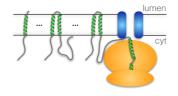
Membrane insertion and topology of the transloconassociated protein (TRAP) gamma subunit

Manuel Bañó-Polo, Carlos A. Martínez-Garay, Brayan Grau, Luis Martínez-Gil and Ismael Mingarro *

Departament de Bioquímica i Biologia Molecular, Estructura de Recerca Interdisciplinar en Biotecnologia i Biomedicina (ERI BioTecMed),

Universitat de València.

E-46 100 Burjassot, Spain



 $^*\ Corresponding\ author.\ Phone:\ Int+34-963543796.\ Fax:\ Int+34-963544635.$

E-mail: Ismael.Mingarro@uv.es

Running title: TRAP-γ insertion and topology

Manuscript information: 26 text pages; 1 table; 4 figures

Abstract

Translocon-associated protein (TRAP) complex is intimately associated with the ER translocon for the insertion or translocation of newly synthesised proteins in eukaryotic cells. The TRAP complex is comprised of three single-spanning and one multiple-spanning subunits. We have investigated the membrane insertion and topology of the multiple-spanning TRAP- γ subunit by glycosylation mapping and green fluorescent protein fusions both *in vitro* and in cell cultures. Results demonstrate that TRAP- γ has four transmembrane (TM) segments, an Nt/Ct cytosolic orientation and that the less hydrophobic TM segment inserts efficiently into the membrane only in the cellular context of full-length protein.

Keywords: Endoplasmic reticulum; membrane protein topology; *N*-linked glycosylation; Sec61 complex; translocon-associated proteins

Abbreviations:

ER, endoplasmic reticulum; GFP, green fluorescent protein; HR, hydrophobic region; Lep, leader peptidase; SDS-PAGE, sodium dodecylsulfate polyacrylamide gel electrophoresis; SRP, signal recognition particle; TM, transmembrane; TRAP, translocon-associated protein.

1. Introduction

In eukaryotic cells insertion of most integral membrane proteins into the endoplasmic reticulum (ER) membrane occurs primarily in a co-translational manner. In this process, targeting of the ribosome-mRNA-nascent chain complex to the ER depends on the signal recognition particle (SRP) and its interaction with the membrane-bound SRP receptor. The ribosome and the nascent chain are then transferred to the translocon, a multi-protein complex that facilitates insertion of integral membrane proteins into the lipid bilayer and translocation of secreted proteins across the lipid bilayer [1]. Translocons are not passive pores in the bilayer, but instead are dynamic complexes that cycle between ribosome-bound and ribosome-free states, and convert between translocation and membrane integration modes of action, while maintaining the membrane permeability barrier [2, 3]. The core components of the mammalian translocon are the Sec61 α , β and γ subunits [4] and the translocating chain-associating membrane protein [5, 6]. However, several other proteins including the signal peptidase, the oligosaccharyltransferase (OST), PAT-10, RAMP4, BAP31 or the translocon-associated protein (TRAP) complex, interact at some point with the core translocon modulating its activity [1].

The TRAP complex is comprised of four membrane protein subunits. The α , β , and δ -subunits are single-spanning proteins with suggested large N-terminal (Nt) luminal and small C-terminal (Ct) cytosolic domains, while γ -subunit was proposed to cross the membrane four times with a putative Nt/Ct luminal orientation [7]. This complex remains stably associated with detergent-solubilised ribosome-translocon complexes [8], being an integral part of the translocon. Despite notorious efforts, the role of TRAP complex during membrane insertion or translocation is not fully understood. It has been proposed that TRAP complex acts, in a substrate-specific manner, facilitating the initiation of protein translocation [9]. It has also been suggested a role of TRAP complex in membrane topology regulation by moderating the 'positive-inside' rule [10] of membrane proteins with weak topogenic determinants [11]. The TRAP complex is also involved in ER associated degradation, where its function might be to recruit misfolded proteins to the translocon [12]. Most of these studies have been carried out

using *in vitro* reconstituted systems, mainly with microsomal membranes. Its role in membrane protein biogenesis has been emphasised *in vivo*, where it has been shown that TRAP- α subunit mutant alters mammalian heart development [13], a mutation in TRAP- δ has been related to human congenital disorders of glycosylation [14], and TRAP- γ subunit plays an important role in the process of pronephros differentiation during *Xenopus* development [15], and appears to be required for vascular network formation in murine placental development [16]. Despite the general importance of the TRAP complex in secreted and membrane protein biogenesis, and the particular involvement of TRAP- γ subunit in development, its membrane topology has not been thoroughly investigated.

In this study, we investigated TRAP- γ subunit biogenesis both *in vitro* and *in vivo* using *N*-glycosylation- and fluorescence-based topological reporters. Our results reveal that TRAP- γ is a tetra-spanning membrane protein with an Nt/Ct cytosolic orientation. Furthermore, we identified the precise position of all four TM segments and established the topology of the protein at multiple stages of its biogenesis by translating a series of TRAP- γ truncations containing appropriate reporter tags. The results suggest that the membrane integration of some poorly hydrophobic transmembrane (TM) segments is only effective within the global context of the full-length TRAP- γ protein in a cellular environment.

2. Materials and methods

2.1. Enzymes and chemicals.

All enzymes, as well as plasmid pGEM1, RiboMAX SP6 RNA polymerase system and rabbit reticulocyte lysate were from Promega (Madison, WI). ER rough microsomes from dog pancreas were from tRNA Probes (College Station, TX, USA). [*S]Met and [*C]-methylated markers were from Perkin Elmer. Restriction enzymes and Endoglycosidase H were from Roche Molecular Biochemicals. Proteinase K was from Sigma-Aldrich (St Louis, MO). The DNA plasmid, RNA clean up and PCR purification kits were from Thermo Fisher Scientific (Ulm, Germany). All oligonucleotides were purchased from IDT (USA).

2.2. Computer-assisted Analysis of TRAP-γ sequence.

Prediction of TM helices was done using a pletora of the most common methods available on the Internet: ΔG Prediction Server [17, 18] (http://www.cbr.su.se/DGpred/), **HMMTOP** [19] (http://www.enzim.hu/hmmtop/), **MEMSAT** [20] (http://www.bioinf.cs.ucl.ac.uk/psipred/psiform.html), **PHD** [21] (http://www.cubic.bioc.columbia.edu/pp/) **SOSUI** [22] **SPLIT** (http://bp.nuap.nagoya-u.ac.jp/sosui/), [23] (http://split.pmfst.hr/split/4/), **TMHMM** [24] (http://www.cbs.dtu.dk/services/TMHMM), **TMPred** (http://www.ch.embnet.org/software/TMPRED_form.html), OCTOPUS [25] (http://octopus.cbr.su.se/), TOPCONS [26] (http://topcons.net/) and TopPRED [27] (http://www.sbc.su.se/~erikw/toppred2/).

2.3. DNA Manipulation.

The full-length TRAP- γ sequence was cloned from a mouse brain cDNA library into pGEM-T plasmid as previously described [6]. Insertion of T111 and mutation G5N were performed using the QuikChange PCR mutagenesis kit from Stratagene (La Jolla, CA) following the manufacturer's protocol. For construction of plasmids encoding glycosylatable GFP (gGFP) fusion proteins TRAP- γ sequence was subcloned into the yeast HA-gGFP vector [28], which contained the GFP sequence with an additional glycosylation site, between the *PstI/BamHI* restriction sites, to obtain a construct containing the TRAP- γ subunit with gGFP fused in frame at the C-terminus. All DNA manipulations were confirmed by sequencing of plasmid DNAs at Macrogen Company (Seoul, South Korea).

2.4. In vitro transcription and translation.

Constructs cloned in pGEM-T were transcribed and translated in the presence of [*S]Met/Cys as previously reported [29, 30]. In all cases, after translation membranes were collected by ultracentrifugation and analysed by sodium-dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Finally, autoradiographies of the gels were visualized on a Fuji FLA3000 phosphorimager and Image Reader 8.1j software.

2.5. In vitro Endoglycosidase H and Proteinase K treatment.

Endoglycosidase H (EndoH, Roche) treatment was done as previously decribed [31]. For the proteinase K protection assay, the translation mixture was supplemented with 1 μ L of 50 mM CaCl₂ and 1 μ L of proteinase K (2 mg/mL), then digested for 40 min on ice. The reaction was stopped by adding 2 mM PMSF before SDS-PAGE analysis [32].

2.6. Protein preparation from yeast and Western blot analysis.

Yeast cells transformed with plasmids mentioned above, were grown in SD-Trp (synthetic dextrose) medium [0.67% yeast nitrogen base, synthetic mixture-TRP (drop-out) and 2% glucose]. Approximately 10^s yeast cells were resuspended in 50 mM Tris/HCl (pH 7.5) supplemented with a mixture of protease inhibitors and broken with vigorous shaking in the presence of glass beads. The lysate was removed from the intact cells and the debris by centrifuged 2000g for 5 min at 4°C. The supernatant was denatured by adding 1 volume of denaturation buffer [1mM Tris/HCl (pH 7.5), 40 m M DTT and 0.5% SDS] and incubated for 5 min at 95°C. Samples were diluted 2-fold (final concentration 0.25% SDS) and adjusted to a final concentration of 0.2 mM ammonium acetate (pH 5.5). The sample was divided into two aliquots: one was treated with 50 mU of EndoH; and the other incubated with buffer at 37°C for 2 h. Proteins were resolved by SDS-PAGE (10% gel), immunoblotted and detected using the corresponding antibody (anti-GFP monoclonal antibody, Roche) and the SuperSignal ® Western blotting detection kit (Thermo Scientific) according to the manufacturer's instructions.

2.7. Fluorescence microscopy.

GFP tagged constructs were analysed in living cells grown overnight on selective medium SD-Trp. Samples were visualized in an Axioskop 2 Fluorescence Microscope (Zeiss Inc.), and pictures were taken using an AxioCam MRm (Zeiss Inc.).

2.8. TRAP-γ expression in mammalian cells.

TRAP- γ sequence was fused to a FLAG-tag sequence and inserted in a pSELECT-puromicin plasmid. Once sequenced verified the plasmid was transfected into HEK 293-T cell using Lipofectamine 2000 (Life Technologies) according to the manufacturer's instructions. Approximately 24 hours post-transfection cells were collected and washed with PBS. After a short centrifugation (1000 rpm for 5 minutes on a table-top centrifuge) cells were

lysed by adding 300 μ L of Lysis Buffer (30 mM Tris-HCl, 150 mM NaCl, 0.5% NP-40, Protease inhibitor (complete Protease Inhibitor Cocktail, Roche) and phosphatase inhibitor (PhosSTOP, Roche). After 20 min incubation on ice, total protein was quantified (Bradford) and equal amounts of protein submitted to Endo H treatment or mock treated, followed by SDS-PAGE analysis and transferred into an PVDF transfer membrane (ThermoFisher Scientific). Protein glycosylation status was analysed by Western Blott using an anti-flag antibody.

3. Results

3.1. Identification and location of TRAP- γ subunit transmembrane segments.

Early work predicted TRAP-γ subunit to span the ER membrane four times with an Nt/Ct lumenal orientation [7]. First, as a preliminary test, we analysed TRAP-γ sequence for TM segments using a plethora of membrane protein prediction algorithms. As shown in Table 1, the predicted outcome displayed variability according to the methods used, likely due to the limited hydrophobicity of the most C-terminal predicted membrane-spanning segment, which contains several hydrophilic residues, particularly amino acids with hydroxyl groups (see Fig. 1A). Next, the membrane insertion capacity for the predicted membrane-spanning segments was analysed using an in vitro assay based on the E. coli leader peptidase (Lep). Lep consists of two TM segments (H1 and H2) connected by a cytoplasmic loop (P1) and a large C-terminal domain (P2), and inserts into ER-derived microsomal membranes with an Nt/Ct luminal orientation (Fig. 1B). This assay allows accurate analysis and quantification of the membrane insertion capability of short sequences. The hydrophobic region (HR) to be tested is inserted into the luminal P2 domain flanked by two N-glycosylation acceptor sites (G1 and G2, tripeptide Asn-X-Ser/Thr sequences, where X can be any amino acid except Pro) [17, 18]. N-linked glycosylation has been extensively used as topological reporter for more than two decades [33]. This post-translational modification occurs only in the lumen of the ER because the active site of OST, a translocon-associated enzyme responsible for oligosaccharide transfer, is located there [34]; thus, no N-linked glycosylation occurs in polypeptide

sequences spanning the membrane or facing the cytosol. When the tested HR region is recognized by the translocon as a TM segment and inserted into the membrane, only G1 (not G2) localizes in the ER lumen rendering a single-glycosylated Lep version (Figure 1B, left). On the other hand, if the tested region is fully translocated across the lipid bilayer both G1 and G2 sites locate in the ER lumen and are, subsequently, glycosylated by the OST (Fig. 1B, right). Glycosylation of an acceptor site increases the molecular mass of the protein by ≈ 2.5 kDa relative to the observed molecular mass in the absence of membranes, allowing monitoring the insertion into the membrane of the tested HR by SDS-PAGE.

Translation of Lep chimeric constructs harbouring the TRAP-γ TM regions predicted by the ΔG Prediction Server [residues 30-51 (HR1, $\Delta G^{\text{pred}}_{\text{app}} =$ -0.89 kcal/mol), 55-77 (HR2, $\Delta G_{\text{pred}_{\text{app}}} = +0.97 \text{ kcal/mol}$), 138-160 (HR3, $\Delta G_{\text{pred}_{\text{app}}} =$ -1.12 kcal/mol) and 163-182 (HR4, $\Delta G^{\text{pred}}_{\text{app}} = +1.78 \text{ kcal/mol}$), Figure 1A] resulted mainly in single-glycosylated forms (HR1, HR2 and HR3) except for HR4 (Figure 1C, lane 11). In this last case, the Lep construct containing HR4 is found mostly doubly-glycosylated. These results were confirmed by proteinase K (PK) treatment. Digestion with PK degrades membrane protein domains located exclusively towards the cytosol, while membrane-embedded or lumenally exposed domains are protected. As expected, Lep chimeras bearing HR1, HR2 and HR3 regions were sensitive to PK digestion (Figure 1C, lanes 3, 6 and 9). However, Lep constructs containing TRAP-γ HR4 sequence were partially resistant to the protease treatment due to its luminal P2 localization (Figure 1C, lane 12, arrowhead, expected size ≈33.5 kDa). It should be mentioned that TRAP- γ sequence includes a native potential Nglycosylation site at position 141, i.e. within HR3 (see Fig. 1A), the membraneembedded location of this site precludes its modification (Fig. 1C, lanes 7-9).

In our *in silico* analysis, the algorithms predicting membrane protein topology agree on an Nt cytoplasmic orientation for TRAP-γ subunit (Table 1). In this scenario, HR1 and HR3 would adopt an Nt-cytosolic/Ct-luminal orientation when inserted into the lipid bilayer. That is, the opposite orientation than in the Lep system, where they are forced to adopt an Nt-luminal/Ct-cytosolic disposition (Fig. 1B). To test the membrane insertion of these HRs in its predicted orientation, we used a Lep-derived system (Lep') (Fig. 2A) [29, 35]. In this assay, the HR to be tested replaces Lep H2 domain.

Then, the second glycosylation site (G2') remains located at P2, while the first acceptor site (G1') has been embedded in an extended N-terminal sequence. Contrarily to the original Lep system a doubly glycosylated Lep' connotes membrane insertion of the tested HR, whilst single glycosylation implies that the tested hydrophobic segment has not been recognized as a TM domain by the translocon machinery. As shown in Figure 2B, TRAP- γ HR1 and HR3 inserted efficiently into the microsomal membranes (lanes 2 and 5 respectively). These results were confirmed by PK treatment. It is worth noting that in Lep' constructs P2 protection from proteinase K treatment (expected size \approx 32.5 kDa) denotes insertion of the tested region (Figure 2B, lanes 3 and 6).

TRAP-γ HR1 and HR2 are connected by a small extra-membranous region of 3 residues, similarly HR3 and HR4 are separated by just 2 amino acid residues (Fig. 1A). Furthermore, the residues in these two short connecting regions display high turn-inducing propensities [36], suggesting the presence of two α -helical TM hairpins. It has been observed that a neighbouring TM helix can promote the insertion of a flanking poorly hydrophobic TM segment [29, 32, 37, 38]. Therefore, the low insertion capacity observed for HR4 might increase in the presence of its preceding HR3 domain. To test this hypothesis, we analysed the insertion of the full-length putative α -helical hairpin regions (residues 30-77 or 138-182) in the Lep' system, since Lep' allows the insertion of both hairpins in the predicted orientations (Table 1 and Fig. 2C). As expected, according to the observed insertion efficiency for isolated HR1 and HR2 in Lep and Lep' systems, translation of Lep' constructs harbouring HR1-HR2 hairpin (residues 30-77) generated single-glycosylated forms, indicating that both HRs are efficiently inserted as an α -helical hairpin into the membrane (Fig. 2D, lane 2). However, Lep' constructs harbouring putative HR3-HR4 hairpin (residues 138-182) were double-glycosylated (Fig. 2D, lane 5), denoting failure of insertion for one of the two HRs, most likely HR4 according to previous results (Figs. 1C and 2B). PK treatment of these translation reaction mixtures produced complete digestion of HR1-HR2 containing molecules (Fig. 2D, lane 3), and a band corresponding to P1-digested molecules (Fig. 2D, lane 6, expected size ≈34.6 kDa), which contained HR3-HR4 region plus the single-glycosylated P2 domain (Fig. 2C, right), supporting the low insertion efficiency observed.

3.2. TRAP- γ subunit inserts into ER-derived membranes with an Nt/Ct cytosolic orientation

To experimentally map the membrane topology of TRAP-γ subunit, we prepared a series of polypeptide truncates containing an added C-terminal glycosylation tag (NST), which has proven to be efficiently modified in the in vitro reconstituted translation system [30, 39]. The constructs used are shown in Fig. 3A. Wild type TRAP-y subunit carries a native potential glycosylation site at position 23 (see Fig. 1A). It has been previously reported that glycosylation occurs when the acceptor Asn is at least ~11-13 residues away from the membrane [40-42]. Therefore, since TM1 starts around position 30 according to all the prediction algorithms (Table 1) and our experimental data, in case of a TM disposition Asn23 should not be modified independently of the location of TRAP-γ N-terminus. However, in case of translocation across the microsomal membrane, both Asn23 and the Cterminal glycosylation tag should be modified rendering doubly-glycosylated forms. As shown in Fig. 3B, translation products containing the N-terminal 54 residues of TRAP-y sequence, including the first predicted TM segment (TM1) plus an optimized glycosylatable C-terminal tag (truncated 54mer polypeptides), were efficiently singly-glycosylated in the presence of microsomal membranes (lane 2). The nature of these higher molecular weight polypeptide species was analysed by endoglycosidase H (EndoH) treatment, a highly specific enzyme that cleaves N-linked oligosaccharides regardless of their localization. Treatment with EndoH eliminated higher molecular mass band (Fig. 3B, lanes 1), confirming the sugar source of their retarded electrophoretic mobility and suggesting 54mer insertion into the microsomal membrane with an N-terminal cytoplasmic orientation (see Fig. 3D for a scheme). To corroborate these results, 54mer translations in the presence of microsomal membranes were performed in parallel either with the above described C-terminal acceptor tag (NST, Fig. 3C lanes 1 and 2) or with a nonglycosylatable tag, in which the acceptor Asn was replaced by a non-acceptor Gln (QST, Fig. 3C lanes 3 and 4). The efficient glycosylation observed only when using the C-terminal acceptor site (Fig. 3C, lane 2), strongly indicate that TM1 in the TRAP-y context is acting as a non-cleavable signal sequence and is properly recognized by the translocon machinery to be inserted into the membrane with its N-terminus facing the cytosol.

Truncated 132mer polypeptides, which include the N-terminal α -helical hairpin, were poorly glycosylated (13 \pm 5% of glycosylation, Fig. 3B lane 4), indicating that the second predicted TM segment (TM2) efficiently integrates into the membrane (see Fig. 3D top for an illustrative scheme). The insertion of the third predicted TM segment (TM3) was tested by translating a 161-residue truncation with the same C-terminal glycosylatable tag (161mer). The glycosylation of this construct (66 \pm 9%, Fig. 3B, lane 6) indicates that TM3 was inserted with its C-terminus oriented towards the ER lumen (see Fig. 3D, top). Finally, the intermediate glycosylation efficiency observed for full-length protein (185mer) translations (38 \pm 3, Fig. 3B, lane 8), supports the existence of an equilibrium between insertion-translocation for this last TM segment, suggesting that the presence of preceding TM segments causes a noticeable increase in the insertion efficiency of HR4, but not enough for full assembly into the microsomal membranes, as illustrated in Fig. 3D (bottom).

3.3. TRAP-y membrane topology in cells.

In order to investigate the overall membrane topology of TRAP-y subunit in cellula we used two different approaches. First, the topology of TRAP-γ subunit was analysed in vivo in the yeast S. cerevisiae using a glycosylatable green fluorescent protein (gGFP) reporter [28]. When oriented towards the ER lumen the glycosylation of the gGFP interferes with the protein folding and maturation of the fluorophore, rendering the protein nonfluorescent. On the contrary, the un-modified gGFP moiety is properly folded and remains fluorescent when oriented towards the cytosol. Since our in vitro results showed an un-efficient integration for the last TM segment, the gGFP topology reporter was C-terminally fused to TRAP-γ sequence at position 161, after TM3 (161mer-gGFP) and at the C-terminal end of the protein (residue 185, 185mer-gGFP) to monitor TRAP-γ subunit topology. Figure 4 shows the glycosylation status (Fig. 4A) and the fluorescence (Fig. 4B) of both protein chimeras. To monitor the molecular weight of the un-glycosylated protein, the samples were treated with EndoH. The 161mer-gGFP chimera was efficiently mono-glycosylated (Fig. 4A, lane 1) and subsequently no fluorescence was observed (Fig. 4B), indicating a lumenal orientation of the reporter moiety. Conversely, 181mer-gGFP was not glycosylated (Fig. 4A) and thus fluorescent (Fig. 4B), demonstrating the cytosolic localisation of fulllength TRAP- γ C-terminus and the concomitant insertion into the yeast membranes of TM4 (see Fig. 4C for a scheme). It should be mentioned that native acceptor asparagine (Asn23) was never modified.

Second, to analyse the protein topology in mammalian cells, a series of TRAP- γ variants were transfected into HEK-293T cells. As expected, wild type TRAP- γ subunit is not glycosylated in mammalian cells (Fig. 4D, lanes 1 and 2), in good agreement with yeast glycosylation data (Fig. 4A). To monitor the location of TRAP- γ N-terminus we designed a highly efficient artificial (Asn-Ser-Ser) glycosylation site [43] at position 5 (G5N mutant, Fig. 1A). As shown in Fig. 4D (lanes 3 and 4), this engineered acceptor site was not modified, denoting an N-terminal cytosolic localisation. Finally, by inserting a Thr at position 111 (T111) we incorporated an additional glycosylation site (Asn-Arg-Thr) in the loop connecting TM2 and TM3 (see Fig. 1A). The cytosolic localisation of this region prevented the glycosylation of the designed acceptor site (Fig. 4D, lane 5). All in all, our *in vivo* results clearly demonstrated that TRAP- γ subunit contains four TM segments and integrates into the eukaryotic membranes with an Nt/Ct cytoplasmic orientation.

4. Discussion

Membrane proteins participate in a vast array of cellular processes. Their correct insertion and orientation within the lipid bilayer is of paramount importance for the cell. In eukaryotes, the insertion of most membrane proteins occurs through the ER translocon, which performs its activity assisted by multiple associated proteins. Among them, the TRAP complex is a stoichiometric component of the ribosome-bound translocon complex [44], which has been postulated to participate in the early stages of protein translocation [9], and in moderating topogenic determinants at a final stage of membrane protein biogenesis *in vivo* [11]. Interestingly, structural data showed that membrane-embedded regions of TRAP complex and Sec61 translocon are in close proximity, which does not block the lateral gate of Sec61 translocon that permits TM segment partition into the lipid bilayer [45]. Furthermore, proteomic analyses have shown that TRAP can form stable complexes with Sec61 and OST even in the absence of bound ribosomes [46].

We have determined the membrane topology of TRAP-γ subunit by utilizing glycosylation mapping as marker for lumenal localization and evaluating the membrane integration efficiency of its TM segments both in microsomal membranes and in eukaryotic cells. Taken together the results demonstrate that TRAP-γ subunit consists of four TM segments and its Nand C-termini are on the cytosolic side, which is consistent with the in silico predictions (Table 1) but reverse to the topology suggested in a seminal work [7]. To support our conclusions we assessed the localization of all protein domains in vitro but also used gGFP as a topological reporter in S. cerevisiae and glycosylation mapping in mammalian cells to allocate both the N- and Ctermini, and the charged segment between the two α -helical hairpins. These offered complementary methodologies results that confirmed our conclusions.

TRAP-γ subunit contains two poorly hydrophobic TM segments (TM2 and TM4), as denoted by the predicted free energy of insertion apparent positive values estimated by the ΔG Prediction algorithm (Fig. 1A). Individual integration of TM2 in a model protein proved to be very efficient (Fig. 1C). However, in these experiments TM4 was not recognised by the translocon as a TM sequence (Fig. 1C). In fact, the glycosylation pattern found in this in vitro system when TM4 was expressed in-block with preceding TM3 (including their native turn region) suggests that HR4 is not efficiently recognized as an α -helical hairpin component during protein biogenesis (Figs. 2D and 3B). Surprisingly, in a cellular context, TM4 becomes efficiently inserted into the membrane, as found in our in vivo experiments (Fig. 4). Nevertheless, the presence of TM1 and/or TM2 might slightly stimulate TM4 insertion despite their distance in the primary sequence. Whether there is an interaction in the tertiary structure between TM4 and TM1 and/or TM2 or there is only a temporary cooperation between these hydrophobic regions at the time of their insertion [47, 48] cannot be concluded from our results. Either scenario the final topology of the protein requires the presence of all TM segments and, likely, some additional cellular components.

The proposed topology of TRAP- γ provides a basis for the functional dissection of this stoichiometric translocon component. Functional studies involved TRAP in the initial stages of substrate transport in a signal sequence-modulated manner [9]. Indeed, TRAP luminal domain and stalk observed in

structural studies are likely composed of the N-terminal extra-membranous regions of the α , β , and δ subunits [45], while the small cytosolic domain that binds to the ribosome correspond to the TRAP- γ N-terminus and hairpin-connecting loop regions [34]. Our results suggest that either signal sequence or ribosome interactions with the TRAP complex might occur through γ -subunit, since TRAP- γ exposes sufficiently large extra-membranous domains towards the cytosolic side. This topological information facilitates the design of protein-protein interaction screenings and focus future structural studies.

Acknowledgments

We thank Silvia Tamborero and Janine Tittel for preliminary work, Prof. Isabel Fariñas (Univ. of Valencia) for providing the mouse cDNA library and Prof. Hyun Kim (Seoul National University) for the yeast HA-gGFP cDNA. This work was supported by grants BFU2016-79487 from the Spanish Ministry of Economy and Competitiveness (MINECO) (to I.M.), and by grants PROMETEOII/2014/061 (Grups d'Excel·lència) and GV/2016/139 (Grups Emergents) from the Generalitat Valenciana (to I.M. and L.M-G, respectively). M.B-P and B.G. are postdoctoral and predoctoral fellows from the Generalitat Valenciana, respectively. L.M-G is a postdoctoral fellow from the Juan de la Cierva program (MINECO).

Figure legends

Fig. 1. Insertion of predicted TM segments into microsomal membranes. (a) Amino acid sequence of TRAP- γ . Predicted TM regions by the ΔG algorithm (http://dgpred.cbr.su.se/)[17, 18] are highlighted in orange boxes and their apparent values shown at the bottom. Positive ΔG_{****} values, indicative of membrane translocation, are shown in red. Negative ΔG_{****} values, indicative of insertion, are shown in green. Glycosylation acceptor sites are underlined, with the mutated or inserted residue shown in bold. A set of oligonucleotides (arrowhead dotted lines) was designed to generate TRAP- γ truncated forms, which are indicated by the length of the truncated polypeptide (ending residue -mer). (B) Schematic representation of the leader peptidase (Lep) construct used to report insertion into ER membrane of TRAP- γ HR1, HR2,

HR3 and HR4. The HR under study is inserted into the P2 domain of Lep flanked by two artificial glycosylation acceptor sites (G1 and G2). Recognition of the HR by the translocon machinery as a TM domain (green) locates only G1 in the luminal side of the ER membrane preventing G2 glycosylation. The Lep chimera will be doubly glycosylated when the HR being tested is translocated (red) into the lumen of the microsomes. (C) In vitro translation in the presence of membranes of the different Lep constructs. Constructs containing the HR1 (residues 30-51, lanes 1-3), HR2 (residues 55-77, lanes 4-6), HR3 (residues 138-160, lanes 7-9) and HR4 (residues 163-182, lanes 10-12) were translated in the presence (+) and absence (-) of rough microsomes (RM) and proteinase K (PK). Bands of non-glycosylated proteins are indicated by a white dot; singly and doubly glycosylated proteins are indicated by one and two black dots, respectively. In the case of Lep-HR3 construct a triply glycosylated band (≈2%) was observed (lane 8) due to the presence of an acceptor NNT site (residues 141-143) within the (translocated) hydrophobic region. The protected doubly-glycosylated H2/HR4/P2 fragment is indicated by an arrowhead. Each gel is representative of three independent experiments.

Fig. 2. Insertion of hydrophobic regions (HRs) using the Lep' construct. (A) Schematic representation of the Lep-derived construct (Lep'). In this Lep' construct tested HRs replaces the H2 domain from Lep. The glycosylation acceptor site (G2') located in the beginning of the P2 domain will be modified only if tested HR inserts into the membrane, while the G1' site, embedded in an extended N-terminal sequence of 24 amino acids, is always glycosylated. (B) *In vitro* translation in the presence of membranes (+) and absence (-) of rough microsomes (RM) and proteinase K (PK) of HR1 (lanes 1-3) and HR3 (lanes 4-6). Bands of non-glycosylated proteins are indicated by a white dot; singly and doubly glycosylated proteins are indicated by one and two black dots, respectively. The protected glycosylated HR/P2 fragment is indicated by an arrowhead. (C) In this Lep' constructs the H2 domain from Lep is replaced by TRAP-γ helical hairpins and the N-terminus elongated. (D) *In vitro* translations of Lep'-HR1-2 and Lep'-HR3-4 were performed and labeled as in panel (B). Each gel is representative of three independent experiments.

Fig. 3. *In vitro* analysis of truncated TRAP- γ constructs. To monitor the membrane orientation of truncated TRAP- γ molecules a glycosylatable

(NSTMSM) tag (white rectangle) was added at position 54 (54mer), 132 (132mer), 161 (161mer) and 185 (185mer). (A) Schematic representation of the constructs used in the assay. The position of the glycosylation sites is marked with a Y symbol. The presence of a TM segment identified by the ΔG prediction server (http://dgpred.cbr.su.se/) in each construct and the required linker sequence preceding the glycosylatable tag to allow glycosylation is also included for 54mer, 161mer and 185mer truncates. (B) In vitro translation of the 54mer, 132mer, 161mer and 185mer truncates in the presence of rough microsomes (RM). Lanes 1, 3, 5 and 7, samples were treated with endoglycosidase H (EndoH) prior to SDS-PAGE analysis. A white dot marks the non-glycosylated form of the protein while a black dot indicates glycosylation of the C-terminal tag. Glycosylation percentages are shown at the bottom. (C) In vitro translations in the presence of RM of 54mer truncated constructs were performed bearing an acceptor (NST, NSTMSM) or nonacceptor (QST, QSTMSM) C-terminal glycosylation tag. Samples were subjected to EndoH digestion or mock treated as indicated. White and black dots indicate non-glycosylated and glycosylated molecules respectively, as in panel (B). (D) Schematic representation of the membrane topology of the 54mer, 132mer, 161mer and 185mer truncates. A hydrophobic region is noted as a green box when inserted in the membrane, or as a red box if it is not recognized by the translocon as a TM domain. The position of the glycosylatable tag (white rectangle) and its glycosylation status (white and black dots, represents non-glycosylated and glycosylated respectively) is also shown. Two potential topologies for the 185mer are represented (bottom). Experiments were done and quantified at least in triplicates.

Fig. 4. TRAP-γ membrane topology in eukaryotic cells. (A-C) TRAP-γ 161mer and 185mer (full length) were fused to gGFP to analyse the membrane topology of the protein in *S. cerevisiae*. The gGFP is properly folded and thus fluorescent in the yeast cytosol, on the other hand once translocated into the ER lumen it is glycosylated and the fluorescence is lost. (A) Glycosylation status of the 161mer-gGFP and 185mer-gGFP constructs. To confirm the glycosylation of the protein samples were treated (+) or mock treated (-) with endoglycosidase H (EndoH). White dot indicate non-glycosylated protein while black dot point out the glycosylated form. (B) Differential Interference Contrast (DIC) and green fluorescent (GFP) images of *S.cerevisiae* cells

expressing the 161mer-gGFP and 185mer-gGFP constructs. (C) Schematic representation of the membrane topology adopted by the 161mer-gGFP and 185mer-gGFP chimeras. The location of TM domains (green), glycosylation acceptor sites (white and black dots, according to their glycosylation status) and the gGFP (gray line and green structure representation when unfolded and properly folded, respectively) is shown. (D) Full-length TRAP-γ subunit was expressed in HEK-293T cells to evaluate its membrane topology utilizing its glycosylation status as a reporter. To monitor the electrophoretic mobility of the non-glycosylated proteins the samples were treated (+) or mock treated (-) with EndoH as in panel (a). Wild type (wt) (lanes 1 and 2), Gly5 to Asn (G5N) mutant (lanes 3 and 4) and the Thr111 (T111) (lanes 5 and 6) insertion were analyzed. Lanes 7 and 8 show the Nipah G protein as an internal glycosylation control. White dot indicate non-glycosylated protein while black dot point out the molecular weight of a glycosylated G protein. The number of modified glycosylation sites is shown between brackets.

Table 1. Computer analysis of TRAP- γ amino acid sequence.

Algorithm	N° TMs (starting-ending position)	Topology
SOSUI	3(26-48)(56-78)(138-157)	NA^a
Split	3(29-49)(55-78)(135-158)	Nt_{cyt}/Ct_{lum}
OCTOPUS	3(29-50)(57-78)(135-157)	Nt_{cyt}/Ct_{lum}
PHD	3(31-51)(54-77)(140-164)	Nt_{cyt}/Ct_{lum}
TOP PRED 2	4(28-48)(55-77)(140-160)(164-184)	Nt_{cyt}/Ct_{cyt}
TMHMM	4(29-51)(55-77)(138-157)(162-184)	Nt_{cyt}/Ct_{cyt}
TM pred	4(30-47)(50-77)(138-160)(165-181)	Nt_{cyt}/Ct_{cyt}
MEMSAT	4(30-48)(55-77)(135-155)(162-179)	Nt_{cyt}/Ct_{cyt}
HMMTOP	4(30-48)(59-77)(135-157)(166-184)	Nt_{cyt}/Ct_{cyt}
TOPCONS	4(30-50)(55-75)(136-156)(163-183)	Nt_{cyt}/Ct_{cyt}
ΔG PRED	4(30-51)(55-77)(138-160)(163-182)	NA^a

^a Not applicable

References

- [1] L. Martinez-Gil, A. Sauri, M.A. Marti-Renom, I. Mingarro, Membrane protein integration into the ER, FEBS J, 278 (2011) 3846-3858.
- [2] A.E. Johnson, M.A. van Waes, The translocon: A dynamic gateway at the ER membrane, Annual Review of Cell and Developmental Biology, 15 (1999) 799-842.
- [3] P. Whitley, I. Mingarro, Stitching proteins into membranes, not sew simple, Biol Chem, 395 (2014) 1417-1424.
- [4] D. Görlich, T.A. Rapoport, Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane, Cell, 75 (1993) 615-630.
- [5] D. Görlich, E. Hartmann, S. Prehn, T.A. Rapoport, A Protein of the Endoplasmic Reticulum Involved Early in Polypeptide Translocation, Nature, 357 (1992) 47-52.
- [6] S. Tamborero, M. Vilar, L. Martinez-Gil, A.E. Johnson, I. Mingarro, Membrane Insertion and Topology of the Translocating Chain-Associating Membrane Protein (TRAM), J Mol Biol, 406 (2011) 571-582.
- [7] E. Hartmann, D. Gorlich, S. Kostka, A. Otto, R. Kraft, S. Knespel, E. Burger, T.A. Rapoport, S. Prehn, A tetrameric complex of membrane proteins in the endoplasmic reticulum, Eur J Biochem, 214 (1993) 375-381.
- [8] J.-F. Menetret, R.S. Hegde, S.U. Heinrich, P. Chandramouli, S.J. Ludtke, T.A. Rapoport, C.W. Akey, Architecture of the Ribosome-Channel Complex Derived from Native Membranes, J Mol Biol, 348 (2005) 445-457.
- [9] R.D. Fons, B.A. Bogert, R.S. Hegde, Substrate-specific function of the transloconassociated protein complex during translocation across the ER membrane, J Cell Biol, 160 (2003) 529-539.
- [10] G. von Heijne, Analysis of the distribution of charged residues in the N-terminal region of signal sequences: implications for protein export in prokaryotic and eukaryotic cells, EMBO J, 3 (1984) 2315-2318.
- [11] N. Sommer, T. Junne, K.U. Kalies, M. Spiess, E. Hartmann, TRAP assists membrane protein topogenesis at the mammalian ER membrane, Biochim Biophys Acta, 1833 (2013) 3104-3111.
- [12] K. Nagasawa, T. Higashi, N. Hosokawa, R.J. Kaufman, K. Nagata, Simultaneous induction of the four subunits of the TRAP complex by ER stress accelerates ER degradation, EMBO Rep, 8 (2007) 483-489.
- [13] K. Mesbah, A. Camus, C. Babinet, J. Barra, Mutation in the Trapalpha/Ssr1 gene, encoding translocon-associated protein alpha, results in outflow tract morphogenetic defects, Mol Cell Biol, 26 (2006) 7760-7771.
- [14] M.E. Losfeld, B.G. Ng, M. Kircher, K.J. Buckingham, E.H. Turner, A. Eroshkin, J.D. Smith, J. Shendure, D.A. Nickerson, M.J. Bamshad, G. University of Washington Center for Mendelian, H.H. Freeze, A new congenital disorder of glycosylation caused by a mutation in SSR4, the signal sequence receptor 4 protein of the TRAP complex, Hum Mol Genet, 23 (2014) 1602-1605.
- [15] D.H. Li, T. Chan, R. Satow, S. Komazaki, K. Hashizume, M. Asashima, The role of XTRAP-gamma in Xenopus pronephros development, Int J Dev Biol, 49 (2005) 401-408.
- [16] Y.L. Yamaguchi, S.S. Tanaka, N. Oshima, H. Kiyonari, M. Asashima, R. Nishinakamura, Translocon-associated protein subunit Trap-gamma/Ssr3 is required for vascular network formation in the mouse placenta, Dev Dyn, 240 (2011) 394-403.
- [17] T. Hessa, H. Kim, K. Bihlmaier, C. Lundin, J. Boekel, H. Andersson, I. Nilsson, S.H. White, G. von Heijne, Recognition of transmembrane helices by the endoplasmic reticulum translocon, Nature, 433 (2005) 377-381.
- [18] T. Hessa, N.M. Meindl-Beinker, A. Bernsel, H. Kim, Y. Sato, M. Lerch-Bader, I. Nilsson, S.H. White, G. von Heijne, Molecular code for transmembrane-helix recognition by the Sec61 translocon, Nature, 450 (2007) 1026-1030.
- [19] G.E. Tusnady, I. Simon, The HMMTOP transmembrane topology prediction server, Bioinformatics, 17 (2001) 849-850.

- [20] D.T. Jones, Improving the accuracy of transmembrane protein topology prediction using evolutionary information, Bioinformatics, 23 (2007) 538-544.
- [21] B. Rost, P. Fariselli, R. Casadio, Topology prediction for helical transmembrane proteins at 86% accuracy, Protein Sci, 5 (1996) 1704-1718.
- [22] T. Hirokawa, S. Boon-Chieng, S. Mitaku, SOSUI: classification and secondary structure prediction system for membrane proteins, Bioinformatics, 14 (1998) 378-379.
- [23] D. Juretic, B. Lee, N. Trinajstic, R.W. Williams, Conformational Preference Functions for Predicting Helices in Membrane Proteins, Biopolymers, 33 (1993) 255-273.
- [24] A. Krogh, B. Larsson, G. von Heijne, E.L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, J Mol Biol, 305 (2001) 567-580.
- [25] H. Viklund, A. Elofsson, OCTOPUS: improving topology prediction by two-track ANN-based preference scores and an extended topological grammar, Bioinformatics, 24 (2008) 1662-1668.
- [26] K.D. Tsirigos, C. Peters, N. Shu, L. Kall, A. Elofsson, The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides, Nucleic Acids Res, 43 (2015) W401-407.
- [27] M.G. Claros, G. von Heijne, TopPred II: An improved software for membrane protein structure prediction, CABIOS, 10 (1994) 685-686.
- [28] H. Lee, J. Min, G. von Heijne, H. Kim, Glycosylatable GFP as a compartment-specific membrane topology reporter, Biochemical and biophysical research communications, 427 (2012) 780-784.
- [29] L. Martinez-Gil, M. Bano-Polo, N. Redondo, S. Sanchez-Martinez, J.L. Nieva, L. Carrasco, I. Mingarro, Membrane integration of poliovirus 2B viroporin, Journal of virology, 85 (2011) 11315-11324.
- [30] P. Donate-Macian, M. Bano-Polo, J.L. Vazquez-Ibar, I. Mingarro, A. Peralvarez-Marin, Molecular and topological membrane folding determinants of transient receptor potential vanilloid 2 channel, Biochem Biophys Res Commun, 462 (2015) 221-226.
- [31] L. Martinez-Gil, A. Sauri, M. Vilar, V. Pallas, I. Mingarro, Membrane insertion and topology of the p7B movement protein of Melon Necrotic Spot Virus (MNSV), Virology, 367 (2007) 348-357.
- [32] M. Bano-Polo, L. Martinez-Gil, B. Wallner, J.L. Nieva, A. Elofsson, I. Mingarro, Charge pair interactions in transmembrane helices and turn propensity of the connecting sequence promote helical hairpin insertion, J Mol Biol, 425 (2013) 830-840.
- [33] I. Mingarro, G. von Heijne, P. Whitley, Membrane-protein engineering, TIBTECH, 15 (1997) 432-437.
- [34] S. Pfeffer, J. Dudek, R. Zimmermann, F. Forster, Organization of the native ribosometranslocon complex at the mammalian endoplasmic reticulum membrane, Biochim Biophys Acta, 1860 (2016) 2122-2129.
- [35] C. Lundin, H. Kim, I. Nilsson, S.H. White, G. von Heijne, Molecular code for protein insertion in the endoplasmic reticulum membrane is similar for Nin/Cout and Nout/Cin transmembrane helices, Proceedings of the National Academy of Sciences, 105 (2008) 15702-15707.
- [36] M. Monne, I. Nilsson, A. Elofsson, G. von Heijne, Turns in transmembrane helices: determination of the minimal length of a "helical hairpin" and derivation of a fine-grained turn propensity scale, J Mol Biol, 293 (1999) 807-814.
- [37] L.E. Hedin, K. Ojemalm, A. Bernsel, A. Hennerdal, K. Illergard, K. Enquist, A. Kauko, S. Cristobal, G. von Heijne, M. Lerch-Bader, I. Nilsson, A. Elofsson, Membrane insertion of marginally hydrophobic transmembrane helices depends on sequence context, J Mol Biol, 396 (2010) 221-229.
- [38] K. Ojemalm, K.K. Halling, I. Nilsson, G. von Heijne, Orientational preferences of neighboring helices can drive ER insertion of a marginally hydrophobic transmembrane helix, Mol Cell, 45 (2012) 529-540.

- [39] M. Bano-Polo, F. Baldin, S. Tamborero, M.A. Marti-Renom, I. Mingarro, N-glycosylation efficiency is determined by the distance to the C-terminus and the amino acid preceding an Asn-Ser-Thr sequon, Protein Sci, 20 (2011) 179-186.
- [40] I. Nilsson, G. von Heijne, Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane, J Biol Chem, 268 (1993) 5798-5801.
- [41] M. Orzaez, J. Salgado, A. Gimenez-Giner, E. Perez-Paya, I. Mingarro, Influence of proline residues in transmembrane helix packing, J Mol Biol, 335 (2004) 631-640.
- [42] S. Pfeffer, J. Dudek, M. Gogala, S. Schorr, J. Linxweiler, S. Lang, T. Becker, R. Beckmann, R. Zimmermann, F. Forster, Structure of the mammalian oligosaccharyl-transferase complex in the native ER protein translocon, Nat Commun, 5 (2014) 3072.
- [43] L. Kasturi, H. Chen, S.H. Shakin-Eshleman, Regulation of N-linked core glycosylation: use of a site-directed mutagenesis approach to identify Asn-Xaa-Ser/Thr sequons that are poor oligosaccharide acceptors, The Biochemical journal, 323 (Pt 2) (1997) 415-419.
- [44] S. Pfeffer, L. Burbaum, P. Unverdorben, M. Pech, Y. Chen, R. Zimmermann, R. Beckmann, F. Forster, Structure of the native Sec61 protein-conducting channel, Nat Commun, 6 (2015) 8403.
- [45] J.F. Menetret, R.S. Hegde, M. Aguiar, S.P. Gygi, E. Park, T.A. Rapoport, C.W. Akey, Single copies of Sec61 and TRAP associate with a nontranslating mammalian ribosome, Structure, 16 (2008) 1126-1137.
- [46] T. Shibatani, L.L. David, A.L. McCormack, K. Frueh, W.R. Skach, Proteomic analysis of mammalian oligosaccharyltransferase reveals multiple subcomplexes that contain Sec61, TRAP, and two potential new subunits, Biochemistry, 44 (2005) 5982-5992.
- [47] H. Sadlish, D. Pitonzo, A.E. Johnson, W.R. Skach, Sequential triage of transmembrane segments by Sec61alpha during biogenesis of a native multispanning membrane protein, Nat Struct Mol Biol, 12 (2005) 870-878.
- [48] A. Sauri, P.J. McCormick, A.E. Johnson, I. Mingarro, Sec61alpha and TRAM are sequentially adjacent to a nascent viral membrane protein during its ER integration, J Mol Biol, 366 (2007) 366-374.

Fig 1

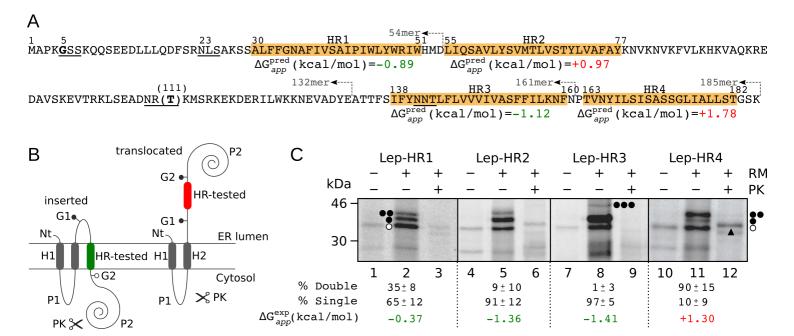
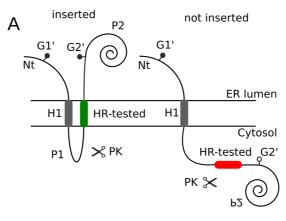
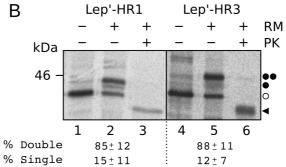
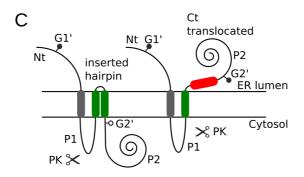
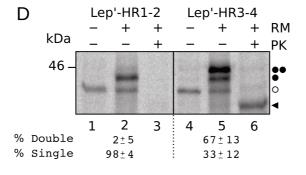


Fig. 2









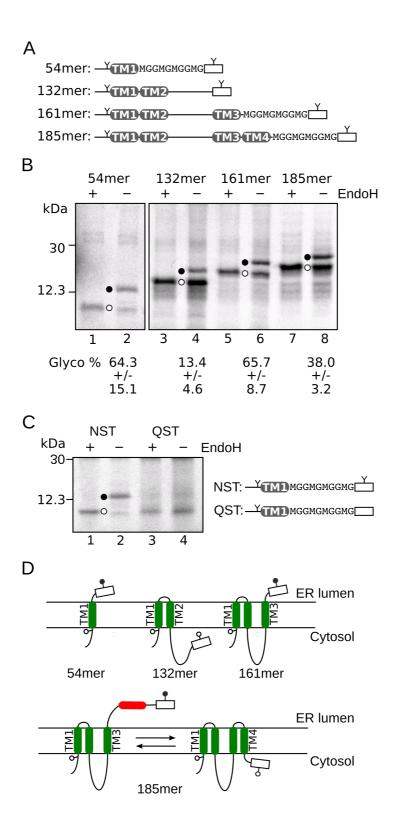


Fig. 4

