Mechanism of Sec61-mediated insertion of proteins into the endoplasmic reticulum membrane

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Lucyna Kocik

aus Boguslawice, Polen

Basel, 2010

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

Auf Antrag von

Prof. Dr. Martin Spiess und Prof. Dr. Anne Spang

Basel, den 19. Oktober 2010

Prof. Dr. Martin Spiess

(Dekan)

Acknowledgements

I would like to express my gratitude to:

Professor Martin Spiess for giving me the opportunity to perform this research in his lab and for his guidance

Tina Junne for teaching me yeast procedures, data contribution and for the great collaboration on this project

Nicole Beuret for introducing me to various biochemical techniques and for technical assistance

All the present and former members of the Spiess group: Barry Shortt, David Hirschmann, Deyan Mihov, Franziska Hasler, Sonja Huser, Simone Kaelin, Cristina Baschong, Gregor Suri, Julia Birk, Michael Friberg, Szymon Kobialka, Pascal Crottet for the nice working atmosphere and fruitful discussions

Dziekuje mojej Rodzinie za wsparcie

Summary

In eukaryotic cells, hydrophobic signal sequences of newly synthesized secretory and membrane proteins target them to the Sec61 translocon in the endoplasmic reticulum (ER) membrane. The translocon forms a hydrophilic pore, in its idle state closed by a lumenal plug domain and a hydrophobic constriction ring. Within the Sec61 channel, transmembrane segments of proteins achieve their proper orientation (topology) and are laterally released into the lipid bilayer. Orientation of signal sequences in the ER membrane is determined by charged residues flanking the hydrophobic core of the signal, hydrophobicity of the signal, the size and folding properties of the N-terminal domain preceding the signal and, in some cases, the length of the C-terminus.

In Part I of this thesis we compared the insertion process of N-terminal versus internal signal-anchors of single-spanning membrane proteins and determined the effect of the N-terminal hydrophilic domain on protein topogenesis. We showed that insertion of these two types of signals occurs via different mechanisms. Transition from N-terminal to internal signals, achieved by extension of the N-domain with hydrophilic residues, was accompanied by loss of C-terminal length dependence and insensitivity to increased hydrophobicity of the signal. It indicated that, in contrast to N-terminal signals, signal-anchors localized internally cannot undergo reorientation within the pore. Furthermore, hydrophilic N-terminal domains sterically hinder N-translocation.

In Part II we analyzed the insertion process of proteins with conflicting signal sequences: type I cleavable hemagglutinin (HA) signal and a type II signal-anchor of H1. We showed that proteins with wild-type HA and H1 signals, connected by a 40-amino acid linker compete for the preferred orientation in the translocon, manifested by a rapid inversion of a fraction of the polypeptides, triggered by the signal-anchor. The process could be slowed down by increasing the hydrophobicity of the H1 signal or manipulating its flanking charges. Under such conditions, topogenesis was interrupted upon termination of translation, like previously observed for N-terminal signal-anchors. In contrast to single-spanning membrane proteins, the topogenesis window is not a constant of the translocation machinery, but rather appears to be substrate-specific.

In Part III we tested the function of the apolar core of the Sec61 translocon. We mutated the ring residues of yeast Sec61p to more hydrophilic, bulky, or even charged amino acids (alanines, glycines, serines, tryptophans, lysines, or aspartates). The translocon turned out to be surprisingly tolerant even to the charge mutations in the constriction ring, since growth and

translocation efficiency were not drastically affected. Ring mutants altered the integration of hydrophobic sequences into the lipid bilayer, which indicated that the translocon does not simply catalyze the partitioning of potential transmembrane segments between an aqueous environment and the lipid bilayer, but that it plays an active role in setting the hydrophobicity threshold for membrane integration.

Table of contents

Abbreviations

aa	amino acid
ADP	adenosine diphosphate
ASGP	asialoglycoprotein
ASGPR	asialoglycoprotein receptor
ATP	adenosine-5`-triphosphate
BiP	binding immunoglobulin protein
bp	base pair
BSA	bovine serum albumin
CHX	cycloheximide
COP	coatamer protein
CPY	carboxypeptidase Y
cryoEM	cryo-electron microscopy
cyt	cytoplasmic
DHFR	dihydrofolate reductase
DMEM	Dulbecco`s modified Eagle`s medium
DMSO	dimethyl sulfoxide
DPAPB	dipeptidyl aminopeptidase B
EF	elongation factor
endoH	endoglycosidase H
ER	endoplasmic reticulum
ERGIC	ER-Golgi intermediate compartment
exo	exoplasmic
FCS	fetal calf serum
Ffh	fifty-four homologue
GDP	glyceraldehyde-3-phosphate dehydrogenase
GTP	guanosine-5`-triphosphate
HA	hemagglutinin
HSD	helical scaffold domain
Hsp	heat shock protein
kDa	kilodalton

Lep	leader peptidase
mRNA	messenger ribonucleic acid
NBD	nucleotide binding domain
NBF	nucleotide-binding fold
NEF	nucleotide exchange factor
OD	optical density
OST	oligosaccharyl transferase
PBS	phosphate-buffered saline
PC	procoat
PCC	protein conducting channel
PCR	polymerase chain reaction
PDB	Protein Data Bank
PE	phosphatidylethanolamine
PEI	polyethylenimine
PIC	protease inhibitor coctail
PM	plasma membrane
PMSF	phenylmethylsulfonyl fluoride
PPXD	pre-protein crosslinking domain
prl	*pr*otein *l*ocalization
RAMP	ribosome-associated membrane protein
RM	rough microsome
RNC	ribosome-nascent chain complex
SA	signal-anchor
SBD	substrate binding domain
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SL RNA	spliced leader ribonucleic acid
SP	signal peptidase
SPC	signal peptidase complex
SR	signal recognition particle receptor
SRP	signal recognition particle
TA	tail-anchored
TM	transmembrane segment

TMD   transmembrane domain

TRAM   translocating chain-associated membrane protein

TRAP   translocon-associated protein

VAMP   vesicle-associated membrane protein

wt   wild-type

YPDA   yeast peptone dextrose adenine

I General introduction

1. Protein sorting pathways

Eukaryotic cells are subdivided into structurally and functionally distinct, membrane-enclosed compartments, called organelles. As membranes act as hydrophobic physical barriers, in order to maintain a unique protein composition of each organelle, the cell developed a sophisticated machinery that enables transport of solutes and macromolecules across these barriers. In eukaryotic cells, the synthesis of most proteins begins in the cytosol, followed by their delivery to the appropriate cellular compartment (**Figure 1**). Protein transfer is guided by sorting signals in the transported protein, which are recognized by complementary sorting receptors in the target organelles. Examples of sorting signals are presented in **Table I**.
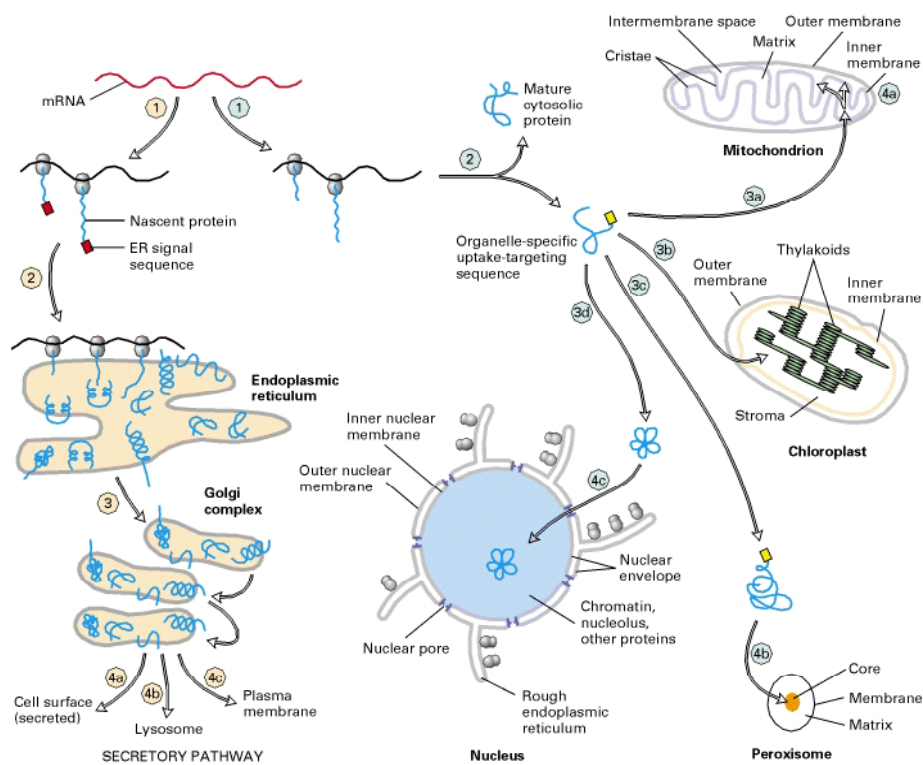
Figure 1. Protein sorting in eukaryotic cells. Proteins encoded by the nuclear mRNA are synthesized on cytosolic ribosomes. Polypeptides entering the secretory pathway (1) contain an ER signal sequence that targets the ribosome-nascent chain complex to the ER membrane (2). Proteins exit the ER in transport vesicles and are delivered to the Golgi apparatus (3). From there, they can be transported out of the cell (4a), delivered to lysosomes (4b) or incorporated into the plasma membrane (4c). Proteins that lack an ER targeting signal are synthesized on free ribosomes in the cytoplasm (1). They can be released into the cytosol (2) or delivered to the mitochondrion (3a), chloroplast (3b), peroxisome (3c) or the nucleus (3d), mediated by an organelle-specific sorting signal. (Lodish, 2000)

Table I. Some typical signal sequences.

Function of signal sequence	Example of signal sequence
Import into nucleus	-Pro-Pro-Lys-Lys-Lys-Arg-Lys-Val-
Export from nucleus	-Leu-Ala-Leu-Lys-Leu-Ala-Gly-Leu-Asp-Ile
Import into mitochodria	$^+$H$_3$N-Met-Leu-Ser-Leu-Arg-Gln-Ser-Ile-Arg-Phe-Phe-Lys-Pro-Ala-Thr-Arg-Thr-Leu-Cys-Ser-Ser-Arg-Tyr-Leu-Leu-
Import into plastid	$^+$H$_3$N-Met-Val-Ala-Met-Ala-Met-Ala-Ser-Leu-Gln-Ser-Ser-Met-Ser-Ser-Leu-Ser-Leu-Ser-Ser-Asn-Ser-Phe-Leu-Gly-Gln-Pro-Leu-Ser-Pro-Ile-Thr-Leu-Ser-Pro-Phe-Leu-Gln-Gly-
Import into peroxisomes	-Ser-Lys-Leu-COO$^-$
Import into ER	$^+$H$_3$N-Met-Met-Ser-Phe-Val-Ser-Leu-Leu-Leu-Val-Gly-Ile-Leu-Phe-Trp-Ala-Thr-Glu-Ala-Glu-Gln-Leu-Thr-Lys-Cys-Glu-Val-Phe-Gln-
Return to ER	-Lys-Asp-Glu-Leu-COO$^-$

Amino acids characteristic for different classes of signal sequences are highlighted in color. When important for the function of the signal, positively charged residues are shown in red, negatively charged- in green, hydrophobic- in white and hydroxylated amino acids- in blue. $^+$H$_3$N and COO$^-$ indicate the N- and C-terminus of the polypeptide, respectively. Redrawn from: (Alberts, 2008)

In eukaryotic cells, targeting to the endoplasmic reticulum (ER) is the first step in the biosynthesis of secretory and membrane proteins, such as resident proteins of the ER, ER-Golgi intermediate compartment (ERGIC), Golgi apparatus, endosomes, lysosomes and the plasma membrane (PM) (Palade, 1975). While secretory proteins are completely translocated into the ER lumen, transport of membrane proteins requires their incorporation into the ER membrane. In prokaryotes, newly synthesized proteins are transported across- or are built into the plasma membrane.

Protein targeting to the ER is mediated by signal sequences located within the nascent chain. They are composed of a positively charged n-region, a h-region containing 8-20 hydrophobic amino acids and a polar c-region (**Table I**) (Gierasch, 1989). Signal sequences carry topogenic information that allows proteins to achieve their proper orientation in the membrane. Based on the composition and mechanism of insertion, proteins can be divided into secretory proteins (fully translocated into the ER lumen), single-spanning membrane proteins (type I, type II, type III and tail-anchored) and multispanning membrane proteins (**Figure 2**). Proteins of type I are targeted to the ER by an N-terminal, cleavable signal sequence, and then anchored in the membrane by a subsequent hydrophobic `stop-transfer` sequence. Their final topology is N_{exo}/C_{cyt} (exoplasmic or lumenal N-terminus and cytoplasmic C-terminus). Proteins of type II possess a signal-anchor sequence, which mediates both targeting and membrane anchoring. They assume an N_{cyt}/C_{exo} orientation. Reverse signal-anchors of type III membrane proteins initiate translocation of the N-terminus across the ER membrane, yielding the opposite topology. In

addition, there is a class of tail-anchored proteins, where the signal sequence is located at the very C-terminus (Goder and Spiess, 2001).
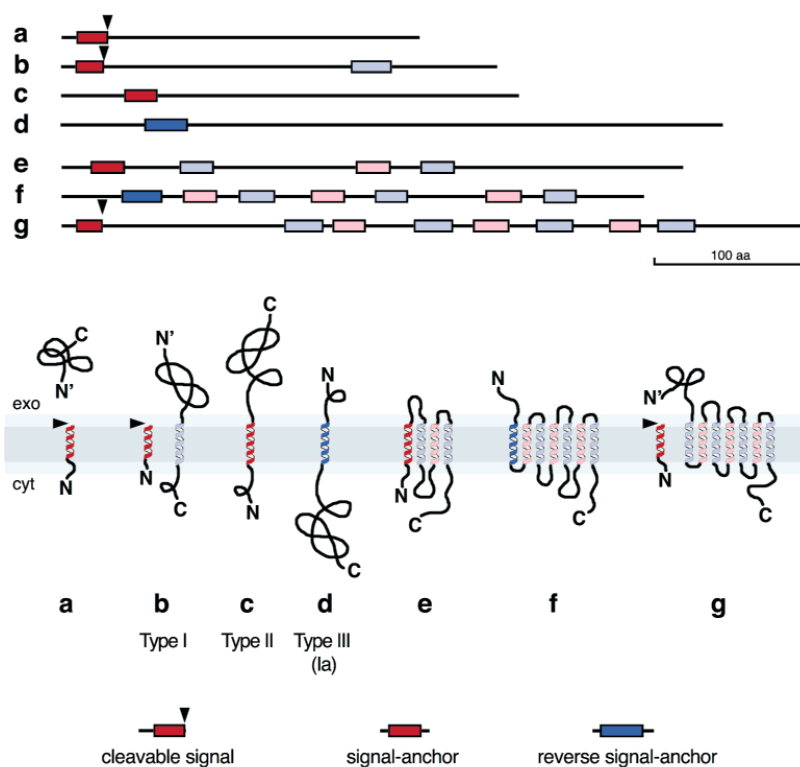
Figure 2. Three types of signals initiate co-translational protein topogenesis. Cleavable signals (red with arrowhead indicating the signal peptidase cleavage site) and uncleaved signal-anchors (red without arrowhead) induce translocation of the C-terminus and assume an N_{cyt}/C_{exo} orientation. Reverse signal anchors (blue) insert with the opposite N_{exo}/C_{cyt} orientation and translocate their N-terminus. Multispanning proteins contain additional transmembrane segments inserting in alternating orientations (light red for N_{cyt}/C_{exo} and light blue for N_{exo}/C_{cyt}). Examples of secretory and single-spanning membrane proteins: a- secretory protein (preprolactin), b- type I membrane protein (cation-dependent mannose-6-phosphate receptor), c- type II membrane protein (asialoglycoprotein receptor), d- type III membrane protein (synaptotagmin I). Examples of multispanning membrane proteins: e- gap junction protein R6, f- vasopressin receptor V2, g- glucagon receptor). (Higy et al., 2004)

The targeting process can occur co-translationally (coupled to protein synthesis on the ribosome) or post-translationally. Both mechanisms merge at the ER membrane, where the heterotrimeric translocon complex (Sec61 complex in eukaryotes, SecYEβ in bacteria, or SecYEG in *E. coli*) forms a pore that enables transfer of polypeptides across the membrane. The translocon is evolutionary conserved and consists of three subunits: the α-subunit, which is composed of a 10-helix bundle and forms the actual channel, and single-spanning β- and γ-subunits. The translocon, along with additional components, such as ER lumenal chaperones (e.g., BiP), nucleotide exchange factors (NEFs), TRAM (translocating chain-associated membrane) protein

(Gorlich and Rapoport, 1993; Snapp et al., 2004) and/or the heterotrimeric TRAP (translocon-associated protein) complex (Fons et al., 2003) participate in membrane integration or translocation of polypeptides into the ER lumen. As they cross the ER membrane, precursor polypeptides are processed on the lumenal side by the signal peptidase complex (SPC) and oligosaccharyl transferase (OST). These components contribute to proper folding of the protein and its eventual ER exit and vesicular transport along the secretory pathway. In the case of misfolding or assembly problems, the polypeptides are retrotranslocated to the cytosol, where they undergo proteolytic degradation by the proteasome. There is evidence that the Sec61 complex along with BiP might also be involved in this process (Schafer and Wolf, 2009; Willer et al., 2008).

2. Co-translational translocation

The co-translational translocation pathway is used for both secretory and membrane proteins and is found in all cells (Halic and Beckmann, 2005). It appears to be the predominant pathway in mammalian cells. During ER insertion, the Sec61 complex associates with the translating ribosome. The signal sequence that emerges from the ribosomal tunnel is recognized by the signal recognition particle (SRP). This interaction slows down the elongation process and allows targeting of the SRP-ribosome-nascent chain complex (RNC) to the ER membrane, where SRP binds its receptor (SR). Upon docking, the RNC is transferred to the translocon, SRP and SR dissociate from each other and the elongating polypeptide chain is inserted into the translocating channel (**Figure 3A**).

2.1 SRP structure and interaction with a signal peptide

In eukaryotic cells, SRP has two functions: transient translation arrest necessary for efficent ER delivery, and targeting of the RNCs to the SRP receptor in the ER membrane (Lakkaraju et al., 2008). Main substrates for the SRP-dependent route are integral membrane proteins that are prone to aggregation in the cytosol. They often contain a noncleavable transmembrane signal sequence, called signal-anchor (SA), which enables SRP binding and anchoring the protein in the membrane.
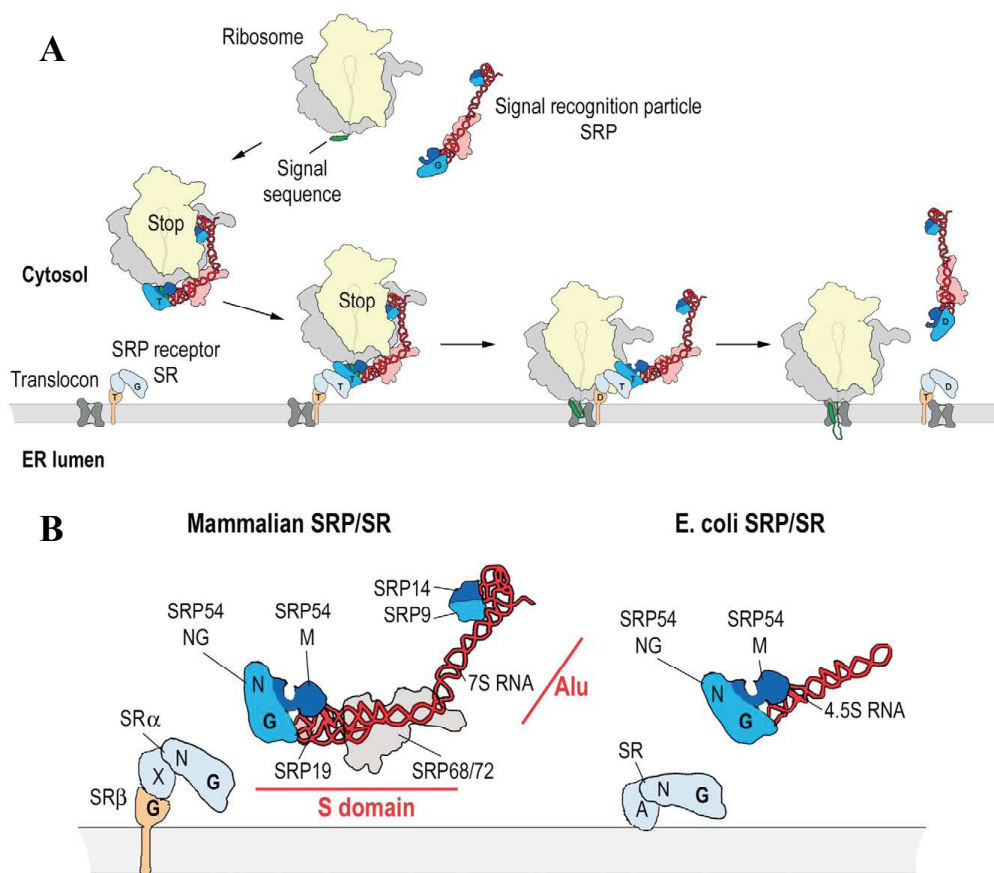
Figure 3. Co-translational targeting in eukaryotes and the SRP systems in mammalian and *E. coli* cells. **(A)** SRP interacts with the signal peptide as it emerges from the ribosomal tunnel. In eukaryotic cells, the formation of the SRP-RNC complex leads to elongation retardation. ER targeting is facilitated by interaction of SRP with SR and both molecules require GTP binding for their activity. The RNC complex is transferred to the translocon and upon GTP hydrolysis in SRP and its receptor, the complex disassembles. **(B)** Schematic overview of the mammalian and *E. coli* SRP and SR. SRP is divided into two main domains: the *Alu* and S domain. X and A in the SRP receptor refer to the corresponding domains in eukaryotic SRα and bacterial SR. G indicates GTPase domains. In eukaryotes, SR contains an additional subunit, SRß, whose transmembrane helix anchors it in the ER membrane. (Halic and Beckmann, 2005)

Mammalian SRP is composed of 7SL RNA and six proteins (SRP54, SRP19, SRP68/72, SRP14 and SRP9). The SRP of bacteria consists only of a 4.5S RNA and Ffh (fifty-four homologue), a homologue of SRP54. SRP54 and Ffh comprise three domains: the N-terminal domain that forms a 4-helix bundle and a Ras-like GTPase domain (G domain), which together form the NG domain, and the C-terminal M domain, rich in methionines, which associates with SRP RNA and the signal sequence. Signal peptide binding to the groove in the M domain can involve an induced fit mechanism to maximize the hydrophobic interactions. Additional hydrophobic amino acids lining the groove are available for interaction with different or longer signal peptides. The minimal length of the h-region is eight residues. It has been suggested that the

SRP RNA might play a role in signal recognition via electrostatic interactions with the positively charged residues in the n-region (Batey et al., 2000). However, the cryo-electron microscopy (cryoEM) models of the SRP-RNC complexes from *E. coli* and mammalian cells indicate no such interaction (Halic et al., 2006). SRP54 and Ffh bind the ribosome through the N domain and can be crosslinked to the ribosomal protein L23 near the nascent chain exit channel (Janda et al., 2010). The heterodimer SRP9/14 and the 5` and 3` ends of the SRP RNA form the *Alu* domain which functions in elongation arrest by preventing the binding of elongation factor 2 (EF2) (**Figure 3B**) (Ogg and Walter, 1995).

It has been shown that during translation eukaryotic SRP can discriminate between cytosolic and membrane proteins (Berndt et al., 2009). The binding process can be divided into three stages. In stage 1 SRP binds weakly to translating ribosomes. This interaction is independent of the length and type of nascent chain, however, SRP affinity to nontranslating ribosomes is higher. In stage 2, upon synthesis of a signal-anchor, SRP binds the ribosome more tightly in an electrostatic manner and localizes close to the exposed portion of the nascent chain. In the last stage, when the SA emerges from the ribosomal tunnel, SRP binds to it with high affinity via hydrophobic interactions. This leads to translation arrest and targeting to the SR in the ER membrane. As yeast cells contain only 1-2 SRP molecules per 100 ribosomes, it is crucial to recognize transmembrane proteins early. It was speculated that *in vivo* this is assured by SRP recruitment to ribosomes translating hydrophobic SAs even before they emerge from the tunnel. The eukayotic ribosome may have evolved to directly recognize a hydrophobic SA and trigger the co-translational translocation route. In contrast, in prokaryotic cells a SA inside the ribosome does not enhance the SRP affinity for ribosomes (Bornemann et al., 2008). In addition, the function of bacterial SRP, which lacks the elongation-arrest domain, may not have to sense a signal sequence as early as in the eukaryotic system.

2.2 Mechanism of SRP binding to SR

In bacteria, co-translational protein targeting must be completed before the nascent polypeptide exceeds approximately 140 amino acids in length (Flanagan et al., 2003). As a consequence, the time window for SRP binding is about 3–5 s and efficient ER targeting requires a rapid SRP–SR interaction. Both molecules contain NG-domains acting as GTPases that directly interact with each other to mediate SRP–SR (FtsY) complex assembly. Formation of a stable complex requires extensive structural rearrangements, such as removal of the steric

hindrance posed by the N-domains of both SRP and SR and alignment of GTP molecules to form a cyclic pair of hydrogen bonds across the dimer interface. The SRP RNA, which is universally conserved, accelerates this otherwise very slow process 200-fold. The effect is purely catalytic as it also accelerates complex disassembly without changing its equilibrium stability. It is the first example of an RNA molecule catalyzing a protein-protein interaction. FtsY–Lys399 residue on the lateral surface of the FtsY G-domain provides a key site that mediates the SRP RNA-induced stimulation of complex assembly (**Figure 4**).
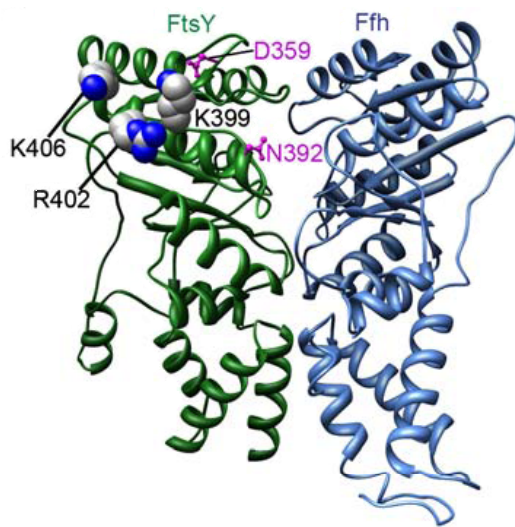
Figure 4. FtsY-Lys399 plays a crucial role in SRP–FtsY complex assembly. The basic residues on the FtsY Gα2-helix are highlighted in spacefill in the crystal structure of the *Thermus aquaticus* Ffh-FtsY NG-domain complex. (Shen and Shan, 2010)

It seems that the positive charges on and surrounding Lys399 are involved in electrostatic interactions with the RNA backbone. FtsY–K399A mutation reduces the kinetics of SRP–FtsY complex formation 82-fold. Lys399 also provides a key link that couples cargo binding by the M domain of SRP to efficient SRP–receptor interactions. It has been estimated that signal peptide binding accelerates this process 100-400 fold. Together, the combined effect of the cargo and the SRP RNA brings the SRP–FtsY interaction kinetics to a range of $>10^6$ M^{-1} s^{-1}, appropriate for cotranslational protein targeting in the cell (Shen and Shan, 2010).

3. Post-translational translocation

3.1 Post-translational translocation in eukaryotic cells

In yeast, there is a considerable number of proteins delivered to the ER via a post-translational route. These are often secreted proteins containing N-terminal, cleavable signal sequences. In mammalian cells, only a few proteins are known to be targeted via this pathway.

Such proteins are usually shorter than 75 amino acids, which is below the minimal size of a nascent polypeptide chain to cotranslationally interact with SRP via its signal sequence. Fully synthesized precursor polypeptides are transported with the help of cytosolic molecular chaperones of the Hsp70 and Hsp40 chaperone families (Ngosuwan et al., 2003). By cycling on and off, they assure that the substrate polypeptide is soluble and competent for interaction with the transport machinery in the ER membrane (**Figure 5**). In yeast, targeting via the post-translational pathway involves a heterotetrameric complex of membrane proteins: Sec62p, Sec63p, Sec71p (also called Sec66p) and Sec72p (also termed Sec67p) that might serve as a signal peptide receptor. In both co- and post-translational insertion, polypeptide translocation across the target membrane is facilitated by the heterotrimeric Sec61 complex: Sec61αβγ in mammalian cells; Sec61p, Sbh1p, Sss1p in yeast, SecYEG in bacteria. In eukaryotes it involves additional components, such as the ER-lumenal chaperone BiP (Kar2p in yeast), its membrane receptor and co-chaperone, Sec63, and its NEFs: Sil1p and Lhs1p (Lyman and Schekman, 1995). In yeast there is an additional heterotrimeric Sec61 complex, comprising Ssh1p, Sbh2p and Sss1p. The genes coding for Sec61p, Sss1p, Sec62p, Sec63p and Kar2p are essential.
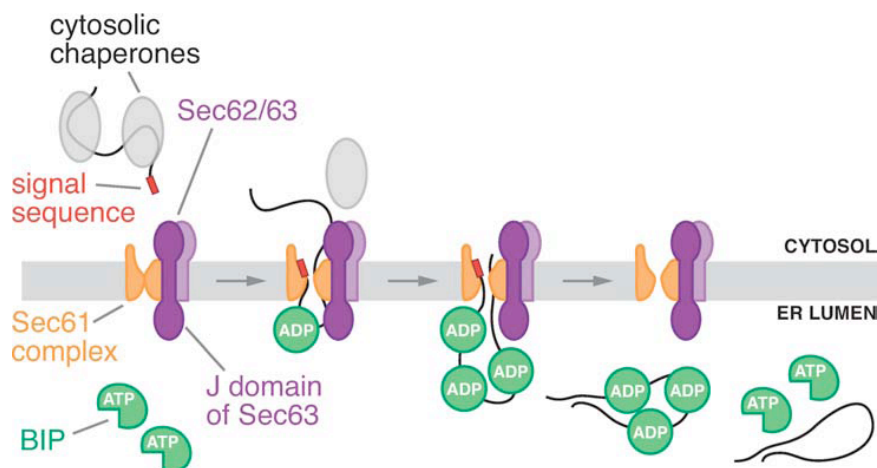
Figure 5. Post-translational translocation in eukaryotic cells. (Osborne et al., 2005)

Molecular chaperones of the Hsp70 family reversibly bind to substrate polypeptides via their substrate binding domains (SBD). Hsp70 binds to hydrophobic stretches of essentially unfolded polypeptides emerging from the ribosomal tunnel. The process of binding and release is modulated by communication between the SBD and the nucleotide binding domain (NBD). The latter is controlled by the ATPase cycle and different Hsp70 interaction partners. BiP in its ATP-

bound state has a low affinity for substrate polypeptides; ATP hydrolysis promotes polypeptide binding. Nucleotide exchange factors (NEFs) stimulate the ADP to ATP exchange, therefore inducing substrate release. Hsp40 proteins, such as Sec63, contain the J-domain that allows interaction with Hsp70. Immediately after the insertion of a precursor polypeptide into the Sec61 complex, BiP binds its N-terminal end, preventing its back-sliding into the cytosol. Subsequently, the polypeptide chain is stepwise transported across the channel by the molecular ratchet mechanism which is described as trapping of portions of the polypeptide in the ER lumen by BiP molecules. Additionally, BiP may be involved in opening of the Sec61 translocon by causing plug displacement (Zimmermann et al., 2010). Mammalian ER membranes contain the additional Hsp40, ERj1, which is related to Sec63 in providing a lumenal J-domain (**Figure 6**, right panel).

In yeast, the negatively charged C-terminus of Sec63p interacts with the overall positively charged N-terminal domain of Sec62p. Recently, it has been demonstrated that this mode of interaction is conserved from yeast to humans (**Figure 6**) and in the course of evolution vertebrate Sec62 has gained a function, i.e., the ability to intract with the ribosomal tunnel exit. The human Sec62/Sec63 complex and human ERj1 are similar in providing a binding site for ribosomes in the cytoplasm and binding BiP on the lumenal side of the ER membrane, meaning they can both be involved in co-translational transport. This is supported by experiments in which Sec62 is protected against externally added antibodies by ribosomes in permeabilized human cells (Muller et al., 2010).
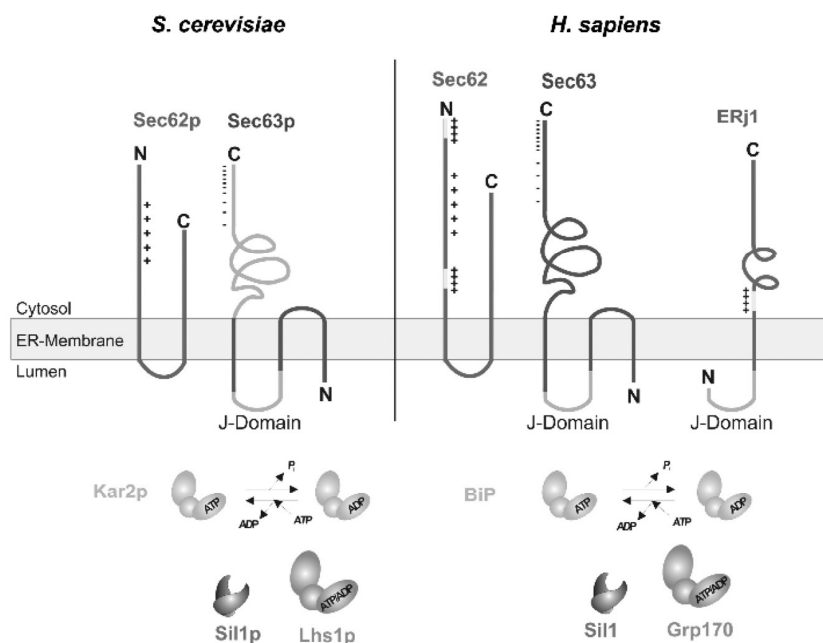
Figure 6. Structural and functional characteristics of yeast and human Sec62/Sec63 complex. Sec63 and ERj1 are membrane resident Hsp40s with ER lumenal J-domains and cochaperones for ER-lumenal Hsp70s (BiP and Kar2p). Sil1, Grp170 and Lhs1p act as NEFs for these Hsp70s. (Muller et al., 2010)

Müller at al. suggested that both Sec62/Sec63 and ERj1 could recruit BiP to Sec61 complexes and/or incoming polypeptides. They observed that BiP binding to the J-domain of ERj1 still allows binding of the ERj1/BiP complex to the ribosome and this heterotrimeric complex does not inhibit translation. A similar mechanism may be in operation for the Sec62/Sec63 complex. There could exist a functional specialization: one system could recruit BiP to ribosomes and the Sec61 complex in order to seal off the channel, while the second system could recruit BiP to act as a molecular ratcher for incoming polypeptides. Alternatively, these two systems may have different substrate specificities.

3.2 Insertion of tail-anchored proteins

Tail-anchored (TA) proteins are a large and diverse class of integral membrane proteins found in all organisms. They comprise nearly 5% of membrane proteins, examples include Sec61ß, Sec61γ, RAMP4 (ribosome-associated membrane protein 4), VAMPs (also called synaptobrevins), cytochrom b5 and the Bcl2 family of apoptotic proteins. TA proteins lack an N-terminal signal sequence. They are anchored to the membrane by a single C-terminal transmembrane domain (TMD), exposing their larger N-terminal (and usually functional) part to the cytosol (Favaloro et al., 2008). The targeting information for TA proteins resides solely within the TMD. Because this region is still within the ribosomal tunnel when the termination codon is reached, SRP binding and co-translational targeting is precluded. Thus, TA proteins must find their correct membrane for insertion post-translationally.

A central component of the TA protein pathway to the ER is a highly conserved cytosolic ATPase termed Asna1 or TRC40. Both mammalian TRC40 and its yeast homologue Get3 (for Guided Entry of Tail-anchored proteins) recognize and bind the TMD of TA proteins in the cytosol in a selective manner. This complex targets to the ER by membrane-bound receptors (Get1 and Get2 in yeast), where the transported protein is released for insertion. The process is regulated by ATP binding and hydrolysis. Mateja et al. in 2009 determined the crystal structure of Get3 from *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* in the open and closed (ADP•AlF$_4^-$ -bound) state, respectively (**Figure 7A** and **B**). Both structures show a symmetric homodimer and each monomer comprises a core ATPase subdomain and an α-helical subdomain. In the Get3 open dimer, the α-helical counterparts are separated, creating a large, charged cleft between the two subunits that is unsuitable for TMD binding. In contrast, the closed dimer state is characterized by a continous, solvent-exposed, hydrophobic groove that spans both monomers and

is proposed to be suitable for binding to an α-helical TMD of ~20 residues. Similar to the M-domain of SRP, the hydrophobic groove of Get3 is rich in methionines that could accommodate diverse TA protein targeting signals.

Recently, Chartron et al. (2010) obtained crystal structures of additional components of TA proteins targeting machinery, Get4 and Get5, which operate upstream of Get3, and proposed a functional model (**Figure 7C**). Get4 and Get5 are highly conserved proteins which form a complex that dimerizes, mediated by the C-domain of Get5. The Get4 N-terminal face forms part of the recognition interface with Get3. Targeting is initiated in the cytosol when ATP binding drives Get3 towards the closed dimer state, facilitating recognition of newly synthesized TA proteins in a TMD-dependent manner. Get4/5 are able to recognize the nucleotide state of Get3 and localize its closed form to the ribosome. Binding of the TA protein to Get3 leads to a conformational change that releases the Get3/TA complex from Get4/5 and the ribosome. Sgt2 and cellular chaperones could either facilitate this transfer or act as parts of an alternate pathway.
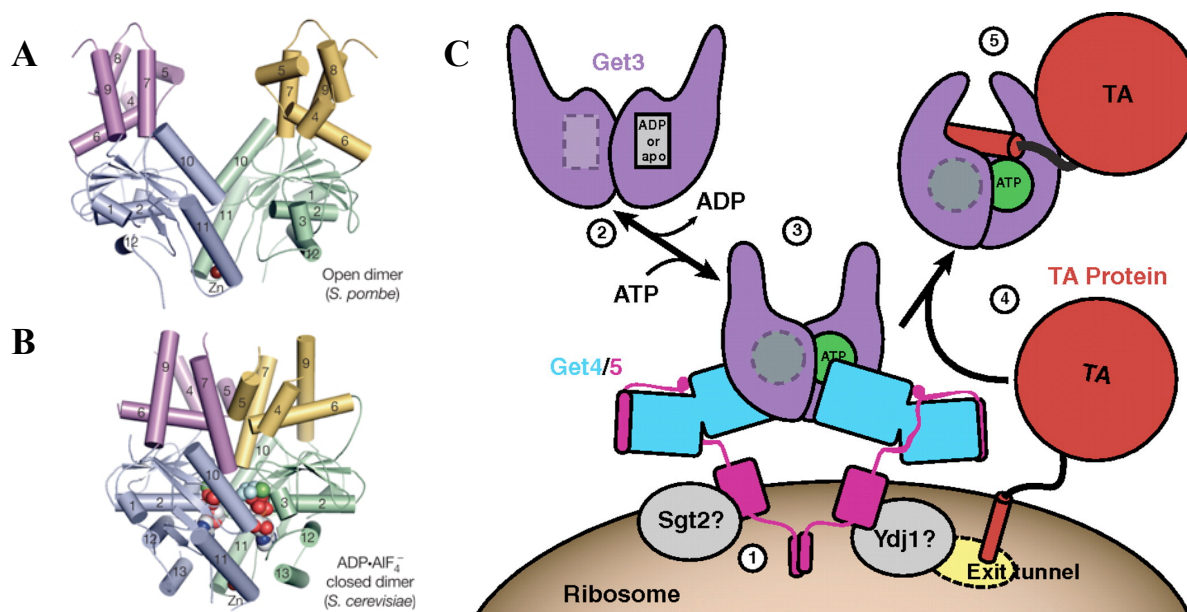
Figure 7. Crystal structures of Get3 in open (A) and closed (B) dimer states. Each monomer comprises a core ATPase subdomain (blue, green) and an α-helical subdomain (magenta, yellow). A tightly bound zinc atom (brown sphere) lies at the dimer interface (Mateja et al., 2009). (C) A model for the role of Get4/5 (Chartron et al., 2010).

3.3 Post-translational translocation in bacteria

In 2007 Osborne and Rapoport proposed the dimer model of post-translational translocation in *E. coli*, where one copy of SecYEG provides a docking site for the cytosolic

ATPase SecA, whereas another copy is used as a translocation channel. A year later Tsukazaki et al. (2008) obtained a crystal structure of SecYEG from *Thermus thermophilus* and via molecular dynamics and disulphide mapping analysis identified the binding sites on SecY and SecA that trigger conformational changes in both molecules on formation of the functional complex (**Figure 8**).
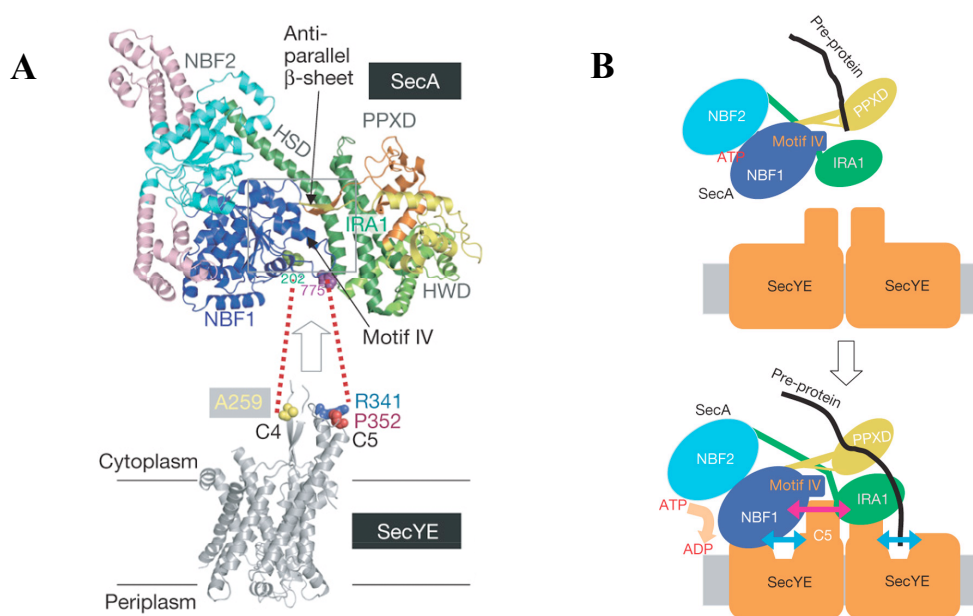
Figure 8. Interactions during post-translation translocation in eubacteria. (A) Contact residues between *Thermus thermophilus* SecA and SecYE. The SecA structure (Protein Data Bank 2IPC) is colour-coded for its domains. (B) According to the dimer model of protein translocation, one copy of SecY serves as a SecA-docking site, while another functions as a translocation pore. Both components undergo conformational changes upon their interaction, as shown by bidirectional arrows. (Tsukazaki et al., 2008)

SecA contains two nucleotide-binding folds (NBFs, also called NBDs), a pre-protein crosslinking domain (PPXD) and a C-terminal translocation domain (HWD and IRA1), which are all connected by a long α-helical scaffold domain (HSD). NBF1 has been shown to be in the physical proximity of SecY (Osborne and Rapoport, 2007). Disulphide crosslinking experiments revealed that residues 775 (**Figure 8A**, purple) and 202 (green) of SecA are adjacent to the C5 and C4 residues of SecY, respectively. An evolutionary conserved region corresponding to the C-terminal half of motif IV interacts with SecY and undergoes a conformational change that is coordinated with the formation of a motor-translocon complex.

The ATPase of SecA is tightly downregulated in the resting state through its interaction with the IRA1 domains. Binding to the channel physically separates NBF1 and IRA1, explaining the translocon-mediated triggering of the membrane ATPase. Motif IV of SecA communicates

with an anti-parallel ß-sheet that is involved in propagation of a pre-protein binding signal to the ATPase domain.

SecA interacts with SecY in at least two different modes: the one involving the SecY C4-C5 domains is probably required for ATPase activation, while the other, involving C6, participates in the actual SecA-driven translocation (Tsukazaki et al., 2008).

4. Structure of the translocon

4.1 Crystallography and X-ray analysis

In 2004 the first crystal structure of a Sec61 complex family member was obtained for the heterotrimeric SecYEß complex (corresponding to eukaryotic α, γ, and β subunits) from the archae *Methanococcus jannaschii* at 3.2 Å resolution (Van den Berg et al., 2004). The translocon has an hourglass shape when viewed from the membrane and a square shape in a view from the cytosol. The α-subunit is divided into two halves: transmembrane segments (TM) 1-5 and TM 6-10. The loop between TM5 and 6 at the back of the α-subunit serves as a hinge, allowing opening at the front - the so called lateral gate. The γ-subunit links the two halves of the α-subunit at the back by extending one transmembrane segment diagonally across their interface. The ß-subunit is in contact only with the periphery of the α-subunit and this probably explains why it is dispensable for the function of the complex (**Figure 9**).

The α-subunit forms the pore of the channel, as was initially shown in experiments that involved incorporation of photoreactive probes into a stalled translocating polypeptide and cross-linking to the α-subunit (Mothes et al., 1994). The channel has an aqueous interior revealed by electrophysiology experiments (Simon and Blobel, 1991) and measurements of the fluorescence lifetime of probes incorporated into a translocating polypeptide (Bol et al., 2007). The ten helices of the α-subunit form an hourglass-shaped pore that consists of cytoplasmic and exoplasmic funnels, whose tips meet about half way across the membrane. In the closed state, the cytoplasmic funnel is empty, while the exoplasmic half of the channel is obstructed by a short helix, called the plug. The constriction of the pore is formed by a ring of six hydrophobic residues that project their side chains radially inward. These are amino acids with bulky, hydrophobic side chains (Rapoport, 2007).

The diameter of the pore ring calculated from the crystal structure is too small to allow the passage of most polypeptide chains. Therefore pore widening was postulated, which could

proceed via movement of the helices to which the pore ring residues are attached. Results from fluorescence-quenching experiments with ER membranes estimated the pore size to be 9-15 Å in the resting state and it widens to 40-60 Å during translocation (Hamman et al., 1997; Liao et al., 1997), which is difficult to reconcile with the crystal structure.
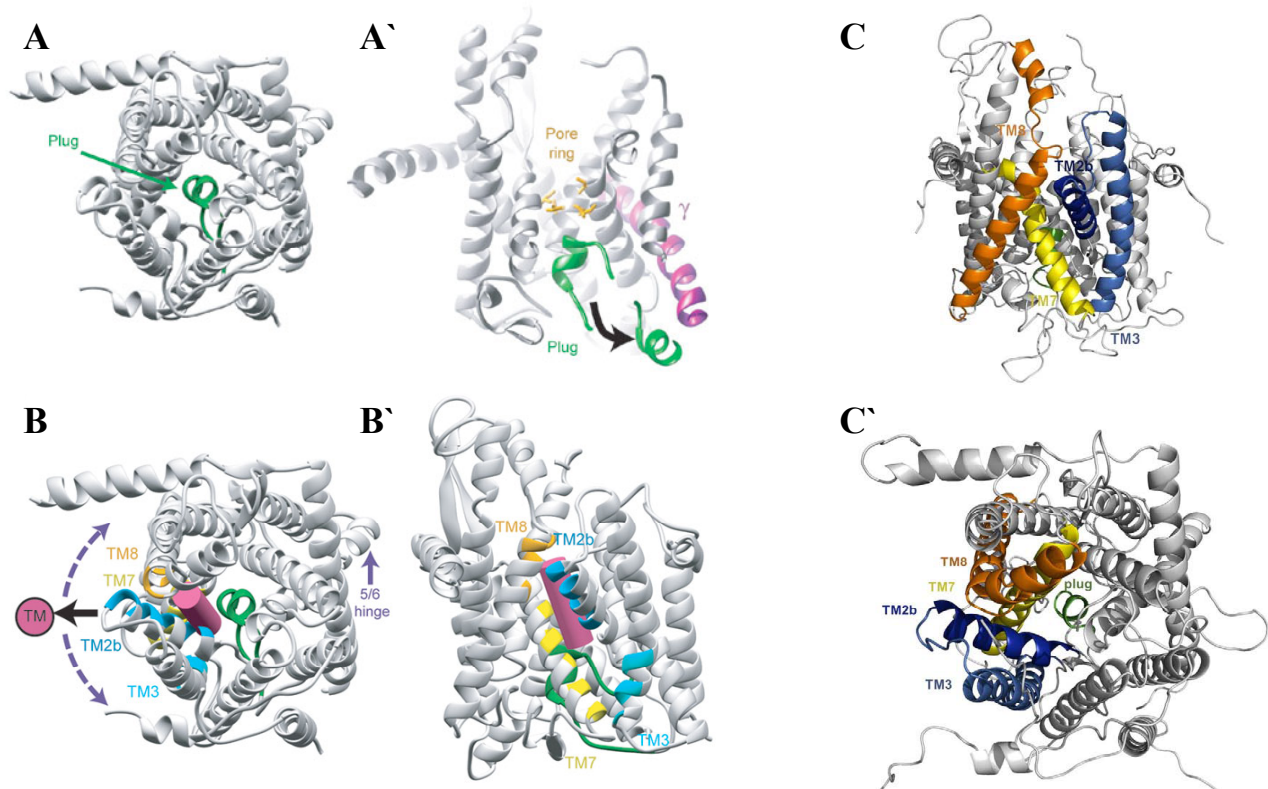
Figure 9. Architecture of the SecY complex from *Methanococcus jannaschii*: view from the top with the transmembrane domain 2a (TM2a, plug) coloured in green (A). View from the side with the front half of the model cut away. The modelled plug movement towards the γ-subunit (magenta) is indicated. The hydrophobic pore ring is shown by the side chains coloured in gold (A`). Signal-sequence-binding site and lateral gate: views from the top (B) and the front (B`), with faces of the helices that form the signal-sequence-binding site and the lateral gate through which TMs of nascent membrane proteins exit the channel into lipid highlighted in colours. The plug is coloured in green. The hydrophobic core of the signal sequence probably forms a helix, modelled as a magenta cylinder, which intercalates between TM2b and TM7 above the plug. Intercalation requires opening the front surface, as indicated by the broken arrows, with the hinge for the motion being the loop between TM5 and TM6 at the back of the molecule (5/6 hinge). A solid arrow pointing to the magenta circle in the top view indicates schematically how a TM of a nascent membrane protein would exit the channel into lipid. (Van den Berg et al., 2004)
Homology model for the human heterotrimeric Sec61 complex: views from the plane of the membrane (C) and from the cytosol (C`). The model was generated using the SecY X-ray structure (PDB 1rhz) by the program MODELLER 9.5, it was optimised with the variable target function and refined using molecular dynamics and simulated annealing. (Zimmermann et al., 2010)

Opening of the channel during protein translocation across the membrane probably occurs in two steps. The first step is binding of the channel to its partner (the ribosome, Sec62/Sec63

complex or SecA), which likely destabilizes the interactions with the plug domain, resulting in transient displacement of the plug and continous opening and closing of the lateral gate. In the second step, the hydrophobic core of the signal sequence intercalates into the walls of the channel between TM2b and TM7 (**Figure 9B** and **B`**). The signal can also be crosslinked to phospholipids, indicating it is located at the channel/lipid interface. The binding of the signal sequence separates TM2b and TM7 and causes the plug to move away from the exoplasmic cavity. Crosslinking studies showed it is located to the proximity of SecE (Tam et al., 2005). During translocation of the polypeptide, the signal sequence stays put, while the rest of the polypeptide chain moves through the pore.

4.2 3D-reconstruction after cryo-EM

There has been a lot of debate concerning the oligomeric state of the translocon. Experiments with solubilized ribosome-Sec61 complex revealed that a ribosome may bind firmly with only one Sec61 channel (Menetret et al., 2008). On the other hand, oligomeric states of the Sec61 and SecY complexes were observed by electron microscopy (Hanein et al., 1996; Mitra et al., 2005) and biochemical analysis (Mori et al., 2003; Schaletzky and Rapoport, 2006). Recent data analyzed with blue-native PAGE shows that oligomeric states of the SecY channel may change depending on the substrate (Boy and Koch, 2009).

Cryo-EM is a powerful technique for structural analysis of complexes between ribosomes and the translocon components that could help solve this issue. Using this technique, Becker et al. in 2009 managed to obtain the structure of the translating ribosome interacting with monomeric yeast and mammalian Sec61 complex (**Figure 10**). For the first time, the nascent polypeptide chain was visualized inside the ribosomal tunnel. Structure determination was based on digitonin-solubilized purified Ssh1 complex from *S. cerevisae*. This complex is active in the co-translational translocation mode only. The complex was reconstituted with 80S ribosomes carrying a nascent polypeptide chain composed of the first 120 amino acids of dipeptidyl amino peptidase B (DP120) together with its signal-anchor sequence. For the analysis of the mammalian translocon, purified Sec61 complex from *Canis familiaris* bound to an active DP120 signal-anchor containing 80S ribosome was used. 3D reconstruction together with biochemical data revealed that only a single copy of the Sec61 complex is recruited to both translating and nontranslating ribosomes. In the yeast Ssh1, as well as in the mammalian translocon, the DPAPB chain accomodated within a single copy of the protein conducting channel (PCC). The lateral gate of the PCC can be in

a closed or nearly closed conformation after polypeptide insertion. The mode of PCC binding to ribosomes appears to be conserved between species and is maintained in the presence or absence of a signal sequence. The main binding site for the channel is the universal adaptor site at the ribosomal tunnel exit and includes mainly the cytoplasmic loop L8 of the α-subunit, whereas loop L6 is also in contact with the emerging nascent chain and may function in sensing or guiding the peptide to the pore of the PCC. This observed mode of Sec61 binding fits with the authors` previous findings that the universal adaptor site also serves to bind SRP and is cleared upon SRP receptor interaction to enable PCC binding (Halic et al., 2004).
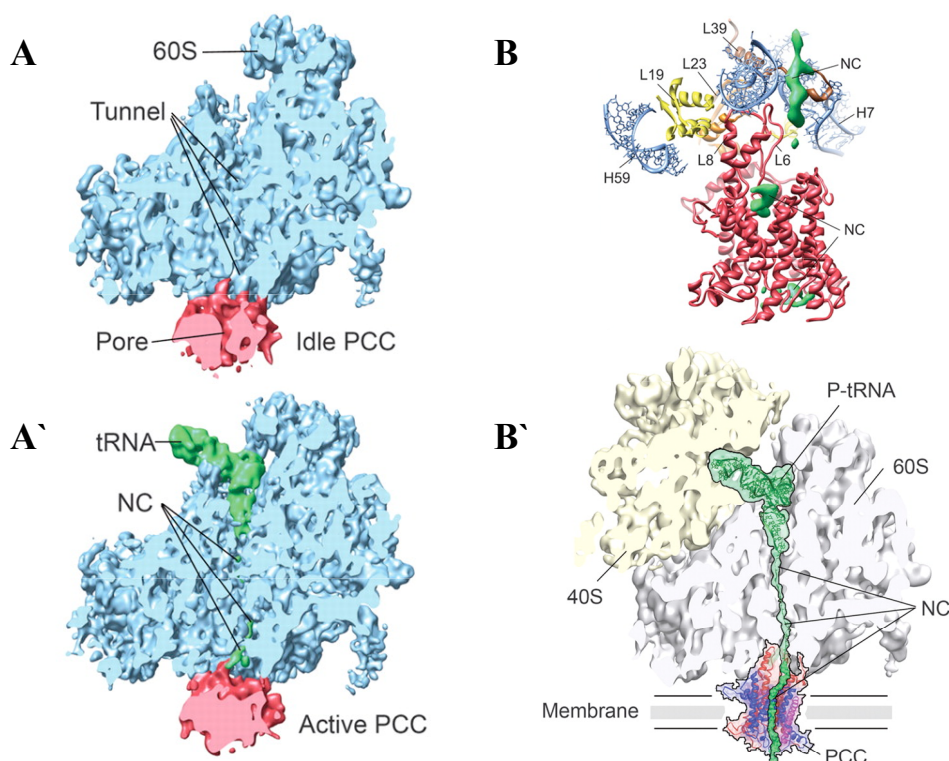
Figure 10. Cryo-EM reconstruction of 80S ribosome-translocon complexes. Visualization of the PCC pore of the Ssh1 complex associating with the idle (A) and active ribosome (A`). The nascent chain (NC) is visible inside the ribosomal tunnel exit. Side view of the mammalian Sec61 model with the nascent chain (green) entering the PCC (B). Schematic representation of an actively translocating eukaryotic ribosome-Sec61 complex with a single copy forming the PCC (B`). (Becker et al., 2009)

5. Co-translational processing of the polypeptide

5.1 Signal peptide cleavage

In eukaryotic cells, many proteins that enter the endoplasmic reticulum for either retention in the ER or for export to the Golgi apparatus, secretory vesicles, plasma membrane, or

vacuole/lysosome are processed by the ER signal peptidase complex (SPC) (Paetzel et al., 2002). In bacteria, there are at least three distinct signal peptidases (SPases) involved in cleaving signal peptides. Among them, the essential SPase I can process protein substrates that are exported by the SecYEG complex. The substrate specificity of the *E. coli* SPase and the eukaryotic SPC is similar and conserved in evolution. In bacteria and eukaryotic cells, preproteins have a common pattern in the c-region of the signal peptide at the -1 and -3 positions (**Figure 11A**). Gunnar von Heijne proposed that the region preceding the cleavage site constitutes the substrate recognition site for the SPase enzyme (von Heijne, 1983). Residues at the -1 position that are tolerated are alanine, glycine, serine, cysteine, and, in some cases, threonine. Residues tolerated at the -3 position are alanine, glycine, serine, cysteine, isoleucine, valine, and leucine. Proper processing of signal peptides requires the cleavage sites to be located in close proximity to the h/c-region border.

Signal peptides released by the SPase usually disappear in a short time period. This may be due to the fact that signal peptides are harmful to cells since they can inhibit protein translocation and may be lytic if they accumulate in the membrane. In both bacteria and the ER, signal peptides are degraded by the signal peptide peptidase, which cleaves within the hydrophobic core region of the signal peptide (**Figure 11B**). In eukaryotic cells, some of the cleaved and released signal peptides are bioactive, affecting pathways such as immune surveillance, virus maturation or cellular signaling (Martoglio, 2003).
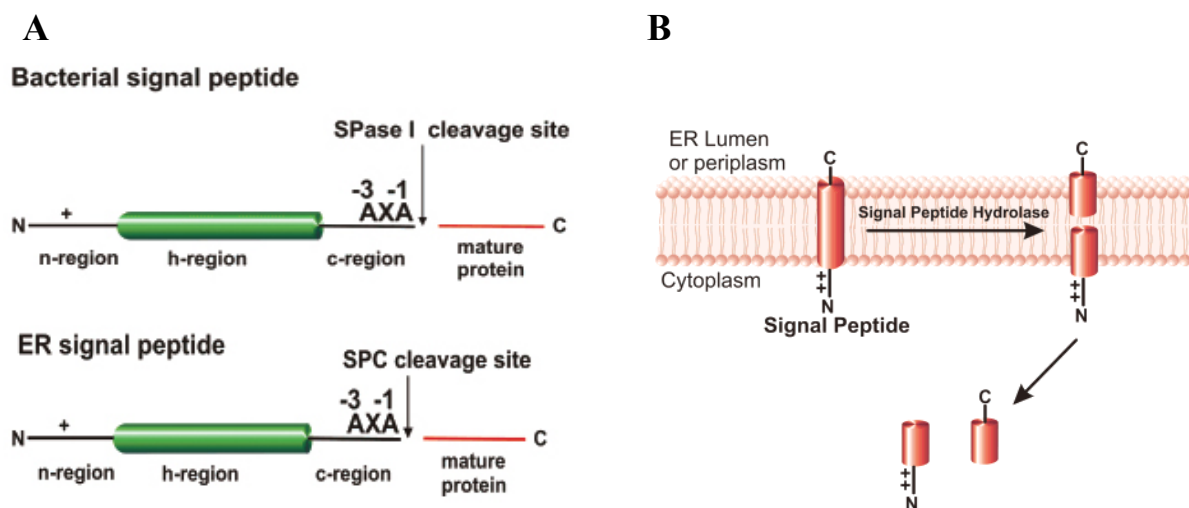
Figure 11. Targeting signal sequences and signal peptidase cleavage sites in bacteria and eukaryotes (A). The cleaved signal peptide resides within the lipid bilayer until it is processed within its transmembrane segment by a signal peptide hydrolase. The smaller cleaved peptides from this reaction most likely recede back into the cytosol where they are degraded by proteases or become involved in signaling events (B). (Paetzel et al., 2002)

5.2 N-linked glycosylation

N-glycosylation occurs in all eukaryotic cells and in a few prokaryotes. In the bacterium *Campylobacter jejuni*, a single protein, PglB mediates N-glycosylation in the periplasm, independently of protein translocation. In eukaryotes, the process is catalyzed by a large transmembrane complex called oligosaccharyl transferase (OST). N-glycosylation is coupled to protein translocation across the ER membrane and OST is in a direct contact with the translocon and the translocating ribosome. Coupling of these events is necessary to prevent partial folding of the polypeptide chain, which would render the N-glycosylation sites inaccessible to OST.

The yeast OST contains consists of nine integral proteins: Ost1p-Ost6p, Stt3p, Swp1p, and Wbp1p. Among these, five subunits (Ost1p, Ost2p, Stt3p, Swp1p, and Wbp1p) are essential for cell viability. The enzyme catalytic site is contained within Stt3p. In mammalian cells, the following homologues of OST proteins are present: ribophorin I, DAD1, N33/IAP, OST4, STT3A/STT3B, Ost48, and ribophorin II.

To facilitate N-glycosylation, OST must bind to the donor and acceptor substrate, cleave and transfer the N-glycan precursor from the dolicholpyrophosphate to the polypeptide and catalyze the formation of a covalent bond between the oligosaccharide and the asparagine of the -N-X-S/T- (where X cannot be proline) acceptor sequence. In 2008 Li et al. presented a cryo-EM structure of the yeast OST at 12 Å resolution. It shows a groove between the lumenal domains of Stt3p, Wbp1p, and Ost1p that is approximately parallel to the ER membrane. The authors proposed it functions as a tunnel through which the nascent polypeptide chain threads during cotranslational glycosylation. Ost1p scans for the -N-X-S/T- sequence as the polypeptide moves through the groove. Once this sequence is detected, the oligosaccharide bound to Wbp1p is transferred onto the acceptor Asn (**Figure 12A**). Peptide threading can enhance the detection efficiency, explaining the fact that most sequences are actually glycosylated (Li et al., 2008).

OST utilizes the activated glucose$_3$-mannose$_9$-N-acetylglucosamine$_2$ (Glc$_3$Man$_9$GlcNAc$_2$) oligosaccharide donor as a substrate for covalent modification of the acceptor Asn side chains (**Figure 12B**). The hydrophilic structure of the carbohydrate affects the solubility and folding of proteins. More importantly, it can be modified by ER glycosyl hydrolases and one glucosyl transferase in a cascade of reactions which gives rise to structures that serve as ligands for carbohydrate binding proteins, lectins. Lectin binding is the prerequisite for protein folding and quality control, which result either in protein export from the ER or degradation. The folding program involves a concerted action of glucosidase I and glucosidase II, which allows immediate association of polypeptides emerging in the ER lumen with ER chaperones, calnexin and

calreticulin, followed by the oxidoreductase ERp57, resulting in co-translational formation of native disulfide bonds. Upon attainment of the native structure, most of cargo proteins are exported from the ER in vesicles coated with cytosolic coatamer protein II (COPII) which bud at ER exit sites. In mammals, transport vesicles undergo homotypic fusion to generate a stationary ER-Golgi intermediate compartment (ERGIC) from which cargo proteins reach the *cis*-Golgi in COPI-coated vesicles. In yeast, COPII-coated cargo vesicles are delivered directly to the Golgi apparatus. ER export of certain glycoproteins is facilitated by leguminous L-type lectins located in the ER (VIPL), cycling between the ER and the ERGIC (ERGIC-53) or between ERGIC and *cis*-Golgi (VIP36). Yeast orthologs of ERGIC-53, Emp47p and Emp46p, have been proposed to act as cargo receptors between the ER and the Golgi (Aebi et al., 2010).
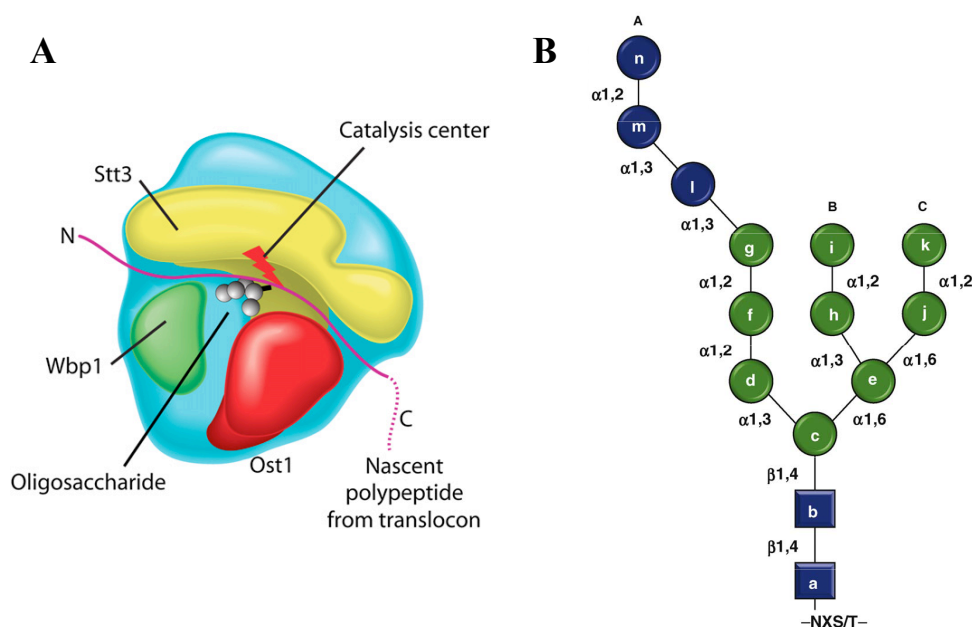
Figure 12. Peptide-threading mechanism for oligosaccharyl transferase. OST is illustrated in its lumenal view (A) (Li et al., 2008). The N-linked core oligosaccharide structure is composed of two N-acetylglucosamine (blue squares), nine mannose (green circles) and three glucose residues (blue circles). A, B and C define the oligosaccharide branch (B) (Aebi et al., 2010).

6. Topogenesis of single-spanning membrane proteins

Signal sequences have a dual function: they enable targeting of the protein to the appropriate organelle and play an important role in protein topogenesis by orienting themselves in the translocation channel prior to their release into the lipid environment of the membrane (Spiess, 1995). A signal sequence can initiate translocation of the C-terminal part of the polypeptide across

the ER (proteins of type I, type II and tail-anchored) or induce transfer of the N-domain (type III membrane proteins) (**Figure 2**). Several factors determine the orientation of the signal in the membrane (**Figure 13**):
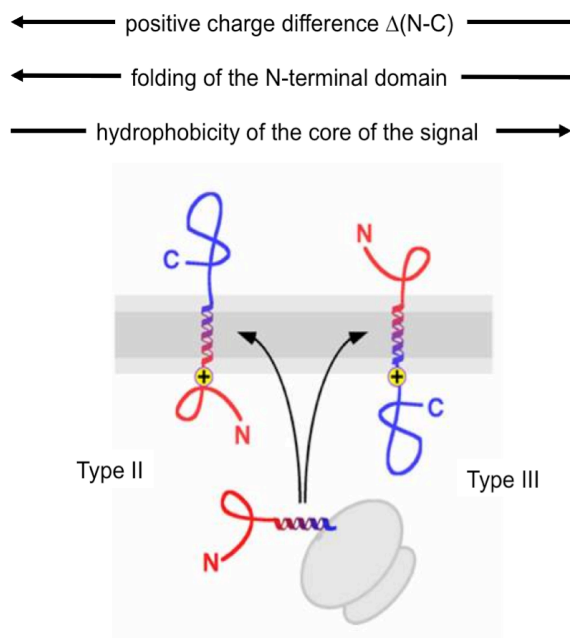
Figure 13. Determinants of signal sequence orientation in the membrane.

a) distribution of charged residues flanking the hydrophobic core of the signal-anchor sequence – the more positive end is retained on the cytosolic side. This phenomenon is known as the `positive-inside rule` and was first described for bacterial proteins, where charged amino acids were found to be more abundant in cytoplasmic than in periplasmic protein loops (Heijne, 1986). For eukaryotic proteins, the charge difference between the N- and C-flank of the signal core, rather than the presence of positive charges influences protein topology: the more positive end is cytosolic (Hartmann et al., 1989). Type III membrane proteins can be converted to type II by engineering their charges and vice versa. Examples include conversion of cytochrome P-450 (type III) to a type II protein by insertion of positively charged amino acids into its short N-terminal domain (Monier et al., 1988; Szczesna-Skorupa et al., 1988; Szczesna-Skorupa and Kemper, 1989) or mutagenesis of flanking charges of the asialoglycoprotein (ASGP) receptor H1 subunit (type II), which causes a fraction of the polypeptides to insert in an inverted topology (Beltzer et al., 1991). Positive charges have a stronger influence on topogenesis than negative ones, the effect depends on their distance from the hydrophobic segment of the signal. However, Kida et al. showed that orientation of

the synaptotagmin signal (type III) can be affected by a cluster of lysines positioned up to 25 residues downstream the hydrophobic segment (Kida et al., 2006). Proper distribution of the flanking charges is not the only requirement a signal sequence must fulfill in order to attain the right topology.

b) folding of the N-terminal domain – during translocation, this portion of the polypeptide is exposed to the cytosol before the targeting signal emerges from the ribosomal tunnel exit. Folding of this domain acts as a steric hindrance that prevents its translocation into the ER lumen and favors its retention in the cytosol instead, irrespective of the flanking charges. This was demonstrated by truncation of the N-terminal domain of H1, which shifted the topology towards N-insertion, whereas fusing additional sequences to the N-terminus abolished N-translocation. The observed effect was caused by protein folding and not simply due to the size of the N-extension (Denzer et al., 1995).

c) hydrophobicity of the core of the signal sequence – an influence of the hydrophobicity of the signal on membrane protein topogenesis was first suggested by Sato et al. and Sakaguchi et al. Truncations of the hydrophobic segment of the cytochrome P-450 signal sequence (type III) resulted partially in C-terminal translocation (Sato et al., 1990). Similarly, an artificial sequence with the core consisting of less than 12 leucines and a negative N-terminal net charge allowed translocation of the C-terminus, whereas with longer hydrophobic segments of 13 or 15 leucines, a fraction of the polypeptides was inserted in type III orientation (Sakaguchi et al., 1992).

The process of insertion of single-spanning membrane proteins into the ER was studied extensively by Goder and Spiess in 2003 (**Figure 14**) and led to a model of signal sequence orientation in the translocon (**Figure 15**). They employed a series of diagnostic constructs derived from the H1 subunit of human ASGP receptor (type II membrane protein), where the topogenic determinants were altered to generate mixed topologies. Namely, the N-terminal hydrophilic domain preceding the signal-anchor was truncated to only four residues (MGPR), which were subsequently mutated to MGPQ, MGPH or MGRR to yield constructs with different flanking charges. The apolar core of the signal was replaced by a stretch of 13-23 leucines. Additionally, the C-terminal portion of the protein was shortened from 230 to 170 and 110 amino acids (aa), or extended to 290, 350, 400, 460, 520 and 580 aa. Glycosylation analysis revealed that topology of constructs with the same signal sequence depends on the length of the polypeptide, with C-terminal translocation increasing with the size of the C-terminus, reaching a maximum for polypeptides of ~300 amino acids or more.
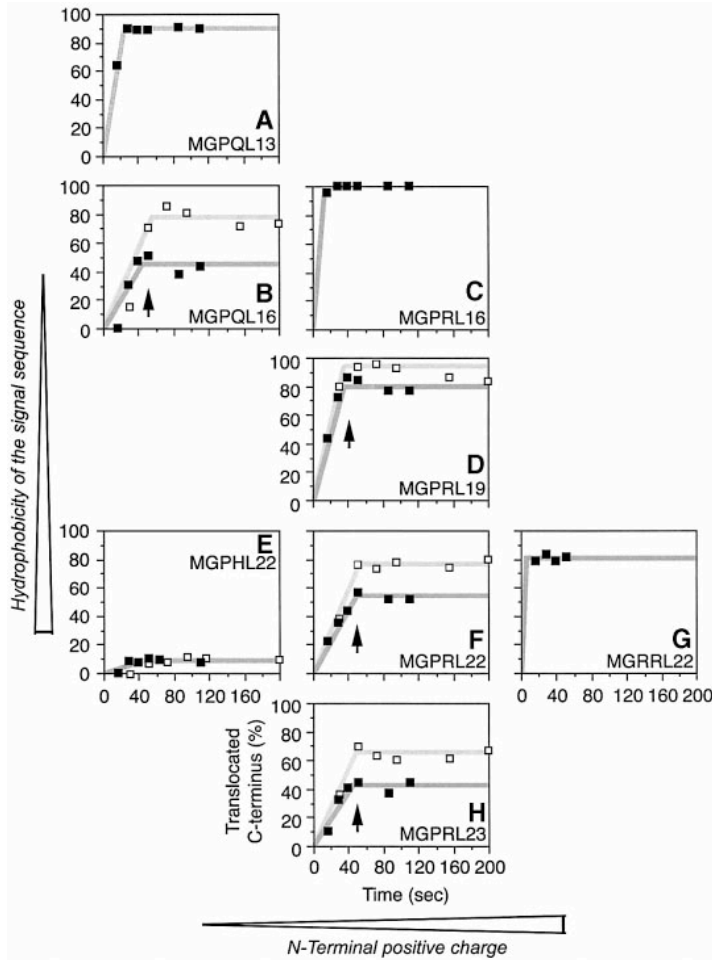
Figure 14. Effect of flanking charges and hydrophobicity on signal inversion. Series of proteins with increasing or decreasing hydrophobicity of the signal and different N-terminal flanking sequences are analyzed. Protein orientation is plotted versus translation time (5 aa/s). Experiments were performed in the absence (filled squares) or presence of 1 µg/ml cycloheximide (open squares). The arrows mark the termination of topogenesis. (Goder and Spiess, 2003)
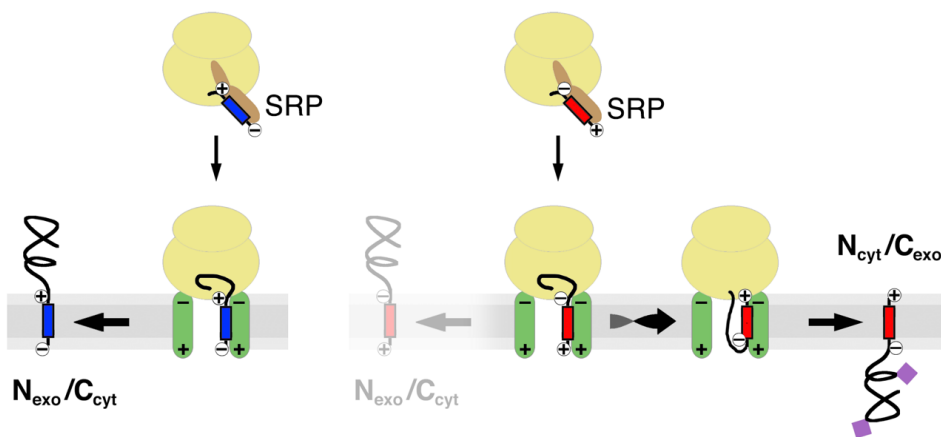
Figure 15. Model of insertion of N-terminal signal anchors (red) and reverse signal-anchors (blue). Both signals initially engage with the translocon in an N_{exo}/C_{cyt} orientation (`head-on` insertion). Signals can invert their orientation over time, driven by flanking charges according to the `positive-inside rule`. The process is slowed down by increased hydrophobicity of the signal core and stopped upon termination of translation or by a signal-independent mechanism after ~50 s. This results in a fraction of the polypeptides inserted in the false, N_{exo}/C_{cyt} topology (grayed-out portion). The reorientation process can be monitored by the glycosylation status of proteins (purple diamons indicate N-linked glycans attached in the ER lumen to the C-terminus of proteins). The SRP receptor was ommited for simplicity. (Goder and Spiess, 2003; Higy et al., 2004)

When the translation rate was reduced with low concentrations of the reversible elongation inhibitor, cycloheximide, the topology ratio shifted in favor of C-terminal translocation. Under both conditions, the maximal amount of C-insertion was reached at the same time, after ~50 s, based on the translation rate in mammalian cells of ~5 aa/s (Hershey, 1991) and the estimation that ~40 residues are hidden within the ribosomal tunnel. Both flanking charges and hydrophobicity influenced the topology, but did not affect the time when topogenesis stopped (**Figure 14**). This suggested that signal sequences initially insert into the channel `head-on`, with N_{exo}/C_{cyt} topology, and undergo reorientation within the translocon that is interrupted by termination of translation (`translation-stop phenomenon`). It explains why short polypeptides showed predominantly N-terminal translocation. If translation is slowed down with cycloheximide, short constructs gain time for signal inversion and produce more N_{cyt}/C_{exo} topology. The rate of inversion depends on flanking charges, which position the signal according to the `positive-inside rule`. Importantly, also the hydrophobicity of the core of the signal-anchor affects the rate of reorientation. As the signal needs to dissociate from the apolar binding site in the translocation pore to reorient itself, long and hydrophobic signals will impair this process (**Figure 15**).

The observation that the hydrophobic core of a signal sequence affects topology by influencing the kinetics of signal reorientation opened the possibility of exploring the environment of signal-anchors during topogenesis. Higy et al. tested the effect of positioning large hydrophobic amino acids, such as tryptophan, phenylalanine and tyrosine, as well as small hydrophobic residues, valine and alanine, throughout an oligoleucine sequence on the topogenesis process (Higy et al., 2005). Tryptophans showed a dramatic position dependence in a symmetric pattern. When positioned at either end of the hydrophobic segment, the signals inserted predominantly with an N_{exo}/C_{cyt} orientation. In contrast, tryptophans placed further inside the hydrophobic core of the signal favored C-insertion with the exception of positions 7 and 8 within the stretch of 16 leucines, where the N-insertion was again increased to ~50%. The symmetry of the position dependence for bulky aromatic amino acids reflects the symmetry of the lipid bilayer and the ease of their accomodation, thus suggesting contact of the signal with the lipid membrane during topogenesis. Tryptophans at the ends of the h-domain would interact with the interphase between the apolar core and the phospholipid headgroups region of the membrane. Positions WW5/6 and WW9/10, which yielded the highest C-insertion, appear to have the lowest affinity for the translocon-bound state. They reflect the situation, when tryptophans are located in the center of the acyl chains of phospholipids – such a highly ordered environment would not favor the

accomodation of large and stiff side chains. In contrast, tryptophans at the center of the bilayer (WW7/8) are more easily accepted. These results support the concept that the signal sequence is in contact with the lipid bilayer during topogenesis, previously formulated by photo-cross-linking of the signal to Sec61α and lipids (Martoglio et al., 1995; McCormick et al., 2003; Mothes et al., 1998).

In an experimental setting, signal sequences can be generated from homooligomers of apolar amino acids that are able to form a helix. They all share the feature of slowing down signal reorientation with its increasing length (Goder and Spiess, 2001). Based on the ability of different homooligomers to promote N-translocation, a hydrophobicity scale could be formed, where I>L>V~W>Y>F>M. Except for oligoalanine, which was not functional as a signal sequence, all homooligomers tested could efficiently target proteins in the co-translational pathway (Rosch et al., 2000). In addition, each amino acid can be ranked with respect to promoting N-terminal translocation when inserted in pairs into an oligoleucine helix: I>V>L~W>F>Y>C>M>A>T>S>G>N>Q>H>P. This ranking resembles the hydrophobicity scale, but in addition is similar to the scale of helix propensities in an apolar environment (Liu and Deber, 1998).

A couple of years ago, Hessa et al. presented the `biological` hydrophobicity scale of amino acids based on the apparent free energy of their membrane insertion (Hessa et al., 2005). A striking feature of this scale is that it predicts the very low cost of membrane solvation for charged residues. Based on molecular simulation results, Johansson et al. suggested that this phenomenom could be explained by the protein content of membranes which influences the solvation properties. Charged amino acids, such as arginine, are surrounded by hydration water in the translocation channel. Protein helices in the membrane significantly reduce the energetic cost of introducing solvation water into the bilayer environment. In contrast, leucine remains dehydrated inside the membrane. The high fluidity of the ER membrane due to its low cholesterol content could be an important factor to allow inserted helices to interact either with lipids or proteins depending on their sequence composition (Johansson and Lindahl, 2009).

7. Topogenesis of multi-spanning membrane proteins

Biogenesis of polytopic membrane proteins requires the coordination of several events: recognition and targeting of the nascent chain to the membrane localized translocation machinery,

integration and orientation of TMs coupled with folding of extramembrane domains, helical packing within the lipid bilayer, and formation of the tertiary structure (Dowhan and Bogdanov, 2009). Topology is determined by an interplay between the topogenic signals residing within the protein sequence, the interaction of the protein with the translocon, internal protein interactions, and protein-lipid interactions during final folding. During topogenesis, these factors act simultaneously or sequentially.

According to the simplest, the so called `linear insertion` model, the most N-terminal TM defines its own orientation as well as the orientations of all subsequent TMs, which insert into the bilayer in alternating orientations (Blobel, 1980). However, there is also evidence against this model, examples include *E. coli* lactose permease (LacY) and maltose transporter carrying deletions of individual TMs (Bibi et al., 1992; Ehrmann and Beckwith, 1991) or the tetracycline/H^+ antiporter with a perturbation of the orientation of its N-terminal segment (Guo et al., 1996) – in all cases the topology of downstream TMs was unaffected.

The hydrophobic properties of TMs allow them to passively partition into the lipid bilayer with flanking charged residues positioned near the aqueous-membrane interface. In many cases, the hydrophobicity of a TM is sufficient to drive the translocation of flanking charges, but it is not the sole determinant of membrane insertion. Asn- or Asp-mediated hydrogen bonding between neighbouring helices of polytopic proteins can enhance the membrane insertion efficiency of a marginally hydrophobic TM and could possibly form during Sec61 translocon-mediated insertion (Meindl-Beinker et al., 2006). Another topologically important feature is the folding state of the extramembrane hydrophilic domains, whose translocation occurs in the unfolded state. Their rapid folding in the cytoplasm may prevent export, whereas tight folding on the lumenal side of the membrane may ensure the location of those domains (Granseth et al., 2005). Glycosylation within the hydrophilic domain can also prevent its transport across the ER membrane and thus affect the topology (Goder et al., 1999).

Polytopic membrane proteins follow the `positive-inside rule`. Positively charged residues not only exert local control over the orientation of TMs, but can also affect the global topology of a protein. A recent study shows that the topology of an *E. coli* inner membrane protein EmrE with four or five transmembrane helices can be controlled by a single positive charge placed in different locations throughout the protein, including the very C-terminus. The C-terminal Lys can reverse the orientation of as many as five upstream TMs. Apparently, the topology of this protein remains undetermined until the last residue has been synthesized (Seppala et al., 2010). This raises important questions regarding the mechanism of insertion and assembly of polytopic proteins, in particular how much protein can the translocation pore accommodate. Negatively charged amino

acids appear to be topologically active only if they are present in high numbers, flank a marginally hydrophobic TM or lie within a window of six residues from a highly hydrophobic TM. Several negative residues are required to translocate a cytoplasmic domain containing even a single positive charge (Nilsson and von Heijne, 1990).

Another determinant of polytopic protein topology is the lipid composition of the membrane. The influence of lipids on membrane protein topogenesis was studied in *E. coli* strains in which the steady-state phosphatidylethanolamine (PE) content can be regulated. Analysis of LacY insertion in PE-lacking cells revealed a dramatic misassembly of the protein, with the N-terminal six-TM helical bundle (TM I-VI) and adjacent membrane domain completely inverted with respect to the plane of the bilayer and the C-terminal five-TM helical bundle (TM VIII-XII). Even more dramatic is that the aberrant topological organization of LacY in PE-deficient cells is nearly completely reversible post-assembly by induction of PE synthesis (Bogdanov et al., 2002; Bogdanov et al., 2008). The simplest interpretation is that introduction of PE into the membrane destabilizes the folded state of the protein, which results in reorientation of most of the N-terminal helical domains in order for the protein to assume its new minimum energy state. The fact that the topological organization of a membrane protein, once established, is dynamic in response to changes in the lipid environment suggests that lipids and proteins have co-evolved to follow a set of interdependent rules governing topogenesis.

8. Thesis goal

The aim of this thesis was to investigate the mechanism of insertion of membrane proteins. We wanted to compare the insertion process of internal versus N-terminal signal-anchors and to characterize the role of the N-terminal domain as a co-determinant of protein topology. We tested several parameters that influence topogenesis, such as the size of the N-domain, flanking charges, signal hydrophobicity and the C-terminal length. We studied the kinetics of signal inversion and looked for evidence for the `translation-stop phenomenon`, previously proposed for N-terminal signal-anchors (Goder and Spiess, 2003).

Furthermore, we wanted to investigate the insertion process of polypeptides with conflicting signal sequences. By altering the signals and their environment, we aimed at observing and characterizing the insertion process driven by the topogenic information of the second signal.

Finally, we studied the function of the plug and the constriction ring of the yeast Sec61p translocon in protein translocation. We mutated the ring residues to more hydrophilic, bulky or charged amino acids, alone or in combination with a point mutation in the plug or a full plug deletion, and analyzed the resulting phenotypes with respect to viability, translocation defects, translocon assembly and stability, as well as its ability to recognize and integrate transmembrane helices. We then transferred the analysis to the mammalian system and tested the effect of prolonged translation time on TM integration.

II Results

Part I: Insertion of polypeptides with internal signal-anchors

SUMMARY

In eukaryotic cells, hydrophobic signal sequences of newly synthesized secretory and membrane proteins target them to the Sec61 translocon in the endoplasmic reticulum membrane. Within the Sec61 channel, transmembrane segments of proteins achieve their proper orientation (topology) and are laterally released into the ER membrane. Cleaved signals of secretory and type I membrane proteins, as well as type II signal-anchors facilitate translocation of the C-terminus, yielding an N_{cyt}/C_{exo} orientation. The so-called `reverse signal-anchors` of type III proteins generate the opposite topology. Orientation of signal sequences in the ER membrane is determined by charged residues flanking the hydrophobic core of the signal, hydrophobicity of the signal, the size and folding properties of the N-terminal domain preceding the signal and, in come cases, the length of the C-terminus. Here, we compared the insertion process of N-terminal versus internal signal-anchors and determined the effect of the N-terminal hydrophilic domain on protein topogenesis. We show that insertion of these two types of signals occurs via different mechanisms. Transition from N-terminal to internal signals, achieved by extension of the N-domain with hydrophilic residues, is accompanied by loss of C-terminal length dependence and insensitivity to increased hydrophobicity of the signal. This indicates that, in contrast to N-terminal signals, signal-anchors localized internally cannot undergo reorientation within the pore and the initial orientation that promotes either N- or C-terminal translocation is favoured. Hydrophilic N-terminal domains contribute to this decision, sterically hindering N-translocation.

INTRODUCTION

In eukaryotic cells, hydrophobic signal sequences of newly synthesized secretory and membrane proteins mediate targeting to the Sec61 translocon in the endoplasmic reticulum membrane. Signal sequences may initiate translocation of the C-terminal sequence, as in the case of the cleaved signals of secretory and type I membrane proteins, and of the uncleaved signal-anchors of type II membrane proteins. Alternatively, type III signal-anchors (`reverse signal-

anchors`) translocate the N-teminal domain into the ER lumen, generating N_{exo}/C_{cyt} topology (exoplasmic or lumenal N-terminus and cytoplasmic C-terminus).

Several factors determine the orientation of cleaved signals and signal-anchor sequences in the ER membrane. Most prominent is the effect of charged amino acids flanking the apolar core of the signal sequence which cause retention of the more positive end on the cytosolic side – the so called `positive-inside rule` (Hartmann et al., 1989; Heijne, 1986). The second factor is the hydrophobicity of the core of the signal-sequence (Sakaguchi et al., 1992; Wahlberg and Spiess, 1997). Goder and Spiess (2003) described the contribution of these two determinants during the insertion process of N-terminal type II signal-anchors. Such signals initially engage with the translocon `head-on`, in N_{exo}/C_{cyt} topology, followed by their reorientation within the channel and translocation of the C-terminus. This process is driven by flanking charges, which position the signal according to the `positive-inside rule`, and it is slowed down by increased hydrophobicity of the signal core. Reorientation of N-terminal signals requires ongoing protein synthesis by the ribosome and it is stopped upon termination of translation (the so called `translation-stop phenomenon`) or by a signal-independent mechanism after ~50 s.

Another topogenic determinant is the size and the folding state of the N-terminal domain. In type III membrane proteins, it is translocated across the membrane after it has been synthesized in the cytoplasm. N-terminal domains that rapidly fold to stable structures are detrimental for type III topology (Denzer et al., 1995). In natural proteins of this type, the N-domain is often short, as in the cytochrome P450 family, but it may also be of considerable length, like in synaptotagmin I (53 residues) or neuregulin (>200 amino acids). Type II membrane proteins can be converted to type III and vice versa not only by mutagenesis of the flanking charges (Beltzer et al., 1991; Monier et al., 1988; Szczesna-Skorupa et al., 1988; Szczesna-Skorupa and Kemper, 1989), but also by alterations of the size of the N-terminal domain. Examples of such experiments include mutagenesis of the ASGP receptor H1, a type II membrane protein. The wild-type form consists of an N-terminal cytoplasmic domain of 40 amino acids, a membrane-spanning segment of 19 residues and a C-terminal exoplasmic domain of 231 amino acids with two sites for N-linked glycosylation (Spiess and Lodish, 1986). Mutation of the two N-terminal positive and the two C-terminal negative flanking charges to amino acids of the opposite charge (construct named A1-4) decreased C-terminal translocation from 100 to 50%. An additional truncation of the N-domain from 40 to 11 residues increased the fraction of type III polypeptides from 50 to >90%. Thus, the natural N-terminal domain of H1 sterically hinders N-translocation. Insertion of a glycosylation site into the N-domain of the mutant with inverted flanking charges (A1-4g) prevented N-terminal translocation and increased the type II population to 70% (Beltzer et al.,

1991). Folded structures in the N-domain strongly inhibited their translocation and thus type III insertion. The 234-amino acid sequence of dihydrofolate reductase (DHFR) fused to the N-terminus of A1g and A1-4g constructs completely blocked translocation of the N-terminal segment, and even a small folded peptide, a zinc finger domain of only 29 amino acids, effectively inhibited type III insertion (Denzer et al., 1995).

In this chapter, we further characterized the role of the N-terminal domain as a co-determinant of protein topology. Our model protein for topology studies, the H1 subunit of the ASGP receptor, contains a 40-aa N-terminal extension whose folding properties are unknown. Therefore, we tested the effect of N-terminal domains of different lengths that can be safely assumed not to fold in any defined structure. As a starting point for mutagenesis we used N-terminal signal-anchors whose insertion process had been previously described (Goder and Spiess, 2003). The transition from N-terminal to internal signals was achieved by a stepwise extension of the N-domain with a cluster of hydrophilic glycines and serines that provide good water solubility and conformational flexibility. Experiments were conducted *in vivo* in COS-1 cells transiently transfected with protein constructs.

We observed that flexible, hydrophilic N-terminal domains promote C-translocation. Our data indicate that in contrast to N-terminal signal-anchors, insertion of polypeptides with N-domains longer than 20 amino acids required opening of the translocon. Such internal signal-anchors were insensitive to increased signal hydrophobicity and showed no C-terminal length dependence, suggesting they orient themselves before contacting the hydrophobic core of the membrane.

MATERIALS AND METHODS

Cloning strategy

9GSMGPQL16 – 24GSMGPQL16

The starting construct MGPQL16 encoding the H1 subunit of the asialoglycoprotein receptor with a truncation of the N-terminal domain has been previously described (Wahlberg and Spiess, 1997). Glycine/serine hydrophilic N-terminal extensions were generated by annealing two pairs of complementary phosphorylated oligonucleotides, termed: BglII-1s with SalI-1a and XhoI-2s with BamHI-2a (all primers are listed in **Table I**). The annealed pairs were mixed, ligated and digested with BamHI, BglII, SalI and XhoI. The digestion products were resolved on an agarose gel, yielding bands of the following size: 45, 90, 135, 180, 225, 270, 315, 360, 405, 450, 495 bp. Odd-

numbered oligomers (45, 135, 225, 315, 405, 495 bp), comprising products of BglII-SalI or XhoI-BamHI cleavage, were selected and ligated into a pECE-vector (Ellis et al., 1986) cut with XhoI and BamHI. The starting codon along with a Hind III site was added with the HindMGSx-s primer.

34GSMGPQL16

Extension of the N-terminal domain from 24GS to 34 GS was generated by PCR in two consecutive steps, adding 5GS at a time (primers GS15+5.s and GS20+5.s, and ECEright). The amplified HindIII→EcoRI fragment was digested with HindIII and BamHI and the N-domain together with the signal-anchor were replaced.

27GS-4L13 – 87GS-4L13

Constructs with inverted flanking charges were essentially prepared the same way as GS with normal charge distribution, followed by the fusion of the C-terminal domain from the A1-4 polypeptide series (Beltzer et al., 1991).

57GS-4L22

Cloning proceeded in two steps. First, the BglII→HindIII fragment from pSA1-4gL22 plasmid containing the L22 signal-anchor (Beltzer et al., 1991) was fused to the HindIII→BglII segment from 57GS-4L13. In the next step, C-terminal tails were replaced by transfering the HindIII→BamHI sequence containing 57GS-4L22 to GS series with inverted charges.

14GSAGPQL16 – 34GSAGPQL16

Mutagenesis of the second methionine was performed by PCR of the HindIII→BamHI fragment, comprising the N-terminal domain together with the signal-anchor, using ECEleft sense primer and 24GSL16MtoA antisense primer.

All constructs were verified by sequencing with ECEleft, ECEleft2 or ECEright (whenever PCR was involved), or by a restriction digestion test.

Table I. PCR primers.

Primer name	Primer sequence	Primer type
24GS16MtoA-a	GCGGGATCCCAAGAGCAACAGCAGGAGCAACAAGAGGAGCAGCAGC AAAAGCAACAGCTGCGGTCCCGCAGATCCTCC	antisense
BamHI-2a	GATCCTCCTGACCCCGAACCAGAGCCCGATCCACTACCGCTCCCAC	antisense
BglII-1s	GATCTGGATCAGGGTCGGGTTCTGGCTCCGGAAGTGGTAGCGGG	sense
ECEleft	GAAGTAGTGAGGAGGC	sense
ECEleft2	CGGCCTCTGAGCTATTCCAG	sense

ECEright	CTACAAATGTGGTATGGC	antisense
GS15+5.S	CGCAAGCTTGCGATGAGCGGAAGCGGGTCGGGAACCTCGAGTGGG	sense
GS20+5.s	CGCAAGCTTACCATGGGTAGCGGATCAGGGAGCGGAAGCGGGTCGGG	sense
HindMGSx-s	GCGAAGCTTACCATGGGAACCTCGAGTGGGA	sense
SalI-1a	TCGACCCGCTACCACTTCCGGAGCCAGAACCCGACCCTGATCCA	antisense
XhoI-2s	TCGAGTGGGAGCGGTAGTGGATCGGGCTCTGGTTCGGGGTCAGGAG	sense

Cell culture

COS-1 cells were cultivated in Dulbecco's modified Eagle's medium (DMEM; Sigma) supplemented with 10% fetal calf serum (FCS), 100 units/ml penicillin, 100 µg/ml streptomycin, 2 mM L-glutamine, in a humidified incubator containing 7.5% CO_2 at 37°C.

Transient transfections

For transient transfections, cells from a confluent culture were split 1:10 into 6-well plates or 60-mm dishes and transfected the next day with one of the following reagents: Lipofectine (Life Technologies), FuGENE HD (Roche) or polyethylenimine (PEI; Sigma-Aldrich). FuGENE HD, although highly efficient, proved to be toxic to COS-1 cells and was replaced with PEI. Cells were processed 2 days after transfection.

Metabolic labeling with ^{35}S-methionine

Transfected cells were incubated for 30 min in starvation medium (DMEM without methionine and cysteine, containing 2 mM L-glutamine; Sigma). Cells were labeled for 40 min with 100 µCi/ml [^{35}S]protein labeling mix consisting of 77% methionine and 23% cysteine (PerkinElmer), then transfered to 4°C and washed twice with cold phosphate-buffered saline (PBS).

Immunoprecipitation

Non-integrated proteins were removed by extraction with 0.1% saponin (in PBS, with 2 mM phenylmethylsulfonylfluoride [PMSF], 1x protease inhibitor coctail [PIC]: 5 mg/ml benzamidine, 1 mg/ml pepstatin A, 1 mg/ml leupeptin, 1 mg/ml antipain, 1 mg/ml chymostatin, in 40% DMSO/60% ethanol) for 30 min, followed by a PBS wash. Cells were lysed in 500 µl of lysis buffer (PBS, 1% TritonX-100, 0.5% deoxycholate, 2 mM PMSF, 1x PIC) for 1 h, then scraped, vortexed and incubated for 30 min on ice. Lysates were cleared by centrifugation and proteins were immunoprecipitated overnight by addition of 500 µl immuno-mix (lysis buffer, 1 mg/ml bovine serum albumin [BSA], 1mM PMSF, 1x PIC) containing 1 µl rabbit anti-serum raised

against a peptide corresponding to residues 277-287 at the C-terminus of H1 (anti-H1C). The immune complexes were pulled down by incubation with 10 µl/sample protein A-Sepharose (Zymed) for 1 h. Samples were washed 4 times with immuno-wash (lysis buffer, 1 mg/ml BSA, 1 mM PMSF) and 2 times with PBS (with 1 mM PMSF).

EndoH treatment

At the last PBS wash samples were split in half for deglycosylation by endoglycosidase H. All samples were boiled in endoH buffer (50 mM sodium citrate, 1% SDS; in PBS), 1 µl of the enzyme (Roche) was added and samples were incubated for 3 h at 37°C.

SDS-PAGE and autoradiography

Reduced samples were resolved on 12.5% - 17.5% acrylamide gels and quantified using a phosphorimager (Molecular Dynamics Inc.). Protein orientation in the ER membrane was determined by calculating the ratio of glycosylated proteins divided by total protein within a lane. This value did not depend on the signal intensity caused by different transfection or protein pull-down efficiency.

RESULTS

Transition from N-terminal to internal signal-anchors is accompanied by loss of C-terminal length dependence

In order to study the insertion process of internal signal-anchors, we tested the effect of non-folding, hydrophilic N-terminal domains on topogenesis. For this purpose, we selected the derivative of the asialoglycoprotein receptor subunit H1 with truncated N-terminus, H1ΔLeu16, and extended its N-terminal part. This protein contains an N-terminal signal-anchor consisting of Leu16 preceded only by a MGPQ sequence and is therefore called MGPQL16 (**Figure 1A**). The C-terminal tails consist of 75 to 460 residues and contain two potential sites for N-linked glycosylation. The constructs were expressed in COS-1 cells and proteins were metabolically labeled with [^{35}S]-methionine. Any soluble proteins were removed by saponin extraction and the remaining, membrane-integrated proteins were immunoprecipitated with an antibody directed against the C-terminus (**Figure 1B**). The MGPQL16 series showed C-terminal length dependence (Goder and Spiess (2003); **Figure 1C**, blue line). The transition from N-terminal to internal signal-anchors was achieved by a stepwise addition of 5 residues made of glycines and serines to

the starting construct MGPQL16. The resulting constructs were named 9GSMGPQL16[#], 14GSMGPQL16[#], 24GSMGPQL16[#] and 34GSMGPQL16[#], where the leading number indicates the total number of residues in the N-domain and [#] indicates various lengths of the C-terminal tail (**Figure 1A**).

Upon expression, we received several different protein forms, representing different glycosylation states of the proteins, as confirmed by deglycosylation with endoH (shown for 34GSMGPQL16[#] in **Figure 1B**). Addition of an N-linked glycan results in an increase in the apparent molecular weight of approximately 3 kDa on the SDS gel. Thus, the low-molecular weight band (0) represents the non-glycosylated product, generated when the C-terminus of the protein is located on the cytosolic side of the ER membrane. The two upper bands (1 and 2) correspond to the once and twice glycosylated polypeptides whose C-terminus was translocated into the ER lumen. One of the glycosylation sites, presumably the one located close to the membrane, is not always modified. Upon digestion with endoH, the bands representing the singly and doubly glycosylated products collapsed into a single band. It migrated slightly more slowly than the non-glycosylated form, which is explained by one N-acetylglucosamine residue remaining after endoH cleavage. Topology of proteins in the ER membrane was assessed by calculating the fraction of C-terminal translocation, e.g., the sum of singly and doubly glycosylated products divided by total protein in each lane of the gel. C-translocation was plotted versus the size of the C-terminal domain, which is the fragment from the end of the signal-anchor to the stop codon (**Figure 1C**).

The topologies of N-terminal signal-anchors of MGPQL16[#] show a C-terminal length dependence – this relationship has been previously interpreted as reorientation of the signal within the translocon, terminated upon chain completion or independently of the substrate after 50 s (Goder and Spiess, 2003). Upon extension of the N-domain to a total of 9 or 14 residues (9GSMGPQL16[#], 14GSMGPQL16[#]), C-translocation was slightly increased, but the prominent dependence on the C-terminal length was largely retained (**Figure 1C**, red and green lines, respectively). For constructs with a longer N-terminus (24GSMGPQ[#]), the dependence on C-terminal length was almost completely lost, while retaining mixed topologies (**Figure 1C**, purple line). Finally, the topologies of proteins with the longest N-terminal domain tested, 34GSMGPQ[#], were quite independent of the size of the C-terminal tail (**Figure 1C**, orange line). Taken together, the results indicate that the transition between N-terminal and internal signal anchors occurs for proteins with N-domain of ~20 residues and is manifested by loss of C-terminal length dependence.

A

signal-anchor

MGPQ LLLLLLLLLLLLLLLL GSQNSQLQEELRGL...   MGPQL16=H1ΔQLeu16

MSGGSMGPQ LLLLLLLLLLLLLLLL GSQNSQLQEELRGL...   9GSMGPQL16

MGSGSGSGGSMGPQ LLLLLLLLLLLLLLLL GSQNSQLQEELRGL...   14GSMGPQL16

MGTSSGSGSGSGSGSGSGGSMGPQ LLLLLLLLLLLLLLLL GSQNSQLQEELRGL...   24GSMGPQL16

MGSGSGSGSGSGSGTSSGSGSGSGSGSGSGGSMGPQ LLLLLLLLLLLLLLLL GSQNSQLQEELRGL...   34GSMGPQL16

B
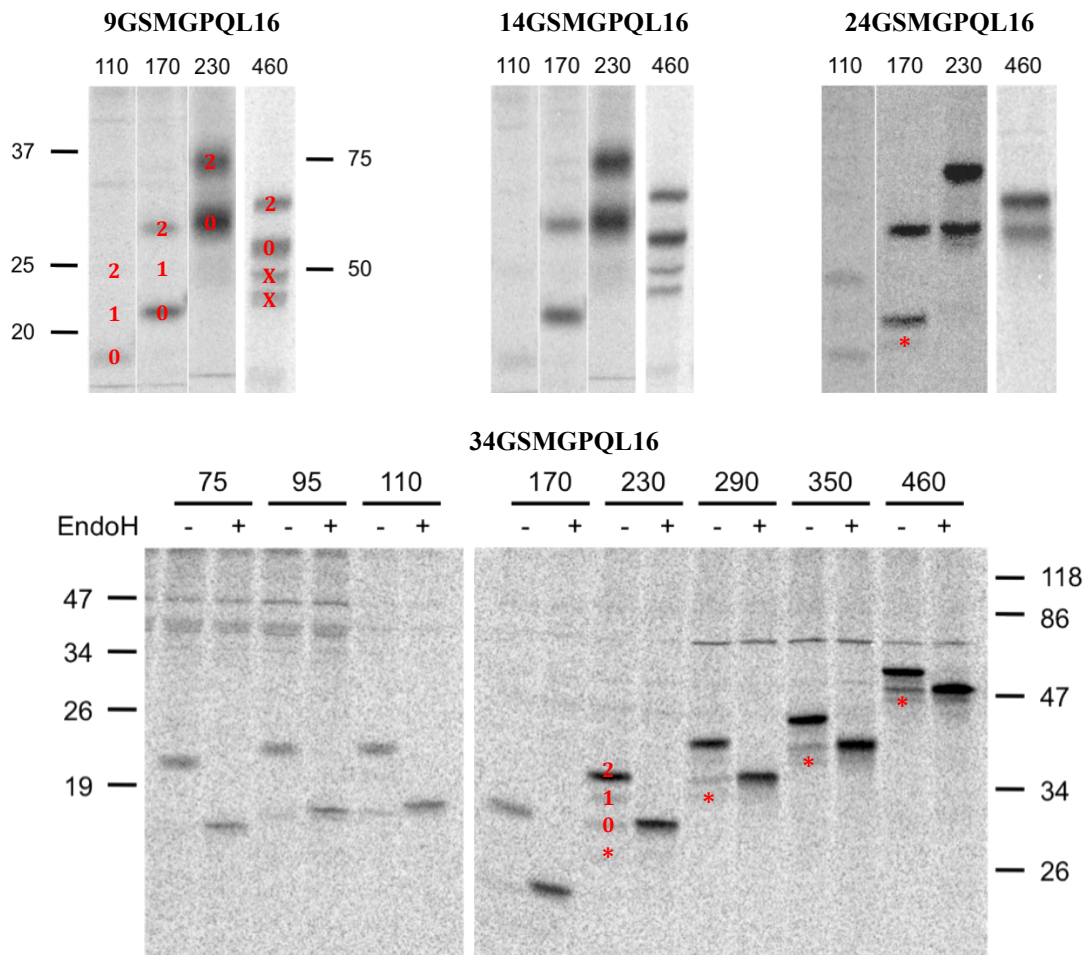
C
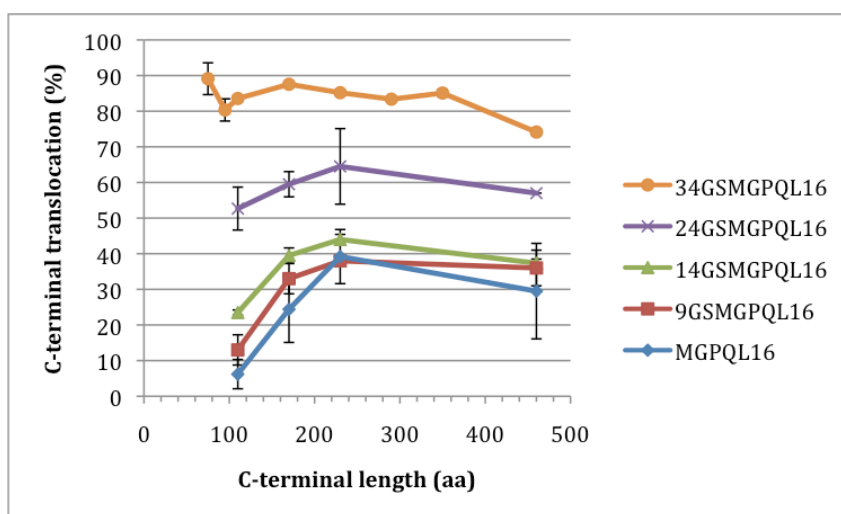
Figure 1. Topology analysis of N-terminal and internal signal-anchors. The transition was achieved by a stepwise extension of the N-domain by clusters of hydrophilic glycines and serines (A). The proteins were expressed in COS-1 cells, pulse-labeled with [^{35}S]methionine for 40 min, immunoprecipitated, separated by gel electrophoresis, and visualized by autoradiography. Prior to immunoprecipitation, soluble proteins were removed by saponin extraction. Expression of the constructs yielded several different protein forms: 0-unglycosylated, 1-singly glycosylated, 2-doubly glycosylated, ∗-truncated form resulting from internal translation initiation, x-background bands. Glycosylation status was confirmed by endoH digestion (B). Protein orientation in the ER membrane was determined by calculating the ratio of C-terminal translocation, i.e., the ratio of glycosylated products divided by total protein within a gel lane. C-translocation is plotted against the size of the C-terminal domain (sequence from the end of the signal-anchor until the stop codon). Topology analysis revealed that transition from N-terminal to internal signals occurs for N-domains of ~20 amino acids and is accompanied by loss of C-terminal length dependence. Long, hydrophilic N-domains sterically hinder N-insertion (C). The average and standard deviations of two independent experiments are shown. The position of molecular weight markers (in kDa) is indicated.

Non-folding N-domains increase C-translocation with increasing length

Increasing the size of the N-domain by addition of hydrophilic amino acids resulted in an increased fraction of proteins with a translocated C-terminus (**Figure 1C**). Thus, even highly flexible, folding-incompetent sequences can prevent N-translocation. In addition to the steric hindrance effect of the N-domain, the insertion process of our model proteins might be potentially influenced by an opposing effect of moving the positive charge of the α-amino group of the protein further away from the signal (thus weakening the `positive-inside rule`). This charge effect decreased with increasing distance from the signal core.

Certain N-domains cause internal initiation of protein synthesis

Extension of the N-terminal domain had another, unexpected consequence: expression of the constructs with 24 and 34 residues in the N-terminus yielded an additional, low-molecular weight form, insensitive to endoH deglycosylation (marked by asterisks in **Figure 1B**). In some experiments the band was prominent and could potentially interfere with quantitation. Those proteins were not the result of degradation of the N-terminal extension, since they did not disappear upon shortening of the time of radiolabeling or the time of lysis and using increased amounts of protease inhibitors (data not shown). Cleavage within the C-domain could be excluded, as proteins were immunoprecipitated with an antibody directed against the very end of the polypeptide. Another possibility was initiation of protein synthesis at the second methionine, positioned next to the signal-anchor, which would produce proteins with a truncated N-terminus. To test it, we mutated the internal methione to alanine, a residue of a similar hydrophobicity to methionine, according to the hydrophobicity scale of Rösch et al. (2000), in order to minimize a potential effect on topogenesis (**Figure 2A**). Upon expression of these constructs, named

24GSAGPQL16[#] and 34GSAGPQL16[#], the short, non-glycosylated form was eliminated, confirming our hypothesis (**Figure 2B**). However, the Met→Ala exchange caused a general shift in topology towards C-terminal translocation (**Figure 2C**, purple and orange lines), suggesting that the alanine introduced into the N-terminal flank of signal-anchor had a topogenic contribution. In order to test whether we can still observe the transition between N-terminal and internal signals and the associated loss of C-terminal length dependence, we modified the next series of polypeptides, containing 14 residues in the N-domain. Glycosylation analysis revealed that these proteins had a similar ratio of C- vs. N-insertion as 24GSMGPQL16[#] (**Figure 2C, green line**) and when compared with MGPQL16[#] and 9GSMGPQL16[#] (**Figure 1C**, blue and red lines, respectively), they showed intermediate topologies.
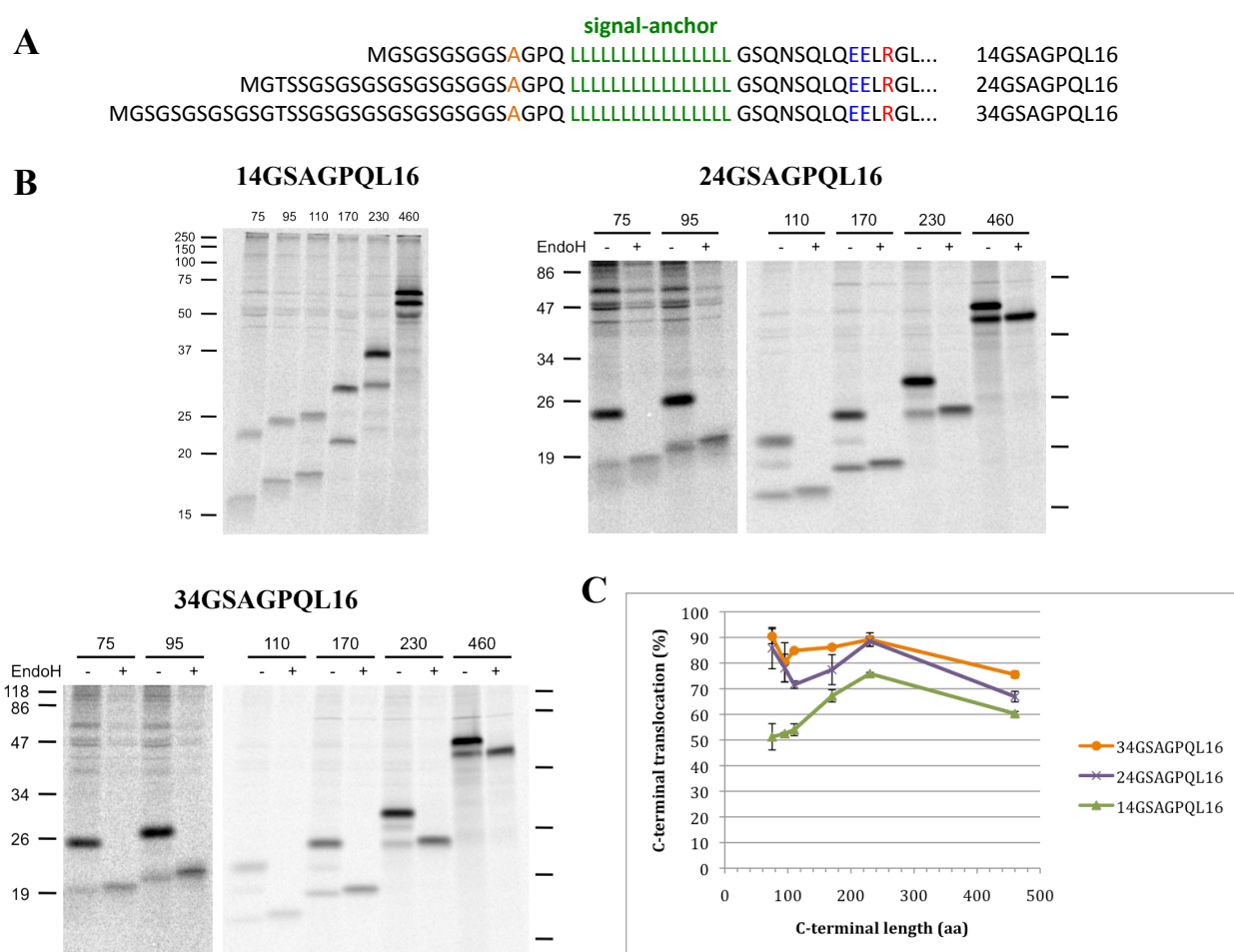
Figure 2. Internal initiation of synthesis of polypeptides with internal signal-anchors. The second methionine located in the N-terminal flank of the signal was replaced by alanine (A) and the resulting constructs were expressed in COS-1 cells and processed as described in Figure 1. Upon mutagenesis, the truncated forms were no longer detected on the gels (B), but the ratio of protein insertion shifted towards C-translocation (C). The average and standard deviations of two to three independent experiments are shown. The position of molecular weight markers (in kDa) is indicated.

Based on these results, we decided not to further modify the remaining constructs and to use the series of polypeptides containing the second methionine. Our quantitation of protein insertion ratio was only minimally affected by the truncated form. Namely, its glycosylated version contaminated the non-glycosylated band of our test constructs with a hydropilic N-domain. However, since the internal translation initiation was not prominent and did not occur in every experiment, and the exchange of methionine to alanine did not change the overall insertion pattern, we considered it negligible.

Internal signal-anchors are insensitive to increased hydrophobicity of the signal

Transition from N-terminal to internal signal-anchors was accompanied by the loss of C-terminal length dependence, which is considered a hallmark of protein reorientation within the translocon. Lack of C-terminal length dependence of signal-anchors with N-domains of >20 residues could indicate that they did not undergo reorientation or that reorientation occurred very rapidly and was completed by the time the shortest constructs were finished. In order to invert itself, the signal needs to dissociate from the hydrophobic environment of the channel or the membrane. The process is slowed down by increased hydrophobicity of the signal core (MGPQL22[#] in **Figure 3C**, blue line) (Goder and Spiess, 2003). In case of signal-anchors with N-domains >20 aa, increased hydrophobicity of the signal should not influence the insertion process. To test this, we exchanged the apolar core of the signal consisting of 16 leucines by a stretch of 22 leucines (**Figure 3A**) and expressed in COS-1 cells (**Figure 3B**). Glycosylation analysis revealed that proteins with 4, 9 and 14 residues in the N-domain were sensitive to increased signal hydrophobicity, manifested by a dramatic decrease of C-terminal translocation (**Figure 3C**, blue, red and green lines, respectively). In contrast, long hydrophobic signals had no influence on the insertion process of the constructs with 24- or 34-residue N-domains – the ratio of C vs. N-terminal translocation remained unchanged (**Figure 3C**, purple and orange lines). This confirms the notion that our model proteins with N-domains longer than 20 amino acids, such as 24GSMGPQL16[#] and 34GSMGPQL16[#], can be considered internal signal-anchors and they do not undergo inversion upon insertion into the pore.

A

signal-anchor

MGPQ LLLLLLLLLLLLLLLLLLLLLL GSQNSQLQEELRGL...   MGPQL22

MSGGSMGPQ LLLLLLLLLLLLLLLLLLLLLL GSQNSQLQEELRGL...   9GSMGPQL22

MGSGSGSGGSMGPQ LLLLLLLLLLLLLLLLLLLLLL GSQNSQLQEELRGL...   14GSMGPQL22

MGTSSGSGSGSGSGSGSGGSMGPQ LLLLLLLLLLLLLLLLLLLLLL GSQNSQLQEELRGL...   24GSMGPQL22

MGSGSGSGSGSGTSSGSGSGSGSGSGSGSGGSMGPQ LLLLLLLLLLLLLLLLLLLLLL GSQNSQLQEELRGL...   34GSMGPQL22

B

C

Figure 3. Effect of increased hydrophobicity of the signal on topogenesis of N-terminal and internal signal-anchors. The apolar core of the signal consisting of 16 leucines was replaced by a Leu_{22} sequence (A). The resulting constructs were expressed in COS-1 cells and processed as described in Figure 1 (B). Topology analysis revealed that in contrast to N-terminal signal-anchors, internal signals are insensitive to increased signal hydrophobicity. For comparison, panel C of Figure 1 was reproduced on the right (C). The average and standard deviations of two to three independent experiments are shown. The position of molecular weight markers (in kDa) is indicated.

Insertion of N-terminal and internal signal-anchors is controlled by flanking charges

Inversion of N-terminal signal-anchors is driven by electrostatic interactions acting on flanking charges. We next tested the sensitivity of internal signals to changes of the charged residues in the N-terminal flank. For this purpose, the glutamine immediately preceding the apolar core in the MGPQL16 series was replaced by arginine, followed by extension of the N-domain. The resulting constructs were named MGPRL16[#], 9GSMGPRL16[#], 14GSMGPRL16[#] and 24GSMGPRL16[#] (**Figure 4A**). Proteins were expressed in COS-1 cells and subjected to glycosylation analysis (**Figure 4B**).

Figure 4. Comparison of the effects of flanking charges and signal hydrophobicity on the insertion process of N-terminal and internal signal-anchors. The neutral glutamine in the N-terminal flank of the signal-anchor was mutated to positively charged arginine, alone (A) or in combination with increased hydrophobicity of the signal (Leu$_{16}$ to Leu$_{22}$; D). The constructs were expressed in COS-1 cells and subjected to glycosylation analysis (B and E, respectively). Quantitation of C-terminal translocation showed that both types of signals are sensitive to flanking charges (C; for comparison, panel C of Figure 1 was reproduced on the right), but only N-terminal signals are affected by increased hydrophobicity of the signal-anchor (F). The average and standard deviations of two to three independent experiments are shown.

The introduced N-terminal charge promoted orienting of the signal according to the `positive-inside rule` and increased C-translocation. The most sensitive constructs, MGPRL16[#], showed almost complete C-translocation (Goder and Spiess 2003; **Figure 4C,** blue line). Polypeptides with N-domains of 9, 14 and even 24 residues also reached a very high level of C-translocation, between ~85 and 100% (**Figure 4C**, red, green and purple lines, respectively). The constructs with the longest N-terminus, consisting of 34 amino acids, were excluded in this experiment, since they had already had a high ratio of C-terminal translocation prior to mutagenesis.

The next step was to study the combined effect of mutagenesis of flanking charges and signal hydrophobicity. Signal-anchors of proteins with a positive N-terminal flank were replaced by a Leu_{22} sequence (**Figure 4D**) and metabolically labeled in COS-1 cells (**Figure 4E**). Again, constructs with short N-domains (4/9/14 aa) were affected inversely proportional to the size of the N-domain (**Figure 4F,** blue, red and green lines, respectively), whereas constructs with 24 residues in the N-domain were insensitive (**Figure 4F**, purple line). The rate of inversion of polypeptides with short N-domains of 4-14 residues decreased to a level that allowed us to observe C-terminal length dependence.

Taken together, the data suggest that constructs with short N-domains (<20 amino acids) behave differently than those with longer N-termini.

Internal signal-anchors with inverted flanking charges show discontinuous C-terminal length dependence

Increase of the size of the N-domain was correlated with increased rate of C-insertion, which reached ~85% for constructs with 34 residues in the N-terminus and, thus, limited further extensions (**Figure 1C**, orange line). Therefore, in order to study proteins with longer non-folding N-terminal domains, we moved to proteins with inverted flanking charges which disfavor type II insertion. We prepared constructs with total N-terminal domain lengths of 27, 57, or 87 hydrophilic residues, signal-anchors composed of 13 leucines, and C-terminal tails containing 75-460 amino acids. The signals were flanked by two negative charges in the N-terminal part and two positive ones in the C-terminus. The constructs were called 27GS-4L13, 57GS-4L13 and 87GS-4L13 (**Figure 5A**). Proteins were labeled *in vivo* with [^{35}S]-methionine in COS-1 cells (**Figure 5B**).

A

signal-anchor   **27GS-4L13**

MGTSSGSGSGSGSGSGSGSGGGSDLCSGPD LLLLLLLLLLLLL GSQNSQLQKK...

57GS-4L13

MGTSSGSGSGSGSGSGSGGGSGSGSGSGSGSGSGSSGSGSGSGSGSGSGGGSDLCSGPD LLLLLLLLLLLLL GSQNSQLQKK...

87GS-4L13

MGTSSGSGSGSGSGSGSGGGSGSGSGSGSGSGSGSSGSGSGSGSGSGSGGGSGSGSGSGSGSGSGSSGSGSGSGSGSGSGGGSDLCSGPD LLLLLLLLLLLLL GSQNSQLQKK...

B

C

D

E

signal-anchor

MGTSSGSGSGSGSGSGSGGGSGSGSGSGSGSGSGSSGSGSGSGSGSGSGGGSDLCSGPD LLLLLLLLLLLLLLLLLLLLLL GSQNSQLQKK...   **57GS-4L22**

F   **57GS-4L22**

G

Figure 5. Insertion process of internal signal-anchors with inverted flanking charges. Constructs with long N-terminal domains and signal-anchors composed of 13 leucines, flanked by two negative charges in the N-terminal part and two positive ones in the C-terminus were studied (A). In addition, the hydrophobicity of the signal was increased to 22 leucines (E). Glycosylation analysis in COS-1 cells (B and F, respectively) revealed C-terminal length dependence that was lost for the constructs with the longest N-domain, 87GS. Hydrophilic amino acids in the N-domain increase C-insertion by ~0.6-1% per residue (C). Length dependence is not coupled to sensitivity to changes in signal hydrophobicity (G). The average and standard deviations of three independent experiments are shown. The position of molecular weight markers (in kDa) is indicated.

Glycosylation analysis revealed that, unexpectedly, the first two series of polypeptides showed discontinuous C-terminal length dependence (**Figure 5C,** blue and red lines) that was largely lost for the constructs with the longest N-domain, 87GS (**Figure 5C,** green line). The increasing length of the N-domain resulted in a quite uniform increase of C-translocation for all lengths of C-domains. Plotting C-terminal translocation vs. the length of the N-terminus revealed the contribution of each additional residue in the N-domain to C-translocation. The increase was ~1% for polypeptides with C-terminal tail of 75-110 amino acids, ~0.8% for [170], ~0.6% for [230], and ~0.7% for [460], (**Figure 5D**).

C-terminal length dependence of internal signal-anchors with type III charge distribution is not associated with signal reorientation

The discovery that topogenesis of internal signal-anchors with inverted flanking charges depends on the size of the C-terminus was surprising. In addition, its discontinuous nature was not compatible with a (continuous) reorientation process. To investigate it further, we selected the 57GS-4L13 series and increased the hydrophobicity of the signal-anchor to Leu_{22} (**Figure 5E**). This should have slowed down or even blocked protein reorientation. The resulting series of constructs, 57GS-4L22[#], was expressed in COS-1 cells (**Figure 5F**). Quantitation of the ratio of C-terminal translocation revealed that increased signal hydrophobicity had no influence on the insertion process. (**Figure 5G**). Therefore, the pattern of C-terminal length dependence does not reflect a continuous reorientation process, but the properties of the individual constructs. The underlying cause is unclear.

DISCUSSION

The process of integration of single-spanning membrane proteins begins when a hydrophobic signal-anchor emerges from the translating ribosome and targets the ribosome-nascent chain complex to the Sec61 translocon in the ER membrane. Within the translocation channel, proteins achieve their topology and are released laterally into the bilayer. Topogenic information encoded in the protein sequence influences the decision whether the C-terminus is retained in the cytosol or translocated across the membrane (Rapoport et al., 2004). Here, we have explored the topogenesis of single-spanning membrane proteins. In particular, we compared the insertion process of N-terminal versus internal signal-anchors and tested effects of several

topogenic parameters, such as the size of the N-terminal domain, hydrophobicity of the signal-anchor, length of the C-terminal tail and flanking charges. The results allowed us to formulate models for orientation of these two types of signal-anchors in the Sec61 channel (**Figure 6**).

The transition from N-terminal to internal signal-anchors was achieved by extending the N-domain from 4 residues to 9, 14, 24 and 34 with hydrophilic, folding-incompetent clusters of glycines and serines. Experiments revealed that proteins with 4, 9 and 14 amino acids in the N-terminus have a different insertion behaviour in response to changes in topogenic determinants compared with proteins with 24 and 34 residues that belonged to internal signals. The characteristic features of N-terminal signal-anchors included continuous C-terminal length dependence and sensitivity to signal hydrophobicity. In contrast, insertion of internal signal-anchors did not depend on the size of the C-terminus, they were insensitive to increased signal hydrophobicity, but similar to N-terminal signals, they reacted to changes of the flanking charges.

In the context of the structure of the SecY/Sec61 translocon, it has been proposed that N-terminal signal-anchors can enter the channel head first (**Figure 6A**, arrow a), reversibly intercalate in the exit site of the translocon and contact lipid (equilibrium arrows b) without fully displacing the plug and opening the pore and while being tethered to the ribosome. The more hydrophobic the TM, the higher its affinity to the membrane and the lower the rate of return into the translocon, where the flanking charges induce reorientation (inversion arrow d). As the polypeptide chain is elongated, the C-terminal sequence accumulating inside the channel will trigger pore opening and C-translocation (arrow e). Termination of translation will block the equilibration of proteins oriented in the translocon in N_{exo}/C_{cyt} topology and trigger their lateral release in this orientation (arrow c), while proteins that have inverted will be released into the membrane in a final N_{cyt}/C_{exo} topology (arrow f). The time-dependence of signal inversion and translocon gating, together with interruption of this process upon chain completion (`translation-stop phenomenon`) explain the observed C-terminal length dependence of topology of N-terminal signals (Goder and Spiess, 2003).

Orientations of internal signal-anchors, with N-domain longer than 20 amino acids, and a positive charge difference between the N- and the C-terminal flank of the signal proceeded independently of the size of the C-terminus. Lack of C-terminal length dependence suggests that the signal does not undergo reorientation inside the pore. Another argument against the inversion process is insensitivity to increased signal hydrophobicity, which inhibits the signal's ability to dissociate from the apolar binding site or the membrane prior to inversion. The simplest interpretation of this behaviour is that internal signal-anchors engage with the translocon depending on the steric effect of the N-domain (**Figure 6B,** arrows a and e) and in addition, they

arrange their orientation according to the flanking charges (`positive-inside rule`, arrow e and/or d). Insertion of either end requires full translocon opening (b and f), a process that is irreversible. Termination of translation thus does not affect the final outcome of topogenesis for internal signal-anchors (c or g).
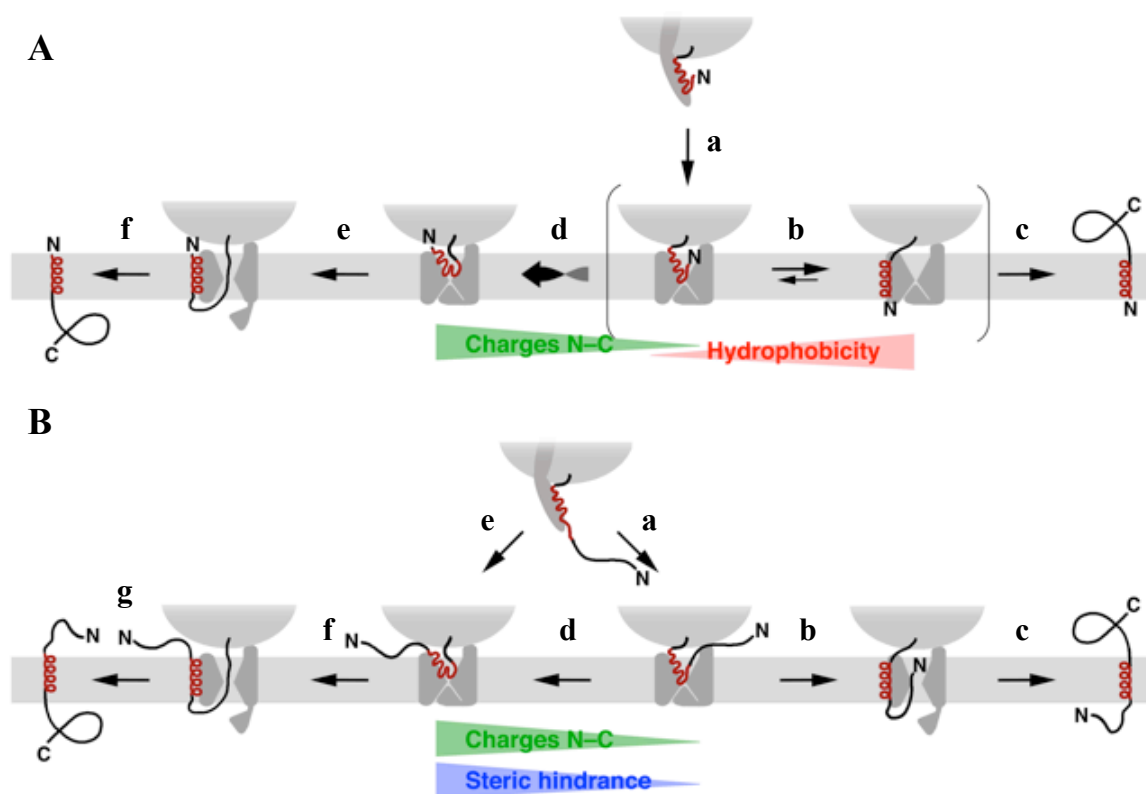
Figure 6. Models for topogenesis of N-terminal and internal signal-anchors with a positive charge difference. (A) N-terminal signal anchors can reversibly intercalate in the exit site of the translocon (area in brackets in panel A) and enter the lipid bilayer (c) in an N_{exo}/C_{cyt} orientation without inducing full pore opening (a-c). Type II signal-anchors initially insert `head-on` (a) and can invert their orientation over time (d). The process requires an ongoing protein synthesis, it is driven by electrostatic forces acting on flanking charges and is inhibited by increased hydrophobicity of the signal. The C-terminal sequence induces pore opening and C-translocation (e). (B) Polypeptides with internal signal-anchors and hydrophilic, folding incompetent N-domains engage with the translocon in an orientation that promotes either N- (a) or C-insertion (e). Flanking charges contribute to orienting the signal according to the `positive-inside rule` (e or/and d), while the long N-terminal domain acts as a steric hindrance, preventing N-insertion (e-g).

Lack of C-terminal length dependence was observed for internal signal-anchors with type II charge distribution (positive N-terminal flank). In the case of proteins with inverted flanking charges, the situation is more complicated. Topogenesis of polypeptides with N-domains consisting of 27 and 57GS depended on the size of the C-terminus in a discontinuous manner. C-terminal length dependence is a hallmark of signal inversion within the channel. However, the

process is electrostatic in nature and requires a positive charge difference between the N- and C-terminal flank of the signal-anchor. If internal signal-anchors were able to invert their orientation despite infavourable charge distribution, the process should have been slowed down by increased signal hydrophobicity. Upon mutagenesis of the signal-anchor, we observed no sensitivity to increased hydrophobicity. However, it was the only criterion of signal inversion that we tested and it needs to be further investigated.

Perhaps the observed C-terminal length dependence of internal signal-anchors with inverted flanking charges is the result of competition between conflicting topogenic factors localized within the protein sequence. One is the steric hindrance effect posed by the N-domain, where each additional residue contributed to ~1% increase in C-translocation. In contrast, negatively charged amino acids in the N-terminal flank and positive charges in the C-terminal flank of the signal core promote N_{exo}/C_{cyt} orientation. During membrane insertion, these topogenic factors act simultaneously. In case of polypeptides with long C-domains, the increased time of translation shifted the ratio in favor of C-translocation. However, it does not explain the discontinuous pattern of C-terminal length dependence.

Part II: Insertion of polypeptides with conflicting signals

SUMMARY

Multispanning membrane proteins are cotranslationally targeted to the ER membrane by the first hydrophobic signal sequence. All subsequent transmembrane segments (TMs) are inserted into the bilayer in alternating orientations, according to the `linear-inserion model`, or they can assume a unique topology defined by topogenic determinants localized within the protein sequence. These include flanking charges surrounding a TM, its overall hydrophobicity, as well as the folding state of hydrophilic spacer sequences. Here, we have analyzed the insertion process of a series of chimeric proteins composed of the cleavable hemagglutinin (HA) signal, followed by a signal-anchor of the H1 subunit of the asialoglycoprotein receptor. These signal sequences are carrying conflicting topogenic information (type I insertion induced by the HA signal vs. type II topology of the H1 signal-anchor). Previous experiments revealed that only when the two signals were sufficiently separated from each other, by a linker of ≥80 residues, the insertion proceeded according to the `linear insertion model`. Here we show that proteins with wild-type HA and H1 signals, connected by a 40-amino acid linker compete for the preferred orientation in the translocon, manifested by a rapid inversion of a fraction of the polypeptides, triggered by the signal-anchor. The process could be slowed down by increasing the hydrophobicity of the H1 signal or manipulating its flanking charges in a way that inhibits positioning of the signal according to the `positive-inside rule`. Under such conditions, topogenesis is interrupted upon termination of translation, like previously observed for N-terminal signal anchors. However, in contrast to single-spanning membrane proteins, the topogenesis window was not a constant of the translocation machinery, but rather appeared to be substrate-specific.

INTRODUCTION

Biogenesis of multispanning membrane proteins in both prokaryotes and eukaryotes requires their co-translational, SRP-dependent targeting and orientation within the membrane. Targeting is generally mediated by the first hydrophobic signal, which is either a cleaved signal sequence or a signal-anchor comprising the first TM of the protein. The topology of polytopic membrane proteins is determined by topogenic sequences in the protein, protein-translocon

interactions, interactions between individual protein domains during folding and between the protein and the lipid environment of the membrane (Dowhan and Bogdanov, 2009). According to the simplest model, the `linear-insertion model`, the first signal defines its own orientation and directs the insertion of all subsequent TMs, which will be incorporated into the membrane in alternating orientations (Blobel, 1980). This could, indeed, be shown in an experimental setting with chimeric proteins composed of two to four TMs separated by linker sequences of ~50-200 residues. The results demonstrated that signal-anchors acted as stop-transfer sequences depending on their position relative to the preceding hydrophobic segments (Lipp et al., 1989; Wessels and Spiess, 1988).

In natural proteins, topogenic information is contained also in internal transmembrane domains, which follow the `positive-inside rule` (Heijne, 1986). The orientation of membrane proteins can be reversed by addition or removal of even a single positive charge or by introduction of negatively charged residues situated close to the ends of TMs (Gafvelin and von Heijne, 1994; Nilsson and von Heijne, 1990; Seppala et al., 2010).

Transmembrane segments are separated by hydrophilic domains, which are either retained in the cytoplasm or translocated into the ER lumen in an unfolded state during protein insertion. Their proper location on the cytoplasmic or exoplasmic side may be assured by rapid folding in the cytosol or glycosylation in the ER lumen, respectively. The extramembrane hydrophilic domains in natural polytopic proteins are often much shorter than those used in the studies supporting the `linear insertion model`. Goder et al. (1999) tested how the topology of two-signal proteins is influenced by the length of the spacer separating the two signal sequences. They constructed a series of chimeric polypeptides containing an N-terminal cleaved signal sequence of influenza virus hemagglutinin (HA, in the constructs called shortly H) and an internal type II signal-anchor of the H1 subunit of the ASGPR (called A), separated by a linker sequence. The spacer was successively truncated to ~20, 40, 60, 80 and 100 residues, and the corresponding constructs were named H20A, H40A, etc. In the wild-type context, both signals initiate translocation of the downstream C-terminal part of the protein. However, in the experimental setup, they carry conflicting topogenic information. According to the `linear insertion model`, the N-terminal HA signal would target the polypeptide to the ER membrane and induce translocation of the spacer. The following signal-anchor would act as a `stop-transfer` sequence, leaving the C-terminus in the cytoplasm. In the opposite situation, when the H1 topogenic determinant is dominant, the carboxyterminal portion would be translocated, whereas the HA signal would be forced to insert in the N_{exo}/C_{cyt} orientation or would fail to insert at all. The insertion process was

monitored by lumenal glycosylation of two aparagines located in the C-terminal portion of the polypeptide.

The experiments showed that only when the two signals were sufficiently separated from each other (≥80 residues), the insertion proceeded according to the `linear insertion model`. Polypeptides with shorter linkers inserted in mixed orientations. The first signal induced opening of the channel and translocation of the spacer by the time the second signal entered the pore and initiated inversion of the protein (**Figure 1A**). For this, the linker had to flip back to the cytosolic side. This was prevented by glycosylation of the spacer, acting as a steric hindrance, trapping the linker in the ER lumen and, thus, blocking the reorientation process.

The ratio of topologies depended also on the characteristics of the first signal. Different cleaved secretory signals, such as those of preprolactin (P) or vasopressin (V) had different `strength`, i.e., the ability to dominate the insertion process, which could be ranked P>H>V (**Figure 1B**).

A   **B**
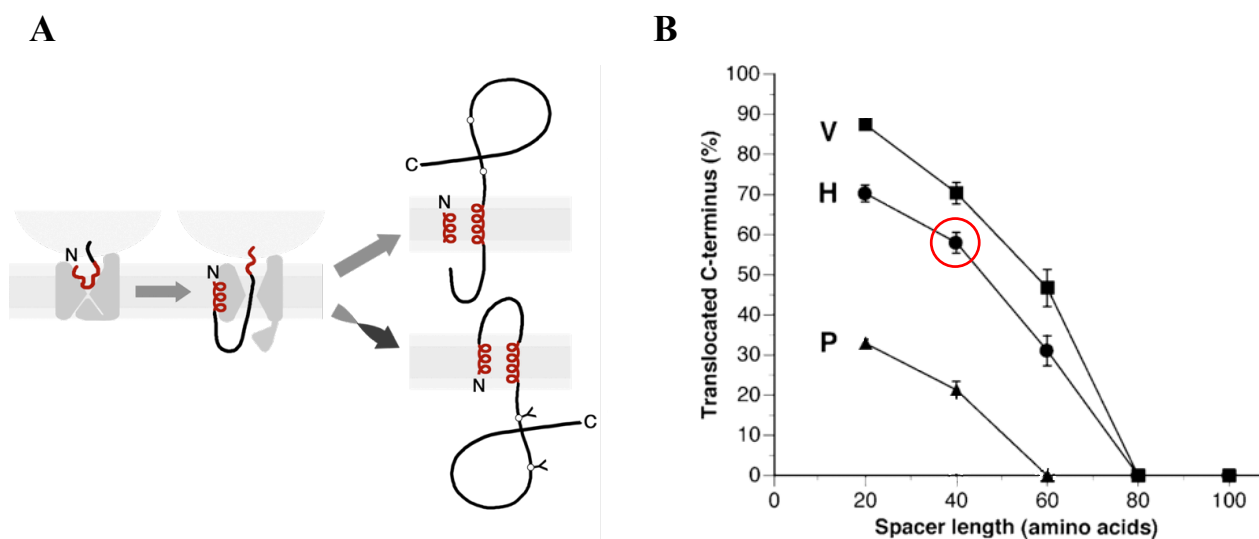
Figure 1. Insertion of proteins with conflicting signals. Two-signal proteins compete for the preferred orientation in the translocon. The first, cleaveable signal induces pore opening and translocation of the spacer. When the signal-anchor enters the channel, it triggers inversion of the protein, relocation of the spacer to the cytosolic side and translocation of the C-terminal domain for a fraction of the polypeptides The reorientation process can be monitored by the glycosylation status of proteins (`Y`s indicate N-linked glycans attached in the ER lumen to the C-terminus of proteins (A). Topogenesis of proteins with conflicting signals dependens on the length of the spacer and the potency of the cleaved signal to initiate type I insertion. V-prepro-vasopressin-neurophysin II, H-hemagglutinin, P-preprolactin. Red circle marks the reference construct, H40A, used in experiments described in this chapter (B) (Goder et al. 1999, modified).

In this chapter, we describe experiments designed to explore the insertion process of polypeptides with conflicting signal sequences. We investigated whether the `translation-stop phenomenon`, previously described for N-terminal signal-anchors, whose reorientation was interrupted upon chain termination (or by an unknown mechanism after ~50 s) (Goder and Spiess, 2003), also applies to two-signal proteins. For this reason, we generated series of model proteins with different lengths of C-terminal domains, which would offer different time windows for the reorientation process. We selected the H40A construct as a starting point for mutagenesis (**Figure 1B,** red circle). It consists of the cleaved hemagglutinin signal, followed by a spacer of 40 amino acids, the H1 signal-anchor and the wild-type C-terminal domain of H1 composed of 230 residues. This protein was found to be inserted into the ER membrane with mixed topologies (60% C-terminal translocation) and could therefore act as a sensor to monitor changes in the insertion process triggered by altering the environment of the signals. This includes topological determinants such as the hydrophobicity of the core of the signal-anchor and the flanking charges.

We found that proteins with competing signals undergo a reorientation within the translocon upon insertion of the signal-anchor. The process is rapid for wild-type H1 signal and can only be monitored when it is slowed down by greatly increasing the hydrophobicity of the second signal or mutagenesis of its flanking charges. Replacement of the HA signal sequence with preprolactin, a stronger signal for type I insertion, altered the final topology of the proteins, according to the `linear-insertion model`.

MATERIALS AND METHODS

Cloning strategy
H40A[50]-[350]

The construction of the H40A[#] series of proteins has been described previously (Goder et al., 1999). Each construct was composed of the cleavable signal sequence of hemagglutinin fused to a linker comprising 40 amino acids of the N-terminal domain of wild-type H1, followed by the signal-anchor of H1 and a tail of 50-350 aa. Because extensive truncations of the C-terminus eliminated some of the methionines located there, four aditional Met were added to the shortest constructs (50-110 aa) to enhance the radioactivity signal (Higy 2005).

H40L25

Exchange of the wild-type H1 signal-anchor into a stretch of 25 leucines required several steps. In the first step, the 40-aa linker was synthesized on the matrix of pEA1Leu25 construct (Wahlberg

and Spiess, 1997) using the short40 sense primer introducing a BglII site, and ECEright antisense primer (all primers are listed in **Table I**). Next, the PCR product, containing also the signal-anchor and the C-terminal domain of 230 amino acids, was cut with BglII and EcoRI and transfered to the pEHC+ plasmid cleaved with BamHI and EcoRI. This vector carried the HA signal sequence flanked by HindIII and BamHI sites and a downstream EcoRI site. Upon ligation with the insert, the BamHI site in the vector was transformed into an XhoII site. Finally, the 230-aa C-terminal BamHI→EcoRI stretch was replaced with different tails from the H40A series.

H40L25-4

The constructs with inverted flanking charges were prepared similarly to H40L25, except that PCR was performed with pSA1-4Leu25 plasmid (Beltzer et al., 1991) as a matrix, using the short40 sense primer and GEMright antisense primer.

H40A-4

The N-terminal part including the signal-anchor sequence of H1 was synthesized on pEA1-4 matrix (Wahlberg and Spiess, 1997) with short40 and ECEright primers, then fused to the HA signal as described above. However, the C-terminal part turned out to be incorrect, i.e., it did not contain the two additional positive charges. To correct this, we amplified the C-terminal domains on the H40A matrix, with a mutagenic sense primer (LysLys-s) and ECEright, then replaced the BamHI→EcoRI carboxyterminal fragments.

H40A-3

The series where the N-terminal positive flank of the signal-anchor was mutated into negatively charged amino acids was prepared similarly to the H40A-4 series, except that C-terminal BamHI→EcoRI tails with normal charge distribution were fused.

H40A Minus VI, Plus VI

Mutations in the apolar core of H1 signal were introduced by PCR of the HindIII→BamHI fragment from H40A series with ECEleft sense primer and MinusVI-a or PlusVI-a mutagenic primers. Then, the wild-type H1 230-aa C-terminus was fused.

H40A Nto D

Point mutation in the C-terminal flank of the signal-anchor was introduced by PCR of the C-terminal tails with the sense NtoD primer and ECEright.

P40A

The HA signal sequence was replaced with bovine preprolactin signal from P40A[230] construct (Goder et al., 1999) (this plasmid was originally called L40A).

All constructs were verified by sequencing with ECEleft, ECEleft2 or ECEright (whenever PCR was involved), or by a restriction digestion test.

Table I. PCR primers.

Primer name	Primer sequence	Primer type
ECEleft	GAAGTAGTGAGGAGGC	sense
ECEleft2	CGGCCTCTGAGCTATTCCAG	sense
ECEright	CTACAAATGTGGTATGGC	antisense
GEMright	GCGAGGAAGCGGAAG	antisense
LysLys-s	CGCGGATCCCAAAACTCCCAGCTGCAGAAGAAGCTGCGGGGCC	sense
MinusVI-a	CGCGGATCCACAGACAACCACAAGC	antisense
NtoD-s	GATCGGATCCCAAGACTCCCA	sense
PlusVI-a	CGCGGATCCTATTACGATCACACAGACAACC	antisense
short40	CCCAGATCTGAGTATCAAGACCTTCAG	sense

Cell culture

COS-1 cells were cultivated in DMEM (Sigma) supplemented with 10% FCS, 100 units/ml penicillin, 100 µg/ml streptomycin, 2 mM L-glutamine, in a humidified incubator containing 7.5% CO_2 at 37°C.

Transient transfections

For transient transfections, cells from a confluent dish were split 1:10 into 6-well plates or 60-mm dishes and transfected the next day with one of the following reagents: Lipofectine (Life Technologies), FuGENE HD (Roche) or polyethylenimine (PEI; Sigma-Aldrich). FuGENE HD, although highly efficient, proved to be toxic to COS-1 cells and was replaced with PEI. Cells were processed 2 days after transfection.

Metabolic labeling with ^{35}S-methionine

Transfected cells were incubated for 30 min in starvation medium (DMEM without methionine and cysteine, containing 2 mM L-glutamine; Sigma). Cells were labeled for 40 min with 100 µCi/ml [^{35}S]protein labeling mix consisting of 77% methionine and 23% cysteine (PerkinElmer), then transfered to 4°C and washed twice with cold PBS.

Immunoprecipitation

Non-integrated proteins were removed by extraction with 0.1% saponin (in PBS, with 2 mM PMSF, 1x PIC) for 30 min, followed by a PBS wash. Cells were lysed in 500 µl of lysis buffer (PBS, 1% TritonX-100, 0.5% deoxycholate, 2 mM PMSF, 1x PIC) for 1 h, then scraped, vortexed and incubated for 30 min on ice. Lysates were cleared by centrifugation and proteins were immunoprecipitated overnight by addition of 500 µl immuno-mix (lysis buffer, 1 mg/ml BSA, 1mM PMSF, 1x PIC) containing 1 µl rabbit anti-serum raised against a peptide corresponding to residues 277-287 at the C-terminus of H1 (anti-H1C). The immune complexes were pulled down by incubation with 10 µl/sample protein A-Sepharose (Zymed) for 1 h. Samples were washed 4 times with immuno-wash (lysis buffer, 1 mg/ml BSA, 1 mM PMSF) and 2 times with PBS (with 1 mM PMSF).

EndoH treatment

At the last PBS wash samples were split in half for deglycosylation by endoglycosidase H. All samples were boiled in endoH buffer (50 mM sodium citrate, 1% SDS; in PBS), 1 µl of the enzyme (Roche) was added and samples were incubated for 3 h at 37°C.

SDS-PAGE and autoradiography

Reduced samples were resolved on 12.5% - 17.5% acrylamide gels and quantified using a phosphorimager (Molecular Dynamics Inc.). Protein orientation in the ER membrane was determined by calculating the ratio of glycosylated proteins divided by total protein within a lane. This value did not depend on the signal intensity caused by different transfection or protein pull-down efficiency.

RESULTS

Topology of the H40A series of proteins shows no C-terminal length dependence

In order to study the insertion process of proteins with conflicting signals, we expressed the H40A[#] series of constructs (**Figure 2A**) (# indicates various lengths of the C-terminal tail; **Figure 2B**) in COS-1 cells and analyzed the ratio of C-terminal translocation. Upon expression, we obtained different protein forms (**Figure 2C**). Constructs with the longest C-terminal tails, 290 and 350 aa, were resolved as two major bands: the upper band corresponds to twice glycosylated

polypeptides (2) representing the state when the C-terminus is translocated into the ER lumen. The lower band represents non-glycosylated products (0). Glycosylation status was confirmed by endoH digestion, which caused the high-molecular weight band to collapse. Because the cells had been subjected to saponin extraction before immunoprecipitation, all products are membrane integrated. The unglycosylated forms therefore correspond to loop-translocated polypeptides (Goder et al., 1999). For all shorter constructs, there is an additional form, sensitive to endoH (1). This represents the once glycosylated product, when the glycosylation site located close to the membrane cannot be efficiently modified.

The shortest constructs produced another, low-molecular weight species, insensitive to deglycosylation (marked by an asterisk in **Figure 2C**). They were previously identified to be the loop-translocated polypeptides with cleaved hemagglutinin signal (Higy 2005). When the signal peptidase cleavage site had been inactivated (mutation of Gly→Leu at position -1 and of Val→Leu at position -3), the short forms were not observed. Signal cleavage is not specific only for the short constructs, but due to limitations of electrophoresis resolution, it cannot de detected for larger proteins under the conditions we tested. For quantitation of the topology ratio, both unglycosylated forms were considered as loop-translocated, while polypeptides with one and two glycans reflect C-terminal translocation.

Glycosylation analysis of the H40A series of constructs revealed no change of the topology with the length of the C-terminus (**Figure 2D**). The same fraction of ~60% C-translocation was obtained over the range of 50 to 350 amino acids downstream of the second signal and it is similar to previous results with a C-domain of 230 amino acids (Goder et al., 1999). This might suggest that the `translation-stop phenomenon`, observed for N-terminal signal-anchors (Goder and Spiess, 2003), does not apply to a situation when the signal-anchor is localized internally. In such a case, the reorientation process would not require ongoing protein synthesis, thus, it would not be blocked upon translation termination and dissociation of the ribosome.

A

HA signal   spacer   H1 signal-anchor

MAIIYLILLFTPVRGDQI RGSEYQDLQHLDNEESDHHQLRKGPPPPQPLLQRLCSGPR LLLLSLGLSLLLLVVVCVI GSQNSQLQEELRGL... **H40A**

B
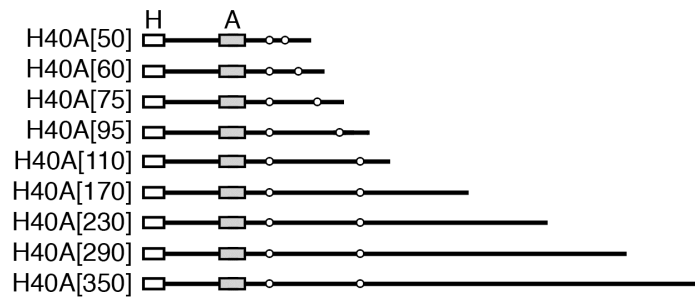
C
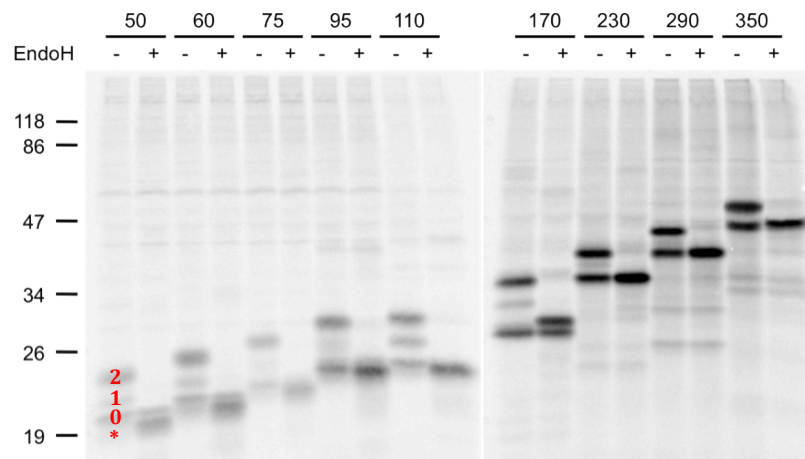
D
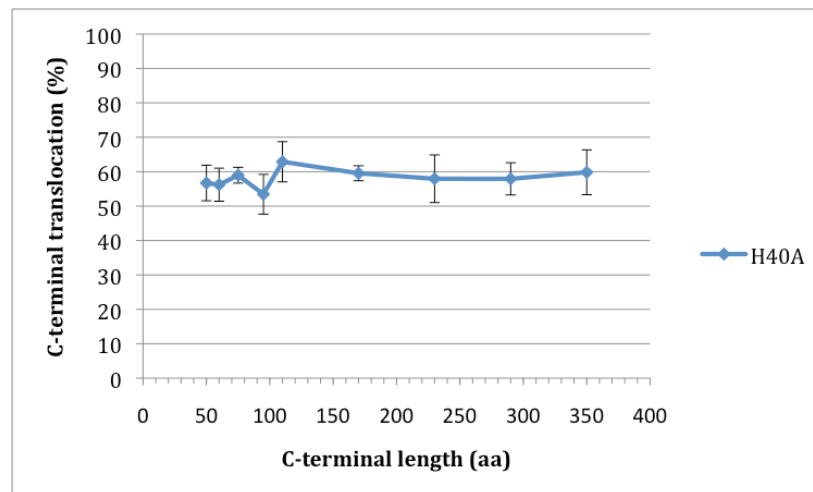
Figure 2. Topology analysis of H40A series of proteins, composed of a cleaved hemagglutin sequence, followed by a linker of 40 amino acids and a type II signal-anchor of H1. H1 signal was flanked by two positive charges at the N-terminus and a net negative charge at the C-terminus (A). The C-terminal domains were composed of 50-350 aa and contained two potential glycosylation sites (B). The proteins were expressed in COS-1 cells, pulse-labeled with [^{35}S]methionine for 40 min, immunoprecipitated, separated by gel electrophoresis, and visualized by autoradiography. Prior to immunoprecipitation, soluble proteins were removed by saponin extraction. 0-unglycosylated products, 1-singly glycosylated, 2-doubly glycosylated, ∗-products with cleaved HA signal. Glycosylation status was confirmed by endoH digestion (C), Glycosylation analysis revealed that topogenesis of these chimeric proteins did not depend on the length of the C-terminus (D). The average and standard deviations of four independent experiments are shown. The marker sizes are in kDa.

Proteins with conflicting signals are only moderately sensitive to changes in hydrophobicity of the signal-anchor

One possibility is that the inversion process occurs much more rapidly than for N-terminal signals and inversion is completed by the time the shortest constructs are completed. Alternatively, there might be no arrest of topogenesis upon termination of translation. We set out to test this by manipulating the topogenic determinants in a way that should slow down polypeptide reorientation within the translocon. One of the most potent factors is the hydrophobicity of the core of the signal-anchor. We decreased the hydrophobicity of the H1 signal by removing the last valine and isoleucine from the signal core (construct named H40A Minus VI), or increased it by extending it with two additional valines and isoleucines (H40A Plus VI). In addition, we dramatically increased the hydrophobicity of the signal-anchor by replacing the wild-type H1 signal sequence with a stretch of 25 leucines (**Figure 3A**). The latter signal had been previously shown to completely block the reorientation of proteins with N-terminal signal-anchors (Wahlberg and Spiess, 1997). Upon expression of the H40A Minus VI/Plus VI constructs with a C-terminus of 230 residues (**Figure 3B**), no significant change of the ratio of glycosylated to unglycosylated polypeptides was observed (**Figure 3C**). Even the signal-anchor composed of 25 leucines (H40L25) produced a reduction of C-terminal translocation of only ~13%. Thus, internal signal-anchors appear to be less sensitive to hydrophobicity changes than N-terminal ones.

Flanking charges affect orientation of two-signal proteins according to the `positive-inside rule`

The rate of inversion of N-terminal signal-anchors strongly depends on the flanking charges, which position the signal in the translocon according to the `positive-inside rule` (Goder and Spiess, 2003). We therefore tested whether signal-anchors contained within two-signal proteins exhibit a similar sensitivity. We reduced the charge difference Δ(N-C) by inversion of the flanking charges: the two positive charges in the N-terminal flank of the H1 signal were replaced by two negative ones (mutation of arginines to aspartates) and the negatively charged glutamates in the C-terminal flank were mutated to positively charged lysines. Inversion of the flanking charges was introduced alone or in combination with an increased hydrophobicity of the signal-anchor (exchange of the H1 signal into a Leu_{25} sequence). The corresponding constructs were named H40A-4 and H40L25-4, respectively. To test the opposite effect, the charge difference of the second signal was increased by replacing the neutral asparagine in the C-terminal flank with a negatively charged aspartate (construct called H40A NtoD) (**Figure 3A**). Expression in COS-1 cells and glycosylation analysis of these proteins (**Figure 3B**) revealed that, as expected, topology of two-signal proteins depended on the charge difference of the signal-anchor.

A

HA signal   spacer   signal-anchor

MAIIYLILLFTPVRGDQI RGSEYQDLQHLDNEESDHHQLRKGPPPPQPLLQRLCSGPR LLLLSLGLSLLLLVVVCVI   GSQNSQLQEELRGL... **H40A**

MAIIYLILLFTPVRGDQI RGSEYQDLQHLDNEESDHHQLRKGPPPPQPLLQRLCSGPR LLLLSLGLSLLLLVVVC   GSQNSQLQEELRGL... **H40A Minus VI**

MAIIYLILLFTPVRGDQI RGSEYQDLQHLDNEESDHHQLRKGPPPPQPLLQRLCSGPR LLLLSLGLSLLLLVVVCVIVI   GSQNSQLQEELRGL... **H40A Plus VI**

MAIIYLILLFTPVRGDQI RGSEYQDLQHLDNEESDHHQLRKGPPPPQPLLQRLCSGPR LLLLLLLLLLLLLLLLLLLLLLLLL GSQNSQLQEELRGL... **H40L25**

MAIIYLILLFTPVRGDQI RGSEYQDLQHLDNEESDHHQLRKGPPPPQPLLQDLCSGPD LLLLSLGLSLLLLVVVCVI   GSQNSQLQKKLRGL... **H40A-4**

MAIIYLILLFTPVRGDQI RGSEYQDLQHLDNEESDHHQLRKGPPPPQPLLQDLCSGPD LLLLLLLLLLLLLLLLLLLLLLLLL GSQNSQLQKKLRGL... **H40L25-4**

MAIIYLILLFTPVRGDQI RGSEYQDLQHLDNEESDHHQLRKGPPPPQPLLQRLCSGPR LLLLSLGLSLLLLVVVCVI   GSQDSQLQEELRGL... **H40A NtoD**

B
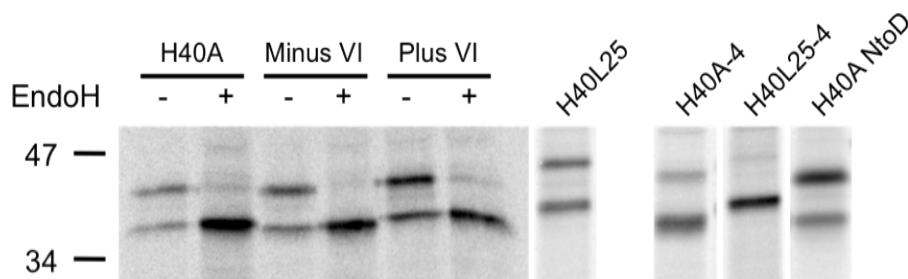
C
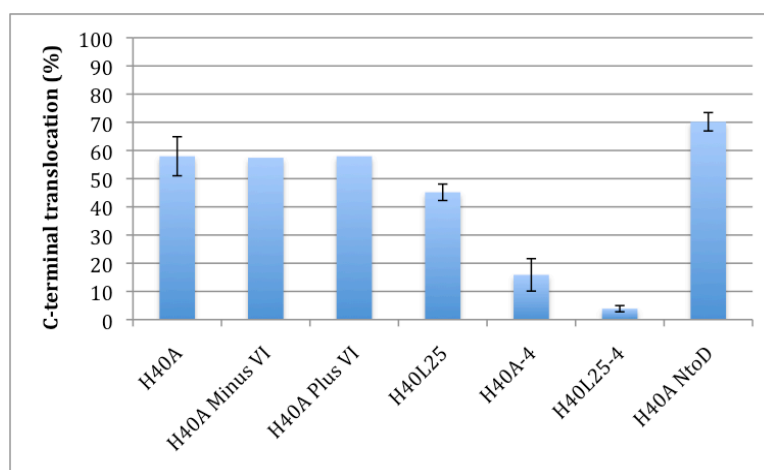
Figure 3. The effect of hydrophobicity of the signal-anchor and the flanking charges on the insertion process of two-signal proteins. Protein constructs carrying mutations in the signal-anchor (H40A Minus VI, H40A Plus VI, H40L25), mutations in the flanking charges (H40A-4, H40A NtoD) or combination of those (H40L25-4) and the control (H40A) (A) were expressed in COS-1 cells (B) and subjected to glycosylation analysis (C). The experiments were conducted for wild-type H1 C-terminus consisting of 230 amino acids. Topology was only midly affected by changes in signal hydrophobicity. Inversion of the flanking charges (H40A-4) strongly reduced C-translocation. Combined mutations of the flanking charges and the signal hydrophobicity resulted in almost exclusive N_{exo}/C_{cyt} topology. An increase of the charge difference between the N- and the C-terminal flank of the signal-anchor (H40A NtoD) produced more C-translocation The plot shows the results of 1-4 independent experiments with standard deviations, where applicable. The marker sizes are in kDa.

Flanking charges proved to be a more powerful topogenic determinant than hydrophobicity of the signal-anchor: inversion of charged residues in the vicinity of the H1 signal reduced the inversion rate by ~45% (compared to ~13% decrease upon replacement of the signal with Leu_{25}). Combined mutations of the flanking charges and the signal hydrophobicity resulted in almost exclusive N_{exo}/C_{cyt} topology. In contrast, an increase of the charge difference by introduction of an additional negative charge at the C-terminal flank (H40A NtoD) produced more C-translocation (**Figure 3C**). These results support the model that the flanking charges (and thus the `positive-inside rule`) are the driving force for signal inversion. Increased hydrophobicity of the signal-anchor reduced C-terminal translocation suggesting that, as for N-terminal signals, a hydrophobic interaction at the translocon slows down the reorientation process.

Two-signal proteins reorient within the Sec61 translocon

In order to investigate the reorientation process of two-signal proteins, we tested the effect of signal hydrophobicity on the insertion of polypeptides with various lengths of the C-terminus, which offer different time windows for signal inversion. We compared the topogenesis of the H40A series of constructs with H40L25[#] (**Figure 4A**). Upon expression in COS-1 cells (**Figure 4B**), we discovered that increased hydrophobicity of the signal-anchor slowed down the reorientation process to a point that allows observation of the kinetics of signal inversion. Topogenesis of the H40L25 series of proteins showed C-terminal length dependence (**Figure 4C, red line**). The fraction of C-terminally translocated polypeptides increased with the increasing length of the C-domain and approached a plateau of ~50%. This indicates that topogenesis of such two-signal proteins is arrested upon termination of translation. It also confirms that the polypeptides reorient upon emergence of the second signal. The data suggest a rapid inversion phase up to ~100 residues (~20 s after emergence of the second signal and ~30 s after the first signal has left the ribosomal tunnel, using a translation rate of 5 aa/s; Hershey 1991), potentially followed by a slow phase of inversion.

A

HA signal   spacer   signal-anchor

MAIIYLILLFTPVRGDQI RGSEYQDLQHLDNEESDHHQLRKGPPPPQPLLQRLCSGPR LLLLSLGLSLLLLVVVCVI   GSQNSQLQEELRGL... **H40A**

MAIIYLILLFTPVRGDQI RGSEYQDLQHLDNEESDHHQLRKGPPPPQPLLQRLCSGPR LLLLLLLLLLLLLLLLLLLLLLLLLL GSQNSQLQEELRGL... **H40L25**

B

C

Figure 4. Kinetics of inversion of two-signal proteins. Proteins with H1 signal (H40A) as well as mutants carrying a strongly hydrophobic signal-anchor (H40L25) and C-terminal tails of 50-350 residues (A) were expressed in COS-1 cells (C-cell lysate after saponin extraction, S-saponin extract containing nonintegrated proteins) (B) and subjected to glycosylation analysis (C). Increased hydrophobicity of the signal-anchor slowed down the reorientation process and revealed the C-terminal length dependence of topology. The plot shows the results of two (H40L25) or four (H40A) independent experiments with standard deviations. The marker sizes are in kDa.

As shown in **Figure 3**, topology of proteins with conflicting signals depends on the charged residues in the vicinity of the signal-anchor. We therefore tested whether mutagenesis of the flanking charges has a similar potential of slowing down the reorientation process as increased signal hydrophobicity. We replaced the two positive charges in the N-terminal flank of the H1 signal (mutation of arginines to aspartates). The C-terminal flank remained unaltered, with one net negative charge. The series of constructs was named H40A-3[#] (**Figure 5A**) and was expressed in COS-1 cells along with the H40A[#] control (**Figure 5B**). Proteins with mutations in the flanking charges showed a significantly reduced ratio of C-terminal translocation: ~10%

C-translocation was obtained for the shortest constructs ([50]-[110]) and it increased up to ~35% for H40A-3[350] (**Figure 5C**). This biphasic insertion pattern could be potentially explained by rapid inversion of ~10% of polypeptides with short C-termini, followed by the C-terminal length dependence of proteins with tails >100 amino acids. Another possibility is targeting to the ER by the second signal, which gives a ″background″ of ~10% C-terminally translocated proteins.

A

HA signal   spacer   H1 signal-anchor

MAIIYLILLFTPVRGDQI RGSEYQDLQHLDNEESDHHQLRKGPPPPQPLLQRLCSGPR LLLLSLGLSLLLLVVVCVI GSQNSQLQEELRGL... **H40A**

MAIIYLILLFTPVRGDQI RGSEYQDLQHLDNEESDHHQLRKGPPPPQPLLQDLCSGPD LLLLSLGLSLLLLVVVCVI GSQNSQLQEELRGL... **H40A-3**

B
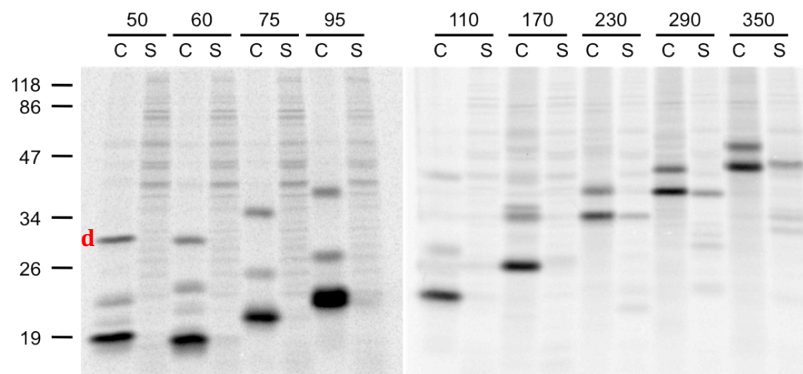
C
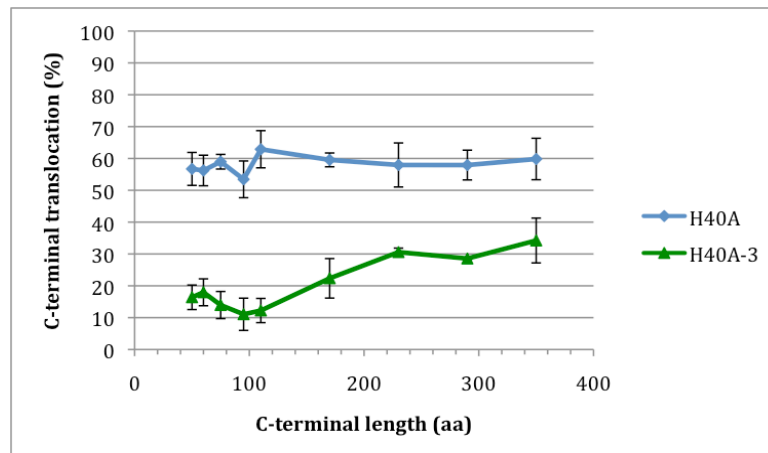
Figure 5. The effect of the flanking charges on the insertion process of two-signal proteins. Proteins containing the H1 signal-anchor with wild-type charge distribution (H40A[#]) as well as mutants where the positively charged residues in the N-terminal flank were replaced by negative charges, H40A-3[#] (A) were expressed in COS-1 cells (C-cell lysate after saponin extraction, S-saponin extract containing nonintegrated proteins; d-oxydative dimer) (B) and subjected to glycosylation analysis (C). Reduction of Δ(N-C) charge difference decreased the rate of signal inversion determined by the `positive-inside rule`. Topogenesis depends on the C-terminal length. The plot shows the results of two independent experiments with standard deviations. The marker sizes are in kDa.

Topology of proteins with conflicting signals is influenced by the translation time

Since the process of protein inversion depends on the size of the growing polypeptide, it can be manipulated by extending the time of translation. This is achieved by addition of the reversible elongation inhibitor, cycloheximide (CHX), during radiolabeling of the newly synthesized proteins. Cycloheximide concentration of 1 µg/ml has been shown to reduce the translation rate in COS cells by a factor of 1.8 (Goder et al., 2000). When applied to proteins with N-terminal signal-anchors, it resulted in a shift of the topologies in favor of C-terminal translocation (Goder and Spiess, 2003). Interestingly, it did not affect the fixed, 50-s time window, after which topogenesis was terminated by the translocation machinery. The higher inversion rate at slow translation could be explained by the fact that a smaller portion of the polypeptide has been synthesized during this 50-s time period and thus, such shorter chains can move more easily through the translocation channel as the signal reorients itself. When we applied cycloheximide to our system (**Figure 6B**), we observed a general increase in C-terminal translocation of the H40A series of proteins (**Figure 6C**). The shift reached ~10-15% for short polypeptides and ~20-25% in case of longer proteins. Since the inversion was already rapid at normal translation rate, we did not expect much change when the polypeptides were offered more time to achieve their proper topology. Only proteins with a long C-terminal domain could benefit in the situation when the Sec61 channel was filled with a smaller portion of the nascent chain.

Figure 6. The effect of translation time on the topology of two-signal proteins. H40A series of constructs (A) was radiolabeled in the presence of the elongation inhibitor, cycloheximide (CHX), at 1µg/ml concentration (B). Prolonged time of protein synthesis resulted in an increased fraction of polypeptides inserted in N_{cyt}/C_{exo} topology (C). The plot shows the results of two (H40A+CHX) or four (H40A) independent experiments with standard deviations. The marker sizes are in kDa.

The ratio of topologies of two-signal proteins depends on the characteristics of the first signal

In the last experiment, we determined the influence of the N-terminal, cleaved signal on topogenesis of two-signal proteins. For this purpose, the hemagglutinin signal in the H40A series was replaced by the cleavable preprolactin signal, and the resulting series of constructs was named P40A (**Figure 7A**). Proteins with both types of cleaveable signal sequences were expressed in COS-1 cells (**Figure 7B**) and subjected to glycosylation analysis (**Figure 7C**).

A

HA signal   spacer   H1 signal-anchor
MAIIYLILLFTPVRGDQI RGSEYQDLQHLDNEESDHHQLRKGPPPPQPLLQRLCSGPR LLLLSLGLSLLLLVVVCVI GSQNSQLQEELRGL...   **H40A**

preprolactin signal   spacer   H1 signal-anchor
MDSKGSSQKGSRLLLLLVVSNLLLCQGVVSTPVCPP RGSEYQDLQHLDNEESDHHQLRKGPPPPQPLLQRLCSGPR LLLLSLGLSLLLLVVVCVI GSQNS...   **P40A**

B

C

Figure 7. The influence of the cleaved signal on the insertion process of two-signal proteins. The HA signal in the H40A series was replaced with preprolactin signal sequence (A) and both series of constructs were expressed in COS-1 cells (B). Glycosylation analysis revealed that the preprolactin signal is a stronger inducer of type I topology than hemagglutinin signal and it is able to dominate the insertion process of the P40A protein series (C). The plot shows the results of two (P40A) or four (H40A) independent experiments with standard deviations. The marker sizes are in kDa.

As shown in **Figure 1B**, preprolactin is a strong inducer of type I topology (Goder et al., 1999). In our experimental setup, its topological information was dominant and only 10-20% of polypeptides were able to translocate their C-terminus (**Figure 7C**). This ratio seemed to be independent of the size of the protein, however, the values were not contained within a range that would allow us to observe the inversion kinetics. This could be feasible by introducing mutations that favour type II topology, such as insertion of additional positive charges preceding the signal-anchor.

DISCUSSION

Topogenesis of multispanning membrane proteins is a complex process mediated by topogenic determinants contained within the protein sequence, interactions of the protein with the translocation machinery, interactions between individual transmembrane segments and protein-membrane phospholipids interactions (Dowhan and Bogdanov, 2009). Here, we have analyzed the process of *in vivo* insertion of two-signal proteins into the mammalian ER membrane. We generated series of chimeric proteins composed of a cleaved hemagglutin signal for type I insertion, followed by a spacer of 40 aa and a type II signal-anchor sequence of H1. These signals contain conflicting topogenic information and thus, they compete for the preferred orientation in the translocon (**Figure 1A**). The topogenic factors encoded in the wild-type H1 sequence are dominant during the insertion of ~60% of the polypeptides belonging to the H40A[#] series. The remaining fraction of ~40% products inserted in an N_{exo}/C_{cyt} topology might be the result of BiP binding to the initially translocated spacer, which blocks its return to the cytosolic side and, thus, renders the protein unable to reorient. Long linkers would increase the chance of capturing by BiP – this is supported by the experiments of Goder et al. (1999), where extension of the loop sequence resulted in a decreased fraction of C-terminally translocated products (**Figure 1B**).

Previous studies of the insertion process of conflicting signals raised a concern that what we are actually observing is not the signal competition for the preferred topology, but competition for SRP recruitment in the cytosol (Goder et al., 1999). In such a situation, SRP binding by the HA signal will initiate loop translocation and will result in N_{exo}/C_{cyt} topology. In contrast, if the signal-anchor first binds SRP, it will induce translocation of the C-terminus, whereas the HA signal will remain in the cytosol or will subsequently insert into the membrane. To investigate

this, Goder et al. tested translocation of polypeptides consisting of the cleaved signal of vasopressin or hemagglutin and a short downstream segment, corresponding to the moment in the translation of V40A or H40A constructs when the second signal starts emerging from the ribosome. The results suggested that the second signal in those constructs is able to compete for SRP binding. Despite this, a larger fraction of the V40A than of the H40A constructs inserted with a translocated C-terminus, even though targeting by the vasopressin signal was faster (**Figure 1B**) (Goder et al., 1999). In our experiments, background of ~10% targeting by the second signal could potentially explain the biphasic insertion pattern of H40A-3[#] (**Figure 5C**) and P40A[#] (**Figure 7C**) series, where short polypeptides (up to ~100 residues in the C-terminus) showed a similar, low level of C-translocation, followed by a phase of inversion associated with C-terminal length dependence. However, in general we can assume that the percentage of polypeptides targeted by the second signal in the series of our diagnostic constructs is insignificant and different topologies presented in the model in **Figure 1A** are indeed the result of signal competition.

Competition between signals is manifested by reorientation of the proteins induced by the signal-anchor. For the wild-type H1 signal it occurs very rapidly and has already completed by the time the shortest of our constructs was synthesized. The kinetics of reorientation can be only observed when the process of signal-anchor inversion is slowed down significantly by increasing signal hydrophobicity or by reducing the charge difference of the signal-anchor. Our experiments revealed that signal-anchors localized internally are far less sensitive to increased hydrophobicity than N-terminal signal-anchors in single spanning membrane proteins (Goder and Spiess, 2003). A possible explanation for the observed behaviour may lie in different functions of the two signals during the insertion process. In single-spanning membrane proteins, insertion of the signal into the translocation channel triggers a series of rearrangements of the structure of the Sec61 complex that facilitate protein translocation. These include widening of the constriction ring and removal of the plug domain sealing the channel, that enable translocation of the C-terminal chain. In addition, the channel has to open laterally to release the transmembrane segment into the lipid environment. In contrast, during insertion of our model two-signal proteins, the translocon is already open and the loop is translocated by the time the second signal enters the pore. Thus, some of the functions of N-terminal signal-anchors do not apply to the internal ones.

It has been shown that N-terminal signal-anchors initially insert into the translocon `head-on`, in an N_{exo}/C_{cyt} orientation, followed by signal inversion (Goder and Spiess, 2003). In order to reorient itself, the signal needs to dissociate from the apolar binding site in the translocation pore (Van den Berg et al., 2004) or the membrane (Higy et al., 2005). Perhaps the observed low sensitivity to hydrophobicity of internal vs. N-terminal signal-anchors is caused by

their different binding strength to TM helices of the channel and/or the time the signal spends inside the pore before the lateral release into the membrane. Internally localized signal-anchors contained within two-signal proteins are potentially not yet in contact with the lipid phase when reorientation occurs. These features would be the consequence of different roles of these two types of signals during protein insertion, as mentioned above.

Topogenesis of single-spanning membrane proteins with an N-terminal signal-anchor is coupled to protein synthesis and is interrupted upon chain completion (`translation-stop phenomenon`). Protein reorientation seemed to stop ~50 s after insertion into the channel, even at reduced translation rate caused by cycloheximide treatment. It was proposed that cells might have developed this mechanism in order to purge the translocon of polypeptides that failed to attain their proper topology within the programmed time frame (Goder and Spiess, 2003). We have looked for evidence of a similar checkpoint during topogenesis of two-signal proteins, however, the results were unclear. For only two series of constructs, H40L25[#] and H40A-3[#], we observed a C-terminal length dependence of topology. In the first series, the relationship between protein inversion and C-terminal length was biphasic, with an initial rapid inversion phase, potentially followed by a slow inversion phase, since no plateau was reached within the diagnostic range of 50 to 350 amino acids in the C-terminal domain. In the case of H40A-3[#] protein series, topogenesis appeared to have come to a halt for polypeptides ≥230 residues. If we assume this is indeed the moment of topogenesis termination by the translocation machinery, we obtain a topogenesis window of 40 s. This is based on the calculation that topogenesis starts when the signal-anchor has just emerged from the ribosome and 30 amino acids of the nascent chain are still hidden within the ribosome tunnel exit (Matlack and Walter, 1995; Morgan et al., 2000). The translation rate in mammalian cells is ~5 aa/s (Hershey, 1991). If one argues that topogenesis starts when the first signal has been exposed to the cytosol, this will increase the time window value by 12 seconds, which is close to the one measured for single-spanning proteins. However, this is calculated for only one series of proteins, H40A-3, and we do not know if other double-spanning membrane proteins show a similar pattern.

Different time windows obtained for two-signal proteins would suggest that the topogenesis window depends on the substrate, rather than being a constant programmed in the translocation machinery. Another possibility is that the 50-s time window is the period during which the translocon or perhaps the ribosome scans the polypeptide in search of hydrophobic transmembrane segments. If a potential TM is detected, the timer will be restarted and the segment will be allowed to orient itself in the pore, mediated by its topogenic determinants. That is why topogenesis of single-spanning proteins was terminated, even though the C-terminal portion was

still being synthesized. The exact mechanism of insertion of multispanning membrane proteins is still unknown, including how many TMs the translocation channel can accomodate and how long it takes to assemble a polytopic protein.

Part III: The hydrophobic core of the Sec61 translocon defines the hydrophobicity threshold for membrane integration

Tina Junne, Lucyna Kocik, and Martin Spiess

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

Lucyna Kocik generated the 6W, 6WN and 6WΔ strains, performed the experiment represented in Figure 4 and provided critical comments on the manuscript.

ABSTRACT

The Sec61 translocon mediates the translocation of proteins across the endoplasmic reticulum membrane and the lateral integration of transmembrane segments into the lipid bilayer. The structure of the idle translocon is closed by a lumenal plug domain and a hydrophobic constriction ring. To test the function of the apolar constriction, we have mutated all six ring residues of yeast Sec61p to more hydrophilic, bulky, or even charged amino acids (alanines, glycines, serines, tryptophans, lysines, or aspartates). The translocon was found to be surprisingly tolerant even to the charge mutations in the constriction ring, since growth and translocation efficiency were not drastically affected. Most interestingly, ring mutants were found to affect the integration of hydrophobic sequences into the lipid bilayer, indicating that the translocon does not simply catalyze the partitioning of potential transmembrane segments between an aqueous environment and the lipid bilayer, but that it plays an active role in setting the hydrophobicity threshold for membrane integration.

Abbreviations used: CPY, carboxypeptidase Y; DPAPB, dipeptidyl aminopeptidase B; ER, endoplasmic reticulum; GPD, glyceraldehyde-3-phosphate dehydrogenase; HA, hemaglutinin; PCR, polymerase chain reaction; TM, transmembrane.

INTRODUCTION

Protein translocation across the endoplasmic reticulum (ER) membrane is initiated by a hydrophobic signal sequence which, mediated by signal recognition particle (SRP) and SRP receptor, is targeted to the Sec61 translocon (Osborne et al., 2005). Here the signal is oriented to transfer one end across the membrane and to integrate itself into the membrane. The translocon provides a pore for hydrophilic polypeptide segments to pass through, while simultaneously facilitating the integration of apolar segments into the lipid bilayer.

Mutagenesis of substrate proteins showed that charged residues flanking the hydrophobic core of a signal or signal-anchor sequence are important to define its final orientation according to the 'positive-inside rule' - generally positioning the more positive end on the cytoplasmic side (Beltzer et al., 1991; Hartmann et al., 1989; Heijne, 1986). The hydrophobicity of the signal influences the orientation process (Goder and Spiess, 2003) and drives integration into the

membrane and insertion of the adjacent hydrophilic segment into the pore (Kida et al., 2009). Furthermore, subsequent apolar segments integrate into the membrane depending on their hydrophobicity. Crosslinking studies led to the proposal that membrane integration is a multistep process involving intermediate binding sites (Do et al., 1996; Ismail et al., 2006). Systematic analysis of potential transmembrane segments (TM) in mammalian *in vitro* and *in vivo* systems, in bacteria, and in yeast (Hessa et al., 2005; Hessa et al., 2007; Hessa et al., 2009; Lundin et al., 2008; Xie et al., 2007) yielded 'biological hydrophobicity scales' and suggested that membrane insertion is fundamentally a thermodynamic partitioning process. Based on this interpretation, it was proposed that the function of the Sec61p channel is to provide a site in the membrane through which TMs can equilibrate between the lipid and aqueous phases (Heinrich et al., 2000; von Heijne, 2006).

Crystal structures of the archaeal SecYEß translocon (Van den Berg et al., 2004) provided a first basis to understand these processes mechanistically. SecY/Sec61α is a compact 10-helix bundle of two halves that may open a lateral gate towards the lipid membrane between TM helices 2/3 and 7/8, as illustrated in **Figure 1** for the model of the yeast Sec61 complex (Junne et al., 2006). In the idle translocon, the central pore is obstructed by a lumenal plug domain (highlighted in **Figure 1A**), but in addition also by a central constriction (**Figure 1, B** and **C**). The latter is generated by six, almost invariably hydrophobic side chains provided by TMs 2, 5, 7, and 10. This constriction ring might be responsible for the good viability of yeast cells and the short-term survival of bacteria with full deletion of the plug domain in Sec61p and SecY, respectively (Junne et al., 2006; Maillard et al., 2007).

Mutations in the plug, in the constriction ring, as well as in the helices forming the lateral gate were found to destabilize the closed state of the translocon, resulting in a *prl* (*pr*otein *l*ocalization) phenotype that suppresses inactivating mutations in signal sequences, both in bacteria (Emr et al., 1981; Li et al., 2007; Veenendaal et al., 2004) and in yeast (Junne et al., 2007). Such mutations in bacterial SecY produced transient channel openings in planar membrane permeability measurements (Saparov et al., 2007). In addition, *prl* mutants were shown to affect signal-anchor topology by premature opening of the translocation pore, before the orientation of the signal is completed (Junne et al., 2007).
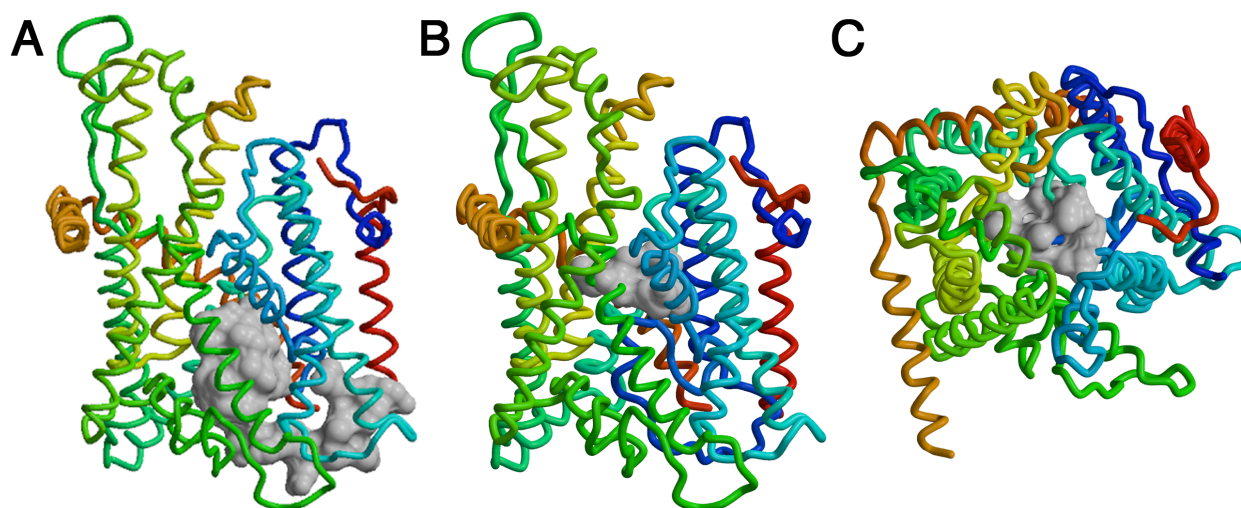
Figure 1. Plug domain and constriction ring of Sec61p. The model of yeast Sec61 complex is shown as the polypeptide backbone (Sec61p in blue to yellow, Sbh1p in red, and Sss1p in orange) with the plug domain (residues 52–74; A) or the residues of the constriction ring (V82, I86, I181, T185, M294, and M450; B and C) in space-filling representation in gray. Views from within the membrane (A and B) or from the cytosol (C) are shown with the lateral exit gate to the front or bottom, respectively.

To specifically analyze the importance and function of the hydrophobic constriction ring, we have mutated all of its contributing residues to more hydrophilic or even charged amino acids, alone or in combination with a point mutation in the plug or a full plug deletion, and analyzed the resulting phenotypes with respect to viability, translocon assembly and stability, translocation defects, and TM integration. The translocon was found to be surprisingly tolerant to even drastic mutations in the constriction ring. Most interestingly, ring mutants were found to affect the integration of hydrophobic sequences into the lipid bilayer, indicating that the translocon does not simply catalyze the partitioning of potential TM segments between an aqueous environment and the lipid bilayer, but that it plays an active part in setting the threshold for lipid integration.

MATERIALS AND METHODS

Yeast strains

Yeast strain VGY61 (Goder et al., 2004) corresponds to RSY1293 (*matα, ura3-1, leu2-3,-112, his3-11,15, trp1-1, ade2-1, can1-100, sec61::HIS3*, [pDQ1]) (Pilon et al., 1997) in which pDQ1 (i.e. YCplac111 (*LEU2 CEN*) containing *SEC61* with codons 2–6 replaced by codons for H$_6$RS and with its own promoter) was exchanged for YCPlac33 (*URA3 CEN*) with the same *SEC61*

gene. This made it possible to introduce mutant *sec61* in YCplac111 (*LEU2 CEN*) by plasmid shuffling using 5-fluoro-orotic acid. The absence of wild-type *SEC61* was confirmed by polymerase chain reaction (PCR) and restriction enzyme digestion of the products. VGY61 with a disruption of *SSH1* was described previously (Goder et al., 2004).

Mutagenesis of Sec61p

Sec61p mutant strains with the mutations L63N and Δplug (residues 52–74 replaced by a glycine) have been previously described (Junne et al., 2006). Ring mutations were introduced sequentially in each of four quarters of ~350 bp of the coding sequence delimited by unique restriction sites *Xba*I, *Sac*I (created by a silent mutation), *Stu*I, *Acc*I, and *EcoR*I (464 bp after the stop codon) at nucleotide positions 27, 343, 710, 1099, and 1916 from the initiation codon, respectively. V82/I86 and I181/T185 were mutagenized simultaneously, whereas mutations of M294 and M450 were generated separately. In proximity of these four loci new silent restriction sites, *Asp718, BamHI, PstI* and again *BamHI,* were created at positions 232, 541, 892, and 1336, respectively. Mutagenesis was performed by PCR using appropriate mutagenic primers and Vent polymerase (New England Biolabs). Ring and plug mutations were combined via *Asp718*. All constructs were verified by sequencing.

Growth analysis and Sec61p levels

For serial dilution experiments, yeast strains were grown in YPDA medium at 30°C to mid-log phase and diluted to 0.1 OD_{600}. Aliquots of 6.6-fold serial dilutions were transferred onto YPDA plates and incubated at 15, 30, or 37°C.

To determine steady-state levels of Sec61p by immunoblot analysis, 10 OD_{600} equivalents of yeast cells were lysed in SDS-sample buffer with glass beads and boiled for 10 min. Aliquots of equal total protein were separated by SDS-gel electrophoresis, blotted onto nitrocellulose, and decorated with a rabbit antiserum against the C terminus of Sec61p. Antibody was detected using horseradish peroxidase-conjugated anti-mouse secondary antibody and the enhanced chemoluminescence kit (GE Healthcare). Equal protein loading was approximated based on Coomassie blue staining of a separate gel.

To analyze the stability of Sec61p mutants in the presence of wild-type Sec61p, the *sec61* coding sequences were extended by a sequence encoding a triple-HA epitope tag, cloned with the original promoter into YCplac111 (*LEU2 CEN*), transformed into VGY61, and grown on SD–Leu–Ura to maintain both wild-type and hemagglutinin (HA)-tagged mutant copy of Sec61p. Translocons

were analyzed by immunoblotting as above using antibodies directed against the C terminus of Sec61p and against the HA epitope, respectively.

Model proteins

The substrate proteins dipeptidyl aminopeptidase B (DPAPB), carboxypeptidase Y (CPY), and CPYΔ3 were described previously (Junne et al., 2007). In CPYΔC, the C-terminal 209 amino acids of CPY were deleted by PCR mutagenesis and fused to a triple-HA tag. To determine the effect of Sec61p mutations on membrane integration, the potential TM segments developed by Hessa et al. (2005) and shown in **Table I** were inserted into the translocated domain of DPAPB replacing codons 170–378, by PCR mutagenesis. The resulting model proteins thus consisted of an N-terminal cytoplasmic domain, a signal-anchor, a spacer sequence, the potential TM segment, and a C-terminal sequence of 29, 16, 124, 27, and 470 residues, respectively. Spacer and C-terminal sequence contain 4 and 3 potential glycosylation sites, respectively. They were expressed in pRS426 (*URA3 2μ*) with a glyceraldehyde-3-phosphate dehydrogenase (GPD) promotor and a C-terminal triple-HA tag.

Labeling and immunoprecipitation

Yeast cells were *in vivo* pulse labeled for 5 min with 150 μCi/ml [^{35}S]methionine/cysteine (PerkinElmer) and, if indicated, chased with 30 μg/ml each of unlabeled methionine and cysteine and 3 mM ammonium sulfate. Cells were lysed with glass beads, heated at 95°C for 5 min with 1% SDS, cleared by centrifugation, subjected to immunoprecipitation, and analyzed by SDS-gel electrophoresis and autoradiography as described previously (Junne et al., 2006). Signals were quantified by phosphorimager.

RESULTS

Sec61p mutants with hydrophilic or even charged constriction residues retain functionality

In order to test the importance of the hydrophobic constriction ring in Sec61p, the six amino acids constituting it, V82, I86, I181, T185, M294, and M450, were first mutated to either alanines (6A) or serines (6S). Since these mutants produced no striking growth defects, we also generated mutants in which the ring residues were replaced by glycines (6G), lacking any side chain. In addition, an opposite mutant with six tryptophans, the most bulky amino acid, was produced (6W). Finally, all six ring residues were also mutated to the charged amino acids

aspartate (6D) or lysine (6K). We expected these latter mutants to be nonfunctional, because they were likely to interfere with the proper insertion of the Sec61p TM segments or because charge repulsion might prevent the formation of the helix bundle. To our surprise, however, all mutants supported growth at 30°C, even in the absence of the second, nonessential Sec61p homolog Ssh1p (**Figure 2**).
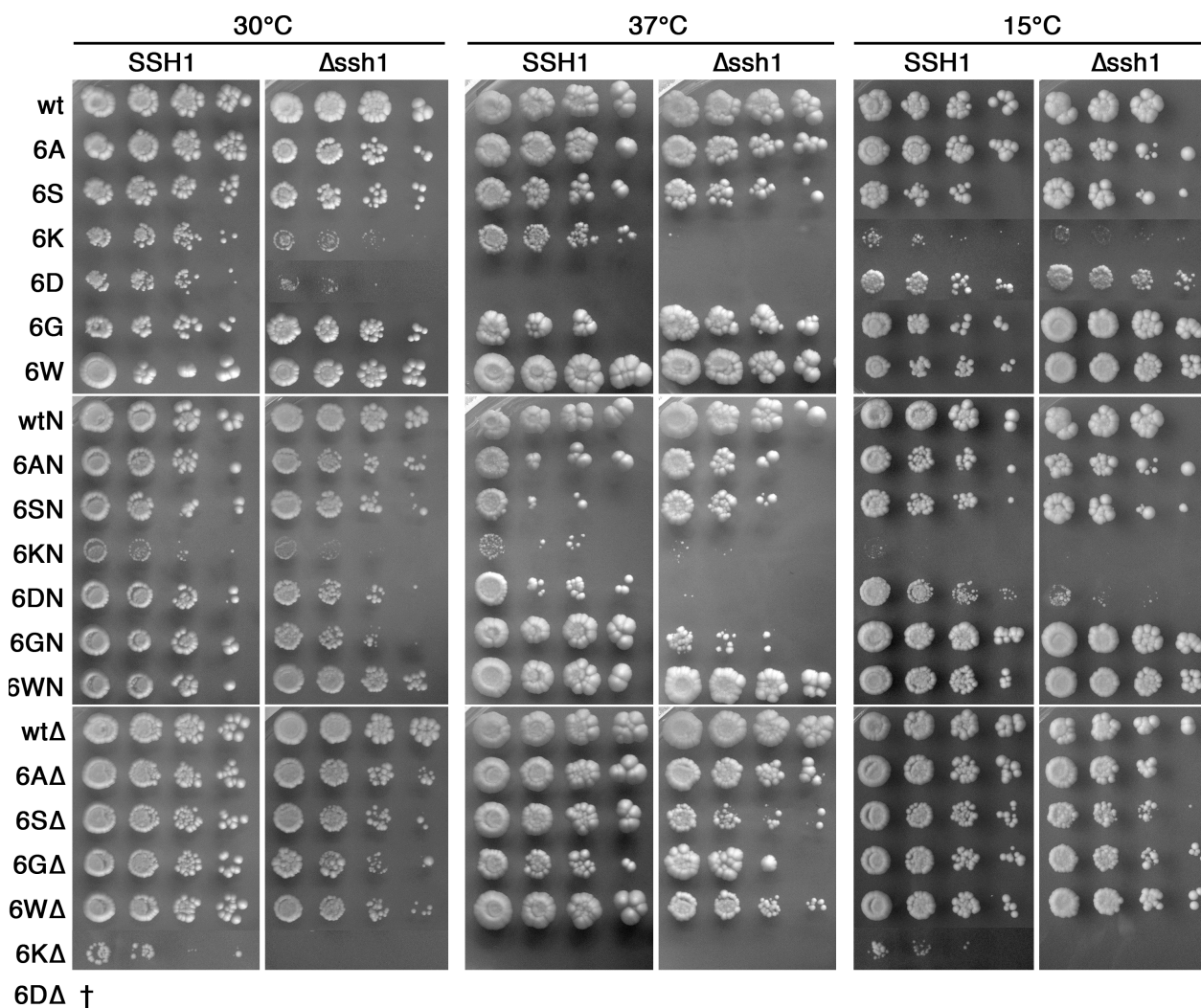

Figure 2. Growth of yeast cells with wild-type or mutant Sec61p in the presence or absence of Ssh1p. *SSH1* or Δ*ssh1* cells expressing the indicated Sec61p mutants were plated at serial dilutions onto YPDA plates and incubated for 3 d at 30°C, 5 d at 37°C, or 11 d at 15°C.

We further constructed Sec61p mutants in which the ring mutations were combined with the L63N point mutation in the plug domain (Junne et al., 2006) named 6XN (X standing for A, S, G, W, D, or K), or with the full plug deletion (replacement of residues 52–74 by a glycine; Junne et al., 2006) named 6XΔ. As is shown in **Figure 2**, yeast cells with any of these mutants in place

of wild-type Sec61p and in the absence of the nonessential *SEC61* homolog *SSH1* were viable except for cells containing 6DΔ, which could not lose the wild-type copy of SEC61. In addition, cells with 6KΔ grew so poorly that they were not yet visible after 3 d. *SSH1* rescued growth of cells with 6KΔ, but not with 6DΔ. Not unexpectedly, it was the charge mutants that showed the severest growth defects: 6K, 6D, 6KN, and 6KΔ had the lowest growth rates, and 6K, 6D, 6KN, 6DN, and 6KΔ showed heat and/or cold sensitivity, in some cases rescued by expression of *SSH1*. Ssh1p is functional only in co-translational translocation, since it does not assemble with the Sec62–Sec63 complex essential for post-translational translocation (Finke et al., 1996), but is found associated with translating ribosomes (Prinz et al., 2000) and co-translational substrate proteins (Wittke et al., 2002). Rescue of growth in the presence of Ssh1p thus suggests that cotranslational translocation was limiting. The growth behavior of the other mutants showed little, if any, deviation from wild-type.

Ring mutations affect translocation efficiency

To test the functionality of mutant translocons with respect to co- and post-translational translocation, the translocation efficiency was tested for DPAPB and CPY (**Figure 3**, **A–C**), established co- and post-translational substrates, respectively (Ng et al., 1996). Rather modest defects were detected for co-translational translocation of DPAPB with less than 30% nonintegration even for the charge mutants. However, ~50% of CPY precursor failed to be translocated by 6K, 6D, and 6KN translocons during the 5-min labeling period, and even more by 6KΔ, whereas the other mutants showed only mild defects in comparison to the respective control (wt, wtN, or wtΔ; **Figure 3**, **A** and **C**).

While unglycosylated full-length products of an obligatory co-translational substrate directly reflect the defect in translocation, unglycosylated products of a post-translational substrate primarily indicate a reduced rate of translocation resulting in an increased pool of cytosolic precursor. To test whether the CPY precurors not translocated after the labeling period can still be translocated later on, we performed pulse-chase experiments. With CPY this is complicated by the fact that mature CPY, or after deglycosylation the ER and Golgi forms, comigrate with the unglycosylated precursor. For this reason, we analyzed CPYΔC, a C-terminally truncated version of CPY that cannot fold and is retained in the ER.
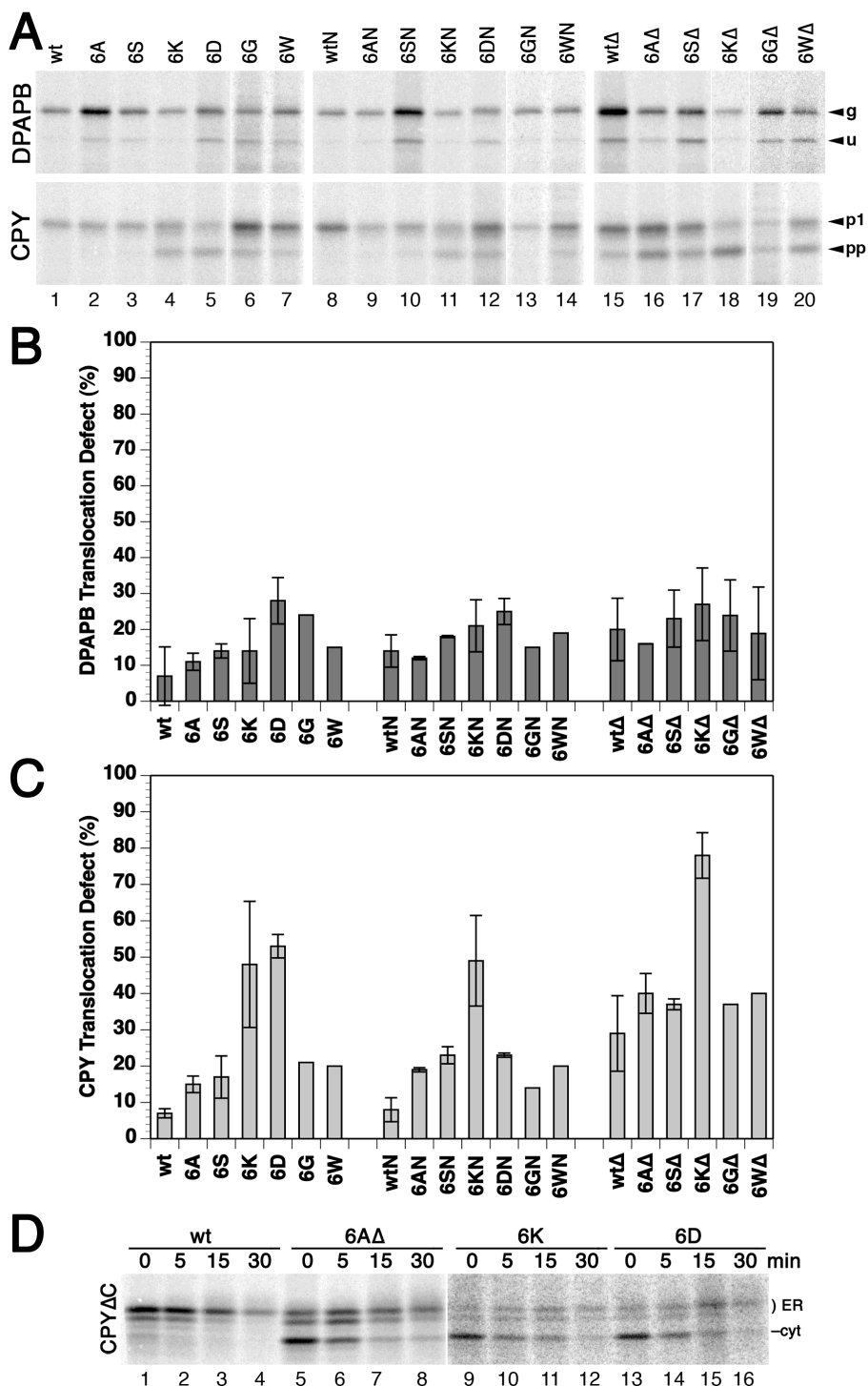
Figure 3. Translocation efficiency of wild-type and mutant Sec61p. Integration of DPAPB as a co-translational and of CPY as a post-translational substrate of the Sec61 translocon was analyzed in a Δ*ssh1* background by pulse labeling for 5 min with [^{35}S]methionine, immunoprecipitation, gel electrophoresis, and autoradiography (A). The products correspond to glycosylated (g) and unglycosylated (u) forms of DPAPB, and to the glycosylated first proform (p1) and the unglycosylated preproform (pp) of CPY. Results were quantified by phosphorimaging and the fraction of untranslocated DPAPB (B) and CPY (C) was plotted (mean and standard deviation of three determinations; single measurements for 6GΔ and 6WΔ in C). In panel D, C-terminally truncated CPYΔC was expressed in cells with the indicated wild-type and mutant translocons, pulse labeled for 5 min and chased with unlabeled methionine for up to 30 min before immunopreciptation, gel electrophoresis and autoradiography to separate the translocated, two- and threefold glycosylated ER forms (ER) from cytosolic precursor (cyt).

Expressed with wild-type Sec61p, CPYΔC was almost completely glycosylated and thus translocated within the pulse period (**Figure 3D**, lane 1). During the chase, the signal was gradually reduced to ~25% within 30 min by degradation (lanes 2–4). In cells with mutant translocons 6AΔ, 6K, or 6D (lanes 5–16), the signal of the glycosylated ER forms initially increased during the chase and then decreased more slowly, indicating that cytosolic precursors continued to be translocated in this period. The post-translational defects observed by pulse-labeling in **Figure 3** (**A** and **C**) reflect reduced translocation rates and not the final loss of translocated protein, which is defined by competition between the rates of translocation and cytosolic degradation.

The hydrophobic constriction ring stabilizes the closed state of the translocon

Single point mutations in the constriction ring have previously been shown to produce a *prl* phenotype, i.e. the suppression of inactivating mutations in signal sequences, both in the bacterial system (Smith et al., 2005) and in yeast (Junne et al., 2007). In general, *prl* mutations are interpreted to destabilize specifically the closed state of the translocon by disturbing the structure of the plug, its binding site in the lumenal cavity, or lateral gate closure, and thus facilitate pore opening. As a result, even marginally hydrophobic signal sequences obtain access to the membrane that are rejected by the wild-type translocon. To test for a *prl* phenotype, translocation of CPYΔ3 was tested, a mutant CPY in which the signal was inactivated by deletion of three apolar residues to yield less than 15% translocation with wild-type Sec61p. Replacement of all six hydrophobic constriction residues by alanines, serines, or glycines showed a clear *prl* effect, since more than 50% of CPYΔ3 was translocated (**Figure 4**). 6W showed a smaller effect, whereas 6D and 6K did not suppress the signal mutation, even when taking into account their general translocation defect. The effect of the charge mutations and in part of 6W thus appears not to specifically destabilize the closed state in favor of the open one, but to disturb the structure in a more general manner. Ring mutations and the L63N plug mutation or the full plug deletion, which both generate a *prl* phenotype on their own, were not additive in suppression of the signal defect (**Figure 4**). The *prl* phenotype was always reduced, only mildly by alanines and most significantly by the charge mutations. This suggests that the effects of the combined mutations are not limited to facilitate pore opening.
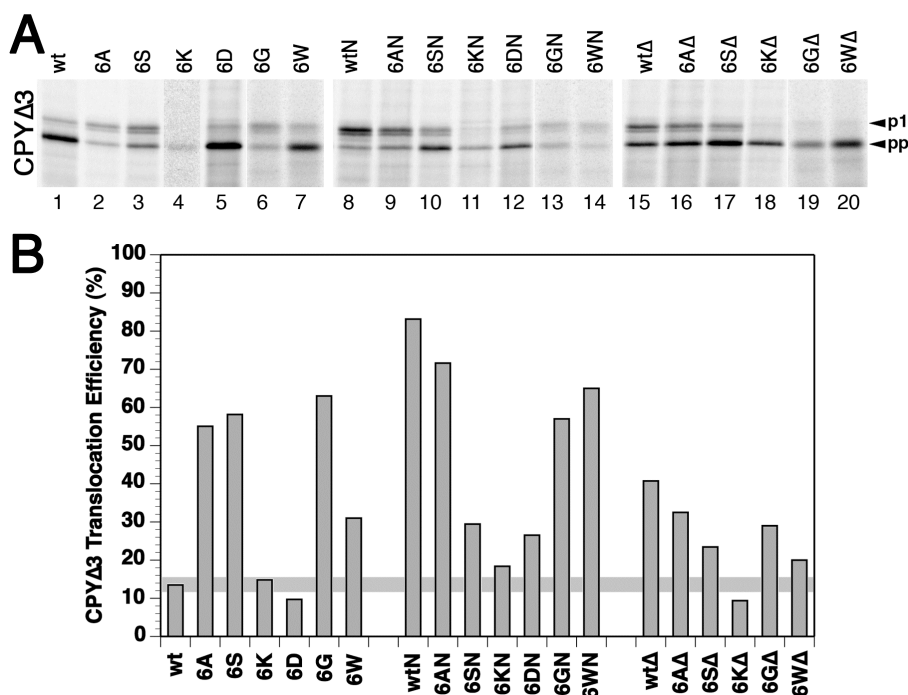
Figure 4. *prl* phenotype of mutant Sec61p. (A) CPYΔ3 (CPY with a signal sequence lacking three apolar residues) was expressed in Δ*ssh1*cells with wild-type (wt) or the indicated mutant Sec61p, labeled and analyzed as in Figure 3A. (B) Translocation efficiency was quantified by phosphorimager. The average of 1–3 determinations is shown. The horizontal line indicates the wild-type levels.

To test the stability of Sec61p mutants, their steady-state levels in cells lacking wild-type *SEC61* were analyzed by immunoblot analysis. Surprisingly, charged residues replacing the hydrophobic constriction residues did not significantly reduce protein levels in comparison to the respective wild-type translocons (wt, wtN, and wtΔ; **Figure 5A**). In contrast, in a heterozygous situation, when coexpressed with a wild-type copy of SEC61, several of the mutant translocons (tagged with an HA-epitope for independent detection) were observed at strongly reduced levels (**Figure 5B**). This phenomenon was previously observed for ΔTM2 (deletion of codons 77–107; (Wilkinson et al., 2000) and Δplug ((Junne et al., 2006) and **Figure 5B**, lane15). It indicates competition of wild-type and mutant Sec61p for limiting interaction partners that are required for stability. In support of this notion, overexpression of the β and γ subunits Sbh1p and Sss1p at least partially rescued Δplug and ΔTM2 (Junne et al., 2007). Reduced levels in a heterozygous situation therefore suggest an altered protein surface with reduced binding affinity to partner molecules. Most affected were the mutations to lysines, aspartates, and glycines, whereas mutations to serines and tryptophans caused minor effects. Only the mutations to alanines showed no reduction of protein levels compared to the corresponding wild-type version of the translocon.
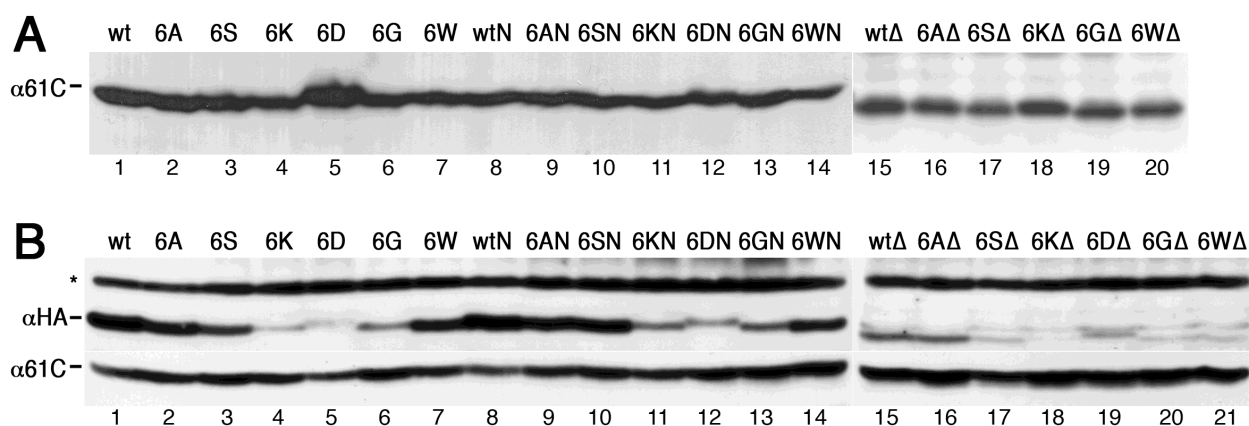
Figure 5. Levels of wild-type or mutant translocons in the absence (A) or presence (B) of a second wild-type copy of Sec61p. A: Steady-state amounts of wild-type and mutant Sec61p were determined in an *SSH1* background by immunoblot analysis of total cell lysate. Equal loading was approximated based on protein determination and Coomassie staining of SDS-gels. B: Yeast cells expressing equal amounts of wild-type Sec61p and the indicated HA-tagged mutants were analyzed by immunoblot analysis using an antiserum against the C terminus of Sec61p (α61C) and an anti-HA antibody (αHA) recognizing wild-type and mutant Sec61p, respectively. Mutation of constriction residues to aspartates consistently resulted in slightly reduced electrophoretic mobility. The asterisk indicates a background band recognized by the anti-HA antibody that also serves as a loading control. Data for *SSH1* cells are shown; the same result was obtained with Δ*ssh1* cells (unpublished data).

The properties of the constriction ring regulate membrane insertion

In addition to offering a passage for polypeptides through the membrane, the translocon also provides a lateral gate for the insertion of TM segments into the lipid bilayer. To analyze the effect of ring mutations on integration of TM sequences into the lipid bilayer, we tested the integration efficiency of moderately hydrophobic H-segments previously used by von Heijne and colleagues to characterize this process in mammalian *in vitro* and *in vivo* systems and in yeast (Hessa et al., 2005; Hessa et al., 2009). They consisted of a 19-alanine host segment in which an increasing number of residues were replaced by leucines (**Table I**), thus creating a series of increasing hydrophobicity. These sequences were inserted into the exoplasmic domain of DPAPB, generating a protein (DPAPB-H) with an uncleaved signal-anchor sequence for co-translational ER targeting and translocation of its C-terminus, and a potential stop-transfer sequence (as illustrated in **Figure 6A**). Depending upon whether this sequence is integrated into the membrane or is translocated, the protein is glycosylated only at sites between the signal-anchor and the H-sequence or also at downstream sites, respectively. The fraction of translocated to integrated H-segments can thus be determined after pulse-labeling, immunoprecipitation, gel electrophoresis, and autoradiography from the intensities of the fully and partially glycosylated forms corresponding to the translocated (T) and integrated H-segments (I), respectively. Unglycosylated

products (U) generated by some of the partially defective mutant translocons were clearly separated and ignored as irrelevant to the process of membrane integration of the H-segments.

Table I. Potential TM segments to test membrane integration behavior of Sec61p mutants.

DPAPB-H nX/(19-n)A	Potential TM sequence*
0L/19A	GGPG AAAAAAAAAAAAAAAAAAA GPGG
1L/18A	GGPG AAAAAAAAALAAAAAAAAA GPGG
2L/17A	GGPG AAAALAAAALAAAAAAAAA GPGG
3L/16A	GGPG AAAALAAAALAAAALAAAA GPGG
4L/15A	GGPG AAAALALAALAAAALAAAA GPGG
5L/14A	GGPG AAAALALAALAALALAAAA GPGG
6L/13A	GGPG AAAALALALALALALAAAA GPGG
1S/18A	GGPG AAAAAAAAASAAAAAAAAA GPGG
2S/17A	GGPG AAAASAAAASAAAAAAAAA GPGG
3S/16A	GGPG AAAASAAAASAAAASAAAA GPGG

* Essentially the same model sequences were used as had previously been described by Hessa et al. (2005; 2007; 2009), including flanking glycine/proline tetrapeptides to 'insulate' the central 19-residue stretch from the surrounding sequence. For simplicity of construction, 2S, 2L, and 4L guest residues were left asymmetric, since position dependence had been shown to be negligible (Hessa et al., 2007).

Oligo-alanine H-segments with no or one leucine were fully translocated by the wild-type translocon, whereas increasing membrane integration was observed with additional leucines (**Figure 6B**, and quantified as the integrated fraction in **Figure 7A**). 50% integration was obtained with ~4 leucines. Cells expressing the Δplug translocon dealt with the DPAPB-H constructs identically, indicating that the plug domain does not affect the outcome of the integration process. Similarly, the Q96R translocon, a Sec61p mutant altered at the cytoplasmic end of TM2 and affecting signal-anchor orientation (Junne et al., 2007), did not affect membrane integration. Also the replacement of the constriction ring by 6 tryptophans, which are quite hydrophobic, had no effect (**Figures 6B** and **7B**).

In contrast, exchange of the ring residues to hydrophilic amino acids, 6G, 6S, 6K, and 6D, clearly affected the insertion of H-segments (**Figures 6B**, and **7B** and **C**). The required number of leucines to allow 50% membrane insertion was reduced to ~2 with 6S, 6G, and 6K, and the 6D translocon already mediated more than 60% integration for a pure 19-alanine sequence. Insertion of 1–3 serines into the oligo-alanine H-segment (**Table I**) was necessary to prevent integration completely (**Figure 6B**, and **7B** and **C**). In addition, the transition from predominantly

translocated to mostly integrated H-segments occurred over a wide range of hydrophobicity for 6K and 6D, rather than within 3 leucines as for the other translocons. Surprisingly, mutation of the constriction residues to 6A had the opposite effects on H-domain insertion: it required 5 leucines for 50% insertion and the transition was completed in a range of only 2 additional leucines.
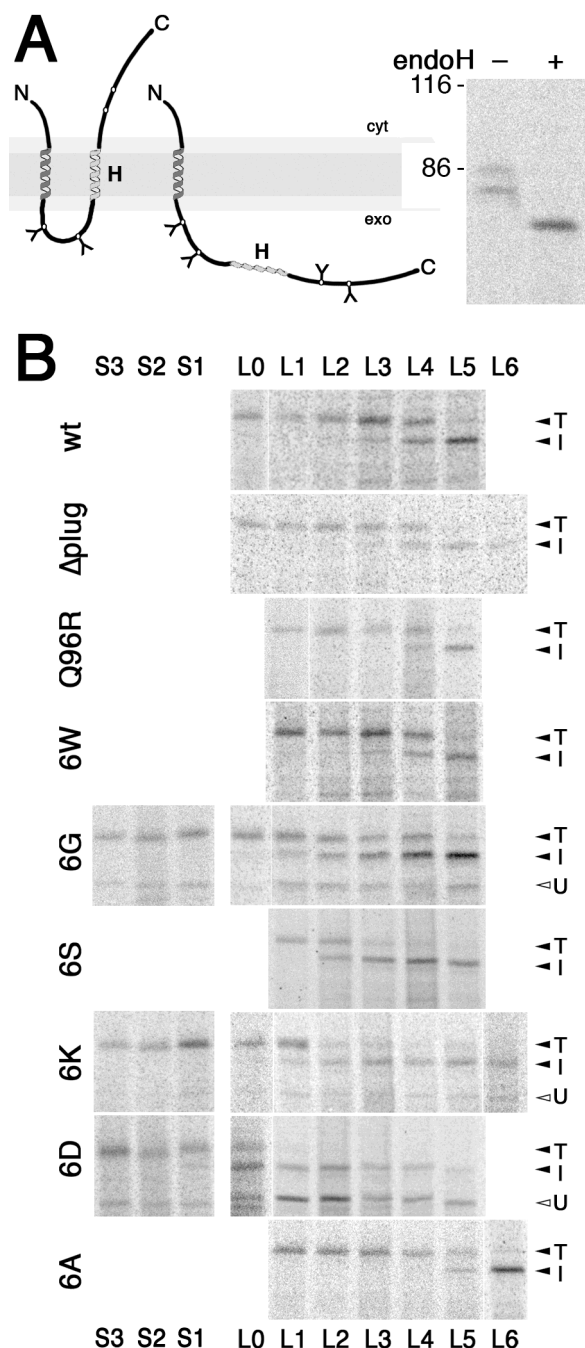

Figure 6. Membrane insertion of H-segments of various hydrophobicities mediated by wild-type and mutant Sec61p. (A) Schematic representation of the DPAPB-H model proteins (*left*) and products of the DPAPB-H substrate with 4 leucines expressed in cells with wild-type Sec61p after [^{35}S]methionine labeling, immunoprecipitation, incubation with (+) or without (–) endoglycosidase H (endoH), and gel electrophoresis (*right*). Integration of the H-segment results in a partially glycosylated double-spanning membrane protein (I), whereas its translocation yields a fully glycosylated type II protein (T). U, unglycosylated form, cyt, cytoplasmic; exo, exoplasmic. The position of molecular weight markers (in kD) is indicated.   (B) *SSH1* cells expressing wild-type or mutant translocons (as indicated on the left) as well as a DPAPB-H substrate (with the number of leucine or serine residues indicated above and below) were pulse-labeled with [^{35}S]methionine, and the substrate products were immunoprecipitated, separated by gel electrophoresis, and visualized by autoradiography.

If the results are interpreted as an equilibration between a membrane inserted and a free state (according to Hessa et al., 2005), the ratio of integrated to translocated fractions, $K_{app} = f_i/f_t$, i.e. the apparent equilibrium constant, can be used to calculate apparent free energies of membrane insertion as $\Delta G_{app} = -RT\ln K_{app}$ (where R is the gas constant and T the absolute temperature of 303 K). The resulting plots (shown in **Figure 7**, **A'–C'**) reveal a good linearity, consistent with the equilibrium assumption.
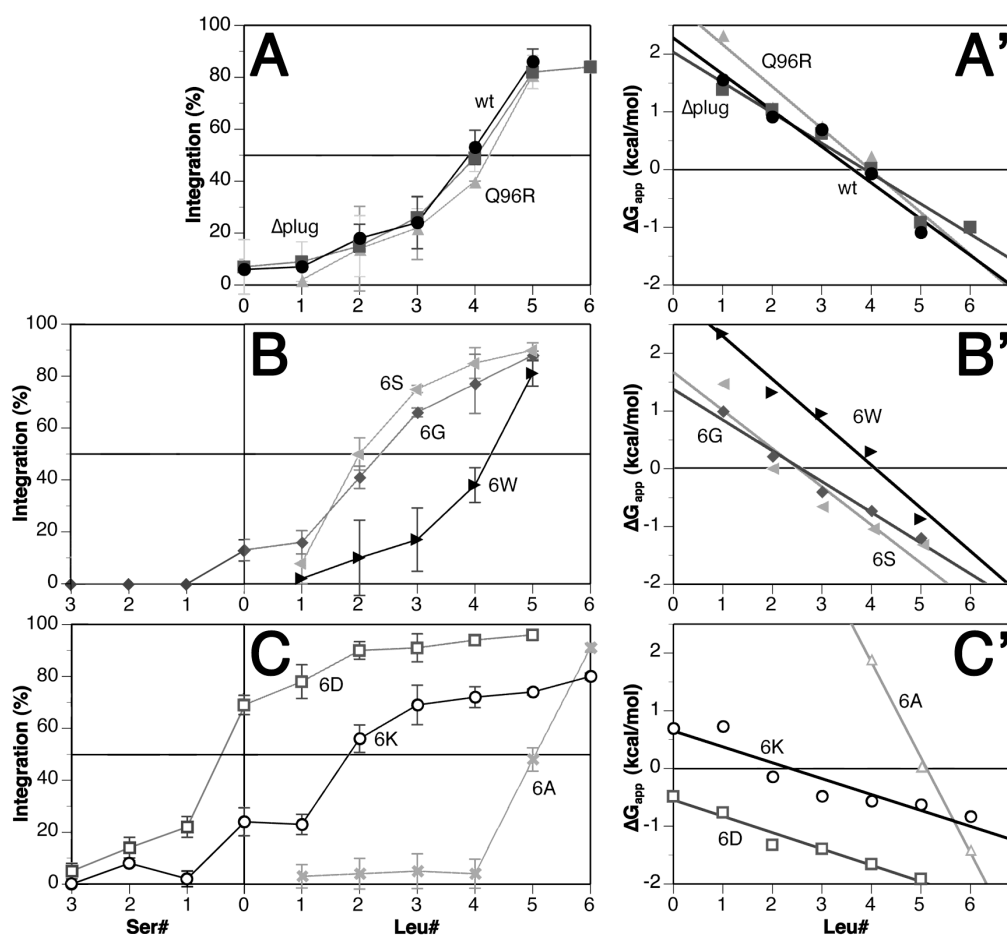

Figure 7. Efficiency of H-segment integration by wild-type and mutant Sec61p. (A–C) The membrane-inserted fraction of 2–5 experiments like those shown in Figure 6 was quantified and plotted (with standard deviations) vs. the number of leucines or serines in the H-segment. (A'–C') The data for the H-segments containing 0–6 leucines were also plotted as apparent free energies of membrane insertion, ΔG_{app} with straight lines determined by linear regression.

The number of leucines in the oligo-alanine host sequence necessary for 50% membrane insertion with each mutant translocon was interpolated from these plots and listed in **Table II**. These values and the apparent free energies of insertion ΔG_{app} are generally reduced with mutant translocons containing hydrophilic or charged ring residues. The amino acids lining the core of the

translocation pore thus clearly influence the hydrophobicity threshold for membrane insertion, most likely by defining the polarity of the environment of the substrate sequence within the pore, i.e. of one of the two compartments between which the H-segment is partitioning. Increased polarity in the more polar compartment is expected to favor membrane integration of a moderately hydrophobic sequence. This is what we observe for the 6G, 6S, 6K and 6D mutant translocons. The behavior of the 6A mutant, however, does not simply correlate with the increase in polarity in the ring residues and requires a different explanation.

Table II: Summary of membrane integration parameters of wild-type and mutant translocons.

Sec61	n for 50% membrane integration of nL/ (19–n)A*
wt	3.6
Δplug	3.8
Q96R	4.0
6W	4.1
6G	2.6
6S	2.5
6K	2.3
6D	-2.0
6A	5.1

*The number of leucines in the oligo-alanine host sequences (shown in **Table I**) for $\Delta G_{app}= 0$ was interpolated from the linear regressions shown in Figure 7 (A'–C').

DISCUSSION

Constriction ring mutants retain translocon functionality

The crystal structure of the idle translocon shows a pore that is closed by the lumenal plug domain and a central constriction formed by six hydrophobic residues. A likely function of this apolar constriction is thus to prevent or reduce ion permeability when the plug is out and during protein translocation. This is supported by the viability of yeast cells expressing plugless mutant translocons (Junne et al., 2006) and the fact that in *E. coli* expression of translocons with their plug locked open, although lethal, did not lead to an immediate membrane depolarization (Harris and Silhavy, 1999). Indeed, it has recently been shown that *prl* mutants of SecY with single hydrophobic-to-asparagine mutations in the constriction ring or with a deletion in the plug domain

caused ion conductance across inner membrane vesicles, but with strong selectivity for anions, particularly chloride (Dalal and Duong, 2009). This selectivity, which preserves the seal for protons, is also functional during protein translocation through the wild-type translocon.

Here, we have tested the effects of replacing all six constriction residues to various less hydrophobic, polar, and even charged amino acids. Only minor defects were detected for mutations to alanines, serines, tryptophans, or glycines, and significant functionality was retained even with charged residues (aspartates or lysines). Mutation to the uncharged amino acids, 6A, 6S, 6G, and 6W, produced a *prl* phenotype, indicating specific destabilization of the closed state and facilitated translocon opening. *Prl* mutations specifically disturb interactions that must be overcome for preprotein insertion. Therefore, not every ring mutation causes a *prl* phenotype (in Sec61p I86T does, but not T185K or M450K; Junne *et al.*, 2007) and multiple ring mutations do not necessarily have a stronger phenotype than single ones (6S showed less suppression than I86T). The charge mutations in 6D and 6K, but also the single mutations T185K or M450K (Junne et al., 2007), caused less specific perturbations and thus no *prl* effect. Similarly, by cumulation of *prl* mutations upon combining mutations in the ring residues and in the plug domain may generally destabilize the structure and thus not enhance or even reduce suppression of signal defects.

Since the constriction ring forms part of the plug binding site, ring mutations are expected to simultaneously disturb plug insertion. This could explain that an additional point mutation or even deletion of the plug did not strongly aggravate the phenotypes. Translocon stability was not significantly compromised by the ring mutations, not even by charged residues. Only in competition with a wild-type copy of Sec61p were the levels of mutant translocons clearly reduced in the order A<S≈W<G<D≈K, suggesting altered or less stable surface binding sites for limiting interaction partners. Again, this is a surprisingly mild effect for considerable alterations in the center of the protein. The ring mutations 6A, 6S, 6G, and 6W also did not significantly affect the orientiation of sensitive diagnostic constructs (as used previously in Junne *et al.*, 2007), and 6D and 6K did not have strong and opposite effects that could be correlated with their charges (unpublished results). It suggests that the 'positive-inside rule' is not dominated by charged residues in the core of the translocon.

The translocon core regulates membrane integration

It is long known that hydrophobicity is the essential property of a sequence for membrane integration (Davis and Model, 1985). Crosslinking experiments with reconstituted proteoliposomes demonstrated that the translocon allows a TM domain to bypass the barrier posed

by the polar head groups of the lipid bilayer and to come into contact with the hydrophobic interior of the membrane (Heinrich et al., 2000). It was proposed that Sec61 provides a site through which a TM domain can dynamically equilibrate between the lipid and aqueous phases, depending on its hydrophobicity. The systematic analyses by von Heijne and colleagues (Hessa et al., 2005; Hessa et al., 2009) yielded a 'biological hydrophobicity scale' for membrane integration consistent with a thermodynamic partitioning process. The observed position dependence of residues like tryptophan or tyrosine in the H-segment reflects the symmetry of the lipid bilayer (Hessa et al., 2007) and is also consistent with equilibration. Based on this interpretation, an apparent free energy contribution (ΔG^{aa}_{app}) for membrane integration could be calculated for each amino acid. Interestingly, the values, although generally similar, were different for different systems (mammalian ER *in vitro* and *in vivo*, yeast, bacteria, and biophysical measurements). From our measurements in yeast, we obtained ΔG^{Leu}_{app} = -0.51 kcal/mol, ΔG^{Ala}_{app} = 0.12 kcal/mol, and an interpolated 3.6 leucines in a 19-residues oligo-alanine sequence for 50% integration (compared to previous measurements in yeast of -0.21 kcal/mol, 0.06 kcal/mol, and 4.4 leucines, respectively (Hessa et al., 2009).

An alternative model to partitioning is a kinetically controlled mechanism in which H-segments trigger the opening of the lateral gate for (irreversible) exit into the lipid phase. Membrane integration would reflect the probability of gate opening, as supported by the molecular modelling study by Zhang (Zhang and Miller, 2010). If gate opening depends on the hydrophobicity of the H-segment, the result might be difficult to distinguish from that of a free equilibration mechanism. Here we found that mutations in the translocon affect TM integration, seemingly in support of a kinetic model, since a pure catalyst of partitioning should not affect the equilibrium. However, the translocon not only mediates the transition between two environments, but also defines the properties of one of them. The hydrophobic constriction ring provides an apolar core that has also been observed experimentally at the center of the SecYEG translocon of *E. coli* (Bol et al., 2007). In addition, the narrowness of the pore (particularly in the presence of a substrate) partially excludes water to create conditions less polar than in bulk solution. The choice for an H-segment is therefore not the lipid environment *vs.* the aqueous solution, but *vs.* a less polar pore environment. Constriction ring mutations change the conditions inside the pore and thus also the outcome of a partitioning process with the membrane. An increase in polarity inside the pore by replacing the hydrophobic constriction residues with polar amino acids reduces the hydrophobicity required for 50% membrane integration (**Table II**). Replacement with serines or glycines reduced it by at least one leucine, replacement with lysines by even more. The effect was most dramatic for aspartates that resulted in predominant membrane integration even of the

Ala_{19} H-segment. The reason is most likely that the charges on the short side chains of aspartates are more concentrated whereas those on the long lysine side chains are more delocalized. Tryptophans, as relatively hydrophobic amino acids and excluding water by their size, did not alter membrane integration significantly.

Interestingly, mutation of the ring residues to alanines had a clear effect of raising the hydrophobicity threshold for membrane integration to ~5 leucines. This is contrary to the expectation for introducing less hydrophobic ring residues. Of course, the effect of the mutations on the conformation of the translocon is not known. It is conceivable that the small side chains allow the pore to contract and thus to further exclude water, resulting in conditions corresponding to increased hydrophobicity. Glycines may not have this effect, because they introduce increased conformational flexibility.

It has been shown by cysteine crosslinking that the plug can move out of its binding cavity to reach SecE and that plug movement is triggered by polypeptide translocation (Tam et al., 2005). Molecular dynamics simulations, however, suggested that the plug is not necessarily fully displaced by a translocation polypeptide, but remains positioned either towards the lateral gate for a hydrophilic substrate or towards the inner side of the pore for a hydrophobic substrate (Zhang and Miller, 2010). In our experiments, full deletion of the plug did not affect H-segment integration, suggesting that the plug does not play an active role in regulating membrane integration.

Yet, our results clearly show that the properties of the residues forming the central constriction in the Sec61 translocon adjust the hydrophobicity threshold at which translocating sequences prefer the apolar environment of the lipid bilayer and thus stop further transfer to be anchored as TM domains.

ACKNOWLEDGMENTS

ADDITIONAL DATA

Translation rate affects integration consistent with the equilibration model

An alternative model to that of H-domains equilibrating between the pore and membrane environments is a kinetically controlled mechanism in which H-segments trigger the opening of the lateral gate. Membrane integration would reflect the probability of gate opening. If gate opening depended on the hydrophobicity of the H-segment, the result might be difficult to distinguish from that of a free equilibration mechanism. The effect of slowing down translation and thus increasing the time an H-segment spends in the translocon could distinguish between kinetic and equilibration mechanisms. Increased time at the translocon should increase the probability of membrane integration of an intermediately hydrophobic substrate, if kinetically controlled. The opposite should happen, if the substrate can equilibrate: the H-segment can only move forward while situated in the pore, but is arrested when in the membrane (**Figure 8A**). Increased translation time thus increases the probability of the equilibrating substrate to move forward and to be translocated.

To test this, a substrate integrating approximately 50% under normal conditions (a 4L H-segment for the wild-type translocon) was analyzed in the presence of the reversible elongation inhibitor cycloheximide at low concentrations (0.5 and 1 µg/ml) that reduce translation rate (five- and tenfold; as estimated by the incorporation of [^{35}S]methionine). As shown in **Figure 8B**, the fraction of integrated products was reduced with increasing translation time, consistent with the equilibration model. The same phenomenon was observed for the 5L substrate with the 6A translocon. The mechanism of membrane integration is thus not altered in this respect by mutation of the constriction ring residues to alanines, despite their different effects on the integration threshold.
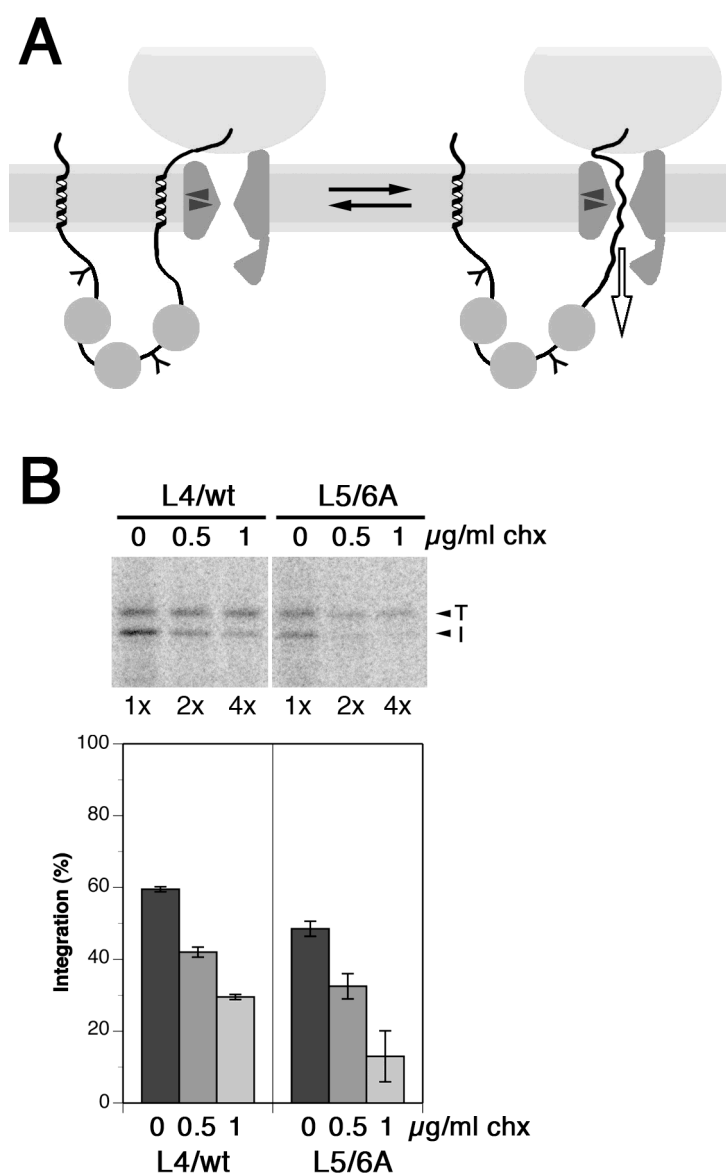
A

B

L4/wt   L5/6A

0 0.5 1   0 0.5 1 *µg/ml* chx

◄T
◄I

1x 2x 4x   1x 2x 4x

0 0.5 1   0 0.5 1 *µg/ml* chx
L4/wt   L5/6A

Integration (%)

Figure 8. Dependence of membrane integration on translation time. A: Schematic representation of the equilibration model. While translation is in progress, the H-segment is equilibrating between the membrane environment and the pore. Forward movement of the polypeptide and trapping by BiP (gray circles) binding can only occur while the substrate is in the pore (open arrow). Extending the time of translation by cycloheximide increases the probablility of translocation. B: Cells expressing wild-type or a mutant translocon as well as a DPAPB-H substrate with the indicated number of leucines were pulse-labeled with $[^{35}S]$methionine in the presence of 0, 0.5, or 1 µg/ml cycloheximide (chx), and the substrate products were immunoprecipitated. To compensate for the reduction in $[^{35}S]$methionine incorporation, the samples were upscaled two- or fourfold for gel electrophoresis and autoradiography of cycloheximide samples as indicated. Integration of the H-segment results in a partially glycosylated double-spanning membrane protein (I), whereas its translocation yields a fully glycosylated type II protein (T). The membrane inserted fraction of duplicate determinations was quantified.

The threshold for membrane integration of H-segments is similar in different eukaryotic organisms

An important question is whether translocons in different organisms have similar characterisitics for the recognition and integration of transmembrane segments localized within the polypeptide chain. In order to provide quantitative data of Sec61-mediated insertion of H-segments in other species, we transfered the analysis to the mammalian *in vivo* system and compared it with our results obtained *in vivo* in yeast and with the results of the group of von Heijne performed in *E. coli*, yeast, dog rough microsomes (RMs) and BHK cells (Hessa et al., 2005; Hessa et al., 2009; Xie et al., 2007).

Unfortunately, the model proteins based on DPAPB sequence and an H-segment of varying hydrophobicity used in yeast (**Figure 9E**, construct d; see also **Table I**) could not be expressed in mammalian cells, most likely due to unfavorable codon usage. Therefore, we inserted the sequences encoding the transmembrane segments containing 1-7 leucines together with flanking regions of ~20 amino acids into the exoplasmic domain of wild-type H1, a type II membrane protein. The resulting proteins, named H1-L1, H1-L2, etc., contained two glycosylation sites preceding the TM and one site following the hydrophobic segment (**Figure 9A**). COS-1 cells were transiently transfected with these constructs, the proteins were labeled with $[^{35}S]$methionine at 37°C, immunoprecipitated with an antibody directed against the C-terminus of H1 (αH1C), subjected to SDS-PAGE and autoradiography. The fraction of translocated to integrated TMs was determined from the intensities of thrice and twice glycosylated forms corresponding to the translocated (T) and integrated (I) H-segments, respectively. Glycosylation status was confirmed by endoH digestion (**Figure 9B**).

Glycosylation analysis revealed a gradual increase of membrane integration of H-segments with their increasing hydrophobicity (~17% integration for H1-L1 up to ~77% for H1-L7) (**Figure 9C**). The apparent free energies of membrane insertion $\Delta G_{app} = -RT\ln K_{app}$ (where R is the gas constant and T the absolute temperature of 310 K) were calculated for each construct. The resulting plot (shown in **Figure 9D**) revealed a good linearity, consistent with the equilibrium assumption. 50% membrane integration was obtained with ~4.8 leucines in the H-segment. This result is different from our yeast system, where the number of leucines required for $\Delta G_{app}= 0$ kcal/mol was ~3.6 for wild-type translocon (**Figure 9F**, blue line). The data obtained by the group of von Heijne for the same H-segments in *E. coli*, yeast, mammalian RMs and in BHK cells are shown in comparison in **Figure 9F** (purple, light blue and green lines, respectively) and their characteristics are listed in **Table III** (Hessa et al., 2005; Hessa et al., 2009; Xie et al., 2007). The slope of the ΔG_{app} as a function of the number of leucines in the H-segment, corresponds to the $\Delta\Delta G_{app}$ value for an Ala→Leu replacement in the H-segment. Assuming a simple additive model for the contributions of Leu and Ala residues to the apparent free energy of insertion, ΔG_{app}^{Leu} and ΔG_{app}^{Ala} were calculated and presented in **Table III**.
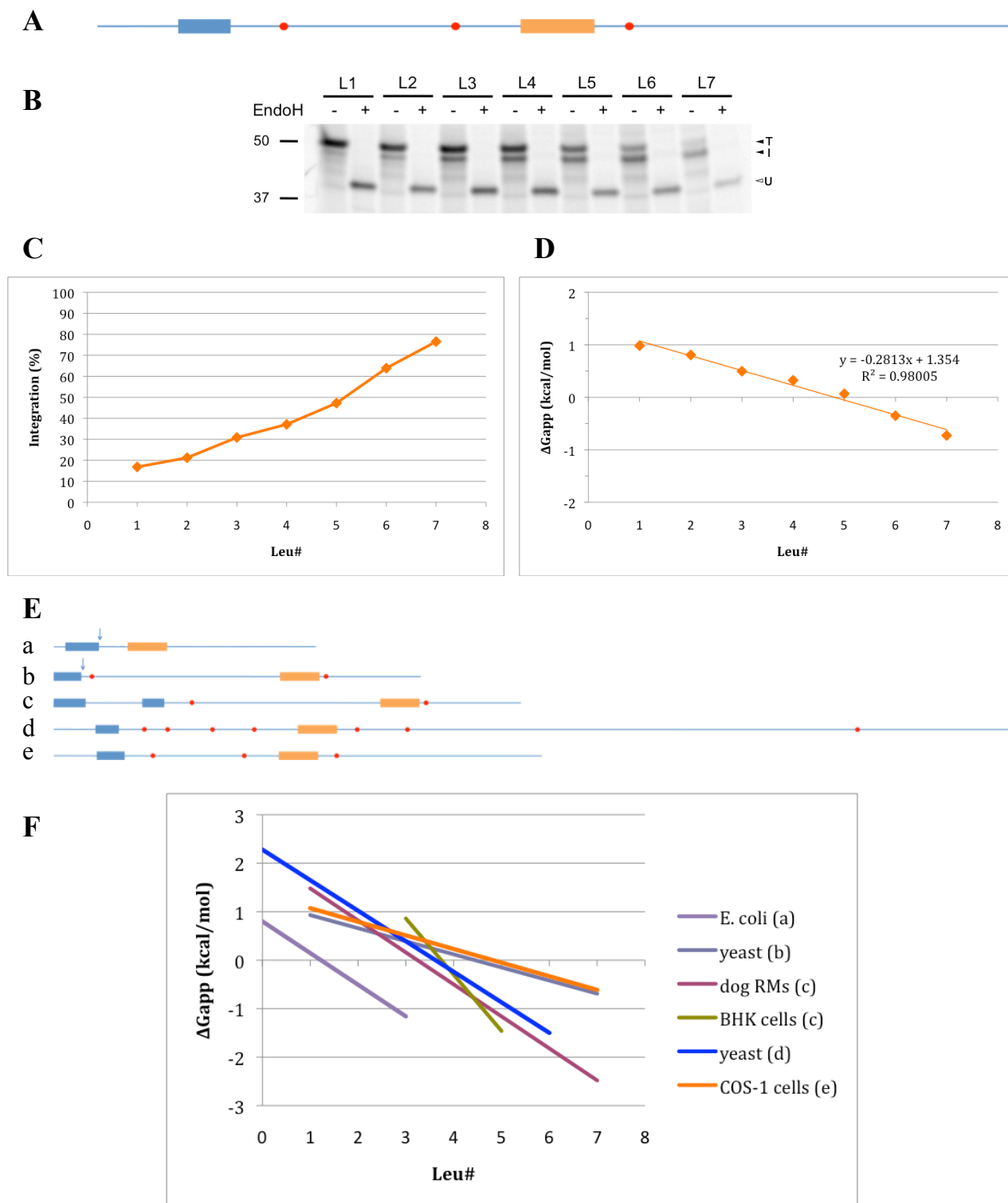
Figure 9. Membrane-insertion efficiency of Leu/Ala H-segments composed of nL/(19-n)A in various organisms. (A) Construction of H1-L1 − L7 proteins used in COS-1 cells. The blue rectangle corresponds to the natural TM segment, the orange rectangle represents the H-segment, red points mark the glycosylation sites. (B) COS-1 cells were transiently transfected with H1-L1 − L7 constructs, labeled with [35S]methionine, immunoprecipitated, subjected to deglycosylation with endoglycosidase H (endoH), electrophoresis and autoradiography. T-thrice glycosylated products corresponding to polypeptides with translocated H-segment, I-twice glycosylated forms representing proteins with integrated H-segment, U-unglycosylated form corresponding to proteins not targeted to the ER or deglycosylated with endoH. The marker sizes are in kDa. (C) The membrane-inserted fraction of a single experiment was quantified and

plotted vs. the number of leucines in the H-segment. (D) The same data plotted as apparent free energies of membrane insertion, ΔG_{app}, with a straight line determined by linear regression. (E) Model proteins used in the insertion studies in various organisms: a-PCLep (*E. coli*), b-Suc2p-Lep (yeast; von Heijne), c-SP-Lep (dog RMs, BHK cells), d-DPAPH-H (our yeast system), e-H1-L1 – L7 (COS-1 cells). Construct elements are marked as in A, with arrows indicating signal peptide cleavage sites. (D) Efficiency of membrane insertion of Leu/Ala-based H-segments. The data obtained from the indicated organism using the construct given in parentheses were plotted as apparent free energies of membrane insertion, ΔG_{app}, as straight lines determined by linear regression.

Experiments performed by the group of von Heijne in *E. coli*, yeast, dog RMs and BHK cells required different substrate proteins in order to achieve good expression levels and proper topology in the membrane. Xie et al. (2007) studied the insertion of H-segments based on the M13 procoat (PC) protein, which is composed of an N-terminal cleavable signal peptide and a C-terminal transmembrane helix. The TM was replaced by the P2 globular domain of leader peptidase (Lep) and H-segments containing 0-3 leucines were inserted. The resulting constructs were called PCLep (**Figure 9E**; a) and were unique amongst other model proteins as they were inserted by the YidC translocon instead of SecYEG. YidC functions in the insertion and assembly of a subset of proteins into the bacterial inner membrane. It can function on its own, or together with the SecYEG complex. YidC is a 60 kDa protein with six transmembrane segments, five of them are conserved among the YidC family of proteins and may provide a platform for binding the hydrophobic regions of the substrates (Luirink et al., 2001). The best-studied substrates that are inserted exclusively by YidC are the M13 procoat and the Pf3 coat proteins (Kiefer and Kuhn, 2007). Integration of H-segments contained within a protein carrying the M13 procoat protein signal is presented in **Figure 9F** (purple line). The studies were based on a protease protection assay instead of a glycosylation assay (Xie et al., 2007).

For analysis of H-segment integration in mammalian systems (*in vitro* in dog RMs and *in vivo* in BHK cells) constructs based on the sequence of leader peptidase were used (SPLep; **Figure 9E**, construct c) (Hessa et al., 2005). In *S. cerevisae* this protein did not attain a unique topology, therefore the two N-terminal TMs in Lep were replaced by the cleavable signal peptide of the yeast secretory protein Suc2p (Suc2p-Lep; **Figure 9E**, construct b) (Hessa et al., 2009).

Taken together, the data obtained by our group and by the group of von Heijne indicated that the overall relation between the Leu/Ala ratio in H-segments composed of nL(19-n)A and the apparent free energy of membrane insertion (ΔG_{app}) is similar in different eukaryotic organisms, with 50% membrane integration (ΔG_{app}=0 kcal/mol) observed for n=3-5 (**Table III**). The *E. coli*

system with a much lower threshold for membrane integration of H-sements is an exception, which can be easily explained by the involvement of a different translocon (YidC). The main difference between the eukaryotic systems tested was the energetic cost of Ala→Leu replacement in the H-segment, $\Delta\Delta G_{app}^{Ala\rightarrow Leu}$, which corresponds to the slope of the ΔG_{app} as a function of the number of leucines. In our experiments, the value obtained for yeast was ~2-fold smaller compared with COS-1 cells. In contrast, $\Delta\Delta G_{app}^{Ala\rightarrow Leu}$ measured by Hessa et al. was lowest for BHK cells, followed by dog RMs and yeast, with ~2-fold difference between each system. Thus, the membrane integration in BHK cells appeared to be the most sensitive to changes in hydrophobicity of the H-segment and each additional leucine resulted in the highest increase of membrane insertion of the transmembrane segment.

Table III. Integration efficiency of H-segments in various systems.

System	Model protein (Figure 9E)	no. of Leu for ΔG_{app}=0	$\Delta\Delta G_{app}^{Ala\rightarrow Leu}$ (kcal/mol)	ΔG_{app}^{Leu} (kcal/mol)	ΔG_{app}^{Ala} (kcal/mol)	Reference
E. coli	a	1.2	-0.65	-0.61	0.04	(Xie et al., 2007)
yeast	b	4.4	-0.27	-0.21	0.06	(Hessa et al., 2009)
dog RMs	c	3.2	-0.66	-0.54	0.11	(Hessa et al., 2005)
BHK cells	c	3.7	-1.16	-0.93	0.22	(Hessa et al., 2005)
yeast	d	3.6	-0.63	-0.51	0.12	this study
COS-1 cells	e	4.8	-0.28	-0.21	0.07	this study

There are several possible explanations for the discrepancies between different eukaryotic organisms and different experiments with the same model organism:

a) lipid and protein composition of the ER membrane: cholesterol content affects the fluidity of the membrane. Protein content of membranes influences the solvation state of transmembrane helices. Differences in these two parameters may play a role during the lateral release of TMs by the translocon.

b) pore sequence: differences in the structure of the Sec61 complex between different organisms may affect the integration of potential transmembrane segments. In particular, the composition of Sec61 ring residues has an influence on the threshold of membrane integration (**Figure 7**). The ring domain of *S. cerevisae*, composed of V82, I86, I181,

T185, M294, and M450, is slightly less hydrophobic than in other organisms, which could result in a higher ratio of membrane insertion.

c) labeling temperature: the polypeptide moves through the channel via Brownian motion, therefore the temperature may affect the equilibration between the lipid phase and the pore. In our experiments, the labeling was performed at 30°C for yeast and 37°C in case of COS-1 cells. The labeling temperatures used by the group of von Heijne were: RT for yeast, 30°C for dog RMs and 37°C for BHK cells. Experiments with yeast were performed at ambient temperature and thus, might not be very well controled and in addition such conditions were suboptimal for yeast. For ΔG_{app} calculations a fixed value of 25°C was used.

d) translation rate: it affects the time the polypeptide spends inside the translocation channel and is available for equilibration or for binding by lumenal chaperones that facilitate transfer. The translation rate is ~5 aa/s in cultured mammalian cells (Hershey, 1991) and ~10 aa/s in yeast. As shown in **Figure 8B**, extension of the translation time with cycloheximide in yeast led to a decreased ratio of insertion of the H-segments.

e) variations caused by different conditions of the same type of experiment: different preparations of dog RMs resulted in changes in H-segment insertion efficiency (Hessa et al., 2005)

f) host proteins carrying the H-segments: the model protein used in studies of insertion into the mammalian ER by Hessa et al. (2005; 2009) was the *E. coli* inner-membrane protein leader peptidase (Lep). In *S. cerevisae* the two natural TMs in Lep were replaced by the cleavable signal peptide of Suc2p. As a control, both model proteins were expressed *in vitro* in dog RMs and it proved that the difference between the mammalian and yeast system was not due to the use of Suc2p signal peptide in the latter. We also experienced a problem with transfering the analysis from yeast to mammals: the model proteins based on the yeast DPAPB sequence could not be expressed in COS-1 cells. Therefore, the H-segment was removed and transfered into the C-terminus of H1, along with flanking regions of ~20 amino acids, which should have prevented an influence of the neighbouring sequence on the insertion of the H-segments.

g) possible use of the post-translation translocation pathway in *Saccharomyces cerevisae*

MATERIALS AND METHODS

Cloning strategy

<u>H1-L1 – H1-L7</u>

The fragments of DPAPB-H containing the H-segments with one to seven leucines were cut out and replaced into the exoplasmic domain of wild-type H1. First, the sequence of H1 was modified in order to introduce a ClaI site. This was achieved by PCR of the fragment from the vector to the natural BstXI site within the C-terminus of H1 with ECEleft and H1DPAPBstopnew-a primers (the primers are listed in **Table IV**). At the same time, the H-segments from DPAPB were amplified by PCR along with the flanking regions and an N-terminal ClaI and a C-terminal BstXI site (primers called DPAPBstopins-s and DPAPBstopins-a). The ClaI site turned out to be Dam-methylated. In order to enable further cloning, the plasmids were amplified in a bacterial strain deficient in Dam methylase, GM48. The ClaI-BstXI fragment was inserted into the modified H1. The last glycosylation sites in the H1 sequence, located after the BstXI site was destroyed by exchanging the serine of the consenseus N-X-S sequence into alanine. The mutagenesis was performed by PCR with H1BstxStoA-s primer and H1XbaEco-a primer. The latter introduced an XbaI site before the EcoRI site, which made possible recloning the whole sequence into pRSV vector.

All constructs were verified by sequencing with ECEleft2.

Table IV. PCR primers.

Primer name	Primer sequence	Primer type
DPAPBstopins-a	GCGCCATTGCCCTGGCTATCAAATGTTTCATTTGC	sense
DPAPBstopins-s	CGCATCGATCTGAAGCGGTTATTAATTAG	sense
ECEleft	GAAGTAGTGAGGAGGC	sense
ECEleft2	CGGCCTCTGAGCTATTCCAG	sense
H1BstXStoA-s	GCGCCAGGGCAATGGCGCAGAAAGGACCTGC	sense
H1DPAPBstopnew-a	GAGCCATTGCCCTGGAGCGCATCGATCATTGTTATATTCATCTGACAGCTCAGGC	antisense
H1XbaEco-a	CGCGAATTCTAGATTAAAGGAGAGGTGGCTC	antisense

Cell culture

COS-1 cells were cultivated in DMEM (Sigma) supplemented with 10% FCS, 100 units/ml penicillin, 100 μg/ml streptomycin, 2 mM L-glutamine, in a humidified incubator containing 7.5% CO_2 at 37°C.

Transient transfections

For transient transfections, cells were split 1:10 into 60 mm dishes and transfected the next day with polyethylenimine (PEI). Cells were processed 2 days after transfection.

Metabolic labeling with ^{35}S-methionine

Transfected cells were incubated for 30 min in starvation medium (DMEM without methionine and cysteine, containing 2 mM L-glutamine; Sigma). Cells were labeled for 40 min with 100 µCi/ml [^{35}S]protein labeling mix consisting of 77% methionine and 23% cysteine (PerkinElmer), then transfered to 4°C and washed twice with cold PBS.

Immunoprecipitation

Cells were lysed in 1 ml of lysis buffer (PBS, 1% TritonX-100, 0.5% deoxycholate, 2 mM PMSF, 1x PIC) for 1 h, then scraped, vortexed and incubated for 30 min on ice. Lysates were cleared by centrifugation and proteins were immunoprecipitated overnight by addition of 2 µl rabbit anti-serum raised against a peptide corresponding to residues at the C-terminal part of H1 (anti-H1C). The immune complexes were pulled down by incubation with 15 µl protein A-Sepharose (Zymed) for 1 h. Samples were washed 4 times with immuno-wash (lysis buffer, 1 mg/ml BSA, 1 mM PMSF) and 2 times with PBS (with 1 mM PMSF).

EndoH treatment

At the last PBS wash 20% of each sample was transferred into a new tube and subjected to deglycosylation with endoglycosidase H. All samples were boiled in endoH buffer (50 mM sodium citrate, 1% SDS; in PBS), 1 µl of the enzyme (Roche) was added and samples were incubated for 3 h at 37°C.

SDS-PAGE and autoradiography

Reduced samples were resolved on a tricine gel for low-molecular weight proteins (1-70 kDa) and quantified using a phosphorimager (Molecular Dynamics Inc.). The fraction of translocated to integrated TMs was determined from the intensities of thrice and twice glycosylated forms corresponding to the translocated and integrated H-segments, respectively.

III General discussion

Topogenesis of membrane proteins is coordinated at the endoplasmic reticulum by the conserved heterotrimeric Sec61 complex, which facilitates the lateral release of transmembrane segments (TMs) into the lipid bilayer during polypeptide translocation. In Part I of this thesis we analyzed the insertion process of single-spanning membrane proteins containing a signal-anchor sequence localized at the N-terminus or internally. We showed that insertion of these two types of signals proceeds via different mechanisms. N-terminal signals insert into the translocon head first and can invert their orientation over time in order to enable translocation of the C-terminal domain. The hallmarks of inversion are C-terminal length dependence, which determines the time the polypeptide spends inside the translocation channel and is available for reorientation, and sensitivity to increased signal hydrophobcity, which affects the signal's ability to dissociate from the apolar binding site prior to inversion (Goder and Spiess, 2003). Upon transition from N-terminal to internal signal-anchors, these two properties are lost, which indicates that proteins with an internal signal-anchor and a non-folding N-terminal domain, acting as a steric hindrance, are unable to reorient upon insertion into the Sec61 pore. Based on these results we proposed a model for signal-anchors` orientation in the channel, where N-terminal signals can enter the pore and be laterally released in an N_{exo}/C_{cyt} topology without full pore opening (Discussion section of Part I of this thesis). In contrast, internal signals insert in an orientation that favours either N- or C-translocation and do not invert within the channel.

The mechanism of insertion of N-terminal signal-anchors was first postulated by Rapoport et al., who suggested that proteins with a long hydrophobic TM and an N-terminus that is not retained in the cytosol would be rapidly released into the membrane. The more hydrophobic the TM, the greater would be its tendency to partition immediately into the lipid phase. The plug domain would be displaced only transiently (Rapoport et al., 2004). Evidence that such a mechanism is possible has just emerged from the crystal structure of SecYEß from the hyperthermophile archaeon *Pyrococcus furiosus* (Egea and Stroud, 2010). In the crystals, two adjacent SecYE complexes face each other through the cytoplasmic loops of SecY. The C-terminal helix (TM10-αC) of one molecule is inserted into the cytoplasmic vestibule of another SecY subunit, mimicking a nascent chain. Entry of this pseudo-protein substrate is accompanied by widening of the cytoplasmic half of the channel and a complete opening of the lateral gate defined by TMs 2/3 and 7/8. The hydrophobic seal provided by the ring region is compromised on lateral gate opening, but the plug domain still occludes the channel. Based on the X-ray structure

and mutagenesis, the authors suggested that the C-terminal helix of SecY acts as a sensor of and a guiding sequence for the incoming nascent chain and relays conformational changes to the lateral gate. Docking of the nascent chain mimic releases the clamping effect that SecE exerts on SecY, which leads to lateral gate opening. As the substrate progresses farther, the ring loosens to accommodate the incoming peptide, but the plug still maintains a central seal, preventing ion leakage. The crevice generated by lateral gate opening is about 11 Å wide and is likely to allow interaction of the signal sequence with host phospholipids. It could also enable reorientation of TM segments in a relatively protected environment and acquisition of tertiary structure. The closed state of the translocon is stabilized by hydrogen bonds between highly conserved amino acids. Perturbations caused by mutations at various locations are transferred to the plug (Bondar et al., 2010), leading to its displacement by increased hydration. The structure of the archaeal translocon obtained by Egea and Stroud (2010) suggests that the lateral gate is exquisitely sensitive to the presence of a nascent chain. It is possible that it is continuously opened during the polypeptide`s transfer through the channel, allowing contact with lipids, which could influence the process.

The process of insertion of N-terminal and internal signal-anchors depends on the flanking charges which position the signal according to the `positive-inside rule`. Inversion of flanking charges had a dramatic effect on N-terminal signals, blocking their ability to reorient and to translocate their C-terminus. During topogenesis of internal signal-anchors with type III charge distribution we could observe a competing effect of the hydrophilic N-domain, acting as a steric hindrance, preventing N-translocation, and charged residues flanking the TM, promoting the opposite topology. Such proteins showed a discontinuous C-terminal length dependence (**Figure 5** in Part I). Surprisingly, for proteins with the the shortest C-terminal tail we obtained a higher ratio of C-translocation compared with the polypeptides with a longer tail within a series. A potential explanation for this behaviour could be the competing effect of the N-terminal and C-terminal domains during SRP targeting to the translocon. When the short C-terminus is completed by the time SRP has bound the signal-anchor and transferred it to the Sec61 complex, it could pose a weaker steric hindrance effect than the longer and flexible N-terminal domain and thus, be preferentially translocated into the ER lumen. An interesting task for the future will be to study the insertion of internal-signal anchors with no flanking charges, but with different sizes of the N- and C-terminal domains. Since such proteins cannot be positioned in the translocon according to the `positive-inside rule`, their topology should depend solely on the steric effects of the N- or C-terminal sequence.

As the basic mechanism of insertion of membrane proteins into the endoplasmic reticulum is understood in outline, the attention focused on the quantitative relationships between the amino acid composition of transmembrane segments and their membrane-insertion efficiency. Up to now, over 130 hydropathy scales have been described (Palliser and Parry, 2001). Many of them are used for prediction of TM insertion by relying, in part, on the concept of threshold hydrophobicity, i.e., the level at which the intrinsic hydrophobicity of a TM will trigger its transfer from the hydrophilic environment into the membrane (Deber et al., 2001; Kyte and Doolittle, 1982; von Heijne, 1992). Recent work by Hessa et al. has made a significant contribution to our understanding of the insertion process, in particular by establishing the `biological`, amino acid-specific free energy scale for TM helix integration (Hessa et al., 2005; Hessa et al., 2007; Hessa et al., 2009). By employing *in vitro* transcription-translation of model proteins into dog rough microsomes (or *in vivo* expression in BHK cells), translocon-mediated helix insertion of potential TMs could be monitored as a function of amino acid sequence. In addition, the effects of overall hydrophobicity, length, flanking sequences and orientation of the H-segments on membrane insertion were studied. A subset of these H-segments was also analyzed in *E. coli* using a different model protein, inserted via YidC translocon, in order to study whether different kinds of translocons, and Sec61 translocons in different organisms have similar characteristics for the recognition and integration of TM segments.

The data on transmembrane segments insertion obtained by Hessa together with the experiments we performed *in vivo* in yeast and mammalian (COS-1 cells) systems were presented as additional material to Part III of this thesis. The data are consistent with the model of thermodynamic partitioning, where the Sec61 translocon provides a site through which a TM segment can dynamically equilibrate between the lipid phase and the hydrophilic interior of the pore, depending on its hydrophobicity. However, different results obtained in different organisms using transmembrane helices of the same composition indicate that, although the thermodynamic model provides a simple explanation of the insertion mechanism, TM integration as a whole is not a simple partitioning process. In all eukaryotic systems tested, the hydrophobicity threshold for 50% membrane insertion was similar, with n=3-5 leucines in H-segments composed of nL(19-n)A. However, various organisms appear to differ significanty in the energetic cost of alanine to leucine replacement in the TM helix.

Upon insertion of a potential transmembrane segment into the translocation channel, it does not remain stationary, but instead, it is subjected to various factors that affect its partitioning between the aqueous and lipid phase and influence the final decision whether it is laterally released into the membrane or translocated into the ER lumen. The latter is facilitated by the

lumenal chaperone BiP or its yeast homolog, Kar2p, members of the Hsp70 family of heat shock proteins, which contain a conserved N-terminal nucleotide-binding domain (NBD), a substrate binding domain (SBD), and a variable C-terminal domain (Bukau and Horwich, 1998). When BiP is bound to ATP, it is in the open conformation, and ATP hydrolysis to ADP causes a tight binding of BiP to its substrate. BiP binds transiently to exposed hydrophobic regions approximately seven amino acids in length, where aromatic and hydrophobic residues occur in alternating positions. Blond-Elguindi et al. developed a computer algorithm to predict the inteaction between BiP and any 7-aa peptide (Blond-Elguindi et al., 1993). The stretches recognized and bound by BiP are predicted to exist quite frequently in protein sequences (about every 36 amino acids) (Rudiger et al., 1997).

The action of BiP can affect the equilibration process in several ways. The substrate protein is available for forward movement only when it is in the translocation channel, but is arrested when in the membrane. Thus, BiP can facilitate polypeptide translocation into the ER upon its removal from the translocon pool only. This process is influenced by the time the polypeptide spends inside the pore. Long proteins which take a long time to synthesize offer a bigger time window for BiP binding. A similar effect has translation attenuation with the reversible elongation inhibitor cycloheximide (see **Figure 8** in Part III). TM insertion analysis in different systems using the same model proteins can lead to incompatibility problems, which we also encountered. Due to codon usage, some proteins cannot be efficiently expressed in different organisms. Rare codons could stall polypeptide elongation and thus increase the time it is located inside the pore. In contrast to the effect caused by cycloheximide, the pauses in translation would be irregular and difficult to control. In addition, the chains of different model proteins could potentially have different affinities for BiP binding.

The process of H-segment equilibration could be also affected by the flanking sequences. Although the TM helices were insulated from the surrounding environment by a hydrophilic GPGG sequence, we do not know the effect of charged residues in the vicinity of the TM and the folding properties of the loop regions on TM motion and BiP binding. All of these factors, together with those discussed in the appendix section of Part III might explain the observed system-dependent differences in the insertion pattern of different model proteins carrying identical H-segments.

While for single-spanning membrane proteins the lateral transfer of TM segments from the translocon into the lipid phase can be correlated with the thermodynamic properties of the segment residues, in the case of polytopic membrane proteins this model is probably an oversimplification.

Several additional factors, such as interactions between individial TM segments and membrane protein assembly, have yet to be fully accounted for (Cross and High, 2009). Experimental data suggested that up to four TMs may be present in the translocation pore at the same time (Enquist et al., 2009; Tu et al., 2000). An interesting phenomenon occuring during polytopic proteins biosynthesis is selective TM-segment retention at the ER translocon site (Cross and High, 2009; Ismail et al., 2008; Pitonzo et al., 2009; Sadlish et al., 2005), which could serve several purposes. TM retention might facilitate the formation of inter-TM contacts necessary for the correct folding of the protein (Pitonzo et al., 2009). In case of membrane transporters, which often contain charged or polar residues in their transmembrane segments, the lateral release of TMs one-by-one would be energetically unfavorable. During assembly in the Sec61 complex, the helices of polytopic proteins can interact with each other via hydrogen bonding involving asparagine, glutamine, aspartic acid and glutamic acid (Meindl-Beinker et al., 2006) or salt bridges (Ciczora et al., 2005; Ciczora et al., 2007).

In Part II we analyzed the insertion process of proteins with two signals, which could represent the simplest polytopic protein. The signal sequences carried a conflicting topogenic information (type I signal of HA vs. type II signal-anchor of H1), and thus, competed for the preferred orientation in the translocon. With wild-type forms of both signals, the signal-anchor exerted its topogenic information immediately after its insertion into the pore, which resulted in a rapid inversion of the protein. The signal-anchor was only moderately sensitive to increased signal hydrophobicity. This would suggest that internal signal-anchors, contained within two-signal proteins, are potentially not yet in contact with the lipid phase when reorientation occurs and they do not equilibrate between the apolar and polar phases. To investigate it, one could perform experiments to assess if the TM contacts the lipid phase during its stay at the translocon, such as chemical cross-linking (McCormick et al., 2003) or by placing bulky residues at the N- or C-terminus of the signal (Higy et al., 2005).

The mechanism of protein translocation into the ER is reasonably well understood in general terms, but several issues require further investigation. It is clear that the Sec61 complex provides a polar path for the entry of soluble polypeptides into the ER lumen, as well as a lateral gate for transmembrane helix integration (Zimmermann et al., 2010). The hydrophobic core of the translocon participates in recognition of transmembrane segments and defines the threshold for membrane integration. Different precursor polypeptides require a different protein transport machinery, manifested by the involvement of additional components, such as BiP and Sec63, TRAM, TRAP or cytosolic chaperones. The precise mechanism of action of these auxiliary

components remains to be elucidated. Another unsolved issue is the origin of the `positive-inside rule` and the involvement of lipids in membrane-protein topogenesis.

IV References

Aebi, M., Bernasconi, R., Clerc, S., and Molinari, M. (2010). N-glycan structures: recognition and processing in the ER. Trends Biochem Sci *35*, 74-82.

Alberts, B. (2008). Molecular biology of the cell, 5th edn (New York, Garland Science).

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

Becker, T., Bhushan, S., Jarasch, A., Armache, J.P., Funes, S., Jossinet, F., Gumbart, J., Mielke, T., Berninghausen, O., Schulten, K.*, et al.* (2009). Structure of monomeric yeast and mammalian Sec61 complexes interacting with the translating ribosome. Science *326*, 1369-1373.

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.

Berndt, U., Oellerer, S., Zhang, Y., Johnson, A.E., and Rospert, S. (2009). A signal-anchor sequence stimulates signal recognition particle binding to ribosomes from inside the exit tunnel. Proc Natl Acad Sci U S A *106*, 1398-1403.

Bibi, E., Stearns, S.M., and Kaback, H.R. (1992). The N-terminal 22 amino acid residues in the lactose permease of Escherichia coli are not obligatory for membrane insertion or transport activity. Proc Natl Acad Sci U S A *89*, 3180-3184.

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

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

Bogdanov, M., Heacock, P.N., and Dowhan, W. (2002). A polytopic membrane protein displays a reversible topology dependent on membrane lipid composition. EMBO J *21*, 2107-2116.

Bogdanov, M., Xie, J., Heacock, P., and Dowhan, W. (2008). To flip or not to flip: lipid-protein charge interactions are a determinant of final membrane protein topology. J Cell Biol *182*, 925-935.

Bol, R., de Wit, J.G., and Driessen, A.J. (2007). The active protein-conducting channel of Escherichia coli contains an apolar patch. J Biol Chem *282*, 29785-29793.

Bondar, A.N., del Val, C., Freites, J.A., Tobias, D.J., and White, S.H. (2010). Dynamics of SecY translocons with translocation-defective mutations. Structure *18*, 847-857.

Bornemann, T., Jockel, J., Rodnina, M.V., and Wintermeyer, W. (2008). Signal sequence-independent membrane targeting of ribosomes containing short nascent peptides within the exit tunnel. Nat Struct Mol Biol *15*, 494-499.

Boy, D., and Koch, H.G. (2009). Visualization of distinct entities of the SecYEG translocon during translocation and integration of bacterial proteins. Mol Biol Cell *20*, 1804-1815.

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

Chartron, J.W., Suloway, C.J., Zaslaver, M., and Clemons, W.M., Jr. (2010). Structural characterization of the Get4/Get5 complex and its interaction with Get3. Proc Natl Acad Sci U S A *107*, 12127-12132.

Ciczora, Y., Callens, N., Montpellier, C., Bartosch, B., Cosset, F.L., Op de Beeck, A., and Dubuisson, J. (2005). Contribution of the charged residues of hepatitis C virus glycoprotein E2 transmembrane domain to the functions of the E1E2 heterodimer. J Gen Virol *86*, 2793-2798.

Ciczora, Y., Callens, N., Penin, F., Pecheur, E.I., and Dubuisson, J. (2007). Transmembrane domains of hepatitis C virus envelope glycoproteins: residues involved in E1E2 heterodimerization and involvement of these domains in virus entry. J Virol *81*, 2372-2381.

Cross, B.C., and High, S. (2009). Dissecting the physiological role of selective transmembrane-segment retention at the ER translocon. J Cell Sci *122*, 1768-1777.

Dalal, K., and Duong, F. (2009). The SecY complex forms a channel capable of ionic discrimination. EMBO Rep *10*, 762-768.

Davis, N.G., and Model, P. (1985). An artificial anchor domain: hydrophobicity suffices to stop transfer. Cell *41*, 607-614.

Deber, C.M., Wang, C., Liu, L.P., Prior, A.S., Agrawal, S., Muskat, B.L., and Cuticchia, A.J. (2001). TM Finder: a prediction program for transmembrane protein segments using a combination of hydrophobicity and nonpolar phase helicity scales. Protein Sci *10*, 212-219.

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.

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.

Dowhan, W., and Bogdanov, M. (2009). Lipid-dependent membrane protein topogenesis. Annu Rev Biochem *78*, 515-540.

Egea, P.F., and Stroud, R.M. (2010). Lateral opening of a translocon upon entry of protein suggests the mechanism of insertion into membranes. Proc Natl Acad Sci U S A.

Ehrmann, M., and Beckwith, J. (1991). Proper insertion of a complex membrane protein in the absence of its amino-terminal export signal. J Biol Chem *266*, 16530-16533.

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.

Emr, S.D., Hanley-Way, S., and Silhavy, T.J. (1981). Suppressor mutations that restore export of a protein with a defective signal sequence. Cell *23*, 79-88.

Enquist, K., Fransson, M., Boekel, C., Bengtsson, I., Geiger, K., Lang, L., Pettersson, A., Johansson, S., von Heijne, G., and Nilsson, I. (2009). Membrane-integration characteristics of two ABC transporters, CFTR and P-glycoprotein. J Mol Biol *387*, 1153-1164.

Favaloro, V., Spasic, M., Schwappach, B., and Dobberstein, B. (2008). Distinct targeting pathways for the membrane insertion of tail-anchored (TA) proteins. J Cell Sci *121*, 1832-1840.

Finke, K., Plath, K., Panzner, S., Prehn, S., Rapoport, T.A., Hartmann, E., and Sommer, T. (1996). A second trimeric complex containing homologs of the Sec61p complex functions in protein transport across the ER membrane of S. cerevisiae. EMBO J *15*, 1482-1494.

Flanagan, J.J., Chen, J.C., Miao, Y., Shao, Y., Lin, J., Bock, P.E., and Johnson, A.E. (2003). Signal recognition particle binds to ribosome-bound signal sequences with fluorescence-detected subnanomolar affinity that does not diminish as the nascent chain lengthens. J Biol Chem *278*, 18628-18637.

Fons, R.D., Bogert, B.A., and Hegde, R.S. (2003). Substrate-specific function of the translocon-associated protein complex during translocation across the ER membrane. J Cell Biol *160*, 529-539.

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

Gierasch, L.M. (1989). Signal sequences. Biochemistry *28*, 923-930.

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

Granseth, E., von Heijne, G., and Elofsson, A. (2005). A study of the membrane-water interface region of membrane proteins. J Mol Biol *346*, 377-385.

Guo, D., Liu, J., Motlagh, A., Jewell, J., and Miller, K.W. (1996). Efficient insertion of odd-numbered transmembrane segments of the tetracycline resistance protein requires even-numbered segments. J Biol Chem *271*, 30829-30834.

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

Halic, M., and Beckmann, R. (2005). The signal recognition particle and its interactions during protein targeting. Curr Opin Struct Biol *15*, 116-125.

Halic, M., Blau, M., Becker, T., Mielke, T., Pool, M.R., Wild, K., Sinning, I., and Beckmann, R. (2006). Following the signal sequence from ribosomal tunnel exit to signal recognition particle. Nature *444*, 507-511.

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.

Hanein, D., Matlack, K.E., Jungnickel, B., Plath, K., Kalies, K.U., Miller, K.R., Rapoport, T.A., and Akey, C.W. (1996). Oligomeric rings of the Sec61p complex induced by ligands required for protein translocation. Cell *87*, 721-732.

Harris, C.R., and Silhavy, T.J. (1999). Mapping an interface of SecY (PrlA) and SecE (PrlG) by using synthetic phenotypes and in vivo cross-linking. J Bacteriol *181*, 3438-3444.

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.

Heijne, G. (1986). The distribution of positively charged residues in bacterial inner membrane proteins correlates with the trans-membrane topology. EMBO J *5*, 3021-3027.

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.

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.

Hessa, T., Meindl-Beinker, N.M., Bernsel, A., Kim, H., Sato, Y., Lerch-Bader, M., Nilsson, I., White, S.H., and von Heijne, G. (2007). Molecular code for transmembrane-helix recognition by the Sec61 translocon. Nature *450*, 1026-1030.

Hessa, T., Reithinger, J.H., von Heijne, G., and Kim, H. (2009). Analysis of transmembrane helix integration in the endoplasmic reticulum in S. cerevisiae. J Mol Biol *386*, 1222-1228.

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

Higy, M., Junne, T., and Spiess, M. (2004). Topogenesis of membrane proteins at the endoplasmic reticulum. Biochemistry *43*, 12716-12722.

Ismail, N., Crawshaw, S.G., Cross, B.C., Haagsma, A.C., and High, S. (2008). Specific transmembrane segments are selectively delayed at the ER translocon during opsin biogenesis. Biochem J *411*, 495-506.

Ismail, N., Crawshaw, S.G., and High, S. (2006). Active and passive displacement of transmembrane domains both occur during opsin biogenesis at the Sec61 translocon. J Cell Sci *119*, 2826-2836.

Janda, C.Y., Li, J., Oubridge, C., Hernandez, H., Robinson, C.V., and Nagai, K. (2010). Recognition of a signal peptide by the signal recognition particle. Nature *465*, 507-510.

Johansson, A.C., and Lindahl, E. (2009). Protein contents in biological membranes can explain abnormal solvation of charged and polar residues. Proc Natl Acad Sci U S A *106*, 15684-15689.

Junne, T., Schwede, T., Goder, V., and Spiess, M. (2006). The plug domain of yeast Sec61p is important for efficient protein translocation, but is not essential for cell viability. Mol Biol Cell *17*, 4063-4068.

Junne, T., Schwede, T., Goder, V., and Spiess, M. (2007). Mutations in the Sec61p channel affecting signal sequence recognition and membrane protein topology. J Biol Chem *282*, 33201-33209.

Kida, Y., Morimoto, F., Mihara, K., and Sakaguchi, M. (2006). Function of positive charges following signal-anchor sequences during translocation of the N-terminal domain. J Biol Chem *281*, 1152-1158.

Kida, Y., Morimoto, F., and Sakaguchi, M. (2009). Signal anchor sequence provides motive force for polypeptide chain translocation through the endoplasmic reticulum membrane. J Biol Chem *284*, 2861-2866.

Kiefer, D., and Kuhn, A. (2007). YidC as an essential and multifunctional component in membrane protein assembly. Int Rev Cytol *259*, 113-138.

Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J Mol Biol *157*, 105-132.

Lakkaraju, A.K., Mary, C., Scherrer, A., Johnson, A.E., and Strub, K. (2008). SRP keeps polypeptides translocation-competent by slowing translation to match limiting ER-targeting sites. Cell *133*, 440-451.

Li, H., Chavan, M., Schindelin, H., and Lennarz, W.J. (2008). Structure of the oligosaccharyl transferase complex at 12 A resolution. Structure *16*, 432-440.

Li, W., Schulman, S., Boyd, D., Erlandson, K., Beckwith, J., and Rapoport, T.A. (2007). The plug domain of the SecY protein stabilizes the closed state of the translocation channel and maintains a membrane seal. Mol Cell *26*, 511-521.

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.

Liu, L.P., and Deber, C.M. (1998). Uncoupling hydrophobicity and helicity in transmembrane segments. Alpha-helical propensities of the amino acids in non-polar environments. J Biol Chem *273*, 23645-23648.

Lodish, H.F. (2000). Molecular cell biology, 4th edn (New York, W.H. Freeman).

Luirink, J., Samuelsson, T., and de Gier, J.W. (2001). YidC/Oxa1p/Alb3: evolutionarily conserved mediators of membrane protein assembly. FEBS Lett *501*, 1-5.

Lundin, C., Kim, H., Nilsson, I., White, S.H., and von Heijne, G. (2008). Molecular code for protein insertion in the endoplasmic reticulum membrane is similar for N(in)-C(out) and N(out)-C(in) transmembrane helices. Proc Natl Acad Sci U S A *105*, 15702-15707.

Lyman, S.K., and Schekman, R. (1995). Interaction between BiP and Sec63p is required for the completion of protein translocation into the ER of Saccharomyces cerevisiae. J Cell Biol *131*, 1163-1171.

Maillard, A.P., Lalani, S., Silva, F., Belin, D., and Duong, F. (2007). Deregulation of the SecYEG translocation channel upon removal of the plug domain. J Biol Chem *282*, 1281-1287.

Martoglio, B. (2003). Intramembrane proteolysis and post-targeting functions of signal peptides. Biochem Soc Trans *31*, 1243-1247.

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.

Mateja, A., Szlachcic, A., Downing, M.E., Dobosz, M., Mariappan, M., Hegde, R.S., and Keenan, R.J. (2009). The structural basis of tail-anchored membrane protein recognition by Get3. Nature *461*, 361-366.

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.

Meindl-Beinker, N.M., Lundin, C., Nilsson, I., White, S.H., and von Heijne, G. (2006). Asn- and Asp-mediated interactions between transmembrane helices during translocon-mediated membrane protein assembly. EMBO Rep *7*, 1111-1116.

Menetret, J.F., Hegde, R.S., Aguiar, M., Gygi, S.P., Park, E., Rapoport, T.A., and Akey, C.W. (2008). Single copies of Sec61 and TRAP associate with a nontranslating mammalian ribosome. Structure *16*, 1126-1137.

Mitra, K., Schaffitzel, C., Shaikh, T., Tama, F., Jenni, S., Brooks, C.L., 3rd, Ban, N., and Frank, J. (2005). Structure of the E. coli protein-conducting channel bound to a translating ribosome. Nature *438*, 318-324.

Monier, S., Van Luc, P., Kreibich, G., Sabatini, D.D., and Adesnik, M. (1988). Signals for the incorporation and orientation of cytochrome P450 in the endoplasmic reticulum membrane. J Cell Biol *107*, 457-470.

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.

Mori, H., Tsukazaki, T., Masui, R., Kuramitsu, S., Yokoyama, S., Johnson, A.E., Kimura, Y., Akiyama, Y., and Ito, K. (2003). Fluorescence resonance energy transfer analysis of protein

translocase. SecYE from Thermus thermophilus HB8 forms a constitutive oligomer in membranes. J Biol Chem *278*, 14257-14264.

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.

Muller, L., de Escauriaza, M.D., Lajoie, P., Theis, M., Jung, M., Muller, A., Burgard, C., Greiner, M., Snapp, E.L., Dudek, J.*, et al.* (2010). Evolutionary gain of function for the ER membrane protein Sec62 from yeast to humans. Mol Biol Cell *21*, 691-703.

Ng, D.T., Brown, J.D., and Walter, P. (1996). Signal sequences specify the targeting route to the endoplasmic reticulum membrane. J Cell Biol *134*, 269-278.

Ngosuwan, J., Wang, N.M., Fung, K.L., and Chirico, W.J. (2003). Roles of cytosolic Hsp70 and Hsp40 molecular chaperones in post-translational translocation of presecretory proteins into the endoplasmic reticulum. J Biol Chem *278*, 7034-7042.

Nilsson, I., and von Heijne, G. (1990). Fine-tuning the topology of a polytopic membrane protein: role of positively and negatively charged amino acids. Cell *62*, 1135-1141.

Ogg, S.C., and Walter, P. (1995). SRP samples nascent chains for the presence of signal sequences by interacting with ribosomes at a discrete step during translation elongation. Cell *81*, 1075-1084.

Osborne, A.R., and Rapoport, T.A. (2007). Protein translocation is mediated by oligomers of the SecY complex with one SecY copy forming the channel. Cell *129*, 97-110.

Osborne, A.R., Rapoport, T.A., and van den Berg, B. (2005). Protein translocation by the Sec61/SecY channel. Annu Rev Cell Dev Biol *21*, 529-550.

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

Palade, G. (1975). Intracellular aspects of the process of protein synthesis. Science *189*, 867.

Palliser, C.C., and Parry, D.A. (2001). Quantitative comparison of the ability of hydropathy scales to recognize surface beta-strands in proteins. Proteins *42*, 243-255.

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.

Pitonzo, D., Yang, Z., Matsumura, Y., Johnson, A.E., and Skach, W.R. (2009). Sequence-specific retention and regulated integration of a nascent membrane protein by the endoplasmic reticulum Sec61 translocon. Mol Biol Cell *20*, 685-698.

Prinz, A., Hartmann, E., and Kalies, K.U. (2000). Sec61p is the main ribosome receptor in the endoplasmic reticulum of Saccharomyces cerevisiae. Biol Chem *381*, 1025-1029.

Rapoport, T.A. (2007). Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. Nature *450*, 663-669.

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.

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.

Rudiger, S., Buchberger, A., and Bukau, B. (1997). Interaction of Hsp70 chaperones with substrates. Nat Struct Biol *4*, 342-349.

Sadlish, H., Pitonzo, D., Johnson, A.E., and Skach, W.R. (2005). Sequential triage of transmembrane segments by Sec61alpha during biogenesis of a native multispanning membrane protein. Nat Struct Mol Biol *12*, 870-878.

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.

Saparov, S.M., Erlandson, K., Cannon, K., Schaletzky, J., Schulman, S., Rapoport, T.A., and Pohl, P. (2007). Determining the conductance of the SecY protein translocation channel for small molecules. Mol Cell *26*, 501-509.

Sato, T., Sakaguchi, M., Mihara, K., and Omura, T. (1990). The amino-terminal structures that determine topological orientation of cytochrome P-450 in microsomal membrane. EMBO J *9*, 2391-2397.

Schafer, A., and Wolf, D.H. (2009). Sec61p is part of the endoplasmic reticulum-associated degradation machinery. EMBO J *28*, 2874-2884.

Schaletzky, J., and Rapoport, T.A. (2006). Ribosome binding to and dissociation from translocation sites of the endoplasmic reticulum membrane. Mol Biol Cell *17*, 3860-3869.

Seppala, S., Slusky, J.S., Lloris-Garcera, P., Rapp, M., and von Heijne, G. (2010). Control of membrane protein topology by a single C-terminal residue. Science *328*, 1698-1700.

Shen, K., and Shan, S.O. (2010). Transient tether between the SRP RNA and SRP receptor ensures efficient cargo delivery during cotranslational protein targeting. Proc Natl Acad Sci U S A *107*, 7698-7703.

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

Smith, M.A., Clemons, W.M., Jr., DeMars, C.J., and Flower, A.M. (2005). Modeling the effects of prl mutations on the Escherichia coli SecY complex. J Bacteriol *187*, 6454-6465.

Snapp, E.L., Reinhart, G.A., Bogert, B.A., Lippincott-Schwartz, J., and Hegde, R.S. (2004). The organization of engaged and quiescent translocons in the endoplasmic reticulum of mammalian cells. J Cell Biol *164*, 997-1007.

Spiess, M. (1995). Heads or tails--what determines the orientation of proteins in the membrane. FEBS Lett *369*, 76-79.

Spiess, M., and Lodish, H.F. (1986). An internal signal sequence: the asialoglycoprotein receptor membrane anchor. Cell *44*, 177-185.

Szczesna-Skorupa, E., Browne, N., Mead, D., and Kemper, B. (1988). Positive charges at the NH2 terminus convert the membrane-anchor signal peptide of cytochrome P-450 to a secretory signal peptide. Proc Natl Acad Sci U S A *85*, 738-742.

Szczesna-Skorupa, E., and Kemper, B. (1989). NH2-terminal substitutions of basic amino acids induce translocation across the microsomal membrane and glycosylation of rabbit cytochrome P450IIC2. J Cell Biol *108*, 1237-1243.

Tam, P.C., Maillard, A.P., Chan, K.K., and Duong, F. (2005). Investigating the SecY plug movement at the SecYEG translocation channel. EMBO J *24*, 3380-3388.

Tsukazaki, T., Mori, H., Fukai, S., Ishitani, R., Mori, T., Dohmae, N., Perederina, A., Sugita, Y., Vassylyev, D.G., Ito, K.*, et al.* (2008). Conformational transition of Sec machinery inferred from bacterial SecYE structures. Nature *455*, 988-991.

Tu, L., Wang, J., Helm, A., Skach, W.R., and Deutsch, C. (2000). Transmembrane biogenesis of Kv1.3. Biochemistry *39*, 824-836.

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.

Veenendaal, A.K., van der Does, C., and Driessen, A.J. (2004). The protein-conducting channel SecYEG. Biochim Biophys Acta *1694*, 81-95.

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

von Heijne, G. (1992). Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. J Mol Biol *225*, 487-494.

von Heijne, G. (2006). Membrane-protein topology. Nat Rev Mol Cell Biol *7*, 909-918.

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.

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.

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.

Willer, M., Forte, G.M., and Stirling, C.J. (2008). Sec61p is required for ERAD-L: genetic dissection of the translocation and ERAD-L functions of Sec61P using novel derivatives of CPY. J Biol Chem *283*, 33883-33888.

Wittke, S., Dunnwald, M., Albertsen, M., and Johnsson, N. (2002). Recognition of a subset of signal sequences by Ssh1p, a Sec61p-related protein in the membrane of endoplasmic reticulum of yeast Saccharomyces cerevisiae. Mol Biol Cell *13*, 2223-2232.

Xie, K., Hessa, T., Seppala, S., Rapp, M., von Heijne, G., and Dalbey, R.E. (2007). Features of transmembrane segments that promote the lateral release from the translocase into the lipid phase. Biochemistry *46*, 15153-15161.

Zhang, B., and Miller, T.F., 3rd (2010). Hydrophobically stabilized open state for the lateral gate of the Sec translocon. Proc Natl Acad Sci U S A *107*, 5399-5404.

Zimmermann, R., Eyrisch, S., Ahmad, M., and Helms, V. (2010). Protein translocation across the ER membrane. Biochim Biophys Acta.

Higy, M. (2005). Dynamic insertion of membrane proteins at the endoplasmic reticulum. PhD thesis. Martin Spiess` group, Biozentrum, University of Basel.

Curriculum vitae

Name: Lucyna Kocik
Date of birth: 10.12.1982
Place of origin: Boguslawice (Poland)
Nationality: Polish

<u>Education:</u>

2006-2010   PhD in biochemistry under the guidance of Prof. Dr. Martin Spiess; Biozentrum, University of Basel. Thesis title: Mechanism of Sec61-mediated insertion of proteins into the endoplasmic reticulum membrane.

2004-2006   MSc in medical biotechnology under the guidance of Prof. Maria-Malicka-Blaszkiewicz; Department of Cell Pathology, University of Wroclaw (Poland). Thesis title: Biological activity of pyridine N-oxides.

February-June 2005   Socrates/Erasmus exchange under the guidance of Dr. Sven van IJzendoorn; Rijksuniversiteit Groningen (The Netherlands), Department of membrane cell biology. Research project: Role of rab11a in HepG2 membrane dynamics and polarity development.

2001-2004   BSc in biotechnology under the guidance of Prof. Maria-Malicka-Blaszkiewicz; Department of Cell Pathology, University of Wroclaw (Poland). Thesis title: Eukaryotic cell motility. Current opinion. Changes in pathology.

1997-2001   Secondary school in Olesnica (Poland)

1989-1997   Elementary school in Olesnica (Poland)

<u>Additional courses</u>

May 2010   Good Clinical Practice and regulatory requirements; University Hospital Basel, Clinical Trial Unit

<u>Languages</u>

Polish   (mother tongue)
English   (fluent)
German   (basic knowledge)

<u>Publications</u>

Junne, T., **Kocik, L.**, and Spiess, M. (2010). The hydrophobic core of the Sec61 translocon defines the hydrophobicity threshold for membrane integration. Mol Biol Cell *21*, 1662-1670