Refined topology model of the STT3/Stt3 protein subunit of the oligosaccharyltransferase complex

Received for publication, February 3, 2017, and in revised form, May 10, 2017 Published, Papers in Press, May 16, 2017, DOI 10.1074/jbc.M117.779421

Patricia Lara[‡1], Karin Öjemalm[‡1], Johannes Reithinger[‡1], Aurora Holgado[‡], You Maojun[‡], Abdessalem Hammed[‡], Daniel Mattle[‡], Hyun Kim[§], and IngMarie Nilsson[‡2]

From the [‡]Department of Biochemistry and Biophysics, Stockholm University, SE-10691 Stockholm, Sweden and [§]Department of Biological Sciences, Seoul National University, Seoul 08826, South Korea

Edited by Thomas Söllner

The oligosaccharyltransferase complex, localized in the endoplasmic reticulum (ER) of eukaryotic cells, is responsible for the *N*-linked glycosylation of numerous protein substrates. The membrane protein STT3 is a highly conserved part of the oligosaccharyltransferase and likely contains the active site of the complex. However, understanding the catalytic determinants of this system has been challenging, in part because of a discrepancy in the structural topology of the bacterial *versus* eukaryotic proteins and incomplete information about the mechanism of membrane integration. Here, we use a glycosylation mapping approach to investigate these questions. We measured the membrane integration efficiency of the mouse STT3-A and yeast Stt3p transmembrane domains (TMDs) and report a refined topology of the N-terminal half of the mouse STT3-A. Our results show that most of the STT3 TMDs are well inserted into the ER membrane on their own or in the presence of the natural flanking residues. However, for the mouse STT3-A hydrophobic domains 4 and 6 and yeast Stt3p domains 2, 3a, 3c, and 6 we measured reduced insertion efficiency into the ER membrane. Furthermore, we mapped the first half of the STT3-A protein, finding two extra hydrophobic domains between the third and the fourth TMD. This result indicates that the eukaryotic STT3 has 13 transmembrane domains, consistent with the structure of the bacterial homolog of STT3 and setting the stage for future combined efforts to interrogate this fascinating system.

N-Linked glycosylation is one of the most common types of eukaryotic protein modifications. The attachment of carbohydrates to asparagine residues in proteins takes place during protein translocation across the ER[3] membrane and is catalyzed by the OST complex, referred to OT in yeast. The OST/OT catalytic site is located on the lumenal side of the ER membrane, and the whole complex is located near the translocon, the protein translocation machinery that is composed of Sec61α, -β, and -γ. The translocon facilitates the translocation of proteins across or into the ER membrane (1, 2). Proteins of the translocon that come into contact with the nascent polypeptide chains during their translocation across the ER membrane have been identified by different cross-linking methods (3, 4). The mechanism of membrane protein integration in these studies is based upon *in vitro* biochemical and biophysical assays (for review, see Ref. 5–7).

The central subunit of the OST complex is the highly conserved STT3 found in eukaryotes and archaea and in some eubacteria. Both the mammalian OST and the yeast OT are composed of eight to nine identified subunits (8, 9). Of these, STT3 (Stt3 protein (Stt3p) in yeast) is believed to be the catalytic subunit that contains the active site (10, 11). It was first discovered as a staurosporine- and temperature-sensitive mutant in *Saccharomyces cerevisiae* (12) and was later found to be a subunit of the heteroligomeric OT complex (11). The 3.4 Å resolution X-ray structure of the bacterial OST, the single-subunit PglB protein from *Campylobacter lari*, reveals the fold of the STT3 proteins. It also indicates the locations of the transmembrane (TM) helices, the connecting loops and conserved residues forming the active site of PglB (13). The sequence alignment between PglB and other STT3 proteins show conserved motifs in loops between helices and in the C-terminal part of the proteins (13–15). Furthermore, the similarity in membrane topology indicates that PglB and eukaryotic STT3s share a common reaction mechanism for *N*-linked glycosylation (8, 10, 16). Mammalian cells express two homologues of the yeast Stt3p (STT3-A and STT3-B) that form tissue-specific complexes with other OST subunits (ribophorins I and II, OST48, DAD1, and Ost4) (9, 17, 18). Stt3p from *S. cerevisiae* has 50% sequence identity to the human homologues of STT3, and highly conserved stretches of residues are found throughout the protein (19).

This work was supported by grants from the Swedish Cancer Foundation (130624; to I. N. and Gunnar von Heijne), from the Swedish Foundation for International Cooperation in Research and Higher Education (STINT) (210/083(12) and KU 2003-4674; to I. N.), from the Swedish Strategic Foundation (A302:200 and SSF-Infection Biology 2012(SB12–0026); to I. N.), and from the National Research Foundation of Korea (NRF-2016R1A2B2013459) and Promising-Pioneering Researcher Program through Seoul National University (to H. K.). The authors declare that they have no conflicts of interest with the contents of this article.

[1] These authors contributed equally to this work.

[2] To whom correspondence should be addressed: Dept. of Biochemistry and Biophysics, Stockholm University, Svante Arrhenius väg 16C, SE-10691 Stockholm, Sweden. Tel.: 46-8-162728; Fax: 46-8-153679; E-mail: ingmarie@dbb.su.se.

[3] The abbreviations used are: ER, endoplasmic reticulum; CRM, column-washed rough microsomes; OST (OT), oligosaccharyltransferase; TM, transmembrane; TMD, transmembrane domain; Endo H, endoglycosidase H; PK, proteinase K; Lep, leader peptidase; SP, signal peptide; SPase, signal peptidase; H, hydrophobic.

Both STT3 and Stt3p are multispanning membrane proteins, with 10–13 predicted TM segments by hydrophobicity plot, whereas the X-ray structure of the single-subunit membrane protein, PglB indicated 13 TM segments (13). STT3 is predicted not to have a cleavable signal sequence. It has a glycosylated C-terminal domain containing the WWD motif in the lumen of the ER, which is suggested to define the polypeptide substrate specificity (13). We earlier showed that the topologies for both yeast Stt3p and mouse STT3A are similar, with the N terminus located in the cytoplasm, the C terminus in the lumen, and the number of TM segments to be 11 with possible alterations of first half of the protein (20).

Although, the crystal structure of a bacterial single-subunit STT3 protein is available (13), the molecular mechanism underlying the STT3 membrane integration is still not fully understood. We, therefore, performed a membrane integration study of yeast Stt3p and mouse STT3-A in the ER membrane. Initially, we defined the STT3 TM domains (TMDs) by the Δ*G* predictor and then we analyzed the insertion efficiency of the predicted TMDs using an *in vitro* assay based on asparagine-linked glycosylation. Our results of the membrane insertion efficiency show that most of the STT3/Stt3p TMDs, when including the natural flanking residues surrounding the TMD, integrate effectively into the membrane. Exceptions are TMD4 and TMD6 of mouse STT3A and TMD2, -3a, -3c, and -6 of yeast Stt3p. Here, we also present experimental data for a refinement of the first half of the mouse STT3-A protein, including two extra TMDs between the third and the fourth TMD. We, therefore, conclude that eukaryotic STT3s have 13 TM segments that are consistent with the structural model of PglB (Fig. 1).

Results

Prediction of the position of the TMDs of mouse STT3-A and yeast Stt3p

Our study started by defining the positions of the mouse STT3-A and yeast Stt3p TMDs using the Δ*G* predictor server (http://dgpred.cbr.su.se/)[4]. The server calculates the theoretical Δ*G* value of all predicted TMDs (Table 1 and 2), which is a measure of the tendency of any sequence to insert into a biological membrane (21). Adding TMD flanking residues allows the program to find the most optimal TMD sequence with the lowest Δ*G* value. We, therefore, calculated and measured the Δ*G* value for each TMD sequence alone and when the flanking loops were included. The loop here is defined as the protein sequence between two adjacent TMDs. For the long loop sequences, we extended the TMD flanks with a maximum of 20 residues. We then compared the predicted TMD positions with the experimentally determined eleven-TM topology model of STT3 (20) and the recently published crystal structure of PglB (13) and found them to agree well (Fig. 1).

[4] Please note that the JBC is not responsible for the long-term archiving and maintenance of this site or any other third party-hosted site.

Table 1

Position of the 13 mouse TMD helices of STT3-A determined by the Δ*G* predictor

Data are *versus* measured Δ*G* values for each of the selected hydrophobic domains in mouse STT3-A. For the predicted values the Δ*G* prediction server was used (http://www.dgpred.cbr.su.se).[4] All TMDs were predicted from the full-length mouse STT3-A protein and selected to be 19-residues long. These 19-residues long TMD segments without and with natural flanking residues (F) were predicted by the Δ*G* predictor using length corrections and sub-sequences with the lowest Δ*G* (red). The numbers in parentheses correlate with the positions of the TMDs. All TMDs without natural flanking residues and with flanks composed of < 10 residues were insulated by a tetra-peptide, GGPG . . . GPGG. Two potential glycosylation sites are underlined. The integration (%) is calculated from results based on at least 3 independent experiments.

Region (position)	Flank	TMD-sequence	Flank	Integration (%)	Predicted ΔG (kcal/mol)	Measured ΔG (kcal/mol)
TMD1 (17-35)	GGPG	LLKLLILSMAAVLSFSTRL	GPGG	78	-0.2	-0.7
TMD1F (1-55)	…KQD	TLLKLLILSMAAVLSFSTRLF AVLRF	E…	85	-1.6	-1.0
TMD2 (113-131)	GGPG	RNVCVFLAPLFSSFTTIVT	GPGG	6	2.7	1.6
TMD2F (93-141)	…H	ITIDIRNVCVFLAPLFSSFTTIV TYHLT	K…	60	1.9	-0.2
TMD3A (142-160)	GGPG	GAGLLAAAMIAVVPGYISR	GPGG	70	1.6	-0.5
TMD3AF (132-165)	…KE	LKDAGAGLLAAAMIAVVPGY ISRSVA	GS…	81	1.3	-0.8
TMD3B (166-184)	GGPG	YDNEGIAIFCMLLTYYMWI	GPGG	34	0.6	0.4
TMD3BF (161-194)	…YDNE	GIAIFCMLLTYYMWI	KAVK…	70	0.6	-0.5
TMD3C (195-213)	GGPG	AAKCALAYFYMVSSWGGYV	GPGG	58	3.2	-0.2
TMD3CF (185-213)	…GS	IYWAAKCALAYFYMVSSWG GYV	GPGG	83	2.1	-0.9
TMD4 (214-232)	GGPG	FLINLIPLHVLVLMLTGRF	GPGG	29	0.1	0.5
TMD4F (214-235)	GGPG	FLINLIPLHVLVLMLTGRF	SHR…	38	0.1	0.3
TMD5 (236-254)	GGPG	IYVAYCTVYCLGTILSMQI	GPGG	84	1.7	-1.0
TMD5F (233-264)	…SH	RIYVAYCTVYCLGTILSMQIS FV	G…	95	1.1	-1.7
TMD6 (269-287)	GGPG	AAFGVFGLCQIHAFVDYLR	GPGG	8	2.5	1.4
TMD6F (254-300)	…SEH	MAAFGVFGLCQIHAFVDYLR	SKLN…	22	2.2	0.7
TMD7 (302-320)	GGPG	VISLVGFVLLTVGALLMLT	GPGG	84	-0.8	-1.0
TMD7F (289-340)	…G	VLFRSVISLVGFVLLTVGALL MLT	GKIS…	87	-2.2	-1.1
TMD8 (361-379)	GGPG	YFDLQLLVFMFPVGLYYCF	GPGG	19	0.0	0.8
TMD8F (341-384)	…TWSS	YYFDLQLLVFMFPVGLYYCF	SNLS…	72	-0.3	-0.6
TMD9 (385-403)	GGPG	ARIFIIMYGVTSMYFSAVM	GPGG	81	1.8	-0.8
TMD9F (381-403)	…SN	LSDARIFIIMYGVTSMYFSAV M	GPGG	83	1.5	-0.9
TMD10 (404-422)	GGPG	VRLMLVLAPVMCILSGIGV	GPGG	34	0.6	0.4
TMD10F (404-441)	GGPG	VRLMLVLAPVMCILSGIGVSQ VL	S…	88	0.0	1.2
TMD11 (456-474)	GGPG	VASGMILVMAFFLITYTFH	GPGG	89	-1.0	-1.2
TMD11F (436-494)	…IKNE	VASGMILVMAFFLITYTF	HSTW…	87	-1.0	-1.1

Membrane integration efficiency of individual TMDs of mouse STT3-A

To measure the efficiency of the translocon-mediated membrane integration of the mouse STT3-A TMDs, we used a well established experimental approach based on *N*-linked glycosylation of a modified version of the *Escherichia coli* inner membrane protein leader peptidase (Lep) (22–25). We used two engineered versions of Lep (LepH2 and LepH3) that allow experimentally based quantitative measurements of the insertion efficiency of STT3-A TMD segments in their natural orientation combined without and with adjacent loops. The LepH2 version is engineered such that Lep TM helix 2 (N_{in}-C_{out}) can be replaced by any sequence of interest (23). Here, we replaced it with all STT3-A TMDs with uneven numbers without (TMD1, -3, -5, etc.) or with flanking residues (TMD1F, -3F, -5F, etc,). Likewise, we introduce the even-numbered TMDs of STT3-A as well as their extended versions into the P2 domain of LepH3 (N_{out}-C_{in}) (21, 22) (Table 1, Fig. 2). The efficiency of the

translocon-mediated membrane integration of the TMD inserts can be monitored by the extent of glycan modification of the second of two engineered acceptor sites for *N*-linked glycosylation in LepH2 (*G1*, *G2*) or LepH3 (*G2*, *G3*) (Fig. 2*A*). We expressed each cloned construct *in vitro* in the presence of ER microsomal membranes and subsequently separated the single and double-glycosylated Lep model proteins containing TMDs of STT3-A by SDS-PAGE. Then we quantified and calculated the insertion efficiency as described under "Experimental procedures" (Fig. 2, *B* and *C*) (21–23).

Table 2

Yeast Stt3p TMDs and their predicted and measured Δ*G* values

The positions of the TMD segments were predicted from full-length yeast Stt3p using the ΔG predictor (21). The TMD length was set to be 19 residues. Δ*G* values in kcal/mol were also predicted with the Δ*G* predictor using length corrections and sub-sequences with lowest Δ*G* (red). The TMDs are either flanked by insulating GGPG/GPGG tetra-peptides or their natural flanks (F), as indicated. The integration efficiency of the TMDs is based on result from at least 3 independent experiments and was used to calculate the "measured" Δ*G* values in kcal/mol.

Region (position)	Flank	TMD-sequence	Flank	Integration (%)	Predicted ΔG (kcal/mol)	Measured ΔG (kcal/mol)
TMD1 (16-34)	GGPG	LKLVIFVAIFGAAISSRLF	GPGG	85	-0.3	-1.1
TMD2 (114-132)	GGPG	VCVLFAPLFSGVTAWATYE	GPGG	11	1.9	1.2
TMD3a (140-158)	GGPG	ASAGLLAAGFIAIVPGYIS	GPGG	12	1.5	1.2
TMD3b (167-185)	GGPG	NEAIAITLLMVTFMFWIKA	GPGG	53	-0.4	-0.1
TMD3bF (159-205)	…GSYD	NEAIAITLLMVTFMFWIKA	QKTG…	95	-0.4	-2.3
TMD3c (195-212)	GGPG	ATCAALFYFYMVSAWGGYV	GPGG	21	1.6	0.8
TMD4 (213-231)	GGPG	FITNLIPLHVFLLILMGRY	GPGG	66	0.2	-0.4
TMD6 (276-285)	GGPG	MAALGVFGLIQIVAFGDFV	GPGG	22	1.0	0.8
TMD7 (300-318)	GGPG	VSLFLILVKGVVGLSALTY	GPGG	75	1.3	-0.7
TMD8 (365-383)	GPGG	FLIWLFPAGVFLLFLDLKD	GPGG	49	-1.1	0.0
TMD8F (345-383)	…FDTH	FLIWLFPAGVFLLFLDLKD	GPGG	44	-1.1	0.2
TMD11 (444-462)	GGPG	ALLAKLIVSGSFIFYLYLF	GPGG	65	-0.4	-0.4

When LepH2 constructs are co-translationally integrated into microsomal membranes, a fraction of the molecules with membrane-integrated TMD is often cleaved by signal peptidase (SPase) into two protein fragments with lower molecular mass (Fig. 2*B*). The extent of cleavage is dependent on the nature of the TMD sequence and can be prevented by the addition of a SPase inhibitor during translation (23, 26). Thus the fraction of cleaved protein can be added to the fraction of proteins with integrated STT3 TMD.

Interestingly, when expressing TMD2 without flanking loops in LepH3 we observed an extra band on the SDS-gel that corresponds to a triple-glycosylated LepH3 protein. We verified the triple glycosylation by endoglycosidase H (*Endo H*) treatment (Figs. 2*B* and 3). An unusual glycosylation site ^{114}NVC (Asn-Val-Cys) was discovered in TMD2 of STT3 (Table 1) (27–31), and we confirmed the site by substituting it to ^{114}QVC (Gln-Val-Cys), which completely abolished the glycosylation. We used this modified construct to avoid complications when quantifying the insertion efficiency of TMD2. The acceptor site was found in the beginning of TMD2 and can only be modified when TMD2 is not inserted into the membrane. It has a lower extent of glycosylation efficiency than the more common Asn-X-Thr/Ser site (~40%).

Our results for 10 of the 11 previously annotated STT3-A TMDs (20) (TMD3 is discussed later) show that most TMDs integrate efficiently into the ER membrane when taken out of their natural context together with the flanking regions and introduced into the LepH2 or LepH3 model proteins. The exceptions are the following. TMD2 integrated to only 6% into the membrane and by including the flanking loops the insertion increased to 60%. TMD4 integrated poorly, even in the presence of its natural flanks (29% as "isolated TMD," 38% including

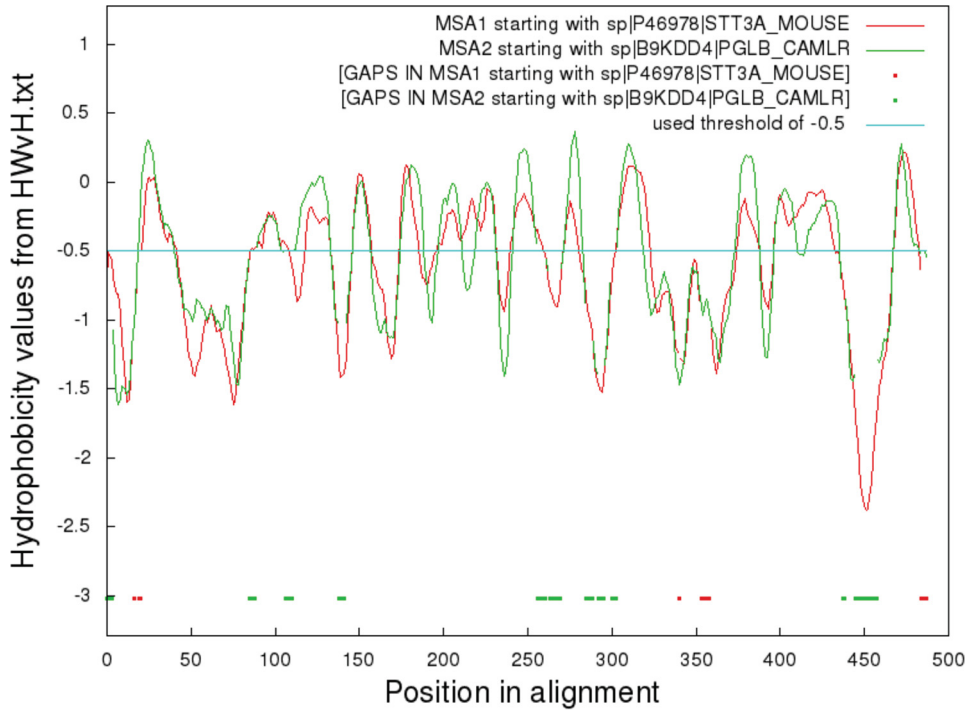
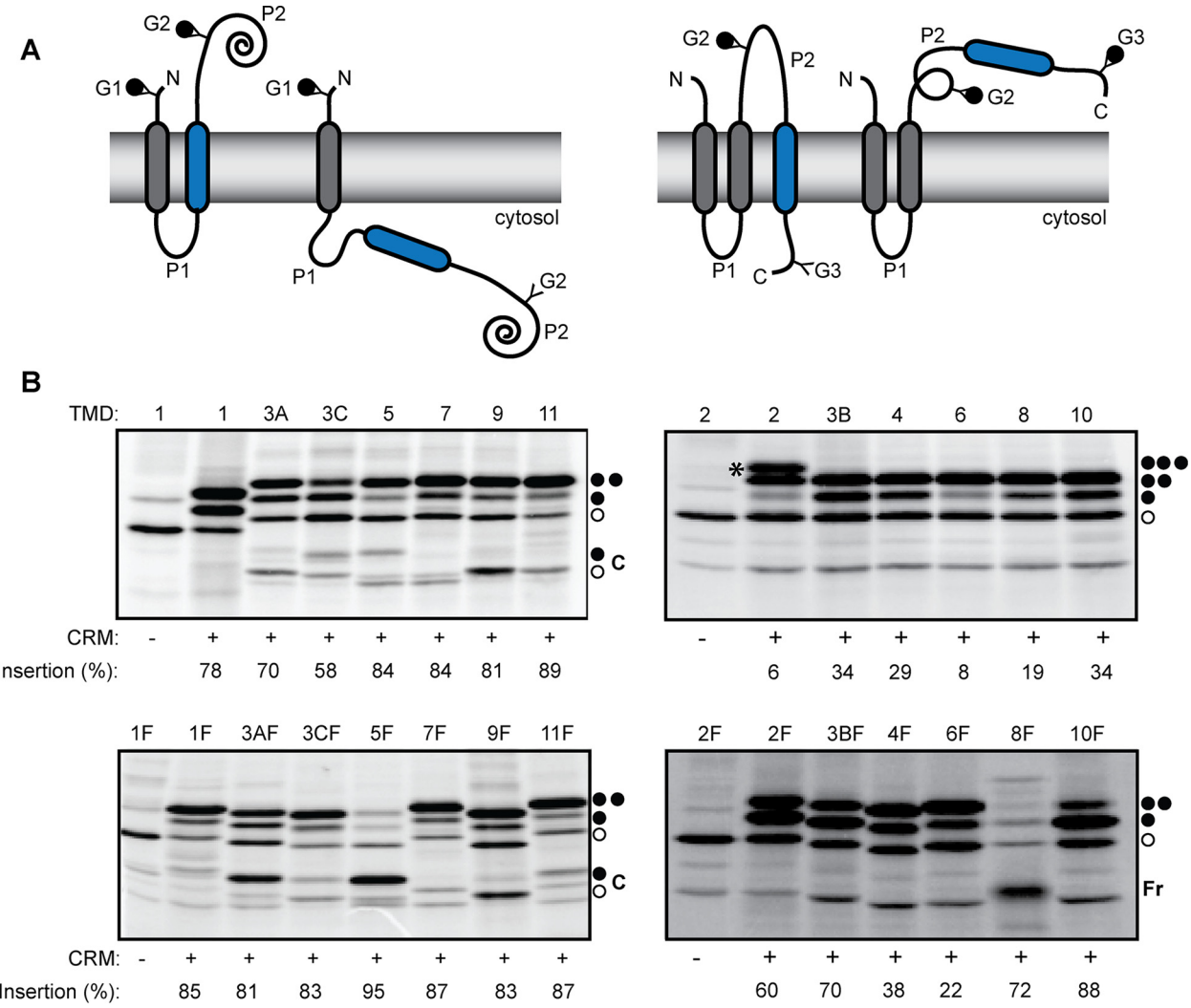
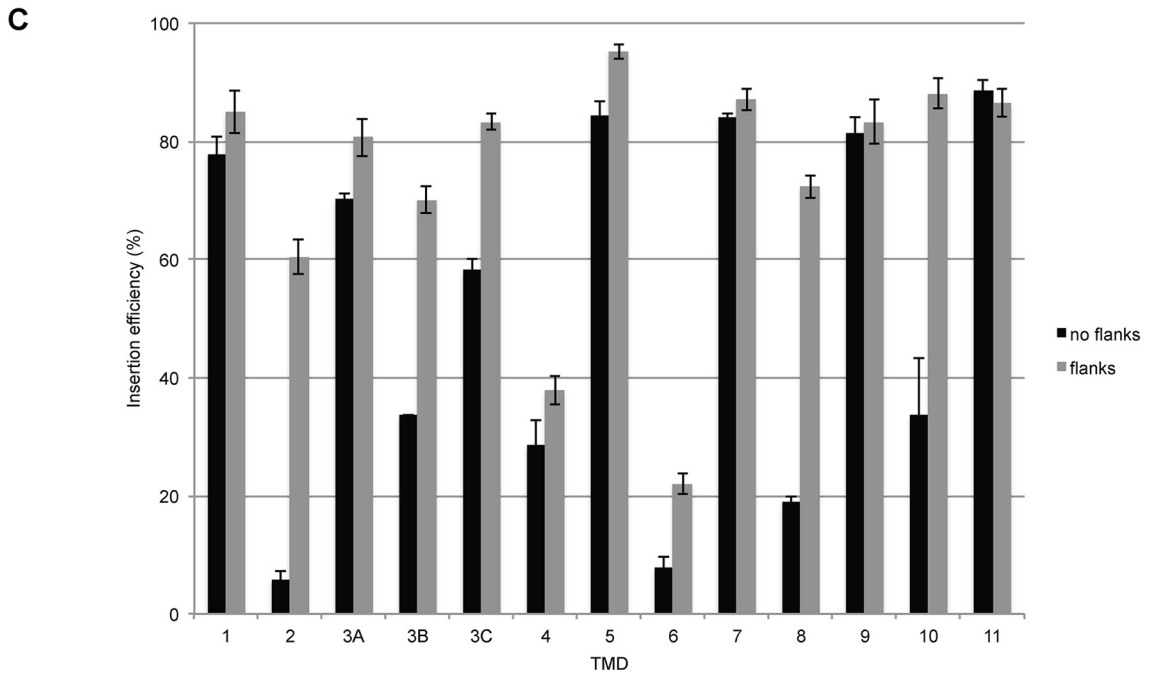


Figure 1. Alignment of the TMD region of STT3 and PglB. The alignment when using hydrophobicity plot suggests that mouse STT3A (UniProtKB ID: P46978) (1–474) and Campylobacter lari PglB (UniProtKB ID: B9KDD4) (1–433) have a similar membrane topology (http://www.bioinfo.mpg.de/AlignMe/AlignMe_MSA.html) (47–49).[4]


A
G2
P2
G1
N
G1
N
cytosol
P1
P1
G2
P2
G2
P2
N
N
P2
G3
G2
C
cytosol
P1
C
G3
P1
B
TMD: 1 1 3A 3C 5 7 9 11
C
CRM: - + + + + + + +
Insertion (%): 78 70 58 84 84 81 89
2 2 3B 4 6 8 10
*
- + + + + + + +
6 34 29 8 19 34
1F 1F 3AF 3CF 5F 7F 9F 11F
CRM: - + + + + + + +
Insertion (%): 85 81 83 95 87 83 87
2F 2F 3BF 4F 6F 8F 10F
Fr
- + + + + + + +
60 70 38 22 72 88


C
Insertion efficiency (%)
100
80
60
40
20
0
1 2 3A 3B 3C 4 5 6 7 8 9 10 11
TMD
no flanks
flanks

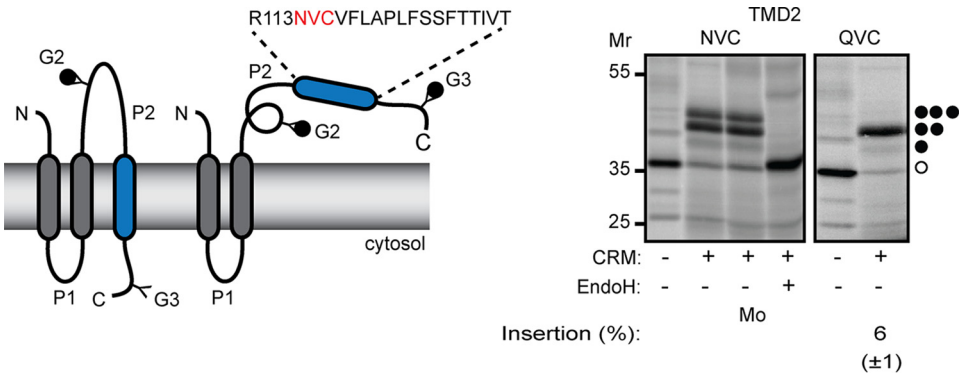


Figure 3. Verification of mouse TMD2 NVC glycosylation site. To verify the glycosylation of the NVC site in mouse TMD2, the sample was Endo H-treated, and as a control a sample with the same buffer conditions but without Endo H (*Mo*) was made. To specifically verify the NVC glycosylation, the site was destroyed by a replacement of the N (Asn) to a Q (Gln). Unglycosylated, single-, double-, and triple-glycosylated products are indicated with an *unfilled circle* and, respectively, *one*, *two*, or *three filled circles*. The percent insertion is calculated from at least three independent experiments.

flanking loops) (Table 1, Fig. 2, *B* and *C*). Finally, TMD6 integrated to 8% in the membrane and by adding its adjacent regions the integration efficiency increased barely to 22%. The results for TMD2 and TMD6 correlated well with their highly positive, predicted ΔG values for the TMDs by themselves (2.7 and 2.5 kcal/mol, respectively) and together with their loop-TMD-loop sequence (1.9 and 2.2 kcal/mol, respectively). For TMD4, on the other hand, the values did not correlate so well with its predicted values (0.1 without and 0.1 kcal/mol with flanking residues).

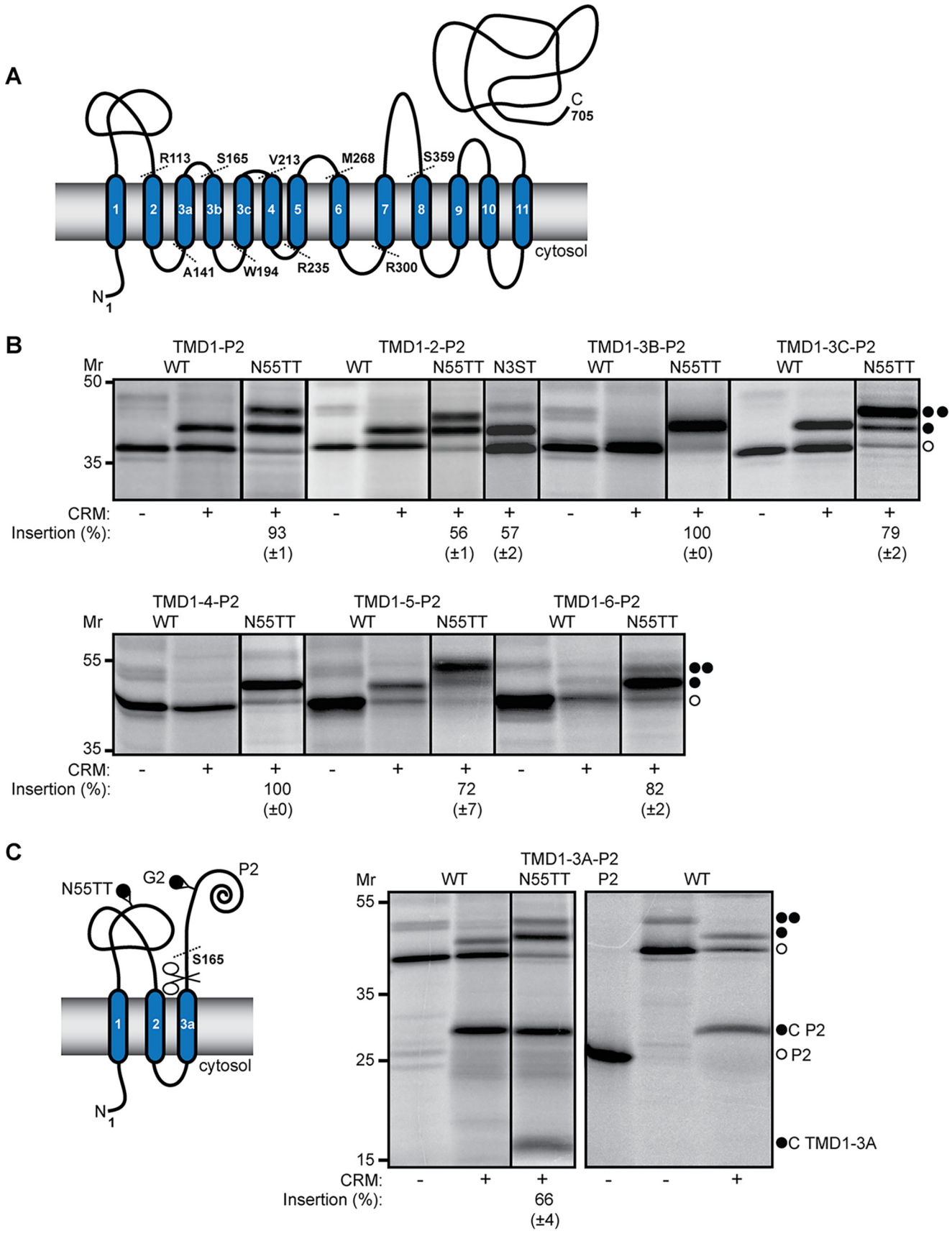
Refining the topology of the N-terminal part of the STT3 protein

The unresolved question of the experimentally determined eleven-TMD topology model of STT3 and Stt3p, is how two hydrophobic stretches could serve as TMD3 (TMD3A, TMD3B) (20). This together with the recently published 13-TMD crystal structure of the bacterial homologue PglB (13) motivated us to experimentally take a closer look at the first half of the STT3 protein. We noted a marginally hydrophobic sequence directly upstream adjacent to TMD4 of STT3-A with a theoretical ΔG value of 2.1 kcal/mol. This encouraged us to hypothesize a 13-TMD topology model also for the mouse STT3 homologues. TMD3A, TMD3B, and the marginally hydrophobic sequence, named here TMD3C, would in this case all be STT3 TMDs. Indeed, TMD3A, -B, and -C, cloned into the appropriate Lep variant, inserted efficiently into microsomal membranes in the presence of their flanking loops (*A*, 81%; *B*, 70%; *C*, 83%; Fig. 2, *B* and *C*).

To strengthen our results, we decided to map the location of the C terminus of STT3-A truncations by fusing them to the Lep-P2 domain, which included an engineered glycosylation site. For this we generated mouse STT3-A truncations that ended directly after the C-terminal loop of the TMDs of interest, *i.e.* right before the next TMD. We included one to seven TMDs from STT3-generating STT3-A-Lep-P2 fusions (Fig. 4*A*). For a sequence with more than two TMDs, multiple topologies are possible. By adding strategically glycosylation sites in the N terminus and in one loop in the STT3-A-Lep fusion series, it was possible to determine the most populated topologies more precisely (Fig. 4*A*). Therefore, one extra site except from the one in the P2 domain of LepH2 at position 120 was engineered at position ^{55}NTT (Asn-Thr-Thr) in the large luminal domain connecting TMD1 and -2 in STT3-A. The other was introduced in the N-terminal loop of the TMD1–2-P2 construct at position ^{3}NST (Asn-Ser-Thr). We reason this would allow us to monitor the location of the P2 domain of every protein fusion construct (Fig. 4*A*). The results of the two series of constructs without and with the glycosylation site in the loop correlate well with each other. This indicates no or little disturbance of the natural protein topology due to creation of additional glycosylation sites.

The glycosylation pattern of STT3-truncated versions reflects to some extent the measured insertion efficiency of the individual TMDs, as for TMD1–2-P2 and TMD2F in LepH3 (56 and 60%). The same was observed for TMD1–3C-P2 compared with TMD3CF in LepH3 (79% and 83%, respectively). Adding a glycosylation site (replacing ^{3}KLG for ^{3}NST) in the N terminus made it possible to determine the most populated topology of TMD1–2-P2 Lep fusion more precisely (Fig. 4*B*). The data confirm that the N terminus of the TMD1–2-P2 is located in the cytosol, and the single glycosylation in construct ^{55}NTT suggests a cytosolic orientation of the P2 domain (~56%). TMD1-P2 inserts with about the same efficiency as for TMD1F (93% and 85%, respectively) (Figs. 2, *B* and *C*, and 4*B*) when calculated with both glycosylated species in the ^{55}NTT construct. Although both TMD1–3A-P2 and TMD1–5-P2 insert ~20% less than the individual TMD3A and TMD5, they insert

Figure 2. Model protein Lep and membrane insertion efficiency of mouse STT3-A. *A*, the leader peptidase model proteins, LepH2 (*left*) and LepH3 (*right*). *B*, *in vitro* translation in the presence and absence of CRM of pancreas from dog of constructs containing all STT3-A TMDs without and with flanking residues (*F*). Odd-numbered TMDs to the *left* were inserted into LepH2, and even numbered TMDs are to the *right* in LepH3. The NVC-glycosylated product is indicated with an asterisk (*), SPase-cleaved fragments are indicated with a *C*, not fully translated full-length protein is indicated with *Fr*, and unglycosylated, single-, double-, and triple-glycosylated products are indicated with an open circle and, respectively, *one*, *two*, or *three filled circles*. The percent insertion is calculated from at least three independent experiments. *C*, insertion efficiency of all mouse TMDs without and with natural flanking residues. TMDs without and with flanking residues are indicated with a *black* and *gray* color, respectively. The *error bars* are calculated from at least three independent experiments.


A
C
705
R113
S165
V213
M268
S359
1
2
3a
3b
3c
4
5
6
7
8
9
10
11
cytosol
A141
W194
R235
R300
N
1
B
TMD1-P2
TMD1-2-P2
TMD1-3B-P2
TMD1-3C-P2
Mr
WT
N55TT
N3ST
50
35
CRM:
Insertion (%):
93 (±1)
56 (±1)
57 (±2)
100 (±0)
79 (±2)
TMD1-4-P2
TMD1-5-P2
TMD1-6-P2
55
100 (±0)
72 (±7)
82 (±2)
C
N55TT
G2
P2
S165
cytosol
TMD1-3A-P2
25
15
C P2
P2
C TMD1-3A
66 (±4)

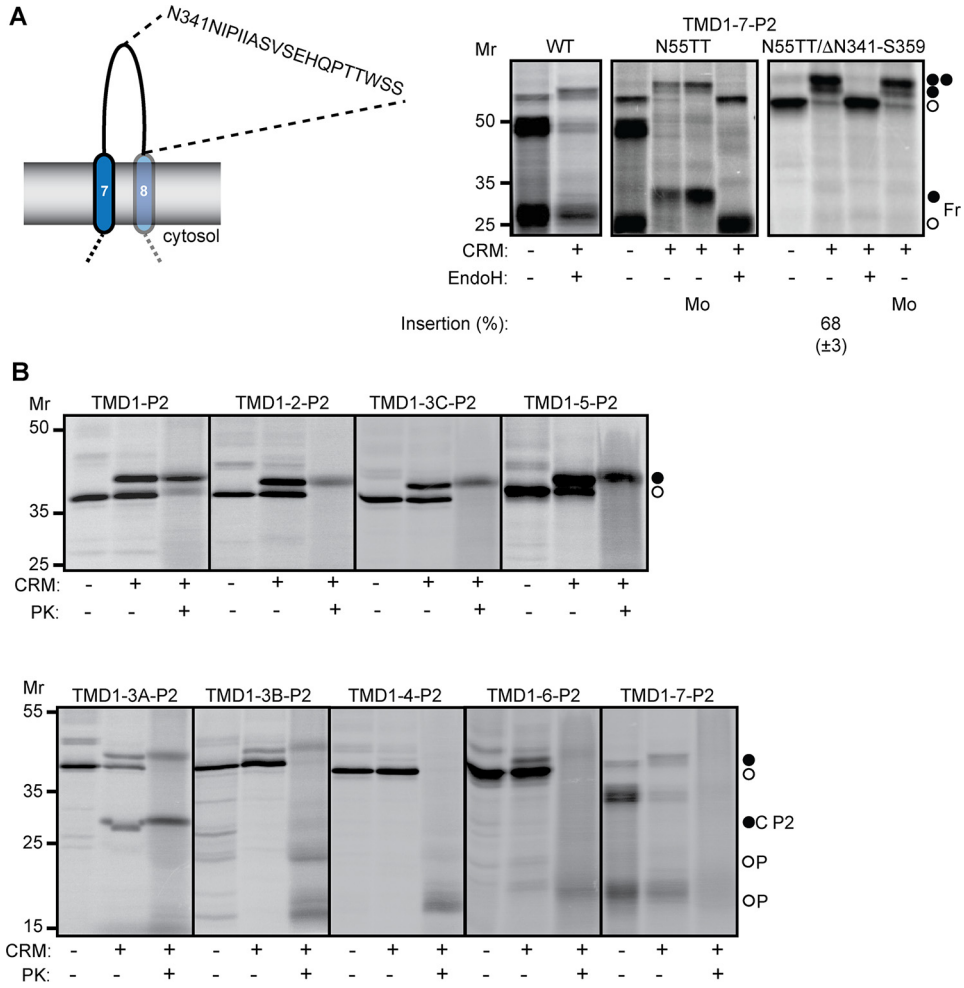


Figure 5. Verifying the N-terminal topology of mouse STT3 using Endo H and protease protection. *A*, *in vitro* translation without and with CRM of the TMD1–7-P2 fusion construct. The appearance of the shorter fragment of TMD1–7-P2 is dependent of the last 19 residues of the truncated TMD1–7, and it was confirmed by deletion of the residues, *N341-S359*. To verify the glycosylation the samples were Endo H-treated and, as a control, the same buffer conditions but without Endo H (*Mo*) was made. *B*, PK treatment after translation of all constructs, TMD1-P2 to TMD1–7-P2. The protease protection assay shows a protected lumenal and glycosylated band for TMD1-P2, TMD, TMD1–3A-P2, TMD1–3C-P2, TMD1–5-P2; only unglycosylated band is completely degraded. In the case of TMD1–3A-P2 the SPase cleaved P2 domain is also protected. The other constructs are more or less completely degraded except for a protected N-terminal fragment (*P*). Not fully translated full-length protein is indicated with *Fr*, the protected fragment is indicated with a *P*, and unglycosylated and single- and double-glycosylated products are indicated with *unfilled circles* and, respectively, *one* or *two filled circles*. The percent insertion is calculated from at least three independent experiments.

strongly into the membrane (66% and 72%, respectively). In the case of TMD1–3A-P2, two faster migrated bands could be seen that correspond to a SPase cleavage of the glycosylated fusion construct on the lumenal side of the membrane after TMD3A (Figs. 2*B* and 4*C*). The two fragments were confirmed by calculating the size of the fragments and by expression of a P2 construct (Fig. 4*C*).

The TMD1–3B-P2 and TMD1– 4-P2 protein pattern reveals a single band of mono-glycosylated protein that corresponds to a protein population with exclusively the expected topology. That is in contrast to the low insertion efficiency of the individual TMD4F (38%) and in agreement with the high insertion of TMD3BF (70%). In a similar manner, TMD1–6-P2 possibly adopts its putative topology to 82% in contrast from the individual TMD6F (22%) (Figs. 2, *B* and *C*, and 4*B*). The TMD1–7-P2 includes the long C-terminal loop containing a possible translational arrest or premature ribosome release resulting in truncated protein forms that have been previously observed (20). Deleting the last 19-amino acid residues of the truncated construct prevents this phenomenon and simplifies the quantification of glycosylated protein (Fig. 5*A*).

The very short cytosolic loops connecting TMDs of the first half of STT3-A make it possible to use a protease protection

Figure 4. A detailed study of the N-terminal topology of mouse STT3. *A*, truncations of mouse TMD1-TMD7 including loop residues between the TMDs and the P2 domain of Lep containing an engineered glycosylation site. The residues at the end of the different truncations are indicated. *B*, *in vitro* translation both without and with CRM of truncation constructs TMD1-P2 to TMD1–6-P2. The constructs with an extra glycosylation site located in the N-terminal loop or in the first lumenal loop of STT3A are indicated with ^{3}NST and ^{55}NTT. *C*, for the construct TMD1–3A-P2, two SPase cleaved fragments (*C*) are indicated dependent on size with P2 and TMD1–3A. The extra glycosylation site located in the first lumenal loop of STT3A is indicated with ^{55}NTT. Unglycosylated and single- and double glycosylated products are indicated with *unfilled circles* and, respectively, with *one* or *two filled circles*. The percent insertion is calculated from at least three independent experiments.

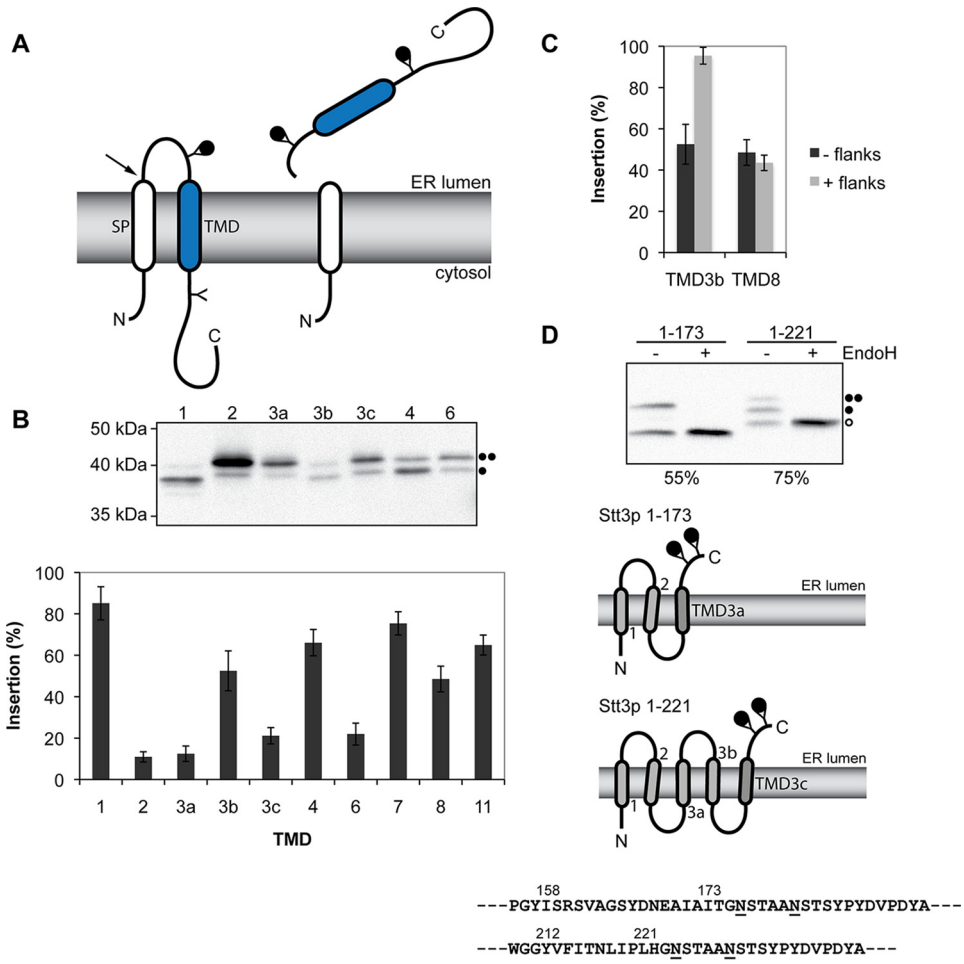


Figure 6. SP-Lep model protein and membrane insertion efficiency of yeast Stt3p TMDs. *A*, the model protein containing a cleavable signal peptide. *B*, insertion efficiency of yeast TMDs. *C*, insertion efficiency of TMD3b and TMD8 without and with flanking residues. *D*, membrane integration of two Stt3p truncations with the length of 1–173 and 1–221 amino acid residues. Sequences of the end of TMD 3a (last residue 158) and 3c (last residue 221) are shown with the *N*-linked glycosylation sites in the linker preceding the HA epitope. Endo H treatment confirmed the glycosylation of both constructs. A schematic shows the membrane topologies of the two truncations. Unglycosylated and single- and double-glycosylated products are indicated with *unfilled circles* and *one* and *two filled circles*, respectively. The percent insertion is calculated from at least three independent experiments.

assay to access if TMDs are certainly integrated into the microsomal membranes and not retained on the cytosolic side of the membrane. We utilized this to verify our results of the STT3-A P2 fusion series (Fig. 5*B*). In this case the second glycosylation site in the large lumenal domain connecting TMD1 and -2 was not present. We observed that nearly all the proteins that had the C terminus on the lumenal side of the membrane were protected, and only unglycosylated, untargeted proteins were degraded (TMD1-P2, TMD1–3A-P2, TMD1–3C-P2, and TMD1–5-P2). In contrast, for the rest of the truncated proteins (TMD1–2-P2, TMD1–3B-P2, TMD1–4-P2, and TMD1–6-P2) the P2 domain was completely degraded because of its location in the cytosol (Fig. 5*B*). In addition, for longer truncations (from TMD1–3B-P2) we observed a lower molecular weight band that corresponded to the protected N-terminal part of the protein, which was long enough to be detectable in the SDS-gel (Fig. 5*B*).

Insertion efficiency of individual TMDs and integration of the N-terminal part of yeast Stt3p

To determine the integration efficiency of individual TMDs of the yeast Stt3p, we utilized the SP-Lep model protein (32). It consists of the cleavable signal peptide (SP) of the yeast secretory protein invertase and the C-terminal domain of Lep. The Stt3p TMDs were individually introduced into the SP-Lep construct, flanked by insulating GGPG/GPGG tetrapeptides, as well as two potential glycosylation sites, up- and downstream of the TMDs (Fig. 6*A*). The N-terminal SP targets the model protein efficiently to the ER membrane and is subsequently cleaved by the membrane resident enzyme SPase. Membrane insertion of the test TMDs was assessed by a glycosylation assay. If the TMD is inserted into the membrane, only the first glycosylation site became modified, whereas translocation of the TMD into the ER lumen led to glycosylation of both glycan acceptor sites (Fig. 6*A*).

Here, we should mention that even the TMD3 of Stt3p was less defined in our earlier model (20), and hence, the three most hydrophobic stretches between TMD2 and -4 were tested in the SP-Lep model protein. SP-Lep model proteins containing TMDs 1, 2, 3a, 3b, 3c, 4, 6, 7, 8, or 11 were expressed in yeast, and the glycosylation state of the model proteins was assayed by Western blotting (Fig. 6*B*). The insertion efficiency of the individual Stt3p TMDs varied greatly, reaching from 11% (TMD2) to 85% (TMD1) (Fig. 6*B*, Table 2). The ΔG_{app} values, calculated

from the experimentally obtained insertion efficiencies, were similar to the predicted ΔG_{app} using the ΔG predictor software (21). Comparing the experimentally ΔG_{app} between mouse and yeast, however, we found some differences between some of the TMDs, such as TMDs 3a (−0.5 *versus* 1.2; mouse *versus* yeast), 3c (−0.2 *versus* 0.8)m and 11 (−1.2 *versus* −0.4) (Tables 1 and 2).

To test whether the addition of the natural flanking regions would also increase the insertion efficiency, as observed for the mouse TMDs, SP-Lep model proteins containing TMD3b or TMD8 including their natural flanks were constructed. We found that the insertion efficiency of TMD3b was altered by the presence of the natural flanks, increasing the membrane insertion from 53% to 95% (Fig. 6*C*). The integration efficiency of TMD8, however, was unaffected by the addition of its natural flanks and remained at ~45%.

To further test the membrane integration of Stt3p, two Stt3p truncations were expressed in yeast with the length of 1–173 and 1–221 amino acid residues. The sequence length 173 contains TMD1–3a, and the length 221 includes TMD1–3c both with two C-terminal NST sites. The idea was to show that both 3a and 3c are real TM segments. When expressed, both constructs became glycosylated, and hence, the C-terminal part is located in the ER, an indicator that 3a and 3c serve as TMDs. From this we can also deduce that when expressed in context of the first TMDs of Stt3p, their flanking regions, 3a and 3c, integrate better than when expressed individually in the Lep model. The glycosylation rate was ~55% for the length of 173 and 75% for amino acid sequence length 221. For Stt3p 1–221 construct, more of the single-glycosylated form appeared (Fig. 6*D*) due to the fact that the first *N*-linked glycosylation site was located too close to TMD3c (11 residues away from the end of TMD3c). If the *N*-linked glycosylation site is located at a distance of <12 residues from the end of the TMD, it is not efficiently glycosylated by the oligosaccharyltransferase (33). The possible topologies of the truncated constructs 173 and 221 are shown in Fig. 6*D*.

Discussion

In this report we present a detailed study on the membrane integration of two homologous proteins from two different organisms, the mouse STT3A and the yeast Stt3p, that have been reported to be the catalytic subunit of the OST/OT enzyme complex (11, 34). The different eukaryotic STT3 proteins are highly conserved, and they are believed to share a common reaction mechanism for *N*-linked glycosylation; even the bacterial single-subunit STT3s share the same mechanism (10, 16).

Most eukaryotic membrane proteins insert and fold in the ER membrane. The insertion process is mediated by the Sec-translocon, and both the hydrophobicity of the TMDs and the charged residues flanking the TMD play important roles during integration (6). Many multi-spanning membrane proteins are predicted to contain TMDs, which are not hydrophobic enough to insert into the ER membrane by themselves (35, 36) and will, therefore, need assistance from their neighboring TMDs to integrate into the membrane. Here, we have taken advantage of a well established *in vitro* translation system combined with ER vesicles to study the targeting, translocation, and integration process in detail (37–39). The individual ΔG-predicted TMDs of mouse STT3-A without and with flanking residues were introduced into the model protein LepH2 or LepH3 depending on their natural orientation (21–23). Remarkably, six isolated TMDs (2, 3B, 4, 6, 8, and 10) inserted poorly into the membrane from 6 to 34%. Four of them (TMD2, 3B, 8, and 10) improved their insertion efficiency by adding their natural flanking loops from 60 to 88%. On the other hand, TMD4 and TMD6 retained their low integration efficiency even after including their flanking residues (38% and 22%, respectively). TMD6 has a very high, predicted ΔG value, which reflects low integration, in contrast to TMD4, which has a rather low ΔG value. However, it is important to keep in mind that these results display how individual TMDs behave in the lipid bilayer. The N-terminal truncated results (shown later) reflect the importance of the adjacent TMDs in assisting their marginally hydrophobic neighbors for integration into the membrane (35).

Previously, we reported an 11 TMDs topology model of STT3 protein, with the uncertainty of where TMD3 was located (20). In this paper we attempt to respond to that question by studying the TMD3 region and also the whole N-terminal part (TMD1-TMD7) of the STT3 protein. We made seven truncations starting from the N-terminal of STT3 and ending at the beginning of each TMD sequence (TMD2-TMD8) and fused it to the P2 domain of LepH2 that contains a glycosylation site. Furthermore, we introduced an extra glycosylation site in the long loop between TMD1 and -2 of the STT3 protein to determine the correct topology of each truncation. Most adopted their correct topology to a high degree (>60%). TMD1-P2 inserted with the same efficiency (nearly 100%) compared with the individual TMD1F, which inserted to 85%, including both single and double glycosylation. Both single and double glycosylation were observed for the ^{55}NTT TMD1-P2-construct. One possible explanation for this could be the position of TMD1 in the membrane, hindering access to the OST so that the ^{55}NTT (Asn-Thr-Thr) site is less efficiently glycosylated, generating the singly glycosylation species (33). Another explanation could be if the very weak hydrophobic domain between TMD1 and TMD2 (Fig. 1) influence the orientation of the very short construct like TMD1-P2 to some extent and thereby also the glycosylation status. The small amounts of unglycosylated species tell us that targeting seems not to be a problem. In two cases the insertion for the truncations were almost the same as for the individual TMDs (TMD1–2-P2 and TMD1–3C-P2). Furthermore, although TMD1–3A-P2, TMD1–5-P2, and TMD1–7-P2 inserted worse than their individual counterparts, their insertion efficiency was >60%. Importantly, our results show that the poorly inserted TMD4 and -6 could fully integrate into the membrane in conjunction with their more hydrophobic TMD neighbors. It is known that marginally hydrophobic TMDs are common in multispanning membrane proteins (35, 36, 39, 40), and helices that do not insert into the membrane by themselves have been identified in both P-glycoprotein and the cystic fibrosis transmembrane conductance regulator (CFTR) (36, 39), in aguaporin-1 (AQP1) (41), and in the plant K_v channel KAT1 (42). When working with more than one TMD the possibilities of acquiring more than one topology

arise, and we can confidently exclude the less dominant topology and assign the most dominant topology by ingeniously placing glycosylation sites in the gene of STT3. Nevertheless, this can be more difficult the more TMDs that are added. Therefore, we confirmed these results by using a proteinase protection assay. If the TMDs are correctly inserted into the membrane, the connecting loops are too short to be digested by the proteinase K (PK) in the cytosol (35). The long loop connecting TMD1 and -2 is on the lumenal side and, therefore, unreachable for digestion. However, the position of the soluble Lep-P2 domain is determined by the last TMD that is studied. Depending on the location of the Lep-P2 domain, we could observe fully protected truncated proteins or protease-sensitive ones. Moreover, we only kept the glycosylation site in the P2 domain and did not add any additional in the STT3 gene. Clearly, unglycosylated species are completely degraded, whereas the inserted glycosylated proteins are protected. TMD1–4-P2 has its C terminus in the cytosol, so in this case the unglycosylated species is the fully correctly inserted protein that has been protected in agreement with the previous result (Figs. 4*B* and 5*B*).

For yeast Stt3p, TMD3 was poorly defined, and we tested membrane insertion efficiencies of three possible TMDs (3a, 3b, and 3c). Although TMD3a and -3c were not membrane-inserted by themselves, TMD3b showed an ~50% membrane insertion and much enhanced membrane efficiency when the natural flanking residues were included. However, when truncated at the residues 173 and 221, the C-terminal glycosylation sites were glycosylated in the lumen of the ER, suggesting that both 3a and 3c TMDs are located in the membrane.

In summary, we have made a thorough analysis of STT3A/Stt3p with regard to topological and structural aspect. We have found that flanking residues and neighboring TMDs have an impact on membrane integration in more than one way and to various extents. Most importantly, our data suggest that eukaryotic STT3s have the same number of TM segments as the bacterial homolog and that our data are consistent with the bacterial structural model. Future studies are warranted to obtain a structure of the eukaryotic STT3.

Experimental procedures

Mouse plasmid construction

Mouse STT3-A constructs were made as previously described (20, 25). For the membrane integration assay, double-stranded oligonucleotides encoded the mouse STT3 TMDs of the *Itm1* gene encoding STT3-A (GenBank™ accession number L34260) (20, 43). The gene sequence for TMDs alone or together with their connecting loops were introduced as amplified and purified PCR fragments (QIAquick PCR Purification kit (Hilden, Germany)) with an N-terminal Spe1 and a C-terminal Kpn1 restriction site into the previously described modified *lepB* gene in pGEM1 (Promega, Madison, WI) (22, 23). The PCR fragments were obtained using Phusion High-Fidelity DNA polymerase (Finnzymes, Espoo, FI) or Expand Long Template PCR System (Roche Diagnostics) with primers complementary to the 5′ and 3′ ends of the selected part of the gene. For the study of the first half of the STT3-A protein, the gene sequence for the truncations was introduced with an N-terminal Xba1 and a C-terminal Kpn1 restriction site into the *lepB* gene. Glycosylation acceptor sites (N*X*(S/T/C) found in both genes (*lepB* and *Itm1*) were mutated to Q*X*(S/T/C), and extra glycosylation acceptor sites were introduced using the QuikChange™ Site-Directed Mutagenesis protocol (Stratagene) by replacing one, two, or three appropriately positioned codons with codons for the acceptor tripeptide N*X*T (Asn-*X*-Thr). All constructs were confirmed by sequencing of plasmid DNA at Eurofins MWG Operon (Ebersberg, Germany) or BM labbet AB (Furulund, Sweden). All cloning steps were done according to standard procedures using restriction enzymes from Promega or Fermentas (Burlington, Ontario, CA).

In vitro expression of mouse STT3-A

STT3-A constructs cloned in pGEM1 vector were transcribed and translated in the TNT® SP6 Quick Coupled System (Promega) in the presence and absence of column-washed rough microsomes (CRM) of pancreas from dog (tRNA Probes, TX) (44). 10 μl of reticulocyte lysate, 150–200 ng of DNA template, 1 μl of [^{35}S]Met (5 μCi), and 0.5 μl of CRM were mixed and incubated at 30 °C for 90 min.

For Endo H treatment 9 μl of the TNT reaction was mixed with 1 μl of 10× glycoprotein denaturing buffer. After the addition of 1 μl of Endo H (500,000 units/ml; New England BioLabs), 7 μl of distilled H_2O, and 2 μl of 10× GlycoBuffer 3 reaction buffer, the sample was incubated at 37 °C for 1 h (23).

PK treatment was performed by adding 1 μl of $CaCl_2$ (200 mM) and 0.2 μl of Proteinase K (4.5 units/μl) to the translation reaction. After incubating on ice for 30 min, 1 μl of PMSF (20 mM ethanolic solution) was added to inactivate PK and further incubated on ice for 5 min (35).

Translation products were analyzed by SDS-PAGE gels and visualized in by Fuji FLA-3000 phosphorimaging (Fujifilm) using the Image Reader V1.8J/Image Gauge V 3.45 software. The MultiGauge (Fujifilm) software was used to generate a one-dimensional intensity profile of each gel lane and the multi-Gaussian fit program (Qtiplot, www.qtiplot.com)[4] was used to calculate the peak areas of the glycosylated protein bands. The membrane insertion of each studied construct was calculated as the quotient between the peak area of the single- or double-glycosylated band and the peak area of double- or single-glycosylated protein band, respectively (depending on which Lep construct was used). On average, the glycosylation levels vary by no more than ±5% between repeated experiments.

Yeast plasmid construction and transformation

All yeast plasmids were constructed from p424GPD-SP-Lep (23, 32). To substitute the sequence coding for the hydrophobic (H) segment of the SP-Lep model protein with various yeast Stt3p TMDs, with or without adjacent loops, a SmaI site was inserted within the H-segment by site-directed mutagenesis. Stt3p TMDs were amplified from genetic DNA using primers which include sequences that complement the sequences upstream and downstream of the H-segment. The PCR products were inserted into the SmaI-linearized p424GPD-SP-Lep by homologous recombination. Yeast transformants were selected on −Trp plates, plasmids were isolated, and the cor-

rect sequence was confirmed by sequencing. Stt3p 1–173 and Stt3p 1–221 constructs were prepared by amplifying the indicated fragments by PCR and subcloning into p424GPDHA vector (23) carrying the engineered two *N*-linked glycosylation sites (Fig. 6*D*). For both cloning and expression the yeast strain W303-1α (*MATα*, *ade2*, *can1*, *his3*, *leu2*, *trp1*, *ura3*) (45) were used.

Western blotting analysis of yeast Stt3p

Protein preparation and Western blotting was carried out as described in Reithinger *et al.* (46). In brief, model protein-expressing yeast cells were grown overnight in the appropriate selective media, harvested by centrifugation, washed with distilled H_2O, and resuspended in SDS-PAGE sample buffer. After incubation at 60 °C for 15 min and centrifugation, the supernatant was loaded onto SDS-gels and subjected to Western blotting using mouse anti-HA antiserum (Covance). To remove glycans, samples of the whole-cell lysates were treated with Endo H (Roche Diagnostics) before SDS-PAGE as previously described (46).

Author contributions—P. L., K. Ö., J. R., H. K., and I. N. planned the experiments. P. L., K. Ö, J. R., A. Ho., Y. M., A. Ha., and D. M. performed the experiments. P. L., K. Ö., J. R., H. K., and I. N. analyzed the data. P. L., K. Ö., J. R., H. K., and I. N. wrote the paper.

Acknowledgments—We gratefully thank Prof. William Lennarz for providing yeast strains and plasmids, Prof. Arthur E. Johnson for providing rough microsomes, and Prof. Joseph Merregaert for providing a mouse STT3 plasmid (Itm1 cDNA).

References

1. Devaraneni, P. K., Conti, B., Matsumura, Y., Yang, Z., Johnson, A. E., and Skach, W. R. (2011) Stepwise insertion and inversion of a type II signal anchor sequence in the ribosome-Sec61 translocon complex. *Cell* **146,** 134–147
2. De Marothy, M. T., and Elofsson, A. (2015) Marginally hydrophobic transmembrane α-helices shaping membrane protein folding. *Protein Sci.* **24,** 1057–1074
3. 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
4. Martínez-Gil, L., Saurí, A., Marti-Renom, M. A., and Mingarro, I. (2011) Membrane protein integration into the endoplasmic reticulum. *FEBS J.* **278,** 3846–3858
5. Shao, S., and Hegde, R. S. (2011) Membrane protein insertion at the endoplasmic reticulum. *Annu. Rev. Cell Dev. Biol.* **27,** 25–56
6. Park, E., and Rapoport, T. A. (2012) Mechanisms of Sec61/SecY-mediated protein translocation across membranes. *Annu. Rev. Biophys.* **41,** 21–40
7. Gilmore, R., and Mandon, E. C. (2012) Understanding integration of α-helical membrane proteins: the next steps. *Trends Biochem. Sci.* **37,** 303–308
8. Kelleher, D. J., and Gilmore, R. (2006) An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* **16,** 47R–62R
9. Kelleher, D. J., Karaoglu, D., Mandon, E. C., and Gilmore, R. (2003) Oligosaccharyltransferase isoforms that contain different catalytic STT3 subunits have distinct enzymatic properties. *Mol. Cell* **12,** 101–111
10. Wacker, M., Linton, D., Hitchen, P. G., Nita-Lazar, M., Haslam, S. M., North, S. J., Panico, M., Morris, H. R., Dell, A., Wren, B. W., and Aebi, M. (2002) *N*-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science* **298,** 1790–1793
11. Zufferey, R., Knauer, R., Burda, P., Stagljar, I., te Heesen, S., Lehle, L., and Aebi, M. (1995) STT3, a highly conserved protein required for yeast oligosaccharyltransferase activity in vivo. *EMBO J.* **14,** 4949–4960
12. Yoshida, S., Ikeda, E., Uno, I., and Mitsuzawa, H. (1992) Characterization of a staurosporine- and temperature-sensitive mutant, stt1, of *Saccharomyces cerevisiae*: STT1 is allelic to PKC1. *Mol. Gen. Genet.* **231,** 337–344
13. Lizak, C., Gerber, S., Numao, S., Aebi, M., and Locher, K. P. (2011) X-ray structure of a bacterial oligosaccharyltransferase. *Nature* **474,** 350–355
14. Igura, M., and Kohda, D. (2011) Selective control of oligosaccharide transfer efficiency for the *N*-glycosylation sequon by a point mutation in oligosaccharyltransferase. *J. Biol. Chem.* **286,** 13255–13260
15. Jaffee, M. B., and Imperiali, B. (2011) Exploiting topological constraints to reveal buried sequence motifs in the membrane-bound *N*-linked oligosaccharyltransferases. *Biochemistry* **50,** 7557–7567
16. Wacker, M., Feldman, M. F., Callewaert, N., Kowarik, M., Clarke, B. R., Pohl, N. L., Hernandez, M., Vines, E. D., Valvano, M. A., Whitfield, C., and Aebi, M. (2006) Substrate specificity of bacterial oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and eukaryotic systems. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 7088–7093
17. Cherepanova, N., Shrimal, S., and Gilmore, R. (2016) *N*-linked glycosylation and homeostasis of the endoplasmic reticulum. *Curr. Opin. Cell Biol.* **41,** 57–65
18. Cherepanova, N. A., and Gilmore, R. (2016) Mammalian cells lacking either the cotranslational or posttranslocational oligosaccharyltransferase complex display substrate-dependent defects in asparagine-linked glycosylation. *Sci. Rep.* **6,** 20946
19. Burda, P., and Aebi, M. (1999) The Dolichol pathway of *N*-linked glycosylation. *Biochim. Biophys. Acta* **1426,** 239–257
20. Kim, H., von Heijne, G., and Nilsson, I. (2005) Membrane topology of the STT3 subunit of the oligosaccharyltransferase complex. *J. Biol. Chem.* **280,** 20261–20267
21. 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
22. 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
23. 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
24. Ojemalm, K., Halling, K. K., Nilsson, I., and von Heijne, G. (2012) Orientational preferences of neighboring helices can drive ER insertion of a marginally hydrophobic transmembrane helix. *Mol. Cell* **45,** 529–540
25. Johansson, M., Nilsson, I., and von Heijne, G. (1993) Positively charged amino acids placed next to a signal sequence block protein translocation more efficiently in *Escherichia coli* than in mammalian microsomes. *Mol. Gen. Genet.* **239,** 251–256
26. Nilsson, I., Johnson, A. E., and von Heijne, G. (2002) Cleavage of a tail-anchored protein by signal peptidase. *FEBS Lett.* **516,** 106–108
27. Sato, C., Kim, J. H., Abe, Y., Saito, K., Yokoyama, S., and Kohda, D. (2000) Characterization of the *N*-oligosaccharides attached to the atypical Asn-X-Cys sequence of recombinant human epidermal growth factor receptor. *J. Biochem.* **127,** 65–72
28. Chi, Y. H., Koo, Y. D., Dai, S. Y., Ahn, J. E., Yun, D. J., Lee, S. Y., and Zhu-Salzman, K. (2010) *N*-Glycosylation at non-canonical Asn-X-Cys sequence of an insect recombinant cathepsin B-like counter-defense protein. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **156,** 40–47
29. Gil, G. C., Velander, W. H., and Van Cott, K. E. (2009) *N*-Glycosylation microheterogeneity and site occupancy of an Asn-X-Cys sequon in plasma-derived and recombinant protein C. *Proteomics* **9,** 2555–2567
30. Miletich, J. P., and Broze, G. J., Jr. (1990) Beta protein C is not glycosylated at asparagine 329: the rate of translation may influence the frequency of usage at asparagine-X-cysteine sites. *J. Biol. Chem.* **265,** 11397–11404
31. Yasuda, D., Imura, Y., Ishii, S., Shimizu, T., and Nakamura, M. (2015) The atypical *N*-glycosylation motif, Asn-Cys-Cys, in human GPR109A is re-

quired for normal cell surface expression and intracellular signaling. *FASEB J.* **29,** 2412–2422
32. 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
33. Nilsson, I. M., and von Heijne, G. (1993) Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. *J. Biol. Chem.* **268,** 5798–5801
34. Nilsson, I., Kelleher, D. J., Miao, Y., Shao, Y., Kreibich, G., Gilmore, R., von Heijne, G., and Johnson, A. E. (2003) Photocross-linking of nascent chains to the STT3 subunit of the oligosaccharyltransferase complex. *J. Cell Biol.* **161,** 715–725
35. Hedin, L. E., Ojemalm, K., Bernsel, A., Hennerdal, A., Illergård, K., Enquist, K., Kauko, A., Cristobal, S., von Heijne, G., Lerch-Bader, M., Nilsson, I., and Elofsson, A. (2010) Membrane insertion of marginally hydrophobic transmembrane helices depends on sequence context. *J. Mol. Biol.* **396,** 221–229
36. Sadlish, H., and Skach, W. R. (2004) Biogenesis of CFTR and other polytopic membrane proteins: new roles for the ribosome-translocon complex. *J. Membr. Biol.* **202,** 115–126
37. Cuviello, F., Tellgren-Roth, Å., Lara, P., Ruud Selin, F., Monné, M., Bisaccia, F., Nilsson, I., and Ostuni, A. (2015) Membrane insertion and topology of the amino-terminal domain TMD0 of multidrug-resistance associated protein 6 (MRP6). *FEBS Lett.* **589,** 3921–3928
38. Wanngren, J., Lara, P., Ojemalm, K., Maioli, S., Moradi, N., Chen, L., Tjernberg, L. O., Lundkvist, J., Nilsson, I., and Karlström, H. (2014) Changed membrane integration and catalytic site conformation are two mechanisms behind the increased Aβ42/Aβ40 ratio by presenilin 1 familial Alzheimer-linked mutations. *FEBS Open Bio.* **4,** 393–406
39. 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
40. Anthony, V., and Skach, W. R. (2002) Molecular mechanism of P-glycoprotein assembly into cellular membranes. *Curr. Protein Pept. Sci.* **3,** 485–501
41. Pitonzo, D., and Skach, W. R. (2006) Molecular mechanisms of aquaporin biogenesis by the endoplasmic reticulum Sec61 translocon. *Biochim. Biophys. Acta* **1758,** 976–988
42. Sato, Y., Sakaguchi, M., Goshima, S., Nakamura, T., and Uozumi, N. (2002) Integration of Shaker-type K+ channel, KAT1, into the endoplasmic reticulum membrane: synergistic insertion of voltage-sensing segments, S3-S4, and independent insertion of pore-forming segments, S5-P-S6. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 60–65
43. Hong, G., Deleersnijder, W., Kozak, C. A., Van Marck, E., Tylzanowski, P., and Merregaert, J. (1996) Molecular cloning of a highly conserved mouse and human integral membrane protein (Itm1) and genetic mapping to mouse chromosome 9. *Genomics* **31,** 295–300
44. Walter, P., and Blobel, G. (1983) Preparation of microsomal membranes for cotranslational protein translocation. *Methods Enzymol.* **96,** 84–93
45. Wilkinson, B. M., Tyson, J. R., and Stirling, C. J. (2001) Ssh1p determines the translocation and dislocation capacities of the yeast endoplasmic reticulum. *Dev. Cell* **1,** 401–409
46. Reithinger, J. H., Kim, J. E., and Kim, H. (2013) Sec62 protein mediates membrane insertion and orientation of moderately hydrophobic signal anchor proteins in the endoplasmic reticulum (ER). *J. Biol. Chem.* **288,** 18058–18067
47. Stamm, M., Staritzbichler, R., Khafizov, K., and Forrest, L. R. (2014) AlignMe–a membrane protein sequence alignment web server. *Nucleic. Acids Res.* **42,** W246–W251
48. Stamm, M., Staritzbichler, R., Khafizov, K., and Forrest, L. R. (2013) Alignment of helical membrane protein sequences using AlignMe. *PLoS ONE* **8,** e57731
49. Khafizov, K., Staritzbichler, R., Stamm, M., and Forrest, L. R. (2010) A study of the evolution of inverted-topology repeats from LeuT-fold transporters using AlignMe. *Biochemistry* **49,** 10702–10713