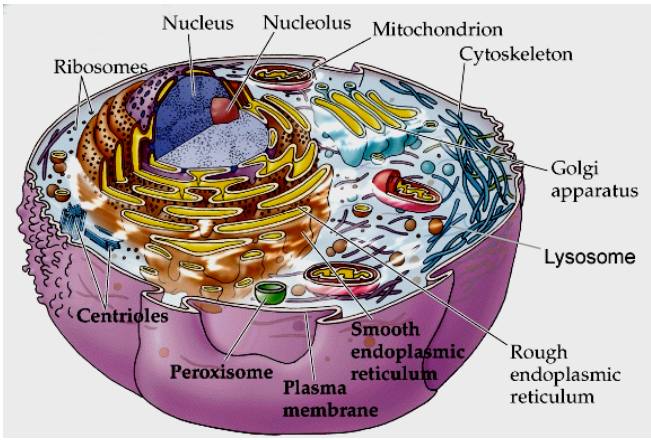

Figure 1: Schematic representation of an eukaryotic cell and its organelles. (Mod. from 1998 Sinaver Associates Inc.)

With the exception of few mitochondrial and chloroplast proteins, all proteins are synthesized on ribosomes in the cytoplasm and are then sorted and transported to the specific compartment of their function. There are three fundamentally different mechanisms of protein transport and sorting to a specific organelle in the cell:

- Transport via a large aqueous passage without the need to cross a membrane. This mechanism is exemplified by import into the nucleus via the nuclear pore complex.
- Transport through or insertion into a lipid bilayer. Of all organelles in the cell, only the endoplasmic reticulum (ER), peroxisomes, and mitochondria (and in plants chloroplasts) possess a general machinery to import proteins through their membranes or to integrate proteins into their membranes.
- Membrane transport. All compartments other than the ER, peroxisomes or mitochondria receive their proteins directly or indirectly from the ER together with lipid membranes, typically by vesicular transport.

Below, these three fundamental transport and sorting mechanisms will be briefly described, with emphasise on the transport through membranes, leading up to the mechanism of protein insertion into the lipid bilayer and membrane protein topogenesis at the ER.

Nuclear import/export

The nucleus, the largest organelle in eukaryotic cells, is surrounded by two membranes. The outer nuclear membrane is continuous with the rough endoplasmic reticulum, and the space between the inner and outer nuclear membranes is continuous with the lumen of the rough endoplasmic reticulum. Embedded in the double membrane of the nuclear envelope (NE), the nuclear pore complex (NPC) connects the cytoplasm and the nucleus of interphase eukaryotic cells (Figure 2). The NPC allows the free diffusion of ions and small molecules across the NE, and it facilitates receptor-mediated nucleocytoplasmic transport of proteins, RNAs and ribonucleoprotein particles.

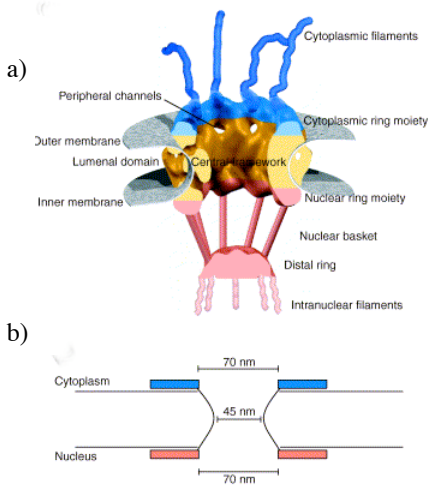

Figure 2: Nuclear pore complex (NPC) architecture and linear dimensions of its central pore. **a)** Consensus model of the 3D architecture of the NPC. **b)** Linear dimensions of the central pore of the NPC. Blue boxes represent the cytoplasmic ring moiety of the NPC, orange boxes represent the nuclear ring moiety of the NPC (Fahrenkrog et al., 2004).

The NPC is composed of a set of proteins that are collectively termed nucleoporins. In both yeast and mammals, the NPC comprises about 30 different nucleoporins (Rout et al., 2000; Cronshaw et al., 2002). Many nucleoporins contain distinct domains of phenylalanine-glycine (FG) repeats, which mediate the main interaction between nucleoporins and soluble transport receptors.

Although the actual translocation mechanism is understood only poorly, the interaction of transport receptors with FG-repeat domains of nucleoporins seems to be essential. On the basis of the affinity and the nature of the interaction, different translocation models have been proposed: the Brownian affinity gating model or virtual gating model (Rout et al., 2000; Cronshaw et al., 2002), the selective phase model (Ribbeck and Gorlich, 2001), and the oily-spaghetti model (Macara, 2001).

Nuclear transport is mediated by short sequence elements in cargo molecules: cargo carrying a nuclear localization sequence (NLS) is imported, whereas a nuclear export sequence (NES) is used for export. The NLS are generally composed of one (monopartite NLS) or two (bipartite NLS) short stretches of basic residues exposed on the surface of the folded protein (Hodel et al., 2001). The monopartite NLS has the core consensus sequence K(K/R)X(K/R), though NLS sequences tend to vary quite a lot (Cokol et al., 2000). These sequence elements are recognized by transport factors, collectively termed karyopherins (also referred to as importins, exportins or transportins), which ferry the cargo.

The karyopherins (Kap) bind to the NLS of a cargo protein, and function only in nuclear import The assembly and disassembly of cargo–kap complexes is governed by Ran, a G protein that comes in two conformationally distinct states depending on whether it is bound to the nucleotide guanosine triphosphate (GTP) or to guanosine diphosphate (GDP) (Scheffzek et al., 1995; Vetter et al., 1999).

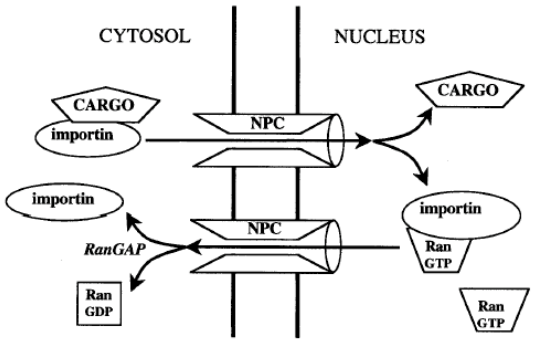

Figure 3: Mechanism of cargo import by direct interaction with an importin carrier protein (Macara, 2001).

Directionality and energetic of transport are provided by an asymmetric distribution of GEFs, GAPs in the nucleus and the cytoplasm (Figure 3). RanGTP is present at high concentrations only in the nucleus, where it disassembles the cargo-importin complex. The importin-RanGTP complex returns to the cytoplasm, where the GTP is hydrolyzed, releasing the RanGDP from the importin.

The nuclear translocation itself can be compared to a facilitated diffusion through the NPC, but experimental data that directly address this mechanism at a molecular level are lacking.

Membrane transport

Description of compartments and pathways

Eukaryotic cells possess an elaborate endomembrane system that makes up the secretory pathway and endocytosis. This network consists of a number of interdependent organelles that function sequentially to effect protein secretion to the extracellular environment, uptake of macromolecules from outside the cell and protein sorting. Each compartment provides a specialized environment that facilitates the various stages in protein biogenesis, modification, sorting, and, ultimately secretion.

The organelles of the secretory pathway are involved in the sorting of proteins to a variety of intracellular membrane compartments and to the cell surface. For example, proteins that are transported within the secretory pathway are either secreted from the cell, integrated into the plasma membrane, sorted to lysosomes, or are retained as "residents" in any of the organelles.

Within the lumen of the endoplasmic reticulum (ER) are a number of chaperones that bind to the polypeptide chain and assist the protein in forming the correct conformation. These chaperones include BiP (Gething, 1999), the lectins calnexin and calreticulin (Freedman, 1994; Zhang et al., 1997). Chaperones assist the nascent polypeptide chain protein by slowing folding, preventing aggregation, and ensuring that the correct disulfide bonds are formed. The chaperone proteins

also have a role in "quality control", directing malformed proteins back through the translocon and to the proteaosome for degradation (Lord et al., 2000).

The majority of plasma membrane and secretory proteins are glycosylated. N-linked oligosaccharides are added to the growing polypeptide chain as it enters the ER (Kornfeld and Kornfeld, 1985). After glycosylation, ER-derived transport vesicles subsequently fuse to form a network of vesicular tubular clusters (VTCs) (Bannykh et al., 1996) also known as pre-Golgi intermediates (Saraste and Kuismanen, 1992) or the ER-Golgi intermediate compartment (ERGIC) (Hauri and Schweizer, 1992).

The ERGIC was identified as a compartment in which cargo accumulates during a block in traffic from the ER to the Golgi at 15°C (Saraste and Kuismanen, 1984; Schweizer et al., 1990). The ERGIC compartment is a major sorting station, recycling ER proteins in retrograde vesicles as well as delivering secretory cargo to the *cis*-Golgi (Warren and Mellman, 1999). Peripheral ERGICs move along microtubules to the Golgi region (Presley et al., 1997; Scales et al., 1997) where they fuse to form the *cis*-Golgi network, (Saraste and Kuismanen, 1992; Presley et al., 1997), a compartment which contains the *cis* most Golgi cisternae (Ladinsky et al., 1999). Secretory proteins are then transported through the Golgi cisternae to the *trans*-Golgi network (TGN), or Golgi exit site.

The Golgi apparatus consists of a series of flattened membrane cisternae, called the Golgi stack, bordered by two tubulo-vesicular networks, the *cis*-Golgi network (CGN) and the *trans*-Golgi network (TGN). Long membrane tubules interconnect multiple Golgi stacks, which are arranged around the nucleus, close to the centrosome (Thyberg and Moskalewski, 1999). The *cis* face of the Golgi receives cargo from the ER, which is transported through the Golgi stack to the TGN.

Further evidence of compartmentalization of the Golgi stack was clearly evident from studies on the location of the glycosylation machinery. Golgi glycosyltransferases were found to have distributions that reflected the order of glycosylation events (Kornfeld and Kornfeld, 1985; Rabouille et al., 1995). The Golgi apparatus also plays a role in lipid biosynthesis with increasing concentrations of cholesterol and sphingolipids present across the stack (Rabouille et al., 1995).

At the TGN, proteins are sorted according to their final destinations. The TGN is also the site where the biosynthetic and endocytic pathways converge. Molecules are internalized from the cell surface in endocytic vesicles and transported to the early endosome where extensive sorting takes place. For example, endocytosed proteins can then be recycled to the plasma membrane (such as recycling receptors), or transported to the TGN or to the lysosome via the late endosomes for degradation. Thus, the TGN and the early endosome represent the two major sorting stations of the cell. Protein transport in the secretory and endocytic pathways is a multi-step process involving the generation of transport carriers loaded with defined sets of cargo, the shipment of the cargo-loaded transport carriers between compartments, and the specific fusion of these transport carriers with a target membrane.

Vesicles made by cytosolic coat proteins

There are a variety of distinct pathways within the secretory system in which protein and lipid cargo can be transported. Each pathway is highly selective for certain cargo. The transport mechanisms that operate from each compartment reflect the requirement to target cargo molecules to specific destinations and yet at the same time maintain the membrane and protein composition of the individual compartments.

For every vesicular trafficking pathway, there are many different types of cargo proteins. Some of these proteins are ligand receptors, such as the LDL and mannose 6-phosphate receptors; the others are proteins which use vesicles to make one-way journeys to their final destination. Proteins use vesicles to get to compartments where they will receive posttranslational modifications, where they will be degraded and where they will find new binding partners or new substrates. Some of the cargo proteins, such as the SNAREs, are essential components of the vesicle fusion machinery. Somehow all of these different types of cargo proteins must be accommodated in the same vesicle. This is the job of the coat proteins.

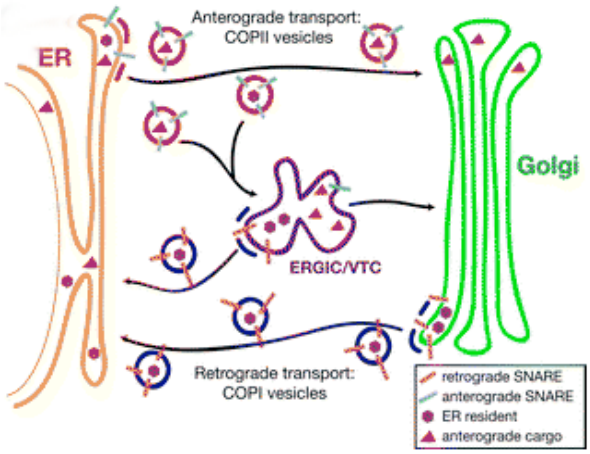

Figure 4: Diagrammatic representation of ER-Golgi transport (Lee et al., 2004).

As represented Figure 4, secretory cargo leaving the ER is packaged into vesicles with a specialized protein coat known as COP (**Co**at **P**rotein) II (Barlowe, 1998) in the early secretory pathway. Following budding from the ER, COPII vesicles loose their coats and fuse to form ERGICs (Aridor et al., 1995). Another COP coat, called COPI, assembles on the membranes of the ERGICs generating vesicles for retrieval of ER residents (Aridor et al., 1995; Scales et al., 1997).

Recent findings indicate that soluble cargo is concentrated by selective removal of ER proteins in retrograde COPI coated vesicles (Martinez-Menarguez et al., 1999).

Proteins are sorted at the TGN for delivery to multiple destinations including: the basolateral and apical plasma membranes; secretory granules; endosomes; and for retrograde transport. Mechanisms for sorting to these different locations are quite distinct but are generally signal dependent (Keller and Simons, 1997). Proteins destined for regulated secretion aggregate in the TGN where they are packaged into immature secretory granules (ISGs) (Thiele et al., 1997).

Clathrin coated vesicles form on these ISGs and recycle TGN proteins. Mature secretory granules are stored in the cytosol until the cell receives a signal to release their contents. Lysosomal enzymes are transported from the TGN in specialized vesicles. These vesicles carry enzymes and their receptors to the early/late endosomes. The clathrin adaptors select cargo for inclusion into clathrin coated vesicles (Figure 5) in the late secretory and endocytic pathways.

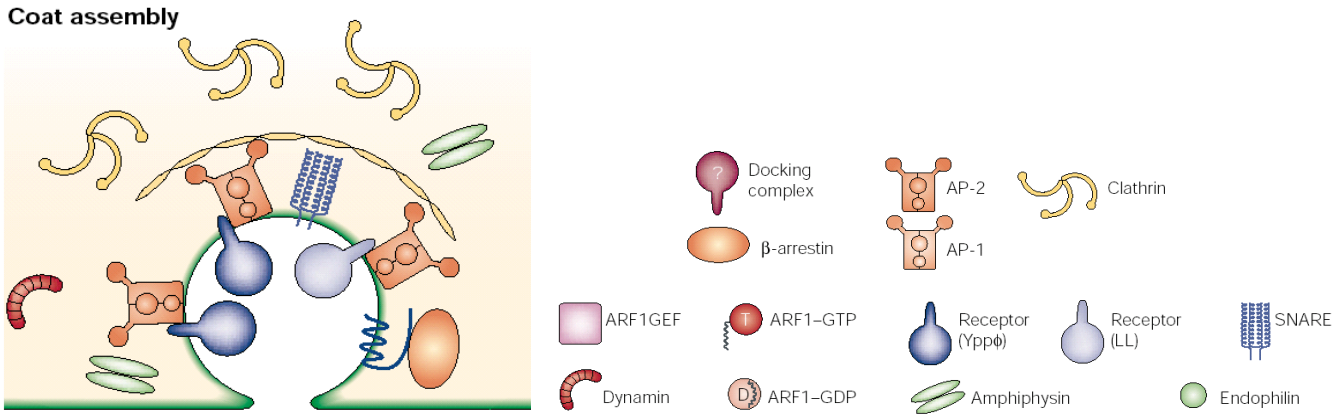

Figure 5: The key steps in the formation of clathrin –coated vesicles. At the TGN network, coat assembly is activated by the recruitment of ARF1 to the membrane. One end of adaptor proteins bind to cargo molecules and the other end to other coat components, including clathrin. Clathrin triskelions polymerise into hexagons and pentagons, forming a cage which leads to membrane deformation. When the coat is almost complete, dynamin pinches off the vesicle. Uncoating requires ATP hydrolysis by Hsc 70 and auxilin (Kirchhausen, 2000).

The first coated vesicle adaptors to be identified were the AP-1 and AP-2 complexes. Both of these adaptors are highly enriched in purified clathrin-coated vesicles, second in abundance only to clathrin itself, and they both promote clathrin assembly *in vitro*. (The name AP was originally introduced as an acronym for assembly polypeptides (Zaremba and Keen, 1983), although conveniently it also stands for 'adaptor protein'). They localize to different membranes: AP-1 is found on the TGN and endosomes, whereas AP-2 is found at the plasma membrane. Two additional adaptor complexes, AP-3 and AP-4, were discovered by searching sequence databases for homologs of the AP-1 and AP-2 subunits (Robinson and Bonifacino, 2001). Like AP-1, AP-3 and AP-4 are found on TGN/endosomal membranes, with AP-3 localized more to endosomes and AP-4 more to the TGN . Both appear to be able to function independently of clathrin.

Another family of clathrin adaptors, the GGAs (Golgi-localized, γ-ear-containing, ARF-binding proteins), was identified more recently, again by searching databases for AP subunit homologs (Boman et al., 2000; Dell'Angelica et al., 2000; Hirst et al., 2000). The GGAs are monomeric and function together with clathrin in both mammalian cells and yeast. They are found on TGN and endosomal membranes, but reports vary as to this distribution between the two (Meyer et al., 2001; Valdivia et al., 2002; Doray et al., 2002; Lui et al., 2003).

Signals for cargo sorting

The localization of proteins to specific organelles within eukaryotic cells depends on discrete targeting signals contained within these proteins. The targeting signals for a number of resident proteins have been identified. However, in many cases the underlying mechanisms responsible for the recognition of these signals remain not totally defined.

By far the best-characterized adaptor–cargo interaction is the one between the subunits of AP complexes and the sorting signal YXXΦ (where Φ is a bulky hydrophobic residue). It binds to the μ-subunit of AP. This signal can act both as an internalization signal at the plasma membrane (e.g. in the transferrin receptor) and as an intracellular sorting signal (e.g. in lysosomal membrane proteins such as LAMP-1).

There is another type of adaptor-dependent sorting signal, which consists of a pair of leucine residues preceded by one or more acidic residues. Again, there are at least two such motifs, which bind to different partners. The first dileucine signal to be described has the consensus sequence [D/E]XXXL[L/I] and is found mainly in proteins resident in late endosomes, lysosomes and lysosome-related organelles such as melanosomes.

A second type of dileucine motif has the consensus sequence DXXLL. This motif is found in both the cation-independent and cation-dependent mannose 6-phosphate receptors for lysosomal enzymes, and it has been shown to be the major sorting determinant for the cation-independent receptor (Chen et al., 1997).

Localization of ER proteins is generally achieved through two complementary mechanisms: retention and retrieval (Teasdale and Jackson, 1996). Particular sequences within ER proteins function to retain the protein in the ER, while additional sequences are required to retrieve escaped ER proteins from downstream compartments. Soluble ER proteins contain the carboxy-terminal tetrapeptide, -KDEL or -HDEL, which is necessary for their retrieval from the Golgi (Munro and Pelham, 1987).

The KDEL receptor (Erd2p) is packaged into COPI-coated vesicles at the Golgi. Like other ER type I transmembrane proteins, Erd2p contains a di-lysine retrieval (Lewis and Pelham, 1990; Semenza et al., 1990). The K(X)KXX

consensus sequence interacts with the COPI coat (Cosson and Letourneur, 1994; Letourneur et al., 1994). Thus, Erd2p binds KDEL tagged proteins and returns them to the ER in COPI vesicles. In addition to its role as a retrieval signal, the di-lysine motif in other membrane proteins may also act as a retention signal. Chimeras containing KKAA tags are actively retained in the ER and this localization is not affected in COPI mutant cell lines (Andersson et al., 1999).

Transport across and into membranes

General

The three main compartments competent to accept proteins are the ER, the mitochondria and the peroxisomes. To import the proteins to the correct organelles the proteins possess a signal in their sequence, which is recognized in the cytosol by specific receptors.

C. Milstein (1972) was the first to experimentally identify a signal peptide. He concluded from his studies that cytoplasmic proteins are made on free polysomes and in contrast, secretory proteins are synthetized on microsomes. The machinery whereby this segregation is achieved was indentified to a large part by G. Blobel (Nobel laureate in Physiology or Medecine in 1999) who discovered that proteins possess intrinsic signals that govern their localization in the cell (Sabatini et al., 1971). The concept that emerged for protein import was found to be largely conserved for the other organelles as well, and in a very basic way for all sorting mechanisms in the cell.

Peroxisome

All animal cells (except erythrocytes) and many plant cells contain peroxisomes (Figure 6), a class of small organelles bounded by a single membrane. Peroxisomes contain several oxidase*s* — enzymes that use molecular oxygen to oxidize organic substances, in the process forming hydrogen peroxide (H_2O_2), a corrosive substance. In most eukaryotic cells, the peroxisome is the principal organelle in which fatty acids are oxidized, thereby generating precursors for important biosynthetic pathways.

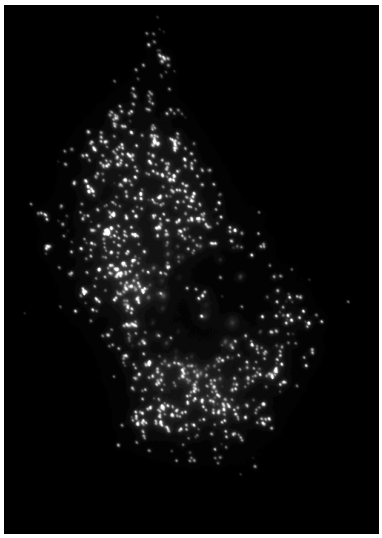

Figure 6: Peroxisomes labeling in COS-1 cell transfected with PST1-GFP (M. Higy).

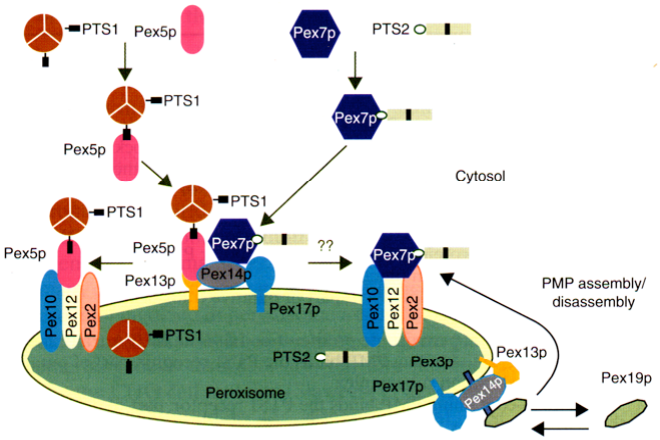

Figure 7: Import machinery of the peroxisome. (Protein Targeting transport and translocation, RE Dalbey and G Von Heijne).

The peroxisome biogenesis involves “peroxins”, proteins encoded by PEX genes (Figure 7). These peroxins are involved in three key stages of peroxisome

development, the import of peroxisomal membrane proteins; the import of peroxisomal matrix proteins and the peroxisome proliferation. Peroxisomes use a posttranslational import pathway, and is the only organelle which can import folded proteins. Proteins which are destined for the peroxisomal matrix or membrane possess distinct targeting signals that engage different signal sequence receptors to drive their transport to their final subcellular destination.

The first class of peroxisome targeting signals (PTSs), PTS1, was discovered in 1987 in firefly luciferase (Gould et al., 1987). It consists of a conserved C-terminal tripeptide SKL or some variants, and is necessary and sufficient for protein targeting into the peroxisome matrix. PTS1 is the most frequent signal involved in protein transport to peroxisome matrix. The second class, PTS2, consist of a nonapeptide located in the N-terminal part of the proteins (Osumi et al., 1991; Swinkels et al., 1991). Its consensus sequence is $XXRLX_5HL$ and some variants.

Peroxisomal membrane proteins use membrane PTSs (mPTSs) for their targeting which have little in common with the other PTSs, except a basic region (Dyer et al., 1996; Baerends et al., 2000; Subramani et al., 2000).The studies of the import of proteins into the peroxisome started in the late 1980s and showed that PTSs are conserved from yeast to human (Gould et al., 1989).

The PTS1 and PTS2 signal are recognized in the cytosol by their respective receptors Pex5p and Pex7p. The complex is then delivered to the peroxisome by interacting with a docking complex (Pex13-14-17, Pex2-10-12) (Subramani et al., 2000). A new theory proposes that the peroxisomal fraction of Pex5p assembles to a large protein complex at the membrane, referred to as importomer. This importomer consists of an oligomirization of several Pex5 which forms a pore in the peroxisomal membrane and so delivers the lumenal protein inside the peroxisome (Schliebs W, data not published). The recycling of Pex5 is an ATP-dependent process (Gouveia et al., 2003) and is mediated by ubiquitination of the receptor (Platta et al., 2004) and depends on the N-terminal half of Pex5p (Gouveia et al., 2003).

For Pex7p, the soluble receptor of the PTS2, recent studies showed that Pex7p is also translocated in and out of the peroxisomes in yeast (Nair et al., 2004).

Mitochondria

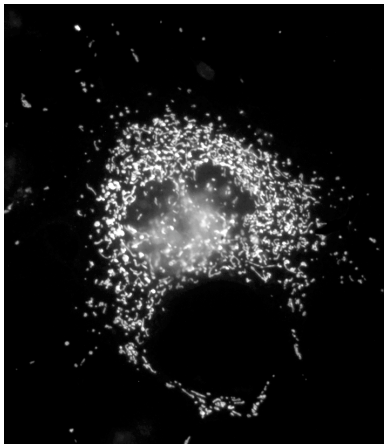

Figure 8: COS-1 cell labeled for mitochondria with antibody against the protein ADP/ATP carrier. (M. Higy).

Mitochondria (Figure 8) consist of two membranes, the outer membrane and the folded inner membrane, and two aqueous compartments, the intermembrane space and the matrix. Mitochondria play crucial roles in cellular energy production and the metabolism of amino acids, iron, and lipids, as well as in apoptosis.

Although mitochondria possess a complete genetic system in the matrix, only about one percent of all mitochondrial proteins are encoded by the mitochondrial genome. The vast majority of mitochondrial proteins are synthesized as precursor proteins in the cytosol and are imported mainly by a posttranslational mechanism (Beddoe and Lithgow, 2002).

The precursor proteins are translocated across the mitochondrial membranes in an unfolded state (Schleyer and Neupert, 1985; Eilers and Schatz, 1986; Rassow et al., 1990). Two processes are at work *in vivo* to minimize aggregation and misfolding of mitochondrial precursor proteins: coupling of translation to translocation, and formation of transient, stabilizing complexes with molecular chaperones and other cytoplasmic factors. Cytosolic chaperones, like 70-kDa heat-shock protein (Hsp70) (Sheffield et al., 1990), Mitochondrial import Stimulation Factor (MSF) (Hachiya et al., 1993), Nascent-Associated polypeptide Complex (NAC) (Wiedmann et al., 1994; Funfschilling and Rospert, 1999), Ribosome-associated complex (RAC) (Gautschi et al., 2001) are involved in guiding the precursor proteins to receptors on the mitochondrial surface.

Mitochondrial precursor proteins can be separated into two main classes (Figure 9). Preproteins which are destined for the mitochondrial matrix, as well as a number of proteins of the inner membrane and intermembrane space, carry N-terminal cleavable extensions, termed presequences. These positively charged extensions function as targeting signals that interact with the mitochondrial import receptors and direct the preproteins across both outer and inner membranes (Schatz and Dobberstein, 1996; Pfanner and Geissler, 2001). The second class of precursor proteins, carrying various internal targeting signals, include all outer membrane proteins along with many intermembrane space and inner membrane proteins. These precursors are synthesized without cleavable extensions with the same primary structure as the mature protein, but their conformation typically differs from the mature protein (Koehler et al., 1999; Pfanner and Geissler, 2001).

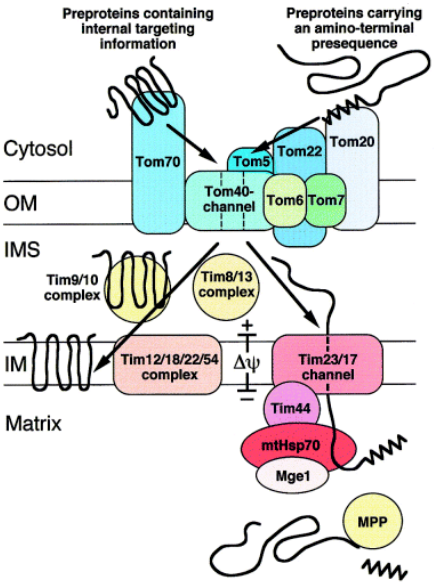

Figure 9: Schematic model of the mitochondrial import machinery. At the outer membrane (OM), preproteins are either targeted to Tom70 or to Tom20. The components of the GIP complex include Tom40, Tom22, Tom 5, Tom6, and Tom7. Transit through the intermembrane space (IMS) is mediated by direct binding to complexes of Tim9 and Tim10. They deliver their cargo to Tim12, a component of a complex in the inner membrane (IM) that comprises Tim22, Tim54, and Tim18. The complex containing Tim23, Tim17, and Tim44 is required for translocation of presequence-carrying preproteins across the inner membrane into the matrix. Tim44 is a hydrophilic protein and serves as an adaptor protein for mitochondrial Hsp70 (mtHsp70). The membrane potential is required for translocation of positively charged presequences through the Tim23/Tim17 channel. Mge1 regulates the interaction of mtHsp70 with nucleotides. The matrix processing peptidase (MPP) cleaves presequences after translocation (Rassow and Pfanner, 2000).

Precursor proteins which are destined for the matrix enter through both mitochondrial membranes. The translocase of the outer mitochondrial membrane, the TOM complex, first transports the signal sequence across the outer membrane. The TOM complex includes import receptors that initially recognize the signal peptide or a signal sequence (these include Tom20 and Tom70). Different proteins

use different receptors. The receptors bring the protein to the region containing the translocator proteins. This is actually a complex of proteins called the **G**eneral **I**mport **P**ore (GIP). This GIP complex consists of Tom40, Tom22 and the three small Tom proteins, Tom7, Tom6 and Tom5 (Kunkele et al., 1998). It facilitates the translocation of the presequence of the protein across the outer membrane. Tom40 appears to be the core element of the pore and forms oligomers (Rapaport and Neupert, 1999).

Once it reaches the intermembrane space, the signal sequence binds to the translocase of the inner membrane, the TIM complex (Tim23-44 complex, Tim22 complex). Insertion of preproteins into the Tim23 channel strictly depends on the presence of the membrane potential across the inner membrane (Truscott et al., 2001).

This opens the channel through which the polypeptide chain enters the matrix or inserts into the inner membrane. Transport of proteins into mitochondria is directional and therefore requires energy. This import is fueled by ATP hydrolysis at two sites: one outside the mitochondrion, when the unfolded precursor protein interacts with the import receptor, and one in the matrix.

Proteins that have to go all the way to the matrix have an NH2 cleavable signal sequence. Many proteins have secondary signals that result in transport to other mitochondrial locations (e.g. inner mitochondrial membrane, inter-membrane space).

Endoplasmic Reticulum (ER)

ER is a compartment comprising a network of interconnected, membrane-bounded tubules (Figure 10).

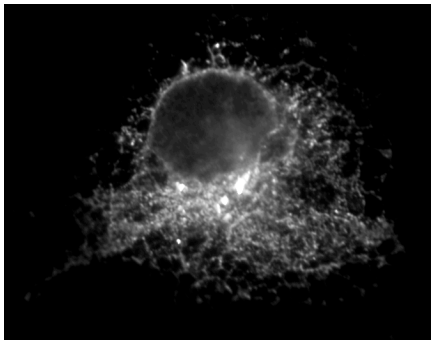

Figure 10: COS-1 cell labeled for ER with an antibody against p53 (N. Beuret).

The synthesis of fatty acids and phospholipids occurs in the ER. The ribosomes that synthesize secretory proteins are bound to the rough ER by the nascent polypeptide chain of the protein. Protein targeting to the ER for secretion or insertion into the membrane in mammalian cells is mostly cotranslational, *i.e.*, it takes place at the same time that they are being synthesized by ribosomes (Walter and Johnson, 1994). Cotranslational transport of secretory and membrane proteins depends on the SRP system, which consists of a cytosolic ribonucleoprotein particle (SRP) and its membrane-bound receptor (SR).

Newly synthesized proteins destined for secretion or membrane insertion carry a hydrophobic signal sequence at their N terminus, a stretch of mainly 7 to 25, mainly apolar residues (Walter and Johnson, 1994). SRP interacts with the signal sequence as soon as it emerges from the ribosomal polypeptide exit tunnel. In eukaryotes, peptide elongation is retarded upon binding of SRP to the ribosome nascent chain complex (RNC). Subsequently, the SRP–RNC complex is targeted to the ER membrane by the interaction with the SR. The RNC is then transferred to the protein-conducting channel in the membrane (the translocon) (Keenan et al., 2001; Koch et al., 2003). The GTP dependency is thought to enable correct targeting by coordinating the presence of a signal sequence on the ribosome with

the availability of a translocon in the membrane (Keenan et al., 2001; Koch et al., 2003).

The proteins captured by the ER are of two types: transmembrane proteins, which are only partly translocated across the ER membrane as they are being synthesized and become imbedded in it; water-soluble proteins, which are fully translocated across the ER membrane as they are synthesized and released into the ER lumen. Although secreted proteins pass through the membrane of the ER into the lumen, proteins that will eventually reside in the membranes of the ER, Golgi, lysosomes or plasma membrane do not. They are "trapped" in the ER membrane as they are synthesized and therefore remain as membrane-associated proteins for their entire lifetime. The difference between proteins that pass all the way through the membrane and those that are retained appears to be the presence in the protein of an internal "stop" sequence, that consists of a series of around 20 hydrophobic amino acids that usually form an α helix in the membrane (it takes about 10-12 amino acids to forms a helix long enough to span the entire lipid bilayer). These amino acids presumably anchor the protein in the membrane by interacting somehow with the lipid bilayer and making further movement of the newly synthesized peptide chain impossible. If the peptide lacks such a stop sequence it passes completely through the pore into the lumen, where the signal peptidase cuts off the amino terminus and releases it from the membrane.

The nascent polypeptide–ribosome complex transfers to the translocation channel. Integration occurs at sites in the ER membrane termed translocons, which are composed of a specific set of membrane proteins (Jensen and Johnson, 1999; Schnell and Hebert, 2003). Simultaneously to this integration, the nascent protein may undergo covalent modifications (*e.g.*, signal sequence cleavage, disulfide bond formation, and *N*-linked glycosylation), folding, and interaction with other proteins (*e.g.* chaperones) that ultimately lead to the assembly of the polypeptide into a functional monomeric or multimeric complex (Johnson and van Waes, 1999; Deutsch, 2003; Schnell and Hebert, 2003).

Membrane protein integration and topogenesis

Higy M, Junne T, Spiess M. Topogenesis of membrane proteins at the endoplasmic reticulum. Biochemistry. 2004 Oct 12;43(40):12716-22. Review.

Topogenesis of membrane proteins at the endoplasmic reticulum†

Marie Higy, Tina Junne and Martin Spiess

Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

† Our own work was supported by grant 31-061579.00 from the Swiss National Science Foundation.

Address for Correspondence:

Dr. Martin Spiess
Biozentrum, University of Basel
Klingelbergstrasse 70
CH-4056 Basel
Switzerland

Phone: +41-61-2672164
FAX: +41-61-2672148

E-mail: Martin.Spiess@unibas.ch

[1]Abbreviations: ER, endoplasmic reticulum; N_{cyt}/C_{exo}, cytoplasmic N- and exoplasmic C-terminus; SRP, Signal recognition particle; TRAM, translocating chain-associated membrane protein.

ABSTRACT:

Most eukaryotic membrane proteins are cotranslationally integrated into the endoplasmic reticulum membrane by the Sec61 translocation complex. They are targeted to the translocon by hydrophobic signal sequences which induce the translocation of either their N- or C-terminal sequence. Signal sequence orientation is largely determined by charged residues flanking the apolar sequence (the positive-inside rule), folding properties of the N-terminal segment, and the hydrophobicity of the signal. Recent in vivo experiments suggest that N-terminal signals initially insert into the translocon head-on to yield a translocated N-terminus. Driven by a local electrical potential, the signal may invert its orientation and translocate the C-terminal sequence. Increased hydrophobicity slows down inversion by stabilizing the initial bound state. In vitro crosslinking studies indicate that signals rapidly contact lipids upon entering the translocon. Together with the recent crystal structure of the homologous SecYEb translocation complex of *Methanococcus jannaschii* which did not reveal an obvious hydrophobic binding site for signals within the pore, a model emerges in which the translocon allows the lateral partitioning of hydrophobic segments between the aqueous pore and the lipid membrane. Signals may return into the pore for reorientation until translation is terminated. Subsequent transmembrane segments in multispanning proteins behave similarly and contribute to the overall topology of the protein.

Few membranes in a eukaryotic cell are competent to translocate and integrate proteins synthesized by the ribosomes in the cytoplasm: mitochondria, chloroplasts, peroxisomes, and – most prominently – the endoplasmic reticulum (ER)[1]. The ER serves as the gateway for proteins destined for all compartments of the secretory pathway, for the plasma membrane and the cell exterior, as well as for the endocytic organelles. The signals for protein targeting to the ER are highly degenerate. Their essence is an uncharged, predominantly hydrophobic stretch of 7–25 amino acids (von Heijne, 1990). They are not only important for targeting to the ER membrane, but play a role in protein topogenesis. In some cases, they also anchor the polypeptide as a transmembrane domain and assemble into helix-bundles, contributing to the structure and function of complex membrane proteins.

In cotranslational targeting, which is the predominant mode of ER sorting in mammalian cells, a signal sequence is first recognized by signal recognition particle (SRP). As it emerges from the ribosome, it binds to a hydrophobic groove or saddle created by a cluster of methionines on the 54-kD subunit (SRP54) (Zopf et al., 1990; Lutcke et al., 1992; Keenan et al., 1998). The ribosome–nascent chain–SRP complex is directed to the ER membrane by interaction with the SRP receptor (Keenan et al., 2001). Both SRP and SRP receptor are GTPases which interact in a unique manner by forming a shared catalytic chamber for the two GTP nucleotides (Egea et al., 2004; Focia et al., 2004). Reciprocal GTPase activation upon the release of the signal from SRP triggers disassembly of the targeting complex. The ribosome docks onto the translocon aligning the ribosomal exit tunnel with the protein-conducting channel (Menetret et al., 2000; Beckmann et al., 2001). The signal enters the translocon and is oriented with respect to the membrane to initiate translocation of its N- or C-terminal sequence across the membrane. The respective hydrophilic portion of the polypeptide is transferred through the channel into the ER lumen and the signal is released laterally into the lipid bilayer. Additional hydrophobic segments may stop or re-initiate protein transfer and integrate as transmembrane domains into the membrane to generate multi-spanning helix-bundle proteins. These processes determine the topology of proteins in the lipid bilayer.

In the thirty years since the discovery of ER signals (Blobel and Dobberstein, 1975a; Blobel and Dobberstein, 1975b), determinants of protein topology have

been characterized by mutagenesis of substrate proteins (Goder and Spiess, 2001) and the components of the translocation machinery have been discovered by genetic and biochemical studies (Johnson and van Waes, 1999). Sophisticated crosslinking experiments identified molecules in contact with various parts of substrate proteins (Martoglio and Dobberstein, 1996). These mostly static data together with the recent first crystal structure of a protein-conducting channel (Van den Berg et al., 2004) lead to new insights into the highly dynamic process of protein topogenesis and membrane integration.

Orienting signal sequences in the membrane

In secretory and single-spanning membrane proteins, topology is determined by the orientation of the signal sequence in the membrane.

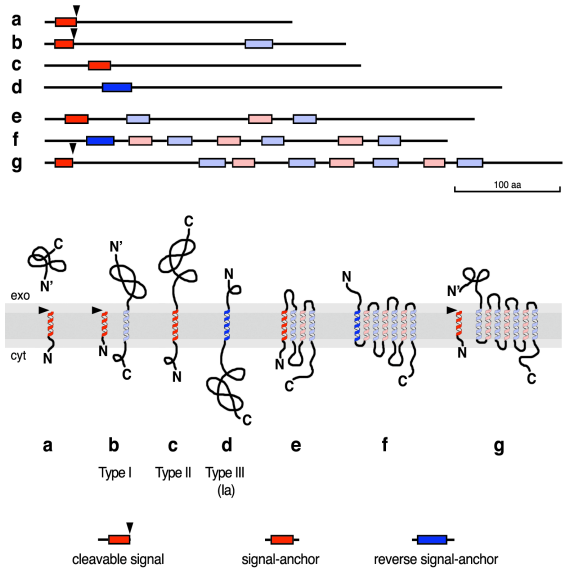

FIGURE 1: Three types of signals initiate cotranslational protein topogenesis. Cleavable signals (red with arrowhead indicating the signal peptidase cleavage site) and uncleaved signal-anchor sequences (red without arrowhead) induce translocation of the C-terminal sequence and assume an N_{cyt}/C_{exo} orientation. Reverse signal-anchors (blue) insert with the opposite N_{exo}/C_{cyt} orientation and translocate their N-terminus. More complex topologies are produced by the combination of the signal with additional transmembrane segments inserting in alternating orientations (light red for N_{cyt}/C_{exo} and light blue for N_{cyt}/C_{exo}). The distribution of hydrophobic signal and transmembrane segments and their orientation in the membrane are shown for a secretory protein (a; preprolactin), a type I membrane protein (d; cation-dependent mannose-6-phosphate receptor), a type II membrane protein (b; asialoglycoprotein receptor), a type III membrane protein (c; synaptotagmin I), and of examples of multispanning membrane proteins with corresponding initial signal sequences (e, gap junction protein a6; f, vasopressin receptor V2; g, glucagon receptor).

Cleavable signals of secretory proteins (Figure 1, a) or type I membrane protein (b) initiate translocation of their C-terminal sequence. Signal peptidase cleaves off these signals and generates new lumenal N-termini (Paetzel et al., 2002). Signal-anchors of type II membrane proteins (c) similarly translocate their C-terminus.

They are not necessarily at the very N-terminus of the protein, remain uncleaved, and have a longer apolar segment to span the hydrophobic core of the bilayer with an N_{cyt}/C_{exo} orientation (cytoplasmic N-terminus, exoplasmic C-terminus) in the completed protein. Very likely, even cleaved signals integrate into the lipid membrane. Signal peptide peptidase, an intramembrane protease, was shown to process them within the membrane producing soluble fragments with potential signalling function (Martoglio et al., 1997; Weihofen et al., 2002; Lemberg and Martoglio, 2004). In contrast to cleavable signals and signal-anchors, reverse signal-anchors of type III proteins (also classified as type Ia) insert with an N_{exo}/C_{cyt} orientation and induce translocation of the N-terminus (Figure 1, d).

Several factors have been shown to determine the orientation of the signal in the membrane. Most prominently, charged residues flanking the hydrophobic core of the signal influence orientation: the more positive end is generally cytosolic, a phenomenon known as the "positive-inside rule" (Hartmann et al., 1989; Beltzer et al., 1991; Parks and Lamb, 1991). Since there is no general electrical potential across the ER membrane, local charges at the translocation apparatus must be involved in orienting the signal sequence (see below). In addition, folding of hydrophilic sequences N-terminal to a signal sterically hinders N-terminal translocation irrespective of the flanking charges (Denzer et al., 1995). The polypeptide needs to be unfolded to be transferred through the translocation channel. A third determinant is the hydrophobicity of the core of the signal sequence (the h-domain) itself. Strongly hydrophobic signals were observed to insert with N_{exo}/C_{cyt} orientation even when the flanking charges were more positive at the N-terminus (Sakaguchi et al., 1992; Wahlberg and Spiess, 1997; Rosch et al., 2000). How hydrophobicity exerts its topogenic effect is less obvious.

The mechanism by which hydrophobicity affects signal orientation was explained in a recent in vivo study by Goder and Spiess (Goder and Spiess, 2003). An N-terminal signal-anchor with a generic h-domain of 22 leucine residues

inserted with mixed orientations despite a positive N-terminus. Surprisingly, the topology depended on the total length of the protein: the fraction of polypeptides with an N_{cyt}/C_{exo} orientation was lowest for a short protein and increased up to ~300 residues following the signal sequence. This result indicated that N-terminal signals initially insert to yield an N_{exo}/C_{cyt} orientation (schematically shown in Figure 2).

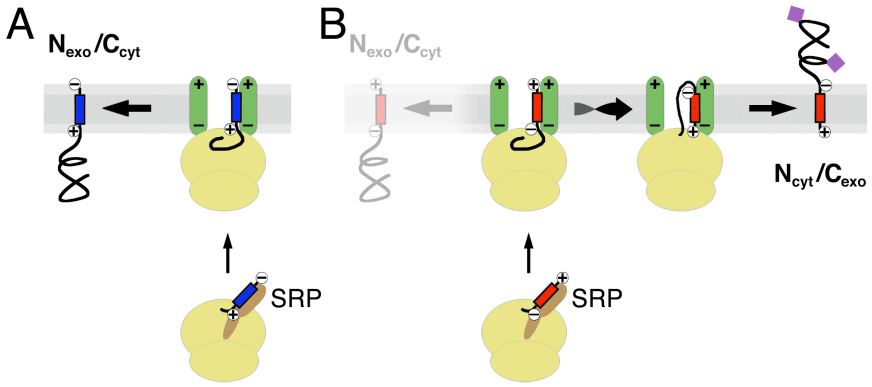

FIGURE 2: Insertion mechanism for N-terminal signal-anchor and reverse signal-anchor sequences. Reverse signal-anchor (panel A; blue) and signal-anchor sequences (panel B; red) initially insert into the translocation apparatus to yield an N_{exo}/C_{cyt} orientation (Goder and Spiess, 2003). The charge distribution of signal-anchors (positive N-terminus, negative C-terminus) drives their inversion to an N_{cyt}/C_{exo} orientation resulting in C-terminal translocation (and potentially glycosylation, shown by purple diamonds). Inversion is accelerated by increased charge difference (ΔN–C), slowed down by increased hydrophobicity of the signal core, and stopped upon translation termination or after ~50 s (Goder and Spiess, 2003). Signal-anchors that have not inverted when reorientation is blocked result in products with the "wrong" N_{exo}/C_{cyt} orientation (grayed-out portion). The process is illustrated schematically. The SRP receptor was omitted for simplicity.

Driven by electrostatic forces, a signal-anchor is inverted until protein synthesis is completed or until further reorientation is stopped after approximately 40–50 s by an as yet unknown mechanism. Increased N-terminal positive charge accelerated the kinetics of signal inversion, whereas reduced charge slowed them down. Increased hydrophobicity of the h-domain, however, diminished the rate of inversion, whereas reduced hydrophobicity accelerated it. Thirteen or more consecutive leucines were necessary to trap a fraction of the polypeptides before they had oriented themselves according to their flanking charges. Most or all natural cleavable signals and signal-anchors are less hydrophobic and thus invert within seconds, long before translation is completed. Hydrophobicity inhibits

inversion because the signal cannot reorient when bound at the translocation apparatus, but only when dissociated. The more hydrophobic the signal, the higher the affinity to the bound state and the lower the inversion rate.

The Sec61 complex: the gateway through and into the membrane

The machinery for the translocation and insertion of proteins into the membrane had first been identified genetically in yeast as Sec61p, a membrane protein with 10 transmembrane domains (Deshaies and Schekman, 1987; Stirling et al., 1992). The mammalian homolog, Sec61α, was found to be part of a complex with two smaller components Sec61β (Sbh1p in yeast) and Sec61γ (Sss1p), which span the membrane only once (Gorlich et al., 1992; Gorlich and Rapoport, 1993). Sec61αβγ is evolutionarily homologous to the bacterial translocation complex SecYEG (Schatz and Dobberstein, 1996). The Sec61 complex is sufficient for translocation and membrane integration of some proteins in reconstituted liposomes, whereas others require an additional component, the translocating chain-associated membrane protein (TRAM), which spans the membrane 8 times (Gorlich and Rapoport, 1993). Although not necessary in the minimal reconstituted system, the lumenal chaperone BiP/Kar2p and the ER membrane protein Sec63p are required for cotranslational translocation in yeast (Brodsky et al., 1995). Sec63p is part of a complex with Sec62p, Sec71p, and Sec72p which together with the Sec61 complex constitutes the machinery for posttranslational translocation in yeast (Panzner et al., 1995). Sec63p contains a J domain that recruits BiP to the lumenal exit site of the translocon. BiP binds to translocating polypeptides and, acting as a molecular ratchet, drives translocation (Matlack et al., 1999). Probably, this mechanism is also functional for efficient cotranslational translocation. In general, binding of chaperones to polypeptide segments emerging into the ER lumen is likely to trap them there and fix the topology of the protein accordingly.

To characterize the machinery components and the environment of nascent polypeptides at various stages of translocation and insertion, photocrosslinking techniques proved extremely powerful (Martoglio and Dobberstein, 1996). Photoreactive probes were incorporated by *in vitro* translation at defined positions

into nascent chains programmed by truncated mRNAs. Upon photolysis, molecules in close proximity to the probe were crosslinked to the arrested polypeptides. As the signal emerged from the ribosome, it was found in contact with SRP54 (Krieg et al., 1986; Kurzchalia et al., 1986). Upon docking of the ribosome–nascent chain complex to the ER membrane, cleavable signals, signal-anchors and reverse signal-anchors were all crosslinked to Sec61αα (High et al., 1991). The N-terminus of a cleavable signal was shown to contact TRAM (Mothes et al., 1994). The hydrophilic polypeptide arrested in translocation was also found close to Sec61α (Mothes et al., 1994) while in an aqueous environment (Crowley et al., 1994), indicating that Sec61α is the major component forming the protein-conducting channel through the bilayer. A stop-transfer sequence was crosslinked to Sec61α and TRAM in what appeared an ordered succession of proteinaceaous environments with increasing polypeptide length (Do et al., 1996). This led to the model of specific proteinaceaous binding sites for signal and transmembrane domains in the translocation apparatus.

However, the signal in an arrested nascent chain that was just long enough to reach into the translocon was found to contact not only Sec61α, but also lipids (Martoglio et al., 1995; Mothes et al., 1997). It was proposed that the signal might be bound at the lateral exit site of the channel, simultaneously exposed to the aqueous pore, Sec61α, and lipids. Crosslinking patterns obtained with a reactive side chain in different positions in the h-domain suggested that the signal was in a helical conformation in stable contact with transmembrane helices 2 and 7 of Sec61α on one side and with lipid on the other (Mothes et al., 1998). However, these findings are also consistent with the signal leaving rapidly into the lipid bilayer upon entering the translocon. Because it is still tethered to the translocon, crosslinking to Sec61α persists. Position-dependent crosslinking may reflect preferred contact surfaces on the outside of the translocation complex. Indeed, stop-transfer sequences were similarly found to crosslink to Sec61α and lipid as soon as they extended into the channel (Mothes et al., 1998; Heinrich et al., 2000). Different sequences were detected in different positions, in some cases adjacent also to TRAM, suggesting that transmembrane segments tethered to the translocation complex associate at various places to the outside of the pore complex (McCormick et al., 2003). It should also be considered that experiments with

arrested nascent polypeptides do not truly represent a time course. In vivo experiments suggested that orientation of a signal-anchor is terminated at the latest approximately 50 s after the signal emerged from the ribosome, even if translation is not yet completed (Goder and Spiess, 2003). This period has certainly passed by the time of in vitro crosslinking. Crosslinking results are thus likely to reflect the situation of the signal after it has left the translocation pore for the lipid membrane, while still closely connected to the translocon via the nascent chain.

Recently, the crystal structure of the SecYEG translocation complex of *Methanococcus jannaschii* has been determined Van den Berg (2004). It suggests that a translocation pore is formed by a single SecYEG complex rather than by three or four complexes as previously proposed based on electron microscopy of the yeast and mammalian translocons (Menetret et al., 2000; Beckmann et al., 2001). As a consequence, the hydrophilic pore is likely to be considerably less spacious than previously expected (Hamman et al., 1997), even considering that the crystal structure is of the closed state. The 10 transmembrane helices of Sec61α form an aqueous channel with a central constriction of hydrophobic residues (Figure 3A).

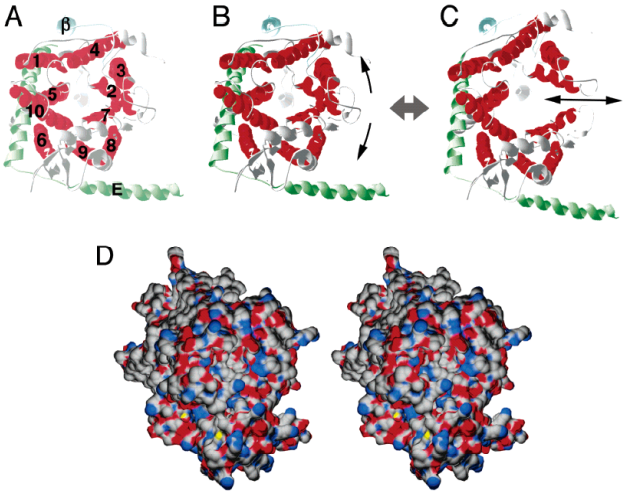

FIGURE 3: The translocation complex and its lateral exit site. The backbone structure of the *M. jannaschii* SecYEb complex (Protein Data Base accession code 1RHZ) (Serek et al., 2004) is shown from the cytosolic side. SecY (corresponding to Sec61α) is shown in gray with its 10 transmembrane domains in red (numbered in panel A). The central hydrophilic pore is blocked by a short luminal helix that must move away to allow passage of a translocating polypeptide. SecE (Sec61γ) and the β subunit are shown in green and blue, respectively. To allow exit of a hydrophobic sequence into the lipid bilayer, the two pseudo-symmetric halves (helices 1–5 and 6–10) must open (arrows in panel B) hinging around the connection between helices 5 and 6 (schematically shown in C). Panel D shows a stereo view of the SecYEG complex from the cytosolic side. The structure is slightly turned down in comparison to the view of panels A–C to better see into the hydrophilic pore, which in the closed state is blocked by the central constriction and the lumenal plug. Atoms are colored gray for C, blue for N, red for O, and yellow for S.

The channel is open to the cytosolic side, i.e. to the ribosome, but plugged by a short helix inserted from the lumenal side. This plug has to move away, probably by turning out as a whole around a flexible hinge in the connecting sequences, to allow passage of a translocating polypeptide. The plug may also play a role in sealing the channel against ion loss from within the ER lumen while idle.

Most interestingly for topogenesis, there is no obvious hydrophobic surface lining the inside of SecY/Sec61α that could serve as a static recognition site for signal sequences (Figure 3D). The translocon is organized in two halves (transmembrane helices 1–5 and 6–10; Figure 3A). To laterally leave the channel towards the lipid membrane, a polypeptide has to pass between helices 2/3 and 7/8. A hydrophobic environment is accessible to a signal sequence in the pore only when the channel laterally opens, for example due to thermal motion within the structure (Figure3, B and C). As the h-domain of the signal exits into the surrounding lipid, it will form a helix, optimizing intramolecular hydrogen bonds. Similarly, the signal might return into the hydrophilic channel where the peptide as a flexible chain may invert its orientation due to the flanking charges. According to the model illustrated in Figure 4, the Sec61 complex allows lateral equilibration of the signal between an aqueous and a transmembrane environment. The translating ribosome may facilitate the transient lateral opening of the pore. Upon termination of translation, re-entry of the signal may be hindered, resulting in the observed block of further topology changes. For translocation of the C-terminal sequence, the lumenal plug is shifted out of the way. This is also necessary for translocation of the N-terminal domain of type III proteins with reverse signal-anchors.

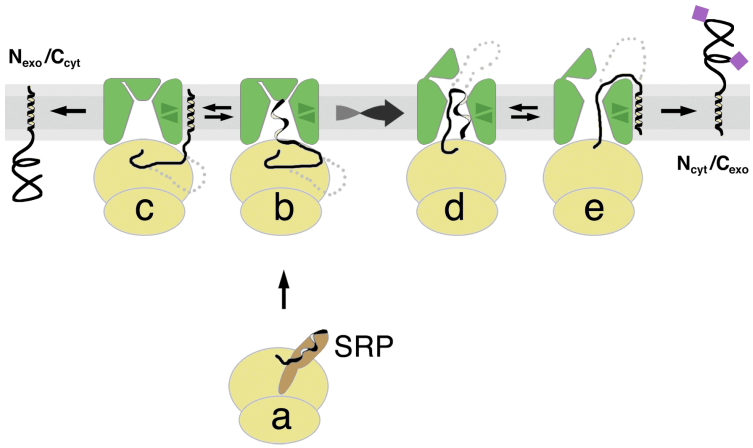

FIGURE 4: Model for signal orientation in the translocation complex. The signal of a nascent polypeptide is recognized by SRP (a) and targeted to the ER membrane via the SRP receptor (not shown). In the translocon, the signal reversibly partitions between the hydrophilic channel and the lipid environment via the lateral exit site (indicated by green arrowheads; b↔c and d↔e). The initial orientation is N_{exo}/C_{cyt}, (b and c) because the polypeptide is too short to loop early in the process. Based on a local electrical potential and depending on the flanking charges, the signal may invert its orientation while in the channel (b → d) and enter the lipid in the N_{cyt}/C_{exo} topology (e). Translocation of hydrophilic sequences requires the opening of the lumenal plug (d and e). The kinetics of inversion are accelerated by increasing charge difference of the signal (ΔN–C) and slowed down by increasing hydrophobicity (shifting the equilibrium to the membrane bound form). With further translation the growing peptide loops out (indicated by dotted lines) into the ER lumen or the cytoplasm depending on the signals orientation.

Charge interactions appear to provide the driving force for signal orientation. The role of conserved charged residues in Sec61p of yeast was tested by mutagenesis. Three mutations were identified that affect the topology of diagnostic substrates as expected when the positive-inside rule is weakened (Goder et al., 2004): two arginines at the lumenal plug and a glutamate at the cytosolic end of transmembrane domain 8. Although these three residues do not account for the entire charge effect in signal orientation, they show that Sec61p contributes to the positive-inside rule.

Topogenesis of multi-spanning membrane proteins

In complex membrane proteins, which span the membrane multiple times, it is generally the first hydrophobic sequence that targets the nascent protein to the ER membrane. This may be a cleavable signal, a signal-anchor, or a reverse signal-anchor, just as in single-spanning proteins (Figure 1, e–g). Some members of the seven-transmembrane receptor family, for example, generate an exoplasmic N-terminus with a reverse signal-anchor, whereas others (particularly those with large translocated N-terminal domains) employ a cleavable signal (f vs. g). Subsequent transmembrane segments insert with alternating orientations. In the simplest case, their orientations are determined by that of the initial signal sequence. Indeed, signal-anchors inserted downstream of a first cleavable signal or signal-anchor can function perfectly as stop-transfer sequences. Artificial proteins spanning the membrane up to four times have been created by tandemly repeating identical copies of a signal-anchor in a polypeptide separated by >100 amino acids from each other (Wessels and Spiess, 1988).

However, the topology of natural proteins is not just dictated by the first transmembrane domain, but appears to be supported by contributions of downstream sequences. Statistically, internal transmembrane domains also follow the charge rule, although less stringently than the most N-terminal signal (von Heijne, 1989). Mutations designed to invert the orientation of the initial transmembrane domain of the glucose transporter Glut1 did not cause inversion of the entire protein but resulted in a local defect (Sato et al., 1998). Similarly, insertion of positive charge clusters into short exoplasmic loops of model proteins caused "frustration" of individual hydrophobic domains, showing that internal charges can be topogenically active, but did not affect the topology of the rest of the protein (Gafvelin et al., 1997). In the case of the seven-transmembrane protein ProW, efficient translocation of the N-terminus of required the presence of at least four transmembrane domains (Nilsson et al., 2000). The topology of multispanning proteins thus seems to be determined by a consensus of its segments.

Cooperation of topogenic determinants throughout the sequence could be accomplished by retaining and assembling the transmembrane domains within the translocation machinery before the completed protein is released as a whole into the membrane, as previously proposed (Borel and Simon, 1996). However, subsequent transmembrane domains could be crosslinked to lipid as soon as they reached the

translocon (Martoglio et al., 1995; Mothes et al., 1997; Heinrich et al., 2000), indicating that they exit into the membrane one-by-one. If downstream sequences overrule the initial topology of a transmembrane segment, it must be able to return back into the translocation pore to reorient itself.

Evidence for substantial polypeptide reorientation was obtained in in vivo experiments using a model protein with two conflicting topogenic sequences, a cleavable signal at the N-terminus and an internal signal-anchor (Goder et al., 1999). When these two signals were separated by ≥80 residues, these spacer residues were translocated and the second hydrophobic segment functioned as a stop-transfer sequence. With shorter spacers, however, an increasing fraction of proteins inserted with a translocated C-terminus as dictated by the second signal. A glycosylation site in the spacer increased translocation of the spacer sequence. This indicates that the second hydrophobic sequence, by inserting in an N_{cyt}/C_{exo} orientation, forces the spacer of up to ~60 residues to return from the ER lumen to the cytosol, unless it is glycosylated. In the process, at least two hydrophilic polypeptide segments have to pass simultaneously through the translocation pore. A similar situation may underlie the generation of the various topologies of the prion protein (in particular those termed ^{Ntm}PrP and ^{Ctm}PrP, where a mildly hydrophobic sequence may integrate in either orientation) (Stewart et al., 2001).

As one transmembrane segment after the other reversibly partitions into the lipid membrane, they may associate with each other and partially assemble before protein synthesis is completed. Membrane integration of weakly hydrophobic sequences was found to be stabilized by a preceding transmembrane segment, and the overall topology became more defined (Ota et al., 2000; Heinrich and Rapoport, 2003). Similarly, complementary charged residues in different transmembrane segments of the K^+ channel KAT1 were found to be required for the correct topology (Sato et al., 2003). Topogenesis and protein folding are thus not necessarily separable events. Helix bundling may start already during protein insertion and influence the resulting topology.

The Sec61 translocon provides multiple functions: it constitutes a gated pore for the passage of hydrophilic polypeptides through the membrane barrier, it allows hydrophobic segments lateral access to the core of the lipid bilayer for integration as transmembrane helices, and it contributes to their orientation. It works with

highly diverse substrate sequences and even de novo designed generic sequences. Other components are likely to act upon the translocation complex and regulate its properties. Regulatory roles have been suggested for the ribosome (Liao et al., 1997) and the lumenal chaperone BiP (Haigh and Johnson, 2002) in sealing the translocation pore either on the cytosolic or the lumenal side to maintain ion gradients at the ER membrane (Alder and Johnson, 2004). It is likely that unassembled transmembrane domains of nascent proteins are taken care of by intramembrane chaperones (potentially TRAM or PAT-10 (Meacock et al., 2002)). There are further indications that specific (particularly non-bilayer) lipids assist protein folding in the membrane as "lipochaperones" (Bogdanov and Dowhan, 1999) and influence translocon function and topogenesis (van Klompenburg et al., 1997; Van Voorst and De Kruijff, 2000). The current challenge is to derive a molecular understanding of a highly dynamic process from relatively static experimental data such as crosslinking snapshots, endpoint topologies of model substrates, and structural data, the most recent milestone being the crystal structure of SecYEG (Van den Berg et al., 2004).

Thesis goal:

In eukaryotic cells, polypeptides destined for membrane proteins are integrated into the membrane of the ER (Johnson and van Waes, 1999). But membrane proteins biogenesis is therefore exceedingly complex, especially because the mechanisms involved are further constrained by the need to maintain the permeability barrier of the membrane.

The most recent advances in understanding the cotranslational integration at the ER membrane are focusing on different areas, like the translocon structural and functional states, the nascent chain topogenesis, the insertion of transmembrane domains (TMs) into the bilayer and the nascent chain regulation of integration.

To help understanding these different points, Goder et al (2000) established a glycosylation assay. In this assay. these authers used as model protein the H1 subunit of the asialoglycoprotein receptor, a C-type lectin of hepatocytes that recognizes desialylated glycoproteins for clathrin dependent endocytosis and lysosomal degradation. H1 is a typical type II single-spanning transmembrane protein. They analyzed *in vivo* how model signal sequences were inserted and oriented in the membrane during cotranslational integration into the ER. Their results were incompatible with the current models of retention of positive flanking charges or loop insertion of the polypeptide into the translocon. Instead they indicated that these N-terminal signals initially insert head-on with a cytoplasmic C-terminus before they invert their orientation to translocate the C-terminal sequence. The rate of inversion increased with more positive N-terminal charges and was reduced with increasing hydrophobicity of the signal. Inversion may proceed for up to approximately 50 s, when it is terminated by a signal-independent process. These findings provided a mechanism for the topogenic effects of flanking charges as well as of signal hydrophobicity (Goder and Spiess, 2003).

To go further in this direction, we decided to use the same model protein to define the environment of the TM domain during topogenesis. It was showed that the TM domain could be in equilibrium between two states inside the translocon a "free" and a "bound" state. Each of these states could influence the topology of the protein. One partners could be the translocon itself (McCormick et al., 2003). More

precisely Rapoport and colleagues showed that the signal sequence interacts with the translocon on a hydrophobic site, this site was not clearly identified with the crystallization of the Sec61 translocon (Van den Berg et al., 2004). On the other hand the TM could also interact with lipids surrounding the translocon (Martoglio et al., 1995; Mothes et al., 1998) .

In order to define the binding partners of the signal sequence during topogenesis, we decided to affect the interaction between the signal sequence and its binding partners inside the translocon. A way to look at the environment of the signal sequence is to modify it, by changing only one parameter, in this model we modified the bulkiness of the signal sequence. For this purpose we designed a series of constructs with a couple of bulky amino acids placed inside an oligoleucine transmembrane domain (Figure 1). Therefore, we looked at the position effect of different bulky amino acids placed through the signal sequence. Previous data (Rosch et al., 2000) already showed that bulky amino acids could have an effect on topology.

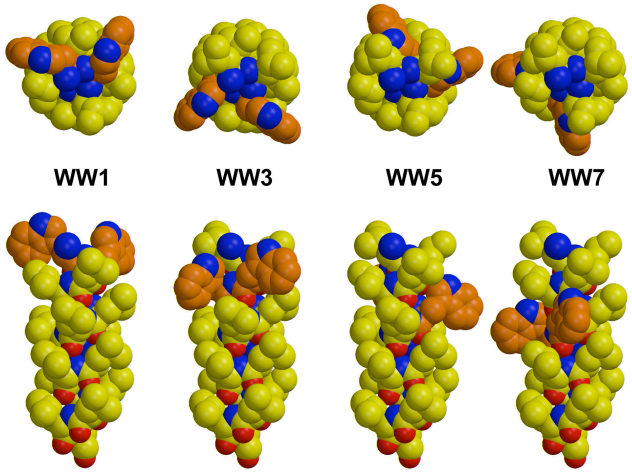

Figure 1: Bulkiness of the Tryptophan. The space-filling models of the signal sequence of the first four tryptophan constructs show the bulkiness of the tryptophan residues (orange) within the oligo-leucine (yelllow) helix in a top and side view for the first four constructs.

Furthermore, Goder (2003) found that one can study the kinetics of topogenesis by making constructs with a different tail lengths. Since topogenesis finishes when translation ends, one can look at different time points of topogenesis with a series of longer and shorter constructs. So for the second part, we looked at the same kinetics of inversion but this time with a double-spanning membrane protein. In this case we would like to determine to the kinetics of inversion of the second transmembrane domain and also the effect of this domain on the topology of the first signal.

CHAPTER I: Probing the environment of the signal during topogenesis

Higy M, Gander S, Spiess. Probing the environment of signal--anchor sequences during topogenesis in the endoplasmic reticulum. Biochemistry. 2005 Feb 15;44(6):2039-47.

Probing the Environment of Signal-Anchor Sequences during Topogenesis in the Endoplasmic Reticulum[†]

*Marie Higy, Stefan Gander, and Martin Spiess**

Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

† This work was supported by Grant 31-061579.00 from the Swiss National Science Foundation.

* To whom correspondence should be addressed: Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland; Phone: +41-61-2672164; FAX: +41-61-2672148; E-mail: Martin.Spiess@unibas.ch

ABSTRACT

Signal sequences for insertion of protein into the mammalian endoplasmic reticulum orient themselves in the translocon on the basis of their flanking charges. It has recently been shown that hydrophobic N-terminal signals initially insert head-on before they invert their orientation to translocate the C-terminus. The rate of inversion is reduced with the increasing hydrophobicity of the signal due to an increased affinity for the initial bound state at the translocon. To probe the environment of the signal while its orientation is determined, different hydrophobic residues were inserted at various positions throughout a uniform oligo-leucine signal sequence and the constructs were expressed in transfected COS-7 cells. The resulting topologies revealed a strikingly symmetric position dependence specifically for the bulky aromatic amino acid, reflecting the structure of a lipid bilayer. Maximal N-translocation was observed when the guest residues were placed at the N- or C-terminus of the hydrophobic sequence or in the very center, corresponding to the positions of highest expected affinity of the signal sequence as a membrane-spanning helix for the bilayer. The results support the model that during topogenesis in vivo the signal sequence is exposed to the lipid membrane.

[1] The abbreviations used are: ER, endoplasmic reticulum; N_{cyt}/C_{exo}, and N_{exo}/C_{cyt}, cytoplasmic N- terminus and exoplasmic C-terminus, and vice versa; respectively; SRP, signal recognition particle.

Hydrophobic signal sequences target secretory and membrane proteins to the endoplasmic reticulum (ER) for translocation into the lumen or for integration into the lipid bilayer (Walter and Johnson, 1994; Rapoport et al., 1996; Johnson and van Waes, 1999). The signal is first recognized by the signal recognition particle (SRP)[1], which directs the nascent chain-ribosome complex to the ER membrane by interaction with the SRP receptor (Keenan et al., 2001). Both SRP and SRP receptor are GTPases which control specific docking to the translocation machinery. The actual translocation pore or translocon is created by the Sec61 complex composed of an a-subunit that spans the membrane ten times, and a single-spanning β- and γ-subunit. This pore complex allows hydrophilic sequences to pass through the membrane and permits signal and transmembrane sequences to laterally exit into the lipid bilayer.

A signal sequence can insert into the translocon and subsequently into the membrane in two orientations (Goder and Spiess, 2001). Cleavable signals and signal-anchor sequences of type II membrane proteins translocate their C-terminal end, acquire an N_{cyt}/C_{exo} orientation (cytoplasmic N-terminus and exoplasmic C-terminus), and they initiate cotranslational transfer of the growing polypeptide across the membrane. In contrast, reverse signal-anchors sequences of type III proteins (also classified as type Ia) insert with an N_{exo}/C_{cyt} orientation and induce translocation of the N-terminus. Examples for type III proteins are synaptotagmin, neuregulin, and the family of cytochromes P450.

Several factors have been shown to determine the orientation of the signal in the membrane. Most prominently, charged residues flanking the hydrophobic core of the signal influence orientation: the more positive end is generally cytosolic, a phenomenon known as the "positive-inside rule" (Hartmann et al., 1989; Beltzer et al., 1991; Parks and Lamb, 1991). Charge interactions at the translocon, including residues of the Sec61 complex itself (Goder et al., 2004) are responsible for orienting the signal sequence. In addition, the folding state of hydrophilic sequences N-terminal to a signal may sterically hinder N-terminal translocation irrespective of the flanking charges (Denzer et al., 1995). A third determinant is the hydrophobicity of the core of the signal sequence (the h-domain) itself. Strongly hydrophobic signals were observed to insert with an N_{exo}/C_{cyt} orientation even when

the charge distribution was more positive at the N-terminus (Sakaguchi et al., 1992; Wahlberg and Spiess, 1997; Rosch et al., 2000).

How hydrophobicity influences signal orientation was not obvious. A recent in vivo study indicated that hydrophobicity acts on the kinetics of signal orientation in the translocation machinery (Goder and Spiess, 2003). An N-terminal signal-anchor sequence with an h-domain of 22 consecutive leucine residues inserted with mixed topologies despite a positive N-terminus. Surprisingly, the topology depended on the total length of the protein: the fraction of polypeptides with an N_{cyt}/C_{exo} orientation was lowest for the shortest constructs and increased up to a length of ~300 residues following the signal sequence. This result indicated that N-terminal signals initially insert "head-on" into the translocon to yield an N_{exo}/C_{cyt} orientation. Then they may invert over time until protein synthesis is completed or until further reorientation is stopped by an as yet unknown mechanism after approximately 40-50 s. Inversion appears to be driven by electrostatic interactions, since the topology change was accelerated by increasing the N-terminal positive charge and slowed by reducing it. The hydrophobicity of the h-domain also affected reorientation: increased hydrophobicity diminished the rate of inversion, whereas reduced hydrophobicity allowed more rapid inversion. This effect could be explained by hydrophobicity stabilizing the binding of the signal to an interaction site at the translocation apparatus.

The hydrophobic core of a signal sequence thus affects topology by influencing the kinetics of signal reorientation, rather than by changing its preference for a final topology. This opens the possibility of exploring the environment of the signal during the orientation process using a series of constructs with identical charge distribution and overall hydrophobicity, but containing guest residues in the h-domain at various positions through the sequence. Here we tested the effect of the large hydrophobic amino acids tryptophan, phenylalanine, and tyrosine, as well as the small hydrophobic residues valine and alanine inserted at various positions in an oligo-leucine sequence. Bulky hydrophobic amino acids showed a striking position dependence of topogenesis in a conspicuously symmetric pattern not observed with the other similarly hydrophobic amino acid. The results suggest the exposure of the signal to the lipid bilayer during topogenesis in vivo.

RESULTS

Protein topology is strongly dependent on the position of double-tryptophans in an H1ΔQLeu16 host sequence.

To explore the properties of the signal's environment while its orientation is determined, we tested the behavior of model proteins based on H1ΔQLeu16. This protein is derived from the H1 subunit of the asialoglycoprotein receptor, a typical type II membrane protein. It has been modified at the N-terminus to start with an artificial signal sequence of 16 consecutive leucine residues with a single positive charge at the N-terminus (the α-amino group) and a net negative C-terminal flanking sequence (Figure 1B).

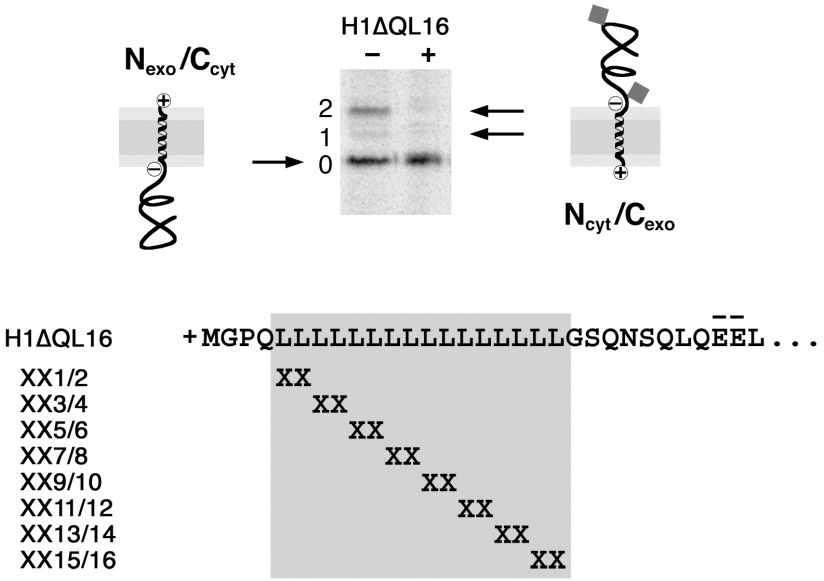

FIGURE 1. Model proteins for studying the position effect of guest residues in an oligo-leucine host signal-anchor sequence. Top Figure: The signal-anchor protein H1ΔQL16 was expressed in COS-7 cells, labeled with [^{35}S]methionine, immunoprecipitated, and analyzed by SDS gel electrophoresis and autoradiography. Three forms were produced corresponding to the unglycosylated protein (0) with the N_{exo}/C_{cyt} orientation, or with one (1) or two (2) glycans of the N_{cyt}/C_{exo} orientation, as schematically illustrated. The latter two forms are sensitive to deglycosylation by endoglycosidase H (+). One of the glycosylation sites is close to the transmembrane sequence and is not completely modified, resulting in some products that are glycosylated only once. Bottom Figure: To probe the signal's environment during topogenesis, two identical guest residues (XX) were systematically introduced at the indicated positions into the hydrophobic domain (gray background) of the host sequence H1ΔQL16. The N-terminal sequence of H1ΔQL16 is shown in single-letter code (charged residues denoted with + or –).

Because of its considerably hydrophobic core, only ~35% of the molecules are able to translocate the C-terminus, before the polypeptide of 230 amino acids is completed and further reorientation of the protein is terminated (ref. Goder and Spiess, 2003, and autoradiograph in Figure 1A). To test the position effect of guest residues throughout the uniform h-domain of this construct, we replaced pairs of leucine residues with tryptophans, the most voluminous amino acid. Two residues were simultaneously altered to potentially generate more significant effects. When two tryptophans had previously been placed at positions 8 and 13 of H1ΔQLeu16, the distribution of topologies had not significantly changed in comparison to the parental construct (Rosch et al., 2000) (hydroscales deleted)

The constructs containing double tryptophans placed throughout the signal sequence were expressed in transfected COS-7 cells and labeled with [^{35}S]methionine for 40 min. The products were immunoprecipitated and analyzed by SDS gel electrophoresis and autoradiography (Figure 2).

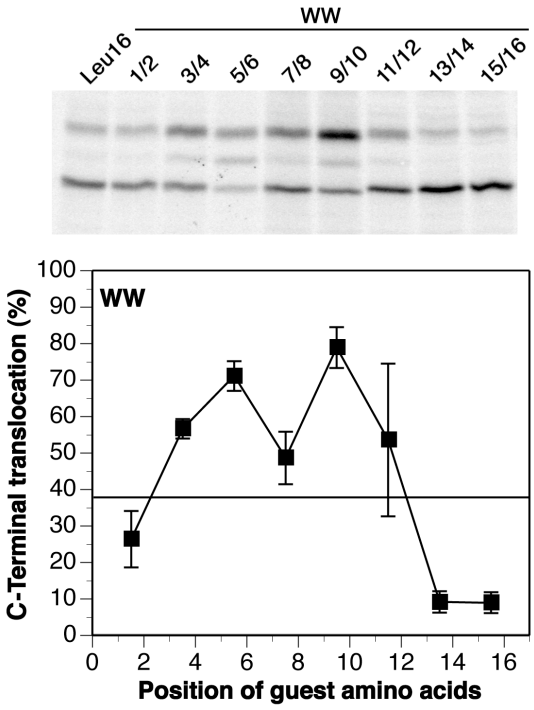

FIGURE 2. Topogenic effect of tryptophan guest residues in an oligo-leucine signal-anchor sequence. A series of constructs with two tryptophans (WW) replacing two leucines throughout the oligo-leucine sequence in H1ΔQL16 (as shown in Figure 1) was expressed in COS-7 cells labeled with [^{35}S]methionine, immunoprecipitated, and analyzed by SDS gel electrophoresis and autoradiography. The glycosylated and unglycosylated forms representing N_{cyt}/C_{exo} and N_{exo}/C_{cyt} orientations, respectively, were quantified and expressed as the fraction of polypeptides with a translocated C-terminus (N_{cyt}/C_{exo} orientation) vs the position of the guest residues in the oligo-leucine sequence. The average and standard deviation of four independent experiments are shown. The horizontal line represents the topology distribution of the original host sequence (Leu16).

Topology was derived from the glycosylation pattern, since modification of the N-glycosylation sites at positions 40 and 108 indicates C-terminal translocation, whereas the unglycosylated form indicates an N_{exo}/C_{cyt} orientation (Figure 1A and refs. Wahlberg) (Wahlberg and Spiess, 1997; Rosch et al., 2000; Goder and Spiess, 2003). The results showed a dramatic position dependence for the tryptophans in a surprisingly symmetric pattern. When positioned at either end of the h-domain, the tryptophan-containing sequences inserted predominantly with an N_{exo}/C_{cyt} orientation. In contrast, tryptophans further inside the h-domain favored C-terminal translocation, except when placed in the center at positions 7 and 8 where the fraction of N_{exo}/C_{cyt} molecules with ~50% was again significantly increased. Constructs with single tryptophans inserted into H1ΔQL16 produced a similar pattern with smaller deviations (data not shown).

Interpretation of the glycosylation pattern in terms of protein topology depends on the assumption that all products are integrated into the membrane and that glycosylation efficiency is not affected by insertion of the guest residues. Glycosylation efficiency at one of the two glycosylation sites (most likely the one closest to the membrane) is indeed somewhat dependent on the sequence, probably reflecting the protein's position in the membrane when the glycans are transferred (Nilsson and von Heijne, 1993). To test for complete membrane integration, WW1/2, WW7/8, and WW9/10, i.e., constructs covering the entire spectrum of different topology distributions, were expressed in COS-7 cells, labeled, and subjected to alkaline extraction and centrifugation (Figure 3A)

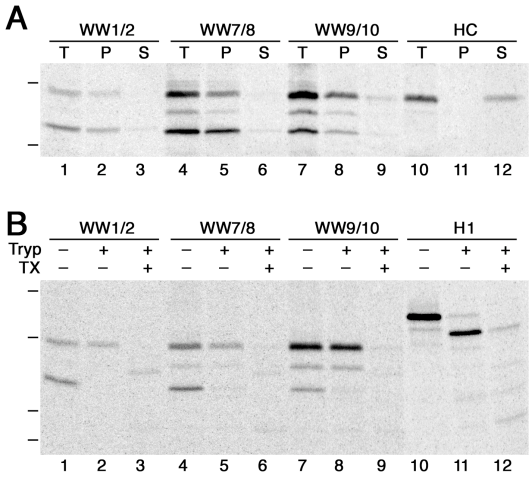

FIGURE 3. Glycosylation patterns reflect protein topology. (A) COS-7 cells expressing selected constructs were labeled with [^{35}S]methionine and subjected to alkaline extraction. After centrifugation, the membrane pellet (P) and the supernatant (S), as well as an equal aliquot of the total starting material (T), were analyzed by immunoprecipitation, SDS gel electrophoresis, and autoradiography. The control construct HC consists of the cleavable signal sequence of influenza hemagglutinin fused to the C-terminal domain of H1. The positions of molecular mass markers of 26 and 37 kDa are indicated. (B) COS-7 cells expressing selected constructs or wild-type H1 were labeled with [^{35}S]methionine, permeabilized by swelling and scraping, and incubated with or without trypsin (Tryp) in the presence or absence of Triton X-100 (TX). The products were then analyzed by immunoprecipitation, SDS gel electrophoresis, and autoradiography. The positions of molecular mass markers of 20, 26, 37, and 50 kDa are indicated.

Glycosylated and unglycosylated forms of all constructs were equally and almost completely recovered in the membrane pellet, whereas a control protein consisting of the cleavable signal of influenza hemagglutinin fused to the C-terminal portion of H1 was extracted into the supernatant. All signal-anchor constructs were thus efficiently targeted to the ER and integrated into the membrane irrespective of the position of the tryptophan residues.

To test whether all unglycosylated products have an N_{exo}/C_{cyt} orientation, cells were labeled, ruptured by swelling and scraping, and incubated at 4 °C with or without trypsin (Figure 3B).

Wild-type H1, which was analyzed as a control protein, shifted its position in SDS gel electrophoresis upon trypsin treatment because of digestion of its 40-amino acid cytoplasmic domain. The unglycosylated forms of all three WW constructs were efficiently digested by protease, whereas the glycosylated ones were resistant. Upon permeabilization of membranes by detergent, all products were sensitive. These control experiments confirm that the glycosylation state of the proteins as analyzed in Figure 2 directly represents their final orientation in the membrane.

Topologies at a fixed protein length reflect the kinetics of signal reorientation.

To analyze whether the observed topologies of different constructs are a measure of the inversion kinetics of the signal-anchor sequence in the translocon as shown before for the parental construct H1ΔQL16 (Goder and Spiess, 2003) we generated a series of constructs with 110, 170, 230, 290, and 460 residues following the signal-anchor sequences of WW1/2, WW7/8, and WW9/10. Upon expression in COS-7 cells, labeling, and immunoprecipitation, the fraction of glycosylated products corresponding to N_{cyt}/C_{exo} polypeptides was determined.

With the increasing length of the proteins, the C-terminally translocated fraction increased (Figure 4). On the basis of a translation rate of ~5 amino acids/s as determined for cultured mammalian cells (Hershey, 1991), the length of the protein can be converted into the time of translation from the moment the signal has fully emerged from the ribosome (with ~40 residues still hidden within the ribosome) to termination (bottom scale in Figure 4).

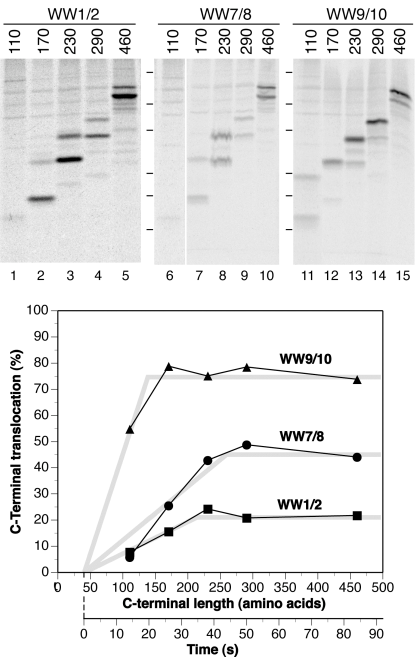

FIGURE 4. Topologies reflect the inversion kinetics of the signal-anchor sequences. For each of the constructs WW1/2, WW7/8, and WW9/10, a series was constructed with polypeptides of 110, 170, 230, 290, and 460 residues following the signal-anchor sequence. These constructs were expressed in COS-7 cells, labeled with [^{35}S]methionine, immunoprecipitated, and analyzed by SDS gel electrophoresis. The positions of molecular mass markers of 15, 20, 26, 37, 50, and 64 kDa are indicated. The glycosylated and unglycosylated forms representing the N_{cyt}/C_{exo} and N_{exo}/C_{cyt} orientations, respectively, were quantified and plotted as the fraction of polypeptides with a translocated C-terminus (N_{cyt}/C_{exo} orientation) vs the length of the C-terminal domain and the time of translation from when the signal completely emerged from the ribosome to termination. The time scale is based on a translation rate of 5 amino acids/s as determined for cultured cell lines (Hershey, 1991) and starts at 40 amino acids, since 30-40 residues are hidden within the ribosome. The average of three independent experiments (including the one presented at the top) is shown.

The rate of increase of N_{cyt}/C_{exo} products was lowest for the WW1/2 series of constructs, highest for WW9/10, and intermediate for WW7/8. The topologies observed in Figure 2 for a fixed length of the protein of 230 amino acids downstream of the signal-anchor sequences therefore reflect the signals' inversion kinetics. Thus, tryptophans at the end of the signal core inhibit inversion compared to the oligo-leucine host sequence, whereas tryptophans between enhance inversion, except in the very center.

Other bulky hydrophobic guest residues generate a similar position dependence of protein topology.

The signal's inversion kinetics could be affected either by changing the flexibility of the polypeptide in its free state in the translocon in comparison to the host sequence or by altering the affinity to the initial bound state. Tryptophan has a bulky aromatic side chain, and it appears to be unlikely that the exchange with leucines by tryptophans at any position would make the polypeptide more flexible. To explain the increased inversion kinetics for guest tryptophans at positions 3-12, it is more plausible that tryptophans at these positions reduce the affinity for the signal binding site in comparison to the parental oligo-leucine sequence.

Because of the nitrogen in its side chain, tryptophan also has polar properties. To address whether the observed position dependence of tryptophans in the host signal is caused by the size of the side chain or its polar contribution, two additional series of constructs with two phenylalanines or two tyrosines as guest residues in the oligo-leucine sequence were prepared and analyzed in transfected COS-7 cells as described above. Both phenylalanine and tyrosine are hydrophobic, aromatic, and rigid. They are large in comparison to most other amino acids, although somewhat smaller than tryptophan. The side chain of phenylalanine lacks a polar group entirely, whereas that of tyrosine contains a polar hydroxyl group.

The resulting topologies for the constructs with double phenylalanines (Figure 5A) presented almost the same symmetric position dependence with minimal N_{cyt}/C_{exo} orientation for phenylalanines placed at the beginning, at the end, or in the center of the signal-anchor sequence. The most notable difference is that

phenylalanines at the ends of the signal yielded more N_{cyt}/C_{exo} orientation than the tryptophans. The polar characteristics of tryptophan might be responsible for the strongly reduced inversion kinetics when positioned at the ends of the signal.

The corresponding series of tyrosine-containing constructs again showed the same basic pattern of topologies with relative minima of N_{cyt}/C_{exo} topology with tyrosines on either end or in the center of the h-domain (Figure 5B). Tyrosines at the C-terminal end inserted almost completely with an N_{exo}/C_{cyt} orientation like the corresponding tryptophan constructs, consistent with tyrosine having polar characteristics as well. However, N-terminal tyrosines (YY1/2) produced relatively high levels of the N_{cyt}/C_{exo} orientation (see Discussion).

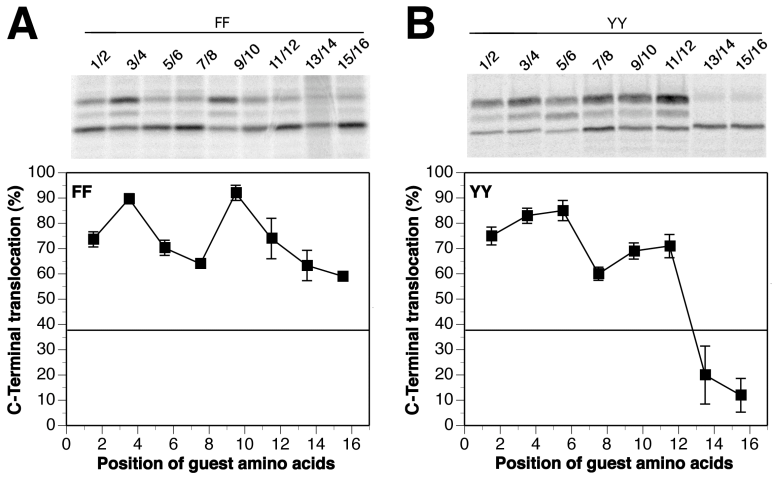

Figure 5. Position dependence of other bulky aromatic guest residues: phenylalanines and tyrosines. Constructs series with two phenylalanines (FF; panel A) or tyrosines (YY; panel B) replacing leucines throughout the oligo-leucine sequence in H1ΔQL16 were expressed in COS-7 cells, labeled with [^{35}S]methionine, immunoprecipitated, and analyzed by SDS gel electrophoresis and autoradiography. The glycosylated and unglycosylated forms were quantified and plotted as the fraction of polypeptides with a translocated C-terminus (N_{cyt}/C_{exo} orientation) vs the position of the guest residues in the oligo-leucine sequence. The average and standard deviation of three independent experiments are shown. The horizontal line represents the topology distribution of the original host sequence.

Symmetric position dependence correlates with side chain bulkiness.

As a control, the effect of valine guest residues was tested with the constructs VV1/2 to VV15/16 (Figure 6A).

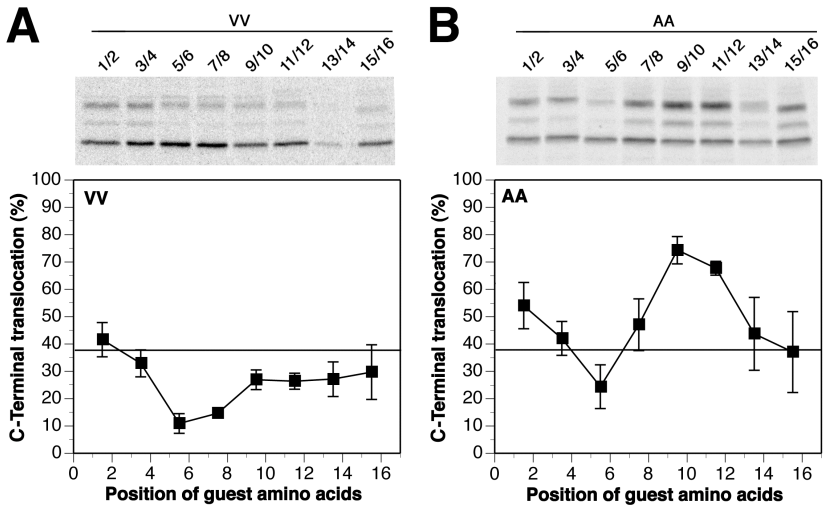

FIGURE 6. Topogenic effect of nonbulky valine and alanine guest residues. Constructs series with two valines (VV; panel A) or alanines (AA; panel B) replacing leucines throughout the oligo-leucine sequence in H1ΔQL16 were expressed in COS-7 cells, labeled with [^{35}S]methionine, immunoprecipitated, and analyzed by SDS gel electrophoresis and autoradiography. The glycosylated and unglycosylated forms were quantified and plotted as the fraction of polypeptides with a translocated C-terminus (N_{cyt}/C_{exo} orientation) vs the position of the guest residues in the oligo-leucine sequence. The average and standard deviation of three independent experiments are shown. The horizontal line represents the topology distribution of the original host sequence.

Valine is smaller than leucine, but of similar in hydrophobicity. The position effects were less distinctive and entirely different from those of the bulky aromatic residues. The fraction of products with the N_{cyt}/C_{exo} orientation was generally somewhat lower than that of the parental sequence H1ΔQLeu16, ~30%, but was slightly increased to 40% when the valines were positioned at the N-terminus and decreased to ~15% at positions 5 and 6 and positions 7 and 8. The pattern was again different when alanines were used as guest residues (Figure 6B). Alanine is small and considerably less hydrophobic than the other residues that were tested. The major effect on topology of the host sequence was a significant increase in the amount of N_{cyt}/C_{exo} orientation when the alanines were placed at positions 9 and 10 and positions 11 and 12. The results with alanine and valine guest residues show that the symmetric M-shaped position dependence of topology is specific for amino acids with bulky side chain

DISCUSSION

Signal sequences are first recognized by SRP on a hydrophobic surface created by a cluster of methionines on the 54 kDa subunit (SRP54) (Zopf et al., 1990; Lutcke et al., 1992; Keenan et al., 1998). Once the ribosome-nascent chain-SRP complex has docked, the signal is transferred to a second recognition site in the translocon (Do et al., 1996; Mothes et al., 1998; McCormick et al., 2003). Topogenic determinants control how the signal is positioned in the translocon: electrostatic forces act on the flanking charges (Hartmann et al., 1989; Goder et al., 2004) and an N-terminal hydrophilic extension sterically hinders insertion of the N-terminus (Denzer et al., 1995). Our previous in vivo studies for N-terminal, very hydrophobic signals showed that their orientation changes with the time of translation (Goder and Spiess, 2003). Initial insertion leads to an N_{exo}/C_{cyt} orientation, but C-terminal translocation is acquired gradually in a process that is accelerated by an increasing charge difference Δ(N-C) and slowed by the increasing hydrophobicity of the signal. With constant charge and overall hydrophobicity, the effect on topology of guest residues inserted at various positions in a generic oligo-leucine signal is expected to reflect changes in the affinity of the signal for the initial bound state and in the flexibility in the free state. The position effect of guest residues on topology may thus provide information about the situation of the signal sequence during topogenesis in vivo.

The position effects of valines and alanines may be best explained by their effects on the flexibility of the polypeptide chain in comparison to the parental oligo-leucine sequence. Valine is a β-branched amino acid that reduces flexibility, whereas the small side chain of alanine allows more conformational freedom. Reduced or increased flexibility particularly at central positions in the apolar sequence will hinder or facilitate, respectively, the inversion of the free polypeptide. Bulky hydrophobic residues generated a strikingly symmetric pattern that cannot be explained by variations in polypeptide flexibility. It is more likely that the properties of the signal binding site dominate the changes in inversion kinetics for different positions of the guest residues. In a proteinaceous signal

binding site, bulky side chains might sterically hinder binding or contribute additional favorable contacts, therefore reducing or increasing the binding affinity, resulting in higher or lower inversion rates, respectively. Without a detailed structure of such a binding site, the position dependence of guest residues is not predictable. Symmetry would be accidental. If the bound signal had a helical conformation and were bound on one side to a protein surface in a fixed position, one would expect to find a periodicity of three to four residues. This is clearly not the case. In addition, the oligo-leucine sequence is uniform and therefore might always position itself in such a way that the bulky residues are facing away from the protein surface.

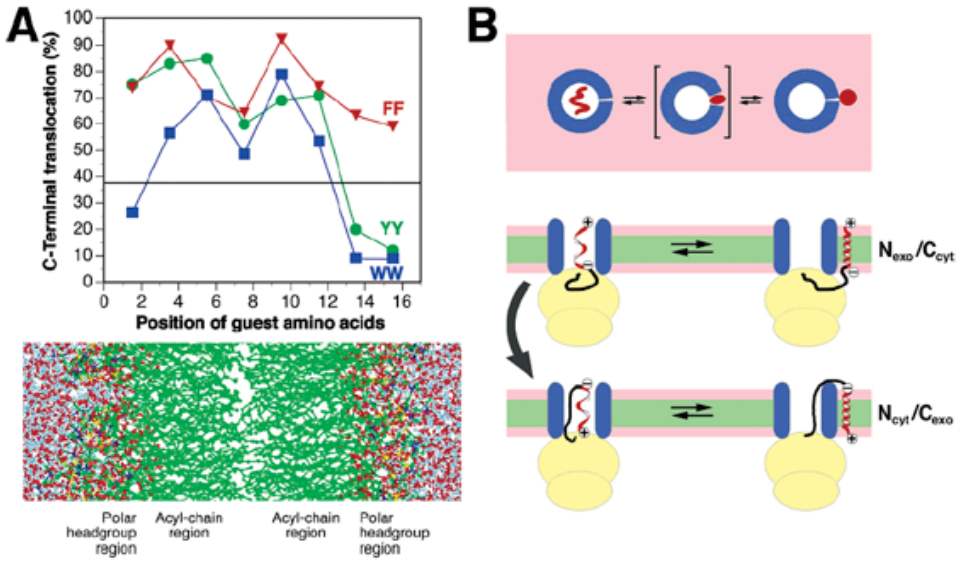

FIGURE 7. The position dependence of bulky guest residues in the signal-anchor sequence of H1ΔLeu16 that reflects the symmetry of the lipid bilayer. (A) The dependence of topology on the position of double tryptophans (WW, blue squares; from Figure 2), phenylalanines (FF, red triangles; from Figure 5A), and tyrosines (YY, green circles; from Figure 5B) are shown with a section of a dipalmitoyl phosphatidylcholine bilayer upon molecular dynamics simulation (Kulinska et al., 2000), clearly visualizing the low density in the center of the membrane and the polar-nonpolar transition at the headgroup interphases (courtesy of A. Lyubartsev, Stockholm University, Sweden). (B) The signal during topogenesis at the translocon is schematically shown from the top (top panel) or from the side (bottom panel). The signal is proposed to be in an equilibrium between a free state in the pore and a bound state in the lipid membrane outside the translocon. The situation of the signal intercalating between the surfaces of the exit site is likely to be an unstable transition state (indicated by brackets). Only the free state is able to invert in a manner driven by charge interactions. See the text also.

However, the symmetry of the position dependence for bulky aromatic guest residues parallels the symmetry of the lipid bilayer (illustrated in Figure 7A) and thus suggests contact of the signal with the lipid membrane during topogenesis.

The low fraction of the N_{cyt}/C_{exo} orientation obtained for tryptophans positioned at the ends of the h-domain could be explained by their favorable interaction with the interphase between the apolar core and the headgroup regions of the lipid bilayer (Wimley and White, 1996). Statistically, tryptophan is enriched in the interphase regions of transmembrane helices (Landolt-Marticorena et al., 1993; Arkin and Brunger, 1998) and of β-barrel proteins (Weiss et al., 1991; Schirmer et al., 1995), probably stabilizing the proteins of the position in the membrane. This is also the case for tyrosine, whereas there is no preferred position for phenylalanine in transmembrane sequences. The highest fractions of N_{cyt}/C_{exo} orientation and thus the lowest apparent affinity for the bound state in the translocon are observed for WW5/6 and WW9/10, in which the tryptophans in a transmembrane helix would be positioned in the center of the acyl chain regions of the two lipid layers, where the membrane is most tightly packed and accommodation of a large and stiff side chain is least favorable. In contrast, at the center of the bilayer, order and density are lowest and bulky side chains are more easily accommodated. Correspondingly, tryptophans at the center of the signal in WW7/8 showed a reduced amount of the N_{cyt}/C_{exo} orientation, i.e., increased affinity in comparison to WW5/6 and WW9/10.

The behavior of phenylalanine guest residues supports this interpretation. Because the phenylalanine side chain is similarly bulky, they produce the same basic pattern as tryptophans. Since the side chain is nonpolar and cannot take advantage of the polarity change at the interphase regions of the bilayer, phenylalanines at the ends of the oligo-leucine sequence showed more N_{cyt}/C_{exo} orientation. In contrast, tyrosines should behave more like tryptophans. They do, except at the very N-terminus where tyrosines yielded a higher fraction of N_{cyt}/C_{exo} topology and the entire topology pattern for tyrosine guest residues appears to be tilted. A potential explanation might be that the polarity axis of the tyrosine side chain, unlike that of the tryptophan side chain, cannot reach into the interphase as steeply at the N-terminus of a helix as at the C-terminus.

Sixteen apolar residues are not quite long enough to completely span the hydrophobic core region of the membrane. They are, however, flanked on both sides by several polar but uncharged residues, which can also be found within natural signal and transmembrane sequences. Part of these flanking sequences thus must be pulled into the apolar phase (Killian and von Heijne, 2000). The symmetry of topology patterns for all three aromatic residues seems to be slightly offset toward the N-terminus, suggesting that the N-terminus of the signal is pulled into the hydrophobic phase more easily. It has also been shown by glycosylation mapping that tryptophans can influence the positioning of a transmembrane helix according to its interface preference (Braun and von Heijne, 1999). However, the presence of aromatic residues is not likely to affect the topogenesis of natural signal sequences, because they orient themselves much more rapidly than the highly hydrophobic model signals used here. They have completely inverted their orientation before topogenesis is terminated.

In summary, our results suggest that the signal contacts the lipid bilayer during topogenesis in vivo. The concept that the signal upon insertion into the translocon is in contact with lipids is of course not new. In vitro, arrested nascent chains that are just long enough for the signal to enter the translocon could be photo-cross-linked not only to Sec61α, but also to lipids (Martoglio et al., 1995; Mothes et al., 1998). Since the extent of cross-linking to the lipid was higher for a more hydrophobic signal-anchor sequence than for a short cleavable signal, it was further proposed that the translocon might open more or less toward the lipid membrane depending on the hydrophobicity of the signal (Martoglio et al., 1995). However, cross-linking experiments provide a snapshot of a dynamic situation, and the possibility that some signals were cross-linked to the protein while still within the translocon could not be ruled out; others had already been integrated into the bilayer reacting with lipids. Cross-linking patterns obtained with a reactive side chain in different positions in the h-domain suggested that the signal was in a helical conformation in stable contact with transmembrane helices 2 and 7 of Sec61α on one side and with the lipid on the other (Mothes et al., 1998). This might reflect the state after the signal had left the translocon upon completion of topogenesis. In vivo, signal reorientation was found to be terminated at the latest approximately 50 s after the signal emerged from the ribosome, even if translation

was not yet completed (ref. (Goder and Spiess, 2003 and Figure 4). This period of signal orientation has certainly passed by the time of in vitro cross-linking. Stop-transfer sequences were similarly found to be cross-linkable in defined positions at the interface of the translocon and lipid membrane as long as the nascent chain remained attached to the ribosome (McCormick et al., 2003). Different sequences were detected in different positions (in some cases adjacent also to TRAM), suggesting that transmembrane segments tethered to the translocation complex associate at various places with the outside of the pore complex (McCormick et al., 2003). In contrast, our current results indicate that the signal is in contact with the lipid bilayer during topogenesis, i.e., while signal reorientation takes place.

Recently, the crystal structure of the SecYEG translocation complex of *Methanococcus jannaschii* has been determined (Van den Berg et al., 2004), which is homologous to the mammalian Sec61αβγ complex. It suggests that the translocation pore is formed by a single Sec61 complex, rather than by three or four complexes as previously proposed on the basis of electron microscopy of the yeast and mammalian translocons (Menetret et al., 2000; Beckmann et al., 2001). As a consequence, the hydrophilic pore generated upon opening of the channel is likely to be less spacious than previously expected (Hamman et al., 1997). In any case, however, there is no obvious hydrophobic surface lining the inside of the pore that could serve as a static binding site for apolar signal sequences.

The translocon is organized in two halves (transmembrane helices 1-5 and 6-10) with a lateral exit site toward the lipid membrane between helices 2 and 3 and 7 and 8 (Van den Berg et al., 2004). It seems that a hydrophobic environment becomes accessible to an entering signal only when the exit site opens, for example, due to thermal motion within the structure. At such a moment, the h-domain of the signal might exit into the surrounding lipid where it will form a helix, optimizing intramolecular hydrogen bonds. Similarly, the signal might return into the hydrophilic channel pore where the peptide is unlikely to remain helical. As a flexible chain, it may invert its orientation due to charge interactions. We propose that the signal is in an equilibrium between a “bound state” as a transmembrane helix outside the translocon and a flexible, free state within the translocon (Figure 7). An intermediate state of a signal helix intercalating between the transmembrane domains of the translocon’s exit site is likely to be unstable, particularly also

considering the multitude of possible signal sequences. The translocon therefore allows lateral equilibration of the signal between an aqueous and a transmembrane environment. The translating ribosome may facilitate the transient lateral opening of the pore. Upon termination of translation, re-entry of the signal may be inhibited, resulting in the observed block of further topology changes.

MATERIALS AND METHODS

DNA constructs.

The starting construct encoding H1ΔQLeu16 has been previously described (Wahlberg and Spiess, 1997). Pairs of leucine residues in the hydrophobic oligo-leucine core of the signal sequence were replaced with other amino acids by polymerase chain reaction using Vent polymerase (New England Biolabs) and appropriate mutagenic oligonucleotide primers similar to the procedure used in ref. (Rosch et al., 2000). For example, to construct H1ΔQL16WW1/2, the sense oligonucleotide CGGGGTACCATGGGACCGCAGTGGTGGCTTTTGC-TGCTGCTC was used (the Kpn I cloning site and the mutated codons for tryptophan are underlined) in combination with a reverse primer corresponding to a sequence in the plasmid vector and with the cDNA of H1ΔQL16 as the template. Mutations in the C-terminal half of the signal sequence were generated with an antisense primer such as CCGGGATCCCAAGAGCAACAGCAGGAGCCTCCTGAGGAGCAGC for H1ΔQL16WW9/10 (the BamH I cloning site and the mutated codons for tryptophan are underlined) in combination with an upstream primer complementary to the vector sequence. With the BamH I site, the polymerase chain reaction products were ligated to the downstream cDNA sequence of H1. The final constructs were subcloned into the expression vector pECE (Ellis et al., 1986). To test the dependence of protein topology on the length of the protein, the sequence C-terminal to the signal sequence of normally 230 amino acids in H1ΔQLeuWW1/2, H1ΔQLeuWW7/8, and H1ΔQLeuWW9/10 was replaced with truncated or extended versions of 110, 170, 290, or 460 residues as described by

Goder and Spiess (Goder and Spiess, 2003). All constructs were verified by sequencing.

Cell culture, transfection, and immunoprecipitation.

Cell culture reagents were from Life Technologies, Inc. COS-7 cells were grown in modified Eagle's minimal essential medium supplemented with 10% fetal calf serum, 2 mM L-glutamine, 100 units/mL penicillin, and 100 μg/mL streptomycin at 37 °C with 7.5% CO_2. Transient transfection was performed in six-well clusters with lipofectin (Life Technologies, Inc.) according to the manufacturer's instructions. The cells were processed the second day after transfection. For in vivo labeling, transfected cells were incubated for 40 min in methionine-free medium, labeled for 40 min at 37 °C with 100 μCi/mL [^{35}S]methionine, transferred to 4 °C, washed twice with phosphate-buffered saline, and finally lysed and immunoprecipitated using a rabbit anti-serum directed against a synthetic peptide corresponding to residues 277-287 near the C-terminus of ASGP receptor H1 (anti-H1C). The immune complexes were isolated with protein A-Sepharose (Amersham Pharmacia Biotech) and analyzed by SDS-polyacrylamide gel electrophoresis and autoradiography. Quantitation was performed using a phosphorimager (Molecular Dynamics Inc.). To determine the fraction of products with the N_{cyt}/C_{exo} topology, the intensity of the glycosylated forms with one and two glycans in the percentage of the total of all glycosylated and unglycosylated forms was calculated. This value proved to be independent of transfection efficiency, which may vary somewhat between experiments.

Alkaline extraction and protease protection assays.

Alkaline extraction was performed as previously described (Wessels et al., 1991). To reduce the viscosity of the sample, the cells suspended in alkaline solution were pipetted up and down through a 25 gauge needle to shear the DNA before it was loaded onto the sucrose cushion. For the protease protection assay, labeled cells were incubated at 4 °C with hypotonic swelling buffer (15 mM Hepes/KOH (pH 7.2) and 15 mM KCl) and scraped with a rubber policeman.

Aliquots were incubated without protease or with 100 μg/mL trypsin in the presence or absence of 0.5% Triton X-100 for 30 min at 4 °C. Trypsin was then inhibited by addition of 500 μg/mL soybean trypsin inhibitor before immunoprecipitation and analysis by SDS gel electrophoresis and autoradiography. As control constructs, HC, encoding the cleavable signal of influenza hemagglutinin fused to the C-terminal portion of H1 (Schmid and Spiess, 1988), and wild-type H1 were used.

OUTLOOK

We showed that a type II protein without a N-terminal domain could contact lipids surrounding the translocon during topogenesis. The translocon therefore allows lateral equilibration of the signal between an aqueous and a transmembrane environment. Upon termination of translation, re-entry of the signal may be inhibited, resulting in the observed block of further topology changes. In this model, the protein shows two states, a "free" state inside the translocon and a "bound" state when the TM is bound to lipids surrounding the translocon. Van den Berg et al (2004) showed by crystallization that the translocon is sealed by a plug in its idle state but nothing is known about this plug role with a translating ribosome. When viewed from the cytosol, the channel has a square shape with three main subunits, the α/γ/β-subunits. On the extracellular side, an additional short helix called the plug, fills the center of the cavity (Van den Berg et al., 2004).

The two large, connected cavities in the structure are likely to form the pore through which proteins pass across the membrane, with the plug acting as a gate. Opening the pore requires the plug to move into a cavity present at the back of the molecule. Recently, Rapoport et al proposed a model for TM domain orientation based on the structure of the Sec61 translocon (Rapoport et al., 2004).

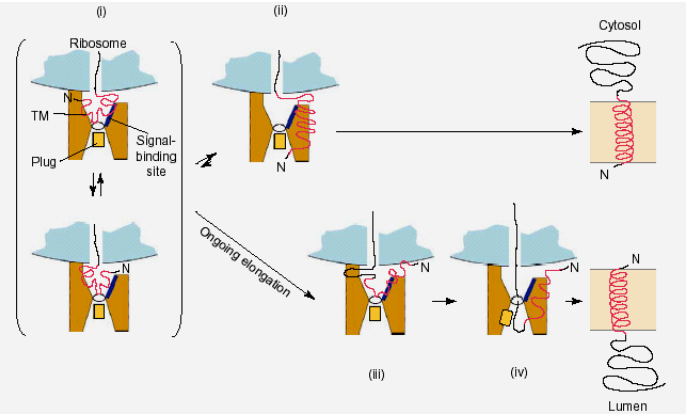

Figure 1: Model of the orientation of transmembrane segments. **(i)** A hydrophobic TM has emerged from the ribosome and binds to the signal-binding site at the front of the channel. The TM can bind reversibly in several conformations and in two different orientations. **(ii)** If the hydrophobic sequence of the TM is long and the N terminus is not retained in the cytosol, it will rapidly flip across the membrane. The plug (yellow) will be displaced only transiently. **(iii,iv)** If the N terminus is retained in the cytosol and the polypeptide chain is elongated, the C terminus can translocate across the membrane. If a polypeptide chain is present in the pore, the plug will be prevented from returning to its closed-state position (Rapoport et al., 2004).

According to the model, If the signal has emerged from the ribosome, the N terminus can be translocated (Figure 1 ii). Translocation of the N terminus requires only a brief displacement of the plug. Once enough of the segment that follows the signal has emerged from the ribosome, the plug moves away completely and the carboxyl (C) terminus could be translocated. In this model, signals have an opportunity to re-orient by multiple binding and release events, and they could even invert their orientation across the membrane (Goder and Spiess, 2003). This model is consistent with the finding that glycosylation of a lumenal domain can fix the orientation of a membrane protein (Goder et al., 1999), and that downstream signals of a multispanning membrane protein can invert a previously integrated TM domain (Nilsson et al., 2000).

We can hypothesize that an N-terminal extension may open the plug during topogenesis and affect the bound state of the TM domain with its bound state, i.e. the lipids. To test this theory, we made different constructs with an N-terminal tail using the following constructs, H1ΔQL16WW7, H1ΔQL16WW9, and H1ΔQL16WW13, previously used to show the contact with lipids. The N-terminal sequence is composed of a repetition of **G**lycine-**S**erine (GS) with final lengths of five, ten or fifteen amino acids. This unnatural, hydrophilic, and flexible N-terminal

extension will not affect topology by folding. As previously described, the folding of sequences N-terminal to an internal signal may sterically prevent translocation of the N-terminus irrespective of the charge distribution (Denzer et al, 1995). Normally, a cotranslational translocation ensures the direct transfer of sequences downstream of cleavable signals and signal-anchors across the membrane and prevents exposition to and folding in the cytosol. In contrast, N-terminal sequences placed before an internal signal are exposed to the cytosol before the targeting signal emerges from the ribosome. Folding of these domains may thus affect their translocation competence and favor retention of the N-terminus in the cytosol. It was indeed demonstrated using mutants of our model protein, the ASGP receptor, that truncation of the N-terminal domain allowed almost complete N-terminal translocation, whereas the full coding sequence of dihydrofolate reductase or a small zinc finger domain fused at the N-terminus hindered or even blocked it (Denzer et al., 1995). These results confirmed that the polypeptide chain needs to be unfolded for translocation and that the folding properties of the N-terminal domain influence protein orientation. Hence, for our constructs the only changing parameter is the size of the N-terminal part.

All constructs were transfected in COS-7 cells, labeled with [35S]-methionine, and analyzed by immunoprecipitation, SDS-gel electrophoresis, and auroradiography. The results are presented in Figure 3a and the quantitation in Figure 3b.

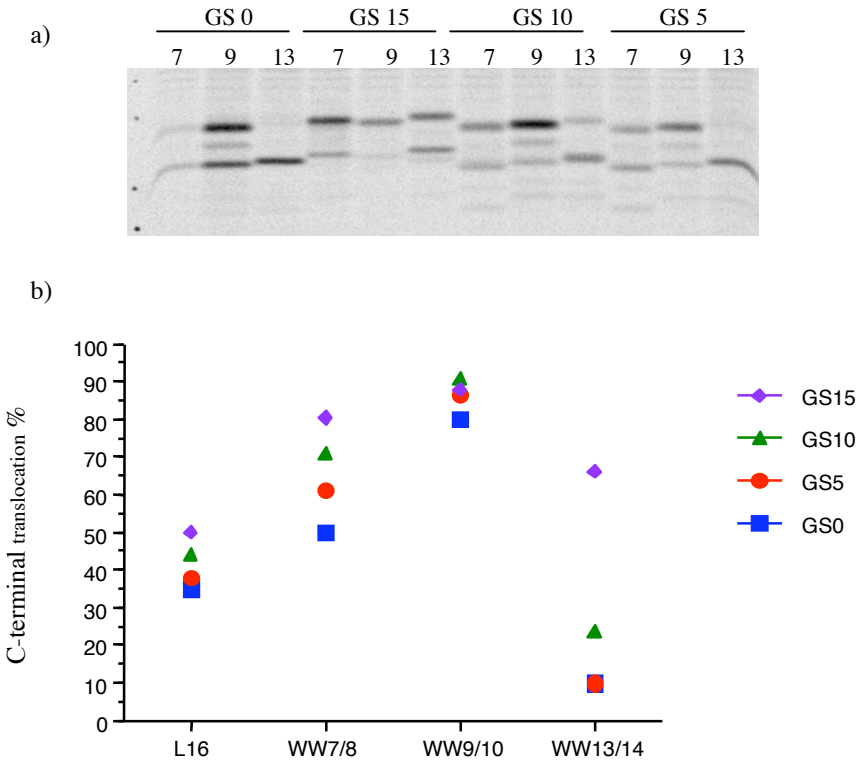

Figure 3: a) SDS page gel with all GS(X)L16ΔQWW(tails) constructs. b) C-terminal translocation plotted against the length of the different GS extension for the constructs GS(X)H1ΔQL16, GS(X)H1ΔQL16WW7, GS(X)H1ΔQL16WW9, and GS(X)H1ΔQL16WW13.

These first results show an effect on C-terminal translocation for each type of construct depending on the size of the added N-terminal tail. All the constructs showed an increase of C-translocation with an increase of the GS sequence length but for the construct GS(X)H1ΔQL16WW13, the effect is more dramatic. For this construct, the C-translocation started around 10% in absence of the glycine-serine tail and for a GS tail of 5 amino acids. Then the C-terminal translocation increased a little bit more for the 10 amino acids tail and reached 65% for the biggest tail of 15 amino acids. The N-terminal tail had a huge effect on final topology. To interpret this effect we should be aware of the reference construct GS(X)H1ΔQL16. This control construct showed that the increase of C-translocation is mainly due to the tryptophan position in the TM and not only to the size of the N-terminal extension.

This might suggests that the N-terminal extension affects the affinity between the TM domain and its partner during the insertion process , i.e. the lipids. In this case the contact of the TM with lipids might be reduced and the protein is able to turn more easily. An experiment to prove this theory would be to look at the

kinetics of inversion of these constructs by varying their C-terminal sequence lengths. This experiment would tell us more directly whether the kinetics of inversion are increased by the N-terminal extensions. Another possibility could be a modification of the mobility in the free state due to these extensions, which will also favor inversion. This could be tested by adding different size of N-terminal extensions. Finally, these results could be explained by the a effect of the plug which may hinder the exit of the protein in a N_{exo}/C_{cyt} orientation. To test this hypothesis, we could use yeast and express these constructs in a Sec61 mutant lacking the plug. All those experiments will help to define the open state of the translocation machinery.

Chapter II: Kinetics of topogenesis for a double-spanning protein

Introduction

Multi-spanning membrane proteins are believed to be targeted to the ER membrane by their first hydrophobic signal sequence. According to the simplest model, this most N-terminal signal defines its own orientation as well as the orientation of all subsequent transmembrane anchors, which insert with alternating orientation and therefore do not need to contain any additional topogenic information. Evidence for this "linear insertion model", initially proposed by Blobel (1980) has been provided by in vitro studies using chimeric proteins with two to four transmembrane segments separated by about 50-200 residues from each other (Wessels and Spiess, 1988; Lipp et al., 1989). The results showed that signal-anchors (normally N_{cyt}/C_{exo}) insert as stop-transfer sequences (N_{exo}/C_{cyt}) depending only on their position relative to the preceding signal sequence.

However, there is also strong evidence against that model of linear insertion. Statistics show that internal transmembrane sequences also insert according to the charge rules, although less stringently in eukaryotic proteins than the most N-terminal signal (von Heijne, 1986; von Heijne, 1989). Experimentally, insertion of clusters of positive charges into short exoplasmic loops of model proteins caused individual hydrophobic domains not to insert at all (Gafvelin and von Heijne, 1994; Gafvelin et al., 1997). Deletion of individual membrane-spanning segments in bacterial proteins did not necessarily affect the topology of the downstream transmembrane domains (Bibi et al., 1991; McGovern and Beckwith, 1991). Similarly, inversion of the charge difference of the first signal of the glucose transporter Glut1 did not affect the topology of the rest of the molecule, but prevented insertion of the first signal (Sato et al., 1998). These studies provided evidence that multi-spanning membrane proteins containing topogenic information throughout their sequence and that insertion is not always strictly linear.

Hydrophilic sequences separating transmembrane anchors in natural proteins are frequently much shorter than those used in the studies supporting the linear insertion model.

Goder and Spiess (1999) systematically analyzed the topogenic influence of an internal signal sequence in relation to its distance from a conflicting N-terminal signal sequence. When the signals were separated by more than 60 residues, linear insertion was observed. With shorter spacers, an increasing fraction of proteins inserted with a translocated C-terminus, as dictated by the second signal. This suggested the following mechanism: the first signal initially translocated its C-terminal sequence. When the second signal entered the translocon, it inverted its orientation due to its flanking charges and induced the translocated spacer to be pulled back to the cytosol.

This model is supported by the effect of glycosylation sites inserted into the spacer sequence: the spacer glycosylation generated more products with a translocated spacer and a cytosolic C-terminus suggesting that the added glycan inhibited retrotranslocation of the spacer sequence. This model also predicts that at least two polypeptide segments need to be simultaneously in the translocation channel.

Recently, Goder ans Spiess (2003) found that kinetics of topogenesis can be studied in vivo. They used a series of diagnostic constructs with an hydrophobic N-terminal and a signal sequence that generated mixed topologies and observed that topogenesis depends on the length of the polypeptide. Further changes in topology appear to be blocked when translation is terminated. This phenomenon opens the possibility to best the above model on the insertion of polypeptides with two competiting signals in more details. Using C-terminal extensions of different lengths following the second signal, topogenesis will be stopped after different times. This should allow to observe a change in topology from the glycosylation pattern of the resulting proteins. In particular, the kinetics of topology changes and the extrapolated start point of inversion is of interest. In addition, it has been observed that topogenesis of a single-signal protein stopped spontaneously around 50 s after signal insertion. This time window for protein reorientation appeared to be constant for several variations, in flanking charges or hydrophobicity, of the signal-anchor. Based on this, it was speculated that this might be a general property

of the translocation apparatus to terminate "unsuccesful" (slow) attempts to acquire a final topology, independent of the substrate proteins. Analysis of the constructs with two conflicting signals, which are quite different frome the subtrates analysed before, could test this concept and might reveal whether this time window of topogenesis starts with the first or perhaps restarts with the second signal.

For these experiments, we used a series of constructs consisting of the H1 subunit of the asialoglycoprotein receptor with an N-terminally fused cleavable signal sequence from hemagglutinin of human Influenza A virus (HA). There is thus a spacer of 40 amino acids separating the cleavable signal of hemagglutinin from the H1 signal-anchor. In their normal environment, both signals mediate translocation of the C-terminal part of the polypeptide across the membrane as illustrated in Figure 1.

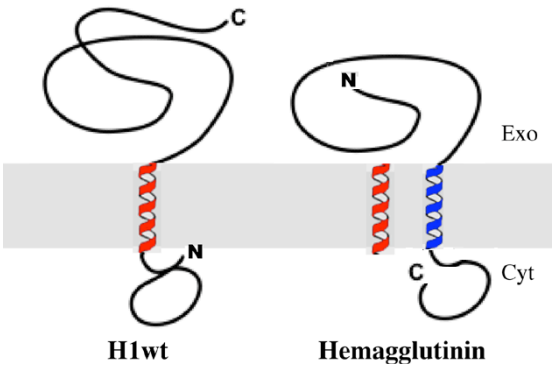

Figure 1: Natural topology of the signals of H1 and hemagglutinin. In their wild-type context the signal-anchor of H1 and the cleavable signal of hemagglutinin translocate their C-terminus across the membrane. Both signals thus have an N_{cyt}/C_{exo} orientation (labeled in red). Hemagglutinin has an additional transmembrane sequence with an N_{exo}/C_{cyt} orientation (labeled in blue).

In the context of the fusion constructs, one signal has to insert in the opposite orientation (N_{cyt}/C_{exo}) to generate a type I membrane protein (Figure 2).

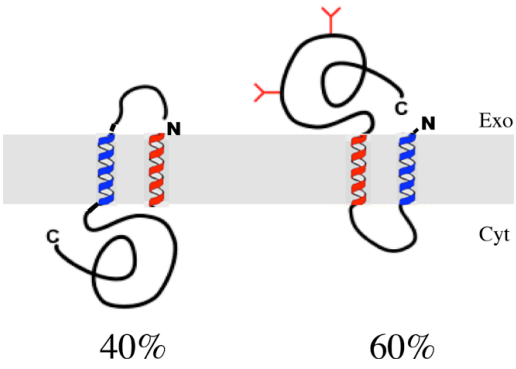

Figure 2: Schematic drawing of the constructs Hg40H1(X). Hemagglutinin signal was fused to wild-type H1 with a spacer of 40 amino acids between the two signals. The construct was found to be inserted with a loop-translocated topology (40%) as well as with a C-terminus translocated orientation (60%). The sequences inserted in their unnatural orientation (N_{exo}/C_{cyt}) are marked in blue.

Results and Discussion

We made Hg40H1(X) constructs with tail lengths (the length from the second signal to the end of the construct) of 50, 60, 110, 170, 230, 290, and 350 amino acids to study different time points in topogenesis (Figure 3a). COS-7 cells were transfected with these constructs and labeled for 40 min with [35S]methionine and analyzed by immunoprecipitation with an antibody against the C-terminal part of H1 (anti-C1), SDS-gel elecrophoresis and autoradiography (Figure3b).

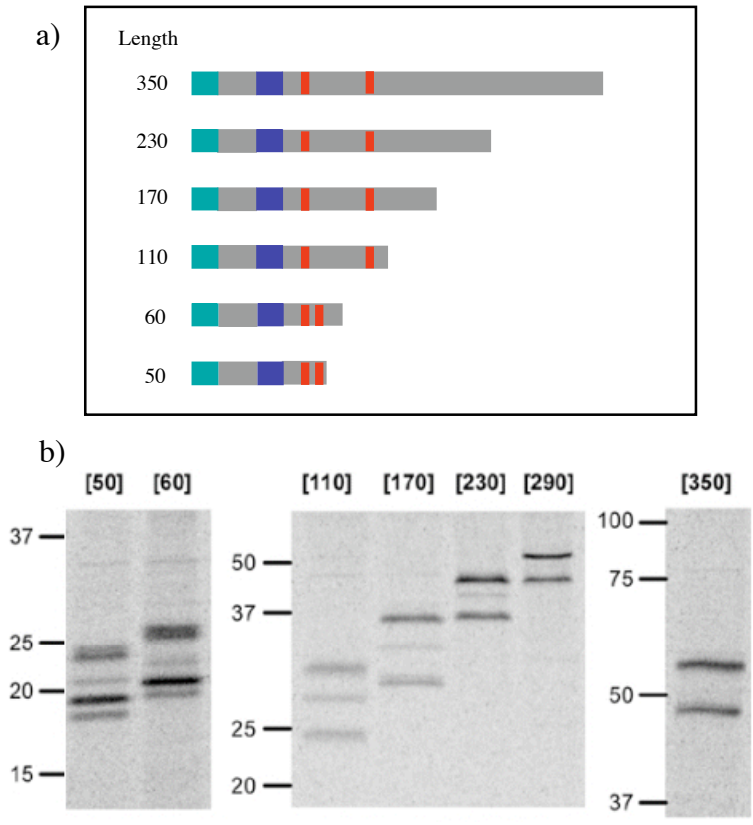

Figure 3: a) Schematic representation of Hg40H1(X). X is the length of the C-terminal sequence following the second signal in amino acids (WT= 230). The signal of Hg (green) and the signal of H1 (blue) are represented in squares. The glycosylation sites are represented in red lines . b) Autoradiography of the SDS-gel electrophoresis with constructs Hg40H1(X). Different percentage of gels were used depending on the constructs lengths. The marker sizes are in kDa.

The longest constructs, 290 and 350, as expected, produced two major forms, the upper one represents the twice glycosylated form indicating that the C-terminal sequence was translocated into the ER lumen. The lower band corresponds to the unglycosylated form when the loop was translocated.

For the shorter constructs an additional form in between the other two appeared. This represents the once glycosylated form, when the glycosylation site close to the membrane was not effectively modified, as it has been observed previously. For the constructs with 50 and 60 amino acid tails (for which a higher percentage gel was used), additional bands were resolved (Figure 3b). There are several possibilities to explain these results. These extra bands are not due to rapid degradation of the constructs, because they were not altered in intensity with shorter labeling times (data not shown). The most obvious test is to analyse the glycosylation state of these bands by deglycosylation using the endoglycosidase endo-ß-N-acetylglucosaminidase H (endo H).

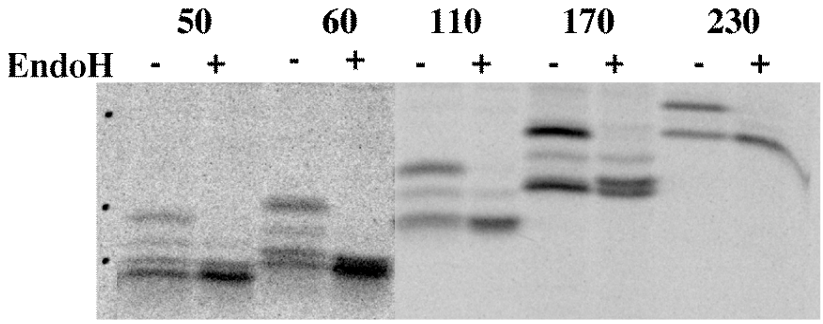

Figure 4: Glycosylation analysis of Hg40H1(X) constructs by 10 to 15% acrylamide gradient gel electrophoresis. Proteins were labeled and analysed as Figure 3. The latter two forms are sensitive to deglycosylation by endoglycosidase H (+)

For this experiment we used a gradient gel of 10 to 15% acrylamide to obtain high resolution for all constructs (Figure 4). Under these conditions, separation of an additional form of lower molecular weight could be achieved also for the constructs with 110 and 170 C-terminal residues. The phenomenon is therefore not specific for the shortest constructs, but is simply more difficult to detect for larger proteins. Endo H treatment showed also that the upper two bands were glycosylated whereas the lower two were not. These results showed that these extra bands are not due to an unexpected glycosylation. Interestingly, with endo H treatment the

upper bands appeared to collapse with the higer band of the non glycosylated forms for the contructs Hg40H1(110) and Hg40H1(170) and with the lower band for the constructs Hg40H(40) to Hg40H1(60).

To understand these results we should precisely determined the nature of the two non-glycosylated bands. A further possibility to account for the additional unglycosylated form is proteolytic processing at the N-terminus (the C-terminal sequence can not be affected, because we used an antibody directed against the very C-terminus of the protein). Indeed, for the topology with a translocated spacer, we expected the hemagglutinin signal to be removed by signal peptidase. Our constructs contain the entire signal peptide including the cleavage site of hemagglutinin sequence (Figure 5a).

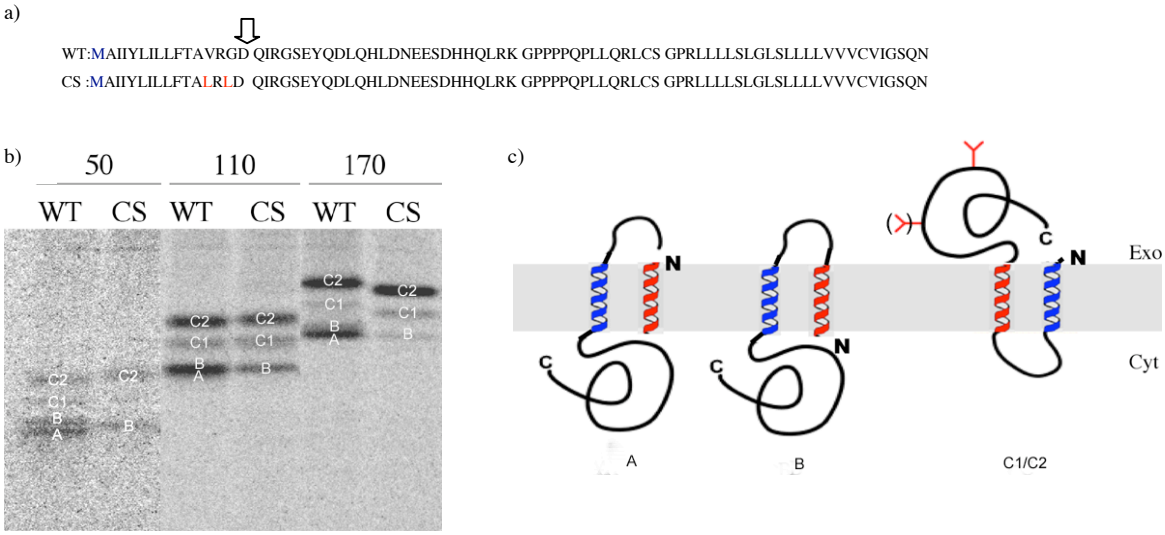

Figure 5: Signal cleavage analysis. a) Representation of the wild type sequences constructs Hg40H1 (WT), the mutated cleavage site construct Hg40H1CS (CS). The cleavage site is indicated by an arrow for the wild type construct. The modified amino acids are labeled in red for the mutant cleavage sitre contruct CS. b) autoradiography of a SDS page electrophoresis. Constructs pHg40H1(X) are labeled WT and pHg40H1CS(X) are labeled CS. c) Model of insertion in the ER membrane. Type A represents the loop translocation with a cleavage after the hemagglutinin signal, type B represents the loop translocation without any cleavage after the hemagglutinin signal and type C1/C2 represent the C-terminal sequence translocation with one or two site glycosylated.

To test this hypothesis, we produced a control construct with a mutated cleavage site (Figure 5a). We decided to use the three following constructs, Hg40H1(50), Hg40H1(110) and Hg40H1(170) where the cleavage site was inactivated by mutation of glycine to leucine at position –1 and of valine to leucine

at position –3. As described by the the so-called (-3,-1) rule (Perlman and Halvorson, 1983; von Heijne, 1983), signal peptidase cleavage requires small amino acids at positions –1 (Ala, Gly or Ser) and –3 (Ala, Gly, Ser, Val or Ile). The mutant constructs were named Hg40H1CS(50), Hg40H1CS(110) and Hg_{40}H1CS(170).

These constructs were transfected, labeled with [35S] methionine and analyzed by immnunoprecipitation, SDS-gel electrophoresis, and autoradiography as shown in Figure 5b. We obtained four different bands, A and B which represent the unglycosylated forms, C1 and C2 which are the one- and two-fold glycosylated forms.

Surprisingly, not always the mobility of the polypeptides corresponds directly to the expected size. While the no cleavable form of Hg40H1CS(50) is larger than the Hg40H1(50) construct. In the case of Hg40H1CS(170) construct, it seems to be slightly faster than those of Hg40H1(170). These are most likely effects of the mutations on SDS binding and/or the shape of the SDS complexes.

Most importantly, however, the cleavage site mutated constructs (CS) always produced a single unglycosylated form. The two unglycosylated forms of Hg40H1 constructs are then directly the results of incomplete signal cleavage as represented in Figure 5c. The unglycosylated cleavage site mutants showed slower mobility than the corresponding constructs without the hemagglutinin signal (in the case of Hg40H1CS(170) after correction for the general shift induced by the mutation). These results imply that the peptidase cut the construct during topogenesis, so the hemagglutinin signal could be cleave during the time the topology and then, the C-terminal sequence could be translocated into the ER lumen. The results also showed that the cleavage is not necessary to obtain either one or the other orientation, because both orientations were found even if the cleavage site was mutated. The Hg40H1CS constructs also suggested a slight increase in glycosylated forms compared to the unglycosylated ones. The cleavage of the protein might influence the final topology of the protein by increasing the spacer-translocation by hindering inversion after cleavage. These results are just preliminary data and should be confirmed by statistical analysis.

For the unglycosylated forms, which have translocated their loops across the membrane, signal cleavage was not efficient. In any case, for the quantitation of

spacer-translocation vs C-translocation topologies, both unglycosylated forms belong to the spacer translocation topology.

Quantitation of the products of the Hg40H1(X) series of constructs (including those shown in Figure 3) yielded the length dependence of topology shown Figure 6. The lengths of the C-terminal sequence defines the time of translation from the moment when the second signal emerges from the ribosome to the end of translocation (and thus of topogenesis). This is also the time during which, according to our model, protein reorientation may take place.

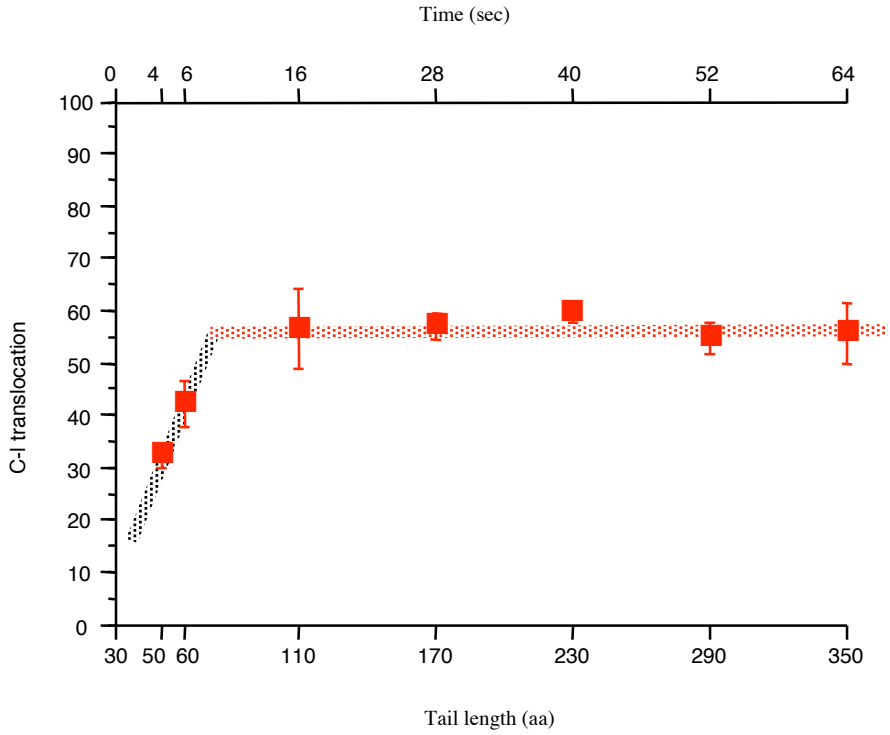

Figure 6: Length dependence of topologies. The average of C-terminal translocation is plotted against tail length of Hg40H1(X) and time with standard deviation (n=5). Translation time was set to 5aa/s. (S Gander and M Higy).

We obtained an increase of C-terminal translocation for the constructs Hg40H1(50), Hg40H1(60) and Hg40H1(110). For all longer constructs than Hg40H1(110), it reached a plateau around 60% C-terminal translocation. These results support the model that protein integration indeed starts with translocation of the spacer sequence, but after the second signal has emerged from the ribosome, it continues by reorientation of the construct.

The data in Figure 6 can also be plotted as a function of the translation time from the moment when the second signal has just emerged from the ribosome to insert into the translocon (i.e. with 30 residues following the hydrophobic core of the signal still hidden within the ribosome; (Matlack and Walter, 1995; Morgan et al., 2000) until the ribosome has reached the stop codon (Figure 6, upper scale). The calculation is based on a translation rate of 5 amino acids/s, as determined for cultured cell lines (Hershey, 1991). The 0% C-terminal translocated topologies should be obtain when the entire second signal has emerged from the ribosome. Our results showed an additional 5 s before the 0% C-terminal translocation, these results are a matter of the inaccuracy of the method, i.e. the quantification of the bands on the gel, or that the second signal already influences topogenesis when it has only partially emerged from the ribosome. Anyway, an extrapolation based on two data points is not accurate and we need more data to define precisely our time window.

For our double-spanning membrane protein further topology changes appear to be stopped after less than 16 s following the emergence of the second signal. Even if one argues that the timer for topogenesis starts after the first signal has emerged from the ribosome which would increase the measured time by 12 additional seconds, topogenesis would still finish less than 27 s after emergence of the first signal. The time window during which topology can change seems to be clearly shorter than previously observed for model proteins with a single N-terminal signal-anchor.

In the case of a single N-terminal signal-anchor the signal inversion can only be observed during the first 50 s after insertion into the translocon (Goder and Spiess, 2003). If it takes longer than this period for the protein to be completed, either because of the length of the protein or because of reduced translation rate, the resulting ratio of topologies does not change further; topogenesis appears to come to a halt. The time when this occurs is not significantly affected by alterations in the signal, neither of the flanking charges nor of the apolar core. It is thus a property of the translocation machinery to commit this type of substrate to its current orientation 50 s after engagement of the protein with the translocon. In the case of a double spanning protein, we obtained a different time window which would imply

that the time window is dependent of the substrate type and is not a constant of the translocation machinery for all subtrates.

Material and Methods

DNA constructs

For the model protein, the cleavable transmembrane signal of hemagglutinine was fused at 40 amino acids from the N-terminus of the wild type transmembrane signal of H1 (Goder and Spiess, 1999). The constructs were named to indicate the origin of the first signal, the length of the hydrophilic spacer sequence, the origin of the second signal and the length of the C-terminal tail (e.g., Hg40H1(230)). For the smallest constructs we added four methionines at the C terminal part to increase the signal. For example, the contract pHg40H1(110)Met4 were made using the constructs pHg40H1(230) as a wild type, and pH1ΔQMet4(110) were cut with BamH1 and EcoRI. In the case of pHg40H1(60)Met4, the PCR reaction was performed using pHg40H1(110)Met4 as a template, with BstXI50-a as an anti-sense primer in combination with an upstream primer complementary to the vector sequence. The PCR product pHg40H1(110)Met4 were cut with HindIII and BstXI. and ligated into the cut vector. For pHg40H1CS(50), pHg40H1CS(110) and pHg40H1CS(170) we used pHg40H1(50), pHg40H1(110) and pHg40H1(170) as a template with the sense oligo HAuncut-s -GCGAAGCTTACCATGGCCATCATTTATCTCATTCTCCTGTTCACAGCACTGAGACTGGACCAG- in combination with a reverse primer complementary to the vector sequence. The PCR product was cut with HindIII and EcoRI. The final constructs for in vivo expression were subcloned into the vector pECE (Ellis et al. 1986) and verified by sequencing.

Cell culture, transfection, and immunoprecipitation

Cell culture reagents were from Life Technologies, Inc. COS-7 cells were grown in modified Eagle's minimal essential medium supplemented with 10% fetal calf serum, 2 mM L-glutamine, 100 U/ml penicillin and 100 μg/ml streptomycin at 37°C with 7.5% CO_2. Transient transfection was performed in 6-well clusters with lipofectin (Life Technologies, Inc.) according to the manufacturer's instructions. The cells were processed the second day after transfection. For in vivo labeling, transfected cells were incubated for 40 min in methionine-free medium, labeled for 40 min at 37°C with 100 μCi/ml [^{35}S] methionine, transfered to 4°C, washed twice with phosphate-buffered saline, and finally lysed and immunoprecipitated using a rabbit anti-serum directed against a synthetic peptide corresponding to residues 277-287 near the C-terminus of the ASGP receptor H1 (anti-H1C). The immune complexes were isolated with protein A-Sepharose (Amersham Pharmacia Biotech) and analyzed by SDS-polyacrylamide gel electrophoresis and autoradiography. Quantitation was performed using a phosphorimager (Molecular Dynamics Inc.). To determine the fraction of products with N_{cyt}/C_{exo} topology, the intensity of the glycosylated forms in percent of the total was calculated. This value proved to be independent of transfection efficiency, which may vary somewhat between experiments.

Endo H treatment

After the last washing step 50µl TNET were left in the tube and 50µl endo H buffer and 0.005 units endo H (50mM K_3PO_4, 25mM EDTA, 2% Triton X-100, 0.2% SDS, 1% 2-mercaptoethanol) were added. The sample was then incubated for 1h at 37°C. After adding of 50µl sample buffer it was immunoprecipitated.

OUTLOOK

We conclude from our experiments that the reorientation after insertion of the second signal is surprisingly rapid and the time window for reorientation is unexpectedly short. Both of these feature make it more difficult to characterise the reorientation process. To improve the quality of our data and determine the time window more precisely, one might prepare additional constructs with tail lengths between 60 and 110 amino acids. Constructs shorter than 50 amino acids following the second signal carry the risk of inefficient glycosylation.

The process of reorientation could be slowed down by increasing the hydrophobicity of the second signal. This could be done by replacing the transmembrane domain of H1 by an oligo-leucone signal. Then, to investigate the effect on the time window of topogenesis one can make constructs with a tail length of 50-230 amino acids. Alternatively, the charges flanking the hydrophobic core of the second signal could be altered. Reducing the N-terminal positive charges should also slow the reorientation. In contrast, increasing the N-terminal positive charges and/or reducing the C-terminal charges is expected to accelerate protein inversion significantly. This could be useful to test whether all polypeptides can be induced to invert or whether there is a population of polypeptides that are unable to reorient, for example because chaperones may have associated with the translocated spacer segment.

General Conclusion

Most membrane proteins in eukaryotic cells are integrated into the membrane of the ER before they are transported in vesicles to the Golgi apparatus, to other compartments of the secretory pathway, and to the endo-lysosomal system. Whereas secretory proteins cross the membrane completely, only some regions of a membrane protein are transferred across the membrane; others remain in the cytosol or stop within the lipid bilayer and must be oriented. Because membrane proteins can have one or more transmembrane segments, the integration machinery must be able to create various topologies, each of which is somehow dictated by the sequence of the protein.

Most membrane proteins are cotranslationally integrated into the endoplasmic reticulum membrane by the Sec61 translocation complex. They are targeted to the translocon by hydrophobic signal sequences which induce the translocation of either their N- or C-terminal sequence. In secretory and signal-anchor membrane proteins, hydrophobic N-terminal signals initially insert head-on before they invert their orientation to translocate the C-terminus.

Several factors have been shown to determine the orientation of the signal in the membrane: charged residues flanking the hydrophobic core of the signal influence orientation according to the "positive-inside rule" (Hartmann et al., 1989; Beltzer et al., 1991; Parks and Lamb, 1991); local charges at the translocation apparatus must be involved in orienting the signal sequence; folding of hydrophilic sequences N-terminal to a signal sterically hinders N-terminal translocation irrespective of the flanking charges (Denzer et al., 1995); hydrophobicity of the core of the signal sequence (the h-domain). Strongly hydrophobic signals were observed to insert with N_{exo}/C_{cyt} orientation even when the flanking charges were more positive at the N-terminus (Sakaguchi et al., 1992; Wahlberg and Spiess, 1997; Rosch et al., 2000); and finally the interaction with partners during topogenesis could also influence the final topology of the protein.

To identify the proteins that decode the topogenic information, one can use a genetic approach by investigating protein topogenesis in the yeast *S.cerevisae*. A variety of mutants of the yeast translocation machinery components has been used

to study protein antero-and retro-translocation (e.g., Pilon et al., 1997; Pilon et al., 1998; Gillece et al., 2000; Wilkinson et al., 2000). In our laboratory, we monitored the effect of Sec61p mutants on signal orientation, by using model proteins that insert with mixed orientations in wild-type yeast cells (Goder et al., 2004).

Sophisticated crosslinking experiments identified molecules in contact with various parts of substrate proteins (Martoglio and Dobberstein, 1996). These mostly static data together with the recent first crystal structure of a protein-conducting channel (Van den Berg et al., 2004) lead to new insights into the highly dynamic process of protein topogenesis and membrane integration. We propose that the signal is in an equilibrium between a "bound state" as a transmembrane helix outside the translocon and a flexible, "free state" within the translocon. The translocon therefore allows lateral equilibration of the signal between an aqueous and a lipidic environment. The translating ribosome may facilitate the transient lateral opening of the pore.

To probe the environment of the signal while its orientation is determined, different hydrophobic residues were inserted at various positions throughout a uniform oligo-leucine signal sequence. The resulting topologies revealed a strikingly symmetric position dependence specifically for the bulky aromatic amino acid, reflecting the structure of a lipid bilayer. Maximal N-translocation was observed when the guest residues were placed at the N- or C-terminus of the hydrophobic sequence or in the very center, corresponding to the positions of highest expected affinity for the bilayer of the signal sequence as a membrane-spanning helix. The results support the model that during topogenesis the signal sequence is exposed to the lipid membrane *in vivo* (Results in Chapter I of this thesis), and that the signal may partition between the aqueous pore and the lipid environment. This is consistent with a very recent report by Hessa et al (2005) where stop-transfer activity appears to depend on a direct protein-lipid interaction. As in our study, tryptophans strongly reduced membrane insertion when placed centrally, but became much less unfavourable as they are moved apart. Their results provided further support for the idea that protein–lipid interactions are central to the recognition of transmembrane helices by the translocon.

The signal-lipid interaction was determined for the simplest single-spanning model protein without N-terminal hydrophilic dromain. To generalize our findings

to more complex proteins, we added an N-terminal extension to the constructs that we have used in our study. This resulted in an increase of the C-terminal translocation with an increase of the extension lengths. The results suggest that the N-terminal extension affects the affinity between the transmembrane domain and its partner during translocation.

An interesting study of the kinetics of TM domain incorporation suggests that for N-terminal signal-anchor proteins the TM domain enters the translocon N-terminus first and then subsequently rearrange over a period of about 50 s to achieve their final TM orientation (Goder and Spiess, 2003). For a double-spanning membrane protein further topology changes appear to be stopped after maximally 27 s if the timer for topogenesis starts after the first signal has emerged from the ribosome (Results Chapter II of this thesis). The time window during which topology can change seems to be surprisingly shorter than previously observed for model proteins with a single N-terminal signal-anchor. This time window is thus different for different types of substrate and is not a constant of the translocation machinery.

The next step is to look at the competition between two signals sequences while topology is determined. As previously suggested by the three-dimensional structure of a protein-conducting channel, in its quiete state, the translocation machinery may acts as a monomer in the incorporation of TM segments into membranes. Futhermore, the TM helix incorporation into membranes is a result of the simple partitioning of potential TM segments between the translocon and the membrane bilayer. Emerging evidence proposes that secondary structure may be formed during the passage of the nascent chain through the ribosome exit tunnel.

All these studies lay a good foundation for an understanding of how the translocon and membrane bilayer work in concert to decode the folding instructions of the membrane protein amino acid sequence. This will be a fruitful area for study in the next several years.

References

Alder, N. N., and Johnson, A. E. (2004). Cotranslational membrane protein biogenesis at the endoplasmic reticulum. J Biol Chem *279*, 22787-22790.

Andersson, H., Kappeler, F., and Hauri, H. P. (1999). Protein targeting to endoplasmic reticulum by dilysine signals involves direct retention in addition to retrieval. J Biol Chem *274*, 15080-15084.

Aridor, M., Bannykh, S. I., Rowe, T., and Balch, W. E. (1995). Sequential coupling between COPII and COPI vesicle coats in endoplasmic reticulum to Golgi transport. J Cell Biol *131*, 875-893.

Arkin, I. T., and Brunger, A. T. (1998). Statistical analysis of predicted transmembrane alpha-helices. Biochim Biophys Acta *1429*, 113-128.

Baerends, R. J., Faber, K. N., Kram, A. M., Kiel, J. A., van der Klei, I. J., and Veenhuis, M. (2000). A stretch of positively charged amino acids at the N terminus of Hansenula polymorpha Pex3p is involved in incorporation of the protein into the peroxisomal membrane. J Biol Chem *275*, 9986-9995.

Bannykh, S. I., Rowe, T., and Balch, W. E. (1996). The organization of endoplasmic reticulum export complexes. J Cell Biol *135*, 19-35.

Barlowe, C. (1998). COPII and selective export from the endoplasmic reticulum. Biochim Biophys Acta *1404*, 67-76.

Beckmann, R., Spahn, C. M., Eswar, N., Helmers, J., Penczek, P. A., Sali, A., Frank, J., and Blobel, G. (2001). Architecture of the protein-conducting channel associated with the translating 80S ribosome. Cell *107*, 361-372.

Beddoe, T., and Lithgow, T. (2002). Delivery of nascent polypeptides to the mitochondrial surface. Biochim Biophys Acta *1592*, 35-39.

Beltzer, J. P., Fiedler, K., Fuhrer, C., Geffen, I., Handschin, C., Wessels, H. P., and Spiess, M. (1991). Charged residues are major determinants of the transmembrane orientation of a signal-anchor sequence. J Biol Chem *266*, 973-978.

Bibi, E., Verner, G., Chang, C. Y., and Kaback, H. R. (1991). Organization and stability of a polytopic membrane protein: deletion analysis of the lactose permease of Escherichia coli. Proc Natl Acad Sci U S A *88*, 7271-7275.

Blobel, G. (1980). Intracellular protein topogenesis. Proc Natl Acad Sci U S A *77*, 1496-1500.

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

Blobel, G., and Dobberstein, B. (1975b). Transfer to proteins across membranes. II. Reconstitution of functional rough microsomes from heterologous components. J Cell Biol *67*, 852-862.

Bogdanov, M., and Dowhan, W. (1999). Lipid-assisted protein folding. J Biol Chem *274*, 36827-36830.

Boman, A. L., Zhang, C., Zhu, X., and Kahn, R. A. (2000). A family of ADP-ribosylation factor effectors that can alter membrane transport through the trans-Golgi. Mol Biol Cell *11*, 1241-1255.

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

Braun, P., and von Heijne, G. (1999). The aromatic residues Trp and Phe have different effects on the positioning of a transmembrane helix in the microsomal membrane. Biochemistry *38*, 9778-9782.

Brodsky, J. L., Goeckeler, J., and Schekman, R. (1995). BiP and Sec63p are required for both co- and posttranslational protein translocation into the yeast endoplasmic reticulum. Proc Natl Acad Sci U S A *92*, 9643-9646.

Chen, H. J., Yuan, J., and Lobel, P. (1997). Systematic mutational analysis of the cation-independent mannose 6-phosphate/insulin-like growth factor II receptor cytoplasmic domain. An acidic cluster containing a key aspartate is important for function in lysosomal enzyme sorting. J Biol Chem *272*, 7003-7012.

Cokol, M., Nair, R., and Rost, B. (2000). Finding nuclear localization signals. EMBO Rep *1*, 411-415.

Cosson, P., and Letourneur, F. (1994). Coatomer interaction with di-lysine endoplasmic reticulum retention motifs. Science *263*, 1629-1631.

Cronshaw, J. M., Krutchinsky, A. N., Zhang, W., Chait, B. T., and Matunis, M. J. (2002). Proteomic analysis of the mammalian nuclear pore complex. J Cell Biol *158*, 915-927.

Crowley, K. S., Liao, S., Worrell, V. E., Reinhart, G. D., and Johnson, A. E. (1994). Secretory proteins move through the endoplasmic reticulum membrane via an aqueous, gated pore. Cell *78*, 461-471.

Dell'Angelica, E. C., Puertollano, R., Mullins, C., Aguilar, R. C., Vargas, J. D., Hartnell, L. M., and Bonifacino, J. S. (2000). GGAs: a family of ADP ribosylation factor-binding proteins related to adaptors and associated with the Golgi complex. J Cell Biol *149*, 81-94.

Denzer, A. J., Nabholz, C. E., and Spiess, M. (1995). Transmembrane orientation of signal-anchor proteins is affected by the folding state but not the size of the N-terminal domain. Embo J *14*, 6311-6317.

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

Deutsch, C. (2003). The birth of a channel. Neuron *40*, 265-276.

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

Doray, B., Ghosh, P., Griffith, J., Geuze, H. J., and Kornfeld, S. (2002). Cooperation of GGAs and AP-1 in packaging MPRs at the trans-Golgi network. Science *297*, 1700-1703.

Dyer, J. M., McNew, J. A., and Goodman, J. M. (1996). The sorting sequence of the peroxisomal integral membrane protein PMP47 is contained within a short hydrophilic loop. J Cell Biol *133*, 269-280.

Egea, P. F., Shan, S. O., Napetschnig, J., Savage, D. F., Walter, P., and Stroud, R. M. (2004). Substrate twinning activates the signal recognition particle and its receptor. Nature *427*, 215-221.

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

Ellis, L., Clauser, E., Morgan, D. O., Edery, M., Roth, R. A., and Rutter, W. J. (1986). Replacement of insulin receptor tyrosine residues 1162 and 1163 compromises insulin-stimulated kinase activity and uptake of 2-deoxyglucose. Cell *45*, 721-732.

Fahrenkrog, B., Koser, J., and Aebi, U. (2004). The nuclear pore complex: a jack of all trades? Trends Biochem Sci *29*, 175-182.

Focia, P. J., Shepotinovskaya, I. V., Seidler, J. A., and Freymann, D. M. (2004). Heterodimeric GTPase core of the SRP targeting complex. Science *303*, 373-377.
Freedman, R. B. (1994). Protein folding. Folding helpers and unhelpful folders. Curr Biol *4*, 933-935.

Funfschilling, U., and Rospert, S. (1999). Nascent polypeptide-associated complex stimulates protein import into yeast mitochondria. Mol Biol Cell *10*, 3289-3299.

Gafvelin, G., Sakaguchi, M., Andersson, H., and von Heijne, G. (1997). Topological rules for membrane protein assembly in eukaryotic cells. J Biol Chem *272*, 6119-6127.

Gafvelin, G., and von Heijne, G. (1994). Topological "frustration" in multispanning E. coli inner membrane proteins. Cell *77*, 401-412.

Gautschi, M., Lilie, H., Funfschilling, U., Mun, A., Ross, S., Lithgow, T., Rucknagel, P., and Rospert, S. (2001). RAC, a stable ribosome-associated complex in yeast formed by the DnaK- DnaJ homologs Ssz1p and zuotin. Proc Natl Acad Sci U S A *98*, 3762-3767.

Gething, M. J. (1999). Role and regulation of the ER chaperone BiP. Semin Cell Dev Biol *10*, 465-472.

Gillece, P., Pilon, M., and Romisch, K. (2000). The protein translocation channel mediates glycopeptide export across the endoplasmic reticulum membrane. Proc Natl Acad Sci U S A *97*, 4609-4614.

Goder, V., Bieri, C., and Spiess, M. (1999). Glycosylation can influence topogenesis of membrane proteins and reveals dynamic reorientation of nascent polypeptides within the translocon. J Cell Biol *147*, 257-266.

Goder, V., Crottet, P., and Spiess, M. (2000). In vivo kinetics of protein targeting to the endoplasmic reticulum determined by site-specific phosphorylation. Embo J *19*, 6704-6712.

Goder, V., Junne, T., and Spiess, M. (2004). Sec61p contributes to signal sequence orientation according to the positive-inside rule. Mol Biol Cell *15*, 1470-1478.

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

Goder, V., and Spiess, M. (2003). Molecular mechanism of signal sequence orientation in the endoplasmic reticulum. Embo J *22*, 3645-3653.

Gorlich, D., Prehn, S., Hartmann, E., Kalies, K. U., and Rapoport, T. A. (1992). A mammalian homolog of SEC61p and SECYp is associated with ribosomes and nascent polypeptides during translocation. Cell *71*, 489-503.

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

Gould, S. G., Keller, G. A., and Subramani, S. (1987). Identification of a peroxisomal targeting signal at the carboxy terminus of firefly luciferase. J Cell Biol *105*, 2923-2931.

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

Gouveia, A. M., Guimaraes, C. P., Oliveira, M. E., Reguenga, C., Sa-Miranda, C., and Azevedo, J. E. (2003). Characterization of the peroxisomal cycling receptor, Pex5p, using a cell-free in vitro import system. J Biol Chem *278*, 226-232.

Hachiya, N., Alam, R., Sakasegawa, Y., Sakaguchi, M., Mihara, K., and Omura, T. (1993). A mitochondrial import factor purified from rat liver cytosol is an ATP-dependent conformational modulator for precursor proteins. Embo J *12*, 1579-1586.

Haigh, N. G., and Johnson, A. E. (2002). A new role for BiP: closing the aqueous translocon pore during protein integration into the ER membrane. J Cell Biol *156*, 261-270.

Hamman, B. D., Chen, J. C., Johnson, E. E., and Johnson, A. E. (1997). The aqueous pore through the translocon has a diameter of 40-60 A during cotranslational protein translocation at the ER membrane. Cell *89*, 535-544.

Hartmann, E., Rapoport, T. A., and Lodish, H. F. (1989). Predicting the orientation of eukaryotic membrane-spanning proteins. Proc Natl Acad Sci U S A *86*, 5786-5790.

Hauri, H. P., and Schweizer, A. (1992). The endoplasmic reticulum-Golgi intermediate compartment. Curr Opin Cell Biol *4*, 600-608.

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

Heinrich, S. U., and Rapoport, T. A. (2003). Cooperation of transmembrane segments during the integration of a double-spanning protein into the ER membrane. Embo J *22*, 3654-3663.

Hershey, J. W. (1991). Translational control in mammalian cells. Annu Rev Biochem *60*, 717-755.

Hessa, T., Kim, H., Bihlmaier, K., Lundin, C., Boekel, J., Andersson, H., Nilsson, I., White, S. H., and von Heijne, G. (2005). Recognition of transmembrane helices by the endoplasmic reticulum translocon. Nature *433*, 377-381.

High, S., Flint, N., and Dobberstein, B. (1991). Requirements for the membrane insertion of signal-anchor type proteins. J Cell Biol *113*, 25-34.

Hirst, J., Lui, W. W., Bright, N. A., Totty, N., Seaman, M. N., and Robinson, M. S. (2000). A family of proteins with gamma-adaptin and VHS domains that facilitate trafficking between the trans-Golgi network and the vacuole/lysosome. J Cell Biol *149*, 67-80.

Hodel, M. R., Corbett, A. H., and Hodel, A. E. (2001). Dissection of a nuclear localization signal. J Biol Chem *276*, 1317-1325.

Jensen, R. E., and Johnson, A. E. (1999). Protein translocation: is Hsp70 pulling my chain? Curr Biol *9*, R779-782.

Johnson, A. E., and van Waes, M. A. (1999). The translocon: a dynamic gateway at the ER membrane. Annu Rev Cell Dev Biol *15*, 799-842.

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

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

Keller, P., and Simons, K. (1997). Post-Golgi biosynthetic trafficking. J Cell Sci *110 (Pt 24)*, 3001-3009.

Killian, J. A., and von Heijne, G. (2000). How proteins adapt to a membrane-water interface. Trends Biochem Sci *25*, 429-434.

Kirchhausen, T. (2000). Three ways to make a vesicle. Nat Rev Mol Cell Biol *1*, 187-198.

Koch, H. G., Moser, M., and Muller, M. (2003). Signal recognition particle-dependent protein targeting, universal to all kingdoms of life. Rev Physiol Biochem Pharmacol *146*, 55-94.

Koehler, C. M., Merchant, S., and Schatz, G. (1999). How membrane proteins travel across the mitochondrial intermembrane space. Trends Biochem Sci *24*, 428-432.

Kornfeld, R., and Kornfeld, S. (1985). Assembly of asparagine-linked oligosaccharides. Annu Rev Biochem *54*, 631-664.

Krieg, U. C., Walter, P., and Johnson, A. E. (1986). Photocrosslinking of the signal sequence of nascent preprolactin to the 54-kilodalton polypeptide of the signal recognition particle. Proc Natl Acad Sci U S A *83*, 8604-8608.

Kulinska, K., Kulinski, T., Lyubartsev, A., Laaksonen, A., and Adamiak, R. W. (2000). Spatial distribution functions as a tool in the analysis of ribonucleic acids hydration--molecular dynamics studies. Comput Chem *24*, 451-457.

Kunkele, K. P., Heins, S., Dembowski, M., Nargang, F. E., Benz, R., Thieffry, M., Walz, J., Lill, R., Nussberger, S., and Neupert, W. (1998). The preprotein translocation channel of the outer membrane of mitochondria. Cell *93*, 1009-1019.

Kurzchalia, T. V., Wiedmann, M., Girshovich, A. S., Bochkareva, E. S., Bielka, H., and Rapoport, T. A. (1986). The signal sequence of nascent preprolactin interacts with the 54K polypeptide of the signal recognition particle. Nature *320*, 634-636.

Ladinsky, M. S., Mastronarde, D. N., McIntosh, J. R., Howell, K. E., and Staehelin, L. A. (1999). Golgi structure in three dimensions: functional insights from the normal rat kidney cell. J Cell Biol *144*, 1135-1149.

Landolt-Marticorena, C., Williams, K. A., Deber, C. M., and Reithmeier, R. A. (1993). Non-random distribution of amino acids in the transmembrane segments of human type I single span membrane proteins. J Mol Biol *229*, 602-608.

Lee, M. C., Miller, E. A., Goldberg, J., Orci, L., and Schekman, R. (2004). Bi-directional protein transport between the ER and Golgi. Annu Rev Cell Dev Biol *20*, 87-123.

Lemberg, M. K., and Martoglio, B. (2004). On the mechanism of SPP-catalysed intramembrane proteolysis; conformational control of peptide bond hydrolysis in the plane of the membrane. FEBS Lett *564*, 213-218.

Letourneur, F., Gaynor, E. C., Hennecke, S., Demolliere, C., Duden, R., Emr, S. D., Riezman, H., and Cosson, P. (1994). Coatomer is essential for retrieval of dilysine-tagged proteins to the endoplasmic reticulum. Cell *79*, 1199-1207.

Lewis, M. J., and Pelham, H. R. (1990). A human homologue of the yeast HDEL receptor. Nature *348*, 162-163.

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

Lipp, J., Flint, N., Haeuptle, M. T., and Dobberstein, B. (1989). Structural requirements for membrane assembly of proteins spanning the membrane several times. J Cell Biol *109*, 2013-2022.

Lord, J. M., Davey, J., Frigerio, L., and Roberts, L. M. (2000). Endoplasmic reticulum-associated protein degradation. Semin Cell Dev Biol *11*, 159-164.

Lui, W. W., Collins, B. M., Hirst, J., Motley, A., Millar, C., Schu, P., Owen, D. J., and Robinson, M. S. (2003). Binding partners for the COOH-terminal appendage domains of the GGAs and gamma-adaptin. Mol Biol Cell *14*, 2385-2398.

Lutcke, H., High, S., Romisch, K., Ashford, A. J., and Dobberstein, B. (1992). The methionine-rich domain of the 54 kDa subunit of signal recognition particle is sufficient for the interaction with signal sequences. Embo J *11*, 1543-1551.

Macara, I. G. (2001). Transport into and out of the nucleus. Microbiol Mol Biol Rev *65*, 570-594, table of contents.

Martinez-Menarguez, J. A., Geuze, H. J., Slot, J. W., and Klumperman, J. (1999). Vesicular tubular clusters between the ER and Golgi mediate concentration of soluble secretory proteins by exclusion from COPI-coated vesicles. Cell *98*, 81-90.

Martoglio, B., and Dobberstein, B. (1996). Snapshots of membrane-translocating proteins. Trends Cell Biol *6*, 142-147.

Martoglio, B., Graf, R., and Dobberstein, B. (1997). Signal peptide fragments of preprolactin and HIV-1 p-gp160 interact with calmodulin. Embo J *16*, 6636-6645.

Martoglio, B., Hofmann, M. W., Brunner, J., and Dobberstein, B. (1995). The protein-conducting channel in the membrane of the endoplasmic reticulum is open laterally toward the lipid bilayer. Cell *81*, 207-214.

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

Matlack, K. E., and Walter, P. (1995). The 70 carboxyl-terminal amino acids of nascent secretory proteins are protected from proteolysis by the ribosome and the protein translocation apparatus of the endoplasmic reticulum membrane. J Biol Chem *270*, 6170-6180.

McCormick, P. J., Miao, Y., Shao, Y., Lin, J., and Johnson, A. E. (2003). Cotranslational protein integration into the ER membrane is mediated by the binding of nascent chains to translocon proteins. Mol Cell *12*, 329-341.

McGovern, K., and Beckwith, J. (1991). Membrane insertion of the Escherichia coli MalF protein in cells with impaired secretion machinery. J Biol Chem *266*, 20870-20876.

Meacock, S. L., Lecomte, F. J., Crawshaw, S. G., and High, S. (2002). Different transmembrane domains associate with distinct endoplasmic reticulum components during membrane integration of a polytopic protein. Mol Biol Cell *13*, 4114-4129.

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

Meyer, C., Eskelinen, E. L., Guruprasad, M. R., von Figura, K., and Schu, P. (2001). Mu 1A deficiency induces a profound increase in MPR300/IGF-II receptor internalization rate. J Cell Sci *114*, 4469-4476.

Milstein, C., Brownlee, G. G., Harrison, T. M., and Mathews, M. B. (1972). A possible precursor of immunoglobulin light chains. Nat New Biol *239*, 117-120.

Morgan, D. G., Menetret, J. F., Radermacher, M., Neuhof, A., Akey, I. V., Rapoport, T. A., and Akey, C. W. (2000). A comparison of the yeast and rabbit 80 S ribosome reveals the topology of the nascent chain exit tunnel, inter-subunit bridges and mammalian rRNA expansion segments. J Mol Biol *301*, 301-321.

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

Mothes, W., Jungnickel, B., Brunner, J., and Rapoport, T. A. (1998). Signal sequence recognition in cotranslational translocation by protein components of the endoplasmic reticulum membrane. J Cell Biol *142*, 355-364.

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

Munro, S., and Pelham, H. R. (1987). A C-terminal signal prevents secretion of luminal ER proteins. Cell *48*, 899-907.

Nair, D. M., Purdue, P. E., and Lazarow, P. B. (2004). Pex7p translocates in and out of peroxisomes in Saccharomyces cerevisiae. J Cell Biol *167*, 599-604.

Nilsson, I., Witt, S., Kiefer, H., Mingarro, I., and von Heijne, G. (2000). Distant downstream sequence determinants can control N-tail translocation during protein insertion into the endoplasmic reticulum membrane. J Biol Chem *275*, 6207-6213.

Nilsson, I. M., and von Heijne, G. (1993). Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. J Biol Chem *268*, 5798-5801.

Osumi, T., Tsukamoto, T., Hata, S., Yokota, S., Miura, S., Fujiki, Y., Hijikata, M., Miyazawa, S., and Hashimoto, T. (1991). Amino-terminal presequence of the precursor of peroxisomal 3-ketoacyl-CoA thiolase is a cleavable signal peptide for peroxisomal targeting. Biochem Biophys Res Commun *181*, 947-954.

Ota, K., Sakaguchi, M., Hamasaki, N., and Mihara, K. (2000). Membrane integration of the second transmembrane segment of band 3 requires a closely apposed preceding signal-anchor sequence. J Biol Chem *275*, 29743-29748.

Paetzel, M., Karla, A., Strynadka, N. C., and Dalbey, R. E. (2002). Signal peptidases. Chem Rev *102*, 4549-4580.

Panzner, S., Dreier, L., Hartmann, E., Kostka, S., and Rapoport, T. A. (1995). Posttranslational protein transport in yeast reconstituted with a purified complex of Sec proteins and Kar2p. Cell *81*, 561-570.

Parks, G. D., and Lamb, R. A. (1991). Topology of eukaryotic type II membrane proteins: importance of N-terminal positively charged residues flanking the hydrophobic domain. Cell *64*, 777-787.

Perlman, D., and Halvorson, H. O. (1983). A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. J Mol Biol *167*, 391-409.

Pfanner, N., and Geissler, A. (2001). Versatility of the mitochondrial protein import machinery. Nat Rev Mol Cell Biol *2*, 339-349.

Pilon, M., Romisch, K., Quach, D., and Schekman, R. (1998). Sec61p serves multiple roles in secretory precursor binding and translocation into the endoplasmic reticulum membrane. Mol Biol Cell *9*, 3455-3473.

Pilon, M., Schekman, R., and Romisch, K. (1997). Sec61p mediates export of a misfolded secretory protein from the endoplasmic reticulum to the cytosol for degradation. Embo J *16*, 4540-4548.

Platta, H. W., Girzalsky, W., and Erdmann, R. (2004). Ubiquitination of the peroxisomal import receptor Pex5p. Biochem J *384*, 37-45.

Presley, J. F., Cole, N. B., Schroer, T. A., Hirschberg, K., Zaal, K. J., and Lippincott-Schwartz, J. (1997). ER-to-Golgi transport visualized in living cells. Nature *389*, 81-85.

Rabouille, C., Hui, N., Hunte, F., Kieckbusch, R., Berger, E. G., Warren, G., and Nilsson, T. (1995). Mapping the distribution of Golgi enzymes involved in the construction of complex oligosaccharides. J Cell Sci *108 (Pt 4)*, 1617-1627.

Rapaport, D., and Neupert, W. (1999). Biogenesis of Tom40, core component of the TOM complex of mitochondria. J Cell Biol *146*, 321-331.

Rapoport, T. A., Goder, V., Heinrich, S. U., and Matlack, K. E. (2004). Membrane-protein integration and the role of the translocation channel. Trends Cell Biol *14*, 568-575.

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

Rassow, J., Hartl, F. U., Guiard, B., Pfanner, N., and Neupert, W. (1990). Polypeptides traverse the mitochondrial envelope in an extended state. FEBS Lett *275*, 190-194.

Rassow, J., and Pfanner, N. (2000). The protein import machinery of the mitochondrial membranes. Traffic *1*, 457-464.

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

Robinson, M. S., and Bonifacino, J. S. (2001). Adaptor-related proteins. Curr Opin Cell Biol *13*, 444-453.

Rosch, K., Naeher, D., Laird, V., Goder, V., and Spiess, M. (2000). The topogenic contribution of uncharged amino acids on signal sequence orientation in the endoplasmic reticulum. J Biol Chem *275*, 14916-14922.

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

Sabatini, D. D., Blobel, G., Nonomura, Y., and Adelman, M. R. (1971). Ribosome-membrane interaction: Structural aspects and functional implications. Adv Cytopharmacol *1*, 119-129.

Sakaguchi, M., Tomiyoshi, R., Kuroiwa, T., Mihara, K., and Omura, T. (1992). Functions of signal and signal-anchor sequences are determined by the balance between the hydrophobic segment and the N-terminal charge. Proc Natl Acad Sci U S A *89*, 16-19.

Saraste, J., and Kuismanen, E. (1984). Pre- and post-Golgi vacuoles operate in the transport of Semliki Forest virus membrane glycoproteins to the cell surface. Cell *38*, 535-549.

Saraste, J., and Kuismanen, E. (1992). Pathways of protein sorting and membrane traffic between the rough endoplasmic reticulum and the Golgi complex. Semin Cell Biol *3*, 343-355.

Sato, M., Hresko, R., and Mueckler, M. (1998). Testing the charge difference hypothesis for the assembly of a eucaryotic multispanning membrane protein. J Biol Chem *273*, 25203-25208.

Sato, Y., Sakaguchi, M., Goshima, S., Nakamura, T., and Uozumi, N. (2003). Molecular dissection of the contribution of negatively and positively charged residues in S2, S3, and S4 to the final membrane topology of the voltage sensor in the K+ channel, KAT1. J Biol Chem *278*, 13227-13234.

Scales, S. J., Pepperkok, R., and Kreis, T. E. (1997). Visualization of ER-to-Golgi transport in living cells reveals a sequential mode of action for COPII and COPI. Cell *90*, 1137-1148.

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

Scheffzek, K., Klebe, C., Fritz-Wolf, K., Kabsch, W., and Wittinghofer, A. (1995). Crystal structure of the nuclear Ras-related protein Ran in its GDP-bound form. Nature *374*, 378-381.

Schirmer, T., Keller, T. A., Wang, Y. F., and Rosenbusch, J. P. (1995). Structural basis for sugar translocation through maltoporin channels at 3.1 A resolution. Science *267*, 512-514.

Schleyer, M., and Neupert, W. (1985). Transport of proteins into mitochondria: translocational intermediates spanning contact sites between outer and inner membranes. Cell *43*, 339-350.

Schmid, S. R., and Spiess, M. (1988). Deletion of the amino-terminal domain of asialoglycoprotein receptor H1 allows cleavage of the internal signal sequence. J Biol Chem *263*, 16886-16891.

Schnell, D. J., and Hebert, D. N. (2003). Protein Translocons. Multifunctional Mediators of Protein Translocation across Membranes. Cell *112*, 491-505.

Schweizer, A., Fransen, J. A., Matter, K., Kreis, T. E., Ginsel, L., and Hauri, H. P. (1990). Identification of an intermediate compartment involved in protein transport from endoplasmic reticulum to Golgi apparatus. Eur J Cell Biol *53*, 185-196.

Semenza, J. C., Hardwick, K. G., Dean, N., and Pelham, H. R. (1990). ERD2, a yeast gene required for the receptor-mediated retrieval of luminal ER proteins from the secretory pathway. Cell *61*, 1349-1357.

Serek, J., Bauer-Manz, G., Struhalla, G., Van Den Berg, L., Kiefer, D., Dalbey, R., and Kuhn, A. (2004). Escherichia coli YidC is a membrane insertase for Sec-independent proteins. Embo J *23*, 294-301.

Sheffield, W. P., Shore, G. C., and Randall, S. K. (1990). Mitochondrial precursor protein. Effects of 70-kilodalton heat shock protein on polypeptide folding, aggregation, and import competence. J Biol Chem *265*, 11069-11076.

Stewart, R. S., Drisaldi, B., and Harris, D. A. (2001). A transmembrane form of the prion protein contains an uncleaved signal peptide and is retained in the endoplasmic Reticulum. Mol Biol Cell *12*, 881-889.

Stirling, C. J., Rothblatt, J., Hosobuchi, M., Deshaies, R., and Schekman, R. (1992). Protein translocation mutants defective in the insertion of integral membrane proteins into the endoplasmic reticulum. Mol Biol Cell *3*, 129-142.

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

Swinkels, B. W., Gould, S. J., Bodnar, A. G., Rachubinski, R. A., and Subramani, S. (1991). A novel, cleavable peroxisomal targeting signal at the amino-terminus of the rat 3-ketoacyl-CoA thiolase. Embo J *10*, 3255-3262.

Teasdale, R. D., and Jackson, M. R. (1996). Signal-mediated sorting of membrane proteins between the endoplasmic reticulum and the golgi apparatus. Annu Rev Cell Dev Biol *12*, 27-54.

Thiele, C., Gerdes, H. H., and Huttner, W. B. (1997). Protein secretion: puzzling receptors. Curr Biol *7*, R496-500.

Thyberg, J., and Moskalewski, S. (1999). Role of microtubules in the organization of the Golgi complex. Exp Cell Res *246*, 263-279.

Truscott, K. N., Kovermann, P., Geissler, A., Merlin, A., Meijer, M., Driessen, A. J., Rassow, J., Pfanner, N., and Wagner, R. (2001). A presequence- and voltage-sensitive channel of the mitochondrial preprotein translocase formed by Tim23. Nat Struct Biol *8*, 1074-1082.

Valdivia, R. H., Baggott, D., Chuang, J. S., and Schekman, R. W. (2002). The yeast clathrin adaptor protein complex 1 is required for the efficient retention of a subset of late Golgi membrane proteins. Dev Cell *2*, 283-294.

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

van Klompenburg, W., Nilsson, I., von Heijne, G., and de Kruijff, B. (1997). Anionic phospholipids are determinants of membrane protein topology. Embo J *16*, 4261-4266.

Van Voorst, F., and De Kruijff, B. (2000). Role of lipids in the translocation of proteins across membranes. Biochem J *347 Pt 3*, 601-612.

Vetter, I. R., Nowak, C., Nishimoto, T., Kuhlmann, J., and Wittinghofer, A. (1999). Structure of a Ran-binding domain complexed with Ran bound to a GTP analogue: implications for nuclear transport. Nature *398*, 39-46.

von Heijne, G. (1983). Patterns of amino acids near signal-sequence cleavage sites. Eur J Biochem *133*, 17-21.

von Heijne, G. (1986). Mitochondrial targeting sequences may form amphiphilic helices. Embo J *5*, 1335-1342.

von Heijne, G. (1989). Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. Nature *341*, 456-458.

von Heijne, G. (1990). The signal peptide. J Membr Biol *115*, 195-201.

Wahlberg, J. M., and Spiess, M. (1997). Multiple determinants direct the orientation of signal-anchor proteins: the topogenic role of the hydrophobic signal domain. J Cell Biol *137*, 555-562.

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

Warren, G., and Mellman, I. (1999). Bulk flow redux? Cell *98*, 125-127.

Weihofen, A., Binns, K., Lemberg, M. K., Ashman, K., and Martoglio, B. (2002). Identification of signal peptide peptidase, a presenilin-type aspartic protease. Science *296*, 2215-2218.

Weiss, M. S., Abele, U., Weckesser, J., Welte, W., Schiltz, E., and Schulz, G. E. (1991). Molecular architecture and electrostatic properties of a bacterial porin. Science *254*, 1627-1630.

Wessels, H. P., Beltzer, J. P., and Spiess, M. (1991). Analysis of protein topology in the endoplasmic reticulum. Methods Cell Biol *34*, 287-302.

Wessels, H. P., and Spiess, M. (1988). Insertion of a multispanning membrane protein occurs sequentially and requires only one signal sequence. Cell *55*, 61-70.

Wiedmann, B., Sakai, H., Davis, T. A., and Wiedmann, M. (1994). A protein complex required for signal-sequence-specific sorting and translocation. Nature *370*, 434-440.

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

Wimley, W. C., and White, S. H. (1996). Experimentally determined hydrophobicity scale for proteins at membrane interfaces. Nat Struct Biol *3*, 842-848.

Zaremba, S., and Keen, J. H. (1983). Assembly polypeptides from coated vesicles mediate reassembly of unique clathrin coats. J Cell Biol *97*, 1339-1347.

Zhang, J. X., Braakman, I., Matlack, K. E., and Helenius, A. (1997). Quality control in the secretory pathway: the role of calreticulin, calnexin and BiP in the retention of glycoproteins with C-terminal truncations. Mol Biol Cell *8*, 1943-1954.

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

Curriculum vitae

Higy Marie

Born : 02.02.1974

City: Mulhouse

Nationality: French

Academic Preparation

1978-1986	Primary school of Dannemarie
1986-1989	College of dannemarie
1989-1993	High school in the College Episcopale of Zillisheim
1993	High School Diploma type D (biology, science) , college JJ Henner, Altkirch
1994-1996	First year of Medecine, Strasbourg
1996-1997	Diploma in Biology, Louis Pasteur University, Strasbourg
1998	Degree in Biochemistry, Louis Pasteur University, Strasbourg, 1998
1999	Master's degree in Cellular Biology and Physiology, Louis Pasteur University, Strasbourg
2000	Post-graduate Degree in Neurosciences, Louis Pasteur University, Strasbourg
2001-2005	Ph.D. in Biochemistry, Biozentrum, Basel University, Basel

Scientific Papers

M Higy, S Gander, M Spiess, Probing the environment of signal-anchor sequences during topogenesis in the endoplasmic reticulum, Biochemistry, 2005, **44**:2039-2047.

M Higy, T Junne, M Spiess, Topogenesis of membrane proteins at the Endoplasmic Reticulum, Biochemistry, 2004, **43**:12716-12722.