Aus dem Bereich Medizinische Biochemie und
Molekularbiologie Theoretische Medizin und
Biowissenschaften der Medizinischen Fakultät
der Universität des Saarlandes, Homburg/Saar

The mammalian Translocon-Associated Protein (TRAP) complex and co-translational protein transport

Dissertation zur Erlangung des Grades eines
Doktors der Naturwissenschaften
der Medizinischen Fakultät
der UNIVERSITÄT DES SAARLANDES

2019

Vorgelegt von: Antonietta Russo
Geb. am: 10.06.1966

Tag der Promotion: 03.06.2020
Dekan: Prof. Dr. M. D. Menger
Berichterstatter: Prof. Dr. Richard Zimmermann
Prof. Dr. Markus Hoth

Abstract

In eukaryotes, many proteins translocate into the endoplasmic reticulum (ER) and some insert themselves into the membrane. In 1971, G. Blobel proposed the “Signal Hypothesis”, explaining how proteins translocate from the cytosol into the ER. The signal peptides (SPs) have a tripartite structure but heterogeneous peptide sequence composition. SP complexity is essential for many processes, such as signal recognition particle binding, translocon gating, early folding prevention, and signal peptidase interaction/cleavage, as well as post-cleavage functions, such as antigen presentation. ER protein translocation takes place through the heterotrimeric Sec61 protein-conducting channel (PCC). Sec61α is a multi-spanning membrane protein that forms a complex with the single-spanning partners Sec61β and Sec61γ. During co-translational translocation, Sec61 is associated with the ribosome (via two cytosolic loops) and many accessory components (ribosome-channel complex), such as the translocon-associated protein (TRAP) complex. TRAP has four subunits—TRAP α (ssr1), TRAP β (ssr2), TRAP γ (ssr3), and TRAP δ (ssr4). The subunits α, β, and δ are single-spanning transmembrane (TM) proteins with luminal and cytosolic domains (type I), while the subunit γ has four TM domains and a prominent cytosolic domain. Recently, microscopic techniques, such as cryo-EM and cryo-ET, have enabled the determination of the translocation machinery structure. However, at present there is a lack of understanding regarding the roles of some of its components and domains. Protein function is determined by many different aspects, including localisation, sequence, structure, expression, post-translational modifications, and interactions. The present study aimed to contribute to the understanding of TRAP functions. Analyses of protein-protein interactions (PPIs), sequences (motifs), and expressions were carried out using experimental and computational methods. Importantly, we found that the TRAP complex interacts with the translocon Sec61 (peptide array). These PPIs may be essential for translocating substrates or stabilising translocon machinery. The PPIs occur in the ER luminal side between Sec61α1 loop 5 and TRAP α/β subunits. The latter also interact with one another, as is expected for elements of a complex (pull-down assays). The computational analysis identified a calcium-binding domain at the N-terminus of the TRAP α subunit, which may have a functional role.

Zusammenfassung

In Eukaryonten müssen viele Proteine in das ER transferiert werden, und einige von ihnen werden in die Membran eingebracht. Ein Meilenstein im Verständnis der Protein-Translokation ist die von G. Blobel 1971 vorgeschlagene "Signalhypothese", die erklärt, wie die Proteine aus dem Cytosol in das ER transloziert werden. Das Signalpeptid (SP) hat typischerweise eine dreiteilige Struktur, ist aber in der Aminosäuresequenz sehr heterogen, was eine schnelle Entwicklung im Verlauf der Evolution impliziert, die wahrscheinlich mit dem reifen Protein verbunden ist. Heutzutage wissen wir, dass die Komplexität des SP mit vielen biologischen Prozessen verbunden ist: Signalerkennungspartikel (SRP)-Bindung, translokale Interaktion (Gating), frühe Faltungsprävention, Signalpeptidase (SPase)-Interaktion und - Spaltung und sogar Post-Cleavage-Funktionen wie Antigenpräsentation. Die Translokation erfolgt über einen heterotrimeren proteinleitenden Kanal (PCC): Sec61α ist ein multi- spannendes Membranprotein und bildet mit den Single-Spanning-Partnern Sec61β und Sec61γ einen Komplex. Während der kotranslationalen Translokation ist Sec61 dem Ribosom (über zwei zytosolische Schleifen) und einer großen Anzahl von weiteren Komponenten zugeordnet (RCC, Ribosom-Kanal- Komplex). Zu diesen zusätzlichen Bestandteilen gehört der heterotetramere Translocon-Associated Protein (TRAP)-Komplex aus TRAP α (ssr1), TRAP β (ssr2), TRAP γ (ssr3) und TRAP δ (ssr4). Die Untereinheiten α, β und δ sind Single-Spanning- Transmembran (TM)-Proteine mit ER-luminalen und zytosolischen Domänen, während die γ Untereinheit vier TM-Domänen und eine prominente zytosolische Domäne aufweist. In letzter Zeit wurden große Fortschritte bei der Untersuchung der Struktur der Translokationsmaschinen erzielt, auch dank der verbesserten mikroskopischen Techniken wie Cryo-EM und Cryo-ET; die Forschung hat jedoch stets gezeigt, dass das Verständnis für die Rolle(n) einiger ihrer Komponenten und Domänen fehlt. Eine Proteinfunktion wird unter Berücksichtigung vieler Aspekte untersucht: intrazelluläre Lokalisation, Sequenz, Struktur, Expressionsprofil, post-translationale Modifikationen, Interaktionen. Der Hauptzweck dieser Forschungsarbeit ist es, zum Verständnis der TRAP- Funktion(en) beizutragen, indem Sequenzen (Motive), Expressionen und vor allem Protein- Protein-Interaktionen innerhalb des Komplexesund mit den umgebenden Strukturen durch computergestützte und experimentelle Methoden analysiert werden. Die relevanteste Erkenntnis ist, dass der TRAP-Komplex mit dem Translokon interagiert (Peptid- Array). Diese Wechselwirkungen könnten für die Translokation einiger Substrate oder auch nur für die Stabilisierung der Translokomaschinerie von wesentlicher Bedeutung sein. Die Interaktionen finden auf der ER-Lumenseite zwischen Sec61α1 loop 5 und den TRAP α /β Untereinheiten statt, letztere interagieren auch untereinander, wie es für Elemente eines Komplexes erwartet wird (Pulldown-Assay). Die Computeranalyse zeigt eine Calcium-Bindungsdomäne am N-Terminus der TRAP alpha Untereinheit auf, die eine funktionelle Rolle spielen könnte.

Table of contents

1. INTRODUCTION 1

1.1 The Endoplasmic reticulum (ER) 1

1.2 Protein synthesis and ER translocation 2

1.3 Translocation regulation and Sec61 accessory components 4

1.4 Signal Peptide: from sequence to sorting 7

1.5 Transmembrane proteins: the connection between two environments 9

1.6 Quality control in the ER: an extensive network 10

1.7 EF-hand calcium-binding domains: structural and regulatory 11

1.8 TRAP complex: a Sec61 associated component 12

1.9 OST and post-translocation modification 27

1.10 Objectives 28

2. MATERIAL AND METHODS 29

2.1 Experimental Methods 29

2.1.1 TRAP alpha, beta, gamma, and delta (most common isoforms) 29

2.1.2 Molecular cloning 35

2.1.3 Quantitative PCR 37

2.1.4 Digestion with restriction enzymes 38

2.1.5 Ligation 38

2.1.6 Plasmid purification miniprep and MIDI 39

2.1.7 Sequencing 39

2.1.8 Protein expression and purification 39

2.1.8.1 Bacterium strains 39

2.1.8.2 Affinity chromatography batch method 40

2.1.8.3 Affinity chromatography column method 41

2.1.8.4 GST purification 42

2.1.8.5 Solubility, concentration, and purity 43

2.1.8.6 SDS-PAGE 43

2.1.8.7 Western blotting 44
2.1.8.8 Liquid Chromatography-Mass Spectrometry (LC-MS) 44
2.1.8.9 GST and HIS pull-down assays 44
2.1.8.10 Peptide array 46
2.2 Computational methods 49

3. RESULTS 50
3.1 Experimental results 50
3.1.1 TRAP alpha, beta, gamma, and delta cloning 50
3.1.2 Recombinant tagged proteins 53
3.1.3 Pull-down assays 63
3.1.4 Peptide array 65
3.2 Computational results 71
3.2.1 Hydrophobicity/TMDs of TRAP subunits 71
3.2.2 TRAP alpha non-canonical EF-hand motif 73
3.2.3 STRING and RaptorX predictions 74
3.2.4 TRAP beta and TM motif retention 79
3.2.5 Signal Peptide 79

4. DISCUSSION 82
4.1 Recombinant proteins 82
4.2 TRAP alpha and TRAP beta interaction 83
4.3 TRAP alpha/beta and Sec61alpha1 loop5 interactions 84
4.4 TRAP functions 86
4.5 Overview and future prospective 87

Abbreviations 91
References 94
Acknowledgements 103
Curriculum Vitae 105

1. Introduction

1.1 Endoplasmic reticulum

The endoplasmic reticulum (ER) consists of a network of tubules and vesicles that makes up a subcellular compartment of eukaryotic cells. The ER is continuous with the nuclear envelope and is the most extensive membrane structure in the cell with a surface size of up to 30 times that of the cellular membrane. ER membranes are less packed than plasma membranes and are made up of many dynamically regulated lipids, the most abundant of which are phosphatidylcholine and phosphatidylethanolamine, with cholesterol and other lipids present in smaller amounts. ER can be smooth or rough and in the latter ribosomes are attached to the membrane (polysomes). In smooth ER, which is made up of a tubule structure, lipid metabolism, calcium release, detoxification, and carbohydrate synthesis take place, while in rough ER, made up of a series of flattened sacs, protein translocation, folding, oligomerisation, glycosylation, and degradation occur (Fig.1.1.1).

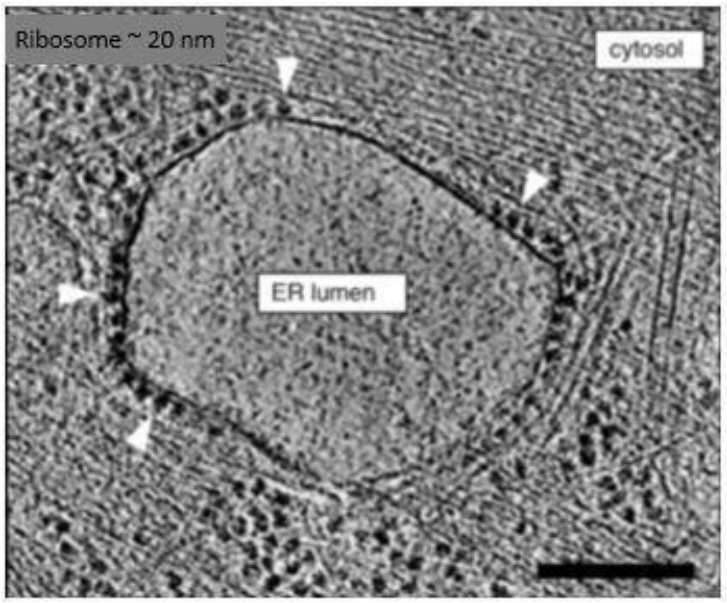


Fig.1.1.1 - Cryo-electron microscopy (Cryo-EM) of mammalian rough ER: membrane- ribosome (arrowheads), cytosolic ribosomes, and cytosolic skeletal filaments. Scale bar: 200 nm (Pfeffer et al., 2012).

Exit sites are present on the ER membranes for the export of newly synthesised proteins into the secretory pathway. The coat protein complex II (COPII), made up of five cytosolic conserved proteins (Sar1, Sec23, Sec24, Sec13, and Sec31), creates small membrane vesicles. The vesicles transport the cargo proteins from the ER to the Golgi apparatus and on to the final destination (Jensen et al., 2011). The ER lumen presents a high concentration of calcium, between 100 and 800 µM, whereas in the cytosol, the concentration is about 100 nM. ER is the first Ca2+ store in cells and an active pump, sarcoplasmic-endoplasmic reticulum Ca2+-ATPase (SERCA), maintains this gradient. The maturation of many proteins into the ER relies on Ca2+ concentration. The concentrations of other electrolytes between ER and cytosol are similar, and the pH is near to neutrality. Many ER resident proteins have functional calcium-binding domains, such as calreticulin, calnexin, and binding immunoglobulin protein (BiP). These proteins are essential chaperones: 1) BiP binds the translocating nascent protein to assist folding but is also involved in ER-associated degradation (ERAD) and unfolded protein response (UPR); 2) calnexin and calreticulin carry out similar functions during folding, UPR, and ERAD

(Ellgard et al., 2003). A currently unknown mechanism imports ATP into the ER. Many processes occur as a result, such as i) the formation of disulphide bonds, phosphorylation, and glycosylation and ii) the dissociation of chaperones involved in UPR.

1.2 Protein synthesis and ER translocation

In eukaryotes, the velocity of mRNA translation in polypeptides is about five residues per second. Ribosomes are protein-RNA complexes (3.6 M Dalton) consisting of two subunits: 40S, which binds and decodes the mRNA, and 60S, for peptide bond formation (peptidyl-transferase). About one-third of the synthesised proteins in the cell translocate or reside in the ER. These proteins consist of soluble intracellular, soluble secreted, type I membrane, type II membrane, and multi-spanning membrane proteins. Translocation into the ER can take place after translation - "post-translational", or during protein synthesis, - "co-translational". In mammalians, proteins consisting of fewer than 120 amino acid residues reach the ER via the post-translational pathway. The following is required for both the co-translational and post-translational pathway: 1) identification of proteins and targeting to the ER; 2) association with the translocation machinery; 3) energy necessary for these processes; and 4) protein folding and maturation. Co-translational translocation is the primary conserved route in all organisms, and the translocon Sec61 complex is the main component. The passive Sec61 channel requires other components and energy provided by guanosine-5'-triphosphate (GTP) hydrolysis. Recently, the improved resolution of cryo-electron microscopy (cryo-EM) and cryo-electron tomography (cryo-ET) has contributed considerably to the understanding of the Sec61 structure. These 3D imaging techniques allow the visualisation of complexes in their physiological environment associated with native membranes when the structure does not exceed a certain thickness (0.5–1 μm). Sec61 spans the membranes multiple times and is made up of three different subunits—α, β, and γ. The subunit α forms a channel via ten transmembrane domains (TMDs)—five α-helix domains in the N-terminal and five in the C-terminal connected via a short hinge helix. The subunits β and γ are at the periphery of the channel with one TMD and a cytosolic N-terminus (type II). In contrast to the β subunit, subunits α and γ are conserved sequences that are essential for cell viability. This difference in essentiality is apparent under induction conditions; during ER stress, there is lower expression of the β subunit (Nagasawa et al., 2007; Linxweiler et al., 2017) (Fig.1.2.1). According to its channel structure, Sec61 has at least two functional states: 1) the non-inserting state (9–15 Å) and 2) the inserting state (diameter 40–60 Å). It is believed that Sec61 achieves the open state by the nascent polypeptide moving the "plug" inside the channel after interaction with a ribosome (Fig.1.2.1), and the interaction between subunits α and γ (Sec61) maintain this open state. The open state can accommodate the unfolded chain and α-helix (Dudek et al., 2015). Inside the channel there is also a "pore ring", the thinnest point where six hydrophobic residues lead to constriction during the closed state resulting in a barrier that prevents the passage of folded proteins. Nonetheless, other studies have proposed BiP as a necessary seal (Van den Berg et al., 2004). The channel is not selective and, therefore, small compounds, such as sucrose and glutamate, can go through, most likely when a non-translating ribosome interacts with the translocon (Lizak et al., 2008). Calcium leakage occurs throughout the channel but is partially prevented by BiP (Schäuble et al., 2012). When the plug is displaced, hydrophobic interactions are interrupted and the polypeptide with the signal peptide (SP) inserts as a loop. Then, Nin-Cout inverts and cleavage of the SP by the signal peptidase (SPase) occurs. The Sec61α TM2, TM3, TM7, and TM8 domains surround the nascent chain. TM7 mutants show defects in co- and post-translational translocation due to delays in channel gating (Trueman et al., 2011).

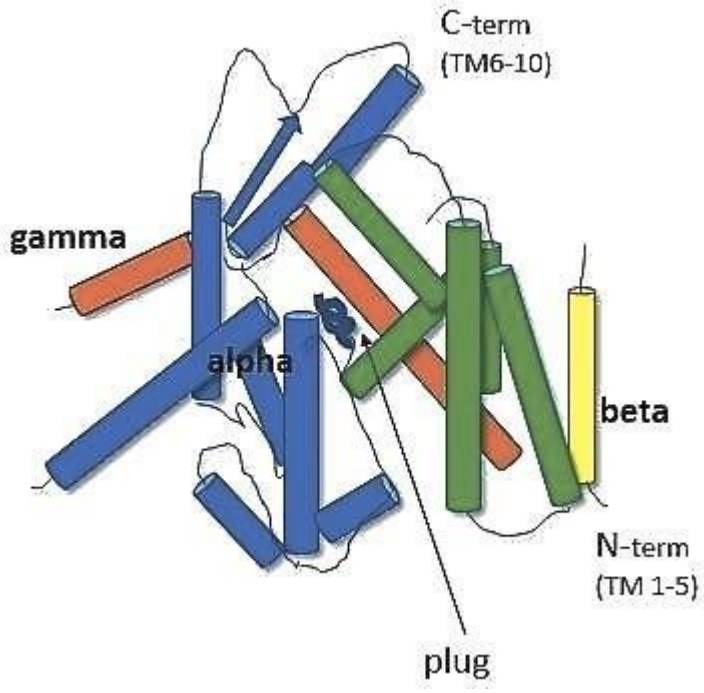


Fig. 1.2.1 – The translocon Sec61 is viewed from outside the ER membrane on the cytosolic side. The complex is made up of three subunits, α, β and γ: the N-terminal of α-subunit (TM1-TM5) (green), the C- terminal of α-subunit (TM6-TM10) (blue), the β-subunit (yellow), γ-subunit (orange), and the plug (blue).

The channel can open in two directions; inside (central pore) and laterally. The nascent protein reaches the membrane through the lateral gate (LG), a gap between two Sec61 TMDs that accommodates TMD α-helices (Fig. 1.2.2). This process occurs via the recognition of SPs/anchor-signals/TMDs, which is characterised by hydrophobic sequences and polar amino acid residues. Mutagenesis and structural analyses showed that the LGs and pores recognise hydrophobic segments (H-segments).

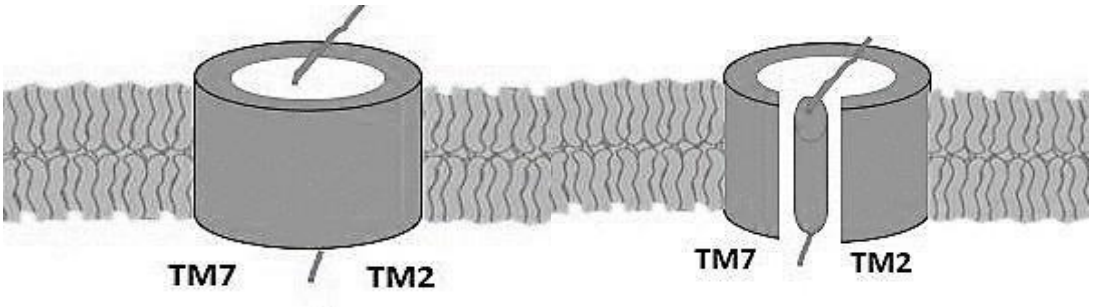


Fig.1.2.2 - Schematic representation of co-translational translocation of soluble and TM proteins: the soluble protein crosses the central pore and reaches the lumen (left). Instead, the lateral gate between TM2 and TM7 accommodates the TM protein that reaches the proper localisation in the ER membrane (right).

Ribosomes, via 28S rRNA backbones and uL23, eL19, and eL39 proteins at the exit-tunnel, interact with C-terminal cytosolic domains of Sec61 α, loops TM6/TM7 and TM8/TM9, and with the N-terminal of Sec γ (Voorhees et al., 2014; Voorhees and Hedge, 2016). Between the translocon and ribosome, there is a space of 10-12 Å. It is likely that these interactions lead to some conformational changes in the channel; the lessening of some contacts and the stabilisation of new conformation by hydrogen bonds. Single substitution at the cytosolic positively charged residues in Sec61 loop TM6/TM7 uncouples the binding with the ribosome and consequently protein translocation and membrane protein integration (Mandon et al., 2018). Specific characteristics of the SPs and mature proteins can further open the channel; non-clients of the translocon are rejected and do not reach the luminal side, even with an appropriate SP (Jungnickel and Rapoport, 1995). When the synthesised chain of 60–70 amino acid residues appears from the large subunit of the ribosome, it interacts with the signal recognition particle (SRP). The SRPs are lower in number compared to the ribosome-nascent-chain complexes (RNCs), and selection of the ribosomes directed to the ER is the primary role of the SRPs. Then, the conserved component of the SRP, SRP54, binds to the SP/SS/TMD, and translation stops until the complex (ribosome-SRP) reaches the SRP receptor (SR) located on the ER membrane (Meyer et al., 1982). The targeting is controlled by changes that follow cargo loading and GTP dimerisation (Lam et al., 2010). The SR consists of two subunits, α and β, and both are GTPases. The former is attached to the membrane via the beta subunit where the Sec61 translocon is also present. SRP and SR intercede on the transfer of the ribosome/polypeptide to Sec61 (Linxweileet al., 2017). Translocation and maturation of the proteins are highly regulated processes that consist of multiple steps: 1) ribosome/SRP/translocon interaction; 2) translocation through the translocon; and 3) post-translocation modifications (PTMs) and folding (Tyedmers et al., 2003). The most common post-translocation modifications are N-glycosylation, disulphide bridge formation, and phosphorylation (Shental-Bechor and Levy, 2009). N-glycosylation is the prevalent modification in eukaryotic cells; it reduces aggregations and increases folding and thermodynamic stability (Price et al., 2012). Asparagine-linked glycosylation (ALG) is carried out by oligosaccharyltransferase (OST) located near the translocon Sec61 (Fig. 1.8.4, p. 19). Protein glycosylation is also essential for other biological processes, such as attachment to the extracellular matrix, protein-protein interactions (PPIs), and homeostasis (Murray et al., 2015). The chaperone BiP and other proteins maintain the polypeptide in an unfolded state allowing the glycosylation step (Lakkaraju et al., 2012).

1.3 Translocation regulation and Sec61 accessory components

Multiple pathways exist for ER protein translocation in addition to co-translational and post-translational pathways. However, all of these routes require; i) recognition by the translocation machinery of the nascent protein; ii) recognition by accessory elements, and iii) cell availability of these auxiliary components.
The proteins and complexes involved in ER protein transport belong to three groups:

1) Targeting components and cytosolic chaperones, such as SRP/SR.
2) Auxiliary components, such as BiP, translocating chain-associating membrane (TRAM), translocon-associated protein (TRAP), Sec62/63, ERj1, calnexin, and calreticulin.
3) Modifying enzymes, such as OST (Fig. 1.3.1).

More than 20 integral membrane proteins are involved in ER protein translocation (Voorhees et al., 2016).

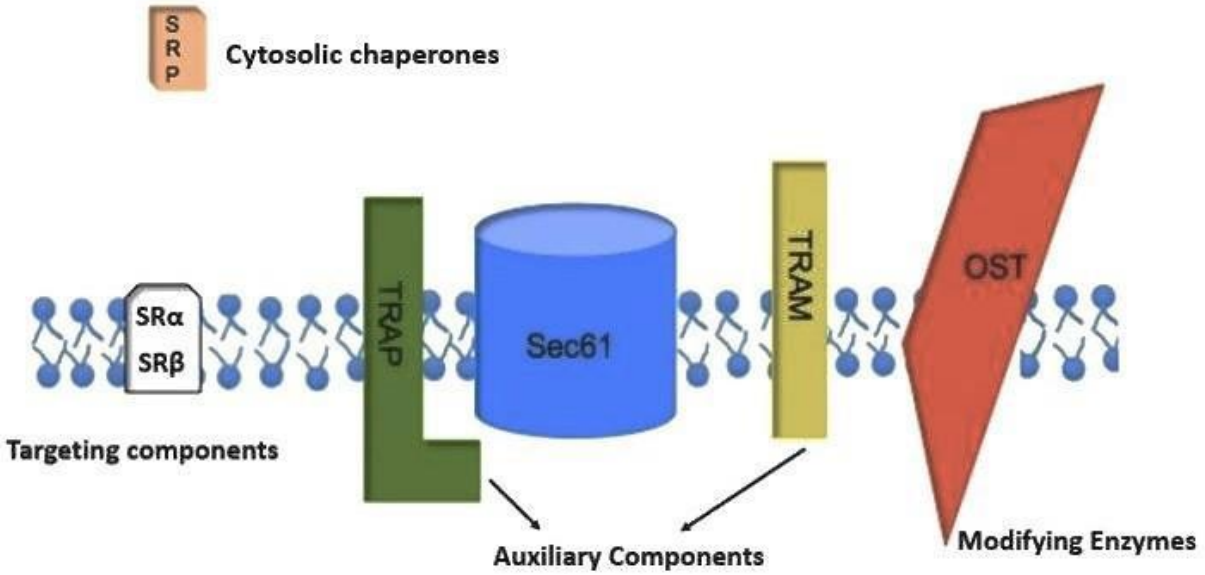


Fig. 1.3.1 – Essential groups of proteins or complexes during ER protein translocation: cytosolic chaperones, targeting components, auxiliary components, and modifying enzymes.

The Sec61 translocon is involved in other translocation pathways in addition to co-translation translocation: i) post-translation translocation and ii) retro-translocation of proteins for degradation, wherein Sec61 interacts with ERAD substrates and the proteasome (Kalies et al., 2005; Scott and Schekman, 2008). In mammalians, during co-translation translocation, Sec61 associates with Sec62 and Sec63, the latter by Sec62 interacts with ribosomes. This interaction is essential for the translocation of some substrates (Muller et al. 2010; Lang et al., 2012). The interaction of Sec63 with the translocon allows the membrane chaperone ERj1 (Hsp40) to recruit BiP via its luminal domain. Then, ERj1 dissociates from the ribosome tunnel and the interactions between ribosomes and Sec61 occur (Blau et al., 2005) (Fig. 1.3.2). The SEC genes are extensively involved in important diseases; for instance, mutated Sec61γ is involved in glioblastoma (Linxweiler et al., 2017) (Fig.1.3.2).

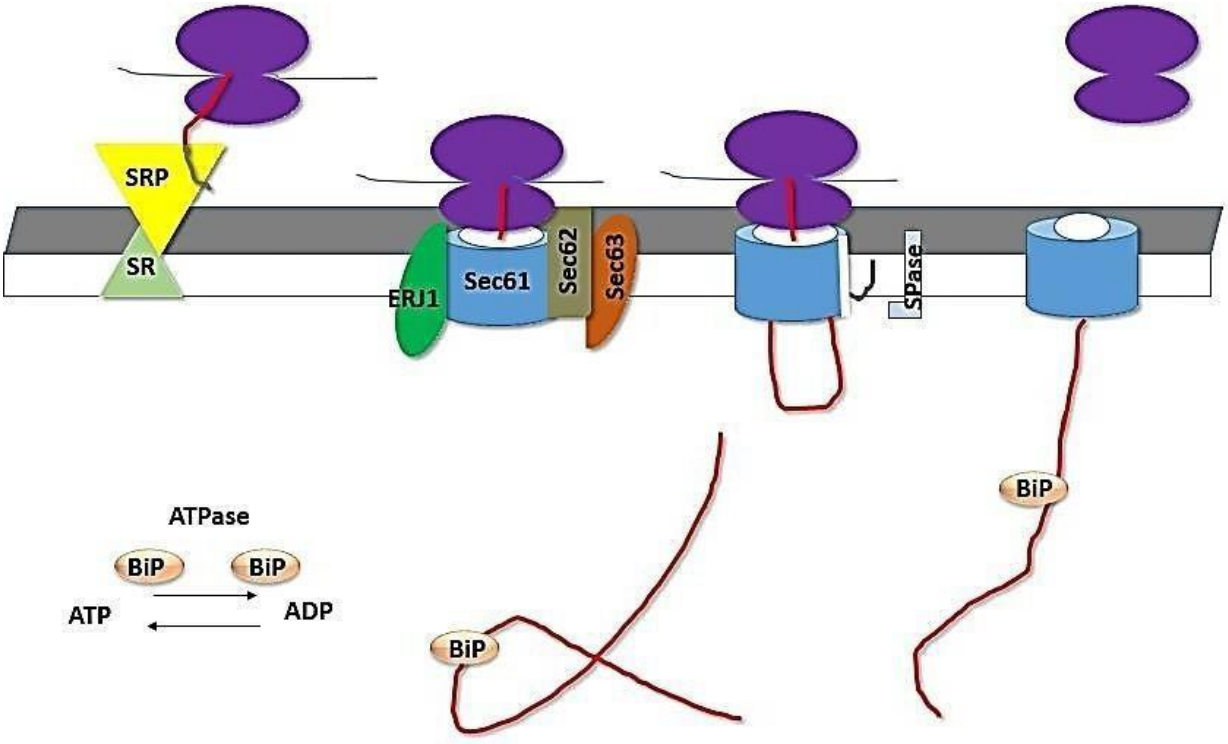


Fig. 1.3.2 – Co-translational translocation: ribosome-SRP complex interacts with SR, then with the translocon. Sec63 and Sec62 interaction allows ERj1 to recruit BiP, the latter binds to the nascent polypeptide in ATP-dependent manner.

The SPs, anchor-signals, and TMDs are recognized by the SRP when they emerge from the ribosome. The SRP has at least three essential roles: 1) recognition of proteins with cleavable SP; 2) recognition of proteins with anchor signals and TMDs; and 3) maintaining the specificity of organelle targeting (ER). Other co-translational translocation targets are possible (mitochondria and chloroplasts). After the SRP-SR dissociates, the nascent polypeptide is inserted into the Sec61 channel via random Brownian ratchet, when BiP binds to the transient polypeptide (Fig. 1.3.2). Cryo-EM analysis have shown that during nascent chain synthesis, an α-helix formation occurs inside the channel with some concomitant folding (Cabrita et al., 2016). Typically, the substrates reach the ER lumen via a loop; the N-terminus tail faces the lumen, the SP is cleaved off by SPase, and the rest of the chain crosses the channel (Hedge et al. 2008), (Fig. 1.3.3). The SP cleavage is a vital function; the accumulation of pre-proteins at the membrane leads to cell death (Auclair et al., 2012).

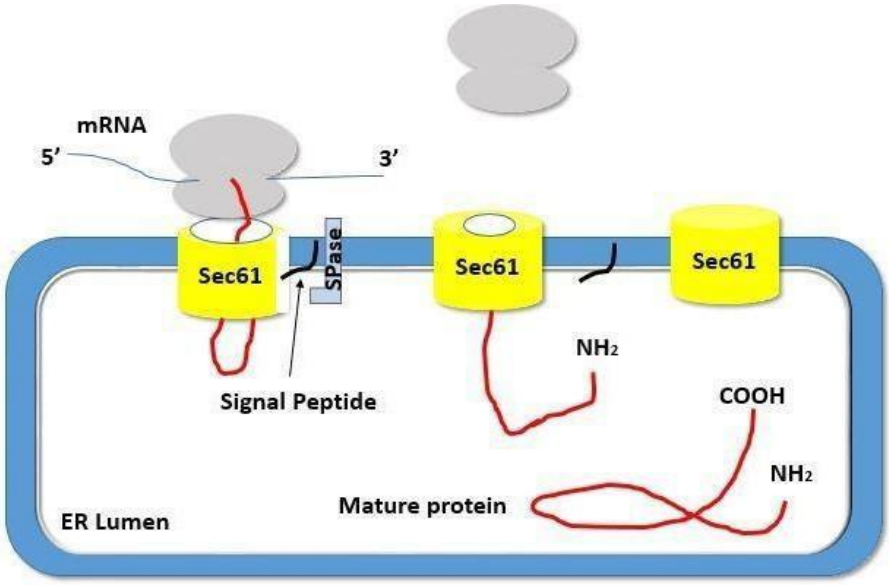


Fig. 1.3.3 – Co-translational translocation of secretory and luminal proteins into the ER: the polypeptide forms a loop, the SP inserts in the membrane through the LG, and is cleaved off by the SPase. The mature protein crosses the channel and reaches the lumen.

The TM proteins reach the membrane through the Sec61 lateral gate, a gap between TM2/3 (N-terminus) and TM7/8 (C-terminus), which allocates the SPs. Anchor-signals, and α-helices of TMDs (Egea et al., 2010). Sec61 can translocate secretory and TM proteins with the same classes of polysomes and not different subclasses, differently to what was previously hyphotesized by G. Palade (1975). The translocon, by recognizing different signal characteristics, such as hydrophobicity and polarity, can discriminate between the different classes of proteins and translocate them with higher or lower efficiency to the desired destination (Kida and Sakaguchi, 2018).

1.4 SP: from sequence to sorting

The signal hypothesis proposed by G. Blobel in 1971 and demonstrated by Blobel and Dobberstein in 1975 explains how membrane and secretory proteins reach the ER. The SP has a variety of different functions, such as the prevention of early folding, interaction with SRP, interaction with the translocon (gating)/accessory components, interaction with SPase, and cleavage timing, as well as post-cleavage functions, such as antigen presentation. The signal sequence (SS) or SP is a short and transient α-helical and beta-sheet sequence present at the amino terminus of many secretory and TM proteins in prokaryotes and eukaryotes. According to the UniProt list, the SP ranges from 16 to 30 residues in 84% of the proteins, and in 99% has fewer than 50 residues (Jarjanazi et al., 2007). The SPs have the same structure but are heterogenous in peptide sequence composition, and the only shared characteristic is a hydrophobic core of at least six amino acid residues.
The SP has a tripartite structure:

- n-region positively charged (or with polar residues).
- h-region, hydrophobic residues.
- c-region, polar, present at the cleavage site (Fig. 1.4.1).

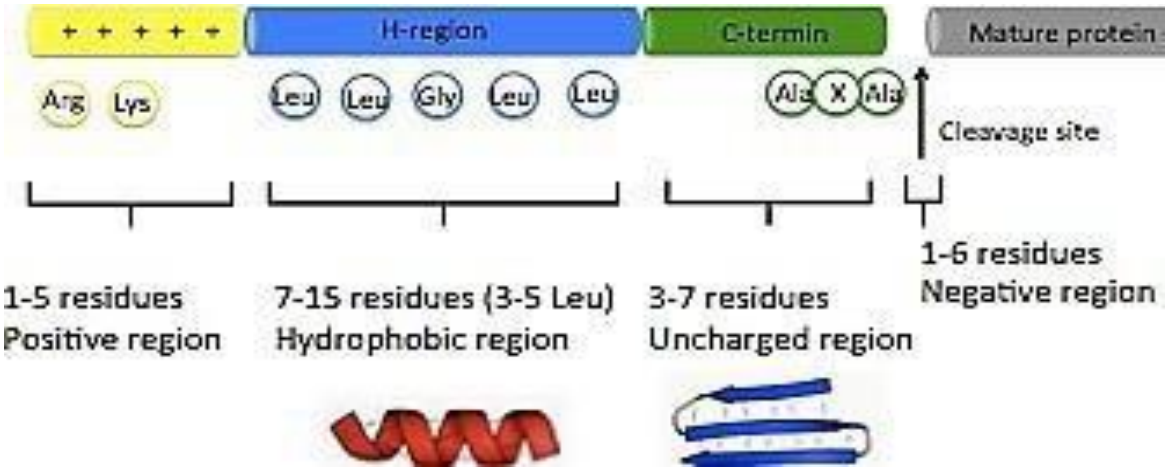


Fig. 1.4.1 - The signal peptide, typically, has a tripartite structure: N-region with positive or polar amino- acid residues (yellow), H-region highly hydrophobic (blue), and C-region where the signal sequence is cleaved off by the signal peptidase (green).

Another SP characteristic is the presence of two RR residues (arginine positively charged) in the upstream h-region (Fig.1.4.2).

Fig. 1.4.2 – The RR residues may be present in the upstream "h region" of the SP (arrow).

The hydrophobicity of the SP is essential for proper protein translocation, which has been confirmed in numerous studies since the 1990s (Jarianazi et al., 2007); Wahlberg et al., 1997; Sakaguchi et al., 1992). SRP binds to the SP via hydrophobic interactions. The amino acid residues in the SP n-region and the SRP RNA phosphate backbone may also play a role either directly or by altering the α-helix length of SP h-region. The basic residues are necessary when hydrophobicity is below a particular threshold (Peterson et al., 2003). Leucine is the most abundant hydrophobic amino acid residue in the wild type (Nilsson et al., 2015). The cleavage site presents a “short-side amino acid” at position -1 an “uncharged amino acid” at position -3 at the C-terminus. However, the SP n- and h-region properties can also influence cleavage. Small neutral residues, such as alanine, glycine, serine, and threonine are present at position -1 and -3 preceding the cleavage site. The sequence AXA (alanine) is present in some SPs, and this domain makes cleavage site recognition easier. Glycine and proline, which interrupt the helices, are present in TMDs but are less common in the hydrophobic core of the SP. In addition, tyrosine and asparagine are found in TMDs but rarely in SPs (Buske et al., 2008).
Approximately 40% of human protein-coding genes (19,000–20,000) contain the SP. Some protein classes present similarities between their SPs, such as human PDGF, VEGF, and neurotrophins (A. Russo, unpublished). This reinforces the hypothesis that SPs may be functionally distinct and optimised

based on their mature protein (Kim et al., 2002).
SP has essential roles during co-translational translocation: i) ability to be recognised by the SRP; ii) a gating step to initiate translocation by the N-terminus with a pulling force; and iii) inversion to acquire Nin-Cout orientation for cleavage (Kriegler et al., 2018; Fons et al., 2003). It is likely that SP is involved in critical checkpoints defined by binding, induced fit, and proofreading kinetic mechanisms (Zhang et al., 2013) as plausible steps when SP interacts with SRP and the translocon.

1.5 Transmembrane proteins: the connection between two environments

In prokaryotes and eukaryotes, 20–30% of genes express integral membrane proteins. The membrane proteins perform different functions, such as signal transduction, conduction, transport, and protein-protein interactions. Membrane proteins connect two environments, such as Sec61, ERj1, and calnexin, that connect cytosol and ER. Many TMDs are α-helices (helix-bundle class) that cross the lipid-bilayer by single- or multiple-spans. There are also β-barrel TM proteins in bacteria, mitochondria, and chloroplasts (Spiess et al., 2019). The thickness of the fatty acyl chain of the membrane lipid bilayer is ~3 nm; consequently, a transmembrane domain of ~20 residues and five or six helical turns are necessary to span the entire membrane. In the majority of genomes, the positive charges in proteins increase at the N-terminus, which interact with the negatively charged ribosome exit tunnel. Additionally, these charges are essential in membrane proteins for TMD topology (Charneski and Hurst, 2014). Transmembrane proteins can assume two different orientations depending on the hydrophobic core, the difference in net charge between the Nterm and Cterm, and the protein length (Spiess, 2019). Long proteins tend to have Nlum and Ccyt (type I), whereas small proteins tend to have the opposite orientation (type II). Usually, proteins with multi TMDs have N- and C-termini in the cytosol (Von Heijne, 2006), such as the TRAP gamma subunit. The internal signal-anchor stops the translocation through the channel until complete synthesis of the polypeptide. Then, the polypeptide moves laterally through the LG until it reaches the phospholipid bilayer. Three classes of SSs are present in membrane proteins: 1) classical SPs with insertion Nlum/Ccyt; 2) signal-anchors with orientation Nlum/Ccyt that function as stop-transfer/TM anchors in the membrane bilayer; and 3) reverse signal-anchors that insert in the opposite orientation Ncyt/Clum. Many positive charges (Lys, Arg) are present in the non-translocated sequences of the membrane proteins (Hessa et al., 2005; Elofsson and von Heijen, 2007). This accounts for the "positive-inside rule"; the lipid bilayer of the ER membrane is asymmetric and contains anionic phospholipids on the cytosolic face (Shao et al., 2011). The TM proteins with cleavable SP always have the N terminus on the luminal side (type I); this orientation is present in three TRAP subunits. The subunits, ssr1, ssr2, and ssr4, are single-spanning integral proteins with a SP between 17 and 23 residues long and a luminal N-terminus. In 1987, a short sequence in the C-terminus of an adenovirus membrane protein was discovered, identifying it as an ER-resident protein (Pääbo et al., 1987). ER TM proteins type I, usually, have a specific retrieval and retention domain; two lysines at positions three and four – X(5)K(4)K(3)X(2)X(1)-C-term or three and five – KXKXX. The lysine in position four can be in position five without compromising the function (Jackson et al., 1990). An arginine can substitute the lysine in position four and the protein will remain in the membrane (Shin et al., 1990). In particular, the most crucial lysine is in the third position. Retrieval and retention ability also depends on the length of the cytosolic domain; a minimum distance between the lysines and the TMD is necessary (Vincent et al., 1997). TRAP β subunits present this retention motif at the C-terminus with the lysines at positions -3-5 (Human and Mus musculus; Results, p. 79) but this domain is absent in the other TRAP subunits. Several mechanisms are responsible for retention when the motif is absent, such as structure, hydrophobicity, and charge. Recent studies have shown a significant bias between the hydrophilic and hydrophobic amino acids at the N- versus C-terminal of TM helices (Park and Helms, 2008). In the ER membrane, protein orientation depends on three factors: 1) N-terminal without a stable tertiary structure; 2) distribution of charged amino acids within the TM domain(s); and 3) length of the hydrophobic sequence that supports the orientation of the

N-terminal into the ER. The single-spanning membrane proteins reach the final localisation by SPs/anchor-signals. The initial insertion of the polypeptide establishes the membrane-protein topology (Lao et al., 2002; Singh et al., 2013). Instead, the multi-spanning proteins for insertion and localisation rely on interactions with the proximal TMDs. When the polypeptide exits the Sec61 LG, helix-helix interactions occur (Cymer et al., 2015) (Fig.1.5.1). Typically, these interactions are caused by conserved domains, such as GXXXG, QXXS, glycine, and leucine zippers. Frequently, polar amino acid residues are also present to form hydrogen bonds. Charged amino acids are essential for the structure (Fink et al., 2012). In addition, many non-polar amino acid residues cause the TMDs to be hydrophobic, which suggests integration in the membrane lipid bilayer. The ER membrane is more hydrophobic than the SSs and TMDs. Localisation in the membrane is a balance between different forces; a thermodynamic equilibrium (Rapoport et al., 2004). This process of integration is called "Lipid Partitioning"; the membrane protein leaves the aqueous channel and moves into the membrane-lipids, where its hydrophobic sequences segregate (Heinrich et al., 2000).

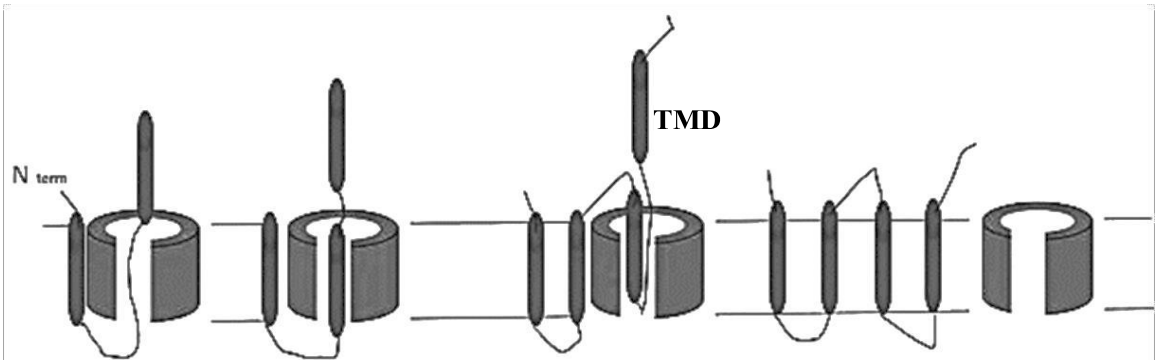


Fig. 1.5.1 - The schematic cartoon shows how the multi-spanning proteins insert in the membrane, for instance, TRAP γ, Sec61α. The interaction with the neighbouring TMD is essential.

1.6 Quality control in the ER: an extensive network

The synthesis of cytosolic and secretory proteins has mechanisms of quality control during different steps, such as transcription, translation, folding, and assembly. In the ER, chaperones and foldases ensure the correct folding of the translocated proteins; the former prevents aggregation and the latter performs the folding steps. The recognition of unfolded proteins activates the UPR, which leads to the ERAD pathways. The proteasome, a prominent structure, degrades the proteins after the attachment of multiple copies of ubiquitin (protein hydrolysis). The ubiquitin-proteasome system is an essential cell component; it controls many other processes beyond degradation (cell cycle, signal transduction, DNA repair, chromatin remodelling, cell death, immune responses, and metabolism) (Demartino and Gilette, 2007; Tanaka, 2009). Three main steps are necessary for ERAD: 1) recognition and targeting, 2) retro-translocation and ubiquitination, and 3) proteasome targeting and degradation. The UPR also occurs due to the production of proteins overcoming the necessity of the cell; the nascent proteins misfold and aggregate because of their high concentrations (300–400 g/L) (Braakman et al., 2013). Moreover, the perturbation of any process in the ER, such as protein synthesis, transport/phospholipid synthesis, and distribution/calcium storage, drives ER stress and UPR. The ubiquitous ER membrane proteins in UPR are Ire1α, PERK, and ATF6; three different pathways that stimulate transcription factors to express the ER chaperones and ERAD components (Fig.1.6.1). The upregulation of folding and degradation, and the downregulation of protein synthesis alleviates stress. UPR also triggers pro-apoptotic pathways that are controlled by calcium concentration in the ER, mitochondria, and cytosol. Under these circumstances, the proteins remain in the ER until apoptosis occurs or they are retro-translocated for proteasome degradation.

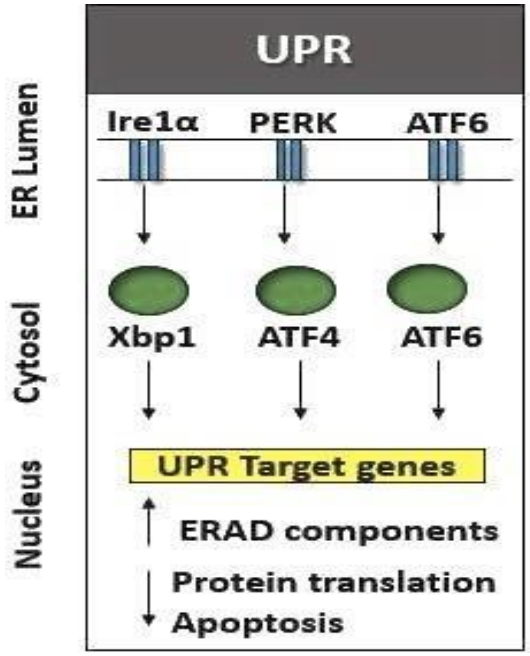


Fig. 1.6.1 – Different UPR pathways: Ire1alpha, PERK, and ATF6 lead to activation of transcription factors. The transcription factors target specific genes for expression of ERAD players.

To differentiate the abnormal proteins from the nascent proteins, hydrophobic sequences on the protein surface are recognised based on their unfolding state. The N-glycosylation and Man8 forms are also determinant. The sugar moiety Glc3Man9GlcNAc2 attaches to the asparagine residues of nascent proteins. The terminal GlcNAc2 is cleaved off by glucosidase I and II. When the protein remains unfolded, mannosidase removes the terminal mannose residue (Man9). The Man8 form is recognised by lectins (ERAD players) that target the protein to the retro-translocation pathway. When the UPR is active, the glycan-dependent chaperones calnexin and calreticulin, both calcium-binding proteins (co-evolution with Asn-linked glycosylated proteins), retain the unfolded proteins. The second glycan independent system associates with BiP regulates the UPR pathway by activating transducers (Ma and Hendershot, 2004). All proteins are molecular chaperones involved in the folding process, in addition to being unfolding sensors. The degradation of the proteins does not always occur in the same manner. Further studies are needed to shed more light on the ERAD pathways.

1.7 EF-hand calcium-binding domains: structural and regulatory

Calcium has regulatory and structural roles and is a crucial element both outside and inside cells. Outside the cells, its concentration is approximately 10-3 M, while inside the cells, it is 104 times lower and mainly concentrated into the ER. Calcium is an essential primary and secondary messenger that influences apoptosis, and many proteins bind Ca2+ to maintain/change their structure and carry out biological functions. Proteins bind calcium via the motif DXDXDG included in a linear sequence of about 30 amino acid residues, where two perpendicular α-helices form the 12-residue Ca2+-binding loop. The binding residues are in positions 1,3,5,7,9, and 12, with the latter always being Glu (E) or Asp (D), which are negatively charged residues that interact with the positively charged Ca2+. These canonical EF-hand domains are located in calmodulin proteins. There are also non-canonical EF-hand domains or EF-hand-like domains that are mostly present in the N-termini of S100 and S100-like proteins. However, canonical and non-canonical domains can be present in the same protein (Results p. 73). The EF-hand domains in their conformation are open or closed and dynamic or static (Denessiouk et al.,2014).

The Ca2+-binding proteins are:
1) Signalling proteins and calcium sensors.
2) Buffering/transport proteins that control Ca2+ levels in cytoplasm.

1.8 TRAP complex: a Sec61-associated component

The presence of accessory structures that carry out specific function(s) during translocation is an essential aspect of translocation machinery. Two auxiliary components are the TRAM protein and the TRAP complex (Snapp et al., 2004). TRAM proteins are involved in co-translational translocation. Some nascent proteins are TRAM-dependent but other substrates do not rely on this protein, and it is likely that this depends on SP characteristics. In particular, when the SP of these substrates is cleaved, the crosslinking is lost (Walter, 1992; Görlich et al., 1993; Voigt et al., 1996). TRAP is a ubiquitous protein complex present in all eukaryotes. In mammalians, it is a heterotetrametric complex with a molecular weight of approximately 150 kDa. All four subunits, previously known as signal sequence receptors (ssr), are membrane proteins: α (ssr1), β (ssr2), γ (ssr3), and δ (ssr4). TRAP α, β, and δ are single-spanning protein type I (Nlum/ Ccyt) with an SP; TRAP γ is a multi-spanning TM protein that crosses the membrane four times and has a conspicuous cytosolic domain and no SP (Fig.1.8.1). Cryo-ET methods were previously employed to compare mammalian and algae complexes (the latter and plants lack the subunits γ and δ), resulting in the determination of the TRAP complex low-resolution structure (Pfeffer et al., 2017) (Fig. 1.8.2).

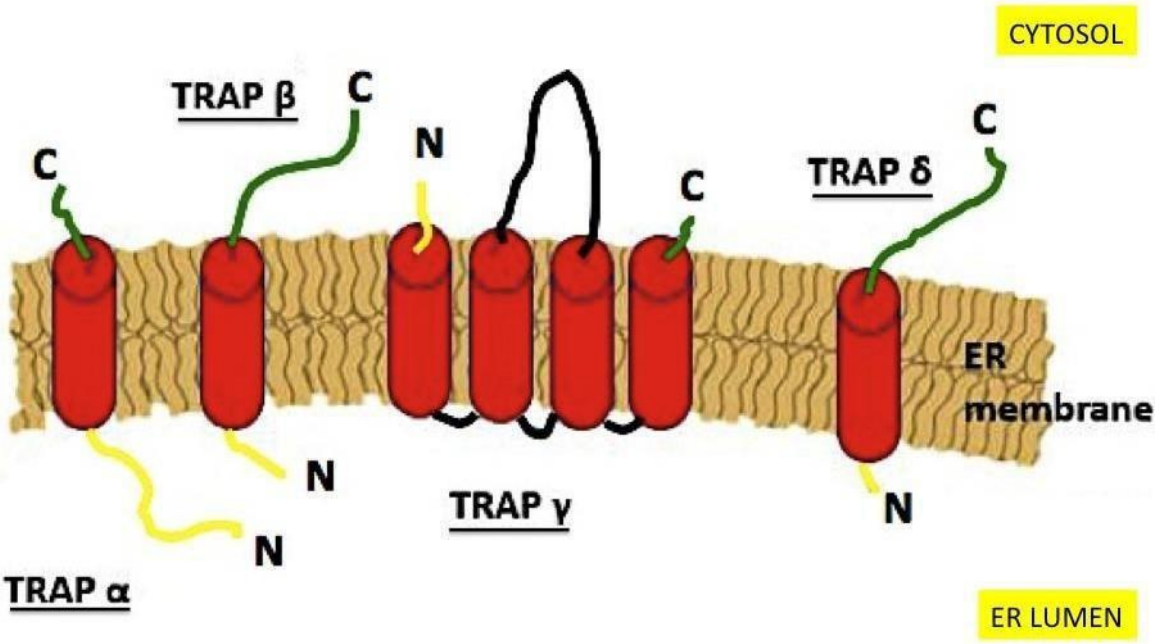


Fig. 1.8.1 – Schematic representation of TRAP complex: four different subunits, α, β, γ and δ. Alpha, beta and delta are single-spanning TM proteins (type I); instead, gamma has four TMDs (UniProt, Bano-Polo et al., 2017). N-terminus in yellow, middle of the sequence in black, and C-terminus in green.

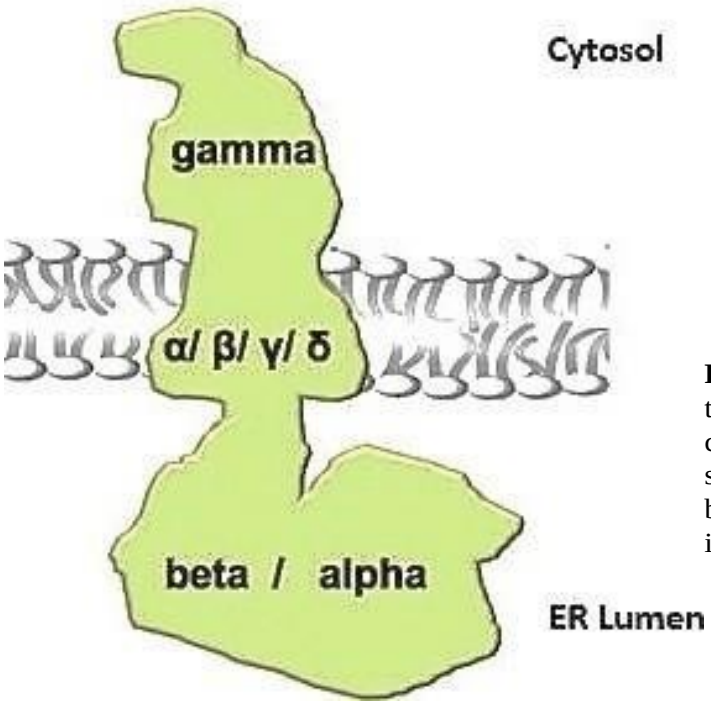


Fig. 1.8.2 – The cartoon shows the structure of the TRAP complex, four subunits, determined by cryo- ET. TRAP alpha and beta present a significant luminal domain, instead, gamma has a big cytosolic domain. Delta is mostly embedded in the membrane.

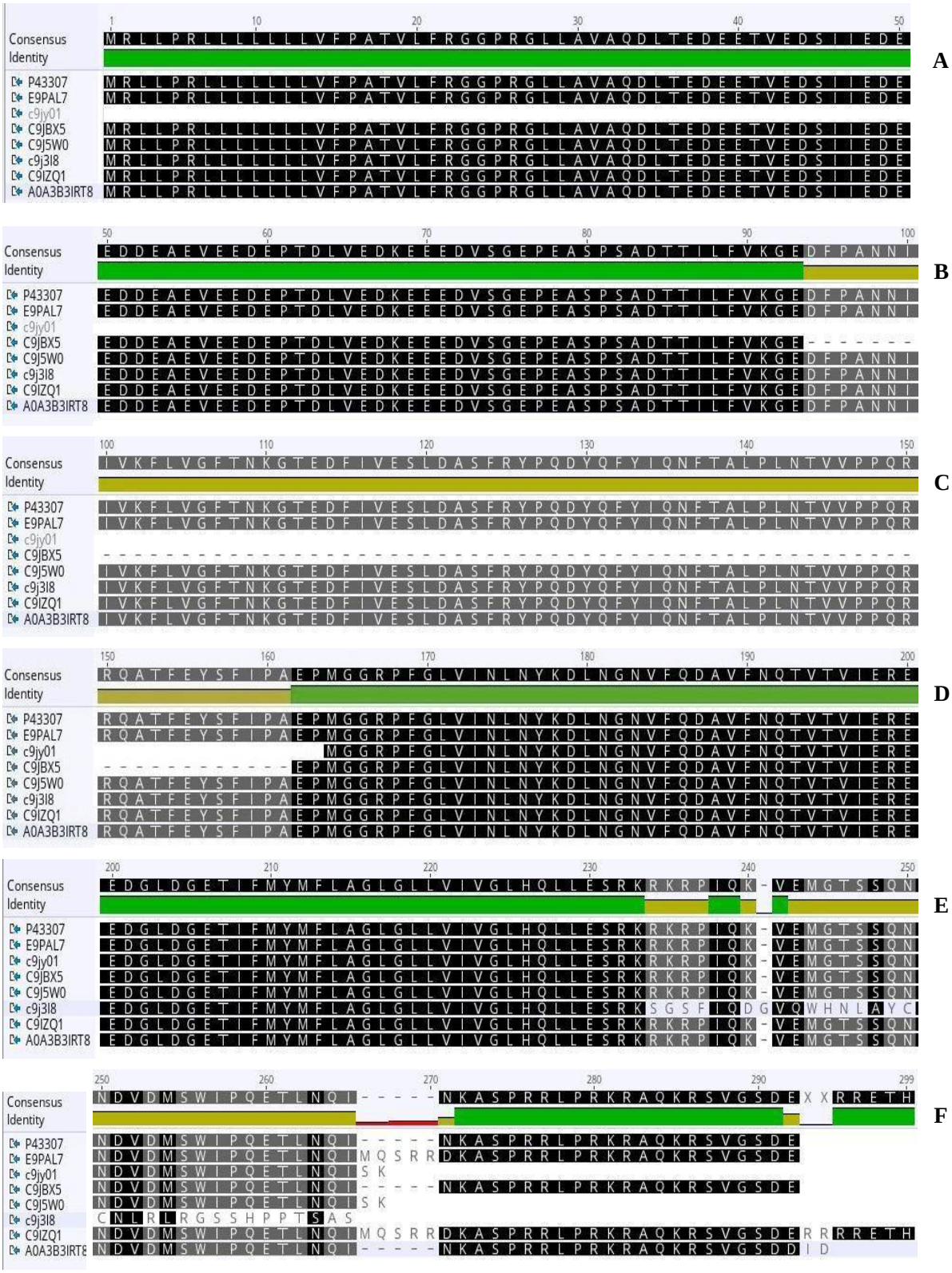
The human/*Mus Musculus* TRAPα subunit is a glycoprotein, and the gene is present on chromosome 6 with many isoforms present, though two are more common than others. The ubiquitous general form, which is conserved between different mammalian organisms, and another form only expressed in skeletal muscle. The general form has two mRNAs, alternative polyadenylation (2.7 kb and 1.2 kb) at the 3' non-coding regions. The mother supplies these until the eight-cell stage, then it is expressed during embryogenesis and in the adult. The other isoform is present in muscle tissue, including cardiac muscle, and is expressed after birth when the general form is turned off. The protein presents a longer C-terminus (1.8 kDa), 35% of which consists of arginine residues. Homozygous mutants die at birth for several cardiac defects. The subunit ssr1 could assist in the translocation of essential factors for heart cushion formation, such as interferon γ (γ-INF) and atrial natriuretic peptide (ANP) (Li et al., 2008). These proteins inhibit the transforming growth factor-β (TGF-β), which downregulates the development of mesenchymal cells in endocardial cushions. This deregulation leads to mouse death (Mesbach et al. 2006). The silencing of the TRAPα general isoform permits embryonic development progression because many cells are unaffected. Then, the defects that arise in the heart lead to mortality.

Other human TRAP α isoforms have been identified and are listed by experimental evidence on the table below (Tab. 1.8.1).

Entry Name	Length	Annotation Score
P43307	286	●●●●●
C9JBX5	218	●●
C9IZQ1	298	●●
E9PAL7	291	●●
C9J5W0	266	●●
C9J3L8	265	●●
C9JY01	103	●
A0A3B3IRT8	288	●

Tab. 1.8.1 – Human TRAP α (ssr1) isoforms listed by experimental evidence (UniProt).

The alignment of TRAP α isoforms shows complete match at the N terminus except for a shorter form, just 103 amino acid residues long (Tab. 1.8.2).

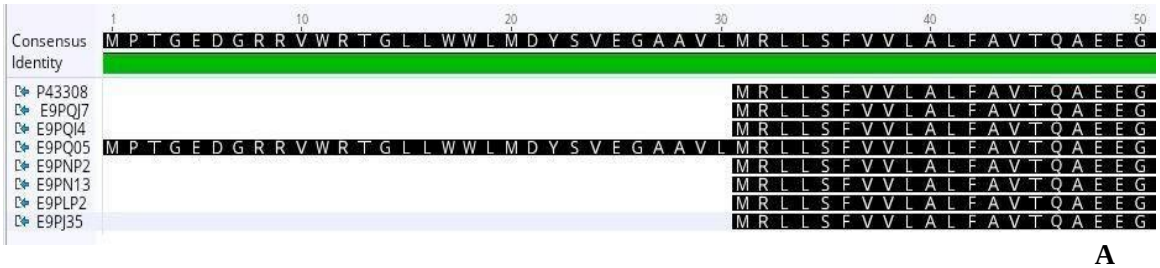


Tab.1.8.2 - Alignment of human TRAP α isoforms: same N-terminus except for the shortest form which is just 103 residues long (C9JY01). A = 1- 50 residues; B = 50 - 100 residues; C = 100 – 150 residues; D= 150-200; E= 200-250; F= 250-299.

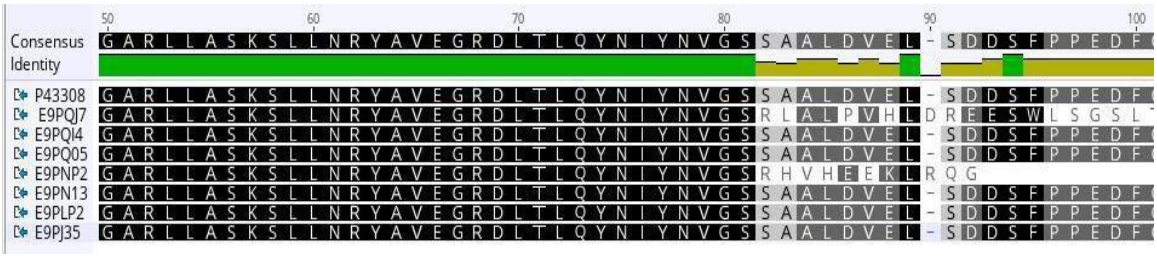
In addition to TRAP α, the transcripts of other subunits undergo to alternative splicing, in the tables below are shown the isoforms of human TRAP ß (ssr2), TRAP γ (ssr3), TRAP δ (ssr4) and the correspondent protein alignments (Tab. 1.8.3-1.8.8).

Entry Name	Length	Annotation Score
P43308	183	●●●●
E9PQI4	147	●
E9PQ05	114	●
E9PJ35	133	●
E9PLP2	90	●
E9PN13	93	●
E9PNP2	62	●
E9PQJ7	86	●

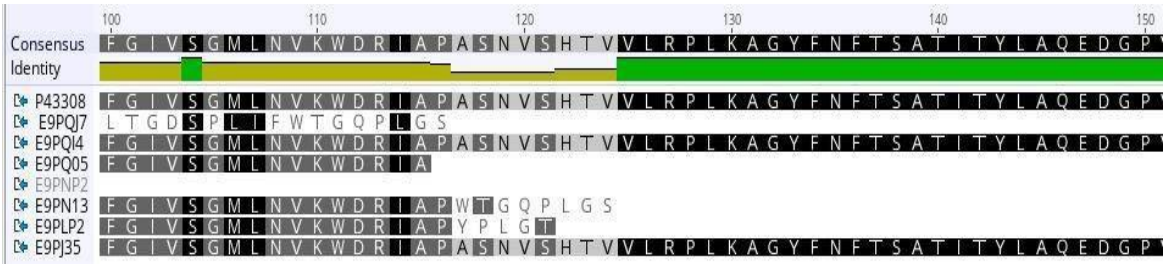
Tab.1.8.3 – Human TRAP ß isoforms listed by experimental evidence (UniProt).



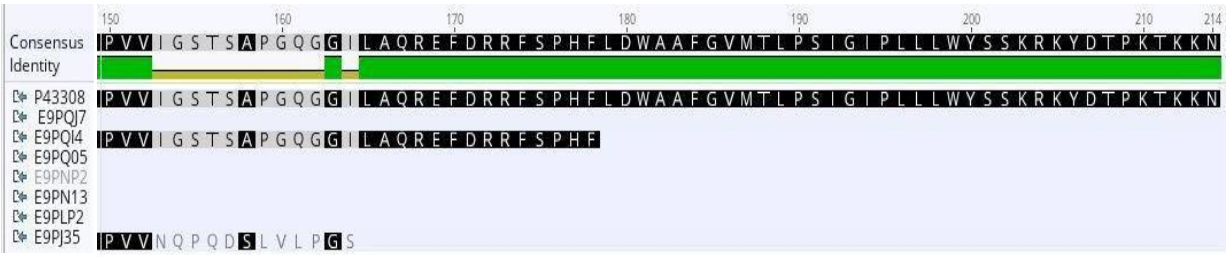
A



B



C

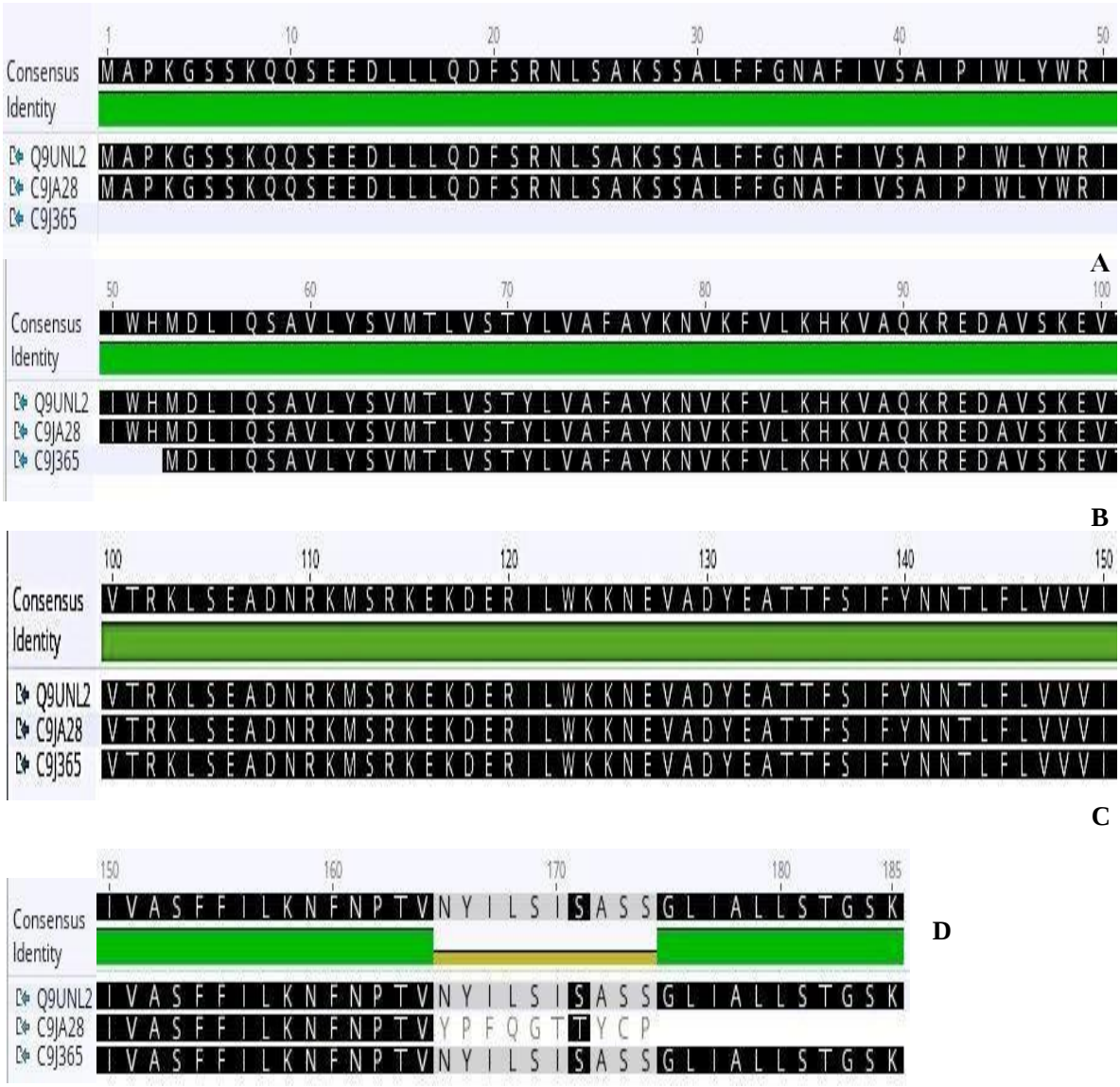


D

Tab.1.8.4 - Alignment of human TRAP ß isoforms: except for the form with 114 amino acid residues (E9PQ05) the N-terminal tail is same. A= 1-50 residues; B = 50 – 100 residues; C=100-150 residues; D= 150-214 residues.

Entry Name	Length	Annotation Score
Q9UNL2	185	●●●●
C9J365	133	●
C9JA28	174	●

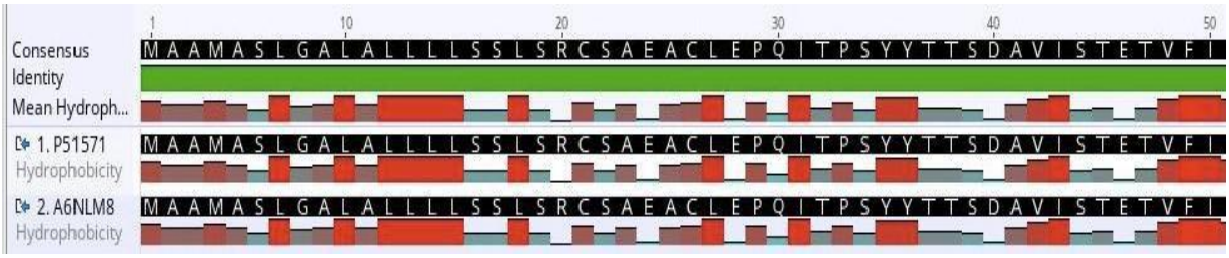
Tab.1.8.5 – Human TRAP γ presents three isoforms which are listed by experimental evidence (UniProt).



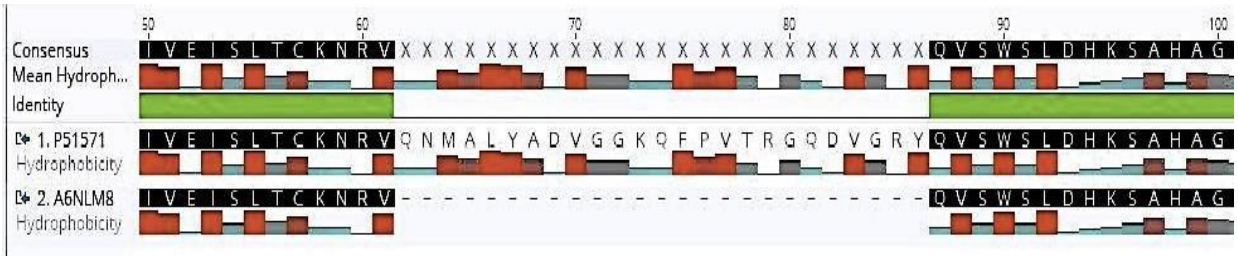
Tab.1.8.6 – The alignment of human TRAP γ isoforms points out a protein with a shorter N-terminus (Q9J365). A= 1-50; B= 50-100; C= 100-150; D= 150-185.

Entry Name	Length	Annotation Score
P51571	173	●●●●●
A6NLM8	148	●

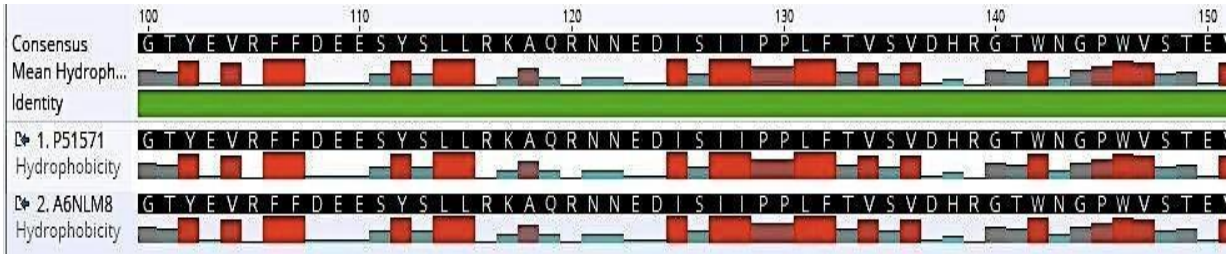
Tab. 1.8.7 – Human TRAP δ (ssr4) presents two isoforms which are listed by experimental evidence (UniProt).



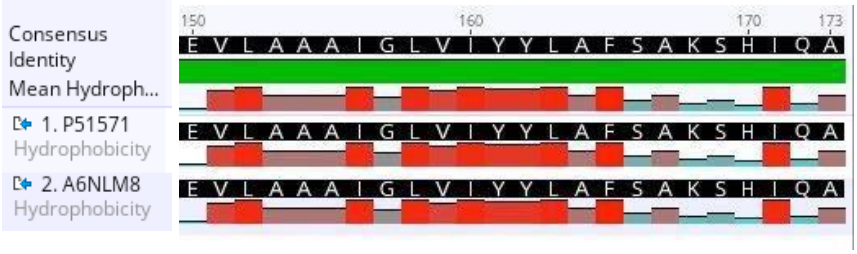
A



B



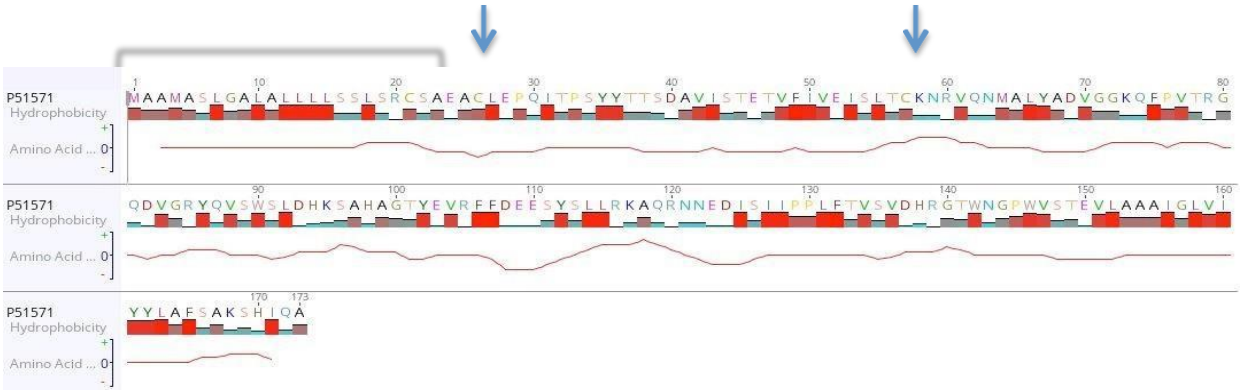
C



D

Tab.1.8.8 - Alignment of the two human TRAP δ isoforms: the short isoform of 148 residues (A6NLM8) lacks middle of the sequence; otherwise, the alignment matches 100%. A= 1-50; B= 50-100; C= 100-150; D= 150-173.

Mus musculus TRAP δ (ssr4) most common isoform forms a disulfide bridge on the ER luminal side, the cysteine residues are present in positions 3 and 34 in the mature protein (Hartmann et al., 1993); two cysteines in the same position are also present in human Trap delta (Tab. 1.8.9).



Tab.1.8.9 – Human TRAP δ protein sequence: signal peptide (bracket), two cysteine residues in the luminal domain form the disulfide bridge (arrows). Hydrophobicity (bars) and amino acid charges (red line).

The TRAP α subunit is crucial for mouse heart development. Other TRAP subunits are essential in some tissues during development. TRAP γ is essential to mouse placenta formation, and the silencing of this subunit leads to embryonic organ defects in the lungs. During placenta development, many secretory proteins, such as growth factors, cytokines, FGF, PDGF, EGF, and correspondent receptors, are expressed. The authors of this review believe that ssr3 is essential to the placenta vascular network, and may have a direct role in translocation, or indirectly by producing an uncoordinated TRAP complex (Yamaguchi et al., 2011). It is likely that TRAP γ interacts with ribosomes via rRNA or ribosomal protein L38 to stabilise the complex structure (Pfeffer et al., 2016). Moreover, TRAP γ is necessary for kidney development in mice (Mesbah et al., 2006) and Xenopus pronephros development (Li et al., 2005).

The ssr3 subunit, similar to other TRAP subunits, is involved in UPR pathways and cellular homeostasis maintenance (Yamaguchi et al., 2011). The IRE1α/XBP1 pathway induces TRAP expression; indeed, IRE1α knockout leads to the suppression of TRAP transcription. Interestingly, UPR inactivation by the IRE1α/XBP1 pathway leads to poor placenta vasculogenesis (Iwawaki et al., 2009). Moreover, the silencing of TRAP leads to reduced ERAD (Nagasawa et al., 2007) and TRAP binds misfolded proteins, such as superoxide-1 dismutase (Miyazaki et al., 2004). Additionally, the granulocyte-macrophage colony-stimulating factor (GM-CSF) induces TRAP α transcripts; another element related to the UPR (Hirama et al., 1999). Together, these studies suggest the role of the TRAP complex in the UPR, recognition of misfolded proteins, and ERAD.

The TRAP δ (ssr4) subunit is associated with a congenital disorder of glycosylation (ssr4 CDG) wherein the X-linked SSR4 gene is mutated. In the fibroblasts of these patients, the proteins are under-glycosylated and the overexpression of ssr4 partially recovers glycosylation. It is likely that the TRAP complex interacts with OST subunits SST3, DAD1, and DDOST. The latter two are essential for OST complex stabilisation. The interactions with ssr4 and DAD1 may play a role in pancreatic beta-cell survival in type 2 diabetes (Sanjay et al., 1998; Singh et al., 2013). In "ssr4 CDG", non-glycosylated proteins induce ER stress but the ERAD response is reduced because of the lower expression of the TRAP subunits.

Systematic microscopy analyses have shown that TRAP is always present at the back of the channels, and represents approximately 25% of the total volume made up of Sec61 and TRAP; the stoichiometry between Sec61 and TRAP is 1:1. In Fig.1.8.3, a comparison of an ER membrane-associated ribosome with and without the TRAP complex determined by cryo-EM is shown (Menetret et al., 2005, 2008). In 2008, Menetret et al. detected single copies of Sec61 and the TRAP complex associated with the non-translating ribosome.

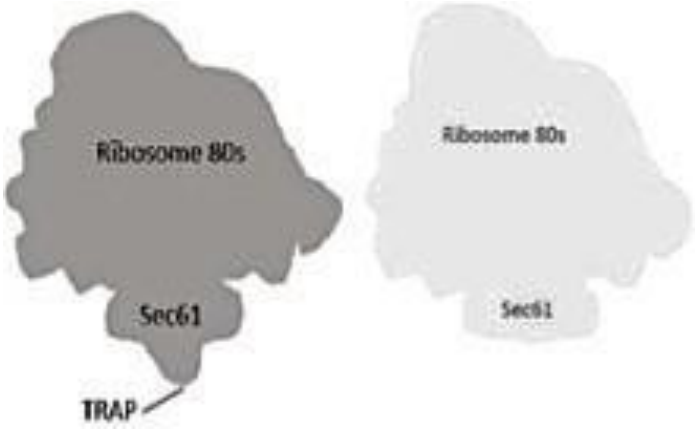


Fig. 1.8.3 – The ribosome-translocon complex (RTC): the frontal view of the ribosome, Sec61, and TRAP under the channel (left); frontal view of the ribosome and Sec61 (right).

In 2015, Pfeffer et al. used rER vesicles isolated from canine pancreases and CET/subtomogram analysis to determine the structure of the ER-membrane-associated ribosomes. It was found that Sec61 is in an open state only when associated with ribosomes. TRAP is always present, and OST is present in 40–70% of the complexes (Pfeffer et al., 2015) (Fig. 1.8.4).

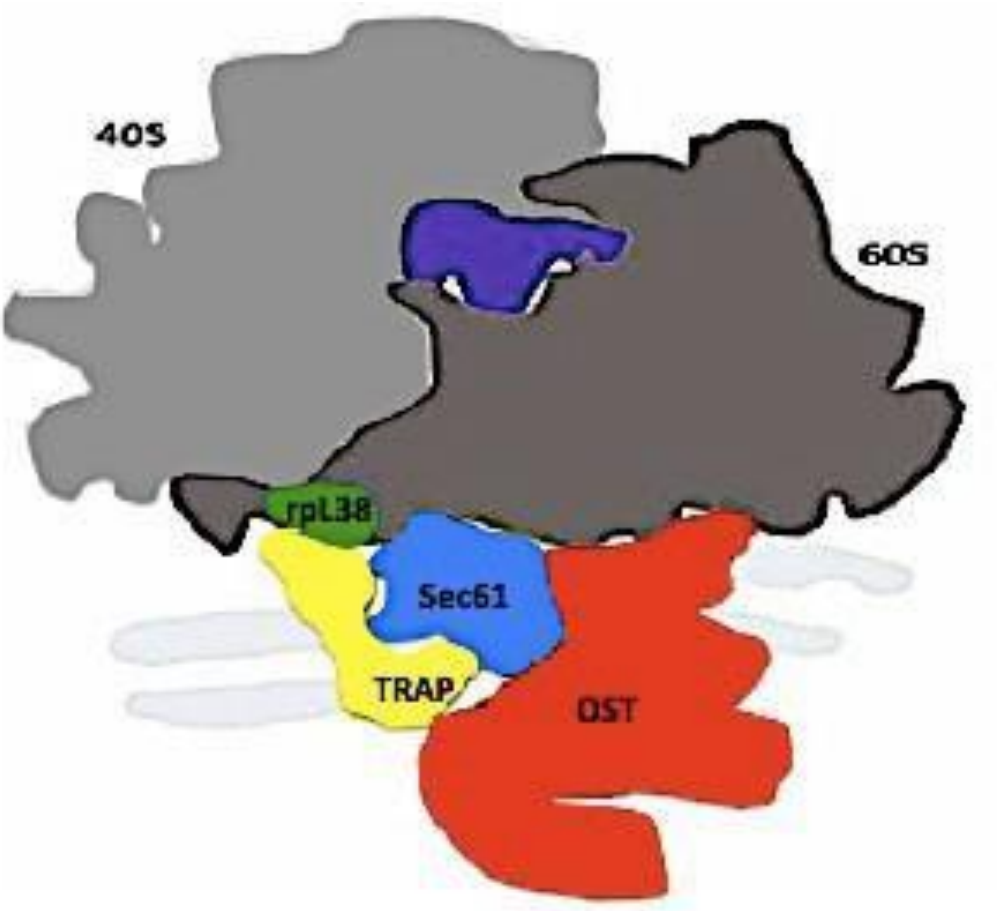
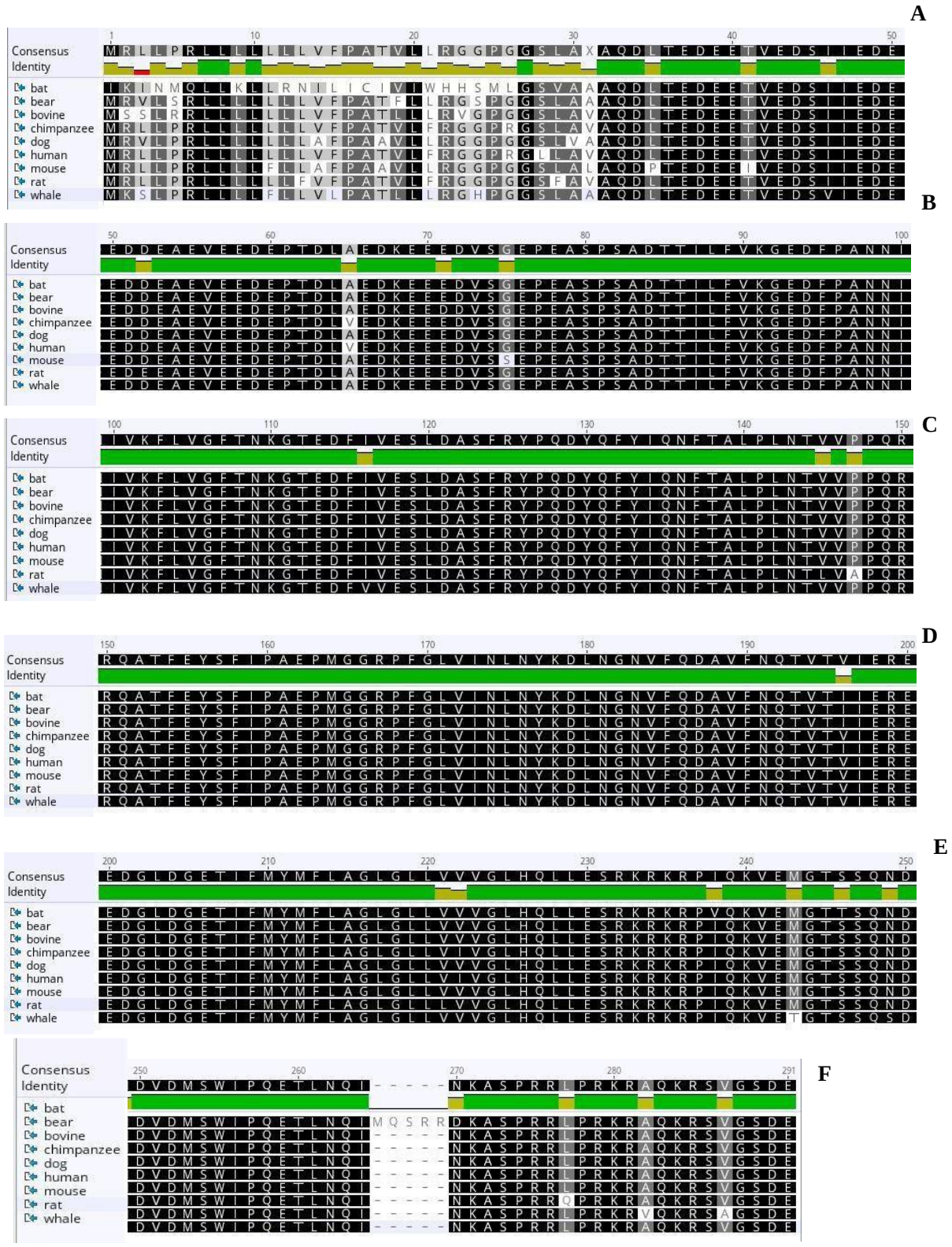
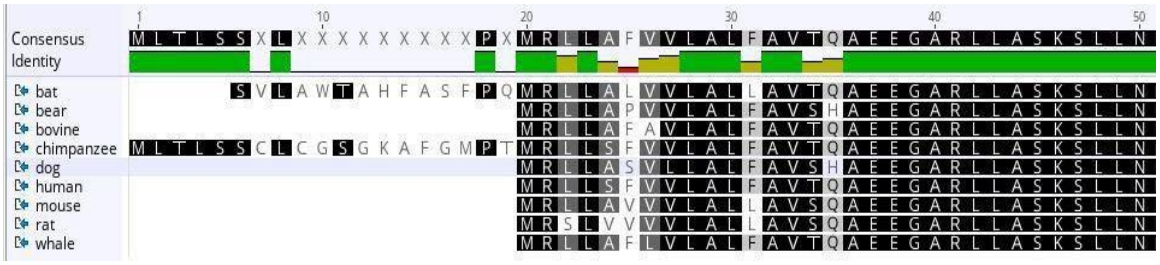


Fig. 1.8.4 – Structure of the ER membrane-associated ribosome determined by cryo-ET: Sec61 (blue), TRAP (yellow) and OST (red). The TRAP complex is under the Sec61 channel and close to the OST complex in the luminal side. In the cytosolic side, TRAP is close to the ribosome (60S) and precisely the ribosomal protein rpL38 (green).

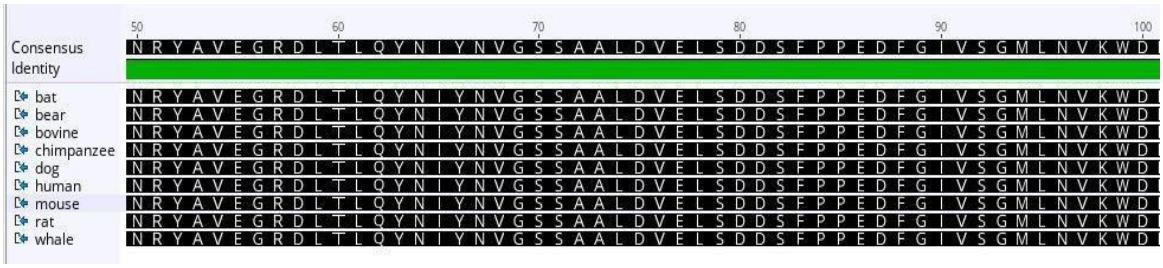
The human TRAP α/β/γ/δ isoforms are very conserved, and the alignment of the most common isoform between different mammalian organisms displays a high identity. The alignments of the most common isoforms of ssr1, ssr2, ssr3, and ssr4 of different species are shown in tables 1.8.10–1.8.13. The mammalian species are bat, bear, bovine, chimpanzee, dog, human, mouse, rat, and whale.



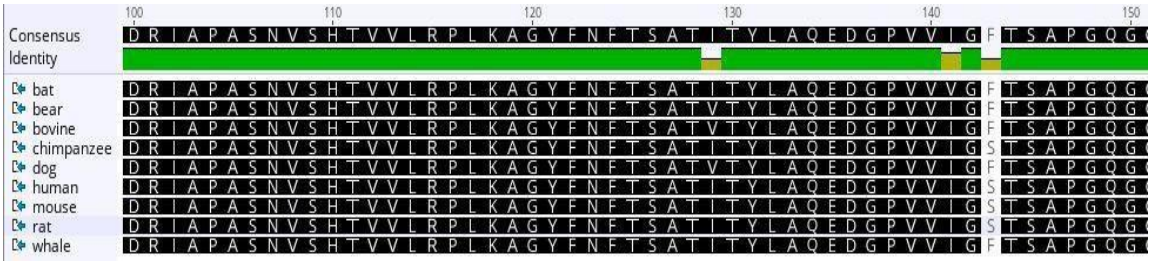
Tab. 1.8.10 - The alignment of the ssr1, most common isoform, between different mammalian species. A = 1-50 residues; B = 50-100; C = 100-150 residues; D= 150-200; E= 200-250; F= 250-291.



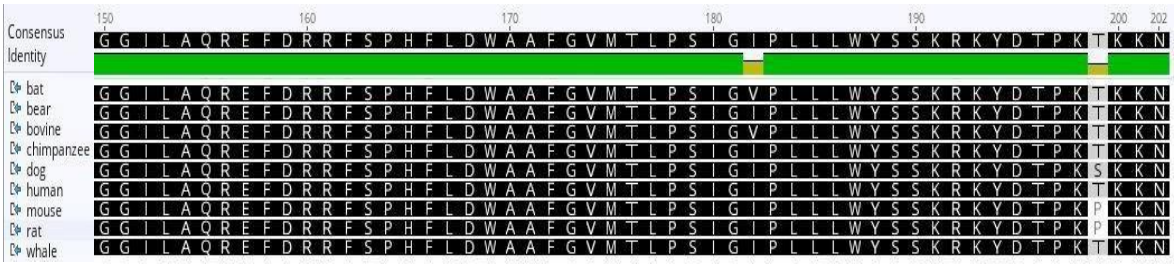
A



B

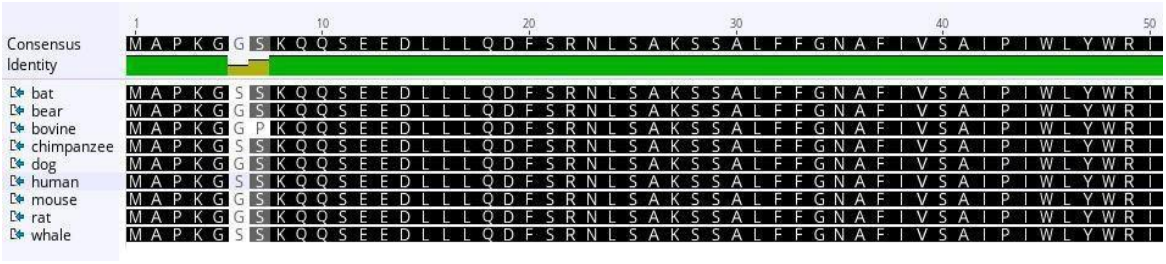


C

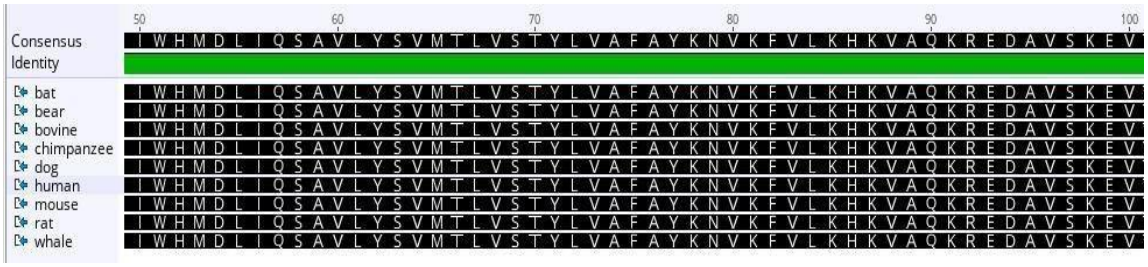


D

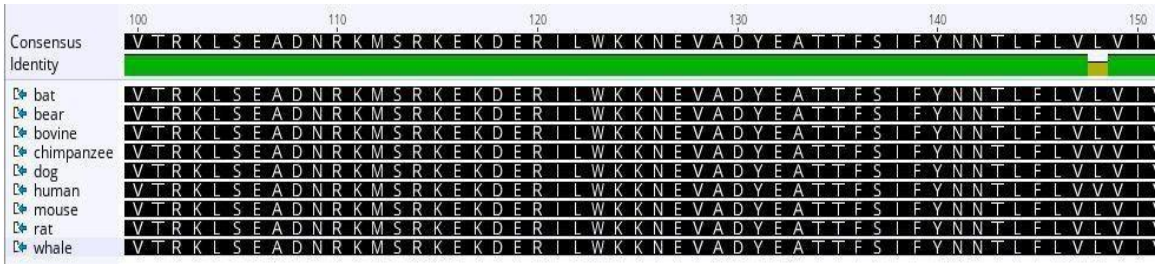
Tab. 1.8.11 – The alignment of the ssr2, most common isoform, between different mammalian species. A = 1-50 residues; B =50 -100 residues; C= 100-150; D= 150-202.



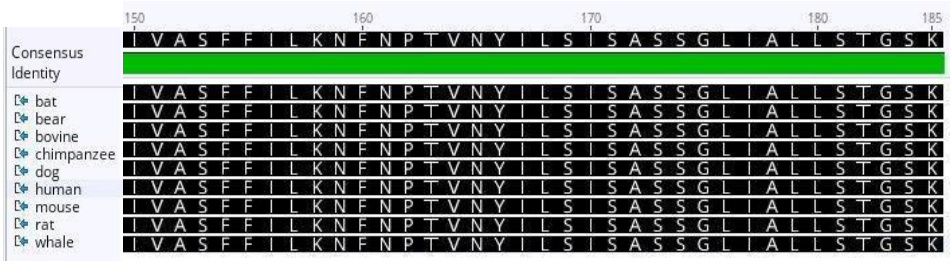
A



B

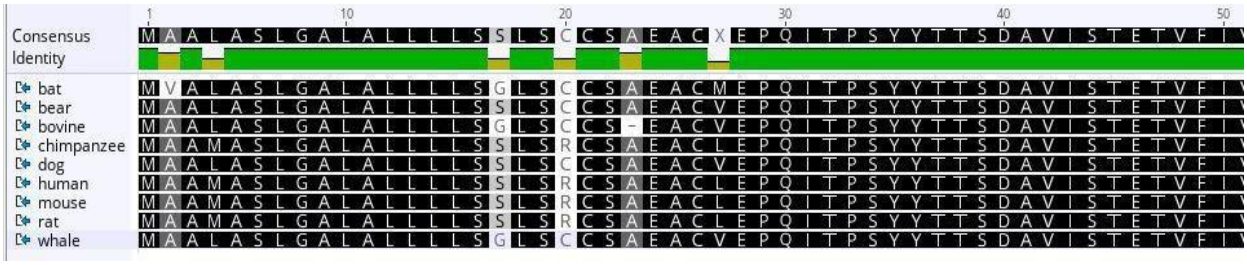


C

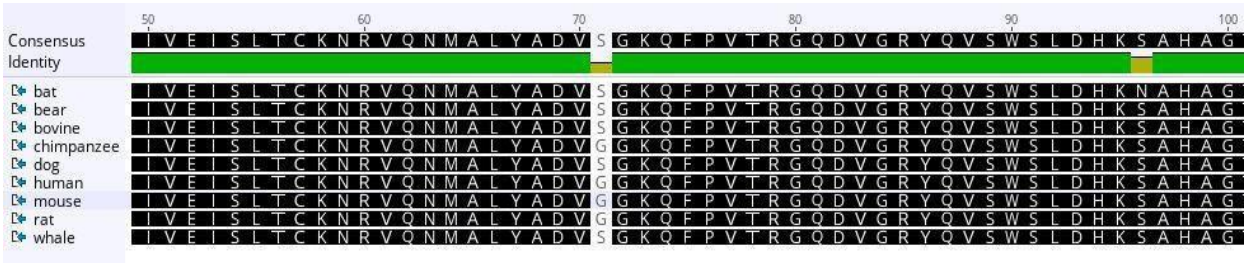


D

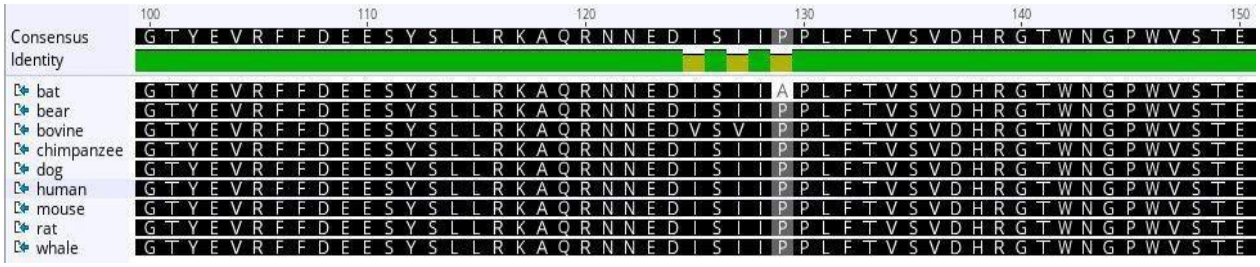
Tab. 1.8.12 – The alignment of the ssr3, most common isoform, between different mammalian species. A = 1-50 residues; B = 50-100 residues; C= 100-150; D= 150-185.



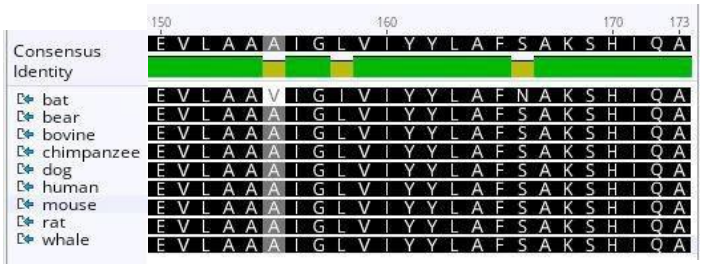
A



B



C



D

Tab. 1.8.13 – The alignment of the ssr4, most common isoform, between different mammalian species. A = 1-50 residues; B = 50-100 residues; C= 100-150; D= 150-173.

Calnexin (90 kDa) is a membrane protein type I, like TRAPα, and both likely bind calcium in the ER lumen (Wada et al., 1991). Ssr1 has a non-canonical EF domain at the N-terminus (Results, p.73). Remarkably, calnexin is also a component of the ribosome-translocon complex and, like TRAP, is close to the translocation polypeptide. Calnexin captures some substrates that acquire N-linked glycans. The palmitoylation of calnexin by DHHC6 permits the interaction with TRAP α. The palmitoylation also recruits the actin cytoskeleton needed for RTC stabilisation (Lakkaraju et al., 2012) (Fig. 1.8.5). Similar to the TRAP subunits, calnexin is involved in the ERAD pathways and the cnx-/- cells have active UPR for acute stress (Coe et al., 2008). In addition, calnexin plays a role in protein folding (Schrag et al., 2001).

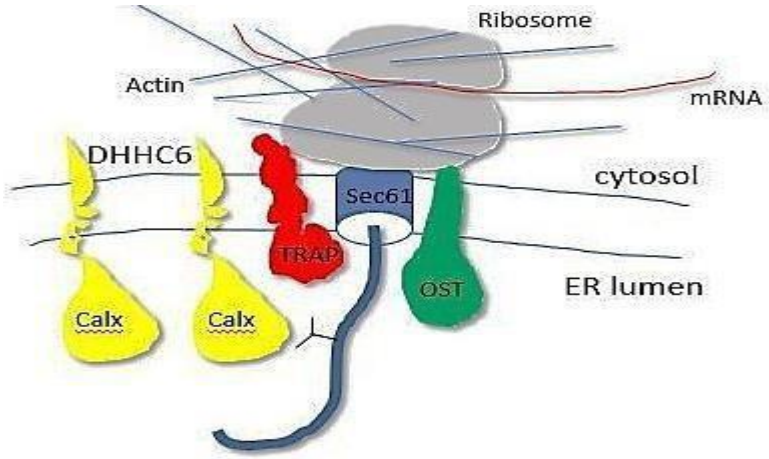


Fig. 1.8.5 – Calnexin like ssr1 is a ribosome translocon complex (RTC) component; both close to the translocating polypeptide. The interaction of Calnexin with TRAPα depends on Calx palmitoylation by DHHC6.

1.9 OST and post-translocation modification

Approximately 90% of the secretory and membrane proteins are N-glycosylated. Glycosylation is the most common protein modification in eukaryotes (Dumax-Vorzet et al., 2013; Bai et al., 2018) and directly affects protein folding in a positive manner (Wang et al., 2008). OST, a multimeric complex of about 200 kDa, catalyses the N-glycosylation into the ER lumen. The complex is part of the RTC, near to Sec61, ribosome 80S subunit, and TRAP complex (Pfeffer et al., 2014; Chawan et al., 2005) (Fig.1.8.8). Similar to other enzymes, an active site allocates the substrate, and the pre-assembled oligosaccharide mannose (glycan) is transferred from the carrier dolichol pyrophosphate to the amino nitrogen of selected protein Asn residues; a sequon Asn-XXX-Ser/Thr or Asn-XXX-Cys, where XXX is any residue except Pro (Fig. 1.91.). The removal of the terminal N-acetylglucosamine from the N-glycan by ER glucosidases I and II permits the calnexin/calreticulin and BiP systems in carrying out protein folding. The compromised biosynthesis of the oligosaccharide substrates leads to CDG, and the TRAP δ subunit is involved in one of these forms.

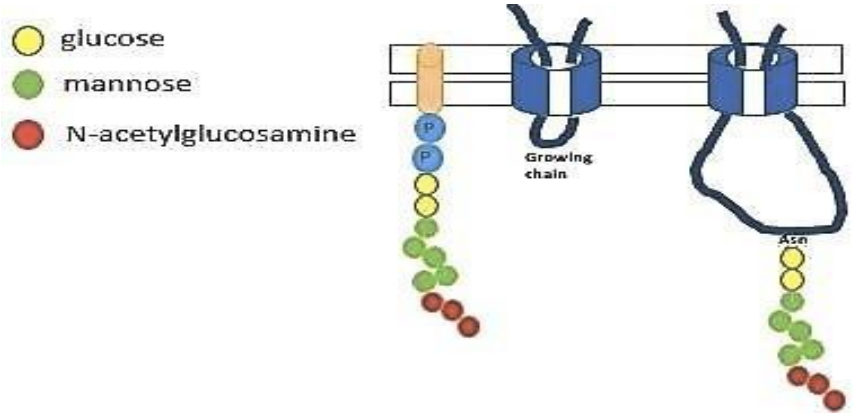


Fig.1.9.1 - N-glycosylation of proteins into the ER lumen: glycan is added to the Asn residues of the nascent protein by the OST complex.

1.10 Objectives

The primary purpose of this study was to contribute to the understanding of the TRAP complex function during ER co-translational translocation. The cryo-EM/ET methods addressed the structure of the ribosome-translocon complex; the luminal domains of TRAP α/ß subunits are close to Sec61α1, and the subunit γ is next to the ribosomal protein rpL38 on the cytosolic side. We cloned these TRAP domains, and the GST and HIS tagged proteins were expressed in bacteria and purified by affinity chromatography. The recombinant proteins were used to carry out experiments that require antibody detection or interaction with glutathione sepharose/Ni-NTA agarose beads. We aimed to address the interaction of TRAP α/ß with the translocon, and hence, we employed a peptide array, which permits the analysis of PPI by using a specific sequence of a protein, in this case, the loop 5 of Sec61α1. Prior to this, we carried out pull-down assays to determine if the two subunits, TRAP α and ß, interact with one another as expected for elements that form a complex. The TRAP complex that interacts with the translocon assists with the stabilisation of the open state. Then, the interactions are transient and established for some substrates, as not all substrates are TRAP-dependent. Nonetheless, we cannot exclude that the interactions between these two complexes (TRAP and Sec61) are stable and require support. Another hypothesis is that TRAP interacts with some precursor polypeptides as demonstrated through crosslinking experiments. The interactions occur when the nascent protein has a length inside the ER lumen of more than 100 residues. It is plausible, and has previously been hypothesised, that TRAP can recognise the mature protein rather than the SP, which would explain its protuberant ER luminal domain under the Sec61 channel. Moreover, we investigated the expression and domains of TRAP subunits, such as the isoforms, calcium-binding domains, and TMDs. As well as the SP properties (hydrophobicity, polarity, and structure), the mature protein features are relevant during translocation. By employing computational methods, we analysed the SPs of some classes of proteins.

2. MATERIAL AND METHODS

2.1 Experimental Methods

2.1.1 TRAP alpha, beta, gamma, and delta (most common isoforms)

Trap-alpha (ssr1)
Uniprot-Q9C50 (*Mus musculus*)
Protein Length: 286
Transmembrane protein Type I
Signal peptide (SP): 21 residues (underline)
Negatively charged luminal N-terminus, Positively charged cytosolic C-terminus
GRAVY value: -0.358

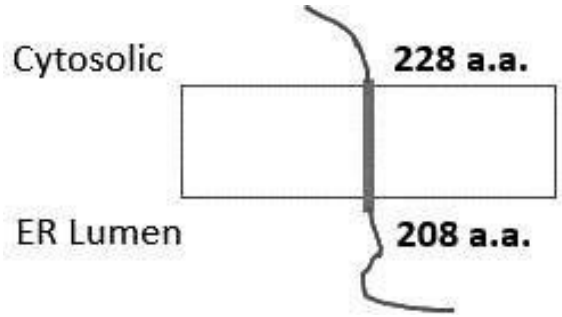
- Nucleotide sequence

ATGAGGCTGCTGCCCAGGCTGCTGCTGCTGTTCCTGCTGGCCTTCCCCGCCGCCGTGCTGCTGAGGG
GCGGCCCCGGCGGCAGCCTGGCCCTGGCCCAGGACCCCACCGAGGACGAGGAGATCGTGGAG
GACAGCATCATCGAGGACGAGGACGACGAGGCCGAGGTGGAGGAGGACGAGCCCACCGACC
TGGCCGAGGACAAGGAGGAGGAGGACGTGAGCAGCGAGCCCGAGGCCAGCCCCAGCGCCGA
CACCACCATCCTGTTCGTGAAGGGCGAGGACTTCCCCGCCAACAACATCGTGAAGTTCCTGGT
GGGCTTCACCAACAAGGGCACCGAGGACTTCATCGTGGAGAGCCTGGACGCCAGCTTCAGGT
ACCCCCAGGACTACCAGTTCTACATCCAGAACTTCACCGCCCTGCCCCTGAACACCGTGGTGC
CCCCCCAGAGGCAGGCCACCTTCGAGTACAGCTTCATCCCCGCCGAGCCCATGGGCGGCAGG
CCCTTCGGCCTGGTGATCAACCTGAACTACAAGGACCTGAACGGCAACGTGTTCCAGGACGCC
GTGTTCAACCAGACCGTGACCGTGATCGAGAGGGAGGACGGCCTGGACGGCGAGACCATCTT
CATGTACATGTTCCTGGCCGGCCTGGGCCTGCTGGTGGTGGTGGGCCTGCACCAGCTGCTGGA
GAGCAGGAAGAGGAAGAGGCCCATCCAGAAGGTGGAGATGGGCACCAGCAGCCAGAACGAC
GTGGACATGAGCTGGATCCCCCAGGAGACCCTGAACCAGATCAACAAGGCCAGCCCCAGGAG
GCAGCCCAGGAAGAGGGCCCAGAAGAGGAGCGTGGGCAGCGACGAG %G ~ C content: 65.3

- Protein sequence

MRLLPRLLLLFLLAFPAAVLLRGGPGGSLALAQDPTEDEEIVEDSIIEDEDDEAEVEEDEPTDLAEDKEE
EDVSSEPEASPSADTTILFVKGEDFPANNIVKFLVGFTNKGTEDFIVESLDASFRYPQDYQFYIQNFTALPL
NTVVPPQRQATFEYSFIPAEPMGGRPFGLVINLNYKDLNGNVFQDAVFNQTVTVIEREDGLDGETIFMYMF
LAGLGLLVVVGLHQLLESRKRKRPIQKVEMGTSSQNDVDMSWIPQETLNQINKASPRRQPRKRAQKRSVGS
DE

1-21 Signal Peptide; Gravy value: 2.04
22-207 Luminal
208-228 Transmembrane
229-286 Cytosolic (Fig. 2.1.1.1)



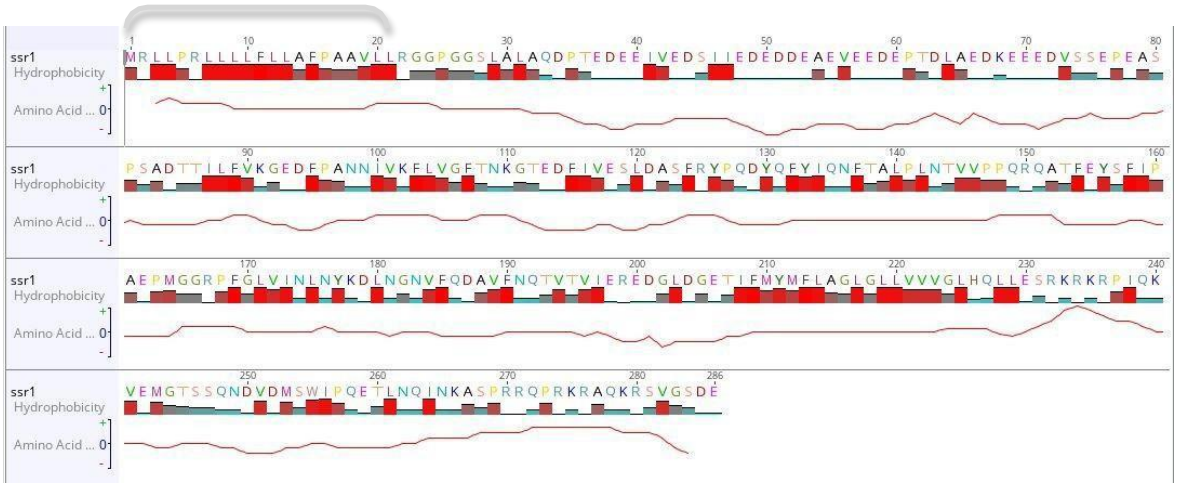


Fig. 2.1.1.1 – *Mus musculus* TRAP α is a TM protein type I with luminal N-terminus. (UniProt). Below: hydrophobicity (red bars) and amino acid charges (red line) of the entire sequence; SP (bracket).

Trap-beta (ssr2)
Uniprot-Q9CPW5 (*Musmusculus*)
Protein Length: 183
Transmembrane protein Type I
SP: 17 residues (underline)
GRAVY value: 0.066

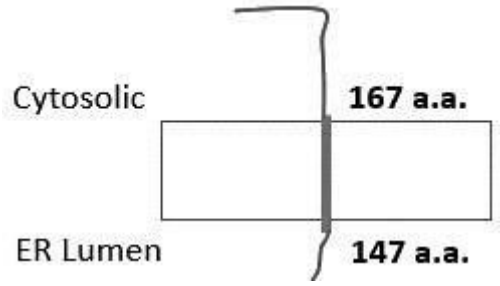
- Nucleotide sequence

ATGAGGCTGCTGGCCGTGGTGGTGCTGGCCCTGCTGGCCGTGAGCCAGGCCGAGGAGGGCGCCAGG
CTGCTGGCCAGCAAGAGCCTGCTGAACAGGTACGCCGTGGAGGGCAGGGACCTGACCCTGCAGTAC
AAATCTACAACGTGGGCAGCAGCGCCGCCCTGGACGTGGAGCTGAGCGACGACAGCTTCCCCCCCG
AGGACTTCGGCATCGTGAGCGGCATGCTGAACGTGAAGTGGGACAGGATCGCCCCCGCCAGCAACG
TGAGCCACACCGTGGTGCTGAGGCCCCTGAAGGCCGGCTACTTCAACTTCACCAGCGCCACCATCAC
CTACCTGGCCCAGGAGGACGGCCCCGTGGTGATCGGCAGCACCAGCGCCCCCGGCCAGGGCGGCAT
CCTGGCCCAGAGGGAGTTCGACAGGAGGTTCAGCCCCCACTTCCTGGACTGGGCCGCCTTCGGCGTG
ATGACCCTGCCCAGCATCGGCATCCCCCTGCTGCTGTGGTACAGCAGCAAGAGGAAGTACGACACC
CCCAAGCCCAAGAAGAAC %G ~ C content: 66.8

- Protein sequence

MRLLAVVVLALLAVSQAEEGARLLASKSLLNRYAVEGRDLTLQYNIYNVGSSAALDVELSDDSFPPEDFGIV
SGMLNVKWDRIAPASNVSHTVVLRPLKAGYFNFTSATITYLAQEDGPVVIGSTSAPGQGGILAQREFDRRFSPH
FLDWAAFGVMTLPSIGIPLLLWYSSKRKYDTPKPKKN

1-17 Signal Peptide; Gravy value: 2.12
18-146 Luminal
147-167 Transmembrane
168-183 Cytosolic (Fig. 2.1.1.2)



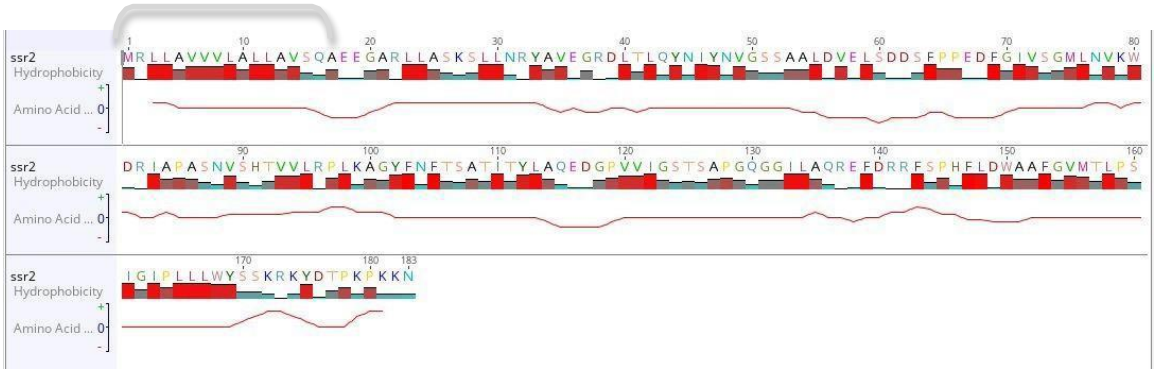


Fig. 2.1.1.2 – *Mus musculus* TRAP ß is a TM protein type I, luminal N-terminus (UniProt). Below: hydrophobicity (red bars) and amino acid charges (red line) of the entire sequence; SP (bracket).

Trap-gamma (ssr3)
Uniprot-Q9DCF9 (*Musmusculus*)
Protein Length:185
Multi-spanning TM protein with cytosolic N- and C-terminus
GRAVY value: 0.066

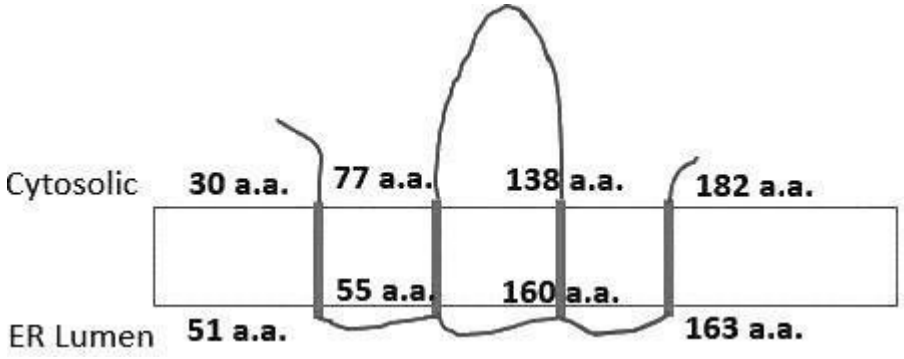
- Nucleotide sequence

ATGGCCCCCAAGGGCGGCAGCAAGCAGCAGAGCGAGGAGGACCTGCTGCTGCAGGACTTCAGCAGGA
ACCTGAGCGCCAAGAGCAGCGCCCTGTTCTTCGGCAACGCCTTCATCGTGAGCGCCATCCCCATC
TGGCTGTACTGGAGGATCTGGCACATGGACCTGATCCAGAGCGCCGTGCTGTACAGCGTGATGAC
CCTGGTGAGCACCTACCTGGTGGCCTTCGCCTACAAGAACGTGAAGTTCGTGCTGAAGCACAAGGTGG
CCCAGAAGAGGGAGGACGCCGTGAGCAAGGAGGTGACCAGGAAGCTGAGCGAGGCCGACAACA
GGAAGATGAGCAGGAAGGAGAAGGACGAGAGGATCCTGTGGAAGAAGAACGAGGTGGCCGACT
ACGAGGCCACCACCTTCAGCATCTTCTACAACAACACCCTGTTCCTGGTGCTGGTGATCGTGGCC
AGCTTCTTCATCCTGAAGAACTTCAACCCCACCGTGAACTACATCCTGAGCATCAGCGCCAGCAG
CGGCCTGATCGCCCTGCTG AGCACCGGCAGCAAG %G~C content: 59.8

- Protein sequence

MAPKGGSKQQSEEDLLLQDFSRNLSAKSSALFFGNAFIVSAIPIWLYWRIWHMDLIQSAVLYSVMTLVSTYLVAF
AYKNVKFVLKHKVAQKREDAVSKEVTRKLSEADNRKMSRKEKDERILWKKNEVADYEATTFSIFYNNTLFLV
LVIVASFFILKNFNPTVNYILSISASSGLIALLSTGSK

1-29 Cytosolic
30-51 Transmembrane
52-54 Luminal
55-77 Transmembrane
78-137 Cytosolic
138-160 Transmembrane
161-162 Luminal
163-182 Transmembrane
183-185 Cytosolic (Fig.2.1.1.3)



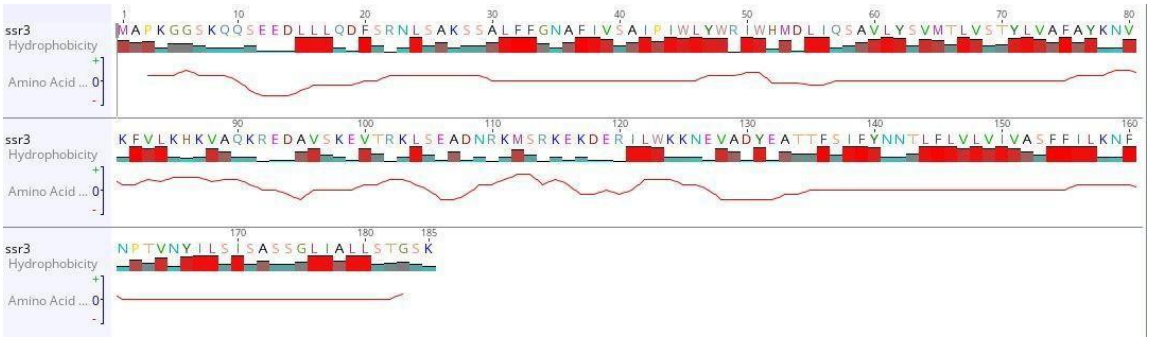


Fig. 2.1.1.3 – *Mus musculus* TRAP γ is a multi-spanning TM protein with cytosolic Nter and Cter, and a prominent cytosolic domain (Bano-Polo et al., 2017). Below: hydrophobicity (red bars) and amino acid charges (red line) of the entire sequence.

Trap-delta (ssr4)
Uniprot- Q62186 (*Mus musculus*)
Protein Length: 173
Transmembrane protein Type I
SP: 23 residues (underline)
GRAVY value: 0.099

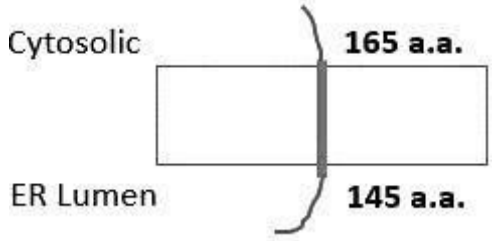
- Nucleotide sequence

ATGGCCGCCATGGCCAGCCTGGGCGCCCTGGCCCTGCTGCTGCTGAGCAGCCTGAGCAGGT
GCAGCGCCGAGGCCTGCCTGGAGCCCCAGATCACCCCCAGCTACTACACCACCAGCGACGC
CGTGATCAGCACCGAGACCGTGTTCATCGTGGAGATCAGCCTGACCTGCAAGAACAGGGTGCAGAA
CATGGCCCTGTACGCCGACGTGGGCGGCAAGCAGTTCCCCGTGACCAGGGGCCAGGACGTGGGCAG
GTACCAGGTGAGCTGGAGCCTGGACCACAAGAGCGCCCACGCCGGCACCTACGAGGTGAGGTT
CTTCGACGAGGAGAGCTACAGCCTGCTGAGGAAGGCCCAGAGGAACAACGAGGACATCAGCATCAT
CCCCCCCCTGTTCACCGTGAGCGTGGACCACAGGGGCACCTGGAACGGCCCCTGGGTGAGCACCG
AGGTGCTGGCCGCCGCCATCGGCCTGGTGATCTACTACCTGGCCTTCAGCGCCAAGAGCCACAT
CCAGGCC %G ~ C content: 66.1

- Protein sequence

MAAMASLGALALLLLSSLSRCSAEACLEPQITPSYYTTSDAVISTETVFIVEISLTCKNRVQNMALYADV
GGKQFPVTRGQDVGRYQVSWSLDHKSAHAGTYEVRFFDEESYSLLRKAQRNNEDISIIPPLFTVSVDHRG TW
NGPWVSTEVLAAAIGLVIYYLAFSAKSHIQA

1-23 Signal peptide; Gravy value: 1.51
24-144 Luminal
145-165 Transmembrane
166-173 Cytosolic (Fig. 2.1.1.4)



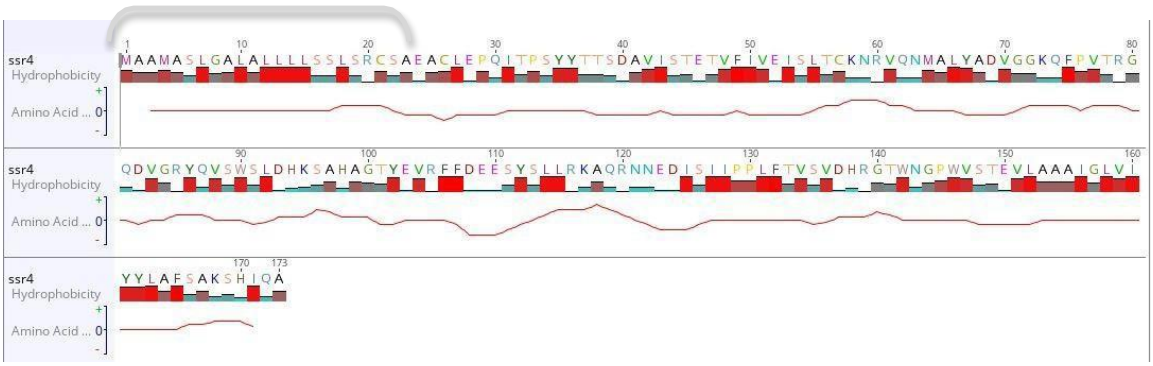


Fig. 2.1.1.4 – *Mus musculus* TRAP δ is a TM protein type I, luminal N-terminus (UniProt). Below: hydrophobicity (red bars) and amino charges (red line) of the entire sequence; SP (bracket).

2.1.2 Molecular cloning

Domains of TRAP complex subunits (*Mus musculus*) were cloned in pEX-N-GST, pEX-C- GST, pGEX- C-GST, and pGEX-C-HIS vectors. The GST tag in N- and C-terminus is 26 kDa from the parasitic helminth Schistosoma japonicum; this tag can increase protein solubility by avoiding inclusion bodies, and permits a natural cleavage. The HIS tag is just six amino acid residues (6 His) which avoids interference with the structure/function of the recombinant protein and provides high yield during purification.

The GST tag is the following:

MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSM
AIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFEDRLCH
KTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQ
ATFGGGDHPPKSDLVPRGSPEFPGR LERPHRD (26 KDa).

The HIS tag is 6 x HIS (Fig. 2.1.2.1)

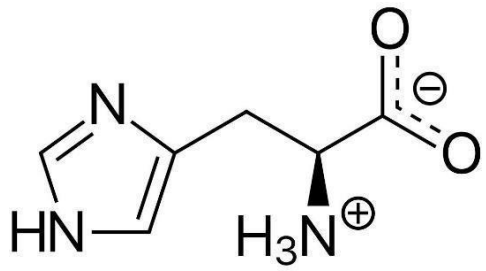


Fig. 2.1.2.1 – Histidine amino acid.

The cloned domains of the TRAP subunits are the following:
Trap-alpha - from 22 to 205 amino acid residues (183-mer), luminal domain;
Trap-beta - from 18 to 147 residues (129-mer), luminal domain;
Trap-gamma - from 78 to 134 residues (56-mer), cytosolic domain;
Trap-delta - from 24 to 144 residues (120-mer), luminal domain.

Trap α domain, 22 to 205 residues/183-mer, was inserted by Origene Biotechnology Company into two different plasmids with Ct GST and Ct HIS tag (Fig. 2.1.2.2):
pEX-C-GST (5.3 kb), Ct GST, TEV cleavage site, ampicillin-resistant;
pEX-C-HIS (4.6 kb), Ct HIS, TEV cleavage site, ampicillin resistant.

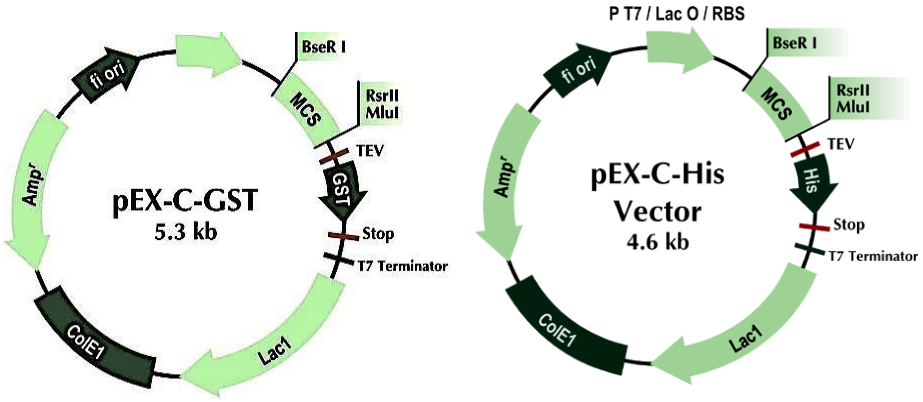


Fig. 2.1.2.2 – Vectors for cloning of TRAP alpha domain: one with Cterm GST tag and one with Cterm HIS tag. The former long 5.3 kb and the latter 4.6 kb (Origene).

Previously, we inserted the domains of TRAP β/γ/δ in pGEX vectors by taking into account the reading frame, orientation, size, and end compatibility. The pGEX vectors, like pEX vectors, present a multiple cloning site (MCS), a tag, and the tac promoter which is induced by the lactose analogue isopropyl β-D thiogalactoside (IPTG).

TRAP beta domain, 18-147 residues/130-mer, was inserted in pGEX-TEV-GST (pJDE) plasmid: XbaI (blunted)-EcoRI hisG-URA3-hisG fragment from pUC19 inserted into MunI (blunted)- and EcoRI- digested pKC8, Nt GST, TEV cleavage site, ampicillin-resistant, 5356 bp (Fig. 2.1.2.3) Trap-delta domain, 24-144 residues/120-mer inserted in pGEX-TEV-GST (pJDE) plasmid: XbaI (blunted)-EcoRI hisG-URA3-hisG fragment from pUC19 inserted into MunI (blunted)- and EcoRI- digested pKC8, Nt GST, TEV cleavage site, ampicillin-resistant, 5356 bp (Fig. 2.1.2.3).

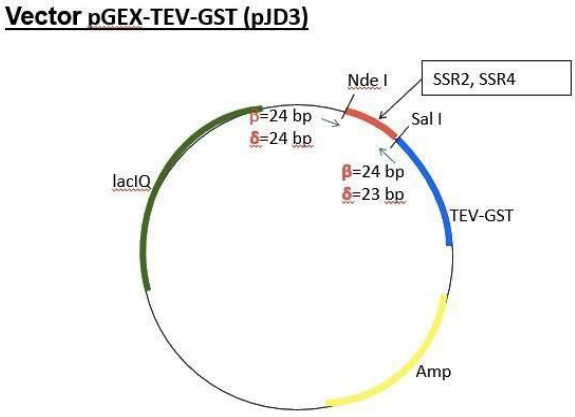


Fig. 2.1.2.3 – In the vector pGEX-TEV- GST (pJD3) were inserted: luminal TRAP β and δ domains; the GST tag is at the Cterm.

Trap-gamma domain, 78-134 residues/56-mer, was inserted in pGEX-4T-TEV (pGS804) vector, which derives from pGEX-4T-1, it contains a TEV cleavage site, Nt GST tag, and ampicillin-resistant gene, 5155 bp (Fig. 2.1.2.4).

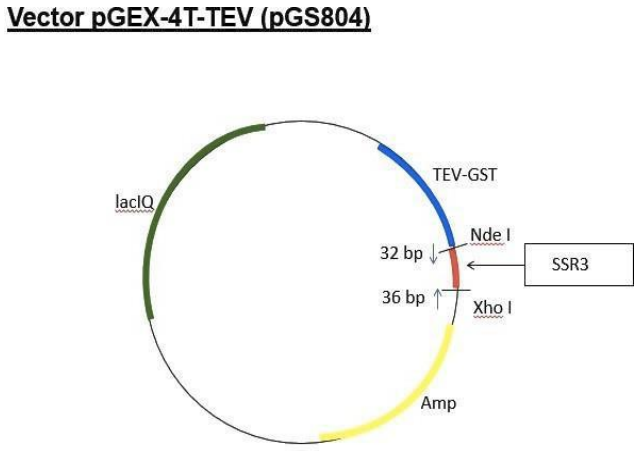


Fig. 2.1.2.4 –In the vector pGES- 4T-TEV (pGS804) was inserted the cytosolic TRAP γ domain; the GST tag is at the Nterm.

2.1.3 Quantitative PCR

The cDNA of *Mus musculus* TRAP subunit domains, inserted in CMV6 plasmids (1µg/µL), were used as templates for qPCR. The qPCR was performed by using Pfu DNA Polymerase (Thermo Scientific) with the following protocol (Tab. 2.1.3.1).

Water, nuclease-free (#R0581)	variable
10X *Pfu* Buffer with $MgSO_4$*	5 µL
dNTP Mix, 2 mM each (#R0241)	5 µL (0.2 mM of each)
Forward primer	0.1-1.0 µM
Reverse primer	0.1-1.0 µM
Template DNA	50 pg - 1 µg
***Pfu* DNA Polymerase**	1.25-2.5 U
Total volume	50 µL

Tab. 2.1.3.1 – qPCR protocol to amplify the TRAP β, γ, δ subunits.

Initial denaturation 95 °C/30 sec, denaturation 95 degrees/30 sec, annealing 53 °C/30 sec, extension 72°C/45 sec, and final extension 72 °C/2 min. A master mix that included all components were prepared and then pipetted in PCR tubes (50 µl). The primers were for TRAP β - RZ623 and RZ624, for TRAP γ - RZ617 and RZ618, for TRAP δ - RZ619 and RZ620. The primers (Eurofins, 100 pmol/µl), presented a 5' extension as a restriction enzyme site, which was not included in the Tm calculation (Tab. 2.1.3.2). After the PCR reaction, each sample was loaded on 1% agarose gel. The amplified products were purified with PCR Purification Kit (Qiagen); then, the enzymatic digestions were performed.

Primer	Sequence
RZ623 (24 nt) forward	cagctgggtggtccgcatctcttg
RZ624 (24 nt) reverse	catgccgtctcgtcacccaggctc
RZ617 (32 nt) forward	atctcgagtcatgtagcttcataatcagcaac
RZ618 (38 nt) reverse	cgccatatgaagaatgtgaaatttgttctcaagcac
RZ619 (24 nt) forward	gtacatatggaggcctgcctggag
RZ620 (23 nt) reverse	gtagtcgaccccgttccaagtgc

Tab. 2.1.3.2 - The primers (forward and reverse) used for qPCR of TRAP beta, gamma, and delta.

2.1.4 Digestion with restriction enzymes

The plasmid pGEX-TEV-GST (pJD3) and TRAP β/ TRAP δ PCR products were digested with NdeI and SalI restriction enzymes (Thermo Scientific) with the following protocol:

3 µl 10x Buffer
0.5-1 µl DNA
1.5 µl NdeI
0.5 µl SalI
up to 16 µl free-nuclease water
Incubate at 37 °C for 2-3 hours

The plasmid pGEX-TEV-GST (pGS804) and TRAP γ PCR product were digested with NdeI and XhoI (Thermo Scientific) with the following protocol:

µl 10x Buffer O
0.5-1 µl DNA
1.5 µl NdeI
0.5 µl XhoI
up to 16 µl free-nuclease water
Incubate at 37 °C for 2-3 hours

The plasmids were dephosphorylated with Shrimp Alkaline Phosphatase (rSAP, Biolabs): 2 µl of phosphatase, 37 degrees, 1 hour. After digestion and dephosphorylation, plasmids and inserts were purified by gel extraction, with 0.8% agarose gel. The gel was run at 200 V for ~ 40 minutes, then the bands carefully cut and centrifuged for 15 minutes at 14,000 rpm (Eppendorf, 5415C) in filter 2ml tubes (Castar, Spin-x Centrifuge Tube Filter). Natriumacet pH 5.2 (1/10 volume) and 96% Analytical Ethanol (2 and ½ volume) were added to the tubes and were frozen at 80 °C for 30 minutes. The tubes were centrifuged for 10 minutes at 4 °C, the supernatant was removed, and the pellet was washed, by centrifugation, with 70% ethanol (150 µl) (2-5 minutes at 14,000 rpm). The pellet after air-drying was resuspended in 10 µl of TE buffer pH 8.

2.1.5 Ligation

The ligation of the inserts (TRAP domains) in the chosen vectors (listed above) was performed with T4 DNA Ligase (Thermo Scientific) at ratio 1:3 or 1:5 (plasmid: insert). The protocol was the following:

20-100 ng plasmid
60-500 ng insert
3µl 10x T4 DNA Ligase Buffer
up to 20 µl free-nuclease water
Incubate at 22 °C for 1 hour

The *E.coli* strains used for transformation were JM101 and DH5α, heat shock was carried out at 42 ˚C for 45 sec. The transformed cells were rolled in the incubator at 37 ˚C for about 1 hour, after adding 500 µL of LB Miller Medium (Fisher BiOReagents). The cultures were spread on the agar plates (Agar- Agar, Biosciences, ROTH) and incubated overnight at 37 ˚C. The first screening was ampicillin resistance, the viable colonies were grown in 5ml of Terrific Broth (TB) Medium (ROTH) overnight, and the next day miniprep/plasmid isolation (small-scale) was performed.

2.1.6 Plasmid purification: miniprep and MIDI

The cells were starved by centrifugation for 15 minutes at 2500 rpm (Beckman GS-6KR Centrifuge), the pellet was resuspended in 100 µl of GTE, then 200 µl of 0.2 M NaOH with 1% SDS was added, and the tubes were inverted 6-10 times. The solution was neutralized with 150µl of 3M NaoAC, pH 5.2, 400 µl of phenol/chloroform/isoamyl alcohol, the tubes were vortexed for 30 sec, then centrifuged for 2 minutes at 16,000 rpm (Eppendorf, 5415 C). The upper phase was transferred in new Eppendorf tubes with 800 µl of 96% ethanol, the tubes were inverted several times, and left for one minute at room temperature. The suspension was centrifuged at 16,000 rpm for 2 minutes, and the supernatant was removed, the pellet was washed with 70% ethanol and air-dried. To digest the RNA present, 50 µl of RNase diluted in TE buffer were added, the tubes were incubated at 55 ˚C for 5 minutes, and at 37 ˚C for 30 minutes. The tubes were stored at -20 ˚C, or the enzymatic digestion was performed (above enzymes and EcoR1) to confirm the presence and direction of the insert. When the cloning was successful, the constructs were grown in JM101, or DH5 alpha strains with 100 ml of LB medium and big-scale purification of the plasmids was performed by MIDI Kit (Qiagen). The samples were sent for DNA sequencing with the appropriate primers to confirm the correct cloning.

2.1.7 Sequencing

The plasmids were sent at the concentration of 50-100 ng/µL, up to 15 µL with free-nuclease water in Eppendorf tubes; 3 µl of primers (Eurofins) at the concentration of 100pmol/µl were dissolved in 12 µl of free nuclease water and sent with the samples (EurofinsGenomics).

TRAPβ was sequenced with the primer RZ623 (forward) and RZ624 (reverse).
TRAPγ was sequenced with a yeast primer – 518 – (WWG-Biotech AG) (forward, 19-mer) – gctggcaagccacgtttgc.
TRAPδ was sequenced with the primer RZ619 (forward) and primer RZ620 (reverse).

2.1.8 Protein Expression and Purification

2.1.8.1 Bacterium strains

We used for cloning and plasmid maintenance the E.coli JM101 and DH5α strains and BL21 (DE3)/BL21 RosettaStar strains for protein expression and purification.

JM101

glnV44 thi-1 Δ(lac-proAB) F'[lacI^{q}ZΔM15 traD36 proAB$^+$]

DH5α

F$^-$ *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20* φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, hsdR17($r_K^-m_K^+$), λ$^-$

BL21(DE3)

E. coli str. B F$^-$ *ompT gal dcm lon hsdS$_B$*($r_B^-m_B^-$) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K-12}$(λ^S)

BL21 RosettaStar: mutatedRnaseE,thedegradationisreduced,mRNAismorestablethan inBL21(DE3)strain.

The transformation was carried out by heat-shock (42 ° C, 45/30 sec), the bacteria were grown in LB medium with the appropriate antibiotics (1:1000), ampicillin (100mg/ml) and chloramphenicol (32mg/ml). The induction was performed with 0.4 mM IPTG (72 mg/L), at 37 °C for 3 hours. After starvation, the cells were lysed, for BL21 (DE3) by freezing (Nitrogen liquid) and thawing. Instead, for BL21 RosettaStar sonication.was performed, three times for 30 sec with an interval of 60 sec in between (Sonics & Materials Inc, VibraCell). The expression of the proteins was checked before at small-scale, by "affinity chromatography batch method", then the proteins were purified by "affinity chromatography column method", by using Glutathione Sepharose 4 Fast Flow beads (GE Healthcare) for GST tagged proteins and Ni-NTA Agarose beads (Qiagen) for HIS tagged proteins.

2.1.8.2 Affinity chromatography batch method

By using *E.coli* as a host, it was possible to analyse a small amount of overnight culture by SDS-PAGE and check for the expression of the target protein. However, a more precise method was "Batch Affinity Chromatography", to the lysate from a few ml of overnight culture are added the Glutathione- Sepharose beads; the tagged proteins bind to the ligand, then washing remove all impurity, and finally, the proteins can be solved by denaturing SDS-PAGE gel after staining with Coomassie brilliant blue. The intensity of bands is proportional to the amount of protein expressed. Precisely, the transformed cells were grown in 20 ml LB medium/antibiotics overnight, with sufficient aeration, no more than 20% of the total flask volume. The next day the culture was split in four flasks with the final volume of 20 ml in LB medium, ampicillin and chloramphenicol, the cells were grown until OD600 reached 0.8 (Pharmacia Biotech, Ultrospec 300, UV/Visible Spectrophotometer). The IPTG induction (0.4 mM) was performed for 3 hours, the cells were starved lysate centrifugation, freezing, and thawing (or sonication) (10 µl + 2x Lämmli); d) the supernatant

after lysate centrifugation (10 µl + 2x Lämmli). The entire sequence of samples was loaded on the SDS-PAGE (15%): -IPTG (a), +IPTG (b), Insoluble fraction (c), Soluble Fraction (d), Eluate (e).

2.8.1.3 Affinity chromatography column method

The cell lysate, from 1liter of overnight culture, is slowly transferred in a column (20 ml, disposable chromatography columns, Bio-Rad) where previously 4 ml of GSH-Sepharose beads (or Ni-NTA Agarose beads) were poured. The column was washed to eliminated no-binding proteins and impurity, and finally, by using the reducing buffer (32mg L-Glutathione red, 0.5ml 1 M Tris pH 8, up to 10 ml H2O), the tagged proteins were eluted (Fig. 2.1.8.3.1).
Precisely, the transformed cells were grown overnight in 20 ml of LB medium with the antibiotics (1:1000), the next day the culture was diluted to 1 liter with LB medium, ampicillin and chloramphenicol (1ml/liter), and the cells were grown until OD600 reached 0.8; then, IPTG induction (0.4 mM) was performed for 3 hours, at 37 °C. Before adding the IPTG a sample was taken (500 µl) and spin down for 30 sec, the pellet was resuspended with 100µl Lämmli buffer and boiled 10 min at 95°C; for the +IPTG sample, based on the OD measure (table) some µl were taken and treated like the - IPTG sample. The cells were starved, 6000 rpm, 10 min, 4 °C (J2-M Centrifuge, Beckman) and lysed by PBS-KMT (100 ml 10x PBS, 1.5ml 2M KCl, 1 ml 1M MgCl2,1 ml TWEEN 20 up to 1 liter with filtered H2O), the Protease Inhibitor Cocktail (PLAC) and Dithiothreitol (DDT) were added. The lysate was centrifuged, 50,000 for 30 min, 4°C (Optima L-80 Ultracentrifuge, Beckman) to remove non- solubilized material; then, the supernatant was stored at -80 ° C. The GSH-Sepharose beads (or Ni-NTA Agarose) were washed three times with PBS- KMT by centrifugation (1500rpm, 4°C). The thawed lysate was rolled with 4 ml of beads for 1 hour at 4°C and poured in the chromatography column. The flow-through sample was collected (10 µl) and 2x Lämmli buffer (10 µL) was added. The column was washed until the OD was 0.1 or below (Nanodrop ND-1000 Spectrophotometer). Afterwards, the elution buffer was added to the column, and the eluates were collected in eight Eppendorf tubes (1ml each) (Fig. 2.1.8.3.1). The samples collected, -IPTG, +IPTG, IF, SF, 8 x E were loaded on SDS-PAGE (15%). The gel was run at 200 V for about 40 minutes in 5X running buffer (75 gr. TRIS, 360 gr.Glycin, 25 gr. SDS up to 5 litres with H2O). Then, it was stained with Coomassie blue (02% Coomassie R250, 0005% Coomassie G250 up to 1 liter with the destained solution I), destained with solution 1 (800ml acetic acid, 3200ml Methanol, 3840 ml H2O, 160 ml 87% Glycerin), and solution 2 (400ml acetic acid, 800 ml Methanol, 6620 ml H2O, 180 ml 87% Glycerin); finally, it was scanned (Image III, GE Healthcare).
1.5ml 2M KCl, 1 ml 1M MgCl2,1 ml TWEEN up to 1 liter with filtered dist. H2O) with Protease Inhibitor Cocktail (PLAC) and Dithiothreitol (DDT). The lysate was frozen, thawed (or sonicated), and centrifuged at 14,000 rpm, 30 min, at 4 °C (Centrifuge 5415R Eppendorf); 100 µl of the supernatant was rolled with 30 µl of beads for 1 hour, at 4 °C. Finally, the Eppendorf tubes were centrifuged, the supernatant removed, to the GSG-Sepharose beads were added 40 µl of 2x Lämmli buffer, and the sample was boiled for 5 minutes (sample e). Other samples were collected: a) before induction (10 µl + 10 µl 2x Lämmli); b) after induction (10 µl + 10 2x Lämmli); c) the pellet, after

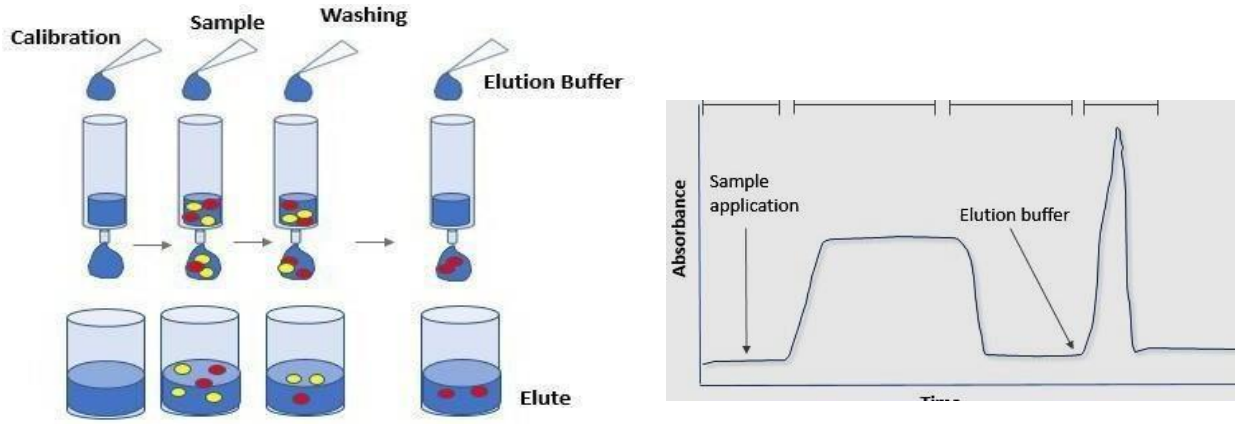


Fig. 2.1.8.3.1 – The workflow for the "Column Affinity Chromatography Method": calibration with lysate buffer, sample loading, washing with lysate buffer, and elution. The graph shows the absorbance (protein concentration) in each workflow step.

2.1.8.4 GST purification

The transformed BL21 cells were inoculated into 50ml LB medium with ampicillin and chloramphenicol (1:1000) and incubated overnight at 37°C. The next day the culture was diluted to 1 liter with LB medium containing antibiotics, and the cells were grown until the OD600 reached 0.8. Then, to the 1L culture was added IPTG (0.4 mM) and it was incubated for 2 hours at 37°C; the cells were starved by centrifugation (10 min 5000 rpm, 4°C, JA10 Beckman), the pellet was resuspended with PBS-KMT, PLAC (20µl/20ml) and DDT (20µl/20ml) were added. The samples were quick-frozen in Nitrogen liquid for 15 min and thawed on the ice. Then, the falcons were spin down in Ti70 Beckman rotor (30 min, 5000rpm, 4°C), the supernatant was frozen at -80°C until utilization. The GSH-Sepharose beads (Sepharose 4 Fast Flow, GE Healthcare) were washed three times with PBS-KMT by centrifugation (1500rpm, 4°C), the supernatant was incubated (rolling) with the beads, 1 hour at 4°C. The solution was poured in the chromatography column; then, the column was washed with 50ml of PBS-KMT until the OD280 reached 0.1 or less. Finally, the elution buffer was poured, and for each collected sample (1ml) 10 µl were taken for the SDS- Page, and 40µl of Lämmli were added to. The elution buffer for all purified proteins was replaced with PBS/KMT or "Peptide Array" buffer by using Zeba Spin Desalting Columns (Thermo Scientific), 2mL, 5mL or 10ml based on the volume of the purified proteins. The Zeba D. S. columns contain a high- performance resin which permits buffer-exchange, useful also for desalting or to remove small molecules. First, the Zeba Spin D.S. columns were centrifuged to remove the storage solution (2 min, 1000 xg), then were washed three times with the new buffer. Finally, the protein with the elution buffer was poured in the column, centrifuged and the sample was collected. The proteins were stored, small aliquots (50 µl) in thin-walled PCR plastic tubes at -80 ° until utilization. In some cases, the proteins were concentrated with Centrifugal Filter Devices (Amicon Centricon, Millipore Corporation), a process that also led to a purer protein, by eliminating, for instance, GST degradation or other impurities with different molecular weight. These results were achieved when the right molecular weight cut-off (MWCO) was chosen; the molecules with low molecular weight (solvent) by pressure (centrifugation) went through the membrane and were eliminated. With these devices, we were able to concentrate more than 10 mL of purified protein solution to 0.5 mL in about one hour.

2.1.8.5 Solubility, concentration, and purity

100 µL of each purified protein was centrifuged for 20 min., at 68,000 rpm and 4°C (Beckman Coulter, Optima Max-E Ultracentrifuge, TLA 120.2). The samples before and after centrifugation were loaded on 15% SDS PAGE gel. The concentration and purity of purified proteins were measured by UV quantification –Vis Nanodrop Spectrophotometer ND1000 (peqLab, Biotechnologie GmbH), optical density (OD) protein at 280 nm. The absorbance is directly proportional to the concentration of the solution, as is explained by Beer-Lambert's law:

$A \propto c$ $A = absorbance$

$c = concentration$

$A = log10 = (I0/I) = L$ $I0 = incident\ intensity$

$I = transmitted intensity$

$A \propto L$

$A \propto c\ L$

$A \propto ecL$ $e = molar\ absorptivity\ costant$

1-2 µL of the sample was released on the nanodrop instrument pedestal, and in a few seconds, the concentration value and the purity (ratio 260:280) were displayed. This method is quantitative, but the purity and the concentration can also be estimated by looking to the SDS-PAGE (qualitative method).

2.1.8.6 SDS-PAGE

A method to analyse the proteins by electrophoresis is the Sodium-Dodecyl-Sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) which discriminates small difference on migration based on the molecular weight. By boiling the samples, the proteins are denatured and by adding SDS, they acquire negative charges. They migrate only in one direction (anode) when the electric field is applied (smaller proteins move faster). The samples are loaded into each well, the migration forms lanes under the well where one band(s), which represents the protein(s), is visible after staining. The proteins are determined based on the molecular weight by using the appropriate marker (Prestained Protein Ladder) also loaded on the gel.

<u>SDS-PAGE 15% Protocol</u>:
Separating Gel (16 ml) (samples run through)
ml ddH2O, 6 ml 40% Acrylamide, 4 ml 1.5M Tris pH 8.8, 160 µl 10%SDS, 160 µl 10%APS, 16µl TEMED

Stacking Gel (10 ml) (well formation and sample loading)
ml ddH2O, 1.5 ml 40% Acrylamide, 2.5 ml 0.5 M Tris pH 6.8, 100 µl 10%SDS, 100 µl APS, 10 µl TEMED

2.1.8.7 Western blotting

First, electrophoresis with SDS-PAGE (200 V, ~ 40 minutes) was performed, then the proteins were transferred on PVDF membrane (Immobilon-P, Merck Millipore) (Towbin et al., 1979) by running the transfer-sandwich in an electrophoresis chamber containing 10x Loading Buffer (30 minutes, 200 V). For GST detection, the membrane was blocked in 5% milk powder solution (ROTH) for 1 hour and incubated for 90 minutes with the primary antibody - rabbit GST (7.2 µg/µl, Lab Collection) at a 1:1000 dilution in 1x PBS and 0.01% Triton. After washing, the membranes were incubated overnight with the secondary antibody - ECL PLEX Goat-α-rabbit IgG, Cy5, (1 µg/µl, GE Healthcare) and then washed. The bound antibody was detected after peroxidase reaction with Vilber Lourmat FUSION SL (Peqlab Biotechnologie). For HIS detection, the membrane was incubated for 90 minutes with primary antibody - sc-804-G, rabbit anti-HIS antibody (100µg/ml, Santa Cruz) at a 1:1000 dilution in 1x PBS and 0.01% Triton. After washing, the membrane was incubated overnight with the secondary antibody - A8275/ rabbit anti-IgG coupled with peroxidase (1:2500 dilution in TBS, 0.5% BSA, SigmaAldrich) at a dilution 1:1000 and then washed. The membrane was scanned with Typhoon TRIO, Variable Mode Imager (GE Healthcare).

2.1.8.8 Liquid Chromatography-Mass Spectrometry (LC-MS)

Proteins were loaded on NuPAGE 10% which was run at 200 V for ~ 30 minutes, incubated for 30 minutes at room temperature in fixation solution (40% ethanol, 10% Acetic acid), washed, and stained with colloidal Coomassie (0.12% Coomassie G250 dye, 10% Ammonium sulfate, 10% Phosphoric acid, 20% Methanol). The gel was destained with dist. H_2O and soaking in water to reach the pH neutrality, the bands were cut out by maximising the ratio of protein to gel, placed into Eppendorf tubes and stored at - 20 °C.

Subsequently, C. Fecher-Trost (General and Clinical Pharmacology Department, UKS, Homburg) performed the LC-MS analysis.

The samples were incubated with 15 µl of trypsin (porcine, 20 ng/µl, Promega) at 37 °C /overnight. The obtained small peptides were concentrated in a vacuum centrifuge andresuspended in 20 µl of 0.1% formic acid. Then, six µl of the tryptic peptide samples were measured by full scan MS, after collision- induced dissociation CID and higher collisional dissociation HCD with an Orbitrap Velos mass spectrometer (Thermo Fisher Scientific, TOP5 method, gradient 60 min). The fragmented peptides were analysed using the software (PROTEOME DISCOVERER, Thermo Fisher Scientific) applying the reviewed protein database (SwissProt release 2018_02) (Fecher-Trost C. et al., 2013).

2.1.8.9 GST and HIS pull-down assays

First, the proteins were partially purified from GST degradation by centrifugation with Centricon Ultracel YM-10 (Amicon).

The reaction buffers for pull-down assays were the following:

GST-Pull Down: 20 mM Tris-HCl (pH 8), 150 mM NaCl, 0.1% Nonidet P-40,10% glycerol,

Glutathione Sepharose 4 Fast Flow (GE Healthcare) beads;

HIS-Pull Down: 50mM Tris-HCl (pH 7.5), 60 mM NaCl, 10 mM β-mercaptoethanol, 0.01 Nonidet P-40, 10mM imidazole, 10% glycerol, Ni-NTA Agarose (Qiagen) beads.

The beads (30 µl) were added to 12 µg of bait protein (TRAPβ:: GST or TRAPα:: HIS) in 100 µl of pull- down buffer, the Eppendorf tubes were rolled at 4 °C for 1 hour, centrifuged, and thepellet was washed several times by centrifugation with the pull-down buffer. Afterwards, the prey protein (TRAPα:: HIS or TRAP β:: GST) was added to the same buffer (same volume) and rolled for 1 hour at 4 ° C (Fig. 2.1.8.9.1). The next day the sample was centrifuged and washed several times. The washing steps were for both incubations the following: 5 min/2,200 rpm/4 ° C, five min/30 sec/3,200 rpm/4 ° C, 30 sec/10,000rpm/4° C (Eppendorf; centrifuge 5415 R). Finally, the samples were boiled at 95 ° C with 2x Lämmli buffer, and 15 µl were loaded on 15% SDS-Page gel. The same steps were performed for GST negative control, by adding GST protein (bait) and TRAPα:: HIS (prey).

.

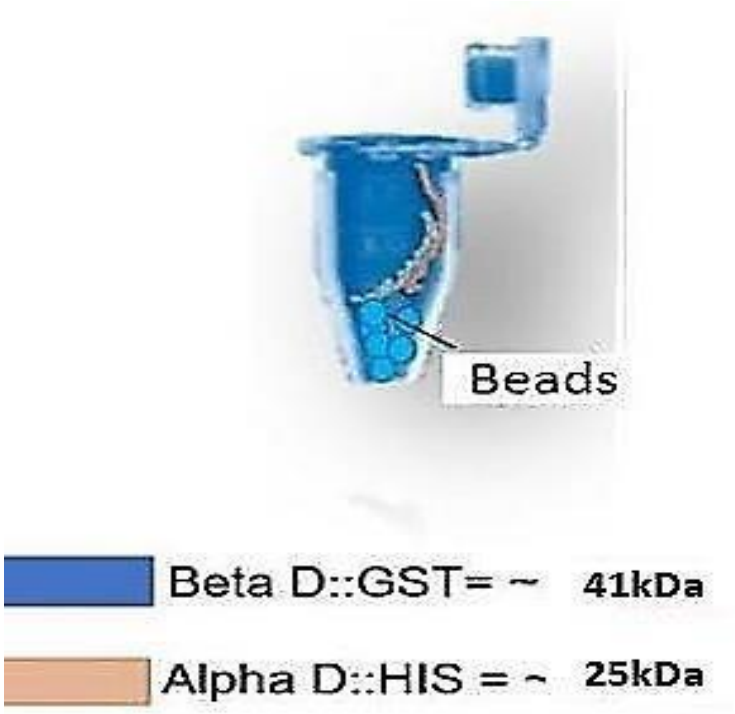


Fig. 2.1.8.9.1 - The second incubation permits the interaction between the bait (blue), already attached to the Glutathione Sepharose 4 Fast beads, and the prey (pink). In the HIS-pull down assay the beads were Ni-NTA Agarose beads, the bait was TRAPα:: HIS and the prey was TRAPβ:: GST.

2.1.8.10 Peptide array

The SPOT technology consists of adding each time an amino acid residue to a growing peptide chain on a cellulose membrane; the synthesizer (INTAVIS, ResPepSL) delivers the reaction-mixture droplet (up to 1
µl) containing the amino acid. When the read-out is via chemiluminescence, the peptides bind via C-term on cellulose membrane by starting with Ala: a) Fmoc-β-Ala-OH, DIC, NMI; b) 20 % piperidine; c) Fmoc-β- X-protection group; d) Ac2 O, DIEA; e) 20% piperidine (X= any residue) (Fig. 2.1.8.10.1). The peptide array is carried out by incubating the cellulose membrane with the partner protein, then with antibody, and finally the signal is detected (Fig. 2.1.8.10.2).

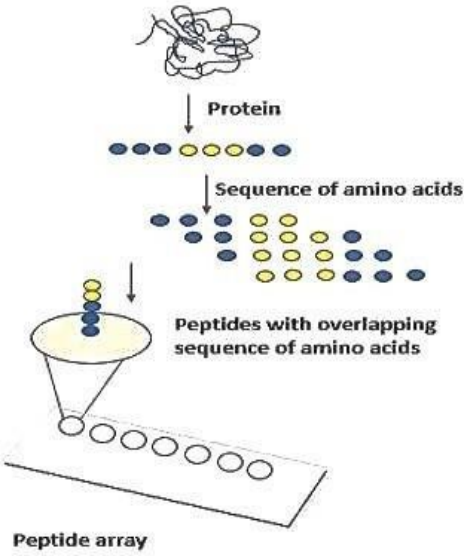


Fig. 2.1.8.10.1 – SPOT technology consists on synthesizing overlapping sequences on cellulose membrane.

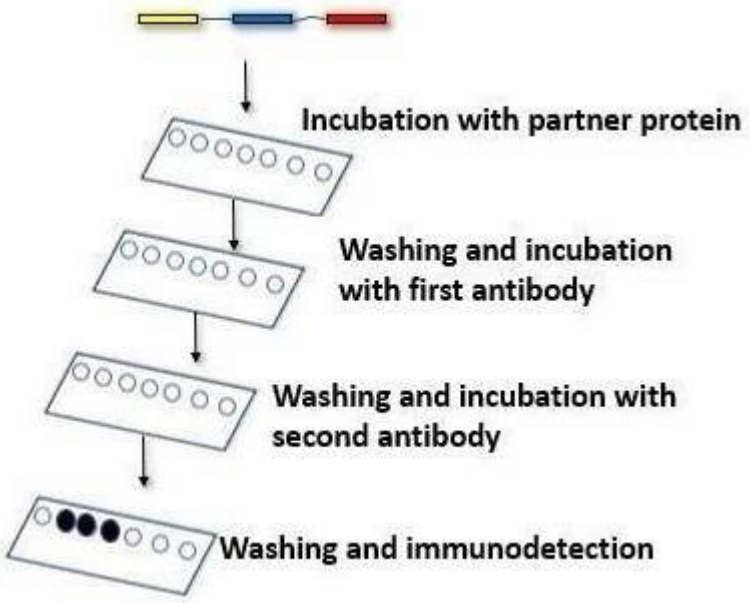


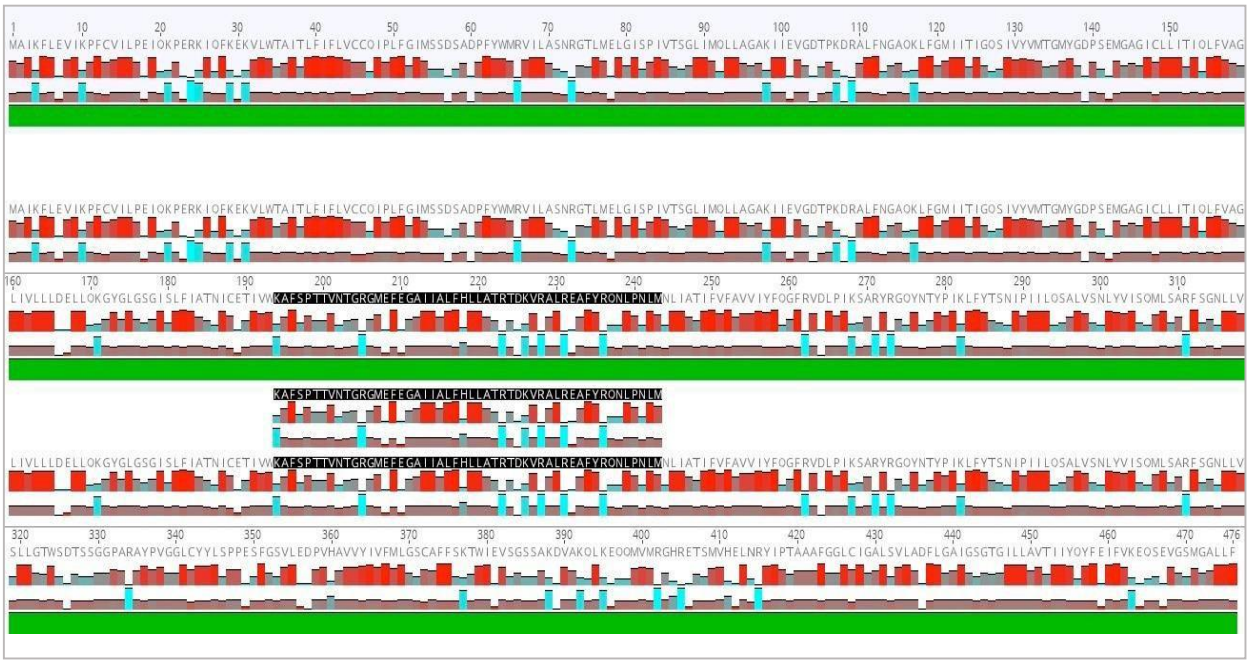
Fig. 2.1.8.10.2 - Workflow for "Peptide Array": incubation with a partner protein, washing, incubation with the first antibody, washing, incubation with the second antibody, and finally immunodetection and analysis.

M. Jung (Medical Biochemistry and Molecular Biology Department, UKS, Homburg) carried out the synthesis of the overlapping spots of the sequence Sec61alpha1 loop 5 on a cellulose membrane. Each spot is a single-peptide 20 amino acid long; in each membrane (5 x 15 cm), there were 30 spots of the wild type sequence loop 5 (upper lane), and 28 spots of mutated sequence loop 5 (lower lane) where an Alanine substituted another residue along the entire sequence (Tab. 2.1.8.10.1- 2.1.8.10.3).

```
        10         20         30         40         50
MAIKFLEVIK PFCVILPEIQ KPERKIQFKE KVLWTAITLF IFLVCCQIPL
        60         70         80         90        100
FGIMSSDSAD PFYWMRVILA SNRGTLMELG ISPIVTSGLI MQLLAGAKII
       110        120        130        140        150
EVGDTPKDRA LFNGAQKLFG MIITIGQSIV YVMTGMYGDP SEMGAGICLL
       160        170        180        190        200
ITIQLFVAGL IVLLLDELLQ KGYGLGSGIS LFIATNICET IVWKAFSPTT
       210        220        230        240        250
VNTGRGMEFE GAIIALFHLL ATRTDKVRAL REAFYRQNLP NLMNLIATIF
       260        270        280        290        300
VFAVVIYFQG FRVDLPIKSA RYRGQYNTYP IKLFYTSNIP IILQSALVSN
       310        320        330        340        350
LYVISQMLSA RFSGNLLVSL LGTWSDTSSG GPARAYPVGG LCYYLSPPES
       360        370        380        390        400
FGSVLEDPVH AVVYIVFMLG SCAFFSKTWI EVSGSSAKDV AKQLKEQQMV
       410        420        430        440        450
MRGHRETSMV HELNRYIPTA AAFGGLCIGA LSVLADFLGA IGSGTGILLA
       460        470
VTIIYQYFEI FVKEQSEVGS MGALLF
```

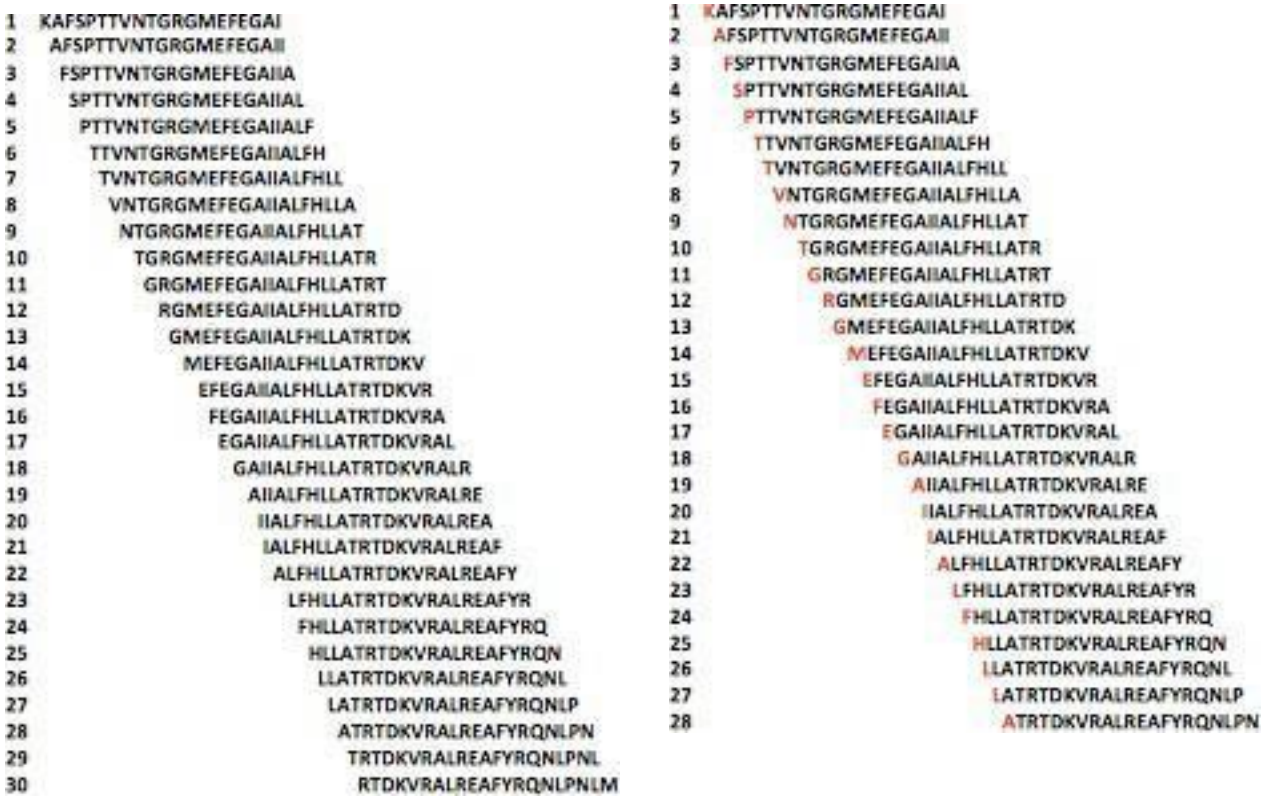
```
194
   KAFSPTT VNTGRGMEFE
GAIIALFHLL ATRTDKVRAL
REAFYRQNLP NLM
            243
```

Tab.2.1.8.10.1 - Sec61α isoform 1 entire sequence (left) and Sec61α1 loop 5 (rigth), sequence from 194 to 243 amino acid residues.



Tab. 2.1.8.10.2 - Alignment of Sec61α1 loop5: loop 5 is 50 residues long, much shorter comparing with the entire sequence of Sec61α1 (476 residues).

The peptide array was performed by activating the membrane with MeOH, which was washed, and equilibrated for two hours with binding buffer (30mM TRIS/HCl, 170mM NaCl, 2 mM CaCl2, 2mM MgCl2, 0.05% Tween 20). The membrane was incubated with the partner protein at the concentration of 5µM: TRAPα:: GST/ TRAPα:: HIS/TRAPβ:: GST, and washed with the binding buffer. Afterwards, the membrane was incubated with the first antibody - rabbit GST (7.2 µg/µl, Lab Collection 1:1000 dilution in 1x PBS, 0.01% Triton), or sc-804-G, rabbit anti-HIS antibody, (100µg/ml, Santa Cruz), and washed with the binding buffer. Then, it was incubated with the second antibody - A8275, rabbit anti- IgG antibody coupled with peroxidase (horseradish) (1:2500 dilution in TBS, 0.5% BSA, SigmaAldrich, dilution 1:1000). Finally, there was detection by chemiluminescence with Typhoon TRIO, Variable Mode Imager (GE Healthcare), after ECL reaction (Pierce) (horseradish peroxidase (HRP) reaction); the results were analysed based on the overlapping peptide sequences (Fig. 2.1.8.10.2).
The membrane can be reused a few times after regeneration by ultrasonic bath (BRANSON 8200): 1) 2x 30 minutes with buffer A (8M Urea/50mM tris HCl/1%SDS/0.5% Mercatoethanol/pH 7 (HCl); 2) 30 minutes with buffer B (50% etOH/10% Acetic acid/40% H2O); 3) rinse with PBS. Finally, chemiluminescent detection test (after ECL reaction) is performed to confirm that the membrane is free of any binding and ready to be stored at -20 ˚C.



Tab. 2.1.8.10.3 - Overlapping sequences of WT loop 5 Sec61α1(left) and mutated loop 5 Sec61α1(right) were synthesized on the cellulose membrane. In each spot, the residues are overlapping except for the first and last amino acid residues. In the mutated version (left), alanine substituted another residue along the sequence (red).

2.2 Computational methods

Geneious: R11- 11.1.2, Copyright 2005-2018, Biomatters Ltd., the software platform was used for organization/analysis of sequence data and domains, search for motifs.

RaptorX: protein structure server predicts 3D structures from protein sequences without close homologs in the Protein Data Bank (PDB).

STRING: a database of known and predicted protein-protein interactions; SIB (Swiss Institute of Bioinformatics), CPR-NNF (Center for Protein Research), EMBL (European Molecular Biology Laboratory). STRING relies also on COG, Ensembl, Intact, RefSeq, PubMed, Reactome, DIP, BioGRID, MINT, KEGG, SGD, FlyBase, SwissProt/UniProt, SwissModel, HUGO, OMIM, NCI/Nature PID, PDB, The Interactive Fly, BioCyc, Gene Ontology, SIMAP, etc.

3. RESULTS

3.1 Experimental results

3.1.1 TRAP alpha, beta, gamma, and delta cloning

ORIGENE cloned TRAPα domains with vectors containing Cterm GST and Cterm HIS tag. The clones were delivered after a long time due to some complications, problems that we also encountered before deciding to order them.

The DNA sequence of TRAPβ, γ and δ subunit domains, inserted in pCMV6 vectors (1µg/µl), were used as templates for qPCR. The qPCR was performed as described in the Methods section, and the DNA amplification of three TRAP subunit domains is visible on agarose gel (1%) (Fig. 3.1.1.1).

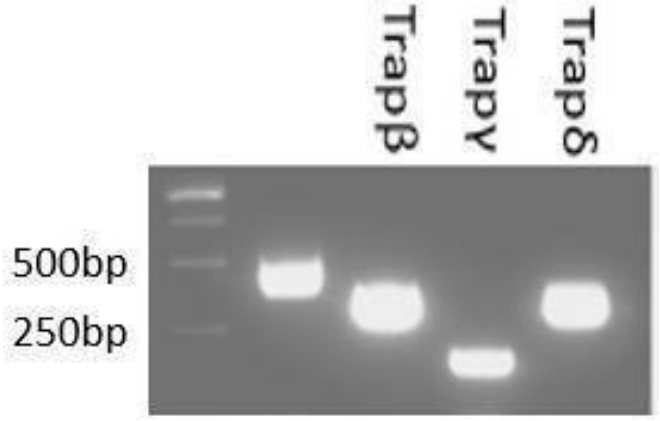


Fig. 3.1.1.1 - Agarose gel (1.2%) shows the amplification by qPCR of TRAP β luminal domain (387 bp), TRAP γ cytosolic domain (168 bp), and TRAP δ luminal domain (360 bp).

The amplified DNA sequences of TRAP β and δ were digested, and then inserted in the C - terminal GST tagged vectors by ligation, as described in the methods section. The amplified sequence of TRAP γ, was digested, and inserted in the N-terminal GST tagged vector. The constructs were used to transform *E.coli* cells, JM101 and DH5α; between the two strains no significant difference was detectable, but for strain availability the plasmids were purified by using the strain JM101. The isolated plasmids (MIDI protocol) were sent for sequencing to "Eurofins Sequencing Company" by using the appropriate primers. The alignments between clone sequence design and DNA sequencing have an exact match, 100% of identity (Tab. 3.1.1.1-3.1.1.3).

```
TRAP BETA sequencing        3 CTCCAAATAATG----AAGATCTCATCACCATCACCATCACTAATGACTGACGATCTGCC   58
                              C C   A A  G     AGATCTCATCACCATCACCATCACTAATGACTGACGATCTGCC
TRAPBETA vector 1-1301    241 CACAGGAAACAGTACATAGATCTCATCACCATCACCATCACTAATGACTGACGATCTGCC  300

TRAP BETA sequencing       59 TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCA  118
                              TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCA
TRAPBETA vector 1-1301    301 TCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCA  360

TRAP BETA sequencing      119 CAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTG  178
                              CAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTG
TRAPBETA vector 1-1301    361 CAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTG  420

TRAP BETA sequencing      179 TTGGCGGGTGTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTATAATT  238
                              TTGGCGGGTGTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTATAATT
TRAPBETA vector 1-1301    421 TTGGCGGGTGTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTATAATT  480

TRAP BETA sequencing      239 CTTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAA  298
                              CTTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAA
TRAPBETA vector 1-1301    481 CTTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAA  540

TRAP BETA sequencing      299 TGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTT  358
                              TGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTT
TRAPBETA vector 1-1301    541 TGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTT  600

TRAP BETA sequencing      359 TATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGC  418
                              TATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGC
TRAPBETA vector 1-1301    601 TATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGC  660

TRAP BETA sequencing      419 TTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTC  478
                              TTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTC
TRAPBETA vector 1-1301    661 TTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTC  720

TRAP BETA sequencing      479 CCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAA  538
                              CCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAA
TRAPBETA vector 1-1301    721 CCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAA  780

TRAP BETA sequencing      539 AAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCG  598
                              AAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCG
TRAPBETA vector 1-1301    781 AAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCG  840

TRAP BETA sequencing      599 GTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAG  658
                              GTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAG
TRAPBETA vector 1-1301    841 GTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAG  900

TRAP BETA sequencing      659 TTCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTCGCC  718
                              TTCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTCGCC
TRAPBETA vector 1-1301    901 TTCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTCGCC  960

TRAP BETA sequencing      719 GCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTA  778
                              GCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTA
TRAPBETA vector 1-1301    961 GCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTA 1020

TRAP BETA sequencing      779 CGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTG  838
                              CGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTG
TRAPBETA vector 1-1301   1021 CGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTG 1080

TRAP BETA sequencing      839 CGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACA  898
                              CGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACA
TRAPBETA vector 1-1301   1081 CGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACA 1140

TRAP BETA sequencing      899 ACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATAC  958
                              ACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATAC
TRAPBETA vector 1-1301   1141 ACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATAC 1200

TRAP BETA sequencing      959 CAAACGACGAGCGTGACACCACGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACTA- 1017
                              CAAACGACGAGCGTGACACCACGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACTA
TRAPBETA vector 1-1301   1201 CAAACGACGAGCGTGACACCACGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACTAT 1260
```

Tab. 3.1.1.1 - Alignment between clone design TRAP β and DNA sequencing (pGEX-TEV-GST - pJDE-vector).

```
TRAP GAMMA Sequ              70 ATATGAAGAATGTGAAATTTGTTCTCAAGCACAAAGTAGCACAGAAGAGGGAGGATGCTG  129
                                ATATGAAGAATGTGAAATTTGTTCTCAAGCACAAAGTAGCACAGAAGAGGGAGGATGCTG
TRAP gamma vector 1-1320     961 ATATGAAGAATGTGAAATTTGTTCTCAAGCACAAAGTAGCACAGAAGAGGGAGGATGCTG 1020

TRAP GAMMA Sequ             130 TTTCCAAAGAAGTGACTCGAAAACTTTCTGAAGCTGATAATAGAAAGATGTCTCGGAAGG  189
                                TTTCCAAAGAAGTGACTCGAAAACTTTCTGAAGCTGATAATAGAAAGATGTCTCGGAAGG
TRAP gamma vector 1-1320    1021 TTTCCAAAGAAGTGACTCGAAAACTTTCTGAAGCTGATAATAGAAAGATGTCTCGGAAGG 1080

TRAP GAMMA Sequ             190 AGAAAGATGAAAGAATCTTGTGGAAGAAGAATGAAGTTGCTGATTATGAAGCTACATGAC  249
                                AGAAAGATGAAAGAATCTTGTGGAAGAAGAATGAAGTTGCTGATTATGAAGCTACATGAC
TRAP gamma vector 1-1320    1081 AGAAAGATGAAAGAATCTTGTGGAAGAAGAATGAAGTTGCTGATTATGAAGCTACATGAC 1140

TRAP GAMMA Sequ             250 TCGAGCGGCCGCATCGTGACTGACTGACGATCTGCCTCGCGCGTTTCGGTGATGACGGTG  309
                                TCGAGCGGCCGCATCGTGACTGACTGACGATCTGCCTCGCGCGTTTCGGTGATGACGGTG
TRAP gamma vector 1-1320    1141 TCGAGCGGCCGCATCGTGACTGACTGACGATCTGCCTCGCGCGTTTCGGTGATGACGGTG 1200

TRAP GAMMA Sequ             310 AAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCG  369
                                AAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCG
TRAP gamma vector 1-1320    1201 AAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCG 1260

TRAP GAMMA Sequ             370 GGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCA  429
                                GGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCA
TRAP gamma vector 1-1320    1261 GGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCA 1320
```

Tab. 3.1.1.2 - Alignment between clone design TRAP γ and DNA sequencing (pGEX-4T- TEV- pGS804-vector).

```
TRAP Delta Sequ        1 ------------------------------------AAATTTCACACAGGAAACAGTACAT   25
                                                              AATTTCACACAGGAAACAGTACAT
TRAP Delta vector    181 TGTTTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACAGTACAT  240

TRAP Delta Sequ       26 ATGGAGGCCTGCCTGGAGCCCCAGATCACCCCTTCCTACTACACCACTTCTGACGCTGTC   85
                         ATGGAGGCCTGCCTGGAGCCCCAGATCACCCCTTCCTACTACACCACTTCTGACGCTGTC
TRAP Delta vector    241 ATGGAGGCCTGCCTGGAGCCCCAGATCACCCCTTCCTACTACACCACTTCTGACGCTGTC  300

TRAP Delta Sequ       86 ATTTCCACTGAGACCGTCTTCATTGTGGAGATCTCCCTGACATGCAAGAACAGGGTCCAG  145
                         ATTTCCACTGAGACCGTCTTCATTGTGGAGATCTCCCTGACATGCAAGAACAGGGTCCAG
TRAP Delta vector    301 ATTTCCACTGAGACCGTCTTCATTGTGGAGATCTCCCTGACATGCAAGAACAGGGTCCAG  360

TRAP Delta Sequ      146 AACATGGCTCTCTATGCTGACGTCGGTGGAAAACAATTCCCTGTCACTCGAGGCCAGGAT  205
                         AACATGGCTCTCTATGCTGACGTCGGTGGAAAACAATTCCCTGTCACTCGAGGCCAGGAT
TRAP Delta vector    361 AACATGGCTCTCTATGCTGACGTCGGTGGAAAACAATTCCCTGTCACTCGAGGCCAGGAT  420

TRAP Delta Sequ      206 GTGGGGCGTTATCAGGTGTCCTGGAGCCTGGACCACAAGAGCGCCCACGCAGGCACCTAT  265
                         GTGGGGCGTTATCAGGTGTCCTGGAGCCTGGACCACAAGAGCGCCCACGCAGGCACCTAT
TRAP Delta vector    421 GTGGGGCGTTATCAGGTGTCCTGGAGCCTGGACCACAAGAGCGCCCACGCAGGCACCTAT  480

TRAP Delta Sequ      266 GAGGTTAGATTCTTCGACGAGGAGTCCTACAGCCTCCTCAGGAAGGCTCAGAGGAATAAC  325
                         GAGGTTAGATTCTTCGACGAGGAGTCCTACAGCCTCCTCAGGAAGGCTCAGAGGAATAAC
TRAP Delta vector    481 GAGGTTAGATTCTTCGACGAGGAGTCCTACAGCCTCCTCAGGAAGGCTCAGAGGAATAAC  540

TRAP Delta Sequ      326 GAGGACATTTCCATCATCCCGCCTCTGTTTACAGTCAGCGTGGACCATCGGGGCACTTGG  385
                         GAGGACATTTCCATCATCCCGCCTCTGTTTACAGTCAGCGTGGACCATCGGGGCACTTGG
TRAP Delta vector    541 GAGGACATTTCCATCATCCCGCCTCTGTTTACAGTCAGCGTGGACCATCGGGGCACTTGG  600

TRAP Delta Sequ      386 AACGGGGTCGACGAAAACCTGTATTTTCAGGGCTCCCCTATACTAGGTTATTGGAAAATT  445
                         AACGGGGTCGACGAAAACCTGTATTTTCAGGGCTCCCCTATACTAGGTTATTGGAAAATT
TRAP Delta vector    601 AACGGGGTCGACGAAAACCTGTATTTTCAGGGCTCCCCTATACTAGGTTATTGGAAAATT  660

TRAP Delta Sequ      446 AAGGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATATCTTGAAGAAAAATATGAAGAG  505
                         AAGGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATATCTTGAAGAAAAATATGAAGAG
TRAP Delta vector    661 AAGGGCCTTGTGCAACCCACTCGACTTCTTTTGGAATATCTTGAAGAAAAATATGAAGAG  720

TRAP Delta Sequ      506 CATTTGTATGAGCGCGATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATTGGGTTTG  565
                         CATTTGTATGAGCGCGATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATTGGGTTTG
TRAP Delta vector    721 CATTTGTATGAGCGCGATGAAGGTGATAAATGGCGAAACAAAAAGTTTGAATTGGGTTTG  780

TRAP Delta Sequ      566 GAGTTTCCCAATCTTCCTTATTATATTGATGGTGATGTTAAATTAACACAGTCTATGGCC  625
                         GAGTTTCCCAATCTTCCTTATTATATTGATGGTGATGTTAAATTAACACAGTCTATGGCC
TRAP Delta vector    781 GAGTTTCCCAATCTTCCTTATTATATTGATGGTGATGTTAAATTAACACAGTCTATGGCC  840

TRAP Delta Sequ      626 ATCATACGTTATATAGCTGACAAGCACAACATGTTGGGTGGTTGTCCAAAAGAGCGTGCA  685
                         ATCATACGTTATATAGCTGACAAGCACAACATGTTGGGTGGTTGTCCAAAAGAGCGTGCA
TRAP Delta vector    841 ATCATACGTTATATAGCTGACAAGCACAACATGTTGGGTGGTTGTCCAAAAGAGCGTGCA  900
```

Tab. 3.1.1.3 – Alignment between clone design TRAP δ and DNA sequencing (pGEX-TEV- GST – pJDE- vector).

E.coli "BL21 (DE3)" and "BL21 RosettaStar" strains were transformed, for protein expression, with the positive clones (plasmids) confirmed by sequencing. The BL21(DE3) cells were appropriate for all TRAP constructs - β:: GST, GST::γ, GST::δ - except for TRAP α (GST and HIS tagged). Nevertheless, the use of BL21 RosettaStar cells increased the yield also for the other subunits. Eventually, the chosen cells were BL21 RosettaStar for all recombinant proteins.

3.1.2 Recombinant tagged proteins

Initially, the cells "BL21 (DE3)" were used for expression of TRAP α- GST and HIS tagged recombinant proteins, but a short version of the proteins was detected. I decided to use the strain "BL21RosettaStar" which derived from BL21 (DE3) strain; it carries a mutation on the rne gene (rne131) that encodes a truncated RNase E that reduces mRNA degradation. The right size of proteins with high yield was achieved for TRAPα:: GST and TRAP α::HIS. The results are shown by the following SDS-PAGEs (Fig.3.1.2.1- 3.1.2.2).

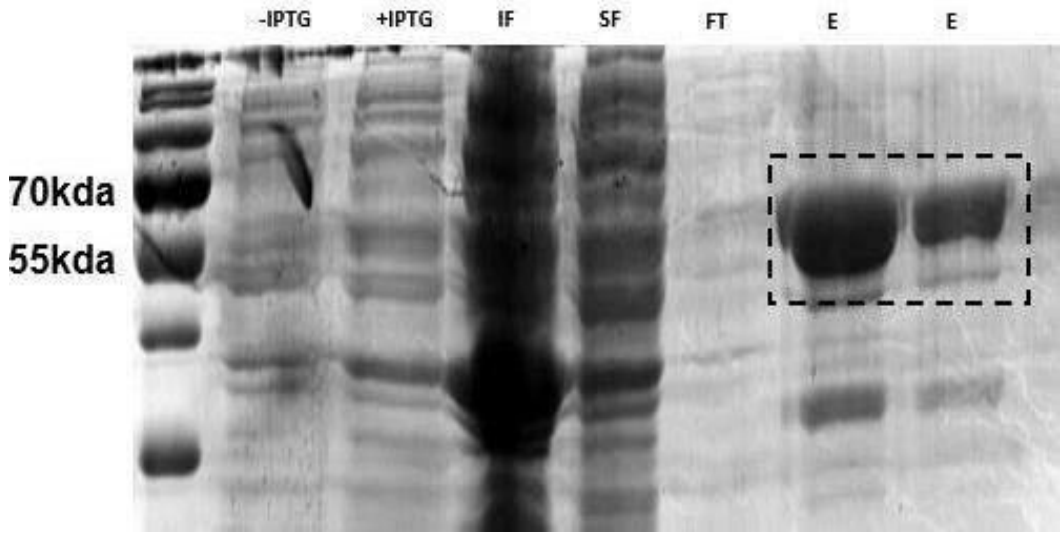


Fig. 3.1.2.1 - SDS-PAGE (15%) of TRAPα::GST (~ 56 kb) expression in E.coli BL21 RosettaStar cells, midiprep from E.coli JM101 cells. Ladder: Prestained Protein 10-180 kDa (PageRuler, Thermo Fisher), Coomassie Blue Staining. -IPTG = before induction, +IPTG= after induction, SF= soluble fraction, IF = insoluble fraction, FT= flow through, E = eluate (dashed rectangle).

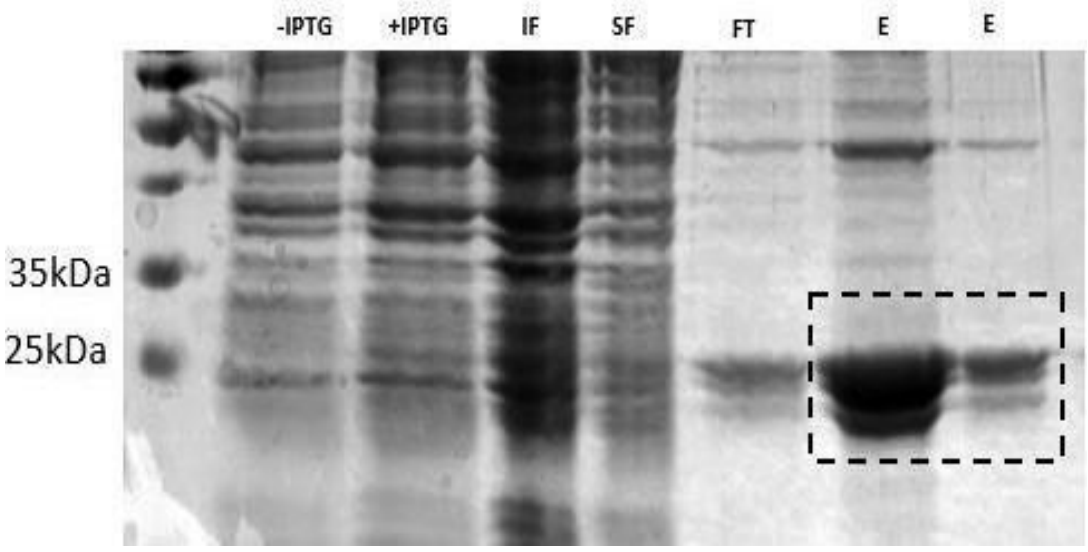


Fig. 3.1.2.2 - SDS-PAGE (15%) of TRAPα::HIS (~ 25 kb) expression in E.coli BL21 RosettaStar cells, midiprep from E.coli JM101 cells. Ladder: Prestained Protein 10-180 kDa, Coomassie Blue Staining. - IPTG = before induction, +IPTG = after induction, SF = soluble fraction, IF = insoluble fraction, FT= flow through, E = eluate (dashed rectangle).

The cells BL21 RosettaStar also increased the yield of TRAP β::GST, GST::TRAP γ and TRAP δ:: GST, and the expression of these three proteins is visible on the following SDS-PAGEs (Fig.3.1.2.3 - 3.1.2.5).

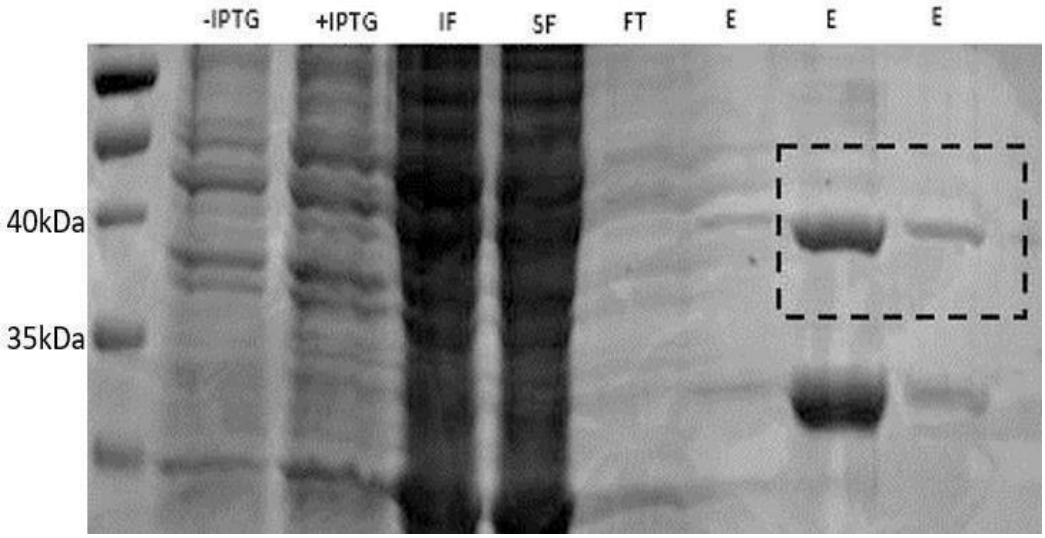


Fig. 3.1.2.3 - SDS-PAGE (15%) of TRAPβ::GST (~ 41 kb) expression in E.coli BL21 RosettaStar cells, midiprep from E.coli JM101 cells. Ladder: Prestained Protein 10-180 kDa, Coomassie Blue Staining. -IPTG = before induction, +IPTG = after induction, SF = soluble fraction, IF = insoluble fraction, FT = flow through, E = eluate (dashed rectangle).

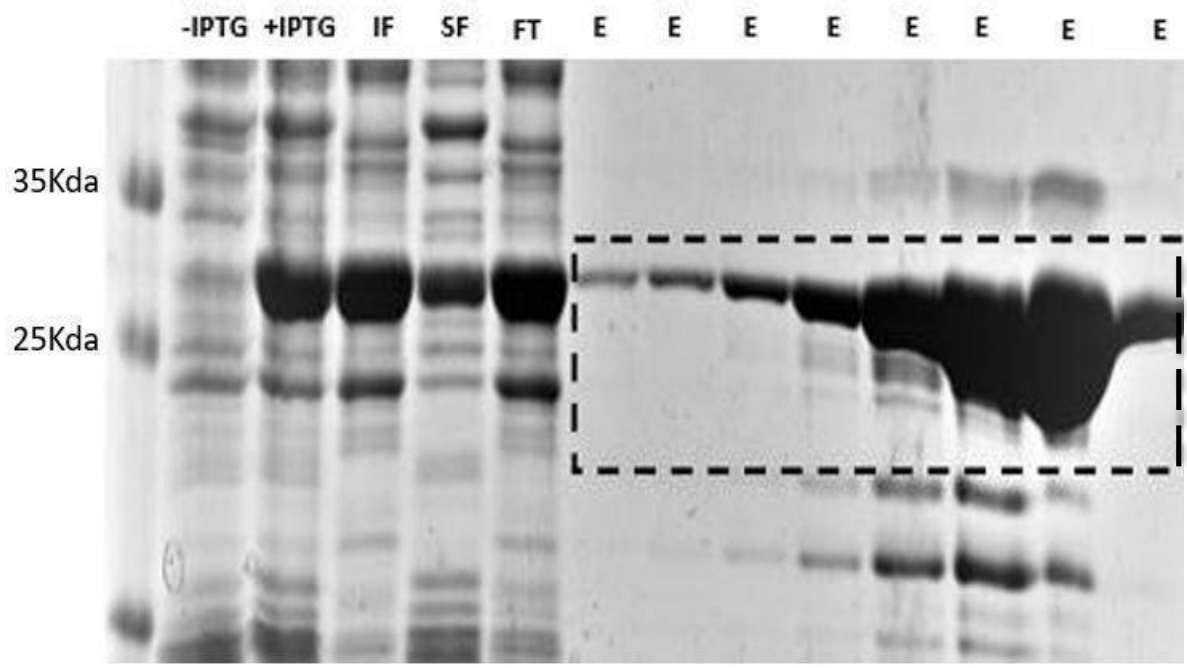


Fig. 3.1.2.4 - SDS-PAGE (15%) of GST::TRAP γ (~ 30 kb) expression in E.coli BL21cells, midiprep from E.coli JM101 cells. Ladder: Prestained Protein 10-180 kDa, Coomassie Staining. -IPTG = before induction, +IPTG = after induction, SF = soluble fraction, IF = insoluble fraction, FT = flow through, E = eluate (dashed rectangle).

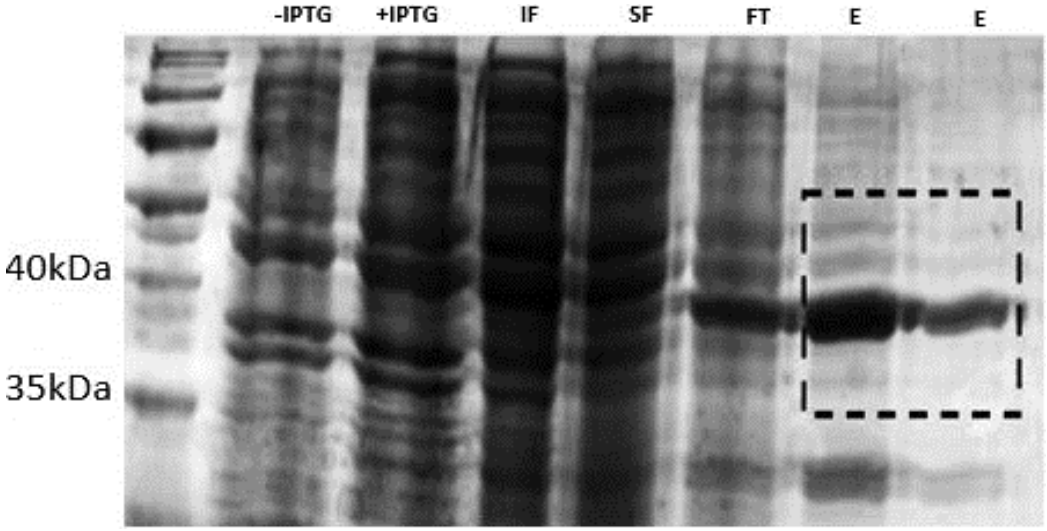


Fig. 3.1.2.5 - SDS-PAGE (15%) of TRAP δ::GST (~ 40 kb) expression in E.coli BL21 RosettaStar cells, midiprep from E.coli JM101 cells. Ladder: Prestained Protein 10-180 kDa, Coomassie Staining. -IPTG = before induction, +IPTG = after induction, SF = soluble fraction, IF = insoluble fraction, FT = flow through, E = eluate (dashed rectangle).

The concentration of the recombinant purified proteins by affinity chromatography was measured by nanodrop spectrophotometer, and the highest yield was achieved with GST:: TRAP γ. The mg of the purified proteins from 1liter culture are the following (Tab. 3.1.2.1; Fig. 3.1.2.6):

PROTEIN	mg /1L
TRAPα::GST	3
TRAPα::HIS	3
TRAPβ::GST	1
GST::TRAPγ	12
TRAPδ::GST	1

Tab. 3.1.2.1 – Concentration of purified recombinant proteins, mg/liter culture.

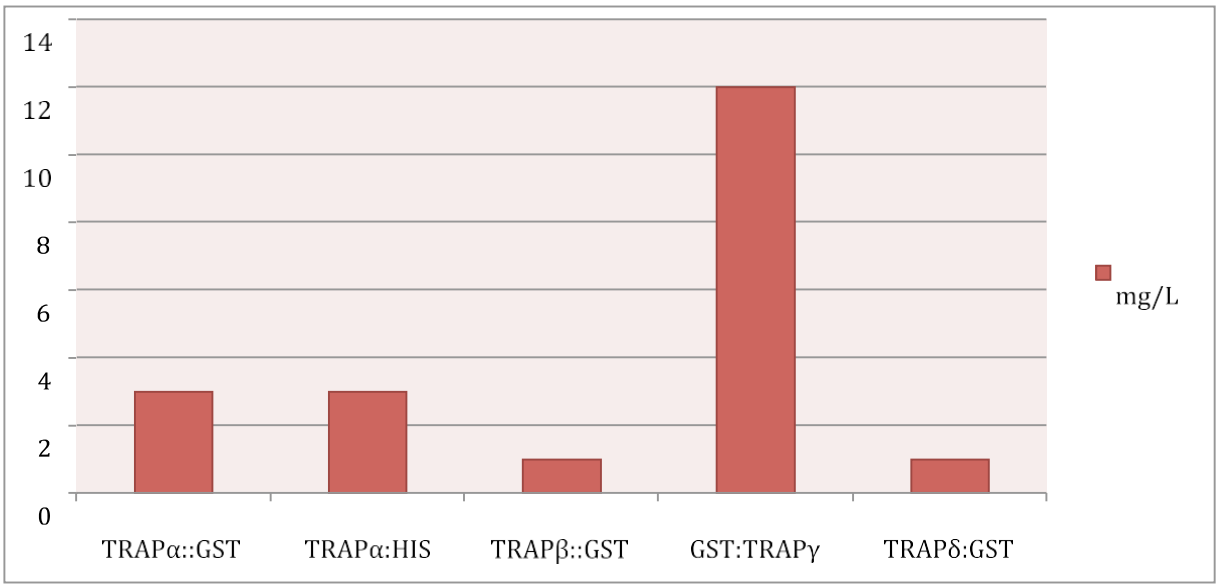
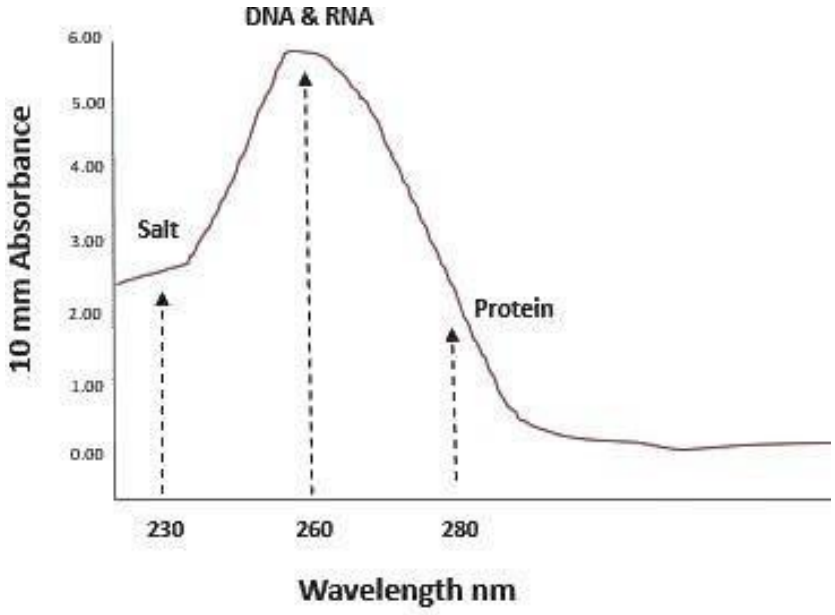


Fig. 3.1.2.6 - The graph points out the difference of concentration of all purified TRAP tagged proteins, mg for 1-liter bacteria culture.

The five TRAP recombinant purified protein shows a high purity level. The ratio 260:280 (λ 280) is between 0.5 and 0.8; by looking on the table below (Tab. 3.1.2.2), it means more than 95% of protein in solution.

% protein	260:280 ratio
100	0.57
95	1.06
90	1.32
70	1.73



Tab. 3.1.2.2 - The table shows the percentage of a protein concerning the OD ratio (260/280 nm). The graph shows the different Absorbance (OD) (y-axis) and wavelength (nm) (x-axis) of DNA/RNA, salt and protein; higher is the OD at 280 nm purer is the protein.

By analysing the SDS-PAGE is also possible to estimate the purity and the protein expression yield (Fig. 3.1.2.7).

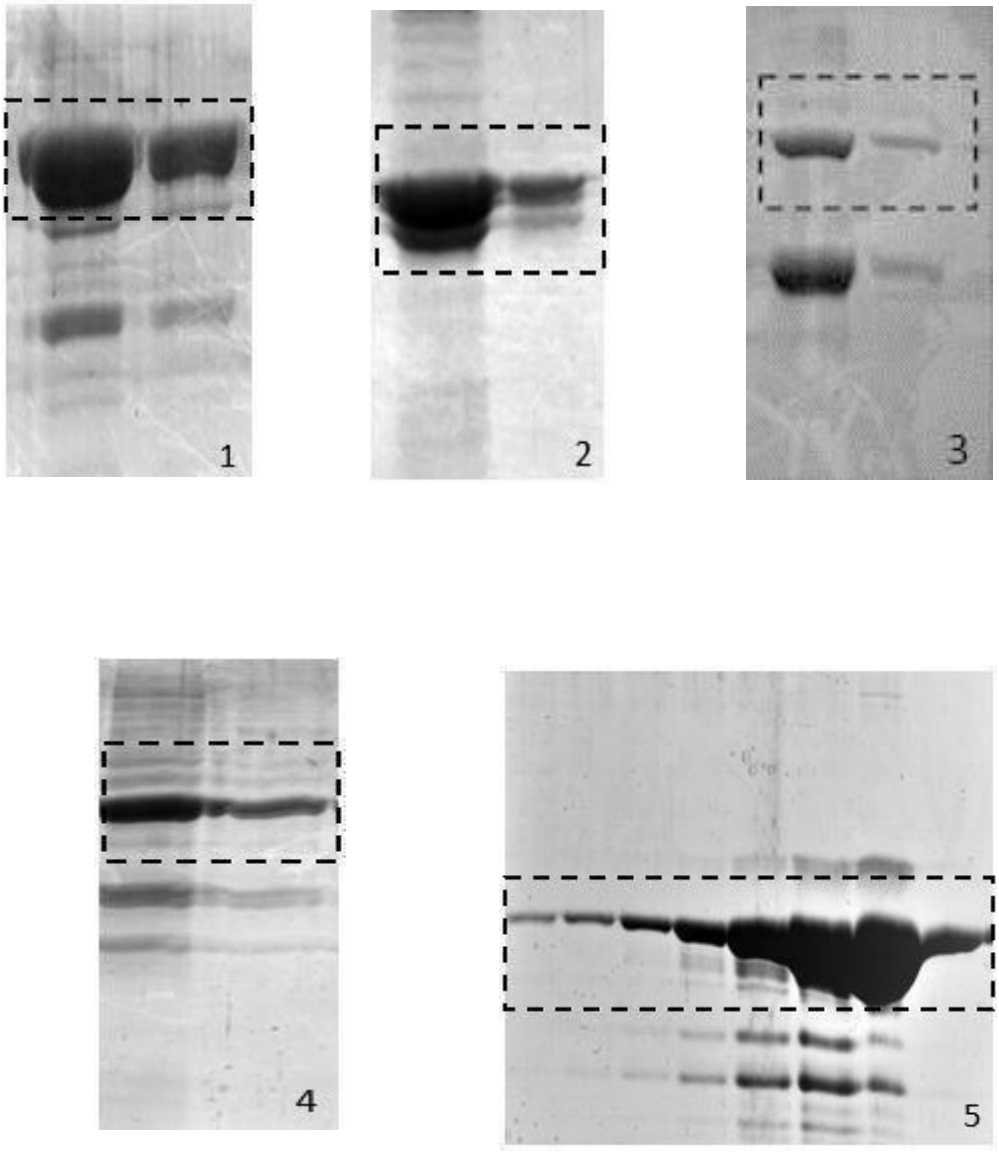


Fig. 3.1.2.7 - SDS-PAGE sections show the eluates (dashed rectangle) of each of five purified proteins. By analysing the size of the bands as well as the presence of extra bands it is possible to evaluate respectively concentration and purity. 1) = TRAPα::GST; 2) = TRAPα::HIS; 3) = TRAPβ::GST; 4) = TRAPδ::GST; 5) = GST::TRAPγ. Some extra bands represent GST tag dueto protein degradation.

The purified proteins are also very soluble, how the following SDS-PAGE points out (Fig. 3.1.2.8). The proteins were loaded before and after ultracentrifugation test, no significant difference is noticeable between the two bands.

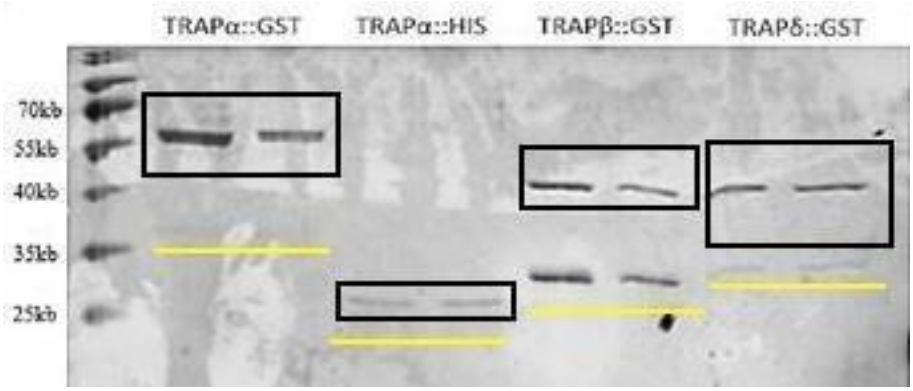


Fig. 3.1.2.8 - SDS PAGE (15%) shows the two bands of each protein, before and after the centrifugation test (rectangle) (Methods). In this gel is not present GST:: TRAP γ. The other bands are GST or HIS tags due to protein degradation (yellow lines). First two lanes: alpha TRAPα:: GST, third and fourth: TRAPα::HIS; fifth and sixth: TRAPβ:: GST; last two lines: TRAPδ:: GST.

The western blot confirmed the purified proteins TRAPα:: GST, TRAP α::HIS, TRAP β:: GST, GST:: TRAP γ and TRAPδ:: GST (GST or HIS antibody). A clear band with the expected size of each protein is visible on the transfer membrane (Fig. 3.1.2.9 - 3.1.2.11).

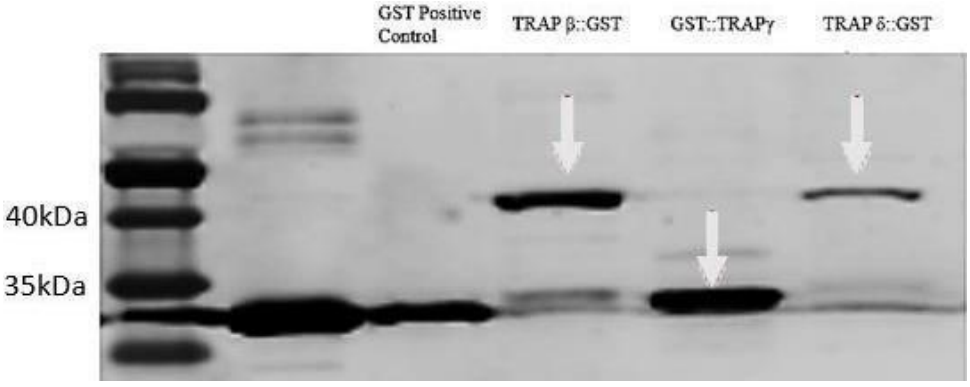


Fig. 3.1.2.9 - Western blot of TRAPβ:: GST, GST:: TRAPγ, TRAPδ:: GST (arrows). Lane 2: GST positive control, lane 3: TRAPβ::GST (~ 41 kDa), lane 4: GST::TRAPγ (~ 31 kDa), and lane 5: TRAδ::GST (~ 40 KDa). The other band, around 26 kDa, in each lane is GST due to protein degradation. GST antibody detection.

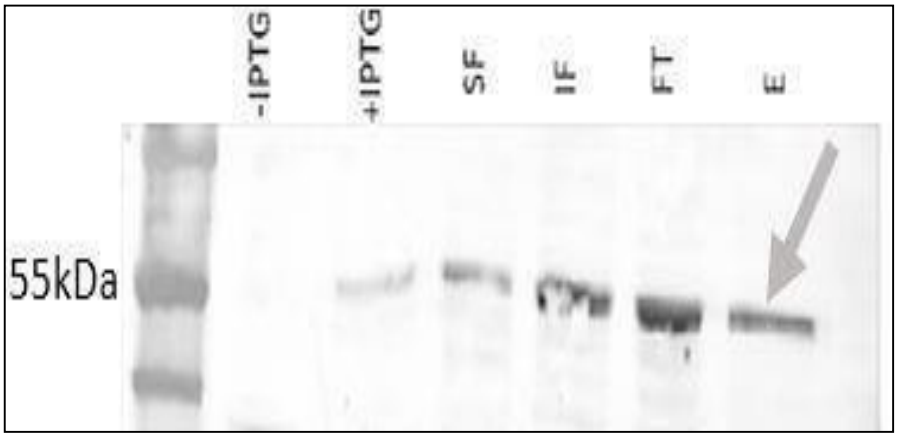


Fig. 3.1.2.10 - TRAPα:: GST Western blot: a band around 56 KDa is visible in all lanes except before IPTG induction. Eluate (arrow). GST antibody detection.

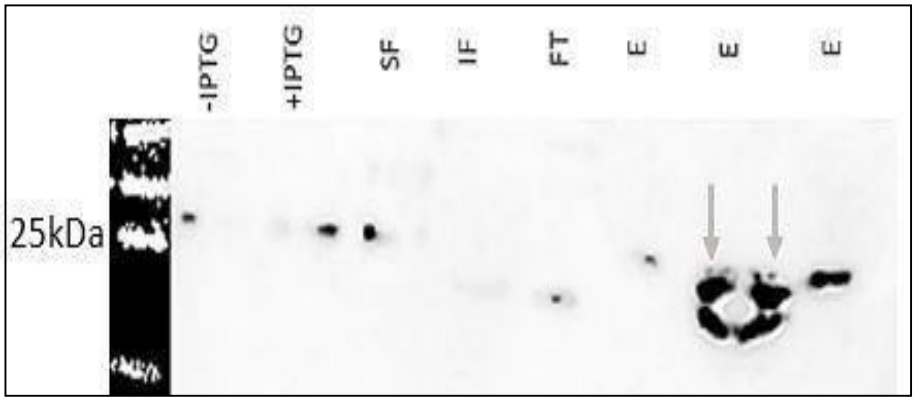
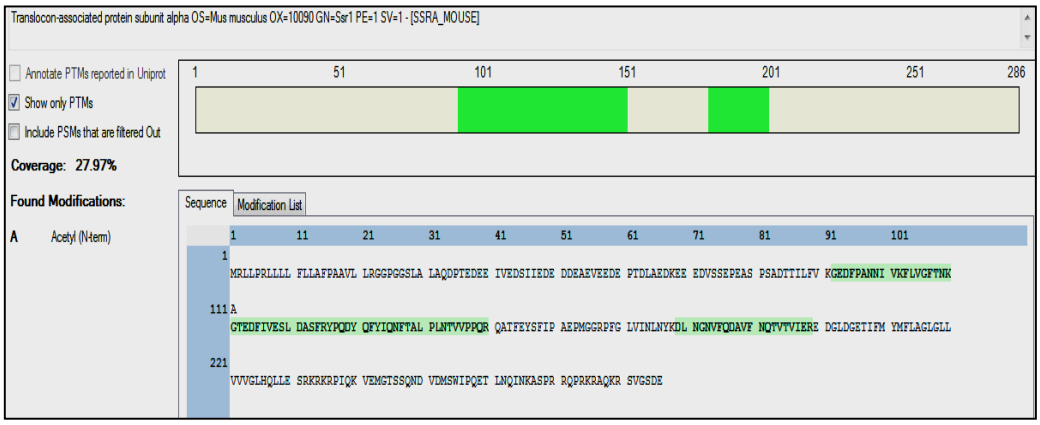
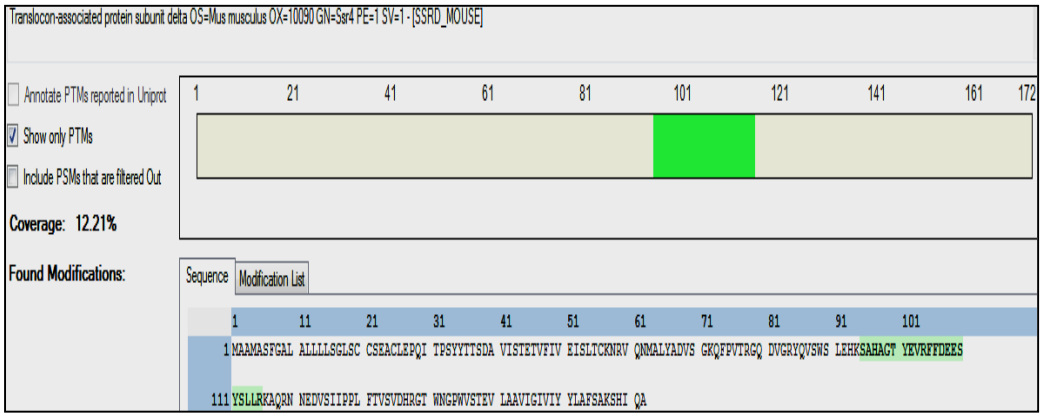


Fig. 3.1.2.11 - TRAPα:: HIS Western blot: clear bands are visible in the eluates (arrows) (~ 25 kDa). HIS antibody detection.

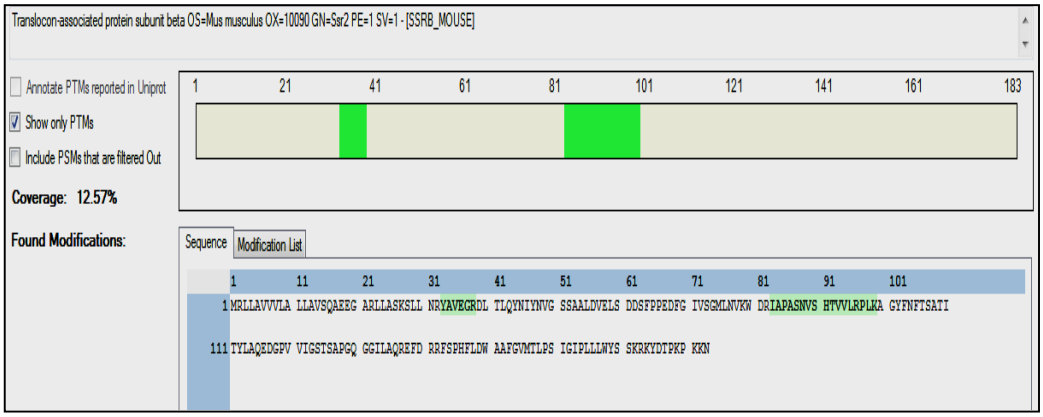
C. Fecher-Trost has analysed the proteins by Liquid Chromatography-Mass Spectrometry (LC-MS), and three TRAP recombinant proteins have been detected. The GST:: TRAP γ protein was not suitable for this method, the protein sequence is just 57 amino acid residues; the digestion with trypsin (cleavage site at the carboxyl side of Lysine or Arginine) leads to very small peptides. The coverage for TRAP α::GST is 27.97 %, two distinct amino acid residue sequences (Tab. 3.1.2.3); the coverage for TRAP β::GST is 12.57%, also two fragments of protein sequence (Tab.3.1.2.4); the coverage for TRAP δ::GST is 12.21%, one fragment in the sequence (Tab. 3.1.2.5).



Tab. 3.1.2.3 – LC-MS of TRAPα::GST: coverage is 27.9, the sequence from 92 to 150 and from 69 to 89 amino acid residues.



Tab. 3.1.2.4 – LC-MS of TRAPβ::GST: coverage is 12.57%, sequence from 33 to 38 and from 83 to 98 amino acid residues.



Tab. 3.1.2.5 – LC-MS of TRAPδ::GST: coverage is 12.21%, sequence from 95 to 115 amino acid residues.

GST protein was expressed by transforming BL21 (DE3) cells with pGEX-4T-PGS10 plasmid (protocol explained in methods). The achieved concentration was 7.35 mg/ml, the total amount of 3 ml in PBS- KMT buffer, was frozen at -80°C (Fig. 3.1.2.12). The ratio OD 260/280 = 0.55, 100% pure protein. The GST protein was used for the next experiments: GST pull-down assay, peptide array, and for western blot.

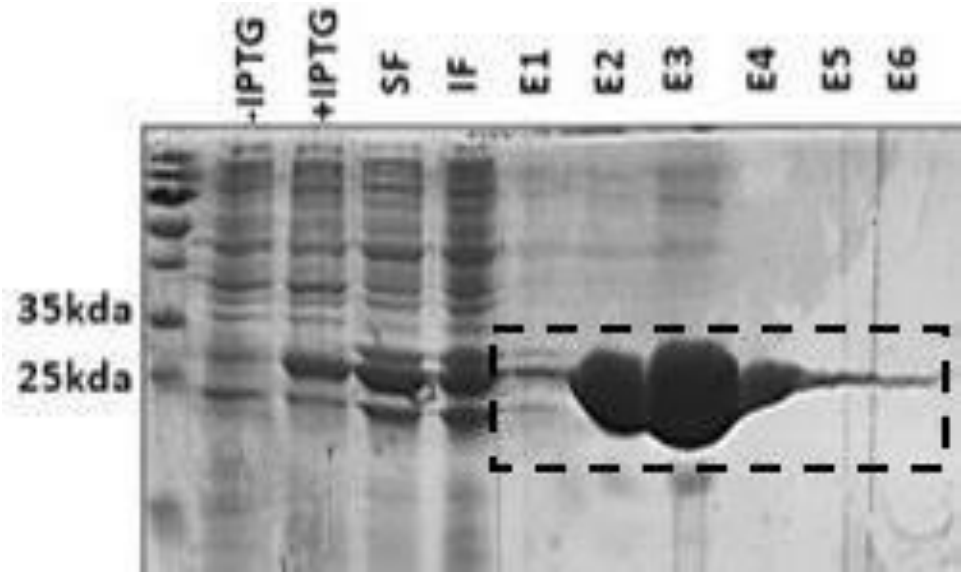


Fig. 3.1.2.12 - SDS-PAGE (15%) shows GST expression (~26 kda), plasmid pGEX-4T- PPGS10, (0.5µg/µl), expression in E.coli BL21 (DE3) cells. Ladder: Prestained Protein 10-180 kDa, Coomassie Blue Staining. -IPTG = before induction, +IPTG = after induction, SF = soluble fraction, IF = insoluble fraction, FT = flow through, E = eluate (dashed rectangle).

3.1.3 Pull-down assays

The pull-down assays were carried out to address the interaction between TRAP α and TRAP β; this investigation was sufficient to confirm this physical PPI. In the first assay, the bait was TRAPβ:: GST and the prey was TRAPα:: HIS, and the beads were GSH-Sepharose. The incubation of bait with the beads because of the GST tagged protein led to coupling, the second incubation with the prey clearly demonstrated the interaction between bait – TRAPβ:: GST and prey – TRAP α:: HIS; two bands of the right size are visible on the SDS-PAGE (15%) (Fig. 3.1.3.1 , 3.1.3.2). The same assay was performed by exchanging bait/prey and using Ni-NTA Agarose beads (bait was HIS tagged), and the same results were achieved (Fig. 3.1.3.3). Besides, the assay was performed as GST negative control, by incubating GST protein (bait) with GSH-Sepharose beads; then, in the second rolling the TRAPα:: HIS was added (prey) (Fig. 3.1.3.1-3.1.3.3)

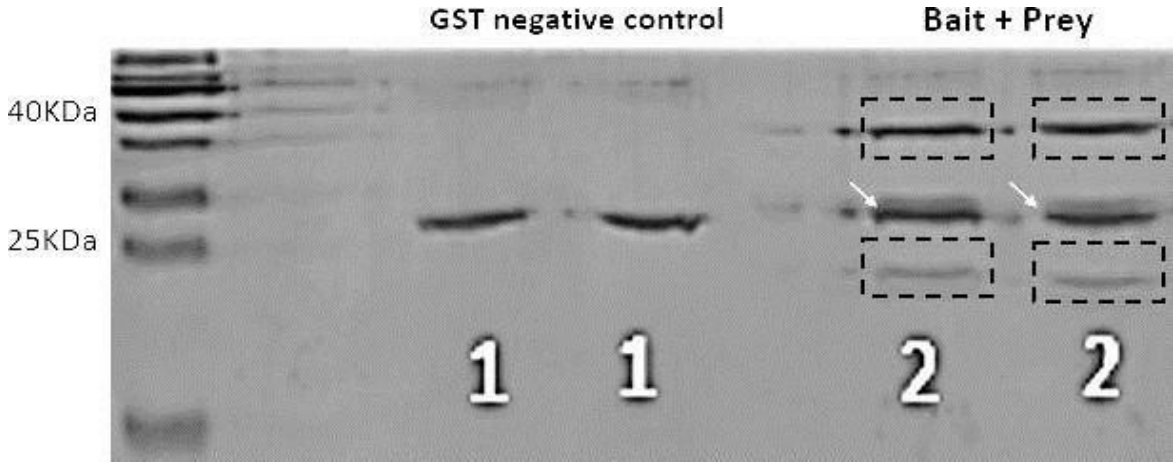


Fig. 3.1.3.1 - SDS-PAGE (15%) of the eluate after GST pull-down assay.

1) GST negative control pull-down assay. 1st rolling: GSH-Sepharose beads and GST protein/bait, 2nd rolling: TRAPα:: HIS/prey was added.

2) TRAPβ:: GST/bait and TRAPα:: HIS/prey pull-down assay. 1st rolling: GSH-Sepharose beads and TRAPβ:: GST/bait; second rolling: TRAPα:: HIS/prey was added.

On the lanes number 2 is visible a band around 40 kDa (TRAPβ:: GST) and a band around 25 kDa (TRAPα:: HIS). The intermediate band represents the GST tag due to protein degradation (arrows).

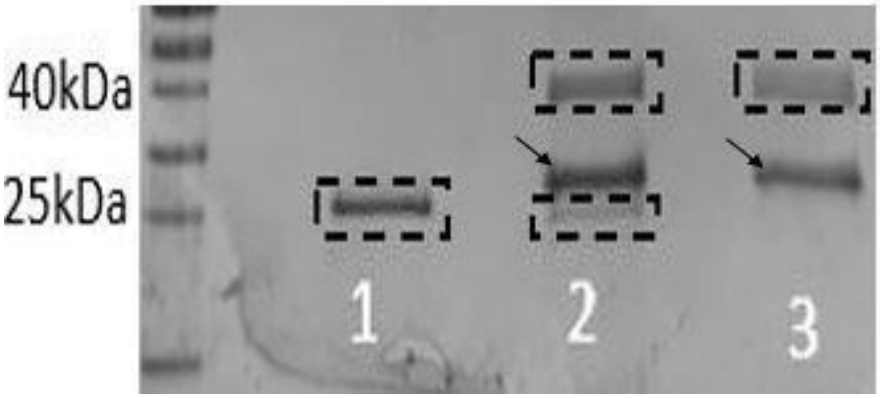


Fig. 3.1.3.2 - SDS-PAGE (15%) of the eluates after GST pull-down assays.

1) GST negative control pull-down assay. 1st rolling: GSH-Sepharose beads and GST protein/bait, 2nd rolling: TRAPα:: HIS/prey was added. GST band (dashed rectangle).

2) TRAPβ:: GST (bait) and TRAPα:: HIS (prey) pull-down assay. 1st rolling: GSH-Sepharose beads and TRAPβ:: GST/bait; 2nd rolling: TRAPα:: HIS/prey was added. TRAPβ:: GST (40 kDa) and TRAPα:: HIS) (dashed rectangles).

3) 1st rolling: GSH-Sepharose beads were incubated with TRAPβ:: GST/bait. TRAPβ:: GST band (dashed rectangle).

On the lane number 2 is visible a band around 40 kDa (TRAPβ:: GST) and a band around 25 kDa (TRAPα:: HIS) (dashed rectangles) which are missed on lane 1) and 3). The intermediate band is GST tag due to protein degradation (arrows).

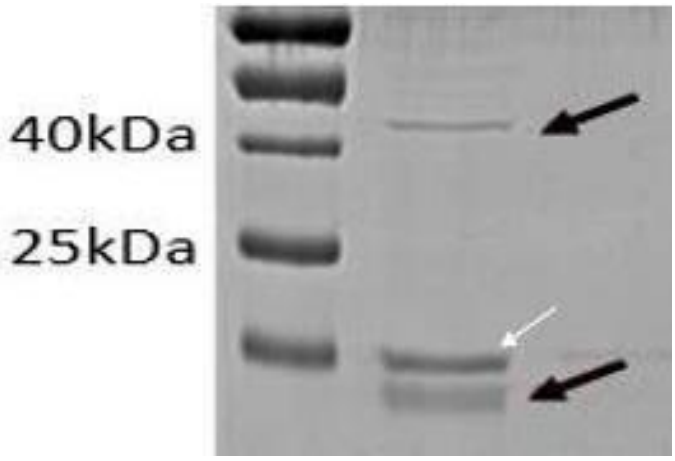


Fig. 3.1.3.3 - SDS-PAGE (15%) of the eluate after HIS pull-down assay. 1st rolling: Ni-NTA Agarose beads were incubated with TRAPα:: HIS/bait; 2nd rolling: TRAPβ:: GST/prey was added. A band around 25 kDa (bait) and 40 kDa (prey) are present on the gel (black arrows). The intermediate band is GST tag due to protein degradation (white arrow).

The protein-protein interaction (PPI) between TRAPα and TRAPβ takes place on the luminal domains of both proteins (cloned sequences); by this assay, it is not possible to establish the exact interface binding sequences.

3.1.4 Peptide array

The peptide array with Sec61α1 loop 5 spots (Fig. 3.1.4.5) was performed by incubating the cellulose membrane with the TRAPα/β:: GST and TRAPα:: HIS recombinant proteins. The results establish the physical interactions between the translocon luminal domain and the luminal domains of TRAP subunits. Cryo-ET analysis already addressed the proximity of these structures (Fig. 3.1.4.1).

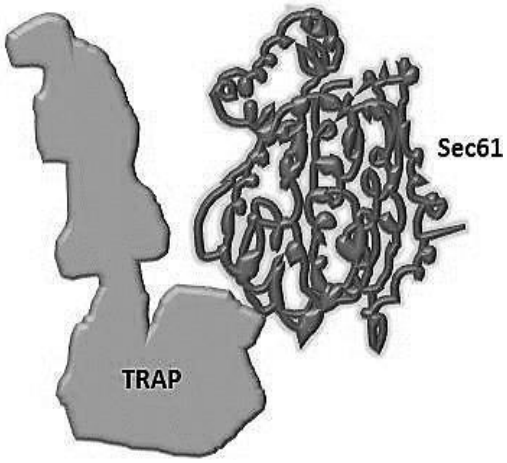


Fig. 3.1.4.1 – The carton shows the proximity of TRAP complex (α/β) with Sec61α1 loop 5 as determined by cryo-ET.

The first assay was performed with the proteins TRAPα:: HIS and TRAPβ:: GST, which were incubated with the membrane containing the Sec61α1 loop 5 spots (WT and mutated). The black dots point out the interaction between the translocon and the two TRAP subunits or one of them (Fig. 3.1.4.2).

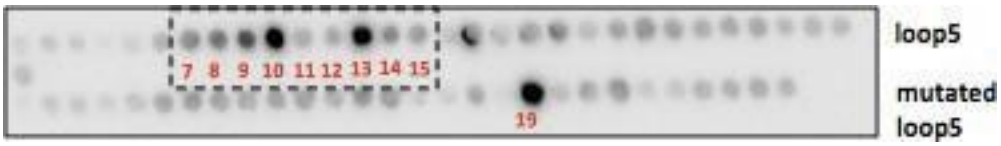


Fig. 3.1.4.2 - Peptide array of Sec61α1 loop5 and TRAPα:: HIS/TRAPβ:: GST (GST antibody detection). The cellulose membrane with 30 overlapping WT Sec61α1 loop5 spots (upper lane) and 28 of mutated ovelapping Sec61α1 loop5 spots (lower lane) (alanine substitutes anotheramino aicds along the sequence) was incubated with the solution (peptide array buffer) contaning TRAPα:: HIS/TRAPβ:: GST; then, antibody incubation and detection was carried out.

The spots that correspond to the black dots in WT Sec61α1 loop5 and TRAPα:: HIS/TRAPβ:: GST array are shown below, as well as the mutated spot where a black dot is present (Tab. 3.1.4.1).

1 KAFSPTTVNTGRGMEFEGAI	1 KAFSPTTVNTGRGMEFEGAI
2 AFSPTTVNTGRGMEFEGAII	2 AFSPTTVNTGRGMEFEGAII
3 FSPTTVNTGRGMEFEGAIIA	3 FSPTTVNTGRGMEFEGAIIA
4 SPTTVNTGRGMEFEGAIIAL	4 SPTTVNTGRGMEFEGAIIAL
5 PTTVNTGRGMEFEGAIIALF	5 PTTVNTGRGMEFEGAIIALF
6 TTVNTGRGMEFEGAIIALFH	6 TTVNTGRGMEFEGAIIALFH
7 TVNTGRGMEFEGAIIALFHLL	7 TVNTGRGMEFEGAIIALFHLL
8 VNTGRGMEFEGAIIALFHLLA	8 VNTGRGMEFEGAIIALFHLLA
9 NTGRGMEFEGAIIALFHLLAT	9 NTGRGMEFEGAIIALFHLLAT
10 TGRGMEFEGAIIALFHLLATR	10 TGRGMEFEGAIIALFHLLATR
11 GRGMEFEGAIIALFHLLATRT	11 GRGMEFEGAIIALFHLLATRT
12 RGMEFEGAIIALFHLLATRTD	12 RGMEFEGAIIALFHLLATRTD
13 GMEFEGAIIALFHLLATRTDK	13 GMEFEGAIIALFHLLATRTDK
14 MEFEGAIIALFHLLATRTDKV	14 MEFEGAIIALFHLLATRTDKV
15 EFEGAIIALFHLLATRTDKVR	15 EFEGAIIALFHLLATRTDKVR
16 FEGAIIALFHLLATRTDKVRA	16 FEGAIIALFHLLATRTDKVRA
17 EGAIIALFHLLATRTDKVRAL	17 EGAIIALFHLLATRTDKVRAL
18 GAIIALFHLLATRTDKVRALR	18 GAIIALFHLLATRTDKVRALR
19 AIIALFHLLATRTDKVRALRE	19 AIIALFHLLATRTDKVRALRE
20 IIALFHLLATRTDKVRALREA	20 IIALFHLLATRTDKVRALREA
21 IALFHLLATRTDKVRALREAF	21 IALFHLLATRTDKVRALREAF
22 ALFHLLATRTDKVRALREAFY	22 ALFHLLATRTDKVRALREAFY
23 LFHLLATRTDKVRALREAFYR	23 LFHLLATRTDKVRALREAFYR
24 FHLLATRTDKVRALREAFYRQ	24 FHLLATRTDKVRALREAFYRQ
25 HLLATRTDKVRALREAFYRQN	25 HLLATRTDKVRALREAFYRQN
26 LLATRTDKVRALREAFYRQNL	26 LLATRTDKVRALREAFYRQNL
27 LATRTDKVRALREAFYRQNLP	27 LATRTDKVRALREAFYRQNLP
28 ATRTDKVRALREAFYRQNLPN	28 ATRTDKVRALREAFYRQNLPN
29 TRTDKVRALREAFYRQNLPNL	
30 RTDKVRALREAFYRQNLPNLM	

Tab. 3.1.4.1 – The spots that correspond to the black dots of array WT Sec61α1 loop5 and TRAPß/ TRAPα:: HIS are in blue (left). The mutated Sec61α1 loop5 spot where there is a black dot is 19 (right, red).

The overlapping peptides for Sec61α1 loop5 /TRAPß/ TRAPα:: HIS array are listed in blue (see Tab. 3.1.4.1).

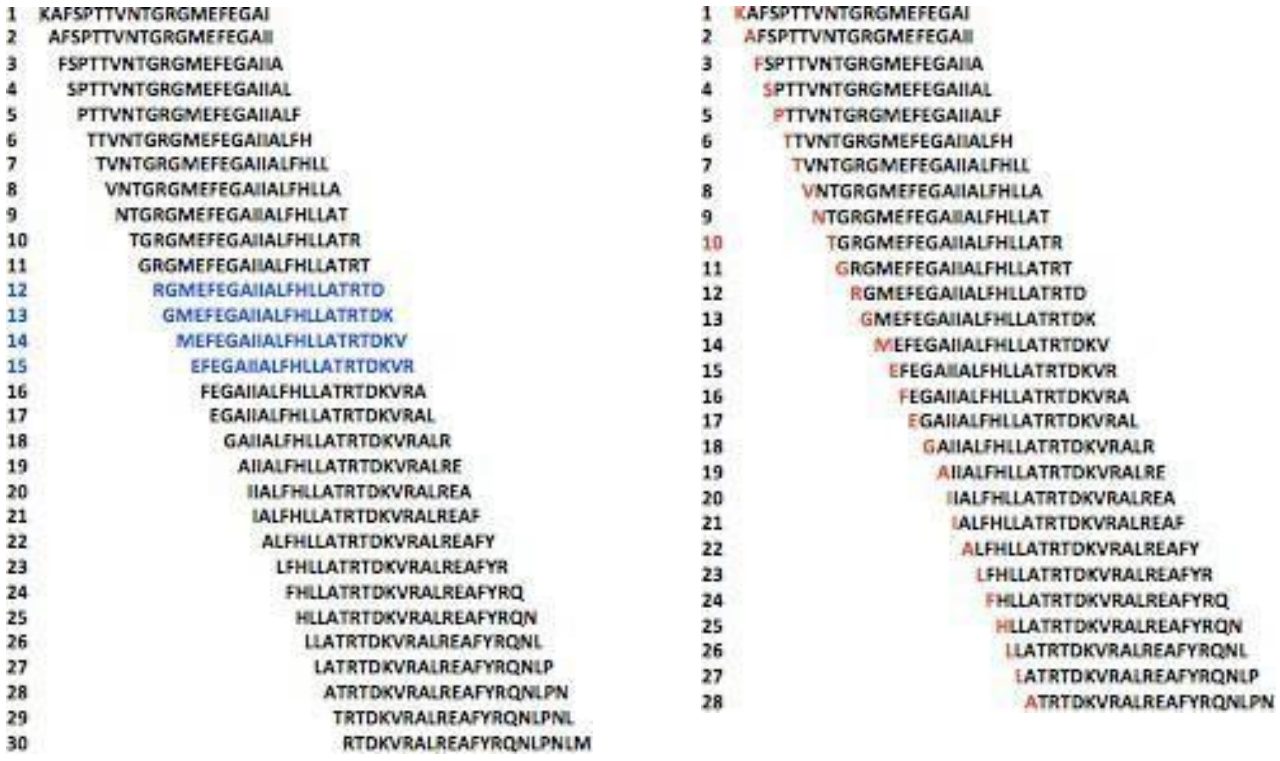
TVNTGRGMEFEGAIIALFHL
VNTGRGMEFEGAIIALFHLL
NTGRGMEFEGAIIALFHLLA
TGRGMEFEGAIIALFHLLAT
GRGMEFEGAIIALFHLLATR
RGMEFEGAIIALFHLLATRT
GMEFEGAIIALFHLLATRTD
MEFEGAIIALFHLLATRTDK
EFEGAIIALFHLLATRTDKR

A black dot on the spot 19 of mutated loop 5 (see Tab. 3.1.4.1): there is not substitution, an alanine (A) is present.

AIIALFHLLATRTDKVRALR

To address, which subunit or if both TRAP subunits interact with the translocon, the membrane was incubated separately with TRAPα:: GST, TRAPα:: HIS, and TRAPβ:: GST. In all three experiments, black dots are detected, which means that Sec61α loop5 interacts with TRAPα luminal domain and TRAPβ luminal domain; the latter, as the pull-down assays demonstrated, interact with one another (Fig. 3.1.4.3, 3.1.4.4) (Tab. 3.1.4.2, 3.1.4.3).

Fig. 3.1.4.3 - Peptide array of Sec61α1 loop5 and TRAPα:: HIS (HIS antibody detection). The cellulose membrane with 30 overlapping WT Sec61α1 loop5 spots (upper lane) and 28 of mutated ovelapping Sec61α1 loop5 spots (lower lane) (alanine substitutes another amino aicds along the sequence) was incubated with the solution (peptide array buffer) contaning TRAPα:: HIS; then, antibody incubation and detection was carried out. (Other spots are not visible when the membrane is reused).



Tab. 3.1.4.2 – The spots that correspond to the black dots of array WT Sec61α1 loop5 and TRAPα:: HIS are in blue (left), spots 12. 13, 14, 15. The mutated WT Sec61α1 loop5 spot where there is a black dot is 10 (right, red).

The overlapping peptides for WT Sec61α1 loop5 TRAPα::HIS array (see Tab. 3.1.4.2) are listed in blue below:

RGMEFEGAIIALFHLLATRT
GMEFEGAIIALFHLLATRTD
MEFEGAIIALFHLLATRTDK
EFEGAIIALFHLLATRTDKR

In the mutated loop5 spot 10, a threonine (T) has been substituted with an alanine (A) (see Tab. 3.1.4.2).

TGRGMEFEGAIIALFHLLAT

By comparing the array TRAPα:: HIS/TRAPß:: GST with the array TRAPα:: HIS the overlapping peptide sequence is EFEGAIIALFHL. However, there is not the same overlapping sequence with TRAPα:: GST array (not shown); consequently, we cannot establish the interface binding sequence.

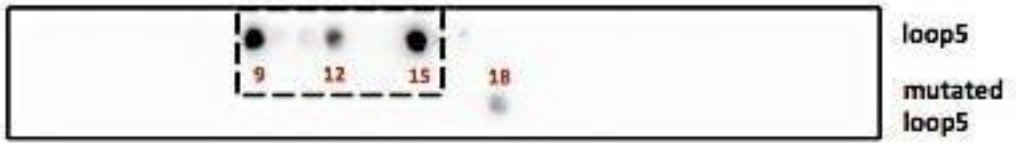


Fig. 3.1.4.4 - Peptide array of Sec61α1 loop5 and TRAPβ:: GST (GST antibody detection). The cellulose membrane with 30 overlapping WT Sec61α1 loop5 spots (upper lane) and 28 of mutated ovelapping Sec61α1 loop5 spots (lower lane) (alanine substitutes another amino aicds along the sequence) was incubated with the solution (peptide array buffer) contaning TRAPβ:: GST; then, antibody incubation and detection was carried out. (Other spots are not visible when the membrane is reused).

1	KAFSPTTVNTGRGMEFEGAI
2	AFSPTTVNTGRGMEFEGAII
3	FSPTTVNTGRGMEFEGAIIA
4	SPTTVNTGRGMEFEGAIIAL
5	PTTVNTGRGMEFEGAIIALF
6	TTVNTGRGMEFEGAIIALFH
7	TVNTGRGMEFEGAIIALFHLL
8	VNTGRGMEFEGAIIALFHLLA
9	NTGRGMEFEGAIIALFHLLAT
10	TGRGMEFEGAIIALFHLLATR
11	GRGMEFEGAIIALFHLLATRT
12	RGMEFEGAIIALFHLLATRTD
13	GMEFEGAIIALFHLLATRTDK
14	MEFEGAIIALFHLLATRTDKV
15	EFEGAIIALFHLLATRTDKVR
16	FEGAIIALFHLLATRTDKVRA
17	EGAIIALFHLLATRTDKVRAL
18	GAIIALFHLLATRTDKVRALR
19	AIIALFHLLATRTDKVRALRE
20	IIALFHLLATRTDKVRALREA
21	IALFHLLATRTDKVRALREAF
22	ALFHLLATRTDKVRALREAFY
23	LFHLLATRTDKVRALREAFYR
24	FHLLATRTDKVRALREAFYRQ
25	HLLATRTDKVRALREAFYRQN
26	LLATRTDKVRALREAFYRQNL
27	LATRTDKVRALREAFYRQNLP
28	ATRTDKVRALREAFYRQNLPN
29	TRTDKVRALREAFYRQNLPNL
30	RTDKVRALREAFYRQNLPNLM

1	KAFSPTTVNTGRGMEFEGAI
2	AFSPTTVNTGRGMEFEGAII
3	FSPTTVNTGRGMEFEGAIIA
4	SPTTVNTGRGMEFEGAIIAL
5	PTTVNTGRGMEFEGAIIALF
6	TTVNTGRGMEFEGAIIALFH
7	TVNTGRGMEFEGAIIALFHLL
8	VNTGRGMEFEGAIIALFHLLA
9	NTGRGMEFEGAIIALFHLLAT
10	TGRGMEFEGAIIALFHLLATR
11	GRGMEFEGAIIALFHLLATRT
12	RGMEFEGAIIALFHLLATRTD
13	GMEFEGAIIALFHLLATRTDK
14	MEFEGAIIALFHLLATRTDKV
15	EFEGAIIALFHLLATRTDKVR
16	FEGAIIALFHLLATRTDKVRA
17	EGAIIALFHLLATRTDKVRAL
18	GAIIALFHLLATRTDKVRALR
19	AIIALFHLLATRTDKVRALRE
20	IIALFHLLATRTDKVRALREA
21	IALFHLLATRTDKVRALREAF
22	ALFHLLATRTDKVRALREAFY
23	LFHLLATRTDKVRALREAFYR
24	FHLLATRTDKVRALREAFYRQ
25	HLLATRTDKVRALREAFYRQN
26	LLATRTDKVRALREAFYRQNL
27	LATRTDKVRALREAFYRQNLP
28	ATRTDKVRALREAFYRQNLPN

Tab. 3.1.4.3 – The spots that correspond to the black dots of array WT Sec61α1 loop5 TRAPβ::GST are in blue (left), 9,12,and 15. The mutated Sec61α1 loop5 spot with the black dot is 18 (right, red).

The overlapping peptides for Sec1α1/TRAPß:: GST array (see Tab. 3.1.4.3) are listed in blue below:

NTGRGMEFEGAIIALFHLLA
RGMEFEGAIIALFHLLATRT
EFEGAIIALFHLLATRTDKR

In the mutated Sec61α1 loop5 spot 18 where there is a black dot, a glycine (G) has been substituted with alanine (A) (see Tab. 3.1.4.3).

GAIIALFHLLATRTDKVR

The interface sequence of the interaction TRAPβ:: GST/Sec61α loop5 seems to be EFEGAIIALFHLLA; but, also in this array the GST tag can compromise the results and analysis.

The same array with the cellulose membrane containing loop 5 Sec61α1 spots (WT and mutated) was performed with GST- antibody and protein as negative controls (incubation), no black dots are present in both membranes.

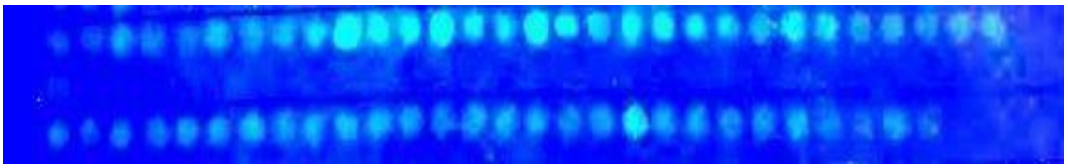
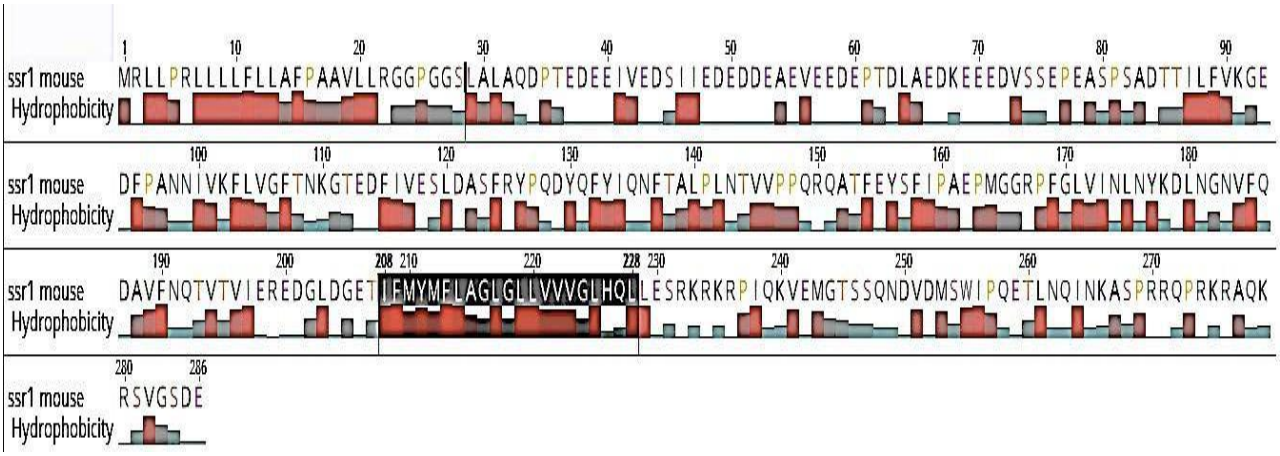


Fig. 3.1.4.5 - Spots on the cellulose membrane visible by UV light (peqLab), the peptides adsorb at approximately 280 nm. Above: WT Sec61α1 loop 5 (30 spots); below: mutated Sec61α1 loop 5 (28 spots).

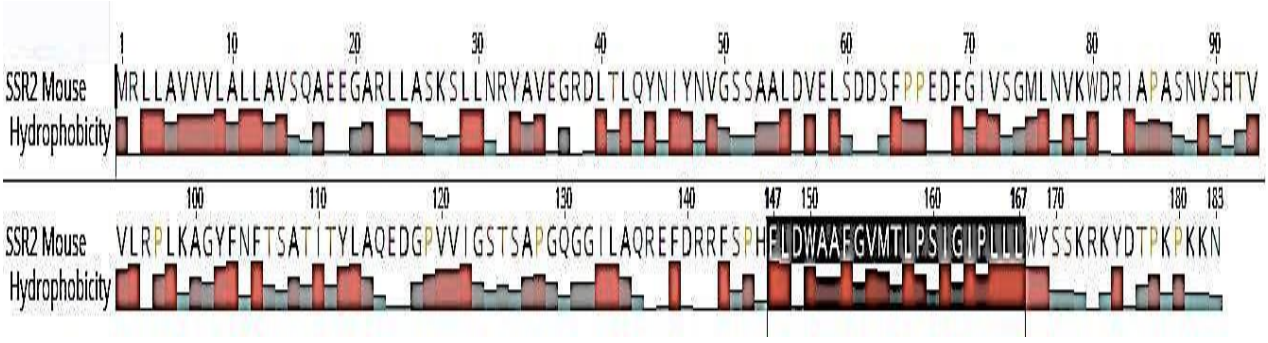
3.2 Computational results

3.2.1 Hydrophobicity/TMDs of TRAP subunits

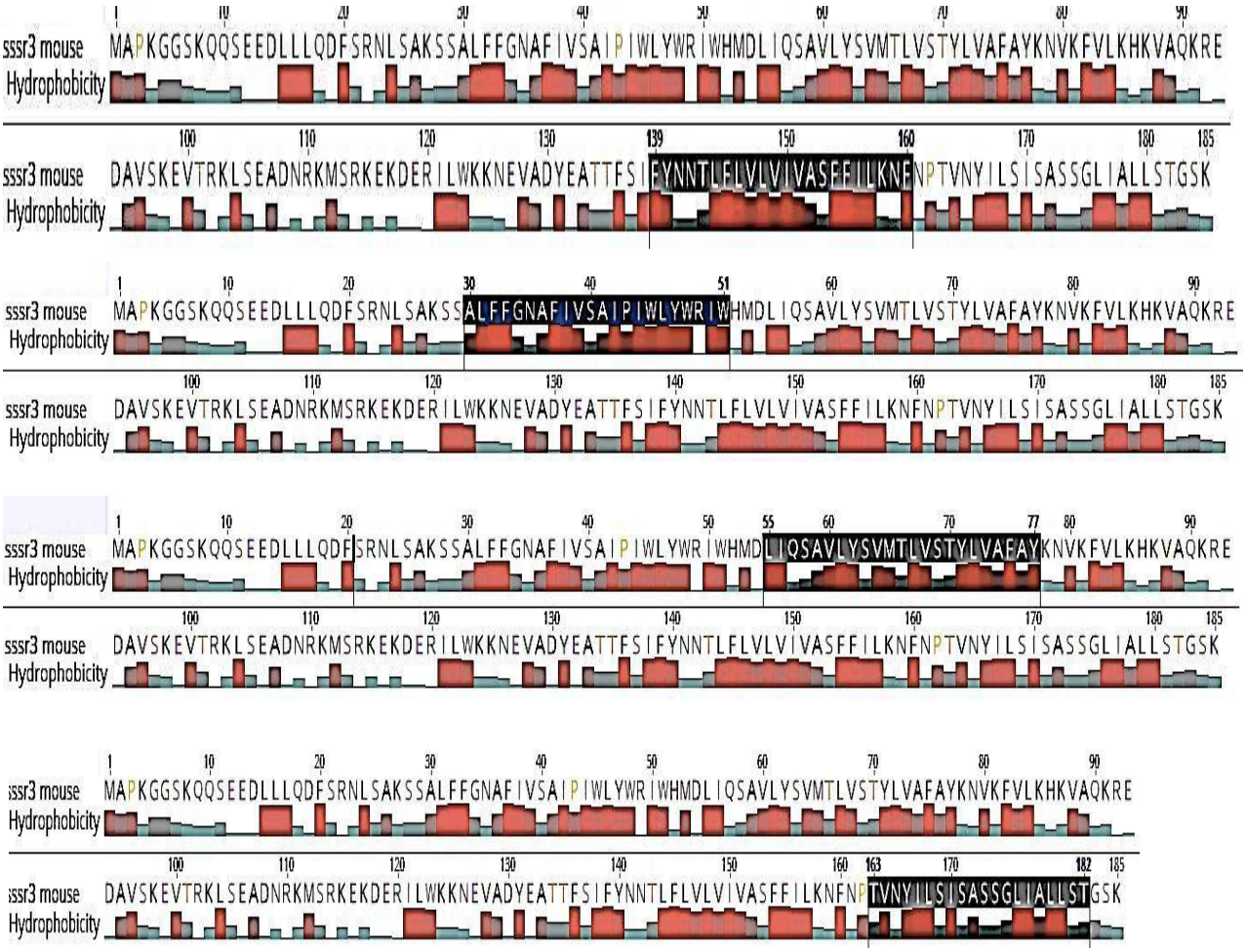
The hydrophobicity of TMDs in TRAP subunits, how expected, is very high. The Geneious analysis shows that the most common amino acids are Ala(A), Ile(I), Val(V), and Leu(L) (Tab. 3.2.1.1 - 3.2.1.4).



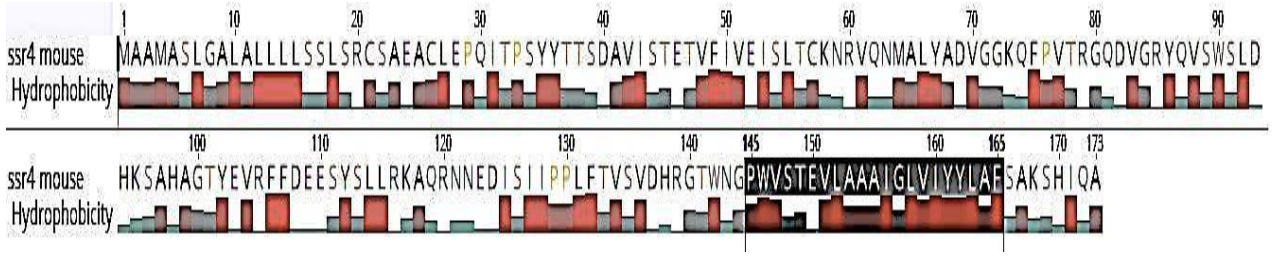
Tab. 3.2.1.1 – *Mus musculus* ssr1 TMD hydrophobicity (red bars), 208-228 residues. A=4.8% (entire sequence=6.3%). I= 4.8% (entire sequence= 4.9%), L= 28.3% (entire sequence=10.5%), V=14.3% (entire sequence= 7.3%).



Tab. 3.2.1.2 – *Mus musculus* ssr2 TMD hydrophobicity (red bars), 147-167 residues. A=9.5% (A entire sequence=9.8%), I= 9.5% (entire sequence=4.4%), L=23.8% (entire sequence= 12.6%), V=4.8% (entire sequence= 8.2%).



Tab. 3.2.1.3 – *Mus musculus* ssr3 hydrophobicity in four TMDs (red bars). A=13.6%, 13%, 4.3%, 10% (entire sequence=8.6%), I=18.2%, 4.3%, 13%, 15% (entire sequence= 6.5), L=9.1%, 17.4%, 17.4%, 20% (entire sequence= 11.4), V=4.5%, 17.4%, 13%, 5% (entire sequence= 8.1%).



Tab. 3.2.1.4 - *Mus musculus* ssr4 TMD hydrophobicity (red bars). A=19% (A entire sequence= 11%), I=9.5% (entire sequence=5.8%), L=14.3% (entire sequence= 9.8%), V=14.3% (entire sequence=8.1%).

3.2.2 TRAP alpha non-canonical EF-hand motif

By aligning TRAPα ubiquitous isoform (most common isoform) protein sequence with a non-canonical EF-hand motif, I have found this motif at the N-terminus in *Mus musculus* and Human (Tab. 3.2.2.2, 3.2.2.3). In the table below (Tab. 3.2.2.1) non-canonical and canonical EF-hand domain are present in human α-Parvalbumin, a protein involved in intracellular calcium signalling (Wang et al., 2013).

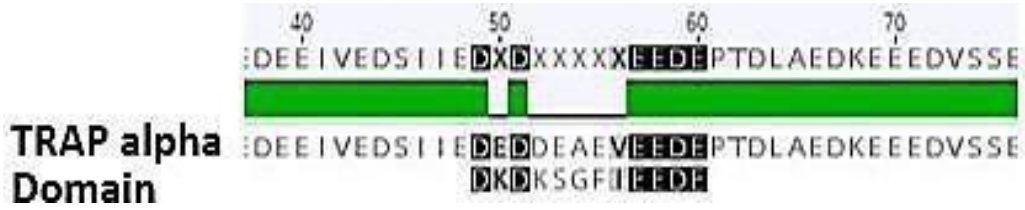
ADDVKKVFHMLDKDKSGFIEEDELGFILKG

non-canonical EF-hand domain
CD domain

KETKMLMAAGDKDGDGKIGVDEFSTLVAE

canonical EF-hand domain

Tab. 3.2.2.1 – Non-canonical EF-hand domain (above) and EF-hand domain (below) found in human α- Palvalbumin (Parv). Parvalbumin is a calcium-binding protein involved in intracellular calcium signalling. CD = calcium-binding.



MRLLPRLLLLFLLAFPAAVLLRGGPGGSLALAQDPTEDEEIVEDSIIEDEDDEAEVEEDE
PTDLAEDKEEEDVSSEPEASPSADTTILFVKGEDFPANNIVKFLVGFTNKGTEDFIVESL
DASFRYPQDYQFYIQNFTALPLNTVVPPQRQATFEYSFIPAEPMGGRPFGLVINLNYKDL
NGNVFQDAVFNQTVTVIEREDGLDGETIFMYMFLAGLGLLVVVGLHQLLESRKRKRPIQK
VEMGTSSQNDVDMSWIPQETLNQINKASPRRQPRKRAQKRSVGSDE

Tab. 3.2.2.2 – Above: the alignment between TRAPα sequence and the non-canonical EF-hand domain (Parv). Below: the entire *Mus musculus* TRAPα sequence and the likely non-canonical EF-hand domain in red. Signal Peptide sequence in grey.

MRLLPRLLLLLLLVFPATVLFRGGPRGLLAVAQDLTEDEETVEDSIIEDEDDEAEVEEDE
PTDLVEDKEEEDVSGEPEASPSADTTILFVKGEDFPANNIVKFLVGFTNKGTEDFIVESL
DASFRYPQDYQFYIQNFTALPLNTVVPPQRQATFEYSFIPAEPMGGRPFGLVINLNYKDL
NGNVFQDAVFNQTVTVIEREDGLDGETIFMYMFLAGLGLLVIVGLHQLLESRKRKRPIQ
K VEMGTSSQNDVDMSWIPQETLNQINKASPRRLPRKRAQKRSVGSDE

Tab. 3.2.2.3 - The sequence of human TRAP α most common isoform: signal peptide (grey), and probable non-canonical EF-hand domains (red).

3.2.3 STRING and RaptorX predictions

STRING server predicts the TRAP α/TRAP β interaction and the interaction between TRAP α/β and Sec61α1; this aspect confirms the results achieved by pull-down assays and peptide array. The TRAP α prediction includes interaction with TRAP β and with Sec61α1, and the same prediction is for TRAP β (Fig. 3.2.3.1, 3.2.3.2).

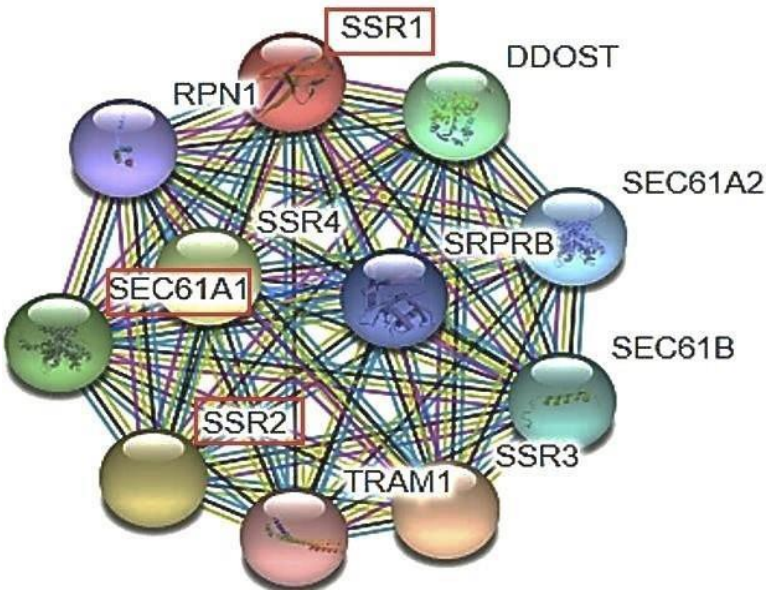


Fig. 3.2.3.1- STRING prediction for TRAP α (ssr1): among the possible PPI partners, there is ssr2 (TRAP β) and Sec61α isoform 1.

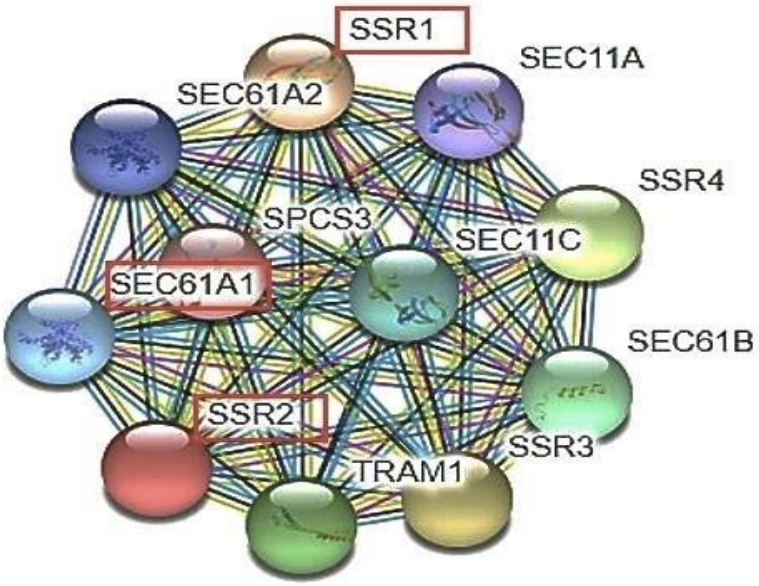


Fig. 3.2.3.2 - STRING prediction for TRAP β (ssr2): among the possible PPI partners, there is ssr1 (TRAPα) and Sec61α isoform 1.

As already described in the Introduction, TRAP δ (ssr4) subunit is associated with a congenital disorder of glycosylation (ssr4 CDG). It is plausible that this subunit interacts with the OST subunit DDOST, an essential subunit for complex stabilization; this interaction is also predicts by STRING (Fig. 3.2.3.3).

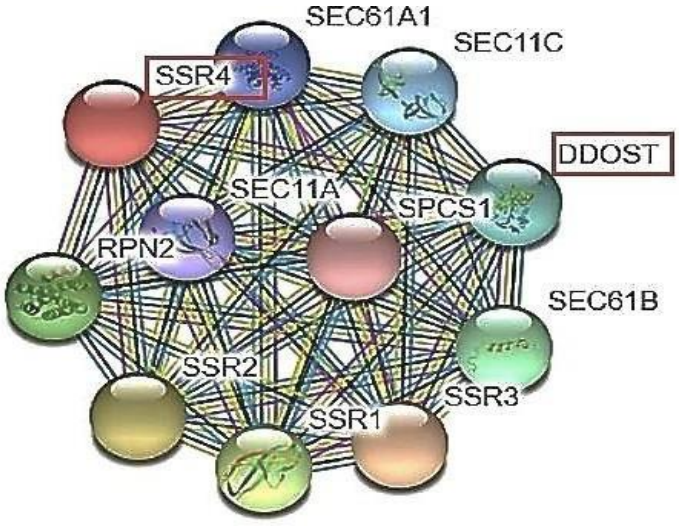


Fig. 3.2.3.3 – STRING interaction prediction for TRAP δ: one possible partner is DDOST subunit of OST complex.

The RaptorX server predicts the protein structure of TRAP α, β, γ, δ subunits without the SP sequences; the results are shown in the following figures (3.2.3.4- 3.2.3.7).
In the secondary structure of the TMDs, as expected, alpha-helices are present in three subunits, for TRAP α TMD there is not a prediction.

- TRAP α prediction: TRAP α protein is 286 residues long (1-21, not included SP), from 64 to 187 luminal residues, mainly, present alpha-helices, close to the ER membrane where the interaction between the TRAP β subunit and the translocon Sec61α1 take place; from 187 to 211 there are beta-sheets, no prediction for TMD.
- TRAP β prediction: TRAP β protein is 183 residues long (1-17, not included SP), alpha-helices are present in the TMD.
- TRAP γ prediction: TRAP γ protein is 185 residues long (no SP), alpha helices are present in the TMDs (four).
- TRAP δ prediction: TRAP δ protein is 173 residues long (1-23, not included SP), alpha-helices are present in the TMD.

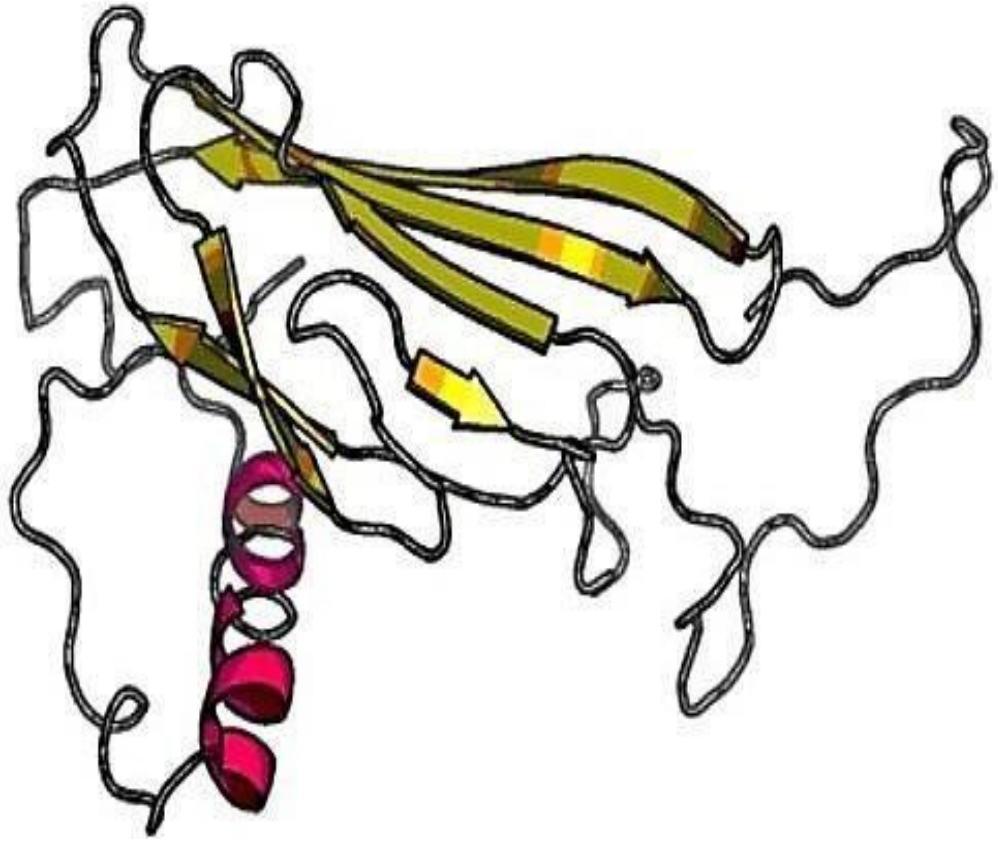


Fig. 3.2.3.4 - RaptorX structure prediction TRAP α: from 64 to 217 luminal residues. The alpha-helices are present from 64 to 187 residue, close to the ER membrane where the interactions with TRAP β and Sec61α1 take place; from 178 to 217 residue mostly are beta-sheets.

Fig. 3.2.3.5 - RaptorX structure prediction TRAP β: from 1 to 118 residue there are beta-sheets, from 119 to 166 there are alpha-helices. The TMD is from 147 to 167 residue.

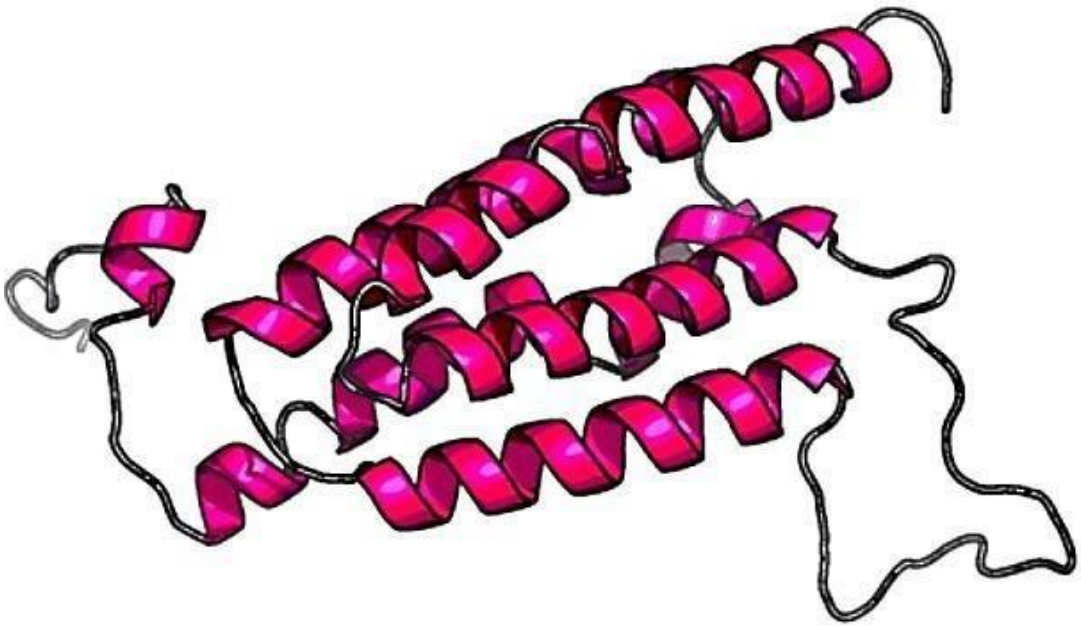


Fig. 3.2.3.6 - RaptorX structure prediction TRAP γ: alpha-helices are present also in the TMDs (four).

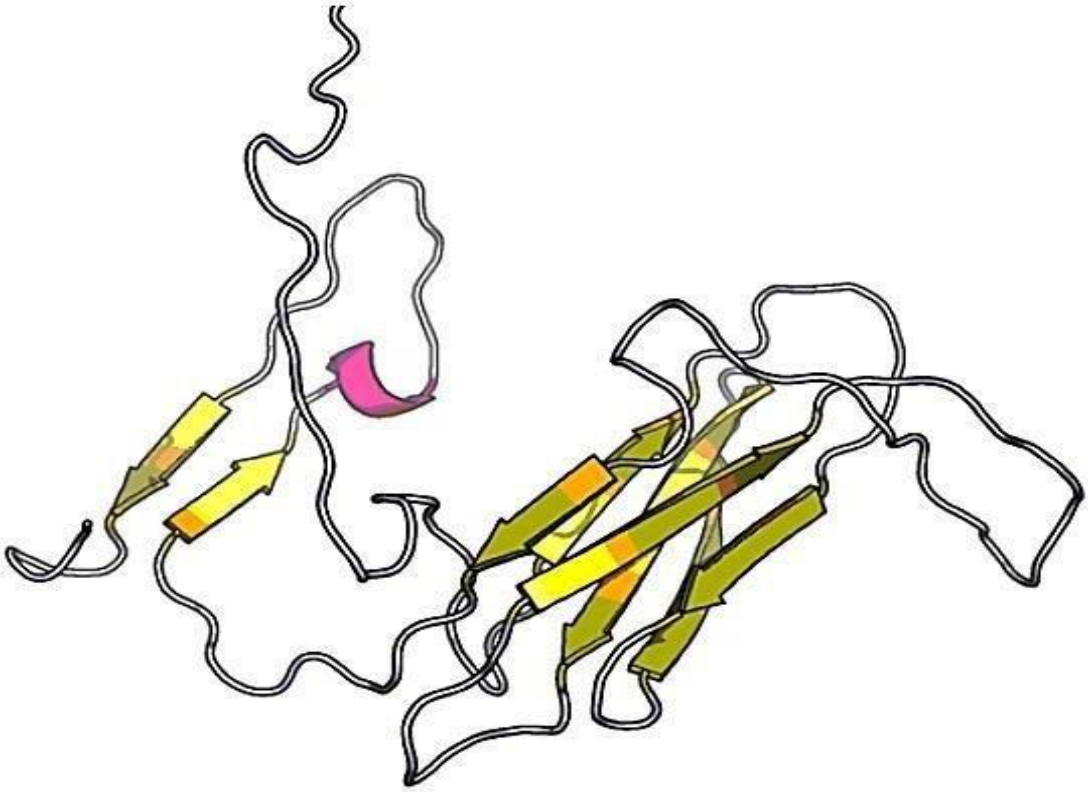


Fig. 3.2.3.7 - RaptorX structure prediction TRAP δ: from 1 to 87 residue/from 109 to 127/from 88 to 108/ are present beta-sheets, and from 128 to 151 there are alpha-helices. The TMD is from 145 to 165 residues.

3.2.4 TRAP beta and TM motif retention

The ER single-spanning membrane proteins type I can present motifs that retain the protein in the membrane, and a typical TM motif is – K(5)X(4)K(3)X(2)X(1) in position -3/-5 at the C-terminus (K=Lysine, X= any residue). This motif is present in TRAP β subunit of *Mus musculus* and Human, but is absent in TRAPα and δ, which are also single-spanning membrane proteins type I (Tab. 3.2.4.1).

```
        10         20         30         40         50
MRLLSFVVLA LFAVTQAEEG ARLLASKSLL NRYAVEGRDL TLQYNIYNVG
        60         70         80         90        100
SSAALDVELS DDSFPPEDFG IVSGMLNVKW DRIAPASNVS HTVVLRPLKA
       110        120        130        140        150
GYFNFTSATI TYLAQEDGPV VIGSTSAPGQ GGILAQREFD RRFSPHFLDW
       160        170        180
AAFGVMTLPS IGIPLLLWYS SKRKYDTPKT KKN
                                5  3
```

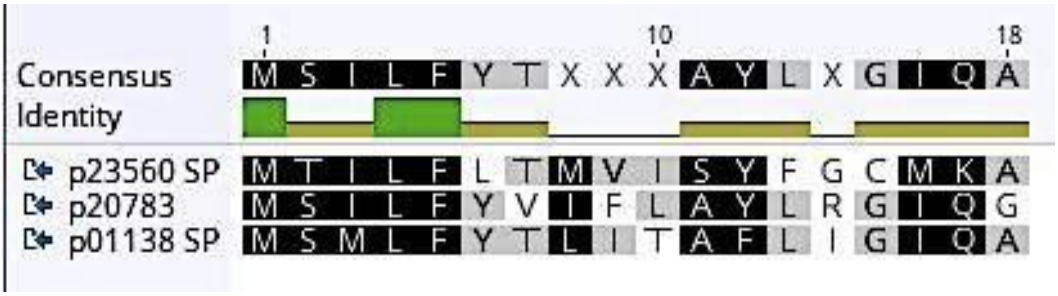
```
        10         20         30         40         50
MRLLAVVVLA LLAVSQAEEG ARLLASKSLL NRYAVEGRDL TLQYNIYNVG
        60         70         80         90        100
SSAALDVELS DDSFPPEDFG IVSGMLNVKW DRIAPASNVS HTVVLRPLKA
       110        120        130        140        150
GYFNFTSATI TYLAQEDGPV VIGSTSAPGQ GGILAQREFD RRFSPHFLDW
       160        170        180
AAFGVMTLPS IGIPLLLWYS SKRKYDTPKP KKN
                                5  3
```

Tab.3.2.4.1 - Retention TM motif – K(5)X(4)K(3)X(2)X(1) - at the C-terminus of TRAP β Human (above) and *Mus Musculus* (down).

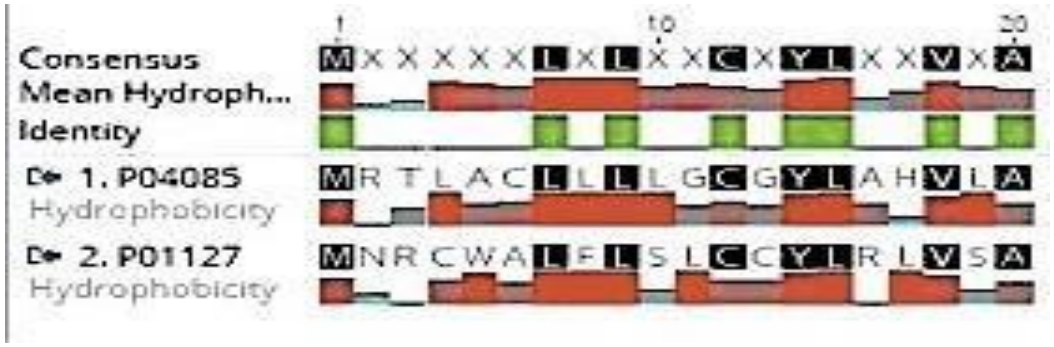
3.2.5 Signal Peptide

The interconnection between the SP and the mature protein is an evolving aspect that deserves further studies. The Geneious analysis points out that some classes of proteins present a significant similarity in their SP sequences. The analysed classes are growth factors (some are involved in placenta development, Introduction p. 20), below some results (Tab. 3.2.5.1)

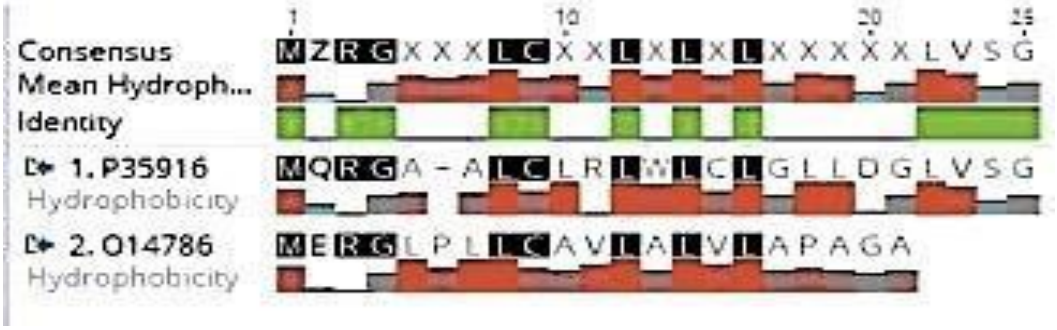
human Neurotrophins



human PDGFs

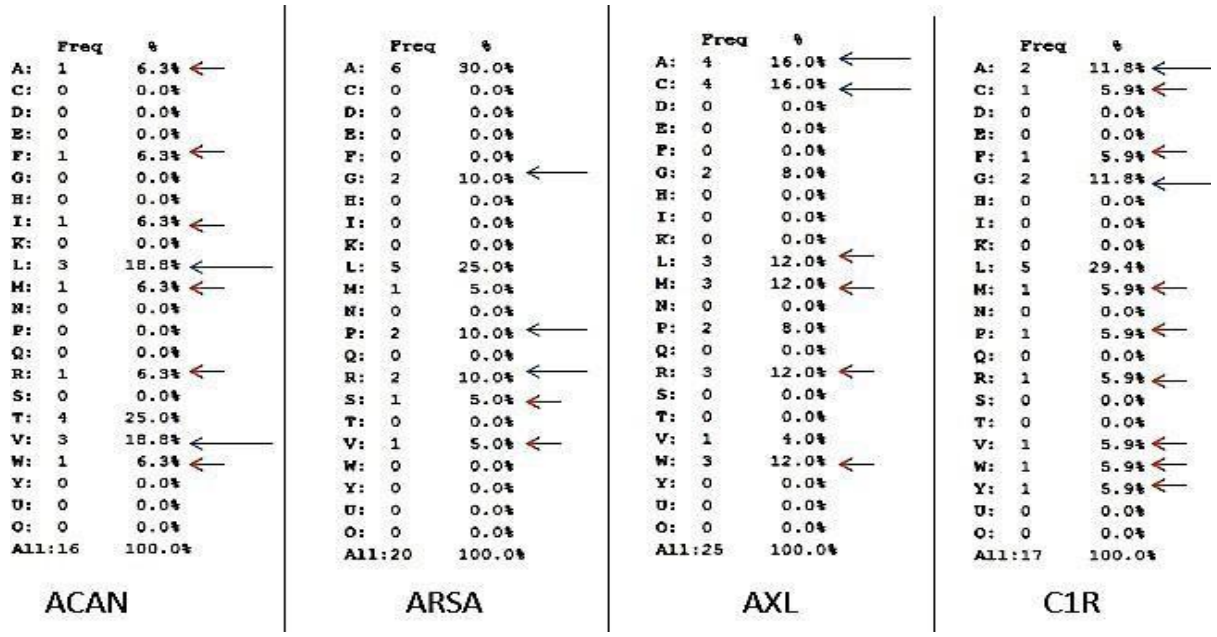


human VEGFs



Tab. 3.2.5.1 - The similarity between SPs of the same class of human proteins: Neurotrophins (P01138, P20783, P23560 UniProt). PDGFs: (Platelet-Derived Growth Factors) (P4085, P01127 UniProt), and VEGFs (P35916, O14786 UniProt).

The analysis of the signal peptide of 119 human secretory proteins from Adipose-Derived Stem Cell (hASC) list has been carried out. An intriguing aspect is that in almost all SP sequences there are two or more residues that are present in the same percentage (same number) (Tab. 3.2.5.2). Probably, these amino acids repeats (AARs) can avoid a bias within the SP and/or with the mature protein sequence.



Tab. 3.2.5.2– Analysis of SP secretory proteins from the hASC list: two or more amino acid residues are present in the same percentage along the SP sequence (red and blue arrows).

4. DISCUSSION

A milestone in the understanding of protein translocation is the "Signal Hypothesis" proposed by Blobel in 1971. Typically, SPs have a tripartite structure but are very heterogenous in peptide sequence composition. SP complexity influences many biological processes. ER membrane components associated with the translocon are essential for translocation. Blobel et al. found in the 1990s that without ER membrane components, the protein precursors move freely into the channel and reach the cytosolic side once more (Nicchitta et al., 1993). Recently, the structure of translocation machinery has been deduced due to improved microscopic techniques, such as Cryo-EM and Cryo-ET. Most of the structures are known, but the roles of some components remain unclear. Among these is the TRAP complex, which is involved in protein translocation, maturation, and degradation. PPIs play fundamental roles in many cellular processes; the identification of binding partners is essential for the analysis of protein functions. We aimed to determine the interactions of TRAP subunits within the complex and with the surrounding structures.

4.1 Recombinant proteins

We attempted to obtain GST and HIS tagged recombinant proteins from TRAP subunits. The cloning design includes luminal and cytosolic domains that are essential to determine the TRAP complex interactions. Two luminal protrusions are present in the ER lumen near Sec61. One is part of the OST complex and the other represents the luminal domain of TRAP α and β subunits. Microscopy analyses have established that loop 5 of Sec61α1 is close to these two TRAP subunits (Pfeffer et al., 2015, 2017). The recombinant proteins of TRAP α and β include the luminal domains, and these domains allow the study of the interaction with the translocon and interactions within the TRAP complex. The domain of TRAP δ recombinant protein is also luminal; the silencing of this subunit destroys the entire complex. It may be beneficial for future experiments to address these interactions. The domain of the TRAP γ recombinant protein is cytosolic and close to the ribosomal protein rpL38. Therefore, an interaction between them is plausible. Microscopy studies have already been undertaken to assess this possibility (Pfeffer et al., 2015; 2017), and one study in particular identified the TRAP γ subunit among the proteins isolated with the mammalian ribosome (Simsek et al., 2017).
We successfully cloned these domains for three subunits—TRAP β, γ, and δ—but the cloning for TRAP α was inconclusive. The constructs TRAPα:: GST and TRAPα:: HIS were synthesized by a company which dealt also with complications. The sequencing of the three clones TRAP β, γ, and δ, matched entirely (100% identity), demonstrating high amplification fidelity with the proper junctions of fused fragments. Therefore, it was determined that the cloning design for these three subunits was appropriate. The extended sequence and high content of guanine/cytosine (G/C) made TRAP α cloning impossible.
Escherichia coli is one of the best hosts for protein expression due to its long experience and easy manipulation and genetic modification. Nevertheless, transmembrane domain purification is challenging due to the use of harsh conditions and strong detergents, which can compromise the structure of the proteins. The TMDs in all four TRAP subunit domains are not present in the cloning design. Protein expression is singular and every protein poses a new problem. Expression can be affected by numerous modifications, such as aggregation, misfolding, random disulphide bridges, and proteolytic cleavage. Therefore, it is necessary to determine the most appropriate solutions. In this study, we faced some challenges, such as low yield and degradation by the host (or instability). By using different media (LB/TB/ 2YT), changing the temperature/growing time/cell density, and IPTG induction, it was possible to optimise protein expression. Nonetheless, we achieved the greatest success by changing the bacteria strain. The "BL21 (DE3)" strain for TRAP α recombinant proteins was found to not be appropriate due to steady degradation (or instability) and incomplete protein synthesis. When

the “BL21 RosettaStar” strain was used, the appropriate size and a high yield was achieved for both TRAP α proteins, GST and HIS tagged. The BL21(DE3) strain enhances eukaryotic protein expression via codons that are rarely present in bacteria, but this was not sufficient. BL21 RosettaStar cells have a gene mutation that reduces mRNA degradation, which likely contributed to the successful results. As the starting point, the BL21 (DE3) strain was used for TRAP β/ γ/ δ recombinant proteins; β and δ showed a low yield. Therefore, the samples were then concentrated using centrifugal filters, but by using the BL21 RosettaStar strain a higher yield was achieved. The highest yield achieved was for TRAP γ with BL21 (DE3) cells. This protein consists of 57 amino acid residues (plus GST tag), making it the shortest of the studied proteins. Therefore, it was more successfully expressed and purified than the others. The concentration was approximately 12 mg/1 L culture, which permitted the performing of X-ray crystallography (my lab and the University of Alberta, Canada). It was concluded that choosing the appropriate bacterial strain is fundamental for correct protein expression and yield; more so than the growing conditions, such as medium composition, incubation time/temperature, or induction. The purified proteins had a high degree of purity, which is essential to a variety of assays, such as pull-down or peptide arrays where non-specific binding can compromise the results. Furthermore, to achieve a higher degree of protein purity, it is advisable to separate the eluates and select them based on SDS-PAGE results, rather than relying purely on nanodrop spectrophotometer measurements. Additionally, this analysis makes it possible to select the eluate with the highest concentration. The recombinant proteins also presented a strong degree of solubility, which is an essential propriety for in vitro experiments and enhances protein purification. The quantitative prediction of protein solubility depends on its condensate form, which is different for each solvent. Some conditions are determinant, such as pH, which affects protonation and deprotonation. More precisely, solubility is directly proportional to the free transfer of energy (*Gsol – Gcon*); wherein a higher value indicates more protein conversion from the condensate form to a soluble one. The equation is “S *(free energy) = exp (Gsol – Gcond) /KBT” (KBT = thermal energy*) (Tjong et al., 2008). In the present study, the proteins during the centrifugation-solubility test were in PBS-KMT buffer, which did not affect the soluble form.
Some GST (HIS) fusion proteins are inclined to degrade, which was the case for the recombinant proteins in this study. Other bands were visible on the SDS-PAGEs in addition to the full length of the tagged protein.
Before performing specific assays with purified proteins, they were checked at different levels. Western blot analysis with GST and HIS antibodies confirmed the correct size of the tagged proteins, and the MS analysis identified the protein sequences. TRAP γ was found to not be appropriate for MS, as enzymatic digestion with trypsin led to small peptides unsuitable for LC-MS. Another digestion method may be adopted, or the use of new methods, such as mechanic dissociation.

4.2 TRAP alpha and beta interaction

PPI is an essential step for protein-function determination; as it provides a large amount of information. The PPIs can be analysed at different levels; kinetics/thermodynamics, structure, and expression. Protein interactions are made up of hydrophobic effects, hydrogen bonds, covalent bonds, and electrostatic interactions. TRAP α and β are subunits of a tetrameric complex. We investigated whether they physically interact by performing GST and HIS pull-down assays, one of the most common in vitro methods for studying PPIs due to its effectiveness caused by its high specificity of GST to reduced GSH (Luo et al., 2014) and His to Ni. By using purified proteins instead of subcellular structures some troubleshooting was avoided (e.g. steric issues).
The bait and prey bands were distinctly visible on the SDS-PAGE gel; the latter permitting to resolve proteins that differed by only 1% in electrophoretic mobility. Furthermore, the prey did not bind to GST (or HIS) beads. An interaction was found between TRAP α and β subunits. The interaction seems to be stable and structural as it is required in a multiprotein complex; a static interaction between two luminal protein domains. The stable interactions led to an increase in internal energy (enthalpy) and a reduction

in disorder (entropy). Further studies are necessary to define the binding interface sequences, such as the 3D structure by x-ray crystallography, peptide array analysis, and computer simulation and modelling. The interface binding residues are more conserved than the other protein sequences, and comparisons with different species may also be helpful. Some forces involved in these physical interactions are electrostatic, which are described by Coulomb's law: F = kc q1 q2 /r2, kc = Coulomb constant, q1 and q2 = magnitudes of the charges, and r2 = distance between the charges. This interaction takes place if the charges are opposite and the distance is short enough; specifically, it takes place between amino acid residues with positive and negative charges and specific geometry. The neighbouring residues to the hot spots (binding interfaces) also influence the interaction, the most common being alanine, aspartic acid, glycine, histidine, isoleucine, asparagine, serine, and tyrosine. It is likely that these residues cluster together to form the appropriate environment for the interaction (Ye et al., 2014). The characteristics of PPIs are endogenous and exogenous factors: 1) specificity, the ability of a protein to bind a single partner; 2) promiscuity, a single protein carries out different functions; 3) selectivity, the protein uses other proteins for binding; and 4) affinity, the strength of the interaction.

The interaction between TRAP α and β is not unique, the same proteins interact with the Sec61α1 subunit, but it is likely that there are not many partners. The number of functions that the TRAP complex performs is currently unknown, and other proteins may be involved in the binding. The affinity seems high because the pull-down assays were not under restricted conditions (pH, concentration, or temperature), yet the same results were achieved for every experiment. The molecular dissociation constant, Kd, establishes the interaction affinity for a general reaction: Ax By ←→xA+ yB, Kd = [A]x [B]y / [Ax By]. Kd is the ratio between the dissociated and interacting states, and the smaller the value, the more the protein interacts. Further studies could investigate this further, and many appropriate methods are available, such as Förster resonance energy transfer (FRET) assays and quantitative mass spectrometry. The former is a quantitative technology based on fluorescence emission. The quantitative MS uses isotopes for labelling and the proteins are precisely quantified. STRING, an essential consortium for PPI prediction based on many resources, predicted an interaction between TRAP α and β. This supports our results along with previous cryo-ET studies that report a short physical distance between TRAP α and β subunits, both inserted in the ER membrane and with luminal domains (Pfeffer et al., 2015, 2017).

4.3 TRAP α/β and Sec61α1 loop5 interactions

The PPIs play different roles beyond complex formation. They allow substrate channelling, the formation of new binding sites for other effector molecules, and the changing of protein specificity for its substrate. All these aspects can be suggested for the interactions between the TRAP complex and the translocon Sec61 detected by "peptide array". Peptides are a selective approach to studying PPIs, and it presents different advantages: i) it focuses on specific binding sites; ii) no secondary structures until binding; and iii) peptides can be selected, mutated, and easily synthesised. The "peptide array" is a qualitative and quantitative technique that uses small peptides generated by SPOT technology (or other methods) on a substrate. It permits the investigation of different processes, such as peptide-metal interaction, peptide-nucleic acid-binding, peptide enzymatic modification, and PPI. This assay is a popular and powerful tool to study PPI as it focuses on a specific interaction site, detects interactions with several proteins, and establishes the binding interface sequence. Moreover, the mutational analysis allows the study of the effect of some mutations, which is useful for medical purposes. By changing the protocol, it is possible to improve the results; for instance, by using different blocking buffers (milk or sucrose instead of BSA), increasing buffer concentrations, using different antibodies, and using different detection procedures, such as chemiluminescence, fluorescence, and electrochemiluminescence (Amartely et al., 2014). The assay is more powerful when a complex, such as Sec61, is involved and the steric aspect is relevant. We aimed to determine whether TRAP α and β

subunits interacted with the Sec61α1 loop 5; the sequence of loop 5 is 50 residues, making it more convenient than analysing the entire Sec61α1 subunit of 476 amino acid residues. The results of this study indicate that loop 5 interacts with ssr1 and ssr2. Previous cryo-EM/cryo-ET studies had hypothesised that the translocon is physically close to these TRAP subunits in the ER luminal side (Pfeffer et al., 2017). Sec61, TRAPα, and TRAP β were found to be in a stoichiometric ratio of 1:1:1; it was demonstrated by quantitative MS by Menetret et al. in 2008. Moreover, the STRING server predicted an interaction between the translocon Sec61α1 and the two TRAP subunits. Furthermore, the TRAPα N-term protein sequence is very conserved among different species, which confirms the relevance of this domain for its interactions. TRAP α is a phosphorylated protein, and phosphorylation plays a significant role during PPIs. Interestingly, proteins typically interact with the partner a limited amount, and because the majority of proteins are complexes (four-fifths in eukaryotes), there are often two or more subunits involved (Raghavachari et al., 2008). Moreover, the PPIs often occur near a cellular membrane, which is the case for interactions between TRAPα / β and Sec61α1. These interactions between the translocon and TRAP subunits could be transitory, defined by a specific function. It is likely that a regulated-switchable binding leads to different conformation, or as a result of different conformations. When co-translational translocation takes place, the Sec61 can recruit the TRAP complex. It is possible that at this stage the channel and TRAP subunits interact. This event could be related to the dynamic properties of Sec61, stabilisation of the open channel state, or the increasing of LG mobility. Different methods stabilise a channel and, for instance, some voltage channels are maintained in an open state by cations that occupy the inner cavity and avoid the closure (Goodchild et al., 2012). During co-translational protein transport, the precursor polypeptides trigger the opening of the Sec61 channel by targeting the ER membrane, which is achieved via GTP hydrolysis. This interaction and ribosome interactions displace the plug inside the channel, but the next steps are unknown.

When the interaction is transient, several amino acid residues are involved. The short linear motif (SLIM) is a conserved sequence that interacts with globular domains. Typically, the proteins that interact transiently undergo conformational changes and state transition (order, disorder). Bioinformatic tools are useful in determining the SLIMs. However, this is not straightforward because of the short length of these sequences (3-20 residues), and because they are rarely conserved among different proteins (Neduva and Russel, 2005). The binding free energy (ΔΔG) of some residues can facilitate sequence identification. Nonetheless, the residues of the binding interface cannot be determined exclusively from these features; the geometry of the molecular surface and its 3D structure is also necessary. That the interactions between the TRAP subunits and Sec61α1 are stable cannot be excluded. TRAP as well as OST, another component that interacts with the translocon, are also observed after translation, even in the absence of ribosomes (Conti et al., 2015; Shibatani et al., 2005; Snapp et al., 2004).

Finally, the black dots of the TRAPα:: GST sample did not entirely overlap with the TRAPα:: HIS sample or TRAPα:: GST/TRAPα:: HIS sample. This is likely due to the different structure and length of these two tagged proteins; different geometry can lead to interactions with different overlapping spots. It was also not possible to establish the binding interface sequence of Sec61α1 loop 5 with TRAPβ:: GST, the GST tag could affect the results. Further studies are needed, for instance, for mutational analysis, as it is possible that the substitution of residues can determine the binding sequences. Black dots are also present where there are the mutated loop5 spots, but are not present on the parallel wild type loop5 spots. These results may be explained by the fact that alanine substituted threonine in the TRAPα:: HIS sample (10) (Fig. 3.1.4.3), and glycine in the TRAP β::GST sample (18) (Fig. 3.1.4.4), leading to artificial bindings. The neighbouring amino acid residues influence the interaction, and, hence, the interactions on the mutated spots occurred. The amino acid residues that are exchanged with alanine and acquire binding are considered key residues to study PPIs (Volkmer and Tapia, 2012).

4.4 TRAP complex functions

Early studies reported the TRAP complex as an unnecessary structure for translocation. However, it is now known that TRAP is a substrate-specific element of the mammalian translocon machinery. Not all substrates are TRAP-dependent, which may be due to the different features of the signal SP and the mature protein of the substrate. A firm or weak perception of the SP and mature protein signals by the translocon can determine which substrates depend on TRAP and which are independent. TRAP is essential for some substrates that have a weak SP (Fons et al., 2013), and some clients of TRAP have a high glycine and proline content (Nyuyen et al., 2018), which is connected with the secondary structure. These residues present to the border of SP h- and c-regions, contributing to the formation of β-barrel. TRAP can interact directly with the Sec61 channel to compensate interaction weakness with the substrate and maintain an open conformation or influence its dynamic. After interaction with ribosomes, it is plausible that different nascent proteins lead to different Sec61 conformations (Voorhees et al., 2014). TRAM, another accessory component of translocation machinery, shows substrate dependence during co-translation translocation. TRAP and TRAM may have similar functions, and may remain next to Sec61 until complete translocation. They may drive the movement of the chain along the Sec61 channel after the initial force made by the ribosomes and GTP hydrolysis. TRAP, similar to TRAM, may also function as a chaperone and carry out a storage step until substrate maturation. Post-translational modifications, such as adding glycans (hydrophilic polymers), phosphorylation (negative charges), and disulphide bridges (covalent bonds), lead to greater solubility, thermal stability, and folding. This is connected with the "translocation pausing" required for the reactions of protein biogenesis. The TRAP complex and substrate crosslinking has been detected in the late stage of translocation. Instead, TRAM seems to interact with the NH2-terminal region (Gorlich et al., 1993; Oliver et al., 1995). This may explain the conspicuous TRAP luminal domain under the channel observed by Menetret et al. in 2005. Specifically, TRAP directly interacted with the substrate to facilitate the translocation and/or maintain its orientation/structure, or with the translocon. These coordinated roles are known for BiP chaperone; the opening of Sec61 (Dierks et al., 1996), the closure of Sec61 channel (Alder et al., 2005), and binding to the nascent polypeptides in transit to complete translocation (Nichitta and Blobel, 1993; Tyedmers et al., 2003; Shaffer et al., 2005). The opening of the channel by BiP is due to nucleotide exchange, while the closure by BiP depends on direct interaction with Sec61α1 loop7 (Schäuble et al., 2012); a function that is also important to avoid calcium leakage (Simon and Blobel, 1991).

Calnexin and TRAP α appear to be calcium-binding ER membrane proteins, while calreticulin is an ER lumen calcium-binding protein. Therefore, TRAP α may have a calcium-binding role in the interaction with the complex Sec61. Additionally, it is possible that the EF-hand motif has a functional role rather than structural. TRAP α calcium-binding affinity also needs to be elucidated. TRAP can undergo different conformation that influences its interactions, such as for calreticulin, by calcium-binding. It is likely that TRAP α binds calcium on the luminal side, where it interacts with Sec61α. This cation can lead to changes in the TRAP α structure, and interaction with the translocon. Previous studies report that the binding of calcium by the C-terminus EF-hand domain of Se62 leads to the dissociation from its interacting partners, such as Sec61 (Ampofo et al., 2013; Linxweiler et al., 2013). Additionally, TRAP may be involved in the topology of TM proteins. These proteins require correct orientation when leaving the LG to be accommodated in the membrane lipid bilayer. Previous investigations demonstrated that the rapid folding of the N-terminus sequence in TM proteins before the signal-anchor sequence restrains translocation (Denzer et al., 1995; Spiess et al., 2019). It is possible that Sec61 is sufficient to translocate TM proteins with a cleavable SP type I (luminal N-terminus) but not TM proteins with a signal-anchor (Oliver et al., 1995). It is possible that TRAP plays a role in these situations. Crosslinking experiments indicated that TRAM is involved in viral TM protein integration into the ER. First, each segment of the chain is associated with Sec61α, then with TRAM when it is

about 100 residues long (Sauri et al., 2007).

Whether the TRAP complex can carry out more than one function remains to be elucidated. Some findings suggest a role of the TRAP complex in the UPR pathway and cellular equilibrium:

- TRAP interacts with some unfolded substrates but not with the wild-type form;
- TRAP induction under ER stress by the IRE1α pathway;
- TRAP indcuction under GM-GSF stimulation, a factor that leads to the transcription of many genes, UPR, and ERAD (Hirama et al., 1999).

Some proteins, such as calreticulin, calnexin, and BiP, have a role in folding and quality control. It is possible that this could also be the case for TRAP.

Some studies have suggested that Hrd1 and Hrd3 retro-translocate abnormal proteins after ubiquitination (Schoebel eta al., 2017; Jarosch et al., 2002). Nonetheless, the translocon Sec61 could retro-translocate the proteins that undergo degradation. Indeed, it interacts with ERAD substrates and the proteoasome. The retro-translocon Sec61 may require support form the luminal side. The over-expression of TRAP subunits during ERAD could be connected with these processes. Which associated components push the substrate through the channel is currently unknown. It is also not known whether ubiquitination is sufficient. Additionally, the role of TRAP δ plays a role in the congenital disorder of glycosylation (ssr4 CDG). The complex interacts with some OST subunits, and the lack of this cooperation leads to OST dysfunctionality. These interactions may modify the OST kinetic properties; indeed, STRING predicts an interaction between TRAP delta and DDOST subunit of OST. TRAP may maintain the newly synthesized chain in al linear structure to permit N-glycosylation. Another hypothesis is the OST is a TRAP client, and its synthesis is compromised. Overall, whether TRAP plays a direct or secondary role in the glycosylation disorder is not currently known, Plants and fungi lack TRAP γ and δ subunits, yet have a coordinated complex. Although extensive research has been conducted, it is not currently possible to form any conclusion regarding the role of the TRAP complex in different tissues. The knockout of TRAP subunits in different tissues and organ leads to different consequences. Each isoform of TRAP subunits can play a different role in different tissues, or the knockout of one subunit can compromise the entire complex.

4.5 Overview and future prospective

This study contributes to the current understanding of TRAP complex functions during co-translational protein transport. The identification of molecular interactions progresses the understanding of cellular processes. The structure of the TRAP complex suggests that the interaction
of TRAP α/β is not unique, as other PPIs likely are present within the complex. The subunit β is very close to the δ subunit, and TRAP α knockout showed β and δ under-expression (Sommer et al., 2013).
While the interaction between TRAP α and β was plausible, they are subunits of a complex. The interaction between TRAP α/β subunits with Sec61 α1 is a more relevant finding. Sec61 is a channel with different conformations and states. The modern resolutive methods make the analysis of channels a promising investigation. The ER co-translational protein translocation relies on general structures: targeting signals, membrane receptors, transmembrane channels, and accessory components. It is not currently known when some accessory components are necessary and the channel is insufficient; the functions of these components require further study. The field limitations are the analysis of subcellular structures during their function. Additionally, separate components from cell fractions require good separation, representation, and conditions (Nichitta and Blobel, 1990).
Undoubtedly, methods, such as cryo-EM/ET, are appropriate for structural analysis in entire cells or lysates, and they have been extensively used to study the TRAP complex (Pfeffer et al., 2017).

However, the assemble of the subsequent snapshots to describe the entire biological mechanism is a major disadvantage. The processes are rapid and consist of real dynamics; for instance, the configuration between the RTC and nascent polypeptide changes over time. Therefore, it is necessary to overcome these weaknesses. Indeed, studies have established TRAP as a cellular component, have determined its structure, and have identified some interacting partners. However, its function at the molecular level and its biological processes are currently unclear.

Future studies should employ traditional approaches, such as following the protein translocation into the ER by, for instance, perceiving the substrate N-glycosylation detectable on SDS-page (different molecular weights). However, new approaches are necessary, such as microarray assays carried out under TRAP siRNA and in different tissues. Finally, analysing the SP of TRAP clients and mature protein, an approach that has already being employed (Nugyen et al., 2018).

At present, it seems that some roles of TRAP are redundant with other components, such as BiP, TRAM, and calnexin. The exact contribution of TRAP is currently unknown. Whether TRAP have an essential role in the clustering and integration of TM proteins requires further investigation. TRAP substrate-dependence may be more connected with the secondary structure than hydrophobic domains. Whether TRAP interacts with the substrate or the translocon to carry out its function is not currently known. TRAP may be needed when Sec61 and TRAM cannot complete substrate translocation. TRAP could not recognise the SP characteristics but could recognise some mature protein features. Indeed, TRAP interacts with the substrates only when they are of a certain length. Preliminary studies are already focusing on the possibility that TRAP interacts with the mature protein rather the SP.

The similarity between SPs of some growth factors (Results) reinforces the hypothesis that the SP is specialised for its substrate. Adding an SP to a mature protein does not always result in translocation taking place. The mature proteins hold additional information that is essential for the translocation and unfolding state (Orfanoudaki et al., 2017). Signals in the mature regions influence the translocation in mitochondria (Backes et al., 2018; Yamamoto et al., 2009) and bacteria (Kajava et al., 2000).

TRAP α has two isoforms and the cardiac/skeletal muscle form is crucial for mouse viability. This implies that many proteins expressed during the development of these tissues rely on TRAP. A future strategy could establish the TRAP clients by following the entire development by microarray or immunoblotting analysis. Furthermore, measuring the mRNA maturation when the gene expression increases may make more splicing forms detectable. By comparing mRNA and protein expression in different tissues at different stages, some TRAP clients necessary during development may be identified.

The functions of TRAP α calcium-binding is not currently known. The presence of the non-canonical EF-hand domain potentially confirms that this subunit binds Ca^{2+} as the domain is present in the luminal N-terminus. Calcium binding could change the TRAP conformation allowing interaction with the neighbouring structures or merely increase rigidity, which is essential during a stable physical interaction. This is another line of study worth investigating.

Further analysis is also necessary to determine whether some interactions take place between TRAP and OST complexes. Reduced glycosylation in a congenital disorder indicates the absence of TRAP δ and reduced expression of other TRAP subunits. These results lead to some conclusions: lack of TRAP complex stability/function and /or a lack of interaction between TRAP and OST; indeed, OST is isolated with Sec61 and TRAP (Shibatami et al., 2005).

It is not currently known why TRAP subunits are overexpressed during ER stress. A future study could investigate this by monitoring the response of TRAP genes under different stress conditions and comparing stressed and unstressed cells (microarrays). Concurrently, the expression of the other genes under the same conditions could be measured.

OST is present in about 50% of isolated ribosome-associated membrane proteins (RAMPs). During glycosylation, OST may acquire different morphology and different interactions with RAMPs. Further studies regarding this topic could shed more light on the functions of TRAP and other structures.

TRAM is essential for some TM proteins. The TRAP complex knockout also compromises the translocation of TM proteins. A future study could list the clients for both structures and compare them.

The study of proteins is a fast-evolving and interdisciplinary field. Some aspects to consider are intracellular localisation, structure, sequence, evolution, motifs, post-translational modifications, and

interactions with proteinss, DNA, and RNA. In addition, it is necessary to take into account expression profiles, isoforms, and tissue-specific expression. The expression of TRAP genes and their isoforms is important to consider. Determining the expression variation in different tissues and organs and establish the isoforms involved is an important topic for further study.

TRAP subunits are transmembrane proteins, and, like other TM proteins, they represent a connection between two different environments (Sjöstrand et al., 2017). In the case of the TRAP complex, some PPIs take place into the ER lumen, and others may present in the cytosol, for instance, with the ribosomal protein rpL38; interaction that may be structural as well as functional. Therefore, a ribosome affinity assay is appropriate to address this aspect, it permits the detection of binding of a single protein with precipitated ribosomes.

In summary, although some studies have been carried out regarding the TRAP complex, no single study exists that adequately addresses its role inside the ER and during protein translocation. Further integration of many uncoordinated and divergent studies is necessary, including the results of the present study. This integration could establish the molecular functions and biological processes beyond the knowledge of the cellular components and structure.

(Elsevier company has edited the language in the following sections: abstract, introduction, and discussion).

Abbreviations

3D = 3 dimensional
AARs = amino acid repeats
ALG = asparagine-linked glycosylation
AMPK = AMP-activated protein kinase
ATP = Adenosine triphosphate
BiP = binding immunoglobulin-heavy-chain-protein
BSA = bovine serum albumin
CD = calcium binding
CDG = congenital disorders of glycosylation
coIP = co-immunoprecipitation
COPII = coat protein complex II
Cryo-EM = cryoelectron microscopy
Cryo-ET = cryoelectron tomography DDT = Dithiothreitol
ER = endoplasmatic reticulum
ERAD = endoplasmatic-reticulum (ER)-associated degradation
FRET = Förster resonance energy transfer
GC = guanine/cytosine
GM-CSF = granulocyte-macrophage colony-stimulating factor
GTP = Guanosine-5'-triphosphate
hASC = Adipose-Derived Stem Cell HRP = horseradish peroxidase
H-segments = hydrophobic segments
IPTG = isopropyl β-D-1-thiogalactopyranoside
IRE1α = 1/inositol-requiring 1α
Kd = dissociation costant LB = Luria-Bertani medium
LC- MS = Liquid Chromatography-Mass Spectrometry
LD = luminal domain
LG = lateral gate
MCS = multiple cloning site
MWCO = molecular weight cut-off
OD = optical density
OST = oligosaccharyltransferase complex
PCC = protein conducting channel
PLAC = Protease Inhibitor Cocktail
PPI = protein-protein interaction
PTM = post-translocation modifications
qPCR = quantitative PCR
RAP = ribosome-associated protein
RCC = ribosome-channel complex
RNC = ribosome-nascent-chain complex
RTC = ribosome-translocon complex
SDS-PAGE = sodium dodecyl sulfate polyacrylamide gel electrophoresis
SERCA = sarcoplasmic-endoplasmic reticulum Ca2+-ATPase
SP = signal peptide
SPase = signal peptidase
SR = signal receptor
SRP = signal recognition particle
SS = signal sequence
TB = terrific broth

TM = transmembrane
TMD = transmembrane domain
TMH = transmembrane helix
TRAM = translocating chain-associating membrane protein
TRAP = translocon-associated protein complex
UPR = unfolded protein response
UPS = ubiquitin-proteasome system

References

Alder, N. N., Shen, Y., Brodsky, J. L., Hendershot, L. M., & Johnson, A. E. (2005). The molecular mechanisms underlying BiP-mediated gating of the Sec61 translocon of the endoplasmic reticulum. *The Journal of Cell Biology*,*168*(3),389–399.doi:10.1083/jcb.200409174

Amartely, H., Iosub-Amir, A., & Friedler, A. (2014). Identifying Protein-protein Interaction Sites Using Peptide Arrays. *Journal of Visualized Experiments*, (93). doi: 10.3791/52097

Ampofo, E., Welker, S., Jung, M., Müller, L., Greiner, M., Zimmermann, R., & Montenarh, M. (2013). CK2 phosphorylation of human Sec63 regulates its interaction with Sec62. *Biochimica Et Biophysica Acta (BBA) - General Subjects*, *1830*(4), 2938–2945. doi: 10.1016/j.bbagen.2012.12.020

Araki, K., & Nagata, K. (2011). Protein Folding and Quality Control in the ER. *Cold Spring Harbor Perspectives in Biology*, *3*(11). doi: 10.1101/cshperspect.a007526

Auclair, S. M., Bhanu, M. K., & Kendall, D. A. (2011). Signal peptidase I: Cleaving the way to mature proteins. *Protein Science*, *21*(1), 13–25. doi: 10.1002/pro.757

Backes, S., Hess, S., Boos, F., Woellhaf, M. W., Gödel, S., Jung, M., … Herrmann, J. M. (2018). Tom70 enhances mitochondrial preprotein import efficiency by binding to internal targeting sequences. *The Journal of Cell Biology*, *217*(4), 1369–1382. doi: 10.1083/jcb.201708044

Bai, L., Wang, T., Zhao, G., Kovach, A., & Li, H. (2018). The atomic structure of a eukaryotic oligosaccharyltransferase complex. *Nature*, *555*(7696), 328–333. doi: 10.1038/nature25755

Bañó-Polo, M., Martínez-Garay, C. A., Grau, B., Martínez-Gil, L., & Mingarro, I. (2017). Membrane insertion and topology of the translocon-associated protein (TRAP) gamma subunit. *Biochimica Et Biophysica Acta (BBA) - Biomembranes*, *1859*(5), 903–909. doi: 10.1016/j.bbamem.2017.01.027

Berg, B. V. D., Clemons, W. M., Collinson, I., Modis, Y., Hartmann, E., Harrison, S. C., & Rapoport, T. A. (2003). X-ray structure of a protein-conducting channel. *Nature*, *427*(6969), 36–44. doi: 10.1038/nature02218

Blau, M., Mullapudi, S., Becker, T., Dudek, J., Zimmermann, R., Penczek, P. A., & Beckmann, R. (2005). ERj1p uses a universal ribosomal adaptor site to coordinate the 80S ribosome at the membrane. *Nature Structural & Molecular Biology*, *12*(11), 1015–1016. doi: 10.1038/nsmb998

Blobel, G. (1975). Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. *The Journal of Cell Biology*, *67*(3), 835–851. doi: 10.1083/jcb.67.3.835

Braakman, I., & Hebert, D. N. (2013). Protein Folding in the Endoplasmic Reticulum. *Cold Spring Harbor Perspectives in Biology*, *5*(5). doi: 10.1101/cshperspect.a013201

Buske, F. A., Maetschke, S., & Bodén, M. (2009). It's about time: Signal recognition in staged models of protein translocation. *Pattern Recognition*, *42*(4), 567–574. doi: 10.1016/j.patcog.2008.09.020

Cabrita, L. D., Cassaignau, A. M. E., Launay, H. M. M., Waudby, C. A., Wlodarski, T., Camilloni, C., Christodoulou, J. (2016). A structural ensemble of a ribosome–nascent chain complex during cotranslational protein folding. *Nature Structural & Molecular Biology*, *23*(4), 278–285. doi: 10.1038/nsmb.3182

Charneski, C. A., & Hurst, L. D. (2013). Positive Charge Loading at Protein Termini Is Due to Membrane Protein Topology, Not a Translational Ramp. *Molecular Biology and Evolution*, *31*(1), 70–84. doi: 10.1093/molbev/mst169

Chavan, M., Yan, A., & Lennarz, W. J. (2005). Subunits of the Translocon Interact with Components of the OligosaccharylTransferase Complex. *Journal of Biological Chemistry*, *280*(24), 22917–22924. doi: 10.1074/jbc.m502858200

Coe, H., Bedard, K., Groenendyk, J., Jung, J., & Michalak, M. (2008). Endoplasmic reticulum stress in the absence of calnexin. *Cell Stress and Chaperones*, *13*(4), 497–507. doi: 10.1007/s12192-008-0049-x

Conti, B. J., Devaraneni, P. K., Yang, Z., David, L. L., & Skach, W. R. (2015). Cotranslational Stabilization of Sec62/63 within the ER Sec61 Translocon Is Controlled by Distinct Substrate-Driven Translocation Events. *Molecular Cell*, *58*(2), 269–283. doi: 10.1016/j.molcel.2015.02.018

Cymer, F., Heijne, G. V., & White, S. H. (2015). Mechanisms of Integral Membrane Protein Insertion and Folding. *Journal of Molecular Biology*, *427*(5), 999–1022. doi: 10.1016/j.jmb.2014.09.014

Demartino, G. N., & Gillette, T. G. (2007). Proteasomes: Machines for All Reasons. *Cell*, *129*(4), 659–662. doi: 10.1016/j.cell.2007.05.007

Denessiouk, K., Permyakov, S., Denesyuk, A., Permyakov, E., & Johnson, M. S. (2014). Two Structural Motifs within Canonical EF-Hand Calcium-Binding Domains Identify Five Different Classes of Calcium Buffers and Sensors. *PLoS ONE*, *9*(10). doi: 10.1371/journal.pone.0109287

Denzer, A. J., Nabholz, C. E., & Spiess, M. (1995). Transmembrane orientation of signal-anchor proteins is affected by the folding state but not the size of the N-terminal domain. *The EMBO Journal*, *14*(24), 6311– 6317. doi: 10.1002/j.1460-2075.1995.tb00321.x

Dierks, T., Volkmer, J., Schlenstedt, G., Jung, C., Sandholzer, U., Zachmann, K., … Zimmermann, R. (1996). A microsomal ATP-binding protein involved in efficient protein transport into the mammalian endoplasmic reticulum. *The EMBO Journal*, *15*(24), 6931–6942. doi: 10.1002/j.1460-2075.1996.tb01085.x

Dudek, J., Pfeffer, S., Lee, P.-H., Jung, M., Cavalié, A., Helms, V., … Zimmermann, R. (2015). Protein Transport into the Human Endoplasmic Reticulum. *Journal of Molecular Biology*, *427*(6), 1159–1175. doi: 10.1016/j.jmb.2014.06.011

Dumax-Vorzet, A., Roboti, P., & High, S. (2013). OST4 is a subunit of the mammalian oligosaccharyltransferase required for efficient N-glycosylation. *Journal of Cell Science*, *126*(12), 2595– 2606. doi: 10.1242/jcs.115410

Egea, P. F., & Stroud, R. M. (2010). Lateral opening of a translocon upon entry of protein suggests the mechanism of insertion into membranes. *Proceedings of the National Academy of Sciences*, *107*(40), 17182–17187. doi: 10.1073/pnas.1012556107

Ellgaard, L., & Helenius, A. (2003). Quality control in the endoplasmic reticulum. *Nature Reviews Molecular Cell Biology*, *4*(3), 181–191. doi: 10.1038/nrm1052

Elofsson, A., & Heijne, G. V. (2007). Membrane Protein Structure: Prediction versus Reality. *Annual Review of Biochemistry*, *76*(1), 125–140. doi: 10.1146/annurev.biochem.76.052705.163539

Fecher-Trost, C., Wissenbach, U., Beck, A., Schalkowsky, P., Stoerger, C., Doerr, J., … Flockerzi, V. (2013). Thein VivoTRPV6 Protein Starts at a Non-AUG Triplet, Decoded as Methionine, Upstream of Canonical Initiation at AUG. *Journal of Biological Chemistry*, *288*(23), 16629–16644. doi: 10.1074/jbc.m113.469726

Fink, A., Sal-Man, N., Gerber, D., & Shai, Y. (2012). Transmembrane domains interactions within the membrane milieu: Principles, advances and challenges. *Biochimica Et Biophysica Acta (BBA) - Biomembranes*, *1818*(4), 974– 983. doi: 10.1016/j.bbamem.2011.11.029

Fons, R. D., Bogert, B. A., & Hegde, R. S. (2003). Substrate-specific function of the translocon-associated protein complex during translocation across the ER membrane. *The Journal of Cell Biology*, *160*(4), 529– 539. doi: 10.1083/jcb.200210095

Goodchild, S. J., Xu, H., Es-Salah-Lamoureux, Z., Ahern, C. A., & Fedida, D. (2012). Basis for allosteric open-state stabilization of voltage-gated potassium channels by intracellular cations. *The Journal of General Physiology*, *140*(5), 495–511. doi: 10.1085/jgp.201210823

Gorlich, D. (1993). Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane. *Cell*, *75*(4), 615–630. doi: 10.1016/0092-8674(93)90483-7

Halperin, L., Jung, J., & Michalak, M. (2014). The many functions of the endoplasmic reticulum chaperones and folding enzymes. *IUBMB Life*, *66*(5), 318–326. doi: 10.1002/iub.1272

Hartmann, E. et al. A tetrameric complex of membrane proteins in the endoplasmic reticulum. Eur. J. Biochem. (1993). doi:10.1111/j.1432-1033.1993.tb17933

Hegde, R. S., & Kang, S.-W. (2008). The concept of translocational regulation. *The Journal of Cell Biology*, *182*(2), 225–232. doi: 10.1083/jcb.200804157

Heijne, G. V. (2006). Membrane-protein topology. *Nature Reviews Molecular Cell Biology*, *7*(12), 909–918. doi: 10.1038/nrm2063

Heinrich, S. U., Mothes, W., Brunner, J., & Rapoport, T. A. (2000). The Sec61p Complex Mediates the Integration of a Membrane Protein by Allowing Lipid Partitioning of the Transmembrane Domain. *Cell*, *102*(2), 233–244. doi: 10.1016/s0092-8674(00)00028-3

Helenius, A., Marquardt, T., & Braakman, I. (1992). The endoplasmic reticulum as a protein-folding compartment. *Trends in Cell Biology*, *2*(8), 227–231. doi: 10.1016/0962-8924(92)90309-b

Hessa, T., Kim, H., Bihlmaier, K., Lundin, C., Boekel, J., Andersson, H., … Heijne, G. V. (2005). Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature*, *433*(7024), 377– 381. doi: 10.1038/nature03216

Hirama, T., Miller, C. W., & Koeffler, H. (1999). Translocon-associated protein α transcripts are induced by granulocyte-macrophage colony-stimulating factor and exhibit complex alternative polyadenylation. *FEBS Letters*, *455*(3), 223–227. doi: 10.1016/s0014-5793(99)00885-6

Iwawaki, T., Akai, R., Yamanaka, S., & Kohno, K. (2009). Function of IRE1 alpha in the placenta is essential for placental development and embryonic viability. *Proceedings of the National Academy of Sciences*, *106*(39), 16657–16662.doi:10.1073/pnas.0903775106

Jackson, M. R., Nilsson, T., & Peterson, P. A. (1990). Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. *The EMBO Journal*, *9*(10), 3153–3162. doi: 10.1002/j.1460- 2075.1990.tb07513.x

Jarjanazi, H., Savas, S., Pabalan, N., Dennis, J. W., & Ozcelik, H. (2007). Biological implications of SNPs in signal peptide domains of human proteins. *Proteins: Structure, Function, and Bioinformatics*, *70*(2), 394– 403. doi: 10.1002/prot.21548

Jarosch, E., Geiss-Friedlander, R., Meusser, B., Walter, J., & Sommer, T. (2002). Protein Dislocation from the Endoplasmic Reticulum - Pulling Out the Suspect. *Traffic*, *3*(8), 530–536. doi: 10.1034/j.1600-0854.2002.30803.x

Jensen, D., & Schekman, R. (2010). COPII-mediated vesicle formation at a glance. *Journal of Cell Science*, *124*(1), 1–4. doi: 10.1242/jcs.069773

Jungnickel, B., & Rapoport, T. A. (1995). A posttargeting signal sequence recognition event in the endoplasmic reticulum membrane. *Cell*, *82*(2), 261–270. doi: 10.1016/0092-8674(95)90313-5

Kajava, A. V., Zolov, S. N., Kalinin, A. E., & Nesmeyanova, M. A. (2000). The Net Charge of the First 18 Residues of the Mature Sequence Affects Protein Translocation across the Cytoplasmic Membrane of Gram-Negative Bacteria. *Journal of Bacteriology*, *182*(8), 2163–2169. doi: 10.1128/jb.182.8.2163-2169.2000

Kalies, K.-U., Allan, S., Sergeyenko, T., Kröger, H., & Römisch, K. (2005). The protein translocation channel binds proteasomes to the endoplasmic reticulum membrane. *The EMBO Journal*, *24*(13), 2284–2293. doi: 10.1038/sj.emboj.7600731

Kastritis, P. L., & Bonvin, A. M. J. J. (2013). On the binding affinity of macromolecular interactions: daring to ask why proteins interact. *Journal of The Royal Society Interface*, *10*(79), 20120835. doi: 10.1098/ rsif. 2012. 0835

Kida, Y., & Sakaguchi, M. (2018). Interaction mapping of the Sec61 translocon identifies two Sec61α regions interacting with hydrophobic segments in translocating chains. *Journal of Biological Chemistry*, *293*(44), 17050– 17060. doi: 10.1074/jbc.ra118.003219

Kim, S. J., Mitra, D., Salerno, J. R., & Hegde, R. S. (2002). Signal Sequences Control Gating of the Protein Translocation Channel in a Substrate-Specific Manner. *Developmental Cell*, *2*(2), 207–217. doi: 10.1016/s1534- 5807(01)00120-4

Kocik, L., Junne, T., & Spiess, M. (2012). Orientation of Internal Signal-Anchor Sequences at the Sec61 Translocon. *Journal of Molecular Biology*, *424*(5), 368–378. doi: 10.1016/j.jmb.2012.10.010

Kriegler, T., Magoulopoulou, A., Marchal, R. A., & Hessa, T. (2018). Measuring Endoplasmic Reticulum Signal Sequences Translocation Efficiency Using the Xbp1 Arrest Peptide. *Cell Chemical Biology*, *25*(7). doi: 10.1016/j.chembiol.2018.04.006

Kyte, J., & Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology*, *157*(1), 105–132. doi: 10.1016/0022-2836(82)90515-0

Lakkaraju, A. K., Abrami, L., Lemmin, T., Blaskovic, S., Kunz, B., Kihara, A., … Goot, F. G. V. D. (2012). Palmitoylated calnexin is a key component of the ribosome-translocon complex. *The EMBO Journal*, *31*(7), 1823– 1835. doi: 10.1038/emboj.2012.15

Lam, V. Q., Akopian, D., Rome, M., Henningsen, D., & Shan, S.-O. (2010). Lipid activation of the signal recognition particle receptor provides spatial coordination of protein targeting. *The Journal of Cell Biology*, *190*(4), 623–635. doi: 10.1083/jcb.201004129

Lang, S., Benedix, J., Fedeles, S. V., Schorr, S., Schirra, C., Schauble, N., … Dudek, J. (2012). Different effects of Sec61 , Sec62 and Sec63 depletion on transport of polypeptides into the endoplasmic reticulum of mammalian cells. *Journal of Cell Science*, *125*(8), 1958–1969. doi: 10.1242/jcs.096727

Lang, S., Pfeffer, S., Lee, P.-H., Cavalié, A., Helms, V., Förster, F., & Zimmermann, R. (2017). An Update on Sec61 Channel Functions, Mechanisms, and Related Diseases. *Frontiers in Physiology*, *8*. doi: 10.3389/ fphys. 2017.00887

Lao, D. M., Arai, M., Ikeda, M., & Shimizu, T. (2002). The presence of signal peptide significantly affects transmembrane topology prediction. *Bioinformatics*, *18*(12),1562–1566.doi: 10.1093/bioinformatics/18.12. 1562

Li, D.-H., Chan, T., Satow, R., Komazaki, S., Hashizume, K., & Asashima, M. (2005). The role of XTRAP-gamma in Xenopus pronephros development. *The International Journal of Developmental Biology*, *49*(4), 401–408. doi: 10.1387/ijdb.052005dl

Li, P., Wang, D., Lucas, J., Oparil, S., Xing, D., Cao, X., … Chen, Y.-F. (2008). Atrial Natriuretic Peptide

Inhibits Transforming Growth Factor β–Induced Smad Signaling and Myofibroblast Transformation in Mouse Cardiac Fibroblasts. *Circulation Research*, *102*(2), 185–192. doi: 10.1161/circresaha.107.157677

Linxweiler, M., Schick, B., & Zimmermann, R. (2017). Let's talk about Secs: Sec61, Sec62 and Sec63 in signal transduction, oncology and personalized medicine. *Signal Transduction and Targeted Therapy*, *2*(1). doi: 10.1038/sigtrans.2017.2

Lizák, B., Csala, M., Benedetti, A., & Bánhegyi, G. (2008). The translocon and the non-specific transport of small molecules in the endoplasmic reticulum (Review). *Molecular Membrane Biology*, *25*(2), 95–101. doi: 10.1080/09687680701670481

Luo, L., King, N. P., Yeo, J. C., Jones, A., & Stow, J. L. (2014). Single-step protease cleavage elution for identification of protein-protein interactions from GST pull-down and mass spectrometry. *Proteomics*, *14*(1), 19–23. doi: 10.1002/pmic.201300315

Ma, Y., & Hendershot, L. M. (2004). ER chaperone functions during normal and stress conditions. *Journal of Chemical Neuroanatomy*, *28*(1-2), 51–65. doi: 10.1016/j.jchemneu.2003.08.007

Mandon, E. C., Butova, C., Lachapelle, A., & Gilmore, R. (2018). Conserved motifs on the cytoplasmic face of the protein translocation channel are critical for the transition between resting and active conformations. *Journal of Biological Chemistry*, *293*(35), 13662–13672. doi: 10.1074/jbc.ra118.004123

Ménétret, J.-F., Hegde, R. S., Aguiar, M., Gygi, S. P., Park, E., Rapoport, T. A., & Akey, C. W. (2008). Single Copies of Sec61 and TRAP Associate with a Nontranslating Mammalian Ribosome. *Structure*, *16*(7), 1126–1137. doi: 10.1016/j.str.2008.05.003

Ménétret, J.-F., Hegde, R. S., Heinrich, S. U., Chandramouli, P., Ludtke, S. J., Rapoport, T. A., & Akey, C. W. (2005). Architecture of the Ribosome–Channel Complex Derived from Native Membranes. *Journal of Molecular Biology*, *348*(2), 445–457. doi: 10.1016/j.jmb.2005.02.053

Mesbah, K., Camus, A., Babinet, C., & Barra, J. (2006). Mutation in the Trap /Ssr1 Gene, Encoding Translocon- Associated Protein, Results in Outflow Tract Morphogenetic Defects. *Molecular and Cellular Biology*, *26*(20), 7760–7771. doi: 10.1128/mcb.00913-06

Miyazaki, K., Fujita, T., Ozaki, T., Kato, C., Kurose, Y., Sakamoto, M., … Nakagawara, A. (2003). NEDL1, a Novel Ubiquitin-protein Isopeptide Ligase for Dishevelled-1, Targets Mutant Superoxide Dismutase-1. *Journal of Biological Chemistry*, *279*(12), 11327–11335. doi: 10.1074/jbc.m312389200

Müller, L., Escauriaza, M. D. D., Lajoie, P., Theis, M., Jung, M., Müller, A., … Zimmermann, R. (2010). Evolutionary Gain of Function for the ER Membrane Protein Sec62 from Yeast to Humans. *Molecular Biology of the Cell*, *21*(5), 691–703. doi: 10.1091/mbc.e09-08-0730

Murray, A. N., Chen, W., Antonopoulos, A., Hanson, S. R., Wiseman, R. L., Dell, A., … Kelly, J. W. (2015). Enhanced Aromatic Sequons Increase Oligosaccharyltransferase Glycosylation Efficiency and Glycan Homogeneity. *Chemistry & Biology*, *22*(8), 1052–1062. doi: 10.1016/j.chembiol.2015.06.017

Nagasawa, K., Higashi, T., Hosokawa, N., Kaufman, R. J., & Nagata, K. (2007). Simultaneous induction of the four subunits of the TRAP complex by ER stress accelerates ER degradation. *EMBO Reports*, *8*(5), 483–489. doi: 10.1038/sj.embor.7400933

Neduva, V., & Russell, R. B. (2005). Linear motifs: Evolutionary interaction switches. *FEBS Letters*, *579*(15), 3342–3345. doi: 10.1016/j.febslet.2005.04.005

Nguyen, D., Stutz, R., Schorr, S., Lang, S., Pfeffer, S., Freeze, H. H., … Zimmermann, R. (2018). Proteomics reveals signal peptide features determining the client specificity in human TRAP-dependent ER protein import. *Nature Communications*, *9*(1). doi: 10.1038/s41467-018-06188-z

Nicchitta, C. V., & Blobel, G. (1993). Lumenal proteins of the mammalian endoplasmic reticulum are required to complete protein translocation. *Cell*, *73*(5), 989–998. doi: 10.1016/0092-8674(93)90276-v

Nicchitta, C. V., & Blobel, G. (1990). Assembly of translocation-competent proteoliposomes from detergent-solubilized rough microsomes. *Cell*, *60*(2), 259–269. doi: 10.1016/0092-8674(90)90741-v

Nilsson, I., Lara, P., Hessa, T., Johnson, A. E., Heijne, G. V., & Karamyshev, A. L. (2015). The Code for Directing Proteins for Translocation across ER Membrane: SRP Cotranslationally Recognizes Specific Features of a Signal Sequence. *Journal of Molecular Biology*, *427*(6), 1191–1201. doi: 10.1016/j.jmb.2014.06.014

Oliver, J., Jungnickel, B., Görlich, D., Rapoport, T., & High, S. (1995). The Sec61 complex is essential for the insertion of proteins into the membrane of the endoplasmic reticulum. *FEBS Letters*, *362*(2), 126–130.

doi: 10.1016/0014-5793(95)00223-v

Orfanoudaki, G., Markaki, M., Chatzi, K., Tsamardinos, I., & Economou, A. (2017). MatureP: prediction of secreted proteins with exclusive information from their mature regions. *Scientific Reports*, *7*(1). doi: 10.1038 / s41598-017-03557-4

Pääbo, S., Bhat, B. M., Wold, W. S., & Peterson, P. A. (1987). A short sequence in the COOH-terminus makes an adenovirus membrane glycoprotein a resident of the endoplasmic reticulum. *Cell*, *50*(2), 311–317. doi: 10.1016/0092-8674(87)90226-1

Palade, G. (1975). Intracellular Aspects of the Process of Protein Synthesis. *Science*, *189*(4206), 867–867. doi: 10.1126/science.189.4206.867-b

Park, Y., & Helms, V. (2008). Prediction of the translocon-mediated membrane insertion free energies of protein sequences. *Bioinformatics*, *24*(10), 1271–1277. doi: 10.1093/bioinformatics/btn114

Peterson, J. H., Woolhead, C. A., & Bernstein, H. D. (2003). Basic Amino Acids in a Distinct Subset of Signal Peptides Promote Interaction with the Signal Recognition Particle. *Journal of Biological Chemistry*, *278*(46), 46155–46162. doi: 10.1074/jbc.m309082200

Pfeffer, S., Brandt, F., Hrabe, T., Lang, S., Eibauer, M., Zimmermann, R., & Förster, F. (2012). Structure and 3D Arrangement of Endoplasmic Reticulum Membrane-Associated Ribosomes. *Structure*, *20*(9), 1508–1518. doi: 10.1016/j.str.2012.06.010

Pfeffer, S., Burbaum, L., Unverdorben, P., Pech, M., Chen, Y., Zimmermann, R., … Förster, F. (2015). Structure of the native Sec61 protein-conducting channel. *Nature Communications*, *6*(1). doi: 10.1038/ncomms 9403

Pfeffer, S., Dudek, J., Schaffer, M., Ng, B. G., Albert, S., Plitzko, J. M., … Förster, F. (2017). Dissecting the molecular organization of the translocon-associated protein complex. *Nature Communications*, *8*(1). doi: 10.1038/ncomms14516

Price, J. L., Culyba, E. K., Chen, W., Murray, A. N., Hanson, S. R., Wong, C.-H., … Kelly, J. W. (2012). N-glycosylation of enhanced aromatic sequons to increase glycoprotein stability. *Biopolymers*, *98*(3), 195–211. doi: 10.1002/bip.22030

Raghavachari, B., Tasneem, A., Przytycka, T. M., & Jothi, R. (2007). DOMINE: a database of protein domain interactions. *Nucleic Acids Research*, *36*(suppl_1). doi: 10.1093/nar/gkm761

Rapoport, T. A., Goder, V., Heinrich, S. U., & Matlack, K. E. (2004). Membrane-protein integration and the role of the translocation channel. *Trends in Cell Biology*, *14*(10), 568–575. doi: 10.1016/j.tcb.2004.09.002

Ruggiano, A., Foresti, O., & Carvalho, P. (2014). ER-associated degradation: Protein quality control and beyond. *The Journal of Cell Biology*, *204*(6), 869–879. doi: 10.1083/jcb.201312042

Rutkowski, D., & Kaufman, R. J. (2004). A trip to the ER: coping with stress. *Trends in Cell Biology*, *14*(1), 20–28. doi: 10.1016/j.tcb.2003.11.001

Sakaguchi, M., Tomiyoshi, R., Kuroiwa, T., Mihara, K., & Omura, T. (1992). Functions of signal and signal-anchor sequences are determined by the balance between the hydrophobic segment and the N-terminal charge. *Proceedings of the National Academy of Sciences*, *89*(1), 16–19. doi: 10.1073/pnas.89.1.16

Saurí, A., Mccormick, P. J., Johnson, A. E., & Mingarro, I. (2007). Sec61α and TRAM are Sequentially Adjacent to a Nascent Viral Membrane Protein during its ER Integration. *Journal of Molecular Biology*, *366*(2), 366–374. doi: 10.1016/j.jmb.2006.11.052

Shaffer, Kelly L., et al. "Regulation of Protein Compartmentalization Expands the Diversity of Protein Function." *Developmental Cell*, vol. 9, no. 4, 2005, pp. 545–554., doi:10.1016/j.devcel.2005.09.001.

Schäuble, Nico, et al. "BiP-Mediated Closing of the Sec61 Channel Limits Ca2 Leakage from the ER." *The EMBO Journal*, vol. 31, no. 15, 2012, pp. 3282–3296., doi:10.1038/emboj.2012.189.

Schoebel, Stefan, et al. "Cryo-EM Structure of the Protein-Conducting ERAD Channel Hrd1 in Complex with Hrd3." *Nature*, vol. 548, no. 7667, June 2017, pp. 352–355., doi:10.1038/nature23314.

Schrag, Joseph D., et al. "The Structure of Calnexin, an ER Chaperone Involved in Quality Control of Protein Folding." *Molecular Cell*, vol. 8, no. 3, 2001, pp. 633–644., doi:10.1016/s1097-2765(01)00318-5.

Scott, Daniel C., and Randy Schekman. "Role of Sec61p in the ER-Associated Degradation of Short-Lived Transmembrane Proteins." *The Journal of Cell Biology*, vol. 181, no. 7, 2008, pp. 1095–1105., doi: 10.1083/ jcb.200804053.

Shao, Sichen, and Ramanujan S. Hegde. "Membrane Protein Insertion at the Endoplasmic Reticulum."

Annual Review of Cell and Developmental Biology, vol. 27, no. 1, Oct. 2011, pp. 25–56., doi:10.1146/ annurev-cellbio-092910-154125.

Shental-Bechor, Dalit, and Yaakov Levy. "Folding of Glycoproteins: toward Understanding the Biophysics of the Glycosylation Code." *Current Opinion in Structural Biology*, vol. 19, no. 5, 2009, pp. 524–533., doi: 10.1016/j.sbi.2009.07.002.

Shibatani, Toru, et al. "Proteomic Analysis of Mammalian Oligosaccharyltransferase Reveals Multiple Subcomplexes That Contain Sec61, TRAP, and Two Potential New Subunits†." *Biochemistry*, vol. 44, no. 16, 2005, pp. 5982–5992., doi:10.1021/bi047328f.

Shin, J., et al. "Signals for Retention of Transmembrane Proteins in the Endoplasmic Reticulum Studied with CD4 Truncation Mutants." *Proceedings of the National Academy of Sciences*, vol. 88, no. 5, Jan. 1991, pp. 1918–1922., doi:10.1073/pnas.88.5.1918.

Simon, Sanford M., and Günter Blobel. "A Protein-Conducting Channel in the Endoplasmic Reticulum." *Cell*, vol. 65, no. 3, 1991, pp. 371–380., doi:10.1016/0092-8674(91)90455-8.

Simsek, Deniz, et al. "The Mammalian Ribo-Interactome Reveals Ribosome Functional Diversity and Heterogeneity." *Cell*, vol. 169, no. 6, 2017, doi:10.1016/j.cell.2017.05.022.

Singh, Snigdha, and Aditya Mittal. "Transmembrane Domain Lengths Serve as Signatures of Organismal Complexity and Viral Transport Mechanisms." *Scientific Reports*, vol. 6, no. 1, 2016, doi:10.1038/srep22352.

Sjöstrand, Dan, et al. "A Rapid Expression and Purification Condition Screening Protocol for Membrane Protein Structural Biology." *Protein Science*, vol. 26, no. 8, 2017, pp. 1653–1666., doi:10.1002/pro.3196.

Snapp, Erik L., et al. "The Organization of Engaged and Quiescent Translocons in the Endoplasmic Reticulum of Mammalian Cells." *The Journal of Cell Biology*, vol. 164, no. 7, 2004, pp. 997–1007., doi:10.1083/jcb.200312079.

Sommer, Nicole, et al. "TRAP Assists Membrane Protein Topogenesis at the Mammalian ER Membrane." *Biochimica Et Biophysica Acta (BBA) - Molecular Cell Research*, vol. 1833, no. 12, 2013, pp. 3104– 3111., doi:10.1016/j.bbamcr.2013.08.018.

Spiess, M., Junne, T., & Janoschke, M. (2019). Membrane Protein Integration and Topogenesis at the ER. *The Protein Journal*, *38*(3), 306–316. doi: 10.1007/s10930-019-09827-6

Tanaka, K. (2009). The proteasome: Overview of structure and functions. *Proceedings of the Japan Academy, Series B*, *85*(1), 12–36. doi: 10.2183/pjab.85.12

Tjong, H., & Zhou, H.-X. (2008). Prediction of Protein Solubility from Calculation of Transfer Free Energy. *Biophysical Journal*, *95*(6), 2601–2609. doi: 10.1529/biophysj.107.127746

Trueman, S. F., Mandon, E. C., & Gilmore, R. (2011). Translocation channel gating kinetics balances protein translocation efficiency with signal sequence recognition fidelity. *Molecular Biology of the Cell*, *22*(17), 2983–2993. doi: 10.1091/mbc.e11-01-0070

Tyedmers, J., Lerner, M., Wiedmann, M., Volkmer, J., & Zimmermann, R. (2003). Polypeptide-binding proteins mediate completion of co-translational protein translocation into the mammalian endoplasmic reticulum. *EMBO Reports*, *4*(5), 505–510. doi: 10.1038/sj.embor.embor826

Vincent, M. J., Martin, A. S., & Compans, R. W. (1998). Function of the KKXXMotif in Endoplasmic Reticulum Retrieval of a Transmembrane Protein Depends on the Length and Structure of the Cytoplasmic Domain. *Journal of Biological Chemistry*, *273*(2), 950–956. doi: 10.1074/jbc.273.2.950

Vishnu, N., Khan, M. J., Karsten, F., Groschner, L. N., Waldeck-Weiermair, M., Rost, R., … Malli, R. (2014). ATP increases within the lumen of the endoplasmic reticulum upon intracellular Ca2 release. *Molecular Biology of the Cell*, *25*(3), 368–379. doi: 10.1091/mbc.e13-07-0433

Voigt, S. (1996). Signal sequence-dependent function of the TRAM protein during early phases of protein transport across the endoplasmic reticulum membrane. *The Journal of Cell Biology*, *134*(1), 25–35. doi: 10.1083/jcb.134.1.25 **Volkmer**, R., Kretzschmar, I., & Tapia, V. (2012). Mapping receptor–ligand interactions with synthetic peptide arrays: Exploring the structure and function of membrane receptors. *European Journal of Cell Biology*, *91*(4), 349– 356. doi: 10.1016/j.ejcb.2011.03.004

Voorhees, R. M., Fernández, I. S., Scheres, S. H., & Hegde, R. S. (2014). Structure of the Mammalian Ribosome- Sec61 Complex to 3.4 Å Resolution. *Cell*, *157*(7), 1632–1643. doi: 10.1016/j.cell.2014.05.024

Voorhees, R. M., & Hegde, R. S. (2016). Toward a structural understanding of co-translational protein translocation. *Current Opinion in Cell Biology*, *41*, 91–99. doi: 10.1016/j.ceb.2016.04.009

Wada, I. et al. (1991) SSRα and associated calnexin are major calcium binding proteins of the endoplasmic reticulum membrane. *The Journal Biological Chemistry*, 266(29):19599-610.

Wahlberg, J. M., & Spiess, M. (1997). Multiple Determinants Direct the Orientation of Signal–Anchor

Proteins: The Topogenic Role of the Hydrophobic Signal Domain. *The Journal of Cell Biology*, *137*(3), 555–562. doi: 10.1083/jcb.137.3.555
Walter, P. (1992). Travelling by TRAM. *Nature*, *357*(6373), 22–23. doi: 10.1038/357022a0
Walter, P. (1994). Signal Sequence Recognition and Protein Targeting to the Endoplasmic Reticulum Membrane. *Annual Review of Cell and Developmental Biology*, *10*(1), 87–119. doi: 10.1146/annurev.cellbio.10.1.87
Wang, S., & Ng, D. T. W. (2008). Lectins sweet-talk proteins into ERAD. *Nature Cell Biology*, *10*(3), 251–253. doi: 10.1038/ncb0308-251
Wang, W., Barnabei, M. S., Asp, M. L., Heinis, F. I., Arden, E., Davis, J., … Metzger, J. M. (2013). Noncanonical EF-hand motif strategically delays Ca2 buffering to enhance cardiac performance. *Nature Medicine*, *19*(3), 305–312. doi: 10.1038/nm.3079
Yamaguchi, Y. L., Tanaka, S. S., Oshima, N., Kiyonari, H., Asashima, M., & Nishinakamura, R. (2011). Translocon- associated protein subunit Trap-γ/Ssr3 is required for vascular network formation in the mouse placenta. *Developmental Dynamics*, *240*(2), 394–403. doi: 10.1002/dvdy.22528
Yamamoto, H., Fukui, K., Takahashi, H., Kitamura, S., Shiota, T., Terao, K., … Endo, T. (2009). Roles of Tom70 in Import of Presequence-containing Mitochondrial Proteins. *Journal of Biological Chemistry*, *284*(46), 31635–31646. doi: 10.1074/jbc.m109.041756
Zhang, X., Rashid, R., Wang, K., & Shan, S. O. (2010). Sequential Checkpoints Govern Substrate Selection During Cotranslational Protein Targeting. *Science*, *328*(5979), 757–760. doi: 10.1126/science.1186743

Another journey, thank you to all have contributed to this achievement.

I want to thank Richard Zimmermann and Gabriele Amoroso for allowing me to start my PhD studies.

Thanks to all people that had contributed to my successful experiments: Richard Zimmermann, Moni Lerner, Martin Jung, Steffi Caesar, Johanna Dudek, Silk Guthöl, and Claudia Fecher-Trost. Thank you to all my lab and the entire Biochemistry Department.

Thanks to Markus Hoth and dissertation committee.

Thank you DAAD!!!

Antonella